Fraunhofer-Institut für Angewandte Polymerforschung IAP
Universität Potsdam, Arbeitskreis Polymermaterialien und Polymertechnologien

Synthesis of Protein-Polymer Conjugates and Block Copolymers via Sortase-Mediated Ligation

Dissertation zur Erlangung des akademischen Grades
"doctor rerum naturalium" (Dr. rer. nat.)
in der Wissenschaftsdisziplin "Polymerchemie"

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Potsdam
von

M.Sc. Johannes Martin

Disputation: Potsdam, 21.06.2024

Huauptbetreuer: Prof. Dr. Alexander Böker
Betreuer: Dr. Ulrich Glebe, Prof. Dr. Heiko Möller
Gutachter: Prof. Dr. Alexander Böker, Prof. Dr. Heiko Möller, Prof. Dr. Katja Loos

Published online on the
Publication Server of the University of Potsdam:
https://doi.org/10.25932/publishup-64566
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-645669

Tell me and I will forget, show me and I may remember; involve me and I will understand.

Confucius

I certify that the content presented within this report is my own original work based on the research I performed using only the means and source materials as noted therein.

This thesis was not submitted to another examination board in this or other countries. There were no unsuccessful examination processes.

Potsdam, 30.11.2023 Johannes Martin

ACKNOWLEDGEMENTS

This thesis would not have been possible without the generous assistance and support of many people. I would like to express my gratitude to everyone who helped me scientifically and mentally throughout my PhD journey.

First and foremost, my supervisors Prof. Dr. Alexander Böker and Dr. Ulrich Glebe. Thank you for the opportunity to work on such an intriguing topic and for providing the “basics” that make research possible. I highly appreciate your ongoing mentoring, new ideas, and invaluable assistance with a variety of problems.

Prof. Dr. Heiko Möller, thank you for agreeing to be my second reviewer and for being an excellent collaboration partner who provided valuable input. Thank you to Prof. Dr. Katja Loos for being the third reviewer of this thesis.

I would like to express my gratitude to Prof. Dr. André Laschewsky, who was always eager to lend a helping hand, discuss problems and provide invaluable input with his vast knowledge of almost everything. Furthermore, I am greatful for your efforts to keep the polymer community together during difficult times, encouraging knowledge transfer, and teaching us how to give proper presentations.

Dr. Matthias Hartlieb made my PhD life much easier by always helping out with all kinds of problems, no matter how strange and difficult the experimental data was. Thank you for inventing my favorite polymerization method: It simply works.

This project would not have been possible without the assistance of skilled biochemists. Thank you very much to Dr. Marcus Michaelis, Prof. Dr. Heiko Möller, Dr. Peijian Zou, Prof. Dr. Michael Sattler, Dr. Robin Zou, Prof. Dr. Ulrich Schwaneberg, Dr. Saša Petrović, and Prof. Dr. Petra Wendler for looking after everything "protein".

No chemical research can exist without proper analytics and the people who operate the instruments. Many thanks to Anne Lehnen, Sascha Prentzel, Angela Krtitschka, Dr. Ines Starke, and Sylvia Fürstenberg for countless SEC, NMR, and MS measurements.

Dr. Sophia Rosencrantz and Dr. Ruben Rosencrantz deserve special thanks for asking the probing questions and directing me to the gaps in my research while always being willing to assist.

Daily work and frustrating research become bearable and often entertaining with the right colleagues in the lab and office. Thank you to Xiaolin Dai, Maria Mathieu-Gaedke, Magnus Schwieters, Yannic Müllers, and Liang Qiu for your help with all of the big and small problems as well as for keeping the spirit high.

Furthermore, colleagues and friends from Fraunhofer IAP were amazingly helpful, expecially Michelle Hechenbichler, Cristiane Henschel, Benjamin Rodriguez Hernandez, Pinar Akarsu, Alejandro Martinez, Sany Chea, Jo Sing Tang, Kristin Schade, Stefan Reinicke, and Hyung Seok Choi. Thank you all for your scientific and mental support, countless very enjoyable lunch breaks, and for patiently listening to my endless ramblings about riding bikes, saving money, and sports.

SCIENTIFIC PUBLICATIONS

Parts of the results were already published, submitted or are still in preparation:

Scientific journals

Application of Sortase-Mediated Ligation for the Synthesis of Block-Copolymers and Protein-Polymer Conjugates
J. Martin, M. Michaelis, M. Hartlieb, S. Petrovic, A.-C. Lehnen, P. Wendler, H. Möller, U. Glebe
in preparation

Conference Posters:

U. Glebe, M. Mathieu, M. Schwieters, J. Martin, H. Charan, X. Dai, A. Böker
Protein-polymer conjugates: synthesis strategies and applications as natural membrane mimics and therapeutic agents
Symposium on innovative polymers for the nanomedicine of the 21st century, Jena, 16.07.2019.

PhD and PostDoc Conference, ProMatLeben Berlin 03.-04.09.2019

J. Martin, X. Dai, M. Schwieters, M. Mathieu, U. Glebe, A. Böker
Synthetic Strategies for Polymer-Protein Conjugates: The Sortase Approach

Macromolecular Colloquium, Albert-Ludwigs-Universität Freiburg 26.02.–28.02.2020

J. Martin, U. Glebe, A. Böker
Creating (Multiple) Connections between Natural and Synthetic Polymers: Sortase-Mediated Ligation to form Protein-Polymer Conjugates

Apart from the scientific work presented in this thesis, I was involved in further research which is published as follows:

B. Oloya, J. Namukobe, M. Heydenreich, W. Ssengooba, J. Martin, H. M. Möller, B. Schmidt, R. Byamukama; Two new compounds and the anti-mycobacterial activity of the constituents from *Zanthoxylum leprieurii* root bark; *Phytochem. Lett.*, **2023**, *54*, 107-113, DOI: 10.1016/j.phytol.2023.02.002.

L. Qiu, H. Zhang, T. Bick, J. Martin, P. Wendler, A. Böker, U. Glebe, C. Xing; Construction of Highly Ordered Glyco-Inside Nano-Assemblies through RAFT Dispersion Polymerization of Galactose-decorated Monomer; *Angew. Chem. Int. Ed.*, **2021**, *60*, 11098-11103. DOI: 10.1002/anie.202015692. (Open Access)

ABSTRACT

During the last decades, therapeutical proteins have risen to great significance in the pharmaceutical industry. As non-human proteins that are introduced into the human body cause a distinct immune system reaction that triggers their rapid clearance, most newly approved protein pharmaceuticals are shielded by modification with synthetic polymers to significantly improve their blood circulation time. All such clinically approved protein-polymer conjugates contain polyethylene glycol (PEG) and its conjugation is denoted as PEGylation. However, many patients develop anti-PEG antibodies which cause a rapid clearance of PEGylated molecules upon repeated administration. Therefore, the search for alternative polymers that can replace PEG in therapeutic applications has become important. In addition, although the blood circulation time is significantly prolonged, the therapeutic activity of some conjugates is decreased compared to the unmodified protein. The reason is that these conjugates are formed by the traditional conjugation method that addresses the protein's lysine side chains. As proteins have many solvent exposed lysines, this results in a somewhat uncontrolled attachment of polymer chains, leading to a mixture of regioisomers, with some of them eventually affecting the therapeutic performance.

This thesis investigates a novel method for ligating macromolecules in a site-specific manner, using enzymatic catalysis. Sortase A is used as the enzyme: It is a well-studied transpeptidase which is able to catalyze the intermolecular ligation of two peptides. This process is commonly referred to as sortase-mediated ligation (SML). SML constitutes an equilibrium reaction, which limits product yield. Two previously reported methods to overcome this major limitation were tested with polymers without using an excessive amount of one reactant.

Specific C- or N-terminal peptide sequences (recognition sequence and nucleophile) as part of the protein are required for SML. The complementary peptide was located at the polymer chain end. *Grafting-to* was used to avoid damaging the protein during polymerization. To be able to investigate all possible combinations (protein-recognition sequence and nucleophile-protein as well as polymer-recognition sequence and nucleophile-polymer) all necessary building blocks were synthesized. Polymerization via reversible deactivation radical polymerization (RDRP) was used to achieve a narrow molecular weight distribution of the polymers, which is required for therapeutic use.

The synthesis of the polymeric building blocks was started by synthesizing the peptide via automated solid-phase peptide synthesis (SPPS) to avoid post-polymerization attachment and to enable easy adaptation of changes in the peptide sequence. To account for the different

functionalities (free N- or C-terminus) required for SML, different linker molecules between resin and peptide were used.

To facilitate purification, the chain transfer agent (CTA) for reversible addition-fragmentation chain-transfer (RAFT) polymerization was coupled to the resin-immobilized recognition sequence peptide. The acrylamide and acrylate-based monomers used in this thesis were chosen for their potential to replace PEG.

Following that, surface-initiated (SI) ATRP and RAFT polymerization were attempted, but failed. As a result, the newly developed method of xanthate-supported photo-iniferter (XPI) RAFT polymerization in solution was used successfully to obtain a library of various peptide-polymer conjugates with different chain lengths and narrow molar mass distributions.

After peptide side chain deprotection, these constructs were used first to ligate two polymers via SML, which was successful but revealed a limit in polymer chain length (max. 100 repeat units). When utilizing equimolar amounts of reactants, the use of Ni^{2+} ions in combination with a histidine after the recognition sequence to remove the cleaved peptide from the equilibrium maximized product formation with conversions of up to 70 %.

Finally, a model protein and a nanobody with promising properties for therapeutical use were biotechnologically modified to contain the peptide sequences required for SML. Using the model protein for C- or N-terminal SML with various polymers did not result in protein-polymer conjugates. The reason is most likely the lack of accessibility of the protein termini to the enzyme. Using the nanobody for C-terminal SML, on the other hand, was successful. However, a similar polymer chain length limit was observed as in polymer-polymer SML. Furthermore, in case of the synthesis of protein-polymer conjugates, it was more effective to shift the SML equilibrium by using an excess of polymer than by employing the Ni^{2+} ion strategy.

Overall, the experimental data from this work provides a good foundation for future research in this promising field; however, more research is required to fully understand the potential and limitations of using SML for protein-polymer synthesis. In future, the method explored in this dissertation could prove to be a very versatile pathway to obtain therapeutic protein-polymer conjugates that exhibit high activities and long blood circulation times.

ZUSAMMENFASSUNG

In den vergangenen Jahrzehnten haben therapeutische Proteine in der pharmazeutischen Industrie mehr und mehr an Bedeutung gewonnen. Werden Proteine nichtmenschlichen Ursprungs verwendet, kann es jedoch zu einer Immunreaktion kommen, sodass das Protein sehr schnell aus dem Körper ausgeschieden oder abgebaut wird. Dadurch sind solche Proteine als Medikament wenig effizient. Um die Zirkulationszeit im Blut signifikant zu verlängern, werden die Proteine mit synthetischen Polymeren modifiziert (Protein-Polymer-Konjugate). Die Proteine aller heute auf dem Markt erhältlichen Medikamente dieser Art tragen eine oder mehrere Polymerketten aus Poly(ethylenglycol) (PEG). Die Konjugation von PEG an ein Protein wird als PEGylierung bezeichnet. Ein Nachteil der PEGylierung ist, dass viele Patienten bei regelmäßiger Einnahme dieser Medikamente Antikörper gegen PEG entwickeln, die den effizienzsteigernden Effekt der PEGylierung wieder aufheben. Aus diesem Grund wurde die Forschung zur Entwicklung von alternativen Polymeren zu PEG angestoßen.

Ein weiterer Nachteil der PEGylierung ist die oftmals deutlich verringerte Aktivität der Konjugate im Vergleich zum nativen Protein. Der Grund dafür ist die Herstellungsmethode der Konjugate, bei der meist die primären Amine der Lysin-Seitenketten und der N-Terminus des Proteins genutzt werden. Da die meisten Proteine mehrere gut zugängliche Lysine aufweisen, werden oft unterschiedliche und teilweise mehrere Lysine mit PEG funktionalisiert, was zu einer Mischung an Regioisomeren führt. Je nach Position der PEG-Kette kann das aktive Zentrum abgeschirmt oder die 3D-Struktur des Proteins verändert werden, was zu einem teilweise drastischen Aktivitätsabfall führt.

In dieser Arbeit wurde eine neuartige Methode zur Ligation von Makromolekülen untersucht. Die Verwendung eines Enzyms als Katalysator zur Verbindung zweier Makromoleküle ist bisher wenig untersucht und ineffizient. Als Enzym wurde Sortase A ausgewählt, eine gut untersuchte Ligase aus der Familie der Transpeptidasen, welche die Ligation zweier Peptide katalysieren kann. Dabei wurde ein Ansatz basierend auf *grafting-from* gewählt, um das Polymer unabhängig vom Protein herstellen zu können.

Ein Nachteil dieser Sortase-vermittelten Ligation ist, dass es sich um eine Gleichgewichtsreaktion handelt, wodurch hohe Ausbeuten schwierig zu erreichen sind. Im Rahmen dieser Dissertation wurden zwei zuvor entwickelte Methoden zur Verschiebung des Gleichgewichts ohne Einsatz eines großen Überschusses von einem Edukt für Makromoleküle überprüft.

Zur Durchführung der Sortase-vermittelten Ligation werden zwei komplementäre Peptidsequenzen verwendet, die Erkennungssequenz und das Nukleophil. Die zu verbindenden Bausteine tragen an

jeweils einem Ende die erforderliche Peptidgruppe. Um eine systematische Untersuchung durchführen zu können, wurden alle nötigen Bausteine (Protein-Erkennungssequenz zur Reaktion mit Nukleophil-Polymer und Polymer-Erkennungssequenz mit Nukleophil-Protein) hergestellt. Als Polymerisationstechnik wurde die radikalische Polymerisation mit reversibler Deaktivierung (im Detail, *Atom Transfer Radical Polymerization, ATRP* und *Reversible Addition-Fragmentation Chain Transfer, RAFT polymerization*) gewählt, um eine enge Molmassenverteilung zu erreichen, welche für die Anwendung als Medikament nötig ist.

Die Herstellung der Bausteine begann mit der Synthese der Peptide via automatisierter Festphasen-Peptidsynthese, um eine einfache Änderung der Peptidsequenz zu gewährleisten und um eine Modifizierung der Polymerkette nach der Polymerisation zu umgehen. Um die benötigte unterschiedliche Funktionalität der zwei Peptidsequenzen (freier C-Terminus bei der Erkennungssequenz bzw. freier N-Terminus bei dem Nukleophil) zu erreichen, wurden verschiedene Linker zwischen Harz und Peptid verwendet. Danach wurde der Kettenüberträger (*chain transfer agent, CTA*) zur Kontrolle der Polymerisation mit dem auf dem Harz befindlichen Peptid gekoppelt. Die für die anschließende Polymerisation verwendeten Monomere basierten auf Acrylamiden und Acrylaten und wurden anhand ihrer Eignung als Alternativen zu PEG ausgewählt. Verschiedene Versuche, die Polymerketten mittels oberflächeninduzierter Polymerisation (*surface-initiated, SI-ATRP und SI-RAFT*) direkt auf dem Harz herzustellen, waren nicht erfolgreich. Stattdessen wurde eine kürzlich entwickelte Technik basierend auf der RAFT-Polymerisation (*xanthate-supported photo-iniferter RAFT, XPI-RAFT*) verwendet um eine Reihe an Peptid-Polymeren mit unterschiedlichen Molekulargewichten und engen Molekulargewichtsverteilungen herzustellen. Nach Entfernung der Schutzgruppen der Peptid-Seitenketten wurden die Peptid-Polymere zunächst genutzt, um mittels Sortase-vermittelter Ligation zwei Polymerketten zu einem Blockcopolymer zu verbinden. Unter Verwendung von Ni^{2+}-Ionen in Kombination mit einer Verlängerung der Erkennungssequenz um ein Histidin zur Unterdrückung der Rückreaktion konnte ein maximaler Umsatz von 70 % erreicht werden. Dabei zeigte sich ein oberes Limit von durchschnittlich 100 Wiederholungseinheiten; die Ligation von längeren Polymeren war nicht erfolgreich.

Danach wurden ein Modellprotein und ein Nanobody mit vielversprechenden medizinischen Eigenschaften mit den für die enzymkatalysierte Ligation benötigten Peptidsequenzen für die Kopplung mit den zuvor hergestellten Peptid-Polymeren verwendet. Dabei konnte bei Verwendung des Modellproteins keine Bildung von Protein-Polymer-Konjugaten beobachtet werden, unabhängig davon ob eine C- oder N-terminale Modifikation angestrebt wurde. Der Grund hierfür ist wahrscheinlich eine nicht ausreichende Zugänglichkeit der Kettenenden des Proteins.

Der Nanobody konnte dagegen C-terminal mit einem Polymer funktionalisiert werden. Dabei wurde eine ähnliche Limitierung in der Polymer-Kettenlänge beobachtet wie zuvor. Die auf Ni-Ionen basierte Strategie zur Gleichgewichtsverschiebung hatte hier keinen ausschlaggebenden Effekt, während die Verwendung von einem Überschuss an Polymer zur vollständigen Umsetzung des Edukt-Nanobody führte.

Die erhaltenen Daten aus diesem Projekt bilden eine gute Basis für weitere Forschung in dem vielversprechenden Feld der enzymkatalysierten Herstellung von Protein-Polymer-Konjugaten und Blockcopolymeren. Langfristig könnte diese Herangehensweise eine vielseitig einsetzbare Herstellungsmethode von ortsspezifischen therapeutischen Protein-Polymer Konjugaten darstellen, welche sowohl eine hohe Aktivität als auch eine lange Zirkulationszeit im Blut aufweisen.

CONTENTS

LIST OF ABBREVIATIONS AND SYMBOLS

2-CT	2-chlorotrityl linker for SPPS
AA	amino acid
Abz-	2-aminobenzoyl moiety
AIBN	2,2′-azobis(2-methylpropionitrile)
ATRP	Atom Transfer Radical Polymerization
BABTC	*tert*-butanoic acid butyl trithiocarbonate
bipy	2,2′-bipyridine
BMPA	2-bromo-2-methylpropionic acid
Boc	*tert*-butoxycarbonyl protection group
BSA	bovine serum albumin
CBM (G-CBM/CBM-LPETGG)	carbohydrate binding module 3b ($CBM3b^{N126W}$) (modified for N-terminal/C-terminal SML)
CPABTC	2-cyano pentanoic acid butyl trithiocarbonate
CRP	Controlled Radical Polymerization
CTA	Chain Transfer Agent
DCC	dicyclohexylcarbodiimide
DCM	dichloromethane
DIC	*N,N'*-diisopropylcarbodiimide
DIPEA	*N,N*-diisopropylethylamine
DMAP	4-dimethylaminopyridine
DMF	*N,N*-dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
Dnp	2,4-dinitrophenol
DP	number-average Degree of Polymerization

Đ	dispersity
EDA	ethylenediamine
EDC-HCl	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloric acid
ESI-MS	ElectroSpray Ionization Mass Spectrometry
FDA	Food and Drug Administration
Fmoc	9-fluorenylmethyloxycarbonyl
FRET	Fluorescence (or Förster) Resonance Energy Transfer
GG	diglycine, used as nucleophile in SML
HBTU	hexafluorophosphate menzotriazole tetramethyl uronium
(Glu)-HEMA	(α-D-glucosyl)-hydroxyethyl methacrylate
HOBt	hydroxybenzotriazole
(RP)-HPLC	(reverse-phase) High Performance Liquid Chromatography
HSQC-NMR	heteronuclear single quantum coherence NMR
IFNα	interferon alpha
iniferter	chemical compound that simultaneously acts as initiator, transfer agent, and terminator
ISC	intersystem crossing
k_{hyd}	reaction rate of hydrolysis in SML
k_{trans}	reaction rate of transpeptidation in SML
LED	light-emitting diode
LPxTG/LPETG/G	sortase A recognition sequence
M_n	number-average molecular weight
MALDI-Tof MS	matrix-assisted laser desorption/ionization time of flight mass spectrometry
MALS	multiangle light scattering
MAM	More Activated Monomer
MeCN	acetonitrile
Me_6TREN	tris(2-dimethylaminoethyl) amine

MeOH	methanol
MS	Mass Spectrometry
MW	molecular weight
NHS	*N*-hydroxysuccinimide
Ni-NTA	Ni^{2+} ions coupled to nitrilotriacetic acid
NMI	*N*-methylimidazole
NMM	*N*-methylmorpholine
NMP	Nitroxide-Mediated Polymerization
NMP	*N*-methyl-2-pyrrolidone
NMR	Nuclear Magnetic Resonance
P (as prefix)	poly
P3T	photoinduced triplet energy transfer
PC	photocatalyst
P(DMA)	poly(dimethylacrylamide)
PEG	poly(ethylene glycol)
PeT	photoinduced electron transfer
PET-RAFT	Photo-Energy/Electron Reversible Addition-Fragmentation Chain-Transfer
PI-RAFT	Photo-Iniferter Reversible Addition-Fragmentation Chain-Transfer
P(NAM)	Poly(n-acryloylmorpholine)
P(NIPAM)	poly(*N*-isopropylacrylamide)
PMDTA	N,N,N',N",N"-pentamethyldiethylenetriamine
P(OEGA)	poly(oligoethyleneglycolacrylate)
P(OEGMA)	poly(oligoethyleneglycolmethacrylate)
RAFT	Reversible Addition-Fragmentation Chain-Transfer
RAM	Rink Amide linker for SPPS
RDRP	Reversible-Deactivation Radical Polymerization

RT	room temperature
SDS-PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SEC	size-exclusion chromatography
SI (as prefix)	surface-initiated
SML	sortase-mediated ligation
SPE	3-[dimethyl-[2-(2-methylprop-2-enoyloxy)ethyl]azaniumyl]propane-1-sulfonate
SPPS	Solid-Phase Peptide Synthesis
SrtA-WT/4M/5M/7M	sortase A variants: wild-type/4/5/7 mutations
TFA	trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
THF	tetrahydrofuran
TIPS	triisopropylsilane
TNBSA	2,4,6 trinitrobenzene sulfonic acid
TTC	trithiocarbonate
Ty1 (Ty1-LPETGG)	single domain antibody fragment (nanobody) against SARS-CoV-2 (modified for C-terminal SML)
UV	ultraviolet
XAN	*O*-ethyl-*S*-(1-carboxy)methyl xanthate
XPI-RAFT	Xanthate-Supported Photo-Iniferter Reversible Addition-Fragmentation Chain-Transfer

1 INTRODUCTION

1.1 PROTEINS IN MEDICINE

Proteins and the sub-class of enzymes are highly important not only in modern life but are also essential tools to enable the correct function of all living organisms. Based on a "simple" chain of amino acids, these molecules can accomplish remarkable feats in nearly every aspect. Many modern-day processes, from simple nutrients to specifically designed drugs, are unthinkable without the use of engineered and mass-produced proteins and enzymes. This introduction focuses on the use of proteins as drugs in the treatment of human diseases.

Protein therapeutics have several distinct advantages over "classical" small-molecule drugs. First, because proteins are highly specific, they are much less likely to disrupt normal biological processes and cause less harmful side effects. Second, because many of the protein drugs administered to patients are naturally produced by the body, they often cause a much lower immune response and are easily degraded without harmful degradation products.[1] In addition, the time required for approval and introduction of proteins into the drug market is much shorter than for small molecules (data for the US market) making the development financially appealing.[2]

Proteins have been used as drugs for humans since 1922, when insulin was first used to successfully treat diabetes.[3] The first approach was largely unpopular, owing to limited availability and very high costs. Fast forward 80 years, after overcoming the challenges of very inefficient and costly insulin extraction and purification from animal sources, insulin was approved as the first therapeutic protein for human use in 1982. This milestone was achieved by mass-producing recombinant insulin using engineered bacteria (*Escherichia coli*) and the human insulin gene. Recombinant insulin is still the primary treatment for type 1 diabetes.

Nowadays, recombinant DNA technology is used to produce the majority of therapeutic proteins approved for human use. The use of bacteria as bioreactors for protein synthesis offers several advantages over the extraction of proteins from natural sources, including simple scalability, a higher specific activity, much lower costs and the ability to tailor the protein to a specific purpose.[1] The outcome from the past decades of research is a large number of different recombinant therapeutic proteins which have been approved for human use, including monoclonal antibodies, natural interferons, vaccines, hormones, and modified enzymes.

Unsurprisingly, there are more challenges to overcome when developing new therapeutic proteins. To begin with, proteins are highly complex and cannot be easily synthesized from basic raw materials. Thus, cell lines, bacteria species, cell culture conditions and purification processes must

all be optimized. Furthermore, most approved protein therapeutics require post-translational modifications that must be carried out without causing harm to the delicate protein structure. Next, *in vivo* half life is a major concern based on the negative impact on patients when multiple injections are required in a short period of time, as well as economic viability (more about this topic in the following chapter).[4]

The US Food and Drug Administration (FDA) has approved an increasing number of drugs of biological origin in recent years. In 2022, 40 % of the FDA-approved drugs were biologics,[5] highlighting that therapeutic proteins are becoming increasingly important as modern drugs.

1.2 Protein-Polymer Conjugates

With the rise of biomolecules in the pharmaceutical industry, new issues arose that needed to be addressed. Especially the limited stability, low solubility, unwanted immunogenic reactions and short in vivo half life of protein drugs were major concerns that limited further development. To address these issues, a synthethic polymer was covalently bound to the proteins, resulting in hybrid molecules: Davis and Abuchowski reported increased protein activity, proteolytic resistance as well as increased thermal and pH stability of bovine serum albumin (BSA) which was modified with polyethylene glycol (PEG) in their groundbreaking work published 1977.[6] Perhaps the most important improvement of PEGylated proteins compared to their unmodified counterparts is the so-called "stealth" effect: PEGylated proteins show vastly improved blood circulation times and overall lower immunogenicity caused by sterically repelling antigenic epitopes, preventing degradation by proteolytic enzymes and reducing renal filtration due to the increased hydrodynamic radius.[6,7] The modification of a protein with a synthetic polymer is not without consequences: research showed that protein-polymer conjugates typically show reduced catalytic efficiencies caused by a stiffer structure[8] and superhydrophilicity[9]. However, the advantages of modification outweigh the drawbacks, which led to a continuous interest of researchers in this field.

For example, PEGylated asparginase (an enzyme used to treat leukaemia) exhibits an impressive plasma half-life of 357 h compared to the unmodified enzyme (20 h). Despite higher costs for the PEGylated drug, cost savings of up to 78 % were possible.[10] Follow-up research and significant improvments in the specificity and efficiency of PEGylation led ultimately to more than 17 PEGylated drugs that are FDA-approved (data from 2018).[11]

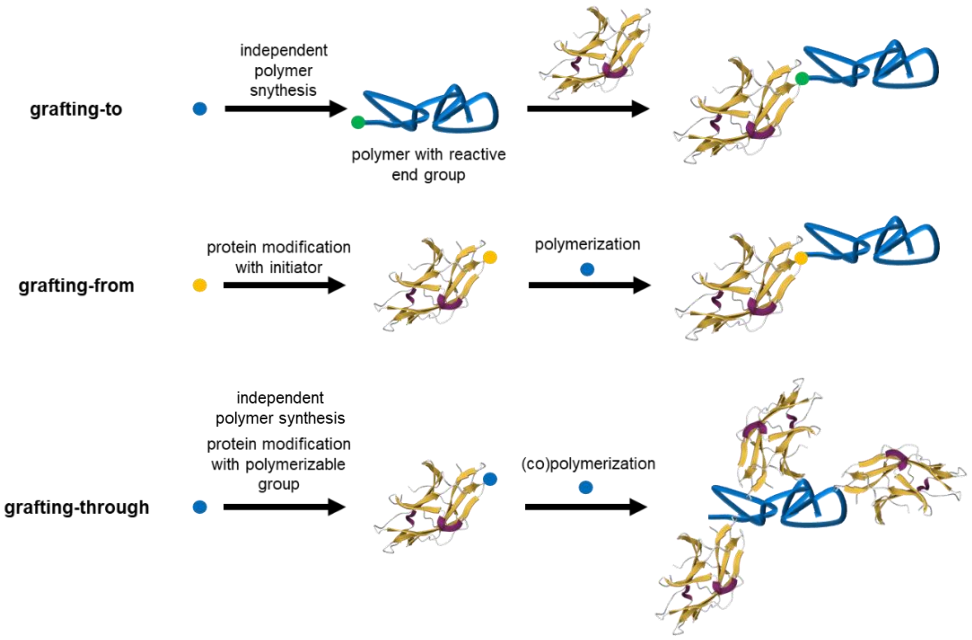
In the past decades, the development of new, highly controllable polymerization techniques (such as ATRP and RAFT polymerization, discussed in chapter 1.5) has expanded the use of polymers for protein conjugation far beyond PEG. As a result, a completely new set of protein-polymer conjugate properties for specific applications such as disease treatments, bioimaging, drug delivery,

bioactive surfaces, and tissue engineering is now possible.[12] In addition, recent research shows that anti-PEG antibodies are produced following a single initial dose of PEGylated drugs. This leads to accelerated blood clearance if continuous administration of (any) PEGylated drug is required. Due to the widespread use of PEG in recent years for numerous applications, not only for pharmaceuticals, many patients develop anti-PEG antibodies prior to initial treatment with a PEGylated drug.[13] In 2016, Yang *et al.* reported that in up to 70 % of untreated patients, anti-PEG antibodies were found.[14] This is not only a concern for drug efficacy, but patients with anti-PEG antibodies may develop PEG hypersensitivity, which could have an impact on their health upon treatment with a PEGylated drug. This phenomenon along with several other factors, promted extensive research into PEG alternatives.[15–17]

A reccurring theme in the search for PEG alternatives apart from the chemical structure of the polymer is the need to synthesize very well-defined polymer chains in order to avoid accumulation of non-biodegradable polymers within the body. Barz *et al.* go into great detail about this issue in their review.[15]

Researchers investigated a wide range of different polymers and polymer architectures as PEG substitutes: From “natural” carbohydrates and polypeptides to “pseudo-peptides” (poly(2-oxazolines) and purely synthetic polyacrylates and -acrylamides, there is a multitude of options.[13,15,18] All of them have advantages and disadvantages and must be carefully inspected for the intended application. From the standpoint of a polymer chemist, the synthethic pathway of controlled radical polymerization (which addresses the field of polyacrylates and -amides) provides the most versatility while maintaining simple synthesis procedures.

In the classical approach, PEGylated proteins are obtained by covalently attaching the previously synthesized PEG via amine-reactive polymer end groups to the lysine residues or the N-terminus of the protein. Other anchor points used to a lesser extend are cysteines (thiol-reactive PEG), and tyrosines (alcohol-reactive PEG).[12] This method is classified as *grafting-to*. With the development of new polymerization techniques compatible with aqueous media and mild reaction conditions, *in-situ* polymerization from the protein surface (*grafting-from*) and incorporating the protein in a growing polymer chain (*grafting-through*) are becoming increasingly popular.[19] Scheme 1 depicts a schematic representation of these methods. Synthesis via *grafting-through* is typically used only in rare cases for nieche applications. *Grafting-to* is still a viable option, especially with the recent development of click chemistry[20] and other (sometimes site-specific) methods.



Scheme 1: Different synthethic pathways to obtain protein-polymer conjugates.

A particularly impressive example of using *grafting-from* to synthesize protein-polymer conjugates was shown by Murata *et al.*.[21] Their approach is based on an insoluble solid support to which they attached the protein via it's N-terminus. Next, the protein was modified with initators for controlled radical polymerization (specifically atom-transfer radical polymerization, ATRP) and polymerization was carried out. Finally, the conjugate was cleaved from the solid support for further use. The use of a solid support greatly facilitates the normally required multiple cleaning and purification steps, resulting in a very quick and easy to perform technique to obtain conjugates. Drawbacks, however, are the limited yield of conjugates after cleavage of the support (51 % under optimal conditions) and the for some proteins very harsh cleavage conditions at pH 3.[21]

Even with some modern, FDA-approved protein-polymer conjugates, the problem of a distinct loss of activity after PEGylation remains. The primary reason is the nonspecific attachment of (several) PEG chains via *grafting-to* (for example at the N-terminus and several lysine side chains), which results in a mixture of positional isomers with varying numbers of PEG chains attached at different sites on each protein molecule.[12,22,23] Researchers have since attempted to solve this problem by using various different methods, such as coupling to specific amino acids within the protein or addressing the C- or N-terminus specifically.[24] More recently, new elegant approaches using enzymatic ligation (see chapter 1.3) of protein and polymer have been developed.[11,25] Enzymatic ligation, unlike click chemistry, does not require any chemical modification of the protein. The link is formed simply by joining two peptide sequences together, making transpeptidases ideal candidates to catalyze this reaction. The enzyme recognizes a specific peptide sequence

(recognition sequence) and catalyzes the intermolecular peptide bond formation with a different peptide.

Hu and coworkers, for example, reported the expression of genetically engineered interferon alpha (IFNα, widely used in viral disease and cancer treatment, PEGylated version FDA-approved as PEGASYS and PEGINTRON) containing the recognition sequence required for ligation by the enzyme sortase A (see chapter 1.3). Following the incorporation of an ATRP initiator via sortase-mediated ligation (SML) site-specifically at the C-terminus, a *grafting-from* polymerization of poly(oligo(ethyleneglycol) methyl ether methacrylate) (POEGMA) was carried out. The researchers reported significant improvements in yield (66 % vs. 1 % in classic *grafting-to* PEGylation) and *in vitro* bioactivity (7.2 fold increase) while maintaining the good pharmacokinetics of PEGylated IFNα.[25]

In a similar report, Popp *et al.* attempted to improve the stability and pharmacokinetics of IFNα with the sortaseA recognition sequence. They synthesized PEGylated IFNα via *grafting-to*, and obtained similar results as the report discussed before.[26]

Because *grafting-to* in contrast to *grafting-from* allows for independent polymer synthesis, special polymerization conditions for fragile proteins are not required. Therefore, a much broader range of polymers is applicable for protein conjugation. Suguri and Olsen demonstrated the use of poly(*N*-isopropylacrylamide) (PNIPAM) in a *grafting-to* approach to form model conjugates to study their self-assembly behavior, broadening the scope of polymers for SML beyond PEG and P(OEGMA).[27]

One of the major disadvantages of the SML-approach to protein-polymer conjugates is the requirement to use a large excess of either polymer or protein to achieve high yields (discussed in more detail in chapter 1.3). During research of this project, Reed and coworkers reported a very elegant method to address this problem. They successfully synthesized a PEGylated protein via *grafting-to* using equimolar amounts of reagents (discussed in chapter 1.3).[28]

Analytical techniques to characterize protein-polymer conjugates consist primarily of mass spectrometry, which provides accurate information about the sample molecular weight, in some cases including polymer chain length distribution.[28–31] When a simple yes/no answer on the synthesis outcome is required, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is commonly used.[26,27,32,33] To obtain more detailed information on the location of the polymer chain as well as the effects of ligation on the protein structure, $^{1}H/^{15}N$ 2D-NMR spectroscopy using isotope-labelled proteins can be applied.[34]

To summarize, the field of protein-polymer conjugates has advanced rapidly in the last decades, overcoming major limitations, making more high-performance protein drugs available. Recent

research continues to look for novel solutions to long-standing issues. With more research, enzymatic ligation could become another route to the mild and efficient synthesis of the next generation of protein-polymer conjugates.

1.3 SORTASE

Enzymes are proteins with catalytic properties. They catalyze a massive amount of biological reactions and are therefore essential for the overall metabolism. Most enzymes are highly efficient and highly specific catalysts that can speed up reactions by millions of times.[26] An enzyme, like a chemical catalyst, is not consumed during the reaction it catalyzes and has no effect on the reaction equilibrium. A specific enzyme can only catalyze a specific reaction due to its highly complex, three-dimensional shape. The catalytic step takes place at a specific site of the protein known as the active center. To enable efficient catalysis, all participating reactants (substrates) must fit precisely into this active center (commonly referred to as the "lock and key model").

Sortases are enzymes that belong to the superfamily of transpeptidases. This class of enzymes catalyzes the cleavage of a peptide bond within the peptide substrate, forming a covalent enzyme-substrate complex, which is then released from the enzyme via the formation of another amide bond with a different peptide.[35]

Sortases are found in Gram-positive bacteria, where they are responsible for attachment of specific surface proteins to the peptidoglycan of the cell wall. This is accomplished by recognizing a specific C-terminal peptide sequence (5 AA) of the substrate-protein, forming the previously mentioned enzyme-substrate complex, followed by the attachment of the protein to the cell wall. Researchers have discovered many different sortases which were categorized into different classes (A-F) depending on their different recognition peptide sequences.[36] Because only sortase A (SrtA) is used in this project, no additional information about the other sortases is provided; instead the reader is referred to an excellent review by Spirig *et al.*[36] Plus, SrtA was solely used *in-vitro* within this study, so readers interested in more details about the use of sortases *in-vivo* are directed to several articles on the subject.[37–40]

Sortase A

Among the different types of sortases, SrtA is by far the most studied so far. A major reason for the high interest of researchers in SrtA is that it is the only enzyme in the sortase family that is active *in vitro*. It has been used in a variety of different applications *in-vivo* and *in-vitro* as a tool to enable the synthesis of many highly complex and advanced molecules.[41] The use of sortase to catalyze an amide bond formation between two peptides is often referred to as sortase-mediated ligation (SML) or sortagging.

The mechanism of SML (shown in Figure 1) is commonly accepted to be as follows: SrtA recognizes the peptide sequence of LPX_1TGX_2 (X_1 stands for any amino acid except proline, with

glutamic acid E being the most commonly used; X_2 is a C-terminal amide or other AA, the peptide with free glycine carboxylic acid is not recognized as substrate).[42] A cysteine residue which is part of the active center attacks and cleaves the peptide between threonine (T) and glycine (G). From now on, LPX_1TGX_2, or more specifically LPETGG, will be referred to as the recognition sequence for SrtA. Although LPETG is commonly referred to as the recognition sequence in the scientific literature, I will use LPETGG because it is the shortest peptide used in this thesis and emphasizes that LPETG with a free carboxyl-C-terminus cannot be used for SML. If additional AA are part of the peptide, they are separated from the recognition sequence by a dash.

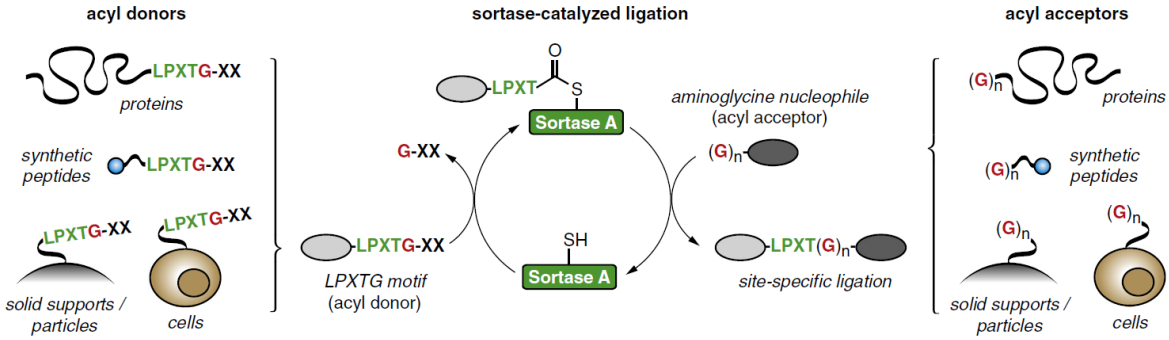


Figure 1: Mechanism of sortase-mediated ligation (SML) and commonly used substrates. Reprinted from reference[45].

Continuing the catalytic cycle, the peptide-sortase thiacyl intermediate contains the LPX_1T sequence, with the cleaved GX_2 remaining as a byproduct. The sortase-substrate complex is then attacked by a nucleophile, forming a new peptide bond and releasing sortase to restart the catalytic cycle.[43] The natural nucleophile for SrtA is a pentaglycine peptide which is located in the cell wall. In addition to that, researchers discovered that SML works well with a variety of nucleophiles, with the only requirement being an accessible primary amine.[44]

Because the LPXTG peptide sequence is not commonly found in proteins, transpeptidation via SML is a very powerful tool for obtaining various new materials for various applications. In addition, SrtA is robust and nowadays commercially available.

During the last decades, researchers quickly recognized the utility of SML and reported numerous used of SML to solve complex problems. The major applications of SML are found in the field of biotechnology for cell surface modifications and protein engineering.[41] Aside from that, ligation of small molecules (mostly dyes) and proteins, protein cyclization, synthesis of fusion-proteins, protein-polymer conjugates and block copolymers, and surface-immobilization of proteins have all been reported.[41] The ability to easily ligate small, synthetic molecules to proteins offers the most new applications for SML of any of these methods. Because the number of literature reports has

increased significantly in recent years, rather than covering recent advances in SML here, the reader is directed to several excellent SML reviews.[41,42,45,46]

In addition, more information about the recent advances in the use of SML to synthesize protein-polymer conjugates can be found in chapter 1.2.

A major drawback of SrtA is a comparably low catalytic efficiency, requiring the use of large amounts of enzyme. Furthermore, SrtA is dependent on Ca^{2+} ions. Both disadvantages have been addressed, and improved variants of SrtA have been developed. Chen *et al.* began by using a yeast display to evolve SrtA and increase its catalytic activity 140 times.[47] Based on this finding, Hirakawa *et al.* engineered a variant that is independent of Ca^{2+}.[48] Both variants (commonly referred to as SrtA-5M and SrtA-7M respectively) are now widely used in applications where the wild-type variant (SrtA-WT) is not producing satisfactory results. Aside from that, many more SrtA variants are known to the scientific community, each designed and engineered for a specific purpose. A detailed review was published recently by Freund and Schwarzer.[49]

A problem that was not successfully solved via bioengineering (so far) is a hydrolysis side-reaction. If low amounts of good nucleophiles (primary amines) are present, the sortase-substrate complex can be cleaved by a nucleophilic attack from water.[50] The result of this irreversible reaction is a

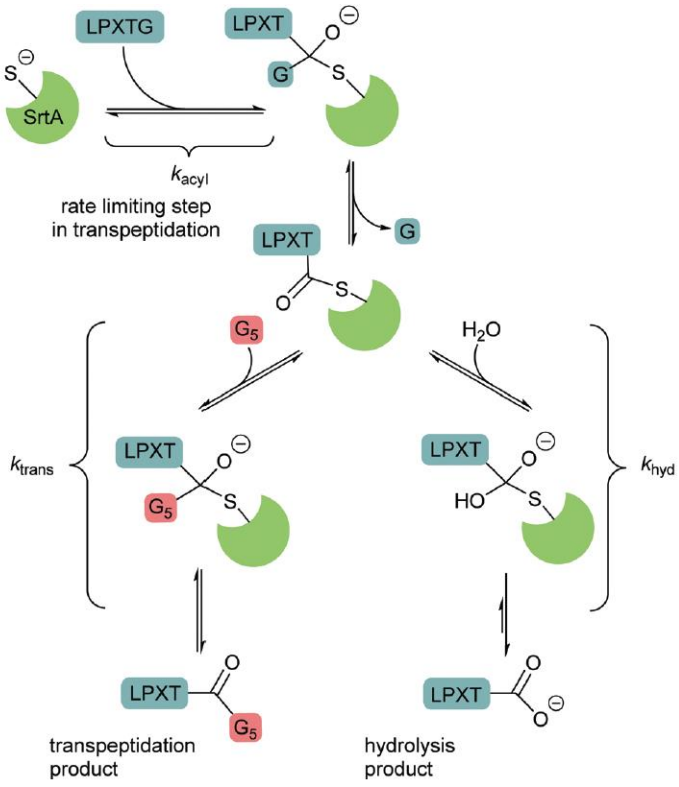


Figure 2: Schematic depiction of competing transpeptidation and hydrolysis in SML. Reprinted from reference[42].

hydrolyzed educt peptide sequence (Figure 2). Under normal conditions, k_{hyd} is significantly smaller than k_{trans}; however, when using sterically demanding reagents that require long reaction times or equimolar educt ratios (see below), hydrolysis may become a significant side reaction. Especially when using enhanced SrtA variants with high substrate turnover, the hydrolysis rate is increased as well and the reaction conditions may need to be optimized.[42]

SML Equlilibrium Reaction

In addition to the disadvantages of SrtA already mentioned, SML has another major drawback that has resulted in a variety of solutions to this problem. Because the peptide sequence in the substrate and the product after SML identical (LPETG, see Figure 1), SrtA also catalyzes the reaction back to the educts. The resulting equilibrium limits product formation to theoretically 50 %. Traditionally, this equilibrium has been shifted to maximize product yield by using an excess of one of the reactants (usually the smaller, less expensive reactant). This method, however, is not sustainable and is not always feasible, especially when using very expensive reagents or when reagent solubility is limited. Therefore, several strategies such as modifying the reagents, using additives or simply removing the initially cleaved glycine side product have been investigated. A recently published review covers this topic in great detail.[42] From the strategies presented in the scientific literature, I want to introduce those applied here in more detail.

A very promising strategy was introduced by Yamamura *et al.*, which is based on additional specific AA at both reagents to form a secondary structure after transpeptidation. Their idea is based on the "L-shape" of the LPXTG recognition sequence, which fits into the active center of SrtA. This shape is modified by inserting AA with aromatic side chains at the N-terminus of LPXTG and the C-terminus of the oligoglycine nucleophile.[51] Following SML, the aromatic side chains are in close proximity, and the peptide forms a β-hairpin structure due to attractive interactions. The shape of the β-hairpin differs significantly from the previous "L-shape" of the substrate and is thus not longer recognized by SrtA, resulting in irreversible SML (Figure 3).

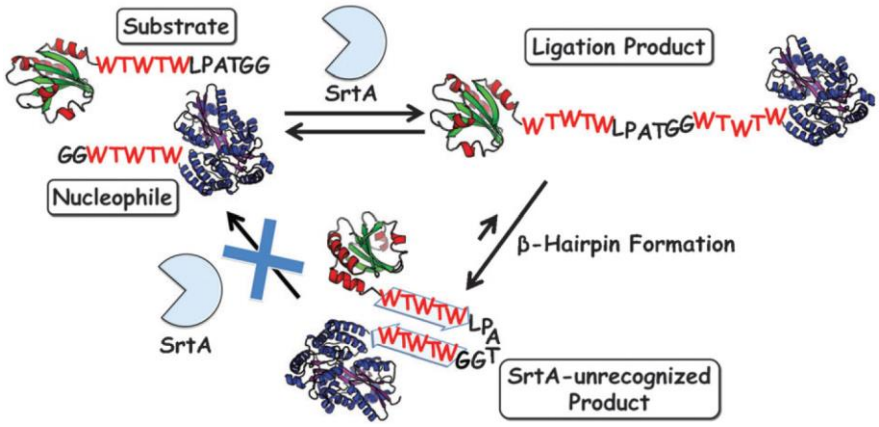


Figure 3: Irreversible SML via β-hairpin formation after transpeptidation. Reprinted from reference[51].

Yamamura and coworkers investigated the ability of various peptide sequences to form stable β-hairpins. The best sequence, based on a "Tryptophan-Zipper" (alternating tryptophan W and threonine T), performed very well when used in SML. They were able to achieve a fusion-protein yield of about 75 % in 24 h using equimolar ratios of substrate- and nucleophile-proteins as well as a catalytic amount of SrtA-WT (8 mol-%). Although both reagents must be modified for this strategy, the modification uses only natural AA which are relatively easy to incorporate into proteins using modern bioengineering techniques. Given the difficulty of SML of two sterically demanding proteins, the impressive results presented by Yamamura *et al.* make this approach highly promising for the use in my project.

The second strategy employed within this thesis is based on a minor change in the sequence of the substrate peptide. The approach, first described as "metal-assisted SML" by J. M. Antos' group, is based on complexation of the cleaved glycine residue following sortase-substrate complex formation. This is made possible by the addition of a histidine (H) AA at the C-terminus of the substrate. Upon cleavage from the substrate, this G_nH peptide is complexed by Ni^{2+} ions which are added to the reaction mixture (Figure 4) and can therefore not act as a nucleophile.[28,52] As a result, SML is irreversible because the G_nH nucleophile is essentially removed from the equilibrium. Because the peptide sequence after transpeptidation remains recognizable by SrtA, the substrate-sortase complex can be formed, but the lack of nucleophile prevents the reaction back to the educts. However, this means hydrolysis remains a competing side-reaction that needs to be closely monitored.

The authors tested various substrates and nucleophiles in equimolar ratios, including a synthetic polymer (PEG). In all of the cases presented, the addition of Ni^{2+} ions in conjunction with the

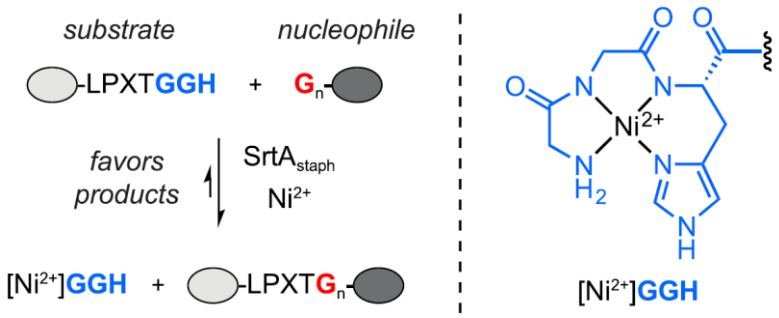


Figure 4: Metal-assisted SML made irreversible by complexation of side product by Ni^{2+} ions. Reprinted from reference[28].

modified substrate provided significantly higher product yields than the control reaction. Furthermore, analysis of the reaction mixtures reveiled very low amounts of hydrolyzed substrate.[28] A distinct advantage of this approach over others is the implementation of histidine as the only modification of the substrate. In most cases, the protein of interest exhibits a "purification handle" that allows for simple purification after expression. This handle is usually made up of a hexahistidine sequence at the C-terminus. Thus, such proteins can even be used for C-terminal metal-assisted SML without further modifications.

This approach to overcome the equilibrium problem of SML was chosen because of its simplicity, tests with a variety of reagents (including synthethic polymers) and excellent in-depth analysis of the reaction outcome.

1.4 SOLID-PHASE PEPTIDE SYNTHESIS (SPPS)

In order to use polymeric building blocks for ligation with SrtA to synthesize protein-polymer conjugates, incorporation of the recognition peptide sequence or the nucleophilic peptide is required. Therefore, the following chapter provides an overview of the most commonly used pathway to synthethic peptides today.

E. Fischer pioneered peptide synthesis research more than a century ago when he reported the synthesis of the dipeptide glycylglycine (GG) in 1907.[53] Several milestones later in 1966, B. Merrifield and A. Marglin reported the synthesis of insulin (51 AA) using their groundbreaking solid-phase approach, for which Merrifield was awarded the Nobel prize 1984.[54] Their method is based around a solid, crosslinked polymer support (resin) that does not dissolve during synthesis. The first AA is coupled via its carboxylic acid moiety followed by subsequent couplings of AA to produce the desired peptide sequence. The ease of purification, which is achieved in SPPS by simple washing of the insoluble resin, is a distinct advantage over previously used solution-based methods. The growing peptide is covalently bound to the resin and thus can be easily separated by filtration from reagents and excess AA in solution. Furthermore, the use of a solid support allows the use of a single reaction vessel for the entire process which includes numerous deprotection, coupling and washing steps. Nowadays, fully automated synthesis robots allow for the simple and quick synthesis of peptides of about 50 AA.[55] Protein synthesis is possible thanks to special protocols and the coupling of pre-synthesized peptide sequences.[56]

The success of SPPS is heavily reliant on the correct combination of orthogonal protecting groups of the N-terminus and the AA side chain functional groups.

There are primarily two protection group approaches for SPPS: To begin, *tert*-butoxycarbonyl (Boc) is used as an N-terminal protection group in combination with benzyl as side chain protection group.[57] Second, the base-labile 9-fluorenylmethyloxycarbonyl (Fmoc) is the N-terminal protection group and various orthogonal (acid-labile) protection groups (e.g. Boc and *tert*-butyl) for the AA side chains.[58] The first strategy has a major drawback in that it requires HF for the final deprotection and resin cleavage. In addition to being highly dangerous, HF deprotection led to significant amounts of side-reactions. This circumstance makes the Fmoc-strategy the preferred and mostly used technique nowadays. Figure 5 depicts an example peptide (tyrosine-glycine) ready for additional AA couplings.

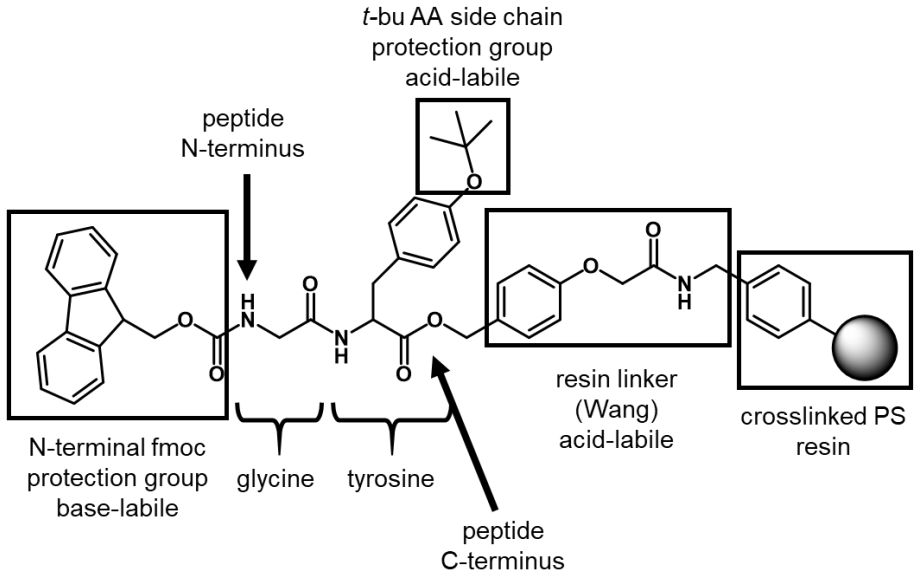


Figure 5: Fmoc-strategy in solid-phase peptide synthesis. Fmoc-strategy in combination with a Wang linker is shown as an example.

