



Site-specific incorporation of fluorinated amino acids into basic pancreatic trypsin inhibitor

Dissertation zur Erlangung des akademischen Grades des Doktors der
Naturwissenschaften (*Dr. rer. nat.*)

eingereicht im Fachbereich Biologie, Chemie, Pharmazie der Freien
Universität Berlin

vorgelegt von

Diplom Biochemiker Shijie Ye

aus Anhui, V. R. China

April 2014

1. Gutachter: Prof. Dr. Beate Korsch
2. Gutachter: Prof. Dr. Nediljko Budisa

Disputation am: 24 June 2014

ERKLÄRUNG

Die vorliegende Arbeit wurde auf Anregung und unter Anleitung von Frau Prof. Dr. Beate Koksch in der Zeit von August 2008 bis März 2014 am Institut für Chemie und Biochemie des Fachbereiches Biologie, Chemie, Pharmazie der Freien Universität Berlin angefertigt.

Hiermit versichere ich, die vorliegende Arbeit mit dem Titel „*Site-specific incorporation of fluorinated amino acids into basic pancreatic trypsin inhibitor*“ selbständig und ohne Benutzung anderer als der zugelassenen Hilfsmittel angefertigt zu haben. Alle angeführten Zitate sind als solche kenntlich gemacht. Die vorliegende Arbeit wurde in keinem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

Aus dieser Dissertation gingen bisher nachfolgende Veröffentlichung bzw. ein Manuskript hervor:

Shijie Ye, Allison Ann Berger, Dominique Petzold, Oliver Reimann, and Beate Koksch, *Chemical aminoacylation of tRNAs with fluorinated amino acids for in vitro protein mutagenesis*. Beilstein Journal of Organic Chemistry, **2010**, 6, 40. doi: 10.3762/bjoc.6.40.

Shijie Ye, Bernhard Loll, Allison Ann Berger, Ulrike Mülow, Markus Wahl, and Beate Koksch, *Fluorine teams up with water molecules to restore the inhibitor properties of BPTI, in preparation*

Berlin

April 2014

Shijie Ye

DANKSAGUNG

Mein ganz besonderer Dank gilt Frau Prof. Dr. Beate Kokschi für die freundliche Aufnahme in ihre Arbeitsgruppe und für die herausfordernde Themenstellung. Ich möchte mich auch für ihre Betreuung, Diskussionsbereitschaft, die wissenschaftlichen Freiräume und die Bereitstellung finanzieller Mittel bedanken.

Herrn Prof. Dr. Nediljko Budisa danke ich für die Übernahme des Zweitgutachtens meiner Dissertation.

I am truly thankful to Dr. Allison A. Berger for all her patience, guidance, and discussion during my doctoral research. Also special thanks to her for critical proofreading of my thesis draft.

In der Arbeitsgruppe möchte ich Dr. Mario Salwiczek, Dr. Elisabeth Nyakatura, Dr. Ulla Gerling, Dipl.-Chem. Jan-Stefan Völler, und M. Sc. Vivian Asante für die tägliche Zusammenarbeit und allgemeine Diskussion über Fluorforschung danken. Dipl.-Chem. Jan-Stefan Völler möchte ich auch für die Expression und Aufreinigung der SUMO-Protease 1 danken. Dr. Cosimo D. Cadicamo möchte ich für seine freundliche Diskussion über die chemischen Synthese und Peptidsynthese danken. M. Sc. Jason Heier möchte ich für das Korrekturlesen meiner Poster und Konferenzabstracts danken. M. Sc. Kristin Folmert bedanke ich mich für das Korrekturlesen der deutschen Teile meiner Dissertation. Dipl.-Chem. Matthias Hakelberg möchte ich für die Einführung in die HPLC danken. Alle Mitglieder der Arbeitsgruppe danke ich für die angenehme Arbeitsatmosphäre.

Prof. Dr. Markus Wahl und seiner Arbeitsgruppe möchte ich für die Bereitstellung der ausgezeichneten Laborausstattung für meine biologischen Arbeiten und die Zusammenarbeit bei der Strukturanalyse von Protease-BPTI-Komplexen danken. Besonderer Dank gilt Claudia Alings und Clemens Langer für ihre Hilfestellung und technische Unterstützung. Darüber hinaus danke ich Claudia Alings für die Zusammenarbeit beim Kristall-Screening. Dr. Sunbin Liu, Dr. Traudy Wandersleben,

Alexander Ulrich, Dr. Haydar Bulut, Dr. Oleg Ganichkin und Christian Becke möchte ich für die Unterstützung im biologischen Labor und bei biologischen Fragestellungen in dem biologischen Lab danken. Dr. Haydar Bulut danke ich auch für die Abgabe des Expressionsvektors pIVEX. Alle Mitglieder der AG Wahl möchte ich für die komfortable Arbeitsatmosphäre danke.

Für die Zusammenarbeit bei der Proteinkristallographie möchte ich mich ganz herzlich bei Dr. Bernhard Loll für die Röntgenstrahlen-Messungen, Strukturanalysen von Protein-Komplexen, Herstellung der Abbildungen und die Diskussion über die Dateninterpretation bedanken.

Ein großer Dank geht an Dr. Eberhard Krause und Heike Stephanowitz (Leibniz-Institut für molekulare Pharmakologie, Berlin) für die Durchführung der MALDI-Messungen und Protein-Fingerprinting-Analyse.

Großer Dank gebührt Prof. Dr. Wolfram Saenger für seine Betreuung und die Hilfe bei meiner biologischen Arbeit in seinem Labor in 2008 – 2009.

Prof. Dr. Volker A. Erdmann und Dr. Jens Peter Fürste möchte ich für die Benutzung des biologischen Labors danken.

Bei Dr. Christoph Weise möchte ich mich für MALDI-Messungen bedanken.

I am grateful to Professor S. M. Hecht of the Department of Chemistry and Biology, Biodesign Institute, Arizona State University, AZ, USA for providing the plasmid DNA PYRNA8, which contains gene of yeast suppressor tRNA^{Phe_{CUA}}.

Weiterer Dank gebührt Prof. Dr. Holger Stark (Institut für Pharmazeutische und Medizinische Chemie, Universität Düsseldorf) für das Geschenk der Expressionsvektoren pET21cHx und pBH4.

Die vorliegende Arbeit wurde von der Freien Universität Berlin, der Dahlem Research School, und dem Graduiertenkolleg „Fluor als Schlüsselement“ der Deutschen Forschungsgemeinschaft (GRK 1582/1) finanziell unterstützt

最后我想感谢我的家人在这些年岁月里对我留学生涯的支持和鼓励。爸爸妈妈在我生化专业学习的最后2年经济上的支持。 特别感谢我的表姐张熠贝， 最初帮助我来到德国， 之后一直在我海外生活的方方面面关心着我。 还有我的太太冯薇洁， 感谢你这几年来生活上的照顾和对我工作的支持和理解。 你们所有这一切是我完成这项工作的动力源泉。

感谢那些在我初到德国时给予我关心和帮助的好朋友们， 吴晓兰， 陈立缤， 孔申芳， 郭轩， 付世立， 吴庆， 感谢你们在我初到德国时给予的帮助， 引导我很快地熟悉德国社会。

还有我的好朋友们， 陆晶明， 陆文菁， 陈伟泉， 徐静， 蒋浩雄， 赵炜， 方婉纯， 感谢你们和我一起在北京度过的岁月， 留下的那段美好的留学回忆。以及自柏林由大学化学系众多中国同砚一起度过的午餐时光和广泛而有趣的讨论。

REFERAT

Aufgrund seiner einzigartigen Eigenschaften erweist sich Fluor als ein vielfältiges Werkzeug in der pharmazeutischen Chemie und in den Materialwissenschaften. Der Einbau der C-F Bindung in Proteine zeigt vielseitige, erstrebenswerte Einflüsse auf die Thermostabilität, metabolische Stabilität, Protein-Protein-Interaktionen und Selbstassoziation. Allerdings ist das aktuelle Verständnis der Eigenschaften von Fluor innerhalb einer natürlichen Proteinumgebung noch sehr eingeschränkt, d. h. dass ein rationales Proteinengineering mit fluorierten Aminosäuren zum gezielten Einfluss auf die strukturellen Eigenschaften und Funktionalitäten bisher noch nicht möglich ist. Im Rahmen der vorliegenden Doktorarbeit wurden die α -Aminobuttersäure und zwei ihrer fluorierten Analoga in den pankreatischen Trypsin-Inhibitor (BPTI) ortsspezifisch eingebaut. Die Wirksamkeiten der mutierten BPTI wurden charakterisiert.

Zwei Methoden wurden für den Einbau der nicht-natürlichen Aminosäuren in P₁-Position des BPTI eingesetzt. In der ersten Methode wurde eine Suppressor-tRNA^{Phe}_{AUC} aus Hefe-Zellen mittels semi-synthetischer Methode mit nicht-natürlichen Aminosäuren beladen. Die Proteinexpression erfolgt anschließend mittels Zell-freier Proteinexpression. Jedoch war die Produktion der mutierten BPTI-Proteine wegen geringer Effizienz der Zell-freien Expression nicht erfolgreich. Deshalb wurden verschiedene mutierte BPTI-Analoga mittels der chemischen Totalsynthese, die auf der Festphasenpeptidsynthese zusammen mit natürlicher chemischer Ligation basiert, synthetisiert. Die globale Struktur und die thermische Stabilität der synthetisierten BPTI-Analoga wurden mittels Circular dichroismus charakterisiert. Im Gegensatz zu den nicht-fluorierten BPTI-Proteinen zeigten die fluorierten BPTI-Analoga reduzierte thermische Stabilität und ein ähnliches Bindungsvermögen zu Proteasen. Die hochaufgelösten Kristallstrukturen der BPTI-Trypsin-Komplexe zeigten die minimale strukturelle Perturbation von nicht-natürlichen Seitenketten bei der S₁-Position, mit Ausnahme der Einlagerung von zwei Wassermolekülen. Die niedrigen B-Faktoren dieser zwei Wassermoleküle in fluorierten Komplexen geben Einblick darin, dass das Inhibitionsvermögen der fluorierten BPTI-Analoga von einem Fluor-induzierten Wasserstoff-Brücken-Netzwerk in der S₁-Bindungstasche unterstützt wird.

ABSTRACT

Fluorine has proven to be a versatile tool in pharmaceutical chemistry and materials science, due to its unique physicochemical properties. The incorporation of C-F bonds into peptides and proteins has also been shown to have desirable effects on thermal and metabolic stability, as well as protein-protein interactions and self-assembly. However, because our understanding of the impact of fluorine within a native protein environment is still very limited, the rational design of such unnatural biopolymers remains impractical. In this doctoral study, aminobutyric acid and two of its fluorinated analogues were site-specifically incorporated into basic (bovine) trypsin inhibitor (BPTI) and the properties of these novel mutants were characterized.

Two approaches were taken to introduce the unnatural amino acids into the P₁ position of BPTI: *in vitro* translation-based amber stop codon suppression and total chemical synthesis. In the first case, tRNAs bearing the unnatural amino acid were successfully generated by means of chemical aminoacylation, but the extremely low efficiency of *in vitro* translation made it necessary to abandon this route to the mutant inhibitors. Total chemical synthesis based on solid phase peptide synthesis and native chemical ligation enabled sufficient quantities of the unnatural BPTI variants to be obtained for further study. Their global structure and thermal stability were characterized by circular dichroism. The inhibitor activity of each was determined by means of a well-established assay. In contrast to the hydrocarbon parent side chain, the fluorinated BPTI variants retain thermal stability and inhibitor activity. High-resolution crystal structures of the three synthetic BPTI variants in complex with β -trypsin show very little overall structural perturbation with the exception of two new water molecules that fill in the space created by the removal of the native lysine side chain at the P₁ position. In the fluorine-containing structures, the *B*-factors of these two water molecules are significantly lower than the values seen in the aminobutyric acid structure, indicating that fluorine-induced noncovalent interactions within an extensive H-bond network account for the observed restoration of inhibition.

for my family and weijie

Nihil est in intellectu quod non prius fuerit in sensu.

“Nothing is in the intellect that was not earlier in the senses”

- *A principle subscribed by*

Aristotle

St. Thomas Aquinas

John Lock

Die Dissertation wurde in englischer Sprache verfasst.

ABBREVIATIONS

AA	amino acid
aa-tRNA	aminoacylated tRNA
AMPs	antimicrobial peptides
APS	ammonium persulfate
ATP	adenosine-5'-triphosphate
Bicine	2-(Bis(2-hydroxyethyl)amino)acetic acid
BPTI	basic (bovine) pancreatic trypsin inhibitor
BSA	bovine serum albumin
Bzl	benzyl
CAT	chloramphenicol acetyltransferase
CD	circular dichroism
CTP	cytidine-5'-triphosphate
CTRB1	chymotrypsinogen B1
DBU	1,8-diazabicycloundec-7-ene
dC-CE	4- <i>N</i> -benzoyl-5'- <i>O</i> -(4,4'-dimethoxytrityl)-2'-deoxycytidinyl-3'[(2-cyanoethyl)-(<i>N,N</i> -diisopropyl)]
DCM	methylene chloride
DEAE	diethylaminoethanol
DHFR	dihydrofolate reductase
DIC	<i>N,N'</i> -diisopropylcarbodiimide
DIEPA	<i>N, N</i> -diisopropylethylamine
DMAP	4-dimethylaminopyridine
DMF	<i>N, N</i> -dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDT	1,2-ethanedithiol
EDTA	ethylenediaminetetraacetic acid
EF	elongation factor
EPL	expressed protein ligation

ESI	electrospray ionization
EtBr	ethidium bromide
FAD	flavin adenine dinucleotide
Fmoc	9-fluorenylmethyloxycarbonyl
FRET	Förster resonance energy transfer
GdmCl	guanidinium chloride
GdmSCN	guanidinium thiocyanate
GFP	green fluoresce protein
GTP	guanosine-5'-triphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HOAt	1-hydroxy-7-azabenzotriazole
HOBt	1-hydroxybenzotriazole
HRP	horseradish peroxidase
5-HT	5-hydroxytryptamine receptor
IC₅₀	half maximal inhibitory concentration
ITC	isothermal titration calorimetry
IF	initial factor
kb	kilobase
kDa	kilodalton
L	Liter
LB	lysogeny broth
MALDI	matrix-assisted laser desorption/ionization
MeCN	acetonitrile
MES	2-(<i>N</i> -morpholino)ethanesulfonic acid
MOPS	3-(<i>N</i> -morpholino)propanesulfonic acid
MPAA	4-mercaptophenylacetic acid
MPD	1-methyl-2,4-pentanediol
mRNA	messenger RNA
NCL	native chemical ligation
NADP	nicotinamide adenine dinucleotide phosphate
NMR	nuclear magnetic resonance
NVOC	nitroveratryloxy carbonyl
OD	optical density

PCR	polymerase chain reaction
pdCpA	5'-phosphoryl-2'-deoxyribocytidylyriboadenosine
PEG	polyethylene glycol
PET	positron emission tomography
PGI₂	prostacyclin
PNA	peptide nucleic acid
ppm	parts per million
PyBOP	benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
RF	release factor
RNA	ribonucleic acid
RP-HPLC	reversed-phase high performance liquid chromatography
rpm	revolutions per minute
SB	super broth
SDS	sodium dodecyl sulfate
SOC	super optimal broth
SPI	selective pressure incorporation
SPPS	solid-phase peptide synthesis
TAE	Tris-acetate-EDTA
TBA	tetra- <i>n</i> -butylammonium
<i>t</i>Boc	<i>tert</i> -butyloxycarbonyl
TBS	Tris-buffered saline
TBTU	<i>O</i> -(Benzotriazol-1-yl)- <i>N,N,N',N'</i> -tetramethyluronium tetrafluoroborate
<i>t</i>(Bu)	<i>tert</i> -butyl
TEA	triethylamine
TECP-HCl	Tris(2-carboxyethyl)phosphine hydrochloride
TEMED	tetramethylethylenediamine
TFA	trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
THF	tetrahydrofuran
TIPS	triisopropylsilane
TLC	thin layer chromatography
TOF	time of flight
Tris	tris(hydroxymethyl)aminomethane

tRNA	transfer RNA
UTP	uridine-5'-triphosphate

Abbreviations of L-amino acids relevant to this thesis are given below.

Canonical amino acids

A	Ala	Alanine	M	Met	Methionine
C	Cys	Cysteine	N	Asn	Asparagine
D	Asp	Aspartic acid	P	Pro	Proline
E	Glu	Glutamic acid	Q	Gln	Glutamine
F	Phe	Phenylalanine	R	Arg	Arginine
G	Gly	Glycine	S	Ser	Serine
H	His	Histidine	T	Thr	Threonine
I	Ile	Isoleucine	V	Val	Valine
K	Lys	Lysine	W	Trp	Tryptophan
L	Leu	Leucine	Y	Tyr	Tyrosine

Non-canonical amino acids

Abu	α -aminobutyric acid, (S)-2-aminobutyric acid
DfeGly	difluoroethylglycine, (S)-2-amino-4,4-difluorobutanoic acid
DfpGly	difluoropropylglycine, (S)-2-amino-4,4-difluoropentanoic acid
2-fTyr	2-fluorotyrosine
3-fTyr	3-fluorotyrosine
HfLeu	5,5,5,5',5',5'-hexafluoroleucine
HfVal	4,4,4,4',4',4'-hexafluorovaline
MfeGly	difluoroethylglycine, (S)-2-amino-4-monofluorobutanoic acid
TfeGly(Atb)	trifluoroethylglycine, (S)-2-amino-4,4,4-trifluorobutanoic acid
TfmAla	trifluoromethylalanine, 2-amino-2-(trifluoromethyl)propanoic acid
TfMet	trifluoromethionine
TfVal	4,4,4-trifluorovaline, (S)-2-amino-3-(trifluoromethyl) butanoic acid

LIST OF TABLES

Table 1.1-1 Representative physical data for selected elements and physical properties of fluoromethanes.....	6
Table 1.1-2 pK_a values of selected fluorinated and non-fluorinated carboxyl acids, alcohols, and alanine.....	9
Table 4.1-1 A short summary of orthogonal tRNA/aaRS pairs, which have been used for the incorporation of non-canonical amino acids into proteins in different expression hosts.....	56
Table 4.2-1 Summary of chemical synthesis results for N-4-pentenoyl protected non-canonical aminoacyl-pdCpA.....	62
Table 4.2-2 Summary of constructs used for cell-free expression tests.....	68
Table 4.2-3 Commercially available cell-free protein expression kits used here.....	69
Table 4.3-1 Reaction preparation for enzymatic linearization of PYRNA8.....	86
Table 4.3-2 Run-off in vitro transcription reactions under various conditions.....	87
Table 4.3-3 Preparation of 7% acidic PAGE containing 8 M urea for tRNA analysis.....	90
Table 4.3-4 Preparation of cell culture medium.....	92
Table 4.3-5 Preparation of site-directed mutagenesis PCR.....	97
Table 4.3-6 PCR primers used for cloning.....	97
Table 4.3-7 Preparation of restriction digests and DNA ligation.....	98
Table 4.3-8 Cell-free reaction preparation by use of S30 T7 High-Yield Expression System.....	99
Table 4.3-9 Cell-free reaction preparation by use of RTS 100 E.coli Disulfide Kit.....	100
Table 4.3-10 Preparation of cell-free premixes according to literatures.....	101
Table 4.3-11 Reconstitution of cell-free reaction with prepared premixes and commercial cell-free lysate according to literature with minor modifications.....	102
Table 4.3-12 Preparation of NPI buffer for purification of 6 X His-tagged protein by use of Ni-NTA spin column.....	103
Table 4.3-13 SDS-PAGE for protein analysis and Western blot.....	104
Table 4.3-14 Buffer and solution composition for SDS-PAGE and Western blot.....	104
Table 5.2-1 Synthesis summary of mutant BPTI variants containing non-canonical amino acids at position 15.....	116
Table 5.3-1 Thermodynamic denaturation parameters for thermal unfolding of BPTI variants containing non-canonical amino acids at P_1 position and data from literature.....	121
Table 5.3-2 Binding disassociation constant (K_d M) of BPTI variants with proteases and data from literature.....	123
Table 5.3-3 RMSD analysis in Å of β -trypsin-BPTI _{mutant} complexes in comparison with the β -trypsin-BPTI _{Lys15wild-type}	130
Table 5.3-4 Summary of distances between side chain atoms of P_1 residues and atoms belonging to S1 binding pocket of β -trypsin or BPTI.....	132
Table 5.3-5 Summary of B-factors of β -trypsin-BPTI complexes.....	135
Table 5.3-6 Summary of distances between P_1 side chains and structural water molecules in S_1 pocket, distances between water molecules and trypsin residues.....	136
Table 5.4-1 Protocol for elongation of peptide fragment BPTI 1-nAA15-37.....	145
Table 5.4-2 Fitting parameters for thermodynamic characterization.....	152
Table 5.4-3 Preparation of crystallization screening for protease-BPTI complexes.....	158
Table 5.4-4 Data collection and refinement statistics.....	159

LIST OF FIGURES

Figure 1.1-1 Naturally occurring organofluorine compounds.....	5
Figure 1.1-2 Schematic representation of gauche effect of 1,2-difluoroethane and N- β -fluoroethylamide and relative free energy difference.....	7
Figure 1.1-3 Optimizations of SCH48461 by fluorination.....	8
Figure 1.1-4 Modulation of pK _a values and bioavailability of 3-(3-(Piperidin-1-yl)propyl)indole derivate by fluorination.....	10
Figure 1.1-5 Hydrolysis of prostacyclin in vivo and redesign of prostacyclin by fluorination.....	11
Figure 1.1-6 A typical three-phase liquid extraction shows fluorous effect.....	13
Figure 1.1-7 Structures of 5-fluorouracil and trifluridine.....	13
Figure 1.1-8 Schematic representation of methyl transfer mechanism of thymidylate synthase and inhibition mechanism of 5-fluorouracil.....	14
Figure 1.1-9 Chemical structure of 2-deoxy-2-[¹⁸ F]fluoro-D-glucose.....	15
Figure 1.2-1 3D representation of dimeric form of fluorinated GCN4-pld.....	18
Figure 1.2-2 Schematic representation of Leu, Val, Ile, and their fluorinated analogues incorporated into coiled-coil model peptides.....	19
Figure 1.2-3 Schematic representation of (S)-2-aminobutyric acid and its fluorinated analogues used in the studies of our group.....	20
Figure 1.2-4 Schematic representation of MSI-78 dimer.....	21
Figure 1.2-5 Schematic representation of the dimeric α_2D model protein, the monomers presented in gray and light cyan; the fluorinated analogues of Phe incorporated into α_2D and their electrostatic potential maps.....	23
Figure 1.2-6 Schematic representation of probing the stability and spectroscopic properties of GFP by the incorporation of fluorinated analogues of tyrosine.....	25
Figure 3.1-1 Ribbon representation of BPTI.....	34
Figure 4.1-1 Illustration of aminoacylation of tRNA catalyzed by aminoacyl-tRNA synthetases.....	41
Figure 4.1-2 Editing mechanisms of aaRSs against misacylated tRNAs.....	42
Figure 4.1-3 Schematic representation of translation elongation and peptidyl transfer.....	44
Figure 4.1-4 Schematic representation of amber suppression based in vitro biosynthetic approaches for protein mutagenesis.....	47
Figure 4.1-5 Schematic representation of T4 RNA ligase mediated chemical aminoacylation.....	48
Figure 4.1-6 Schematic representations of nitroveratryloxy carbonyl and N-4-pentenoyl protecting groups and their removal mechanisms.....	49
Figure 4.1-7 Schematic representation of diversity of tRNA ^{Tyr} from different organisms and general approach of orthogonal tRNA selection.....	54
Figure 4.1-8 Schematic representation of an efficient system for the evolution of orthogonal aaRSs.....	55
Figure 4.1-9 Schematic representation of an orthogonal expression network.....	58
Figure 4.2-1 Structures of non-canonical amino acids used for chemical aminoacylation.....	59
Figure 4.2-2 General approaches for the synthesis of N-4-pentenoyl amino acid cyanomethyl ester.....	60
Figure 4.2-3 Scheme of synthesis of mono-2' (3')-O-[N-(4-pentenoyl)aminoacyl]-pdCpAs and 2' -3' -bis-O-[N-(4-pentenoyl)aminoacyl]-pdCpAs.....	61
Figure 4.2-4 Agarose gel electrophoresis analysis of linearized plasmid PYRNA8.....	62
Figure 4.2-5 Urea denaturing acidic PAGE analysis of run-off in vitro transcription and purification of truncated suppressor tRNAs by diethylaminoethanol sepharose anion-exchange column.....	63
Figure 4.2-6 Urea denaturing PAGE for analysis of T4 RNA ligase mediated ligation.....	64
Figure 4.2-7 Urea denaturing acidic PAGE analysis for stability of N-4-pentenoyl-aa-tRNA and deprotected aa-tRNA.....	66
Figure 4.2-8 Characterization of cell-free expression vectors by endonuclease digestion and agarose gel electrophoresis.....	68

Figure 4.2-9 Western blot analysis of protein expression by use of RTS 100 E.coli Disulfide Kit and S30 T7 High-Yield Protein Expression System.....	70
Figure 4.2-10 Western blot and SDS-PAGE analysis of protein expression by use of RTS 100 E.coli Disulfide Kit.....	71
Figure 4.2-11 SDS PAGE analysis of purified SUMO-BPTI and BPTI produced by RTS 100 E.coli Disulfide Kit and protein sequence identified by protein fingerprinting.....	72
Figure 4.2-12 SDS PAGE and Western blot analysis of SUMO fusion tag cleavage.....	73
Figure 4.3-1 Structure of synthetic intermediates pdCpA-1 and (iPr) ₂ N-P-(OCH ₂ CH ₂ CN) ₂	75
Figure 4.3-2 Structure of synthetic intermediates pdCpA-2 and pdCpA-3	76
Figure 4.3-3 Structure of synthetic intermediate pdCpA-4 and pdCpA.....	78
Figure 4.3-4 Structure of tetrabutylammonium salt of pdCpA.....	80
Figure 4.3-5 Schematic illustration of assembly for tank blotting.....	105
Figure 5.1-1 Schematic representation of the principles of SPPS.....	110
Figure 5.1-2 Schematic representation of the principles of native chemical ligation	112
Figure 5.2-1 Schematic representation of synthetic strategy for total chemical synthesis of mutant BPTI variants containing non-canonical amino acids at position 15 and three incorporated non-canonical amino acids.....	114
Figure 5.2-2 Monitoring the NCL reaction by RP-HPLC.....	115
Figure 5.2-3 Monitoring BPTI _{Lys15Abu} folding by RP-HPLC	116
Figure 5.2-4 Circular dichroism analysis of synthetic BPTI species in Tris buffer at pH 7.4.....	117
Figure 5.3-1 Thermal unfolding transition curves of BPTI variants at pH 2 in the presence of 6 M GdmCl and 8 M urea	120
Figure 5.3-2 Schematic representation of inhibitory activity of synthetic BPTI variants to α -chymotrypsin, β -trypsin, and plasmin from human plasma	122
Figure 5.3-3 ITC measurement of wild-type BPTI with β -trypsin.....	125
Figure 5.3-4 ITC measurement of wild-type BPTI with α -chymotrypsin in which equilibrium was not reached	125
Figure 5.3-5 Size-exclusion spectrum of α -chymotrypsin-BPTI _{Lys15wild-type} complex and SDS PAGE analysis of appropriate fractions obtained from size-exclusion chromatography	126
Figure 5.3-6 Size-exclusion spectrum of β -trypsin-BPTI _{Lys15wild-type} complex and SDS PAGE analysis of appropriate fractions obtained from size-exclusion chromatography	127
Figure 5.3-7 Crystals of β -trypsin-BPTI complexes under microscopy.....	128
Figure 5.3-8 Crystals of α -chymotrypsin-BPTI _{Lys15wild type} complexes under microscopy	128
Figure 5.3-9 Structural context of the P ₁ residues in the S ₁ binding pocket.....	129
Figure 5.3-10 Superimposition of residues 15 of BPTI and structural water molecules.....	130
Figure 5.3-11 Top view of S ₁ pocket of β -trypsin occupied with non-canonical residues at P ₁ position and distance analysis in he context β -trypsin residues for each fluorine atom	131
Figure 5.3-12 Stereoview of the interacting loop of BPTI _{Lys15TfGly} in S ₁ pocket.....	134
Figure 5.3-13 Schematic representation of P ₁ residues in the S ₁ pocket and distances between potential interacting side chains.....	138
Figure 5.4-1 CD spectra of polypeptides and proteins showing secondary structures.....	149
Figure 5.4-2 Schematic representation of a reversible denaturation of BPTI.....	150
Figure 5.4-3 Schematic representation of a melting curve showing the linear baseline for folded and denatured state of protein	152
Figure 5.4-4 Van't Hoff plot of ΔH_m against T _m for BPTI variants.....	153
Figure 5.4-5 Schematic representation of the ITC set up	156
Figure 6.2-1 Ribbon representation of BPTI	166

TABLE OF CONTENTS

Preface.....	1
Chapter 1 Unique properties of fluorine and fluorine in protein design.....	3
1.1 Unique properties of fluorine	5
1.1.1 The effects of fluorine on molecular size and conformation.....	6
1.1.1.1 Mimic effect and steric perturbation.....	6
1.1.1.2 Effects of fluorine on molecular conformation	7
1.1.2 The effects of fluorine on physicochemical properties.....	8
1.1.2.1 Block effect.....	8
1.1.2.2 Inductive effects of fluorine.....	9
1.1.2.3 Lipophilicity of fluorinated compounds.....	11
1.1.3 Participation of fluorine in weak interactions.....	11
1.1.3.1 Hydrogen bond.....	11
1.1.3.2 Orthogonal dipolar interactions.....	12
1.1.4 Fluorous effect.....	12
1.1.5 Orthogonal reactivity of fluorine.....	13
1.1.6 Fluorine as label.....	14
1.1.6.1 Positron emission tomography (PET).....	14
1.1.6.2 ¹⁹ F nuclear magnetic resonance	15
1.2 Fluorine in protein design.....	17
1.2.1 Effects of fluorine in coiled-coil peptides	17
1.2.2 Secondary structure propensities of fluorinated amino acids	20
1.2.3 Fluorination of antimicrobial peptides.....	21
1.2.4 An example of fluorine in aromatic stacking effects	22
1.2.5 Effects of global substitution with fluorinated amino acids	24
Chapter 2 Aim of the work	27
Aim of the work	29
Chapter 3 Basic (bovine) pancreatic trypsin inhibitor	31
3.1 General properties of BPTI.....	33
3.2 Protein stability and basis of proteases-BPTI interaction	34
Chapter 4 Probing the site-specific incorporation of fluorinated amino acids into BPTI by <i>in vitro</i> non-sense suppression	37
4.1 Protein engineering with non-canonical amino acids.....	39
4.1.1 Fidelity of protein translation.....	39
4.1.1.1 Aminoacyl-tRNA synthetases and amino acid activation	40
4.1.1.2 Selection by decoding.....	43
4.1.1.3 Selection by kinetic proofreading and induced fit.....	43
4.1.2 Residue-specific incorporation of non-canonical amino acids into proteins.....	45
4.1.3 <i>in vitro</i> non-sense suppression	46
4.1.3.1 Chemical aminoacylation of suppressor tRNA.....	48
4.1.3.2 <i>in vitro</i> protein synthesis	50
4.1.3.3 Limitations and improvements of <i>in vitro</i> non-sense suppression.....	52
4.1.4 Expanding the genetic code	52
4.1.4.1 General approaches to the selection of orthogonal tRNA/aaRS pairs	53
4.1.4.2 Limitations on expanding the genetic code	57
4.1.5 Orthogonal systems and reprogramming the genetic code.....	58
4.2 Site-specific incorporation of fluorinated amino acids into BPTI	59

4.2.1	Chemical misacylation of fluorinated amino acids	59
4.2.1.1	Synthesis of hybrid dinucleotide pdCpA.....	59
4.2.1.2	Synthesis of <i>N</i> -4-pentenoyl amino acid cyanomethyl esters	59
4.2.1.3	Synthesis of mono-2' (3')- <i>O</i> -[<i>N</i> -(4-pentenoyl)aminoacyl]-pdCpAs and 2' -3' -bis- <i>O</i> -[<i>N</i> -(4-pentenoyl)aminoacyl]-pdCpAs	61
4.2.1.4	T4 RNA ligase-mediated ligation.....	62
4.2.1.5	Characterization of misacylated suppressor tRNA ^{Phe} _{CUA}	65
4.2.2	Construct design and cloning for expression of BPTI.....	66
4.2.2.1	Design of fluorinated BPTI.....	66
4.2.2.2	Use of small ubiquitin-like modifier as expression fusion tag.....	67
4.2.2.3	Expression test of chymotrypsinogen B1 and trypsin 1	67
4.2.2.4	Cloning of expression constructs	68
4.2.3	<i>in vitro</i> protein expression.....	69
4.2.3.1	Protein expression by use of commercially available kits	69
4.2.3.2	Cleavage of the SUMO fusion tag.....	72
4.2.3.3	Protein expression by reconstitution of cell-free reactions	73
4.3	Methods and experimental procedures	74
4.3.1	Chemical aminoacylation of non-canonical amino acids with yeast suppressor tRNA ^{Phe} _{CUA}	74
4.3.1.1	General information about chemical synthesis, purification, and chemical characterization.....	74
4.3.1.2	Synthesis of pdCpA.....	75
4.3.1.3	Synthesis of <i>N</i> -(4-pentenoyl)-amino acid cyanomethyl ester	80
4.3.1.4	Synthesis of mono-2' (3')- <i>O</i> -[<i>N</i> -(4-pentenoyl)aminoacyl]-pdCpAs and 2' -3' -bis- <i>O</i> -[<i>N</i> -(4-pentenoyl)aminoacyl]-pdCpAs	83
4.3.1.5	T4 RNA ligase-mediated ligation.....	86
4.3.2	Cloning and protein expression	90
4.3.2.1	General methods for molecular biology and protein chemistry.....	90
4.3.2.2	Cloning of plasmid DNA for protein expression	95
4.3.2.3	Protein expression and <i>in vitro</i> amber suppression.....	99
4.3.2.4	Protein purification and analysis	103
Chapter 5 Total chemical synthesis and full characterization of BPTI containing fluorinated amino acids		107
5.1	Total chemical synthesis of proteins.....	109
5.1.1	Solid phase peptide synthesis	109
5.1.2	Native chemical ligation.....	111
5.1.3	Fmoc-based SPPS synthesis of peptide α -thioester	112
5.2	Total chemical synthesis of BPTI containing fluorinated amino acids	114
5.3	Characterization of BPTI containing of fluorinated amino acids	119
5.3.1	Protein stability regarding substitution at P ₁ position of BPTI	119
5.3.2	Studies of mutant BPTI-protease interactions.....	122
5.3.2.1	Studies of inhibitor activity.....	122
5.3.2.2	Probing inhibitor-protease binding by means of ITC.....	125
5.3.3	Protein crystallographic analysis of protease-BPTI complexes	126
5.3.3.1	Purification of protease-BPTI complexes	126
5.3.3.2	Crystallization of proteases-BPTI complexes	127
5.3.3.3	Crystallographic analysis of proteases-BPTI complexes	129
5.4	Methods and experimental procedures	139
5.4.1	General information about chemical synthesis and analysis	139
5.4.2	Synthesis of Fmoc-(<i>S</i>)-2-amino-4,4,4-trifluorobutyric acid.....	141
5.4.3	Peptide synthesis and purification.....	144
5.4.3.1	Synthesis of BPTI C-terminal fragment	144

5.4.3.2	Synthesis of BPTI N-terminal fragment; preparation of α -thioester	144
5.4.4	Native chemical ligation	146
5.4.5	Refolding of synthetic mutant BPTI variants	147
5.4.6	Determination of protein concentration	148
5.4.7	Protein structural analysis and thermal stability analysis by circular dichroism ...	148
5.4.7.1	Basis of circular dichroism spectroscopy	148
5.4.7.2	Structural analysis of BPTI variants by use of CD	149
5.4.7.3	Thermal stability analysis of BPTI variants	150
5.4.8	Binding studies of BPTI variants with proteases	153
5.4.8.1	Competitive inhibition assay	153
5.4.8.2	Binding studies by means of isothermal titration calorimetry	156
5.4.9	Protein crystallographic analysis	157
5.4.9.1	Purification of proteins and protein complexes for crystallization	157
5.4.9.2	Crystallization of protease-BPTI complexes	157
5.4.10	Data collection, structural refinement, and data analysis	158
Chapter 6	Summary and outlook	161
6.1	Site-specific incorporation of fluorinated amino acids into BPTI by <i>amber</i> suppression	163
6.2	Total chemical synthesis and full characterization of BPTI containing fluorinated amino acids	165
Literature	169
Appendix	189
Curriculum vitae		

PREFACE

Proteins are biological macromolecules consisting of one or several units, which are assembled into linear chains and naturally fold into unique three-dimensional structures; there are 20 common canonical amino acids and rare selenocysteine and pyrrolysine. With the exception of certain functional ribonucleic acids, proteins are the cellular actors that carry out a remarkable range of functions, including building up the cell scaffold, catalyzing chemical reactions, and transducing signals. Most proteins require post-translational modifications or cofactors for their function. Therefore, one of the essential tasks of advanced biochemical research is protein engineering, to develop a new chemical toolkit to produce proteins with novel or enhanced bioactivities, which meet the needs of human societies, particularly as potential pharmaceuticals and biomaterials for healthcare.

Recently, fluorine has been emerging as a promising tool for protein design. This is mainly due to the fact that incorporation of fluorine into molecules can result in unique effects, which cannot be achieved by the use of any other elements. The incorporation of fluorine into proteins might increase protein stability and proteolytic resistance, which could result, in an extreme case, in an artificial “*Teflon protein*” that could be advantageous for pharmaceutical and biotechnological purposes. Fluorination of proteins might also modulate membrane permeability and induce novel interactions. Moreover, fluorine can be an excellent bioorthogonal biophysical marker, and can be used to monitor dynamic biological processes. Though numerous studies of protein engineering by fluorination have been done, a general interpretation of fluorine’s behavior within a native protein environment cannot be made based on our current knowledge.

In the context of this doctoral thesis, the general properties of fluorine, protein engineering by fluorination, and the methodologies used for the incorporation of non-canonical amino acids into proteins, based on literature reports, are discussed by way of example and briefly summarized. Experimentally, the incorporation of

Preface

several fluorinated amino acids and their non-fluorinated analogues into a proteinogenic inhibitor by means of *in vitro* amber stop codon suppression and total chemical synthesis was attempted and accomplished, respectively. The results obtained from biophysical characterization and protein structural analysis provide the scientific community with a novel insight into the effect of fluorine within a native protein environment.

Chapter 1

Unique properties of fluorine and fluorine in protein design

1.1 Unique properties of fluorine

As the thirteenth most common element in earth's crust, the natural abundance of fluorine exists mainly as fluorite (CaF_2), fluorapatite ($\text{Ca}_5(\text{PO}_4)_3\text{F}$) and cryolite (Na_3AlF_6). Surprisingly, as the most abundant halogen in the earth's crust, and compared to the some 3000 natural products containing chlorine, bromine, and iodine that are known, only 13 natural organic compounds contain fluorine, of which eight are ω -fluorinated fatty acids from the same plant.¹

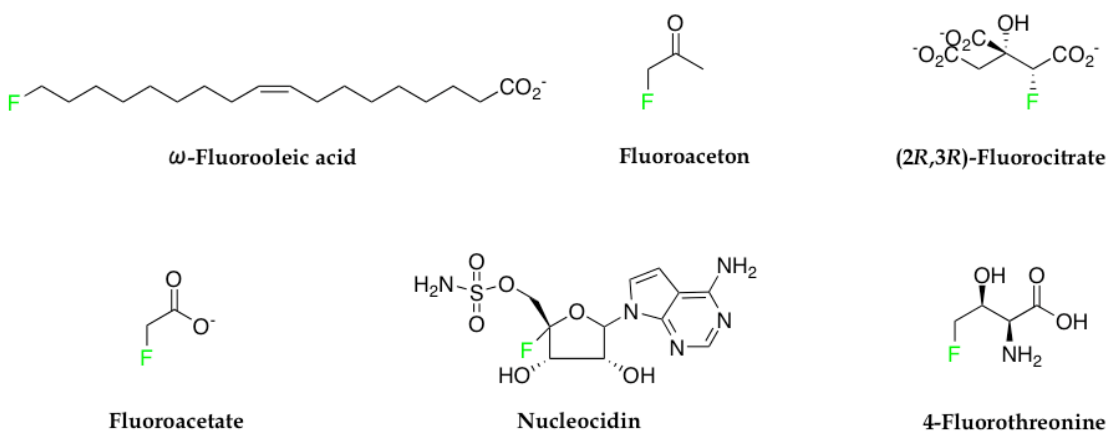


Figure 1.1-1 Naturally occurring organofluorine compounds

In spite of such scarcity, synthetic fluorine-containing compounds have been widely used in various fields. For instance, prior to 1957 no fluorine-containing pharmaceuticals had been developed, but now about 20% of all pharmaceuticals on the market contain at least one fluorine atom. Commercial agrochemicals give an even higher figure with more than 30% of fluorine-containing products.² Furthermore, organic fluorine compounds are also widely used in materials science to design materials with unique properties.³ Recently, fluorinated analogues of the canonical amino acids have attracted widespread attention as building blocks for protein design, to produce proteins with advantageous biochemical and biophysical properties.⁴

1.1 The unique properties of fluorine

1.1.1 The effects of fluorine on molecular size and conformation

The properties of organofluorine compounds are influenced by the size and electronegativity of the fluorine constituent. The C-F bond is a highly polarized bond with great bond strength and very low polarizability. *Table 1.1-1* summarizes the relevant physical properties of fluorine in comparison with other selected elements, and the physical properties of fluoromethanes.

Table 1.1-1 Representative physical data for selected elements and physical properties of fluoromethanes

	H	C	N	O	F	Cl	Br	I
van der Waals radius (Å) ^a	1.20	1.70	1.55	1.52	1.47	1.75	1.85	1.98
C-X bond lengths (Å) ^a	1.09	1.54	1.47	1.43	1.35	1.77	-	-
Electronegativity ^a	2.2	2.6	3.0	3.4	4.0	3.2	3.0	2.7
C-X bond strength (kcal/mol) ^b	98.8	83.1	69.7	84.0	105.4	78.5	65.9	57.4
Ionization potential (Kcal/mol) ^a	313.6	240.5	335.1	314.0	401.8	299.0	272.4	241.2
Electronaffinity (Kcal/mol) ^a	17.7	29.0	-6.2	33.8	79.5	83.3	72.6	70.6
Polarizability (Å) ^a	0.67	1.76	1.10	0.82	0.56	2.18	3.05	4.7
	CH₄	CH₃F	CH₂F₂	CHF₃	CF₄			
C-F bond length (Å) ^b	-	1.39	1.36	1.33	1.32			
C-F bond strength (kcal/mol) ^b	-	107	109.6	114.6	116			
Dipole moment (μ)/D ^b	0	1.85	1.97	1.65	0			
Boiling point (°C) ^b	-161	-78	-52	-83	-128			

Reference: a⁵ b⁶.

1.1.1.1 Mimic effect and steric perturbation

The van der Waals radius of fluorine is only about 20% larger than hydrogen, and the C-F bond length is 21% greater than the C-H bond in CH₄. It has been shown that enzymes often ignore the difference in substrates when hydrogen is substituted by fluorine. Therefore, substitution with fluorine in enzyme substrates or inhibitors can be directly adapted to biochemical systems for probing of molecular recognition.⁷ Besides the mimic effect, the introduction of a trifluoromethyl group results in important conformational changes within organic molecules. It is worth noting that

1.1 The unique properties of fluorine

introducing one fluorine atom into a methyl group leads to a higher axial preference compared to the hydrocarbon parent. The incorporation of two fluorine atom into a methyl group leads to a very small increase in the *A* value (conformational free energy, kcal/mol), while the substitution of three fluorine atoms into a methyl group causes a significant change in the *A* value.⁸ It is generally accepted that the bulkiness of a CF₃ group is close to that of a (CH₃)₂CH- group.⁹

1.1.1.2 Effects of fluorine on molecular conformation

The combination of steric effects and high electron withdrawing effect of fluorine sometimes leads to unpredictable molecular conformations. For instance, methoxybenzene and trifluoromethoxybenzene do not share similar structures. In trifluoromethoxybenzene the dihedral angle of C-C-O-C for the -OCF₃ group is about 90°, rather than a planar conformation as in the -OCH₃ group in methoxybenzene.¹⁰ This phenomena was further confirmed by molecular modeling and applied to an inhibitor of cholesteryl ester transfer protein, in which the out-of-plane orientation of the fluoromethyl group fits better in the binding site and enhances binding affinity.¹¹

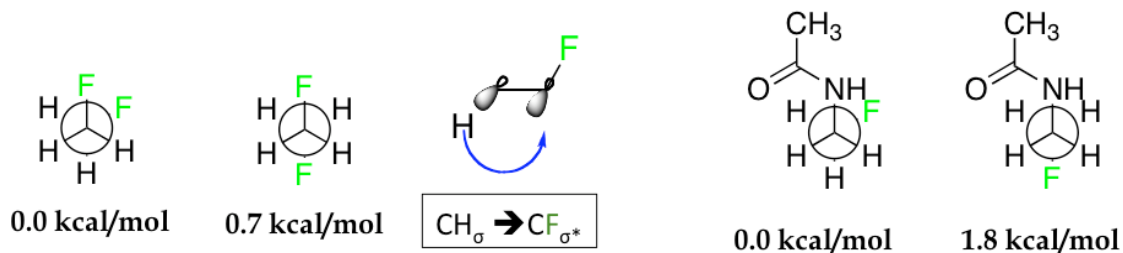


Figure 1.1-2 Schematic representation of *gauche* effect of 1,2-difluoroethane and *N*-β-fluoroethylamide and relative free energy difference

Another well know conformational effect due to fluorination is the *gauche* effect, exemplified in chemicals such as 1,2-difluoroethane, which prefers a *gauche* rather than an *anti* conformation. *Figure 1.1-2* depicts the *gauche* and *anti* conformation and

1.1 The unique properties of fluorine

their relative free energy difference for 1,2-difluoroethane and *N*- β -fluoroethylamide. The more favorable *gauche* conformation can be explained by hyperconjugation involving the highly polarized C-F bond. Both electronegative fluorine atoms are anti-periplanar to a C-H bond and stabilize this conformation by hyperconjugative interaction.^{6,12} The *gauche* effect was also observed in the conformational preference of 4-fluoroprolines. NMR data demonstrated that (4*R*)-fluoroproline prefers the *C^γ-exo* conformation while (4*S*)-fluoroproline prefers the *C^γ-endo* conformation. This difference can be applied for protein design to produce proteins containing 4-fluoroproline with different biophysical characteristics.¹³

1.1.2 The effects of fluorine on physicochemical properties

1.1.2.1 Block effect

The block effect refers to the observation that the substitution of hydrogen with fluorine can significantly block the metabolic processes that result in the hydroxylation of C-H bonds. This is mainly due to the greater C-F bond strength.

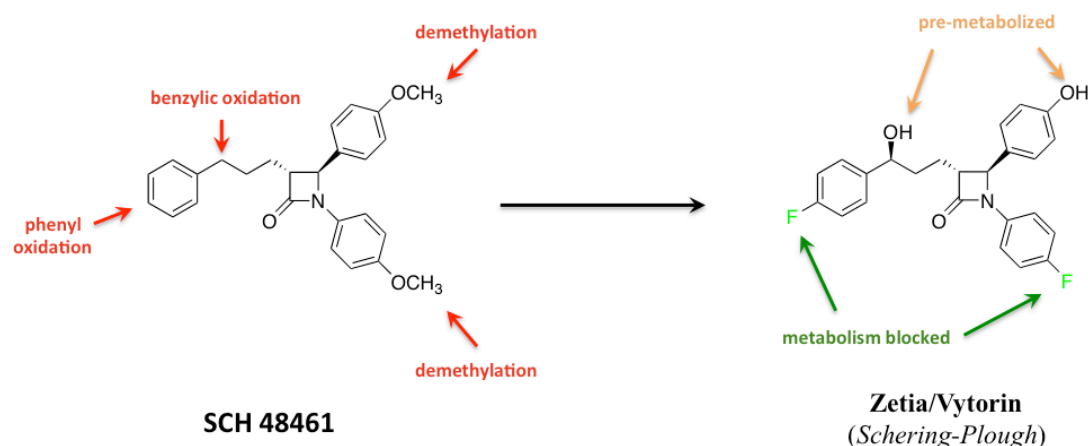


Figure 1.1-3 Optimizations of SCH48461 by fluorination

Replacement of hydrogen with fluorine on aromatic ring inductively deactivates the phenyl groups towards oxidation

1.1 The unique properties of fluorine

This effect can be applied for drug design to increase the metabolic stability of drugs, particularly resistance to the enzyme-like cytochrome P-450. *Figure 1.1-3* schematically depicts the metabolic degradation sites of SCH 48461 and the redesign of SCH 48461 by introducing fluorine atoms. The incorporation of fluorine protects the phenyl group towards oxidization and results in the drug Zetia. Zetia was found to be ~400 times more potent than SCH48461, and is effective even at low doses.¹⁴

1.1.2.2 Inductive effects of fluorine

Because fluorine is the most electronegative element, its incorporation can induce electronic effects on the neighboring groups by withdrawing electron density and subsequently affecting the pK_a value, Lewis acidity/basicity, and stability of the neighboring functional groups. Such effects can be utilized for drug design to generate pharmaceutical compounds with increased binding affinity for the target receptors, improved biocompatibility, and enhanced pharmacokinetics. *Table 1.1-2* presents pK_a values of selected carboxylic acids, alcohols, and alanine with or without fluorine substituents.

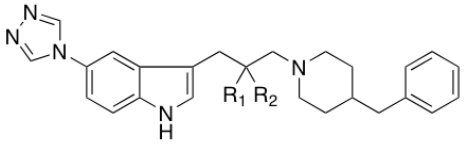
Table 1.1-2 pK_a values of selected fluorinated and non-fluorinated carboxyl acids, alcohols, and alanine

Carboxylic acid	pK_a	Alcohol	pK_a	Amino acid ^b	pK_{COO^-}	$pK_{NH_3^+}$
CH ₃ COOH	4.76	CH ₃ CH ₂ OH	15.93	NH ₃ ⁺ -CH(CH ₃)-COO ⁻	2.3	9.9
CH ₂ F ₂ COOH	2.59	CF ₃ CH ₂ OH	12.93	NH ₃ ⁺ -CH(CH ₂ F)-COO ⁻	2.4	9.8
CHF ₂ COOH	1.33			NH ₃ ⁺ -CH(CF ₃)-COO ⁻	1.2	5.4
CF ₃ COOH	0.50	(CH ₃) ₂ CHOH	17.10			
		(CF ₃) ₂ CHOH	9.30			
CH ₂ ClCOOH	2.87					
CH ₂ BrCOOH	2.90	(CH ₃) ₃ COH	19.00			
		(CF ₃) ₃ COH	5.40			
CH ₃ CH ₂ OOH	4.87					
CF ₃ CH ₂ COOH	3.06	C ₆ H ₅ OH	9.99			
		C ₆ F ₅ COH	5.50			
C ₆ H ₅ COOH	4.21					
C ₆ F ₅ COOH	1.70					

Reference: a⁷ b¹⁵

1.1 The unique properties of fluorine

3-(3-(Piperidin-1-yl)propyl)indoles bind 5-hydroxytryptamine receptors (5-HT) receptors with high selectivity. The introduction of fluorine into these molecules results in dramatic reductions in pK_a values of these compounds and leads to different efficiencies of oral absorption and bioavailability. The pK_a values, half-inhibition concentration (IC_{50}), and bioavailability data of 3-(3-(Piperidin-1-yl)propyl)indoles are summarized in Figure 1.1-4.^{14, 16}



3-(3-(Piperidin-1-yl)propyl)indoles derivate
5HT_{1D} agonists for migraine relief

R1	R2	pK_a	IC_{50} of 5-HT _{1D} receptor (nM)	Bioavailability
H	H	9.7	0.3	Low
F	H	8.7	0.9	Good
F	F	6.7	78	-

Figure 1.1-4 Modulation of pK_a values and bioavailability of 3-(3-(Piperidin-1-yl)propyl)indole derivate by fluorination

Besides the block effect by direct fluorination, the introduction of fluorine can also indirectly affect metabolic stability by influencing neighboring functional groups. Prostacyclin (PGI_2) belongs to a family of lipids, and it was initially used to inhibit platelet aggregation. Recently, PGI_2 became an attractive candidate for the treatment of inflammatory and vascular diseases. *In vivo* PGI_2 easily undergoes hydrolysis to afford 6-keto- PGI_2 (Figure 1.1-5). In order to increase the half-life of this pharmaceutical, fluorine was introduced, as its strong electron withdrawing effects influence the enol ether group and reduce the hydrolysis rate. Experimental data showed that the half-life of fluorinated PGI_2 analogues was increased from 10 minutes to ~1 months by monofluorination, and ~3 months by difluorination at the neighboring position.^{14, 17}

1.1 The unique properties of fluorine

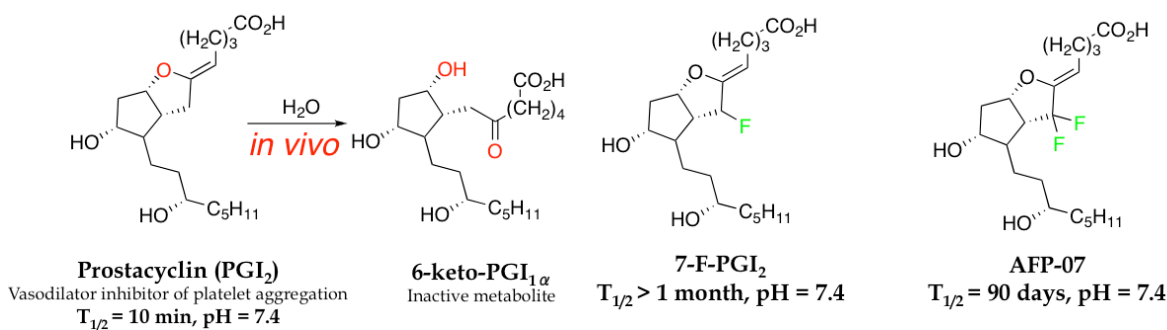


Figure 1.1-5 Hydrolysis of prostacyclin *in vivo* and redesign of prostacyclin by fluorination

1.1.2.3 Lipophilicity of fluorinated compounds

Since most orally administered drugs are absorbed and distributed via passive transport, in drug design it is very important to ensure that molecules penetrate the cell membrane but not become trapped. Fluorine has been proven to be effective in modulating the lipophilicity of molecules; however, the nature of its impact remains difficult to predict. In general, aromatic fluorination or per/polyfluorination increases lipophilicity, while introduction of fluorine into aliphatic compounds leads to a decrease in lipophilicity.¹⁵

1.1.3 Participation of fluorine in weak interactions

1.1.3.1 Hydrogen bond

Although the C-F bond is strongly polarized, three lone pairs of electrons on the fluorine atom could enable interactions with a proton; however, statistical analysis of structural data indicates that organic fluorine hardly accepts hydrogen bonds.¹⁸ Numerous studies by means of theoretical calculation, crystallography, and experimental investigations have been done to reveal evidence for intermolecular or intramolecular hydrogen bonding.¹⁹ A conventional hydrogen bond, such as $\text{ROH} \cdots \text{O}=\text{C}$, has a contact distance of $\sim 1.9 \text{ \AA}$ and contributes 5-10 kcal/mol. In

1.1 *The unique properties of fluorine*

isolated cases, the shortest observed $\text{CF}\cdots\text{HX}$ contacts are in the range 2.0-2.2 Å. But many more instances have revealed that distances of $\text{CF}\cdots\text{HX}$ are typically beyond the van der Waals contact distance between fluorine and hydrogen (~2.65 Å). Thus, due to its high electron affinity and low polarizability, fluorine hardly ever acts as hydrogen acceptor. Several researchers have suggested that the weak interaction between $\text{CF}\cdots\text{HX}$ might more appropriately be categorized as a dipolar interaction rather than a true hydrogen bond.

1.1.3.2 Orthogonal dipolar interactions

In contrast to its participation in a hydrogen bond, the fluorine-induced dipole-dipole interaction is widely accepted. Examples include the attractive non-covalent dipole-dipole interaction between $\text{C}_{\text{sp}2}\text{-F}$ and amide group and fluorine interaction with a fluorophilic environment provided by H-C-C=O fragment.²⁰ Such interactions are weak, but accounts for many biologically significant phenomena. Therefore, the fluorine-induced dipolar interactions have been exploited for structure-based inhibitor design.

1.1.4 Fluorous effect

Fluorous effect refers to the behavior of perfluoroalkyl chains in avoiding unfavorable interactions between fluorine atoms and other elements, due to the very low polarizability of fluorine. Based on this phenomenon, many technologies for synthesis and separation were able to be developed.²¹ Recently, several new concepts based on the fluorous effect for biological applications have been developed; e.g., Pohl and coworkers introduced fluorous small molecule microarrays for screening carbohydrates.²² Peters and coworkers developed a fluorous affinity tagging technique to enrich peptide samples for MS analysis, which could be applied for

1.1 The unique properties of fluorine

proteomics in the future.²³ The potential fluorous effect in the self-sorting of peptides is discussed in the next section.

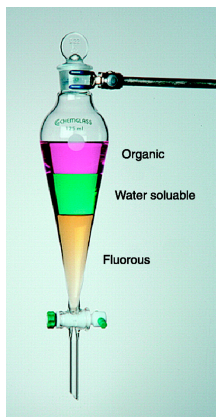


Figure 1.1-6 A typical three-phase liquid extraction shows fluorous effect

Organic solvent (rose), aqueous solvent (green), and perfluorinated solvent (yellow). (Adapted with permission from DeSimone,²⁴ Copyright © 2002, American Association for the Advancement of Science)

1.1.5 Orthogonal reactivity of fluorine

Trifluoridine (trade name: *Viroptic*®) is an antiviral drug. 5-fluorouracil (trade name: *Efudex*) is widely used as a chemotherapeutic agent for malignancies (*Figure 1.1-7*). The design of both mechanism-based enzyme inhibitors is based on the orthogonal reactivity of fluorine.

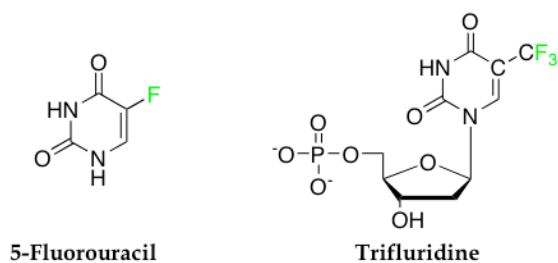


Figure 1.1-7 Structures of 5-fluorouracil and trifluridine

Both drugs cause the irreversible inhibition of thymidylate synthase, which is an important enzyme in nucleic acid metabolism, transferring the methyl group onto deoxyuridine monophosphate to form deoxythymidine monophosphate. *Figure 1.1-8*

1.1 The unique properties of fluorine

illustrates the methylation mechanism of thymidylate synthase and inhibition mechanism of 5-fluorouracil. Catalysis begins with a Michael addition of a cysteine residue in the enzyme onto 5-fluorouracil. The next step is the transfer of a methyl group from a cofactor via β -elimination, forming tetrahydrofolic acid. The 5-fluorouracil interacts with thymidylate synthase and the cofactor in a similar way, but the reaction stops at the elimination step due to the absence of a proton. Thus, the resulting lack of thymidine monophosphate as a precursor for DNA biosynthesis results in blockage of DNA replication (Figure 1.1-8). Trifluridine acts on thymidylate synthase in the same manner.¹⁴

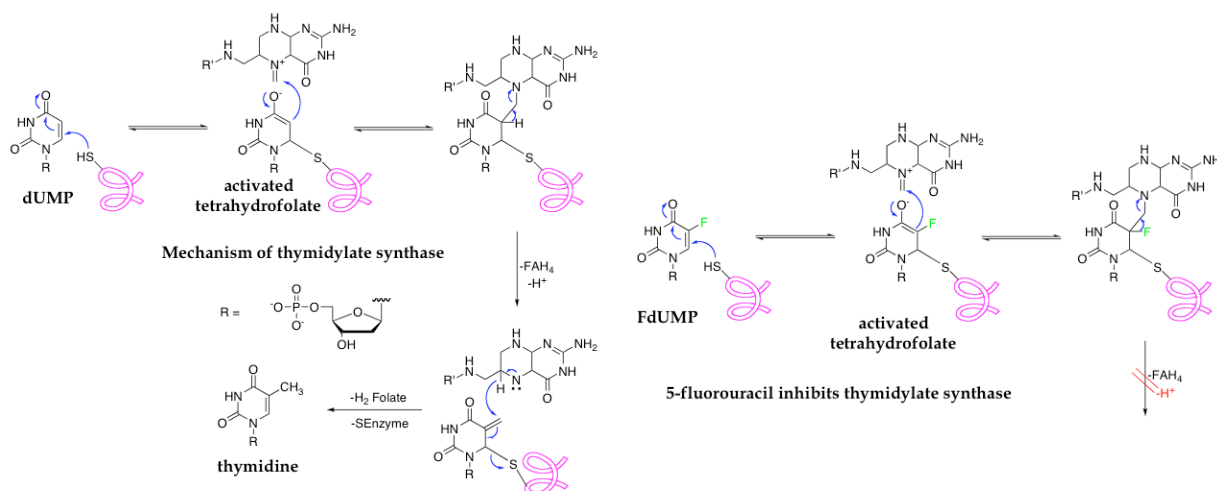


Figure 1.1-8 Schematic representation of methyl transfer mechanism of thymidylate synthase and inhibition mechanism of 5-fluorouracil

1.1.6 Fluorine as label

In addition to fluorine's effects on conformation and physiochemical properties, it is also widely used as a probe for biophysical investigations and clinical studies.

1.1.6.1 Positron emission tomography (PET)

Positron emission tomography (PET) is a powerful nuclear medical imaging technique, which enables the collection of dynamic three-dimensional images of *in*

1.1 The unique properties of fluorine

in vivo processes for diagnostics and clinical studies. Among the common radionuclides, ^{18}F has a relatively long half-life of 110 minutes (^{11}C ($t_{1/2} = 20.4$ minutes), ^{13}N ($t_{1/2} = 9.97$ minutes) and ^{15}O ($t_{1/2} = 2.04$ minutes)) and has better physical characteristics for imaging due to its low positron energy.²⁵ 2-deoxy-2- ^{18}F fluoro-D-glucose (^{18}F FDG) is the best known and most successful commercial radiopharmaceutical for PET (Figure 1.1-9).

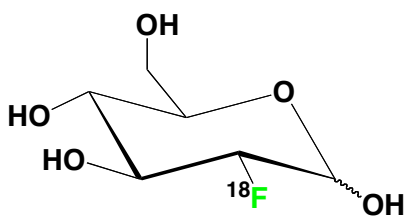


Figure 1.1-9 Chemical structure of 2-deoxy-2- ^{18}F fluoro-D-glucose

The major clinical application of ^{18}F FDG is the diagnosis of tumors, which is based on the increased metabolic activity of cancer cells; recently, PET has also emerged as an important tool for the development of pharmaceuticals. PET images can provide quantitative information on the *in vivo* distribution, pharmacokinetics, and efficiency of receptor binding of potential pharmaceuticals. Because fluorine has drawn much attention in the pharmaceutical industry, and because it can be used directly as a label for clinical studies, a rapidly expanding repertoire of synthetic methods has been established to optimize late-stage fluorination with ^{18}F and avoid unproductive decay.²⁶

1.1.6.2 ^{19}F nuclear magnetic resonance

Due to its natural abundance of 100%, nuclear spin of $1/2$, and high magnetization sensitivity ($\sim 83\%$ of ^1H), ^{19}F is an attractive nucleus for NMR investigation. In particular, ^{19}F NMR is suitable for the study of protein structure and dynamic conformational changes, as well as protein-ligand/protein interactions.²⁷ The CF_3 group is the most common label for ^{19}F NMR used in structural biophysical studies. Incorporation of CF_3 group into peptides and proteins can reduce problems

1.1 The unique properties of fluorine

associated with HF-elimination, racemization, and slow coupling. In addition to applications in ^{19}F NMR, ^{19}F is also used as a probe for magnetic resonance imaging (MRI). Besides the above mentioned nucleus properties of ^{19}F , several advantages, including a broad chemical shift range, bioorthogonality, and the well understood nature of its biodistribution and toxicity, gained through extensive studies of perfluorocarbons as blood substitutes, make ^{19}F attractive as an MRI tracer for cell labeling and *in vivo* imaging.²⁸ ^{19}F has been used for the non-invasion tracing of fluorinated drugs, including 5-fluorouracil for chemotherapy and anesthetics such as isoflurane and methoxyflurane. Quantitative data, including chemical shift, exchangeability, and relaxation processes can be obtained in biological systems.²⁹

This section of the introductory chapter has summarized the unique properties of fluorine and its current biomedical applications. It has become more and more popular to use fluorine in the design of novel small molecules or materials for which extraordinary chemical and physiological properties are desirable. Furthermore, ^{18}F and ^{19}F have been widely used as biophysical markers for diagnostics and drug development.

1.2 Fluorine in protein design

The unique properties of fluorine were discussed in section 1.1. In this section, the effects of fluorination in peptides and proteins are discussed in the context of several examples. Due to the fact that the introduction of fluorine can provide diverse benefits, particularly in pharmaceuticals, agrochemicals, and materials science, the incorporation of fluorine into proteins has emerged as a standard strategy for protein design. Successful synthetic strategies of fluorinated analogues of amino acids³⁰ have enabled various studies for probing protein stability, conformational effects, proteolytic resistance, enzyme mechanisms, and noncovalent interactions.

1.2.1 Effects of fluorine in coiled-coil peptides

The leucine-zipper motif, comprising an α -helical coiled coil, is widely accepted as a model system for investigating protein folding.³¹ It can be simplified as a repetitive sequence or heptad repeat $(abcdefg)_n$, in which hydrophobic amino acids are located at the *a* and *d* positions, charged amino acids at the *e* and *g* positions, and *b*, *c*, and *f* are occupied by polar amino acids. Such systems can be designed *de novo* to investigate various aspects of the self-assembly of helical bundles.³²

Residue-specific or global, replacement of leucine (Leu) with its highly fluorinated analogues 5,5,5-trifluoroleucine (TfLeu), 5,5,5,5',5',5'-hexafluoroleucine (HfLeu), or 4,4,4-trifluorovaline (TfVal) at *a* and *d* positions was carried out by Tirrell and coworkers to study the thermodynamic stability and self-sorting of a rationally designed coiled coil.³³ The fluorinated DNA-binding protein GCN4-bZip showed enhanced protein stability and retention of DNA-binding specificity and affinity (Figure 1.2-1).^{33e}

Kumar and coworkers investigated the substitution of Leu and valine (Val) with TfLeu and TfVal in the hydrophobic core. Protein stability and the formation of oligomers within a membrane environment were surveyed. It was observed that the

1.2 Fluorine in protein design

protein gained stability, and the researchers concluded that this is due to the greater hydrophobicity of the fluorinated side chain.³⁴ In helical bundles studied within a membrane environment, the favorable dimerization of fluorinated helices in micells was demonstrated. Thus, the investigators envision that such fluorination could provide the opportunity to modulate biological processes.³⁵

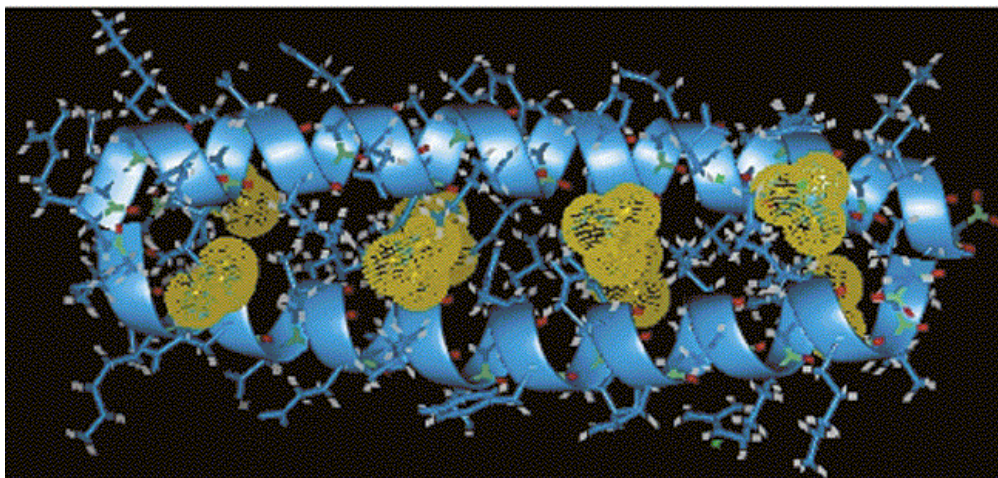


Figure 1.2-1 3D representation of dimeric form of fluorinated GCN4-pld

d positions in the helix are substituted with TfLeu, Yellow spheres represent the van der Waals radii of fluorine atoms (Adapted with permission from Tang *et al.*,^{33e} Copyright © 2001, American Chemical Society).

