Synthesis and modification of abiotic sequence-defined poly(phosphodiester)s

Niklas Felix König

UNIVERSITÉ DE STRASBOURG

École Doctorale des Sciences Chimiques ED 222

Institut Charles Sadron, CNRS

Thèse présentée par :

Niklas Felix KÖNIG

soutenue le : 03 septembre 2018

pour obtenir le grade de : **Docteur de l'Université de Strasbourg**

Spécialité : Chimie des Polymères

Synthesis and Modification of Abiotic Sequence-defined Poly(phosphodiester)s

THÈSE dirigée par :

M. Jean-François LUTZ	Directeur de Recherche CNRS, Institut Charles Sadron, Strasbourg

RAPPORTEURS :

M. Renaud NICOLAŸ	Professeur, ESPCI ParisTech
M. Michael A. R. MEIER	Professeur, Karlsruhe Institute of Technology

AUTRES MEMBRES DU JURY :

Mme Jutta RIEGER	Chargée de recherche, Université de Pierre et Marie Curie Paris VI
Mme Emilie MOULIN	Chargée de recherche, Institut Charles Sadron, Strasbourg
M. Jean-François NIERENGARTEN	Directeur de Recherche CNRS, École européenne de chimie, polymères et matériaux, Strasbourg

Meiner Mutter

Annegret König

gewidmet.

Acknowledgment

Three intense years have passed. The majority of the time's scientific output is covered in this manuscript. But the time I have spent in Strasbourg means much more to me. I have met great people and found new friends. We experienced joy, frustration and excitement together. That is why I want to express my gratitude and appreciation to all those who have accompanied me during this journey—in Strasbourg or from far.

*First and foremost, dear **Jean-François Lutz**, I want to thank you for your steadfast optimism throughout the last three years. Apart from other valuable skills like simple graphical representation of complex scientific problems, I hope I can say that you as my doctoral supervisor taught me a new attitude of tackling problems: confidence.*

*Dear **Aziz**, roi de la colonne, maître, chef, endurant and with stoic patience you have accompanied my efforts in the lab. We discussed the project designs and practical implementations that form the base of this thesis. You are an amazing chemist with knowledge in so many areas. Even more importantly, you are a very supportive, friendly and motivating lab mate. I will miss working together with you as I already miss our coffee breaks.*

*Dear **Gianni** and **Benoît**, together with "our" postdoc Aziz, we shared the lab and I want to thank all of you for this turbulent and eventful time. You are great guys, we were a strong team and I much appreciated also our intellectually less stimulating conversations while working. **Giannibananni**, you were my «binôme» from the first until the last day of our PhD studies. Luckily, we like eachother… Together, we discovered Strasbourg, shared uncounted beers and spend so much time on different trips through Europe together. I am very happy having had you at my side all that time. **Sneb**, we also became friends and I am grateful our paths crossed at the ICS. I will never forget how we followed our passion for gastronomic delights together as well as our late hour discussions. Istanbul, Berlin, … what is next?*

*Research is team work. Gratefully, I acknowledge the supportive spirit among postdoctoral researchers and PhD students.Thank you, **Aziz**, **Roza**, **Tito**, **Clothilde**, **Evgeniia**, **Gianni**, **Benoît**, **Guillaume**, **Denise**, **Jeroen**, **Eline**, **Chloé** and **Sofia**! At least the hoodie makes us more than just a research team… As part of the "third floor" **Eliott**, **Swann**, **Clément** and **Cécile**, you were part of this. For support and advices I want to especially thank **Laurence** and **Paul** as well as all other members of the **PMC group**.*

*I want to thank **Catherine Foussat** and **Mélanie Legros** for conscientiously analyzing samples as well as members of **SAMS**, **SYCOMMOR** and **PECMAT** for helping out when things had to be solved unbureaucratically. I also enjoyed the scientific exchange with researchers from other domains. **Bert Meijer**, **Anja Palmans**, **Beatrice Adelizzi**, **Marcin Ślęczkowski**, **Laurence Charles**, **Salomé Poyer**, **Jean-Arthur Amalian**, **Jan Behrends, Mordjane Boukhet**, **Albrecht Ott**, and all PIs and ESRs of **ITN EUROSEQUENCES**, you enriched my research experience very much.*

*Cher «hawk eye» **Guillaume**, sans toi il n'y aurait peut-être jamais eu cette thèse. Sans tes traductions innombrables, ta maîtrise des délais et toute ton aide en face des formalités administratives, sans ta présence pendant la phase finale de l'écriture et mon déménagement, sans toi … rien. Merci !*

*Léif **Pit**, mäi Lëtzebuergesch Frënd, dass ich dich in der Chemikercommunity Straßburgs kennenlernen durfte, ist ein besonderes Geschenk meiner Promotionszeit. Danke für die gemeinsame Zeit, für deinen motivierenden Input in so vielen Bereichen und die Hunderten von Kilometern auf dem Vëlo.*

Und am Ende danke ich all meinen Freunden aus Deutschland, die unseren Kontakt weiter pflegten, die mich besuchen kamen und die ich besuchen durfte oder mit denen ich wunderschöne Urlaube verbrachte. Diese Gewissheit der Freundschaft hat die Promotion im Ausland sehr erleichtert. Auch meiner Familie möchte ich danken, die mich immer begleitet hat. Dass ihr, Familie und Freunde, so zahlreich bei der Promotionsverteidigung wart, bedeutet mir sehr viel und hat mich stolz gemacht.

*Ich danke dir innigst, lieber **Norbert**, für die Gewissheit, dass ich mich auf dich als Vater und deine bedingungslose Hilfe immer verlassen kann.*

*Liebe **Elisabeth**, dass ich so vieles mit dir teilen darf, ist mein größtes Glück. Du bist wundervoll.*

Viel Vergnügen beim Lesen. Bonne lecture. Enjoy reading!

Table of Contents

List of Abbreviations

Ac	acetyl
AcOH	acetic acid
ADMET	acyclic diene metathesis
α-HL	α-hemolysin
AMA	ammonia/methylamine
anh.	anhydrous
ARGET	activatorsregenerated by electron transfer
ASCII	American standard code for information interchange
ATRP	atom transfer radical polymerization
Bz	benzyl
CE	2-cyanoethyl
CPG	controlled pore glass
CuAAC	copper(I)-assisted alkyne-azide cycloaddition
Đ	dispersity
dA	deoxyadenosine
DBTL	dibutyltin dilaurate
DCC	*N*,*N*'-dicyclohexylcarbodiimide
DCM	dichloromethane
DIPEA	*N*,*N*-diisopropylethylamine
DMAP	4-dimethylaminopyridine
DMF	dimethylformamide
dmf	dimethylformamidyl
DMSO	dimethylsulfoxide
DMT	4,4'-dimethoxytrityl
DMTrCl	4,4'-dimethoxytrityl chloride
DNA	deoxyribonucleic acid
dNbipy	4,4'-di-*n*-nonyl-2,2'-bipyridine
DP	degree of polymerization
Dppp	1,3-bis(diphenylphosphino)propane
DSC	*N*,*N*'-disuccinimidyl carbonate
dT	deoxythymidine
ER	endoplasmatic reticulum
ESI	electrospray ionization
EtOAc	ethyl acetate
FITC	fluorescein 4-isothiocyanate
Fmoc	fluorenylmethyloxycarbonyl
Fmoc-Cl	fluorenylmethyloxycarbonyl chloride
F-SPE	fluorous solid phase extraction
Gly	glycine
HBTU	2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HD	high-defintion
HL	high-loading
HOBt	hydroxybenzotriazole
HPLC	high-performance liquid chromatography
HRMS	high-resolution mass spectrometry
ICAR	initiators for continuous activator regeneration

*i*Pr-Pac	4-*iso*-propylphenoxyacetyl
IR	infrared
Isb	*iso*-butyryl
KO*t*Bu	potassium *tert*-butanolate
lcaa	long-chain alkyl amine
Me_6TREN	tris[2-(dimethylamino)ethyl]amine
MeCN	acetontrile
MeLi	methyllithium
MeOH	methanol
MMT	monomethoxytrityl
MPPS	aminomethylated macroporous polystyrene
mRNA	messenger RNA
MS	mass spectrometry
MS/MS or MS_2	tandem mass spectrometry (tandem-MS)
MsCl	mesyl chloride / methanesulfonyl chloride
MSNT	1-(2-mesitylenesulfonyl)-3-nitro-1*H*-1,2,4-triazole
MspA	mycobacterium smegmatis porin A
MWCO	molecular weight cut-off
n.d.	not detected
n.e.	not expected
n-BuLi	*n*-butyllithium
NGS	next-generation sequencing
NMP	nitroxide-mediated polymerization
NMR	nuclear magnetic resonance
NPPOC	(2-nitrophenyl)propoxycarbonyl
NPPOC-Cl	2-(2-nitrophenyl)propyl chloroformate
NVOC	nitroveratryloxycarbonyl
oa-TOF	orthogonal acceleration time-of-flight
P-3CR	Passerini three-component reaction
Pac	phenoxyacetyl
PCR	polymerase chain reaction
PEG	polyethylene glycol
PET	polyethylene terephthalate
PS	polystyrene
*p*TsCl	4-toluenesulfonyl chloride
*p*TsOH	4-toluenesulfonic acid
PVA	polyvinyl alcohol
PVC	polyvinyl chloride
PyBOP	benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
Pyr	pyridine
RACP	radical addition-coupling polymerization
RAFT	reversible addition-fragmentation chain transfer
RNA	ribonucleic acid
ROMP	ring-opening metathesis polymerization
ROP	ring-opening polymerization
RT	room temperature
SARA	supplemental activator and reducing agent
SBS	Society for Biomolecular Screening
SEC	size-exclusion chromatography

SET-LRP	single electron transfer living radical polymerization
SPPS	solid-phase peptide synthesis
TBA	tetrabutylammonium
TBAF	tetrabutylammonium fluoride
TBDMS	*tert*-butyldimethylsilyl
TBTA	tris((1-benzyl-4-triazolyl)methyl)amine
*t*BuOH	*tert*-butyl alcohol
TCA	trichloroacetic acid
TEA	triethylammonium
TEMPO	2,2,6,6-tetramethylpiperidin-1-yl)oxyl
TFA	trifluoroacetic acid
TFAA	trifluoroacetic anhydride
THF	tetrahydrofuran
THP	tetrahydropyran
TIPS	triisopropylsilyl
TIPS-Cl	triisopropylsilyl chloride
TLC	thin layer chromatography
TMS	trimethylsilyl
TOM	tri-*iso*-propylsilyloxymethyl
Tr	trityl
*t*RNA	transfer RNA
U-4CR	Ugi four-component reaction
UV	ultraviolet
UV-Vis	ultraviolet/visible

List of Figures

List of Schemes

List of Tables

Introduction Générale

La vie, telle que nous la connaissons, est basée sur les polymères. Les biopolymères–comme les protéines–gouvernent tous les processus d'un organisme tandis qu'une grande proportion des matériaux créés par l'Homme sont construits à partir de polymères. Ils facilitent la vie quotidienne depuis le commencement de l'Age du plastique, il y a environ un siècle. Les nouvelles générations de polymères sont actuellement conçues et il sera montré dans cette thèse que la chimie macromoléculaire applique des concepts et techniques issus d'une grande variété de domaines scientifiques voisins.

La diversité des biopolymères est riche. Certains montrent une complexité exceptionnelle et semblent parfaitement conçus pour leur fonction, malgré la variété limitée des monomères à partir desquels ils sont fabriqués. Une de leurs caractéristiques importante est la régulation de la séquence des monomères dans la chaîne. Certains biopolymères, comme l'ARN et l'ADN, possèdent des monomères précisément ordonnés dans leur structure primaire. Cette dernière contrôle la structure secondaire–qui fait référence à l'appariement des bases formant des arrangements en hélices et en boucles–de même qu'elle régit la structure tertiaire. Par conséquent, les hélices et boucles de l'ARNt sont positionnées les unes par rapport aux autres d'une manière définie. En plus de la conception des monomères, leur séquence détermine intrinsèquement la fonction du polymère : l'ARNt transporte les acides aminés jusqu'au ribosome pour la synthèse de peptides, alors que l'ADN, composé de deux brins complémentaires, stocke l'information permettant la synthèse d'un peptide en particulier.

Il est évident que le contrôle des séquences de monomères a suscité une attention croissante dans la science des polymères en tant qu'outils majeur pour la conception de systèmes et matériaux artificiels complexes.[1] Des procédés permettant d'atteindre des degrés élevés de contrôle des séquences ont été développés pour les polymères synthétiques. De nombreuses approches régulent la microstructure des polymères obtenus avec un mécanisme par étapes. Des structures de copolymères périodiques ou à blocs peuvent être synthétisées par des polymérisations en chaîne. En particulier, les polymérisations radicalaires contrôlées comme l'ATRP, la NMP et la RAFT sont des techniques importantes qui permettent un contrôle assez fin du degré de polymérisation et mènent par exemple au positionnement précis d'unités monomères le long de chaînes polystyrène présentant une faible dispersité.[2]

Le but principal–les séquences définies–peut être atteint par des mécanismes de polymérisations multi-étapes, qui reposent le plus souvent sur des synthèses supportées.[3] Le concept sous-jacent est l'interruption de la polymérisation par étapes après chaque addition de monomère et le contrôle des conditions d'addition du monomère suivant. Cela peut être réalisé par des approches orthogonales ou en ayant recours à des groupement protecteurs. Récemment, les outils nécessaires à la synthèse de polymères avec une structure primaire précise se sont rapidement étendus.[4] Conçue originellement par Caruthers pour la synthèse d'oligonucléotides (des courts fragments d'ADN), la chimie des phosphoramidites est une stratégie de synthèse déjà bien établie.[5] Comme décrit sur la Figure 1B, elle n'en est pas moins une approche à la fois très efficace et façonnable à souhait grâce à une grande variété de monomères qui peuvent être imaginés pour accéder à des poly(phosphodiester)s sur mesure avec une procédure automatisée.[6]

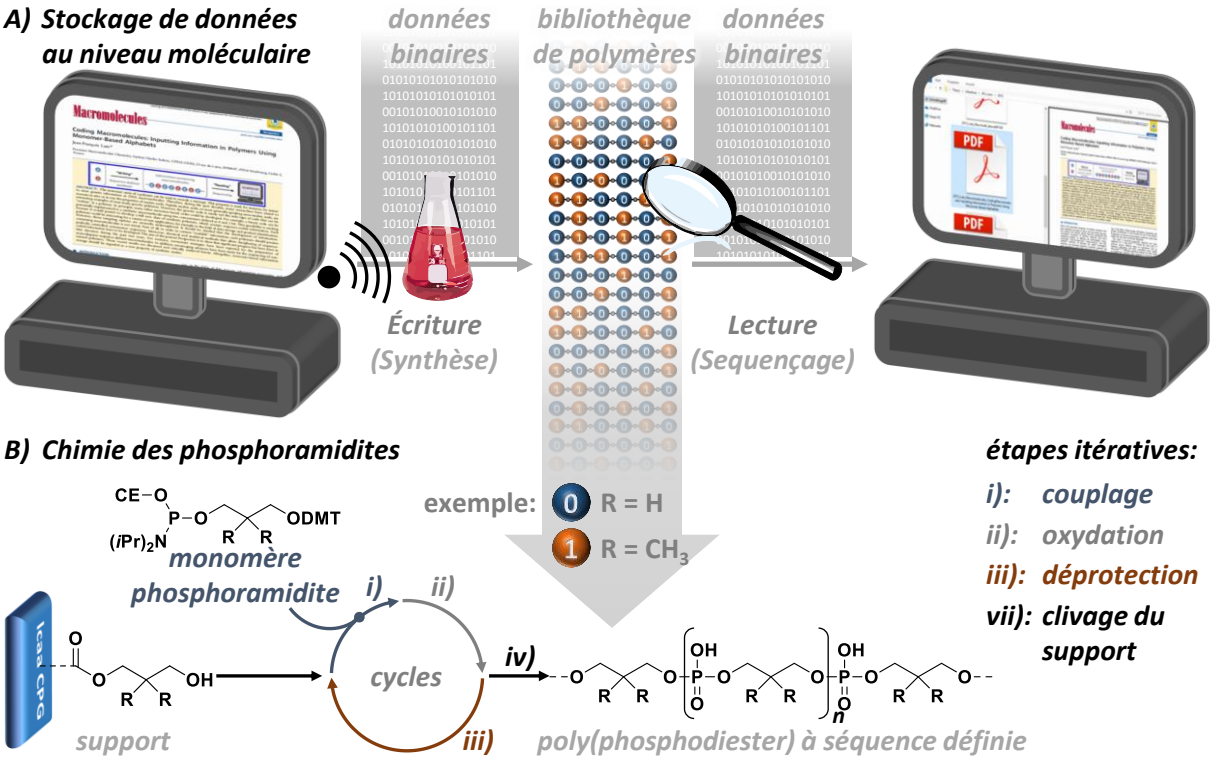


Figure 1: A) Stockage de données au niveau moléculaire par écriture de données binaires sur des polymères et leurs lecture ; B) Chimie des phosphoramidites pour la synthèse de poly(phosphodiester)s numériques (montré : code binaire).

De manière analogue aux macromolécules contenant de l'information biologique comme l'ADN, le stockage de données à l'échelle moléculaire est une application prometteuse pour les polymères à séquences définies synthétiques : différents comonomères peuvent être définis comme les bits 0 ou 1, et la structure primaire régule leurs positions afin de produire l'information binaire encodée.[7] Comme présenté par la Figure 1A, un fichier d'ordinateur pourrait ainsi être traduit en un flux de données binaires, consistant en une suite de bits 0 ou 1, et directement encodé dans une bibliothèque de polymères par un procédé de polymérisation multi-étapes automatisé. Ces polymères numériques doivent être pensés de sorte que des protocoles automatiques et contrôlés par des logiciels puissent encoder l'information puis la convertir à nouveau en son fichier d'origine. Étant donné le nombre d'outils biotechnologiques à disposition, le stockage d'information sur des brins d'ADN artificiels a récemment beaucoup évolué.[8-9] Cependant, il peut être imaginé que l'orientation vers les exemples biologiques pourrait également limiter les développements et applications futurs. La recherche sur les polymères contenant de l'information synthétiques est donc logiquement nécessaire pour prouver leurs capacités et démontrer que des possibilités chimiques infinies sont à disposition pour la conception de systèmes de stockage d'information aux attributs supérieurs.

Les travaux de recherche de cette thèse se situent à cette nouvelle frontière. La chimie des phosphoramidites est utilisée pour la synthèse efficace de nouveaux poly(phosphodiester)s à séquences définies. Leur design devrait permettre d'accéder à de nouveaux systèmes de stockage d'information au niveau moléculaire artificiels présentant des propriétés avantageuses.

Ce manuscrit est organisé en différentes sections. En premier lieu, une introduction détaillée de plusieurs aspects du domaine de recherche est donnée. Les chapitres qui suivent sont indépendants les uns des autres et couvrent les activités de recherche qui ont été réalisées dans le cadre du doctorat. Toutes les informations concernant les détails expérimentaux des projets de recherche peuvent être trouvées dans la partie expérimentale qui est elle-même organisée selon la structure des chapitres.

Enfin, les annexes contiennent les données additionnelles et discussions qui n'ont pas pu être incorporées au texte principal.

Le **Chapitre I** fournit au lecteur un résumé de l'état de l'art de la chimie macromoléculaire de précision et des disciplines avoisinantes. Le contrôle des séquences est discuté dans le contexte des systèmes biologiques, de même que les chimies synthétiques. Le parcours débute par une présentation des biopolymères à séquences définies, leur biosynthèse et leur accès synthétique. Ensuite, les progrès dans les polymérisations de séquences contrôlées pour obtenir des polymères de précision abiotiques sont couverts. Une attention particulière est portée sur les synthèses multi-étapes qui peuvent mener à des polymères à séquences définies abiotiques. Un résumé du développement historique de la chimie des phosphoramidites–la stratégie chimique sophistiquée pour la synthèse de poly(phosphodiester)s à séquences définies–et de ses variations et applications modernes est ensuite offert. Enfin, l'état de l'art des technologies de stockage d'information au niveau moléculaire est récapitulé afin de décrire les applications des polymères à séquences définies ayant de l'impact.

Le **Chapitre II** couvre un projet de recherche dont l'objectif est de permettre l'analyse de molécules uniques de poly(phosphodiester)s à séquences définies spécialement adaptés. Alors que la synthèse de poly(phosphodiester)s contenant de l'information avait déjà été établie,[6,10] ce n'est que très récemment que le déchiffrage de l'information encodée moléculairement a connu ses premiers balbutiements. Le séquençage en spectrométrie de masse en tandem a énormément évolué.[11-12] Néanmoins, la spectrométrie de masse est une technique d'analyse destructive et d'autres techniques de séquençage de polymères synthétiques sont développées en parallèle. Un de ces outils prometteurs est le séquençage par les nanopores, qui est basé sur le passage de l'analyte à travers une cavité de taille nanométrique. En bref, les variations de courant au cours de la translocation sont observées et peuvent être reliées à la structure primaire de l'analyte.[13] Une meilleure précision de la mesure est atteinte en utilisant un jeu de monomères qui montrent des interactions différenciables avec le pore. Pour la découverte d'une nouvelle classe de poly(phosphodiester)s séquençables, un grand nombre de monomères avec différentes chaînes latérales doivent être analysées selon leurs interactions avec un dispositif contenant un nanopore. C'est pourquoi un protocole de modification binaire se devait d'être développé dans une première phase, afin de permettre un contrôle des chaînes latérales post-polymérisation par des cycloadditions alcyne-azoture catalysées par le cuivre(I) successives. Ce concept est présenté sur la Figure 2A.

Dans le **Chapitre III**, le développement d'une voie de synthèse photocontrôlée des oligo(phosphodiester)s à séquences définies abiotiques est décrit. Le concept reposant sur un support soluble est dévoilé sur la Figure 2B. Afin de faire progresser les polymères numériques, les méthodes de synthèse et de séquençage efficaces sont cruciales, mais la manière dont l'information est organisée est tout aussi importante. Dans ce contexte, la miniaturisation des systèmes de stockage à l'échelle moléculaire est un aspect clé. Des capacités de stockage très élevées peuvent être atteintes en organisant la bibliothèque de chaînes polymères sur un microréseau. Par exemple, des biopuces d'ADN sont obtenues en synthétisant des centaines de milliers d'oligonucléotides différents en parallèle sur une lame de microscope en verre, en utilisant un protocole de chimie des phosphoramidites photocontrôlée.[14] Dans ce cas, des miroirs guident le stimulus lumineux sur des points prédéfinis de la surface, activant un contrôle spatial des cycles de réactions itératifs. Tandis que la chimie des phosphoramidites standard repose sur un groupement protecteur DMT clivable en conditions acides qui permet le couplage phosphoramidite suivant par la libération d'une fonction hydroxy, la fabrication des microréseaux d'ADN repose sur le groupement nitrophénylpropoxycarbonyle (NPPOC) qui protège la fonction alcool et peut être clivé par irradiation à $\lambda = 365$ nm. Ainsi, la démonstration de l'utilisation de protocoles phosphoramidites basés sur le groupement NPPOC pour la synthèse de microréseaux de polymères numériques a été visée.

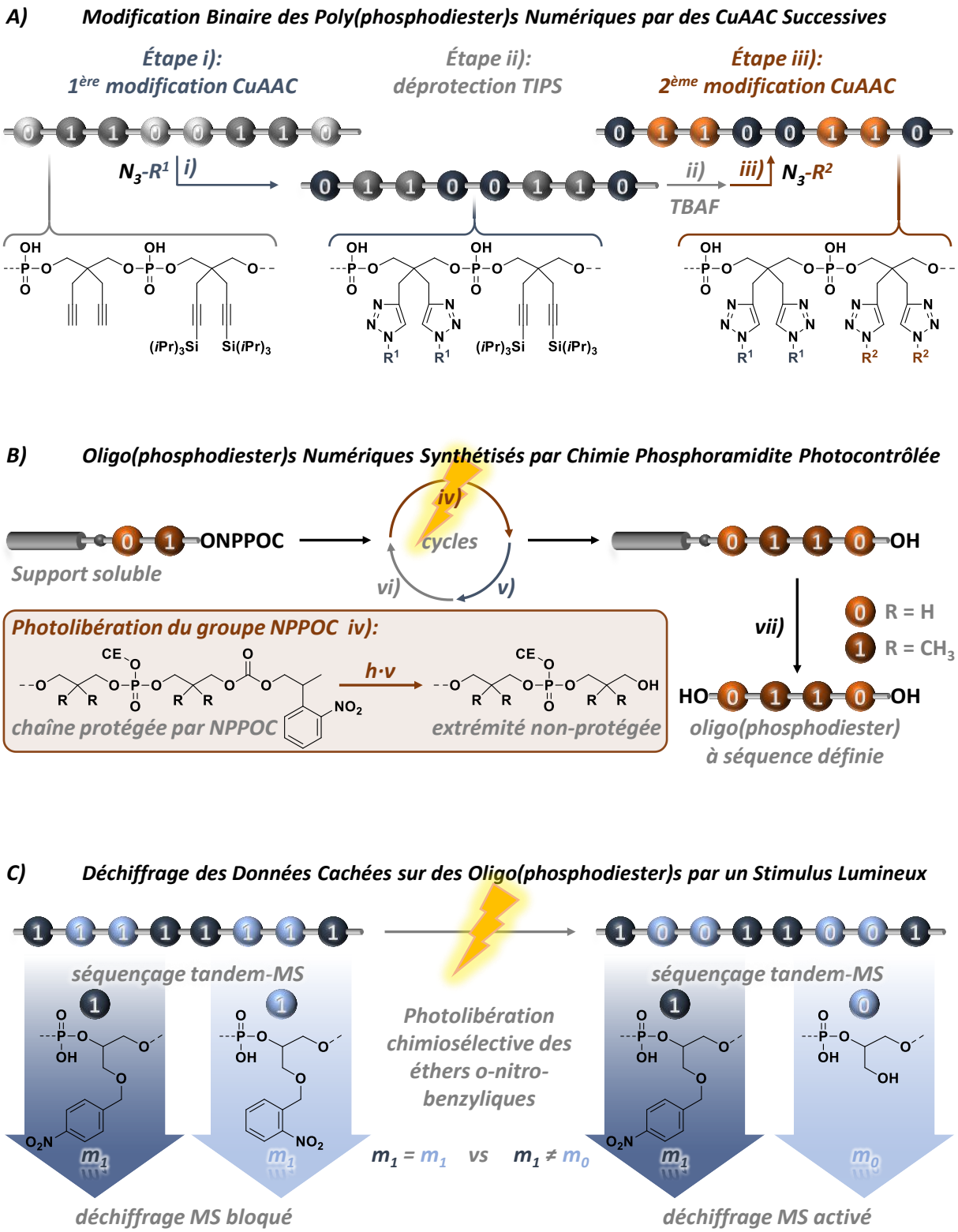


Figure 2: A) Concept de la modification successive binaire pour contrôler le caractère des chaînes latérales ; B) Concept de la synthèse phocontrôlée d'oligo(phosphodiester)s abiotiques pour la fabrication de microréseaux numériques ; C) Concept du stockage de données sur différentes couches et déchiffrage grâce à un stimulus.

Le **Chapitre IV** montre qu'il est facile de façonner les systèmes de stockage moléculaire abiotiques pour obtenir des propriétés difficilement atteignables avec des systèmes biologiques. Des possibilités chimiques pratiquement infinies sont disponibles pour les chimistes des polymères pour moduler les systèmes de stockage de données au niveau moléculaire et insérer des attributs additionnels. Une demande potentielle pour de tels systèmes est de stocker de l'information sur différentes couches.

Celles-ci seraient atteintes après traitement du système par des stimuli prédéfinis. Un tel exemple est connu dans les systèmes basés sur l'ADN qui reposent sur des monomères artificiels.[15] Une polymérase reconnaît ces derniers comme des monomères naturels et, après traitement par un agent chimique, les monomères modifiés sont reconnus comme d'autres monomères naturels. Dans un système basé sur la reconnaissance par une polymérase, les possibilités sont certainement limitées. C'est pourquoi un polymère contenant de l'information d'un genre nouveau devait être imaginé pour démontrer que l'emploi du séquençage par spectrométrie de masse en tandem pouvait donner accès à plusieurs niveaux d'information avec des systèmes répondant aux stimuli. Dans un premier exemple, l'information binaire a été encodée sur un niveau moléculaire caché qui ne pouvait être déchiffré qu'après traitement de la macromolécule dans certaines conditions. Un stimulus lumineux a été choisi pour induire une modification quantitative de l'une des deux fonctionnalités structurellement très proches. Les groupements ortho-nitrobenzyles sont des groupements protecteurs pour les alcools, amines, amides et acides carboxyliques.[16] Lorsque l'une des fonctionnalités est protégée par ce groupement ortho-nitrobenzyle sur un monomère qui est copolymérisé avec son monomère analogue protégé par un groupement para-nitrobenzyle, le séquençage par spectrométrie de masse en tandem– qui est sensible aux différences de masses–ne peut pas extraire l'information puisque les monomères ortho et para sont détectées comme des espèces identiques. En revanche, après irradiation avec la lumière, seuls les monomères avec le groupe ortho sont déprotégés et les unités monomères peuvent être distinguées en spectrométrie de masse. Le séquençage donne alors accès à l'information encodée binairement comme montré sur la Figure 2C.

General Introduction

Life as we know it is based on polymers. Biopolymers like proteins govern all processes in an organism and a large share of man-made materials are composed of engineered polymers. They facilitate everyday life since the plastic age has begun a hundred years ago. Next-generation polymers are currently designed and it will be shown that the discipline within macromolecular chemistry applies concepts and techniques from a variety of neighboring scientific domains.

The diversity of biopolymers is rich. Some of them show exceptional complexity and seem to be perfectly designed for their very purpose – despite the modest diversity of available building blocks they are made of. An important feature is sequence definition. Some biopolymers, like RNA and DNA, have (four) building blocks precisely ordered in their primary structure. The latter governs the secondary structure—which refers to base paring, helix formation and loop arrangement—as well as it defines the tertiary structure. Accordingly, in tRNA helices and loops are positioned to another in a particular way. Besides monomer design, their sequence inherently determines the function of the polymer: tRNA transports amino acids to the ribosome for peptide synthesis while DNA, composed of two complimentary strands, stores the information of how to synthesize a particular peptide.

It is obvious that sequence control has attained growing attention in polymer science as a key tool to design artificial complex materials and systems.[1] Procedures that yield high degrees of sequence control were developed for synthetic polymers. Numerous approaches regulate the microstructure of polymers obtained by a step-growth mechanism. Periodic microstructures or block copolymers can be synthesized by chain-growth polymerizations. Especially controlled radical polymerizations like ATRP, NMP, and RAFT are important techniques that enable an additional fine control over the degree of polymerization that allow for example the rather precise positioning of single monomers along a low-dispersity polystyrene chain.[2]

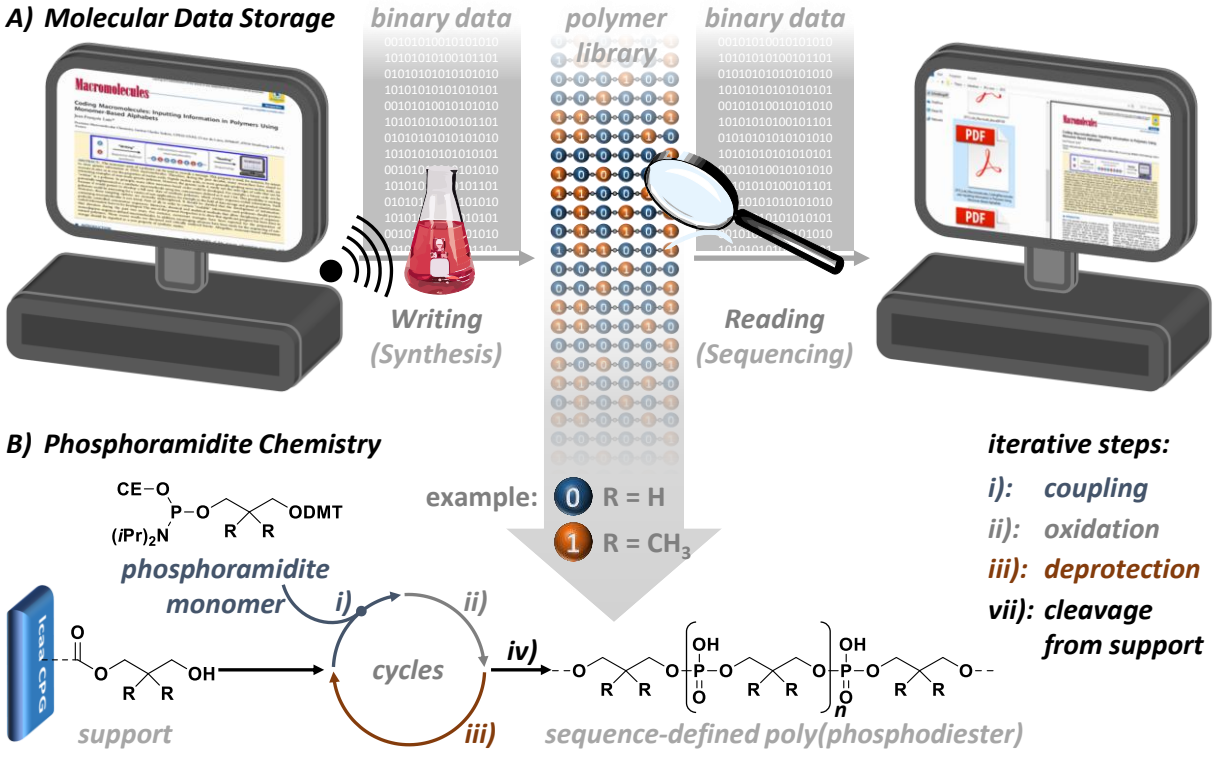


Figure 1: A) Data storage on the molecular level by writing binary information on polymers and reading them. B) Phosphoramidite chemistry to obtain information-containing poly(phosphodiester)s (shown: binary code).

The ultimate goal—sequence definition—can be accomplished by multistep-growth polymerization, often relying on support-assisted syntheses.[3] The underlying concept is to interrupt a step-growth polymerization after each monomer addition and to define under which conditions the next monomer addition is to follow. This can be achieved by orthogonal approaches or by using protective groups. Recently, the tool box for the synthesis of polymers with a precise primary structure has rapidly grown.[4] Originally conceived by Caruthers for the synthesis of oligonucleotides (short DNA fragments) phosphoramidite chemistry is an already established synthetic strategy.[5] Nevertheless, as depicted in Figure 1B, it is a very efficient approach and a broad variety of phosphoramidite monomers can be envisioned to access tailor-made sequence-defined poly(phosphodiester)s in an automated procedure.[6]

Analogous to the biological information storing macromolecule DNA, molecular data storage is a promising application of sequence-defined synthetic polymers: different comonomers can be defined as bit-0 and bit-1, the primary structure organizing their order is thus exhibiting encrypted binary information.[7] Visualized in Figure 1A, a computer file could be translated into a bitstream. The bitstream consists of binary data, a stream of 0-bits and 1-bits that is directly encrypted in a polymer library by automated multi-step growth polymerization. The digital polymers have to be designed in a way that automated and software-driven protocols could encrypt the information and convert it back to the original file. Given the number of biotechnological tools, artificial DNA-based data storage has recently considerably evolved.[8-9] Yet, it can be hypothesized that the close orientation on the biological example can limit future applicability as well. Research on synthetic information-containing polymers is hence needed to proof their conceptual capability and demonstrate that almost limitless chemical space is available to design molecular data storage systems with superior attributes.

The doctoral research is situated at this new frontier. Phosphoramidite chemistry is used to efficiently synthesize new sequence-defined poly(phosphodiester)s. Their design should give access to new artificial molecular data storages having advantageous features.

The present manuscript is organized in several sections. First, an in-depth introduction to various aspects of the research field is given. The following chapters are independently from another and cover the research activities that had been executed in the frame of the doctoral research. All information regarding experimental details of the research projects can be found in the experimental part that itself is organized according to the structure of the chapters. Finally, the annex contains additional details and discussions that could not be covered in the main document.

Chapter I supplies the reader with a summary of the state-of-the-art in precision macromolecular chemistry and neighboring disciplines. Sequence control is discussed in the context of biological systems as well as synthetic chemistry. The journey starts with the presentation of sequence-defined biopolymers, their biosynthesis and their synthetic access. Followingly, progress in sequence-controlled polymerizations to obtain precise abiotic polymers is covered. A focus is thereby set on multi-step growth synthesis that can lead to sequence-defined abiotic polymers. A summary of the historical development of phosphoramidite chemistry—the sophisticated chemical strategy for sequence-defined poly(phosphodiester) synthesis—is further given and its modern variations and applications are summarized. Finally, the state-of-the-art in molecular data storage technologies is summarized to describe an impactful application of sequence-defined polymers.

Chapter II covers a research project whose ultimate goal it is to enable single-molecule analysis of particularly designed sequence-defined poly(phosphodiester)s. While the synthesis of information-containing poly(phosphodiester)s had already been established,[6, 10] it was not until very recently that the deciphering of molecularly encoded information was still in its infancy. Sequencing by tandem mass spectrometry has dramatically evolved.[11-12] Yet, mass spectrometry is a destructive analysis technique and other sequencing techniques should be developed for synthetic polymers in parallel. One such

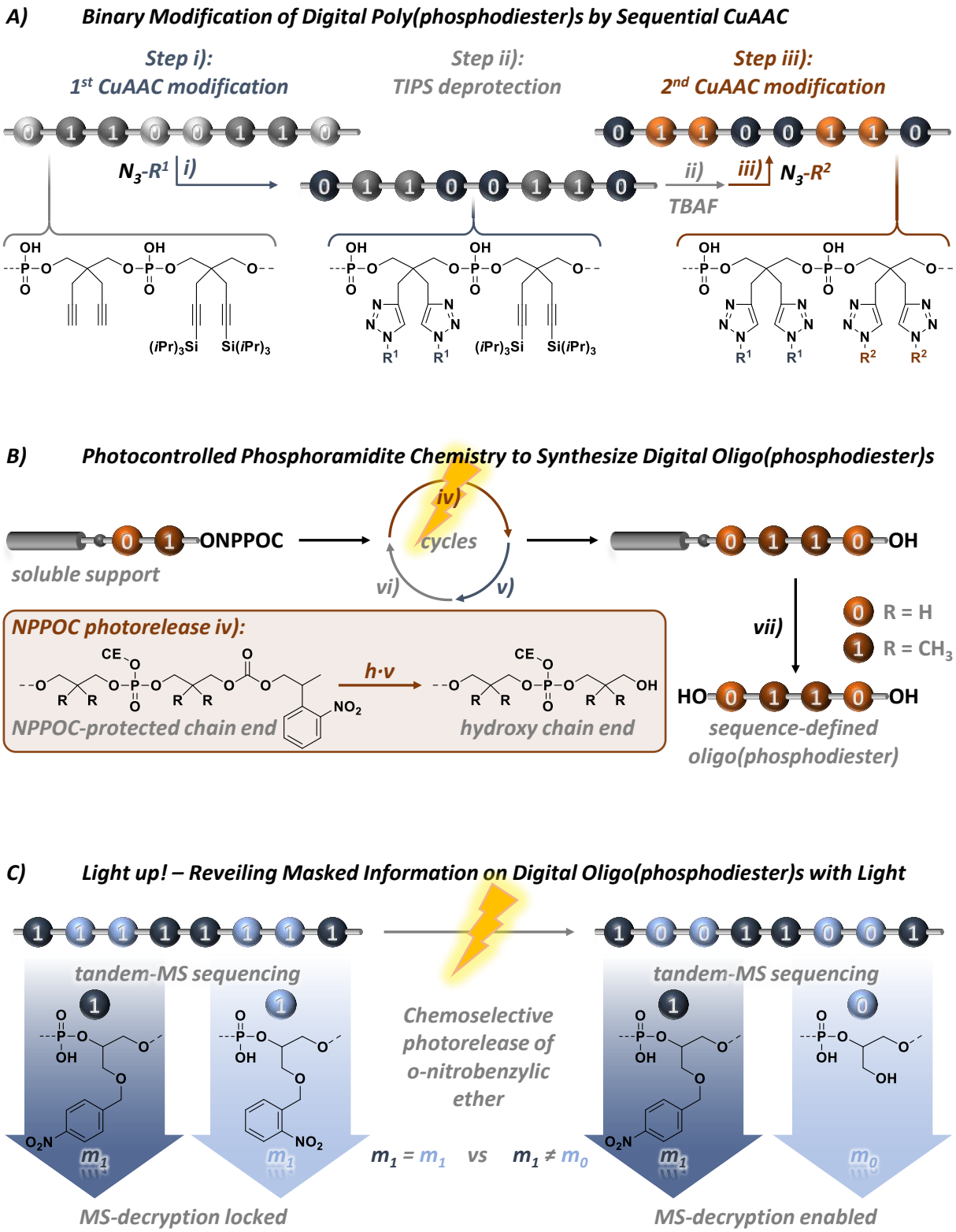


Figure 2: A) Concept of binary sequential post-polymerization modification of a digital poly(phosphodiester) to control side chain information; B) Concept of a photocontrolled synthesis of abiotic oligo(phosphodiester)s to access digital microarrays; C) Concept of coding in different layers and stimuli-dependent MS-readability.

promising tool is nanopore sequencing that is based on the threading of the analyte through a nanosized hole. In brief, current variations are observed during translocation and can be traced back to the primary structure of the analyte.[13] Accuracy is achieved by use of a set of monomers that show distinguishable interactions with the pore. If a new sequencable poly(phosphodiester) is to be found, many monomers with different side chains have to analyzed for their interaction in a nanopore set-up.

Hence, a binary modification protocol had to be developed in a first phase allowing an efficient post-polymerization side chain control by a sequential copper(I)-catalyzed alkyne-azide cycloaddition. The design is depicted in Figure 2A.

In **Chapter III** the development of a photocontrolled synthesis of defined abiotic oligo(phosphodiester)s is described. The concept benefitting from a soluble support is depicted in Figure 2B. To progress digital polymers, efficient synthesis and sequencing techniques are pivotal, but it is equally important how information is organized. Miniaturization of molecular data storages is a key aspect in this context. Very high storage capacities could be achieved when organizing a polymer library on a microarray. For example, DNA biochips are obtained by synthesizing some hundred thousands of different oligonucleotides in parallel on a glass microscope slide making use of a photocontrolled phosphoramidite protocol.[14] Here, micromirrors guide a light stimulus to predefined spots on the surface enabling a spatial control over iterative phosphoramidite reaction cycles. While standard phosphoramidite chemistry is relying on an acid-cleavable DMT protective group enabling the following phosphoramidite coupling by the liberation of a hydroxy group, DNA microarray fabrication is relying on the 2-(2-nitrophenyl)propoxycarbonyl (NPPOC) alcohol protecting group that can be cleaved *iv)* by irradiation at λ = 365 nm. It was thus aimed for a demonstration that NPPOC-based phosphoramidite protocols could also be employed for the synthesis of digital polymer microarrays.

In **Chapter IV** it is shown that abiotic molecular data storages can be easily designed showing features that are difficult to obtain in biological systems. An almost limitless chemical space is available for polymer chemists to tune molecular data storage systems and insert additional attributes. A potential demand to such systems is to store information in several logic layers. The latter would be accessed after treating the molecular data storage to predefined stimuli. One such example is known from a DNA-based system relying on artificial monomers.[15] A polymerase recognizes the latter as a sort of natural monomer and upon stimuli with a chemical reagent, the modified monomer is recognized as another natural monomer. The possibilities are certainly limited in a polymerase-based recognition system, hence a completely new information-containing polymer should be designed to demonstrate that using tandem mass spectrometry sequencing, stimuli response-systems could easily give access to different information layers. In a first example binary information should be encoded on the molecular level that could only be read-out when treating the macromolecule with certain conditions. It was concluded that a light stimulus could be used to elegantly induce a quantitative modification of one of two structurally closely related moieties. *ortho*-nitrobenzyl groups (like NPPOC) are protective groups for alcohols, amines, amides, and carboxylic acids.[16] When such moiety is *ortho*-nitrobenzyl-protected in one monomer and it is copolymerized with an analogue *para*-nitrobenzyl monomer, tandem mass spectrometry sequencing – that is sensitive towards mass differences – could not extract information because the *ortho*- and *para*-monomers are detected as the same species. Upon irradiation with light only the *ortho*-monomer would be modified and the monomeric units of the sequence can be distinguished by mass spectrometry. Sequencing will give access to binary coded information as it is depicted in Figure 2C.

Chapter I:

Sequence Control in Polymer Chemistry & Poly(phosphodiester)s as Digital Polymers

1 Introduction

Polymers are everywhere. Life as we know it is based on polymers. We eat them, we metabolize them and we create new ones. Biopolymers like proteins govern all processes in an organism. Also a large share of man-made materials are composed of engineered polymers facilitating everyday life since the *plastic age* has begun a hundred years ago. Yet, when comparing biopolymers with plastic commodity polymers, we obviously realize a vast difference in the degree of their complexity. However, this difference is about to decrease because strategies can be employed for a more precise polymer design gaining control of a polymer's microstructure as it is known from biological systems.

The following subchapters are dedicated to some recent development in the field of polymer science whose goal it is to investigate sequence control in synthetic polymers and its observable consequences. At first, sequence-defined biopolymers are presented in full detail including their biosynthesis and synthetic accessibility. Biopolymers are interesting because concepts employed in living matter can also be inspiring for polymer chemists. The progress and state-of-the-art of sequence control in synthetic polymers is presented afterwards, emphasizing strategies to obtain synthetic sequence-defined polymers. Phosphoramidite chemistry is a convenient tool to synthetically access DNA which is a sequence-defined poly(phosphodiester). A whole subchapter is revealing the strategy's *chemical evolution* during 30 years of steady improvements and how phosphoramidite chemistry is used nowadays. Finally, a promising application of synthetic sequence-defined polymers is presented. Absolute control of a polymer's sequence allows storing information on them. Thus, an overview of different molecular data storage systems is provided.

2 Polymers and Sequence Control

Polymers are macromolecules composed of repeating subunits. Consistently, the term *polymer* originates from the ancient Greek words πολύς (*polus*, English: *many* or *much*) and μέρος (*meros*, English: *parts*). The subunits forming the macromolecule during the process of *polymerization* are called monomers (ancient Greek: μόνος; *monos*, English: *one* or *sole*). Synthetic as well as biopolymers play a vital role in everyday life. Their diverse properties govern material characteristics of Teflon coatings and wood. They do also enable highly specific catalytic activity of proteins or facile decoding of synthetic digital polymers. While for some application in biology and synthetic chemistry a rough control over polymer length and the choice of (co)monomers and their ratio readily satisfies the demands on microscopic and macroscopic properties, there are requirements in more complex systems that call for higher levels of sequence control.

Being subject to evolution and thus relentless optimization during millions of years on our planet, living matter developed the chemical means for synthesis of polymers characterized by absolute sequence control. These biomolecules set a high yardstick when tailor-made properties or complex interaction are required. However, our growing understanding of the molecular machinery in living systems enables exploiting and manipulating the latter to obtain e.g. synthetic sequence-defined biomolecules. Moreover, fully synthetic protocols have been developed for the same purpose. Humankind furthermore asks for polymer properties that never before had to be challenged by nature. Research on chemical tools for diverse controlled polymerizations and the analysis of new engineered sequence-controlled abiotic macromolecules can offer us new interesting features.[1]

The multidisciplinary field of polymer chemistry in the context of sequence control is followingly described discussing relevant examples of biopolymers and modern synthetic polymers.

2.1 Sequence Definition in Biological Systems

Deoxyribonucleic acids (DNA), ribonucleic acids (RNA), proteins and certain oligosaccharides are biomolecules characterized by strict sequence control. Their functioning relies—at least in parts—on their exact primary structure and derivations from the original sequence can cause fatal malfunction. DNA, RNA, proteins and defined oligosaccharides are closely interacting in living matter as depicted in Figure I-1. Employing optimized molecular machinery, DNA can replicate itself as it can be transcribed into RNA. Messenger RNA (mRNA), in turn, can be translated into an amino acid sequence. After folding to a specialized protein, it can eventually be glycosylated with a defined oligosaccharide to e.g. demonstrate a sugar code as an identification tag. The whole building plan of living matter is encoded into DNA and information can be replicated and transcribed on the nucleic acid level. During protein synthesis this building plan is translated from mRNA unidirectionally into the amino acid sequence. The concept for this flow of genetic information is visualized with bold arrows in Figure I-1 and was phrased by Francis Crick:[17]

> *The central dogma of molecular biology deals with the detailed residue-by-residue transfer of sequential information. It states that such information cannot be transferred back from protein to either protein or nucleic acid.*

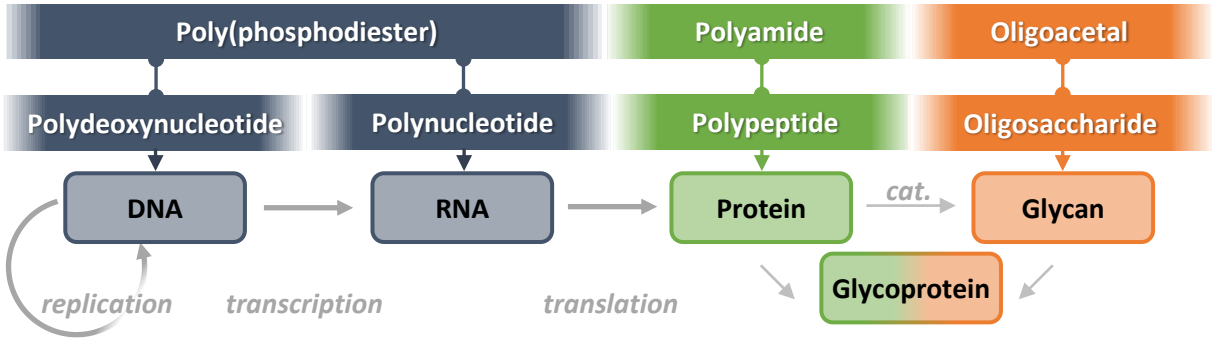
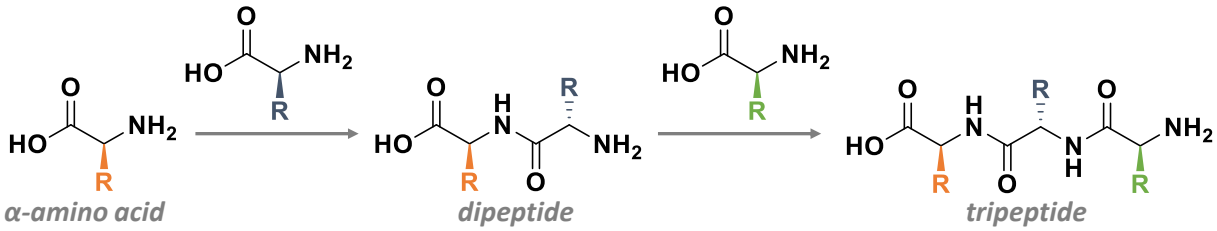


Figure I-1: Overview of sequence-defined biomolecules and their relation in molecular biology.

While DNA, RNA and proteins are made of long and linear macromolecules, sequence-defined oligosaccharides are in contrast short and often branched structures. For the sake of completeness, all representative classes of sequence definition in biological systems will be described. The focus is set on polymerizations leading efficiently to long sequence-defined polymers highlighting chemical processes from a polymer chemist's point of view rather than that of a biochemist.

2.1.1 Proteins



Scheme I-1: Schematic representation of a tripeptide synthesis by stepwise condensation of α-amino acids.

Proteins are polypeptides whose amide bonds are formed by condensation of α-amino acids as shown schematically for a tripeptide in Scheme I-1. The primary structure is defined by the sequence of the 22 proteinogenic building blocks (two of them not covered by the standard genetic code). Local regular structures such as α-helices or β-sheets are stabilized by hydrogen bonding of the amide backbone and are called secondary structure. The packing of such structure elements in domains localizes all atoms of the sequence and defines the tertiary structure whereas the quaternary structure is the eventual

local arrangement with other proteins. It is noteworthy that the primary structure of a protein inherently contains all information for its three-dimensional structure and complex interaction, yet, *in silico* prediction of folding is so far unaccomplished.[18]

2.1.1.1 Biosynthesis

Biosynthesis of proteins is characterized by the cooperation of almost 300 different macromolecules, many of these organized into the complex architecture of the ribosome. In brief, a mRNA molecule is delivering the sequential information that is then translated by the molecular machinery into the amino acid sequence of a protein; translation speed at 37 °C is about twenty residues per second. Figure I-2 (left) shows charged transport RNA (tRNA) that organizes the specific amino acid monomers at the ribosome in the order that is prescribed by the mRNA chain (shown in Figure I-2, right). mRNA serves thereby as a template containing the genetic information in trimer codes—termed codons—which are recognized by the anticodon of the specific tRNA. Amide formation during polymerization is further facilitated due to the activation of the carboxylic acid as an ester at the 3-OH of the tRNA's adenosyl 3'-end.[19]

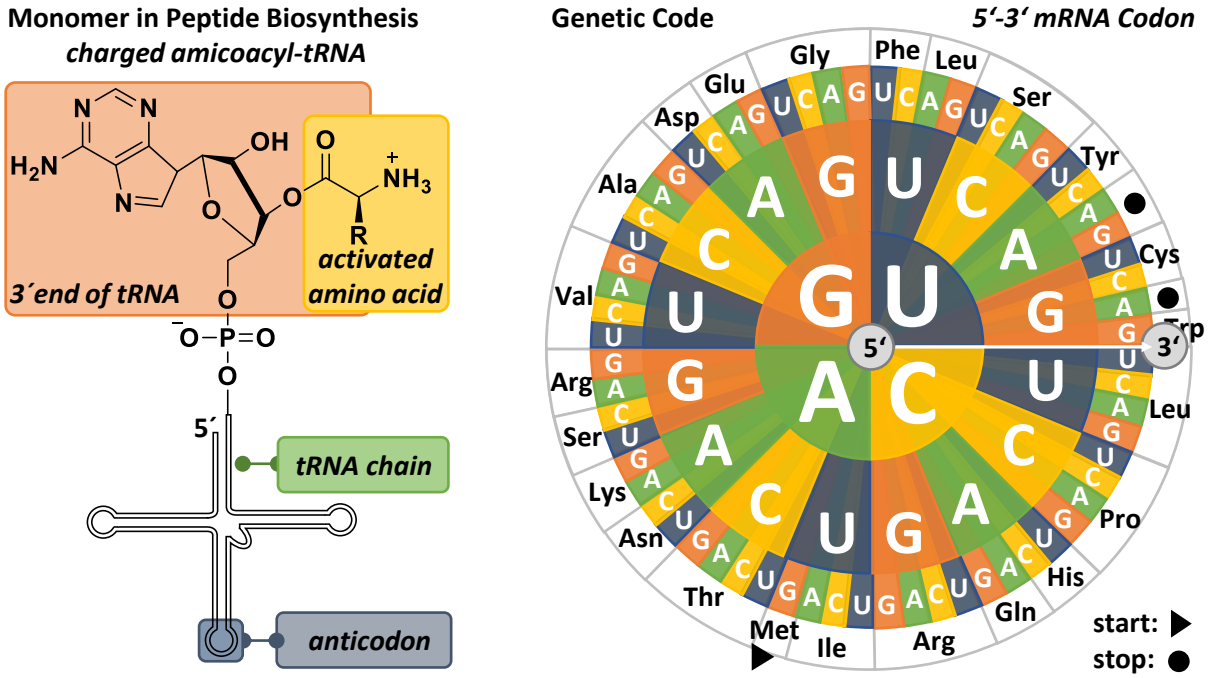


Figure I-2: Charged amicoacyl-tRNA, the active monomer for peptide biosynthesis (left) and genetic code assignment for mRNA-to-protein translation.

In sum, peptide biosynthesis requires a template (mRNA), guiding molecular machinery (ribosome) and amino acid specific macromolecules that read out the template and activate the monomeric units (tRNA). Two approaches are described below, mastering the challenges of sequence definition in different manners.

2.1.1.2 Solid-phase Peptide Synthesis

Synthesis of a tripeptide (e.g. Gly-Gly-Gly; see Scheme I-1, all R = H) in solution was first reported by Emil Fischer in 1903.[20] Longer peptides were synthesized by stepwise activation of the carboxy terminus to an acid halogenide and reaction with the next added amino acid.[21-22] Further development of protective group and activation strategies[23] led to the synthesis of bioactive peptides.[24-28] However, peptide synthesis in solution remained rather difficult and time-consuming. R. Bruce Merrifield who revolutionized the field commented during his Nobel lecture in 1984:[29]

> *It soon became clear […] that such syntheses were difficult and time-consuming and that a new approach was needed if large numbers of peptides were required or if larger and more complex peptides were to be made.*

In 1963 he reported on the *new approach* termed *solid-phase peptide synthesis (SPPS)*.[30] It takes advantage of attaching the growing polymer chain on an insoluble solid particle which allows for washing it free of reagents and by-products instead of recrystallizing the growing peptide after each step. A schematic representation of an example reactor used during the approach is depicted in Figure I-3 on the left. A filter ensures the solid support particles stay inside the reactor while diverse solutions can pass through.

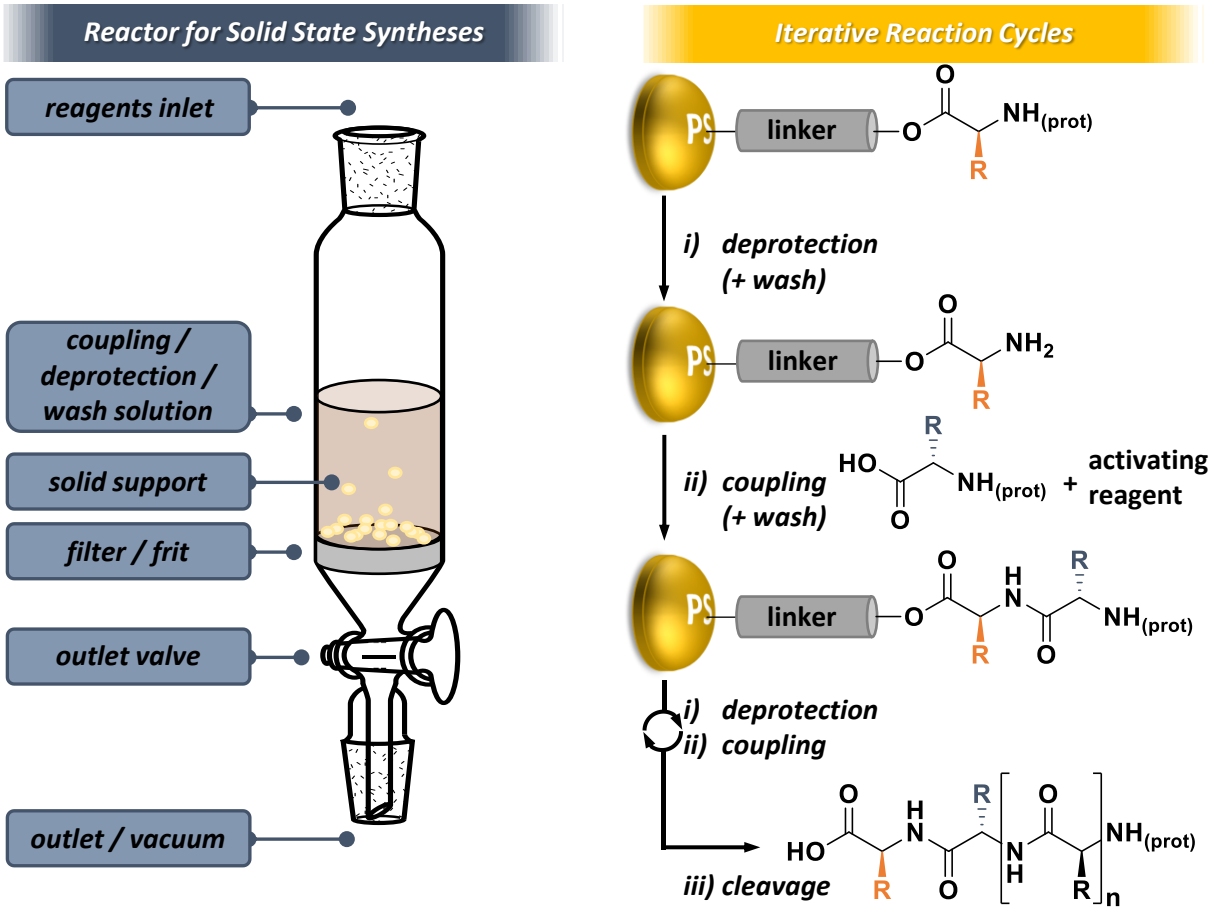


Figure I-3: Reactor for manual SPPS (left) and description of iterative reaction cycles on a solid polystyrene (PS) support (right).

The first employed solid support was a chloromethylated copolymer of styrene and divinylbenzene (200 – 400 mesh beads) immobilizing the carboxybenzyl (Cbz) *N*-protected amino acid via a benzyl ester on the resin. Nitration or bromination of the solid support further improves hydrolysis resistance of the linkage. Addition of a monomeric unit is described in Figure I-3 on the right and comprises at first Cbz-deprotection *i)* to obtain a terminal amine followed by condensation *ii)* with the added amine-protected amino acid upon activation with *N*,*N*'-dicyclohexylcarbodiimide (DCC). Cleavage from the support *iii)* yields the desired peptide sequence.[30] Envisioned since SPPS' inception, automation of the process[31] soon resulted in the synthesis of longer peptides like bovine pancreatic ribonuclease A (a 124-mer).[32]

Automated SPPS, pioneered by R. Bruce Merrifield, is commonplace today and a standard method for synthesizing peptides and proteins. It further enables the synthesis of artificial peptide sequences as well as the synthesis of other peptides which are difficult to express in bacteria.[33] The efficiency of

SPPS has continued to improve. The amino acid monomer enters the reaction cycles with a protected amine. The temporary protective group is commonly the acid-sensitive *tert*-butoxycarbonyl (Boc) group[34] or the base-sensitive 9-fluorenylmethyloxycarbonyl (Fmoc) group[35] which both can be removed efficiently. The condensation reaction has been optimized further. While initially employed carbodiimides like DCC have shown to be too reactive and can lead to epimerization, new activator generations of the benzotriazole family allow an efficient coupling under mild conditions.[36-39] New solid supports were developed to increase the length of accessible peptides[40] and linkers are available that allow for efficient liberation of the peptide sequence under various (orthogonal) conditions.[41] Progress in protective group chemistry gives furthermore the tools and strategies needed to reliably block amino acid side chains' functional groups from unwanted side reactions.[38, 42]

The development of native chemical ligation (NCL) further facilitates the access of long peptide sequences.[43] Here, a SPPS-made sequence with a thioester at the α-carboxyl group is connected via the formation of a *native* peptide bond to the cysteine residue of the amino terminus of another SPPS-made sequence. The ligation reaction is enabled by a transthioesterification *i)* that leads to a thioester intermediate that undergoes in a second step *ii)* a *S*,*N*-acyl shift[44] *ii)* as depicted in Figure I-4. Relying on a similar concept, other more recently developed ligation strategies allow cysteine-independent peptide couplings.[45]

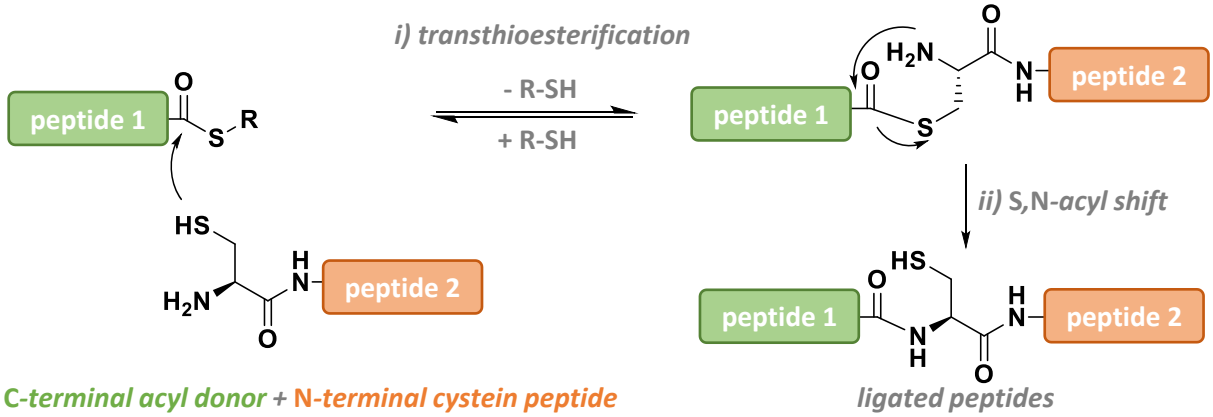


Figure I-4: Native chemical ligation of a C-terminal thioester with a N-terminal cysteine peptide.

2.1.1.3 A Supramolecular Peptide Synthesizer

In 2013 Leigh's group first reported on the design of a ribosome-inspired supramolecular peptide synthesizer that is shown in Figure I-5.[46-47] The artificial system presents the amino acid monomers as phenolic esters preordered along an axis. When activated, a macrocycle that is linked to a tripeptide and that can move freely between the end stopper and the first amino acid residue, reacts with its cysteine thiol in a transesterification reaction *i)* with the first amino acid. Subsequently, native chemical ligation transfers the amino acid to the *N*-terminus of the immobilized cysteine-glycine-glycine tripeptide *ii)*. This *S*,*N*-acyl shift regenerates the catalytic cysteine thiol and grows the tripeptide to a tetrapeptide. Moreover, the amino acid barrier has been removed, allowing the macrocycle to reach until the second amino acid that is presented on the axis. Transesterification *i)* and *S*,*N*-acyl shift *ii)* are repeated until the macrocycle is liberated from the axis *iii)* and after hydrolysis *iv)* the desired oligopeptide can be detected. Tandem mass spectrometry (MS/MS) confirms the defined sequence. The *machine* is rather slow compared to a ribosome and an amide formation takes approximately twelve hours while the oligomer's length is limited by the efficiency of long-range *S*,*N*-acyl shifts and the increasing migration range along the axis restricts the reactivity of the macrocycle additionally. However, it convincingly shows that humankind can mimic biological machinery and the conceptional design is generally applicable and can be adapted to synthesize e.g. oligoamides based on β-amino acid monomers.[48] Very recently it was additionally shown that the artificial ribosome can synthesize short

oligopeptides using a chain-growth-made axis.[49] By definition the dispersity of the axis is thereby translated into the dispersity of the attained oligopeptide sequence. Yet, the progress that has been made in controlled chain-growth polymerization and that is presented in subchapter *Chain-growth Polymerization* (pp. 26ff.) should theoretically enable the translation of very precise encoded synthetic sequences into precise oligopeptides using a man-made molecular machine.

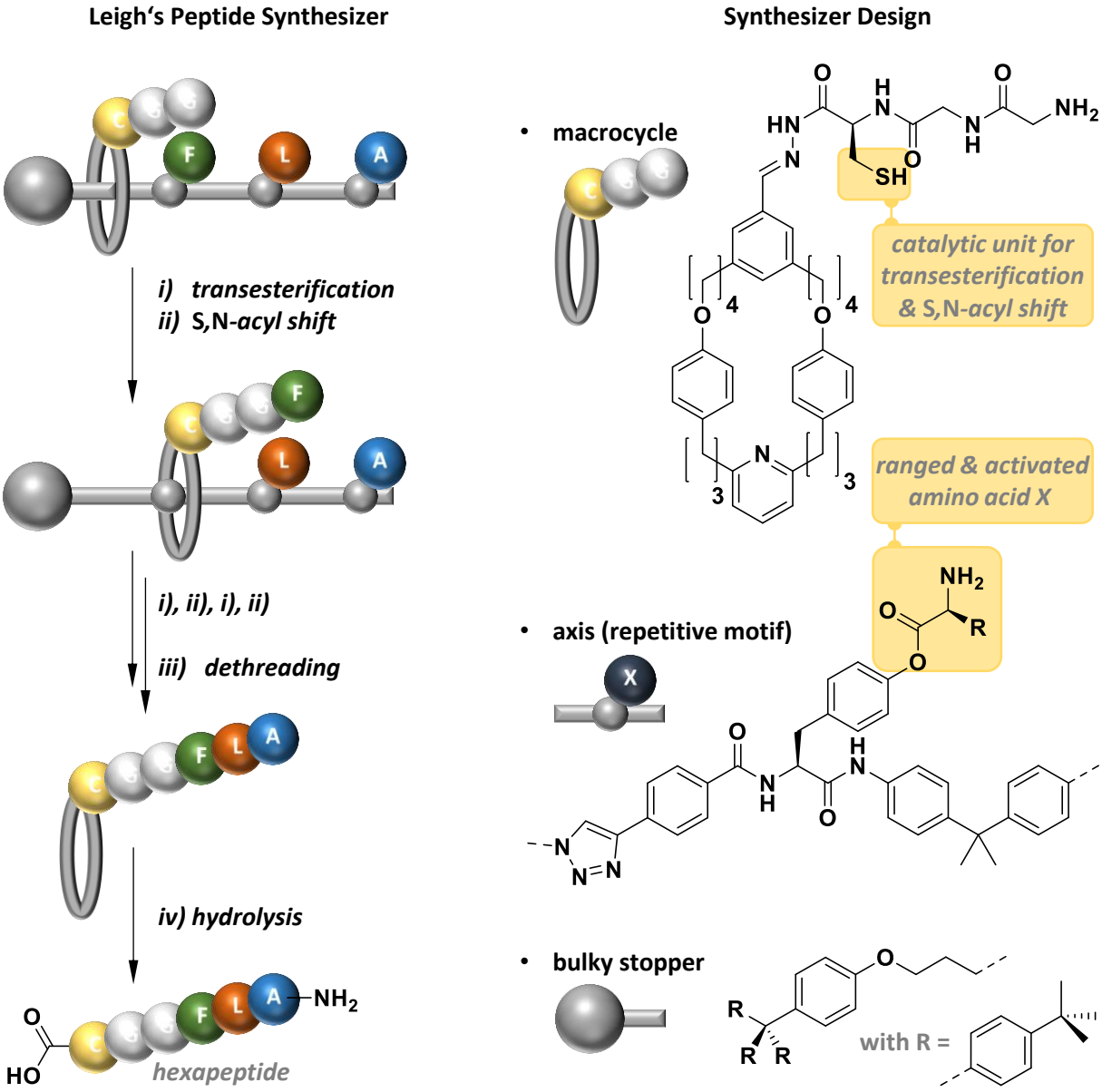


Figure I-5: Design of Leigh's supramolecular peptide synthesizer. Definition of abbreviation for amino acids: cysteine (C), glycine (G), phenylalanine (F), leucine (L), alanine (A).

2.1.2 Sequence-defined Oligosaccharides

In living matter, monosaccharides serve as building blocks in several forms and applications. Hexose sugars are polymerized to structural polysaccharides like cellulose by forming β-1,4 glycosidic bonds resulting in a rather stiff, rod-like conformation. Energy and carbon storage polymers as starch (α-1,4 glycosidic linkage) and branched glycogen (α-1,4 and α-1,6 glycosidic linkage) show a coiled architecture.[18] Moreover, oligosaccharides are used as informational molecules. While a hexapeptide that is built from the 20 common amino acids allows for (20^6 =) some million different primary structures and a hexanucleotide assembled with four nucleotide subuntis allows for (4^6 =) 4096 different structures, a hexameric oligosaccharide has a variability of many billions of different shapes.

20 different monosaccharide subunits are available that can be linked (1→2), (1→3), (1→4), (1→6), (2→3) and (2→6) each with either α or β configuration (see Figure I-6). These sequence-defined oligosaccharides are covalently linked to lipids or proteins and bind in intra- and intercellular recognition processes highly specific to proteins with process-dependent optimized affinity.[50]

2.1.2.1 Biosynthesis

Biosynthesis of sequence-defined oligosaccharides is supported by a complex molecular machinery. For example, during the synthesis of the core oligosaccharide that is to be linked to an asparagin side chain of a peptide forming a *N*-glycan, monosaccharides activated as phosphodiesters are added stepwise to a dolicholphosphate-immobilized structure on the cytosolic membrane side of the endoplasmatic reticulum (ER). Later, the synthesis switches to the ER lumen where dolicholphosphate-immobilized monomers are further added in a stepwise manner.[51]

2.1.2.2 Ongoing Development of Synthetic Access

Tools for the synthetic production of sequence-defined oligosaccharides are needed for the development of vaccines and antagonists. The monomeric building blocks are polyalcohols (see Figure I-6) and when forming an interglycosidic bond, regio- as well as stereoselectivity has to be considered. Development of protective group strategies and careful synthesis planning is therefore required.[52] Next to persistent protective groups, orthogonal temporary protective groups have to be used to allow the synthesis of branches. Participating neighboring protecting groups can, moreover, help to define the *trans*-stereospecific formation of a glycosidic bond as depicted in Figure I-6.

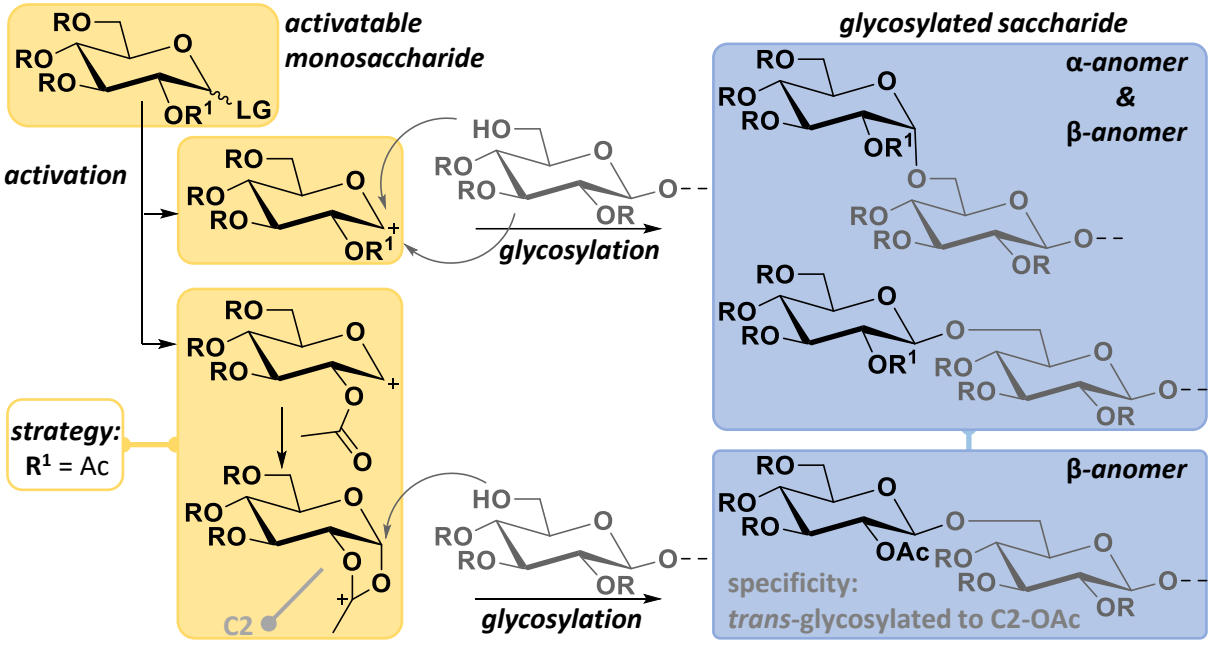


Figure I-6: Activation and protective group strategy for monosaccharide addition to a growing saccharide. (adapted from [54])

Yet, methods are available to stereoselectively generate glycosidic linkages. The yield and stereospecifity of the process depend on the steric and electronic nature of the monomer as it depends on the nucleophile and has to be reflected by the choice of reaction conditions.[53-54] Much effort was spent to synthesize sequence-defined oligosaccharides in solution.[55-57] Remarkably, in depth reactivity studies allowed the design of programmable one-pot oligosaccharide assembly strategies whose targets remain, however, relatively small oligosaccharides.[58-60] Automated synthesis on a solid support means certainly a break-through for the research field and advances have recently been

achieved.[61-64] Contrary to the introductory discussion concerning the immense theoretical structural variability, mammalian sequence-defined oligosaccharides occupy only a fraction of the available glycospace and a huge share of 75 % of the latter could be synthesized using a set of 36 monomers.[65] The first glycan synthesizer is commercially available and a routine synthesis of selected but complex sequence-defined polysaccharides is feasible.[66]

2.1.3 Nucleic Acids

From a polymer chemist's point of view, DNA and RNA are particular classes of sequence-defined poly(phosphodiester)s, whose monomeric units and their precise sequence induce manifold supramolecular properties intermolecularly as well as intramolecularly.

The carrier of genetic information, DNA, was first isolated and identified by Avery and coworkers[67] in 1944; its conformation was deduced nine years later by Watson and Crick[68] in 1953, relying not least on works by Franklin[69] and Wilkins:[70] DNA is found in biological systems as a double stranded helix as schematically represented in Figure I-7 on the left. Mother (in red) and daughter (in blue) strands interact via hydrogen bonding between the nucleobase pairs A=T (two bonds) and G≡C (three bonds) as indicated in Figure I-7 (center); the phosphodiester-linked ribose backbones show thereby a different sense of direction. In aqueous medium, the double helix between a strand and its anti-strand is guided by base recognition and held together not only by hydrogen bonding but also due to the hydrophobic effect in the non-polar DNA core and π-π interactions between the nucleobases that are arranged - contrary to the representation in Figure I-7 – in a plane perpendicular to the backbone. Ribonucleic acid (RNA) consists of the same building blocks, solely the pyrimidine base thymine (T) is replaced by its non-methylated analogue uracil (U) and the 2-hydroxy function on the ribose moieties is present. RNA single strands can fold to complex structures and are known to have catalytic function far beyond the already introduced mRNA and tRNA and are thus considered an important species in a world before today's *Protein World*.[19, 71]

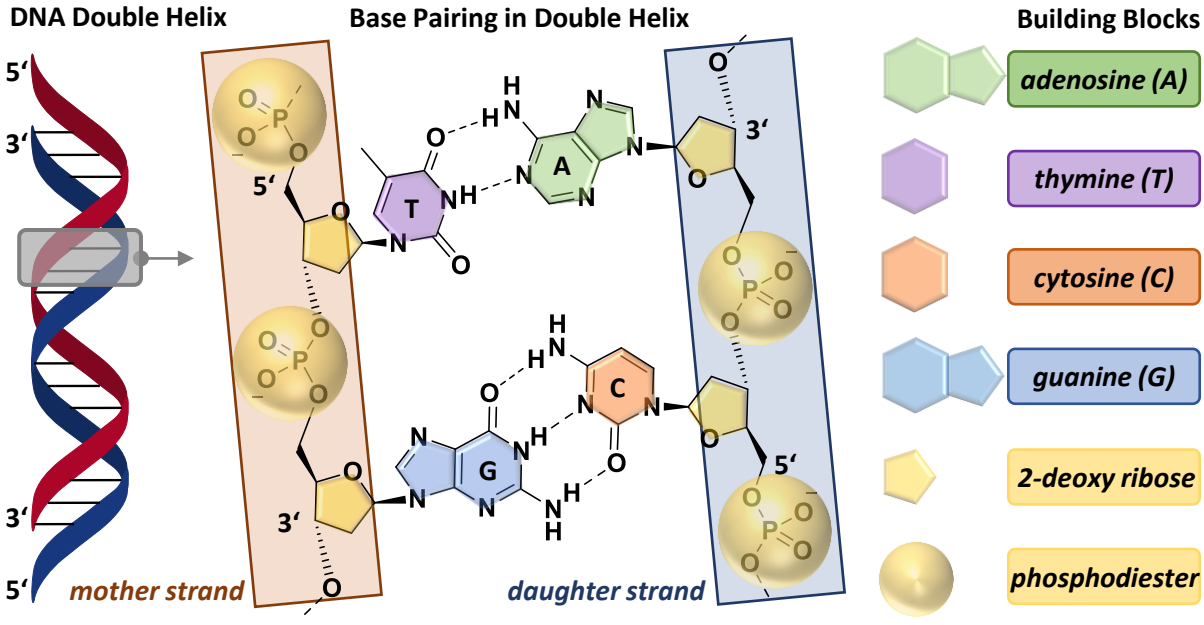


Figure I-7: Building principle of double stranded DNA helix.

2.1.3.1 Biosynthesis

In rare cases, DNA is synthesized RNA-dependent.[71] Usually, DNA is obtained by replication in a semiconservative manner, which means that each DNA strand serves as a template for the synthesis of a new strand, leading to two new DNA helices – each composed of a newly synthesized and its old

complementary strand[72] as it is schematically described in Figure I-8 on the left. The double helix is unwound by specific enzymes and a polymerase uses the single strands as templates for the catalyzed polymerization of the complementary strand. The activated monomer is the specific deoxynucleoside 5'-triphosphate providing pyrophosphate (diphosphate) as an attractive leaving group upon the nucleophilic attack of the 3'-hydroxy group of the growing chain as depicted in Figure I-8 on the right. The DNA polymerization occurs strictly in 5'→3' direction, thus along the lagging strand DNA is synthesized discontinuously in *Okazaki fragments* that are later interconnected.[73]

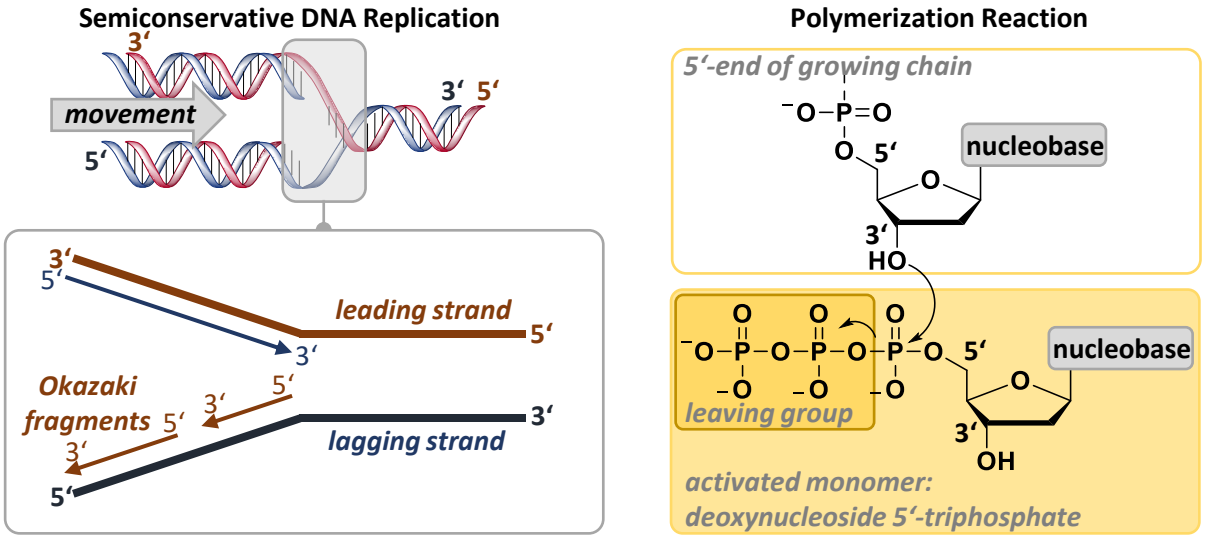


Figure I-8: Semiconservative DNA replication (left) and polymerization with an activated monomer (right)

2.1.3.2 Synthetic Access

Fully synthetic access of DNA and RNA is commonly achieved using phosphoramidite chemistry. Being a key tool of the present work in obtaining sequence-defined poly(phosphodiester)s, the approach is devoted the whole subchapter *Chemical Synthesis of Poly(phosphodiester)s* (pp. 33ff.).

2.2 Synthetic Polymers

Research on sequence-controlled polymers has become an emerging field in polymer chemistry. Called *embryonic*[74] and associated with the *next Holy Grail in polymer science*[75] earlier this decade, the development of synthetic polymers with a controlled monomer sequence has recently gained rising attention.[1, 3, 76-81]

Synthetic polymers are all polymers intentionally synthesized by humans. To date living matter demonstrates the most achieved strategies for the synthesis of sequence-controlled polymers. As presented in the preceding subchapter *Sequence Definition in Biological* Systems (pp. 16ff.), sequence control in biological systems reaches the ultimate goal of sequence definition where needed. Research on the synthetic access to sequence-defined biopolymers and the related development of universal strategies and tools were fueled by the impact of life sciences. From a polymer chemist's perspective these are early and very important contributions to the research area *sequence-controlled polymers* – even if the key players where not primarily considering themselves polymer chemists. The development of polymer chemistry and the development of polymers for life sciences, consequently, occurred mostly independently from each other during the 20th century due to the fact that they did not ask the same questions.[77]

It has been found that already a rough control over polymer microstructure results in polymer materials with unprecedented properties.[80] And it can indeed be expected that further research and development of techniques does not only deepen our understanding of global structure-property

relationships, it will rather lead to tailored properties beyond today's imagination as synthetic polymers and resulting properties will be much more diverse than those of today's yardstick, biopolymers which are limited by a confined availability of building blocks.[82]

There are three main classes of polymerization mechanisms that serve polymer chemists for synthesizing sequence-controlled polymers. According to a recently suggested terminology,[3] approaches in *step-growth polymerization*, *chain-growth polymerization*, and *multistep-growth synthesis* can be used to control the microstructure of resulting polymers as indicated in Figure I-9.

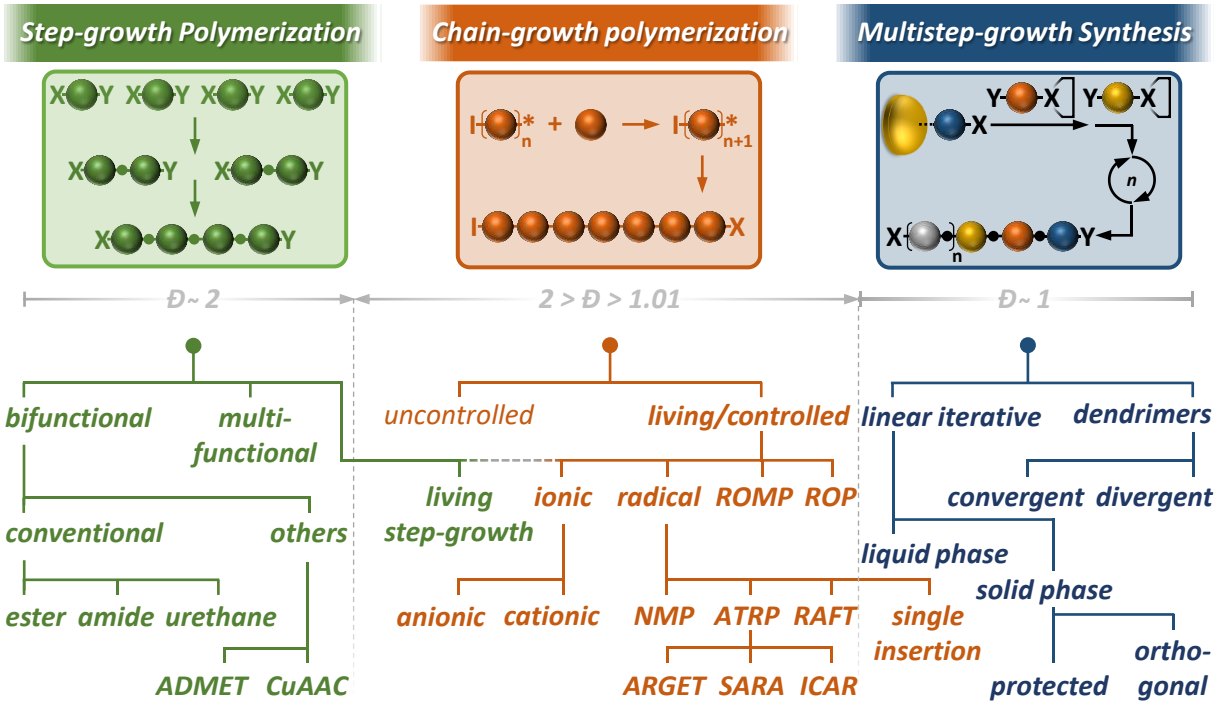


Figure I-9: Overview (non-exhaustive) of classes of mechanisms leading to sequence-controlled synthetic polymers. (adapted from literature[3] suggesting the terminology multistep-growth*). Definition of abbreviations: Step-growth polymerization: acyclic diene metathesis (ADMET), copper-catalyzed alkyne-azide cycloaddition (CuAAC); chain-growth polymerization: ring-opening metathesis polymerization (ROMP), ring-opening polymerization (ROP), nitroxide-mediated polymerization (NMP), atom transfer radical polymerization (ATRP), reversible addition-fragmentation chain-transfer polymerization (RAFT), activators regenerated by electron transfer (ARGET), supplemental activator and reducing agent (SARA), initiators for continuous activator regeneration (ICAR). Multistep-growth synthesis is shown with functional groups X and Y for clarity. Yet, more complex strategies require utilization of another letter code when presented in depth.*

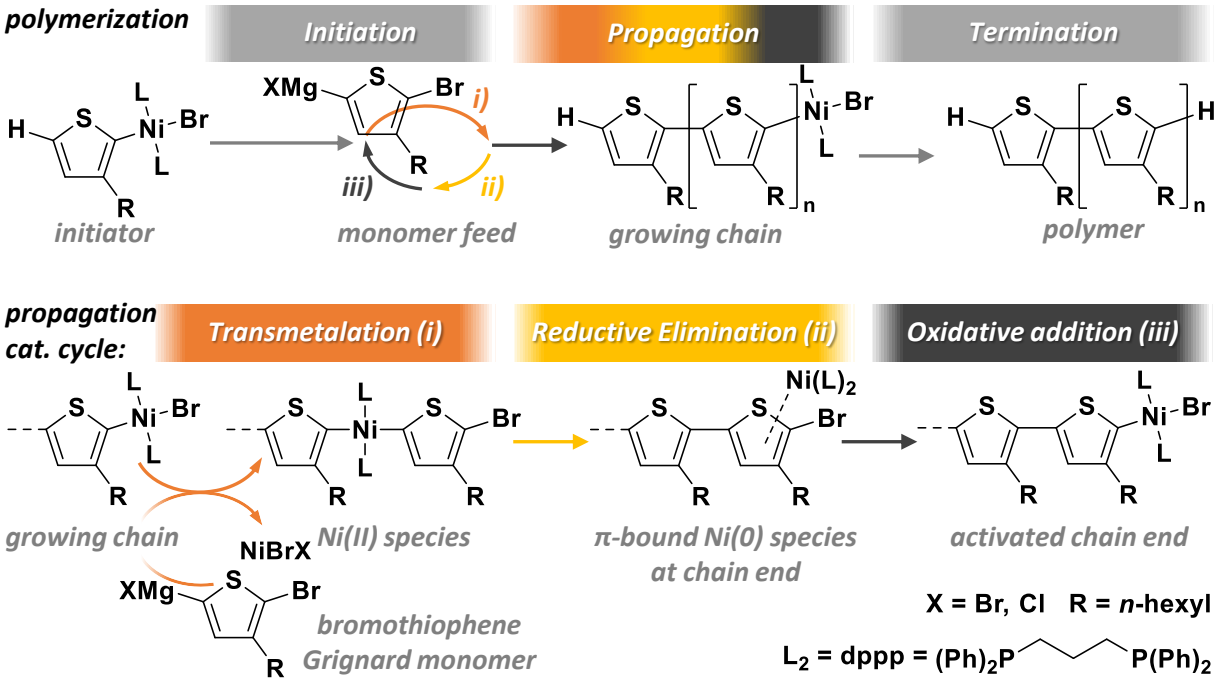
The degree of sequence control during a polymer synthesis is reflected by the product's *dispersity*, *Đ*. Different macromolecules are usually attained during a polymerization and the resulting polymer (mixture) is non-uniform. The dispersity can be based on the degree of polymerization (polymer length) or the molar mass. For a homopolymer or a long alternating copolymer these are proportional to each other; thus $Đ_X$ and $Đ_M$ are equal and simply referred to as *Đ*. Sequence-defined polymers are uniform and characterized by a defined degree of polymerization as well as a particular primary structure, hence a defined molecular weight. For their dispersity holds theoretically, $Đ = 1$.[83]

Depicted in Figure I-9, the different strategies for the synthesis of sequence-controlled polymers characterized by different degrees of microstructure control are described in the following subchapters and relevant examples are briefly described. In order to present sequence control for synthetic polymers lucidly, polymerization reactions relying on a template polymer (e.g. in protein engineering)—although a very promising and advanced strategy—will not be covered. The focus is set explicitly on initial sequence control.

2.2.1 Step-growth Polymerization

As depicted in Figure I-9 bifunctional monomers (XY) can randomly react with each other forming dimers and oligomers characterized by the same reactive end-groups as the initial monomers. In a stepwise fashion smaller fragments fuse and the length of chains is hence growing. The same principle is at work when copolymerizing monomers XX and YY. Multifunctional monomers do analogously form branched structures. The emergence of the first synthetic polymer bakelite[84] developed by Baekeland until 1907[85] was starting the *plastic age*. It was synthesized as other early materials like polyesters,[86] polyamides,[87] polyurethanes,[88] etc. during a step-growth polymerization. It can be derived from *Carothers' equation*[89] that such polymers show a rather high level of dispersity with $Đ \sim 2$. However, sequence control is easily obtained with respect to the arrangement of monomers. Copolymerization of XX and YY monomers leads imperatively to an alternating polymer $(AB)_n$ that, nevertheless, could be interpreted as a homopolymer, too. When telechelic polymers (and oligomers) or di-end-functional (heterotelechelic) analogues are prepared by other techniques, step-growth polymerization represents a convenient tool to access high molecular weight polymers with a periodic microstructure, e.g. $(ABC)_n$. Important studies performing the step-growth polymerization of precursors by means of ADMET,[90-92] carboxy-alcohol condensation,[93] metal-catalyzed step-growth radical polymerization,[94] CuAAC,[95] radical addition-coupling polymerization (RACP),[96-97] thiol-yne,[98] and thiol-ene coupling[99-100] yield sequence-controlled copolymers. In an elegant approach, You and coworkers report a one pot sequential monomer addition synthesis by primarily forming a telechelic DCBABCD (dithiomaleimide) oligomer that yields after polymerization with a diamine a $(DCBABCDE)_n$ polymer with a number-average molar mass, $\overline{M}_N$, of 25600 and a dispersity, $Đ_M$, of 1.24.[101] In all these examples the microstructure is precisely controlled regarding the interdistance of functional sites on a polymer chain. If reported, features differ severely from their more random analogues.

Step-growth polymerization is relying on the prerequisite that growing polymer chains have the same end-group reactivity as the monomers. Depending on the strength of particular effects, the chain end of an oligomer can have a higher reactivity than the monomer and thus a higher probability to be elongated. By tuning the applied conditions, a typical step-growth polymerization can thus be shifted to a chain-growth polymerization. This exotic strategy may exceptionally be treated in this subchapter as it historically developed when step-growth polymerization should be performed but a reduced monomer reactivity as compared to the growing chain was observed, either by a substituents effect[102] favoring the polymer-end to react or a specific catalyst transfer[103] to the chain-end after each monomer addition. The latter is particularly interesting because biosynthesis of sequence-defined polymers like ribonucleotides (compare Chapter I: 2.1.3.1 *DNA Biosynthesis*, p. 22) and peptides (compare Chapter I: 2.1.1.1 *Protein Biosynthesis*, pp. 17f.) is depending on a catalyst (here: enzyme) transfer to the growing chain's end activating the condensation reaction that is under catalyst-free conditions not occurring between the monomers.[104] It has to be added that the sequence-definition is in both cases derived from a template, however, the transformation of a random step-growth process to a controlled chain-growth polymerization is the key achievement. A man-made example of the very strategy was reported by the groups of Yokozawa[105-106] and McCullough:[107-108] Kumada-Tamao catalyst-transfer condensation polymerization occurs when a Ni(0) catalyst is added to a bromothiophene Grignard monomer. The reaction starts with an oxidative addition of Ni(0) to the C-Br bond. After reaction with a Grignard monomer in a transmetalation *i) trans-cis* isomerization (not shown) and subsequently reductive elimination occurs *ii)* resulting in a C-C bond formation and a π-bound Ni(0) species at the chain end that closes the reaction cycle with an oxidative addition *iii)*.[109] The initiation, propagation and termination with hydrochloric acid is depicted on top while the full catalytic cycle is shown on the buttom of Scheme I-2 including the crucial Ni(0) complex that is formed at the chain end after reductive elimination *ii)*. It has further been found that only when an initiator ArNi(dppp)X (X = Br or Cl) complex is used, unidirectional chain growth is ensured.[110-112]



Scheme I-2: Catalytic cycle of a Kumada-Tamao catalyst-transfer condensation polymerization of a Grignard monomer. Abbreviation ligand L_2: 1,3-bis(diphenylphosphino)propane (dppp).

2.2.2 Chain-growth Polymerization

As shown in Figure I-9, a chain-growth polymerization starts with an initiation reaction, the reactive transient species at the chain end—usually a radical, anion or cation—reacts repeatedly with a monomer transmitting the reactivity to the growing chain end during the propagation. Manifold termination reactions finally yield the polymer. The sequence control during a common copolymerization process is based on the reactivity ratios of the monomers present in the reaction mixture. Depicted in Figure I-10, mainly random, block and alternating copolymers are hence accessible.[75, 113] Particular differences in monomer reactivity ratios can lead to alternating primary structures[114] using e.g. donor monomer styrene and acceptor monomer maleic anhydride to synthesize a $(AB)_n$ copolymer[115] or tetrahydrofuran, epichlorohydrin and phthalic anhydride to synthesize a $(ABC)_n$ terpolymer.[116] The group of Sawamoto reported the use of template-prearranged monomer dyads and triads leading to alternating $(AB)_n$[117] and $(ABA)_n$[118] copolymers, respectively.

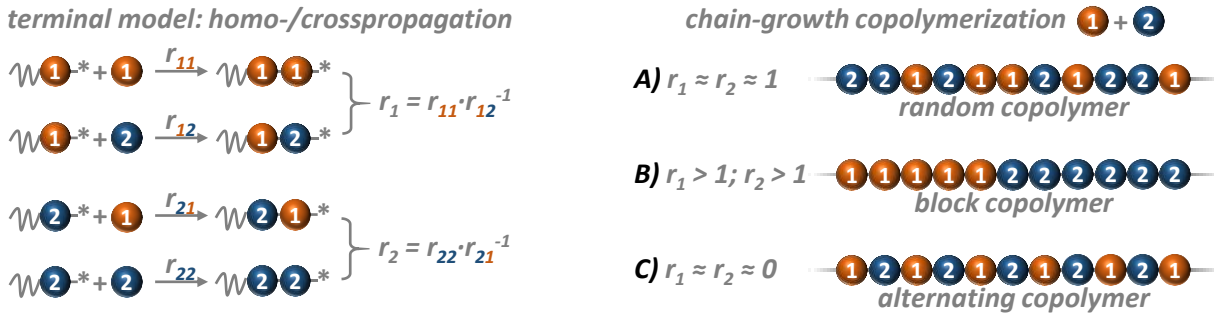


Figure I-10: Reactivity ratio-based copolymer sequence control. Non-exhaustive cases: A) no preference for homo- or crosspropagation; B) homopropagation favored; C) crosspropagation favored.

However valuable monomer reactivity-based and statistical concepts proof for copolymerization, an uncontrolled polymerization inherently implies chain termination and chain transfer reactions that lead to a high degree-of-polymerization dispersity, $Đ_X$. This is overcome by *controlled/living polymerizations* first reported by Szwarc in 1956[119] for anionic polymerization following mechanisms where the …

> *… process does not involve a termination step. […] The polymeric molecules then "live" for an indefinite period of time [however] […] any growth requires food. And the food for a growing polymer is the monomer. Consequently, if the supply of monomer is exhausted the growth is interrupted, although the living ends are potentially able to grow further if an additional amount of monomer is available.*[119]

Moreover, the rate of chain initiation is fast compared with the rate of chain propagation and the number of growing chains stays constant.[120] The polymer chains grow with the same rate and the chain lengths thereby stay similar resulting in a relatively low degree-of-polymerization dispersity, $Đ_X$. Theoretically, it is approaching $Đ_X = 1$.[121] Next to different generations of anionic polymerization[119, 122] the concept was reported for cationic polymerization[123-125] and for ring-opening metathesis polymerization (ROMP)[126-127] and a broad range of living polymer systems for synthesis of high molecular weight and low dispersity were known until the 1990's.[128-129]

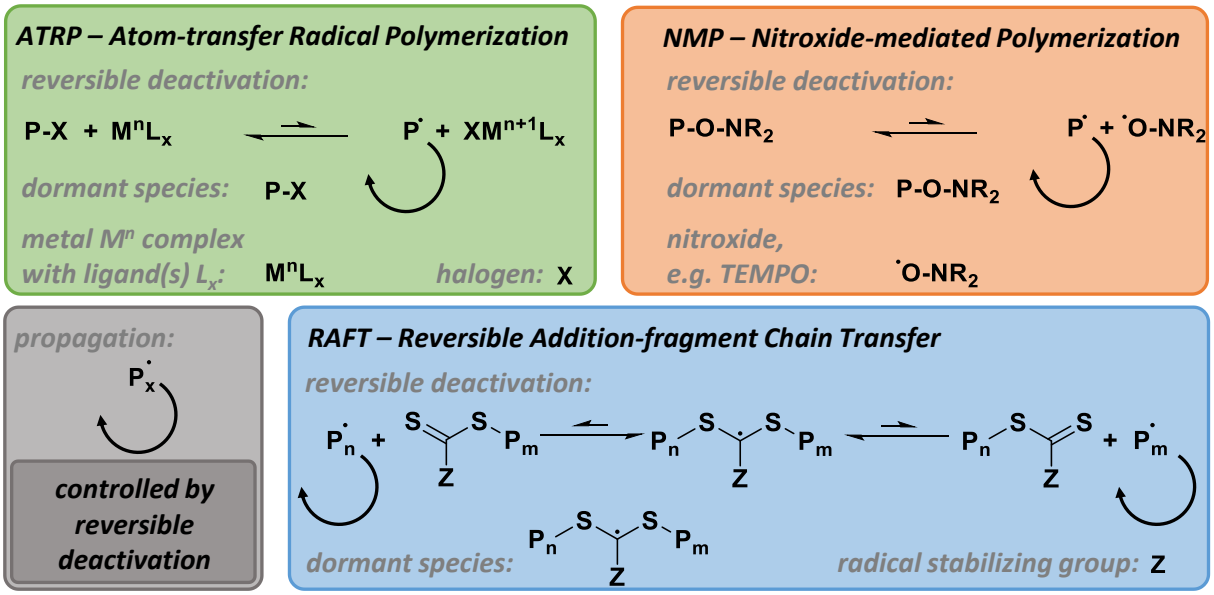


Figure I-11: Mechanisms of controlled/living polymerizations or reversible-deactivation polymerizations. A type of metal-catalyzed ATRP is shown.

Eagerly awaited[128] development of analogous controlled/living radical polymerization gave polymer chemists the tools to obtain sequence-controlled polymers in a process that is more tolerant toward a variety of functional groups und thus, the diversity and means to tune the polymer properties were greatly enhanced.[130-132] Namely ATRP,[133-134] NMP,[135] and RAFT[136] are processes that make radical polymerizations controlled and living by intermediately capping the growing polymer chain as a dormant species thereby reducing the overall concentration of the propagating radical chain end significantly. Dormant species and the reversible reactions leading to activated chain ends are depicted in Figure I-11. It is obvious that the initiator and the terminus of the dormant species are retained and sequence-control as well as control over chain end functionalities are obtained. This is a particularly important aspect for already above mentioned usage as (hetero)telechelic entries in a subsequent step-growth polymerization to access periodic polymers.

Thanks to particular differences in monomer reactivity ratios referred to already above, it is not only possible to control chain length distribution during living polymerizations. Hawker's group reported the one-step synthesis of functionalized block copolymers. During NMP of a 9 : 1 mixture of styrene and maleic anhydride at first preferential incorporation of maleic anhydride is observed leading to an initial styrene-*r*-maleic anhydride block. After reaching rather low maleic anhydride concentrations in the monomer feed, a terminal *quasi* homopolystyrene block accrues.[137] An elegant strategy pushing the conceptually related aforementioned approach to its limits yields tailor-made microstructures and

was recently reported by Lutz.[138] It was demonstrated that a variety of *N*-substituted maleimides could be precisely positioned on a (homo)polystyrene chain by time-controlled addition to an ATRP styrene feed.[139] As in Hawker's one-step synthesis, it is the acceptor monomer maleimide that has a low tendency of homopolymerization and shows a favored cross-propagation even if the comonomer is present in large excess.[140] Conceptually depicted in Figure I-12, *ultra-precise* insertion of one acceptor monomer per chain can be attained in situations when the donor-to-acceptor comonomer ratio is very low. Here, consecutive feeds of donor and acceptor comonomers are used in a NMP and other living polymerizations (activator regenerated by electron transfer (ARGET) ATRP and single electron transfer living radical polymerization (SET-LRP)).[2] With this strategy in hand, heterotelechelic copolymers for further CuAAC step-growth polymerization were attained.[95] Moreover, it could be shown that moieties can be inserted that result in a predefined functionality[141-143] and precise architecture[144-147] of the polymer chains. The predefined order of particular maleimide monomers along a polystyrene chain was apart from its impact in material science also discussed as a system for coding information on the molecular level.[148]

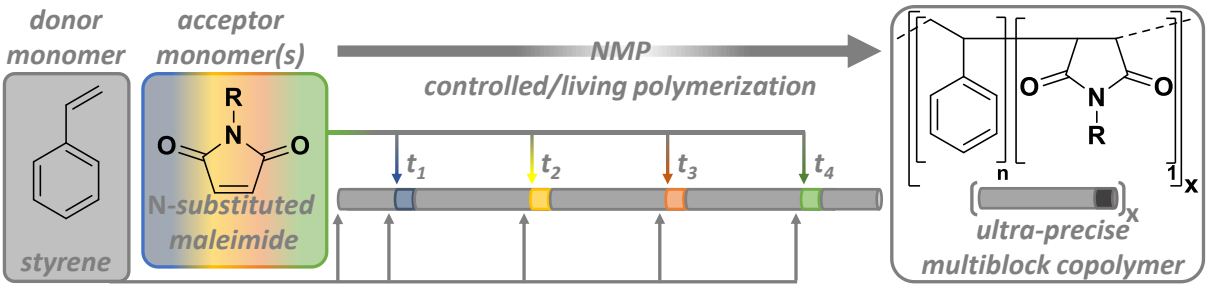


Figure I-12: Concept of ultra-precise N*-substituted maleimide insertion (colors, black) into a controlled growing polystyrene chain (grey).*

Conceptually related, it was shown by Perrier that sequence-controlled multiblock copolymers are accessible thanks to the high retention of livingness during RAFT polymerization. Acrylamido monomers were polymerized by RAFT until almost complete conversion was reached to establish the first homopolymer block. Subsequent reinitiation with another monomer feed—that again was consumed until almost complete conversion—added thereby one block after another in a sequence-controlled fashion.[149-150] The concept was further applied to RAFT polymerization of acrylamido and acrylate monomers at 25 °C yielding e.g. a dispersity of $Đ_M$ = 1.15 for an acrylate-based heptablock copolymer[151] and to ultrafast polymerizations without protection from air where multiblock acrylamidocopolymers were obtained within minutes with satisfying dispersities.[152] Very promising, the RAFT multiblock copolymerization process was performed in a looped flow process that can be scaled and that ensured reproducibly 99 % monomer conversion per feed and so-formed hexablock acrylamidocopolymers were characterized by a low dispersity ($Đ_M$ = 1.11) at a moderate molecular weight (M_n = 8100 g·mol^{-1}).[153]

2.2.3 Multistep-growth Synthesis

Multistep-growth synthesis is to-date the only available strategy granting the initial access to sequence-defined polymers. Other strategies are relying on sequence-defined templates (or—in the case of the supramolecular peptide synthesizer—on a sequence-defined axis) that had to be synthesized in a multistep-growth synthesis in advance. As depicted in Figure I-9, multistep-growth indicates that monomers are added stepwise during chain elongation enabling the definition of an exact primary sequence composed of a freely selectable number of monomers. Sequence definition comes with the price of additional purification steps that need to be added at least once per monomer addition cycle. Several examples from the recently progressing discipline[4] are discussed in the following sections.

2.2.3.1 Monomer Design and Stepwise Purification Strategy

Linear polymers can be synthesized using (AB) monomers. To prevent homopolymerization the B function enters the reaction as a protected species and only the growing chain's terminal B group can react with a monomer. An analogous orthogonal approach would employ (AA) and (BB) monomers that are alternatingly added. In both examples A can only react with B and *vice versa*. Called AB+AB and AA+BB approach, respectively, resulting polymers are formed by one repeating intermonomer linkage. The AB+CD approach relies on (AB) and (CD) monomers that are alternatingly added where functional group B reacts solely with C, and D with A, respectively. A derivation is represented by the AA+BC approach where the selected conditions (or chemoselectivities) result in a reaction between A and B or C and A. So-built polymers can but do not need to comprise different intermonomer linkages. Furthermore, a few examples exist using more than two monomer classes to construct a sequence-defined polymer chain.

The iterative steps in multistep-growth syntheses have to be separated from each other. E.g. before another $(AB)^y$ monomer can be added, all remaining $(AB)^x$ from a previous step has to be removed. That is the reason why—apart from approaches in which no support is used—the majority of strategies relies on solid or soluble support materials[1] that immobilize the growing polymer chain and facilitate the necessary purification steps dramatically. Recent examples that do not rely on supports have the advantage of possible scale-up. In terms of applying the strategies to obtain longer sequences and subsequently tuning the properties, some strategies remain at this stage questionable when reasonable yields and global applicability is wanted.[154] Nevertheless, promising support-free strategies are reported using stepwise chromatographic purification,[155-156] semi-automatic flow synthesis with in-line purification,[157] stepwise precipitation,[158] and preparative size-exclusion chromatography (SEC).[159]

The vast majority of reported multistep-growth syntheses benefits from fast purification by making use of supports. The general concept is explained in detail in subchapter *Solid-phase Peptide Synthesis* (pp. 17ff.): briefly, the solid support particles immobilize the growing chain and washing steps are employed to purify the growing sequence from excess reagents or soluble side products. Soluble supports, on the other hand, allow isolation of the supported growing chain by e.g precipitation in a non-solvent for linear polystyrene supports,[160] or by means of fluorous solid-phase extraction (F-SPE) when a fluorous purification tag[161] is used.

The range of reactions that can be employed in multistep-growth syntheses is infinite, yet, coupling times and yields are to be considered. The theoretical overall time t_{pol} needed for the synthesis of a polymer with chain length n is the sum of n-1 repeated cycle durations t_c consisting of x steps with reaction times of t_i, and is expressed in (1).

$$t_{pol} = (n-1)t_c = (n-1)\sum_{i=1}^{x} t_i \qquad (1)$$

[1] A much criticized downside of such support-assisted approaches is the limited synthesis scale related to the fact that material properties are an important subject to the polymer chemists' community. Justification usually points to the fact that the field is young and a new generation of polymer chemists is developing chemical methods and general strategies that can be used as a toolbox to find new polymer properties related to the defined primary structure. The here presented vision should convince the reader that it is indeed not anticipated to replace PET with sequence-defined polymers. Also living matter does still rely on roughly micro-controlled polymers like cellulose. As only a handful of DNA molecules governs the reader's retina bipolar cells' activity in this moment, very few tailor-made sequence-defined molecules are needed for already elaborated applications, such as data storage that is presented in chapter *Sequence-defined Polymers for Information Storage* (pp. 32-33). However, it can be kept in mind that scale-up is still a future option if needed.

The theoretical overall yield y_{pol} of the synthesis of a polymer with chain length n is the product of n-1 repeated cycle yields y_c consisting of x steps with stepwise yields y_i, and is expressed in (2).

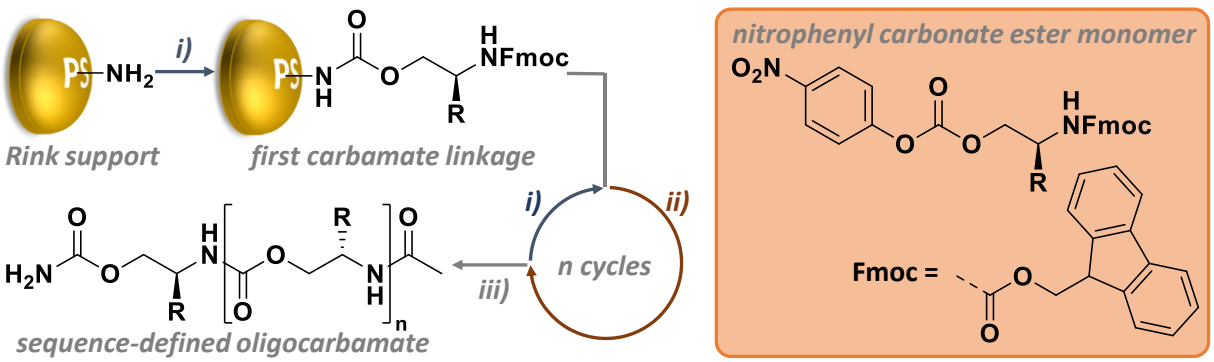
$$y_{pol} = {y_c}^{n-1} = \left(\prod_{k=1}^{x} y_k\right)^{n-1} \qquad (2)$$

The corresponding values for the cleavage from the support and, importantly if applicable, final or repeated purification steps have to be added to (1) and (2). Hence it is advantageous to choose reactions that are known to be fast and efficient in terms of product selectivity and hence their yield. E. g. if a special monomer to be inserted only a few times requires prolonged reaction times it will be tolerated in case of a satisfying yield. If long polymers are aimed for, reactions have to be performed near-quantitatively due to the exponential influence of polymer length on the synthesis yield, y_{pol}. Moreover, if non-quantitative reactions are employed, strategies should be designed precedingly that allow a purification of the desired sequence from missing-step sequences or other side products.

Phosphoramidite chemistry is a strategy for AB+AB multistep-growth synthesis that meets all the listed criteria. Its exceptional speed and efficiency make it the strategy of choice for the synthesis of sequence-defined poly(phosphodiester)s, the title polymer class of the present thesis. The strategy is covered in detail in the following subchapter *Chemical Synthesis of Poly(phosphodiester)s* (pp. 33ff.) to avoid repetition. The interested reader is furthermore kindly directed to the subchapter *Sequence-defined Polymers for Information Storage* (pp. 46ff.) for a more detailed report on multistep-growth synthesis of oligo(triazole amide)s (p. 48), oligo(alkoxyamine amide)s (pp. 49f.), oligocarbamates (pp.51f.), and poly(alkoxyamine phosphodiester)s (pp. 52f.).

