

**Expanding the Scope of Sequence-Defined
Oligo(Phosphodiester)s with Novel Building Blocks**

By

Donatien de Rochambeau

*A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of
Doctor of Philosophy*

Department of Chemistry, McGill University

Montréal, Québec, Canada

May, 2019

© Donatien de Rochambeau, 2019

I dedicate this thesis to my parents, Édith et Hubert

Abstract

Biopolymers such as proteins and nucleic acids adopt three-dimensional (3D) shapes dictated by their monomer sequence. This well-defined structure is responsible for the functions of biopolymers in biological systems. The field of sequence-controlled polymers (SCPs) seeks to mimic the precision of biopolymer composition, but with synthetic polymers incorporating a greater variety of artificial monomers. The potential to outperform biopolymers in several applications has been a driving force in the development of novel SCPs. Numerous examples reported in the past few years highlight the promises of precision polymers in areas such as information encoding, self-assembly and folding of nanostructures, as well as drug discovery, delivery and catalysis. However, few artificial polymers have been shown to achieve the degree of efficiency and selectivity of biopolymers. While novel synthetic approaches have broadened the structural and chemical diversity of SCPs, the challenges of generating polymers with (i) total sequence-control, (ii) relatively high degree of polymerization and (iii) versatile monomer content remain an obstacle to the emergence of novel applications. Automated solid-phase phosphoramidite chemistry has the advantage of addressing the first two challenges. This thesis focuses on the third challenge, by expanding the alphabet of available phosphoramidite monomers. These would introduce a variety of potential supramolecular interactions and direct the self-assembly of sequence-defined oligo(phosphodiester)s. In Chapter 2, a perfluorocarbon (PFC) containing monomer is designed and successfully incorporated into nucleic acid strands. PFC chains introduce the “fluorous” supramolecular interaction, leading to the formation of micelles that are detectable by ^{19}F NMR, nuclease-resistant DNA duplexes with high melting temperatures, and a potent perfluorocarbon-modified RNA strand for gene silencing applications. Chapter 3 describes the development of a naphthalene-based monomer, which drives the formation of nanostructures of different shapes and sizes, including spherical micelles and 2D nanosheets, upon incorporation into various precision oligomers. The significant impact of monomer sequence on self-assembly outcome was assessed by making subtle sequence modifications. Furthermore, Chapter 4 reports a two-step protocol to access a large library of monomers from two tertiary amine-based achiral molecular platforms. The numerous building blocks accessible through this strategy yield oligomers with a variety of side-chains such as amino acids and carbohydrates. Chapter 5 describes a combinatorial approach based on DNA-encoded non-nucleosidic aptamers designed to leverage monomer diversity. This methodology yields a large library of sequence-

defined oligomers that structurally resemble the thrombin-binding aptamer. Overall, this thesis reports on the development of numerous monomers compatible with standard automated solid-phase phosphoramidite chemistry which can be readily incorporated into sequence-defined oligo(phosphodiester)s. This work has yielded novel materials and tools to identify sequences of interest, and should pave the way to expanded applications of SCPs in drug delivery, self-assembly of structurally diverse materials, as well as discovery of novel aptamers and catalysts made of artificial building blocks.

Résumé

Les biopolymères tels que les protéines ou les acides nucléiques, adoptent une structure bien définie, dictée par leur séquence de monomères. Cet agencement en trois dimensions est crucial pour que les biopolymères assument leur fonction au sein des systèmes biologiques. Les polymères à séquence contrôlée (PSCs) sont des macromolécules synthétiques dont la séquence de monomères tend à imiter la composition précise des biopolymères. Ils sont façonnés à partir d'une grande diversité de monomères artificiels et représentent un domaine de recherche en pleine expansion. En effet, de tels polymères ont le potentiel d'être plus efficaces que les biopolymères dans certains domaines. L'utilité de nombreux PSCs a déjà été démontrée en catalyse, dans le codage et stockage de données, l'auto-assemblage et le pliage de nanostructures ou encore la découverte de nouveaux médicaments et de vecteurs de transport de ces derniers. Cependant, peu de ces exemples ont montré une efficacité de l'ordre de ce qu'accomplissent les biopolymères. De nombreuses approches synthétiques ont permis de créer des PSCs d'une grande diversité chimique et structurelle. Néanmoins, les efforts se poursuivent pour trouver des méthodes de synthèse permettant d'obtenir : (i) un contrôle total de la séquence, (ii) un degré de polymérisation cohérent avec les applications pressenties et (iii) une grande variété de monomères compatibles. La synthèse automatisée en phase solide avec des phosphoramidites remplit les deux premiers critères. En agrandissant l'alphabet de phosphoramidites disponibles et donc le nombre d'interactions supramoléculaires disponibles, influencer l'auto-assemblage des oligomères à séquences définies de type phosphodiester serait possible. C'est l'objet de cette thèse. En premier lieu, un monomère de type perfluorocarbure (PFC) est synthétisé et incorporé efficacement dans des brins d'ADN. Cela permet d'instaurer un nouveau type d'interactions supramoléculaires spécifique aux PFC. Ainsi, des micelles détectables par RMN du fluor, des duplex d'ADN à la température de fusion élevée et résistants aux nucléases ou encore un brin d'ARN modifié avec un PFC et capable d'extinction de gène sont décrits. Dans un deuxième temps, le développement d'un nouveau monomère composé de naphthalène est présenté. Ce dernier permet la synthèse de polymères à séquence définie formant des micelles sphériques de différentes taille ou des films à l'épaisseur nanoscopique. Grâce à l'introduction de différences subtiles dans la composition des oligomères, l'impact de la séquence sur l'auto-assemblage a pu être étudié. Par la suite, deux amines tertiaires non-chirales sont utilisées comme plateformes moléculaires pour synthétiser un grand nombre de

nouveaux monomères en deux étapes seulement. Cette stratégie permet d'obtenir des oligomères à séquence contrôlée contenant des chaînes latérales très variées comme des acides aminés ou des monosaccharides. Finalement, cette diversité est mise à profit avec la synthèse combinatoire d'aptamères faits à partir de ces nouveaux monomères et liés à des codes-barres en ADN. Ainsi, une grande bibliothèque d'aptamères dont la structure moléculaire est très différente de celle des oligonucléotides classiques est synthétisée. Cette bibliothèque a été construite dans le but de trouver de nouveaux ligands pour la thrombine. En général, cette thèse décrit le développement de nombreux monomères compatibles avec la synthèse automatisée en phase solide avec des phosphoramidites, pour leur incorporation dans des oligomères à séquence définie de type phosphodiester. Ces travaux ont permis la découverte de nouveaux nanomatériaux et d'outils pour identifier des séquences d'intérêt. Celles-ci devraient mener à de nombreuses opportunités en auto-assemblage de divers matériaux ainsi qu'à la découverte de vecteurs de transport de médicaments, de nouveaux aptamères et de catalyseurs faits de monomères artificiels.

Acknowledgments

First, I would like to thank Dr. Hanadi Sleiman for her supervision. From day one when I was a trainee, her excitement regarding a new idea or a good result was contagious and helped me to commit for the extraordinary experience that is a PhD. Her creativity and resilience during the last five years were essential for the completion of my PhD.

I would like to thank my high-school and prep school chemistry teachers, Mrs. Ribes and Mr. Hanachi as well as Dr. Zard and Dr. Grimaud from Ecole Polytechnique for having seeded the love for chemistry in me.

All Sleimanites are valuable members of the Sleiman group to me and all contributed to this thesis in a way. First, I want to acknowledge Michael Dore and Aurélie Lacroix whose unwavering support and friendship helped me so much along this journey. I feel really grateful that we joined the lab almost at the same time and could overcome obstacles together. Tuan Trinh has also started with me and I am grateful to have worked with such a funny and hard-working friend.

Dr. Karina Carneiro and Dr. Tom Edwardson were my mentors and their training (in science but also in beer knowledge and in Scottish traditions) was so helpful. I would like to especially thank Dr. Kai Lin Lau whose personality was very refreshing in the lab environment. Dr. Nicole Avakyan and Dr. Katherine Bujold were the sunshine Québécois of the Sleiman lab and I am grateful for all the great memories together. I would like to thank Hammond Sun. Having him as my student was a very happy experience. A special thanks to Dr. Violeta Toader with whom I was happy to collaborate with on many projects and shared nice conversations. Dr. Empar Vengut-Climent and Dr. Mohammad Askari, thanks for having been the most secret Santa Claus. Dr. Danny Bousmail and Dr. Johans Fakhoury's cheerful disposition was highly appreciated especially during lunch time. I acknowledge Dr. Amani Hariri for bringing joy wherever she goes. I thank Dr. Justin Conway for all the great music in the lab and parties in the South Shore. I am grateful to Dr. Pongphak Chidchob who was always so helpful in the lab. I would like to thank Dr. Sonia Romero: working with her was a fulfilling experience. Dr. Alain Li, Dr. Pierre Quérard, Sosthène Ung, thanks for being so mythical. I thank the numerous people among them who proofread my thesis.

Of course, I am also very grateful to the people who joined after me: Felix Rizzuto for having been such a funny and cool friend for the last year, Casey Platnich for her awesome Calgarian-ness, Hassan Fakeeh for his bro-ness, Xin Luo for his constant helpfulness, Daniel Saliba for the nice

collaboration and the poem about my bike, Kamila Mustafina for her beautiful voice, Christophe Lachance-Brais, Alex Prinzen and many others.

I would like to thank all the people who helped me collecting data and analyzing them. First, Dr. Robin Stein, for the hard work of maintaining so many NMR spectrometers in perfect shape and for the nice collaboration, thank you! I thank Dr. Nadim Saade and Dr. Alexander Wahba for their great help with MS characterization, Dr. Beatrice Lego and Dr. Mohini Ramkaran for the AFM and LVEM, Dr. Masad Damha and Leonora Abdullahu for their help in setting up the levulinyll strategy in Chapter 5, Dr. Maureen McKeague for being so enthusiastic about my last project and for her expertise in aptamer discovery, Dr. Pierre Lepage and Julie Boudreau from Génome Québec. A special thanks to Chantal Marotte who has always been so helpful. I am also grateful to everybody working in the department office and the chemistry stores.

My experience as a TA was very fulfilling and made me realize how much I like teaching. For this, I would like to thank Dr. Laura Pavelka whose cheerful energy and organizational skills were highly appreciated. I also thank the TAs I had the chance to work with: Michael Dore, Samantha Gateman, Alicia Wu, Igor Huskic, Siba Moussa.

A special thanks to Jean-Michel Blais, TR/ST, Arcade Fire, Daniel Balavoine, Étienne Daho, Patrick Watson, Rone, Radiohead and so many others for their art and the motivation they inspired. Merci à Lise et Yvon pour leur accueil et leur soutien. Merci à mon équipe de Dans la rue. Je n'en serai pas là sans mes amis que je remercie infiniment. Amis de Montréal, de l'X (notamment la Kès et les volleyeurs), de Fermat, de St Jo, des scouts... Notamment un grand merci aux déglings de l'apéro, ma deuxième famille, dont je suis fier d'être le CEO. Un merci tout particulier à Lisa, Claire, Manon qui ont été de vrais piliers tout au long de mon PhD. Je vous aime tellement!

Bien sûr, j'aimerais remercier ma famille. Gustave et ma filleule Margaux, vous voir en vrai et sur Skype a toujours été un grand moment de joie simple. Merci à Nicolas et ma grande sœur Maÿlis, leurs merveilleux parents, qui m'ont largement soutenu dans cette aventure.

J'aimerais remercier Jean-Michel qui m'a soutenu et fait tellement grandir au cours des trois dernières années. Ta présence a engendré en moi de nombreux changements qui m'aiguillent sur la route du bonheur. Je t'aime.

Enfin, Papa, Maman, merci pour votre soutien et votre amour. Vous êtes des parents extraordinaires. J'espère vous rendre fiers avec cette thèse.

Contribution of authors

Dr. Hanadi Sleiman provided funding for the work presented in this thesis.

Chapter 2, **Donatien de Rochambeau** codesigned the project and performed all experiments unless listed below, analyzed the results and cowrote the manuscript. **Dr. Maciej Barłóg** synthesized compounds **1**, **2** and **3**. **Dr. Thomas Edwardson** provided training and helped for RNA synthesis. **Dr. Johans Fakhoury** did the cell work and the luciferase assay data collection. **Dr. Robin Stein** helped design and perform the NMR experiments. **Dr. Hanadi Sleiman** codesigned the project, guided interpretation of data and discussion of results and cowrote the manuscript.

Chapter 3, **Donatien de Rochambeau** codesigned the project and performed all experiments unless listed below, analyzed the results and cowrote the manuscript. **Dr. Maciej Barłóg** synthesized compounds **NAP''**, **NAP'** and **C12'**. **Dr. Violeta Toader** synthesized compound **C12**. **Dr. Hanadi Sleiman** codesigned the project, guided interpretation of data and discussion of results and cowrote the manuscript.

Chapter 4, **Donatien de Rochambeau** codesigned the project and performed all experiments unless listed below, analyzed the results and cowrote the manuscript. **Yuanye Sun** synthesized compound **5'** and **5''** and prepared the first batch of **6** and its precursors. **Dr. Maciej Barłóg** synthesized compounds **1'**, **2'**, **4'**, **9'** and their precursors. **Dr. Hanadi Sleiman** codesigned the project, guided interpretation of data and discussion of results and cowrote the manuscript.

Chapter 5, **Donatien de Rochambeau** codesigned the project and performed all experiments unless listed below, analyzed the results and cowrote the manuscript. **Michael Dore** synthesized the DMT-protected precursor of **Bal** and **Dr. Violeta Toader** performed the phosphoramidite synthesis step for **Bal** and **C12**. **Daniel Saliba** performed and optimized the PCR amplification experiments. **Dr. Hanadi Sleiman** codesigned the project, guided interpretation of data and discussion of results and cowrote the manuscript.

Table of Contents

1	Introduction	1
1.1	Sequence control towards functional macromolecules	3
1.2	Repurposing biopolymers	5
1.2.1	DNA synthesis	5
1.2.2	Applications of Nucleic acids	7
1.2.3	Conclusions of Section 1.2	13
1.3	Synthetic routes towards sequence-controlled polymers made of unnatural building blocks.....	14
1.3.1	Polymerization of sequence-defined macromonomers	15
1.3.2	Timely incorporation of monomers in a polymer chain.....	17
1.3.3	Templated synthesis.....	18
1.3.4	Iterative strategies	21
1.3.5	Overview of SCP synthetic approaches.....	29
1.4	Applications of sequence-controlled polymers made of unnatural building blocks.....	30
1.4.1	Information Coding, Decoding and Storage with sequence-controlled polymers	30
1.4.2	The impact of sequence control on self-assembly and folding	33
1.4.3	Sequence-Controlled Polymers for biomedical applications	39
1.4.4	Sequence-controlled polymers design	41
1.4.5	Overview of SCP applications	45
1.5	Context and scope of this thesis	46
1.6	References.....	47
2	“DNA-Teflon” Sequence-Defined Oligomers.....	67
2.1	Preface.....	69
2.2	Contribution of authors	70
2.3	Abstract	70
2.4	Introduction.....	70
2.5	Results and discussion.....	72
2.5.1	Synthesis of PFC-containing sequence-defined oligomers.....	72
2.5.2	“DNA-Teflon” oligomers self-assembly studies	75
2.5.3	¹⁹ F NMR detection of “DNA-Teflon” oligomers	78
2.5.4	“Fluorous” effect in PFC-containing DNA duplexes.....	79
2.5.5	Nuclease resistance of PFC-containing DNA duplexes.....	82
2.5.6	PFC-containing siRNA silencing properties.....	83
2.6	Conclusions.....	84

2.7	Experimental section	85
2.7.1	Chemicals	85
2.7.2	Instrumentation	85
2.7.3	Small molecule synthesis	86
2.7.4	Solid-phase synthesis	89
2.7.5	LC-ESI-MS characterization	92
2.7.6	Atomic force microscopy.....	101
2.7.7	Dynamic Light Scattering	104
2.7.8	Gel electrophoresis	106
2.7.9	NMR study.....	108
2.7.10	Melting curves	110
2.7.11	Serum stability assay	112
2.7.12	Gene silencing assay.....	114
2.7.13	pKa of N[PFC]	115
2.8	References.....	115
[3]	Self-Assembled Nanostructure Library from Monodisperse Sequence-Defined Oligo(Phosphodiester)s 119	
3.1	Preface.....	121
3.2	Contribution of authors	122
3.3	Abstract	122
3.4	Introduction.....	122
3.5	Results and discussion.....	124
3.5.1	Phosphoramidite monomers from diols	124
3.5.2	Synthesis of sequence-defined block co-oligomers with a DNA synthesizer.....	125
3.5.3	Purification of sequence-defined oligo(phosphodiester)s.....	128
3.5.4	Self-assembly characterization	130
3.5.5	Influence of sequence on self-assembly.....	135
3.6	Conclusions.....	137
3.7	Experimental section.....	138
3.7.1	Chemicals	138
3.7.2	Instrumentation	139
3.7.3	Small molecule synthesis.....	140
3.7.4	Determination of absorption coefficient of 2,7-dihydroxynaphthalene and oligomers	144
3.7.5	Solid-phase synthesis.....	144
3.7.6	Gel electrophoresis purification.....	148
3.7.7	LC-ESI-MS characterization	148

3.7.8	Atomic force microscopy.....	151
3.7.9	Transmission Electron Microscopy.....	159
3.7.10	Dynamic Light Scattering.....	160
3.7.11	Gel electrophoresis analysis.....	161
3.8	References.....	163
[4]	Modular Strategy to Expand the Chemical Diversity of Sequence-Defined Oligo(Phosphodiester)s .	167
4.1	Preface.....	169
4.2	Contribution of authors.....	170
4.3	Abstract.....	170
4.4	Introduction.....	170
4.5	Results and discussion.....	172
4.5.1	Tertiary amine backbones.....	172
4.5.2	Synthesis of the optimized monomer platforms.....	177
4.5.3	Sequence-defined oligo(phosphodiester)s.....	180
4.5.4	Design of an additional primary amine-containing platform.....	182
4.6	Conclusion.....	183
4.7	Experimental section.....	184
4.7.1	Chemicals.....	184
4.7.2	Instrumentation.....	185
4.7.3	Small molecule synthesis.....	185
4.7.4	Solid-phase Synthesis.....	202
4.7.5	Post solid-phase synthesis functionalization of alkyne containing DNA strands.....	204
4.7.6	Gel electrophoresis.....	205
4.7.7	RP-HPLC purification and analysis.....	207
4.7.8	LC-ESI-MS characterization.....	210
4.8	References.....	216
[5]	DNA-Encoded Functionalized Aptamers.....	221
5.1	Preface.....	223
5.2	Contribution of authors.....	224
5.3	Abstract.....	224
5.4	Introduction.....	224
5.5	Results and discussion.....	227
5.5.1	Finding an orthogonal protecting group.....	227
5.5.2	Branching unit and branched oligonucleotides.....	230
5.5.3	Aptamer design.....	235
5.5.4	Synthesis of model compounds and the library.....	239

5.5.5	PCR and sequencing	241
5.6	Conclusions	243
5.7	Experimental section	244
5.7.1	Chemicals	244
5.7.2	Instrumentation	245
5.7.3	Small molecule synthesis	245
5.7.4	Solid-Phase Synthesis	258
5.7.5	Reverse-phase HPLC	261
5.7.6	Gel electrophoresis	263
5.7.7	LC-MS	266
5.7.8	PCR	271
5.7.9	Sequencing	273
5.8	References	274
6	Conclusions and Future Work	279
6.1	Contributions to original knowledge	281
6.2	Future work	283
6.3	Publications	285
6.4	References	286
7	NMR Spectra	287
7.1	Molecules from Chapter 2	289
7.1.1	Compound 1	289
7.1.2	Compound 2	289
7.1.3	Compound 3	290
7.1.4	Monomer N[PFC]	291
7.2	Molecules from Chapter 3	292
7.2.1	Compound C12'	292
7.2.2	Monomer C12	293
7.2.3	Compound NAP''	294
7.2.4	Compound NAP'	295
7.2.5	Monomer NAP	296
7.3	Molecules from Chapter 4	297
7.3.1	Compound 1''	297
7.3.2	Compound 1'	298
7.3.3	Compound 3''	299
7.3.4	Compound 3'	299
7.3.5	Monomer 3	300

7.3.6	Compound 4''	301
7.3.7	Compound 4'	302
7.3.8	Monomer 4	303
7.3.9	Compound 5''	304
7.3.10	Compound 5'	305
7.3.11	Monomer 5	305
7.3.12	Compound PT1a	306
7.3.13	Platform PT1	307
7.3.14	Monomer 6	308
7.3.15	Compound 7'	309
7.3.16	Monomer 7	310
7.3.17	Compound PT2a	311
7.3.18	Compound PT2b	311
7.3.19	Platform PT2	312
7.3.20	Compound 8'	313
7.3.21	Monomer 8	314
7.3.22	Platform PT3	315
7.3.23	Compound 11'	316
7.3.24	Monomer 11	317
7.4	Molecules from Chapter 5.....	318
7.4.1	Compound 1	318
7.4.2	Compound 2	319
7.4.3	Compound 3	319
7.4.4	Compound 5	320
7.4.5	Compound 6	320
7.4.6	Compound 7	321
7.4.7	Branching unit BU	322
7.4.8	Compound 8	323
7.4.9	Compound 9	324
7.4.10	Monomer Ant	325
7.4.11	Compound 10	326
7.4.12	Monomer His	327
7.4.13	Compound 11	328
7.4.14	Monomer Trp	329
7.4.15	Compound Bal'	330
7.4.16	Monomer Bal	330

List of abbreviations

ACE	2'-Acetoxy ethyl orthoester
ACN	Acetonitrile
AGE	Agarose gel electrophoresis
AROP	Anionic ring opening polymerization
ASO	Antisense oligonucleotide
ATRA	Atom transfer radical addition
ATRP	Atom transfer radical polymerization
BCP	Block copolymers
br.	Broad
CDI	1,1'-Carbonyldiimidazole
CEP	<i>N,N'</i> -Diisopropylcyanoethylphosphoramidite
CeNA	Cyclohexene nucleic acid
CMC	Critical Micelle Concentration
CPG	Controlled pore glass
CRISPR	Clustered regularly interspaced short palindromic repeats
CV	Column volume
Cye	Cyanine
d	Doublet
DCM	Dichloromethane
dd	Doublet of doublets
DEL	DNA-encoded library
DIC	<i>N,N'</i> -Diisopropylcarbodiimide
DIPEA	Diisopropylethylamine
DMEM	Dulbecco's modified Eagle's minimal essential medium
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DMT	Dimethoxytrityl
DNA	Deoxyribonucleic acid
DP	Degree of polymerization
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
equiv.	Equivalent
EM	Electron microscopy
ESI	Electrospray ionization

EtOAc	Ethyl acetate
ETT	5-(Ethylthio)- <i>1H</i> -tetrazole
FANA	Fluoroarabino nucleic acid
FBS	Fetal bovine serum
Fmoc	Fluorenylmethyloxycarbonyl
FRNA	2'-Fluoro RNA
GNA	Glycerol nucleic acid
hdDNA	2',3'-Dideoxy- β -D-glucopyranose DNA
HEG	Hexaethylene glycol
HOBt	Hydroxybenzotriazole
HPLC	High-performance liquid chromatography
HRMS	High-resolution mass spectrometry
HTS	High-throughput screening
IEG	Iterative exponential growth
<i>J</i>	Coupling constant
LCAA	Long chain alkylamine
LC-MS	Liquid Chromatography-Mass spectrometry
LNA	Locked Nucleic Acid
LRMS	Low resolution mass spectrometry
LVEM	Low voltage electron microscopy
m	Multiplet
mRNA	Messenger RNA
MRI	Magnetic resonance imaging
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
NHS	<i>N</i> -hydroxysuccinimide
NMP	Nitroxide-mediated radical polymerization
NMR	Nuclear magnetic resonance
PAGE	Polyacrylamide gel electrophoresis
PEG	Polyethylene glycol
PET	Photo-induced electron transfer
PCR	Polymerase chain reaction
PFC	Perfluorocarbon
PMO	Phosphorodiamidate morpholino oligomer
PNA	Peptide nucleic acid
PPI	Protein-protein interaction

q	Quadruplet
qRT-PCR	Quantitative real-time polymerase chain reaction
quint	Quintuplet
RAFT	Reversible addition-fragmentation chain-transfer
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
ROMP	Ring opening metathesis polymerization
ROP	Ring opening polymerization
s	Singlet
SCP	Sequence-controlled polymers
SELEX	Systematic evolution of ligands by exponential enrichment
siRNA	Silencing RNA
SOMAmer	Slow off-rate modified aptamer
SPS	Solid-phase synthesis
SUMI	Single-unit monomer insertion
t	Triplet
TALEN	Transcription activator-like effector nuclease
TAMg	Tris Acetate buffer supplemented with magnesium ions
TA5Mg	Tris Acetate buffer supplemented with 5 times more magnesium than TAMg
TBA	Thrombin-binding aptamer
TBAF	Tetrabutylammonium fluoride
TBE	Tris-Boric acid buffer with EDTA
TBDMS	<i>tert</i> -Butyldimethylsilyl
TBTU	<i>O</i> -(Benzotriazol-1-yl)- <i>N,N,N',N'</i> -tetramethyluronium tetrafluoroborate
TEA	Triethylamine
TEAA	Triethylammonium acetate buffer
TEM	Transmission electron microscopy
TEMED	Tetramethylethylenediamine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TNA	Threose nucleic acid
TOM	<i>O</i> -Triisopropylsilyloxymethyl
XNA	Xeno-nucleic acid
X-SELEX	SELEX for XNA
ZFN	Zinc finger nuclease

List of Figures

Figure 1.1. Sequence leads to 3D structure that induces function.	4
Figure 1.2. Examples of DNA nanostructures.	9
Figure 1.3. Gene silencing mechanism and chemical modifications in siRNA and ASO.	10
Figure 1.4. SELEX process and chemical modifications used in aptamers.	12
Figure 1.5. The different sequence-controlled polymers.	15
Figure 1.6. Step-growth polymerization with sequence-defined macromonomers.....	16
Figure 1.7. Chain-growth polymerization with sequence-defined macromonomers.	16
Figure 1.8. Accurate positioning of monomers in chain-growth polymerization.	17
Figure 1.9. DNA-templated synthesis of sequence-defined oligomers.....	18
Figure 1.10. Parallel DNA-templated synthesis using PNA adapters.....	19
Figure 1.11. Sequential autonomous DNA-templated synthesis using NHS ester-amine couplings.....	20
Figure 1.12. Molecular machine performing sequence-defined synthesis along a rotaxane thread.....	20
Figure 1.13. Chain-growth stepwise polymerization through SUMI.....	22
Figure 1.14. Solid-phase synthesis with protection/deprotection strategy.	23
Figure 1.15. Solid-phase synthesis: submonomer strategy.	25
Figure 1.16. Solid-phase synthesis involving different types of chemistry.	26
Figure 1.17. Liquid-phase synthesis involving different chemistry and purification processes.	28
Figure 1.18. Iterative exponential growth.	29
Figure 1.19. Information-encoded sequence-controlled polymers.....	32
Figure 1.20. Information storage in SCP: improvements and applications.....	33
Figure 1.21. Self-assembly of sequence-defined oligomers.	35
Figure 1.22. Applications of foldamers.....	36
Figure 1.23. Single-chain folding in sequence-controlled polymers.	38
Figure 1.24. Biomedical applications with SCP.	40
Figure 1.25. Combinatorial synthesis strategies for peptides, small molecules and artificial sequence-defined oligomers.....	43
Figure 1.26. Biopolymer-unnatural SCP hybrid structures.....	45
Figure 2.1. Reverse-phase-HPLC traces from crude mixtures of “DNA-Teflon” oligomers (B_n).....	75
Figure 2.2. Dry AFM images of B₅ and B₁₀	75
Figure 2.3. Effect of sequence-control and Mg ²⁺ addition for “DNA–Teflon” co-oligomers assembly....	77
Figure 2.4. Agarose gel electrophoresis analysis of “DNA-Teflon” co-oligomers.....	78
Figure 2.5. ¹⁹ F NMR spectra of B₄ (left) and B₁₀ (right) in D ₂ O.....	79

Figure 2.6. Representative melting curves of PFC-modified DNA duplexes in Mg ²⁺ containing buffer. .	80
Figure 2.7. Thermodynamics data for PFC-modified DNA duplexes extracted from melting curves.....	82
Figure 2.8. Serum stability assay for PFC-modified DNA duplexes.	83
Figure 2.9. Silencing efficiency of PFC-modified siRNA.	84
Figure 2.10. Reverse-phase HPLC traces from crude mixtures of A _{n,m} oligomers.....	91
Figure 2.11. MS data for A _{n,m} oligomers.	94
Figure 2.12. LC/MS data for DNA-“Teflon” oligomers (B _n).	96
Figure 2.13. LC/MS data for DNA with internal modifications or 3’end modification.	100
Figure 2.14. LC/MS data for RNA with 3’end N[PFC] modification.....	100
Figure 2.15. Dry AFM images of unmodified DNA control B ₀	101
Figure 2.16. Dry AFM images of AT-(N[PFC]) ₅ micelles (B ₅).	102
Figure 2.17. Dry AFM images of AT-(N[PFC]) ₁₀ micelles (B ₁₀).....	103
Figure 2.18. Dry AFM images of B ₂	103
Figure 2.19. Dry AFM images of C ₆ . White bars represent 400 nm.....	104
Figure 2.20. DLS intensity correlation functions for sequence-defined oligomers samples.....	106
Figure 2.21. 15 % denaturing PAGE analysis of “DNA-Teflon” oligomers.	107
Figure 2.22. 8 % native PAGE of PFC-modified duplexes.	107
Figure 2.23. ¹⁹ F NMR spectra of “DNA-Teflon” oligomers.....	109
Figure 2.24. Representative melting curves of PFC-modified DNA duplexes at several concentrations in TAMg.....	111
Figure 2.25. Average of DNA-duplexes degradation curves.	114
Figure 3.1. Strategy to obtain a library of precision oligomers with a DNA synthesizer.	127
Figure 3.2. HPLC purification of DMT-on oligomers.	129
Figure 3.3. Gel electrophoresis analysis and purification of unnatural oligo(phosphodiester)s.	130
Figure 3.4. Dry AFM images of HEG ₄ -NAP ₄ from 100 μM sample in TAMg.	131
Figure 3.5. Dry AFM images of control and HEG ₄ -NAP ₄ from 10 μM sample in TA5Mg+NiCl ₂	131
Figure 3.6. AFM images of HEG ₄ -NAP ₄ and HEG ₆ -NAP ₆ from 50 μM sample in TA5Mg+NiCl ₂	132
Figure 3.7. TEM images of HEG ₄ -NAP ₄ from 25 μM sample in TA5Mg	133
Figure 3.8. Representative AFM images of A ₈ , B ₃ and B ₅ forming spherical micelles.....	133
Figure 3.9. 2.5 % AGE in TAMg of some of the oligomers made.	135
Figure 3.10. Absorbance as a function of 2,7-dihydroxynaphtalene concentration.....	144
Figure 3.11. Reverse-phase HPLC traces (UV detection, 260 nm) from oligomers.....	147
Figure 3.12. MS data for sequence-controlled oligomers, negative mode.	151
Figure 3.13. Supplementary dry AFM images of oligomer A ₄ , 50μM.	152

Figure 3.14. Supplementary dry AFM images of oligomer A7 , 50 μ M	153
Figure 3.15. Dry AFM images of spherical micelles forming oligomers.	154
Figure 3.16. Liquid AFM image of oligomer A8 , 10 μ M.....	155
Figure 3.17. Supplementary liquid AFM images of oligomer B3 , 50 μ M.....	155
Figure 3.18. Liquid AFM images of oligomer B2	156
Figure 3.19. Liquid AFM images of oligomer B6 , 10 μ M.	157
Figure 3.20. Supplementary liquid AFM images of oligomer A4 , 50 μ M.	158
Figure 3.21. Supplementary TEM images of oligomer A4	159
Figure 3.22. Representative DLS intensity correlation curves for 25 μ M solutions of oligomers.	160
Figure 3.23. Dynamic light scattering on oligomers B5 and B6 in water after annealing	161
Figure 3.24. Denaturing gel (PAGE 20 %) with oligomers.	162
Figure 3.25. Agarose gel (2.5 % AGE) with oligomers A1 to A7	162
Figure 3.26. Agarose gel (2.5 % AGE) with oligomers from the A and B series.	163
Figure 4.1. Synthesis of versatile monomers for sequence-defined oligo(phosphodiester)s.	172
Figure 4.2. Attachment of 3 , 4 and 5 to DNA and oligomers.	176
Figure 4.3. Attachment of PT1 , PT2b , 7 and 8 to DNA.	179
Figure 4.4. RP-HPLC chromatograms of sequence-defined oligomers with novel monomers.	181
Figure 4.5. Post SPS functionalization of alkyne monomer 6 with β -D-glycosyl azide.....	182
Figure 4.6. Attachment of amine-functionalized containing monomer in a DNA 21mer.....	183
Figure 4.7. 18 % PAGE in denaturing conditions, of modified DNA strands.	205
Figure 4.8. 15 % PAGE in denaturing conditions of 21mer modified with 11	206
Figure 4.9. 18 % PAGE in denaturing conditions of modified DNA strands with post SPS click chemistry.	206
Figure 4.10. RP-HPLC traces from crude mixtures of 5' modified DNA 19mers.....	208
Figure 4.11. RP-HPLC traces from crude mixtures of sequence-defined oligomers.....	209
Figure 4.12. MS data for modified DNA strands.....	214
Figure 4.13. MS data for sequence-defined oligomers.	216
Figure 5.1. Selection principle of DNA-encoded unnatural aptamer libraries.....	227
Figure 5.2. General design of DNA-encoded unnatural aptamers.	228
Figure 5.3. Synthesis of DNA strands with 5'-Lev and 5'-DMT phosphoramidites.	230
Figure 5.4. Analysis of BR3 synthesis.....	234
Figure 5.5. Synthesis of modified DNA 21mers and analysis of hydrazine influence.	238
Figure 5.6. Design and synthesis of DNA-encoded thrombin-binding aptamers made of unnatural monomers.....	240

Figure 5.7. PCR with DNA-encoded aptamers.....	242
Figure 5.8. RP-HPLC chromatogram of AT-2-Nap	261
Figure 5.9. RP-HPLC chromatogram of BR3	262
Figure 5.10. RP-HPLC chromatogram of AT, AT-His, AT-Trp	262
Figure 5.11. RP-HPLC chromatogram of AT-Ant (top), AT-Bal (bottom)	263
Figure 5.12. 15 % PAGE, denaturing conditions, HYD1, LEV1, LEV2	264
Figure 5.13. 12 % PAGE, denaturing conditions, T10, T30, BR1, BR2, BR3	264
Figure 5.14. 15 % PAGE denaturing conditions, AT-Ant, AT, AT-Bal, AT-His, AT-Trp	265
Figure 5.15. 12 % PAGE, denaturing conditions, TBA1, TBA2, TBA3 and LIB	265
Figure 5.16. MS data for modified DNA strands.....	271
Figure 5.17. 12 % PAGE, denaturing conditions of amplicons after PCR with BR1, BR2, BR3, TBA1, TBA2, TBA3	272
Figure 5.18. 8 % PAGE, native conditions of amplicons after PCR with BR1, BR2, BR3, TBA1, TBA2, TBA3	272

List of Schemes

Scheme 1.1. Synthetic cycle in an automated DNA synthesizer.....	6
Scheme 2.1. Synthetic pathway to N[PFC]	72
Scheme 2.2. Synthesis of versatile sequence-defined oligomers with HEG , N[PFC] and nucleoside phosphoramidites.	73
Scheme 2.3. pKa of protonated counterpart of perfluorocarbon containing amines found on SciFinder.	115
Scheme 3.1. General strategy to obtain a DNA synthesizer-compatible monomer from a diol.....	125
Scheme 3.2. Synthetic pathway to make naphthalene-containing phosphoramidite monomer NAP	125
Scheme 4.1. Different backbones explored.	173
Scheme 4.2. Synthetic route for 1'	174
Scheme 4.3. Potential degradation mechanism of diethanolamine based monomers.....	174
Scheme 4.4. Synthetic route for 3	175
Scheme 4.5. Synthetic route for 4	175
Scheme 4.6. Synthetic route for 5	175
Scheme 4.7. Synthetic route for platform PT1 , alkyne phosphoramidite 6 and a β -D-glucose-containing phosphoramidite 7	177
Scheme 4.8. Synthetic route for platform PT2 and a phenylalanine-functionalized phosphoramidite 8	178
Scheme 4.9. Synthesis of platform 3 (PT3) and phosphoramidite 11	183
Scheme 5.1. Synthesis of 5'-Teoc protected thymidine and attachment to DNA.	229
Scheme 5.2. Synthesis of branching unit BU	231
Scheme 5.3. Parallel synthesis of branched oligo(phosphodiester)s.	233
Scheme 5.4. Phosphoramidite monomers used in this chapter.....	237
Scheme 5.5. Synthesis of Ant , His , Trp	237
Scheme 5.6. Split and pool strategy for the synthesis of a DNA-encoded unnatural aptamer library.....	241

List of Tables

Table 2.1. Yields and ESI-MS characterization of amphiphilic oligomers A_{n,m}	74
Table 2.2. Yields and ESI-MS characterization of “DNA-Teflon” hybrids B_n	74
Table 2.3. Summary of DLS results for “DNA-Teflon” co-oligomers.	76
Table 2.4. ¹⁹ F NMR relaxation times and micelle concentrations measured.	79
Table 2.5. DNA, RNA and conjugates sequence.	92
Table 2.6. LC/MS results for DNA with internal modifications or 3’ end modification.	97
Table 2.7. Melting temperature of PFC-containing DNA duplexes.....	110
Table 2.8. Average thermodynamic values found for PFC-modified duplexes	112
Table 2.9. Summary of exponential decay analysis for duplexes degradation in serum.....	113
Table 3.1. Yields and MS characterization of the library of oligomers.	127
Table 3.2. Self-assembly characterization of the oligomers made.	134
Table 4.1. Yields and ESI-MS characterization of oligomers containing monomers 3, 4 and 5 ...	177
Table 4.2. Yields and ESI-MS characterization of DNA strands modified with PT1, PT2b, 7 and 8	180
Table 4.3. Yields and ESI-MS characterization of sequence-defined oligo(phosphodiester)s.	181
Table 5.1. Sequencing of amplicons from BR1, BR2, BR3 and TBA1 to TBA3 templates.....	243
Table 5.2. Sequence of strands used, and special synthetic conditions.....	259
Table 5.3. Monomers used during library synthesis.	261
Table 5.4. ESI-MS characterization of the strands synthesized.....	266
Table 5.5. Master Mix components for barcoding step.....	273

|1|

Introduction

1.1 Sequence control towards functional macromolecules

Enzymes are amongst the most efficient catalysts known. They are able to accelerate chemical reactions up to 10^{19} times.¹ Nature primarily relies on enzymes in life-sustaining processes. Indeed, food digestion, synthesis of high-value natural products, energy storage and conversion through nutrient and oxygen metabolisms, toxins and metabolic waste degradation, and healing processes occur with the help of enzymatic reactions. Enzymatic functions and efficiency can be transposed from biological reactions to chemical systems, where they can be used to tackle emerging issues, such as the green synthesis of chemical products, energy conversion from renewable resources, energy storage, non-toxic and low-polluting waste products management, and the healing of damaged ecosystems. Studying and mimicking enzymes paves the way to a very large panel of discoveries that have already started to revolutionize chemistry.

The specific sequence of the enzyme's comprising components is the key feature for efficient biocatalysis. Most enzymes are made of one or several polypeptide chains, peptides, composed of 20 different building blocks, amino acids. Once placed at specific positions on the peptidic backbone, the different chemical functionalities of the amino acids may interact with each other through weak or covalent bonds. Well-defined 2D and 3D shapes called secondary structures may result from these interactions. The final peptide shape is called tertiary structure and is a direct consequence of the initial sequence of monomers. Peptides can also associate to form the final form of proteins known as the quaternary structure. The enzyme shape allows it to be specific for its substrate. Once substrates are bound, the close location of specific moieties decreases the kinetic reaction barrier, enabling rapid turnover to generate the product. As a conclusion, the function of enzymes, as well as their efficiency and selectivity, rely on their 3-dimensional structure, which is itself dictated by the primary sequence of monomer constituents. This relationship, that sequence induces shape which itself induces function, probably is one of the most important lessons learnt from natural catalytic systems (**Figure 1.1**).

Nature also relies on sequence control in other biopolymers. Polyphenols, oligosaccharides, ribonucleic acid (RNA) and deoxyribonucleic acid (DNA) are of fundamental importance in life-sustaining processes. The first of these has an important role in plant biology,² while oligosaccharides are mostly used for cell recognition and adhesion.³ RNA is known as an information relay molecule, but it is also capable of catalysis (ribozymes) and gene regulation.⁴

DNA is the molecule of genetic storage. As such, it has been broadly studied in the past few decades. DNA is composed of nucleotide monomers linked together through phosphodiester bonds. Each nucleotide is composed of a 5-carbon deoxyribose sugar with a nitrogenous base (Figure 1.1). There are four types of nucleobases that can pair with one another – adenine with thymine and guanine with cytosine to form a double helix. The structure was elucidated in 1953 by Francis Crick and James Watson using X-ray diffraction data obtained by Rosalind Franklin and Raymond Gosling.⁵ The sequence of monomers that forms the complementary strands in the DNA double helix constitutes the genetic information. Human cells contain DNA composed of about 3 billion bases,⁶ which are equivalent to 0.75 gigabytes of information. The density of information storage in DNA (theoretically $>10^7$ times denser than in state-of-the-art hard drives)⁷ can be reliably read, erased and copied. Complex synthetic machinery ensures that DNA can be replicated with a very low error rate and transcribed to RNA for further translation into proteins.

Once again, the key concept behind the effectiveness of biopolymers is the fact that fine sequence regulation confers on them specific 3D structures responsible for their function. In other words, Nature achieves crucial processes like energy-effective and selective catalysis, accurate molecular recognition or dense data storage with decoding and copying properties using biopolymers with specific sequences. Can we apply Nature's sequence control strategy to deal with 21st century challenges in chemistry?

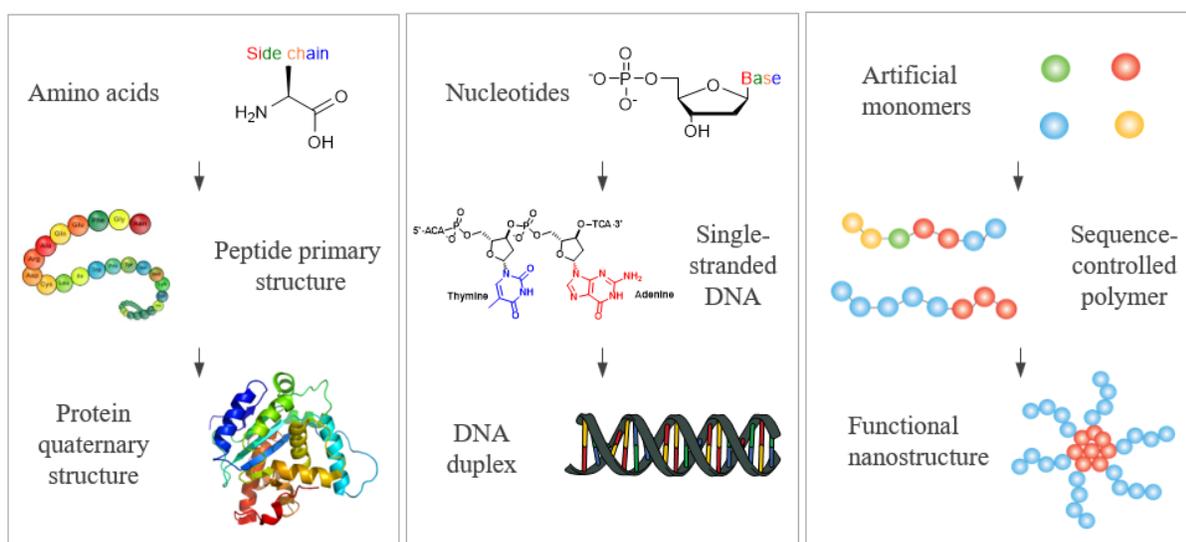


Figure 1.1. Sequence leads to 3D structure that induces function.

In this introduction, some of the most advanced applications with nucleic acids as an example of biopolymer will be described. Challenges due to the structural limitations of nucleic acids will be discussed. In a second section, the most recent synthetic routes towards precision polymers with advanced to total sequence-control will be broadly covered. Applications with these polymers are numerous and most of the state-of-the-art advances in the field are described in the third section. Finally, the scope of this thesis in the context of sequence-controlled polymers is discussed.

1.2 Repurposing biopolymers

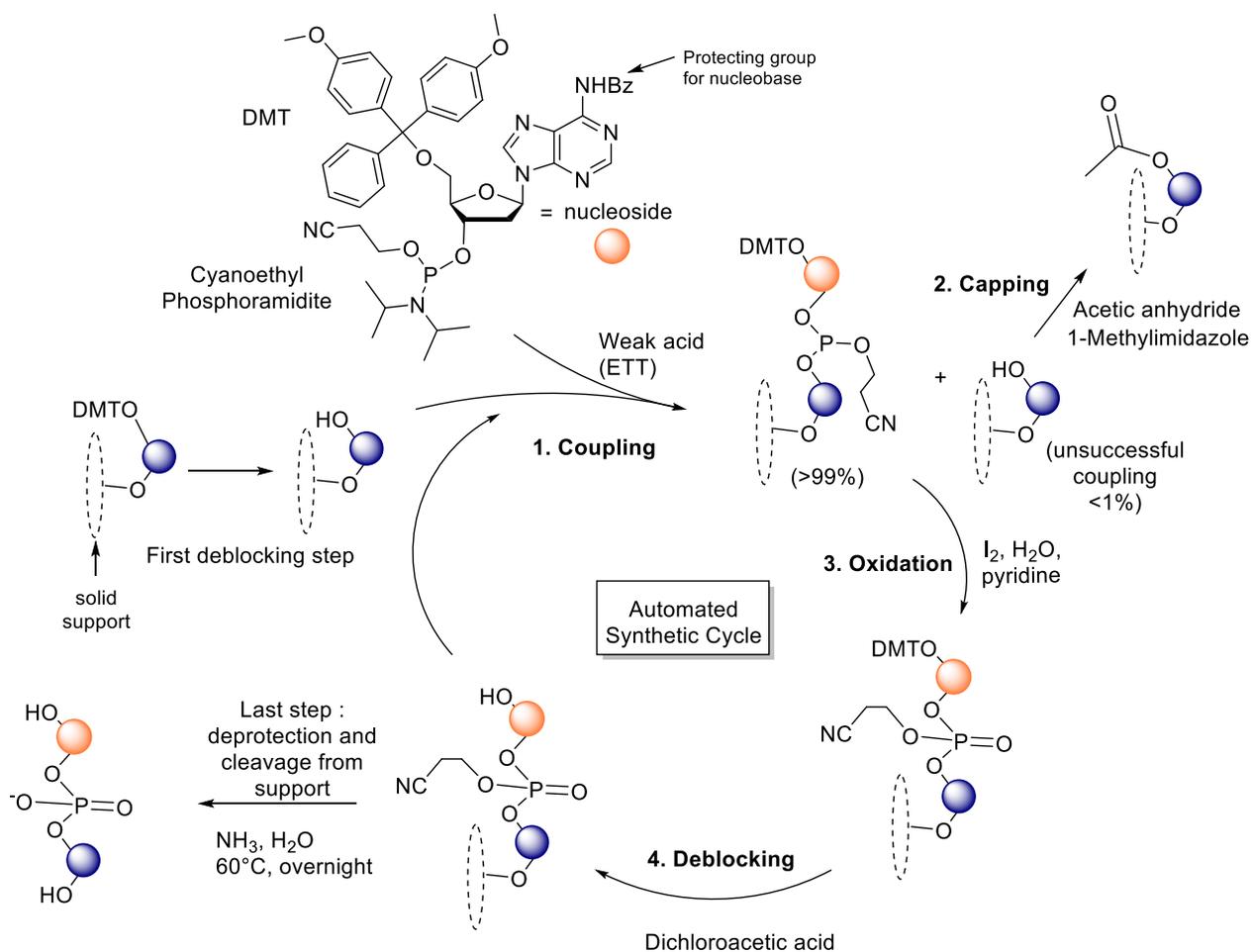
Since Nature has been developing its own biopolymers, the simplest approach to access functional macromolecules is to use them. The following section specifically provides an overview of the uses and limitations of natural and modified nucleic acids. Indeed, nucleic acids have a broad range of applications in data storage, catalysis, therapeutics and molecular recognition. The two latter applications are more broadly covered to facilitate the understanding of the other chapters of this thesis.

1.2.1 DNA synthesis

1.2.1.1 Solid-Phase Synthesis

Ease of synthesis of a class of molecules has always been closely related with the number and importance of applications reported. When Merrifield came up with the idea of solid-phase synthesis (SPS),⁸ the development of peptide- and DNA-based technologies started to thrive. SPS relies on the use of solid polystyrene beads or controlled-pore glass (CPG) support on which biopolymers can be grown using an iterative strategy shown in Scheme 1.1. As poly(phosphodiester)s, the building blocks for DNA synthesis were designed using organophosphorus chemistry. Phosphoramidites have been found by Beaucage and Caruthers to be efficient to make synthetic oligonucleotides.⁹ The diisopropylamine group on the phosphorus center is labile in the presence of a weak acid, enabling the efficient nucleophilic attack of an alcohol. This is the coupling reaction, first step of the DNA SPS synthetic cycle. At each step, excess reagents and byproducts in solution are washed away in a simple filtration purification step. Since coupling efficiency is not 100 %, coupling failure can occur at each step, leading to a complex mixture of side products. When a coupling is unsuccessful, the capping step prevents

further couplings. The oxidizing step is required to oxidize phosphorus(III) to more stable phosphorus(V). To allow single-unit insertion, the 5'-hydroxyl group of nucleotide building blocks is protected with a dimethoxytrityl (DMT) moiety. It provides the advantage of being labile in acidic conditions, forming a strong orange DMT cation detectable with a photodetector. The deprotected 5'-hydroxyl group is ready for a new coupling step at this point and a new cycle can start. When the designed oligomer length is reached, the cyanoethyl protection on the phosphodiester, as well as the necessary protecting groups on DNA bases, are cleaved during the final deprotection step, where the oligomer can be recovered free in solution.



Scheme 1.1. Synthetic cycle in an automated DNA synthesizer. ETT = Ethylthiotetrazole, DMT = dimethoxytrityl. The yields for the coupling step need to be over 99 % to ensure that long sequences can be recovered. For a 20mer, if coupling yields were 90 %, global yields would only be $(90 \%)^{20}=12 \%$.

Solid-phase allows complete control over the sequence produced. The broad number of applications for peptides and oligonucleotides has led to a large decrease in cost for this process

(DNA is now ~2\$/mg on average).¹⁰ However, the main drawback of solid-phase synthesis is the inability to make very long sequences. They have a maximum length of about 200mers for DNA and not more than 50mers for peptides. Therefore, chemical ligation and enzymatic methods are also broadly applied to access longer DNA, peptides and proteins.

1.2.1.2 Enzymatic synthesis

DNA strands in the range of 1 to 30 kbp can be made enzymatically through the use of shorter synthetic oligonucleotides assembled together with ligation enzymes or through multiple recombination steps in yeast.¹¹⁻¹³ These methods, however, are usually time-intensive (~week scale) to perform for strands of more than 1kb. For longer strands (~100-1000 kbp), natural bacterial genomic DNA can be useful for many applications like protein production.¹⁴ Recombinases and genome editing machineries such as zinc finger nucleases (ZFN), transcription activator-like effector nucleases (TALEN) and the more broadly applicable clustered regularly interspaced short palindromic repeats (CRISPR-Cas) are used to finely modify DNA sequences.¹⁵ Examples of fully tailored long DNA sequences have also been reported. For example, a 1Mbp long fully artificial genome was synthesized and assembled.¹⁶

1.2.2 Applications of Nucleic acids

1.2.2.1 Digital storage with DNA

Scientists have tried to repurpose DNA to store non-genetic information. Indeed, due to its long-term, high-density storage abilities along with the capacity to be readily sequenced, DNA has naturally been used to write and store data. For example, Church and coworkers encoded a whole book representing more than 5 Mb in DNA and showed it was easily read through next-generation sequencing.¹⁷ More recently, the Ceze and Strauss groups demonstrated the encoding of 200Mb of information in 35 distinct files.¹⁸ Using a specific library of primers, they could “read” the different files according to a similar principle as random access memories. While significant advances in this field have been realized, the synthesis time, cost and yield issues associated with long DNA strands are limiting factors. Moreover, DNA only allows information storage in base 4 (because there are four possibilities – A, T, C or G – at each position). To avoid the synthesis of long sequences, it would be valuable to write information in base n (where $n > 4$) in a fast and

inexpensive manner. Unnatural polymers with well-defined sequences could represent such an alternative.¹⁹

1.2.2.2 DNA Nanotechnology

Seeman envisioned the possible use of DNA strands as building blocks to make 2-dimensional and 3-dimensional objects through self-assembled branched junctions.²⁰ He showed the synthesis of such a junction with chemically synthesized oligonucleotides in 1983.²¹ This idea started the field now known as DNA nanotechnology. 2D and 3D extended networks as well as discrete structures of arbitrary shapes have been assembled.^{22,23} Among them, the DNA origami method relies on folding a long viral DNA strand into highly programmable shapes through the use of multiple short staple strands.²⁴ Recent advances show the possibility of making a micrometer-sized Mona Lisa portrait made of DNA (**Figure 1.2**).²⁵ Other groups have taken inspiration over decades of supramolecular chemistry to pioneer the expansion of DNA nanotechnology to different types of interactions.²⁶ For instance, they paved the way to the use of small molecule geometry inducers,^{27,28} metal-DNA junctions²⁹ (**Figure 1.2**) as well as DNA amphiphiles.³⁰ The latter are polymers with a hydrophilic section made of DNA and a hydrophobic part made of small molecules or polymers. The hydrophobic effect triggers new modes of assembly of DNA nanostructures. For example, cylindrical and spherical micelles with a DNA corona,³¹ and DNA nanocubes with a hydrophobic core capable of drug encapsulation (**Figure 1.2**)³² have been reported.

Many applications have resulted from these studies. For example, the high degree of precision obtained with DNA origami allows the formation of “nanorulers” –distance measurement standards at the nanoscale.³³ Gold nanoparticles were encapsulated in a linear DNA nanotube and released only when a specific DNA strand was added.³⁴ Many biomedical applications,³⁵ including DNA nanostructures that encapsulate nucleic acid therapeutic cargo have also been reported.³⁶

As a conclusion, with only 4 different building blocks, DNA has been shown to be a highly programmable material. DNA nanostructures are now easier to design than proteins due to the small numbers of rules and interactions governing the assembly. However, the limited number of building blocks also reduces function in DNA as compared to proteins, for example. To introduce functionality in this biomolecule, artificial building blocks need to be added to DNA.

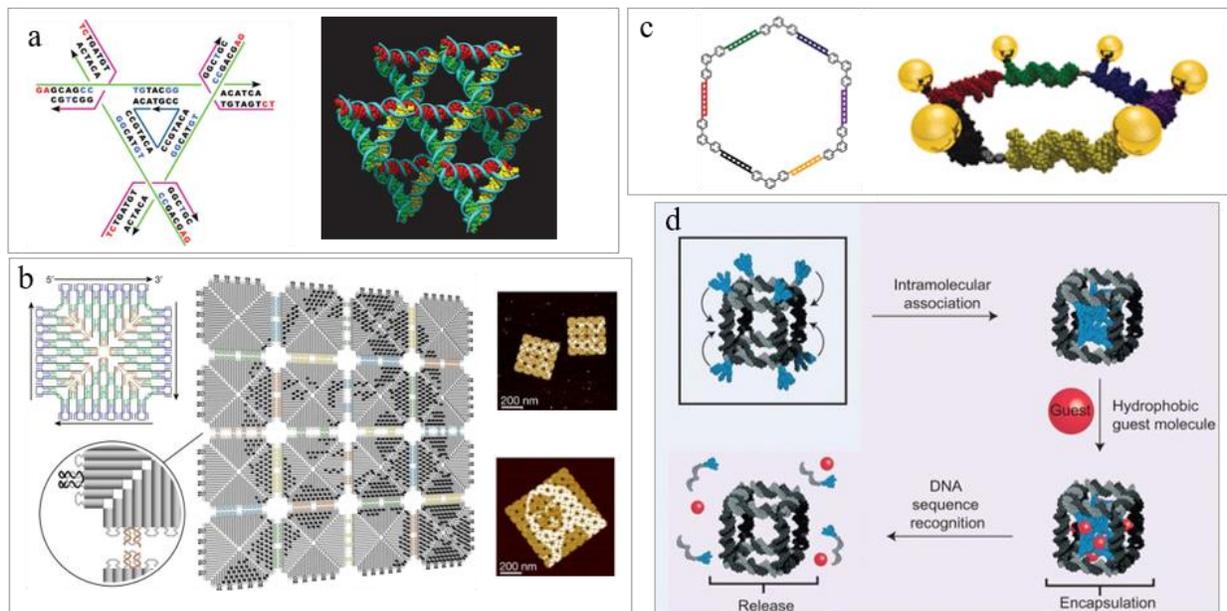


Figure 1.2. Examples of DNA nanostructures. (a) DNA tensegrity triangles 3D crystals. Reproduced with permission from ref. 23. (b) Fractal-assembly of DNA origami. Reproduced with permission from ref. 25. (c) Gold nanoparticle decorated DNA-hexagon. Reproduced with permission from ref. 26. (d) DNA nanocube with a hydrophobic core. Reproduced with permission from ref. 32.

1.2.2.3 Gene silencing

Oligonucleotides have recently emerged as highly promising therapeutics. Partially complementary sequences to messenger RNA (mRNA) called antisense oligonucleotides (ASO) can repress gene expression (**Figure 1.3**).³⁷ They are generally 12 to 25 nucleotides long single-stranded DNA or RNA that recognize specific mRNA, leading to their cleavage through an enzymatic process.³⁸ Similar size double-stranded RNA oligonucleotides called small interfering RNA (siRNA) can also reduce gene expression. They act through a different mechanism. Once in cells, they form an RNA-induced silencing complex (RISC) together with several proteins. This complex is capable of mRNA recognition and degradation resulting in reduced gene expression. The therapeutic potential of gene silencing oligonucleotides is theoretically immense since they can trigger specific gene knock-down.³⁸ However, only a few gene silencing oligonucleotide drugs have made it to the market so far.³⁹ The Foods and Drugs Administration (FDA) approved the first ASO drug in 1998⁴⁰ and the first siRNA in 2018.⁴¹ The two approved oligonucleotides have a point in common: they both contain non-natural modifications. As such, they are called xeno-nucleic acids (XNA) or genetic polymers.⁴²

Indeed, the main limitation of oligonucleotides for gene silencing is their poor cell-membrane permeability and susceptibility to nuclease degradation. Extensive research has been carried out to address these drawbacks, by using drug delivery vehicles⁴³ or chemical modifications of the strands (**Figure 1.3**).⁴⁴ In the latter case, the main challenge⁴⁴ is to increase cell internalization of the oligonucleotide and to avoid nuclease degradation. However, it should still bind the mRNA of interest and be recognized by the silencing enzymatic machinery. Many base, sugar and backbone modifications have been developed. Representative examples are presented in Figure 1.3. Specific positioning of several of these novel monomers is usually necessary for efficient gene silencing. So far, rational design of silencing oligonucleotides is difficult. Therefore, the choice of modifications and positions requires tedious testing of numerous oligonucleotides.⁴⁴

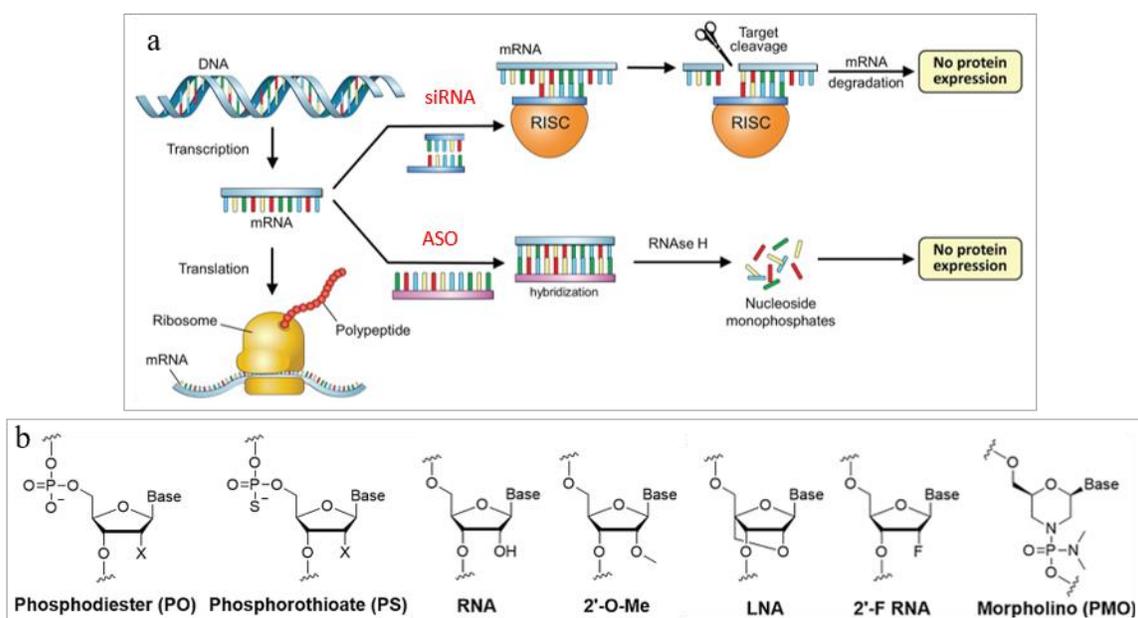


Figure 1.3. Gene silencing mechanism and chemical modifications in siRNA and ASO. (a) Different mechanisms induce reduction of protein expression with siRNA and ASO. Adapted with permission from ref. 45. (b) Some of the most common modifications in ASO and siRNA in major clinical trials. LNA stands for locked nucleic acids and PMO for phosphorodiamidate morpholino oligomer. Adapted with permission from ref. 39.