As a solid support, Merrifield used polystyrene crosslinked with 1 % divinylbenzene. To date, most of the resins are based either on crosslinked polystyrene, PEG or a combination of both. The resin used by Merrifield was functionalized for AA attachment by chlorination of the benzyl groups, yielding a carboxylic acid C-terminal peptide after cleavage.[57] There are now numerous resins that use different linker molecules to produce different C-terminal functionalities after resin cleavage. The most commonly used are Wang- and 2-chlorotrityl (2-CT)-resins to yield carboxylic acids, and Rink amide (RAM) resin to produce amides. A more detailed discussion about linker chemistry can be found in chapter 3.1.1.

Because of the highly stable ammonium carboxylate salt formed by the direct reaction of two AA, the formation of a peptide bond is not favored thermodynamically. Therefore, an efficient method to activate the carboxylic acid prior to reaction with the primary amine of the previously coupled AA on the resin is needed. This is typically achieved by *in-situ* activation via carbodiimide, anhydride or active ester. After much optimization during several decades of research, the active ester approach, employing phosphonium or aminium salts is now widely used. These reagents are capable of forming active esters very quickly, even with sterically hindered AA, and exhibit very few side reactions and little racemization. Plus, all reaction products are soluble in *N*,*N*-dimethylformamide (DMF, the most commonly used solvent for SPPS), allowing for quick and efficient washing steps. To achieve high yields in SPPS, a near quantitative coupling efficiency of AA must be reached, which requires careful selection of reactants and reaction conditions. Coupling efficiencies can be improved by a second coupling with the same AA before removal of

the N-terminal protection group. In addition, the amount of peptides missing one or more AA in the sequence (deletion peptides) can be significantly reduced by capping unreacted N-termini, thus preventing further growth of those peptides and enabling easier purification. For a more in-depth look at peptide coupling chemistry, the reader is referred to the excellent review by S. Han and Y. Kim.[59] The steps of N-terminal protection group removal without affecting side chain protection groups, followed by carboxylic acid activation of the next AA in the desired sequence, (double) coupling to the free N-terminus on-resin, and capping of unreacted N-termini are repeated to synthesize the final peptide. In between those steps, rigorous washing of the resin is performed.

Finally, the deprotection of AA side chains and cleavage of the peptide from the resin is accomplished either with HF (Boc-strategy) or trifluoroacetic acid (TFA, Fmoc-strategy). Different linkers allow a selective resin cleavage without side chain deprotection (for example, 2-CT, see chapter 3.1.1). When side chain protection groups are removed, they form stabilized carbocations that can react with electron-rich AA side chains (for example, the thiol moiety in cysteine), resulting in unwanted side products. This is avoided by incorporating silane-based scavengers which are added to the cleavage cocktail.[60] After resin cleavage, the peptides are usually precipitated, lyophilized and, if nessecary, purified using preparative high-performance liquid chromatography (HPLC).

1.5 Reversible Deactivation Radical Polymerization

Reversible deactivation radical polymerization (RDRP) is the most versatile technique to date to synthesize a variety of polymers that are considered canditates to replace PEG in protein-polymer conjugates in the future.[13,15–17]

The concept of RDRP combines the convenience of free radical polymerization with the excellent control over the molecular weight of a living polymerization. This polymerization technique produces polymers with controlled molar masses, comparably low dispersities, a high number of defined end groups, and precise molecular architectures in topology and composition.

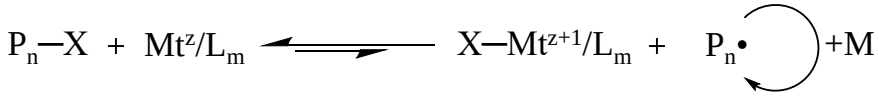
To achieve these attibutes, a fast and dynamic equilibrium between active propagating radicals and a dormant species is required. To maintain good control and ensure equal growth of all polymer chains, the reaction mixture should contain a low number of active radicals and a high number of dormant species at any given time. The reactivation of dormant radicals can be achieved via three major pathways: Either via a redox process (as in ATRP), via homolytic bond cleavage (nitroxide-mediated polymerization NMP) or via degenerative transfer (RAFT polymerization).[61]

Since the discovery of RDRP many different specialized subcategories of this technique have been invented. In this introduction, I will only discuss the ones used within this project (ATRP and RAFT polymerization). The reader is directed to several recent reviews about the development of RDRP for additional information.[62–64]

The following subchapters explain mechanisms, benefits and drawbacks of the polymerization techniques used in this project.

1.5.1 Atom Transfer Radical Polymerization (ATRP)

For reversible activation/deactivation of growing polymer chain ends, ATRP employs a metal ion catalyst. An alkyl halide P_n–X reacts with a transition metal (most commonly Cu^I) complex Mt^z/L_m with low oxidation state z to form an active radical P_n• and the oxidized metal complex X–Mt^{z+1}/L_m. The active radical P_n• reacts with the monomer M to form a polymer before it is deactivated (P_n–X) again (Scheme 2).[65] Because the halide remains always at the chain end of a dormant chain, the polymerization can be reinitiated to polymerize a different monomer and produce block copolymers. Almost all of the reagents involved can be changed to achieve the best results.



Scheme 2: ATRP mechanism.

Initiators for ATRP are typically alkylhalides (mostly Br and Cl), particularly secondary or tertiary alkylester halides. Alkyl-iodides are sometimes used, but the stability of the Cu^{II}-I bond is very low, resulting in more active radicals and less control. When using the same halide, the activation constant of an α-substituted ester depends on the grade of substitution of the halide-carbon: a tertiary carbon stabilizes the radical best and thus has the highest activation constant, whereas a primary carbon has the lowest activation constant. Plus, substituents on the halide-carbon may provide additional radical stability through mesomeric effects (for example, a higher activation constant when using nitril or phenyl instead of methyl moieties).[66]

Other redox-active transition metal complexes besides Cu^{I} are based on Co, Ni, Ru, Fe, etc.[67] The activity of the metal complex is largely controlled by the structure of the ligands. Most often, chelating ligands are used, since they contribute significantly more to the stabilization the Cu^{II} complex than monodentate ligands. The use of nitrogen-containing tetradentate ligands such as tris(2-dimethylaminoethyl)-amine (Me_6TREN) strongly activates the complex, whereas bidentate ligands such as 2,2′-bipyridine (bipy) result in much less active catalysts.[65,68]

Polar solvents increase the equilibrium constants and thus speed up the reaction because the more polar Cu^{II} complex is better stabilized by the solvent then the less polar Cu^{I}. Very polar solvents, such as water, significantly increase reaction speed, resulting in a lack of control for certain polymerizations.[65]

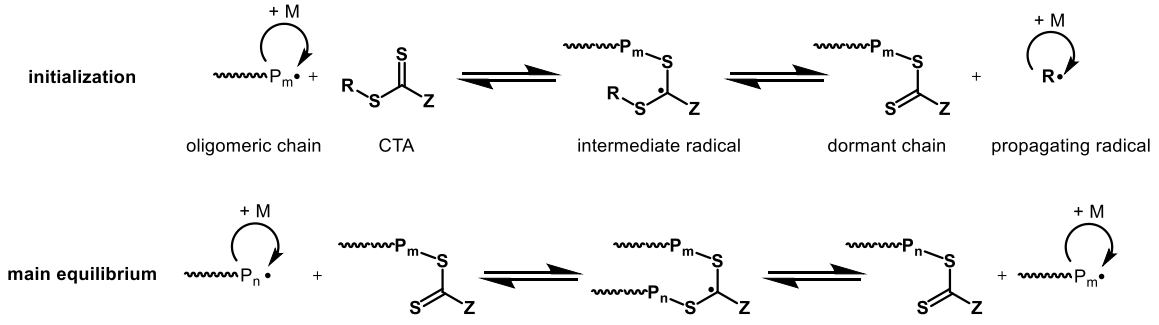
Higher temperatures favor active radicals due to the lower activation energy required for termination, which speeds up the polymerization. However, increasing the temperature may cause unwanted side reactions.[65]

An advantage of ATRP is the tunability and flexibility of the method. Teodorescu *et al*., for example, discovered that polymerization of acrylamides, particularly *N*,*N*-dimethylacrylamide (DMA), was very challenging, resulting in very limited monomer conversion and poor control.[69] They attributed the poor results to catalyst deactivation, most likely caused by competitive complexation of the transition metal. When different reaction conditions were tested, it was discovered that polymerization of DMA was very fast and controlled when polymerizing in very polar solvents (e. g. pure water). Another advantage of ATRP over thermally induced RAFT polymerization are the very mild reaction conditions: temperatures can be kept very low, even below room temperature is possible when using polar solvents.[65] Some monomers, on the other hand, are extremely difficult to polymerize via ATRP and necessitate extensive optimization of the initiator, transition metal, ligands, and solvent (see above). Furthermore, traces of transition metals can remain in the polymer after polymerization which are difficult to remove and might be problematic for certain (pharmaceutical) applications.[65]

1.5.2 THERMALLY INDUCED REVERSIBLE ADDITION-FRAGMENTATION CHAIN-TRANSFER (RAFT) POLYMERIZATION

Moad, Rizzardo, and Thang independently developed the RAFT concept in 1998.[70] The mechanism of RAFT polymerization is similar to conventional free radical polymerization: initiation, propagation and termination are all the same. Using RAFT, an additional activation/deactivation equilibrium is introduced via degenerative chain transfer. To enable this equilibrium formation, the addition of a chain transfer agent (CTA) is required. CTAs are typically thiocarbonylthio compounds: The reactive site is the C-S double bond, which is attacked by the active radical, causing the attacking radical to become dormant while releasing the other polymer chain attached to it, allowing propagation.[71]

In traditional RAFT polymerization, initiation is induced thermally by the decomposition of a radical starter compound (typically azo-compounds such as azobisisobutyronitrile AIBN). The formed radicals attack monomer molecules, forming short oligomeric chains. These chains attach to a CTA via its C-S double bond, resulting in an intermediate radical. This radical can fragment, releasing a new radical (R-group) that starts reacting with monomer, forming a new chain. Following the consumption of all initial CTA, radicals attack CTA moieties that contain dormant polymer chains, resulting in chain transfer (Scheme 3). This equilibrium ensures that radicals are distributed evenly, enabling that all chains are growing similarly, resulting in low dispersities of the final polymer.



Scheme 3: RAFT mechanism, reversible deactivation via degenerative chain transfer.

To provide good control over the polymerization, the CTA needs to be finetuned depending on the type of monomer used and the desired end group of the polymer chain. For more activated monomers (producing more stabilized radicals, resulting in slower polymerization) such as meth(acrylates) and (meth)acrylamides, highly reactive CTAs such as dithioesters and trithiocarbonates (TTC) are required. Less reactive CTAs based on xanthates or dithiocarbamates

are used for less activated monomers (producing highly reactive radicals, resulting in very fast polymerization), such as vinyl ethers or vinyl esters.[72]

Finetuning of CTA is achieved by variation of R- and Z-groups. Because the R-group is responsible for both CTA fragmentation and reinitiation, a balance must be found between R-radical stability for efficient fragmentation and reactivity for efficient reinititation. The Z-group determines the reactivitiy of the C-S double bond to radicals as well as the fragmentation of the intermediate radical species. Overall, to obtain polymers with low dispersities, chain transfer must be much faster (>10 fold) than propagation.[73] More information about the correct selection of R- and Z-groups can be found in the scientific literature.[72]

Aside from the nature of the CTA, the ratio of CTA to monomer needs to be adjusted. In contrast to free radical polymerization, the degree of polymerization (or chain length) is determined by the CTA/monomer ratio rather than the number of external radicals. Next, initiator to CTA ratio needs to be considered. This ratio determines the livingness of the reaction: Every initiator radical that reacted with a monomer results in a polymer chain with an (unwanted) radical starter molecule at the α chain end. Thus, the amount of external initiator needs to be as small as possible to ensure high livingness while still providing enough radicals to achieve high monomer conversion.[74] Polymer chains with either a radical starter or a destroyed CTA (caused by side reactions) at the end are referred to as "dead chains" because they cannot participate in the RAFT equilibrium.

Overall, the RAFT polymerization process is very robust, tolerating a wide range of functional groups and providing good adjustability to control challenging monomers. Plus, the demand for special equipment and experimental techniques is very low compared to ionic polymerizations, because RAFT polymerization is only critically affected by radical inhibitors (e. g. oxygen) and functional groups capable of destroying the CTA.[71]

Surface-Initiated (SI)-RAFT Polymerization

Many different strategies can be used to covalently attach a polymer to a surface. Some of the more elegant ones use in-situ polymerization to or on the surface. Apart from attaching a (thermal) radical starter, this can be realized using a surface-bound CTA. Attaching the CTA via its R-group is defined as a *grafting-from* approach, whereas attachment via Z-group is considered *grafting-to*. The two methods differ in that the polymer grows in solution and reattaches to the surface when the Z-group approach is used. When using the R-group approach, however, the polymer is always bound to the surface.[75] Both approaches have advantages and disavantages that must be considered for the desired application. Chapter 4 contains a more detailed discussion of the Z- vs. R-group approach in this project.

After extensive research, the R-group approach is widely used today. Ranjan and Brittain compared R- and Z-group approaches on silica particles and found that the R-group approach achieved much higher grafting densities, whereas the molecular weight (MW) of the chains synthesized via Z-group approach was higher due to less steric hindrance.[76–78]

The R-group approach requires the addition of free CTA that ensures a good control via rapid transfer between free and surface-bound chains. The polymer chains formed in solution can be analyzed and provide a good indication of MW and dispersity of the surface-bound chains.[79]

Today, many different surfaces have been successfully employed for SI-RAFT, including silica particles, metal-oxide, graphene, and cellulose. These hybrid materials have been used in various applications.[75]

1.5.3 Light-Induced RAFT-Polymerization

In recent years, new approaches to initiating the RAFT polymerization have been explored. Because of the high spatial and temporal control, there has been increased interest in using light to create radicals to drive polymerization. The research resulted in three different pathways: Starting the polymerization via energy/electron transfer (PET-RAFT)[80] using a photocatalyst which is activated by light, or the use of the CTA itself als photoiniferter (a molecule that can act as initiator, transfer agent and termination agent, PI-RAFT)[74]. Ironically, PI-RAFT was first used in 1982 by Otsu and Yoshida, more than 10 years before the development of RDRP.[81] Unfortunatly, its potential was not recognized at the time, most likely due to high polymer dispersities caused by the incorrect combination of monomer and CTA.[82]

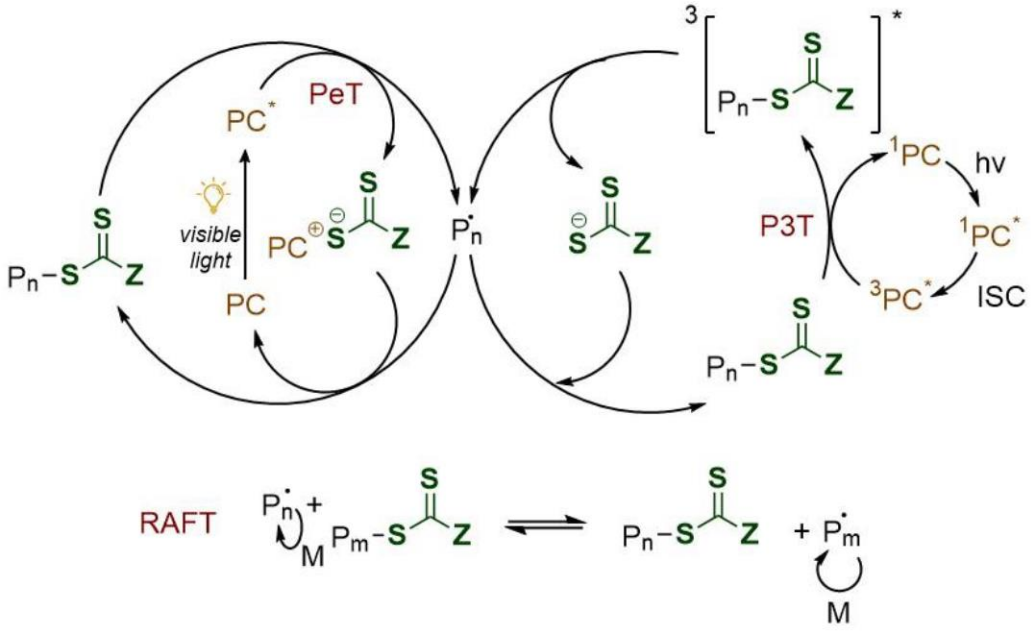
The third light-initiated strategy employs photo-initiators that generate radicals upon irradiation instead of the classical thermal initiation. While some benefits of light-induced RAFT polymerization remain, the amount of dead chains caused by the photo-initiator as well as decreased livingness due to termination via recombination with the photo-initiator radicals are the same as in thermally induced RAFT polymerization. Because of that, this method was not applied here.

The other two strategies, PET-RAFT and PI-RAFT, have been used for this project and are thus discussed in more detail in the following subchapters.

Photo-Energy/Electron Transfer (PET) RAFT Polymerization

PET-RAFT was developed with the goal in mind to create a robust and efficient living/controlled method to polymerize a wide variety of monomers, even in the presence of molecular oxygen. To achieve this goal, Boyer *et al.* used a photocatalyst (PC) which is added to the reaction mixture.[83]

It harvests visible light and enters an excited state. The consumed energy is transferred to the CTA either via a photoredox process or via photoinduced energy transfer. The electronically excited CTA undergoes fragmentation, generating a propagating radical and a CTA radical. Both radicals are capable of recombination, resulting in the formation of the main RAFT equilibrium (Scheme 4).



Scheme 4: Proposed mechanisms for PeT (left) and P3T (right) RAFT polymerization. PeT: photoinduced electron transfer; P3T: photoinduced triplet energy transfer; ISC: intersystem crossing. Reprinted from reference[80].

Compared to thermally induced RAFT polymerization, PET-RAFT has several advantages. First and foremost the avoidance of an external radical source, which results in dead chains (polymer chains with no CTA located at the chain end, which are unable to participate in the RAFT equilibrium), improves the livingness of the process. Second, the PET-RAFT process is not dependent on a temperature that is high enough to ensure decomposition of a thermal radical initiator, thus potentially harmful high temperatures can be avoided without using highly explosive initators. Moreover, the polymerization can be applied in presence of molecular oxygen. Finally, using light as activator for the radical source provides the possibility to switch the reaction on and off as well as fine-tune reaction speeds by controlling the energy input, and it is orthogonal to other RDRP techniques.[80]

Several different PCs are in use nowadays, the majority of which are based on transition metal complexes, which can be toxic and must be removed for most medical applications. In an attempt to move the method towards a greener non-toxic approach, organic dyes have been tested and were

found to be partially very potent PCs. When using organic dyes, however, additional sacrificial electron donors (for example, tertiary amines) are required to run the polymerization in the presence of oxygen.[84]

In an impressive demonstration of the method's versatility, Li *et al.* "printed" a complex 3-dimensional gray-scale picture from a surface using several subsequent PET-RAFT polymerizations in the presence of oxygen.[85]

To date, PET-RAFT polymerization has been employed successfully in a wide range of applications, including nanocomposite materials, bioconjugates, and materials with antimicrobial and antibiofouling properties.[80]

Xanthate-Supported Photo-Iniferter (XPI) RAFT Polymerization

The XPI-RAFT process is the most recent advancement in the field of PI-RAFT. In general, PI-RAFT relies on homolytic bond dissociation of a CTA to generate both persistent CTA radicals (thiocarbonylthio-radicals) and transient radicals to start the polymerization directly by irradiation. As a result, the CTA serves as both a chromophore and a radical source to initiate polymerization and the main RAFT equilibrium. This process is successful mostly because of the relatively high stability of the thiocarbonylthio-radicals: they are stable (persistent) enough to be unable to start the polymerization, but are reactive enough to reversibly deactivate active chain ends and thus contribute to the RAFT equilibrium.[74]

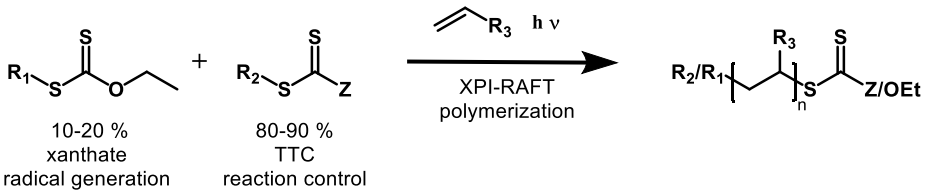
In addition to the benefits of photo-induced RAFT polymerization mentioned in the previous chapter, PI-RAFT offers another advantage: There is no need to remove any potentially toxcic compound from the obtained polymer because no external photocatalyst is used.

Several different light sources have been used to generate radicals for PI-RAFT over the years. γ-radiation proved to be useful in some cases, although the unselective nature led to radical formation by reaction with monomer and solvent. In addition, long reaction times were needed, and monomer conversion was rather low.[74] Following that, UV-light was employed for the activation of various types of CTAs, including dithioesters, dithiocarbamates, trithiocarbonates (TTC), and xanthates. It was found that in order to achieve a well controlled polymerization, the combination of wavelength, CTA and monomer had to be carefully selected. Non-optimal combinations led to uncontrolled reactions, low end group fidelity due to CTA photolysis, low conversion, and long reaction times. In general, it was found that TTC-based CTAs are more resistant to photodegradation than others, but exhibit a low efficiency of activation, resulting in long reaction times to achieve high conversions. CTAs based on xanthates and dithiocarbonates, on the other

hand, have a much higher activation efficiency but typically have a low chain-transfer constant, resulting in poor control over the polymerization of more activated monomers.[74]

Light in the visible spectrum was used successfully as well. Despite having less energy than UV-light, the same problems with irreversible CTA photodegradation occurred. This required careful selection of CTA and wavelength based on the desired monomer system. Plus, using visible light resulted in significantly longer reaction times when compared to UV-light if the input-power of the light source was not drastically increased.[86]

Hartlieb *et al.* recently published a simple solution to this complex problem.[87] They achieved impressive results in terms of ease of use, monomer versatility, high conversions, and short reaction times while maintaining excellent reaction control by combining xanthate-based CTAs (fast but uncontrolled) and TTC-CTAs (controlled but slow). The xanthate is added in minor amounts during this process, primarily as a photo-initiator due to its superior activation efficiency under UV-light. After the initial (rapid) start of the polymerization, the TTC-CTA reacts with the active chains, controlling the polymerization (Scheme 5). Here, the excess of TTC becomes important, because the xanthate is still part of the RAFT equilibrium but provides little control. Hartlieb and coworkers found that the optimal amount of xanthate is 10-20 % of total CTA loading, resulting in low dispersity polymers for various monomers after several hours of irradiation using a standard UV-lamp with a power of 2 W.[87]



Scheme 5: Schematic representation of the XPI-RAFT process using two different CTAs to generate radicals and to control the polymerization.

Of course, optimization of CTA-types and ratios, as well as light intensity, is required to achieve the best results, but XPI-RAFT offers very promising results with minimal optimization required. Using XPI-RAFT results in very high end group fidelities, allowing the synthesis of up to dodecablock copolymers.[87]

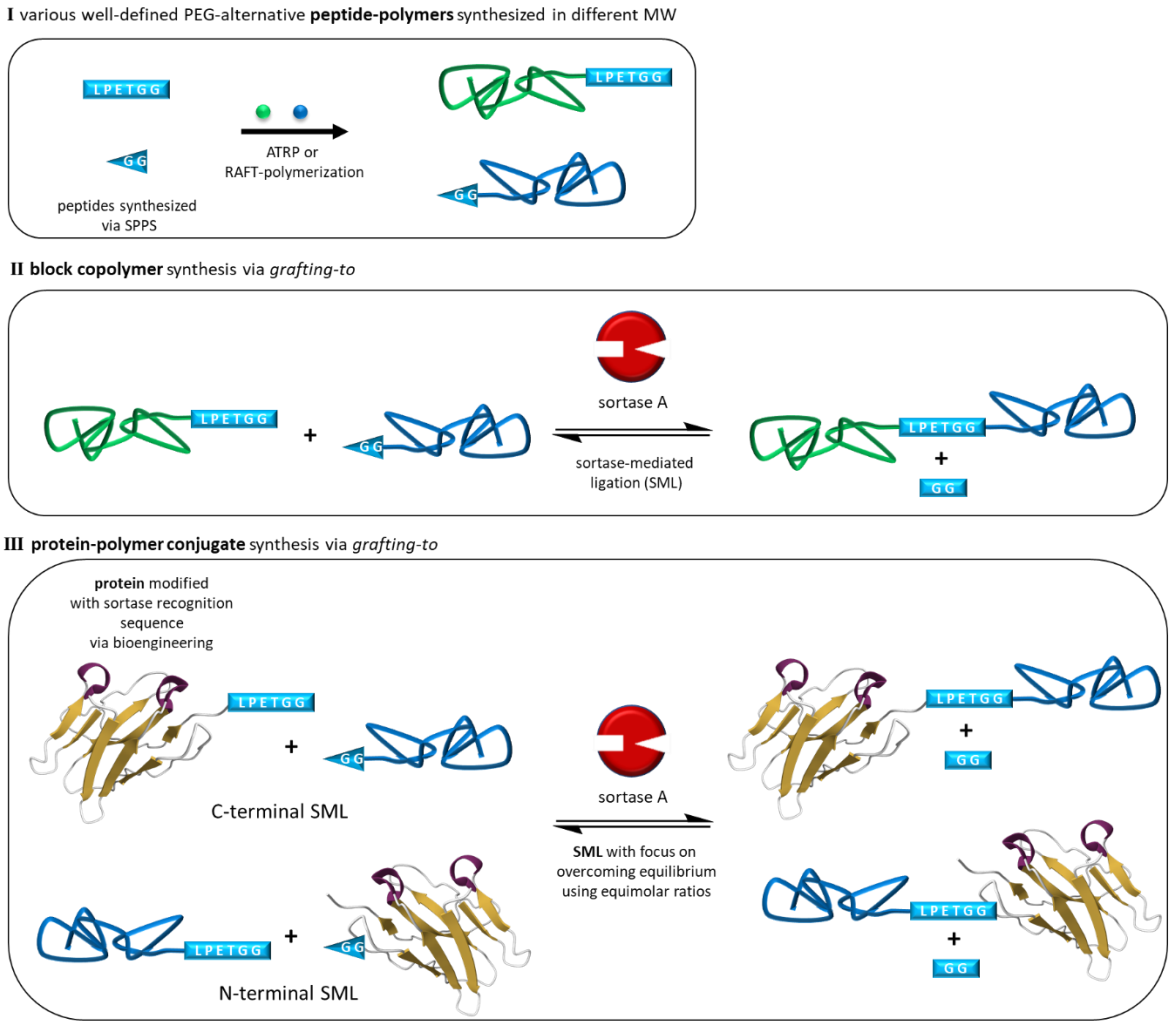
This new method solves previous problems by decoupling radical generation from reaction control and eliminating the need for external radical sources. It is also very user-friendly, can tolerate moderate levels of oxygen, and requires no special equipment.[87]

2 OBJECTIVES AND MOTIVATION

Protein-polymer conjugates, as discussed in earlier chapters, are critical for driving innovation in the field of biopharmaceutics by overcoming major limitations of medicines of purely biological origin. However, the protein-polymer conjugates currently offered to patients still have flaws that and can be significantly improved. With this project, I sought to address several of these issues. To begin, the exclusive use of PEG as synthethic polymer poses a number of issues.[13] As a result, instead of using PEG, I chose to explore promising alternative synthethic polymers (chapter 4) and test different chain lengths. I employed reversible deactivation radical polymerization (RDRP) to achieve well-defined polymers with low polymer dispersity while being able to cover a large variety of monomers. Second, the commonly used synthesis method of protein-polymer conjugates is mostly realized in an unspecific fashion, sometimes causing drastic loss of protein activity. Research has shown that site-specific conjugation has considerably less impact on protein performance[16] and structure[34]. Therefore, I decided to use the recently developed site-specific synthesis method of enzymatically catalyzed ligation. Sortase A (SrtA) was chosen as the enzyme, because it is well characterized, readily available and has been utilized before in similar studies. I chose *grafting-to* in favor of *grafting-from* to be able to synthesize the polymers independently, thus not being restricted to mild polymerization conditions compatible with the fragile protein. Finally, most earlier investigations utilizing sortase-mediated ligation (SML) to synthesize protein-polymer conjugates employed a large excess (10-30 fold) of one reagent and/or large amounts of sortase (up to 5 eq),[33] to maximise product yields,[27] resulting in a substantial amount of waste and poor sustainability. I picked two different promising strategies to solve the SML-equilibrium issue and compared them directly using equimolar concentrations of reagents to determine which one produces the best results.

The overall aim of this project is to gain insight into the ligation of macromolecules by using an enzymatic synthesis approach, thus opening a synthethic pathway to synthetic block copolymers and protein-polymer conjugates.

A graphical representation of these goals is depicted in Scheme 6.



Scheme 6: Schematic representation of the goals of this project.

To the best of my knowledge, there is very little literature on the synthesis of protein-polymer conjugates via SML and *grafting-to*.[26,27,33,88,89] Those employ a limited range of polymers, using mostly PEG. So far, Suguri and Olsen were the only researchers who employed RDRP to synthesize well-defined polymers.[27] Furthermore, only one paper addresses the need to use large amounts of one reactant to achieve acceptable yields in the context of protein-polymer conjugate synthesis (while still using PEG).[28] This report was published during ongoing research of my project.

Besides testing PEG-alternatives for conjugation, SML may prove to be a viable pathway to classical PEGylated proteins: Given that to date PEG is the only polymer which is clinically approved for protein-polymer conjugates and the synthesis of PEGylated proteins is only possible via *grafting-from*, resource-efficient and site-specific ligation is highly desirable. SML may prove to be a viable method to achieve these objectives; however, more research on avoiding excessive amounts of reagents besides the report mentioned above of needs to be conducted.

I used a divergent synthesis approach[90] to obtain the necessary polymers with a peptide end group, beginning with the synthesis of the peptide. Using this procedure allowed me to utilize a small amount of peptide rather than a large excess that would be required for post-polymerization modification of the end group. The Börner group used this method frequently and with great success.[91–94] Furthermore, it permits elegant surface-initiated synthesis of the polymer on the solid support used for peptide synthesis (see chapter 4.2.1). To establish which polymerization technique is more suited, the two most versatile RDRP techniques, atrom-transfer-radical polymerization (ATRP) and reversible addition-fragmentation chain-transfer (RAFT) polymerization, were studied in parallel.

The following synthesis plan (Figure 6) was devised to meet the aforementioned objectives. The initial step, synthesis of the peptide precursor either as recognition sequence, or as nucleophile for SML was carried out employing solid-phase peptide synthesis (SPPS). While the recognition sequence peptide was still linked to the resin, a chain transfer agent (CTA) for RAFT polymerization or an initiator for ATRP was coupled to it. This macro-CTA/initiator-peptide construct was subsequently used to obtain various polymers with different chain lengths, resulting in synthethic peptide-polymers with either recognition sequence or nucleophilic end groups. Next, the synthesis of block-copolymers via SML was investigated, potentially opening up an additional pathway to block copolymers besides classical click reaction chemistry or one-pot synthesis.Finally, SML was used to couple a polymer and a protein bearing complementary peptidic end groups.

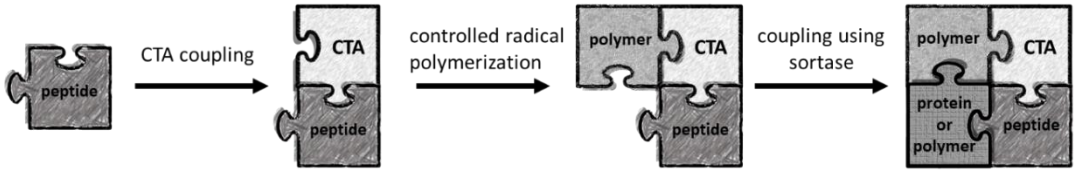


Figure 6: Schematic display of synthesis steps required to synthesize block copolymers or protein-polymer conjugates via SML.

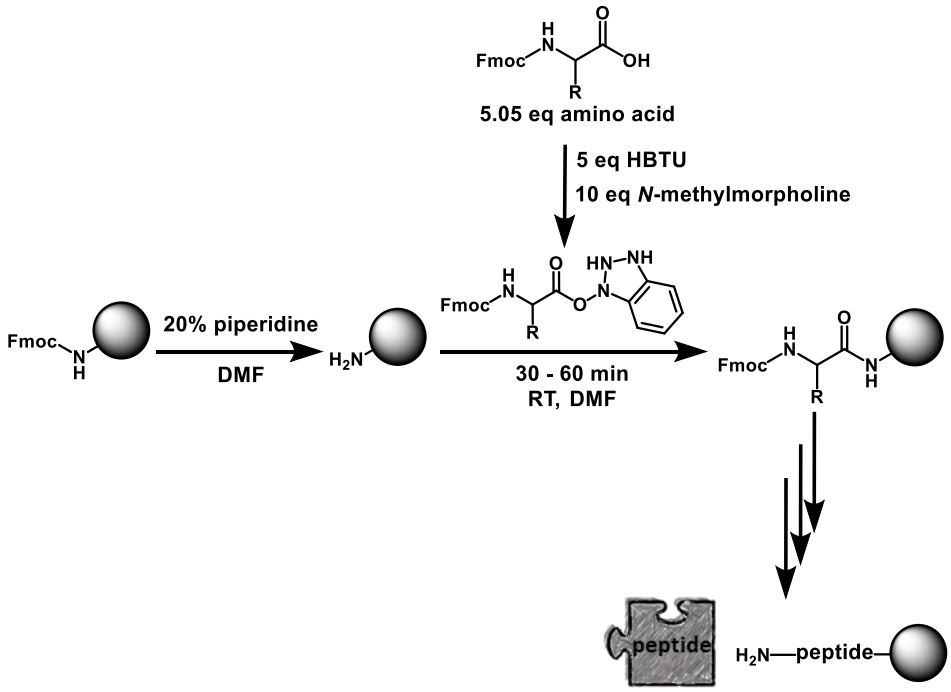
Two different proteins, a model protein and a functional nanobody were biotechnologically modified to contain the required peptide sequences for SML and utilized to synthesize protein-polymer conjugates. In addition, studies have been conducted to determine which combination of peptidic end groups provides the best results: either nucleophile-polymer + recognition sequence-protein or recognition sequence-polymer + nucleophile-protein.

3 PEPTIDE SYNTHESIS AND CTA/INITIATOR COUPLING

The next subchapters discuss the synthesis of the various building blocks necessary to synthesize polymers with specified end groups suitable for SML. Polymerization is discussed in chapter 4. Two fundamentally different peptides, recognition sequence and nucleophile, are needed with subsequent attachment of CTA (for RAFT polymerization) or initiator (for ATRP). The synthesis routes for the building blocks are different, because the recognition sequence peptide requires a free C-terminus (chapter 3.1) whereas the nucleophile requires a free N-terminus (chapter 3.2).

3.1 RECOGNITION SEQUENCE PEPTIDE-CTA

Scheme 7 depicts the synthesis plan for obtaining peptide constructs including the SML recognition sequence. SPPS, a standardized approach to synthesize peptides with high yields and purities was utilized (introduced in chapter 1.4). Experimental details can be found in chapter 7.3; only a few minor alterations to the conventional conditions, such as extended reation times and double couplings, have been implemented to maximise yield and purity. All peptide syntheses were performed using an automated peptide synthesizer.



Scheme 7: SPPS carried out using a crosslinked polystyrene resin (grey spheres), employing the "Fmoc-strategy".

The strategy employed here is known as “Fmoc-strategy”. To prevent uncontrolled reactions, the amino acid that needs to be coupled to the resin has a Fmoc protection group at the primary amine (see Scheme 7). The carboxylic acid moiety of the first (Fmoc-protected) amino acid was activated by hexafluorophosphate benzotriazole tetramethyl uronium (HBTU) in the presence of *N*-methylmorpholine (NMM), which formed an active ester that reacted with the primary amine groups on the resin to form the peptide bond. This reagent combination is widely used in (automated) peptide synthesis and is well-known for its mild properties, among other benefits.[95]

To attach the next amino acid, the Fmoc protection group was cleaved off using basic conditions (piperidine), releasing the first amino acid’s primary amine. The active ester was provided in excess (usually 5 eq), ensuring that the coupling reaction is completed, which is especially important for longer peptides. To obtain the final peptide with the desired amino acid sequence, the coupling and deprotection steps were repeated with all planned amino acids. Several washings of the resin were performed in between coupling steps to ensure that all reagents from the previous reaction are rinsed off. In addition, unreacted, free primary amines that remained after removal of the reaction solution were capped with acetic anhydride. They were rendered unreactive for further couplings by the formed acetamide end group.

Each amino acid coupling requires near quantitative conversion to achieve high yields of the final peptide. In SPPS, conversion is usually monitored by colorimetric detection of amines. 2,4,6-trinitrobenzene sulfonic acid (TNBSA), which forms a bright orange dye when reacting with amines, is a very fast and sensitive reagent for checking the reaction mixture for free amines.[96]

Complete conversion of the amino acid coupling was achieved by prolonging reaction times or repeating the coupling step for couplings that were found to be non-quantitative using standard reaction conditions (such as coupling of L to P and G to G). A qualitative TNBSA assay was used to confirm complete conversion.

In the case of this work, an additional functionality that allows the growth of a polymer chain from the peptide is required. As a result, after synthesis, the peptide was not cleaved from the resin and was used without further modification for CTA or ATRP initiator coupling (see chapters 3.1.3 and 3.1.4).

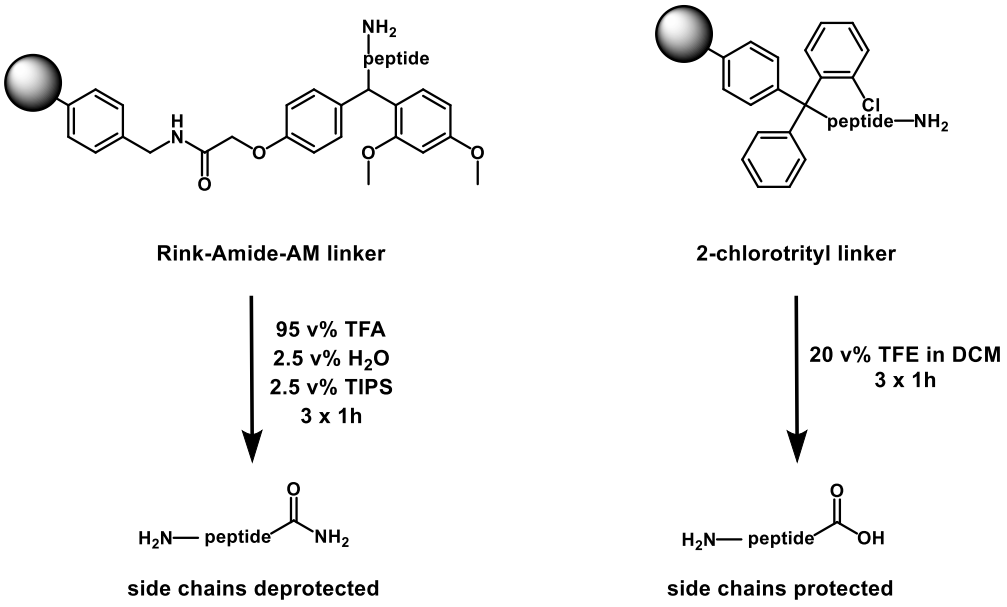
3.1.1 Resin Linker Choice

A linker molecule is used between the resin (usually crosslinked polystyrene) and the growing peptide to allow resin cleavage. This linker is cleaved as a designated weak point to ensure that the peptidic C-terminus functions properly after separation from the solid support. Scheme 8 contains all of the linkers used in this work.

The most common SPPS resins contain a Rink Amide (aminomethyl) RAM linker molecule between the resin surface and the peptide. This linker can be cleaved under highly acidic conditions, typically with a mixture of mostly trifluoroacetic acid (TFA, 95 v%) and small amounts of water (2.5 v%) and triisopropylsilane (TIPS 2.5 v%). While the ATRP initiator remained intact during the standard cleavage procedure (see below), both RAFT agents used in this study degenerated quickly and could not be isolated (see chapter 3.1.4)

To address this issue, CTA peptide constructs were prepared using a resin linker that can be cleaved under very mild conditions. The linker 2-chlorotrityl (2-CT) was cleaved using a mixture of 20 v% trifluoroethanol (TFE) in dichloromethane (DCM). Under these mild conditions, resin cleavage yields a peptide with the amino acid side chain protection groups still attached. In this case, this meant that after polymerization, a final deprotection step was required to remove the side chain protection groups and activate the peptide end group for SML.

A 2-CT resin was used to synthesize all CTA-containing peptides. This resin was commercially available and preloaded with the first glycine.



Scheme 8: Resin linker molecules and cleavage procedure.

3.1.2 PEPTIDE SEQUENCE

Sortase's essential peptidic recognition sequence is LPxTG, as previously stated. To maximize SML yields, E has been mostly used as x because it is the most abundant in natural substrates for sortase A.[97,98] Another vital aspect of the sortase A recognition sequence is the need for a C-terminal amide or several other hydrophilic amino acids. The use of a carboxylic acid C-terminus does not lead to a successful reaction.[99,100] Furthermore, the LPxTG sequence must be able to fit into the catalytic center of sortase A, which leads to the common use of a short, flexible linker at the N-terminus.[99] However, in most cases, this linker is not required, and its necessity must be determined on a case-by-case basis.

The following peptide sequences (Table 1) were synthesized using the previously mentioned recognition sequence information. Specific spacers that add flexibility to provide better accessibility have not been incorporated because the peptide will be connected to a hydrophilic polymer which should be flexible enough in theory. Also, longer peptide sequences necessitate more (expensive) reagents, take longer to synthesize, and typically result in a lower yield than shorter peptides.

Table 1: Peptide sequences and yields used as recognition sequences for SML.

#	Peptide sequence	Yield
1	**LPETG**-G	85-90 %
2	**LPETG**-GG	85-90 %
3	**LPETG**-GHHHH	68 %
4	**LPETG**-GHHHHHH	very low – not isolated
5	WTWTW-**LPETG**-G	81 %
6	FLFG-**LPETG**-GHG	79 %

Peptide **1** was chosen as the essential recognition sequence with only an additional C-terminal glycine for efficient SML as stated above. In **2**, another glycine is added for potentially improved performance as seen in the work from Liu *et al.*[101] **3** and **4** are equipped with a purification handle (H4-H6) to facilitate the removal of unreacted starting material via Ni-affinity chromatography. This strategy has been used several times, but usually when the recognition sequence was incorporated into a protein.[42,102] Because it is difficult to achieve high yields in SPPS when coupling several of the same amino acids,[103] 4 histidines were chosen as the minimal amount for successful Ni-affinity chromatography, with 6 beeing the standard amount. Specific amino acids were incorporated in **5** and **6**, to be able to shift the SML-equilibrium towards the products without the use of excess reagents. More details can be found in chapter 1.3. In case of peptide **6**, additional aromatic amino acids (phenylalanine, F) were added to improve detection during HPLC analysis

due to their absorbance at 280 nm. To avoid making the final peptide too hydrophobic and/or stiff for SML, hydrophilic L and G were added.

All peptides were isolated in high purity (at least 80%) and analyzed using RP-HPLC and ESI-MS. Following synthesis, the peptides were either cleaved from the resin, analyzed, and used for SML test reactions without further purification (see chapter 5.1), or a polymerization initiator (for ATRP) or CTA (for RAFT polymerization) was attached to the resin-bound peptide (see following subchapters).

Because the following synthesis steps required resin-bound peptide, resin cleavage and peptide analysis were performed once per peptide to ensure the correct molecular weight and sufficient purity.

3.1.3 Choice of ATRP Initiator and CTA

To synthesize protein-polymer conjugates suitable for pharmaceutical applications, the polymer chain length must be precisely controlled and exhibit a narrow molecular weight dispersion.[11] This requirement, combined with the availability of a wide range of PEG-substitute polymers via radical polymerization,[15] led to the decision to use controlled radical polymerization (CRP) to synthesize peptide-polymers.

ATRP and RAFT polymerization were chosen from the toolbox of CRP due to their easy-to-couple and relatively stable polymerization-inducing groups.

The majority of polymers classified as PEG-alternatives available through radical polymerization (see chapter 1.5)[13,15] are derived from "more activated" monomers such as acryl- and methacrylamides. Because these polymers are important in the context of this work, two trithiocarbonate (TTC)-based CTAs were chosen (Scheme 9). TTCs are more useful than other types of RAFT agents for maintaining good control and achieving low dispersity polymers due to their high reactivity towards "more activated" chain end radicals.[72] Another distinct advantage of TTCs over dithiobenzoate-based CTAs is their greater resistance to hydrolysis or aminolysis, which is a concern because coupling to the resin, resin cleavage and deprotection after polymerization all require (strongly) acidic reagents.[104] Cate *et al.* found significant amounts of side-product resulting from attack of the peptidic N-terminus's primary amine at the dithioester moiety, cleaving the CTA.[92] According to Hentschel *et al*., this side reaction occurs to a much lesser extent when using a TCC-based CTA.[93]

To avoid side reactions while accepting the slightly lower transfer constant, a linear alkyl chain (butyl) was used as the Z-group. To allow for the use of different solvents for polymerization of

the macro-CTA, the alkyl chain was kept in order to avoid making the CTA too hydrophobic. The latter is usually the case with the more common C_{12}-chain variant.

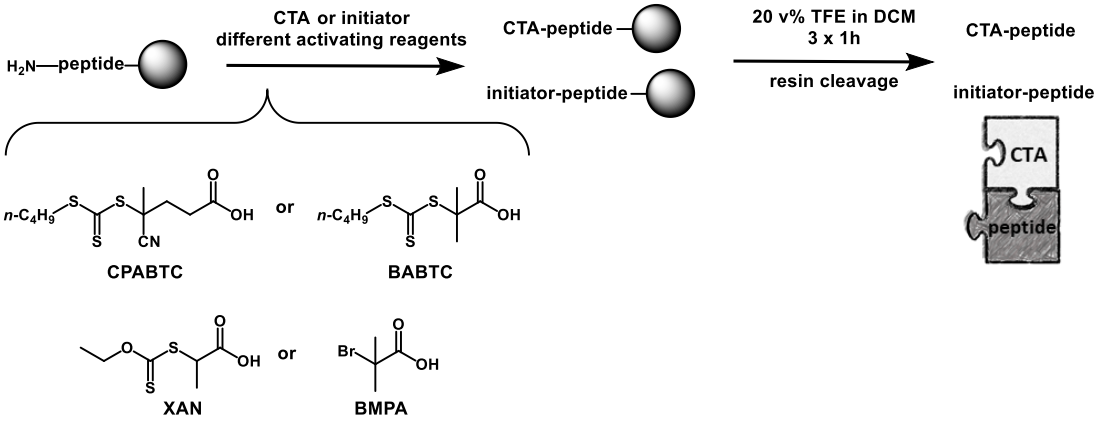
To ensure that the polymer has a peptidic end group, even if the trithiocarbonate moiety is lost during or after polymerization, the connection between CTA and peptide must be at the R group. As a result, the R group used to connect CTA and peptide had to carry a carboxylic acid moiety. The choice of attaching the CTA via its R-group will be discussed further in chapter 4 (polymerization).

A tertiary alkyl (*tert*-butanoic acid butyl trithiocarbonate, BABTC) and a tertiary cyano-alkyl (2-cyano pentanoic acid, CPABTC) R-group were chosen for their high transfer rates to cover as many monomers as possible that are considered PEG-alternatives.

Except for minor changes to the R- and Z-groups, both CTAs have been used to successfully polymerize peptidic macro-CTAs.[105] Both CTAs were synthesized with a slightly modified literature protocol (chapter 7.4).[93,105]

Another CTA based on a xanthate-moiety was required at a later stage of the project (see chapter 4.2.2 for details). Conveniently, the CTA used by Lehnen *et al.*,[87] (*O*-ethyl-*S*-(1-carboxy)methyl xanthate, XAN, Scheme 9), contains a carboxylic acid moiety as part of the R-group. According to their findings, combining BABTC and XAN resulted in very well controlled polymerizations of some of the monomers used here. As a result, it was chosen for this project as well. XAN was synthesized by J. Kurki following a published procedure.[106]

2-Bromo-2-methylpropionic acid (BMPA) was chosen as ATRP initiator because it has been used to successfully obtain peptide-polymer conjugates.[94] In addition, it is frequently used for ATRP for various applications and is commercially available.



Scheme 9: Synthesis of macro-CTA/initiator-peptide and cleavage of solid support.

3.1.4 Connection of Peptide and CTA

A CTA was attached to the N terminus of a peptide construct to make it suitable for RAFT polymerization. This was accomplished on-resin, similarly to the attachment of another amino acid, to facilitate workup and eliminate the need for extensive purification of the final product (Scheme 9). To ensure an efficient and quantitative coupling to the resin-immobilized peptide, an excess (3 eq vs. peptide) of polymerizable group and activating reagents was used, as in SPPS.

The reaction mixture was repeatedly monitored until complete conversion was achieved. In contrast to the previously used TNBSA assay, a different method was required because the resin turned orange-yellow when CTA was added, causing in a false-positive result. Another widely used method is very well suited in this case: the Kaiser test. The test is carried out by adding ninhydrin (among other additives) to a few resin beads and then heating the mixture. If there are traces of unreacted primary amines, ninhydrin will form a brightly colored dye (Ruhemann's Purple), turning the beads blue.[107] All CTA-peptide syntheses were monitored until the beads remained yellow (after 2 to 3 h reaction time).

ESI-MS analysis of the cleaved CTA-peptide from RAM-resin did not result in the expected molecular weight. Because peptide synthesis with the same resin was successful, the presence of CTA had to be the cause of this outcome.

Degradation of the TTC moiety of the CTA was deemed most likely, since it is the most fragile functional group. Degradation already during coupling, possibly caused by the base (NMM), is unlikely, because NMM and other bases used for carboxylic acids activation have very low nucleophilicity. Nonetheless, if a nucleophile (either a coupling reagent or the primary amine from the peptide) attacks the TTC moiety, a thiol is formed.

ThioGlo® (methyl maleimidobenzochromenecarboxylate) was specifically developed to monitor thiols in biological samples using fluorescence and can easily detect this thiol.[108] In short, thiols react with the maleimide moiety in ThioGlo®, converting the previously non-fluorescent molecule into a strong fluorophor. This change was monitored using a plate-reader.

Neither the ThioGlo® assay nor ESI-MS revealed any CTA-peptide fragments containing a thiol in the cleaved residue. This finding contradicts reports in the work of X. Dai facing a similar problem.[103] Several literature reports using standard reagent combinations and successfully isolating the desired CTA-peptide while using more sensitive dithioester-based CTAs are another indication that conditions during carboxylic acid activation are not damaging the CTA.[92,93]

NMR spectroscopy studies of mixtures of both CTAs with the respective reagents, either during peptide coupling or resin cleavage, were performed to investigate the cause of this degradation in greater detail. The results of exposure to TFA are depicted in Figure 7 and Figure 8.

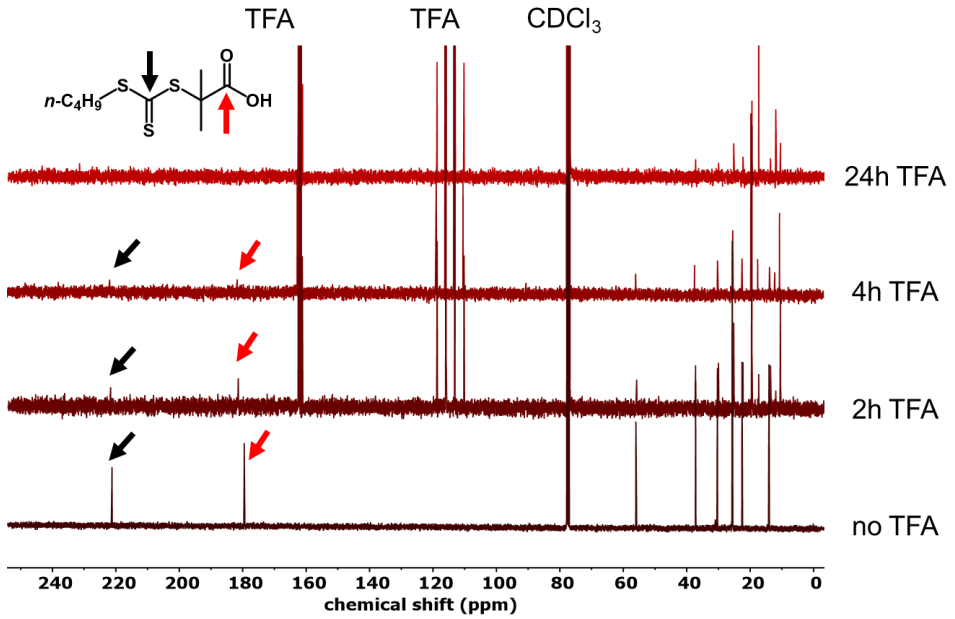


Figure 7: Stability study of CPABTC in TFA via ^{13}C-NMR spectroscopy. Complete disappearance of carbonyl and thiocarbonyl peaks after 2 h in TFA.