2.2.3.2 Coupling Strategies for the AB+AB Approach

During the AB+AB approach, protective groups are used. It is thus important that the coupling conditions, the deblocking conditions and—if used—the cleavage of the linkage to the support material are orthogonal to each other. An early example of such system not targeting a biopolymer but a biomimetic polymer was reported by Schultz and coworkers in 1993. Using a solid Rink support an oligocarbamate backbone was build comprising side chains equal to those of the canonic amino acids because the interaction with biomolecules was under investigation. Activated by HOBt, the synthesis begins with the carbonate ester coupling by a nucleophilic attack on the carbonyl carbon by the amine producing *para*-nitrophenol as leaving group. The next step is the deprotection of the terminal amine with a weak base. The coupling *i)* and deprotection *ii)* steps are repeated until the desired sequence is attained. Yields of 99 % per cycle are reported. Finally the amine terminus is acetylated and cleaved from the resin *iii)* under acidic conditions as shown in in Scheme I-3.[162]



Scheme I-3: Early multistep-growth synthesis of sequence defined oligocarbamate by Schultz and coworkers. Carbamate coupling i): monomer, HOBt, diisopropylethylamine, N*-methyl-2-pyrrolidon; deprotection ii): piperidine,* N*-methyl-2-pyrrolidon; cleavage from Rink resin iii) a. acetic anhydride,* N*-methyl-2-pyrrolidon, b. TFA, triethylsilane. Abbreviations: 2,2',2''-trifluoroacetic acid (TFA).*

The strategy was also used with the photolabile protective group nitroveratryloxycarbonyl (NVOC) to synthesize an oligocarbamate library. The underlying principle is covered in subchapter

Phosphoramidite Chemistry for *Microarray* Production *(*pp. 44f.). Structurally related to carbamates, sequence-defined oligoureas are synthesized. The analogue monomers are *N*-protected carbamates with a good leaving group[163-165] or e.g. *in situ* generated isocyanates.[166] Other classes of sequence-defined oligomers use the AB+AB approach to build non-peptidic oligoamides[167-170] using the SPPS strategies, oligoamines by various strategies,[171] oligostesters,[172-173] and conjugated oligoarylenes.[174-178]

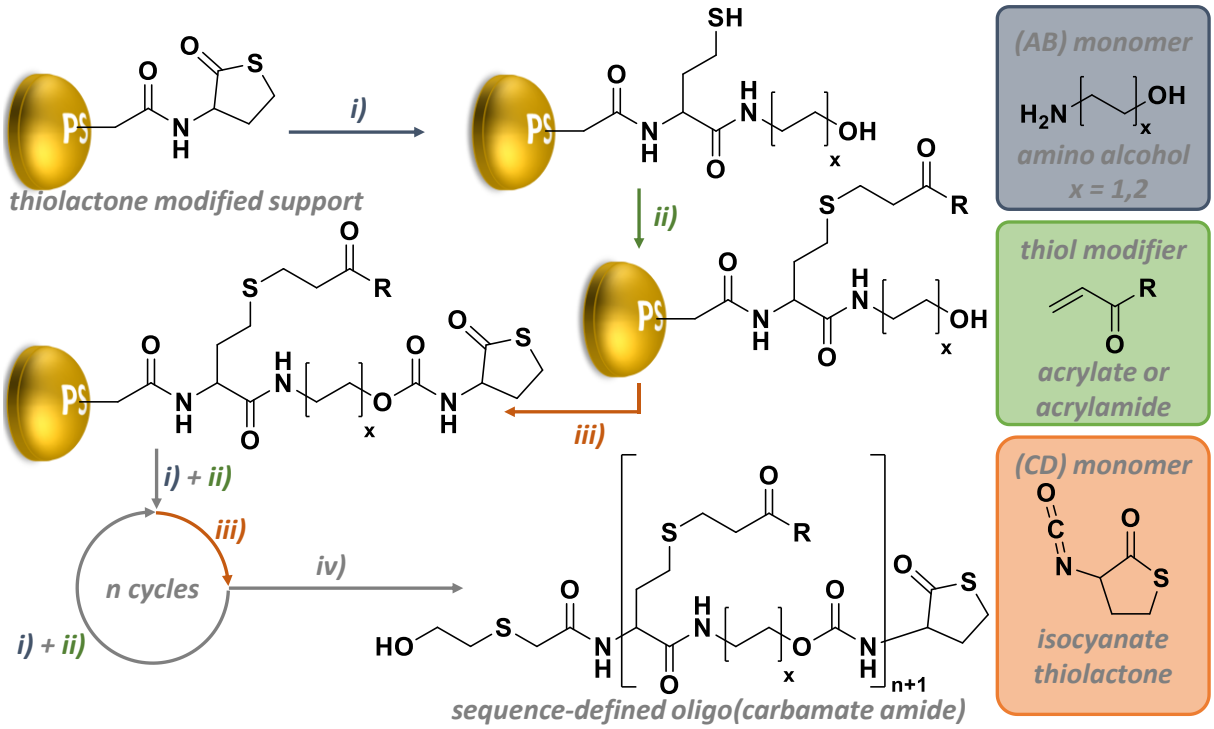
2.2.3.3 Coupling Strategies for Orthogonal Approaches

Orthogonal approaches are protective group-free in the sense that at least during the repeated steps of polymer backbone formation no protective group is employed. That can reduce the overall reaction time and is economically beneficial. Some strategies that rely on protective groups can be transformed to an approach using latent groups. E.g. amines and azides can be transformed into each other. Instead of protecting the amines in a polyurea synthesis, an azide monomer can be used that is efficiently converted to an amine in the following step.[179-180] Oligo(1,3-triazole)s would be accessible when monomers exhibiting an azide and a protected alkyne moiety are used. Here, an amine monomer can be used that is convertible to the azide upon demand.[181]

An interesting submonomer strategy was introduced to the field of peptoid polymers which are peptidomimetic poly(*N*-substituted glycine)s. Before, it was the state-of-the-art to synthesize polyglycine by means of SPPS (AB+AB approach) and *N*-substitute the formed amide after each coupling cycle. Zuckerman proposed to perform the amidification with bromoacetic acid. The formed bromide can be reacted with a primary amine of choice to regenerate a terminal amine. This protocol that follows an AA+BC approach results in cycle yields above 99 % and can be performed on an adapted peptide synthesizer.[182]

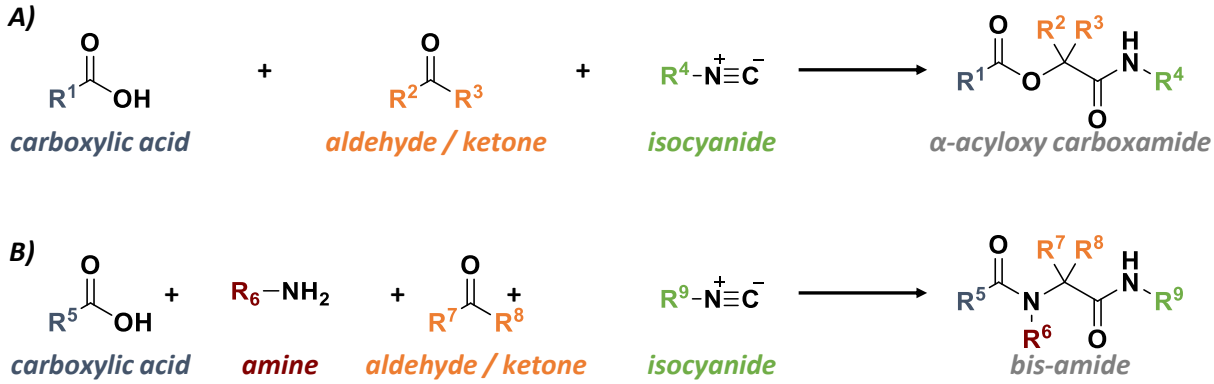
AA+BB, AA+BC and AB+CD approaches are particularly interesting to the field of sequence-controlled polymers because they enable polymer growth in every step. If the monomers have at least three of the functional groups A, B, C, and D locally separated from each other it is possible to easily introduce two different chemical linkages into a polymer chain. Some examples of these approaches are presented in detail in subchapter *Sequence-defined Polymers for Information Storage* (pp. 46ff.). At present a large variety of examples are reported that combine two fast and ideally quantitative reaction steps to yield sequence-defined polymers.[76]

An automated AB+CD approach was reported by the group of Du Prez. The same monomers are repeatedly coupling. Tailored functionalization was achieved by adding a thiol modifier to the (AB) monomer coupling mixture. The immobilized thiolactone ring is opened *i)* chemoselectively with the amine function of the (AB) monomer. The resulting thiol immediately reacts with the acrylic double bond of the modifier in a thiol-ene reaction *ii)*. This step allows sequence definition by addition of several different thiol modifiers in following cycles. A thiolactone is retrieved by conversion of the terminal alcohol with α-isothiocyanato-γ-thiolactone via a carbamate linkage. The cycles can be repeated—adding the thiol modifier of choice during the ring-opening reaction—until the desired sequence is attained. In the automated approach that is shown in Scheme I-4, each reaction cycle lasts for about 90 minutes to yield highly pure oligomers.[183] In general, several reactions can be performed at the same time without intermediate reactions steps. The important condition for that is perfect chemoselectivity.[184] Du Prez exploited this feature of the multicomponent amine-thiol-ene conjugation to introduce a side chain modification. Furthermore, the concept can be used to rapidly construct a sequence-defined backbone. This was shown by Meier and colleagues employing the



Scheme I-4: Synthesis of sequence-defined oligo(carbamate amide)s including a amine-thiol-ene conjugation in one step. Amine-thiol-ene conjugation i) + ii): chloroform, (AB) monomer, thiol modifier (ratio amine to acryl compound: 1 to 2); iii) chloroform, (CD) monomer, DBTL; cleavage from resin iv): TFA, DCM. Abbreviations: dibutyltin dilaurate (DBTL).

Passerini three-component reaction (P-3CR). The latter is the reaction of an aldehyde or ketone with a carboxylic acid and an isocyanide, forming an α-acyloxy carboxamide as it is shown in Scheme I-5.A. The indicated substituents have to be chosen in a way that P-3CR can be incorporated into a multistep-growth approach. E.g. if the aldehyde (R^3 = H) comprises a terminal double bond at the substituent (R^2), the P-3CR can be followed by a thiol-ene reaction that introduces a carboxy end to the growing chain. The reaction cycle would be terminated by adding the aldehyde and an isocyanate of choice for the next P-3CR. The so-attained sequences vary by the different substituents R^4 from the isocyanate insertions.[185]



Scheme I-5: Overview A) Passerini three-component reaction (P-3CR); B) Ugi four-component reaction (U-4CR).

A modification of the strategy employed different aldehydes for sequence-coding and an isocyanate was used that comprises a benzyl-protected carboxylic acid on the chain end. Thus, each P-3CR was followed by carboxylic acid deprotection to regenerate the acid before the coding aldehyde and the *terminating* isocyanate were added again.[155] A mixed multicomponent AB+AB approach is then realized. Interestingly, higher yields were obtained compared to P-3CR/thiol-ene strategy. As the P-

3CR is tolerant towards several functional groups, it can be employed together with other orthogonal protective group-free chemistries to yield sequence-defined polymers.[186-187]

Depicted in Scheme I-5.B, an aldehyde (R^7 = H) or ketone can condense together with a carboxylic acid, an amine and an isocyanide during the Ugi four-component reaction (U-4CR) to form a bis-amide. The reaction was applied to precision macromolecular chemistry as described before connecting the multi-component reaction steps by a thiol-ene coupling (R^5 = CH_2CH_2SH and R^7 = H; R^8 = $(CH_2)_8CH{=}CH_2$). It is just feasible to incorporate two coding monomers—namely amine and isocyanate components, varying R^6 and R^9, respectively—per U-4CR step.[188] The straight-forward access to a bis-amide motif is particularly interesting for peptidomimetics. Consequently, the groups of Meier[189] and Becer[190] combined different U-4CRs with iterative peptide synthesis.

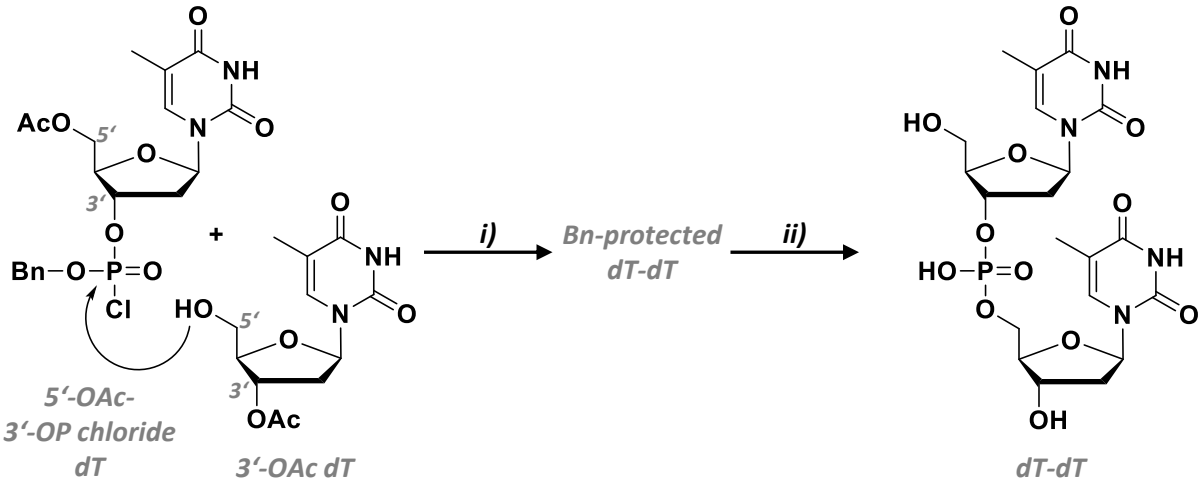
3 Chemical Synthesis of Poly(phosphodiester)s

After the understanding of DNA structure, 30 years of intensive research had to pass until a commodity approach for synthetic access to sequence-defined poly(phosphodiester)s was developed: Phosphoramidite Chemistry.[191] The development was fueled by parallel progress in other scientific areas and while early synthetic protocols already allowed important breakthroughs in chemical biology, it was not until the 1980's that phosphoramidite chemistry was available.[192] It meant such an immense improvement of the state-of-the-art that benefited biologists, biochemists, and molecular biologists for their research that the contributors called their common achievement a *Gift to Science*.[193] The evolution of phosphoramidite chemistry is described below covering the historic development of relevant chemical approaches and incorporation of neighboring disciplines' findings. Then, phosphoramidite chemistry is described in detail and modern variations and applications are presented.

3.1 Historical Development and Progress

Taking into account the chemical structure of DNA (see Figure I-7), the formation of the phosphodiester bond between the 3'-hydroxy function of one monomer and the 5'-hydroxy function of the other is identified as the most promising reaction for polynucleotide formation. Nature, too, employs an activated 5'-triphosphate monomer and guides the phosphodiester formation by a nucleophilic attack by the 3'-hydroxy function on a P(V) center as described in subchapter *DNA Biosynthesis* (p. 22; see also Figure I-8).

3.1.1 First Dinucleotide

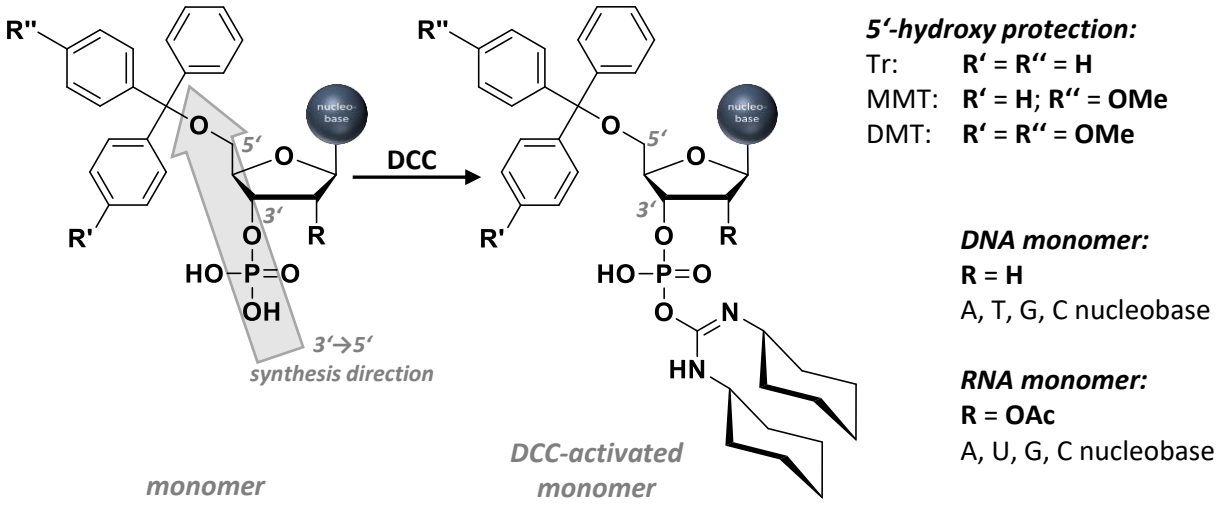


Scheme I-6: Synthesis of the first dinucleotide by Todd. coupling i): 2,6-lutidine; cleavage ii): hydrolysis.

In 1955 Michelson and Todd reported on the first successful dinucleotide synthesis. Dideoxythymidine dinucleotide was obtained by condensing 3'-OAc-protected deoxythymidine with 5'-OAc-protected deoxythymidine exhibiting phosphorochloridate at the 3'-position. As depicted in Scheme I-6, the obtained phosphodiester after coupling *i)* could be deprotected *ii)* to yield the first man-made dinucleotide. The experiment was a historic milestone because the dT-dT product was found to show the same characteristics as dideoxythymidine dinucleotide isolated from natural product digestion experiments and, thus, ultimately confirmed the structure of the 3'-5'-internucleotide phosphodiester linkage in DNA.[194]

3.1.2 Phosphodiester Method

Being a former member of the Todd research group, Khorana was investigating carbodiimides and their impact for synthesizing biologically relevant pyrophosphates.[195] By chance he found a new method for linking two deoxynucleotides in 1956,[196] later termed the phosphodiester method. A phosphomonoester, 5'-phosphorylated deoxythymidine can be activated with *p*-toluenesulfonic acid (TsOH)[196] or DCC[197]—the latter method resulting in higher yields—for esterification with the 3'-hydroxy group of another dT monomer. Subsequently, the Khorana group developed a strategy to access DNA and RNA oligonucleotides using the approach. Contrary to the initial discovery, synthesis is then performed in 3'-5'-direction enabling the more active primary 5'-hydroxy group to perform a



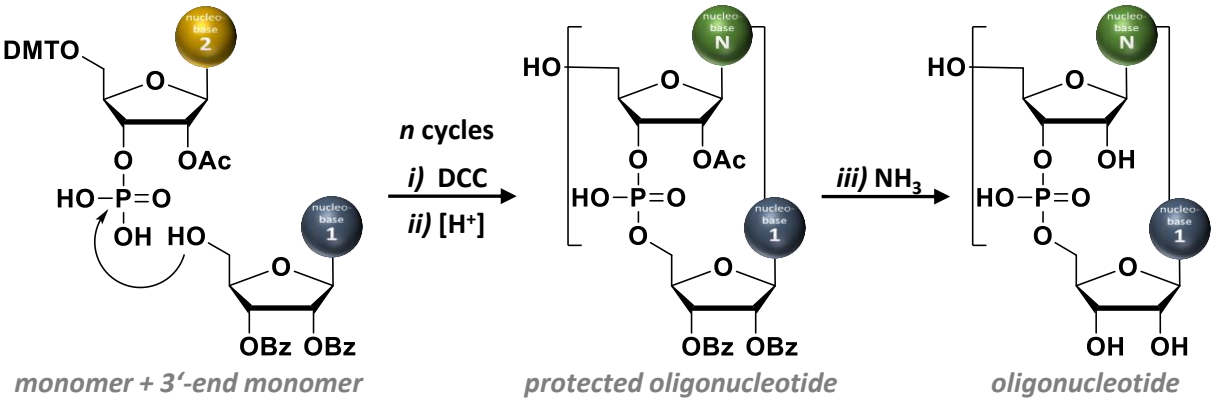
Scheme I-7: Monomer activation with DCC for phosphodiester method.

nucleophilic attack on the DCC activated 3'-phosphate monoester. The acid activation with DCC that is usually performed *in situ* is shown below in Scheme I-7. The leaving group is (insoluble) *N*,*N*'-dicyclohexylurea.

The 2'-hydroxy functions of RNA monomers are protected as semi-permanent acetic esters. Also the particular nucleobase is protected with a base-labile protective group while, importantly, the 5'-hydroxy function is temporarily protected with an acid-labile (substituted) trityl group.[198] Trityl (Tr), 4-methoxytrityl (MMT), and 4,4'-dimethoxytrityl (DMT) groups can be efficiently cleaved from the alcohol with acid. The latter DMT group shows a characteristic absorbance when cleaved that allows for a quantitative determination of the cleavage process.[199] But also does the DMT group allow for fast and complete cleavage under mild acidic conditions (pH 4-5) which is crucial due to nucleobase lability under more acidic conditions.

The first monomer with the 3'-terminus exhibits the 3'-hydroxy function (and the 2' hydroxy function for the synthesis of RNA analogues) permanently protected as a benzoyl ester. The primary 5'-alcohol can react with an incoming monomer in step *i)* upon activation with DCC. In step *ii)* DMT-deprotection

with acid regenerates the primary alcohol at the 5'-terminus. Monomer addition *i)* and alcohol deprotection *ii)* are repeated with different RNA or DNA monomers until the desired sequence is attained and finally, the oligonucleotide is obtained after deprotection *iii)* in base.[198, 200-201] The concept is shown in Scheme I-8.



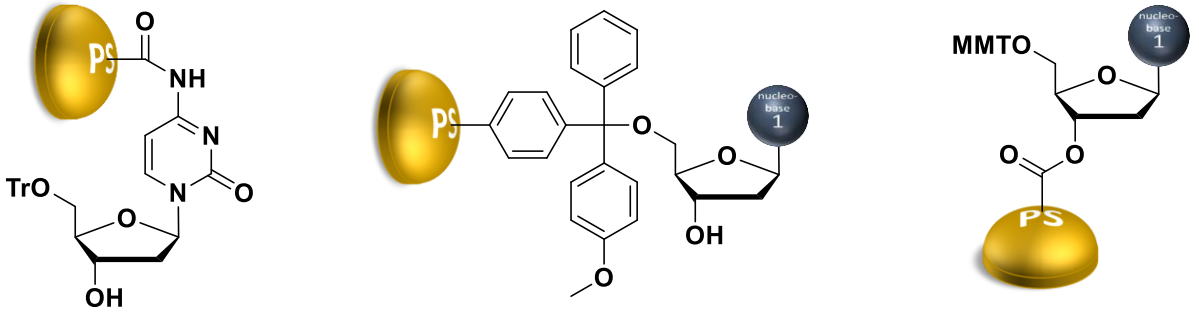
Scheme I-8: Oligonucleotide synthesis using the phosphodiester method.

Synthesis of oligonucleotides employing the phosphodiester method is time-consuming and characterized by an inherent shortcoming. During iterative synthesis the growing amount of phosphate groups leads to an increasing tendency of hyperbranching as the phosphodiester moieties can react with the monomer in a competing reaction, too. For that reason each iterative cycle (step *i)* + step *ii)*) includes a chromatographic purification of the oligonucleotide. With increasing chain length the purification effort becomes laborious.

The very long-lasting benefit of Khorana's contribution was the elaboration of useful protective groups, in particular nucleobase protective groups. (Subhapter *Appropriate Protective Groups* (pp. 42ff.) summarizes the state-of-the-art.) It was, thus, possible to employ the phosphodiester method for the synthesis of any desired RNA or DNA oligomer as protective group strategies for all nucleotide monomers were developed and orthogonal 5'-hydroxy deprotection was optimized. These impressive achievements were paving the way for Khorana and colleagues to contribute to the complete deciphering of the genetic code,[202] the full assignment of codons in mRNA (see Figure I-2 in subchapter *Protein Biosynthesis*, pp. 17f.) and the group used the tool to first synthesize full genes artificially.[203-204]

3.1.3 Advent of Solid Supports in Oligonucleotide Synthesis

The second player who next to Khorana moved the synthesis of oligonucleotides tremendously forward was Letsinger. Independently from Merrifield, he was interested in synthesizing peptides on solid supports.[205] His findings were published in 1963, few months later than his competitor's. He consequently left the field of condensing amino acids as he envisioned the impact that his laboratory's experience in solid phase syntheses could have on the synthesis of oligonucleotides.[206-207] Letsinger's group reported the first internucleotide linkage performed on a solid support that is shown on the left in Figure I-13. The protective group strategy was inspired by Khorana's findings. The first monomeric unit, 5'-O-trityldeoxycytidine, is here immobilized by amidification of the nucleobase's amine function. Cleavage from the support could later be achieved by amide hydrolysis with sodium hydroxide.[208-209] Also Khorana contributed with an acid-cleavable support, where the 5'-O-trityl group is attached to the soluble polymer support.[210] The analogous solid support[211] was reported by others and is shown in the middle of Figure I-13. Interestingly, another major player in the field, Caruthers, contributed during his early years as Letsinger's PhD student with an optimized support that immobilizes any 5'-OMMT-protected deoxynucleotide on the 3'-terminus (see Figure I-13 on the right).[212]



Cytidine-Support
- ***cytidine-specific***
- ***nucleobase bound***
- ***base cleavable***

MMT-Support
- ***universal***
- ***5'-O-bound***
- ***acid cleavable***

3'-ester Support
- ***universal***
- ***3'-O-bound***
- ***base cleavable***

Figure I-13: First generation of solid supports for oligonucleotide synthesis.

The 3'-ester support was shown to tolerate the conditions of the phosphodiester approach (see Scheme I-8) and a deoxythymidine trimer could be synthesized. DCC activation of the monophosphate (see Scheme I-7) was presumably not beneficial because the *N*,*N*'-dicyclohexylurea leaving group would precipitate with the support and a reagent that can be washed out is more beneficial. Thus, the activator 2,4,6-triisopropylbenzenesulfonyl chloride was used and it turned out that the synthesis was faster and more efficient compared to the outcome when using the cytidine support or the MMT-support (compare Figure I-13). Still, performing the same synthesis with Khorana's soluble MMT-support in the solution phase resulted in better yields compared with employing any of the solid supports.[212]

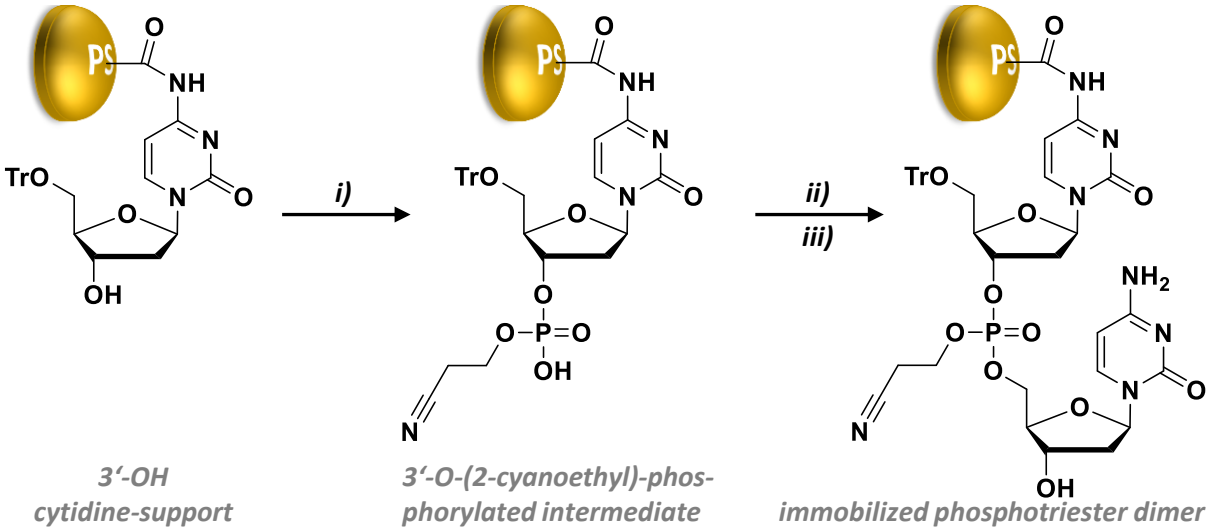
The advent of solid supports in oligonucleotide synthesis facilitated the synthesis of short sequences but it was neither a tool against the hyperbranching side reaction, nor did it result in higher synthesis yields. More potent protocols had to be developed to make the process on the solid phase convenient and efficient.

3.1.4 Phosphotriester Method

When Letsinger entered the field of oligonucleotide synthesis, he came up with a new approach to avoid hyperbranching. The phosphodiester backbone that is formed during the phosphodiester method preserves a certain reactivity. Implementation of a phosphate protective group, e.g. the 2-cyanoethyl group, can avert the side reaction. Thus, instead of a ribosyl monophosphate, the analogue protected diester gets activated to react with another ribosyl alcohol as shown in Scheme I-9. The so-formed unreactive phosphotriester backbone intermediate is the eponym for the approach.

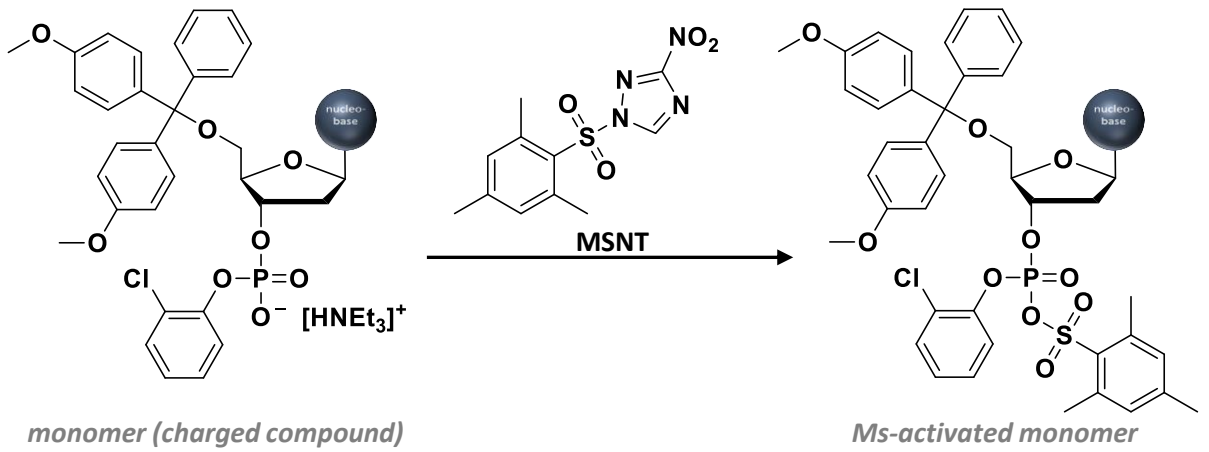
The phosphotriester method was first executed in a stepwise fashion that is shown in Scheme I-9. Here, the 3'-OH terminus of a sequence is phosphorylated in step *i)* using e.g. 2-cyanoethyl phosphate and DCC as an activator. The intermediate phosphodiester is then activated in step *ii)* with mesitylenesulfonyl chloride (MsCl) to convert a hydroxy function into a good leaving group. The iterative reaction cycle is completed by addition of the desired nucleotide to form the phosphotriester in step *iii)*.[208] Cycles of steps *i)*, *ii)*, and *iii)* can be repeated with different nucleotides until the desired sequence is obtained.[209] Treatment with a weak base releases the 2-cyanoethyl phosphate protective group to convert the backbone motif into a phosphodiester; subsequent treatment with strong base cleaves the sequence from the support. The phosphotriester method successfully suppresses side reaction during oligonucleotide synthesis and renders very difficult purification steps unnecessary, yet the early method as pioneered by Letsinger suffered from low reactivities and—especially on solid polystyrene supports—reaction times were days to weeks for each step.[208-209] Transferring the system to the solution phase increased cycle yields (from 44 %[209] to 64 %[213] in selected examples) and

reduced the reaction time per cycle to less than two days. However, the choice of 2,4,6-triisopropylbenzenesulfonyl chloride as activator[214] and the development of 2'-hydroxy and 3'-hydroxy protective groups[215] was granting the access to DNA as well as RNA oligonucleotides[216] at the end of the 1960's.



Scheme I-9: Phosphotriester method using a solid support. Phosphorylation i): 2-cyanoethyl phosphate, DCC; phosphate activation ii): MsCl; phosphotriester formation iii): nucleoside (3'-OH and 5'-OH, non-protected).

Further chemical and technical perfection led to the *charged-compound coupling method* developed mainly by Itakura and Gait until 1980. As depicted in Scheme I-10, the protected monomer enters the reaction cycle as a triethylammonium salt of the phosphodiester, the condensing agent 1-(mesitylenesulfonyl) 3-nitrotriazole (MSNT) allows a very efficient coupling to the 5'-hydroxy end of the growing chain, even when immobilized on polystyrene or polyamide supports.[217-218] The activation process is depicted in Scheme I-10 and was employed in a semiautomatic oligonucleotide synthesizer.[219] The duration of a coupling cycle was shifted to the minute time-scale and yields of 95 % for each monomer elongation were feasible.[217]

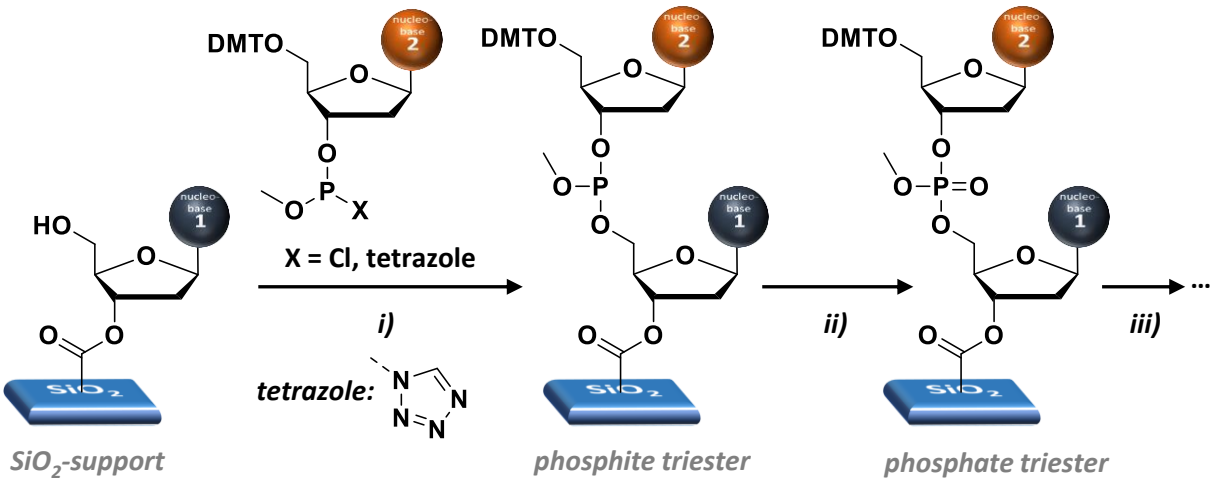


Scheme I-10: Monomer activation with MSNT for phosphotriester method (charged-compound coupling).

3.1.5 Phosphite Triester Method

It was again Letsinger's group that after experimenting with aryl phosphorodichloridates,[220] pioneered an important step forward. When using instead the P(III) analogue, an aryl phosphorodichloridite, remarkable reactivity was observed, leading to an internucleotide phosphite triester linkage that could be oxidized to the corresponding P(V) phosphotriester within seconds using iodine and water.[221] The repeating reaction cycles of the phosphite triester method comprise thus

phosphite triester formation *i)*, oxidation to the phosphotriester *ii)*, and terminal alcohol deprotection *iii)*. By repeating the cycles it proved to be a robust new method.[222] What Letsinger demonstrated in solution was further improved by his former student Caruthers. The latter understood that slow diffusion rates, irreversible reagent absorption, etc. were limiting the scope of various polymeric supports. Overcoming some of these obstacles, he proposed a new macroporous silica support that proved compatible with Letsinger's phosphite triester method.[223] The synthetic approach is described below in Scheme I-11. The support material is discussed in subchapter *Solid Support Material* (pp. 39f.)



Scheme I-11: Oligonucleotide synthesis using the phosphite triester method on a novel silica support. Coupling i): monomer, -78 °C; oxidation ii): iodine, water, base, RT; deprotection iii): TsOH or $ZnBr_2$, RT.

When using the phosphitechloridite monomer (X = Cl in Scheme I-11), cycle yields above 90 % were obtained and a cycle took about four hours.[223] The chemistry was compatible with all protected nucleobases and a simple methyl group was used as phosphate protective group. When tetrazole analogues were employed as monomers under the same conditions (X = tetrazole in Scheme I-11) the coupling reaction *i)* was found to be complete after 60 min, a cycle lasted less than two hours, while the coupling yield per cycle was above 95 %. The phosphate protective group could be removed with triethylammonium thiophenoxide at reduced temperature, followed by oligonucleotide cleavage from the support by treatment with concentrated ammonia at ambient temperature.[192] In the latter described approach from 1981, an apparatus was designed performing automated synthesis by passing the reagents through a column that is filled with the silica support. The approach was nevertheless not convenient for non-chemists because the phosphite chloride or phosphite tetrazolide monomers had to be stored at low temperatures due to their enhanced reactivity. Moreover, the extreme conditions during couplings that were performed at -78 °C rendered the commercialization of the method at this stage unsuccessful.

3.1.6 Development of the Phosphoramidite Method

Inspired by the discovery of reactive phosphite tetrazolides, Caruthers' group drew fruitful conclusions in the same year. They found that analogue *N*,*N*-dimethylaminophosphoramidites are stable enough to be stored at ambient temperature but can readily be activated by treatment with a weak acid (making dimethylamine a leaving group). They could show that a *N*,*N*-dimethylaminophosphoramidite is transformed into the corresponding and very reactive phosphitechloride when treated with *N*,*N*-dimethylaniline hydrochloride (DMAP·HCl). A screening for the most suitable weak acid for activating a phosphoramidite monomer identified the activator 1*H*-tetrazole. The phosphoramidite method allows the activatable monomer to be stored at room temperature and to be used during the chemical conversion at the same temperature. The *in situ* activation with a weak base reduced side reactions and as the activated intermediate species is present at room temperature, the chemical conversion

completes even faster.[5] The process was hereinafter optimized for the use of more sterically hindered dialkylamine residues at the phosphite center identifying diisopropylamine useful at this place to increase stability in solution.[224-225] The 2-cyanoethyl phosphate protective group furthermore experienced a renaissance. Compatible with the phosphoramidite method it can readily be cleaved with weak base.[226] Together with a new silica support with a longer linker,[225] mild conditions were generating near-quantitative coupling yields within minutes. The optimized protocol is presented in Scheme I-12.

Successful commercialization of automated synthesizers relying on the phosphoramidite method in following years made convenient bench-top oligonucleotide synthesis available to non-chemists. Phosphoramidite chemistry was consequently established as a complete concept. The following subchapter describes relevant aspects in detail.

3.2 Phosphoramidite Chemistry

Developed in the 1980's phosphoramidite chemistry is today's first choice when synthesizing DNA fragments. Being such a relevant technique to non-chemists, modern oligonucleotide synthesizers can basically be used as a black box tool and all reagents are commercially available. The following paragraphs provide a brief summary of support materials, the robust underlying chemistry, relevant compatible protective groups used during the protocol, and further variation and application of phosphoramidite chemistry.

3.2.1 Solid Support Material

Inorganic silica has early been identified to be a suitable support matrix.[227] With the availability of high performance liquid chromatography (HPLC) grade controlled pore glass (CPG) that was designed to be a uniform stationary phase, the advantages of silica supports were tremendous.[192, 228] Due to the rigid network, the material is incompressible and optimized for this sake. High flow rates of solvents and reagents are tolerated. Rigidity also means that there is no support swelling and surface accessibility is independent from the solvent selection. The narrow pore size distribution implies similar conditions for all surface-immobilized chains. And finally the chemical nature of the silica entails the support's chemical stability over a large range of conditions as well as no unintended interactions between the support and the growing chain.[229] CPG's are available with different average pore sizes. For oligonucleotide synthesis pore diameter sizes from some ten nanometers to some hundred nanometers are used. The bigger the pore size, the longer is the oligonucleotide length that can be efficiently synthesized. It has been shown that immobilization on long chain alkylamine (lcaa) CPG allows an efficient accessibility of reagents.[225] Lcaa CPG is depicted below in Figure I-14.

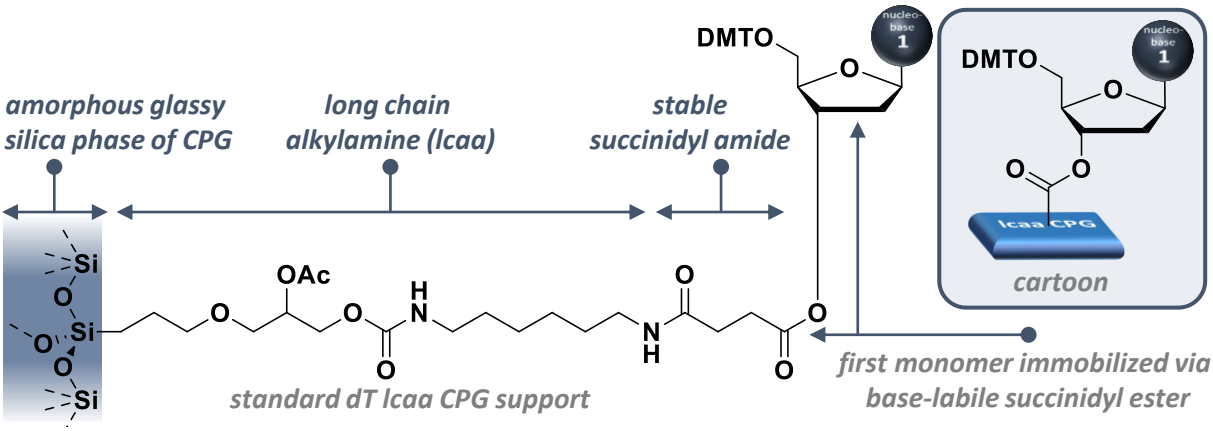


Figure I-14: Standard lcaa CPG support with immobilized deoxynucleotide.

It is noteworthy that also modern polystyrene supports are in use in oligonucleotide synthesis that have overcome initial shortcomings. Aminomethylated macroporous polystyrene (MPPS) is a low-swellable and highly crosslinked support whose macropores are generated by use of a porogeneous agent during polymerization. Containing anhydrous conditions is the material's main advantage due to its hydrophobicity.[230]

3.2.2 Iterative Reaction Cycles in Detail

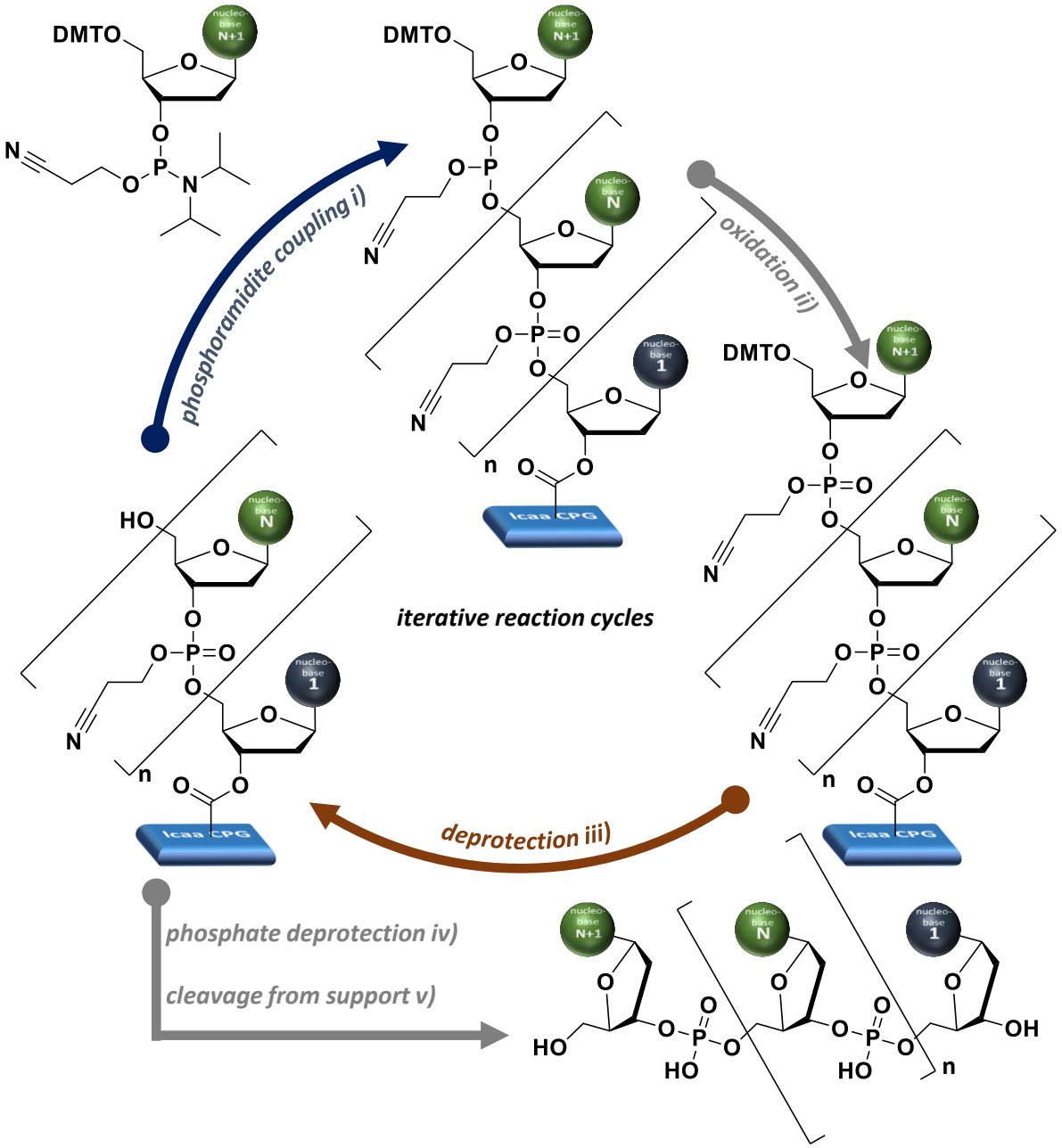
Phosphoramidite chemistry is usually performed using an oligonucleotide synthesizer. The machine delivers reagents in optimal amounts and in precisely controlled time intervals to a column with the support material. Among others protocols can be adapted to the synthesis scale and a software allows typing in a primary sequence and the machine does then automatically perform polymer synthesis according to the phosphoramidite monomers charged in the machine's reservoirs.

Iterative reaction cycles are performed and comprise the following steps (compare Scheme I-12):

- **Phosphoramidite coupling** *i)*: The growing chain comprises a 5'-hydroxy terminus that reacts with the particular 1*H*-tetrazole activated phosphoramidite monomer. Activator and monomer thereby come in contact by mixing inside the reaction column. An activated monomer reacts immediately and the approximately tenfold monomer excess ensures a yield above 98 %. Coupling time including the subsequent washing step does usually not take longer than 2 min.
- **Oxidation** *ii)*: The so-formed phosphite triesters are transformed to the more stable P(V) species, the phosphotriester, by oxidation with iodine and water. The oxidizer solution is flushed through the support matrix at maximum speed. The oxidation and subsequent washing step takes around 30 sec.
- **Capping step**: Phosphoramidite coupling is near-quantitative. Yet, the non-reacted chains comprise reactive 5'-OH ends that should not enter the next phosphoramidite coupling steps. By capping them, only the desired sequences reach the full length—a feature that plays an important role in purification protocols. Acetic anhydride activated by DMAP or *N*-methylimidazole is usually used during the capping step that can be performed both before and after oxidation.
- **Deprotection step** *iii)*: The growing chain end comprises a DMT-protected alcohol that is unblocked with trichloroacetic acid (TCA). The cleaved protective group is characterized by strong absorption of visible light around 500 nm wavelength. That can serve for a quantitative yield determination after each iterative reaction cycle. Deprotection and extensive washing take less than a minute.

Each monomer is added to the growing chain following the above described steps of the cycle. The reaction cycles are repeated until the desired sequence-defined poly(phosphotriester) is attained. Steps to finally obtain the desired sequence are:

- **Backbone deprotection** *iv)*: The phosphate protective group, 2-cyanoethyl, can be cleaved from the backbone with a weak base and is then washed away while the sequence is still immobilized. The same can be done for potential nucleobase protective groups. Procedures differ here depending on strategies and purification protocols employed.
- **Cleavage from the support** *v)*: The succinic ester connecting the 3'-terminus of the sequence with the support material is cleaved with ammonia and the sequence is washed from the support.



*Scheme I-12: Phosphoramidite Chemistry. Iterative reaction cycle comprises phosphoramidite coupling i): MeCN, phosphoramidite monomer, 1*H*-tetrazole; oxidation ii): THF, iodine, water, 2,6-lutidine; deprotection iii): TCA, DCM. Desired sequence attained after phosphate deprotection (and nucleobase deprotection) iv): MeCN, piperidine; and cleavage from support v): (methylamine) ammonia.*

Different purification protocols are used to separate the desired sequence from the capped fractions. *DMT-ON purification* requires stopping the iterative reaction cycles after the last oxidation step and retaining the last added monomer DMT-protected. This feature renders the desired sequences particularly different from the unwanted truncated sequences and allows for an efficient separation by chromatography. The truncated oligonucleotides (also called capped fractions or shortmers) are polar and can be washed through a reverse phase column. Also protective groups eventually just removed during sequence's cleavage from the support are washed away. The terminal DMT-group of the desired oligonucleotide is less polar and its interactions with the column material result in a much slower movement of the whole molecule. Separation is thus possible by a preparative HPLC using an appropriate solvent gradient from polar to less polar. In another elegant manner, the oligonucleotide support cleavage solution is applied manually to a disposable reverse phase purification kit, polar

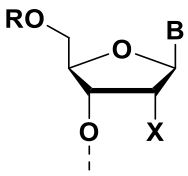
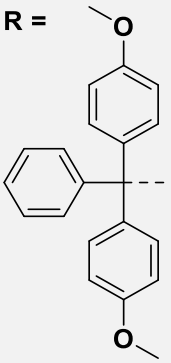
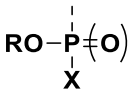
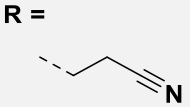
contaminants are washed out and the remaining desired oligonucleotide is treated with acid when still being on the column and after DMT-deprotection the desired and fully deprotected oligonucleotide can be isolated.

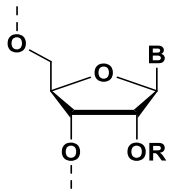
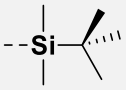
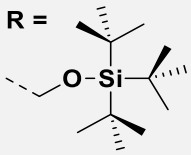
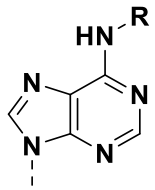
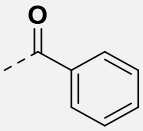
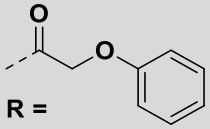
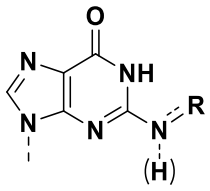
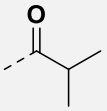
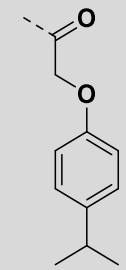
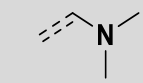
3.2.3 Appropriate Protective Groups

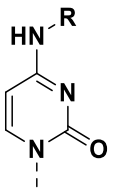
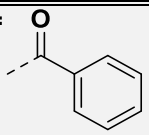
A DMT-based protection strategy has proven most beneficial in combination with phosphoramidite chemistry on a solid support system.[231] Protection group strategies are thus presented relying on a DMT-protected 5'-hydroxy function. Exocyclic amine functions of the nucleobases as well as 2'-alcohols in the case of RNA monomers would react with phosphoramidites and are hence protected orthogonally to the 5'-hydroxy function. The choice of protective group is delicate because it changes other chemical properties of the compound. In the case of RNA monomers' 2'-hydroxy protection sterics next to the coupling site and thus coupling times are strongly affected. Protection of exocyclic nucleobase amine functions has a tremendous impact on the electronic structure and depurination rates under e.g. the acidic conditions used for DMT-removal have to be taken into account. Considering the protective groups' stability during and the influence on oligonucleotide synthesis using phosphoramidite chemistry, their quantitative and fast removal after synthesis of sequences is another pivotal feature.

In Table I-1 relevant protective groups are presented. Protective groups for a standard and robust strategy are presented in light grey fields. They can be removed within minutes using an aqueous ammonia/methylamine (AMA) cleavage solution.[232] For specialty applications, when base-labile non-natural monomers (e.g. dyes) are added during oligonucleotide synthesis, a less harsh deprotection strategy has to be performed. A standard protocol termed *ultra-mild cleavage* requires exocyclic amine protective groups that cleave at reasonable rates using potassium carbonate in methanol.[233] The adequate protective groups for this protection strategy are highlighted in dark grey. Thymine and uracil do not exhibit an exocyclic amine function and are thus not included in the list.

Table I-1: Relevant protective groups for phosphoramidite chemistry

Protected Species	Protective Group	Comment
 5'-hydroxy group	 DMT group	The DMT protective group[198] can be efficiently cleaved with a weak acid like 3 w% TCA in DCM. Cleavage is facilitated by stabilization of the central carbocation by the methoxy electron-donating groups. Tr is thus too stable for an efficient cleavage under mild conditions whereas 4,4',4''-trimethoxytrityl (TMT) is already too unstable. The DMT cation's strong absorption serves to spectroscopically evaluate the yield after each reaction cycle.[231]
 phosphite (phosphate)	 CE group	The 2-cyanoethyl (CE) phosphate protective group was first used[208] and then identified as stable during phosphoramidite chemistry.[226] It can be selectively cleaved using a weak base like 20 v% piperidine in acetonitrile or it is removed during cleavage from support. The leaving group acrylonitrile is volatile.

Protected Species	Protective Group	Comment
 2'-hydroxy group	R =  TBDMS group	The tert-butyldimethylsilyl (TBDMS) protective group[234-235] is orthogonal to DMT and base-labile nucleobase protective groups as it is cleaved with fluoride ions by using e.g. tetrabutylammonium fluoride (TBAF). The 2'-alcohol is the last to be unblocked because it is responsible for phosphodiester bond (3'→2') migration and cleavage under basic conditions.
	R =  TOM group	The tri-iso-propylsilyloxymethyl (TOM) protective group[236-237] is as TBDMS orthogonally cleaved with fluoride ions. The smaller sterical hindrance close to the phosphoramidite moiety, however, allows a fast coupling step (approx. tenfold faster compared to TBDMS analogues).
 adenosine amine	R =  Bz group	Amines can be protected with the benzoyl (Bz) group.[199] This protective group is removed at the same time when performing the cleavage from the support with methylamine/ammonia.
	 Pac group	The phenoxyacetyl (Pac) is a fast-cleaving amine protective group suitable for adenosine protection.[238]
 guanosine amine	R =  isb group	The iso-butyryl (isb) protective group is a robust protective group of the guanosine exocyclic amine group.[239]
	R =  *i*Pr-Pac group	The 4-iso-propylphenoxyacetyl (iPr-Pac) amine protective group can be cleaved under mild conditions. When used to protect the amine function of guanosine, care has to be taken during the capping step. Use of acetic anhydride leads to a substitution and the resulting acetamide would need less mild conditions for cleavage. It is hence suggested to use trimethylacetic anhydride as a milder capping agent.[233]
	R =  dmf group	The dimethylformamidyl (dmf) amine protective group can be cleaved under mild conditions and, advantageously, is stable enough to not require an adapted capping protocol.[240]

Protected Species	Protective Group	Comment
 cytosine amine	R =  Bz group R =  Ac group	Historically, the cytosine amine was protected with the benzoyl (Bz) group.[199] However, base modification occurs during ultrafast deprotection/cleavage when methylamine is added.[232] Using the acetyl (Ac) protective group overcomes the above described side product formation. Compared to the Bz group, the increased electropositivity of the carbonyl carbon also increases the rate of deprotection and the attack on the pyrimidine ring is prevented.[232]

3.2.4 Phosphoramidite Chemistry for Microarray Production

Phosphoramidite chemistry can be applied to a more complex task, the production of an oligonucleotide microarray. Such a microarray consists of a number of microspots each immobilizing one particular oligonucleotide species that is characterized by its primary structure. Also called DNA chips, these oligonucleotide-presenting surfaces reconstruct (artificial) genes and e.g. unknown analytes can be characterized by their interaction with oligonucleotide sequences at specific sets of microspots. Hybridization can be spotted by fluorescence with high accuracy. Research in genomics and molecular biology as well as pharmacogenomic research, infectious and genetic disease and cancer diagnostics among other disciplines are benefitting from this technology.[241] However, the scope of the present work is to describe the chemistry of *in situ* DNA microarray production.

For the production of efficient DNA microarrays it is the key challenge to have a maximum number of spots with a reasonable oligonucleotide density on the smallest possible surface area. The oligonucleotide's primary sequence has to be guaranteed as well as the grid of microspots is to be precisely localized. Historically, the challenge was met using an adapted version of phosphoramidite chemistry allowing for photolithographic massive parallel *in situ* oligonucleotide synthesis on a two-dimensional surface.[242] Standard phosphoramidite chemistry relies on an acid-labile DMT protective group that regenerates a reactive hydroxy group at the terminus of the growing chain prior to coupling a new phosphoramidite monomer as it is depicted in Scheme I-12. By changing the 5'-OH protective group to a photolabile moiety, light can be employed as a spatially addressable chemical stimulus in the deprotection step. The high resolution of photolithography enables miniaturization of the synthesis scale and thousands of sequences can be synthesized in parallel in a compact design. Masks can be used to define which spot is to be activated for the following coupling step. Microspots in the shade of the mask contain the fraction of oligonucleotides that are prevented from elongation during the following phosphoramidite coupling step. The position of holes in the mask defines which microspots contain the oligonucleotides that are determined to react during the next phosphoramidite monomer addition. The parallel oligonucleotide synthesis relying on four different nucleotide monomers requires four phosphoramidite reaction cycles per monomer addition in this way. A maximum of 4^n oligonucleotides with the length n is synthesized in $4n$ cycles.[243] E.g. more than a million decamers can be synthesized with a defined sequence allocated to a defined spot applying only 40 phosphoramidite reaction cycles. The development of maskless array synthesis[244] where software-controlled micromirrors automatically and precisely guide the irradiation stimulus resulted in further miniaturization of DNA microarrays. Optimization of the crucial photodeprotection step by using the (2-nitrophenyl)propoxycarbonyl[245-246] (NPPOC) alcohol protective group rendered the synthetic

protocol highly efficient.[247] The NPPOC deprotection is schematically described in Figure I-15. Besides spatial control of the deprotection reaction by means of photolithography with pre-made masks or dynamic micromirror devices, other technologies are also used relying classically on an acid-labile alcohol protective DMT group (shown in Figure I-15). Inkjet printing technologies allow the spatial reagent delivery as well.[248] Electrochemical stimuli on microelectrode arrays do allow spatially controlled generation of acid to control a parallel *in situ* array synthesis.[249-250]

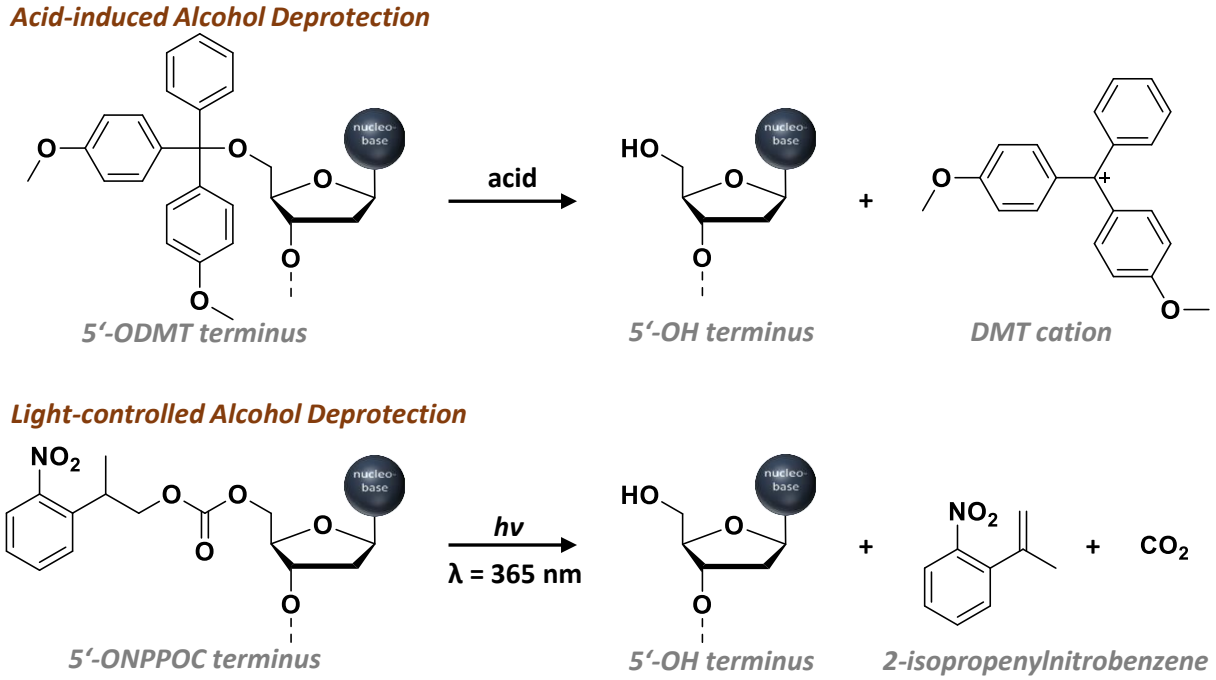


Figure I-15: Acid-induced alcohol deprotection used in standard phosphoramidite chemistry and for DNA microarray fabrication by inkjet printing reagent delivery or by acid generation at microelectrodes (top) and light-controlled alcohol deprotection used in photolithographic DNA microarray fabrication (bottom).

3.2.5 Phosphoramidite Chemistry for Synthesis of Abiotic Poly(phosphodiester)s

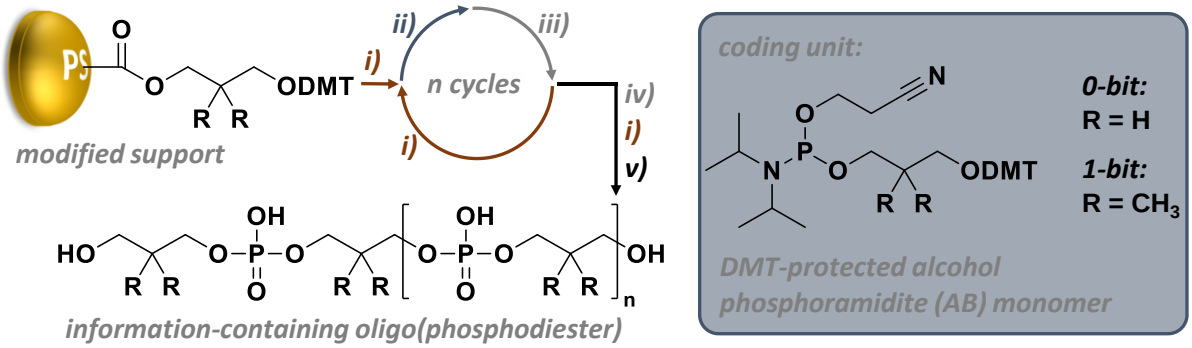
In 1987, Seela and coworkers found that the insertion of an abiotic monomeric unit—namely a propan-1,3-diol moiety—into an oligonucleotide's sequence would stabilize a hairpin motif in its secondary structure.[251] Other abiotic *spacer* monomers comprising different length hydrocarbon[252-253] and oligo(ethyleneglycol) chains[254-255] were developed that result in a finely controlled hairpin stabilization. A design including two insertions of such spacers would even stabilize a triple helix.[256] Similarly, rigid spacers can predefine angles in DNA nano-assemblies to form polygons[257] or defined assemblies on the micrometer scale.[258] Numerous examples confirm that the insertion of abiotic monomeric units at distinct locations within an oligonucleotide's sequence is an important tool in DNA-related biotechnology. DNA-hybrids are further of scientific interest because they are well-defined systems that give access to tunable architectures, e. g. distinct distances between moieties to investigate photochemical processes.[259]

By increasing the ratio of abiotic monomer content, Sleiman and coworkers designed triblock DNA hybrids. The middle block of such duplex is hybridized DNA, the terminal blocks are composed of an oligoethylene glycol dendron resulting in a macromolecular amphiphilic structure. These triblock duplexes assembled in MeCN/water to fibers.[260] When the abiotic part of the DNA amphiphiles was given a more apolar character, micelles were formed in water.[261] By using dodecyl and hexaethylene glycol phosphoramidite monomers, a series of abiotic blocks was added to an oligonucleotide block. These precision diblock DNA-hybride copolymers were investigated for the influence of the abiotic block's sequence (of hydrophilic/hydrophobic monomers) on the micelle formation in water.[262] As phosphoramidite chemistry—employed for the automated precision synthesis of such DNA hybrides—

tolerates *"many functional groups, such as amides, aryl halides, ethers, disulfides, maleimides, alkenes, alkynes, and protected amines",*[262] the core structure of such almost uniform micelles could further be varied for tuning particular properties.[263-264]

Fully abiotic defined sequences were investigated by Häner and coworkers. They showed that driven by a hydrophobic effect and π-π stacking, *oligopyrenotides* hybridize in water similar to DNA forming helices with axial chirality and several self-assembled structures could be designed.[265-267] Amphiphilic pyrenotide trimers could further be assembled as sheet-like two-dimensional supramolecular polymeric networks.[268-269] The directionality of the latter was thereby governed by temperature gradients or the sense of stirring during the supramolecular polymerization process.[270] Changes in positioning of the linker unit on the central pyrene moiety enabled the supramolecular trimer assembly in form of tubes.[271] When the central moiety was in majority replaced by phenanthrene and a doping with pyrene units was achieved, the resulting supramolecular nanotubes showed light-harvesting properties.[272]

Abiotic sequence-defined poly(phosphodiester)s were recently investigated for their information storage potential by Lutz and coworkers. By defining two comonomers as 0-bit and 1-bit, respectively, a binary code could be written along the polymer chain.[10] Relying on a polystyrene support, the manual synthesis of a sequence-defined oligomer is shown below in Scheme I-13.



Scheme I-13: Manual synthesis of uniform information-containing oligo(phosphodiester)s. DMT-deprotection i): TCA, DCM; phosphoramidite coupling ii): MeCN, 1H-tetrazole, (AB) monomer; oxidation iii): THF, iodine, 2,6-lutidine, water; phosphate deprotection iv): MeCN, piperidine; cleavage from resin v): methylamine, ammonia.

Phosphoramidite chemistry has proven particularly useful for this approach due to its high fidelity: as every truncated sequence would inevitably result in an erroneous code, absolute sequence-definition is crucial. Important for the amount of data to be encrypted, the synthetic strategy furthermore allows automation that resulted in the synthesis of sequence-defined digital poly(phosphodiester)s with a degree of polymerization above 100. The unprecedently long sequence-defined polymer contained the binary-encoded message *Macromolecule*.[6] thus a data storage capacity of 13 bytes (= 104 bits) per sequence could be achieved. Yet, the information read-out remains challenging when pure poly(phosphodiester)s are used as digital polymers.[273] This central aspect of digital polymers is object to the following subchapter.

4 Sequence-defined Polymers for Information Storage

In a *vivid* example, valuable information was kept for some 10,000 years well-preserved on a sequence-defined polymer. The DNA of some long-extinct wooly mammoths was isolated. Despite the fact that the whole species has died, with its DNA in hand, mammoth proteins—optimized to help organisms master a cold climate—can nowadays be artificially expressed and be brought back to *life*.[274] It is thus hardly surprising that the biotic storage medium was a prime candidate also to archive digital data since the function of DNA was understood. An interesting aspect is the immense miniaturization and

thus enlarged data capacity when moving information from classical storage devices like magnetic hard drives to the molecular level.[9, 275]

DNA received the code for *life* unintentionally by trial and error during the process of evolution. While a cell's DNA encodes one amino acid in an enzyme with a base triad (see Figure I-2, right) and the vast majority of DNA material is a non-coding playground for evolution, artificial DNA could encode information by defining the four DNA monomers as particular small data packages. Their order would reveal the digital code. The progress that has been made in the field is discussed in subchapter *DNA-based* (p. 47). Indeed, using the data storing biomolecule is advantageous. The synthetic access is convenient and progress in chemical biology has provided the tools to replicate synthetic DNA efficiently during polymerase chain reaction (PCR). That is equivalent to writing and copying digital data. Reading the data is equally important. DNA can be sequenced by intentional termination reactions during PCR called *Sanger sequencing*[276-277] and several modern *next-generation sequencing*[278-279] (NGS) techniques. A promising technique is nanopore sequencing. During the latter a DNA chain passes through a biological transmembrane pore protein and induces measurable voltage variation motifs that can be traced back to the specific primary structure.[280] Real time analysis and destruction-free single-molecule sequencing is thus achievable.[281]

Inspired by the progress that polymer chemists had recently achieved in the field of sequence-controlled polymers, Lutz envisioned in 2013 that instead of being restricted to DNA as a molecular storage medium, the full range of chemical possibilities may be employed to obtain tailor-made information containing polymers.[148] Virtually all techniques leading to sequence-controlled polymers presented in subchapter *Multistep-growth Synthesis* (pp. 28ff.) could be used to encode data on the molecular level. Hence, on the one hand, the chemical means for writing information on diverse polymer chains are available. On the other hand, sequencing techniques relying on biomolecules like enzymes are limited to DNA sequencing. For that reason the parallel development of applicable (new) sequencing techniques is required. Besides *NMR tweezer*[282] usage, tandem mass spectrometry (MS/MS) is such a promising tool, that has not been able to be exploited for DNA sequencing. Careful polymer design enables software-based automated milliseconds sequencing of such coded polymers.[283]

After presentation of DNA-based data storages other classes of information-containing polymers that were recently developed in Lutz' and others' laboratories are described followingly. The corresponding subsections can be read like a steady improvement of *MS sequencing* through design-analysis feedback loops.

4.1 DNA-based Data Storage

After a first proof-of-concept in 1988[284] and other tiny improvements,[285-291] it was a contribution by Church that systematically applied modern next-generation techniques to tackle data storage on DNA. The html data of a book, eleven JPG images and a JavaScript program were converted into a 5.27 Mbit (0.66 MB) bitstream. With nucleotides A and C coding for zero and G and T for one, the data was inkjet printed on DNA microchips (see Chapter I: 3.2.4 *Phosphoramidite Chemistry for Microarray Production*, pp. 44f.) in form of a library of information-containing oligonucleotides. Each so-formed oligomer on the surface contained 96 coding monomers (twelve bytes) as well as 19 monomers describing the location of the data block within the bitstream and 22 monomers were added to allow later amplification and sequencing. Assuming 100 synthesized oligonucleotides per spot on the chip an unprecedented information density of 5.5 petabits·mm^{-3} (Peta, 10^{15}) was reported.[9] Comparing with an estimated annual data generation in the range of some zetabytes (ZB; Zeta, 10^{21}),[292] theoretically the world's data accumulation would annually grow only by a shoe box of so-stored data. Based on the principle of dividing data into small blocks in an oligonucleotide library, others reported an enhanced

error correcting scheme. The bitstream is written on *overlapping* DNA chains thereby creating fourfold redundancy against systematic failure and data loss. The cost was reported to be 12,400 $\$\cdot MB^{-1}$ for synthesis and 220 $\$\cdot MB^{-1}$ for sequencing. Long term data storages that do not need to be read frequently could thus soon be commercially beneficial.[293] However, it should be noted that the system with the highest storage capacity reported to date (as of March 2018) covers just 200 MB of data in a proof-of-concept study.[8, 294] Ultra-long term DNA data storage could be designed to be free of errors for thousands of years even at ambient temperatures if preserved in an inorganic silica matrix.[295] Benefiting from informatics, advanced codes can be used to write arbitrary data into DNA chains using the near-maximum coding space but preventing at the same time biochemically disadvantageous microstructures like homopolymer blocks. With such techniques redundancy for error corrections can be reduced to access even higher information densities.[296] CRISPR-Cas methodology-based insertion of artificial DNA, coding for a short GIF movie, into a population of living bacteria was very recently reported by Church.[297] It is also discussed whether such artificial DNA-segments could be used as tags in patent-protected seeds to follow their use in unauthorized breeding programs.[298]

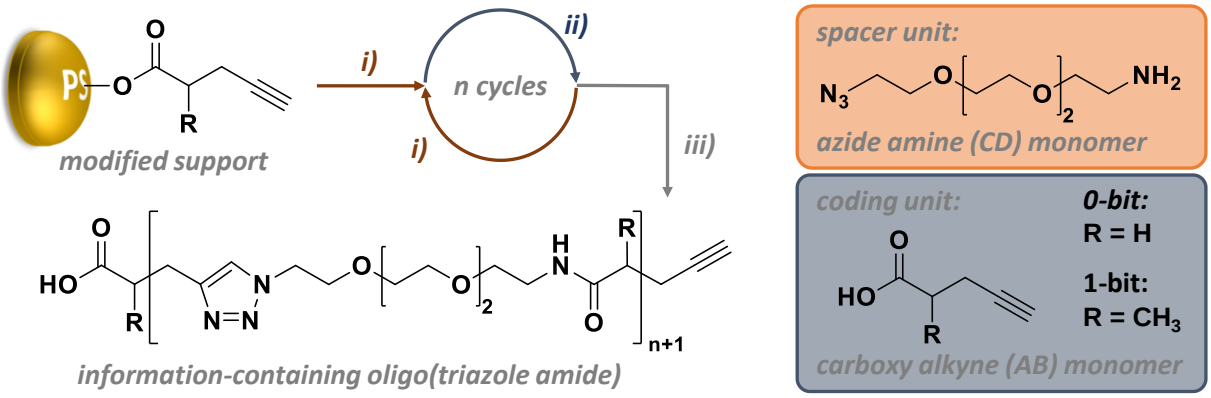
Balasubramanian and coworkers reported an elegant concept on how to use a DNA double helix to code in three different *layers*.[15] Only one strand, say the mother strand, is relevant for data storage. C and A code for a 0-bit, G and T for a 1-bit. The first *layer* is accessible by PCR amplification and standard sequencing. Post-polymerization modification, transforms all cytosine nucleobases into uracil (RNA building block) when treated with sodiumbisulfite. Subsequent PCR amplification does not translate the mismatch into following generations but at the position of the initial (C≡G) base pair there is besides the latter a 50 % possibility to find a (T=A) base pairing. An initial bit 0 C-position in the mother strand becomes a mixed C/T-position that is defined as a bit 1. Just the same, an initial bit 1 G-position becomes a mixed G/A-position and is defined as bit 0. All in all, bits that should be retained between the information layers have to be coded with A and T while bits that are supposed to change have to be coded with C and G in the mother strand. Even a third layer could be introduced in an analogue but more complex approach using additionally two artificial cytosine derivatives that can be chemoselectively addressed by two orthogonal chemical stimuli.[15]

The example shows that DNA-based data storages are to date advantageous due to the accessibility of tools by molecular biologists like PCR amplification. Nevertheless, the diversity of chemical reactions and reaction conditions is limited when applied to biopolymers. It can hence be expected that tailor-made information-containing polymers will show more elaborated data modification in future works.

4.2 Oligo(triazole amide)s

Following the AB+CD approach oligo(triazole amide)s are accessible by interconnecting repeatedly a coding 4-pentynoic acid (AB) and an 11-azido-3,6,9-trioxaundecane spacer (CD) by formation of triazole and amide linkages.[299] The solid or soluble[160] Wang support exhibits the first coding monomer attached as an ester; the present functional group is thus an alkyne. The (CD) spacer unit is linked in a CuAAC *i)* forming a 1,3-triazole linkage. The cycle is completed by an amidification *ii)* with a coding (AB) monomer as it is shown in Scheme I-14. By repeating the iterative reaction cycles and choosing which coding monomer is entering step *ii)* namely 4-pentynoic acid coding for bit 0 or its 2-methylated analogue coding for bit 1, binary information is written on the oligomer chain. Finally, the sequence-defined oligo(triazole amide)s are cleaved and washed from the Wang resin. The synthetic protocol was further optimized by synthesis of a dyad library with two coding (AB) units already connected via a spacer resulting in four larger (AB) monomers coding for bit pairs 00, 01, 10, and 11. With the same effort during multistep-growth synthesis an oligomer of double length is thus accessible.[300]

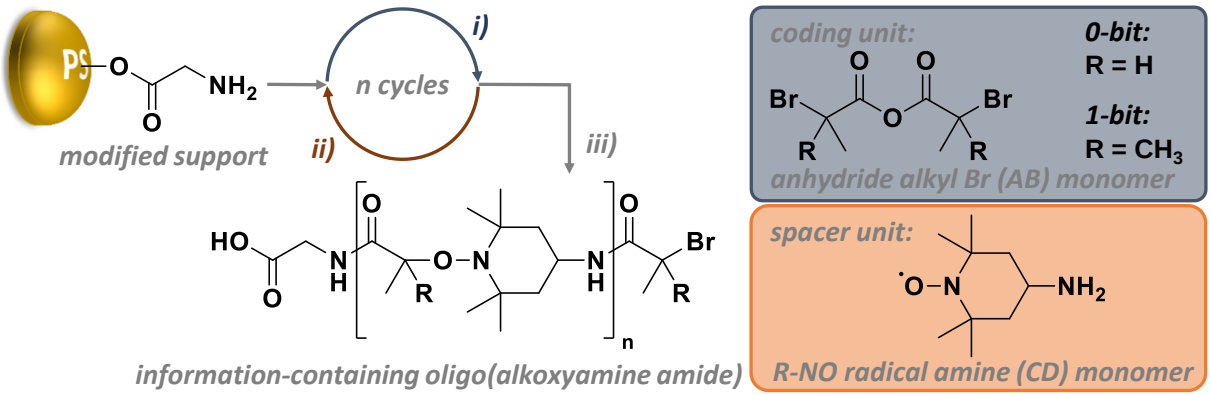
The attained sequences can be analyzed by electrospray ionization mass spectrometry (ESI-MS). The encoded binary data can further be deducted from tandem mass spectrometry (MS/MS).[301] Yet, the cleaving patterns were numerous and sequences had to be designed that allow for a more convenient sequencing methodology.[302]



Scheme I-14: Synthesis of uniform information-containing oligo(triazole amide)s. CuAAC i): CuBr, dNbipy, (CD) monomer; amidification ii): PyBOP, DIPEA, N-methyl-2-pyrrolidone, (AB) monomer, 60°C; cleavage from Wang support iii): TFA, DCM. Abbreviations: 4,4'-di-n-nonyl-2,2'-bipyridine (dNbipy), benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP), N,N-diisopropylethylamine, or DIPEA.

4.3 Oligo(alkoxyamine amide)s

Sequencing by means of MS/MS could be greatly facilitated by the intentional incorporation of a labile bond. With an optimal chemical design and under optimized electrospray ionization conditions the information-containing chain would be analyzed by MS and then—after MS/MS ionization—it would statistically cleave solely at inter-bit linkages. The entirety of so-attained fragments would reveal the total encoded information by their differences in mass. Oligo(alkoxyamine amide)s were shown to master these criteria. They can be synthesized in an AB+CD approach using as the (AB) monomer a coding carboxylic acid that is used in the form of a reactive symmetric anhydride that also exhibits an alkyl bromide function. The (CD) monomer is a 3-amino-2,2,6,6-tetramethylpiperidinyloxy (amino-TEMPO) spacer. The synthesis is performed on different glycine-modified supports starting with an amidification *i)* between the amine terminus of the growing chain and the anhydride (AB) monomer. The alkyl bromide is then converted into the corresponding radical that is then prone to radical-radical coupling *ii)* in presence of (CD) monomer amino-TEMPO as it is known from NMP. The iterative reaction



Scheme I-15: Synthesis of uniform information-containing oligo(alkoxyamine amide)s. amidification i): THF, HBTU, DIPEA, coding (AB) monomer; radical-radical coupling ii): CuBr, Me_6TREN, DMSO, spacer (CD) monomer; cleavage from Wang support iii): TFA, DCM. Abbreviations: 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium-hexafluorophosphat (HBTU), tris[2-(dimethylamino)ethyl]amine (Me_6TREN).

cycles characterized by step *i)* and *ii)* are repeated until the desired sequence is attained. Cleavage *iii)* from the support yields the information-containing sequence-defined oligo(alkoxyamine amide) in excellent yields.[303] The synthetic access is described below in Scheme I-15.

It has to be noted that quantitative yields during the radical-radical coupling *ii)* were only obtained when—depending on the nature of the (AB) monomer—an optimized amount of catalyst was used. The reactivity of the singly methylated bromide species of coding 0-bit is lower compared to the dimethylated analogue as the radical intermediate is less stabilized. An advantageous strategy was later reported. Described in the following subsection *Oligo(alkoxyamine phosphodiester)s* (pp. 52ff.), the coding site of the moiety should not drastically influence the monomer reactivity.

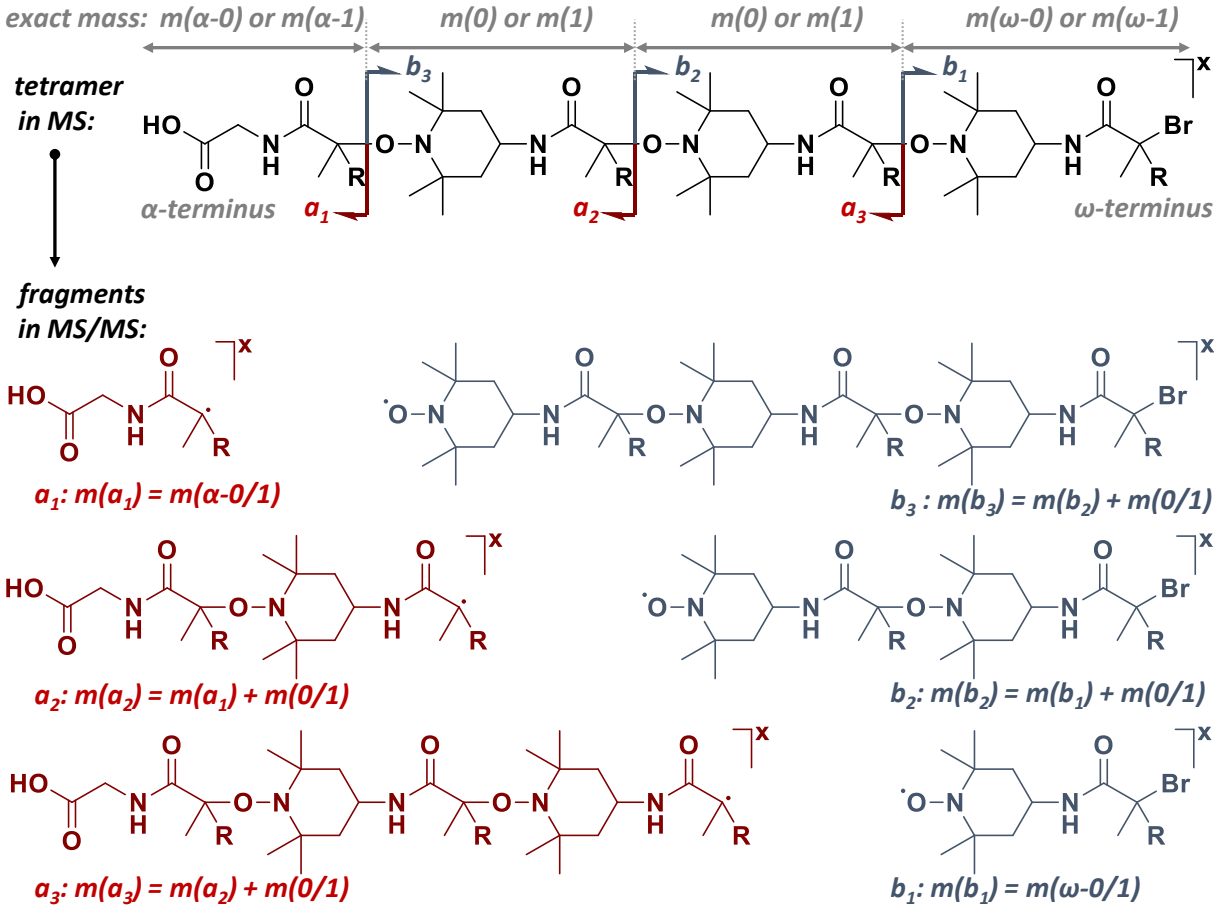


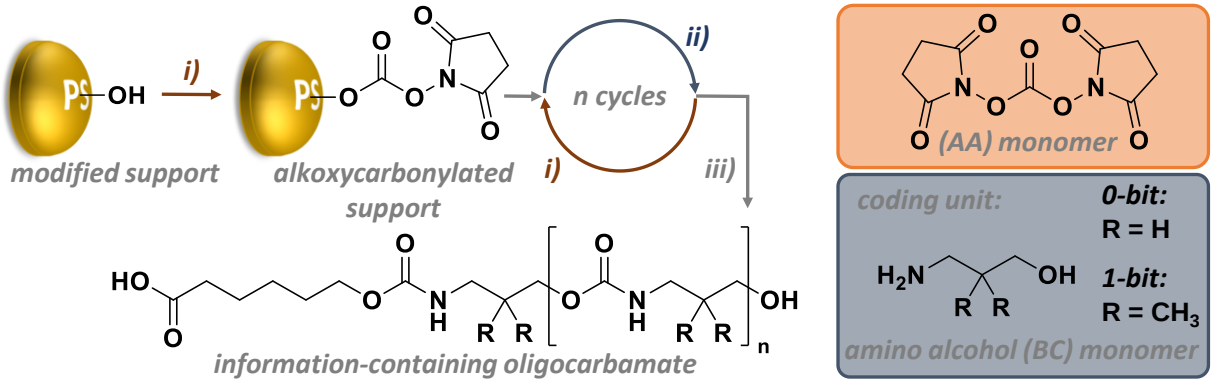
Figure I-16: Concept of MS/MS sequencing of linear macromolecules containing labile alkoxyamine bonds.

Same as for oligo(triazole amide)s, it could be shown that longer oligo(alkoxyamine amide)s are rapidly accessible using a dyad encoded (AB) monomer library being all combinations of the two coding moeties connected by a spacer unit.[304] Because alkoxyamines are labile and can again disproportionate along the bonds formed during radical-radical couplings, the information containing sequences can be irreversibly destroyed by heating above 60 °C entailing a very easy and fast way to erase the coded information.[303] The incorporation of labile alkoxyamine bonds successfully leads to defined fragments that can be detected by MS/MS. Figure I-16 shows the different patterns attained in MS/MS and the logic of deciphering the binary code thereof. The C-ON bond is the most labile in the molecule and that is why MS/MS conditions can be adjusted to see solely statistical C-ON cleavage when fragmenting the *x*-charged product peak (here: $x = \pm 1$). Hence, depending on the position of cleavage and charge, a limited number of fragments can be detected. Chains without charge are mute. The expected information-containing fragments' peaks are ordered relating to the chain end of the corresponding species. In Figure I-16 the structures with α-chain ends are shown in red, those with ω-chain ends in blue. As it is indicated, the fragments' exact masses relate to each other and when

ordering them from small to larger values, the differences in mass *m*(0) or *m*(1) reveal the 0/1-bit sequence in direction from the chosen chain end. E.g. the difference between b3 and b2 be *m*(0). The third coding monomer viewed from the ω-end is therefore a bit-0.[305] When mass spectrometry is performed in negative mode and *x* = -1, the sequencing is further facilitated because—thanks to the carboxylic acid—only the α-chain ends can stabilize a negative charge and thus the a_i-fragments are exclusively detectable rendering the analysis even more straight forward as redundant information is surpressed.[306] A recent report shows that several oligo(alkoxyamine amide) sequences of different lengths can be blended and sequencing by MS is still conveniently achieved. Called 2D-coding, this proof-of-concept of parallel analysis in one sample highlights the versatility and accuracy of the analytical method.[307]

4.4 Oligocarbamates

Possible application of digital polymers includes their use as identification tags in a (polymer) matrix, e.g. as a means against product piracy. The low heat resistance of the before described polymer class is a drawback for such application. Oligocarbamates overcome this limitation and still allow efficient deciphering by MS/MS. Sequence-defined oligocarbamates can be synthesized following a newly developed AA+BC approach that is described below in Scheme I-16.



Scheme I-16: Synthesis of uniform information-containing oligocarbamates. alkoxycarbonylation i): MeCN, Et_3N, DSC (AA) monomer, microwave, 60 °C; carbamate formation ii): DMF, Et_3N, coding (BC) monomer; cleavage from modified Wang support iii): TFA, DCM. Abbreviations: N,N'*-disuccinimidyl carbonate (DSC), triethylamine (Et_3N)*

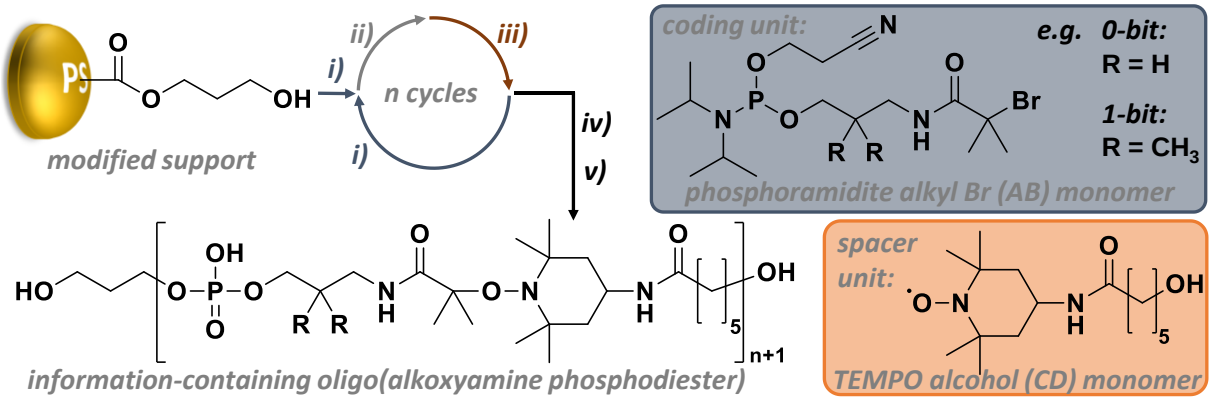
A Wang resin is modified in an esterification with 6-hydroxyhexanoic acid. The hydroxy function presented by the modified support is converted to an active alkoxycarbonyl by reaction with the (AA) monomer, DSC in step *i)*. When adding the amino alcohol coding (BC) monomer, the carbamate formation *ii)* is selective. The corresponding carbonate is not obtained.[308] After repeating the iterative reaction cycles until the desired sequence is formed, the sequence-defined oligocarbamate is cleaved from the resin and obtained in excellent yields.[309] Information-containing oligocarbamates allow analogously to oligo(alkoxyamine amide)s convenient code deciphering by MS/MS as explained exemplary in Figure I-16. The labile site is the R_1O-C(O)NHR_2 bond and the α-chain end exhibits a carboxylic group stabilizing the negative charge of MS/MS fragments. Starting from the carboxy terminus, the binary information is efficiently *read out*.

The here described synthetic approach was used to develop oligocarbamate-coded ATRP macroinitiators[310] and as anticounterfeit agents information-containing sequences were blended in methacrylate-based intraocular implants.[311] It was shown that weight fractions below 0.5 w% of added identification tag does neither impair the properties of the intraocular lenses, nor is the sequence-defined oligocarbamate significantly extracted from its matrix in experimets simulating *in vivo* conditions. Furthermore, 2D-coding (and MS-reading) as mentioned above for oligo(alkoxyamine amide)s showed to be applicable to different commodity polymers, such as PS, polyvinylchloride (PVC)

and polyethylene terephthalate (PET).[312] For polyvinyl alcohol (PVA) model implants, it was recently reported, that their labeling was stable for several month in *in vivo* experiments. The individual oligocarbamate labels were extracted and read in steady quality from materials that had been implanted for a week, a month and three months.[313] In conclusion, sequence-defined oligocarbamates represent a very promising class of digital polymers thanks to their convenient synthetic access, their behavior in commodity polymer blends and their straight-forward MS/MS sequencing.

4.5 Oligo(alkoxyamine phosphodiester)s

Combining phosphoramidite chemistry and radical-radical coupling resulted in a novel orthogonal oligophosphodiester synthesis. The intentional insertion of labile alkoxyamines is achieved by following an AB+CD approach that is depicted in Scheme I-17.



Scheme I-17: Synthesis of uniform information-containing oligo(alkoxyamine phosphodiester)s. phosphoramidite coupling i): MeCN, 1H-tetrazole, (AB) monomer; oxidation ii): THF, iodine, 2,6-lutidine, water; radical-radical coupling iii): DMSO, CuBr, Me_6TREN, (CD) monomer; phosphate deprotection iv): MeCN, piperidine; cleavage from resin v): methylamine, ammonia.

A modified support's terminal alcohol is coupled *i)* with a coding (AB) phosphoramidite monomer. After oxidation to the intermediate phosphotriester the alkyl bromide chain end is reacted with a (CD) monomer in a radical-radical coupling *iii)*. The three steps can be repeated until attaining the desired sequence whose phosphate groups are finally deprotected *iv)* before the information-containing oligo(alkoxyamine phosphodiester) is cleaved from the solid support *v)*.[12] The synthesized sequence-defined oligomers show the expected property when subject to MS sequencing. The inter-bit alkoxyamine linkage can selectively be cleaved during MS/MS fragmentation allowing for an easy information read-out of an oligomer based on phosphodiester linkages.[302] When using different spacer units with carefully chosen differences in their molecular weight, it has recently been revealed that the information density could be drastically increased.[314] In this way, MS/MS sequencing yields—instead of a bit resolution—two-bit information packages that code for 0-0, 0-1, 1-0, or 1-1 dyads.

4.6 Poly(phosphodiester)s With Intentional Alkoxyamine Insertion

The above described approach of inserting labile alkoxyamine bonds into an oligo(phosphodiester) was, in turn, reapplied to an automated synthesis solely based on phosphoramidite chemistry to be able to synthesize long sequences. To increase the efficiency of MS sequencing instead of inter-bit (or inter two-bit) alkoxyamine linkages, such linkages should separate bytes (8 bit packages) from each other. Each byte block does additionally exhibit an identification mass tag to distinguish them from another. Next to deciphering the 8-bit code the inter-byte order of the binary code is thus easy to conclude. The concept's successful implementation constitutes the present day's gold standard in terms of conveniently writing information on a polymer chain and deciphering the molecularly stored

data by means of MS.[11] The employed phosphoramidite monomers are given in Figure I-17 and the design as well as MS sequencing is shown in detail.

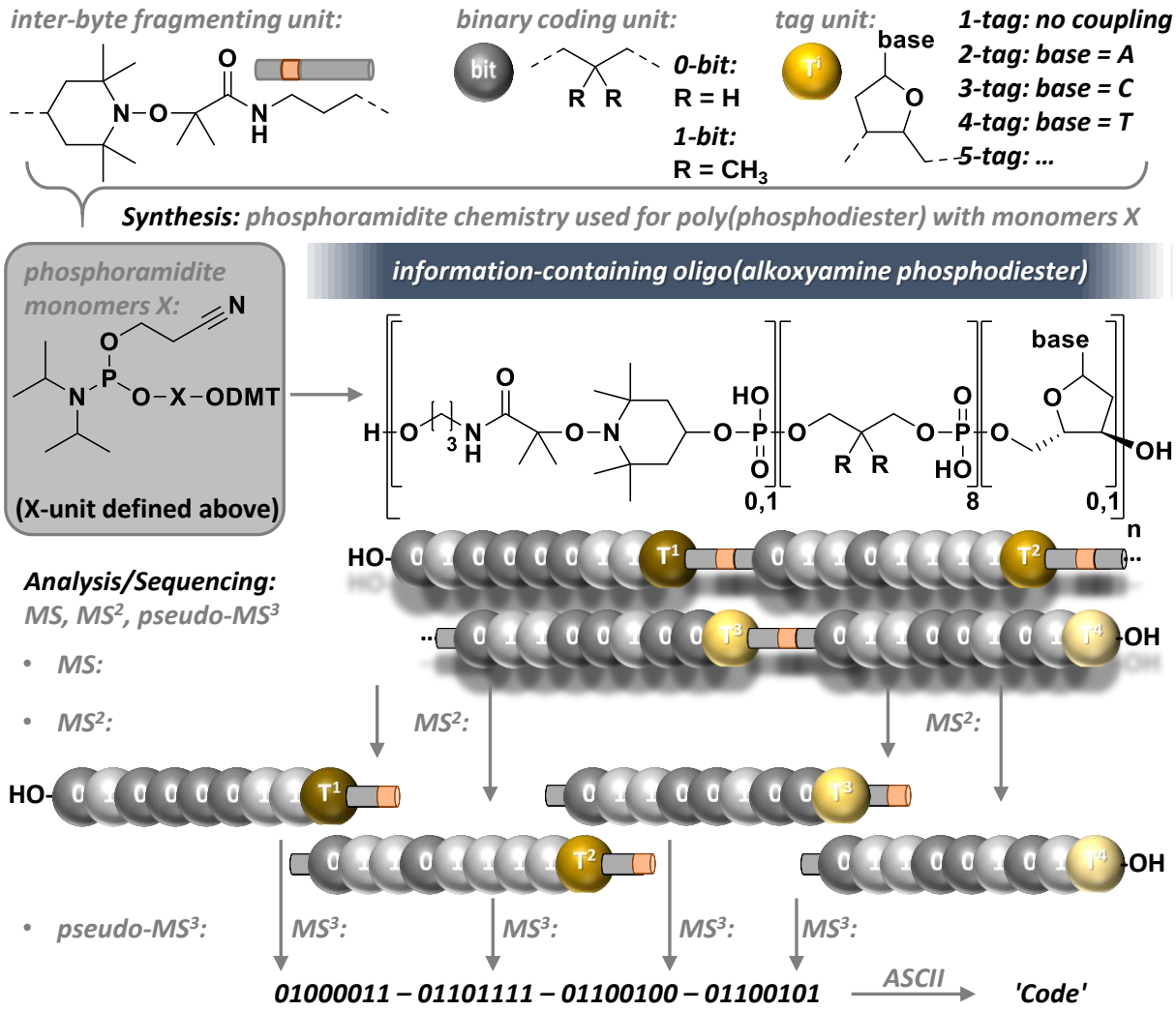
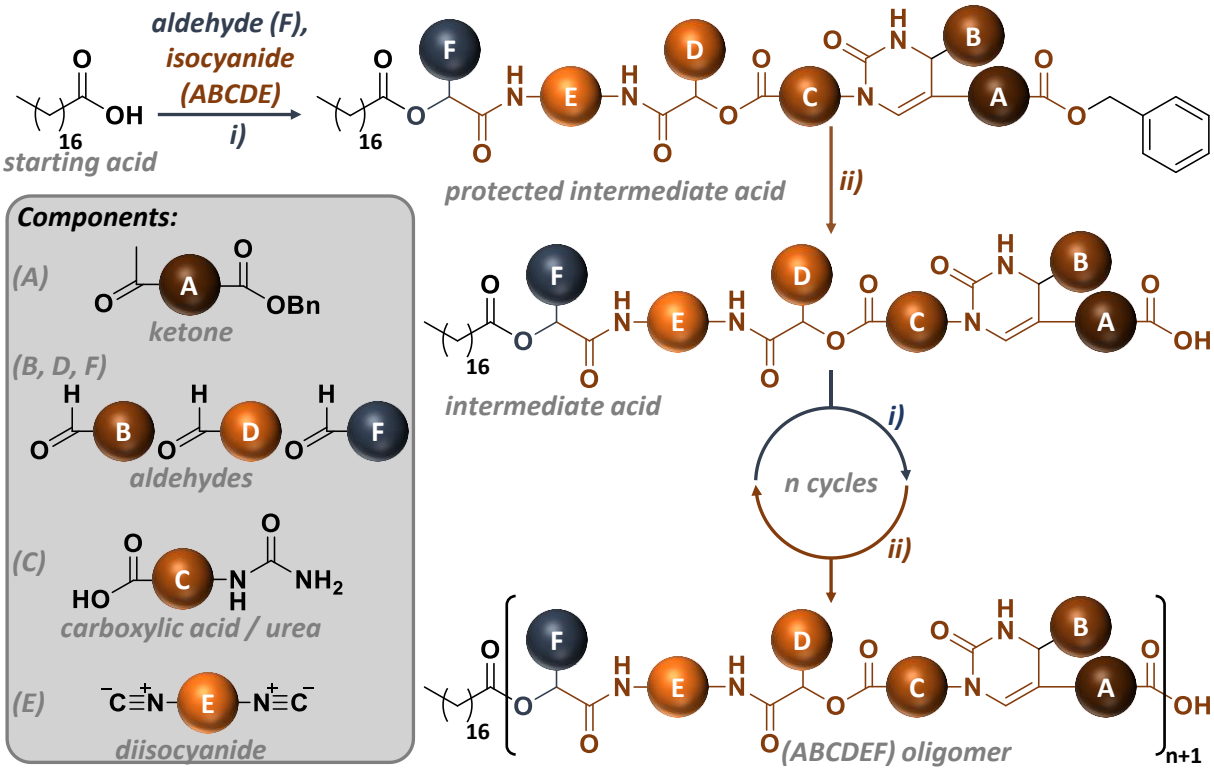


Figure I-17: Automated synthesis of uniform information-containing oligo(alkoxyamine phosphodiester)s and concept of long digital polymers' MS sequencing facilitated by programmed inter-byte fragmentation. Abbreviation: American Standard Code for Information Interchange (ASCII).

4.7 Molecular Data Storage Systems from Multi-component Reactions

Recent works by Meier and coworkers were contributing with a new conceptual approach to molecular data storage systems. Benefitting from multi-component reactions and the diversity of components that can be introduced at each step, it was shown that immense information density is theoretically accessible within a repeating unit.[315] A new synthetic strategy is reported that allows for fusion of six components to a repeating motif. During the oligomerization—that is shown in Scheme I-18—an already established protocol[155] was used including iterative P-3CR *i)* and benzyl deprotection steps *ii)*. During the repeated P-3CR the growing chain carboxylic acid is reacted with an aldehyde (F) and a macromonomeric isocyanate (ABCDE) resulting in the characteristic α-acyloxy carboxamide as it was already shown in Scheme I-5A. Novelty lies in the preparation of a library of protected macromonomers (ABCDE). The latter is accessible via a Biginelli reaction creating a carboxylic acid-bearing 3,4-dihydropyrimidin-2-one from an acetoacetic acid ester (A), an aldehyde (B) and a urea (C). The so-formed acid is reacted subsequently in a P-3CR with an aldehyde (D) and an diisocyanate (E), forming the isocyanate macromonomer (ABCDE).[316] The diversity of the strategy was demonstrated by the commercial or synthetic availability of elven ureido-carboxylic acids (C), six acetoacetate-benzyl

esters (A) and 18 aromatic aldehydes (B) for the Biginelli reaction. 26 different aldehydes (D) and 29 diisocyanides (E) could be used for the Passerini reaction to yield the isocyanate macromonomers (ABCDE). Similarly, 26 aldehydes (F) are available for each Passerini linking step. While the proof-of-concept was underpinned using 14 different ABCDE isocyanate monomers, using the theoretical chemical space from the before described variety of components A, B, C, D, E, and F, 24 bits per repeating unit could be stored. MS/MS serves as a convenient information read-out.[316] The reported data density is clearly advantageous to the prior art. The scalability of the process due to the synthesis in solution can be an important aspect for potential applications. On the other hand, the synthetic access including chromatographic purification steps is tedious and—for now—limited to tetramers (96 bits·macromolecule^{-1}). Conditions and yields are depending on the synthesized sequence. However, the here-described work is an inspiring example of the diversity of abiotic digital polymers and their potential in bearing features that are superior to the biotic molecular data storage systems.



Scheme I-18: Synthesis of information-containing macromolecules from several multi-component reactions. Passerini three-component reaction i): carboxylic acid (intermediate or starting), aldehyde (F), (ABCDE) isocyanate macromonomer, DCM, rt, 3 d; benzyl deprotection ii): H_2, Pd/C, EtOAc (or THF) rt, 24 h.

5 Conclusion and Aim of the Thesis

It has been shown that sequence control in synthetic polymers is an emerging field in polymer science. The study of sequence-defined biomolecules can be a source of inspiration for polymer chemists. It can be deduced what a powerful concept sequence definition is and what is still to be achieved in artificial systems. At the same time it is possible to postulate limitations of biopolymers in some applications and overcome them with an artificial polymer design. It has been shown that the toolbox for sequence control in synthetic polymer science has enlarged with increasing speed. Radical living/controlled polymerizations were certainly a breakthrough in this context. More recently, manifold strategies of multistep-growth syntheses have proven to yield sequence-defined polymers that can be constituted of a variety of functional monomers. A less recent approach, yet newly discovered by polymer chemists, is phosphoramidite chemistry that is tolerant to a wide range of functional monomers and gives a convenient access to long sequence-defined poly(phosphodiester)s.

With these tools, concepts and strategies in hand, a new generation of functional polymers is accessible and fascinating systems can be discovered in the future. One such example is without doubt the current research on molecular data storage systems. Tailor-made functionality can be inserted to obtain digital polymers with features that may be superior to state-of-the-art DNA-based storage systems. That has already been reported for digital polymers with increased information density and a new information read-out that is—pivotal to specialty application—relying on tandem mass spectrometry. In the frame of the present thesis these very challenges are to be addressed. Digital poly(phosphodiester)s are investigated due to their convenient accessibility using phosphoramidite chemistry.

It is the first goal of the present work to broaden the available functional space of such polymers. That could theoretically be done by the synthesis of a variety of phosphoramidite monomers that would constitute the polymers. Instead, a binary modification protocol should be developed that would drastically reduce synthetic efforts because it would be the same sequence-defined precursor that could be modified with a variety of different moieties. A possible application of the methodology to gain side chain control is the screening for digital poly(phosphodiester)s that could be sequenced by nanopore technology.

Information density is another aspect that should be aimed for. While new strategies were already leading to synthetic digital polymers with an increased information density, the organization of digital polymer libraries can be miniaturized by a photocontrolled digital microarray fabrication. A photocontrolled synthetic strategy to obtain abiotic oligo(phosphodiester)s should thus be developed in this context.

Finally, it should be shown that polymer chemists can benefit from the almost unlimited chemical space that is available for the synthesis of tailor-made digital polymers to induce features that cannot as easily be achieved with biopolymers. A system should be developed that allows to code information in two different logic layers of a digital polymer. While the first information layer is directly accessible and can be read out by tandem mass spectrometry sequencing, light should serve here as the stimulus to reach the second information layer of the polymer that would—after loss of photoremovable groups—reveal another data set.

Chapter II:

Control over Side Chain Information of Sequence-defined Poly(phosphodiester)s

1 Introduction

This chapter presents how the side chain information of sequence-defined polymers can be controlled after polymer synthesis. That is an important aspect for researching digital polymers when a screening for special features is required and a scaffold with different side chains is investigated. To the best of the author's knowledge, the here-presented work is the first example of such approach: a method is described allowing to address all monomeric units of a defined sequence to a quantitative post-polymerization modification. The sequential approach enables control over side chain information while homogeneity of the sample as well as its binary encoded information are preserved.

Finding new sequencing techniques to access molecular encoded information is clearly the motivation of the research project. In the following paragraphs, the motivation for this research project as well as its context are described in detail. The aims are further explained and relevant state-of-the-art is briefly summarized.

1.1 Motivation & Context

Research on sequence-defined polymers beyond biopolymers is still in its infancy. Consequently, it is believed that systems or materials with features relying on the defined spatial organization of monomeric units are yet to be discovered. The motivation for the studies presented in this chapter is to supply polymer chemists with a tool and a reliable methodology to access a variety of polymers to screen for a particular combination of comonomers that shows desired features. Such screenings would normally require the researchers to synthesize the same variety of monomers. That could potentially be a tedious and sometimes chemically impossible task when required side chains were not tolerant towards the conditions of polymerization.

Followingly, a particular motivation to investigate such tool for side chain information is presented. It is one of the goals of Lutz' research group to investigate sequence-defined polymers and their use as digital polymers that store binary data at the molecular level. It has been presented in Chapter I: 4 *Sequence-defined Polymers for Information Storage* (pp. 46ff.) that the synthesis of such polymers—and poly(phosphodiester)s in particular—is already much advanced. In other words: *writing* a message that is encrypted on the molecular level is contained in the state-of-the-art and different tools are known for implementation. In addition, the sequencing by mass spectrometry has already tremendously progressed during recent years.[11] Though MS-reading of molecularly encoded data is relatively easy[12] and can be processed automatically,[283] it is a destructive analysis tool and the *reading* process consumes material each time molecular data is accessed. Therefore, the development of a non-destructive sequencing tool that is compatible with abiotic poly(phosphodiester)s is needed. Nanopore sequencing represents such a smart technique and, as some decades of research were required to find conditions to use it during DNA-sequencing,[13, 317] it was considered that abiotic poly(phosphodiester)s could be tuned to make also this polymer class readable by means of biological or solid-state nanopores. The tuning should affect the side chains of a binary sequence to optimize the design of two monomers that represent two distinguishably readable repeating units within a defined sequence. Improving the sequencing of abiotic poly(phosphodiester)s is hence the particular motivation to developing this quantitative and sequential post-polymerization modification protocol—the subject of the present chapter. The concept of nanopore-sequencing is described in Figure II-1 and followingly on page 62.

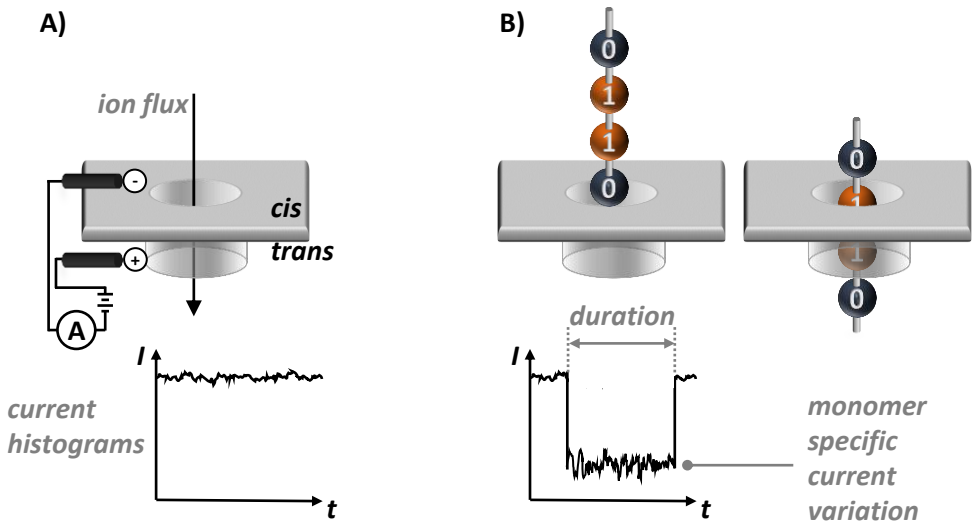


Figure II-1: Nanopore Sequencing. A) open nanopore and maximum ion flux; B) closed nanopore: threading/translocation of a polymer sequence results in a current variation.

Depicted in Figure II-1A, a nanopore system comprises of two electrolytic solutions—separated from another—on the *cis*-side and the *trans*-side of a membrane. A nanometer size hole constitutes the only connection between said chambers and when a constant electric field is applied, an electric current is observed in the system. Usually, potassium chloride is used and a constant flux of chloride ions through the nanopore from *cis* to *trans* is established. When a polyelectrolyte, like a poly(phosphodiester), is applied to the *cis*-side of the membrane, it is affected by the same electric field. In this case anaphoresis transports the negatively charged sequence to the site of the nanopore. When it threads through—the situation is depicted in Figure II-1B—the *open pore current* is hampered for the time the sequence interacts with the pore. The maximum *open pore current* is reestablished after complete translocation of the molecule. When the sequence interacts with the pore, it occupies space inside the pore channel and less ions can flow through the hole. Sequencing is then based on the development of different interactions between (bundles of) monomers and the pore. A distinct ion current can then be associated to a (bundle of) monomer that is present inside the pore at a certain time. With an adequate observable current resolution and data processing tools, the primary structure of a sequence is directly read out when it passes through the pore. Given the speed of translocation and the non-destructive single-molecule analysis, nanopore sequencing is a popular sequencing tool.[280, 318]

Nanopores can be produced from a variety of materials for solid-state nanopores.[319] Historically, biological nanopores were first investigated: a bilayer of biological phospholipids is produces in aqueous media. Transmembrane proteins penetrate the bilayer, assemble and form thereby the desired pores. α-Hemolysin (α-HL)[320] and *mycobacterium smegmatis* porin A (MspA)[321] are among the best-suited biological nanopores for DNA sequencing. It is, however, not a trivial task as more than one monomeric unit is passing the restriction zone of the pores in a microsecond.[321] Protein engineering placing e.g. a polymerase at the entry of the pore is being used to slow down the translocation of samples through the pore to result in an increase of the technique's accuracy.[320, 322] Another limitation is the length of the restriction zone within the transmembrane protein channels and it has been found that it is in fact the specific interaction of quadromers that is being measured when passing a sequence through a pore.[323]

In summary, nanopore sequencing of DNA is an already much developed technique and given the importance of this class of biomolecules in life sciences, methods have been optimized for discrimination of sequence information defined by the order of the four chemically quite similar DNA nucleotides. Translocation of abiotic sequence-defined poly(phosphodiester)s had not been reported before. A plausible way to benefit from nanopore sequencing during the read-out of digitally encoded

polymers is thus to *write* binary data in a defined polyelectrolyte and screen different side chain motifs for their influence on the polymer's sequencability using state-of-the-art nanopores. In other words, the motivation for the present project is to follow an inverse strategy to present research on nanopores: rather than tuning the nanopore itself, the analyte can be tuned taking into account the wide range of accessible chemical space. A higher degree of chemical difference within a set of two monomers is hereby thought to overcome limitations observed for DNA nanopore sequencing.

1.2 State-of-the Art in Precision Post-Polymerization Modification

In brief, a quantitative and sequential post-polymerization modification protocol should be developed for sequence-defined polymers, in particular for digital poly(phosphodiester)s. Nowadays, polymer chemists have an increasing tool box at their disposal that allows post-polymerization modification to access e.g. functional polymers.[324] Among them, the copper-catalyzed alkyne–azide cycloaddition (CuAAC) performes particularly well. The reaction between the highly energetic moieties is characterized by its narrow chemoselectivity that makes quantitative conversions likely while the moieties are tolerant towards a broad range of conditions during polymerization of monomers. For these reasons CuAAC is a popular tool in macromolecular engineering[325] and was the reaction of choice as the basis of a sequential post-polymerization modification methodology.