1.2.2.4 Nucleic acid aptamers

Nucleic acid aptamers are DNA and RNA strands capable of binding to a specific molecular target, such as a protein or small molecule. They usually fold into well-defined structures leading to high binding affinity to the molecule of interest. Their applications range from biosensing to diagnostics and targeted therapies.⁴⁶ DNA and RNA aptamers were discovered through the combinatorial

synthesis of oligonucleotide libraries. In a first step, these undergo a selection step against targets such as proteins. The binding oligomers from the library are then isolated and subjected to amplification. At this point, mutations are introduced in the sequences obtained in order to improve the binding affinity to the target in a stepwise manner (**Figure 1.4**). This process is called systematic evolution of ligands by exponential enrichment (SELEX) or *in vitro* selection and was developed independently by two groups in the early 1990s.^{47,48} Since then, potent binders to many biologically relevant targets were discovered.⁴⁶ Similar to silencing oligonucleotides, aptamers are susceptible to nuclease degradation. Moreover, aptamers are limited in the type of interactions they create with their target because they are made of only 4 structurally similar building blocks (ATGC or AUGC). This partly explains why only one aptamer drug has been approved by the FDA so far.⁴⁶ Accessing similar or broader functionalities than in antibodies made of 20 amino acids would palliate this issue. Therefore, chemical modifications have been introduced.

Contrary to gene-silencing applications, aptamers do not need to be recognized by some enzymatic machinery apart from the polymerase used in the SELEX process. This allows the use of more numerous modifications or components that are less structurally similar to natural nucleotides (**Figure 1.4**).^{49,50} As an example, the Gold group developed slow off-rate aptamers (SOMAmers) where a few nucleotides were decorated with amino-acid like moieties to resemble antibodies (**Figure 1.4**).⁵¹ Similarly, Mayer and coworkers developed click-SELEX.⁵² In this process, 5-ethynyl-deoxyuridine is used instead of thymidine in DNA libraries generation. Before selection, the alkyne-containing nucleobase is modified through copper-catalyzed azide-alkyne cycloaddition (CuAAC). After selection, the modification is removed during the amplification process and added again for the next SELEX cycle. The Liu and Hili groups developed the synthesis of highly modified nucleic acids that can bind to specific targets.^{53,54} They made short modified codons that hybridize on a regular DNA template and are further enzymatically ligated (**Figure 1.4**). The modified strand is then separated from the template and tested against a target of interest. Modifications not only increased the half-life of the oligomers in serum but also led to higher binding affinities than with unmodified oligonucleotides. Increasing the number of available monomers in aptamers have made them closer to antibodies in terms of binding affinity reached.⁵⁰ These studies pave the way to the use of even more structurally different oligomers for specific binding to biologically relevant targets.

reactions such as amino acids condensation in the ribosome.⁴ Apart from this reaction, the reactions performed by ribozymes *in vivo* involve the cleavage or formation of a phosphodiester bond.⁴ SELEX was also applied to ribozymes. With this tool, a number of artificial catalysts made of RNA and DNA (DNAzymes) were discovered and shown to catalyze diverse reactions.⁵⁶ For example, a Diels–Alder reaction of anthracene derivatives with maleimides catalyzed by RNA,⁵⁷ a peroxidase-like DNAzyme,⁵⁸ an RNA phosphorothioate bond cleavage using a DNAzyme⁵⁹ were reported.

Similarly to aptamers and gene silencing, the use of unnatural nucleotide-like modifications allowed the engineering of novel DNAzymes.⁴ For example, including amino-acid like moieties on deoxyuridine allowed the discovery of a DNAzyme capable of aliphatic amide cleavage.⁶⁰ From a different perspective, Holliger and coworkers developed a reverse-transcriptase and a polymerase compatible with multiple types of XNAs.⁶¹ This discovery allowed the direct evolution of XNA through a modified SELEX process called cross-chemistry selective enrichment by exponential amplification (X-SELEX).⁶² This process is similar to SELEX but cycles start with the synthesis of a library of XNAs from a library of regular DNA sequences. After the selection of potent XNAzymes, a reverse-transcriptase is used to make complementary regular single-stranded DNA and proceed to sequencing or mutagenesis before the next cycle. For example, XNAzymes with four different types of chemistry were found to have RNA endonuclease activity.⁶³ These findings further demonstrate how exploring a larger chemical space broadens the scope of nucleic acid applications.

1.2.3 Conclusions of Section 1.2

In general, novel methods to obtain synthetic nucleic acids allow their use for a variety of applications ranging from data storage to molecular recognition, drug discovery and catalysis. Chemical modifications of natural molecules are key to overcoming inherent biopolymer limitations such as nuclease susceptibility. Shifting from biopolymer-like structures to fully artificial polymers with specific sequences can pave the way to further improvements and new applications.

The design of such polymers is a difficult task since synthetic chemistry offers a large number of possible polymeric backbones, lengths and sequence control. Studying a biopolymer like DNA

and its most recent applications helped us extract the main criteria for making the most efficient oligomers.

- (i) *Achieving total sequence control.* Indeed, a single DNA mutation can have a major impact on the related protein efficiency, leading to serious genetic diseases. Similarly, an aptamer with one mutation can have a drastically lower affinity for its target.
- (ii) *Accessing the right degree of polymerization.* For example, the degree of polymerization (DP) is an essential parameter for information storage in DNA.
- (iii) *Expanding the chemical diversity of monomers.* The DNA nanotechnology field shows how only four monomers can lead to structurally diverse nanoobjects. However, the addition of different materials (hydrophobic polymers or nanoparticles, for instance) and monomers (metal-binding ligands for instance) were required to introduce functionality. Similarly, aptamers half-life and affinity for biological targets was improved with the addition of new monomers.

1.3 Synthetic routes towards sequence-controlled polymers made of unnatural building blocks

The terminology “sequence-controlled polymer” is used for a polymer for which the order of monomers is controlled to some degree (**Figure 1.5**). Pioneers in the introduction of sequence control in classical polymers, Mayo and Lewis synthesized polymers with two alternating monomers in 1944.⁶⁴ In block copolymers, polymeric segments of different compositions are attached together to form one chain. This type of polymer appeared later in the work of Melville⁶⁵ and then Szwarc⁶⁶ but they are now broadly studied. They are used in a variety of applications such as drug delivery, nanolithography, photoactive structures and porous materials.^{67,68} Continuously varying the chemical composition along a polymer chain is also possible and such macromolecules are called gradient polymers.⁶⁹ Despite the immense potential of such control in polymerization, we will here focus on more advanced sequence control as it paves the way to more functional macromolecules.^{70–72} This includes periodic copolymers where a complex sequence of monomers is repeated. Chain positioning polymers feature regions along the chain that are

chemically functionalized with specific monomers. Polymers for which sequence control is total down to the last monomer, like biopolymers, are called sequence-defined and will be covered in greater detail.

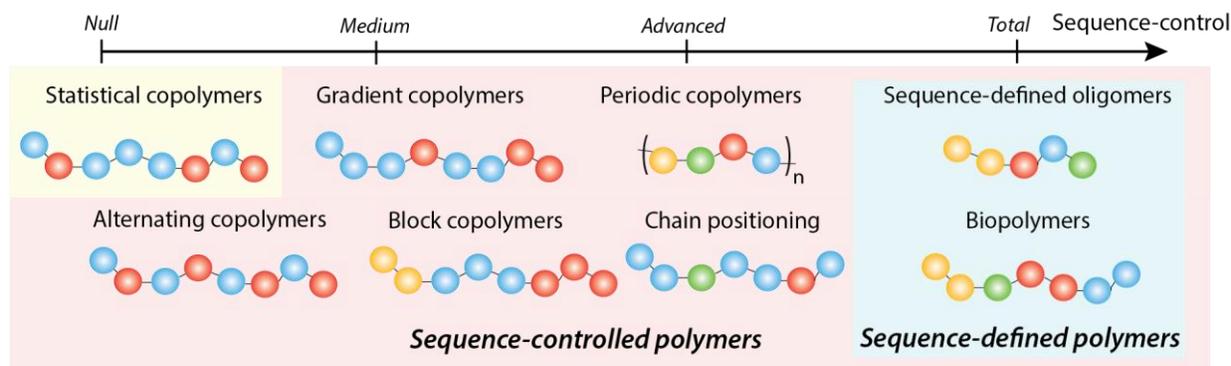


Figure 1.5. The different sequence-controlled polymers. Figure inspired from ref. 71.

1.3.1 Polymerization of sequence-defined macromonomers

It was shown that a sequence-defined “macromonomer” (*i.e.*, a monomer unit that is itself comprised of multiple parts) could be used as the repeat unit of a periodic copolymer (**Figure 1.6**). Wagener and coworkers pioneered this approach using step-growth polymerization. They controlled the position of ethyl branches⁷³ and charged 1-methylimidazolium groups by using aliphatic chains of predetermined length.⁷⁴ Similarly, the Lutz group made oligostyrene chains containing one functional group of interest and polymerized these “oligomonomers” through CuAAC.⁷⁵ To precisely tune the stereochemistry of some poly(lactic-co-glycolic acid), Meyer and coworkers prepared “segmers”, made of 4 monomers with perfect sequence control, using conventional chemistry.⁷⁶ They polymerized the “segmers” through step-growth polymerization. Kamigaito and coworkers showed the preparation of complex monomers made of multiple functional parts (**Figure 1.6**). Two of them have a vinyl and a chloride used for metal-catalyzed radical polymerization.⁷⁷ The macromonomer can also be generated *in situ* in multicomponent polymerization to implement periodic sequence control.⁷⁸ Using SPS or classic organic chemistry, several recent reports have shown the synthesis of sequence-defined macromonomers followed by step-growth polymerization.^{79–83} It is advantageous in terms of molecular weights achieved but can lead to high polydispersity indexes.

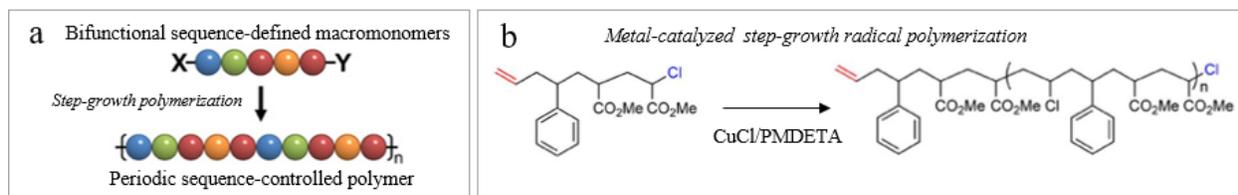


Figure 1.6. Step-growth polymerization with sequence-defined macromonomers. (a). Principle of step-growth polymerization for periodic polymers. Adapted with permission from ref 80. (b) Metal-catalyzed step-growth radical polymerization. Adapted with permission from ref 77.

Chain-growth polymerization approaches have also been employed with sequence-controlled “macromonomers”. The Ouchi and Sawamoto groups showed the radical polymerization of chemically-templated monomers (**Figure 1.7**).^{84,85} The template could be removed after the polymerization, leading to ABA or AB periodic polymers. More recently, Kamigaito and coworkers used the atom transfer radical addition (ATRA) reaction to make sequence-regulated oligomonomers that they polymerized in a radical process.⁸⁶ Ring-Opening Polymerization (ROP) has been extensively studied to introduce sequence-defined monomers into polymers. The Hillmyer group first showed the synthesis of sequence-regulated vinyl-copolymers using a cyclooctene in Ring Opening Metathesis Polymerization (ROMP).⁸⁷ The cyclooctadiene was decorated with three different moieties. The use of a second or third generation Grubb’s catalyst allowed regiospecificity of the monomer introduction and was key to accurate sequence control. Hawker and coworkers adapted this method to unstrained macrocycles (**Figure 1.7**).⁸⁸ More recently, the difference of reactivity of *cis*- and *trans*-olefins has been explored by the Meyer group allowing the polymerization of a 9-monomer-long repetition motif.⁸⁹

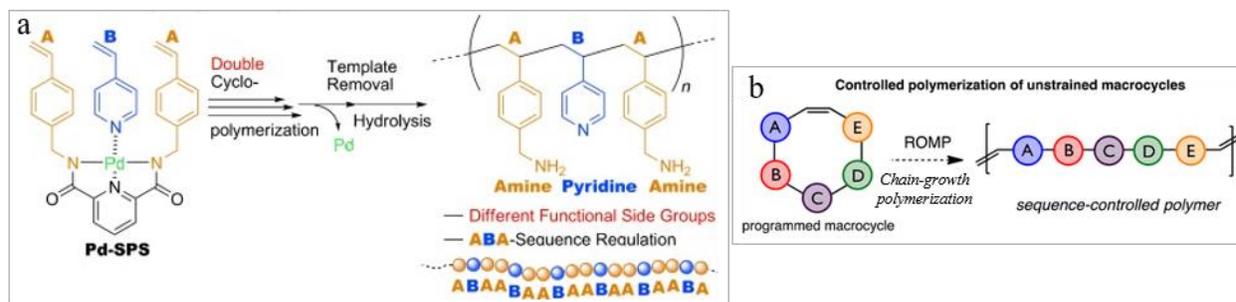


Figure 1.7. Chain-growth polymerization with sequence-defined macromonomers. (a). Palladium-templated macromonomer for the synthesis of periodic polymers. Adapted with permission from ref 84 (b) ROMP for the synthesis of periodic copolymers, Adapted with permission from ref. 88.

1.3.2 Timely incorporation of monomers in a polymer chain

A valuable strategy for positioning monomers in a polymer chain using atom transfer radical polymerization (ATRP) has been developed by the Lutz group in 2007 (**Figure 1.8**).⁹⁰ Their strategy relies on the use of two types of co-monomers, electron donors and acceptors. The first, usually styrene, is the main constitutive element of the polymer chain while the other, maleimide, is added at specific moments during the polymerization. Due to the opposite electronic demand of the two double bonds, maleimide and styrene copolymerize quickly. Once all maleimides are consumed, the polymerization resumes with the remaining styrene molecules. This strategy led to polystyrene with small regions tagged with a maleimide that contains chemical information. The Lutz group further studied the timely controlled incorporation of special monomers in a polymer chain.¹⁹ They notably developed a library of 20 different maleimides that are compatible with this method⁹¹ and subsequently automated the process.⁹² The region where the maleimide monomer can be introduced is not perfectly well-defined due to the probabilistic nature of the polymerization process. Indeed, even if the active site would preferentially catch a monomer with an inverse electron demand, there is still the possibility that it reacts with similar electron demand monomer. This drawback was partly addressed by starving the polymerization from the donor monomers before each maleimide insertion,⁹³ enhancing the location accuracy of maleimides along the chains. While this method relies on the manual incorporation of a fast-adding monomer, stimuli-responsive polymerization is of interest. For example, light-induced timely incorporation of another type of monomer was achieved in homopolymers made through anionic ring-opening polymerization (AROP) (**Figure 1.8**).⁹⁴

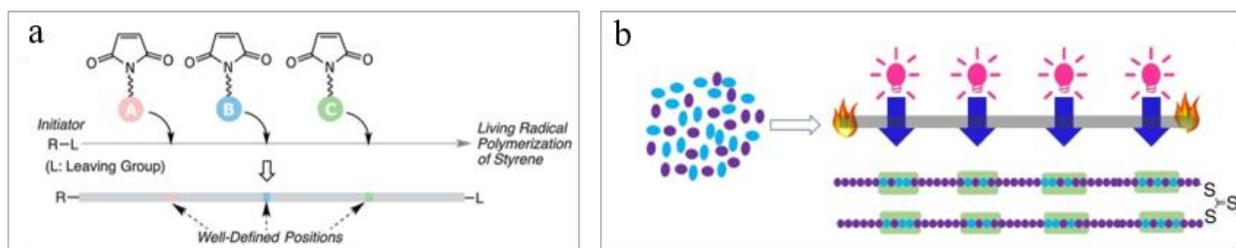


Figure 1.8. Accurate positioning of monomers in chain-growth polymerization. Timely incorporation of fast-adding maleimides in a polystyrene chain. Adapted with permission from ref 70 (b) Photoinduced electron transfer – radical addition-fragmentation chain transfer (PET-RAFT) addition of an acrylate monomer during the homopolymerization of a thiocycle containing monomer through AROP. PET-RAFT is light-responsive while AROP is thermo-responsive Adapted with permission from ref. 94.

1.3.3 Templated synthesis

1.3.3.1 DNA-templated synthesis

DNA-templated synthesis is a strategy where hybridization of an oligomer to a DNA template would allow a reaction to occur between two moieties in close proximity. The inspiration for nucleic acid templated synthesis comes from natural biopolymer synthesis. Polymerases are able to synthesize a new DNA strand from nucleoside triphosphates and a DNA template. Similarly, proteins are synthesized in a ribosome following an RNA template. DNA-templated synthesis can be parallel, sequential or sequential autonomous as described on Figure 1.9.⁹⁵

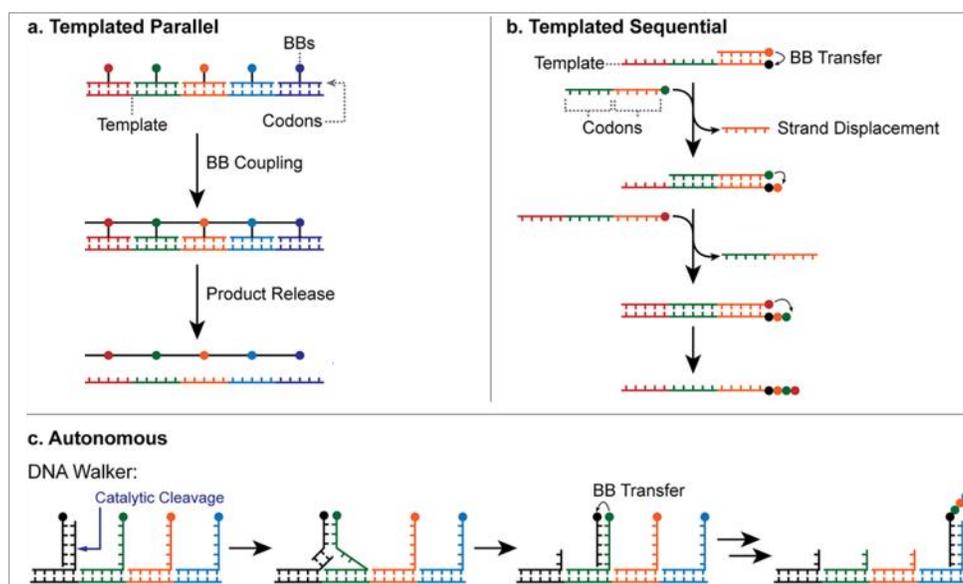


Figure 1.9. DNA-templated synthesis of sequence-defined oligomers. BB stands for building block. (a) DNA-templated synthesis with polymerization occurring in parallel or (b) in a stepwise manner. (c) Example of an autonomous sequential polymerization method with a “DNA-walker” Adapted with permission from ref. 95.

The parallel strategy consists in templating and reacting all monomers at the same time. The synthesis of Peptide-Nucleic Acid (PNA), a nucleic-acid like polymer with a polyamide backbone, through DNA-templated synthesis was first reported by Orgel and coworkers in 1995.⁹⁶ Similarly, short polyamine tetramers with thymidine and adenosine were synthesized in a perfectly sequence-defined fashion using a DNA-template.⁹⁷ The Liu group then reported the DNA-templated synthesis of a PNA strand modified with non-nucleosidic side-chains.⁹⁸ In a more recent report, they succeeded in synthesizing a sequence-defined 16mer structurally different to nucleic acids using a templated parallel approach (**Figure 1.10**).⁹⁹ This strategy was based on the attachment of

small PNA-adapters along a DNA strand. These PNA adapters bore di-alkyne or di-azide synthetic monomers. Once all the adapters were positioned on the template, CuAAC was used to attach two monomers in proximity. In a last step, the newly formed oligomer was released from the PNA adapters and the DNA template. This translation system is promising for the synthesis of fully artificial sequence-defined oligomers. Another strategy involves attaching monomers directly to DNA bases. For example, Schuster and coworkers modified bases with aniline and 2,5-bis(2-thienyl)pyrrole, and formed cyclic and linear conducting copolymers upon DNA strands assembly.¹⁰⁰ Sequence-controlled oxidative ligation with horseradish peroxidase was then applied to attach the monomers together.

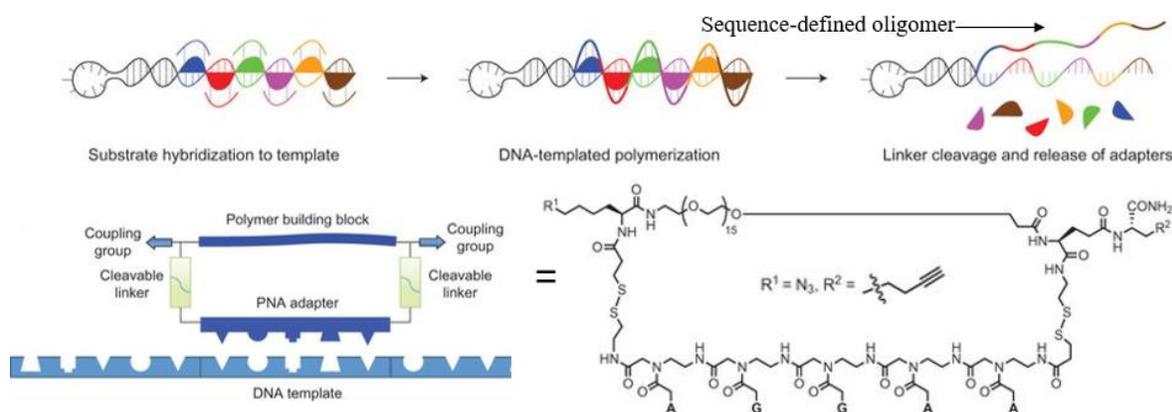


Figure 1.10. Parallel DNA-templated synthesis using PNA adapters. Adapted with permission from ref. 99.

Sequential DNA-templated synthesis wherein monomers are attached one after the other (**Figure 1.9**) have been reported and can involve Wittig reactions,¹⁰¹ amine coupling^{102–104} and CuAAC¹⁰⁵. So far, these methods led to a maximum degree of polymerization of 6.¹⁰³ Applying similar processes repetitively, O’Reilly, Turberfield and coworkers could make longer sequences at the expense of sequence control with a A(BC)₄D oligomer.¹⁰⁶ The sequential aspect of this process can be tedious for longer oligomers and two autonomous strategies have been reported. The first one is based on enzymatic cleavage reactions to make a “DNA walker” that moves along a track made of DNA (**Figure 1.11**).¹⁰⁷ N-hydroxysuccinimide (NHS) ester-amine couplings occur between the different monomers placed along the track. In another report, O’Reilly, Turberfield and coworkers use the hybridization chain reaction where hairpins unfold to make duplex DNA in a cascade reaction.¹⁰⁸

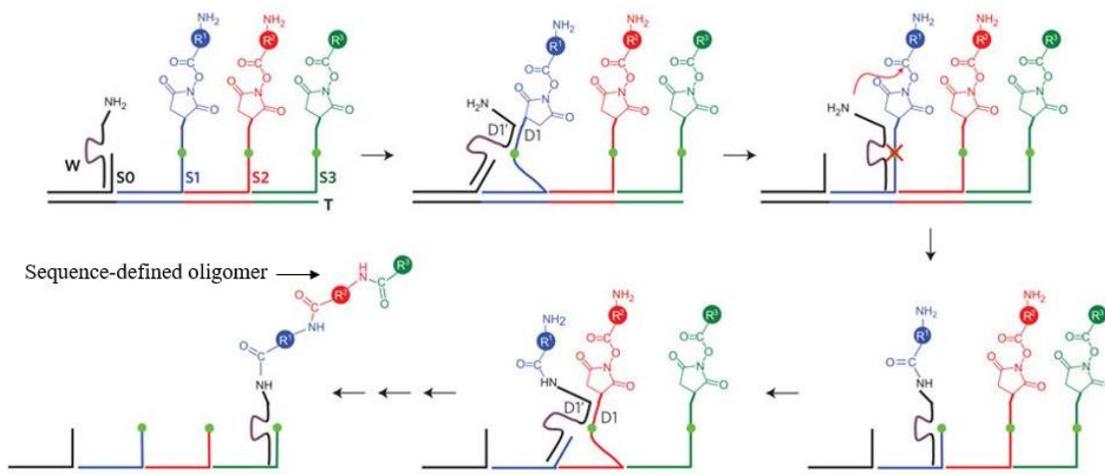


Figure 1.11. Sequential autonomous DNA-templated synthesis using NHS ester-amine couplings. Adapted with permission from ref. 107.

1.3.3.2 Non-DNA templates

The template for sequence-defined polymers is not necessarily DNA or RNA. Recently, Leigh and coworkers succeeded in building a molecular machine that synthesizes sequence-defined oligoamides (**Figure 1.12**).^{109,110} To achieve this, a pseudo-rotaxane molecule was synthesized, where monomers were spatially distributed along the thread and the sequence-defined oligomer was built as the macrocycle de-threaded, leading to a sequential reaction with each precisely-placed monomer. This example shows how autonomous molecular machines may be designed for the synthesis of artificial sequence-defined oligomers.

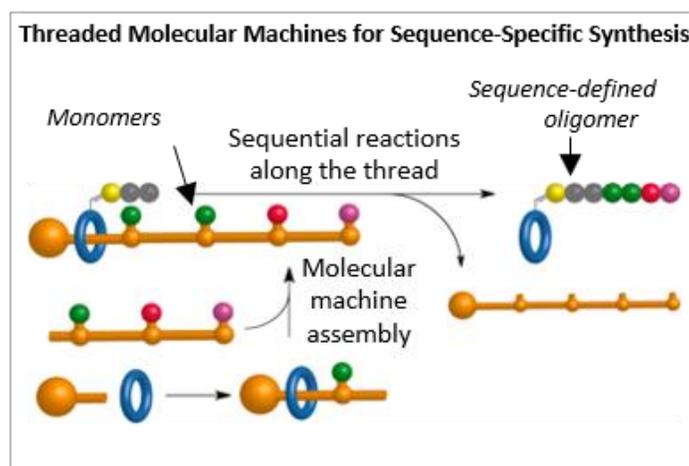


Figure 1.12. Molecular machine performing sequence-defined synthesis along a rotaxane thread. Adapted with permission from ref 110.

In general, templated strategies are valuable to achieve perfect sequence control and make sequence-defined oligomers. However, nucleic acid templates may only be used in water, no oligomers with DP>10 have been made and artificial templates require involved organic synthesis.

1.3.4 Iterative strategies

The method of choice to make oligonucleotides and peptides is an iterative strategy, wherein each monomer is added one after the other. This naturally leads to perfect sequence control and has been explored by scientists through a variety of polymerization methods and reactions.

1.3.4.1 Chain growth

Chain-growth polymerization tools can be used in an iterative way to attach one monomer at a time on a growing polymer chain. This strategy was first reported in 1990 with the synthesis of sequence-defined trimers and tetramers through living cationic polymerization.¹¹¹ The key challenge is single unit monomer insertion (SUMI) at every step. Atom transfer radical addition (ATRA) was used with low-activity allyl alcohol.¹¹² In this case, SUMI occurred thanks to the low reactivity of the adduct made. The latter had to be further oxidized to be functionalized and reactive for subsequent ATRA. In another report, sequence-defined acrylate trimers were made through photoinduced copper-mediated radical monomer insertions.¹¹³ In this example, addition of a monomer only occurs under the right light irradiation. However, purification after each monomer insertion remained necessary. Sawamoto, Ouchi and coworkers used a sterically hindered acrylate monomer to avoid the insertion of several units at once (**Figure 1.13**). A bulky adamantane group had to get cleaved to proceed to further steps.¹¹⁴ The same group also elegantly achieved SUMI via cyclization of the reactive radical into a ‘closed’, unreactive structure.¹¹⁵ However, only short dimers or trimers could be produced through these methods.

RAFT polymerization has also been investigated by several groups and showed great promise for making longer oligomers. The single monomer insertion problem was overcome by the Junkers group by purifying the sequence-defined oligomer after each step with an in-line SEC purification system.¹¹⁶ Moad and coworkers developed scalable methods for single monomer insertion into RAFT agents.^{117,118} This work led to the synthesis of trimers using a highly selective PET-RAFT system.¹¹⁹ The authors further optimized this method to make stereospecific pentamers using indenones and *N*-substituted maleimides as monomers (**Figure 1.13**).¹²⁰

As a conclusion, chain-growth polymerization holds great promise for the high-scale synthesis of discrete oligomers with a variety of side-chains. However, the low degree of polymerization achieved remains a limitation to overcome.

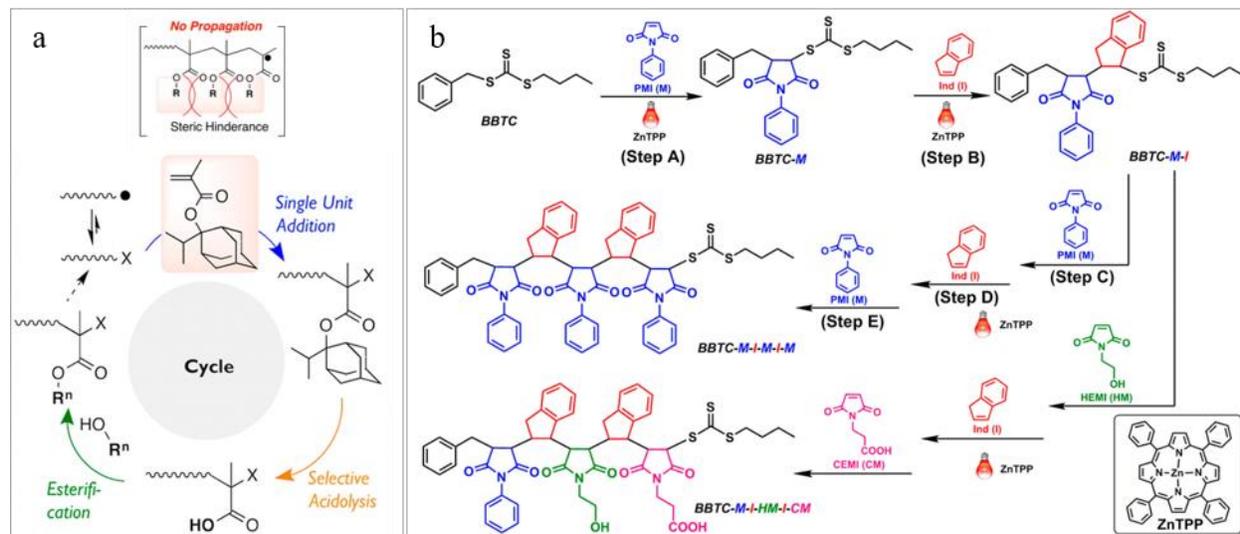


Figure 1.13. Chain-growth stepwise polymerization through SUMI. (a) Bulky substituents prevent growing chain from propagation in metal-catalyzed living radical polymerization. Adapted with permission from ref. 114. (b) Alternating monomers electron demand allows for RAFT SUMI. Adapted with permission from ref. 120.

1.3.4.2 Solid-Phase Synthesis

Peptides and DNA strands containing less than 50 amino acids and 200 bases respectively are mostly made through solid-phase synthesis. For biopolymers, this process affords perfect sequence control, large monomer versatility (20 amino acids) and substantial polymer length (about 50mers for peptides and 200mers for oligonucleotides). Moreover, the process is fast, reliable thanks to the introduction of automated synthesizers and scalable.¹⁰ Therefore, solid-phase strategy was investigated for making artificial sequence-defined oligomers. A number of different types of chemical reactions combined with SPS were investigated. We counted three main coupling strategies that are detailed further thereafter.

The first SPS strategy relies on one sort of attachment chemistry (**Figure 1.14**). For example, one monomer end is reactive (phosphoramidite/carboxylic acid) and ready to attach to a growing polymer chain while the other one is protected (DMT/Fmoc) or unreactive.

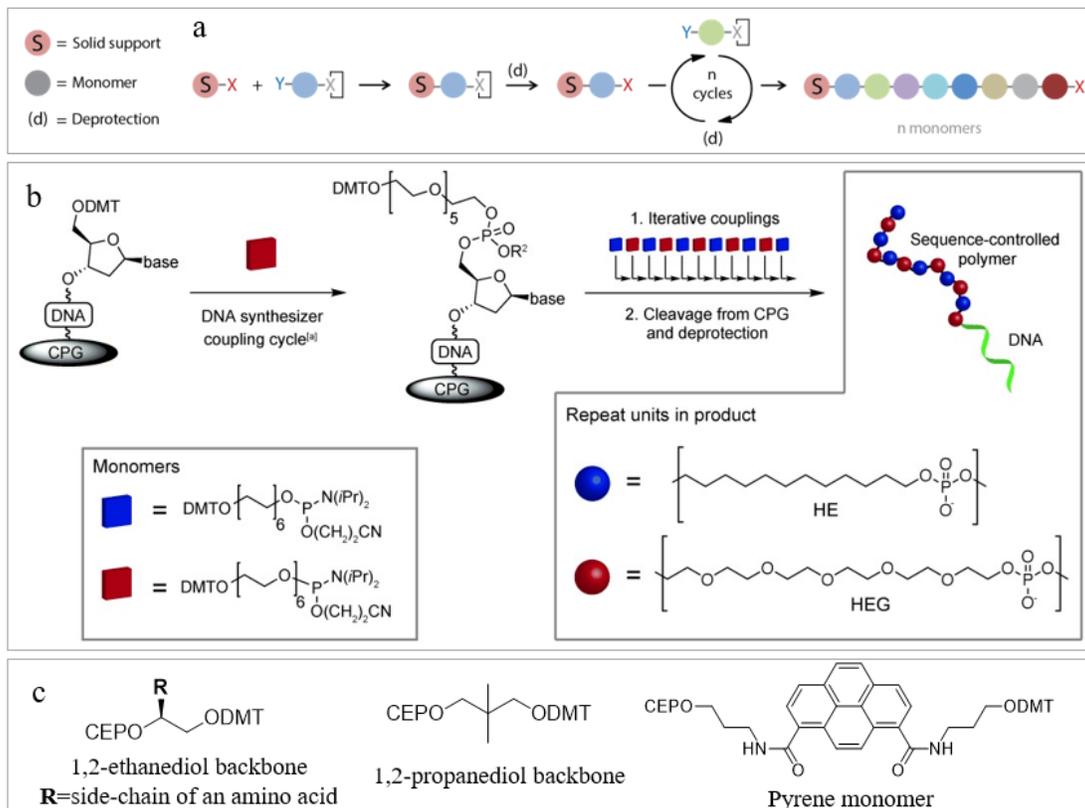


Figure 1.14. Solid-phase synthesis with protection/deprotection strategy. (a) Principle of the protection-deprotection strategy Adapted with permission from ref. 70. (b) Strategy to append an artificial sequence-defined oligomer to DNA through phosphoramidite chemistry. Adapted with permission from ref 121. (c) other types of phosphoramidite monomers for the preparation of artificial oligo(phosphodiester)s. Structures are reported in ref. 122, 123 and 124 respectively. DMT=dimethoxytrityl, CEP=cynoethylphosphoramidite.

The first examples of oligo(phosphodiester)s with side-chains not structurally related to nucleotides are found in the early 1990s by scientists aiming to label DNA strands with multiple moieties.^{125,126} However, the first time this strategy was employed without DNA attached was in 1998 by Ganesan and coworkers.¹²² This study generated 12mer peptidomimetics with amino acid side-chains bearing phosphoramidites (**Figure 1.14c**). This chemistry was also used for the synthesis of “oligopyrenotides” which are homo-oligomers. The monomers are pyrene-based phosphoramidites (**Figure 1.14c**) and oligomers have a DP of 14.¹²⁴ Our group showed the synthesis of unnatural precision oligo(phosphodiester)s attached to DNA.¹²¹ Up to 12 hydrophobic aliphatic carbon chains and hydrophilic hexa(ethylene glycol) chains were attached in a perfectly sequence-controlled manner to a DNA 19-mer (**Figure 1.14b**). The polymers made were all water-soluble due to the negative charge of the phosphate groups. The use of two monomers with

different solubility in water led to sequence-dependent self-assembly behaviours. Lutz and coworkers then reused the 1,3-propanediol backbone reported previously¹²⁶ (**Figure 1.14c**) to make artificial sequence-defined oligo(phosphodiester)s with up to 25 units in a row.¹²³ They built up on this work with the synthesis of a 100mer using an automated synthesizer.¹²⁷ Phosphoramidite chemistry combined with SPS can thus afford long unnatural sequence-defined polymers with up to 100 units. This finding highlights the efficiency of this technique compared to all other methods reported in this introduction. However, substantial side-chain diversity remains to be explored.

Unnatural sequence-defined polyamides were synthesized through this strategy. Up to 9 *N*-methyl imidazole, pyrroles and other groups could be introduced to the backbone of sequence-defined oligoamides.¹²⁸ In another report, a discrete oligobenzamide (7mer) was synthesized and attached to polyethylene glycol (PEG) using a peptide synthesizer.¹²⁹ Sequence-defined polymers that have a non-peptidic or phosphodiester backbone have also been synthesized. For instance, an oligocarbamate was reported by the Schultz group.¹³⁰ Similar oligomers that contained guanidinium moieties were then synthesized by Wender and coworkers.¹³¹ Oligoureas were also investigated as peptidomimetics,¹³² and their synthesis was improved with microwave irradiation.¹³³ Cyclic and linear oligoamines,^{134,135} oligoesters^{136,137} and oligotriazoles¹³⁸ were also made through SPS. Two state-of-the-art methods were successful for the synthesis of oligoarylenes¹³⁹ and oligo(arylthiophenes) through SPS.¹⁴⁰ All these unnatural oligomers have been synthesized for specific applications that are further detailed in section 1.4.

The second SPS strategy is called the submonomer strategy (**Figure 1.15**). It consists in attaching a monomer that has a convenient chemical handle to the growing chain. The moiety of interest is then reacted with the chemical handle and forms a side-chain. The next step is the chain extension with the attachment of another monomer.

Peptoids are polyamides like peptides but the side chain diversity rests on the amide nitrogen instead of the α -carbon. They can be synthesized similarly to peptides¹⁴¹ but the most applied method to make them is through the submonomer approach. The first peptoids were prepared in 1992 using solid-phase synthesis.¹⁴² Many following reports have greatly expanded the diversity of side-chain functions available (>100), as well as the amide backbone (from α - to β -amino acids for example).^{143–145} The most common strategy relies on cycles starting with the coupling of 2-

bromo acetic acid to grow a peptoid chain attached to resin. Then, the bromide is substituted with broadly available functionalized amines (**Figure 1.15**). Polymers up to 50 monomers long were achieved using this method.¹⁴⁶ Versatile sequence-defined trimers with a γ -peptidic backbone were also synthesized.¹⁴⁷ In this study, the main building block used was a Trans- γ -Fmoc-amino- α -Boc-l-proline where the Fmoc-protected nitrogen was used for SPS peptide elongation and the Boc-protected group was used for the stepwise coupling of different chromophores.

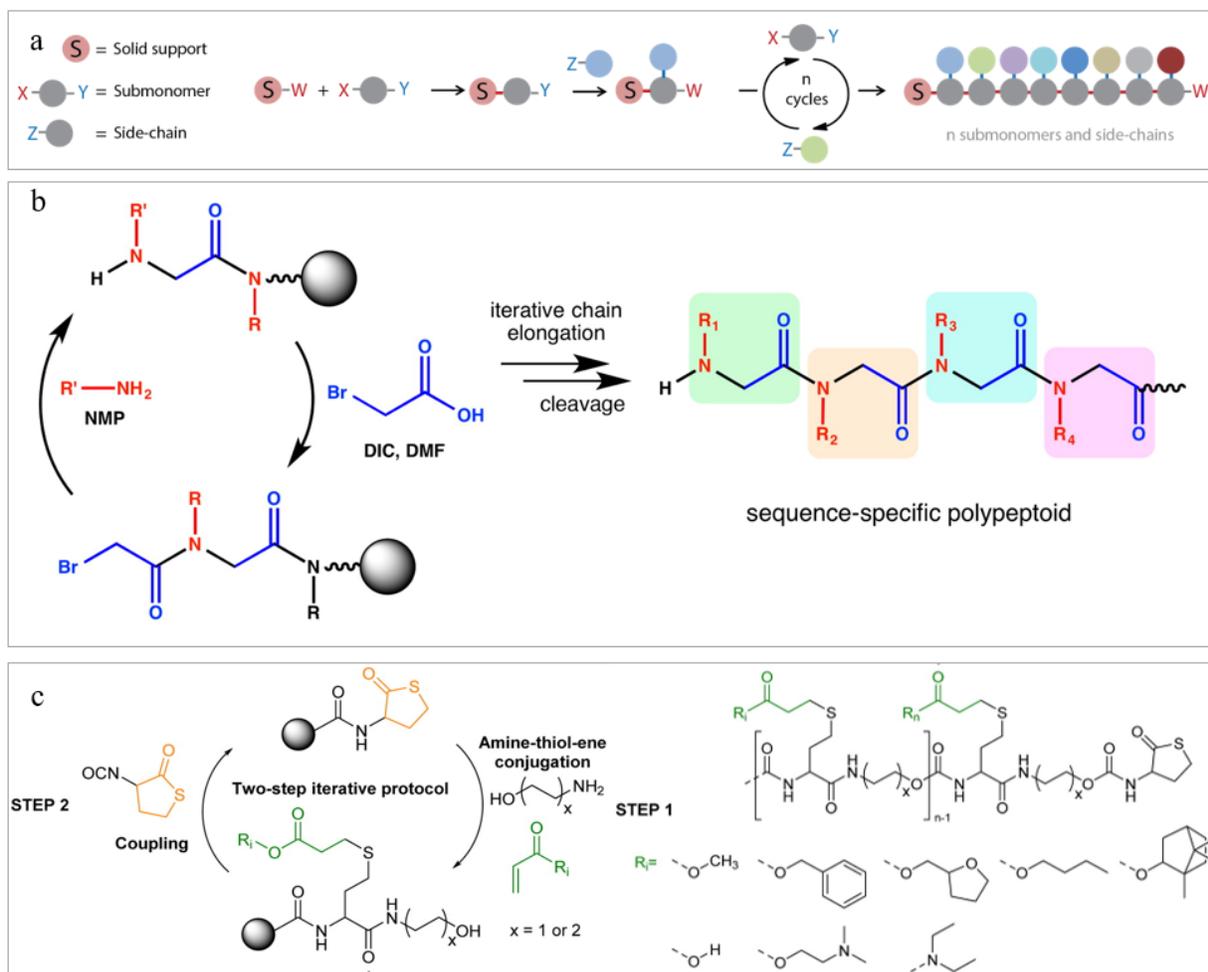


Figure 1.15. Solid-phase synthesis: submonomer strategy. Principle of the synthetic strategy. Adapted with permission from ref. 70. (b) Most common peptoid synthetic strategy. DIC is *N,N'*-diisopropylcarbodiimide, DMF is dimethylformamide. Adapted with permission from ref 148. (c) Isocyanate-containing lactones and acrylic esters as building blocks in sequence-defined polymer iterative synthesis. Adapted from ref. 149.

Thiolactone chemistry has also been investigated. For example, sequence-defined oligomers were made through cycles of thiolactone aminolysis followed by thiol addition on a thiolactone-containing building block.¹⁵⁰ Side-chain functionality was introduced during the aminolysis step.

Perfect sequence control was obtained for the synthesis of tetramers. The same group further elaborated on this method by automating the process and introducing a readily available isocyanate containing thiolactone (**Figure 1.15**).¹⁴⁹ This molecule can react with an alcohol-bearing growing oligomer in high yields. With this improvement, the authors were able to make up to a decamer.

To make sequence-defined oligomers using SPS, strategies involving several types of monomer attachment chemistry have been explored (**Figure 1.16**). These strategies produce less waste than protection/deprotection methods and have therefore attracted much interest in the past few years.

CuAAC has been investigated along with amine coupling to make up to a 6-mer by Lutz and coworkers.¹⁵¹ The same group made poly(alkoxyamine amide)s by associating amine coupling and copper mediated radical-radical coupling of an alkyl bromide and a nitroxide.¹⁵² The method was still efficient after 24 synthetic cycles showing its great potential for long oligomers with two different monomers. Oligocarbamates¹⁵³ and oligo(alkoxyamine phosphodiester)s¹⁵⁴ were also synthesized by the Lutz group and further showed the potential of SPS without protecting groups. Daily and coworkers have shown the synthesis of triazine based sequence-defined 6-mers in high yields.¹⁵⁵ The Niu group combined sulfur-fluoride exchange reaction and CuAAC to make up to a 9-mer (**Figure 1.16**).⁸¹

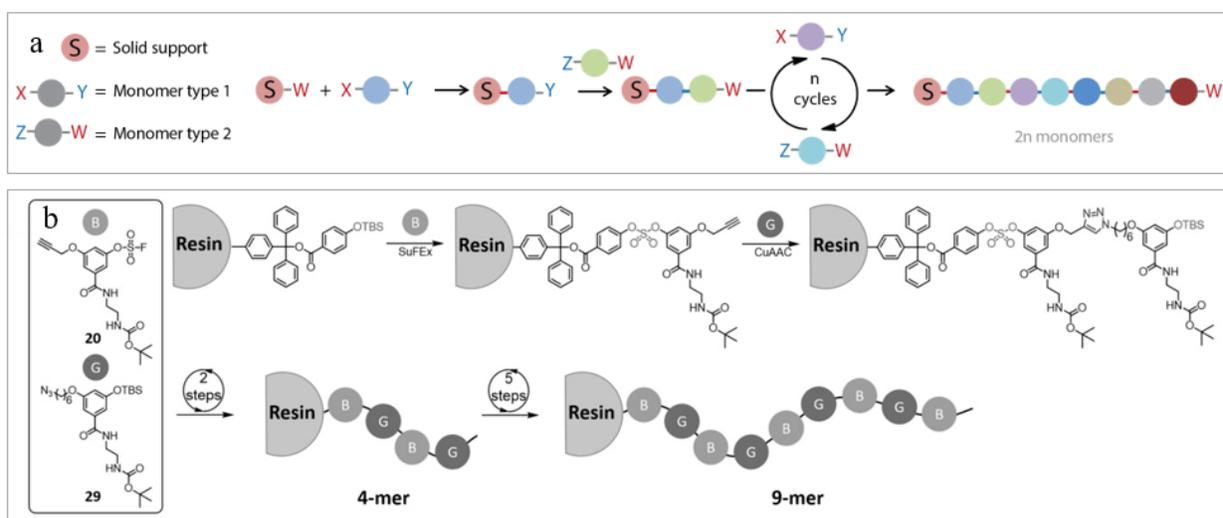


Figure 1.16. Solid-phase synthesis involving different types of chemistry. (a) Principle of the strategy with two monomer types. Adapted with permission from ref. 70. (b) Sequence-defined oligomers made through sulfur-fluoride exchange and CuAAC. Adapted with permission from ref. 81.

1.3.4.3 Liquid-Phase Synthesis

While progress has been made towards SPS scalability,¹⁰ liquid-phase synthesis remains valuable for higher scale synthesis. Moreover, NMR spectroscopy or spectrophotometry are solution-based characterization methods, enabling reaction efficiency to be probed at each step.

Multicomponent reactions are of interest because they reduce the number of steps required in synthesis. They have been widely used in SCP synthesis.¹⁵⁶ This strategy has been notably applied by Meier and coworkers. They used the Passerini reaction combined with thiol-ene chemistry to make a sequence-defined tetramer.¹⁵⁷ Replacing the thiol-ene reaction by a deprotection step¹⁵⁸ or using thiolactone chemistry (**Figure 1.17**)¹⁵⁹ led to some improvements in product scale and oligomer length (up to a 15mer). In another publication, the Goldup group made a pentarotaxane in high yields.¹⁶⁰ Five macrocycles were added along the same thread in a sequence-defined manner. However, while efficient, these methods involve tedious purification steps.

Many reports developed elegant methods to circumvent this limitation. Liquid-liquid extraction and filtration through molecular sieving may be used as a simpler purification step than column chromatography. This strategy has been applied to make up to 12mers with 6 sidechains on a sequence-defined polymer star.¹⁶¹ “Fluorous” affinity purification can be considered as a fast selective alternative to classic chromatography.^{162,163} Perfluorocarbon chains have very low affinity for classical aqueous and organic solvents. Therefore, compounds having a perfluorocarbon chain would not be easily washed away from a fluorous affinity column. Alabi and coworkers showed the rapid synthesis of oligo(thioetheramide)s using fluorous purification in between each synthetic step (**Figure 1.17**).¹⁶⁴ The synthesis of up to a 16-mer highlights the efficiency of the fluorous tag strategy. This method was improved to make sequence-defined macrocycles with a large-variety of side-chains.¹⁶⁵ Tetramers made of hydroxyproline blocks could also be synthesized applying the fluorous tag strategy.¹⁶⁶ The Burke group developed a synthesizer for making a variety of polyenes in the liquid phase through iterative cross-coupling reactions (**Figure 1.17**).¹⁶⁷ After each monomer coupling, growing chains contain a N-methyliminodiacetic acid (MIDA) boronate moiety. Such a functional group has unusual binary affinity for silica gel in specific elution conditions. Thus, growing chains are “caught” on silica gels while byproducts elute and are then “released” using the right combination of solvents. This “catch and release” technique allowed the full automation of polyenes synthesis.

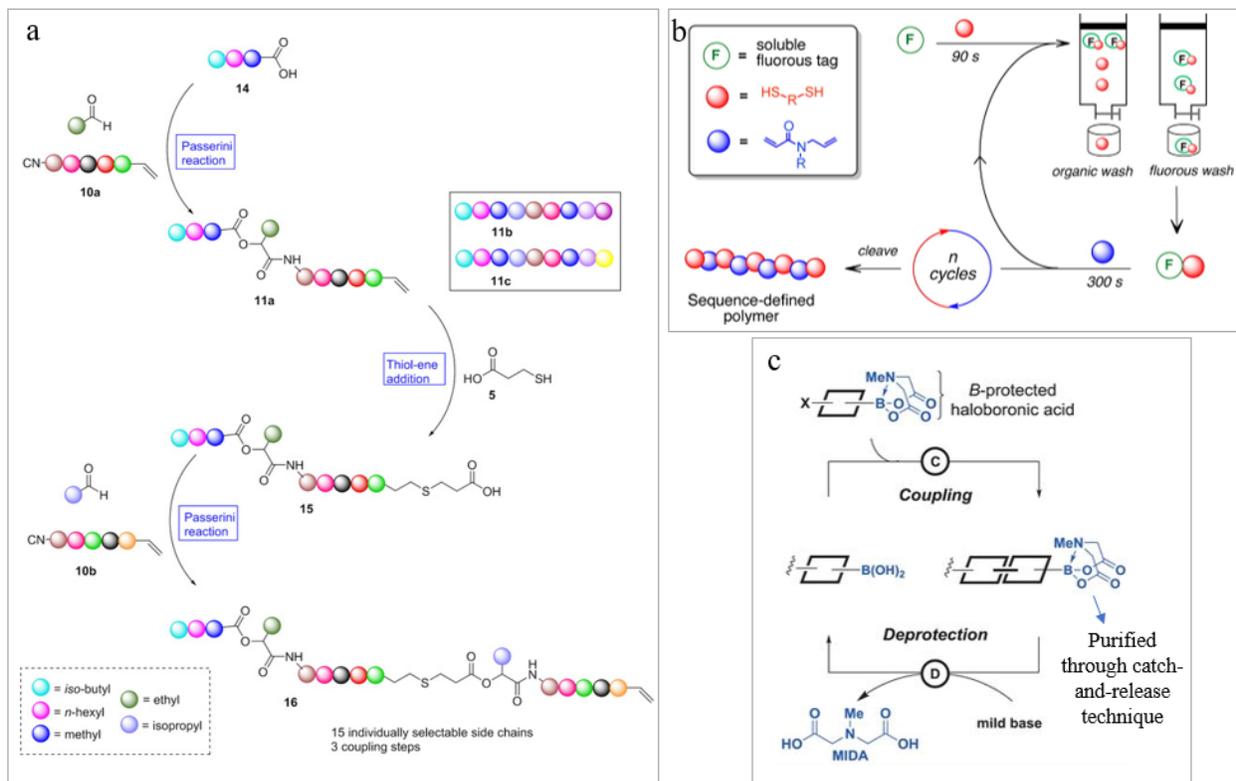


Figure 1.17. Liquid-phase synthesis involving different chemistry and purification processes. (a) Use of Passerini multicomponent reaction and thiolactone chemistry leads to compound 10a, itself involved in multiple Passerini and thiol-ene reactions to afford a 15mer. Adapted with permission from ref. 159. (b) Use of a fluororous tag for purification. Adapted with permission from ref. 164. (c) MIDA-boronates in polyenes synthesis. Adapted with permission from ref 167.

Soluble supports have originally been developed for oligonucleotide¹⁶⁸ and peptide synthesis.¹⁶⁹ Some of them are insoluble in solvent systems where synthesis reagents are soluble, allowing purification through precipitation. Few examples of modest sequence control in unnatural polymers using a soluble polymer support were reported.^{170,171} For actual sequence-definition, previously described poly(alkoxyamine amide)s¹⁵² and polymers made through the Passerini 3-component reaction¹⁵⁷ were respectively made with a polystyrene and a PEG polymer support precipitating in selected solvents. Conjugated sequence-defined tetramers were also synthesized using a soluble polymer support.¹⁷² Soluble supports are valuable methods to scale-up the synthesis of sequence-defined polymers, although they often lead to poor degrees of polymerization and are therefore not broadly used.

The Barner-Kowollik group showed the possibility of making sequence-defined oligomers that grew on both ends at the same time.¹⁷³ To increase further the speed and degree of polymerization,

Iterative Exponential Growth (IEG), first reported in the early 1980s, may be the method of choice (**Figure 1.18**).¹⁷⁴ The Tour group used this approach to make conjugated oligo(1,4-phenylene ethynylene)s.¹⁷⁵ Similar strategies were used to make discrete polycaprolactone¹⁷⁶ and conjugated polymers.^{177,178} The Monteiro group synthesized sequence-defined comb-polymers through the IEG strategy with CuAAC.¹⁷⁹ Zhang and coworkers combined thiol-maleimide coupling and deprotection steps for applying IEG to the synthesis of discrete block-copolymers with up to 128 units.¹⁸⁰ The Johnson group included stereocontrol in the IEG and called it IEG+ (**Figure 1.18**).¹⁸¹ Through the stereospecific opening of epoxides with sodium azide combined with CuAAC, a variety of sequence-defined block-cooligomers with two different monomeric units were synthesized. This strategy was also adapted to flow chemistry,¹⁸² allowing the synthesis of sequence-defined oligomers at the gram scale without the need of manual purifications. IEG allows to access long oligomers or polymers; however, it is at the expense of sequence control. Indeed, the polymers obtained are repetitive or palindromic.

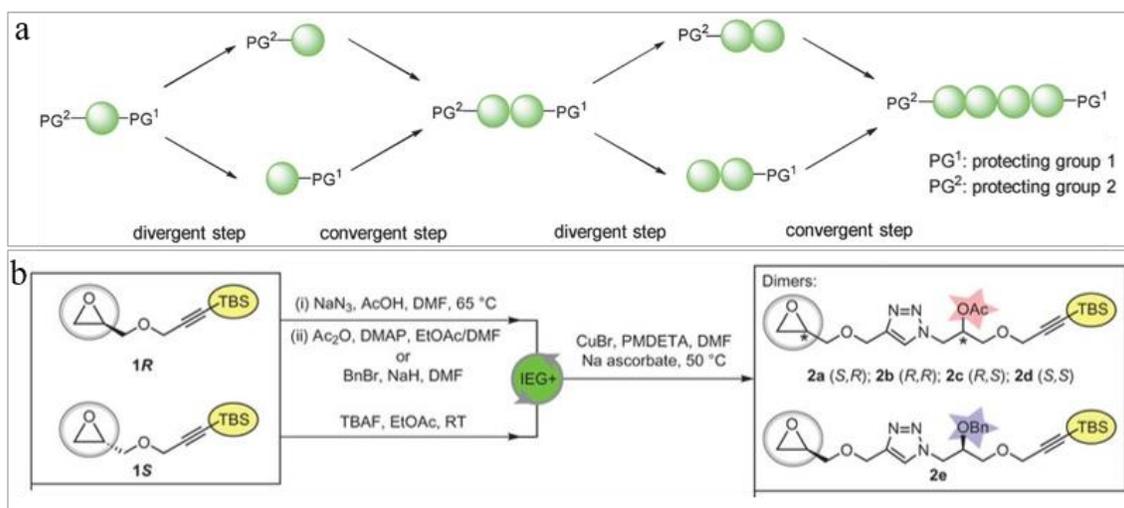


Figure 1.18. Iterative exponential growth. Principle of IEG. Adapted with permission from ref. 72. (b) Synthesis of sequence-defined oligomers through IEG+. Adapted with permission from ref. 181.

1.3.5 Overview of SCP synthetic approaches

Many different methods are available for the synthesis of unnatural sequence-controlled polymers. Ultimately, total control over the sequence, high degrees of polymerization and versatile monomers are the three criteria to reach. Improving one of these conditions usually works at the

detriment to another. Thus, periodic polymers have high degrees of polymerization (>100), are made of versatile monomers but involve imperfect sequence-control. On the other hand, methods such as templated chemistry allow the synthesis of perfectly defined heterofunctional oligomers but are limited to very short oligomers. Among the most accomplished sequence-defined polymers, peptoids answer the three criteria but to a certain extent ($DP < 50$). Unnatural sequence-defined poly(phosphodiester)s can be synthesized with high DPs (~ 100) but very few reports tend to significantly enlarge the library of monomers. This thesis notably focuses on this challenge.

1.4 Applications of sequence-controlled polymers made of unnatural building blocks

Artificial sequence-controlled polymers synthesis has been made possible through a variety of methods affording various types of macromolecules. Parallel to this synthetic work, applications in information storage, self-assembly, folding and medicine have been developed.^{183–186} This section provides an overview of the applications of SCPs.

1.4.1 Information Coding, Decoding and Storage with sequence-controlled polymers

DNA is the biopolymer of information storage. It can be read, copied, stored and degraded. Artificial sequence-controlled polymers offer the opportunity to improve information storage in macromolecules. Indeed, SCPs may be less expensive than DNA, quicker to synthesize and decode, and more stable to specific storage conditions. Moreover, DNA can exclusively be used for quaternary encoding due to its four letters code. An artificial polymer may allow more data storage in less space than DNA. This is a major concern in modern information technology.