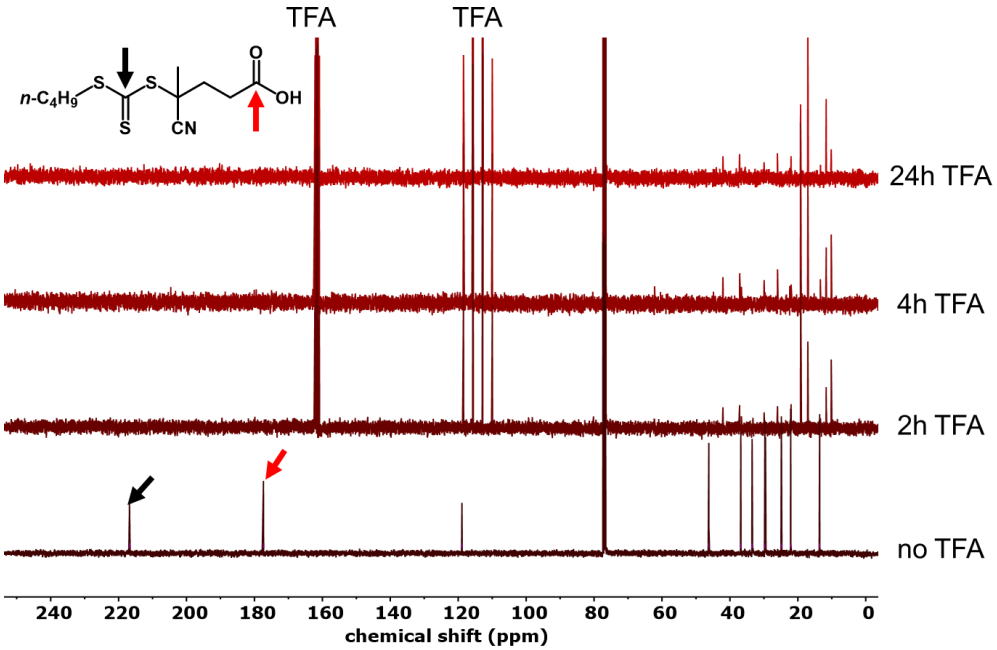


Figure 8: Stability study of BABTC in TFA via ^{13}C-NMR spectroscopy. Complete disappearance of carbonyl and thiocarbonyl peaks after 24 h in TFA.

The complete disappearance of the carbonyl- and thiocarbonyl-peaks strongly suggest that CTA decomposition occurs during resin cleavage rather than peptide coupling Both CTAs remain intact when exposed to the bases used in carboxylic acid activation but decompose quickly when in contact with the cleavage mixture or pure TFA. The findings are summarized in Table 2.

Table 2: Results of stability study of CTAs under different conditions used for the synthesis of CTA-peptide. Large excess of reagents vs. CTA.

reagent	BABTC	CPABTC
cleavage mixture 95 v% TFA, 2.5 v% H_2O, 2.5 v% TIPS[a]	not stable	not stable
pure TFA[b]	stable for up to 4 h	not stable
NMM[b]	stable for up to 24 h	stable for up to 4 h
DIPEA[b]	stable for up to 24 h	stable for up to 24 h

a Result based on complete discoloration and ESI-MS afer several hours of exposure. **b** Result based on ^{13}C-NMR spectra, TTC signal at ~220 ppm (see Figure 8).

This is surprising because the TTC moiety should be stable under strongly acidic and non-nucleophilic conditions. M. Mertoğlu discovered in his dissertation that a TTC-based CTA was stable in aqueous environment at pH 1 and 40 °C.[109] The reason for the rapid degradation observed here is unknown.

Facing a similar problem, Chen *et al.* discovered hydrolysis of the cyano-group after treating a very similar CTA (C_{12} chain instead of C_4 chain as Z-group) with cleavage mixture[110] but found no decomposition of the TTC-moiety. The hydrolysis of only the cyano-group was not observed in this study.

Large peaks corresponding to the unmodified peptide as well as several other unidentified peaks were observed in the cleaved product's ESI-MS spectra (see example in Figure 32, appendix). The presence of unmodified peptide contradicts the previously negative Kaiser test (no free amines means no unmodified peptide). It is unlikely that the observed peak is a fragment of the desired CTA-peptide because the ESI ionization technique produces very few fragments or none at all.[111]

To address the degradation issue, a different resin linker (see chapter 3.1.1) had to be used. The chosen 2-CT linker can be cleaved using TFE instead of TFA. It is important to remember that cleavage of a 2-CT linker results in a peptide with side chain protection groups still attached. If the standard RAM linker could have been used, resin cleavage as well as cleavage of all side chain protection groups could have been accomplished in a single step. According to literature reports, the functional groups in the peptide side chains are unlikely to affect polymerization.[110] This means

that the 2-CT linker method requires an additional deprotection step with TFA after polymerization. Since keeping the CTA intact is not important after polymerization, this is not a problem.
From this point forward, only 2-CT resin was used to synthesize CTA-peptide constructs.

Optimization of CTA Coupling

Various carboxylic acid activation reagents have been used in the scientific literature to couple the CTA.[91–93,112] These reagents are usually different from those used for carboxylic acid activation in SPPS, but the reasons for their selection are not given in the literature.
To optimize the purity and yield of the final construct, I tested various combinations of activating reagents, including HBTU/NMM (used for SPPS). The reaction was stopped when the resin beads had not changed color after performing the Kaiser test. The results are summarized in Table 3.

Table 3: Optimization of CTA-peptide coupling on resin support. Purity determination via RP-HPLC analysis. CTA-LPETGG on 2-CT resin.

Reaction conditions	BABTC		CPABTC	
	Yield	Purity	Yield	Purity
DIC/NMI	74-84 %	69-87 %	30 %	45%
DIC/HOBt	57-69 %	77-80 %	10 %	18 %
HBTU/NMM	61-62 %	74-80 %	-	-

Aside from changing the coupling reagents, omitting product precipitation after cleavage was an important step in increasing the yield. Typically, a peptide is precipitated in diethylether after cleavage to remove small-molecule side products and then freeze-dried to obtain a powder for easy handling. Here, precipitation of the CTA-peptide construct in diethylether was not particularly successful and some product remained in solution. Comparison of RP-HPLC elugrams of the precipitated product and the non-precipitated product showed no difference. By omitting precipitation, the yield increased significantly. To maintain comparability, none of the peptide-CTAs in Table 3 have been precipitated.
The CTA-peptides were analyzed using RP-HPLC to obtain accurate purity information (see Figure 9). The absorbtion signal at 310 nm (not shown), where the TTC moiety exhibits strong absorption, was used to determine the product peak. The results from HPLC analysis were later confirmed using ESI-MS of the product-containing fraction. Purity calculations were performed using short-wavelength absorption UV-signal (205 nm) to include all organic molecules present in the sample, yielding accurate results.

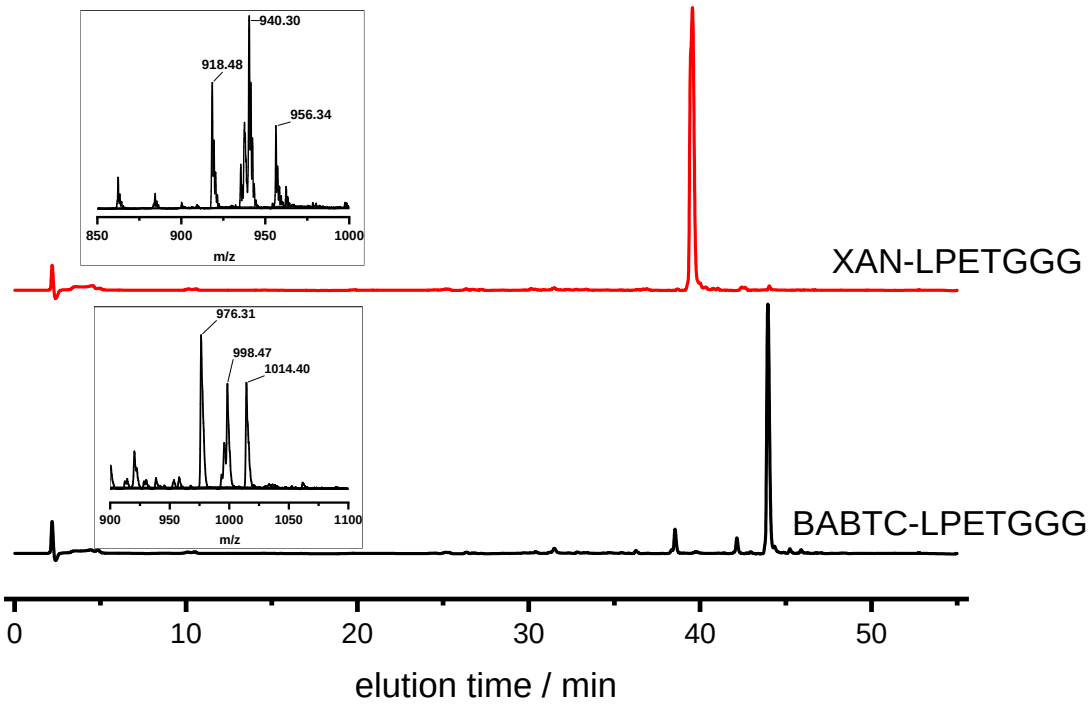


Figure 9: RP-HPLC elugrams (absorption at 220 nm, C_{18}, 5-95% MeOH/H_2O, 40 min with 0.1% TFA) and ESI-MS spectra (inserts) after optimization of reaction conditions. Expected m/z values: XAN-LPETGGG: $[M+H]^+$ 918.43, $[M+Na]^+$ 940.41, $[M+K]^+$ 956.39; BABTC-LPETGGG: $[M+H]^+$ 975.45, $[M+Na]^+$ 998.44, $[M+K]^+$ 1014.41.

These results show, that it is – at least in the case presented here – benefitial to switching carboxylic acid activation reagents from those used in standard peptide synthesis. The lower yields compared to peptide synthesis are most likely due to side-reactions, primarily the amine of the peptide N-terminus attacking not only the active ester of the CTA, but also the electrophilic carbon at the TTC-moiety.

In this regard, the cyano-moiety containing CPABTC proved to be much more sensitive/unstable. One possible explanation is that CPABTC is more vulnerable to the peptide's primary amine, due to the cyano group's rather strong -M and -I effects. This would pull electrons from the carbon atom of the TTC moiety even more strongly than in BABTC, making it more vulnerable to a nucleophilic attack.

The reaction product was tested for thiols using ThioGlo® and ESI-MS as before to determine if a nucleophilic attack on the TTC moiety was the cause for the degradation. Again, both methods revealed no thiol formation. Based on this, it is unclear why the yield and purity of CPABTC are so low when compared to BABTC.

Unfortunately, yield information is missing from literature reports using a similar synthesis approach,[93,110] so the results obtained here cannot be compared to published research.

A different coupling approach could be used in future to improve yields further: Zhao *et al.* published several different RAFT CTAs based on amino acids that were able to control the polymerization of common monomers very well.[113] These CTAs were synthesized with almost quantitative yields, and when used to couple to an existing resin-immobilized peptide, higher yields could be expected due to the CTA structure being much closer to an amino acid than the ones used here.

Because the peptide-BABTC coupling reaction produced high purities of close to 90 % (see Table 3 and Figure 9), further purification was deemed unnecessary. Aside from that, the only viable purification technique at this point would be preparative HPLC,[110] which was not available at the time of experiments.

Because the stability (see Table 2), purity and yields (Table 3) of CPABTC-peptides were significantly lower, only BABTC-peptides were used for future experiments. To increase the yield and purity of CPABTC to an acceptable level, much more optimization would be required, whereas the – for some monomers potentially suboptimal – BABTC was readily available in good yields and high purities. Furthermore, research on cyano-containing RAFT CTAs from Fuchs *et al.* revealed that these CTAs are highly likely to hydrolyze even when stored at -20 °C. The use of the (partially) hydrolyzed form resulted in a significant increase in polymer dispersity.[114]

The differences in yield and purity of CTA or initiator coupling between the different peptide sequences mentioned in Table 1 were minor. Longer peptides on the other hand, resulted mostly in higher purities than the shortest sequence (LPETGG).

Coupling of XAN to resin-bound peptide

Because the need for XAN as an additional CTA arose at a later stage of the experiments, the coupling of XAN was facilitated by the previously optimized reaction conditions discussed above. Coupling reactions carried out under optimized conditions produced high yields (up to 82 %) and purities (up to 91 %, as determined by RP-HPLC, confirmed by ESI-MS, Figure 9).

3.1.5 Coupling of BMPA to Resin-Bound Peptide

To be able to introduce the peptide as an end group in a polymer obtained by ATRP, BMPA was linked to the resin-bound peptide in the same way that the CTA was in the previous chapter. Using a combination of BMPA (5 eq), HBTU (5 eq) and NMM (10 eq) resulted in a negative Kaiser test after 4 h reaction time. Similarly to the previous chapter, precipitation of the BMPA-peptide failed and lyophilization was used to remove the solvents instead.

Because the BMPA-peptide construct was unaffected by TFA, so the RAM-resin based as well as the 2-CT-resin based constructs were used for polymerization with no additional workup or purification (see chapter 4).

Unlike CTAs, attachment of the ATRP initator resulted in near-quantiative yields with very high purities. For more information, see Figure 31 in the appendix. As a result, no further optimization of this reaction was required.

To summarize, the first chapter describes the synthesis of macro RAFT CTA and macro ATRP initiator containing the sortase A recognition peptide sequence.

A 2-CT linker was used for CTA-peptide synthesis because the standard RAM-linker resulted in construct degradation upon resin cleavage. Peptide synthesis was achieved via the “Fmoc strategy”. CTA or initiator was attached to the peptide on the resin via carboxylic acid activation, followed by resin cleavage under mild conditions.

The synthesis of CTA-peptide proved to be far more difficult than its counterpart for ATRP. As a result, the peptide-CTA coupling reaction was optimized. Due to low yields and purities, one of the two tested TTC-based CTA-peptides was not used for polymerization.

Following optimization, both CTA-peptide and initiator-peptide were obtained in high yields and good purities, necessitating no additional purification prior to polymerization. CTA and initator coupling were only slightly affected by the peptide sequence used. Longer peptides, on the other hand, resulted in slightly higher purities.

3.2 NUCLEOPHILIC CTA-PEPTIDE

Literature reports have shown that nucleophiles for SML can be as simple as any primary (but not secondary)[115] amine; however, natural nucleophiles (pentaglycine) resulted in significantly higher yields than "artificial" nucleophiles.[44] The mechanistic study by Huang *et al.* revealed that at least a diglycine with a free N-terminus is required to successfully apply SML. Longer oligo-glycines (especially 4-5) have no effect on reaction speed or efficiency.[50] Therefore, all peptidic nucleophiles used in this study consisted of 2-3 N-terminal glycines.

To overcome the problem of SML being an equilibrium reaction, some strategies use a nucleophile change to make the reaction irreversible. As an alternative nucleophile, a Trp-Zipper (GG-WTWTW)[51] was used (introduced in chapter 1.3). The outcomes of this strategy are discussed in chapter 5.1.

3.2.1 SYNTHESIS OF PEPTIDE PRECURSOR

In general, the synthesis of a peptide-polymer containing the nucleophilic peptide sequence for SML is much more difficult than the recognition sequence peptide, because a glycine with a free primary amine is required at the N-terminus of the peptide. As a result, the polymerization-inducing group (CTA or ATRP initiator) cannot be easily attached to the N-terminus of the growing peptide as shown in the previous chapter. Another challenge is the inherent instability of CTAs towards nucleophiles, which was discussed in the previous chapter. The N-terminal primary amine needed to be protected to solve this problem. This protection group was then cleaved after polymerization to activate the peptide-polymer for SML.

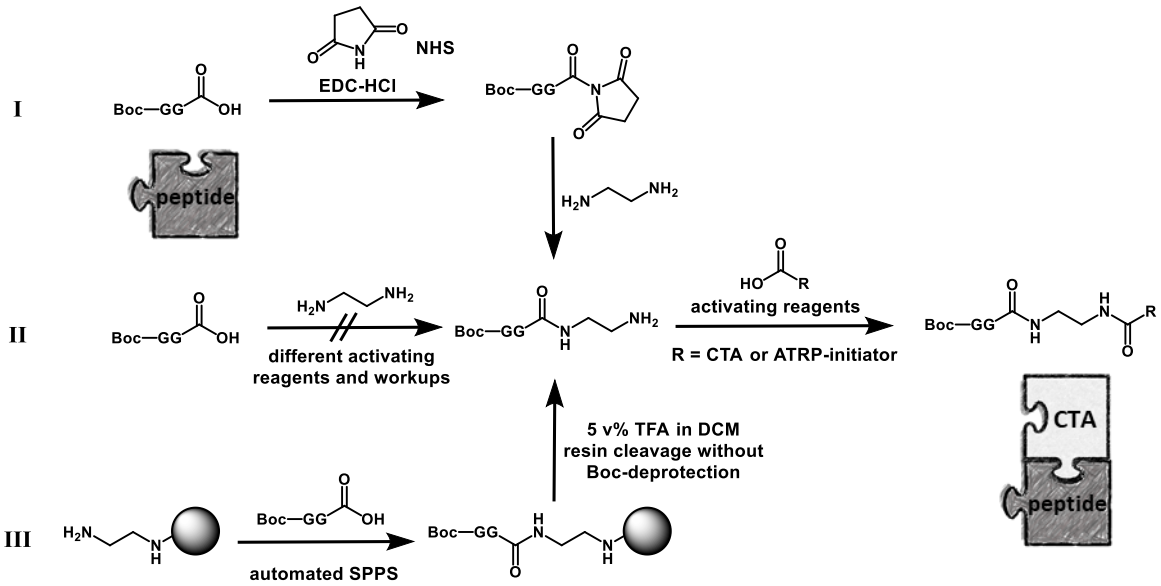
Thus, the first building block was a Boc-protected diglycine (Boc-GG) that was commercially available. Boc was chosen as the protection group over Fmoc, which is typically used in peptide synthesis, because Fmoc is cleaved under basic conditions used during carboxylic acid activation. The next step in the synthesis was to attach a bifunctional primary amine (ethylenediamine, EDA) to either the C-terminus of Boc-GG or the CTA/ATRP initiator, allowing Boc-GG and CTA/initiator to be connected.

X. Dai's previous work on this topic took a slightly different approach using ester bonds to connect SML-nucleophile to the CTA.[103,116] This synthesis approach worked well, resulting in acceptable yields. I decided against ester bonds for this project, because they are susceptible to hydrolysis in aqueous media with high and low pH values. Keeping in mind the overall goal of this project,

extreme pH values may occur during application, posing a risk of cleaveing protein from polymer or separating polymer blocks.

The use of amide bonds results in more challenging conditions (i. e. the presence of primary amines), particularly for the CTA (as seen previously in chapter 3.1.4), but was deemed worthwile to avoid potential hydrolysis of the final construct.

Several pathways have been investigated in order to synthesize the desired constructs. They are depicted in Scheme 10. Longer peptides other than diglycine are not shown for clarity.



Scheme 10: Synthesis pathways to obtain nucleophilic sequence for SML with polymerizable group attached. **I**: NHS-ester approach. **II**: Solution-based approach, direct coupling. **III**: Resin-based approach.

First, synthesis pathway **I** was followed, which used an N-hydroxysuccinimide (NHS) ester as an activated species to couple ethylenediamine (EDA) as a bifunctional amine. This strategy has been reported in a patent,[117] though it uses triglycine rather than diglycine. Unfortunately, even when using triglycine, I was unable to reproduce the results, partly due to the lack of access to preparative HPLC equipment for purification. The desired product was formed according to ESI-MS analysis, but significant amounts of impurities were present. Due to the fact that EDA is added directly to the NHS-ester, double coupling on a single EDA molecule is possible, resulting in a significant amount of unwanted side product. Purification with silica-based column chromatography was not successful. Based on the work from X. Dai[103] and Suguri *et al.*[27] (both using the ester-equivalent) EDA coupling with various other activating reagents for carboxylic acids (such as DCC/DMAP, DIC/NMI, HBTU/NMM, pathway **II**, Scheme 10) was attempted. Again, the desired product was

detected on occasion, but solution behavior (spontaneous, non-reproducible precipitation) rendered purification impossible. Attempts to avoid side-reactions by using an Fmoc-protection group to render one primary amine of EDA unreactive were made, but the purity was not significantly improved despite the addition of two laborious protection/deprotection steps.

Finally, pathway **III** was used, which made use of a commercially available resin that had been prefunctionalized with EDA. Because of its two primary amines, EDA can act as precursor for SPPS and will yield a primary amine after resin cleavage. In addition, pathway **III** is much more adaptable than pathways **I** and **II,** when alternative nucleophiles or longer nucleophilic peptide sequences are required.

A trityl-linker which can be cleaved using mildly acidic conditions (see chapter 3.1, Resin Linker Choice) was used by the supplier to connect resin and EDA. It was important to avoid strongly acidic conditions to not deprotect the primary amine of Boc-GG during resin cleavage. The Boc-protection group is required to ensure that the CTA coupling happens at the primary amine of EDA which is accessible after resin cleavage. Trials to cleave the resin conducted under very mild conditions (20 v% TFE in DCM, used primarily for cleavage of a 2-CT linker) yielded little to no product. Slightly more acidic conditions using 5 v% TFA in DCM resulted in efficient resin cleavage. In addition, the resin needed to be extracted with acetonitrile (MeCN) to efficiently remove the product because of the product's poor solubility in the cleavage mixture. Product precipitation in diethylether was not possible, as a consequence, the solvents were removed *in vacuo*. The addition of MeCN besides its function to dissolve the cleaved peptide efficiently, proved to be critical: During evaporation without additional MeCN, TFA was concentrated because it evaporated slower than DCM. As a result, the concentration of TFA in the sample rose significantly, which caused Boc-group cleavage (as seen by ^{1}H-NMR-spectroscopy, see Figure 10, integral of Boc group 3.4 without MeCN vs. 8.5 with MeCN). The addition of MeCN prevents concentration of TFA because TFA is more volatile than MeCN and is thus removed from the sample first. This phenomenon has previously been reported, and the published method was used to avoid it here.[118]

Despite method adjustment, RP-HPLC analysis using a C_{18} column yielded no useful information due to the very hydrophilic product. To confirm the successful synthesis and product purity, NMR spectroscopy (Figure 10) and ESI-MS (Figure 33, appendix) were used. Even with the addition of MeCN, the NMR spectrum shows that a partial removal of the Boc group occured, as indicated by the lower than expected (9.0) integral of the Boc-group peak at 1.45 ppm. Nonetheless, when compared to the experiment without addition of MeCN, this result is significantly better.

Further purification was not performed, because after CTA coupling, purification via column chromatography was necessary. The Boc-deprotected compound would be removed, regardless of whether or not it reacted with the CTA at the N-terminus.

Because the obtained product was very hygroscopic, it was used immediatly for CTA coupling after all solvents were removed to avoid contamination with water. Coupling of BMPA was not attempted because parallel ATRP experiments yielded no polymer and ATRP trials were discontinued (see chapter 4.1).

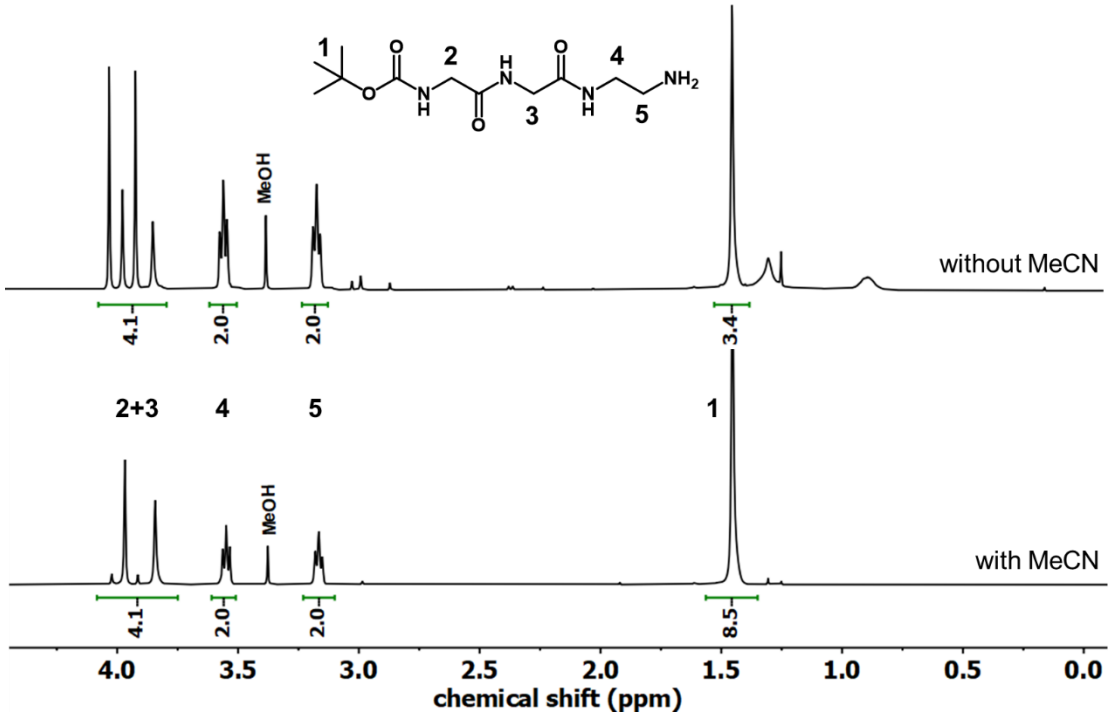


Figure 10: ^{1}H-NMR spectra of Boc-GG-EDA with and without addition of MeCN during solvent removal. Integration relative to signal 4.

3.2.2 Coupling of CTA and Peptide Precursor

Because the coupling reaction parameters for the CTAs used here (BABTC and XAN) had already been optimized (see chapter 3.1), the same carboxylic activation reagents (DIC/NMI) were reused. The cyano-group containing CTA CPABTC was not used based on the findings in chapter 3.1. Because all reagents were in solution and not on solid support, the only difference was a much lower excess of CTA and reagents (1.1 eq vs. 3 eq previously).

The reactions were monitored using the Kaiser test and the pure product was obtained in good (55 % BABTC), and acceptable (28 % XAN) yields after solvent removal and column chromatography. Surprisingly, this more complex synthesis resulted in higher yields than previously reported for the combination of Boc-GG-BABTC (43 %[103] and 37 %[27]) via ester-bond.

For future improvement, a different strategy could be used. The final coupling of CTA and peptide produces the least amount of product for both recognition sequence and nucleophilic constructs. Instead of coupling an already completed CTA, the on-resin peptide synthesis could be changed to produce a usable building block at C- or N-terminus for assembling the TTC-moiety. The coupling of the R-group via a good leaving group (alkylhalide) is a common technique for obtaining unsymmetrical TTC-based CTAs.[119]

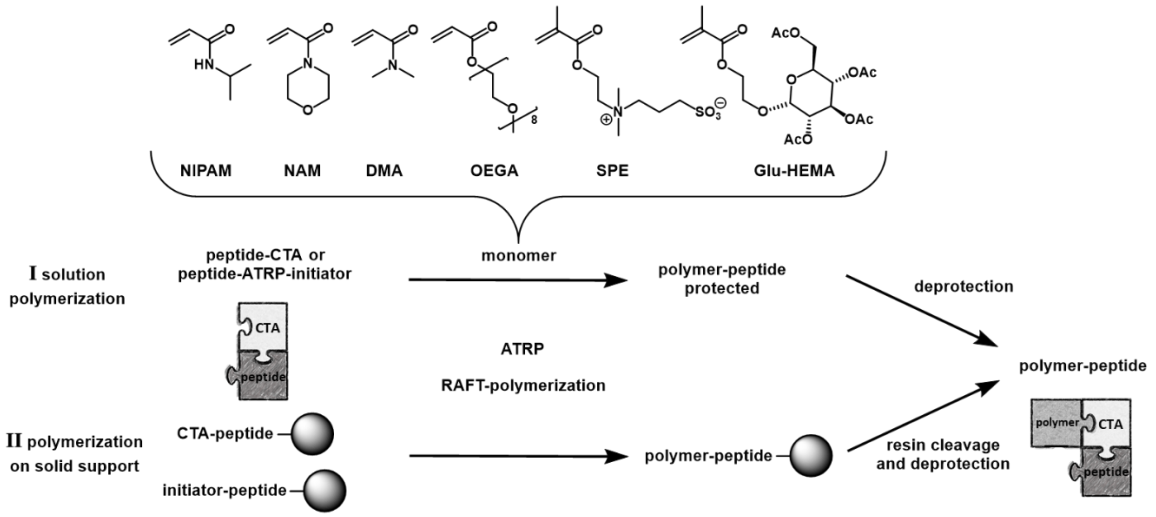
In case of the recognition sequence peptide (CTA attachment at the N-terminus), a peptide containing a secondary or tertiary alkylhalide at the N-terminus is easily obtained as it has been shown here with the coupling of the ATRP initiator (chapter 3.1.5). In fact, the conversion of a peptide-based ATRP initiator to a peptide-CTA has been demonstrated by Cate *et al.*[92,112] The reaction was carried out on-resin, resulting in almost quantitative yields without the need for purification. Despite the differences in CTA structure (dithiobenzoate vs. TTC), this strategy should be applicable to the system used here.

Again, because SPPS proceeds from C- to N-terminus, the synthesis of the nucleophilic building block with a Boc-protected N-terminus carrying the CTA at the C-terminus is more difficult and laborious. A possibility would be the attachment of the aforementioned ATRP initator as CTA precursor to the N-terminus after cleavage of the peptide from the EDA-preloaded resin. The TTC-formation off-resin, which would necessitate (likely difficult) purification of the final product, is a significant disadvantage. Nonetheless, as demonstrated in chapter 3.1.4, attaching an ATRP initator to a peptide yields more product than the TTC-based CTA counterpart.

In summary, this chapter discusses the results and findings of synthesizing the nucleophilic peptide sequence (diglycine) required for SML with an attached CTA. One goal of this newly developed synthesis route was to avoid the use of ester couplings in order to avoid fragility towards hydrolysis. After several failed trials, a resin-based approach involving EDA-prefunctionalized resin, followed by resin cleavage and CTA coupling was developed, resulting in pure product and high yields.

4 POLYMERIZATION AND DEPROTECTION

Polymers with various peptidic end groups suitable for SML were synthesized using the peptide-CTAs and peptide-ATRP initiators discussed in chapter 3. In accordance with the objectives of this work presented in chapter 2, the monomers used for polymerization were under ongoing research to be used as PEG-alternatives. As a result, a variety of monomers polymerizable via ATRP or RAFT polymerization were chosen from the portfolio of promising PEG-alternatives.[15] Scheme 11 depicts their structures as well as the general synthesis approach. The acrylamide-based monomers (*N*-acryloylmorpholine NAM; *N*,*N*-dimethylacrylamide DMA) were chosen because they have all been thoroughly studied and are regarded as "standards" in the field of biocompatible polymers. Furthermore, poly-(*N*-*iso*propylacrylamide) (P(NIPAM)) was chosen because of its thermoresponsive solution behavior in water, which could result in interesting effects for protein-polymer conjugates.[27,120] Polymeric oligoethyleneglycol acrylate (P(OEGA)) was chosen for its similarity to PEG. It is, however, considered a valid replacement for the latter because it does not result in the same antibody production.[13,121] Moreover, polymerization of OEGA will yield so-called bottle-brush polymers, which are sterically much more complex and may pose a challenge for SML.



Scheme 11: Synthesis scheme to obtain peptide-polymer constructs ready for SML. All the monomers used in this work are shown.

The zwitterionic monomer 3-[dimethyl-[2-(2-methylprop-2-enoyloxy)ethyl]azaniumyl]propane-1-sulfonate (SPE) was chosen as commercially available, well understood representative of the promising class of polyzwitterions as PEG-alternatives.[122] Because of their complex solution behavior, zwitterionic polymers will provide a challenge for both peptide-polymer synthesis and

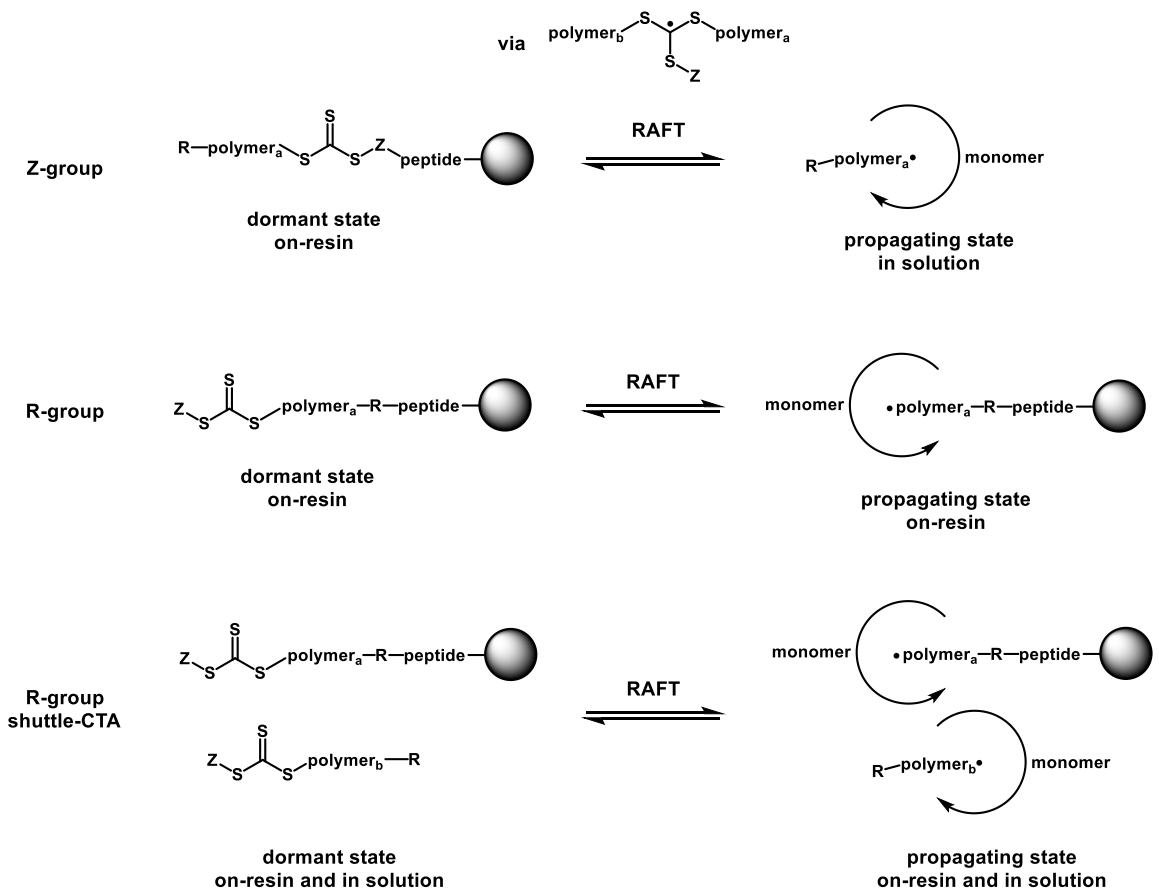
SML. Finally, the class of glycopolymers is represented by the glycomonomer Glu-HEMA, which is based on glucose and was made polymerizable by the addition of hydroxyethyl methacrylate (HEMA). Because this monomer, and glycomonomers in general, are not readily commercially available, it was synthesized using a literature report (for experimental data, see chapter 7.6).[123] Glycopolymers, with their demanding sterics and strong attractive interactions caused by many hydrogen bonds, could pose a challenge for SML.

The peptide-polymer conjugates were obtained via two distinct pathways, as shown in Scheme 11. First, a classical solution polymerization approach (pathway **I**) was used, followed by deprotection of the peptide side chains to activate the end group for SML (see chapter 3.1.1). Second, there was the more elegant and much less labor-intensive approach, in which the polymer was grown from the immobilized peptide with a CTA or ATRP initiator attached (*grafting-from*, pathway **II**). ATRP and RAFT polymerizations have both been conducted successfully on Merrifield-type resin[124–128] (as used here) before, though the number of reports is very limited given the method's enormous potential.

The main benefits of surface-initiated (SI) polymerization (see introduction chapter 1.5) are higher purity (impurities and catalysts can be removed with simple washing steps) and the ability to use the resin-bound macro CTA/initiator immediatly after synthesis without resin cleavage. TFA can be used for the final resin cleavage after polymerization and simultaneous peptide side chain deprotection because the CTA's intactness is no longer required.

Pathway **II** can only be used to synthesize polymers with recognition sequence end groups because its nucleophilic counterpart had to be synthesized off-resin (chapter 3.2.2). To synthesize the nucleophilic peptide-polymer, solution polymerization via pathway **I** is required.

Scheme 12 depicts yet another decision that needed to be made: There are several ways to attach a CTA to a solid support.[75] Either via the radical-stabilizing group (**Z-group**), or the leaving- and reinitiating group (**R-group**). It is now widely accepted that without the addition of free CTA in solution, neither the R-group nor the Z-group approaches provide good control over SI-RAFT polymerization.[125,129] The addition of this "shuttle CTA" improves reaction control by allowing for easy and efficient chain transfer between polymer chains in solution and on the support. In Scheme 12, bottom line, the use of a "shuttle CTA" is depicted. Only the "shuttle CTA" R-group strategy is shown for clarity.



Scheme 12: Comparison of different attachment options of CTA onto solid support and the effect on the RAFT polymerization mechanism.

Anchoring the CTA with its **Z-group** results in the propagation of the active polymer chain in solution, making this approach technically a *grafting-to* method, according to the commonly accepted mechanism for RAFT polymerization. This has several advantages, including the fact that only living chains are attached to the resin, allowing for easy separation of unwanted byproducts and monomer. Furthermore, since propagation occurss in solution, monomer diffusion and thus accessibility are much better than when the growing polymer is fixed in position on the support. The synthesis of long polymer chains, on the other hand, may become very inefficient because the polymer chain becomes bulkier throughout the reaction, making attachment back to the resin (dormant state) difficult. Several studies[75,130–132] have found that using the **Z-group** approach results in much lower grafting densities due to bulky polymer chains.

Furthermore, synthesis or attachment of the CTA via **Z-group** on the peptide-loaded resin would necessitate the development of a novel synthesis path in comparison to the carboxylic acid moiety at the R-group found in all the CTAs of this study.

In contrast, the **R-group** strategy produces a polymer that remains attached to the peptide (and thus the solid support) throughout all stages of the polymerization (*grafting-from*). That means, after polymerization and rinsing of the resin and cleavage, dead polymer chains and other side products can be removed In addition, it has been reported that controlling the reaction may be difficult on densely grafted surfaces due to slow diffusion of growing polymer chains, which may inhibit chain transfer.[75]

The decision to only use the R-group strategy for polymer-peptide construct synthesis was made not only because CTA attachment to the resin-bound peptide is easier. The separation of peptide and polymer by the TTC-moiety is the primary disadvantage of the Z-group approach. If this (somewhat fragile) moiety was destroyed, the polymer would be unable to be used for SML or, worse, would separate the polymer from the protein in the final conjugate (see Scheme 12).

4.1 ATRP

SI-ATRP was chosen as the first method to be tested because it is less complicated than SI-RAFT polymerization (requires the use of sacrificial CTA, see Chapter 4.2.1 for details).

Not many literature studies that included in-depth analytical data and detailed descriptions using initiators that were immobilized on Merrifield-type resins were found, but the advantages of very easy initiator-coupling to the peptide compared to CTA coupling for RAFT polymerization (see chapter 3.1.4) and very promising reports[126–128,133] led to this decision.

Experiments were started based on the most recent report on the subject by Trzebicka *et al.*, employing CuBr and Me_6TREN as catalyst system and DMF/H_2O (3:1) as reaction medium.[126] In 21 h reaction time, they polymerized NIPAM, resulting in a polymer with slightly higher than expected molecular weights and very good dispersity (D = 1.14).

Unfortunately, using the previously described resin-bound macro-initiator based on the peptide sequence LPETGG, polymerizing with the conditions described by Trzebicka *et al.* resulted in no polymer. Only the unreacted macro-initiator was found after resin cleavage and ^{1}H-NMR spectroscopy. A change in catalyst amount, type (CuCl instead of CuBr), or ligand (PMDTA instead of Me_6TREN) had no effect on the outcome of the experiment.

On first glance, a possible explanation is the use of a different resin than in both literature reports, however the chemical composition of all resin matrices (crosslinked polystyrene with 1 % divinylbenzene) is the same, the only difference is resin prefunctionalization. This leads to another potential reason: a different peptide sequence as well as resin prefunctionalization (with a dye)[126] were used. The use of an (here inappropriate) peptide sequence could have limited the accessibility of initiators to radicals. On the other hand, it is highly unlikely that the short peptide sequence (LPETGG) forms secondary structures that would limit the initiator's radical accessibility. Additional reasons are discussed in chapter 4.2.1 together with the results of surface-initiated RAFT polymerization.

Following the failed SI-polymerization, classical ATRP in solution was attempted. Despite numerous literature reports on peptide-polymers synthesized via solution-ATRP,[90,134,135] no combination of solvent, catalyst system and monomer yielded usable polymers. The reasons for this are unknown. Because RAFT polymerization was carried out in parallel and produced promising results, no further attempts at optimizing ATRP were made.

4.2 RAFT POLYMERIZATION

4.2.1 SURFACE-INITIATED RAFT POLYMERIZATION

Exhibiting the same advantages as SI-ARTP, SI-RAFT polymerization was tested next. Again, detailed studies on the use of Merrifield-type resins as solid supports are lacking, whereas surface-initiated polymerizations from a variety of (smooth) supports have been extensively researched.[75] To the best of my knowledge, the group of S. Perrier published the only two papers describing successful SI-RAFT polymerization using Merrifield resins.[125,131]

Based on the information in these reports, an attempt was made to transfer their method to the system used here. Perrier *et al.* used the Z-group approach (see above) and reported good control only when free CTA was used as "sacrificial CTA" or "CTA-shuttle" in addition to the resin-bound CTA. Free CTA is needed to maintain reaction control, because otherwise the (much faster) free radical polymerization occurring in solution and the RAFT polymerization on resin compete for monomer. As a result, the on-resin polymer has a much shorter chain length than the polymer formed in solution if no free CTA is used. Furthermore, they reported limited conversion and potentially limited molecular weights (MW), because steric hindrance of large molecules renders recombination with the CTA on the resin difficult.[125]

This problem is not expected here, because I used the R-group approach, in which the growing polymer remains attached to the resin at all times. In addition, the part of the CTA carrying the Z-group (in solution when the polymer chain is propagating) is a small molecule with very little steric hindrance.

Several polymerizations were carried out using the literature conditions, including the "shuttle CTA" (1:1 BABTC vs. resin-bound CTA-peptide). As a radical source, AIBN was used, and the solvent was freshly distilled 1,4-dioxane or DMF. Because the majority of the CTA-peptide is attached inside the porous bead rather than on the surface of the resin bead, special care was taken to ensure that the resin swelled well in the solvent used. This was realized by swelling the resin in the desired solvent at least 10 min prior to polymerization

To monitor monomer conversion, samples from the reaction solution were collected and analyzed using ^{1}H-NMR spectroscopy. Monomer conversions for DMA and NIPAM reached 65-70 % within 5 h and did not increase further until the reaction was stopped after 11 h. SEC-analysis of the polymer formed in solution revealed good control ($Đ$ = 1.18) with higher than expected MW of M_n = 16 kg/mol (theoretical MW 12 kg/mol for 100 % monomer conversion). Considering the limited monomer conversion, this difference is significant, despite the use of a chemically different polymer standard for SEC calibration. After resin cleavage, it was found that only minimal amounts

of substance could be isolated. ESI-MS analysis revealed that only unmodified macro-CTA was isolated, indicating that no polymerization occured on-resin. NMR analysis was not possible due to the small amount of residue. If the polymerization had worked as expected, the amount of polymer on the resin would have been sufficient for NMR analysis.

Both the SI-ATRP and SI-RAFT polymerization results indicate that the initially formed radicals did not reach the polymerizable groups on and within the resin bead. The inability of the starting radicals to reach the CTA or ATRP initiator may be due to diffusion limitation caused by steric hindrance and very limited space within the resin beads. In the case of RAFT polymerization, the presence of "shuttle CTA" slows the competing polymerization in solution by switching the polymerization mechanism from free radical polymerization (without "shuttle CTA") to a CRP. This effect appears to be insufficient to allow efficient chain transfer to the resin-bound CTA.

Because all starting radicals have been generated by thermal decomposition thus far, a different approach to start the polymerization was tested.

The recently developed photoinduced electron/energy transfer RAFT (PET-RAFT, introduced in chapter 0) polymerization uses a direct electron or energy transfer from a photocatalyst to the CTA which then initiates polymerization.[136] PET-RAFT was successfully initiated on various surfaces[85,137] including hydrogels,[138] but not on porous Merrifield resin.

The key difference to thermally induced SI-RAFT is that no free radicals are present in solution and thus no competing polymerization is happening in solution when the R-group is immobilized on the resin. Unfortunately, the same rules concerning good control of thermally induced SI-RAFT polymerization apply: The use of a "shuttle CTA" in solution is required. Of course, this negates the main advantage of PET-RAFT, which is that the chain grows exclusively from the surface.

Surprisingly, Seo *et al.* found that using low-energy green light rather than high-energy blue light resulted in good control without the use of CTA in solution.[137] Their proposed explanation for this finding is that using blue light irradiation causes multiple initiation mechanisms, resulting in poor control in the absence of CTA. These encouraging results enabled me to test the resin-bound macro-CTA without competing solution-polymerization while keeping the reaction under control.

The organic dye Eosin Y was chosen from among the many different photocatalytic systems available thus far, owing to its excellent performance in catalyzing the polymerization of similar monomers to those used here.[136]

Based on the findings of Xu *et al.*[84] and Seo *et al.*,[137] a setup using green light LED strips (515-525 nm, 4.8 W/m) or blue light LED strips (460-465 nm, 14.4 W/m, dimmed to 50 % intensity) was employed (see Figure 11). Because PET-RAFT polymerization proceeded much faster in

Figure 11: SI-PET-RAFT reaction setup using blue and green LED strips, Schlenk flasks, and a magnetic stirrer.

DMSO than in other solvents, it was chosen over some of the more commonly used solvents.[139] The resin used as solid support showed very good swelling properties in DMSO.[139] Vigorous stirring resulted in a good mixing of the reaction dispersion that ensured equal light exposure for all resin-beads.

To avoid complicating the reaction setup further, no oxygen-tolerant PET-RAFT (by addition of a tertiary amine[140] or "enzyme degassing"[137]) was used, and the reaction mixture was degassed before starting the reaction by exposing it to light.

After 16 h of irradiation, washing, and resin cleavage, very little residue was isolated and confirmed by ESI-MS analysis to be CTA-peptide. This result was obtained regardless of the monomer or light source used. As expected, the ^{1}H-NMR analysis of the reaction solution showed no polymer formation.

Following the negative results obtained with SI-PET-RAFT, no further research into SI polymerization was conducted. Despite its great potential for peptide-polymer synthesis, the very small number of literature reports using Merrifield-based SI-polymerization may indicate general

problems with the method. In addition, to the best of my knowledge, no report on successful SI-RAFT polymerization using the R-group approach has been published.

Based on the large number of reports of successful SI-polymerization on various surfaces, it appears that the type of solid support used here is not suitable for SI-polymerization.

The reason that comes to mind is the high porosity that all Merrifield-type resins share. They are typically crosslinked with a very small amount of crosslinker (typically 1 %), making them swellable and, due to the large available surface, very efficient in SPPS. The majority of heterogeneous polymerizations reported in the scientific literature use flat, nonporous surfaces.[75]

This could be a major issue for the system used here. Because it is only possible to anchor the CTA to the resin via the R-group, propagating radicals are always bound to the surface and never in solution (as it is the case when using the Z-group approach, see Scheme 12).

In the case of SI-PET-RAFT polymerization when using the R-group approach, the starting radical would be formed exclusively on the surface. Given that no polymer was found and ESI-MS revealed that not even oligomers were formed, one could conclude that the initiating photocatalyst never even reached the CTA-peptide to form radicals.

The reason for this is most likely that the vast majority of macro-CTAs are located on the inside of the resin bead rather than on the surface.[137] Because the lifetime of the excited state of the photocatalyst is very short, visible light penetration into the resin pores is insufficient to create a significant amount of radicals to initiate polymerization. Plus, the addition of “shuttle CTA” should facilitate chain transfer between resin and solution. The inclusion of "shuttle CTA" should aid in chain transfer between resin and solution. Still, polymerization was not taking place, which could be due to diffusion limitations of the monomer and polymer chains within the resin, the formation of dead chains through recombination, or simply a low concentration of initiating radicals within the resin as discussed before.

In conclusion, contrary to some reports in the literature, neither thermally induced SI-ATRP nor SI-RAFT polymerization, nor light-induced SI-PET-RAFT polymerization using BMPA-LPETGG or -CTA-LPETGG yielded polymers. The reason why SI-ATRP did not yield polymers remains unknown since very similar reaction conditions were reported in literature,[126–128] showing successful and well-controlled polymerization using Merrifield resin as solid support.

Concerning the failed attempts of SI-RAFT, the reason is most likely the porosity of the Merrifield resin combined with the attachment of CTA via the R-group. There are no literature reports on R-group SI-RAFT polymerization from porous supports and there are only a few reports on SI-RAFT on Merrifield resin (Z-group approach).[125,131]

None of the aforementioned circumstances are easily changeable. The use of a nonporous support for SPPS would result in low amounts of peptide, requiring large amounts of resin to synthesize usable amounts of peptide-polymer and potentially necessitating the development of a new reaction procedure. Because of the reasons stated above, attaching the CTA via Z-group is not worthwhile. As a result, macro-CTA polymerization was carried out in solution.

4.2.2 RAFT Polymerization in Solution

Because none of the SI-polymerizations and ATRP in solution produced usable polymers (see previous chapters), the previously resin-bound macro-CTA suitable for synthesis of polymers bearing the recognition sequence end group (e. g. BABTC-LPETGG) was cleaved from the resin (for details see chapter 3.1.4) and used in solution-based RAFT polymerization. Because the synthesis of macro-CTA with a nucleophilic end group (e. g. GG-BABTC) required resin cleavage prior to CTA attachment, RAFT polymerization in solution was the only option (see chapter 3.2 for more information).

Many improvements to the initial reaction conditions, namely radical formation by thermal decomposition of an initiator molecule, have been made since the beginning of the development of the RAFT process. Very promising in the context of this project is light-induced generation of radicals, directly using the CTA to initiate polymerization.[74] This method eliminates the need for an external radical source resulting in higher amounts of usable polymer. The external initiator (typically used in a 1:10 molar ratio vs. CTA) initiates polymer chains that lack a peptidic end group and are thus unusable in SML, resulting in 10 % of polymer chains being unusable.

Another advantage of this method over SI-RAFT is that the same polymerization parameters can be used to polymerize both the nucleophilic and the recognition sequence CTA-peptides.

PET-RAFT

The previously described PET-RAFT method (chapter 4.2.1, Figure 11) was tested to polymerize CTA-peptide. It is was based on a photocatalytic system using the dye Eosin Y. The photocatalytic energy/electron transfer ensures, that all polymer chains have the appropriate peptidic end group.

Both monomers used for initial testing (DMA and NIPAM) have been shown to be suitable for PET-RAFT: In the first publication about PET-RAFT, Xu *et al.* reported 93-95 % monomer conversion, MWs close to the theoretical values and very narrow MW-distributions of $Đ = 1.08\text{-}1.09$.[83]

Test polymerizations of DMA and NIPAM with BABTC (no peptide) initiated by green or blue light resulted in high monomer conversions (85-87 ‘% after 22 h reaction time). In contrast to the findings of Seo *et al.*,[137] the wavelength of the light source had no effect on the polymer.

Polymerization of DMA resulted in a MW (M_n = 8.7 kg/mol, SEC in THF, expected 16.8 kg/mol) much lower than the theoretical value with relativley high dispersities Đ = 1.34. The measured MW values may be inaccurate due to the polystyrene standard used for SEC calibration being chemically very different. Surprisingly, the MW values obtained for PNIPAM were close to the expected MW, albeit with a broad MW distribution with Đ = 1.56.

Similar monomer conversions were obtained with BABTC-LPETGG reactions, and the use of DMA resulted in a much lower than expected MW and a high dispersity (Đ = 1.47). Using NIPAM on the other hand, resulted in polymers with higher than theoretical MW, which is to be expected given that CTA-peptide purity was less than 100 %. However, a high dispersity of Đ = 1.49 indicates that the reaction was not particularly well controlled. SEC data (Figure 34, appendix) reveiled a high MW tailing of all samples, indicating a loss of control.

One possible explanation is the nature of the R-group: In the literature report used for these experiments,[83] the R-group was secondary, whereas the R-group used here was tertiary. Other studies found that a photoiniferter-RAFT (PI-RAFT) polymerization initiated by blue light and mediated by a secondary R-group did not propagate, whereas the reaction proceeded when a tertiary R-group was used.[141] The reported dispersities were high, similar to those ones found here and were attributed to CTA decomposition under blue light, as confirmed by another report.[142] This suggests that by using a tertiary R-group, the polymerization might have been initiated directly by blue light rather than the photocatalytic system, resulting in the relatively high dispersities. However, this does not explain why using green light (no photoinitiation of tertiary R-groups was reported) did not result in lower dispersities.