It has already been shown earlier by Lutz's research group that monomers bearing alkynes protected with different moieties could be used for a sequential post-polymerization modification. Those different monomers were addressable by CuAAC in a predefined order regarding the alkyne deprotection conditions that were applied between the modification of two different alkyne monomers.[141] In this way, a variety of functional polymers was potentially accessible with a limited number of monomers benefitting from the ultraprecise insertion strategy (see Figure I-12 in Chapter I: 2.2.2 Chain-growth Polymerization on pp. 26f.) rather than synthesizing a large library[139-140] of functional monomers.

Carell and coworkers reported a sequential post-polymerization methodology of oligonucleotides.[326] They synthesized analogues of the conventional dT and dC phosphoramidite monomers that exhibit chains with a terminal, TMS-protected or TIPS-protected alkyne attached to the pyrimidine's 5-position. With these *modifiers*, up to three different labels could be inserted in a given oligonucleotide chain after phosphoramidite chemistry-based polymerization. Interestingly, motifs too unstable to sustain the harsh cleavage conditions (cleavage of an oligonucleotide from the support using ammonia or TMS-deprotection with acid), were incorporated by sequential CuAAC in solution. For the final modification, a TIPS-protected moiety was used as its alkyne can be liberated under the mild conditions of treatment with tetra-*n*-butyl ammonium fluoride. More tolerant labels were inserted directly when the desired oligonucleotide sequence was still bound to the solid support by first reacting the monomer with a terminal alkyne with the azide of choice and then liberating the TMS-protected alkyne moiety with diluted acetic acid and performing the second CuAAC with another azide on the solid phase. Afterwards, the desired doubly modified oligonucleotide was cleaved and washed from the support material. As the TMS and TIPS protective groups are orthogonal, using all three moieties—namely a terminal alkyne, a TMS-protected alkyne, and a TIPS-protected alkyne—allows the insertion of three different labels by a sequential post-polymerization modification methodology. The here-described literature report is the closest state-of-the-art system that had been found. It is, however, dealing with oligonucleotide labeling and intrinsically contains limitations that does not make it a globally applicable methodology. At first, it was not shown that a quantitative modification would be achievable when a higher density of *modifiers* were to be used. During each modification step only one alkyne moiety reacted while the total number of monomers in the sequence was 15 or 16. And maybe more importantly, if not completely synthesized on a support material when tolerant side chains were to be introduced, the purification of modified oligonucleotide sequences was based on precipitation

in ethanol. Practical solutions for the latter described limitations during purification would be required in a more global methodology.

Recently, Lutz reported the parallel CuAAC modification of several terminal alkyne-bearing monomeric units of abiotic sequence-defined oligo(phosphodiester)s.[10] That finding clearly indicates that CuAAC modification is a tool that can be globally expanded to sequence-defined poly(phosphodiester)s. Yet, a sequential modification had not been reported until now.

1.3 Aims

In the frame of data storage using digital polymers, the superordinate long-time goal is to improve the sequencing of such polymers. The synthesis has already advanced. Though remarkably precise, the best developed read-out—*sequencing by mass spectrometry* (see Chapter I: 4 Sequence-defined Polymers for Information Storage, pp. 46ff.)—involves disadvantages like a destructive analysis and limitations in the analyzed sequences' lengths. For the earlier described advantages linked to *nanopore sequencing*, the technique is the choice for further investigations. A digital polymer—namely a sequence-defined poly(phosphodiester) encoding binary information—should be designed and it should be readable by passing it through a nanopore.

This superordinate goal was to be tackled step by step and in a parallel approach. A collaboration with experts in nanopore sequencing had to be established. Simple abiotic poly(phosphodiester)s from Lutz's laboratories should be investigated with state-of-the-art nanopores to gain a basic understanding of their behavior in the system. This task was to be supported synthetically but the responsibility for the actual studies would be with the collaborators. Ideally, based on the finding the next stage would require a screening for those abiotic monomers with desirable properties. In other words, monomers' side chains have to be tuned until a pair of such optimized monomers is found whose interactions with the nanopore when translocating through it are that different that the induced current variations would clearly indicate which monomer is passing the pore at a certain point in time.

During the time the collaborators are gaining a general understanding of abiotic poly(phosphodiester) translocation through their nanopores, a methodology should be developed that will then allow for an efficient screening. This methodology should comprise a platform that is building on the two closest state-of-the-art systems by Carrell[326] and Lutz,[10] fusing them to a globally applicable sequential post-polymerization modification methodology. The system should thus be based on efficient phosphoramidite chemistry-derived abiotic sequence-defined poly(phosphodiester)s as the backbones of the platform. Only two monomers should thereby be employed to form a binary information-containing precursor. Those two monomers should comprise either alkynes or protected alkynes to guarantee a sequential modification approach. Model compounds should be synthesized that are suggestive of the global applicability of the so-developed methodology including as well the introduction of less tolerant side chains.

In a third phase the platform should be used for screening a library of polymers with different side chains and a potential *hit* would deliver a new sort of poly(phosphodiester)s that is easy-to-read by means of nanopore sequencing.

2 Results and Discussion

This subchapter presents and discusses the results that have been obtained during the investigations that led to the development of a sequential binary post-polymerization modification protocol. First, the design of the underlying platform is justified. Based on that, new monomers, other reagents and support materials were needed. Their synthetic access is discussed before results from polymerizations

are presented. The strategy of sequential CuAAC modification is finally presented and the results are discussed in-depth.

2.1 Conceptual Design

In order to meet the needs that were defined in the description of the project's aims before, a poly(phosphodiester) platform should be synthesized that is composed of monomers that can be sequentially modified using CuAAC and different organic azides. The concept is visualized on the left in Figure II-2. Based on the required features of the system, the chemical identity of the system was defined and the monomer design was envisioned accordingly.

The platform's design is adapted from an earlier published system[10] where an abiotic phosphodiester sequence is composed of repeatedly linked 2,2-dipropargylpropyl units. These 0-units can undergo a cycloaddition with organic azides converting the terminal alkynes to a 1,2,3-triazole ring.

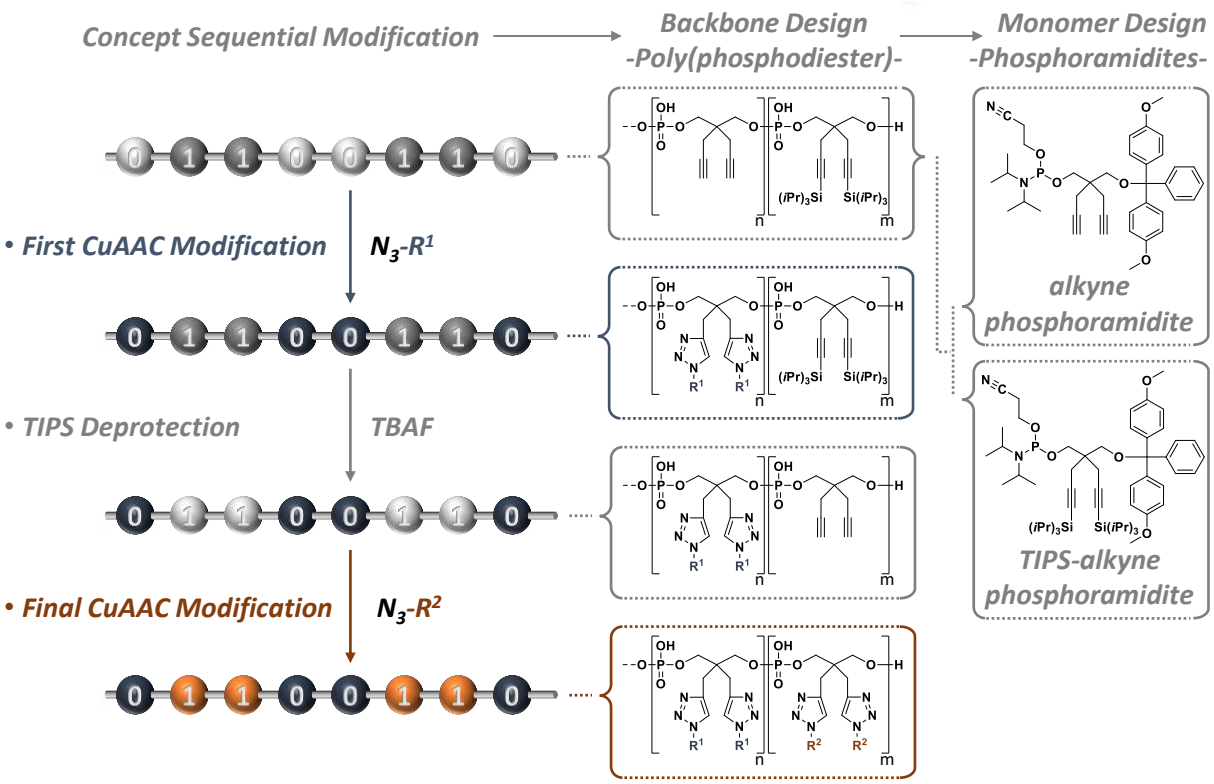


Figure II-2: Overview of conceptualization and design.

To enable the sequential binary CuAAC modification aimed at during this project, the platform has to be designed in such way that additional sequence-forming 1-units do not react during modification of the 0-units. Hence, the 1-units were designed to be the TIPS-protected analogues of the reported 0-units. The strategy relies on the report by Carell and coworkers[326] revealing that the TIPS-protective group does not only protect the alkyne moiety from a conversion during CuAAC conditions, it can also be removed with TBAF which allows the insertion of a broad range of side chains during the modification step before. In conclusion, a platform is designed that is composed of 0-units and 1-units that bear terminal alkyne moieties and TIPS-protected alkyne moieties, respectively. A first CuAAC inserts desired side chains at the sites of 0-units. Followingly, 1-units can be deprotected using TBAF to create terminal alkynes that can then undergo a final CuAAC modification inserting another sort of side chain. The initial sequence-defined platform is accessible using phosphoramidite chemistry to copolymerize the alkyne and the TIPS-protected alkyne phosphoramidite monomers that are depicted on the right of Figure II-2.

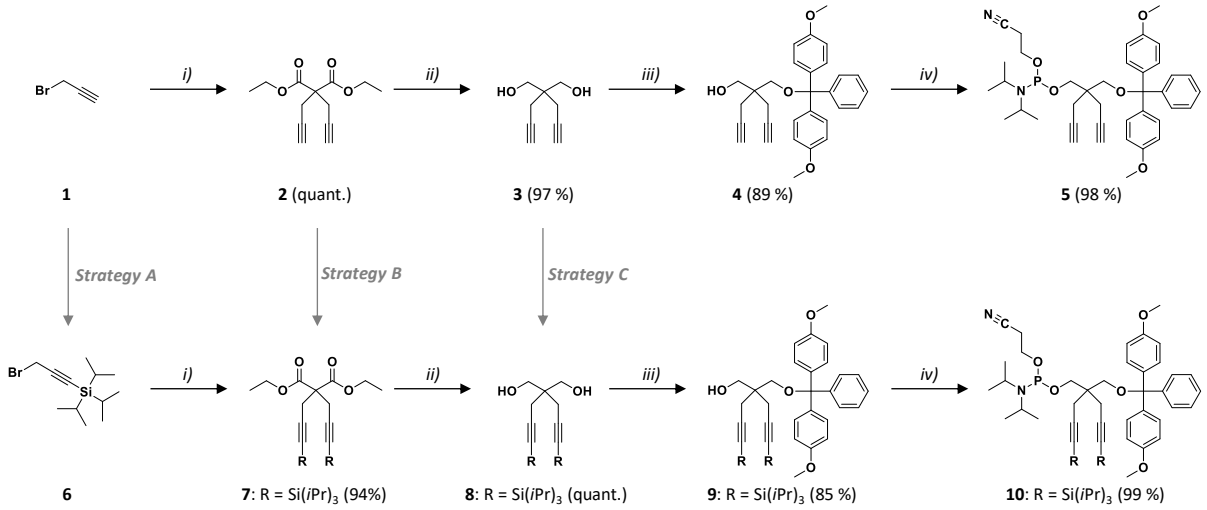
2.2 Monomers, Organic Azides and Solid Supports

The following section presents and discusses results from organic synthesis. Being an essential part for the realization of the project, it has, however, not been key object in the published report dealing with the topic.[327] A deeper discussion of different chemical routes under investigation is thus presented herein. Synthetic protocols for all monomers, other reagents and solid supports synthesized during the method development are detailed in the Experimental Section (Experimental Protocols Chapter II, pp. 137ff.).

2.2.1 Monomer Synthesis

Phosphoramidite chemistry is an efficient tool for linking different small molecule diols in a sequence-defined manner to another. The link between them is a phosphodiester bond that is formed during the polymerization as discussed in subchapter *Chemical Synthesis of Poly(phosphodiester)s* (pp. 33ff.). This principle of design is also reflected in the way monomers are to be synthesized. The key part that makes the monomers different and that already has to exhibit the desired features is the diol core. After isolation of the diol a hydroxy group has to be protected with e.g. DMT to ensure non-random polymerization of the monomer. The other hydroxy group is designated to exhibit a reactive precursor for the phosphodiester: the phosphoramidite that gave the method its name.

Phosphoramidite monomers exhibiting the features for a sequential binary modification were synthesized by an initial DMT mono-protection of diol **3** and its TIPS-protected analogue **8**, respectively, followed by conversion of the primary alcohol of DMT mono-protected species **4** and **9** to the corresponding phosphoramidites **5** and **10**, respectively (see Scheme II-1). While the synthesis of diol **3** is already reported[328] and DMT mono-protection of diol **3** with subsequent conversion to the phosphoramidite monomer was developed earlier by Lutz and coworkers[10], the synthesis of diol **8** with TIPS-protected alkyne moieties was investigated following different synthetic routes.



*Scheme II-1: Synthesis of phosphoramidite monomers **5** and **10**: i) 1. THF, diethylmalonate, NaH, 0 °C; 2. (TIPS-protected) propargyl bromide; ii) THF, $LiAlH_4$, -10 °C; iii) THF, pyr, DMTrCl; iv) DCM, DIPEA, 2-cyanoethyl-N,N-diisopropylchlorophosphoramidite, <0 °C. Strategies A, B, and C were investigated for insertion of the TIPS protective group and are detailed in Scheme II-2.*

Diethylmalonate is an excellent propan-1,3-diol precursor. The hydrogen atom in α-position to both carbonyl groups shows an increased acidity through mesomeric carbanion stabilization. The base sodium hydride that is added in step *i)* readily deprotonates diethylmalonate to enable the nucleophilic attack on propargyl bromides. Quantitative conversion is determined for the synthesis of malonate **2** in this work (see Experimental Section 2.1.1 on page 138). In the literature[328] a yield of 90 % is reported. The lower literature yield was possibly caused by the losses during purification by distillation.

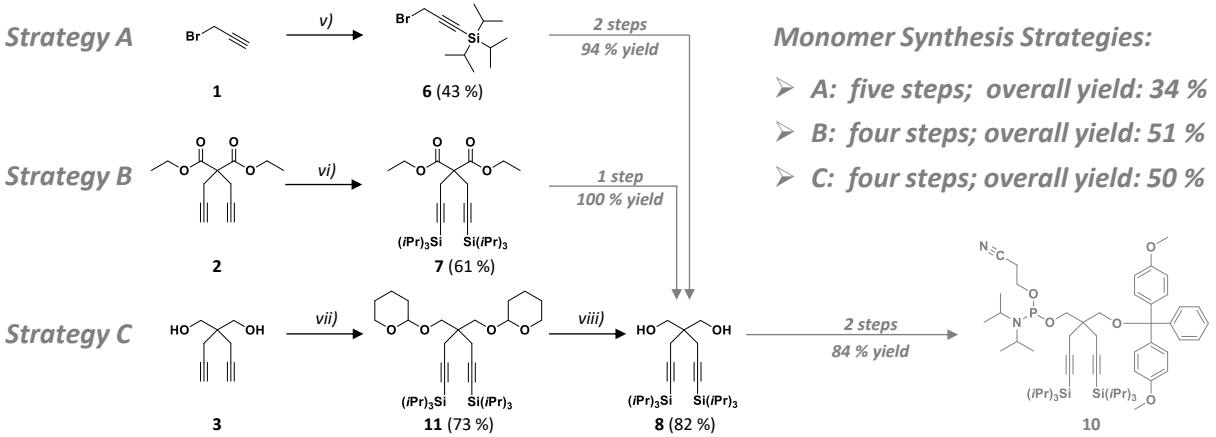
The following reduction of the ester moieties to alcohols using lithium aluminum hydride in step *ii)* yielded diol **3** near-quantitatively (see Experimental Section 2.1.2 on page 138). ^{1}H NMR analysis showed that the product was pure after the aqueous work-up. The literature protocol[328] includes Kugelrohr distillation that usually results in losses that explain the relatively low reported yield of 60 %. During the present work it was observed that crystallization from a warm diethylether solution gave crystals relatively big in size that were easy to handle during the preparation of the following conversion. During step *iii)* the diol **3** should be protected with the acid-labile 4,4'-dimethoxytrityl (DMT) protective group (see Table I-1 on page 42). DMT is added as its electrophilic chloride DMTrCl. Pyridine in the reaction mixture scavenges hydrochloric acid that is formed during the reaction between an organic chloride and an alcohol. A side reaction that cannot be completely avoided is the DMT double-protection. Favorable conditions for mono-protection are created by keeping the concentration of the chloride low to guarantee a maximum excess of hydroxy functions and hence render a mono-protection statistically more favorable. That is why DMTrCl is added in several small portions. Continuous addition with e.g. a syringe pump could not be achieved because of the low solubility of DMTrCl in common solvents. Another strategy includes a general diol excess in the overall experimental design resulting in more constant reaction kinetics due to the diol excess that is kept even near the end of the reaction (last DMTrCl addition). The latter strategy could increase the yield with respect to the chloride used. Yet, this is a beneficial approach when cheap commercial diols are in use. Here, however, diol **3** is product of a two-step reaction and thus scarce. Taking the discussed factors into account a 10 mol% excess was employed to result in a yield near 90 % for precursor **4** with respect to the chloride that was added (see Experimental Section 2.1.3 on page 139). In step *iv)* the DMT-protected precursor **4** was converted to the corresponding phosphoramidite in a similar reaction with 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite and diisopropylethylamine as a base. The yield is near-quantitative (see Experimental Section 2.1.4 on page 139). Step *iv)* and *v)* were already reported with similar yields.[10]

Steps *i-iv)* were analogously applied to the counterparts **6** - **9** with TIPS-protected alkynes. Just as satisfying yields were hereby determined (see 2.1.6 - 2.1.9, pages 140 - 141). Nevertheless, this synthetic route—described as *Strategy A* in Scheme II-1—was not the first choice from the beginning. The synthesis of the new TIPS-protected phosphoramidite **10** was investigated following three different synthetic routes as depicted in Scheme II-1. The three strategies are detailed below in Scheme II-2. The TIPS protective group is inserted at the earliest stage in strategy A. Here, propargyl bromide **1** is protected with TIPS in step *v)* leading to the TIPS analogue **6** with a 43 % yield (see Experimental Section 2.1.5 on page 140). Compared to the literature protocol[329] reporting a 41 % yield and a side reaction between the base *n*-BuLi with the product, the synthetic access could be dramatically simplified and a pure product was isolated. Yet, the efficiency remained unsatisfyingly low. It is probable that the deprotonated propargyl bromide is not stable at the reaction temperatures around -78 °C and undergoes a (slow) nucleophilic substitution reaction between the alkyne carbanion and another alkylbromide before it reacts with the added TIPS chloride.

The TIPS protective group is inserted in step *vi)* a step later in the scheme following strategy B. The double-protected malonate **7** was isolated with 61 % yield (see Experimental Section 2.1.10, p. 142). The TIPS-insertion *vi)* on the malonate **2** is probably more efficient than *v)* at the propargyl bromide **1** because the malonate **2** does selectively react with the added TIPS-chloride as other functions are not reacting at the applied conditions.

Strategy C's sense of purpose can be questioned. It starts from diol **3** that has been synthesized from the diol precursor malonate **2**. As the alcohol moieties are more acidic than the alkynes and thus silyl ethers were to be mainly formed upon conversion with a base and TIPS chloride, the diol (that has been created the step before) has to be protected again. Protection of the alcohols as a tetrahydropyranyl (THP) ether and subsequent alkyne-protection with TIPS was performed in step *vii)*

to give **11** with a 73 % yield (see Experimental Section 2.1.11, p. 142). Acid catalyzed hydrolysis *viii)* led to diol **8** with TIPS-protected alkynes (see Experimental Section 2.1.12, p. 143).



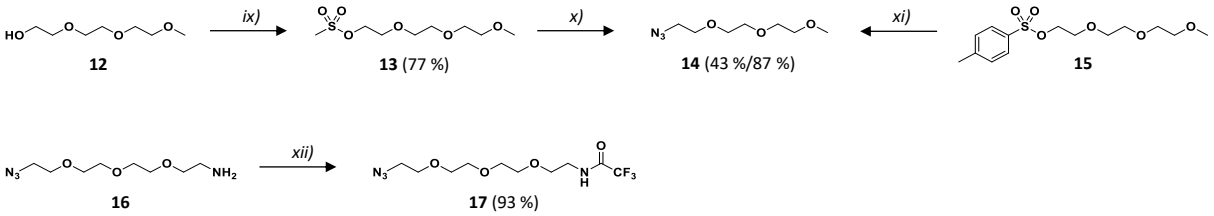
Scheme II-2: Monomer Synthesis Strategies A, B, and C. v) 1. THF, n*-BuLi, -77 °C; 2. TIPS-Cl, -77 °C; vi) 1. THF,* n*-BuLi, -77 °C; 2. TIPS-Cl, -77 °C; vii) 1. DCM, 3,4-dihydro-2*H*-pyran, cat. TsOH, 2. THF,* n*-BuLi, -77 °C; 3. TIPS-Cl, -77 °C; viii) MeOH, TsOH.*

When complex syntheses are planned, diversification should occur as late as possible in the scheme to avoid long parallel routes. Strategy C branches late in the scheme but compared to strategy B, the demand in resources for this reaction path is not reduced. In contrary, compound **2** is first *deprotected* to yield diol **3** only to insert again a protective group afterwards in the first reaction of step *vii)*. Strategy C is thus not recommended. When comparing the numbers, strategy A is obviously outperformed by strategy B. The latter is a step shorter and results in a 51 % overall monomer synthesis yield compared to 34 % for strategy A. On the other hand, also practical aspects are important in research laboratories and strategy A was characterized by less difficult purifications. Step *vi)* in strategy B was tedious and a relatively high amount of resources was needed to isolate malonate **7** from side products and excess reagents. Strategy A, in contrast, relies on an early TIPS insertion step *v)* that is much more simple to perform than step *vi)* when protecting the malonate **2** with TIPS. Yet, synthesis of bromide **6** resulted in a lower yield. But here, the losses referred to commercially available starting materials and not to precious intermediates. Followingly, also the isolation of TIPS malonate **7** in subsequent step *i)* was—as described before—efficient and required much less time and material resources due to a relatively simple purification. For practical reasons strategy A is thus the recommended synthetic pathway to access monomer **10** even though it requires one more reaction step and does not deliver the highest overall yield.

2.2.2 Synthesis of Organic Azides

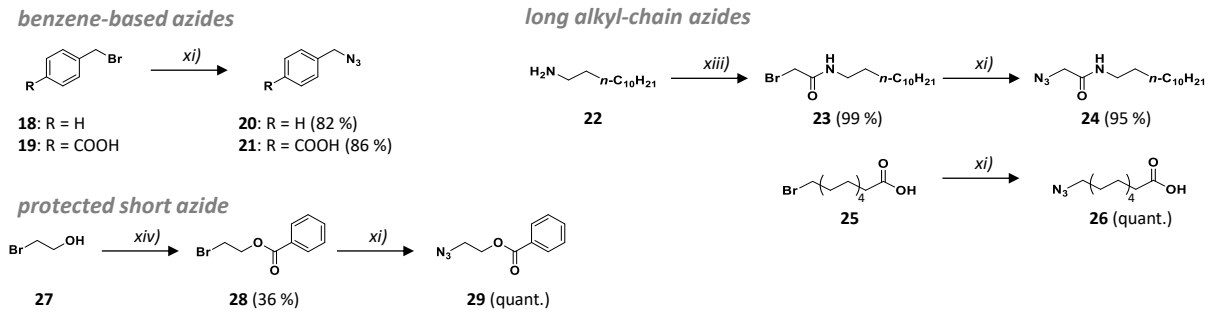
Different azides were synthesized that could potentially be coupled to terminal alkyne-bearing phosphodiester sequences. Among them were ethylene glycol-based chains that were considered to be optimal model compounds due to their solubility in a variety of solvents. Azide **14** was obtained starting from the corresponding alcohol triethylene glycol monomethyl ether **12**. In a first step *ix)* the primary alcohol was converted into a leaving group by either mesylation or tosylation to obtain mesylate **13** and tosylate **15**, respectively. The latter reaction was not performed by the author but in the same laboratory. Followingly, sulfonate groups in **13** and **15** can be substituted by an azide in *x)* and *xi)*. The yields presented in Scheme II-3 are unsatisfactory. Yet, the syntheses were already reported[330-332] and further optimization was not considered necessary as a sufficient quantity of azide **14** was easily accessed. It is plausible that most reagent losses occurred during the aqueous work-up of reaction mixtures and further extraction would have delivered the products in higher yields. The second azide **17** was designed to have additional features. It can be cleaved under basic conditions,

delivering a terminal amine that potentially allowed for subsequent and diverse modifications. In this regard trifluoroacetamide azide **17** is the model base-labile azide whose successful insertion would prove the global applicability of the method. By incorporation of a 2,2,2-trifluoroacetamide a new signal could be conveniently detected by ^{19}F NMR after CuAAC modification evidencing the intactness of the labile side chain. The trifluoroacetamide was obtained by reaction of the corresponding amine **16** with trifluoroacetic anhydride (TFAA) in *xii)*. Symmetric anhydrides are highly reactive transfer agents for carbonyl groups. Among nucleophilic attack of the carbonyl, a carboxylic acid is liberated – fueling the process due to its high leaving group quality. Unsurprisingly, the literature[333] protocol's low efficiency for this reaction was greatly surpassed.



Scheme II-3: Synthesis of oligo(ethylene glycol) azides. ix) DCM, Et_3N, MsCl, 0 °C; x) DMF, NaN_3, 80 °C; xi) DMSO, NaN_3, RT; xii) DCM, Et_3N, TFAA.

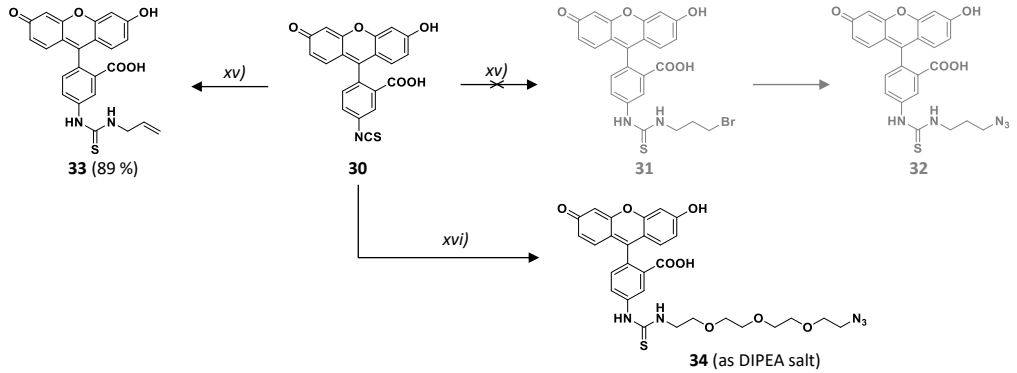
Besides model compounds **14** and **17**, several other azides were synthesized for future side chain screening. They should introduce differing properties to a sequence upon insertion. That is why benzene-based azides **20** and **21** with different polarities and azides **24** and **26** with long alkyl chains and different functional groups were among the target structures. A short azide **29** was synthesized bearing a base-cleavable protective group. As low molecular weight azides are explosive, the protective benzoate (Bz) group was used to decrease the ratio of azide nitrogen atoms to the rest of atoms in the molecule. Also a fluorescent azide **34** was synthesized.



Scheme II-4: Overview Synthesis of further azides to allow diverse side chain tuning. Azide substitution xi): DMSO, NaN_3, RT; amidification xiii) K_2CO_3, bromoacetyl bromide, DCM, 0 °C; Bz-protection xiv) DIPEA, benzoic anhydride, DCM, 0 °C.

Azides **20**, **21**, **24**, **26**, and **29** were obtained by treating them in step *v)* with an excess of sodium azide in DMSO at room temperature overnight. The synthetic routes are depicted in Scheme II-4. The reported yields are satisfying and it is worth mentioning that clean bromides **18**, **19**, **23**, **25**, and **28** yielded pure azides after removal of excess sodium azides, the reaction product sodium bromide and importantly DMSO during an aqueous work-up. Further purification was not needed. As no trace of unreacted bromide could be detected after the work-up, it is obvious that reaction *xi)* is quantitative, yet the extraction from the aqueous phase has not been complete in every case. Reaction control relied on ^{13}C NMR spectroscopy where a characteristic downfield shift of the signal of the CH_2Br to that of the CH_2N_3-carbon could be detected. Bromide **23** was obtained by a simple amidification *xiii)* of *n*-dodecylamine with bromoacetyl bromide. The acid halide contains the much more electrophilic carbon atom when compared to the alkyl halide. The nucleophilic attack by the primary amine **22** is assisted with the base potassium carbonate and is due to the reactivity differences in bromoacetyl bromide highly selective for amide formation. The amide function in **23** is stabilized by mesomeric effects and

not nucleophilic enough for intermolecular attack on another alkyl bromide. Bz-protection of 2-bromoethanol **27** with the corresponding anhydride in step *xiv)* shows an unsatisfying yield. The literature protocol reports a slightly higher yield of 58 %.[334]



*Scheme II-5: Attempts for synthesis of fluorescent azide **34**. Thiourea formation xv) 3-bromopropyl amine hydrobromide, DIPEA, MeOH, DCM; thiourea formation II xvi): 11-azido-3,6,9-trioxaundecan-1-amine **16**, DIPEA, $CHCl_3$.*

The insertion of a fluorescent side chain can be crucial for some applications. Thus the archetypical fluorescent dye fluorescein should be linked with an azide. A commercially available fluorescein derivative is its 4-isothiocyanate (FITC) **30** depicted in Scheme II-5. Isothiocyanates react efficiently to amines thus forming an unsymmetrical thiourea derivative. The link to a bromide was first attempted in step *xv)*. The desired thiourea bromide **31** was not isolated, instead the slightly basic conditions required for the reaction also led to an elimination reaction resulting in a terminal double bond. The undesired product **33** was isolated as a pure compound (see Scheme II-5) suggesting that both thiourea formation and the elimination side reaction are highly efficient. The strategy to conveniently circumvent an elimination was the choice of an amine already bearing the desired azide.

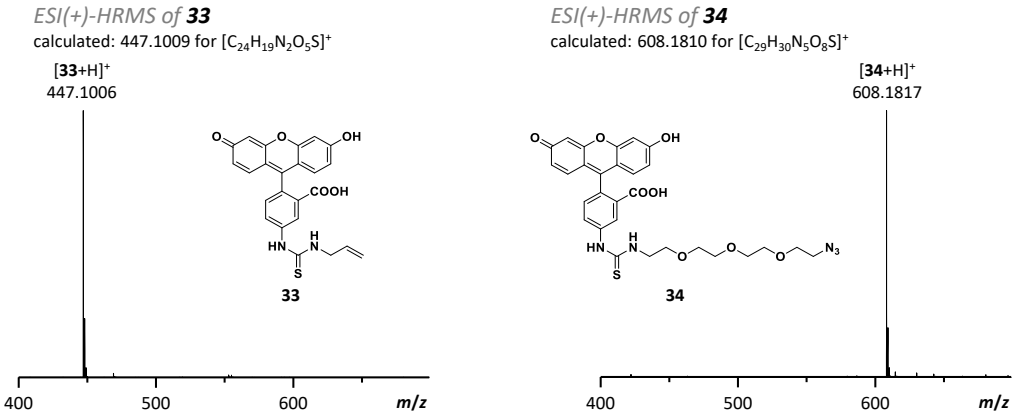
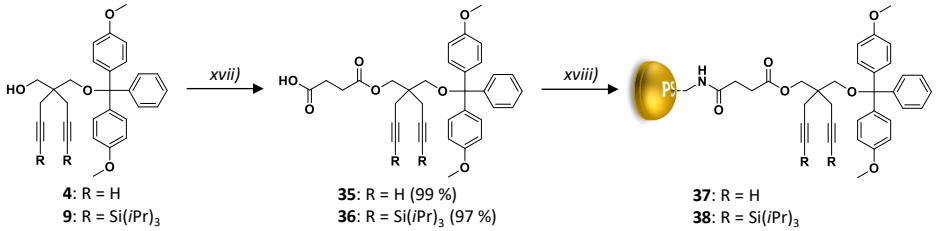


*Figure II-3: ESI(+)-HRMS of **33** (left) and **34** (right) shown in the m/z 400-700 range.*

Using one equivalent of 11-azido-3,6,9-trioxaundecan-1-amine **16** lead to the formation of fluorescent azide **34**. Also here in step *xvi)* DIPEA was used as a base. The excess DIPEA could be removed under vacuum; azide **34** was finally recovered as a bis-diisopropylethylammonium salt. The described reactions *xv)* and *xvi)* can be followed by NMR, yet, the spectra are complex and thus analysis by mass spectrometry could deliver evidence for the formation of **33** and **34** as depicted in Figure II-3. The fact fluorescent azide **34** was isolated as an ammonium salt should not negatively influence its use in CuAAC.

2.2.3 Solid Support Synthesis

Support materials are crucial for the synthesis of sequence-defined polymers (see 2.2.3 Multistep-growth Synthesis, pp. 28 ff.). The linker chemistry has to be chosen with regard to the reaction conditions during polymerization to avoid simultaneous cleavage of the growing sequence from the support. Base-labile linkers relying on esters meet this criterion for phosphoramidite chemistry (see Chapter I: 3.2.1 Solid Support Material, pp. 39ff.). For the realization of this project, it was thus decided to build upon the state-of-the-art that is presented in subchapter *Phosphoramidite Chemistry for Synthesis of Abiotic Poly(phosphodiester)s* (pp. 45f.): these support materials are composed of a first DMT-protected monomeric unit that is immobilized on a polystyrene matrix via a succinic ester. In a first step *xvii)* a precursor monomer **4** or **9** with a DMT-protected and a free alcohol is reacted with succinic anhydride, thereby forming the succinic esters **35** or **36**, respectively. The terminal carboxylic acid can be activated with DMAP and DCC for the following amidation step *xviii)* to react with the aminomethylated polystyrene resin to yield supports **37** or **38**. Ensuring that the support material is chemically inert under phosphoramidite chemistry conditions potentially remained amine moieties are capped with acetic anhydride (see Scheme II-6).



*Scheme II-6: Synthesis of solid polystyrene supports **37** and **38**. xvii) pyr, succinic anhydride, DMAP, 35 °C; xviii) 1. DCM, DMAP, DCC, aminomethylated polystyrene; 2. pyr, acetic anhydride.*

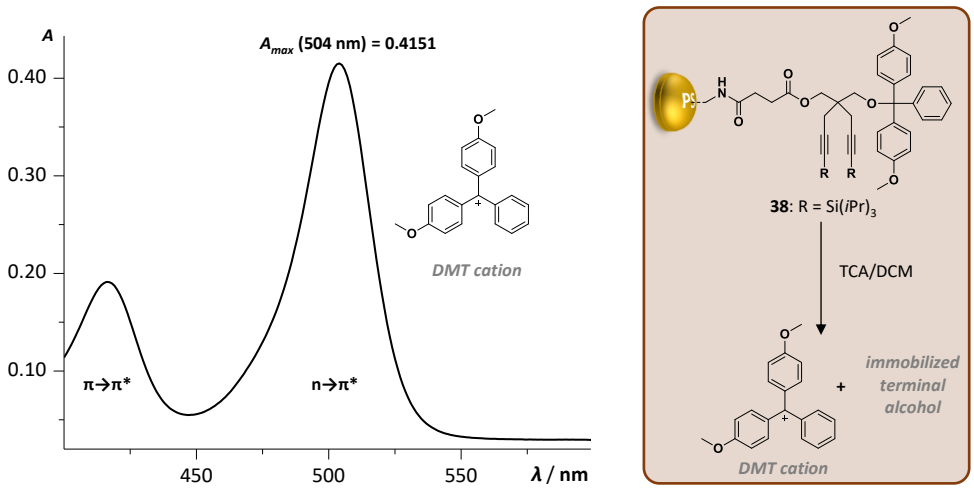


*Figure II-4: Determination of support loading by UV-Vis absorption. In the depicted example 78.8 mg of support **38** were cleaved with TCA in DCM. The support material was washed with further DCM and the filtrate was filled to 50 mL. An aliquot was diluted by the factor of 200 and the absorbance at 504 nm was determined to 0.4151. Given the extinction coefficient $\varepsilon_{504,DCM}$ = 76,000 the DMT cation concentration in the initial solution was determined to be 1.09 mM. Given the volume of 50 mL (n = c V) 54.6 µmol of DMT group were cleaved from 78.8 mg support **38**. That yields finally the reported 0.69 mol g^{-1} loading.*

Succinates **35** and **36** were obtained in satisfying yields due to the efficient base-catalyzed ring opening of the cyclic succinic anhydride. Likewise the immobilization of succinates was highly efficient. During amidation, the support material increases its weight by the weight of immobilized succinates. Taking also into account the weight loss of one water molecule per amidation, a theoretical maximum loading can be determined. The manufacturer indicates a loading of 1.41 mmol amine moieties per gram of resin. For clarity, this translates into the statement that a mole of amine moieties is contained in

709.2 g resin. If this mole of functionalities entirely condensates with the succinates, a mole of the immobilized monomer is contained in (709.2 g + 1 mol·$M_{succinate}$ – 1 mol·M_{water}) of modified resin. That results in a maximum loading of 0.80 mmol·g^{-1} for support **37** and 0.64 mmol·g^{-1} for support **38**, respectively. These loadings were achieved and interestingly, the reaction was as efficient when only one equivalent of succinate was used as it is shown for support **38** (see 2.3.4 on page 148). Future support syntheses can hence be performed without wasting precious succinates.

Reaction control by UV-Vis absorption is shown in Figure II-4. A share of the support was washed with TCA in DCM. An intensive red coloration was observed immediately stemming from the absorption of the cleaved DMT cation as shown on the right of Figure II-4. The red filtrate is then analyzed by UV-Vis spectroscopy after being subject to a defined dilution series. The spectrum that is depicted on the left of Figure II-4 shows two characteristic absorption maxima at 416 nm for the $\pi\rightarrow\pi^*$-transition and at 504 nm for the $n\rightarrow\pi^*$-transition, respectively. The latter maximum was determined and with the defined extinction coefficient[335] in DCM the concentration of the sample can be determined.

2.3 Phosphodiester Synthesis

Oligo- and poly(phosphodiester)s were synthesized using monomers **5** and **10** to obtain defined precursor sequences that are sequentially modifiable. Required sequence control is guaranteed by a multistep-growth polymerization (see 2.2.3, pp. 28ff.) relying on phosphoramidite chemistry (see 3.2, pp. 39ff.) using different solid support materials. The results obtained during manual oligomer synthesis are discussed first followed by discussion of the results for automated polymer synthesis.

2.3.1 Manual Oligomer Synthesis

Manual precursor oligomer synthesis following essentially a procedure from the literature[10] was performed on solid polystyrene support **37** when a leading alkyne monomer was desired. Respectively, support **38** was used to obtain sequences with a leading TIPS-protected monomer (see Figure II-5). Syntheses were usually performed on an approximately 50 µmol scale and after charging the support **37** or **38** to the reactor, the exact scale was determined: all phases from consecutive TCA/DCM washings resulting from the first DMT-deprotection step *i)* were collected and diluted to a certain volume. Analysis of the UV absorption is used to determine the concentration of the DMT cation that is directly linked to the reaction scale. Here, it has been found that additional TCA has to be added because small impurities could already quench the cation's absorption.

As it is detailed in Figure II-5, after DMT-deprotection *i)* a terminal alcohol is immobilized on the support. It must be washed with a low concentration of base to quench the access acid and further washes with anhydrous solvents ensure optimal dry conditions for the following phosphoramidite coupling *ii)*. Here, two equivalents of monomer **5** or **10** were added to the resin and subsequently activated by addition of four equivalents of the weak acid 1*H*-tetrazole. The so-formed phosphite bond is converted to the more stable phosphate bond during oxidation *iii)* with iodine and water. Steps *i)*, *ii)*, and *iii)* are repeated. During each phosphoramidite cycle the choice of monomer in *ii)* defines the identity of the sequence at this very position. After having constructed the desired sequence, the phosphate protective 2-cyanoethyl groups are cleaved *iv)* using diluted piperidine. The moderately basic conditions ensure selective phosphate deprotection while the succinic ester link between resin and sequence is retained. The terminal DMT-group is cleaved *i)* after having washed away acrylonitrile formed in *iv)* before. The order of steps *iv)* and *i)* is considered important because the terminal alcohol could potentially react with acrylonitrile. A wash with sodium thiosulfate *v)* is followed by cleavage *vi)* of the succinic ester linkage allowing to wash the desired sequence from the support material.

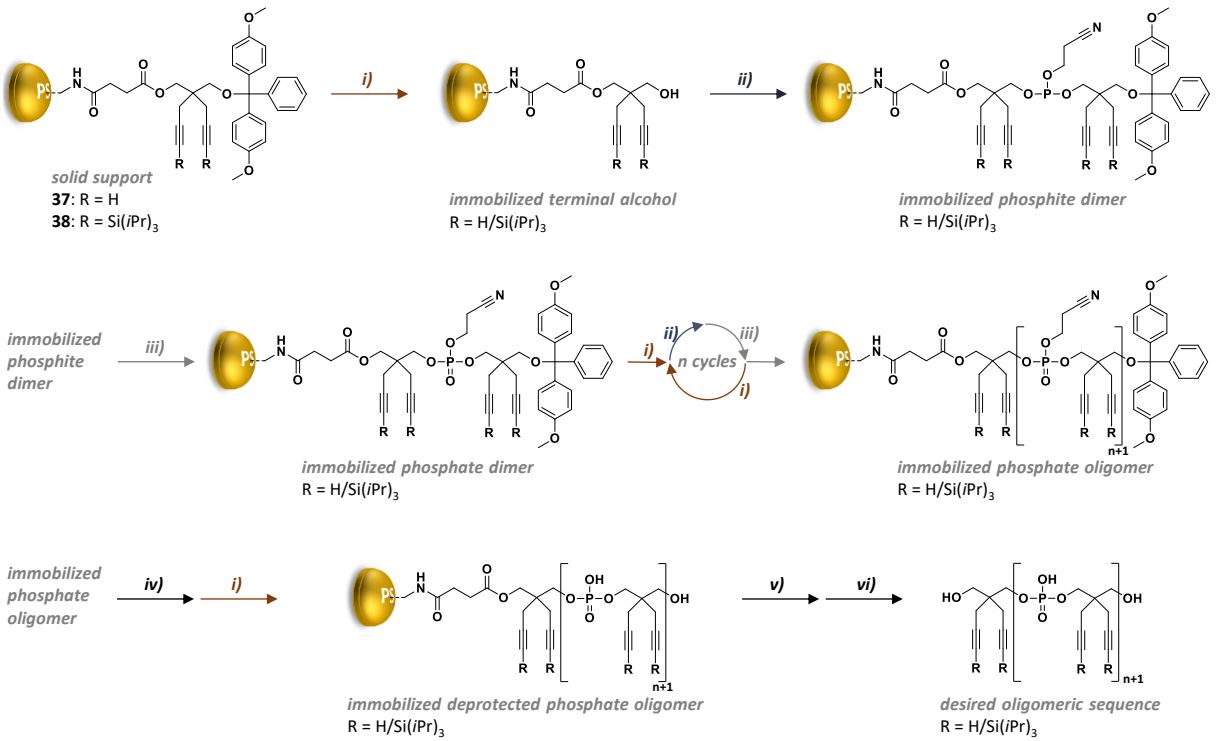


*Figure II-5: Manual synthesis of defined oligo(phosphodiester) precursor sequences. DMT-deprotection i): TCA, DCM; phosphoramidite coupling ii): MeCN, 1H-tetrazole, monomer **5** or **10**; oxidation iii): THF, iodine, 2,6-lutidine, water; phosphate deprotection iv): MeCN, piperidine; iodine reduction v): $Na_2S_2O_3$, water; cleavage from resin vi): methylamine, ammonia.*

Precursor sequences **S1**–**S4** were manually synthesized following the strategy depicted in Figure II-5. The sequences were analyzed by ^{1}H NMR, ^{31}P NMR and ESI(-)-HRMS. Details are given in Table II-1 and the full detail synthetic protocol is reported in the Experimental Section: 2.4.1 (pp. 149f.). It can be concluded that the desired sequences were obtained in moderate yields using both support **37** and **38** and there was no hint the sterically demanding monomer **10** would require an adapted protocol. The results suggest that any binary sequence can hence be synthesized using monomer **5** and **10**. A slight polydispersity owing to missed coupling steps was, however, found for all manually synthesized sequences. Bearing into mind the immobilized sequences were washed from the support without further purification, this observation harmonizes with the expectations because phosphoramidite couplings are near complete though not quantitative. Longer sequences synthesized with automatized protocols include capping steps and a final purification routine that ensures homogeneity of the isolated sequence. The here presented protocol is not simply cloning already reported conditions.[10] Interestingly, a side reaction was observed that could be suppressed by adding step *v)* to the synthetic protocol. The investigations are discussed in the next section.

*Table II-1: Overview and ESI-HRMS characterization of manually synthesized precursor sequences **S1**-**S4**[a].*

	Support	Sequence	Yield	m/z_{th}	m/z_{exp}
S1	**37**	01100110[b]	53 %	966.1367[c]	966.1375
S2	**37**	0110[b]	73 %	708.8620[d]	708.8626
S3	**38**	1001[b]	68 %	708.8620[d]	708.8638
S4	**38**	1001	69 %	708.8620[d]	708.8628

[a] The numbers 0 and 1 denote for phosphate linked repetitive units derived from diols **3** and **8**, respectively. [b] step *v)* not performed, side reaction: species with masses that were higher by increments of 126 Da detected. [c] $[M-3H]^{3-}$. [d] $[M-2H]^{2-}$.

2.3.2 A Side Reaction and its Suppression

Initially, the literature protocol[10] was strictly followed. It does not include a sodium thiosulfate wash. Interestingly, as indicated in Table II-1, so-attained sequences **S1**-**S3** were contaminated with substances that were shown by ESI-HRMS to have masses that were higher by increments of 126 Da than the desired sequences. A theoretical random proton iodine exchange within a sequence would provoke such observation. Additionally, such species had never been detected after polystyrene support-based synthesis of other sequences that did not contain free alkyne moieties. Hence, it was hypothesized that—despite the extensive washing steps—traces of iodine that remain entrapped in the polystyrene support matrix can react with terminal alkyne units under the conditions of the cleavage reaction *vi)*. Indeed, use of aqueous ammonia and elemental iodine are reported to lead to a direct iodination of monosubstituted alkynes.[336]

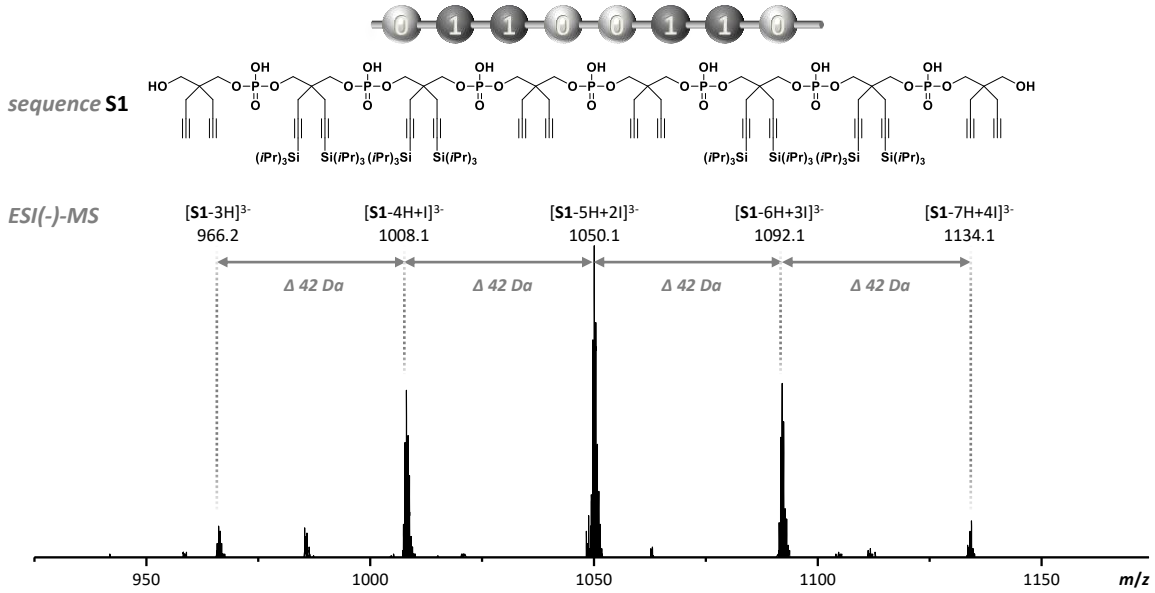


Figure II-6: ESI(-)-MS spectrum (the m/z *925 – 1175 range) recorded for the precursor sequence **S1**.*

A representative example is depicted above in Figure II-6. In the *m*/*z* region 925 – 1175 of the mass spectrum the desired sequence is detected as a triply deprotonated species at *m*/*z* = 966.2 (for a broader m/z range see Figure V-25, p. 201). The species with higher *m*/*z* values are detected with increments of 42 Da at 1008.1, 1050.1, 1092.1, and 1134.1 corresponding to up to four hypothesized direct alkyne iodinations. The difference between the detected peaks is only a triple of the 126 Da weight gain calculated for a iodination because the triply charged species (*z* = 3) are observed in this example. The intensity of the peaks shown in Figure II-6 represents the count of detected species, yet, it does not directly reflect the ratio of iodinated to remained alkynes because the ionizability of different species could be affected. Qualitatively, a statistical iodination can be deduced from the results.

It has been shown that 1-iodo alkynes do as well undergo alkyne-azide cycloadditions.[337-338] Homogeneity of the samples was nevertheless aimed at to develop a protocol that is valuable for diverse applications as the storage of binary information at the molecular level. Random iodination should therefore be prevented and an obvious strategy implies the removal of iodine before bringing the alkyne-bearing sequence in contact to strong base. Iodometry is relying on quantitative iodine reduction using aqueous sodium thiosulfate.[339] This efficient conversion was used to eliminate remaining traces of iodine in the polystyrene support material by addition of a sodium thiosulfate washing step *v)* before immobilized sequences were cleaved from the support. The same binary code '1001' was once coded into a sequence without additional iodine reduction *v)* to yield **S3** and once it

was obtained after iodine reduction *v)* to yield **S4** (see Table II-1). From the analysis by ESI(-)-MS depicted in Figure II-7 the efficient suppression of side product formation by an additional sodium thiosulfate wash *v)* is evidenced.

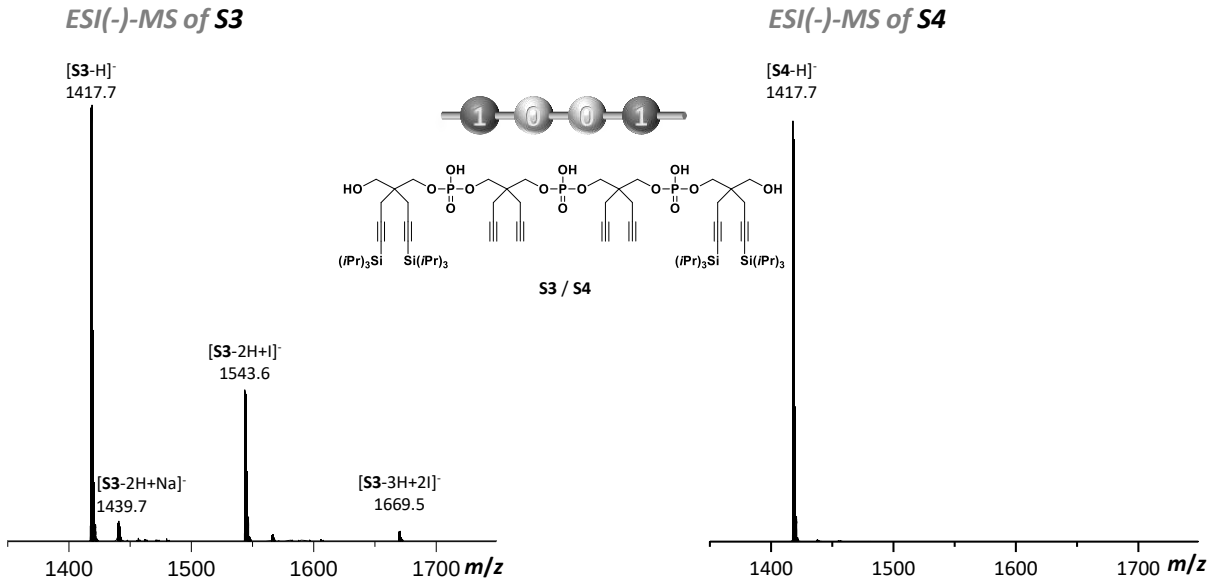


*Figure II-7: Comparison of ESI(-)-HRMS for **S3** and **S4**. Negatively charged (-1) species are detected in the* m/z *1350–1750 range.*

2.3.3 Automated Polymer Synthesis

Robotic precursor oligomer and polymer synthesis was performed following a procedure from the literature.[6] A commercially available dT-CPG support **39** (see 3.2.1, p. 39ff.) was used with a 1000 Å pore size that had already been shown to allow efficient synthesis of abiotic poly(phosphodiester)s with a degree of polymerization above 100. A 1 µmol reaction scale was chosen. The first monomer is already immobilized on the support, namely it is 5'-DMT-protected deoxythymidine linked via a succinate ester bond at the 3'-alcohol position. The support **39** is trapped in powder form in a column between two frits. Attached to the oligonucleotide synthesizer, reagents can be delivered to the inlet and stream vertically through the column from the bottom to the top outlet thereby producing turbulences and dispersing the support material in this simple reaction chamber. As a result, different reagents will mix there and a maximum surface of porous glass support is in contact with reagents.

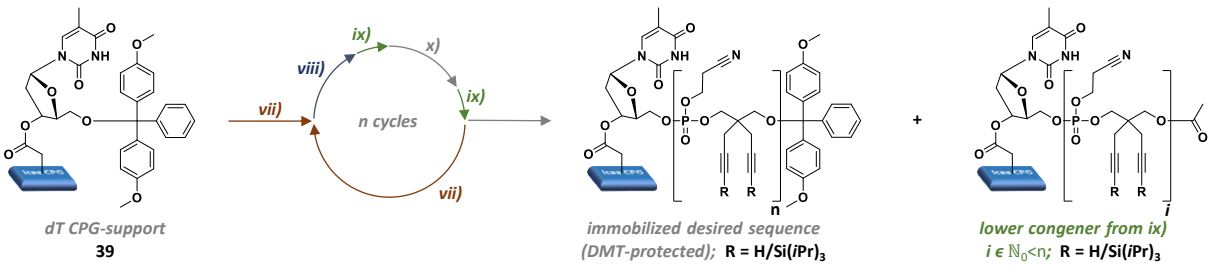


Figure II-8: Automated DMT-ON *synthesis of defined poly(phosphodiester) precursor sequences. DMT-deprotection vii): TCA, DCM; phosphoramidite coupling viii): MeCN, 5-ethylthio tetrazole, monomer **5** or **10**; capping ix): acetic anhydride, pyr,* N*-methylimidazole, THF; oxidation x): THF, iodine, pyr, water.*

Phosphoramidite chemistry is performed benefitting from the described system analogously to the manual phosphodiester synthesis, yet reagent delivery is software-controlled. The exact commands are given in the Experimental Section: 2.4.2 (pp. 149ff.). At the beginning, step *vii)* is applied to cleave the terminal DMT-group. After washing with MeCN the desired monomer **5** or **10** was delivered

together with activator 5-ethylthio tetrazole during the phosphoramidite coupling *viii)*. Non-reacted fractions were capped in *ix)* using acetic anhydride and *N*-methylimidazole as an activator. Oxidation *x)* with iodine converted the intermediate phosphite bond to a phosphate and another capping *ix)* step ensures maximum chain-termination for non-reacted chains from *viii)*. These cycles were repeated automatically delivering monomer **5** and **10** in a predefined order until the desired sequence was attained. The *Expedite* synthesizer automatically quantifies the UV absorption of the cleaved DMT group during *vii)* and produces a *trityl histogram* that gives a qualitative impression of the efficiency of the poly(phosphodiester) synthesis.

As described in Figure II-8, besides the desired sequence after *n* cycles, lower (capped) congeners with *i* added monomeric units are immobilized on the same support. Note that $i < n$ which is why only the desired sequence shows a terminal DMT group. During the synthesis of a (polar) oligonucleotide all sequences are cleaved from the support with base also removing the 2-cyanoethyl protective group in the same step. Subsequently, chromatographic means are applied to isolate the desired sequence relying on the relatively apolar DMT-group. In the case of sequences synthesized with TIPS monomer **10**, it was found that such purification was not applicable. The apolar TIPS-groups are present in both the desired sequence and the congeners and the change in overall polarity that stems from the presence of a terminal DMT-group is too small to ensure an efficient chromatographic purification.

For that reason, the success of automated precursor sequence synthesis was monitored without purification. To the *DMT-ON* synthesis depicted in Figure II-8 a final DMT deprotection step *vii)* was added and the mixture of desired sequence and its lower congeners was analyzed. For the synthesis of precursor sequences **S5**-**S8** the results are given in Table II-2.

*Table II-2: Overview and ESI-HRMS characterization of precursor sequences **S5**-**S8** synthesized using a robot.[a]*

	Support	Sequence	Yield	m/z_{th}	m/z_{exp}
S5	**39**	T 01000001	54 %	1288.9641[b]	1288.9659
S6	**39**	T 01001110 01001011	70 %	880.4022[c]	880.4030
S7	**39**	T 01000001 01000001 01000011	49 %	945.2324[d]	845.2326
S8	**39**	T 01001000 01000001 01010000 01010000 01011001 01000000 01010000 01001000 01000100	-	1079.0947[e]	1079.0873

[a] The numbers 0 and 1 denote for phosphate linked repetitive units derived from diols **3** and **8**, respectively. T refers to a deoxythymidine monomeric unit. [b] $[M\text{-}2H]^{2-}$. [c] $[M\text{-}7H]^{7-}$. [d] $[M\text{-}8H]^{8-}$. [e] $[M\text{-}20H]^{20-}$

Satisfying yields were obtained and analysis by ESI(-)-HRMS detected the desired sequences in their multicharged states. Due to the introduction of capping steps *ix)* during the automated synthesis the detected impurities refer to low congeners while missed steps sequences were not observed. This is a particularly important fact because congeners would later be removed by standard *DMT-ON* purification once the TIPS-protective groups are cleaved during the sequential modification procedure.

In Figure II-9 an example analysis by mass spectrometry is shown for **S8** the longest sequence synthesized for the project. The more phosphodiester moieties are present in a sequence, the more multiple times negatively charged species can be detected. For **S8** charge states between -17 to -29 can be seen in the *m*/*z* 740 - 1350 range. At lower *m*/*z* values, only very low congeners from missed coupling steps during the first eight cycles were detected (not shown). Congeners from other missed couplings were under the detection limit. As there is no plausible explanation for why phosphoramidite couplings should perform with different quality depending on the numbers of cycles that were already run, it is assumed that all possible congeners are formed with the same probability and all reaction products are quantitatively detected as a function of their abundance and ionizability. That gives rise to the speculation that very low congeners are usually *overrepresented* when compared to a desired

long sequence in a mass spectrum. An analysis tool better suited to the class of compound could have been a high-pressure liquid chromatography (HPLC) with a UV detector. The UV trace over volume of eluent at the absorbance wavelength of the thymidine nucleobase could have given a more quantitatively comparable result with respect to the ratio of lower congeners to the desired sequence. Such HPLC system had not been available when the project was conducted.

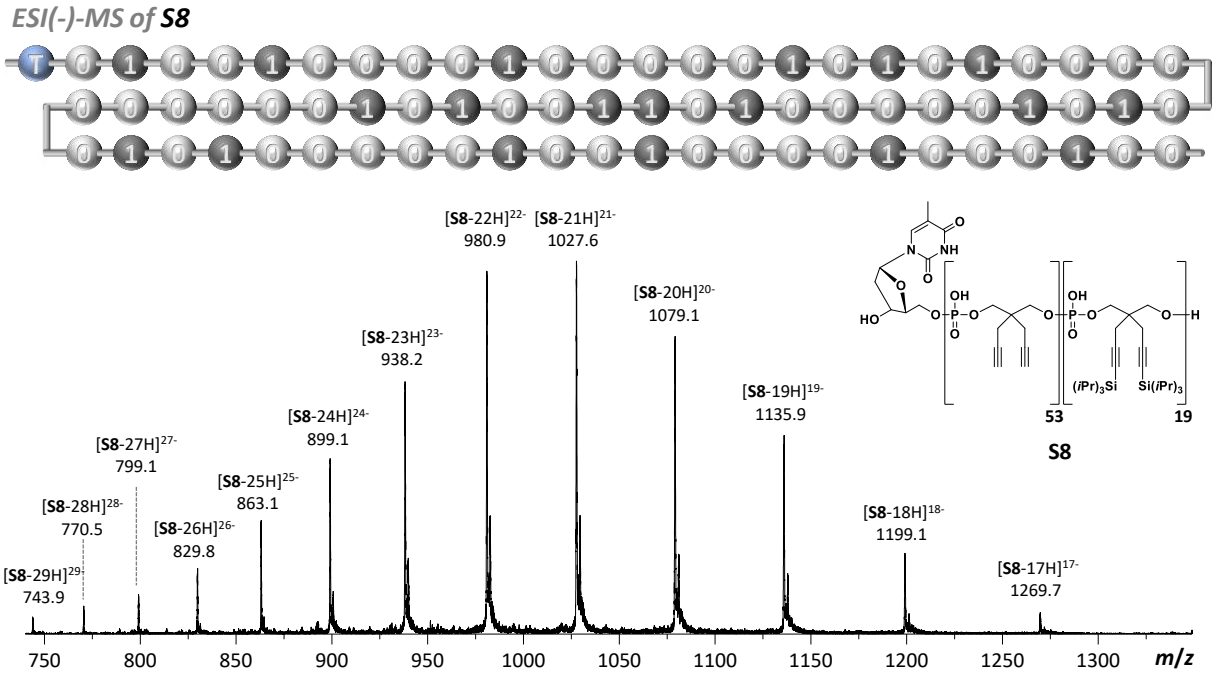


*Figure II-9: ESI(-)-MS of **S8** shown in the m/z 740-1350 range.*

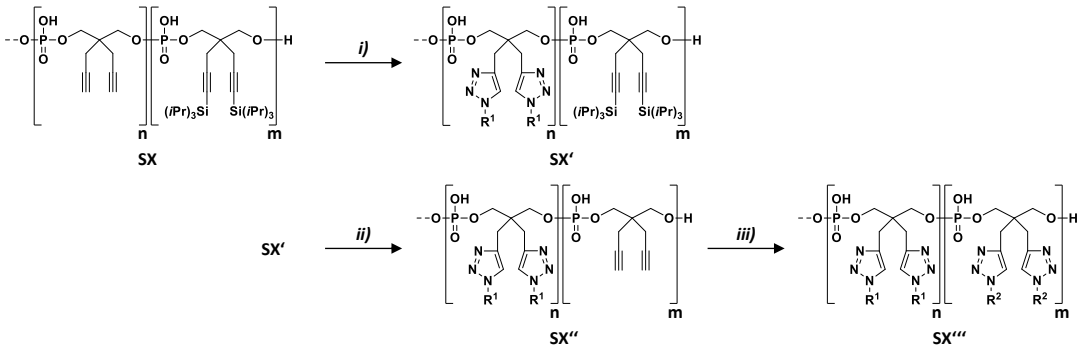
Regarding the side reaction that was observed for sequences **S1**-**S3** that were synthesized on polystyrene supports, there was no indication it could have taken place for sequences **S5**-**S8** that were synthesized on porous glass supports. It is another indicator that underpins the hypothesis of iodine traces that could get stuck in a polystyrene resin and then fuel an iodination during the resin cleavage conditions. CPG is a less attractive matrix for iodine traces that were apparently removable during conventional washing steps. During phosphodiester synthesis on a CPG support the automated protocol had thus not been combined with a sodium thiosulfate wash.

2.4 Optimization of Sequential Post-polymerization Modification Using CuAAC

With a successful synthesis of monomers, solid supports, and precursor sequences obtained both manually and by an automated approach, the tools of the desired platform were successfully developed. In the last stage of the project the sequential platform modification had to be investigated. A method was developed for sequential modification of small oligomers on a relatively large scale. Another method was found for the sequential modification of longer sequences on a relatively small scale.

2.4.1 Sequential Modification of Oligomers

The design of the project requires three quantitative reactions, namely a first CuAAC *i)* with terminal alkyne moieties of a precursor sequence **SX**, a subsequent liberation of retained TIPS-protected alkynes in deprotection step *ii)* yields terminal alkyne moieties and a second CuAAC *iii)* with the latter is used to complete the sequential modification of the precursor **SX** (see Scheme II-7). Different reaction conditions were screened for a first modification step *i)* in solution.



Scheme II-7: Sequential modification of a precursor sequence. First modification i): azide N_3-R^1, CuBr, TBTA, sodium ascorbate, DMSO, tBuOH; TIPS-deprotection ii): TBAF, THF; second modification iii): azide N_3-R^2, CuBr, TBTA, sodium ascorbate, DMSO, tBuOH.

Conditions reported in the context of DNA modification[326] were found to be as well suitable for the here investigated system. There are two main differences between the present and the reported DNA-based system from which the reaction conditions were adapted. For oligonucleotide labeling only one alkyne was modified per macromolecule. However, CuAAC was robust enough to ensure quantitative conversion of multiple and densely arranged functional groups. The second difference was more challenging, oligonucleotides or DNA are insoluble in ethanol and can thus be precipitated and purified from reagents after a conversion by dropping the reaction mixture into the nonsolvent. This was not possible for the oligo- and poly(phosphodiester)s that were investigated here due to the fact that these sequences were characterized by an exceptional solubility in multiple common solvents like water, methanol, acetonitrile, DCM, DMSO, THF, DMF, etc. including also ethanol. This is why a reliable purification method had to be found. Reverse phase chromatography would be an obvious choice. However, a globally applicable method was aimed at that would be independent from the polarity of inserted side chains. A parameter that would always distinguish the modified sequence from all reagents is the size or molecular weight. Dialysis was therefore considered an optimal choice for sequence purification. Indeed, using regenerated cellulose dialysis membranes with a molecular weight cut-off (MWCO) of 1000 Da versus methanol was found to be a globally applicable purification method.

The sequential modification methodology was investigated in-depth following the conditions described above for precursor sequence **S4**. Moderate yields were achieved for each step due to unavoidable losses during each purification by dialysis. As displayed in Figure II-10, the terminal alkynes of precursor sequence **S4** were first reacted with model azide 1-azido-2-(2-(2-methoxyethoxy)ethoxy)-ethane **14**. The copper catalyst complex was formed in situ by adding Cu(I) bromide as a copper source and two equivalents (with respect to the metal) of copper(I)-stabilizing TBTA polydentate ligand to the reaction mixture. Additionally added sodium ascorbate preserved an environment that would reduce the eventually formed copper(II) species back to their catalytically active +I-oxidation state.

The reaction success can be conveniently determined comparing the ^{1}H NMR spectra of precursor sequence **S4** with the modified sequence **S4'**: upon modification the signal of terminal alkyne protons (e) of 0-units at 2.3 ppm completely disappeared and in return a new signal for the triazole proton (i) appears at 8.1 ppm and methylene protons adjacent to the triazole rings (f, h) were observed at 4.5 ppm and 2.8 ppm, respectively. Detection of methylenoxy protons (g) (3.75–3.4 ppm) clearly indicated the introduction of the side chain at all 0-units. Shown in Figure II-10, complete disappearance of the peaks for the triisopropylsilyl protons (d) at 1 ppm and the appearance of new signals for alkyne protons (e) at 3.4 ppm confirms the subsequent quantitative removal of alkyne protective groups on the 1-units by using fluoride ions. Additionally to the here-discussed ^{1}H NMR spectroscopic data of **S4**, **S4'**, and **S4''**, mass spectra showing the detected desired species are given in the Annex in Figure V-28, Figure V-29, and Figure V-30 on page 202. The final CuAAC modification was performed under the

same conditions as the first. Yet, the second model azide *N*-(2-(2-(2-(2-azidoethoxy)ethoxy)-ethoxy)ethyl)-2,2,2-trifluoroacetamide **17** was used resulting in the detection of a ^{19}F NMR signal (see Figure II-10B). Moreover, the quantitative disappearance of the terminal alkyne protons (e) and a significant increase in the intensity of the broad peak of the methylenoxy protons (g) was registered in ^{1}H NMR. And not least is the formation of fully binary modified sequence **S4'''** evidenced because in ESI(-)HRMS the desired species were detected with all 0- and 1-units quantitively modified in steps *i)* and *iii)*, respectively (see Figure II-10C).

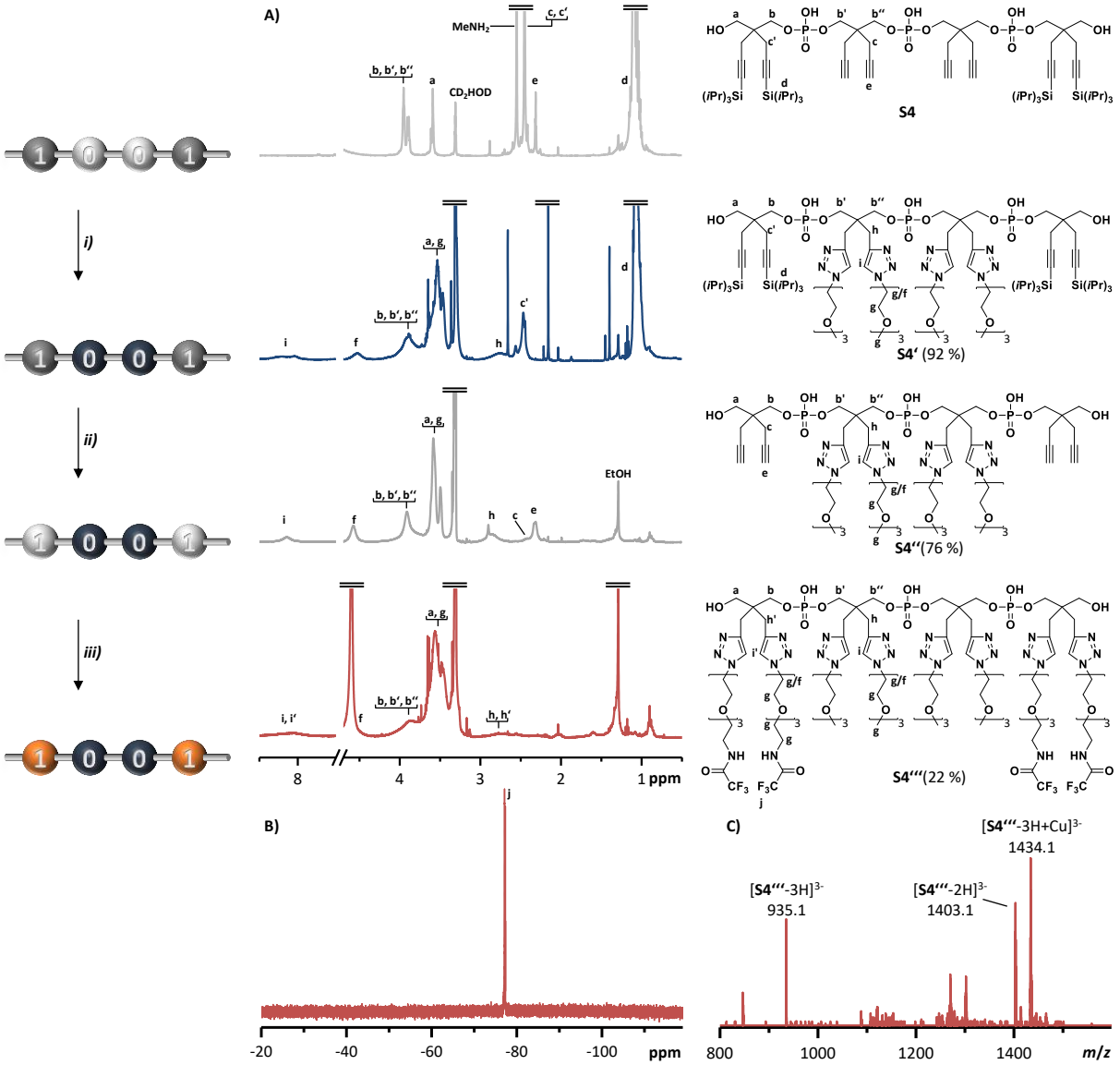
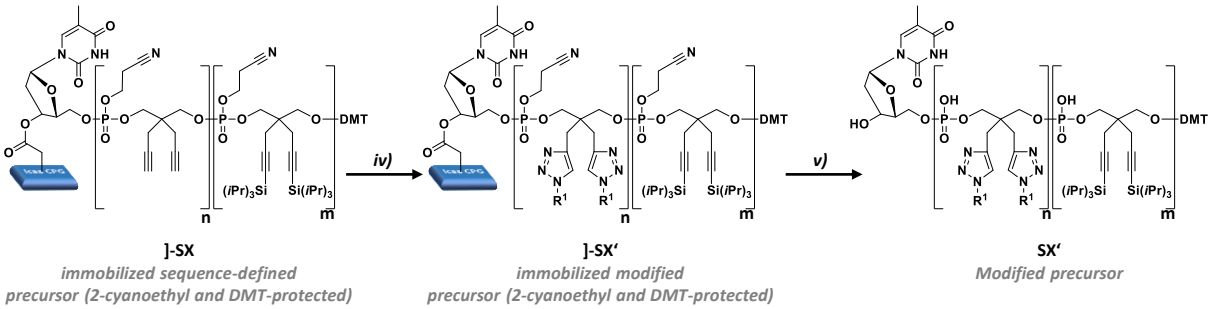


*Figure II-10: Characterization of binary encoded tetrameric phosphodiesters **S4** at different stages of the sequential modification protocol: after solid-phase synthesis **S4** (light gray), after first CuAAC modification of the 0-units **S4'** (blue), after subsequent TIPS-deprotection at the 1-units **S4''** (dark gray), and after the final CuAAC modification of the 1-units **S4'''** (red). A) ^{1}H NMR spectra for **S4**, **S4'**, **S4''**, **S4'''** recorded in d_4-MeOH. The displayed regions are 8.5–7.5 ppm and 4.7–0.5 ppm. The omitted region (7.5–4.7 ppm) contains the water peak of the deuterated solvent. B) ^{19}F NMR spectrum of fully modified sequence **S4'''**. C) ESI(-)-HRMS of fully modified sequence **S4'''**.*

2.4.2 Sequential Modification of Polymers

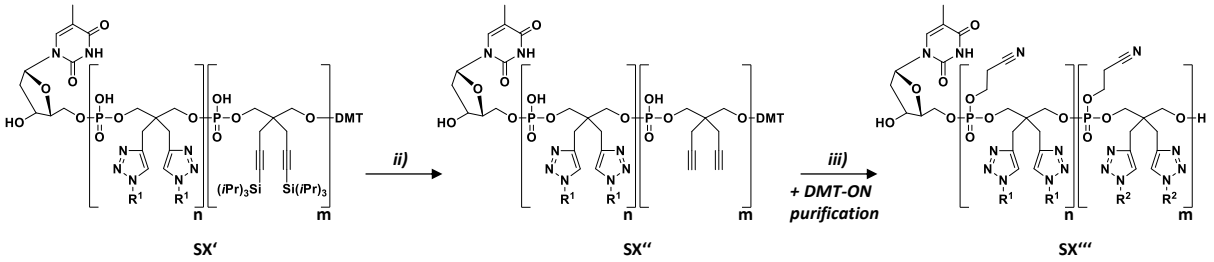
After having optimized the sequential binary modification of oligomeric model compounds, the same protocol was applied to longer sequences like **S5**–**S8** that had been synthesized on a robot on a 1 µmol scale. Yet, the purification by dialysis was at least with the equipment present in the laboratories at that time not successful. The small amount of matter that resulted from the small scale synthesis was

entirely lost during dialysis. That is why another strategy had to be developed for a small scale binary modification. Here, the first CuAAC modification *iv)* was performed on the support material by passing the reaction mixture between two syringes that were connected by a synthesis column. The chosen azide N_3-R^1 together with the copper catalyst was hence passing repeatedly through the CPG powder that was trapped between two frits in the latter column. As it is shown in Figure II-8, the desired precursor sequence **SX** was immobilized on this CPG support together with its lower congeners. During the first modification of immobilized **SX** *iv)* it has to be taken into account that also those congeners are modified even though they are for clarity not shown for the visualization of the process in Scheme II-8. To be precise, in addition to the practical details of a modification on a support, the modification is chemically not fully analogous to the before reported modification of smaller sequences in solution in the sense that the phosphate groups and the terminal hydroxy function of the polymer are protected. Phosphate deprotection is performed together with the subsequent cleavage from the resin *v)* when the modified precursor **SX'** is recovered together with modified lower congeners. As the DMT-group is retained, a DMT-ON purification from congeners is still possible and until then, strictly speaking intermediate crude products are handled.



Scheme II-8: First modification of sequence-defined precursor on a support. CuAAC modification on CPG support iv): azide N_3-R^1, CuBr, TBTA, sodium ascorbate, DMSO, tBuOH; cleavage from support v): methylamine, ammonia.

The final steps of the sequential modification of precursor sequences **SX'** are shown in Scheme II-9 below and were performed analogously to the before reported short sequences' modifications in solution. The important distinction is the choice of purification method. After TIPS-deprotection *ii)* reverse phase chromatography was used to purify **SX''** and its lower congeners from excess tetrabutylammonium ions. After the second CuAAC modification with an azide N_3-R^2 so-formed **SX'''** could be separated from the congeners by performing a *DMT-ON* purification.



Scheme II-9: Final steps for the sequential modification of a precursor sequence. TIPS-deprotection ii): TBAF, THF; second modification iii): 1) azide N_3-R^2, CuBr, TBTA, sodium ascorbate, DMSO, tBuOH; 2) DMT-ON purification including TCA treatment.

Relying on the precursor sequence T 01000001 01000001 01000011 **S7** (see Table II-2), the sequential binary modification methodology was performed using azide **14** in *iv)* and azide **17** during the final CuAAC *iii)*, respectively. While intermediates **S7'** and **S7''** were analyzed by ^{1}H NMR and could be detected by ESI(-)-HRMS (see Figure V-2, pp. 186f.), the ionizability of fully modified **S7'''** was probably too low to make it analyzable by MS. However, its ^{1}H NMR spectrum that is depicted below in Figure II-11 shows analogously to that of **S4'''** in Figure II-10.A retention of signals from monomers 0 and

complete disappearance of signals at 2.4 ppm and 2.3 ppm characteristic for methylene protons adjacent to an alkyne and the terminal alkyne proton, respectively. A comprehensive discussion of analytical data for different stages of the sequential modification protocol is presented in Annex: 1.2 Sequential Modification of Poly(phosphodiester)s (pp. 186ff.). Despite the efforts, a considerable amount of impurities was detected and it must be critically commented that the DNA purification kits used here for reverse phase purification were not an appropriate tool for this difficult isolation task. However, it can legitimately be anticipated that the use of a preparative HPLC would have resulted in pure sequences in higher yields and also detection of binary modified polymers by mass spectrometry would have been feasible. Unreacted azide **17** (marked with # in Figure II-11) was actually dominating the spectra.

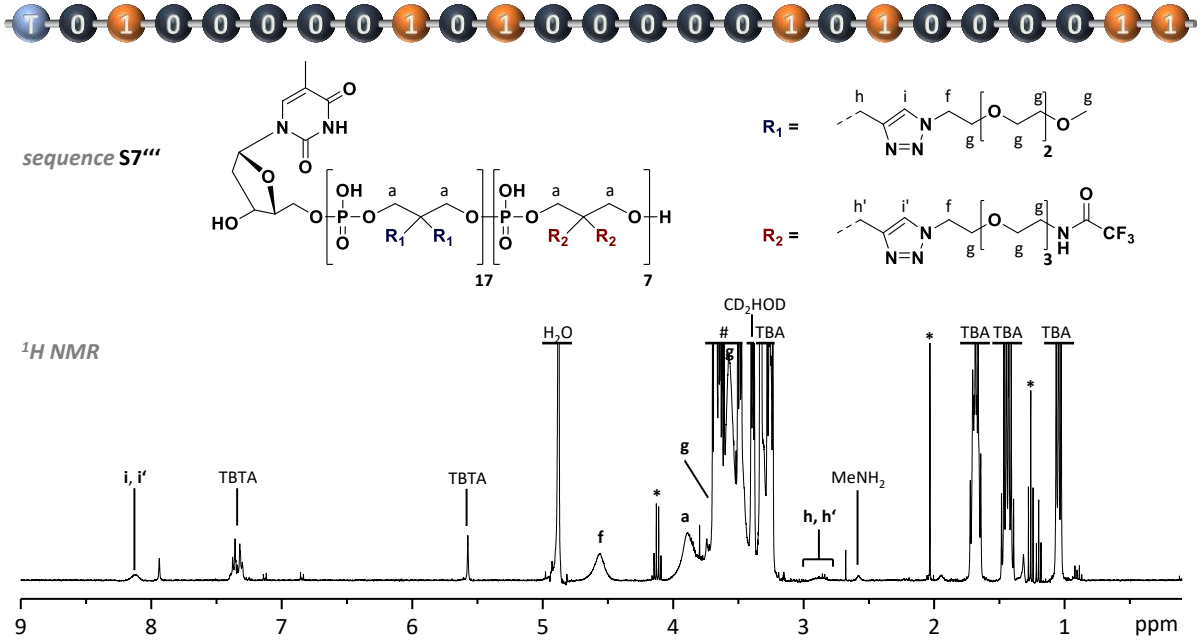


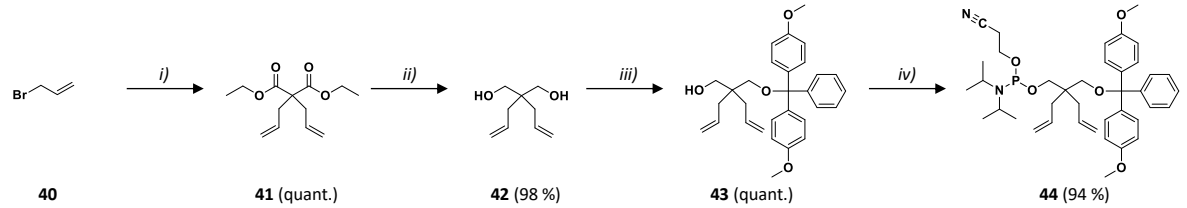
*Figure II-11: ^{1}H NMR spectrum of binary modified sequence **S7'''** that was sisolated as a tetrabutylammonium (TBA) salt, recorded in deuterated methanol. Impurities: TBTA ligand, unreacted azide **17** (#), $MeNH_2$, and ethyl acetate (*).*

2.5 Perspective for the Second Post-polymerization Modification

The development of a platform for a sequential post-polymerization modification of sequence-defined poly(phosphodiester)s was described exhaustively in the above sections. For sake of completeness, this section is devoted to another approach that did not contribute to the already described methodology although it is closely connected. A sequential post-polymerization platform could also be synthesized using the alkyne phosphoramidite monomer **5** and an alkene-containing monomer. This approach had been followed in parallel but was not considered anymore when the desired methodology was successfully developed in the earlier described fashion.

In the beginning of the investigations, the TIPS deprotection step—crucial for the initially presented platform design (see 2.1, p. 65f.)—required tedious optimization work and it prevented the second modification to be performed with an immobilized sequence. Thus, it was additionally followed a strategy adapting the initial platform design to circumvent the deprotection step. In brief, the strategy would have been based on a platform sequence composed of alkyne bearing monomers and alkene bearing monomers. After chemoselective CuAAC modification of the alkyne moieties, the alkene moieties could be modified afterwards (even being still immobilized on the support) using thiol-ene chemistry. For this approach, alkene phosphoramidite monomer **44** was successfully synthesized. Yet, the strategy had never been tested as the initial design was working well after successful optimization.

The design and the synthesis of the new alkene bearing monomer **44** corresponds to the reported monomer syntheses (see Chapter II: 2.2.1, pp. 66ff.) and is depicted in Scheme II-10. From a practical perspective, all compounds were very efficiently synthesized and purifications were—if necessary—easy.



*Scheme II-10: Overview synthesis of alkene phosphoramidite monomer **44**. i) 1. THF, diethylmalonate, NaH, 0 °C; 2. allyl bromide **40**; ii) THF, $LiAlH_4$, -10 °C; iii) THF, pyridine, DMT-Cl; iv) DCM, DIPEA, 2-cyanoethyl*-N,N-*diisopropylchlorophosphoramidite, 0 °C.*

During the conversion of the malonate in step *i)* two allyl groups were quantitatively inserted at the central carbon atom. After aqueous work-up of the reaction mixture, the excess allyl bromide **40** was efficiently washed away and no further purification was needed to obtain malonate **41** quantitatively in excellent purity as controlled by ^{1}H NMR. In analogy to (TIPS) alkyne monomer synthesis an excess of the bromide was used. The same reaction efficiency is reported using a stoichiometric amount of bromide.[340-341] The following reduction step *ii)* to the corresponding diol **42** did also occur near-quantitively as it is described in the literature[341] and did not require a further purification after aqueous work-up.[342] The DMT-protected alcohol precursor **43** was obtained quantitatively (after chromatographic purification). As the synthesis of starting material **42** was highly efficient it was not considered especially scarce and it was used with a 50 % excess with respect to DMT chloride in step *iii)*. Consequently, the high yield was expected when put in context of the discussion of step *iii)* during the analogous monomer synthesis (see Chapter II: 2.2.1, pp. 66f.). In step *iv)* the desired phosphoramidite monomer **44** was obtained with a satisfying yield. It is interesting, that the synthetic route delivered all intermediates as well as the target monomer **44** in excellent yields and exceptionally low purification efforts were realized throughout the scheme.

3 Abiotic Poly(phosphodiester)s Interacting with Biological Nanopores

This section should supply the reader with a brief summary of the results that have been obtained regarding the interaction between biological nanopores α-hemolysin, aerolysin, and MspA (Figure II-12) with simple abiotic poly(phosphodiester)s **S9**-**S14** (Figure II-13).

The analyzed sequences were synthesized following the standard protocol for automated poly(phosphodiester) synthesis using commercially available dT-CPG support **39** (see Experimental Section: 1.2.3.1 Automated Poly(phosphodiester) Synthesis, pp. 135f.). That is why every tested sequence begins with at least one thymidine nucleotide with a 3'-alcohol chain terminus. The study is (until today) restricted to a group of sequences that differ by their monomer composition, length of the abiotic part of the sequence and the number of leading dT monomers was varied between one and three. It has been found that these sequences translocate through aerolysin inducing events that can be analyzed. Using α-hemolysin, blocking of the pore occurred, while interaction with MspA is on such a small time scale that does not allow a comprehensive interpretation of the data. However, the interaction with aerolysin reveals that the length of the leading dT block plays a major role to slow

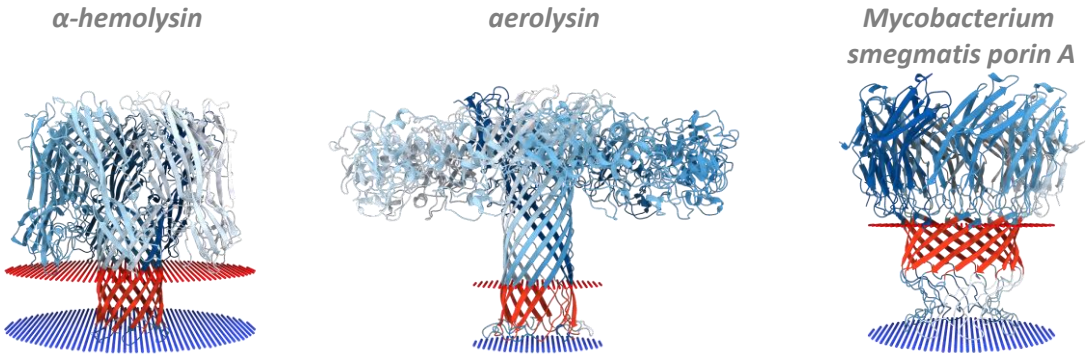


Figure II-12: Transmembrane proteins. Left: α-hemolysin from staphylococcus aureus (structure reported by Gouaux and coworkers[343]); middle: aerolysin from aeromonas hydrophila (structure reported by Zaber and coworkers[344]); right: mutated mycobacterium smegmatis porin A (structure reported by Schulz and coworkers[345]).

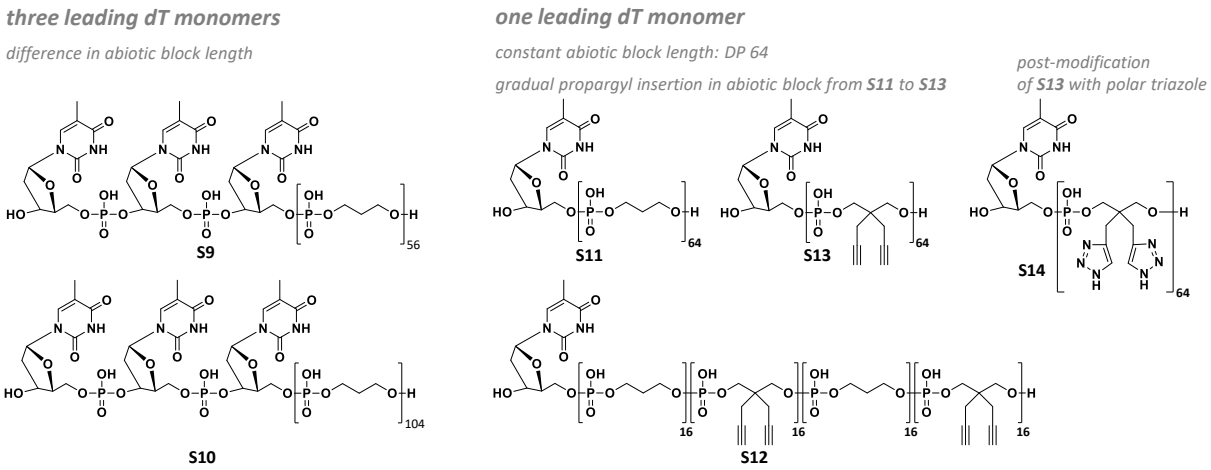


*Figure II-13: Abiotic poly(phosphodiester) sequences **S9**-**S14** tested for their interaction with biological nanopores.*

down translocation, while the length of the abiotic block has a minor influence on that feature. Moreover does the gradual insertion of propargyl moieties to the scaffold and finally the modification to the corresponding triazole additionally lead to stronger interaction of the poly(phosphodiester) when passing through the pore. Already the variation of a few parameters delivered interesting insights and can inspire future attempts for nanopore sequencing of abiotic poly(phosphodiester)s. More details of this interesting study are presented in the Annex: 2 Nanopore Sequencing (pp. 188ff.).

The experiments demonstrate the necessity of polymer side chain modification and how it can help understanding the relation between the nature of side chain and resulting nanopore interaction.

4 Conclusion & Outlook

A binary modification protocol for digital poly(phosphodiester)s was developed. By sequence-defined copolymerization of two analogue phosphoramidite monomers bearing a terminal alkyne group and a TIPS-protected alkyne group, respectively, sequence-defined precursors **S1**-**S8** were accessible. Oligomeric precursors were obtained by manual synthesis on a polystyrene support and a robotic approach on a commercial CPG support gave access to longer sequences. Manual synthesis was thereby performed on a larger scale. This is why the binary modification protocol consisting of a first CuAAC, a TIPS-deprotection with TBAF and a final CuAAC was separately optimized for oligomers and the longer sequences, respectively. Oligomers were for example modified in solution and purified by dialysis while the robot-made sequences were first modified being CPG-bound, then cleaved from the

support and further treated in solution. Final isolation was achieved by DMT-ON chromatographic purification.

It was further shown for the first time that abiotic poly(phosphodiester)s interact with nanopores, namely with the transmembrane protein aerolysin. Two different events were thereby detected that are interpreted as a pore-blocking and the desired translocation, respectively. It was not possible to cleanly sequence the investigated polymers. Yet, trends were observed. Polymer-nanopore interaction could be influenced. The addition of dT-monomers, for example, lengthens the time the open-pore current is interrupted. Also substitution of propyl monomeric units by its more bulky 2,2-dipropargyl analogues in the polymer's backbone provokes this effect. Post-modification converting the alkyne moieties to polar 1*H*-triazoles increases this effect drastically.

The latter example shows that new insights can be generated by post-polymerization modification and the study of so-attained analogous polymers that differ in their side chain information. An important multidisciplinary collaboration delivered already some interesting details and with the protocol for a sequential binary modification developed, further side chain variation will lead to a more comprehensive study It will further be the goal to introduce preferably bulky and polar side chains to slow down the polymer translocation and, importantly, determine a comonomer combination that allows to reveal primary structure order with a high precision.

Chapter III:

Photocontrolled Phosphoramidite Chemistry For the Synthesis of Digital Oligo(phosphodiester)s

1 Introduction

In the following chapter a photocontrolled synthetic access to digital oligo(phosphodiester)s is reported. Relying on efficient phosphoramidite chemistry, the protective group strategy is to be adapted compared to the standard protocol. The concept's implementation to the toolbox of precision polymer chemistry is the aim of the here-described project. Using light as the trigger for polymer growth is particularly powerful as the stimulus can address the substrates with a high degree of temporal and spatial control. In the context of molecular data storage, it should therefore be possible to miniaturize digital polymer libraries. Analogously to DNA chips, a large library of digital polymers should be arranged with high density on a tiny surface.

In order to improve the relevance of digital polymers as molecular information storage systems, synthesis and sequencing have to be optimized. When large data storage capacities are aimed at, miniaturization of digital polymer libraries is clearly the next step to achieve a competitive information density. The here-described approach, is—to the best of the author's knowledge—the first reported attempt investigating the synthetic basis for high-capacity digital polymer data storage.

1.1 Motivation & Context

Large scale data storage on digital polymers is based on converting an information bitstream into chemical information that is represented by the primary structure of a sequence-defined polymer. The degree of polymerization (polymer length) is limited by the synthesis efficiency. This is why a bitstream is usually divided into information packages that can chemically be managed during translation onto the molecular level.[9] High information density is a key parameter when such libraries are envisioned. It can be tackled optimizing different factors: polymer length is obviously important to decrease the number of information packages to be converted into molecules. To date the longest information-containing abiotic polymer has a DP of 105 and codes binary data equivalent to 13 bytes.[6] Another approach is to increase data density in the repeat unit of a polymer.[314] The highest data density per repeat unit was achieved by assembling a large library of components and reached 3 bytes per unit. The synthetic protocol limited the synthesis to rather short sequences. Such tetramers consequently coded for 12 bytes.[316] For the design of high-density libraries, we could assume to combine such approaches. A hypothetical fusion of the described high-capacity digital macromolecules would reach a data density of 312 bytes·chain^{-1}. Being based on that strategy, a conventionally synthesized polymer library could be stored in a standard SBS-format (SBS: Society for Biomolecular Screening) 1536 microwell plate (1536 positions à 10 µL) giving access to a 480,000 byte digital polymer library with a data density of 3.06 kB·cm^{-3}. Even though the design of the microwell could further be miniaturized, such data densities were already obtained with standard 3.5 inch floppy disks at the end of the 1980s.[2]

As described in subchapters *Phosphoramidite Chemistry for Microarray Production* (pp. 44f.) and *DNA-based Data Storage* (pp. 47f.), microarrays organize large numbers of different sequences with high density on a surface. Going back to a poly(phosphodiester) encrypting 13 bytes of information, a standard microarray containing a photolithographically and *in situ* synthesized polymer library with roughly a million individual entries, 13 MB of data were to be stored. The data density achieved is then 8.1 MB·cm^{-2}. Protecting such digital polymer library with a generous housing, volumetric data densities of >15 MB·cm^{-3} are hypothetically accessible. Yet, the strength of microarray production is the massive in parallel precision synthesis of polymers. When the digital polymer library is isolated from the support, the information density is enormous (depending on the number of copies of each polymer

[2] The 3.5 inch floppy disk has housing dimensions of 90 mm x 94 mm x 3.3 mm. Such HD floppy disk has a storage capacity of 1,440 KiB (= 1,440 x 1024 bit = 1,475 kB). The density is thus 52.85 kB·cm^{-3}.

present) and potentially higher than that of DNA-based systems. However, a convenient sequencing technique is still to be developed when handling mixtures of digital polymers on a sub-picomolar scale.

1.2 Aims

The aim of the here reported study is, thus, to preliminary examine in which way a photocontrolled multistep-growth polymerization can be designed that could later serve as the basis for digital polymer microarray fabrication in a follow-up project. It should further be critically investigated whether a sufficient degree of polymerization and sample homogeneity—factors that are crucial for data storage application—are accessible. Nevertheless, the study should be a first step towards a greater goal. Consequently, the fundamental aspects of a photocontrolled synthesis of digital polymers is to be understood to render the approach valuable to the field of precision polymer chemistry.

2 Results and Discussion

In the following paragraphs the results are reported and discussed that were obtained during the development of a photocontrolled synthetic access of abiotic oligo(phosphodiester)s. After presentation of the chemical strategies and design of the project, monomer synthesis and the role of different support materials during photocontrolled phosphoramidite chemistry are detailed. Important observations during the various attempts to access oligo(phosphodiester)s are described and explained. Finally, the optimized chemical strategy is reported and oligomer yields are critically discussed.

2.1 Conceptual Design

The present investigations are a preliminary study toward the synthesis of digital microarrays by photolithographic means. Photocontrol of chemical reactions is an important tool and in theory, diverse very efficient photoligations are reported[346] that could be employed to obtain sequence-defined polymers from a multi-step growth synthesis. One such approach relying on photochemically induced Diels–Alder reactions was investigated by Barner-Kowollik and coworkers.[154] Reportedly, diene deprotection by a retro-Diels-Alder reaction required rough reaction conditions and long reaction times while overall yields were low. Presuming that such features should be strictly avoided for microarray fabrication, it was thus decided to employ a variation of the mild and efficient phosphoramidite chemistry that had already been adapted for the fabrication of DNA chips and would be a new technique for precision macromolecular chemistry.

The underlying idea is to render the alcohol protection step during the iterative protocol a photocontrolled step as it is depicted in Figure III-1A. As discussed in subchapter *Phosphoramidite Chemistry for Microarray Production* (pp. 44f.), the NPPOC photoreleasable group protects the terminal hydroxy group during the phosphoramidite coupling steps. It was decided to study a simple model system that would be an analogue to the first reported oligo(phosphodiester) from Lutz' laboratory.[10] Consequently, the new phosphoramidite monomers that were needed for the study are composed of a central 1,3-propandiol moiety with different substitution patterns at the 2-position having one hydroxy group NPPOC-protected and the other converted to the imperative phosphoramidite moiety as it is shown in Figure III-1B.

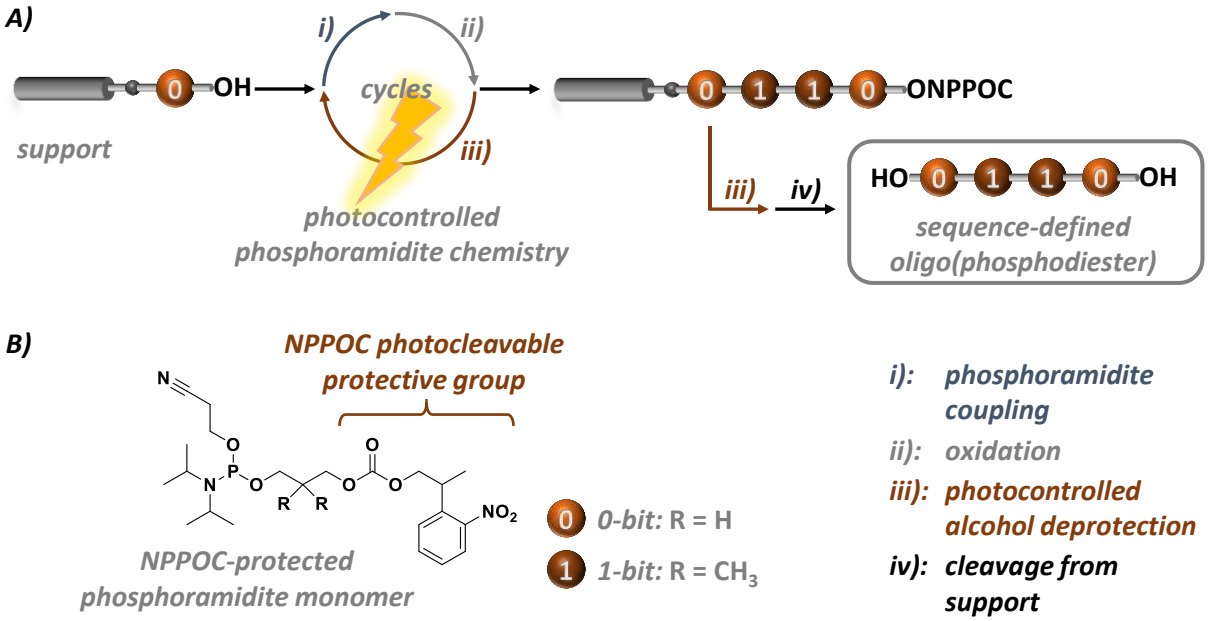
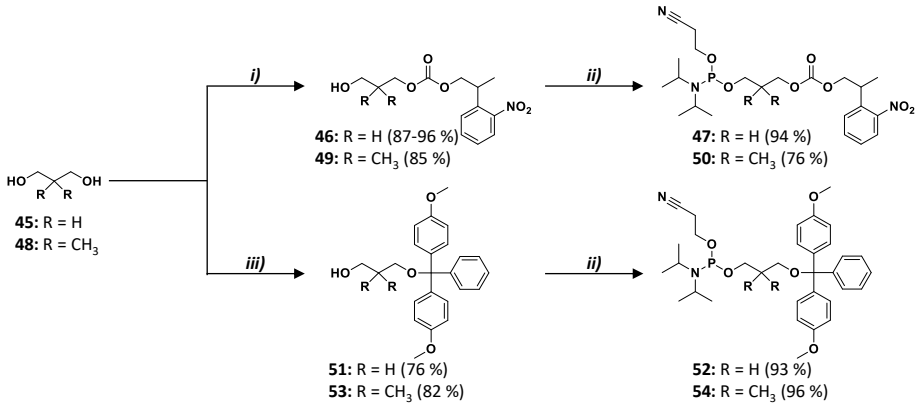


Figure III-1: Project design. A) photocontrolled phosphoramidite chemistry for synthesis of oligo(phosphodiester)s; B) coding phosphoramidite monomers protected with photolabile NPPOC group.

2.2 Monomer Synthesis

The monomers **47** and **50** depicted in Figure III-1B were synthesized by NPPOC monoprotection of corresponding diols **45** and **48** followed by conversion of the remaining hydroxy group to a reactive phosphoramidite. The statistical monoprotection *i)* of diols was efficient. Due to the commercial availability of the diols, an excess of substrate could be used. A 20 mol% excess with respect to the protective group chloride resulted in yields below 90 %. When using a 400 mol% excess of diol **45**, a near-quantitative yield was accomplished. Moreover, advantageous to DMT protection, the corresponding NPPOC chloride is soluble in DCM and was slowly and continuously added via a dropping funnel keeping the concentration of the reactive reagents constantly at a low level to avoid a full diol protection. Phosphoramidite monomers **47** and **50** were isolated with different yields. While full conversion could be detected by TLC, a partial degradation on the silica column during purification could not be further suppressed for monomer **50**. As shown in Scheme III-1, the overall yields were nevertheless satisfying. It was necessary to synthesize the DMT-protected analogues of NPPOC phosphoramidites **47** and **50** in a late stage of the investigation. Already reported in the literature,[10, 347] DMT phosphoramidite monomers **52** and **54** were analogously synthesized from diols **45** and **48** by

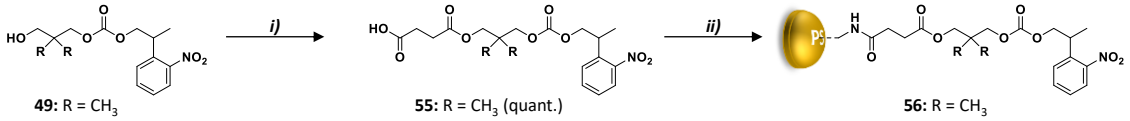


*Scheme III-1: Synthesis of phosphoramidite monomers **49** and **50**. i): DCM, pyr, NPPOC-Cl, rt; ii): DCM, DIPEA, 2cyanoethyl-N,N-diisopropylchlorophosphoramidite, 0 °C; iii) THF, pyr, DMTrCl, rt.*

DMT monoprotection *iii)* and following conversion *ii)* of the remaining hydroxy group to the corresponding phosphoramidite.

2.3 Synthesis of Solid Supports

Analogous to previously reported support syntheses (see Chapter II: 2.2.3 Solid Support Synthesis, pp. 71f.), a solid polystyrene support **56** was obtained by synthesizing a succinic ester **55** from the protected alcohol **49** that was subsequently coupled in an amidation to an aminomethylated polystyrene resin as it is shown in Scheme III-2. The conditions for NPPOC photorelease were to that date object of the investigation. That is why the loading could not be determined by the usual liberation of the terminal protective group followed by calculation of concentration via characteristic UV absorption. Loading determination based on the added weight after amidation resulted in a loading of 0.65 mol·g^{-1}.



*Scheme III-2: Synthesis of solid polystyrene support **56**. i) pyr, succinic anhydride, DMAP, 35 °C; ii) 1. DCM, DMAP, DCC, aminomethylated polystyrene; 2. pyr, acetic anhydride.*

2.4 Oligomer Synthesis Using a Solid Support

Support **56** was employed with the aim of synthesizing from monomer **50** a tetrameric model sequence **S15** that is shown in the upper part of Figure III-2. Iterative phosphoramidite reaction cycles were performed like already described for polystyrene support-assisted oligo(phosphodiester) synthesis in *Manual Oligomer Synthesis* (pp. 72f.). Instead of the acid-induced DMT deprotection, a NPPOC photorelease had to be achieved. For the attempted photodeprotection of NPPOC-terminated sequences, the resin was first swollen in anhydrous DCM and different solvent mixtures were added. The light source were two 150 W high-pressure mercury lamps with a maximum energy at 365 nm. During the irradiation, an orange coloration was observed and after washing the resin, it became increasingly red with extended irradiation times. This was interpreted as the liberation of the *ortho*-nitro compound from the chain terminus. Although the photolability of the NPPOC group during solid-phase peptide synthesis using a polystyrene support has been reported,[348] it was assumed that photodeprotection was not complete. In order to exclude systematic errors during phosphoramidite coupling and oxidation, a sequence was successfully synthesized using the classical DMT-protected monomer **54** and acid-induced DMT protection steps. It was hence indirectly shown that an incomplete NPPOC photorelease prevented oligo(phosphodiester) synthesis. Thus, the new photodeprotection step had to be further optimized. The highest degree of oligomerization using support **56** and monomer **50** was obtained when performing the irradiation of NPPOC-terminated chains for two hours in 50 mM piperidine in DCM. After cleavage from support, the mass of desired sequence **S15** was detected in the noise of the corresponding mass spectrum (negative mode) that is shown in Figure III-2. More prominent were the peaks of sequences with missed steps **S15'** and **S15''** (and their MS fragmentation peaks). Despite the extensive irradiation times with high intensity, photorelease of NPPOC remained marginal. This can be explained by the high light absorption of polystyrene. Growing chains inside the crosslinked support are thus not accessible by the light trigger, stay in the *shadow*, and photodeprotection cannot be driven to completion in this system. This is why a soluble support was further investigated where the growing chain ends are equally distributed in solution and are accessible by light.

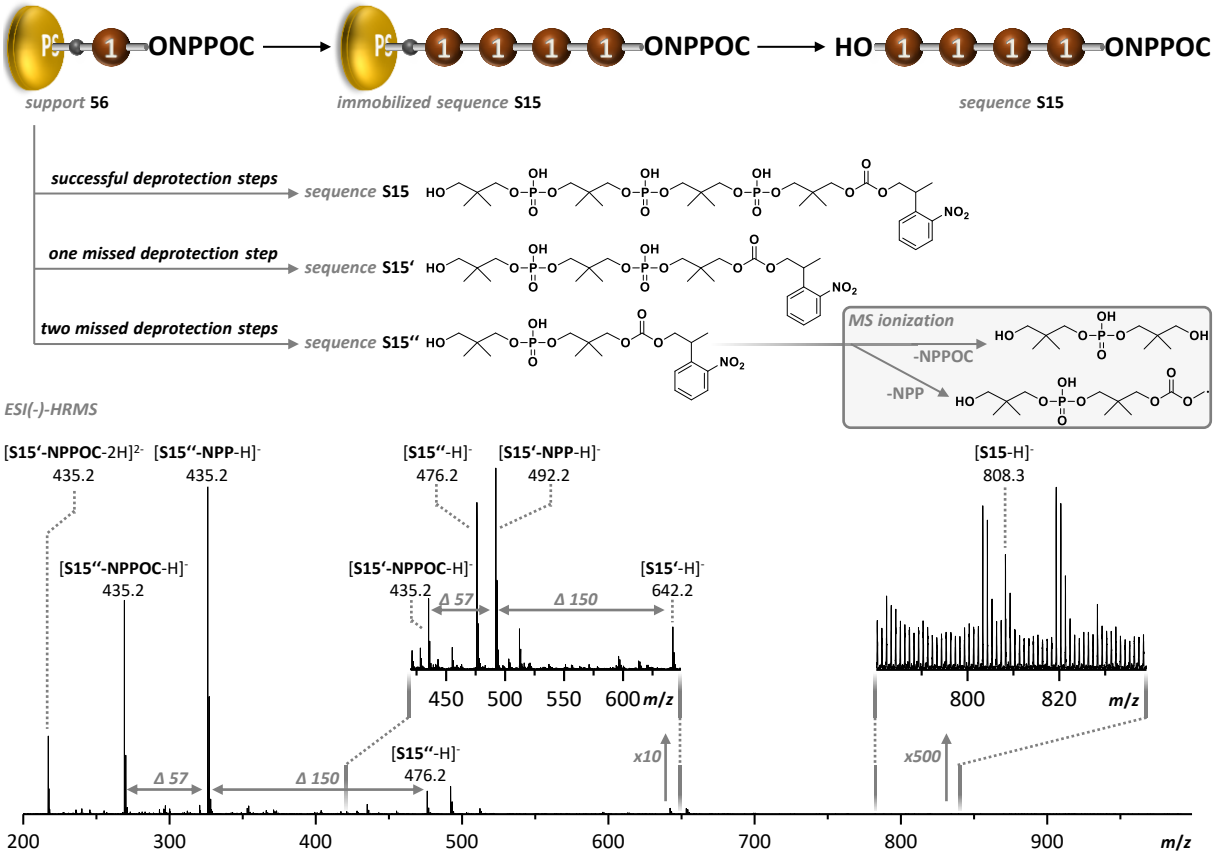


*Figure III-2: Attempted synthesis of **S15** using solid polystyrene support **56**. ESI(-)-MS spectrum shown in the 200 - 1000* m/z range. Area 420 – 650 tenfold and area 780 – 840 500-fold magnified, respectively.

2.5 Oligomer Synthesis Using a Soluble Support

It has recently been shown that ATRP-made linear polystyrene chains with optimized chain length could be used as soluble supports for the synthesis of abiotic oligo(phosphodiester)s using phosphoramidite monomers **52** and **54**. While a solid support material is washed to remove excess reagents from previous steps, soluble supports can be precipitated in a non-solvent. The here-described methodology describes that all reactions were performed in DCM and precipitation occurred in cold methanol. Phosphoramidite coupling and oxidation were reported to be fused to one step without intermediate purification.[349] Conditions for phosphoramidite coupling and oxidation were thus already reported and the novelty when transferring the method to a photocontrolled synthesis with monomers **47** and **50** was the photodeprotection step. During optimization of the latter it was highly advantageous manipulating a soluble support system because more analytical tools were at hand to evaluate the success of employed conditions.

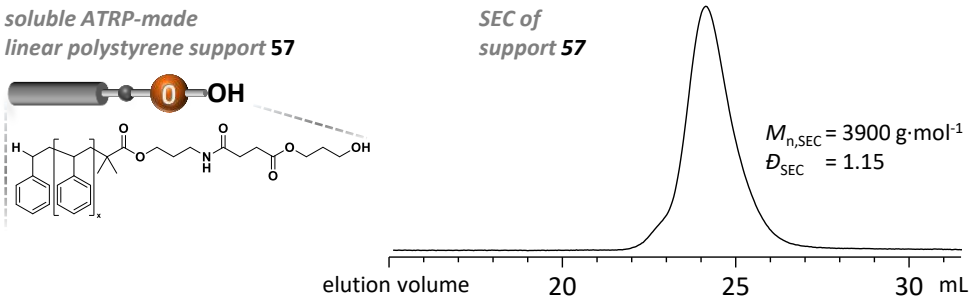


*Figure III-3: SEC characterization of soluble ATRP-made linear polystyrene support **57** (SEC UV trace shown in the range of 15 – 31.5 mL).*

It has been reported that the addition of an amine base enhances the NPPOC photocleavage and ensures a reaction pathway via a β-elimination that results in an unreactive nitrosostyrene moiety rather than reactive nitroso byproducts.[246-247, 350-353] Yet, high base concentration would thereby result in a considerable rate of background dark reaction, too low concentrations are not sufficient for rate enhancement. A 50 mM concentration of piperidine in acetonitrile has been found to considerably increase NPPOC photolability without promoting the background reaction.[247] As the support **57** that is shown in Figure III-3 is not soluble in acetonitrile,[349] it was first investigated whether chloroform could be used during the photocleavage instead. Using two 15 W UV lamps (I = 1100 µW·cm^{-2} at 15 cm distance to filter surface, λ = 365 nm) solutions with different concentrations of NPPOC-terminated sequences in 50 mM piperidine in chloroform were irradiated. Kept under argon, glass vials contained thereby the solutions and were directly placed on the filter of the UV lamp (I > 1100 µW·cm^{-2}). By following the reaction progress by means of ^{1}H NMR and UV-Vis spectroscopy it was found that irradiating NPPOC-terminated support with a concentration of 2 mg·mL^{-1} led to full conversion after roughly 60 min under the above described conditions. However, after isolation of the immobilized hydroxy-terminated sequence by removal of the solvent and precipitation in ice-cold methanol, partial phosphate deprotection was revealed by ^{31}P NMR spectroscopy. As neutralization of the piperidine-containing solution by titration with 3w% TCA in DCM prior to removal of the solvent was yielding fully protected phosphates, it is hypothesized that high piperidine concentrations were temporarily reached during solvent removal that resulted in conditions basic enough to cleave the cyanoethyl protective group (see Table I-1 in Chapter I: 3.2.3 Appropriate Protective Groups, pp. 42ff.). After optimizing support concentration, irradiation time and the work-up conditions, a protocol for photocontrolled phosphoramidite chemistry using a soluble polystyrene support was developed.