From the precise incorporation of maleimides through chain-growth polymerization¹⁸⁷ to actual sequence-defined polymers, Lutz and coworkers have designed their polymers with the objective to store information. They chose to work with binary systems. In other words, their sequence-defined polymers were mostly made of two chemically different monomers. While this makes coding/decoding easier than with more monomers, it *a priori* requires polymers with very high DP to store large amounts of information. They showed the first example of an artificial oligomer that

could be coded, read and erased in 2015.¹⁵² This specific polymer possesses a C–ON bond that would get preferentially cleaved during electrospray ionization (ESI). It made sequencing through mass spectrometry (MS) significantly easier and shows how polymer design is crucial for sequencing application. Using tandem mass spectrometry (MS/MS), they could recognize degradation patterns of the polymer and deduce the sequence of trimers. This method was extended to poly(alkoxyamine phosphodiester)s¹⁵⁴ (**Figure 1.19**) and to denser digital storage without having to increase the polymers DP.¹⁸⁸ They used oligomers dispersity to code in two dimensions: the polymer size would encode for a number and its sequence for more data. Poly(phosphodiester)s with DP>100,¹²⁷ were also used by the same group to enhance information storage capacity. With these polymers, the negatively charged backbone directed the layer-by-layer assembly of poly(phosphodiester)s and polyamines.¹⁸⁹ Such a structure contains information from the polymers' sequence but also from the position of each polymer in the assembly, resulting in another type of 2-dimensional coding system.

Other groups have introduced denser data encoding in shorter sequences. The Meier and du Prez groups independently reported the use of n monomers to encode in base n instead of base 2. Using a library of monomers made through the multicomponent Passerini reaction, Meier showed the possibility of encoding 97 bits of information in a tetramer.¹⁹⁰ More than 15 side-chain functionalities were used by du Prez and coworkers in order to write a sentence and even a QR-code (**Figure 1.19**).¹⁹¹ The latter was encoded in 71 sequence-defined oligomers of different size. A special decoding algorithm was developed to read the data using MS/MS.

in methacrylate based ocular implants tags¹⁹⁷ and even tested in poly(vinyl alcohol) (PVA) films that were implanted in rats.¹⁹⁸ After three months, the films could be explanted and the poly(carbamate)s tags were sequenced efficiently (**Figure 1.20**).

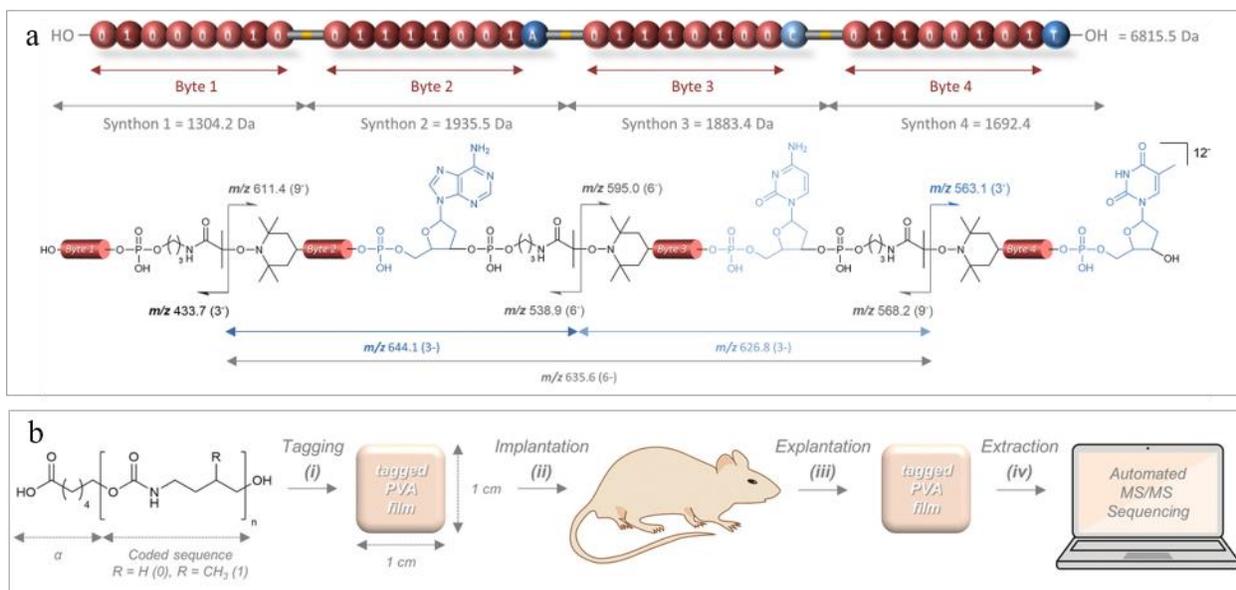


Figure 1.20. Information storage in SCP: improvements and applications. (a). Separation of bytes with cleavable spacers allows easier sequencing through MS/MS. Adapted with permission from ref 193 (b) Polycarbamates as molecular tags of implants. Adapted with permission from ref 198.

Despite these promising strides, challenges remain. The amount of information encoded increases with the degree of polymerization and no artificial sequence-defined polymers with DP>100 have been synthesized so far. The speed of synthesis and the development of accurate decoding techniques for long polymers are yet to be addressed.

1.4.2 The impact of sequence control on self-assembly and folding

The function and efficiency of biopolymers mainly comes from their 3D structure. Hence, it is not surprising that actual research fields (foldamers, single-chain nanoparticles (SCNP)) are dedicated to the elucidation of the folding and self-assembly of precision polymers. Here, we will focus on some of the most recent and relevant applications to give an overview of the state of the art.

1.4.2.1 Self-assembly with sequence-controlled polymers

Block-copolymers highlight how even modest sequence control can lead to a variety of self-assembly behaviours and applications.¹⁹⁹ Herein, we will show how advanced sequence control can influence self-assembly.

Peptoids probably are the most broadly studied sequence-defined polymers. Zuckerman and coworkers discovered that amphiphilic sequence-defined peptoids with a charged hydrophilic end and an aromatic ring-containing hydrophobic portion can self-assemble into nanosheets.^{148,200} Novel secondary structural motifs called peptoid Σ -sheets were characterized with the help of molecular dynamics simulations (**Figure 1.21**).²⁰¹ Finding new structural motifs with artificial sequence-defined polymers highlights the novel modes of assembly that such polymers can offer. A variety of studies show sequence-structure relationship for peptoid nanosheets. For instance, rheology measurements were realized to predict if a peptoid could form nanosheets.²⁰² Segalman and coworkers recently used a hybrid block copolymer-sequence-defined peptoid and showed how domain spacing and order-disorder transition temperature were sequence-dependent.²⁰³ These sheets are well-defined enough to be used for a variety of applications. For instance, sugars were displayed on a precisely engineered peptoid nanosheet for protein recognition.²⁰⁴ Peptoids are not only used for 2D nanosheets. For example, the Zuckerman group showed the formation of well-defined nanotubes.²⁰⁵ Another report highlights the use of peptoids to direct the assembly of hyperbranched gold nanorods for plasmonics applications.²⁰⁶

The variety of monomers types, polymer backbone, sequence, and DP in SCP gives rise to very different self-assembling systems. Using oligo(phosphodiester)s made of two sorts of aromatic flat monomers, the Häner group could make large 2-dimensional nanosheets²⁰⁷ and supramolecular nanotubes for light-harvesting applications (**Figure 1.21**).²⁰⁸ Silicon-based nanocages positioned precisely along a polymeric chain led to various self-assembly behaviours.²⁰⁹

On the frontier between block copolymers and sequence-defined oligomers, the self-assembly of monodisperse relatively short (DP<100) block co-oligomers has recently been the subject of multiple studies. For example, dimethylsiloxane-lactic acid diblock co-oligomers reproduced some BCP self-assembly behaviors: cylindrical, gyroid, and lamellar nanostructures with small-feature size but long-range organization were obtained (**Figure 1.21**).²¹⁰ IEG-made sequence-defined block co-oligomers were shown to self-assemble into similar nanostructures.²¹¹ In this

case, monomers stereochemistry had a major influence on the type of self-assembly observed. The Meijer and the Hawker group independently showed that the discrete aspect of sequence-defined co-oligomers has a great influence on their self-assembly.^{212–214} For example, discrete oligo(ethylene glycol) methyl ether *block*- oligo(L-lactic acid) revealed crystal-driven gelation whereas similar disperse block co-oligomers did not.²¹⁴

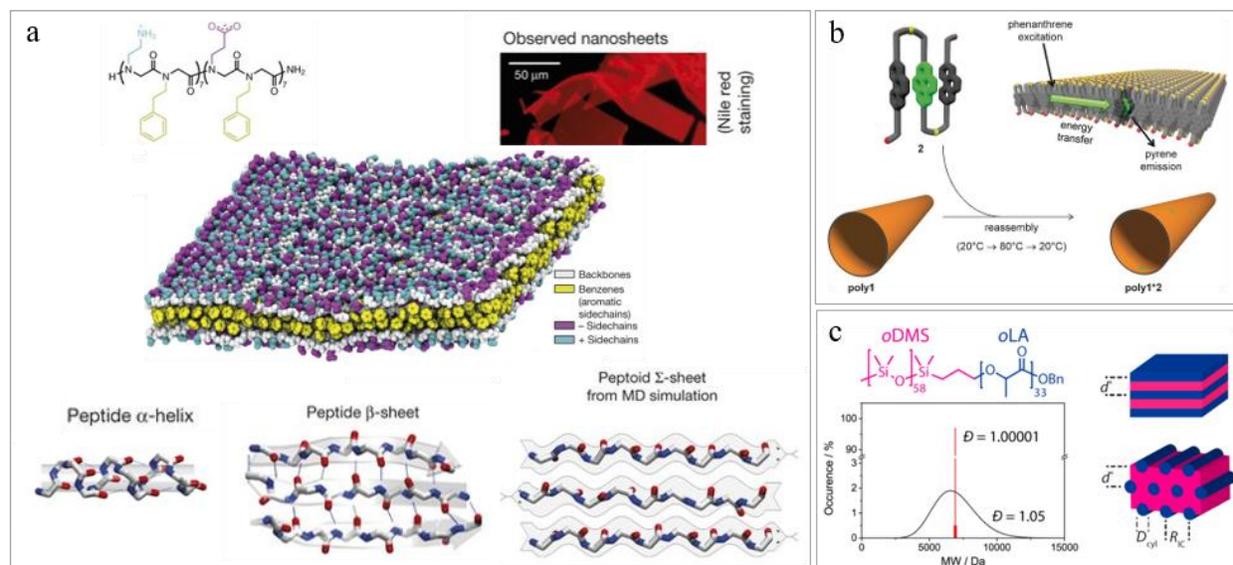


Figure 1.21. Self-assembly of sequence-defined oligomers. (a). Nanosheet-forming peptoids exhibited a new secondary structure motif called peptoid Σ -sheets Adapted with permission from ref. 201. (b) Light-harvesting nanotubes made of sequence-defined triphosphodiester containing phenanthrene and pyrene units. Adapted with permission from ref. 208. (c) Self-assembly of monodisperse block co-oligomers. Adapted with permission from ref. 210.

1.4.2.2 Foldamers

Foldamers are short sequence-defined oligomers that fold into well-defined 3-dimensional structures.²¹⁵ The field of foldamers is older²¹⁶ than sequence-controlled polymers, though foldamers themselves could be considered as sequence-defined oligomers. Scientists initially used SPS or classic organic chemistry to make foldamers. The growing number of synthetic routes towards sequence-defined polymers should open major opportunities in the foldamers field. Herein, we only cover representative examples of the different types of foldamers reported.

Helix-forming oligoamide foldamers are among the most broadly studied.^{217–219} For example, strong host-guest interactions were designed between a small molecule and such foldamers of different sizes.²²⁰ Directional molecular motion was designed in a similar system where the helix

could shuttle along a rod-like guest (**Figure 1.22**).²²¹ Advances in the field have allowed the iterative design of aromatic oligoamide foldamers that bind a specific molecule such as fructose.²²² This example shows the potential of foldamers to be rationally designed for specific applications. Higher order self-assembly was obtained with oligoureia foldamers that afforded helix-forming oligomers assembling into a well-defined six-helix bundle.²²³ This structure was further used for the encapsulation of primary alcohols in water.²²⁴ Foldamers could also be envisioned for catalytic applications. For instance, an oligoamide foldamer made with *trans*-2-aminocyclohexanecarboxylic acid showed aldolase activity.²²⁵ A peptoid foldamer was used to oxidize 1-phenylethanol in an enantioselective way.²²⁶ Similarly, enantioselective C-C bond formation was reported with an oligoureia foldamer.²²⁷

Dynamic foldamer chemistry aims at introducing some conformational freedom into foldamers to make them stimuli-responsive and reconfigurable.²²⁸ For example, Clayden and coworkers designed a helix-forming oligoamide that could be switched from left- to right-handed orientations through photoirradiation.²²⁹ This had a dramatic impact on the enantiospecificity of a C-C bond formation reaction (**Figure 1.22**).

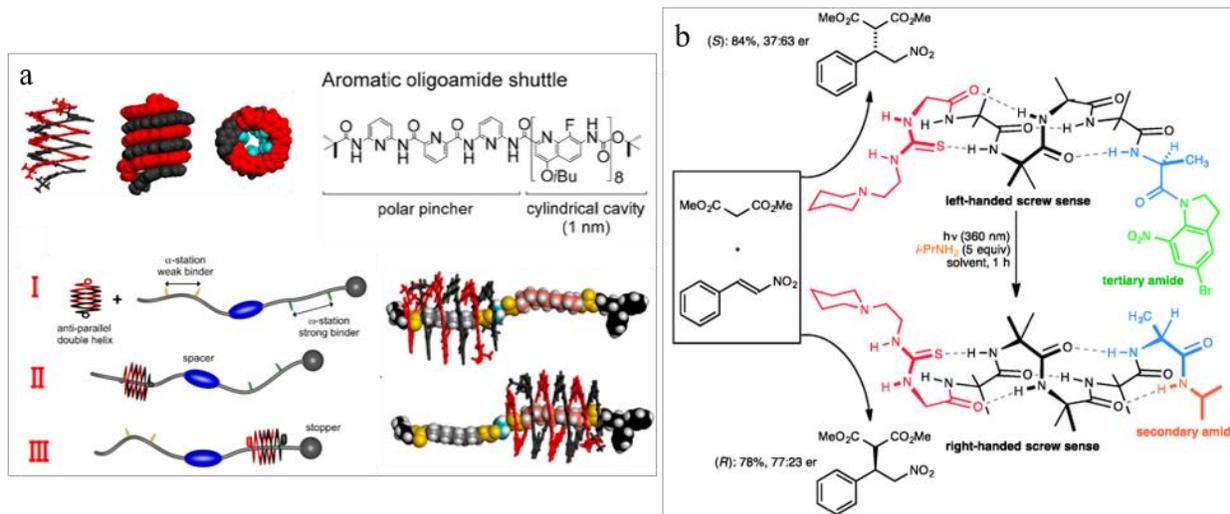


Figure 1.22. Applications of foldamers. (a). Helix-forming oligoamide shuttling along a polycarbamate rod-like structure. Adapted with permission from ref. 221. (b) Photoirradiation induces helix screw-sense switching and product stereochemistry outcome. Adapted with permission from ref. 228.

Aromatic monomers containing oligo(phosphodiester)s were also shown to be able to fold. For instance, perylenetetracarboxylic diimide units arranged on a sequence-defined oligomer backbone were shown to π -stack.²³⁰ The NMR ring-current effect indicated aromatic rings nearly

coaxial alignment. Such nanostructures are good models for studying energy transfer between chromophore neighbours. Sequence-definition is of great interest for precise energy-transfer studies in folded conjugated materials.

1.4.2.3 Single-Chain folding

While foldamers focus on short oligomers forming well-defined secondary structures, single-chain nanoparticles are formed through interactions between remote moieties.²³¹ In other words, distant interactions will induce single-chain collapse into a globular tertiary structure. Synthetic routes allowing high DP and moderate sequence control are privileged. Therefore, the 3D structure of SCNP is usually disordered.

The strategies used for intramolecular chain collapse are based on covalent interactions, metal-binding or weak interactions. Lutz and coworkers showed the possibility of precisely positioning alkyne- and azide-containing maleimide units in a polymer chain.²³² CuAAC or Glaser coupling performed in very dilute conditions allowed specific folding (**Figure 1.23**). The Meijer group showed the synthesis of a block copolymer containing benzene-1,3,5-tricarboxamide units.²³³ These units self-assembled into a helical structure and their formation led to intramolecular chain collapse. The addition of ligands to the polymer chain resulted in the placement of catalytically active sites in the final nanoparticle. Similarly, the same authors showed in a later report the attachment of different ligands to impart more functions to the nanoparticles (**Figure 1.23**), such as catalysis in water.²³⁴ Metal-binding interactions can play the role of glue in intramolecular interactions and provide the active site for potential catalysis.²³⁵ For example, Pomposo and coworkers used a metallo-folded polymer containing complexed Cu(II) to catalyze oxidative couplings.²³⁶ SCNPs adopt a globular shape and can catalyze reactions in water. As such, they can be considered as enzyme mimics. Selectivity and efficiency of these enzymes could be improved with the introduction of more sequence control and more ordered structures. More recently, the Berda group worked towards that direction by adding sequence-defined fragments into long polymer chains.²³⁷ These have two roles : (i) they introduce secondary well-ordered structures in the final SCNP and (ii) they undergo a multicomponent cross-linking reaction in dilute conditions to induce polymer chain collapse into a globular structure.

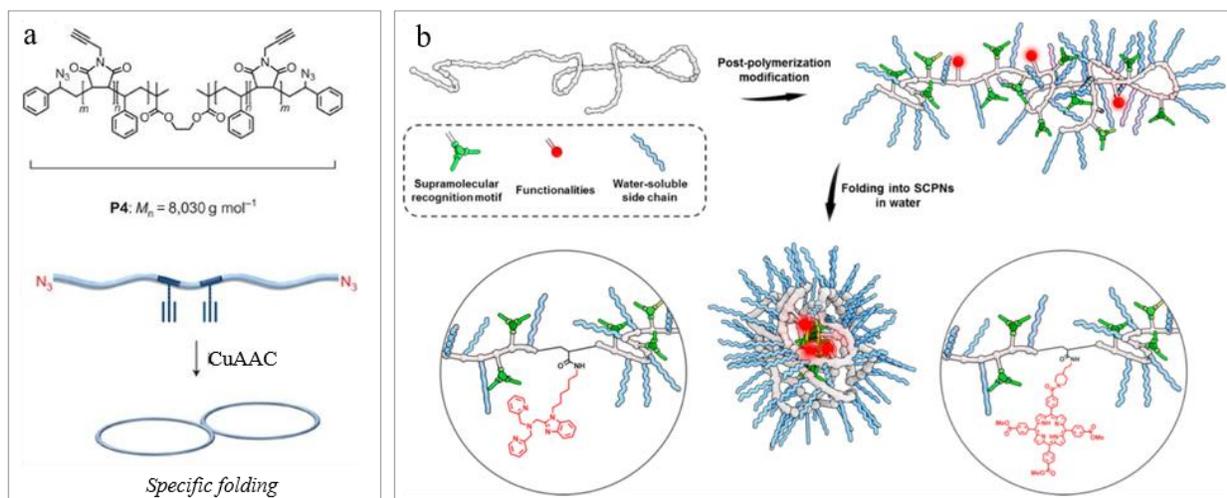


Figure 1.23. Single-chain folding in sequence-controlled polymers. (a). Precise positioning of reactive cross-linkers during chain-growth polymerization. Adapted with permission from ref. 232. (b) Benzene-1,3,5-tricarboxamide units self-assembly triggers single-chain collapse in sequence-controlled polymers. SCPN formed can be functionalized with ligands or photosensitizers. Adapted with permission from ref. 234.

1.4.2.4 Impact of sequence on macroscopic properties

Self-assembly and folding are responsible for the microstructures of polymer materials. Therefore, sequences may have a dramatic effect on the macroscopic properties of some materials. Meyer and coworkers synthesized conjugated sequence-defined oligomers with an electron-donor and an electron-acceptor monomer.²³⁸ Sequence variations had a large impact on the optical and electronic properties of the assembly. This work highlights the great potential of SCP for optoelectronic applications such as organic light-emitting diodes (OLEDs), molecular wires and solar cells. In another report, the Meyer group showed how sequence influenced the swelling and biodegradation patterns of poly(lactic-co-glycolic acid) (PLGA) matrices.²³⁹ According to application needs, materials properties can be engineered through sequence control, paving the way to materials bottom-up design with SCP.

Self-assembly, folding and macroscopic polymer properties have thus been studied with either short, perfectly defined oligomers or longer chains with poorer sequence-control. Both polymer types lead to a large variety of applications. Long, artificial, sequence-defined polymers (DP>100) remain to be synthesized. These materials promise to ‘fill the gap’ between proteins and their unnatural counterparts.

1.4.3 Sequence-Controlled Polymers for biomedical applications

1.4.3.1 SCP in drug discovery

Protein-protein interactions (PPIs) are being extensively studied for a better understanding of biological systems and as a new target for drug delivery. However, this type of interaction is difficult to target with small molecules due to its complexity. In the case of α -helix-mediated PPI, a class of molecules that would be able to structurally mimic α -helices are of great interest.²⁴⁰ Due to their reliable ability to form 3D structures, foldamers are an attractive class of molecules in this context. For example, β -peptide foldamers were used due to their propensity to form α -helices and withstand proteolysis compared to natural amino acids (**Figure 1.24**).²⁴¹ Such a foldamer was found to mimic the Bak (pro-apoptotic protein) BH3 helix that usually forms PPI with Bcl-x_L (anti-apoptotic protein). Similarly, fully artificial oligobenzamide and oligopicolinamide foldamers were shown to form α -helix mimetics that inhibited Bcl-x_L.²⁴²

SCP have also shown the potential to be antimicrobial agents. Due to the emergence of many multi-drug-resistant infections, new classes of antibiotics are being extensively studied. For example, helical peptoids were investigated using a quantitative structure-activity relationship model and a potent antimicrobial agent was synthesized.²⁴³ The automated synthesis of peptoids enabled the synthesis of a number of sequences to be made. This strategy further allowed a decrease in non-specific cytotoxicity of the active peptoid described.²⁴⁴ The Alabi group showed in another report how sequence-defined oligo(thioetheramide)s can be used as macrocyclic antimicrobial agents.¹⁶⁵ Activity measured was similar to the one of ampicillin, an established antibiotic. Total sequence control enabled changing the macrocycle structure one monomer at a time. It allowed to rationally collect insights into the compound mode of action. Unnatural sequence-controlled polymers allow the exploration of the pharmaceutical potential of novel classes of macromolecules.

1.4.3.2 Cell-penetrating sequence-controlled polymers

Cell-penetrating peptides (CPPs) are promising molecules for the intracellular delivery of an active ingredient.²⁴⁵ Unnatural sequence-defined oligomers are less susceptible to proteases than peptides and would therefore be an attractive alternative for intracellular delivery. For instance, guanidinium-bearing oligocarbamates were shown to be efficient in transporting a probe molecule into skin cells.¹³¹ Oligo(thioetheramide)s have also been used in this context.²⁴⁶ Flow cytometry

measurements were used to demonstrate the capability of uncharged sequence-defined oligomers to cross cell membranes as efficiently as some cationic CPPs (**Figure 1.24**). Precise modification of the oligomer sequence adjusts the hydrophobic and cationic content of SCPs. Thus, cell-penetrating sequences for specific cargos can be rationally designed.

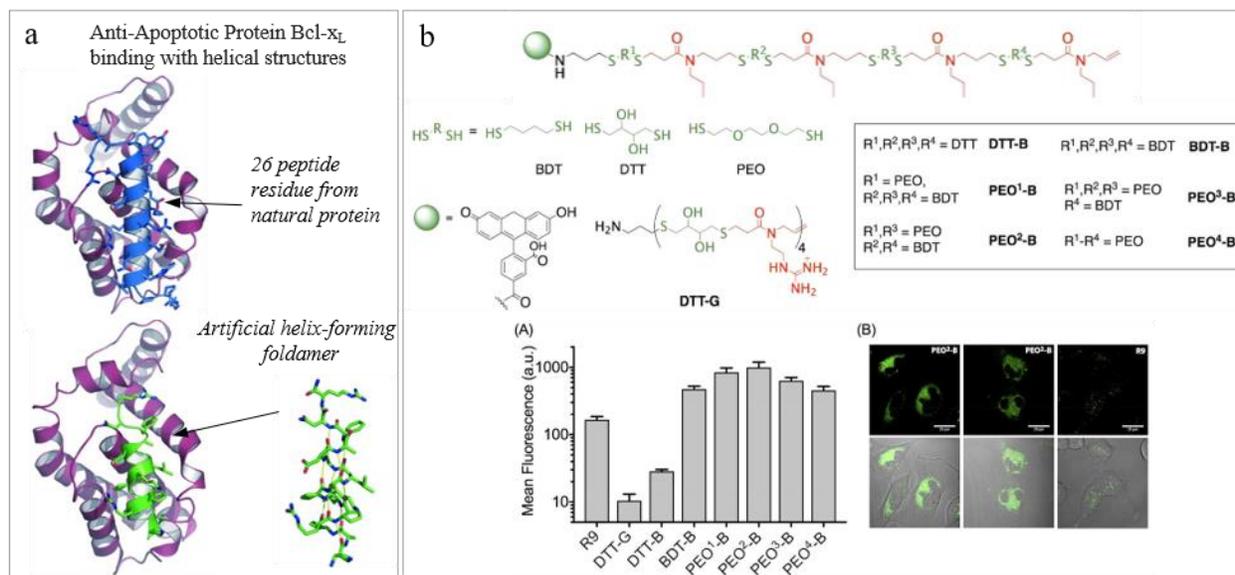


Figure 1.24. Biomedical applications with SCP. (a) An artificial β -peptide helix-forming foldamer is shown to form similar PPIs with anti-apoptotic-protein BCL-x_L as with its natural protein counterpart. Adapted with permission from ref. 241. (b) Versatile uncharged oligo(thioetheramide)s are able to enter cells efficiently compared to R9, a CPP as evidenced by flow-cytometry measurements in HeLa cells (A) and by live-cell confocal fluorescence imaging. Adapted with permission from ref. 246.

The main limitation in gene delivery is the poor cell permeability of nucleic acids. Several reports highlight the role that cationic sequence-defined oligomers can play in enhancing gene delivery therapies' efficiency.²⁴⁷ For example, sequence-defined poly(amido amine)s were synthesized using solid-phase synthesis.²⁴⁸ The positively-charged oligomers induced electrostatic interactions with negatively-charged siRNA and formed nanostructures called polyplexes. EG5 siRNA silences the anti-apoptosis EG5 protein and can therefore be used as an anti-tumor therapy. Polyplexes of EG5 siRNA with cationic sequence-defined polymers resulted in reduced tumor growth. More recent studies also highlight the potential of cationic sequence-defined oligomers for polyplexes formation and gene delivery in cells.^{249,250}

1.4.4 Sequence-controlled polymers design

The ability to design SCPs for a specific application may be more limiting than the SCP synthesis itself. Proteins exemplify this statement with protein folding remaining an active area of research.²⁵¹ Similar challenges (methods to predict the assembly of SCPs) slow down the emergence of advanced applications for artificial sequence-controlled polymers. The toolbox available to look for functional peptidic sequences can be adapted to artificial SCPs.¹⁸⁴ It is composed of the following concepts and research directions: (i) “bioabstraction” where a biopolymer is decomposed into its essential functional units; (ii) *de novo* design where primary sequences that were not previously synthesized are engineered according to folding predictions; (iii) combinatorial methods in which the law of large numbers guarantees to converge to the right polymer; and (iv) hybrid constructs made of a well-defined and understood entity and a variant artificial SCP.

1.4.4.1 Bioabstraction

Peptide chemists have widely applied the concept of “bioabstraction”. For example, antibody fragments have been developed in lieu of the full antibodies for molecular recognition. In the case of unnatural SCPs, the already mentioned use of α -helix mimics for modulating PPIs is a great example of bioabstraction.²⁴⁰ Indeed, only the main structural elements of the protein in contact with the other protein is copied. DNA mimetics with classic nucleobases but a different backbone such as PNA is another example where only the principal feature of a biopolymer is maintained to get enhanced performance with an unnatural SCP. In general, “bioabstraction” is an underlying principle in the design of many SCPs. However, this strategy requires deep understanding of a biopolymer mode of action and is therefore not applicable to all processes.

1.4.4.2 De novo design

De novo design has been showing some success with few protein examples but folding predictions are challenging.²⁵² However, for short sequences or a limited palette of interactions, computer-aided predictions have been shown to be very powerful. For example, *in silico* methods were used to design nanostructures made of nucleic acids (*i.e.* SCPs with 4 types of monomers).²² As explained earlier, DNA nanotechnology led to a large diversity of well-defined nanostructures. With unnatural building blocks, only peptoids have shown to lead to general design rules.^{201,253} *De novo* design is still in its infancy stage with artificial sequence-controlled polymers. Lessons from

the protein folding and nucleic acid nanotechnology fields could ensure rapid growth of this area of research.

1.4.4.3 Combinatorial strategies

Combinatorial synthesis of peptides finds sequences of interest amongst large libraries. This strategy relies on three steps: synthesis, selection and identification. SPOT and split-and-pool (or split-and-mix) strategies can be applied to unnatural sequence-defined polymers (**Figure 1.25**). SPOT is a parallel synthetic strategy that allows peptides to grow on a cellulose membrane or on glass.²⁵⁴ Many thousands of peptides can be grafted to the same surface and the spatial location of each peptide is associated to its sequence. Split-and-pool libraries are made during the iterative synthesis of a peptide.²⁵⁵ At every step, the support is divided in n parts that undergo an amine coupling. After the coupling, the supports are mixed together and split again for another amine coupling, thus obtaining n^x peptides, where x is the peptide length.

Peptoid libraries have already been successfully synthesized through a SPOT strategy leading to compounds having antimicrobial activity.²⁵⁶ Peptoid combinatorial libraries have also been synthesized using a laser-assisted methodology.²⁵⁷ As soon as 1998, the Schultz group reported the split-and-pool synthesis of large oligocarbamate libraries leading to protein ligand discovery.²⁵⁸ Very recently, Borner, du Prez and coworkers showed the combinatorial synthesis of a 8000-member unnatural precision polymer library (**Figure 1.25**).²⁵⁹ They could identify good drug solubilizers, showing the potential of their SCP in drug formulation.

A limitation in peptide library size is their sequencing. Indeed, peptide sequencing is based on degradation and mass spectrometry mostly. This method is sensitive but is restricted to short oligomers. To remediate this issue, Brenner and Lerner proposed in 1992 to tag peptides with oligonucleotides during a split-and-mix synthesis.²⁶⁰ The DNA section would be coding for the peptide sequence. Yet, DNA can be accurately amplified and sequenced from only few copies. This idea started the field of DNA-encoded libraries, where peptides^{104,261,262} and small molecules^{263,264} are tagged with a DNA strand serving as a bar code for the molecule structure (**Figure 1.25**). Recently, this idea came back within the realm of sequence-controlled oligomers.²⁶⁵ Using DNA-templated methods, reports highlight that the template could also be used as a bar code ready to be sequenced.^{104,108} This strategy yielded potent kinase inhibitors in a DNA-

templated library of sequence-defined trimers.²⁶⁶ However, oligomers made so far are short and the field would greatly benefit of longer DNA-encoded sequence-defined oligomers.

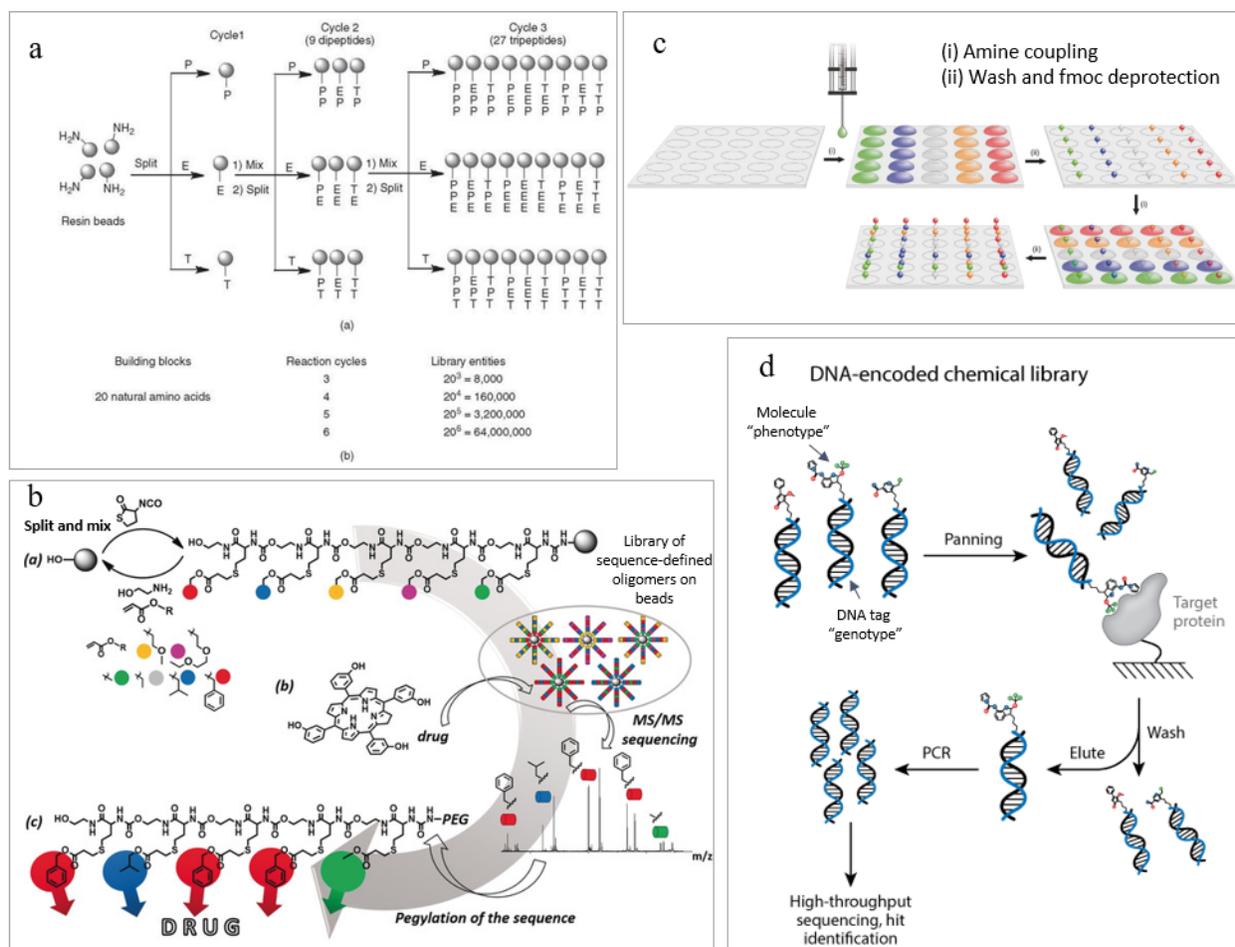


Figure 1.25. Combinatorial synthesis strategies for peptides, small molecules and artificial sequence-defined oligomers. (a) Principle of split-and-pool synthesis with peptides. Adapted with permission from ref. 267. (b) Split-and-pool combinatorial synthesis of non-natural oligomers allowed the discovery of drug solubilizers. Adapted with permission from ref. 259. (c) Principle of peptides SPOT synthesis where each location corresponds to a specific sequence. This has been applied to unnatural peptoids synthesis. Adapted with permission from ref. 184. (d) Principle of DNA-encoded libraries. The DNA serves as a barcode of the small-molecule structure. Adapted with permission from ref. 264.

1.4.4.4 Hybrid constructs

Mimicking the efficiency of biopolymers with SCP in catalysis is ambitious. Spontaneous folding of an unnatural polymer into a designed shape is very difficult to program. A strategy towards this goal would be to use known proteins or DNA nanostructures as scaffolds and position artificial SCP on them. Peptide and oligonucleotide-polymer conjugates can lead to a large variety of

applications.²⁶⁸ Here, we will focus on strategies where unnatural SCP imparted new structural or functional information to a biopolymer.

The most representative examples probably lie in the field of artificial proteins. For example, peptoid-peptide macrocycles were designed *in silico* to inhibit a protein-protein interaction.²⁶⁹ This report represents one of the most successful designs of an unnatural sequence-defined oligomer. Rules from peptoids assembly were used along with the Rosetta suite of computational tools, a program developed for protein modeling and sequence-optimization. The unnatural macrocycles synthesized were capable of inhibiting a relevant PPI and had a therapeutic effect in prostate cancer models.

Two other strategies where a full artificial foldamer is introduced in a protein are worth mentioning. Using a flexible tRNA-acylation ribozyme, aromatic oligoamide foldamers were used to initiate translation in a ribosome.²⁷⁰ This technique afforded foldamer-peptide conjugates synthesized enzymatically, paving the way to the synthesis of artificial foldamer-peptides hybrids. In another report, the Guichard group successfully replaced a peptidic α -helix of a native zinc-finger motif with an artificial α -helix forming oligourea through purely synthetic methods (**Figure 1.26**).²⁷¹ The artificial oligourea-peptide conjugate had similar binding affinity for zinc and data suggests that it can still bind to its target sequence. This clearly shows the potential of artificial foldamers to expand the structural, chemical and functional diversity of natural proteins in hybrid structures.

Oligonucleotide-artificial SCP conjugates have also been studied. For example, DNA strands modified with a sequence-defined trimer made of phenanthrene was reported and showed to self-assemble as a vesicle with light-harvesting properties.²⁷² Similarly, with a pentamer of phenanthrene attached to a DNA strand, Häner and coworkers introduced a photonic wire made of DNA modified with cyanine (Cye) dyes in a light-harvesting supramolecular polymer (**Figure 1.26**).²⁷³ Remarkable energy transfer efficiency of 59 % was observed. Our group used DNA discrete nanostructures such as a DNA cube as a template for SCP (**Figure 1.26**).²⁷⁴ It triggered higher-order assembly. For example, spherical aggregates with a hydrophobic SCP core and DNA nanocubes as the corona were formed. Specific functionalization of the opposite face of the cubes with fluorophores turned these “super-micelles” into light-harvesting structures. In a later report, more parameters such as the number of precision polymers on the cubic scaffold, their spatial

orientation, the number of monomers in each SCP and the sequence of two different monomers in each oligo were studied systematically.²⁷⁵ Finally, the binding affinity of multivalent sequence-defined dendrimer-bearing DNA nanocubes to albumin was tuned based on the spatial organization of the SCP.²⁷⁶ Thus, DNA nanotechnology-based nanostructures can be used as directional scaffolds for the attachment of SCPs. This spatial organization has a great influence on the conjugates self-assembly, opening new opportunities in materials design.

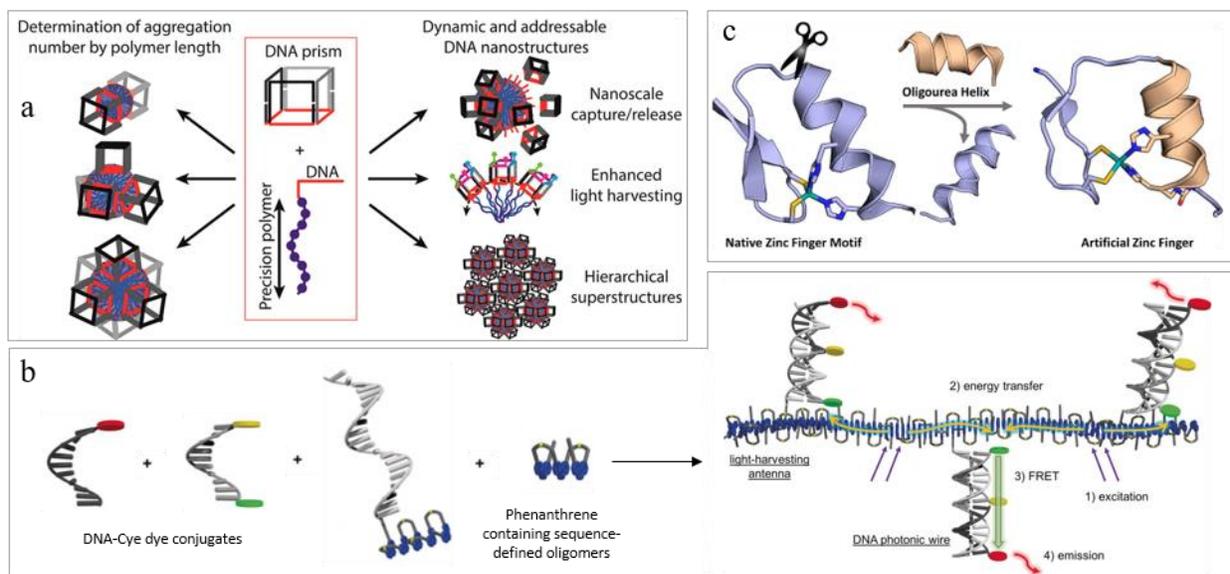


Figure 1.26. Biopolymer-unnatural SCP hybrid structures. (a) Light-harvesting micelles of addressable DNA nanocubes. Adapted with permission from ref. 274. (b) DNA photonic wire on a light-harvesting supramolecular structure made of sequence-defined phenanthrene trimers. Adapted with permission from ref. 273. (c) Replacement of a zinc finger α -helix with an unnatural oligourea of similar size. Adapted with permission from ref. 271.

1.4.5 Overview of SCP applications

Current technologies have applications in information storage, catalysis, molecular recognition, drug discovery, drug delivery, materials design and sensing platforms. Advanced sequence design tools such as combinatorial strategies and the use of hybrid structures enable the synthesis of more complex and functional nanostructures. However, in key fields such as catalysis, unnatural SCP are not outperforming their natural counterparts. This thesis aims at bringing sequence-controlled polymers slightly closer to their ultimate goal.

1.5 Context and scope of this thesis

In terms of synthesis, the goal with SCPs is to access a method that combines perfect sequence control, high DP and versatile building blocks. The two first criteria are fulfilled by phosphoramidite chemistry on solid-phase supports. Therefore, further investigation of sequence-defined oligo(phosphodiester)s synthesis, self-assembly and applications is of particular interest. Moreover, phosphoramidite chemistry is the strategy of choice to synthesize oligonucleotides. Therefore, phosphoramidite-based oligomers are easily attached to nucleic acids, enabling the concomitant use of DNA-amphiphiles, DNA nanotechnology and DNA-encoding tools.

The third criterium is not fulfilled: most sequence-defined oligo(phosphodiester)s structurally different from nucleic acids were made with only two or three different phosphoramidite monomers.^{121,123,124} In this thesis, we aim at exploring the gradual extension of the monomer alphabet for sequence-defined oligo(phosphodiester)s and at showing it leads to novel supramolecular constructs and applications.

In Chapter 2, a novel perfluorocarbon (PFC) containing phosphoramidite was developed to build amphiphilic unnatural sequence-defined oligomers and DNA-amphiphiles. High-yielding synthesis of sequence-defined oligomers made of up to 10 units of this novel building block was achieved and led to “DNA-Teflon” hybrids. These were shown to self-assemble in aqueous solvents. The unique properties of PFC allowed to impart new properties to DNA-Teflon micelles, PFC-modified dsDNA and siRNA.

Chapter 3 describes how the introduction of a novel naphthalene monomer further increased the range of nanostructures accessible with sequence-defined oligo(phosphodiester)s. From four unnatural monomers, a library of sequence-defined oligomers could be designed and synthesized in high yields. They were shown to self-assemble as micelles of different sizes and 2D nanosheets. The balance of the hydrophobic and fluororous effects as well as π - π stacking could be studied to deduce self-assembly rules of precision oligo(phosphodiester)s in aqueous solvent. These conclusions are of fundamental importance for further design of self-assembling sequence-defined polymers.

Expansion of the number of monomers available for sequence-defined oligo(phosphodiester)s is shown in Chapter 4. From two modular platform molecules, a variety of novel phosphoramidite

monomers were made in two steps and high yields. These novel building blocks are fully compatible with DNA and sequence-defined oligo(phosphodiester) automated synthesis. Examples of carbohydrate, alkyne and amino acid containing monomers were synthesized and attached to oligomers, proving the versatility of this strategy.

The new variety of monomers available opens up a large number of opportunities with sequence-defined oligo(phosphodiester)s. In Chapter 5, a combinatorial synthetic method to be able to find sequences of interest for molecular recognition applications is described. The split-and-pool strategy allowed the synthesis of a 300,000 member DNA-encoded sequence-defined oligomer library. Contrary to most DNA-encoded libraries, it is made of oligomers rationally designed to resemble an existing thrombin binding aptamer. This method paves the way to the discovery of numerous unnatural sequence-defined oligomers that can recognize relevant biological or chemical targets.

1.6 References

- (1) Wolfenden, R.; Snider, M. J. The Depth of Chemical Time and the Power of Enzymes as Catalysts. *Acc. Chem. Res.* **2001**, *34* (12), 938–945.
- (2) Rasouli, H.; Farzaei, M. H.; Khodarahmi, R. Polyphenols and Their Benefits: A Review. *Int. J. Food Prop.* **2017**, 1–42.
- (3) Panza, M.; Pistorio, S. G.; Stine, K. J.; Demchenko, A. V. Automated Chemical Oligosaccharide Synthesis: Novel Approach to Traditional Challenges. *Chem. Rev.* **2018**, *118* (17), 8105–8150.
- (4) Müller, S.; Appel, B.; Balke, D.; Hieronymus, R.; Nübel, C. Thirty-Five Years of Research into Ribozymes and Nucleic Acid Catalysis: Where Do We Stand Today? *F1000Research* **2016**, *5*.
- (5) J. D. Watson; Crick, F. H. C. A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171*, 737–738.
- (6) Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. Chromosomal DNA and Its Packaging in the Chromatin Fiber. In *Molecular Biology of the Cell. 4th edition*; Garland Science, 2002.
- (7) Carmean, D.; Ceze, L.; Seelig, G.; Stewart, K.; Strauss, K.; Willsey, M. DNA Data Storage and Hybrid Molecular–Electronic Computing. *Proc. IEEE* **2019**, *107* (1), 63–72.
- (8) Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85* (14), 2149–2154.
- (9) Beaucage, S. L.; Caruthers, M. H. Deoxynucleoside Phosphoramidites—A New Class of Key

- Intermediates for Deoxypolynucleotide Synthesis. *Tetrahedron Lett.* **1981**, 22 (20), 1859–1862.
- (10) Pinheiro, A. V.; Han, D.; Shih, W. M.; Yan, H. Challenges and Opportunities for Structural DNA Nanotechnology. *Nat. Nanotechnol.* **2011**, 6 (12), 763–772.
 - (11) Cello, J.; Paul, A. V.; Wimmer, E. Chemical Synthesis of Poliovirus CDNA: Generation of Infectious Virus in the Absence of Natural Template. *Science* **2002**, 297 (5583), 1016–1018.
 - (12) Kodumal, S. J.; Patel, K. G.; Reid, R.; Menzella, H. G.; Welch, M.; Santi, D. V. Total Synthesis of Long DNA Sequences: Synthesis of a Contiguous 32-Kb Polyketide Synthase Gene Cluster. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, 101 (44), 15573–15578.
 - (13) Hamblin, G. D.; Rahbani, J. F.; Sleiman, H. F. Sequential Growth of Long DNA Strands with User-Defined Patterns for Nanostructures and Scaffolds. *Nat. Commun.* **2015**, 6, 7065.
 - (14) Gräslund, S.; Nordlund, P.; Weigelt, J.; Hallberg, B. M.; Bray, J.; Gileadi, O.; Knapp, S.; Oppermann, U.; Arrowsmith, C.; Hui, R.; et al. Protein Production and Purification. *Nat. Methods* **2008**, 5 (2), 135–146.
 - (15) Ausländer, S.; Ausländer, D.; Fussenegger, M. Synthetic Biology-The Synthesis of Biology. *Angew. Chemie Int. Ed.* **2017**, 56 (23), 6396–6419.
 - (16) Gibson, D. G.; Glass, J. I.; Lartigue, C.; Noskov, V. N.; Chuang, R.-Y.; Algire, M. A.; Benders, G. A.; Montague, M. G.; Ma, L.; Moodie, M. M.; et al. Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome. *Science* **2010**, 329 (5987), 52–56.
 - (17) Church, G. M.; Gao, Y.; Kosuri, S.; Clelland, C. T. Next-Generation Digital Information Storage in DNA. *Science* **2012**, 337 (6102), 1628.
 - (18) Organick, L.; Ang, S. D.; Chen, Y.-J.; Lopez, R.; Yekhanin, S.; Makarychev, K.; Racz, M. Z.; Kamath, G.; Gopalan, P.; Nguyen, B.; et al. Random Access in Large-Scale DNA Data Storage. *Nat. Biotechnol.* **2018**, 36 (3), 242–248.
 - (19) Lutz, J.-F. Writing on Polymer Chains. *Acc. Chem. Res.* **2013**, 46 (11), 2696–2705.
 - (20) Seeman, N. C. Nucleic Acid Junctions and Lattices. *J. Theor. Biol.* **1982**, 99 (2), 237–247.
 - (21) Kallenbach, N. R.; Ma, R.-I.; Seeman, N. C. An Immobile Nucleic Acid Junction Constructed from Oligonucleotides. *Nature* **1983**, 305 (5937), 829–831.
 - (22) Seeman, N. C.; Sleiman, H. F. DNA Nanotechnology. *Nat. Rev. Mater.* **2017**, 3, 17068.
 - (23) Zheng, J.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R.; Constantinou, P. E.; Ginell, S. L.; Mao, C.; Seeman, N. C. From Molecular to Macroscopic via the Rational Design of a Self-Assembled 3D DNA Crystal. *Nature* **2009**, 461 (7260), 74–77.
 - (24) Rothmund, P. W. K. Folding DNA to Create Nanoscale Shapes and Patterns. *Nature* **2006**, 440 (7082), 297–302.
 - (25) Tikhomirov, G.; Petersen, P.; Qian, L. Fractal Assembly of Micrometre-Scale DNA Origami Arrays with Arbitrary Patterns. *Nature* **2017**, 552 (7683), 67–71.
 - (26) McLaughlin, C. K.; Hamblin, G. D.; Sleiman, H. F. Supramolecular {DNA} Assembly. *Chem. Soc. Rev.* **2011**, 40 (12), 5647–5656.

- (27) Shi, J.; Bergstrom, D. E. Assembly of Novel DNA Cycles with Rigid Tetrahedral Linkers. *Angew. Chemie Int. Ed. English* **1997**, *36* (12), 111–113.
- (28) Aldaye, F. A.; Sleiman, H. F. Sequential Self-Assembly of a {DNA} Hexagon as a Template for the Organization of Gold Nanoparticles. *Angew. Chemie Int. Ed.* **2006**, *45* (14), 2204–2209.
- (29) Yang, H.; Sleiman, H. F. Templated Synthesis of Highly Stable, Electroactive, and Dynamic Metal–DNA Branched Junctions. *Angew. Chemie Int. Ed.* **2008**, *47* (13), 2443–2446.
- (30) Jia, F.; Li, H.; Chen, R.; Zhang, K. Self-Assembly of DNA-Containing Copolymers. *Bioconjug. Chem.* **2019**, acs.bioconjchem.9b00067.
- (31) Chien, M.-P.; Rush, A. M.; Thompson, M. P.; Gianneschi, N. C. Programmable Shape-Shifting Micelles. *Angew. Chem. Int. Ed. Engl.* **2010**, *49* (30), 5076–5080.
- (32) Edwardson, T. G. W.; Carneiro, K. M. M.; McLaughlin, C. K.; Serpell, C. J.; Sleiman, H. F. Site-Specific Positioning of Dendritic Alkyl Chains on {DNA} Cages Enables Their Geometry-Dependent Self-Assembly. *Nat. Chem.* **2013**, *5* (10), 868–875.
- (33) Raab, M.; Jusuk, I.; Molle, J.; Buhr, E.; Bodermann, B.; Bergmann, D.; Bosse, H.; Tinnefeld, P. Using DNA Origami Nanorulers as Traceable Distance Measurement Standards and Nanoscopic Benchmark Structures. *Sci. Rep.* **2018**, *8* (1), 1780.
- (34) Lo, P. K.; Karam, P.; Aldaye, F. A.; McLaughlin, C. K.; Hamblin, G. D.; Cosa, G.; Sleiman, H. F. Loading and Selective Release of Cargo in {DNA} Nanotubes with Longitudinal Variation. *Nat. Chem.* **2010**, *2* (4), 319–328.
- (35) Bujold, K. E.; Lacroix, A.; Sleiman, H. F. DNA Nanostructures at the Interface with Biology. *Chem* **2018**, *4* (3), 495–521.
- (36) Bujold, K. E.; Hsu, J. C. C.; Sleiman, H. F. Optimized DNA “Nanosuitcases” for Encapsulation and Conditional Release of siRNA. *J. Am. Chem. Soc.* **2016**, *138* (42), 14030–14038.
- (37) Paterson, B. M.; Roberts, B. E.; Kuff, E. L. Structural Gene Identification and Mapping by DNA-mRNA Hybrid-Arrested Cell-Free Translation. *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74* (10), 4370–4374.
- (38) Bennett, C. F. Therapeutic Antisense Oligonucleotides Are Coming of Age. *Annu. Rev. Med.* **2019**, *70* (1), 307–321.
- (39) Shen, X.; Corey, D. R. Chemistry, Mechanism and Clinical Status of Antisense Oligonucleotides and Duplex RNAs. *Nucleic Acids Res.* **2018**, *46* (4), 1584–1600.
- (40) Mulamba, G. B.; Hu, A.; Azad, R. F.; Anderson, K. P.; Coen, D. M. Human Cytomegalovirus Mutant with Sequence-Dependent Resistance to the Phosphorothioate Oligonucleotide Fomivirsen (ISIS 2922). *Antimicrob. Agents Chemother.* **1998**, *42* (4), 971–973.
- (41) Wood, H. FDA Approves Patisiran to Treat Hereditary Transthyretin Amyloidosis. *Nat. Rev. Neurol.* **2018**, *14* (10), 570–570.
- (42) Ma, Q.; Lee, D.; Tan, Y. Q.; Wong, G.; Gao, Z. Synthetic Genetic Polymers: Advances and Applications. *Polym. Chem.* **2016**, *7* (33), 5199–5216.

- (43) Riley, M.; Vermerris, W.; Riley, M. K.; Vermerris, W. Recent Advances in Nanomaterials for Gene Delivery—A Review. *Nanomaterials* **2017**, *7* (5), 94.
- (44) Alagia, A.; Eritja, R. siRNA and RNAi Optimization. *Wiley Interdiscip. Rev. RNA* **2016**, *7* (3), 316–329.
- (45) Sperry, B. W.; Tang, W. H. W. Amyloid Heart Disease: Genetics Translated into Disease-Modifying Therapy. *Heart* **2017**, *103* (11), 812–817.
- (46) Dunn, M. R.; Jimenez, R. M.; Chaput, J. C. Analysis of Aptamer Discovery and Technology. *Nat. Rev. Chem.* **2017**, *1* (10), 0076.
- (47) Ellington, A. D.; Szostak, J. W. In Vitro Selection of RNA Molecules That Bind Specific Ligands. *Nature* **1990**, *346* (6287), 818–822.
- (48) Tuerk, C.; Gold, L. Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase. *Science* **1990**, *249* (4968), 505–510.
- (49) Rohloff, J. C.; Gelinas, A. D.; Jarvis, T. C.; Ochsner, U. A.; Schneider, D. J.; Gold, L.; Janjic, N. Nucleic Acid Ligands With Protein-like Side Chains: Modified Aptamers and Their Use as Diagnostic and Therapeutic Agents. *Mol. Ther. - Nucleic Acids* **2014**, *3*, e201.
- (50) Anosova, I.; Kowal, E. A.; Dunn, M. R.; Chaput, J. C.; Van Horn, W. D.; Egli, M. The Structural Diversity of Artificial Genetic Polymers. *Nucleic Acids Res.* **2016**, *44* (3), 1007–1021.
- (51) Gold, L.; Ayers, D.; Bertino, J.; Bock, C.; Bock, A.; Brody, E. N.; Carter, J.; Dalby, A. B.; Eaton, B. E.; Fitzwater, T.; et al. Aptamer-Based Multiplexed Proteomic Technology for Biomarker Discovery. *PLoS One* **2010**, *5* (12), e15004.
- (52) Tolle, F.; Brändle, G. M.; Matzner, D.; Mayer, G. A Versatile Approach Towards Nucleobase-Modified Aptamers. *Angew. Chemie Int. Ed.* **2015**, *54* (37), 10971–10974.
- (53) Chen, Z.; Lichtor, P. A.; Berliner, A. P.; Chen, J. C.; Liu, D. R. Evolution of Sequence-Defined Highly Functionalized Nucleic Acid Polymers. *Nat. Chem.* **2018**, *10* (4), 420–427.
- (54) Kong, D.; Yeung, W.; Hili, R. In Vitro Selection of Diversely Functionalized Aptamers. *J. Am. Chem. Soc.* **2017**, *139* (40), 13977–13980.
- (55) Technology Platforms - Creative Biogene <https://www.creative-biogene.com/Services/Aptamers/Technology-Platforms.html> (accessed May 1, 2019).
- (56) Jäschke, A. Artificial Ribozymes and Deoxyribozymes. *Curr. Opin. Struct. Biol.* **2001**, *11* (3), 321–326.
- (57) Seelig, B.; Keiper, S.; Stuhlmann, F.; Jäschke, A. Enantioselective Ribozyme Catalysis of a Bimolecular Cycloaddition Reaction. *Angew. Chemie Int. Ed.* **2000**, *39* (24), 4576–4579.
- (58) Travascio, P.; Li, Y.; Sen, D. DNA-Enhanced Peroxidase Activity of a DNA Aptamer-Hemin Complex. *Chem. Biol.* **1998**, *5* (9), 505–517.
- (59) Jimmy Huang, P.-J.; Moon, W. J.; Liu, J. Instantaneous Iodine-Assisted DNAzyme Cleavage of Phosphorothioate RNA. *Biochemistry* **2019**, *58* (5), 422–429.
- (60) Zhou, C.; Avins, J. L.; Klauser, P. C.; Brandsen, B. M.; Lee, Y.; Silverman, S. K. DNA-

- Catalyzed Amide Hydrolysis. *J. Am. Chem. Soc.* **2016**, *138* (7), 2106–2109.
- (61) Pinheiro, V. B.; Taylor, A. I.; Cozens, C.; Abramov, M.; Renders, M.; Zhang, S.; Chaput, J. C.; Wengel, J.; Peak-Chew, S.-Y.; McLaughlin, S. H.; et al. Synthetic Genetic Polymers Capable of Heredity and Evolution. *Science* **2012**, *336* (6079), 341–344.
- (62) Taylor, A. I.; Holliger, P. Directed Evolution of Artificial Enzymes (XNAzymes) from Diverse Repertoires of Synthetic Genetic Polymers. *Nat. Protoc.* **2015**, *10* (10), 1625–1642.
- (63) Taylor, A. I.; Pinheiro, V. B.; Smola, M. J.; Morgunov, A. S.; Peak-Chew, S.; Cozens, C.; Weeks, K. M.; Herdewijn, P.; Holliger, P. Catalysts from Synthetic Genetic Polymers. *Nature* **2015**, *518* (7539), 427–430.
- (64) Mayo, F. R.; Lewis, F. M. Copolymerization. I. A Basis for Comparing the Behavior of Monomers in Copolymerization; The Copolymerization of Styrene and Methyl Methacrylate. *J. Am. Chem. Soc.* **1944**, *66* (9), 1594–1601.
- (65) Dunn, A. S.; Melville, H. W. Synthesis of ‘Block’ Copolymers. *Nature* **1952**, *169* (4304), 699–700.
- (66) Szwarc, M.; Levy, M.; Milkovich, R. Polymerization Initiated by Electron Transfer to Monomer. A New Method of Formation of Block Polymers. *J. Am. Chem. Soc.* **1956**, *78* (11), 2656–2657.
- (67) Schacher, F. H.; Rupa, P. A.; Manners, I. Functional Block Copolymers: Nanostructured Materials with Emerging Applications. *Angew. Chemie Int. Ed.* **2012**, *51* (32), 7898–7921.
- (68) Bates, F. S.; Hillmyer, M. A.; Lodge, T. P.; Bates, C. M.; Delaney, K. T.; Fredrickson, G. H. Multiblock Polymers: Panacea or Pandora’s Box? *Science* **2012**, *336* (6080), 434–440.
- (69) Matyjaszewski, K.; Ziegler, M. J.; Arehart, S. V.; Greszta, D.; Pakula, T. Gradient Copolymers by Atom Transfer Radical Copolymerization. *J. Phys. Org. Chem.* **2000**, *13* (12), 775–786.
- (70) Lutz, J.-F.; Ouchi, M.; Liu, D. R.; Sawamoto, M. Sequence-Controlled Polymers. *Science* **2013**, *341* (6146), 1238149.
- (71) Lutz, J.-F. Defining the Field of Sequence-Controlled Polymers. *Macromol. Rapid Commun.* **2017**, *38* (24), 1700582.
- (72) Solleder, S. C.; Schneider, R. V.; Wetzlar, K. S.; Boukis, A. C.; Meier, M. A. R. Recent Progress in the Design of Monodisperse, Sequence-Defined Macromolecules. *Macromol. Rapid Commun.* **2017**, *38* (9), 1600711.
- (73) Sworen, J. C.; Smith, J. A.; Berg, J. M.; Wagener, K. B. Modeling Branched Polyethylene: Copolymers Possessing Precisely Placed Ethyl Branches. *J. Am. Chem. Soc.* **2004**, *126* (36), 11238–11246.
- (74) Aitken, B. S.; Buitrago, C. F.; Heffley, J. D.; Lee, M.; Gibson, H. W.; Winey, K. I.; Wagener, K. B. Precision Ionomers: Synthesis and Thermal/Mechanical Characterization. *Macromolecules* **2012**, *45* (2), 681–687.
- (75) Berthet, M.-A.; Zarafshani, Z.; Pfeifer, S.; Lutz, J.-F. Facile Synthesis of Functional Periodic

- Copolymers: A Step toward Polymer-Based Molecular Arrays. *Macromolecules* **2010**, *43* (1), 44–50.
- (76) Stayshich, R. M.; Meyer, T. Y. New Insights into Poly(Lactic-Co-Glycolic Acid) Microstructure: Using Repeating Sequence Copolymers to Decipher Complex NMR and Thermal Behavior. *J. Am. Chem. Soc.* **2010**, *132* (31), 10920–10934.
- (77) Satoh, K.; Ozawa, S.; Mizutani, M.; Nagai, K.; Kamigaito, M. Sequence-Regulated Vinyl Copolymers by Metal-Catalysed Step-Growth Radical Polymerization. *Nat. Commun.* **2010**, *1*, 6.
- (78) Wei, B.; Li, W.; Zhao, Z.; Qin, A.; Hu, R.; Tang, B. Z. Metal-Free Multicomponent Tandem Polymerizations of Alkynes, Amines, and Formaldehyde toward Structure- and Sequence-Controlled Luminescent Polyheterocycles. *J. Am. Chem. Soc.* **2017**, *139* (14), 5075–5084.
- (79) Xi, W.; Pattanayak, S.; Wang, C.; Fairbanks, B.; Gong, T.; Wagner, J.; Kloxin, C. J.; Bowman, C. N. Clickable Nucleic Acids: Sequence-Controlled Periodic Copolymer/Oligomer Synthesis by Orthogonal Thiol-X Reactions. *Angew. Chemie Int. Ed.* **2015**, *54* (48), 14462–14467.
- (80) Gerke, C.; Siegfeld, P.; Schaper, K.; Hartmann, L. Enabling Directional Sequence-Control via Step-Growth Polymerization of Heterofunctionalized Precision Macromonomers. *Macromol. Rapid Commun.* **2019**, *40* (3), 1800735.
- (81) Yang, C.; Flynn, J. P.; Niu, J. Facile Synthesis of Sequence-Regulated Synthetic Polymers Using Orthogonal SuFEx and CuAAC Click Reactions. *Angew. Chemie Int. Ed.* **2018**, *57* (49), 16194–16199.
- (82) Yu, L.; Zhang, Z.; You, Y.-Z.; Hong, C.-Y. Synthesis of Sequence-Controlled Polymers via Sequential Thiol-Ene and Amino-Yne Click Reactions in One Pot. *Eur. Polym. J.* **2018**, *103*, 80–87.
- (83) Zhang, Z.; Tan, Z.-B.; Hong, C.-Y.; Wu, D.-C.; You, Y.-Z. One-Pot Sequential Multicomponent Reaction and a Multicomponent Polymerization Method for the Synthesis of Topologically Different Polymers. *Polym. Chem.* **2016**, *7* (7), 1468–1474.
- (84) Hibi, Y.; Ouchi, M.; Sawamoto, M. Sequence-Regulated Radical Polymerization with a Metal-Templated Monomer: Repetitive ABA Sequence by Double Cyclopolymerization. *Angew. Chemie Int. Ed.* **2011**, *50* (32), 7434–7437.
- (85) Hibi, Y.; Tokuoka, S.; Terashima, T.; Ouchi, M.; Sawamoto, M. Design of AB Divinyl “Template Monomers” toward Alternating Sequence Control in Metal-Catalyzed Living Radical Polymerization. *Polym. Chem.* **2011**, *2* (2), 341–347.
- (86) Soejima, T.; Satoh, K.; Kamigaito, M. Main-Chain and Side-Chain Sequence-Regulated Vinyl Copolymers by Iterative Atom Transfer Radical Additions and 1:1 or 2:1 Alternating Radical Copolymerization. *J. Am. Chem. Soc.* **2016**, *138* (3), 944–954.
- (87) Zhang, J.; Matta, M. E.; Hillmyer, M. A. Synthesis of Sequence-Specific Vinyl Copolymers by Regioselective ROMP of Multiply Substituted Cyclooctenes. *ACS Macro Lett.* **2012**, *1* (12), 1383–1387.