Marsch and coworkers have systematically studied the fluorous effect in a native protein environment by incorporating HfLeu into a 4- α -helix bundle. The fluorinated peptide exhibits remarkable stability, resistance to proteolysis, and self-sorting behavior. Increased protein stability, with a $\Delta\Delta G_{\text{unfold}}$ of 0.3 kcal/mol per HfLeu residue, was observed, which is in good agreement with the predicted contribution to $\Delta\Delta G_{\text{unfold}}$ of 0.4 kcal/mol per residue from enhanced hydrophobicity. Therefore, the researchers concluded that the enhancement in protein stability is due to the hydrophobic effect rather than the fluorous effect.³⁶ Recently, the packing efficiency of fluorinated side chains in the hydrophobic core of a 4- α -helix bundle was

1.2 Fluorine in protein design

investigated by NMR and X-ray crystallography. The crystallographic data indicate that the packing of fluorinated side chains in the hydrophobic core is indeed more important than the fluororous effect in enhancing protein stability. The increased stability of the helix bundle can most likely be explained by an increase in buried hydrophobic surface area in the fluorinated variant.³⁷

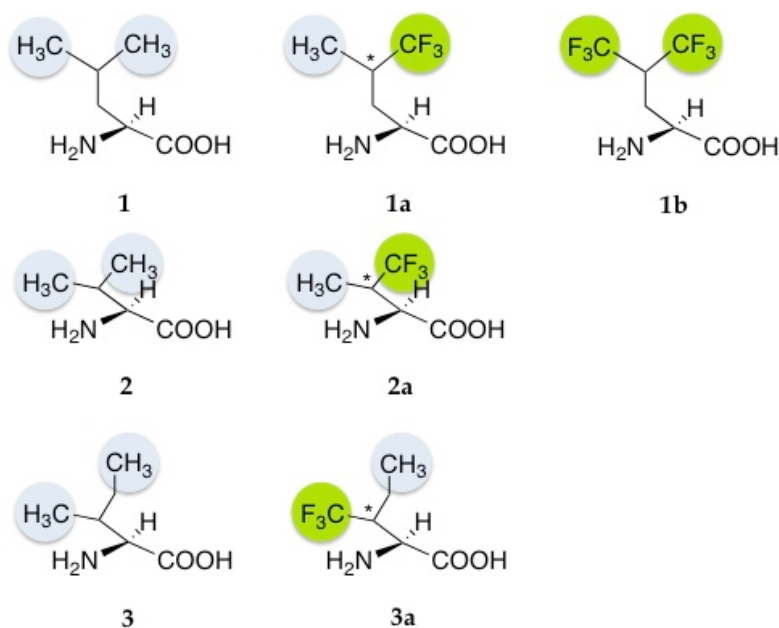


Figure 1.2-2 Schematic representation of Leu, Val, Ile, and their fluorinated analogues incorporated into coiled-coil model peptides

1) Leu, 1a) TfLeu, 1b) HfLeu, 2) Val, 2a) TfVal, 3) Ile, 3a) TfIle

In our group, the effects regarding the single incorporation of fluorinated amino acids into coiled-coil model peptides have been systematically investigated. (*S*)-2-aminobutyric acid (Abu) and its fluorinated analogues (*S*)-2-amino-4,4-difluorobutanoic acid (DfeGly), (*S*)-2-amino-4,4,4-trifluorobutanoic acid (TfeGly), (*S*)-2-amino-4,4-difluoropentanoic acid (DfpGly) have been incorporated into both antiparallel homodimers and parallel heterodimers to probe interactions involving fluorine.³⁸ Initially, these studies were based on the assumption that the steric size of

1.2 Fluorine in protein design

a trifluoromethyl group is comparable to that of an isopropyl group. The results of thermal denaturation experiments and MD simulations indicate that the effects of partially fluorinated side chains are strongly dependent on the packing and orientation of the fluorinated side chains within the hydrophobic core. The peptides containing these fluorinated amino acids were further characterized in phage display screening for the preferred interaction partners composed of only canonical amino acids. It was found that in spite of their different fluorine content, all three fluorinated amino acids tested favor pairing with hydrophobic amino acids.³⁹

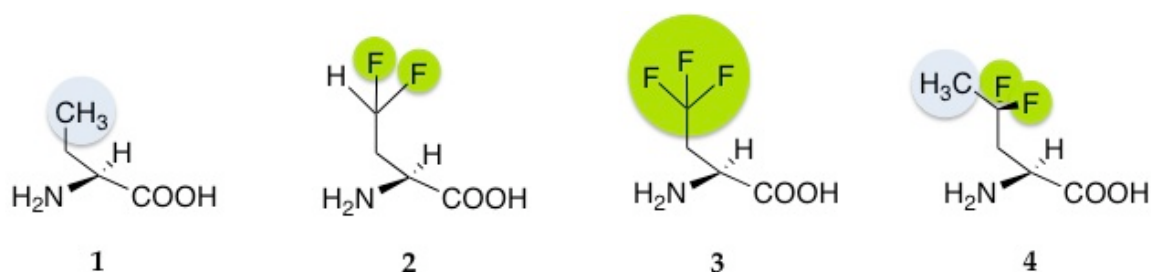


Figure 1.2-3 Schematic representation of (*S*)-2-aminobutyric acid and its fluorinated analogues used in the studies of our group

1) Abu, 2) DfeGly, 3) TfeGly, 4) DfpGly

1.2.2 Secondary structure propensities of fluorinated amino acids

As discussed in section 1.2.1, the introduction of extensively fluorinated amino acids into the hydrophobic core of coiled-coil model peptides can enhance protein stability, which is probably due to elevated hydrophobicity and increased buried hydrophobic surface area. Nevertheless, the overall conformational stability of proteins is not only determined by the hydrophobic effect, but also by other factors including electrostatic interactions between charged residues, hydrogen bonds, and the secondary structure propensities of amino acids. It has been argued that some fluorinated amino acids exhibit low helix propensities and might be more suitable for stabilizing a β -sheet. Cheng and coworkers demonstrated that HfLeu, (*S*)-2-amino-4,4,4-trifluorobutyric acid (Atb), and (*S*)-pentafluorophenylalanine have lower helix propensities and higher β -sheet propensities by incorporating these fluorinated

1.2 Fluorine in protein design

amino acid analogues into a solvent-exposed position on an internal strand of protein G B1 domain.⁴⁰ Raleigh and coworkers incorporated diastereomeric TfVal into the β -sheet core structure of the N-terminal domain of the ribosomal protein L9 (NTL9) and observed increased protein stability; however, it was unclear in this study whether this was due to more favorable hydrophobic interactions or greater β -sheet propensities of the fluorinated amino acids.⁴¹ Recently, our group determined the helix propensities of a series of fluorinated amino acids and probed the effects on α -helix and β -sheet formation. The results show that (3S)-TfV, (3R)-TfV, and (3S)-TfIle have very low helix propensities, and Abu, MfeGly, DfeGly, and TfeGly exhibit decreasing α -helix propensities with increasing numbers of fluorine atoms.⁴²

1.2.3 Fluorination of antimicrobial peptides

Fluorinated amino acids have been incorporated into antimicrobial peptides (AMPs) to probe their biological activities. AMPs are class of genetically encoded short cationic amphipathic peptide that act on bacterial cells by selectively disrupting cell membranes.⁴³

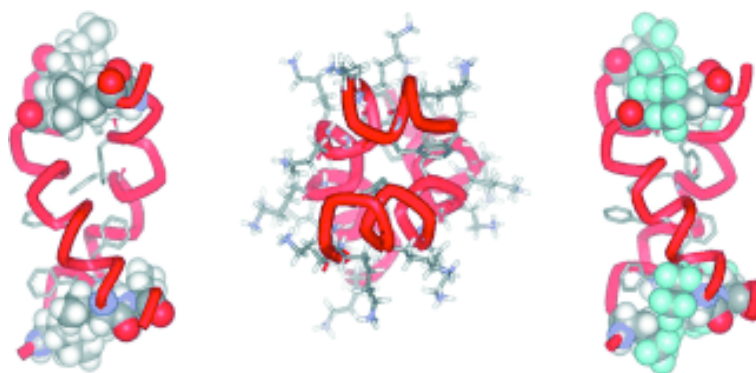


Figure 1.2-4 Schematic representation of MSI-78 dimer

Side view, view along the helical axis, and model of fluorinated MSI-78, Leu and Ile are shown in CPK rendering. (Adapted with permission from Gottler *et al.*,⁴⁴ copyright © 2008 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim)

1.2 Fluorine in protein design

Marsh and coworkers synthesized and biologically characterized a fluorinated AMP, MSI-78, by substituting Leu and Ile with HfLeu. MSI-78 adopts a random coil conformation when free in solution, but forms a dimeric anti-parallel coiled-coil in the presence of lipid bilayers (Figure 1.2-4). Upon fluorination, MSI-78 shows a great increase in proteolytic resistance and retains its broad spectrum of antimicrobial activity. Moreover, the fluorinated MSI-78 acts selectively upon certain bacterial strains. It was assumed that the incorporation of fluorine into MSI-78 enhances hydrophobic interactions in the dimer, which results in observed enhancement of proteolytic resistance.⁴⁴

In another study from the same group, two valines were substituted with Leu and HfLeu in a member of the protegrin family of antimicrobial peptides. The experimental data show that the Leu-containing AMP was significantly more active than the parent peptide, while HfLeu-containing AMP lost its antibiotic activity. Furthermore, the latter substitution seems to alter the stoichiometry of peptide-lipid interaction, which was confirmed by ITC measurements.⁴⁵

1.2.4 An example of fluorine in aromatic stacking effects

Recently, Gao and coworkers systematically investigated the stacking effects of fluorinated aromatic side chains by the use of a designed α_2D model protein.⁴⁵⁻⁴⁶ α_2D is a *de novo* designed polypeptide containing 35 residues that folds into a dimeric helix bundle with a remarkably cooperative, reversible folding behavior. Two phenylalanine (Phe) side chains located in the middle of the helix bundle stack face to face.⁴⁷ Analogues of Phe with various numbers of fluorine atoms on the aromatic ring were used to replace these Phe residues to study aromatic stacking energetics (Figure 1.2-5). In earlier studies, Gao and coworkers had demonstrated that the quadrupolar interaction between the pentafluor and hydrocarbon phenyl rings stabilizes heterodimer formation.^{46c} In recent systematic studies, the thermodynamic stabilities of α_2D double mutants (both positions 10 and 29 substituted with various

1.2 Fluorine in protein design

fluorinated Phe residues) and single mutants (position 29 substituted with various fluorinated Phe residues) were characterized. Among the double mutants, the species substituted with the pentafluoro phenyl ring showed the highest stability. The experimental data also revealed that α_2D variants harboring regioisomers of tetrafluoro phenyl rings have different thermodynamic stabilities ($Z^o > Z^m > Z^P$) (Figure 1.2-5). Because the thermal stability of α_2D varies directly with the magnitude of the dipole moment of the phenyl side chain, it was concluded that dipole-dipole coupling stabilizes the formation of homodimers. The thermal stabilities of the single mutants revealed a preference for stacking interactions between fluorinated Phe analogues and native Phe, with Z^o showing the highest stability, indicating a synergistic effect of hydrophobicity and dipolar interactions in a protein environment.^{46a}

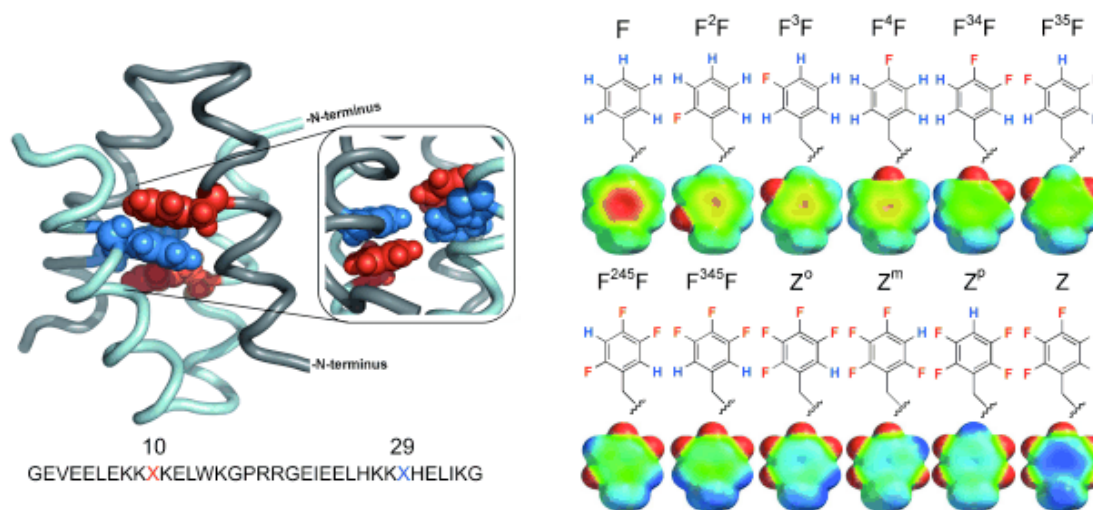


Figure 1.2-5 Schematic representation of the dimeric α_2D model protein, the monomers presented in gray and light cyan (left); the fluorinated analogues of Phe incorporated into α_2D and their electrostatic potential maps (blue is positive, red is negative) (right)

(Adapted with permission from Pace *et al.*,^{46a} Copyright © 2012, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim)

1.2.5 Effects of global substitution with fluorinated amino acids

Fluorinated analogues of amino acids have also been incorporated into whole proteins residue-specifically to explore the effects of fluorination on the overall conformation and stability of proteins, as well as the catalytic properties of enzymes.

Diastereomeric TfLeu was substituted for Leu in chloramphenicol acetyltransferase (CAT). A slight enhancement in protein stability in secondary structure, but decrease in trimeric structure, and misaggregation of the protein by an increased temperature were observed.⁴⁸ Similarly, protein misfolding was observed when diastereomeric TfLeu was incorporated into green fluorescent protein (GFP).⁴⁹ Recently, Tirrell and Montcalé demonstrated that the thermostability of CAT could be increased through a combination of introducing fluorinated building blocks and applying directed evolution.^{33b} These studies indicated that the global substitution of hydrophobic side chains with fluorinated analogues in proteins does not always lead to stabilization, even though a high degree of fluorination is commonly thought to confer hydrophobicity and the size of fluorine is slightly larger than hydrogen. The intrinsic properties of protein sequence determine the protein microenvironment, and fluorine is not found in the natural occurring proteins. Not all proteins and primary structures are tolerant to fluorination. Nevertheless, a reassignment of protein sequence through directed protein evolution can compensate for the loss in protein stability and activity.

2-fluorotyrosine (2-fTyr) and 3-fluorotyrosine (3-fTyr) were incorporated into GFP globally to study protein stability and spectroscopic properties (*Figure 1.2-6*). Though fluorination did not change the overall protein stability, it did lead to changes in the absorbance and fluorescence spectra.⁵⁰ A similar spectroscopic changing was also observed when fluorinated analogues of tryptophan (Trp) were globally introduced into GFP. Protein stability was slightly enhanced, and the mutant protein possessed a red shift in emission. These results indicate that the fluorination can reorganize microenvironments of protein chromophores and lead to spectral shifts.⁵¹

1.2 Fluorine in protein design

Marx and coworkers investigated the effects of the global substitution of Pro with (4*R*)-fluoroproline,⁵² and Met with trifluoromethionine (TfMet)⁵³ in the KlenTaq DNA polymerase from *Thermus aquaticus*. In both studies, the mutant KlenTaq retained catalytic activity, fidelity, and stability. The incorporation of TfMet into the enzyme provides the opportunity to monitor the catalytic process of the enzyme by ¹⁹F NMR.

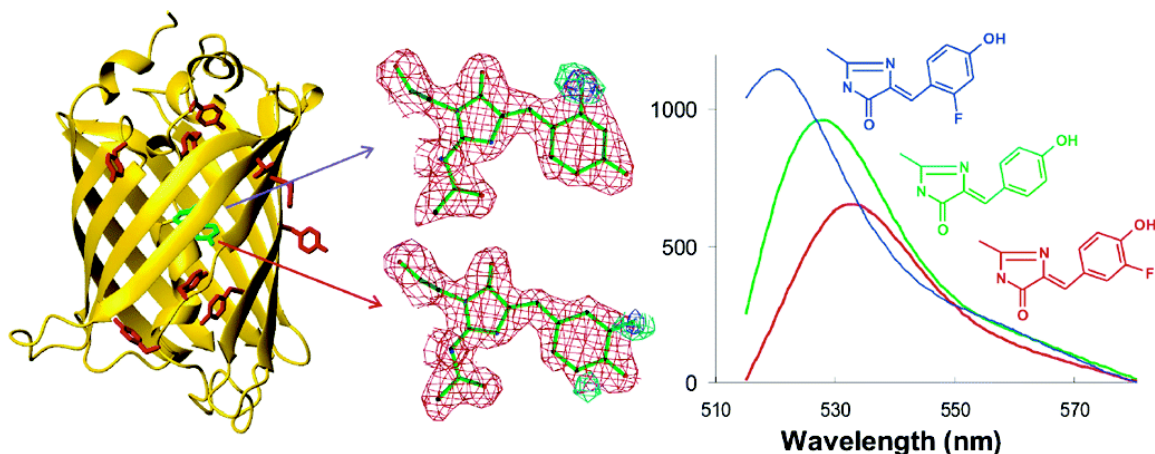


Figure 1.2-6 Schematic representation of probing the stability and spectroscopic properties of GFP by the incorporation of fluorinated analogues of tyrosine

The general structure of GFP, which contains 11 tyrosine residues; tyrosine 66 is an integral part of the chromophore; structure of 2-fluorotyrosine and 3-fluorotyrosine used in this study (**left**), fluorination results in spectral shift (**right**). (Adapted with permission from Pal *et al.*,^{50b} Copyright © 2015, American Chemical Society)

In summary, a great deal of effort has been made to understand the general effects of fluorine on protein stability. Although the incorporation of fluorinated aliphatic amino acids into the hydrophobic core of proteins generally enhances protein stability, it remains difficult to predict how arbitrary numbers of fluorinated building blocks at arbitrary locations will influence the properties of whole proteins.^{4, 54} The question of improving stability in the context of global substitution relies on whether the primary and tertiary structures of the protein environment and assignment of intramolecular interactions is distributed in space. The secondary structure propensities of fluorinated amino acid analogues must also be considered an

1.2 Fluorine in protein design

important factor in determining protein stability, especially when the residue being substituted lies within a well-defined secondary structural element. The incorporation of fluorine into membrane-active peptides can dramatically enhance their biological activities. Fluorination is a powerful tool for modulating proteolytic resistance and membrane-interacting behavior. Such initial studies can be further applied for peptide-based therapeutic purposes. The mechanisms of such effects are probably due to an enhanced hydrophobic effect, for substitutions that lie within the hydrophobic core of proteins or dipole-dipole or dipole-induced dipole or fluororous interactions, i.e. preferred van der Waals type interactions between fluorocarbon residues. Fluorinated aromatic amino acids are a useful tool for modulating stacking interactions. Protein design involving fluorination extends the scope of protein-protein interactions and more systematic studies are needed to clearly define which types of noncovalent interaction contribute to what extent regarding key protein properties.

Numerous studies have used fluorinated aromatic and aliphatic amino acids, fewer have relied on fluorinated analogues of polar and charged amino acids. This is mainly due to the relative difficulty of the synthesis of such amino acids, and their low efficiency of incorporation in solid-phase peptide synthesis (SPPS). Obviously, charged and polar amino acids also play important roles in protein stability and function via their respective electrostatic and polar interactions. For example, the active sites of enzymes frequently present polar and charged amino acids that are important for catalysis. Therefore, it is relevant to explore the novel interactions of fluorine within such environments.

Chapter 2

Aim of the work

Aim of the work

Bioorthogonal fluorine-containing analogues of aliphatic hydrophobic amino acids, derivatives of aminobutyric and aminopentanoic acid, had been previously incorporated into model peptides to systematically investigate how this unique element influences numerous properties. These results showed that residues containing fluorinated side chains often, but not always, impart peptides with desirable characteristics, and that they do not perturb the native structure. Their effects depend upon the microenvironment in which they are found, and the number of C-F bonds present.

To expand our understanding of the fundamental properties of fluorine within a native protein environment, these amino acids must also be incorporated into larger globular proteins with diverse tertiary structures. The aim of the current study was to site-specifically incorporate these amino acids into a well characterized protein by means of two different methods, namely *in vitro* translation based non-sense suppression and total chemical synthesis. We chose basic (bovine) trypsin inhibitor (BPTI) and introduced α -aminobutyric acid (Abu) and its fluorinated derivatives into its P₁ position. The produced mutant inhibitors will be characterized by biological and biophysical methods to determine thermal stability, inhibitory activity, and high-resolution three-dimensional structure. A novel insight into fluorine's behavior within a natural tertiary protein environment will therefore be gained.

Chapter 3

Basic (bovine) pancreatic trypsin inhibitor

3.1 General properties of BPTI

Basic (bovine) pancreatic trypsin inhibitor (BPTI), also known as aprotinin, trade name Trasylol (*Bayer AG*), belongs to the Kunitz-type serine protease inhibitor family. BPTI is a small globular protein consisting of 58 amino acids that folds into a compact pear-shaped structure with two β -strands (Ile18 - Asn24 and Leu29 - Tyr35) forming a twisted anti-parallel β -sheet and two α -helical regions (Pro2 - Glu7 and Ser47 - Gly56). Three disulfide bonds, Cys5 - Cys55, Cys14 - Cys38, and Cys30 - Cys51 stabilize the overall conformation of the protein. BPTI shows broad inhibitory activity, particularly against serine proteases such as α -chymotrypsin, β -trypsin, plasmin, and elastase-like enzymes. In addition, BPTI also impairs K^+ transport by Ca^{2+} activated K^+ channels, and the nitric oxide biosynthesis pathway.⁵⁵

As a model protein, BPTI has been extensively studied by structural biologists. It was one of the proteins for which a crystal structure was solved in the early days of protein crystallography, and later its crystal structure was again determined at high resolution.⁵⁶ In the late 1970s, BPTI was the first macromolecule to which molecular dynamics simulations were applied.⁵⁷ In the early 1980s, the structure of BPTI was confirmed by NMR spectroscopy, and this was also a first for a larger biomolecule.⁵⁸ Moreover, BPTI is the protein for which folding studies initially revealed general information on the arrangement of the peptide chain in three dimensions during the folding process.⁵⁹

Therapeutically, BPTI, also known as aprotinin, was initially used in the treatment of acute pancreatitis, inhibiting the enzymes of this gland and halting the self-destruction of the pancreas. Later, as an antifibrinolytic agent, BPTI was used to reduce bleeding during cardiac surgery and liver transplantation. Due to the increased risk of death associated with its use, Bayer AG withdrew aprotinin (trade name Trasylol) in 2008.⁶⁰

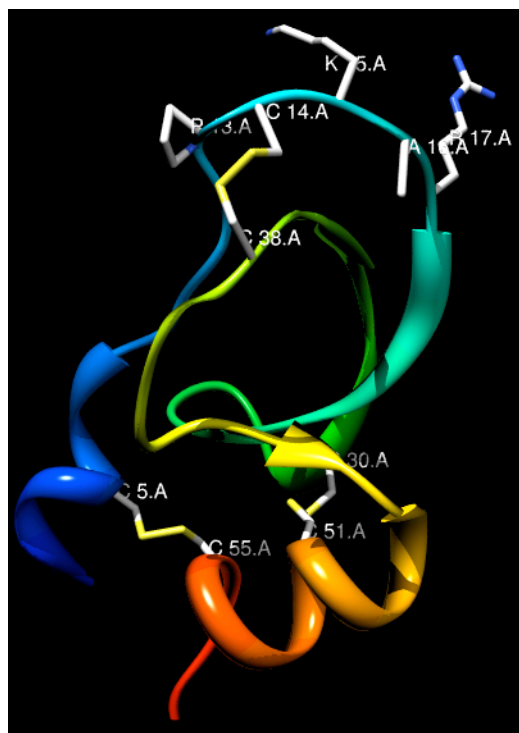


Figure 3.1-1 Ribbon representation of BPTI

The protease interacting loop and three disulfide bridges are labeled. The picture was generated by use of UCSF Chimera with crystal data of protein data bank, (PDB code 1bpi).⁶¹

3.2 Protein stability and basis of proteases-BPTI interaction

BPTI is a protein with unusually high thermal stability (very high specific stability $\Delta G_{st} = -8.0$ J/g at 25 °C) and highly resistant towards common protein denaturants, for example, 6 M guanidinium chloride (GdmCl).⁶² Similar to many globular proteins, the unfolding of BPTI is well approximated by a two-state transition. The thermal denaturation of BPTI at neutral pH up to 100 °C is not sufficient to completely unfold the protein, but results in hydrolysis of the peptide bond (between Pro2 - Asp3 and Asp3 - Phe4).⁶³ Efficient chemical denaturation requires guanidinium thiocyanate (GdmSCN) at high concentrations, 3-4 M.⁶⁴ Studies in which Lys15 (P₁ position) was substituted with other canonical amino acids indicated a possible folding mechanism regarding this solvent-exposed residue. The reverse hydrophobic effect and secondary structure propensities of the amino acids might contribute to protein stability regarding the P₁ position.⁶⁵

3 Basic (bovine) pancreatic trypsin inhibitor

The serine protease-BPTI complexes were also extensively investigated to gain an understanding of protein-protein interactions.⁶⁶ BPTI interacts rapidly with serine proteases and forms stable complexes. It has been reported that the trypsin-BPTI complex has a dissociation constant of 6×10^{-14} and a half-life of about 13 weeks at 25 °C and pH 8. At room temperature, the trypsin-BPTI complex is even resistant to 8 M urea and 6 M GdmCl.⁶⁷ Structural analysis has revealed that the binding loop (Pro13 (P₃) – Arg17 (P₂')) of BPTI is the main contact region with the protease. This interacting loop recognizes the active site of the enzyme and forms an anti-parallel β -sheet. The P₁ position (Lys15) of the interacting loop is a true “hot spot” of the proteases-BPTI contact area. Crystallographic data has revealed that the charged side chain of Lys15 forms a tight polar interaction with the negatively charged Asp189 residue of trypsin, while in the chymotrypsin-BPTI complex Lys15 forms instead hydrogen bonds with Gly216 and Ser217 of the enzyme.^{66b,68} Studies in which Lys15 was replaced with other canonical amino acids by site-directed mutagenesis also confirmed the importance of Lys15. For instance, replacement of Lys15 with Gly decreases the K_a for trypsin 10^9 fold, and for chymotrypsin 10^3 fold, respectively.⁶⁹ Krowarsch *et al.* demonstrate that electrostatic interaction between P₁ side chain of BPTI and Asp189 of trypsin at the bottom of S₁ pocket and additional binding force from the burial hydrophobicity of the P1 side chain are the main sources of binding energy in trypsin-BPTI interaction. In the case of chymotrypsin-BPTI interaction, the association energy originates predominantly from the hydrophobicity of P1 side chain of BPTI. And it is more complicated with branching, charging, and polarity of side chain.^{69b,70}

Chapter 4

Probing the site-specific incorporation of fluorinated amino acids into BPTI by *in vitro* non-sense suppression

4.1 Protein engineering with non-canonical amino acids

With the two exceptions of pyrrolysine and selenocysteine, the genetic code enables the incorporation of only the 20 canonical amino acids in all organisms.⁷¹ It is clear that proteins need additional functional groups to carry out their natural functions. In the cell this is realized by numerous post-translational modifications including glycosylation, phosphorylation, hydroxylation, and collaboration with coenzymes and metal ion cofactors in catalytic processes. Therefore, it is useful to expand the chemical tool-kit for proteins and to explore protein structure and function by means of *in vitro* and *in vivo* methods, e.g., incorporation of non-canonical amino acids to enhance the protein stability, introduction of novel functional group at catalytic center of enzymes to modulate the catalytic activity, and incorporation of biophysical label into proteins. Both biological and chemical strategies have been developed to achieve this goal.⁷² In this section, the fidelity of protein translation and the recently developed methodologies for incorporation of non-canonical amino acids into proteins by means of biochemical procedures are discussed. The incorporation of several non-canonical amino acids in which we are interested was carried out by means of *in vitro* non-sense (*amber*) suppression.

4.1.1 Fidelity of protein translation

The central *dogma* of molecular biology refers to the flow of genetic information within a biological system from deoxyribonucleic acid (DNA) through messenger ribonucleic acid (mRNA) to three-dimensional folded functional proteins. This information flow requires three fundamental polymerization reactions with remarkable accuracy, having misincorporation rates in 10^{-10} - 10^{-8} in DNA semi-replication, 10^{-4} in transcription, and 10^{-4} - 10^{-3} in protein translation.⁷³ In general, the specificity of enzymatic reactions is caused by favorable molecular interactions, which take advantage of differences in free energy by recognition and interaction with cognate and non-cognate substrates, at catalytic sites. The great accuracy of

4.1 Protein engineering with non-canonical amino acids

DNA-replication is determined by the correct recognition of Watson-Crick base pairs and the proofreading functions of DNA polymerases.^{73c} Compared to the known detailed mechanisms of DNA polymerases from numerous studies, the fidelity mechanisms of RNA polymerases remain poorly understood. However, results from earlier studies also reveal that several mechanisms including selection of correct substrates (error prevention), induction of off-line states by mismatches (error detection), and backtracking and RNA cleavage (error removal) are also involved in the fidelity of transcription.⁷⁴ The ribosome, a large and complex molecular machine in cells in which the bulk of protein synthesis takes place, differs in its catalytic mechanism totally from DNA or RNA polymerases. For instance, the catalytic site and substrate recognizing site are far away from each (the peptidyl transferase site is about 70 Å away from the decoding site).⁷⁵ Experimental data indicate that the ribosome alone is unable to distinguish between amino acids; however, the overall translation fidelity is ascertained in several critical steps during the whole protein synthesis processes. These include: the activation of amino acids by highly specific enzymes, aminoacyl-tRNA synthetases (aaRSs), aminoacyl-tRNA selection by kinetic proofreading, and decoding through Watson-Crick base pairing between tRNA and mRNA.⁷⁵⁻⁷⁶

4.1.1.1 Aminoacyl-tRNA synthetases and amino acid activation

aaRSs catalyze the attachment of amino acids onto the 3'-end of tRNAs. It is worth noting that only eukaryotes and a handful of bacteria have the full set of 20 aaRSs.⁷⁷ The aaRSs can be categorized in two classes: class I, whose members have a catalytic domain that contains a dinucleotide or Rossmann fold, e.g., TyrRS;⁷⁸ and class II, whose members share a common domain that is organized as a seven-stranded β -sheet flanked by α -helices. The structures of class II enzymes were first determined with SerRS and AspRS.⁷⁹ The catalytic mechanisms of aaRSs are thought to be similar and divisible into two steps: 1) an α -carboxylate oxygen of amino acid attacks the α -phosphorus of ATP (Mg^{2+} as cofactor) and forms the mixed anhydride linkage of the

4.1 Protein engineering with non-canonical amino acids

aminoacyl adenylate with the release of pyrophosphate; 2) either the 2'- or 3'-hydroxyl group of the *cis*-diol at the 3'-terminal of tRNA attacks the carbonyl carbon of the adenylate and forms aminoacyl-tRNA with the release of AMP (Figure 4.1-1). Class I and II aaRSs differ in the aminoacylation site. The synthetases of class I aminoacylate the 2'-OH of the ribonucleotide at the 3'-end of tRNAs, while, with the exception of PheRS, members of class II aminoacylate the 3'-OH.




 Ling J, et al. 2009.
Annu. Rev. Microbiol. 63:61–78

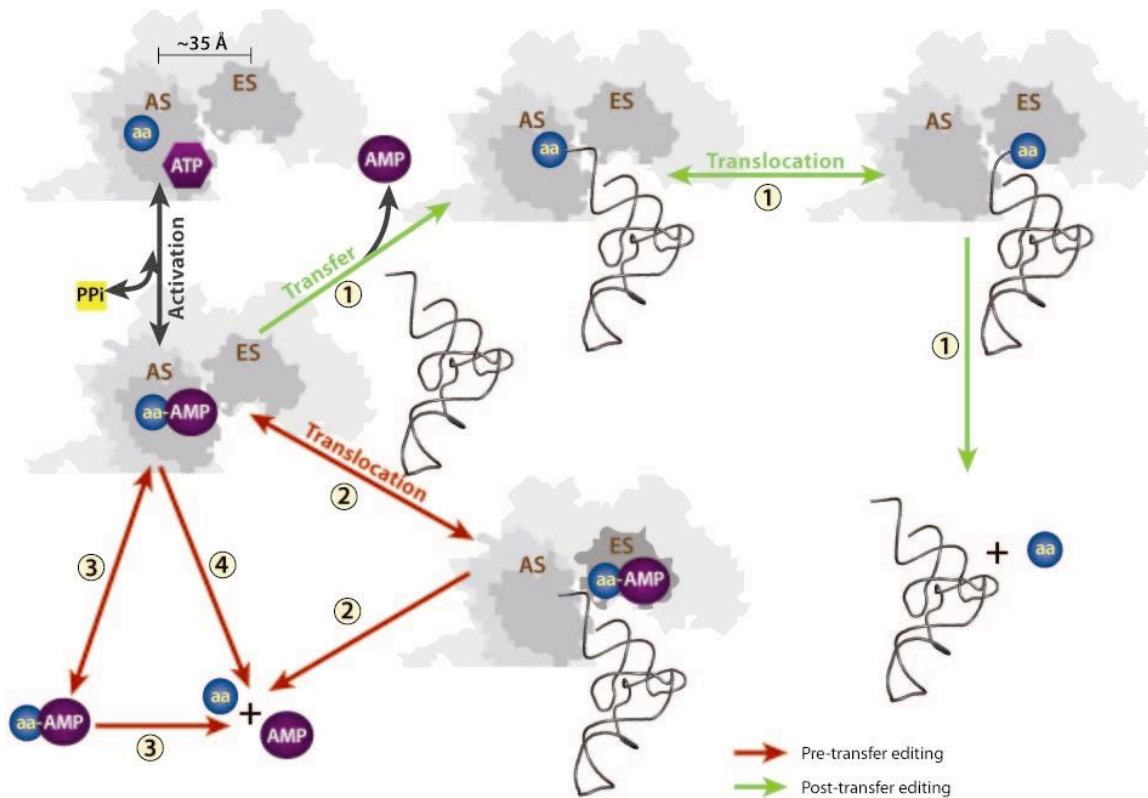
Figure 4.1-1 Illustration of aminoacylation of tRNA catalyzed by aminoacyl-tRNA synthetases (aaRS)

The amino acids are first activated at the aaRS active site with ATP to form aa-AMP, then transferred to the 3'-end of tRNA. (Adapted with permission from Ling *et al.*,^{76c} Copyright © 2009, Annual Reviews)

Actually, many aaRSs have difficulty in distinguishing near-cognate amino acids, and thus a strong proofreading function was evolved to decrease the misacylation rate. The proofreading mechanism is either tRNA-dependent or tRNA-independent and occurs pre- or post-transfer in a manner that is either *trans* or *cis*.^{76c, 77} As presented in Figure 4.1-2, the editing pathway includes: (1) posttransfer editing: the incorrect amino acid is first charged onto tRNA and then translocated from the activation site to the editing site for hydrolysis; (2) the misacylated tRNAs are directly translocated to the editing site for hydrolysis; (3) selective hydrolysis in which the incorrect aa-AMP is subjected to spontaneous hydrolysis; (4) active site hydrolysis.^{76c} The correct aminoacylation and hydrolysis of misaminoacylation are

4.1 Protein engineering with non-canonical amino acids

mainly controlled by the specificity of the aaRSs to tRNAs and also the discrimination in geometry and physical properties of amino acid side chains such as size, polarity, and hydrophobicity. For IleRS, ValRS, LeuRS, ProRS, ThrRS, and AlaRS, because the side chains of these amino acids share very similar size and hydrophobicity, a “double-sieve” mechanism was evolved. The binding sites of posttransfer editing domains are generally built to exclude the larger cognate amino acids.^{76c, 80}




 Ling J, et al. 2009.
Annu. Rev. Microbiol. 63:61–78

Figure 4.1-2 Editing mechanisms of aaRSs against misacylated tRNAs

AS = activation site, ES = editing site; 1) posttransfer editing, the incorrect amino acid is first attached to tRNA, then translocated from active site to editing site for hydrolysis; 2) the misacylated aa-AMP is directly translocated to the editing site of aaRS for hydrolysis; 3) selective hydrolysis, the misacylated aa-AMP is expelled into solution and subjected for spontaneous hydrolysis; 4) active site hydrolysis, the misacylated aa-AMP is hydrolyzed at the active site before release. (Adapted with permission from Ling *et al.*,^{76c} Copyright © 2009, Annual Reviews)

4.1 Protein engineering with non-canonical amino acids

The tRNA identity defines specific nucleotide regions particular to each amino acid's isoaccepting group. The nucleotide sequence of tRNAs have been extensively investigated. The experimental data indicate that identity regions are often found in the anticodon and in acceptor-stem base pairs at both extremities of the tRNAs.⁸¹ And in rare cases, the post-translational modification of a nucleobase can also be crucial to recognition between an aaRS and a tRNA.⁸²

4.1.1.2 Selection by decoding

Correct base pairing between the anti-codon loop of tRNA and mRNA is the fundamental contributing factor in the selection of aa-tRNAs to participate in protein synthesis. Although the energy difference in base pairing of cognate tRNAs and near-cognate tRNAs (single mismatch) is minimal, the accuracy of this selection has an error rate of 10^{-4} - 10^{-3} . It has been suggested that the "decoding site" of the ribosome recognizes the geometry of codon-anti-codon base pairing and discriminates sterically against mismatches.⁸³

4.1.1.3 Selection by kinetic proofreading and induced fit

The proposal of "kinetic proofreading" was introduced by Ninio and Hopfield in 1970s independently.⁸⁴ This theory is based on the principal that aa-tRNAs are delivered to the A-site of the ribosome as a ternary complex with GTP and the elongation factor EF-Tu, and that the hydrolysis of GTP enables the irreversible reaction (*Figure 4.1-3*). The tRNAs can be dissociated either in the initial selection or the proofreading process. During the initial selection process, the ternary complex translocates to the A-site of the ribosome. Experimental data have shown that all cognate aa-tRNAs exhibit similar binding affinities to the A-Site while near-cognate and non-cognate aa-tRNAs bind less well. This difference in binding free energy results in different stability of substrates to the A-site of ribosome.⁸⁵ It was assumed that a conformational change in the ribosome is induced by cognate aa-tRNAs.^{76d, 86}

4.1 Protein engineering with non-canonical amino acids

This is supported by experimental studies that have shown that cognate tRNAs have lower dissociation rates and much faster forward rates in terms of GTPase activation and accommodation than the near-cognate tRNAs.⁸⁷

Hydrolysis of GTP and dissociation of EF-Tu are involved in the next step. Discrimination arises from the different dissociation rates between the cognate-aa-tRNAs and non-cognate-aa-tRNAs in the context of conformational change of the active site of EF-Tu upon GTP hydrolysis. A detailed model of tRNA selection, which is supported by numerous experiments including FRET and single-molecule approaches have been summarized and discussed in Green *et al.*⁷⁵

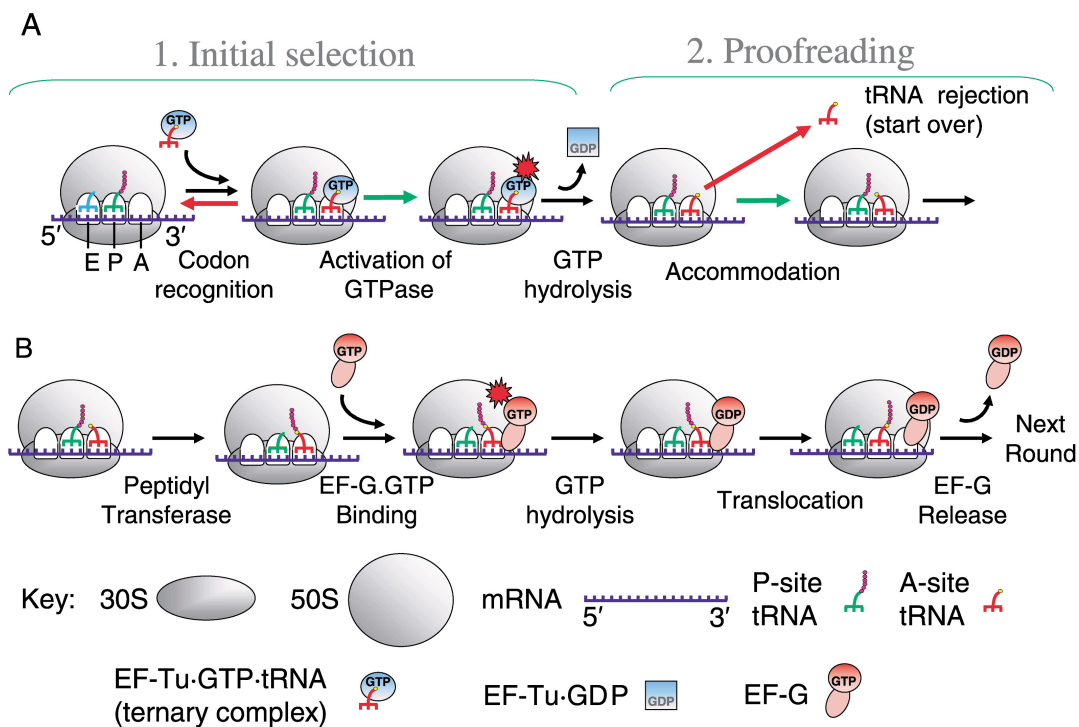


Figure 4.1-3 Schematic representation of translation elongation and peptidyl transfer

(A) initial selection and proofreading, (B) peptidyl transfer and translocation (Adapted with permission from Ramakrishnan,^{76d} Copyright © 2002, Elsevier)

In addition to the kinetic proofreading hypothesis, amino acid specificity and tRNA diversity are also critical factors in determining the binding of aa-tRNAs to EF-Tu. It was expected that the non-specific binding protein EF-Tu should bind to all cognate

4.1 Protein engineering with non-canonical amino acids

aa-tRNAs with similar affinities.⁸⁸ However, additional experimental data showed that four cognate aa-tRNAs bind EF-Tu within a narrow range of binding affinities while 12 misacylated-tRNAs have highly differentiated binding affinities to EF-Tu. It was proposed that during the evolution process each aa-tRNA was selected to bind uniformly to EF-Tu, but the tRNA sequences have evolved to adjust their affinities for EF-Tu according to each canonical amino acid.^{76b} This phenomenon was also observed in aa-tRNA binding to the A-site of the ribosome. The dissociation rates of eight different tRNAs from the A-site were identical when aminoacylated but quite different when deacylated.^{85, 89}

4.1.2 Residue-specific incorporation of non-canonical amino acids into proteins

Residue-specific incorporation refers to the substitution of all or a fraction of a particular canonical amino acid with structurally similar non-canonical amino acids. Practically, this can be achieved by the use of auxotrophic bacteria starved for the canonical amino acid to which the analogue is added as a supplement.⁹⁰ On the one hand, organisms have evolved to adapt to their living conditions; on the other hand, the structural similarity of the non-natural analogues may enable circumvention of the fidelity mechanisms. A fraction of these misacylated-tRNAs can survive and even be delivered to the ribosome and incorporated into peptide chains. To minimize the global effects on the expression host and increase the substitution rate in the desired protein, a selective pressure incorporation (SPI) method has been developed.^{72a, 91} In this way, the amino acid pool and fermentation parameters can be carefully controlled, and reassignment of the sense-codon enables a good incorporation rate and high yield of the mutant protein. The SPI method can be further optimized through enhancing enzymes of the translation system, e.g., by altering the aaRS activity.⁹²

4.1 Protein engineering with non-canonical amino acids

This approach has been used to substitute about 60 non-canonical amino acids for their canonical counterparts, e.g., fluorophenylalanine for phenylalanine (Phe), selenomethionine for methionine (Met), tryptazan for tryptophan (Trp), trifluoroleucine for Leu, and canavanine for arginine (Arg).^{72c} One disadvantage of this method is that all such amino acids in the protein are replaced by the analogue, and such a multiplicity of replacement can't provide information about one specific site. Furthermore, in most cases, the canonical amino acid is also partially incorporated into the protein, and separating these mixed populations can be challenging.

4.1.3 *in vitro* non-sense suppression

One method that enables the site-specific incorporation of non-canonical amino acids is a stop codon-suppression based *in vitro* approach. This approach takes advantage of the recognition tolerance of the ribosome to aa-tRNAs, and also the degeneracy of the three stop codons: TAG (*amber*), TAA (*ochre*), and TGA (*opal*). Indeed, these three stop codons do not encode any amino acids and any one of them is able to terminate the protein synthesis when they bind to release factors.⁹³ Thus, when one stop codon is used for translation, another two can be used as “blank codons”.

Practically, a blank codon (non-sense codon) was introduced into the target protein by site-directed mutagenesis. Non-canonical amino acids are charged onto an engineered suppressor tRNA in which the anti-codon loop has been modified to form Watson-Crick base pairs with the non-sense codon. The aminoacylation of non-canonical amino acids of suppressor tRNA can be achieved by a semi-synthetic approach or by the use of engineered aaRSs. Protein expression can be carried out by means of *in vitro* synthesis (cell-free protein synthesis), and has been shown to be compatible with microinjection of chemically aminoacylated suppressor tRNAs into *Xenopus* oocytes (Figure 4.1-4).⁹⁴

4.1 Protein engineering with non-canonical amino acids

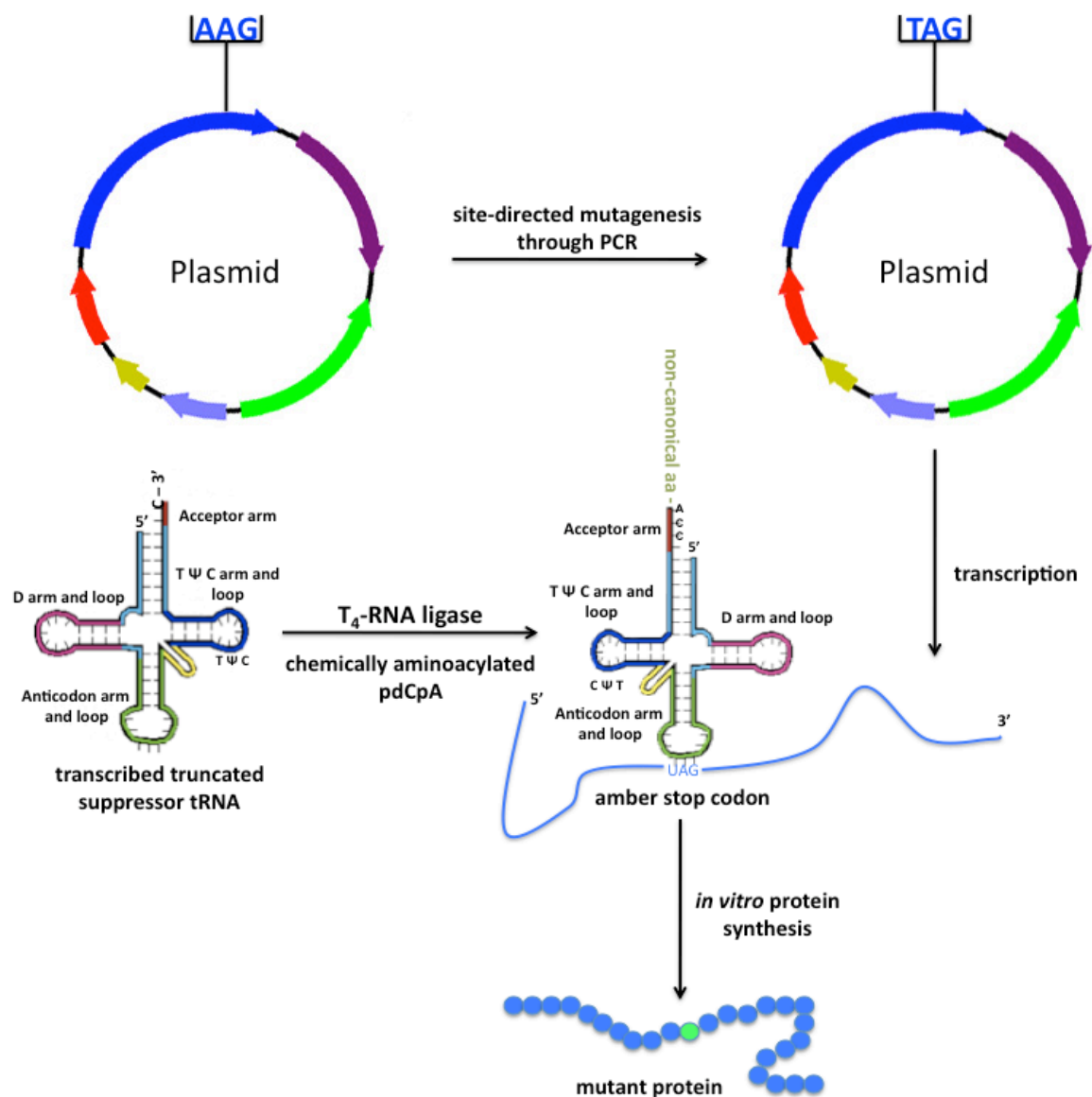


Figure 4.1-4 Schematic representation of *amber* suppression based *in vitro* biosynthetic approaches for protein mutagenesis

An *amber* stop codon is introduced into gene at desired position. A constructed tRNA in which the anti-codon loop presents an anti-*amber* codon, is chemically aminoacylated. Protein expression is performed by means of cell-free protein synthesis. (Adapted from Ye *et al.*⁹⁵, Beilstein Journal of Organic chemistry, open access journal)

One requirement of this selective incorporation of non-canonical amino acids into proteins is that a tRNA can uniquely recognize the non-sense codon and efficiently deliver amino acids to the ribosome for polypeptide synthesis. This tRNA must be

4.1 Protein engineering with non-canonical amino acids

orthogonal to endogenous aaRSs and avoid deacylation by the proofreading of any endogenous aaRSs. Therefore, for experiments with extracts from *E. coli*, an orthogonal *amber* suppressor tRNA derived from yeast Phe tRNA (tRNA^{Phe}_{CUA}) was constructed for this purpose.⁹⁶ Furthermore, the development of suitable methods to couple non-canonical amino acids and optimize cell-free protein synthesis are also crucial parameters for success. Techniques for aminoacylation and cell-free protein synthesis are separately discussed in the following sections.

4.1.3.1 Chemical aminoacylation of suppressor tRNA

As described in section 4.1.1, aaRSs are enzymes with very high specificity. The misaminoacylation of suppressor tRNA with non-canonical amino acids by the use of naturally occurring aaRSs is problematic. In general, directed evolution is required to isolate enzymes that recognize the amino acid analogue as a substrate.⁹⁷ Direct chemical aminoacylation is also difficult, because the numerous reactive groups present in the tRNA molecule make regioselectivity impossible. Therefore, semi-synthetic approaches have been developed in which a mono- or dinucleotide representing the 3'-end of a suppressor tRNA is chemically synthesized and aminoacylated, and subsequently ligated to a truncated suppressor tRNA produced by run-off *in vitro* transcription (Figure 4.1-5).

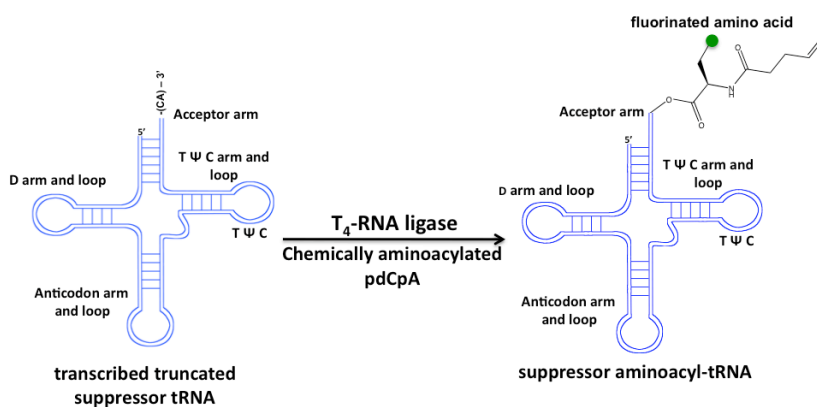


Figure 4.1-5 Schematic representation of T4 RNA ligase mediated chemical aminoacylation

(Adapted from Ye *et al.*⁹⁵, Beilstein Journal of Organic chemistry, open access journal)

4.1 Protein engineering with non-canonical amino acids

Hecht and coworkers demonstrated this method by using a pCpA dinucleotide.⁹⁸ Later, Schultz and coworkers improved the method by the use of a hybrid pdCpA dinucleotide that eliminates the need for a 2'-OH protecting group during synthesis.⁹⁹ During chemical aminoacylation, the α -amino group of the amino acid is protected to increase the half-life of the aminoacylated dinucleotide, and the carboxyl group of the amino acid is activated to a cyanomethyl ester, which facilitates the selective aminoacylation of the dinucleotide.^{99a} The nitroveratryloxy carbonyl (NVOC) and *N*-4-pentenoyl groups are the most widely used α -amino protecting groups for chemical aminoacylation.^{99a, 100} The removal of the NVOC group is performed under irradiation with 350 nm light. The *N*-4-pentenoyl group can be removed under mild aqueous reaction conditions with iodine (Figure 4.1-6). The use of the *N*-4-pentenoyl group even enables the preparation of caged proteins.¹⁰¹ It should be noted that higher reaction temperatures and longer incubation times result in both mono-aminoacylated products and bis-aminoacylated products. Hecht and coworkers systematically studied the suitability of bis-aminoacylated suppressor tRNAs on protein translation. Their results indicate that bis-aminoacylated products have greater stability than mono-aminoacylated products, and that the ligated tRNAs are able to deliver both types of activated amino acids for protein synthesis.¹⁰²

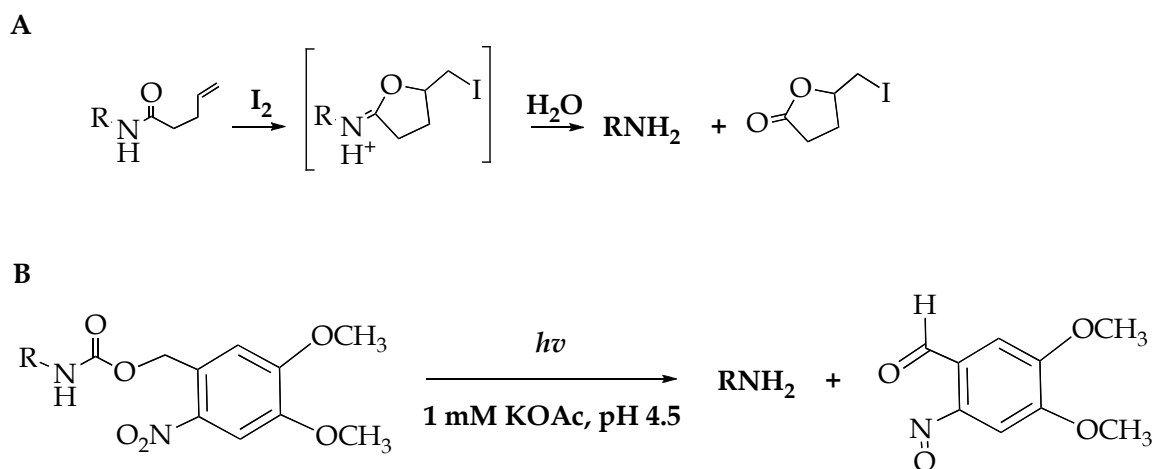


Figure 4.1-6 Schematic representations of nitroveratryloxy carbonyl and *N*-4-pentenoyl protecting groups and their removal mechanisms

(A) nitroveratryloxy carbonyl group (NVOC), (B) *N*-4-pentenoyl protecting group

4.1 Protein engineering with non-canonical amino acids

Several optimization steps and new methods have been developed to reduce the amount of time consumed by the chemical aminoacylation of suppressor tRNAs.¹⁰³ Sisido and coworkers have demonstrated *in situ* chemical aminoacylation through amino acid thioester linked peptide nucleic acids (PNA) and the preparation of misacylated tRNAs in cationic micelles and under ultrasonic agitation.^{103b, 104} Recently, Suga and coworkers have developed and optimized an alternative system for highly flexible tRNA aminoacylation, called flexizymes.¹⁰⁵ The system is based on the evolution and selection of programmable ribozymes, which can recognize the 3'-end of a tRNA and a benzylic moiety on the leaving group. Therefore, flexizymes can catalyze the aminoacylation of amino acid benzyl esters as aminoacyl acceptors and donors with a wide variety of tNRAs. The flexizymes can be further immobilized on agarose resin, which facilitates the isolation and purification of aa-tRNAs.¹⁰⁶ More recently, a method for the direct aminoacylation of full-length tRNAs through a La³⁺ mediated acylation was established. Although the acylation yield of this method is lower than that of the T4 RNA ligase-mediated method, the aminoacylated tRNAs do efficiently participate in protein translation.¹⁰⁷

4.1.3.2 *in vitro* protein synthesis

Cell-free protein expression (*in vitro* protein synthesis) takes advantage of the fact that protein synthesis is carried out in cells by soluble components, and that cell integrity is not necessary for protein synthesis to occur. The most important enzymes and factors are ribosome, aaRSs, initiation factors (IF), elongation factors (EF), tRNAs, amino acids, and ribonucleotide triphosphates as energy sources.¹⁰⁸

Cell-free protein expression offers numerous unique advantages compared to traditional cell-based protein expression methods: 1) cell-free protein expression meets the demands of single-protein expression by only adding the gene of the target protein, which facilitates downstream protein purification steps; 2) the open environment of the reaction enables direct control over the reaction conditions, and

4.1 Protein engineering with non-canonical amino acids

the possibility of providing target-oriented cofactors; 3) once the cell-free lysate is prepared, protein synthesis can be conducted rapidly; 4) non-canonical amino acids can be site-specifically introduced;^{94a, 94c} 5) cell-free protein expression systems are not sensitive to toxic proteins, enabling the synthesis of such targets; 6) cell-free protein expression systems also enable the production of membrane proteins by tolerating the presence of detergents;¹⁰⁹ 7) cell-free protein expression can be directly adopted for *in vitro* display technologies such as ribosome display and mRNA display, which allow *in vitro* selection from a combinatorial library;¹¹⁰ 8) cell-free protein expression systems can be applied to protein array technologies.¹¹¹ Recent achievements in the development of cell-free systems with optimized energy consumption,¹¹² optimized cell lysate preparation,¹¹³ amino acid stabilization,¹¹⁴ utilization of a continuous exchange system,¹¹⁵ and those that allow the addition of cofactors¹¹⁶ or mimic the cytoplasmic environment,¹¹⁷ have enabled longer reaction times and the production of folded functional proteins with high yields.

The most common cell-free translation systems are extracts from *Escherichia coli* (*E. coli*), wheat germ, rabbit reticulocytes, and insect cells.^{108c} In general, the prokaryotic system uses a DNA template for coupled transcription-translation and is typically high yielding, while the eukaryotic systems use an mRNA template resulting in lower yields but providing the possibility for post-translational modifications of the expressed proteins. Though any organism is a potential source of cell-free protein lysates, it is challenging to establish a new system from a new organism. Recently, Alexandrov and coworkers demonstrated a universal cell-free protein expression system for protein synthesis from multiple organisms by designing an unstructured translation-initiation sequence that presumably operates independently of initiation factors.^{115, 118}

4.1 Protein engineering with non-canonical amino acids

4.1.3.3 Limitations and improvements of *in vitro* non-sense suppression

By means of *in vitro* non-sense suppression approaches more than 100 non-canonical amino acids have been incorporated into proteins with the aim of generating novel chemical and biophysical properties.^{94b, 119} However, this method is limited by poor yields, suppression efficiency is commonly less than 20%; targets have included dehydrofolate reductase (DHFR), streptavidin, luciferase, lysozyme, green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT), and β -lactamase.^{72a, 72c} The low suppression efficiency is due to both intrinsic factors, which are relevant to the high fidelity of translation as discussed in section 3.1, and competitive release factors (RFs). RFs are small proteins that recognize stop codons, bind to the A-site of the ribosome, and trigger the termination of translation by inducing hydrolysis of the ester bond linking the peptide chain to the tRNA at the P-site.¹²⁰ Several improvements have been made to this approach, to minimize the competition effects of RFs: heat treatment of cell-free extracts to deactivate RF1;¹²¹ the use of RNA aptamers that bind to RF1;¹²² and frameshift suppression of four-base codons.¹²³ These strategies have resulted in slight improvements in suppression efficiency, and high yields of target proteins are not always a requirement. For example, Dougherty and co-workers take advantage of *in vitro* suppression (microinjection of suppressor aa-tRNA into *Xenopus oocytes*) in mapping ion channel functions. In their studies, electrical signals can be obtained from 10 amol ion channel proteins.^{7, 94b, 124}

4.1.4 Expanding the genetic code

A potential general approach for the site-specific incorporation of non-canonical amino acids is based on the fact that the “21st” amino acid selenocysteine is naturally incorporated into proteins. Selenocysteine is encoded by an in-frame TGA non-sense codon and charged onto a Sec-tRNA, which acts in concert with a special elongation factor and a *cis*-acting element in mRNA.^{71a} Several requirements must be fulfilled to

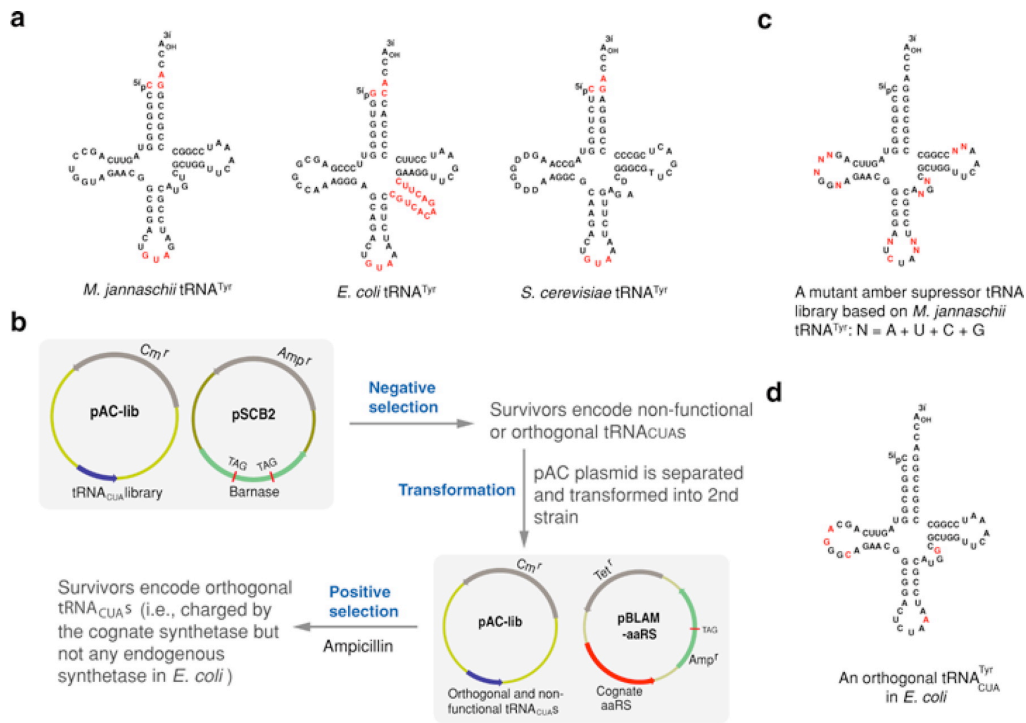
4.1 Protein engineering with non-canonical amino acids

generate an *in vivo* suppression system: 1) the non-canonical amino acid must be efficiently transported into cells, and may not be the substrate of any endogenous aaRSs; 2) a novel tRNA must be constructed, one that can efficiently deliver non-canonical amino acids in response to a unique codon but is not recognized by any host aaRSs (it must also avoid proofreading mechanisms); 3) an orthogonal aaRS must be designed and evolved, which aminoacylates only the novel tRNA, and not any other endogenous tRNA, with the non-canonical amino acid.

4.1.4.1 General approaches to the selection of orthogonal tRNA/aaRS pairs

Two strategies can be used to generate tRNA/aaRS pairs to expand the genetic code: 1) generating an orthogonal tRNA/aaRS pair from existing tRNA/aaRS pairs through directed evolution; 2) importing an orthogonal tRNAs/aaRS pair from a different organism.¹²⁵ tRNA^{Tyr}_{CUA}/TyrRS from *Methanococcus Jannaschii* (*M. Jannaschii*) is the first efficient tRNA/aaRS pair for the incorporation of non-canonical amino acids into proteins in an *E. coli* host.¹²⁶ This orthogonal pair was generated by a combined positive and negative selection. In the negative selection, a tRNA library, which was based on the consensus sequence of major identity elements in tRNAs, was introduced into an *E. coli* host. The *E. coli* host contained a reporter gene encoding the protein barnase, which is toxic. Into the reporter gene of barnase an amber stop codon was introduced. Thus, in case a member of the suppressor tRNA library was charged by endogenous aaRSs, full-length barnase is expressed, which leads to cell death; the hosts containing orthogonal or non-functional tRNAs survive. The positive selection was conducted in the presence of heterologous aaRSs and a β -lactamase gene containing the amber stop codon in *E. coli*, and this antibiotic was also present in the growth medium. Once an aaRS library member efficiently aminoacylates the orthogonal tRNAs, the cells survive by producing β -lactamase. In this way an effective orthogonal tRNA can be selected.^{72d, 126d} The selection procedures and identity elements from three organisms are illustrated in *Figure 4.1-7*.