A different (secondary) CTA may result in well-defined polymers, according to this potential cause of poor control. However, changing the CTA would have been time-consuming because the connection to the peptide and resin cleavage would have had to be optimized again.

In retrospect, there is a potentially very simple solution to the problem of high dispersities caused by CTA degradation when exposed to blue light: Qiao and colleagues performed well-controlled PI-RAFT under UV light using TTC-based CTAs (with a tertiary R-group) in combination with a tertiary amine.[143] Because the same homolytic bond dissociation is expected when using blue light, the amine’s effect of reversibly reducing the formed TTC radical to a more stable anion should work in the same way for PET-RAFT as it does for PI-RAFT.

With mixed results obtained from the initial tests and the option to use the very recently developed method of XPI-RAFT, no further optimization of PET-RAFT was initiated.

XPI-RAFT

XPI-RAFT, described in chapter 1.5.3, was developed in the group of M. Hartlieb during the experimental phase of this work and was thus only used at a later stage of polymerization experiments.

The combination of TTC- and xanthate-based CTAs in an XPI-RAFT process provides three distinct advantages (among several others when compared to thermal RAFT) in this work: To begin, the reaction setup is very user-friendly and does not require any special equipment or reagents other than the CTAs. Second, because no external radical source is required, all polymer chains have a peptidic end group. Third, the TTC-based CTA can be exchanged with little effort to make a wider range of monomers accessible and changing the TTC-CTA should have no effect on the other polymerization parameters because the main radical source (XAN-based CTA) remains constant.

A.-C. Lehnen and coworkers showed impressive results when testing the oxygen tolerance of the XPI-RAFT process.[87] However, when the reaction was exposed to air in some of their experiments, they observed a loss of control. Based on these findings, the reaction mixture was degassed prior to illumination by purging with nitrogen. Pictures of the reaction setup are shown in Figure 12.

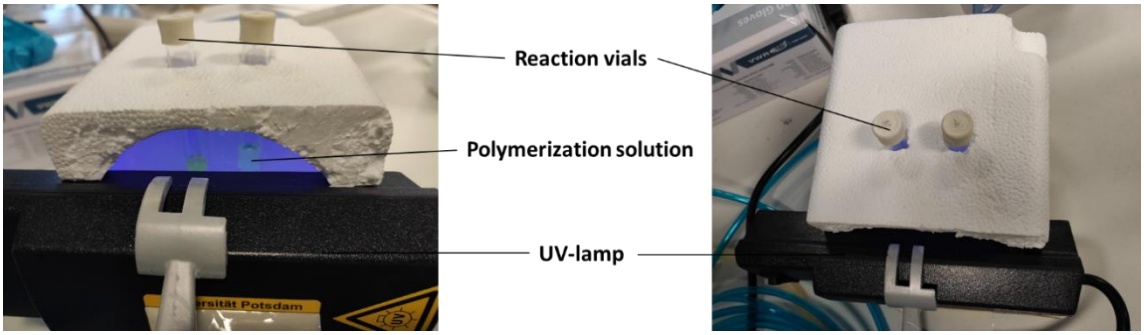


Figure 12: XPI-RAFT polymerization setup (side and top view).

All of the previously mentioned monomers (NAM, NIPAM, DMA, OEGA, Glu-HEMA, SPE) were polymerized using the nucleophilc and the recognition sequence macro-CTA after the first test polymerizations yielded very promising results. A peptide-XAN to peptide-BABTC molar ratio of 2:8 was used for all reactions because it provided the best balance of reaction control and short reaction times.[87] All polymerizations were carried out with the aid of a standard 365 nm UV-lamp (9.13 mW cm^{-2}) Based on the very good solubility of CTA-peptides and the literature report,[87]

DMSO was used as the solvent for all but the polymerization of the zwitterionic SPE, which was done in TFE.

In general, all monomers tested were consumed quickly, with polymerization reaching nearly quantitative conversion within 3-5 h of irradiation. The reaction progress was monitored using ^{1}H-NMR spectroscopy. Following polymerization, DMSO was removed under reduced pressure, the polymer was dissolved in acetone and precipitated to remove remaining monomer. Precipitation of the liquid polymers based on OEGA was not possible. Instead, dialysis against water for several days was performed to remove small molecules and oligomers. Following solvent removal, the polymers were dissolved in water, lyophilized, and analyzed. Some polymers, particularly those based on NIPAM and those with long peptide sequences, were discovered to be poorly water soluble. This is not a concern for SML at this point, because all functional groups of the peptide are still masked by hydrophobic protecting groups.

Table 4 contains more information about the various polymers obtained via XPI-RAFT polymerization. SEC elugrams of all polymers can be found in Figure 35, appendix.

Table 4: Overview of the various peptide-polymer constructs obtained via XPI-RAFT polymerization.

Monomer	#	Peptide Sequence[a]	Target DP	$M_{n\ (theoret.)}$[b] in g/mol	M_n in g/mol[c] protected	Đ protected	M_n in g/mol[d] deprotected	Đ deprotected
NAM	1[e]	LPETGG	10	1500	-	-	-	-
	2	LPETGG	100	15000	9700^{NMP}	1.60	9200^{NMP}	1.42
					640^{H2O}	1.79	-	-
	3	FLFG-LPETGG-HG	50	8900	8900^{NMP}	1.57	9600^{NMP}	1.41
					-	-	4700^{H2O}	1.21
	4	GG	20	3300	2800^{NMP}	1.32	2600^{NMP}	1.34
					-	-	2400^{THF}	1.16
	5	GG	100	14500	12700^{NMP}	1.26	12200^{NMP}	1.34
					-	-	5600^{H2O}	1.26
DMA	6	LPETGG	50	5900	4600^{THF}	1.13	-	-
					-	-	5800^{NMP}	1.30
	7	LPETGG	200	20700	11000^{THF}	1.26	-	-
					-	-	26000^{NMP}	1.18

					-	-	16 200^{H2O}	1.32
	8	FLFG-LPETGG-HG	200	21 600	34 000^{NMP}	1.18	32 000^{NMP}	1.18
					-	-	4100^{H2O}	1.27
	9	FLFG-LPETGG-HG	1000	100 900	99 000^{NMP}	1.40	96 000^{NMP}	1.42
					-	-	77 000^{H2O}	1.43
	10	GG	20	2400	1500^{THF}	1.29	-	-
					2200^{NMP}	1.15	2000^{NMP}	1.21
NIPAM[f]	11	LPETGG	50	6600	5200^{THF}	1.20	4600^{THF}	1.20
					-	-	7200^{NMP}	1.49
	12	LPETGG	100	12 300	10 000^{THF}	1.32	10 000^{THF}	1.26
	13	LPETGG	200	23 700	12 900^{THF}	1.35	12 500^{THF}	1.37
					-	-	23800^{NMP}	1.35
	14	WTWTW-LPETGG	50	7800	5300^{THF}	1.36	830^{THF}	2.7
					-	-	6500^{NMP}	2.2
	15	WTWTW-LPETGG	200	24 900	17 900^{THF}	1.34	10 600^{THF}	1.49
					-	-	33 800^{NMP}	1.31
	16	FLFG-LPETGG	50	7500	7300^{NMP}	1.95	8900^{NMP}	1.61
	17	GG	50	6200	7000^{NMP}	1.25	7800^{NMP}	1.22
	18	GG	100	11 900	12 500^{NMP}	1.34	12 200^{NMP}	1.26
OEGA	19	LPETGG	50	24 900	20 000^{THF}	1.20	19 700^{THF}	1.17
					-	-	24 000^{NMP}	1.15
					-	-	3400^{H2O}	1.67
	20	LPETGG	200	96 900	48 000^{THF}	1.21	47 100^{THF}	1.19
					-	-	52 800^{NMP}	1.51
					-	-	40 000^{H2O}	1.72
	21	GG	20	10 100	9600^{NMP}	1.20	9200^{NMP}	1.24
					-	-	6800^{H2O}	1.14
(Glu)-HEMA	22	LPETGG	50	23 900	47 000^{THF}	2.02	-	-
					40 000^{NMP}	2.40	-	-
SPE	23	LPETGG	50	28 900	114 200^{HFIP}	2.40	-	-

a: peptide side chains were always protected, GG exhibited a Boc group at its N-terminus. **b**: Theoretical values of M_n based on 100 % monomer conversion, including the MW of the respective peptide end group. **c**: SEC calibrations were performed using standards of PS (NMP and THF), PEG (H_2O) and PMMA (HFIP). **d**: After deprotection of peptide side chains in TFE for 1-2 h. **e**: Sample did not result in usable SEC data because of low MW, analysis via ESI-MS, see below. **f**: No H_2O-SEC data due to polymer precipitation because of thermoresponsive behavior.

The targeted DPs of the polymers were chosen to correspond with literature reports on (partially commercially available) protein-polymer conjugates,[11,144] which mostly used PEG with a MW of 10-20 kg/mol. In addition, longer polymers were synthesized to test the limits of SML due to probably less accessible end groups and a generally high sterical hindrance. Shorter polymers, especially with the nucleophilic end group, were used to facilitate SML, because long chains were expected to be the method's main limitation. Most SML reactions have traditionally used proteins with the recognition sequence peptide on the C-terminus (C terminal protein ligation). This strategy necessitates the use of an excess of the (usually small) nucleophile to shift the equilibrium towards the product. Short polymers that are easy to couple were required to replicate the literature strategy and obtain benchmark data.

For the majority of the entries in Table 4, the MW data obtained by SEC is slightly lower than the expected MW for full monomer conversion including the MW of the peptide end group. As the purity of the CTAs used was not 100 % (see chapter 3.1.4 and 3.2.2), the polymerization was expected to result in slightly higher than expected MWs, leading to a lower amount of intact CTAs and thus longer chains. This emphasizes the importance of viewing MW values obtained by SEC using a calibration standard with a different chemical composition as the analyte as an estimate rather than an absolute value.

When using different solvents for SEC, there are sometimes large differences in M_n (e. g. entries 5 and 8, Table 4). This observation can be attributed to a number of factors. First, the calibration standards are different for THF/NMP and H_2O, resulting in incomparable data. Second, especially for short polymers, the peptide end group is expected to have a significant impact on SEC data. Peptides can form strong hydrogen bonds, resulting in intra- and inter-chain interactions as well as potentially interactions with the column material. These interactions are strongly influenced by the surrounding solvent, resulting in sometimes large differences in MW when different solvents are used. The peptide's side chain protection groups are another factor that influences MW determination. This is reflected in the sometimes significant differences in M_n before and after the removal of those protection groups, resulting in a shift from very hydrophobic (*tert*-butyl) to very hydrophilic (carboxylic acid and alcohol). When comparing entries 14 and 15 (7 side-chain protecting groups) with 17 and 18 (one N-terminal protection group), the differences in MW between protected and deprotected peptide-polymers become much more pronounced.

The measured dispersities are relatively high compared to the excellent values reported by Lehnen and colleagues ($Đ \leq 1.1$ for P(NAM)),[87] but they do not indicate an uncontrolled reaction. This is most likely due to the use of macro-CTAs with limited accessibility in comparison to the unfunctionalized CTAs. Furthermore, peptides exhibit a complex solution behavior, potentially

leading to an incomplete solvation in the reaction medium. Measuring the chain transfer constants of the macro-CTAs would be useful in learning more about this particular finding, but it was beyond the scope of this work.

Nonetheless, the monomodal nature of all samples with the exception of samples 22 and 23 suggests a well controlled polymerization with intact RAFT equilibrium.

SEC data (Figure 36, appendix) reveals that the reaction is not controlled, for the most challenging monomers (Glu)HEMA and SPE (entries 22 and 23), resulting in very broad ((Glu)HEMA) or multimodal (SPE) curves. Based on the data from Lehnen *et al.*,[87] less control over the polymerization of slow-propagating (tertiary) radicals, in this case methacrylates, was expected. This could be due to the solvent (TFE) in the case of SPE: The initial report on XPI-RAFT revealed that conversion was highly dependent on the solvent used. TFE has not yet been tested for XPI-RAFT, but reports in the literature show successful (thermally initiated) RAFT polymerization of SPE in TFE.[145]

The reason for the poor control is most likely not the bulky side chain of (Glu)HEMA or SPE since the polymerization of OEGA proceeded with good control. The tertiary nature of the propagating radicals in combination with XAN and BABTC, both of which exhibit low chain transfer constants for methacrylates, is much more likely to be the cause. This hypothesis is supported further by the observation of low MW tailing in the SEC elugram, indicating that XAN reactivation occurs via photolysis, resulting in dead chains.[87] A different combination of CTAs may be worth trying in future experiments to improve reaction control using monomers forming tertiary radicals.

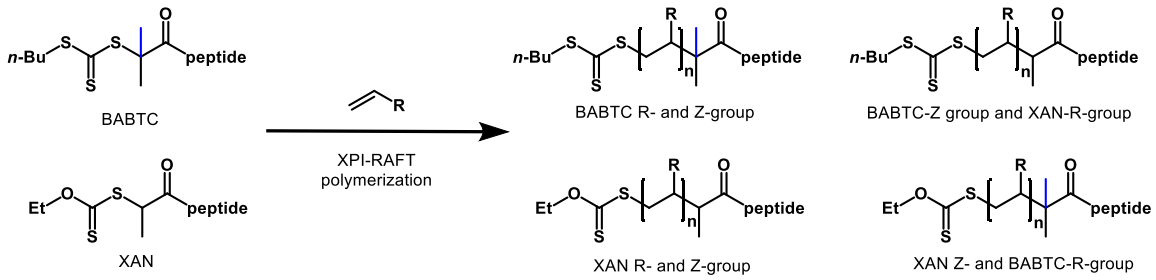
Because SEC yielded inconclusive results, MALDI-ToF was used to try and gain more information about the peptide-polymers. Unfortunately, no useful information was obtained for any of the polymers tested. This might be due to the dispersity being too high to obtain good spectra. It has been reported that MALDI-ToF becomes difficult once a polymer's dispersity exceeds $Đ = 1.2$.[146] Due to many overlapping signals, analysis of the MW using ^{1}H-NMR spectroscopy proved to be very inaccurate, especially when using long peptide end groups.

Different - preferably absolute - analytical methods would be required to accurately determine the MW of the polymers. Because the overall shape of the SEC traces is monomodal, there should be no unwanted interaction of the peptide-polymers with the column material. Thus, SEC combined with a viscosimetry detector and/or a multi-angle light scattering (MALS) detector may be useful for accurately determining the MW and learning more about the sample's solution behavior. Analysis of MW by viscosimetry is especially useful for samples that do not have a chemically equivalent narrowly dispersed standard, as is the case here. The Universal Calibration used for viscosimetry is based solely on the hydrodynamic radius, decoupling the MW-calibration from the

chemical composition of the calibration standard. Furthermore, the Mark-Houwink plot of viscosimetry provides information about the solution behavior of the peptide-polymer.[147] Additional MALS detection can be used to gather more information about the solution's behavior. Light scattering data, on the other hand, is very sensitive to system impurities (dust particles), yields limited information about block copolymers, and works best for large MW polymers (radius of gyration 10 nm, about 200 000 g/mol polystyrene in a good solvent).[147] With this in mind, additional analysis of the obtained peptide-polymers could be performed using SEC equipped with a viscosity detector. No further characterization was performed due to time constraints and because the ESI-MS analysis demonstrated successful end group retention (see below).

Because of overlapping signals, ^{1}H-NMR spectroscopy to analyze end group fidelity and MW proved difficult and inaccurate. Thus, ESI-MS was used to confirm end group retention after polymerization. Because of the relatively low mass cutoff (m/z 3000) of the quadrupole detector in the MS instrument, combined with the characteristic of electrospray ionization to produce mostly ions with a single charge,[111] this was only possible for very short polymers (entries 1, 4 and 10 in Table 4). Nonetheless, the information gathered was useful.

There are four possible end group combinations, as shown in Scheme 13. Assuming efficient chain transfer, it was expected to find a low number of XAN-Z and BABTC-R-group species (due to the 2:8 molar ratio of XAN:BABTC) and very few XAN-R and Z-group species (the most unlikely combination).



Scheme 13: Possible polymer end group combinations using BABTC and Xan within a XPI-RAFT process. The methyl group of the BABTC-R-group is highlighted because it marks the only difference between the R-groups of XAN and BABTC.

The major peaks in the case of $PNAM_{20}$-LPETGG, correspond very well to BABTC end groups and their respective aggregates with Na^+ and K^+ (Figure 13). As expected, the peaks corresponding to XAN end groups in combination with a BABTC-R-group are not always detectable and are very small in all cases (not highlighted in Figure 13). The most unlikely combination of XAN R- and Z-groups was not found in the spectrum. Interestingly, peaks separated by the mass of the NAM

repeat unit (marked in Figure 13 at m/z 2444.01 and 2472.67) could not be assigned to an expected end group or its common aggregates. This suggests the presence of decomposed CTA at the chain end ("dead chains"), although the peaks do not correspond to the expected values of a thiol end group, which would be the most likely result of CTA decomposition. This finding contradicts the report from A.-C. Lehnen *et al.*,[87] who found no dead chain ends using the same reagents, but without using peptide-modified CTA.

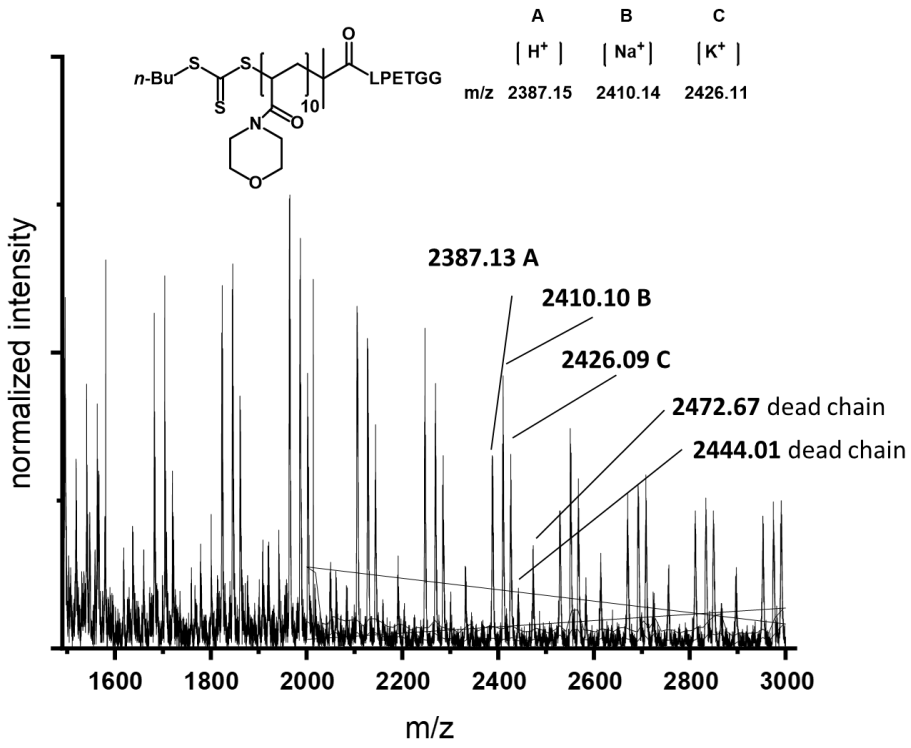


Figure 13: Mass spectrum of P(NAM)$_{10}$-LPETGG (Table 4, entry 1, peptide side chains protected) for end group analysis. Values depicted and assigned to theoretical values exemplary for DP 10.

Because the peptide is attached via the R-group, some dead chains after polymerization are not a major concern in this case. However, when it comes to more complex polymers (for example, block-copolymers), a large number of active chain ends is critical.

The ESI-MS of the short nucleophilic GG-polymer (Figure 14) is very similar to its counterpart. Signals were assigned to peptide-BABTC end groups and the Na^+ as well as the K^+ adducts once more. Furthermore, peaks assigned to XAN-Z group and the BABTC-R-group (methyl group highlighted in blue) were visible. In contrast to the polymer in Figure 13, no large signals were found that could not be assigned to an expected end group, implying that the polymerization of the oligoglycine macro-CTAs results in fewer dead chain ends.

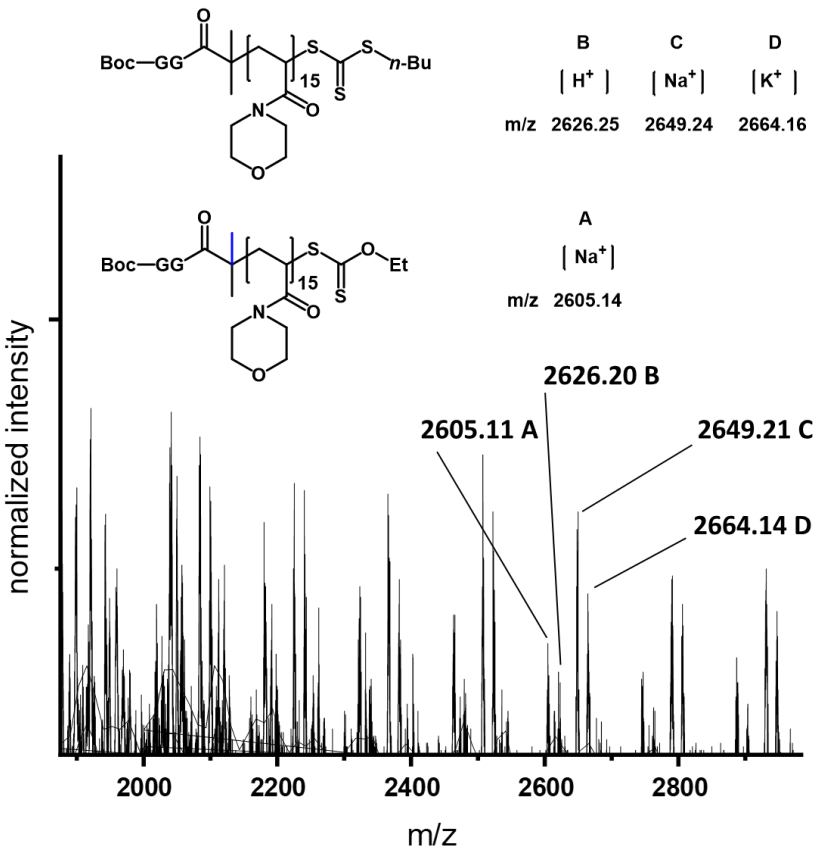


Figure 14: Mass spectrum of Boc-GG-P(NAM)$_{20}$ (Table 4, entry 4) for end group analysis. Values depicted and assigned to theoretical values exemplary for DP 15.

In summary, the results of RAFT polymerization of macro-CTAs containing both nucleophilic and recognition sequence peptides in solution are presented in the two preceding subchapters. The promising PET-RAFT approach did not result in a well-controlled reaction. The photolysis under blue light of the (tertiary) CTA used here could explain this finding. Using a different (secondary) CTA or adding a tertiary amine to oxidize the formed CTA radical into a less reactive anion could potentially optimize the reaction.

Finally, XPI-RAFT, a method developed in part parallel to this work, was used to generate a library of different polymers with all the peptidic end groups required for SML. The XPI-method was simple to use and produced polymers with acceptable dispersities. SEC as an analytical method produced mixed results, most likely due to peptidic end group. End group analysis by ESI-MS revealed that short polymers had very good end group fidelity after polymerization, with very few (LPETGG) or no (GG) detectable dead chain ends.

Overall, the polymer with the recognition sequence end group is much easier to synthesize than its nucleophilic counterpart. The main reason for this is that CTA coupling is possible on resin, which

greatly reduces the purification effort prior to polymerization. Second, a change in the peptide sequence is much easier, because changing the nucleophilic sequence (for example, to incorporate a flexible spacer) would require a change in the CTA-peptide purification procedure (either a change of solvents for column chromatography or the use of preparative HPLC, including method development). Furthermore, because EDA is highly hygroscopic, the nucleophilic peptide-EDA compound must be handled quickly to avoid excessive water uptake.
The final step before SML was deprotection of the peptidic end groups, which is covered in the following chapter.

4.3 Polymer Deprotection

The side chain protection groups (for LPETGG-containing peptides) and the N-terminal Boc-protection group (for GG) had to be cleaved off in order to activate the peptide-polymer for SML. All acid-labile protection groups were cleaved by dissolving the polymer in pure TFA for up to 3 h. Because all of the cleavage products were volatile, precipitation was unnecessary, and the polymer was lyophilized directly after TFA removal. Attempts to lyophilize from aqueous solution were unsuccessful for most polymers except P(OEGA) because the polymers formed a dispersion rather than completely dissolving. This is surprising because all of the components (TFA and polymers) should be water soluble. Instead, a 50:50 v:v MeCN:benzene mixture was required. To prevent thawing of the mixture during lyophilization, benzene was chosen because it has a higher melting point than many other organic solvents, and MeCN was used to dissolve the polymer. The Kaiser test was used to ensure that GG-polymers have a free amine end group, which is required for SML, and it yielded positive results in all cases. Following that, the polymers were analyzed using ^{1}H-NMR spectroscopy and used for SML without further purification. ^{1}H-NMR spectroscopy showed the complete removal of all protecting groups (see Figure 39, appendix).
As previously stated in chapter 3.1.4, the CTA was expected to decompose at least partially after treatment with TFA. As a result, the short polymers were analyzed again with ESI-MS. It is unimportant whether the CTA is intact or not for the use in SML, but it may be important in future experiments when determining the construct's biocompatibility.

Figure 15 (P(NAM)-LPETGG) and Figure 16 (GG-P(NAM)) depict the results of the same polymer samples as discussed before in chapter 4.2.2.

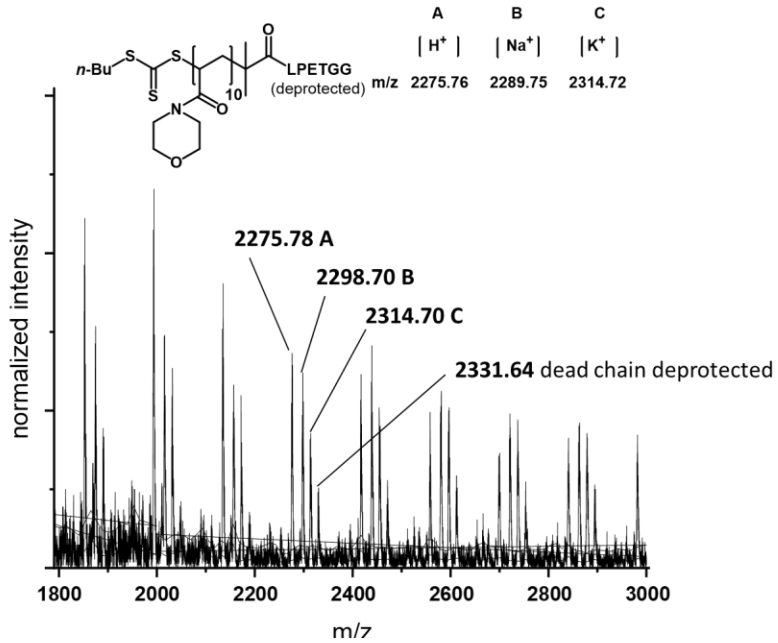


Figure 15: ESI-MS spectrum of short P(NAM) with LPETGG (peptide side chains deprotected) end groups. Values depicted and assigned to theoretical values exemplary for DP 10.

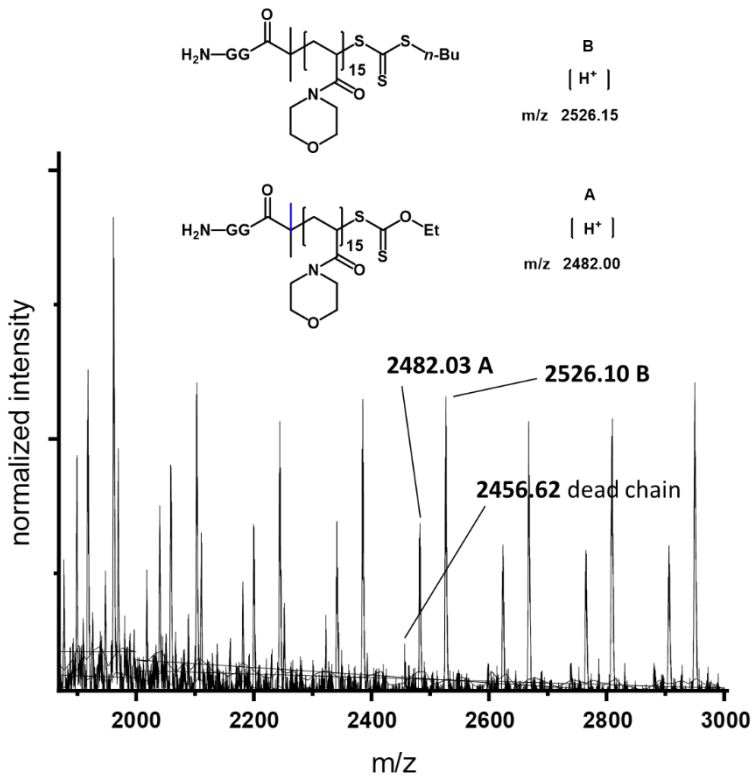


Figure 16: ESI-MS spectrum of short P(NAM) with deprotected GG end groups. Values depicted and assigned to theoretical values exemplary for DP 15.

For both end groups, the m/z values correspond well to the expected values of the deprotected peptides. All expected combinations of R- and Z-groups were found except XAN R- and Z-group (Scheme 13, for clarity, the less-intense peaks were not marked in the ESI-MS spectra). Peaks corresponding to dead chains were found for the P(NAM)-LPETGG polymer, but this does not indicate CTA decomposition upon TFA treatment because these peaks were found in similar intensities in the protected polymer spectrum (see Figure 13).

For this analysis, is was assumed that recurring peaks in the mass difference of the NAM repeat unit that could not be assigned to a commonly observed aggregate (such as $[M+Na]^+$ or $[M+K]^+$) corresponded to dead chain ends with a fragment of the decomposed CTA. Moreover, only one of the dead chain end types found in the protected polymer was detected here.

Dead chain ends were detected in the GG-PNAM sample (marked once in Figure 16 as m/z 2456.62), which do not correspond to any end group combination in the protected polymer.

This suggests that end group decomposition occurs during deprotection. According to the results from the ^{1}H-NMR spectroscopy analysis in chapter 3.1.4, decompositon is expected, but was not detectable in in P(NAM)-LPETGG (Figure 15). Surprisingly, the m/z values of polymers with either peptidic end group do not correspond to the expected thiol decomposition products, indicating that a different (unidentified) decomposition product was formed.

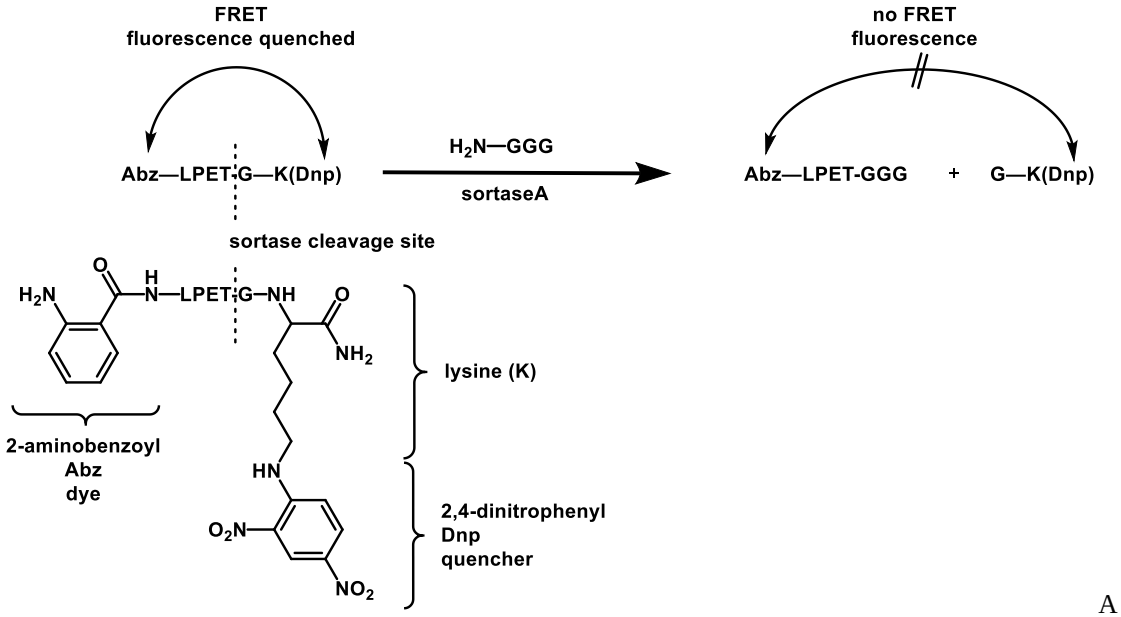
SEC was used to analyze the deprotected polymers again, yielding mostly similar molecular weights and dispersities as the protected polymers. The difference between pre- and post-deprotection was much more pronounced for polymers with WTWTW-LPETGG end groups (entries 14 and 15 in Table 4). This was attributed to a change in hydrodynamic radius caused by a much larger change of the end group characteristics due to the fact that threonine (T) and tryptophan (W) carried protection groups (5 more than in LPETGG). Table 4 contains all SEC data of the deprotected polymers used for SML in comparison to the protected polymers. Again, MALDI-ToF MS analysis produced no useful spectra. The reasons for this problem, as well as potential solutions, have been discussed in chapter 4.2.2.

5 SORTASE-MEDIATED REACTIONS

Out of the many different variations of sortase A known to the scientific community,[49] the wild-type (SrtA-WT) and an engineered version based on the very popular "5M" variant (SrtA-5M)[47] were used for SML initially. All sortases used within this work were obtained from the group of Prof. Schwaneberg (RWTH Aachen, SrtA-WT) and Prof. Möller (University of Potsdam, SrtA-WT and SrtA-4M). During the experiments, the enhanced sortase was changed from SrtA-5M to a 4M variant with one less mutation.[148] According to P. Zou (Prof. Sattler, Helmholtz Zentrum München) who generously provided the plasmids for SrtA-4M, the 5th mutation, located far away from the catalytic site, was thus deemed unnecessary. Chapter 7.10 contains information on sortase amino acid sequences, expression, and purification. All experiments with SrtA were conducted in SML buffer (50 mM tris(hydroxymethyl)aminomethane (TRIS), 150 mM NaCl, 5 mM $CaCl_2$, pH 7.5).[102]

5.1 PEPTIDE ASSAYS

SML was initially performed only with peptides without attached polymers to test both sortase activity and the suitability of the synthesized peptide sequences. A very convenient fluorescence-based assay to determine and quantify the sortase activity was developed by Schneewind and coworkers (Scheme 14).[149]



A

Scheme 14: Reaction scheme using Abz-LPETG-K(Dnp) as FRET-based fluorescence quenched substrate for sortase assays.

modified peptide with a fluorescent dye (2,4-dinitrophenol, Dnp) at the C-terminus and a quencher (2-aminobenzoyl-, Abz) at the N-terminus of an LPETGK sequence is used in the assay. Because of Förster resonance energy transfer (FRET) between dye and quencher, this peptide is not fluorescent. Once SrtA cleaves the peptide between T and G, separating Abz and Dnp, the sample begins to fluoresce because efficient FRET is only possible over very short distances. The intensity of the fluorescence signal corresponds to the amount of cleaved peptide.

This assay is popular among researchers working with sortase to quickly gain information about the enzyme activity. Kruger *et al.* discovered later, however, that the system exhibits a substrate-induced inner filter effect that quenches fluorescence, rendering the kinetic data from these assays incomparable.[150]

A detailed quantitative analysis of reaction kinetics in SML with various substrates has previously been conducted.[43,50,150,151] Therefore, the emphasis of the assays performed here was on determining a yes/no answer on sortase activity, comparing different sortase variants, and comparing SML-performance of different substrates. As a well-known, commercially available benchmark substrate the Abz-LPETG-K(Dnp) sequence was used. It was not used to measure fluorescence because this only monitors the cleavage of the peptide. Therefore, no information can be gained on resolving the enzyme-substrate intermediate and, consequently, side-reactions (particularly hydrolysis, see chapter 1.3) cannot be detected by fluorescence. Instead, HPLC analysis of the reaction mixture was performed, based on a protocol developed by Aulabaugh *et al.*[43] In short, the substrate peptides were dissolved in sortase reaction buffer to prepare stock solutions, aliquots were mixed in the targeted molar ratios, sortase was added to start the reaction, aliquots were taken, sortase was quenched by denaturation with aqueous HCl and the samples were analyzed.

Figure 17 depicts an HPLC elugram obtained from a typical benchmark experiment to test SrtA activity. Information about the HPLC methods is not mentioned here and can be found in chapter 7.2. Unless otherwise stated, all HPLC elugrams shown, were obtained using the same method parameters.

In order to shift the reaction equilibrium and maximize product formation, 100 molar equivalents of nucleophile (GGG) were used in relation to recognition sequence peptide (Abz-LPETG-K(Dnp)) in benchmark experiments. Unless otherwise stated, all sortase assays were performed with 4 mol-% SrtA based on the recognition sequence peptide to maintain comparability.

The elugrams show the expected peaks for the different compounds. The reaction progressed to quantitative conversion, shown by no residual educt after 2 h of reaction time. The product was already formed after “0 h” reaction time, highlighting the high reaction speed and the sensitivity of

the analytical method. The “0 h” sample was generally taken and quenched 10-30 s after sortase addition.

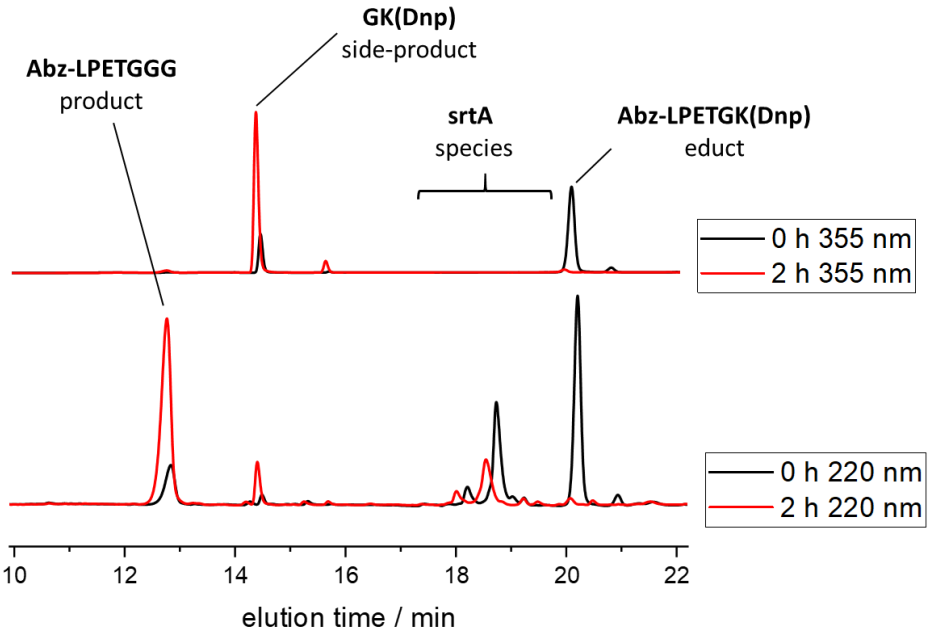


Figure 17: RP-HPLC elugrams showing a typical benchmark test using Abz-LPETGK(Dnp) (0.05 mM, 1 eq) as substrate and GGG (5 mM, 100 eq) as nucleophile. 4 mol-% SrtA-WT was used. GGG results in a very high peak due to the excess used and is not shown for reasons of clarity.

When measurements at 355 nm are considered, an advantage of the FRET-active substrate peptide is revealed. The quencher (Dnp) absorbs light and can be detected at this wavelength. Because the product no longer contains Dnp and thus does not absorb light at 355 nm, peak identification is facilitated without the need for an MS-detector. Nonetheless, confirmation of the correct peak identification was achieved after collection of the relevant HPLC fractions and analysis via ESI-MS (Figure 40, appendix).

SrtA-WT exhibits multiple peaks that change during the reaction, as expected. According to Aulabaugh and colleagues, those peaks represent the various species of sortase present during the reaction.[43]

Figure 18 shows the HPLC elugrams with LPETGG and GGG as reactants. These peptides represent the end groups used for SML of proteins and polymers later on. Neither LPETGG synthesized via SPPS nor commercially acquired LPETGG showed any clear signs of product formation. This implies that when LPETGG is used as a substrate, either no product is formed or the reaction occurs but the product is not detectable by HPLC, potentially because the difference to the educt is only one G.

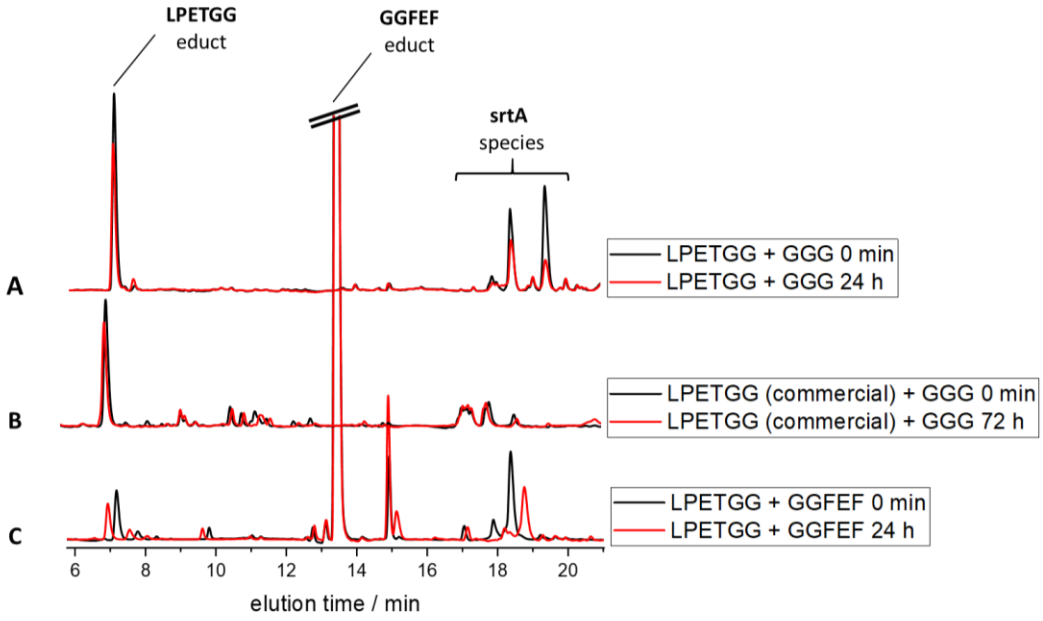


Figure 18: RP-HPLC elugrams showing SML using LPETGG as recognition sequence peptide. **A**: LPETGG synthesized via SPPS. **B**: Commercially aquired LPETGG. **C**: LPETGG synthesized via SPPS with alternative GGFEF nucleophile.
Ratios for all reactions: LPETGG (0.05 mM, 1 eq), GGG/GGFEF (5 mM, 100 eq) and 4 mol-% SrtA-WT was used. GGG results in a very high peak due to the excess used and is not shown for clarity.

However, when LPETGG was combined with GGFEF as nucleophile, still no reaction product was observed despite a much greater difference in polarity between product and educt and thus, an expected easy separation via HPLC. Because the combination of Abz-LPETG-K(Dnp) and GGFEF produced a clearly separated product peak (data not shown), it was concluded that no reaction occurs when using LPETGG as substrate.

As a result, a new peptide containing the recognition sequence needed to be designed and synthesized. All other assays performed using different peptides (see below) indicated that the amino acid sequence besides LPETG is not important for successful SML. Thus, a new peptide, FLFG-LPETGG-HG was designed and synthesized using SPPS. Aromatic phenylalanine (F), absorbing light at 260 nm, was added to facilitate HPLC peak detection. The selection of the other amino acids is explained in chapter 5.1.

As shown in Figure 19, when combined with one molar equivalent of nucleophile (GGFEF), the new recognition sequence peptide produced 50 % product (compound confirmed by ESI-MS) as expected. The equilibrium-state was quickly reached after 30 min of reaction time, as evidenced by no change in peak integral after this timepoint.

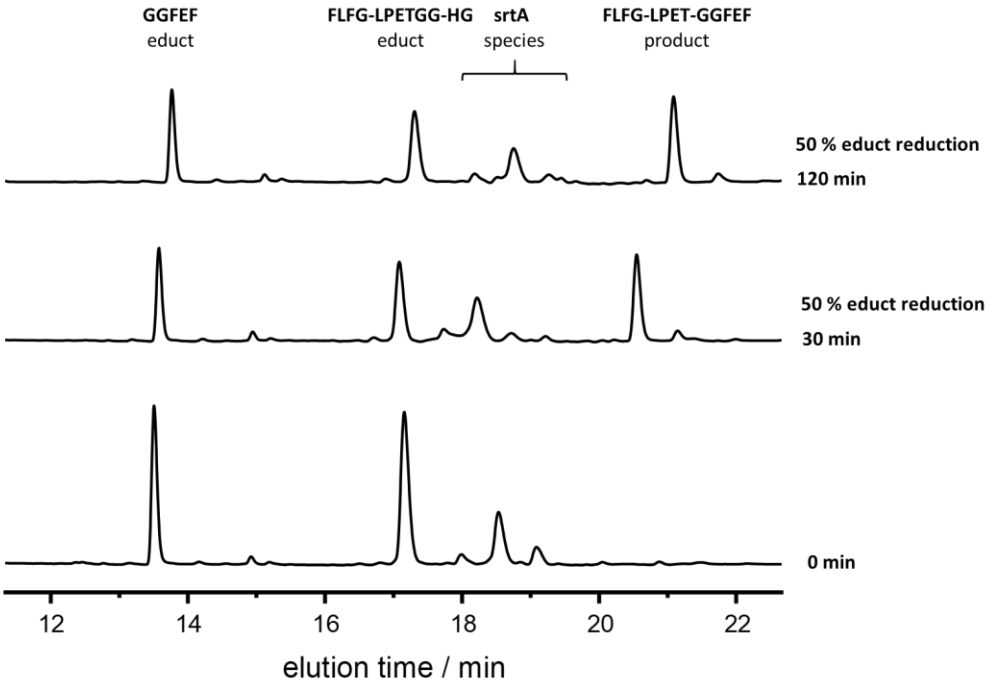


Figure 19: RP-HPLC elugrams of SML using FLFG-LPETGG-HG (1 eq), GGFEF (1 eq) and 4 mol-% SrtA-WT at different reaction times. Reaction yield calculated by comparison of peak integrals. Absorption detection at 220 nm.

On a side note, using FLF-LPETGG-HG as the recognition sequence peptide takes much longer to reach equilibrium than using Abz-LPETGG-K(Dnp) (10 min, same reagent ratios, data not shown). This indicates that longer, more complex peptides containing the recognition sequence require longer reaction times, and it is expected that the required reaction times will increase even more when peptide-polymers are used. The scientific literature contains evidence to support this hypothesis: reaction times ranging from minutes to a few hours have been widely used for SML of short peptides[47,152,153] or the modification of a protein with a small molecule.[102,154–156] Synthesis of fusion proteins[51,157] or protein-polymer conjugates,[26,33,89] on the other hand, required much longer reaction times of up to 24 h.

It should be noted that a ~0.5 min shift of the product peak comparing 30 min and 120 min was found repeatedly in my HPLC elugrams. Both peaks were analyzed via ESI-MS and showed the same m/z values corresponding to the expected product. Small column temperature fluctuations

between measurements could explain the peak shift, resulting in slightly different elution times for certain compounds.

These findings strongly support the hypothesis that a free leucin N-terminus in LPETG is not recognized as a substrate by SrtA. The N-terminus of leucine must be functionalized before it can be used by SrtA. This functionalization is not limited to the addition of (natural) amino acids. Similar to the assay substrate Abz-LPETG-K(Dnp), my own experiments using an ATRP initator on the leucine N-terminus were also leading to product formation (data not shown).

To the best of my knowledge, this has not been reported before, most likely because most research is focused on functional peptides and thus the minimal recognition sequence is rarely used.

5.1.1 SORTASE VARIANTS

The differences between the "native" SrtA-WT and the engineered variant SrtA-4M were examined. A 140-fold increase in reaction speed is claimed for SrtA-5M,[47] and a similar improvement is expected for SrtA-4M according to the supplier of the SrtA-4M plasmids.

The assays were carried out with the benchmark peptide Abz-LPETGG-K(Dnp) with triglycine (100 eq) as nucleophile and 4 mol-% of freshly thawed batches of the respective sortase variant. The HPLC analysis reveiled no significant differences. This is most likely due to the sample intervals. The first sample was taken after 10 min reaction time and showed already full Abz-LPETGG-K(Dnp) conversion for both sortase variants. Even in the reference sample taken directly after sortase addition (quenched after approx. 15-30 s), significant amounts of product were detected (data shown in Figure 42, appendix). The accuracy of this reaction kinetics analysis might have been improved by taking more samples within a few minutes after reaction start. However, due to the method's relatively high margin of error (as demonstrated by the first reaction sample), this was inaccurate.

A much larger difference between the two sortase variants was seen when both were tested for side-reactions. The intermediate LPET-SrtA complex is known to become hydrolysed over long reaction times if no good nucleophile is present (introduced in chapter 1.3). Water can act as a nucleophile in this case, releasing the recognition sequence peptide from sortase.

The analysis of possible side reactions is especially important here because both recognition-sequence peptide and nucleophile are linked to large molecules, rendering the end groups somewhat inaccessible compared to the surrounding water.

Hydrolysis tests with the peptide sequence FLFG-LPETGG-HG resulted in a relatively quick consumption of the educt, while a new peak was emerging (see Figure 20). Unfortunately, due to its overlap with the peaks corresponding to sortase and the small amount of sample, this peak could not be analyzed via ESI-MS.

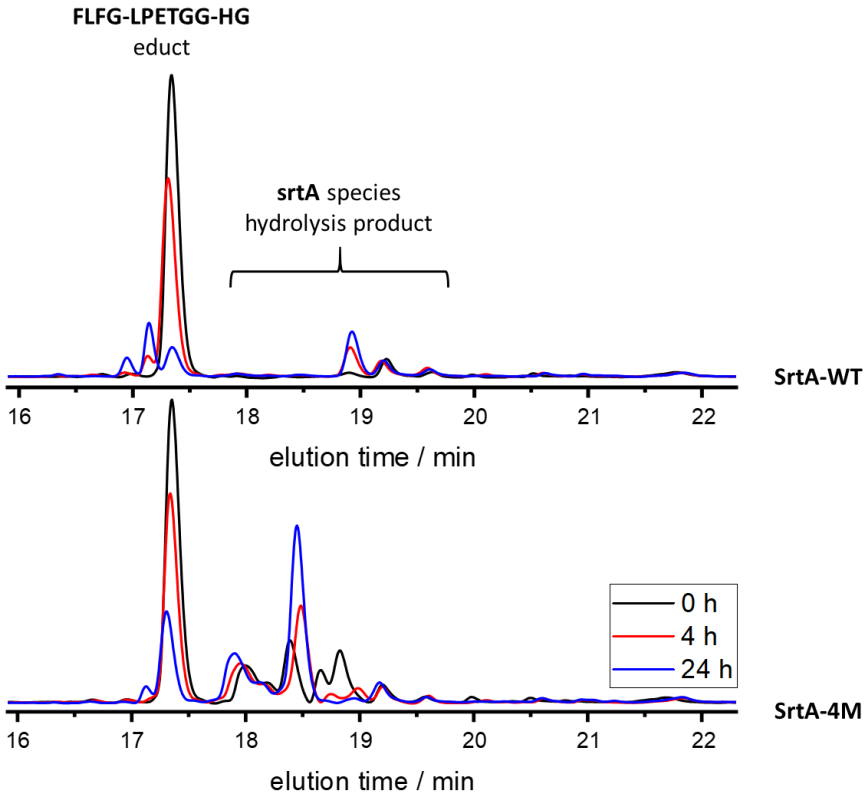


Figure 20: RP-HPLC elugrams of experiments using Srt-WT or Srt-4M (4 mol-%) and FLFG-LPETGG-HG without a nucleophile to force side reactions. Absorption signals measured at 220 nm.

Nonetheless, the sample taken after 24 h reaction time clearly shows that SrtA-WT consumes the educt faster than the mutated SrtA-4M. As no good nucleophile is present, hydrolysis of the SrtA-peptide intermediate can be expected. Based on the assumption that hydrolysis is the primary reaction in the performed assay, one would expect that the much faster SrtA-4M results in quicker educt hydrolysis. However, the hydrolysis seems to be more pronounced in my experiments with WT than with 4M. This finding contradicts literature reports, which found that long reaction times using SrtA-5M lead to higher amounts of hydrolyzed side-product.[42] The reason for this contradiction remains unknown, although this behavior was reported for SrtA-5M and not the SrtA-4M variant used here. However, it seems unlikely that his change in behavior is caused by a single AA (the only difference between SrtA-4M and SrtA-5M).

Based on the much quicker educt consumption (probably hydrolysis) of SrtA-WT with no good nucleophile present, Srt-4M was chosen as catalyst for the reactions with polymers and proteins.

Another critical aspect is the degradation of sortase when stored in a fridge-cooled buffer solution. Some protocols mention that SrtA can be stored for several weeks at 4°C.[102] It was found here that SrtA-WT resulted in much lower reaction speeds after 3 weeks of continuous storage at 4°C when compared to a freshly thawed aliquot from the same sortase batch stored at -80 °C. When Abz-LPETG-K(Dnp) and GGG assays were performed (100 eq of nucleophile, 4 mol-% SrtA-WT, benchmark test conditions), the educt conversion reached its maximum after 4 h (66 %). Using the same reaction parameters, freshly thawed SrtA-WT resulted in quantitative educt conversion within 10 min. (data shown in Figure 41, appendix).
It is therefore recommended that either always freshly thawed SrtA is used, or the performance of aged batches is checked before applying for SML with valuable substrates.