- (88) Gutekunst, W. R.; Hawker, C. J. A General Approach to Sequence-Controlled Polymers Using Macrocyclic Ring Opening Metathesis Polymerization. *J. Am. Chem. Soc.* **2015**, *137* (25), 8038–8041.
- (89) Short, A. L.; Fang, C.; Nowalk, J. A.; Weiss, R. M.; Liu, P.; Meyer, T. Y. Cis -Selective Metathesis to Enhance the Living Character of Ring-Opening Polymerization: An Approach to Sequenced Copolymers. *ACS Macro Lett.* **2018**, *7* (7), 858–862.
- (90) Pfeifer, S.; Lutz, J.-F. A Facile Procedure for Controlling Monomer Sequence Distribution in Radical Chain Polymerizations. *J. Am. Chem. Soc.* **2007**, *129* (31), 9542–9543.
- (91) Pfeifer, S.; Lutz, J.-F. Development of a Library of *N*-Substituted Maleimides for the Local Functionalization of Linear Polymer Chains. *Chem. - A Eur. J.* **2008**, *14* (35), 10949–10957.
- (92) Chan-Seng, D.; Zamfir, M.; Lutz, J.-F. Polymer-Chain Encoding: Synthesis of Highly Complex Monomer Sequence Patterns by Using Automated Protocols. *Angew. Chemie Int. Ed.* **2012**, *51* (49), 12254–12257.
- (93) Zamfir, M.; Lutz, J.-F. Ultra-Precise Insertion of Functional Monomers in Chain-Growth Polymerizations. *Nat. Commun.* **2012**, *3*, 1138.
- (94) Zhang, Z.; Zeng, T.-Y.; Xia, L.; Hong, C.-Y.; Wu, D.-C.; You, Y.-Z. Synthesis of Polymers with On-Demand Sequence Structures via Dually Switchable and Interconvertible Polymerizations. *Nat. Commun.* **2018**, *9* (1), 2577.
- (95) O'Reilly, R. K.; Turberfield, A. J.; Wilks, T. R. The Evolution of DNA-Templated Synthesis as a Tool for Materials Discovery. *Acc. Chem. Res.* **2017**, *50* (10), 2496–2509.
- (96) Böhler, C.; Nielsen, P. E.; Orgel, L. E. Template Switching between PNA and RNA Oligonucleotides. *Nature* **1995**, *376* (6541), 578–581.
- (97) Li, X.; Zhan, Z. Y. J.; Knipe, R.; Lynn, D. G. DNA-Catalyzed Polymerization. *J. Am. Chem. Soc.* **2002**, *124* (5), 746–747.
- (98) Kleiner, R. E.; Brudno, Y.; Birnbaum, M. E.; Liu, D. R. DNA-Templated Polymerization of Side-Chain-Functionalized Peptide Nucleic Acid Aldehydes. *J. Am. Chem. Soc.* **2008**, *130* (14), 4646–4659.
- (99) Niu, J.; Hili, R.; Liu, D. R. Enzyme-Free Translation of DNA into Sequence-Defined Synthetic Polymers Structurally Unrelated to Nucleic Acids. *Nat. Chem.* **2013**, *5* (4), 282–292.
- (100) Chen, W.; Schuster, G. B. Precise Sequence Control in Linear and Cyclic Copolymers of 2,5-Bis(2-Thienyl)Pyrrole and Aniline by DNA-Programmed Assembly. *J. Am. Chem. Soc.* **2013**, *135* (11), 4438–4449.
- (101) McKee, M. L.; Milnes, P. J.; Bath, J.; Stulz, E.; Turberfield, A. J.; O'Reilly, R. K. Multistep DNA-Templated Reactions for the Synthesis of Functional Sequence Controlled Oligomers. *Angew. Chemie Int. Ed.* **2010**, *49* (43), 7948–7951.
- (102) Hansen, M. H.; Blakskjær, P.; Petersen, L. K.; Hansen, T. H.; Højfeldt, J. W.; Gothelf, K. V.; Hansen, N. J. V. A Yoctoliter-Scale DNA Reactor for Small-Molecule Evolution. *J. Am. Chem.*

- Soc.* **2009**, *131* (3), 1322–1327.
- (103) He, Y.; Liu, D. R. A Sequential Strand-Displacement Strategy Enables Efficient Six-Step DNA-Templated Synthesis. *J. Am. Chem. Soc.* **2011**, *133* (26), 9972–9975.
- (104) Usanov, D. L.; Chan, A. I.; Maianti, J. P.; Liu, D. R. Second-Generation DNA-Templated Macrocyclic Libraries for the Discovery of Bioactive Small Molecules. *Nat. Chem.* **2018**, *10* (7), 704–714.
- (105) Cao, C.; Zhao, P.; Li, Z.; Chen, Z.; Huang, Y.; Bai, Y.; Li, X. A DNA-Templated Synthesis of Encoded Small Molecules by DNA Self-Assembly. *Chem. Commun.* **2014**, *50* (75), 10997–10999.
- (106) Milnes, P. J.; McKee, M. L.; Bath, J.; Song, L.; Stulz, E.; Turberfield, A. J.; O'Reilly, R. K. Sequence-Specific Synthesis of Macromolecules Using DNA-Templated Chemistry. *Chem. Commun. (Camb)*. **2012**, *48* (45), 5614–5616.
- (107) He, Y.; Liu, D. R. Autonomous Multistep Organic Synthesis in a Single Isothermal Solution Mediated by a DNA Walker. *Nat. Nanotechnol.* **2010**, *5* (11), 778–782.
- (108) Meng, W.; Muscat, R. A.; McKee, M. L.; Milnes, P. J.; El-Sagheer, A. H.; Bath, J.; Davis, B. G.; Brown, T.; O'Reilly, R. K.; Turberfield, A. J. An Autonomous Molecular Assembler for Programmable Chemical Synthesis. *Nat. Chem.* **2016**, *8* (6), 542–548.
- (109) Lewandowski, B.; De Bo, G.; Ward, J. W.; Pappmeyer, M.; Kuschel, S.; Aldegunde, M. J.; Gramlich, P. M. E.; Heckmann, D.; Goldup, S. M.; D'Souza, D. M.; et al. Sequence-Specific Peptide Synthesis by an Artificial Small-Molecule Machine. *Science* **2013**, *339* (6116), 189–193.
- (110) De Bo, G.; Kuschel, S.; Leigh, D. A.; Lewandowski, B.; Pappmeyer, M.; Ward, J. W. Efficient Assembly of Threaded Molecular Machines for Sequence-Specific Synthesis. *J. Am. Chem. Soc.* **2014**, *136* (15), 5811–5814.
- (111) Minoda, M.; Sawamoto, M.; Higashimura, T. Sequence-Regulated Oligomers and Polymers by Living Cationic Polymerization. *Polym. Bull.* **1990**, *23* (2), 133–139.
- (112) Tong, X.; Guo, B.; Huang, Y. Toward the Synthesis of Sequence-Controlled Vinyl Copolymers. *Chem. Commun.* **2011**, *47* (5), 1455–1457.
- (113) Vandenberg, J.; Reekmans, G.; Adriaensens, P.; Junkers, T. Synthesis of Sequence-Defined Acrylate Oligomers via Photo-Induced Copper-Mediated Radical Monomer Insertions. *Chem. Sci.* **2015**, *6* (10), 5753–5761.
- (114) Oh, D.; Ouchi, M.; Nakanishi, T.; Ono, H.; Sawamoto, M. Iterative Radical Addition with a Special Monomer Carrying Bulky and Convertible Pendant: A New Concept toward Controlling the Sequence for Vinyl Polymers. *ACS Macro Lett.* **2016**, *5* (6), 745–749.
- (115) Hibi, Y.; Ouchi, M.; Sawamoto, M. A Strategy for Sequence Control in Vinyl Polymers via Iterative Controlled Radical Cyclization. *Nat. Commun.* **2016**, *7* (1), 11064.
- (116) Vandenberg, J.; Reekmans, G.; Adriaensens, P.; Junkers, T. Synthesis of Sequence Controlled Acrylate Oligomers via Consecutive RAFT Monomer Additions. *Chem. Commun. (Camb)*.

- 2013**, 49 (88), 10358–10360.
- (117) Houshyar, S.; Keddie, D. J.; Moad, G.; Mulder, R. J.; Saubern, S.; Tsanaktsidis, J. The Scope for Synthesis of Macro-RAFT Agents by Sequential Insertion of Single Monomer Units. *Polym. Chem.* **2012**, 3 (7), 1879.
- (118) Aerts, A.; Lewis, R. W.; Zhou, Y.; Malic, N.; Moad, G.; Postma, A. Light-Induced RAFT Single Unit Monomer Insertion in Aqueous Solution-Toward Sequence-Controlled Polymers. *Macromol. Rapid Commun.* **2018**, 39 (19), 1800240.
- (119) Xu, J.; Fu, C.; Shanmugam, S.; Hawker, C. J.; Moad, G.; Boyer, C. Synthesis of Discrete Oligomers by Sequential PET-RAFT Single-Unit Monomer Insertion. *Angew. Chemie Int. Ed.* **2017**, 56 (29), 8376–8383.
- (120) Huang, Z.; Noble, B. B.; Corrigan, N.; Chu, Y.; Satoh, K.; Thomas, D. S.; Hawker, C. J.; Moad, G.; Kamigaito, M.; Coote, M. L.; et al. Discrete and Stereospecific Oligomers Prepared by Sequential and Alternating Single Unit Monomer Insertion. *J. Am. Chem. Soc.* **2018**, 140 (41), 13392–13406.
- (121) Edwardson, T. G. W.; Carneiro, K. M. M.; Serpell, C. J.; Sleiman, H. F. An Efficient and Modular Route to Sequence-Defined Polymers Appended to {DNA}. *Angew. Chem. Int. Ed. Engl.* **2014**, 53 (18), 4567–4571.
- (122) Lin, P.; Ganesan, A. Solid-Phase Synthesis of Peptidomimetic Oligomers with a Phosphodiester Backbone. *Bioorg. Med. Chem. Lett.* **1998**, 8 (5), 511–514.
- (123) Al Ouahabi, A.; Charles, L.; Lutz, J.-F. Synthesis of Non-Natural Sequence-Encoded Polymers Using Phosphoramidite Chemistry. *J. Am. Chem. Soc.* **2015**, 137 (16), 5629–5635.
- (124) Häner, R.; Garo, F.; Wenger, D.; Malinovskii, V. L. Oligopyrenotides: Abiotic, Polyanionic Oligomers with Nucleic Acid-like Structural Properties. *J. Am. Chem. Soc.* **2010**, 132 (21), 7466–7471.
- (125) Will, D. W.; Pritchard, C. E.; Brown, T. The Synthesis of Oligonucleotides That Contain 2,4-Dinitrophenyl Reporter Groups. *Carbohydr. Res.* **1992**, 216, 315–322.
- (126) Guzaev, A.; Salo, H.; Azhayev, A.; Lönnberg, H. Novel Non-Nucleosidic Building Blocks for the Preparation of Multilabeled Oligonucleotides. *Bioconjug. Chem.* **1996**, 7 (2), 240–248.
- (127) Ouahabi, A. Al; Kotera, M.; Charles, L.; Lutz, J.-F. Synthesis of Monodisperse Sequence-Coded Polymers with Chain Lengths above DP100. *ACS Macro Lett.* **2015**, 4 (10), 1077–1080.
- (128) And, E. E. B.; Dervan, P. B. Solid Phase Synthesis of Polyamides Containing Imidazole and Pyrrole Amino Acids. *J. Am. Chem. Soc.* **1996**, 118 (26), 6141–6146.
- (129) König, H. M.; Gorelik, T.; Kolb, U.; Kilbinger, A. F. M. Supramolecular PEG-Co-Oligo(p-Benzamide)s Prepared on a Peptide Synthesizer. *J. Am. Chem. Soc.* **2007**, 129 (3), 704–708.
- (130) Cho, C. Y.; Moran, E. J.; Cherry, S. R.; Stephans, J. C.; Fodor, S. P.; Adams, C. L.; Sundaram, A.; Jacobs, J. W.; Schultz, P. G. An Unnatural Biopolymer. *Science* **1993**, 261 (5126), 1303–1305.

- (131) Wender, P. A.; Rothbard, J. B.; Jessop, T. C.; Kreider, E. L.; Wylie, B. L. Oligocarbamate Molecular Transporters: Design, Synthesis, and Biological Evaluation of a New Class of Transporters for Drug Delivery. *J. Am. Chem. Soc.* **2002**, *124* (45), 13382–13383.
- (132) Burgess, K.; Ibarzo, J.; Linthicum, D. S.; Russell, D. H.; Shin, H.; Shitangkoon, A.; Totani, R.; Zhang, A. J. Solid Phase Syntheses of Oligoureas. *J. Am. Chem. Soc.* **1997**, *119* (7), 1556–1564.
- (133) Douat-Casassus, C.; Pulka, K.; Claudon, P.; Guichard, G. Microwave-Enhanced Solid-Phase Synthesis of *N*, *N'*-Linked Aliphatic Oligoureas and Related Hybrids. *Org. Lett.* **2012**, *14* (12), 3130–3133.
- (134) Jönsson, D.; Undén, A. Repetitive Solid-Phase Synthesis of Polyamines. *Tetrahedron Lett.* **2002**, *43* (17), 3125–3128.
- (135) Umezawa, N.; Horai, Y.; Imamura, Y.; Kawakubo, M.; Nakahira, M.; Kato, N.; Muramatsu, A.; Yoshikawa, Y.; Yoshikawa, K.; Higuchi, T. Structurally Diverse Polyamines: Solid-Phase Synthesis and Interaction with DNA. *ChemBioChem* **2015**, *16* (12), 1811–1819.
- (136) Brümmer, O.; Clapham, B.; Janda, K. D. Solid-Phase Synthesis of Oligoesters Using a JandaJel™ Resin. *Tetrahedron Lett.* **2001**, *42* (12), 2257–2259.
- (137) Fyles, T. M.; Hu, C. W.; Luong, H. Solid-Phase Synthesis of Oligoester Ion Channels. *J. Org. Chem.* **2006**, *71* (22), 8545–8551.
- (138) Angelo, N. G.; Arora, P. S. Nonpeptidic Foldamers from Amino Acids: Synthesis and Characterization of 1,3-Substituted Triazole Oligomers. *J. Am. Chem. Soc.* **2005**, *127* (49), 17134–17135.
- (139) Nelson, J. C.; Young, J. K.; Moore, J. S. Solid-Phase Synthesis of Phenylacetylene Oligomers Utilizing a Novel 3-Propyl-3-(Benzyl-Supported) Triazene Linkage. *J. Org. Chem.* **1996**, *61* (23), 8160–8168.
- (140) Briehn, C. A.; Kirschbaum, T.; Bäuerle, P. Polymer-Supported Synthesis of Regioregular Head-to-Tail-Coupled Oligo(3-Arylthiophene)s Utilizing a Traceless Silyl Linker. *J. Org. Chem.* **2000**, *65* (2), 352–359.
- (141) Simon, R. J.; Kania, R. S.; Zuckermann, R. N.; Huebner, V. D.; Jewell, D. A.; Banville, S.; Ng, S.; Wang, L.; Rosenberg, S.; Marlowe, C. K.; et al. Peptoids: A Modular Approach to Drug Discovery. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (20), 9367–9371.
- (142) Zuckermann, R. N.; Kerr, J. M.; Kent, S. B. H.; Moos, W. H. Efficient Method for the Preparation of Peptoids [Oligo(*N*-Substituted Glycines)] by Submonomer Solid-Phase Synthesis. *J. Am. Chem. Soc.* **1992**, *114* (26), 10646–10647.
- (143) Knight, A. S.; Zhou, E. Y.; Francis, M. B.; Zuckermann, R. N. Sequence Programmable Peptoid Polymers for Diverse Materials Applications. *Adv. Mater.* **2015**, *27* (38), 5665–5691.
- (144) Morimoto, J.; Fukuda, Y.; Sando, S. Solid-Phase Synthesis of β -Peptoids with Chiral Backbone Substituents Using Reductive Amination. *Org. Lett.* **2017**, *19* (21), 5912–5915.
- (145) Sable, G. A.; Lee, K. J.; Shin, M.-K.; Lim, H.-S. Submonomer Strategy toward Divergent Solid-Phase Synthesis of α -ABpeptoids. *Org. Lett.* **2018**, *20* (9), 2526–2529.

- (146) Murphy, J. E.; Uno, T.; Hamer, J. D.; Cohen, F. E.; Dwarki, V.; Zuckermann, R. N.; Schultz, P. G. A Combinatorial Approach to the Discovery of Efficient Cationic Peptoid Reagents for Gene Delivery. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 1517–1522.
- (147) Farrera-Sinfreu, J.; Aviñó, A.; Navarro, I.; Aymamí, J.; Beteta, N. G.; Varón, S.; Pérez-Tomás, R.; Castillo-Avila, W.; Eritja, R.; Albericio, F.; et al. Design, Synthesis and Antiproliferative Properties of Oligomers with Chromophore Units Linked by Amide Backbones. *Bioorg. Med. Chem. Lett.* **2008**, *18* (7), 2440–2444.
- (148) Robertson, E. J.; Battigelli, A.; Proulx, C.; Mannige, R. V.; Haxton, T. K.; Yun, L.; Whitelam, S.; Zuckermann, R. N. Design, Synthesis, Assembly, and Engineering of Peptoid Nanosheets. *Acc. Chem. Res.* **2016**, *49* (3), 379–389.
- (149) Martens, S.; Van den Begin, J.; Madder, A.; Du Prez, F. E.; Espeel, P. Automated Synthesis of Monodisperse Oligomers, Featuring Sequence Control and Tailored Functionalization. *J. Am. Chem. Soc.* **2016**, *138* (43), 14182–14185.
- (150) Espeel, P.; Carrette, L. L. G.; Bury, K.; Capenberghs, S.; Martins, J. C.; Du Prez, F. E.; Madder, A. Multifunctionalized Sequence-Defined Oligomers from a Single Building Block. *Angew. Chem. Int. Ed. Engl.* **2013**, *52* (50), 13261–13264.
- (151) Trinh, T. T.; Oswald, L.; Chan-Seng, D.; Lutz, J.-F. Synthesis of Molecularly Encoded Oligomers Using a Chemoselective “AB + CD” Iterative Approach. *Macromol. Rapid Commun.* **2014**, *35* (2), 141–145.
- (152) Roy, R. K.; Meszynska, A.; Laure, C.; Charles, L.; Verchin, C.; Lutz, J.-F. Design and Synthesis of Digitally Encoded Polymers That Can Be Decoded and Erased. *Nat. Commun.* **2015**, *6*, 7237.
- (153) Gunay, U. S.; Petit, B. E.; Karamessini, D.; Al Ouahabi, A.; Amalian, J.-A.; Chendo, C.; Bouquey, M.; Gigmes, D.; Charles, L.; Lutz, J.-F. Chemoselective Synthesis of Uniform Sequence-Coded Polyurethanes and Their Use as Molecular Tags. *Chem* **2016**, *1* (1), 114–126.
- (154) Cavallo, G.; Al Ouahabi, A.; Oswald, L.; Charles, L.; Lutz, J.-F. Orthogonal Synthesis of “Easy-to-Read” Information-Containing Polymers Using Phosphoramidite and Radical Coupling Steps. *J. Am. Chem. Soc.* **2016**, *138* (30), 9417–9420.
- (155) Grate, J. W.; Mo, K.-F.; Daily, M. D. Triazine-Based Sequence-Defined Polymers with Side-Chain Diversity and Backbone-Backbone Interaction Motifs. *Angew. Chem. Int. Ed. Engl.* **2016**, *55* (12), 3925–3930.
- (156) Zhang, Z.; You, Y.; Hong, C. Multicomponent Reactions and Multicomponent Cascade Reactions for the Synthesis of Sequence-Controlled Polymers. *Macromol. Rapid Commun.* **2018**, *39* (23), 1800362.
- (157) Solleder, S. C.; Meier, M. A. R. Sequence Control in Polymer Chemistry through the Passerini Three-Component Reaction. *Angew. Chem. Int. Ed. Engl.* **2014**, *53* (3), 711–714.
- (158) Solleder, S. C.; Zengel, D.; Wetzl, K. S.; Meier, M. A. R. A Scalable and High-Yield Strategy for the Synthesis of Sequence-Defined Macromolecules. *Angew. Chem. Int. Ed. Engl.* **2016**, *55* (3), 1204–1207.

- (159) Solleder, S. C.; Martens, S.; Espeel, P.; Du Prez, F.; Meier, M. A. R. Combining Two Methods of Sequence Definition in a Convergent Approach: Scalable Synthesis of Highly Defined and Multifunctionalized Macromolecules. *Chem. - A Eur. J.* **2017**, *23* (56), 13906–13909.
- (160) Lewis, J. E. M.; Winn, J.; Cera, L.; Goldup, S. M. Iterative Synthesis of Oligo[*n*]Rotaxanes in Excellent Yield. *J. Am. Chem. Soc.* **2016**, *138* (50), 16329–16336.
- (161) Dong, R.; Liu, R.; Gaffney, P. R. J.; Schaepertoens, M.; Marchetti, P.; Williams, C. M.; Chen, R.; Livingston, A. G. Sequence-Defined Multifunctional Polyethers via Liquid-Phase Synthesis with Molecular Sieving. *Nat. Chem.* **2019**, *11* (2), 136–145.
- (162) Zhang, W. Fluorous Tagging Strategy for Solution-Phase Synthesis of Small Molecules, Peptides and Oligosaccharides. *Curr. Opin. Drug Discov. Devel.* **2004**, *7* (6), 784–797.
- (163) Miriyala, B. Fluorous Methods for the Synthesis of Peptides and Oligonucleotides. *Top. Curr. Chem.* **2012**, *308*, 105–133.
- (164) Porel, M.; Alabi, C. A. Sequence-Defined Polymers via Orthogonal Allyl Acrylamide Building Blocks. *J. Am. Chem. Soc.* **2014**, *136* (38), 13162–13165.
- (165) Porel, M.; Thornlow, D. N.; Phan, N. N.; Alabi, C. A. Sequence-Defined Bioactive Macrocycles via an Acid-Catalysed Cascade Reaction. *Nat. Chem.* **2016**, *8* (6), 590–596.
- (166) Kanasty, R. L.; Vegas, A. J.; Ceo, L. M.; Maier, M.; Charisse, K.; Nair, J. K.; Langer, R.; Anderson, D. G. Sequence-Defined Oligomers from Hydroxyproline Building Blocks for Parallel Synthesis Applications. *Angew. Chemie Int. Ed.* **2016**, *55* (33), 9529–9533.
- (167) Li, J.; Ballmer, S. G.; Gillis, E. P.; Fujii, S.; Schmidt, M. J.; Palazzolo, A. M. E.; Lehmann, J. W.; Morehouse, G. F.; Burke, M. D. Synthesis of Many Different Types of Organic Small Molecules Using One Automated Process. *Science* **2015**, *6227*, (347), 1221–1226.
- (168) Hayatsu, H.; Khorana, H. G. Studies on Polynucleotides. LXXII. Deoxyribooligonucleotide Synthesis on a Polymer Support. *J. Am. Chem. Soc.* **1967**, *89* (15), 3880–3887.
- (169) Bayer, E.; Mutter, M. Liquid Phase Synthesis of Peptides. *Nature* **1972**, *237* (5357), 512–513.
- (170) Gravert, D. J.; Datta, A.; Wentworth, P.; Janda, K. D. Soluble Supports Tailored for Organic Synthesis: Parallel Polymer Synthesis via Sequential Normal/Living Free Radical Processes. *J. Am. Chem. Soc.* **1998**, *120* (37), 9481–9495.
- (171) Pfeifer, S.; Zarafshani, Z.; Badi, N.; Lutz, J.-F. Liquid-Phase Synthesis of Block Copolymers Containing Sequence-Ordered Segments. *J. Am. Chem. Soc.* **2009**, *131* (26), 9195–9197.
- (172) Szweda, R.; Chendo, C.; Charles, L.; Baxter, P. N. W.; Lutz, J.-F. Synthesis of Oligoarylacetylenes with Defined Conjugated Sequences Using Tailor-Made Soluble Polymer Supports. *Chem. Commun.* **2017**, *53* (59), 8312–8315.
- (173) Zydziak, N.; Feist, F.; Huber, B.; Mueller, J. O.; Barner-Kowollik, C. Photo-Induced Sequence Defined Macromolecules via Hetero Bifunctional Synthons. *Chem. Commun.* **2015**, *51* (10), 1799–1802.
- (174) Paynter, O. I.; Simmonds, D. J.; Whiting, M. C. The Synthesis of Long-Chain Unbranched

- Aliphatic Compounds by Molecular Doubling. *J. Chem. Soc. Chem. Commun.* **1982**, 0 (20), 1165.
- (175) Jones, L. R.; Schumm, J. S.; Tour, J. M. Rapid Solution and Solid Phase Syntheses of Oligo(1,4-Phenylene Ethynylene)s with Thioester Termini: Molecular Scale Wires with Alligator Clips. Derivation of Iterative Reaction Efficiencies on a Polymer Support. *J. Org. Chem.* **1997**, 62 (5), 1388–1410.
- (176) Takizawa, K.; Tang, C.; Hawker, C. J. Molecularly Defined Caprolactone Oligomers and Polymers: Synthesis and Characterization. *J. Am. Chem. Soc.* **2008**, 130 (5), 1718–1726.
- (177) Binauld, S.; Damiron, D.; Connal, L. A.; Hawker, C. J.; Drockenmuller, E. Precise Synthesis of Molecularly Defined Oligomers and Polymers by Orthogonal Iterative Divergent/Convergent Approaches. *Macromol. Rapid Commun.* **2011**, 32 (2), 147–168.
- (178) Wang, Q.; Qu, Y.; Tian, H.; Geng, Y.; Wang, F. Iterative Binomial Synthesis of Monodisperse Polyfluorenes up to 64-Mers and Their Chain-Length-Dependent Properties. *Macromolecules* **2011**, 44 (6), 1256–1260.
- (179) Amir, F.; Jia, Z.; Monteiro, M. J. Sequence Control of Macromers via Iterative Sequential and Exponential Growth. *J. Am. Chem. Soc.* **2016**, 138 (51), 16600–16603.
- (180) Huang, Z.; Zhao, J.; Wang, Z.; Meng, F.; Ding, K.; Pan, X.; Zhou, N.; Li, X.; Zhang, Z.; Zhu, X. Combining Orthogonal Chain-End Deprotections and Thiol-Maleimide Michael Coupling: Engineering Discrete Oligomers by an Iterative Growth Strategy. *Angew. Chemie Int. Ed.* **2017**, 56 (44), 13612–13617.
- (181) Barnes, J. C.; Ehrlich, D. J. C.; Gao, A. X.; Leibfarth, F. A.; Jiang, Y.; Zhou, E.; Jamison, T. F.; Johnson, J. A. Iterative Exponential Growth of Stereo- and Sequence-Controlled Polymers. *Nat. Chem.* **2015**, 7 (10), 810–815.
- (182) Leibfarth, F. A.; Johnson, J. A.; Jamison, T. F. Scalable Synthesis of Sequence-Defined, Unimolecular Macromolecules by Flow-IEG. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, 112 (34), 10617–10622.
- (183) Cole, J. P.; Hanlon, A. M.; Rodriguez, K. J.; Berda, E. B. Protein-like Structure and Activity in Synthetic Polymers. *J. Polym. Sci. Part A Polym. Chem.* **2017**, 55 (2), 191–206.
- (184) ten Brummelhuis, N.; Wilke, P.; Börner, H. G. Identification of Functional Peptide Sequences to Lead the Design of Precision Polymers. *Macromol. Rapid Commun.* **2017**, 38 (24), 1700632.
- (185) Meier, M. A. R.; Barner-Kowollik, C. A New Class of Materials: Sequence-Defined Macromolecules and Their Emerging Applications. *Adv. Mater.* **2019**, 1806027.
- (186) Austin, M. J.; Rosales, A. M. Tunable Biomaterials from Synthetic, Sequence-Controlled Polymers. *Biomater. Sci.* **2019**, 7 (2), 490–505.
- (187) Lutz, J.-F. Writing on Polymer Chains. *Acc. Chem. Res.* **2013**, 46 (11), 2696–2705.
- (188) Laure, C.; Karamessini, D.; Milenkovic, O.; Charles, L.; Lutz, J.-F. Coding in 2D: Using Intentional Dispersity to Enhance the Information Capacity of Sequence-Coded Polymer

- Barcodes. *Angew. Chemie Int. Ed.* **2016**, *55* (36), 10722–10725.
- (189) Szweda, R.; Tschopp, M.; Felix, O.; Decher, G.; Lutz, J. F. Sequences of Sequences: Spatial Organization of Coded Matter through Layer-by-Layer Assembly of Digital Polymers. *Angew. Chemie - Int. Ed.* **2018**, *57* (48), 15817–15821.
- (190) Boukis, A. C.; Meier, M. A. R. Data Storage in Sequence-Defined Macromolecules via Multicomponent Reactions. *Eur. Polym. J.* **2018**, *104*, 32–38.
- (191) Martens, S.; Landuyt, A.; Espeel, P.; Devreese, B.; Dawyndt, P.; Du Prez, F. Multifunctional Sequence-Defined Macromolecules for Chemical Data Storage. *Nat. Commun.* **2018**, *9* (1), 4451.
- (192) Cavallo, G.; Poyer, S.; Amalian, J.-A.; Dufour, F.; Burel, A.; Carapito, C.; Charles, L.; Lutz, J.-F. Cleavable Binary Dyads: Simplifying Data Extraction and Increasing Storage Density in Digital Polymers. *Angew. Chemie Int. Ed.* **2018**, *57* (21), 6266–6269.
- (193) Amalian, J.-A.; Al Ouahabi, A.; Cavallo, G.; König, N. F.; Poyer, S.; Lutz, J.-F.; Charles, L. Controlling the Structure of Sequence-Defined Poly (Phosphodiester)s for Optimal MS/MS Reading of Digital Information. *J. Mass Spectrom.* **2017**, *52* (11), 788–798.
- (194) Zhu, Z.; Cardin, C. J.; Gan, Y.; Colquhoun, H. M. Sequence-Selective Assembly of Tweezer Molecules on Linear Templates Enables Frameshift-Reading of Sequence Information. *Nat. Chem.* **2010**, *2* (8), 653–660.
- (195) Boukhet, M.; König, N. F.; Ouahabi, A. Al; Baaken, G.; Lutz, J.-F.; Behrends, J. C. Translocation of Precision Polymers through Biological Nanopores. *Macromol. Rapid Commun.* **2017**, *38* (24), 1700680.
- (196) Kasianowicz, J. J.; Brandin, E.; Branton, D.; Deamer, D. W. Characterization of Individual Polynucleotide Molecules Using a Membrane Channel. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93* (24), 13770–13773.
- (197) Karamessini, D.; Petit, B. E.; Bouquey, M.; Charles, L.; Lutz, J.-F. Identification-Tagging of Methacrylate-Based Intraocular Implants Using Sequence Defined Polyurethane Barcodes. *Adv. Funct. Mater.* **2017**, *27* (3), 1604595.
- (198) Karamessini, D.; Simon-Yarza, T.; Poyer, S.; Konishcheva, E.; Charles, L.; Letourneur, D.; Lutz, J.-F. Abiotic Sequence-Coded Oligomers as Efficient In Vivo Taggants for the Identification of Implanted Materials. *Angew. Chemie Int. Ed.* **2018**, *57* (33), 10574–10578.
- (199) Mai, Y.; Eisenberg, A. Self-Assembly of Block Copolymers. *Chem. Soc. Rev.* **2012**, *41* (18), 5969.
- (200) Nam, K. T.; Shelby, S. A.; Choi, P. H.; Marciel, A. B.; Chen, R.; Tan, L.; Chu, T. K.; Mesch, R. A.; Lee, B.-C.; Connolly, M. D.; et al. Free-Floating Ultrathin Two-Dimensional Crystals from Sequence-Specific Peptoid Polymers. *Nat. Mater.* **2010**, *9* (5), 454–460.
- (201) Mannige, R. V.; Haxton, T. K.; Proulx, C.; Robertson, E. J.; Battigelli, A.; Butterfoss, G. L.; Zuckermann, R. N.; Whitlam, S. Peptoid Nanosheets Exhibit a New Secondary-Structure Motif. *Nature* **2015**, *526* (7573), 415–420.

- (202) Robertson, E. J.; Nehls, E. M.; Zuckermann, R. N. Structure–Rheology Relationship in Nanosheet-Forming Peptoid Monolayers. *Langmuir* **2016**, *32* (46), 12146–12158.
- (203) Patterson, A. L.; Danielsen, S. P. O.; Segalman, R. A.; Yu, B.; Davidson, E. C.; Fredrickson, G. H. Sequence Effects on Block Copolymer Self-Assembly through Tuning Chain Conformation and Segregation Strength Utilizing Sequence-Defined Polypeptoids. *Macromolecules* **2019**, *52* (3), 1277–1286.
- (204) Battigelli, A.; Kim, J. H.; Dehigaspitiya, D. C.; Proulx, C.; Robertson, E. J.; Murray, D. J.; Rad, B.; Kirshenbaum, K.; Zuckermann, R. N. Glycosylated Peptoid Nanosheets as a Multivalent Scaffold for Protein Recognition. *ACS Nano* **2018**, *12* (3), 2455–2465.
- (205) Sun, J.; Jiang, X.; Lund, R.; Downing, K. H.; Balsara, N. P.; Zuckermann, R. N. Self-Assembly of Crystalline Nanotubes from Monodisperse Amphiphilic Diblock Copolypeptoid Tiles. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (15), 3954–3959.
- (206) Yan, F.; Liu, L.; Walsh, T. R.; Gong, Y.; El-Khoury, P. Z.; Zhang, Y.; Zhu, Z.; De Yoreo, J. J.; Engelhard, M. H.; Zhang, X.; et al. Controlled Synthesis of Highly-Branched Plasmonic Gold Nanoparticles through Peptoid Engineering. *Nat. Commun.* **2018**, *9* (1), 2327.
- (207) Vybornyi, M.; Rudnev, A.; Häner, R. Assembly of Extra-Large Nanosheets by Supramolecular Polymerization of Amphiphilic Pyrene Oligomers in Aqueous Solution. *Chem. Mater.* **2015**, *27* (4), 1426–1431.
- (208) Bösch, C. D.; Langenegger, S. M.; Häner, R. Light-Harvesting Nanotubes Formed by Supramolecular Assembly of Aromatic Oligophosphates. *Angew. Chemie Int. Ed.* **2016**, *55* (34), 9961–9964.
- (209) Zhang, W.; Huang, M.; Su, H.; Zhang, S.; Yue, K.; Dong, X.-H.; Li, X.; Liu, H.; Zhang, S.; Wesdemiotis, C.; et al. Toward Controlled Hierarchical Heterogeneities in Giant Molecules with Precisely Arranged Nano Building Blocks. *ACS Cent. Sci.* **2016**, *2* (1), 48–54.
- (210) van Genabeek, B.; de Waal, B. F. M.; Gosens, M. M. J.; Pitet, L. M.; Palmans, A. R. A.; Meijer, E. W. Synthesis and Self-Assembly of Discrete Dimethylsiloxane–Lactic Acid Diblock Co-Oligomers: The Dononacotamer and Its Shorter Homologues. *J. Am. Chem. Soc.* **2016**, *138* (12), 4210–4218.
- (211) Golder, M. R.; Jiang, Y.; Teichen, P. E.; Nguyen, H. V.-T.; Wang, W.; Milos, N.; Freedman, S. A.; Willard, A. P.; Johnson, J. A. Stereochemical Sequence Dictates Unimolecular Diblock Copolymer Assembly. *J. Am. Chem. Soc.* **2018**, *140* (5), 1596–1599.
- (212) van Genabeek, B.; de Waal, B. F. M.; Ligt, B.; Palmans, A. R. A.; Meijer, E. W. Dispersity under Scrutiny: Phase Behavior Differences between Disperse and Discrete Low Molecular Weight Block Co-Oligomers. *ACS Macro Lett.* **2017**, *6* (7), 674–678.
- (213) Oschmann, B.; Lawrence, J.; Schulze, M. W.; Ren, J. M.; Anastasaki, A.; Luo, Y.; Nothling, M. D.; Pester, C. W.; Delaney, K. T.; Connal, L. A.; et al. Effects of Tailored Dispersity on the Self-Assembly of Dimethylsiloxane–Methyl Methacrylate Block Co-Oligomers. *ACS Macro Lett.* **2017**, *6* (7), 668–673.

- (214) Das, A.; Petkau-Milroy, K.; Klerks, G.; van Genabeek, B.; Lafleur, R. P. M.; Palmans, A. R. A.; Meijer, E. W. Consequences of Dispersity on the Self-Assembly of ABA-Type Amphiphilic Block Co-Oligomers. *ACS Macro Lett.* **2018**, *7* (5), 546–550.
- (215) Guichard, G.; Huc, I. Synthetic Foldamers. *Chem. Commun. (Camb)*. **2011**, *47* (21), 5933–5941.
- (216) Gellman, S. H. Foldamers: A Manifesto. *Acc. Chem. Res.* **1998**, *31* (4), 173–180.
- (217) Rodriguez, J. M.; Hamilton, A. D. Benzoylurea Oligomers: Synthetic Foldamers That Mimic Extended α Helices. *Angew. Chemie Int. Ed.* **2007**, *46* (45), 8614–8617.
- (218) Bao, C.; Kauffmann, B.; Gan, Q.; Srinivas, K.; Jiang, H.; Huc, I. Converting Sequences of Aromatic Amino Acid Monomers into Functional Three-Dimensional Structures: Second-Generation Helical Capsules. *Angew. Chemie Int. Ed.* **2008**, *47* (22), 4153–4156.
- (219) Guo, L.; Almeida, A. M.; Zhang, W.; Reidenbach, A. G.; Choi, S. H.; Guzei, I. A.; Gellman, S. H. Helix Formation in Preorganized β/γ -Peptide Foldamers: Hydrogen-Bond Analogy to the α -Helix without α -Amino Acid Residues. *J. Am. Chem. Soc.* **2010**, *132* (23), 7868–7869.
- (220) Gan, Q.; Ferrand, Y.; Bao, C.; Kauffmann, B.; Grelard, A.; Jiang, H.; Huc, I. Helix-Rod Host-Guest Complexes with Shuttling Rates Much Faster than Disassembly. *Science* **2011**, *331* (6021), 1172–1175.
- (221) Wang, X.; Wicher, B.; Ferrand, Y.; Huc, I. Orchestrating Directional Molecular Motions: Kinetically Controlled Supramolecular Pathways of a Helical Host on Rodlike Guests. *J. Am. Chem. Soc.* **2017**, *139* (27), 9350–9358.
- (222) Chandramouli, N.; Ferrand, Y.; Lautrette, G.; Kauffmann, B.; Mackereth, C. D.; Laguerre, M.; Dubreuil, D.; Huc, I. Iterative Design of a Helically Folded Aromatic Oligoamide Sequence for the Selective Encapsulation of Fructose. *Nat. Chem.* **2015**, *7* (4), 334–341.
- (223) Lombardo, C. M.; Collie, G. W.; Pulka-Ziach, K.; Rosu, F.; Gabelica, V.; Mackereth, C. D.; Guichard, G. Anatomy of an Oligourea Six-Helix Bundle. *J. Am. Chem. Soc.* **2016**, *138* (33), 10522–10530.
- (224) Collie, G. W.; Bailly, R.; Pulka-Ziach, K.; Lombardo, C. M.; Mauran, L.; Taib-Maamar, N.; Dessolin, J.; Mackereth, C. D.; Guichard, G. Molecular Recognition within the Cavity of a Foldamer Helix Bundle: Encapsulation of Primary Alcohols in Aqueous Conditions. *J. Am. Chem. Soc.* **2017**, *139* (17), 6128–6137.
- (225) Müller, M. M.; Windsor, M. A.; Pomerantz, W. C.; Gellman, S. H.; Hilvert, D. A Rationally Designed Aldolase Foldamer. *Angew. Chemie Int. Ed.* **2009**, *48* (5), 922–925.
- (226) Maayan, G.; Ward, M. D.; Kirshenbaum, K. Folded Biomimetic Oligomers for Enantioselective Catalysis. *Proc. Natl. Acad. Sci.* **2009**, *106* (33), 13679–13684.
- (227) Bécart, D.; Diemer, V.; Salaün, A.; Oiarbide, M.; Nelli, Y. R.; Kauffmann, B.; Fischer, L.; Palomo, C.; Guichard, G. Helical Oligourea Foldamers as Powerful Hydrogen Bonding Catalysts for Enantioselective C–C Bond-Forming Reactions. *J. Am. Chem. Soc.* **2017**, *139* (36), 12524–12532.
- (228) Le Bailly, B. A. F.; Clayden, J.; Patchornik, A.; Leito, I.; Morris, G. A.; Webb, S. J.; Clayden,

- J.; Beadle, J. D.; Clayden, J.; Rossi, F.; et al. Dynamic Foldamer Chemistry. *Chem. Commun.* **2016**, 52 (27), 4852–4863.
- (229) Le Bailly, B. A. F.; Byrne, L.; Clayden, J. Refoldable Foldamers: Global Conformational Switching by Deletion or Insertion of a Single Hydrogen Bond. *Angew. Chemie Int. Ed.* **2016**, 55 (6), 2132–2136.
- (230) Shaller, A. D.; Wang, W.; Li, A.; Moyna, G.; Han, J. J.; Helms, G. L.; Li, A. D. Q. Sequence-Controlled Oligomers Fold into Nanosolenoids and Impart Unusual Optical Properties. *Chem. - A Eur. J.* **2011**, 17 (30), 8350–8362.
- (231) Lyon, C. K.; Prasher, A.; Hanlon, A. M.; Tuten, B. T.; Tooley, C. A.; Frank, P. G.; Berda, E. B. A Brief User's Guide to Single-Chain Nanoparticles. *Polym. Chem.* **2015**, 6 (2), 181–197.
- (232) Schmidt, B. V. K. J.; Fechler, N.; Falkenhagen, J.; Lutz, J.-F. Controlled Folding of Synthetic Polymer Chains through the Formation of Positionable Covalent Bridges. *Nat. Chem.* **2011**, 3 (3), 234–238.
- (233) Terashima, T.; Mes, T.; De Greef, T. F. A.; Gillissen, M. A. J.; Besenius, P.; Palmans, A. R. A.; Meijer, E. W. Single-Chain Folding of Polymers for Catalytic Systems in Water. *J. Am. Chem. Soc.* **2011**, 133 (13), 4742–4745.
- (234) Liu, Y.; Pauloehrl, T.; Presolski, S. I.; Albertazzi, L.; Palmans, A. R. A.; Meijer, E. W. Modular Synthetic Platform for the Construction of Functional Single-Chain Polymeric Nanoparticles: From Aqueous Catalysis to Photosensitization. *J. Am. Chem. Soc.* **2015**, 137 (40), 13096–13105.
- (235) Mavila, S.; Rozenberg, I.; Lemcoff, N. G. A General Approach to Mono- and Bimetallic Organometallic Nanoparticles. *Chem. Sci.* **2014**, 5 (11), 4196–4203.
- (236) Sanchez-Sanchez, A.; Arbe, A.; Colmenero, J.; Pomposo, J. A. Metallo-Folded Single-Chain Nanoparticles with Catalytic Selectivity. *ACS Macro Lett.* **2014**, 3 (5), 439–443.
- (237) Berda, E. B.; Cole, J. P.; Hanlon, A. M.; Mancinelli, J. P.; Reville, E. K.; Rodriguez, K. J.; Lessard, J. J. Single-Chain Nanoparticles Containing Sequence-Defined Segments: Using Primary Structure Control to Promote Secondary and Tertiary Structures in Synthetic Protein Mimics. *Polym. Chem.* **2017**, 8 (38), 5829–5835.
- (238) Norris, B. N.; Zhang, S.; Campbell, C. M.; Auletta, J. T.; Calvo-Marzal, P.; Hutchison, G. R.; Meyer, T. Y. Sequence Matters: Modulating Electronic and Optical Properties of Conjugated Oligomers via Tailored Sequence. *Macromolecules* **2013**, 46 (4), 1384–1392.
- (239) Washington, M. A.; Swiner, D. J.; Bell, K. R.; Fedorchak, M. V.; Little, S. R.; Meyer, T. Y. The Impact of Monomer Sequence and Stereochemistry on the Swelling and Erosion of Biodegradable Poly(Lactic-Co-Glycolic Acid) Matrices. *Biomaterials* **2017**, 117, 66–76.
- (240) Azzarito, V.; Long, K.; Murphy, N. S.; Wilson, A. J. Inhibition of α -Helix-Mediated Protein–protein Interactions Using Designed Molecules. *Nat. Chem.* **2013**, 5 (3), 161–173.
- (241) Lee, E. F.; Sadowsky, J. D.; Smith, B. J.; Czabotar, P. E.; Peterson-Kaufman, K. J.; Colman, P. M.; Gellman, S. H.; Fairlie, W. D. High-Resolution Structural Characterization of a Helical α/β -

- Peptide Foldamer Bound to the Anti-Apoptotic Protein Bcl-x_L. *Angew. Chemie Int. Ed.* **2009**, *48* (24), 4318–4322.
- (242) Yap, J. L.; Cao, X.; Vanommeslaeghe, K.; Jung, K.-Y.; Peddaboina, C.; Wilder, P. T.; Nan, A.; MacKerell, A. D.; Smythe, W. R.; Fletcher, S. Relaxation of the Rigid Backbone of an Oligoamide-Foldamer-Based α -Helix Mimetic: Identification of Potent Bcl-XL Inhibitors. *Org. Biomol. Chem.* **2012**, *10* (15), 2928.
- (243) Czyzewski, A. M.; Jenssen, H.; Fjell, C. D.; Waldbrook, M.; Chongsiriwatana, N. P.; Yuen, E.; Hancock, R. E. W.; Barron, A. E. In Vivo, In Vitro, and In Silico Characterization of Peptoids as Antimicrobial Agents. *PLoS One* **2016**, *11* (2), e0135961.
- (244) Lee, J.; Kang, D.; Choi, J.; Huang, W.; Wadman, M.; Barron, A. E.; Seo, J. Effect of Side Chain Hydrophobicity and Cationic Charge on Antimicrobial Activity and Cytotoxicity of Helical Peptoids. *Bioorg. Med. Chem. Lett.* **2018**, *28* (2), 170–173.
- (245) Lau, J. L.; Dunn, M. K. Therapeutic Peptides: Historical Perspectives, Current Development Trends, and Future Directions. *Bioorg. Med. Chem.* **2018**, *26* (10), 2700–2707.
- (246) Phan, N. N.; Li, C.; Alabi, C. A. Intracellular Delivery via Noncharged Sequence-Defined Cell-Penetrating Oligomers. *Bioconjug. Chem.* **2018**, *29* (8), 2628–2635.
- (247) Troiber, C.; Wagner, E. Nucleic Acid Carriers Based on Precise Polymer Conjugates. *Bioconjug. Chem.* **2011**, *22* (9), 1737–1752.
- (248) Fröhlich, T.; Edinger, D.; Kläger, R.; Troiber, C.; Salcher, E.; Badgular, N.; Martin, I.; Schaffert, D.; Cengizeroglu, A.; Hadwiger, P.; et al. Structure–activity Relationships of SiRNA Carriers Based on Sequence-Defined Oligo (Ethane Amino) Amides. *J. Control. Release* **2012**, *160* (3), 532–541.
- (249) Morys, S.; Urnauer, S.; Spitzweg, C.; Wagner, E. EGFR Targeting and Shielding of PDNA Lipopolyplexes via Bivalent Attachment of a Sequence-Defined PEG Agent. *Macromol. Biosci.* **2018**, *18* (1), 1700203.
- (250) Müller, K.; Klein, P. M.; Heissig, P.; Roidl, A.; Wagner, E. EGF Receptor Targeted Lipo-Oligocation Polyplexes for Antitumoral SiRNA and MiRNA Delivery. *Nanotechnology* **2016**, *27* (46), 464001.
- (251) Dill, K. A.; MacCallum, J. L. The Protein-Folding Problem, 50 Years On. *Science* **2012**, *338* (6110), 1042–1046.
- (252) Huang, P.-S.; Boyken, S. E.; Baker, D. The Coming of Age of de Novo Protein Design. *Nature* **2016**, *537* (7620), 320–327.
- (253) Wetzler, M.; Barron, A. E. Commentary Progress in the de Novo Design of Structured Peptoid Protein Mimics. *Biopolymers* **2011**, *96* (5), 556–560.
- (254) Frank, R. Spot-Synthesis: An Easy Technique for the Positionally Addressable, Parallel Chemical Synthesis on a Membrane Support. *Tetrahedron* **1992**, *48* (42), 9217–9232.
- (255) Furka, Á.; Sebestyén, F.; Asgedom, M.; Dibo, G. General Method for Rapid Synthesis of Multicomponent Peptide Mixtures. *Int. J. Pept. Protein Res.* **1991**, *37* (6), 487–493.

- (256) Schneider, A. C.; Fritz, D.; Vasquez, J. K.; Vollrath, S. B. L.; Blackwell, H. E.; Bräse, S. Microwave-Facilitated SPOT-Synthesis of Antibacterial Dipeptoids. *ACS Comb. Sci.* **2017**, *19* (12), 715–737.
- (257) Mattes, D. S.; Streit, B.; Bhandari, D. R.; Greifenstein, J.; Foertsch, T. C.; Münch, S. W.; Ridder, B.; v. Bojničić-Kninski, C.; Nesterov-Mueller, A.; Spengler, B.; et al. Combinatorial Synthesis of Peptoid Arrays via Laser-Based Stacking of Multiple Polymer Nanolayers. *Macromol. Rapid Commun.* **2018**, 1800533.
- (258) Cho, C. Y.; Youngquist, R. S.; Paikoff, S. J.; Beresini, M. H.; Hebert, A. R.; Berleau, L. T.; Liu, C. W.; Wemmer, D. E.; Keough, T.; Schultz, P. G. Synthesis and Screening of Linear and Cyclic Oligocarbamate Libraries. Discovery of High Affinity Ligands for GPIIb/IIIa. *J. Am. Chem. Soc.* **1998**, *120* (31), 7706–7718.
- (259) Celasun, S.; Remmler, D.; Schwaar, T.; Weller, M. G.; Du Prez, F.; Börner, H. G. Digging into the Sequential Space of Thiolactone Precision Polymers: A Combinatorial Strategy to Identify Functional Domains. *Angew. Chem. Int. Ed.* **2019**, *58* (7), 1960–1964.
- (260) Brenner, S.; Lerner, R. A. Encoded Combinatorial Chemistry. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (12), 5381–5383.
- (261) Needels, M. C.; Jones, D. G.; Tate, E. H.; Heinkel, G. L.; Kochersperger, L. M.; Dower, W. J.; Barrett, R. W.; Gallop, M. A. Generation and Screening of an Oligonucleotide-Encoded Synthetic Peptide Library. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (22), 10700–10704.
- (262) Zhu, Z.; Shaginian, A.; Grady, L. C.; O’Keeffe, T.; Shi, X. E.; Davie, C. P.; Simpson, G. L.; Messer, J. A.; Evindar, G.; Bream, R. N.; et al. Design and Application of a DNA-Encoded Macrocyclic Peptide Library. *ACS Chem. Biol.* **2018**, *13* (1), 53–59.
- (263) Mannocci, L.; Leimbacher, M.; Wichert, M.; Scheuermann, J.; Neri, D. 20 Years of DNA-Encoded Chemical Libraries. *Chem. Commun.* **2011**, *47* (48), 12747.
- (264) Neri, D.; Lerner, R. A. DNA-Encoded Chemical Libraries: A Selection System Based on Endowing Organic Compounds with Amplifiable Information. *Annu. Rev. Biochem.* **2018**, *87* (1), 479–502.
- (265) O’Reilly, R. K.; Turberfield, A. J.; Wilks, T. R. The Evolution of DNA-Templated Synthesis as a Tool for Materials Discovery. *Acc. Chem. Res.* **2017**, *50* (10), 2496–2509.
- (266) Petersen, L. K.; Blakskjær, P.; Chaikuad, A.; Christensen, A. B.; Dietvorst, J.; Holmkvist, J.; Knapp, S.; Kořínek, M.; Larsen, L. K.; Pedersen, A. E.; et al. Novel P38 α MAP Kinase Inhibitors Identified from YoctoReactor DNA-Encoded Small Molecule Library. *Medchemcomm* **2016**, *7* (7), 1332–1339.
- (267) Liu, R.; Lam, K. S. Peptide combinatorial libraries https://schoolbag.info/chemistry/chemical_biology/127.html.
- (268) Carlini, A. S.; Adamiak, L.; Gianneschi, N. C. Biosynthetic Polymers as Functional Materials. *Macromolecules* **2016**, *49* (12), 4379–4394.
- (269) Schneider, J. A.; Craven, T. W.; Kasper, A. C.; Yun, C.; Haugbro, M.; Briggs, E. M.; Svetlov,

- V.; Nudler, E.; Knaut, H.; Bonneau, R.; et al. Design of Peptoid-Peptide Macrocycles to Inhibit the β -Catenin TCF Interaction in Prostate Cancer. *Nat. Commun.* **2018**, *9* (1), 4396.
- (270) Rogers, J. M.; Kwon, S.; Dawson, S. J.; Mandal, P. K.; Suga, H.; Huc, I. Ribosomal Synthesis and Folding of Peptide-Helical Aromatic Foldamer Hybrids. *Nat. Chem.* **2018**, *10* (4), 405–412.
- (271) Lombardo, C. M.; Kumar M. V., V.; Douat, C.; Rosu, F.; Mergny, J.-L.; Salgado, G. F.; Guichard, G. Design and Structure Determination of a Composite Zinc Finger Containing a Nonpeptide Foldamer Helical Domain. *J. Am. Chem. Soc.* **2019**, *141* (6), 2516–2525.
- (272) Bösch, C. D.; Jevric, J.; Bürki, N.; Probst, M.; Langenegger, S. M.; Häner, R. Supramolecular Assembly of DNA-Phenanthrene Conjugates into Vesicles with Light-Harvesting Properties. *Bioconjug. Chem.* **2018**, *29* (5), 1505–1509.
- (273) Kownacki, M.; Langenegger, S. M.; Liu, S.-X.; Häner, R. Integrating DNA Photonic Wires into Light-Harvesting Supramolecular Polymers. *Angew. Chemie* **2019**, *131* (3), 761–765.
- (274) Serpell, C. J.; Edwardson, T. G. W.; Chidchob, P.; Carneiro, K. M. M.; Sleiman, H. F. Precision Polymers and 3D DNA Nanostructures: Emergent Assemblies from New Parameter Space. *J. Am. Chem. Soc.* **2014**, *136* (44), 15767–15774.
- (275) Chidchob, P.; Edwardson, T. G. W.; Serpell, C. J.; Sleiman, H. F. Synergy of Two Assembly Languages in DNA Nanostructures: Self-Assembly of Sequence-Defined Polymers on DNA Cages. *J. Am. Chem. Soc.* **2016**, *138* (13), 4416–4425.
- (276) Lacroix, A.; Edwardson, T. G. W.; Hancock, M. A.; Dore, M. D.; Sleiman, H. F. Development of DNA Nanostructures for High-Affinity Binding to Human Serum Albumin. *J. Am. Chem. Soc.* **2017**, *139* (21), 7355–7362.