4.1 Protein engineering with non-canonical amino acids



AR Wang L., et al. 2006.
Annu. Rev. Biophys. Biomol. Struct. 35:225–49

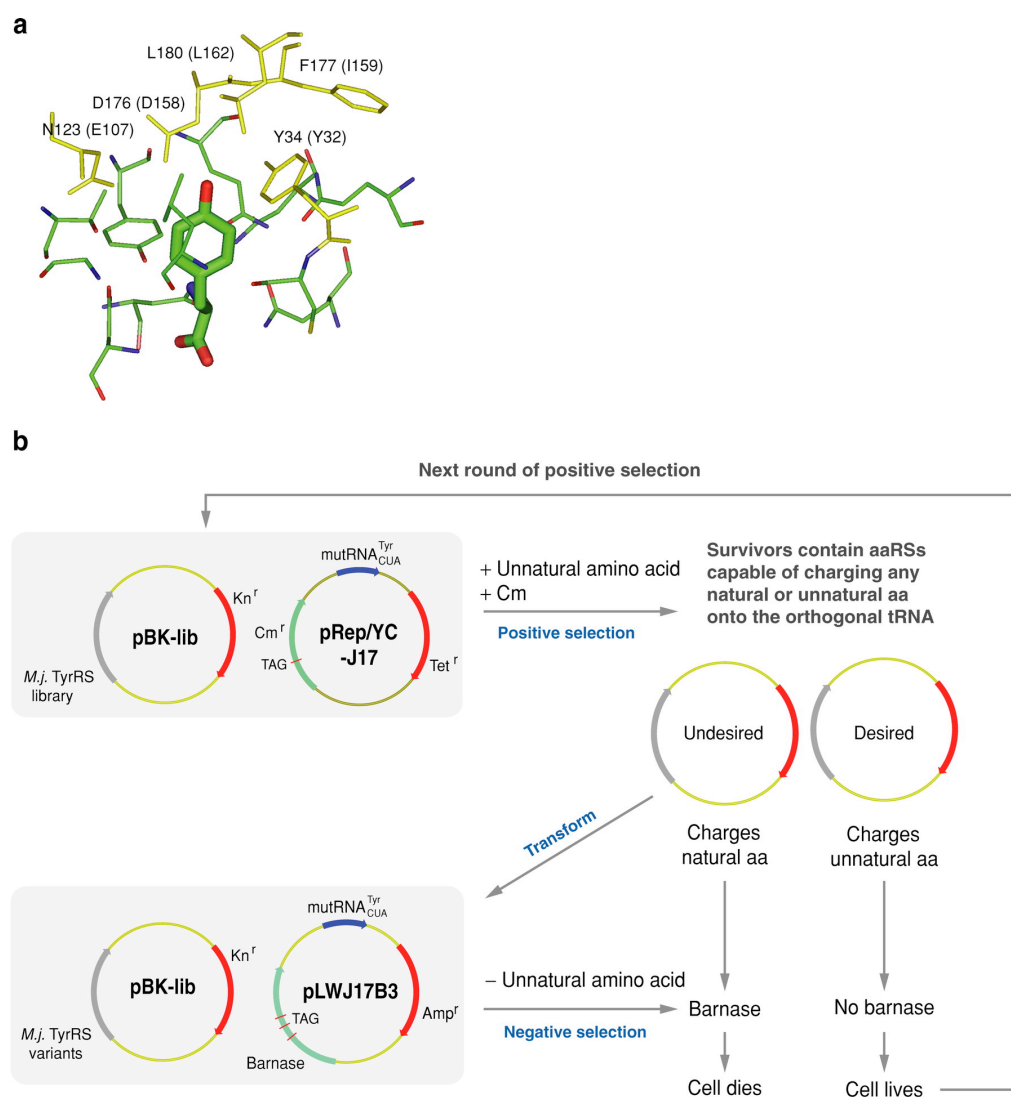
Figure 4.1-7 Schematic representation of diversity of tRNA^{Tyr} from different organisms and general approach of orthogonal tRNA selection

a) The tRNA^{Tyr} sequence from three different species *M. Jannaschii*, *E. coli*, and *S. cerevisiae*, the major identity elements are labeled in red; b) schematic representation of negative and positive selection, c) generation of tRNA^{Tyr} library, the randomizing nucleotide labeled in red; d) the orthogonal tRNA (Adapted with permission from Wang *et al.*,^{72d} Copyright © 2006, Annual Reviews)

In order to enrich the catalytic specificity of aaRSs for certain non-canonical amino acids, a combined positive and negative selection is also conducted in this case. Thus, a library of aaRSs based on the structural analysis of aaRS-amino acid complexes was generated. In the positive selection, a reporter gene of CAT with an internal amber stop codon was used with chloramphenicol present in the cell culture media. Efficient aminoacylation by an orthogonal aaRS and the subsequent incorporation of any amino acid results in the production of full-length CAT, which neutralizes the chloramphenicol and enables the bacterial host to survive. These cells are subjected to the negative selection, again based on the translation of barnase; thus, all *E. coli* strains containing aaRSs that charge tRNAs with canonical amino acids and result in

4.1 Protein engineering with non-canonical amino acids

the production of toxic full-length barnase do not survive. The selection can be repeated in multiple rounds to obtain aaRSs with high specificity. Figure 4.1-8 presents the binding model of Tyr-aaRS^{Tyr} with schemes for the selections.^{72d, 126b, 127}



Wang L, et al. 2006.
Annu. Rev. Biophys. Biomol. Struct. 35:225–49

Figure 4.1-8 Schematic representation of an efficient system for the evolution of orthogonal aaRSs

a) binding model of Tyr and aaRS^{Tyr}, b) selection scheme of combined negative and positive selection (Adapted with permission from Wang *et al.*,^{72d} Copyright © 2006, Annual Reviews)

4.1 Protein engineering with non-canonical amino acids

The general selection approach can be adopted for the incorporation of non-canonical amino acids into proteins by use of other organisms, e.g., yeast and mammalian cells. However, critical points regarding the different cellular mechanisms must be considered. For instance, in eukaryotic expression systems, the transcription of tRNA takes place in the nucleus and needs an internal A- and B-box sequence. After transcription, the tRNA must be exported to the cytoplasm via an exportin-tRNA-dependent process.^{72d, 128} In addition, different reporter genes can be used for the selection of aaRSs. GAL4, a transcriptional activator protein that drives the transcription of genomic GAL4-responsive genes including *his 3*, *ura 2*, and *lacZ*, has been used in positive selections for aaRSs in *S. cerevisiae*. 5-fluoroorotic acid has been used for negative selections. Also GFP and *lacZ* can be used as chromogenic reporters for selections.^{72d, 129} The orthogonal tRNA/aaRS pairs established thus for the successful incorporation of non-canonical amino acids are summarized in *Table 4.1-1*.

Table 4.1-1 A short summary of orthogonal tRNA/aaRS pairs, which have been used for the incorporation of non-canonical amino acids into proteins in different expression hosts

tRNA	Organism	aaRS	Organism	Expression host	Literature
tRNA ^{Tyr}	<i>M. jannaschii</i>	TyrRS	<i>M. jannaschii</i>	<i>E. coli</i>	126c, 130
tRNA ^{Asp}	<i>S. cerevisiae</i>	AspRS	<i>S. cerevisiae</i>	<i>E. coli</i>	126a
tRNA ^{Gln}	<i>S. cerevisiae</i>	GlnRS	<i>S. cerevisiae</i>	<i>E. coli</i>	127
tRNA ^{Tyr}	<i>S. cerevisiae</i>	TyrRS	<i>S. cerevisiae</i>	<i>E. coli</i>	131
tRNA ^{Phe}	<i>S. cerevisiae</i>	PheRS	<i>S. cerevisiae</i>	<i>E. coli</i>	132
tRNA ^{fMet}	<i>E. coli</i>	TyrRS	<i>S. cerevisiae</i>	<i>E. coli</i>	133
tRNA ^{Glu}	<i>P. horikoshii</i>	GluRS	<i>P. horikoshii</i>	<i>E. coli</i>	126b
tRNA ^{Leu}	<i>Halobacterium sp. NCR1</i>	LeuRS	<i>M. thermoautotrophicum</i>	<i>E. coli</i>	134
tRNA ^{Lys}	<i>P. horikoshii</i>	LysRS	<i>P. horikoshii</i>	<i>E. coli</i>	135
tRNA ^{PyI}	<i>M. barkeri</i>	PyIRS	<i>M. barkeri</i>	<i>E. coli</i>	71b, 136
tRNA ^{PyI}	<i>M. mazei</i>	PyIRS	<i>M. mazei</i>	<i>E. coli</i>	137
tRNA ^{PyI}	<i>D. hafniense</i>	PyIRS	<i>D. hafniense</i>	<i>E. coli</i>	138
Init. tRNA	<i>Homo sapiens</i>	GlnRS	<i>E. coli</i>	<i>S. cerevisiae</i>	133
tRNA ^{Leu}	<i>E. coli</i>	LeuRS	<i>E. coli</i>	<i>S. cerevisiae</i>	139
tRNA ^{Tyr}	<i>E. coli</i>	TyrRS	<i>E. coli</i>	<i>S. cerevisiae</i>	129a, 140
tRNA ^{Tyr}	<i>E. coli</i>	TyrRS	<i>E. coli</i>	<i>Pichia pastoris</i>	141
tRNA ^{PyI}	<i>M. mazei</i>	PyIRS	<i>M. mazei</i>	Mammalian cell ^a	142
tRNA ^{Tyr}	<i>B. stearrowthermophilus</i>	TyrRS	<i>E. coli</i>	Mammalian cell ^a	143

4.1 Protein engineering with non-canonical amino acids

tRNA ^{Pyl}	<i>M. mazei</i>	PylRS	<i>M. mazei</i>	<i>C. elegans</i>	144
tRNA ^{Tyr}	<i>E. coli</i>	TyrRS	<i>E. coli</i>	<i>C. elegans</i>	145
tRNA ^{Leu}	<i>E. coli</i>	LeuRS	<i>E. coli</i>	<i>C. elegans</i>	145

Init. = Initiator, *M. jannaschii* = *Methanocaldococcus jannaschii*, *S. cerevisiae* = *Saccharomyces cerevisiae*, *E. coli* = *Escherichia coli*, *M. thermoautotrophicum* = *Methanobacterium thermoautotrophicum*, *P. horikoshii* = *Pyrococcus horikoshii*, *M. barkeri* = *Methanosarcina barkeri*, *M. mazei* = *Methanosarcina mazei*, *D. hafniense* = *Desulfitobacterium hafniense*, *B. stearothermophilus* = *Bacillus stearothermophilus*, *C. elegans* = *Caenorhabditis elegans* a) CHO cells.

4.1.4.2 Limitations on expanding the genetic code

> 70 non-canonical amino acids have been incorporated into proteins by the use of orthogonal tRNA/aaRS pairs in *E. coli*, yeast, and mammalian cells.¹⁴⁶ This remarkable achievement facilitates the incorporation of biophysical and biochemical probes into proteins for functional and structural studies *in vitro* and in living cells. However, the most frequently incorporated non-canonical amino acids are aromatic amino acids, with the exception of certain Lys derivatives.^{72c, d, 146-147} Nevertheless, aliphatic amino acids are also important for the study of protein folding and conformational changes. Because the size and hydrophobicity of aliphatic side chains are more similar to one another than are those of the aromatic side chains, it has been assumed that the aaRSs that recognize the aliphatic amino acids are likely to have greater cross-reactivity, that is, comparably lower selectivity. Therefore, the site-specific incorporation of natural aliphatic analogues by means of orthogonal tRNA/aaRS pairs may prove to be more challenging than in the case of the aromatics. Furthermore, because this strategy is principally based on amber suppression, RF-mediated protein chain termination will always limit the efficiency of synthesis of full-length protein to 20-30%.¹⁴⁸ Recently, Wang and coworkers generated RF1 deletion strains of *E. coli*, which can produce dramatically increased yields of protein containing non-canonical amino acids; this is in line with another recent study that demonstrates that RF1 is non-essential for *E. coli*.¹⁴⁹ However, the precise functions of RF2 and RF3 in suppression-based methods remain unclear. More recently, Park *et al.* demonstrated that engineering of EF-Tu facilitated the incorporation of O-phosphoserine by use of the *amber* stop codon based suppression method.¹⁵⁰

4.1 Protein engineering with non-canonical amino acids

4.1.5 Orthogonal systems and reprogramming the genetic code

Recently, Chin and coworkers developed an orthogonal gene expression network (Figure 4.1-9).¹⁵¹ In the orthogonal expression system, besides the orthogonal tRNA/aaRS pairs used for the aminoacylation of non-canonical amino acids, orthogonal ribosome/mRNA pairs are also generated in such a way that the orthogonal mRNA is not a substrate for the endogenous ribosome, only for the orthogonal ribosome. The efficiency of the site-specific incorporation of one non-canonical amino acid from a single amber codon can be increased from 20% to 60%, and double incorporation from two suppressor codons up to 20%.^{148, 152}

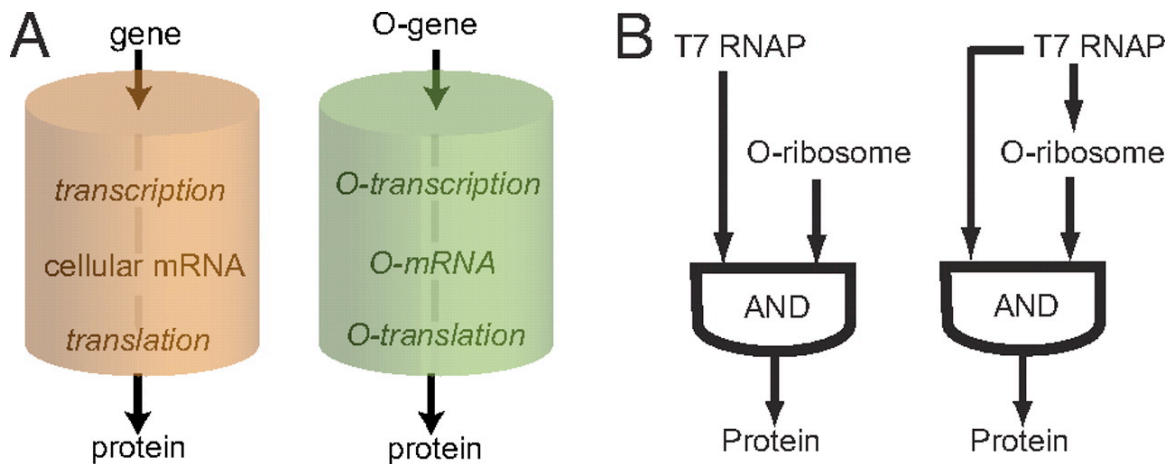


Figure 4.1-9 Schematic representation of an orthogonal expression network

(Adapted with permission from An et al.,¹⁵¹ Copyright © 2009, Proceeding of National Academy of Science of the United States of America)

4.2 Site-specific incorporation of fluorinated amino acids into BPTI

4.2.1 Chemical misacylation of fluorinated amino acids

4.2.1.1 Synthesis of hybrid dinucleotide pdCpA

Hybrid Dinucleotide pdCpA was synthesized following the established protocol^{99b} to give an overall yield of 35%; purification was accomplished by HPLC with a C18 column. The analytic RP-HPLC spectrum of *tetra-n*-butylammonium-pdCpA (TBA-pdCpA) showed a very broad peak (*Appendix; 4.3 Methods and Experimental Procedures*).

4.2.1.2 Synthesis of *N*-4-pentenoyl amino acid cyanomethyl esters

The misacylated suppressor-tRNAs are the key intermediate compounds in *in vitro* amber suppression. In order to probe the incorporation of fluorinated amino acids into BPTI protein, these non-canonical amino acids (*Figure 4.2-1*) were first chemically attached to yeast suppressor tRNA^{Phe}_{CUA}. Following the established protocols, the α -amino group of amino acids was first protected with a *N*-4-pentenoyl group; subsequently, the carboxyl group was activated by forming the cyanomethyl ester.

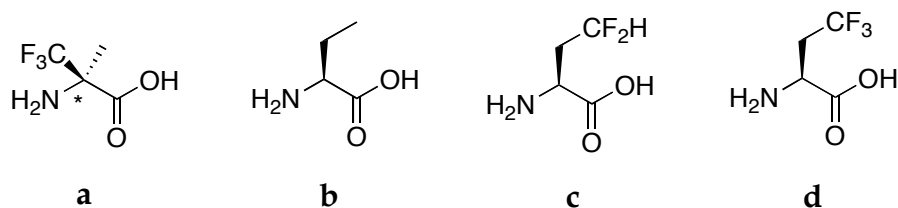


Figure 4.2-1 Structures of non-canonical amino acids used for chemical aminoacylation

a) (*R,S*)-2-amino-2-methyl-3,3,3-trifluoropropanoic acid, b) (*S*)-ethylglycine, c) (*S*)-2-amino-4,4-difluorobutanoic acid, d) (*S*)-2-amino-4,4,4-trifluorobutanoic acid.

4.2 Site-specific incorporation of fluorinated amino acids into BPTI

The general approach for the protection of the α -amino group and the activation of the carboxyl group is presented schematically in *Figure 4.2-2*. Because of the strong electron-withdrawing character of the fluorine atom, the α -trifluoromethyl group in compound **a** strongly influences the reactivity of the neighboring amino and carboxyl groups. A steric effect can also be observed in the protection step. Thus, protection of the amino group in compound **a** is achieved by treatment with the more reactive reagent *N*-4-pentenoyl chloride, while for compounds **b**, **c**, and **d** the less reactive reagent *N*-4-pentenoyloxy succinimide suffices.¹⁵³ It is worth noting that for the protection of compound **a** dimethylformamide (DMF) and pyridine were tested as reaction solvents. When DMF was used as the solvent, 4-dimethylaminopyridine (DMAP) was added as the base. Though pyridine performs double duty as a base in the reaction (pK_a 5.21), the combination of DMF and DMAP gave higher yields. The yields of protection and activation for each amino acid are summarized in *Table 4.2-1*.

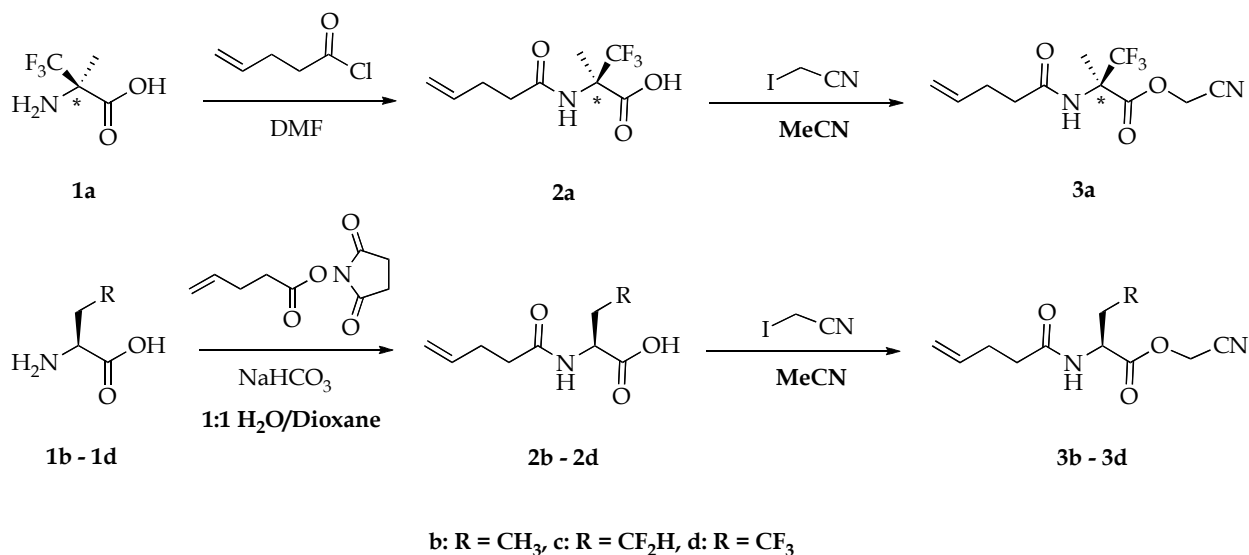


Figure 4.2-2 General approaches for the synthesis of *N*-4-pentenoyl amino acid cyanomethyl ester

4.2 Site-specific incorporation of fluorinated amino acids into BPTI

4.2.1.3 Synthesis of mono-2' (3')-O-[N-(4-pentenoyl)aminoacyl]-pdCpAs and 2'-3'-bis-O-[N-(4-pentenoyl)aminoacyl]-pdCpAs

The misacylation of *N*-4-pentenoyl amino acid to pdCpA was performed following the established protocol. The pdCpA must be in the form of its *tetra-n*-butylammonium (TBA) salt to increase its solubility in DMF, and this is accomplished by means of ion-exchange chromatography. It is important to keep the ratio of TBA to pdCpA larger than 2.2.^{99a} The reaction was carried out in anhydrous DMF with triethylamine (TEA) as the base (Figure 4.2-3).

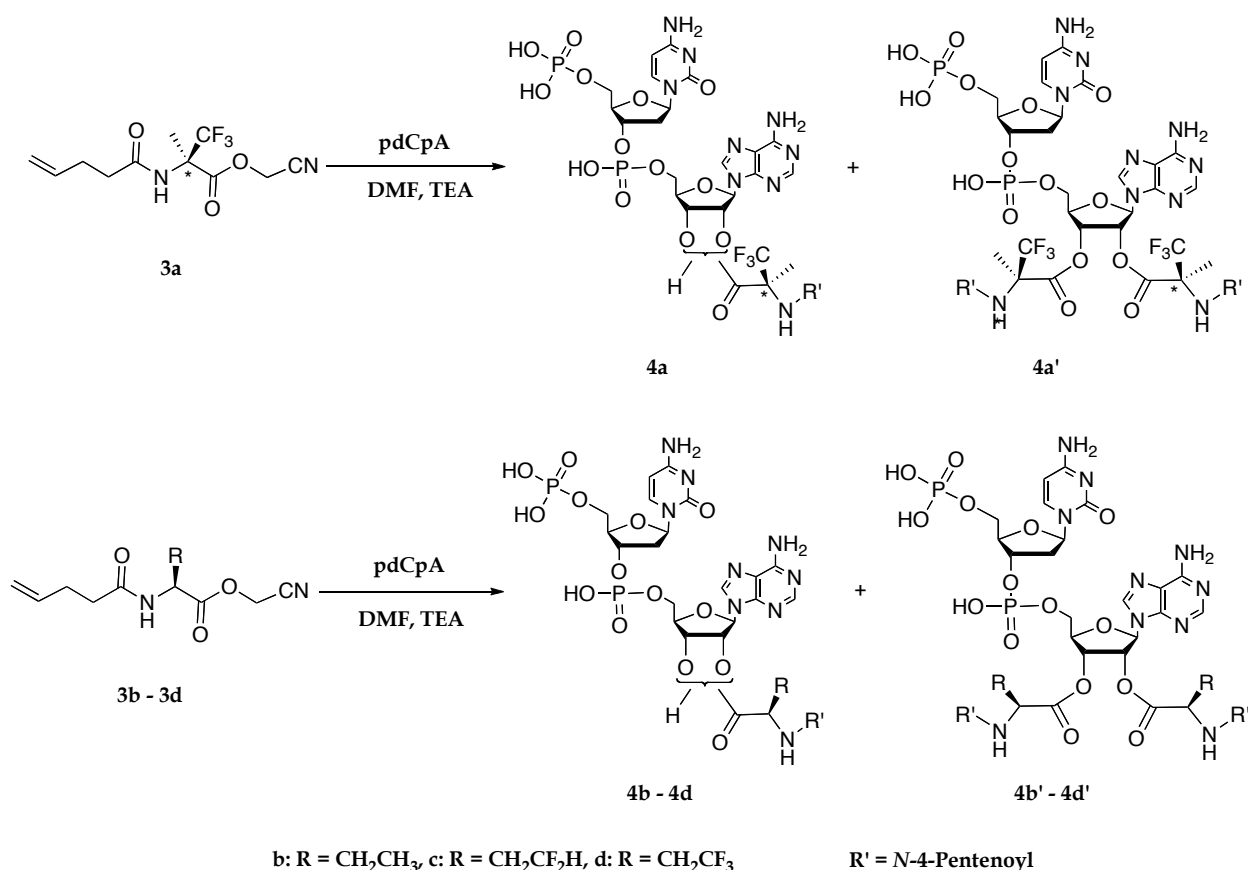


Figure 4.2-3 Scheme of synthesis of mono-2' (3')-O-[N-(4-pentenoyl)aminoacyl]-pdCpAs and 2'-3'-bis-O-[N-(4-pentenoyl)aminoacyl]-pdCpAs

Because the fluorinated amino acids are precious, we extended the incubation time and elevated the reaction temperature to obtain the maximal yield. Bis-acylated

4.2 Site-specific incorporation of fluorinated amino acids into BPTI

products of compounds **a**, **b**, and **d** were also synthesized and purified; based on the results from Hecht and coworkers demonstrating that bis-acylated tRNAs can efficiently participate in protein synthesis, the bis-acylated products were also used downstream in protein expression experiments.^{102a, 102d} The yields of misacylation for each non-canonical amino acid are summarized in *Table 4.2-1*.

Table 4.2-1 Summary of chemical synthesis results for *N*-4-pentenoyl protected non-canonical aminoacyl-pdCpA

Compound	Yield of protection/activation (%)	Ratio of <i>N</i> -4-pentenoyl-aa cyanomethyl ester/pdCpA	Yield of mono-/bis-acylated aa-pdCpA (%)
a	22	1.8 : 1	45/51
b	81	2.5 : 1	34/10
c	54	2.9 : 1	65/-
d	75	3.1 : 1	13/49

4.2.1.4 T4 RNA ligase-mediated ligation

The truncated suppressor tRNA^{Phe}_{CUA} (-CA) was synthesized by run-off *in vitro* transcription. First, the plasmid PYRNA8 containing a T7 promotor was linearized with endonuclease FokI for the production of truncated tRNA, and linearized with BstNI for the production of full-length tRNA as control, respectively (for sequence of tRNA^{Phe}_{CUA}, see *Appendix*).

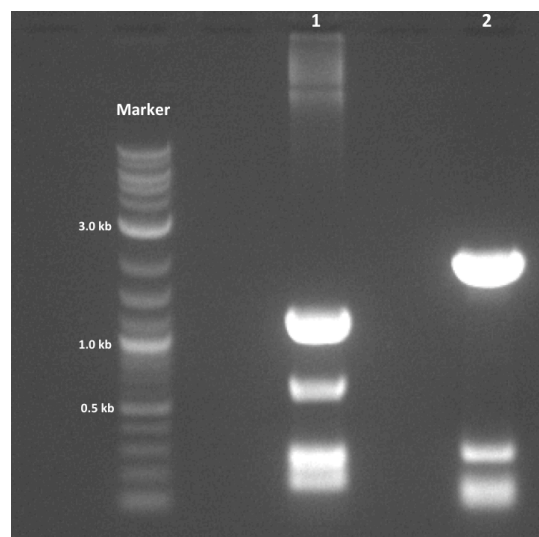


Figure 4.2-4 Agarose gel electrophoresis analysis of linearized plasmid PYRNA8

lane 1, treated with FokI; lane 2, treated with BstNI

4.2 Site-specific incorporation of fluorinated amino acids into BPTI

In order to obtain the truncated suppressor tRNA^{Phe}_{CUA} (-CA) with high yield and quality, the run-off *in vitro* transcriptions were optimized by varying the concentration of GTP and adding/excluding formamide and GMP. Choudhury *et al.* reported that the addition of 5% formamide can increase the quality of transcribed tRNA, and that the addition of 20 mM GMP ensures the correct size of the tRNA product.¹⁵⁴ However, our results indicate that the addition of formamide decreases the yield of transcription (Figure 4.2-5, left panel, lane 4 and 5). Addition of GMP did not influence the yield of run-off *in vitro* transcription (Figure 4.2-5, left panel, lane 1 and 2). GMP was used to prime the *in vitro* transcription to replace the 5'-terminal triphosphate with 5'-monophosphate. Sampson *et al.* demonstrated that yeast tRNA^{Phe} transcripts bearing either 5' monophosphate or triphosphate were comparable to natural yeast tRNA^{Phe} for aminoacylation.¹⁵⁴ However, it is not clear in this case whether the nature of the 5'-terminus of the suppressor tRNA has any effect on chemical misacylation. In the studies of other groups including Schultz, Sisido, and Dougherty, the effects regarding a 5'-mono- or triphosphate were not described.

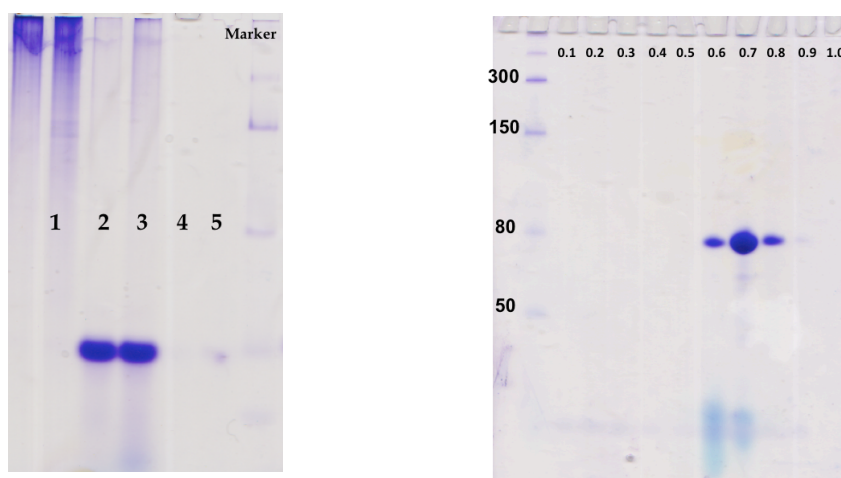


Figure 4.2-5 Urea denaturing acidic PAGE analysis of run-off *in vitro* transcription and purification of truncated suppressor tRNAs by diethylaminoethanol (DEAE) sepharose anion-exchange column

left, run-off *in vitro* transcription, lane 1, positive control by use of plasmid provided from Kit, **lane 2,** normal run-off *in vitro* transcription without additional formamide and GMP, **lane 3,** + 20 mM GMP, **lane 4,** + 5% formamide, **lane 5,** + 20 mM GMP and 5% formamide; **right, purification of tRNA,** the elution of tRNA preformed with NaCl gradient of 0.1-1 M in 0.1 M NaOAc solution. Typically, tRNA eluted in fractions of NaCl concentration with 0.6-0.8 M.

4.2 Site-specific incorporation of fluorinated amino acids into BPTI

The transcribed truncated $\text{tRNA}^{\text{Phe}}_{\text{CUA}}$ (-CA) was purified by ion exchange chromatography (Figure 4.2-5, right panel). Typically, a yield of 3 μg truncated $\text{tRNA}^{\text{Phe}}_{\text{CUA}}$ (-CA) per 1 μL transcription reaction can be obtained after the purification. It must be mentioned that, in general, run-off *in vitro* transcription typically delivers an RNA product of the correct size, but it is not unusual to observe products having one or more additional nucleotides at the 3' terminus. This is referred to as the $n+1$ activity of RNA polymerase.¹⁵⁵ Gerrits systematically investigated the effects on end heterogeneity of T7 transcripts in amber suppression. The $n+1$ activity is dependent on the temperature and RNA-polymerases. The T7 RNA-polymerase and a higher incubation temperature at ~ 40 $^{\circ}\text{C}$ can reduce the $n+1$ activity.¹⁵⁶ In our studies, the run-off *in vitro* transcription reactions were incubated at 42 $^{\circ}\text{C}$. And we did not observe heterogenetic tRNA products in gel by use of AmpliScribe™ T7 High Yield Transcription Kits (EPICENTRE® Biotechnologies).

The chemically misacylated *N*-4-pentenoyl-aa-pdCpAs were ligated to truncated $\text{tRNA}^{\text{Phe}}_{\text{CUA}}$ (-CA) by means of T4 ligase. The ligation efficiency was analyzed by means of denaturing acidic PAGE.¹⁵⁷

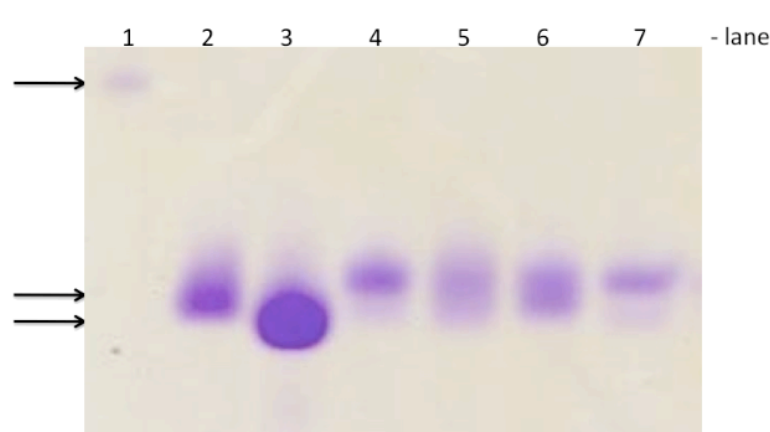


Figure 4.2-6 Urea denaturing PAGE for analysis of T4 RNA ligase mediated ligation

lane 1, Marker (100 base RNA), **lane 2**, full-length yeast suppressor $\text{tRNA}^{\text{Phe}}_{\text{CUA}}$, **lane 3**, truncated yeast suppressor $\text{tRNA}^{\text{Phe}}_{\text{CUA}}$ (-CA), **lane 4**, *N*-4-pentenoyl-Abu-tRNA, **lane 5**, *N*-4-pentenoyl-DfeGly-tRNA, **lane 6**, *N*-4-pentenoyl-TfeGly-

tRNA, **lane 7**, *N*-4-pentenoyl-TfmAla-tRNA. (Adapted from Ye *et al.*,⁹⁵ Beilstein Journal of Organic Chemistry, open access journal)

From our practical experience, the *N*-4-pentenoyl-Abu-pdCpA results in higher ligation efficiencies than the fluorinated *N*-4-pentenoyl-aa-pdCpA species.

4.2 Site-specific incorporation of fluorinated amino acids into BPTI

Furthermore, bis-aminoacylated pdCpAs show higher ligation efficiencies and higher stabilities than the mono-aminoacylated counterparts. We also observed that ligations were typically incomplete upon scaling up the ligation reaction (> 30 μ g truncated tRNA used).

4.2.1.5 Characterization of misacylated suppressor tRNA^{Phe}_{CUA}

Because the stability of misacylated suppressor tRNA^{Phe}_{CUA} is a key issue in incorporation efficiency in cell-free translation, these were characterized within the framework of this thesis. Commonly, the ligation efficiency is analyzed by means of denaturing acidic PAGE. Petersson *et al.* reported the characterization of semi-synthesized suppressor tRNA by means of matrix-assisted laser desorption/ionization (MALDI-TOF) mass spectrometry.¹⁵⁸ Following this literature procedure, we tested MALDI-MS measurements on our system with an AB SCIEX 5800 TOF/TOF (*Applied Biosystems*). Unfortunately, we could not detect any mass corresponding to our misacylated suppressor tRNAs. The aminoacyl bond is an energy-rich bond with a standard free energy $\Delta G^\circ = -29$ kJ/mol,¹⁵⁹ and Schultz and coworkers reported that misacylated suppressor tRNAs undergo spontaneous hydrolysis even at -20 °C. Thus, the misacylated tRNAs are generally stored at -80 °C after lyophilization.^{154, 160}

We tested the aminoacyl bond's stability under different conditions including H₂O, HEPES buffer, and a cell-free translation premix (*Promega*, pH > 8, measured with pH indicator paper) for incubation times of 15, 30, or 60 minutes at 37 °C (*Figure 4.2-7*). As presented in *Figure 4.2-7*, in spite of incomplete ligation, all misacylated tRNAs bearing an amino protecting group are stable, at least over the course of the experiment. The deprotected misacylated tRNAs show instability, particularly in the presence of cell-free premix. These results indicated that our misacylated suppressor tRNAs could be used for cell-free protein expression. However, it was estimated that the cell lysate used for cell-free protein synthesis contains other endogenous factors,

4.2 Site-specific incorporation of fluorinated amino acids into BPTI

which can degrade RNA and hydrolyze the aminoacyl bond, and a low yield of protein expression was anticipated.

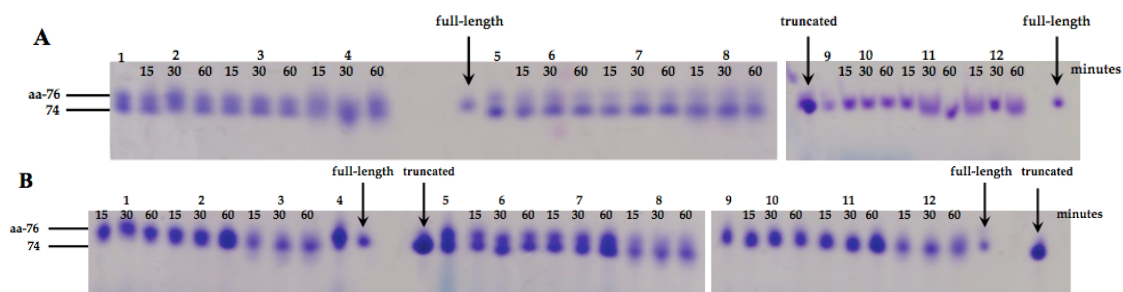


Figure 4.2-7 Urea denaturing acidic PAGE analysis for stability of *N*-4-pentenoyl-aa-tRNA and deprotected aa-tRNA

(A) *N*-4-pentenoyl-aa-tRNAs, (B) deprotected aa-tRNAs, the lanes 15, 30, and 60 represent aliquots incubated under different buffer conditions for 15, 30, or 60 minutes (at 37 °C). (A) All samples are *N*-4-pentenoyl protected. 1. Abu-tRNA, 2. Abu-tRNA in H₂O, 3. Abu-tRNA in HEPES buffer (pH 7.5), 4. Abu-tRNA in cell-free premix, 5. Bis-Abu-tRNA, 6. Bis-Abu-tRNA in H₂O, 7. Bis-Abu-tRNA in HEPES buffer (pH 7.5), 8. Bis-Abu-tRNA in cell-free premix, 9. TfeGly-tRNA 10. TfeGly-tRNA in H₂O, 11. TfeGly-tRNA in HEPES buffer (pH 7.5), 12. TfeGly-tRNA in cell-free premix. (B) All samples are amino group deprotected. 1. Abu-tRNA in H₂O, 2. Abu-tRNA in HEPES buffer (pH 7.5), 3. Abu-tRNA in cell-free premix, 4. Abu-tRNA, 5. Bis-Abu-tRNA, 6. Bis-Abu-tRNA in H₂O, 7. Bis-Abu-tRNA in HEPES buffer (pH 7.5), 8. Bis-Abu-tRNA in cell-free premix, 9. TfeGly-tRNA 10. TfeGly-tRNA in H₂O, 11. TfeGly-tRNA in HEPES buffer (pH 7.5), 12. TfeGly-tRNA in cell-free premix.

4.2.2 Construct design and cloning for expression of BPTI

4.2.2.1 Design of fluorinated BPTI

As described in Chapter 2, BPTI is a model protein that has been used for studies of protein folding and protein-protein interactions. Lys15 is present in the middle of its interacting loop and plays a crucial role in protease-BPTI interactions. In free BPTI, the side chain of Lys15 is fully solvent exposed. Upon surveying the literature, we found that fluorinated amino acids have been primarily incorporated into the hydrophobic cores of proteins to evaluate their effects on protein folding and stability. In contrast, the effects of fluorine on protein stability at solvent exposed positions have been rarely investigated. Therefore, we set out to determine the impact of substituting Lys15 of BPTI with fluorinated aliphatic amino acids. This design provides the opportunity not only to study protein folding and stability

4.2 Site-specific incorporation of fluorinated amino acids into BPTI

regarding fluorination at a solvent-exposed position, but also to assess its impact at a biologically relevant protein-protein interface.

4.2.2.2 Use of small ubiquitin-like modifier as expression fusion tag

Because BPTI consists of only 58 amino acids, it is challenging to express and purify, and therefore we opted to use a removable expression tag. The small ubiquitin-like modifier (SUMO) fusion protein has proven to be a powerful tool for protein expression, as it has been shown to enhance the solubility and expression level of various proteins. The secondary structure of the SUMO fusion is recognized by SUMO protease 1, which is able to robustly and specifically cleave the SUMO domain from the target protein.¹⁶¹ The SUMO fusion-tag had previously been successfully used for the expression of BPTI (Aprotinin).¹⁶² We designed a protein sequence for the expression of BPTI with a N-terminal SUMO-tag in which the codon for Lys15 was mutated to an *amber* stop codon (for sequence, see *Appendix*). It was expected that subsequent to producing full-length BPTI containing non-canonical amino acids at position 15, the SUMO-tag would be cleaved by SUMO protease 1. The mutant BPTI would then be subjected to structural and functional studies.

4.2.2.3 Expression test of chymotrypsinogen B1 and trypsin 1

Our long-term goal is not only to investigate the supramolecular interactions of fluorinated side chains in native protein environments, but also to probe the extent to which intermolecular “fluorous effects” can be exploited as a protein engineering principle. Thus, our future plans include the substitution of both the BPTI inhibitor and the protease in the S₁ site. To this end, the genes encoding chymotrypsinogen B1 and trypsin 1 (*Homo sapiens*, for sequence see *Appendix*) were cloned into multiple cell-free expression vectors for test expression. Gly216 of chymotrypsin, which plays

4.2 Site-specific incorporation of fluorinated amino acids into BPTI

crucial role in the chymotrypsin-BPTI interaction, was also substituted by an *amber* stop codon.

4.2.2.4 Cloning of expression constructs

Since the protein production yield from cell-free expression is strongly dependent on the promoter (vector) efficiency, DNA encoding SUMO-BPTI, BPTI, chymotrypsinogen B1, and trypsin 1 was cloned into three different vectors by means of standard molecular biological methods. The vectors used for our studies are summarized in *Table 4.2-2*. The characterization of cell-free expression vectors pBH4, pET21cHx, and pIVEX by endonuclease treatment is represented in *Figure 4.2-8*.

Table 4.2-2 Summary of constructs used for cell-free expression tests

Vector ^a	Promoter	His-tag	Cloning site	Insert (protein sequence)
pBH4 ^b	T7 promoter	N-terminus	BamHI/XohI	SUMO-BPTI, SUMO-BPTI _{amber15} ,
pET21cHx ^c	T7 promoter	C-terminus	BamHI/XohI	BPTI, BPTI _{amber15} , Trypsin 1
pIVEX ^d	T7 promoter	N-terminus	XohI/BamHI	Chymotrypsinogen B1, Chymotrypsinogen B1 _{amber216}

These vectors have been used for successful cell-free expression protein production. a) The cloning and expression region of expression vectors are schematically represented in the *Appendix*. References: b) ¹⁶³, c) ¹⁶⁴ d) ¹⁶⁵.

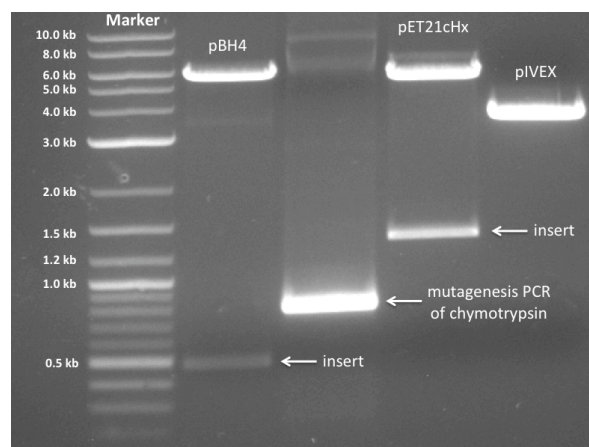


Figure 4.2-8 Characterization of cell-free expression vectors by endonuclease digestion and agarose gel electrophoresis

4.2 Site-specific incorporation of fluorinated amino acids into BPTI

4.2.3 *in vitro* protein expression

4.2.3.1 Protein expression by use of commercially available kits

In order to facilitate our protein expression and reduce the project duration, the protein expression experiments were conducted by the use of commercially available cell lysates or kits. Kits based on *E.coli*, wheat germ, and rabbit reticulocytes are commercially available nowadays. Four different cell-free protein synthesis kits were tested for the site-specific incorporation of fluorinated amino acids (Table 4.2-3).

Table 4.2-3 Commercially available cell-free protein expression kits used here

Commercially available cell-free kit	Cell lysate	Reaction type	Supplier
S30 T7 High-Yield Protein Expression System	<i>E.coli</i>	Batch reaction	Promega
TNT® T7 Quick Coupled Transcription/Translation System	Rabbit reticulocytes	Batch reaction	Promega
RTS 100 <i>E.coli</i> Disulfide Kit	<i>E.coli</i>	Continuous-exchange	5 PRIME
PURExpress Δ RF123 Kit	<i>E.coli</i>	Batch reaction	New England Biolabs

S30 T7 High-Yield Protein expression System enables batch reaction of *in vitro* translation in protein producing with high yield. TNT® T7 Quick System is based on rabbit reticulocytes, which is designed for production of proteins with potential posttranslational modifications. RTS 100 *E.coli* Disulfide Kit is a continuous-exchange cell-free kit that has been optimized particularly for proteins containing disulfide bonds. Because BPTI contains three disulfide bonds, it was assumed that the use of this kit would ensure proper folding, and that extension of the reaction time would result in higher yields. The PURExpress Δ RF123 kit is based on the PURE system, in which all the components for protein synthesis have been purified and reconstituted.¹⁶⁶ In the PURExpress Δ RF123 system, the release factors 1, 2, and 3 are not added, which can eliminate competitive effects between the release factors and suppressor aa-tRNAs and increase the production of full-length protein.

Constructs encoding protein sequence in different expression vectors were test by use of these commercial kits. Protein expression was analyzed by sodium dodecyl polyacrylamide gel electrophoresis (SDS PAGE) and Western blot. The experimental

4.2 Site-specific incorporation of fluorinated amino acids into BPTI

data indicate that native SUMO-BPTI could be expressed from the pIVEX and pBH4 vectors by use of the S30 T7 High Yield Protein Expression System and the RTS 100 *E.coli* Disulfide Kit (Figure 4.2-9 and 4.2-10). Without the SUMO fusion tag, BPTI was only produced from the pBH4 vector with the RTS 100 *E.coli* Disulfide Kit. Among the three tested expression vectors, pBH4 appears to be the most efficient for the *in vitro* translation of BPTI. Owing to the recommended long reaction time (overnight) of the RTS 100 *E.coli* Disulfide Kit, compared to the recommended batch reaction time of 1-2 hours for the S30 T7 High Yield Protein Expression System, a higher yield from the former was expected. Protein expression with and without the SUMO fusion tag showed that coexpression of the SUMO fusion tag can increase expression levels from *in vitro* protein synthesis.

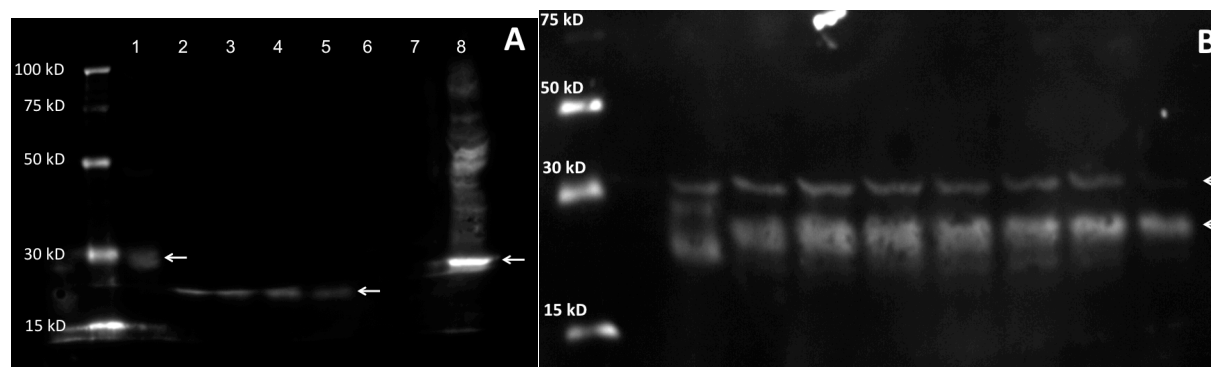


Figure 4.2-9 Western blot analysis of protein expression by use of RTS 100 *E.coli* Disulfide Kit and S30 T7 High-Yield Protein Expression System

(A) **RTS100 *E.coli* Disulfide Kit**, lane 1, positive control sample of urease (plasmid available in kit), lane 2, pIVEX-SUMO-BPTI_{amber15} (addition of suppressor Abu-tRNA), lane 3, pIVEX-SUMO-BPTI_{amber15} (addition of suppressor Bis-Abu-tRNA), lane 4, pIVEX-SUMO-BPTI_{amber15} (addition of suppressor TfeGly-tRNA), lane 5, pIVEX-SUMO-BPTI_{amber15} (addition of suppressor Bis-TfeGly-tRNA), lane 6, pIVEX-chmyotrypsinogen_{amber216}, lane 7, pET21cHx-chmyotrypsinogen_{amber216}, lane 8, pBH4-chmyotrypsinogen_{amber216}. (B) **S30 T7 High-Yield Protein Expression System**, lane 1, pBH4-SUMO-BPTI (full-length protein), lane 2, pBH4-SUMO-BPTI_{amber15} (addition of suppressor TfeGly-tRNA), lane 3, pBH4-SUMO-BPTI_{amber15} (addition of suppressor Abu-tRNA), lane 4, pBH4-SUMO-BPTI_{amber15} (addition of suppressor Bis-TfeGly-tRNA), lane 5 and 6, pBH4-SUMO-BPTI_{amber15} (addition of suppressor Bis-Abu-tRNA), lane 7 and 8, pBH4-SUMO-BPTI_{amber15} (addition of suppressor Bis-TfeGly-tRNA).

4.2 Site-specific incorporation of fluorinated amino acids into BPTI

Afterwards, *amber* suppression for producing of SUMO-BPTI was carried out. Deprotected misacylated suppressor-tRNAs were supplied to *in vitro* translation reactions. Band corresponding to the size of full-length SUMO-BPTI, which may contain non-canonical amino acids, was detected in Western blot (Figure 4.2-9, right panel, **B**). Nevertheless, due to the low yield, the successful incorporation of non-canonical amino acids was not verified with other analytic methods, e.g., mass spectrometry.

Test expressions of chymotrypsinogen B1 and trypsin 1 were carried out as well. The production of chymotrypsinogen B1 and trypsin 1 could be detected only by use of the constructs cloned into pBH4 vector and by use of the RTS 100 *E.coli* Disulfide Kit (Figures 4.2-9 A and 4.2-10 A). However, due to low yields the protein samples were not amenable to purification and isolation.

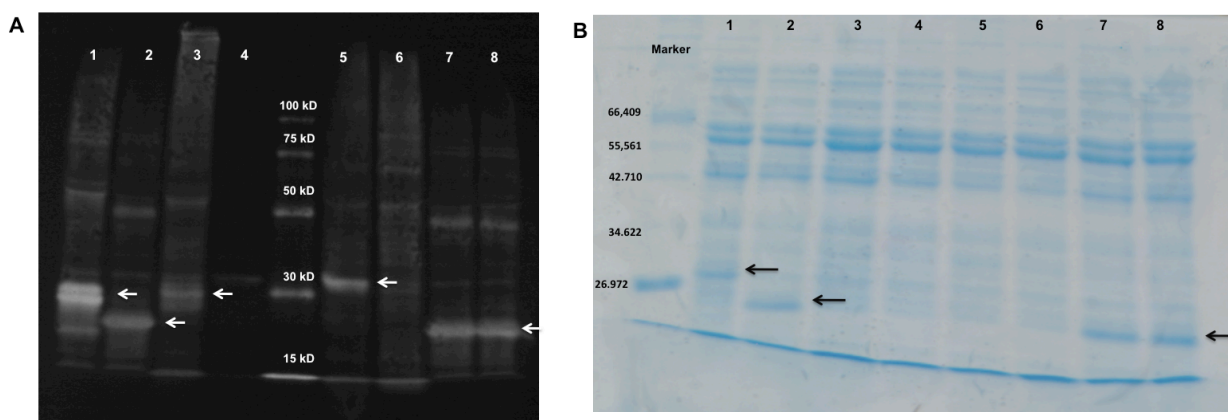


Figure 4.2-10 Western blot and SDS-PAGE analysis of protein expression by use of RTS 100 *E.coli* Disulfide Kit

(A) **Western Blot analysis**, (B) **SDS-PAGE analysis**, the number in (A) and (B) represent the same protein sample, **lane 1**, pBH4-SUMO-BPTI, **lane 2**, pBH4-SUMO-BPTI_{amber15} (addition of suppressor aa-tRNA), **lane 3**, pBH4-trypsin, **lane 4**, pET21cHx-trypsin, **lane 5**, pBH4-chymotrypsinogen B1, **lane 6**, pBH4-BPTI, **lane 7**, pBH4-SUMO-BPTI_{amber15} (addition of suppressor aa-tRNA), **lane 8**, pBH4-SUMO-BPTI_{amber15} (addition of suppressor aa-tRNA).

SUMO-BPTI in its full-length or truncated form could be produced in good yields. The protein could be further purified by means of Ni-NTA chromatography. After purification, expression yields of ~100 µg per mL reaction could be achieved. The successful expression of SUMO-BPTI was further confirmed by protein

4.2 Site-specific incorporation of fluorinated amino acids into BPTI

fingerprinting (Figure 4.2-11). Full-length SUMO-BPTI gives more than one band in the SDS PAGE and Western blot. The results of protein fingerprinting indicate that both bands are full-length SUMO-BPTI (for detected peptide fragments see Appendix). We propose that this may be due to different folding morphologies of BPTI.

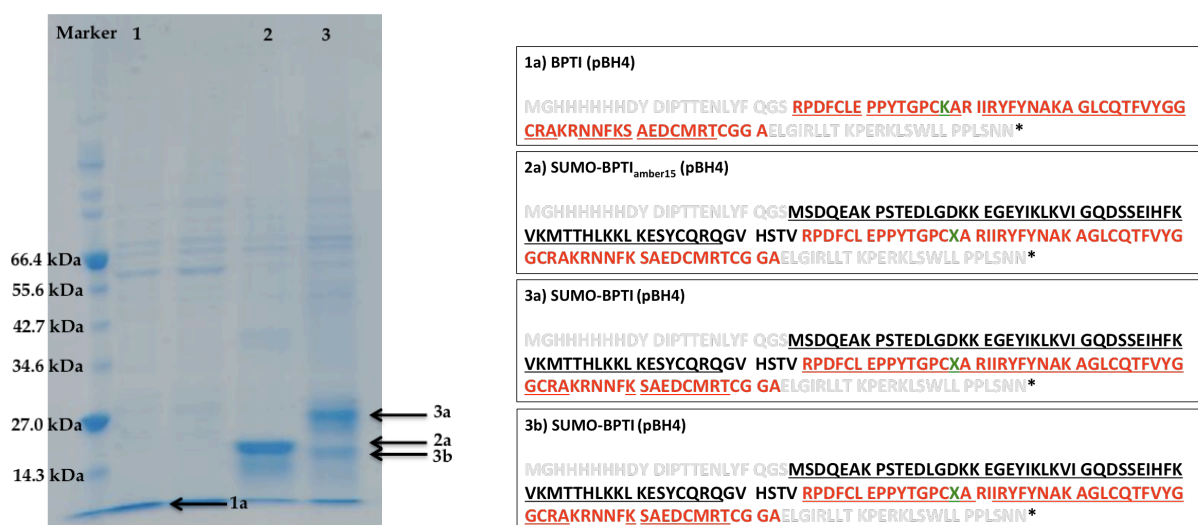


Figure 4.2-11 SDS PAGE analysis of purified SUMO-BPTI and BPTI produced by RTS 100 *E.coli* Disulfide Kit and protein sequence identified by protein fingerprinting

SDS-PAGE (left), lane 1, BPTI, lane 2, production of truncated SUMO-BPTI (addition of suppressor aa-tRNA), lane 3, full-length SUMO-BPTI; **sequence coverage (right)**, the BPTI sequence shown in red, SUMO sequence shown in black, His-tag and sequence of expression region shown in gray, position 15 is labeled as green X, the peptide fragments detected by LC-MS after in gel digestion are underlined, the sample numbers correspond to those presented in the SDS PAGE image and sequence coverage panel.

4.2.3.2 Cleavage of the SUMO fusion tag

Cleavage of the SUMO fusion tag was carried out by means of the SUMO protease 1. The results of cleavage indicate that the SUMO fusion tag can be cleaved after *in vitro* protein expression.

4.2 Site-specific incorporation of fluorinated amino acids into BPTI

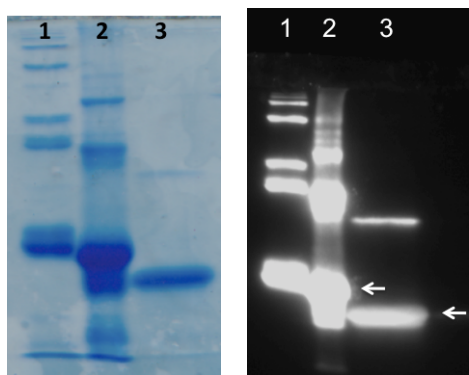


Figure 4.2-12 SDS PAGE and Western blot analysis of SUMO fusion tag cleavage

Left, SDS-PAGE analysis, **right**, Western Blot analysis, the numbering of SDS-PAGE and Western Blot indicates same samples. **lane 1**, SUMO protease 1, **lane 2**, purified SUMO-BPTI (arrow), **lane 3**, SUMO-tag (arrow) after the SUMO cleavage.

4.2.3.3 Protein expression by reconstitution of cell-free reactions

Due to the low yield of products according to the full-length size of protein, the incorporation of non-canonical amino acids into BPTI can not be confirmed. It is estimated that semi-synthesized misacylated tRNAs could have low compatible to the commercial kits. For instance, the kit reactions generally have much higher pH value, which qualifies translation but can results in deacyltion of suppressor aa-tRNAs. Therefore, the cell-free reactions were reconstituted by use of commercial available cell lysate and prepared reaction premix. The *amber* suppression was then tested in these reconstituted *in vitro* translation systems.^{119b, 123} Three different reaction premixes were prepared according to the literature Hecht *et al*, Schultz *et al*, and Sisido *et al*. (preparation see *Table 4.4-10* and *Table 4.4-11*).^{94c, 121, 123, 154, 160, 167} Unfortunately, all of these reconstituted cell-free reactions are failed to produce SUMO-BPTI.

4.3 Methods and experimental procedures

4.3.1 Chemical aminoacylation of non-canonical amino acids with yeast suppressor tRNA^{Phe}_{CUA}

4.3.1.1 General information about chemical synthesis, purification, and chemical characterization

All reagents of synthetic grade were used as supplied. Thin layer chromatography (TLC) was carried out on *Merck* silica gel 60 F254 plates. The fluorinated amino acids TfmAla, DfeGly, and TfeGly were synthesized according to literature procedures.^{30b, 168} Acetonitrile (MeCN) for synthesis and HPLC were purchased from *Acros Organics*. Deionized water for buffer solutions was prepared using the Milli-Q Advantage A10-System (*Millipore*). Standard chemicals that are not listed here were purchased from *Sigma-Aldrich* and *Fluka*.

Nuclear magnetic resonance

¹H, ¹³C, ¹⁹F, and ³¹P nuclear magnetic resonance (NMR) spectra were obtained on a Bruker ECX 400 MHz NMR spectrometer. Proton chemical shifts are reported in parts per million (ppm) and referenced to the residual proton peak of chloroform-d₁. Spectral coupling patterns are reported as follows, b: broad, s: singlet, t: triplet, and m: multiplet.

Mass spectrometry

Electrospray-ionization time-of-flight high-resolution mass spectrometry (ESI-TOF) was performed on an Agilent 6210 ESI-TOF (*Agilent Technologies*).

Reversed-phase high-performance liquid chromatography

Analytical reversed-phase high-performance liquid chromatography (RP-HPLC) was carried out on an Elite Lachrom HPLC system (*VWR & HITACHI*) equipped with an organizer, a L-2200 autosampler, two L-2130 pumps, and a L-2455 diode array detector. A capcel pak C18 column (type: SG 120 5 μm, size: 4.6 mmΦ x 250 mm,

4.3 Methods and experimental procedures

Shisheido) was used. Preparative RP-HPLC was carried out on a RP-HPLC (*Knauer GmbH*) equipped with a smartline manager 5000 system, two smartline pumps 1000, and an ultraviolet detector 2500.

4.3.1.2 Synthesis of pdCpA

6-*N*, 6-*N*, 2'-*O*, 3'-*O*-tetrabenzoyladenine (**pdCpA-1**) was synthesized by Dr. Allison Ann Berger. Di(2-cyanoethyl)*N,N*-diisopropylaminophosphoramidite ((iPr)₂N-P-(OCH₂CH₂CN)₂) and *N*-4-pentenoyloxy succinimide was synthesized by Benjamin Matt.

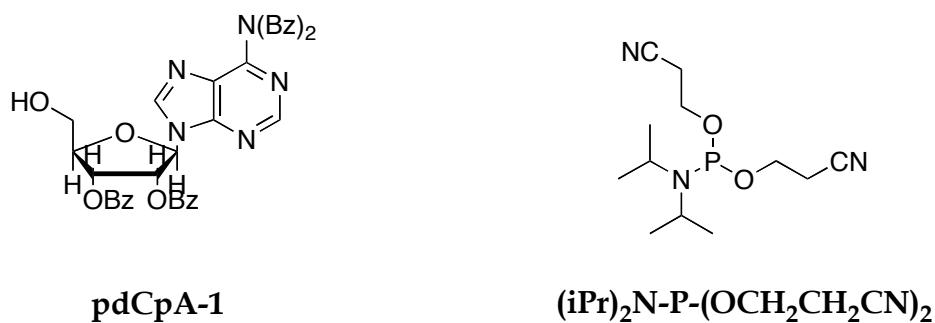


Figure 4.3-1 Structure of synthetic intermediates **pdCpA-1** and ((iPr)₂N-P-(OCH₂CH₂CN)₂)

4-*N*-benzoyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxycytidylyl (3'-5')-P-(2-cyanoethyl)-6-*N*, 6-*N*, 2'-*O*, 3'-*O*-tetrabenzoyladenine (**pdCpA-2**)

2088.0 mg **pdCpA-1** (3.04 mmol) and 502.9 mg tetrazole (7.04 mmol) were added into a dry 100 mL flask equipped with a magnetic stirrer and evaporated twice with toluene. 16 mL dry methylene chloride (DCM) was added and the solution was cooled to 0 °C in an ice bath. 2 g (2.40mmol) of 4-*N*-benzoyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxycytidinylyl-3'[(2-cyanoethyl) (*N,N*-diisopropyl)] phosphoramidite (dC-CE phosphoramidite, *Eurogentec*) was dissolved in cooled anhydrous DCM and added to this solution. The mixture was stirred and allowed to warm to room temperature. TLC was performed in acetone/hexane (v/v, 1/1) to monitor the reaction. After two hours the reaction was diluted in 60 mL DCM and extracted twice

4.3 Methods and experimental procedures

with 60 mL saturated NaCl solution. The organic layer was dried with MgSO₄ and concentrated under vacuum to give 3974.1 mg yellow foam. The crude material was dissolved in 2 mL acetone (another 2 mL acetone used for flask washing) and applied to a silica column (25 x 3 cm) packed with 30% acetone in *n*-hexane containing 0.2% triethylamine (TEA). Elution was accomplished with a gradient of 30-70% acetone in *n*-hexane (0.2% TEA). Fractions collected from 50-60% acetone in *n*-hexane contained the product; these were pooled together and concentrated under vacuum to give 2352.1 mg **pdCpA-2** (1.66 mmol) as a white foam (69%). ³¹P NMR (400 MHz, acetone-d₆) δ 139.33; mass spectrum (ESI-TOF): *m/z* 1438.43 (M+Na)⁺, 1454.41 (M+K)⁺, (C₇₈H₆₆N₉O₁₆P, calculated 1415.44).

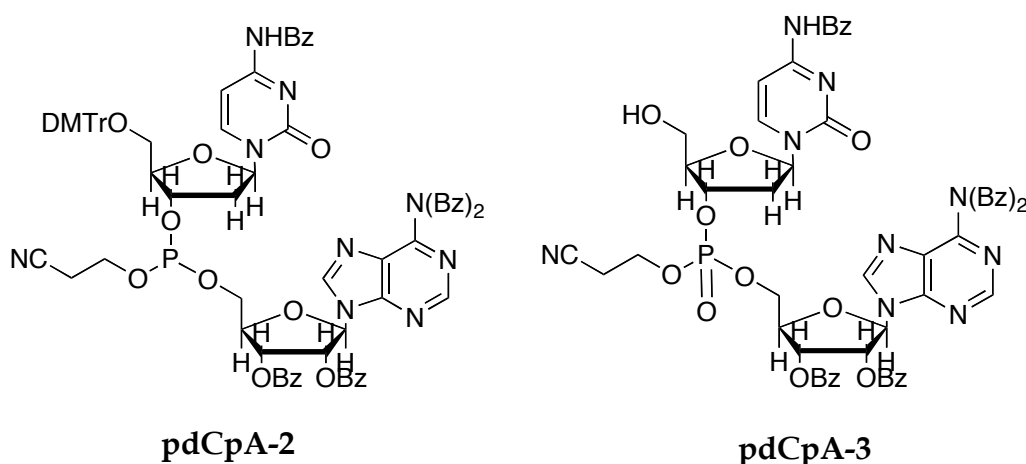


Figure 4.3-2 Structure of synthetic intermediates **pdCpA-2** and **pdCpA-3**

4-N-benzoyl-2'-deoxycytidyl(3'-5'-cyanoethylphosphate)-6-N, 6-N, 2'-O, 3'-O-tetrabenzoyladenine (pdCpA-3)

2352.1 mg **pdCpA-2** (1.66 mmol) was dissolved in 10 mL tetrahydrofuran (THF) in a 50 mL flask equipped with a magnetic stir bar. 643.4 mg (2.53 mmol) I₂ was dissolved in 4 mL of THF/H₂O/pyridine (v/v/v, 2/1/0.1) and added to the solution containing **pdCpA-2**. After stirring for 20 minutes, the reaction mixture was first concentrated under vacuum and then dissolved in 40 mL chloroform and extracted with 100 mL 0.2% sodium bisulfite solution. The aqueous phase was extracted with

4.3 Methods and experimental procedures

50 mL chloroform. The combined organic phase was washed with 50 mL saturated NaCl solution. A solution containing 617.3 mg (3.58 mmol) *p*-toluenesulfonic acid in 30 mL of MeOH/CH₂Cl₂, (v/v, 1/1) was added. After 30 minutes the reaction was quenched by the addition of 100 mL ice cold saturated sodium bicarbonate solution. The organic phase was washed first with 100 mL H₂O, and then with 100 mL saturated NaCl solution, and finally dried with MgSO₄ and concentrated under vacuum to give 1869.6 mg of crude product as a white/yellow foam. The crude material was dissolved in a small amount of acetone and applied to a silica column (25 x 3 cm) packed in 40% acetone in *n*-hexane. The product was eluted with a gradient of 40-100% acetone in *n*-hexane (100 mL per each concentration). The fractions collected at 75-100% acetone in *n*-hexane contained the product. The appropriate fractions were pooled together and concentrated under vacuum to give 1102.2 mg **pdCpA-3** (0.98 mmol) as a white foam (59% yield). ¹H NMR (400 MHz, acetone-d₆) 2.05 (t, 3H), 2.00-2.10 (m), 2.85-3.05 (m), 5.05 (s, 1H), 5.60 (s, 1H), 6.80-6.90 (m), 7.10-7.20 (m), 7.30-7.40 (m), 7.5 (d); ³¹P NMR (400 MHz, acetone-d₆) δ -2.0; mass spectrum (ESI-TOF): *m/z* 1130.31 (M+H)⁺, 1152.29 (M+Na)⁺, (C₅₇H₄₈N₉O₁₅P, calculated 1129.30).