The analysis of an exemplary photodeprotection step prior and after irradiation with light of 365 nm is shown in Figure III-4. It illustrates that the photodeprotection step could be driven to completion when optimized conditions were employed. While Figure III-4A shows the cartoon representation for the photodeprotection step, it is shown in Figure III-4B that the disappearance of the aromatic proton signals stemming from the NPPOC protective group in the region of 7.3 - 7.5 ppm is achieved during photodeprotection *i)*. Analysis by SEC—depicted in Figure III-4C—indicates that the dispersity of the hybrid of ATRP-made support and immobilized growing chain does not change. The dispersity originates from ATRP polystyrene synthesis of the support (see Figure III-3) and should not dramatically change when homogenous oligo(phosphotriester)s are synthesized on it. The derived M_p values are calculated from polystyrene references. As the analyzed sequences are not pure polystyrene chains but contain a minor but much more polar block of linker and growing sequence, the so-attained mass values are not reported as quantitative results. They do serve as a rather qualitative reaction control.

Yet, after cleaving model sequence **S16** from the support it became apparent that to a minor degree photocleavage was not quantitative. Missed step sequences **S16*** (missing monomer 0), **S16†** (missing monomer 1) and **S16†*** (missing monomer 0 and 1) were detected in ESI(-)-MS which is shown in Figure III-5 on the left. It was nevertheless shown by tandem MS fragment analysis (Figure III-5, right) that the dominant species was the desired sequence **S16**. Finally, it was found that a systematic error had been committed after the photodeprotection step. The author was initially recovering immobilized sequences from the inner side of the septum that was sealing the glass vial during irradiation. When the septum had come in contact to the irradiation solution and the solvent was subsequently evaporated, some NPPOC-terminated sequences were not in contact to the solution anymore, consequently did not release the protective group and could not add another monomer during the proceeding phosphoramidite coupling.

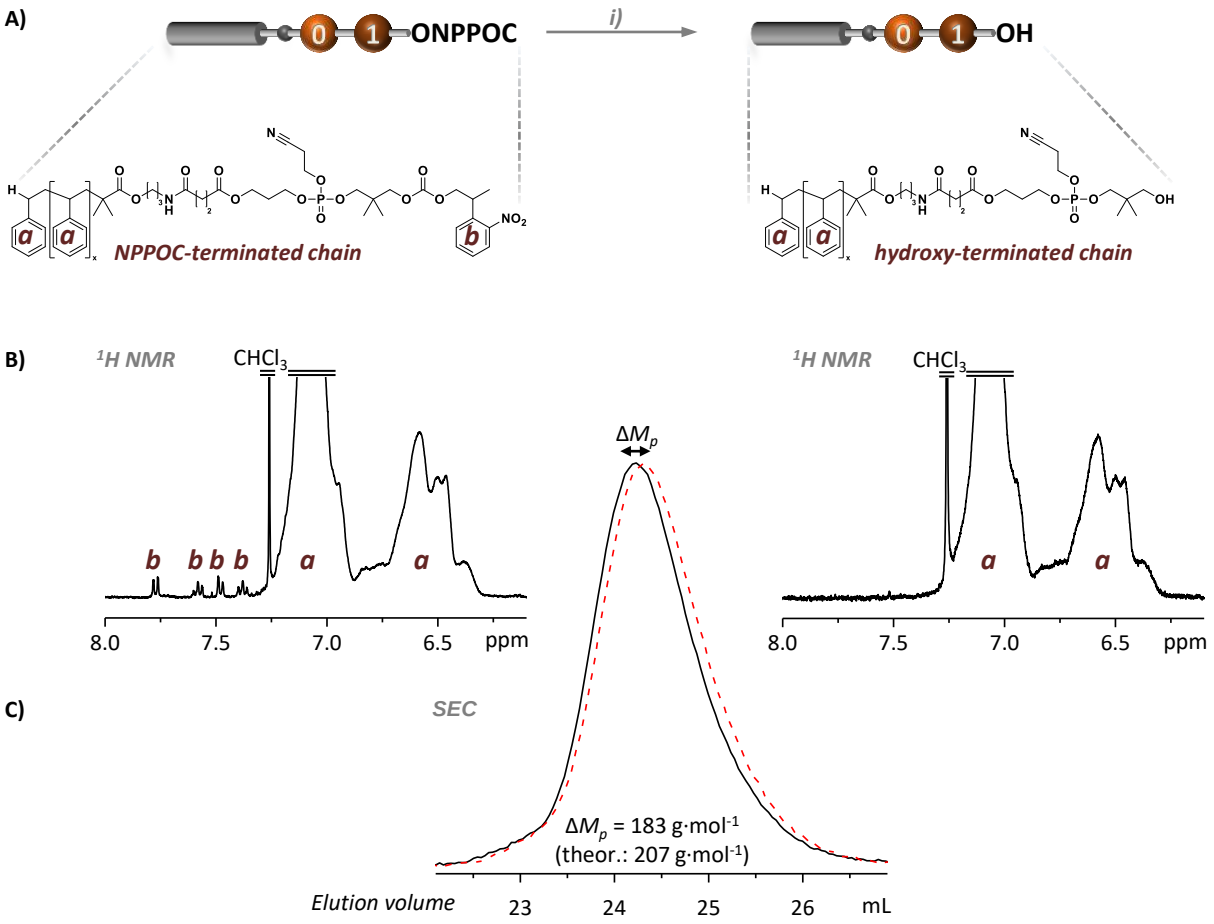


Figure III-4: NPPOC photodeprotection step. A) Cartoon; photodeprotection i): 2 mg·mL^{-1}, 50 mM piperidine in $CHCl_3$, λ = 365 nm, 120 min; B) ^{1}H NMR (zoom of the aromatic region 6.1 – 8.0 ppm) of NPPOC (left)- and hydroxy-terminated chains (right); C) SEC UV traces of NPPOC (black)- and hydroxy-terminated (red, dotted) chains.

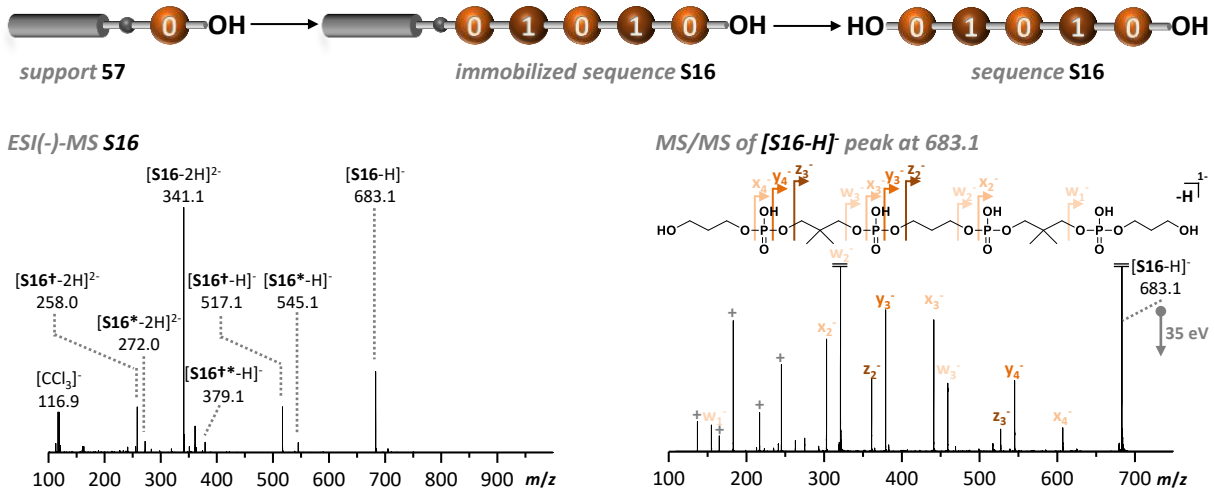


*Figure III-5: MS analysis of attempted **S16** synthesis. Left: **S16*** denotes for a missing 0-unit, **S16†** for a missing 1-unit, and **S16†*** for a missing 0- and 1-unit, respectively. Right: Tandem-MS of [**S16**-H]$^-$ peak. Peaks labeled with + denote for internal fragments, e.g. monomer 0 (m/z 137.0) or 1 (m/z 165.0) were generated in secondary dissociation reactions. Fragmentation peak annotation (m/z): w_1^-: 155.0; x_2^-: 303.0; w_2^-: 321.0; z_2^-: 361.1; y_3^-: 379.1; x_3^-: 441.0; w_3^-: 459.1; z_3^-: 527.1; y_4^-: 545.1; x_4^-: 607.1.*

During the synthesis of the next test sequence **S18** it was first carefully ensured that the septa did not come in contact to the irradiation solution and secondly, the solution was recovered by needle and syringe after irradiation without further washings to guarantee that only the molecules were recovered

that were in solution during the whole irradiation step. Verified by ESI mass spectrometry, this adaptation of the protocol ensures the synthesis of sequence-defined oligomers. The spectrum given in Figure III-6 on the left shows the desired sequence **S18** in its singly (*m/z* = 545.1) and a doubly charged (*m/z* = 272.1) states and compared to the synthesis of **S16** (see Figure III-5), the sample purity is much increased. Consequently, several sequences could be synthesized following the developed synthetic protocol to confirm the reproducibility. The results are given in Table III-1. The reported characterization by HRMS is based on measurements in positive mode. That explains the discrepancy (±2 Da) to the masses of detected species in negative mode as depicted below in Figure III-6.

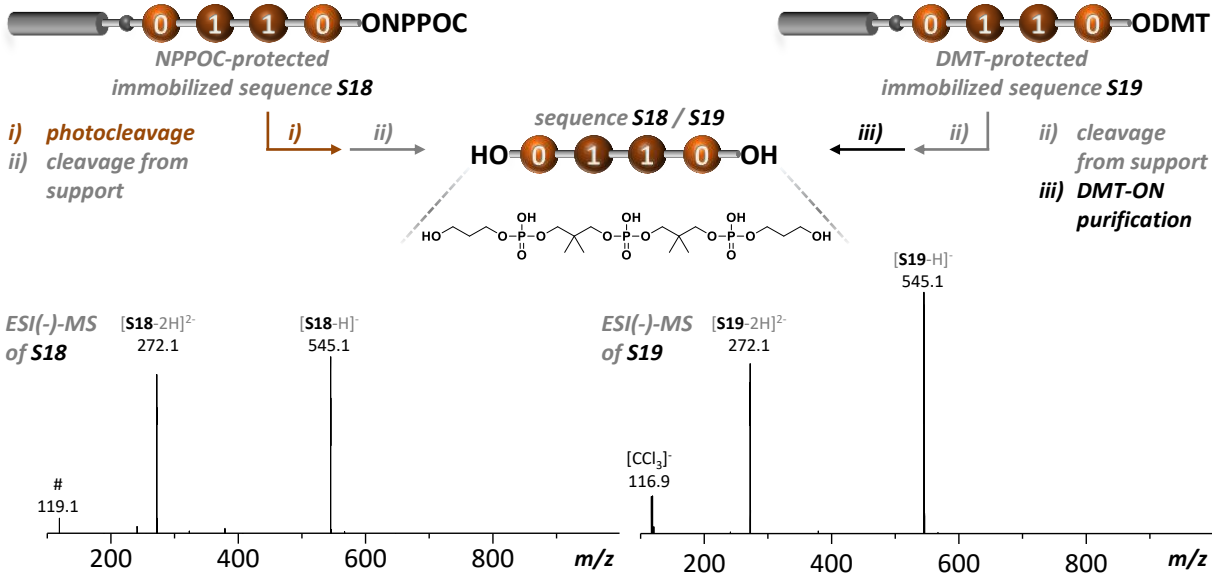


*Figure III-6: Analysis by MS of identical sequences **S18** (left) and **S19** (right) synthesized from NPPOC and DMT precursors, respectively. (Spectra shown in the m/z range of 100-1000; denoted with #, the impurity is structurally not related to sequence **S18**.)*

*Table III-1: Overview and ESI-HRMS characterization of manually synthesized uniform sequences **S17**-**S23**[a].*

	Sequence[a]	Stepwise yield[b]	m/z_{th}	m/z_{exp}
S17	0000	96 %	489.0697[d]	489.0700
S18	0110	90 %	547.1469[d]	547.1478
S19[c]	0110	90 %	547.1469[d]	547.1477
S20[c]	0111	88 %	575.1782[d]	575.1779
S21[c]	0100	85 %	519.1156[d]	519.1150
S22[c]	0101	81 %	547.1469[d]	547.1475
S23	000110	95 %	410.0707[e]	410.0712

[a] The numbers 0 and 1 denote for phosphate linked repetitive units derived from diols **45** an **48**. [b] The stepwise yield is calculated as the mean average of all involved coupling/oxidation and photodeprotection steps. [c] In those syntheses, the last coupling step was performed using DMT-protected monomers **52** and **54**. The cleaved oligomers were purified on a reverse phase column [d] $[M+H]^+$. [e] $[M-2H]^{2-}$.

Sequences **S19**-**S22** (see Table III-1) were synthesized using an additional modification of the protocol that is shown on the right side of Figure III-6. As described before, the photocontrolled NPPOC release step during phosphoramidite chemistry was equally employed here prior to each phosphoramidite coupling. As a special feature, the last monomer that was added to the growing chain was a DMT-protected phosphoramidite monomer **52** or **54**. The desired sequence **S19** was thus first obtained as a DMT-protected immobilized precursor that is shown on the top of Figure III-6. While the NPPOC analogue was deprotected *i)* on the support and then cleaved *ii)* to obtain the desired sequence **S18**, the DMT-protected **S19** precursor was first cleaved from the support *ii)* retaining the protective group.

The last step *iii)* was a DMT-ON purification by reverse phase chromatography. In the course of the latter, the protective DMT group is cleaved yielding the desired sequence **S19**. When comparing the mass spectra shown in Figure III-6, both end group strategies yield sequence-defined oligo(phosphodiester)s of high purity. Yet, the DMT-ON strategy can potentially prevent contaminations stemming from experimental errors. For example, the unknown impurity **#** is not abundant during analysis of **S19**. Instead, a species at *m/z* = 116.9 is detected which indicates that traces of 2,2,2-trichloroacetic acid are present. This acid is used during the DMT-ON purification (see Experimental Section: 1.2.3.2 Manual DMT-ON Purification, p. 137) to finally cleave the DMT protective group.

Demonstrating a photocontrolled DMT-ON oligo(phosphodiester) synthesis was important for two reasons. First, the precipitation of the support matrix after cleavage of the oligo(phosphodiester) block was tedious because the remaining linker-polystyrene conjugate showed an increased solubility. A DMT-ON purification was here additionally used as a filtration to remove the fraction of polystyrene support that was before not easily removable. Moreover, the possibility to employ DMT-ON purification can have some impact on the samples homogeneity when longer sequences should be accessed and additional capping steps are included. In this case only the desired sequence would contain the DMT-terminus that could be efficiently isolated following this strategy.

When analyzing the synthesis yields that are detailed for each cycle step of the synthetic access to sequences **S17**-**S23** in Table V-1 and Table V-2 in the Experimental Section: 3.3 *Photocontrolled Liquid-phase Oligo(phosphodiester) Synthesis* (pp. 159ff.), it becomes apparent that the yields were depending on the synthesis scale. Using more than 500 mg of polystyrene-*b*-oligo(phosphodiester) conjugate in a synthesis resulted in almost quantitative yields. When more than 100 mg were employed, the majority of syntheses could be performed with yields considerably above 90 %. Below 100 mg of converted conjugate, the yields became, however, unsatisfying and are in part even below 50 %. Analysis by mass spectrometry (see Figure III-6 on page 96) confirms sequence definition of the synthesized oligo(phosphodiester) block. The reduced yield is thus not connected to coupling failure but does solely stem from losses during precipitation. During syntheses on all scales small losses could have occurred when a fraction of the support stayed attached to the used equipment. On a relative scale, this error is the smaller the higher the synthesis scale was chosen. For the synthesis on small scales it is probable that the ratio between good solvent and non-solvent were not perfectly complied with or too much methanol was used. The dominant losses are consequently due to imperfect precipitation on synthesis scales below 100 mg.

Further it must be critically mentioned that the use of a soluble polystyrene support also sets limits to the broad applicability of the developed strategy. During the growth of an oligo(phosphodiester) on a polystyrene-*b*-oligo(phosphodiester) conjugate, the polarity of such system increases gradually. At some point, the precipitation of the conjugate would render very inefficient and even large scale syntheses would show drastically reduced yields. If longer sequences were to be synthesized in a photocontrolled way, a soluble fluorous support could thus be advantageous and instead of precipitation, a chromatographic isolation of the support conjugate would be more beneficial.

Very interestingly, the purity of samples was high and for sequences **S18** and **S19** considerable amounts of missed step sequences were not detected (compare Figure III-5 on page 95 with Figure III-6 on page 96). Usually, phosphoramidite couplings using solid support matrices are reported to not result in quantitative conversions and especially NPPOC photodeprotection using micromirror devices guiding light to flat surfaces is reportedly characterized by non-quantitative yields of approximately 95 %.[14, 241] The use of a soluble support system is apparently accompanied by much more efficient conversion rates. Together with the ease to analyze each immobilized intermediate as it was shown in

Figure III-4 on page 95, it makes the ATRP-made soluble support **57** an interesting system to study in detail new strategies in oligo(phosphodiester) synthesis.

3 Conclusion and Outlook

A new method to synthesize abiotic sequence-defined oligo(phosphodiester)s with a photocontrolled version of phosphoramidite chemistry has been developed. Several digital oligomers were synthesized using the new procedure that relies on a soluble ATRP-made polystyrene support. Sample purity was confirmed by NMR and MS.

The synthesis of a new set of NPPOC-protected phosphoramidite monomers could successfully be achieved in two steps. First attempts polymerizing them assisted by a standard solid polystyrene support were suffering from low conversions during the photorelease of the NPPOC alcohol protective group. When switching to a soluble support system, the new photodeprotection step could be followed by UV-Vis spectrophotometry, ^{1}H NMR, and ^{31}P NMR and optimization of reaction conditions was thus straightforward. When the means for quantitative photorelease of the NPPOC protective group were developed and a subsequent work-up that was free from side reactions was established, the accordingly obtained oligo(phosphodiester)s were homogenous and considerable amounts of missed step sequences were not detected by mass spectrometry. Moreover, it could be demonstrated that the protocol could be simplified by a final DMT-ON purification when standard DMT-protected monomers were used during the last phosphoramidite coupling of an oligomer synthesis.

Yields were very satisfying when working on large scale. Small scale syntheses gave unsatisfying yields because the stepwise purification by precipitation was less efficient under these conditions. Another drawback of the use of a soluble polystyrene support system is the limit it sets to the degree of polymerization. The shorter the grown polar oligo(phosphodiester) chain, the more efficient is the precipitation of the conjugate. Nevertheless, the ATRP-made polystyrene support system has proven to be a valuable tool to study in much detail a new synthetic multistep-growth polymerization strategy which would not have been that straightforward by use of a solid support material.

In summary, the fundamental aspects of a photocontrolled synthesis of digital polymers have been understood and it is probable that the new approach is of added value to the field of precision polymers. Future investigations by Lutz' group will show whether large polymer libraries would be accessible using an automated version of photocontrolled phosphoramidite chemistry. The same phosphoramidite monomers coding for bits 0 and 1 can potentially be used for the fabrication of digital polymer microarrays that would comprise a large data density that has not been achieved with synthetic polymers before.

Chapter IV:

Erasing and Revealing Sequence Information by Photoinduced Side Chain Modification

1 Introduction

In this chapter a new conceptual design of information-containing oligo(phosphodiester)s will be described. A photoinduced side chain modification is used to either erase the sequence-information from a polymer backbone or to reveal a molecularly encoded message with a light stimulus.

In the frame of this work, different nitrobenzylamines, carbamates, and ethers are studied as motifs in oligo(phosphodiester)s for their potential as inert or releasable groups upon irradiation with light. To the best of the author's knowledge, the here-described synthetic approach is the first of its kind. Synthetic digital polymers have never before been shown to contain erasable information or hidden a hidden code that can be revealed by a post-polymerization modification. This study is moreover a good example of how synthetic chemists can benefit from the almost infinite chemical space that is at hand. A relatively easily design can thus be the basis of systems that meet demands that had not been tackled before in such simplicity.

1.1 Motivation & State-of-the-Art

Digital polymers can be used for secret information transmission. As the carrier of information they could potentially serve for steganography technology. For example, information-containing carbamates were blended into different polymer matrices without resulting into apparent changes in material properties.[311-313] For an uninformed recipient, the encoded data would thus remain untraceable. In the same way, an encrypted macromolecule could be hidden behind a postage stamp.[315] Examples using DNA as the carrier of information are numerous.[354-355] Here, bacteria colonies[356] or DNA microdots[291] could serve as unsuspicious carriers of digital information. The secret message can be further protected by cryptography. In the case of molecularly encoded information, the recipient has to know which technical analysis tool should be used for deciphering the raw data and how it can be converted to the intended message.[315, 355] For some application, the controlled destruction of a secret molecularly carried message could be of interest. This has been shown for thermally instable poly(alkoxyamine amide) carriers that degrade to unreadable polydisperse analytes upon heating.[303]

In a very innovative study, Balasubramanian and coworkers reported *an epigenetics-inspired DNA-based data storage system* that was already referred to in Chapter I: 4.1 *DNA-based Data Storage* (pp. 47f.). In brief, data was written on DNA chains using besides the set of four nucleobases additional artificial nucleobase monomers. These alien monomers were identified as conventional DNA monomers by polymerase during PCR amplification. Depending on defined stimuli like oxidations or treatments with base, the recognition changes and the alien monomers lead to another signature in the complementary DNA strand after PCR. In this way, it was possible to code information in several layers in a DNA-based storage system.[15]

When post-polymerization modification is used to change the information output, near-quantitative conversion should be aimed at. Like in the above described system, oxidations or efficiently controllable equilibrium reactions are predestined modifications for this purpose. A very handy strategy is—if for example the recipient of a molecularly encoded message does not want to rely on the availability of chemical reagents—the use of light as the stimulus for a quantitative post-polymerization modification, e.g. a side chain photorelease. Used to erase or expose information by transforming a digital polymer to a form with different read-out features, the treatment with UV-light could also be interpreted as an additional level of data protection in the context of cryptography.

A prominent class of photocleavable groups[16] in polymer chemistry are *ortho*-nitrobenzyl groups.[357] They are used to protect hydroxy functions to enable polymerizations[358] or polar carboxylic groups to obtain photoresponsive micelles.[359] When they are used to protect an amide bond, their release can

for example start a supramolecular polymerization as hydrogen bonding is *switched on*.[360-361] *ortho*-Nitrobenzyl protective groups are as well compatible with phosphoramidite chemistry. That has been shown for the synthesis of oligonucleotides that respond with the creation of abasic sites[362-363] and direct strand breaks[364] upon irradiation with light.

1.2 Aims

Inspired by the state-of-the-art, it is the aim of the investigations to design two new systems for molecular data storage. Conceptually related to the work of Balasubramanian *et al.*, post-polymerization modification of side chains should thereby influence the information read-out by tandem mass spectrometry sequencing.

The first system should allow to erase sequence information and hence the encrypted binary code from a digital oligo(phosphodiester) upon irradiation. The original information-containing analyte would be efficiently decryptable by means of MS/MS. Yet, after the triggered photorelease of all *ortho*-nitrobenzyl side chain motifs, the monomeric units should become identical entities which would result in the complete loss of sequence information. The molecularly encoded information would thus be erased upon irradiation.

The second system should allow to hide molecularly encoded information. The digital oligo(phosphodiester) would be composed of two regioisomers as monomeric units that cannot be distinguished from each other by mass spectrometry analysis. The sequence information is thus not accessible. Yet, upon irradiation the *ortho*-nitrobenzyl motif is removed from all side chains while the *para*-nitrobenzyl motif remains intact. The digital oligo(phosphodiester) would thus be composed of two monomeric units with different weight and the sequence information would consequently be accessible by means of tandem mass spectrometry sequencing.

Thus, molecular data storage devices should be developed whose coded message can be erased or whose information can only be exposed after breaking its protection with a defined stimulus.

2 Results and Discussion

Results that were obtained during the study will be presented in detail and are critically discussed. The subchapter is ordered regarding the linker chemistry between the trident platform and the nitrobenzyl moieties and is thereby following the historical evolution of the project. Importantly, understanding of unexpected results is discussed that led to a multiple redesign of the target oligomers and thus the crucial monomeric structures. After discussion of the synthetic access of phosphoramidite monomers and the synthesis of oligomers, the photocleavage of test substances or oligomers will be presented for each design. Finally, the two systems that showed the target features will be studied in detail.

2.1 Conceptual Design

The declared objective was translated into a molecular design relying on nitrobenzyl groups for photoinduced post-polymerization modification. The approach that was under investigation to erase sequence information is depicted below in Figure IV-1 and the strategy to hide and reveal molecularly encoded binary information is shown in Figure IV-2 on page 106.

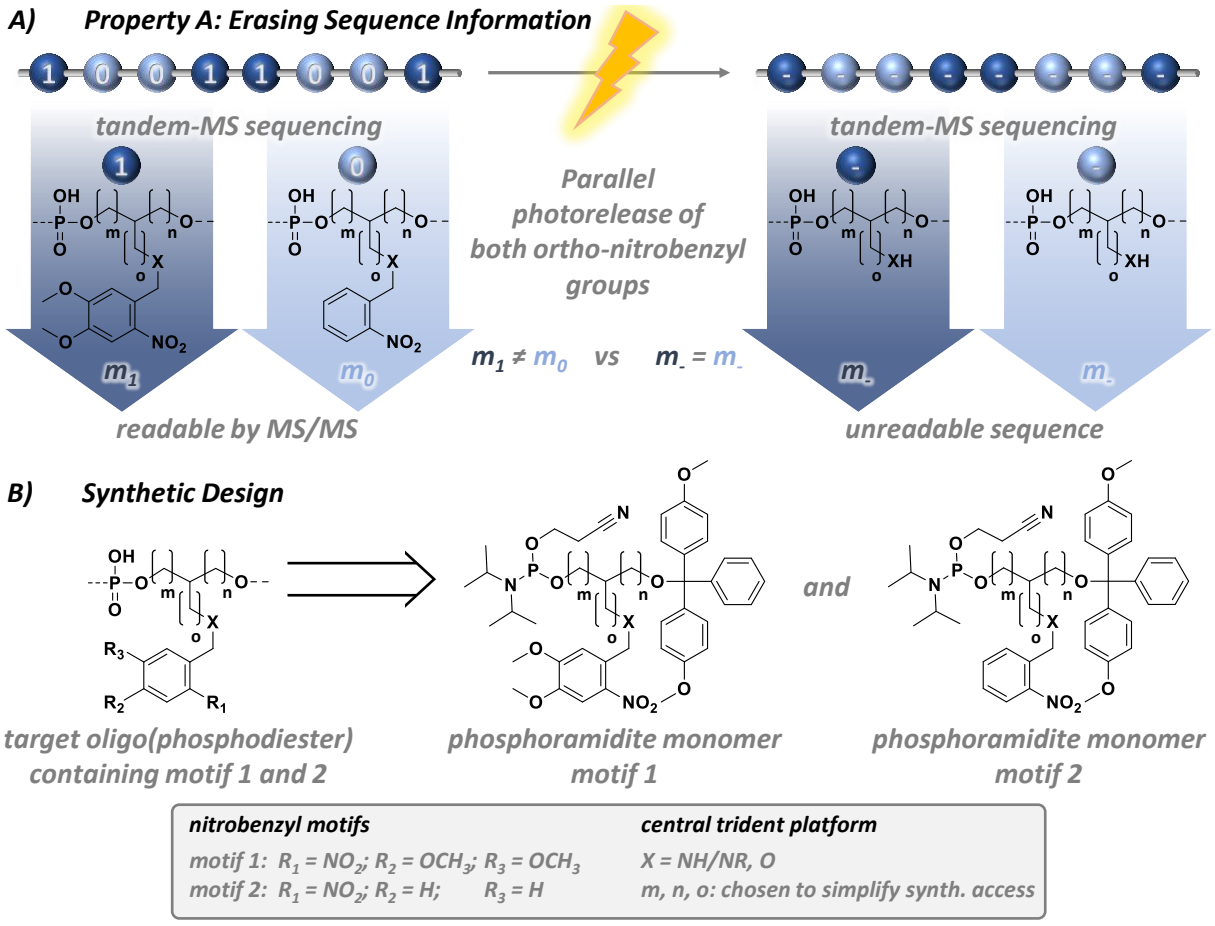


Figure IV-1: Conceptual design to erase molecularly encoded information. A) Property A: erasing information in digital polymers using light; B) Design of oligo(phosphodiester)s and respective phosphoramidite monomers.

Already presented in the state-of-the-art discussion, nitrobenzyl groups can be inserted as ethers (X = O) or they protect amines (X = NH, NR), amides or carbamates (X = NR). To enable the synthesis of an oligo(phosphodiester), corresponding phosphoramidite monomers have to be synthesized. A requirement is the synthesis of a diol intermediate to insert a DMT protective group and the reactive phosphoramidite moiety (see Figure IV-1B and Figure IV-2B). The platform under investigation is thus required to have at least three functional groups, two of them being hydroxy groups to form a classic monomer for phosphoramidite chemistry and the third functional group would be the nitrobenzylated X function. The identity of the trident platform—described by the factors *m*, *n*, and *o* in Figure IV-1B and Figure IV-2B—can thus be adapted to the needs for an efficient synthetic access.

For the implementation of a strategy to enable the removal of sequence information (see Figure IV-1A), a digital oligo(phosphodiester) composed of two comonomers with different *ortho*-nitrobenzyl groups attached to their side chains was envisioned. The comonomers had to have a different mass to allow sequencing by tandem mass spectrometry and should both be photoreleasable. Shown in Figure IV-1B, comonomers exhibiting motif 1 (*ortho*-nitroveratrylbenzyl) and motif 2 (*ortho*-nitrobenzyl), respectively, would meet the criteria. The target digital sequence should be a well-readable information-containing oligo(phosphodiester) and after removal of both *ortho*-nitrobenzyl groups by irradiation with light, identic structures would remain on the oligomer backbone. The encoded information in the target structure is consequently erasable with light.

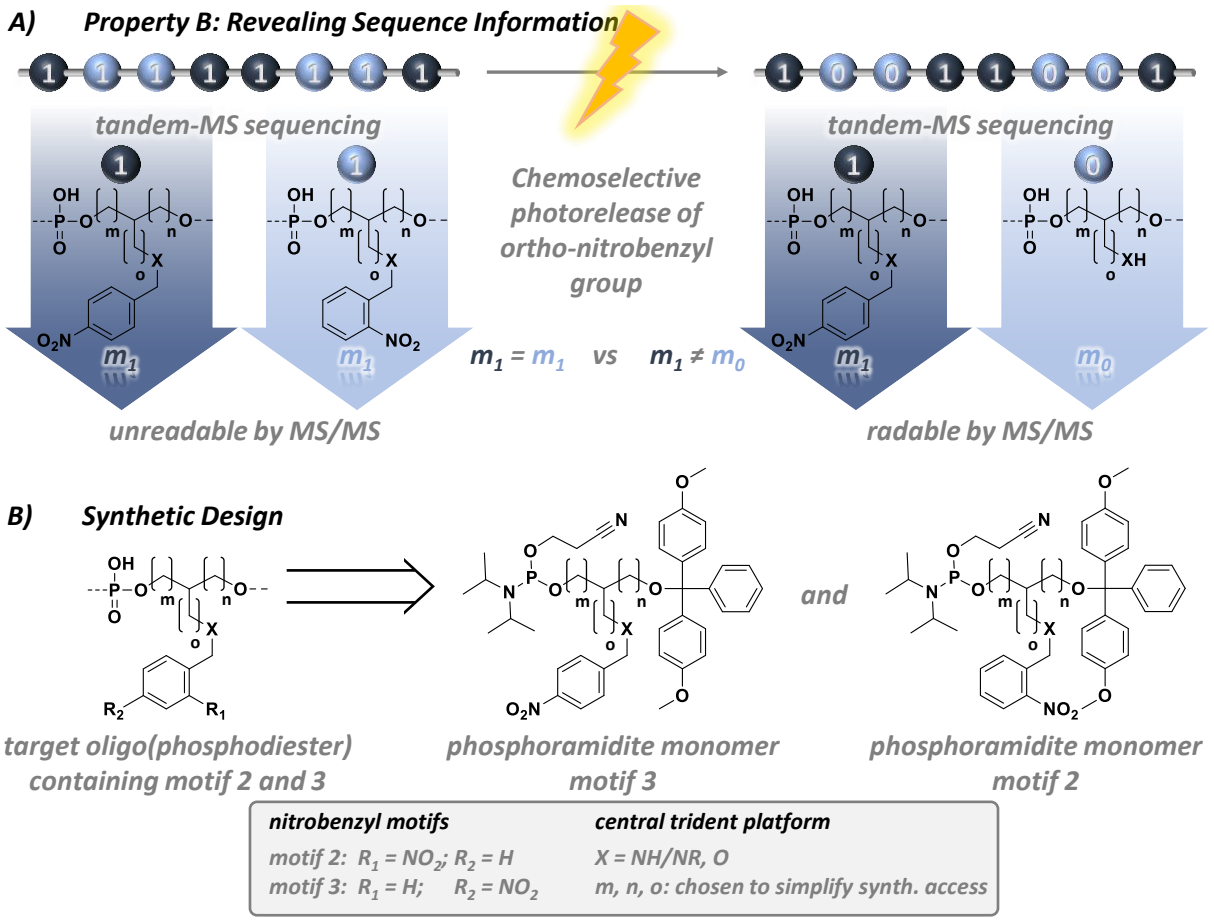


Figure IV-2: Conceptual design to reveal molecularly encoded information. A) Property B: revealing information in digital polymers using light; B) Design of oligo(phosphodiester)s and respective phosphoramidite monomers.

When the concept of revealing hidden sequence information of digital oligo(phosphodiester)s was implemented (see Figure IV-2A), the approach was based on two regioisomers as monomeric units. As depicted in Figure IV-2A+B, the comonomers should have motif 2 (*ortho*-nitrobenzyl) and motif 3 (*para*-nitrobenzyl), respectively, attached to their side chains while the trident platform would be the same. The regioisomers are characterized by the same weight and are thus not distinguishable by mass spectrometry analysis. The encoded binary information of an oligomer—defined by the order of comonomers with motif 2 and 3—can hence not be deduced by MS/MS. Yet, the *ortho*-nitrobenzyl group can be addressed by a light stimulus and is chemoselectively cleavable upon irradiation while the *para*-nitrobenzyl group would remain bound to the side chain. As a result, the so-treated digital oligomer would be composed of comonomers with a different weight. Consequently, the target digital oligomer incorporate a hidden sequence information that can be exposed by irradiation with light.

The implementation of the two concepts *Erasing* and *Revealing* molecularly encoded information was envisioned in a way that in total only three different monomers would be needed. The investigations were thus focused on how motifs 1 to 3 can be introduced to the central trident platform, whether corresponding phosphoramidite monomers were leading to digital oligomers and whether the required photorelease of *ortho*-nitrobenzyl groups was successful.

2.2 Studies on Monomer and Oligomer Design

2.2.1 Study I: Nitrobenzylated Amines

In a first attempt, monomeric units should comprise a nitrobenzylated amine function as it is shown for the phosphoramidite monomer and the oligo(phosphodiester) in Figure IV-3. Referring to the

project design depicted in Figure IV-1B and Figure IV-2B the factors for the platform are chosen as *m* = *n* = 1, *o* = 0 as serinol **58** is a readily available precursor for it and the protected moiety is an amine (X = NH) because nitrobenzyl groups can easily be introduced by reductive amination using the corresponding benzaldehydes **67**, **68**, and **69**.

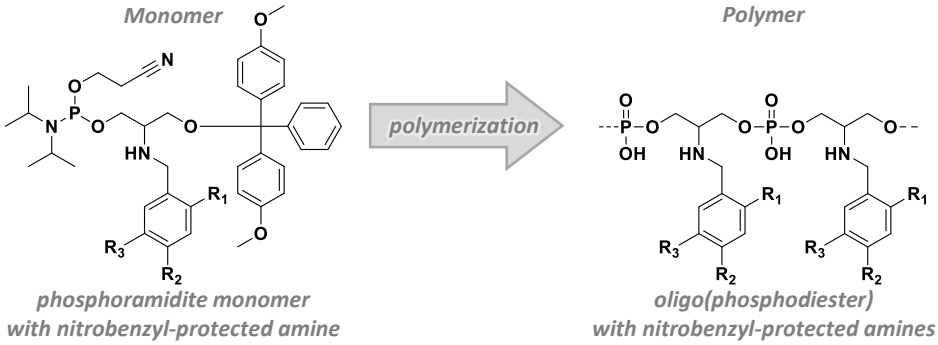


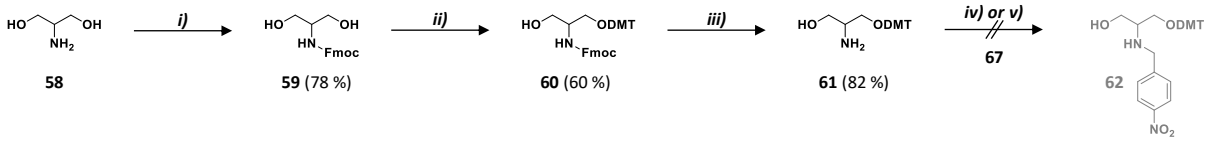
Figure IV-3: Overview phosphoramidite monomer for the synthesis of oligo(phosphodiester)s with nitrobenzyl-protected amine moieties.

It is known from oligonucleotide synthesis that exocyclic (non-aromatic) amines of monomers have to be protected to allow for the sensitive phosphoramidite coupling. Amines would possibly interact inter- or intramolecularly with the tetrazole phosphoramidite complex which prevents efficient coupling to the hydroxy end of a growing oligomer chain. Here, it should be evaluated whether a nitrobenzylamine would be an already sufficiently deactivated secondary amine species. Due to a strong *-I* effect the nucleophilicity should be reduced as when compared to secondary alkyl amines. Yet, when comparing with the usually employed amine protective groups listed in Table I-1 on pages 42ff. a simple benzyl protection is not commonly used for phosphoramidite monomers.

In the following paragraphs the monomer synthesis and attempted oligomer synthesis are reported and the results are evaluated for an optimized redesign of the approach that targets digital oligo(phosphodiester)s with a series nitrobenzylamino moieties.

According to an already published protocol,[365] a DMT-protected serinol precursor **61** was synthesized with satisfying yields as shown in Scheme IV-1. It was the underlying idea to insert the DMT protective group required for phosphoramidite chemistry already early in the scheme to diversify the synthesis by insertion of different nitrobenzyl moieties strategically to a later point of the multistep phosphoramidite monomer synthesis.

To be able to selectively protect a hydroxy function with a DMT group, the primary amine of serinol **58** was temporarily Fmoc-protected *i).* Then, a DMT monoprotection *ii)* could be realized. A simple base treatment *ii)* of intermediate **60** to cleave the 9-fluorenylmethyl carbamate yielded the desired precursor amine **61**.

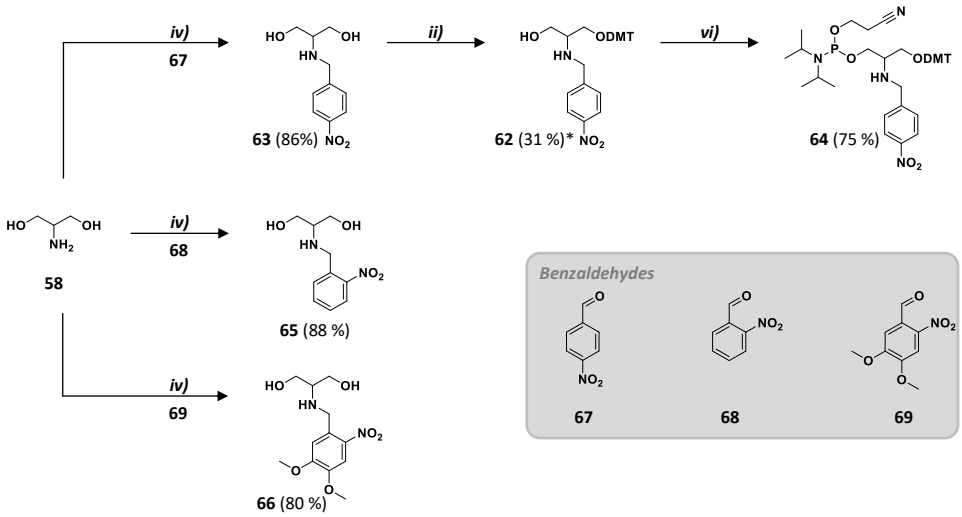


*Scheme IV-1: Synthesis of a DMT-protected precursor **61** for a strategically late chemical diversification. Fmoc protection i): Na_2CO_3 (aq., 10w %), 1,4-dioxane, Fmoc-Cl, RT; DMT protection ii): pyr, THF, DMTrCl, RT; Fmoc deprotection iii): 20 v% piperidine in DCM, RT; reductive amination iv): 1) MeOH, Na_2SO_4, RT, 2) $NaBH_4$, RT; reductive amination v): 1) toluene, reflux in Dean-Stark; 2) $NaBH_4$, RT.*

Yet, subsequent reductive amination was not successful. Usually, primary amines and benzaldehydes do readily form imines by reacting them in a protic solvent.[366] The equilibrium should further be shifted in the desired direction by addition of excess sodium sulfate to bind water and thereby remove

it from the chemical equilibrium. As controlled by TLC, the imine formation could not be observed under the chosen conditions *iv)*. More drastic conditions under reflux *v)* were unsuccessful as well. The same reaction was, however, very efficient under mild conditions *iv)* when serinol **58** was used without a protective group. Depicted in Scheme IV-2, reductive amination of benzaldehydes **67**, **68**, and **69** with serinol **58** resulted in satisfying yields. Synthesis of benzylamine **62** was possibly not successful on this synthetic path because of elevated sterical hindrance at the DMT-protected precursor **61**.

To further synthesize a phosphoramidite monomer, benzylamine **63** was first protected with a DMT group *ii)* to yield precursor **62** that had not been obtained following the synthetic strategy depicted in Scheme IV-1. It was anticipated that the benzylamine function—despite its decreased nucleophilicity—could also react and form side products. That is why the diol was reacted in two-fold excess. With respect to the DMTr chloride that was added, only a 31 % yield was obtained. The only isolated side product was hydrolyzed DMTr chloride (DMT-OH) that is usually obtained due to water contamination in the reaction mixture. The reaction was just performed once and that detail was not further investigated. Though, it is noteworthy that non-reacted starting material **63** could be quantitatively recovered. Shown in Scheme IV-2 above, the conversion of the remained hydroxy function of intermediate **62** to the reactive phosphoramidite moiety *vi)* gave the phosphoramidite monomer **64** in an acceptable yield.



*Scheme IV-2: Synthetis of phosphoramidite monomers with a nitrobenzyl-protected amine starting with a reductive amination. Reductive amination iv): 1) MeOH, Na_2SO_4, RT, 2) $NaBH_4$, RT; DMT protection ii): pyr, THF, DMTrCl, RT; phosphoramidite formation vi): DCM, DIPEA, 2cyanoethyl-*N,N*-diisopropylchloro-phosphoramidite, 0 °C. marked with "*": unreacted starting material quant. recovered.*

It was not sure that oligomers could be efficiently synthesized from nitrobenzylated amine-containing phosphoramidite monomers like reagent **64** as discussed in the introduction to this subchapter. The test reaction was performed employing the standard automated protocol on an oligonucleotide synthesizer and using the commercial dT-CPG support **39**. The target oligomer of the attempt was a sequence **S24** with a leading dT monomer and an octameric block presenting *para*-nitrobenzylamino moieties. The structure is depicted in Figure IV-4. A mass spectrum of the product of the attempted oligomer synthesis is provided in the bottom of the picture. **S24** with an exact mass of 2546.5 would be detectable in an ESI(-) mass spectrum at *m/z* = 2546.6 (*z* = 1), 1272.8 (*z* = 2), 848.2 (*z* = 3), 635.9 (*z* = 4), etc. in its (multiply deprotonated) charged states. The dominant species is however detected at *m/z* = 241.1 (see mass spectrum in Figure IV-4) which could refer to the deprotonated dT monomeric unit from support **39**. It could not be understood in which way the oligomer synthesis did not work but it is obvious that the desired oligomer **S24** could not be obtained using the standard phosphoramidite protocol and phosphoramidite monomer **64**. A following approach should thus make use of a phosphoramidite monomer with a protected benzylamine function.

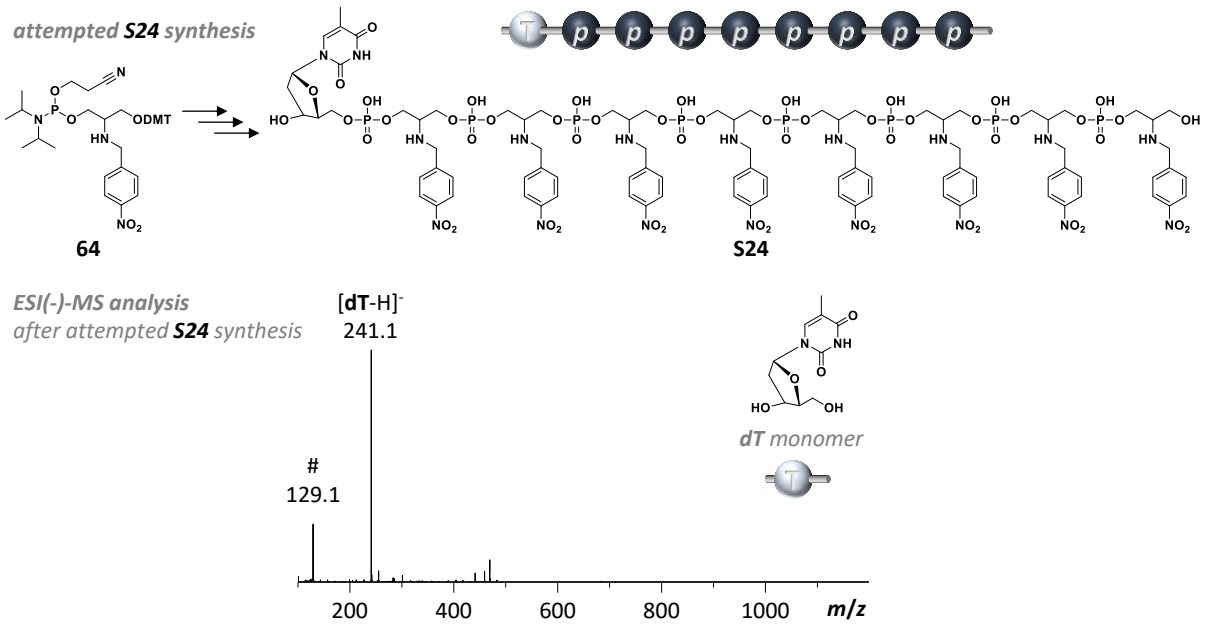


*Figure IV-4: Attempted synthesis of test sequence **S24** using nitrobenzylated amine monomer **64**. (ESI(-)-MS spectrum shown in the m/z range 100 – 1200; non-identified contamination #)*

2.2.2 Study II: Nitrobenzylated 9-Fluorenylmethyl Carbamates

Based on the result discussed before, the redesign of the project obliged to protect the nitrobenzylamine functionality to enable a successful oligo(phosphodiester) synthesis. The Fmoc protective group was chosen for that purpose. The X-unit of the trident platform that is depicted in the project's overview in Figure IV-1B and Figure IV-2B was thus changed to a 9-fluorenymethyl carbamate. The target phosphoramidite monomers are consequently the Fmoc-protected analogues of before discussed monomeric structures. Due to its lability under basic conditions, the chosen protective group would be removed during oligomer cleavage from the CPG support. The target oligomers that are given in Figure IV-5 are hence identical to those of the strategy described in Figure IV-3 on page 107.

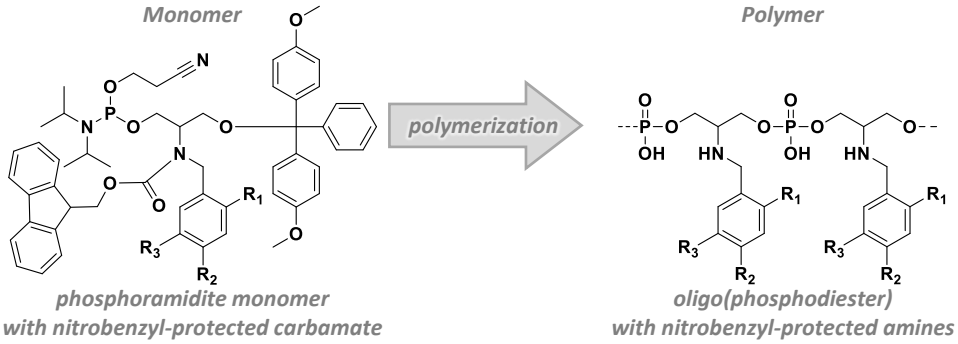
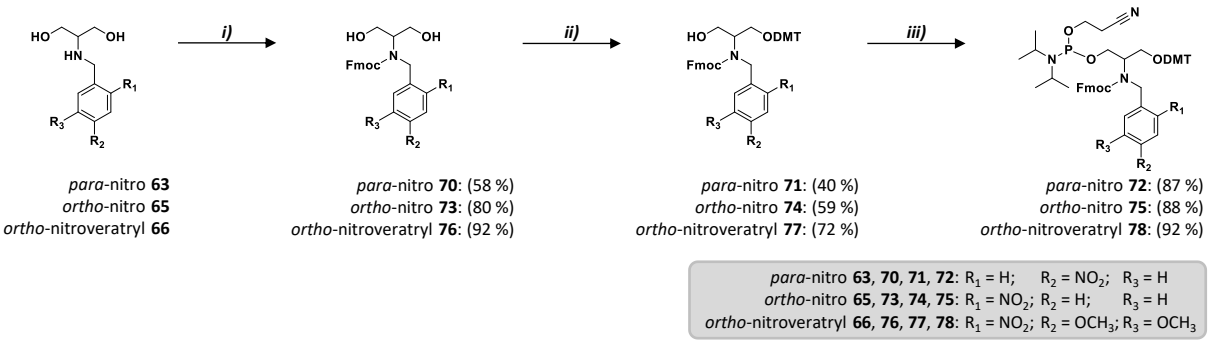


Figure IV-5: Overview carbamate phosphoramidite monomer for the synthesis of oligo(phosphodiester)s with nitrobenzyl-protected amine moieties.

The synthesis of Fmoc-protected phosphoramidite monomers is visualized in Scheme IV-3. Starting from serinol-based nitrobenzylamines **63**, **65**, and **66** that were synthesized by reductive amination, a Fmoc protective group was introduced *i)* with satisfying yields to obtain the desired series of diols **70**, **73**, and **76**. After DMT monoprotection *ii)* and subsequent introduction of the phosphoramidite moiety *iii)* the series of phosphoramidite monomers **72**, **75**, and **78** was obtained in good overall yields.



para-nitro **63**
ortho-nitro **65**
ortho-nitroveratryl **66**

para-nitro **70**: (58 %)
ortho-nitro **73**: (80 %)
ortho-nitroveratryl **76**: (92 %)

para-nitro **71**: (40 %)
ortho-nitro **74**: (59 %)
ortho-nitroveratryl **77**: (72 %)

para-nitro **72**: (87 %)
ortho-nitro **75**: (88 %)
ortho-nitroveratryl **78**: (92 %)

para-nitro **63**, **70**, **71**, **72**: R_1 = H; R_2 = NO_2; R_3 = H
ortho-nitro **65**, **73**, **74**, **75**: R_1 = NO_2; R_2 = H; R_3 = H
ortho-nitroveratryl **66**, **76**, **77**, **78**: R_1 = NO_2; R_2 = OCH_3; R_3 = OCH_3

*Scheme IV-3: Phosphoramidite monomer synthesis including an amine protection with a 9-fluorenylmethyl carbamate (Fmoc) group. Fmoc protection i): Na_2CO_3 (aq., 10w %), 1,4-dioxane, Fmoc-Cl, RT; DMT protection ii): pyr, THF, DMTrCl, RT; phosphoramidite formation iii): DCM, DIPEA, 2cyanoethyl-*N,N*-diisopropylchloro-phosphoramidite, 0 °C.*

With Fmoc-protected phosphoramidite monomers **72**, **75**, and **78** in hand, oligomers with corresponding octameric blocks consisting of the same repeating motif were aimed at. It is shown in Figure IV-6 that the sequence **S24** that could not be obtained using the nitrobenzyl phosphoramidite **64** could indeed be synthesized from the analogue protected monomer **72**. The mass spectrum that is represented on the central left of Figure IV-6 clearly shows the expected peaks for two times and three times deprotonated **S24** species. The triply charged species at *m/z* = 635.9 was also detected but is found in a region together with several series of lower congeners with mass defects stemming from undesired but not understood side reactions. When the same octameric block was synthesized on a dA-CPG support, the intensity of the so-attained sequence **S25** with a leading dA monomer was higher, but the same series of lower congeners was isolated together with the desired sequence. When using phosphoramidite monomers **75** and **78**, the analogue oligomers **S26** and **S27**, respectively, were obtained. Also they were not isolated with the desired purity and degenerated lower congeners were detected in ESI(-)-MS which is depicted at the bottom of Figure IV-6.

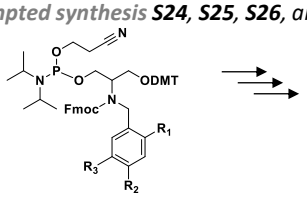
Interestingly, sequences **S24** and **S26** are regioisomers and consequently the detected ions have the same mass. Also their main impurities that can be analyzed by mass spectrometry are regioisomers. That leads to the conclusion that whatever renders the phosphoramidite chemistry inefficient, it does not stem from the different benzylic substitution pattern. It is hypothesized that the Fmoc protective group together with the nitrobenzyl and the bulky DMT group hinder a fast and efficient phosphoramidite coupling for steric reasons. It could have been an interesting task to study the underlying chemistry by increasing the reaction time during phosphoramidite activation and coupling or by repeating the synthesis with a longer linker between phosphoramidite moiety and the other functionalities. Yet, this was not the original purpose of the project.

It was a great success that the redesign of phosphoramidite monomers had the formation of the desired products **S24**, **S25**, **S26**, and **S27** as a result. Yet, it has to be critically noted that the introduction of a bulky temporal protective group was not necessary. In the same way, a persistent small amide could have been formed from nitrobenzylamines **63**, **65**, and **66** to obtain monomers and later oligo(phosphodiester)s with nitrobenzyl-substituted amides. That would still have been an analogue strategy to already reported polymeric systems[360-361] and at the same time it would have avoided extensive sterical hindrance in close proximity to the reactive moiety during polymerization.

First photocleavage tests on sequences **S26** and **S27** were performed. The analytes were dissolved in methanol and irradiated with UV light of the wavelength 365 nm. The slightly yellow solutions thereby turned to intensive red coloration. The before detected peaks shown in Figure IV-6 could not be detected anymore but with the addition that neither the fully deprotected oligomer nor a partially deprotected oligomer could be detected by means of mass spectrometry. It was obvious that a

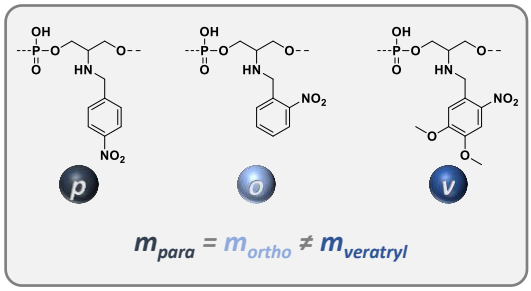
reaction had occurred and in order to analyze what chemical transformation was induced by irradiation it was moved to a less complex test system.

*attempted synthesis **S24**, **S25**, **S26**, and **S27***



S24**, **S25**, **S26**, and **S27

72: R_1 = H; R_2 = NO_2; R_3 = H.
75: R_1 = NO_2; R_2 = H; R_3 = H.
78: R_1 = NO_2; R_2 = OCH_3; R_2 = OCH_3.



*ESI(-)-MS analysis of reaction prducts **S24**, **S25**, **S26**, and **S27***

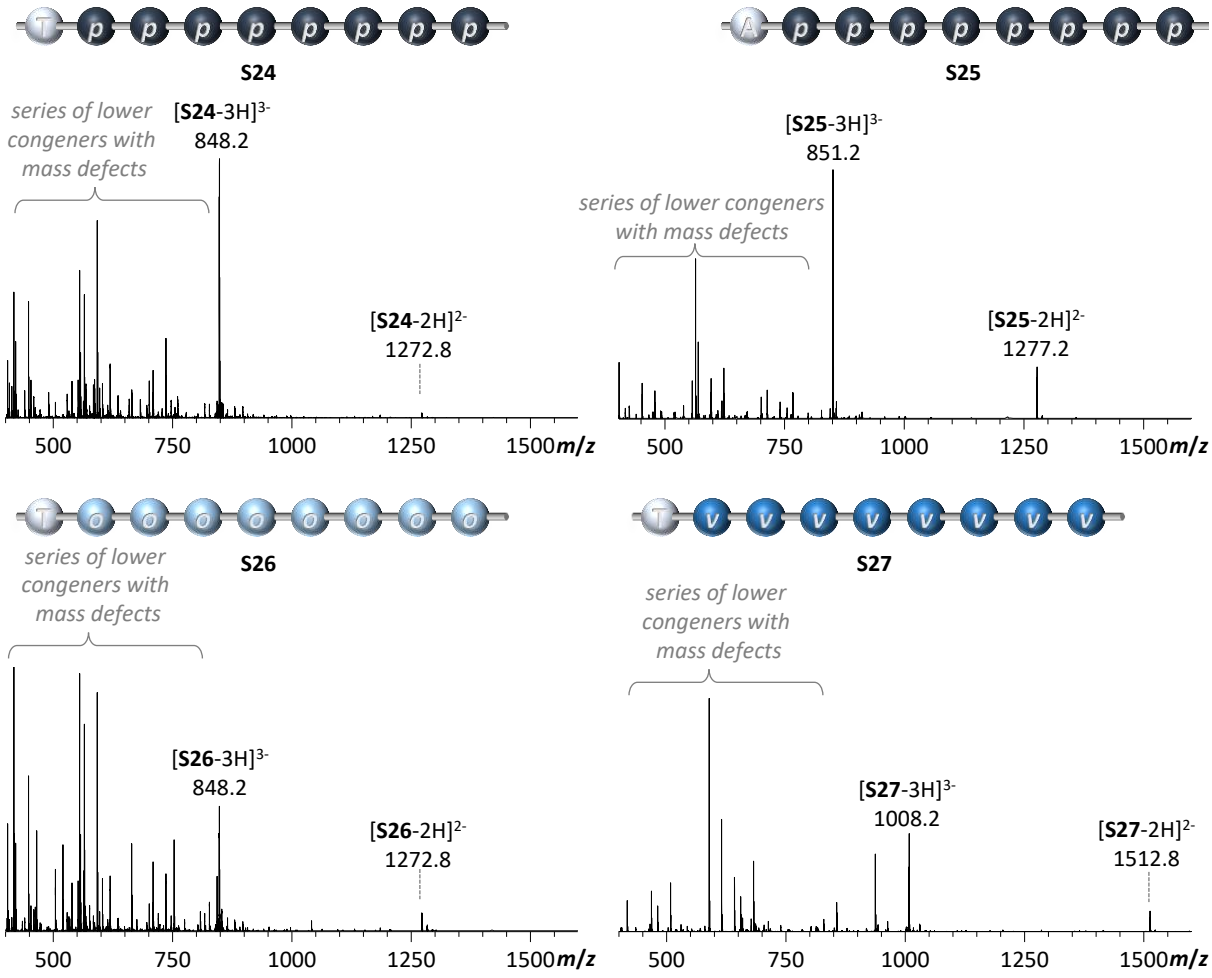


*Figure IV-6: Attempted synthesis of test sequences **S24**, **S25**, **S26**, and **S27** using nitrobenzylated carbamate monomers **72**, **75**, and **78**. (ESI(-)-MS spectra shown in the m/z range 400 – 1600; range 100 – 400 cut for clarity: TCA contamination at 116.9, 160.9)*

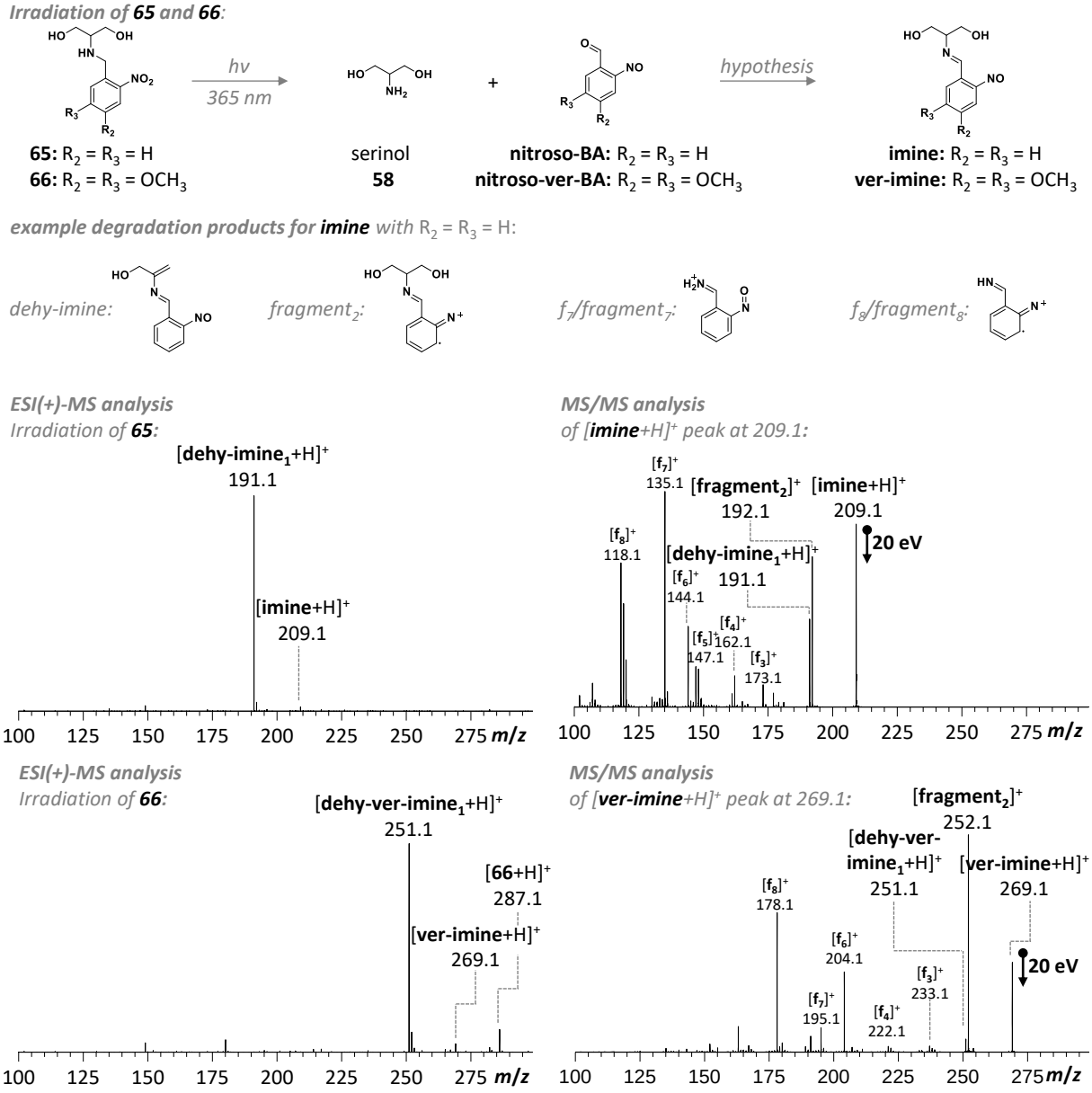


*Figure IV-7: Photoinduced cleavage of nitrobenzylamines **65** and **66** and (tandem) MS analysis of reaction products.*

Under the same conditions nitrobenzylamine **65** and **66** were irradiated. Consumption of starting material could be followed by ^{1}H NMR that was performed in time intervals of minutes. The appearing signals were however too complex to understand the process. When all starting material was consumed—as determined by NMR—the reaction mixture was analyzed by mass spectrometry. An overview of the analysis and the hypothesis derived from it is depicted in Figure IV-7. After irradiation of **65** and **66** MS analysis of the reaction mixtures results in almost identical mass spectra with a shift of 60.0 Da harmonizing with the additionally added weight by two methoxy groups in the case of the nitroveratrylphenyl motif. The intermediately built nitroso benzaldehydes (nitroso-BA) are not detected but it is probable that both cleavage products, serinol and aldehyde, are again converted to the corresponding imine. Species with the mass of the imine and its analogue after water elimination are detected in ESI(+)-MS. Further fragmentation of the imine species leads to the detection of several fragments that can be traced back to a logic decay of an initial imine in the gas phase. The hypothetic path of imine decay is shown for the *ortho*-nitro compound in the Annex in Figure V-14 on page 194.

An imine formation subsequent to the photorelease of a nitroso compound was most probably also occurring when sequences **S26** and **S27** were investigated. It is possible to suppress this reaction by

performing the reaction in an aprotic solvent or by adding an excess of a scavenger amine. But as the oligomer synthesis was not optimal yet and the photorelease of nitrobenzyl groups was showing complex side reactions it was considered more straight forward to redesign the project for another time.

Again, it has to be noticed that a similar approach relying on nitrobenzylated amide moieties in the target oligomers could also have resulted in a less complex photochemistry due to the fact that an amide would not readily react with the formed nitroso benzaldehyde.

2.2.3 Study III: Nitrobenzylated Alcohols

Finally, it was decided to aim for a series of nitrobenzyl ethers in the phosphoramidite monomers as well as the corresponding oligo(phosphodiester)s. The target motifs are depicted in Figure IV-8. Same as amides that were discussed before, a liberated hydroxyl would not react with a benzaldehyde after photodeprotection.

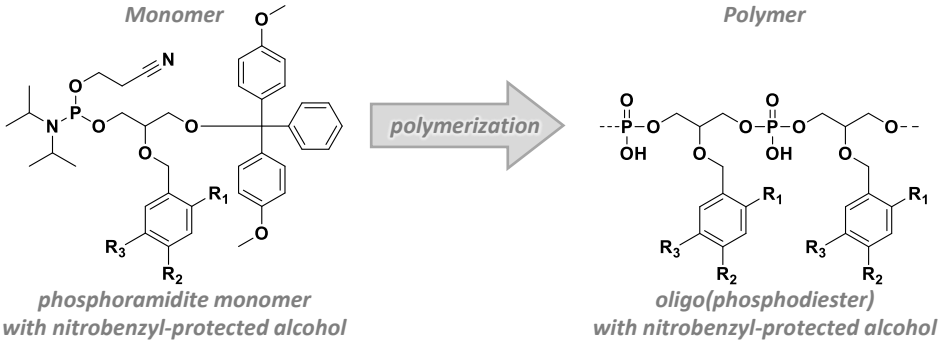
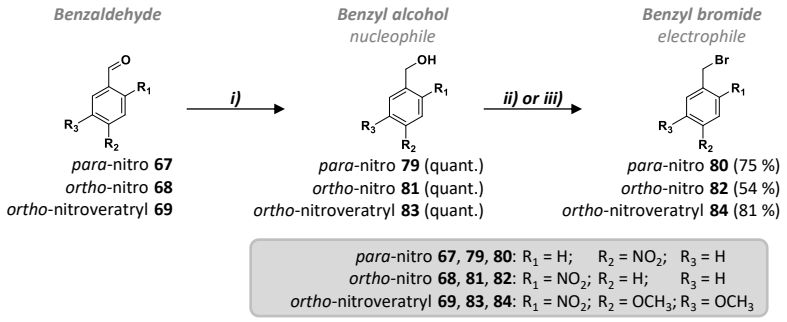


Figure IV-8: Overview phosphoramidite monomer for the synthesis of oligo(phosphodiester)s with nitrobenzyl-protected hydroxy groups.

Glycerol **85** is the hydroxy analogue of serinol **58**. It was identified as a suitable commercial reagent to build the trident platform of the target phosphoramidite monomers. The series of nitrobenzyls should be inserted as ethers. A strategy for an etherification was consequently developed. In Scheme IV-4 it is shown how corresponding nucleophiles **79**, **81**, and **83** and electrophiles **80**, **82**, and **84** can be obtained from the parent benzaldehydes. Even though commercially available they were quickly synthesized for the first test reactions.

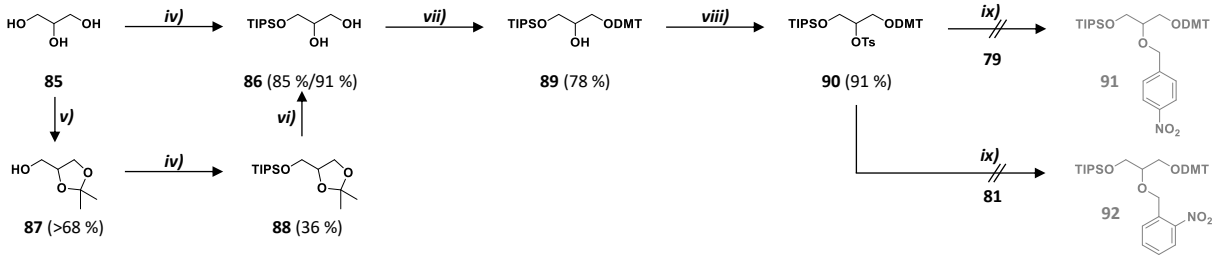


*Scheme IV-4: Synthesis of nucleophiles **79**, **81**, and **83** and electrophiles **80**, **82**, and **84** for nitrobenzylation reactions. Aldehyde reduction i): MeOH, DCM, $NaBH_4$, RT; bromination for **80** and **82** ii): THF, PBr_3, 0°C; Appel-like reaction for **84** iii): THF, CBr_4, PPh_3, 0°C.*

Nitrobenzyl alcohols **79**, **81**, and **83** were quantitatively obtained by reduction of corresponding benzaldehydes **67**, **68**, and **69** with sodium borohydride *i)*. Subsequent conversion of the hydroxyl function to a bromide gave nitrobenzyl bromides **80**, **82**, and **84** in good yields. When phosphorus tribromide was used for bromination *ii)* the yields were slightly lower compared to the same conversion that was achieved by an adapted Appel reaction with tetrabromo methane *iii)*. As a

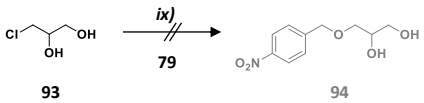
quantitative consumption of benzyl alcohol was detected before reaction work-up, the lower yield using phosphorus tribromide can be explained by the amount of aqueous sodium bicarbonate solution that was added to the reaction mixture to quench excess tribromide and the so-formed hydrogen bromide. Nitrobenzyl bromides are quite polar and were not quantitatively extracted from the aqueous phase. After the Appel reaction in contrary, an aqueous work-up can be avoided and after removal of the solvent the crude can be applied to a short column and purified by chromatography in an efficient way.

In Scheme IV-5 the first strategy for the synthesis of nitrobenzyl ethers on a trident platform is shown. In order to synthesize a tosylate nucleophile **90** that could form an ether with the nitrobenzyl nucleophiles shown in Scheme IV-4, glycerol **85** was first monoprotected on a primary alcohol with TIPS *iv)*. That could be achieved with 85 % yield when glycerol was used in excess. Statistical double protection also occurred but the purification of TIPS-protected glycerol **86** was efficient. An alternative route converted the triol **85** first to the dioxolane **87** *v)* whose remaining hydroxy function was protected with TIPS *iv)* followed by acidic acetal opening *vi)* to obtain TIPS-protected glycerol **86**. The latter was protected with DMT *vii)* at the remaining primary alcohol moiety and the secondary alcohol **89** was converted to a tosylate.



*Scheme IV-5: Synthesis of a DMT-protected tosylate precursor **90** for nitrobenzylether formation by nucleophile insertion. TIPS protection iv): pyr, DCM, TIPS-Cl, 0°C; acetal formation v): toluene, reflux in Dean-Stark, pTsOH, acetone; acetal hydrolysis vi): AcOH (aq., 80 v%), DCM, 80 °C; DMT protection vii): pyr, THF, DMTrCl, RT; tosylate formation viii): pyr, pTsCl, 0 °C; ether formation ix): various conditions.*

Tosylation of alcohols is used to convert them to good leaving groups for nucleophilic substitutions. That is a necessary prerequisite for a successful etherification with an alcohol nucleophile. Yet, all etherification attempts with tosylate **90** remained unsuccessful. The nucleophile was prepared from **79** and **81** by deprotonation with sodium hydride or potassium *tert*-butanolate in THF. When potassium hydroxide in DMF was used, the only conversion detected was the loss of TIPS and tosyl group from starting material **90**. It can be concluded that the etherification strategy was a high-risk approach because the literature's experimental descriptions[367] it was based on were of low quality. Another literature procedure[358] reporting the ether formation between benzyl alcohol **79** and epichlorohydrine (the epoxide analogue of diol 93 in Scheme IV-6) was neither reproducible. When the same reaction was performed with 3-chloropropane-1,2-diol **93** under various conditions, no conversion was observed. Only when using potassium *tert*-butanolate in anhydrous THF, the formation of a side product could be traced. Instead of target product benzyl ether **94**, a benzylic ester **94-ox** was probably formed. The structure of **94-ox**, MS analysis and hypothesis explaining the observations are presented in Figure IV-9. The MS/MS fragmentation of the main product peak harmonizes with the hypothesis. However, it remains unclear at which stage the oxidation could have occurred.



*Scheme IV-6: Attempt of a nucleophilic attack on 3-chloropropane-1,2-diol **93**. Ether formation ix): various conditions.*

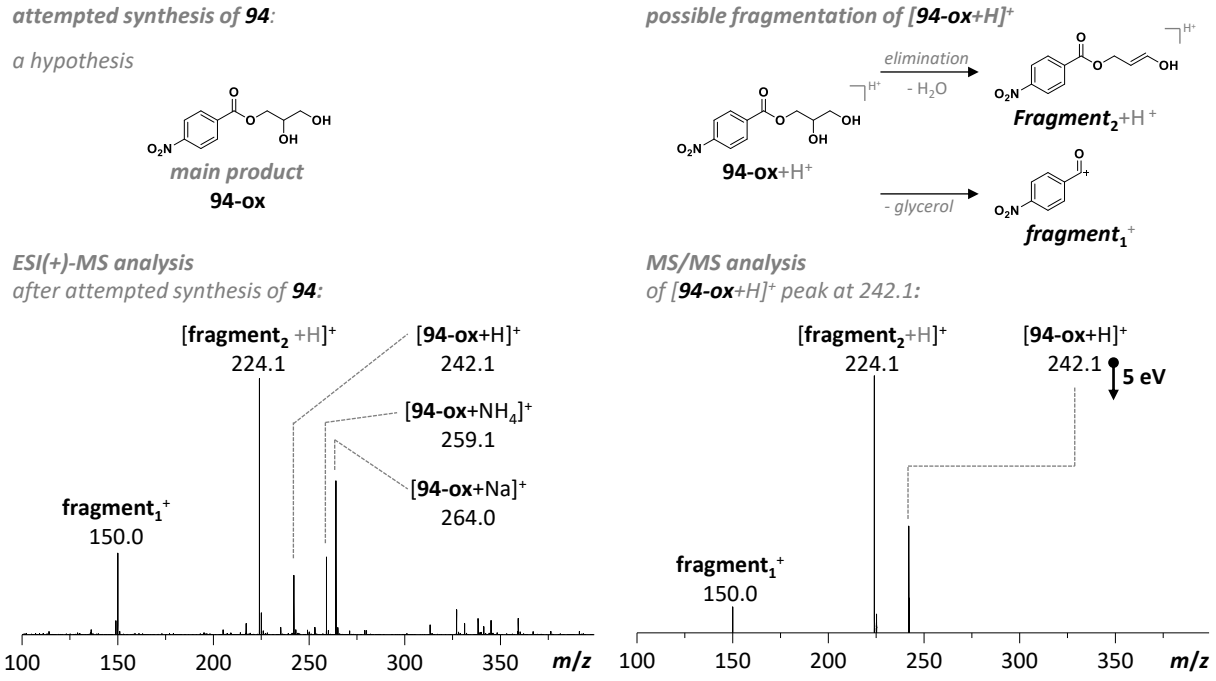
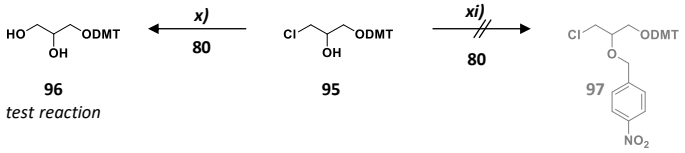


*Figure IV-9: Analysis of main product after attempted synthesis of nitrobenzyl ether **94**.*

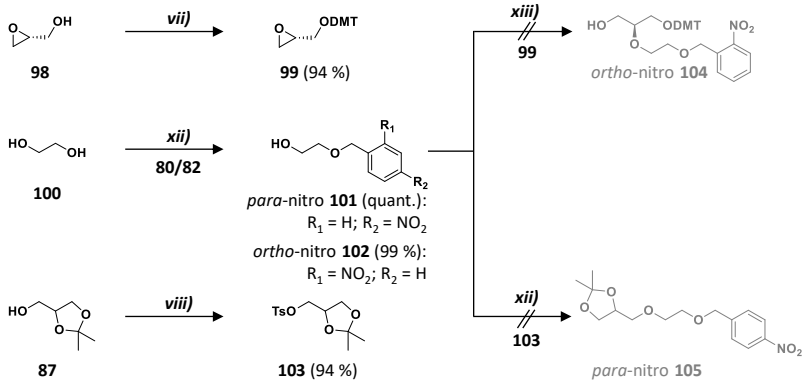
It was concluded that nitrobenzyl alcohols were not suitable nucleophiles for etherifications. The negative inductive effect of the nitro group is probably too strong and renders the benzylic alcohol a poor nucleophile as the negative charge after deprotonation can be stabilized.

The strategy was consequently adapted and electrophilic bromides **80**, **82**, and **84** should be used for the insertion of nitrobenzyl groups during etherification. In a first attempt, a DMT-protected analogue of chloride **93** was investigated for its potential. The test reaction *x)* depicted in Scheme IV-7 showed that the chloride of **95** is a synthon for the necessary alcohol as it could be hydrolyzed efficiently using triethylamine hydroxide as a base. Yet, the etherification with nitrobenzyl bromide **80** was not successful. Neither deprotonation of **95** in THF or DMF[368] followed by addition of bromide **80**, nor the use of silver oxide[369-370] as reagents was giving access to ether **97** and no conversion was observed.



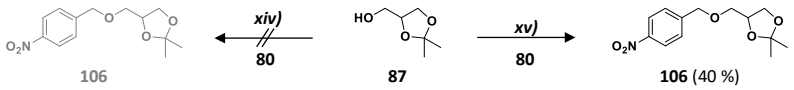
*Scheme IV-7: Attempted use of DMT-protected 3-chloropropane-1,2-diol **95** as a nucleophile. Alkyl chloride hydrolysis x): TEA hydroxide (aq., 35 w%), THF, reflux; ether formation xi) various conditions.*

The ether formation worked best when no additional solvent was added. The reaction between ethylene glycol and nitrobenzyl bromides **80** and **82** was (almost) quantitative when ethylene glycol was used in great excess and sodium hydride was added as the base. Subsequent conversion with the bromide resulted in ethers **101** and **102** that are depicted in Scheme IV-8 and that were obtained with considerably higher yields than the literature reported.[371-372] In an analogous approach using dioxolane **87** instead of ethylene glycol without addition of a solvent, no conversion was observed. It was further tried to make use of the ethylene linked ether moieties of **101** and **102**. They should be coupled as nucleophiles to diol precursors **99** and **103** that are shown in Scheme IV-8. Attempts to synthesize corresponding structures **104** and **105** following similar reported systems[373] were not successful.



Scheme IV-8: Synthesis of an ethylene glycol-linked nitrobenzylic ether and attempted coupling to a diol precursor. DMT protection vii): pyr, THF, DMTrCl, RT; tosylate formation viii): pyr, pTsCl, 0 °C; solvent-free ether formation xii): NaH; nucleophilic epoxide opening xiii): DMF, KOt-Bu, 100 °C;

It was already mentioned that neither the solvent-free conversion of dioxolane **87** to the target compound **106** was achieved. Nor was it possible to synthesize **106** by a classic etherification from the dioxolane's sodium salt in THF *xiv)* shown on the left of Scheme IV-9. When the solvent was changed to acetonitrile, however, the desired product **106** was formed. The procedure was adapted from a similar reported system[374] with dioxolane moieties and a primary alcohol that was converted to a 2-nitrobenzyl ether. A yield of 88 % was reported when using a microwave and elevated temperatures. And indeed, the procedure *xv)* allowed the synthesis of the whole target series of nitrobenzyl ethers **106**, **110**, and **114** (shown in Scheme IV-10 on page 117). Yet, it was found that microwave irradiation decreased the yield. Conversion at 0 °C to room temperature was beneficial.



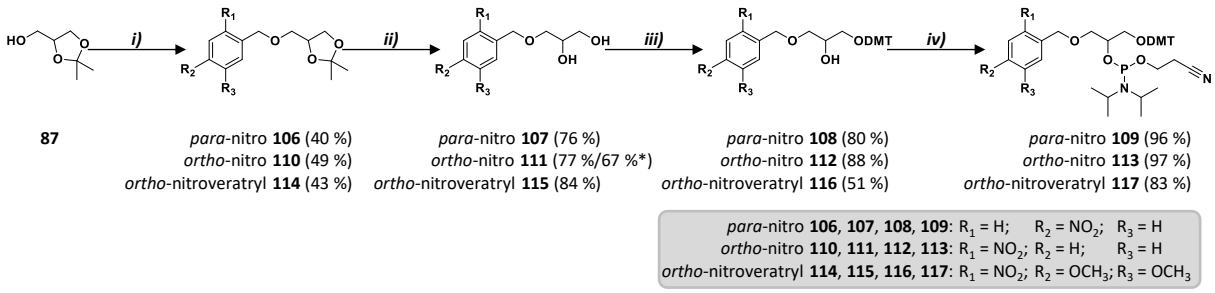
*Scheme IV-9: Synthesis of ether **106**. Attempted ether formation xiv): THF, NaH, 0 °C; successful ether formation xv): MeCN, NaH, 0 °C.*

As the synthetic acces of nitrobenzyl ethers on a diol precursor was found and subsequent monomer synthesis, oligomer synthesis and photoindinduced side chain modification was promising, the further progress is not described as part of a preliminary study anymore but will be reported in the following separated subchapters.

2.3 Monomer Synthesis

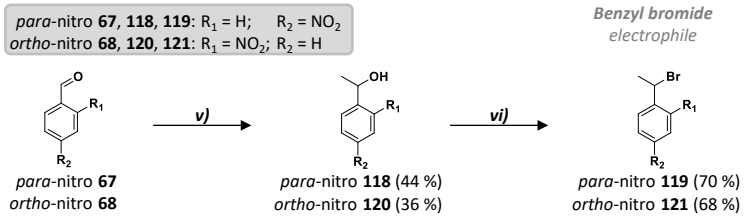
The complete synthesis of nitrobenzyl ether phosphoramidite monomers is depicted in Scheme IV-10. After etherification in acetonitrile *i)*, the series of nitrobenzyl ether dioxolanes was hydrolyzed *ii)* in acetic acid to obtain the series of diols **107**, **111**, and **115**. The latter were protected with DMT *iii)* and the remaining alcohol was converted *iv)* to the reactive phosphoramidite function to obtain the phosphoramidite monomers **109**, **113**, and **117** in good yields. It was furthermore shown for diol **111** that etherification *i)* and hydrolysis *ii)* of the acetal could be performed without the isolation of the intermediate to obtain the desired product in a very satisfying yield of 67 % over two steps.

If the obtained monomers from this attempt are compared with the general design that was initially aimed for (see Figure IV-1B and Figure IV-2B), the criteria are still met with factors $m = 0$, $n = o = 1$, and X = O. The synthesis of oligomers from monomers **109**, **113**, and **117** and their irradiation will be presented and discussed in the subsequent subchapters.



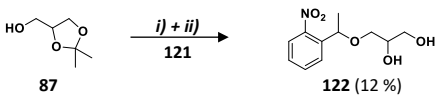
*Scheme IV-10: Synthesis of phosphoramidite monomers **109**, **113**, and **117** with nitrobenzyl-protected hydroxy groups starting with a successful etherification in acetontrile. Ether formation i): MeCN, NaH, 0 °C; acetal hydrolysis ii): AcOH (aq., 80 v%), DCM, 80 °C; DMT protection iii): pyr, THF, DMTrCl, RT; phosphoramidite formation iv): DCM, DIPEA, 2cyanoethyl-*N,N*-diisopropylchlorophosphoramidite, 0 °C. Marked with "*": over two steps, without purification of intermediate dioxolane.*

The introduction of a methyl group at the benzylic position of *ortho*-nitrobenzyls is known to increase quantum yields of the photochemical cleavage reaction.[16] There are also less complex reaction cascades of byproducts after cleavage reported[375] which could be beneficial—for instance when monitoring a photoirradiation by UV or NMR—for more in-depth studies of the photodeprotection reaction. To be able to show that perspectively even this more efficient photoreleasable group could be used, corresponding benzyl bromides **119** and **121** were synthesized. The synthesis is shown in Scheme IV-11 and starts with the *Grignard*-alike insertion *xvii)* of the methyl group in benzaldehydes



*Scheme IV-11: Synthesis of nitrobenzyl bromide electrophiles **119** and **121** for the synthesis of 1-(2-nitrophenyl)ethyl and 1-(4-nitrophenyl)ethyl ethers, respectively. Methylation v): THF, $TiCl_4$, MeLi, -78 °C; bromination vi): THF, PBr_3, 0°C.*

67 and **68**. The corresponding secondary alcohols **118** and **120** were obtained in acceptable yields. In the literature[376] much higher yields are reported and it is assumed that the available *anhydrous* THF was not of sufficient quality and should better have been freshly prepared. Subsequent conversion of the alcohols to electrophilic bromides **119** and **121** using phosphorus tribromide in *ii)* was already discussed before for bromides **80** and **82**. Depicted in Scheme IV-12, the synthesis of diol **122** analogous to **111** was performed in one step *xv) + vi)* but resulted in a lower yield. That was possibly due to a competing elimination reaction of the bromide **121** under basic conditions.



*Scheme IV-12: Synthesis of 3-(1-(2-nitrophenyl)ethoxy)propane-1,2-diol **122**. Ether formation i): MeCN, NaH, 0 °C; acetal hydrolysis ii): AcOH (aq., 80 v%), DCM, 80 °C; without purification of intermediate dioxolane.*

When a methyl group is introduced at the benzylic position, the crucial etherification is less efficient but does work in principle. The reaction should be further optimized, but in case of future needs, the chemistry is not to be completely revised. Even a reaction upscale could at this point already lead to an acceptable amount of diol to obtain the corresponding phosphoramidite monomers on a sufficient scale for subsequent oligo(phosphodiester) synthesis.

2.4 Realization of Property A: Erasing Sequence Information

With nitrobenzyl ether phosphoramidite monomers **113** and **117** in hand, test sequences **S28** and **S29** were synthesized. As the test sequences studied before, they contained a leading dT monomer and an octameric block of a repeating monomeric unit derived from one of the monomers. They are depicted together with their MS analysis in Figure IV-10. While the sequence **S29** was detected by ESI(-)-MS in its multiple charged states from -2 to -5, the *ortho*-nitroveratrylbenzyl-containing sequence **S28** was detected in its multiple charged states from -3 to -5. Additionally, not identified species **#** were also observed in the spectrum.

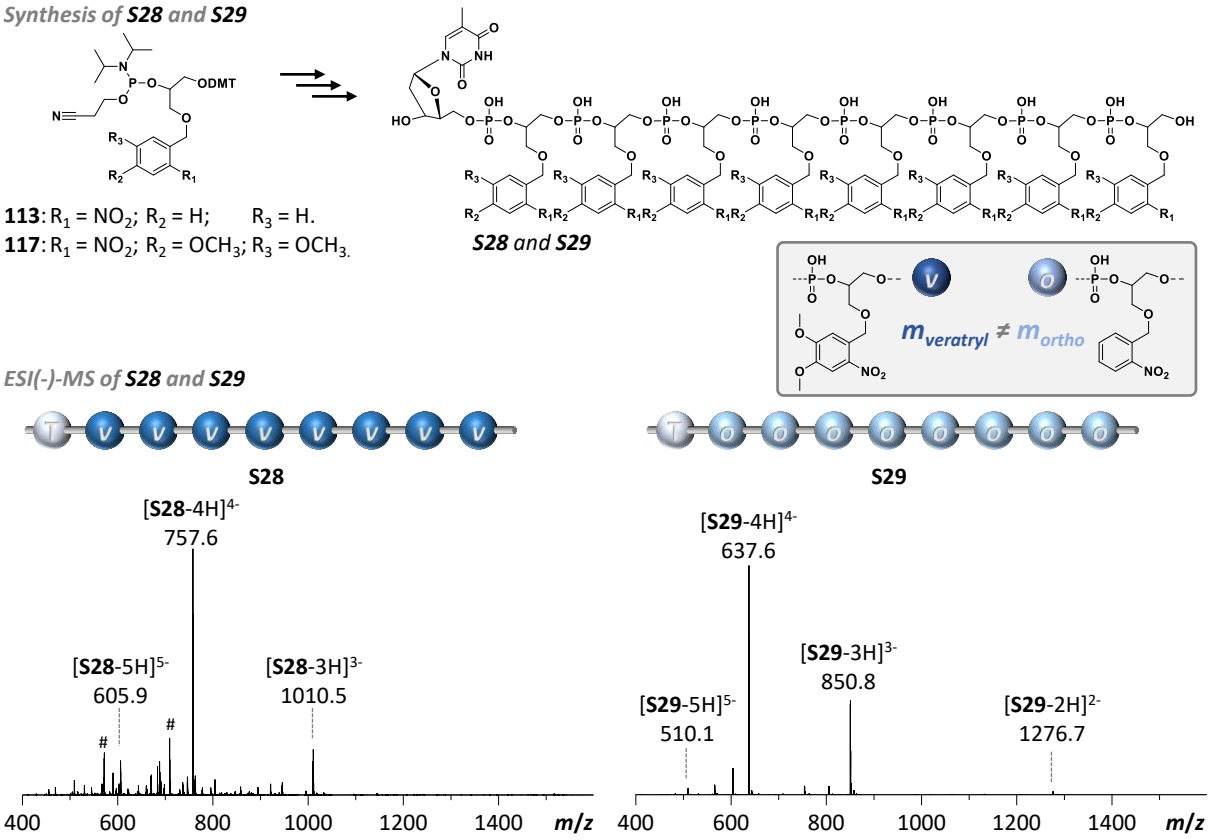


*Figure IV-10: Synthesis of test sequences **S28** and **S29** using nitrobenzyl ether monomers **113** and **117**. Indicated with **#** are unknown contaminations. (ESI(-)-MS spectra shown in the m/z range 400 – 1600.)*

Sequences **S28** and **S29** were analyzed by ^{1}H and ^{31}P NMR in methanol-d^4. The results are shown in Figure IV-11. Most prominent differences between the analogue oligomers are observed in the pattern of aromatic protons *(e)* that directly reflect the phenyl substitution. The *ortho*-nitrobenzyl-containing oligomer **S29** shows four sets of such signals in the ^{1}H NMR spectrum; the spectrum of the *ortho*-nitroveratryl analogue **S28** shows a less sharp resolution of the aromatic protons, possibly due to interaction with neighboring groups in the close proximity. More prominent are here the signals for the veratryl methoxy protons *(f)* at 3.7 to 3.9 ppm. It is also evidenced by NMR that the oligo(phosphodiester)s are isolated as triethylammonium salts. Corresponding signals are marked with *TEA* in Figure IV-11. ^{31}P NMR confirms that phosphate groups are present reflected by signals around 0 ppm.

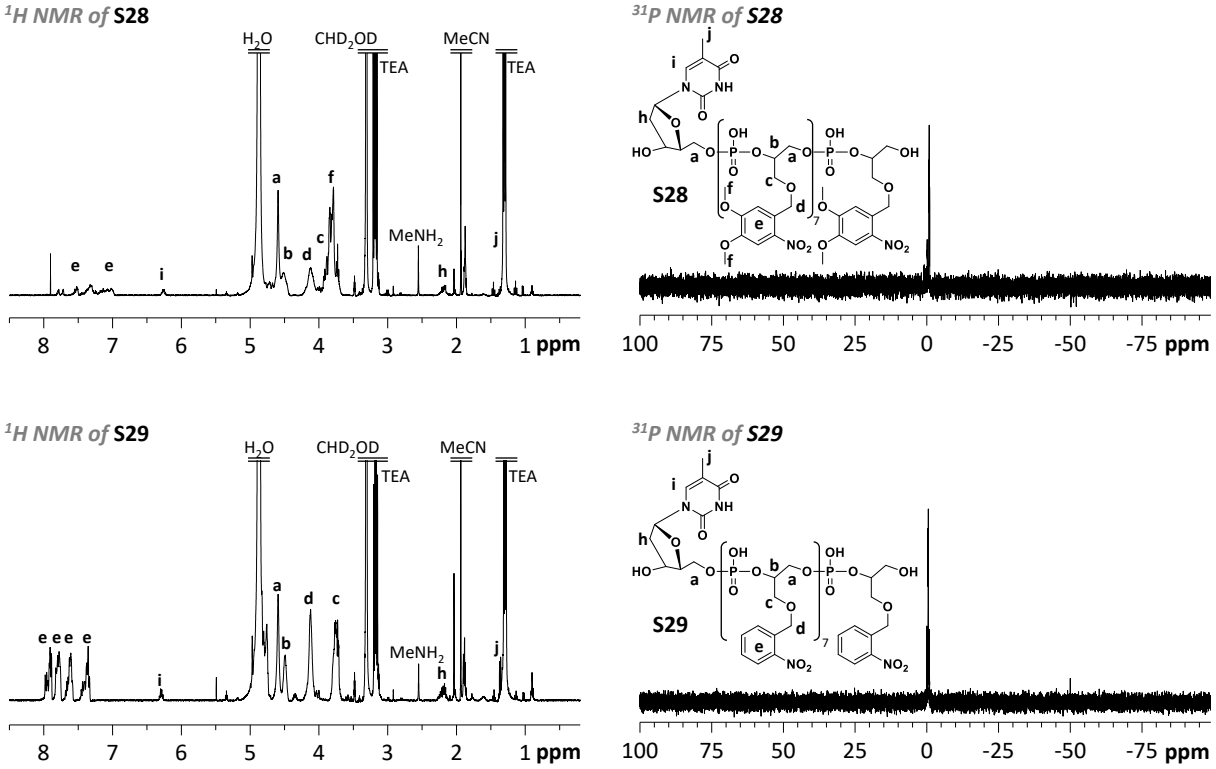


*Figure IV-11: ^{1}H NMR and ^{31}P NMR spectra of sequences **S28** and **S29**, recorded in methanol-d^4, shown in the range of 8.5 to 0.0 ppm and 100 to -100 ppm, respectively.*

Sequence **S29** was clearly characterized by MS and NMR. For sequence **S28**, the additional detection of contaminats **#** by mass spectrometry analysis does not harmonize with its rather accurate NMR analysis. In the same way, the results obtained for sequences **S30**-**S34** containing the *ortho*-nitroveratrylbenzyl ether motif are until now not excellent. It should also in the following paragraphs be taken into account that those sequences gave NMR spectra with the expected signals but MS analysis of this series showed some unexpected peaks of of contaminants or degradation products. It has to be noted that the *ortho*-nitrobenzyl ether-containing sequence **S29** the series of sequences **S28**-**S34** containing *ortho*-nitroveratrylbenzyl ether motifs that were studied for the concept *Erasing sequence information* were characterized by mass spectrometry with varying relative amount of contaminants. Currently, investigations are ongoing to evaluate whether the MS detection of sequences **S28**-**S34** can be improved and in which way different solvent mixtures lead to more or less detected contaminants. These crucial insights could not be covered by the present manuscript anymore due to time restrictions. Yet, the obtained results will be presented and contaminants are marked with **#** in MS spectra. While for sequences **S28**-**S34** the detection by mass spectrometry had already been improved, the MS analysis for sequences **S28'**-**S34'** is still shown following less optimized solubilization protocols.

The first test photocleavage was performed on *ortho*-nitrobenzyl ether-containing sequence **S29**. A NMR tube filled with 0.5 mL of a 1.0 mM solution of **S29** in deuterated methanol was irradiated for 60 min without further precautions. The so-treated solution developed a slightly yellow coloration and it was analyzed by ESI(-) mass spectrometry. The result is shown in Figure IV-12. The dominant peak at *m/z* = 736.1 stems from the double deprotonated desired sequence **S29'**. Besides the detected product, also species are detected that show in their non-charged form an exact mass that is multiples of 149 Da higher than that of the desired product **S29'**.

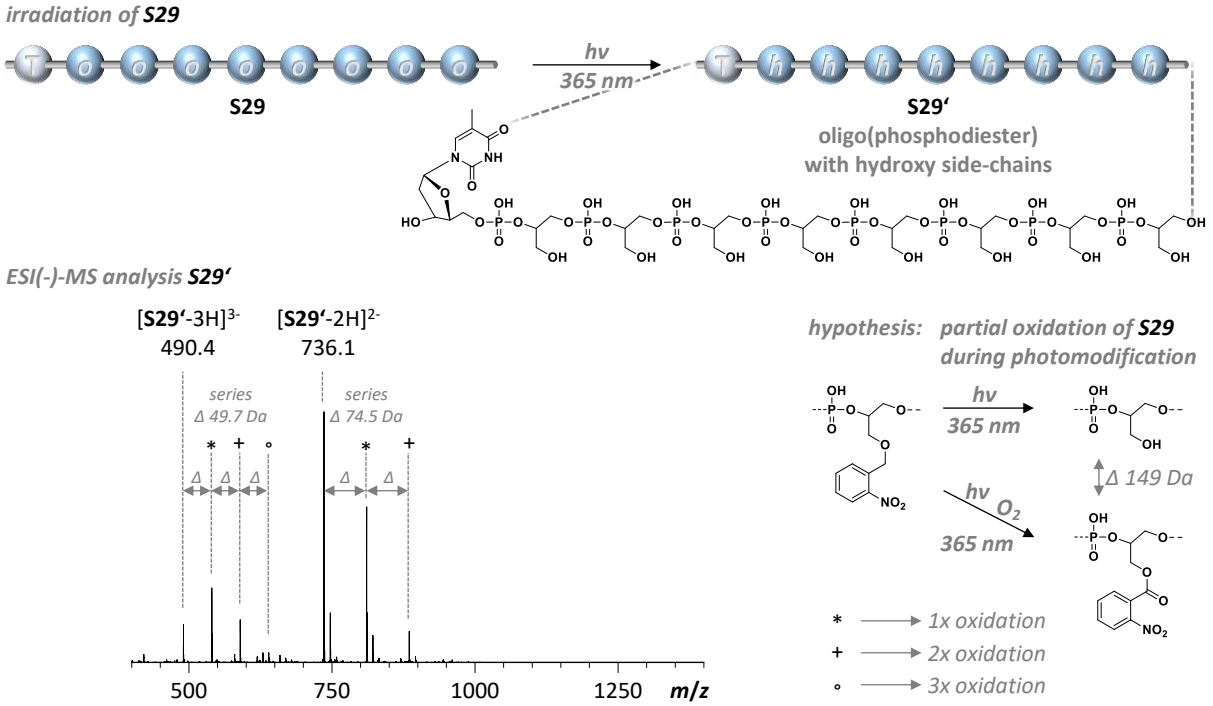


*Figure IV-12: First photomodification of **S29** to **S29'**. (MS spectrum shown in m/z range 400 – 1400; side products detected at m/z = 540.0 (*), 589.7 (+), 639.4 (∘) as triply deprotonated species and m/z = 810.6 (*), 885.1 (+) as double deprotonated species, respectively)*

Visualized on the bottom right of Figure IV-12, it was hypothesized that the benzylic position could have been statistically oxidized forming the corresponding benzylic ester moieties in presence of oxygen. These esters are followingly rendered stable under the irradiation conditions. Indeed, accurate mass measurement of species at *m/z* = 540.0 (*), 589.7 (+), 810.6 (*), and 885.1 (+) confirmed the elemental distribution of the oxidation hypothesis. Furthermore, tandem mass spectrometry analysis of these species showed for side product ***** that +149 Da-fragments were detected together with fragments without the defect. For species **+**, there were fragments additionally found with a +298 Da (= 2x 149 Da) defect. These findings strongly suggest that oxidations occurred statistically along the chain.

In order to obtain a clean sequence **S29'**, an additional base treatment could potentially be employed to hydrolyze the benzylic esters. It was, however, decided to perform the *ortho*-nitrobenzyl photorelease under exclusion of oxygen to suppress the side-reaction in the first place. In Figure IV-13 it is shown that **S29** was detected by MS in its -2, -3, and -4 charge states together with contaminations **#** before irradiation. After irradiation, these species are not detectable anymore. However, sequence **S29'** with hydroxyl side chains (symbolized by *h* in the cartoon representation in Figure IV-13) was observed in its -2 and -3 charge states. The relative amount of oxidation side product was drastically decreased when compared with the irradiation with contact to air (see Figure IV-12). Only species with a single statistical oxidation ***** were detected. The oxidation could probably occur because the NMR solvent methanol that was used to solubilize sequence **S29** contained oxygen. The photoirradiation should thus be repeated in degassed methanol under exclusion of oxygen.

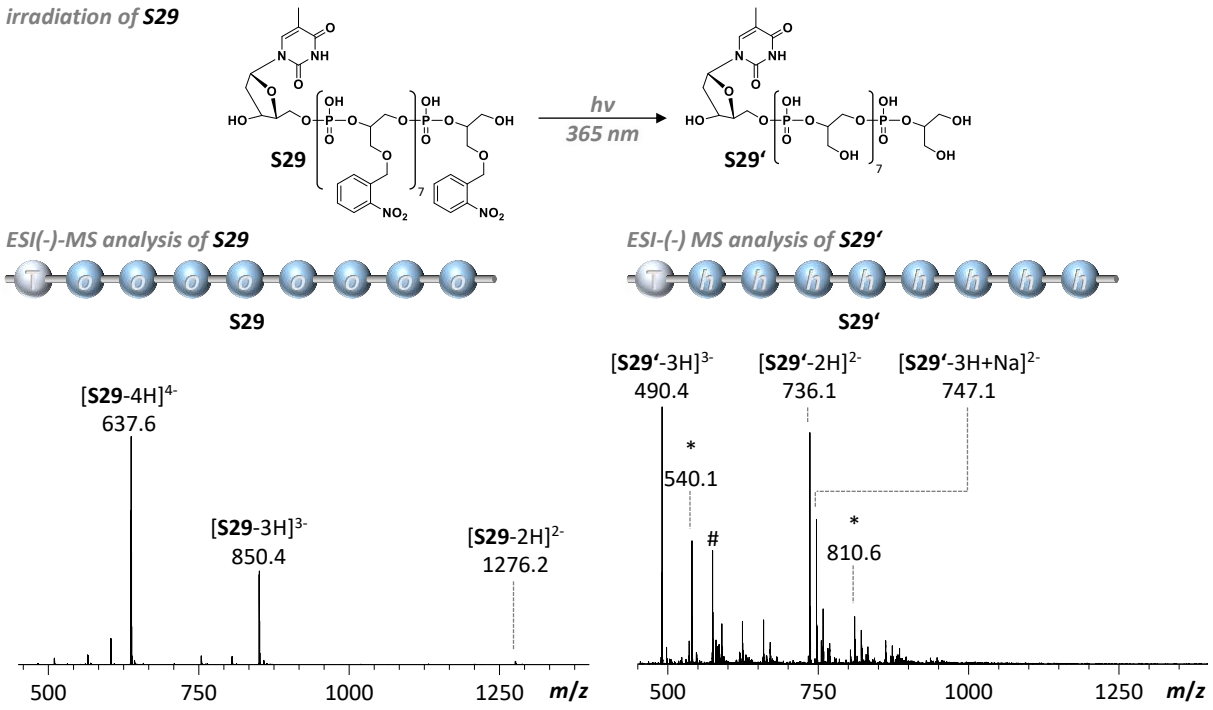


*Figure IV-13: Photomodification of **S29** to **S29'**. MS spectra of **S29** and **S29'** shown in m/z range 450 – 1400; oxidation side product * detected at m/z = 540.1 as triply deprotonated species and m/z = 810.6 as double deprotonated species, respectively.*

The irradiation of *ortho*-nitroveratrylbenzylic ether-containing oligomer **S28** did also result in a complete conversion. The photomodified sequence **S28'** that is identic to **S29'** was, however, not detected by MS as the spectrum was dominated by a contamination #. This was only the case for irradiated sequence **S28'**.

When a whole series of information-containing oligo(phosphodiester)s **S30**-**S34** was synthesized from monomers **113** and **117**, the sequences were characterized by ^{1}H NMR, ^{31}P NMR and MS (see Annex, pp. 209ff.). After irradiation with UV light, the original species were not detected anymore but the identical photomodified oligomers **S30'**-**S34'** were clearly evidenced by MS. Composed of a leading dT monomer and a block exhibiting both monomers with *ortho*-nitrobenzyl motif and the veratryl analogue motif, each original oligomer **S30**-**S34** coded for a letter in the ASCII alphabet. For this purpose, the monomeric units derived from the *ortho*-nitroveratrylbenzyl motif 1 and the *ortho*-nitrobenzyl motif 2 were intentionally defined as 1 and 0 bits, respectively.

Samples were—until very recently—not delivering a sufficient peak intensity to perform tandem mass spectrometry sequencing as the analytes were containing too much of contaminations **#**. Nevertheless, the property A of *erasing sequence information* is exemplary shown in Figure IV-14 for sequence **S33** coding for the letter *s*. Based on the mass differences of coding ortho-nitro (0 bit) and ortho-nitroveratryl monomers (1 bit) the binary code of sequence **S33** can potentially be deciphered by tandem mass spectrometry sequencing when optimized conditions are employed. The ESI(-)-MS spectrum is shown on the bottom of Figure IV-14. After photocontrolled conversion of **S33** to **S33'**, a homooligomeric block is obtained. The side chains comprise identical hydroxyl ends. Hence, the sequence information of the initial sequence is irreversibly lost. Sequence **S33'** was detected together with contaminations # by ESI(-)-MS, also shown on the bottom of Figure IV-14.

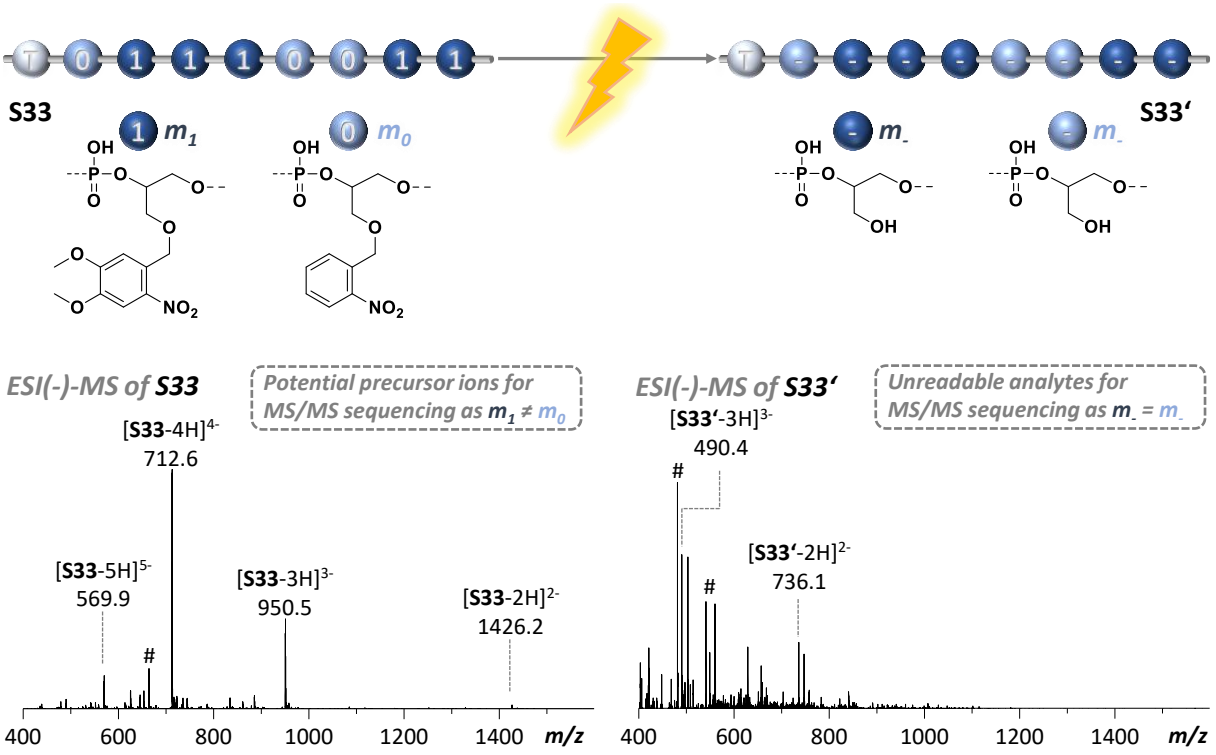


*Figure IV-14: Erasing sequence information from sequence **S33** by quantitative photomodification to obtain* unreadable *sequence **S33'**. ESI(-)-MS spectra of **S33** and **S33'** shown in m/z range 400 – 1600.*

An overview of sequence series **S28**-**S34** (before irradiation) and identic **S28'**-**S34'** (after irradiation) as well as their analysis by ESI(-)-MS is given in Table IV-1.

*Table IV-1: Overview and ESI(-)-MS characterization of sequences **S28**-**S34** synthesized from phosphoramidite monomers **113** and **117**.*

	before irradiation					after irradiation			
	Sequence[a]	ASCII	m/z_{th}[b]	m/z_{exp}		Sequence[a]	ASCII	m/z_{th}[b]	m/z_{exp}
S28	T 11111111	—	1010.5134	1010.5	**S28'**	T - - - - - - - -	—	490.3717	*n. d.*
S29	T 00000000	—	850.4499	850.4	**S29'**	T - - - - - - - -	—	490.3717	490.4
S30	T 01000101	**E**	910.4783	910.5	**S30'**	T - - - - - - - -	—	490.3717	490.4
S31	T 01110010	**r**	930.4853	930.5	**S31'**	T - - - - - - - -	—	490.3717	490.4
S32	T 01100001	**a**	910.4783	910.5	**S32'**	T - - - - - - - -	—	490.3717	490.4
S33	T 01110011	**s**	950.4923	950.5	**S33'**	T - - - - - - - -	—	490.3717	490.4
S34	T 01100101	**e**	930.4853	930.5	**S34'**	T - - - - - - - -	—	490.3717	490.4

[a] The characters 0, 1 and "-" denote for phosphate linked repetitive units derived from diols **111**, **115** and **87**, respectively; [b] $[M-3H]^{3-}$.

2.5 Realization of Property B: Revealing Sequence Information

Analogous to before described concept, phosphoramidite monomers **109** and **113** were used to synthesize oligomeric test samples. Those two oligo(phosphodiester)s contained a leading dT monomer and an octameric block of a repeating monomeric unit derived from one of the monomers. **S35** contained thereby the *para*-nitrobenzyl (motif 3) and **S29** the *ortho*-nitrobenzyl (motif 2) in the side chains of the monomeric units. They are depicted together with their MS analysis in Figure IV-15

below. Both sequences, **S35** and **S29** were detected by ESI(-)-MS in their multiple charged states from -2 to -4 or -5, respectively. Elevated amount of byproducts was not observed. Sequence **S29** has already been studied in the preceding subchapter but is discussed here again for clarity.