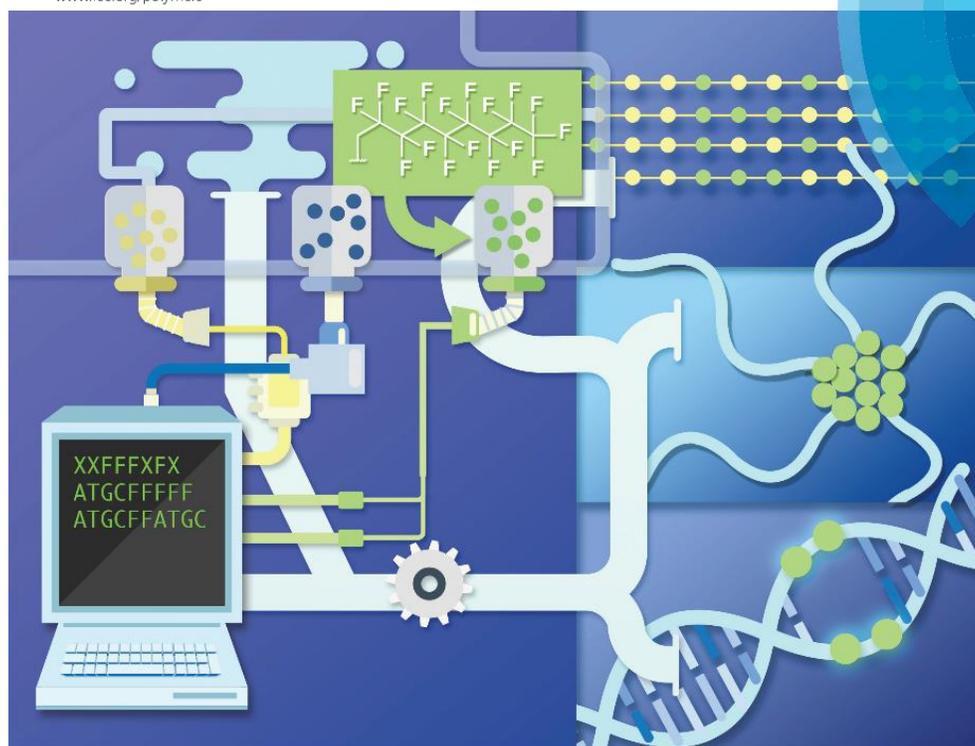
|2|

“DNA-Teflon” Sequence-Defined Oligomers

Volume 7 Number 31 21 August 2016 Pages 4987–5080

Polymer Chemistry

www.rsc.org/polymers



ISSN 1759-9954



PAPER
Hanaei F. Sleiman et al.
“DNA-Teflon” sequence-controlled polymers

175
YEARS

Artwork by Edouard Coune.

2.1 Preface

In the introduction, we highlighted the potential of increasing the number of interactions available in sequence-defined oligo(phosphodiester)s. So far, the introduction of the hydrophobic effect using a 12-carbon alkyl chain monomer in DNA-polymer conjugates has been extensively studied by our group. It led to several applications including the discovery of oligomers that self-assemble depending on their sequence,¹ the higher-order assembly of DNA nanostructures through the hydrophobic interaction,^{2,3} and the synthesis of spherical RNA nanoparticles with gene silencing activity.⁴ Here, we aim at exploring the opportunities related to the introduction of a novel type of interaction.

Perfluorocarbons (PFC) – alkyl chains wherein all C-H bonds have been replaced by C-F bonds – are hydrophobic and oleophobic. Indeed, the highly polarized C-F bonds prevent PFC from forming London intermolecular interactions. Therefore, such molecules tend to show unique supramolecular interactions in a number of self-assembling systems.⁵ The synthesis of a PFC-containing monomer for sequence-defined oligo(phosphodiester)s may confer new properties upon them.

This chapter is mostly composed of work published as «“DNA-Teflon” sequence-controlled polymers» by D de Rochambeau, M Barlóg, TGW Edwardson, JJ Fakhoury, RS Stein, HS Bazzi & HF Sleiman; Polym. Chem. 7, 4998, (2016). The section on DNA duplexes thermodynamics has not been published.

2.2 Contribution of authors

Donatien de Rochambeau codesigned the project and performed all experiments unless listed below, analyzed the results and cowrote the manuscript. **Dr. Maciej Barłóg** synthesized compounds **1**, **2** and **3**. **Dr. Thomas Edwardson** provided training and helped for RNA synthesis. **Dr. Johans Fakhoury** did the cell work and the luciferase assay data collection. **Dr. Robin Stein** helped design and perform the NMR experiments. **Dr. Hanadi Sleiman** codesigned the project, guided interpretation of data and discussion of results and cowrote the manuscript.

2.3 Abstract

Perfluorocarbons (PFCs) are a promising class of molecules for medical applications: they are detectable through ^{19}F nuclear magnetic resonance (NMR) and they assemble separately from water or lipophilic phases, thus bringing unique supramolecular interactions into nanostructures. We report the ready insertion of PFCs into nucleic acids, as well as non-natural oligomers in a sequence-defined fashion. This is achieved via an automated and efficient synthetic pathway using phosphoramidite chemistry. Modulating the PFC tail length of “DNA–Teflon” block co-oligomers resulted in micelles that are almost monodisperse, have a low critical micelle concentration (CMC), are traceable by ^{19}F NMR and are responsive to external stimuli. Strong fluorine–fluorine interactions in DNA duplexes provided nuclease resistance and allowed remarkable melting temperature increases. The Gibbs free energy variations due to the interactions between PFC units could be further studied using van’t Hoff plots. Finally, PFC insertion into siRNA was achieved, and the conjugates were efficient for gene silencing, outlining that these modifications are highly suitable for oligonucleotide therapeutics and bioimaging tools. Perfluorocarbon chains can impart new supramolecular interactions in DNA nanostructures and duplexes and are therefore a valuable addition to the monomers available to make sequence-defined oligo(phosphodiester)s.

2.4 Introduction

Perfluorocarbons (PFCs) offer significant advantages for biological applications. They can be used as highly efficient oxygen carriers,^{6–8} and as labels for non-invasive imaging with magnetic resonance imaging (MRI).^{9,10} Additionally, they can be tailored to improve the efficiency of drug delivery^{11–13} and affinity for cell membranes.¹⁴ Oligonucleotides have emerged as a powerful class

of potential therapeutics, because of their specificity, their ability to address undruggable targets and their inherent biocompatibility.¹⁵ However, their clinical applications have been hampered by poor cellular penetration and stability to nucleases.¹⁶ Perfluorocarbons can help overcome some of these shortcomings. Indeed, it was shown that DNA¹⁷ and peptide-nucleic acid (PNA)¹⁸ modified with a single perfluorocarbon chain exhibit enhanced cellular penetration ability. Given their MRI potential and the possibility to influence nucleic acid biological properties, the efficient synthesis of perfluorocarbon–DNA conjugates would be of great interest towards the design of theranostic DNA structures.¹⁹

As supramolecular building blocks, perfluorocarbons tend to minimize interactions with other compounds by creating a separate phase from both hydrophilic and lipophilic phases; this phenomenon is termed the “fluorous effect”.^{5,20,21} Introducing these units into DNA nanostructures can impart them with the fluorous effect as an orthogonal interaction to Watson–Crick base-pairing, thus increasing the structural range of DNA nanotechnology. ¹⁹F MRI in the case of perfluoro-amphiphiles is known to be very challenging. To achieve high signal intensity, it requires a large number of fluorine atoms. However, aggregation of PFCs into an immobile phase can diminish or even completely suppress the signal.^{22–25} The ability to tune the fluorine location and content in PFC-containing polymers would allow fine adjustment of their assembly and optimization of MRI signals.

Sequence-defined polymer synthesis is a growing area that has recently attracted significant attention.²⁶ Current strategies rely on insertion of specific monomer units into polymer chains through chain-growth polymerization,^{27–29} or using short oligomers with defined sequences as monomers.^{30,31} For complete sequence control, the use of molecular machines,³² DNA-templated synthesis³³ or sequential attachment of monomers^{1,34–36} has been reported. As stated in the first chapter of this thesis, the discovery of novel building blocks to increase the range of materials obtained with sequence-controlled polymers is of great interest. PFC is a good candidate since it could broaden the supramolecular interactions available in sequence-defined oligo(phosphodiester)s.

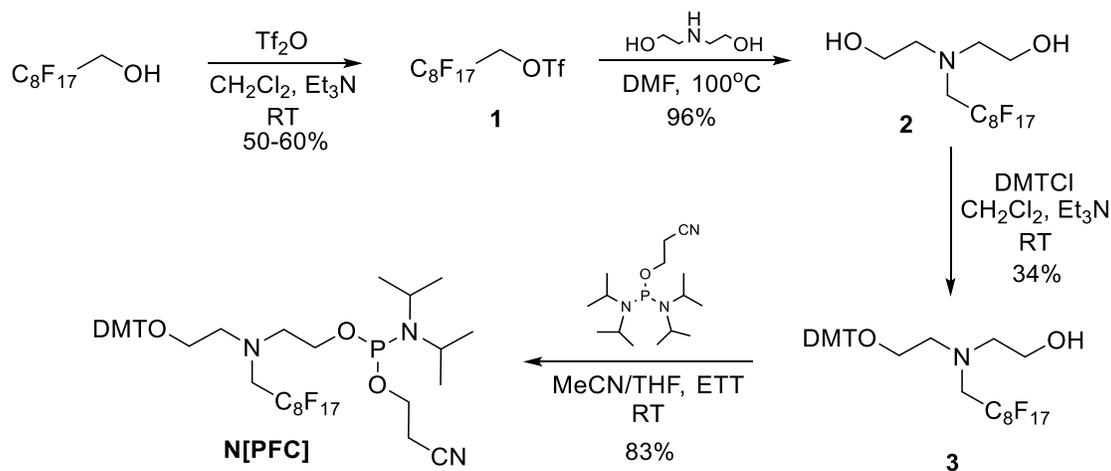
In this chapter, we report a versatile synthesis of sequence-defined oligomers containing perfluorocarbons, using automated phosphoramidite chemistry. This method yields non-natural oligomers containing hydrophilic chains and PFCs in a precise sequence order. We then show the

sequence-defined incorporation of PFC units into DNA strands. Due to the fluorophilic effect, DNA–PFC precision oligomers form micelles at low critical micelle concentrations (CMC) in aqueous media. The micelles have low polydispersity, their size can be switched by adding divalent cations in solution and, surprisingly, they can be readily detected through ^{19}F NMR. We find that the inclusion of two PFCs in a face-to-face arrangement within a DNA duplex results in remarkably high stabilization of DNA towards thermal denaturation.³⁷ The detailed contribution of enthalpy and entropy in the duplex formation is studied in more detail by building van't Hoff plots. PFC units in DNA duplexes also significantly increased nuclease resistance. Finally, we synthesized siRNA (short interfering RNA)–PFC conjugates and showed their ability to effect gene silencing in mammalian cells.

2.5 Results and discussion

2.5.1 Synthesis of PFC-containing sequence-defined oligomers

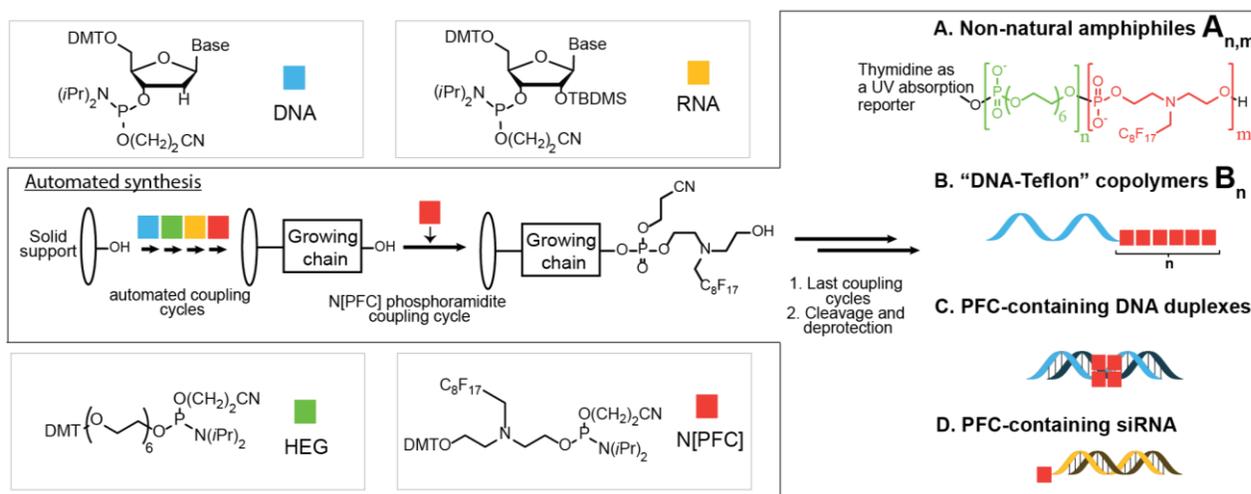
The synthesis of the phosphoramidite **N[PFC]** (**Scheme 2.1**) started with the conversion of a commercial perfluorinated alcohol into its triflate, followed by heating with diethanolamine to give **2**. Monoprotection with dimethoxytrityl chloride and conversion to the phosphoramidite led to **N[PFC]**.



Scheme 2.1. Synthetic pathway to **N[PFC]**.

We first used the **N[PFC]** phosphoramidite to prepare non-natural co-oligomers with a hexaethylene glycol (**HEG**) hydrophilic block and a perfluorocarbon block. **HEG** phosphoramidite (commercially available) and **N[PFC]** were attached sequentially on a solid

support using an automated DNA synthesizer (**Scheme 2.2**, oligomers **A_{n,m}**). In order to make our oligomers UV detectable, we started the synthesis with a thymidine nucleotide on the hydrophilic extremity. Cleavage and deprotection were carried out using ammonium hydroxide at 60 °C for 10 hours. Isolation of the final compounds was performed using reverse phase high performance liquid chromatography (RP-HPLC) and purity was checked by liquid chromatography- mass spectroscopy (LC-MS). Yields were measured with RP-HPLC (**Table 2.1**). The synthetic and purification methods are facile and extremely efficient, and lead to amphiphilic monodisperse polyphosphates with precise numbers and sequences of **HEG** and **PFC** units, and masses up to 5 kDa (**A_{8,4}**) in very good yields.



Scheme 2.2. Synthesis of versatile sequence-defined oligomers with **HEG**, **N[PFC]** and nucleoside phosphoramidites.

We then incorporated **N[PFC]** on the 5'-end of a DNA strand on the automated DNA synthesizer (oligomers **B_n**). Interestingly, we could successfully attach up to ten **N[PFC]** units (equivalent to 6 kDa oligomer) on a single-stranded DNA 19-mer. Cleavage from the solid support, isolation, and determination of yield and purity were carried out as above (**Table 2.2** and **Figure 2.1**). A single **N[PFC]** addition showed excellent coupling yields (>95 %) and grafting several **N[PFC]** was very efficient. This is noteworthy, considering that the attachment of highly hydrophobic units to DNA is often challenging.^{38,39} 1 to 3 **N[PFC]** units were also incorporated at the internal positions of two complementary DNA strands and 1 **N[PFC]** at the 3' end of an RNA strand (sequences and characterization available in the experimental section).

Table 2.1. Yields and ESI-MS characterization of amphiphilic oligomers **A_{n,m}**

Number of HEG (n) and N[PFC] units (m)	HPLC yields ^a (%)	Calculated exact mass [(M-2)/2] (g/mol)	Found exact mass (g/mol)
n=6, m=0	81	2305.83 ^b	2305.82 ^b
n=6, m=2	78	1751.43	1751.44
n=6, m=3	75	2050.94	2050.95
n=8, m=0	83	1496.54	1496.55
n=8, m=2	73	2095.56	2095.57
n=8, m=4	68	2694.57	2694.91

a. Calculated through the integration of the peak associated to the expected product (260 nm detection). b. (M-1) is reported

Table 2.2. Yields and ESI-MS characterization of “DNA-Teflon” hybrids **B_n**.

Number of PFC units (n)	HPLC yields ^a (%)	Calculated exact mass (g mol ⁻¹)	Found exact mass (g mol ⁻¹)
1	95	6364.01	6364.00
2	86	6963.02	6963.00
4	76	8161.05	8161.00
6	73	9362.89 (MW ^[b])	9362.13 (MW)
10	74	11759.61 (MW)	11758.95 (MW)

a. Calculated through the integration of the peak associated to the expected product (260 nm detection). b. MW stands for molecular weight.

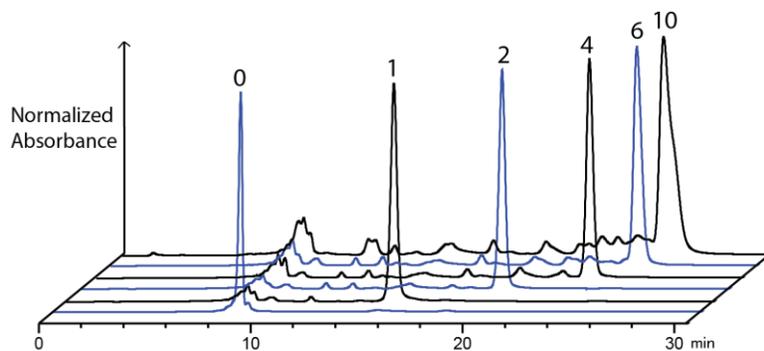


Figure 2.1. Reverse-phase-HPLC traces from crude mixtures of “DNA-Teflon” oligomers (B_n). (UV detection, 260 nm). Numbers on the peaks are the number n of N[PFC] attached onto a single-stranded DNA 19-mer.

2.5.2 “DNA-Teflon” oligomers self-assembly studies

Inspired by nucleic acids nanostructures self-assembly,¹ we decided to first study the potential self-assembly properties of our oligomers in magnesium containing buffers. Figure 2.2 shows atomic force microscopy (AFM) images of the DNA conjugate with 5 and 10 N[PFC] units. Spherical structures of uniform sizes and average diameters (D_{AFM}) of respectively 19.4 ± 3.1 nm and 20.9 ± 3.3 nm were observed. Because of their monodisperse nature and unlike regular polymers, B_5 and B_{10} conjugates likely assemble into spherical micelles of a narrow size distribution, with a perfluorocarbon core and a DNA corona. The features observed by AFM possibly arise from partial collapse of the DNA strands on the mica surface, and protrusion of the perfluoro block above the micelle due to repulsion from the mica.

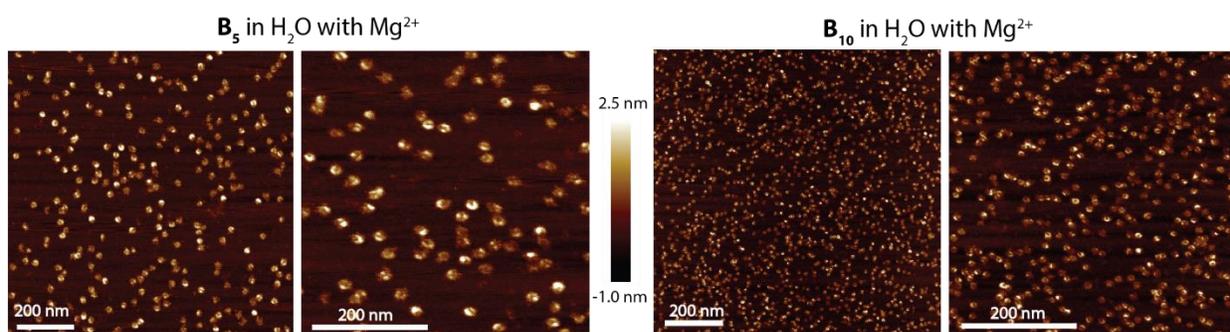


Figure 2.2. Dry AFM images of B_5 and B_{10} . These images have been obtained on mica surface, from solutions of B_5 and B_{10} in a Mg^{2+} containing buffer.

We decided to analyze further the series of B_n conjugates through dynamic light scattering (DLS). It indicated the self-assembly of strands B_4 , B_5 and B_{10} in water at low concentrations (10 μ M,

Table 2.3). Unlike oligo(phosphodiester) micelles reported earlier,^{1,40} “DNA-Teflon” micelles do not require the addition of divalent cation which is of interest for biological applications. Strikingly, the nanostructures are almost monodisperse (polydispersity index, PDI < 16 %). We also found that the addition of magnesium ions significantly alters their assembly. While **B**₂ does not aggregate in pure water, addition of Mg²⁺ triggers assembly. Upon Mg²⁺ addition (7.6 mM), micelles from DNA strands with 4 or 5 PFC units undergo a significant radius decrease within one minute ($D_{DLS} \sim 16$ nm to 13 nm, **Table 2.3**. Summary of DLS results for “DNA-Teflon” co-oligomers. The Mg²⁺ cations likely coordinate the negatively charged phosphate groups between the PFC units and on the DNA strands, thus decreasing their repulsive electrostatic interactions (**Figure 2.3**). It is noteworthy that the precise control over the perfluorocarbon block size of our constructs allows modulation of their assembly properties. DLS allowed us to determine that the critical micellar concentration (CMC) of the conjugates is small (below 10 μ M).

Table 2.3. Summary of DLS results for “DNA-Teflon” co-oligomers. Standard deviations are reported. At 37 °C, the dynamic radius of AT-(N[PFC]₄) micelles is 7.5 ± 0.4 nm in H₂O. NO stands for not observed.

Number of N[PFC] units (n)	Presence of Mg ²⁺ (7.6 mM)	Measured dynamic radius (nm)
0-1	y/n	NO
2	n	NO
	y	12.4 ± 0.2
4	n	16.2 ± 0.8
	y	12.6 ± 0.6
5	n	17.4 ± 0.6
	y	13.0 ± 0.4
10	n	17.0 ± 0.3
	y	17.0 ± 0.1

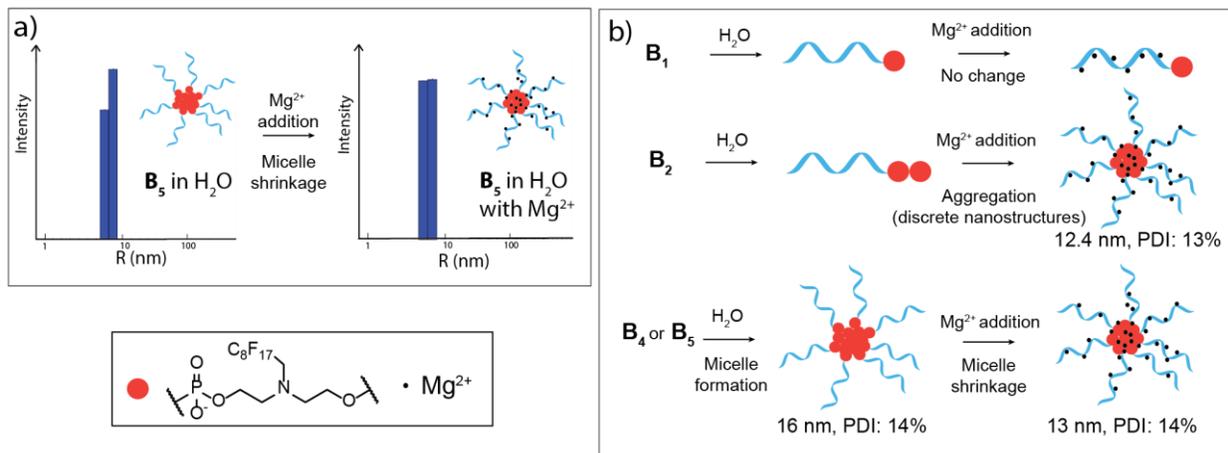


Figure 2.3. Effect of sequence-control and Mg^{2+} addition for “DNA–Teflon” co-oligomers assembly. Left: DLS raw data before and after addition of Mg^{2+} . Right: Schematic representation of DLS data. Micelle diameters and PDI obtained by DLS are reported under the micelles. B_{10} did not show shrinkage upon Mg^{2+} addition.

In addition, we used agarose gel electrophoresis (AGE) to gather more information about our oligomer self-assembly. Indeed, agarose gels have larger pore size than polyacrylamide gels allowing larger constructs such as self-assembled structures to go through the gel. Results suggest that micelles still form at 1 μ M (**Figure 2.4**) from 4 N[PFC] units in non-native conditions (no Mg^{2+}) and from 3 N[PFC] units in native conditions. CMC of “DNA–Teflon” co-oligomers is therefore thought to be under 1 μ M. Interestingly, having 2 or 3 N[PFC] are limit cases in which a smearing band or non-penetrating materials are seen. In these cases, gel electrophoresis is not adapted to gather information on self-assembly. As a conclusion, due to the DNA–Teflon micelles’ narrow polydispersity and low CMC, they are likely to be useful for drug delivery applications of fluorine containing drugs.

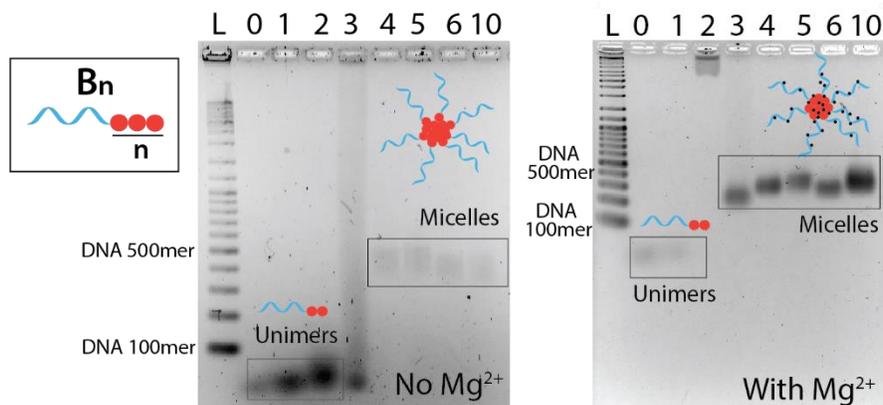


Figure 2.4. Agarose gel electrophoresis analysis of “DNA-Teflon” co-oligomers. Gels are 2.5 % agarose. The ladder (L) used shows the size of DNA strands. The lane number represents the number n of PFC chains in the conjugate. Unimers run faster than a DNA 100mer while micelles have similar mobility than a DNA 500mer.

2.5.3 ^{19}F NMR detection of “DNA-Teflon” oligomers

Given their unique architecture, we examined the ^{19}F NMR properties of our conjugates. Spectra were acquired in D_2O for (i) $\text{N}[\text{PFC}]$ diol (in DMSO) (**2**), (ii) **B**₁, (iii) **B**₄ micelles, and (iv) **B**₁₀ micelles. Surprisingly, even in the core of the micelle structures, $\text{N}[\text{PFC}]$ units were detected through their CF_3 moiety at a low strand concentration ($46 \mu\text{M}$). ^{19}F spin–lattice relaxation (T_1) and spin–spin relaxation (T_2) were measured at 11.7 T (**Table 2.4**). As expected, self-assembly led to a decrease of T_2 from 180 ms for **B**₁ to 7 ms for **B**₄ and 2.0 ms for **B**₁₀. However, these values are still acceptable for potential MRI applications compared to those described elsewhere.²⁵ At the same time, we observed a decrease of T_1 with the number of $\text{N}[\text{PFC}]$ units: $T_1 = 740$ ms, 160 ms and 150 ms respectively for **B**₁, **B**₄ and **B**₁₀. The short T_1 of the micelle samples allows for the possibility of acquiring many scans in short time periods, therefore increasing the signal intensity per unit time.

We were interested in specifically comparing the NMR signal of micelles with a large number of $\text{N}[\text{PFC}]$ per strand (**B**₁₀) to the sensitivity of **B**₄. A higher signal-to-noise ratio (S/N) was observed for **B**₁₀ (24.3) than for **B**₄ (14.4) after fifteen minutes of acquisition, when both samples were analyzed under identical conditions (**Figure 2.5**). Moreover, due to shorter relaxation times for **B**₁₀, signal to noise S/N could be increased to 32.6 in the same amount of time. Modulating the number of $\text{N}[\text{PFC}]$ units through sequence-control and having a poly(phosphodiester) backbone

are particularly promising approaches to control sensitivity for MRI applications. Since the low fluorine background from biological samples allows for ^{19}F MRI quantification,⁴¹ we also determined micelle concentrations using quantitative ^{19}F NMR. We measured the concentrations of NMR samples (ii), (iii), and (iv) with regard to an external reference made with molecule **2**. The results were close to the concentrations obtained by UV absorption even in the case of micelles ($\pm 15\%$, **Table 2.4**) showing that self-assembly does not prevent the retrieval of most of the CF_3 signal. These results highlight the potential of the N[PFC] DNA modifications for bio-imaging with possible micelle quantification.

Table 2.4. ^{19}F NMR relaxation times and micelle concentrations measured.

Molecule	T_1 (s)	T_2 (s)	$[\mathbf{B}_n]_{\text{UV}}^{\text{a}}$ (μM)	$[\mathbf{B}_n]_{\text{NMR}}^{\text{c}}$ (μM)	Ratio $[\mathbf{B}_n]_{\text{NMR}}/[\mathbf{B}_n]_{\text{UV}}$
2 (in DMSO)	1.28 ± 0.09	-	800 ^b	800	1
B₁	0.7 ± 0.1	0.185 ± 0.02	71	81	1.1
B₄	0.161 ± 0.05	0.007 ± 0.002	39	35	0.9
B₁₀	0.148 ± 0.02	0.0020 ± 0.0004	46	40	0.9

a. Concentration determined by UV absorption at 260nm except for **2**. The UV measurements could be influenced by self-assembly. b. Determined by weighing and precisely diluting the compound. c. Concentration determined by NMR.

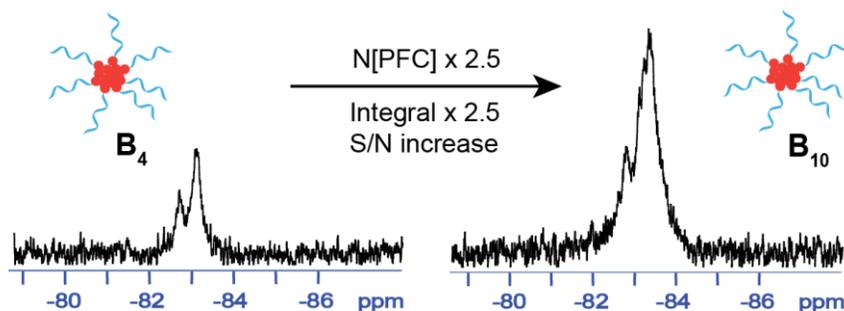


Figure 2.5. ^{19}F NMR spectra of **B₄** (left) and **B₁₀** (right) in D_2O . (S/N = signal to noise ratio).

2.5.4 “Fluorous” effect in PFC-containing DNA duplexes

We assessed the effect of the N[PFC] insertions on DNA thermal denaturation. With a single insertion on each strand of a complementary 19-mer duplex, such that they face each other in the

duplex (**C₂**, **Figure 2.6**), a slight increase of ~ 0.5 °C in the thermal denaturation temperature (T_M) is observable. On the other hand, when two adjacent **N[PFC]** units are incorporated within each strand, a dramatic T_M increase of about 8 °C is detected (**C₃**). This stabilization is similar to that of the replacement of two nucleotides with locked nucleic acid (LNA) nucleotides, considering the length of our DNA strand.³⁷ If the two modifications on each strand are not adjacent (**C₄**), the T_M increase is much less pronounced. The PFC chains thus likely merge into a fluororous environment, avoiding unfavorable interactions in water, and leading to significant melting temperature increase for DNA duplexes. When residing on the duplex end, a single modification on each strand (**C₅**) increases the T_M by 2 °C. Interestingly, a remarkable increase of 20 °C is obtained with two modifications on each strand end (**C₆**) (**Figure 2.6**). DLS data shows that this duplex self-assembles (**Figure 2.20**). Although clean AFM images could not be obtained, we hypothesize these aggregates are also spherical micelles due to the similarity of the DLS signal and the size of the aggregates (14.8 ± 0.4 nm). The **N[PFC]** insertions and the fluororous effect provide a simple method to modulate DNA melting temperatures, and will be valuable for building complex structures from DNA.

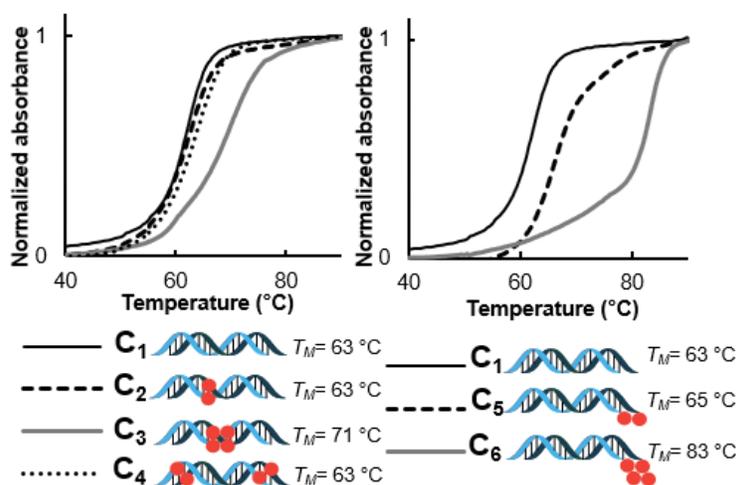


Figure 2.6. Representative melting curves of PFC-modified DNA duplexes in Mg^{2+} containing buffer.

We performed melting temperature measurements at pH 8.0 with different total concentrations (C_T) of DNA. Numbers extracted for these curves were used to construct a van't Hoff plot ($1/T_M$ versus $\ln[DNA]$)⁴² and to extract thermodynamics parameters. Indeed, the linear relation between $1/T_M$ and $\ln[C_T/4]$ under such conditions (**Figure 2.7**) allows to estimate the enthalpy, entropy and Gibbs energy of hybridization of the modified duplexes. Compared to the unmodified DNA strand

C₁, duplexes with only one **N[PFC]** (**C**₂ and **C**₅) showed a less favourable enthalpy of hybridization compensated by a slightly less disfavoured entropic cost (**Figure 2.7**). This resulted in a slightly negative $\Delta\Delta G$ at 37 °C, supporting the stabilizing effect of the PFC chains as expected from the melting curves shown before. Two modifications at internal positions have the best stabilizing effect ($\Delta\Delta G = -3.1$ kcal/mol). Again, it seems that the fluorine effect induced a hybridization entropy increase compared to unmodified DNA that largely compensates for a slight enthalpy increase (**Figure 2.7**). It shows that the **N[PFC]** can be used to significantly stabilize DNA duplexes. Compared to unmodified dsDNA, the main differences in thermodynamics data are visible in the case of **C**₆ where both strands of the duplex are modified with **N[PFC]** at their extremity (**Figure 2.7**). Interestingly, we observed a slight global destabilization of the duplex at 37 °C ($\Delta\Delta G=0.8$ kcal/mol) in spite of a large T_M increase (+20 °C). This observation can be explained by significant differences in hybridization enthalpy ($\Delta\Delta H=60$ kcal/mol) and entropy ($\Delta\Delta S=191$ cal/mol/K). As stated before, **C**₆ and its ssDNA counterpart (**B**₂) are self-assembling in magnesium containing buffer at 20 °C. We hypothesize that **C**₆ hybridization process is significantly less entropically disfavored than **C**₁ due to the entropic gain in micelles size increase from ssDNA (**B**₂, $D_{DLS}=12.4$ nm) to dsDNA (**C**₆, $D_{DLS}=14.8$ nm). On the contrary the fact that it is less enthalpically favored may be explained by some steric hindrance at the interface of the DNA corona and hydrophobic core of the micelles. Due to the close proximity of the duplexes and the core, they might not be able to adopt their most stable conformation. As a conclusion, in-depth thermodynamics analysis allowed to conclude on the stabilizing/destabilizing effect of perfluorocarbon chains in dsDNA when placed at different positions. Moreover, it revealed the enthalpic and entropic contributions in the process, providing useful insight into the hybridization mechanism of PFC-containing strands.

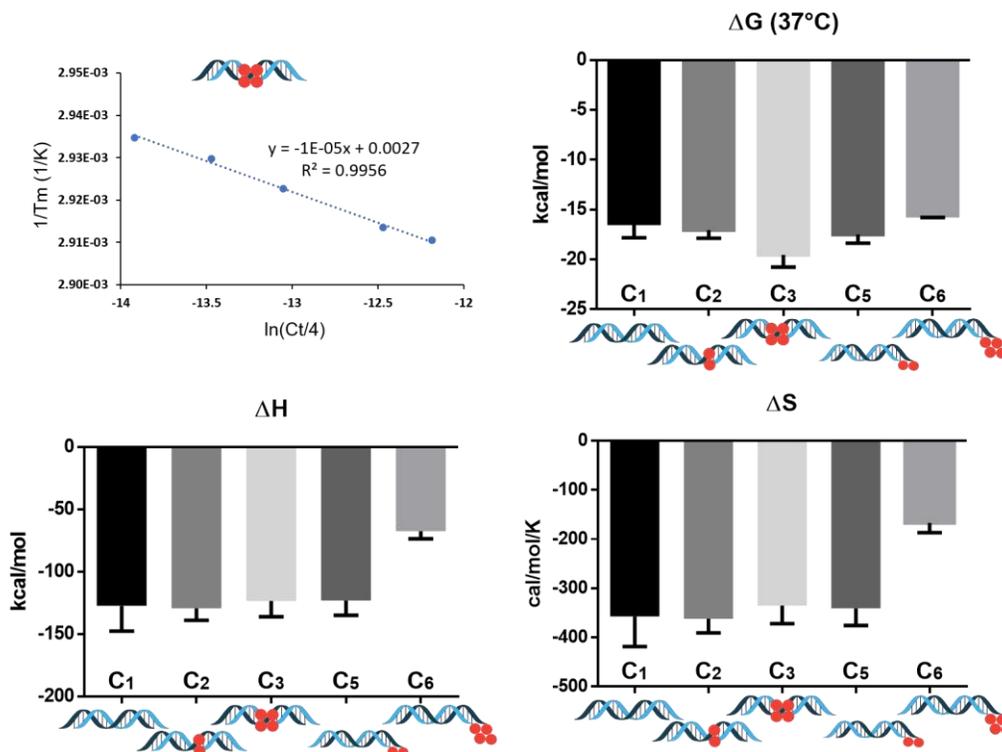


Figure 2.7. Thermodynamics data for PFC-modified DNA duplexes extracted from melting curves. On the top left, representative van't Hoff plot for strand C3; top right, Gibbs free energy of hybridization at $37^\circ C$, bottom left, enthalpy of hybridization, bottom right, entropy of hybridization.

2.5.5 Nuclease resistance of PFC-containing DNA duplexes

Having observed the stabilization they can impart on DNA hybridization, we examined the influence of N[PFC] insertions on nuclease resistance. For this, we exposed an unmodified 19-mer duplex and DNA-PFC conjugates to foetal bovine serum (FBS) in Dulbecco's modified Eagle's minimal essential medium (DMEM) at $37^\circ C$. Aliquots were collected at specific time-points and the nucleases were quenched at low temperature as previously reported.⁴³ DNA in each aliquot was analyzed through quantitative gel electrophoretic mobility assays. Good fit was obtained with a one-phase decay model in the case of C1 to C4 allowing to determine the half-life of the different strands in FBS (**Figure 2.8**). The C2 duplex with a single insertion in each strand showed better nuclease resistance than the unmodified DNA; interestingly, two adjacent modifications even enhanced the DNA half-life by a factor of 3. Due to the incompatibility of a one phase decay model for C5 and C6 which shows non-penetrating bands even in denaturing conditions (**Figure 2.8**), we chose to quantify the total detectable amount of oligonucleotides after

24h incubation. With two modifications on the end of each DNA strand, **C**₆ shows impressive nuclease resistance properties, as ~50 % of the initially detected DNA was still present after 24 h incubation, while only 3 % of **C**₁ was detected after 24 h. Thus, in addition to the capacity for bioimaging by ¹⁹F MRI, the perfluorocarbon modifications can thermally stabilize DNA duplexes and significantly increase their nuclease resistance.

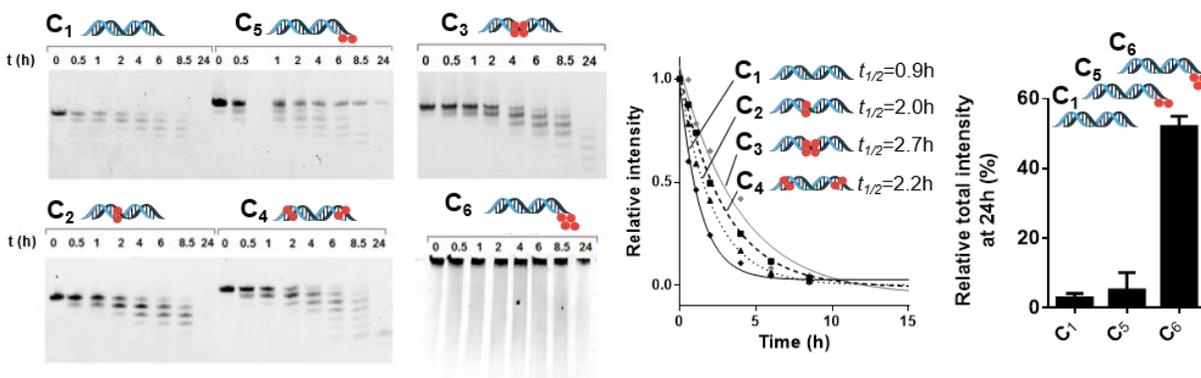


Figure 2.8. Serum stability assay for PFC-modified DNA duplexes. Left: 18 % denaturing polyacrylamide gel electrophoresis (PAGE). In each lane, the strand indicated on top was incubated in FBS for n hours where n is indicated on top of each lane. In the case of **C**₆, non-penetrating bands are observed due to self-assembly. Middle: Half-lives were extracted from a one-phase decay model fit of normalized full-length product band intensity on gels. Graphs with error bars are presented on Figure 2.25. Right: Percentage of DNA detected on gel after 24 h relative to the amount detected initially.

2.5.6 PFC-containing siRNA silencing properties

We were interested to examine if PFC could also be introduced into small interfering RNA strands (siRNA) and whether these constructs would still be therapeutically active. For this experiment, we chose siRNA for apolipoprotein B⁴⁴ and modified the 3'-end of the sense strand with a single **N[PFC]** unit. Yields were similar to those usually obtained for unmodified RNA synthesis and LC-MS characterization confirmed the conjugate's identity (**Figure 2.14**). We transfected unmodified siRNA and PFC-modified siRNA into HepG2 cells (liver carcinoma) and incubated the cells for 24 hours. Total RNA was then collected from the cells and reverse transcribed to cDNA, followed by quantification of ApoB mRNA (messenger RNA) using qRT-PCR (quantitative real-time polymerase chain reaction). We observed that the modified siRNA is able to cause a 50 % reduction in ApoB mRNA levels, and retains its activity compared to unmodified siRNA (**Figure 2.9**). Thus, PFC chains can be incorporated into therapeutic oligonucleotides and

do not interfere with their potency. With the ability to control the sequence, position and length of PFC-insertions, we anticipate being able to further optimize this silencing effect; importantly, we expect that the significant DNA stability, nuclease resistance, nanoscale micelle size and altered cellular penetration^{17,18} provided by the PFC units significantly enhance the *in vivo* delivery of nucleic acid therapeutics.

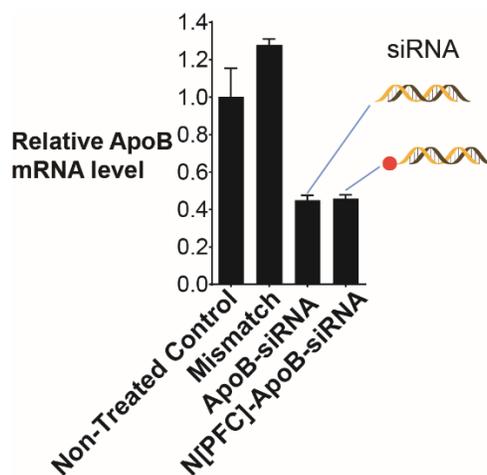


Figure 2.9. Silencing efficiency of PFC-modified siRNA. Mismatch is a negative control: its sequence should not induce gene silencing.

2.6 Conclusions

In conclusion, we highlight the efficiency of automated phosphoramidite chemistry to generate sequence-defined amphiphiles with perfluorocarbon units. Monodisperse oligomers with hydrophilic, fluorophilic and DNA blocks were synthesized. Precise control of the number and location of each of the monomers on the polymer chain was possible and gave rise to different properties in each case. The fluorophilic effect imparted by PFC chains led to the formation of well-defined, almost monodisperse micelles with a hydrophobic PFC core and a DNA corona. Such micelles do not require divalent cations to self-assemble and are detectable and quantifiable through ¹⁹F NMR, outlining their potential for bioimaging applications. At specific locations, PFC chains can also significantly increase the thermodynamic stability of DNA duplexes through the fluorophilic effect inducing an entropically less disfavoured hybridization process. Such duplexes also showed improved nuclease resistance. Perfluorocarbons were also introduced into siRNA, and the conjugate remained potent as a gene silencing agent. A full study of toxicity and *in vivo*

bioaccumulation⁴⁵ has yet to be performed since it is the main concern with PFC-based materials in biology. However, PFCs are inert and usually considered as safe.⁴⁶ Thus, the new PFC-containing monomer available for the synthesis of novel sequence-defined oligo(phosphodiester)s has great potential to broaden the scope of potential applications. For example, DNA- and RNA-perfluorocarbon conjugates are highly likely to be valuable for therapeutic and theranostic applications, given their stability, monodispersity, nanometer sized micelle formation, ability for gene silencing, and their potential for *in vivo* detection by magnetic resonance imaging.

2.7 Experimental section

2.7.1 Chemicals

All starting materials were obtained from commercial suppliers and used without further purification unless otherwise noted. Magnesium chloride, triethylamine, tris(hydroxymethyl)-aminomethane (Tris), urea, EDTA, glycerol, formamide, diethylpyrocarbonate (DEPC), triethylamine tetrahydrofluoride and solvents were used as purchased from Sigma-Aldrich. Acetic acid and boric acid were purchased from Fisher Scientific and used without further purification. GelRed™ nucleic acid stain was purchased from Biotium Inc. Concentrated ammonium hydroxide, acrylamide/Bis-acrylamide (40 % 19:1 solution) and TEMED were obtained from Bioshop Canada Inc. and used as supplied. 1 μmol Universal 1000Å LCAA-CPG supports and standard reagents used for automated DNA and RNA synthesis were purchased through Bioautomation. DMT-hexaethyloxy glycol (cat.# CLP-9765) phosphoramidites were purchased from Chemgenes. Sephadex G-25 (super fine, DNA grade) and sulfurizing reagent II were purchased from Glen Research. AFM cantilevers (model SCANASYST-AIR) were purchased from Bruker and RubyRed mica from Electron Microscopy Sciences. TAMg buffer is composed of 40 mM Tris and 7.6 mM MgCl₂·6H₂O with pH adjusted to 8.0 using glacial acetic acid. TBE buffer is 90 mM Tris, 90 mM boric acid and 1.1 mM EDTA with a pH of 8.0.

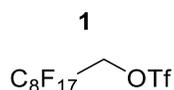
2.7.2 Instrumentation

Standard automated solid-phase synthesis was performed on a Mermade MM6 synthesizer from Bioautomation. HPLC purification was carried out on an Agilent Infinity 1260. DNA

quantification measurements were performed by UV absorbance with a NanoDrop Lite spectrophotometer from Thermo Scientific. A Varian Cary 300 Bio spectrophotometer was used for melting temperature studies. Polyacrylamide Gel Electrophoresis (PAGE) experiments were carried out on a 20 X 20 cm vertical Hoefer 600 electrophoresis unit while agarose gel electrophoresis (AGE) were performed with an Owl Mini gel electrophoresis unit. Gel images were captured using a ChemiDoc™ MP System from Bio-Rad Laboratories. Mass determination of the phosphoramidite was carried out using Electron-Spray Ionization – Ion Trap - Mass Spectrometry (MS) on a Finnigan LCQ Duo device. Liquid Chromatography Electrospray Ionization Mass Spectrometry (LC-ESI-MS) was carried out using Dionex Ultimate 3000 coupled to a Bruker MaXis Impact™ QTOF. Dynamic Light Scattering (DLS) experiments were carried out using a DynaPro™ Instrument from Wyatt Technology. AFM was performed with a MultiMode™ MM8 SPM connected to a Nanoscope™ controller, from the Digital Instruments Veeco Metrology Group. Oxygen and Moisture sensitive experiments were carried out in a Vacuum Atmospheres Co. glove box. The NMR spectra were recorded at 400 or 500 MHz for ¹H and ¹³C at 100.6 or 125 MHz, with chloroform-*d*₁ (δ 7.26, ¹H; δ 77.0, ¹³C), acetone-*d*₆ (δ 2.04, ¹H; δ 29.8, ¹³C) as internal lock solvent and chemical shift standard unless otherwise indicated. ¹⁹F spectra were acquired using a Bruker AVIIIHD spectrometer equipped with a BBFO+ Smartprobe operating at 470.7 MHz. Chemical shift referencing used the lock solvent.

2.7.3 Small molecule synthesis

2.7.3.1 2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-Heptafluoro-1-nonanol trifluoromethanesulfonate (1)



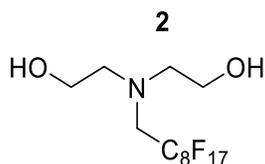
The fluorinated alcohol (3g, 6.67 mmol, 1 equiv.) was dissolved in dry dichloromethane (25 mL) in a round bottom flask (RBF), followed by dry triethylamine (3 mL) and cooled down to 0°C. Trifluoromethanesulphonate anhydride (2.82g, 10 mmol, 1.5 equiv.) was added dropwise over 15 minutes turning the reaction mixture dark. The reaction mixture was stirred at 0°C for 1 hour, than allowed to warm up to RT and stirred for 1 more hour. The reaction was quenched with sat.

NaHCO₃ (50 mL) and product extracted with dichloromethane (3 x 50 mL). Organic fractions were combined, dried with MgSO₄ and solvent removed under vacuum to produce dense black oil. Crude material was purified by flash chromatography with 0-5 % EtOAc/hexanes mixture to produce dense colourless/yellowish oil solidifying on standing. Typical yield 50-60 %.

¹H NMR (400 MHz, CDCl₃) δ 4.82 (t, *J* = 12.3 Hz, 2H).

Note: Synthesis of the mesyl and tosyl equivalents of molecule **1** led to lower yields.

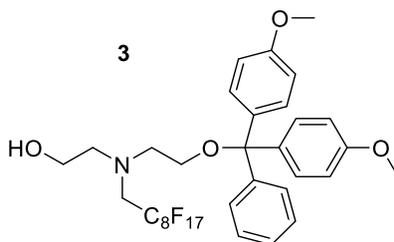
2.7.3.2 2,2'-(2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-Heptafluoro-1-nonylazanediy)diethanol



1 (2.7 g, 4.64 mmol, 1 equiv.) was added to a solution of diethanolamine (1.53 g, 9.28 mmol, 2 equiv.) in dry DMF (6 mL). The reaction mixture was placed in 100°C oil bath and stirred for 2 hours. It was cooled down, dissolved in water (50 mL) and the product extracted with dichloromethane (3 x 50 mL). The organic fractions were combined, dried with MgSO₄ and the solvent was removed under vacuum to produce a dense yellowish oil solidifying on standing (2.42 g, 96 %). The crude product was analyzed with ¹H NMR, determined as analytically pure and used in the next step without any further purification.

¹H NMR (400 MHz, CDCl₃) δ 2.68 (br. s, 2H) 2.92 (t, *J* = 5.1 Hz, 4H), 3.34 (t, *J* = 16.9 Hz, 2H), 3.65 (t, *J* = 5.1 Hz, 4H);

2.7.3.3 2-((2-(Bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)(2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-heptafluorononyl)amino)ethan-1-ol



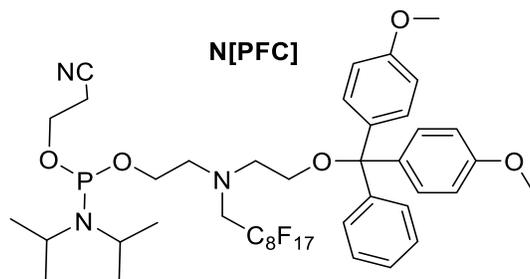
2 (2.42 g, 4.45 mmol, 1 equiv.) was dissolved in dry dichloromethane (15 mL) and triethylamine (2 mL). DMTC1 (1.51 g, 4.45 mmol, 1 equiv.) was added portionwise and the reaction mixture

was stirred at RT for 2 hours. The solvents were evaporated under vacuum resulting a yellow oil purified by column chromatography on triethylamine pre-treated silica with slow gradient of EtOAc/hexanes (0-15 %) mixture to produce a dense yellow oil 1.31g (34 %). Unreacted **2** can be recovered from the column by washing with pure EtOAc.

^1H NMR (400 MHz, acetone- d_6) δ 2.87 (t, $J = 5.9$ Hz, 2H), 3.02 (t, $J = 5.7$ Hz, 2H), 3.26 (t, $J = 5.7$ Hz, 2H), 3.46 (br. s, 1H), 3.57 (t, $J = 17.1$ Hz, 2H), 3.62 (t, $J = 6.0$ Hz, 2H), 3.77 (s, 6H), 6.88 (d, $J = 8.8$ Hz, 4H), 7.22 (d, $J = 7.4$ Hz, 1H), 7.30 (t, $J = 7.6$ Hz, 2H), 7.36 (d, $J = 8.8$ Hz, 4H), 7.49 (d, $J = 7.7$ Hz, 2H).

^{13}C NMR (100 MHz, acetone- d_6) δ 48.1, 56.4, 57.2, 59.4, 61.9, 63.9, 88.2, 114.8, 128.5, 129.5, 130.0, 131.9, 138.2, 147.4, 160.6.

2.7.3.4 2-((2-(Bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)(2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,9-heptafluorononyl)amino)ethyl (2-cyanoethyl) diisopropylphosphoramidite (N[PFC])



An oven-dried round bottom flask was charged with **3** (860 mg, 1.02 mmol, 1 equiv.) dissolved in dry THF (3.3 mL). 5-(ethylthiotetrazole) (ETT) (6.2 mL, 0.25M in acetonitrile, 1.54 mmol, 1.5 equiv.) was added under argon and stirred vigorously followed by addition of 3-((bis(diisopropylamino)phosphanyl)oxy)propanenitrile (0.81 mL, 2.55 mmol, 2.5 equiv.). The reaction mixture was left stirring for 6 hours at room temperature under argon. The solvent was evaporated under reduced pressure. The crude product was purified by column chromatography under a positive pressure of argon with mobile phase: degassed Hexanes/ethyl acetate/triethylamine TEA (90:10:2). **4** was isolated as a colourless oil: 880 mg. Yield: 83 %.

TLC (Hexanes/ethyl acetate/TEA: 90:10:2): R_f =0.16.

LRMS: Calc.exact mass: 1039.30 g/mol. Measured (positive mode): 1062.18 (M+23), 1078.16 (M+39) g/mol.

^1H NMR (500 MHz, acetone- d_6): δ (ppm) = 7.49-7.47 (m, 2H), 7.37-7.34 (m, 4H), 7.32-7.29 (m, 2H), 7.24-7.21 (m, 1H), 6.88 (d, $J=9\text{Hz}$, 4H), 3.81-3.56 (m, 14H), 3.26 (t, $J=6\text{Hz}$, 2H), 3.05 (t, $J=6\text{Hz}$, 2H), 3.00 (t, $J=6\text{Hz}$, 2H), 2.71 (t, $J=6\text{Hz}$, 2H), 1.20-1.15 (m, 12H).

^{13}C NMR (125 MHz, acetone- d_6) δ 159.6, 146.4, 137.2, 130.9, 129.0, 128.6, 127.5, 118.9, 113.9, 87.2, 63.2, 62.9, 62.8, 59.5, 59.4, 56.9, 56.8, 56.8, 56.1, 55.5, 43.8, 43.7, 24.9, 24.9, 20.8, 20.7.

^{31}P NMR (400 MHz, acetone- d_6) δ (ppm) = 147.4 (s).

2.7.4 Solid-phase synthesis

DNA synthesis was performed on a 1 μmole scale, starting from a universal 1000 Å LCAA-CPG solid support. Coupling efficiency was monitored after removal of the dimethoxytrityl (DMT) 5'-OH protecting groups. DMT-hexaethyloxy glycol amidite (**HEG**) and **N[PFC]** were respectively dissolved in acetonitrile and acetonitrile/THF (90:10) under a nitrogen atmosphere in a glove box (<0.04 ppm oxygen and <0.5 ppm trace moisture). For DMT-hexaethyloxy glycol (0.1M) and **N[PFC]** (0.08M) amidites, extended coupling times of 10 minutes were used. For the addition of each RNA nucleoside phosphoramidite extended coupling time of 6 minutes was used. For 2' OMe modified RNA phosphoramidites, under a nitrogen atmosphere, coupling was done using the 'syringe' technique: the amidite solution (200 μl , 0.1 M) is mixed with the usual activator solution (200 μl , 0.25 M) in presence of the CPG using syringes. After twenty minutes, the solution was removed from the columns and the strands underwent capping, oxidation and deblocking steps in the synthesizer. Removal of the DMT protecting group was carried out using 3 % dichloroacetic acid in dichloromethane on the DNA synthesizer. Sulfurizing reagent was used according to standard procedures.

Deprotection procedure (except RNA): Completed syntheses were cleaved from the solid support and deprotected in 28 % aqueous ammonium hydroxide solution for 16-18 hours at 60 °C. The crude product solution was separated from the solid support and concentrated under reduced pressure at 60 °C. This crude solid was re-suspended in 1 mL Millipore water. Filtration with 0.22

µm centrifugal filter was then performed prior to HPLC purification. The resulting solution was quantified by absorbance at 260 nm.

RNA deprotection procedure: Completed 1 µmol syntheses were deprotected in 1 mL of a 1:1 v/v mixture of 40 % aqueous Methylamine and 28 % aqueous ammonium hydroxide solution for 30 minutes at room temperature, followed by 3 hours at 65° C. The crude product solution was separated from the solid support and concentrated under reduced pressure at 60°C. This crude solid was re-suspended in 150 µL of a desilylation solution containing triethylamine, N-methylpyrrolidone, and triethylaminetrihydrofluoride (3:2:1.5) and heated to 65°C for 2 hours, to remove the 2'-OH *tert*-butyldimethylsilyl protecting groups. This desilylation step was then quenched by the addition of 100 µL of 3 M sodium acetate (pH 5.5) and vortexed. RNA precipitation was induced by addition of 1 mL of cold butanol and left for 30 minutes at -20°C. A pellet appeared through centrifugation (20 minutes, 12000 x g, 4°C) and the supernatant was removed. The pellet was washed a second time with 500 µL butanol, and dried under reduced pressure at 60 °C. This crude product was resuspended in DEPC-treated sterile water and quantified by absorbance at 260nm.

HPLC purification: Solvents (0.22 µm filtered): 50 mM triethylammonium acetate (TEAA) buffer (pH 8.0) and HPLC grade acetonitrile. Elution gradient: 3-95 % acetonitrile over 40 minutes at 60 °C. Column: Hamilton PRP-C18 5 µm 100 Å 2.1 x 150 mm. For each analytical separation approximately 0.5 OD₂₆₀ of crude DNA was injected as a 20-50 µL solution in Millipore water. Detection was carried out using a diode-array detector, monitoring absorbance at 260 nm.

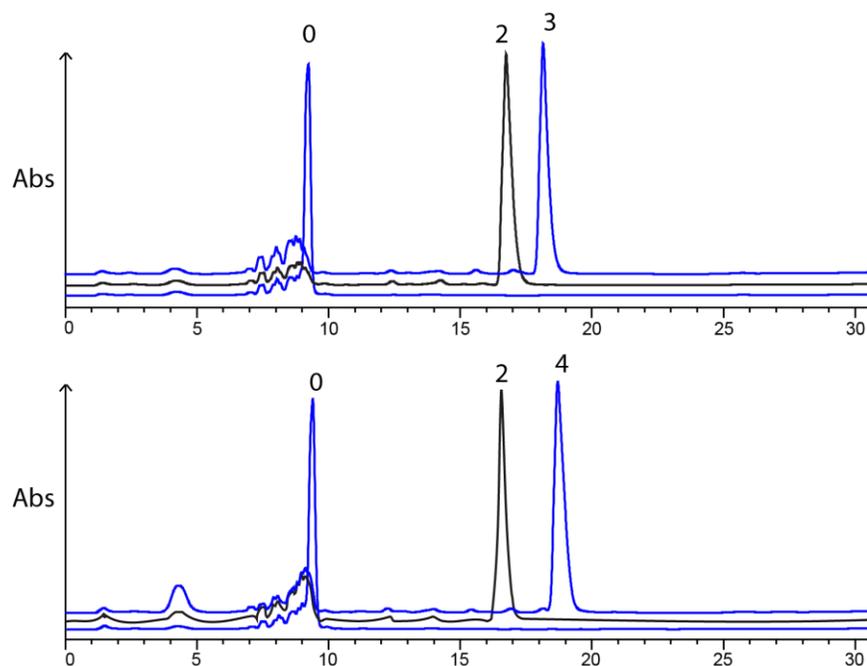


Figure 2.10. Reverse-phase HPLC traces from crude mixtures of $A_{n,m}$ oligomers. (UV detection, 260 nm) Top: $n=6$, bottom: $n=8$. Numbers on the peaks are the number (m) of N[PFC]. Byproducts are in low quantities. They are almost exclusively related to HEG coupling as the byproducts peaks are visible before the polymer retention time for $m=0$.