4-*N*-benzoyl-5'-(di-2-cyanoethylphosphate) 2'-deoxycytidylyl (3'-5'-cyanoethylphosphate)-6-*N*, 6-*N*, 2'-*O*, 3'-*O*-tetrabenzoyladenine (**pdCpA-4**)

1102.2 mg (0.98 mmol) **pdCpA-3** and 140.8 mg (2.0 mmol) tetrazole were added into a 50 mL flask and coevaporated four times with toluene. The mixture was dissolved in 12 mL anhydrous DCM and cooled to 0 °C. 353.3 mg (1.3 mmol) of di(2-cyanoethyl)-*N,N*-diisopropylaminophosphoramidite ((iPr)₂-N-P-(OCH₂CH₂CN)₂), was added via syringe. The reaction was allowed to warm to room temperature and monitored by TLC. After 45 minutes, a solution containing 613.2 mg (2.4 mmol) I₂ in THF/H₂O/pyridine (v/v/v, 2/1/0.1) was added and the mixture was stirred for a further 15 minutes. The reaction mixture was then concentrated under vacuum. The black residue was dissolved in 20 mL chloroform and extracted with 100 mL 0.05% sodium bisulfite solution. The organic layer was washed once with H₂O and once

4.3 Methods and experimental procedures

with saturated NaCl solution. The organic phase was then dried with MgSO₄ and concentrated under reduced pressure to give 1248.3 mg **pdCpA-4** as white/yellow foam. The crude product was carried into the cyanoethyl deprotection step without further purification. ³¹P NMR and MS were performed to confirm the product. ³¹P NMR (400 MHz, acetone-d₆) δ: -0.66, -1.19; mass spectrum (ESI-TOF): *m/z* 1316.28 (M+H)⁺, 1338.29 (M+Na)⁺, 1354.26 (M+K)⁺, (C₆₃H₅₅N₁₁O₁₈P₂, calculated 1315.32).

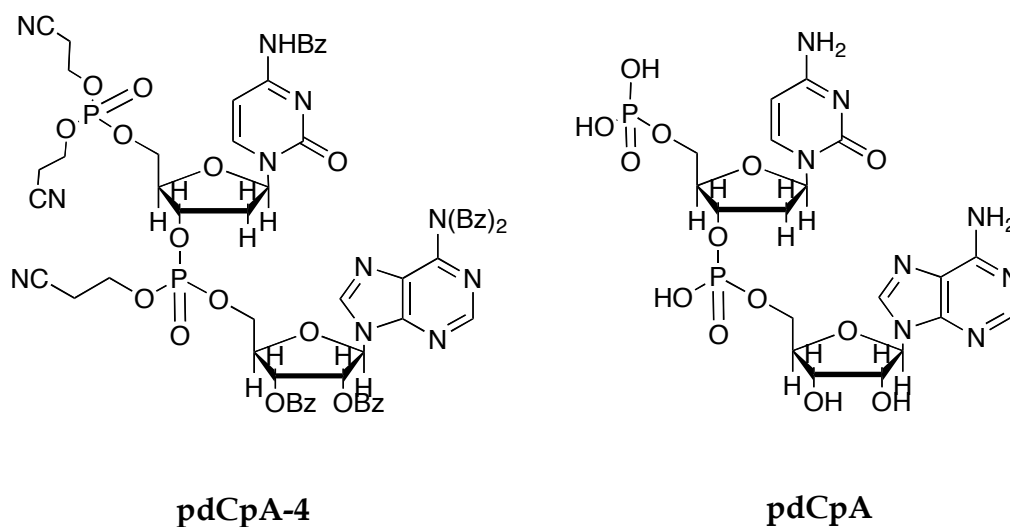


Figure 4.3-3 Structure of synthetic intermediate **pdCpA-4** and **pdCpA**

5'-phospho-2'-deoxycytidylyl (3'-5')adenosine (**pdCpA**)

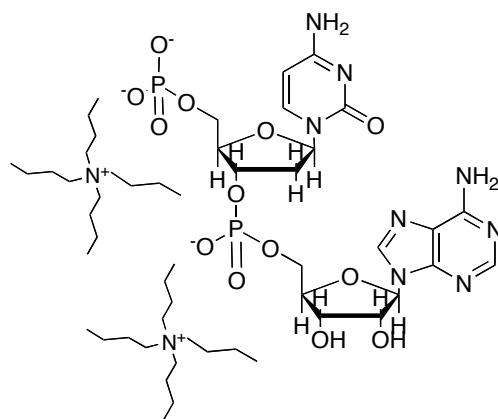
1248.3 mg **pdCpA-4** was dissolved in 10 mL dioxane and 40 mL methanol and added to 100 mL 30% ammonium hydroxide in a sealed flask. The reaction mixture was stirred at 55 °C overnight, then concentrated under reduced pressure. The residue was dissolved in 25 mM aqueous ammonium acetate buffer (pH 5.0) and purified by RP-HPLC with a flow rate of 20 mL/min and a gradient of 10-80% acetonitrile in 25 mM ammonium acetate (pH 5) over 30 minutes to give 314 mg ammonium **pdCpA** as a white solid. ¹H NMR (400 MHz, D₂O) δ: 1.60-1.74 (m, 1H), 2.15-2.25 (m, 1H), 3.86 (s, 2H), 3.98 (s, 2H), 4.11 (s, 1H), 4.22 (s, 1H), 4.42 (t, 1H, *J* = 8 Hz), 5.80-6.02 (m, 3H), 7.60 (d, *J* = 8), 8.05 (s, 1H), 8.34 (s, 1H); ³¹P NMR (400 MHz, D₂O) δ: -0.44, 2.45; mass

4.3 Methods and experimental procedures

spectrum (ESI-TOF): m/z 637.13 (M+H)⁺, 1273.24 (2M+H)⁺, (C₁₉H₂₆N₈O₁₃P₂, calculated 636.11).

Ion exchange chromatography of pdCpA

The *tetra-n*-butylammonium (TBA) salt of pdCpA was prepared by means of ion exchange chromatography. The ion-exchange resin (Dowex 50 W x 8, 20-50 mesh, *Fluka*) was first washed with H₂O, then stirred in a 20% aqueous TBA solution (from a stock solution: 1. 0 M aqueous solution, *Alfa Aesar*, or tris(*tetra*n-butylammonium)-hydrogen pyrophosphate, *Sigma-Aldrich*[®]) for 2 hours. Afterward, the beads were packed into a column (1 x 10 cm) and washed with H₂O until pH of washing elute is neutral. A small amount of pdCpA (100-200 mg) was dissolved in 0.5 mL H₂O and loaded onto the column. pdCpA-TBA was eluted in H₂O and monitored by TLC with UV detection. The ratio of pdCpA and TBA of lyophilized pdCpA-TBA salt was determined by NMR. Normally, an additional amount of TBA hydroxide was directly added to an aqueous solution of pdCpA-TBA to ensure that the ratio of TBA to pdCpA is greater than 2.2. The solution was lyophilized and yielded the product as a fluffy white powder. ¹H NMR (400 MHz, D₂O) δ: 0.80 (t, 3H, *J* = 8.0 Hz), 1.21 (dt, 2H, *J* = 8, *J* = 8), 1.40-1.58 (m, 2H), 1.70-1.80 (td, *J* = 4, *J* = 8), 1H, 2.22-2.32 (m, 1H), 2.98-3.10 (m, 2H), 3.88 (s, 2H), 3.99 (s, 2H), 4.14 (s, 1H), 4.22, (s, 1H), 4.41 (t, 1H, *J* = 4), 5.92 (d, 3H), 7.71 (d, 1H, *J* = 8), 8.05 (s, 1H), 8.34 (s, 1H); ³¹P NMR (400 MHz, D₂O) δ: -0.77, 0.44. The product was dissolved in anhydrous DMF (*Sigma-Aldrich*) and stored in a sealed vial filled with argon at -20 °C (compound is stable for several weeks).



tetrabutylammonium-pdCpA

Figure 4.3-4 Structure of tetrabutylammonium salt of pdCpA

4.3.1.3 Synthesis of *N*-(4-pentenoyl)-amino acid cyanomethyl ester

Synthesis of *N*-pentenoyl (*R,S*)-2-amino-2-methyl-3,3,3-trifluoropropanoic acid cyanomethyl ester (**3a**)

361.8 mg TfmAla (2.3 mmol) was dissolved in 10 mL DMF and treated with 0.5 mL *N*-4-pentenoyl chloride (4.6 mmol) and 847.9 mg 4-dimethylaminopyridine (DMAP, *Acros Organics*). After stirring at room temperature for 5 days, the mixture was diluted in 50 mL of 1 M aqueous NaHSO₄ and extracted three times with ethyl acetate. The combined organic phase was dried with MgSO₄ and concentrated under vacuum to give 240.1 mg compound **2a** (Figure 4.2-2) as a brown oil. The crude product was carried forward without any further purification. 149 mg crude compound **2a** (ca. 0.6 mmol) was dissolved in 3 mL MeCN and treated with 0.2 mL iodoacetonitrile (2.7 mmol) (*Fluka*) and 0.25 mL triethylamine (TEA) (1.8 mmol). The mixture was stirred at room temperature overnight. If the reaction did not reach completion, as determined by TLC, 1 equivalent of iodoacetonitrile and 1 equivalent of TEA were added and the reaction allowed more time; after the reaction was complete, the mixture was diluted in 25 mL ethyl acetate and washed twice with 25

4.3 Methods and experimental procedures

mL 1M NaHSO₄ aqueous solution. The organic layer was dried with MgSO₄ and concentrated under reduced pressure. The crude product was purified on a silica gel column (13 x 2 cm) that was packed and eluted with ethyl acetate and *n*-hexanes (v/v 1/1). Appropriate fractions were pooled together and concentrated under vacuum to give compound **3a** as a yellow oil: yield 87.4 mg (22%); ¹H NMR (400 MHz, CDCl₃) δ: 1.70 (s, 3H), 2.28-2.42 (m, 4H), 4.76 (d, 1H, *J* = 15.7 Hz), 4.83 (d, 1H, *J* = 15.7 Hz), 4.95-5.13 (m, 2H), 5.71-5.83 (m, 1H), 6.30 (s, H); ¹³C NMR (400 MHz, CDCl₃) δ: 18.58 (-CH₃), 29.10 (-CH₂-CH₂-), 34.92 (-CH₂-CH₂-), 49.87 (-CH₂-CN), 61.23 (-NH-CH-), 113.75 (-CH₂-CN) 116.31 (CH₂=CH-), 123.73 (q, *J* = 284.0 Hz, -CH₂-CF₃), 136.31 (CH₂=CH-), 165.29 (-CO-O-), 172.28 (-CO-NH-); ¹⁹F NMR (400 MHz, CDCl₃) δ: -76.2267 (s); mass spectrum (ESI-TOF) *m/z* 279.0929 (M+H)⁺ (C₁₁H₁₃F₃N₂O₃, calculated 278.0878).

Synthesis of *N*-(4-pentenoyl)-(S)-2-aminobutyric acid cyanomethyl ester (**3b**)

A mixture of 6 mL H₂O and 6 mL dioxane containing 226.4 mg Abu (2.2 mmol) (*Bachem*) and 357.2 mg NaHCO₃ (4.3 mmol) was treated with 465.1 mg *N*-4-pentenoyloxy succinimide (2.4 mmol) and stirred at room temperature for 17 hours. The reaction mixture was diluted with 50 mL NaHSO₄ aqueous solution (1 M) and extracted three times with ethyl acetate (50 mL). The combined organic phase was dried with MgSO₄ and concentrated under reduced pressure. Compound **2b** was obtained as a brown oil. The crude product was dissolved in 4 mL MeCN and treated with 1.5 mL TEA (10 mmol) and 0.72 mL iodoacetonitrile (10 mmol). After stirring at room temperature for 20 hours, the reaction mixture was dissolved in 50 mL ethyl acetate and washed twice with 1 M NaHSO₄ aqueous solution. The organic layer was dried with MgSO₄ and concentrated under reduced pressure. The crude product was purified on a silica gel column (12 x 3 cm) with ethyl acetate and *n*-hexane (v/v, 4/1). Compound **3b** was obtained as a gray solid: yield 399.8 mg (81.25%); ¹H NMR (400 MHz, CDCl₃) δ: 0.94 (t, 3H, *J* = 7.4 Hz), 1.66-1.78 (m, 1H), 1.83-1.95 (m, 1H), 2.29-2.41 (m, 4H), 4.58 (td, 1H, *J* = 7.4 Hz, *J* = 5.4 Hz), 4.70 (d, 1H, *J* = 15.7 Hz), 4.82 (d, 1H, *J* = 15.7 Hz), 4.98-5.10 (m, 2H), 5.74-5.86 (m, 1H), 6.06 (bd, 1H, *J* = 7.0 Hz); ¹³C NMR

4.3 Methods and experimental procedures

(400MHz, CDCl₃) δ : 9.74 (-CH₃), 25.30 (-CH₂CH₃), 29.43 (CH₂=CH₂-CH₂-), 35.45 (-CH₂-CH₂-CO-), 48.90 (-CH₂-CN), 53.07 (-NH-CH-), 114.05 (-CN), 115.92 (CH₂=CH-), 136.81 (CH₂=CH-), 171.26 (-CO-O-), 172.45 (-CO-NH-); mass spectrum (ESI-TOF), m/z 225.1263 (M+H)⁺ (C₁₁H₁₆N₂O₃, calculated 224.1161).

Synthesis of *N*-(4-pentenoyl)-(*S*)-2-amino-4,4-difluorobutyric acid cyanomethyl ester (**3c**)

60 mg DfeGly (0.43 mmol) was dissolved in a mixture containing 1.5 mL water and 1.5 mL dioxane. To this mixture was added 72.3 mg NaHCO₃ (0.86 mmol) and 100 mg *N*-4-pentenoyloxy succinimide (0.51 mmol), and the reaction was stirred at room temperature for 18 hours. The mixture was treated with 12 mL 1M NaHCO₃ aqueous solution and extracted three times with 12 mL ethyl acetate. The combined organic phase was dried with MgSO₄ and concentrated under reduced pressure. The crude product was dissolved in 1.5 mL MeCN and treated with 0.35 mL TEA (2.5 mmol) and 0.16 mL iodoacetonitrile (2.2 mmol). After stirring at room temperature for 20 hours, the reaction mixture was diluted in 12 mL ethyl acetate and washed three times with 12 mL 1M aqueous NaHSO₄. The organic layer was dried with MgSO₄ and concentrated under reduced pressure. The crude product was purified on a silica gel column (20 x 1.5 cm) and eluted stepwise with 50 mL ethyl acetate in *n*-hexane (v/v, 1/2) and 50 mL ethyl acetate in *n*-hexane (v/v, 2/3) to give compound **3c** as a yellow oil: yield 66 mg (59%). ¹H NMR (400 MHz, CDCl₃) δ : 2.24-2.56 (m, 6H), 4.74 (d, 1H, J = 15.7 Hz), 4.80 (d, 1H, J = 15.7 Hz), 4.71-4.83 (m, 1H), 4.97-5.09 (m, 2H), 5.73-5.85 (m, 1H), 5.93 (tt, 1H, J = 55.8 Hz, J = 4.3 Hz), 6.59 (bd, 1H, J = 7.5 Hz); ¹³C NMR (400 MHz, CDCl₃) δ : 29.26 (CH₂=CH-CH₂-), 35.23 (-CH₂-CO-), 35.63 (t, J = 22.0 Hz, CH₂-CF₂H), 47.43 (t, J = 5.7 Hz, -NH-CH-), 49.46 (-O-CH₂-), 113.86 (-CH₂-CN), 115.15 (t, J = 238.5 Hz, -CH₂-CF₂H), 115.99 (CH₂=CH-), 136.59 (CH₂=CH-) 169.69 (-CO-O-), 172.76 (-CO-NH-); ¹⁹F NMR (400 MHz, CDCl₃) δ : -116.46 (tdd, J = 320.0 Hz, J = 58.6 Hz, J = 17.1 Hz), -115.70 (tdd, J = 320.0 Hz, J = 58.6 Hz, J = 17.1 Hz); mass spectrum (ESI-TOF), m/z 261.1045 (M+H)⁺ (C₁₁H₁₄F₂N₂O₃, calculated 260.0972).

4.3 Methods and experimental procedures

Synthesis of *N*-(4-pentenoyl)-(*S*)-2-amino-4,4,4-trifluorobutyric acid cyanomethyl ester (**3d**)

60.1 mg NaHCO₃ (0.72 mmol) and 53.5 mg TfeGly (0.34 mmol) were dissolved in a mixture containing 1.5 mL water and 1.5 mL dioxane to which 87.3 mg *N*-4-pentenoyloxy succinimide (0.44 mmol) was added. After stirring for 19 hours, the reaction mixture was diluted with 15 mL of 1 M aqueous NaHSO₄ solution and extracted three times with 15 mL ethyl acetate. The combined organic phase was dried with MgSO₄ and concentrated under reduced pressure. The crude product was dissolved in 1.5 mL MeCN, and 0.25 mL TEA (1.8 mmol) and 0.15 mL iodoacetonitrile (2 mmol) were added. The resulting solution was stirred for 48 hours at room temperature, then diluted with 12 mL ethyl acetate and washed twice with 12 mL of 1 M NaHSO₄ solution. The organic layer was dried with MgSO₄ and concentrated under reduced pressure. The crude product was purified on a silica gel column (13 x 2 cm) and eluted with 4:1 ethyl acetate:*n*-hexane. Appropriate fractions were pooled together and concentrated under vacuum to give compound **3d** as a white solid: yield 71.2 mg (75.3%). ¹H NMR (400 MHz, CDCl₃) δ: 2.29-2.43 (m, 4H), 2.65-2.89 (m, 2H), 4.77 (d, 1H, *J* = 15.8 Hz), 4.83 (d, 1H, *J* = 15.8 Hz), 4.80-4.84 (m, 1H), 5.00-5.12 (m 2H), 5.74-5.88 (m, 1H), 6.32 (bd. 1H, *J* = 7.4 Hz); ¹³C NMR (400 MHz, CDCl₃) δ: 29.16 (CH₂=CH-CH₂-), 34.89 (-NH-CH-), 35.03 (q, *J* = 28.0 Hz, -CH₂-CH₂-CO-), 47.29 (q, *J* = 2.7 Hz, -CH₂-CF₃), 49.63 (-O-CH₂-), 113.52 (-CH₂-CN), 116.10 (CH₂=CH-), 125.50 (q, *J* = 276.0 Hz, -CH₂-CF₃), 136.50 (CH₂=CH-), 168.67 (-CO-O-), 172.44 (-CO-NH-); ¹⁹F NMR (400 MHz, CDCl₃) δ: -62.98 (t, *J* = 9.8 Hz); mass spectrum (ESI-TOF), *m/z* 279.0929 (M+H)⁺ (C₁₁H₁₃F₃N₂O₃, calculated 278.0878).

4.3.1.4 Synthesis of mono-2' (3')-O-[*N*-(4-pentenoyl)aminoacyl]-pdCpAs and 2' - 3' -bis-O-[*N*-(4-pentenoyl)aminoacyl]-pdCpAs

Synthesis of 2' (3')-O-[*N*-(4-pentenoyl)-(*R,S*)-2-amino-2-methyl-3,3,3-trifluoropropyl]-pdCpA (**4a**) and bis-2',3'-O-[*N*-(4-pentenoyl)-(*R,S*)-2-amino-2-methyl-3,3,3-trifluoropropyl]-pdCpA (**4a'**)

4.3 Methods and experimental procedures

300 μ L anhydrous DMF containing 20 mg TBA-pdCpA (31.4 μ mol) and 50 μ L TEA were added under argon into a flame-dried conical vial containing 16 mg compound **3a** (57.6 μ mol). The reaction mixture was stirred at 40 °C overnight. The reaction was monitored by analytical RP-HPLC. Upon completion, the reaction mixture was diluted with a mixture consisting of 50 mM NH₄OAc and MeCN (v/v, 2/1, pH 4.5) to a total volume of 500 μ L and purified using RP-HPLC. The compounds were eluted with a gradient of 0-75% MeCN in 50 mM aqueous NH₄OAc over a period of 40 minutes at a flow rate of 20 mL/min (detection at 260 nm). The appropriate fractions were combined and lyophilized to give compound **4a** as a white solid (retention time 15.5 minutes on preparative reverse-phase C18 HPLC column): yield 12 mg (45%), mass spectrum (ESI-TOF): m/z 858.1796 (M+H)⁺ (C₂₈H₃₆F₃N₉O₁₅P₂, calculated 857.1758). Compound **4a'** was obtained as a white solid (retention time 21.5 min on preparative HPLC, C18 column): yield 15.4 mg (50.5%), mass spectrum (ESI-TOF): m/z 1079.2445 (M+H)⁺ (C₃₇H₄₆F₆N₁₀O₁₇P₂, calculated 1078.2421).

Synthesis of 2'(3')-O-[N-(4-pentenoyl)-(S)-2-amino butyl]-pdCpA (**4b**) and bis-2',3'-O-[N-(4-pentenoyl)-(S)-2-amino butyl]-pdCpA (**4b'**)

Under an argon atmosphere, 200 μ L solution containing 25.5 μ mol pdCpA in anhydrous DMF was added to a flame-dried vial equipped with a magnetic stir bar. To this solution was added 14.4 mg compound **3b** (64.3 μ mol) and 50 μ L TEA. The reaction mixture was stirred at 40 °C overnight. 5 μ L aliquots of the reaction mixture were diluted with 45 μ L 2:1 50 mM NH₄OAc: MeCN solution (pH 4.5) and analyzed by RP-HPLC. Once the reaction was complete, the reaction mixture was diluted with 500 μ L 2:1 50mM NH₄OAc solution:MeCN, pH 4.5, and purified by means of RP-HPLC (C18 column) with a gradient of 10-80% MeCN in 50 mM NH₄OAc buffer (pH 4.5) over a period of 40 minutes at a flow rate of 20 mL/min (detection at 260 nm). The appropriate fractions were lyophilized and gave compound **4b** as a colorless solid (retention time 7.4 minutes by preparative reverse-phase C18 HPLC column): yield 6.9 mg (33.7%), mass spectrum (ESI-TOF): m/z 804.2135 (M+H)⁺ (C₂₈H₃₉N₉O₁₅P₂, calculated 803.2041). Compound **4b'** was obtained as a colorless solid (retention time

4.3 Methods and experimental procedures

13.1 min on preparative HPLC, C18 column): yield 2.4 mg (9.6%), mass spectrum (ESI-TOF): m/z 971.2965 (M+H)⁺ (C₃₇H₅₂N₁₀O₁₇P₂, calculated 970.2987).

Synthesis of 2'(3')-O-[N-(4-pentenoyl)-(S)-2-amino-4,4-difluorobutyl]-pdCpA (**4c**)

To a flame-dried conical vial containing 12.6 mg compound **3c** (48.5 μmol) was added 200 μL of a solution of 10.8 mg pdCpA (16.98 μmol) in anhydrous DMF to which 50 μL TEA had been added. The reaction mixture was stirred at 40 °C overnight; upon completion, as determined by RP-HPLC, the reaction mixture was diluted with 2:1 50 mM aqueous NH₄OAc:MeCN, pH 4.5, to a total volume of 500 μL and purified by RP-HPLC. The compound was eluted with a gradient of 0-75% MeCN in 50 mM NH₄OAc over a period of 40 minutes at a flow rate of 20 mL/min (monitored at 260 nm). The appropriate fractions were lyophilized to give compound **4c** as a colorless solid (retention time 19 minutes on preparative RP-HPLC, C18 column): yield 9.3 mg (65%), mass spectrum (ESI-TOF): m/z 840.1857 (M+H)⁺ (C₂₈H₃₇F₂N₉O₁₅P₂, calculated 839.1852).

Synthesis of 2'(3')-O-[N-(4-pentenoyl)-(S)-2-amino-4,4,4-trifluorobutyl]-pdCpA (**4d**) and bis-2',3'-O-[N-(4-pentenoyl)-(S)-2-amino-4,4,4-trifluorobutyl]-pdCpA (**4d'**)

To a flame-dried conical vial containing 16.7 mg **3d** (60.1 μmol) was added 200 μL of a solution containing 12.35 mg pdCpA (19.4 μmol) in anhydrous DMF to which 50 μL of TEA had been added. The reaction mixture was stirred at 40 °C overnight. Upon completion, as determined by analytical RP-HPLC, the reaction mixture was diluted with 2:1 50 mM aqueous NH₄OAc:MeCN, pH 4.5, to a total volume of 500 μL and purified by RP-HPLC. The compound was eluted with 10-80% MeCN in 50 mM aqueous NH₄OAc over a period of 45 minutes at a flow rate of 20 mL/min (monitored at 260 nm). The appropriate fractions were lyophilized and gave compound **4d** as a white solid (retention time 9.5 minutes on preparative reverse-phase C18 HPLC column): yield 6.7 mg (13%), mass spectrum (ESI-TOF): m/z 858.1796 (M+H)⁺ (C₂₈H₃₆F₃N₉O₁₅P₂, calculated 857.1758). Compound **4d'** was also obtained as a white solid (retention time 13.1 minutes on preparative HPLC, C18

4.3 Methods and experimental procedures

column): yield 10.2 mg (48.7%), mass spectrum (ESI-TOF): m/z 1079.2612 (M+H)⁺ (C₃₇H₄₆F₆N₁₀O₁₇P₂, calculated 1078.2421).

Determination of concentration for misacylated *N*-4-pentenoyl-aminoacyl-pdCpAs

The concentration of the solution containing pdCpA or *N*-4-pentenoyl-aminoacyl-pdCpA can be determined by UV absorption, assuming $\epsilon_{264} = 24,500 \text{ cm}^{-1}\text{M}^{-1}$ for pdCpA.¹⁵⁴

4.3.1.5 T4 RNA ligase-mediated ligation

Linearization of plasmid DNA PYRNA8 and DNA analysis by restriction enzymes

Plasmid DNA PYRNA8 encoding yeast phenylalanine tRNA, which contains the T7 promoter, was linearized by treatment with the restriction enzyme FokI (*New England Biolabs*) to obtain a substrate for the run-off *in vitro* transcription of truncated suppressor tRNA. Linearization with BstNI (*New England Biolabs*) results in templates for the production of full-length suppressor tRNA as control. Table 4.3-1 represents typical analytical scale restriction digestion reactions. For the preparative scale, all components and the overall volume of reaction were increased according to the amount of plasmid DNA.

Table 4.3-1 Reaction preparation for enzymatic linearization of PYRNA8

For production of full-length tRNA		For production of truncated tRNA	
Components	Volume	Components	Volume
DNA	1 μg	DNA	1 μg
BstNI	1 μL	FokI	1 μL
BSA	0.5 μL	BSA	0.5 μL
NEBuffer 2 (10 X)	2 μL	NEBuffer 4 (10 X)	2 μL
H ₂ O	X μL	H ₂ O	X μL
Total	20 μL		20 μL
Incubate at 60 °C for 1 hour		Incubate at 37 °C for 1 hour	

4.3 Methods and experimental procedures

Run-off *in vitro* transcription of truncated suppressor tRNA^{Phe}_{CUA}

Table 4.3-2 Run-off *in vitro* transcription reactions under various conditions^a

Components	Variant 1 (μL)	Variant 2 (μL)	Variant 3 (μL)	Variant 4 (μL)
10 X Buffer	2	2	2	2
DTT (100 mM)	2	2	2	2
ATP (100 mM)	1.5	1.5	1.5	1.5
CTP (100 mM)	1.5	1.5	1.5	1.5
UTP (100 mM)	1.5	1.5	1.5	1.5
GTP (100 mM)	1.5	1	1.5	1
GMP^b (400 mM)	0	1	0	1
Formamide^c	0	0	1	1
DNA template^d	1	1	1	1
T7 RNA polymerase	2	2	2	2
RNase inhibitor	0.5	0.5	0.5	0.5
H₂O	6.5	6	5.5	5
Total	20			

a) Based on Choudhury *et. al.* and short *et. al.*,^{121, 167} a standard reaction has final volume with 20 μL, in practical, the reactions were prepared to final volume with 100 μL. b) 5'-guanosine monophosphate (GMP disodium salt *AppliChem*), sterilization by filtration with 200 nm membrane, c) formamide (deionized, molecular biology grade, *AppliChem*) d) concentration of DNA template is generally about 1μg/μL (for kit manual, see *Appendix*).

The truncated yeast suppressor tRNA^{Phe}_{CUA}(-CA) and the full-length tRNA^{Phe}_{CUA} were synthesized in run-off *in vitro* transcription reactions that made use of the AmpliScribe™ T7 High Yield Transcription Kit (*EPICENTRE® Biotechnologies*). A typical reaction was incubated at 42 °C for at least 6 hours, and it was observed that incubation overnight maximized yields.

Purification of truncated suppressor tRNA_{CUA}(-CA)

After overnight incubation RNase-free DNase (supplied with transcription kit) was added to the reaction mixture, according to the manufacturer's recommendations, and incubated for a further 15 minutes at 37 °C. The reaction solution was then extracted with an equal volume of Roti®-Phenol/Chloroform/Isoamylalcohol (25:24:1, pH 7.5-8.0, *Carl Roth*; centrifugation at 13,300 x g for one minute). The aqueous phase was transferred to a new 1.5 mL tube and extracted with an equal volume of chloroform (centrifugation at 13,300 x g for one minute). The aqueous

4.3 Methods and experimental procedures

phase was then applied to a diethylaminoethanol (DEAE) sepharose anion-exchange column for further purification. 1-1.5 mL DEAE sepharose (*Sigma-Aldrich*[®]), was packed into a small column that was washed with 1 mL H₂O and equilibrated with 1 mL 0.1 M NaOAc (pH 5.2). The chloroform extracted tRNA solution was loaded onto the column and eluted with NaOAc solution (pH 5.2) containing NaCl: stepwise gradient of 0.1 – 1.0 M (500 µL for each NaCl concentration). tRNA in each fraction was precipitated by adding 750 µL isopropanol and cooling at -80 °C for at least 30 minutes; pellets were obtained by centrifugation at 13,300 x g at 4 °C for 15 minutes. Precipitates were washed with 70% ethanol, centrifuged at 13,300 x g at 4 °C for 10 minutes and allowed to air dry for 5 minutes. tRNAs were dissolved in 0.1 M NaOAc and analyzed by acidic urea PAGE. Fractions containing truncated suppressor tRNA were pooled together.

Determination of tRNA concentration

The concentration of tRNA was determined by means of a NanoDrop ND-2000 at 260 nm.

General procedure for synthesis of N-4-pentenoyl-aminoacyl-tRNAs

A mixture (final volume 100 µL for a standard reaction) containing 40 nmol aminoacyl-pdCpA, 10 µg truncated tRNA (0.4 nmol), 15 mM MgCl₂, 1 mM ATP, 50 ng/µL BSA, 10% DMSO (v/v), 50 mM HEPES buffer (pH 7.5) and 100 units of T4-RNA ligase (T4-RNA ligase 1, ssRNA ligase, *New England Biolabs*) was incubated at 37 °C for 30 minutes. The reaction was quenched by the addition of 5 µl 3 M NaOAc solution (pH 5.2) To remove protein contaminants from the misacylated aminoacyl-tRNAs, an additional extraction step with phenol-chloroform (Roti[®]-aqua phenol pH 4.5-5.0, *Carl Roth*) was carried out. N-4-pentenoyl-aminoacyl-tRNA was precipitated by addition of 3 M NaOAc solution (pH 5.2) to a final concentration of 0.3 M and 2.5 volume equivalents of ice-cold ethanol and incubated at -80 °C for 30 minutes. Centrifugation at 4 °C at 13,000 x g for 10 minutes yielded a pellet that was washed

4.3 Methods and experimental procedures

with 250 μ L 70% ethanol. After centrifugation for 5 minutes at 4 $^{\circ}$ C at 13,000 \times g, the pellet was immediately lyophilized and stored at -80 $^{\circ}$ C.¹⁵⁴

Stability test of misacylated suppressor tRNAs

tRNA solution (3 μ L of a stock solution of 2 μ g/ μ L in water) was added into 6 μ L water or buffer and incubated at 37 $^{\circ}$ C. HEPES buffer (pH 7.5) and cell-free premix solution (*Promega*) were used for test. As a control sample, 1 μ L of the tRNA solution was kept on ice. At time points 15 min, 30 min, and 1 hour, a 3 μ L aliquot was taken and quenched with NaOAc (pH 5.2). All samples were mixed with RNA sample buffer and immediately analyzed by denaturing acidic PAGE.

Deprotection of N-4-pentenoyl-aminoacyl-tRNAs

About 20 μ g N-4-pentenoyl aminoacyl-tRNA was dissolved in 10 μ L water (or 1 mM sodium acetate pH 5.2) and mixed with 10 μ L 25 mM iodine in THF/water (v/v, 1/1). The solution was incubated at 25 $^{\circ}$ C for 15 minutes. Misacylated tRNA was precipitate precipitated by addition of 3 M NaOAc solution (pH 5.2) to a final concentration of 0.3 M and 2.5 volume equivalents of ice-cold ethanol and incubated at -80 $^{\circ}$ C for 30 minutes. Centrifugation at 4 $^{\circ}$ C at 13,000 \times g for 10 minutes yielded a pellet that was washed with 250 μ L 70% ethanol. After centrifugation for 5 minutes at 4 $^{\circ}$ C at 13,000 \times g, pellet was dissolved in \sim 5 μ L water and immediately added to *in vitro* protein translation reaction.

Acidic urea polyacrylamide gel electrophoresis

Acidic urea PAGE was used to analyze both the truncated tRNAs produced by run-off *in vitro* transcription and the products of T₄-RNA ligase-mediated RNA ligation.¹⁵⁷ The tRNA sample to be analyzed was mixed with an equal volume of sample buffer (supplied with Low Range ssRNA Ladder, *New England Biolabs*) and heated at 70 $^{\circ}$ C for ten minutes. Afterward, samples were briefly chilled on ice, spun down and loaded onto a 1-mm-thick 7% gel (acrylamide:bisacrylamide, 19:1) containing 8 M urea. Electrophoresis was carried out in 0.1 M sodium acetate (pH 5.2) for 3-4 hours at 200 volts (the gel was normally prewarmed at 200 volts for 1

4.3 Methods and experimental procedures

hour). Gels were stained with Stains-All (*Sigma-Aldrich*) overnight in the dark and destained with water in the light. RiboRuler™ Low Range RNA Ladder (*Fermentas*) and Low Range ssRNA Ladder (*New England Biolabs*) were used as markers.

Table 4.3-3 Preparation of 7% acidic PAGE containing 8 M urea for tRNA analysis

Components	Volume
Gel A ^a	5.542 mL
Gel B ^a	4.375 mL
Urea	9.6 g
10% APS	1.75 mL
4.5 M NaOAc	556 µL
TEMED	37.5 µL
H ₂ O	X mL
Total	25 mL

a) Gel A: 30% acrylamide solution, Gel B: 2% bisacrylamide solution (Rotiphorese® Gel, *Carl Roth*), crosslinking of gel is 5%. (calculated with the formula: http://www.carlroth.com/website/en-nl/pdf/PAGE_E.pdf)

Stock solution of Stains-All: 0.1% Stains-All in formamide

Working solution of Stains-All (100 mL for one gel): 10 mL stock solution, 25 mL isopropanol, 750 µL 2 M Tris-HCl pH 8.8, and 65 mL H₂O

4.3.2 Cloning and protein expression

4.3.2.1 General methods for molecular biology and protein chemistry

Water and buffers used for molecular biology were autoclaved, if not otherwise stated. ClearPAGE Precast Gels (12% C. B. S. *SCIENTIFIC*) were used for sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) analysis of proteins and Western blotting.

Instruments used for molecular biology and protein chemistry

Avanti J-26 XP Centrifuge (*BECKMAN COULTER*)

4.3 Methods and experimental procedures

Allegra® X-15 R Centrifuge (BECKMAN COULTER)

Fresco™ Microcentrifuge (Thermo SCIENTIFIC)

Combispin FVL-2400 mit Vortex (peqlab Biotechnologie GmbH)

Primus 25 advanced® thermocycler (peqlab Biotechnologie GmbH)

Thriller Thermo-Inkubationsmischer (peqlab Biotechnologie GmbH)

NanoDrop ND-2000 (Thermo SCIENTIFIC) (equipped with software)

BioPhotometer Plus (eppendorf)

PerfectBlue Doppelgelsystem Twin M (peqlab Biotechnologie GmbH)

PerfectBlue Gel System Mini (peqlab Biotechnologie GmbH)

Vertikales Electrophoreses System CTV400 (VWR)

C. B. S. SCIENTIFIC DCX-700 system (C. B. S. SCIENTIFIC)

Electrophoresis power supply (EV222) (peqlab Biotechnologie GmbH)

ChemiSmart 5000 digital imaging system (Vilber-Lourmat) (equipped with software)

Mini Rocker MR-1 (peqlab Biotechnologie GmbH)

peqPETTE 5er Set (peqlab Biotechnologie GmbH)

E.coli cell culture

For cloning purposes, *E.coli* cell culture was carried out in different solid and liquid media, the compositions of which are summarized in *Table 4.3-4*. Millipore water was used, and the final pH was adjusted to 7.0. Prior to the addition of ampicillin, the autoclaved medium was first cooled to about 50 °C, then a sterile ampicillin stock solution 1000 X (ampicillin sodium salt *AppliChem*) was added to a final concentration of 100 µg/mL.

4.3 Methods and experimental procedures

Table 4.3-4 Preparation of cell culture medium (for 1 liter)

Medium	Compositions
LB medium ^a	10 g tryptone, 5 g yeast extract, 10 g NaCl
LB plate	10 g tryptone, 5 g yeast extract, 10 g NaCl, 15 g agar
SOC medium ^b	20 g tryptone, 5 g yeast extract, 0.5 g NaCl, 0.18 g KCl, 0.952 g MgCl ₂ , 2.4 g MgSO ₄
SB medium ^c	30 g tryptone, 20 g yeast extract, 10 g MOPS ^d , 20 g glucose

a) LB = Lysogeny broth, b) SOC = super optimal broth, c) SB = super broth, MOPS = 3-(N-morpholino)propanesulfonic acid, tryptone, yeast extract, agar, and all chemicals purchased from *Carl Roth*.

In general, amplification of plasmid DNA for analysis purpose was performed in 3 mL overnight culture with antibiotic pressure. The large scale cell culture for mediprep. of plasmid DNA was carried out in 100-200 mL culture medium with ampicillin (volume of cell culture depends on the plasmid is low-copy or high-copy).

Preparation of electrocompetent cells

E.coli K12 ER2738 (*New England Biolabs*) was streaked onto LB agar plates and incubated at 37 °C overnight. A single colony was used to an overnight starter culture in LB medium, and cells were incubated overnight. The overnight culture was used to inoculate 500 mL of media (1000 fold dilution), and this culture was incubated with shaking (200 rpm) at 37 °C until an optical density (OD) at 595 nm of 0.5 was reached. Cells were immediately centrifuged at 3,000 × g at 4 °C for 10 minutes and the cell pellet was resuspended in 200 mL ice-cold water. Cells were again pelleted, then resuspended in 10 mL ice-cold 10% glycerol. Finally, cells were pelleted and resuspended in 1 mL ice-cold 10% glycerol and diluted with ice-cold 10% glycerol until OD = 1. Cells were divided into 60 μL aliquots, snap-frozen in liquid nitrogen, and stored until use at -80 °C. Transformation efficiencies were determined with a control plasmid, and by means of following equation.

$$\text{competence} = \frac{\text{number of clones} \times \text{dilution factor}}{\text{spread volume } (\mu\text{L})} \times \frac{1 \times 10^4 \mu\text{L}}{10 \text{ pg}} \times \frac{1 \times 10^6 \text{ pg}}{1 \mu\text{L}}$$

4.3 Methods and experimental procedures

Transformation of plasmid DNA

Transformation by heat shock

Plasmid DNA (1 μ L, about 50 ng) was added to 50 μ L NEB 5-alpha competent cells (subcloning efficiency, *New England Biolabs*) in a 1.5 mL tube on ice. The suspension was well mixed and kept on ice for 30 minutes. Heat shock was applied by transferring the cells to a water bath at 42 °C for 45 seconds, then transferring them back to the ice and incubating for a further 2 minutes. Prewarmed antibiotic-free LB or SOC medium (1 mL) was added and the cells allowed to recover by incubating with shaking (200 rpm) at 37 °C for 1 hour, after which 100 μ L of the transformation mixture was spread onto LB plates containing ampicillin.

Transformation by electroporation

Plasmid DNA (1 μ L, about 50 ng) was added to 50 μ L electrocompetent cells (see above for protocol) in an electroporation cuvette on ice (1 mm electrode gap, *peqlab*). Electroporation was carried out at 2500 V. 1 mL SOC medium was added immediately into the cuvette, and the cell suspension was allowed to recover by incubation at 37 °C with shaking at 200 rpm for 1 hour. Afterward, 100 μ L of the transformation mixture was spread onto LB plates containing ampicillin.

DNA isolation and purification

In general, plasmid DNA and cloning intermediates were isolated and purified by commercially available DNA isolation and purification kits, which utilize a silica-based membrane technology (based on ion-exchange). DNA from lysed cells or from nuclease digestions bind to the membrane. Impurities can be removed by washing steps, and the DNA is eluted by water or appropriate buffer solutions.

Plasmid miniprep.

For analytical purposes, plasmid DNA from *E.coli* cells was isolated by means of the peqGold Plasmid Miniprep Kit I (*PEQLAB Biotechnologie GmbH*). 3 mL of an *E. coli* overnight culture was centrifuged at 10,000 \times g (room temperature) for one minute.

4.3 Methods and experimental procedures

The cell pellet was resuspended in 250 μ L Solution I/RNase A (resuspension buffer) and vortexed. After the addition of 250 μ L Solution II (lysis buffer) the tube was gently inverted 8 times, and then incubated at room temperature for 2 minutes. 350 μ L Solution III (neutralization buffer) was added and the reaction mixed gently by inverting the tube 8 times. Afterward, the mixture was centrifuged at 10,000 \times g (room temperature) for 10 minutes. The supernatant was loaded onto a PerfectBind DNA column centrifuged at 10,000 \times g at 4 °C for one minute. The column was washed twice with 750 μ L DNA wash buffer I (centrifugation at 10,000 \times g at 4 °C for one minute). In order to obtain higher purity, an optional washing step with wash buffer II can be performed. The PerfectBind DNA column was dried by centrifugation at 10,000 \times g at 4 °C for two minutes. DNA was eluted with 50 μ L water (centrifugation at 5,000 \times g at 4 °C for one minute).

Plasmid midiprep.

Plasmid DNA from *E. coli* cells was isolated by use of the Pure[®] HiPure Plasmid Midiprep Kit (*Invitrogen*). 100 mL overnight *E. coli* culture (high-copy plasmid) or 200 mL overnight *E. coli* culture (low-copy plasmid) was centrifuged at room temperature at 4,000 rpm for 10 minutes. The cell pellet was resuspended in 4 mL resuspension buffer and then treated with 4 mL lysis buffer and mixed gently. The mixture was incubated at room temperature for 4-5 min. After the addition of 4 mL neutralization buffer, the mixture was centrifuged at 13,300 \times g at 4 °C for 20 minutes. In the meantime, the spin column was equilibrated with 10 mL equilibration buffer. The supernatant was loaded onto the column and allowed to flow through by gravity. The column was washed twice with wash buffer. DNA was eluted with 5 mL elution buffer. Subsequently, 3.5 mL isopropanol was added, the mixture well mixed and centrifuged at 13,300 \times g at 4 °C for 30 minutes. The DNA pellet was washed with 3 mL 70% ethanol and centrifuged at 13,300 \times g at 4 °C for 5 minutes. DNA pellet was air dried for about 5 minutes and then dissolved in water.

Purification of DNA fragments from agarose gel

4.3 Methods and experimental procedures

DNA fragments were isolated after agarose gel electrophoresis by excising the band of interest with a scalpel under UV light; the gel slice was transferred into a 2 mL tube. Binding Buffer (1 mL for 1 g gel) was added and the mixture incubated with shaking (200 rpm) at 65 °C until the gel was completely dissolved. The solution was acidified by the addition of 3 M sodium acetate (5-10 µL, depends on the volume of gel solution). The solution was loaded onto a PerfectBind DNA column (*PEQLAB Biotechnologie GmbH*) and centrifuged at 10,000 x g (room temperature) for 1 minute. The column was washed twice with 750 µL CG Wash Buffer (centrifugation at 10,000 x g at 4 °C for one minute). An optional washing step with 750 µL Binding Buffer can be performed prior to use of the CG Wash Buffer to increase DNA purity. The column was dried by centrifugation at 10,000 x g at 4 °C for two minutes, and the DNA fragment was eluted with water (centrifugation at 5,000 x g at 4°C).

Determination of DNA concentration:

The concentration of isolated plasmid DNA and DNA fragments was determined by means of a NanoDrop ND-2000 (260 nm), and the purity confirmed by the ratio of absorbance at 260 nm versus 280 nm (in general, the ration of 260/280 is >1.8).

4.3.2.2 Cloning of plasmid DNA for protein expression

Cloning of constructs containing SUMO-BPTI

The DNA sequence encoding SUMO-BPTI_{amber15} containing a 5'-terminal XhoI recognition site and 3'-terminal BamHI recognition site was synthesized by *eurofins mwg | operon* and cloned into the pCR® 2.1 (3.9 kb) vector. This DNA sequence was also directly cloned into the pIVEX vector after digestion with XhoI/BamHI. In order to insert this DNA sequence into the vectors pBH4 and pET21cHx, the BamHI and XhoI sites were swapped by PCR with mutagenic primers. For cloning SUMO-BPTI without an amber stop codon, but rather with the wild-type Lys15, as a control, mutagenesis using PCR with mutagenic primers was also carried out.

4.3 Methods and experimental procedures

Cloning of constructs containing BPTI

The DNA sequence encoding SUMO-BPTI was obtained by double digestion with XhoI/BamHI and extraction after agarose gel electrophoresis. Then, the mutagenesis PCR with designed primers was performed to remove the N-terminal SUMO fusion tag and introduce the appropriate endonuclease recognition sites, either BamHI or XhoI, depending on the vector, as described above. The DNA sequence of BPTI was then cloned into the respective vectors and amplified in *E.coli* cells.

Cloning of constructs containing chymotrypsinogen B1 and trypsin 1

cDNA encoding *Homo sapiens* chymotrypsinogen B1 (CTRB1, transfection-ready DNA, pCMV6-XL5 vector; for protein sequence, see *Appendix*) and cDNA encoding *Homo sapiens* trypsin 1 (transfection-ready DNA, pCMV6-XL4 vector; for protein sequence, see *Appendix*) were purchased from ORIGENE. The codon corresponding to Gly216 in CTRB1 was replaced with the *amber* stop codon through mutagenesis PCR. The convenient cleavage sites of restriction enzymes were introduced at the N-terminus or C-terminus of DNA sequences of CTRB1. For cloning of trypsin 1, the convenient cleavage sites of restriction enzymes were also introduced into cloning site. Then the DNA sequence was cloned into expression vectors.

Site-directed mutagenesis by means of PCR

PCR is a standard and powerful biochemical method used to amplify DNA. The principle of the PCR reaction is based on the accurate semi-replication of DNA, which is catalyzed by DNA polymerase. A basic PCR set up requires a DNA template, a forward primer, a reverse primer, dNTPs, buffer, and DNA polymerase. Site-directed mutagenesis can be achieved by designing synthetic primer that are complementary to the template DNA in the vicinity of the mutation site. In this study, *Pfu* DNA polymerase (*Fermentas*) and *Phusion*, high-fidelity DNA polymerase (*New England Biolabs*) were used for the cloning of expression constructs. *Table 4.3-5* represents a typical mutagenesis PCR reaction. The primers designed for the cloning of these constructs are summarized in *Table. 4.3-6*.

4.3 Methods and experimental procedures

Table 4.3-5 Preparation of site-directed mutagenesis PCR

Components	Volume (μL)	Cycle program	
		Temperature	Time
10 x <i>Pfu</i> buffer (+Mg ²⁺)	5	95 °C	2 minutes
dNTPs (2.5 mM)	4		
DNA template ^a	1	95 °C	0.5 minute
Forward primer ^b	0.5	65 °C	1 minute
Reverse primer ^b	0.5	72 °C	5 minutes
<i>Pfu</i> polymerase	1		
H ₂ O		72 °C	10 minutes
		4 °C	End

a) A typical amount of template DNA is 1 μg, b) the final concentration of primer is 0.1-1 μM.

Table 4.3-6 PCR primers used for cloning

Short Name	Sequence	Description
Chy-BamHI	5'-GGC GGA TCC ATG GCT TTC CTC TGG CT-3'	Introduction of BamHI site for chymotrypsinogen B1 (pBH4/pET21cHx)
Chy-Mut	5'-GGC ATT GTG TCC TGG TAG AGC GAC ACC T-3'	Introduction of <i>amber</i> codon for chymotrypsinogen B1
Chym - XhoI	5'-GGC CTC GAG GTT GGC AGC CAG GAT CTT-3'	Introduction of XhoI site for chymotrypsinogen B1 (pET21/pBH4)
Chy-pIVEX-Xoh ^a	5'-C CTC GAG CAT CAT CAT CAT CAT ATG GCT TTC CTC- 3'	Introduction of XhoI site and 6xhis-tag for chymotrypsinogen B1 (pIVEX)
Chy-pIVEX-BamH ^a	5'-GGC GGA TCC GTT GGC AGC CAG GA-3'	Introduction of BamHI site for chymotrypsinogen B1 (pIVEX)
Chy-pIVEX-Xoh-2	5'-C CTC GAG C CAT CAT CAT CAT CAT ATG G-3'	Introduction of His-tag for chymotrypsinogen B1 (pIVEX)
Chy-mut2	5'-TGG AGC AGG TGT CGC TCT ACC AGG AC-3'	Introduction of <i>amber</i> codon for chymotrypsinogen B1
CpX-2	5'-CCT CGA GCC ATC ATC ATC ATC ATA TG-3'	Correction of shift-reading in chymotrypsinogen
SUBP-pET-BamH ^a	5'-GGC GGA TCC ATG TCT GAC CAG GAG GCA-3'	Introduction of BamHI site for SUMO-BPTI (pET21/pBH4)
SUBP-pET-Xoh-1 ^a	5'-GCC GAG CTC TCA AGC ACC ACC ACA GGT - 3'	Introduction of XhoI site for SUMO-BPTI (pET21/pBH4)
SUBP-pET-Xoh-2 ^a	5'-GCC GAG CTC AGC ACC ACC ACA GGT CCT-3'	Introduction of XhoI site (pET21/pBH4)
BPTI for.	5'-ATG TCT GAC CAG GAG GCA AAA CCT TCA-3'	Removal of SUMO
BPTI rev.	5'-AGC ACC ACC ACA GGT CCT CAT GCA GTC - 3'	Removal of SUMO
BPTI-lys	5'-GGT CCC TGC AAG GCC AGA ATT ATC AG-3'	Removal of the <i>amber</i> codon
BPTI-WS-Xoh	5'-CTC GAG CCA TCA TCA TCA TCA TCG GCC TGA C - 3'	Removal of SUMO and introduction of XhoI site (pIVEX)
BPTI-WS-Bam	5'-GGC GGA TCC CGG CCT GAC TTC TGC CTA-3'	Removal of SUMO and introduction of BamHI site (pBH4/pET21)
Try-for	5'-GGCGGATCCATGAATCCACTCCTG-3'	Introduction of BamHI site for trypsin (pBH4/pET21cHx)

4.3 Methods and experimental procedures

Try-rev

5'-GGCCTCGAGTTAGCTATTGGCAGCTAT-3

Introduction of XhoI site for
trypsin (pBH4/pET21cHx)

a) These primers were synthesized and supplied by *Biomers.net*; other primers were synthesized by *Eurofins MEG Operon*.

Restriction digestion and DNA ligation

The restriction enzymes BamHI and XhoI were used for the cloning of constructs for protein expression (for cleavage sites of BamHI and XhoI see *Appendix*). DpnI was used for the digestion of cellular methylated DNA that was used as a PCR template. T4 DNA ligase can recognize the sticky ends of DNA fragments and ligates these by catalyzing the formation of a phosphodiester bond. The preparation of enzymatic digestion and T4 DNA ligation reactions are represented in *Table 4.3-7*. BamHI, XhoI, T4 DNA ligase and DpnI were purchased from *New England Bioabs*.

Table 4.3-7 Preparation of restriction digests and DNA ligation

Restriction enzyme digestion ^a		DNA ligation ^a	
Composition	Volume (μL)	Composition	Volume (μL)
10 x NEBbuffer 3	2	T4 DNA ligase buffer	2
Plasmid DNA	1	Plasmid vector	1
BamHI	1	DNA insert	1
XhoI	1	T4 DNA ligase	1
BSA (10 mg/mL)	0.2	H ₂ O	15
H ₂ O	14.8		
Total	20		20
at 37 °C for 1 hour		RT overnight	

a) Typical amount of DNA for enzymatic digestion is 1-5 μg in 20 μL of total reaction volume, for preparative purpose the reaction volume is increased to 50 μL. b) The overall concentration of vector and insert is 1-10 ng/μL, the molar ratio of insert to vector is 2~6 for optimal ligation.

Agarose gel electrophoresis

Agarose gel electrophoresis is a technique to separate and analyze nucleotide and protein samples. The negatively charged nucleotide samples can be separated

4.3 Methods and experimental procedures

according to their size as they migrate through the agarose matrix. The gel was run in Tris-acetate-EDTA (TAE) buffer. For preparing 50 × TAE buffer, 242.2 g Tris was dissolved in water, and 57.1 mL of glacial acetic acid and 100 mL of 0.5 M EDTA were added; water was added to bring the final volume to 1 liter (pH adjusted to 8.0). 2-Log DNA Ladder (0.1-10.0kb, *New England Biolabs*) was used as a marker.

DNA-sequencing

To identify positive clones, those containing accurate constructs, plasmid DNA isolated from a mini-prep kit was analyzed by means of DNA-sequencing. DNA-sequencing was carried out with primers complimentary to the T7 promotor and T7 terminator (*Seqlab Sequence Laboratories, Göttingen, Germany*).

4.3.2.3 Protein expression and *in vitro* amber suppression

Cell-free protein expression by means of commercial kits was carried out according to the manual with small modifications. (internet link to kit manual is given in *Appendix*). Expression reactions for *in vitro* amber suppression were prepared as shown in *Tables 4.3-8* and *4.3-9*.

Table 4.3-8 Cell-free reaction preparation by use of S30 T7 High-Yield Expression System

Composition	Positive Control ^a	Samples ^b
S 30 Premix Plus (μL)	20	20
T7 S30 extract (μL)	18	18
Plasmid DNA (μL) ^c	2	1-2
Suppressor aa-tRNA ^{Phe} _{CUA} ^d		X
H ₂ O	10	
Total (μL)	50	50

a) Control DNA of luciferase is supplied in kit, b) plasmid constructs containing DNA sequence of SUMO-BPTI, BPTI and chymotrypsinogen B1 were tested for expression, c) the concentration of plasmid DNA is general 0.8-1.5 μg/μL, d) 10-20 μg of deprotected aa-tRNA added into the reaction. The batch reaction was incubated at 37 °C at 300 rpm for 1 hour.

4.3 Methods and experimental procedures

Table 4.3-9 Cell-free reaction preparation by use of RTS 100 *E.coli* Disulfide Kit

Reaction mixture	Positive control ^a	Samples ^b
Activated Lysate (μL)	25	25
Reaction mixture (μL)	7	7
Amino acids (-Met) (μL)	7	7
Met (μL)	1	1
DNA (μL) ^c		1-2
Suppressor aa-tRNA ^d		X
H ₂ O (μL)		
Total (μL)		50
Feeding mixture		
Feeding mixture (μL)		640
Amino acids (-Met) (μL)		140
Met (μL)		20
H ₂ O (μL)		200
Total (μL)		1000

a) Control DNA of urease is supplied in kit, b) plasmid constructs containing DNA sequence of SUMO-BPTI, BPTI, chymotrypsinogen B1, and trypsin were tested for expression, c) the concentration of plasmid DNA is generally 0.8-1.5 μg/μL, d) 10-20 μg of deprotected aa-tRNA added into the reaction. The continuous-exchange reaction was incubated at 30 °C at 600 rpm from 2 hours to overnight.

The cell-free reactions for *in vitro amber* suppression by reconstitution of the reaction mixtures were prepared as shown in Tables 4.3-10 and 4.3-11.

4.3 Methods and experimental procedures

Table 4.3-10 Preparation of cell-free premixes according to literatures

Composition	Hecht <i>et al.</i>			Sisido <i>et al.</i>			Schultz <i>et al.</i>		
	L. con.	V (μl)	F. con.	L. con.	V (μl)	F. con.	L. con.	V (μl)	F. con.
Tris-acetate (pH 7.0) (1 M)	35	70	70						
HEPES-KOH (pH 7.4) (1 M)				55	220	110			
Tris-acetate (pH 7.4) (1 M)							56.4	112.8	112.8
Potassium glutamate (pH 7.0) (1M)	190	380	380	210	420	380	72	144	144
Ammonium acetate (pH 6.6) (1 M)	30	60	60	6.9	13.8	13.8	36	72	72
Magnesium acetate (pH 7.0) (1 M)	11	22	22	7.9	15.4	15.4	12	24	24
DTT (1.15 M)	2	3.2	4	1.7	2.72	3.4	1.76	2.8	3.5
Phosphoenol pyruvate (313.76 mM)	20	127.5	40	26	165.73	52	27	172	54
PEG 8000 (400 mg/mL)				1.90%	95	3.8%	19	95	38
Folinic acid (7.5 mg/mL) ^a				35	9.4	70	35	9.3	70
Spermidine (0.1377 M)				1	14.5	2			
Pyridoxine HCl (3.6 mg/mL) ^a							27	0.67	54
<i>p</i> -amino benzoic acid (1 mg/mL) ^a							11	22	22
Calcium acetate (1 M)							9.7	19.4	19.4
FAD (14.5 mg/mL) ^a							27	3.7	54
NADP (1.2 mg/mL) ^a							27.1	2.25	54
H ₂ O (μL)		337.3			43.45			320.08	
Total (μL)		1000			1000			1000	

L. con. = Literature concentration, F. con. = Final concentration, **a**) the concentration of folinic acid, pyridoxine HCl, *p*-amino benzoic acid, FAD, and NADP is in μg/mL; for other compounds units are in mM. L-glutamic acid (potassium salt monohydrate) and *p*-amino benzoic acid (*Sigma-Aldrich*), phosphoenol pyruvate (mono potassium salt) and FAD (disodium dehydrate) (*BioChemica*), pyridoxine hydrochloride (99%, *Alfa Aesar*), folinic acid (calcium salt hydrate, *Fluka*), spermidine (VWR), DTT, and PEG 8000 (*New England Biolabs*), NADP sodium salt (*AppliChem*)

4.3 Methods and experimental procedures

Table 4.3-11 Reconstitution of cell-free reaction with prepared premixes and commercial cell-free lysate according to literature with minor modifications

	<i>Control</i>	<i>Hecht et al.</i>		<i>Sisido et al.</i>		<i>Schultz et al.</i>	
		L. con./F.con	Volume (μL)	L. con./F.con	Volume (μL)	L. con./F.con	Volume (μL)
Premix ^a	20	50		40		40	
Template (μL) ^b	2	3		3		3	
S30 extract (μL) ^c	18	19		19		19	
ATP (20 mM) ^d		20/7.7	7.5	1.2/1.94	1.5	1.22/1.94	1.5
GTP (5 mM) ^d		5/2.1	2	0.28/0.97	0.75	0.85/0.97	0.75
UTP (5 mM) ^d		5/2.1	2		0.75	0.85/0.97	0.75
CTP (5 mM) ^d		5/2.1	2		0.75	0.85/0.97	0.75
cAMP		N. A.					
Mg (OAc) ₂ (1 M)					0.5		
Isopropyl-β-D-thiogalactopyranoside		N. A.					
RNA Polymerase ^d		1200 Units	2	1200 Units	2	1200 Units	2
Amino acids mixtures (1 mM) ^e							
<i>E. coli</i> tRNAs (100 mg/mL) ^f		0.8 mg/mL	1	0.16 mg/mL	1	0.17 mg/mL	1
H ₂ O	10						
Total (μL)	50		96.5		77.25		77.25

L. con = literature concentration (mM for dNTPs), F. con = final concentration used for our expression (mM for dNTPs), **a**) for control reaction, premix (*Promega*) was used, for other reactions the premix prepared as in *Table 3.14* were used, **b**) both DNA and RNA templates were used for test, in case of RNA, 10 μg of purified RNA produced by *in vitro* transcription was added, **c**) cell lysate from S30 T7 High-Yield Expression System (*Promega*), dNTPs and RNA polymerase from AmpliScribe™ T7 High Yield Transcription Kits (*EPICENTRE® Biotechnologies*), **e**) complete amino acid mixture (*Promega*), **f**) complete *E. coli* tRNAs from MRE 600 (*Roche*), N. A. = not added.

SUMO cleavage

SUMO protease was expressed and purified by coworker Jan-Stefan Völler and supplied at a concentration of 2 mg/mL in buffer containing 50 mM Tris and 150 mM NaCl. The cleavage test was carried out in buffer containing 50 mM tris, 150 mM NaCl, 1 mM DTT (pH 8.0) at 30 °C for 4 hours.

4.3 Methods and experimental procedures

4.3.2.4 Protein purification and analysis

Protein purification

Proteins expressed in cell-free reactions were purified by means of a Ni-NTA Spin Kit (*Qiagen*) under native conditions. The Ni-NTA spin column was equilibrated with 600 μ L NPI-10 buffer (centrifugation at 2900 rpm at 4 °C for 2 minutes). The cell-free reaction mixture (50 μ L) was diluted with 560 μ L NPI-20 buffer and loaded onto the spin column (centrifugation at 1600 rpm for 5 minutes; if the mixture flowed slowly through the membrane, the centrifugation time was extended). The column was washed twice with 600 μ L NPI-20 buffer. Protein containing the 6 X His-tag was eluted by adding 100-150 μ L NPI-500 buffer and centrifuged at 2900 rpm twice.

Table 4.3-12 Preparation of NPI buffer for purification of 6 X His-tagged protein by use of Ni-NTA spin column

NPI buffer	Description	Composition
NPI-10	Lysis buffer	50 mM NaHPO ₄ , 300 mM NaCl, 10 mM imidazole pH 8.0
NPI-20	Wash buffer	50 mM NaHPO ₄ , 300 mM NaCl, 20 mM imidazole pH 8.0
NPI-500	Elution buffer	50 mM NaHPO ₄ , 300 mM NaCl, 500 mM imidazole pH 8.0

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE is a widely used technique in biochemistry and molecular biology for the analysis of protein samples. SDS is an anionic detergent that denatures and linearizes proteins, when heated to 95 °C in the presence of DTT. Denatured samples are loaded onto the prepared gel, and an electric field is applied. The proteins, surrounded by negatively charged SDS molecules, migrate through the gel towards the anode at a rate that inversely relates to their size. Coomassie brilliant blue is commonly used for protein visualization after electrophoresis. In our studies, 12% SDS PAGE was used for protein analysis. Typically, electrophoresis was initially conducted at 80 volts until the sample fully entered the separation gel. Afterward, gels were run at 120 volts for a further 60-90 minutes. The gels for SDS-PAGE and Western blot were prepared as indicated in *Table 4.3-13*. The preparation of running

4.3 Methods and experimental procedures

buffer and buffers for staining are summarized in Table 4.3-14. A broad range (2-212 kDa) protein marker (*New England Biolabs*) was used.

Table 4.3-13 SDS-PAGE for protein analysis and Western blot (for two gels)

Components	Stacking gel Volume	Separation gel Volume
Tris buffer	1.25 mL	2.5 mL
Gel A	0.5 mL	4 mL
Gel B	0.2 mL	1.6 mL
10% APS	25 μ L	50 μ L
10% SDS	50 μ L	100 μ L
TEMED	10 μ L	15 μ L
H ₂ O	2.965 mL	1.735 mL
Total	5 mL	10 mL

Rotiphorese® Gel A (30% acrylamide solution), Gel B (2% bisacrylamide solution), tetramethylethylenediamine (TEMED), ammonium persulfate (APS) purchased from *Carl Roth*.

Table 4.3-14 Buffer and solution composition for SDS-PAGE and Western blot

Buffer/solution	Composition
SDS-PAGE running buffer (1 x)	0.025 M Tris, 0.192 M glycine, 0.1% (w/v) SDS, pH 8.3
Coomassie stock solution	12 g coomassie brilliant blue, 300 mL MeOH, 60 mL CH ₃ COOH
Coomassie working solution (1 liter)	30 mL coomassie stock solution, 500 mL MeOH, 100 mL CH ₃ COOH, 370 mL H ₂ O
SDS-PAGE fix solution (1 liter)	500 mL MeOH, 100 mL CH ₃ COOH, 400 mL H ₂ O
Destaining solution (1 liter)	450 mL MeOH, 100 mL CH ₃ COOH, 450 mL H ₂ O
Tank blotting buffer (1 x)	0.025 M Tris, 0.192 M glycine, 0.1% (w/v) SDS, 20% (v/v) methanol
Ponceau S staining solution (500 mL)	0.5 g ponceau S, 25 mL CH ₃ COOH, 475 mL H ₂ O
TBS buffer (1 liter)	1.21 g Tris, 8.78 g NaCl, pH 7
TBS-Tween buffer (1 liter)	2.42 g Tris, 29.25 g NaCl, 0.05% (v/v) Tween 20, pH 7

Coomassie brilliant blue and chemicals purchased from *Carl Roth*

Western Blot

The Western blot is a widely used analytical technique to detect specific proteins in a sample. In our studies, Western blotting was used to detect protein samples containing a 6 × His-tag produced by cell-free protein synthesis. Proteins present in

4.3 Methods and experimental procedures

the cell-free reaction mixture were first separated according to their size by means of SDS-PAGE. The resolved proteins were subsequently transferred onto a nitrocellulose membrane (0.2 μm , binding capacity 80-90 $\mu\text{g}/\text{cm}^2$, *Carl Roth*) by tank blotting. The membrane was pre-activated in blotting buffer for > 20 minutes. *Figure 4.3-5* represents the assembly of components for tank blotting. Typically, tank blotting was performed by means of a mini-vertical lab gel/blotting system DCX-700 (*C. B. S. SCIENTIFIC*) equipped with two freezer blocks at 180 voltages for two hours.

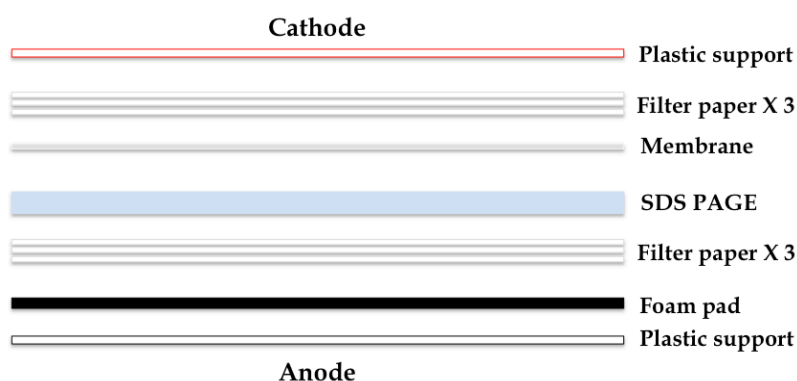


Figure 4.3-5 Schematic illustration of assembly for tank blotting

After transfer to the membrane, it was incubated in ponceau S solution with gentle shaking for 2 min. 6 X His protein ladder (*Qiagen*) and transferred proteins were visible in cases of successful transfer. The membrane was destained in water for ten minutes; an additional destaining step for another ten minutes in fresh water was carried out if necessary. Then the membrane was washed twice with TBS buffer for ten minutes each and incubated in Anti-His HRP Conjugate blocking buffer (Penta-His HRP Conjugate Kit, *Qiagen*) at room temperature for one hour. Afterward, the membrane was washed twice for ten minutes each with TBS-Tween 20 buffer, and once with TBS buffer. 5 μL anti-His HRP conjugate antibody (*Qiagen*) was diluted in 20 mL blocking buffer, and the membrane was incubated in this solution at room temperature overnight. The membrane was washed twice for ten minutes each with TBS-Tween 20 buffer, and once with TBS buffer. Chemiluminescence substrate for

4.3 Methods and experimental procedures

blotting with peroxidase (*Interchim*) was added and the membrane analyzed with a luminescence detector.