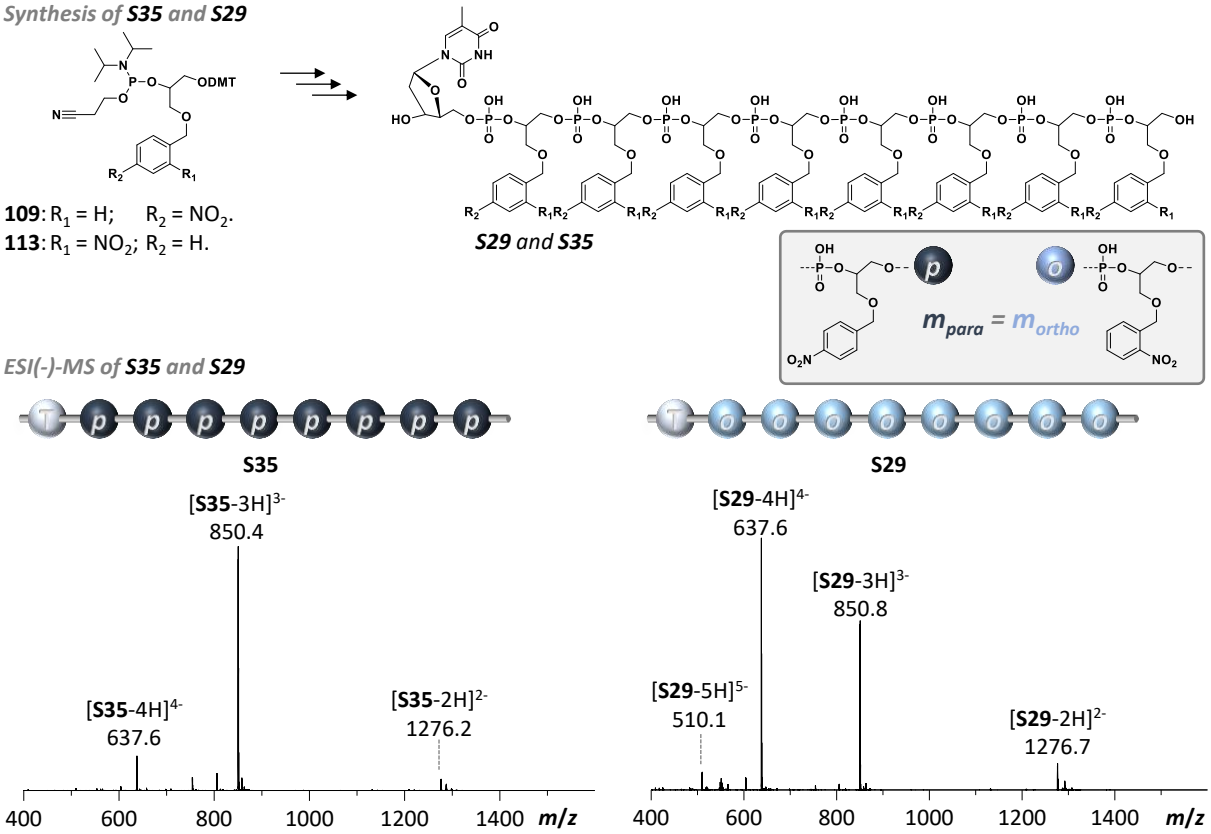


*Figure IV-15: Synthesis of test sequences **S35** and **S29** using nitrobenzyl ether monomers **109** and **113**. (ESI(-)-MS spectra shown in the m/z range 400 – 1600.)*

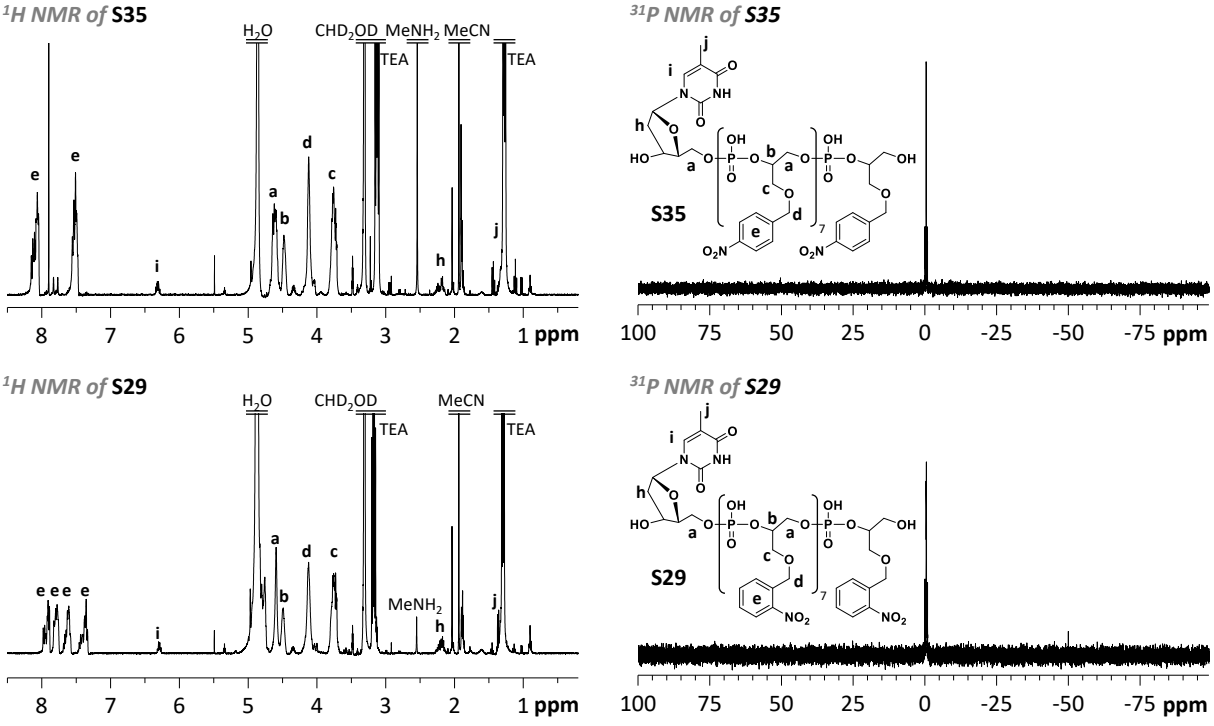


*Figure IV-16: ^{1}H NMR and ^{31}P NMR spectra of sequences **S35** and **S29**, recorded in methanol-d^{4}, shown in the range of 8.5 to 0.0 ppm and 100 to -100 ppm, respectively.*

Results from ^{1}H and ^{31}P NMR analysis of sequences **S35** and **S29** in methanol-d^4 are shown in Figure IV-15. Obviously, differences between the analogue oligomers are observed in the pattern of aromatic protons *(e)* that directly reflect the phenyl substitution. The *para*-nitrobenzyl-containing oligomer **S35** is characterized of two sets of aromatic signals in the ^{1}H NMR spectrum due to the symmetric phenyl substitution. The spectrum of the *ortho*-nitrobenzyl analogue **S29** clearly shows four sets of aromatic protons because the aromatic protons at positions 3, 4, 5, and 6 of the phenyl ring are in chemically different environment. It is also evidenced by NMR that the oligo(phosphodiester)s are isolated as triethylammonium salts. Corresponding signals are marked with *TEA* in Figure IV-15. ^{31}P NMR confirms that phosphate groups reflected by signals around 0 ppm are present.

After successful synthesis of test sequences **S35** and **S29**, their behavior under irradiation with UV light should be investigated. While **S35** is expected not to show a side chain modification, the *ortho*-nitrobenzyl analogue **S29** is known to release the nitrobenzyl group quantitatively from its side chains.

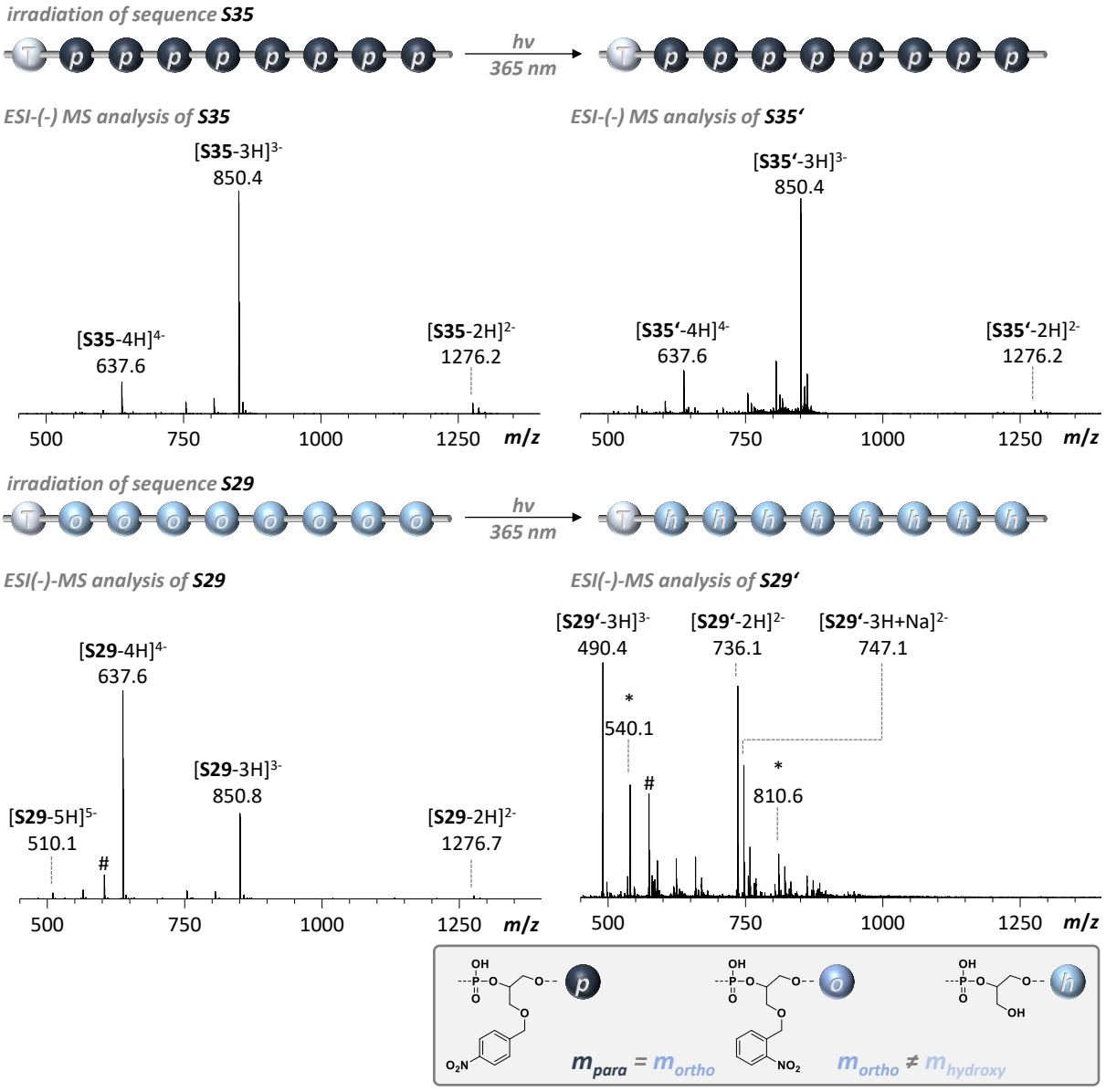


*Figure IV-17: Photomodifiation test of sequences **S35** and **S29**. (ESI(-)-MS spectra shown in the m/z range 450 – 1400.)*

For sake of completeness and to demonstrate this comparative study, analysis of photoirradiation is shown for both sequences **S35** and **S29** in Figure IV-17. It could indeed be confirmed by MS that **S35**

remained intact when triggered with light and no considerable side chain modification was detected. Irradiated **S35'** is thus the same sequence as non-irradiated **S35**.

These very promising results allowed to go on to the implementation of the property B to *reveal sequence information*. A series of information-containing oligo(phosphodiester)s **S36**-**S41** with a leading dT monomer was synthesized from phosphormaidites **109** and **113**. By intentionally defining *ortho*-nitrobenzyl (motif 2) and the *para*-nitrobenzyl ether side chains (motif 3) as **1** (bold) and 1 bits, respectively, each digital oligomers of the series was hiding the binary code (ASCII) of a letter of the alphabet. Analysis by ^{1}H NMR, ^{31}P NMR and MS (see Annex: 6 Analysis of Key Compounds from Chapter IV, pp. 209ff.) showed that sufficiently pure analytes were isolated. Analysis by means of MS/MS gave identical spectra for each sequence of the series **S29**+**S35**-**S41**. It was consequently not possible to distinguish between a **1** bit (bold) and a 1 bit because the underlying monomeric units were regioisomers with the same weight. Yet, it could be shown that the *ortho*-nitrobenzyl groups could be chemoselectively removed from the sequences upon irradiation with light. The modified side chains showed a short hydroxy chain end. These modified monomeric units were defined as 0 bits. Thus, the **1** bit (bold) units were transformed to 0 bit units during irradiation with UV light and the series of sequences **S36'**-**S41'** could be efficiently analyzed by MS/MS to decipher the encoded binary data by determination of the exposed sequence information.

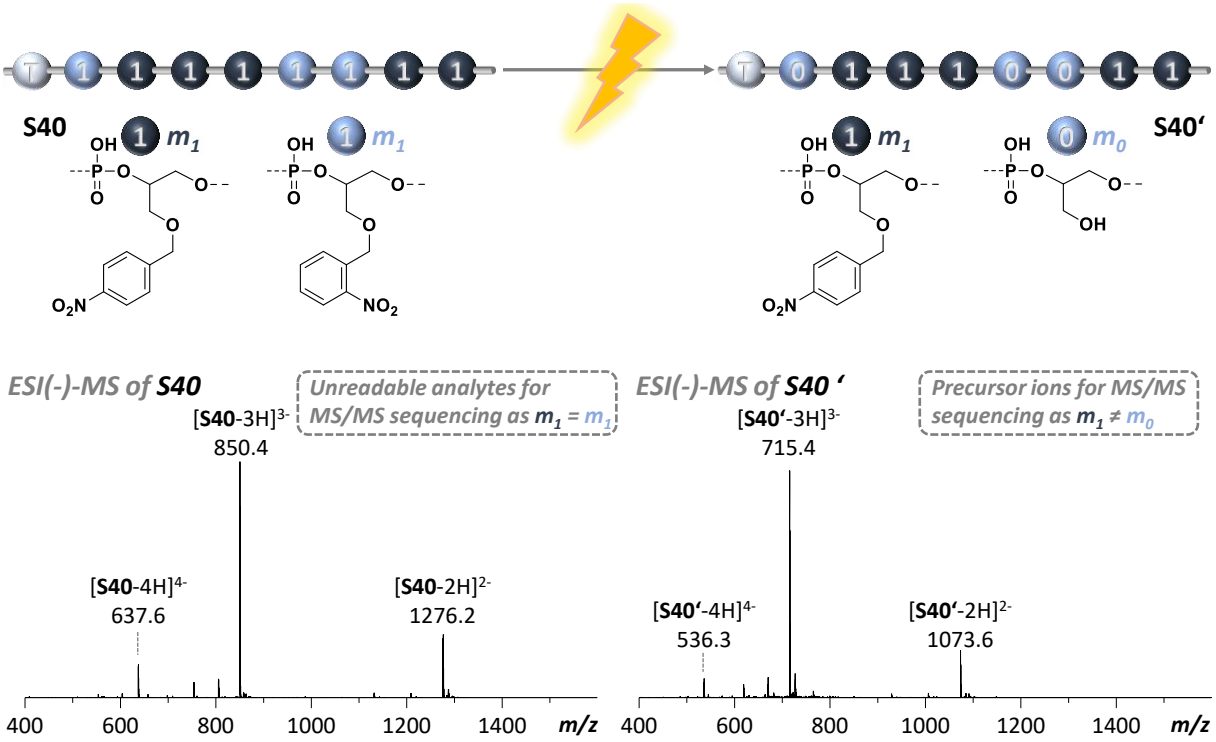


*Figure IV-18: Revealing sequence information from sequence **S40** by chemoselective photomodification to obtain* readable *sequence **S40'**. ESI(-)-MS spectra of **S40** and **S40'** shown in m/z range 400 – 1600.*

The whole procedure is exemplary shown for sequence **S40** in Figure IV-18. On the top of the figure, a cartoon gives an overview of the selective modification of the **1** bit-coding *ortho*-nitrobenzyl units to yield 0 bit-coding hydroxy side chain motifs upon irradiation. The 1 bit-coding *para*-nitrobenzyl units are inert under these conditions and are thus retained. The light-triggered modification can be followed by mass spectrometry as it is shown for **S40** and **S40'** in the center of Figure IV-18. While the MS/MS analysis of the [**S40**-3H]$^{3-}$ peak at 850.4 resulted in the identical and trivial spectrum as for the other entities of the digital oligomer series **S29**+**S35**-**S41**, tandem mass spectrometry sequencing was enabled for modified **S40'**. The respective MS/MS analysis of the [**S40'**-3H]$^{3-}$ peak at 715.4 shown in the Annex in Figure V-21 on page 199 was used to deduce the primary sequence. It could thus be

evidenced that the primary structure of sequences **S40'** was T01110011 which is coding in the ASCII alphabet for the letter *s*. The in-depth analysis of MS/MS fragments and a brief description of tandem mass spectrometry sequencing is given for sequences **S36'**-**S41'** in the Annex: 3.2 (pp. 194ff.).

An overview of sequence series **S29**+**S35**-**S41** (before irradiation) and **S29'**+**S35'**-**S41'** (after irradiation) as well as their analysis by ESI(-)-HRMS is given in Table IV-2.

*Table IV-2: Overview and ESI(-)-HRMS characterization of sequences **S29** and **S35**-**S41** synthesized from phosphoramidite monomers **109** and **113**.*

before irradiation					after irradiation				
	Sequence[a]	ASC II	m/z_{th}[b]	m/z_{exp}		Sequence[a]	ASC II	m/z_{th}[b]	m/z_{exp}
S35	T 11111111	—	850.4499	850.4520	**S35'**	T 11111111	—	850.4499	850.4484
S29	T **11111111**	—	850.4499	850.4[C]	**S29'**	T 00000000	—	490.3717	490.4[c]
S36	T **1**1111**1**1	—	850.4499	850.4504	**S36'**	T 01000101	**E**	625.3965	625.3934
S37	T **1**1111**111**	—	850.4499	850.4527	**S37'**	T 01111000	**x**	670.4072	670.4061
S38	T **1**1111**111**	—	850.4499	850.4589	**S38'**	T 01110000	**p**	625.3965	625.3924
S39	T **1**1111111	—	850.4499	850.4538	**S39'**	T 01101111	**o**	760.4285	760.4254
S40	T **1**1111**1**11	—	850.4499	850.4537	**S40'**	T 01110011	**s**	715.4178	715.4148
S41	T **1**1111**1**11	—	850.4499	850.4522	**S41'**	T 01100101	**e**	670.4072	670.4073

[a] The numbers 1, **1** (bold) and 0 denote for phosphate linked repetitive units derived from diols **107**, **111** and **85**, respectively; [b] $[M-3H]^{3}$ measured in their respective MS/MS spectra using an internal fragment as an internal standard; [c] determined by ESI(-)-MS (not high-resolution MS).

3 Conclusion and Outlook

A new conceptual design of information-containing oligo(phosphodiester)s was described in the present chapter. A photoinduced side chain modification was used to either erase the sequence information from a polymer backbone or to reveal a molecularly encoded message with a light stimulus. The design of the approach was based on tandem mass spectrometry sequencing to retrieve the information that was stored in the primary sequence of digital oligo(phosphodiester)s. Sequences were containing different nitrobenzyl moieties in their side chains and could be analyzed by MS/MS before and after the photorelease of defined side chain motifs.

In several studies—also including nitrobenzyl-substituted amines and carbamates—it was found that nitrobenzylic ethers are the better-suited motifs to be insert on side chains of monomeric units. Such nitrobenzylic ether phosphoramidite monomers could be efficiently synthesized and could be used to yield pure sequence-defined oligo(phosphodiester)s. The obtained digital oligo(phosphodiester)s showed a chemoselective photorelease of different *ortho*-nitrobenzyl groups which was the fundamental basis of the described concepts.

It was reported that digital oligo(phosphodiester)s being built from analogue *ortho*-nitrobenzyl and *ortho*-nitroveratrylbenzyl-containing monomeric units—defined as bit 0 and bit 1, respectively—could be used as temporary molecular data storage. Their sequence information could be erased upon irradiation. Thus, an analyte being composed of two different coding units was sequencable by means of tandem mass spectrometry. Yet, after modification of all side chains to identical hydroxy motifs, the analyte loses its sequence information and the information read-out becomes trivial. This concept was demonstrated with a series of sequences **S30**-**S34** each coding for a letter in the ASCII alphabet. The

results are, however, not satisfying until now because the efficient analysis by mass spectrometry is hindered by abundant contaminations.

Very good results were reported for the implementation of a concept to secretly transmit information via digital oligo(phosphodiester)s. Their sequence information is to be exposed by irradiation with light. Hence, an unreadbale analyte becomes sequencable after triggering it with UV light. This feature is achieved by synthesizing information-containing oligo(phosphodiester)s from regioisomeric *ortho*-nitrobenzyl and *para*-nitrobenzyl-containing comonomers. Both monomeric units are thus characterized by the same weight and cannot be distinguished by analysis using mass spectrometry. The hidden sequence information can, however, be exposed when chemoselectively releasing the *ortho*-nitrobenzyl motifs yielding hydroxy-terminated side chains at these positions. As a result, the monomeric units—being hydroxy side chains and *para*-nitrobenzylic ether side chains, intentionally defined as bit 0 and bit 1, respectively—are distinguishable by their different weight after photomodification. MS/MS analysis clearly evidenced that a non-readable analyte could be rendered MS/MS-readable upon irradiation with UV light. The concept was demonstrated with a series of sequences **S36**-**S41** each coding for a letter in the ASCII alphabet.

Hiding and erasing information does in each case imply that one of the two forms of a digital polymer (before modification or after) is *filled* with trivial information like a repeating 1 bit. Future works could aim for a simple system using the light-trigger and another orthogonal reaction relying on four comonomers A, B, C, and D. If all parameters are well designed, it should be possible to have monomeric units A and D, that persistently code for bit 0 and bit 1, respectively. Comonomer B converts from a bit 0 to a bit 1 upon treatment with trigger I. Trigger II converts selectively comonomer C from bit 1 to a bit 0. In this way, arbitrary binary information could be stored in the original and in the modified digital polymer. In the context of espionage, one of the two states could thereby serve as a jammer or contain the encryption key for the other state.

General Conclusion

Abiotic sequence-defined oligo- and poly(phosphodiester)s have been extensively investigated in the frame of the doctoral studies. The research was situated in precision macromolecular chemistry, a relatively young discipline within polymer science. Methods and tools of neighboring fields could be employed to develop the research area of digital polymers further.

Digital polymers are defined macromolecules whose sequence exhibits a code that can be processed to digital information. A sequence-defined copolymer—composed of two different comonomers that are intentionally defined as 0 bit and 1 bit—for instance, is an information-containing polymer. Its primary sequence can thus be interpreted as a sequence of 0 and 1 bits. That is a bitstream and the prerequisite for digital data storage and transmission. A very efficient, mild and user-friendly strategy to obtain such sequence-defined polymers relatively fast and in high purity is phosphoramidite chemistry. It was originally conceived by biochemists to synthesize long oligonucleotides with high fidelity. Phosphoramidite chemistry was consequently adapted to the tool box of polymer chemistry and new phosphoramidite monomers were polymerized with ultimate precision.

The latter monomers are freely select abiotic motifs that can be precisely designed to meet required criteria. In the context of digital polymers, poly(phosphodiester)s were envisioned that have implemented new beneficial features for data storage.

The study presented in **Chapter II** (pages 61ff.) shows how the side chain information in digital poly(phosphodiester)s can be controlled. A binary modification protocol was developed to precisely insert two sets of arbitrary chosen side chains into a defined poly(phosphodiester) precursor sequence. The latter was composed of two information-coding comonomers that bear terminal alkyne units and TIPS-protected alkyne units, respectively. By performing a first CuAAC post-polymerization modification with an organic azide, the monomers characterized by the terminal alkyne moieties can be quantitatively converted to side chain-attached 1,4-triazole moieties. The process is chemoselective because the TIPS-protected alkyne functions are inert under the conditions. After removal of said protective groups from the remaining alkynes, a second CuAAC modification can be performed to attach the other side chain of choice to the monomeric units that have not been modified before. When following this new methodology, diverse binary modified poly(phosphodiester)s can be obtained.

The efficient synthesis of the phosphoramidite monomers required for precursor synthesis was first investigated and the optimized synthetic strategy reported. Followingly, the manual synthesis of phosphodiester tetramers and octamers on self-made solid polystyrene supports was studied and means were found to suppress the undesired alkyne iodination side reaction. Oligomers could then be sequentially modified with a short methoxy-PEG chain and a trifluoroacetamide-functionalized PEG chain. All intermediates that occurred during the process could be characterized by NMR and mass spectrometry. When longer poly(phosphodiester)s were targeted, the precursor synthesis was performed on an automated oligonucleotide synthesizer using a commercial CPG support. The obtained precursor sequences were in fact modified in the first step when still bound to the support. So, the purification by washing the support prior to cleaving the modified sequence from the carrier is rendering the reaction very efficient. The target binary modified poly(phosphodiester) was then obtained by TIPS deprotection and a final CuAAC modification in solution. Also here, the intermediates were analyzed by NMR and mass spectrometry.

In a collaborative work with experts in nanopore sequencing it was examined how abiotic poly(phosphodiester)s interact with several transmembrane proteins. It was further investigated in which way digital poly(phosphodiester)s could be designed—in particular by screening several side

chain motifs benefiting from the described post-polymerization modification methodology—to ultimately yield an analyte whose sequence-information could be deciphered efficiently by using nanopores.

When storing large amounts of data on the molecular level, it is of outstanding importance that the information-containing molecules can be sequenced with high fidelity and in relatively short time. Nanopore sequencing meets this criterion. Hence, making digital polymers readable by this technique would have a tremendous impact on this research area.

Besides improvements in polymer sequencing it is of equal importance to develop schemes organizing molecularly stored data. For obvious reasons like a limited degree of polymerization, many digital polymers would have to be organized in libraries, where each polymeric entity would contain a fraction of the archived bitstream. A way to organize large amounts of non-identical polymers at defined locations is the fabrication of digital microarrays. Microarrays are 2D objects below the mm scale that contain hundreds of thousand defined positions or spots who each express a group of defined substrate-bound polymeric species. A digital polymer microarray would consequently contain on each of the microspots a fraction of the data that is stored in the library. Microarrays can be synthesized by photolithographic means benefitting from the chemical trigger light. The latter can be addressed to distinct spots on the array surface with high temporal and spatial resolution.

In this context a preliminary study was executed. It was thus the task to develop a synthetic strategy that allows the light-dependent synthesis of digital poly(phosphodiester)s. Inspired by the state-of-the-art in DNA microarray fabrication, an adapted photocontrolled phosphoramidite strategy was employed for the synthesis of abiotic sequence-defined oligo(phosphodiester)s. It is described in **Chapter III** (pages 89ff.) how simple coding monomers could be synthesized that were equipped with a NPPOC photoreleasable protective group. Instead of using the standard acid-labile DMT protective group, this protection strategy is a prerequisite for the photocontrolled polymerization. While phosphoramidite coupling and oxidation would be performed under the commonly used conditions, the liberation of a terminal hydroxy function on the growing chain end would occur by a terminal NPPOC photorelease.

When the oligomer synthesis was performed on a polystyrene support, it was evidenced that the photodeprotection step remained incomplete even when a high intensity UV source was used to facilitate NPPOC cleavage. The support material is possibly absorbing too strongly the UV light. An inefficient photodeprotection step does imperatively yield samples with a high degree of dispersity which is not acceptable for data storage application. The use of a soluble linear ATRP-made polystyrene support was found to allow for a quantitative photodeprotection of NPPOC groups. The underlying optimization of this step was greatly facilitated as the soluble conjugate polystyrene-*b*-oligo(phosphodiester) system could be analyzed by UV, NMR and SEC at each stage of reaction. After each reaction step (coupling/oxidation or photodeprotection), the immobilized growing chain was purified from excess reagents by precipitation in a polar non-solvent. It was found that the higher the reaction scale the better recovery yields were obtained after precipitation. Working above a 100 mg or above a 500 mg scale allowed for a stepwise yield above 90 % or above 95 %, respectively. The oligomers were obtained after cleavage from the soluble support with high purity confirming that abiotic sequence-defined oligo(phosphodiester)s could be synthesized using a photocontrolled phosphoramidite strategy. It was furthermore shown that classic DMT-protected phosphoramidite monomers coupled to the growing oligomer as the last monomers enabled a DMT-ON purification of oligo(phosphodiester)s. That perspective could be beneficial when longer sequences should be synthesized.

Future investigations will show whether the abiotic oligo(phosphodiester)s could be synthesized in parallel on a microarray using a maskless array synthesizer. The fabrication of those digital polymer

microarrays would be a milestone as it would efficiently organize digital polymer libraries with high information density.

A new concept is demonstrated in **Chapter IV** (pages 103ff.). It allows to either hide information in a dormant state of a digital poly(phosphodiester)s. Or the concept is applied to provide freely readable information in the original analyte that can easily be erased to render the molecular data storage unreadable. This project showed that a wide chemical space is available when envisioning new digital polymers and that very interesting features can thus be induced by a sophisticated monomer and polymer design.

The read-out should be achieved for the here-described binary-coded oligo(phosphodiester)s using tandem mass spectrometry sequencing. The distinction between a 0 bit and a 1 bit monomer is consequently based on their weight. When information should be hidden in the primary structure of a sequence, regioisomers were chosen as coding monomer units. Hence, tandem mass spectrometry sequencing cannot discriminate between the two coding units. Yet, the regioisomers were chosen to be orthogonally addressable by a certain stimulus. Here, it was shown that an *ortho*-nitrobenzyl moiety could be cleaved from one of the coding monomers by irradiation with UV light while the analogue *para*-nitrobenzyl moiety stayed intact. The digital oligomer would contain two monomeric units with a distinguishable weight and the encoded information could be exposed by tandem mass spectrometry sequencing. When using as an alternative to the inert *para*-nitrobenzyl motif a comonomer with an *ortho*-nitroveratrylbenzyl motif, the binary coded oligomer constitutes a MS-readable analyte. After photoremoval of both the *ortho*-nitro and the *ortho*-nitroveratrylbenzyl groups, the previously different 0 bit and 1 bit monomers are converted to identical entities. After the treatment with light, the analyte has lost the encoded sequence information. The encoded information is thus erased.

It has been shown that the series of nitrobenzyl groups could be attached to monomers at nitrogen atoms or analogous to oxygen atoms in their side chains. Consequently, a series of *para*-nitro-, *ortho*-nitro-, and *ortho*-nitroveratrylbenzylamines and corresponding ethers was under investigation. The nitrobenzylamine phosphoramidite monomers had to bear an additional amine protective group to allow their use in phosphoramidite chemistry. Nevertheless, impure sequences were obtained in a first attempt and when studying the photorelease in smaller test systems, a complex photochemistry was observed. Optimization of both, oligomer synthesis and a clean photoremoval of nitrobenzyl groups was believed to be feasible. Yet, the series of nitrobenzyl ethers showed excellent purity for oligomers and the treatment with light followed the desired path. Finally, the implementation of the conceptional approach of erasing sequence information by light irradiation was shown for a series of one byte-coding oligomers. Exposure of before hidden information was shown for an equivalent series of digital oligomers

Overall, new perspectives for sequence-defined poly(phosphodiester)s were described in chapter II, III, and IV. Recapitulating the research projects, it is demonstrated that this class of polymers is very versatile and makes a promising candidate for molecular data storage. Digital poly(phosphodiester)s can be synthesized efficiently in a process that tolerates manifold structural motifs. As a result, the information-containing polymers can be tailor-made for their very purpose. It can hence be reasonably anticipated that future demands for specialty application in molecular data storage will be satisfied by tailor-made poly(phosphodiester)s.

Experimental Section

1 Materials and Methods

1.1 Materials and Reagents

1.1.1 Materials and Reagents of Commercial Source

Acetic acid (99.8 %, Fischer Chemical), acetic anhydride (>99 %, Alfa Aesar), acetone (Carlo Erba), acetonitrile (99.8 %, MeCN, Carlo Erba), anhydrous acetonitrile (99.8 %, anh. MeCN, Sigma-Aldrich), allyl bromide (99 %, stab. with 300 - 1000 ppm propylene oxide), aminomethylated polystyrene (HL, 100-200 mesh, Merck/Novabiochem), aqueous ammonia (28 w% NH_3 in H_2O, VWR), ammonium chloride (99.5 %, VWR), 11-azido-3,6,9-trioxaundecan-1-amine (93 %, TCI), benzoic anhydride (98 %, Alfa Aesar), benzyl bromide (98 %, Sigma-Aldrich), bromoacetyl bromide (98 %, Alfa Aesar), 4-(bromomethyl)benzoic acid (97 %, Alfa Aesar), 11-bromoundecanoic acid (>97 %, TCI), 1-bromo-3-propanol (97 %, Sigma-Aldrich), tetra-*n*-butylammonium fluoride (1 M in THF, TBAF, Aldrich), *n*-butyllithium (2.5 M in hexane, *n*-BuLi, Aldrich), *tert*-butyl alcohol (extra pure, *t*BuOH, Fisher Scientific/Acros Organics), carbon tetrabromide (99 %, TCI), chloroform (99.8 %, UV/IR-Grade, Roth), chloroform (Carlo Erba), 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (97 %, ABCR), cyclohexane (Carlo Erba), dichloromethane (99.95 %, DCM, cont. amylene as stabilizer, Carlo Erba), anhydrous dichloromethane (99.8 %, anh. DCM, cont. amylene as stabilizer, Sigma-Aldrich), *N*,*N*'-dicyclohexylcarbodiimide (99 %, DCC, Alfa Aesar), *N*,*N*-diisopropylethylamine (>99.0 %, DIPEA, TCI), diethyl malonate (99 %, Alfa Aesar), diethylether (Carlo Erba), anhydrous diethyl ether (99.7 %, 1 ppm BHT as inhibitor, anh. Et_2O, Sigma-Aldrich), 3,4-dihydro-2*H*-pyran (97 %, Aldrich), 4,4'-dimethoxytrityl chloride (DMTrCl, ChemGenes), 4-(dimethylamino)pyridine (99 %, DMAP, Aldrich), DMSO (99.6 %, Aldrich), anhydrous dimethyl sulfoxide (99.9 %, anh. DMSO, Alfa Aesar), 2,2-dimethyl-1,3-propanediol (99 %, Alfa Aesar), 1,4-dioxane (99.5 %, Alfa Aesar), diphenyl phosphoryl azide (97 %, Aldrich), *n*-dodecylamine (99 %, Aldrich), anhydrous ethylene glycol (99.8 %, AcroSeal, Acros Organics), ethyl acetate (EtOAc, Carlo Erba), fluorescein 5-isothiocyanate (95 %, isomer I, TCI Europe), 9-fluorenylmethoxycarbonyl chloride (98 %, Fmoc-Cl, Roth), glycerol (1,2,3-propanetriol, cont. water, Merck), iodine (99 %, Prolabo), *L*-ascorbic acid (99 %, Sigma-Aldrich), *L*-ascorbic acid sodium salt (99 %, Alfa Aesar), lithium aluminum hydride (2.0 M in THF, Aldrich), 2,6-lutidine (98 %, Alfa Aesar), methanesulfonyl chloride (98 %, MsCl, Alfa Aesar), methanol (MeOH, 99.9 %, Carlo Erba), anhydrous methanol (99.8 %, anh. MeOH, Sigma-Aldrich), aqueous methylamine (40 w% $MeNH_2$ in H_2O, Fluka), methyllithium (1.6 M in Et_2O, Sigma-Aldrich), molecular sieves (3 Å, Merck), 2-nitrobenzaldehyde (98 %, Alfa Aesar), 4-nitrobenzaldehyde (98 %, Alfa Aesar), 2-nitrobenzyl bromide (98 %, Alfa Aesar), 4-nitrobenzyl bromide (97 %, Alfa Aesar), 2-(2-nitrophenyl)propyl chloroformate (95 % chlorine basis, NPPOC chloride, Sigma), 2-nitroveratraldehyde (98 %, Alfa Aesar), *para*-toluenesulfonic acid monohydrate (99 %, Sigma-Aldrich), *para*-toluenesulfonyl chloride (98 %, TsCl, Alfa Aesar), phosphorus tribromide (99 %, Sigma-Aldrich), piperidine (99 %, Alfa Aesar), potassium dihydrogen phosphate (99.9 %, KH_2PO_4, VWR), 1,3-propanediol (99 %, Alfa Aesar), propargyl bromide (80 w% in toluene, stabilized with MgO, Alfa Aesar), anhydrous pyridine (99.8 %, anh. pyr, in Sure/Seal™, Sigma-Aldrich), *R*-glycidol (98 %, TCI), serinol (98 %, 2-amino-1,3-propanediol, TCI), silica gel (high-purity grade, pore size 60 Å, 230 - 400 mesh, for flash chromatography, SiO_2, Fluka), sodium azide (99 %, Alfa Aesar), sodium bicarbonate (99 %, SDS), sodium borohydride (98 %, Aldrich), sodium chloride (for brine solution, ESCO), sodium hydride (60 w% suspension in mineral oil, Aldrich), anhydrous sodium sulfate (99.6 %, VWR), sodium thiosulfate (Fisher Scientific), succinic anhydride (99 %, Sigma-Aldrich), 1*H*-tetrazole (~0.45 M in MeCN, Sigma), titanium tetrachloride (98 %, Fluka), toluene (Carlo Erba), anhydrous

toluene (99.8 %, anh. toluene, Sigma-Aldrich), trichloroacetic acid (99 %, TCA, Sigma-Aldrich), triethylamine (97 %, TEA, Acros Organics), triethylene glycol monomethyl ether (95 %, Sigma-Aldrich) and triisopropylsilyl chloride (97 %, TIPS-Cl, Aldrich) were used as received without further purification.

For synthesis using an oligonucleotide synthesizer anhydrous acetonitrile (phosphoramidite diluent & dry washings, ChemGenes), acetonitrile (≥99.9 %, washings, Roth), activation reagent (0.25 M 5-ethylthio tetrazole in MeCN, ChemGenes), Cap A (acetic anhydride/pyridine/THF, ChemGenes), Cap B (10 % *N*-methylimidazole in THF, ChemGenes), DMT removal reagent (3 w% TCA in DCM, Roth), drying traps (small, 10 - 15 mL, ChemGenes), dT-CPG 1000 (1 µmol in cartridge, 1000 Å pore size, Glen Research), oxidizer (0.02 M iodine/pyridine/H_2O/THF, ChemGenes) were used as received.

1.1.2 Reagents Synthesized In-House

Copper (I) bromide (98 %, CuBr, Alfa Aesar) was purified by stirring in acetic acid and rinsed with ethanol and diethyl ether, then dried. The synthesis of the TBTA ligand was adapted from the literature.[377] A 0.78 M solution of fluoride quencher methoxytrimethylsilane in diethylether was prepared by reacting methanol with trimethylsilyl chloride in diethylether and in the presence of triethylamine and subsequent filtration. Synthesis of hydroxyl-terminated soluble support **57** was described earlier.[349] DMT-protected glycerol chloride **95** was synthesized by stepwise addition of DMTr chloride to a solution of **93** in anh. pyridine and anh. THF.

1.2 Measurement & Analysis, Laboratory Equipment, General Methods

1.2.1 Measurement & Analysis

1.2.1.1 Nuclear Magnetic Resonance (NMR)

All NMR spectra were recorded on a Bruker Avance 400 spectrometer equipped with an Ultrashield magnet. Chemical shifts (δ) are reported in parts per million (ppm) against solvent residual signal (^{1}H NMR $CDCl_3$: δ = 7.26 ppm; ^{1}H NMR CD_3OD: δ = 3.31 ppm; ^{1}H NMR $(CD_3)_2CO$: 2.05 ppm; ^{13}C NMR, $CDCl_3$: δ = 77.16 ppm; ^{13}C NMR CD_3OD: δ = 49.00 ppm; ^{13}C NMR $(CD_3)_2CO$: δ = 29.84 ppm, 206.26 ppm). ^{1}H NMR spectra were recorded at 400.13 MHz, ^{13}C NMR spectra at 100.62 MHz. ^{19}F NMR spectra were recorded at 376.46 MHz and were externally referenced to trichlorofluoromethane. The NMR solvents deuterated acetone (99.8 %, acetone-d^6), deuterated chloroform (99,8 %, chloroform-d^1), deuterated methanol (99,8 %, methanol-d^4) were purchased from Aldrich.

1.2.1.2 Electrospray Ionization Mass Spectrometry (ESI-MS)

High resolution MS and MS/MS experiments were performed using a QqTOF mass spectrometer (QStar Elite, Applied Biosystems SCIEX, Concord, ON, Canada) equipped with an ESI source operated in the negative ion mode (capillary voltage: –4200 V; cone voltage: –75 V). Accurate mass measurements in the MS mode were most performed in the positive ion mode due to the limited availability of standards that can be used for proper internal calibration, in the negative mode, of the orthogonal acceleration time-of-flight (oa-TOF) mass analyzer. MS/MS experiments were conducted in collision-induced dissociation conditions, where precursor ions were selected in a quadrupole mass analyzer prior entering a collision cell filled with nitrogen, and products ions were measured in the oa-TOF. In this instrument, air was used as nebulizing gas (10 psi) while nitrogen was used as curtain gas (20 psi). Instrument control, data acquisition and data processing were achieved using Analyst software (QS 2.0) provided by Applied Biosystems.

Each polymer/oligomer (1-2 mg) was dissolved in 300 µL of water/acetonitrile (50/50, v/v), and further diluted (1/102 to 1/103, v/v) in a methanolic solution of ammonium acetate (3 mM) prior to being injected in the ESI source at a 10 $\mu L \cdot min^{-1}$ flow rate using a syringe pump.

1.2.1.3 Size Exclusion Chromatography (SEC)

Analytes were solubilized in THF and filtered (0.45 µm filter, Captiva by Agilent). SEC analysis was performed at 30 °C in THF (flow rate: 1 $mL \cdot min^{-1}$; flow rate marker: toluene) using a Shimadzu LC20-AD pump and a Shimadzu CT0-10AC column oven. The set-up was equipped with four monoporosity columns (granulometry: 5 mm), PL Gel 30 cm (internal diameter = 7.5 mm) of 50 Å, 100 Å, 500 Å, and 1000 Å (separation range: 100 – 20000 $g \cdot mol^{-1}$) and polymers were detected with a Shimadzu RID 10A refractometer and a Shimadzu SPD-10Avp UV detector. PS calibration was performed with 10 linear polystyrene standards (Polymer Laboratories) with a range of 162 - 20650 $g \cdot mol^{-1}$.

1.2.1.4 UV-Vis Spectrophotometry

UV-Vis spectra were recorded on a Perkin Elmer Lambda 25 UV-Vis spectrometer using the Perkin Elmer UV WinLab software.

1.2.2 Laboratory Equipment

1.2.2.1 Shaker

An IKA KS 130 basic shaker was used to agitate polystyrene supports in reagents solutions in a home-made glass reactor.

1.2.2.2 Centrifuge

An ExtraGene mini centrifuge, Model: 6K, was used to separate small amounts of solids from suspensions in Eppendorf tubes.

1.2.2.3 UV Lamps for Sample Irradiation

Samples were irradiated using a lamp by Vilber Lourmat (VL-115.L, 2x 15W, I = 1100 $\mu W \cdot cm^{-2}$ at 15 cm distance to surface, λ = 365 nm).

For higher intensity two HPK 125 W high-pressure mercury vapor lamps from Heraeus Noblelight (maximum energy at 365 nm, substantial radiation also at 435, 404, 313, and 253 nm; continuum 200 – 600 nm peaking at 260 nm) were used in an aluminum housing under air ventilation.

1.2.2.4 Oligonucleotide Synthesizer

An Expedite DNA synthesizer (Perseptive Biosystem 8900) was employed.

1.2.3 General Methods

1.2.3.1 Automated Poly(phosphodiester) Synthesis

All needed commercial reagents (see Experimental Section: 1.1.1 Materials and Reagents of Commercial Source, pp. 133f.) were used as received and installed in the Expedite DNA synthesizer. The synthesized phosphoramidite monomers were charged to the synthesizer as 100 mM solutions in anhydrous MeCN under argon counter flow. The reservoirs were equipped with a drying trap. Before the start of each automated synthesis, all reagents were primed twice and argon was flushed through

the utilized supply line. Then, a solid support-filled column (dT-CPG, 1 µmol scale) was connected to the synthesizer.

The method for the iterative phosphoramidite cycles deviated from standard protocols by prolonged capping steps (2x 90 s). The command lines that were applied are followingly given with comments between quotation marks or after an asterisk. The first number in each line defines the reagent positions (or commands for UV detection), the second defines the number of pump pulses that deliver reagents to the column (CPG support cartridge) and the third number defines in which time frame (in seconds) these pulses should be performed. In the case of a "0" the pumps work at maximum speed. The protocol reveals that the dead volume between reagent pump and the column is the equivalent of seven pulses (each pulse comprises a volume of 16 µL). The protocol is shown exemplary for the coupling of a monomer at position 8 (25):

$Deblocking

144	/*Index Fract. Coll.	*/ NA	1	0	"Event out ON"
0	/*Default	*/ WAIT	0	1.5	"Wait"
141	/*Trityl Mon. On/Off	*/ NA	1	1	"START data collection"
16	/*Dblk	*/ PULSE	10	0	"Dblk to column"
16	/*Dblk	*/ PULSE	50	49	"Deblock"
38	/*Diverted Wsh A	*/ PULSE	80	0	"Flush system with Wsh A"
141	/*Trityl Mon. On/Off	*/ NA	0	1	"STOP data collection"
38	/*Diverted Wsh A	*/ PULSE	80	0	"Flush system with Wsh A"
144	/*Index Fract. Coll.	*/ NA	2	0	"Event out OFF"

$Coupling

1	/*Wsh	*/ PULSE	5	0	"Flush system with Wsh"
2	/*Act	*/ PULSE	5	0	"Flush system with Act"
25	/*8 + Act	*/ PULSE	6	0	"Monomer + Act to column"
25	/*8 + Act	*/ PULSE	1	8	"Couple monomer"
2	/*Act	*/ PULSE	4	32	"Couple monomer"
1	/*Wsh	*/ PULSE	7	56	"Couple monomer"
1	/*Wsh	*/ PULSE	8	0	"Flush system with Wsh"

$Capping

12	/*Wsh A	*/ PULSE	20	0	"Flush system with Wsh A"
13	/*Caps	*/ PULSE	8	0	"Caps to column"
12	/*Wsh A	*/ PULSE	6	90	"Cap"
12	/*Wsh A	*/ PULSE	14	0	"Flush system with Wsh A"

$Oxidizing

15	/*Ox	*/ PULSE	15	0	"Ox to column"
12	/*Wsh A	*/ PULSE	60	0	"Flush system with Wsh A"

$Capping

13	/*Caps	*/ PULSE	7	0	"Caps to column"
12	/*Wsh A	*/ PULSE	6	90	"Cap"
12	/*Wsh A	*/ PULSE	30	0	"End of cycle wash"

A synthesis can be performed in DMT-ON mode. Thus, the synthesis stops after the last monomer addition with its last capping step and finally a wash with MeCN. Consequently, the desired sequences exhibit a DMT-protected terminus. In the DMT-OFF mode, an additional DMT deprotection step ($Deblocking) is added after the final monomer coupling to result in hydroxy-terminated sequences on the CPG support.

The synthesized polymers were cleaved from the CPG support by passing 0.5 mL of a solution of ammonia (aq., 28 w%) and methylamine (aq., 40 w%) (1 : 1, v : v) for 30 min between two syringes that are installed at the ends of the support cartridge. After recovering the solution another 0.5 mL of said ammonia/methylamine (AMA) solution was passed through the cartridge and the unified poly(phosphodiester)-containing solutions were removed from solvent under reduced pressure to obtain the desired polymers after a DMT-OFF synthesis. After a DMT-ON synthesis the latter solution was entering the DMT-ON purification methodology described below.

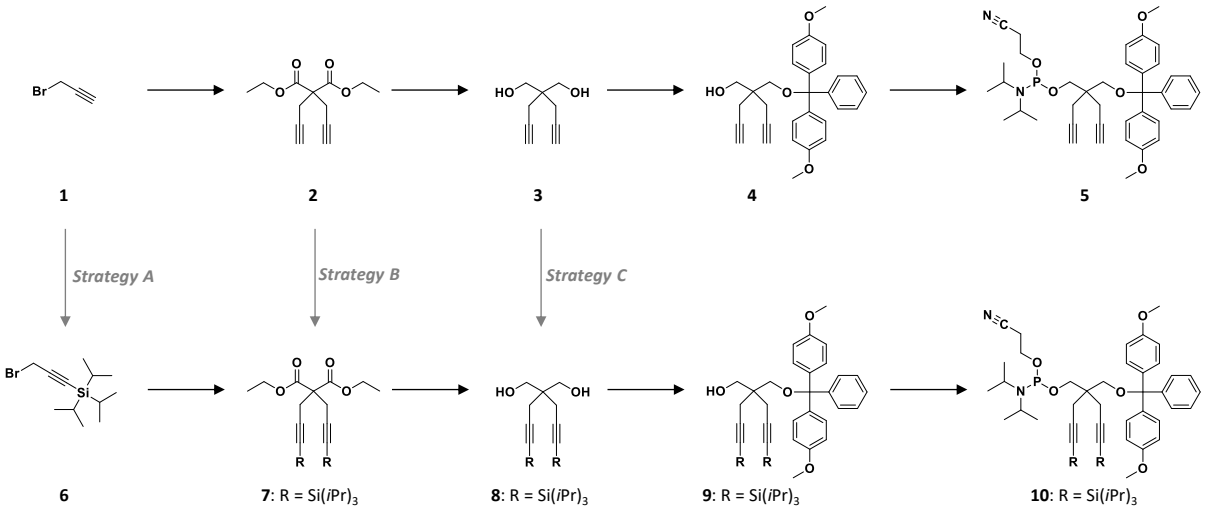
1.2.3.2 Manual DMT-ON Purification

A reverse phase chromatography cartridge (DNA purification kit) was flushed with 0.5 mL MeCN and 1 mL of a 2 M aqueous triethylammonium acetate buffer solution Then, the poly(phosphodiester)-containing AMA solution was mixed with 1 mL of a NaCl solution (aq., 10 w%) and applied to prepared purification device. The cartridge was washed with 1 mL of a salt wash solution (5 v% MeCN in aq., 10 w% NaCl solution) to wash the failure sequences from the cartridge. The cartridge was then washed with 2 mL of 2 w% TCA in H_2O to remove the last terminal DMT group and it was further washed with 2 mL H_2O to remove the acid. The polymer was then washed out with 1 mL of a 1 : 1 (v : v)-mixture of MeCN and H_2O containing an additional 0.5 v% ammonia (aq., 28 w%). The solution was thereby collected in several fractions. The first six drops did usually not contain the desired polymer, however, all fractions showing UV absorption were dried under reduced pressure and analyzed. The principle of DMT-ON purification is explained in detail in the Annex, subchapter 1.1 DMT-ON Purification on page 185.

2 Experimental Protocols Chapter II

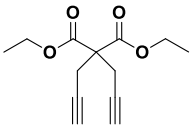
2.1 Monomer Synthesis

The synthesis of monomers **5** and **10** is described herein. The synthetic route following strategy A is first described as the recommended strategy (see Scheme V-1). Synthesis of **7** from **2** and **8** from **3** (via **11**) is detailed afterwards.



Scheme V-1: Overview synthetic routes of monomer synthesis in Chapter II

2.1.1 Malonate 2

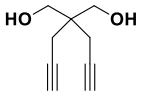


Under argon atmosphere, sodium hydride (4.822 g, 60 wt% in mineral oil, 121 mmol, 2.02 eq.) was suspended in 200 mL THF (anh.). Diethylmalonate (9.556 g, 60 mmol, 1.0 eq.) was added dropwise to the ice-cold reaction mixture that was afterwards stirred for 2 h and allowed to warm to RT. Then, propargylbromide (22.650 g, 80 w% in toluene, stabilized with MgO, 152 mmol, 2.5 eq.) was added dropwise at 0 °C and the reaction mixture was stirred overnight and allowed to reach RT. The reaction mixture was quenched with NH_4Cl (aq., sat.), THF was removed. The slurry was taken in 100 mL EtOAc and 100 mL NH_4Cl (aq., sat.) were added and it was further extracted with EtOAc (3x 150 mL). The combined organic layer was washed with brine (3x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The crude yellow oil was purified by column chromatography (SiO_2 vs 10 v% EtOAc in cyclohexane) to obtain 14.192 g (60 mmol, full conversion) of the product as colorless crystals.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.23 (t, ^3J = 7.1 Hz, 6H, $CH_2\underline{CH_3}$), 2.02 (t, ^4J = 2.7 Hz, 2H, $CH_2C{\equiv}\underline{CH}$), 2.99 (d, ^4J = 2.7 Hz, 4H, $\underline{CH_2}C{\equiv}CH$), 4.23 (q, ^3J = 7.1 Hz, 4H, $\underline{CH_2}CH_3$).

The analysis data harmonizes with the reported values in the literature.[328]

2.1.2 Diol 3

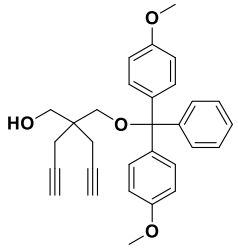


Under argon atmosphere, lithium aluminum hydride solution (15 mL, 2 M in THF, 30 mmol, 1.2 eq.) was diluted in 120 mL THF (anh.). At -10 °C a solution of malonate **2** (6.013 g, 25.5 mmol, 1.0 eq.) in 20 mL THF (anh.) was added dropwise to the solution. The reaction mixture was stirred overnight and allowed to reach RT. At 0 °C the reaction was quenched with 10.0 mL brine. The slurry was taken in 50 mL Et_2O and 50 mL NH_4Cl (aq., sat.) were added and it was further extracted with Et_2O (3x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed to yield 3.773 g (24.8 mmol, 97 %) of the product as colorless crystals.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 2.05 (t, ^4J = 2.7 Hz, 2H, $CH_2C\equiv\underline{CH}$), 2.10 (bs, 2H, $CH_2\underline{OH}$), 2.37 (d, ^4J = 2.7 Hz, 4H, $\underline{CH_2}C\equiv CH$), 3.76 (s, 4H, $\underline{CH_2}OH$).

The analysis data harmonizes with the reported values in the literature.[328]

2.1.3 Precursor Alcohol 4

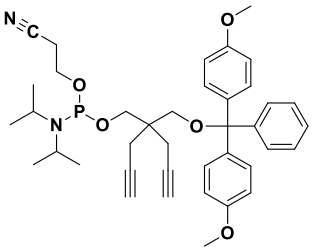


To a stirred solution of diol **3** (1.060 g, 7.0 mmol, 1.1 eq.) in 75 mL THF (anh.) : pyridine (anh.) = 2 :1 (v : v) 4,4'-dimethoxytrityl chloride (2.155 g, 6.3 mmol, 1.0 eq.) was added in four intervals of each 40 min at RT and under argon atmosphere. The reaction mixture was stirred overnight until the reaction was quenched by the addition of MeOH. The solvent was removed, 200 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (4x 100 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 20 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 2.566 g (5.6 mmol, 89 %) of the product as a pale yellow oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.90 (t, ^3J = 6.2 Hz, 1H, $CH_2\underline{OH}$), 1.95 (m, 4H, $CH_2C\equiv\underline{CH}$), 2.40 (m, 4H, $\underline{CH_2}C\equiv CH$), 3.20 (s, 2H, $\underline{CH_2}$ODMT), 3.62 (d, ^3J = 6.2 Hz, 2H, $\underline{CH_2}OH$), 3.79 (s, 6H, $ArO\underline{CH_3}$), 6.83 (m, 4H, Ar$\underline{H}$), 7.23 (m, 1H, Ar$\underline{H}$), 7.29 (m, 2H, Ar$\underline{H}$), 7.33 (m, 4H, Ar$\underline{H}$), 7.44 (m, 2H, Ar$\underline{H}$) ppm.

The analysis data harmonizes with the reported values in the literature[10] where spectra were recorded in deuterated acetone.

2.1.4 Phosphoramidite monomer 5

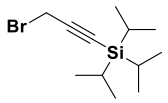


To a solution of mono-protected diol **4** (1058.9 mg, 2.33 mmol, 1.0 eq.) in 25 mL DCM (anh.) was added DIPEA (2.5 mL, 14.3 mmol, 6.1 eq.) and 200 mg of molecular sieves (3 Å) under argon atmosphere at 0 °C. To the stirring reaction mixture 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (562.5 mg, 2.38 mmol, 1.02 eq.) was then added. The reaction mixture was stirred at 0 °C for another hour and then allowed to warm to RT for 1 h. The solvent was removed and the residue purified by column chromatography (SiO_2 vs 30 v% EtOAc in cyclohexane + 2 v% DIPEA) to obtain 1494.1 mg (2.28 mmol, 98 %) of the product as a colorless foam.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.14 (m, 12H, $N(CH_2(\underline{CH_3})_2)_2$), 1.89 (m, 2H, $CH_2C\equiv\underline{CH}$), 2.42 (m, 4H, $\underline{CH_2}C\equiv CH$), 2.54 (m, 2H, $CH_2\underline{CH_2}C\equiv N$), 3.17 (m, 2H, $\underline{CH_2}$ODMT), 3.62 (m, 6H, $\underline{CH_2}CH_2C\equiv N$, $\underline{CH_2}OP$, $N(\underline{CH_2}(CH_3)_2)_2$), 3.79 (s, 6H, $ArO\underline{CH_3}$), 6.81 (m, 4H, Ar$\underline{H}$), 7.20 (m, 1H, Ar$\underline{H}$), 7.26 (m, 2H, Ar$\underline{H}$), 7.33 (m, 4H, Ar$\underline{H}$), 7.44 (m, 2H, Ar$\underline{H}$).

The analysis data harmonizes with the reported values in the literature[10] where spectra were recorded in deuterated acetone.

2.1.5 TIPS-protected propargyl bromide 6

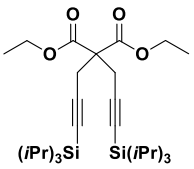


At -78 °C and under argon atmosphere *n*-butyllithium (2.5 M in hexanes, 32.0 mL, 80 mmol, 1 eq.) was added dropwise to 150 mL of THF (anh.). Propargyl bromide (80 w% in toluene, stabilized with MgO, 14.9 g, 100 mmol, 1.25 eq.) was then added dropwise under continuous stirring. After 15 min triisopropylsilyl chloride (15.4 g, 80 mmol, 1 eq.) was added dropwise at -78 °C. After stirring the reaction mixture overnight and allowing it to warm to RT, it was quenched with NH_4Cl (aq., sat.) and the solvent was removed. The slurry was taken in 50 mL of DCM, 100 mL NH_4Cl (aq., sat.) were added and it was further extracted with DCM (3x 50 mL). The org. phase was washed with brine (3x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil (m = 17.94 g) was purified by column chromatography (SiO_2 vs 5 v% DCM in cyclohexane) to obtain 9.4610 g (34.4 mmol, 43 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.07 (m, 21H, Si(CH(CH_3)$_2$)$_3$), 3.95 (s, 2H, CH_2Br) ppm.

The analysis data harmonizes with the reported values in the literature.[329]

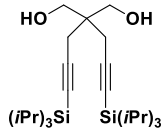
2.1.6 TIPS malonate 7



Under argon atmosphere, sodium hydride (1.10 g, 60 wt% in mineral oil, 27.5 mmol, 2.1 eq.) was suspended in 100 mL THF (anh.). Diethylmalonate (2.10 g, 13.1 mmol, 1.0 eq.) was added dropwise to the ice-cold reaction mixture that was afterwards stirred for 2 h and allowed to warm to RT. Then, 3-bromoprop-1-yn-1-yl)triisopropylsilane **6** (9.00 g, 32.7 mmol, 2.5 eq.) was added dropwise at 0°C and the reaction mixture was stirred overnight and allowed to reach RT. The reaction mixture was quenched with NH_4Cl (aq., sat.), THF was removed. The slurry was taken in 50 mL EtOAc and 100 mL NH_4Cl (aq., sat.) were added and it was further extracted with EtOAc (3x 100 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The crude yellow oil (m = 8.99 g) was purified by column chromatography (SiO_2 vs 10 v% EtOAc in cyclohexane) to obtain 6.7672 g (12.3 mmol, 94 %) of the product as a colorless crystalline powder.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.04 (m, 42H, Si(CH(CH_3)$_2$)$_3$), 1.25 (t, ^3J = 7.0 Hz, 6H, CH_2CH_3), 3.08 (s, 4H, CH_2C≡CTIPS), 4.19 (q, ^3J = 7.0 Hz, 4H, CH_2CH_3); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 11.3, 14.1, 18.6, 24.1, 56.8, 61.9, 77.2, 84.1, 102.6, 168.8.

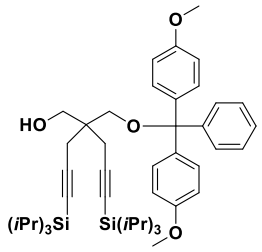
2.1.7 TIPS diol 8



Under argon atmosphere, lithium aluminum hydride solution (7.1 mL, 2 M in THF, 14.2 mmol, 1.2 eq.) was diluted in 50 mL THF (anh.). At -6 °C a solution of diethyl malonate **7** (6.5048 g, 11.85 mmol, 1.0 eq.) in 20 mL THF (anh.) was added dropwise to the solution. The reaction mixture was stirred overnight and allowed to reach RT. At 0 °C the reaction was quenched with 10.0 mL brine. The slurry was taken in 50 mL Et_2O and 50 mL NH_4Cl (aq., sat.) were added and it was further extracted with Et_2O (3x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed to yield 5.5096 g (11.85 mmol, full conversion) of the product as a red oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.06 (m, 42H, Si(CH(CH_3)$_2$)$_3$), 2.14 (bs, 2H, CH_2OH), 2.43 (s, 4H, CH_2C≡CTIPS), 3.76 (s, 4H, CH_2OH); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 11.4, 18.7, 23.2, 42.7, 66.7, 83.6, 104.5.

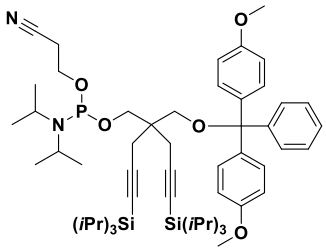
2.1.8 TIPS precursor alcohol 9



To a stirred solution of diol **8** (2.74 g, 5.9 mmol, 1.0 eq.) in 30 mL THF (anh.) : pyridine (anh.) = 2 :1 (v : v) 4,4'-dimethoxytrityl chloride (2.00 g, 5.9 mmol, 1.0 eq.) was added in four intervals of each 40 min at RT and under argon atmosphere. The reaction mixture was stirred overnight until the reaction was quenched by the addition of MeOH. The solvent was removed, 200 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (4x 100 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 10 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 3.84 g (5.0 mmol, 85 %) of the product as a pale yellow oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.05 (m, 42H, Si(CH(CH_3)$_2$)$_3$), 2.01 (m, 1H, CH_2OH), 2.55 (m, 4H, CH_2C≡CTIPS), 3.29 (m, 2H, CH_2ODMT), 3.70 (m, 2H, CH_2OH), 3.79 (s, 6H, ArOCH_3), 6.85m (m, 4H, ArH), 7.23 (m, 1H, ArH), 7.29 (m, 2H, ArH), 7.33 (m, 4H, ArH), 7.44 (m, 2H, ArH) ppm; ^{13}C NMR (100.6 MHz, $CDCl_3$, *δ*, ppm): 11.5, 18.8, 23.8, 42.9, 55.3, 66.3, 67.2, 83.4, 86.6, 104.2, 113.4, 127.0, 128.1, 128.2, 130.2, 135.9, 144.9, 158.7.

2.1.9 TIPS phosphoramidite monomer 10

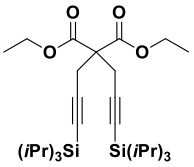


To a solution of mono-protected diol **9** (320.0 mg, 0.42 mmol, 1.0 eq.) in 15 mL DCM (anh.) was added DIPEA (0.44 mL, 2.6 mmol, 6.1 eq.) and 105 mg of molecular sieves (3 Å) under argon atmosphere at 0 °C. To the stirring reaction mixture 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (101.7 mg, 0.43 mmol, 1.0 eq.) was then added. The reaction mixture was stirred at 0 °C for another hour and then allowed to warm to RT for 30 min. The solvent was removed and the residue purified by column

chromatography (SiO_2 vs 10 v% EtOAc in cyclohexane + 2 v% DIPEA) to obtain 398.7 mg (0.41 mmol, 99 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $(CH_3)_2CO$, δ, ppm): 1.04 (m, 42H, Si(CH(CH_3)$_2$)$_3$), 1.17 (m, 12H, N(CH_2(CH_3)$_2$)$_2$), 2.59 (m, 4H, CH_2CΞCTIPS), 2.67 (m, 2H, CH_2CH_2C≡N), 3.28 (m, 2H, CH_2OP), 3.75 (m, 2H, CH_2CH_2CΞN), 3.77 (m, 2H, CH_2ODMT), 3.79 (s, 6H, ArOCH_3), 6.88 (m, 4H, ArH), 7.22 (m, 1H, ArH), 7.32 (m, 6H, ArH), 7.47 (m, 2H, ArH); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 11.4, 18.8, 20.8, 23.8, 27.1, 42.9, 48.7, 55.3, 66.3, 67.1, 83.3, 86.5, 104.7, 113.1, 113.3, 127.0, 128.0, 128.2, 130.2, 130.5, 135.9, 144.9, 158.7; ^{31}P NMR (161.9 MHz, $(CH_3)_2CO$, δ, ppm): 147.9.

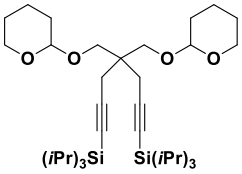
2.1.10 TIPS malonate 7 (Strategy B)



N-BuLi (26.5 mL, 2.5 M in hexanes, 66.3 mmol, 3.0 eq.) was added dropwise to a stirred solution of malonate **2** (5.145 g, 21.8 mmol) in 50 mL THF (anh.) at – 77 °C. After 90 min TIPS chloride (9.22 g, 47.8 mmol, 2.2 eq.) was added dropwise. The reaction mixture was stirred overnight and was thereby allowed to reach RT. After quenching with brine, 100 mL Et_2O were added and the organic phase was washed with 50 mL $NaHCO_3$ (sat., aq.) and 50 mL brine, dried over Na_2SO_4, filtered and the solvent was removed to obtain a dark red crude oil (m = 13 g). The residue was purified by column chromatography (SiO_2 vs 5→10v% EtOAc in cyclohexane) to obtain 7.242 g (13.2 mmol, 61 %) of the product as a red oil.

The analysis data corresponds to the values previously reported for TIPS malonate **7** in 2.1.6 on page 140.

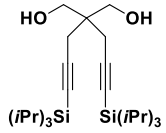
2.1.11 Intermediate 11 (Strategy C)



To a stirred solution of diol **3** (1.618 g, 10.6 mmol, 1.0 eq.) in 30 mL DCM (anh.) was added 3,4-dihydro-2*H*-pyran (4.471 g, 53.2 mmol, 5.0 eq.), a 1 mol% of *para*-toluenesulfonic acid monohydrate (24.1 mg, 0.1 mmol) and 80 mg of molecular sieves (3 Å) under argon atmosphere. After 3 h the reaction mixture was diluted with 75 mL Et_2O and washed with 50 mL water, 50 mL $NaHCO_3$ (sat., aq.) and 10 mL brine, dried over Na_2SO_4, filtered and the solvent was removed to obtain the intermediate tetrahydropyranyl (THP) ether as a colorless oil that is then without further purification dissolved in 30 mL THF (anh.) and stirred under argon atmosphere. At – 77 °C, *n*-BuLi (13 mL, 2.5 M in hexanes, 32 mmol, 3.0 eq.) is added dropwise to the solution that is then stirred for another 90 min until TIPS chloride (4.461, 23.1 mmol, 2.2 eq.) is added dropwise. The reaction mixture was stirred overnight and was thereby allowed to reach RT. After quenching with brine, 100 mL Et_2O were added and the organic phase was washed with 50 mL $NaHCO_3$ (sat., aq.) and 50 mL brine, dried over Na_2SO_4, filtered and the solvent was removed to obtain a yellow crude oil (m = 7.881 g). The residue was purified by column chromatography (SiO_2 vs 5 v% EtOAc in cyclohexane) to obtain 4.885 g (7.7 mmol, 73 %) of the product as a colorless oil.

^{1}H NMR (400 MHz, $CDCl_3$, δ, ppm): 1.06 (m, 42H, Si(CH(CH_3)$_2$)$_3$), 1.65 (m, 12H, $CH_2CH_2CH_2$), 2.48 (s, 4H, CH_2CΞCTIPS), 3.59 (m, 4H, CH_2OTHP), 3.69 (m, 4H, OCH_{2THP}), 4.64 (m, 2H, (RO)$_2$CHR).

2.1.12 TIPS diol 8 (Strategy C)

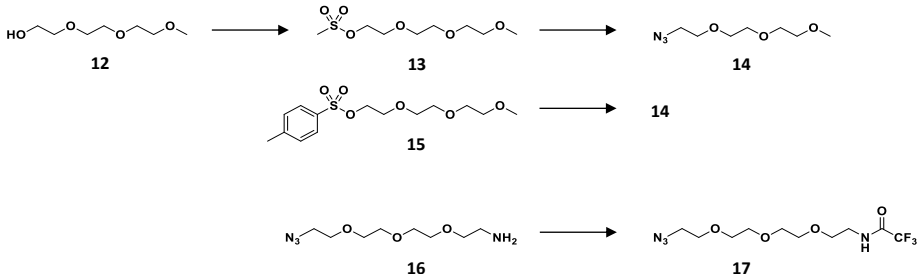


To a stirred solution of TIPS-protected THP-ether **11** (432.0 mg, 0.68 mmol, 1.0 eq.) in 20 mL MeOH was added *para*-toluenesulfonic acid monohydrate (167.7 mg, 0.88 mmol, 1.3 eq.) at RT. After 1 h the solvent was removed and the residue was taken in 20 mL Et_2O and 20 mL $NaHCO_3$ (sat., aq.) were added. It was extracted with Et_2O (2x 20 mL) and the organic phase was washed with 20 mL brine, dried over Na_2SO_4, filtered and the solvent was removed to obtain 260.1 mg (0.56 mmol, 82 %) of the product as a pale yellow oil.

The analysis data corresponds to the values previously reported for TIPS diol **8** in 2.1.7 on page 140.

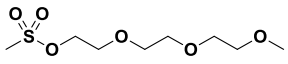
2.2 Synthesis of other Reagents

Synthetic access to model azides **14** and **17** following Scheme V-2 is described as well as the access to azides **20**, **21**, **24**, **26**, and **34.**



Scheme V-2: Overview Synthesis of other reagents for Chapter II.

2.2.1 Mesylate 13

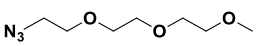


To a stirred solution of triethylene glycol monomethyl ether **12** (8.22 g, 50 mmol, 1.0 eq.) and triethylamine (15.29 g, 150 mmol, 3.0 eq.) in 100 mL DCM anh. was added dropwise methanesulfonyl chloride (5.73 g, 50 mmol, 1.0 eq.) at 0 °C. The reaction mixture was stirred overnight and thereby allowed to reach RT and then filtered. The solvent of the filtrate was removed, the crude was taken in 50 mL EtOAc, 50 mL $NaHCO_3$ (sat., aq.) was added and it was further extracted with EtOAc (4x 100 mL). The organic phase was washed with 50 mL brine, dried over Na_2SO_4, filtered and the solvent was removed to obtain 9.28 g (38.3 mmol, 77 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 3.00 (s, 3H, $SO_2\underline{CH_3}$), 3.30 (s, 3H, $O\underline{CH_3}$), 3.47 (m, 2H, $\underline{CH_2}O$), 3.59 (m, 6H, $\underline{CH_2}O$), 3.69 (m, 2H, $\underline{CH_2}O$), 4.30 (m, 2H, $\underline{CH_2}OMs$).

The analysis data harmonizes with the reported values in the literature.[330]

2.2.2 Azide 14 from mesylate 13



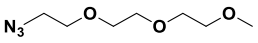
Mesylate **13** (2.51 g, 10.4 mmol, 1 eq.) and sodium azide (3.37 g, 51.8 mmol, 5 eq.) were dissolved in 50 mL DMF. The reaction mixture was stirred at 80 °C overnight and then poured in 300 mL ice-cold water. The product was extracted with Et_2O (5x 100 mL) and the organic phase was washed with H_2O

(2x 100 mL), brine (2x 100 mL), dried over Na_2SO_4, filtered and the solvent was removed to obtain 850 mg (4.5 mmol, 43 %) of the product as a pale yellow oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 3.38 (s, 3H, O$\underline{CH_3}$), 3.39 (m, 2H, $\underline{CH_2}N_3$), 3.55 (m, 2H, $\underline{CH_2}O$), 3.66 (m, 8H, $\underline{CH_2}O$), 4.30 (m, 2H, $\underline{CH_2}O$).

The analysis data harmonizes with the reported values in the literature.[331]

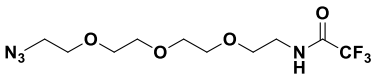
2.2.3 Azide 14 from tosylate 15



Tosylate **15** (6.00 g, 18.7 mmol, 1 eq.) and sodium azide (6.12 g, 93.5 mmol, 5 eq.) were stirred in 30 mL DMSO overnight at RT. The reaction mixture was diluted with 100 mL ice-cold water and extracted with Et_2O (4x 50 mL) and the organic phase was washed with 20 mL H_2O, brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed to obtain 3.07 g (16.2 mmol, 87 %) of the product as a pale yellow oil.

The analysis data harmonizes with the reported values in 2.2.2 above and in the literature.[331]

2.2.4 Azide 17

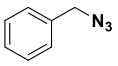


To a solution of 11-Azido-3,6,9-trioxaundecan-1-amine **16** (2.33 g, 10.2 mmol, 1.0 eq.) and triethylamine (1.21 g, 12.2 mmol, 1.2 eq.) in 3 mL DCM anh. was added dropwise 2,2,2-trifluoroacetic anhydride (2.36, 11.2 mmol, 1.1 eq.) at 0 °C. The reaction mixture was stirred overnight and thereby allowed to reach RT. The solvent removed and the crude was purified by column chromatography (SiO_2 vs EtOAc) to obtain 2.97 g (9.5 mmol, 93 %) of the product as a pale yellow oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 3.35 (m, 2H, $\underline{CH_2}N_3$), 3.53 (m, 2H, $\underline{CH_2}NH$), 3.55 (m, 2H, $\underline{CH_2}O$), 3.63 (m, 12H, $\underline{CH_2}O$), 7.13 (bs, 1H, $\underline{NH}COCF_3$) ppm; ^{19}F NMR (376.5 MHz, $CDCl_3$, δ, ppm): -75.9.

The ^{1}H NMR data harmonizes with the reported values in the literature.[333]

2.2.5 Azide 20

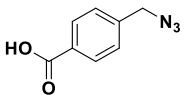


Benzyl bromide **18** (1241.5 mg, 7.3 mmol, 1.0 eq.) and sodium azide (591.1 mg, 9.1 mmol, 1.25 eq.) were dissolved in 20 mL DMSO. The reaction mixture was stirred at RT overnight and then poured in 300 mL NH_4Cl (aq., sat.) The product was extracted with Et_2O (3x 100 mL) and the organic phase was washed with H_2O (5x 100 mL), brine (3x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed to obtain 792.8 mg (6.0 mmol, 82 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 4.52 (s, 2H, $\underline{CH_2}N_3$), 7.37 (m, 5H, ArH).

The analysis data harmonizes with the reported values in the literature.[378]

2.2.6 Azide 21

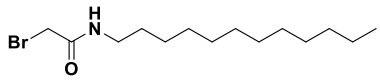


4-(Bromomethyl)benzoic acid **19** (1327.7 mg, 6.2 mmol, 1.0 eq.) and sodium azide (584.5 mg, 9.0 mmol, 1.45 eq.) were dissolved in 20 mL DMSO. The reaction mixture was stirred at RT overnight and then poured in 300 mL NH_4Cl (aq., sat.) The product was extracted with Et_2O (3x 100 mL) and the organic phase was washed with H_2O (5x 100 mL), brine (3x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed to obtain 939.5 mg (5.3 mmol, 86 %) of the product as a colorless crystalline solid.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 4.45 (s, 2H, $\underline{CH_2}N_3$), 7.44 (d, ^3J = 10.9 Hz, 2H, ArH), 8.13 (d, ^3J = 10.9 Hz, 2H, ArH).

The analysis data harmonizes with the reported values in the literature.[379]

2.2.7 Bromide 23

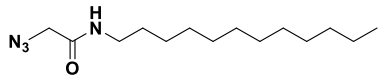


To a solution of *n*-dodecylamine **22** (5577.0 mg, 30 mmol, 1.0 eq.) in 50 mL DCM was first added K_2CO_3 (6226.5 mg, 45 mmol, 1.5 eq.) in 50 mL H_2O. To the stirred reaction mixture was added dropwise bromoacetyl bromide (4.0 mL, 45 mmol, 1.5 eq.) in 40 mL DCM at 0 °C. The reaction mixture was then stirred for 2 h and allowed to react RT until the organic phase was separated and the aqueous layer was extracted with DCM (3x 50 mL). The organic phase was then washed with 50 mL H_2O, brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed to obtain 9.08 g (29.6 mmol, 99 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 0.78 (m, 3H, $CH_2\underline{CH_3}$), 1.19 (m, 18H, $\underline{CH_2}$), 1.44 (m, 2H, $\underline{CH_2}$), 3.17 (m, 2H, $\underline{CH_2}$NHCO), 3.77 (CO$\underline{CH_2}$Br), 7.05 (m, 1H, $\underline{NH}$CO).

The analysis data harmonizes with the reported values in the literature.[380]

2.2.8 Bromide 24

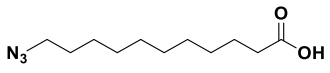


2-Bromo-*N*-dodecylethanamide **23** (5.00 mg, 16.3 mmol, 1.0 eq.) and sodium azide (5.07 g, 78.0 mmol, 4.8 eq.) were dissolved in 20 mL DMSO. The reaction mixture was stirred at RT overnight and then poured in 300 mL NH_4Cl (aq., sat.) The product was extracted with Et_2O (3x 100 mL) and the organic phase was washed with H_2O (5x 100 mL), brine (3x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed to obtain 4.16 g (15.5 mmol, 95 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 0.74 (m, 3H, $CH_2\underline{CH_3}$), 1.14 (m, 18H, $\underline{CH_2}$), 1.36 (m, 2H, $\underline{CH_2}$), 3.11 (m, 2H, $\underline{CH_2}$NHCO), 3.76 (CO$\underline{CH_2}N_3$), 6.94 (m, 1H, $\underline{NH}$CO).

The analysis data harmonizes with the reported values in the literature.[381]

2.2.9 Azide 26



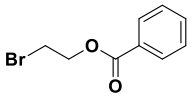
11-Bromoundecanoic acid **25** (5302.1 mg, 20 mmol, 1.0 eq.) and sodium azide (6505.0 mg, 100 mmol, 5 eq.) were dissolved in 20 mL DMSO. The reaction mixture was stirred at RT overnight and then poured in 300 mL NH_4Cl (aq., sat.) The product was extracted with Et_2O (3x 100 mL) and the organic phase was

washed with H_2O (5x 100 mL), brine (3x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed to obtain 4556.6 mg (full conversion) of the product as a colorless powder.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.31 (m, 12H, $\underline{CH_2}$), 1.59 (m, 4H, $\underline{CH_2}$), 2.23 (m, 2H, $\underline{CH_2}COOH$), 3.24 (m, 2H, $\underline{CH_2}N_3$), 11.15 (bs, 1H, $\underline{COOH}$).

The analysis data harmonizes with the reported values in the literature.[382]

2.2.10 Bromide 28

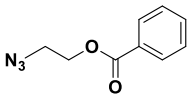


1-bromo-3-propanol **27** (1.2530 mg, 10.0 mmol, 1.0 eq.) and DIPEA (6.56 g, 50 mmol, 5 eq.) were dissolved in 25 mL anh. DCM. To the stirred reaction mixture was added dropwise benzoic anhydride (3395.8 mg, 15 mmol, 1.5 eq.) at 0 °C. After 3 h the reaction mixture was stirred under reflux for 8 h. The solvent was removed, 100 mL NH_4Cl (aq., sat.) were added and it was extracted with EtOAc (3x 100 mL) and the organic phase was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The crude was purified by column chromatography (SiO_2 vs 20 v% acetone in cyclohexane) to obtain 831.5 mg (3.63 mmol, 36 %) of the product as a pale yellow oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 3.65 (t, ^3J = 6.1 Hz, 2H, CH_2Br), 4.63 (t, ^3J = 6.1 Hz, 2H, CH_2O), 7.46 (m, 2H, Ar$\underline{H}$), 7.58 (m, 1H, Ar$\underline{H}$), 8.06 (m, 2H, Ar$\underline{H}$).

The analysis data harmonizes with the reported values in the literature.[334]

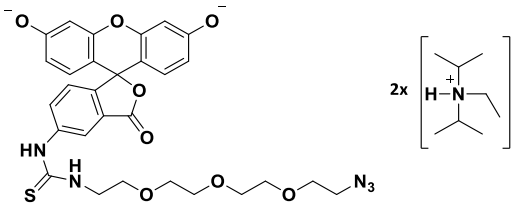
2.2.11 Azide 29



2-Bromoethylbenzoate **28** (632.7 mg, 2.8 mmol, 1.0 eq.) and sodium azide (903.0 mg, 13.8 mmol, 5.0 eq.) were dissolved in 10 mL DMSO. The reaction mixture was stirred at RT overnight and then poured in 100 mL NH_4Cl (aq., sat.) The product was extracted with Et_2O (3x 50 mL) and the organic phase was washed with H_2O (5x 50 mL), brine (3x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed to obtain 537.7 mg (full conversion) of the product as a yellow oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 3.60 (t, ^3J = 5.1 Hz, 2H, CH_2N_3), 4.50 (t, ^3J = 5.1 Hz, 2H, CH_2O), 7.46 (m, 2H, Ar$\underline{H}$), 7.58 (m, 1H, Ar$\underline{H}$), 8.07 (m, 2H, Ar$\underline{H}$).

2.2.12 Fluorescent Azide 34

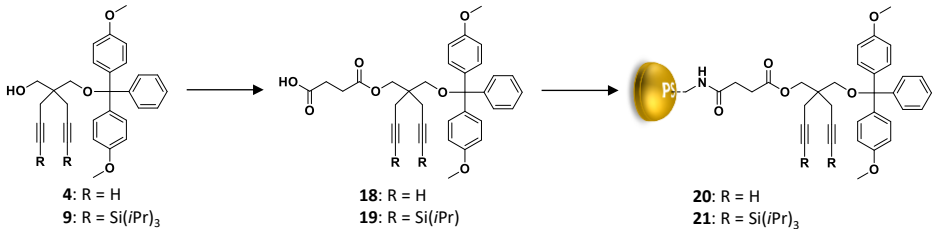


To a stirred solution of 11-azido-3,6,9-trioxaundecan-1-amine **16** (52.6 mg, 270 µmol, 1.0 eq.) and DIPEA (87.5 mg, 680 µmol, 2.5 eq.) in 8 mL $CDCl_3$ was added fluorescein isothiocyanate **30** (104.5 mg, 270 µmol, 1.0 eq.) and it was stirred overnight at RT. It was coevaporated with chloroform to remove excess DIPEA. The DIPEA-salt of the product was quantitatively recovered as a red oil.

HRMS m/z: $[M+H]^+$ calculated for $C_{29}H_{30}N_5O_8S^+$ 608.1810; found 608.1817; ^{1}H NMR (400.1 MHz, CD_3OD, δ, ppm): 1.22 (m, 30H, $HN(CH_2\underline{CH_3})(CH(\underline{CH_3})_2)_2^+$), 3.03 (m, 4H, $HN(\underline{CH_2}CH_3)(CH(CH_3)_2)_2^+$), 3.32 (m, 2H, $\underline{CH_2}N_3$), 3.52 (m, 4H, $HN(CH_2CH_3)(\underline{CH}(CH_3)_2)_2^+$), 3.66 (m, 14H, $\underline{CH_2}\underline{CH_2}O$), 6.58 (m, 6H, Ar$\underline{H}$), 7.14 (m, 3H, Ar$\underline{H}$), 7.88 (m, 2H, $(\underline{NH})_2C=S$).

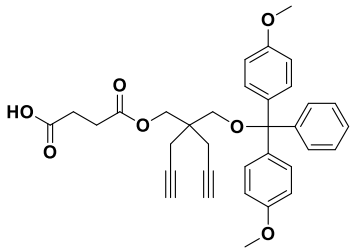
2.3 Support Synthesis

Synthetic access to solid polystyrene supports **37** and **38** following Scheme V-3 is described below.



Scheme V-3: Overview of solid polystyrene support synthesis.

2.3.1 Succinate 35

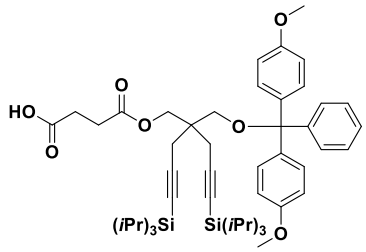


Precursor alcohol **4** (273.0 mg, 0.6 mmol, 1.0 eq.) was co-evaporated with 5 mL of pyridine (anh.), dried under vacuum for 30 min, and reacted with succinic anhydride (91.6 mg, 0.9 mmol, 1.5 eq.) and DMAP (110.7 mg, 0.9 mmol, 1.5 eq.) in 4 mL of pyridine (anh.) at 35 °C for 5 h and allowed to cool to RT. The reaction was quenched with methanol, and the solvent was removed. The resulting slurry oil was taken in 30 mL EtOAc, washed with 50 mL H_2O, 50 mL K_2HPO_4 (sat., aq., pH 5.0) and brine (3x 20 mL). The organic phase was dried over Na_2SO_4, filtered and the solvent was removed to obtain 329.9 mg (0.6 mmol, 99 %) of the product as a pale yellow oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.93 (m, 2H, $CH_2C\equiv\underline{CH}$), 2.41 (m, 4H, $\underline{CH_2}C\equiv CH$), 2.54 (m, 4H, $\underline{CH_2}COO$), 3.15 (s, 2H, $\underline{CH_2}ODMT$), 3.78 (s, 6H, $ArO\underline{CH_3}$), 4.15 (s, 2H, $\underline{CH_2}OCO$), 6.82 (m, 4H, ArH), 7.20 (m, 1H, ArH), 7.27 (m, 2H, ArH), 7.31 (m, 4H, ArH), 7.42 (m, 2H, ArH).

The analysis data has not been reported along with the literature protocol.[10]

2.3.2 Succinate 36

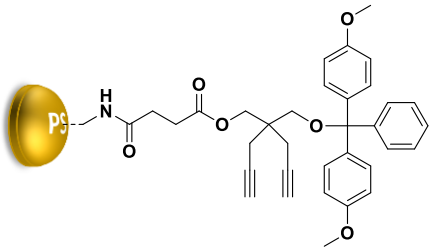


TIPS precursor alcohol **9** (834.0 mg, 1.1 mmol, 1.0 eq.) was co-evaporated with 5 mL of pyridine (anh.), dried under vacuum for 30 min, and reacted with succinic anhydride (175.6 mg, 1.8 mmol, 1.6 eq.) and DMAP (207.8 mg, 1.7 mmol, 1.6 eq.) in 4 mL of pyridine (anh.) at 35 °C overnight and allowed to cool

to RT. The reaction was quenched with methanol, and the solvent was removed. The resulting slurry oil was taken in 50 mL EtOAc, washed with 50 mL of H_2O, 50 mL of K_2HPO_4 (sat., aq., pH 5.0) and brine (3x 20 mL). The organic phase was dried over Na_2SO_4, filtered and the solvent was removed to obtain 917.7 mg (1.1 mmol, 97 %) of the product as a pale yellow oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.05 (m, 42H, Si(CH(CH$_3$)$_2$)$_3$), 2.53 (m, 4H, CH$_2$C≡CTIPS), 2.59 (m, 4H, CH$_2$COO), 3.24 (s, 2H, CH$_2$ODMT), 3.79 (s, 6H, ArOCH$_3$), 4.27 (s, 2H, CH$_2$OCO), 6.83 (m, 4H, ArH), 7.21 (m, 1H, ArH), 7.29 (m, 6H, ArH), 7.42 (m, 2H, ArH); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 11.3, 18.7, 23.8, 28.8, 28.9, 41.7, 55.2, 63.7, 66.4, 83.5, 86.0, 103.8, 113.1, 126.8, 127.8, 128.3, 130.3, 136.0, 145.0, 158.5, 171.4, 178.2.

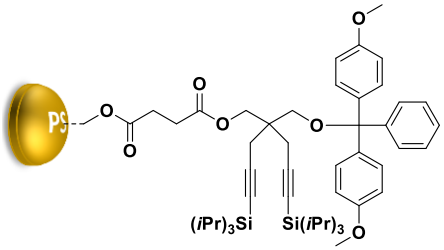
2.3.3 Solid Polystyrene Support 37



To a home-made peptide synthesis reaction vessel equipped with a fritted glass plate and a stopcock, aminomethylated polystyrene resin (1135.2 mg, loading: 1.4 mmol·g^{-1}, 1.6 mmol functionalities, 1.0 eq.) was charged and washed with DCM (anh.) (2x 10 mL), dried under vacuum and purged with argon. A solution of succinate **35** (1327.4 mg, 2.4 mmol, 1.5 eq.), DMAP (439.9 mg, 3.6 mmol, 2.25 eq.), and DCC (748.2 mg, 3.6 mmol, 2.25 eq.) in 15 mL DCM (anh.) was prepared under argon atmosphere and transferred to the reaction vessel which was then shaken overnight at RT. The resin was then washed with DCM (3x 5 mL), MeOH (3x 5 mL), MeCN (3x 5 mL), and Et_2O (3x 5 mL) and dried under vacuum for 2 h. The unreacted amino groups were capped with 3 mL acetic anhydride in 10 mL pyr (anh.) while shaking for 3 h. The resin was finally washed with DCM (3x 5 mL), MeCN (3x 5 mL), and Et_2O (3x 5 mL) and dried under vacuum overnight.

The loading of the support was measured by the removal of the DMT group of an aliquot and analyzing its absorbance at 504 nm to find a loading of 0.77 mmol·g^{-1}.

2.3.4 Solid Polystyrene Support 38



To a homemade peptide synthesis reaction vessel equipped with a fritted glass plate and a stopcock, aminomethylated polystyrene resin (734.9 mg, loading: 1.4 mmol·g^{-1}, 1.0 mmol functionalities, 1.0 eq.) was charged and washed with DCM (anh.) (2x 5 mL), dried under vacuum and purged with argon. A solution of succinate **36** (900.0 mg, 1.0 mmol, 1.0 eq.), DMAP (192.2 mg, 1.6 mmol, 1.5 eq.), and DCC (328.9 mg, 1.6 mmol, 1.5 eq.) in 10 mL DCM (anh.) was prepared under argon atmosphere and transferred to the reaction vessel which was then shaken overnight at RT. The resin was then washed with DCM (3x 5 mL), MeOH (3x 5 mL), MeCN (3x 5 mL), and Et_2O (3x 5 mL) and dried under vacuum for 2 h. The unreacted amino groups were capped with 2 mL acetic anhydride in 8 mL pyr (anh.) while

shaking for 2 h. The resin was finally washed with DCM (3x 5 mL), MeCN (3x 5 mL), and Et_2O (3x 5 mL) and dried under vacuum for 3 h.

The loading of the support was measured by the removal of the DMT group of an aliquot and analyzing its absorbance at 504 nm to find a loading of 0.69 mmol·g^{-1}.

2.4 Phosphodiester Synthesis

2.4.1 Manual Synthesis on Solid Polystyrene Supports

Sequence-coded precursors **S1**-**S4** were prepared following a previously reported procedure.[10] In brief, anhydrous solvents were used and a homemade peptide synthesis reaction vessel equipped with a fritted glass plate and a stopcock was employed to ensure phosphoramidite couplings under dry conditions using vacuum-argon purges. Depending on the desired sequence's leading monomer, one equivalent of modified support **37** or **38** was charged to the reactor and the terminal DMT protective group was removed by repeated treatment with 3 w% TCA solution in DCM until non-occurrence of red coloration. Remaining acid was removed by washing with 2 v% pyridine in MeCN and with pure MeCN. In the coupling step, the resin was first swollen in DCM for 5 min and then shaken at RT for 40 min with a solution of two equivalents of the desired phosphoramidite monomer **5** or **10** in acetonitrile (≈100 mM) and four equivalents of tetrazole in MeCN (0.45 M). The solution was then removed and the resin washed with MeCN. The immobilized phosphite triester was then oxidized to the corresponding phosphate triester by treating it two times for 5 min with an iodine solution (H_2O : 2,6-lutidine : THF = 2 : 20 : 80 (v : v : v), 0.1 M iodine). Remaining reagents were removed by washing with acetonitrile and DCM. DMT deprotection, phosphoramidite coupling and oxidation were iteratively repeated until the formation of the desired sequence. To avoid random iodination when monomer **5** exhibiting a terminal alkyne moiety was used, remaining iodine could be removed by shaking the resin with a saturated aqueous sodium thiosulfate solution for 30 min and washing the resin subsequently with a 1 : 1 (v : v)-mixture of H_2O and THF. Then, the 2-cyanoethyl phosphate protective groups were removed by treatment with 10 v% piperidine in MeCN for 5 min to form the corresponding phosphodiester bonds. The terminal DMT group was removed in a final deprotection step and the desired polymers were cleaved from the polystyrene resin by adding a solution of ammonia (aq., 28 w%) and methylamine (aq., 40 w%) (1 : 1, v : v) and heating the mixture for 2 h to 50 °C. The mixture was filtered and washed with 1,4-dioxane, ammonia was removed by argon bubbling and the solvent was removed under reduced pressure to obtain the desired polymers.

2.4.2 Automated Synthesis on an Expedite Oligonucleotide Synthesizer

Polymers **S5**-**S8** were prepared following a reported procedure[6] described in the Experimental Section: Automated Poly(phosphodiester) Synthesis on pages 135f. **S5**-**S8** were synthesized in the DMT-OFF mode and recovered together with defect-containing capped chains (see explanation in the Annex, subchapter 1.1 DMT-ON Purification, p. 185). However, when the precursors were kept for subsequent modification steps on the CPG support, the DMT-ON mode was used.

2.5 Sequential Modification

2.5.1 Oligomer Modification

37.4 mg **S4** (26.4 µmol), 0.9 mg CuBr (6.3 µmol), 9.5 mg TBTA (17.9 µmol) and 293.9 mg azide **14** (1.5 mmol) were added to a small round bottom flask. After removal of oxygen by subsequent vacuum-argon purges, 1.2 mL of a degassed 3 : 1 (v : v)-mixture of DMSO and *tert*-BuOH were added and the mixture was stirred overnight under argon. When a color change to green was observed, three drops of a 100 mM sodium ascorbate in degassed H_2O were added until a slightly yellow solution was

maintained. All solvents were removed under reduced pressure. The crude was taken in MeOH and dialyzed overnight in MeOH (dialysis membrane: MWCO: 1000 Da, regenerated cellulose, flat width: 45 mm, Roth SpectraPor). The solvent was then removed under reduced pressure to obtain 53.0 mg (24.4 µmol, 92 %) of modified polymer **S4'** as a slightly yellow oil.

26.2 mg TIPS protected polymer **S4'** (11.7 µmol) was dissolved in 1 mL 250 mM TBAF in THF and stirred for 2 h at 70 °C. The reaction mixture was allowed to cool to RT, fluoride ions were quenched with 1 mL of a solution of methoxytrimethylsilane in diethylether and all solvents were removed under reduced pressure. The crude was taken in MeOH and dialyzed for 8 h in MeOH (dialysis membrane: MWCO: 1000 Da, regenerated cellulose, flat width: 45 mm, Roth SpectraPor). The solvent was then removed under reduced pressure to obtain 13.8 mg (8.9 µmol, 76 %) of TIPS-deprotected polymer **S4"** as a slightly yellow oil.

13.7 mg **S4"** (8.9 µmol), 1.0 mg CuBr (6.9 µmol), 7.5 mg TBTA (14.1 µmol) and 471.4 mg azide **17** (1.5 mmol) were added to a small round bottom flask. After removal of oxygen by subsequent vacuum-argon purges, 1.2 mL of a degassed 3 : 1 (v : v)-mixture of DMSO and *tert*-BuOH were added and the mixture was stirred overnight under argon. All solvents were removed under reduced pressure. The crude was taken in MeOH and dialyzed overnight in MeOH (dialysis membrane: MWCO: 1000 Da, regenerated cellulose, flat width: 45 mm, Roth SpectraPor). The solvent was then removed under reduced pressure to obtain 5.6 mg (2.0 µmol, 22 %) of binary modified polymer **S4'''** as a slightly yellow oil.

2.5.2 Polymer Modification

7.4 mg CuBr (52 µmol), 55.8 mg TBTA (105 µmol), 768.9 mg azide **14** (4.0 mmol) and 41.7 mg sodium ascorbate (210 µmol) were added to a round bottom flask. After removal of oxygen by subsequent vacuum-argon purges 4.0 mL of a degassed 3 : 1 (v : v)-mixture of DMSO and *tert*-BuOH were added under argon. 0.5 mL of the freshly prepared solution were taken in a syringe together with 2.5 mL argon. During the connection to a CPG column immobilized with precursor **S7** (synthesized in DMT-ON mode, see previous page), an argon counter flow was maintained from the syringe as well as the column side. The connected syringe was flushed with argon again and the CPG column was connected to another argon flushed syringe in an argon counter flow. The solution was then passed repeatedly between the syringes and the system was gently shaken overnight at RT. The syringes were then removed and the solid support-filled column was washed with each 6 mL of a 3 : 1 (v : v)-mixture of DMSO and *tert*-BuOH, Et_2O, H_2O, and MeCN and finally flushed with argon. Modified polymer **S7'** was cleaved from the support by pressing repeatedly 0.5 mL of a solution of ammonia (aq., 28 w%) and methylamine (aq., 40 w%) (1 : 1, v : v) for 30 min through the column, from a connected syringe on the one end to a syringe on the other end. The support material was another time washed with 0.5 mL of said cleavage solution and the collected solutions were unified and the solvent was removed under reduced pressure.

In total, four CPG columns with immobilized precursor **S7** (4.0 µmol) were subject to the procedure that yielded 29.3 mg (2.0 µmol, 51 %) of **S7'** as a blue oil.

Note: Reverse phase column purification was not performed after this step because of the presence of TIPS-side groups chains (see explanation in the Annex, subchapter 1.1 DMT-ON Purification, p. 185). Nevertheless, all reagents (with the exception of coordinated copper explaining the blue appearance of the product) were removed during the washing steps.

17.6 mg TIPS protected polymer **S7'** (1.23 µmol) were dissolved in 2 mL 250 mM TBAF in THF and stirred for 2 h at 70 °C. The reaction mixture was allowed to cool to RT, fluoride ions were quenched with 2 mL of a solution of methoxytrimethylsilane and all solvents were removed under reduced

pressure. The dry reaction mixture was solubilized in 2 mL NaCl solution (aq., 10 w%) and filtered. The filtrate was recovered and applied to a modified DMT-ON purification method (see Experimental Section: 1.2.3.2 Manual DMT-ON Purification, p. 137) leaving out the TCA wash. The polymer-containing solution was collected in 0.5 mL fractions. After lyophilization 10.0 mg (0.8 µmol, 67 %) of TIPS-deprotected **S7''** were recovered as a slightly yellow oil.

Note: Reverse phase column chromatography was here mostly used to remove excess tetrabutylammonium ions.

9.5 mg **S7''** (0.8 µmol), 2.3 mg CuBr (16.0 µmol), 17.4 mg TBTA (32.8 µmol), 11.3 mg sodium ascorbate (57.0 µmol) and 316.9 mg azide **17** (1.0 mmol) were added to a small round bottom flask. After removal of oxygen by subsequent vacuum-argon purges, 1.0 mL of a degassed 3 : 1 (v : v)-mixture of DMSO and *tert*-BuOH was added and the mixture was stirred overnight under argon. All solvents were removed under reduced pressure. The dry reaction mixture was solubilized in 2 mL NaCl solution (aq., 10 w%) and filtered. The filtrate was recovered and applied to a full DMT-ON purification method (see Experimental Section: 1.2.3.2 Manual DMT-ON Purification, p. 137). After lyophilzation 4.9 mg (0.3 µmol, 39 %) of binary modified **S7'''** were recovered as a slightly yellow oil.

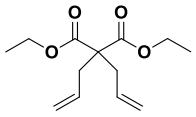
Note: Since **S7'''** does not contain TIPS groups, reverse phase column chromatography was employed at this final stage to remove the capped fractions left from the synthesis of **S7** (see explanation in the Annex, subchapter 1.1 DMT-ON Purification, p. 185).

NMR and MS analysis of the described intermediates is provided in the Annex: 1.2 *Sequential Modification of Poly(phosphodiester)s* on pages 186f. together with further details.

2.6 Perspective Synthesis of an Alkene-bearing Phosphoramidite monomer

The synthetic access to alkene monomer **44** is described herein in detail.

2.6.1 Malonate 41

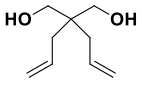


Under argon atmosphere, sodium hydride (4.80 g, 60 wt% in mineral oil, 120 mmol, 2.0 eq.) was suspended in 80 mL THF (anh.). Diethylmalonate (9.56 g, 60 mmol, 1.0 eq.) was added dropwise to the ice-cold reaction mixture that was afterwards stirred for 2 h and allowed to warm to RT. Then, allyl bromide (18.38 g, 152 mmol, 2.5 eq.) was added dropwise at 0°C and the reaction mixture was stirred for 3 h and allowed to reach RT. The reaction mixture was quenched with NH_4Cl (aq., sat.), THF was removed. The slurry was taken in 100 mL EtOAc and 100 mL NH_4Cl (aq., sat.) were added and it was further extracted with EtOAc (3x 150 mL). The combined organic layer was washed with brine (3x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. Without further purification 14.85 g (full conversion) of the product were obtained as colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.24 (t, ^3J = 7.1 Hz, 6H, $CH_2\underline{CH_3}$), 2.63 (m, 4H, $\underline{CH_2}CH=CH_2$), 4.18 (q, ^3J = 7.1 Hz, 4H, $\underline{CH_2}CH_3$), 5.10 (m, 4H, $CH_2CH=\underline{CH_2}$), 5.65 (m, 2H, $CH_2\underline{CH}=CH_2$).

The ^{1}H NMR data harmonizes with the reported values in the literature.[341]

2.6.2 Diol 42

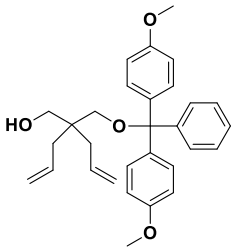


Under argon atmosphere, lithium aluminum hydride solution (37 mL, 2 M in THF, 74 mmol, 1.2 eq.) was diluted in 50 mL THF (anh.). At -10 °C a solution of malonate **41** (14.82 g, 59.6 mmol, 1.0 eq.) in 10 mL THF (anh.) was added dropwise to the solution. The reaction mixture was stirred overnight and allowed to reach RT. At 0 °C the reaction was quenched with 10.0 mL brine. The slurry was taken in 50 mL Et_2O and 50 mL NH_4Cl (aq., sat.) were added and it was further extracted with Et_2O (3x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed to yield 9.10 g (58.3 mmol, 98 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 2.09 (bs, 2H, $CH_2\underline{OH}$), 2.10 (m, 4H, $\underline{CH_2}CH=CH_2$), 3.59 (m, 4H, $\underline{CH_2}OH$), 5.11 (m, 4H, $CH_2CH=\underline{CH_2}$), 5.85 (m, 2H, $CH_2\underline{CH}=CH_2$).

The ^{1}H NMR data harmonizes with the reported values in the literature.[342]

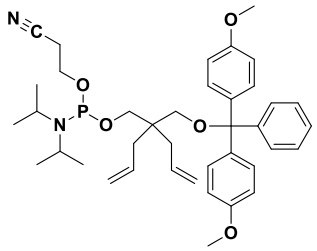
2.6.3 Alcohol precursor 43



To a stirred solution of diol **42** (4.43 g, 28.4 mmol, 1.5 eq.) in 150 mL THF (anh.) : pyridine (anh.) = 2 :1 (v : v) 4,4'-dimethoxytrityl chloride (6.47 g, 19.1 mmol, 1.0 eq.) was added in six intervals of each 30 min at RT and under argon atmosphere. The reaction mixture was stirred overnight until the reaction was quenched by the addition of MeOH. The solvent was removed, 200 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (4x 100 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 20 v% EtOAc in cyclohexane + 2 v% DIPEA) to obtain 8.83 g (full conversion) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.99 (t, ^3J = 6.5 Hz, 1H, $CH_2\underline{OH}$), 2.10 (m, 4H, $\underline{CH_2}CH=CH_2$), 3.04 (s, 2H, $\underline{CH_2}$ODMT), 3.48 (d, ^3J = 6.5 Hz, 2H, $\underline{CH_2}OH$), 3.80 (s, 6H, ArO$\underline{CH_3}$), 5.02 (m, 4H, $CH_2CH=\underline{CH_2}$), 5.69 (m, 2H, $CH_2\underline{CH}=CH_2$), 6.84 (m, 4H, Ar$\underline{H}$), 7.22 (m, 1H, Ar$\underline{H}$), 7.29 (m, 2H, Ar$\underline{H}$), 7.32 (m, 4H, Ar$\underline{H}$), 7.42 (m, 2H, Ar$\underline{H}$) ppm.

2.6.4 Phosphoramidite Monomer 44



To a solution of mono-protected diol **43** (3.54 g, 7.7 mmol, 1.0 eq.) in 30 mL DCM (anh.) was added DIPEA (8.5 mL, 47.2 mmol, 6.1 eq.) and 100 mg of molecular sieves (3 Å) under argon atmosphere at 0 °C. To the stirring reaction mixture 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (1.89 g, 8.0 mmol, 1.04 eq.) was then added. The reaction mixture was stirred at 0 °C for another hour and then allowed to warm to RT for 1 h. The solvent was removed and the residue purified by column chromatography (SiO_2 vs 20 v% EtOAc in cyclohexane + 2 v% DIPEA) to obtain 4.80 g (7.3 mmol, 94 %) of the product as a colorless foam.

^{1}H NMR (400.1 MHz, $(CD_3)_2CO$, δ, ppm): 1.18 (m, 12H, $N(CH_2(\underline{CH_3})_2)_2$), 2.14 (m, 4H, $\underline{CH_2}CH=CH_2$), 2.71 (m, 2H, $CH_2\underline{CH_2}C\equiv N$), 3.01 (m, 2H, $\underline{CH_2}ODMT$), 3.69 (m, 6H, $\underline{CH_2}CH_2C\Xi N$, $\underline{CH_2}OP$, $N(\underline{CH_2}(CH_3)_2)_2$), 3.79 (s, 6H, $ArO\underline{CH_3}$), 4.98 (m, 4H, $CH_2CH=\underline{CH_2}$), 5.69 (m, 2H, $CH_2\underline{CH}=CH_2$), 6.89 (m, 4H, Ar$\underline{H}$), 7.22 (m, 1H, Ar$\underline{H}$), 7.32 (m, 2H, Ar$\underline{H}$), 7.35 (m, 4H, Ar$\underline{H}$), 7.49 (m, 2H, Ar$\underline{H}$); ^{31}P NMR (161.9 MHz, $(CD_3)_2CO$, δ, ppm): 146.9.

2.7 Synthesis of Sequences S9 – S14 for Analysis by Nanopore

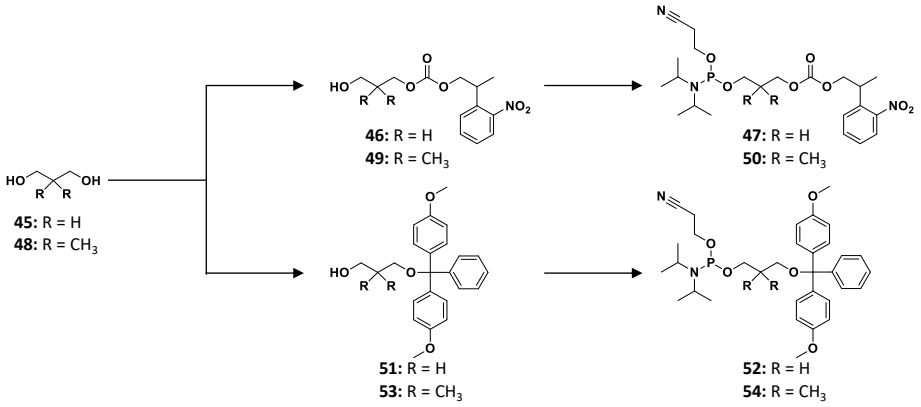
Poly(phosphodiester)s **S9**-**S13** were synthesized according to the general procedure described in *Automated Poly(phosphodiester) Synthesis* (pp. 135ff.) and purified by DMT-ON purification described in subchapter *Manual DMT-ON Purification* (p. 137).

S14 was synthesized starting from CPG-immobilized precursor **S13** that was modified with diphenyl phosphoryl azide in an adapted experimental setup from the procedure described before in *Polymer Modification* pp. (150f.). The cleavage conditions using AMA were thereby sufficient to quantitatively hydrolyze the P-N bond. A DMT-ON purification described in subchapter *Manual DMT-ON Purification* (p. 137) yields the modified sequence **S14**.

3 Experimental Protocols Chapter III

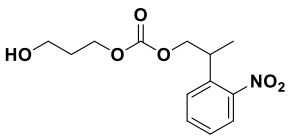
3.1 Monomer Synthesis

The synthesis of monomers **47**, **50**, **52**, and **54** following the strategy depicted in Scheme V-4 is described herein.



Scheme V-4: Overview synthetic routes of monomer synthesis in Chapter III

3.1.1 NPPOC-protected alcohol 46

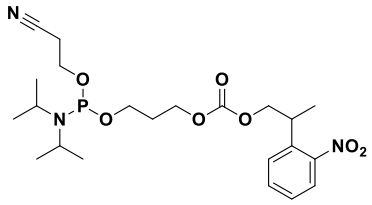


Protocol 1: To a stirred solution of 1,3-propanediol **45** (920.8 mg, 12.1 mmol, 1.2 eq.) in 10 mL anhydrous pyridine, NPPOC chloride (2506.3 mg, 9.8 mmol, 1.0 eq.) in 20 mL anh. DCM was added dropwise at RT and under argon atmosphere. The reaction mixture was stirred overnight until the reaction was quenched by addition of MeOH. The solvent was removed, 50 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (3x 50 mL). The combined organic layer was washed with 50 mL $NaHCO_3$ (sat., aq.), brine (2x 25 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 20 v% EtOAc in cyclohexane → 100 v% EtOAc) to obtain 2418.0 mg (8.5 mmol, 87 %) of the product as a pale yellow oil.

^{1}H NMR (400 MHz, $CDCl_3$, δ, ppm): 1.30 (m, 3H, $OCH_2CH\underline{CH_3}$), 1.80 (m, 2H, $HOCH_2\underline{CH_2}CH_2ONPPOC$), 2.58 (bs, 1H, $\underline{HO}CH_2$), 3.60 (m, 2H, $HO\underline{CH_2}$), 3.66 (m, 1H, $OCH_2\underline{CH}CH_3$), 4.17 (m, 2H, $\underline{CH_2}ONPPOC$), 4.24 (m, 2H, $O\underline{CH_2}CHCH_3$), 7.31 (m, 1H, Ar$\underline{H}$), 7.43 (m, 1H, Ar$\underline{H}$), 7.52 (m, 1H, Ar$\underline{H}$), 7.70 (m, 1H, Ar$\underline{H}$); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 17.6, 31.5, 33.2, 58.7, 65.1, 71.4, 124.2, 127.5, 128.2, 132.7, 136.8, 150.1, 155.1.

Protocol 2: To a stirred solution of 1,3-propanediol **45** (3.97 g, 50 mmol, 5.1 eq.) in 10 mL anhydrous pyridine : anhydrous DCM (1 : 1; v : v), NPPOC chloride (2.48 g, 9.7 mmol, 1.0 eq.) in 40 mL anh. DCM was added dropwise at RT and under argon atmosphere. The reaction mixture was stirred overnight until the reaction was quenched by addition of MeOH. The solvent was removed, 50 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (3x 50 mL). The combined organic layer was washed with 100 mL $NaHCO_3$ (sat., aq.), brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 33 v% → 50 v% EtOAc in cyclohexane) to obtain 2632.8 mg (9.3 mmol, 96 %) of the product as a pale yellow oil.

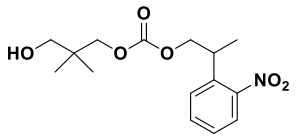
3.1.2 Phosphoramidite monomer 47



To a solution of mono-protected diol **46** (2.63 g, 9.3 mmol, 1.0 eq.) in 20 mL anhydrous DCM, were added DIPEA (13.80 g, 58.3 mmol, 6 eq.) and molecular sieves (3 Å) under argon atmosphere at 0 °C. To the stirring reaction mixture 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (2.30 g, 9.7 mmol, 1.05 eq.) was then added. The reaction mixture was stirred at 0 °C for 1 h and then allowed to warm to RT for 30 min. The solvent was removed and the residue purified by column chromatography (SiO_2 vs 33 v% EtOAc in cyclohexane + 2 v% DIPEA) to obtain 4.1856 mg (8.7 mmol, 94 %) of the product as a pale yellow oil.

HRMS m/z: $[M+H]^+$ calculated for $C_{22}H_{35}N_3O_7P^+$ 484.2207; found, 484.2206; ^{1}H NMR (400 MHz, $CDCl_3$, *δ*, ppm): 1.11 (m, 12H, N(CH($\underline{CH_3}$)$_2$)$_2$), 1.31 (m, 3H, $OCH_2CH\underline{CH_3}$), 1.88 (m, 2H, $POCH_2\underline{CH_2}CH_2ONPPOC$), 2.57 (m, 2H, $POCH_2\underline{CH_2}CN$), 3.65 (m, 7H, $PO\underline{CH_2}CH_2CN$, $OCH_2\underline{CH}CH_3$, N($\underline{CH}(CH_3)_2$)$_2$, $PO\underline{CH_2}CH_2CH_2ONPPOC$), 4.16 (m, 2H, $POCH_2CH_2\underline{CH_2}ONPPOC$), 4.24 (m, 2H, $O\underline{CH_2}CHCH_3$), 7.32 (m, 1H, Ar$\underline{H}$), 7.43 (m, 1H, Ar$\underline{H}$), 7.52 (m, 1H, Ar$\underline{H}$), 7.70 (m, 1H, Ar$\underline{H}$); ^{13}C NMR (100.6 MHz, $CDCl_3$, *δ*, ppm): 17.7, 20.3, 24.5, 30.2, 33.2, 42.9, 58.3, 59.6, 64.9, 71.2, 117.6, 124.2, 127.5, 128.2, 132.7, 136.8, 150.2, 154.9; ^{31}P NMR (161.9 MHz, $CDCl_3$, *δ*, ppm): 147.7.