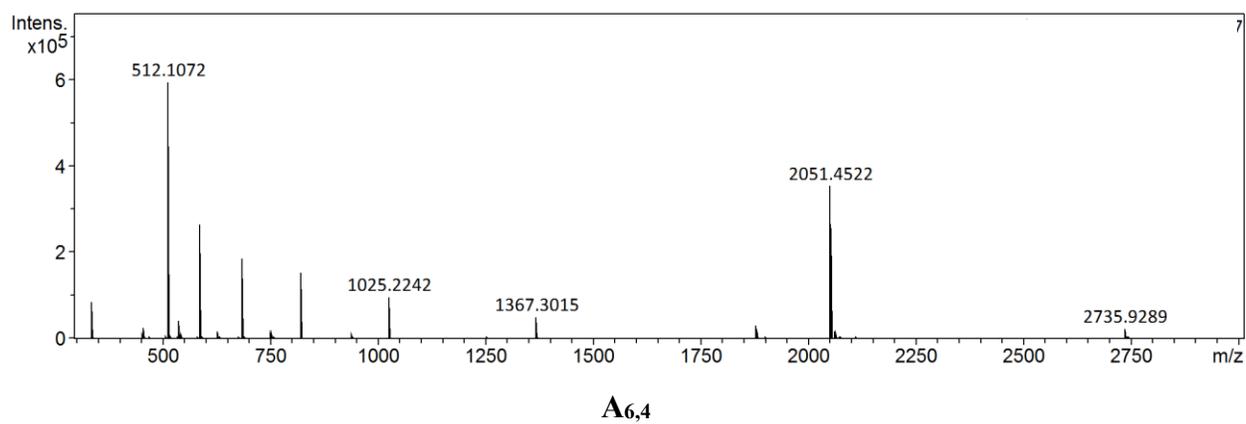
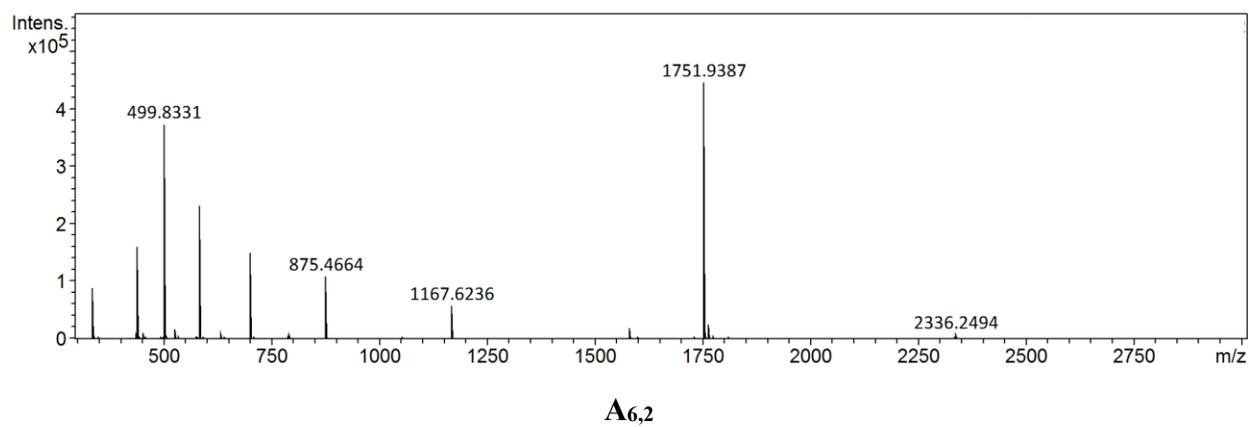
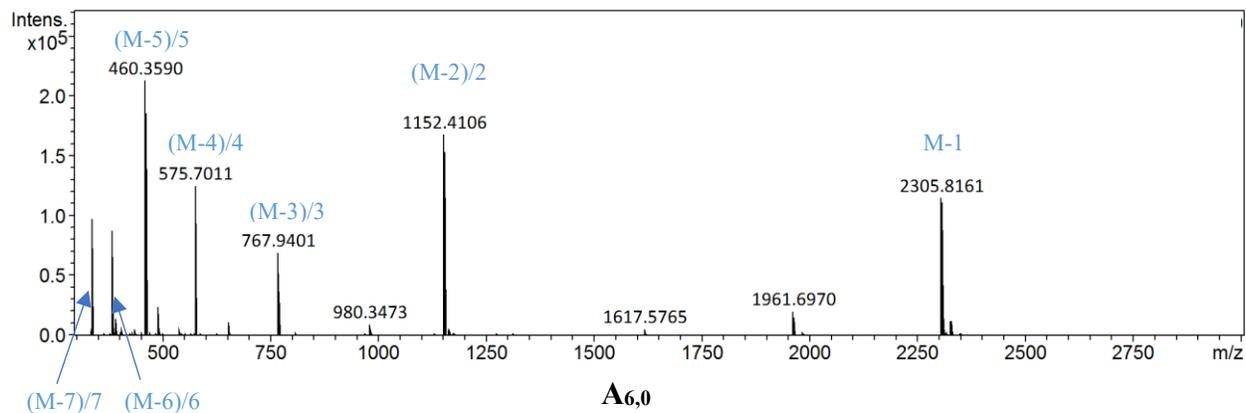
Gel electrophoresis purification: In the case of DNA/RNA internal or 3' end modification, purification was carried out through gel electrophoresis instead of RP-HPLC. In that case, crude products were purified on 19 % polyacrylamide gels, supplemented with 8 M urea (loading up to 20 OD260 of crude DNA per gel, 500 V field applied). Electrophoresis was run at lower voltage for the first 30 minutes. Following electrophoresis, the gel was wrapped in plastic and visualized by UV shadowing over a fluorescent TLC plate. The full-length product was quickly excised, then crushed and incubated in ~10 mL of autoclaved water (treated with DEPC in the case of RNA purification) at 55 °C overnight. The supernatant was then concentrated to 1.0 mL, and desalted using size exclusion chromatography (Sephadex G-25). Sephadex was treated with DEPC in the case of RNA purification. Strands were then quantified (OD260) and converted to micromolar concentrations using the extinction coefficients obtained on IDT technology website (<http://www.idtdna.com/calc/analyzer>).

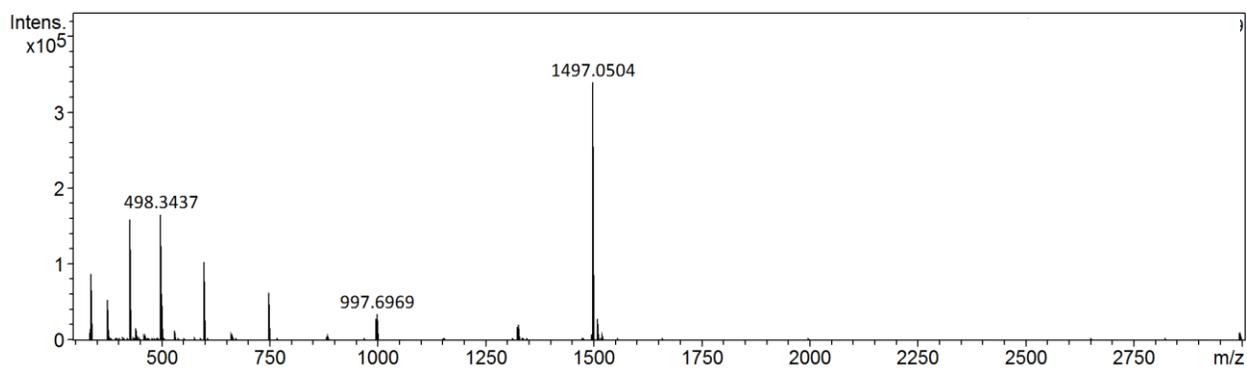
Table 2.5. DNA, RNA and conjugates sequence. Lower case letters indicate RNA, uppercase letters indicate DNA. CpAT is complementary to the AT sequence. * indicates a phosphorothioate linkage. Letters in bold indicate 2' OMe modifications. **C₁** is composed of AT and cpAT; **C₂** is composed of AT-(**N[PFC]**)_{p10} and cpAT-(**N[PFC]**)_{p11}; **C₃** is composed of AT-(**N[PFC]₂**)_{p10} and cpAT-(**N[PFC]₂**)_{p11}; **C₄** is composed of AT-**N[PFC]**_{p7p14} and cpAT-**N[PFC]**_{p7p14}, **C₅** is composed of **B₁** and (**N[PFC]**)-cpAT; **C₆** is composed of **B₂** and (**N[PFC]₂**)-cpAT.

Molecule	Sequence (5'-xx-3')
AT	TTTTTCAGTTGACCATATA
cpAT	TATATGGTCAACTGAAAAA
AT-(N[PFC])_n or B_n	(N[PFC]) _n TTTTTCAGTTGACCATATA
(N[PFC])_n-cpAT	TATATGGTCAACTGAAAAA(N[PFC]) _n
AT-(N[PFC]_n)p10	TTTTTCAGT(N[PFC]) _n TGACCATATA
cpAT-(N[PFC]_n)p11	TATATGGTCA(N[PFC]) _n ACTGAAAAA
AT-N[PFC]_{p7p14}	TTTTTC(N[PFC])AGTTGAC(N[PFC])CATATA
cpAT-N[PFC]_{p7p14}	TATATG(N[PFC])GTCAACT(N[PFC])GAAAAA
AT-GCp10	TTTTTCAGTGCTGACCATATA
cpAT-GCp11	TATATGGTCAGCACTGAAAAA
ApoB-anti	auugguauucagugaugac* a *c
ApoB-sense	gucaucacacugauaccaa*u
ApoB-sense-N[PFC]	gucaucacacugauaccaa*u* N[PFC]

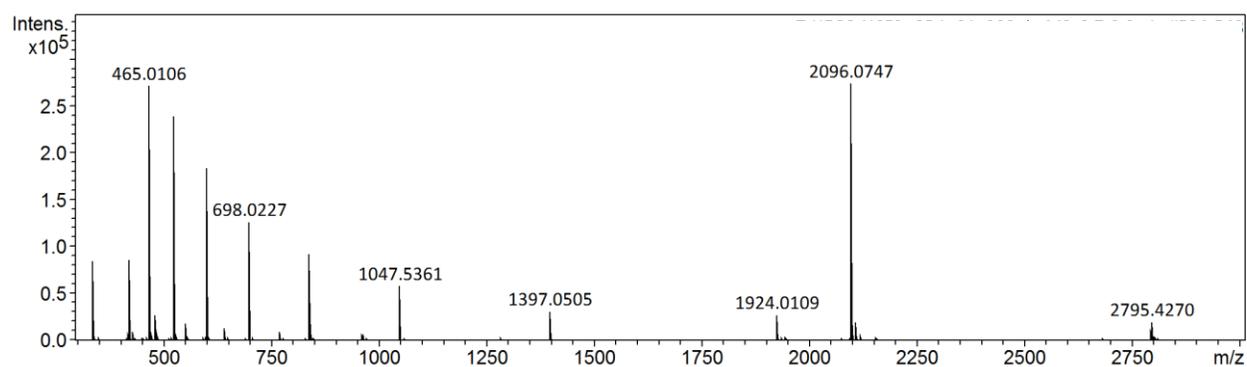
2.7.5 LC-ESI-MS characterization

The oligomers were analyzed by LC-ESI-MS in negative ESI mode. Samples were run through an Acclaim RSLC 120 C18 column (2.2µm, 120Å 2.1 x 50 mm) using a gradient of mobile phase A (100 mM 1,1,1,3,3,3-hexafluoro-2-propanol and 5 mM triethylamine in water) and mobile phase B (Methanol) in 8 minutes (2 % to 100 % B). For each run, ~ 250 pmols of artificial oligomer or ~ 65 pmols of oligonucleotide were injected.

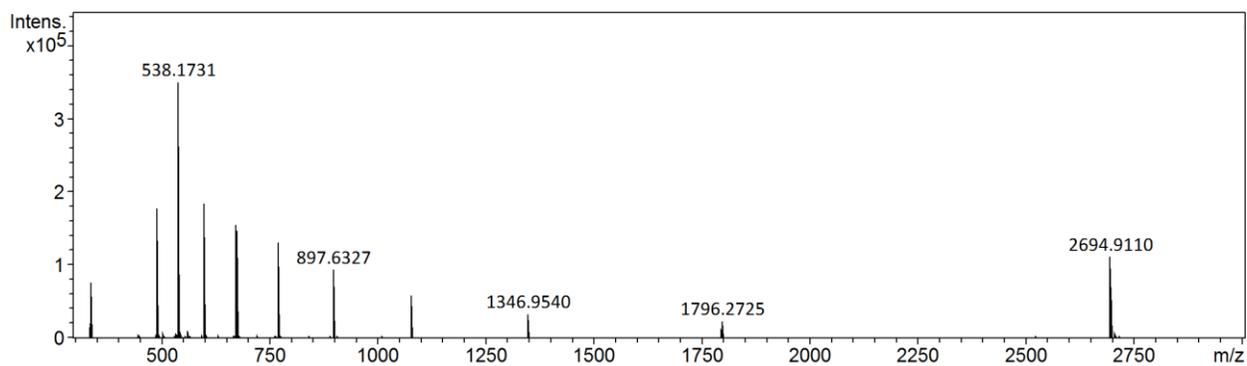




A8,0

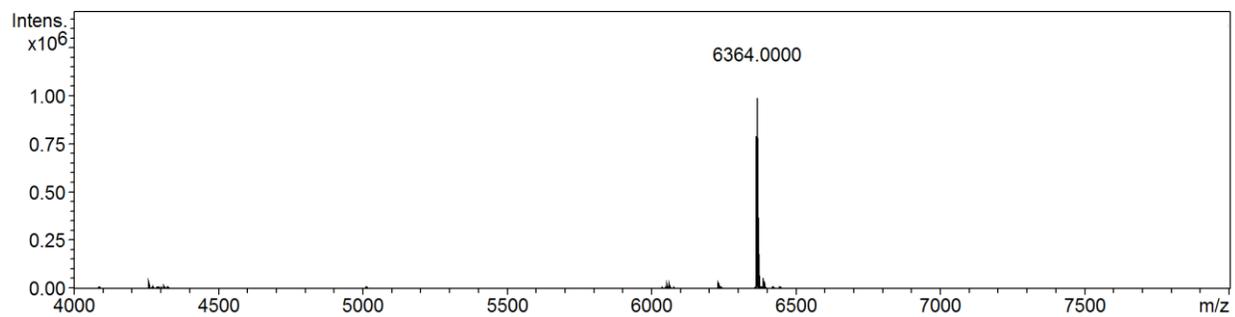


A8,2

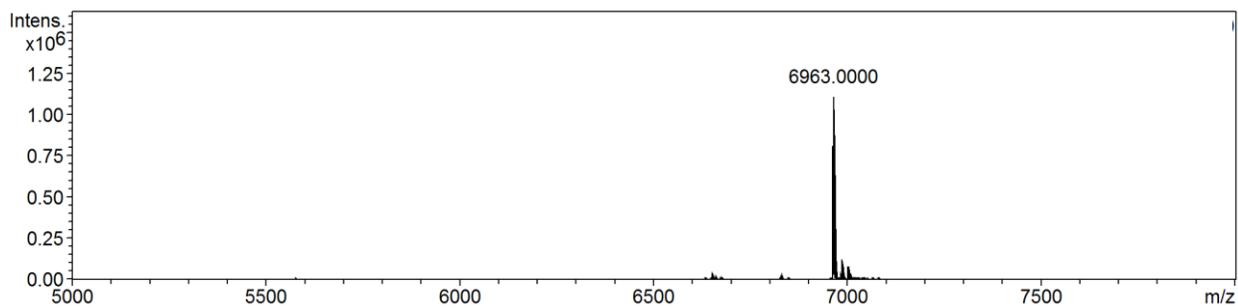


A8,4

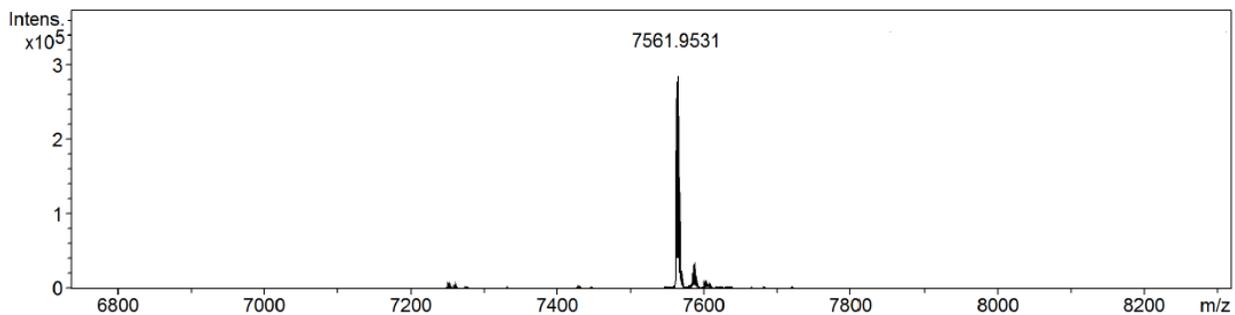
Figure 2.11. MS data for $A_{n,m}$ oligomers. Almost all peaks can be associated with a $(M-x)/x$ anion, as illustrated on the first spectrum. $A_{n,m}$ stands for T- HEG_n -N[PFC] $_m$.



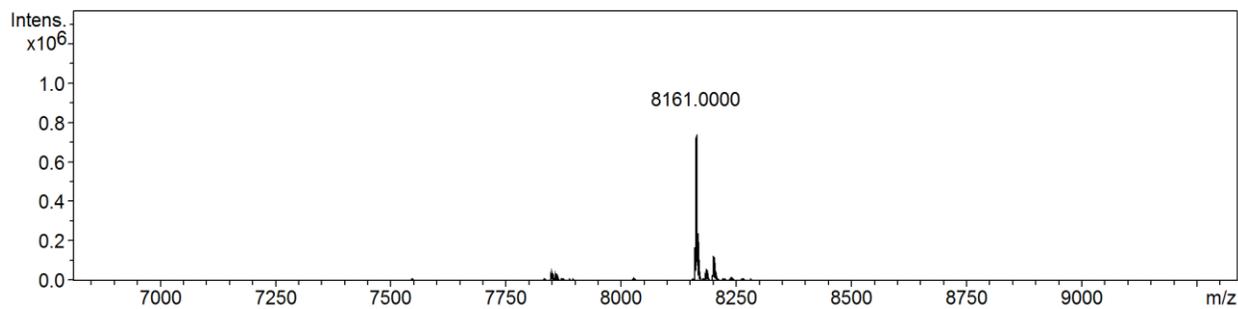
B₁



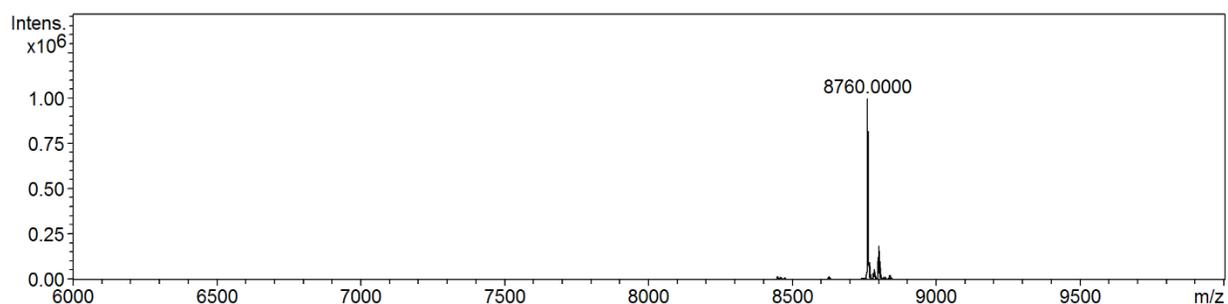
B₂



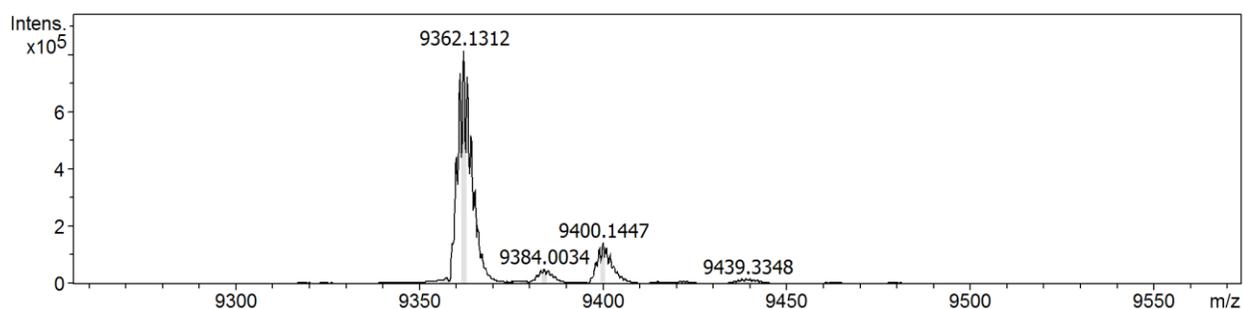
B₃



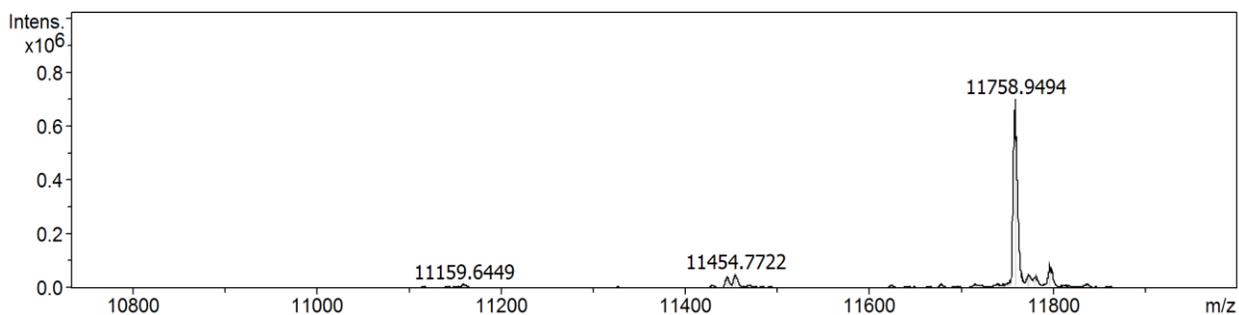
B₄



B₅



B₆

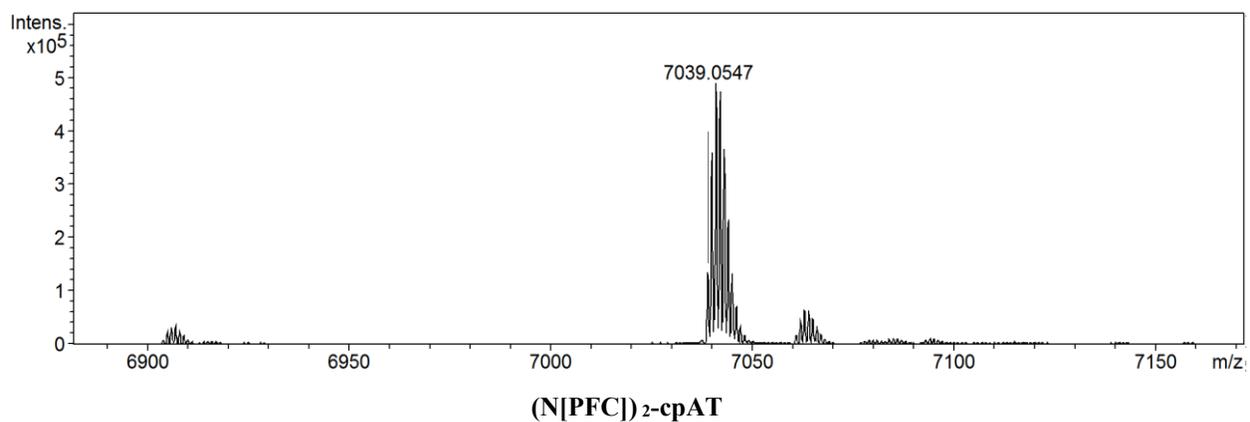
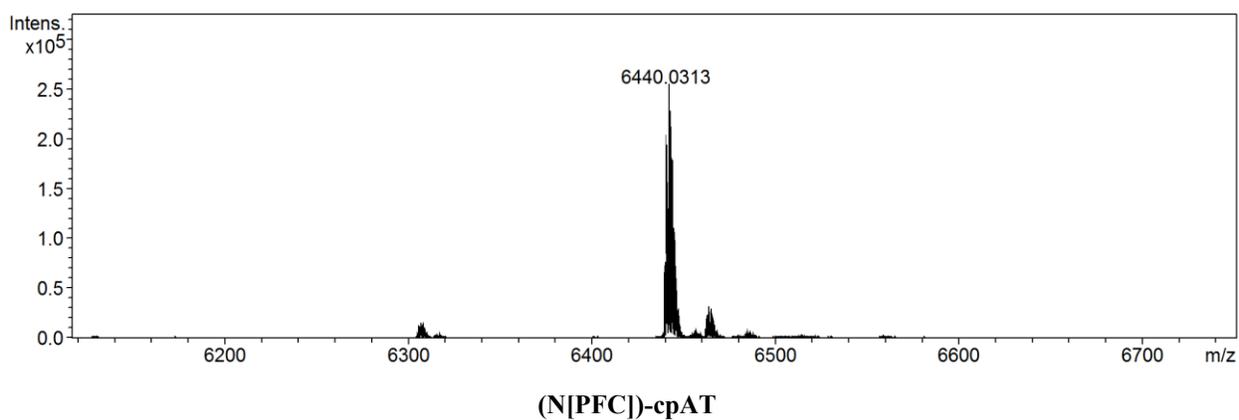


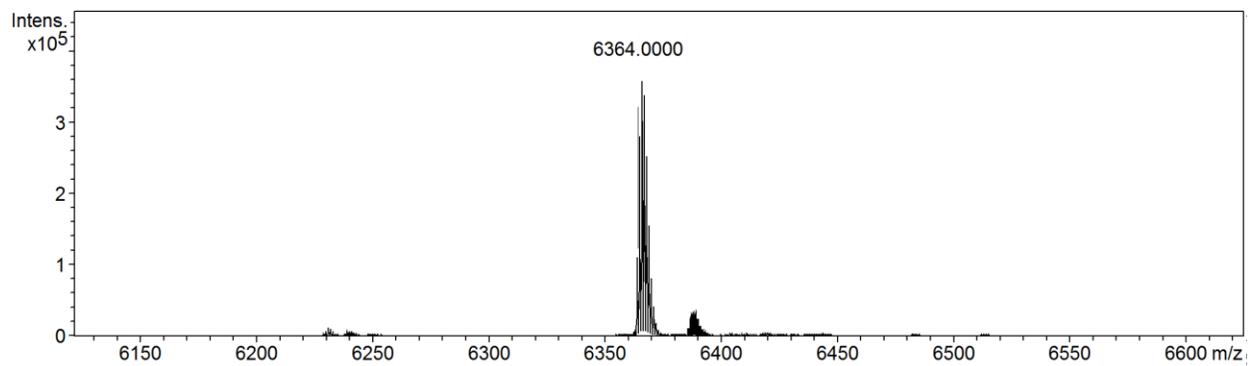
B₁₀

Figure 2.12. LC/MS data for DNA-“Teflon” oligomers (**B_n**). The data was processed and deconvoluted using the Bruker DataAnalysis software version 4.1. Masses reported are exact masses except for big conjugates (>10kDa).

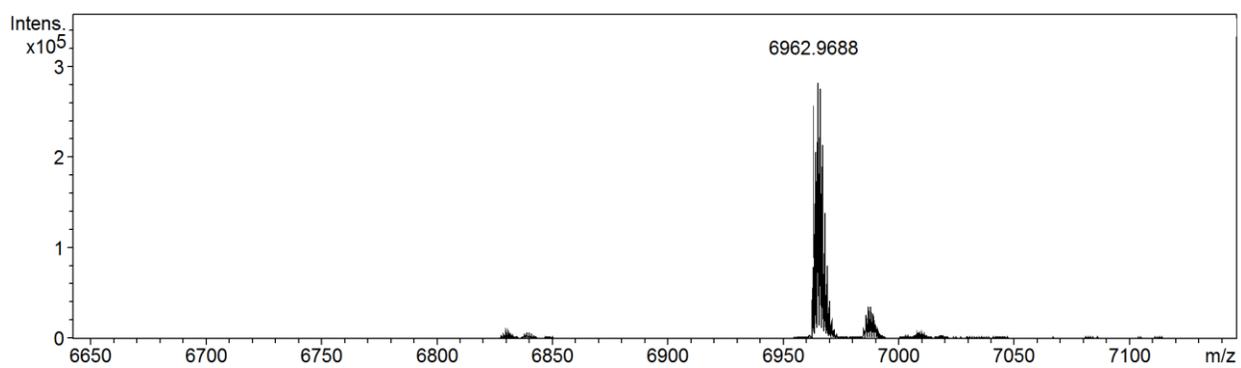
Table 2.6. LC/MS results for DNA with internal modifications or 3' end modification.

Strand name	Calculated exact mass (g/mol)	Measured mass (g/mol)
(N[PFC])-cpAT	6440.05	6440.03
(N[PFC] ₂)-cpAT	7039.06	7039.05
AT-(N[PFC])p10	6364.01	6364.00
AT-(N[PFC] ₂)p10	6963.02	6962.97
AT-(N[PFC] ₃)p10	7562.03	7561.95
cpAT-(N[PFC])p11	6440.05	6440.03
cpAT-(N[PFC] ₂)p11	7039.06	7039.09
cpAT-(N[PFC] ₃)p11	7638.07	7638.01
AT-N[PFC]p7p14	6963.02	6962.97
cpAT-N[PFC]p7p14	7039.06	7039.09

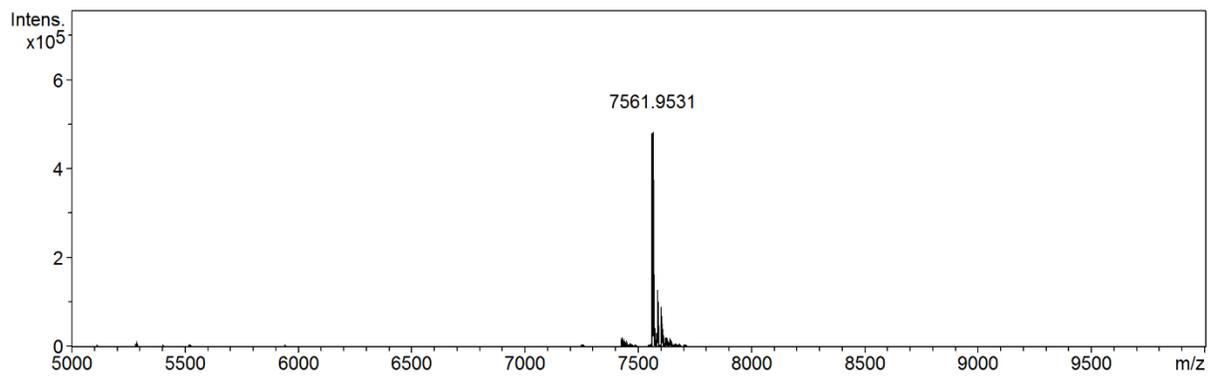




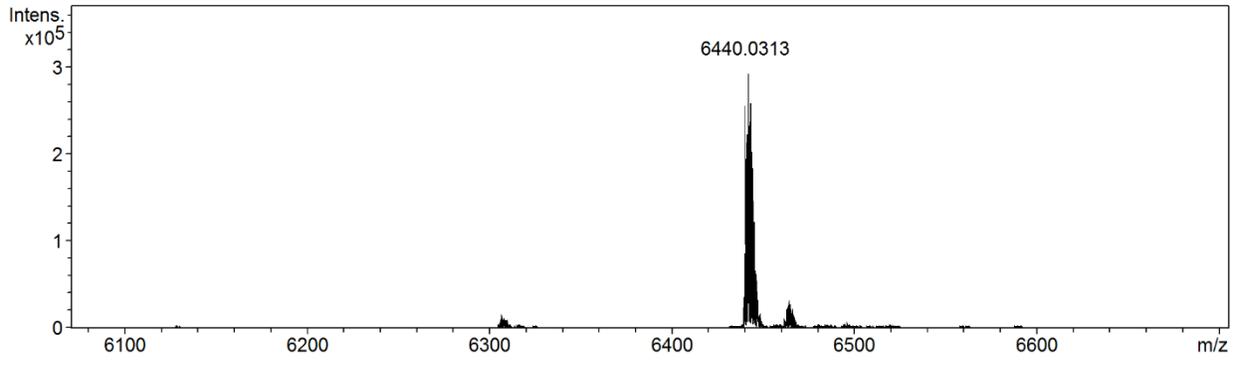
AT-(N[PFC])p10



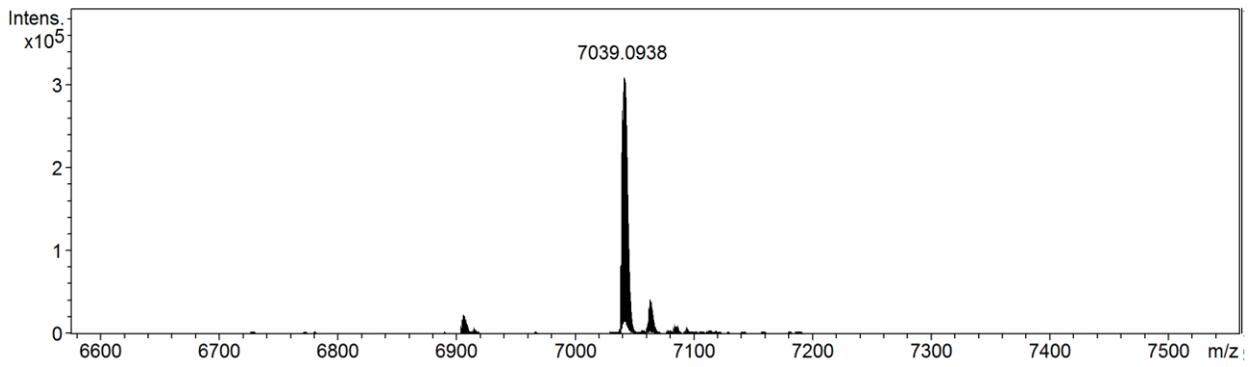
AT-(N[PFC])₂p10



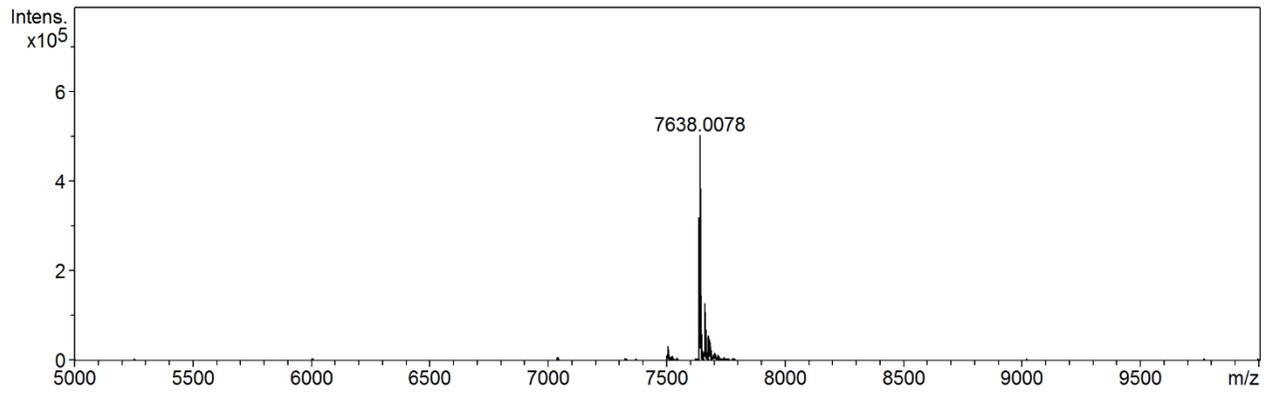
AT-(N[PFC])₃p10



cpAT-(N[PFC])p11



cpAT-(N[PFC])₂p11



cpAT-(N[PFC])₃p11

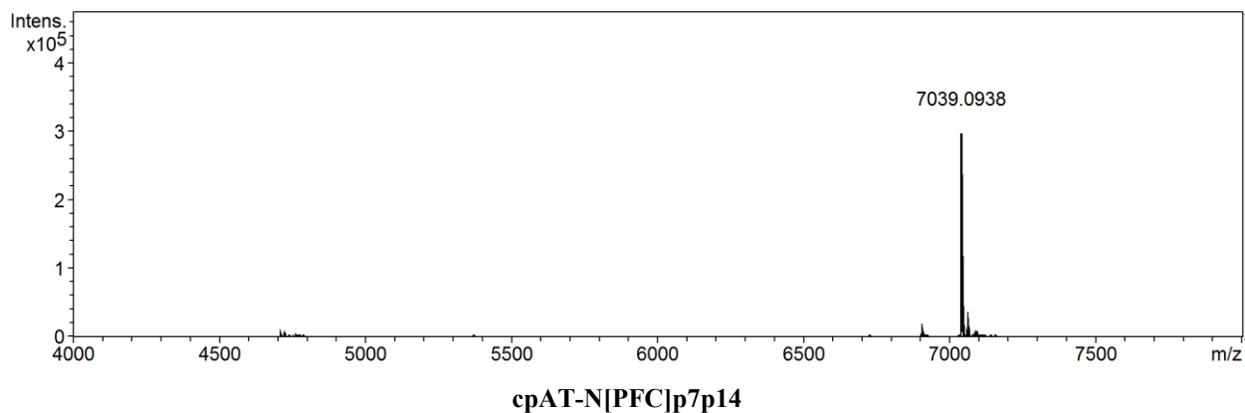
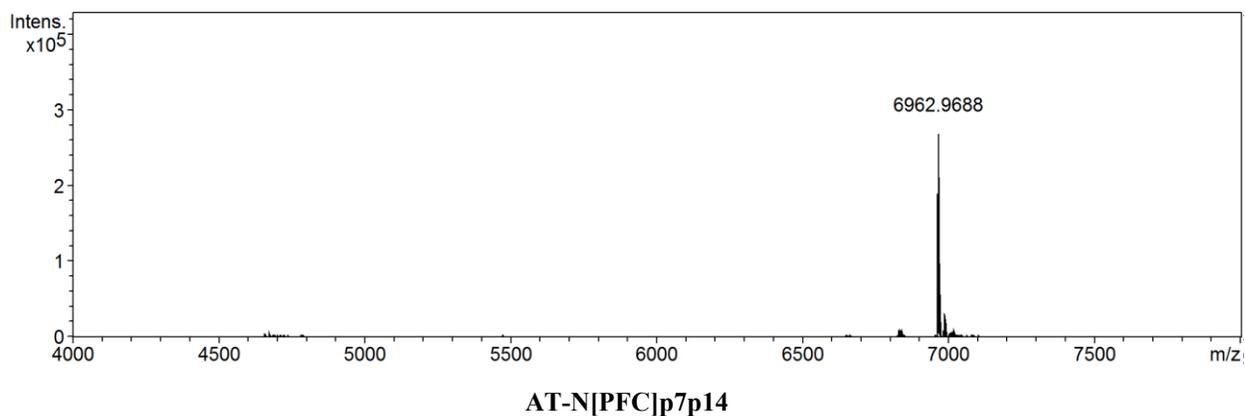


Figure 2.13. LC/MS data for DNA with internal modifications or 3'end modification. The data was processed and deconvoluted using the Bruker DataAnalysis software version 4.1.

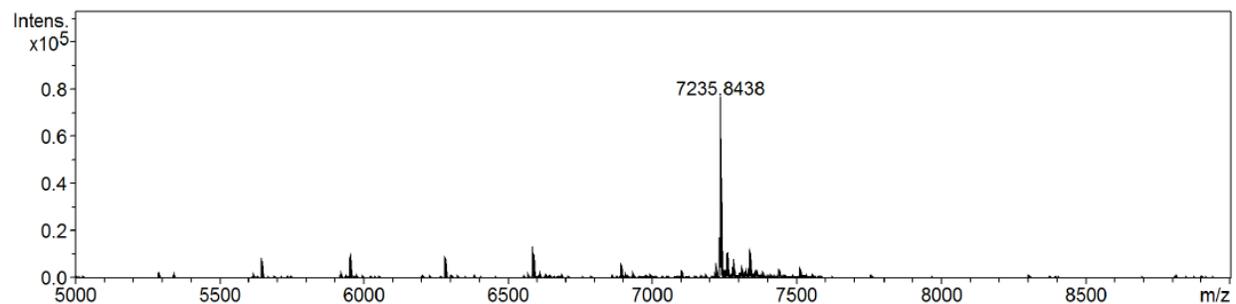


Figure 2.14. LC/MS data for RNA with 3'end N[PFC] modification. The data was processed and deconvoluted using the Bruker DataAnalysis software version 4.1. Calculated exact mass: 7235.94, Obtained: 7235.84.

2.7.6 Atomic force microscopy

Samples were diluted to 1 μM in TAMg buffer and 4 μL of this solution was deposited on a freshly cleaved mica surface (ca. 7 x 7 mm) and allowed to adsorb for 1-2 seconds. Then 50 μL of 0.22 μm filtered Millipore water was dropped on the surface and instantly removed with filter paper. The surface was then washed four times with 80 μL of water and the excess removed with a strong flow of nitrogen. Samples were dried under vacuum for 3 hours prior to imaging. Imaging of samples was performed on a MultiMode 8 microscope with a Nanoscope V controller (Bruker) in ScanAsyst mode. Silicon nitride levers with a nominal spring constant of 0.4 N/m, resonant frequency of 70 kHz and a 2 nm tip radius were used (ScanAsyst Air). All images were captured at a 1.10 to 1.40 Hz scan rate and a resolution of 512 x 512 pixels.

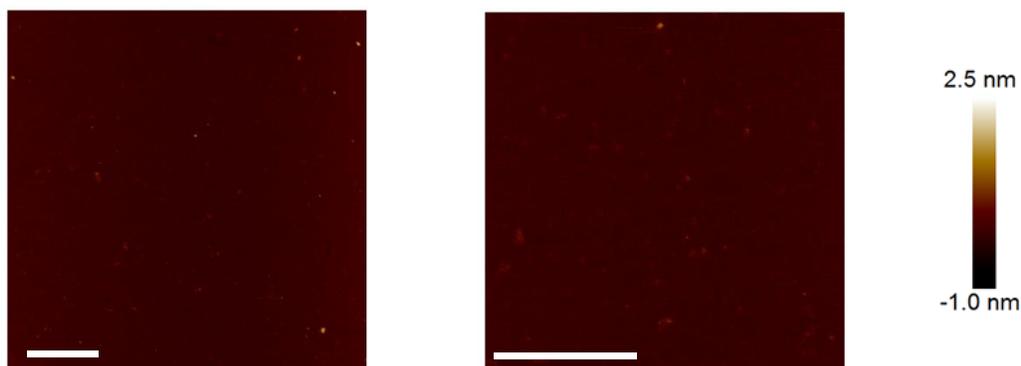


Figure 2.15. Dry AFM images of unmodified DNA control **B₀**.

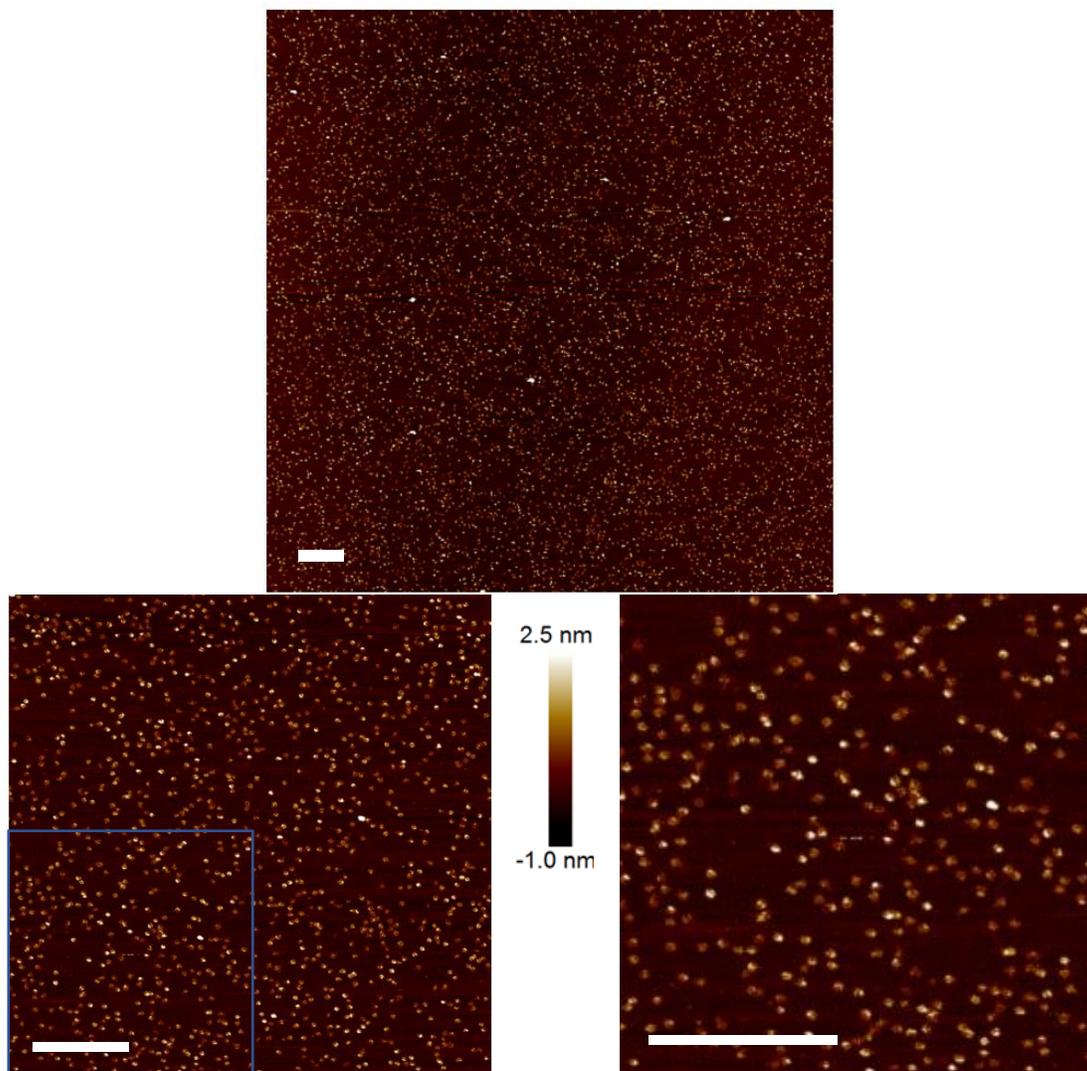


Figure 2.16. Dry AFM images of AT-(N[PFC])₅ micelles (**B₅**). The image in the bottom-right corner is a zoom-in of the bottom-left image (blue square). We found an average diameter of 19.4 ± 3.1 nm for AT-(N[PFC])₅ micelles using the software image J.

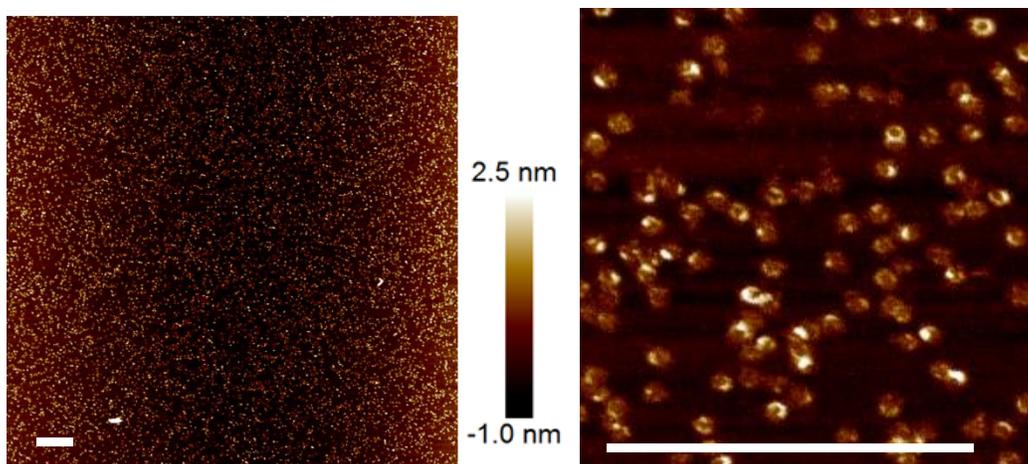


Figure 2.17. Dry AFM images of AT-(N[PFC])₁₀ micelles (**B**₁₀). The image in the right is a zoom-in of the right image of Figure 2.2. White bars represent 400 nm. We found an average diameter of 20.9 ± 3.3 nm. Statistics were realized by hand using the software Image J.

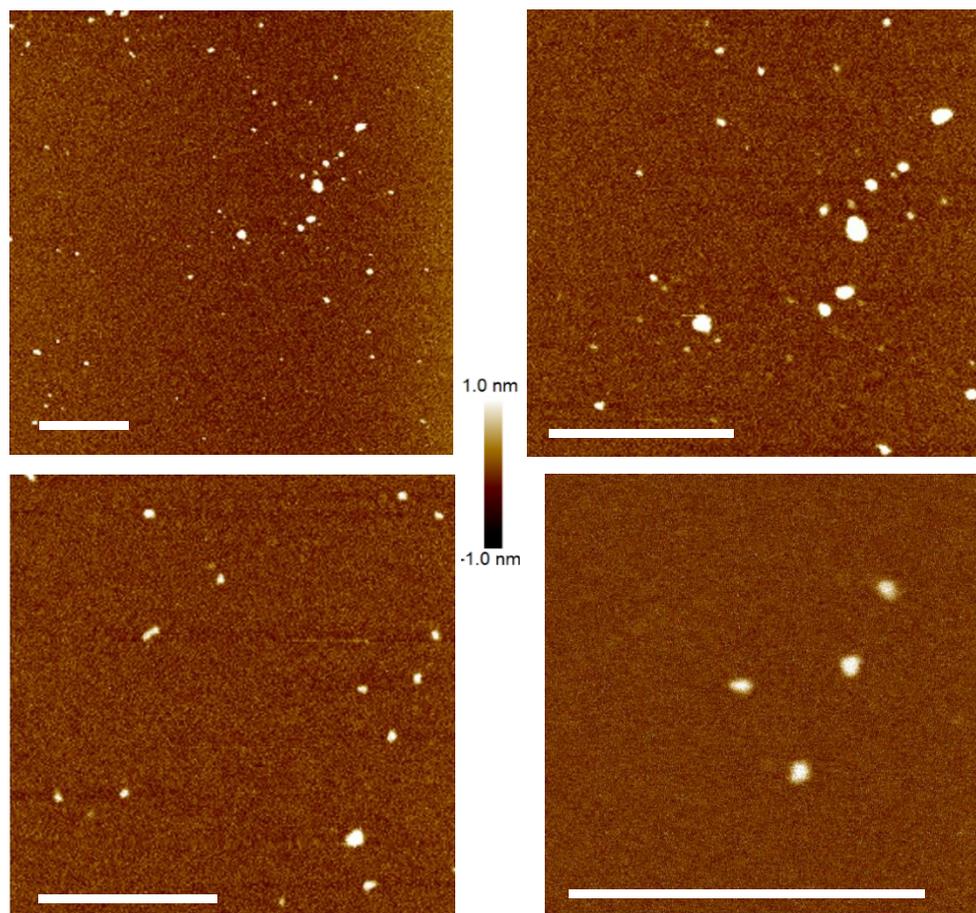


Figure 2.18. Dry AFM images of **B**₂. White bars represent 400 nm. Aggregation is visible but we hypothesized that the smaller PFC hydrophobic core does not maintain structure integrity during sample drying.

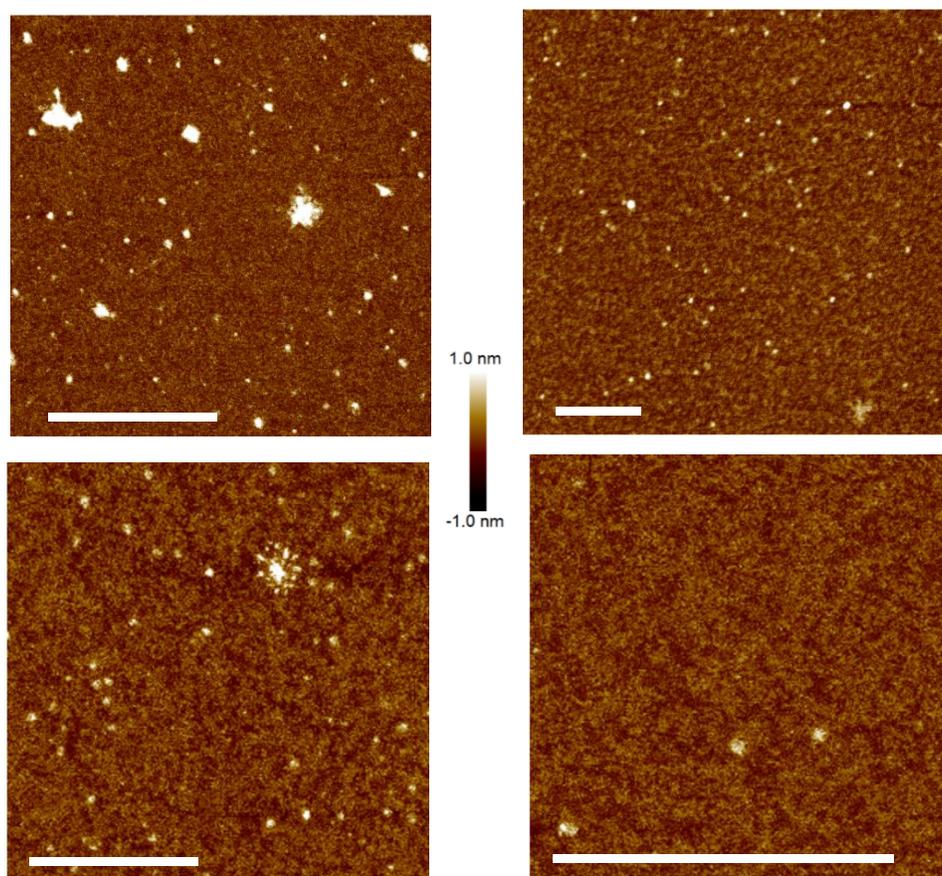
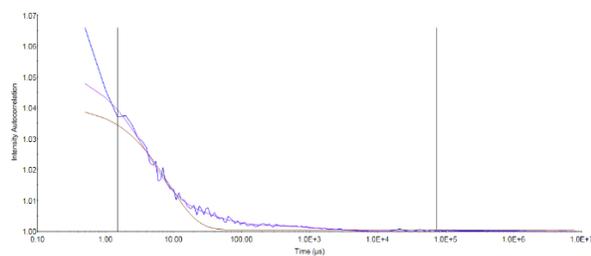


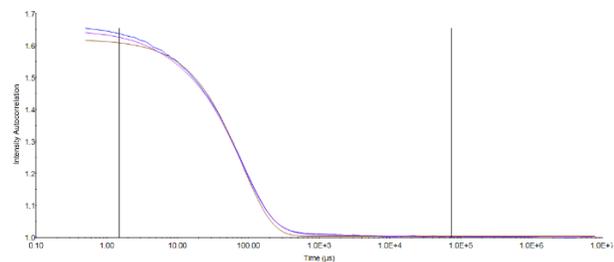
Figure 2.19. Dry AFM images of C_6 . White bars represent 400 nm. Aggregation is visible but we hypothesized that the smaller PFC hydrophobic core does not maintain structure integrity during sample drying.

2.7.7 Dynamic Light Scattering

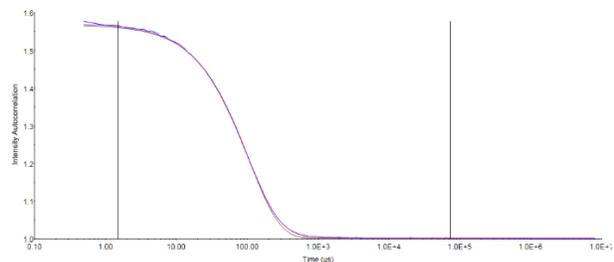
A cumulants fit model was used to confirm the presence and determine the size of a monomodal population of micellar aggregates. Sterile water and TAMg were filtered using a 0.2 μm nylon syringe filter before use for DLS sample preparation. All measurements were carried out at 20 $^{\circ}\text{C}$. Concentration of the sample is 10 μM if not specified. All the measurements were at least triplicated. For experiments with Mg^{2+} , samples were directly diluted in TAMg or Mg^{2+} ions were added after solubilizing our constructs in sterile water: both methods led to similar results.



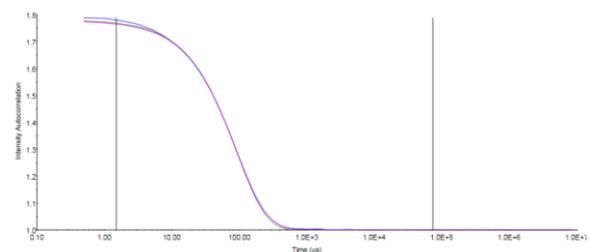
AT (Mg^{2+})



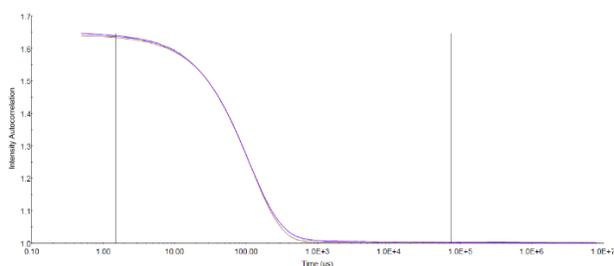
B₂ : AT-(N[PFC])₂ (Mg^{2+})



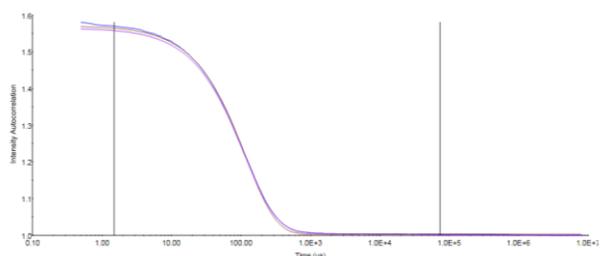
B₄ : AT-(N[PFC])₄



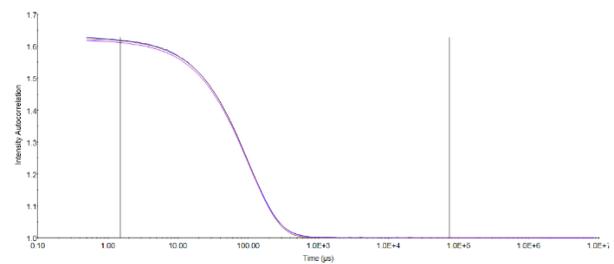
B₄ : AT-(N[PFC])₄ (Mg^{2+})



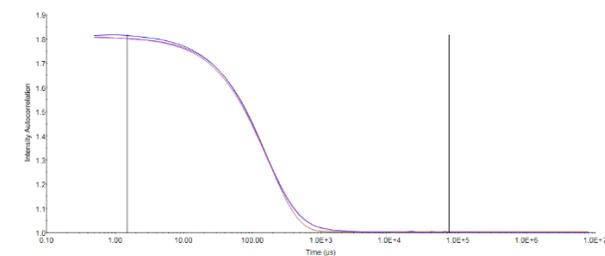
B₅ : AT-(N[PFC])₅



B₅ : AT-(N[PFC])₅ (Mg^{2+})



B₁₀ : AT-(N[PFC])₁₀



B₁₀ : AT-(N[PFC])₁₀ (Mg^{2+})

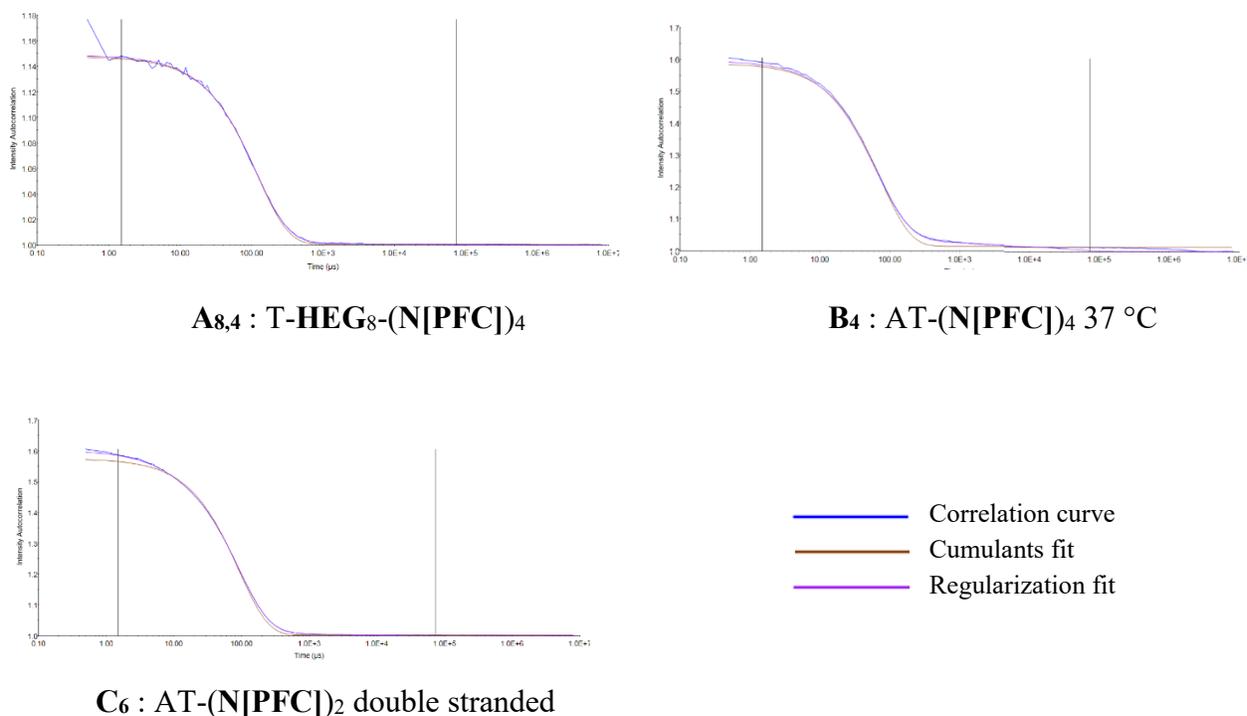


Figure 2.20. DLS intensity correlation functions for sequence-defined oligomers samples. Representative curves for 10 μ M solutions. The low scattering intensity and poor correlation functions measured for AT is characteristic of individual molecules in solution. In contrast the data for self-assembling oligomers reveals good to excellent correlation. Polydispersity in the case of self-assembled material was never exceeding 16 % showing the narrow polydispersity of the observed structures. (Mg²⁺) means Mg²⁺ has been added to reach a concentration of 7.6 mM.