Protein fingerprinting analysis

Protein fingerprinting analysis was carried to identify whether the expressed and purified proteins contained non-canonical amino acids. Protein samples were prepared as follows. Sample buffer contained 5 mM DTT. In order to reduce all disulfide bonds, mixtures of protein and sample buffer were incubated at 55 °C for 30 minutes (common protocol use 98 °C for 5-10 minutes). Iodoacetamide (*AppliChem*) was added to a final concentration of 15 mM and the reaction incubated in the dark for 20 minutes. Subsequently SDS-PAGE was performed to separate the proteins. The interesting protein bands were cut out and transferred to a 1.5 mL tube. Further experimental preparation including in-gel trypsin digestion, nano-LC-MS, and identification and characterization of the protein fragments was carried out by Heike Stephanowitz (Mass spectrometry group, Leibniz-Institut für molekulare Pharmakologie, Berlin).

Chapter 5

Total chemical synthesis and full characterization of BPTI containing fluorinated amino acids

5.1 Total chemical synthesis of proteins

As discussed in *Chapter 4*, great efforts were made to site-specifically incorporate non-canonical amino acids into proteins to determine potential novel structural and functional properties. Still, these techniques based on *in vitro* and *in vivo* approaches show limitations, which need to be overcome in the future. Other methods, which can be used for the site-specific incorporation of non-canonical amino acids into proteins include the total chemical synthesis of proteins, and semi-synthesis by means of diverse chemoselective ligation strategies. In this chapter an overview of the total chemical synthesis of proteins is given. The total chemical synthesis of mutant BPTI variants containing non-canonical amino acids, and their biochemical characterization, are then described.

5.1 Total chemical synthesis of proteins

Synthetic chemical approaches to the synthesis of peptides and proteins have been evolving for decades.¹⁶⁹ In the 1960s, the development of orthogonal protecting groups led to the most widely used technique in peptide and protein chemistry - solid-phase peptide synthesis (SPPS).¹⁷⁰ Nowadays, linear peptide chains can be assembled reliably by means of SPPS. In combination with ligation techniques including native chemical ligation (NCL) and expressed protein ligation (EPL), the synthesis of larger more complex proteins is possible.¹⁷¹

5.1.1 Solid phase peptide synthesis

Briefly, SPPS is carried out on a functional unit (linker) associated with a resin (solid-phase) to which the carboxy terminal amino acid is attached; thus the peptide chain is elongated from C- to N-terminus. Reactive side chains and the N^α amino group of each amino acid are protected. The carboxyl group of each amino acid to be coupled is activated by treatment with activation reagents. After the attachment of the first amino on the resin, the N^α protecting group is removed. In contrast to peptide

5.1 Total chemical synthesis of proteins

synthesis in solution, the solid-phase support enables the removal of all byproducts and chemical reagents by filtration and washing steps. The repetitive synthetic cycle consists of: deprotection-washing-coupling-washing (Figure 5.1-1). The use of an insoluble (heterogeneous) support also enables full-automation of the synthesis process. Over several decades of development numerous types of resins have been made commercially available for the synthesis of different peptide acids, peptide amides, protected peptide fragments, and carboxyl-modified peptides.¹⁷²

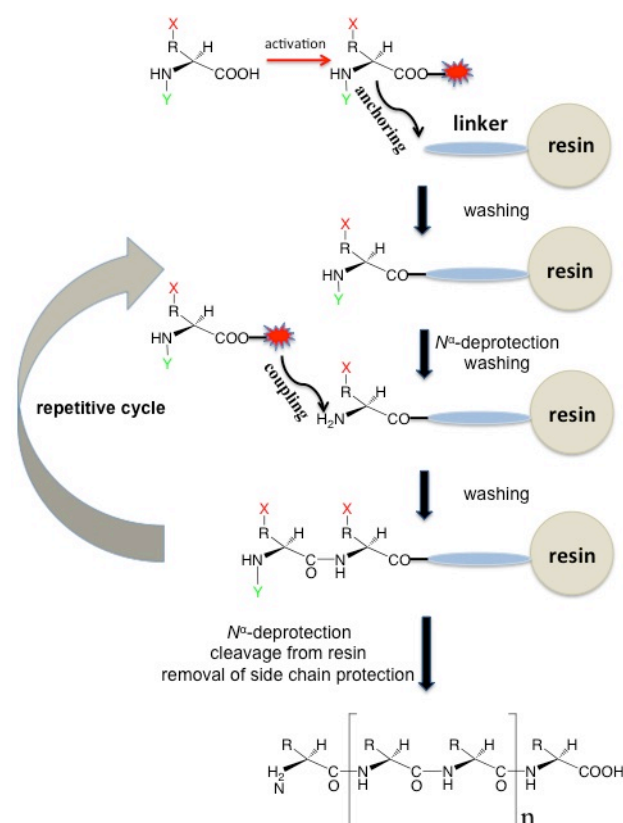


Figure 5.1-1 Schematic representation of the principles of SPPS

X presents the side chain protecting group, Y presents the N^α amino protecting group.

The most widely used protecting groups for the the N^α amino moiety are *tertiary*-butylcarbonyl (*t*Boc) and fluorenylmethyloxycarbonyl (Fmoc).¹⁷³ Typically the *t*Boc group is removed by trifluoroacetic acid (TFA) and the protonated α -ammonium species of the peptide fragment needs to be neutralized. The Fmoc protecting group is removed under basic conditions, commonly with piperidine or 1,8-diazabicycloundec-7-ene (DBU). The functional groups of the amino acid side chains

5.1 Total chemical synthesis of proteins

are usually protected by benzyl (Bzl) in the Boc strategy and *tert*-butyl (*t*Bu) in the Fmoc strategy. However, numerous other side chain protecting groups have been developed that can be removed selectively and conditionally.¹⁷⁴ Both peptide cleavage from the resin for the Boc synthesis protocol and the removal of the Bzl group require harsh conditions, namely anhydrous HF.

Another important feature of SSPS is the effectiveness of amide bond formation by coupling. Among the numerous coupling reagents, there are two major reagent classes that require either *in situ* activation of the carboxyl acid or previously prepared precursors. The most widely used coupling reagents for SPPS are cabodiimides in combination with phosphonium and uronium salts.¹⁷⁵

5.1.2 Native chemical ligation

Kent and coworkers first demonstrated chemoselective ligation at an N-terminal cysteine (Cys) containing peptide with an unprotected peptide fragment containing a thioester at the C-terminal α -carboxyl group. This strategy and EPL, which is based on a naturally occurring enzymatic protein splicing reaction, rapidly became widely used methods in protein chemistry.¹⁷⁶ A significant number of proteins have now been synthesized by means of these strategies.^{171b, 177} In aqueous buffer solution containing the thiol catalyst at neutral pH, the N-terminal Cys peptide and peptide α -thioester first form the thioester-linked intermediate through a reversible thiol exchange reaction. Subsequently, an irreversible intramolecular rearrangement forms the native peptide bond (*Figure 5.1-2*). Initially, the NCL reaction required a N-terminal Cys peptide fragment. Recently, the NCL technique has been expanded to include other N-terminal amino acids, in combination with desulfurization, and chemical modifications at the ligation site, including removable auxiliary groups and non-natural linkages.¹⁷⁶

5.1 Total chemical synthesis of proteins

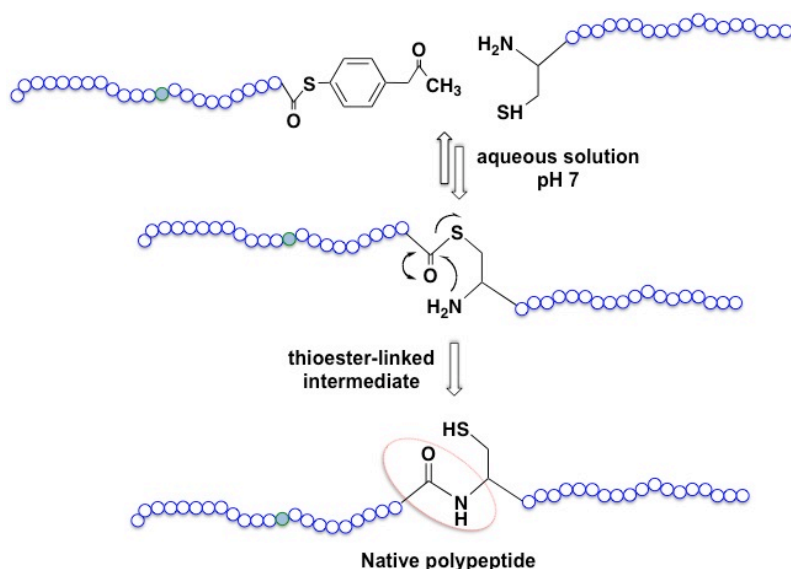


Figure 5.1-2 Schematic representation of the principles of native chemical ligation

5.1.3 Fmoc-based SPPS synthesis of peptide α -thioester

Synthesis of the peptide α -thioester is a key issue in standard NCL and the total chemical synthesis of proteins. Typically, the peptide α -thioester can be efficiently synthesized on thioester resin linkers by means of a Boc/Bzl strategy.¹⁷⁸ Compared to the Boc/Bzl based strategy, the milder Fmoc/*t*Bu method is more limited for the generation of an α -thioester. This is due to the utilization of basic condition for Fmoc-deprotection, which the thioester moiety is labile. Though direct synthesis of the peptide α -thioester by replacing piperidine with non-nucleophilic reagents for Fmoc removal is possible, the yield of product would suffer.¹⁷⁹ Widely used procedures include the thioesterification of fully protected peptide acids. The peptide fragment is synthesized on acid-sensitive resins such as 2Cl-Trt and HMPB resin. After assembly of the peptide sequence, the fully protected product can be obtained by subjecting the peptide-resin to mild acidic conditions. Afterwards, the α -thioester can be synthesized at the N-terminus of the peptide.¹⁸⁰ Another important method to synthesize the α -thioester at the N-terminus of the peptide fragment is direct conversion from the resin-bound peptide. Hilvert and coworkers have used alkyl

5.1 Total chemical synthesis of proteins

aluminum thiolates to cleave the peptide from the resin and obtain a peptide α -thioester simultaneously.¹⁸¹ Also, side chain-anchored and backbone-anchored peptide α -allyl esters as precursors can be used for the synthesis of peptide α -thioesters.¹⁸² More recently, the design of “safety-catch” linkers by means of specifically functionalized resins and synthetic precursors for generating α -thioesters offer alternative synthetic strategies. These include sulfonamide and aryl hydrazine linkers as well as α -aryl benzimidazolones and backbone pyroglutamylimides as precursors.¹⁸³

5.2 Total chemical synthesis of BPTI containing fluorinated amino acids

Mutant BPTI species containing non-canonical amino acids: Abu, DfeGly, and TfeGly at position 15 were synthesized by standard Fmoc-based SPPS and a native chemical ligation.

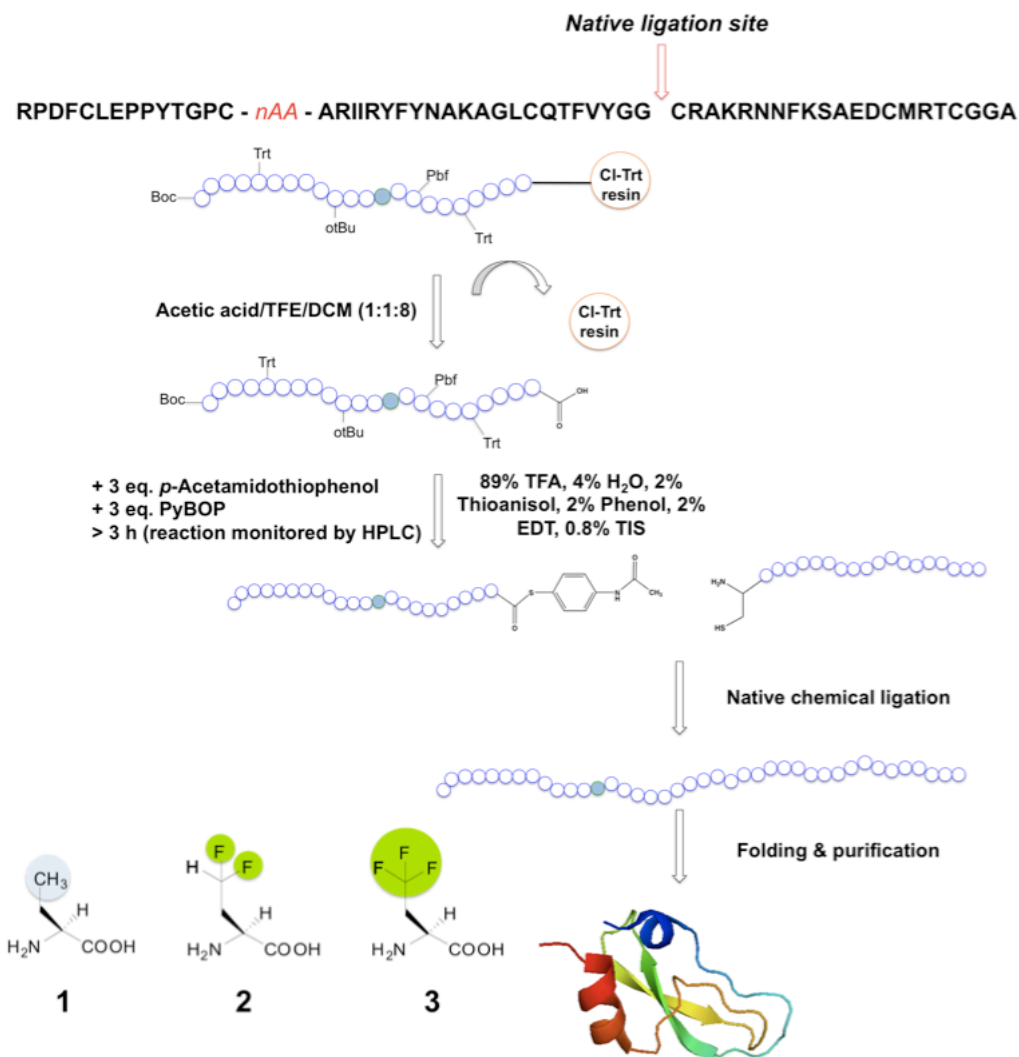


Figure 5.2-1 Schematic representation of synthetic strategy for total chemical synthesis of mutant BPTI variants containing non-canonical amino acids at position 15 and three incorporated non-canonical amino acids

1. Abu, 2. DfeGly, and 3. TfeGly

5.2 Total chemical synthesis of BPTI containing fluorinated aa

The synthetic strategy for the total chemical synthesis of BPTI containing non-canonical amino acids at position 15 is schematically represented in *Figure 5.2-1*.^{181, 184} The N-terminal fragment was synthesized on Fmoc-Gly-2Cl-Trt resin. The fully protected peptide fragment was cleaved from the resin by treatment with acetic acid/trifluoroethanol (TFE)/DCM (1/1/8, v/v/v). The α -thioester of the N-terminal fragment was synthesized through esterification by treating the fully protected peptide fragment with 5 equivalents of benzotriazol-1-yl-*N*-oxy-tris-(pyrrolidino)phosphonium hexafluorophosphate (PyBOP), *p*-acetamidothiophenol, and *N,N*-diisopropylethylamine (DIPEA).^{180a} The side chain protecting groups were removed subsequently. The C-terminal fragment of BPTI, residues 38-58, was synthesized on Fmoc-Ala-NovaSyn® TGT resin. Peptide fragments were purified by means of RP-HPLC and used directly for the native chemical ligation.

The full-length mutant BPTI species were synthesized via NCL between the N-terminal peptide fragment α -thioester and the C-terminal fragment with an N-terminal cysteine. The NCL reaction was monitored by RP-HPLC (*Figure 5.2-2*).

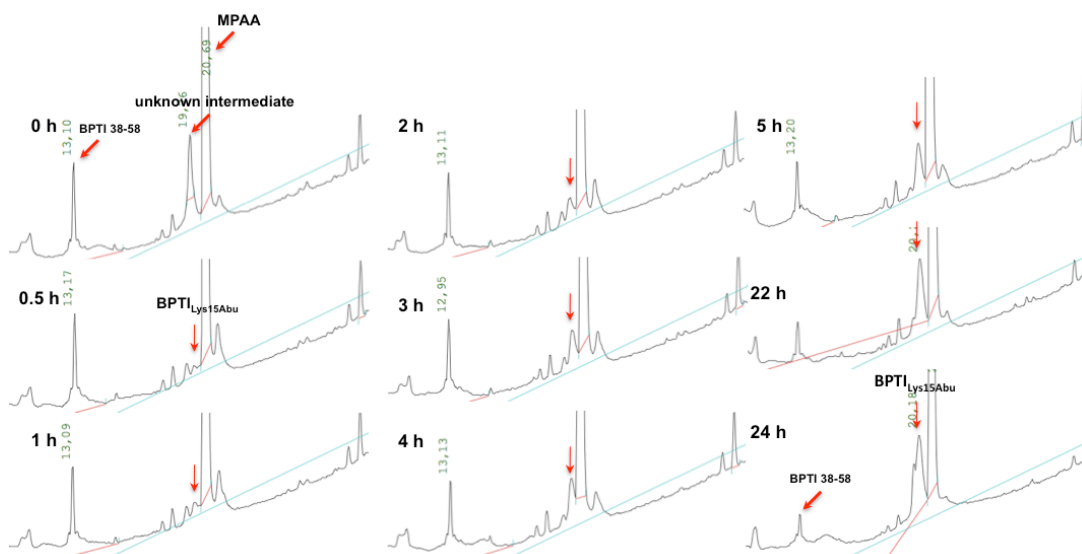


Figure 5.2-2 Monitoring the NCL reaction by RP-HPLC

After a certain incubation time, the peak intensity of the C-terminal fragment (BPTI 38-58) decreases while the peak intensity of the product (full-length BPTI) increases; these representative chromatograms were obtained from the NCL of BPTI_{Lys15Abu}.

5.2 Total chemical synthesis of BPTI containing fluorinated aa

After purification by means of RP-HPLC, mutant full-length BPTI species were allowed to fold by means of air oxidation (Figure 5.2-3). Yields from the key synthetic steps are summarized in Table 5.2-1.

Figure 5.2-3 Monitoring BPTI_{Lys15Abu} folding by RP-HPLC

These representative traces were obtained from the refolding of BPTI_{Lys15Abu}; unfolded BPTI has a longer retention time than folded BPTI, due to its greater hydrophobicity.

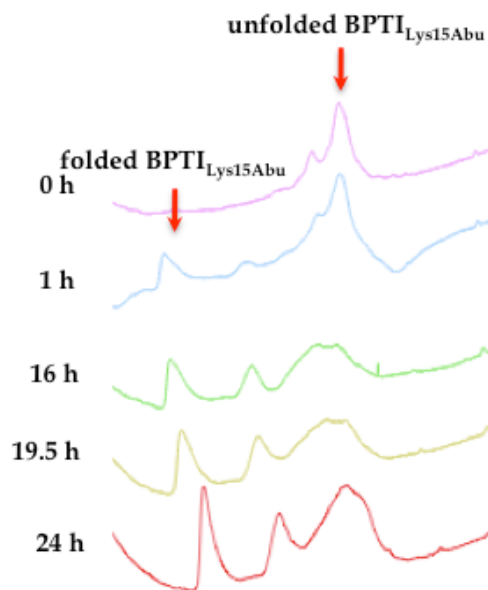


Table 5.2-1 Synthesis summary of mutant BPTI variants containing non-canonical amino acids at position 15

Resin	C-terminal fragment <i>Fmoc-Ala-NovaSyn® TGT</i>	N-terminal fragment- α -thioester <i>Fmoc-Gly-2Cl-Trt</i>	NCL	Protein folding	
Yield	18%	Lys15Abu Lys15DfeGly Lys15TfeGly	6% 6% 7%	26% 17% 20%	17% 13% 11%
Yield from literature (wild-type)					
181	n. r.		8% (Fmoc)	63%	n.r.
184	25% (Boc)		15% (Boc)	60%	45%

NCL: native chemical ligation, n. r.: not reported, Fmoc: synthesis by use of Fmoc based SPPS, Boc: synthesis by use of Boc based SPPS.

5.2 Total chemical synthesis of BPTI containing fluorinated aa

Because the BPTI sequence contains six cysteines, and such cysteine-rich peptides are generally difficult to synthesize by an Fmoc strategy, the overall yields of the three mutant BPTI species are poor. Nevertheless, the yields of the C-terminal peptide fragment and the N-terminal peptide fragment α -thioester are comparable to those reported in literature (Table 5.2-1). Due to the poor solubility of the N-terminal peptide thioester, the yield of the NCL reaction was also poor. It had been previously reported that the theoretical yield of folded BPTI is expected to be less than 50% because of the existence of a kinetic trap dead form in which a disulfide bond between 30-51 lacks.¹⁸⁵ In our studies, due to the small scale of folding reactions, sufficient yields of 11-17% were obtained, although these were about three-fold lower than the reported yield. It is not clear to what extent the substitutions at position 15 may play a role in reducing the yields from the NCL and folding reactions.

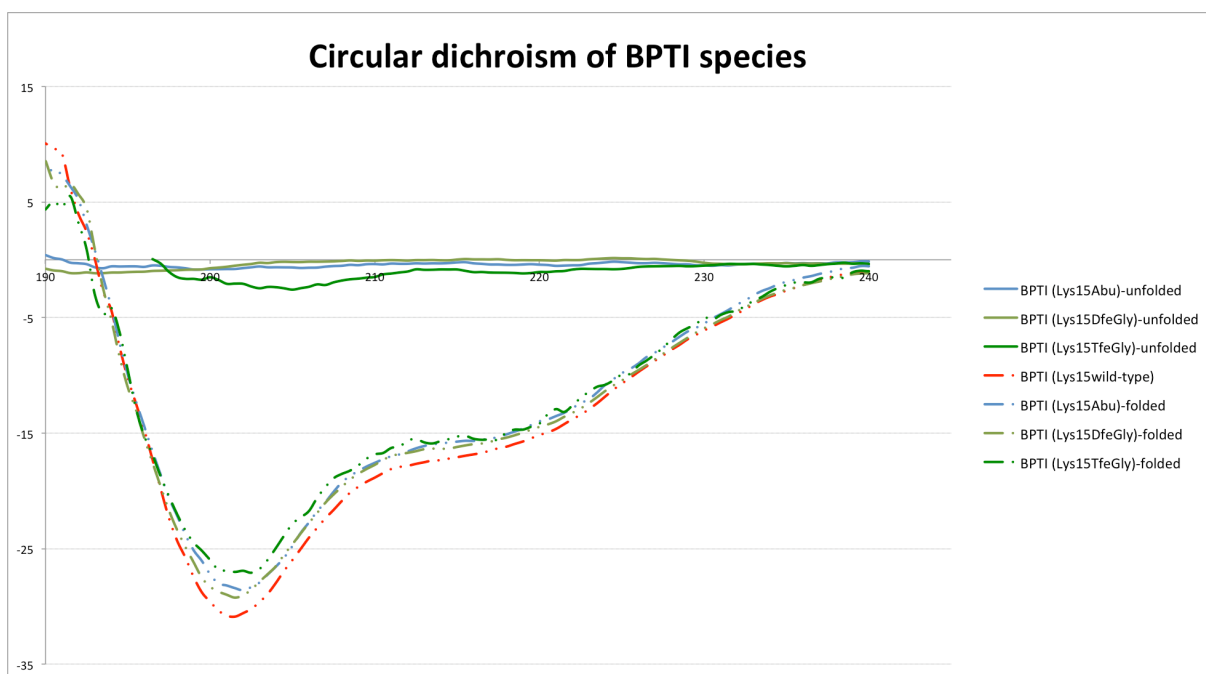


Figure 5.2-4 Circular dichroism analysis of synthetic BPTI species in Tris buffer at pH 7.4

Red line: BPTI_{Lys15wild-type} (*Sigma-Aldrich*); blue line: BPTI_{Lys15Abu}; forest green line: BPTI_{Lys15DfeGly}; green line: BPTI_{Lys15TfeGly}.

5.2 Total chemical synthesis of BPTI containing fluorinated aa

The presence of three disulfide bonds in each synthetic variant was confirmed by ESI-MS (*Appendix*). The global conformation of synthetic BPTI species was surveyed by means of circular dichroism (CD) spectroscopy. All synthetic mutant BPTI variants result in CD spectra that are similar in sharpness and intensity to the commercially available parent protein.^{65, 186} In contrast, unfolded full-length BPTI did not yield spectra indicative of specific secondary structural elements (*Figure 5.2-4*). High-resolution crystal structures of the synthetic mutant BPTI variants in complex with β -trypsin were also obtained, and also reveal a lack of global structural perturbation induced by the single amino acid substitutions (*section 5.3.3*).

5.3 Characterization of BPTI containing of fluorinated amino acids

5.3.1 Protein stability regarding substitution at P₁ position of BPTI

As was discussed in *section 1.2*, the incorporation of highly fluorinated amino acids into the hydrophobic core of coiled-coil peptide model systems increases protein stability, generally speaking.^{33e, 34b, 36b, 38d} In contrast, the effects on the stability of globular proteins as a result of fluorination at solvent-exposed positions have been seldom investigated. Position 15 of BPTI is fully solvent exposed. The protein stability of mutant BPTI species was examined and compared to wild-type BPTI, which contains the amino acid Lys that has a formal charge of +1. Because BPTI is a protein with unusual high stability that resists denaturation under conditions that are sufficient to unfold other proteins, in our studies we combined thermal and chemical protein denaturation. BPTI species were thermally denatured by increasing the temperature from 20 °C to 100 °C in the presence of either 6 M GdmCl or 8 M urea at pH 2.^{62, 65} At higher pH, BPTI is stabilized and hydrolysis of the peptide bond at Pro2 – Asp3 – Phe4 site is observed. It is known that urea and GdmCl denature proteins by distinct mechanisms: whereas urea disrupts hydrogen bonding, GdmCl is a salt that can also mask electrostatic interactions. Therefore, the use of both sets of conditions may yield valuable information regarding to what extent electrostatic interactions contribute to the stability of all variants of BPTI.¹⁸⁷

Under both sets of denaturation conditions, the BPTI_{Lys15Abu} variant has a significantly lower denaturation temperature than the wild-type BPTI_{Lys15}: in GdmCl lower by 3.1 °C, and in urea lower by 4.1 °C. Also, under both sets of conditions, the fluorinated variants have significantly higher melting temperatures than BPTI_{Lys15Abu}: in GdmCl higher by 4.9 °C (BPTI_{Lys15TfeGly}) or 5.4 °C (BPTI_{Lys15DfeGly}), and in urea higher by 2.1 °C (BPTI_{Lys15TfeGly}) or 3.1 °C (BPTI_{Lys15DfeGly}). The two data sets differ primarily in the relative stabilities of the wild-type and fluorinated species: protein stability increases in the order Abu < Lys < TfeGly <≅ DfeGly (in the presence of

5.3 Characterization of BPTI containing fluorinated amino acids

GdmCl) and Abu<TfeGly<DfeGly<Lys (in the presence of urea). That is, in the presence of GdmCl, there is no significant difference between BPTI_{Lys15TfeGly} and BPTI_{Lys15DfeGly}, and these fluorinated variants are the most stable; the wild-type is the second most stable (lower by 1.8 °C and 2.3 °C, respectively). In the presence of urea, the wild-type BPTI is the most stable (75.8 °C), the BPTI_{Lys15DfeGly} variant the second most stable, and the BPTI_{Lys15TfeGly} the third most stable, although it should be noted that these differences are at the limit of significance.

Our results show that the Lys15Abu substitution, a solvent-exposed position of the highly conserved interaction loop of BPTI, significantly destabilizes the folded state of the inhibitor, whereas replacement with partially fluorinated analogues restores (urea) or even enhances (GdmCl) the thermal stability of the BPTI. The Abu data agree well with previous studies of P₁ mutants of BPTI, in which differential scanning calorimetry (DSC) experiments revealed that introducing hydrophobic aliphatic or aromatic residues (Trp, Ile, Phe, Leu, Val, or Ala) at P₁ position is always destabilizing.⁶⁵ In contrast, the stabilizing effect of DfeGly and TfeGly suggest that these unnatural building blocks behave in a comparatively hydrophilic manner.

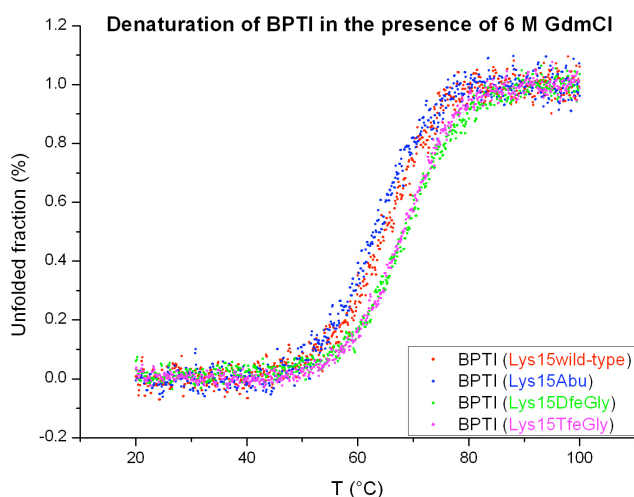


Figure 5.3-1 Thermal unfolding transition curves of BPTI variants at pH 2 in the presence of 6 M GdmCl and 8 M urea

(1) thermal denaturation of BPTI variants in the presence of 6 M GdmCl, (2) thermal denaturation of BPTI variants in the presence of 8 M urea (page 121)

5.3 Characterization of BPTI containing fluorinated amino acids

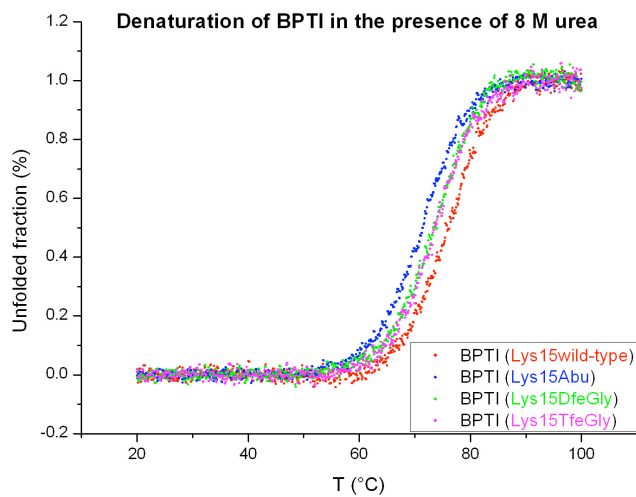


Table 5.3-1 Thermodynamic denaturation parameters for thermal unfolding of BPTI variants containing non-canonical amino acids at P₁ position and data from literature

	T_{den} in 6 M GdmCl (°C) ^a	ΔH_{vH} (kcal/mol) ^b	ΔG° (25°C) ^c (kcal/mol) ^c	T_{den} in 8 M Urea (°C) ^a	ΔH_{vH} (kcal/mol) ^b	ΔG° (25°C) ^c (kcal/mol) ^c	
Data from our studies							
	BPTI _{Lys15wild-type}	66.5	45.7	12.0	75.8	56.2	14.6
	BPTI _{Lys15Abu}	63.4	42.4	11.3	71.7	54.3	13.8
Data	BPTI _{Lys15DfeGly}	68.8	46.8	12.4	74.7	54.6	14.3
	BPTI _{Lys15TfeGly}	68.3	46.0	12.1	73.8	58.3	14.6
Data from literature^d							
	BPTI _{Lys15wild type}	~65	-	-	-	-	-
	BPTI _{Lys15wild type}	58.9	45.0	-	-	-	-
	BPTI _{Lys15His}	61.5	41.0	-	-	-	-
Data	BPTI _{Lys15Arg15}	59.2	44.3	-	-	-	-
	BPTI _{Lys15Thr}	60.1	44.3	-	-	-	-
	BPTI _{Lys15Trp}	54.5	40.4	-	-	-	-
	BPTI _{Lys15Asp}	61.4	46.5	-	-	-	-
	BPTI _{Lys15Met}	58.8	45.1	-	-	-	-
	BPTI _{Lys15Phe}	56.3	41.7	-	-	-	-

a) T_{den} is defined as the temperature at which the fraction of folded is 0.5. Errors are typically not higher than 0.5 °C, b) Errors are typically not higher than 0.1 kcal/mol, c) ΔG° values were calculated for 1 M standard at 25 °C, errors are typically not higher than 0.2 kcal/mol, d) Data was obtained by performing experiments in the presence of 6 M GdmCl in 10 mM glycine-HCl buffer, pH 2^{62, 65, 187}.

5.3 Characterization of BPTI containing fluorinated amino acids

5.3.2 Studies of mutant BPTI-protease interactions

5.3.2.1 Studies of inhibitor activity

BPTI strongly inhibits serine proteases. An estimated K_a (association constant) of BPTI with β -trypsin is about 10^{14} , and only a few groups have actually been able to measure it. Because the k_{off} of BPTI-trypsin complex is too low, an impractical amount of time would be required to determine it.¹⁸⁴ The inhibition profile of synthetic mutant BPTI species with respect to α -chymotrypsin, β -trypsin, and plasmin from human plasma was tested and compared with wild type BPTI.

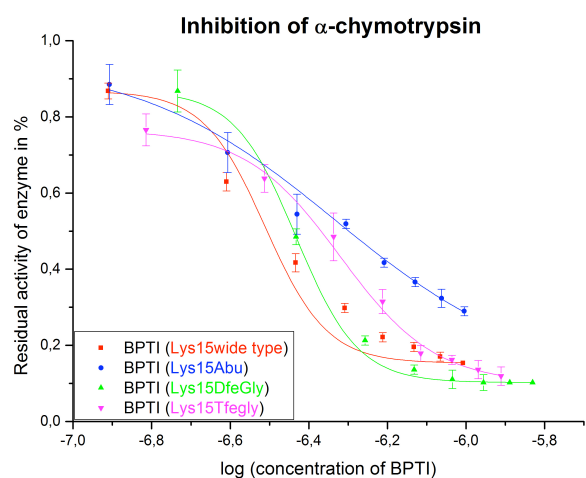
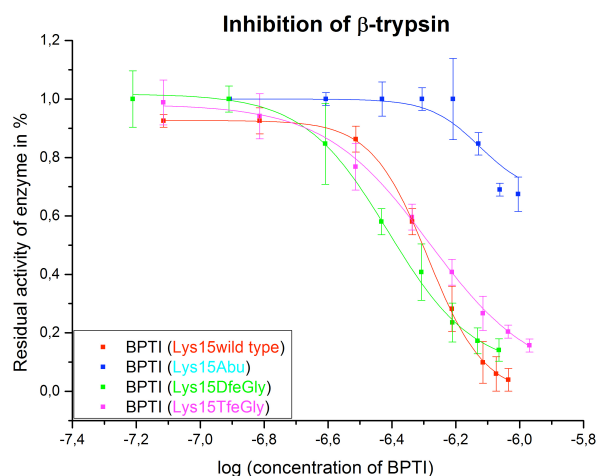
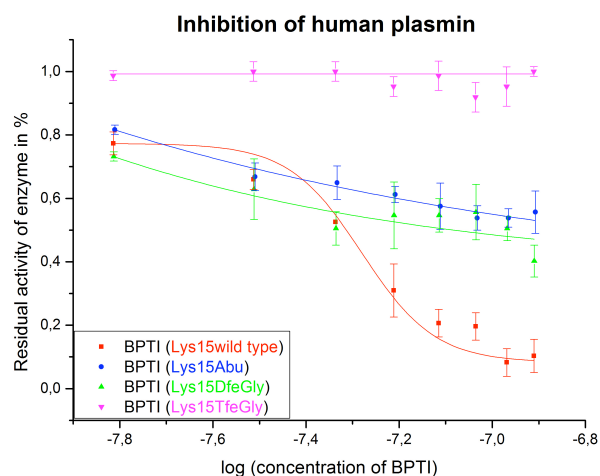


Figure 5.3-2 Schematic representation of inhibitory activity of synthetic BPTI variants to α -chymotrypsin, β -trypsin, and plasmin from human plasma



5.3 Characterization of BPTI containing fluorinated amino acids

Table 5.3-2 Binding disassociation constant (K_d M) of BPTI variants with proteases and data from literature^a

BPTI [Xaa15]	β -trypsin	α -chymotrypsin	Human plasmin
Data from our studies			
Lys (wild-type)	5.17×10^{-7}	3.08×10^{-7}	5.25×10^{-8}
Abu	n. d.	n. d.	(n. d.)
DfeGly	3.88×10^{-7}	3.69×10^{-7}	(n. d.)
TfeGly	5.20×10^{-7}	4.79×10^{-7}	(n. d.)
Data from literature^b			
Wild type (Lys)	5.88×10^{-14}	1.1×10^{-8}	1.45×10^{-10}
Gly	6.67×10^{-5}	1.27×10^{-5}	
Ala	3.70×10^{-6}	4.76×10^{-7}	
Val	2.44×10^{-5}	4.35×10^{-7}	
Ile	9.1×10^{-6}	1.69×10^{-6}	
Ser	2.7×10^{-8}	3.70×10^{-6}	
Thr	3.45×10^{-6}	4.00×10^{-7}	
Asp	1.59×10^{-5}	8.3×10^{-5}	
Asn	4.55×10^{-8}	9.2×10^{-6}	
Met	2.56×10^{-8}	9.8×10^{-8}	
Leu	1.85×10^{-7}	1.3×10^{-9}	
Glu	4.7×10^{-7}	1.2×10^{-5}	
Gln	4.35×10^{-7}	5.2×10^{-7}	
His	1.6×10^{-7}	7.4×10^{-7}	
Arg	n. d.	2.5×10^{-8}	
Phe	8.3×10^{-9}	2.5×10^{-9}	
Tyr	6.25×10^{-9}	7.6×10^{-9}	
Trp	1.32×10^{-9}	5.6×10^{-9}	
Data from literature^c			
Abu ³	+	+	
Ape	1.3×10^{-7}	1.4×10^{-8}	
Ahx	2.0×10^{-8}	6.7×10^{-9}	
tLeu	-	-	
Neo	+	+	

a) Different K_d and K_a values of BPTI interaction with enzyme can be found in literature; these results depend on enzyme activity and assay conditions. Literature values of synthetic BPTI include: $K_d = 1.19 \times 10^{-8}$ ¹⁸⁴ and 1.3×10^{-8} ;¹⁸¹ b) Krowarsch *et al.*,^{69b} c) Beckmann *et al.*,¹⁸⁸ n. d. = not determined, “+”: weak interaction, “-”: no interaction.

The results are schematically represented in *Figure 5.3-2* and summarized in *Table 5.3-2*, with comparison to literature values. In the case of α -chymotrypsin, all three

5.3 Characterization of BPTI containing fluorinated amino acids

mutant BPTI variants show reduced inhibition, however the inhibitory activity of both fluorinated variants are higher than that of Abu containing BPTI. In the case of human plasmin, BPTI variants containing Abu and DfeGly can weakly inhibit the protease while the one containing TfeGly can not inhibit the enzyme at a detectable level. The results from BPTI-trypsin inhibition studies indicate a similar tendency to the one found in the case of α -chymotrypsin; that is, the variant containing Abu inhibits the enzyme to a significantly lesser extent, but both fluorinated variants retain inhibitory ability. These results reveal that fluorine-induced effects are involved in the protein-protein interactions.

It has been widely accepted, upon complex formation, inhibitor reactive loop (Pro13 (P₃) - Arg17 (P₂')) of BPTI forms an anti-parallel β -sheet with the polypeptide of the proteases. In the binding model of BPTI-trypsin interaction, the Lys15 of wild type BPTI interacts with Asp189 of β -trypsin through an electrostatic interaction. In the binding model of BPTI toward α -chymotrypsin, NH₃⁺ of Lys15 forms hydrogen bonds with side chain of Gly216 and Ser217 of α -chymotrypsin. Krowarsch *et al.* performed the binding studies of BPTI to β -trypsin and α -chymotrypsin by substitution of Lys15 with canonical amino acids. They demonstrated that beside the ionic strength between Lys15 and Asp189, an additional binding force, a burial hydrophobicity from -(CH₂)₄- group of Lys in S₁ pocket contributes to the binding as well. In the case of α -chymotrypsin, the burial hydrophobicity of P₁ side chain serves predominately as binding source to the inhibition.^{69b} It is worth noting that crystal structures of the complexes between β -trypsin and ten BPTI mutant variants (P₁ mutation with canonical amino acids) and between α -chymotrypsin and five BPTI mutant variants (P₁ mutation with canonical amino acids) are available.¹⁸⁹ Beckmann *et al.* reported that mutant BPTI containing Abu at P₁ position can still form the protein-protein complex with β -trypsin with a lower binding affinity, which is agreeable to our observation.¹⁸⁸ Our inhibition assay can not directly give detail mechanism of binding between fluorinated BPTI species to the proteases. The protein

5.3 Characterization of BPTI containing fluorinated amino acids

crystallography was carried out for binding studies of fluorinated BPTI variants to proteases.

5.3.2.2 Probing inhibitor-protease binding by means of ITC

Isothermal titration calorimetry (ITC) is a physical technique, which has been used extensively to determine protein-ligand/protein interactions. When two components bind, the change in thermodynamic potential, including ΔG , ΔH , and ΔS , can be measured directly. In order to obtain more precise and direct information about the thermodynamic parameters of the binding of mutant BPTI variants to the proteases, ITC experiments were conducted.

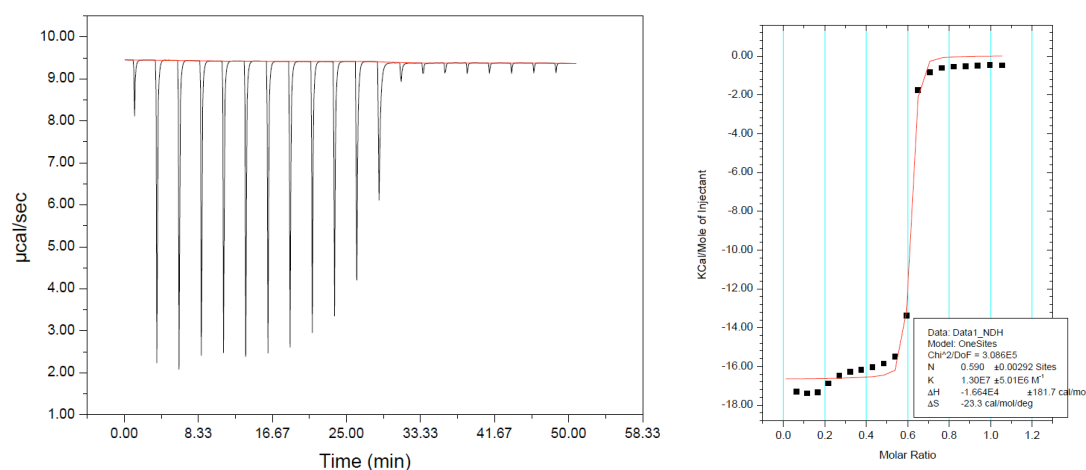


Figure 5.3-3 ITC measurement of wild-type BPTI with β -trypsin

$$k_a = 1.30E7 \pm 5.01E6 \text{ M}^{-1}, \Delta H = -1.664E4 \pm 181.7 \text{ cal/mol}, \Delta S = -23.3 \text{ cal/mol/deg}$$

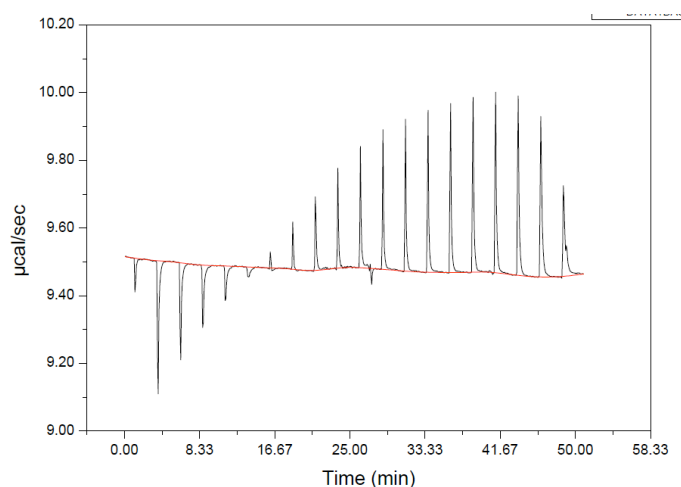


Figure 5.3-4 ITC measurement of wild-type BPTI with α -chymotrypsin in which equilibrium was not reached

5.3 Characterization of BPTI containing fluorinated amino acids

As Figure 5.3-3 and 5.3-4 show that the ITC experiments yielded results in the case of wild-type BPTI interacting with β -trypsin, but that the thermodynamic values of interaction between wild-type BPTI and α -chymotrypsin could not be determined because equilibrium was not reached upon titration. In the former case, the value for the binding constant is in the same range as the value obtained from the enzyme inhibition assay (From ITC, $K_a = 1.30 \times 10^7 \pm 5.01 \times 10^6 \text{ M}^{-1}$, from inhibition assay: $K_a = 1.9 \times 10^6 \text{ M}^{-1}$) In the latter case, the experiment was perhaps complicated by the equilibrium between monomer and dimer for α -chymotrypsin alone, which has been shown to influence the outcome of ITC.¹⁹⁰

5.3.3 Protein crystallographic analysis of protease-BPTI complexes

To further characterize the impact of fluorine on the interaction between a protease and BPTI, protein crystallography was undertaken.

5.3.3.1 Purification of protease-BPTI complexes

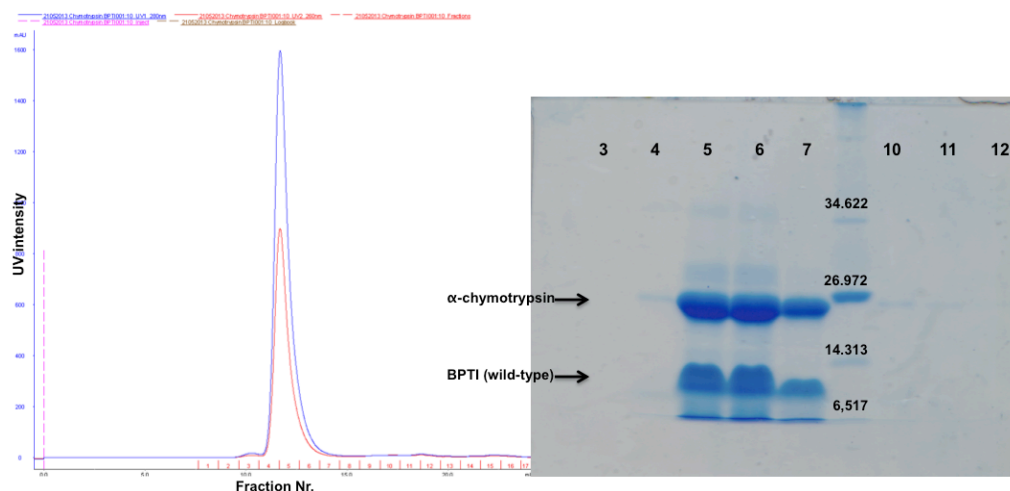


Figure 5.3-5 Size-exclusion spectrum of α -chymotrypsin-BPTI_{Lys15wild-type} complex and SDS PAGE analysis of appropriate fractions obtained from size-exclusion chromatography

Left, size-exclusion spectrum of α -chymotrypsin-BPTI_{Lys15wild-type}, red line indicates the UV detection at 260 nm, blue line indicates the UV detection at 280 nm; **right**, SDS-PAGE analysis of fractions obtained from size-exclusion chromatography

5.3 Characterization of BPTI containing fluorinated amino acids

In order to obtain β -trypsin-BPTI and α -chymotrypsin-BPTI complexes of sufficiently high purity for crystallographic analysis, the purification of protein complexes was carried out by means of size-exclusion (gel-filtration) chromatography under different buffer conditions. The size-exclusion spectra and SDS-PAGE analysis of β -trypsin-BPTI (wild-type) and α -chymotrypsin-BPTI (wild-type) complexes are shown in Figures 5.3-5 and 5.3-6.

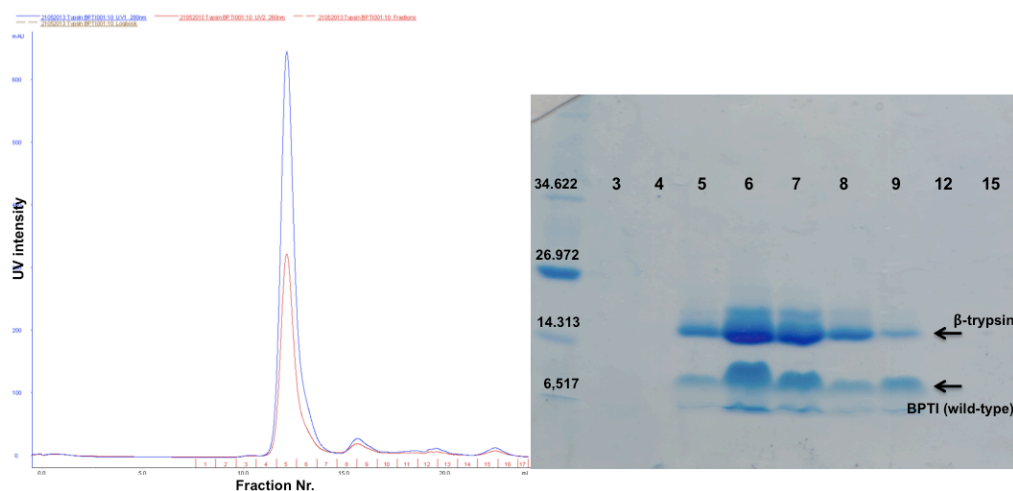


Figure 5.3-6 Size-exclusion spectrum of β -trypsin-BPTI_{Lys15wild-type} complex and SDS PAGE analysis of appropriate fractions obtained from size-exclusion chromatography

Left, size-exclusion spectrum of α -chymotrypsin-BPTI_{Lys15wild-type}, red line indicates the UV detection at 260 nm, blue line indicates the UV detection at 280 nm; **right**, SDS-PAGE analysis of fractions obtained from size-exclusion chromatography

5.3.3.2 Crystallization of proteases-BPTI complexes

Crystals of protease-BPTI complexes were prepared by a screening technique that relies on sitting drop vapor diffusion. Well-formed crystals were selected by optical microscopy and used to collect X-ray scattering data. Figure 5.3-7 and Figure 5.3-8 show typical crystal morphologies of β -trypsin-BPTI complexes and α -chymotrypsin complexes, respectively.

5.3 Characterization of BPTI containing fluorinated amino acids

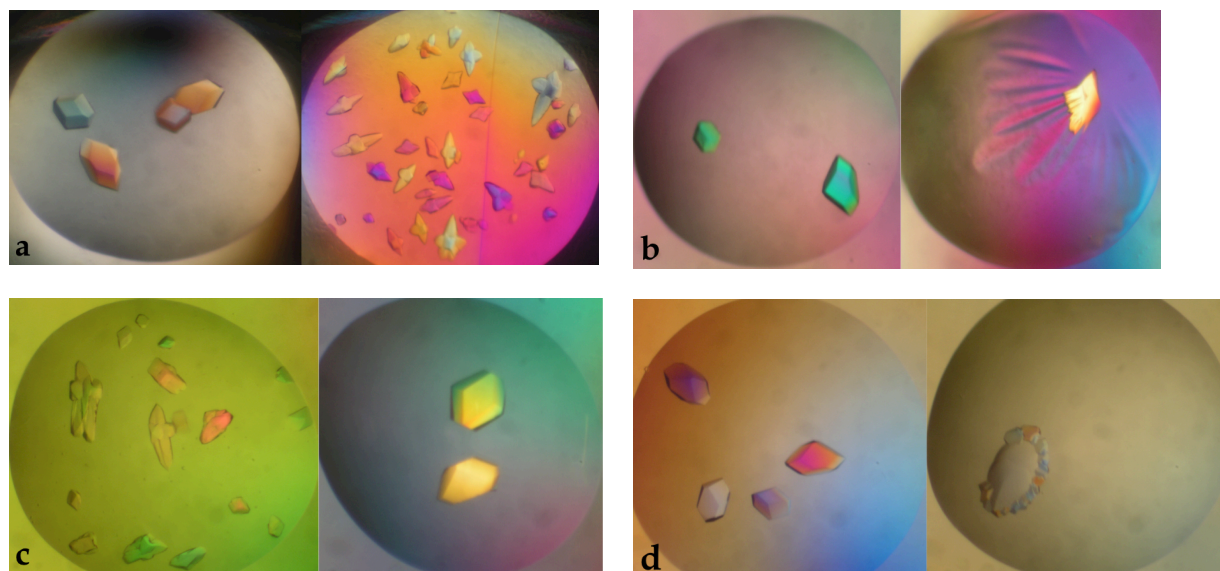


Figure 5.3-7 Crystals of β -trypsin-BPTI complexes under microscopy ($\sim 60 \times$)

a) β -trypsin-BPTI_{Lys15wild-type} b) β -trypsin-BPTI_{Lys15Abu}, c) β -trypsin-BPTI_{Lys15DfeGly}, d) β -trypsin-BPTI_{Lys15TfeGly}



Figure 5.3-8 Crystals of α -chymotrypsin-BPTI_{Lys15wild type} complexes under microscopy ($\sim 60 \times$)

Data for the protease-BPTI_{Lys15wide-type} complex were obtained with resolutions of 1.25 Å for β -trypsin and 2.08 Å for α -chymotrypsin. Data for the β -trypsin-BPTI_{mutant} complexes were obtained with resolutions of 1.37 Å for the BPTI_{Lys15Abu} variant, 1.37 Å for the BPTI_{Lys15DfeGly} variant, and 1.30 Å for the BPTI_{Lys15TfeGly} variant.

5.3 Characterization of BPTI containing fluorinated amino acids

5.3.3.3 Crystallographic analysis of proteases-BPTI complexes

The crystal structural data of the four β -trypsin-BPTI complexes were refined and analyzed. The binding pocket of β -trypsin (S_1 pocket) occupied by the interaction loop of BPTI containing four different residues at P_1 position, and water molecules, is depicted in *Figure 5.3-9* (side view).

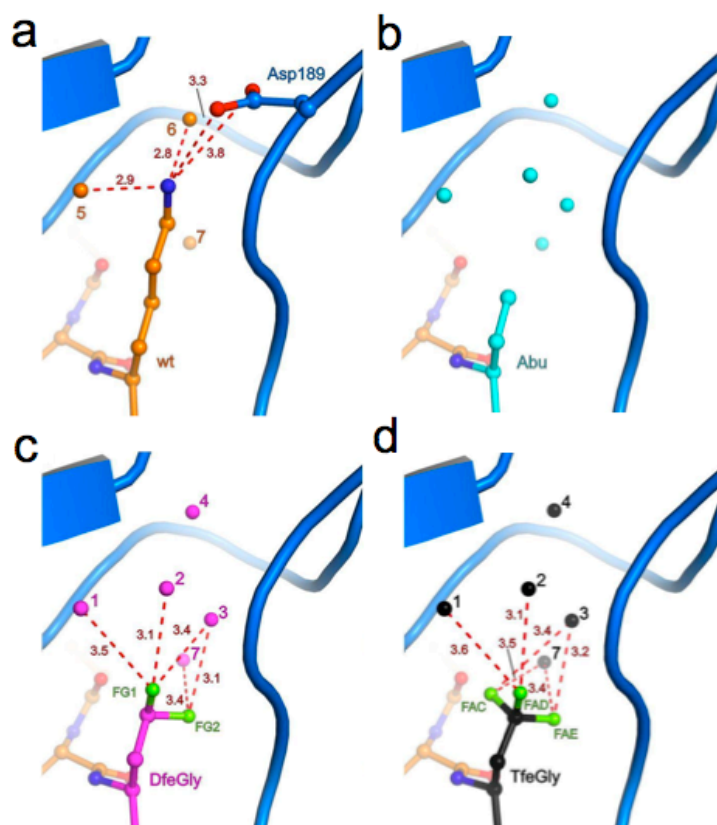


Figure 5.3-9 Structural context of the P_1 residues in the S_1 binding pocket

a) β -trypsin-BPTI_{Lys15}wide-type, **b)** β -trypsin-BPTI_{Lys15}Abu, **c)** β -trypsin-BPTI_{Lys15}DfeGly, **d)** β -trypsin-BPTI_{Lys15}TfeGly; structural water molecules were assigned unique numerical identifiers 1-7, distance between the P_1 side chain and water molecules are given in Å (distance showed are smaller than 4 Å).

Remarkably, as presented in *Figure 5.3-10*, the β -trypsin-BPTI complexes containing the non-canonical amino acids are virtually identical in conformation and S_1 binding site to the β -trypsin-BPTI_{Lys15}wide-type complex. Also the root-mean-square deviation (RMSD) analysis indicates negligible differences in the superimposed complexes (*Table 5.3-3*). The only significant difference among the structures is the number of

5.3 Characterization of BPTI containing fluorinated amino acids

water molecules in the proximity of the side chain at position 15 in BPTI: two water molecules occupy the space where the side chain of lysine has been substituted with the shorter neutral alkyl or fluoroalkyl groups (Figure 5.3-10).

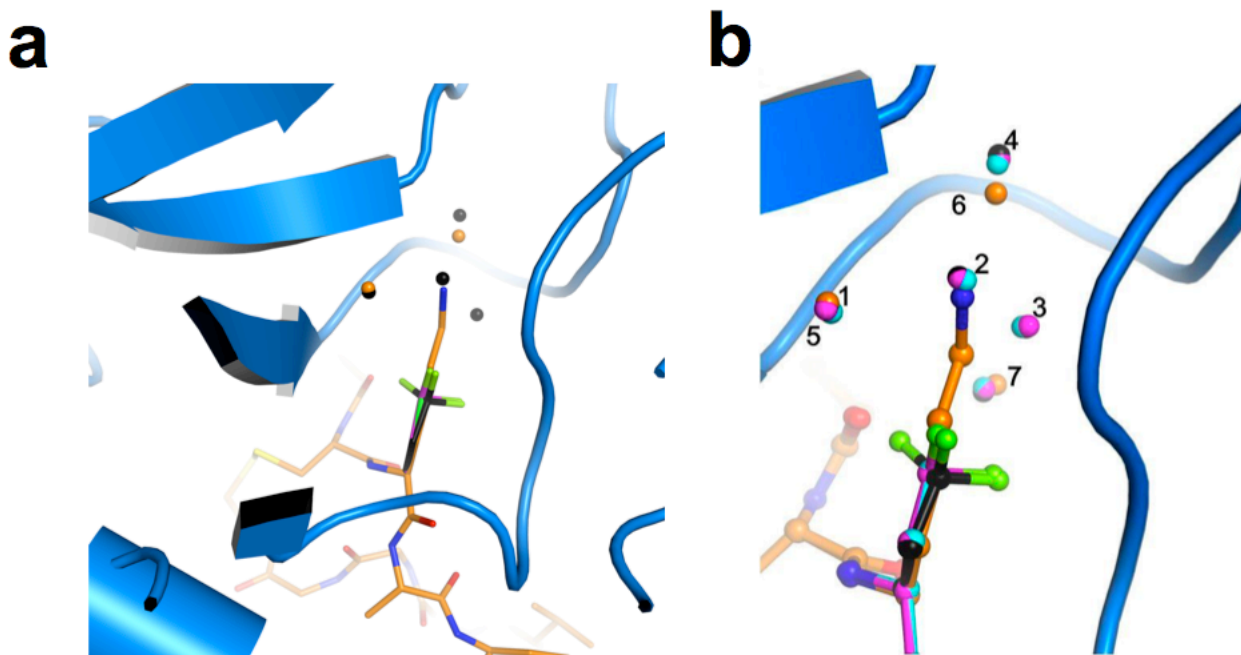


Figure 5.3-10 Superimposition of residues 15 of BPTI and structural water molecules

a) Superimposition of the BPTI binding loop (Pro13 - Arg17) in four structural coordinates, b) superimposition of water molecules in S1 pocket from four structural coordinates (the P₁ side chain and water molecules are labeled with different color, which are according to the Figure 5.3-9, Lys15 in orange, Abu15 in cyan, DfeGly15 in magenta, TfeGly15 in black).

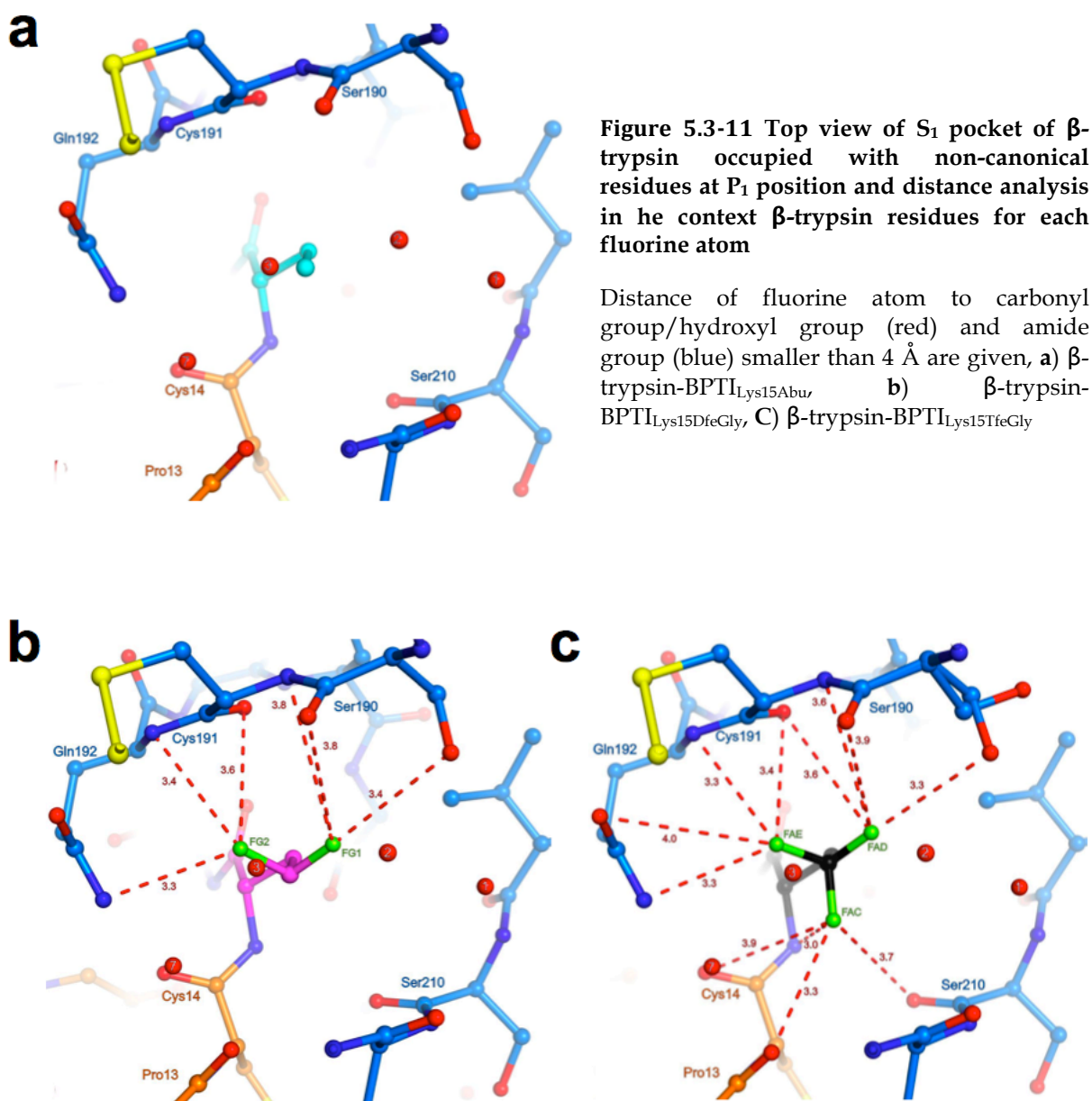
Table 5.3-3 RMSD analysis in Å of β -trypsin-BPTI_{mutant} complexes in comparison with the β -trypsin-BPTI_{Lys15wild-type}

Complexes	β -trypsin all	β -trypsin-BPTI _{Lys15wild-type}	
		BPTI all	BPTI (Pro13 - Arg17) all/main chain
Trypsin-BPTI _{Lys15Abu}	0.50	0.78	0.11/0.06
Trypsin-BPTI _{Lys15DfeGly}	0.38	0.90	0.09/0.05
Trypsin-BPTI _{Lys15TfeGly}	0.45	0.85	0.11/0.06

The binding pockets of β -trypsin with fluorinated side chain of mutant BPTI variants were precisely analyzed. Figure 5.3-11 schematically presented occupancy of non-

5.3 Characterization of BPTI containing fluorinated amino acids

canonical side chain in the S_1 pocket and distance analysis in the context of β -trypsin residues for each fluorine atom. The distances between fluorine atoms and the residues of β -trypsin smaller than 4 Å were summarized also in Table 5.3-4. It worthy noting, short distances between fluorine atom and carbonyl carbon of Cys191 residue were observed (3.2 Å for DfeGly and 3.1 Å for TfeGly). These distances indicate potential orthogonal $C-F\cdots C=O$ interactions.^{20a}



5.3 Characterization of BPTI containing fluorinated amino acids

Table 5.3-4 Summary of distances between side chain atoms of P₁ residues and atoms belonging to S1 binding pocket of β -trypsin or BPTI^a

aa at P ₁ position	Protein and residue		Distance (Å)	
Abu	ABA ^{HG1}	β -trypsin	Ser190 ^{OG}	3.8
	ABA ^{HG1}		Ser190 ^O	4.0
	ABA ^{HG2}		Gln192 ^N	3.4
	ABA ^{HG2}		Cys191 ^o	3.4
	ABA ^{HG3}		Ser210 ^O	3.6
	ABA ^{HG3}		Trp211 ^N	3.9
DfeGly	ABA ^{HG3}	BPTI	Pro13 ^O	3.3
	ABA ^{HG3}		Cys14 ^O	3.8
	OBF ^{FG1}	β -trypsin	Ser190 ^{OG}	3.4
	OBF ^{FG1}		Ser190 ^O	3.8
	OBF ^{FG1}		Ser190 ^C	3.8 ^b
	OBF ^{FG2}		Cys191 ^C	3.2 ^b
	OBF ^{FG2}		Gln192 ^{NE2}	3.3
	OBF ^{FG2}		Gln192 ^N	3.4
	OBF ^{FG2}	Cys191 ^O	3.6	
	OBF ^{FG2}	Gln192 ^{CD}	3.7 ^b	
	OBF ^{HG}	BPTI	Ser210 ^O	3.6
	OBF ^{HG}		Trp211 ^N	3.9
OBF ^{HG}	OBF ^N		2.8	
OBF ^{HG}	Pro13 ^O		3.2	
OBF ^{HG}	Cys14 ^O		3.8	
OBF ^{HG}				
TfeGly	3EG ^{FAC}	β -trypsin	Ser210 ^O	3.7
	3EG ^{FAC}		Trp211 ^N	3.9
	3EG ^{FAC}	BPTI	3EG ^N	3.0
	3EG ^{FAC}		Pro13 ^O	3.3
	3EG ^{FAC}		Cys14 ^C	3.4 ^b
	3EG ^{FAC}		Cys14 ^O	3.9
	3EG ^{FAD}	β -trypsin	Ser190 ^{OG}	3.3
	3EG ^{FAD}		Ser190 ^C	3.6 ^a

5.3 Characterization of BPTI containing fluorinated amino acids

3EG ^{FAD}		Ser190 ^O	3.6
3EG ^{FAD}		Cys191 ^N	3.6
3EG ^{FAD}		Cys191 ^O	3.8
3EG ^{FAD}	β -trypsin	Cys191 ^C	3.9 ^b
3EG ^{FAE}		Cys191 ^C	3.1 ^b
3EG ^{FAE}		Gln192 ^{NE2}	3.3
3EG ^{FAE}		Cys191 ^O	3.4
3EG ^{FAE}		Gln192 ^{CD}	3.7 ^b
3EG ^{FAE}		Gln192 ^{OE2}	4.0

a) Distances present here are smaller than 4 Å, representation of side chain is according to the atoms nomenclature of amino acids, the nomenclature of non-canonical amino acids are according to name existed in protein data bank, ABA=Abu, OBF=DfeGly, 3EG=TfeGly; b) distances between fluorine atom to carbon of carbonyl group.