5.1.2 Equilibrium Strategies

The final set of sortase assays was performed to determine the efficacy of the different methods for shifting the sortase reaction equilibrium toward the products. All the strategies used within this work have been introduced in chapter 1.3.

β-Hairpin Formation

First, the method developed by Yamamura *et al.* was tested.[51] It is based on the formation of a rigid β-hairpin structure after product formation that does not fit into the active center of SrtA and thus can no longer be cleaved. Based on the best performing peptide sequence (WTWTW, "tryptophan-zipper"),[51] both required peptides WTWTW-LPETGG and GG-WTWTW were synthesized (see chapters 3.1.2 and 3.2.1) and tested in a sortase assay with 4 mol-% SrtA-WT. Furthermore, to rule out any potential issues with peptides synthesized in-house, a commercial sample of both peptides was acquired and tested.
After the initial use of a 1:1 molar ratio of recognition sequence peptide and nucleophile did not result in complete educt consumption, a small excess of nucleophile (5 eq) was tried. As shown in Figure 21 (in-house peptides) and Figure 22 (commercial peptides), all reactions resulted in a decrease in educts peaks and the formation of a product peak. However, neither the reaction time nor the quantitative educt consumption were observed to be as fast as the benchmark.

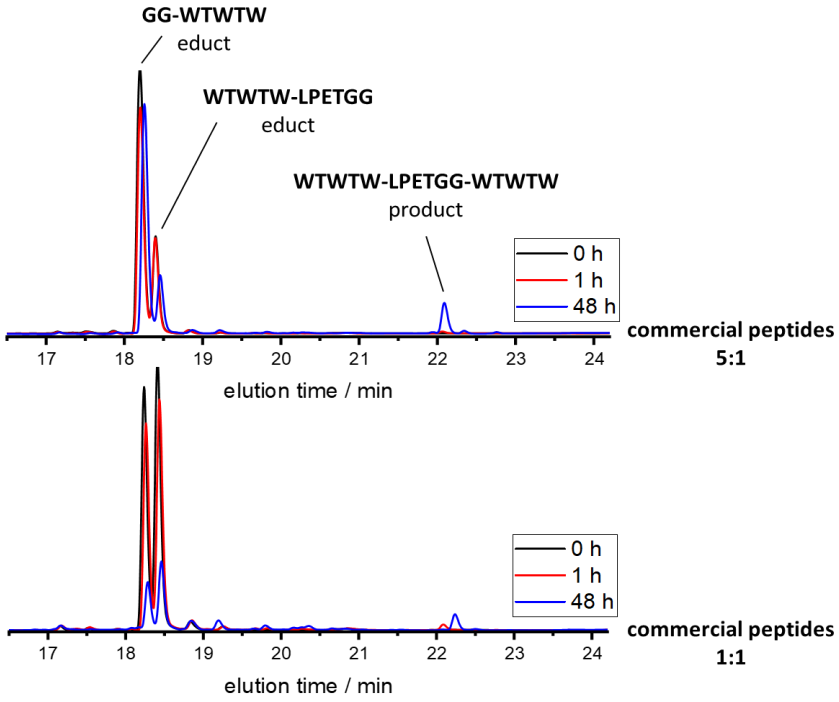


Figure 21: RP-HPLC elugrams of sortase assay to test hairpin-forming peptide sequences with synthesized peptides. GG-WTWTW (0.05 mM or 0.2 mM), WTWTW-LPETGG (0.05 mM), SrtA-WT (4 mol-%). Absorption signals measured at 220 nm.

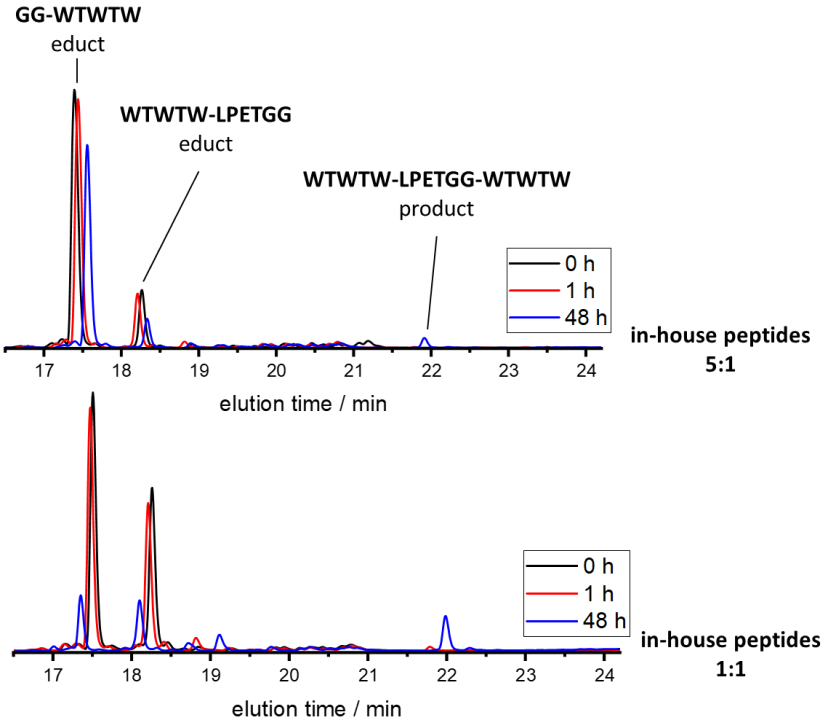


Figure 22: HPLC elugrams of sortase assay to test hairpin-forming peptide sequences with commercial peptides. GG-WTWTW (0.05 mM or 0.2 mM), WTWTW-LPETGG (0.05 mM), SrtA-WT (4 mol-%). Absorption signals measured at 220 nm.

This decrease in reaction speed is to be expected: Baer *et al.* observed a distinct change in reaction speed and overall SML-performance when comparing different peptides and proteins carrying the recognition peptide sequence.[44]
The intensity of the product peak was surprisingly low when compared to the reduction of the educt peaks. One possible explanation is that the β-hairpin peptides are very hydrophobic, making it difficult to dissolve the starting materials in high concentrations in the reaction buffer. Because the product peptide is expected to be even more hydrophobic, it may have partially precipitated from the reaction solution and could be thus undetectable by RP-HPLC analysis.
Following the initial period, educt conversion plateaued and educt levels remained nearly constant for up to 48 h. This implies that the product does not inhibit the attack of sortase, resulting in the formation of an equilibrium reaction. Educt hydrolysis is presumably not a problem, because it is expected to lead to a relatively quick educt consumption as seen before (almost quantitative educt consumption within 24 h, Figure 20). Apart from a slightly cleaner HPLC elugram due to the higher purity, there is no discernible difference between in-house and commercial peptides.

This finding is in contrast with the published report,[51] despite the fact that the peptide sequence used here contains only the functional sequence for the β-hairpin and the literature report lacks sortase assays using short peptides. The use of short peptides could explain why no increased product formation was observed. These short peptides might form aggregates in solution before SML, preventing sortase from accessing the recognition sequence. When using large molecules with the hairpin sequence at the end, this effect is very unlikely because the chain end is most likely sterically shielded from too much aggregation with other educt chains. Furthermore, educt aggregation provides an additional explanation for the observed low intensity of the product peak.

In conclusion, the assay of short peptides showed no benefitial effect on product formation. While the β-hairpin formation and suppression of the back-reaction may be transferable to the proteins and polymers used here, the synthesis of complete polymer and protein building blocks with the β-hairpin sequence would require a significant amount of time and resources, which could in part only be performed by my cooperation partners, with the outcome being questionable. As a consequence, this method to overcome the SML equilibrium was not used in subsequent experiments.

Metal-Assisted SML

The next strategy tested was to use Ni^{2+}-ions to remove the amino acids cleaved from the recognition sequence peptide via complexation, thus suppressing the back reaction.[28,52] The required extended recognition sequence (FLF-LPETGG-HG, see chapter 3.1.2) was synthesized and paired with one molar equivalent GGFEF as nucleophile, SrtA-4M, and $NiSO_4$.

The reaction proceeded without equilibrium formation, resulting in complete educt consumption within 60 min, according to RP-HPLC analysis (Figure 23). The time needed to complete the reaction is much longer than the benchmark peptide, which could be due to the different (longer and sterically more demanding) peptides, but is most likely due to the much lower amount of nucleophile (1 eq vs. 100 eq in benchmark assay).

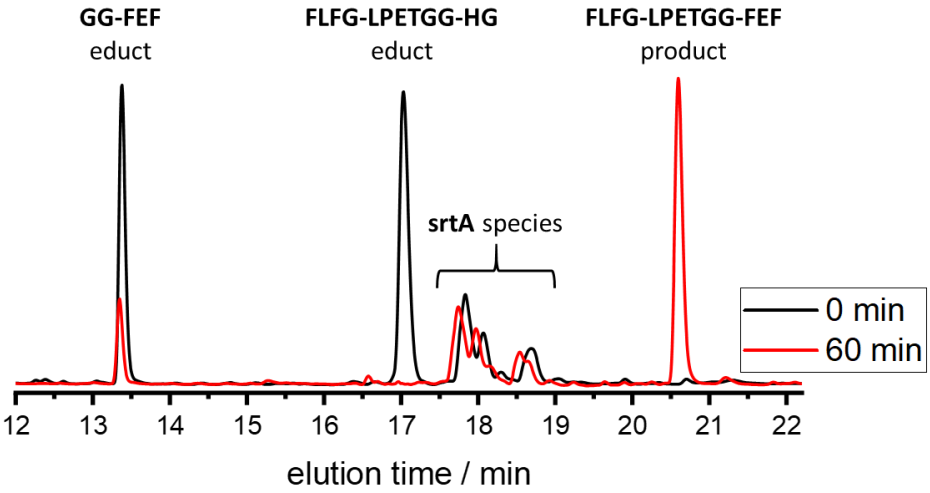


Figure 23: HPLC elugram of an assay testing the effectiveness of preventing equilibrium formation using FLFG-LPETGG-HG (0.05 mM), GGFEF (0.05 mM), $NiSO_4$ (0.2 mM) and 4 mol-% SrtA-4M. Absorption signals measured at 220 nm.

The control reaction which used the same parameters but no $NiSO_4$ resulted in the expected equilibrium formation, maximizing educt consumption at 50 % (see Figure 19).

The hydrolysis product was not detected, presumably the presence of an equimolar amount of (good) nucleophile is enough to prevent hydrolysis. Interestingly, GGFEF educt was not completely consumed, indicating that educts were not used in exactly equimolar amounts. This can be explained with a weighing error, made possible by the small amounts of peptide required for this assay. Due to the small amount of nucleophile left, and the significant difference to the control reaction, the interpretation of this result was assumed to be valid nonetheless.

This result is consistent with the literature reports, making this strategy very promising for efficient SML reactions. Another significant benefit of this method is that proteins are often expressed with

a purification handle (typically hexahistidine), which has been tested and proven to work well with SML and the Ni-salt strategy.[28] Therefore, C-terminal protein modification does not require any additional changes of the educt protein besides the recognition sequence for SrtA.

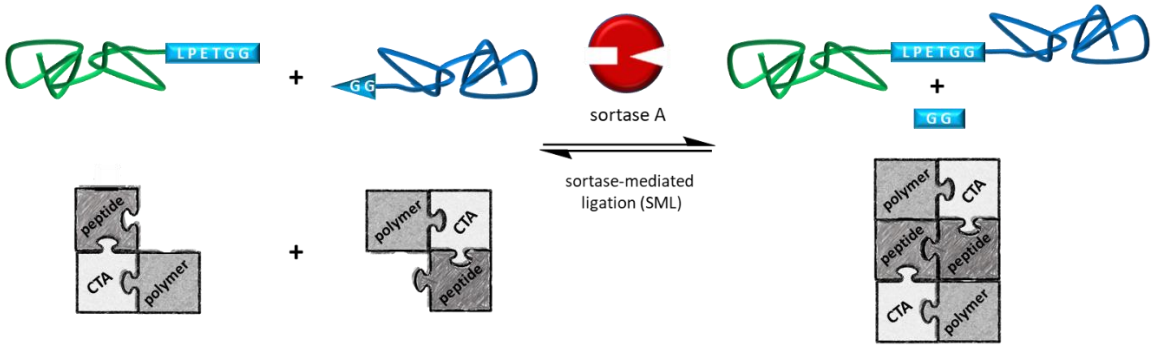
In summary, this chapter describes the development of an HPLC-based assay to evaluate SrtA performance. Several short peptide sequences were tested for SML effectiveness and ability to avoid equilibrium formation when equimolar amounts of recognition sequence peptide and nucleophile were used. It was found that a free N-terminus of the essential recognition sequence (LPETGG) did not result in successful ligation, which had not previously been reported. Thus, the shortest possible sequence that can be used in SML must have a functionalized N-terminus. This functionalization can be (proposedly) any amino acid or another organic molecule.
The hydrolysis side reaction analysis revealed differences between the two sortase variants tested: When no good nucleophile is present, the engineered SrtA-4M proved to be superior, with a slower educt consumption than SrtA-WT. This is surprising because the faster SrtA-4M was expected to hydrolyze the educt much quicker. SrtA deteriorated significantly during storage, with a significant drop in performance after several weeks at 4 °C.
The use of functional peptides to prevent equilibrium formation had no measurable effect when using a β-hairping-forming peptide, but produced excellent results when a histidine-containing peptide was combined with nickel-salt.
The final ligation experiments connecting protein and polymer (chapter 5.3) or two polymers (chapter 5.2) were carried out based on the results obtained in this chapter.

5.2 Polymer-Polymer Coupling via SML

Attempts were made to couple two synthethic polymers containing the respective nucleophilic and recognition sequence end groups as a test of SML using large molecules. These experiments were significant for two reasons: first, the MW and thus steric demand were much easier to vary than with proteins, and second, this approach may make previously difficult-to-achieve block-copolymers accessible.

To the best of my knowledge, the proof of concept shown by our group, specifically X. Dai *et al.*,[116] is the only successful example of SML connecting two purely synthethic polymers. Their approach involves post-polymerization modification of commercially available polymers bearing specific end groups with separately synthesized peptides and using a 50-fold excess of one polymer to increase the yield. There are several other reports on SML of purely artificial polymeric building blocks, but they all are focused on crosslinking.[158–161] To achieve this, the polymers were synthesized with many reaction sites for SML per polymer chain. This limits comparability to the approach presented here.

The polymer building blocks were synthesized with the peptidic end groups already quantiatively incorporated (see Scheme 15), eliminating the need for polymers with specific reactive end groups and a resource-intensive post-polymerization modification. This approach significantly extends the pool of polymers available for SML and facilitates research on various chain lengths, polymer architectures, and peptidic end groups. Furthermore, an attempt was made here to greatly increase the efficiency of SML by suppressing equilibrium formation with Ni-salt, allowing the use of polymeric building blocks in equimolar amounts.[28]



Scheme 15: Schematic display of SML connecting two polymers with complementary peptidic end groups to form block-copolymers.

The experiments were conducted by dissolving a polymer containing a nucleophilic end group, the counterpart-polymer with a recognition sequence and Ni-salt in SML-buffer. SrtA-4M (10 mol-%) was added to start the reaction, and aliquots were taken for analysis after various reaction times for up to 72 h. Unlike the peptide assays in Chapter 5.1, 0 h samples contained only a mixture of the two polymers without SrtA to ensure that the enzyme does not affect the analysis. These aliquots were lyophilized, and the resulting solids were analyzed with SEC using water and NMP as eluents. Additional MALDI-ToF MS analysis was unsuccessful as it had been when analyzing the polymers separately. The amount of sortase was increased compared to the small molecule assays to account for the much higher steric hindrance of the polymers.

It should be noted that the calculation of polymer ratios for SML was based on the results obtained by SEC of the deprotected polymers (chapter 4.3, Table 4) and thus may be not accurate for all polymers tested, resulting in potentially different polymer ratios than calculated. If different SEC datasets were available, the result closest to the theoretical MW was used. Data from the NMR spectra was not used due to overlapping peaks leading to a very inaccurate MW determination.

The parameters tested within this set of experiments were: different chemical compositions, polymer architectures (linear and brush), chain lengths, and excess of the nucleophilic polymer.

Unfortunately, the data obtained by NMP-SEC was not usable for analysis. The formation of a new polymer species was not detectable in all cases, even though the SEC with water as eluent showed a distinct new peak. This result suggests that the column resolution used for NMP-SEC is insufficient in the MW-range investigated here. For reference, a direct comparison of NMP and aqueous SEC datasets can be found in Figure 38, appendix.

Figure 24 A depicts a typical SEC analysis of SML using two different polymers P(NAM)$_{50}$-FLFG-LPETGG-HG and GG-P(OEGA)$_{20}$. The first polymer exhibited the extended recognition sequence containing His, which made the addition of $NiSO_4$ possible to remove cleaved GG-HG from the equilibrium. As a consequence, equimolar amounts of polymers were used for this experiment. The SEC traces are clearly showing the formation of a new polymer peak at higher MW than the educt peaks. Because of the rather small differences in MW, the educt polymer peaks are not visible as two distinct peaks and deconvolution did not result in useful fit functions using three separate peaks. The large shoulder observed in the 0 h sample is unexpected and its origin is unknown. It was not investigated further because the other samples do not show this shoulder and the other peak fits well to the expected values. In order to be comparable, all datasets were normalized to 1 at the educt polymer peak. The expected decrease in the educt polymer peak throughout the reaction is not visible due to normalization. Although this data is valuable, offsets in peak intensity caused by slightly different amounts of sample used for analysis necessitated normalization.

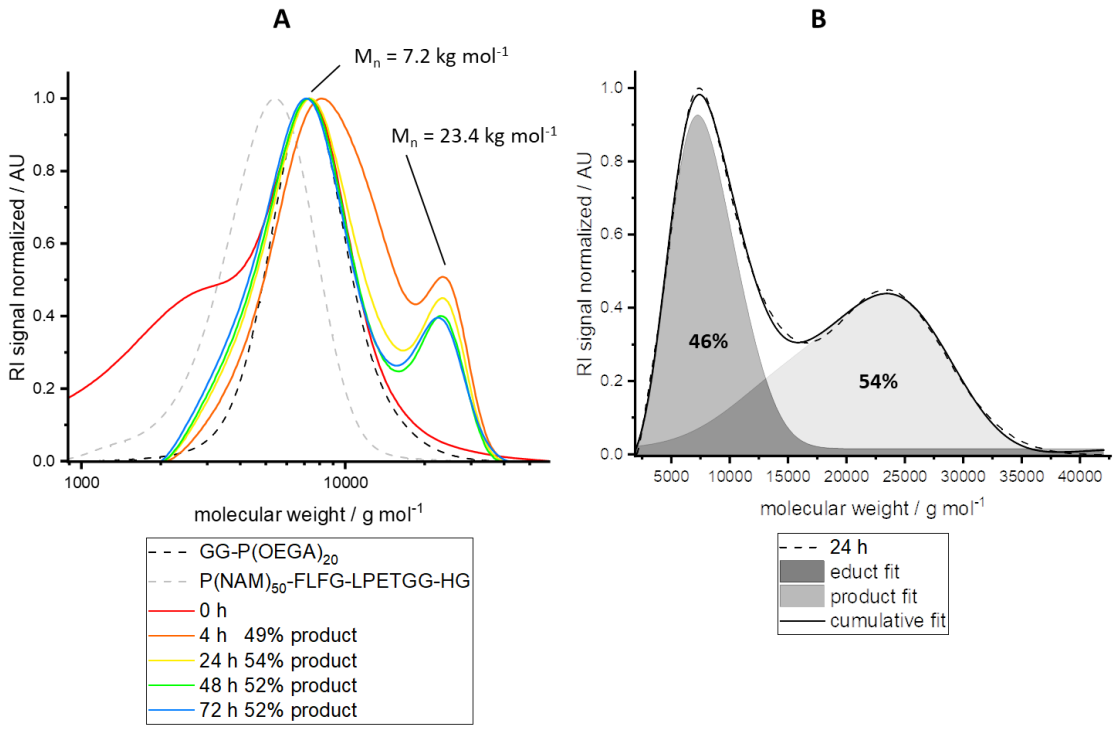


Figure 24: **A**: SEC curves at different reaction times for SML of 0.1 mM GG-P(OEGA)$_{20}$ and 0.1 mM P(NAM)$_{50}$ FLF-LPETGG-HG, 10 mol-% SrtA-4M and 0.2 mM $NiSO_4$ with conversion calculated via deconvolution. Data normalized to 1 at educt peak. **B**: exemplary display of deconvolution and conversion calculation for 24 h sample. Gaussian fit function was used to fit peaks.

For this particular experiment, 1:1 molar ratios of the respective polymers were used in combination with Ni salt to suppress the back reaction.

The SEC data was deconvoluted to separate educt and product peaks, and the area under the respective curves was calculated to determine the amount of product formed (Figure 24 B). Because the refractive indices (signal measured in SEC) of educts and product may not be the same, and elugram deconvolution is purely based on mathematical models, this method of conversion calculation is not completely accurate, but it provides a useful approximation.

In order to fully characterize the product polymers and gain accurate information about yields, separation of product and educt polymers would be required. Due to time constraints, no additional purification or detailed analysis was performed.

Figure 25 depicts the analysis of SML of two P(NAM) polymers with varying chain lengths. Similar to the experiment before, the educt polymer peaks are not separated. The SEC data after 4 h reaction time clearly shows the formation of a new peak with greatly increased MW.

Comparing the SEC data of both experiments, the reaction reached maximum product formation quickly, resulting in a maximum of 54 % and 66-67 % product after 24 h.

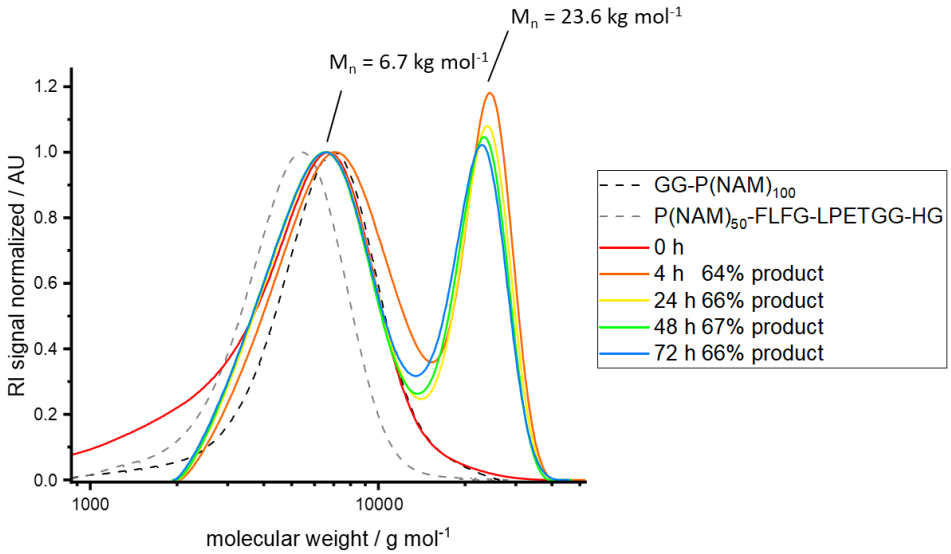


Figure 25: SEC curves at different reaction times for SML of 0.1 mM GG-P(NAM)$_{100}$ and 0.1 mM P(NAM)$_{50}$-FLF LPETGG-HG, 10 mol-% SrtA-4M and 0.2 mM $NiSO_4$ with educt conversions calculated via deconvolution. Data normalized to 1 at the educt peak.

After that, product levels dropped very slightly to 52 % after 72 h in the experiment using P(OEGA)$_{20}$ and P(NAM)$_{50}$. It is unclear whether this is due to side reactions (hydrolysis), which result in an irreversible side product as part of the equilibrium, or whether this finding is within the error of the quantification method. Increased amounts of hydrolysis side product have been reported when running SML for extended periods of time, particularly when using enhanced variants of SrtA.[42]

However, because the cleaved peptide sequence (GG-HG) is expected to be complexed by Ni^{2+} ions, no back reaction should occur. If the complexation of the cleaved peptide sequence by Ni^{2+} did not work as expected, the reaction equilibrium would remain intact, yielding separated polymers, which would eventually lead to a distinct decrease in the product polymer peaks due to the competing irreversible hydrolysis side-reaction as seen in the SrtA assays (chapter 5.1). This decrease is expected to be visible in the SEC data (before normalization). Since no such distinct decrease was visible, it is assumed that the complexation by nickel is working as expected.

It is obvious, that the educt polymers are not completely consumed in both experiments. This may be attributed to a generally limited access of SrtA to the polymer end groups, which are potentially contained within the coil most of the time.

Small differences in conversion have been observed when using different polymer architectures (brush-type P(OEGA) 52 % conversion and linear P(NAM) 67 % conversion), however, direct comparison is difficult due to varying MW of the polymers used.

The MW values obtained by SEC of educt and product polymers in both experiments do not match the expected values of approximately double MW, which is most likely due to a significant change in the hydrodynamic radius of the product polymers. Again, separation and further analysis of the product polymers - preferably by an absolute method - would be required for a more in-depth analysis of the product polymers. For more information about suitable methods, see chapter 4.2.2.
All experiments conducted with polymer chains longer than 100 repeat units did not result in successful ligation, irrespective of the monomer used (DMA, OEGA, NIPAM). A summary of the analytical data of all polymer-polymer SML experiments can be found in Table 10, appendix. The reason is most likely the further decreased accessibility of the end groups to SrtA when employing long polymer chains.
In conclusion, SML of two synthethic polymers using equimolar amounts is possible but quantitative conversion was not achieved and a limitation in polymer chain length has been observed. With the limited amount of available data, it is not possible to provide definite reasons for these limitations. Possible reasons are the increased steric demand of synthetic polymers with increased chain lengths and the potentially limited accessibility of the end groups. Another important factor may be the polymer architecture as observed when comparing conversions using brush-type polymers and linear chains. A comparison with literature data is not possible because polymer-polymer SML has not been reported yet. There are several studies on the synthesis of protein-polymer conjugates via grafting-to SML[27,28], but directly comparing proteins and their somewhat rigid structure with flexible polymer chains is not viable. It is possible that protein end groups are much more exposed due to their rigid 3D structure compared to the ever-changing random coil structure of a synthetic polymer chain.
Other experiments were conducted without the use of Ni salts, but no ligated product was obtained. In these experiments, 10-fold excess of the polymers containing the nucleophilic end group was used to shift the equilibrium and increase product formation (see Figure 37, appendix for exemplary SEC data). There are two possible explanations for this finding: First, the polymers used for these experiments were much longer (DP 100 and higher), and based on the findings from the previous experiments, SML of long polymer chains is not successful. According to the proposed mechanism of SML transpeptidation, the formation of the covalent LPET-SrtA complex is the rate-limiting step.[151] This means that if there is no or very little LPETGG accessible to SrtA, a large excess of nucleophile will not result in more product. Second, all polymers used in the successful experiments contained a FLFG-LPETGG-HG end group, whereas the experiments without Ni-salts had the shortest peptide sequence recognizable by SrtA (LPETGG). As discussed in chapter 5.1, this sequence by itself was found to be not recognizable by SrtA, a modification of the free leucine (L)

N-terminus was required. This modification was found to be not limited to an addition of an amino acid, which would theoretically make a polymer-LPETGG a valid sortase substrate. Nonetheless, it is possible that this construct is not recognized by sortase because it was not tested successfully in SML. More experimental data is required to determine whether a high MW or the need for another spacer-peptide between LPETGG and polymer is the cause of unsuccessful SML.

In summary, this chapter presents the proof-of-concept for SML of two synthethic polymers to form block-copolymers. The synthesis of two different copolymers, $P(OEGA)_{20}$-b-$P(NAM)_{50}$ and $P(NAM)_{100}$-b-$P(NAM)_{50}$ was successful using 1:1 molar ratios of the educt polymers. The obtained data indicates that the maximum chain length for polymer-polymer SML is DP 100, which is most likely due to the increased steric demand of longer polymers and the fact that, unlike with proteins, the polymer end group is not accessible to sortase most of the time. This can also be attributed to the low educt conversion despite the use of Ni-salt to suppress equilibrium formation.
With the amount of data available to me, I am unable to draw additional conclusions and demonstrate the full scope of polymer-polymer SML. More research using a broader range of polymers, chain lengths, and strategies to avoid equilibrium formation is needed to fully explore the possibilities and limitations of this method to synthesize block copolymers.
In the future, SML could provide another synthesis pathway to obtain post-polymerization block copolymers in addition to click eactions (azide-alkyne cycloaddition, thiol-ene and others). The two main issues in classical click chemistry of polymers are the not necessarily quantitative introduction of clickable moieties in a post polymerization reaction and the possibility of influencing polymerization when those moieties are part of initiator or chain-transfer agent. Both lead to unusable polymer chains that must be removed after the click reaction. Furthermore, one major drawback of commonly used click reaction chemistries is the use of toxic reagents (azides) and catalysts (heavy metals), which are very difficult to completely remove, limiting application of the block copolymers in biomedicine.[162] Alternative techniques for avoiding hazardous reagents have been developed, but are far less thoroughly researched than the commonly used ones.[20]
The approach presented here (SPPS and CTA attachment followed by XPI-RAFT polymerization) ensures that all polymer chains have the desired end group and can thus be used for SML. In terms of toxicity, SML uses only naturally occurring building blocks except for the CTA, which can be easily removed in a post-polymerization reaction. It has already been demonstrated that SML can be performed in living animals without adverse effects.[163]
In order to be a valuable tool in post-polymerization block-copolymer synthesis, improvements in educt conversion to avoid the need to separate products from unreacted educts are required. In

addition, a much broader pool of usable polymers – particularly in terms of chain length – must be achieved.

However, it is possible that the physical limitations of sterical demand of the polymer chains and end group accessibility prove insurmountable, limiting polymer-polymer SML to short polymers. This limitation combined with the substantial effort required to obtain polymer-peptide constructs would prevent SML from competing with traditional copolymer synthesis or click chemistry.

5.3 Protein-Polymer Coupling via SML

5.3.1 Protein Choice and Expression

The proteins used for the ligation with polymers were chosen based on their availability, general robustness, and accessibility of the N- and C-termini for SrtA. In addition, a protein of interest in current biomedical research was used.

Because most biotechnological tasks were carried out in collaboration with M. Michaelis (University of Potsdam, group of Prof. H. Möller), a well-understood model protein was used that was available and established in their group. The chosen protein is a site-specific N126W mutant of the carbohydrate binding module 3b (CBM3b^{N126W}, from now on simply CBM, the structure is shown in Figure 26) found in a protein domain of the cellobiohydrolase 9a (Cbh9A) from *Clostridium thermocellum*.[164] CBM was chosen primarily because its structure is well understood; both termini are not part of secondary structures, making them flexible and sortase-accessible. The analysis of the flexibility of the termini was performed by M. Davari (Leibniz Institute for Plant Biochemistry, Halle), using temperature factor computation.

Furthermore, modifications to obtain the variants needed for SML with either N-terminal glycine (G-CBM) or C-terminal LPETGG (CBM-LPETGG) were easily possible.

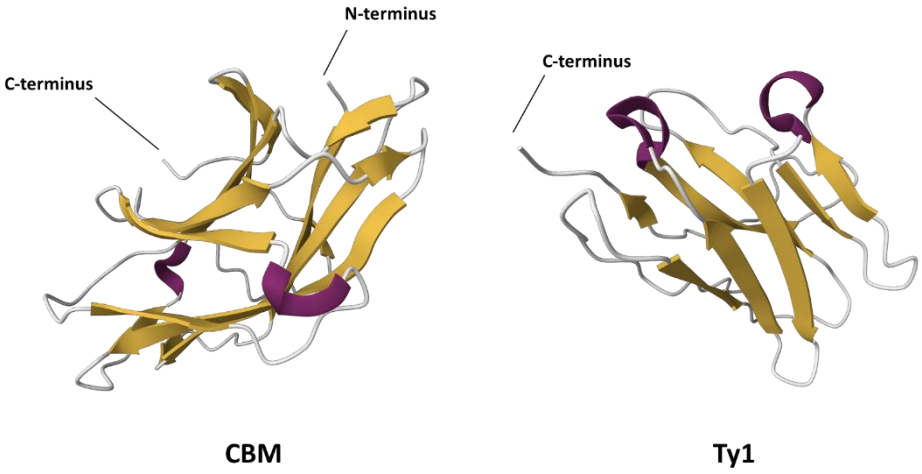


Figure 26: Structures of the proteins used within this work. The coloration highlights secondary structures, thus showing the flexiblitiy of the relevant termini. The illustration is based on CBM (PDB ID 2YLK)[157] and Ty1 (PDB ID 6ZXN)[154] and was created using the RSCB PDB online viewer.

A functional nanobody was chosen as a second protein for SML. Hanke and coworkers reported he alpaca-derived single domain antibody fragment Ty1 (structure shown in Figure 26) which showed promising properties in the neutralization of SARS-CoV-2 by blocking the angiotensin converting

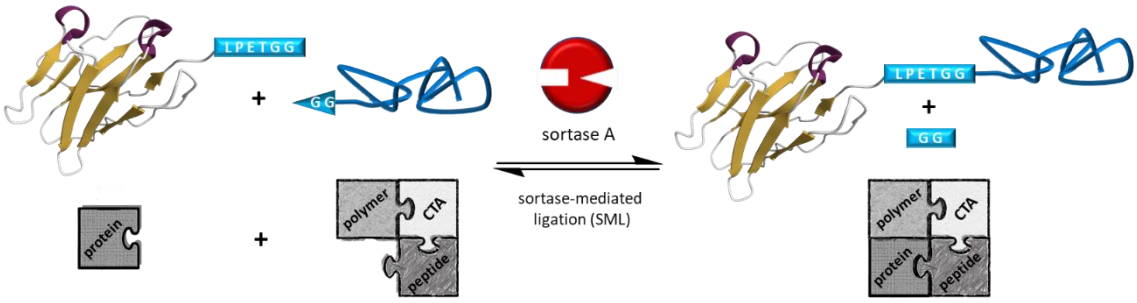
enzyme 2 receptor interaction.[156] In general, nanobodies have distinct advantages compared to the larger antibodies. Because of their smaller size compared to full-length antibodies, they exhibit deeper tissue penetration and can be produced easier and more cost-effective. However, this comes at the expense of a shorter serum half-life.[165] The latter makes nanobodies highly interesting for my work because polymer conjugation is one of the most important methods for improving protein serum half-life. As a result, Ty1 is an excellent model to investigate SML with polymers that are being considered PEG alternatives.

Hanke *et al.* improved the SARS-CoV-2 neutralization potency of Ty1 significantly by introducing multivalency based on a construct of up to four Ty1 molecules connected by multi-arm PEG. They synthesized this Ty1-PEG construct by attaching a functional group capable of azide/alkyne click chemistry to Ty1 using SML. Ty1 had been modified previously with the sortase recognition sequence.[88] This work clearly shows that the C-terminus of Ty1 is accessible to sortase, and thus Ty1-LPETGG (from now on simply Ty1) is well suited for my work to study the ligation of long polymer chains. S. Petrović (University of Potsdam, group of Prof. P. Wendler) kindly provided samples of Ty1 containing C-terminal LPETGG.

Both CBM-LPETGG and Ty1 contained a C-terminal His-tag which was used for protein purification and is useful in SML to separate unreacted protein from the reaction mixture. Ni-affinity chromatography, which temporarily binds histidine-containing proteins to the column material, can be used to achieve this separation.[166] Because the His-tag is cleaved off during SML, the column does not retain the ligated protein-polymer conjugate. In addition, the His-tag enables the use of Ni^{2+} ions to shift the SML equilibrium via complexation of the peptide side-product (see chapters 1.3 and 5.1).

5.3.2 Protein-Polymer SML

To synthesize protein-polymer conjugates, a protein (G-CBM, CBM-LPETGG, or Ty1), equipped with the peptide sequence required for SML at C- or N-terminus, was paired with a synthetic polymer carrying the complementary peptidic end group. A schematic representation of the reaction is depicted in Scheme 16.



Scheme 16: Schematic display of SML connecting a protein and synthethic polymer to form protein-polymer conjugates. The protein illustration is based on Ty1 (PDB ID 6ZXN)[156] and was created using the RSCB PDB online viewer.

In general, the desired amounts of protein and polymer were dissolved in SML buffer and $NiSO_4$ salt (0.2 mM) was added if required Then, SrtA was added to start the reaction. Based on the successful results in polymer-polymer coupling (chapter 5.2) and a report of successful polymer-protein-SML,[28] 10 mol-% of SrtA-4M was used to account for the much larger steric hindrance caused by the large molecules compared to the SML assays in chapter 5.1. Aliquots were taken at regular intervals for 24 hours, and the reaction was quenched with HCl. The reaction samples were then analyzed using SDS-PAGE. Based on the positive sortase assays results, RP-HPLC analysis was attempted but failed because (presumably denatured) CBM precipitated in the column. Unfortunately, MS-analysis was not available to me during the experiments.

Based on previous experience in our group on analyzing proteins with a single synthethic polymer chain using SDS-PAGE, I expected product bands to be slightly blurred and not exactly at, but close to the expected MW.[32] This distinctive effect observed in SDS-PAGE analysis of protein-(PEG) polymer conjugate was discovered to be due to the formation of a PEG-SDS complex, which causes the previously uncharged polymer to become negatively charged.[167] This complex interferes with the local protein charge and alters the dragging force by which the protein-polymer conjugate is dragged through the electrophoresis gel. This is causing the band to appear at slightly different positions than expected in relation to the MW marker bands.[167] The literature report contains only information about PEG, but because the polymers used here were either similar to PEG (P(OEGA)) or contained strongly complexating nitrogen (all other polymers), the same effect was anticipated.

Analysis of the SDS-PAGE results from samples containing G-CBM or CBM-LPETGG in combination with different polymers in different ratios (up to 30 molar equivalents of polymer) and Ni-salt to inhibit back-reaction revealed that no distinct new bands assignable to protein-polymer conjugates were formed (data not shown). Some very broad bands were visible for some reactions, but because the educt band intensity did not decrease during the reaction, I concluded that SML was not successful.

Although SDS-PAGE does not indicate product formation, protein-polymer SML experiments were analyzed using ^{1}H-^{15}N-heteronuclear single quantum coherence (HSQC)-NMR spectroscopy to possibly gain a better understanding of the reaction. NMR measurements were performed by M. Michaelis (group of Prof. H. Möller, University of Potsdam). Previous experiments taught us that in order to obtain usable NMR spectra, one must use a very short polymer (if nitrogen-containing) to avoid introducing too many ^{15}N-atoms and thus "diluting" the signals produced by CBM.[32] Furthermore, introduction of long polymer chains limited diffusion, resulting in broad peaks. This is why very short polymers (oligomers, theoretical DP = 7-10) were used for NMR analysis of SML. Because the chain length was too short to yield useful SEC data, these oligomers were not fully characterized, and ESI-MS analysis to analyze peptide-polymer conjugates was only used at a later stage of the project. Since the positions of the coherence signals of the unmodified CBM were previously known, signal changes after SML would indicate successful ligation. The differences between native and allegedly polymer-equipped CBM were mapped and assigned to the amino acids in the 3-D structure based on the appearance, disappearance, and change in chemical shift of certain coherence peaks. More information on the use of NMR spectroscopy to analyze protein-polymer conjugates can be found in the perspective by M. Mathieu-Gaedke *et al*.[34]

Figure 27 depicts a typical HSQC-NMR spectrum and the corresponding 3-D structure with highlighted changes before and after SML. Because the polymer chain is located at the C-terminus, medium to strong changes in this region are expected for successful C-terminal protein SML.

Based on previous experience in our group, the changes in HSQC-NMR spectra were less than expected.[32] In addition, mixing free polymer chains with native CBM produced similar shifts, implying that the free polymer chain interacts with the protein without being covalently attached to it.

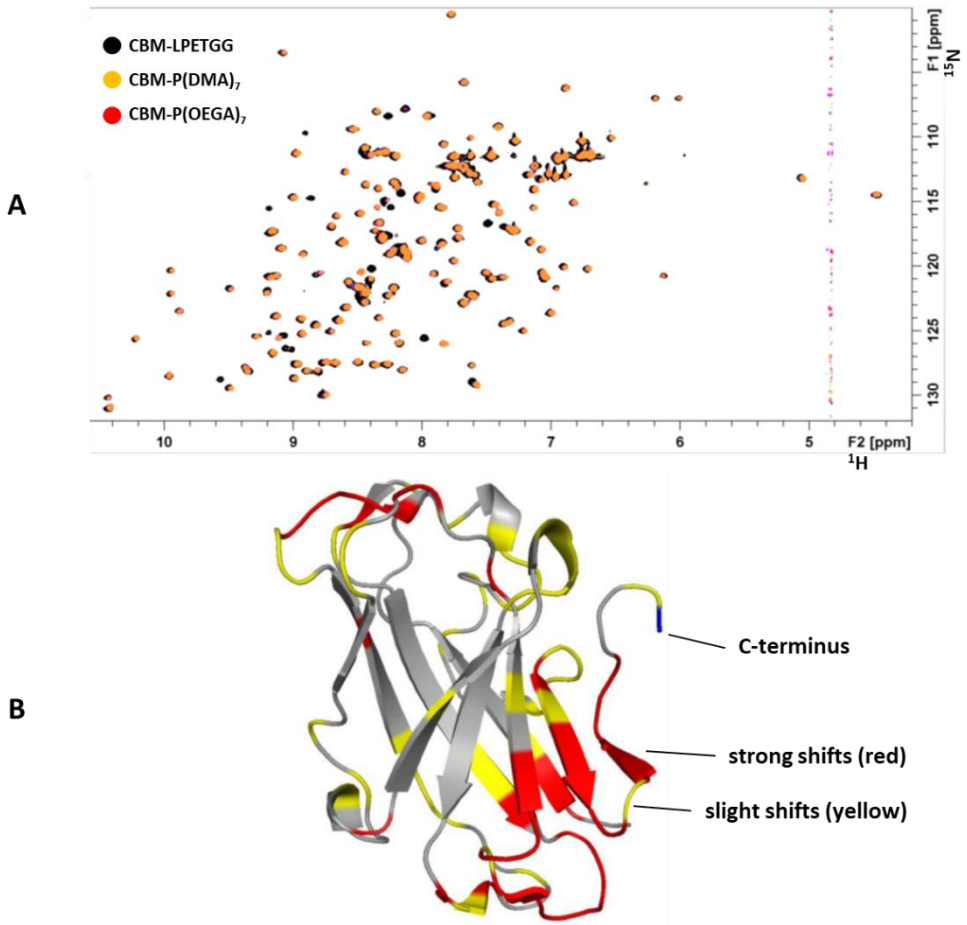


Figure 27: **A**: Exemplary ^{1}H-^{15}N-HSQC-NMR analysis of native CBM in comparison with expected SML reaction products containing $P(DMA)_7$ and $P(OEGA)_7$. **B**: 3-D structure of CBM with shifts highlighted, derived from NMR-spectra in A, data from CBM-$P(DMA)_7$. Figure courtesy of M. Michaelis, University of Potsdam.

Purification of the reaction using a His-trap column was performed to determine whether any reaction occurred when using CBM-LPETGG for SML. All species containing a His-tag are bound to the column using this method, while all other molecules are flushed out. Sortase and unmodified CBM both had a His-tag in this case. CBM's His-tag was located at the C-terminus, resulting in cleavage upon reaction with sortase. Native CBM and SrtA were theoretically removed from the reaction mixture, leaving unreacted polymer and modified CBM in solution.

Judging from the absorption signal at 280 nm of the column outflow, the solution contained molecules. Because the polymer does not absorb UV-light at 280 nm, I deducted that the solution contained CBM that had reacted with sortase, cleaving the His-tag. Taking the other results into consideration, it is very likely that CBM reacts with sortase, but the nucleophilic polymer end group is not accessible, resulting in hydrolysis of CBM.

In order to determine the reaction outcome in more detail, MS-analysis will be required; preferably combined with removal of His-tagged species and protein-SEC.
Based on the preliminary findings presented here, CBM may not be suitable for SML. There are no literature studies available that show successful SML of CBM, and despite the computational analysis of the termini, it is possible that neither terminus is accessible to SrtA. This phenomenon has been reported before when Row *et al.* received unsatisfactory results when attempting to use SML to modify bovine insulin with a peptide.[52] It is unlikely that the polymer is the cause of this issues because SML-experiments with polymers exhibiting similar chain lengths were successful when combined with other polymers (see chapter 5.2) or a different protein (see next paragraph).
Overall, analysis of SML with CBM did not lead to conclusive results. Purification of the reaction mixture by Ni-affinity chromatography suggested successful cleavage of the recognition peptide sequence (and the His-tag) which can only be caused by SrtA, but no conjugates were detected via SDS-PAGE and NMR spectroscopy.

Next, the nanobody Ty1 was employed for C-terminal SML. Since only very limited amounts were available and for reasons of time, only SDS-PAGE and MALDI-ToF MS were used to analyze the results. Unfortunately, the latter method did not provide useful data.
Molliner-Morro *et al.* recently reported successful SML using Ty1, attaching a small molecule to its C-terminus.[88] Because of the certainty that Ty1 can be used for SML, as well as the positive results obtained for GG-P(NAM)$_{20}$ in polymer-polymer SML in chapter 5.2, an attempt was made to ligate these two building blocks using 10 mol-% SrtA-4M as before. With the other goal of this work in mind, making SML as efficient as possible, an experiment was carried out using the building blocks in 1:1 molar ratio with addition of $NiSO_4$ to suppress equilibrium formation (see chapter 5.1.2 for details of this method).
Figure 28 depicts the results of the SDS-PAGE, which show that the classic approach to SML using a large excess of nucleophile (30 eq) yields the desired product (blurred band in approximately the expected position) within 17 h and complete consumption of Ty1 educt. However, using a 1:1 ratio and $NiSO_4$ did not result in the formation of as much product and incomplete Ty1 consumption (Figure 28 B). Quantification of the results using image processing software[51] was not attempted due to the (allegedly) slighty different sample concentrations. MS analysis would be required for final confirmation of product formation.

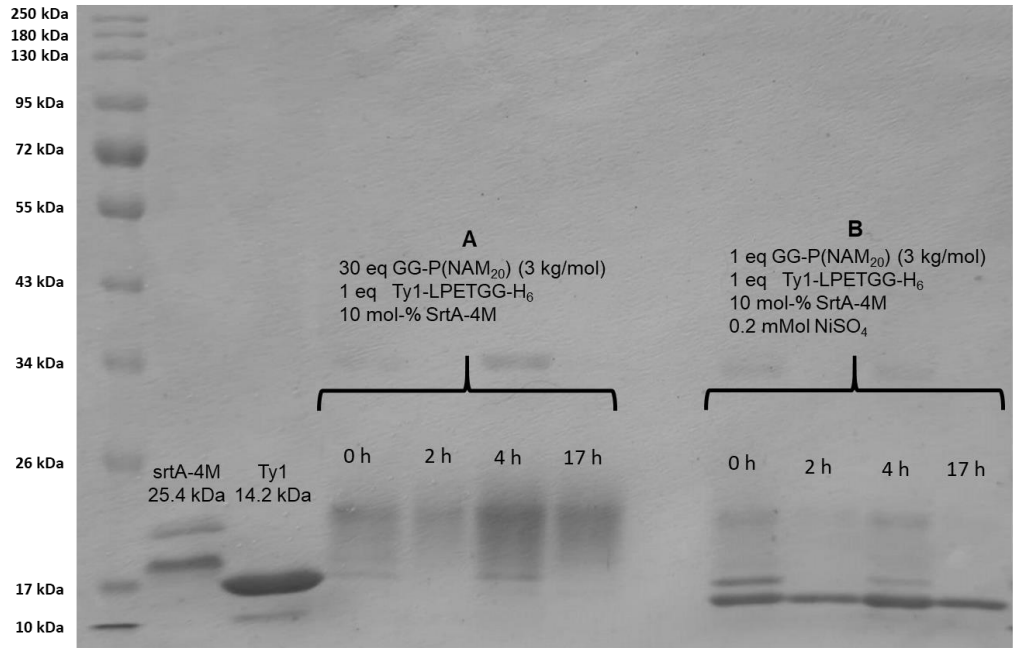


Figure 28: SDS-PAGE of SML of Ty1 and GG-P(NAM)$_{20}$, using an excess of polymer (**A**) or equimolar amounts of educts in combination with $NiSO_4$ (**B**). Samples were stained using Coomassie blue.

Qualitative comparison is possible with the reported findings of Reed and coworkers, who used a functional protein to couple a short (DP 11) but branched GG-PEG polymer to the protein C-terminus. They achieved very good estimated yields of 90 % after 12 h using 1:1 molar ratios, 4 eq $NiSO_4$ and 20 mol-% SrtA-WT.[28] This yield was obviously not possible here, which could be attributed to either the C-terminus of Ty1 or the peptidic end group of GG-P(NAM)$_{20}$ being less accessible. In addition, it is possible that the reagent ratios were not equimolar due to underestimation of the polymer MW by SEC (discussed in detail in chapter 4.2.2). More experimental data is required to fully understand and optimize Ni-assisted SML with Ty1.

An attempt to use GG-P(NAM)$_{100}$ in excess (50 eq) with the same reaction conditions as before did not result in conjugate formation (analyzed by SDS-PAGE, data not shown). This finding is not surprising taking previous results from polymer-polymer SML into consideration, and underlines the assumption that sterics may be a limiting factor for SML.

In conclusion, the results obtained from using G-CBM, CBM-LPETGG and various different polymers with varying MW for SML indicate that no reaction is taking place. Since protein-polymer conjugates via SML have been reported for various proteins,[27,28,33] and considering positive results using Ty1, it is likely that despite the appearance that the C- and N-termini of CBM are both flexible and on the protein surface, neither G-CBM nor CBM-LPETGG is accessible for SrtA.

SML of Ty1 and $P(NAM)_{20}$, on the other hand, was successful, albeit much more so when a 30-fold excess of polymer was used compared to an equimolar approach using Ni-assisted SML.

This chapter presents a foundation upon which more research can be conducted to optimize efficient SML to obtain a variety of protein-polymer conjugates. To fully explore the possibilities and limits of this approach, more research and more in-depth analysis needs to be done in the future. Before attempting to synthesize protein-polymer conjugates, it is critical to determine a protein's suitability for SML (for example, by SML with a small-molecule dye).

6 SUMMARY AND OUTLOOK

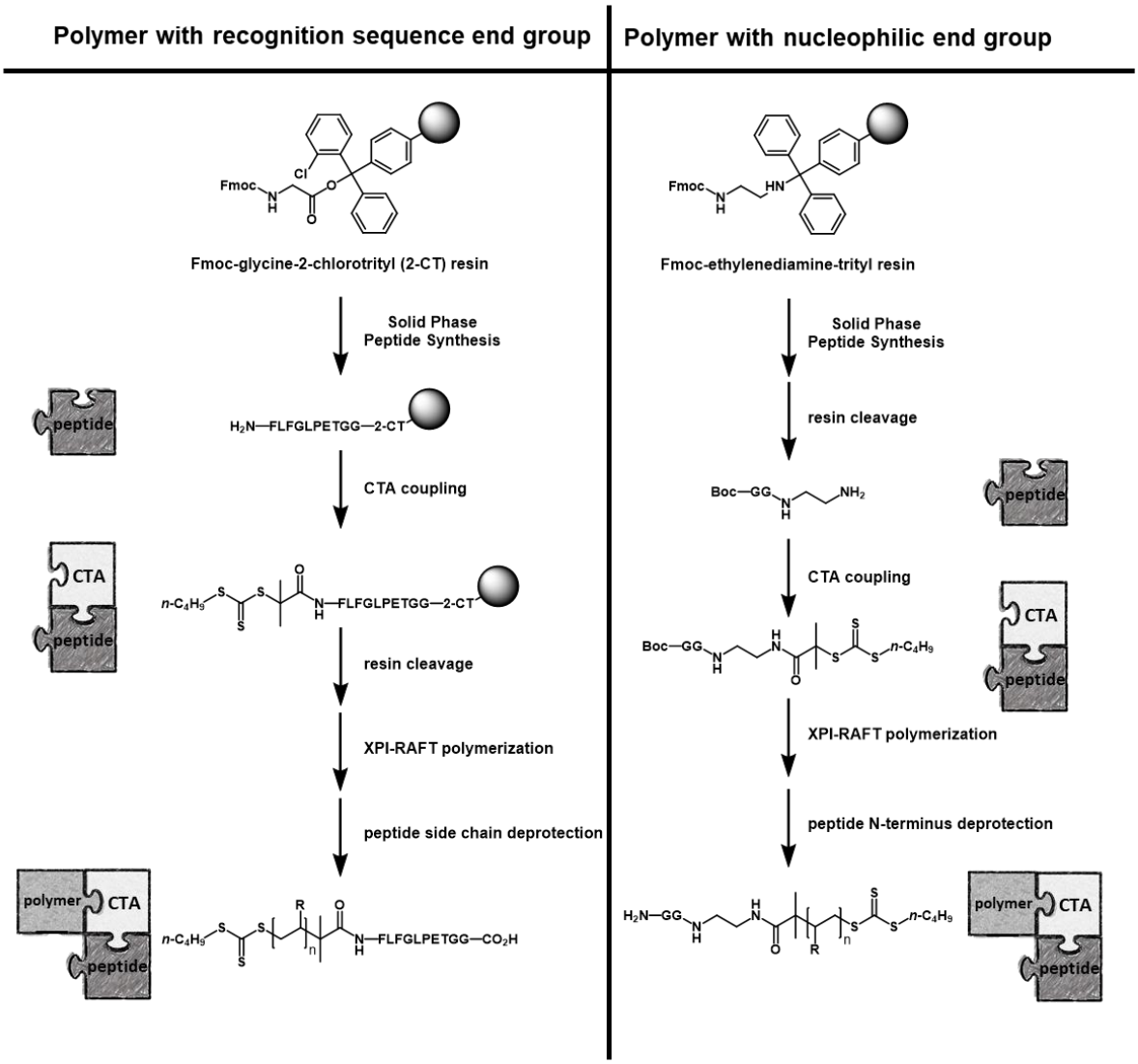
Summary

To overcome some disadvantages of protein-polymer conjugates (mostly performance loss due to unspecific attachment of synthethic polymer chains and immunogenicity of PEGylated proteins) used in modern medicine, a new pathway to this highly important class of phamaceuticals was explored. Sortase A, an enzyme capable of ligating two peptides together, was used as a catalyst to achieve the synthesis of site-specific protein-polymer conjugates and block copolymers.