2.7.8 Gel electrophoresis

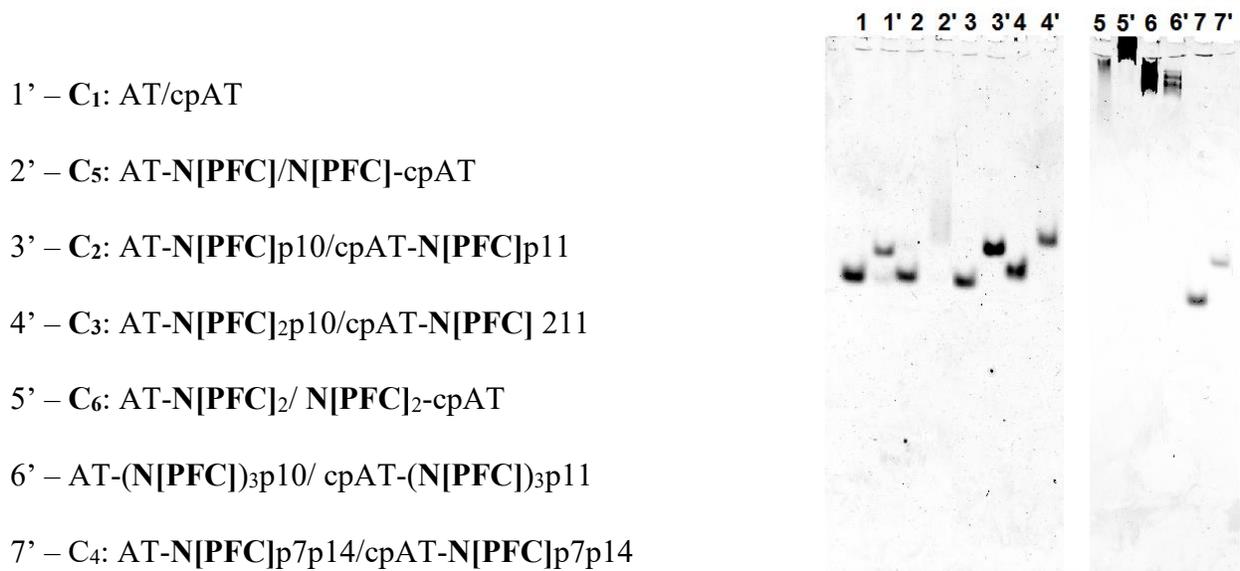
2.5 % agarose gel electrophoresis (AGE) was carried out in two different buffers. The first one, TAE, does contain EDTA whereas the second one, TAMg contains Mg²⁺ cations. They were carried out at 4 °C for respectively 2h and 2h15 at 80V. Gel was cast in the appropriate buffer and the samples were at a concentration of 1 μ M in the appropriate buffer. 2 μ l of glycerol were added to the samples before loading. The DNA bands for all gels were visualized by incubation with GelRed™.

15 % denaturing Polyacrylamide Gel Electrophoresis (PAGE) was carried out at room temperature for 30 minutes at 250V followed by 1 hour at 500V. TBE buffer (1X) was used and the concentration of urea in the gel was 7M. For each lane 5 μ L of sample (2 μ M) in water was added to 5 μ L of 8M urea. The DNA bands for all gels were visualized by incubation with GelRed™.

8 % Native PAGE was carried out at room temperature for 2.5 hours at a constant voltage of 250V in 1X TAMg buffer. Sample loading was 0.01 nmol ssDNA or 0.002 nmol dsDNA per lane (12 μ l samples in 1 X TAMg, including 2 μ l of glycerol).



Figure 2.21. 15 % denaturing PAGE analysis of “DNA-Teflon” oligomers. The DNA is the AT sequence and the length of the PFC tail vary from 0 to 5 PFC units (B_n polymers). From $n=4$, self-assembly occurs and the material does not seem able to penetrate in the gel (non-penetrating band).



- 1' – C₁: AT/cpAT
- 2' – C₅: AT-N[PFC]/N[PFC]-cpAT
- 3' – C₂: AT-N[PFC]_{p10}/cpAT-N[PFC]_{p11}
- 4' – C₃: AT-N[PFC]_{2p10}/cpAT-N[PFC]₂₁₁
- 5' – C₆: AT-N[PFC]₂/ N[PFC]₂-cpAT
- 6' – AT-(N[PFC]₃)_{p10}/ cpAT-(N[PFC]₃)_{p11}
- 7' – C₄: AT-N[PFC]_{p7p14}/cpAT-N[PFC]_{p7p14}

Figure 2.22. 8 % native PAGE of PFC-modified duplexes. Lanes without ' are ssDNA. The ssDNA and duplexes behave well except in lanes 2', 5, 5', 6 and 6' for which self-assembly occurs. For duplexes 2 and 5, we hypothesize aggregation occurs but DNA hybridization remains (cf. melting curves).

2.7.9 NMR study

^{19}F NMR spectra of the following molecules were recorded for quantitation using the PULCON method: molecule **2** in DMSO- d_6 (external reference), and **B1**, **B4**, and **B10** in D_2O . The concentrations of **B1**, **B4**, and **B10** were also determined by UV absorption at 260 nm. The NMR analysis focused on the CF_3 signal of molecule **2** related to the CF_3 end group of the C_8F_{17} chain as the CF_3 signals were the most intense. T_1 and T_2 values were measured using an inversion recovery sequence and a CPMG sequence, respectively, using between 5 and 8 appropriate tau values. Fitting was done using the Bruker Dynamics Center.

All quantitative ^{19}F spectra were acquired using the same experimental conditions, with the exception of the transmitter offset, which was always placed on the CF_3 peak (-80.1 ppm for molecule **2**, -80.8 ppm for **B1** and -83.3 ppm for **B4**, and **B10**). The DEPTH method⁴⁷ along with a pre-acquisition delay of 35 μs and the Bruker baseopt method for baseline correction at acquisition time were used to suppress the large ^{19}F background arising from probe components. A slight baseline hump was still present in the samples with the lowest S/N ratios (**B4** and **B10**) and the Bruker routine cryoprocl1d was used to replace the first points of those FIDs with backwards linearly predicted points. Automatic tuning and matching were performed before each experiment. Because signal-to-noise was poor, the ^{19}F 90° pulse was not measured before each experiment, but the ^1H 90° pulse was measured for all samples and found not to vary, so it was assumed that the ^{19}F was also consistent for all samples. In the quantitative NMR experiments, a recycle delay of 6 s was used for best reproducibility (the T_1 of molecule **2** was found to be 1.3 s). The temperature was controlled at 300K during the experiments. The receiver gain was kept constant at 203 for all experiments.

Concentration calculations were done within the ERETIC 2 implementation in Bruker TopSpin 3.5 pl 2 software, based on the PULCON (pulse length based concentration determination) method.⁴⁸ The external reference, molecule **2**, was prepared using careful weighing and dilution. The results of the measurement as applied to the CF_3 peaks of each sample are shown in **Table 2.4**.

Sensitivity measurements: We measured the signal-to-noise ratios (S/N) of **B4** and **B10** in samples of the same concentration (46 μM). Conditions were the same as explained above except that the

acquisition time was reduced to 28 ms, the recycle delay was reduced to 0.46 s, and the number of scans was set to 1800, with 8 dummy scans, for a total experimental time of 15 min 33s. To further increase the S/N of the 46 μM **B**₁₀ sample, the acquisition time was reduced to 14 ms, the recycle delay was reduced to 0.43 s, and the number of scans was set to 1900 with 8 dummy scans, for a total experiment time of 15 min 10 s.

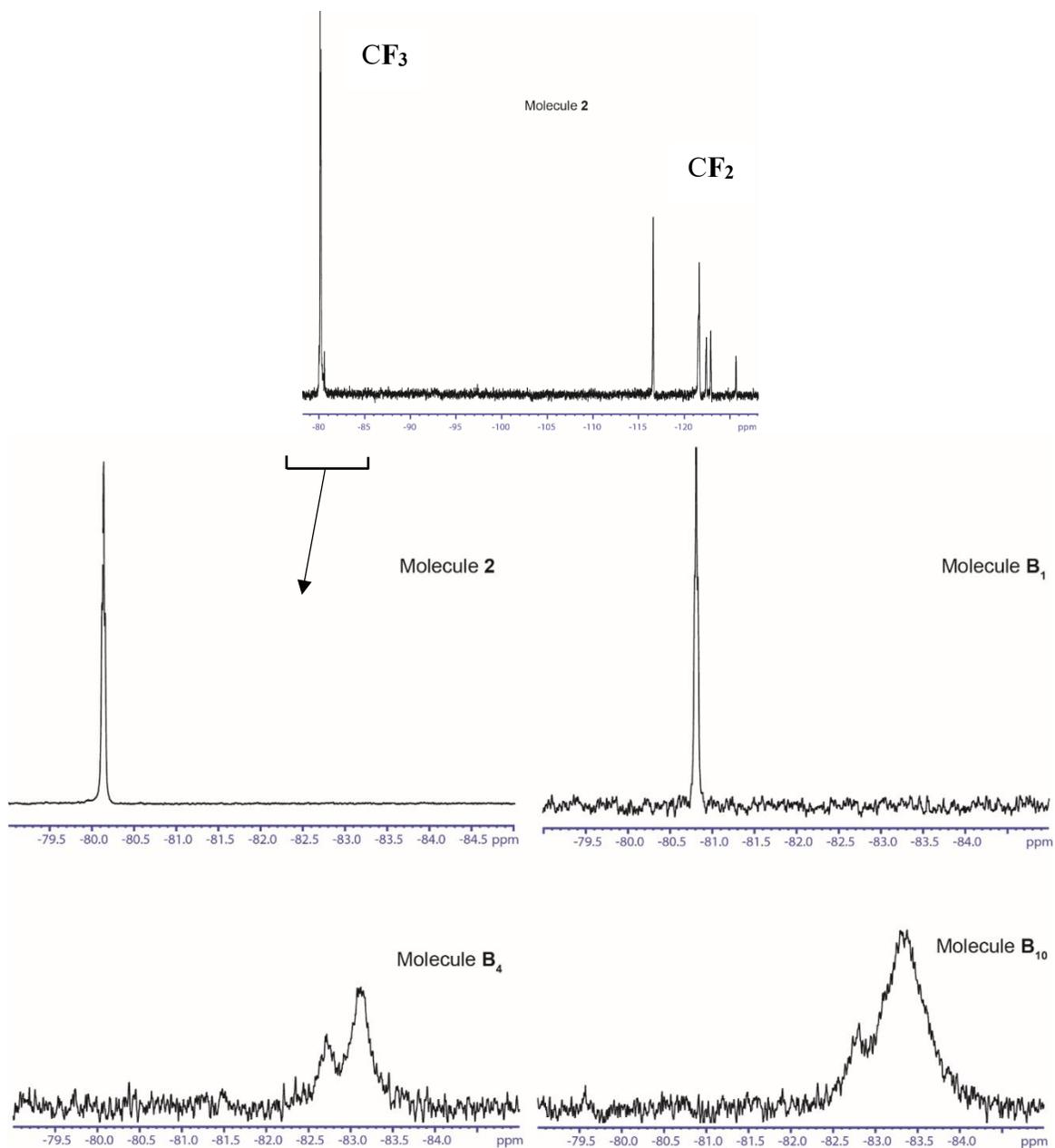


Figure 2.23. ^{19}F NMR spectra of “DNA-Teflon” oligomers. Top: ^{19}F NMR spectrum of molecule **2** (80 μM in $\text{DMSO-}d_6$). Bottom: ^{19}F NMR spectra of molecules **2** (800 μM in $\text{DMSO-}d_6$), **B**₁ (71 μM in D_2O), **B**₄ (39 μM in D_2O), **B**₁₀ (46 μM in D_2O). Only the strong signal from CF_3 is visible at low concentrations.

2.7.10 Melting curves

Experiments were carried out in quartz cuvettes (rectangular, 10 mm, 80 μ L) in triplicate. Absorbance was measured at 260 nm and detected in increments of 1 $^{\circ}$ C from 20 $^{\circ}$ C to 50 $^{\circ}$ C and from 75 $^{\circ}$ C to 90 $^{\circ}$ C of 0.4 $^{\circ}$ C from 50 $^{\circ}$ C to 75 $^{\circ}$ C. Concentration of DNA duplexes was 4 μ M in 1XTAMg buffer. The heating rate used was 1 $^{\circ}$ C/min.

T_M are calculated by taking the temperatures corresponding to the derivative maxima of the curves obtained and are reported in Table 2.7.

Table 2.7. Melting temperature of PFC-containing DNA duplexes.

Duplex name	Strands in the duplex	Melting temperature ($^{\circ}$ C)
C ₁	AT/cpAT	62.6 \pm 0.5.
C ₂	AT-(N [PFC])p10/cpAT-(N [PFC])p11	63.1 \pm 0.4
C ₃	AT-(N [PFC]) ₂ p10/cpAT-(N [PFC]) ₂ p11	70.6 \pm 0.5
C ₄	AT-(N [PFC])p7p14/cpAT-(N [PFC])p7p14	63.3 \pm 0.5
C ₅	AT-(N [PFC])/N[PFC]-cpAT	65.3 \pm 0.7
C ₆	AT-(N [PFC]) ₂ /N[PFC] ₂ -cpAT	82.6 \pm 0.8

To draw van't Hoff plots, we measured the melting temperature of the duplexes **C**₁ to **C**₆ (except for **C**₄) and of an unmodified duplex lengthened with 2 bases (GC) at position 10 at the concentrations: 4, 6, 9, 12, 16 and 20 μ M (total concentration of DNA strands) in triplicates. Representative curves are shown thereafter. We then used these numbers to plot 1/ T_M as a function of $\ln(C_T/4)$. Thermodynamic data can be extracted from the data using the following equation:⁴²

$$Tm^{-1} = R \ln(Ct/4)/\Delta H^{\circ} + [\Delta S^{\circ}/\Delta H^{\circ}].$$

All R squared values of linear fits were found to be over 0.99 when removing at most one T_M value per sample. The average thermodynamic values are reported in Table 2.8.

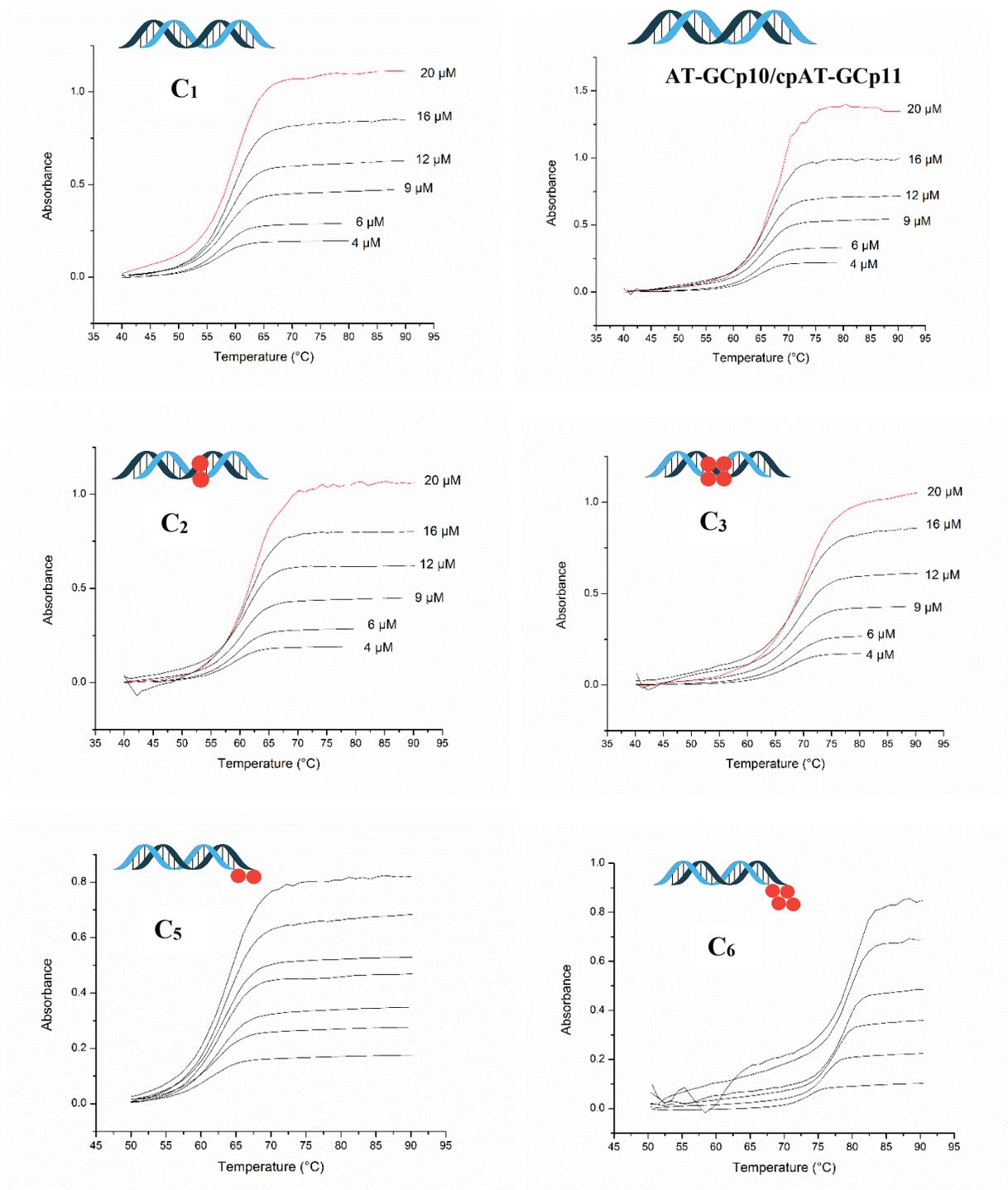


Figure 2.24. Representative melting curves of PFC-modified DNA duplexes at several concentrations in TAMg.

Table 2.8. Average thermodynamic values found for PFC-modified duplexes

Duplex	ΔH (Mcal/mol)	ΔS (kcal/mol/K)	ΔG (kcal/mol)	$\Delta\Delta H$ (Mcal/mol)	$\Delta\Delta S$ (kcal/mol/K)	$\Delta\Delta G$ (kcal/mol)
C₁	-0.13 ± 0.02	-0.36 ± 0.05	-16 ± 1	0	0	0
ATGCp10	-0.16 ± 0.02	-0.44 ± 0.05	-21 ± 2	-0.03	-0.09	-5
C₂	-0.129 ± 0.008	-0.36 ± 0.03	-17.1 ± 0.7	-0.002	-0.004	-0.7
C₃	-0.12 ± 0.01	-0.34 ± 0.03	-19.5 ± 1	0.004	0.02	-3
C₅	-0.12 ± 0.01	-0.34 ± 0.03	-17.5 ± 0.7	0.004	0.02	-1
C₆	-0.067 ± 0.005	-0.17 ± 0.02	-15.6 ± 0.1	0.06	0.2	0.8

2.7.11 Serum stability assay

For degradation studies, DNA duplexes were concentrated to a stock solution of 40 μ M in 1XTAMg buffer. As an example, duplex AT/cpAT (40 μ M, 2 μ L) was first diluted with DMEM media (88 μ L). To this mixture was added a fresh sample of undiluted FBS (10 μ L) with slight mixing to make the overall % of FBS 10 % (v/v). An aliquot was immediately taken out (10 μ L), formamide (10 μ L) added and then stored at -20 °C as the t=0 h time point. The remaining sample was then incubated at 37 °C and similar aliquots were removed and treated as described above at time points of 35 min, 1, 2, 4, 6, 8.5 and 24 hr. Digested products were analyzed by denaturing PAGE (20 %, 15 mA, 250V during 30 min followed by 500V 2 hr).

Higher band intensity was extracted and divided by the intensity at t=0. These experiments were run in triplicate. Data was analyzed using GraphPad Prism with a one phase decay exponential fit. The general equation is : $Y = (Y_0 - \text{Plateau}) * \exp(-K * X) + \text{Plateau}$. Here, Y is the relative band intensity, X is time, the plateau is supposed to be 0 while Y₀ is equal to one. Half-life is defined as $\ln(2)/K$.

Table 2.9. Summary of exponential decay analysis for duplexes degradation in serum.

Duplex name	R squared	Decay constant k (h ⁻¹)	Half-life (h)
C ₁	0.99	0.75	0.9
C ₂	0.98	0.34	2.0
C ₃	0.90	0.26	2.7
C ₄	0.97	0.45	2.2
C ₅	0.96	0.57	1.2
C ₆	NO	NO	NO

The one-phase decay model may not be fully adapted to a few strands with poor R squared but this fit was used in all cases allowing comparisons. Everytime, Y0 and plateau are respectively very close to 1 and 0 except for strand C₅ for which plateau is equal to 0.12. This result highlights that the model chosen is not adapted to this strand.

Only non-penetrating material was observed in the case of C₆ analysis. We hypothesized that the absorbance in the well is proportional to the total amount of oligonucleotides. This quantity was compared to the total intensity in each line during C₁ analysis. It appeared that almost 50 % of the intensity remains after 24h while 3 % was detected after 24h in the case of C₁ and 5 % in the case of C₆. After 8.5h, more than 75 % of the intensity remains for C₆ compared to 20 % in the case of C₁.

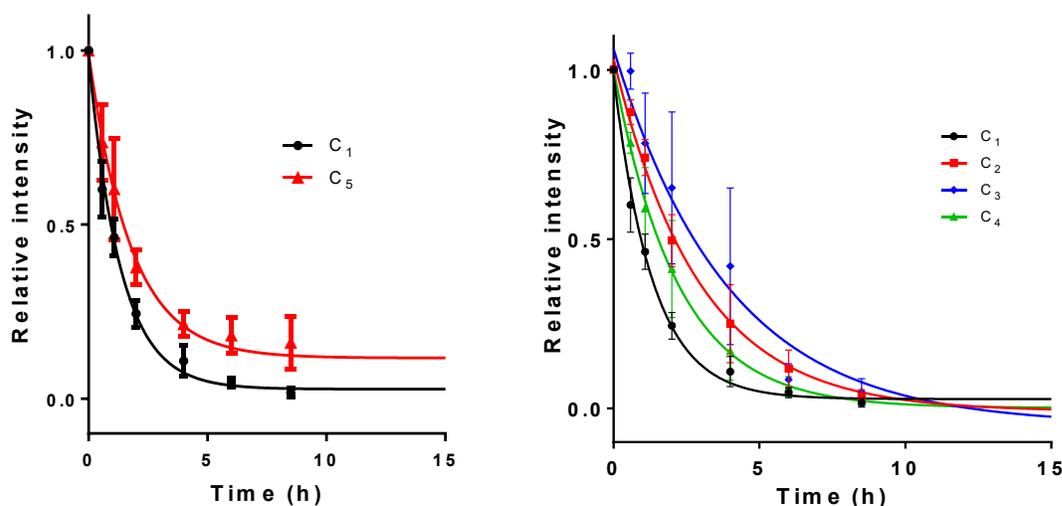


Figure 2.25. Average of DNA-duplexes degradation curves. The full length product is still visible after 24h in the case of C_5 which makes the one phase decay model inaccurate.

2.7.12 Gene silencing assay

2.7.12.1 Transfection of ApoB-siRNA and N[PFC]-ApoB-siRNA

HepG2 cells (Human hepatocellular carcinoma) were seeded in a 24-well plate at a density of 5×10^4 cells. Cells were allowed to adhere overnight. Transfection was achieved by diluting $2 \mu\text{M}$ stocks of pre-annealed siRNA with OptiMEM (Life) and Oligofectamine (Life) to attain a final concentration of 20 nM of nucleic acid therapeutics. After an incubation period of 24 hours, we proceeded to isolate the RNA.

2.7.12.2 RNA isolation

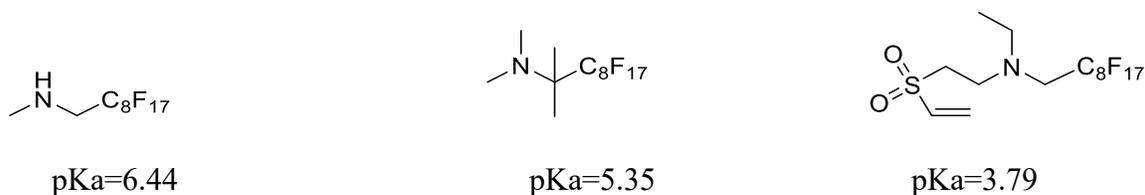
Total RNA was isolated from the 24-well plate by using an RNeasy mini kit (QIAGEN, Hilden, Germany) as described by the manufacturer. Genomic DNA was eliminated by RNase-free DNase I treatment during the isolation procedure. Reverse transcription was performed using the iScript Advanced cDNA Synthesis Kit (Bio-Rad) according the manufacturer's protocol. In a typical reaction 500 ng of RNA was mixed with $4 \mu\text{l}$ of reaction buffer and $1 \mu\text{l}$ of reverse transcriptase and the volume made up to $20 \mu\text{l}$ with nuclease-free water. Reverse transcription was performed at $42 \text{ }^\circ\text{C}$ for 30 min and inactivated at $85 \text{ }^\circ\text{C}$ for 5 minutes.

2.7.12.3 Quantitative real-time PCR.

Gene quantification was performed with a Step-One Plus (Life). Primers for ApoB used: Forward - 5'-TTTGCCCTCAACCTACCAAC-3' and Reverse - 5'-TGCGATCTTGTTGGCTACTG-3'. GAPDH was used as an endogenous control with the following primers: Forward -5'-GGAGCGAGATCCCTCCAAAAT-3' and Reverse - 5'-GGCTGTTGTCATACTTCTCATGG-3'. Each PCR was performed in a 20 μ l reaction mixture containing 10 μ l of SsoAdvanced universal SYBR Green supermix (Bio-Rad) and 250 nM of each primer. The thermal cycling conditions were as follows: 30 sec at 95 $^{\circ}$ C, followed by 40 cycles of 15 s at 95 $^{\circ}$ C, 30 s at 60 $^{\circ}$ C. Data collection was performed during each extension phase. A negative control (distilled water), and RT-negative controls (total RNA sample) were included in each run. For each of the RNA extractions, measurements of gene expression were obtained in triplicate, and the mean of these values was used for further analysis using $\Delta\Delta$ Ct method for relative quantification.

2.7.13 pKa of N[PFC]

pKa of protonated counterpart of molecule **2** should be under 6.5 after comparison with SciFinder[®] theoretical values for very similar compounds:



Scheme 2.3. pKa of protonated counterpart of perfluorocarbon containing amines found on SciFinder.

2.8 References

- (1) Edwardson, T. G. W.; Carneiro, K. M. M.; Serpell, C. J.; Sleiman, H. F. An Efficient and Modular Route to Sequence-Defined Polymers Appended to {DNA}. *Angew. Chem. Int. Ed. Engl.* **2014**, *53* (18), 4567–4571.
- (2) Serpell, C. J.; Edwardson, T. G. W.; Chidchob, P.; Carneiro, K. M. M.; Sleiman, H. F. Precision Polymers and 3D DNA Nanostructures: Emergent Assemblies from New Parameter Space. *J. Am. Chem. Soc.* **2014**, *136* (44), 15767–15774.
- (3) Chidchob, P.; Edwardson, T. G. W.; Serpell, C. J.; Sleiman, H. F. Synergy of Two Assembly Languages in DNA Nanostructures: Self-Assembly of Sequence-Defined Polymers on DNA Cages. *J. Am. Chem. Soc.* **2016**, *138* (13), 4416–4425.

- (4) Dore, M. D.; Fakhoury, J. J.; Lacroix, A.; Sleiman, H. F. Templated Synthesis of Spherical RNA Nanoparticles with Gene Silencing Activity. *Chem. Commun.* **2018**, 54 (80), 11296–11299.
- (5) Tomalia, D. A. Supramolecular Chemistry: Fluorine Makes a Difference. *Nat. Mater.* **2003**, 2 (11), 711–712.
- (6) Clark, L. C.; Gollan, F. Survival of Mammals Breathing Organic Liquids Equilibrated with Oxygen at Atmospheric Pressure. *Science* **1966**, 152 (3730), 1755–1756.
- (7) Spahn, D. R. Blood Substitutes. Artificial Oxygen Carriers: Perfluorocarbon Emulsions. *Crit. Care* **1999**, 3 (5), R93-7.
- (8) Castro, C. I.; Briceno, J. C. Perfluorocarbon-Based Oxygen Carriers: Review of Products and Trials. *Artif. Organs* **2010**, 34 (8), 622–634.
- (9) Chen, J.; Lanza, G. M.; Wickline, S. A. Quantitative Magnetic Resonance Fluorine Imaging: Today and Tomorrow. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **2010**, 2 (4), 431–440.
- (10) Ruiz-Cabello, J.; Barnett, B. P.; Botto mLey, P. A.; Bulte, J. W. M. Fluorine (¹⁹F) MRS and MRI in Biomedicine. *NMR Biomed.* **2011**, 24 (2), 114–129.
- (11) Vierling, P.; Santaella, C.; Greiner, J. Highly Fluorinated Amphiphiles as Drug and Gene Carrier and Delivery Systems. *J. Fluor. Chem.* **2001**, 107 (2), 337–354.
- (12) Wang, M.; Liu, H.; Li, L.; Cheng, Y. A Fluorinated Dendrimer Achieves Excellent Gene Transfection Efficacy at Extremely Low Nitrogen to Phosphorus Ratios. *Nat. Commun.* **2014**, 5, 3053.
- (13) Wang, H.; Wang, Y.; Wang, Y.; Hu, J.; Li, T.; Liu, H.; Zhang, Q.; Cheng, Y. Self-Assembled Fluorodendrimers Combine the Features of Lipid and Polymeric Vectors in Gene Delivery. *Angew. Chem. Int. Ed. Engl.* **2015**.
- (14) Kasuya, M. C. Z.; Nakano, S.; Katayama, R.; Hatanaka, K. Evaluation of the Hydrophobicity of Perfluoroalkyl Chains in Amphiphilic Compounds That Are Incorporated into Cell Membrane. *J. Fluor. Chem.* **2011**, 132 (3), 202–206.
- (15) Dias, N.; Stein, C. A. Antisense Oligonucleotides: Basic Concepts and Mechanisms. *Mol. Cancer Ther.* **2002**, 1 (5), 347–355.
- (16) Shen, X.; Corey, D. R. Chemistry, Mechanism and Clinical Status of Antisense Oligonucleotides and Duplex RNAs. *Nucleic Acids Res.* **2018**, 46 (4), 1584–1600.
- (17) Godeau, G.; Arnion, H.; Brun, C.; Staedel, C.; Barthélémy, P. Fluorocarbon Oligonucleotide Conjugates for Nucleic Acids Delivery. *Medchemcomm* **2010**, 1 (1), 76.
- (18) Ellipilli, S.; Vasudeva Murthy, R.; Ganesh, K. N. Perfluoroalkylchain Conjugation as a New Tactic for Enhancing Cell Permeability of Peptide Nucleic Acids (PNAs) via Reducing the Nanoparticle Size. *Chem. Commun. (Camb).* **2016**, 52 (3), 521–524.
- (19) Pei, H.; Zuo, X.; Zhu, D.; Huang, Q.; Fan, C. Functional DNA Nanostructures for Theranostic

- Applications. *Acc. Chem. Res.* **2014**, *47* (2), 550–559.
- (20) Cametti, M.; Crousse, B.; Metrangolo, P.; Milani, R.; Resnati, G. The Fluorous Effect in Biomolecular Applications. *Chem. Soc. Rev.* **2012**, *41* (1), 31–42.
- (21) Berger, R.; Resnati, G.; Metrangolo, P.; Weber, E.; Hulliger, J. Organic Fluorine Compounds: A Great Opportunity for Enhanced Materials Properties. *Chem. Soc. Rev.* **2011**, *40* (7), 3496.
- (22) Oishi, M.; Sumitani, S.; Nagasaki, Y. On-off Regulation of ¹⁹F Magnetic Resonance Signals Based on PH-Sensitive PEGylated Nanogels for Potential Tumor-Specific Smart ¹⁹F MRI Probes. *Bioconjug. Chem.* **2007**, *18* (5), 1379–1382.
- (23) Takaoka, Y.; Sakamoto, T.; Tsukiji, S.; Narazaki, M.; Matsuda, T.; Tochio, H.; Shirakawa, M.; Hamachi, I. Self-Assembling Nanoprobes That Display off/on ¹⁹F Nuclear Magnetic Resonance Signals for Protein Detection and Imaging. *Nat. Chem.* **2009**, *1* (7), 557–561.
- (24) Wang, H.; Raghupathi, K. R.; Zhuang, J.; Thayumanavan, S. Activatable Dendritic ¹⁹F Probes for Enzyme Detection. *ACS Macro Lett.* **2015**, 422–425.
- (25) Bo, S.; Song, C.; Li, Y.; Yu, W.; Chen, S.; Zhou, X.; Yang, Z.; Zheng, X.; Jiang, Z.-X. Design and Synthesis of Fluorinated Amphiphile as (¹⁹F) MRI/Fluorescence Dual-Imaging Agent by Tuning the Self-Assembly. *J. Org. Chem.* **2015**, *80* (12), 6360–6366.
- (26) Lutz, J.-F. Defining the Field of Sequence-Controlled Polymers. *Macromol. Rapid Commun.* **2017**, *38* (24), 1700582.
- (27) Pfeifer, S.; Lutz, J.-F. A Facile Procedure for Controlling Monomer Sequence Distribution in Radical Chain Polymerizations. *J. Am. Chem. Soc.* **2007**, *129* (31), 9542–9543.
- (28) Zamfir, M.; Lutz, J.-F. Ultra-Precise Insertion of Functional Monomers in Chain-Growth Polymerizations. *Nat. Commun.* **2012**, *3*, 1138.
- (29) Terashima, T.; Mes, T.; De Greef, T. F. A.; Gillissen, M. A. J.; Besenius, P.; Palmans, A. R. A.; Meijer, E. W. Single-Chain Folding of Polymers for Catalytic Systems in Water. *J. Am. Chem. Soc.* **2011**, *133* (13), 4742–4745.
- (30) Li, Z.-L.; Li, L.; Deng, X.-X.; Zhang, L.-J.; Dong, B.-T.; Du, F.-S.; Li, Z.-C. Periodic Vinyl Copolymers Containing γ -Butyrolactone via ADMET Polymerization of Designed Diene Monomers with Built-in Sequence. *Macromolecules* **2012**, *45* (11), 4590–4598.
- (31) Satoh, K.; Ozawa, S.; Mizutani, M.; Nagai, K.; Kamigaito, M. Sequence-Regulated Vinyl Copolymers by Metal-Catalysed Step-Growth Radical Polymerization. *Nat. Commun.* **2010**, *1*, 6.
- (32) Lewandowski, B.; De Bo, G.; Ward, J. W.; Pappmeyer, M.; Kuschel, S.; Aldegunde, M. J.; Gramlich, P. M. E.; Heckmann, D.; Goldup, S. M.; D'Souza, D. M.; et al. Sequence-Specific Peptide Synthesis by an Artificial Small-Molecule Machine. *Science* **2013**, *339* (6116), 189–193.
- (33) O'Reilly, R. K.; Turberfield, A. J.; Wilks, T. R. The Evolution of DNA-Templated Synthesis as a Tool for Materials Discovery. *Acc. Chem. Res.* **2017**, *50* (10), 2496–2509.

- (34) Zuckermann, R. N.; Kerr, J. M.; Kent, S. B. H.; Moos, W. H. Efficient Method for the Preparation of Peptoids [Oligo(N-Substituted Glycines)] by Submonomer Solid-Phase Synthesis. *J. Am. Chem. Soc.* **1992**, *114* (26), 10646–10647.
- (35) Sun, J.; Zuckermann, R. N. Peptoid Polymers: A Highly Designable Bioinspired Material. *ACS Nano* **2013**, *7* (6), 4715–4732.
- (36) Ouahabi, A. Al; Kotera, M.; Charles, L.; Lutz, J.-F. Synthesis of Monodisperse Sequence-Coded Polymers with Chain Lengths above DP100. *ACS Macro Lett.* **2015**, *4* (10), 1077–1080.
- (37) Braasch, D. A.; Corey, D. R. Locked Nucleic Acid (LNA): Fine-Tuning the Recognition of DNA and RNA. *Chem. Biol.* **2001**, *8* (1), 1–7.
- (38) Kedracki, D.; Safir, I.; Gour, N.; Ngo, K. X.; Vebert-Nardin, C. {DNA}–Polymer Conjugates: From Synthesis, Through Complex Formation and Self-Assembly to Applications. In *Bio-synthetic Polymer Conjugates*; Schlaad, H., Ed.; Advances in Polymer Science; Springer Berlin Heidelberg, 2013; pp 115–149.
- (39) Kwak, M.; Herrmann, A. Nucleic Acid Amphiphiles: Synthesis and Self-Assembled Nanostructures. *Chem. Soc. Rev.* **2011**, *40* (12), 5745–5755.
- (40) Bousmail, D.; Amrein, L.; Fakhoury, J. J.; Fakih, H. H.; Hsu, J. C. C.; Panasci, L.; Sleiman, H. F. Precision Spherical Nucleic Acids for Delivery of Anticancer Drugs. *Chem. Sci.* **2017**, *8* (9), 6218–6229.
- (41) Ahrens, E. T.; Zhong, J. In Vivo MRI Cell Tracking Using Perfluorocarbon Probes and Fluorine-19 Detection. *NMR Biomed.* **2013**, *26* (7), 860–871.
- (42) Marky, L. A.; Breslauer, K. J. Calculating Thermodynamic Data for Transitions of Any Molecularity from Equilibrium Melting Curves. *Biopolymers* **1987**, *26* (9), 1601–1620.
- (43) Conway, J. W.; McLaughlin, C. K.; Castor, K. J.; Sleiman, H. DNA Nanostructure Serum Stability: Greater than the Sum of Its Parts. *Chem. Commun. (Camb)*. **2013**, *49* (12), 1172–1174.
- (44) Soutschek, J.; Akinc, A.; Bra mLage, B.; Charisse, K.; Constien, R.; Donoghue, M.; Elbashir, S.; Geick, A.; Hadwiger, P.; Harborth, J.; et al. Therapeutic Silencing of an Endogenous Gene by Systemic Administration of Modified SiRNAs. *Nature* **2004**, *432* (7014), 173–178.
- (45) Lau, C.; Butenhoff, J. L.; Rogers, J. M. The Developmental Toxicity of Perfluoroalkyl Acids and Their Derivatives. *Toxicol. Appl. Pharmacol.* **2004**, *198* (2), 231–241.
- (46) Hintzer, K.; Zipplies, T.; Carlson, D. P.; Schmiegel, W. Fluoropolymers, Organic. In *Ullmann's Encyclopedia of Industrial Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2014; pp 1–55.
- (47) Cory, D. .; Ritchey, W. . Suppression of Signals from the Probe in Bloch Decay Spectra. *J. Magn. Reson.* **1988**, *80* (1), 128–132.
- (48) Wider, G.; Dreier, L. Measuring Protein Concentrations by NMR Spectroscopy. *J. Am. Chem. Soc.* **2006**, *128* (8), 2571–2576.

|3|

**Self-Assembled Nanostructure Library from
Monodisperse Sequence-Defined
Oligo(Phosphodiester)s**

3.1 Preface

We previously reported that hydrophilic hexaethylene glycol and hydrophobic 12 carbon alkyl chain monomers can be efficiently introduced in precision oligomers.¹ The amphiphobic perfluorocarbon (PFC) building block developed in Chapter 2 is a valuable addition to our unnatural monomers collection. Indeed, we have shown the great input PFC chains impart in sequence-defined oligo(phosphodiester)s. An artificial unit capable of π - π stacking would further complement the monomer alphabet for sequence-defined oligo(phosphodiester)s.

With such a palette of interactions available, self-assembling materials may be rationally designed to further explore opportunities with precision polymers. However, such *a priori* predictions are difficult to perform, as illustrated by the protein folding problem.² In-depth analysis of the rules governing the folding and self-assembly of precision polymers are required. Systematic studies of sequence variation impact in polymers self-assembly would give an insight into these rules and are therefore of great interest. In this chapter, we take advantage of the practical automated synthesis of sequence-defined oligo(phosphodiester)s to realize such a study.

This chapter is mostly composed of work posted online as « Self-assembled Nanostructure Library from Monodisperse Sequence-Defined Oligo(phosphodiester)s » by D de Rochambeau, M Barlóg, HS Bazzi & HF Sleiman; ChemRxiv. Preprint. February 2019.

3.2 Contribution of authors

Donatien de Rochambeau codesigned the project and performed all experiments unless listed below, analyzed the results and cowrote the manuscript. **Dr. Maciej Barłóg** synthesized compounds **NAP''**, **NAP'** and **C12'** **Dr. Violeta Toader** synthesized compound **C12**. **Dr. Hanadi Sleiman** codesigned the project, guided interpretation of data and discussion of results and cowrote the manuscript.

3.3 Abstract

Natural biopolymers achieve information storage, molecular recognition and catalysis efficiently through sequence-control. To be able to mimic such properties, self-assembly studies of artificial sequence-defined oligomers are of great interest. In this chapter, we synthesized a naphthalene containing monomer and showed its use along with hydrophilic, lipophilic, and fluorophilic monomers to make a large library of truly monodisperse sequence-defined block co-oligo(phosphodiester)s. Automated and accurate control over the sequence allowed the rational study of how the degree of polymerisation, block ratio, chemical composition and orthogonal supramolecular interactions influence self-assembly. Interestingly, our studies revealed remarkable morphological changes (spheres to nanosheets) caused by very small differences between polymers, e.g., oligomers differing by a single monomer unit. Inverting block sequence in multi-block co-oligomers also caused an increase in micelle size. Conventional polymerization does not allow the exploration of these subtle variations in polymer sequence or composition. Therefore, fast synthesis and purification of a variety of oligomers with slightly different sequences enables the study of the supramolecular chemistry of precision oligomers in a systematic way. It paves the way to the rational design of functional sequence-defined polymers.

3.4 Introduction

Nature uses sequence-defined polymers for molecular recognition, catalysis and information coding. DNA and proteins are biopolymers of specific lengths and sequences, allowing them to fold into highly functional macromolecules. To mimic such bio-macromolecules with artificial components, many new synthetic routes have been developed in the past few years.^{3,4} They rely on state-of-the-art polymerisation techniques,⁵⁻⁷ biopolymer templated synthesis⁸⁻¹⁰ or iterative

strategies.^{1,11-18} Among these, iterative approaches remain the method of choice to achieve the finest sequence-control. As demonstrated in the chapters 1 and 2, phosphoramidite chemistry on solid-phase has been shown to combine the advantages of high sequence precision,^{1,19-21} high degree of polymerisation,²² and water solubility of the resulting oligo(phosphodiester)s.²³ Very high coupling yields and facile purification result in monodisperse oligomers and polymers. In addition, this chemistry has been the method of choice to make oligonucleotides for decades, thus leading to cost-effectiveness and ready scale-up.²⁴ Few studies reported self-assembled structures with sequence-defined oligo(phosphodiester)s made of artificial building blocks.^{19,25-28} However, they do not integrate the use of multiple orthogonal interactions engineered in a sequence defined manner. Much like protein folding, these supramolecular interactions can result in controlled folding of single polymer chains and highly specific inter-chain assembly.²⁹ Studying the influence of such interactions is of great interest in order to design functional precision macromolecules.

The self-assembly of block copolymers (BCPs) has been extensively investigated for numerous applications including drug delivery, nanolithography, photoactive structures and porous materials.³⁰ BCPs can self-assemble into a range of structures: from simple spherical micelles to semi-crystalline cylindrical structures, vesicles or nanosheets.³¹⁻³⁵ The latter allow the 2D organization of organic and inorganic molecules with applications ranging from catalytic arrays, semiconductor materials or solar cells.³⁶ Among BCPs, diblock copolymers (AB) have been the most studied, leading to a detailed understanding of their self-assembly by microphase separation between the blocks.³⁷ The degree of polymerisation (N), the volume fraction of each block (f) and the Flory-Huggins parameter (χ), representative of the degree of incompatibility of both blocks, were shown to influence the polymer assembly at equilibrium in a non-solvent for one of the blocks. Block co-oligomers are block copolymers with low N . Some were shown to follow the assembly rules of BCPs but they sometimes have unusual behaviours especially when they are monodisperse.³⁸⁻⁴¹ Systematic sequence variations of sequence-defined block co-oligomers could give further insight into their self-assembly, and would be valuable for the ultimate rational design of functional macromolecules.

Herein, we report a method to readily synthesize sequence-defined block oligo(phosphodiester)s, rationally varying the three parameters (N , f_A and χ) shown to influence self-assembly of block copolymers. The automated synthetic strategy was key to simultaneously generate several

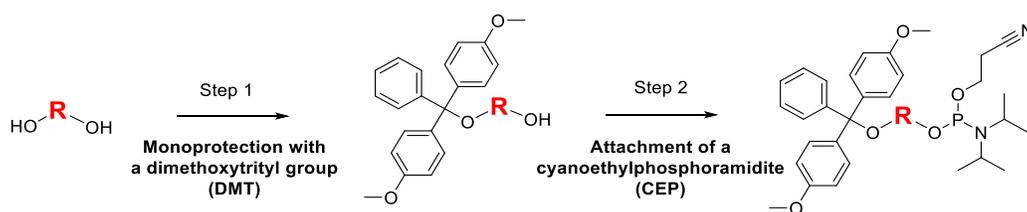
sequences with different degree of polymerisation (N) and blocks ratio (f_A). To implement π - π stacking interactions in the oligomers, we report the synthesis of a naphthalene containing monomer. Along with the hydrophilic and hydrophobic monomers used in an earlier report,¹ and the “fluorophilic” monomer of Chapter 2, we could precisely engineer multiple orthogonal interactions in oligomers and vary the Flory-Huggins parameter (χ). A library of 15 monodisperse block-oligomers with various sizes and sequences was synthesized. Due to their anionic nature, they were easily purified and analyzed using methods that have been already developed for nucleic acids, such as gel electrophoresis. The oligomers self-assembled in magnesium containing buffers. Variation of monomer type, blocks ratio, and length of the oligomers resulted in spherical micelles of different size, and nanosheets. Interestingly, our studies revealed the dramatic morphological effect caused by very subtle differences in the oligomer sequence. For example, adding a single naphthalene monomer unit on the same block co-oligomer chain changed the morphology from spheres to nanosheets, and inverting the sequence of two blocks on the same oligomer caused an increase in micelle size. A number of trends in precision oligo(phosphodiester)s self-assembly were deduced from this investigation. They highlight the valuable insights into supramolecular behaviour that can be obtained with sequence-defined oligomers.

3.5 Results and discussion

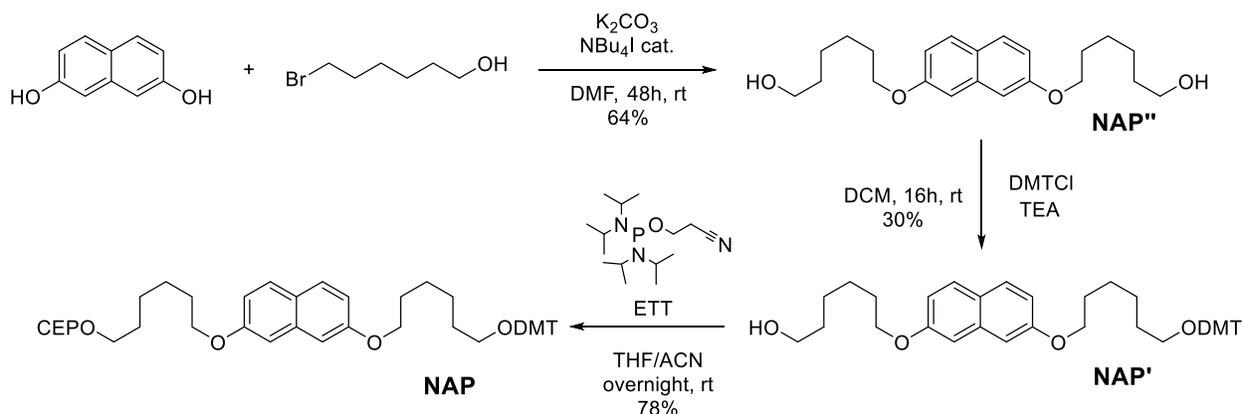
3.5.1 Phosphoramidite monomers from diols

We specifically chose four monomers that would allow us to engineer the hydrophobic effect, π - π stacking and the “fluorophilic” effect. The latter occurs when several perfluorocarbon chains show low affinity for surrounding molecules due to their amphiphobicity (ie hydrophobic and oleophobic). In the PFC-containing phosphoramidite (**PFC**) described in Chapter 2, the moiety of interest is a side chain. The other moieties used here are part of the oligomer backbone. They were made in a cost-effective and scalable manner through a simple synthetic strategy based on diol precursors. First, the diols were converted into mono DMT-protected alcohol and then turned into phosphoramidites as shown in Scheme 3.1. A large variety of monomers can be conceivably made through this simple method. Our group previously reported the synthesis of sequence-controlled polymers appended to DNA using monomers **C12** (12 carbon alkyl chain) and **HEG** (hexaethylene glycol).¹ Both were made from diol precursors and DNA oligomers containing these moieties were shown to self-assemble in water.^{19,25} We chose these two monomers as simple hydrophobic and

hydrophilic units respectively. For π - π stacking, we chose to synthesize a new monomer (**NAP**) that contains a naphthalene moiety. **NAP** synthesis started from 2,7 dihydroxynaphthalene which was turned to a suitable diol and transformed into a DNA synthesizer-compatible phosphoramidite (**Scheme 3.2**). **NAP** allows the oligomers to be UV-detectable, allowing their quantification through spectrophotometry. High efficiency couplings are achieved with **C12**, **HEG**, **NAP** and **PFC** showing that designing monomers using diols generally is an efficient strategy to generate sequence-controlled oligomers (**Figure 3.2**).



Scheme 3.1. General strategy to obtain a DNA synthesizer-compatible monomer from a diol.



Scheme 3.2. Synthetic pathway to make naphthalene-containing phosphoramidite monomer **NAP**.

3.5.2 Synthesis of sequence-defined block co-oligomers with a DNA synthesizer

Monomers **HEG**, **C12**, **NAP** and **PFC** were used to rationally design a library of sequence-controlled oligomers with different sequences and degrees of polymerisation in an automated fashion (**Figure 3.2**). The four phosphoramidites were used in a standard DNA synthesizer along with commercially available DNA synthesis reagents. Coupling cycles were similar to DNA synthesis, but the coupling time was increased to 10 minutes and phosphoramidite concentration

was kept at ca. 0.1 M in dry dichloromethane. With full synthetic cycles of 15 minutes, the synthesis of a 10-mer takes no longer than 2.5 h. Some polymers were made from a divergent method. For example, oligomers **A2** (**HEG₄-NAP₂**) and **A3** (**HEG₄-NAP₃**) were made from the same Controlled Pore Glass (CPG) support. After 4 coupling cycles with **HEG** and two with **NAP**, half of the CPG beads were deprotected and half of it underwent a last coupling cycle with a **NAP** unit. This divergent technique allowed to synthesize twice more oligomers. As a conclusion, considering the synthesizer preparation, handling of the CPG and deprotection (2.5 h in a basic aqueous medium), the synthesis of a 15-member oligomer library shown here can take less than 9 h with a common six ports MerMade® synthesizer. The phosphate groups of our polymers make them water-soluble, enhancing their great ease of manipulation. Similarly to nucleic acids, they can be quantified easily using UV absorbance at 260 nm and the extinction coefficient of **NAP** (see Section 3.7.4). The first series of oligomers (**A**) (**Table 3.1**) shows the possibility of making diblock copolymers with varying length and hydrophilic/hydrophobic ratios with monomers **HEG** and **NAP**. For example, **A2** (**HEG₄-NAP₂**) and **A4** (**HEG₄-NAP₄**) have different lengths and ratios while **A4** (**HEG₄-NAP₄**) and **A7** (**HEG₆-NAP₆**) only differ by the degree of polymerisation. Another series (**B**) shows sequences with up to four different artificial monomers (**Table 3.1**). While the **A** series involved more π - π stacking, the main supramolecular interactions for **B** oligomers are the hydrophobic and fluorophilic effect. For example, **B2** (**HEG₄-NAP-C12₄**) and **B4** (**HEG₄-NAP-PFC₄**) have similar length and sequence but the solvent incompatible block is mostly made of hydrophobic **C12** in one case and fluorophilic **PFC** units in the other. Synthesis of co-oligomers where fluorophilic, hydrophobic and π - π stacking should influence self-assembly was also achieved in high yields with **B5** (**HEG₆-NAP-PFC₄-C12₄**) and **B6** (**HEG₆-NAP-C12₄-PFC₄**).

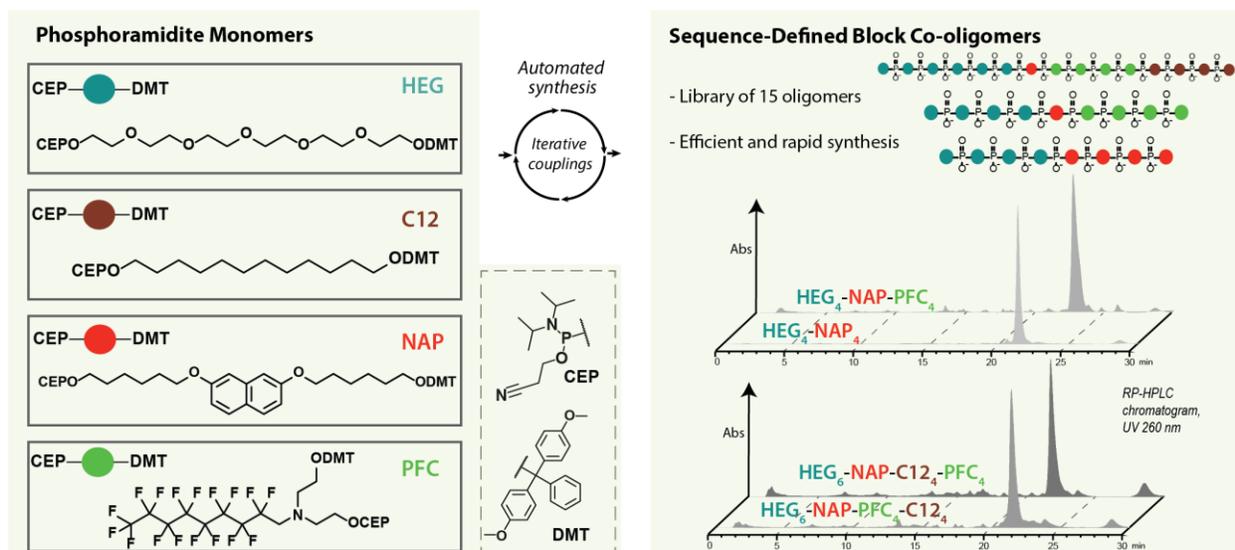


Figure 3.1. Strategy to obtain a library of precision oligomers with a DNA synthesizer. Chromatograms are from Reverse-Phase High Performance Liquid Chromatography (RP-HPLC) with detection at 260 nm. Gradients are detailed in **Figure 3.11**.

Table 3.1. Yields and MS characterization of the library of oligomers.

Oligomer	Sequence	HPLC Yields (%) ^a	Calculated exact mass (g mol ⁻¹)	Found exact mass ^a (g mol ⁻¹)
A1	HEG ₄ -NAP	75	1736.72	1736.73
A2	HEG ₄ -NAP ₂	72	2158.91	2158.93
A3	HEG ₄ -NAP ₃	71	2581.10	2581.11
A4	HEG ₄ -NAP ₄	91	3003.28	3003.04
A5	HEG ₄ -NAP ₆	41	3847.67	3847.67
A6	HEG ₆ -NAP ₃	65	3269.35	3269.36
A7	HEG ₆ -NAP ₆	44	4535.92	4535.94
A8	NAP ₈ -HEG ₁₂	- ^b	7445.01	7445.12
A9	NAP ₂ -HEG ₄ -NAP ₂	48	3003.28	3003.24
B1	HEG ₄ -C12 ₄ -NAP	86	2793.32	2793.27
B2	HEG ₄ -NAP-C12 ₄	83	2793.32	2793.34
B3	HEG ₆ -NAP-C12 ₄	87	3481.57	3481.51
B4	HEG ₄ -NAP-PFC ₄	76	4132.79	4132.52
B5	HEG ₆ -NAP-PFC ₄ -C12 ₄	79	5877.63	5877.72
B6	HEG ₆ -NAP-C12 ₄ -PFC ₄	70	5877.63	5877.73

^a Yields were calculated through product peak area by RP-HPLC (260 nm detection) considering the relative absorbance of by-products. Yields could only be calculated from the first naphthalene introduction. ^b Yields could not be measured by HPLC since the hydrophilic **HEG** monomers were added after the **NAP** monomers.

3.5.3 Purification of sequence-defined oligo(phosphodiester)s

All oligomers except **A8** were purified through reverse-phase high-performance liquid chromatography. Indeed, the addition of a hydrophobic phosphoramidite (**C12**, **NAP** or **PFC**) during the last coupling cycle led to a clear shift between the full-length product and the by-products on the chromatogram. Global yields were found to be very high in most cases (e.g. oligomer **B5**, 79 % for the last 8 couplings meaning an average of 97% per coupling, **Table 3.1**). Other types of sequence-defined polymers found in the literature are mostly isolated through reverse-phase HPLC.^{13,14} However, this method is not adapted to hydrophilic or long oligomers. For example, in the case of oligomer **A8** (**NAP**₈-**HEG**₁₂), hydrophilic monomer **HEG** was added last. Thus, the full-length oligomer and the n-1 impurity (minus one **HEG** unit) would have similar retention times on a reverse-phase chromatogram. In the polymer chemistry field, Size Exclusion Chromatography (SEC) is used as a major characterization and sometimes purification method. However, in the case of sequence-defined polymers, the resolution achieved by SEC is not good enough to obtain good separation between a polymer and a n-1 impurity.¹⁷ Our oligo(phosphodiester)s have the same anionic backbone as nucleic acids. Therefore, we tried to apply typical procedures used for DNA purification. Anionic exchange chromatography was considered but did not lead to satisfactory results in our case due to the oligomer self-assembly at high salt concentration. We also wanted to apply the “DMT-on” purification method. This strategy relies on leaving the hydrophobic DMT group on the last monomer added so that a clear shift is observed using reverse-phase chromatography. In the case of a model DNA 19mer with a DMT moiety at the 5' end, the DMT stayed on under usual HPLC conditions. However, it fell off from another strand modified at the 5' end with 6 **C12**. Therefore, this strategy is difficult to apply to a variety of oligomers (**Figure 3.2**).