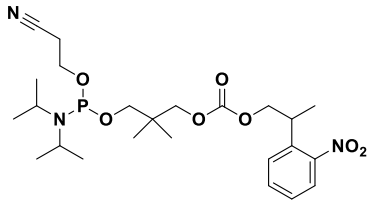
3.1.3 NPPOC-protected alcohol 49



To a stirred solution of 2,2-dimethyl-1,3-propanediol **48** (1249.0 mg, 12.0 mmol, 1.2 eq.) in 10 mL anhydrous pyridine, NPPOC chloride (2558.5 mg, 10.0 mmol, 1.0 eq.) in 20 mL anhydrous DCM was added dropwise at RT and under argon atmosphere. The reaction mixture was stirred overnight until the reaction was quenched by addition of MeOH. The solvent was removed, 50 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (3x 50 mL). The combined organic layer was washed with 50 mL $NaHCO_3$ (sat., aq.), brine (2x 25 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 10 v% → 50 v% EtOAc in cyclohexane) to obtain 2604.6 mg (8.4 mmol, 84 %) of the product as a pale yellow oil.

HRMS m/z: $[M+H]^+$ calcd for $C_{15}H_{22}N_3O_6^+$ 321.1442; found, 321.1452; ^{1}H NMR (400 MHz, $CDCl_3$, *δ*, ppm): 0.91 (s, 6H, $HOCH_2C(\underline{CH_3})_2CH_2ONPPOC$), 1.38 (m, 3H, $OCH_2CH\underline{CH_3}$), 1.99 (bs, 1H, $\underline{HO}CH_2$), 3.31 (m, 2H, $HO\underline{CH_2}$), 3.77 (m, 1H, $OCH_2\underline{CH}CH_3$), 3.95 (m, 2H, $\underline{CH_2}ONPPOC$), 4.31 (m, 2H, $O\underline{CH_2}CHCH_3$), 7.38 (m, 1H, $\underline{ArH}$), 7.48 (m, 1H, Ar$\underline{H}$), 7.58 (m, 1H, Ar$\underline{H}$), 7.78 (m, 1H, Ar$\underline{H}$); ^{13}C NMR (100.6 MHz, $CDCl_3$, *δ*, ppm): 17.8, 21.4, 33.4, 36.6, 68.2, 71.8, 73.1, 124.5, 127.8, 128.4, 132.9, 137.0, 150.5, 155.8.

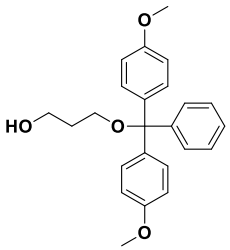
3.1.4 Phosphoramidite monomer 50



To a solution of mono-protected diol **49** (2600.0 mg, 8.4 mmol, 1.0 eq.) in 15 mL anhydrous DCM, were added DIPEA (6.69 mL, 51.8 mmol, 6 eq.) and molecular sieves (3 Å) at 0 °C under argon atmosphere. 2-Cyanoethyl-N,N-diisopropylchlorophosphoramidite (2107.7 mg, 8.8 mmol, 1.05 eq.) was then added to the stirring mixture. The reaction mixture was stirred at 0 °C for another hour and then allowed to warm to RT for 30 min. The solvent was removed and the residue purified by column chromatography (SiO_2 vs 33 v% EtOAc in cyclohexane + 2 v% DIPEA) to obtain 3238.8 mg (6.3 mmol, 76 %) of the product as a colorless oil.

HRMS m/z: $[M+H]^+$ calcd for $C_{24}H_{39}N_3O_7P^+$ 512.2520; found, 512.2509; ^{1}H NMR (400 MHz, $CDCl_3$, *δ*, ppm): 0.93 (s, 6H, $POCH_2C(\underline{CH_3})_2CH_2ONPPOC$), 1.17 (m, 12H, $N(CH(\underline{CH_3})_2)_2$), 1.38 (m, 3H, $OCH_2CH\underline{CH_3}$), 2.62 (m, 2H, $OCH_2\underline{CH_2}CN$), 3.37 (m, 2H, $PO\underline{CH_2}C(CH_3)_2CH_2ONPPOC$), 3.70 (m, 5H, $PO\underline{CH_2}CH_2CN$, $OCH_2\underline{CH}CH_3$, $N(\underline{CH}(CH_3)_2)_2$), 3.92 (m, 2H, $POCH_2C(CH_3)_2\underline{CH_2}ONPPOC$), 4.29 (m, 2H, $O\underline{CH_2}CHCH_3$), 7.38 (m, 1H, Ar<u>H</u>), 7.48 (m, 1H, Ar<u>H</u>), 7.57 (m, 1H, Ar<u>H</u>), 7.78 (m, 1H, Ar<u>H</u>); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 18.0, 20.5, 21.7, 24.7, 33.5, 36.1, 43.3, 58.4, 69.1, 71.5, 73.1, 117.8, 124.4, 127.7, 128.4, 132.8, 137.1, 150.4, 155.3; ^{31}P NMR (161.9 MHz, $CDCl_3$, δ, ppm): 148.0.

3.1.5 DMT-protected alcohol 51

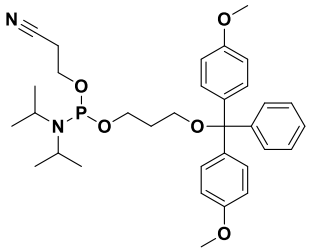


To a stirred solution of 1,3-propanediol **45** (915.0 mg, 12.0 mmol, 1.2 eq.) in 75 mL THF (anh.) : pyridine (anh.) = 3 :2 (v : v) 4,4'-dimethoxytrityl chloride (3.39 g, 10.0 mmol, 1.0 eq.) was added in four intervals of each 40 min at RT and under argon atmosphere. The reaction mixture was stirred overnight until the reaction was quenched by the addition of MeOH. The solvent was removed, 200 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (4x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 20 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 2.566 g (5.6 mmol, 89 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.86 (m, 2H, $\underline{CH_2}CH_2OH$), 2.25 (m, 1H, $CH_2\underline{OH}$), 3.29 (m, 2H, $\underline{CH_2}ODMT$), 3.78 (m, 2H, $\underline{CH_2}OH$), 3.79 (s, 6H, $ArO\underline{CH_3}$), 6.84 (m, 4H, Ar<u>H</u>), 7.22 (m, 1H, Ar<u>H</u>), 7.30 (m, 2H, Ar<u>H</u>), 7.33 (m, 4H, Ar<u>H</u>), 7.44 (m, 2H, Ar<u>H</u>).

The analysis data harmonizes with the reported values in the literature.[347]

3.1.6 Phosphoramidite Monomer 52

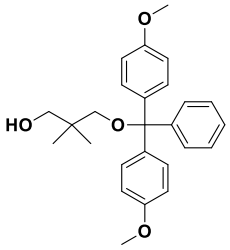


To a solution of mono-protected diol **51** (2.47 g, 6.5 mmol, 1.0 eq.) in 15 mL anhydrous DCM, were added DIPEA (5.07 g, 39.2 mmol, 6 eq.) and molecular sieves (3 Å) under argon atmosphere at 0 °C. To the stirring reaction mixture 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (1.62 g, 6.8 mmol, 1.05 eq.) was then added. The reaction mixture was stirred at 0 °C for 1 h and then allowed to warm to RT for 30 min. The solvent was removed and the residue purified by column chromatography (SiO_2 vs 25 v% EtOAc in cyclohexane + 2 v% DIPEA) to obtain 3524.0 mg (6.1 mmol, 93 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.15 (m, 12H, N(CH($\underline{CH_3}$)$_2$)$_2$), 1.93 (m, 2H, $\underline{CH_2}CH_2$ODMT), 2.55 (m, 2H, OCH$_2\underline{CH_2}$CN), .317 (m, 2H, $\underline{CH_2}$ODMT), 3.50-3.86 (m, 6H, O$\underline{CH_2}CH_2$CN, N($\underline{CH}$(CH$_3$)$_2$)$_2$, PO$\underline{CH_2}$C(CH$_3$)$_2$), 3.79 (s, 6H, ArO$\underline{CH_3}$), 6.82 (m, 4H, Ar$\underline{H}$), 7.20 (m, 1H, Ar$\underline{H}$), 7.28 (m, 2H, Ar$\underline{H}$), 7.32 (m, 4H, Ar$\underline{H}$), 7.44 (m, 2H, Ar$\underline{H}$).

The analysis data harmonizes with the reported values in the literature.[347]

3.1.7 DMT-protected alcohol 53

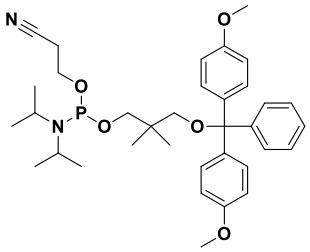


To a stirred solution of 2,2-dimethyl-1,3-propanediol **48** (1.25 g, 12.0 mmol, 1.2 eq.) in 20 mL THF (anh.) : pyridine (anh.) = 3 :2 (v : v) 4,4'-dimethoxytrityl chloride (3.39 g, 10.0 mmol, 1.0 eq.) was added in four intervals of each 40 min at RT and under argon atmosphere. The reaction mixture was stirred overnight until the reaction was quenched by the addition of MeOH. The solvent was removed, 200 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (4x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 20 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 2.86 g (7.6 mmol, 76 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 0.93 (s, 6H, C(CH$_3$)$_2$), 2.23 (t, ^3J = 6.1 Hz, 1H, CH$_2\underline{OH}$), 2.96 (s, 2H, $\underline{CH_2}$ODMT), 3.42 (d, ^3J = 6.1 Hz, 2H, $\underline{CH_2}$OH), 3.79 (s, 6H, ArO$\underline{CH_3}$), 6.84 (m, 4H, Ar$\underline{H}$), 7.21 (m, 1H, Ar$\underline{H}$), 7.28 (m, 2H, Ar$\underline{H}$), 7.32 (m, 4H, Ar$\underline{H}$), 7.43 (m, 2H, Ar$\underline{H}$).

The analysis data harmonizes with the reported values in the literature[10] where spectra were recorded in deuterated acetone.

3.1.8 Phosphoramidite Monomer 54



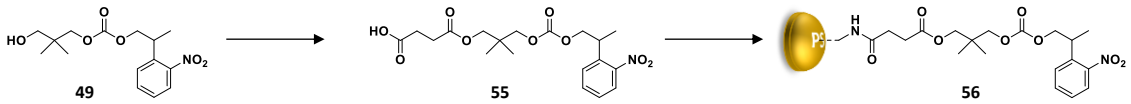
To a solution of mono-protected diol **53** (2.62 g, 6.4 mmol, 1.0 eq.) in 16 mL anhydrous DCM, were added DIPEA (5.1 g, 39.3 mmol, 6.1 eq.) and molecular sieves (3 Å) at 0 °C under argon atmosphere. 2-Cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (1.60 g, 6.8 mmol, 1.05 eq.) was then added to the stirring mixture. The reaction mixture was stirred at 0 °C for another hour and then allowed to warm to RT for 30 min. The solvent was removed and the residue purified by column chromatography (SiO_2 vs 20 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 3.76 g (6.2 mmol, 96 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 0.93 (m, 6H, C($\underline{CH_3}$)$_2$), 1.14 (m, 12H, N(CH($\underline{CH_3}$)$_2$)$_2$), 2.52 (m, 2H, OCH$_2\underline{CH_2}$CN), 2.88 (m, 2H, $\underline{CH_2}$ODMT), 3.35-3.73 (m, 6H, O$\underline{CH_2}$CH$_2$CN, N($\underline{CH}$(CH$_3$)$_2$)$_2$, PO$\underline{CH_2}$C(CH$_3$)$_2$), 3.79 (s, 6H, ArO$\underline{CH_3}$), 6.81 (m, 4H, Ar$\underline{H}$), 7.19 (m, 1H, Ar$\underline{H}$), 7.27 (m, 2H, Ar$\underline{H}$), 7.31 (m, 4H, Ar$\underline{H}$), 7.43 (m, 2H, Ar$\underline{H}$).

The analysis data harmonizes with the reported values in the literature[10] where spectra were recorded in deuterated acetone.

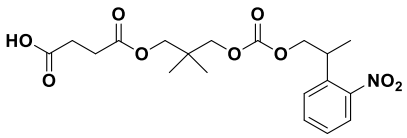
3.2 Support Synthesis

Synthetic access to solid polystyrene support **56** following Scheme V-5 is described below.



Scheme V-5: Overview of solid polystyrene support synthesis.

3.2.1 Succinic ester 55

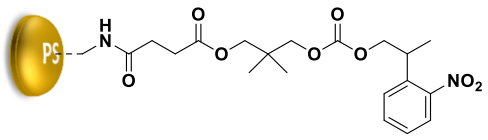


Precursor alcohol **49** (357.8 mg, 1.2 mmol, 1.0 eq.) was co-evaporated with 5 mL of pyridine (anh.), dried under vacuum for 30 min, and reacted with succinic anhydride (273.9 mg, 2.7 mmol, 2.5 eq.) and DMAP (211.5 mg, 1.5 mmol, 1.3 eq.) in 5 mL of pyridine (anh.) at 35 °C overnight and allowed to cool to RT. The reaction was quenched with methanol, and the solvent was removed. The resulting slurry oil was taken in 40 mL EtOAc, washed with 20 mL H_2O, 20 mL K_2HPO_4 (sat., aq., pH 5.0) and brine (3x 20 mL). The organic phase was dried over Na_2SO_4, filtered and the solvent was removed to obtain 481.9 mg (1.2 mmol, quant.) of the product as a yellow oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 0.88 (s, 6H, CH$_2$C($\underline{CH_3}$)$_2$CH$_2$), 1.30 (m, 3H, OCH$_2$CH$\underline{CH_3}$), 2.59 (m, 4H, $\underline{CH_2}$COO), 3.66 (m, 1H, OCH$_2\underline{CH}$CH$_3$), 3.64 (m, 4H, $\underline{CH_2}$OCO), 4.23 (m, 2H, O$\underline{CH_2}$CHCH$_3$), 7.31 (m, 1H, ArH), 7.43 (m, 1H, ArH), 7.52 (m, 1H, ArH), 7.69 (m, 1H, ArH), 10.77 (bs, 1H, $\underline{COOH}$); ^{13}C NMR

(100.6 MHz, $CDCl_3$, δ, ppm): 17.63, 21.3, 28.7, 28.8, 33.2, 34.8, 69.2, 71.4, 72.6, 124.2, 127.5, 128.2, 132.7, 136.7, 150.2, 155.0, 171.7, 177.8.

3.2.2 Polystyrene Support 56



To a homemade peptide synthesis reaction vessel equipped with a fritted glass plate and a stopcock, aminomethylated polystyrene resin (795.0 mg, loading: 1.4 mmol·g^{-1}, 1.1 mmol functionalities, 1.0 eq.) was charged and washed with DCM (anh.) (2x 5 mL), dried under vacuum and purged with argon. A solution of succinic ester **55** (481.9 mg, 1.2 mmol, 1.1 eq.), DMAP (210.3 mg, 1.7 mmol, 1.5 eq.), and DCC (380.0 mg, 1.7 mmol, 1.5 eq.) in 10 mL DCM (anh.) was prepared under argon atmosphere and transferred to the reaction vessel which was then shaken overnight at RT. The resin was then washed with DCM (3x 5 mL), MeOH (3x 5 mL), MeCN (3x 5 mL), and Et_2O (3x 5 mL) and dried under vacuum for 3 h. The unreacted amino groups were capped with 2 mL acetic anhydride in 8 mL pyr (anh.) while shaking for 2 h. The resin was finally washed with DCM (3x 5 mL), MeCN (3x 5 mL), and Et_2O (3x 5 mL) and dried under vacuum overnight.

The loading of the support was determined by the added weight during immobilization to find a loading of 0.65 mmol·g^{-1}.

3.3 Photocontrolled Liquid-phase Oligo(phosphodiester) Synthesis

The coupling and oxidation steps were adapted from a previous procedure[349] and reagents and solutions that were described earlier (see Experimental Section: 2.4.1 Manual Synthesis on Solid Polystyrene Supports, p. 149f.) were employed. In brief, the hydroxyl-terminated polymer **57** was charged to a round-bottom flask under argon atmosphere. 2 Equiv. of a phosphoramidite monomer (either monomer **47**, **50**, **52**, or **54**) dissolved in anhydrous DCM (1 mL per 100 mg polymer) were added under argon. The resulting solution was stirred and 8 equiv. of tetrazole (0.45 M in MeCN) were added and the reaction mixture was stirred at RT for 15 min. Then, oxidizing solution (1 mL per 100 mg polymer) was added and the reaction mixture was left stirring for another 15 min. After removal of solvent, the slurry was taken in DCM (1 mL per 100 mg polymer) and precipitated in ice-cold methanol (10 mL per 100 mg polymer). The mixture was centrifuged, the supernatant liquid decanted off and the precipitate was washed with ice-cold MeOH.

For the photodeprotection step, a solution of 2 mg·mL^{-1} NPPOC-terminated polymer in 50 mM piperidine in $CHCl_3$ was prepared under argon atmosphere. Under exclusion of air, vials were charged with such solution to a max. 3 cm fill level. The vials were placed on the glass surface of a UV lamp (Vilber Lourmat). The solution was irradiated for 180 min, then combined and neutralized by dropwise adding of a solution of 3 w% TCA in DCM. The solvent was removed and the slurry was taken in DCM (1 mL per 100 mg polymer). The formed polystyrene-*b*-oligo(phosphodiester) conjugate was then precipitated in MeOH as described above. Cycles of coupling, oxidation, and photodeprotection steps were repeated until the desired sequences were obtained.

The polystyrene-*b*-oligo(phosphodiester) conjugate was dissolved in 1,4-dioxane (1 mL per 100 mg polymer) and stirred. Then, a 1 : 1 (v : v) mixture of NH_3 (aq., 28 w%) and $MeNH_2$ (aq., 40 w%) (1 mL per 100 mg polymer) was added. After 30 min, the solvent was removed and by means of sonication the slurry was taken in acetonitrile (1 mL per 100 mg polymer) and precipitated in ice-cold methanol (10 mL per 100 mg polymer). During centrifugation/methanol washing steps, the supernatant liquid

was collected and the solvent of the combined phase was removed to obtain the desired sequences **S17**, **S18**, and **S23**.

When using DMT-protected monomer **52** or **54** in the last coupling step of the iterative process, the formed polystyrene-*b*-oligo(phosphodiester) conjugate contains a terminal DMT group allowing reverse phase purification. After cleavage from the soluble support, the DMT-capped oligo(phosphodiester) was dissolved in a mixture of 1 mL NH_3 (aq., 28 w%) and 1 mL NaCl solution (aq., 10 w%) and it was subject to a DMT-ON purification (see Experimental Section: 1.2.3.2 Manual DMT-ON Purification, p. 137) to obtain purified sequences **S19**-**S22**.

The scale of synthesis as well as the yields that were obtained during the synthesis of **S17**-**S23** are listed below in Table V-1 and Table V-2.

*Table V-1: Yields obtained during synthesis of **S17**-**S22** for coupling/oxidation i)/ii), photodeprotection iii) and cleavage iv).*

	cycle 1		**cycle 2**		**cycle 3**		**cleavage**
	i) ii)	*iii)*	*i) ii)*	*iii)*	*i) ii)*	*iii)*	*iv)*
S17	662.7 mg 97 %	605.7 mg 97 %	606.8 mg 96 %	545 mg 97 %	133.8 mg 90 %	117.7 mg 97 %	4.3 mg 53 %
S23	*same as **S17***	*same as **S17***	*same as **S17***	*same as **S17**.*	409.6 mg 95 %	360.4 mg 95 %	-
S18	309.0 92 %	353.8 mg 90 %	241.0 mg 93 %	204.7 mg 90 %	56.0 mg 84 %	45.9 mg 89 %	4.3 mg 84 %
S19	*same as **S18***	*same as **S18***	*same as **S18***	*same as **S18***	65.3 mg 84 %	-	7.0 mg 99 %
S20	*same as **S18***	*same as **S18***	*same as **S18***	*same as **S18***	58.8 mg 76 %	-	5.2 mg 80 %
S21	*same as **S18**.*	*same as **S18***	98.0 mg 76 %	78.0 mg 90 %	30.0 mg 75 %	-	2.4 mg 83 %
S22	*same as **S18***	*same as **S18***	*same as **S21***	*same as **S18***	24.4 mg 57 %	-	2.1 mg 87 %

*Table V-2: Yields obtained during synthesis of **S23** for coupling/oxidation i)/ii), photodeprotection iii) and cleavage iv).*

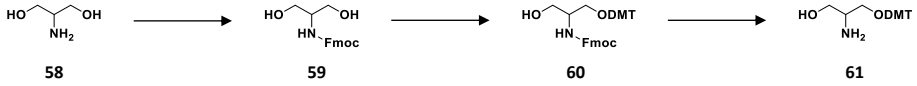
	cycle 4		**cycle 5**		**cleavage**
	i) ii)	*iii)*	*i) ii)*	*iii)*	*iv)*
S23	326.4 mg 99 %	297.7 mg 99 %	253.5 mg 80 %	215.7 mg 96 %	8.3 mg 80 %

4 Experimental Protocols Chapter IV

4.1 Monomer Synthesis Targeting Nitrobenzylamines

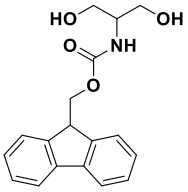
4.1.1 Synthesis of a DMT-protected Precursor

The synthesis of a DMT-protected precursor molecule **61** is shown in Scheme V-6 and followingly described.



*Scheme V-6: Overview synthesis of a DMT-protected precursor **61** for a subsequent nitrobenzylation reaction.*

4.1.1.1 Fmoc-protected Serinol 59

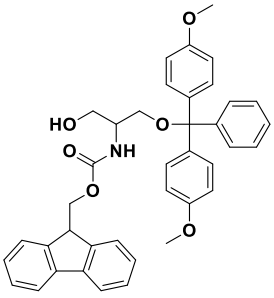


To a stirred solution of serinol **58** (2.77 g, 30.4 mmol, 1.0 eq.) in 75 mL Na_2CO_3 (aq., 10 w%) and 40 mL 1,4-dioxane was added Fmoc-Cl (7.96 g, 30.4 mmol) at 0 °C and it was stirred overnight and thereby the reaction mixture was allowed to reach RT. The solvent was removed and the residue was taken in 50 mL EtOAc, 100 mL H_2O was added and it was extracted with EtOAc (5x 100 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The product was purified by recrystallization in 1,4-dioxane over cyclohexane to obtain 7.42 g (23.7 mmol, 78 %) of the product as colorless cotton-like crystals.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 2.23 (bs, 2H, $CH_2\underline{OH}$); 3.73 (m, 1H, $\underline{CH}N$), 3.83 (m, 4H, $\underline{CH_2}OH$), 4.22 (t, ^3J = 6.7 Hz, 1H, $\underline{CH}CH_2OCON$), 4.45 (d, ^3J = 6.7 Hz, 2H, $CH\underline{CH_2}OCON$), 5.43 (bs, 1H, $OCO\underline{NH}CH$), 7.32 (m, 2H, ArH), 7.41 (m, 2H, ArH), 7.60 (m, 2H, ArH), 7.77 (m, 2H, ArH).

The analysis data harmonizes with the reported values in the literature.[383]

4.1.1.2 Fmoc-protected Precursor 60



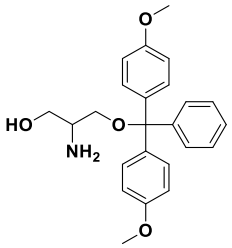
To a stirred solution of Fmoc-protected serinol **59** (1147.6 mg, 3.7 mmol, 1.0 eq.) in 15 mL THF (anh.) : pyridine (anh.) = 3 : 2 (v : v) 4,4'-dimethoxytrityl chloride (1254.6 mg, 3.7 mmol, 1.0 eq.) was added in three intervals of each 40 min at RT and under argon atmosphere and it was stirred for 2.5 h. The solvent was removed, 100 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (4x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs

50 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 1361.6 mg (2.2 mmol, 60 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.05 (m, 1H, $CH_2\underline{OH}$), 3.33 ($\underline{CH_2}$ODMT), 3.77 (m, 6H, $Ar_{DMT}O\underline{CH_3}$), 3.84 (m, 3H, $\underline{CH_2}$OH, $\underline{CH}$NCOO), 4.22 (t, ^3J = 7.2 Hz, 1H, $\underline{CH}CH_2OCON$), 4.39 (d, ^3J = 7.2 Hz, 2H, $CH\underline{CH_2}OCON$), 5.34 (m, 1H, $CH\underline{NH}COO$), 6.83 (m, 4H, $Ar_{DMT}H$), 7.22 (m, 1H, $Ar_{DMT}H$), 7.27 (m, 2H, $Ar_{DMT}H$), 7.31 (m, 6H, $Ar_{Fmoc}H$, $Ar_{DMT}H$), 7.40 (m, 4H, $Ar_{Fmoc}H$, $Ar_{DMT}H$), 7.60 (m, 2H, $Ar_{Fmoc}H$), 7.77 (m, 2H, $Ar_{Fmoc}H$).

The analysis data harmonizes with the reported values in the literature.[365]

4.1.1.3 DMT-protected Precursor 61



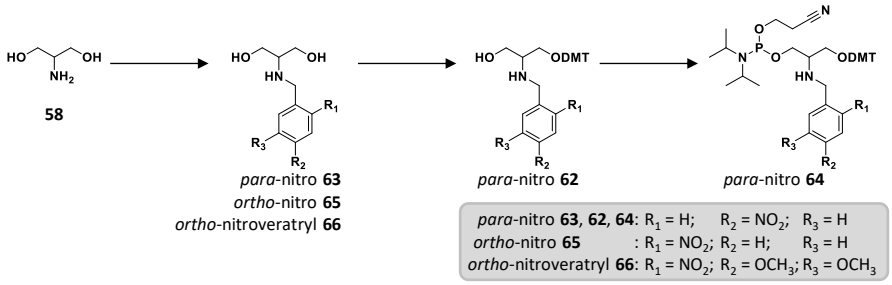
Fmoc-protected precursor **60** (490 mg, 0.8 mmol) was dissolved in 20 v% piperidine in DCM and stirred for 60 min at RT. The solvent was removed and the residue purified by column chromatography (SiO_2 vs 2.5 v% → 10 v% MeOH in DCM + 3 v% DIPEA) to obtain 257.9 mg (0.7 mmol, 82 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 3.08 (m, 1H, $\underline{CH}NH_2$); 3.13 ($\underline{CH_2}$ODMT), 3.56 (m, 2H, $\underline{CH_2}$OH), 3.78 (m, 6H, $Ar_{DMT}O\underline{CH_3}$), 6.83 (m, 4H, $Ar_{DMT}H$), 7.21 (m, 1H, $Ar_{DMT}H$), 7.27 (m, 2H, $Ar_{DMT}H$), 7.30 (m, 4H, $Ar_{DMT}H$), 7.41 (m, 2H, $Ar_{DMT}H$).

The analysis data harmonizes with the reported values in the literature.[365]

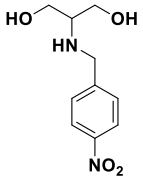
4.1.2 Synthesis of Nitrobenzylamine-containing Phosphoramidite Monomers

The synthesis of nitrobenzylamine-containing phosphoramidite monomers shown in Scheme V-6 is described in the following section.



*Scheme V-7: Overview synthesis of nitrobenzylamine-containing phosphoramidite monomer **64** and different nitrobenzylated serinols **63**, **65**, and **66**.*

4.1.2.1 Nitrobenzylated Serinol 63

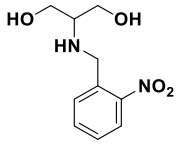


A suspension of serinol **58** (538.1 mg, 5.9 mmol, 1.0 eq.), 4-nitrobenzaldehyde **67** (890.7 mg, 5.9 mmol, 1.0 eq.) and Na_2SO_4 (3382.6 mg, 23.8 mmol, 4.0 eq.) in MeOH was stirred at RT overnight. At 0 °C $NaBH_4$ (238.0 mg, 5.9 mmol, 1.0 eq.) was added to that reaction mixture and after 60 min (after full conversion of the intermediate imine as monitored by TLC) the reaction mixture was filtered. The solvent was removed from the filtrate and the residue was purified by column chromatography (SiO_2 vs 3 v% → 10 v% MeOH in DCM + 2 v% DIPEA) to obtain 1149.6 mg (5.1 mmol, 86 %) of the product as a yellow crystalline solid.

^{1}H NMR (400.1 MHz, CD_3OD, *δ*, ppm): 2.72 (m, 1H, C<u>H</u>CH_2OH), 3.60 (m, 4H, CH<u>CH_2</u>OH), 3.99 (s, 2H, N<u>CH_2</u>Ph), 7.63 (m, 2H, ArH), 8.20 (m, 2H, ArH); ^{13}C NMR (100.6 MHz, CD_3OD, δ, ppm): 51.4, 61.3, 62.5, 124.5, 130.2, 148.4, 149.4.

The analysis data harmonizes with the reported values in the literature[384] where the NMR was recorded in $CDCl_3$.

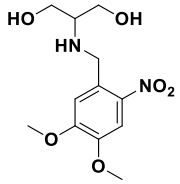
4.1.2.2 Nitrobenzylated Serinol 65



A suspension of serinol **58** (548.0 mg, 6.0 mmol, 1.0 eq.), 2-nitrobenzaldehyde **68** (927.7 mg, 6.0 mmol, 1.0 eq.) and Na_2SO_4 (3426.2 mg, 24.0 mmol, 4.0 eq.) in MeOH was stirred at RT overnight. At 0 °C $NaBH_4$ (250.6 mg, 6.6 mmol, 1.1 eq.) was added to that reaction mixture and after stirring overnight, the reaction mixture was filtered. The solvent was removed from the filtrate and the residue was purified by column chromatography (SiO_2 vs 3 v% → 10 v% MeOH in DCM + 2 v% DIPEA) to obtain 1203.6 mg (5.3 mmol, 88 %) of the product as a yellow crystalline solid.

^{1}H NMR (400.1 MHz, CD_3OD, *δ*, ppm): 2.72 (m, 1H, C<u>H</u>CH_2OH), 3.57 (m, 4H, CH<u>CH_2</u>OH), 4.11 (s, 2H, N<u>CH_2</u>Ph) , 7.50 (m, 1H, ArH) , 7.66 (m, 1H, ArH) , 7.70 (m, 1H, ArH), 7.99 (m, 1H, ArH); ^{13}C NMR (100.6 MHz, CD_3OD, δ, ppm): 49.2, 61.3, 62.5, 125.8, 129.5, 132.9, 134.4, 135.8, 150.5.

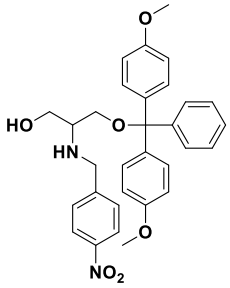
4.1.2.3 Nitrobenzylated Serinol 66



A suspension of serinol **58** (1132.7 mg, 12.4 mmol, 1.1 eq.), 2-nitroveratraldehyde **69** (2392.7 mg, 11.3 mmol, 1.0 eq.) and Na_2SO_4 (6639.2 mg, 46.7 mmol, 4.1 eq.) in MeOH was stirred at RT overnight. At 0 °C $NaBH_4$ (463.3 mg, 12.2 mmol, 1.1 eq.) was added to that reaction mixture and after 60 min (after full conversion of the intermediate imine as monitored by TLC) the reaction mixture was filtered. The solvent was removed from the filtrate and the residue was purified by column chromatography (SiO_2 vs 3 v% → 10 v% MeOH in DCM + 2 v% DIPEA) to obtain 2578.1 mg (9.0 mmol, 80 %) of the product as a yellow crystalline solid.

^{1}H NMR (400.1 MHz, CD_3OD, *δ*, ppm): 2.74 (m, 1H, C<u>H</u>CH_2OH), 3.59 (m, 4H, CH<u>CH_2</u>OH), 3.90 (s, 3H, ArO<u>CH_3</u>), 3.97 (s, 3H, ArO<u>CH_3</u>), 4.12 (s, 2H, N<u>CH_2</u>Ph) , 7.27 (s, 1H, ArH) , 7.69 (s, 1H, ArH); ^{13}C NMR (100.6 MHz, CD_3OD, δ, ppm): 49.5, 56.8, 56.9, 61.3, 62.7, 109.5, 114.7, 131.4, 142.4, 149.4, 154.8.

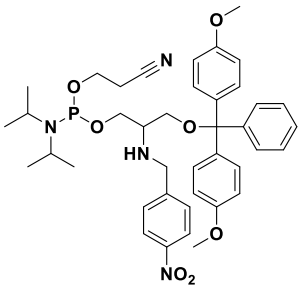
4.1.2.4 DMT-protected Nitrobenzylated Serinol 62



To a stirred solution of nitrobenzylated serinol **63** (502.2 mg, 2.2 mmol, 2.0 eq.) in 7.5 mL THF (anh.) : pyridine (anh.) = 3 : 2 (v : v) 4,4'-dimethoxytrityl chloride (376.8 mg, 1.1 mmol, 1.0 eq.) was added in four intervals of each 40 min at RT and under argon atmosphere and it was stirred overnight. The solvent was removed, 20 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (4x 20 mL). The combined organic layer was washed with brine (2x 20 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 20 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 180.4 mg (0.3 mmol, 31 %) of the product as a colorless oil. Non-reacted starting material **63** was thereby quantitatively recovered.

^{1}H NMR (400.1 MHz, CD_3OD, *δ*, ppm): 2.78 (m, 1H, $\underline{CH}CH_2OH$), 3.17 (m, 2H, $\underline{CH_2}ODMT$), 3.61 (m, 2H, $CH\underline{CH_2}OH$), 3.77 (s, 6H, $ArO\underline{CH_3}$), 3.83 (bs, 2H, $N\underline{CH_2}Ph$), 6.82 (m, 4H, $Ar_{DMT}H$), 7.19 (m, 1H, $Ar_{DMT}H$), 7.26 (m, 6H, $Ar_{DMT}H$), 7.39 (m, 2H, $Ar_{DMT}H$), 7.51 (m, 2H, $Ar_{NB}H$), 8.15 (m, 2H, $Ar_{NB}H$); ^{13}C NMR (100.6 MHz, CD_3OD, δ, ppm): 51.4, 55.7, 59.9, 63.2, 64.2, 87.5, 114.0, 124.4, 127.7, 128.7, 129.2, 130.0, 131.1, 137.2, 146.4, 148.1, 149.4, 159.9.

4.1.2.5 Nitrobenzylamine Phosphoramidite Monomer 64

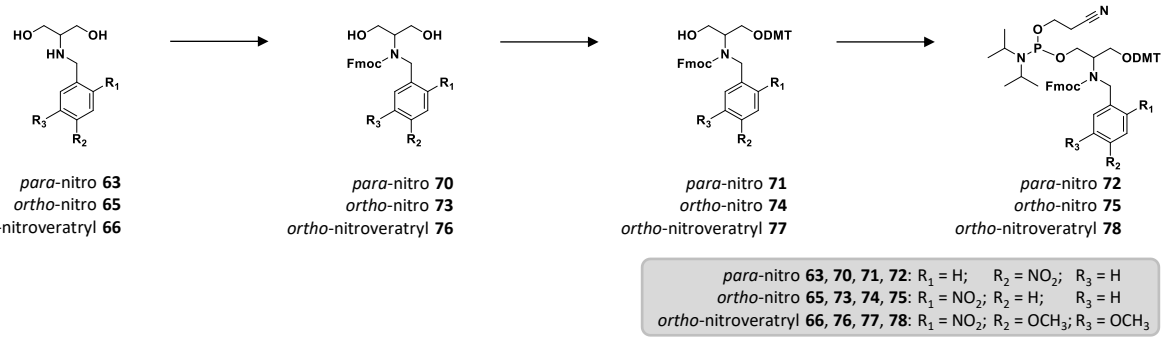


To a solution of DMT-protected nitrobenzylated Serinol **62** (600.0 mg, 1.1 mmol, 1.0 eq.) in 3 mL DCM (anh.) was added DIPEA (980.0 mg, 7.6 mmol, 6.7 eq.) and molecular sieves (3 Å) under argon atmosphere at 0 °C. To the stirring reaction mixture 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (310.5 mg, 1.3 mmol, 1.15 eq.) was then added. The reaction mixture was stirred at 0 °C for 30 min and then allowed to warm to RT for 30 min. The solvent was removed and the residue purified by column chromatography (SiO_2 vs 33 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 624.2 mg (0.9 mmol, 75 %) of the product as a colorless foam.

^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.14 (m, 12H, $N(CH(\underline{CH_3})_2)_2$), 2.55 (m, 2H, $CH_2\underline{CH_2}C{\equiv}N$), 2.91 ($\underline{CH}NHBn$), 3.19 (m, 2H, $\underline{CH_2}ODMT$), 3.65 (m, 6H, $\underline{CH_2}CH_2C{\equiv}N$, $N(\underline{CH}(CH_3)_2)_2$, $\underline{CH_2}OP$), 3.79 (s, 6H, $ArO\underline{CH_3}$), 3.84 (bs, 2H, $Ph\underline{CH_2}O$), 6.80 (m, 4H, $Ar_{DMT}H$), 7.20 (m, 1H, $Ar_{DMT}H$), 7.27 (m, 2H, $Ar_{DMT}H$), 7.29 (m, 4H, $Ar_{DMT}H$), 7.40 (m, 2H, $Ar_{DMT}H$), 7.46 (m, 2H, $Ar_{NB}H$), 8.14 (m, 2H, $Ar_{NB}H$); ^{31}P NMR (161.9 MHz, $CDCl_3$, δ, ppm): 148.1.

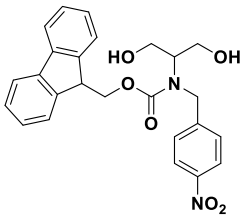
4.2 Monomer Synthesis Targeting Nitrobenzyl-protected 9-Fluorenyl Carbamates

Details to the synthesis of nitrobenzylated 9-fluorenyl carbamate phosphoramidite monomers depicted in Scheme V-8 are followingly presented.



*Scheme V-8: Overview synthesis of nitrobenzylated 9-fluorenyl carbamate phosphoramidite monomers **72**, **75**, and **78**.*

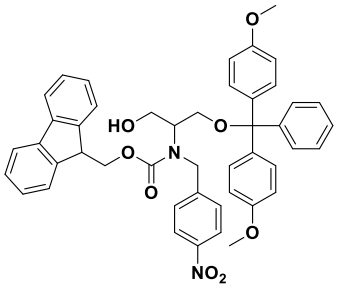
4.2.1 Nitrobenzylated Carbamate 70



To a stirred solution of nitrobenzylated serinol **63** (2.25 mg, 10 mmol mmol, 1.0 eq.) in 50 mL Na_2CO_3 (aq., 10 w%) and 100 mL 1,4-dioxane was added Fmoc-Cl (2.6 g, 10 mmol, 1.0 eq.) at 0 °C and it was stirred overnight and thereby the reaction mixture was allowed to reach RT. The solvent was removed and the residue was taken in 50 mL EtOAc, 100 mL H_2O was added and it was extracted with EtOAc (4x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the residue was purified by column chromatography (SiO_2 vs 50 v% → 80 v% EtOAc in cyclohexane) to obtain 2.59 g (5.8 mmol, 58 %) of the product as a colorless foam.

^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 3.32 (m, 1H, C<u>H</u>CH_2OH), 3.70 (m, 4H, CH<u>CH_2</u>OH), 3.78 (bs, 2H, CH_2<u>OH</u>), 4.05 (m, 1H, C<u>H</u>CH_2OCON), 4.23 (bs, 2H, N<u>CH_2</u>Ph), 4.56 (m, 2H, CH<u>CH_2</u>OCON), 7.00 (m, 2H, Ar_{Fmoc}H), 7.20 (m, 2H, Ar_{Fmoc}H), 7.33 (m, 4H, Ar_{Fmoc}H), 7.61 (m, 2H, Ar_{NB}H), 7.95 (m, 2H, Ar_{NB}H); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 47.0, 49.7, 60.7, 61.5, 66.9, 119.8, 123.4, 124.2, 126.9, 127.1, 127.6, 141.1, 143.3, 145.7, 146.7, 157.0.

4.2.2 DMT-protected Precursor 71

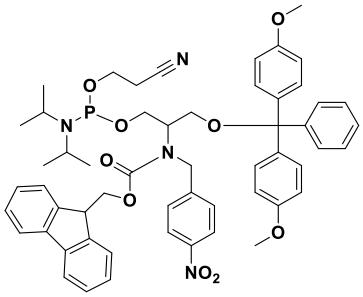


To a stirred solution of nitrobenzylated carbamate **70** (2.45 g, 5.5 mmol, 1.0 eq.) in 15 mL THF (anh.) : pyridine (anh.) = 3 : 2 (v : v) 4,4'-dimethoxytrityl chloride (1852.6 mg, 5.5 mmol, 1.0 eq.) was added in three intervals of each 40 min at RT and under argon atmosphere and it was stirred

overnight. The solvent was removed, 50 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (3x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 50 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 1629.4 mg (2.2 mmol, 40 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, CD_3OD, *δ*, ppm): 2.34 (m, 1H, $CH_2\underline{OH}$), 2.93 (m, 1H, $\underline{CH}CH_2OH$), 3.30 (m, 2H, $\underline{CH_2}ODMT$), 3.70 (m, 2H, $CH\underline{CH_2}OH$), 3.77 (s, 6H, $Ar_{DMT}O\underline{CH_3}$), 3.96 (m, 1H, $\underline{CH}CH_2OCON$), 4.20 (m, 2H, $N\underline{CH_2}Ph$), 4.64 (m, 2H, $CH\underline{CH_2}OCON$), 6.62 (m, 4H, $Ar_{DMT}H$), 6.89 (m, 2H, $Ar_{Fmoc}H$), 7.12 (m, 1H, $Ar_{DMT}H$), 7.25 (m, 12H, $Ar_{Fmoc}H$, $Ar_{DMT}H$), 7.36 (m, 2H, $Ar_{DMT}H$), 7.61 (m, 2H, $Ar_{NB}H$), 7.90 (m, 2H, $Ar_{NB}H$).

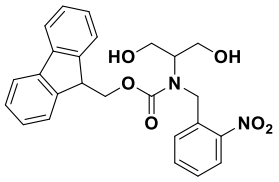
4.2.3 Nitrobenzylated Carbamate Phosphoramidite Monomer 72



To a solution of DMT-protected precursor **71** (1459.9 mg, 1.9 mmol, 1.0 eq.) in 10 mL DCM (anh.) was added DIPEA (1557.9 mg, 12.0 mmol, 6.2 eq.) and molecular sieves (3 Å) under argon atmosphere at 0 °C. To the stirring reaction mixture 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (543.1 mg, 2.3 mmol, 1.2 eq.) was then added. The reaction mixture was stirred at 0 °C for 30 min and then allowed to warm to RT for 30 min. The solvent was removed and the residue purified by column chromatography (SiO_2 vs 50 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 1602.0 mg (1.7 mmol, 87 %) of the product as a colorless foam.

HRMS m/z: $[M+H]^+$ calculated for $C_{55}H_{60}N_4O_9P^+$ 951.4092; found, 951.4091; ^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.08 (m, 12H, $N(CH(\underline{CH_3})_2)_2$), 2.50 (m, 2H, $CH_2\underline{CH_2}C{\equiv}N$), 3.07 (m, 1H, $\underline{CH}NHBn$), 3.25 (m, 2H, $\underline{CH_2}ODMT$), 3.65 (m, 6H, $\underline{CH_2}CH_2C{\equiv}N$, $N(\underline{CH}(CH_3)_2)_2$, $\underline{CH_2}OP$), 3.77 (s, 6H, $ArO\underline{CH_3}$), 4.07 (m, 1H, $\underline{CH}CH_2OCON$), 4.20 (m, 2H, $N\underline{CH_2}Ph$), 4.57 (m, 2H, $CH\underline{CH_2}OCON$), 6.77 (m, 4H, $Ar_{DMT}H$), 6.96 (m, 2H, $Ar_{Fmoc}H$), 7.18 (m, 9H, $Ar_{Fmoc}H$, $Ar_{DMT}H$), 7.32 (m, 4H, $Ar_{DMT}H$), 7.41 (m, 2H, $Ar_{DMT}H$), 7.60 (m, 2H, $Ar_{NB}H$), 7.87 (m, 2H, $Ar_{NB}H$); ^{31}P NMR (161.9 MHz, $CDCl_3$, δ, ppm): 148.3.

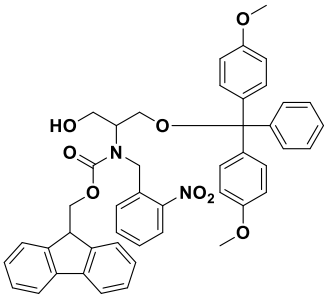
4.2.4 Nitrobenzylated Carbamate 73



To a stirred solution of nitrobenzylated serinol **65** (1169.2 mg, 5.2 mmol mmol, 1.0 eq.) in 10 mL Na_2CO_3 (aq., 10 w%) and 20 mL 1,4-dioxane was added Fmoc-Cl (1470.1 mg, 5.7 mmol, 1.1 eq.) at 0 °C and it was stirred overnight and thereby the reaction mixture was allowed to reach RT. The solvent was removed and the residue was taken in 50 mL EtOAc, 100 mL H_2O was added and it was extracted with EtOAc (4x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the residue was purified by column chromatography (SiO_2 vs 50 v% → 80 v% EtOAc in cyclohexane) to obtain 1845.0 mg (4.1 mmol, 80 %) of the product as a colorless foam.

^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 3.12 (bs, 2H, $CH_2\underline{OH}$), 3.31 (m, 1H, $\underline{CH}CH_2OH$), 3.82 (m, 4H, $CH\underline{CH_2}OH$), 4.05 (m, 1H, $\underline{CH}CH_2OCON$), 4.50 (bs, 2H, $N\underline{CH_2}Ph$), 4.74 (m, 2H, $CH\underline{CH_2}OCON$), 7.15 (m, 2H, $Ar_{Fmoc}H$), 7.28 (m, 4H, $Ar_{Fmoc}H$), 7.37 (m, 1H, $Ar_{NB}H$), 7.47 (m, 1H, $Ar_{NB}H$), 7.60 (m, 2H, $Ar_{Fmoc}H$), 7.78 (m, 1H, $Ar_{NB}H$), 8.03 (m, 1H, $Ar_{NB}H$); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 46.7, 47.7, 60.3, 61.4, 67.3, 119.7, 124.3, 124.4, 126.7, 127.4, 127.5, 127.8, 140.9, 143.3, 147.0, 157.2.

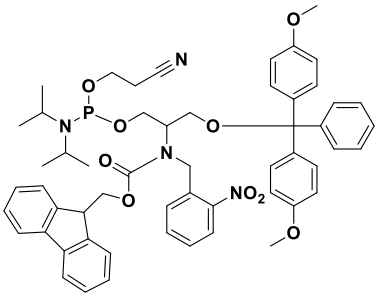
4.2.5 DMT-protected Precursor 74



To a stirred solution of nitrobenzylated carbamate **73** (1639.1 mg, 3.7 mmol, 1.0 eq.) in 15 mL THF (anh.) : pyridine (anh.) = 3 : 2 (v : v) 4,4'-dimethoxytrityl chloride (1238.4 mg, 3.7 mmol, 1.0 eq.) was added in four intervals of each 40 min at RT and under argon atmosphere and it was stirred overnight. The solvent was removed, 50 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (3x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 20 v% → 50 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 1628.1 mg (2.2 mmol, 59 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, CD_3OD, *δ*, ppm): 2.51 (m, 1H, $CH_2\underline{OH}$), 2.98 (m, 1H, $\underline{CH}CH_2OH$), 3.37 (m, 2H, $\underline{CH_2}ODMT$), 3.76 (s, 6H, $Ar_{DMT}O\underline{CH_3}$), 3.78 (m, 2H, $CH\underline{CH_2}OH$), 4.03 (m, 1H, $\underline{CH}CH_2OCON$), 4.45 (m, 2H, $N\underline{CH_2}Ph$), 4.76 (m, 2H, $CH\underline{CH_2}OCON$), 6.76 (m, 4H, $Ar_{DMT}H$), 7.10 (m, 3H, $Ar_{DMT}H$, $Ar_{Fmoc}H$), 7.23 (m, 4H, $Ar_{Fmoc}H$, $Ar_{DMT}H$), 7.30 (m, 6H, $Ar_{DMT}H$), 7.37 (m, 1H, $Ar_{NB}H$), 7.60 (m, 2H, $Ar_{Fmoc}H$), 7.63 (m, 1H, $Ar_{NB}H$), 7.76 (m, 1H, $Ar_{NB}H$), 8.06 (m, 1H, $Ar_{NB}H$).

4.2.6 Nitrobenzylated Carbamate Phosphoramidite Monomer 75

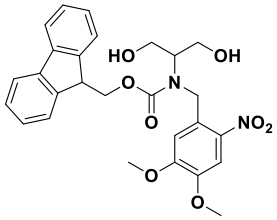


To a solution of DMT-protected precursor **74** (1623.1 mg, 2.2 mmol, 1.0 eq.) in 10 mL DCM (anh.) was added DIPEA (1.80 g, 13.9 mmol, 6.4 eq.) and molecular sieves (3 Å) under argon atmosphere at 0 °C. To the stirring reaction mixture 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (549.3 mg, 2.3 mmol, 1.1 eq.) was then added. The reaction mixture was stirred at 0 °C for 30 min and then allowed to warm to RT for 30 min. The solvent was removed and the residue purified by column chromatography (SiO_2 vs 33 v% → 50 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 1815.0 mg (1.9 mmol, 88 %) of the product as a colorless foam.

HRMS m/z: $[M+H]^+$ calculated for $C_{55}H_{60}N_4O_9P^+$ 951.4092; found, 951.4132; ^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.05 (m, 12H, $N(CH(\underline{CH_3})_2)_2$), 2.50 (m, 2H, $CH_2\underline{CH_2}C{\equiv}N$), 3.05 (m, 1H, $\underline{CH}NHBn$), 3.20 (m, 2H,

$\underline{CH_2}$ODMT), 3.61 (m, 6H, $\underline{CH_2}CH_2C{\equiv}N$, $N(\underline{CH}(CH_3)_2)_2$, $\underline{CH_2}OP$), 3.77 (s, 6H, $Ar_{DMT}O\underline{CH_3}$), 4.01 (m, 1H, $\underline{CH}CH_2OCON$), 4.45 (m, 2H, $N\underline{CH_2}Ph$), 4.62 (m, 2H, $CH\underline{CH_2}OCON$), 6.76 (m, 4H, $Ar_{DMT}H$), 7.24 (m, 13H, $Ar_{DMT}H$, $Ar_{Fmoc}H$), 7.42 (m, 1H, $Ar_{NB}H$), 7.59 (m, 2H, $Ar_{Fmoc}H$), 7.66 (m, 1H, $Ar_{NB}H$), 7.78 (m, 1H, $Ar_{NB}H$), 8.04 (m, 1H, $Ar_{NB}H$); ^{31}P NMR (161.9 MHz, $CDCl_3$, δ, ppm): 148.3.

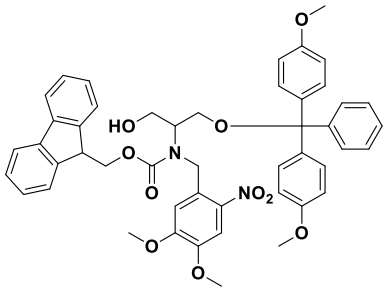
4.2.7 Nitrobenzylated Carbamate 76



To a stirred solution of nitrobenzylated serinol **66** (1547.1 mg, 5.4 mmol, 1.0 eq.) in 10 mL Na_2CO_3 (aq., 10 w%) and 20 mL 1,4-dioxane was added Fmoc-Cl (1537.7 mg, 5.9 mmol, 1.1 eq.) at 0 °C and it was stirred overnight and thereby the reaction mixture was allowed to reach RT. The solvent was removed and the residue was taken in 50 mL EtOAc, 100 mL H_2O was added and it was extracted with EtOAc (3x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the residue was purified by column chromatography (SiO_2 vs 50 v% EtOAc in cyclohexane → EtOAc) to obtain 2536.0 mg (5.0 mmol, 92 %) of the product as a slightly yellow foam.

^1H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 2.60 (bs, 2H, $CH_2\underline{OH}$), 3.64 (m, 1H, $\underline{CH}CH_2OH$), 3.72 (s, 3H, $Ar_{NB}O\underline{CH_3}$), 3.86 (m, 4H, $CH\underline{CH_2}OH$), 4.00 (s, 3H, $Ar_{NB}O\underline{CH_3}$), 4.07 (m, 1H, $\underline{CH}CH_2OCON$), 4.57 (m, 4H, $N\underline{CH_2}Ph$, $CH\underline{CH_2}OCON$), 6.73 (m, 1H, $Ar_{NB}H$), 7.18 (m, 2H, $Ar_{Fmoc}H$), 7.30 (m, 4H, $Ar_{Fmoc}H$), 7.52 (m, 2H, $Ar_{Fmoc}H$), 7.69 (m, 1H, $Ar_{NB}H$); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 47.1, 49.4, 56.4, 56.5, 61.4, 62.5, 67.3, 108.4, 109.1, 119.8, 134.2, 127.0, 127.7, 129.9, 139.1, 141.3, 143.5, 147.5, 153.7, 157.7.

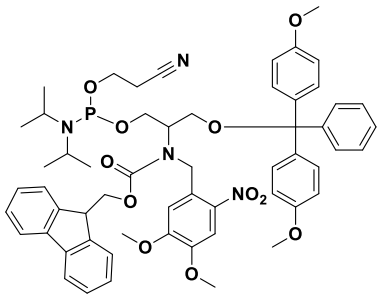
4.2.8 DMT-protected Precursor 77



To a stirred solution of nitrobenzylated carbamate **76** (2506.0 mg, 4.9 mmol, 1.0 eq.) in 15 mL THF (anh.) : pyridine (anh.) = 3 : 2 (v : v) 4,4'-dimethoxytrityl chloride (1683.2 mg, 5.0 mmol, 1.0 eq.) was added in four intervals of each 40 min at RT and under argon atmosphere and it was stirred for 60 min. The solvent was removed, 50 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (3x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 50 v% EtOAc in cyclohexane + 3 v% DIPEA → EtOAc + 3 v% DIPEA) to obtain 2862.1 mg (2.5 mmol, 72 %) of the product as a yellow oil.

^1H NMR (400.1 MHz, CD_3OD, *δ*, ppm): 3.30 (m, 2H, $\underline{CH_2}$ODMT), 3.67 (m, 1H, $\underline{CH}CH_2OH$), 3.54 (s, 3H, $Ar_{NB}O\underline{CH_3}$), 3.77 (s, 6H, $Ar_{DMT}O\underline{CH_3}$), 3.78 (m, 2H, $CH\underline{CH_2}OH$), 3.92 (m, 1H, $\underline{CH}CH_2OCON$), 4.01 (s, 3H, $Ar_{NB}O\underline{CH_3}$), 4.52 (m, 2H, $N\underline{CH_2}Ph$), 4.63 (m, 2H, $CH\underline{CH_2}OCON$), 6.57 (s, 1H, $Ar_{NB}H$), 6.76 (m, 6H, $Ar_{DMT}H$, $Ar_{Fmoc}H$), 7.15 (m, 5H, $Ar_{DMT}H$, $Ar_{Fmoc}H$), 7.24 (m, 6H, $Ar_{DMT}H$), 7.49 (m, 2H, $Ar_{Fmoc}H$), 7.69 (m, 1H, $Ar_{NB}H$).

4.2.9 Nitrobenzylated Carbamate Phosphoramidite Monomer 78



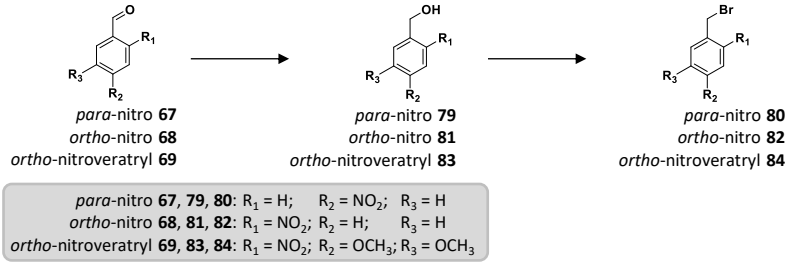
To a solution of DMT-protected precursor **77** (2.70 g, 3.3 mmol, 1.0 eq.) in 10 mL DCM (anh.) was added DIPEA (2.74 g, 21.2 mmol, 6.4 eq.) and molecular sieves (3 Å) under argon atmosphere at 0 °C. To the stirring reaction mixture 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (830.6 mg, 3.5 mmol, 1.05 eq.) was then added. The reaction mixture was stirred at 0 °C for 30 min and then allowed to warm to RT for 30 min. The solvent was removed and the residue purified by column chromatography (SiO_2 vs 50 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 3107.8 mg (3.1 mmol, 92 %) of the product as a yellow oil.

HRMS m/z: $[M+H]^+$ calculated for $C_{57}H_{64}N_4O_{11}P^+$ 1011.4304; found, 1011.4332; ^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.06 (m, 12H, N(CH(C$\underline{H}_3$)$_2$)$_2$), 2.51 (m, 2H, $CH_2\underline{CH_2}C{\equiv}N$), 3.22 (m, 1H, $\underline{CH}NHBn$), 3.42 (m, 2H, $\underline{CH_2}ODMT$), 3.43 (s, 3H, $Ar_{NB}O\underline{CH_3}$), 3.65 (m, 6H, $\underline{CH_2}CH_2C{\equiv}N$, N($\underline{CH}(CH_3)_2)_2$, $\underline{CH_2}OP$), 3.76 (s, 6H, $ArO\underline{CH_3}$), 3.91 (m, 1H, $\underline{CH}CH_2OCON$), 3.99 (s, 3H, $Ar_{NB}O\underline{CH_3}$), 4.48 (m, 2H, $N\underline{CH_2}Ph$), 4.62 (m, 2H, $CH\underline{CH_2}OCON$), 6.35 ($Ar_{NB}H$), 6.73 (m, 4H, $Ar_{DMT}H$), 7.08 (m, 1H, $Ar_{DMT}H$), 7.20 (m, 12H, $Ar_{DMT}H$, $Ar_{Fmoc}H$), 7.47 (m, 2H, $Ar_{Fmoc}H$), 7.66 (m, 1H, $Ar_{NB}H$); ^{31}P NMR (161.9 MHz, $CDCl_3$, δ, ppm): 148.1.

4.3 Monomer Synthesis Targeting Nitrobenzyl Ethers

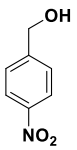
4.3.1 Synthesis of Nitrobenzyl Nucleophiles and Electrophiles for Etherifications

The synthesis of benzyl alcohols **79**, **81**, and **83** as well as benzyl bromides **80**, **82**, **84** according to the strategy shown in Scheme V-9 is followingly described.



Scheme V-9: Overview synthetic route to different nitrobenzyl alcohols and bromides.

4.3.1.1 Nitrobenzyl Alcohol 79



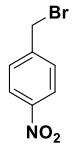
To a stirred suspension of 4-nitrobenzaldehyde **67** (1552.7 mg, 10.3 mmol, 1.0 eq.) in 10 mL MeOH (anh.) at 0 °C was dropped DCM (anh.) until a clear solution was obtained. Then, $NaBH_4$ (574.6 mg, 15.2 mmol, 1.5 eq.) was added at 0 °C and under argon atmosphere. The reaction mixture was stirred for 4.5 h and thereby allowed to reach RT. The solvent was removed, the crude was taken in 50 mL

EtoAc and 100 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (4x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed to obtain 1596.4 mg (quant.) of the product as a yellow crystalline solid.

^{1}H NMR (400.1 MHz, CD_3OD, *δ*, ppm): 4.71 (bs, 2H, Ph$\underline{CH_2}$OH), 7.50 (m, 2H, ArH), 8.08 (m, 2H, ArH); ^{13}C NMR (100.6 MHz, CD_3OD, δ, ppm): 63.8, 124.2, 127.9, 148.0, 150.5.

The analysis data harmonizes with the reported values in the literature.[385]

4.3.1.2 Nitrobenzyl Bromide 80

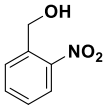


To a stirred solution of 4-nitrobenzyl alcohol **79** (1535.9 mg, 10.0 mmol, 1.0 eq.) in 100 mL THF (anh.) at 0 °C was added dropwise PBr_3 (1.9 mL, 5.4 g, 20.0 mmol, 2.0 eq.) at 0 °C and under argon atmosphere. The reaction mixture was stirred for further 60 min at 0°C and then overnight and thereby allowed to reach RT. The reaction was quenched with $NaHCO_3$ (sat., aq.) until a neutral pH was obtained. The organic solvent was then removed and the slurry was extracted with EtOAc (3x 150 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed to obtain an orange crude oil that was purified by column chromatography (SiO_2 vs 10 v% → 20 v% EtOAc in cyclohexane) to obtain 1611.3 mg (7.5 mmol, 75 %) of the product as a colorless, cotton-like foam.

^{1}H NMR (400.1 MHz, CD_3OD, *δ*, ppm): 4.52 (s, 2H, Ph$\underline{CH_2}$Br), 7.56 (m, 2H, ArH), 8.21 (m, 2H, ArH); ^{13}C NMR (100.6 MHz, CD_3OD, δ, ppm): 31.1, 124.0, 129.9, 144.8, 147.6.

The analysis data harmonizes with the reported values in the literature.[386]

4.3.1.3 Nitrobenzyl Alcohol 81

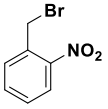


To a stirred solution of 2-nitrobenzaldehyde **68** (3043.0 mg, 20.1 mmol, 1.0 eq.) in 10 mL MeOH (anh.) $NaBH_4$ (1148.9 mg, 30.4 mmol, 1.5 eq.) was added at 0 °C and under argon atmosphere. The reaction mixture was stirred overnight and thereby allowed to reach RT. The solvent was removed, the crude was taken in 50 mL EtoAc and 50 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (4x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed to obtain 3085.2 mg (quant.) of the product as a slightly yellow crystalline solid.

^{1}H NMR (400.1 MHz, CD_3OD, *δ*, ppm): 4.92 (s, 2H, Ph$\underline{CH_2}$OH), 7.40 (m, 1H, ArH), 7.63 (m, 1H, ArH), 7.79 (m, 1H, ArH), 7.95 (m, 1H, ArH); ^{13}C NMR (100.6 MHz, CD_3OD, δ, ppm): 61.8, 125.3, 128.7, 129.3, 134.6, 138.7, 148.2.

The analysis data harmonizes with the reported values in the literature.[387]

4.3.1.4 Nitrobenzyl Bromide 82

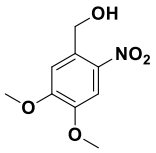


To a stirred solution of 2-nitrobenzyl alcohol **81** (1538.8 mg, 10.0 mmol, 1.0 eq.) in 100 mL THF (anh.) at 0 °C was added dropwise PBr_3 (1.9 mL, 5.4 g, 20.0 mmol, 2.0 eq.) at 0 °C and under argon atmosphere. The reaction mixture was stirred for further 60 min at 0°C and then overnight and thereby allowed to reach RT. The reaction was quenched with $NaHCO_3$ (sat., aq.) until a neutral pH was obtained. The organic solvent was then removed and the slurry was extracted with EtOAc (3x 150 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed to obtain a yellow crude oil that was purified by column chromatography (SiO_2 vs 20 v% → 33 v% EtOAc in cyclohexane) to obtain 1157.6 mg (5.4 mmol, 54 %) of the product as a pale yellow powder.

HRMS m/z: $[M+NH_4]^+$ calculated for $C_7H_{10}N_2O_2Br^+$ 232.9920; found 232.9924; ^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 4.84 (s, 2H, Ph$\underline{CH_2}$Br), 7.50 (m, 1H, ArH), 7.59 (m, 1H, ArH), 7.62 (m, 1H, ArH), 8.05 (m, 1H, ArH); ^{13}C NMR (100.6 MHz, CD_3OD, δ, ppm): 29.0, 125.2, 129.5, 132.3, 132.4, 133.6, 147.6.

The analysis data harmonizes with the reported values in the literature.[388]

4.3.1.5 Nitrobenzyl Alcohol 83

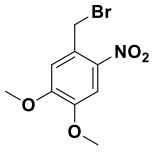


To a stirred suspension of 4,5-dimethoxy-2-nitrobenzaldehyde **69** (1346.0 mg, 6.4 mmol, 1.0 eq.) in 10 mL MeOH(anh.) at 0 °C was dropped DCM (anh.) until a clear solution was obtained. Then, $NaBH_4$ (364.3 mg, 9.6 mmol, 1.5 eq.) was added at 0 °C and under argon atmosphere. The reaction mixture was stirred overnight and thereby allowed to reach RT. The solvent was removed, the crude was taken in 50 mL EtoAc and 50 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (4x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed to obtain 1424.8 mg (quant.) of the product as a yellow crystalline solid.

^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 2.42 (bs, 1H, PhCH$_2\underline{OH}$), 3.95 (s, 3H, ArO$\underline{CH_3}$), 4.00 (s, 3H, ArO$\underline{CH_3}$), 4.96 (s, 2H, Ph$\underline{CH_2}$OH), 7.17 (m, 1H, ArH), 7.70 (m, 1H, ArH); ^{13}C NMR (100.6 MHz, CD_3OD, δ, ppm): 56.5, 56.6, 62.9, 108.4, 111.2, 132.4, 140.0, 148.2, 154.1.

The analysis data harmonizes with the reported values in the literature.[389]

4.3.1.6 Nitrobenzyl Bromide 84



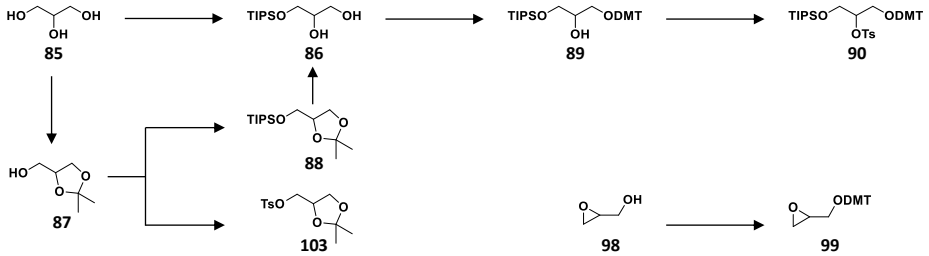
To a stirred solution of 4,5-dimethoxy-2-nitrobenzyl alcohol **83** (1424.8 mg, 6.4 mmol, 1.0 eq.) and CBr_4 (3170.9 mg, 9.6 mmol, 1.5 eq.) in 40 mL THF (anh.) at 0 °C was added PPh_3 (2.504.9 mg, 9.6 mmol, 1.5 eq.) at 0 °C and under argon atmosphere. The reaction mixture was stirred overnight and thereby allowed to reach RT. The solvent was removed and directly purified by column chromatography (SiO_2 vs 33 v% EtOAc in cyclohexane) to obtain 1432.2 mg (5.2 mmol, 81 %) of the product as a pale yellow crystalline solid.

^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 3.96 (s, 3H, ArO$\underline{CH_3}$), 4.00 (s, 3H, ArO$\underline{CH_3}$), 4.87 (s, 2H, Ph$\underline{CH_2}$Br), 6.94 (m, 1H, ArH), 7.67 (m, 1H, ArH); ^{13}C NMR (100.6 MHz, CD_3OD, δ, ppm): 30.1, 56.6, 56.7, 108.8, 113.9, 127.6, 140.5, 149.2, 153.4.

The analysis data harmonizes with the reported values in the literature.[389]

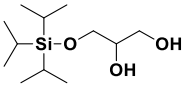
4.3.2 Synthesis of Electrophilic Diol Precursors

The synthesis of electrophilic diol precursors **90**, **99**, **103** shown in Scheme V-10 is followingly described in detail.



*Scheme V-10: Overview synthesis of electrophilic diol precursors **90, 99, 103**.*

4.3.2.1 TIPS-protected Glycerol 86



Procedure starting from glycerol **85**:

Glycerol **85** (4.84 g, 50 mmol, 3.0 eq.) was coevaporated with pyr (anh.) (3x 10 mL), taken in 15 mL pyr (anh.) under argon atmosphere, stirred at 0 °C and then TIPS-Cl (3.31 g, 17.2 mmol, 1.0 eq.) was added. The reaction mixture was stirred overnight and thereby allowed to reach RT. The solvent was removed, 100 mL H_2O was added and it was extracted with DCM (5x 50 mL). The combined organic layer was washed dried over Na_2SO_4, filtered and the crude oil was purified by column chromatography (SiO_2 vs 50 v% EtOAc in cyclohexane) to obtain 3641.1 mg (14.7 mmol, 85 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.08 (m, 21H, Si(CH(CH_3)$_2$)$_3$), 2.12 (bs, 2H, CHOHCH$_2$OH), 2.58 (bs, 1H, CHOH), 3.74 (m, 4H, CH_2O); ^{13}C NMR (100.6 MHz, CD_3OD, δ, ppm): 12.0, 18.1, 64.3, 65.3, 71.8.

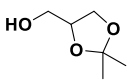
The analysis data harmonizes with the reported values in the literature.[390]

Procedure starting from dioxolane **88**:

To dioxolane **88** (3.2 g, 11.0 mmol) was added 10 mL AcOH (aq., 80 w%) and the suspension was stirred in a closed system under gentle reflux. 10 mL EtOAc were added until a clear solution was maintained. The reaction mixture was stirred at 80 °C overnight. After removal of the solvent and coevaporation with H_2O, toluene and NEt_3, the crude was taken in EtOAc and purified by column chromatography (SiO_2 vs 50 v% EtOAc in cyclohexane) to obtain 2481.0 mg (10.0 mmol, 91 %) of the product as a colorless oil.

The analysis data harmonizes with the values reported above.

4.3.2.2 Dioxolane 87



A solution of glycerol **85** (5.1 g, < 55 mmol) in 50 mL toluene (anh.) was refluxed in a Dean-Stark apparatus for 7 h. After the Dean-Stark trap was emptied, 40 mL acetone and *p*TsOH (monohydrate, 172 mg, 1.0 mmol, cat.) were added to the reaction mixture. The Dean-Stark trap was filled with

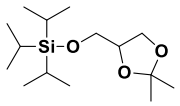
molecular sieves (3 Å) and the reaction mixture was stirred under gentle reflux overnight. The solvent was removed and the crude was purified by column chromatography (SiO_2 vs Et_2O) to obtain 4.99 g (37.8 mmol, >68 %) of the product as a slightly yellow oil.

As the water contamination in glycerol **85** was not determined, an exact yield could not be calculated.

^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.37 (s, 3H, C$\underline{CH_3}$), 1.44 (s, 3H, C$\underline{CH_3}$), 1.96 (bs, 1H, $CH_2\underline{OH}$), 3.67 (m, 2H, $\underline{CH_2}$OH), 3.91 (m, 2H, CH$\underline{CH_2}$O), 4.23 (m, 1H, $\underline{CH}CH_2O$).

The analysis data harmonizes with the reported values in the literature.[391]

4.3.2.3 TIPS-protected dioxolane 88

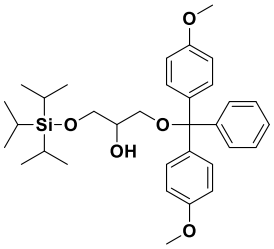


A solution of dioxolane **87** (4.99 g, 37.8 mmol, 1.0 eq.) in 10 mL pyr (anh.) and 8 mL DCM (anh.) was stirred at 0 °C under an argon atmosphere and TIPS-Cl (7.70 g, 39.7 mmol, 1.05 eq.) was added. The reaction mixture was stirred overnight and thereby allowed to reach RT. The solvent was removed, 100 mL H_2O was added and it was extracted with EtOAc (3x 100 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the crude oil was purified by column chromatography (SiO_2 vs 10 v% → 20 v% EtOAc in cyclohexane) to obtain 3184.6 mg (11.0 mmol, 29 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 0.99 (m, 21H, Si($\underline{CH}(\underline{CH_3})_2)_3$), 1.26 (s, 3H, C$\underline{CH_3}$), 1.31 (s, 3H, C$\underline{CH_3}$), 3.65 (m, 2H, $\underline{CH_2}$OTIPS), 3.88 (m, 2H, CH$\underline{CH_2}$O), 4.06 (m, 1H, $\underline{CH}CH_2O$); ^{13}C NMR (100.6 MHz, CD_3OD, δ, ppm): 12.0, 17.9, 25.4, 26.8, 64.4, 67.0, 67.3, 109.0.

The only available literature protocol[392] does not report analysis data.

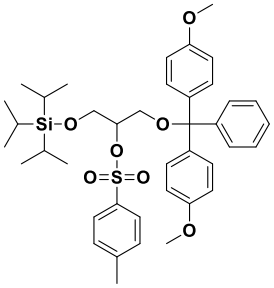
4.3.2.4 DMT and TIPS-protected Glycerol 89



To a stirred solution of TIPS-protected glycerol **86** (3641.1 mg, 14.7 mmol, 1.0 eq.) in 30 mL THF (anh.) : pyridine (anh.) = 3 : 2 (v : v) 4,4'-dimethoxytrityl chloride (4967.8 mg, 14.7 mmol, 1.0 eq.) was added in three intervals of each 40 min at RT and under argon atmosphere and it was stirred overnight. The solvent was removed, 50 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (3x 100 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 10 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 6279.2 mg (11.4 mmol, 78 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.02 (m, 21H, Si($\underline{CH}(\underline{CH_3})_2)_3$), 2.50 (m, 1H, CH$\underline{OH}CH_2$ODMT), 3.20 (m, 2H, $\underline{CH_2}$ODMT), 3.76 (m, 2H, $\underline{CH_2}$OTIPS), 3.79 (s, 6H, Ar$\underline{OCH_3}$), 3.84 (m, 1H, $\underline{CH}$OH),), 6.82 (m, 4H, $Ar_{DMT}H$), 7.20 (m, 1H, $Ar_{DMT}H$), 7.28 (m, 2H, $Ar_{DMT}H$), 7.32 (m, 4H, $Ar_{DMT}H$), 7.43 (m, 2H, $Ar_{DMT}H$); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 12.0, 18.1, 55.2, 64.4, 64.7, 71.4, 86.1, 113.2, 126.8, 127.8, 128.2, 130.1, 136.2, 145.1, 158.5.

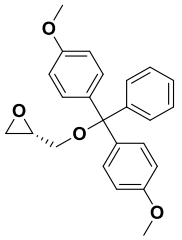
4.3.2.5 Electrophilic Diol Precursor 90



To stirred a solution of DMT and TIPS-protected glycerol **89** (5.5 g, 10 mmol, 1.0 eq.) in 12 mL pyr (anh.) was added *p*TsCl (1.9 g, 10 mmol, 1.0 eq.) at 0 °C. The reaction mixture was stirred overnight and thereby allowed to reach RT. The solvent was removed, 50 mL H_2O were added and it was extracted with EtOAc (3x 100 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The crude oil was purified by column chromatography (SiO_2 vs 20 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 6467.2 mg (9.1 mmol, 91 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.03 (m, 21H, Si(CH(CH_3)$_2$)$_3$), 2.45 (s, 3H, $Ar_{Ts}CH_3$), 3.40 (2H, CH_2ODMT), 3.79 (s, 6H, $Ar_{DMT}OCH_3$), 3.79 (m, 2H, CH_2OTIPS), 4.70 (m, 1H, CHOTs), 6.84 (m, 4H, Ar_{DMT}H), 7.24 (m, 1H, Ar_{DMT}H), 7.28 (m, 2H, Ar_{DMT}H), 7.31 (m, 6H, Ar_{DMT}H, Ar_{Ts}H), 7.44 (m, 2H, Ar_{DMT}H), 7.88 (m, 2H, Ar_{Ts}H); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 11.7, 17.8, 21.5, 55.3, 62.2, 81.6, 86.1, 113.0, 126.6, 127.7, 127.8, 128.0, 129.1, 129.7, 130.0, 134.2, 135.7, 139.6, 144.5, 158.4.

4.3.2.6 Electrophilic Diol Precursor 99

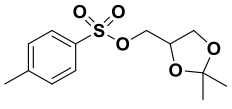


To stirred a solution of *R*-glycidol **98** (1481.0 mg, 20 mmol, 1.0 eq.) in 50 mL DCM (anh.) and Et_3N (5915.2 mg, 60 mmol, 3.0 eq.) 4,4'-dimethoxytrityl chloride (7457.8 mg, 22 mmol, 1.1 eq.) was added at RT and under argon atmosphere and it was stirred for 3 h. The solvent was removed, 100 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (3x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 10 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 7094.9 mg (18.9 mmol, 94 %) of the product as a colorless oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 2.72 (m, 2H, CH_2ODMT), 3.20 (m, 2H, CHCH_2O), 3.42 (m, 1H, $CHCH_2$O), 3.79 (s, 6H, $Ar_{DMT}OCH_3$), 6.90 (m, 4H, Ar_{DMT}H), 7.25 (m, 1H, Ar_{DMT}H), 7.34 (m, 2H, Ar_{DMT}H), 7.45 (m, 4H, Ar_{DMT}H), 7.57 (m, 2H, Ar_{DMT}H); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 44.2, 51.0, 55.0, 64.6, 86.1, 113.1, 126.7, 127.8, 128.1, 130.0, 135.9, 144.8, 158.5.

The analysis data harmonizes with the reported values in the literature.[393]

4.3.2.7 Electrophilic Diol Precursor 103

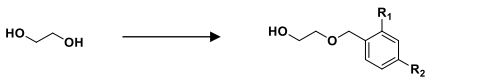


To stirred a solution of dioxolane **87** (2.63 mg, 20 mmol, 1.0 eq.) in 10 mL pyr (anh.) was added TsCl (4.17 g, 22 mmol, 1.1 eq.) at 0 °C. The reaction mixture was stirred overnight and thereby allowed to reach RT. The solvent was removed, the crude was taken in 50 mL EtOAc, filtered and the solid was washed with 300 mL EtOAc. The filtrate was washed with 100 mL NH_4Cl (aq., sat.), 100 mL $NaHCO_3$ (aq., sat.), brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The crude oil was purified by column chromatography (SiO_2 vs 10 v% → 30 v% EtOAc in cyclohexane) to obtain 5328.3 mg (18.6 mmol, 94 %) of the product as a colorless solid.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.31 (s, 3H, C$\underline{CH_3}$), 1.34 (s, 3H, C$\underline{CH_3}$), 2.45 (s, 3H, Ar$\underline{CH_3}$), 3.90 (m, 2H, $\underline{CH_2}$O), 4.00 (m, 2H, CH$\underline{CH_2}$OTs), 4.28 (m, 1H, $\underline{CH}CH_2O$), 7.35 (m, 2H, ArH), 7.80 (m, 2H, ArH); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 21.7, 25.2, 26.7, 66.2, 69.6, 73.0, 77.2, 110.0, 128.0, 130.0, 132.7, 145.1.

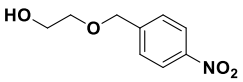
The analysis data harmonizes with the reported values in the literature.[394]

4.3.3 Nucleophilic Linker Insertion



100 *para*-nitro **101**: R_1 = H; R_2 = CH_3
ortho-nitro **102**: R_1 = CH_3; R_2 = H

4.3.3.1 Nitrobenzyl Ether 101

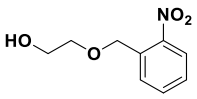


Ethylene glycol (anh.) **100** (10 mL, exc.) was stirred at RT under argon atmosphere and NaH (512.4 mg, 60 w%, 12.5 mmol, 1.25 eq.) was added. A solution of nitrobenzyl bromide **80** (2179.1 mg, 10.0 mmol, 1.0 eq.) in 10 mL THF (anh.) was added dropwise when hydrogen formation had stopped and it was stirred overnight. The reaction mixture was then poured into 300 mL ice-water and it was extracted with EtOAc (3x 100 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 5 v% MeOH in DCM) to obtain 1980.1 mg (quant.) of the product as a yellow oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.84 (bs, 1H, $CH_2\underline{OH}$), 3.66 (t, ^3J = 4.6 Hz, 2H, BnO$\underline{CH_2}CH_2OH$), 3.82 (t, ^3J = 4.6 Hz, 2H, BnO$CH_2\underline{CH_2}$OH), 4.68 (m, 2H, Ph$\underline{CH_2}$O), 7.51 (m, 2H, ArH), 8.21 (m, 2H, ArH); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 60.8, 71.0, 71.6, 122.7, 127.0, 145.7, 146.4.

The analysis data harmonizes with the reported values in the literature.[371]

4.3.3.2 Nitrobenzyl Ether 102



Ethylene glycol (anh.) **100** (10 mL, exc.) was stirred at RT under argon atmosphere and NaH (122.3 mg, 60 w%, 3.0 mmol, 1.5 eq.) was added. A solution of nitrobenzyl bromide **82** (435.5 mg, 2.0 mmol, 1.0 eq.) in 1 mL THF (anh.) was added dropwise when hydrogen formation had stopped and it was stirred overnight. The reaction mixture was then poured into 300 mL ice-water and it was extracted with EtOAc (3x 100 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column

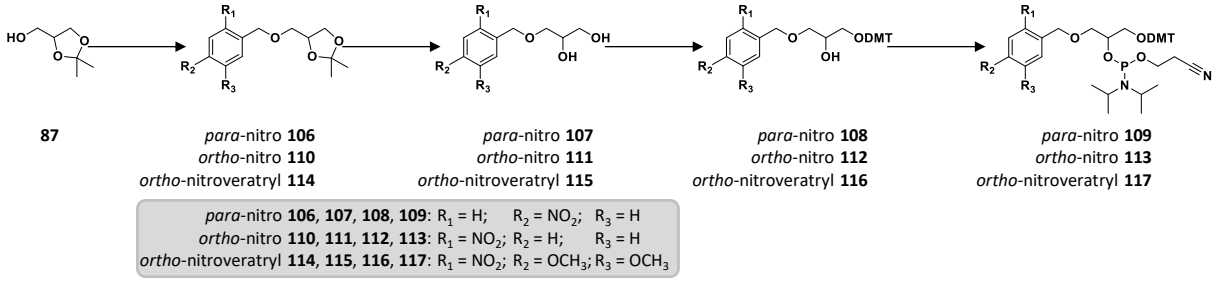
chromatography (SiO_2 vs 50 v% EtOAc in cyclohexane) to obtain 393.4 mg (2.0 mmol, 99 %) of the product as a yellow oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.83 (bs, 1H, $CH_2\underline{OH}$), 3.71 (t, ^3J = 4.5 Hz, 2H, $BnO\underline{CH_2}CH_2OH$), 3.83 (t, ^3J = 4.5 Hz, 2H, $BnOCH_2\underline{CH_2}OH$), 4.95 (m, 2H, $Ph\underline{CH_2}O$), 7.45 (m, 1H, ArH), 7.65 (m, 1H, ArH), 7.78 (m, 1H, ArH), 8.06 (m, 1H, ArH); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 61.7, 69.7, 72.4, 124.6, 128.1, 128.8, 133.6, 134.6, 147.3.

The analysis data harmonizes with the reported values in the literature.[372]

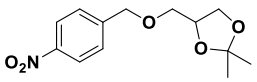
4.3.4 Successful Monomer Synthesis

The synthesis of phosphoramidite monomers **109**, **113**, and **117** is shown in Scheme V-11 and further detailed.



Scheme V-11: Overview synthetic route to different nitrobenzyl ether phosphoramidite monomers.

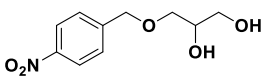
4.3.4.1 Dioxolane 106



To a stirred solution of 2,2-dimethyl-4-hydroxymethyl-1,3-dioxolane **87** (1327.8 mg, 10.0 mmol, 1.0 eq.) and 4-nitrobenzyl bromide **80** (3247.1 mg, 15.0 mmol, 1.5 eq.) in 40 mL MeCN (anh.) at 0 °C was added NaH (440.6 mg, 60 w%, 11.0 mmol, 1.1 eq.) at 0 °C and under argon atmosphere. The reaction mixture was stirred for 30 min at 0°C and for further 30 min allowed to reach RT until the reaction mixture was filtered. The solvent was removed from the filtrate and the crude directly purified by column chromatography (SiO_2 vs 33 v% EtOAc in cyclohexane) to obtain 1077.4 mg (4.0 mmol, 40 %) of the product as a yellow solid.

HRMS m/z: $[M+H]^+$ calculated for $C_{13}H_{18}NO_5^+$ 268.1179; found, 268.1180; ^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.38 (s, 3H, $C\underline{CH_3}$), 1.43 (s, 3H, $C\underline{CH_3}$), 3.58 (m, 2H, $\underline{CH_2}OBn$), 3.93 (m, 2H, $CH\underline{CH_2}O$), 4.34 (m, 1H, $\underline{CH}CH_2O$), 4.68 (m, 2H, $Ph\underline{CH_2}O$), 7.51 (m, 2H, ArH), 8.21 (m, 2H, ArH); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 25.2, 26.6, 66.4,71.6, 72.0, 74.6, 109.4, 123.4, 127.6, 145.7, 147.2.

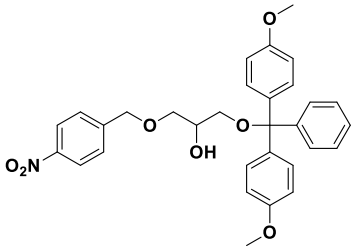
4.3.4.2 Diol 107



To dioxolane **106** (1031.4 mg, 3.9 mmol) was added 20 mL AcOH (aq., 80 w%) and the suspension was stirred in a closed system under gentle reflux. 5 mL DCM were added until a clear solution was maintained. The reaction mixture was stirred at 80 °C overnight. After removal of the solvent and coevaporation with H_2O, toluene and NEt_3, the crude was taken in EtOAc and purified by column chromatography (SiO_2 vs EtOAc) to obtain 670.2 mg (3.0 mmol, 76 %) of the product as a yellow oil.

HRMS m/z: $[M+Na]^+$ calculated for $C_{10}H_{13}NO_5Na^+$ 250.0686; found, 250.0687; ^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.60 (bs, 2H, HC<u>OH</u>C<u>H$_2$OH</u>), 3.58 (m, 4H, <u>CH$_2$</u>OBn, <u>CH$_2$</u>OH), 3.96 (m, 1H, <u>CH</u>CH$_2$O), 4.67 (s, 2H, Ph<u>CH$_2$</u>O), 7.50 (m, 2H, ArH), 8.22 (m, 2H, ArH); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 64.0, 70.9, 72.3, 72.4, 123.8, 127.9, 145.5, 147.5.

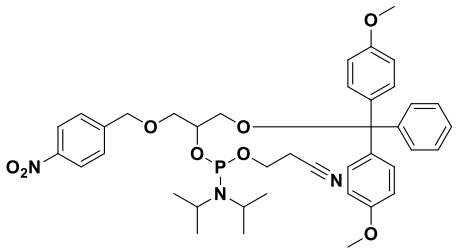
4.3.4.3 DMT Precursor 108



To a stirred solution of diol **107** (660.5 mg, 2.9 mmol, 1.0 eq.) in 30 mL THF (anh.) : pyridine (anh.) = 3 : 2 (v : v) 4,4'-dimethoxytrityl chloride (1427.7 mg, 4.2 mmol, 1.5 eq.) was added in five intervals of each 40 min at RT and under argon atmosphere. The reaction mixture was thereby monitored by TLC for the undesired double DMT protection that did not occur. After another 40 min the solvent was removed, 100 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (4x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 33 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 1223.8 mg (2.3 mmol, 80 %) of the product as a colorless foam.

HRMS m/z: $[M+Na]^+$ calculated for $C_{31}H_{31}NO_7Na^+$ 552.1993; found, 552.1992; ^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.60 (m, 1H, HC<u>OH</u>), 3.27 (m, 2H, <u>CH$_2$</u>ODMT), 3.64 (m, 1H, <u>CH</u>CH$_2$O), 3.78 (s, 6H, Ar$_{DMT}$OCH$_3$), 4.03 (s, 2H, Ph<u>CH$_2$</u>O), 6.83 (m, 4H, Ar$_{DMT}$H), 7.22 (m, 1H, Ar$_{DMT}$H), 7.27 (m, 2H, Ar$_{DMT}$H), 7.33 (m, 4H, Ar$_{DMT}$H), 7.43 (m, 4H, Ar$_{DMT}$H, Ar$_{NB}$H), 8.17 (m, 2H, Ar$_{NB}$H); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 55.27, 64.3, 70.0, 72.1, 72.2, 86.3, 113.2, 123.7, 126.9, 127.7 127.9, 128.2, 130.1, 135.9, 144.8, 145.8, 147.4, 158.6.

4.3.4.4 Phosphoramidite Monomer 109

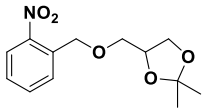


To a solution of DMT precursor **108** (1148.4 mg, 2.2 mmol, 1.0 eq.) in 10 mL DCM (anh.) was added DIPEA (1.85 g, 13.7 mmol, 6.6 eq.) and 100 mg of molecular sieves (3 Å) under argon atmosphere at 0 °C. To the stirring reaction mixture 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (600.1 mg, 2.5 mmol, 1.15 eq.) was then added. The reaction mixture was stirred at 0 °C for 30 min and then allowed to warm to RT for 30 min. The solvent was removed and the residue purified by column chromatography (SiO_2 vs 33 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 1520.7 mg (2.1 mmol, 96 %) of the product as a colorless foam.

HRMS m/z: $[M+H]^+$ calculated for $C_{40}H_{49}N_3O_8P^+$ 730.3252; found, 730.3251; ^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.13 (m, 12H, N(CH(<u>CH$_3$</u>)$_2$)$_2$), 2.52 (m, 2H, CH$_2$<u>CH$_2$</u>C≡N), 3.26 (m, 2H, <u>CH$_2$</u>ODMT), 3.70 (m, 6H, <u>CH$_2$</u>CH$_2$C≡N, N(<u>CH</u>(CH$_3$)$_2$)$_2$, <u>CH$_2$</u>OPh), 3.78 (s, 6H, Ar$_{DMT}$O<u>CH$_3$</u>), 4.18 (m, 1H, <u>CH</u>OP), 4.62 (m, 2H, Ph<u>CH$_2$</u>O), 6.80 (m, 4H, Ar$_{DMT}$H), 7.21 (m, 1H, Ar$_{DMT}$H), 7.27 (m, 2H, Ar$_{DMT}$H), 7.32 (m, 4H, Ar$_{DMT}$H), 7.40 (m, 2H,

$Ar_{DMT}H$), 7.44 (m, 2H, $Ar_{NB}H$), 8.15 (m, 2H, $Ar_{NB}H$); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 20.4, 24.8, 43.3, 55.3, 58.4, 63.9, 71.9, 72.6, 86.1, 113.2, 117.8, 123.6, 126.9, 127.6, 127.9, 128.2, 128.3, 130.2, 136.2, 145.0, 146.2, 147.4, 158.6; ^{31}P NMR (161.9 MHz, $CDCl_3$, δ, ppm): 149.4.

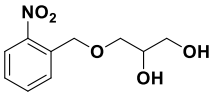
4.3.4.5 Dioxolane 110



To a stirred solution of 2,2-dimethyl-4-hydroxymethyl-1,3-dioxolane **87** (800.5 mg, 6.1 mmol, 1.2 eq.) and 2-nitrobenzyl bromide **82** (1087.5 mg, 5.0 mmol, 1.0 eq.) in 40 mL MeCN (anh.) at 0 °C was added NaH (479.1 mg, 60 w%, 12.0 mmol, 2.4 eq.) at 0 °C and under argon atmosphere. The reaction mixture was stirred for 30 min at 0°C and for further 30 min allowed to reach RT and the solvent was removed. 50 mL NH_4Cl (aq., sat.) was added and it was extracted with EtOAc (3x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 10 v% → 20 v% EtOAc in cyclohexane) to obtain 658.2 mg (2.5 mmol, 49 %) of the product as a pale yellow solid.

^1H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.38 (s, 3H, $C\underline{CH_3}$), 1.44 (s, 3H, $C\underline{CH_3}$), 3.64 (m, 2H, $\underline{CH_2}OBn$), 3.95 (m, 2H, $CH\underline{CH_2}O$), 4.36 (m, 1H, $\underline{CH}CH_2O$), 4.97 (m, 2H, $Ph\underline{CH_2}O$), 7.44 (m, 1H, ArH), 7.65 (m, 1H, ArH), 7.80 (m, 1H, ArH), 8.06 (m, 1H, ArH); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 25.5, 26.8, 66.7,70.1, 72.2, 74.7, 109.6, 124.7, 128.1, 128.8, 133.7, 134.8, 147.3.

4.3.4.6 Diol 111



Synthesis from dioxolane **110**:

To dioxolane **110** (378.1 mg, 1.4 mmol) was added 3.5 mL AcOH (aq., 80 w%) and the solution was stirred under gentle reflux for 2 h. After removal of the solvent and coevaporation with H_2O, toluene and NEt_3, the crude was taken in EtOAc and purified by column chromatography (SiO_2 vs EtOAc) to obtain 247.4 mg (1.1 mmol, 77 %) of the product as a yellow oil.

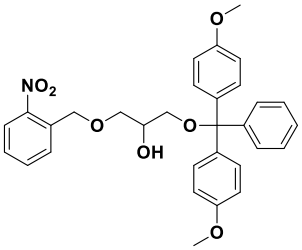
^1H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 3.52 (m, 2H, $\underline{CH_2}OBn$), 3.60 (m, 2H, $\underline{CH_2}OH$), 3.89 (m, 1H, $\underline{CH}CH_2O$), 3.90 (bs, 2H, $HC\underline{OH}CH_2\underline{OH}$), 4.79 (s, 2H, $Ph\underline{CH_2}O$), 7.33 (m, 1H, ArH), 7.54 (m, 1H, ArH), 7.65 (m, 1H, ArH), 7.90 (m, 1H, ArH); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 63.7, 69.9, 70.9, 72.2, 124.5, 128.1, 128.7, 133.5, 134.3, 147.2.

Synthesis from dioxolane **87**:

As for the synthesis of dioxolane **110**, to a stirred solution of 2,2-dimethyl-4-hydroxymethyl-1,3-dioxolane **87** (661.3 mg, 5.0 mmol, 1.0 eq.) and 2-nitrobenzyl bromide **82** (1296.4 mg, 6.0 mmol, 1.2 eq.) in 10 mL MeCN (anh.) at 0 °C was added NaH (240.7 mg, 60 w%, 6.0 mmol, 1.2 eq.) at 0 °C and under argon atmosphere. The reaction mixture was stirred for 30 min at 0°C and for further 30 min allowed to reach RT and the solvent was removed. The residue was taken in 2 mL NH_4Cl (aq., sat.) and 50 mL EtOAc, dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil with a solid fraction (1.5 g) was solubilized in 4ml AcOH (aq., 80 w%) and filtered. The filtrate was stirred under gentle reflux for 3 h. After removal of the solvent and coevaporation with H_2O, toluene and NEt_3, the crude was taken in EtOAc and purified by column chromatography (SiO_2 vs EtOAc) to obtain 763.5 mg (3.4 mmol, 57 %) of the product as a yellow oil.

The analysis data harmonizes with the values reported above.

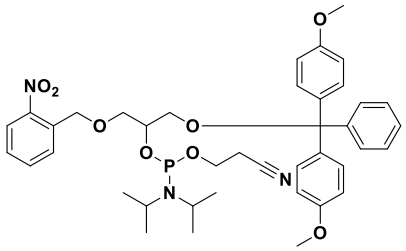
4.3.4.7 DMT Precursor 112



To a stirred solution of diol **111** (240.5 mg, 1.1 mmol, 1.0 eq.) in 5 mL THF (anh.) : pyridine (anh.) = 3 : 2 (v : v) 4,4'-dimethoxytrityl chloride (360.4 mg, 1.1 mmol, 1.0 eq.) was added in three intervals of each 40 min at RT and under argon atmosphere and stirred overnight. Te solvent was removed, 20 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (3x 50 mL). The combined organic layer was washed with brine (2x 25 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 20 v% → 33 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 493.3 mg (0.9 mmol, 88 %) of the product as a colorless foam.

HRMS m/z: $[M+Na]^+$ calculated for $C_{31}H_{31}NO_7Na^+$ 552.1993; found, 552.1996; ^1H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 2.34 (m, 1H, HCOH), 3.27 (m, 2H, CH_2ODMT), 3.69 (m, 1H, $CHCH_2O$), 3.78 (s, 6H, $Ar_{DMT}OCH_3$), 4.91 (s, 2H, $PhCH_2O$), 6.82 (m, 4H, $Ar_{DMT}H$), 7.21 (m, 1H, $Ar_{DMT}H$), 7.28 (m, 2H, $Ar_{DMT}H$), 7.32 (m, 4H, $Ar_{DMT}H$), 7.43 (m, 2H, $Ar_{DMT}H$), 7.44 (m, 1H, $Ar_{NB}H$), 7.60 (m, 1H, $Ar_{NB}H$), 7.69 (m, 1H, $Ar_{NB}H$), 8.07 (m, 1H, $Ar_{NB}H$).

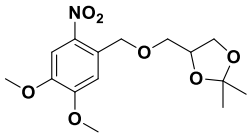
4.3.4.8 Phosphoramidite Monomer 113



To a solution of DMT precursor **112** (1135.9 mg, 2.1 mmol, 1.0 eq.) in 10 mL DCM (anh.) was added DIPEA (1.66 g, 13.7 mmol, 6.0 eq.) and molecular sieves (3 Å) under argon atmosphere at 0 °C. To the stirring reaction mixture 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (581.4 mg, 2.5 mmol, 1.15 eq.) was then added. The reaction mixture was stirred at 0 °C for 30 min and then allowed to warm to RT for 30 min. The solvent was removed and the residue purified by column chromatography (SiO_2 vs 33 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 1508.0 mg (2.1 mmol, 97 %) of the product as a colorless foam.

HRMS m/z: $[M+H]^+$ calculated for $C_{40}H_{49}N_3O_8P^+$ 730.3252; found 730.3244; ^1H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.12 (m, 12H, $N(CH(CH_3)_2)_2$), 2.51 (m, 2H, $CH_2CH_2C{\equiv}N$), 3.27 (m, 2H, CH_2ODMT), 3.70 (m, 6H, $CH_2CH_2C{\equiv}N$, $N(CH(CH_3)_2)_2$, CH_2OPh), 3.78 (s, 6H, $Ar_{DMT}OCH_3$), 4.20 (m, 1H, CHOP), 4.91 (m, 2H, $PhCH_2O$), 6.80 (m, 4H, $Ar_{DMT}H$), 7.20 (m, 1H, $Ar_{DMT}H$), 7.25 (m, 2H, $Ar_{DMT}H$), 7.32 (m, 4H, $Ar_{DMT}H$), 7.43 (m, 2H, $Ar_{DMT}H$), 7.44 (m, 1H, $Ar_{NB}H$), 7.58 (m, 1H, $Ar_{NB}H$), 7.73 (m, 1H, $Ar_{NB}H$), 8.08 (m, 1H, $Ar_{NB}H$); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 20.1, 24.5, 43.1, 55.0, 58.3, 63.8, 69.8, 71.8, 72.5, 85.9, 113.0, 117.6, 124.4, 126.6, 127.6, 128.1, 128.1, 128.5, 130.0, 133.6, 135.3, 136.0, 144.9, 146.8, 158.4; ^{31}P NMR (161.9 MHz, $CDCl_3$, δ, ppm): 149.4.

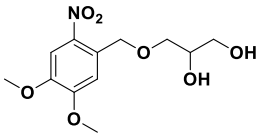
4.3.4.9 Dioxolane 114



To a stirred suspension of 2,2-dimethyl-4-hydroxymethyl-1,3-dioxolane **87** (539.8 mg, 4.1 mmol, 1.0 eq.) and 4,5-dimethoxy-2-nitrobenzyl bromide **84** (1432.2 mg, 5.2 mmol, 1.3 eq.) in 20 mL MeCN (anh.) at 0 °C was added 8 mL DCM (anh.) until a clear solution was maintained. Then, NaH (194.1 mg, 60 w%, 4.8 mmol, 1.2 eq.) was added at 0 °C and under argon atmosphere. The reaction mixture was stirred for 30 min at 0°C and for further 3 h allowed to reach RT and the solvent was removed. 20 mL H_2O was added and it was extracted with DCM (3x 50 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 20 v% → 33 v% EtOAc in cyclohexane) to obtain 565.7 mg (1.7 mmol, 43 %) of the product as a yellow solid. 1.1 mmol of bromide **84** were also recovered.

^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.33 (s, 3H, C$\underline{CH_3}$), 1.39 (s, 3H, C$\underline{CH_3}$), 3.62 (m, 2H, $\underline{CH_2}$OBn), 3.89 (s, 3H, ArO$\underline{CH_3}$), 3.91 (m, 2H, CH$\underline{CH_2}$O), 3.95 (s, 3H, ArO$\underline{CH_3}$), 4.33 (m, 1H, $\underline{CH}CH_2O$), 4.91 (m, 2H, Ph$\underline{CH_2}$O), 7.28 (m, 1H, ArH), 7.64 (m, 1H, ArH); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 25.4, 26.8, 56.3, 56.4, 66.4, 70.1, 71.9, 74.8, 108.0, 109.5, 109.6, 130.7, 139.2, 147.7, 153.9.

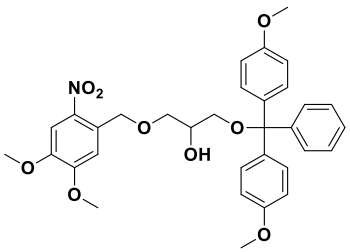
4.3.4.10 Diol 115



To dioxolane **114** (560.0 mg, 1.7 mmol) was added 5 mL AcOH (aq., 80 w%) and the suspension was stirred in a closed system under gentle reflux. 2 mL THF were added until a clear solution was maintained. The reaction mixture was stirred at 80 °C for 2.5 h. After removal of the solvent and coevaporation with H_2O, toluene and NEt_3, the crude was taken in EtOAc and purified by column chromatography (SiO_2 vs EtOAc) to obtain 411.1 mg (1.4 mmol, 84 %) of the product as a yellow solid.

HRMS m/z: $[M+NH_4]^+$ calculated for $C_{12}H_{21}N_2O_7^+$ 305.1343; found, 305.1349; ^{1}H NMR (400.1 MHz, CD_3OD, *δ*, ppm): 3.63 (m, 4H, $\underline{CH_2}$OBn, $\underline{CH_2}$OH), 3.86 (m, 1H, $\underline{CH}CH_2O$), 3.90 (s, 3H, ArO$\underline{CH_3}$), 3.97 (s, 3H, ArO$\underline{CH_3}$), 4.91 (s, 2H, Ph$\underline{CH_2}$O), 7.43 (m, 1H, ArH), 7.72 (m, 1H, ArH); ^{13}C NMR (100.6 MHz, CD_3OD, δ, ppm): 56.8, 56.9, 64.4, 71.0, 72.3, 73.5, 109.2, 111.2, 131.7, 140.7, 149.2, 155.3.

4.3.4.11 DMT Precursor 116

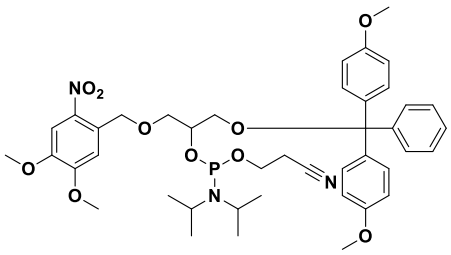


To a stirred solution of diol **115** (403.7 mg, 1.4 mmol, 1.0 eq.) in 10 mL THF (anh.) : pyridine (anh.) = 3 : 2 (v : v) 4,4'-dimethoxytrityl chloride (483.2 mg, 1.4 mmol, 1.0 eq.) was added in four intervals of each 40 min at RT and under argon atmosphere and stirred overnight. The solvent was removed, 20 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (3x 50 mL). The combined organic layer

was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was purified by column chromatography (SiO_2 vs 33 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 419.7 mg (0.7 mmol, 51 %) of the product as a yellow foam.

HRMS m/z: $[M+Na]^+$ calculated for $C_{33}H_{35}NO_9Na^+$ 612.2204; found, 612.2203; ^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 2.51 (bs, 1H, CHOH), 3.27 (m, 2H, CH_2ODMT), 3.69 (m, 2H, CH_2OBn), 3.77 (s, 6H, $Ar_{DMT}OCH_3$), 3.88 (s, 3H, $Ar_{NB}OCH_3$), 3.94 (s, 3H, $Ar_{NB}OCH_3$), 4.08 (m, 1H, CHOH), 4.92 (s, 2H, $PhCH_2O$), 6.81 (m, 4H, $Ar_{DMT}H$), 7.18 (m, 1H, $Ar_{DMT}H$), 7.23 (m, 1H, $Ar_{NB}H$), 7.27 (m, 2H, $Ar_{DMT}H$), 7.31 (m, 4H, $Ar_{DMT}H$), 7.43 (m, 2H, $Ar_{DMT}H$), 7.70 (m, 1H, $Ar_{NB}H$); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 55.01, 56.1, 56.2, 64.5, 69.8, 69.9, 72.5, 86.1, 107.7, 109.3, 113.0, 126.7, 127.7 128.0, 129.9, 130.7, 135.8, 138.8, 144.7, 147.4, 153.7, 158.4.

4.3.4.12 Phosphoramidite Monomer 117

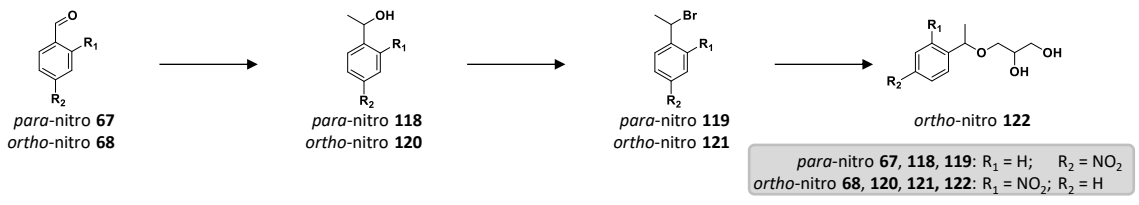


To a solution of DMT precursor **116** (419.7 mg, 0.7 mmol, 1.0 eq.) in 5 mL DCM (anh.) was added DIPEA (584.3 mg, 4.5 mmol, 6.0 eq.) and molecular sieves (3 Å) under argon atmosphere at 0 °C. To the stirring reaction mixture 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (188.7 mg, 0.8 mmol, 1.12 eq.) was then added. The reaction mixture was stirred at 0 °C for 30 min and then allowed to warm to RT for 30 min. The solvent was removed and the residue purified by column chromatography (SiO_2 vs 30 v% EtOAc in cyclohexane + 3 v% DIPEA) to obtain 463.4 mg (0.6 mmol, 83 %) of the product as a yellow foam.

HRMS m/z: $[M+H]^+$ calculated for $C_{42}H_{53}N_3O_{10}P^+$ 790.3463; found 790.3471; ^{1}H NMR (400.1 MHz, $CDCl_3$, *δ*, ppm): 1.12 (m, 12H, $N(CH(CH_3)_2)_2$), 2.51 (m, 2H, $CH_2CH_2C{\equiv}N$), 3.24 (m, 2H, CH_2ODMT), 3.70 (m, 6H, $CH_2CH_2C{\equiv}N$, $N(CH(CH_3)_2)_2$, CH_2OPh), 3.69 (m, 3H, $Ar_{NB}OCH_3$), 3.77 (m, 6H, $Ar_{DMT}OCH_3$), 3.95 (m, 3H, $Ar_{NB}OCH_3$), 4.23 (m, 1H, , CHOP), 4.92 (m, 2H, $PhCH_2O$), 6.79 (m, 4H, $Ar_{DMT}H$), 7.18 (m, 1H, $Ar_{DMT}H$), 7.21 (m, 1H, $Ar_{NB}H$), 7.24 (m, 2H, $Ar_{DMT}H$), 7.32 (m, 4H, $Ar_{DMT}H$), 7.44 (m, 2H, $Ar_{DMT}H$), 7.71 (m, 1H, $Ar_{NB}H$); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 20.0, 24.3, 42.9, 54.9, 56.0, 56.1, 58.1, 58.3, 63.9, 69.9, 72.3, 85.8, 107.5, 109.2, 112.8, 117.6, 126.5, 126.5, 127.5, 127.9, 129.9, 131.0, 135.8, 138.6, 144.8, 147.2, 153.6, 158.2; ^{31}P NMR (161.9 MHz, $CDCl_3$, δ, ppm): 149.4.

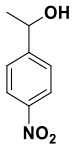
4.3.5 Etherification of Nitrophenyl Ethyl Electrophiles

The synthesis of (1-bromoethyl)-nitrobenzenes **119** and **121** as well as the synthesis of corresponding diol **122** is shown in Scheme V-12 and is described below.



*Scheme V-12: Overview synthesis of nitrophenyl ethyl electrophiles **119** and **121** and ether formation to diol **122**.*

4.3.5.1 1-Nitrophenylethan-1-ol 118

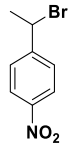


$TiCl_4$ (0.85 mL, 7.5 mmol, 0.25 eq.) was added to 50 mL THF (anh.) at -78 °C and under argon atmosphere and stirred for 60 min. MeLi (1.6 M in Et_2O, 19 mL, 30.3 mmol, 1.0 eq.) was added dropwise to the reaction mixture and it was further stirred at -78 °C for 60 min until the solution remains grass-green. A solution of 4-nitrobenzaldehyde **67** (4535.6 mg, 33.3 mmol 1.1 eq) in 15 mL THF (anh.) was added dropwise at – 78°C and the reaction mixture was stirred overnight and thereby allowed to slowly reach RT. The reaction was quenched by the addition of 20 mL H_2O, the solvent was removed and 50 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (5x 100 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The crude oil was purified by column chromatography (SiO_2 vs 30 v% EtOAc in cyclohexan) to obtain 2228.6 mg (13.3 mmol, 44 %) of the product as a slightly yellow crystalline solid.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.53 (d, ^3J = 6.5 Hz, 3H, PhCHOHC$\underline{H}_3$), 1.83 (bs, 1H, PhCHO$\underline{H}$CH$_3$), 5.03 (q, ^3J = 6.5 Hz, 1H, Ph$\underline{C}$HOHCH$_3$), 7.55 (m, 2H, ArH), 8.21 (m, 2H, ArH); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 25.2, 69.3, 123.5, 126.1, 146.9, 153.4.

The analysis data harmonizes with the reported values in the literature.[395]

4.3.5.2 (1-Bromoethyl)nitrobenzene 119

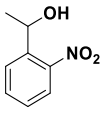


To a stirred solution of 1-(4-nitrophenyl)ethan-1-ol **118** (848.5 mg, 5.1 mmol, 1.0 eq.) in 5 mL DCM (anh.) at 0 °C was added dropwise PBr_3 (5 mL, exc.) at 0 °C and under argon atmosphere. The reaction mixture was stirred for further 60 min at 0°C and then overnight and thereby allowed to reach RT. The reaction was quenched with $NaHCO_3$ (sat., aq.) until a neutral pH was obtained. It was extracted with DCM (3x 100 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed to obtain 801.4 mg (3.5 mmol, 70 %) of the product as a yellow oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 2.06 (d, ^3J = 6.9 Hz, 3H, PhCHBrC$\underline{H}_3$), 5.20 (q, ^3J = 6.9 Hz, 1H, Ph$\underline{C}$HBrCH$_3$), 7.60 (m, 2H, ArH), 8.21 (m, 2H, ArH); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 26.2, 46.6, 123.7, 127.7, 147.2, 149.9.

The analysis data harmonizes with the reported values in the literature.[396]

4.3.5.3 1-Nitrophenylethan-1-ol 120



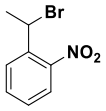
$TiCl_4$ (0.85 mL, 7.5 mmol, 0.25 eq.) was added to 50 mL THF (anh.) at -78 °C and under argon atmosphere and stirred for 60 min. MeLi (1.6 M in Et_2O, 19 mL, 30.3 mmol, 1.0 eq.) was added dropwise to the reaction mixture and it was further stirred at -78 °C for 60 min until the solution remains grass-green. A solution of 2-nitrobenzaldehyde **68** (4538.1 mg, 33.3 mmol 1.1 eq) in 15 mL THF (anh.) was

added dropwise at – 78°C and the reaction mixture was stirred overnight and thereby allowed to slowly reach RT. The reaction was quenched by the addition of 20 mL H_2O, the solvent was removed and 50 mL $NaHCO_3$ (sat., aq.) were added and it was extracted with EtOAc (5x 100 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The crude oil was purified by column chromatography (SiO_2 vs 33 v% EtOAc in cyclohexan) to obtain 1807.4 mg (10.8 mmol, 36 %) of the product as a slightly yellow crystalline solid.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.58 (d, ^3J = 6.4 Hz, 3H, PhCHOHC<u>H$_3$</u>), 1.13 (bs, 1H, PhCH<u>OH</u>CH$_3$), 5.42 (q, ^3J = 6.4 Hz, 1H, Ph<u>CH</u>OHCH$_3$), 7.42 (m, 1H, ArH), 7.65 (m, 2H, ArH), 7.84 (m, 1H, ArH), 7.90 (m, 2H, ArH).

The analysis data harmonizes with the reported values in the literature.[397]

4.3.5.4 (1-Bromoethyl)nitrobenzene 121

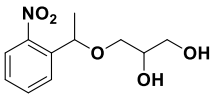


To a stirred solution of 1-(2-nitrophenyl)ethan-1-ol **120** (526.6 mg, 3.2 mmol, 1.0 eq.) in 5 mL DCM (anh.) at 0 °C was added dropwise PBr_3 (4.5 mL, exc.) at 0 °C and under argon atmosphere. The reaction mixture was stirred for further 60 min at 0°C and then stirred for another 4 h and thereby allowed to reach RT. The reaction was quenched with $NaHCO_3$ (sat., aq.) until a neutral pH was obtained. It was extracted with EtOAc (3x 100 mL). The combined organic layer was washed with brine (2x 50 mL), dried over Na_2SO_4, filtered and the solvent was removed. The crude oil was purified by column chromatography (SiO_2 vs 20 v% EtOAc in cyclohexane) to obtain 495.2 mg (2.2 mmol, 68 %) of the product as a yellow oil.

^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 2.08 (d, ^3J = 6.8 Hz, 3H, PhCHOHC<u>H$_3$</u>), 5.81 (q, ^3J = 6.8 Hz, 1H, Ph<u>CH</u>OHCH$_3$), 7.43 (m, 1H, ArH), 7.64 (m, 2H, ArH), 7.83 (m, 1H, ArH), 7.89 (m, 2H, ArH); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 26.7, 41.7, 124.1, 128.8, 129.5, 133.3, 137.2, 147.3.