A main part of this thesis addressed the synthesis of peptide-polymers as building blocks for sortase-mediated ligation (SML). These were subsequently utilized to investigate possibilities and limitations when using synthethic macromolecules in SML.

The building blocks for SML consist of a peptide, a CTA to control RAFT polymerization, and a synthethic polymer. SML requires two different peptide sequences, a recognition sequence (LPETGG) with a free C-terminus and a nucleophilic sequence (e.g., GG) exhibiting a free N-terminus. Thus, two cases have to be distinguished when synthesizing the building blocks for SML: the polymer needs to be linked to the N-terminus of the recognition sequence LPETG or to the C-terminus of the oligoglycine nucleophile. As a result, two fundamentally different synthethic pathways were used. Scheme 17 depicts all synthesis steps needed to form these building blocks.

For the polymer-peptide conjugate with the recognition sequence (Scheme 17 left side), a peptide containing the recognition motif was synthesized by solid-phase peptide synthesis (SPPS) at first. The peptides were obtained using an automated peptide synthesizer and their synthesis went smoothly, resulting in high yields and pure products. After I found that the leucine N-terminus has to be functionalized with either another amino acid or a different organic molecule to be recognized by SrtA, several (aromatic) amino acids were included to facilitate peak assignment during HPLC analysis. After testing different acid-labile resin linkers, a 2-CT linker was chosen to allow mild



Scheme 17: Synthesis pathways used in this project to obtain polymeric building blocks for the use in SML with recognition sequence peptide end group (left) or nucleophilic end group (right).

cleavage of the peptide-CTA from the resin to avoid decomposition of the CTA. Next, the CTA to control RAFT polymerization was attached to the peptide N-terminus via carboxylic acid activation. This reaction was based on a literature report[112] but required optimization of the coupling reagents and ultimately resulted in acceptable yields and very high purities of peptide-CTA. Two robust tertiary CTAs based on trithiocarbonates were tested: The more labile of the two (with a cyano moiety) produced very low yields and was therefore not used for further experiments. In addition, an ATRP initiator was connected to the peptide on the resin.

Subsequent attempts of surface-initiated RAFT polymerization with the CTA-peptide immobilized on the resin were unsuccessful. Neither thermal nor photochemical (PET-RAFT) initiation with addition of free CTA resulted in polymer growth from the resin surface. Surface-initiated ATRP

was also tested, but led to the same results. The porosity of the Merrifield resin, which limited the accessibility of CTA or the initiator for radicals, was most likely the cause of the unsuccessful polymerization. Consequently, the peptide CTA was cleaved from the resin and was used to control RAFT polymerization in solution.

The synthesis of the peptide-polymer containing the nucleophilic peptide (Scheme 17, right) required some modifications in comparison to the recognition sequence peptide, because a free N-terminus was required. Therefore, an ethylenediamine (EDA)-prefunctionalized resin with a trityl-linker was used for SPPS. Following resin cleavage of the obtained Boc-GG-EDA, a primary amine was obtained, which was immediately coupled to the CTA. In contrast to some other approaches, this method allows for independent peptide synthesis and the product contains no ester bonds in the backbone which are prone to hydrolysis under certain conditions. This peptide CTA was then used for RAFT polymerization followed by deprotection of the N-terminus.

Reaction control using peptide CTAs was poor for both thermally induced RAFT and PET-RAFT polymerization. As a result, the recently developed method of XPI-RAFT polymerization[87] was tested and led to good results. A major advantage of this method is that each polymer chain contains the desired peptide end group, as opposed to thermally induced RAFT-polymerization, where the radical starter is incorporated as an end group as well. The XPI-RAFT polymerization method requires two CTAs, a thrithiocarbonate and a xanthate (XAN). The synthesis of the XAN-peptide proceeded without problems and resulted in acceptable yields and very high purities under the reaction conditions as optimized with the trithiocarbonate.

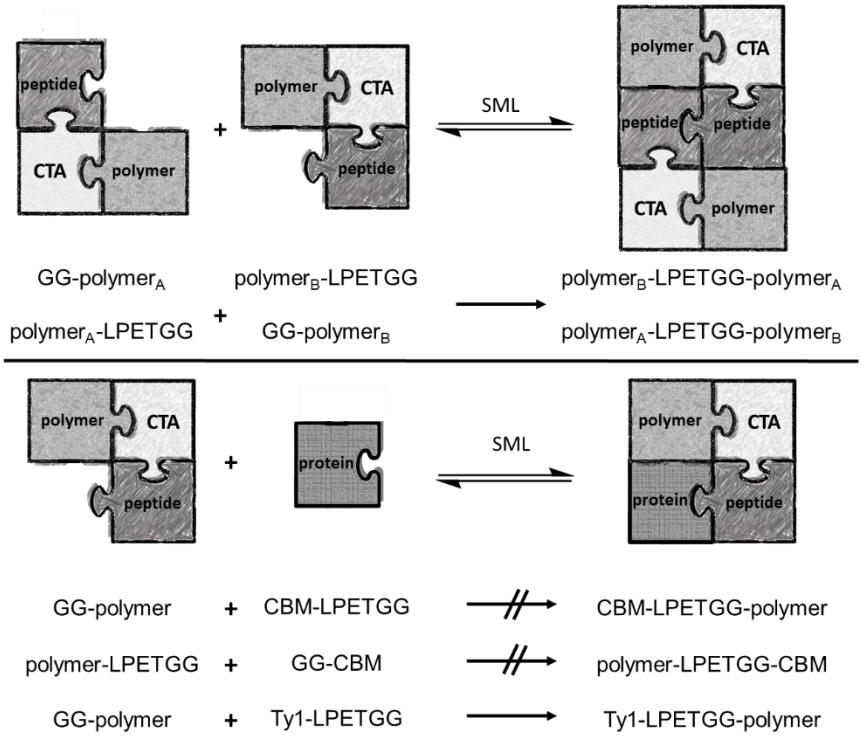
Several (meth)acrylate and acrylamide-based monomers (NAM, NIPAM, DMA, OEGA, SPE, Glu-HEMA) were chosen for their potential as PEG replacements in medical applications of protein-polymer conjugates. Polymers with varying chain lengths were synthesized using XPI-RAFT polymerization to be tested in SML. SEC with refractive index or UV detection was used to determine the MW distribution of the polymers with varied results due to the peptide end groups. For short polymers, ESI-MS was used successfully to prove high end group fidelity and the expected MW. Except for the zwitterionic SPE and the methacrylate-based Glu-HEMA, the polymerization of all monomers produced good results with narrow dispersities.

Because resin cleavage yielded peptides with protected side chains, these protection groups had to be removed before SML. Deprotection was carried out using TFA and worked as expected.

Next, the performance of SrtA was tested in assays using short peptides. Out of the two tested variants, the engineered SrtA-4M was chosen over the wild-type SrtA. Furthermore, SrtA exhibited quick aging, leading to a significant drop in performance when stored in buffer at 4 °C for extended

periods of time (several weeks). As a consequence, freshly thawed SrtA-4M was used for subsequent SML experiments.

Because SML is an equilibrium reaction, two methods of suppressing the back reaction were tested to maximize yield without using a large excess of one reagent. The method of using a Trp-zipper peptide, described to lead to a β-hairpin after SML, which is not recognized by SrtA, did not produce the anticipated results on the basis of experiments with peptides. In contrast, "metal-assisted SML" using Ni^{2+} ions to complex the histidine-containing side-product formed during SML produced very promising results, resulting in complete educt peptide conversion despite using equimolar amounts of reactants.



Scheme 18: Building blocks used for the synthesis of block copolymers (top), or protein-polymer conjugates (bottom) via SML.

To gain a first understanding of SML using large molecules, block copolymers were synthesized using two polymers with complementary peptide end groups (see Scheme 18, top). SEC analysis proved that SrtA was capable of ligating such polymers. Product yields of 50-70 % were achieved when P(NAM)$_{50}$-FLFGLPETGGHG/GG-P(OEGA)$_{20}$ and P(NAM)$_{50}$-FLFG-LPETGG-HG/GG-P(NAM)$_{100}$ were ligated using equimolar amounts with the addition of Ni^{2+} ions to suppress the back reaction. The yield of the reaction was found to be largely independent of polymer composition: a brush polymer (P(OEGA)) yielded comparable yields to a linear polymer (P(NAM)).

Comparing the obtained data from all experiments, it appears that the steric demand of long polymer chains is significantly impeding SML, because using polymers with a (theoretical) chain length of more than 100 repeat units did not result in block copolymers regardless of chemical composition.

To synthesize protein-polymer conjugates, two proteins were chosen: a well-known model protein (CBM) and a functional nanobody (Ty1) with promising properties for use against SARS-CoV-2. Cooperation partners carried out the expression of protein variants with either a SrtA recognition sequence at the C-terminus and/or nucleophilic glycine at the N-terminus. Both proteins were used for SML in combination with various polymers (see Scheme 18, bottom). No conjugate formation was observed in the case of N- or C-terminal ligation of CBM when the reaction mixture was analyzed via SDS-PAGE. Analysis via 2D-NMR spectroscopy of a C-terminal SML experiment with CBM revealed minor changes in the chemical surroundings of the C-terminus of the protein, but these changes were much smaller than expected. Despite temperature factor calculations indicating that both termini are not part of a rigid secondary structure, the obtained experimental results suggest that neither CBM terminus is accessible to SrtA and thus not usable for SML.

On the other hand, Ty1 with LPETGG at the C-terminus in combination with short polymers (theoretical DP = 20) resulted in protein-polymer conjugates. Longer polymers (DP100), as seen previously when synthesizing block copolymers, did not result in successful ligation. When the short polymer was used in excess, the educt protein was consumed quickly and completely according to SDS-PAGE monitoring of the reaction progress. The use of Ni^{2+} ions and equimolar amounts of reactants did not result in quantitative educt consumption. This could be due to steric hindrance of either protein or polymer, as previously observed when synthesizing block copolymers.

The data collected during this study does not provide a comprehensive view of the capabilities and limitations of polymer-polymer and protein-polymer ligation via SML. Especially a thorough and complete characterization of both building blocks and ligation products is still required. It does, however, provide a foundation for further research in the field of protein-polymer conjugate synthesis catalyzed by SrtA.

Outlook

In order to facilitate future research, I would like to offer potential solutions to some of the issues that I discovered during my research.

The first area of further investigations should be the detailed analysis of the obtained conjugates, particularly polymer-protein constructs. As Reed and colleagues demonstrated,[28] mass spectrometry can be a powerful tool for provinding a wealth of information about conjugates as

well as (unwanted) side products. It may be necessary to purify the reaction mixture before analysis, for example by Ni-NTA column chromatography (to remove all molecules containing His-tags such as SrtA, unreacted protein, and the cleaved GG-H_6 peptide sequence). Because unreacted polymer cannot be removed from the reaction mixture using this method, ion exchange chromatography, SEC, or simply dialysis may be useful.

After purification, the yield of protein-polymer SML can be determined more accurately. However, when using sample absorption at 260 or 280 nm to calculate protein concentration, keep in mind that the polymer may change the extincition coefficient of the conjugate depending on its chemical composition. To address this issue, one could use the absorption of the CTA at 320 nm, where no (native) protein absorption is observed. Morgenstern *et al.* developed a very elegant protocol to determine accurate concentrations of protein-polymer conjugates using CTA absorption.[16] However, this method produces accurate results only when high end group fidelity is maintained (as demonstrated here by the ESI-MS data).

To gain additional insight, one could employ the analytical techniques mentioned in the perspective by Mathieu-Gaedke *et al.*[34] In particular, 2D-NMR spectroscopy has been used with great success to monitor the changes in the protein 3D-structure when a polymer chain is attached.[34,168] So far, this has mostly been done using PEGylated proteins that have been synthesized by addressing lysine side chains. Expanding this field of research by offering the ability to study the effect of various polymers connected exclusively at one of the protein's termini, thus eliminating effects occurring when mixtures of protein-polymer conjugates with different attachment points are used, would be worthwile.

Furthermore, more analytical data for the peptide-polymer building blocks is required. Because NMR spectroscopy and MALDI-ToF MS did not provide the desired information about the molecular weight (distribution) and SEC using a simple UV or refractive index detector is greatly influenced by the peptide end groups, additional analysis is required. SEC equipped with a viscosity detector and/or a MALS detector can provide detailed and accurate information about MW distribution, including solution behavior. Because of the size limit for renal clearance (the maximum hydrodynamic radius for renal filtration is approximately 3.8 nm), detailed information about the MW distribution of the polymers is required for use in pharmaceutic applications.[15]

Finally, the immunogenicity of the peptide sequences used (potentially in combination with the proteins and polymers) needs to be analyzed because (poly)peptides are an important trigger for (human) immune system response.[169] Studies of SML in living animals suggested that no negative effects are to be expected,[163] however, long-term studies are not yet available. In addition, the cytotoxicity of the polymers is an important factor. Recently reported data suggests that the

cytotoxicity of polymers obtained by RAFT polymerization, especially using trithiocarbonate-CTAs, is not a concern. On the other hand, the cytotoxicity of polymers synthesized via ATRP largely depends on the efficient removal of the copper catalyst.[170] Nonetheless, cytotoxicity tests of the polymers need to be performed.

Aside from additional analytical data, the limitations of SML discovered in this study need to be explored more thoroughly. Protein and peptide-polymer end group accessibility and sterical demand are both important factors in SML success. Because the connection of two large molecules at one specific point is entropically not favored, sterics may be an overall limiting factor in *grafting-to* SML. End group accessibility might be improved by introducing a flexible spacer such as a short PEG chain,[171] or a peptide sequence such as $(G_4S)_n$.[155,172] Of course, this adds to the complexity of SML, but it may be essential to successfully use longer polymers in SML. SML will most likely be unable to compete with existing methods for synthesizing site-specific protein-polymer conjugates unless incorporation of long polymer chains (the limit found here was 100 repeat units) is possible.

Finally, methodical changes are worthwhile to investigate.

Using the solid support from SPPS for surface-initiated polymerization to synthesize peptide-polymer building blocks containing the recognition sequence would significantly facilitate the procedure (simple peptide-polymer purification, among others). Experiments conducted using SI-ATRP and SI-RAFT polymerization were not successful here (see Chapter 4.2.1). A potential reason is the Merrifield resin used for these experiments, which may not be suited for SI-polymerization. Reports in the scientific literature suggest that SI-polymerization using similar resin types provides good results, but is highly dependent on the correct monomer-resin-solvent combination.[173–176] As a consequence, I think it is worthwile to test SI-polymerization with different solid supports. A promising starting point may be the use of a hydrophilic Merrifield resin based on crosslinked PEG, or agarose beads, as used by Murata *et. al.* for their approach to solid-phase synthesis of protein-polymer conjugates via *grafting-from*.[21]

As previously stated, sterics may be the biggest limiting factor of SML, so a *grafting-from* approach may be successful. Hu *et al.* reported the use of SML to attach an ATRP initiator to a protein. They used ATRP to obtain a site-specific protein-polymer that outperformed the established PEGylated variant of the same protein.[25] In this case, the polymerization conditions had to be carefully chosen so as not to harm the protein and the solvent options were very limited. However, with the various new methods of light-induced RAFT polymerization (PET-RAFT or XPI-RAFT), high reaction temperatures are not of concern any more and this concept could be expanded to various polymers.

They also compared the conjugate yields of *grafting-to* (via SML) and *grafting-from* and discovered that the latter was far superior (1 % yield vs. 66 %).

More experimental data using Ni^{2+} ions to suppress the back reaction is needed. This very promising strategy should be further explored with various polymers and proteins to confirm the excellent performance reported in the scientific literature.[28] Other very promising strategies to maximize product formation are modifications of the recognition sequence using depsipeptides[100,177] or isoacyl peptides[101]. However, because these synthethic amino acids cannot be (easily) incorporated into proteins, the use of the modified recognition sequence is limited to the polymeric building block and, consequently, N-terminal protein SML.

Besides SrtA as the leading enzyme in the field of peptide ligases, other enzymes capable of transpeptidation are emerging.[42] A very promising peptidyl asparaginyl ligase, butelase 1, has several advantages over SrtA: To begin, the peptide recognition sequence consists of only three AA as opposed to 5 of SrtA and after ligation, only one AA of the recognition sequence remains in the product. Second, a large number of dipeptides are accepted as nucleophiles for transpeptidation. Third, the catalytic efficiency is much higher: Nguyen *et al.* estimated a polypeptide cyclization using butelase 1 instead of enhanced SrtA-5M to use 250-fold less enzyme while still being 800 times faster.[178] Although protein cyclization with butelase 1 is irreversible, intermolecular peptidation has been discovered to be reversible.[179] However, a year later Nguyen *et al.* reported irreversible intermolecular peptidation of a substrate based on a thiodepsipeptide (thioester instead of amide).[180] This substrate can be easily incorporated in both Boc- and Fmoc-based SPPS. The different ligation strategies employing butlease 1 are depicted in Figure 29.

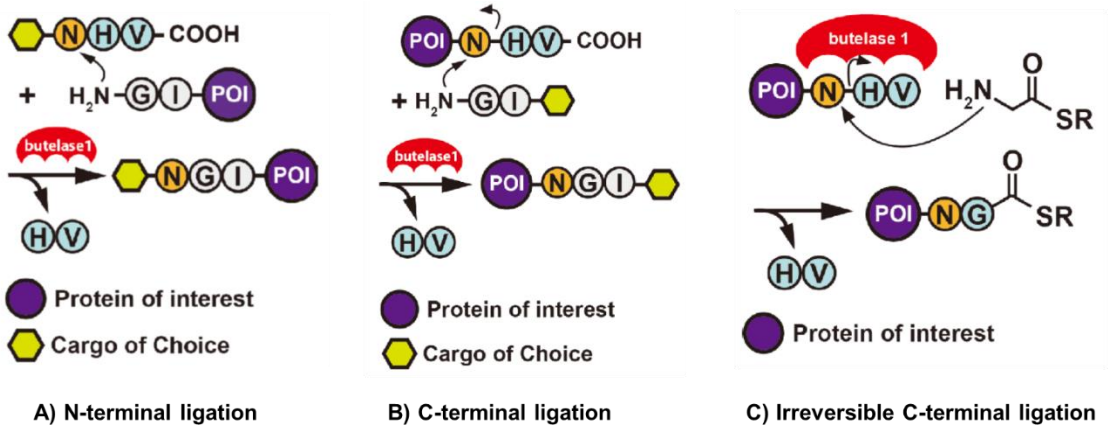


Figure 29: Butelase-mediated ligation containing the C-terminal NHV recognition sequence either at the cargo (A) or the protein (B). Irreversible transpeptidation by using a thiodepsipeptide as nucleophile (C). Reprinted from reference[179].

Research on butelase 1 was very cumbersome until recently, because expression of the recombinant protein remained unsuccessful. In 2019, however, the groups of Mylne and Su reported two independent pathways to recombinant butelase 1.[181,182] So far, research on butelase 1 was mostly focused on peptide and protein cyclization (its natural function),[183–185] however, given the now available recombinant enzyme, development of many more applications either in combination with[186] or as replacement for SrtA are to be expected.

7 EXPERIMENTAL PART

7.1 CHEMICALS

Table 5: Utilized materials.

Name	Purity	Additives	Supplier
1-butanethiol	99 %	stabilized with BHT	Merck
1,4-dioxane	reagent grade		Roth
2,2′-azobis(2-methylpropionitrile) (AIBN)	98 %		Merck
[2-(methacryloyloxy)ethyl]dimethyl-(3-sulfopropyl)ammonium hydroxide (SPE)	95 %		Merck
2-methyl-2-bromopropanoic acid (BMPA)	98 %		Merck
2,2,2- trifluoroethanol (TFE)	≥ 99.8 %		Roth
3-[bis(dimethylamino)methyliumyl]-3H-benzotriazol-1-oxide hexafluorophosphate (HBTU)	98 %		Carbolution
4,4'-azobis(4-cyanopentanoic acid) (ABCVA)	≥ 98.0 %		Merck
4-dimethylaminopyridine (DMAP)	≥ 99.0 %		Fluka
acetic anhydride	≥ 99.0 %		Merck
acetobromoglucose	≥ 95 %		Merck
acetone	≥ 99 %		vwr
acetonitrile (MeCN)	LC-MS grade		vwr
acrylamide	≥ 99 %		vwr
aqueous hydrochloric acid	1 N		Th. Geyer
Bacterial Sortase Substrate III, Abz/DNP TFA	98 %		Med Chem Express
benzene	99.5 %		Roth
calcium chloride	≥ 98 %		Roth
carbon disulfide	99.9 %		Fisher Sci

chloroform-d	99.8 atom% D		Roth
deuterium oxide	99.8 atom% D		Roth
dichloromethane	≥ 99.8 %	stabilized with amylene	Th. Geyer
diethylether	≥ 99.5%		Th. Geyer
dimethylsulfoxide (DMSO)	> 99.0 %		Roth
dimethylsulfoxide-d_6	99.8 atom% D		Roth
ethanol	absolute, > 99.8 %		Merck
ethyl acetate (EA)	99.9 %		vwr
ethylenediamine	≥ 99.5 %		Roth
Eosin Y	99 %		Merck
hydroxybenzotriazole (HOBt) hydrate	98 %		Merck
hydrochloric acid, concentrated			Th. Geyer
imidazole	≥ 99.5 %		vwr
iodine	≥ 99.8 %		Roth
iso-propanol	100 %		vwr
magnesium sulfate	≥ 99 %		Merck
methanol	LC-MS grade		vwr
N-acryloylmorpholine (NAM)	97 %	monomethyl ether hydroquinone as inhibitor	Merck
N,*N*'-diisopropylcarbodiimide (DIC)	≥ 99 %		Roth
N,*N*-dimethylacrylamide (DMA)	99 %	monomethyl ether hydroquinone as inhibitor	Merck
N,*N*-dimethylformamide (DMF)	peptide grade		vwr
n-hexane (Hex)	≥ 95.0 %		Chemsolute
nickel(II)-sulfate hexahydrate	≥ 99 %		Roth

N-isopropylacrylamide (NIPAM)	97 %		Merck
N, *N*-methylene bisacrylamide	99 %		vwr
N-methylmorpholine (NMM)	> 98.5 %		Roth
pentane	> 95 %		Roth
petrol ether 60-80 °C	analytical grade		Chemsolute
piperidine	99 %		Roth
pyridine	≥ 99.5 %		Roth
silica gel 60 for column chromatography			Roth
silver triflate	≥ 99.0 %		Merck
sodium bicarbonate	≥ 99.5 %		Roth
sodium chloride	≥ 99.0 %		Roth
sodium dodecyl sulfate (SDS)	≥ 99.0 %		Roth
sodium hydroxide	≥ 98.8 %		Chemsolute
tetrahydrofurane	pure		Merck
tetrahydrofurane	≥ 99.9 %	stabilized with BHT	Chemsolute
Thioglo® 1	≥ 99 %		Merck
triethylamine	≥ 99.5 %		Roth
trifluoroacetic acid (TFA)	≥ 99.9 %		Roth
triisopropylsilane (TIPS)	≥ 98 %		Merck
tris(hydroxymehyl)aminomethane (TRIS)	≥ 99.5 %		Roth

NIPAM was recrystallized from *n*-heptane prior to use. DMA was distilled to remove the inhibitors. NAM and $OEGA_9$ were passed through a short Al_2O_3 column twice to remove inhibitors. 1,4-dioxane for polymerization was distilled prior to use to remove inhibitors. AIBN was recrystallized from *n*-hexane prior to use. Deionized water for dialysis or buffer preparation was further purified by a Millipore Milli-Q Plus water purification system (resistivity 18 MΩ cm).

7.2 METHODS AND INSTRUMENTATION

NMR spectroscopy

NMR spectra were recorded using a Bruker Avance 300 NMR spectrometer operating at 300 MHz for ^{1}H measurements and 75 MHz for ^{13}C measurements or a Bruker Avance 400 NMR operating at 100 MHz for ^{13}C measurements. Chemical shifts δ are given in ppm referring to the respective solvent peaks at δ (^{1}H) 7.26 ppm and δ (^{13}C) 77.16 ppm for $CHCl_3$, δ (^{1}H) 2.50 ppm for DMSO, and at δ (^{1}H) 4.79 ppm for H_2O.

Reverse-Phase High-Performance Liquid Chromatography (RP-HPLC)

Measurements were performed using an Agilent 1200 system (Agilent Technolgies Inc., USA), equipped with a degasser, autosampler, multi-wavelength DAD detector and automated fraction collector. The column used for analytical runs was a Brownlee SPP2.7 (C18, 2.1x150 mm) from PerkinElmer, USA. For preparative application, a Nucleosil 120-5 (C18, 10x250 mm) manufactured by Machery-Nagel, Germany has been used. Solvent systems were comprised of Milli-Q water with 0.1 % TFA (solvent A) and acetonitrile (solvent B) or methanol (solvent C) containing 0.1 % TFA. Solvent purity was HPLC gradient grade or higher. For instrument control and data collection ChemStation B03.02 (Agilent Technolgies, Inc. USA) was used. Samples for ESI-MS analysis were collected using an automated fraction collector. Data for absorption at 205 (all organic molecules), 220 (peptide bonds), 280 (aromatic amino acids) and 355 (CTAs and Abz-containing peptides) nm was recorded.

Method for peptidic sortase assay (chapter 5.1 and 7.9): 5 % B to 50 % B within 20 min, linear gradient, isocratic 50 % B for 10 min, 95 % B for 10 min, 5 % B for 15 min for reequilibration. Flow rate 0.2-0.3 mL/min.

Method for peptide analysis (chapter 3.1.2): 5 % B to 95 % B within 40 min, linear gradient, isocratic 95 % B for 10 min, 5 % B for 15 min for reequilibration. Flow rate 0.2-0.3 mL/min.

Method for CTA-peptides and peptide-ATRP initiator analysis (chapter 3.1.4): since analyte was not soluble in A and B, methanol (C) had to be used. 95 % A, 5 % C to 95 % C within 40 min, linear gradient, isocratic 95 % C for 10 min, 5 % C for 15 min for reequilibration. Flow rate 0.2 mL/min.

Raw data was processed using OriginPro 2021b (64-bit) SR2.

Electrospray Ionization Mass Spectromety (ESI-MS)

Spectra were recorded using a single quadrupole Flexar SQ 300 MS detector manufactured by PerkinElmer, USA. The instrument was controlled using SingleQuad 2.2. Before measurement, the sample was dissolved in HPLC-MS grade methanol, acetonitrile or water containing 0.1 % TFA. Samples collected after HPLC analysis were measured directly or diluted with HPLC-MS grade methanol. All spectra were recorded in positive ion mode unless stated otherwise. The measurement range was set manually to fit with the expected m/z values for different analytes. Raw data was processed using OriginPro 2021b (64-bit) SR2.

Size Exclusion Chromatography (SEC)

Polymers were analyzed using size exclusion chromatography (SEC) in *N*-methyl-2-pyrrolidone (NMP) + 0.5 % LiBr with simultaneous UV and RI detection at room temperature (flow rate 0.5 mL/min). The stationary phase used was a 300 x 8 mm PSS GRAM linear M column (7 µm particle size). Other samples were analyzed by SEC in THF with simultaneous UV and RI detection at room temperature (flow rate 0.5 mL/min). The stationary phase was a 300 x 8 mm^2 PSS SDV linear M column (3 µm particle size).

Aqueous SEC used 0.1 M NaCl + 0.3 % formic acid as mobile phase with a flowrate of 1.0 ml/min. Stationary phase was a PSS NOVEMA Max VS, (1000 Å particle size) at 40 °C. Data was collected using an RI detector.

All samples were filtered through 0.45 µm filters and the injected volume was always 100 µL. Narrowly distributed polystyrene standards were used for calibration of NMP and THF SEC, aqueous SEC was calibrated with a narrowly distributed PEG standard (all standards from PSS, Mainz, Germany).

P(SPE) (chapter 4.2.2, Table 4) was analyzed using a SEC-3010 system by WGE Dr. Bures, Germany. Two PL HFIPgel (300 x 7.5 mm) columns were used with 1,1,1,3,3,3-hexafluorisopropanol (HFIP) + 0.05 M $NaO_2C_2F_3$ (40 °C, 0.8 mL/min). The RI signal was collected for data processing and a narrowly distributed PMMA standard was used for calibration. Raw data was processed using OriginPro 2021b (64-bit) SR2. For analysis of polymer-polymer SML (chapter 5.2), the calibrated data was normalized to the highest peak to enable comparison of different samples. To separate educts and products, the deconvolution function of Origin was used (Gaussian fit function, sometimes parameter adjustment needed for good fit) and the areas under the respective curves was calculated.

Protein Concentration

Concentrations of all proteins were measured after dilution to approximately 1 mg/mL using a NanoVue Plus™ (GE Healthcare, USA) in protein mode (A280). For some proteins, concentration was measured in 1 cm pathlength disposable cuvettes using a SPECORD 210 (Analytik Jena) UV-Vis spectrometer. Concentration calculation was based on the respective extinction coefficients ε which were calculated using the ProtParam tool from Expasy. The extinction coefficients were as follows: G-CBM ε = 27390 $M^{-1}cm^{-1}$, CBM-LPETGG ε = 27390 $M^{-1}cm^{-1}$, SrtA-WT ε = 14440 $M^{-1}cm^{-1}$, SrtA-4M ε = 18910 $M^{-1}cm^{-1}$, Ty-1 ε = 18450 $M^{-1}cm^{-1}$. Computed coefficients were based on the assumption that all cysteine residues are reduced.

SDS-PAGE

The expression, purification steps and the purity of the protein solutions as well as SML reaction solutions were monitored by tricine-sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)[187] employing a 16 % polyacrylamide (97 % acrylamide and 3 % *N*,*N*- methylene bisacrylamide) gel. For SDS-PAGE, first the 10 % gels and 16 % were prepared according to the composition listed in Table 6.

Table 6: Composition of the SDS-gel page.

ingredients	stacking gel	resolving gel	
		10 % gel	16 % gel
acrylamide solution	0.25 mL	—	2 mL
gel buffer 3x	0.75 mL	1 mL	2.5 mL
60 % glycerol in H_2O	—	—	—
urea	—	1 mL	1.65 g
deionized water	2 mL	1 mL	1 mL
TEME	2 µL	3 µL	3 µL
10 % (w/v) APS	20 µL	30 µL	30 µL

A sample of 10 µL (different protein concentrations depending on reaction conditions) was mixed with the same amount of loading buffer (Table 7) and incubated for 10 min at 95 °C to denature the protein. As the loading buffer contains sodium dodecyl sulfate (SDS), the denatured state of the protein will be solubilized and stabilized by SDS. The gel was run at 15 W approximately 40 min

in the presence of cathode and anode buffer (Table 7). The gel chamber was filled with cathode buffer and the electrophoresis chamber with anode buffer. All buffers used are listed in Table 7.

Table 7: SDS-PAGE buffers

ingredients	loading buffer	gel buffer	anode buffer	cathode buffer
bromophenol blue	0.04 % (w/v)	—	—	—
glycerol	20 % (v/v)	—	—	—
DTT	2 % (w/v)	—	—	—
EDTA	25 mM	—	—	—
SDS	4 % (w/v)	0.3 % (w/v)	—	1 % (w/v)
TRIS	100 mM	300 mM	200 mM	100 mM
Tricin	—	—	—	100 mM
pH	6.8	8.45	8.90	8.25

Broad range molecular weight standard was used to estimate the molecular weights. We used a protein-molecular weight marker from New England Biolabs® (Unstained Protein Standard, Broad Range 10-200 kDa) with following molecular weights: 200, 150, 100, 85, 70, 60, 50, 40, 30, 25, 20, 15, and 10 kDa)

Protein bands on the gel were first stained with imidazole-zinc method.[188] The gel was briefly rinsed in water followed by shaking 10 min in 0.2 M imidazole. After imidazole was discarded, the gel was placed in 0.3 M $ZnCl_2$ solution for 30 s. The solution was discarded and bands were visible against dark background. For destaining, the gel was placed for 10 min in 2 % aqueous citric acid. If coomassie stainig was necessary, the gel was placed in a solution of 20 % (v) methanol, 10 % (v) phosphoric acid, 10 % (wt) ammonium sulphate and 0.12 % (wt) Coomassie brilliant blue G-250. After the staining process, the band intensity was further enhanced by destaining a gel in deionized water.

All SDS-PAGE analyses were conducted by M. Michaelis or Y. Müllers.

7.3 SOLID-PHASE PEPTIDE SYNTHESIS (SPPS)

Resin Choice

To obtain peptides with fully deprotected side chains and C-termini bearing an amide functionality, RAM resin (PS AM RAM, loading 0.79 mmol/g, Rapp Polymere) was used. This resin was only used to establish reaction conditions for peptide synthesis and not to synthesize functional building blocks for polymerization.

For peptides exhibiting protected side chains and carboxylic acid functionality at the C-terminus, Fmoc-glycine-prefunctionalized 2-CT-resin (Fmoc-Gly-2CT, loading 0.7 mmol/g, Iris Biotech) was used. For the synthesis of peptides containing ethylenediamine groups at the C-terminus, ethylenediamine-prefunctionalized trityl-resin (1,2-diaminomethane-trityl, loading 0.94 mmol/g, Iris Biotech) was used.

Synthesis Methods

Peptide synthesis was carried out using an automated peptide synthesizer MultiPep RS manufactured by INTAVIS Bioanalytical Instruments AG.

The synthesis scale was 50-100 µM, carried out in single-use 6 mL syringes equipped with a porous (10 µm) frit to retain the resin.

Synthesis procedure was started by swelling the resin in DMF (3 x 5 min), followed by cleavage of the resin-bound Fmoc protection group (2x3 min, 20 v-% piperidine in DMF, 32 eq of piperidine per deprotection step). Amino acid building blocks were bought from Carbolution as *N*-Fmoc and complementary (Boc or *t*-bu) side chain protected versions and were used as received. In order to synthesize the nucleophilic sequence for sortase A, *N*-Boc-protected GG was used as amino acid building block.

Coupling of amino acids (0.5 M in DMF, 5.05 eq per coupling) was facilitated by HBTU (0.6 M in DMF, 5 eq per coupling) and NMM (4 M in DMF, 10 eq per coupling). Coupling time was 30-60 min per coupling, 60 min was used for difficult couplings (G to G and L to P) to ensure complete conversion. After several rinsings with DMF, capping solution (5 v-% acetic anhydride in DMF, 0.8 eq) was applied to render unreacted free amines unable to participate further in the synthesis and thus result in deletion peptides.

For difficult couplings, the Kaiser test was performed after coupling to ensure complete conversion. The test was carried out following a published procedure.[107] In case of a positive Kaiser test result, coupling of the respective amino acid was repeated.

The Fmoc-deprotection, coupling of amino acids and capping procedure was repeated for all amino acid building blocks required in the peptide. Finally, cleavage of the N-terminal Fmoc-protection group was carried out as before, the resin was rinsed several times with DMF, followed by DCM and dried using reduced pressure.

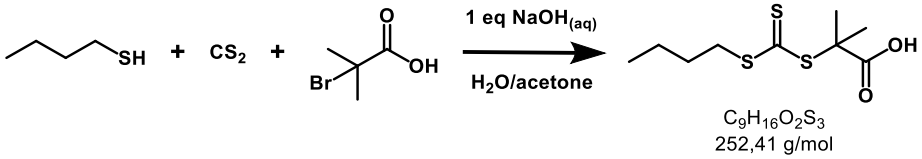
To ensure that the correct peptide sequence was obtained, microcleavage followed by ESI-MS analysis was done. Thus, a small amount of resin was dispersed into pure TFA for 10-30 min. Afterwards, 50 µL of the supernatant was taken, diluted with methanol (1 mL, MS grade) and analyzed in positive ion mode.

For the synthesis of recognition sequence peptides, resin cleavage to isolate the peptide was done only once per peptide to analyze yield and purity (via RP-HPLC, see chapter 7.2) since the CTA was coupled to the peptide on-resin (see chapter 7.5, cleavage procedure included).

In case of the synthesis of nucleophilic sequence peptides, resin cleavage was achieved using 5 v-% TFA in DCM (2 x15 min, 6 mL total) and the resin extracted with MeCN (2 x 3 mL). H_2O (10 mL) was added to the supernatant and the mixture was lyophilized to yield a hygroscopic solid.

7.4 CTA SYNTHESES

Synthesis of BABTC



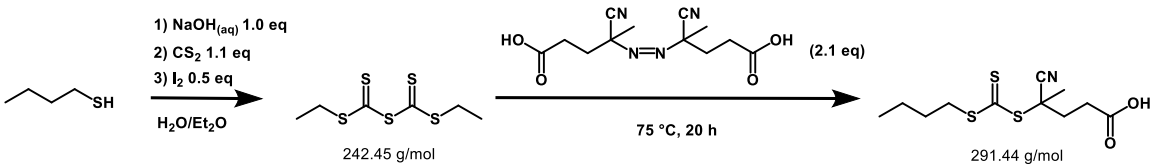
The synthesis was based on a report from Bray *et al.*[189]

Butanethiol (6.0 mL, 55.5 mmol) was mixed with aqueous NaOH (17 wt-%, 4.4 mL, 55.5 mmol, 1.0 eq) in acetone (5 mL) and stirred for 30 min at RT. Next, CS_2 (3.8 mL, 61 mmol, 1.1 eq) was added, and the mixture was stirred for another 30 min at RT, resulting in a clear orange solution. The mixture was cooled in an ice bath and 2-methyl-2-bromopropanoic acid (9.5 g, 56.8 mmol, 1.0 eq) was added step by step while keeping the solution temperature below 30 °C. To dissolve the formed yellow precipitate, more aqueous NaOH (17 wt-%, 4.4 mL, 55.5 mmol, 1.0 eq) was added carefully while monitoring the temperature (<30 °C). The ice bath was removed and the solution was stirred over night at RT. Then, the reaction mixture was diluted with H_2O (50 mL) and washed with *n*-hexane (2 x 50 mL). The aqueous phase was cooled in an ice bath and aqueous HCl (1 N, 50 mL) was added dropwise until a yellow precipitate formed. The precipitate was filtered off and washed with cold water (2 x 50 mL). The residue was dissolved in DCM (50 mL), dried over $MgSO_4$ and filtered. Then, the solvent was evaporated and the remaining yellow solid was recrystallized twice from hexane. After drying, BABTC (10.1 g, 39 mmol 74 %) was obtained as yellow powder which was stored at -20 °C.

^{1}H-NMR (400 MHz, $CDCl_3$): δ = 0.93 (t, 3H, J = 7.3 Hz, -CH_2-CH_3), 1.35-1.52 (m, 2H, -CH_2-CH_3), 1.61-1.69 (m, 2H, S-CH_2-CH_2), 1.72 (s, 6H, C(CH_3)$_2$), 3.29 (t, 2H, J = 7.4 Hz, S-CH_2) ppm.

ESI-MS (MeOH): m/z = 253.4 $[M+H]^+$, expected 253.4.

Synthesis of CPABTC



The synthesis was adapted based on a report from Larnaudie *et al.*[105]

Buthanethiol (7.2 mL, 67.5 mmol) was mixed with aqueous NaOH (33 wt-%, 8 mL, 67.5 mmol, 1.0 eq) in diethylether (130 mL) and stirred for 30 min at RT. Next, CS_2 (4.6 mL, 75.9 mmol, 1.1 eq) was added, and the mixture was stirred for another 30 min at RT, resulting in a clear orange solution. Subsequently, iodine (8.6 g, 33.8 mmol, 0.5 eq) was added carefully and the mixture was stirred for 1 h at RT. The organic phase was washed with aqueous $Na_2S_2O_3$ (2 x 100 mL), water (100 mL), dried over $MgSO_4$ and filtered. After removal of the solvent in vacuo, the bis-trithiocarbonate as yellow oil (7.88 g, 32.3 mmol, 98%) was used in the next step without further purification.

Bis-(butylsulfanylthiocarbonyl) disulfide was dissolved in freshly distilled 1,4-dioxane (180 mL) and 4,4'-azobis(4-cyanopentanoic acid) (ABCVA, 19.13 g, 68.3 mmol, 2.1 eq) was added. The dispersion was heated to 75 °C for 20 h and became a clear orange solution. After cooling to RT and solvent removal *in vacuo*, the orange oil was purified by column chromatography (1:4 v:v EA : Hex, changed to 2:1 v:v EA:Hex after separation of first spot) and yielded pure CPABTC (13.2 g, 45.2 mmol, 70 %) as yellow oil which crystallized after storage at -20 °C for 7 days.

^{1}H-NMR (400 MHz, $CDCl_3$): δ = 0.96 (t, 3H, J = 7.3 Hz, -CH_2-CH_3), 1.34-1.52 (m, 2H, -CH_2-CH_3), 1.63-1.76 (m, 2H, S-CH_2-CH_2), 1.90 (s, 3H, C-CH_3) 2.35-2.62 (m, 2H, S-C-CH_2-), 2.66-2.78 (m, 2H, CH_2-CO_2H), 3.36 (t, 2H, J = 7.4 Hz, S-CH_2) ppm.

ESI-MS (MeOH): m/z = 292.48 $[M+H]^+$, expected 292.44.

Synthesis of XAN

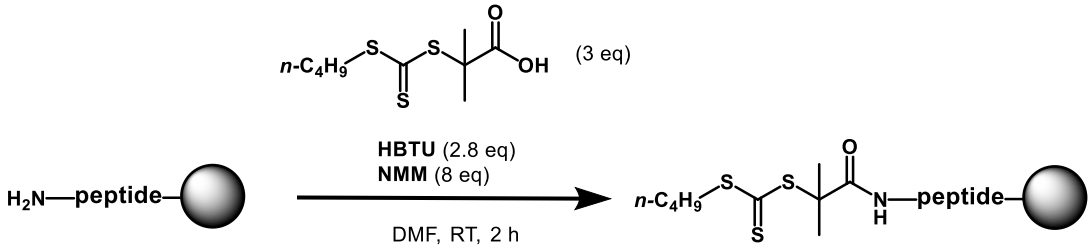
XAN was synthesized by J. Kurki following a published procedure.[106] Product characterization was in accordance with the published data.

7.5 CTA/Initiator Coupling to Peptide

On-Resin Synthesis

The general method for the coupling of CTA and ATRP initiator to an existing peptide on solid support (Merrifield resin) has been reported by Chen *et al.*[110] This method has been modified to yield the products needed here.

The procedure described here is exemplary for BABTC and was conducted in the same manner for CPABTC, XAN and BMPA. This method was used to obtain CTA-peptides bearing the recognition sequence for sortase A.



Dry resin with peptide (amount typically based on 0.1 mmol of peptide loading, N-terminus deprotected, see chapter 7.3) was dispersed for 10 min in NMP (4 mL) to swell within the syringe for SPPS. BABTC (75 mg, 0.30 mmol, 3.0 eq), HBTU (106 mg, 0.28 mmol, 2.8 eq) and NMM (88 µL, 0.80 mmol, 8.0 eq) were dissolved in NMP (2 mL) and stirred in a flask under N_2 atmosphere for 10 min. After removal of NMP from the syringe, the clear yellow reaction solution of BABTC, NMM and HBTU was taken up with the syringe. The syringe was closed with a stopper and left on the shaker for 2 h. To ensure complete reaction, the syringe was opened carefully, a small amount of resin beads were removed and analyzed for primary amines via Kaiser test (examination of tested beads under a digital microscope). Upon a negative test result, the reaction solution was discarded and the resin was washed with DMF (3 x 5 mL) and DCM (3 x 5mL). Remaining DCM was removed *in vacuo*. Dried CTA-peptide-resin was stored at -20 °C.

Yields and purities (determined after resin cleavage, see next paragraph):

BABTC-LPETGG: 62 % yield, 80 % pure

CPABTC-LPETGG: 30 % yield, 45 % pure

XAN-LPETGG: 80 % yield, 87 % pure

BMPA-LPETGG: 96 % yield, 92 % pure

Resin Cleavage

To isolate the CTA-peptide for further use in polymerization, the resin was cleaved. All volumes are given per 0.1 mmol resin batch size (6 mL disposable syringe). Different cleavage mixtures were used for RAM (pure TFA, 3 x 1 h, total 6 mL) and 2-CT (20 v-% TFE in DCM, 3 x 1 h, total 6 mL) resins.

The combined TFA-supernatants (RAM resin) containing the cleaved peptide were precipitated into cold diethylether (40 mL) and the mixture was centrifuged (4000 rpm, 4 min). The supernatant was discarded and the peptide pellet was redispersed in cold diethylether (40 mL) and centrifuged again. After evaporation of remaining diethylether, the peptide was dissolved/dispersed in MilliQ-water (5 mL) and lyophilized to yield a colourless powder. After analysis via ESI-MS and RP-HPLC, peptides were stored at -20 °C.

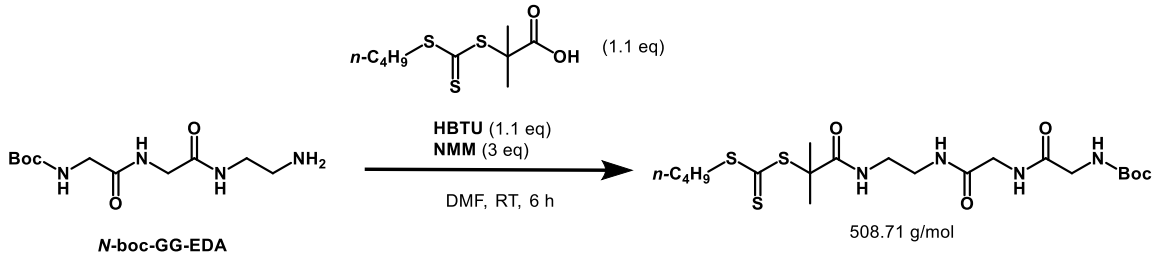
Since precipitation in diethylether was unsuccessful when using 2-CT resin, the combined TFE/DCM-supernatants were mixed with MeCN (5 mL) and benzene (5 mL) and lyophilized on a Schlenk-line equipped with a cooling trap (liquid N_2) to collect the solvents. The remaining partly oily yellow CTA-peptide was redissolved in MeCN (2 mL), benzene (5 mL) was added and lyophilization was repeated to yield yellow powder. The intactness of the trithiocarbonate group was quickly checked via Thioglo-1 JM143 assay for thiols.[108] After analysis via ESI-MS and RP-HPLC, peptides were stored at -20°C.

Synthesis in Solution (*N*-Boc-GG-EDA-BABTC/XAN)

Since modification of the peptide C-terminus is not possible on-resin, the coupling of CTA to the nucleophilic recognition sequence was done in solution.

The reagent *N*-Boc-GG-EDA was synthesized via SPPS and cleaved from the resin as described before.

The procedure described here is exemplary for BABTC and was conducted in the same manner for XAN.



N-boc-GG-EDA

BABTC (349 mg, 1.39 mmol, 1.1 eq), HBTU (525 mg, 1.39 mmol, 1.1 eq) and NMM (416 µL, 3.78 mmol, 3.0 eq) were dissolved in DMF (20 mL) and stirred at RT for 15 min unter N_2

atmosphere. A degassed solution of *N*-Boc-GG-EDA (345 mg, 1.26 mmol) in DMF (4 mL) was added under N_2 atmosphere. The reaction mixture was stirred at RT until results from the Kaiser test remained negative (6 h). After that, water (40 mL) was added to remove DMF and NMM. The mixture was extracted with DCM (3 x 100 mL) and the combined organic phases were dried over Na_2SO_4 and filtered. The solvent was removed in vacuo and the residue was purified via column chromatography (8 v-% MeOH in DCM, three yellow substances collected, product determination via ^{1}H-NMR spectroscopy). The product (BABTC 186 mg, 0.37 mmol, 29 %, XAN 25 mg, 0.06 mmol, 28 %) was dissolved in 1:1 v:v H_2O:MeCN (5 mL) and lyophilized. The yellow, hygroscopic foam was stored at 4 °C and used quickly for polymerization.

^{1}H-NMR (BABTC, 400 MHz, $CDCl_3$): δ = 0.90 (t, 3H, J = 7.3 Hz, -CH_2-CH_3), 1.33-1.42 (m, 2H, -CH_2-CH_3), 1.46 (s, 9H, C(CH_3)$_3$), 1.61-1.66 (m, 2H, S-CH_2-CH_2-), 1.69 (s, 6H, C-(CH_3)$_2$), 3.28 (t, 2H, J = 7.3 Hz, S-CH_2), 3.38 (m_c, 4H, N-CH_2-CH_2-N), 3.38 (br s, 2H, CONH-CH_2-NHCO), 4.01 (d, 2H, HNOC-CH_2-NH-Boc) ppm.

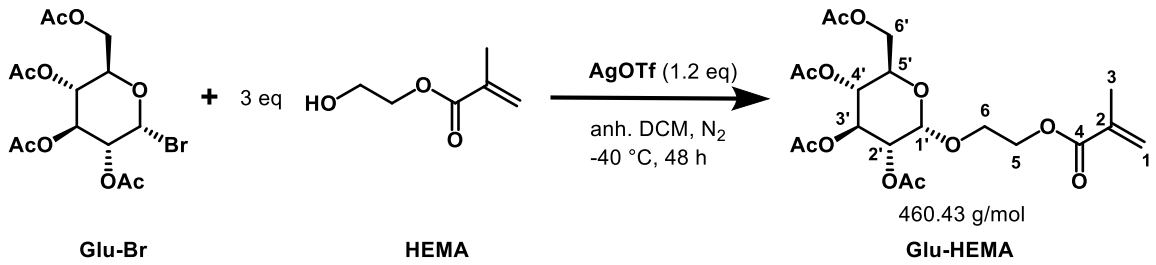
ESI-MS (BABTC, MeOH): m/z = 531.19 $[M+Na]^+$ expected 531.70.

^{1}H-NMR (XAN, 400 MHz, $CDCl_3$): δ = 1.43 (t, 3H, J = 7.3 Hz, -CH_2-CH_3), 1.46 (s, 9H, C(CH_3)$_3$), 1.55-1.61 (m, 2H, o-CH_2-CH_3), 1.87 (br s, 3H, CH-CH_3), 3.28 (t, 2H, 7.3 Hz, S-CH_2), 3.42 (m, 4H, N-CH_2-CH_2-N), 3.91-4.00 (m, 4H, HNOC-CH_2-NHCO-CH_2-NH-Boc) ppm.

ESI-MS (XAN, MeOH): m/z = 473.18 $[M+Na]^+$ expected 473.56

7.6 SYNTHESIS OF GLU-HEMA

The synthesis of Glu-HEMA was based on a literature report.[123] Slight modifications have been made to the purification via column chromatography to improve the yield.



DCM was dried over $MgSO_4$ and filtered before use. 2,3,4,6-Tetra-O-acetyl-α-D-glucopyranosylbromid (acetobromoglucose, Glu-Br, 1.64 g, 4.0 mmol) was dissolved in anhydrous DCM (50 mL) in a flask containing powdered molecular sieve (3 g, 3 Å). The solution was cooled to -40 °C using dry ice in MeCN. Oxygen was removed by purging with N_2 for 10 min. Next, silver triflate (AgOTf, 1.23 g, 4.8 mmol, 1.2 eq) was added, the flask was sealed and the dispersion was stirred for 48 h under continuous cooling (max. temperature -20 °C). Then, the flask was slowly warmed up to RT and all solid residues were filtered off. The solvent was removed *in vacuo* and the residue was dissolved in pyridine (50 mL). Next, acetic anhydride (10 mL) was added and the solution was stirred over night under N_2 atmosphere. Acetic anhydride in combination with pyridine was used to acetylate excess HEMA to facilitate the workup. DCM (50 mL) was added to the dark red solution and the mixture was washed repeatedly with aq. HCl (1 N, 400 mL total) until no more pyridine smell was noticed. The organic phase was washed with sat. $NaHCO_3$ (3 x 50 mL), brine (50 mL), dried over $NaSO_4$ and filtered. The solvent was removed *in vacuo* and the residue was purified by column chromatography (EA:Hex 1:2 v:v, 3rd spot isolated). The pure product (0.66 mg, 1.5 mmol, 37 %, lit. 72 %[123]) was obtained as a viscous oil and stored at -20 °C.