Moreover, interestingly, we notice a close distance (2.8 Å) between hydrogen atom OBF^{HG} of difluoroethyl group and nitrogen atom OBF^N of polypeptide backbone, which indicates a potential intramolecular hydrogen bond. Due to the strong electron withdrawing effect of both fluorine atoms, the OBF^{HG} hydrogen can serve as a good hydrogen donor. In case of Lys15TfeGly, a distance of 3.0 Å between fluorine atom 3EG^{FAC} and nitrogen atom 3EG^N of peptide backbone was observed as well. It should be noted, because of low electron density, the hydrogen atom on the 3EG^N could not be refined. Nevertheless, the distance also indicates a possible intramolecular interaction between fluorine and hydrogen. Such intramolecular interaction might be relevant to the thermal stability of fluorinated BPTI, that is, both fluorinated amino acids show higher thermal stability.

Electrostatic surface potential of trypsin S₁ binding site was mapped, agreeable to the previous studies that S₁ binding pocket is quite hydrophilic. The electron density map of trifluoromethyl group and its occupancy in S₁ pocket are showed in *Figure 5.3-12*.

5.3 Characterization of BPTI containing fluorinated amino acids

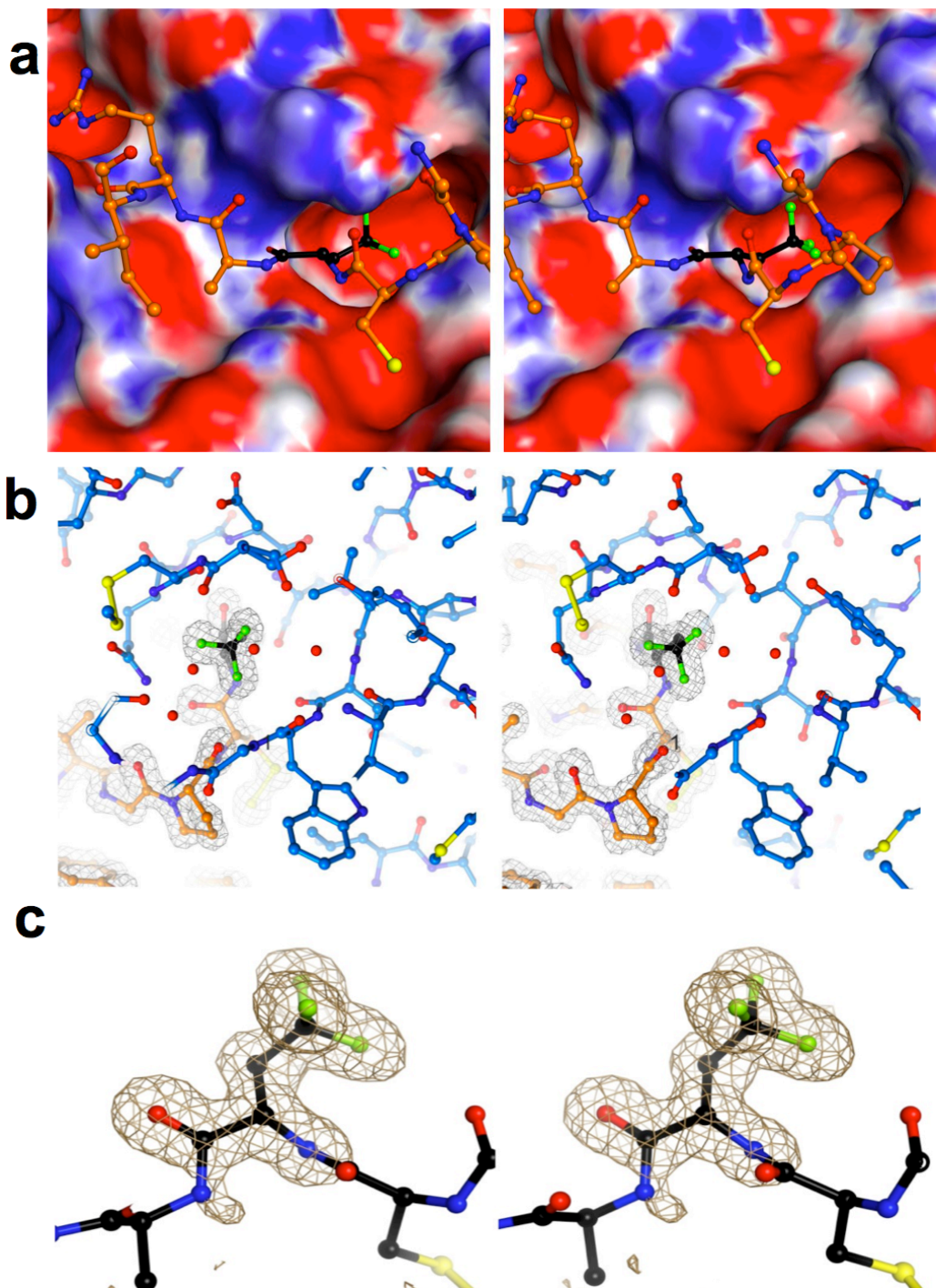


Figure 5.3-12 Stereoview of the interacting loop of BPTI_{Lys15TfGly} in S₁ pocket

a) stereoview of interacting loop in S₁ pocket, S₁ pocket showed as surface with different properties of amino acid residues, the negative charged residues are labeled in red, the positive charged residues are labeled in blue, hydrophobic residues are in white; b) stereoview of interacting loop of S₁ pocket with electron density, c) stereoview of trifluoromethyl group that has been omitted for calculation of a simulated annealing omit map contoured at 3.0 σ shown as brown mesh.

5.3 Characterization of BPTI containing fluorinated amino acids

As revealed by the crystallographic analysis, all three mutant BPTI variants display similar binding modes. Aside from the difference in stoichiometry of structural waters, an important difference can also be seen in the *B*-factors of water molecules in the neighborhood of the substituted side chain. The *B*-factor is generally used to indicate the relative vibrational motion of different residues in the crystal structures. As summarized in Table 5.3-5, the average *B*-factor values of all structural water molecules are greater than 30 Å². The *B*-factors of water molecules in the S₁ pocket are much lower than the average *B*-factor values of overall structural water molecules, which indicate the reduced dynamics of these water molecules in S₁ binding site. Furthermore, the *B*-factor values of water molecules 1, 2, 3, and 7 proximal to fluorinated side chains are significantly lower than those proximal to Abu.

Table 5.3-5 Summary of *B*-factors of β-trypsin-BPTI complexes

<i>B</i> -factor (Å ²)	β-trypsin-BPTI complexes			
	Wild type	Lys15Abu	Lys15DfeGly	Lys15TfeGly
Average B-factor of protein complex	20.7	21.3	19.6	21.2
All water molecules on average	31.0	33.8	31.2	32.0
BPTI Pro13 - Arg17				
All atoms	13.0	13.2	12.4	13.4
Main chain	12.0	12.2	11.1	12.3
Side chain	14.0	14.6	13.9	14.6
BPTI Xaa15^a				
All atoms	12.1	11.9	13.2	12.9
Main chain	11.6	11.5	11.0	11.8
Side chain	12.5	12.8	15.5	13.7
Water molecules in S1 pocket^b				
1	-	21.8	17.5	18.4
2	-	27.7	20.0	20.5
3	-	31.2	22.2	20.7
4	-	20.7	19.1	19.4
5	13.5	-	-	-
6	14.1	-	-	-
7	15.4	23.0	17.8	17.9

5.3 Characterization of BPTI containing fluorinated amino acids

Kawamura *et al.* investigated the network of hydrogen bonds in the S₁ pocket of trypsin-BPTI_{Lys15wild-type} complex by means of H/D-exchange and neutron diffraction and demonstrated that the hydrogen bonds contributing the protease-BPTI interaction include protonated NH₃⁺ of Lys15 to Ser190 by bifurcated hydrogen bonds and two water molecule-mediated hydrogen bonds, namely Lys15-H₂O-Val227 and Lys15-H₂O-Asp189.¹⁹¹ To explore the fluorine induced interactions in our studies, the distances between the structural water molecules and trypsin residues and distances between water molecules and fluorine atoms were analyzed (Table 5.3-6). With the exception that two distance values between the fluorine atoms and water molecules 2 and 3 are 3.1 Å, observed with both fluorinated side chains, other distance values between fluorine atoms and structural water molecules are greater than 3.5 Å. It is likely that the closest distances of 3.1 Å correspond to the lower *B*-factor values of water molecules 2 and 3. The distances between structural water molecules and trypsin residues around 3.0 Å were observed, which revealed the evidence of structural water molecule mediated interactions (hydrogen bond network).

Table 5.3-6 Summary of distances between P₁ side chains and structural water molecules in S₁ pocket, distances between water molecules and trypsin residues (distances given in Å)^a

Residues/ water		Water molecules in S ₁ pocket						
		1	2	3	4	5	6	7
Distances between side chain atom of P₁ position and structural water molecules in S₁ pocket								
Lys	Lys ^{NZ}	-	-	-	-	2.9	2.8	5.1
Abu	ABA ^{HG1}	3.6	3.2	3.4				4.4
	ABA ^{HG2}	5.1	4.2	3.1				3.4
	ABA ^{HG3}	4.3	4.2	3.6				3.5
DfeGly	OBF ^{FG1}	3.5	3.1	3.4	-	-	-	4.1
	OBF ^{FG2}	5.5	4.3	3.1	-	-	-	3.4
	OBF ^{HG}	4.5	4.3	3.7				3.5
TfeGly	3EG ^{FAC}	4.3	3.9	3.4	-	-	-	3.4
	3EG ^{FAD}	3.6	3.1	3.5	-	-	-	4.8
	3EG ^{FAE}	5.5	4.4	3.2	-	-	-	3.5

5.3 Characterization of BPTI containing fluorinated amino acids

Distances between the structural water molecules											
Water molecules in S ₁ pocket of trypsin-BPTI _{wild-type}	5	-	-	-	-	-	5.1	7.1			
	6	-	-	-	-	-	-	4.9			
	7	-	-	-	-	-	-	-			
Water molecules in S ₁ pocket of trypsin-BPTI _{Lys15Abu}	1	-	2.9	5.0	5.5	-	-	6.7			
	2	-	-	2.6	3.3	-	-	5.1			
	3	-	-	-	3.3	-	-	2.8			
	4	-	-	-	-	-	-	5.2			
Water molecules in S ₁ pocket of trypsin-BPTI _{Lys15DfeGly}	1	-	2.9	5.2	5.7	-	-	6.9			
	2	-	-	2.7	3.4	-	-	5.1			
	3	-	-	-	3.4	-	-	2.9			
	4	-	-	-	-	-	-	5.2			
Water molecules in S ₁ pocket of trypsin-BPTI _{Lys15TfeGly}	1	-	2.9	5.2	5.7	-	-	6.8			
	2	-	-	2.8	3.4	-	-	5.2			
	3	-	-	-	3.5	-	-	2.9			
	4	-	-	-	-	-	-	5.4			
Distances between structural water molecules and the residues of trypsin/BPTI, decorated with Lys15Abu											
β-Trypsin		1		2		3		4		7	
		Ser190 ^{OG}	3.0	Asp189 ^{OD1}	3.0	Ser190 ^O	3.4	Asp189 ^{OD2}	2.8	Gln192 ^{NE}	3.1
		Trp211 ^O	3.0	Asp189 ^{OD2}	3.5	Gly214 ^O	2.6	Ser213 ^N	3.3		
		Val227 ^N	3.4	Ser190 ^O	3.1			Gly214 ^O	3.2		
BPTI		Val227 ^O	2.9	Ser190 ^{OG}	3.0			Lys220 ^O	3.2		
										Pro13 ^O	2.8
Distances between structural water molecules and the residues of trypsin/BPTI, decorated with Lys15DfeGly											
β-Trypsin		1		2		3		4		7	
		Ser190 ^{OG}	3.0	Asp189 ^{OD1}	2.9	Ser190 ^O	3.3	Asp189 ^{OD2}	2.8	Gln192 ^{NE}	3.1
		Trp211 ^O	3.1	Asp189 ^{OD2}	3.4	Gly214 ^O	2.6	Ser213 ^N	3.2		
		Val227 ^N	3.4	Ser190 ^O	3.2			Gly214 ^O	3.3		
BPTI		Val227 ^O	2.9	Ser190 ^{OG}	3.0			Lys220 ^O	3.1		
										Pro13 ^O	2.8
Distances between structural water molecules and the residues of trypsin/BPTI, decorated with Lys15TfeGly											
β-Trypsin		1		2		3		4		7	
		Ser190 ^{OG}	3.0	Asp189 ^{OD1}	2.9	Ser190 ^O	3.3	Asp189 ^{OD2}	2.8	Gln192 ^{NE}	3.1
		Trp211 ^O	3.0	Asp189 ^{OD2}	3.5	Gly214 ^O	2.7	Ser213 ^N	3.2		
	Val227 ^N	3.3	Ser190 ^O	3.2			Gly214 ^O	3.3			

5.3 Characterization of BPTI containing fluorinated amino acids

Val227^O 2.9 Ser190^{OG} 2.9 Lys220 3.0

BPTI Pro13^O 2.8
 a) Distances given here between water molecules and trypsin residues are cut-off in 3.5 Å

Based on the distances and *B*-factors analysis, obviously, structural water molecules participate in a hydrogen bond network at S₁ pocket and fluorinated side chains weakly interact with water molecules referred to lower *B*-factor values. Conclusion can be drawn that fluorine atoms team up with water molecules contributing to the interaction between fluorinated BPTI species and β-Trypsin. Thus, the fluorinated BPTI can restore the inhibitory activity to the BPTI containing Abu.

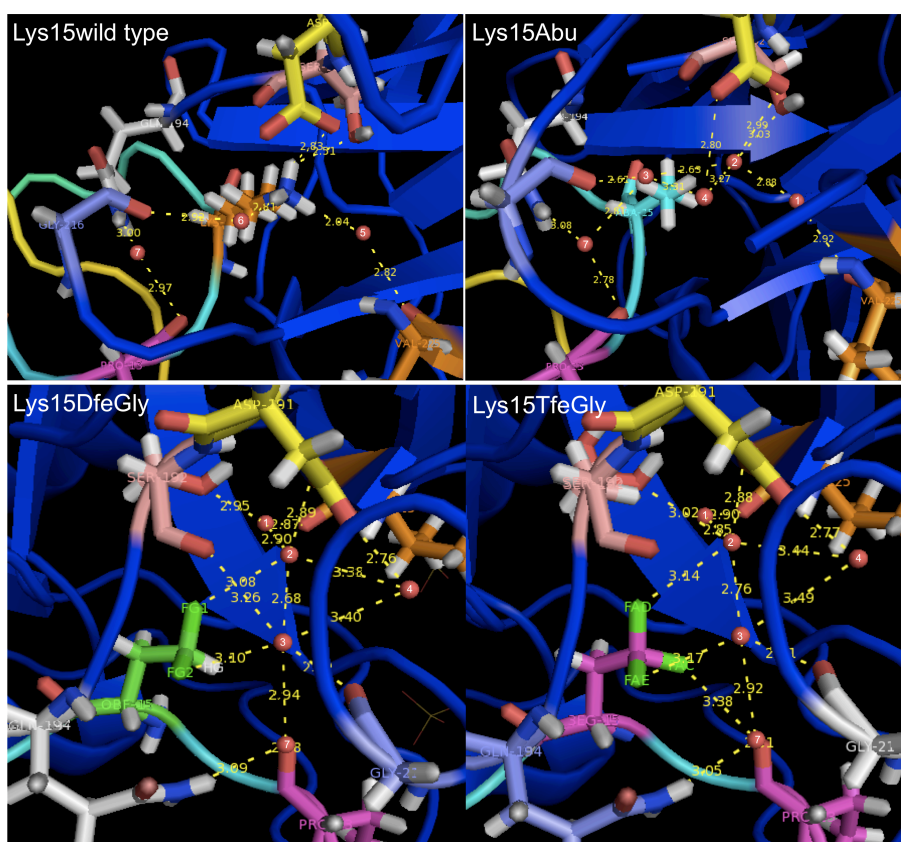


Figure 5.3-13 Schematic representation of P₁ residues in the S₁ pocket and distances between potential interacting side chains

5.4 Methods and experimental procedures

5.4.1 General information about chemical synthesis and analysis

All reagents of synthetic grade were used as supplied. Thin layer chromatography (TLC) for monitoring reactions was carried out on *Merck* silica gel 60 F254 plates. Acetonitrile (MeCN) for reversed-phase high-performance liquid chromatography (RP-HPLC) was purchased from *Acros Organics*. Building blocks and chemicals for peptide synthesis and cleavage were purchased from the following suppliers. Fmoc-L-amino acids, 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate (TBTU), and 1-hydroxybenzotriazole (HOBT) were purchased from *Fa. Gerhardt*; dimethylformamide for synthesis, DIPEA (98+ %), *N,N*-diisopropylcarbodiimide (DIC, 99 %), piperidine (99 %, extra pure), TFA for synthesis (99 %), triisopropylsilane (TIPS 99 %), and 1,2-ethanedithiol (EDT, 95%) were purchased from *Acros*; phenol (99%) was purchased from *Janssen Chimica*; thioanisole (99%) was purchased from *Chimica*; and Fmoc-(*S*)-2-aminobutyric acid (Abu) was purchased from *Bachem*. (*S*)-2-amino-4,4-difluorobutyric acid (DfeGly) and TfeGly were synthesized according to literature procedures.¹⁶⁸ Standard chemicals that are not listed here are from *Sigma-Aldrich* and *Fluka*. Deionized water for buffer solutions was prepared using the MilliQ-Advantage A10-System (*Millipore*).

Nuclear magnetic resonance

¹H and ¹⁹F nuclear magnetic resonance (NMR) spectra were obtained on a Bruker AVANCE 500 MHz NMR spectrometer. Proton chemical shifts are reported in parts per million (ppm) and referenced to the residual proton peak of chloroform-d₁. Spectral coupling patterns are reported as follows: b, broad; s, singlet; t, triplet; and m, multiplet.

Mass spectrometry

5.4 Methods and experimental produces

Electrospray-ionization time-of-flight high-resolution mass spectrometry (ESI-TOF) was performed on an Agilent 6210 ESI-TOF (*Agilent Technologies*).

Reversed-phase high-performance liquid chromatography

RP-HPLC was performed with appropriate gradients of acetonitrile (solvent **B**) in water (solvent **A**) and supplemented with 0.1 % TFA (Uvasol® TFA, *Millipore Corporation*)

Analytical RP-HPLC

LaChrom RP-HPLC system (*MERCK HITACHI*) equipped with a D-7000 interface, an L-7200 autosampler, two L-7100 pumps, and a diode array detector.

Elite Lachrom RP-HPLC system (*VWR & HITACHI*) equipped with an organizer, a L-2200 autosampler, two L-2130 pumps, and a L-2455 diode array detector.

Analytical columns

Capcel pak C18 column, SG 120 5 µm, size: 4.6 mm Φ x 250 mm (*Shisheido*)

Kinetex C18 column, 5 µm, 4.6 mm Φ x 250 mm (*phenomenex*)

Preparative RP-HPLC system

Preparative RP-HPLC was carried out on a RP-HPLC (*Knauer GmbH*) equipped with smartline manager 5000 system, two smartline pumps 1000, and an ultraviolet detector 2500.

Preparative column

Gemini® 10 µm C18 column, 110 Å, AXIA™ packed Ea, 250 x 21.2 mm, (*phenomenex*), with SecurityGuard PREP Cartridge C18 15 x 21.2 mm ID (*phenomenex*).

Composition of cleavage cocktail

Cocktail **A**: 82% TFA, 3% DTT, 5% H₂O, 5% TIPS and 5% thioanisol

Cocktail **B**: 89% TFA, 2% phenol, 4% H₂O, 1% TIPS, 2% thioanisol, 2% EDT

5.4 Methods and experimental produces

250 μ L of cleavage cocktail was used for test cleavage

5 mL of cleavage cocktail was used for cleavage of 0.5 g resin

Instruments used for peptide synthesis and protein chemistry

SyroXP-I peptide synthesizer (*MultiSynTech GmbH*)

Activo P11 Synthesizer (*Activotec*)

Jasco J-715 Circular Dichroism Spectropolarimeter equipped with Jasco PTC-348WI peltier thermostat and software (*Jasco*)

MicroCal iTC200 system (*GE Healthcare*), equipped with software

NanoDrop ND-2000 (*Thermo SCIENTIFIC*), equipped with software

Äkta Explorer (*GE Healthcare*)

Cartesian Dispensing System (*GENOMIC Solution®*, a *Harvard Bioscience Company*) equipped with an EYELA CCA-1110 low temperature circulator (*EYELA*)

MICROSYS Liquid Handling (*DIGILAB®*) equipped with an EYELA CCA-111a low temperature circulator (*EYELA*)

BioPhotometer Plus and Eppendorf UVette® (*Eppendorf*)

C. B. S. SCIENTIFIC DCX-700 system (*C. B. S. SCIENTIFIC*)

MilliQ-AdvantageA10-System (*Millipore*)

5.4.2 Synthesis of Fmoc-(S)-2-amino-4,4,4-trifluorobutyric acid (TfeGly)

Synthesis of diethyl- α -(N-acetylamino)- α -(2,2,2)-trifluoroethyl malonate

8.0 g (0.037 mol) of diethyl acetamidomalonate (>99%, *Acros Organics*) and 4.1 g (0.037 mol) of potassium *tert*-butoxide (>98%, *Acros Organics*) were dissolved in 100 mL anhydrous THF in a 250 mL flask fitted with a reflux condenser. The mixture was stirred under argon at 60 °C for 2 hours. The mixture was cooled to room temperature and 6 mL (1.1 equivalent) of 2,2,2-trifluoroethyl trifluoromethanesulfonate (97%, *ABCR GmbH & Co. KG*) was added, and the mixture

5.4 Methods and experimental produces

further stirred at 70 °C for two days. The THF was removed by rotary evaporation and the remaining mixture dissolved in 200 mL 1 M HCl and extracted 4 times with 100 mL ethyl acetate. The combined organic phase was washed twice with water, dried with MgSO₄, and concentrated under reduced pressure. The crude product was further purified by silica gel chromatography (column size 25 x 4 cm, packed in 3:1 *n*-hexane:ethyl acetate). Fractions were analyzed by means of TLC (KMnO₄ staining), and appropriate ones were pooled together and concentrated under reduced pressure to give a pale yellow solid. Recrystallization was carried out in 2:1 diethyl ether:*n*-hexane at -20 °C overnight. The resulting white needles were carefully washed several time with *n*-hexane and dried under vacuum to give 3.6 g final product (32%). ¹H NMR (500 MHz, chloroform-d₁) δ: 1.25 (t, 6H, *J* = 10 Hz), 2.04 (s, 2H), 3.34 (dd, 2H, *J* = 10 Hz), 4.26 (dd, 4H, *J* = 10 Hz); ¹⁹F NMR (500 MHz, chloroform-d₁) δ: -62.01 (t, 3F, *J* = 10 Hz).

Synthesis of (*R,S*)-2-amino-4,4,4-trifluorobutyric acid

3.6 g diethyl- α -(*N*-acetylamino)- α -(2,2,2)-trifluoroethyl malonate was dissolved in 30 mL concentrated HCl in a 250 mL flask equipped with a reflux condenser. The solution was stirred at 80 °C overnight. The reaction mixture was concentrated under vacuum, and the product lyophilized. ¹H NMR (500 MHz, D₂O) δ: 3.05-3.45 (m, 2H), 4.54-4.72 (m, 1H), 4.94-5.26 (m, 5H); ¹⁹F NMR (500 MHz, D₂O) δ: -63.58 (t, 3F, *J* = 10 Hz)

Synthesis of (*R,S*)-2-aminoacetyl-4,4,4-trifluorobutyric acid

(*R,S*)-2-amino-4,4,4-trifluorobutyric acid was dissolved in 8 mL of water (pH = 0.2) and the pH adjusted to 9 with 1M NaOH and 1M HCl. The solution was kept cool on ice and 2 mL freshly distilled acetic anhydride was added, upon which the pH dropped to about 4.5. The pH of the solution was readjusted to 8.2. The reaction mixture was stirred on ice for 30 minutes and then at room temperature for a further 2 hours. Afterward, the reaction mixture was acidified with concentrated HCl to pH 2 and extracted five times with ethyl acetate. The combined organic phase was dried

5.4 Methods and experimental produces

with MgSO_4 and concentrated under vacuum to give 1.32 g (6.62 mmol, 17.9%) of product as a white solid. ^1H NMR (500 MHz, acetone- d_6) δ : 2.00 (s, 3H), 2.62-2.92 (m, 2H), 4.64-4.72 (dd, 1H, $J = 5$ Hz); ^{19}F NMR (500 MHz, acetone- d_6) δ : -64.20 (t, 3F, $J = 10$ Hz).

Synthesis of enantiomerically pure (S)-2-amino-4,4,4-trifluorobutyric acid

(*R,S*)-2-aminoacetyl-4,4,4-trifluorobutyric acid was dissolved in water to give a solution with a final concentration of 50 mM. The pH of the solution was first adjusted to 11 with 1 M NaOH, and then adjusted with 10% AcOH to 7.0. Acylase I (from porcine kidney, grade II, *Sigma-Aldrich*) was added (for each 100 mg amino acid, 5 mg enzyme). The pH of the solution was adjusted again to 7 with 0.1% AcOH and 0.01 M NaOH. The solution was stirred at 37 °C overnight. The enzymatic reaction was quenched by adjusting the pH to 4 with 10% AcOH. Ion-exchange beads (8 g, Dowex 50 W x 8, 100-200 mesh, *Acros Organics*) were added and the mixture stirred for at least 4 hours. Afterward, the resin was packed into a small column and washed with water until a neutral pH was reached. The amino acid was eluted with 1 M ammonium hydroxide solution. Appropriate fractions, which were determined by a ninhydrin test, were pooled together and lyophilized to give 361.2 mg (2.3 mmol, 6%) of the product as a white powder. ^1H NMR (500 MHz, D_2O) δ : 2.68-3.02 (m, 2H), 3.97 (dd, 1H, $J = 5\text{Hz}$, $J = 10\text{Hz}$), 4.64-4.92 (m, 2H); ^{19}F NMR (500 MHz, D_2O) δ : -63.96 (t, 3F, $J = 10$ Hz).

9-fluorenylmethoxycarbonyl (S)-2-amino-4,4,4-trifluorobutyric acid

(*S*)-2-amino-4,4,4-trifluorobutyric acid was dissolved in 10% Na_2CO_3 solution and incubated on ice. Dioxane was added (1 mL per 100 mg amino acid), and the mixture stirred in an ice bath while 1.1 equivalents of 9-fluorenylmethyl *N*-succinimidyl carbonate (*Acros Organics*) was added. The mixture was initially stirred for 2 hours on ice, then overnight at room temperature. Afterward, the mixture was diluted with water and extracted once with diethyl ether. The aqueous solution was acidified with concentrated HCl and extracted three times with DCM. The combined

5.4 Methods and experimental produces

organic layer was washed once with water, dried over MgSO_4 , and concentrated under reduced pressure. The product was purified by RP-HPLC to give a white solid. ^1H NMR (500 MHz, D_2O) δ : 4.22 (t, 1H, $J = 5$ Hz), 4.36-4.48 (m, 2H), 5.61 (d, 1H, $J = 5$ Hz), 7.35 (dt, 4H, $J = 10$ Hz $J = 5$ Hz), 7.56 (td, 2H, $J = 5$ Hz), 7.75 (d, 2H, $J = 5$ Hz); ^{19}F NMR (500 MHz, D_2O) δ : -62.99 (t, 3F, $J = 10$ Hz); mass spectrum (ESI-TOF), m/z 378.91 (M-H)⁻ ($\text{C}_{19}\text{H}_{16}\text{F}_3\text{NO}_4$, calculated m/z 379.10).

5.4.3 Peptide synthesis and purification

5.4.3.1 Synthesis of BPTI C-terminal fragment

The C-terminal peptide segment of BPTI, amino acid residues 38-58, was synthesized on preloaded Fmoc-Ala-NovaSyn®TGT resin (0.2 mmol/g, Novabiochem) by using a SyroXP-I peptide synthesizer at a 0.05 mmol scale according to standard Fmoc/*t*Bu chemistry. For standard couplings a four-fold excess of amino acids and coupling reagents (TBTU/HOBt), as well as an eight-fold excess of DIPEA relative to resin loading, was used. All couplings were performed as double couplings for 30 minutes each. Cleavage and removal of the side chain protecting groups was carried out by treatment with 4 mL of cleavage cocktail A. The peptides were precipitated with cold diethylether. The peptide was purified by RP-HPLC using a gradient of 0-80% MeCN in water (with supplement of 0.1% TFA) over a period of 30 minutes at flow rate of 20 mL/minute (monitoring at 220 nm).

5.4.3.2 Synthesis of BPTI N-terminal fragment; preparation of α -thioester

The peptide fragment of BPTI, residues 16-37, was first synthesized on a preloaded H-Gly-2-Chlorotrityl resin (~0.69 mmol/g, Novabiochem) by using a SyroXP-I peptide synthesizer at a 0.1 mmol scale according to standard Fmoc/*t*Bu chemistry. For standard couplings, a four-fold excess of amino acids and coupling reagents (TBTU/HOBt), as well as an eight-fold excess of DIPEA relative to resin loading,

5.4 Methods and experimental produces

were used. All couplings were performed as double couplings for 30 minutes each. Product was analyzed by test cleavage using cleavage cocktail **B**. The precipitated peptide was subsequently analyzed by RP-HPLC (0-80% B over 30 min) and ESI-MS.

Afterward, the Fmoc-protected non-canonical amino acid was incorporated into the peptide by means of hand coupling following the standard HOBt/DIC protocol. 1.25 equivalent of Fmoc-L-amino acid/HOBt/DIC relative to resin loading was used and incubated overnight. The efficiency of manual coupling was analyzed by test cleavage, RP-HPLC, and ESI-MS. If the first coupling was not complete, a second coupling was performed. Elongation of the peptide fragments that contained Fmoc-Abu and Fmoc-DfeGly were carried out on the Activo P11 Synthesizer. 10 equivalents of Fmoc-L-amino acids/HOBt/DIC were used. Because the sequence of BPTI is extremely difficult, the duration and number of coupling were set up according to an optimized protocol (Table 5.4-1).¹⁹² The sequence elongation of BPTI 15-37 containing Fmoc-TfeGly was carried out on a SyroXP-I peptide synthesizer. The coupling conditions were as described above but with double coupling of 60 minutes each. Boc-Arg (Pbf) (*Sigma-Aldrich*) was used for coupling of the last residue.

Table 5.4-1 Protocol for elongation of peptide fragment BPTI 1-*n*AA15-37

Cycle No.	Residue ^a	Coupling duration (min)	Cycle No.	Residue ^a	Coupling duration (min)
0	BPTI 15-37	-	12		72
1	C14	90	13	L6	72
2	P13	90	14		72
3	G12	90	15	C5	120
4	T11	90	16	F4	120
5		72	17	D3	120
6	Y10	72	18	P2	72
7	P9	72	19		72
8		72	20	R1 ^b	240
9	P8	72	21		72
10		72	22		72
11	E7	72			72

a) Amino acid in 1-letter abbreviation, b) Boc-Arg (Pbf) (*Sigma-Aldrich*) was used for coupling of last residue.

5.4 Methods and experimental produces

The peptide fragment was first analyzed by test cleavage with cleavage cocktail **B**. The precipitated peptide was analyzed by means of RP-HPLC and MS. The successfully assembled fully protected peptide was cleaved from 2-Cl-Trt resin by treating the resin with a solution containing acetic acid/TFE/DCM (1:1:8 v/v/v). In order to remove the acetic acid, cleavage solution was extracted with half-saturated NaHCO₃ three times. The DCM phase containing the fully protected peptide fragment was pooled together and concentrated under vacuum. A small amount of oil-like residue was taken for side chain deprotection using cleavage cocktail **B** and analyzed by HPLC/ESI-MS. The fully protected peptide fragment was dissolved in about 5 mL DCM (for ~0.1 mmol) and treated with 5 equivalent of *p*-acetamidothiophenol (98%, *Alfa Aesar*), DIEPA (*Acros*), and benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP, *Novabiochem*) and stirred overnight. The thioesterification reaction was monitored by RP-HPLC and MS (side chain protecting groups were removed by the use of cleavage cocktail **B**). Once the reaction was complete, the organic solvent was evaporated. Side chain protecting groups were removed by treatment with cleavage cocktail **B**. The BPTI 1-*nAA15-37-α*-thioester is poorly soluble in H₂O/MeCN. A suitable amount of dimethyl sulfoxide (DMSO) was added to completely dissolve the thioester in H₂O/MeCN for the purification. Purification was performed on a RP-HPLC (C18 column) using a gradient of 5-70% MeCN in water at a flow rate of 20 mL/minute over a period of 30 minutes (monitored at 260 nm).

5.4.4 Native chemical ligation

4-mercaptophenylacetic acid (MPAA, 97% *Alfa Aesar*), and tris-(2-carboxyethyl)phosphine hydrochloride (TECP, *Carl Roth*) were added into 5 mL of freshly prepared NCL buffer: 50 mM MPAA and 20 mM TECP containing 6 M GdmCl and 0.1 M Na₂HPO₄. This NCL buffer was degassed with argon for 2 minutes. The pH of this buffer was adjusted to 7.0 with 2 M NaOH solution (pH

5.4 Methods and experimental produces

meter). C-terminal and N-terminal fragments of BPTI were added to this solution to give a final concentration of peptides between 1-5 mM. The pH of the reaction mixture was adjusted to 7.0 if necessary with diluted NaOH solution (0.2 M). The reaction mixture was incubated and stirred at room temperature. Aliquots of 10 μ L from the reaction were taken at time points 0 h, 1 h, 2 h, 4 h, 8 h, 16 h and diluted to 100 μ L with MeCN/H₂O (1/1, v/v, 0.1 % TFA) for analysis by RP-HPLC and ESI-MS. After the formation of product halted, the NCL reaction mixture was directly loaded onto a PD-10 desalting column (Sephadex G-25 M, *GE Healthcare*) to remove small compounds such as MPAA, which give a huge peak in the RP-HPLC chromatogram and hinder downstream purification. Afterward, full-length peptides were purified by RP-HPLC on a C 18 column using a gradient of 5-70% MeCN in water at a flow rate of 20 mL/minute over a period of 30 minutes (monitored at 260 nm).

5.4.5 Refolding of synthetic mutant BPTI variants

Reduced full-length BPTI variants were dissolved in folding buffer containing 0.6 M Tris-HCl, 6 mM EDTA and 6 M GdmCl at pH 8.7 to a final protein concentration of 1 mg/mL. This solution was diluted 6-fold with water all at once and stirred in an open-air container for 1-2 days. Aliquots of 10 μ L were taken at certain time points and diluted to 100 μ L with MeCN/H₂O (1/1, v/v, 0.1 % TFA) for HPLC/ESI-MS analysis. Refolding was generally complete after two days. Folded BPTI mutants were purified by RP-HPLC on a C 18 column using a gradient of 5-70% MeCN in water at a flow rate of 20 mL/min over a period of 30 minutes (monitored at 260 nm).

5.4.6 Determination of protein concentration

BPTI variants were dissolved in the appropriate buffer for CD measurements or enzyme inhibition assays and dialyzed overnight against the same buffer. α -chymotrypsin and β -trypsin were also dissolved in the appropriate buffer. Protein concentration was determined with a Nanodrop 2000 (*Thermo*) by entering the corresponding molecular weights and extinction coefficients. Extinction coefficients were calculated using the online ProtParam tool into which the protein sequence was entered (<http://web.expasy.org/protparam/>, ExPASy Bioinformatics Resource Portal). ϵ (BPTI) = 6335 M⁻¹cm⁻¹, ϵ (α -chymotrypsin) = 50585 M⁻¹cm⁻¹, ϵ (β -trypsin) = 37650 M⁻¹cm⁻¹.

5.4.7 Protein structural analysis and thermal stability analysis by circular dichroism

5.4.7.1 Basis of circular dichroism spectroscopy

Circular dichroism (CD) spectroscopy is a valuable technique that is widely used for determining protein secondary structure elements, as well as studying the folding and binding properties of proteins. In order to understand the principle of CD, some physical background is first to be contemplated. Electromagnetic radiation consists of a magnetic and an electric field, which oscillate perpendicular to each other and to the propagating direction. Linearly polarized light refers to the case in which the electric field vector oscillates only in one plane. In CD spectroscopy, light is first polarized by passage through suitable prisms and filters. This plane-polarized light is further split into two circularly polarized components.¹⁹³ The differential absorption of these circularly polarized light components is measured. When this light source interacts with asymmetric molecules, the CD signal gives the difference in absorbance of the left-handed and right-handed components as $\Delta A = A_L - A_R$ or as the ellipticity (θ) in degrees ($\tan \theta = (E_R - E_L) / (E_R + E_L)$).

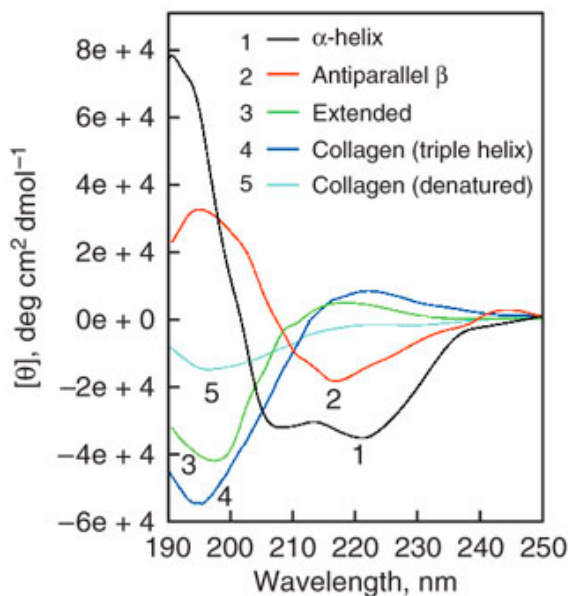


Figure 5.4-1 CD spectra of polypeptides and proteins showing secondary structures

1) α -helix (spectra of poly-L-lysine at pH 11.1), 2) antiparallel β -sheet, 3) extended conformation, 4) native triple-helical structure, 5) denatured collagen structure. (Adapted with permission from Greenfield *et al.*,^{193a} Copyright ©2007, Nature Publishing Group)

In protein structure analysis, the most important chromophore for CD spectroscopy is the amide group. In the far ultraviolet region, the spectrum of the protein is dominated by the dipolar transition of the amide group, which is influenced by the geometry of the protein backbone.¹⁹⁴ Figure 4.5-1 presents typical CD spectra indicating certain secondary structure elements in proteins.

5.4.7.2 Structural analysis of BPTI variants by use of CD

CD spectra were recorded on a Jasco J-715 spectropolarimeter at 25 °C. Overall peptide concentrations were 20 μ M at pH 7.4 (10 mM Tris buffer). CD-spectra were obtained in the far-UV range (190–240 nm) using an 0.1 cm Quartz Suprasil cuvette (*Hellma*) equipped with a stopper. The nitrogen flow rate was set to 3.0 L/min. The spectra were recorded using 0.2 nm intervals, a 2 nm bandwidth, and a response times of 3 per sec. After baseline correction, the measured ellipticities were converted to molar ellipticities per residue ($\text{degree cm}^2 \text{dmol}^{-1} \text{residue}^{-1}$) by normalizing for the concentration of the peptide, the number of amino acids in the sequence, and the path length.

5.4 Methods and experimental produces

5.4.7.3 Thermal stability analysis of BPTI variants

The thermal stability of BPTI variants was determined by means of CD spectroscopic measurements. Protein samples were dissolved in buffer either containing 10 mM Tris-HCL and 6 M GdmHCl at pH 2 or 10 mM Tris-HCL and 8 M urea at pH 2. Melting curves were recorded using the signal at 222 nm applying a heating rate of 3 K/min from 20 °C to 100 °C. Each sample was measured independently three times and the melting curves were averaged.

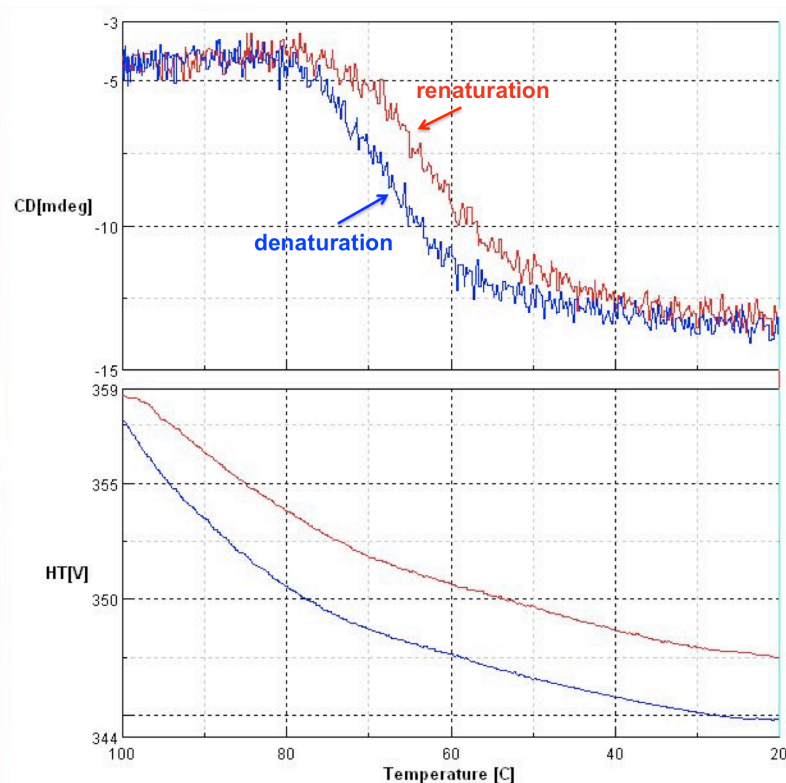


Figure 5.4-2 Schematic representation of a reversible denaturation of BPTI

Denaturation and renaturation curves were recorded by the use of a CD-spectroscopy, here shows BPTI_{Lys15TfeGly} in the presence of GdmCl as example.

Thermodynamic parameters were determined by non-linear least square fitting of the normalized CD melting curves with six parameters (a , b , ΔH_m , T_m , $[\theta]_F$, and $[\theta]_U$) as a two-state folding-unfolding equilibrium.⁶³ The fits were performed in Microsoft Excel as follows:¹⁹⁵

At any temperature, T , the folding constant, K , is:

5.4 Methods and experimental produces

$$K = [F] / [U] \quad (1)$$

[F] and [U] are the concentration of the folded and unfolded proteins, respectively.

The free energy of folding is

$$\Delta G = -RT \ln K \quad (2)$$

R is the Boltzmann Gas constant = 1.986 cal mol⁻¹K⁻¹, T is the absolute temperature (Kelvin, K).

The fraction unfolded at any temperature is f_U .

$$f_U = [U] / ([F] + [U]) \quad (3)$$

$$f_U = ([\theta]_o - [\theta]_F) / ([\theta]_F + ([\theta]_U)) \quad (4)$$

Where $[\theta]_o$ is the observed ellipticity at any given time, $[\theta]_U$ is the linear temperature dependence of the ellipticity of the fully unfolded proteins, $[\theta]_F$ is the linear temperature dependence of the ellipticity of the fully folded proteins:

$$[\theta]_o = f_U \times ([\theta]_F + [\theta]_U) + [\theta]_F \quad (5)$$

$$[\theta]_F = a \times T + [\theta]_F \text{ (at 0 K)} \quad (6)$$

$$[\theta]_U = b \times T + [\theta]_U \text{ (at 0 K)} \quad (7)$$

To fit the change in intensity at 222 nm in the CD spectra as a function of temperature, T, the Gibbs-Helmholtz equation is used to describe folding as a function of temperature. The following equations are appropriate for the unfolding of a monomer:

$$\Delta G = \Delta H (1 - T / T_M) - \Delta C_p ((T_M - T) + T \ln (T / T_M)) \quad (8)$$

$$K = \exp (-\Delta G / (RT)) \quad (9)$$

$$f_U = K / (1 + K) \quad (10)$$

5.4 Methods and experimental produces

T_M (melting temperature, also referred to as T_{den} , denature temperature) is the temperature where $f_U = 0.5$ and ΔC_p is the change in heat capacity in going from a folded to an unfolded state. ΔC_p is usually assumed to be zero for initial calculations of the thermodynamics of the folding of monomeric proteins. Equations (5) – (10) were combined and the data fitted directly.

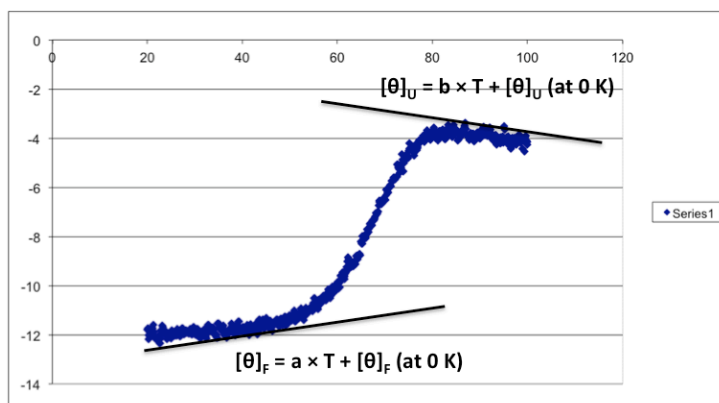


Figure 5.4-3 Schematic representation of a melting curve showing the linear baseline for folded and denatured state of protein

The linear temperature dependence of ellipticity is according to equations (6) and (7).

Table 5.4-2 Fitting parameters for thermodynamic characterization

	BPTI Variant	<i>a</i>	<i>b</i>	$[\theta]_F$	$[\theta]_U$	T_m (0°)	ΔH_m (KJ/mol)	ΔG°
6 M GdmCl	BPTI _{Lys15wild-type}	0.002	-0.007	-6.529	-1.625	66.48	191.92	11.99
	BPTI _{Lys15Abu}	0.002	-0.012	-6.272	-1.129	63.44	177.94	11.26
	BPTI _{Lys15DfeGly}	0.009	-0.028	-8.634	-0.208	68.77	196.47	12.42
	BPTI _{Lys15TfeGly15}	0.020	-0.040	-12.545	-0.248	68.27	193.22	12.05
8 M urea	BPTI _{Lys15wild-type}	0.010	-0.002	-6.595	-2.083	75.78	236.11	14.62
	BPTI _{Lys15Abu}	0.013	-0.015	-8.862	-1.664	71.68	228.00	13.78
	BPTI _{LysDfeGly}	0.011	-0.001	-8.238	-2.841	74.75	229.35	14.25
	BPTI _{Lys15TfeGly}	0.011	-0.023	-7.691	-0.707	73.80	244.97	14.64

a) in $10^3 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ residue}^{-1}$, b) T_m is defined as the temperature at which the fraction of folded is 0.5. Errors are typically not higher than 0.5 °C, c) Errors are typically not higher than 1 kcal/mol, d) ΔG° values were calculated for standard conditions of a 1 M solution at 25 °C, errors are not higher than 0.2 kcal/mol.

BPTI undergoes unfolding with a very small heat capacity change. An approximate value for ΔC_p was calculated from a plot of the ΔH_m against the melting temperature T_m (Van't hoff plot). The slope of the plot was calculated to be $3.29 \text{ kcal mol}^{-1}\text{K}^{-1}$ for

5.4 Methods and experimental produces

denaturation in the presence of 6 M GdmCl, and 1.31 kcal mol⁻¹K⁻¹ for denaturation in the presence of 8 M urea, respectively. These values are in agreement with the range reported in the literature with a ΔC_p of 3.0 at 25 °C, 2.6 at 50 °C, 2.1 at 75 °C, and 1.3 at 100 °C.⁶³

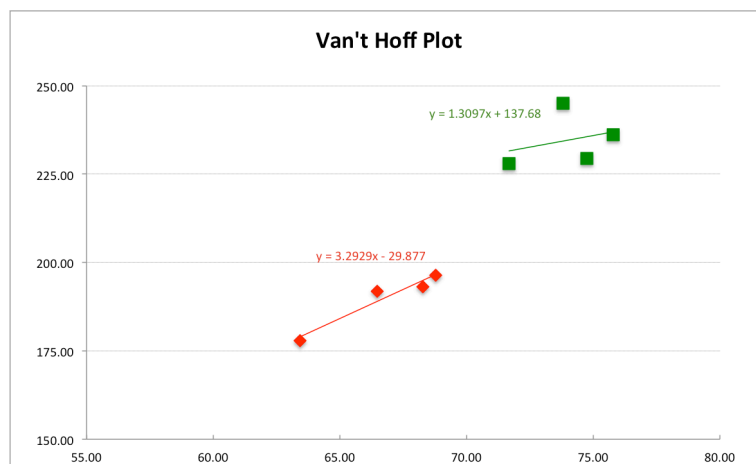


Figure 5.4-4 Van't Hoff plot of ΔH_m against T_m for BPTI variants

The slope of which yields ΔC_p denaturation in the presence of 6 M GdmCl (red), denaturation in the presence of 8 M urea (green).

5.4.8 Binding studies of BPTI variants with proteases

5.4.8.1 Competitive inhibition assay

The inhibitory activities of synthetic BPTI variants and wild-type BPTI (*Sigma-Aldrich*) were tested by means of a well-established enzymatic assay relying on continuous spectrophotometric determination. To optimize enzymatic activity, the appropriate amounts of enzyme and BPTI were first determined. A series of mixtures in which the protease concentration was held constant and the inhibitor with varied concentration were prepared and well mixed. After incubation for 45 min, the enzyme's substrate was added and the reaction quickly mixed. Absorbance was recorded on a photometer (*BioPhotometer Eppendorf*) every 15 seconds over four minutes. The enzymes, assay buffers, substrates, and absorption coefficients are

5.4 Methods and experimental produces

listed below (based on the protocol of Sigma-Aldrich and the Worthington Enzyme Online Manual):

α -chymotrypsin (Cat. 27272, Lot. S43979058, Fluka) assay:

Assay buffer: 0.08 M Tris-HCl and 0.1 M CaCl₂, pH 7.8

Substrate: *N*-Benzoyl-L-tyrosine ethyl ester (BTEE, *Sigma-Aldrich*) was dissolved in a solution containing 50% methanol in water (v/v) to a final concentration of 1 mM

Wavelength of detection: 260 nm

Absorption coefficients: $\epsilon = 964 \text{ M}^{-1}\text{cm}^{-1}$ ¹⁹⁶

Plasmin from human plasma (*Sigma-Aldrich*) assay¹⁹⁷:

Assay buffer: 0.05 M Tris-HCl, pH 7.4

Substrate: D-Ile-Phe-Lys *p*-nitroanilide (*Sigma-Aldrich*) was dissolved in a solution of 10% methanol in H₂O (v/v) to a final concentration of 0.5 mM

Wavelength of detection: 405 nm

Absorption coefficient: $\epsilon = 10600 \text{ M}^{-1}\text{cm}^{-1}$ for plasmin assay

β -trypsin (DPCC-behandelt, Cat. 93611, Lot. S39283258, Fluka) assay:

Assay buffer: 0.046 M Tris-HCl and 0.0115 M CaCl₂, pH 8.1

Substrate: *N* _{α} -Benzoyl-L-arginine 4-nitroanilide hydrochloride (*Sigma-Aldrich*) was dissolved in a solution of 1:1 (v/v) acetone and water to a final concentration of 1 mM

Wavelength of detection: 260 nm

Absorption coefficient: $\epsilon = 8800 \text{ M}^{-1}\text{cm}^{-1}$ ¹⁹⁸

The residual activity of the enzyme was calculated as described below.

5.4 Methods and experimental produces

According to the Beer-Lambert law: $A = \epsilon * l * c$ (1)

A refers to absorbance, ϵ is the absorption coefficient, c is the molar concentration, l is the path length. The equation (1) can be written as

$$c = A / (\epsilon * l) \quad (2)$$

$$\Delta c = \Delta A / (\epsilon * l) \quad (3)$$

In an enzymatic reaction, given that the reaction is independent of the concentration of substrate and product, the velocity (V) is

$$V = \pm dc / dt \quad (4)$$

Combining equations (3) and (4), the velocity of an enzymatic reaction can be written:

$$\Delta c / \Delta t = \Delta A / (\epsilon * l * \Delta t) \quad (5)$$

Equation (5) can be written as

residual activity in terms of volume

$$\Delta c / \Delta t = \Delta A * V / (\epsilon * l * \Delta t * v) \text{ (U/mL)} \quad (6)$$

Where V is the total volume of the enzymatic reaction and v is the volume of enzyme.

In order to calculate the association constant, the experimental data were fitted to the following equation:⁶⁹

$$[E] = 0.5 \times \left\{ [E_0] - F \times [I_0] - K_a^{-1} + \sqrt{([E_0] + F \times [I_0] + K_a^{-1})^2 - 4 \times [E_0] \times F \times [I_0]} \right\} \quad (7)$$

Where $[E_0]$ and $[I_0]$ are the final concentrations of enzyme and inhibitor, respectively, $[E]$ is the residual concentration of the enzyme, and F is the enzyme-inhibitor equimolarity factor ($F = 1$).

5.4 Methods and experimental produces

5.4.8.2 Binding studies by means of isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is a powerful experimental method to investigate protein-ligand and protein-protein interactions. Among the methods that can be used to determine the parameters of molecular interactions, ITC is the only one that enables direct determination of the thermodynamic parameters ΔG , ΔH , ΔS , and ΔC_p .

ITC makes use of two identical lollipop-shaped cells, which are contained within an adiabatic jacket that is made of a material that is a highly efficient thermal conductor, yet chemically inert. One of the cells is a sample cell used for titration of the interaction partner, the other cell acts as a reference cell. Low power is applied to continuously heat the reference cell. As long as the interactions take place in the sample cell, the thermopile/thermocouple detectors sense the difference in temperature between the sample and reference cells. Depending on the type of interaction, whether it is exothermic or endothermic, a feedback circuit will reduce or increase the power supplied to the sample cell to maintain an equal temperature with the reference cell. The heat per certain unit of time provided by the sample cell is directly measured (Figure 5.4-5).^{190, 199}

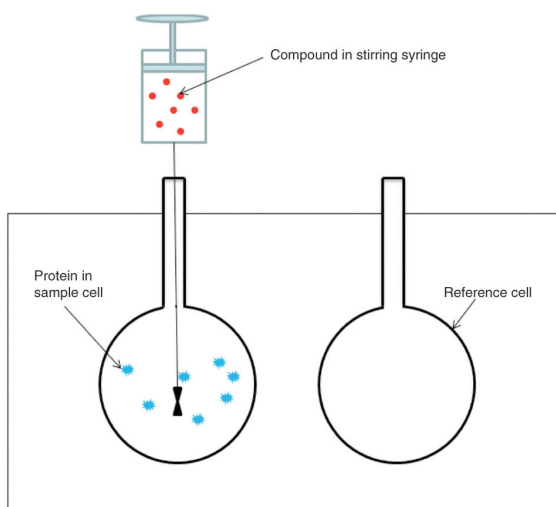


Figure 5.4-5 Schematic representation of the ITC set up

(Adapted with permission from Zhou *et al.*,²⁰⁰ copyright © 2011, Nature Publishing Group)

For our binding studies, the proteases and wild-type BPTI were extensively dialyzed

5.4 Methods and experimental produces

against a buffer containing 25 mM Tris-HCl and 10 mM CaCl₂ at pH 7.5 for three days (dialysis buffer changed for twice). Samples at the appropriate concentration were introduced into the sample cell and filled in stirring syringe for titrant. In general, the concentration of titrants was ten-fold higher than that of the sample. The ITC instrument was operated according to the manual, set up for 20 injections of 2 μ L titrant each, except the first injection of 0.2 μ L, conducted as a test.

5.4.9 Protein crystallographic analysis

5.4.9.1 Purification of proteins and protein complexes for crystallization

The synthetic BPTI variants were purified by dialysis for 2-3 days with buffer change (3 kD standard grade regenerated cellulose dialysis membrane, *Spectra/Por*[®]). Purification of the α -chymotrypsin-BPTI_{Lys15wild-type} and β -trypsin-BPTI_{Lys15wild-type} complexes was performed by means of size-exclusion (gel filtration) chromatography with a Superdex 75 10/300 GL column (*GE Healthcare*) in buffer containing 10 mM Tris and 200 mM NaCl at pH 7.4. Appropriate fractions were analyzed by means of SDS-PAGE. Protein solutions were concentrated by using Amicon[®] Ultra-4 and Amicon[®] centrifugal filter devices (3 kD, *Millipore*) for large volumes, and Ultra-0.5 devices for small volumes.

5.4.9.2 Crystallization of protease-BPTI complexes

Crystallization screening was performed by means of a sitting drop vapor diffusion method using the AmSO₄ Suite Kit (*Qiagen*). Drops of 100 nL protein solution and a 100 nL reservoir were dispensed by a robot with eight channels into 96-well crystal screening plates (*SWISSCI*) and incubated at room temperature. Purification methods, ratios of protease to BPTI, and final concentrations of complexes used for crystal screening are summarized in *Table 5.4-3*.

5.4 Methods and experimental produces

Table 5.4-3 Preparation of crystallization screening for protease-BPTI complexes

Protein complex	Concentration (mg/mL)	Ratio of protease/BPTI	Buffer	Purified method	Robot
β-trypsin-BPTI					
Wild-type	56.0	1: 1.1			1
Lys15Abu	25.6	a	25 mM Tris, 10 mM CaCl ₂	Dialysis ^b	2
Lys15DfeGly	54.6	a			2
Lys15TfeGly	31.4	1: 1.1			2
α-chymo-trypsin-BPTI (wild-type)					
	57.7	1: 1.1	25 mM Tris, 10 mM CaCl ₂	Dialysis ^b	1
	29.7	1: 1.1	25 mM Tris, 10 mM NaCl	Gel-filtration	2
	26.9	1: 1.1	25 mM Tris, 200 mM NaCl	Gel-filtration	2

1. Cartesian Dispensing System (*GENOMIC Solution*[®], a *Harvard Bioscience Company*) equipped with an EYELA CCA-1110 low temperature circulator (*EYELA*). 2. MICROSYS Liquid Handling (*DIGILAB*[®]) equipped with an EYELA CCA-111a low temperature circulator (*EYELA*). a) Samples of trypsin-BPTI (Abu 15) and trypsin-BPTI (DfeGly 15) were taken from ITC measurement, ratio of protease:BPTI were greater than 1. b) Trypsin and chymotrypsin were purchased as lyophilized powders and dissolved in buffer. Synthetic mutant BPTI variants were purified by means of RP-HPLC. Proteases and BPTI variants were dialyzed against buffer.

5.4.10 Data collection, structural refinement, and data analysis

Dr. Bernhard Loll performed data collection, structural refinement, and crystallographic data analysis.

Well-formed crystals were soaked in 30% glycol plus reservoir solution and frozen in liquid nitrogen. Data was collected in Berlin on the BESSYII synchrotron, beamline 14-2. X-ray data collection was performed at 100 K. Diffraction data were processed with the XDS in space group *I*222.²⁰¹ The structure of the wild-type trypsin/BPTI complex was solved by molecular replacement with the coordinates of single polypeptide chains of trypsin and BPTI (PDB: 2FTL; ²⁰²) as search models using PHASER.²⁰³ The structure of protein complexes were determined by isomorphous replacement. For calculation of the free R-factor, a randomly generated set of 5 % of the reflections from the diffraction data set was used and excluded from the refinement. The structure was initially refined by applying a simulated annealing protocol and in later refinement cycles by maximum-likelihood restrained refinement using PHENIX²⁰⁴ followed by iterative, manual model building cycles with

5.4 Methods and experimental produces

COOT²⁰⁵. Model quality was evaluated with MolProbity²⁰⁶. Figures were prepared using PyMOL²⁰⁷.

Table 5.4-4 Data collection and refinement statistics¹

Data collection	Trypsin/BPTI wt	Trypsin/BPT Abu	Trypsin/BPTI DfeGly	Trypsin/BPTI TfeGly
PDB entry ²	XXXX	XXXX	XXXX	XXXX
Space group	<i>I</i> 222	<i>I</i> 222	<i>I</i> 222	<i>I</i> 222
Wavelength [Å]	0.91841	0.91841	0.91841	0.91841
Unit cell a; b; c [Å]	74.8; 81.5; 124.2	74.9; 82.3; 123.4	74.7; 81.9; 123.6	75.0; 82.1; 123.6
α; β; γ [°]	90.0; 90.0; 90.0	90.0; 90.0; 90.0	90.0; 90.0; 90.0	90.0; 90.0; 90.0
Resolution [Å] ^a	50.00 - 1.25 (1.28 - 1.25)	50.00 - 1.37 (1.45 - 1.37)	50.00 - 1.37 (1.48 - 1.37)	50.00 - 1.30 (1.38 - 1.30)
Unique reflections	104119 (7297)	79271 (12329)	79871 (12620)	92992 (14647)
Completeness ^a	99.4 (94.8)	98.4 (95.6)	99.6 (98.6)	99.3 (97.9)
<I/σ(I)> ^a	18.7 (1.7)	16.4 (1.9)	19.2 (2.4)	16.81 (1.9)
R _{meas} ^{a, b}	0.043 (0.880)	0.067 (1.118)	0.072 (1.004)	0.052 (1.920)
CC1/2 ^a	100.0 (66.2)	99.9 (69.6)	100.0 (81.8)	99.9 (74.4)
Redundancy ^a	4.8 (3.3)	5.9 (5.9)	8.0 (8.0)	4.4 (4.1)
Refinement				
Non-hydrogen atoms	2698	2669	2602	2605
R _{work} ^{a, c}	0.139 (0.304)	0.133 (0.244)	0.138 (0.229)	0.1334 (0.260)
R _{free} ^{a, d}	0.157 (0.319)	0.160 (0.289)	0.166 (0.261)	0.1565 (0.290)
Average B-factor [Å ²]	20.7	21.3	19.6	21.2
Trypsin residues	243 / 21.6	243 / 21.8	243 / 20.2	243 / 22.5
BPTI residues	58 / 19.2	57 / 18.9	57 / 17.6	57 / 19.4
Water molecules	344 / 31.00	347 / 33.8	346 / 31.2	301 / 32.0
Buffer molecules	10 / 35.00	13 / 33.9	11 / 31.3	12 / 32.5
r.m.s.d. ^e bond le	0.016	0.016	0.018	0.016

5.4 Methods and experimental produces

bond ^a	1.731	1.657	1.756	1.687
Ramachandran outliers [%]	0.0	0.0	0.0	0.0
Ramachandran favoured [%]	98.7	98.4	98.6	98.3

1) Refinement and crystal structure analysis was done by Dr. Bernhard Loll. 2) The PDB codes for protein complexes are available after the submission of coordinates to protein data bank.

^a values in parentheses refer to the highest resolution shell.

^b $R_{\text{meas}} = \frac{\sum_h [n/(n-1)]^{1/2} \sum_i |I_h - I_{h,i}|}{\sum_h \sum_i I_{h,i}}$, where I_h is the mean intensity of symmetry-equivalent reflections and n is the redundancy.

^c $R_{\text{work}} = \frac{\sum_h |F_o - F_c|}{\sum F_o}$ (working set, no σ cut-off applied).

^d R_{free} is the same as R_{cryst} , but calculated on 5% of the data excluded from refinement.

^e Root-mean-square deviation (r.m.s.d.) from target geometries.

Chapter 6

Summary and outlook

6.1 Site-specific incorporation of fluorinated amino acids into BPTI by *amber* suppression

In order to incorporate fluorinated amino acids into proteins to study the fundamental effects of fluorine within a native protein environment, *in vitro* *amber* suppression has been tested for the non-canonical amino acids Abu, DfeGly, and TfeGly. These non-canonical amino acids were successfully chemically misacylated onto yeast suppressor tRNA^{Phe}_{CUA} by means of T4 RNA ligase mediated semi-synthesis. From practical work experience, the misacylation of fluorinated amino acids is more difficult than in the case of Abu; this may be due to the strong electronegativity of fluorine, which could destabilize the aminoacyl bond. Furthermore, the T4 RNA ligation reactions were occasionally incomplete. The stability of misacylated full-length suppressor aa-tRNAs was characterized under different buffer conditions at 37 °C. These results indicate that our misacylated suppressor aa-tRNAs should maintain their integrity under these buffer conditions, at least for a 1-hour incubation time.

We designed expression constructs for BPTI containing an N-terminal SUMO fusion tag to test the incorporation of non-canonical amino acids by the use of misacylated suppressor tRNAs. Commercially available cell-free protein synthesis kits were used. Full-length protein SUMO-BPTI was able to be produced on an analytical scale. Nevertheless, there was very poor evidence for the production of full-length protein containing the non-canonical amino acids. Several potential explanations may exist for the poor expression yields and/or unsuccessful incorporation: 1) although the misacylated suppressor tRNAs are stable under the test buffer conditions, the stability of these tRNAs in commercial cell lysates, which contain numerous cellular components and endogenous factors, could be compromised; 2) suppressor tRNA^{Phe}_{CUA} has proven to be an efficient carrier for non-canonical amino acids for a broad range of substrates; however, the incorporation of DfeGly and TfeGly by means of this suppressor tRNA had not been previously reported.^{119a, 119c, 121, 167}

6 Summary and outlook

Kearney *et al.* reported the successful incorporation of Abu into nicotinic receptor for neuron studies by *amber* suppression through microinjection into oocyte cells.²⁰⁸ The yield of protein production was not reported. And in general in such neuronal studies even limited protein expression is enough to record the neuronal physiological signal; 3) the presence of release factors in the cell-free reaction might decrease the expression yield of full-length protein. Unfortunately, the PURExpress Δ RF123 Kit failed to produce protein, and therefore could not deliver any information regarding the effects of release factors in our studies; 4) as discussed in *section 4.1.1*, though the semi-synthesis of misacylated tRNAs is a crucial step in this method, the misacylated tRNAs may not deliver the non-canonical amino acids into the ribosome, and several critical mechanisms in protein synthesis, such as binding affinities of the aminoacyl-tRNAs to the EF-Tu and the ternary complex formed by aa-tRNA and EF-Tu to the A-site of the ribosome can strongly influence incorporation efficiency.

In general, *in vitro* suppression gave very poor yields of full-length proteins in our hands. This method has been all but replaced by *in vivo* approaches, in which large cell culture volumes enable the production of mutant proteins on a preparative scale. Recently, the misacylation activities of ValRS, LeuRS, and IleRS have been tested for TfeGly and DfeGly in our group. These results indicate that ValRS and LeuRS can activate TfeGly with low synthetase specificity.²⁰⁹ Thus, the possibility to generate novel tRNA/aaRS pairs with specificity for DfeGly and TfeGly should be considered. Moreover, the tRNA^{Pyl}/PylRS from archaea shows very broad specificity toward non-canonical amino acids. A general selection based on Wang *et al.*^{126d, 130} might provide an opportunity to enrich the suppressor efficiency of the tRNA^{Pyl}/PylRS pair for incorporating our fluorinated amino acids.

The fundamental advantages of cell-free protein synthesis are the accessible environment of the reaction and the ability to express proteins that are difficult or impossible to produce in cell hosts; for instance, toxic proteins and membrane proteins on a large scale. Therefore, *in vitro* suppression remains an excellent choice for the production of such proteins containing non-canonical amino acids. The

expression yield of full-length proteins might be improved by following considerations: 1) the release factors 1, 2, and 3 should be eliminated from the cell-free reaction to decrease the competitive effects between these and the suppressor tRNAs; 2) regarding the instability of the energy-rich aminoacyl bond, depletion of the pool of misacylated suppressor tRNAs over time may be another big issue that contributes to poor yields. Because the continuous-exchange cell-free system enables long reaction times (up to 24 hours), we should consider strategies for spiking the reaction with this intermediate periodically. Of course, the non-canonical amino acids must be misacylated onto suppressor tRNAs through more promising methods than the tedious semi-synthesis method. The optimization of the flexizyme system and its combination with cell-free expression system might be considered.