The analysis data harmonizes with the reported values in the literature.[398]

4.3.5.5 Diol 122



To a stirred solution of 2,2-dimethyl-4-hydroxymethyl-1,3-dioxolane **87** (270.7 mg, 2.0 mmol, 1.0 eq.) in 5 mL MeCN (anh.) at 0 °C was added NaH (85.4 mg, 60 w%, 2.2 mmol, 1.05 eq.) at 0 °C and under argon atmosphere. The reaction mixture was stirred for 45 min at 0°C and for further 30 min allowed to reach RT. A solution of 1-(1-bromoethyl)-nitrobenzene **121** (495.2 mg, 2.2 mmol, 1.05 eq.) in 2 mL MeCN (anh.) added dropwise at 0 °C. The reaction mixture was stirred overnight and thereby allowed to reach RT and the solvent was removed. The residue was taken in 2 mL NH_4Cl (aq., sat.) and 50 mL EtOAc, dried over Na_2SO_4, filtered and the solvent was removed. The yellow crude oil was taken in 7 mL AcOH (aq., 80 w%) and stirred under gentle reflux for 90 min. After removal of the solvent and coevaporation with H_2O, toluene and NEt_3, the crude was taken in EtOAc and purified by column chromatography (SiO_2 vs EtOAc) to obtain 58.3 mg (0.2 mmol, 12 %) of the product as a brown oil.

HRMS m/z: $[M+NH_4]^+$ calculated for $C_{11}H_{19}N_2O_5^+$ 259.1288; found, 259.1291; ^{1}H NMR (400.1 MHz, $CDCl_3$, δ, ppm): 1.50 (m, 3H, PhCHC<u>H$_3$</u>), 3.09 (bs, 2H, HC<u>OH</u>CH$_2$<u>OH</u>), 3.32 (m, 2H, <u>CH$_2$</u>OBn), 3.58 (m, 2H, <u>CH$_2$</u>OH), 3.82 (m, 1H, <u>CH</u>CH$_2$O), 5.00 (m, 1H, Ph<u>CH</u>CH$_3$), 7.39 (m, 1H, ArH), 7.62 (m, 1H, ArH), 7.70 (m,

1H, ArH), 7.86 (m, 1H, ArH); ^{13}C NMR (100.6 MHz, $CDCl_3$, δ, ppm): 23.3, 64.0, 70.7, 70.9, 74.0, 124.3, 127.7, 128.3, 133.7, 139.1, 148.6.

4.4 Oligomer Synthesis

Oligo(phosphodiester)s **S24**-**S41** were synthesized according to the general procedure described in *Automated Poly(phosphodiester) Synthesis* (pp. 135ff.) using phosphoramidite monomers **64**, **72**, **75**, **78**, **109**, **113**, and **117**. The sequences were purified by DMT-ON purification described in subchapter *Manual DMT-ON Purification* (p. 137). Sequences **S28**-**S41** were obtained as triethylammonium salts (see NMR spectra on pages 209ff.) and were analyzed by ^{1}H NMR, ^{31}P NMR, ESI(-)-MS, ESI(-)-HRMS, and ESI(-)-MS/MS (see pages 194ff. and 209ff.).

4.5 Photomodification

A 1.0 mM solution of oligo(phosphodiester)s **S28**-**S41** in methanol-d^4 was prepared under argon atmosphere. 100 µL od said solution were filled into a small glass vial under exclusion of air and this vial was placed on the glass surface of a UV lamp (Vilber Lourmat; λ = 365 nm). The solution was irradiated for 60 min and the solvent was afterwards removed to recover the modified sequences **S28'**-**S41'**.

Annex

1 Comments on Experimental Protocols

1.1 DMT-ON Purification

Sequence definition is guaranteed during phosphoramidite chemistry by additional capping steps that prevent hydroxy-terminal immobilized sequences that did not react during the phosphoramidite coupling from participating in further coupling cycles as it is explained in Chapter I: *Phosphoramidite Chemistry* (pp. 39ff.). When washing sequences from the support after a DMT-ON synthesis, the desired sequence is thus obtained with a DMT terminus together with hydroxy-terminated lower congeners. This mixture is further treated during the *DMT-ON purification* (see procedure in Experimental Section: 1.2.3.2 Manual DMT-ON Purification, p. 137) to isolate the desired sequence from contaminants using a reverse phase column. As shown in Figure V-1 it is the crucial step 2 when a salt wash solution is passed through the purification kit containing 5 v% MeCN. While alone the DMT group of the desired sequence is apolar enough to establish sufficient interaction with the static phase to anchor the whole sequence, lower congeners that do not contain the apolar DMT group are washed out with the mobile phase.

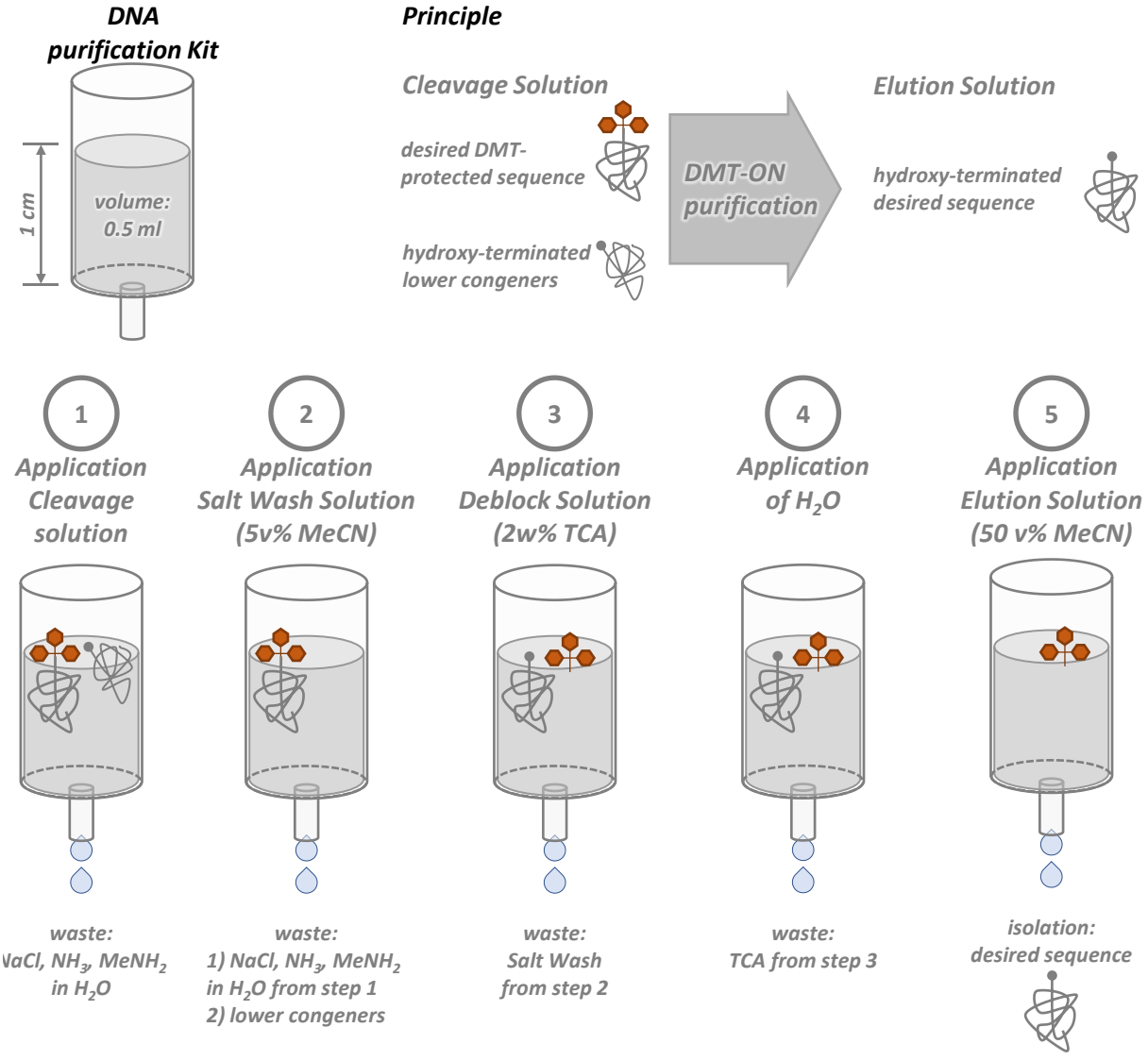


Figure V-1: Overview of DMT-ON purification.

Afterwards the DMT group is cleaved off with TCA (step 3) in 100 % H_2O. Under these conditions the desired sequence is not polar enough to be washed out. Finally in step 5, 50 v% MeCN are applied to wash the purified desired sequence from the reverse phase column.

This protocol was originally conceived for the purification of oligonucleotides. Yet, abiotic poly(phosphodiester)s exhibiting apolar TIPS protective groups were investigated as precursors in Chapter II. Here, the global polarity of the sequence was not governed by the DMT group anymore and step 2 (see Figure V-1) did not result in the separation from lower congeners. This is why a DMT-ON purification could only be applied when all TIPS groups were removed from the intermediates during the sequential post-polymerization modification. This fact also implies that it is not until the final purification that sequence-defined polymers were isolated. Strictly speaking, polymer mixtures were isolated before.

1.2 Sequential Modification of Poly(phosphodiester)s

NMR and mass spectra are presented in this section which were allowing to follow the reaction progress of sequential post-polymerization modifications. As discussed in the paragraph above, the sequences **S7**, **S7'**, and **S7''** are strictly speaking not isolated intermediate products because they do still contain the (modified) lower congeners from the precursor polymer synthesis. Moreover, the first modification was performed when **S7** was immobilized on the support. That is why **S7** was synthesized in another batch and cleaved from the support to access analytic data of that precursor species. In Figure V-2 the ESI(-) mass spectrum of **S7** is depicted. It is striking that only the very low congeners could be detected by mass spectrometry. During each phosphoramidite coupling cycle the chances of a missed step and subsequent chain capping are the same. It is consequently supposed that the potential ionization in MS decreases for poly(phosphodiester)s with increasing length. And as the desired sequence **S7** is detected as the dominant species in multicharged states, **S7** was clearly isolated as the main product.

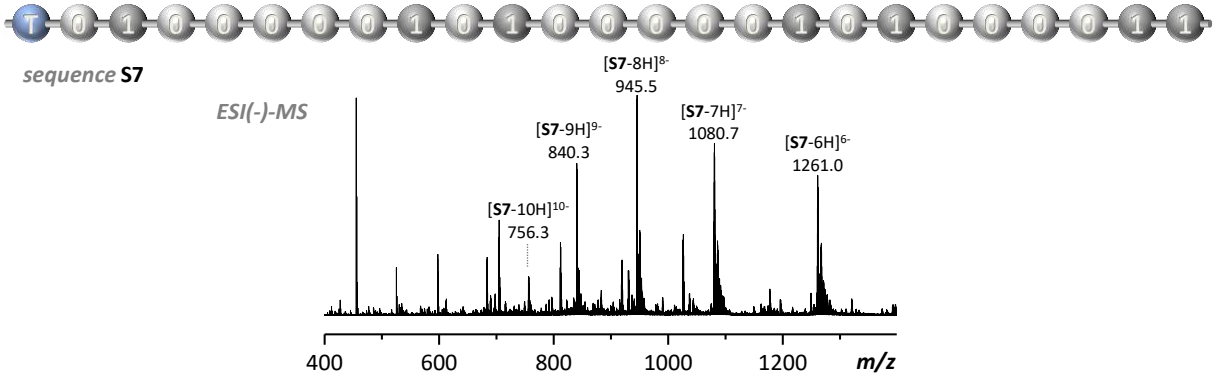


*Figure V-2: ESI(-) mass spectrum of **S7**. Sequence **S7** detected as [**S7** – 6H]$^{6-}$ at m/z = 1261.0, [**S7** – 7H]$^{7-}$ at m/z = 1080.7, [**S7** – 8H]$^{8-}$ at m/z = 945.5, [**S7** – 9H]$^{9-}$ at m/z = 840.3, and [**S7** – 10H]$^{10-}$ at m/z = 756.3. Some lower congeners were detected as deprotonated molecules and identified as T**0**$_1$ (m/z = 455.1; z = 1), T**0**$_2$**1**$_1$ m/z = 1195.4; z = 1 and 597.7; z = 2), T**0**$_3$**1**$_1$ (m/z = 704.7; z = 2), T**0**$_4$**1**$_1$ (m/z = 811.7; z = 2), T**0**$_5$**1**$_1$ (m/z = 919.3; z= 2), and T**0**$_6$**1**$_1$ (m/z = 1026.3; z = 2 and m/z = 683.9; z = 3). Multiple H/Na exchanges account for the peak series observed at higher masses next to the main signals.*

After the CuAAC modification of immobilized DMT-protected **S7** and its congeners, the support material was washed extensively before cleaving the modified sequences from the support. Shown in Figure V-3, target sequence **S7'** was detected by mass spectrometry and ^{1}H NMR. It is apparent that it is even more difficult to analyze the compound by ESI(-)-MS with the insertion of 34 PEG side chains at the 0 bit monomers. Yet, sequence **S7'** was identified in different multicharged states and additionally, H/Cu exchanges were observed. The presence of copper ions is probably also the reason for isolating **S7'** as a blue oil as well as it can explain the very broad peaks observed in NMR. Importantly,

methyleneoxy protons *(g)* are clearly detected as well as the signal for methylene protons adjacent to the newly formed triazole ring *(f)* and the signal for the aromatic proton at the triazole ring *(i)*.

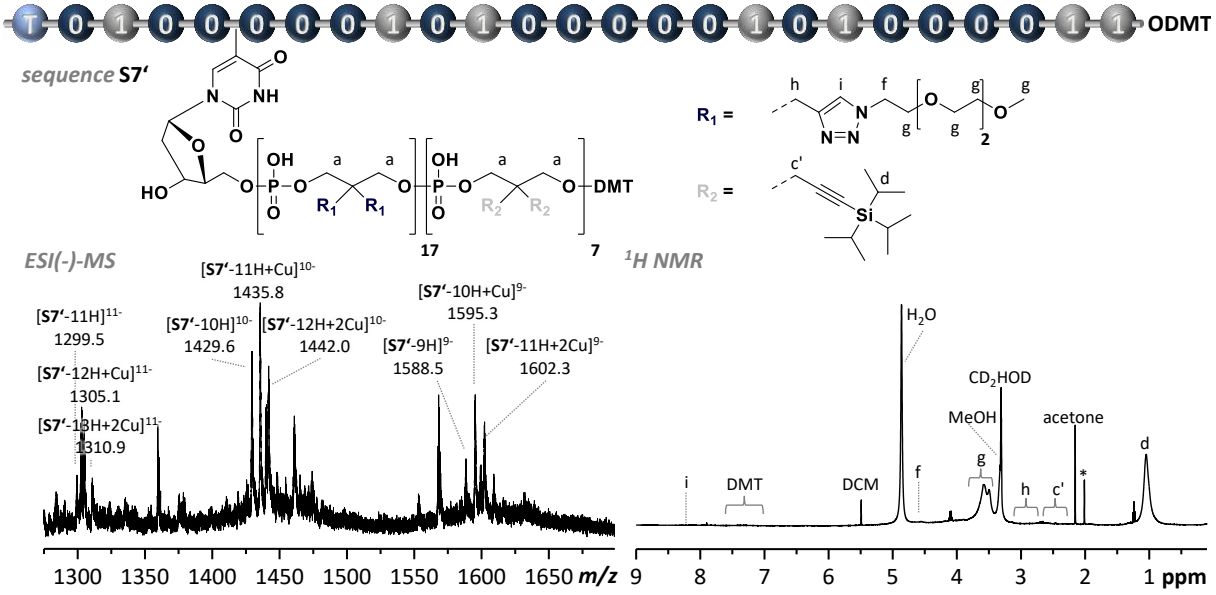


*Figure V-3: Analysis of oligomer **S7'**. Left: ESI(-) mass spectrum; **S7'** detected as [**S7'** – 9H]$^{9-}$ at m/z = 1588.5, [**S7'** – 10H]$^{10-}$ at m/z = 1429.6, and [**S7'** – 11H]$^{11-}$ at m/z = 1299.5, as well as after one H/Cu exchange ([**S7'** - (z+1) H + Cu]$^{z-}$ at m/z = 1595.3 (z = 9), m/z = 1435.8 (z = 10), and m/z = 1305.1 (z = 10) or two H/Cu exchanges ([**S7'** – (z + 2) H + 2 Cu]$^{z-}$ at m/z = 1602.3 (z = 9), m/z = 1442.0 (z = 10), and m/z = 1310.9 (z = 11) in the polymer. Right: ^{1}H NMR spectrum recorded in methanol-d^4 shown in the range of 9.0 – 0.0 ppm; TIPS protons signals d detected at 1.0 ppm, the region 3.4 - 3.8 ppm as well as a weak signal at 4.5 ppm indicate side chain modification through the presence of methylenoxy protons g and methylene protons f adjacent to N1 of triazole rings, respectively.*

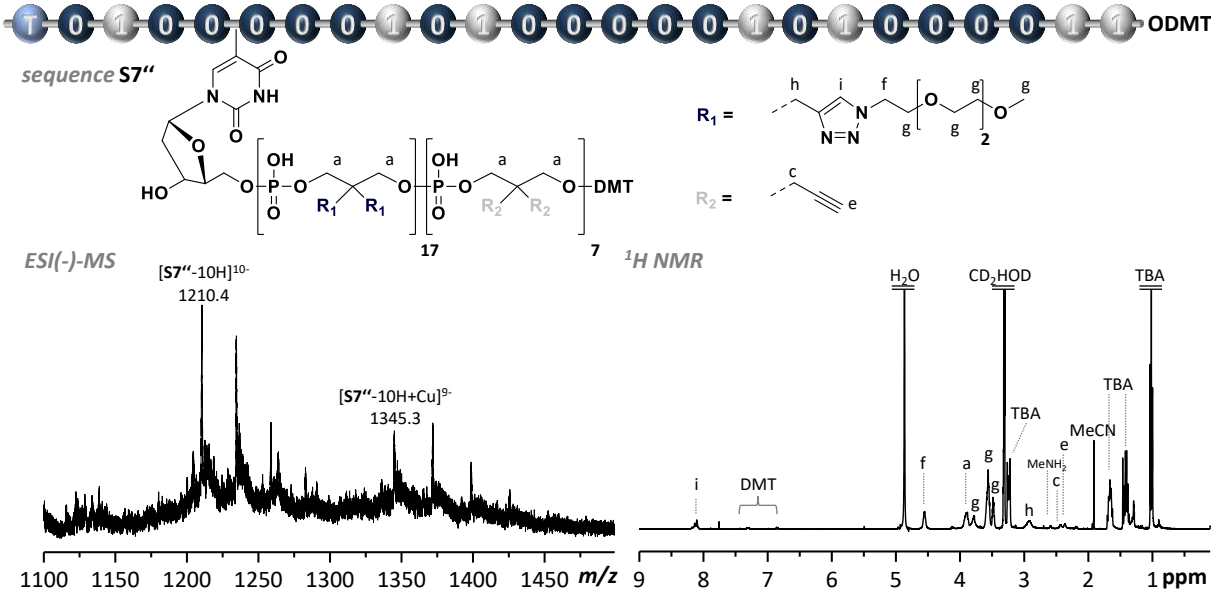


*Figure V-4: Analysis of **S7''**. Left: ESI(-) mass spectrum; **S7''** detected as [**S7''** – 9H]$^{9-}$ at m/z = 1345.3, and [**S7''** – 10H]$^{10-}$ at m/z = 1210.4. Right: : ^{1}H NMR spectrum recorded in methanol-d^4 shown in the range of 9.0 – 0.0 ppm; modified monomers 0 are characterized by signals for triazole protons i at 8.1 ppm, methylene protons f adjacent to N1 of triazole rings at 4.5 ppm, methylenoxy protons g at 3.75 ppm, 3.6 ppm, and 3.4 ppm, and methylene protons h adjacent to C4 of triazole rings at 2.8 ppm; monomers 1 result in signals for terminal alkyne protons e at 2.3 ppm and methylene protons neighboring the terminal alkyne c at 2.4 ppm.*

Sequence **S7''** was obtained after TIPS deprotection with TBAF as indicated by mass spectrometry shown in Figure V-4 on the left. The ^{1}H NMR spectrum on the right of the same figure shows that **S7''** was—even after chromatographic purification from the reagent—isolated as a tetrabutylammonium

(TBA) salt. The signals at 3.3 ppm, 2.6 ppm, 1.4 ppm, and 1.0 ppm originate from the TBA alkyl chains. Yet, copper had been removed and relatively sharp peaks indicate the presence of the attained alkyne proton *(e)*. Quantitative TIPS-deprotection of monomers 1 is indicated by disappearance of signals for TIPS-protons *(d)* at 1.0 ppm that were prominent for **S7'** in Figure V-3.

Finally, the second modification yielded **S7'''** that could be purified from its lower congeners by a DMT-ON purification. The NMR spectrum is shown in Figure II-11 on page 81. The signal at 2.3 ppm for free alkyne protons *(e)* detected for **S7''** (see Figure V-4) disappeared, indicating a quantitative conversion. Binary modified sequence **S7'''** could not be detected by ESI(-) mass spectrometry.

2 Nanopore Sequencing

In the following sections the behavior of abiotic poly(phosphodiester)s is briefly discussed regarding the interactions with each of the three biological nanopores depicted before in Figure II-12 on page 83.[3] For experimental details, the interested reader is directed to the related publication[399] for further reading.

2.1 Interaction with α-Hemolysin

In a first attempt, a suitable biological pore had to be found and analyzed for its interaction with a simple abiotic poly(phosphodiester) **S9**. The first tested biological nanopore was α-hemolysin, a homo-oligomeric heptamer. It comprises a solvent filled channel that is 100 Å in length. Running along the central sevenfold axis of the protein, its diameter ranges from 14 Å to 46 Å. Model sequence **S9** consisted of three leading dT monomers and a block of 56 1,3-propandiol-based units as it is depicted in Figure V-5A. When **S9** was injected at the cis-side of the membrane and a voltage of +100 mV was applied, the characteristic open-pore current I_0 was registered. Negatively charged species ions are translocating through the channel. Apart from the initial open-pore current, a new slightly lower current level was detected. Shown in Figure V-5B, it stayed constant over seconds. This observation is possibly due to strong interaction between **S9** and α-hemolysin. The open-pore current I_0 was not retrieved as long as an electric field was applied. As a control the electric field was inverted by applying a voltage of -100 mV. As it is shown in Figure V-5C, following the expectations the inverse open-pore current stays constant. There is no species at the trans-side of the membrane that could be electrophoreticly transported to the pore and therefore no current variation is registered. In conclusion, the applied electric field induces migration of **S9** to the cis-side of the pore. Having arrived there, the sequence blocks the entry of the pore channel to a certain degree in a constant manner as long as an electric field is applied. α-Hemolysin was, thus, not considered a suitable biological nanopore for sequencing the supplied class of abiotic poly(phosphodiester)s.

[3] Hereby, it is clearly stated that the author of the thesis did supply the collaborators with several analytes and gave advice regarding chemistry-related questions. Practical experiments employing nanopores, data treatment and data interpretation as well as visualization was performed by the collaborators M.Sc. Mordjane Boukhet and her scientific supervisor Jan C. Behrends (both University of Freiburg, Germany). Obtained results are discussed and printed with their permission.

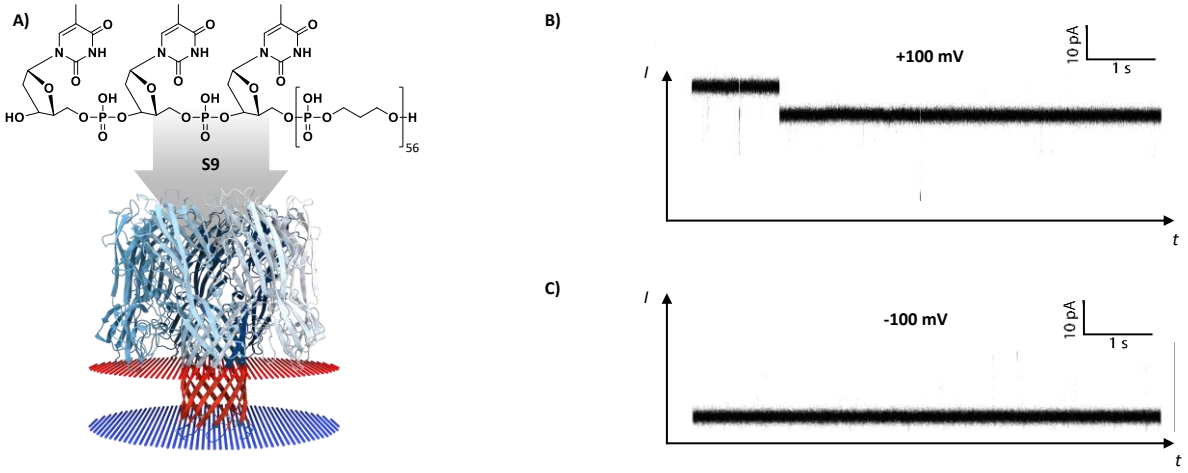


*Figure V-5: Interaction between S9 and α-hemolysin. A) visualization of sequence **S9** and pore; B) current histogram at positive voltage (+100 mV); C) current histogram at negative voltage (-100 mV).*

2.2 Interaction with Aerolysin

The experiment that is depicted above in Figure V-5 was repeated with aerolysin. The transmembrane protein is also a homo-oligomeric heptamer that comprises a solvent filled channel that shows a narrowest diameter with 17 Å in the channel's constriction zone.[400] When applying **S9** to the cis-side

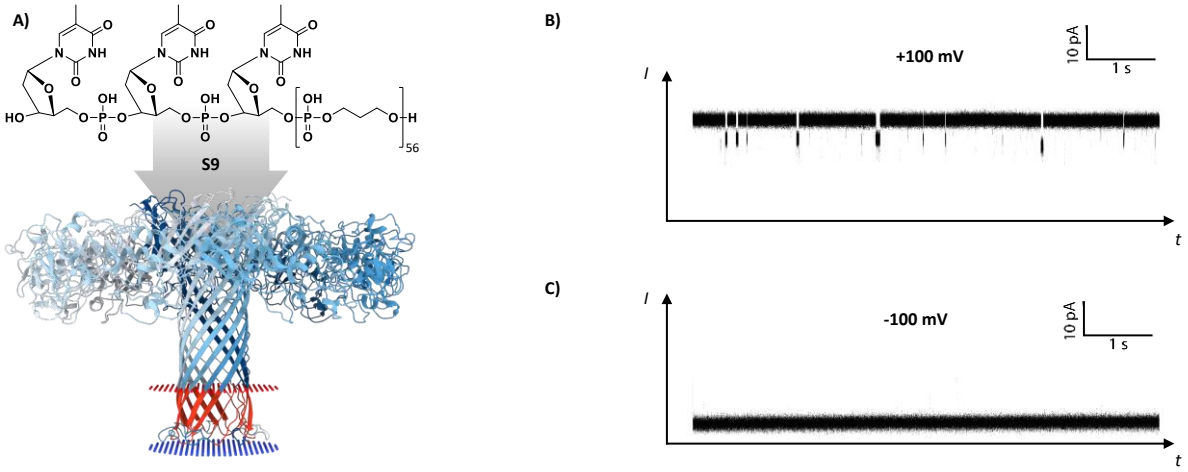


*Figure V-6: Interaction between **S9** and aerolysin. A) visualization of sequence **S9** and pore; B) current histogram at positive voltage (+100 mV); C) current histogram at negative voltage (-100 mV).*

of the membrane (see Figure V-6A) an open pore current I_0 was observed while short events induced two different levels of lower currents as it can be seen in the histogram in Figure V-6B. The control experiment (Figure V-6C) showed that those events are not observed when an inverse electric field is applied. Hence, **S9** was found to interact in short reversible events with aerolysin.

The events that could indirectly reveal information about the sequence were then to be analyzed in depth. The dwell time of each event was plotted over its current level in Figure V-7A when a 300 mM concentration of KCl was used. Event 1 with current I_1 permits less ions to pass and shows a dwell time of up to 100 ms while event 2 causes a current I_2 that is closer to the open-pore current and less defined as I_1. Events of type 2 last considerably shorter and are in the range of milliseconds. The situation was also investigated with higher KCl concentrations. Figure V-7B shows that the relative current of events 1 and 2 does not change, however, dwell times are much longer.

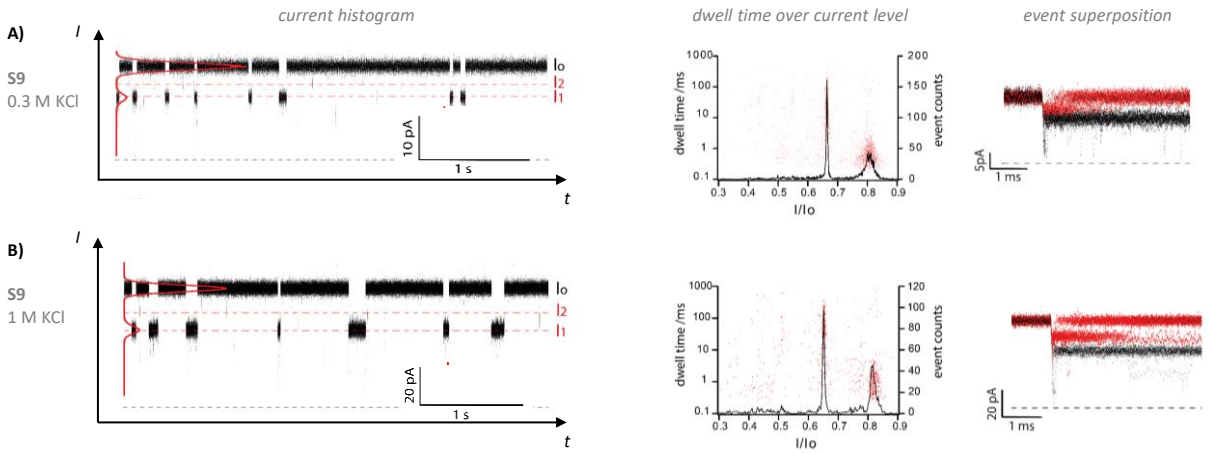


*Figure V-7: Event analysis for **S9** with aerolysin at 120 mV. A) using 0.3 M KCl; B) using 1.0 M KCl.*

When a polymer **S10** with a considerably longer abiotic block of 104 phosphate linked 1,3-propandiol moieties was applied to the *cis*-side of the same nanopore setup, a very similar characteristics was observed (see Figure V-8A and B). Both, event 1 and event 2 block the open-pore current slightly less than it was observed for the shorter **S9**. Dwell times are similar and apparently not strongly depending on the abiotic block's length in the poly(phosphodiester)s.

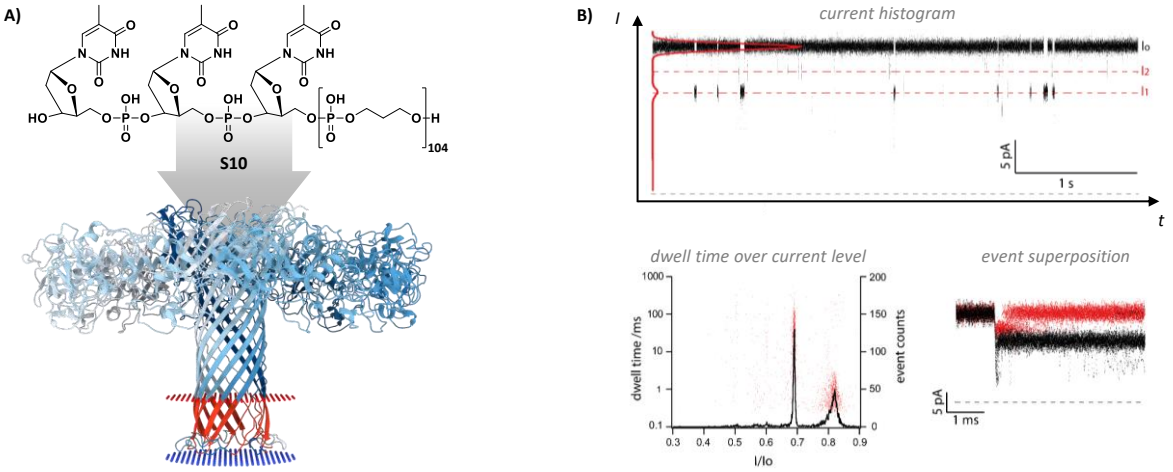


*Figure V-8: Interaction between **S10** and aerolysin at 120 mV. A) Visualization of sequence **S10** and pore protein; B) event analysis using 0.3 M KCl.*

Very interestingly, a different voltage dependence of dwell time was observed for events 1 and 2. Relative currents I_1 and I_2 were measured, identifying events 1 and 2, respectively, at different applied electric fields. Dwell time analysis of the events shows that event 1 is characterized by dwell times that first increase with the applied voltage and after passing a maximum at approximately 150 mV, the dwell time decreases again with the applied voltage as it is shown in Figure V-9 (A and B, middle). According to the working hypothesis explaining this observation, at low electric fields a polymer bundle can block the pore to some degrees and leave the pore again but it is staying at the *cis*-side of the membrane. Higher fields increase the force *sucking* the polymer into its cavity against entropic forces that would lead to migration on the *cis*-side of the membrane: The dwell time of events is prolonged. If a certain critical voltage is reached, the electrical field will be high enough to compensate losses of entropy when the polymer is forced into a conformation that allows it to pass through the membrane channel reaching the *trans*-side of the membrane. The higher the applied voltage is set from this point, the faster is the induced translocation of the polymer and thus, the dwell time reduces with an increasing electrical field. The situation for events 2 is different (see Figure V-9A and B, right). The dwell time increases linearly with the applied voltage, suggesting that the polymer is in a conformation or

orientation that does not permit translocation (at the range of applied voltages). The working hypothesis is visualized in Figure V-10.

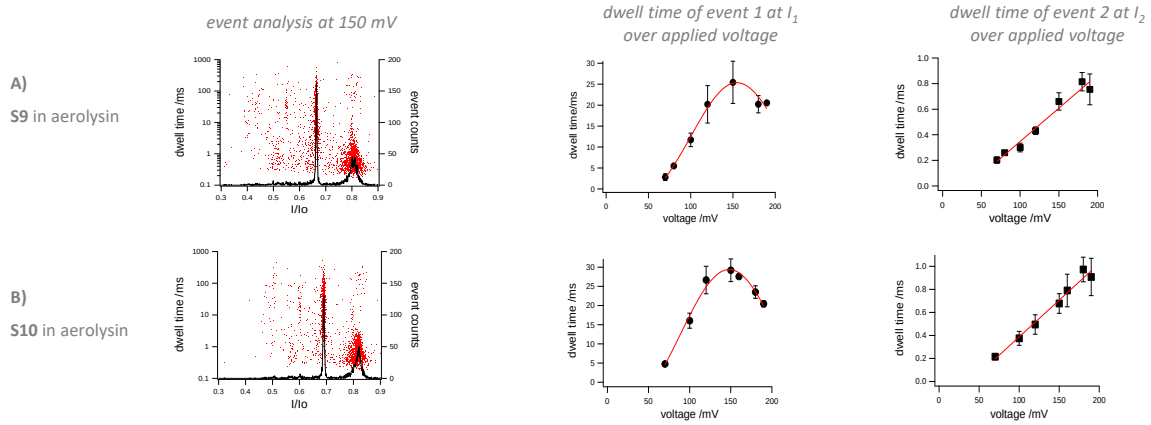


*Figure V-9: In-depth comparison between events 1 and 2 of **S9** (A) and **S10** (B) in aerolysin: Voltage dependence of dwell time.*

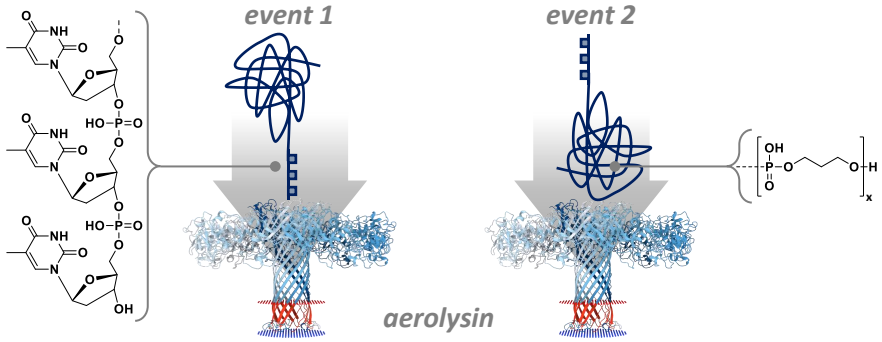


Figure V-10: Hypothesis: Different interactions cause distinguishable events 1 and 2.

First insights about the interaction between the biological nanopore aerolysin and abiotic poly(phosphodiester)s were revealed for **S9** and **S10**. The length of the abiotic part of the polymer with 56 subunits for **S9** and 104 for **S10**, respectively, does not have a qualitative impact on the observation of two distinguishable events. The next analogue polymer under investigation, **S11**, was chosen to exhibit an intermediate length of the abiotic part with 64 subunits while the number of leading deoxythymidine units was reduced to one. As depicted in Figure V-11.A much shorter dwell times were observed during the experiment while the distribution of current levels was less sharp. It is suggested that a decreasing number of deoxythymidine units leads to an overall reduced interaction between polymer and nanopore. When the 1,3-propandiol motif constitutes the abiotic block like in polymers **S9**, **S10**, and **S11**, its contribution to interaction with the pore is small compared to the deoxythymidine chain end. Hence, in analytes **S12** and **S13** 1,3-propandiol moieties were gradually replaced by 2,2-dipropargyl analogues. For block copolymer **S12** that was composed of an equal number of the two comonomers organized in four blocks with each 16 subunits, the dwell time of events increased slightly with respect to analyte **S11** and two discrete current levels were observed (see Figure V-11A). The analogue **S12** exhibiting an alkyne homopolymer block showed two discrete current levels as well and the dwell time of events was considerably increased to the range of some hundred microseconds (see Figure V-11B). When comparing aerolysin interaction of **S11** with that of **S13** it can be concluded that the replacement of the unbranched monomeric units with its 2,2-dipropargyl analogues results in a stronger interaction that is yet still too weak to enable a precise read-out (see Figure V-11C). Thus, sequence **S14**, the triazole analogue of **S11** and **S13**, was tested for its interaction with the pore. Shown in Figure V-12, events with two current levels were observed that showed dwell times above a

millisecond. Inside the transmembrane protein, the triazole moieties can potentially contribute to hydrogen bonds as donors and acceptors.

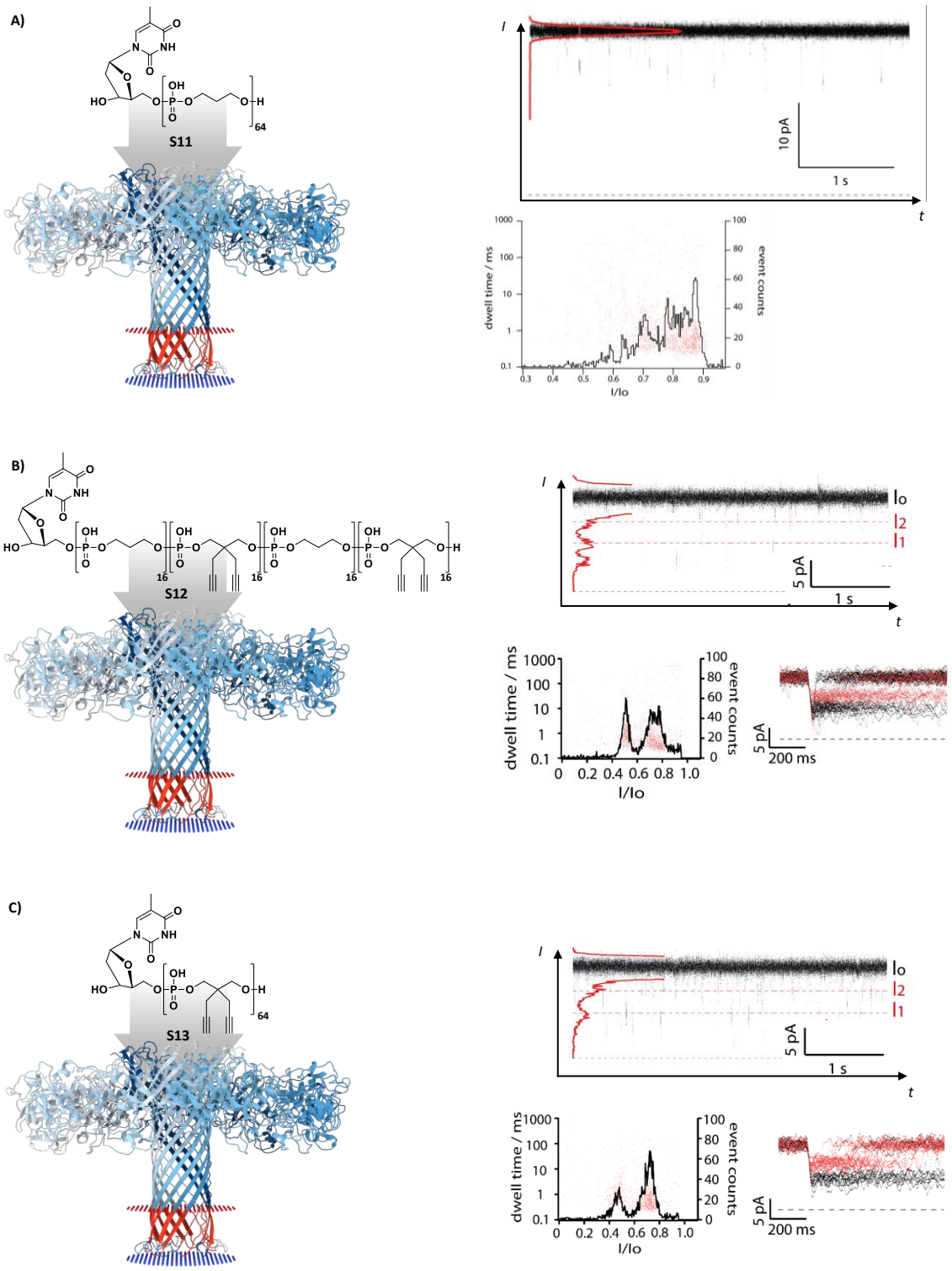


*Figure V-11: Overview interaction of **S11** (A), **S12** (B), and **S13** (C) with aerolysin.*

S14 was synthesized using a post-polymerization modification by CuAAC and it demonstrates that starting from this point, nanopore interaction can be studied as a function of factors like sterical demand, polarity, etc. introduced by the side chains.

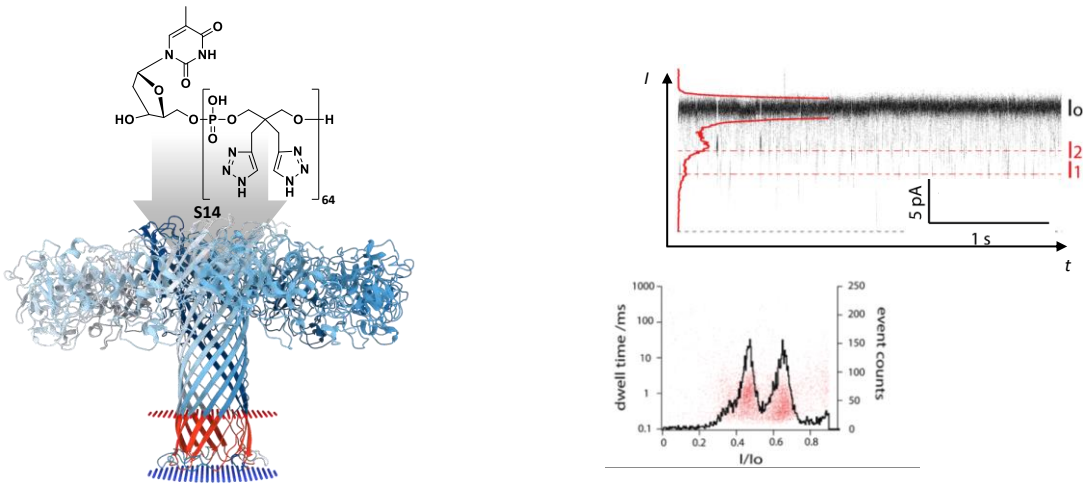


*Figure V-12: Overview interaction of **S14** with aerolysin.*

2.3 Interaction with Mycobacterium Smegmatis Porin A

Interaction of sequence **S11** with MSpA did not result in comprehensive events as can be seen in Figure V-13.

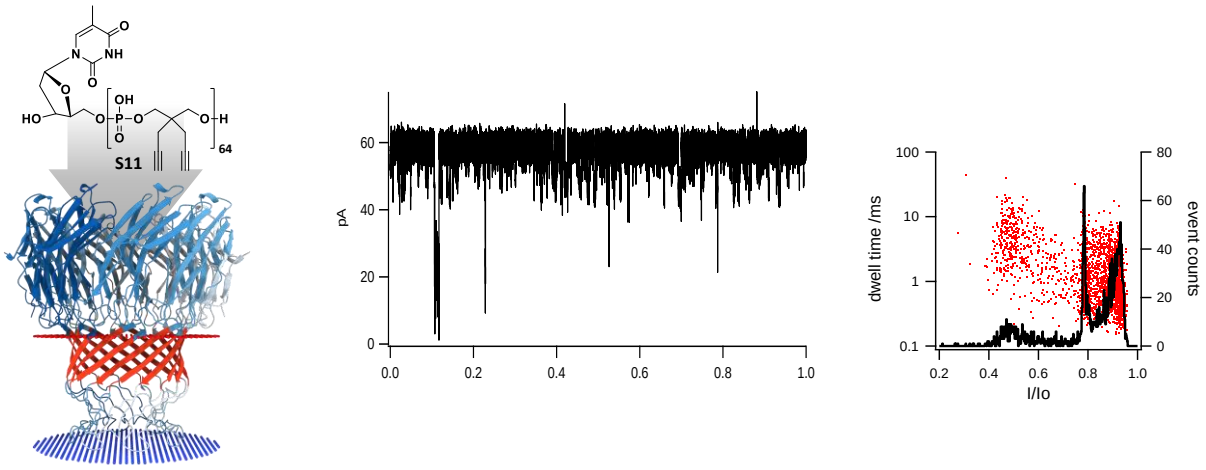


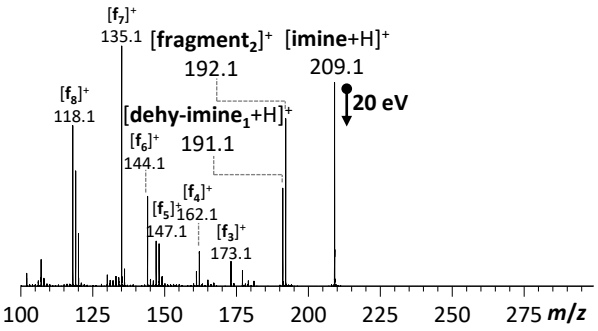
*Figure V-13: Interaction of **S11** with MSpA.*

3 Additional Discussion and Data to Chapter IV

3.1 Proposed Mechanism for Imine Fragmentation

The proposed mechanism for imine fragmentation during tandem mass spectrometry is depicted in Figure V-14B giving a possible structure to fragments **f_1-f_8** detected during the MS/MS analysis shown in Figure V-14A.

A) *MS/MS analysis of [**imine**+H]⁺ peak at 209.1*



B) *Proposed mechanisms for fragmentation of [**imine**+H]⁺ peak at 209.1 during MS/MS*

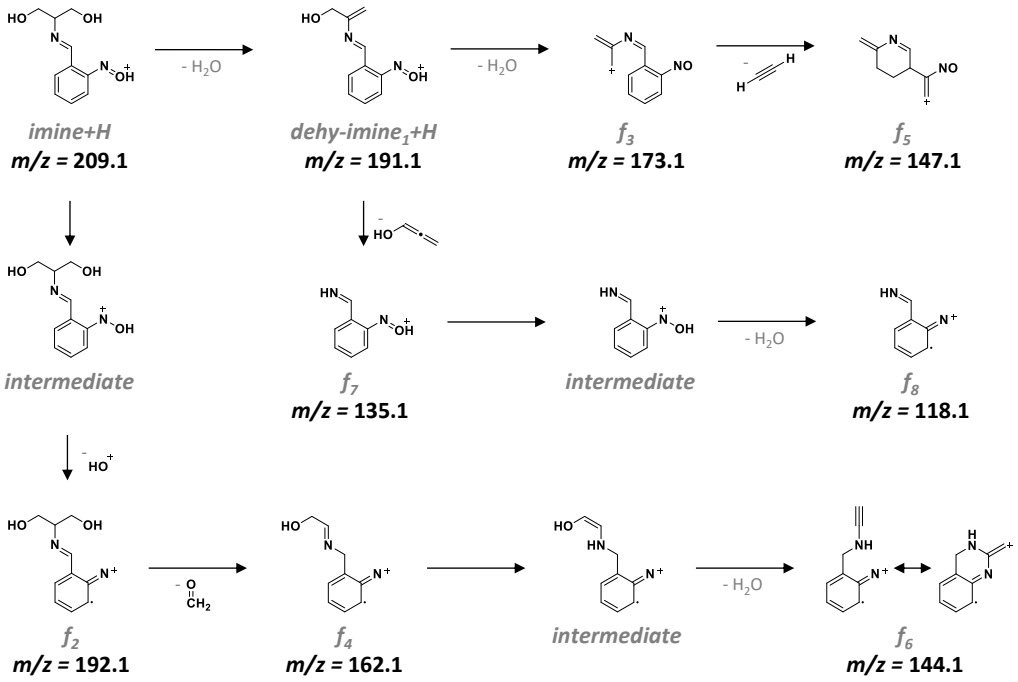


Figure V-14: A) Tandem mass spectrum of [imine+H]⁺ peak; B) proposed fragmentation mechanism.

3.2 Tandem Mass Spectrometry Sequencing of Oligomers S35-S41 and S35'-S41'

The sequences **S35**-**S41** that were synthesized for the concept *Reveal sequence information* were analyzed by mass spectrometry. Tandem mass spectrometry of the [**SX**-3H]$^{3-}$ peak at 850.4 yielded identical spectra to the one of **S35** that is depicted on the left of Figure V-15. That harmonizes with the expectations as the monomeric units were regioisomers and thus not distinguishable by mass spectrometry. Only for sequence **S35**, the analyte is not modified upon irradiation due to the fact that it is only exhibiting inert *para*-nitrobenzylic ethers. Hence, the MS/MS spectrum of the [**S35'**-3H]$^{3-}$ ion is identical to the MS/MS spectrum of the [**S35**-3H]$^{3-}$ ion as it is shown in Figure V-15.

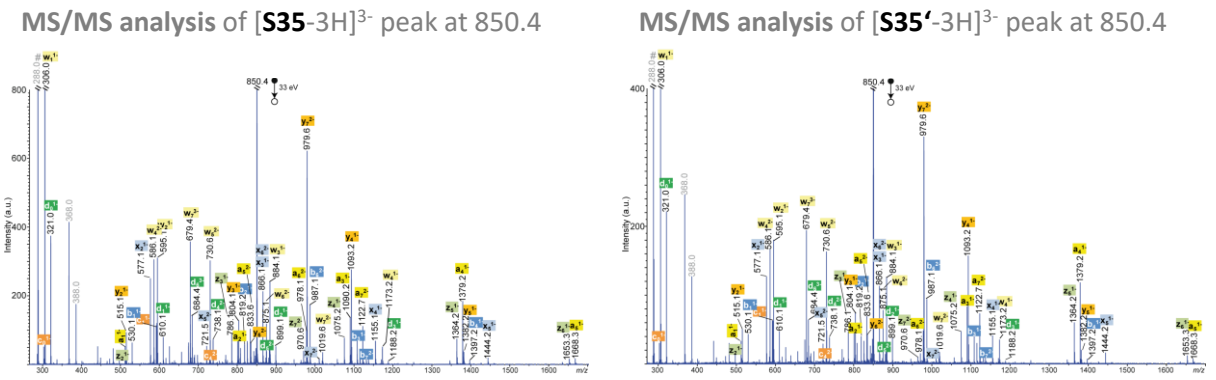


*Figure V-15: MS/MS spectra of the [**S35**–3H]$^{3-}$ ion (left) and the [**S35'**–3H]$^{3-}$ ion at m/z 850.4.*

A brief introduction to the interpretation of the MS/MS spectra should be given here. To deduce the sequence information, the detected ions have to be assigned to several possible fragmentation pattern. The result is shown in Table V-1 below. In Figure V-16 the fragmentation for series *d* and *w* are depicted. These fragments d_0-d_6 and w_1-w_7 were detected in their multiple charged states in both MS/MS spectra (compare above Figure V-15). Either in α→ω direction for series *d* or in ω→α direction for series *w* the mass difference between the entities of the series reveals the primary structure. In fact that is only the case if each monomeric unit is defined by its weight. Here, it is found that the repeating mass difference of 289 Da can only stem from a *para*-nitrobenzylic ether containing monomer after irradiation. That allows for this case the reconstruction of the primary structure α-T11111111-ω for sequence **S35'**. This result is already reflected by the middle line in Table V-1. As *ortho*- and *para*-nitrobenzylic ether containing monomers are characterized by the same mass, the sequence information of **S35** as well as the sequence information of all non-irradiated sequence **S36-S41** cannot be deduced.

*Table V-1: List of m/z values measured for fragments of the [**S35'**–3H]$^{3-}$ ion at m/z 850.4. n.e.: not expected; n.d.: not detected.*

a ↓	b ↓	c ↓	d ↓		w ↑	x ↑	y ↑	z ↑
n.e.	*n.e.*	303.0^{1-}	321.0^{1-}	**α**	*n.e.*	*n.e.*	*n.e.*	*n.e.*
512.1^{1-}	530.1^{1-}	592.1^{1-}	610.1^{1-}	**1**	*n.e.*	769.8^{3-}	749.1^{3-}	*n.d.*
801.1^{1-}	819.2^{1-}	881.1^{1-}	899.1^{1-}	**1**	679.4^{3-}	1010.6^{2-}	979.6^{2-}	970.6^{2-}
1090.2^{1-}	1108.2^{1-}	1170.2^{1-}	1188.2^{1-}	**1**	875.1^{2-}	866.1^{2-}	835.1^{2-}	1653.3^{1-}
1379.2^{1-}	1397.2^{1-}	729.1^{2-}	738.1^{2-}	**1**	730.6^{2-}	1444.2^{1-}	1382.2^{1-}	1364.2^{1-}
1668.3^{1-}	842.6^{2-}	873.6^{2-}	882.6^{2-}	**1**	586.1^{2-}	1155.1^{1-}	1093.2^{1-}	1075.2^{1-}
978.1^{2-}	987.1^{2-}	1018.1^{2-}	684.4^{3-}	**1**	884.1^{1-}	866.1^{1-}	804.1^{1-}	786.1^{1-}
1122.7^{2-}	1131.7^{2-}	774.8^{3-}	*n.e.*	**1**	595.1^{1-}	577.1^{1-}	515.1^{1-}	497.1^{1-}
n.e.	*n.e.*	*n.e.*	*n.e.*	**1-ω**	306.0^{1-}	*n.e.*	*n.e.*	*n.e.*

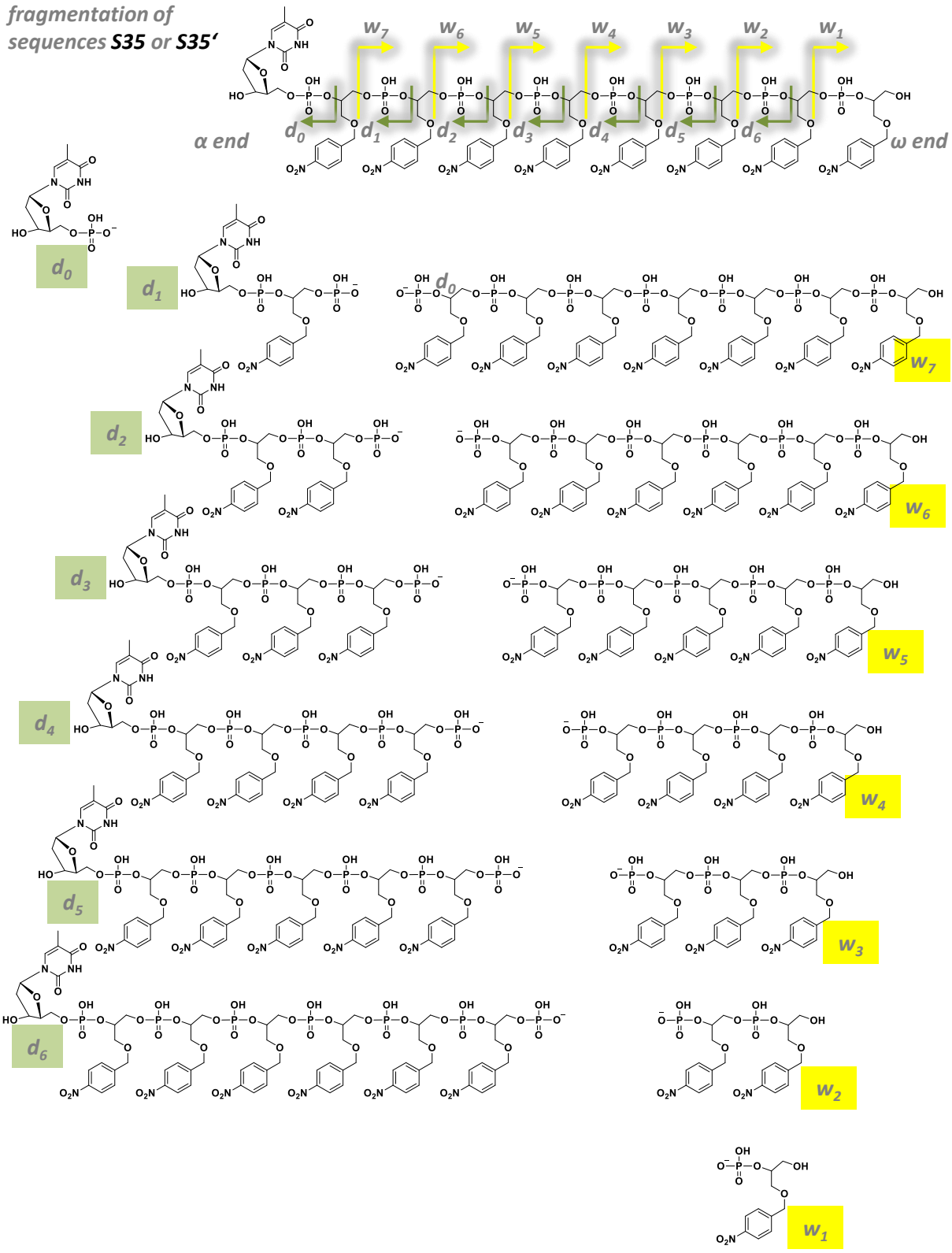


*Figure V-16: Possible fragmentation pattern of the triply charged ion of sequence **S35** or identical **S35'**.*

Followingly, the MS/MS spectra of the [**SX'**-3H]$^{3-}$ peak and analysis tables to deduce the sequence information of irradiated analytes are given for the rest of the series of sequences.

Deduced from MS/MS analysis (Figure V-17) and shown in Table V-2, the primary structure of sequence **S36'** is determined to be α-T01000101-ω, coding for ***E*** in the ASCII alphabet.

MS/MS analysis of [**S36'**-3H]$^{3-}$ peak at 625.4

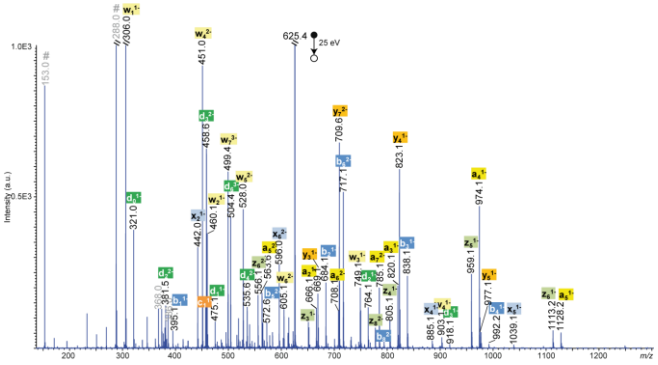


*Figure V-17: MS/MS spectrum of the [**S36'**–3H]$^{3-}$ ion at m/z 625.4.*

*Table V-2: List of m/z values measured for fragments of the [**S36'**–3H]$^{3-}$ ion at m/z 625.4 after irradiation. n.e.: not expected; n.d.: not detected.*

a ↓	b ↓	c ↓	d ↓		w ↑	x ↑	y ↑	z ↑
n.e.	*n.e.*	303.0^{1-}	321.0^{1-}	**α**	*n.e.*	*n.e.*	*n.e.*	*n.e.*
377.1^{1-}	395.1^{1-}	457.1^{1-}	475.1^{1-}	**0**	*n.e.*	544.7^{3-}	*n.d.*	777.6^{2-}
666.1^{1-}	684.2^{1-}	746.1^{1-}	764.1^{1-}	**1**	499.4^{3-}	740.6^{2-}	709.6^{2-}	700.6^{2-}
820.2^{1-}	838.2^{1-}	900.1^{1-}	918.1^{1-}	**0**	605.1^{2-}	596.0^{2-}	565.1^{2-}	1113.2^{1-}
974.1^{1-}	992.2^{1-}	526.6^{2-}	535.6^{2-}	**0**	528.0^{2-}	1039.1^{1-}	977.1^{1-}	959.1^{1-}
1128.2^{1-}	572.6^{2-}	603.6^{2-}	612.6^{2-}	**0**	451.0^{2-}	885.1^{1-}	823.1^{1-}	805.1^{1-}
708.1^{2-}	717.1^{2-}	748.1^{2-}	504.4^{3-}	**1**	749.1^{1-}	731.1^{1-}	669.1^{1-}	651.1^{1-}
785.1^{2-}	794.1^{2-}	549.7^{3-}	*n.e.*	**0**	460.1^{1-}	442.0^{1-}	380.1^{1-}	362.1^{1-}
n.e.	*n.e.*	*n.e.*	*n.e.*	**1-ω**	306.0^{1-}	*n.e.*	*n.e.*	*n.e.*

Deduced from MS/MS analysis (Figure V-18) and shown in Table V-3, the primary structure of sequence **S37'** is determined to be α-T01111000-ω, coding for ***x*** in the ASCII alphabet.

MS/MS analysis of [**S37'**-3H]$^{3-}$ peak at 670.4

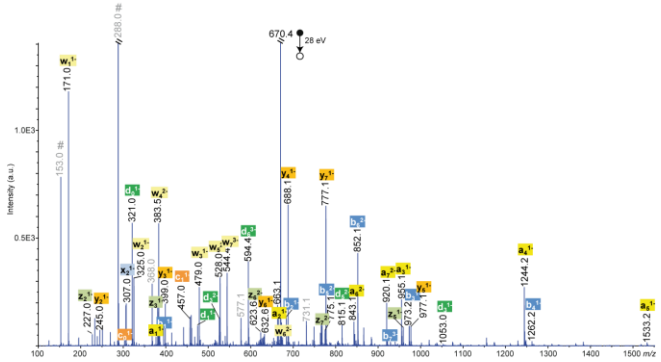


*Figure V-18: MS/MS spectrum of the [**S37'**–3H]$^{3-}$ ion at m/z 670.4.*

*Table V-3: List of m/z values measured for fragments of the [**S37'**–3H]$^{3-}$ ion at m/z 670.4 after irradiation. n.e.: not expected; n.d.: not detected.*

a ↓	b ↓	c ↓	d ↓		w ↑	x ↑	y ↑	z ↑
n.e.	*n.e.*	303.0^{1-}	321.0^{1-}	**α**	*n.e.*	*n.e.*	*n.e.*	*n.e.*
377.1^{1-}	395.1^{1-}	457.1^{1-}	475.1^{1-}	**0**	*n.e.*	589.7^{3-}	*n.d.*	845.1^{2-}
666.1^{1-}	684.2^{1-}	746.1^{1-}	764.1^{1-}	**1**	544.4^{3-}	808.1^{2-}	777.0^{2-}	768.1^{2-}
955.1^{1-}	973.2^{1-}	1035.1^{1-}	1053.1^{1-}	**1**	672.6^{2-}	*n.d.*	632.6^{2-}	623.6^{2-}
1244.2^{1-}	1262.2^{1-}	661.6^{2-}	670.5^{2-}	**1**	528.0^{2-}	519.0^{2-}	977.1^{1-}	959.1^{1-}
1533.2^{1-}	775.1^{2-}	603.6^{2-}	815.1^{2-}	**1**	383.5^{2-}	474.5^{2-}	688.1^{1-}	670.1^{1-}
843.1^{2-}	852.1^{2-}	806.1^{2-}	594.4^{3-}	**0**	479.0^{1-}	461.0^{1-}	399.0^{1-}	381.0^{1-}
920.1^{2-}	929.1^{2-}	549.7^{3-}	*n.e.*	**0**	325.0^{1-}	307.0^{1-}	245.0^{1-}	227.0^{1-}
n.e.	*n.e.*	*n.e.*	*n.e.*	**0-ω**	171.0^{1-}	*n.e.*	*n.e.*	*n.e.*

Deduced from MS/MS analysis (Figure V-19) and shown in Table V-4, the primary structure of sequence **S38'** is determined to be α-T01110000-ω, coding for ***p*** in the ASCII alphabet.

MS/MS analysis of [**S38'**-3H]$^{3-}$ peak at 625.4

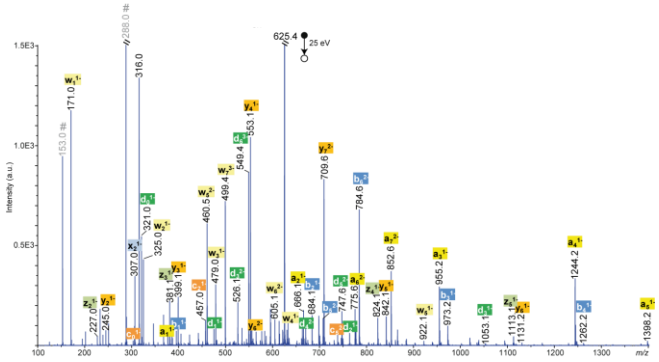


*Figure V-19: MS/MS spectrum of the [**S38'**–3H]$^{3-}$ ion at m/z 625.4.*

*Table V-4: List of m/z values measured for fragments of the [**S38'**–3H]$^{3-}$ ion at m/z 625.4 after irradiation. n.e.: not expected; n.d.: not detected.*

a ↓	b ↓	c ↓	d ↓		w ↑	x ↑	y ↑	z ↑
n.e.	*n.e.*	303.0^{1-}	321.0^{1-}	**α**	*n.e.*	*n.e.*	*n.e.*	*n.e.*
377.1^{1-}	395.1^{1-}	457.1^{1-}	475.1^{1-}	**0**	*n.e.*	544.7^{3-}	*n.d.*	777.6^{2-}
666.1^{1-}	684.2^{1-}	746.1^{1-}	764.1^{1-}	**1**	499.3^{3-}	740.5^{2-}	709.6^{2-}	700.6^{2-}
955.1^{1-}	973.2^{1-}	1035.1^{1-}	526.1^{2-}	**1**	605.1^{2-}	596.0^{2-}	1131.2^{1-}	1113.1^{1-}
1244.2^{1-}	1262.2^{1-}	661.6^{2-}	670.6^{2-}	**1**	460.5^{2-}	451.5^{2-}	845.1^{1-}	824.1^{1-}
1398.2^{1-}	707.6^{2-}	738.6^{2-}	747.6^{2-}	**0**	633.0^{1-}	*n.d.*	553.1^{1-}	535.1^{1-}
775.6^{2-}	784.6^{2-}	815.6^{2-}	549.4^{3-}	**0**	479.0^{1-}	461.0^{1-}	399.0^{1-}	381.0^{1-}
852.6^{2-}	861.6^{2-}	594.7^{3-}	*n.e.*	**0**	325.0^{1-}	307.0^{1-}	245.0^{1-}	227.0^{1-}
n.e.	*n.e.*	*n.e.*	*n.e.*	**0-ω**	171.0^{1-}	*n.e.*	*n.e.*	*n.e.*

Deduced from MS/MS analysis (Figure V-20) and shown in Table V-5, the primary structure of sequence **S39'** is determined to be α-T01110000-ω, coding for ***o*** in the ASCII alphabet.

MS/MS analysis of [**S39'**-3H]$^{3-}$ peak at 760.4

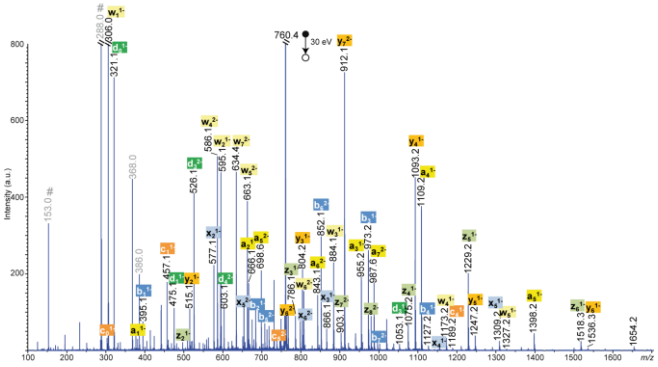


*Figure V-20: MS/MS spectrum of the [**S39'**–3H]$^{3-}$ ion at m/z 760.4.*

*Table V-5: List of m/z values measured for fragments of the [**S39'**–3H]$^{3-}$ ion at m/z 760.4 after irradiation. n.e.: not expected; n.d.: not detected.*

a ↓	b ↓	c ↓	d ↓		w ↑	x ↑	y ↑	z ↑
n.e.	*n.e.*	303.0^{1-}	321.0^{1-}	**α**	*n.e.*	*n.e.*	*n.e.*	*n.e.*
377.1^{1-}	395.1^{1-}	457.1^{1-}	475.1^{1-}	**0**	*n.e.*	679.7^{3-}	*n.d.*	980.1^{2-}
666.1^{1-}	684.1^{1-}	746.1^{1-}	764.1^{1-}	**1**	634.4^{3-}	943.1^{2-}	912.1^{2-}	903.1^{2-}
955.1^{1-}	973.2^{1-}	1035.1^{1-}	526.1^{2-}	**1**	807.6^{2-}	798.6^{2-}	1536.3^{1-}	1518.3^{1-}
1109.2^{1-}	1127.2^{1-}	594.1^{2-}	603.1^{2-}	**0**	663.1^{2-}	1309.2^{1-}	1247.2^{1-}	1229.2^{1-}
1398.2^{1-}	707.6^{2-}	738.6^{2-}	747.6^{2-}	**1**	586.1^{2-}	1155.2^{1-}	1093.2^{1-}	1075.2^{1-}
843.1^{2-}	852.1^{2-}	883.1^{2-}	549.4^{3-}	**1**	884.1^{1-}	866.1^{1-}	804.2^{1-}	786.1^{1-}
986.7^{2-}	996.6^{2-}	684.7^{3-}	*n.e.*	**1**	595.1^{1-}	577.1^{1-}	515.1^{1-}	497.1^{1-}
n.e.	*n.e.*	*n.e.*	*n.e.*	**1-ω**	306.0^{1-}	*n.e.*	*n.e.*	*n.e.*

Deduced from MS/MS analysis (Figure V-21) and shown in Table V-6, the primary structure of sequence **S40'** is determined to be α-T01110000-ω, coding for ***s*** in the ASCII alphabet.

MS/MS analysis of [**S40'**-3H]$^{3-}$ peak at 715.4

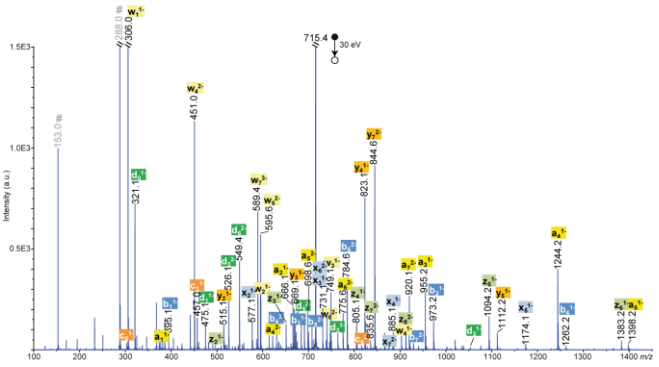


*Figure V-21: MS/MS spectrum of the [**S40'**–3H]$^{3-}$ ion at m/z 715.4.*

*Table V-6: List of m/z values measured for fragments of the [**S40'**–3H]$^{3-}$ ion at m/z 715.4 after irradiation. n.e.: not expected; n.d.: not detected.*

a ↓	b ↓	c ↓	d ↓		w ↑	x ↑	y ↑	z ↑
n.e.	*n.e.*	303.0^{1-}	321.1^{1-}	**α**	*n.e.*	*n.e.*	*n.e.*	*n.e.*
377.1^{1-}	395.1^{1-}	457.0^{1-}	475.1^{1-}	**0**	*n.e.*	634.7^{3-}	*n.d.*	912.6^{2-}
666.1^{1-}	684.2^{1-}	746.1^{1-}	764.1^{1-}	**1**	589.4^{3-}	875.6^{2-}	844.6^{2-}	835.6^{2-}
955.1^{1-}	973.2^{1-}	1035.1^{1-}	526.1^{2-}	**1**	740.1^{2-}	731.1^{2-}	700.1^{2-}	1383.2^{1-}
1244.2^{1-}	1262.2^{1-}	661.6^{2-}	670.6^{2-}	**1**	595.6^{2-}	1174.1^{1-}	1112.2^{1-}	1094.2^{1-}
1398.2^{1-}	707.6^{2-}	738.6^{2-}	747.6^{2-}	**0**	451.1^{2-}	885.1^{1-}	823.1^{1-}	805.1^{1-}
775.6^{2-}	784.6^{2-}	815.6^{2-}	549.4^{3-}	**0**	749.1^{1-}	731.1^{1-}	669.1^{1-}	651.1^{1-}
920.1^{2-}	929.1^{2-}	639.7^{3-}	*n.e.*	**1**	595.1^{1-}	577.1^{1-}	515.1^{1-}	497.1^{1-}
n.e.	*n.e.*	*n.e.*	*n.e.*	**1-ω**	306.0^{1-}	*n.e.*	*n.e.*	*n.e.*

Deduced from MS/MS analysis (Figure V-22) and shown in Table V-7, the primary structure of sequence **S41'** is determined to be α-T01110000-ω, coding for ***e*** in the ASCII alphabet.

MS/MS analysis of [**S41'**-3H]$^{3-}$ peak at 670.4

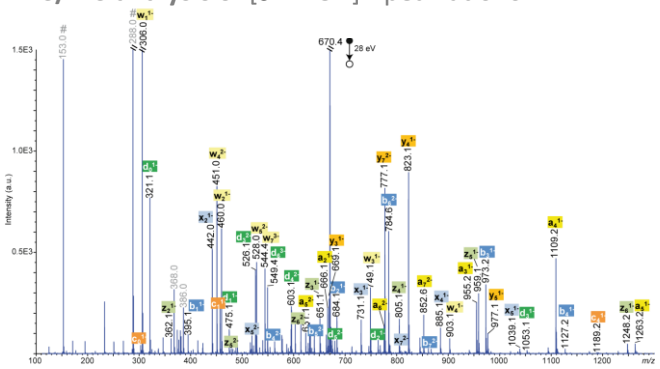


*Figure V-22: MS/MS spectrum of the [**S41'**–3H]$^{3-}$ ion at m/z 670.4.*

*Table V-7: List of m/z values measured for fragments of the [**S41'**–3H]$^{3-}$ ion at m/z 670.4 after irradiation. n.e.: not expected; n.d.: not detected.*

a ↓	b ↓	c ↓	d ↓		w ↑	x ↑	y ↑	z ↑
n.e.	*n.e.*	303.0^{1-}	321.1^{1-}	**α**	*n.e.*	*n.e.*	*n.e.*	*n.e.*
377.1^{1-}	395.1^{1-}	457.0^{1-}	475.1^{1-}	**0**	*n.e.*	589.7^{3-}	*n.d.*	845.1^{2-}
666.1^{1-}	684.1^{1-}	746.1^{1-}	764.1^{1-}	**1**	544.4^{3-}	808.1^{2-}	777.1^{2-}	768.1^{2-}
955.2^{1-}	973.2^{1-}	1035.1^{1-}	526.1^{2-}	**1**	672.6^{2-}	663.6^{2-}	632.6^{2-}	1248.2^{1-}
1109.2^{1-}	1127.2^{1-}	1189.2^{1-}	603.1^{2-}	**0**	528.0^{2-}	1039.1^{1-}	977.1^{1-}	959.1^{1-}
1263.2^{1-}	640.1^{2-}	671.5^{2-}	680.1^{2-}	**0**	451.0^{2-}	885.1^{1-}	823.1^{1-}	805.1^{1-}
775.6^{2-}	784.6^{2-}	815.6^{2-}	549.4^{3-}	**1**	749.0^{1-}	731.1^{1-}	669.1^{1-}	651.1^{1-}
852.6^{2-}	861.6^{2-}	594.7^{3-}	*n.e.*	**0**	460.0^{1-}	442.0^{1-}	380.0^{1-}	362.1^{1-}
n.e.	*n.e.*	*n.e.*	*n.e.*	**1-ω**	306.0^{1-}	*n.e.*	*n.e.*	*n.e.*

4 Analysis of Key Compounds from Chapter II

***Phosphoramidite monomer* 5**

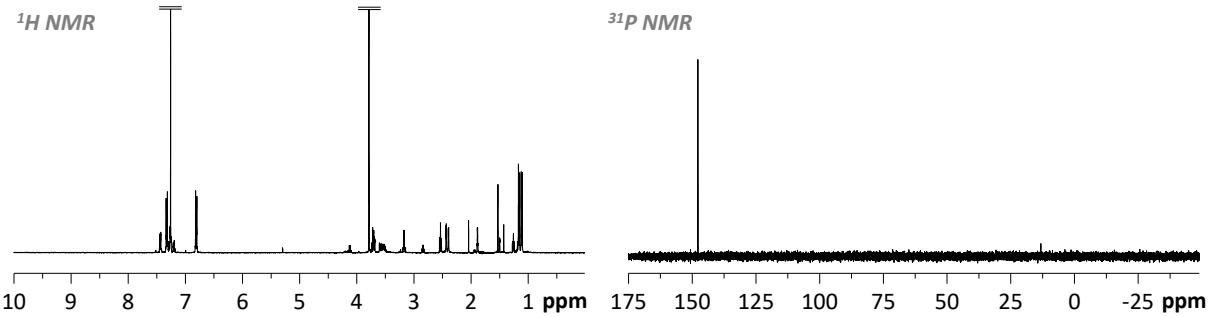


*Figure V-23: Characterizationo of phosphoramidite monomer **5**.*

Phosphoramidite monomer 10

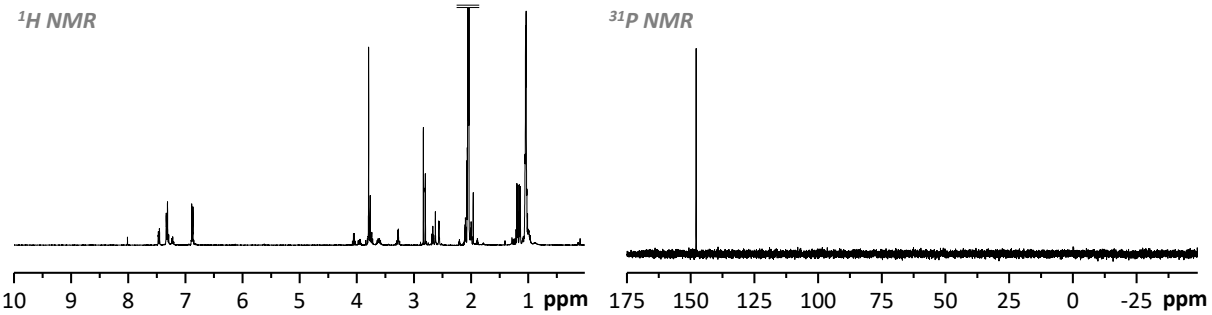


*Figure V-24: Characterization of phosphoramidite monomer **10**.*

Sequence S1 ***ESI(-)-MS***

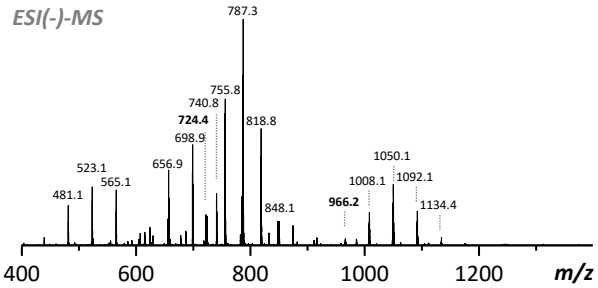


*Figure V-25: Characterization of sequence **S1**.*

Sequence S2

ESI(-)-MS ***MS/MS of [S2-2H]$^{2-}$ peak at m/z = 708.4***

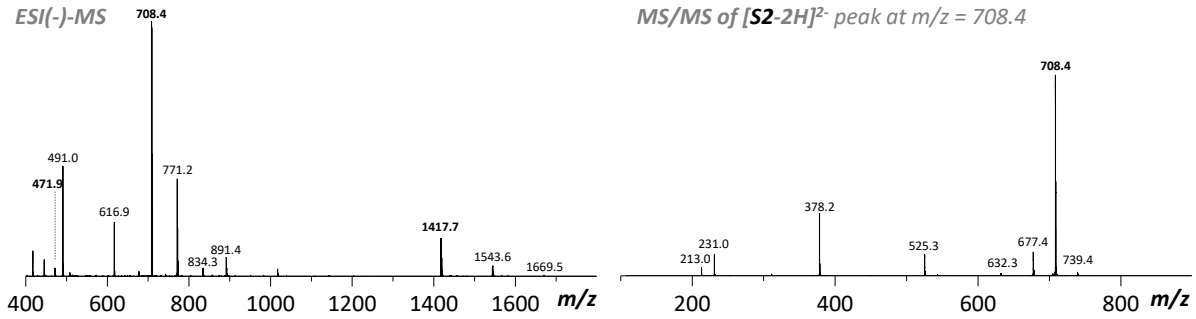


*Figure V-26: Characterization of sequence **S2**.*

Sequence S3

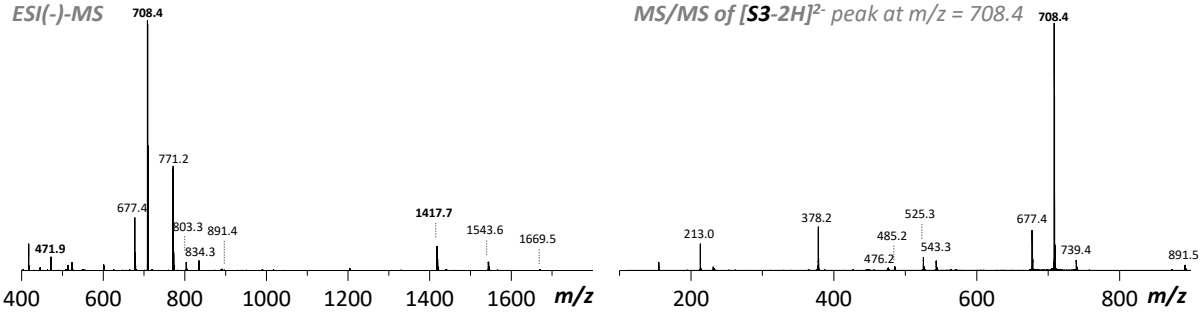


*Figure V-27: Characterization of sequence **S3**.*

Sequence S4

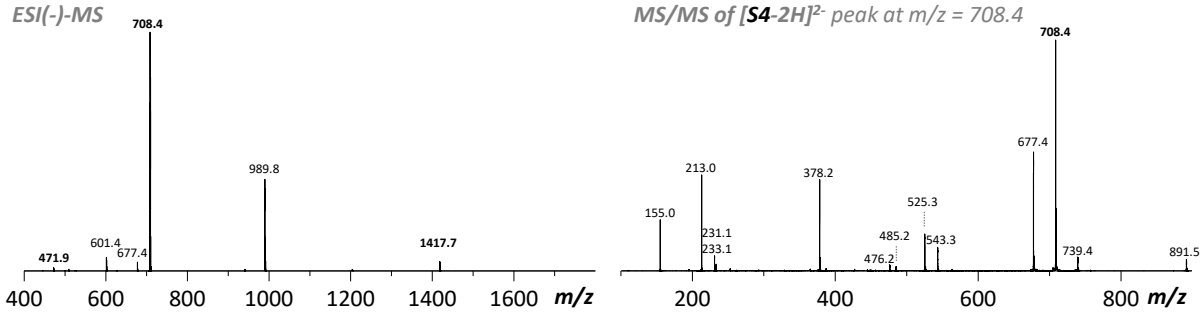


*Figure V-28: Characterization of sequence **S4**.*

Sequence S4'

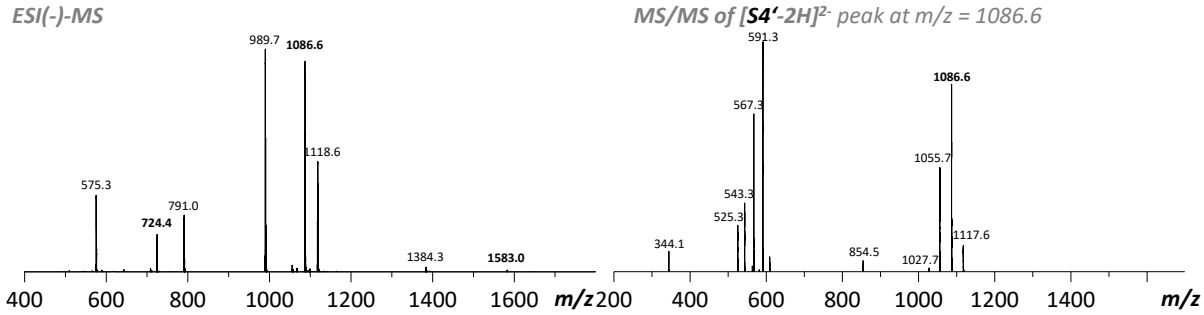


*Figure V-29: Characterization of sequence **S4'**.*

Sequence S4''

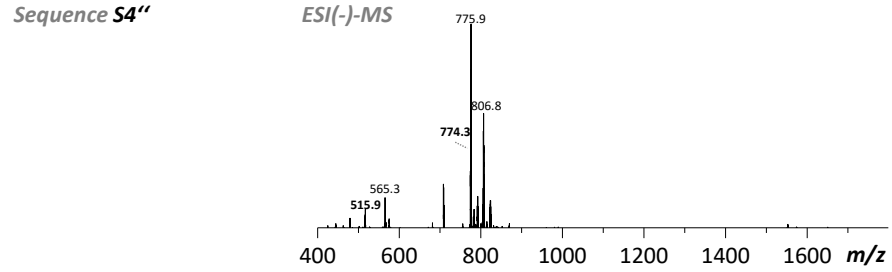


*Figure V-30: Characterization of sequence **S4''**.*

*Sequence **S5***

ESI(-)-MS

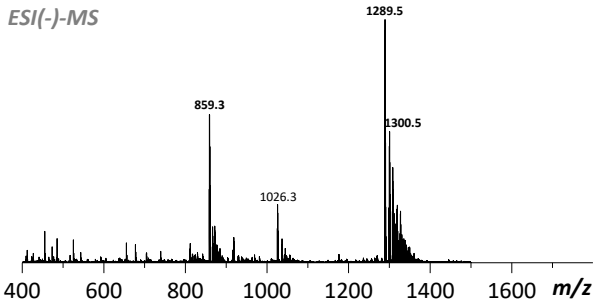
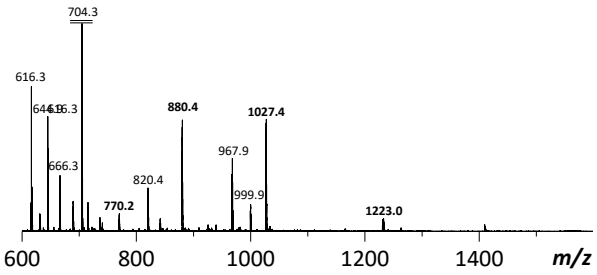


*Figure V-31: Characterization of sequence **S5**.*

*Sequence **S6***

ESI(-)-MS



*MS/MS of [**S6**-6H]$^{6-}$ peak at m/z = 1027.4*

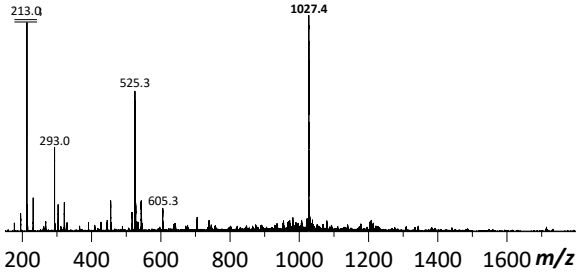


*Figure V-32: Characterization of sequence **S6**.*

*Sequence **S7***

ESI(-)-MS

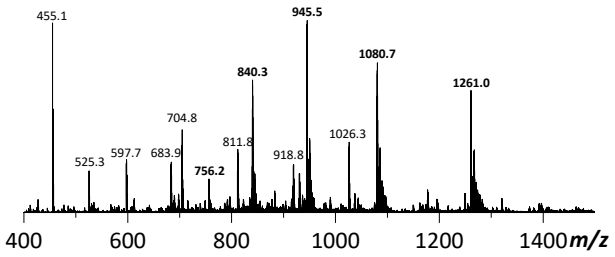


*Figure V-33: Characterization of sequence **S7**.*

*Sequence **S11***

ESI(-)-MS

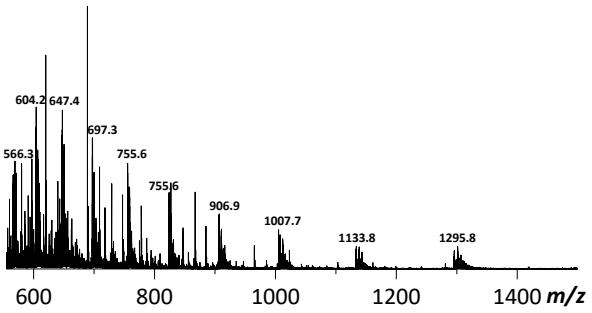


*Figure V-34: Characterization of sequence **S11**.*

Sequence ***S12*** *ESI(-)-MS*

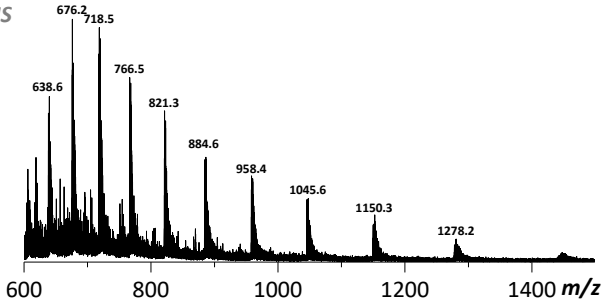


Figure V-35: Characterization of sequence ***S12****.*

Sequence ***S13*** *ESI(-)-MS*

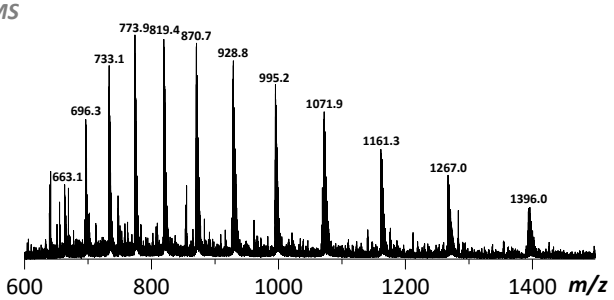


Figure V-36: Characterization of sequence ***S13****.*

Sequence ***S14*** *ESI(-)-MS*

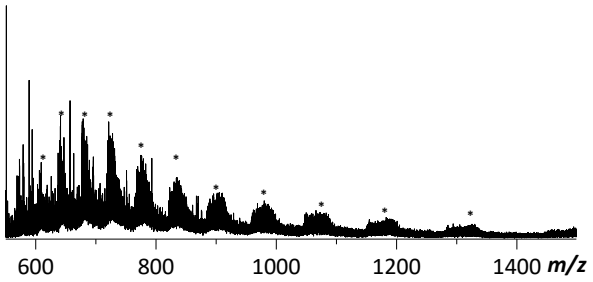


Figure V-37: Characterization of sequence ***S14****.*

5 Analysis of Key Compounds from Chapter III

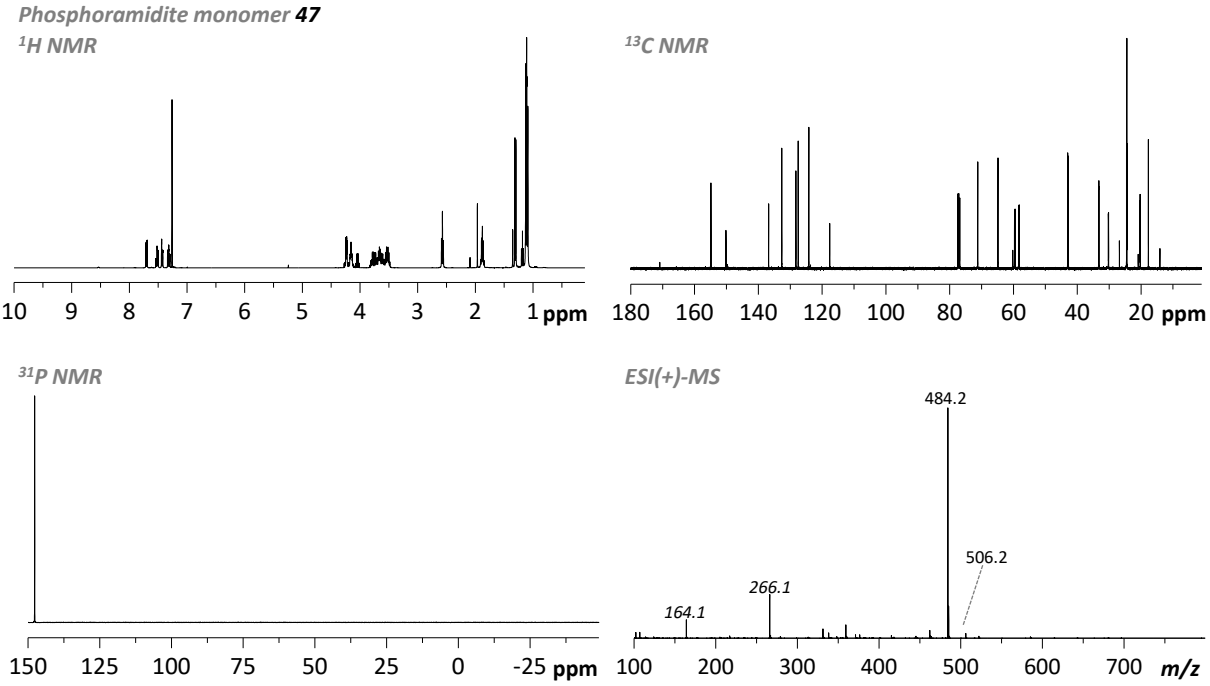


*Figure V-38: Characterization of phosphoramidite monomer **47**.*

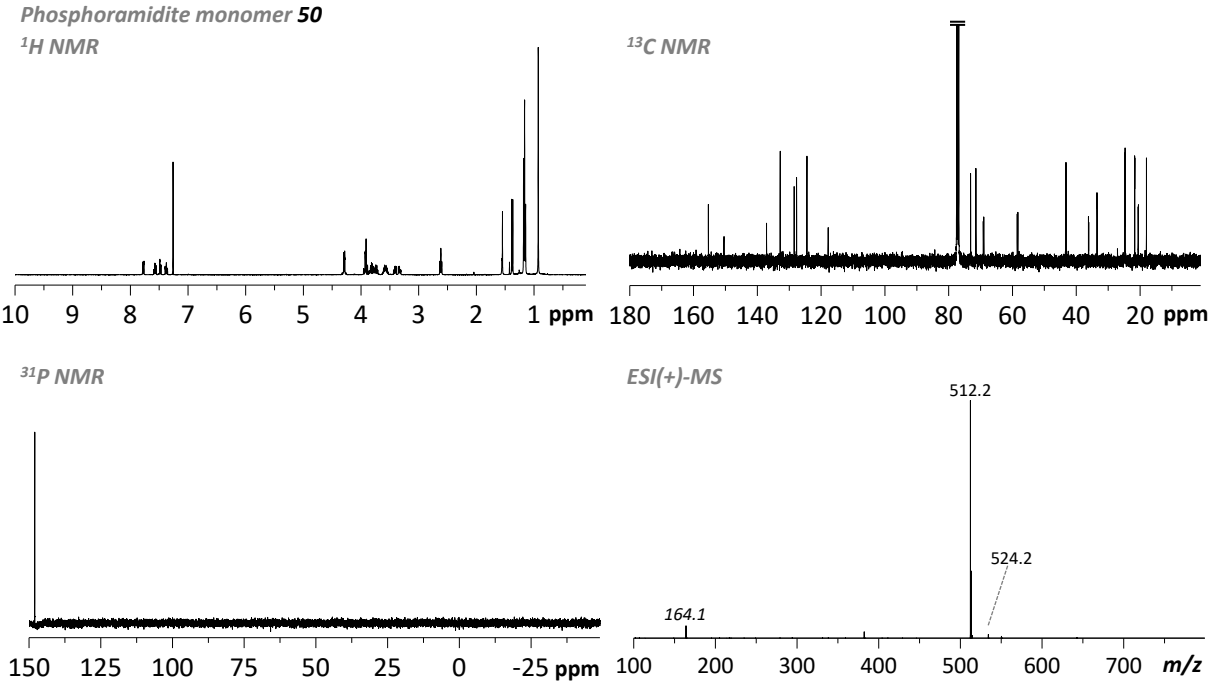


*Figure V-39: Characterization of phosphoramidite monomer **50**.*

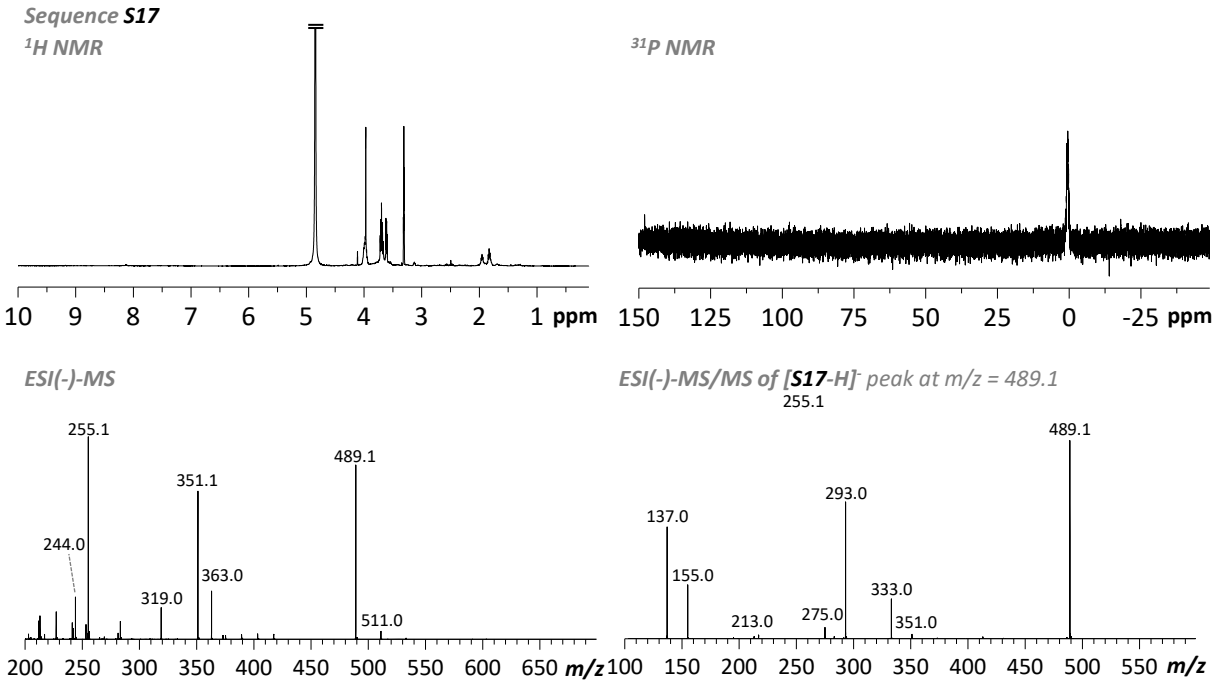


*Figure V-40: Characterization of sequence **S17**.*

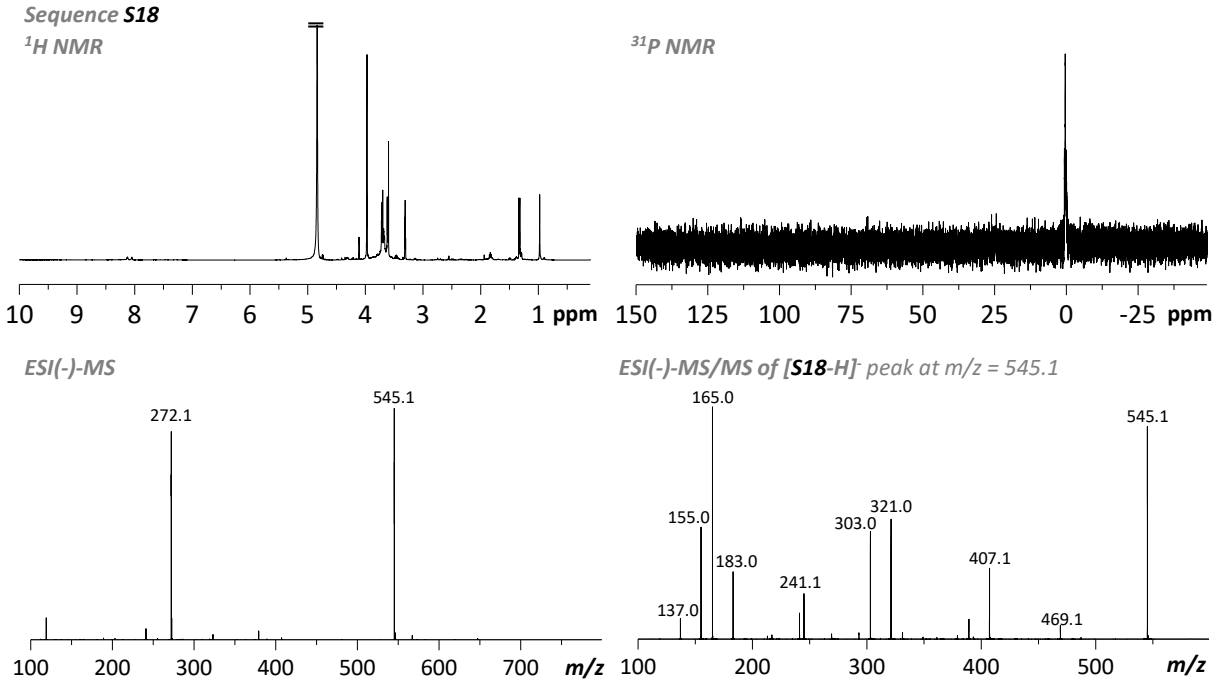


*Figure V-41: Characterization of sequence **S18**.*

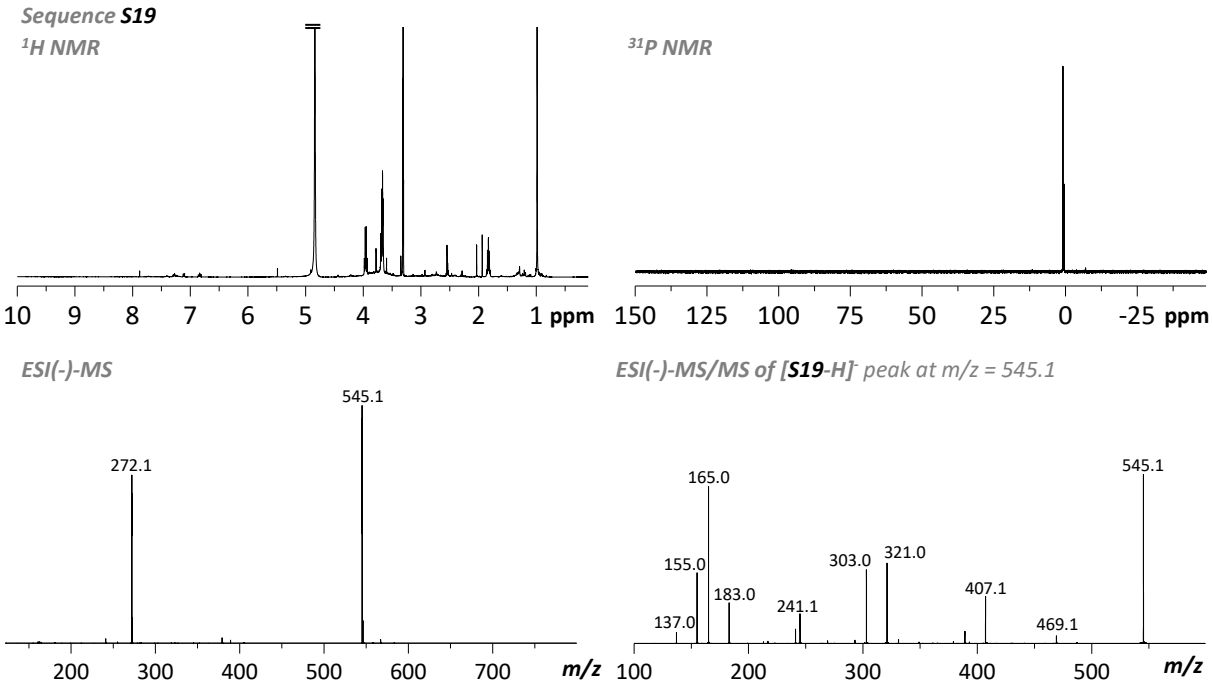


*Figure V-42: Characterization of sequence **S19**.*

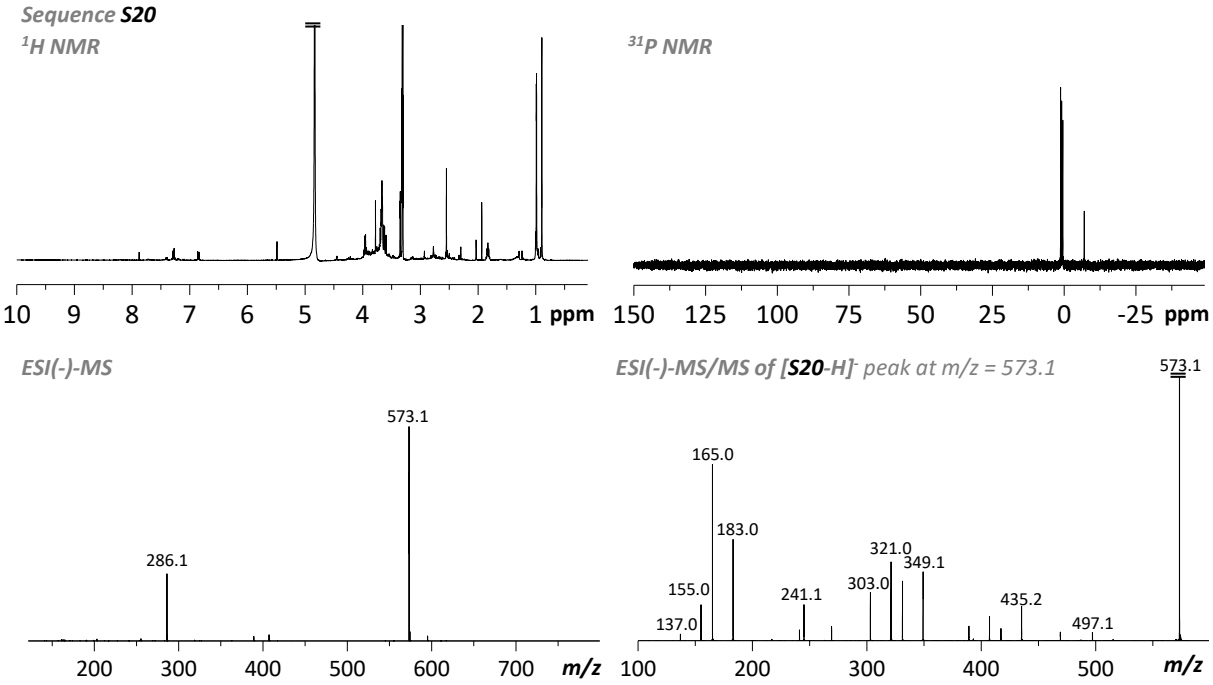
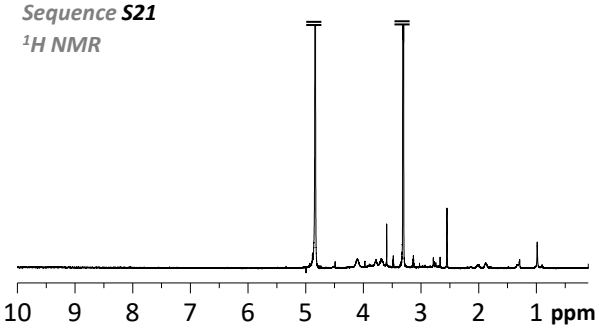
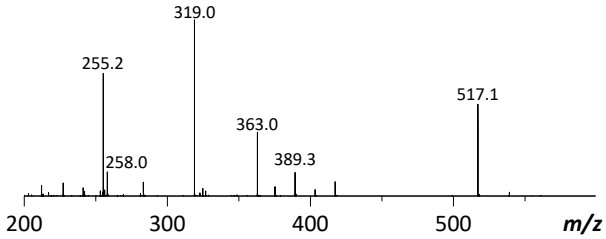


*Figure V-43: Characterization of sequence **S20**.*

Sequence S21
^{1}H NMR



ESI(-)-MS



ESI(-)-MS/MS of [S21-H]$^-$ peak at m/z = 517.1

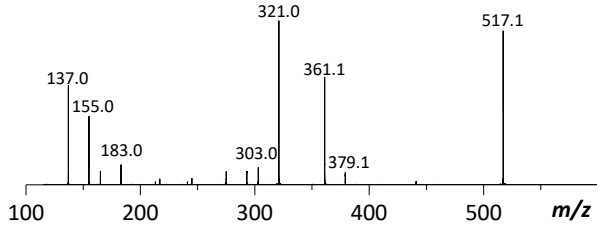
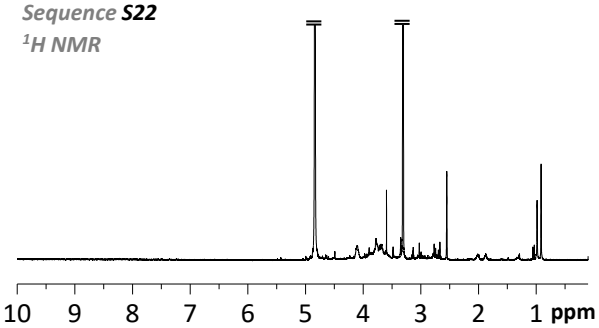
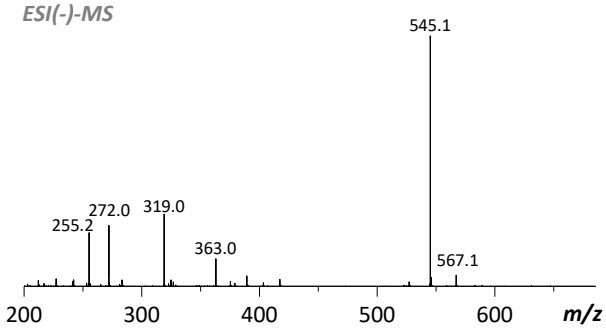