^{1}H-NMR (400 MHz, $CDCl_3$): δ = 1.92 (s, 3H, 3-H_3), 1.97, 1.98, 1.99, 2.06 (4 x s, 4 x 3H, 4 x Ac), 3.66 (ddd, $J_{4',5'}$ 10.3 Hz, $J_{5',6'a}$ 2.5 Hz, $J_{5',6'b}$ 4.3 Hz, 5'-H), 3.79 (ddd, 1H, $J_{6'a,6'b}$ 11.3 Hz, $J_{5'a,6'a}$ 6.8 Hz, $J_{5b,6a}$ 3.3 Hz, 6'-H), 4.02 (ddd, 1H, $J_{6a,6b}$ 11.3 Hz, $J_{5a,6b}$ 4.8 Hz, $J_{5b,6b}$ 4.8 Hz, 6b-H), 4.07-4.13 (m, 1H, 6'-H), 4.19-4.33 (m, 3H, 6'b-H, 5-H_2), 4.54 (d, 1H, $J_{1',2}$ 7.6 Hz), 4.97 (m_c, 1H, 2'-H), 5.05 (t, 1H, 4'-H), 5.17 (t, 1H, 3'-H), 5.56 (m_c, 1H, 1-Hb), 6.09 (m_c, 1H, 1-Ha) ppm.

ESI-MS (MeOH): m/z = 483.4 $[M+Na]^+$ expected 483.4; 499.5 $[M+K]^+$ expected 499.4

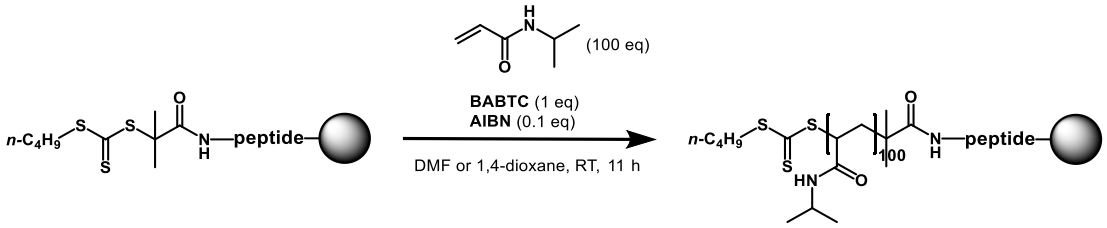
7.7 POLYMERIZATION

7.7.1 SURFACE-INITIATED RAFT POLYMERIZATION

To synthesize polymers carrying the recognition sequence peptide for sortase A, previously synthesized resin-bound CTA-peptides (see chapter 7.5) were used. In theory, a *grafting-from* polymerization would lead to polymers with peptidic end groups linked to the resin. To facilitate chain transfer and analysis of reaction progress, free CTA (shuttle-CTA) was used.[125,131] Two monomers (NIPAM and DMA) and two resin-bound CTAs (BABTC and CPABTC) were tested in this approach. The following description is based on the polymerization of NIPAM with BABTC, the other combinations of monomers and CTAs were realized in the same manner.

Since quantification of resin-bound CTA-peptide was difficult, the basis for calculation was taken from analysis of yield and purity (determined via RP-HPLC) of a batch of CTA-peptide which was cleaved from the resin for analysis.

For a typical experiment, the amount of CTA per resin-batch was estimated to 20 µmol. Based on that, the molar ratios of resin-bound CTA:free CTA:monomer:AIBN were 1:1:100:0.1.



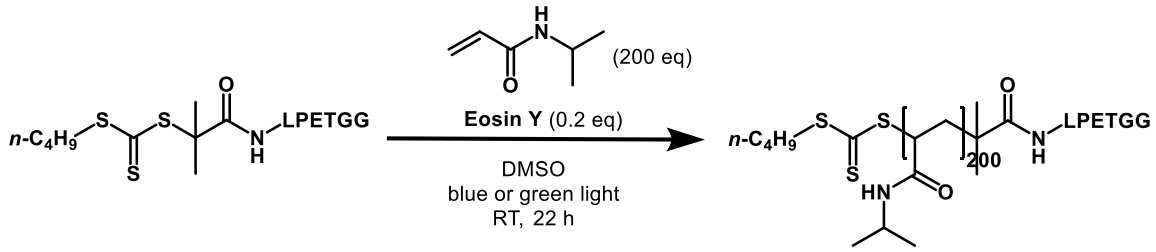
NIPAM (226 mg, 2.0 mmol, 100 eq) was dissolved in freshly distilled 1,4-dioxane or DMF (0.75 mL, 30 % solid content) in a Schlenk tube and the resin-bound CTA-peptide (0.2 mmol of CTA) was added. The mixture was stirred for 10 min at RT to allow resin swelling. Then, AIBN (0.33 mg, 0.2 mmol, 0.1 eq, via stock solution) was added and freeze-pump-thaw (3 x 5 min) was performed. The polymerization was started upon placing the tube in a preheated oil bath (60 °C). The reaction mixture was vigorously stirred throughout the polymerization. A sample was taken after 5 h reaction time to analyze monomer conversion (~70 %) via ^{1}H-NMR spectroscopy. After 11 h the reaction was stopped by removing the oil bath and exposure to air. Resin and supernatant were separated and the supernatant was precipitated into a cold mixture of 5:1 v:v Hex:diethylether. After drying the precipitate was dissolved in acetone and precipitated again, dried and analyzed via ^{1}H-NMR spectroscopy and SEC.

The resin was washed with DMF or 1,4-dioxane (3 x10 mL) and DCM (3 x10 mL). Then, the resin was cleaved using the commonly used cleavage mixture (95:2.5:2.5 TFA:H_2O:TIPS v:v:v, 3 x 1 h,

total 6 mL) and precipitated into the same solvent mixture as the supernatant. The very small amounts of residue were dried and analyzed via ESI-MS.

7.7.2 PET-RAFT POLYMERIZATION

This polymerization approach was carried out with NIPAM and DMA as monomers. The procedure and the amount of Eosin Y is based on literature data.[84]

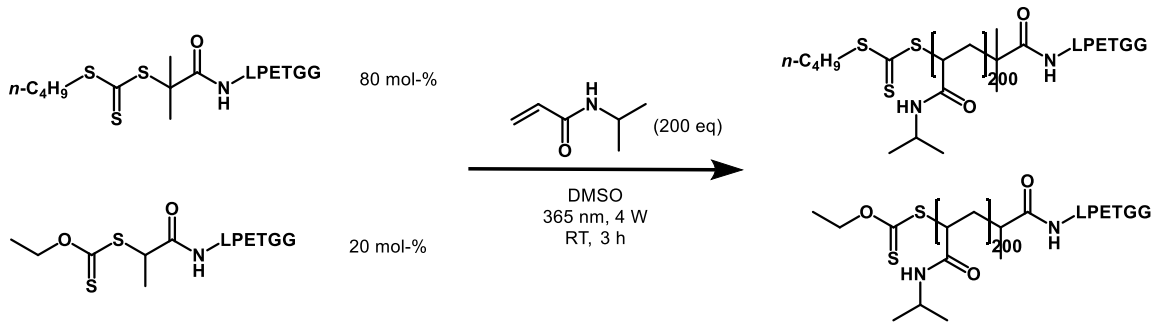


In a typical experiment, BABTC-LPETGG (previously synthesized and cleaved from resin, 4.6 mg, 5.0 µmol) and NIPAM (113.2 mg, 1.0 mmol, 200 eq) were dissolved in DMSO (0.36 mL, 30% solid content) using a Schlenk tube. After turning off all lights, a stock solution of Eosin Y in DMSO (0.06 mg, 0.1 µmol, 0.02 eq) was added. Oxygen was removed by freeze-pump-thaw (3 x 10 min) and the reaction was started by turning on LED strips (green 515-525 nm, 4.8 W/m or blue 460-465 nm 14.4 W/m, both dimmed to half power, reaction setup in Figure 11). The reaction mixture was stirred vigorously for 22 h at RT. Samples for ^{1}H-NMR analysis were taken after 4 h, 16 h and 22 h. After that, the reaction was stopped by turning of the lights and exposure to air. H_2O (4 mL) was added and all solvents were removed by lyophilization. The remaining solid was dissolved in DCM and precipitated in a mixture of Hex and diethylether (5:1, v:v). After drying, the residues were analyzed via ^{1}H-NMR spectroscopy and SEC.

7.7.3 XPI-RAFT POLYMERIZATION

All XPI-RAFT polymerizations were conducted using a method developed by A. Lehnen and coworkers.[87]

The monomers used for XPI-RAFT polymerization were NAM, DMA, NIPAM, HEMA, $OEGA_9$, Glu-HEMA, and SPE. The following procedure describes a typical polymerization of NIPAM, all other monomers were polymerized in the same manner except for SPE (solvent change to TFE). Both types of CTA-peptides (nucleophilic Boc-GG and LPETGG recognition sequence) were used in the same manner for this approach.



BABTC-LPETGG (7.4 mg, 8 µmol, 0.8 eq), XAN-LPETGG (1.7 mg, 2 µmol, 0.2 eq) and NIPAM (226.3 mg, 2 mmol, 200 eq) were dissolved in DMSO (0.7 mL, 30% solid content) in a glass vial (5 mL, Pyrex). The vial was sealed with a rubber septum and the mixture was degassed by purging with N_2 gas for 10 min. Polymerization was started by placing the vial next to a UV-lamp (365 nm, 9.13 mW/cm^2, UVP Hand Lamp, PL Compact UVL-23). Taking samples for ^{1}H-NMR analysis periodically, the reaction was run without stirring for 3 h (95 % monomer conversion). The reaction was stopped by removing the light source and exposure to oxygen. DMSO was removed *in vacuo*, the residue was dissolved in acetone and precipitated in a cold mixture of *n*-hexane:diethylether (5:1, v:v). After drying, the polymers were analyzed via ^{1}H-NMR spectroscopy and SEC.

Precipitation was not possible for the liquid P(OEGA)-based polymers. Instead, dialysis (membrane molecular weight cutoff 1-10 kDa depending on MW of the polymer) against deionized water was performed (4 days, water change every 12-24 h).

7.8 Polymer Deprotection

In order to be used in SML, the side chain protection groups (recognition sequence) and the N-terminal Boc-group (nucleophilic sequence) of the peptides needed to be cleaved. This was achieved by dissolving the polymer in excess TFA for 1 h. Since all cleaved molecules are volatile, precipitation was not necessary. TFA was removed on a Schlenk line equipped with a liquid-N_2 cooling trap, the residue was dissolved in MeCN and benzene (final solvent mixture 50 v-% MeCN, 50 v% benzene) was added. This mixture was lyophilized to yield the deprotected peptide-polymer ready to be used in SML. Confirmation of complete removal of the protection groups was achieved via ^{1}H-NMR analysis and the polymers were analyzed by SEC.

7.9 SORTASE ASSAY

To quickly compare SML performance of different peptide sequences, sortase variants and reaction conditions, small peptides were used and the reaction outcome was analyzed by RP-HPLC. The method used here is based on a literature procedure.[43]

In a typical assay, stock solutions of GGFEF and FLFG-LPETGG-HG in buffer A (see Table 8) were mixed in a micro reaction tube (1.5 mL), and a solution of $NiSO_4$ in buffer A (14.0 µL, 0.2 mM in final reaction solution) was added. To reach the calculated concentrations of reactants, buffer A (300.6 µL) was added. The reaction was started by addition of SrtA-4M and shaking of the reaction tube. The total reaction volume was 350 µL, enabling analysis of up to 7 timepoints. Immediately after addition of sortase, the first 50 µL aliquot (0 h) was taken and quenched by addition of HCl (40 µL, 1 N). The mixture was centrifuged to remove potentially precipitated sortase and the supernatant was analyzed via RP-HPLC. Educt/product ratios were calculated using the area under the respective peaks. More aliquots were taken and quenched after certain reaction times.

Table 8: Volumes and ratios of reactants and sortase used in a typical sortase assay.

	GGFEF	FLFG-LPETGG-HG	SrtA-4M
concentration in reaction solution	0.05 mM	0.05 mM	2 µM
volume of stock solution used	17.5 µL	8.8 µL	9.1 µL
concentration of stock solution	1 mM	2 mM	77 µM
molar ratios	1 eq	1 eq	4 mol-%

Buffer A, sortase reaction buffer: 50 mM TRIS, 150 mM NaCl, 5 mM $CaCl_2$, pH adjusted with conc. HCl to 7.5.

7.10 Protein Expression and Purification

SrtA-WT was provided by Z. Zou (DWI – Leibnitz-Institute for Interactive Materials, expression described in[190]), P. Zou (plasmid, Helmholtz Zentrum München) and Marcus Michaelis (expression, University of Potsdam).

The plasmid for SrtA-4M was provided by P. Zou and the protein was expressed by M. Michaelis following a published procedure.[148]

The different CBM variants (G-CBM and CBM-LPETGG) were expressed by M. Michaelis using a known procedure.[191]

The nanobody Ty1 (exhibiting C-terminal LPETGG-H_6) was provided by S. Petrovic (University of Potsdam) and expressed via a modified a published procedure.[88,156] A pET28a plasmid carrying the DNA for Ty1-LPETGG-H_6 was transformed into E. coli Rosetta 2 (DE3) and a large-scale culture was grown in LB medium at 37 °C to optical density of 0.6. Gene expression was induced with 1 mM isopropyl-β-D-thio-galactopyranoside (IPTG) at 18 °C overnight. After 16 hours the bacterial pellet was collected by centrifugation at 5,000 g for 15 min and frozen at -80 °C. Prior to purification, the bacterial pellet was thawed and resuspended in lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM imidazole, 20 µg/mL DNase I, 1x Roche Complete Protease Inhibitor Cocktail). Bacterial cells were disrupted by four passes through the French Press, and the lysate was centrifuged at 15,000 g for 45 min. The supernatant was loaded on a column containing equilibrated Ni-NTA resin (Qiagen) and washed three times with 5 CV (column volumes) of wash buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 20 mM imidazole). The sample was eluted three times with 5 CV of elution buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 250 mM imidazole), concentrated and loaded on a Superose 6 XK 16 70 column (GE Healthcare) pre-equilibrated with SEC Buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl). Fractions containing pure Ty1-LPETGG-H_6 were pooled and concentrated to ~2.0 mg/ml (see Figure 30). Protein aliquots were flash-frozen in liquid nitrogen and stored at -80 °C.

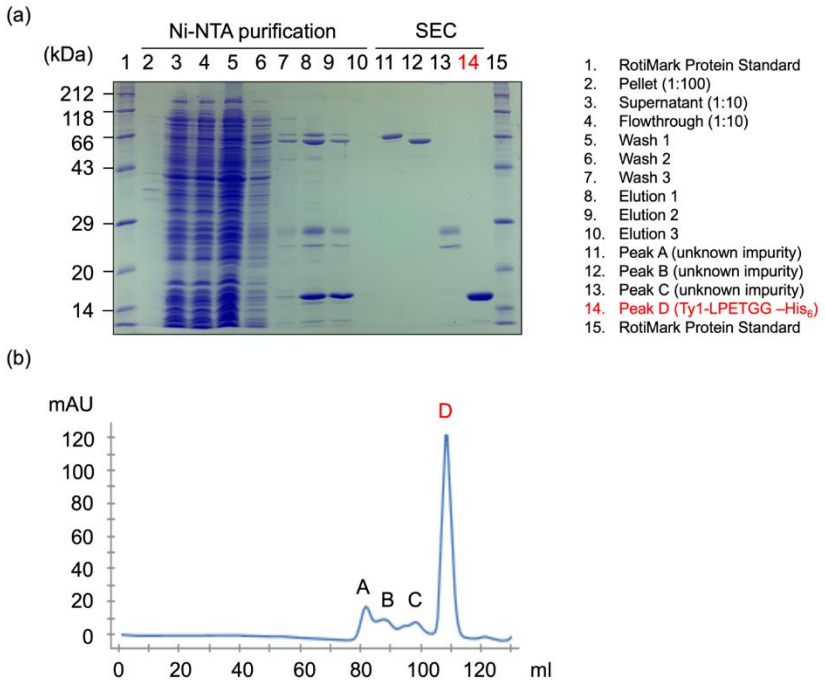


Figure 30: Purification of Ty1-LPETGG-H_6 from *E. coli*. SDS-PAGE of the affinity chromatography (Ni-NTA) and SEC samples (a). SEC profile of the Ty1-LPETGG-H_6 purification on a Superose 6 XK 16 70 column, where Peak D corresponds to fractions containing pure Ty1-LPETGG-H_6 (b). Figure courtesy of S. Petrovic (University of Potsdam).

The amino acid sequences of all proteins used in this thesis are presented below.

G-CBM: GSHMDVKVQY LCENTQTSTQ EIKGKFNIVN TGNRDYSLKD IVLRYYFTKE HNSQLQFICY YTPIGSGNLI PSFGGSGDEH YLQLEFKDVK LPAGGQTGEI QFVIRYADWS FHDQSNDYSF DPTIKAFQDY GKVTLYKNGE LVWGTPPGG

CBM-LPETGG: MDVKVQYLCE NTQTSTQEIK GKFNIVNTGN RDYSLKDIVL RYYFTKEHNS QLQFICYYTP IGSGNLIPSF GGSGDEHYLQ LEFKDVKLPA GGQTGEIQFV IRYADWSFHD QSNDYSFDPT IKAFQDYGKV TLYKNGELVW GTPPGGLPET GGHHHHHH

Ty1-LPETGG: MQVQLVETGG GLVQPGGSLR LSCAASGFTF SSVYMNWVRQ APGKGPEWVS RISPNSGNIG YTDSVKGRFT ISRDNAKNTL YLQMNNLKPE DTALYYCAIG LNLSSSSVRG QGTQVTVSSG GLPETGGHHH HHH

SrtA-WT (after TEV cleavage)**:** SAMAYLFAKP HIDNYLHDKD KDEKIEQYDK NVKEQASKDK KQQAKPQIPK DKSKVAGYIE IPDADIKEPV YPGPATPEQL NRGVSFAEEN ESLDDQNISI AGHTFIDRPN YQFTNLKAAK KGSMVYFKVG NETRKYKMTS IRDVKPTDVG VLDEQKGKDK QLTLITCDDY NEKTGVWEKR KIFVATEVK

SrtA-4M (after TEV-cleavage): SAMAKPQIPK DKSKVAGYIE IPDADIKEPV YPGPATSEQL NRGVSFAEEN ESLDDQNISI AGHTFIDRPN YQFTNLKAAK KGSMVYFKVG NETRKYKMTS IRNVKPTAVG VLDEQKGKDK QLTLITCDDY NEKTGVWETR KIFVATVK

7.11 SORTASE-MEDIATED LIGATION

SML was conducted to ligate two polymers (chapter 5.2) or to synthesize protein-polymer conjugates (chapter 5.3.2). The reaction setup and conditions were similar for all SML. Thus, the following paragraph describes an exemplary synthesis of a protein-polymer conjugate.

A solution of Ty1-LPETGG and GG-P(NAM)$_{20}$ (from stock solutions, see Table 9) and $NiSO_4$ in buffer A (16 µL, 0.2 mM final concentration) was prepared in a 0.5 mL micro reaction tube. To start SML, a solution of SrtA-4M (see Table 9) in buffer A was added, the tube was sealed and shaken thoroughly.

The total reaction volume was 400 µL, enabling analysis of up to 4 timepoints using 100 µL aliquots. Samples were taken until reaction times of up to 17. Immediately after addition of sortase A, the first aliquot (0 h) was taken and quenched by addition of HCl (50 µL, 1 N). Then, the solution was analyzed via SDS-PAGE and MALDI-ToF mass spectrometry.

Table 9: Volumes and ratios of reactants and sortase used in a typical SML reaction to synthesize protein-polymer conjugates.

	GG-P(NAM)$_{20}$	Ty1-LPETGG	SrtA-4M
concentration in reaction solution	0.04 mM	0.04 mM	5 µM
volume of stock solution used	32.0 µL	94.1 µL	26.0 µL
concentration of stock solution	0.5 mM	0.2 mM	77 µM
molar ratios	1 eq	1 eq	10 mol-%

8 REFERENCES

(1) Benjamin Leader, Quentin J. Baca, David E. Golan. Protein therapeutics: a summary and pharmacological classification. *Nat. Rev. Drug Discovery* **2008**, 21–39.

(2) Reichert, J. M. Trends in development and approval times for new therapeutics in the United States. *Nat. Rev. Drug Discovery* **2003**, *2*, 695–702.

(3) Banting, F. G., Best, C. H., Collip, J. B., Campbell, W. R. A. A. Pancreatic extracts in the treatment of diabetes mellitus: preliminary report. *J. Assoc. Med. Can. CMAJ* **1922**, 1281–1286.

(4) Lagassé, H. A. D.; Alexaki, A.; Simhadri, V. L.; Katagiri, N. H.; Jankowski, W.; Sauna, Z. E.; Kimchi-Sarfaty, C. Recent advances in (therapeutic protein) drug development. *F1000Research* **2017**, *6*, 113.

(5) La Torre, B. G. de; Albericio, F. The Pharmaceutical Industry in 2022: An Analysis of FDA Drug Approvals from the Perspective of Molecules. *Molecules* **2023**, *28*.

(6) Abuchowski, A.; van Es, T.; Palczuk, N. C.; Davis, F. F. Alteration of immunological properties of bovine serum albumin by covalent attachment of polyethylene glycol. *J. Biol. Chem.* **1977**, *252*, 3578–3581.

(7) Caliceti, P.; Veronese, F. M. Pharmacokinetic and biodistribution properties of poly(ethylene glycol)-protein conjugates. *Adv. Drug Delivery Rev.* **2003**, *55*, 1261–1277.

(8) Rodríguez-Martínez, J. A.; Solá, R. J.; Castillo, B.; Cintrón-Colón, H. R.; Rivera-Rivera, I.; Barletta, G.; Griebenow, K. Stabilization of alpha-chymotrypsin upon PEGylation correlates with reduced structural dynamics. *Biotechn. and Bioeng.* **2008**, *101*, 1142–1149.

(9) Cummings, C. S.; Campbell, A. S.; Baker, S. L.; Carmali, S.; Murata, H.; Russell, A. J. Design of Stomach Acid-Stable and Mucin-Binding Enzyme Polymer Conjugates. *Biomacromolecules* **2017**, *18*, 576–586.

(10) Peters BG, Goeckner BJ, Ponzillo JJ, Velasquez WS , Wilson AL. Pegaspargase versus asparaginase in adult ALL: a pharmacoeconomic assessment. *Formulary* **1995**, 388–393.

(11) Kaupbayeva, B.; Russell, A. J. Polymer-enhanced biomacromolecules. *Prog. Polym. Sci.* **2019**, 101194.

(12) Wright, T. A.; Page, R. C.; Konkolewicz, D. Polymer conjugation of proteins as a synthetic post-translational modification to impact their stability and activity. *Polym. Chem.* **2019**, *10*, 434–454.

(13) Hoang Thi, T. T.; Pilkington, E. H.; Nguyen, D. H.; Lee, J. S.; Park, K. D.; Truong, N. P. The Importance of Poly(ethylene glycol) Alternatives for Overcoming PEG Immunogenicity in Drug Delivery and Bioconjugation. *Polymers* **2020**, *12*.

(14) Moreno, A.; Pitoc, G. A.; Ganson, N. J.; Layzer, J. M.; Hershfield, M. S.; Tarantal, A. F.; Sullenger, B. A. Anti-PEG Antibodies Inhibit the Anticoagulant Activity of PEGylated Aptamers. *Cell Chem. Biol.* **2019**, *26*, 634-644.e3.

(15) Barz, M.; Luxenhofer, R.; Zentel, R.; Vicent, M. J. Overcoming the PEG-addiction: well-defined alternatives to PEG, from structure–property relationships to better defined therapeutics. *Polym. Chem.* **2011**, *2*, 1900.

(16) Morgenstern, J.; Gil Alvaradejo, G.; Bluthardt, N.; Beloqui, A.; Delaittre, G.; Hubbuch, J. Impact of Polymer Bioconjugation on Protein Stability and Activity Investigated with Discrete Conjugates: Alternatives to PEGylation. *Biomacromolecules* **2018**, *19*, 4250–4262.

(17) Pelegri-O'Day, E. M.; Lin, E.-W.; Maynard, H. D. Therapeutic protein-polymer conjugates: advancing beyond PEGylation. *J. Am. Chem. Soc.* **2014**, *136*, 14323–14332.

(18) Qi, Y.; Simakova, A.; Ganson, N. J.; Li, X.; Luginbuhl, K. M.; Özer, I.; Liu, W.; Hershfield, M. S.; Matyjaszewski, K.; Chilkoti, A. A brush-polymer conjugate of exendin-4 reduces blood glucose for up to five days and eliminates poly(ethylene glycol) antigenicity. *Nat. Biomed. Eng.* **2016**, *1*.

(19) Messina, M. S.; Messina, K. M.; Bhattacharya, A.; Montgomery, H. R.; Maynard, H. D. Preparation of biomolecule-polymer conjugates by grafting-from using ATRP, RAFT, or ROMP. *Prog. Polym. Sci.* **2020**, *100*, 101186.

(20) Liu, B.; Ianosi-Irimie, M.; Thayumanavan, S. Reversible Click Chemistry for Ultrafast and Quantitative Formation of Protein-Polymer Nanoassembly and Intracellular Protein Delivery. *ACS nano* **2019**, *13*, 9408–9420.

(21) Murata, H.; Carmali, S.; Baker, S. L.; Matyjaszewski, K.; Russell, A. J. Solid-phase synthesis of protein-polymers on reversible immobilization supports. *Nature communications* **2018**, *9*, 845.

(22) Gauthier, M. A.; Klok, H.-A. Polymer–protein conjugates: an enzymatic activity perspective. *Polym. Chem.* **2010**, *1*, 1352.

(23) Treetharnmathurot, B.; Ovartlarnporn, C.; Wungsintaweekul, J.; Duncan, R.; Wiwattanapatapee, R. Effect of PEG molecular weight and linking chemistry on the biological activity and thermal stability of PEGylated trypsin. *Int. J. Pharm.* **2008**, *357*, 252–259.

(24) Ekladious, I.; Colson, Y. L.; Grinstaff, M. W. Polymer-drug conjugate therapeutics: advances, insights and prospects. *Nat. Rev. Drug Discovery* **2019**, *18*, 273–294.

(25) Hu, J.; Wang, G.; Zhao, W.; Liu, X.; Zhang, L.; Gao, W. Site-specific in situ growth of an interferon-polymer conjugate that outperforms PEGASYS in cancer therapy. *Biomaterials* **2016**, *96*, 84–92.

(26) Popp, M. W.; Dougan, S. K.; Chuang, T.-Y.; Spooner, E.; Ploegh, H. L. Sortase-catalyzed transformations that improve the properties of cytokines. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 3169–3174.

(27) Suguri, T.; Olsen, B. D. Topology effects on protein–polymer block copolymer self-assembly. *Polym. Chem.* **2019**, *10*, 1751–1761.

(28) Reed, S. A.; Brzovic, D. A.; Takasaki, S. S.; Boyko, K. V.; Antos, J. M. Efficient Sortase-Mediated Ligation Using a Common C-Terminal Fusion Tag. *Bioconjugate Chem.* **2020**, *31*, 1463–1473.

(29) Ashutosh Chilkoti, Guohua Chen, Patrick S. Stayton, and Allan S. Hoffman. Site-Specific Conjugation of a Temperature-Sensitive Polymer to a Genetically-Engineered Protein. *Bioconjugate Chem.* **1994**, 504–507.

(30) Boyer, C.; Bulmus, V.; Liu, J.; Davis, T. P.; Stenzel, M. H.; Barner-Kowollik, C. Well-defined protein-polymer conjugates via in situ RAFT polymerization. *J. Am. Chem. Soc.* **2007**, *129*, 7145–7154.

(31) Pelosi, C.; Duce, C.; Wurm, F. R.; Tinè, M. R. Effect of Polymer Hydrophilicity and Molar Mass on the Properties of the Protein in Protein-Polymer Conjugates: The Case of PPEylated Myoglobin. *Biomacromolecules* **2021**, *22*, 1932–1943.

(32) Mathieu-Gaedke, M. Grafting-to and Grafting-from Proteins - Synthesis and Characterization of Protein-Polymer Conjugates on the Way to Biohybrid Membrane Materials. Dissertation, University of Potsdam, Germany, 2021.

(33) Shi, H.; Shi, Q.; Oswald, J. T.; Gao, Y.; Li, L.; Li, Y. Site-specific PEGylation of Human Growth Hormone by Mutated Sortase A. *Chem. Res. Chin. Univ.* **2018**, *34*, 428–433.

(34) Mathieu‐Gaedke, M.; Böker, A.; Glebe, U. How to Characterize the Protein Structure and Polymer Conformation in Protein‐Polymer Conjugates – a Perspective. *Macromol. Chem. Phys.* **2023**, *224*, 2200353.

(35) Fonzé, E.; Vermeire, M.; Nguyen-Distèche, M.; Brasseur, R.; Charlier, P. The crystal structure of a penicilloyl-serine transferase of intermediate penicillin sensitivity. The DD-transpeptidase of streptomyces K15. *J. Biol. Chem.* **1999**, *274*, 21853–21860.

(36) Spirig, T.; Weiner, E. M.; Clubb, R. T. Sortase enzymes in Gram-positive bacteria. *Mol. Microbiol.* **2011**, *82*, 1044–1059.

(37) Bradshaw, W. J.; Davies, A. H.; Chambers, C. J.; Roberts, A. K.; Shone, C. C.; Acharya, K. R. Molecular features of the sortase enzyme family. *The FEBS journal* **2015**, *282*, 2097–2114.

(38) Clancy, K. W.; Melvin, J. A.; McCafferty, D. G. Sortase transpeptidases: insights into mechanism, substrate specificity, and inhibition. *Biopolymers* **2010**, *94*, 385–396.

(39) Schneider, T.; Senn, M. M.; Berger-Bächi, B.; Tossi, A.; Sahl, H.-G.; Wiedemann, I. In vitro assembly of a complete, pentaglycine interpeptide bridge containing cell wall precursor (lipid II-Gly5) of Staphylococcus aureus. *Mol. Microbiol.* **2004**, *53*, 675–685.

(40) Vollmer, W.; Blanot, D.; Pedro, M. A. de. Peptidoglycan structure and architecture. *FEMS Microbiol. Rev.* **2008**, *32*, 149–167.

(41) Dai, X.; Böker, A.; Glebe, U. Broadening the scope of sortagging. *RSC Adv.* **2019**, *9*, 4700–4721.

(42) Morgan, H. E.; Turnbull, W. B.; Webb, M. E. Challenges in the use of sortase and other peptide ligases for site-specific protein modification. *Chem. Soc. Rev.* **2022**, *51*, 4121–4145.

(43) Aulabaugh, A.; Ding, W.; Kapoor, B.; Tabei, K.; Alksne, L.; Dushin, R.; Zatz, T.; Ellestad, G.; Huang, X. Development of an HPLC assay for Staphylococcus aureus sortase: evidence for the formation of the kinetically competent acyl enzyme intermediate. *Anal. Biochem.* **2007**, *360*, 14–22.

(44) Baer, S.; Nigro, J.; Madej, M. P.; Nisbet, R. M.; Suryadinata, R.; Coia, G.; Hong, L. P. T.; Adams, T. E.; Williams, C. C.; Nuttall, S. D. Comparison of alternative nucleophiles for Sortase

A-mediated bioconjugation and application in neuronal cell labelling. *Org. Biomol. Chem.* **2014**, *12*, 2675–2685.

(45) Antos, J. M.; Truttmann, M. C.; Ploegh, H. L. Recent advances in sortase-catalyzed ligation methodology. *Curr. Opin. Struct. Biol.* **2016**, *38*, 111–118.

(46) Ritzefeld, M. Sortagging: a robust and efficient chemoenzymatic ligation strategy. *Chem. - Eur. J.* **2014**, *20*, 8516–8529.

(47) Chen, I.; Dorr, B. M.; Liu, D. R. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 11399–11404.

(48) Hirakawa, H.; Ishikawa, S.; Nagamune, T. Design of Ca2+-independent Staphylococcus aureus sortase A mutants. *Biotechnol. Bioeng.* **2012**, *109*, 2955–2961.

(49) Freund, C.; Schwarzer, D. Engineered Sortases in Peptide and Protein Chemistry. *Chembiochem* **2021**, *22*, 1347–1356.

(50) Huang, X.; Aulabaugh, A.; Ding, W.; Kapoor, B.; Alksne, L.; Tabei, K.; Ellestad, G. Kinetic mechanism of Staphylococcus aureus sortase SrtA. *Biochemistry* **2003**, *42*, 11307–11315.

(51) Yamamura, Y.; Hirakawa, H.; Yamaguchi, S.; Nagamune, T. Enhancement of sortase A-mediated protein ligation by inducing a β-hairpin structure around the ligation site. *Chem. Commun. (Cambridge, U. K.)* **2011**, *47*, 4742–4744.

(52) David Row, R.; Roark, T. J.; Philip, M. C.; Perkins, L. L.; Antos, J. M. Enhancing the efficiency of sortase-mediated ligations through nickel-peptide complex formation. *Chem. Commun. (Cambridge, U. K.)* **2015**, *51*, 12548–12551.

(53) Goodman, M.; Cai, W.; Smith, N. D. The bold legacy of Emil Fischer. *J. Pept. Sci.* **2003**, *9*, 594–603.

(54) B. Marglin and R. B. Merrifield. The Synthesis of Bovine Insulin by the Solid Phase Method. *J. Am. Chem. Soc.* **1966**, 5051.

(55) Coin, I.; Beyermann, M.; Bienert, M. Solid-phase peptide synthesis: from standard procedures to the synthesis of difficult sequences. *Nat. Protoc.* **2007**, *2*, 3247–3256.

(56) Kulkarni, S. S.; Sayers, J.; Premdjee, B.; Payne, R. J. Rapid and efficient protein synthesis through expansion of the native chemical ligation concept. *Nat. Rev. Chem.* **2018**, *2*.

(57) Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, 2149-2154.

(58) Louis A. Carpino, Grace Y. Han. 9-Fluorenylmethoxycarbonyl function, a new base-sensitive amino-protecting group. *J. Am. Chem. Soc.* **1970**, 5748–5749.

(59) Han, S.-Y.; Kim, Y.-A. Recent development of peptide coupling reagents in organic synthesis. *Tetrahedron* **2004**, *60*, 2447–2467.

(60) Daniel A. Pearson, Mary Blanchette, Mary Lou Baker, Cathy A. Guindon. Trialkylsilanes as scavengers for the trifluoroacetic acid deblocking of protecting groups in peptide synthesis. *Tetrahedron Lett.* **1989**, 2739–2742.

(61) Corrigan, N.; Jung, K.; Moad, G.; Hawker, C. J.; Matyjaszewski, K.; Boyer, C. Reversible-deactivation radical polymerization (Controlled/living radical polymerization): From discovery to materials design and applications. *Prog. Polym. Sci.* **2020**, *111*, 101311.

(62) Braunecker, W. A.; Matyjaszewski, K. Controlled/living radical polymerization: Features, developments, and perspectives. *Prog. Polym. Sci.* **2007**, *32*, 93–146.

(63) Parkatzidis, K.; Wang, H. S.; Truong, N. P.; Anastasaki, A. Recent Developments and Future Challenges in Controlled Radical Polymerization: A 2020 Update. *Chem* **2020**, *6*, 1575–1588.

(64) Nghia P. Truong, Glen R. Jones, Kate G. E. Bradford, Dominik Konkolewicz, Athina Anastasaki. A comparison of RAFT and ATRP methods for controlled radical polymerization. *Nat. Rev. Chem.* **2021**, 859–869.

(65) Matyjaszewski, K. Atom Transfer Radical Polymerization (ATRP): Current Status and Future Perspectives. *Macromolecules* **2012**, *45*, 4015–4039.

(66) Tang, W.; Kwak, Y.; Braunecker, W.; Tsarevsky, N. V.; Coote, M. L.; Matyjaszewski, K. Understanding atom transfer radical polymerization: effect of ligand and initiator structures on the equilibrium constants. *J. Am. Chem. Soc.* **2008**, *130*, 10702–10713.

(67) Fetzer, L.; Toniazzo, V.; Ruch, D.; di Lena, F. Transition-Metal Catalysts for Controlled Radical Polymerization: A First Update. *Isr. J. Chem.* **2012**, *52*, 221–229.

(68) Tang, W.; Matyjaszewski, K. Effect of Ligand Structure on Activation Rate Constants in ATRP. *Macromolecules* **2006**, *39*, 4953–4959.

(69) M. Teodorescu; Krysztof Matyjaszewski. Controlled polymerization of (meth)acrylamides by atom transfer radical polymerization. *Macromol. Rapid Commun.* **2000**, 190–194.

(70) John Chiefari, Y. K. (Bill) Chong, Frances Ercole, Julia Krstina, Justine Jeffery, Tam P. T. Le, Roshan T. A. Mayadunne, Gordon F. Meijs, Catherine L. Moad, Graeme Moad, Ezio Rizzardo and San H. Thang. Living Free-Radical Polymerization by Reversible Addition-Fragmentation Chain Transfer: The RAFT Process. *Macromolecules* **1998**, 5559–5562.

(71) Perrier, S. 50th Anniversary Perspective : RAFT Polymerization—A User Guide. *Macromolecules* **2017**, *50*, 7433–7447.

(72) Keddie, D. J.; Moad, G.; Rizzardo, E.; Thang, S. H. RAFT Agent Design and Synthesis. *Macromolecules* **2012**, *45*, 5321–5342.

(73) Derboven, P.; van Steenberge, P. H. M.; Reyniers, M.-F.; Barner-Kowollik, C.; D'hooge, D. R.; Marin, G. B. Chain Transfer in Degenerative RAFT Polymerization Revisited: A Comparative Study of Literature Methods. *Macromol. Theory Simul.* **2016**, *25*, 104–115.

(74) Hartlieb, M. Photo-Iniferter RAFT Polymerization. *Macromol. Rapid Commun.* **2022**, *43*, 2100514.

(75) Zhao, Y.; Perrier, S. Reversible Addition-Fragmentation Chain Transfer Polymerization from Surfaces. *Adv. Polym. Sci.* **2015**, 77–106.

(76) Ranjan, R.; Brittain, W. J. Combination of Living Radical Polymerization and Click Chemistry for Surface Modification. *Macromolecules* **2007**, *40*, 6217–6223.

(77) Ranjan, R.; Brittain, W. J. Tandem RAFT Polymerization and Click Chemistry: An Efficient Approach to Surface Modification. *Macromol. Rapid Commun.* **2007**, *28*, 2084–2089.

(78) Ranjan, R.; Brittain, W. J. Synthesis of High Density Polymer Brushes on Nanoparticles by Combined RAFT Polymerization and Click Chemistry. *Macromol. Rapid Commun.* **2008**, *29*, 1104–1110.

(79) Cash, B. M.; Wang, L.; Benicewicz, B. C. The preparation and characterization of carboxylic acid-coated silica nanoparticles. *J. Polym. Sci. A Polym. Chem.* **2012**, *50*, 2533–2540.

(80) Bellotti, V.; Simonutti, R. New Light in Polymer Science: Photoinduced Reversible Addition-Fragmentation Chain Transfer Polymerization (PET-RAFT) as Innovative Strategy for the Synthesis of Advanced Materials. *Polymers* **2021**, *13*.

(81) Takayuki Otsu, M. Y. Role of initiator-transfer agent-terminator (iniferter) in radical polymerizations: Polymer design by organic disulfides as iniferters. *Macromol. Rapid Commun.* **1982**, 127–132.

(82) Bray, C.; Li, G.; Postma, A.; Strover, L. T.; Wang, J.; Moad, G. Initiation of RAFT Polymerization: Electrochemically Initiated RAFT Polymerization in Emulsion (Emulsion eRAFT), and Direct PhotoRAFT Polymerization of Liquid Crystalline Monomers. *Aust. J. Chem.* **2021**, *74*, 56.

(83) Xu, J.; Jung, K.; Atme, A.; Shanmugam, S.; Boyer, C. A robust and versatile photoinduced living polymerization of conjugated and unconjugated monomers and its oxygen tolerance. *J. Am. Chem. Soc.* **2014**, *136*, 5508–5519.

(84) Xu, J.; Shanmugam, S.; Duong, H. T.; Boyer, C. Organo-photocatalysts for photoinduced electron transfer-reversible addition–fragmentation chain transfer (PET-RAFT) polymerization. *Polym. Chem.* **2015**, *6*, 5615–5624.

(85) Li, M.; Fromel, M.; Ranaweera, D.; Rocha, S.; Boyer, C.; Pester, C. W. SI-PET-RAFT: Surface-Initiated Photoinduced Electron Transfer-Reversible Addition–Fragmentation Chain Transfer Polymerization. *ACS Macro Lett.* **2019**, *8*, 374–380.

(86) Lewis, R. W.; Evans, R. A.; Malic, N.; Saito, K.; Cameron, N. R. Ultra-fast aqueous polymerisation of acrylamides by high power visible light direct photoactivation RAFT polymerisation. *Polym. Chem.* **2018**, *9*, 60–68.

(87) Lehnen, A.-C.; Gurke, J.; Bapolisi, A. M.; Reifarth, M.; Bekir, M.; Hartlieb, M. Xanthate-supported photo-iniferter (XPI)-RAFT polymerization: facile and rapid access to complex macromolecules. *Chem. Sci.* **2023**, *14*, 593–603.

(88) Moliner-Morro, A.; J. Sheward, D.; Karl, V.; Perez Vidakovics, L.; Murrell, B.; McInerney, G. M.; Hanke, L. Picomolar SARS-CoV-2 Neutralization Using Multi-Arm PEG Nanobody Constructs. *Biomolecules* **2020**, *10*, 1661.

(89) Qu, Z.; Krishnamurthy, V.; Haller, C. A.; Dorr, B. M.; Marzec, U. M.; Hurst, S.; Hinds, M. T.; Hanson, S. R.; Liu, D. R.; Chaikof, E. L. Immobilization of actively thromboresistant assemblies on sterile blood-contacting surfaces. *Adv. Healthcare Mater.* **2014**, *3*, 30–35.

(90) Dehn, S.; Chapman, R.; Jolliffe, K. A.; Perrier, S. Synthetic Strategies for the Design of Peptide/Polymer Conjugates. *Polym. Rev.* **2011**, *51*, 214–234.

(91) Cate, M. G. J. ten; Börner, H. G. Synthesis of ABC-Triblock Peptide-Polymer Conjugates for the Positioning of Peptide Segments within Block Copolymer Aggregates. *Macromol. Chem. Phys.* **2007**, *208*, 1437–1446.

(92) Cate, M. G. J. ten; Rettig, H.; Bernhardt, K.; Börner, H. G. Sequence-Defined Polypeptide−Polymer Conjugates Utilizing Reversible Addition Fragmentation Transfer Radical Polymerization. *Macromolecules* **2005**, *38*, 10643–10649.

(93) Hentschel, J.; Bleek, K.; Ernst, O.; Lutz, J.-F.; Börner, H. G. Easy Access to Bioactive Peptide−Polymer Conjugates via RAFT. *Macromolecules* **2008**, *41*, 1073–1075.

(94) Lutz, J.-F.; Börner, H. G. Modern trends in polymer bioconjugates design. *Prog. Polym. Sci.* **2008**, *33*, 1–39.

(95) Adam, S. HBTU: a mild activating agent of muramic acid. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 571–574.

(96) Cayot, P.; Tainturier, G. The quantification of protein amino groups by the trinitrobenzenesulfonic acid method: a reexamination. *Anal. Biochem.* **1997**, *249*, 184–200.

(97) Litou, Z. I.; Bagos, P. G.; Tsirigos, K. D.; Liakopoulos, T. D.; Hamodrakas, S. J. Prediction of cell wall sorting signals in gram-positive bacteria with a hidden markov model: application to complete genomes. *J. Bioinf. Comput. Biol.* **2008**, *6*, 387–401.

(98) Jos Boekhorst; Mark W. H. J. de Been; Michiel Kleerebezem; Roland J. Siezen. Genome-Wide Detection and Analysis of Cell Wall-Bound Proteins with LPxTG-Like Sorting Motifs. *J. Bacteriol.* **2005**, 4928–4934.

(99) Antos, J. M.; Ingram, J.; Fang, T.; Pishesha, N.; Truttmann, M. C.; Ploegh, H. L. Site-Specific Protein Labeling via Sortase-Mediated Transpeptidation. *Curr. Protoc. Protein Sci.* **2017**, *89*, 15.3.1-15.3.19.

(100) Williamson, D. J.; Fascione, M. A.; Webb, M. E.; Turnbull, W. B. Efficient N-terminal labeling of proteins by use of sortase. *Angew. Chem., Int. Ed. Engl.* **2012**, *51*, 9377–9380.

(101) Liu, F.; Luo, E. Y.; Flora, D. B.; Mezo, A. R. Irreversible sortase A-mediated ligation driven by diketopiperazine formation. *J. Org. Chem.* **2014**, *79*, 487–492.

(102) Theile, C. S.; Witte, M. D.; Blom, A. E. M.; Kundrat, L.; Ploegh, H. L.; Guimaraes, C. P. Site-specific N-terminal labeling of proteins using sortase-mediated reactions. *Nat. Prot.* **2013**, *8*, 1800–1807.

(103) Xiaolin Dai. Synthesis of Artificial Building Blocks for Sortase-Mediated Ligation and Their Enzymatic Linkage. Dissertation, University of Potsdam, Germany, 2019.

(104) Sigma-Aldrich Co. LLC. Controlled Radical Polymerization Guide. https://www.sigmaaldrich.com/deepweb/assets/sigmaaldrich/marketing/global/documents/716/722/crp-guide-br5077en-mk.pdf (accessed September 29, 2023).

(105) Larnaudie, S. C.; Brendel, J. C.; Jolliffe, K. A.; Perrier, S. Cyclic peptide-polymer conjugates: Grafting-to vs grafting-from. *J. Polym. Sci. Part A: Polym. Chem.* **2016**, *54*, 1003–1011.

(106) Lehnen, A.-C.; Kurki, J. A. M.; Hartlieb, M. The difference between photo-iniferter and conventional RAFT polymerization: high livingness enables the straightforward synthesis of multiblock copolymers. *Polym. Chem.* **2022**, *13*, 1537–1546.

(107) E. Kaiser, R.L. Colescott, C.D. Bossinger, P.I. Cook. Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal. Biochem.* **1970**, *2*, 595–598.

(108) Jun-Rui Yang, Margaret E. Langmuir. Synthesis and properties of a maleimide fluorescent thiol reagent derived from a naphthopyranone. *J. Heterocyclic Chem.* **1991**, 1177.

(109) Murat Mertoğlu. The Synthesis of Well-Defined Functional Homo- and Block Copolymers in Aqueous Media via Reversible Addition-Fragmentation Chain Transfer (RAFT) Polymerization. Dissertation, University of Potsdam, Germany, 2004.

(110) Chen, C.; Kong, F.; Wei, X.; Thang, S. H. Syntheses and effectiveness of functional peptide-based RAFT agents. *Chem. Commun. (Cambridge, U. K.)* **2017**, *53*, 10776–10779.

(111) Vijlder, T. de; Valkenborg, D.; Lemière, F.; Romijn, E. P.; Laukens, K.; Cuyckens, F. A tutorial in small molecule identification via electrospray ionization-mass spectrometry: The practical art of structural elucidation. *Mass Spectrom. Rev.* **2018**, *37*, 607–629.

(112) Hentschel, J.; Cate, M. G. J. ten; Börner, H. G. Peptide-Guided Organization of Peptide−Polymer Conjugates: Expanding the Approach from Oligo- to Polymers. *Macromolecules* **2007**, *40*, 9224–9232.

(113) Zhao, Y.; Perrier, S. Synthesis of well-defined conjugated copolymers by RAFT polymerization using cysteine and glutathione-based chain transfer agents. *Chem. Commun.* **2007**, *411*, 4294.

(114) Fuchs, A. V.; Thurecht, K. J. Stability of Trithiocarbonate RAFT Agents Containing Both a Cyano and a Carboxylic Acid Functional Group. *ACS Macro Lett.* **2017**, *6*, 287–291.

(115) Glasgow, J. E.; Salit, M. L.; Cochran, J. R. In Vivo Site-Specific Protein Tagging with Diverse Amines Using an Engineered Sortase Variant. *J. Am. Chem. Soc.* **2016**, *138*, 7496–7499.

(116) Dai, X.; Mate, D.; Glebe, U.; Mirzaei Garakani, T.; Körner, A.; Schwaneberg, U.; Böker, A. Sortase-Mediated Ligation of Purely Artificial Building Blocks. *Polymers* **2018**, *10*, 151.

(117) Wayne C. Widdison. Conjugates Comprising Cell-Binding Agents and Cytotoxic Agents PCT/US2014/019434, Sep 4, 2014.

(118) Orlandin, A.; Formaggio, F.; Toffoletti, A.; Peggion, C. Cotton functionalized with peptides: characterization and synthetic methods. *J. Pept. Sci.* **2014**, *20*, 547–553.

(119) Fallah-Mehrjardi, M. Review of the organic trithiocarbonates synthesis. *Monatsh. Chem.* **2018**, *149*, 1931–1944.

(120) Kakwere, H.; Chun, C. K. Y.; Jolliffe, K. A.; Payne, R. J.; Perrier, S. Polymer-peptide chimeras for the multivalent display of immunogenic peptides. *Chem. Commun. (Cambridge, U. K.)* **2010**, *46*, 2188–2190.

(121) Zhang, P.; Sun, F.; Liu, S.; Jiang, S. Anti-PEG antibodies in the clinic: Current issues and beyond PEGylation. *J. Controlled Release* **2016**, *244*, 184–193.

(122) Shih, Y.-J.; Chang, Y. Tunable blood compatibility of polysulfobetaine from controllable molecular-weight dependence of zwitterionic nonfouling nature in aqueous solution. *Langmuir* **2010**, *26*, 17286–17294.

(123) Moira Ambrosi, Andrei S. Batsanov, Neil R. Cameron, Benjamin G. Davis, Judith A. K. Howard and Rob Hunter. Influence of preparation procedure on polymer composition: synthesis and characterisation of polymethacrylates bearing β-D-glucopyranoside and β-D-galactopyranoside residues. *J. Chem. Soc., Perkin Trans. 1* **2002**, 45–52.

(124) Nguyen, D. H.; Wood, M. R.; Zhao, Y.; Perrier, S.; Vana, P. Solid-Supported MADIX Polymerization of Vinyl Acetate. *Macromolecules* **2008**, *41*, 7071–7078.

(125) Perrier, S.; Takolpuckdee, P.; Mars, C. A. Reversible Addition−Fragmentation Chain Transfer Polymerization Mediated by a Solid Supported Chain Transfer Agent. *Macromolecules* **2005**, *38*, 6770–6774.

(126) Trzebicka, B.; Robak, B.; Trzcinska, R.; Szweda, D.; Suder, P.; Silberring, J.; Dworak, A. Thermosensitive PNIPAM-peptide conjugate – Synthesis and aggregation. *Eur. Polym. J.* **2013**, *49*, 499–509.

(127) Becker, M. L.; Liu, J.; Wooley, K. L. Functionalized micellar assemblies prepared via block copolymers synthesized by living free radical polymerization upon peptide-loaded resins. *Biomacromolecules* **2005**, *6*, 220–228.

(128) Ayres, N.; Haddleton, D. M.; Shooter, A. J.; Pears, D. A. Synthesis of Hydrophilic Polar Supports Based on Poly(dimethylacrylamide) via Copper-Mediated Radical Polymerization from a Cross-Linked Polystyrene Surface: Potential Resins for Oligopeptide Solid-Phase Synthesis. *Macromolecules* **2002**, *35*, 3849–3855.

(129) Zheng, Z.; Ling, J.; Müller, A. H. E. Revival of the R-group approach: a "CTA-shuttled" grafting from approach for well-defined cylindrical polymer brushes via RAFT polymerization. *Macromol. Rapid Commun.* **2014**, *35*, 234–241.

(130) Tsujii, Y.; Ejaz, M.; Sato, K.; Goto, A.; Fukuda, T. Mechanism and Kinetics of RAFT-Mediated Graft Polymerization of Styrene on a Solid Surface. 1. Experimental Evidence of Surface Radical Migration. *Macromolecules* **2001**, *34*, 8872–8878.

(131) Takolpuckdee, P.; Mars, C. A.; Perrier, S. Merrifield resin-supported chain transfer agents, precursors for RAFT polymerization. *Org. Lett.* **2005**, *7*, 3449–3452.

(132) Ohno, K.; Ma, Y.; Huang, Y.; Mori, C.; Yahata, Y.; Tsujii, Y.; Maschmeyer, T.; Moraes, J.; Perrier, S. Surface-Initiated Reversible Addition–Fragmentation Chain Transfer (RAFT) Polymerization from Fine Particles Functionalized with Trithiocarbonates. *Macromolecules* **2011**, *44*, 8944–8953.

(133) Mei, Y.; Beers, K. L.; Byrd, H. C. M.; VanderHart, D. L.; Washburn, N. R. Solid-phase ATRP synthesis of peptide-polymer hybrids. *J. Am. Chem. Soc.* **2004**, *126*, 3472–3476.

(134) González-Toro, D. C.; Thayumanavan, S. Advances in Polymer and Polymeric Nanostructures for Protein Conjugation. *Eur. Polym. J.* **2013**, *49*, 2906–2918.

(135) Grover, G. N.; Maynard, H. D. Protein-polymer conjugates: synthetic approaches by controlled radical polymerizations and interesting applications. *Curr. Opin. Chem. Biol.* **2010**, *14*, 818–827.

(136) Allegrezza, M. L.; Konkolewicz, D. PET-RAFT Polymerization: Mechanistic Perspectives for Future Materials. *ACS Macro Lett.* **2021**, *10*, 433–446.