6.2 Total chemical synthesis and full characterization of BPTI containing fluorinated amino acids

Three mutant BPTI variants containing non-canonical amino acids Abu, DfeGly, and TfeGly at a “hot spot” position were synthesized by means of SPPS, native chemical ligation, and renaturation. Thermal stability measurements, inhibitory assays, and biochemical structural analyses were performed to characterize the synthetic BPTI variants and compare their properties to wild type BPTI.

Interestingly, our results show that the incorporation of DfeGly and TfeGly into a solvent exposed position can enhance protein stability of BPTI compared to Abu. This result is actually contrary to the generally accepted view in protein design that fluorination increases protein stability by virtue of its greater hydrophobicity. Because there are not many studies based on the incorporation of fluorine into solvent exposed positions, we cannot draw conclusions regarding whether the phenomena we observed are common properties of fluorine or particular to the case of BPTI.

6 Summary and outlook

We believe that our results indicate the existence of an alternative strategy to enhance protein stability by fluorination at solvent exposed positions, and we propose to continue this line of investigation by broadening our systematic studies on fluorine's effect at solvent exposed positions in proteins. BPTI contains 4 Lys residues, all of which are located in the loop element, not within an α -helix or β -sheet structure. Moreover, the charged side chains of these four Lys residues are oriented into the solvent (*Figure 6.2-1*). Given that the substitution of Lys15 with fluorinated amino acids does not perturb the overall conformation of BPTI, it would be very interesting to explore the substitution effects by replacing Lys with fluorinated amino acids one by one, in a sort of "fluorine scan", and by global substitution. Thus, a series of fluorinated BPTI mutants could be generated and their thermal stabilities, structures, and inhibitor activities determined and compared.

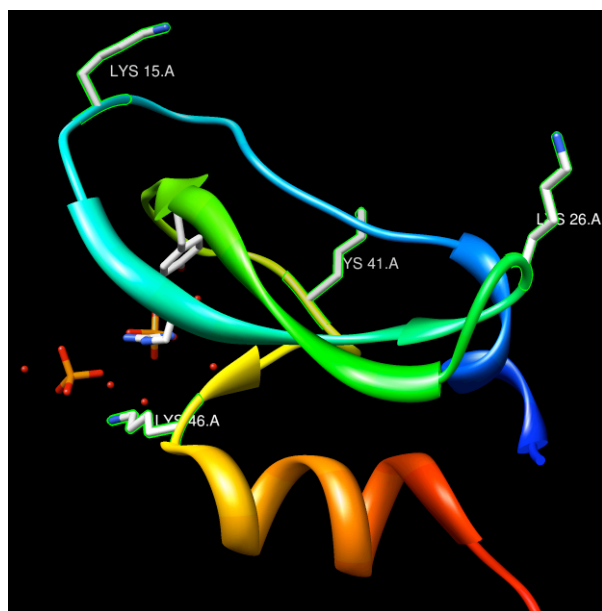


Figure 6.2-1 Ribbon representation of BPTI

The lysines located in the loop structure are labeled. The picture was generated with UCSF Chimera with crystal data from the protein data bank, (PDB code 1bpi).⁶¹

Our studies have shown that fluorine is able to restore inhibitory activity in a mutant BPTI background in which Lys15 has been replaced by Abu; this was true for both DfeGly and TfeGly variants. Crystallographic analysis revealed that fluorine interacts weakly to the water molecule in S_1 pocket and team up with water

6 Summary and outlook

molecules to maintain the protease-inhibitor interaction. This type of binding interaction has not previously been described. Because the crystallization procedures for complexes between BPTI and β -trypsin and α -chymotrypsin have been optimized and established in the course of this thesis, this line of investigation can be continued and the crystallographic data used to determine how general this type of fluorine and water mediated protein-protein interaction may be. The structural water molecules mediated protease-BPTI interaction can be also analysis in gas phase by means of mass spectrometry.

Moreover, due to the unique effects of fluorine, it may be also interesting to rationally incorporate these amino acid analogues into the S_1 pocket of the protease, to study a potential fluorine-fluorine interaction in a native inhibitor-enzyme context.

In brief summary, our studies of the single replacement of an amino acid at a key position in BPTI indicate that fluorine is a powerful tool for rational protein design with the goal of enhancing stability and modulating protein-protein molecular recognition.

LITERATURE

1. David; B. Harper, D., Fluorine-containing natural products. *Journal of Fluorine Chemistry* **1999**, *100* (1–2), 127-133.
2. Müller, K.; Faeh, C.; Diederich, F., Fluorine in Pharmaceuticals: Looking Beyond Intuition. *Science* **2007**, *317* (5846), 1881-1886.
3. (a) Babudri, F.; Farinola, G. M.; Naso, F.; Ragni, R., Fluorinated organic materials for electronic and optoelectronic applications: the role of the fluorine atom. *Chemical Communications* **2007**, (10), 1003-1022; (b) Berger, R.; Resnati, G.; Mentrangolo, P.; Weber, E.; Hulliger, J., Organic fluorine compounds: a great opportunity for enhanced materials properties. *Chemical Society Reviews* **2011**, *40* (7), 3496-3508.
4. Salwiczek, M.; Nyakatura, E. K.; Gerling, U. I. M.; Ye, S.; Kocsch, B., Fluorinated amino acids: compatibility with native protein structures and effects on protein-protein interactions. *Chemical Society Reviews* **2012**, *41* (6), 2135-2171.
5. Bégué, J.-P.; Bonnet-Delpon, D., General Remarks on Structural, Physical, and Chemical Properties of Fluorinated Compounds. In *Bioorganic and Medicinal Chemistry of Fluorine*, John Wiley & Sons, Inc.: 2007; pp 1-22.
6. O'Hagan, D., Understanding organofluorine chemistry. An introduction to the C-F bond. *Chemical Society Reviews* **2008**, *37* (2), 308-319.
7. Nowak, M. W.; Kearney, P. C.; Sampson, J. R.; Saks, M. E.; Labarca, C. G.; Silverman, S. K.; Zhong, W.; Thorson, J.; Abelson, J. N.; Davidson, N.; Schultz, P. G.; Dougherty, D. A.; Lester, H. A., Nicotinic receptor binding site probed with unnatural amino acid incorporation in intact cells. *Science* **1995**, *268* (5209), 439-442.
8. (a) Carcenac, Y.; Tordeux, M.; Wakselman, C.; Diter, P., Experimental determination of the conformational free energies (A values) of fluorinated substituents in cyclohexane by dynamic ¹⁹F NMR spectroscopy. Part 2. Extension to fluoromethyl, difluoromethyl, pentafluoroethyl, trifluoromethylthio and trifluoromethoxy groups. *New Journal of Chemistry* **2006**, *30* (3), 447-457; (b) Hirsch, J. A., Table of Conformational Energies—1967. In *Topics in Stereochemistry*, John Wiley & Sons, Inc.: 2007; pp 199-222.
9. Leroux, F., Atropisomerism, Biphenyls, and Fluorine: A Comparison of Rotational Barriers and Twist Angles. *ChemBioChem* **2004**, *5* (5), 644-649.
10. Klocker, J.; Karpfen, A.; Wolschann, P., On the structure and torsional potential of trifluoromethoxybenzene: an ab initio and density functional study. *Chemical Physics Letters* **2003**, *367* (5–6), 566-575.
11. Massa, M. A.; Spangler, D. P.; Durley, R. C.; Hickory, B. S.; Connolly, D. T.; Witherbee, B. J.; Smith, M. E.; Sikorski, J. A., Novel heteroaryl replacements of aromatic 3-tetrafluoroethoxy substituents in trifluoro-3-(tertiaryamino)-2-propanols as potent inhibitors of cholesteryl ester transfer protein. *Bioorganic & Medicinal Chemistry Letters* **2001**, *11* (13), 1625-1628.
12. O'Hagan, D.; Bilton, C.; Howard, J. A. K.; Knight, L.; Tozer, D. J., The preferred conformation of N-[small beta]-fluoroethylamides. Observation of the fluorine amide gauche effect. *Journal of the Chemical Society, Perkin Transactions 2* **2000**, (4), 605-607.
13. DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines, R. T.; Markley, J. L., Collagen Stability: Insights from NMR Spectroscopic and Hybrid

Density Functional Computational Investigations of the Effect of Electronegative Substituents on Prolyl Ring Conformations. *Journal of the American Chemical Society* **2002**, *124* (11), 2497-2505.

14. Purser, S.; Moore, P. R.; Swallow, S.; Gouverneur, V., Fluorine in medicinal chemistry. *Chemical Society Reviews* **2008**, *37* (2), 320-330.

15. Smart, B. E., Fluorine substituent effects (on bioactivity). *Journal of Fluorine Chemistry* **2001**, *109* (1), 3-11.

16. van Niel, M. B.; Collins, I.; Beer, M. S.; Broughton, H. B.; Cheng, S. K. F.; Goodacre, S. C.; Heald, A.; Locker, K. L.; MacLeod, A. M.; Morrison, D.; Moyes, C. R.; O'Connor, D.; Pike, A.; Rowley, M.; Russell, M. G. N.; Sohal, B.; Stanton, J. A.; Thomas, S.; Verrier, H.; Watt, A. P.; Castro, J. L., Fluorination of 3-(3-(Piperidin-1-yl)propyl)indoles and 3-(3-(Piperazin-1-yl)propyl)indoles Gives Selective Human 5-HT_{1D} Receptor Ligands with Improved Pharmacokinetic Profiles. *Journal of Medicinal Chemistry* **1999**, *42* (12), 2087-2104.

17. (a) Bannai, K.; Toru, T.; Ōba, T.; Tanaka, T.; Okamura, N.; Watanabe, K.; Hazato, A.; Kurozumi, S., Synthesis of chemically stable prostacyclin analogs. *Tetrahedron* **1983**, *39* (22), 3807-3819; (b) Chang, C.-S.; Negishi, M.; Nakano, T.; Morizawa, Y.; Matsumura, Y.; Ichikawa, A., 7,7-Difluoroprostacyclin derivative, AFP-07, a highly selective and potent agonist for the prostacyclin receptor. *Prostaglandins* **1997**, *53* (2), 83-90.

18. (a) Dunitz, J. D., Organic Fluorine: Odd Man Out. *ChemBioChem* **2004**, *5* (5), 614-621; (b) Dunitz, J. D.; Taylor, R., Organic Fluorine Hardly Ever Accepts Hydrogen Bonds. *Chemistry – A European Journal* **1997**, *3* (1), 89-98.

19. (a) Barbarich, T. J.; Rithner, C. D.; Miller, S. M.; Anderson, O. P.; Strauss, S. H., Significant Inter- and Intramolecular O–H···FC Hydrogen Bonding. *Journal of the American Chemical Society* **1999**, *121* (17), 4280-4281; (b) Dalvit, C.; Vulpetti, A., Intermolecular and Intramolecular Hydrogen Bonds Involving Fluorine Atoms: Implications for Recognition, Selectivity, and Chemical Properties. *ChemMedChem* **2012**, *7* (2), 262-272; (c) Giuffredi, G. T.; Gouverneur, V.; Bernet, B., Intramolecular OH···FC Hydrogen Bonding in Fluorinated Carbohydrates: CHF is a Better Hydrogen Bond Acceptor than CF₂. *Angewandte Chemie International Edition* **2013**, n/a-n/a; (d) Li, C.; Ren, S.-F.; Hou, J.-L.; Yi, H.-P.; Zhu, S.-Z.; Jiang, X.-K.; Li, Z.-T., F···H N Hydrogen Bonding Driven Foldamers: Efficient Receptors for Dialkylammonium Ions. *Angewandte Chemie* **2005**, *117* (35), 5871-5875; (e) Li, C.; Zhu, Y.-Y.; Yi, H.-P.; Li, C.-Z.; Jiang, X.-K.; Li, Z.-T.; Yu, Y.-H., Strong Stacking between F···H N Hydrogen-Bonded Foldamers and Fullerenes: Formation of Supramolecular Nano Networks. *Chemistry – A European Journal* **2007**, *13* (35), 9990-9998; (f) Pham, M.; Gdaniec, M.; Połoński, T., Three-Center CF···HN Intramolecular Hydrogen Bonding in the 2,6-Bis(2,6-difluorophenyl)piperidine Systems¹. *The Journal of Organic Chemistry* **1998**, *63* (11), 3731-3734; (g) Takahashi, L. H.; Radhakrishnan, R.; Rosenfield Jr, R. E.; Meyer Jr, E. F.; Trainor, D. A.; Stein, M., X-ray diffraction analysis of the inhibition of porcine pancreatic elastase by a peptidyl trifluoromethylketone. *Journal of Molecular Biology* **1988**, *201* (2), 423-428; (h) Takahashi, L. H.; Radhakrishnan, R.; Rosenfield, R. E.; Meyer, E. F.; Trainor, D. A., Crystal structure of the covalent complex formed by a peptidyl .alpha.,.alpha.-difluoro-.beta.-keto amide with porcine pancreatic elastase at 1.78 Å resolution. *Journal of the American Chemical Society* **1989**, *111* (9), 3368-3374.

20. (a) Fischer, F. R.; Schweizer, W. B.; Diederich, F., Molecular Torsion Balances: Evidence for Favorable Orthogonal Dipolar Interactions Between Organic Fluorine and Amide Groups. *Angewandte Chemie International Edition* **2007**, *46* (43), 8270-8273; (b) Hof, F.; Scofield, D. M.; Schweizer, W. B.; Diederich, F., A Weak Attractive Interaction between Organic Fluorine and an Amide Group. *Angewandte Chemie International Edition* **2004**, *43* (38), 5056-5059; (c) Olsen, J. A.; Banner, D. W.; Seiler, P.; Obst Sander, U.; D'Arcy, A.; Stihle, M.; Müller, K.; Diederich, F., A Fluorine Scan of Thrombin Inhibitors to Map the Fluorophilicity/Fluorophobicity of an Enzyme Active Site: Evidence for C F ··· C O Interactions. *Angewandte Chemie International Edition* **2003**, *42* (22), 2507-2511; (d) Olsen, J. A.; Banner, D. W.; Seiler, P.; Wagner, B.; Tschopp, T.; Obst-Sander, U.; Kansy, M.; Müller, K.; Diederich, F., Fluorine Interactions at the Thrombin Active Site: Protein Backbone Fragments H C α C O Comprise a Favorable C F Environment and Interactions of C F with Electrophiles. *ChemBioChem* **2004**, *5* (5), 666-675.
21. (a) Cametti, M.; Crousse, B.; Metrangolo, P.; Milani, R.; Resnati, G., The fluorous effect in biomolecular applications. *Chemical Society Reviews* **2012**, *41* (1), 31-42; (b) Curran, D. P., Fluorous Tags Unstick Messy Chemical Biology Problems. *Science* **2008**, *321* (5896), 1645-1646; (c) Studer, A.; Hadida, S.; Ferritto, R.; Kim, S.-Y.; Jeger, P.; Wipf, P.; Curran, D. P., Fluorous Synthesis: A Fluorous-Phase Strategy for Improving Separation Efficiency in Organic Synthesis. *Science* **1997**, *275* (5301), 823-826; (d) Zhang, W., Fluorous Synthesis of Heterocyclic Systems. *Chemical Reviews* **2004**, *104* (5), 2531-2556.
22. Ko, K.-S.; Jaipuri, F. A.; Pohl, N. L., Fluorous-Based Carbohydrate Microarrays. *Journal of the American Chemical Society* **2005**, *127* (38), 13162-13163.
23. Brittain, S. M.; Ficarro, S. B.; Brock, A.; Peters, E. C., Enrichment and analysis of peptide subsets using fluorous affinity tags and mass spectrometry. *Nat Biotech* **2005**, *23* (4), 463-468.
24. DeSimone, J. M., Practical Approaches to Green Solvents. *Science* **2002**, *297* (5582), 799-803.
25. (a) Ametamey, S. M.; Honer, M.; Schubiger, P. A., Molecular Imaging with PET. *Chemical Reviews* **2008**, *108* (5), 1501-1516; (b) Miller, P. W.; Long, N. J.; Vilar, R.; Gee, A. D., Synthese von ¹¹C-, ¹⁸F-, ¹⁵O- und ¹³N-Radiotracer für die Positronenemissionstomographie. *Angewandte Chemie* **2008**, *120* (47), 9136-9172.
26. Liang, T.; Neumann, C. N.; Ritter, T., Introduction of Fluorine and Fluorine-Containing Functional Groups. *Angewandte Chemie International Edition* **2013**, *52* (32), 8214-8264.
27. (a) Danielson, M. A.; Falke, J. J., Use of ¹⁹F NMR to Probe Protein Structure and Conformational Changes. *Annual Review of Biophysics and Biomolecular Structure* **1996**, *25* (1), 163-195; (b) Kitevski-LeBlanc, J. L.; Prosser, R. S., Current applications of ¹⁹F NMR to studies of protein structure and dynamics. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2012**, *62* (0), 1-33.
28. (a) Chen, J.; Lanza, G. M.; Wickline, S. A., Quantitative magnetic resonance fluorine imaging: today and tomorrow. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology* **2010**, *2* (4), 431-440; (b) Lowe, K. C., Perfluorinated blood substitutes and artificial oxygen carriers. *Blood Reviews* **1999**, *13* (3), 171-184; (c)

Srinivas, M.; Heerschap, A.; Ahrens, E. T.; Figdor, C. G.; Vries, I. J. M. d., 19F MRI for quantitative in vivo cell tracking. *Trends in Biotechnology* **2010**, *28* (7), 363-370.

29. Wolf, W.; Presant, C. A.; Waluch, V., 19F-MRS studies of fluorinated drugs in humans. *Advanced Drug Delivery Reviews* **2000**, *41* (1), 55-74.

30. (a) March, T. L.; Johnston, M. R.; Duggan, P. J.; Gardiner, J., Synthesis, Structure, and Biological Applications of α -Fluorinated β -Amino Acids and Derivatives. *Chemistry & Biodiversity* **2012**, *9* (11), 2410-2441; (b) Smits, R.; Cadicamo, C. D.; Burger, K.; Kokschi, B., Synthetic strategies to [small alpha]-trifluoromethyl and [small alpha]-difluoromethyl substituted [small alpha]-amino acids. *Chemical Society Reviews* **2008**, *37* (8), 1727-1739.

31. (a) Lumb, K. J.; Kim, P. S., Measurement of interhelical electrostatic interactions in the GCN4 leucine zipper. *Science* **1995**, *268* (5209), 436-439; (b) O'Shea, E. K.; Klemm, J. D.; Kim, P. S.; Alber, T., X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **1991**, *254* (5031), 539-544; (c) Wendt, H.; Berger, C.; Baici, A.; Thomas, R. M.; Bosshard, H. R., Kinetics of Folding of Leucine Zipper Domains. *Biochemistry* **1995**, *34* (12), 4097-4107.

32. Woolfson, D. N., The Design of Coiled-Coil Structures and Assemblies. In *Advances in Protein Chemistry*, David, A. D. P.; John, M. S., Eds. Academic Press: 2005; Vol. Volume 70, pp 79-112.

33. (a) Montclare, J. K.; Son, S.; Clark, G. A.; Kumar, K.; Tirrell, D. A., Biosynthesis and Stability of Coiled-Coil Peptides Containing (2S,4R)-5,5,5-Trifluoroleucine and (2S,4S)-5,5,5-Trifluoroleucine. *ChemBioChem* **2009**, *10* (1), 84-86; (b) Montclare, J. K.; Tirrell, D. A., Evolving Proteins of Novel Composition. *Angewandte Chemie* **2006**, *118* (27), 4630-4633; (c) Son, S.; Tanrikulu, I. C.; Tirrell, D. A., Stabilization of bzip Peptides through Incorporation of Fluorinated Aliphatic Residues. *ChemBioChem* **2006**, *7* (8), 1251-1257; (d) Tang, Y.; Ghirlanda, G.; Petka, W. A.; Nakajima, T.; DeGrado, W. F.; Tirrell, D. A., Fluorinated Coiled-Coil Proteins Prepared In Vivo Display Enhanced Thermal and Chemical Stability. *Angewandte Chemie International Edition* **2001**, *40* (8), 1494-1496; (e) Tang, Y.; Ghirlanda, G.; Vaidehi, N.; Kua, J.; Mainz, D. T.; Goddard, W. A.; DeGrado, W. F.; Tirrell, D. A., Stabilization of Coiled-Coil Peptide Domains by Introduction of Trifluoroleucine†. *Biochemistry* **2001**, *40* (9), 2790-2796; (f) Tang, Y.; Tirrell, D. A., Biosynthesis of a Highly Stable Coiled-Coil Protein Containing Hexafluoroleucine in an Engineered Bacterial Host. *Journal of the American Chemical Society* **2001**, *123* (44), 11089-11090.

34. (a) Bilgiçer, B.; Fichera, A.; Kumar, K., A Coiled Coil with a Fluorous Core. *Journal of the American Chemical Society* **2001**, *123* (19), 4393-4399; (b) Bilgiçer, B.; Xing, X.; Kumar, K., Programmed Self-Sorting of Coiled Coils with Leucine and Hexafluoroleucine Cores. *Journal of the American Chemical Society* **2001**, *123* (47), 11815-11816.

35. Bilgiçer, B.; Kumar, K., De novo design of defined helical bundles in membrane environments. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101* (43), 15324-15329.

36. (a) Gottler, L. M.; de la Salud-Bea, R.; Marsh, E. N. G., The Fluorous Effect in Proteins: Properties of α 4F6, a 4- α -Helix Bundle Protein with a Fluorocarbon Core†. *Biochemistry* **2008**, *47* (15), 4484-4490; (b) Lee, K.-H.; Lee, H.-Y.; Slutsky, M. M.; Anderson, J. T.; Marsh, E. N. G., Fluorous Effect in Proteins: De Novo Design and

Characterization of a Four- α -Helix Bundle Protein Containing Hexafluoroleucine†. *Biochemistry* **2004**, *43* (51), 16277-16284.

37. (a) Buer, B. C.; de la Salud-Bea, R.; Al Hashimi, H. M.; Marsh, E. N. G., Engineering Protein Stability and Specificity Using Fluorous Amino Acids: The Importance of Packing Effects. *Biochemistry* **2009**, *48* (45), 10810-10817; (b) Buer, B. C.; Meagher, J. L.; Stuckey, J. A.; Marsh, E. N. G., Structural basis for the enhanced stability of highly fluorinated proteins. *Proceedings of the National Academy of Sciences* **2012**, *109* (13), 4810-4815.

38. (a) Jäckel, C.; Salwiczek, M.; Kokschi, B., Fluorine in a Native Protein Environment—How the Spatial Demand and Polarity of Fluoroalkyl Groups Affect Protein Folding. *Angewandte Chemie International Edition* **2006**, *45* (25), 4198-4203; (b) Jäckel, C.; Seufert, W.; Thust, S.; Kokschi, B., Evaluation of the Molecular Interactions of Fluorinated Amino Acids with Native Polypeptides. *ChemBioChem* **2004**, *5* (5), 717-720; (c) Salwiczek, M.; Kokschi, B., Effects of Fluorination on the Folding Kinetics of a Heterodimeric Coiled Coil. *ChemBioChem* **2009**, *10* (18), 2867-2870; (d) Salwiczek, M.; Samsonov, S.; Vagt, T.; Nyakatura, E.; Fleige, E.; Numata, J.; Cölfen, H.; Pisabarro, M. T.; Kokschi, B., Position-Dependent Effects of Fluorinated Amino Acids on the Hydrophobic Core Formation of a Heterodimeric Coiled Coil. *Chemistry – A European Journal* **2009**, *15* (31), 7628-7636; (e) Samsonov, S. A.; Salwiczek, M.; Anders, G.; Kokschi, B.; Pisabarro, M. T., Fluorine in Protein Environments: A QM and MD Study. *The Journal of Physical Chemistry B* **2009**, *113* (51), 16400-16408.

39. Vagt, T.; Nyakatura, E.; Salwiczek, M.; Jackel, C.; Kokschi, B., Towards identifying preferred interaction partners of fluorinated amino acids within the hydrophobic environment of a dimeric coiled coil peptide. *Organic & Biomolecular Chemistry* **2010**, *8* (6), 1382-1386.

40. (a) Chiu, H.-P.; Kokona, B.; Fairman, R.; Cheng, R. P., Effect of Highly Fluorinated Amino Acids on Protein Stability at a Solvent-Exposed Position on an Internal Strand of Protein G B1 Domain. *Journal of the American Chemical Society* **2009**, *131* (37), 13192-13193; (b) Chiu, H.-P.; Suzuki, Y.; Gullickson, D.; Ahmad, R.; Kokona, B.; Fairman, R.; Cheng, R. P., Helix Propensity of Highly Fluorinated Amino Acids. *Journal of the American Chemical Society* **2006**, *128* (49), 15556-15557.

41. Horng, J.-C.; Raleigh, D. P., Φ -Values beyond the Ribosomally Encoded Amino Acids: Kinetic and Thermodynamic Consequences of Incorporating Trifluoromethyl Amino Acids in a Globular Protein. *Journal of the American Chemical Society* **2003**, *125* (31), 9286-9287.

42. (a) Erdbrink, H.; Peuser, I.; Gerling, U. I. M.; Lentz, D.; Kokschi, B.; Czekelius, C., Conjugate hydrotrifluoromethylation of [small alpha],[small beta]-unsaturated acyl-oxazolidinones: synthesis of chiral fluorinated amino acids. *Organic & Biomolecular Chemistry* **2012**, *10* (43), 8583-8586; (b) Gerling, U. I. M.; Salwiczek, M.; Cadicamo, C. D.; Erdbrink, H.; Czekelius, C.; Grage, S. L.; Wadhvani, P.; Ulrich, A. S.; Behrends, M.; Haufe, G.; Kokschi, B., Fluorinated amino acids in amyloid formation: a symphony of size, hydrophobicity and [small alpha]-helix propensity. *Chemical Science* **2014**.

43. Zasloff, M., Antimicrobial peptides of multicellular organisms. *Nature* **2002**, *415* (6870), 389-395.

44. Gottler, L. M.; Lee, H.-Y.; Shelburne, C. E.; Ramamoorthy, A.; Marsh, E. N. G., Using Fluorous Amino Acids to Modulate the Biological Activity of an Antimicrobial Peptide. *ChemBioChem* **2008**, *9* (3), 370-373.

45. Gottler, L. M.; de la Salud Bea, R.; Shelburne, C. E.; Ramamoorthy, A.; Marsh, E. N. G., Using Fluorous Amino Acids To Probe the Effects of Changing Hydrophobicity on the Physical and Biological Properties of the β -Hairpin Antimicrobial Peptide Protegrin-1†. *Biochemistry* **2008**, *47* (35), 9243-9250.
46. (a) Pace, C. J.; Zheng, H.; Mylvaganam, R.; Kim, D.; Gao, J., Stacked Fluoroaromatics as Supramolecular Synthons for Programming Protein Dimerization Specificity. *Angewandte Chemie International Edition* **2012**, *51* (1), 103-107; (b) Zheng, H.; Comeforo, K.; Gao, J., Expanding the Fluorous Arsenal: Tetrafluorinated Phenylalanines for Protein Design. *Journal of the American Chemical Society* **2008**, *131* (1), 18-19; (c) Zheng, H.; Gao, J., Highly Specific Heterodimerization Mediated by Quadrupole Interactions. *Angewandte Chemie International Edition* **2010**, *49* (46), 8635-8639.
47. Hill, R. B.; DeGrado, W. F., Solution Structure of α 2D, a Nativelike de Novo Designed Protein. *Journal of the American Chemical Society* **1998**, *120* (6), 1138-1145.
48. Panchenko, T.; Zhu, W. W.; Montclare, J. K., Influence of global fluorination on chloramphenicol acetyltransferase activity and stability. *Biotechnology and Bioengineering* **2006**, *94* (5), 921-930.
49. Yoo, T. H.; Link, A. J.; Tirrell, D. A., Evolution of a fluorinated green fluorescent protein. *Proceedings of the National Academy of Sciences* **2007**, *104* (35), 13887-13890.
50. (a) Bae, J. H.; Paramita Pal, P.; Moroder, L.; Huber, R.; Budisa, N., Crystallographic Evidence for Isomeric Chromophores in 3-Fluorotyrosyl-Green Fluorescent Protein. *ChemBioChem* **2004**, *5* (5), 720-722; (b) Pal, P. P.; Bae, J. H.; Azim, M. K.; Hess, P.; Friedrich, R.; Huber, R.; Moroder, L.; Budisa, N., Structural and Spectral Response of *Aequorea victoria* Green Fluorescent Proteins to Chromophore Fluorination†. *Biochemistry* **2005**, *44* (10), 3663-3672; (c) Veettil, S.; Budisa, N.; Jung, G., Photostability of green and yellow fluorescent proteins with fluorinated chromophores, investigated by fluorescence correlation spectroscopy. *Biophysical chemistry* **2008**, *136* (1), 38-43.
51. (a) Budisa, N.; Pal, P. P.; Alefelder, S.; Birle, P.; Krywcun, T.; Rubini, M.; Wenger, W.; Bae, J. H.; Steiner, T., Probing the role of tryptophans in *Aequorea victoria* green fluorescent proteins with an expanded genetic code. *Biological chemistry* **2004**, *385* (2), 191-202; (b) Hyun Bae, J.; Rubini, M.; Jung, G.; Wiegand, G.; Seifert, M. H. J.; Azim, M. K.; Kim, J.-S.; Zumbusch, A.; Holak, T. A.; Moroder, L.; Huber, R.; Budisa, N., Expansion of the Genetic Code Enables Design of a Novel "Gold" Class of Green Fluorescent Proteins. *Journal of Molecular Biology* **2003**, *328* (5), 1071-1081.
52. Holzberger, B.; Marx, A., Replacing 32 Proline Residues by a Noncanonical Amino Acid Results in a Highly Active DNA Polymerase. *Journal of the American Chemical Society* **2010**, *132* (44), 15708-15713.
53. Holzberger, B.; Rubini, M.; Möller, H. M.; Marx, A., A Highly Active DNA Polymerase with a Fluorous Core. *Angewandte Chemie International Edition* **2010**, *49* (7), 1324-1327.
54. (a) Buer, B. C.; Marsh, E. N. G., Fluorine: A new element in protein design. *Protein Science* **2012**, *21* (4), 453-462; (b) Merkel, L.; Budisa, N., Organic fluorine as a polypeptide building element: in vivo expression of fluorinated peptides, proteins and proteomes. *Organic & Biomolecular Chemistry* **2012**, *10* (36), 7241-7261; (c) Yoder, N. C.; Kumar, K., Fluorinated amino acids in protein design and engineering. *Chemical Society Reviews* **2002**, *31* (6), 335-341.

55. Ascenzi, P.; Bocedi, A.; Bolognesi, M.; Spallarossa, A.; Coletta, M.; De Cristofaro, R.; Menegatti, E., The bovine basic pancreatic trypsin inhibitor (Kunitz inhibitor): a milestone protein. *Current protein & peptide science* **2003**, *4* (3), 231-51.
56. (a) Deisenhofer, J.; Steigemann, W., Crystallographic refinement of the structure of bovine pancreatic trypsin inhibitor at 1.5 Å resolution. *Acta Crystallographica Section B* **1975**, *31* (1), 238-250; (b) Huber, R.; Kukla, D.; Rühlmann, A.; Epp, O.; Formanek, H., The basic trypsin inhibitor of bovine pancreas. *Naturwissenschaften* **1970**, *57* (8), 389-392.
57. McCammon, J. A.; Gelin, B. R.; Karplus, M., Dynamics of folded proteins. *Nature* **1977**, *267* (5612), 585-590.
58. (a) Havel, T. F.; Wüthrich, K., An evaluation of the combined use of nuclear magnetic resonance and distance geometry for the determination of protein conformations in solution. *Journal of Molecular Biology* **1985**, *182* (2), 281-294; (b) Wagner, G.; Wüthrich, K., Sequential resonance assignments in protein 1H nuclear magnetic resonance spectra: Basic pancreatic trypsin inhibitor. *Journal of Molecular Biology* **1982**, *155* (3), 347-366.
59. (a) Creighton, T., The disulfide folding pathway of BPTI. *Science* **1992**, *256* (5053), 111-114; (b) Weissman, J.; Kim, P., Reexamination of the folding of BPTI: predominance of native intermediates. *Science* **1991**, *253* (5026), 1386-1393; (c) Weissman, J. S.; Kim, P. S., A kinetic explanation for the rearrangement pathway of BPTI folding. *Nat Struct Mol Biol* **1995**, *2* (12), 1123-1130.
60. (a) Gardiner, H., F.D.A Says Bayer Failed to Reveal Drug Risk Study. *The New Yorks Times* **2006**; (b) Schneeweiss, S.; Seeger, J. D.; Landon, J.; Walker, A. M., Aprotinin during Coronary-Artery Bypass Grafting and Risk of Death. *New England Journal of Medicine* **2008**, *358* (8), 771-783; (c) Shaw, A. D.; Stafford-Smith, M.; White, W. D.; Phillips-Bute, B.; Swaminathan, M.; Milano, C.; Welsby, I. J.; Aronson, S.; Mathew, J. P.; Peterson, E. D.; Newman, M. F., The Effect of Aprotinin on Outcome after Coronary-Artery Bypass Grafting. *New England Journal of Medicine* **2008**, *358* (8), 784-793.
61. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **2004**, *25* (13), 1605-1612.
62. Moses, E.; Hinz, H.-J., Basic pancreatic trypsin inhibitor has unusual thermodynamic stability parameters. *Journal of Molecular Biology* **1983**, *170* (3), 765-776.
63. Makhatadze, G. I.; Kim, K.-S.; Woodward, C.; Privalov, P. L., Thermodynamics of bpti folding. *Protein Science* **1993**, *2* (12), 2028-2036.
64. Chang, J.-Y.; Ballatore, A., The structure of denatured bovine pancreatic trypsin inhibitor (BPTI). *FEBS Letters* **2000**, *473* (2), 183-187.
65. Krowarsch, D.; Otlewski, J., Amino-acid substitutions at the fully exposed P1 site of bovine pancreatic trypsin inhibitor affect its stability. *Protein Science* **2001**, *10* (4), 715-724.
66. (a) Bode, W.; Huber, R., Natural protein proteinase inhibitors and their interaction with proteinases. *European Journal of Biochemistry* **1992**, *204* (2), 433-451; (b) Bode, W.; Huber, R., Structural basis of the endoproteinase-protein inhibitor interaction. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **2000**, *1477* (1-2), 241-252; (c) Marquart, M.; Walter, J.; Deisenhofer, J.; Bode, W.; Huber, R., The geometry of the reactive site and of the peptide groups in

trypsin, trypsinogen and its complexes with inhibitors. *Acta Crystallographica Section B* **1983**, 39 (4), 480-490.

67. Borjigin, J.; Nathans, J., Bovine pancreatic trypsin inhibitor-trypsin complex as a detection system for recombinant proteins. *Proceedings of the National Academy of Sciences* **1993**, 90 (1), 337-341.

68. Blow, D. M.; Wright, C. S.; Kukla, D.; Rühlmann, A.; Steigemann, W.; Huber, R., A model for the association of bovine pancreatic trypsin inhibitor with chymotrypsin and trypsin. *Journal of Molecular Biology* **1972**, 69 (1), 137-144.

69. (a) Buczek, O.; Koscielska-Kasprzak, K.; Krowarsch, D.; Dadlez, M.; Otlewski, J., Analysis of serine proteinase-inhibitor interaction by alanine shaving. *Protein Science* **2002**, 11 (4), 806-819; (b) Krowarsch, D.; Dadlez, M.; Buczek, O.; Krokoszynska, I.; Smalas, A. O.; Otlewski, J., Interscaffolding additivity: binding of P1 variants of bovine pancreatic trypsin inhibitor to four serine proteases. *Journal of Molecular Biology* **1999**, 289 (1), 175-186.

70. Czapinska, H.; Otlewski, J.; Krzywda, S.; Sheldrick, G. M.; Jaskólski, M., High-resolution structure of bovine pancreatic trypsin inhibitor with altered binding loop sequence. *Journal of Molecular Biology* **2000**, 295 (5), 1237-1249.

71. (a) Böck, A.; Forchhammer, K.; Heider, J.; Leinfelder, W.; Sawers, G.; Veprek, B.; Zinoni, F., Selenocysteine: the 21st amino acid. *Molecular Microbiology* **1991**, 5 (3), 515-520; (b) 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), 1459-1462.

72. (a) Budisa, N., Prolegomena to Future Experimental Efforts on Genetic Code Engineering by Expanding Its Amino Acid Repertoire. *Angewandte Chemie International Edition* **2004**, 43 (47), 6426-6463; (b) Chin, J. W., Reprogramming the Genetic Code. *Science* **2012**, 336 (6080), 428-429; (c) Wang, L.; Schultz, P. G., Expanding the Genetic Code. *Angewandte Chemie International Edition* **2005**, 44 (1), 34-66; (d) Wang, L.; Xie, J.; Schultz, P. G., EXPANDING THE GENETIC CODE. *Annual Review of Biophysics and Biomolecular Structure* **2006**, 35 (1), 225-249.

73. (a) Bouadloun, F.; Donner, D.; Kurland, C. G., Codon-specific missense errors in vivo. *The EMBO journal* **1983**, 2 (8), 1351-1356; (b) Edelman, P.; Gallant, J., Mistranslation in *E. coli*. *Cell* **1977**, 10 (1), 131-137; (c) Kunkel, T. A.; Bebenek, K., DNA REPLICATION FIDELITY1. *Annual Review of Biochemistry* **2000**, 69 (1), 497-529; (d) Rosenberger, R. F.; Foskett, G., An estimate of the frequency of in vivo transcriptional errors at a nonsense codon in *Escherichia coli*. *Molec. Gen. Genet.* **1981**, 183 (3), 561-563.

74. Sydow, J. F.; Cramer, P., RNA polymerase fidelity and transcriptional proofreading. *Current Opinion in Structural Biology* **2009**, 19 (6), 732-739.

75. Zaher, H. S.; Green, R., Fidelity at the Molecular Level: Lessons from Protein Synthesis. *Cell* **2009**, 136 (4), 746-762.

76. (a) Cochella, L.; Green, R., Fidelity in protein synthesis. *Current Biology* **2005**, 15 (14), R536-R540; (b) Dale, T.; Uhlenbeck, O. C., Amino acid specificity in translation. *Trends in Biochemical Sciences* **2005**, 30 (12), 659-665; (c) Ling, J.; Reynolds, N.; Ibba, M., Aminoacyl-tRNA Synthesis and Translational Quality Control. *Annual Review of Microbiology* **2009**, 63 (1), 61-78; (d) Ramakrishnan, V., Ribosome Structure and the Mechanism of Translation. *Cell* **2002**, 108 (4), 557-572.

77. Banerjee, R.; Chen, S.; Dare, K.; Gilreath, M.; Praetorius-Ibba, M.; Raina, M.; Reynolds, N. M.; Rogers, T.; Roy, H.; Yadavalli, S. S.; Ibba, M., tRNAs: Cellular barcodes for amino acids. *FEBS Letters* **2010**, *584* (2), 387-395.
78. (a) Brick, P.; Bhat, T. N.; Blow, D. M., Structure of tyrosyl-tRNA synthetase refined at 2.3 Å resolution: Interaction of the enzyme with the tyrosyl adenylate intermediate. *Journal of Molecular Biology* **1989**, *208* (1), 83-98; (b) Perona, J. J.; Hadd, A., Structural Diversity and Protein Engineering of the Aminoacyl-tRNA Synthetases. *Biochemistry* **2012**, *51* (44), 8705-8729.
79. (a) Cusack, S.; Berthet-Colominas, C.; Hartlein, M.; Nassar, N.; Leberman, R., A second class of synthetase structure revealed by X-ray analysis of Escherichia coli seryl-tRNA synthetase at 2.5 [angst]. *Nature* **1990**, *347* (6290), 249-255; (b) Ruff, M.; Krishnaswamy, S.; Boeglin, M.; Poterszman, A.; Mitschler, A.; Podjarny, A.; Rees, B.; Thierry, J.; Moras, D., Class II aminoacyl transfer RNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with tRNA(Asp). *Science* **1991**, *252* (5013), 1682-1689.
80. Fersht, A. R., Editing mechanisms in protein synthesis. Rejection of valine by the isoleucyl-tRNA synthetase. *Biochemistry* **1977**, *16* (5), 1025-1030.
81. (a) Beuning, P. J.; Musier-Forsyth, K., Transfer RNA recognition by aminoacyl-tRNA synthetases. *Biopolymers* **1999**, *52* (1), 1-28; (b) Giegé, R.; Sissler, M.; Florentz, C., Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Research* **1998**, *26* (22), 5017-5035.
82. (a) Muramatsu, T.; Nishikawa, K.; Nemoto, F.; Kuchino, Y.; Nishimura, S.; Miyazawa, T.; Yokoyama, S., Codon and amino-acid specificities of a transfer RNA are both converted by a single post-transcriptional modification. *Nature* **1988**, *336* (6195), 179-181; (b) Putz, J.; Florentz, C.; Benseler, F.; Giege, R., A single methyl group prevents the mischarging of a tRNA. *Nat Struct Mol Biol* **1994**, *1* (9), 580-582.
83. (a) Davies, J.; Gilbert, W.; Gorini, L., STREPTOMYCIN, SUPPRESSION, AND THE CODE. *Proceedings of the National Academy of Sciences* **1964**, *51* (5), 883-890; (b) Potapov, A. P., A stereospecific mechanism for the aminoacyl-tRNA selection at the ribosome. *FEBS Letters* **1982**, *146* (1), 5-8.
84. (a) Hopfield, J. J., Kinetic Proofreading: A New Mechanism for Reducing Errors in Biosynthetic Processes Requiring High Specificity. *Proceedings of the National Academy of Sciences* **1974**, *71* (10), 4135-4139; (b) Ninio, J., Kinetic amplification of enzyme discrimination. *Biochimie* **1975**, *57* (5), 587-595.
85. Fahlman, R. P.; Dale, T.; Uhlenbeck, O. C., Uniform Binding of Aminoacylated Transfer RNAs to the Ribosomal A and P Sites. *Molecular Cell* **2004**, *16* (5), 799-805.
86. (a) Nierhaus, K. H., The allosteric three-site model for the ribosomal elongation cycle: features and future. *Biochemistry* **1990**, *29* (21), 4997-5008; (b) Pape, T.; Wintermeyer, W.; Rodnina, M., Induced fit in initial selection and proofreading of aminoacyl-tRNA on the ribosome. *EMBO J* **1999**, *18* (13), 3800-3807.
87. Ogle, J. M.; Brodersen, D. E.; Clemons, W. M.; Tarry, M. J.; Carter, A. P.; Ramakrishnan, V., Recognition of Cognate Transfer RNA by the 30S Ribosomal Subunit. *Science* **2001**, *292* (5518), 897-902.
88. (a) Louie, A.; Journak, F., Kinetic studies of Escherichia coli elongation factor Tu-guanosine 5'-triphosphate-aminoacyl-tRNA complexes. *Biochemistry* **1985**, *24* (23), 6433-6439; (b) Ott, G.; Schiesswohl, M.; Kiesewetter, S.; Förster, C.; Arnold, L.; Erdmann,

- V. A.; Sprinzl, M., Ternary complexes of Escherichia coli aminoacyl-tRNAs with the elongation factor Tu and GTP: Thermodynamic and structural studies. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1990**, *1050* (1-3), 222-225.
89. LaRiviere, F. J.; Wolfson, A. D.; Uhlenbeck, O. C., Uniform Binding of Aminoacyl-tRNAs to Elongation Factor Tu by Thermodynamic Compensation. *Science* **2001**, *294* (5540), 165-168.
90. Link, A. J.; Mock, M. L.; Tirrell, D. A., Non-canonical amino acids in protein engineering. *Current Opinion in Biotechnology* **2003**, *14* (6), 603-609.
91. Budisa, N.; Steipe, B.; Demange, P.; Eckerskorn, C.; Kellermann, J.; Huber, R., High-level Biosynthetic Substitution of Methionine in Proteins by its Analogs 2-Aminohexanoic Acid, Selenomethionine, Telluromethionine and Ethionine in Escherichia coli. *European Journal of Biochemistry* **1995**, *230* (2), 788-796.
92. Kiick, K. L.; van Hest, J. C. M.; Tirrell, D. A., Expanding the Scope of Protein Biosynthesis by Altering the Methionyl-tRNA Synthetase Activity of a Bacterial Expression Host. *Angewandte Chemie* **2000**, *112* (12), 2232-2236.
93. Crick, F. H. C.; Barnett, L.; Brenner, S.; Watts-Tobin, R. J., General Nature of the Genetic Code for Proteins. *Nature* **1961**, *192* (4809), 1227-1232.
94. (a) Bain, J. D.; Diala, E. S.; Glabe, C. G.; Dix, T. A.; Chamberlin, A. R., Biosynthetic site-specific incorporation of a non-natural amino acid into a polypeptide. *Journal of the American Chemical Society* **1989**, *111* (20), 8013-8014; (b) Beene, D. L.; Dougherty, D. A.; Lester, H. A., Unnatural amino acid mutagenesis in mapping ion channel function. *Current Opinion in Neurobiology* **2003**, *13* (3), 264-270; (c) Noren, C.; Anthony-Cahill, S.; Griffith, M.; Schultz, P., A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* **1989**, *244* (4901), 182-188.
95. Ye, S.; Ann Berger, A.; Petzold, D.; Reimann, O.; Matt, B.; Kokschi, B., Chemical aminoacylation of tRNAs with fluorinated amino acids for in vitro protein mutagenesis. *Beilstein Journal of Organic Chemistry* **2010**, *6*, 40.
96. (a) Ayer, D.; Yarus, M., The context effect does not require a fourth base pair. *Science* **1986**, *231* (4736), 393-395; (b) Bossi, L.; Roth, J. R., The influence of codon context on genetic code translation. *Nature* **1980**, *286* (5769), 123-127; (c) Kleina, L. G.; Masson, J.-M.; Normanly, J.; Abelson, J.; Miller, J. H., Construction of Escherichia coli amber suppressor tRNA genes: II. Synthesis of additional tRNA genes and improvement of suppressor efficiency. *Journal of Molecular Biology* **1990**, *213* (4), 705-717; (d) Normanly, J.; Kleina, L. G.; Masson, J.-M.; Abelson, J.; Miller, J. H., Construction of Escherichia coli amber suppressor tRNA genes: III. Determination of tRNA specificity. *Journal of Molecular Biology* **1990**, *213* (4), 719-726; (e) Normanly, J.; Masson, J. M.; Kleina, L. G.; Abelson, J.; Miller, J. H., Construction of two Escherichia coli amber suppressor genes: tRNAPheCUA and tRNACysCUA. *Proceedings of the National Academy of Sciences* **1986**, *83* (17), 6548-6552.
97. (a) Hartman, M. C. T.; Josephson, K.; Szostak, J. W., Enzymatic aminoacylation of tRNA with unnatural amino acids. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103* (12), 4356-4361; (b) Josephson, K.; Hartman, M. C. T.; Szostak, J. W., Ribosomal Synthesis of Unnatural Peptides. *Journal of the American Chemical Society* **2005**, *127* (33), 11727-11735.
98. (a) Hecht, S. M.; Alford, B. L.; Kuroda, Y.; Kitano, S., "Chemical aminoacylation" of tRNA's. *Journal of Biological Chemistry* **1978**, *253* (13), 4517-20; (b) 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.

99. (a) Robertson, S. A.; Ellman, J. A.; Schultz, P. G., A general and efficient route for chemical aminoacylation of transfer RNAs. *Journal of the American Chemical Society* **1991**, *113* (7), 2722-2729; (b) Robertson, S. A.; Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G., The use of 5' -pbospho-2 deoxyribocytidylylriboadenosine as a facile route to chemical aminoacylation of tRNA. *Nucleic Acids Research* **1989**, *17* (23), 9649-9660.

100. (a) Lodder, M.; Golovine, S.; Laikhter, A. L.; Karginov, V. A.; Hecht, S. M., Misacylated Transfer RNAs Having a Chemically Removable Protecting Group. *The Journal of Organic Chemistry* **1998**, *63* (3), 794-803; (b) Madsen, R.; Roberts, C.; Fraser-Reid, B., The Pent-4-enoyl Group: A Novel Amine-Protecting Group That Is Readily Cleaved under Mild Conditions. *The Journal of Organic Chemistry* **1995**, *60* (24), 7920-7926.

101. Lodder, M.; Wang, B.; Hecht, S. M., The N-pentenoyl protecting group for aminoacyl-tRNAs. *Methods* **2005**, *36* (3), 245-251.

102. (a) Duca, M.; Chen, S.; Hecht, S. M., Aminoacylation of transfer RNAs with one and two amino acids. *Methods* **2008**, *44* (2), 87-99; (b) Duca, M.; Chen, S.; Hecht, S. M., Modeling the reactive properties of tandemly activated tRNAs. *Organic & Biomolecular Chemistry* **2008**, *6* (18), 3292-3299; (c) Duca, M.; Trindle, C. O.; Hecht, S. M., Structural Basis for the Exceptional Stability of Bisaminoacylated Nucleotides and Transfer RNAs. *Journal of the American Chemical Society* **2011**, *133* (29), 11368-11377; (d) Wang, B.; Zhou, J.; Lodder, M.; Anderson, R. D.; Hecht, S. M., Tandemly Activated tRNAs as Participants in Protein Synthesis. *Journal of Biological Chemistry* **2006**, *281* (20), 13865-13868.

103. (a) Cropp, T. A.; Anderson, J. C.; Chin, J. W., Reprogramming the amino-acid substrate specificity of orthogonal aminoacyl-tRNA synthetases to expand the genetic code of eukaryotic cells. *Nat. Protocols* **2007**, *2* (10), 2590-2600; (b) Ninomiya, K.; Kurita, T.; Hohsaka, T.; Sisido, M., Facile aminoacylation of pdCpA dinucleotide with a nonnatural amino acid in cationic micelle. *Chemical Communications* **2003**, (17), 2242-2243.

104. (a) Hashimoto, N.; Ninomiya, K.; Endo, T.; Sisido, M., Simple and quick chemical aminoacylation of tRNA in cationic micellar solution under ultrasonic agitation. *Chemical Communications* **2005**, (34), 4321-4323; (b) Ninomiya, K.; Minohata, T.; Nishimura, M.; Sisido, M., In Situ Chemical Aminoacylation with Amino Acid Thioesters Linked to a Peptide Nucleic Acid. *Journal of the American Chemical Society* **2004**, *126* (49), 15984-15989.

105. (a) Bessho, Y.; Hodgson, D. R. W.; Suga, H., A tRNA aminoacylation system for non-natural amino acids based on a programmable ribozyme. *Nat Biotech* **2002**, *20* (7), 723-728; (b) Lee, N.; Bessho, Y.; Wei, K.; Szostak, J. W.; Suga, H., Ribozyme-catalyzed tRNA aminoacylation. *Nat Struct Mol Biol* **2000**, *7* (1), 28-33; (c) Ohuchi, M.; Murakami, H.; Suga, H., The flexizyme system: a highly flexible tRNA aminoacylation tool for the translation apparatus. *Current Opinion in Chemical Biology* **2007**, *11* (5), 537-542; (d) Ramaswamy, K.; Saito, H.; Murakami, H.; Shiba, K.; Suga, H., Designer

- Ribozymes: Programming the tRNA Specificity into Flexizyme. *Journal of the American Chemical Society* **2004**, *126* (37), 11454-11455; (e) Xiao, H.; Murakami, H.; Suga, H.; Ferre-D'Amare, A. R., Structural basis of specific tRNA aminoacylation by a small in vitro selected ribozyme. *Nature* **2008**, *454* (7202), 358-361.
106. Murakami, H.; Bonzagni, N. J.; Suga, H., Aminoacyl-tRNA Synthesis by a Resin-Immobilized Ribozyme. *Journal of the American Chemical Society* **2002**, *124* (24), 6834-6835.
107. (a) Duffy, N. H.; Dougherty, D. A., Preparation of Translationally Competent tRNA by Direct Chemical Acylation. *Organic Letters* **2010**, *12* (17), 3776-3779; (b) Tzvetkova, S.; Kluger, R., Biomimetic Aminoacylation of Ribonucleotides and RNA with Aminoacyl Phosphate Esters and Lanthanum Salts. *Journal of the American Chemical Society* **2007**, *129* (51), 15848-15854.
108. (a) Carlson, E. D.; Gan, R.; Hodgman, C. E.; Jewett, M. C., Cell-free protein synthesis: Applications come of age. *Biotechnology Advances* **2012**, *30* (5), 1185-1194; (b) Harris, D. C.; Jewett, M. C., Cell-free biology: exploiting the interface between synthetic biology and synthetic chemistry. *Current Opinion in Biotechnology* **2012**, *23* (5), 672-678; (c) Katzen, F.; Chang, G.; Kudlicki, W., The past, present and future of cell-free protein synthesis. *Trends in Biotechnology* **2005**, *23* (3), 150-156.
109. (a) Co-translational association of cell-free expressed membrane proteins with supplied lipid bilayers. *Molecular Membrane Biology* **2013**, *30* (1), 75-89; (b) Junge, F.; Schneider, B.; Reckel, S.; Schwarz, D.; Dötsch, V.; Bernhard, F., Large-scale production of functional membrane proteins. *Cell. Mol. Life Sci.* **2008**, *65* (11), 1729-1755; (c) Katzen, F.; Peterson, T. C.; Kudlicki, W., Membrane protein expression: no cells required. *Trends in Biotechnology* **2009**, *27* (8), 455-460; (d) Schwarz, D.; Dötsch, V.; Bernhard, F., Production of membrane proteins using cell-free expression systems. *PROTEOMICS* **2008**, *8* (19), 3933-3946.
110. (a) Amstutz, P.; Forrer, P.; Zahnd, C.; Plückthun, A., In vitro display technologies: novel developments and applications. *Current Opinion in Biotechnology* **2001**, *12* (4), 400-405; (b) He, M.; Taussig, M. J., Ribosome display: Cell-free protein display technology. *Briefings in Functional Genomics & Proteomics* **2002**, *1* (2), 204-212; (c) Lipovsek, D.; Plückthun, A., In-vitro protein evolution by ribosome display and mRNA display. *Journal of Immunological Methods* **2004**, *290* (1-2), 51-67; (d) Schaffitzel, C.; Hanes, J.; Jermutus, L.; Plückthun, A., Ribosome display: an in vitro method for selection and evolution of antibodies from libraries. *Journal of Immunological Methods* **1999**, *231* (1-2), 119-135.
111. He, M.; Stoevesandt, O.; Taussig, M. J., In situ synthesis of protein arrays. *Current Opinion in Biotechnology* **2008**, *19* (1), 4-9.
112. (a) Calhoun, K. A.; Swartz, J. R., An Economical Method for Cell-Free Protein Synthesis using Glucose and Nucleoside Monophosphates. *Biotechnology Progress* **2005**, *21* (4), 1146-1153; (b) Calhoun, K. A.; Swartz, J. R., Energizing cell-free protein synthesis with glucose metabolism. *Biotechnology and Bioengineering* **2005**, *90* (5), 606-613; (c) Jewett, M. C.; Calhoun, K. A.; Voloshin, A.; Wu, J. J.; Swartz, J. R., An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol Syst Biol* **2008**, *4*; (d) Jewett, M. C.; Miller, M. L.; Chen, Y.; Swartz, J. R., Continued Protein Synthesis at Low [ATP] and [GTP] Enables Cell Adaptation during Energy Limitation. *Journal of Bacteriology* **2009**, *191* (3), 1083-1091.

113. Liu, D. V.; Zawada, J. F.; Swartz, J. R., Streamlining Escherichia Coli S30 Extract Preparation for Economical Cell-Free Protein Synthesis. *Biotechnology Progress* **2005**, *21* (2), 460-465.
114. Calhoun, K. A.; Swartz, J. R., Total amino acid stabilization during cell-free protein synthesis reactions. *Journal of Biotechnology* **2006**, *123* (2), 193-203.
115. Shirokov, V.; Kommer, A.; Kolb, V.; Spirin, A., Continuous-Exchange Protein-Synthesizing Systems. In *In Vitro Transcription and Translation Protocols*, Grandi, G., Ed. Humana Press: 2007; Vol. 375, pp 19-55.
116. (a) Goerke, A. R.; Swartz, J. R., Development of cell-free protein synthesis platforms for disulfide bonded proteins. *Biotechnology and Bioengineering* **2008**, *99* (2), 351-367; (b) Kim, D.-M.; Swartz, J. R., Efficient production of a bioactive, multiple disulfide-bonded protein using modified extracts of Escherichia coli. *Biotechnology and Bioengineering* **2004**, *85* (2), 122-129; (c) Yin, G.; Swartz, J. R., Enhancing multiple disulfide bonded protein folding in a cell-free system. *Biotechnology and Bioengineering* **2004**, *86* (2), 188-195.
117. (a) Jewett, M. C.; Swartz, J. R., Mimicking the Escherichia coli cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnology and Bioengineering* **2004**, *86* (1), 19-26; (b) Jewett, M. C.; Swartz, J. R., Substrate replenishment extends protein synthesis with an in vitro translation system designed to mimic the cytoplasm. *Biotechnology and Bioengineering* **2004**, *87* (4), 465-471.
118. Swartz, J. R., Universal cell-free protein synthesis. *Nat Biotech* **2009**, *27* (8), 731-732.
119. (a) Gilmore, M.; Steward, L.; Chamberlin, A. R., Incorporation of Noncoded Amino Acids by In Vitro Protein Biosynthesis. In *Implementation and Redesign of Catalytic Function in Biopolymers*, Schmidtchen, F.; Baltzer, L.; Chamberlin, A. R.; McDonnell, K. A.; Famulok, M.; Gilmore, M. A.; Imperiali, B.; Jenne, A.; Shogren-Knaak, M.; Steward, L. E., Eds. Springer Berlin Heidelberg: 1999; Vol. 202, pp 77-99; (b) Hohsaka, T.; Sisido, M., Incorporation of non-natural amino acids into proteins. *Current Opinion in Chemical Biology* **2002**, *6* (6), 809-815; (c) Mendel, D.; Cornish, V. W.; Schultz, P. G., Site-Directed Mutagenesis with an Expanded Genetic Code. *Annual Review of Biophysics and Biomolecular Structure* **1995**, *24* (1), 435-462.
120. Noble, C. G.; Song, H., Structural studies of elongation and release factors. *Cell. Mol. Life Sci.* **2008**, *65* (9), 1335-1346.
121. Short, G. F.; Golovine, S. Y.; Hecht, S. M., Effects of Release Factor 1 on in Vitro Protein Translation and the Elaboration of Proteins Containing Unnatural Amino Acids†. *Biochemistry* **1999**, *38* (27), 8808-8819.
122. Szkaradkiewicz, K.; Nanninga, M.; Nesper-Brock, M.; Gerrits, M.; Erdmann, V. A.; Sprinzl, M., RNA aptamers directed against release factor 1 from Thermus thermophilus. *FEBS Letters* **2002**, *514* (1), 90-95.
123. Hohsaka, T.; Ashizuka, Y.; Taira, H.; Murakami, H.; Sisido, M., Incorporation of Nonnatural Amino Acids into Proteins by Using Various Four-Base Codons in an Escherichia coli in Vitro Translation System. *Biochemistry* **2001**, *40* (37), 11060-11064.
124. Dougherty, D. A., Unnatural amino acids as probes of protein structure and function. *Current Opinion in Chemical Biology* **2000**, *4* (6), 645-652.
125. Wang, L.; Schultz, P. G., Expanding the genetic code. *Chemical Communications* **2002**, (1), 1-11.

126. (a) Pastrnak, M.; Magliery, T. J.; Schultz, P. G., A New Orthogonal Suppressor tRNA/Aminoacyl-tRNA Synthetase Pair for Evolving an Organism with an Expanded Genetic Code. *Helvetica Chimica Acta* **2000**, *83* (9), 2277-2286; (b) Santoro, S. W.; Wang, L.; Herberich, B.; King, D. S.; Schultz, P. G., An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nat Biotech* **2002**, *20* (10), 1044-1048; (c) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G., Expanding the Genetic Code of Escherichia coli. *Science* **2001**, *292* (5516), 498-500; (d) Wang, L.; Schultz, P. G., A general approach for the generation of orthogonal tRNAs. *Chemistry & Biology* **2001**, *8* (9), 883-890.
127. Liu, D. R.; Magliery, T. J.; Pastrnak, M.; Schultz, P. G., Engineering a tRNA and aminoacyl-tRNA synthetase for the site-specific incorporation of unnatural amino acids into proteins in vivo. *Proceedings of the National Academy of Sciences* **1997**, *94* (19), 10092-10097.
128. Arts, G.-J.; Kuersten, S.; Romby, P.; Ehresmann, B.; Mattaj, I. W., The role of exportin-t in selective nuclear export of mature tRNAs. *EMBO J* **1998**, *17* (24), 7430-7441.
129. (a) Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z.; Schultz, P. G., An Expanded Eukaryotic Genetic Code. *Science* **2003**, *301* (5635), 964-967; (b) Chin, J. W.; Cropp, T. A.; Chu, S.; Meggers, E.; Schultz, P. G., Progress Toward an Expanded Eukaryotic Genetic Code. *Chemistry & Biology* **2003**, *10* (6), 511-519.
130. Wang, L.; Magliery, T. J.; Liu, D. R.; Schultz, P. G., A New Functional Suppressor tRNA/Aminoacyl-tRNA Synthetase Pair for the in Vivo Incorporation of Unnatural Amino Acids into Proteins. *Journal of the American Chemical Society* **2000**, *122* (20), 5010-5011.
131. Ohno, S.; Yokogawa, T.; Fujii, I.; Asahara, H.; Inokuchi, H.; Nishikawa, K., Co-Expression of Yeast Amber Suppressor tRNA^{Tyr} and Tyrosyl-tRNA Synthetase in Escherichia coli: Possibility to Expand the Genetic Code. *Journal of Biochemistry* **1998**, *124* (6), 1065-1068.
132. Furter, R., Expansion of the genetic code: Site-directed p-fluoro-phenylalanine incorporation in Escherichia coli. *Protein Science* **1998**, *7* (2), 419-426.
133. Kowal, A. K.; Köhrer, C.; RajBhandary, U. L., Twenty-first aminoacyl-tRNA synthetase-suppressor tRNA pairs for possible use in site-specific incorporation of amino acid analogues into proteins in eukaryotes and in eubacteria. *Proceedings of the National Academy of Sciences* **2001**, *98* (5), 2268-2273.
134. Anderson, J. C.; Schultz, P. G., Adaptation of an Orthogonal Archaeal Leucyl-tRNA and Synthetase Pair for Four-base, Amber, and Opal Suppression†. *Biochemistry* **2003**, *42* (32), 9598-9608.
135. 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), 7566-7571.
136. Krzycki, J. A., The direct genetic encoding of pyrrolysine. *Current Opinion in Microbiology* **2005**, *8* (6), 706-712.
137. Yanagisawa, T.; Ishii, R.; Fukunaga, R.; Kobayashi, T.; Sakamoto, K.; Yokoyama, S., Multistep Engineering of Pyrrolysyl-tRNA Synthetase to Genetically Encode N ϵ -(o-Azidobenzoyloxycarbonyl) lysine for Site-Specific Protein Modification. *Chemistry & Biology* **2008**, *15* (11), 1187-1197.

138. Nozawa, K.; O'Donoghue, P.; Gundllapalli, S.; Araiso, Y.; Ishitani, R.; Umehara, T.; Soll, D.; Nureki, O., Pyrrolysyl-tRNA synthetase-tRNA^{Pyl} structure reveals the molecular basis of orthogonality. *Nature* **2009**, *457* (7233), 1163-1167.
139. (a) Brustad, E.; Bushey, M. L.; Brock, A.; Chittuluru, J.; Schultz, P. G., A promiscuous aminoacyl-tRNA synthetase that incorporates cysteine, methionine, and alanine homologs into proteins. *Bioorganic & Medicinal Chemistry Letters* **2008**, *18* (22), 6004-6006; (b) Lee, H. S.; Guo, J.; Lemke, E. A.; Dimla, R. D.; Schultz, P. G., Genetic Incorporation of a Small, Environmentally Sensitive, Fluorescent Probe into Proteins in *Saccharomyces cerevisiae*. *Journal of the American Chemical Society* **2009**, *131* (36), 12921-12923; (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. *Nat Chem Biol* **2007**, *3* (12), 769-772; (d) Summerer, D.; Chen, S.; Wu, N.; Deiters, A.; Chin, J. W.; Schultz, P. G., A genetically encoded fluorescent amino acid. *Proceedings of the National Academy of Sciences* **2006**, *103* (26), 9785-9789; (e) Tippmann, E. M.; Schultz, P. G., A genetically encoded metallocene containing amino acid. *Tetrahedron* **2007**, *63* (27), 6182-6184; (f) 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.
140. (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) Edwards, H.; Schimmel, P., A bacterial amber suppressor in *Saccharomyces cerevisiae* is selectively recognized by a bacterial aminoacyl-tRNA synthetase. *Molecular and Cellular Biology* **1990**, *10* (4), 1633-1641.
141. Young, T. S.; Ahmad, I.; Brock, A.; Schultz, P. G., Expanding the Genetic Repertoire of the Methylotrophic Yeast *Pichia pastoris*†. *Biochemistry* **2009**, *48* (12), 2643-2653.
142. (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 Edition* **2009**, *48* (22), 4052-4055; (b) 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.
143. (a) Liu, W.; Brock, A.; Chen, S.; Chen, S.; Schultz, P. G., Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat Meth* **2007**, *4* (3), 239-244; (b) Sakamoto, K.; Hayashi, A.; Sakamoto, A.; Kiga, D.; Nakayama, H.; Soma, A.; Kobayashi, T.; Kitabatake, M.; Takio, K.; Saito, K.; Shirouzu, M.; Hirao, I.; Yokoyama, S., Site - specific incorporation of an unnatural amino acid into proteins in mammalian cells. *Nucleic Acids Research* **2002**, *30* (21), 4692-4699.
144. Greiss, S.; Chin, J. W., Expanding the Genetic Code of an Animal. *Journal of the American Chemical Society* **2011**, *133* (36), 14196-14199.
145. Parrish, A. R.; She, X.; Xiang, Z.; Coin, I.; Shen, Z.; Briggs, S. P.; Dillin, A.; Wang, L., Expanding the Genetic Code of *Caenorhabditis elegans* Using Bacterial Aminoacyl-tRNA Synthetase/tRNA Pairs. *ACS Chemical Biology* **2012**, *7* (7), 1292-1302.
146. Liu, C. C.; Schultz, P. G., Adding New Chemistries to the Genetic Code. *Annual Review of Biochemistry* **2010**, *79* (1), 413-444.