*Figure V-44: Characterization of sequence **S21**.*

Sequence S22
^{1}H NMR



ESI(-)-MS



ESI(-)-MS/MS of [S22-H]$^-$ peak at m/z = 545.1

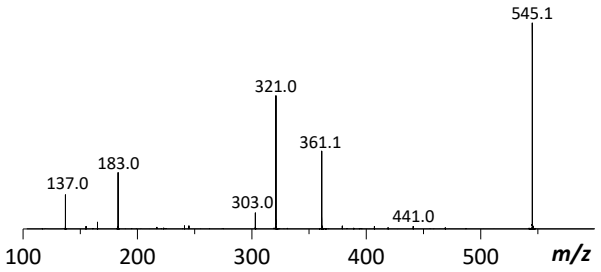


*Figure V-45: Characterization of sequence **S22**.*

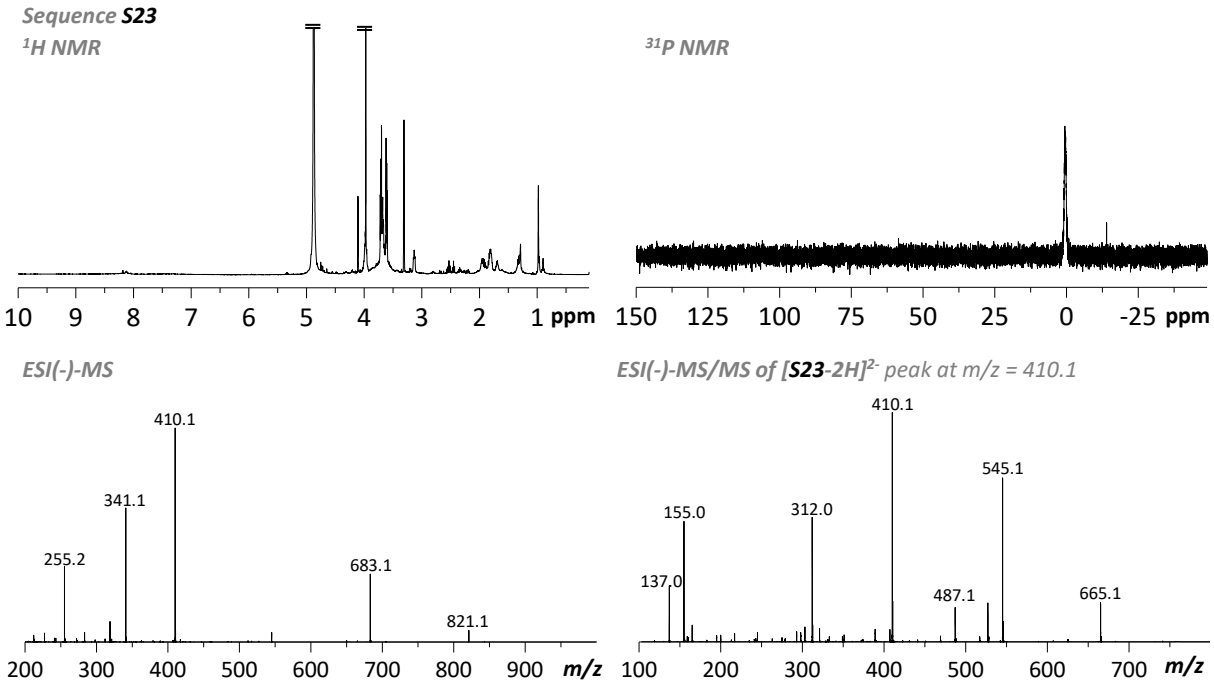


*Figure V-46: Characterization of sequence **S23**.*

6 Analysis of Key Compounds from Chapter IV

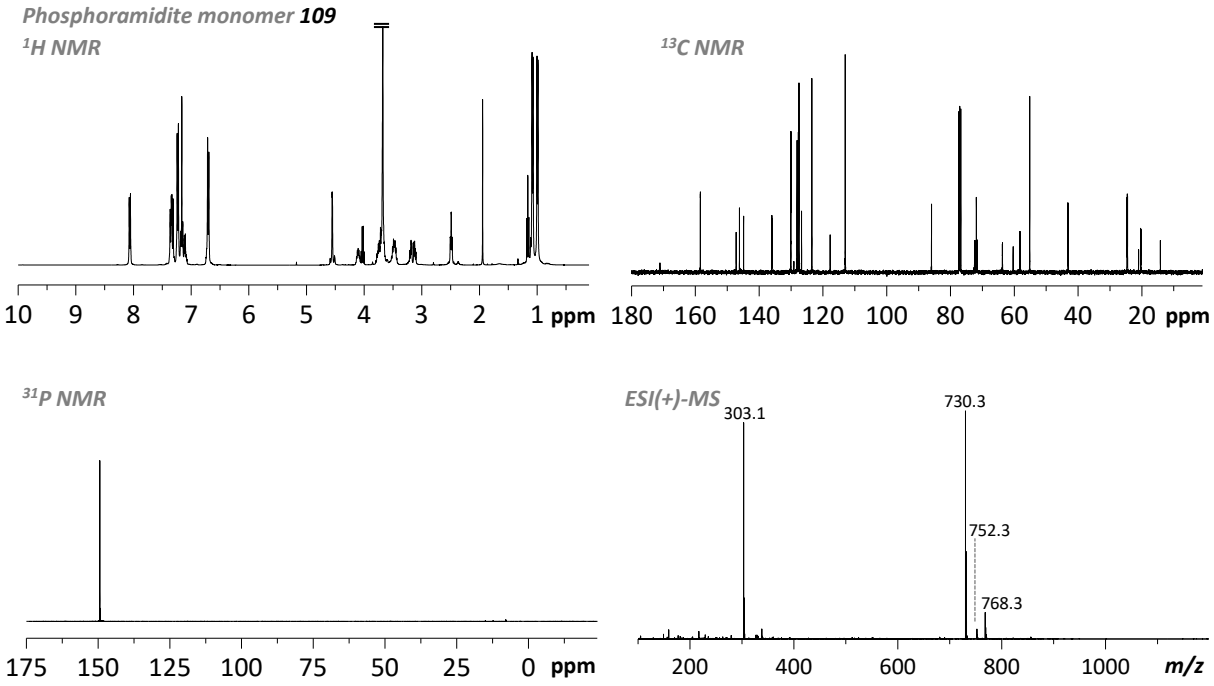


*Figure V-47: Characterization of phosphoramidite monomer **109**.*

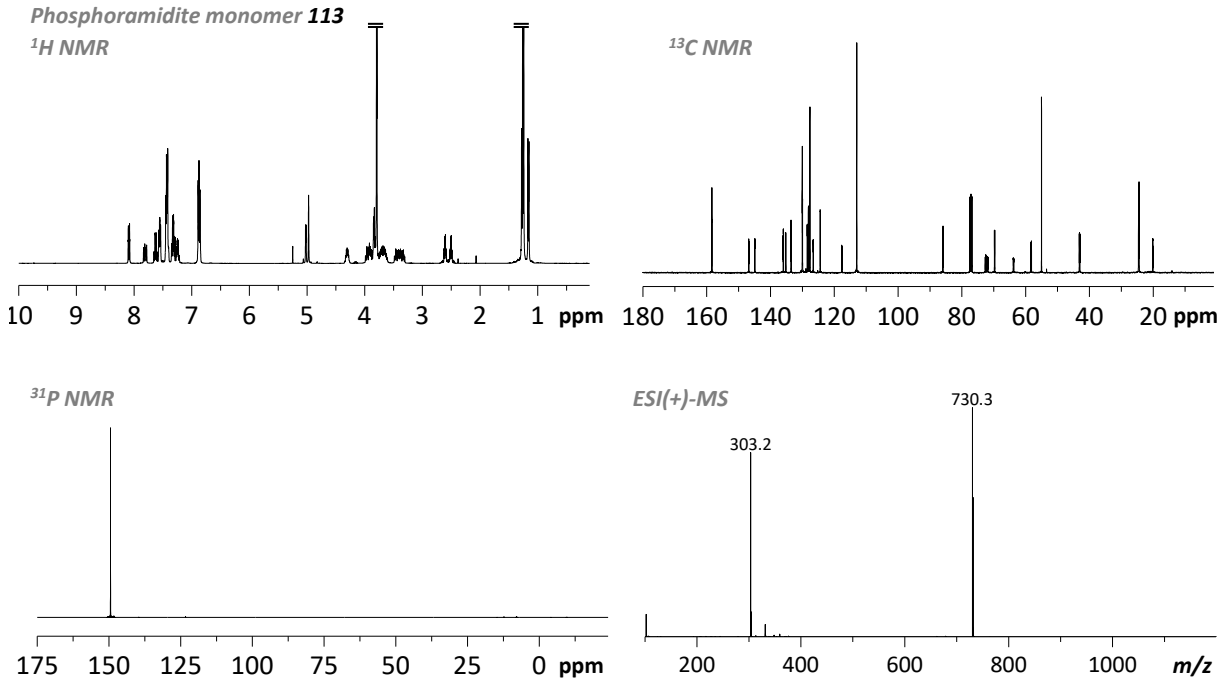


*Figure V-48: Characterization of phosphoramidite monomer **113**.*

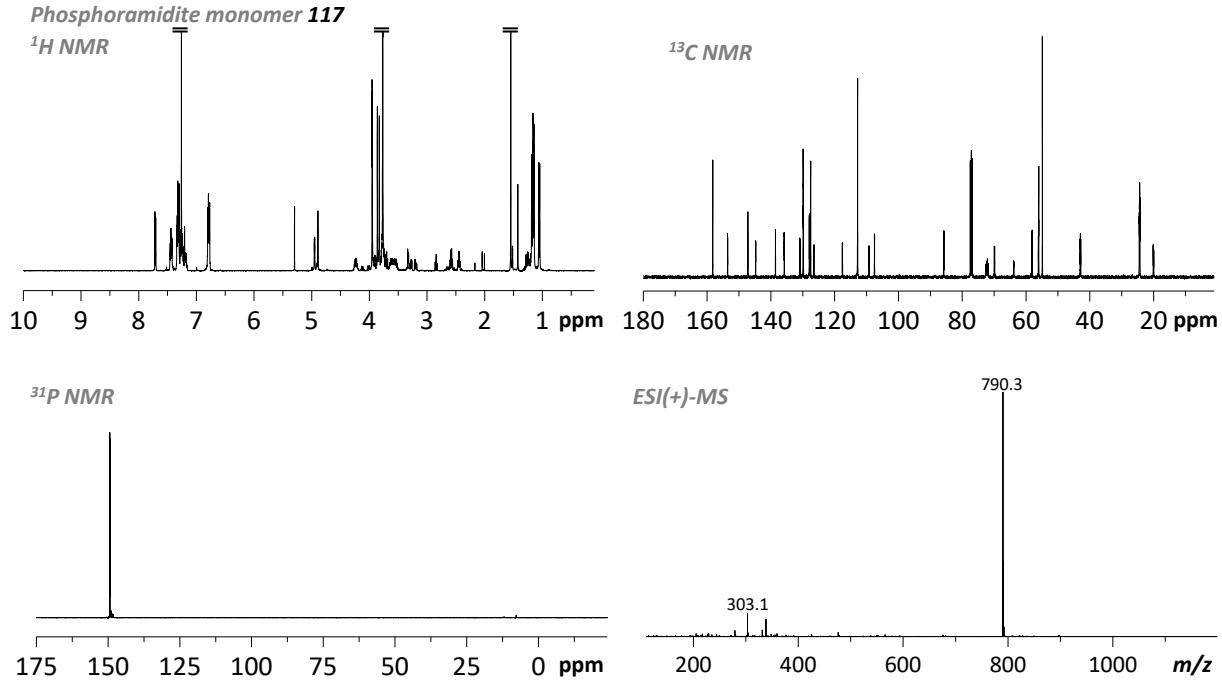


*Figure V-49: Characterization of phosphoramidite monomer **117**.*

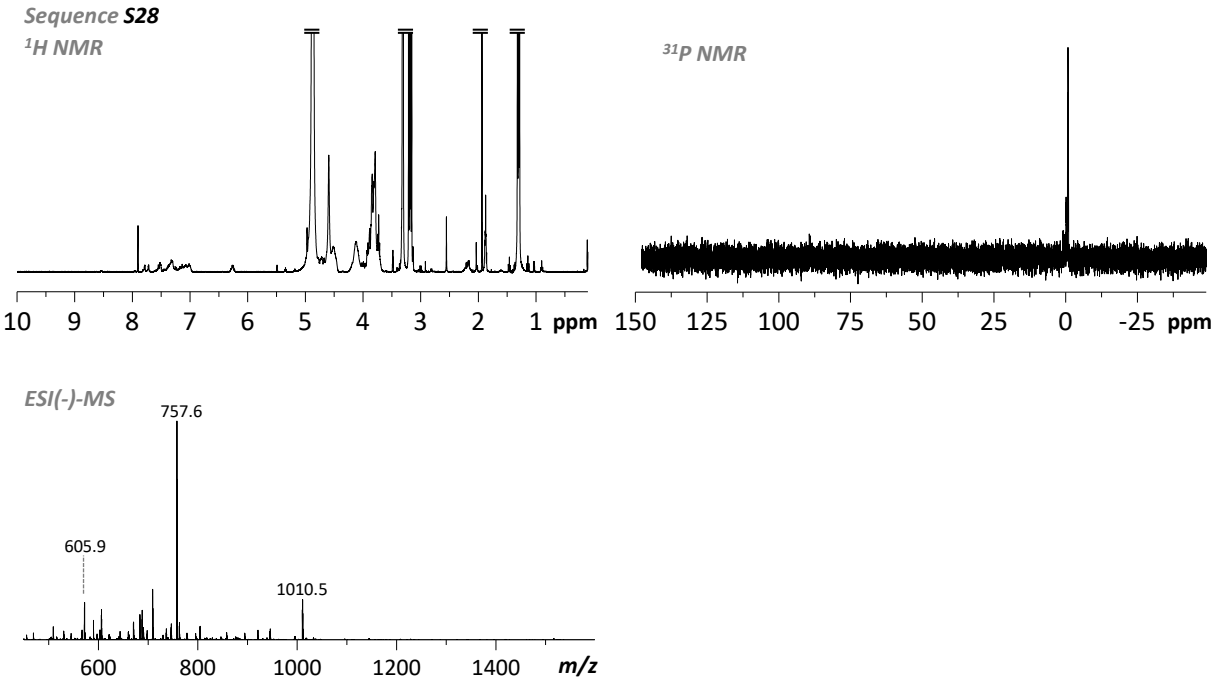


*Figure V-50: Characterization of sequence **S28**.*

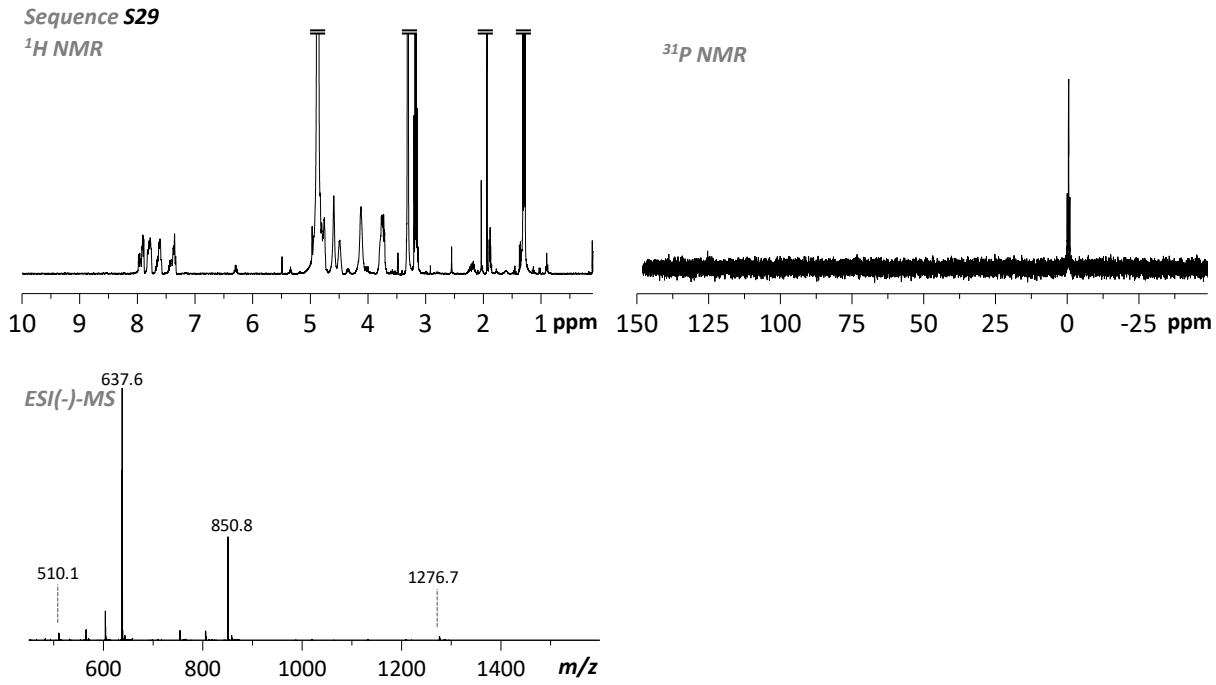


*Figure V-51: Characterization of sequence **S29**.*

Sequence S29'

ESI(-)-MS

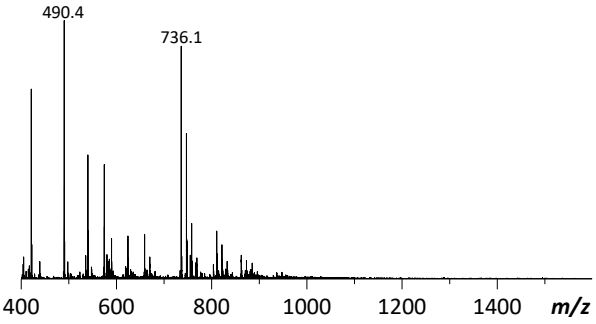
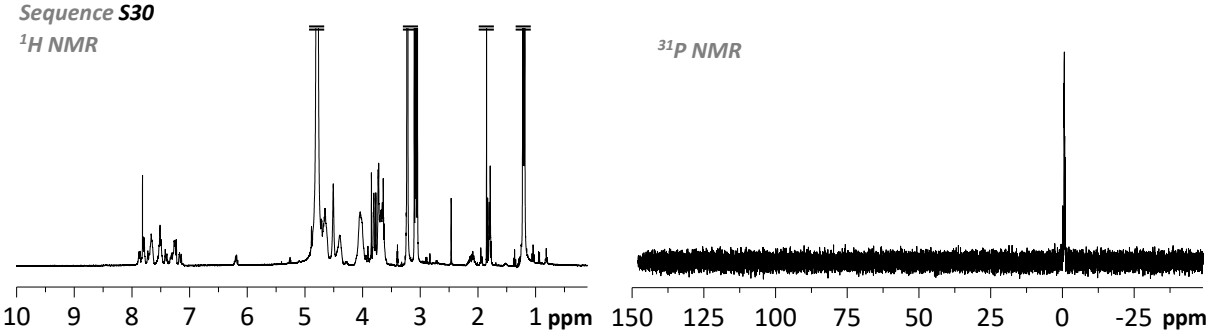


*Figure V-52: Characterization of photomodified sequence **S29'** by ESI(-)-MS.*

Sequence S30

^{1}H NMR

^{31}P NMR



ESI(-)-MS

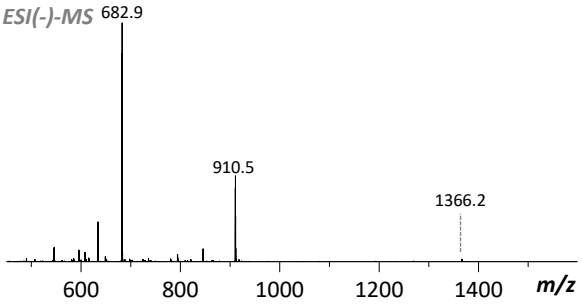


*Figure V-53: Characterization of sequence **S30**.*

Sequence S30'

ESI(-)-MS

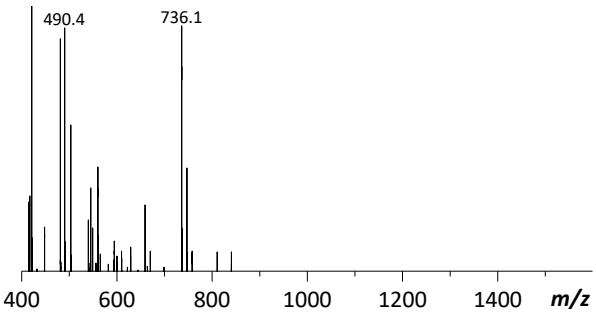


*Figure V-54: Characterization of photomodified sequence **S30'** by ESI(-)-MS.*

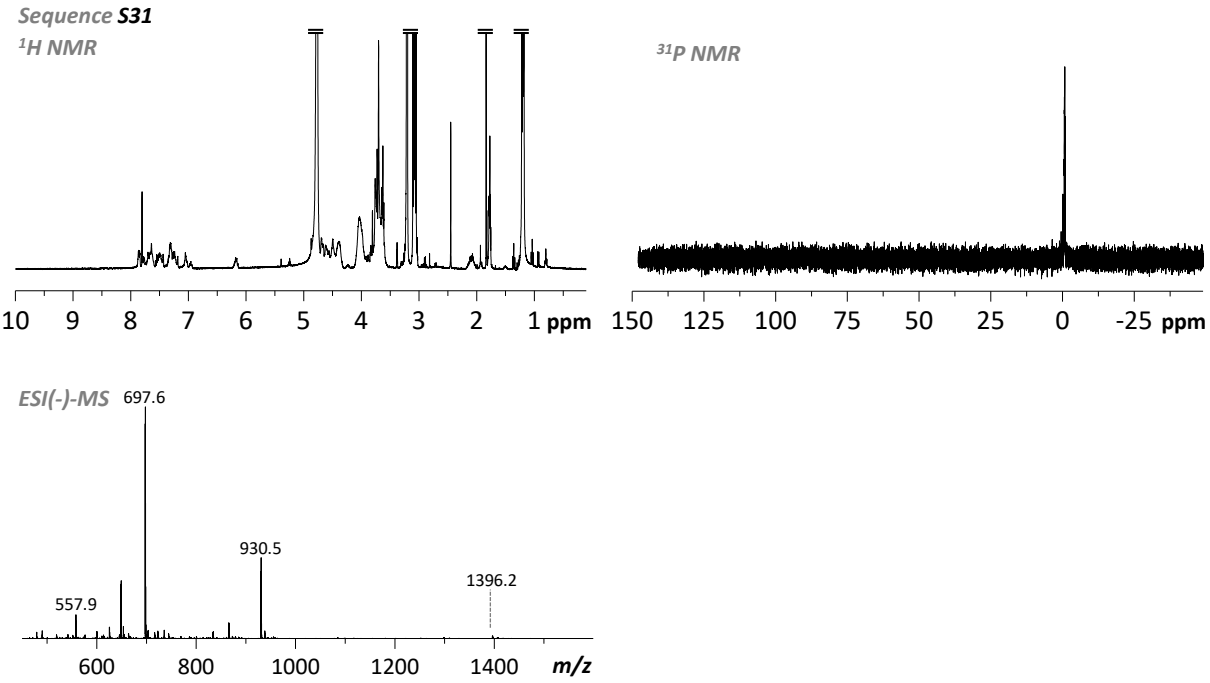


*Figure V-55: Characterization of sequence **S31**.*

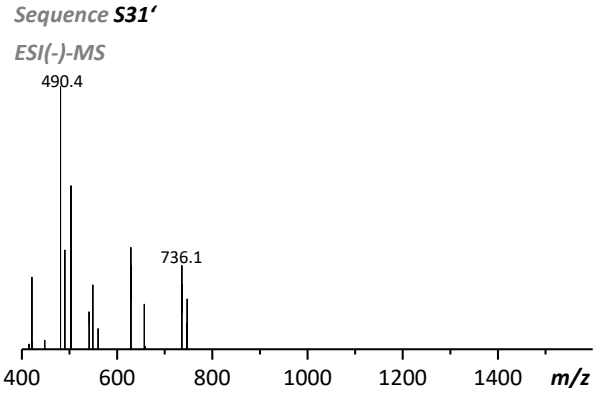
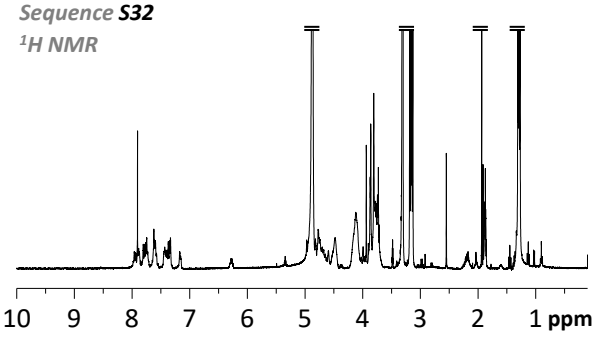


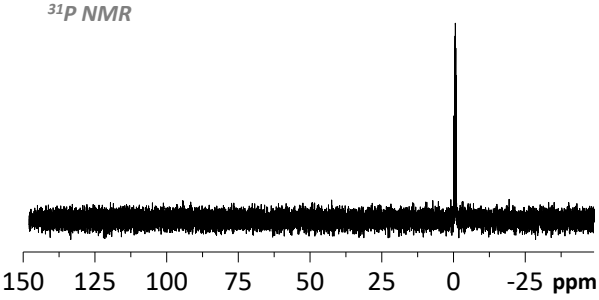
*Figure V-56: Characterization of photomodified sequence **S31'** by ESI(-)-MS.*

*Sequence **S32***

^{1}H NMR



^{31}P NMR



ESI(-)-MS

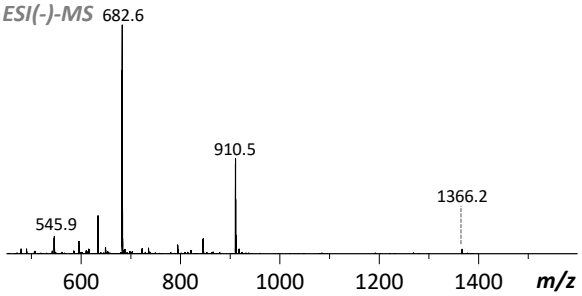


*Figure V-57: Characterization of sequence **S32**.*

*Sequence **S32'***

ESI(-)-MS

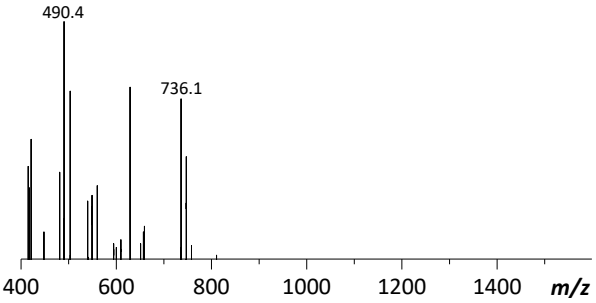


*Figure V-58: Characterization of photomodified sequence **S32'** by ESI(-)-MS.*

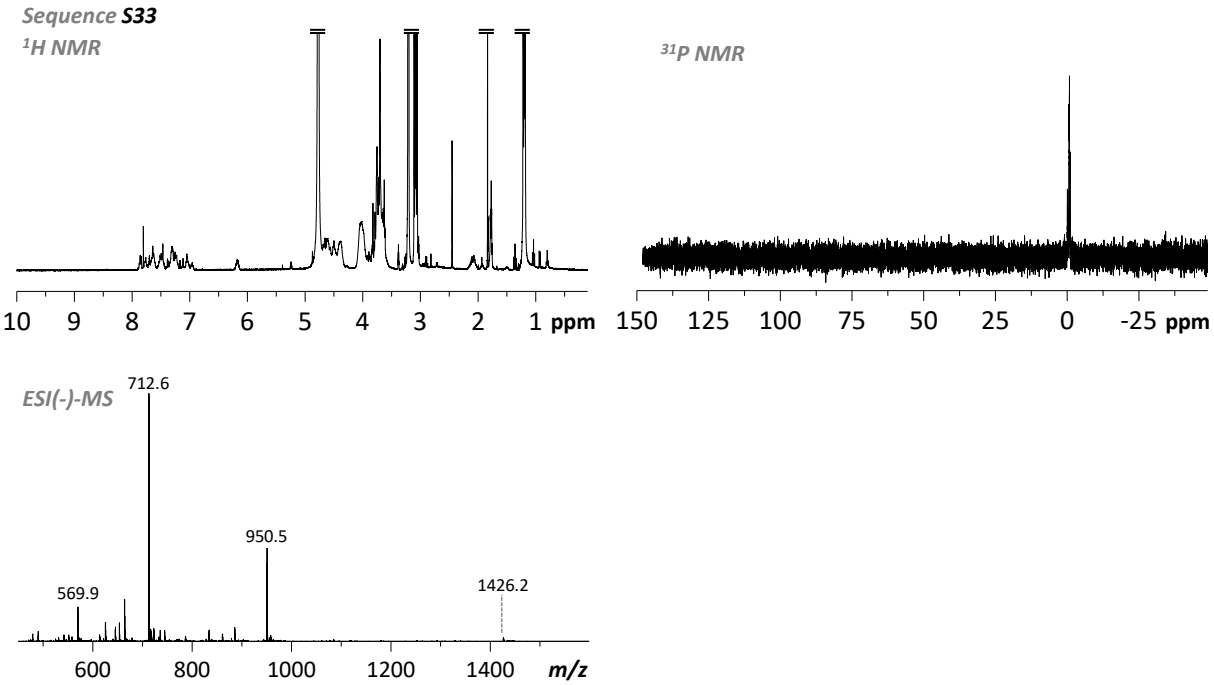


*Figure V-59: Characterization of sequence **S33**.*

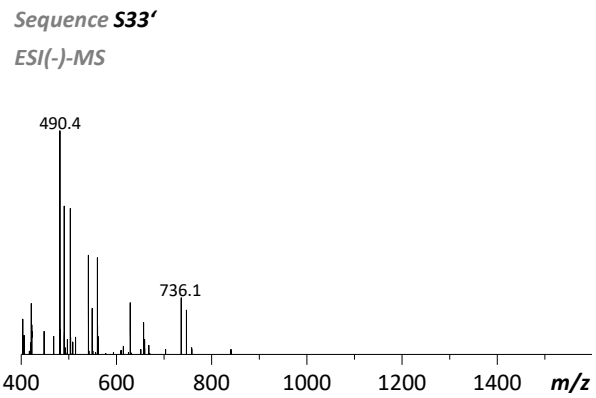
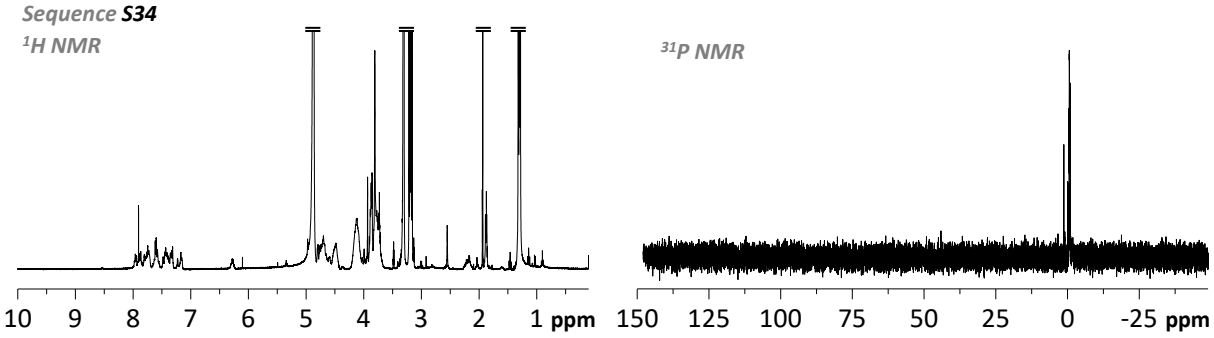


*Figure V-60: Characterization of photomodified sequence **S33'** by ESI(-)-MS.*

*Sequence **S34***



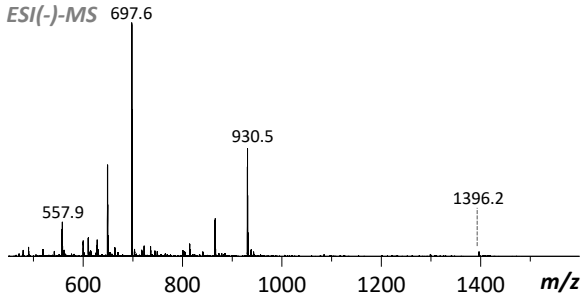


*Figure V-61: Characterization of sequence **S34**.*

*Sequence **S34'***

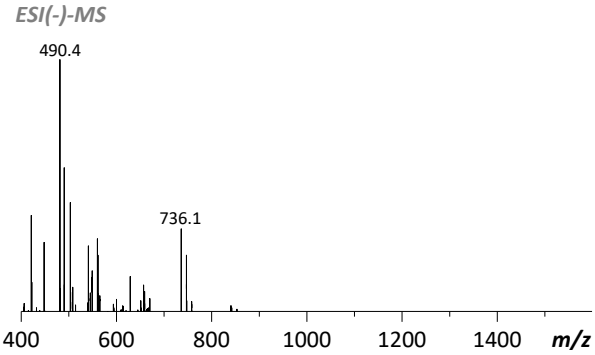
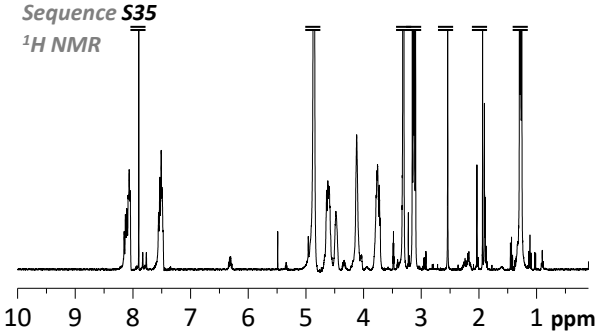


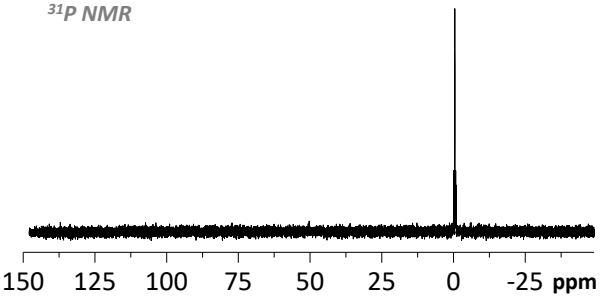
*Figure V-62: Characterization of photomodified sequence **S34'** by ESI(-)-MS.*

*Sequence **S35***

^{1}H NMR



^{31}P NMR



ESI(-)-MS

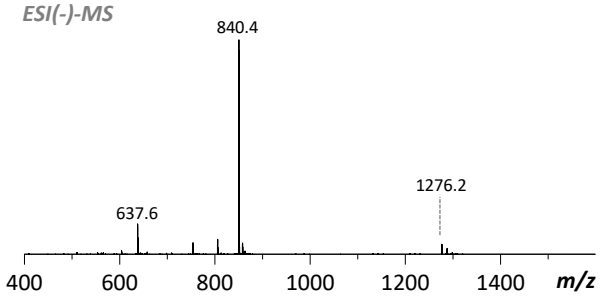


*Figure V-63: Characterization of sequence **S35**.*

*Sequence **S35'***

ESI(-)-MS

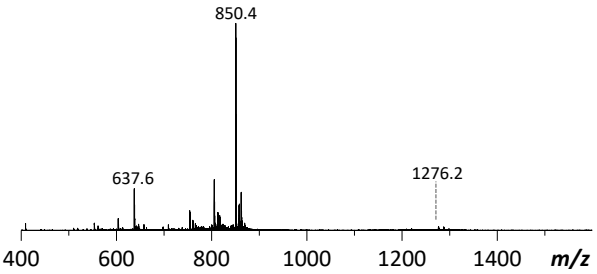
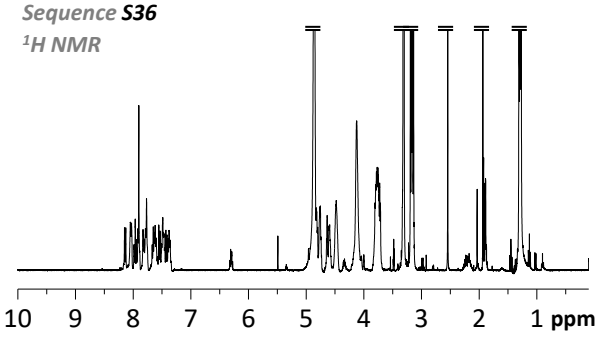


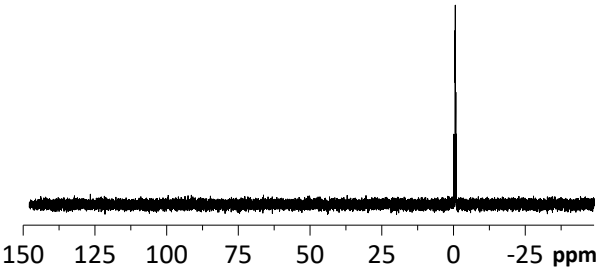
*Figure V-64: Characterization of photomodified sequence **S35'** by ESI(-)-MS.*

*Sequence **S36***

^{1}H NMR



^{31}P NMR



ESI(-)-MS

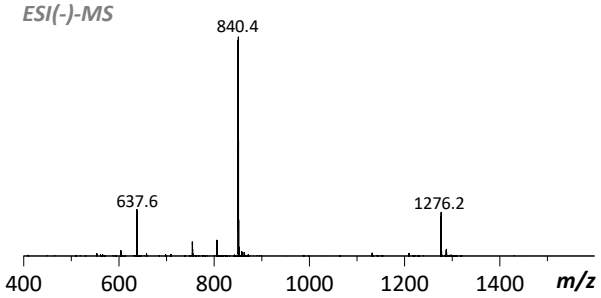


*Figure V-65: Characterization of sequence **S36**.*

*Sequence **S36'***

ESI(-)-MS

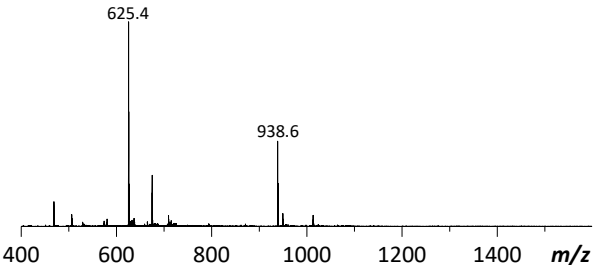
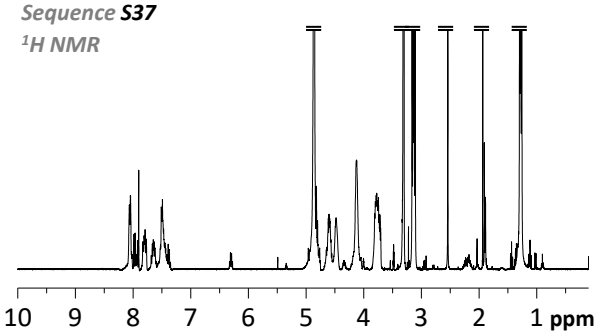


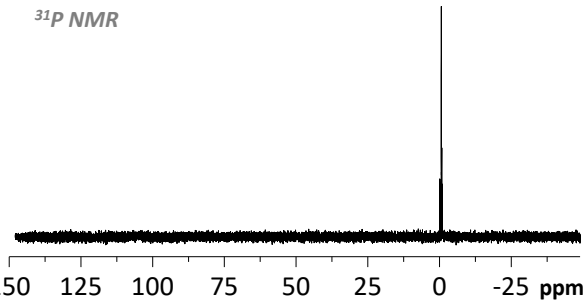
*Figure V-66: Characterization of photomodified sequence **S36'** by ESI(-)-MS.*

*Sequence **S37***

^{1}H NMR



^{31}P NMR



ESI(-)-MS

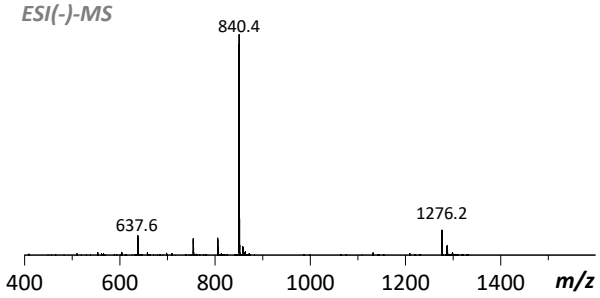


*Figure V-67: Characterization of sequence **S37**.*

*Sequence **S37'***

ESI(-)-MS

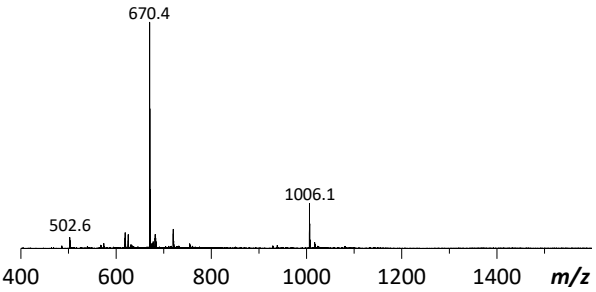
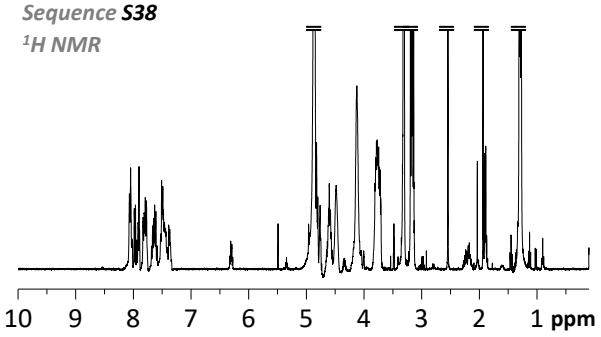


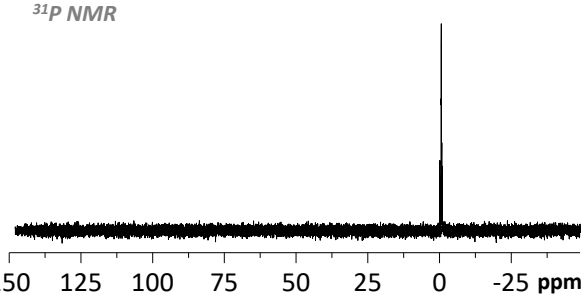
*Figure V-68: Characterization of photomodified sequence **S37'** by ESI(-)-MS.*

*Sequence **S38***

^{1}H NMR



^{31}P NMR



ESI(-)-MS

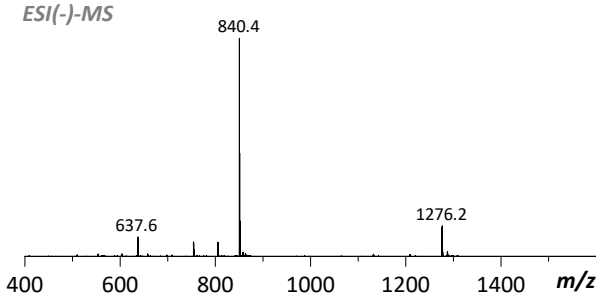


*Figure V-69: Characterization of sequence **S38**.*

*Sequence **S38'***

ESI(-)-MS

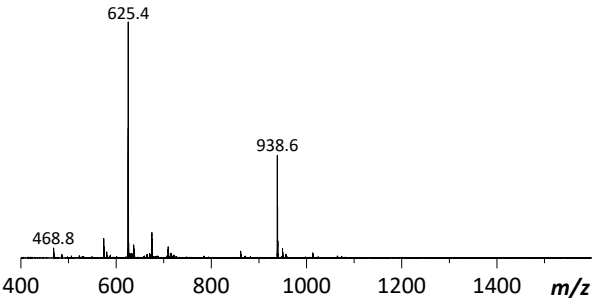
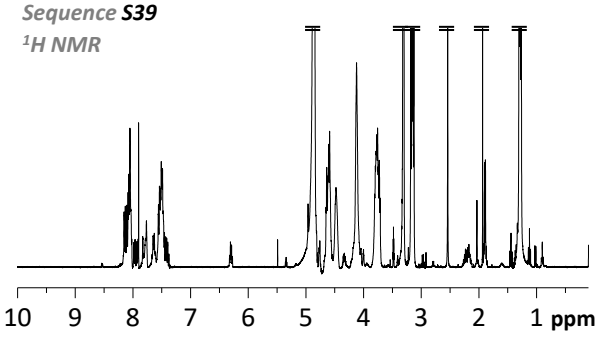


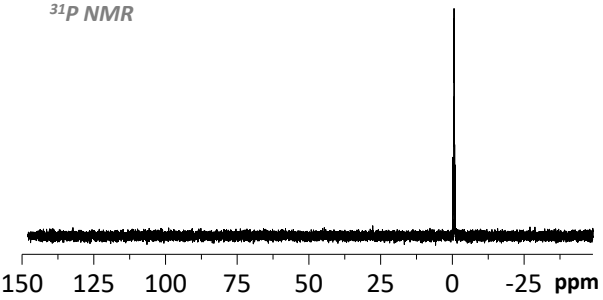
*Figure V-70: Characterization of photomodified sequence **S38'** by ESI(-)-MS.*

*Sequence **S39***

^{1}H NMR



^{31}P NMR



ESI(-)-MS

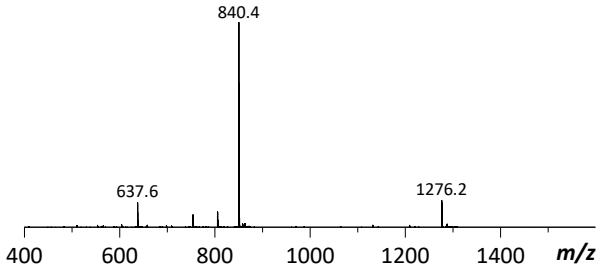


*Figure V-71: Characterization of sequence **S39**.*

*Sequence **S39'***

ESI(-)-MS

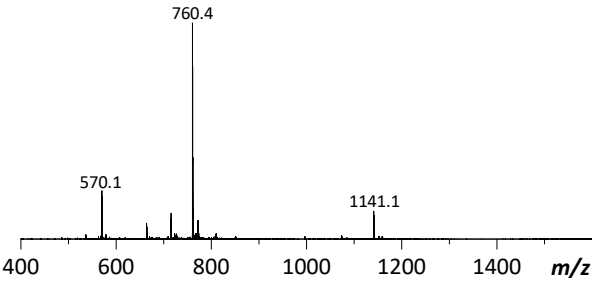
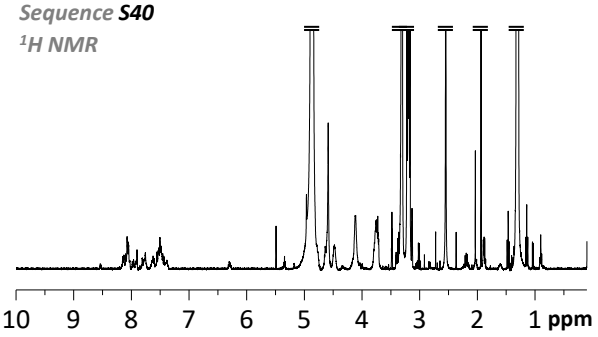


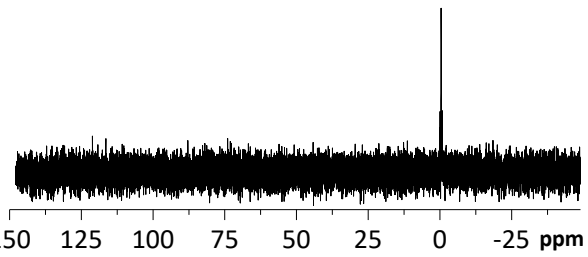
*Figure V-72: Characterization of photomodified sequence **S39'** by ESI(-)-MS.*

*Sequence **S40***

^{1}H NMR



^{31}P NMR



ESI(-)-MS

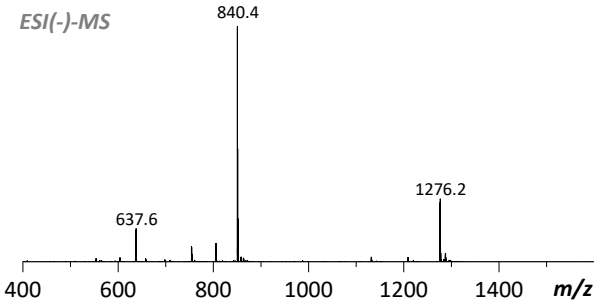


*Figure V-73: Characterization of sequence **S40**.*

*Sequence **S40'***

ESI(-)-MS

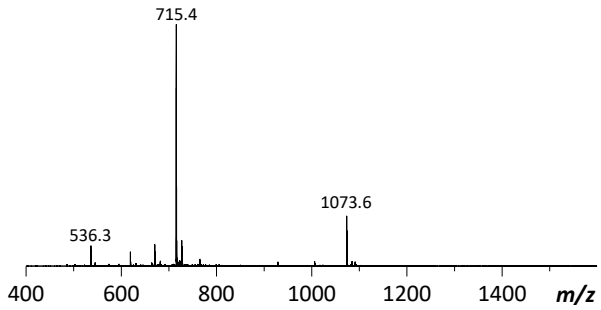


*Figure V-74: Characterization of photomodified sequence **S40'** by ESI(-)-MS.*

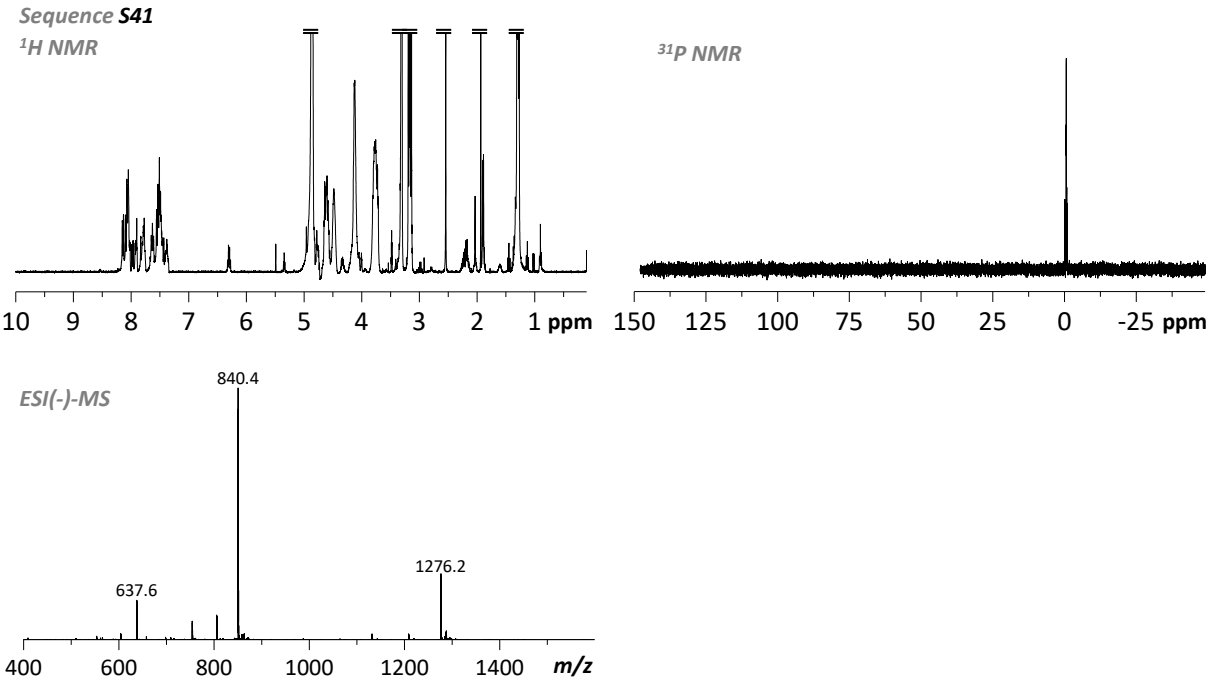


*Figure V-75: Characterization of sequence **S41**.*

*Sequence **S41'***

ESI(-)-MS

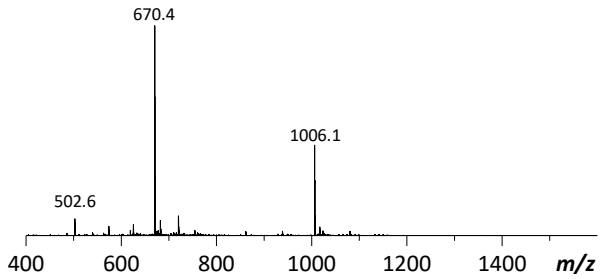


*Figure V-76: Characterization of photomodified sequence **S41'** by ESI(-)-MS.*

Literature

[1] J.-F. Lutz, J.-M. Lehn, E. W. Meijer, K. Matyjaszewski, *Nat. Rev. Mat.* **2016**, *1*, 16024.
[2] M. Zamfir, J.-F. Lutz, *Nat. Commun.* **2012**, *3*, 1138.
[3] J.-F. Lutz, *Macromol. Rapid Commun.* **2017**, 1700582.
[4] S. C. Solleder, R. V. Schneider, K. S. Wetzel, A. C. Boukis, M. A. Meier, *Macromol. Rapid Commun.* **2017**, *38*, 1600711.
[5] S. L. Beaucage, M. H. Caruthers, *Tetrahedron Lett.* **1981**, *22*, 1859-1862.
[6] A. Al Ouahabi, M. Kotera, L. Charles, J.-F. Lutz, *ACS Macro Lett.* **2015**, *4*, 1077-1080.
[7] J.-F. Lutz, *Macromolecules* **2015**, *48*, 4759-4767.
[8] L. Organick, S. D. Ang, Y.-J. Chen, R. Lopez, S. Yekhanin, K. Makarychev, M. Z. Racz, G. Kamath, P. Gopalan, B. Nguyen, C. N. Takahashi, S. Newman, H.-Y. Parker, C. Rashtchian, K. Stewart, G. Gupta, R. Carlson, J. Mulligan, D. Carmean, G. Seelig, L. Ceze, K. Strauss, *Nat. Biotechnol.* **2018**, *36*, 242.
[9] G. M. Church, Y. Gao, S. Kosuri, *Science* **2012**, 1226355.
[10] A. Al Ouahabi, L. Charles, J.-F. Lutz, *J. Am. Chem. Soc.* **2015**, *137*, 5629-5635.
[11] A. Al Ouahabi, J. A. Amalian, L. Charles, J.-F. Lutz, *Nat. Commun.* **2017**, *8*, 967.
[12] G. Cavallo, A. Al Ouahabi, L. Oswald, L. Charles, J.-F. Lutz, *J. Am. Chem. Soc.* **2016**, *138*, 9417-9420.
[13] D. Branton, D. W. Deamer, A. Marziali, H. Bayley, S. A. Benner, T. Butler, M. Di Ventra, S. Garaj, A. Hibbs, X. Huang, S. B. Jovanovich, P. S. Krstic, S. Lindsay, X. S. Ling, C. H. Mastrangelo, A. Meller, J. S. Oliver, Y. V. Pershin, J. M. Ramsey, R. Riehn, G. V. Soni, V. Tabard-Cossa, M. Wanunu, M. Wiggin, J. A. Schloss, *Nat. Biotechnol.* **2008**, *26*, 1146-1153.
[14] C. Agbavwe, C. Kim, D. Hong, K. Heinrich, T. Wang, M. M. Somoza, *J. Nanobiotechnol.* **2011**, *9*, 57.
[15] C. Mayer, G. R. McInroy, P. Murat, P. Van Delft, S. Balasubramanian, *Angew. Chem. Int. Ed.* **2016**, *55*, 11144-11148.
[16] P. Klán, T. Šolomek, C. G. Bochet, A. Blanc, R. Givens, M. Rubina, V. Popik, A. Kostikov, J. Wirz, *Chem. Rev.* **2013**, *113*, 119-191.
[17] F. Crick, *Nature* **1970**, *227*, 561.563.
[18] R. R. Crichton, in *Biological Inorganic Chemistry*, 2nd ed., Elsevier, Oxford, **2012**, pp. 35-68.
[19] D. L. Nelson, A. L. Lehninger, M. M. Cox, in *Lehninger Principles of Biochemistry*, 5th ed., W. H. Freeman and Company, New York, **2008**, pp. 1065-1113.
[20] E. Fischer, *Ber. Dtsch. Chem. Ges.* **1903**, *36*, 2982-2992.
[21] E. Fischer, *Ber. Dtsch. Chem. Ges.* **1904**, *37*, 2486-2511.
[22] E. Fischer, *Ber. Dtsch. Chem. Ges.* **1907**, *40*, 1754-1767.
[23] M. Bergmann, L. Zervas, *Ber. Dtsch. Chem. Ges.* **1932**, *65*, 1192-1201.
[24] J. Meienhofer, E. Schnabel, H. Bremer, O. Brinkhoff, R. Zabel, W. Sroka, H. Klostermeyer, D. Brandenburg, T. Okuda, H. Zahn, *Z. Naturforsch., B: Anorg. Chem., Org. Chem., Biochem., Biophys., Biol.* **1963**, *18*, 1120-1121.
[25] V. du Vigneaud, C. Ressler, J. M. Swan, C. W. Roberts, P. G. Katsoyannis, *J. Am. Chem. Soc.* **1954**, *76*, 3115-3121.
[26] P. G. Katsoyannis, A. M. Tometsko, C. Zalut, K. Fukuda, *J. Am. Chem. Soc.* **1966**, *88*, 5625-5635.
[27] R. B. Merrifield, D. W. Woolley, *J. Am. Chem. Soc.* **1956**, *78*, 4646-4649.
[28] H. Schwarz, F. M. Bumpus, I. H. Page, *J. Am. Chem. Soc.* **1957**, *79*, 5697-5703.
[29] L. M. Gierasch, *Pept. Sci.* **2006**, *84*, 433-434.
[30] R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149-2154.
[31] R. B. Merrifield, J. M. Stewart, N. Jernberg, *Anal. Chem.* **1966**, *38*, 1905-1914.
[32] B. Gutte, R. B. Merrifield, *J. Biol. Chem.* **1971**, *246*, 1922-1941.
[33] B. L. Nilsson, M. B. Soellner, R. T. Raines, *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 91-118.

[34] M. Schnölzer, P. Alewood, A. Jones, D. Alewood, S. B. H. Kent, *Int. J. Pept. Res. Ther.* **2007**, *13*, 31-44.
[35] D. A. Wellings, E. Atherton, *Methods Enzymol.* **1997**, *289*, 44-67.
[36] R. Knorr, A. Trzeciak, W. Bannwarth, D. Gillessen, *Tetrahedron Lett.* **1989**, *30*, 1927-1930.
[37] L. A. Carpino, *J. Am. Chem. Soc.* **1993**, *115*, 4397-4398.
[38] L. P. Miranda, P. F. Alewood, *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 1181-1186.
[39] F. Albericio, L. A. Carpino, *Methods Enzymol.* **1997**, *289*, 104-126.
[40] M. Meldal, *Methods Enzymol.* **1997**, *289*, 83-104.
[41] M. F. Songster, G. Barany, *Methods Enzymol.* **1997**, *289*, 126-174.
[42] M. Amblard, J.-A. Fehrentz, J. Martinez, G. Subra, in *Peptide Synthesis and Applications, Vol. 298* (Ed.: J. Howl), Humana Press, Totowa, NJ, USA, **2005**, pp. 3-24.
[43] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776-779.
[44] W. Theodor, B. Ekkehart, B. Lieselotte, L. H. Ulrich, L. Hans, *Justus Liebigs Ann. Chem.* **1953**, *583*, 129-149.
[45] L. R. Malins, R. J. Payne, *Curr. Opin. Chem. Biol.* **2014**, *22*, 70-78.
[46] B. Lewandowski, G. De Bo, J. W. Ward, M. Papmeyer, S. Kuschel, M. J. Aldegunde, P. M. E. Gramlich, D. Heckmann, S. M. Goldup, D. M. D'Souza, A. E. Fernandes, D. A. Leigh, *Science* **2013**, *339*, 189-193.
[47] G. De Bo, S. Kuschel, D. A. Leigh, B. Lewandowski, M. Papmeyer, J. W. Ward, *J. Am. Chem. Soc.* **2014**, *136*, 5811-5814.
[48] G. De Bo, M. A. Y. Gall, M. O. Kitching, S. Kuschel, D. A. Leigh, D. J. Tetlow, J. W. Ward, *J. Am. Chem. Soc.* **2017**, *139*, 10875-10879.
[49] G. De Bo, M. A. Y. Gall, S. Kuschel, J. De Winter, P. Gerbaux, D. A. Leigh, *Nat. Nanotechnol.* **2018**, *13*, 381-385.
[50] D. L. Nelson, A. L. Lehninger, M. M. Cox, in *Lehninger Principles of Biochemistry*, 5th ed., W. H. Freeman and Company, New York, **2008**, pp. 235-270.
[51] P. Stanley, H. Schachter, N. Taniguchi, in *Essentials of Glycobiology [Internet]*, 3rd ed. (Eds.: A. Varki, R. D. Cummings, J. D. Esko), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, **2017**.
[52] J. D. C. Codée, A. Ali, H. S. Overkleeft, G. A. van der Marel, *C. R. Chim.* **2011**, *14*, 178-193.
[53] R. R. Schmidt, K.-H. Jung, in *Carbohydrates in Chemistry and Biology* (Eds.: B. Ernst, G. W. Hart, P. Sinaý), Wiley-VCH Verlag GmbH, Weinheim, Germany, **2008**, pp. 5-59.
[54] P. H. Seeberger, H. S. Overkleeft, in *Essentials of Glycobiology [Internet]*, 3rd ed. (Eds.: A. Varki, R. D. Cummings, J. D. Esko, e. al), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, **2017**.
[55] T. G. Mayer, B. Kratzer, R. R. Schmidt, *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 2177-2181.
[56] J. Seifert, M. Lergenmüller, Y. Ito, *Angew. Chem. Int. Ed.* **2000**, *39*, 531-534.
[57] Z.-G. Wang, X. Zhang, M. Visser, D. Live, A. Zatorski, U. Iserloh, K. O. Lloyd, S. J. Danishefsky, *Angew. Chem. Int. Ed.* **2001**, *40*, 1728-1732.
[58] Z. Zhang, I. R. Ollmann, X.-S. Ye, R. Wischnat, T. Baasov, C.-H. Wong, *J. Am. Chem. Soc.* **1999**, *121*, 734-753.
[59] B. Yang, K. Yoshida, X. Huang, in *Glycochemical Synthesis* (Eds.: S.-C. Hung, M. M. L. Zulueta), John Wiley & Sons, Inc., Hoboken, NJ, USA, **2016**, pp. 155-187.
[60] C.-Y. Wu, C.-H. Wong, in *Reactivity Tuning in Oligosaccharide Assembly. Topics in Current Chemistry, Vol. 301* (Eds.: B. Fraser-Reid, J. Cristóbal López), Springer Berlin Heidelberg, Berlin, Heidelberg, **2011**, pp. 223-252.
[61] O. J. Plante, E. R. Palmacci, P. H. Seeberger, *Science* **2001**, *291*, 1523-1527.
[62] P. Sears, C.-H. Wong, *Science* **2001**, *291*, 2344-2350.
[63] P. H. Seeberger, *Chem. Soc. Rev.* **2008**, *37*, 19-28.
[64] C.-H. Hsu, S.-C. Hung, C.-Y. Wu, C.-H. Wong, *Angew. Chem. Int. Ed.* **2011**, *50*, 11872-11923.
[65] D. B. Werz, R. Ranzinger, S. Herget, A. Adibekian, C.-W. von der Lieth, P. H. Seeberger, *ACS Chem. Biol.* **2007**, *2*, 685-691.
[66] P. H. Seeberger, *Acc. Chem. Res.* **2015**, *48*, 1450-1463.

[67] O. T. Avery, C. M. MacLeod, M. McCarty, *J. Exp. Med.* **1944**, *79*, 137-158.
[68] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 964-967.
[69] R. E. Franklin, R. G. Gosling, *Nature* **1953**, *171*, 740-741.
[70] M. H. F. Wilkins, A. R. Stokes, H. R. Wilson, *Nature* **1953**, *171*, 738-740.
[71] D. L. Nelson, A. L. Lehninger, M. M. Cox, in *Lehninger Principles of Biochemistry*, 5th ed., W. H. Freeman and Company, New York, **2008**, pp. 1021-1064.
[72] M. Meselson, F. W. Stahl, *Proc. Natl. Acad. Sci. U. S. A.* **1958**, *44*, 671-682.
[73] D. L. Nelson, A. L. Lehninger, M. M. Cox, in *Lehninger Principles of Biochemistry*, 5th ed., W. H. Freeman and Company, New York, **2008**, pp. 975-1020.
[74] N. Badi, J.-F. Lutz, *Chem. Soc. Rev.* **2009**, *38*, 3383-3390.
[75] J.-F. Lutz, *Polym. Chem.* **2010**, *1*, 55-62.
[76] T. T. Trinh, C. Laure, J.-F. Lutz, *Macromol. Chem. Phys.* **2015**, *216*, 1498-1506.
[77] J.-F. Lutz, in *Sequence-Controlled Polymers: Synthesis, Self-Assembly, and Properties, Vol. 1170* (Eds.: J.-F. Lutz, M. Sawamoto), American Chemical Society, Washington, DC, **2014**, pp. 1-11.
[78] J.-F. Lutz, *ACS Macro Lett.* **2014**, *3*, 1020-1023.
[79] J.-F. Lutz, M. Ouchi, D. R. Liu, M. Sawamoto, *Science* **2013**, *341*, 1238149.
[80] N. Badi, D. Chan-Seng, J. F. Lutz, *Macromol. Chem. Phys.* **2013**, *214*, 135-142.
[81] C. Barner-Kowollik, J.-F. Lutz, S. Perrier, *Polym. Chem.* **2012**, *3*, 1677-1679.
[82] R. Jones, *Nat. Nanotechnol.* **2008**, *3*, 699-700.
[83] R. F. Stepto, *Pure Appl. Chem.* **2009**, *81*, 351-353.
[84] D. Crespy, M. Bozonnet, M. Meier, *Angew. Chem. Int. Ed.* **2008**, *47*, 3322-3328.
[85] L. H. Baekeland, U.S. Patent 942,699, **1907**.
[86] W. H. Carothers, J. W. Hill, *J. Am. Chem. Soc.* **1932**, *54*, 1559-1566.
[87] W. H. Carothers, U.S. Patent 2,130,523, **1935**.
[88] O. Bayer, *Angew. Chem.* **1947**, *59*, 257-272.
[89] W. H. Carothers, *Trans. Faraday Soc.* **1936**, *32*, 39-49.
[90] J. C. Sworen, J. A. Smith, J. M. Berg, K. B. Wagener, *J. Am. Chem. Soc.* **2004**, *126*, 11238-11246.
[91] M. E. Seitz, C. D. Chan, K. L. Opper, T. W. Baughman, K. B. Wagener, K. I. Winey, *J. Am. Chem. Soc.* **2010**, *132*, 8165-8174.
[92] B. S. Aitken, C. F. Buitrago, J. D. Heffley, M. Lee, H. W. Gibson, K. I. Winey, K. B. Wagener, *Macromolecules* **2012**, *45*, 681-687.
[93] R. M. Stayshich, T. Y. Meyer, *J. Am. Chem. Soc.* **2010**, *132*, 10920-10934.
[94] K. Satoh, S. Ozawa, M. Mizutani, K. Nagai, M. Kamigaito, *Nat. Commun.* **2010**, *1*, 6.
[95] M.-A. Berthet, Z. Zarafshani, S. Pfeifer, J.-F. Lutz, *Macromolecules* **2010**, *43*, 44-50.
[96] C. Zhang, Q. Wang, *Macromol. Rapid Commun.* **2011**, *32*, 1180-1184.
[97] C. Zhang, J. Ling, Q. Wang, *Macromolecules* **2011**, *44*, 8739-8743.
[98] O. Türünç, M. A. R. Meier, *J. Polym. Sci., Part A: Polym. Chem.* **2012**, *50*, 1689-1695.
[99] F. Driessen, F. E. Du Prez, P. Espeel, *ACS Macro Lett.* **2015**, *4*, 616-619.
[100] W. Xi, S. Pattanayak, C. Wang, B. Fairbanks, T. Gong, J. Wagner, C. J. Kloxin, C. N. Bowman, *Angew. Chem. Int. Ed.* **2015**, *54*, 14462-14467.
[101] J.-J. Yan, D. Wang, D.-C. Wu, Y.-Z. You, *Chem. Commun.* **2013**, *49*, 6057-6059.
[102] R. W. Lenz, C. E. Handlovits, H. A. Smith, *J. Polym. Sci.* **1962**, *58*, 351-367.
[103] W. Koch, W. Risse, W. Heitz, *Makromol. Chem.* **1985**, *12*, 105-123.
[104] T. Yokozawa, Y. Ohta, *Chem. Rev.* **2016**, *116*, 1950-1968.
[105] A. Yokoyama, R. Miyakoshi, T. Yokozawa, *Macromolecules* **2004**, *37*, 1169-1171.
[106] R. Miyakoshi, A. Yokoyama, T. Yokozawa, *Macromol. Rapid Commun.* **2004**, *25*, 1663-1666.
[107] E. E. Sheina, J. Liu, M. C. Iovu, D. W. Laird, R. D. McCullough, *Macromolecules* **2004**, *37*, 3526-3528.
[108] M. C. Iovu, E. E. Sheina, R. R. Gil, R. D. McCullough, *Macromolecules* **2005**, *38*, 8649-8656.
[109] R. Miyakoshi, A. Yokoyama, T. Yokozawa, *J. Am. Chem. Soc.* **2005**, *127*, 17542-17547.

[110] R. Tkachov, V. Senkovskyy, H. Komber, J.-U. Sommer, A. Kiriy, *J. Am. Chem. Soc.* **2010**, *132*, 7803-7810.
[111] M. Verswyvel, F. Monnaie, G. Koeckelberghs, *Macromolecules* **2011**, *44*, 9489-9498.
[112] P. Kohn, S. Huettner, H. Komber, V. Senkovskyy, R. Tkachov, A. Kiriy, R. H. Friend, U. Steiner, W. T. S. Huck, J.-U. Sommer, M. Sommer, *J. Am. Chem. Soc.* **2012**, *134*, 4790-4805.
[113] B. Tieke, *Makromolekulare Chemie: Eine Einführung*, 3rd ed., WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, **2014**.
[114] J. Huang, S. R. Turner, *Polymer* **2017**, *116*, 572-586.
[115] B. Klumperman, *Polym. Chem.* **2010**, *1*, 558-562.
[116] H. L. Hsieh, U.S. Patent 3,661,865, **1972**.
[117] Y. Hibi, S. Tokuoka, T. Terashima, M. Ouchi, M. Sawamoto, *Polym. Chem.* **2011**, *2*, 341-347.
[118] Y. Hibi, M. Ouchi, M. Sawamoto, *Angew. Chem.* **2011**, *123*, 7572-7575.
[119] M. Szwarc, *Nature* **1956**, *178*, 1168-1169.
[120] A. Jenkins, P. Kratochvil, R. Stepto, U. Suter, *Pure Appl. Chem.* **1996**, *68*, 2287-2311.
[121] P. J. Flory, *J. Am. Chem. Soc.* **1940**, *62*, 1561-1565.
[122] S. K. Varshney, J. Hautekeer, R. Fayt, R. Jérôme, P. Teyssié, *Macromolecules* **1990**, *23*, 2618-2622.
[123] T. Higashimura, O. Kishiro, *Polym. J.* **1977**, *9*, 87-93.
[124] R. Faust, J. P. Kennedy, *J. Polym. Sci., Part A: Polym. Chem.* **1987**, *25*, 1847-1869.
[125] M. Miyamoto, M. Sawamoto, T. Higashimura, *Macromolecules* **1984**, *17*, 265-268.
[126] L. R. Gilliom, R. H. Grubbs, *J. Am. Chem. Soc.* **1986**, *108*, 733-742.
[127] C. W. Bielawski, R. H. Grubbs, *Prog. Polym. Sci.* **2007**, *32*, 1-29.
[128] O. W. Webster, *Science* **1991**, *251*, 887-893.
[129] M. Ouchi, M. Sawamoto, *Macromolecules* **2017**, *50*, 2603-2614.
[130] K. Matyjaszewski, J. Xia, *Chem. Rev.* **2001**, *101*, 2921-2990.
[131] C. J. Hawker, A. W. Bosman, E. Harth, *Chem. Rev.* **2001**, *101*, 3661-3688.
[132] G. Moad, E. Rizzardo, S. H. Thang, *Aust. J. Chem.* **2009**, *62*, 1402-1472.
[133] J.-S. Wang, K. Matyjaszewski, *J. Am. Chem. Soc.* **1995**, *117*, 5614-5615.
[134] M. Kato, M. Kamigaito, M. Sawamoto, T. Higashimura, *Macromolecules* **1995**, *28*, 1721-1723.
[135] D. Solomon, E. Rizzardo, P. Cacioli, U.S. Patent 4,581,429, **1986**.
[136] J. Chiefari, Y. K. Chong, F. Ercole, J. Krstina, J. Jeffery, T. P. T. Le, R. T. A. Mayadunne, G. F. Meijs, C. L. Moad, G. Moad, E. Rizzardo, S. H. Thang, *Macromolecules* **1998**, *31*, 5559-5562.
[137] D. Benoit, C. J. Hawker, E. E. Huang, Z. Lin, T. P. Russell, *Macromolecules* **2000**, *33*, 1505-1507.
[138] S. Pfeifer, J.-F. Lutz, *J. Am. Chem. Soc.* **2007**, *129*, 9542-9543.
[139] S. Pfeifer, J.-F. Lutz, *Chem. Eur. J.* **2008**, *14*, 10949-10957.
[140] J.-F. Lutz, B. V. K. J. Schmidt, S. Pfeifer, *Macromol. Rapid Commun.* **2011**, *32*, 127-135.
[141] N. Baradel, S. Fort, S. Halila, N. Badi, J. F. Lutz, *Angew. Chem. Int. Ed.* **2013**, *52*, 2335-2339.
[142] S. Srichan, H. Mutlu, N. Badi, J. F. Lutz, *Angew. Chem. Int. Ed.* **2014**, *53*, 9231-9235.
[143] S. Srichan, H. Mutlu, J.-F. Lutz, *Eur. Polym. J.* **2015**, *62*, 338-346.
[144] N. Baradel, O. Gok, M. Zamfir, A. Sanyal, J.-F. Lutz, *Chem. Commun.* **2013**, *49*, 7280-7282.
[145] R. K. Roy, J.-F. Lutz, *J. Am. Chem. Soc.* **2014**, *136*, 12888-12891.
[146] O. Shishkan, M. Zamfir, M. A. Gauthier, H. G. Börner, J.-F. Lutz, *Chem. Commun.* **2014**, *50*, 1570-1572.
[147] S. Srichan, N. Kayunkid, L. Oswald, B. Lotz, J.-F. Lutz, *Macromolecules* **2014**, *47*, 1570-1577.
[148] J.-F. Lutz, *Acc. Chem. Res.* **2013**, *46*, 2696-2705.
[149] G. Gody, T. Maschmeyer, P. B. Zetterlund, S. Perrier, *Macromolecules* **2014**, *47*, 3451-3460.
[150] G. Gody, T. Maschmeyer, P. B. Zetterlund, S. Perrier, *Nat. Commun.* **2013**, *4*, 2505.
[151] L. Martin, G. Gody, S. Perrier, *Polym. Chem.* **2015**, *6*, 4875-4886.
[152] G. Gody, R. Barbey, M. Danial, S. Perrier, *Polym. Chem.* **2015**, *6*, 1502-1511.
[153] A. Kuroki, I. Martinez-Botella, C. H. Hornung, L. Martin, E. G. L. Williams, K. E. S. Locock, M. Hartlieb, S. Perrier, *Polym. Chem.* **2017**, *8*, 3249-3254.

[154] N. Zydziak, W. Konrad, F. Feist, S. Afonin, S. Weidner, C. Barner-Kowollik, *Nat. Commun.* **2016**, *7*, 13672.
[155] S. C. Solleder, D. Zengel, K. S. Wetzel, M. A. R. Meier, *Angew. Chem. Int. Ed.* **2016**, *55*, 1204-1207.
[156] Y. Jiang, M. R. Golder, H. V. T. Nguyen, Y. Wang, M. Zhong, J. C. Barnes, D. J. C. Ehrlich, J. A. Johnson, *J. Am. Chem. Soc.* **2016**, *138*, 9369-9372.
[157] F. A. Leibfarth, J. A. Johnson, T. F. Jamison, *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 10617-10622.
[158] W. Zhang, M. Huang, H. Su, S. Zhang, K. Yue, X.-H. Dong, X. Li, H. Liu, S. Zhang, C. Wesdemiotis, B. Lotz, W.-B. Zhang, Y. Li, S. Z. D. Cheng, *ACS Cent. Sci.* **2016**, *2*, 48-54.
[159] F. Amir, Z. Jia, M. J. Monteiro, *J. Am. Chem. Soc.* **2016**, *138*, 16600-16603.
[160] L. Oswald, T. T. Trinh, D. Chan-Seng, J.-F. Lutz, *Polymer* **2015**, *72*, 341-347.
[161] W. Zhang, D. P. Curran, *Tetrahedron* **2006**, *62*, 11837-11865.
[162] C. Y. Cho, E. J. Moran, S. R. Cherry, J. C. Stephans, S. P. Fodor, C. L. Adams, A. Sundaram, J. W. Jacobs, P. G. Schultz, *Science* **1993**, *261*, 1303-1305.
[163] G. Guichard, V. Semetey, M. Rodriguez, J.-P. Briand, *Tetrahedron Lett.* **2000**, *41*, 1553-1557.
[164] A. Boeijen, R. M. J. Liskamp, *Eur. J. Org. Chem.* **1999**, *1999*, 2127-2135.
[165] A. Boeijen, J. van Ameijde, R. M. J. Liskamp, *J. Org. Chem.* **2001**, *66*, 8454-8462.
[166] K. Burgess, H. Shin, D. S. Linthicum, *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 907-909.
[167] J. Farrera-Sinfreu, A. Aviñó, I. Navarro, J. Aymamí, N. G. Beteta, S. Varón, R. Pérez-Tomás, W. Castillo-Avila, R. Eritja, F. Albericio, M. Royo, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2440-2444.
[168] H. M. König, T. Gorelik, U. Kolb, A. F. M. Kilbinger, *J. Am. Chem. Soc.* **2007**, *129*, 704-708.
[169] N. R. Wurtz, J. M. Turner, E. E. Baird, P. B. Dervan, *Org. Lett.* **2001**, *3*, 1201-1203.
[170] E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1996**, *118*, 6141-6146.
[171] V. Kuksa, R. Buchan, P. K. T. Lin, *Synthesis* **2000**, *2000*, 1189-1207.
[172] O. Brümmer, B. Clapham, K. D. Janda, *Tetrahedron Lett.* **2001**, *42*, 2257-2259.
[173] T. M. Fyles, H. Luong, *Org. Biomol. Chem.* **2009**, *7*, 725-732.
[174] R. Szweda, C. Chendo, L. Charles, P. N. W. Baxter, J.-F. Lutz, *Chem. Commun.* **2017**, *53*, 8312-8315.
[175] N. F. Utesch, F. Diederich, *Org. Biomol. Chem.* **2003**, *1*, 237-239.
[176] C. A. Briehn, T. Kirschbaum, P. Bäuerle, *J. Org. Chem.* **2000**, *65*, 352-359.
[177] L. Jones, J. S. Schumm, J. M. Tour, *J. Org. Chem.* **1997**, *62*, 1388-1410.
[178] J. C. Nelson, J. K. Young, J. S. Moore, *J. Org. Chem.* **1996**, *61*, 8160-8168.
[179] C. Douat-Casassus, K. Pulka, P. Claudon, G. Guichard, *Org. Lett.* **2012**, *14*, 3130-3133.
[180] J.-M. Kim, Y. Bi, S. J. Paikoff, P. G. Schultz, *Tetrahedron Lett.* **1996**, *37*, 5305-5308.
[181] N. G. Angelo, P. S. Arora, *J. Org. Chem.* **2007**, *72*, 7963-7967.
[182] R. N. Zuckermann, J. M. Kerr, S. B. H. Kent, W. H. Moos, *J. Am. Chem. Soc.* **1992**, *114*, 10646-10647.
[183] S. Martens, J. Van den Begin, A. Madder, F. E. Du Prez, P. Espeel, *J. Am. Chem. Soc.* **2016**, *138*, 14182-14185.
[184] A. Dömling, I. Ugi, *Angew. Chem. Int. Ed.* **2000**, *39*, 3168-3210.
[185] S. C. Solleder, M. A. R. Meier, *Angew. Chem. Int. Ed.* **2014**, *53*, 711-714.
[186] W. Konrad, F. R. Bloesser, K. S. Wetzel, A. C. Boukis, M. A. R. Meier, C. Barner-Kowollik, *Chem. Eur. J.* **2018**, *24*, 3413-3419.
[187] S. C. Solleder, S. Martens, P. Espeel, F. Du Prez, M. A. R. Meier, *Chem. Eur. J.* **2017**, *23*, 13906-13909.
[188] S. C. Solleder, K. S. Wetzel, M. A. R. Meier, *Polym. Chem.* **2015**, *6*, 3201-3204.
[189] B. Ridder, D. S. Mattes, A. Nesterov-Mueller, F. Breitling, M. A. R. Meier, *Chem. Commun.* **2017**, *53*, 5553-5556.
[190] M. Hartweg, C. J. C. Edwards-Gayle, E. Radvar, D. Collis, M. Reza, M. Kaupp, J. Steinkoenig, J. Ruokolainen, R. Rambo, C. Barner-Kowollik, I. W. Hamley, H. S. Azevedo, C. R. Becer, *Polym. Chem.* **2018**, *9*, 482-489.
[191] M. H. Caruthers, *Science* **1985**, *230*, 281-285.

[192] M. D. Matteucci, M. H. Caruthers, *J. Am. Chem. Soc.* **1981**, *103*, 3185-3191.
[193] M. H. Caruthers, *J. Biol. Chem.* **2013**, *288*, 1420-1427.
[194] A. M. Michelson, A. R. Todd, *J. Chem. Soc.* **1955**, 2632-2638.
[195] M. Caruthers, R. Wells, *Science* **2011**, *334*, 1511-1511.
[196] H. Khorana, G. Tener, J. Moffatt, E. Pol, *Chem. Ind.* **1956**, 1523-1523.
[197] H. Khorana, W. Razzell, P. Gilham, G. Tener, E. Pol, *J. Am. Chem. Soc.* **1957**, *79*, 1002-1003.
[198] M. Smith, D. Rammler, I. Goldberg, H. Khorana, *J. Am. Chem. Soc.* **1962**, *84*, 430-440.
[199] H. Schaller, G. Weimann, B. Lerch, H. Khorana, *J. Am. Chem. Soc.* **1963**, *85*, 3821-3827.
[200] Y. Lapidot, H. Khorana, *J. Am. Chem. Soc.* **1963**, *85*, 3852-3857.
[201] Y. Lapidot, H. Khorana, *J. Am. Chem. Soc.* **1963**, *85*, 3857-3862.
[202] D. Söll, E. Ohtsuka, D. S. Jones, R. Lohrmann, H. Hayatsu, S. Nishimura, H. G. Khorana, *Proc. Natl. Acad. Sci. U. S. A.* **1965**, *54*, 1378-1385.
[203] K. L. Agarwal, H. Büchi, M. H. Caruthers, N. Gupta, H. G. Khorana, K. Kleppe, A. Kumar, E. Ohtsuka, U. L. Rajbhandary, J. H. Van De Sande, V. Sgaramella, H. Weber, T. Yamada, *Nature* **1970**, *227*, 27-34.
[204] H. Khorana, *Science* **1979**, *203*, 614-625.
[205] M. H. Caruthers, *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 18098-18099.
[206] R. L. Letsinger, M. J. Kornet, *J. Am. Chem. Soc.* **1963**, *85*, 3045-3046.
[207] R. L. Letsinger, M. J. Kornet, V. Mahadevan, D. M. Jerina, *J. Am. Chem. Soc.* **1964**, *86*, 5163-5165.
[208] R. L. Letsinger, V. Mahadevan, *J. Am. Chem. Soc.* **1965**, *87*, 3526-3527.
[209] R. L. Letsinger, V. Mahadevan, *J. Am. Chem. Soc.* **1966**, *88*, 5319-5324.
[210] H. Hayatsu, H. Khorana, *J. Am. Chem. Soc.* **1966**, *88*, 3182-3183.
[211] L. R. Melby, D. R. Strobach, *J. Am. Chem. Soc.* **1967**, *89*, 450-453.
[212] R. L. Letsinger, M. H. Caruthers, D. M. Jerina, *Biochemistry* **1967**, *6*, 1379-1388.
[213] R. L. Letsinger, K. K. Ogilvie, *J. Am. Chem. Soc.* **1967**, *89*, 4801-4803.
[214] R. L. Letsinger, K. K. Ogilvie, P. S. Miller, *J. Am. Chem. Soc.* **1969**, *91*, 3360-3365.
[215] R. L. Letsinger, M. H. Caruthers, P. S. Miller, K. K. Ogilvie, *J. Am. Chem. Soc.* **1967**, *89*, 7146-7147.
[216] G. W. Grams, R. L. Letsinger, *J. Org. Chem.* **1970**, *35*, 868-870.
[217] J. E. Davies, H. G. Gassen, *Angew. Chem., Int. Ed. Engl.* **1983**, *22*, 13-31.
[218] M. J. Gait, M. Singh, R. C. Sheppard, M. D. Edge, A. R. Greene, G. R. Heathcliffe, T. C. Atkinson, C. R. Newton, A. F. Markham, *Nucleic Acids Res.* **1980**, *8*, 1081-1096.
[219] R. Cook, O. Hudson, E. Mayran, J. Ott, in *Chemical and Enzymatic Synthesis of Gene Fragments: A Laboratory Manual* (Eds.: H. G. Gassen, A. Lang), Wiley-VCH Verlag GmbH, Weinheim, Germany, **1982**, pp. 111-130.
[220] R. C. Pless, R. L. Letsinger, *Nucleic Acids Res.* **1975**, *2*, 773-786.
[221] R. L. Letsinger, J. L. Finnan, G. A. Heavner, W. B. Lunsford, *J. Am. Chem. Soc.* **1975**, *97*, 3278-3279.
[222] R. L. Letsinger, W. B. Lunsford, *J. Am. Chem. Soc.* **1976**, *98*, 3655-3661.
[223] M. D. Matteucci, M. H. Caruthers, *Tetrahedron Lett.* **1980**, *21*, 719-722.
[224] L. J. McBride, M. H. Caruthers, *Tetrahedron Lett.* **1983**, *24*, 245-248.
[225] S. P. Adams, K. S. Kavka, E. J. Wykes, S. B. Holder, G. R. Galluppi, *J. Am. Chem. Soc.* **1983**, *105*, 661-663.
[226] N. D. Sinha, J. Biernat, J. McManus, H. Köster, *Nucleic Acids Res.* **1984**, *12*, 4539-4557.
[227] H. Köster, *Tetrahedron Lett.* **1972**, *13*, 1527-1530.
[228] G. R. Gough, M. J. Brunden, P. T. Gilham, *Tetrahedron Lett.* **1981**, *22*, 4177-4180.
[229] H. Köster, A. Stumpe, A. Wolter, *Tetrahedron Lett.* **1983**, *24*, 747-750.
[230] C. McCollum, A. Andrus, *Tetrahedron Lett.* **1991**, *32*, 4069-4072.
[231] S. L. Beaucage, R. P. Iyer, *Tetrahedron* **1992**, *48*, 2223-2311.
[232] M. P. Reddy, N. B. Hanna, F. Farooqui, *Nucleosides Nucleotides* **1997**, *16*, 1589-1598.
[233] Q. Zhu, M. O. Delaney, M. M. Greenberg, *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1105-1107.
[234] N. Usman, K. K. Ogilvie, M. Y. Jiang, R. J. Cedergren, *J. Am. Chem. Soc.* **1987**, *109*, 7845-7854.

[235] K. K. Ogilvie, N. Theriault, K. L. Sadana, *J. Am. Chem. Soc.* **1977**, *99*, 7741-7743.
[236] S. Pitsch, P. A. Weiss, L. Jenny, A. Stutz, X. Wu, *Helv. Chim. Acta* **2001**, *84*, 3773-3795.
[237] S. Pitsch, P. A. Weiss, X. Wu, D. Ackermann, T. Honegger, *Helv. Chim. Acta* **1999**, *82*, 1753-1761.
[238] J. C. Schulhof, D. Molko, R. Teoule, *Nucleic Acids Res.* **1987**, *15*, 397-416.
[239] H. Köster, K. Kulikowski, T. Liese, W. Heikens, V. Kohli, *Tetrahedron* **1981**, *37*, 363-369.
[240] L. J. McBride, R. Kierzek, S. L. Beaucage, M. H. Caruthers, *J. Am. Chem. Soc.* **1986**, *108*, 2040-2048.
[241] M. J. Heller, *Annu. Rev. Biomed. Eng.* **2002**, *4*, 129-153.
[242] S. Fodor, J. Read, M. Pirrung, L. Stryer, A. Lu, D. Solas, *Science* **1991**, *251*, 767-773.
[243] S. P. A. Fodor, R. P. Rava, X. C. Huang, A. C. Pease, C. P. Holmes, C. L. Adams, *Nature* **1993**, *364*, 555.
[244] S. Singh-Gasson, R. D. Green, Y. Yue, C. Nelson, F. Blattner, M. R. Sussman, F. Cerrina, *Nat. Biotechnol.* **1999**, *17*, 974.
[245] A. Hasan, K.-P. Stengele, H. Giegrich, P. Cornwell, K. R. Isham, R. A. Sachleben, W. Pfleiderer, R. S. Foote, *Tetrahedron* **1997**, *53*, 4247-4264.
[246] H. Giegrich, S. Eisele-Bühler, C. Hermann, E. Kvasyuk, R. Charubala, W. Pfleiderer, *Nucleosides Nucleotides* **1998**, *17*, 1987-1996.
[247] M. Beier, J. D. Hoheisel, *Nucleic Acids Res.* **2000**, *28*, e11-e11.
[248] C. Lausted, T. Dahl, C. Warren, K. King, K. Smith, M. Johnson, R. Saleem, J. Aitchison, L. Hood, S. R. Lasky, *Genome Biol.* **2004**, *5*, 58.
[249] R. D. Egeland, E. M. Southern, *Nucleic Acids Res.* **2005**, *33*, e125.
[250] K. Maurer, J. Cooper, M. Caraballo, J. Crye, D. Suciu, A. Ghindilis, J. A. Leonetti, W. Wang, F. M. Rossi, A. G. Stöver, *PLoS One* **2006**, *1*, e34.
[251] F. Seela, K. Kaiser, *Nucleic Acids Res.* **1987**, *15*, 3113-3129.
[252] M. Hariharan, K. Siegmund, F. D. Lewis, *J. Org. Chem.* **2010**, *75*, 6236-6243.
[253] M. J. Doktycz, T. M. Paner, A. S. Benight, *Biopolymers* **1993**, *33*, 1765-1777.
[254] W. Pils, R. Micura, *Nucleic Acids Res.* **2000**, *28*, 1859-1863.
[255] M. Durand, K. Chevrie, M. Chassignol, N. T. Thuong, J. C. Maurizot, *Nucleic Acids Res.* **1990**, *18*, 6353-6359.
[256] M. Durand, S. Peloille, N. Thuong, J. Maurizot, *Biochemistry* **1992**, *31*, 9197-9204.
[257] F. A. Aldaye, H. F. Sleiman, *Angew. Chem. Int. Ed.* **2006**, *45*, 2204-2209.
[258] F. A. Aldaye, P. K. Lo, P. Karam, C. K. McLaughlin, G. Cosa, H. F. Sleiman, *Nat. Nanotechnol.* **2009**, *4*, 349.
[259] N. Appukutti, C. J. Serpell, *Polym. Chem.* **2018**, *9*, 2210-2226.
[260] K. M. M. Carneiro, F. A. Aldaye, H. F. Sleiman, *J. Am. Chem. Soc.* **2010**, *132*, 679-685.
[261] T. G. W. Edwardson, K. M. M. Carneiro, C. K. McLaughlin, C. J. Serpell, H. F. Sleiman, *Nat. Chem.* **2013**, *5*, 868.
[262] T. G. W. Edwardson, K. M. M. Carneiro, C. J. Serpell, H. F. Sleiman, *Angew. Chem. Int. Ed.* **2014**, *53*, 4567-4571.
[263] D. de Rochambeau, M. Barlog, T. G. W. Edwardson, J. J. Fakhoury, R. S. Stein, H. S. Bazzi, H. F. Sleiman, *Polym. Chem.* **2016**, *7*, 4998-5003.
[264] T. Trinh, P. Chidchob, H. S. Bazzi, H. F. Sleiman, *Chem. Commun.* **2016**, *52*, 10914-10917.
[265] R. Häner, F. Garo, D. Wenger, V. L. Malinovskii, *J. Am. Chem. Soc.* **2010**, *132*, 7466-7471.
[266] A. L. Nussbaumer, D. Studer, V. L. Malinovskii, R. Häner, *Angew. Chem.* **2011**, *123*, 5604-5608.
[267] F. Simona, A. L. Nussbaumer, R. Häner, M. Cascella, *J. Phys. Chem. B* **2013**, *117*, 2576-2585.
[268] M. Vybornyi, A. V. Rudnev, R. Häner, *Chem. Mater.* **2015**, *27*, 1426-1431.
[269] M. Vybornyi, A. V. Rudnev, S. M. Langenegger, T. Wandlowski, G. Calzaferri, R. Häner, *Angew. Chem. Int. Ed.* **2013**, *52*, 11488-11493.
[270] N. Micali, M. Vybornyi, P. Mineo, O. Khorev, R. Häner, V. Villari, *Chemistry – A European Journal* **2015**, *21*, 9505-9513.

[271] M. Vybornyi, Y. Bur-Cecilio Hechevarria, M. Glauser, A. V. Rudnev, R. Häner, *Chem. Commun.* **2015**, *51*, 16191-16193.
[272] C. D. Bösch, S. M. Langenegger, R. Häner, *Angew. Chem.* **2016**, *128*, 10115-10118.
[273] J.-A. Amalian, A. Al Ouahabi, G. Cavallo, N. F. König, S. Poyer, J.-F. Lutz, L. Charles, *J. Mass Spectrom.* **2017**, *52*, 788-798.
[274] E. Callaway, *Nature* **2015**, *521*, 18-19.
[275] H. Colquhoun, J.-F. Lutz, *Nat. Chem.* **2014**, *6*, 455-456.
[276] F. Sanger, A. R. Coulson, *J. Mol. Biol.* **1975**, *94*, 441-448.
[277] L. M. Smith, J. Z. Sanders, R. J. Kaiser, P. Hughes, C. Dodd, C. R. Connell, C. Heiner, S. B. H. Kent, L. E. Hood, *Nature* **1986**, *321*, 674-679.
[278] D. C. Koboldt, K. M. Steinberg, D. E. Larson, R. K. Wilson, E. R. Mardis, *Cell* **2013**, *155*, 27-38.
[279] S. Goodwin, J. D. McPherson, W. R. McCombie, *Nat. Rev. Genet.* **2016**, *17*, 333-351.
[280] D. Deamer, M. Akeson, D. Branton, *Nat. Biotechnol.* **2016**, *34*, 518.
[281] H. Mutlu, J. F. Lutz, *Angew. Chem. Int. Ed.* **2014**, *53*, 13010-13019.
[282] Z. Zhu, C. J. Cardin, Y. Gan, H. M. Colquhoun, *Nat. Chem.* **2010**, *2*, 653-660.
[283] A. Burel, C. Carapito, J.-F. Lutz, L. Charles, *Macromolecules* **2017**, *50*, 8290-8296.
[284] J. Davis, *Art J. (N. Y.)* **1996**, *55*, 70-74.
[285] C. Bancroft, T. Bowler, B. Bloom, C. T. Clelland, *Science* **2001**, *293*, 1763-1765.
[286] C. Gustafsson, *Nature* **2009**, *458*, 703.
[287] M. Ailenberg, O. D. Rotstein, *BioTechniques* **2009**, *47*, 747-754.
[288] N. G. Portney, Y. Wu, L. K. Quezada, S. Lonardi, M. Ozkan, *Langmuir* **2008**, *24*, 1613-1616.
[289] Y. Nozomu, S. Kazuhide, S. Junichi, O. Yoshiaki, T. Masaru, *Biotechnol. Progr.* **2007**, *23*, 501-505.
[290] P. C. Wong, K.-K. Wong, H. Foote, *Commun. ACM* **2003**, *46*, 95-98.
[291] C. T. Clelland, V. Risca, C. Bancroft, *Nature* **1999**, *399*, 533-534.
[292] J. M. Tien, *J. Syst. Sci. Syst. Eng.* **2013**, *22*, 127-151.
[293] N. Goldman, P. Bertone, S. Chen, C. Dessimoz, E. M. LeProust, B. Sipos, E. Birney, *Nature* **2013**, *494*, 77-80.
[294] R. Heckel, *Nat. Biotechnol.* **2018**, *36*, 236.
[295] R. N. Grass, R. Heckel, M. Puddu, D. Paunescu, W. J. Stark, *Angew. Chem. Int. Ed.* **2015**, *54*, 2552-2555.
[296] Y. Erlich, D. Zielinski, *Science* **2017**, *355*, 950-954.
[297] S. L. Shipman, J. Nivala, J. D. Macklis, G. M. Church, *Nature* **2017**, *547*, 345-349.
[298] K. Fister, I. Fister, J. Murovec, B. Bohanec, *Transgenic Res.* **2017**, *26*, 87-95.
[299] T. T. Trinh, L. Oswald, D. Chan-Seng, J.-F. Lutz, *Macromol. Rapid Commun.* **2014**, *35*, 141-145.
[300] T. T. Trinh, L. Oswald, D. Chan-Seng, L. Charles, J.-F. Lutz, *Chem. Eur. J.* **2015**, *21*, 11961-11965.
[301] J. A. Amalian, T. T. Trinh, J.-F. Lutz, L. Charles, *Anal. Chem.* **2016**, *88*, 3715-3722.
[302] L. Charles, G. Cavallo, V. Monnier, L. Oswald, R. Szweda, J.-F. Lutz, *J. Am. Soc. Mass. Spectrom.* **2017**, *28*, 1149-1159.
[303] R. K. Roy, A. Meszynska, C. Laure, L. Charles, C. Verchin, J.-F. Lutz, *Nat. Commun.* **2015**, *6*, 7237.
[304] R. K. Roy, C. Laure, D. Fischer-Krauser, L. Charles, J.-F. Lutz, *Chem. Commun.* **2015**, *51*, 15677-15680.
[305] L. Charles, C. Laure, J.-F. Lutz, R. K. Roy, *Macromolecules* **2015**, *48*, 4319-4328.
[306] L. Charles, C. Laure, J.-F. Lutz, R. K. Roy, *Rapid Commun. Mass Spectrom.* **2016**, *30*, 22-8.
[307] C. Laure, D. Karamessini, O. Milenkovic, L. Charles, J.-F. Lutz, *Angew. Chem. Int. Ed.* **2016**, *55*, 10722-10725.
[308] A. K. Ghosh, T. T. Doung, S. P. McKee, W. J. Thompson, *Tetrahedron Lett.* **1992**, *33*, 2781-2784.
[309] Ufuk S. Gunay, Benoît É. Petit, D. Karamessini, A. Al Ouahabi, J.-A. Amalian, C. Chendo, M. Bouquey, D. Gigmes, L. Charles, J.-F. Lutz, *Chem* **2016**, *1*, 114-126.
[310] S. Telitel, B. É. Petit, S. Poyer, L. Charles, J.-F. Lutz, *Polym. Chem.* **2017**, *8*, 4988-4991.

[311] D. Karamessini, B. É. Petit, M. Bouquey, L. Charles, J.-F. Lutz, *Adv. Funct. Mater.* **2017**, *27*, 1604595.
[312] D. Karamessini, S. Poyer, L. Charles, J.-F. Lutz, *Macromol. Rapid Commun.* **2017**, *38*, 1700426.
[313] D. Karamessini, T. Simon-Yarza, S. Poyer, E. Konishcheva, L. Charles, D. Letourneur, J.-F. Lutz, *Angew. Chem. Int. Ed.* **2018**, *57*, 10574-10578.
[314] G. Cavallo, S. Poyer, J. A. Amalian, F. Dufour, A. Burel, C. Carapito, L. Charles, J.-F. Lutz, *Angew. Chem. Int. Ed.* **2018**.
[315] A. C. Boukis, K. Reiter, M. Frölich, D. Hofheinz, M. A. R. Meier, *Nat. Commun.* **2018**, *9*, 1439.
[316] A. C. Boukis, M. A. R. Meier, *Eur. Polym. J.* **2018**, *104*, 32-38.
[317] Y. L. Ying, J. Zhang, R. Gao, Y. T. Long, *Angew. Chem. Int. Ed.* **2013**, *52*, 13154-13161.
[318] H. Bayley, *Clin. Chem.* **2015**, *61*, 25-31.
[319] C. Dekker, *Nat. Nanotechnol.* **2007**, *2*, 209-215.
[320] G. M. Cherf, K. R. Lieberman, H. Rashid, C. E. Lam, K. Karplus, M. Akeson, *Nat. Biotechnol.* **2012**, *30*, 344.
[321] I. M. Derrington, T. Z. Butler, M. D. Collins, E. Manrao, M. Pavlenok, M. Niederweis, J. H. Gundlach, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 16060-16065.
[322] E. A. Manrao, I. M. Derrington, A. H. Laszlo, K. W. Langford, M. K. Hopper, N. Gillgren, M. Pavlenok, M. Niederweis, J. H. Gundlach, *Nat. Biotechnol.* **2012**, *30*, 349-353.
[323] A. H. Laszlo, I. M. Derrington, B. C. Ross, H. Brinkerhoff, A. Adey, I. C. Nova, J. M. Craig, K. W. Langford, J. M. Samson, R. Daza, K. Doering, J. Shendure, J. H. Gundlach, *Nat. Biotechnol.* **2014**, *32*, 829-833.
[324] K. A. Günay, P. Theato, H. A. Klok, *J. Polym. Sci., Part A: Polym. Chem.* **2013**, *51*, 1-28.
[325] J.-F. Lutz, *Angew. Chem. Int. Ed.* **2007**, *46*, 1018-1025.
[326] P. M. E. Gramlich, S. Warncke, J. Gierlich, T. Carell, *Angew. Chem. Int. Ed.* **2008**, *47*, 3442-3444.
[327] N. F. König, A. Al Ouahabi, S. Poyer, L. Charles, J.-F. Lutz, *Angew. Chem. Int. Ed.* **2017**, *56*, 7297-7301.
[328] J. Mandal, S. Krishna Prasad, D. S. S. Rao, S. Ramakrishnan, *J. Am. Chem. Soc.* **2014**, *136*, 2538-2545.
[329] M. Qi, M. Hülsmann, A. Godt, *J. Org. Chem.* **2016**, *81*, 2549-2571.
[330] L. J. O'Driscoll, D. J. Welsh, S. W. D. Bailey, D. Visontai, H. Frampton, M. R. Bryce, C. J. Lambert, *Chem. Eur. J.* **2015**, *21*, 3891-3894.
[331] H. B. Nulwala, C. N. Tang, B. W. Kail, K. Damodaran, P. Kaur, S. Wickramanayake, W. Shi, D. R. Luebke, *Green Chem.* **2011**, *13*, 3345-3349.
[332] M. R. Molla, P. Prasad, S. Thayumanavan, *J. Am. Chem. Soc.* **2015**, *137*, 7286-7289.
[333] J. ter Maat, R. Regeling, C. J. Ingham, C. A. G. M. Weijers, M. Giesbers, W. M. de Vos, H. Zuilhof, *Langmuir* **2011**, *27*, 13606-13617.
[334] W. Bornmann, M. Alauddin, J. Gelovani, P. Ghosh, L. Guo, D. Han, D. Maxwell, U. Mukhopadhyay, Z. Peng, A. Shavrin, WO Patent 2008024826, **2008**.
[335] M. P. Reddy, P. J. Voelker, *Int. J. Pept. Protein Res.* **1988**, *31*, 345-348.
[336] T. H. Vaughn, J. A. Nieuwland, *J. Am. Chem. Soc.* **1933**, *55*, 2150-2153.
[337] J. E. Hein, J. C. Tripp, L. B. Krasnova, K. B. Sharpless, V. V. Fokin, *Angew. Chem. Int. Ed.* **2009**, *48*, 8018-8021.
[338] D. N. Barsoum, N. Okashah, X. Zhang, L. Zhu, *J. Org. Chem.* **2015**, *80*, 9542-9551.
[339] H. Schwarz, in *Lehrbuch der chemisch-analytischen Titrirmethode, 2nd Ed.* (Ed.: F. Mohr), Friedrich Vieweg und Sohn, Braunschweig, **1853**, p. 117.
[340] V. A. Schmidt, E. J. Alexanian, *J. Am. Chem. Soc.* **2011**, *133*, 11402-11405.
[341] J. Escudero, V. Bellosta, J. Cossy, *Angew. Chem. Int. Ed.* **2018**, *57*, 574-578.
[342] A. Kamimura, Y. Kawakami, *Phosphorus Sulfur Silicon Relat. Elem.* **2016**, *191*, 259-262.
[343] L. Song, M. R. Hobaugh, C. Shustak, S. Cheley, H. Bayley, J. E. Gouaux, *Science* **1996**, *274*, 1859-1866.
[344] I. Iacovache, S. De Carlo, N. Cirauqui, M. Dal Peraro, F. G. van der Goot, B. Zuber, *Nat. Commun.* **2016**, *7*, 12062.

[345] M. Faller, M. Niederweis, G. E. Schulz, *Science* **2004**, *303*, 1189-92.
[346] M. A. Tasdelen, Y. Yagci, *Angew. Chem. Int. Ed.* **2013**, *52*, 5930-5938.
[347] H. Masuda, N. Watanabe, H. Naruoka, S. Nagata, K. Takagaki, T. Wada, J. Yano, *Biorg. Med. Chem.* **2010**, *18*, 8277-8283.
[348] G. Wang, T. Peng, S. Zhang, J. Wang, X. Wen, H. Yan, L. Hu, L. Wang, *RSC Adv.* **2015**, *5*, 28344-28348.
[349] L. Oswald, A. Al Ouahabi, L. Charles, J.-F. Lutz, *Chem. Eur. J.* **2016**, *22*, 3462-3469.
[350] H. Yi, S. Maisonneuve, J. Xie, *Org. Biomol. Chem.* **2009**, *7*, 3847-3854.
[351] D. Wöll, J. Smirnova, M. Galetskaya, T. Prykota, J. Bühler, K.-P. Stengele, W. Pfleiderer, U. E. Steiner, *Chem. Eur. J.* **2008**, *14*, 6490-6497.
[352] S. Bühler, I. Lagoja, H. Giegrich, K.-P. Stengele, W. Pfleiderer, *Helv. Chim. Acta* **2004**, *87*, 620-659.
[353] S. Walbert, W. Pfleiderer, U. E. Steiner, *Helv. Chim. Acta* **2001**, *84*, 1601-1611.
[354] T. Anwar, S. Paul, S. K. Singh, *Int. J. Bio-Sci. Bio-Technol.* **2014**, *6*, 215-222.
[355] G. Xiao, M. Lu, L. Qin, X. Lai, *Chin. Sci. Bull.* **2006**, *51*, 1413-1420.
[356] M. A. Palacios, E. Benito-Peña, M. Manesse, A. D. Mazzeo, C. N. LaFratta, G. M. Whitesides, D. R. Walt, *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 16510-16514.
[357] H. Zhao, E. S. Sterner, E. B. Coughlin, P. Theato, *Macromolecules* **2012**, *45*, 1723-1736.
[358] X. Wu, H. Zhao, B. Nörnberg, P. Theato, G. A. Luinstra, *Macromolecules* **2014**, *47*, 492-497.
[359] Y. Huang, R. Dong, X. Zhu, D. Yan, *Soft Matter* **2014**, *10*, 6121-6138.
[360] P. D. Sankar, K. Haridas, G. Suhrit, *Chem. Eur. J.* **2016**, *22*, 16872-16877.
[361] T. Mes, R. van der Weegen, A. R. A. Palmans, E. W. Meijer, *Angew. Chem.* **2011**, *123*, 5191-5195.
[362] M. Aso, M. Fukuda, K. Usui, H. Suemune, *Nucleic Acids Symp. Ser.* **2004**, *48*, 59-60.
[363] J. Kim, J. M. Gil, M. M. Greenberg, *Angew. Chem.* **2003**, *115*, 6062-6065.
[364] R. Roger, M. Brendan, A. Anna, G. Raimundo, E. Ramon, *Helv. Chim. Acta* **2009**, *92*, 613-622.
[365] A. Galeone, L. Mayol, G. Oliviero, D. Rigano, M. Varra, *Bioorg. Med. Chem. Lett.* **2001**, *11*, 383-386.
[366] A. F. Abdel-Magid, K. G. Carson, B. D. Harris, C. A. Maryanoff, R. D. Shah, *J. Org. Chem.* **1996**, *61*, 3849-3862.
[367] T. László, B. István, Á. Béla, H. Zoltán, L. Ernő, T. Klára, H. Mária, H. Salman, P. Ernő, *Liebigs Ann. Chem.* **1988**, *1988*, 549-554.
[368] Q. Jing, H. Li, L. J. Roman, P. Martásek, T. L. Poulos, R. B. Silverman, *Bioorg. Med. Chem. Lett.* **2014**, *24*, 4504-4510.
[369] P. Regenass, J.-F. Margathe, A. Mann, J. Suffert, M. Hibert, N. Girard, D. Bonnet, *Chem. Commun.* **2014**, *50*, 9657-9660.
[370] Y. Hirayama, M. Iwamura, T. Furuta, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 905-908.
[371] M. R. Pitts, J. R. Harrison, C. J. Moody, *J. Chem. Soc., Perkin Trans. 1* **2001**, 955-977.
[372] A. Kenji, K. Hiroyuki, K. Tetsuya, O. Masahiro, O. Yoshiro, Y. Masayuki, N. Manabu, S. Hideo, *J. Labelled Compd. Radiopharm.* **2013**, *56*, 562-572.
[373] E. Manzo, M. Carbone, E. Mollo, C. Irace, A. Di Pascale, Y. Li, M. L. Ciavatta, G. Cimino, Y.-W. Guo, M. Gavagnin, *Org. Lett.* **2011**, *13*, 1897-1899.
[374] K. Mannerstedt, O. Hindsgaul, *Carbohydr. Res.* **2008**, *343*, 875-881.
[375] J. E. T. Corrie, A. Barth, V. R. N. Munasinghe, D. R. Trentham, M. C. Hutter, *J. Am. Chem. Soc.* **2003**, *125*, 8546-8554.
[376] M. T. Reetz, S. H. Kyung, M. Hüllmann, *Tetrahedron* **1986**, *42*, 2931-2935.
[377] T. R. Chan, R. Hilgraf, K. B. Sharpless, V. V. Fokin, *Org. Lett.* **2004**, *6*, 2853-2855.
[378] B. Ourri, O. Tillement, T. Tu, E. Jeanneau, U. Darbost, I. Bonnamour, *New J. Chem.* **2016**, *40*, 9477-9485.
[379] H. Ankati, E. Biehl, *Tetrahedron Lett.* **2009**, *50*, 4677-4682.
[380] J. Hoque, M. M. Konai, S. Gonuguntla, G. B. Manjunath, S. Samaddar, V. Yarlagadda, J. Haldar, *J. Med. Chem.* **2015**, *58*, 5486-5500.
[381] M. Gosangi, H. Rapaka, V. Ravula, S. V. Patri, *Bioconjugate Chem.* **2017**, *28*, 1965-1977.

[382] C.-W. Huang, P.-W. Wu, W.-H. Su, C.-Y. Zhu, S.-W. Kuo, *Polym. Chem.* **2016**, *7*, 795-806.
[383] R. Tomlinson, M. Klee, S. Garrett, J. Heller, R. Duncan, S. Brocchini, *Macromolecules* **2002**, *35*, 473-480.
[384] Y. Zhang, G. B. Evans, K. Clinch, D. R. Crump, L. D. Harris, R. Fröhlich, P. C. Tyler, K. Z. Hazleton, M. B. Cassera, V. L. Schramm, *J. Biol. Chem.* **2013**, *M113*, 521955.
[385] L. Koren-Selfridge, H. N. Londino, J. K. Vellucci, B. J. Simmons, C. P. Casey, T. B. Clark, *Organometallics* **2009**, *28*, 2085-2090.
[386] X. Shu, Z. Lu, J. Zhu, *Chem. Mater.* **2010**, *22*, 3310-3312.
[387] A. K. Patra, S. K. Kundu, D. Kim, A. Bhaumik, *ChemCatChem* **2015**, *7*, 791-798.
[388] S. Adimurthy, S. Ghosh, P. U. Patoliya, G. Ramachandraiah, M. Agrawal, M. R. Gandhi, S. C. Upadhyay, P. K. Ghosh, B. C. Ranu, *Green Chem.* **2008**, *10*, 232-237.
[389] T. Šolomek, S. Mercier, T. Bally, C. G. Bochet, *Photochem. Photobiol. Sci.* **2012**, *11*, 548-555.
[390] Z. X. Giustra, K. L. Tan, *Chem. Commun.* **2013**, *49*, 4370-4372.
[391] A. P. Dieskau, B. Plietker, *Org. Lett.* **2011**, *13*, 5544-5547.
[392] G. G. Haraldsson, T. Stefansson, H. Snorrason, *Acta Chem. Scand.* **1998**, *52*, 824-826.
[393] L. Zhang, A. E. Peritz, P. J. Carroll, E. Meggers, *Synthesis* **2006**, *4*, 645-653.
[394] K. Jeyakumar, D. K. Chand, *Synthesis* **2008**, *5*, 807-819.
[395] T. Saito, Y. Nishimoto, M. Yasuda, A. Baba, *J. Org. Chem.* **2006**, *71*, 8516-8522.
[396] L. Yin, J. Wu, J. Xiao, S. Cao, *Tetrahedron Lett.* **2012**, *53*, 4418-4421.
[397] T. Zweifel, J.-V. Naubron, T. Büttner, T. Ott, H. Grützmacher, *Angew. Chem. Int. Ed.* **2008**, *47*, 3245-3249.
[398] I. Seven, T. Weinrich, M. Gränz, C. Grünewald, S. Brüß, I. Krstić, T. F. Prisner, A. Heckel, M. W. Göbel, *Eur. J. Org. Chem.* **2014**, *2014*, 4037-4043.
[399] M. Boukhet, N. F. König, A. A. Ouahabi, G. Baaken, J.-F. Lutz, J. C. Behrends, *Macromol. Rapid Commun.* **2017**, *38*, 1700680.
[400] M. W. Parker, J. T. Buckley, J. P. M. Postma, A. D. Tucker, K. Leonard, F. Pattus, D. Tsernoglou, *Nature* **1994**, *367*, 292-295.

Résumé :
Récemment, la chimie des phosphoramidites s'est montrée efficace et polyvalente en tant que plateforme pour accéder à des poly(phosphodiester)s à séquence définies abiotiques. Grâce à cette stratégie, les monomères peuvent être placés dans la chaîne à des positions choisies, ouvrant la voie à de nombreuses possibilités pour la préparation de macromolécules fonctionnelles.
Ici, la méthode phosphoramidite a été explorée pour la synthèse de polymères dits numériques, qui contiennent des séquences de monomères encodées binairement. Des polymères dont les longueurs de chaînes et les séquences numériques sont contrôlées ont été préparés en utilisant une stratégie phosphoramidite classique impliquant des groupements protecteurs diméthoxytrityles, ou bien un procédé photocontrôlé faisant intervenir des groupements nitrophénylpropyloxycarbonyles clivables à la lumière. En outre, plusieurs stratégies pour modifier l'information contenue dans les chaînes latérales ont été étudiées dans cette thèse. Une modification binaire post-polymérisation à travers deux cycloadditions alcyne-azoture catalysées par le cuivre(I) consécutives a été examinée pour optimiser les chaînes latérales des poly(phosphodiester)s à séquences définies. De plus, la libération photocontrôlée de différents motifs éthers ortho-nitrobenzyliques latéraux a été étudiée. Ces fonctions ont permis la conception d'oligo(phosphodiester)s numériques dont les séquences d'information peuvent être effacées ou révélées grâce à la lumière.

Mots-clés :
polymères encodés numériquement · polymères à séquences définies · chimie des phosphoramidites · synthèse en phase solide · modification post-polymérisation · CuAAC · synthèse de poly(phosphodiester)s photocontrôlée · effacer/révéler des séquences d'information.

Abstract:
Phosphoramidite chemistry has recently been evidenced to be an efficient and versatile platform to access sequence-defined abiotic poly(phosphodiester)s. Using this strategy, monomers can be placed at defined positions positions in a chain, thus opening up wide possibilities for the preparation of functional macromolecules.
Here, the phosphoramidite platform was explored to synthesize so-called digital polymers, which contain monomer-coded binary sequences. Polymers with controlled chain lengths and digital sequences were prepared using either a standard phosphoramidite strategy involving dimethoxytrityl protective groups or a photo-controlled process involving light-cleavable nitrophenylpropyloxycarbonyl protective groups. Additionally, several strategies to modify the side chain information were investigated in this thesis. A binary post-polymerization modification by means of sequential copper(I)-catalyzed alkyne-azide cycloadditions was investigated for tuning the side chain functionality of sequence-defined poly(phosphodiester)s. Moreover, the photo-controlled release of several *ortho*-nitrobenzylic ether side chain motifs was studied. These moieties allowed the design of digital oligo(phosphodiester)s whose sequence information can be erased or revealed with light as a trigger.

Keywords:
digitally encoded polymers · sequence-defined polymers · phosphoramidite chemistry · solid-phase synthesis · post-polymerization modification · CuAAC · photocontrolled phosphodiester synthesis · erase/reveal sequence information