(137) Seo, S. E.; Discekici, E. H.; Zhang, Y.; Bates, C. M.; Hawker, C. J. Surface‐initiated PET‐RAFT polymerization under metal‐free and ambient conditions using enzyme degassing. *J. Polym. Sci.* **2020**, *58*, 70–76.

(138) Zhou, J.; Ye, L.; Lin, Y.; Wang, L.; Zhou, L.; Hu, H.; Zhang, Q.; Yang, H.; Luo, Z. Surface modification PVA hydrogel with zwitterionic via PET‐RAFT to improve the antifouling property. *J Appl Polym Sci* **2019**, *136*, 47653.

(139) Kurek, P. N.; Kloster, A. J.; Weaver, K. A.; Manahan, R.; Allegrezza, M. L.; Alwis Watuthanthrige, N. de; Boyer, C.; Reeves, J. A.; Konkolewicz, D. How Do Reaction and Reactor Conditions Affect Photoinduced Electron/Energy Transfer Reversible Addition–Fragmentation Transfer Polymerization? *Ind. Eng. Chem. Res.* **2018**, *57*, 4203–4213.

(140) Kuzmyn, A. R.; Nguyen, A. T.; Teunissen, L. W.; Zuilhof, H.; Baggerman, J. Antifouling Polymer Brushes via Oxygen-Tolerant Surface-Initiated PET-RAFT. *Langmuir* **2020**, *36*, 4439–4446.

(141) J. Xu, S. Shanmugam, N. A. Corrigan, C. Boyer. Catalyst-Free Visible Light-Induced RAFT Photopolymerization, ACS Symposium Series Vol. 1187. *Controlled Radical Polymerization: Mechanisms* **2015**, 247–267.

(142) McKenzie, T. G.; Da Costa, L. P. M.; Fu, Q.; Dunstan, D. E.; Qiao, G. G. Investigation into the photolytic stability of RAFT agents and the implications for photopolymerization reactions. *Polym. Chem.* **2016**, *7*, 4246–4253.

(143) Fu, Q.; McKenzie, T. G.; Tan, S.; Nam, E.; Qiao, G. G. Tertiary amine catalyzed photo-induced controlled radical polymerization of methacrylates. *Polym. Chem.* **2015**, *6*, 5362–5368.

(144) Alconcel, S. N. S.; Baas, A. S.; Maynard, H. D. FDA-approved poly(ethylene glycol)–protein conjugate drugs. *Polym. Chem.* **2011**, *2*, 1442.

(145) Hildebrand, V.; Laschewsky, A.; Zehm, D. On the hydrophilicity of polyzwitterion poly (N,N-dimethyl-N-(3-(methacrylamido)propyl)ammoniopropane sulfonate) in water, deuterated water, and aqueous salt solutions. *J. Biomater. Sci., Polym. Ed.* **2014**, *25*, 1602–1618.

(146) Michelle, B. H. C.; McEwen, C. N. The limitations of MALDI-TOF mass spectrometry in the analysis of wide polydisperse polymers. *Anal. Chem.* **2000**, *72*, 4568–4576.

(147) Agilent Technologies. A guide to multi-detector gel permeation chromatography. https://www.agilent.com/cs/library/primers/Public/5990-7196EN.pdf (accessed July 19, 2023).

(148) Freiburger, L.; Sonntag, M.; Hennig, J.; Li, J.; Zou, P.; Sattler, M. Efficient segmental isotope labeling of multi-domain proteins using Sortase A. *J. Biomol. NMR* **2015**, *63*, 1–8.

(149) Ton-That, H.; Mazmanian, S. K.; Faull, K. F.; Schneewind, O. Anchoring of surface proteins to the cell wall of Staphylococcus aureus. Sortase catalyzed in vitro transpeptidation reaction using LPXTG peptide and NH(2)-Gly(3) substrates. *J. Biol. Chem.* **2000**, *275*, 9876–9881.

(150) Kruger, R. G.; Dostal, P.; McCafferty, D. G. Development of a high-performance liquid chromatography assay and revision of kinetic parameters for the Staphylococcus aureus sortase transpeptidase SrtA. *Anal. Biochem.* **2004**, *326*, 42–48.

(151) Frankel, B. A.; Kruger, R. G.; Robinson, D. E.; Kelleher, N. L.; McCafferty, D. G. Staphylococcus aureus sortase transpeptidase SrtA: insight into the kinetic mechanism and evidence for a reverse protonation catalytic mechanism. *Biochemistry* **2005**, *44*, 11188–11200.

(152) Zou, Z.; Mate, D. M.; Rübsam, K.; Jakob, F.; Schwaneberg, U. Sortase-Mediated High-Throughput Screening Platform for Directed Enzyme Evolution. *ACS Comb. Sci.* **2018**, *20*, 203–211.

(153) Tran, H. N. T.; Tran, P.; Deuis, J. R.; Agwa, A. J.; Zhang, A. H.; Vetter, I.; Schroeder, C. I. Enzymatic Ligation of a Pore Blocker Toxin and a Gating Modifier Toxin: Creating Double-Knotted Peptides with Improved Sodium Channel NaV1.7 Inhibition. *Bioconjugate Chem.* **2020**, *31*, 64–73.

(154) Crowe, S. O.; Pham, G. H.; Ziegler, J. C.; Deol, K. K.; Guenette, R. G.; Ge, Y.; Strieter, E. R. Subunit-Specific Labeling of Ubiquitin Chains by Using Sortase: Insights into the Selectivity of Deubiquitinases. *Chembiochem* **2016**, *17*, 1525–1531.

(155) Antos, J. M.; Popp, M. W.-L.; Ernst, R.; Chew, G.-L.; Spooner, E.; Ploegh, H. L. A straight path to circular proteins. *J. Biol. Chem.* **2009**, *284*, 16028–16036.

(156) Hanke, L.; Vidakovics Perez, L.; Sheward, D. J.; Das, H.; Schulte, T.; Moliner-Morro, A.; Corcoran, M.; Achour, A.; Karlsson Hedestam, G. B.; Hällberg, B. M.*; et al.* An alpaca nanobody neutralizes SARS-CoV-2 by blocking receptor interaction. *Nat. Commun.* **2020**, *11*, 4420.

(157) Witte, M. D.; Cragnolini, J. J.; Dougan, S. K.; Yoder, N. C.; Popp, M. W.; Ploegh, H. L. Preparation of unnatural N-to-N and C-to-C protein fusions. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 11993–11998.

(158) Arkenberg, M. R.; Lin, C.-C. Orthogonal enzymatic reactions for rapid crosslinking and dynamic tuning of PEG-peptide hydrogels. *Biomat. Sci.* **2017**, *5*, 2231–2240.

(159) Arkenberg, M. R.; Moore, D. M.; Lin, C.-C. Dynamic control of hydrogel crosslinking via sortase-mediated reversible transpeptidation. *Acta Biomater.* **2019**, *83*, 83–95.

(160) Broguiere, N.; Formica, F. A.; Barreto, G.; Zenobi-Wong, M. Sortase A as a cross-linking enzyme in tissue engineering. *Acta Biomater.* **2018**, *77*, 182–190.

(161) Trachsel, L.; Broguiere, N.; Rosenboom, J.-G.; Zenobi-Wong, M.; Benetti, E. M. Enzymatically crosslinked poly(2-alkyl-2-oxazoline) networks for 3D cell culture. *J. Mater. Chem. B* **2018**, *6*, 7568–7572.

(162) Z. Geng, J.J. Shin, Y. Xi, C.J. Hawker. Click chemistry strategies for the accelerated synthesis of functional macromolecules. *J. Polym. Sci.* **2021**, *59*.

(163) Ham, H. O.; Qu, Z.; Haller, C. A.; Dorr, B. M.; Dai, E.; Kim, W.; Liu, D. R.; Chaikof, E. L. In situ regeneration of bioactive coatings enabled by an evolved Staphylococcus aureus sortase A. *Nat. Commun.* **2016**, *7*, 11140.

(164) Yaniv, O.; Petkun, S.; Shimon, L. J. W.; Bayer, E. A.; Lamed, R.; Frolow, F. A single mutation reforms the binding activity of an adhesion-deficient family 3 carbohydrate-binding module. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *68*, 819–828.

(165) Salvador, J.-P.; Vilaplana, L.; Marco, M.-P. Nanobody: outstanding features for diagnostic and therapeutic applications. *Anal. Bioanal. Chem.* **2019**, *411*, 1703–1713.

(166) Block, H.; Maertens, B.; Spriestersbach, A.; Brinker, N.; Kubicek, J.; Fabis, R.; Labahn, J.; Schäfer, F. Immobilized-metal affinity chromatography (IMAC): a review. *Methods Enzymol.* **2009**, *463*, 439–473.

(167) Zheng, C.; Ma, G.; Su, Z. Native PAGE eliminates the problem of PEG-SDS interaction in SDS-PAGE and provides an alternative to HPLC in characterization of protein PEGylation. *Electrophoresis* **2007**, *28*, 2801–2807.

(168) Hodgson, D. J.; Aubin, Y. Assessment of the structure of pegylated-recombinant protein therapeutics by the NMR fingerprint assay. *J. Pharm. Biomed. Anal.* **2017**, *138*, 351–356.

(169) Baker, M. P.; Reynolds, H. M.; Lumicisi, B.; Bryson, C. J. Immunogenicity of protein therapeutics: The key causes, consequences and challenges. *Self/nonself* **2010**, *1*, 314–322.

(170) Broyer, R. M.; Grover, G. N.; Maynard, H. D. Emerging synthetic approaches for protein-polymer conjugations. *Chem. Commun. (Cambridge, U. K.)* **2011**, *47*, 2212–2226.

(171) Nauka, P. C.; Lee, J.; Maynard, H. D. Enhancing Conjugation Yield of Brush Polymer-Protein Conjugates by Increasing Linker Length at the Polymer End-Group. *Polym. Chem.* **2016**, *7*, 2352–2357.

(172) Chen, Q.; Sun, Q.; Molino, N. M.; Wang, S.-W.; Boder, E. T.; Chen, W. Sortase A-mediated multi-functionalization of protein nanoparticles. *Chem. Commun. (Cambridge, U. K.)* **2015**, *51*, 12107–12110.

(173) Angot, S.; Ayres, N.; Bon, S. A. F.; Haddleton, D. M. Living Radical Polymerization Immobilized on Wang Resins: Synthesis and Harvest of Narrow Polydispersity Poly(methacrylate)s. *Macromolecules* **2001**, *34*, 768–774.

(174) Vaino, A. R.; Goodin, D. B.; Janda, K. D. Investigating resins for solid phase organic synthesis: the relationship between swelling and microenvironment as probed by EPR and fluorescence spectroscopy. *J. Comb. Chem.* **2000**, *2*, 330–336.

(175) Vaino, A. R.; Janda, K. D. Solid-phase organic synthesis: a critical understanding of the resin. *J. Comb. Chem.* **2000**, *2*, 579–596.

(176) Virender K. Sarin, Stephen B. H. Kent, R. B, Merrifield. Properties of Swollen Polymer Networks. Solvation and Swelling of Peptide-Containing Resins in Solid-Phase Peptide Synthesis. *J. Am. Chem. Soc.* **1980**, 5463–5470.

(177) Williamson, D. J.; Webb, M. E.; Turnbull, W. B. Depsipeptide substrates for sortase-mediated N-terminal protein ligation. *Nat. Protoc.* **2014**, *9*, 253–262.

(178) Nguyen, G. K. T.; Kam, A.; Loo, S.; Jansson, A. E.; Pan, L. X.; Tam, J. P. Butelase 1: A Versatile Ligase for Peptide and Protein Macrocyclization. *J. Am. Chem. Soc.* **2015**, *137*, 15398–15401.

(179) Nguyen, G. K. T.; Wang, S.; Qiu, Y.; Hemu, X.; Lian, Y.; Tam, J. P. Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nat. Chem. Biol.* **2014**, *10*, 732–738.

(180) Nguyen, G. K. T.; Cao, Y.; Wang, W.; Liu, C. F.; Tam, J. P. Site-Specific N-Terminal Labeling of Peptides and Proteins using Butelase 1 and Thiodepsipeptide. *Angew. Chem., Int. Ed. Engl.* **2015**, *54*, 15694–15698.

(181) Pi, N.; Gao, M.; Cheng, X.; Liu, H.; Kuang, Z.; Yang, Z.; Yang, J.; Zhang, B.; Chen, Y.; Liu, S.*; et al.* Recombinant Butelase-Mediated Cyclization of the p53-Binding Domain of the

Oncoprotein MdmX-Stabilized Protein Conformation as a Promising Model for Structural Investigation. *Biochemistry* **2019**, *58*, 3005–3015.

(182) James, A. M.; Haywood, J.; Leroux, J.; Ignasiak, K.; Elliott, A. G.; Schmidberger, J. W.; Fisher, M. F.; Nonis, S. G.; Fenske, R.; Bond, C. S.*; et al.* The macrocyclizing protease butelase 1 remains autocatalytic and reveals the structural basis for ligase activity. *The Plant journal* **2019**, *98*, 988–999.

(183) Nguyen, G. K. T.; Qiu, Y.; Cao, Y.; Hemu, X.; Liu, C.-F.; Tam, J. P. Butelase-mediated cyclization and ligation of peptides and proteins. *Nat. Protoc.* **2016**, *11*, 1977–1988.

(184) Cao, Y.; Nguyen, G. K. T.; Chuah, S.; Tam, J. P.; Liu, C.-F. Butelase-Mediated Ligation as an Efficient Bioconjugation Method for the Synthesis of Peptide Dendrimers. *Bioconjugate Chem.* **2016**, *27*, 2592–2596.

(185) Nguyen, G. K. T.; Hemu, X.; Quek, J.-P.; Tam, J. P. Butelase-Mediated Macrocyclization of d-Amino-Acid-Containing Peptides. *Angew. Chem., Int. Ed. Engl.* **2016**, *55*, 12802–12806.

(186) Harmand, T. J.; Bousbaine, D.; Chan, A.; Zhang, X.; Liu, D. R.; Tam, J. P.; Ploegh, H. L. One-Pot Dual Labeling of IgG 1 and Preparation of C-to-C Fusion Proteins Through a Combination of Sortase A and Butelase 1. *Bioconjugate Chem.* **2018**, *29*, 3245–3249.

(187) Schägger, H. Tricine-SDS-PAGE. *Nat. Protoc.* **2006**, *1*, 16–22.

(188) C. Fernandez-Patron, L. Castellanos-Serra, P. Rodriguez. Reverse staining of sodium dodecyl sulfate polyacrylamide gels by imidazole-zinc salts: sensitive detection of unmodified proteins. *Biotechniques* **1992**, 564–573.

(189) Bray, C.; Peltier, R.; Kim, H.; Mastrangelo, A.; Perrier, S. Anionic multiblock core cross-linked star copolymers via RAFT polymerization. *Polym. Chem.* **2017**, *8*, 5513–5524.

(190) Zou, Z.; Alibiglou, H.; Mate, D. M.; Davari, M. D.; Jakob, F.; Schwaneberg, U. Directed sortase A evolution for efficient site-specific bioconjugations in organic co-solvents. *Chem. Commun. (Cambridge, U. K.)* **2018**, *54*, 11467–11470.

(191) Jindou, S.; Petkun, S.; Shimon, L.; Bayer, E. A.; Lamed, R.; Frolow, F. Crystallization and preliminary diffraction studies of CBM3b of cellobiohydrolase 9A from Clostridium thermocellum. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2007**, *63*, 1044–1047.

9 APPENDIX

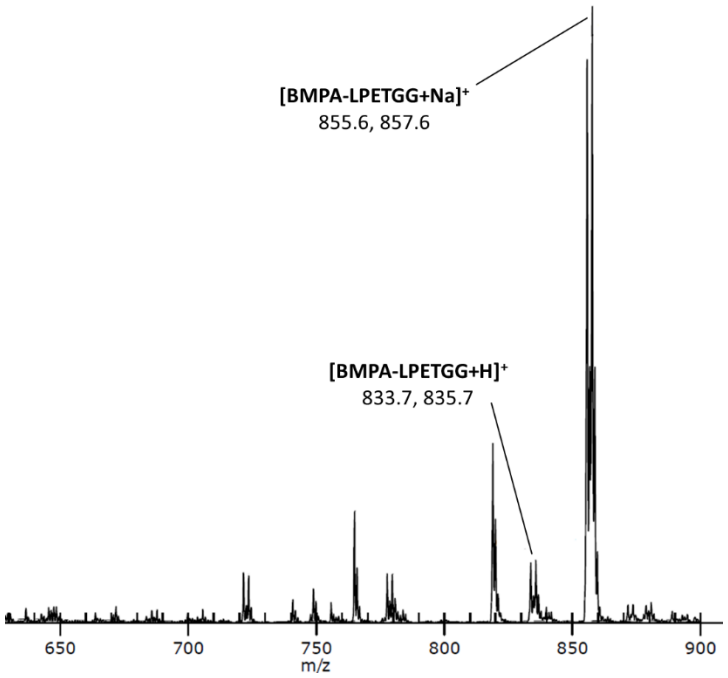


Figure 31: ESI-MS spectrum of BMPA-LPETGG. Expected m/z $[M+H]^+$ 833.6 and 835.6; $[M+Na]^+$ 855.7 and 857.7 (two peaks with similar intensities due to different Br-isotopes).

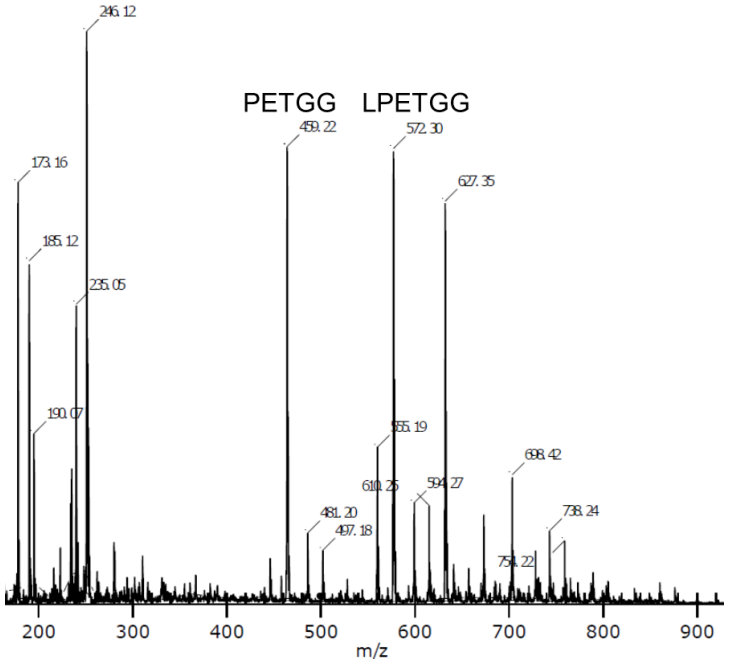


Figure 32: ESI-MS spectrum of unsuccessful CTA coupling after cleavage with TFA/H_2O/TIPS, expected m/z 806.27.

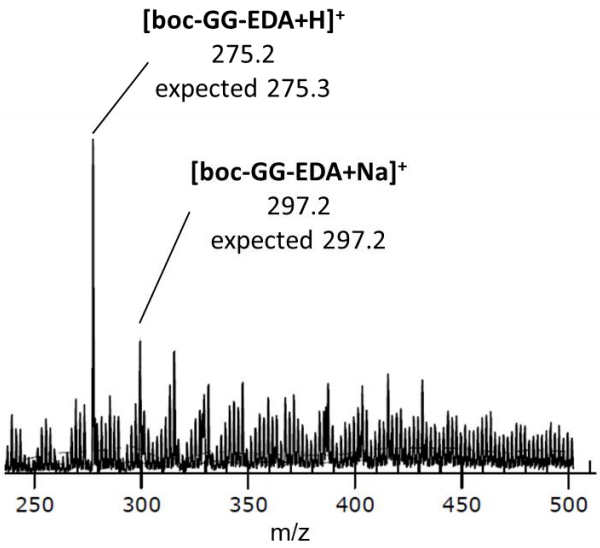


Figure 33: ESI-MS spectrum of Boc-GG-EDA used in synthesis of Boc-GG-BABTC.

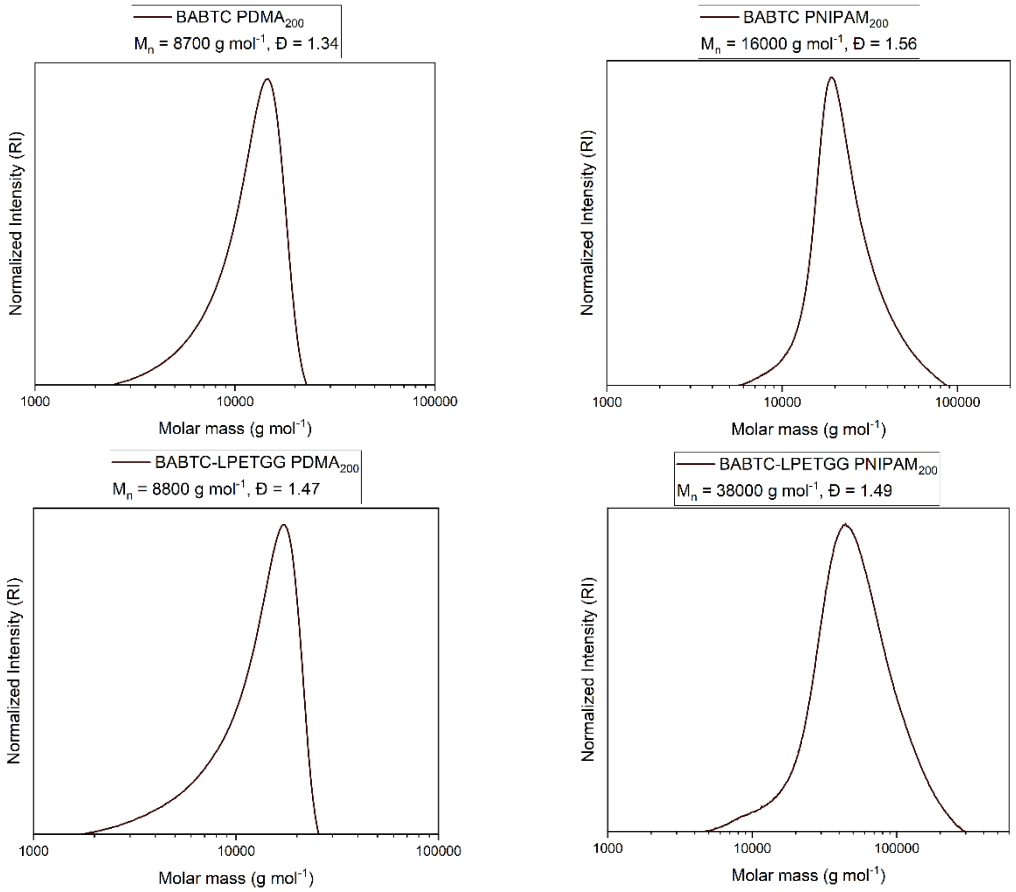
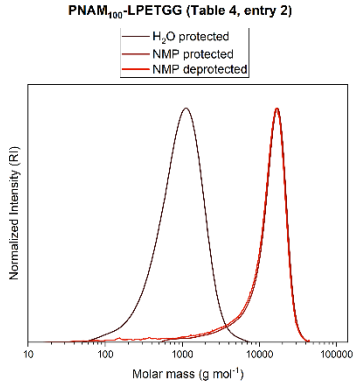
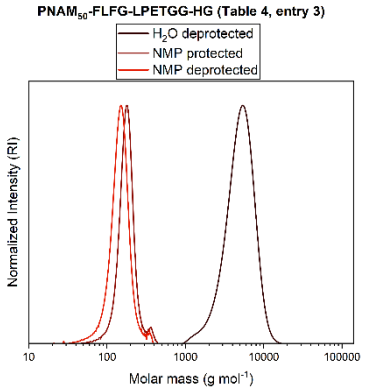
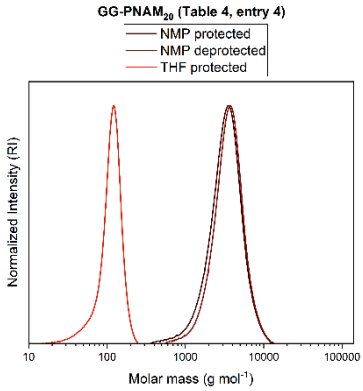
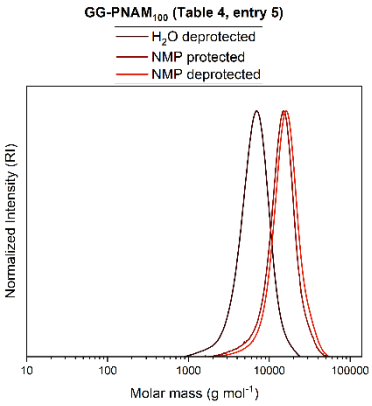


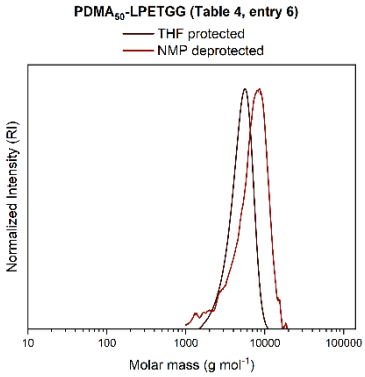
Figure 34: SEC data of PET-RAFT polymerizations with NIPAM and DMA using BABTC or BABTC-LPETGG as CTA/macroCTA.

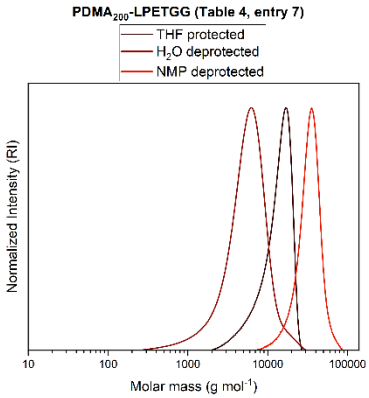

PNAM100-LPETGG (Table 4, entry 2)
H2O protected
NMP protected
NMP deprotected
Normalized Intensity (RI)
Molar mass (g mol-1)

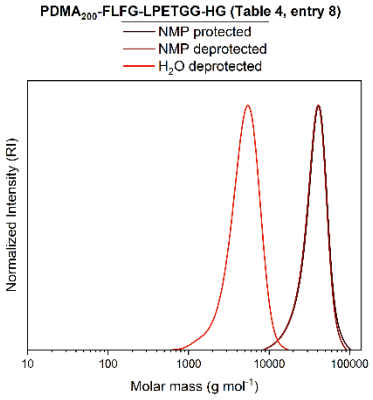

PNAM50-FLFG-LPETGG-HG (Table 4, entry 3)
H2O deprotected
NMP protected
NMP deprotected
Normalized Intensity (RI)
Molar mass (g mol-1)

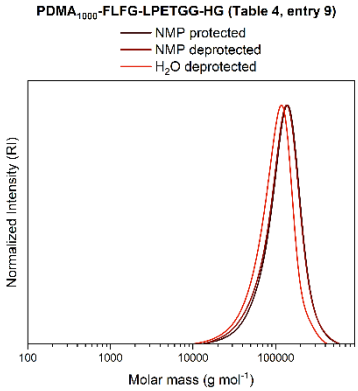

GG-PNAM20 (Table 4, entry 4)
NMP protected
NMP deprotected
THF protected
Normalized Intensity (RI)
Molar mass (g mol-1)

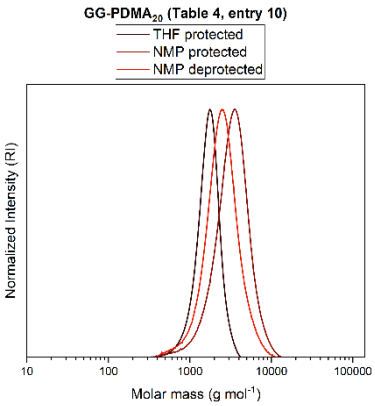

GG-PNAM100 (Table 4, entry 5)
H2O deprotected
NMP protected
NMP deprotected
Normalized Intensity (RI)
Molar mass (g mol-1)

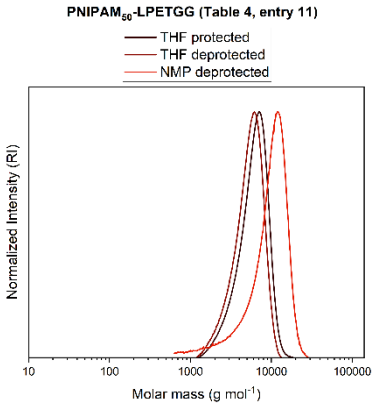

PDMA50-LPETGG (Table 4, entry 6)
THF protected
NMP deprotected
Normalized Intensity (RI)
Molar mass (g mol-1)

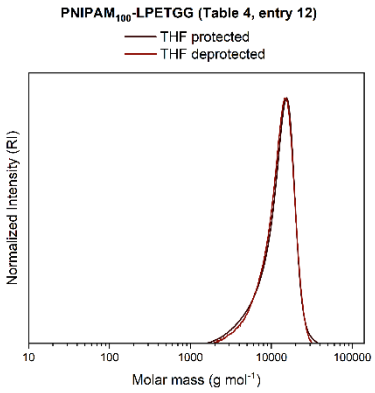

PDMA200-LPETGG (Table 4, entry 7)
THF protected
H2O deprotected
NMP deprotected
Normalized Intensity (RI)
Molar mass (g mol-1)

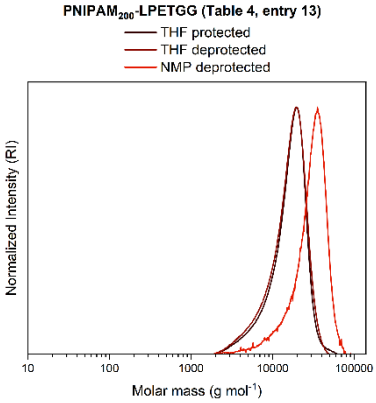

PDMA200-FLFG-LPETGG-HG (Table 4, entry 8)
NMP protected
NMP deprotected
H2O deprotected
Normalized Intensity (RI)
Molar mass (g mol-1)

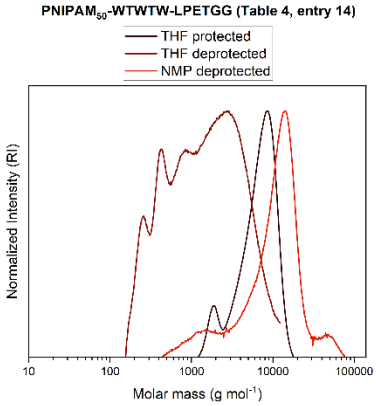

PDMA1000-FLFG-LPETGG-HG (Table 4, entry 9)
NMP protected
NMP deprotected
H2O deprotected
Normalized Intensity (RI)
Molar mass (g mol-1)

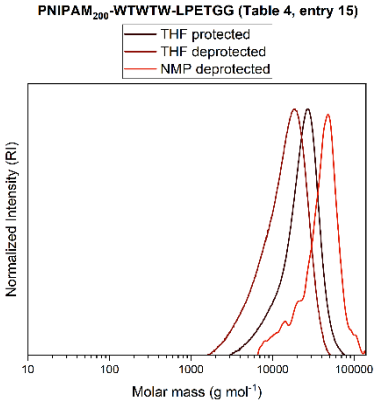

GG-PDMA20 (Table 4, entry 10)
THF protected
NMP protected
NMP deprotected
Normalized Intensity (RI)
10
100
1000
10000
100000
Molar mass (g mol-1)

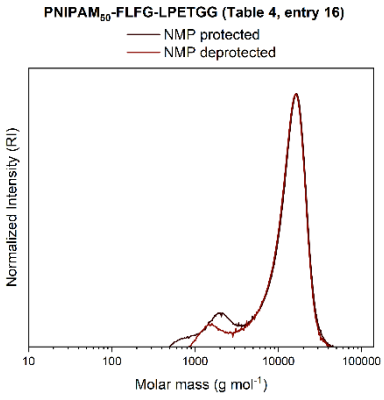

PNIPAM50-LPETGG (Table 4, entry 11)
THF protected
THF deprotected
NMP deprotected
Normalized Intensity (RI)
10
100
1000
10000
100000
Molar mass (g mol-1)

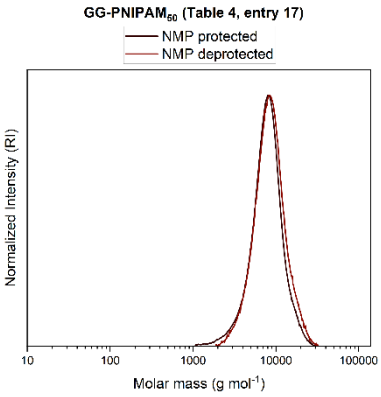

PNIPAM100-LPETGG (Table 4, entry 12)
THF protected
THF deprotected
Normalized Intensity (RI)
10
100
1000
10000
100000
Molar mass (g mol-1)


PNIPAM200-LPETGG (Table 4, entry 13)
THF protected
THF deprotected
NMP deprotected
Normalized Intensity (RI)
10
100
1000
10000
100000
Molar mass (g mol-1)


PNIPAM50-WTWTW-LPETGG (Table 4, entry 14)
THF protected
THF deprotected
NMP deprotected
Normalized Intensity (RI)
10
100
1000
10000
100000
Molar mass (g mol-1)


PNIPAM200-WTWTW-LPETGG (Table 4, entry 15)
THF protected
THF deprotected
NMP deprotected
Normalized Intensity (RI)
10
100
1000
10000
100000
Molar mass (g mol-1)


PNIPAM50-FLFG-LPETGG (Table 4, entry 16)
NMP protected
NMP deprotected
Normalized Intensity (RI)
10
100
1000
10000
100000
Molar mass (g mol-1)


GG-PNIPAM50 (Table 4, entry 17)
NMP protected
NMP deprotected
Normalized Intensity (RI)
10
100
1000
10000
100000
Molar mass (g mol-1)

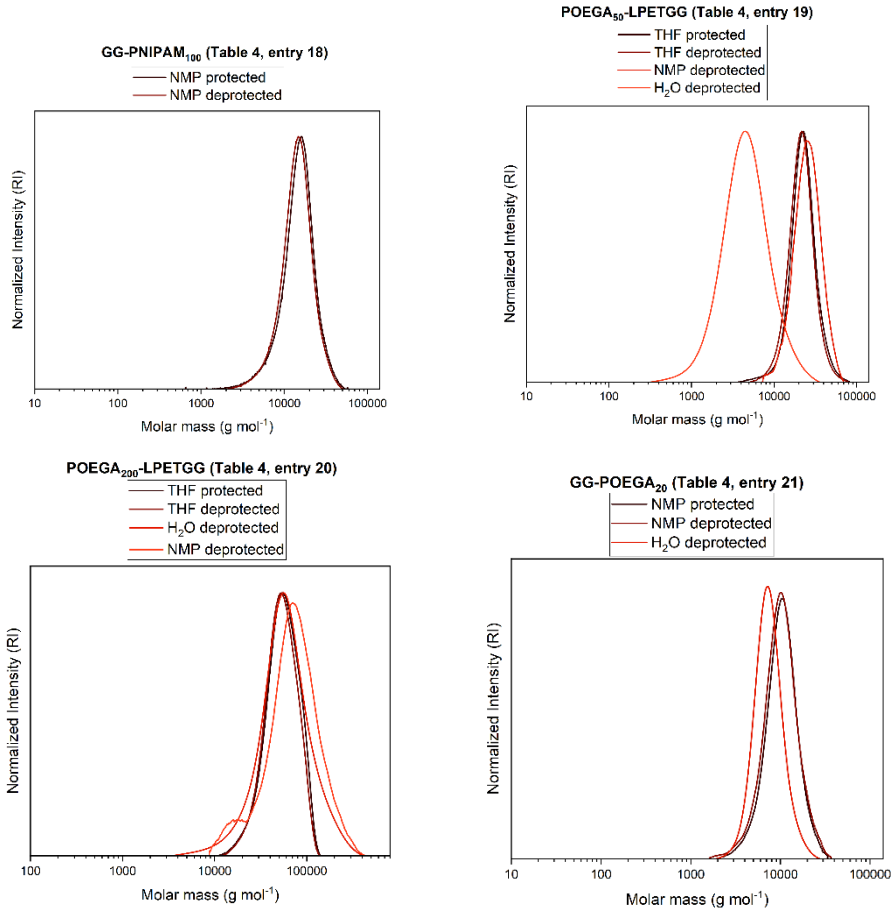


Figure 35: SEC elugrams of all polymers synthesized via XPI-RAFT.

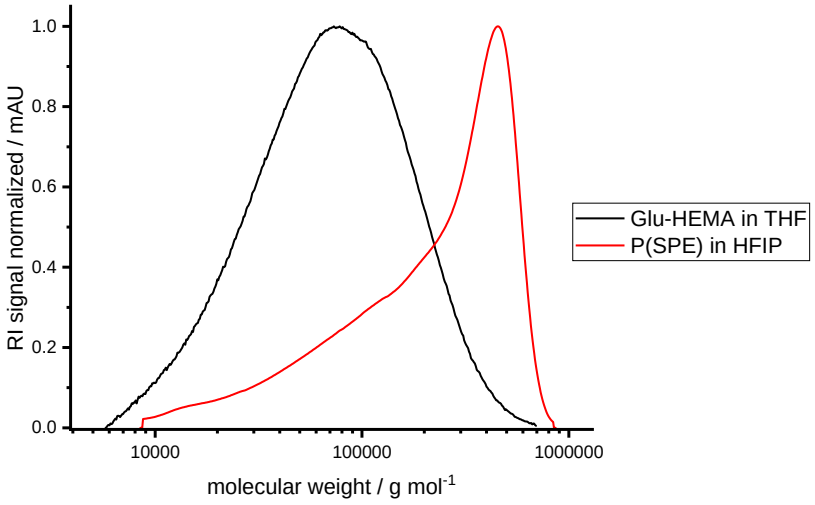


Figure 36: SEC data of P(Glu-HEMA) and P(SPE) synthesized via XPI-RAFT.

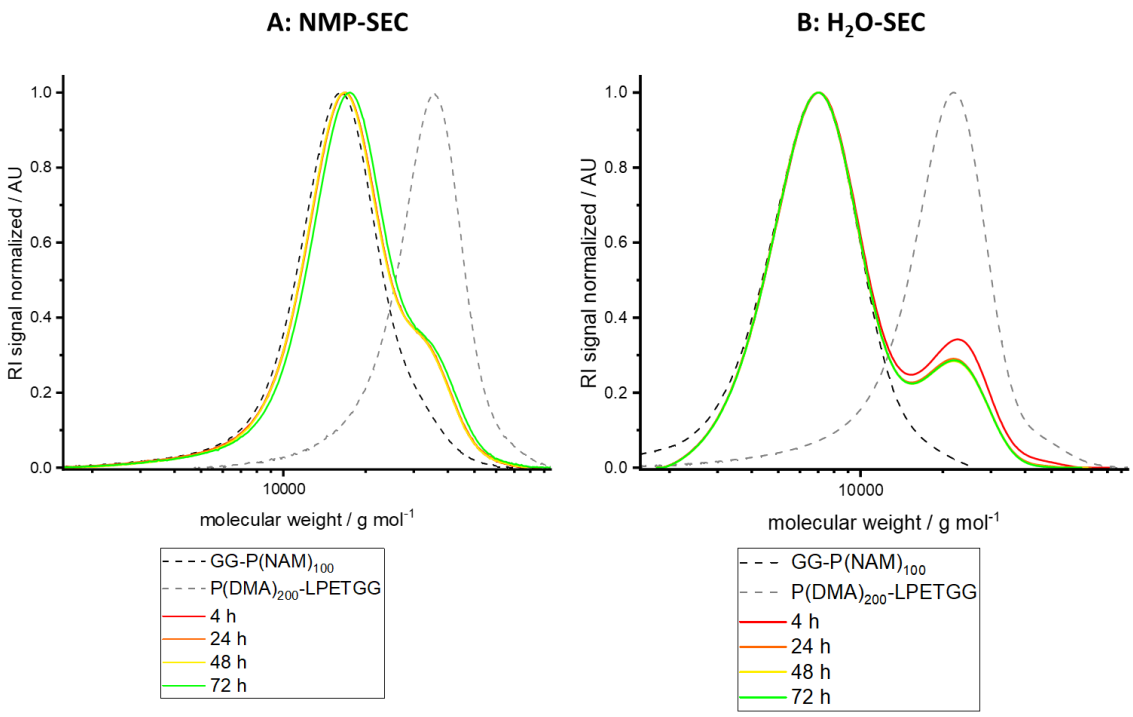


Figure 37: Comparison of SEC data in NMP (A) and aqueous SEC (B): SML of 1 mM GG-P(NAM)$_{100}$ and 0.1 mM P(DMA)$_{200}$-LPEGG, 10 mol-% SrtA-4M. Data normalized to 1 at educt peak.

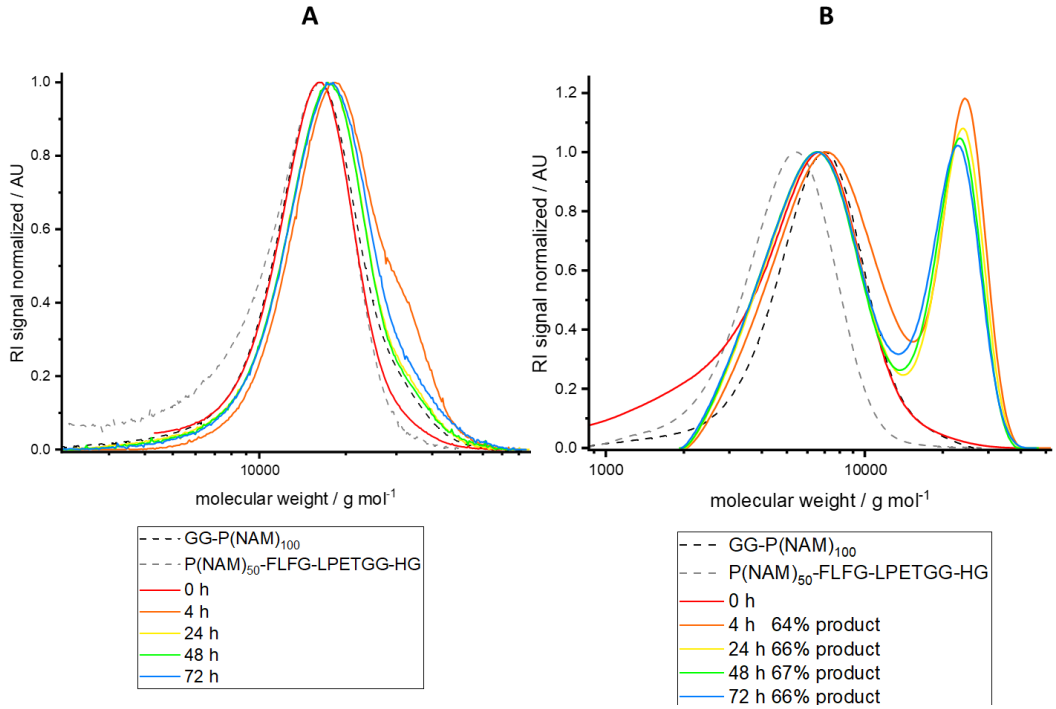


Figure 38: Comparison of SEC data in NMP (A) and aqueous SEC (B): SML of 0.1 mM GG-P(NAM)$_{100}$ and 0.1 mM P(NAM)$_{50}$-FLFG-LPEGG-HG, 10 mol-% SrtA-4M, 0.2 mM $NiSO_4$. Data normalized to 1 at educt peak.

Table 10: Experimental data for all polymer-polymer SML using 10 mol-% SrtA-4M. a: conversion calculated by deconvolution of aqueous SEC data. b: only SEC in NMP available, aqueous SEC not possible due to thermoresponsive behavior of P(NIPAM)

Nucleophile	Nucleophile concentration / mM	Recognition sequence	Recognition sequence concentration / mM	Maximum conversion[a] / %	$NiSO_4$ (0.2 mM)
GG-$P(NAM)_{100}$	0.1	$P(NAM)_{50}$-FLFG-LPETGG-HG	0.1	67	Yes
GG-$P(NAM)_{100}$	0.1	$P(DMA)_{200}$-FLFG-LPETGG-HG	0.1	0	Yes
GG-$P(OEGA)_{20}$	0.1	$P(NAM)_{50}$-FLFG-LPETGG-HG	0.1	54	Yes
GG-$P(NIPAM)_{100}$	0.1	$P(DMA)_{200}$-FLFG-LPETGG-HG	0.1	-[b]	Yes
GG-$P(NIPAM)_{100}$	0.1	$P(NAM)_{50}$-FLFG-LPETGG-HG	0.1	-[b]	Yes
GG-$P(NAM)_{100}$	0.1	$P(DMA)_{1000}$-FLFG-LPETGG-HG	0.1	0	Yes
GG-$P(NAM)_{100}$	0.1	$P(NIPAM)_{50}$-FLFG-LPETGG-HG	0.1	-[b]	Yes
GG-$P(NAM)_{100}$	1.0	$P(DMA)_{200}$-LPETGG	0.1	0	No
GG-$P(NAM)_{100}$	1.0	$P(OEGA)_{200}$-LPETGG	0.1	0	No

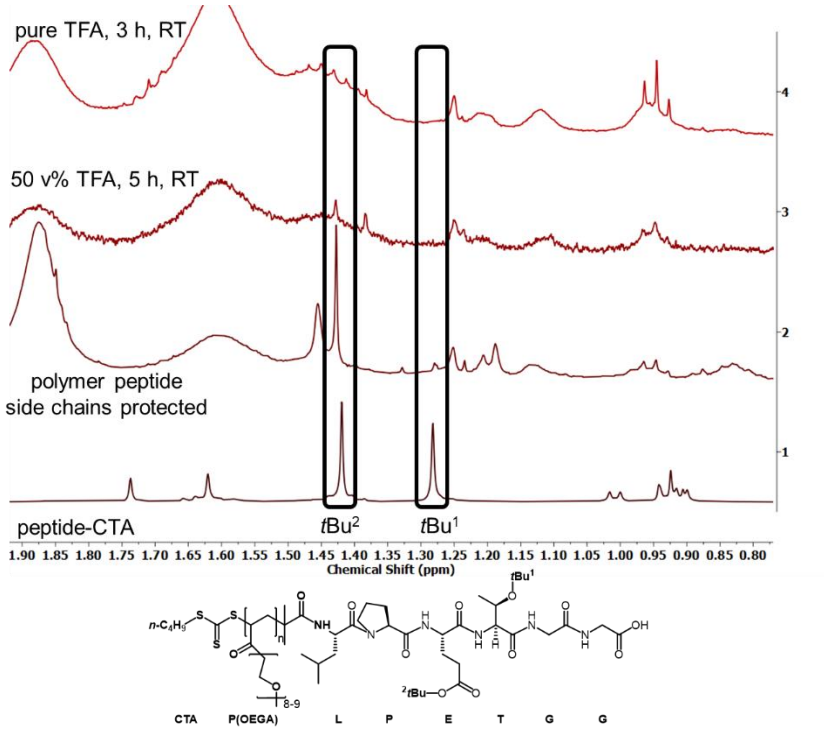


Figure 39: ^{1}H-NMR spectra for comparison of protected and deprotected peptide side chains in an exemplary P(OEGA)-LPETGG peptide-polymer building block, showing complete removal of peaks corresponding to *t*Bu only after treatment with pure TFA.

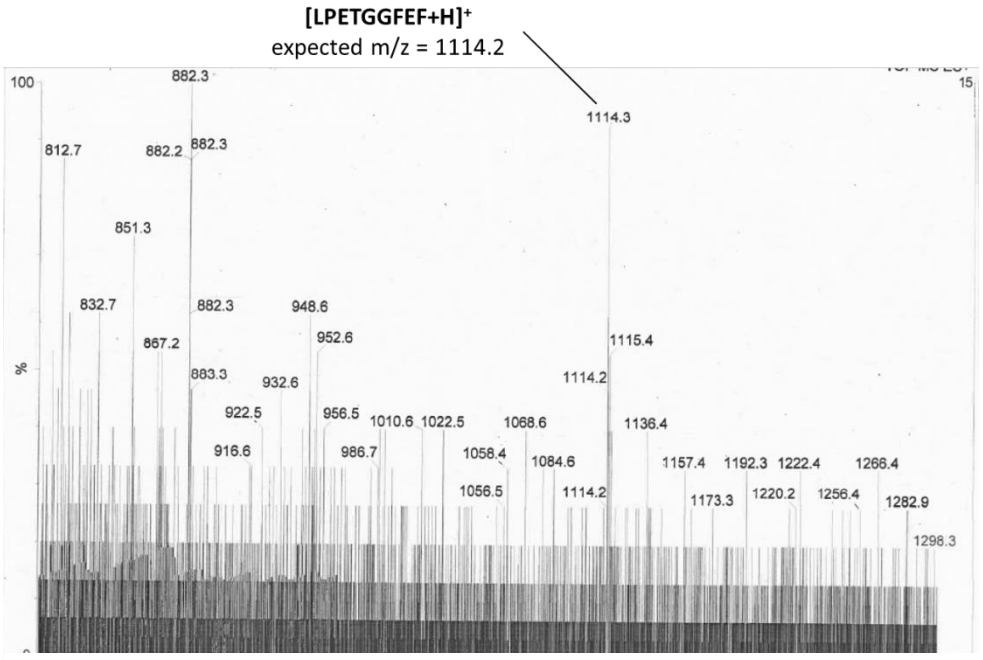


Figure 40: ESI-MS spectrum of HPLC fraction containing SML product using LPETGG and GGFEF as reactants. Data was received as printout from the analytical department.

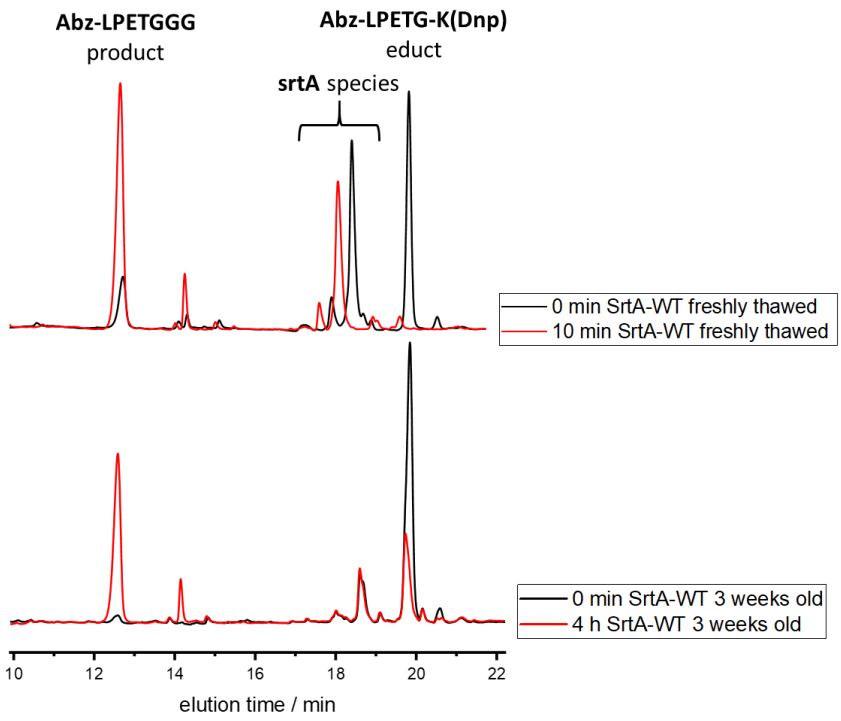


Figure 41: HPLC elugrams (absorption at 205 nm) showing a benchmark test using Abz-LPETGK(Dnp) (0.05 mM, 1 eq) as substrate and GGG (5 mM, 100 eq) as nucleophile. 4 mol-% SrtA-WT (fresh or 3 weeks old) were used. Fresh SrtA-WT consumed all substrate within 10 min, old SrtA-WT did not consum all substrate even after4 h.

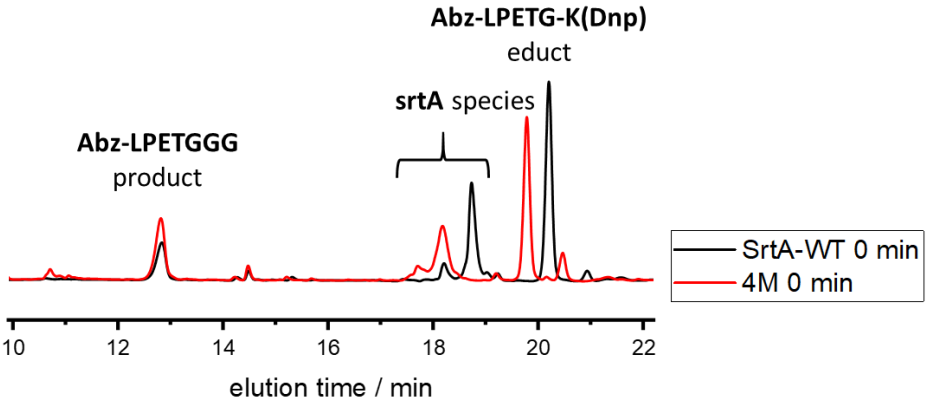


Figure 42: HPLC elugrams (absorption at 220 nm) showing a benchmark test using Abz-LPETGK(Dnp) (0.05 mM, 1 eq) as substrate and GGG (5 mM, 100 eq) as nucleophile. 4 mol-% SrtA-WT or SrtA-4M were used. Already after “0 min” reaction time (sample quenched after 15-30 s) product is formed.