147. (a) Xie, J.; Schultz, P. G., A chemical toolkit for proteins [mdash] an expanded genetic code. *Nat Rev Mol Cell Biol* **2006**, *7* (10), 775-782; (b) Young, T. S.; Schultz, P. G., Beyond the Canonical 20 Amino Acids: Expanding the Genetic Lexicon. *Journal of Biological Chemistry* **2010**, *285* (15), 11039-11044.
148. Wang, K.; Neumann, H.; Peak-Chew, S. Y.; Chin, J. W., Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. *Nat Biotech* **2007**, *25* (7), 770-777.
149. (a) Johnson, D. B. F.; Wang, C.; Xu, J.; Schultz, M. D.; Schmitz, R. J.; Ecker, J. R.; Wang, L., Release Factor One Is Nonessential in Escherichia coli. *ACS Chemical Biology* **2012**, *7* (8), 1337-1344; (b) Johnson, D. B. F.; 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), 779-786.
150. 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), 1151-1154.
151. An, W.; Chin, J. W., Synthesis of orthogonal transcription-translation networks. *Proceedings of the National Academy of Sciences* **2009**, *106* (21), 8477-8482.
152. (a) Rackham, O.; Chin, J. W., Cellular Logic with Orthogonal Ribosomes. *Journal of the American Chemical Society* **2005**, *127* (50), 17584-17585; (b) Rackham, O.; Wang, K.; Chin, J. W., Functional epitopes at the ribosome subunit interface. *Nat Chem Biol* **2006**, *2* (5), 254-258.
153. Sewald, N.; Hollweck, W.; Mütze, K.; Schierlinger, C.; Seymour, L. C.; Gaa, K.; Burger, K.; Koksche, B.; Jakubke, H. D., Peptide modification by introduction of α -trifluoromethyl substituted amino acids. *Amino Acids* **1995**, *8* (2), 187-194.
154. Thorson, J.; Cornish, V.; Barrett, J.; Cload, S.; Yano, T.; Schultz, P., A Biosynthetic Approach for the Incorporation of Unnatural Amino Acids into Proteins. In *Protein Synthesis*, Martin, R., Ed. Springer New York: 1998; Vol. 77, pp 43-73.
155. (a) Kao, C.; Zheng, M.; Rüdiger, S., A simple and efficient method to reduce nontemplated nucleotide addition at the 3' terminus of RNAs transcribed by T7 RNA polymerase. *RNA* **1999**, *5* (9), 1268-1272; (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 Research* **1987**, *15* (21), 8783-8798.
156. Gerrits, M., Funktion und Effizienz von amber-Suppressor-tRNA in der zellfreien Proteinbiosynthese. *Dissertation Freie Universität Berlin* **2002**.
157. Köhrer, C.; RajBhandary, U. L., The many applications of acid urea polyacrylamide gel electrophoresis to studies of tRNAs and aminoacyl-tRNA synthetases. *Methods* **2008**, *44* (2), 129-138.
158. Petersson, E. J.; Shahgholi, M.; Lester, H. A.; Dougherty, D. A., MALDI-TOF mass spectrometry methods for evaluation of in vitro aminoacyl tRNA production. *RNA* **2002**, *8* (4), 542-547.
159. M, N. D. L. a. C. M., Lehninger Principles of Biochemistry, Fourth Edition. **2004**.
160. Ellman, J.; Mendel, D.; Anthony-Cahill, S.; Noren, C. J.; Schultz, P. G., [15] Biosynthetic method for introducing unnatural amino acids site-specifically into proteins. In *Methods in Enzymology*, John, J. L., Ed. Academic Press: 1991; Vol. Volume 202, pp 301-336.

161. (a) Butt, T. R.; Edavettal, S. C.; Hall, J. P.; Mattern, M. R., SUMO fusion technology for difficult-to-express proteins. *Protein Expression and Purification* **2005**, *43* (1), 1-9; (b) Malakhov, M.; Mattern, M.; Malakhova, O.; Drinker, M.; Weeks, S.; Butt, T., SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *J Struct Func Genom* **2004**, *5* (1-2), 75-86; (c) Panavas, T.; Sanders, C.; Butt, T. R., SUMO Fusion Technology for Enhanced Protein Production in Prokaryotic and Eukaryotic Expression Systems #. In *T SUMO Protocols*, 2008; Vol. 497, pp 303-317; (d) Zuo, X.; Li, S.; Hall, J.; Mattern, M.; Tran, H.; Shoo, J.; Tan, R.; Weiss, S.; Butt, T., Enhanced Expression and Purification of Membrane Proteins by SUMO Fusion in Escherichia coli. *J Struct Func Genom* **2005**, *6* (2-3), 103-111.
162. Sun, Z.; Lu, W.; Jiang, A.; Chen, J.; Tang, F.; Liu, J.-N., Expression, purification and characterization of aprotinin and a human analogue of aprotinin. *Protein Expression and Purification* **2009**, *65* (2), 238-243.
163. Klammt, C., Functional and structural analysis of cell-free produced transporters and G-protein coupled receptors: Development of new techniques for the fast and efficient production of integral membrane proteins. *DíSSERTATION Johann Wolfgang Goethe Universität in Frankfurt am Main* **2006**.
164. Klammt, C.; Schwarz, D.; Eifler, N.; Engel, A.; Piehler, J.; Haase, W.; Hahn, S.; Dötsch, V.; Bernhard, F., Cell-free production of G protein-coupled receptors for functional and structural studies. *Journal of Structural Biology* **2007**, *158* (3), 482-493.
165. Betton, J.-M.; Miot, M., Cell-Free Production of Membrane Proteins in the Presence of Detergents. In *Cell-Free Protein Synthesis*, Wiley-VCH Verlag GmbH & Co. KGaA: 2008; pp 165-178.
166. Shimizu, Y.; Inoue, A.; Tomari, Y.; Suzuki, T.; Yokogawa, T.; Nishikawa, K.; Ueda, T., Cell-free translation reconstituted with purified components. *Nat Biotech* **2001**, *19* (8), 751-755.
167. Choudhury, A. K.; Golovine, S. Y.; Dedkova, L. M.; Hecht, S. M., Synthesis of Proteins Containing Modified Arginine Residues†. *Biochemistry* **2007**, *46* (13), 4066-4076.
168. (a) Winkler, D. B., K., Synthesis of Enantiomerically Pure D- and L-Armentomycin and its Difluoro Analogues from Aspartic Acid. *Synthesis* **1996**, (12), 1419-1421; (b) Tsushima, T.; Kawada, K.; Ishihara, S.; Uchida, N.; Shiratori, O.; Higaki, J.; Hirata, M., Fluorine containing amino acids and their derivatives. 7. Synthesis and antitumor activity of α - and γ -substituted methotrexate analogs. *Tetrahedron* **1988**, *44* (17), 5375-5387.
169. Kimmerlin, T.; Seebach, D., '100 years of peptide synthesis': ligation methods for peptide and protein synthesis with applications to β -peptide assemblies*. *The Journal of Peptide Research* **2005**, *65* (2), 229-260.
170. Merrifield, R. B., Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *Journal of the American Chemical Society* **1963**, *85* (14), 2149-2154.
171. (a) Dawson, P.; Muir, T.; Clark-Lewis, I.; Kent, S., Synthesis of proteins by native chemical ligation. *Science* **1994**, *266* (5186), 776-779; (b) Kent, S. B. H., Total chemical synthesis of proteins. *Chemical Society Reviews* **2009**, *38* (2), 338-351; (c) Muir, T. W.; Sondhi, D.; Cole, P. A., Expressed protein ligation: A general method for protein engineering. *Proceedings of the National Academy of Sciences* **1998**, *95* (12), 6705-6710;

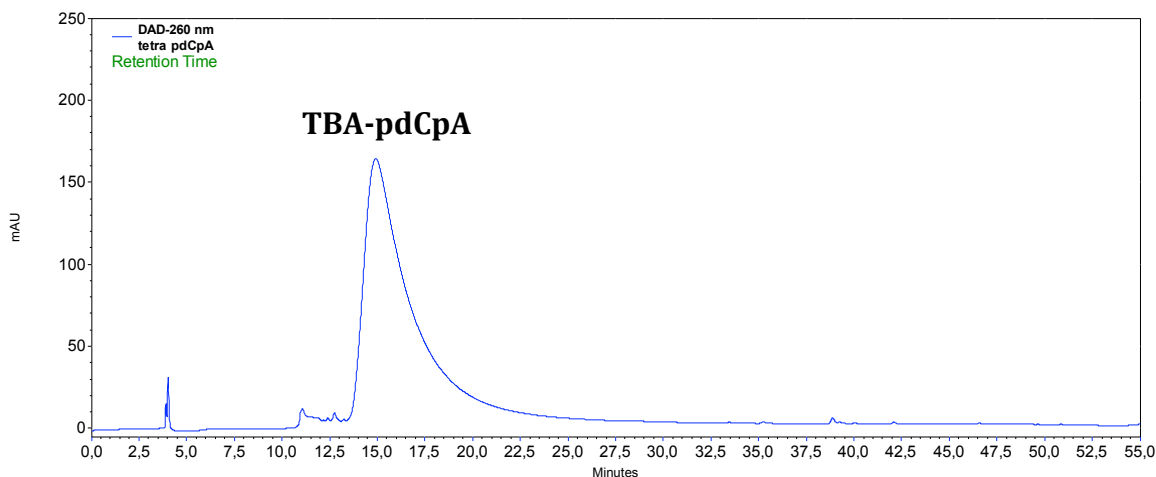
- (d) Muralidharan, V.; Muir, T. W., Protein ligation: an enabling technology for the biophysical analysis of proteins. *Nat Meth* **2006**, *3* (6), 429-438.
172. Meldal, M., [6] Properties of solid supports. In *Methods in Enzymology*, Gregg, B. F., Ed. Academic Press: 1997; Vol. Volume 289, pp 83-104.
173. (a) Carpino, L. A.; Han, G. Y., 9-Fluorenylmethoxycarbonyl function, a new base-sensitive amino-protecting group. *Journal of the American Chemical Society* **1970**, *92* (19), 5748-5749; (b) McKay, F. C.; Albertson, N. F., New Amine-masking Groups for Peptide Synthesis. *Journal of the American Chemical Society* **1957**, *79* (17), 4686-4690.
174. Isidro-Llobet, A.; Álvarez, M.; Albericio, F., Amino Acid-Protecting Groups. *Chemical Reviews* **2009**, *109* (6), 2455-2504.
175. (a) Albeicicio, F.; Chinchilla, R.; Dodsworth, D. J.; Nájera, C., NEW TRENDS IN PEPTIDE COUPLING REAGENTS. *Organic Preparations and Procedures International* **2001**, *33* (3), 203-303; (b) El-Faham, A.; Albericio, F., Peptide Coupling Reagents, More than a Letter Soup. *Chemical Reviews* **2011**, *111* (11), 6557-6602; (c) Han, S.-Y.; Kim, Y.-A., Recent development of peptide coupling reagents in organic synthesis. *Tetrahedron* **2004**, *60* (11), 2447-2467.
176. Hackenberger, C. P. R.; Schwarzer, D., Chemoselective Ligation and Modification Strategies for Peptides and Proteins. *Angewandte Chemie International Edition* **2008**, *47* (52), 10030-10074.
177. Dawson, P. E.; Kent, S. B. H., SYNTHESIS OF NATIVE PROTEINS BY CHEMICAL LIGATION 1. *Annual Review of Biochemistry* **2000**, *69* (1), 923-960.
178. (a) Canne, L. E.; Walker, S. M.; Kent, S. B. H., A general method for the synthesis of thioester resin linkers for use in the solid phase synthesis of peptide- α -thioacids. *Tetrahedron Letters* **1995**, *36* (8), 1217-1220; (b) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E., Protein synthesis by native chemical ligation: Expanded scope by using straightforward methodology. *Proceedings of the National Academy of Sciences* **1999**, *96* (18), 10068-10073.
179. Bu, X.; Xie, G.; Law, C. W.; Guo, Z., An improved deblocking agent for direct Fmoc solid-phase synthesis of peptide thioesters. *Tetrahedron Letters* **2002**, *43* (13), 2419-2422.
180. (a) von Eggelkraut-Gottanka, R.; Klose, A.; Beck-Sickinger, A. G.; Beyermann, M., Peptide α thioester formation using standard Fmoc-chemistry. *Tetrahedron Letters* **2003**, *44* (17), 3551-3554; (b) Zhu, J.; Wan, Q.; Ragupathi, G.; George, C. M.; Livingston, P. O.; Danishefsky, S. J., Biologics through Chemistry: Total Synthesis of a Proposed Dual-Acting Vaccine Targeting Ovarian Cancer by Orchestration of Oligosaccharide and Polypeptide Domains. *Journal of the American Chemical Society* **2009**, *131* (11), 4151-4158.
181. Sewing, A.; Hilvert, D., Fmoc-Compatible Solid-Phase Peptide Synthesis of Long C-Terminal Peptide Thioesters. *Angewandte Chemie International Edition* **2001**, *40* (18), 3395-3396.
182. Alsina, J.; Yokum, T. S.; Albericio, F.; Barany, G., Backbone Amide Linker (BAL) Strategy for $N\alpha$ -9-Fluorenylmethoxycarbonyl (Fmoc) Solid-Phase Synthesis of Unprotected Peptide p-Nitroanilides and Thioesters¹. *The Journal of Organic Chemistry* **1999**, *64* (24), 8761-8769.
183. Mende, F.; Seitz, O., 9-Fluorenylmethoxycarbonyl-Based Solid-Phase Synthesis of Peptide α -Thioesters. *Angewandte Chemie International Edition* **2011**, *50* (6), 1232-1240.

184. Lu, W.; Starovasnik, M. A.; Kent, S. B. H., Total chemical synthesis of bovine pancreatic trypsin inhibitor by native chemical ligation. *FEBS Letters* **1998**, *429* (1), 31-35.
185. Weissman, J. S.; Kim, P. S., The pro region of BPTI facilitates folding. *Cell* **1992**, *71* (5), 841-851.
186. Ferrer, M.; Woodward, C.; Barany, G., Solid-phase synthesis of bovine pancreatitis trypsin inhibitor (BPTI) and two analogues. A chemical approach for evaluating the role of disulfide bridges in protein folding and stability. *International Journal of Peptide and Protein Research* **1992**, *40* (3-4), 194-207.
187. Monera, O. D.; Kay, C. M.; Hodges, R. S., Protein denaturation with guanidine hydrochloride or urea provides a different estimate of stability depending on the contributions of electrostatic interactions. *Protein Science* **1994**, *3* (11), 1984-1991.
188. Beckmann, J.; Mehlich, A.; Schröder, W.; Wenzel, H. R.; Tschesche, H., Preparation of chemically 'mutated' aprotinin homologues by semisynthesis. *European Journal of Biochemistry* **1988**, *176* (3), 675-682.
189. (a) Czapinska, H.; Helland, R.; Smalås, A. O.; Otlewski, J., Crystal Structures of Five Bovine Chymotrypsin Complexes with P1 BPTI Variants. *Journal of Molecular Biology* **2004**, *344* (4), 1005-1020; (b) Helland, R.; Otlewski, J.; Sundheim, O.; Dadlez, M.; Smalås, A. O., The crystal structures of the complexes between bovine β -trypsin and ten P1 variants of BPTI. *Journal of Molecular Biology* **1999**, *287* (5), 923-942.
190. Velazquez-Campoy, A.; Leavitt, S.; Freire, E., Characterization of Protein-Protein Interactions by Isothermal Titration Calorimetry. In *Protein-Protein Interactions*, Fu, H., Ed. Humana Press: 2004; Vol. 261, pp 35-54.
191. Kawamura, K.; Yamada, T.; Kurihara, K.; Tamada, T.; Kuroki, R.; Tanaka, I.; Takahashi, H.; Niimura, N., X-ray and neutron protein crystallographic analysis of the trypsin-BPTI complex. *Acta crystallographica. Section D, Biological crystallography* **2011**, *67* (Pt 2), 140-8.
192. Barany, G.; Gross, C. M.; Ferrer, M.; Barbar, E.; Pan, H.; Woodward, C., Optimized methods for chemical synthesis of bovine pancreatic trypsin inhibitor (BPTI) analogues. In *Techniques in Protein Chemistry*, Daniel, R. M., Ed. Academic Press: 1996; Vol. Volume 7, pp 503-514.
193. (a) Greenfield, N. J., Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protocols* **2007**, *1* (6), 2876-2890; (b) Kelly, S. M.; Price, N. C., The use of circular dichroism in the investigation of protein structure and function. *Current protein & peptide science* **2000**, *1* (4), 349-84.
194. (a) Sreerama, N.; Woody, R. W., Computation and Analysis of Protein Circular Dichroism Spectra. In *Methods in Enzymology*, Ludwig, B.; Michael, L. J., Eds. Academic Press: 2004; Vol. Volume 383, pp 318-351; (b) Whitmore, L.; Wallace, B. A., Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* **2008**, *89* (5), 392-400.
195. (a) Greenfield, N. J., Analysis of the kinetics of folding of proteins and peptides using circular dichroism. *Nat. Protocols* **2007**, *1* (6), 2891-2899; (b) Greenfield, N. J., Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat. Protocols* **2007**, *1* (6), 2527-2535.

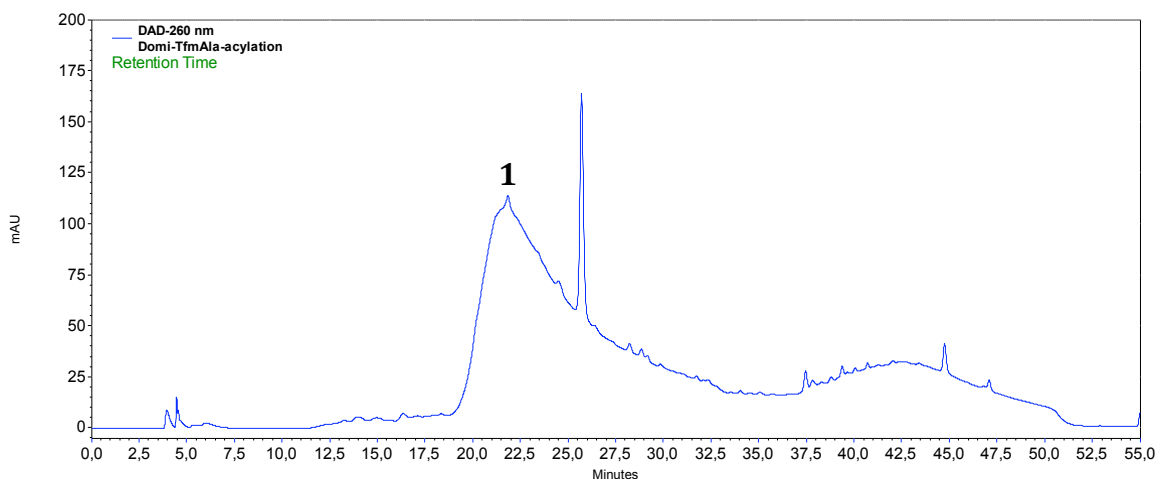
196. Sigma-Aldrich, Enzymatic Assay of Chymotrypsin (EC 3.4.21.1). *Sigma-Aldrich Protocol* <http://www.sigmaaldrich.com/technical-documents/protocols/biology/enzymatic-assay-of-chymotrypsin.html>.
197. Cs.-szabó, G.; Pozsgay, M.; Elovics, P., Investigation of the substrate-binding site of human plasmin using tripeptidyl-p-nitroanilide substrates. *Thrombosis Research* **1980**, *20* (2), 199-206.
198. Gravett, P. S.; Viljoen, C. C.; Oosthuizen, M. M. J., A steady-state kinetic analysis of the reaction between arginine esterase E-I from *Bitis gabonica* venom and synthetic arginine substrates and the influence of pH, temperature and solvent deuterium isotope. *International Journal of Biochemistry* **1991**, *23* (10), 1085-1099.
199. Pierce, M. M.; Raman, C. S.; Nall, B. T., Isothermal Titration Calorimetry of Protein-Protein Interactions. *Methods* **1999**, *19* (2), 213-221.
200. Zhou, X.; Kini, R. M.; Sivaraman, J., Application of isothermal titration calorimetry and column chromatography for identification of biomolecular targets. *Nat. Protocols* **2011**, *6* (2), 158-165.
201. Kabsch, W., XDS. *Acta Crystallographica Section D* **2010**, *66* (2), 125-132.
202. Hanson, W. M.; Domek, G. J.; Horvath, M. P.; Goldenberg, D. P., Rigidification of a Flexible Protease Inhibitor Variant upon Binding to Trypsin. *Journal of Molecular Biology* **2007**, *366* (1), 230-243.
203. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. *Journal of Applied Crystallography* **2007**, *40* (4), 658-674.
204. (a) Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **2010**, *66* (Pt 2), 213-21; (b) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D., Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **2012**, *68* (Pt 4), 352-67.
205. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **2010**, *66* (Pt 4), 486-501.
206. Chen, V. B.; Arendall, W. B., 3rd; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **2010**, *66* (Pt 1), 12-21.
207. DeLano, The PyMOL Molecular Graphics System. **2002**.
208. Kearney, P. C.; Zhang, H.; Zhong, W.; Dougherty, D. A.; Lester, H. A., Determinants of Nicotinic Receptor Gating in Natural and Unnatural Side Chain Structures at the M2 9' Position. *Neuron* **1996**, *17* (6), 1221-1229.
209. Völler, J. K., B., Unpublished results.

APPENDIX

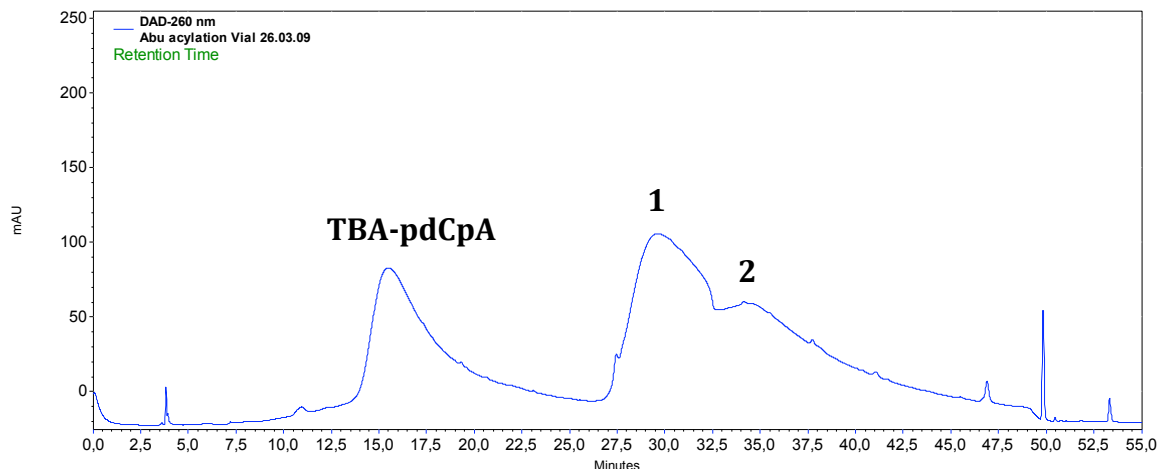
RP-HPLC spectrum of chemical aminoacylation of non-canonical amino acids



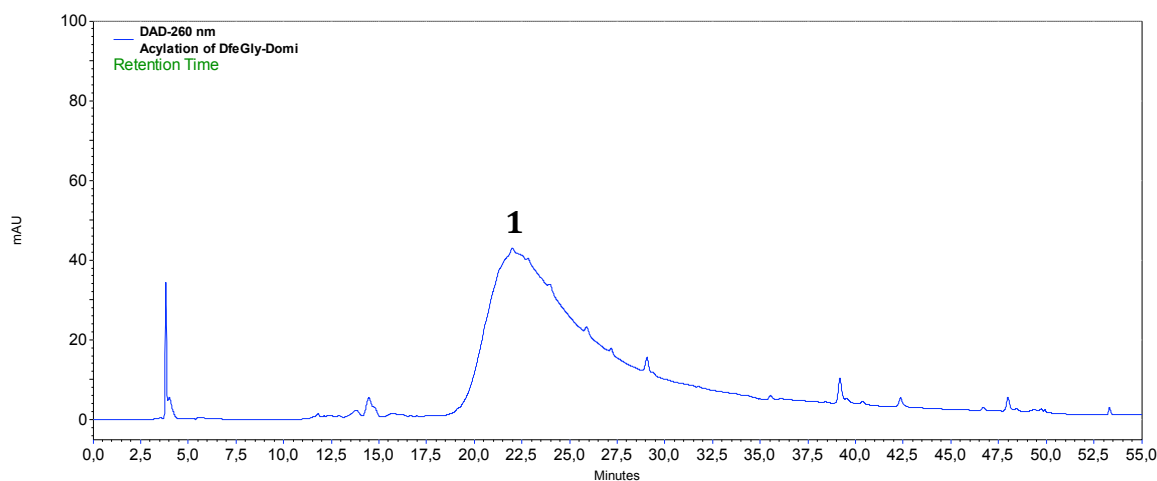
Analytic HPLC of *n*-tetrabutylammonium-pdCpA on C18 column (capcel pak, *Shisheido*), washed with 10-80% CH₃CN in 50 mM NH₄OAc (pH 4.5) over a period of 45 minutes at a flow rate of 1 mL/minute (monitoring at 260 nm).



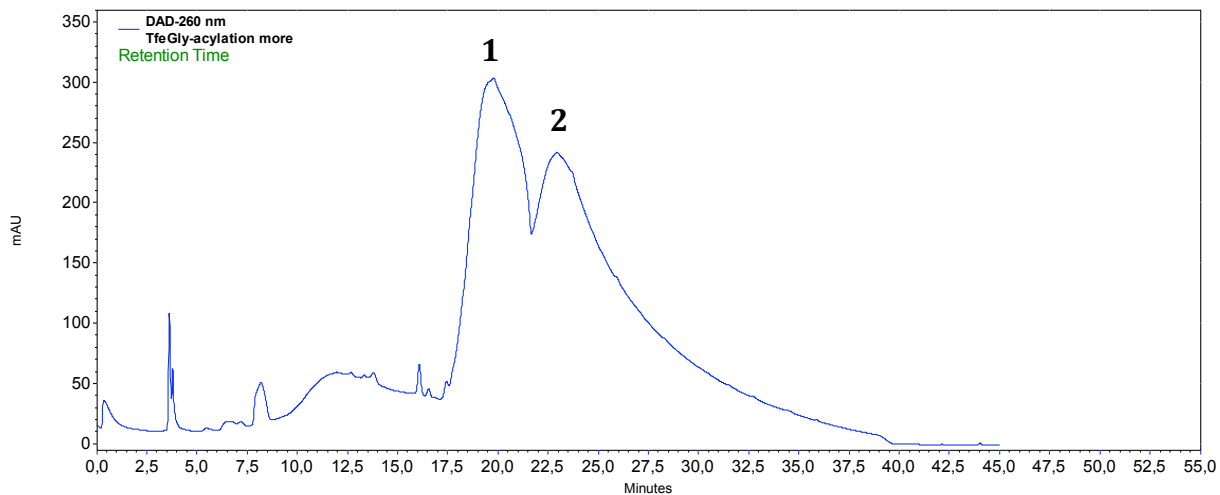
Analytic HPLC of misacylation of *N*-pentenoyl-TfmAla-pdCpA on C18 column (capcel pak, *Shisheido*), washed with 10-80% CH₃CN in 50 mM NH₄OAc (pH 4.5) over a period of 45 minutes at a flow rate of 1 mL/minute (monitoring at 260 nm); peak 1: *N*-pentenoyl-TfmAla-pdCpA.



Analytic HPLC of misacylation of *N*-pentenoyl-Abu-pdCpA on C18 column (capcel pak, *Shisheido*), washed with 10-80% CH₃CN in 50 mM NH₄OAc (pH 4.5) over a period of 45 minutes at a flow rate of 1 mL/minute (monitoring at 260 nm); peak **1**: *N*-pentenoyl-Abu-pdCpA, peak **2**: bis-*N*-pentenoyl-Abu-pdCpA.



Analytic HPLC of misacylation of *N*-pentenoyl-DfeGly-pdCpA on C18 column (capcel pak, *Shisheido*), washed with 10-80% CH₃CN in 50 mM NH₄OAc (pH 4.5) over a period of 45 minutes at a flow rate of 1 mL/minute (monitoring at 260 nm); peak **1**: *N*-pentenoyl-DfeGly-pdCpA.



Analytic HPLC of misacylation of *N*-pentenoyl-TfeGly-pdCpA on C18 column (capcel pak, *Shisheido*), washed with 10-80% CH₃CN in 50 mM NH₄OAc (pH 4.5) over a period of 35 minutes at a flow rate of 1 mL/minute (monitoring at 260 nm); peak **1**: *N*-pentenoyl-TfeGly-pdCpA, peak **2**: bis-*N*-pentenoyl-TfeGly-pdCpA.

DNA, protein sequences, cloning site of vectors used for *in vitro* suppression

Sequence of yeast phenylalanine suppressor tRNA:

5'- GTA CCG CTG CAG TAA TAC GAC TCA CTA TAG CGG ATT TAG CTC

3'- CAT GGC GAC GTC ATT ATG CTG AGT GAT ATC GCC TAA ATC GAG

T7 Promoter

AGT TGG GAG AGC GCC AGA CTC TAA ATC TGG AGG TCC TGT GTT CGA

TCA ACC CTC TCG CGG TCT GAG ATT TAG ACC TCC AGG ACA CAA GCT

amber stop codon

TCC ACA GAA TTC GCA CCA GGT GAT CCA TCC AAG CT -3'

AGG TGT CTT AAG CGT GGT CCA CTA GGT AGG TTC GA -5'

BstN I site FokI recognition site

Designed SUMO-BPTI sequence:

ATG TCT GAC CAG GAG GCA AAA CCT TCA ACT GAG GAC

M S D Q E A K P S T E D

TTG GGG GAT AAG AAG GAA GGT GAA TAT ATT AAA CTC

L G D K K E G E Y I K L

AAA GTC ATT GGA CAG GAT AGC AGT GAG ATT CAC TTC

K V I G Q D S S E I H F

AAA GTG AAA ATG ACA ACA CAT CTC AAG AAA CTC AAA

K V K M T T H L K K L K

GAA TCA TAC TGT CAA AGA CAG GGT GTT CCA ATG AAT

E S Y C Q R Q G V P M N

TCA CTC AGG TTT CTC TTT GAG GGT CAG AGA ATT GCT
 S L R F L F E G Q R I A
 GAT AAT CAT ACT CCA AAA GAA CTG GGA ATG GAG GAA
 D N H T P K E L G M E E
 GAA GAT GTG ATT GAA GTT TAT CAG GAA CAA ACG GGG
 E D V I E V Y Q E Q T G
 GGT CAT TCA ACA GTT CGG CCT GAC TTC TGC CTA GAG
 G H S T V R P D F C L E
 CCT CCA TAT ACG GGT CCC TGC TAG GCC AGA ATT ATC
 P P Y T G P C nAA A R I I
 AGA TAC TTC TAC AAC GCC AAG GCT GGG CTC TGC CAG
 R Y F Y N A K A G L C Q
 ACC TTT GTA TAT GGC GGC TGC AGA GCT AAA AGA AAC
 T F V Y G G C R A K R N
 AAT TTC AAG AGC GCA GAG GAC TGC ATG AGG ACC TGT
 N F K S A E D C M R T C
 GGT GGT GCT TAA
 G G A Stop

Protein sequence of Homo sapiens trypsin 1 ([UniProtKB: P07477](http://UniProtKB.org/entry/UniProtKB:P07477)):

MNPLLLITFV	AAALAAPFDD	DDKIVGGYNC	EENSVPYQVS
LNSGYHFCGG	SLINEQWVVS	AGHCYKSRIQ	VRLGEHNIEV
LEGNEQFINA	AKIIRHPQYD	RKTLNNDIML	IKLSSRAVIN
ARVSTISLPT	APPATGTKCL	ISGWGNTASS	GADYPDELQC
LDAPVLSQAK	CEASYPGKIT	SNMFCVGFLE	GGKDSCQGDS
GGPVVCNGQL	QGVVSWGDC	AQKNKPGVYT	KVYNYVKWIK
NTIAANS			

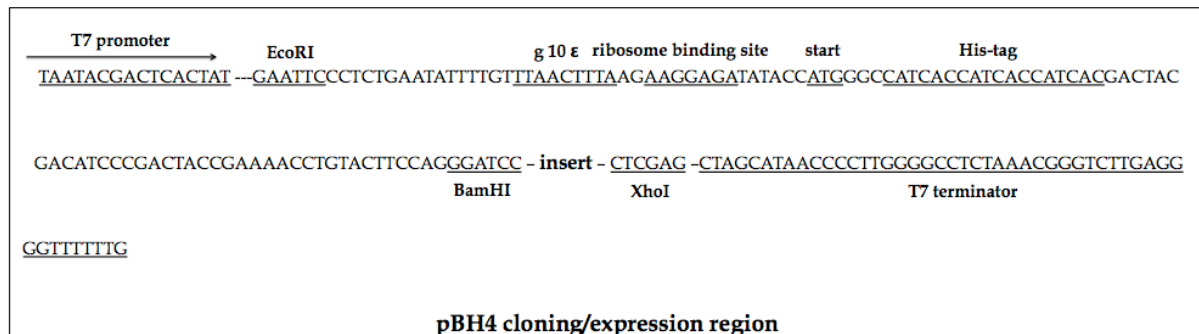
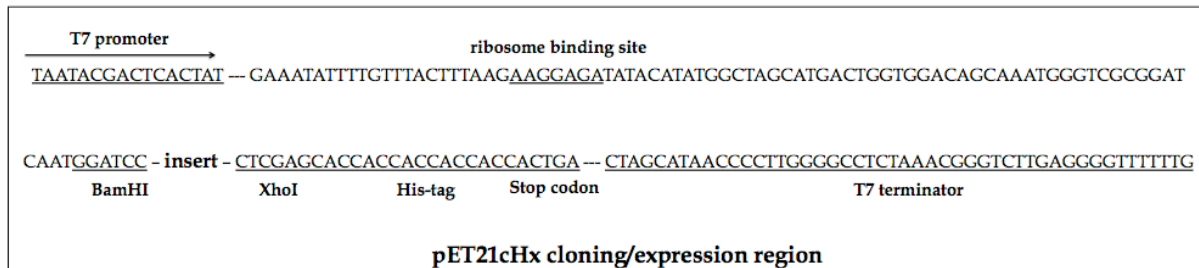
Protein sequence of Homo sapiens chymotrypsinogen B1 (UniProtKB: P17538):

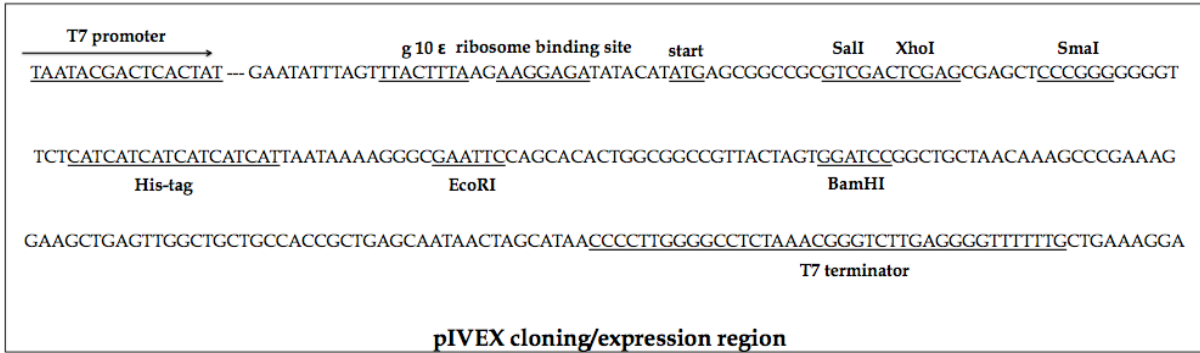
MAFLWLLSCW	ALLGTTFGCG	VPAIHPVLSG	LSRIVNGEDA
VPGSWPWQVS	LQDKTGFHFC	GGSLISEDWV	VTAAHCGVRT
SDVVVAGEFD	QGSDEENIQV	LKIAKVFKNP	KFSILTVNND
ITLLKLATPA	RFSQTVSAVC	LPSADDDFPA	GTLCATTGWG
KTKYNANKTP	DKLQQAALPL	LSNAECKKSW	GRRITDVMIC
AGASGVSSCM	GDSGGPLVCQ	KDGAWTLVGI	VSWGSDTCST
SSPGVYARVT	KLIPWVQKIL	AAN	

Vector diagram of pCMV6-XL4/XL5 vectors:

<http://www.origene.com/cdna/vectordiagram.msp> (Accessed 20 March 2014)

Cloning and expression region of expression vectors:





Cleavage site of restriction endonucleases used in the studies:

BstNI recognition and cleavage site:

5'... CC ↓X GG ... 3'

3'... GG X ↑CC ... 5'

FokI recognition and cleavage site:

5' ... GGATG (N)₉ ↓ ... 3'

3' ... CCTAC (N)₁₃ ↑ ... 5'

BamHI recognition and cleavage site:

5'... G ↓GATCC ... 3'

3'... CCTAG ↑G ... 5'

XhoI recognition and cleavage site:

5' ... C ↓TCGAG ... 3'

3' ... GAGCT ↑C ... 5'

DpnI recognition and cleavage site:

5'... G ↓ATC ... 3'

3'... CTA ↑G ... 5' The adenine is methylated.

Manual of commercially available kits used for studies:

Run-off *in vitro* transcription kit:

AmpliScribe™ T7 High Yield transcription Kit (*EPICENTRE® Biotechnologies*)

<http://www.epibio.com/docs/default-source/protocols/ampliscribe-t7-high-yield-transcription-kit.pdf?sfvrsn=6> (Accessed 20 march 2014)

DNA purification kit:

peqGOLD Plasmid Mini Kit (*PEQLAB Biotechnologie GmbH*):

http://www.peqlab.de/wcms/de/pdf/12-6942-02_m.pdf (Accessed 20 march 2014)

peqGOLD Gel Extraction Kit (*PEQLAB Biotechnologie GmbH*):

http://www.peqlab.de/wcms/de/pdf/12-2501-02_m.pdf (Accessed 20 march 2014)

Pure® HiPure Plasmid Midprep Kit (*Invitrogen*):

http://tools.invitrogen.com/content/sfs/manuals/purelink_hipure_plasmid_dna_purification_man.pdf (Accessed 20 march 2014)

Cell-free protein synthesis kits:

S30 T7 High-Yield Protein Expression System (*Promega*):

<http://www.promega.de/~media/Files/Resources/Protocols/Technical%20Manuals/0/S30%20T7%20High%20Yield%20Protein%20Expression%20System%20Protocol.pdf> (Accessed 20 march 2014)

RTS™ 100 *E.coli* Disulfide Kit (*5 PRIME*)

http://www.5prime.com/media/404635/rts%20100%20ecoli%20disulfide%20manual_5prime_1059726_%20091210.pdf (Accessed 20 march 2014)

TNT® Quick Coupled Transcription/Translation System (*Promega*)

<http://www.promega.de/~media/Files/Resources/Protocols/Technical%20Manuals/0/TNT%20Quick%20Coupled%20Transcription%20Translation%20Systems%20Protocol.pdf> (Accessed 20 march 2014)

PURExpress® ΔRF123 Kit (*New England Biolabs*)

<https://www.neb.com/products/e6850-purexpress-rf123-kit#tabselect2> (Accessed 20 march 2014)

Protein purification and analysis kits:

Ni-NTA Spin Columns (*Qiagen*)

<http://www.qiagen.com/resources/Download.aspx?id={3FC8C76D-6D21-4887-9BF8-F35F78FCC2F2}&lang=en&ver=1> (Accessed 20 march 2014)

QIAexpress® Detection and Assay Handbook for Anti•His HRP Conjugates

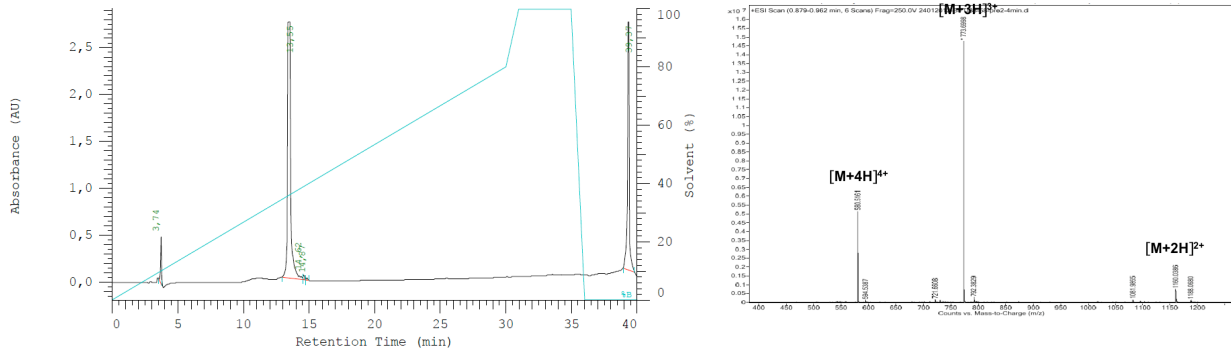
<http://www.qiagen.com/resources/Download.aspx?id={A8D20BF6-B88C-436B-A7BE-DE66B5DD70FD}&lang=en&ver=1> (Accessed 20 march 2014)

Table Peptide fragments of SUMO-BPTI detected by protein fingerprinting

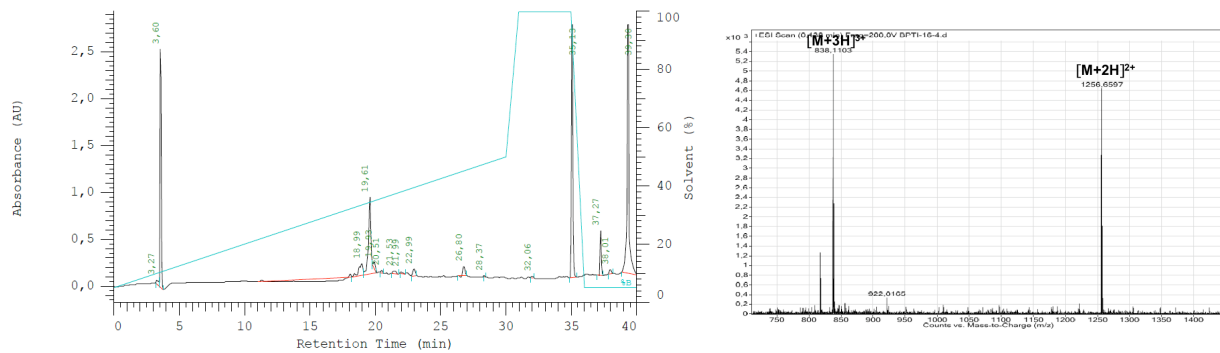
Start	End	Observed	Mr(expt)	Mr(calc)	Miss	Sequence	Ion Score	
							Abu	TfeGly
31	39	481.229	960.444	960.440	0	K.PSTEDLGDK.K	67	66
31	40	545.274	1088.533	1088.535	1	K.PSTEDLGDKK.E	51	50
31	46	603.635	1807.883	1807.884	2	K.PSTEDLGDKKEGEYIK.L	45	70
40	46	433.734	865.453	865.455	1	K.KEGEYIK.L	20	15
41	48	490.276	978.538	978.539	1	K.EGEYIKLK.V	40	42
47	60	534.294	1599.860	1599.862	1	K.LKVGQDSSEIHFK.V	58	62
47	62	610.015	1827.022	1827.025	2	K.LKVGQDSSEIHFKVK.M	54	
49	60	453.901	1358.682	1358.683	0	K.VIGQDSSEIHFK.V	54	59
49	62	529.622	1585.844	1585.846	1	K.VIGQDSSEIHFKVK.M	43	
61	68	479.281	956.547	956.548	1	K.VKMTTHLK.K	51	48
63	69	429.747	857.480	857.479	1	K.MTTHLKK.L	30	29
63	69	437.743	873.472	873.474	1	K.MTTHLKK.L Oxidation	26	27
69	77	404.545	1210.612	1210.613	2	K.KLKESYCQR.Q Carbamidomethyl	30	
70	77	513.755	1025.496	1025.496	1	K.LKESYCQR.Q	43	38
70	77	542.266	1082.517	1082.518	1	K.LKESYCQR.Q Carbamidomethyl	52	39
70	77	549.274	1096.534	1096.534	1	K.LKESYCQR.Q Propionamide	52	23
72	77	421.678	841.341	841.339	0	K.ESYCQR.Q Carbamidomethyl	30	20
72	77	428.686	855.358	855.355	0	K.ESYCQR.Q Propionamide	21	27
105	110	403.198	804.381	804.381	0	R.YFYNAK.A	18	9

Start	End	Observed	Mr(expt)	Mr(calc)	Miss	Sequence	Ion Score	
							Abu	TfeGly
1	17	626.963	1877.866	1877.868	2	-.MSDQEAKPSTEDLGDKK.E	59	52
1	17	632.295	1893.863	1893.863	2	-.MSDQEAKPSTEDLGDKK.E Oxidation (M)	46	
2	17	583.283	1746.827	1746.827	2	M.SDQEAKPSTEDLGDKK.E	54	
55	63	501.263	1000.511	1000.512	0	R.QGVPMNSLR.F	47	45
55	63	509.260	1016.506	1016.507	0	R.QGVPMNSLR.F Oxidation (M)	56	46
64	70	448.734	895.454	895.455	0	R.FLFEGQR.I	46	45
71	78	448.234	894.453	894.456	0	R.IADNHTPK.E	54	49

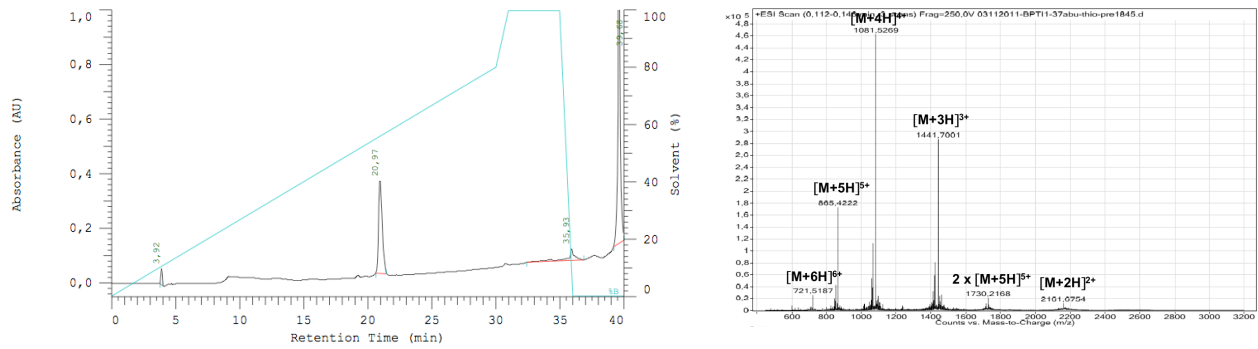
RP-HPLC spectrum and ESI-MS spectrum of total chemical synthesis of BPTI species



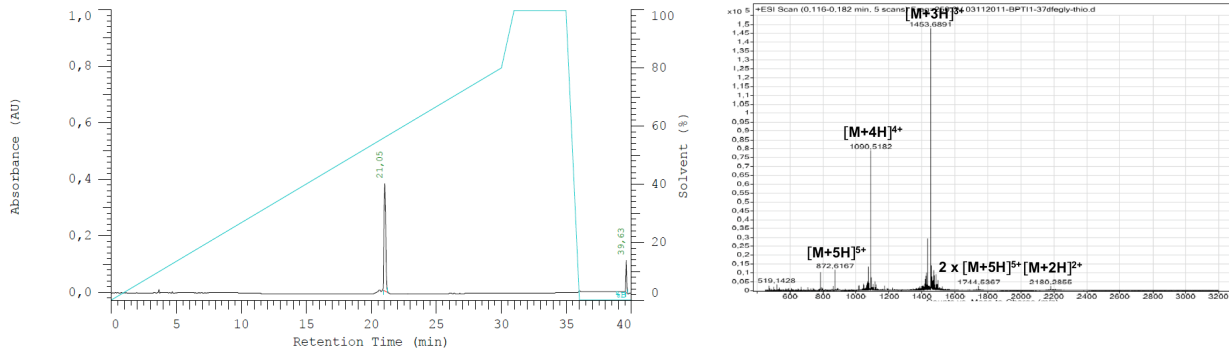
Left, analytic HPLC of purified BPTI 38-58 (retention time: 13.6 minute, 0-80% B in 30 minutes); **right**, ESI-MS analysis of purified BPTI 38-58



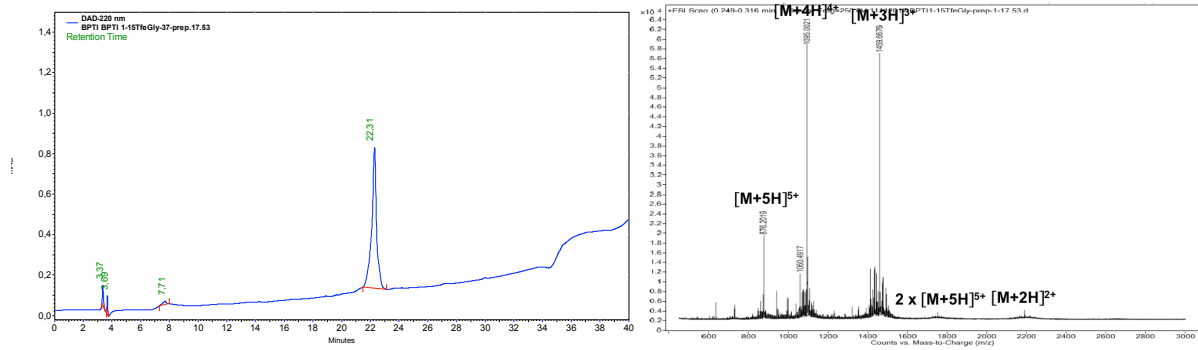
Left, Analytic HPLC of crude BPTI 16-37 (retention time: 19.6 minute, 0-80% B in 30 minutes); **right**, ESI-MS analysis of crude BPTI 16-37



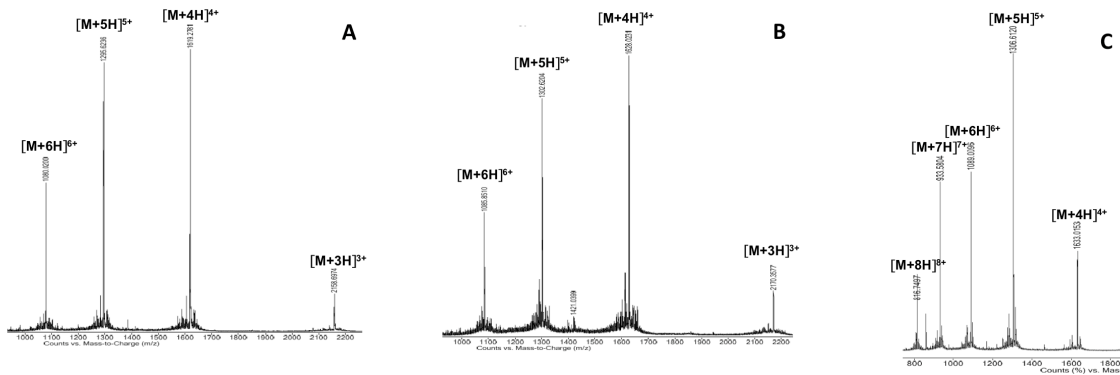
Left, Analytic HPLC of purified BPTI 1-*Abu15-37-SR* (retention time: 21.0 minute, 5-70% B in 30 minutes); **right**, ESI-MS analysis of purified BPTI 1-*Abu15-37-SR*



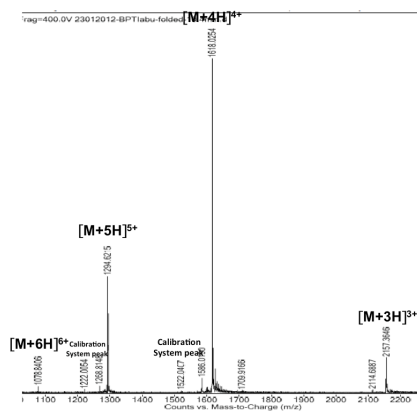
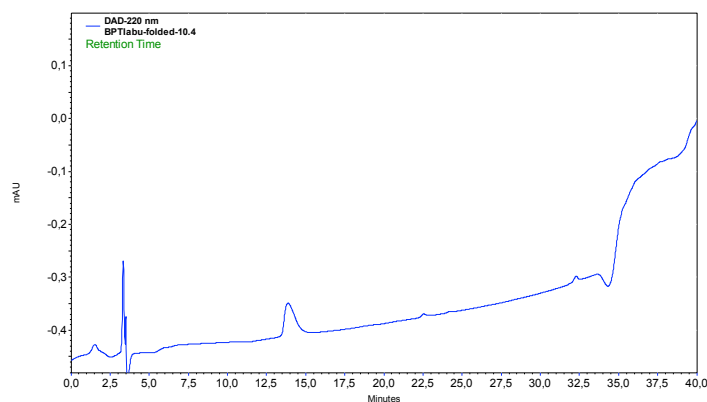
Left, Analytic HPLC of purified BPTI 1-*DfeGly*15-37-SR (retention time: 21.1 minute, 5-70% B in 30 minutes); **right**, ESI-MS analysis of purified BPTI 1-*DfeGly*15-37-SR



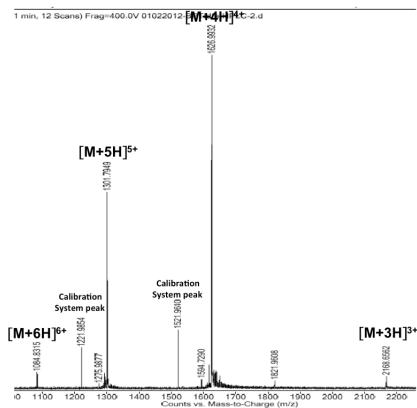
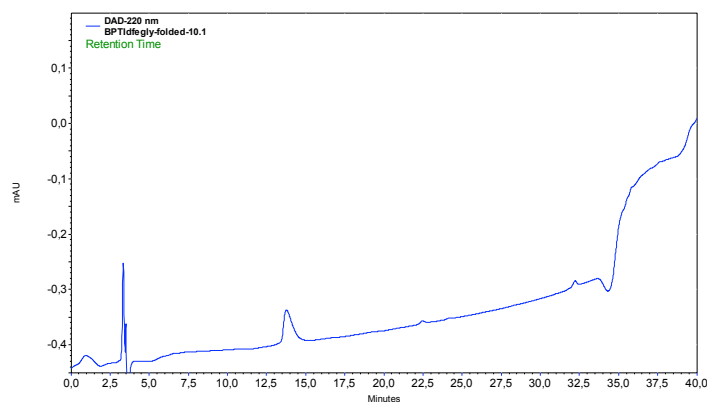
Left, Analytic HPLC of purified BPTI 1-*TfeGly*15-37-SR (retention time: 22.3 minute, 0-80% B in 30 minutes); **right**, ESI-MS analysis of purified BPTI 1-*TfeGly*15-37-SR



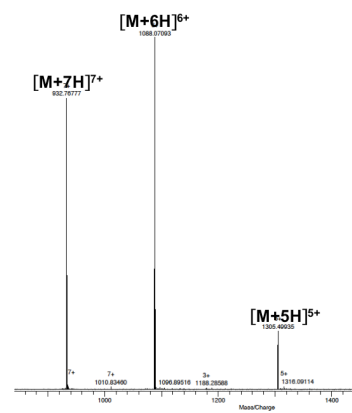
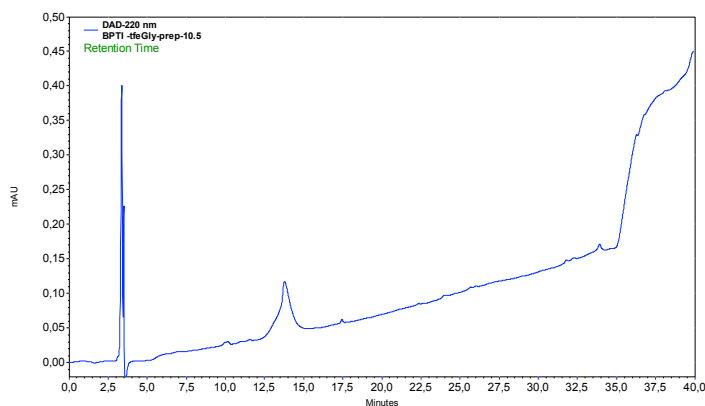
ESI-MS analysis of unfolded full-length BPTI mutants, **A**) unfolded BPTI_{Lys15Abu}, **B**) unfolded BPTI_{Lys15DfeGly}, **C**) unfolded BPTI_{Lys15TfeGly}



Left, analytic HPLC of purified folded BPTI_{Lys15Abu} (retention time: 13.9 minute, 5-70% B in 30 minutes); **right**, ESI-MS analysis of purified folded BPTI_{Lys15Abu}



Left, analytic HPLC of purified BPTI_{Lys15DfeGly} (retention time: 13.7 minute, 5-70% B in 30 minutes); **right**, ESI-MS analysis of purified folded BPTI_{Lys15DfeGly}



Left, analytic HPLC of purified BPTI_{Lys15TfeGly} (retention time: 14.1 minute, 5-70% B in 30 minutes); **right**, ESI-MS analysis of purified folded BPTI_{Lys15TfeGly}

Table Observed and calculated mass of synthetic BPTI fragments and BPTI species (m/z)

Peptide fragment		[M+2H] ²⁺	[M+3H] ³⁺	[M+4H] ⁴⁺	[M+5H] ⁵⁺	[M+6H] ⁶⁺	[M+7H] ⁷⁺	[M+8H] ⁸⁺
BPTI 38-58	ob. MS	1160.0386	773.6998	580.5161				
	cal. MS	1159.5226	773.3510	580.2652				
BPTI 16-37	ob. MS	1256.6597	838.1103					
	cal. MS	1256.1524	837.7708					
BPTI 1-Abu 15-37-SR	ob. MS	2161.6754	1441.7001	1081.5269	865.4222	721.5187		
	cal. MS	2161.1699	1441.1158	1081.0888	865.0726	721.0618		
BPTI 1-DfeGly 15-37-SR	ob. MS	2180.2855	1453.6891	1090.5182	872.8167			
	cal. MS	2179.1299	1453.0892	1090.0688	872.2566			
BPTI 1-TfeGly 15-37-SR	ob. MS		1459.6676	1095.0021	876.2019			
	cal. MS		1459.0892	1094.5688	875.8566			
BPTI_{Lys15Abu} - unfolded	ob. MS		2158.6974	1619.2781	1295.6236	1080.0200		
	cal. MS		2157.7122	1618.5361	1295.0304	1079.3600		
BPTI_{Lys15Abu} - folded	ob. MS		2157.3646	1618.0254	1294.6215	1078.8406		
	cal. MS		2156.0138	1617.2623	1294.0114	1078.0511		
BPTI_{Lys15DfeGly} - unfolded	ob. MS		2170.3577	1628.0231	1302.6204	1085.8510		
	cal. MS		2169.6855	1627.5161	1302.2144	1085.3467		
BPTI_{Lys15DfeGly} - folded	ob. MS		2168.6562	1626.9932	1301.7949	1084.8315		
	cal. MS		2167.9872	1626.2423	1301.1954	1084.4975		
BPTI_{Lys15TfeGly} - unfolded	ob. MS			1633.0153	1306.6120	1089.0096	933.5804	816.7497
	cal. MS			1632.0161	1305.8144	1088.3476	933.0125	816.5119
BPTI_{Lys15TfeGly} - folded	ob. MS				1305.4994	1088.0709	932.7678	
	cal. MS				1304.6514	1087.3775	932.1818	

ob. = observed, cal. = calculated

Table Crystal screening conditions

	Salt	Buffer	Precipitant	Trypsin +BPTI			
				wt	Abu	Dfe	Tfe
A1			2.2 M Ammonium sulfate	+	✓		✓
B1	0.2 M Ammonium sulfate		2.2 M Ammonium sulfate		+	+	+
C1	0.2 M Ammonium chloride		2.2 M Ammonium sulfate	✓	+	+	✓
D1	0.2 M Ammonium phosphate		2.2 M Ammonium sulfate	+	+	+	✓
E1	0.2 M Ammonium fluoride		2.2 M Ammonium sulfate				
F1	0.2 M Ammonium formate		2.2 M Ammonium sulfate	✓	✓	+	✓
G1	0.18 M tri-Ammonium citrate		2.2 M Ammonium sulfate			✓	
H1	0.2 M di-Ammonium phosphate		2.2 M Ammonium sulfate				
A2	0.2 M Ammonium iodide		2.2 M Ammonium sulfate	+	+	✓	✓
B2	0.2 M Ammonium nitrate		2.2 M Ammonium sulfate		✓	✓	
C2	0.2 M di-Ammonium tartrate		2.2 M Ammonium sulfate	+	✓	+	+
D2	0.2 M Cadmium chloride		2.2 M Ammonium sulfate	✓		✓	+
E2	0.2 M Cadmium sulfate		2.2 M Ammonium sulfate			✓	✓
F2	0.2 M Cesium chloride		2.2 M Ammonium sulfate	✓			✓
G2	0.2 M Cesium sulfate		2.2 M Ammonium sulfate	✓			✓
H2	0.2 M Ammonium bromide		2.2 M Ammonium sulfate		+		✓
A3	0.2 M Lithium acetate		2.2 M Ammonium sulfate	✓	+	+	
B3	0.2 M Lithium chloride		2.2 M Ammonium sulfate		✓	+	+
C3	0.2 M tri-Lithium citrate		2.2 M Ammonium sulfate				
D3	0.2 M Lithium nitrate		2.2 M Ammonium sulfate	+	+	+	+
E3	0.2 M Lithium sulfate		2.2 M Ammonium sulfate				
F3	0.2 M Potassium acetate		2.2 M Ammonium sulfate		✓	+	✓
G3	0.2 M Potassium bromide		2.2 M Ammonium sulfate				✓
H3	0.2 M Potassium chloride		2.2 M Ammonium sulfate				
A4	0.2 M tri-Potassium phosphate		2.2 M Ammonium sulfate	+	✓	+	
B4	0.2 M potassium phosphate		2.2 M Ammonium sulfate			+	✓
C4	0.2 M Potassium fluoride		2.2 M Ammonium sulfate	+	+	+	✓
D4	0.2 M Potassium formate		2.2 M Ammonium sulfate				
E4	0.2 M di-potassium phosphate		2.2 M Ammonium sulfate				
F4	0.2 M Potassium iodide		2.2 M Ammonium sulfate	+	+	+	+
G4	0.2 M Potassium nitrate		2.2 M Ammonium sulfate				
H4	0.2 M K/Na tartrate		2.2 M Ammonium sulfate			+	+
A5	0.2 M Potassium sulfate		2.2 M Ammonium sulfate		+		
B5	0.2 M Potassium thiocyanate		2.2 M Ammonium sulfate		+	+	✓
C5	0.2 M Sodium acetate		2.2 M Ammonium sulfate	+		✓	+
D5	0.2 M Sodium bromide		2.2 M Ammonium sulfate	+		+	+
E5	0.2 M Sodium chloride		2.2 M Ammonium sulfate				
F5	0.2 M tri-Sodium citrate		2.2 M Ammonium sulfate	+	+	+	+
G5	0.2 M Sodium phosphate		2.2 M Ammonium sulfate				✓
H5	0.2 M Sodium fluoride		2.2 M Ammonium sulfate		+	+	+
A6	0.2 M Sodium formate		2.2 M Ammonium sulfate	+	+	+	✓
B6	0.2 M di-Sodium phosphate		2.2 M Ammonium sulfate			+	+
C6	0.2 M Sodium iodide		2.2 M Ammonium sulfate	✓			
D6	0.2 M Sodium malonate		2.2 M Ammonium sulfate	+	+	✓	+
E6	0.2 M Sodium nitrate		2.2 M Ammonium sulfate	+			
F6	0.2 M Sodium sulfate		2.2 M Ammonium sulfate	+	✓	+	+
G6	0.2 M di-Sodium tartrate		2.2 M Ammonium sulfate	+			✓
H6	0.2 M Sodium thiocyanate		2.2 M Ammonium sulfate			+	+

A7		0.1 M Citric acid pH 4.0	0.8 M Ammonium sulfate	✓	+	+	✓
B7		0.1 M Citric acid pH 5.0	0.8 M Ammonium sulfate		✓	+	✓
C7		0.1 M MES pH 6.0	0.8 M Ammonium sulfate	+	✓	+	+
D7		0.1 M HEPES pH 7.0	0.8 M Ammonium sulfate	✓			+
E7		0.1 M Tris pH 7.0	0.8 M Ammonium sulfate	✓			
F7		0.1 M Bicine pH 9.0	0.8 M Ammonium sulfate	+		+	
G7		0.1 M Citric acid pH 4.0	1.6 M Ammonium sulfate				
H7		0.1 M Citric acid pH 5.0	1.6 M Ammonium sulfate		+	+	+
A8		0.1 M MES pH 6.0	1.6 M Ammonium sulfate	+	+	+	+
B8		0.1 M HEPES pH 7.0	1.6 M Ammonium sulfate		✓	+	+
C8		0.1 M Tris pH 7.0	1.6 M Ammonium sulfate	+	✓	+	+
D8		0.1 M Bicine pH 9.0	1.6 M Ammonium sulfate	+	✓	+	+
E8		0.1 M Citric acid pH 4.0	2.4 M Ammonium sulfate		✓		✓
F8		0.1 M Citric acid pH 5.0	2.4 M Ammonium sulfate	+		+	
G8		0.1 M MES pH 6.0	2.4 M Ammonium sulfate	✓			✓
H8		0.1 M HEPES pH 7.0	2.4 M Ammonium sulfate			+	+
A9		0.1 M Tris pH 7.0	2.4 M Ammonium sulfate	✓			
B9		0.1 M Bicine pH 9.0	2.4 M Ammonium sulfate			+	+
C9		0.1 M Citric acid pH 4.0	3.2 M Ammonium sulfate	+	+	+	+
D9		0.1 M Citric acid pH 5.0	3.2 M Ammonium sulfate	+	+	+	✓
E9		0.1 M MES pH 6.0	3.2 M Ammonium sulfate	✓			✓
F9		0.1 M HEPES pH 7.0	3.2 M Ammonium sulfate	+		+	+
G9		0.1 M Tris pH 7.0	3.2 M Ammonium sulfate	+	+	+	+
H9		0.1 M Bicine pH 9.0	3.2 M Ammonium sulfate		+	+	+
A10	0.1 M tri-Sodium citrate		0.5 M Ammonium sulfate	+	✓	+	+
B10			1.0 M Lithium Sulfate				
C10			1.0 M Ammonium sulfate			+	+
D10		0.1 M Sodium acetate pH 4.6	1.0 M Ammonium sulfate				
E10		0.1 M HEPES sodium salt pH 7.5	2.4 M Ammonium sulfate	+	✓	+	+
F10	0.05 M tri-Sodium citrate		2% (w7v) PEG 400				
G10		0.1 M Tris•HCl pH 8.5	1.0 M Ammonium sulfate	+		+	✓
H10			1.2 M Ammonium sulfate	+	+	+	+
A11	0.5 M Lithium chloride		3% (w/v) Isopropanol				
B11	1.0 M Lithium sulfate		1.6 M Ammonium sulfate	✓			
C11	0.2 M Sodium chloride	0.1 M HEPES sodium salt pH 7.5	15% (w/v) Glycerol				
D11		0.1 M HEPES sodium salt pH 7.5	1.6 M Ammonium sulfate	+	✓	+	✓
E11			1.6 M Ammonium sulfate	+	✓	✓	✓
F11		0.1 M HEPES sodium salt pH 7.5	1.6 M Ammonium sulfate	+	✓	✓	+
G11		0.1 M MES sodium salt pH 6.5	2% (w/v) PEG 1000				
H11			1.8 M Ammonium sulfate	+	+	+	+
A12	2.0 M Sodium chloride		2.0 M Ammonium sulfate	+			✓
B12		0.1 M Sodium acetate pH 4.6	2.0 M Ammonium sulfate	+	+	+	+
C12		0.1 M MES sodium salt pH 6.5	2.0 M Ammonium sulfate	+			
D12			5% PEG 400				
E12		0.1 M Tris•HCl pH 8.5	2.0 M Ammonium sulfate				+
F12			2.2 M Ammonium sulfate	+	✓	+	+
G12			2.2 M Ammonium sulfate		+	+	✓
H12	0.1 M tri-Sodium citrate		20% (w/v) Glycerol				
A13			2.4 M Ammonium sulfate	+	+	✓	✓
B13			3.0 M Ammonium sulfate	✓		✓	✓
C13			1% (w/v) MPD				
D13			3.0 M Ammonium sulfate	+			✓
E13			10% (w/v) Glycerol				
F13		0.1 M HEPES sodium salt pH 7.5	3.5 M Ammonium sulfate	+	+	+	+
G13		0.1 M MES sodium salt pH 6.5	3.5 M Ammonium sulfate	✓		+	✓
H13			1% (w/v) MPD				
I13			3.5 M Ammonium sulfate				

w.t. = wild-type BPTI; Abu, Dfe, and Tfe denotes BPTI variant, “+” indicates the observation of crystals, “✓” indicates the condition resulting in well-formed crystals (Figure 5.3-7).

For data security reasons the curriculum vitae has been omitted from the published version.