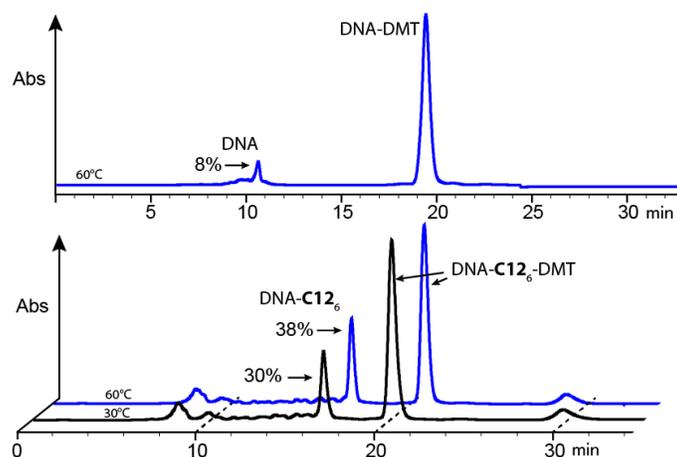


Figure 3.2. HPLC purification of DMT-on oligomers. RP-HPLC traces (UV 260 nm). Top: DNA with DMT-on. Peak area corresponding to DMT-off product is only 8 % the one for DMT-on. Run performed at 60 °C. Gradient: 3 to 50% ACN in 30 minutes. Bottom: DNA-C12₆ with DMT on. At 60 °C, the peak area corresponding to DMT-off product is 38 % the one for DMT-on. Moreover, this percentage decreased (30 %) if the temperature is set at 30 °C showing that some DMT cleavage is due to the HPLC run itself. Gradient: 3 to 80 % in 30 minutes. DNA sequence: 5'-TTTTTCAGTTGACCATATA-3' with 6 C12 at the 5' end.

We considered another classic method for DNA chemists more adapted to anionic oligomers: electrophoresis. This method was recently applied by Sutton and coworkers for the separation of oligo(acrylic acid)s through free solution capillary electrophoresis.⁴² We observed that oligomers containing **NAP** could be visualized after running polyacrylamide gel electrophoresis (PAGE) and staining with GelRed®, a molecule used to stain DNA (**Figure 3.3**). The cationic and aromatic GelRed® molecules most likely interact electrostatically with the anionic oligo(phosphodiester)s, and with **NAP** units via π - π stacking. This experiment showed clear electrophoretic mobility differences between oligomers of different size and sequence. With oligo(phosphodiester)s, PAGE appeared as a way to differentiate oligomers similar in size. Therefore, it can be used more accurately than SEC. Oligomers **A6**, **A7** and **A8** were purified via PAGE. They were loaded and run on a 20 % polyacrylamide gel and visualized with UV light (**Figure 3.3**). Main bands were excised, extracted from the gel using the “crush and soak” technique (see Section 3.7.6) and salt-purified using a small SEC filter. Hence, we showed that such polymers can be purified through gel electrophoresis due to their anionic nature leading to a truly monodisperse population of oligomers.

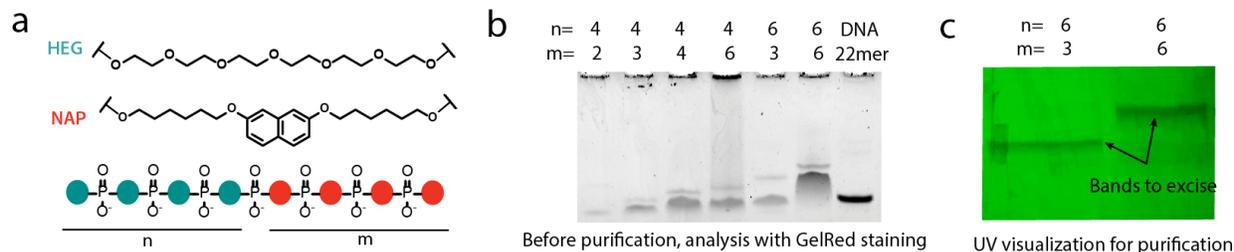


Figure 3.3. Gel electrophoresis analysis and purification of unnatural oligo(phosphodiester)s. (a) Schematic representation of the repeat units and the oligomers. (b) PAGE analysis (20 % in denaturing conditions) of the crude mixtures of oligomers obtained after cleavage from solid support. Interestingly, **HEG₄-NAP₆** seems to self-assemble even in denaturing conditions (non-penetrating band). (c) Visualization of oligomers **A6** and **A7** through UV 254 nm illumination of a 20 % polyacrylamide gel on a fluorescent TLC plate.

3.5.4 Self-assembly characterization

The amphiphilic nature of our oligomers made them good candidates for self-assembly in aqueous solvents. Inspired by previous work in our group,¹ we decided to mainly study the self-assembly properties of our oligomers in magnesium containing buffers (TAMg, see Section 3.7.1). Divalent magnesium cations interact with the phosphate negative charges minimizing repulsion between anionic oligomers. This allows the formation of micelles with a phosphate-containing hydrophobic core like in Chapter 2. Strikingly, the high solubility of all oligomers **A** and **B** in water due to the phosphate groups allows them to directly self-assemble from the dry state to the aqueous state. To ensure clean self-assembly of our oligomers, we annealed them by heating to 95 °C and cooling in 1 hour to 4 °C. Annealing prevents the formation of kinetically trapped nanostructures and favors the thermodynamic product. To characterize our structures, we performed Atomic Force Microscopy (AFM), Dynamic Light Scattering (DLS) and Agarose Gel Electrophoresis (AGE). AFM was performed in dry conditions first. We started our study with oligomer **A4** (**HEG₄-NAP₄**). Since typical procedures such as the ones used for the “DNA-Teflon” micelles did not lead to consistent results, we mainly explored three parameters: (i) the number of washing steps after sample deposition, (ii) the concentration of nanostructure and (iii) the buffer used. All these parameters play a crucial role in the nanostructure integrity and its ability to interact with the negatively charged mica surface. First, we tried to vary the buffer and the concentration of sample between 1 and 100 μM. Images obtained were not reproducible which made us hypothesize that the washing step with water following the sample deposition was disrupting the assembly. In dry

conditions, salts from buffers can sometimes form aggregates that may be interpreted as nanostructures from the sample. To allow the use of unwashed samples, we decided to image a buffer control at the beginning of each AFM session. We never observed structures similar to the ones described in this chapter. While low nanostructure concentration still did not lead to consistent results, raising the sample concentration to 100 μM , we started to observe reproducible patterns (**Figure 3.4**).

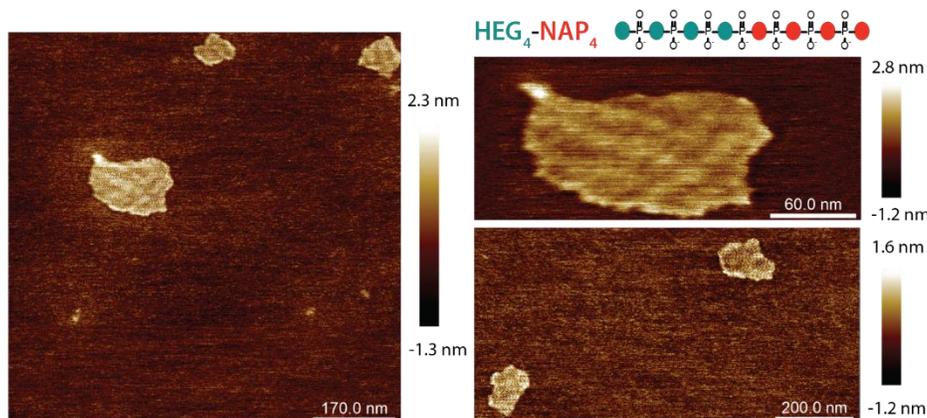


Figure 3.4. Dry AFM images of $\text{HEG}_4\text{-NAP}_4$ from 100 μM sample in TAMg.

Increasing the concentration of Mg^{2+} 5-fold (TA5Mg buffer) and using Nickel (II) to favour the interaction of the negatively charged phosphates with the mica⁴³ was the most successful strategy to image our nanostructures. Under such conditions, we were able to observe our structures consistently, provided that a salt layer was formed (**Figure 3.5**).

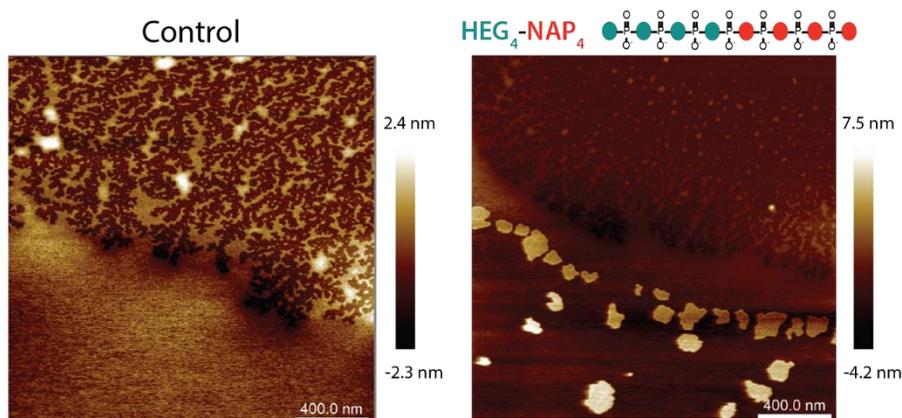


Figure 3.5. Dry AFM images of control and $\text{HEG}_4\text{-NAP}_4$ from 10 μM sample in TA5Mg+NiCl₂. We observed the formation of a salt layer as shown on the control. Nanostructures are deposited on the salt layers, most probably due to electrostatic interactions with Ni^{2+} and Mg^{2+} .

These conditions allowed us to conclude that **A4** (**HEG₄-NAP₄**) and **A7** (**HEG₆-NAP₆**), formed nanosheets of uniform height (**Figure 3.6**). The average height in dry images was difficult to evaluate, as magnesium and nickel salt deposits resulted in a non-uniform background layer on mica. Liquid AFM was also successfully performed on **A4** confirming self-assembly as nanosheets on mica. The average height of **A4** is 5.7 nm.

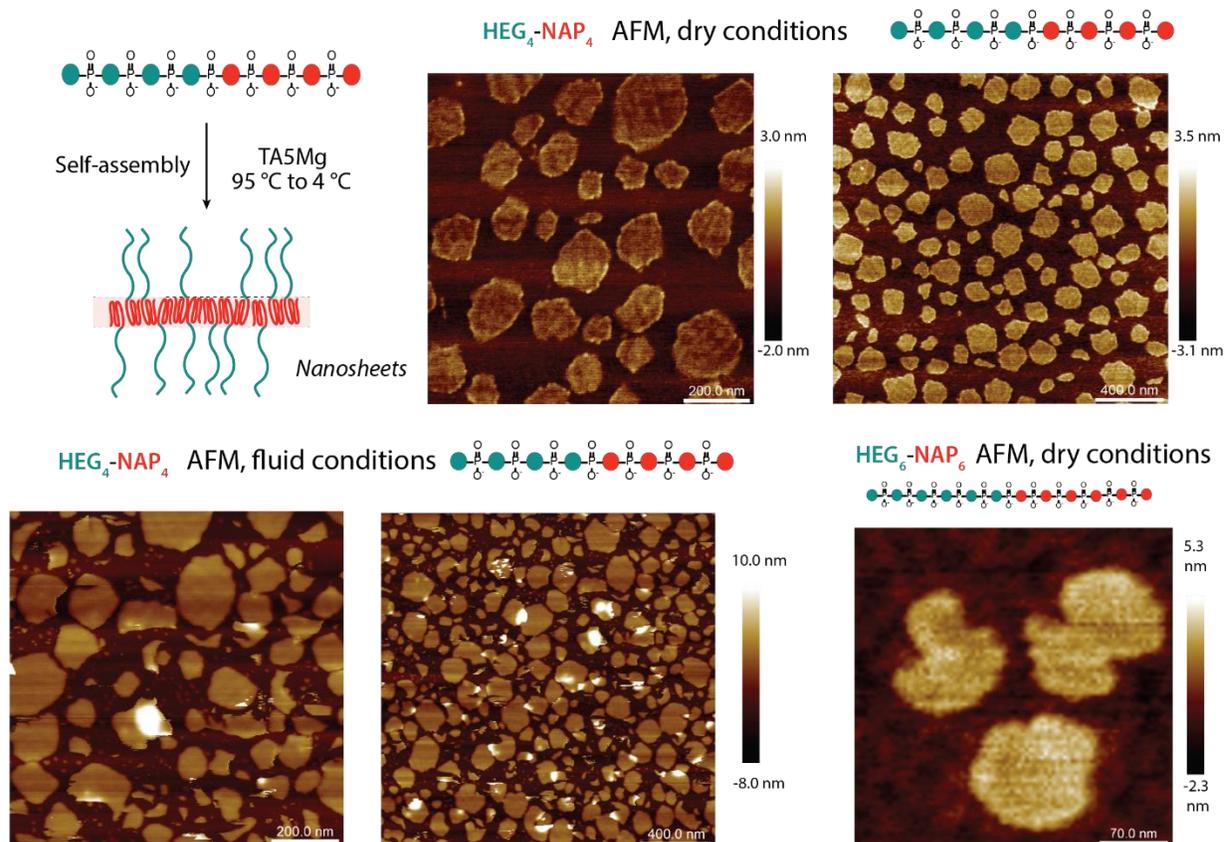


Figure 3.6. AFM images of **HEG₄-NAP₄** and **HEG₆-NAP₆** from 50 μ M sample in TA5Mg+NiCl₂.

Low voltage Transmission Electron Microscopy (TEM) was also performed with **A4** (**HEG₄-NAP₄**). Uniform nanosheets were observed further confirming the assembly observed by AFM (**Figure 3.7**). Sequence-defined peptoids and pyrene-based oligo(phosphodiester)s nanosheets have attracted significant interest due to potential applications in organic materials templated growth.^{44,45} However, examples of sequence-defined nanosheets are rare and usually involve only rigid monomers and short sequences. In our case, the possibility of making such sheets in the presence of the flexible **HEG** corona highlights the great stability of the sheets core.

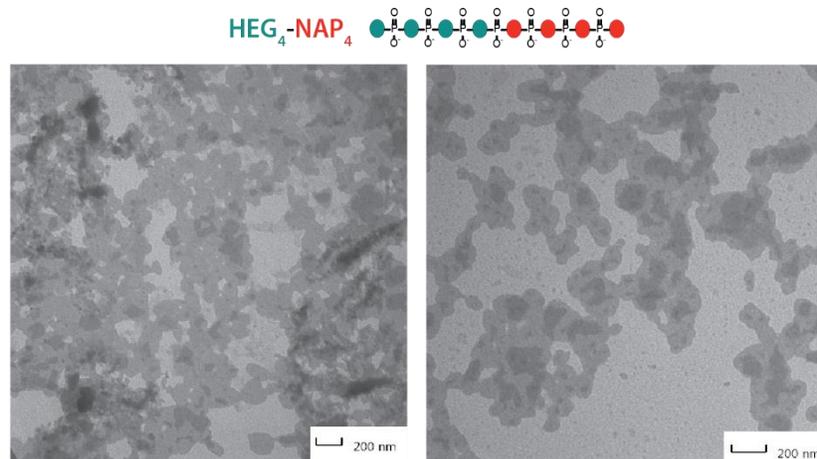


Figure 3.7. TEM images of **HEG₄-NAP₄** from 25 μM sample in TA5Mg

Under similar conditions, AFM showed the formation of spherical micelles with oligomers **A8** (**HEG₁₂-NAP₈**), **B2** (**HEG₄-NAP-C12₄**), **B3** (**HEG₆-NAP-C12₄**), **B5** (**HEG₆-NAP-PFC₄-C12₄**) and **B6** (**HEG₆-NAP-C12₄-PFC₄**) despite the smaller contact area between the mica and the nanostructures (**Figure 3.8** and Section 3.7.8).

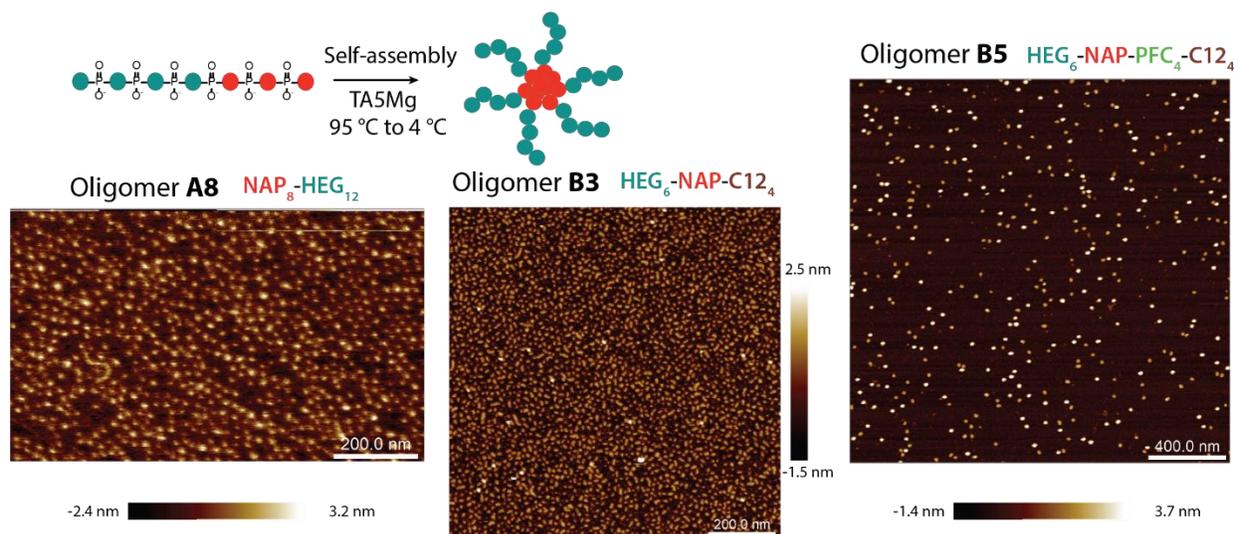


Figure 3.8. Representative AFM images of **A8**, **B3** and **B5** forming spherical micelles. Dry conditions for **A8** from a 10 μM sample in TA5Mg+NiCl₂, fluid conditions for **B3** and **B5** from a 50 μM and 1 μM respectively in TA5Mg +NiCl₂.

To study our oligomers in solution, we also performed Dynamic Light Scattering (DLS) measurements. The technique allowed us to estimate the diffusion coefficient and polydispersity index of our structures (**Table 3.2**). Most spherical micelle forming oligomers led to signals

confirming their self-assembly into discrete highly monodisperse nanostructures. Hydrodynamic diameters were calculated using a globular protein model (**Table 3.2**). This gives a good estimation of the size of the spherical micelles in solution (see section 3.5.5 for analysis). As expected, nanosheet forming structures had a greater polydispersity index and a smaller diffusion coefficient than the other structures. **A1 (HEG₄-NAP)** and **A2 (HEG₄-NAP₂)** had similar signal as the buffer control, showing that they do not self-assemble under the conditions explored. Over-scattering was often observed in the case of **A5 (HEG₄-NAP₆)**, **A9 (NAP₂-HEG₄-NAP₂)**, **B1 (HEG₄-C12₄-NAP)**, **B2 (HEG₄-NAP-C12₄)** and **B4 (HEG₄-NAP-PFC₄)**. We hypothesize their self-assembly may induce the formation of some large aggregates at the concentration (25 μM) used for DLS.

Table 3.2. Self-assembly characterization of the oligomers made.

Name	Sequence	Morphology in aqueous medium ^a	Diffusion coefficient (10 ⁻⁷ cm ² .s ⁻¹) ^b	Hydro-dynamic diameter ^c (nm)	Poly-Dispersity Index ^b (%)
A1	HEG₄-NAP	Unimers (D)	- ^d	-	-
A2	HEG₄-NAP₂	Unimers (D)	- ^d	-	-
A3	HEG₄-NAP₃	Spherical (A/D/G)	3.00 ± 0.04	16.0 ± 0.2	7 ± 2
A4	HEG₄-NAP₄	Nanosheets (A/D/G)	0.44 ± 0.02	109 ± 5 ^e	17 ± 5
A5	HEG₄-NAP₆	Not determined	- ^f	-	-
A6	HEG₆-NAP₃	Spherical (D/G)	3.10 ± 0.03	15.5 ± 0.2	12 ± 3
A7	HEG₆-NAP₆	Nanosheets (A/D/G)	1.0 ± 0.1	50 ± 5 ^e	38 ± 8
A8	HEG₁₂-NAP₈	Spherical (A/D)	1.14 ± 0.03	42 ± 1	12 ± 6
A9	NAP₂-HEG₄-NAP₂	Not determined	- ^f	-	-
B1	HEG₄-C12₄-NAP	Spherical (G)	- ^f	-	-
B2	HEG₄-NAP-C12₄	Spherical (G)	- ^f	-	-
B3	HEG₆-NAP-C12₄	Spherical (A/D/G)	3.7 ± 0.1	13.0 ± 0.5	7 ± 4
B4	HEG₄-NAP-PFC₄	Spherical (G)	- ^f	-	-
B5	HEG₆-NAP-PFC₄-C12₄	Spherical (A/D/G)	2.77 ± 0.08	17.4 ± 0.5	12 ± 6
B6	HEG₆-NAP-C12₄-PFC₄	Spherical (A/D/G)	2.16 ± 0.09	22 ± 1	15 ± 6

^a Methods of morphology determination is between parentheses. A=AFM, D=DLS, G=Gel electrophoresis.

^b Measured by DLS, 25 μM ^c Obtained from calculations from DLS data using a globular protein model

^d Results seem to show the presence of only very small particles. ^e Hydrodynamic diameters are reported for comparisons even if these structures are not necessarily spherical in solution. ^f Over-scattering was observed.

We performed a similar Agarose Gel Electrophoresis (AGE) study as in Chapter 2 to gather more information about our oligomers self-assembly. As explained before, agarose gels large pores size allows constructs such as self-assembled nanostructures to go through the gel. In the case of the library of oligomers that we describe here, diblock oligomers **A3** ($\text{HEG}_4\text{-NAP}_3$), **A6** ($\text{HEG}_6\text{-NAP}_3$) as well as **B1** ($\text{HEG}_4\text{-C12}_4\text{-NAP}$) to **B6** ($\text{HEG}_6\text{-NAP-C12}_4\text{-PFC}_4$) have similar mobility behavior as DNA-polymer micelles reported in Chapter 2 which is in accordance with the AFM and DLS results (**Figure 3.9**). Oligomers **A4** ($\text{HEG}_4\text{-NAP}_4$), **A5** ($\text{HEG}_4\text{-NAP}_6$) and **A7** ($\text{HEG}_6\text{-NAP}_6$) do not show discrete bands on these gels but smearing bands or non-penetrating material. This observation is consistent with the sheets observed by AFM and TEM for **A4** and **A7**. Hence, AGE is a valuable inexpensive method to qualitatively characterize the self-assembly behavior of various nanostructures at once. This is of great interest for relating the influence of degree of polymerisation, sequence and chemical composition with self-assembly.

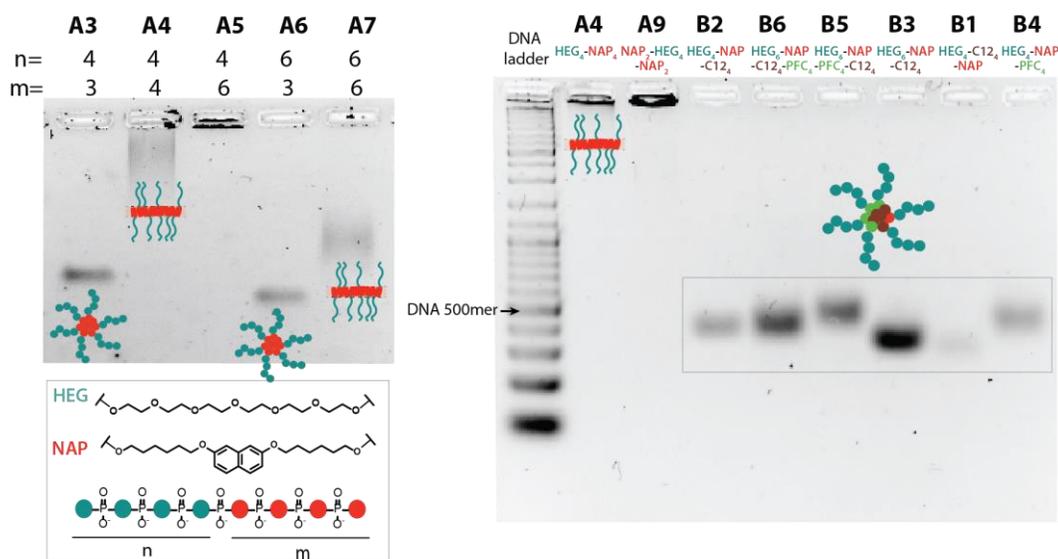


Figure 3.9. 2.5 % AGE in TAMg of some of the oligomers made. Oligomers that run with similar mobility than a DNA 500mer are thought to form spherical micelles. **A4** and **A6** were proven to form higher order assembly, leading to a different behavior on gel.

3.5.5 Influence of sequence on self-assembly

Taken together, our results can lead to multiple conclusions regarding the influence of parameters such as size, composition and supramolecular interactions involved in the self-assembly of sequence-defined oligomers. **A1** ($\text{HEG}_4\text{-NAP}$) to **A5** ($\text{HEG}_4\text{-NAP}_6$) were designed to gradually

increase the hydrophobic content and π - π stacking interactions while keeping the hydrophilic part constant. This series shows the potential of precision polymers, which allow the systematic analysis of very subtle differences of single monomer units between oligomers. We showed through Dynamic Light Scattering that **A1** (**HEG₄-NAP**) and **A2** (**HEG₄-NAP₂**) do not self-assemble (**Table 3.2**). This result was expected considering their low hydrophobic content. **A3** (**HEG₄-NAP₃**), with one additional **NAP** monomer, formed spherical micelles with a **NAP**-containing core and a **HEG** corona. **A2** and **A3** only differ by one **NAP** monomer. Interestingly, adding another **NAP** monomer (**A4**, (**HEG₄-NAP₄**)) radically changed the mode of assembly of the oligomer: nanosheets were observed on mica with AFM (**Figure 3.6**) and on carbon films with TEM (**Figure 3.7**). DLS and AGE also suggested the formation of higher-order structures (**Table 3.2**, **Figure 3.9**). This result is in accordance with the subtle **NAP/HEG** ratio increase from **A3** (3:4) to **A4** (4:4). This drastic change in assembly between two similar oligomers is striking, compared to conventional BCPs. In the case of sequence-defined oligomers, only one monomer can trigger self-assembly from spherical micelles to nanosheets.

A6 (**HEG₆-NAP₃**) was designed to have similar hydrophobic content as **A3** (**HEG₄-NAP₃**) but two more hydrophilic **HEG** units. We hypothesized that **A6** would form spherical micelles with a larger hydrophilic corona. However, DLS and AGE seem to show that **A6** micelles are comparable or even smaller in size than **A3**, suggesting that **A6** possibly assembles with a denser chain-packing than **A3** (**Table 3.2**, **Figure 3.9**). On the other hand, **A8** (**HEG₁₂-NAP₈**) was designed to have a similar **HEG** to **NAP** ratio as **A3** but a higher degree of polymerization. We observed that both oligomers form spherical micelles, but the diameter of **A8** is much larger than **A3** (**Table 3.2**). Thus, a small increase in hydrophilic content did not lead to a micelle diameter increase, but a higher degree of polymerization led to larger micelles.

In the case of nanosheets, **A4** (**HEG₄-NAP₄**) and **A7** (**HEG₆-NAP₆**) have the same **HEG:NAP** ratio, and only differ by their degree of polymerization. As expected, they formed similar lamellae structures (**Figure 3.6**). However, contrary to the case of **A8** (**HEG₁₂-NAP₈**) and **A3** (**HEG₄-NAP₃**), the higher DP nanostructures have a larger diffusion coefficient, implying that **A7** forms smaller structures than **A4** (**Table 3.2**).

The **B** series of oligomers was designed to gather information on the influence of different supramolecular interactions on self-assembly. We could change the type of interaction by changing

the nature of the monomers involved. For example, **B2** (HEG₄-NAP-C12₄) and **B4** (HEG₄-NAP-PFC₄) differ by the type of monomers involved but have similar degrees of polymerization and block copolymer ratios. Using agarose gel electrophoresis, **B2** appears to have a different size than **B4**, indicating that using **PFC** leads to different micelle morphology than **C12** (Figure 3.9).

We designed **B5** (HEG₆-NAP-PFC₄-C12₄) and **B6** (HEG₆-NAP-C12₄-PFC₄) because lipophilic and fluorophilic triblocks copolymers can sometimes induce unusual morphologies as reported elsewhere.⁴⁶ In our case, AFM, DLS and AGE results showed the presence of spherical micelles. The significant diameter difference of **B5** (HEG₆-NAP-PFC₄-C12₄) (17.4 nm) and **B6** (HEG₆-NAP-C12₄-PFC₄) (21.6 nm) observed by DLS in solution (Table 3.2) highlights the crucial role of the oligomer sequence. On the contrary, by AGE, it seems that **B5** is actually bigger in size compared to **B6**. We hypothesize that AGE and DLS give indications about different parameters of the micelles. For example, the continuous current used for AGE may create slight micelle deformation during the run. Indeed, these two oligomers have similar degree of polymerisation and composition but the fluorophilic and lipophilic blocks positions are inverted.

Through the design and synthesis of oligomers **A** and **B**, we observed self-assembly trends depending on the size, blocks ratio, blocks position and monomer nature of each oligomer. Some parameters dictated very different behaviours. For example, oligomers **A3** (HEG₄-NAP₃) and **A4** (HEG₄-NAP₄) only differ by one monomer but have radically different modes of assembly. These conclusions give insight into the self-assembly of oligo(phosphodiester)s and can be used to rationally design nanostructures for specific applications.

3.6 Conclusions

In conclusion, we demonstrated the rapid synthesis of a new class of self-assembling sequence-defined block oligomers with rational variation of the degree of polymerisation – i.e., increasing the number of monomers one by one in a monodisperse chain, block ratio, monomer type and supramolecular interactions. We designed four monomers to introduce hydrophobic and fluorophilic effects as well as π - π stacking in the final oligomers. Purification was achieved using reverse-phase chromatography or simple electrophoresis techniques. Block co-oligomers obtained

were truly monodisperse. Self-assembly experiments revealed the formation of spherical micelles of different sizes as well as two-dimensional nanosheets. This diversity demonstrates how the variety of interactions introduced through several monomers is of great interest for self-assembly purposes. For a given oligomer, our observations were systematically compared to the ones obtained with oligomers of different length, sequence or chemical composition. They revealed how these parameters influenced the constructs size and modes of assembly. Perfect sequence- and length-control was found to be a determinant parameter in the self-assembly process. Indeed, a small modification in the sequence can have a drastic effect in the nanostructure output (size or shape).

Automated synthesis provides the possibility to rationally study new parameters in block copolymer assembly, such as addition of a single monomer to a chain, or inversion of polymer sequence. It can significantly broaden our understanding of sequence-defined oligomers and paves the way to the design of functional precision polymers.

3.7 Experimental section

3.7.1 Chemicals

All starting materials were obtained from commercial suppliers and used without further purification unless otherwise noted. Acetic acid, Boric acid, solvents were purchased from Fisher Scientific. Chloroform-*d*1 was purchased from Cambridge Isotope Laboratories. Importantly, it was stored on molecular sieves in order to keep it neutral. If used as sold, hydrolysis of phosphoramidite (fast) as well as DMT deprotection (slow) may be observed. GelRed™ nucleic acid stain was purchased from Biotium Inc. Concentrated ammonium hydroxide, ammonium persulfate, acrylamide/Bis-acrylamide (40 % 19:1 solution) and TEMED were obtained from Bioshop Canada Inc. and used as supplied. 1 μmol Universal 1000Å LCAA-CPG supports and standard reagents used for automated DNA and RNA synthesis were purchased through Bioautomation. N,N-diisopropylamino Cyanoethyl phosphonamidic-chloride (CEP-Cl) and DMT-hexaethoxy glycol (cat.# CLP-9765) phosphoramidites were purchased from Chemgenes. Sephadex G-25 (super fine, DNA grade) was purchased from Glen Research. All other reagents were obtained from Sigma-Aldrich. TEAA (triethylammonium acetate) buffer is composed of 50

mM TEA with pH adjusted to 8.0 using glacial acetic acid. TBE buffer is 90 mM Tris, 90 mM boric acid and 1.1 mM EDTA with a pH of 8.0. TAMg buffer is 40 mM Tris, 7.6 mM magnesium chloride and 1.4 mM acetic acid. TA5Mg buffer is similar than TAMg but with 38 mM magnesium chloride. Muscovite Ruby mica sheets (grade 2) were used as substrate for all AFM imaging studies. Scan-Asyst Fluid + AFM cantilevers were purchased from Bruker. Electron microscopy (EM) grids were purchased from Pacific Grid Tech.

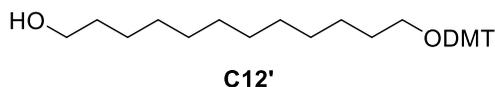
3.7.2 Instrumentation

Standard automated solid-phase synthesis was performed on a Mermade MM6 synthesizer from Bioautomation. HPLC purification was carried out on an Agilent Infinity 1260. DNA and oligomers quantification measurements were performed by UV absorbance with a NanoDrop Lite spectrophotometer from Thermo Scientific. For structure assembly, Eppendorf Mastercycler 96-well thermocycler and Bio-Rad T100™ thermal cycler were used to anneal all oligomers. Polyacrylamide Gel Electrophoresis (PAGE) experiments were carried out on a 20 X 20 cm vertical Hoefer 600 electrophoresis unit while agarose gel electrophoresis (AGE) experiments were performed with an Owl Mini gel electrophoresis unit. Gel images were captured using a ChemiDoc™ MP System from Bio-Rad Laboratories. Gel images were captured using a ChemiDoc™ MP System from Bio-Rad Laboratories. Dry solvents were taken from an Innovation Technology device. High Resolution mass determination was achieved using a Bruker Maxis API (Atmospheric pressure ionization) QTOF or a THERMO Exactive Plus Orbitrap-API. Liquid Chromatography Electrospray Ionization Mass Spectrometry. Liquid Chromatography Electrospray Ionization Mass Spectrometry (LC-ESI-MS) of oligomers was carried out using Dionex Ultimate 3000 coupled to a Bruker MaXis Impact™ QTOF. Dynamic Light Scattering (DLS) experiments were carried out using a DynaPro™ Instrument from Wyatt Technology. AFM was performed with a MultiMode™ MM8 SPM connected to a Nanoscope™ controller, from the Digital Instruments Veeco Metrology Group. TEM was performed with a LVEM5 from Delong America. The NMR spectra were recorded on Bruker 400 MHz, 500 MHz, Varian 300 MHz or 400 MHz for ^1H , ^{13}C and ^{31}P with chloroform-*d*1 (δ 7.26, ^1H ; δ 77.16, ^{13}C) or acetone-*d*6 (δ 2.04, ^1H ; δ 29.8, ^{13}C) as internal lock solvents and chemical shift standard.

3.7.3 Small molecule synthesis

Synthesis of PFC has been reported in Chapter 2.

3.7.3.1 12-(Bis(4-methoxyphenyl)(phenyl)methoxy)dodecan-1-ol, C12'.



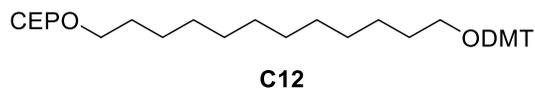
1,12-dodecanodiol (6 g, 29.65 mmol) was dissolved in anhydrous pyridine (30 mL) and DMTCI (10.04 g, 29.65 mmol) was added in small portions over 30 minutes. Reaction mixture was stirred for 2h at RT and pyridine was removed in vacuo, leaving dark orange dense oil. Crude material was purified by chromatography on triethylamine pre-treated silica gel with slow gradient of EtOAc/n-hexane with 1 % of NEt₃. Target material was collected at 15-20 % EtOAc/Hexane, solvent evaporated and product obtained as a pale yellow very viscous oil. (4.17 g, 31 %). Characterization data matched those reported elsewhere.⁴⁷

HRMS (ESI-QTOF) m/z: [M + Na]⁺ Calcd for C₃₃H₄₄O₄Na 527.3132; Found 527.3124.

¹H NMR (400 MHz, Acetone-*d*₆) δ 1.35-1.42 (m, 16H), 1.48-1.55 (m, 2H), 1.58-1.65 (m, 2H), 3.06 (t, *J* = 6.5Hz, 2H), 3.42 (br. s, 1H), 3.53 (t, *J* = 6.5Hz, 2H), 3.77 (s, 6H), 6.87 (d, *J* = 8.9 Hz, 4H), 7.18-7.22 (m, 1H), 7.25-7.37 (m, 6H), 7.45-7.48 (m, 2H).

¹³C NMR (100 MHz, Acetone-*d*₆) δ 26.7-33.8, 55.4, 62.5, 63.9, 86.4, 113.7, 127.3, 128.4, 128.9, 130.8, 137.4, 146.6, 159.4.

3.7.3.2 12-(Bis(4-methoxyphenyl)(phenyl)methoxy)dodecyl (2-cyanoethyl) diisopropylphosphoramidite, C12.



Monoprotected diol C12' was suspended in toluene and solvent was evaporated under reduced pressure (60 °C). The operation was repeated once and the dried compound was kept under high vacuum for at least 5 hours. In a 50 mL oven-dried round bottom flask, molecule C12' (191 mg, 0.38 mmol, 1 equiv.) is dissolved in 1.2 mL of THF. 5-(ethylthiotriazole) (ETT, 2.25 mL, 0.25 M in acetonitrile, 0.94 mmol, 1.5 equiv.) is added under vigorous stirring. 3-

((bis(diisopropylamino)phosphanyl)oxy)propanenitrile (210 μ l, 0.84 mmol, 2.2 equiv.) is added on the reaction medium. The reaction is left under stirring during 5 hours. Solvent is evaporated under reduced pressure. Molecule **C12** is purified by column chromatography. The mix of solvents used was hexanes/ethyl acetate/TEA (90:10:2) and was degassed before use. Compound C12 was obtained in a yield of 226 mg as a colorless oil. Yield: 85 %. Characterization data matched those reported elsewhere.⁴⁷

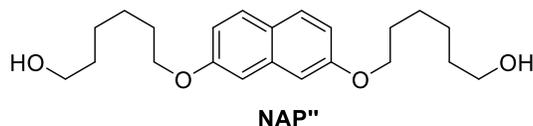
HRMS (ESI-QTOF) m/z : $[M + H]^+$ Calcd for $C_{42}H_{62}O_5N_2P$ 705.4391; Found 705.4382.

1H NMR (500 MHz, Acetone-*d*6) δ 7.47-7.44 (m, 2H), 7.34-7.28 (m, 6H), 7.23-7.19 (m, 1H), 6.89-6.86 (m, 4H), 3.89-3.80 (m, 2H), 3.78 (s, 6H), 3.70-3.59 (m, 4H), 3.05 (t, $J = 7$ Hz, 2H), 2.74 (t, $J = 6$ Hz, 2H), 1.64-1.58 (m, 4H), 1.41-1.36 (m, 4H), 1.35-1.28 (m, 12H), 1.19 (dd, $J = 4, 7$ Hz, 12H).

^{13}C NMR (126 MHz, Acetone-*d*6) δ 159.5, 146.7, 137.5, 130.9, 129.0, 128.5, 127.4, 113.8, 86.5, 64.4, 64.2, 63.9, 59.5, 59.3, 55.5, 43.8, 43.7, 32.1, 32.0, 30.8, 30.4, 30.3, 30.2, 27.1, 26.7, 24.9, 24.9, 20.8, 20.7.

^{31}P NMR (203 MHz, Acetone-*d*6): δ (ppm) 147.0.

3.7.3.3 6,6'-(Naphthalene-2,7-diylbis(oxy))bis(hexan-1-ol), NAP''.

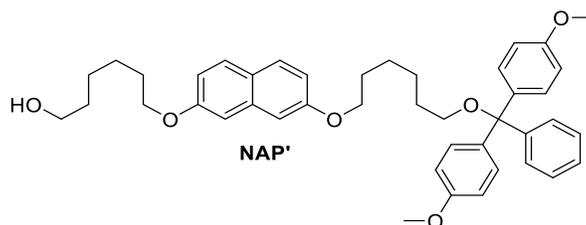


An oven dried round-bottom flask was charged with 2,7-dihydroxynaphthalene (2 g, 12.5 mmol), K_2CO_3 (5.18 g, 37.5 mmol) and Bu_4NI (461 mg, 1.25 mmol), atmosphere was exchanged to argon and anhydrous DMF (20 mL) was added followed by 6-bromohexanol (5.65 g, 31.25 mmol). Reaction mixture was stirred at RT for 48h until TLC analysis confirmed consumption of starting material, quenched with water (100 mL) and product extracted with CH_2Cl_2 (3 x 50 mL), dried over $MgSO_4$ and solvent evaporated under vacuum. Crude product was purified by chromatography with n-hexane/EtOAc eluent, collecting desired material at 40-60 % of EtOAc. Product was isolated as a white solid: 2.2 g, 64 %. Characterization data matched those reported elsewhere.⁴⁸

^1H NMR (400 MHz, CDCl_3) δ 1.23 (t, $J = 5.3$ Hz, 2H), 1.44-1.90 (m, 16H), 3.68 (dt, $J = 6.3, 5.4$ Hz, 4H), 4.07 (t, $J = 6.5$ Hz, 4H), 6.98 (dd, $J = 8.8, 2.4$ Hz, 2H), 7.03 (d, $J = 2.3$ Hz, 2H), 7.64 (d, $J = 8.8$ Hz, 2H).

^{13}C NMR (100 MHz, CDCl_3) δ 25.5, 25.9, 29.2, 32.6, 62.9, 67.8, 106.0, 116.2, 124.1, 129.0, 135.9, 157.6.

3.7.3.4 6-(7-(6-(Bis(4-methoxyphenyl)(phenyl)methoxy)hexyloxy)naphthalen-2-yloxy)hexan-1-ol, NAP'.



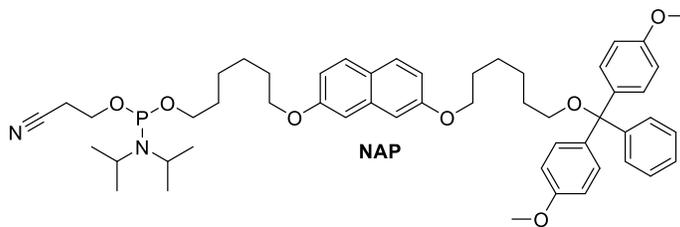
NAP'' (2.0 g, 5.55 mmol) was dissolved in anhydrous CH_2Cl_2 (50 mL). Freshly distilled triethylamine (2.5 mL) was added followed by DMTCI (2.1 g, 5.55 mmol) added portionwise over 30 minutes. Reaction mixture was stirred at RT for 16 hours. Solvents were removed under vacuum to form yellow dense oil which was purified by chromatography on triethylamine pre-treated silica gel with n-hexane/EtOAc with 1 % of NEt_3 . Desired material was collected at 40-50 % EtOAc/n-hexane as very viscous yellow oil: 1.1 g (30 %).

HRMS (ESI-QTOF) m/z : $[\text{M} + \text{K}]^+$ Calcd for $\text{C}_{43}\text{H}_{50}\text{O}_6\text{K}$ 701.3239; Found 701.3254.

^1H NMR (400 MHz, Acetone- d_6) δ 1.42-1.87 (m, 16H), 3.08 (t, $J = 6.4$ Hz, 2H), 3.42 (t, $J = 5.2$ Hz, 1H), 3.54-3.59 (m, 2H), 3.76 (s, 6H), 4.05-4.09 (m, 4H), 6.86 (d, $J = 8.9$ Hz, 4H), 6.96 (ddd, $J = 8.6, 6.0, 2.4$ Hz, 2H), 7.16-7.35 (m, 9H), 7.47 (d, $J = 7.5$ Hz, 2H), 7.68 (dd, $J = 8.9, 1.2$ Hz, 2H).

^{13}C NMR (100 MHz, Acetone- d_6) δ 26.5-33.7, 55.5, 62.4, 63.8, 68.4, 68.5, 86.5, 106.9, 113.8, 117.0, 125.1, 127.4, 128.5, 129.0, 129.8, 130.8, 137.3, 137.4, 146.6, 158.6, 159.5.

3.7.3.5 6-((7-((6-(Bis(4-methoxyphenyl)(phenyl)methoxy)hexyl)oxy)naphthalen-2-yl)oxy)hexyl (2-cyanoethyl) diisopropylphosphoramidite, NAP.



Monoprotected diol **NAP'** was suspended in toluene and solvent was evaporated under reduced pressure (60 °C). The operation was repeated once and the dried compound was kept under high vacuum for at least 5 hours. In the glovebox, 416 mg of **NAP'** (0.63 mmol) were then dissolved in anhydrous THF (1 mL) and 0.23 mL of ETT activator, 0.25 M in dry acetonitrile, (0.94 mmol, 1.5 equiv., 4.8 mL) was added to the reaction medium under stirring. 2-Cyanoethyl N,N,N',N'-tetraisopropyl phosphoramidite (0.50 mL, 1.57 mmol, 2.5 equiv.) was added slowly and the reaction was allowed to stir under inert gas at room temperature overnight. Solvent was evaporated under reduced pressure (40 °C). The crude product was loaded on a 25 g SiO₂ column and purified using a 90/10/1 Hexanes/Ethyl acetate/Triethylamine mixture in ~3 column volumes (CV). A clear transparent oil was isolated (424 mg, 78 %). NMR purity >95 %.

HRMS (ESI-QTOF) m/z: [M + H]⁺ Calcd for C₅₂H₆₈N₂O₇P 863.4759; Found 863.4745.

¹H NMR (500 MHz, CDCl₃): δ (ppm) 7.63 (dd, *J*=2, 9Hz, 2H), 7.45-7.43 (m, 2H), 7.33 (d, *J*=9Hz, 4H), 7.29-7.26 (m, 2H), 7.21-7.18 (m, 1H), 7.02 (s, 1H), 7.01 (s, 1H), 6.98-6.95 (m, 2H), 6.81 (d, *J*=9Hz, 4H), 4.04 (q, *J*=7Hz, 4H), 3.88-3.76 (m, 8H), 3.71-3.66 (m, 1H), 3.63-3.57 (m, 3H), 3.06 (t, *J*=7Hz, 2H), 2.62 (t, *J*=7Hz, 2H), 1.88-1.80 (m, 4H), 1.70-1.63 (m, 4H), 1.54-1.45 (m, 8H), 1.18 (dd, *J*=3, 7Hz, 12H).

¹³C NMR (126 MHz, CDCl₃): δ (ppm) 158.4, 157.8, 157.7, 145.6, 139.8, 136.9, 136.1, 130.2, 129.2, 128.3, 127.8, 126.7, 124.3, 117.8, 116.4, 116.3, 113.1, 106.1, 85.8, 68.0, 67.9, 63.8, 63.6, 63.4, 58.5, 58.3, 55.3, 43.2, 43.1, 31.3, 31.3, 30.2, 29.4, 29.4, 26.3, 26.2, 26.0, 25.9, 24.8, 24.8, 24.7, 20.5, 20.5.

³¹P NMR (203 MHz, CDCl₃): δ (ppm) 147.3.

3.7.4 Determination of absorption coefficient of 2,7-dihydroxynaphthalene and oligomers

2,7-dihydroxynaphthalene was dissolved in water at the concentrations 306, 312 and 409 μM . Absorbance at 260 nm of several dilutions of each sample was measured with a spectrophotometer.

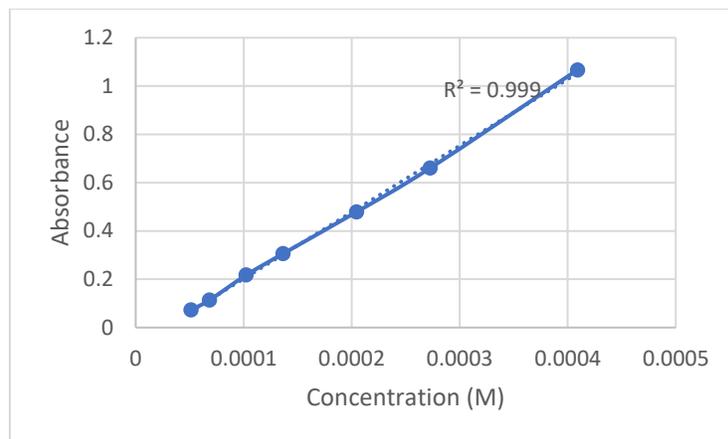


Figure 3.10. Absorbance as a function of 2,7-dihydroxynaphthalene concentration. One representative sample is shown.

Correlation coefficient (R^2) being close to .0.999 for the three samples, extinction coefficient could be calculated thanks to the Beer-Lambert law: $A = \epsilon * l * C$. A being the absorbance of the sample, ϵ the extinction coefficient of 2,7-dihydroxynaphthalene ($\text{M}^{-1} \cdot \text{cm}^{-1}$), l the width of the cuvette (cm) and C the concentration of the sample (M). Taking the average value for each sample, it was concluded that $\epsilon = 2.80 \text{ mM}^{-1} \cdot \text{cm}^{-1}$.

3.7.5 Solid-phase synthesis

Standard oligomer synthesis was performed on a 1 μmol scale, starting from a universal 1000 \AA LCAA-CPG solid support. Amidite **HEG** was dissolved in dry acetonitrile and all other amidites were dissolved in dry DCM to obtain 0.1 M solutions. Extended coupling times of 10 minutes were used. Other reagents (oxidizing solution, capping solutions, activator solution) are the same than for standard DNA synthesis. Removal of the DMT protecting group was carried out using 3 % dichloroacetic acid in dichloromethane on the DNA synthesizer.

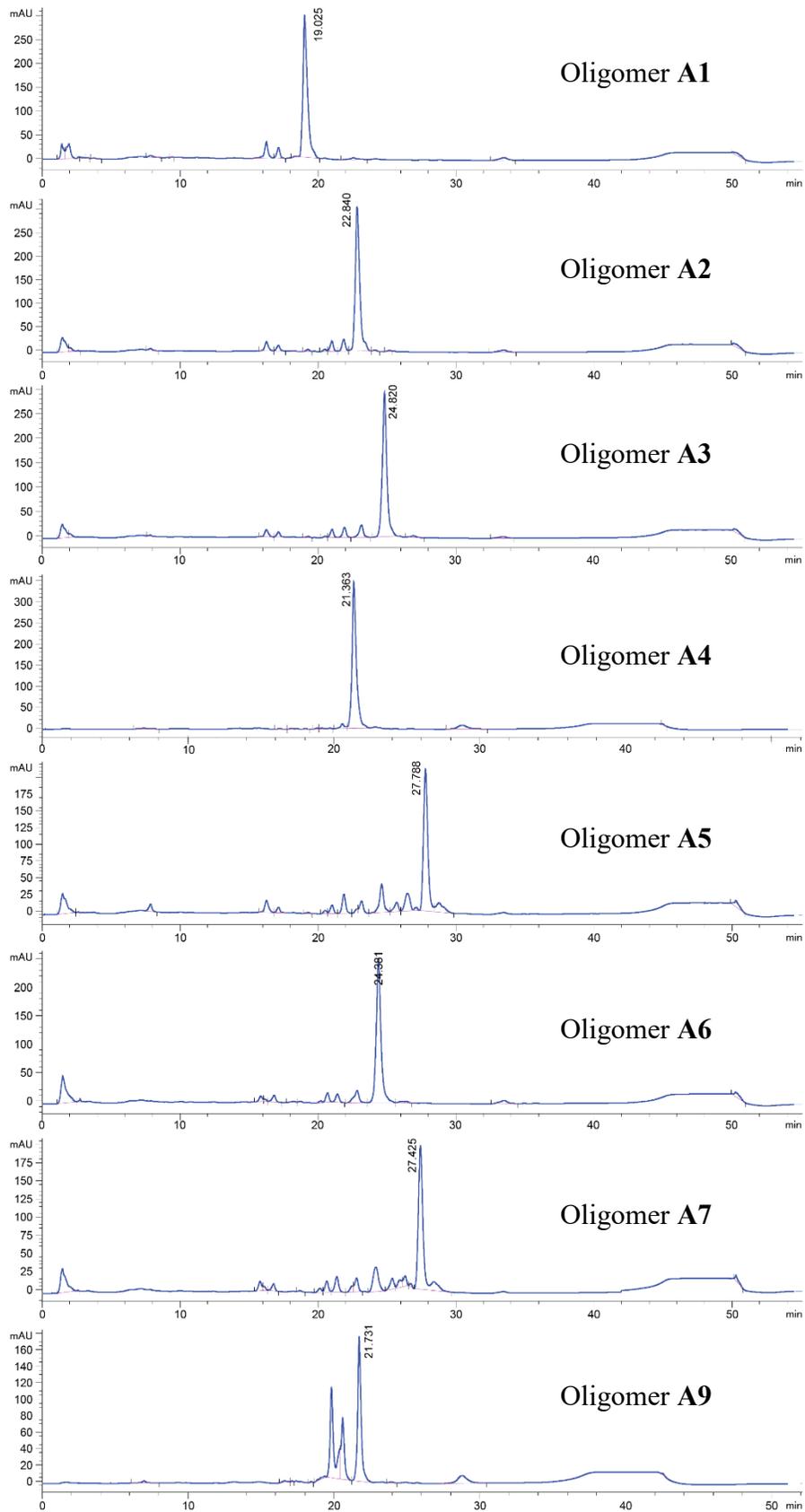
Completed syntheses were cleaved from the solid support and deprotected in a 1:1 mixture 28 % aqueous ammonium hydroxide solution/methylamine 40 wt.% in H_2O for 2.5 h at 60 $^\circ\text{C}$. The crude product solution was separated from the solid support and concentrated under reduced pressure at 60 $^\circ\text{C}$. This crude solid was re-suspended in 1 mL Millipore water. Filtration with 0.22 μm

centrifugal filter was performed prior to HPLC purification. The resulting solution was quantified by absorbance at 260 nm.

To quantify the oligomers in solution, absorbance at 260 nm was measured using a Nanodrop Lite spectrophotometer. Extinction coefficients were calculated according to the following formula:

$\epsilon = \epsilon_{\text{NAP}} \times \text{number of NAP in the sequence}$ with $\epsilon_{\text{NAP}} = 2.80 \text{ mM}^{-1} \cdot \text{cm}^{-1}$. This coefficient has been measured through standard methods detailed in Section 3.7.4.

For HPLC purification, solvents were (0.22 μm filtered): 50 mM triethylammonium acetate (TEAA) buffer (pH 8.0) and HPLC grade acetonitrile. Column: Hamilton PRP-C18 5 μm 100 Å 2.1 x 150 mm was used at 60 °C. For each analytical separation approximately 0.5 OD₂₆₀ of crude oligomer was injected as a 20-50 μL solution in Millipore water. Detection was carried out using a diode-array detector, monitoring absorbance at 260 nm.



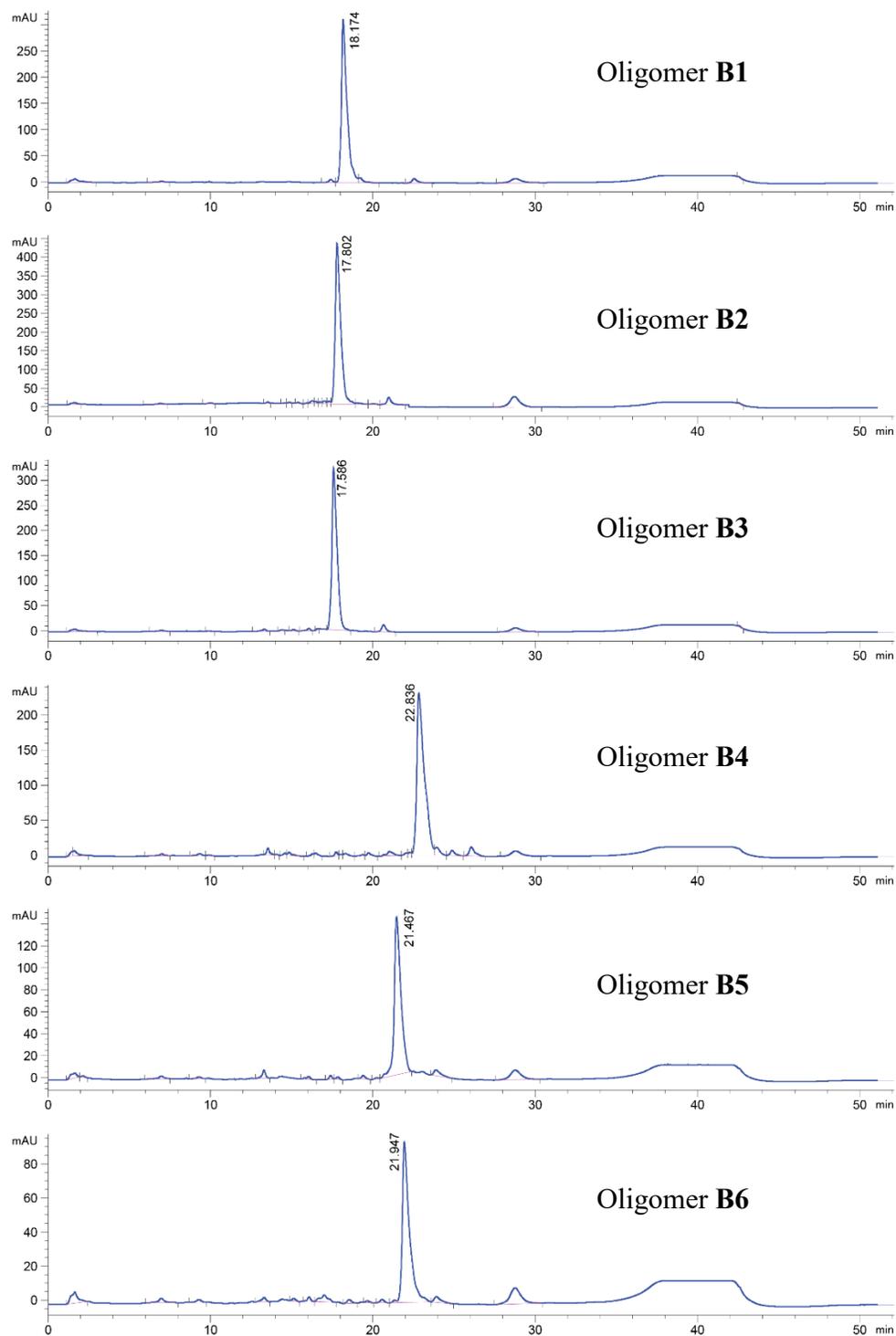


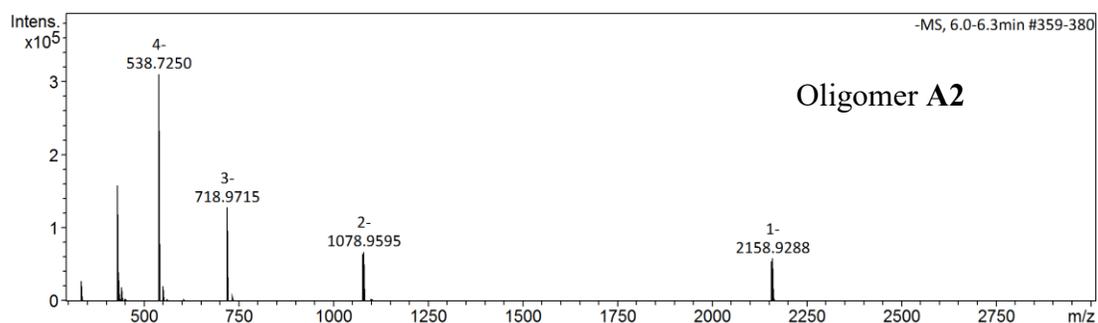
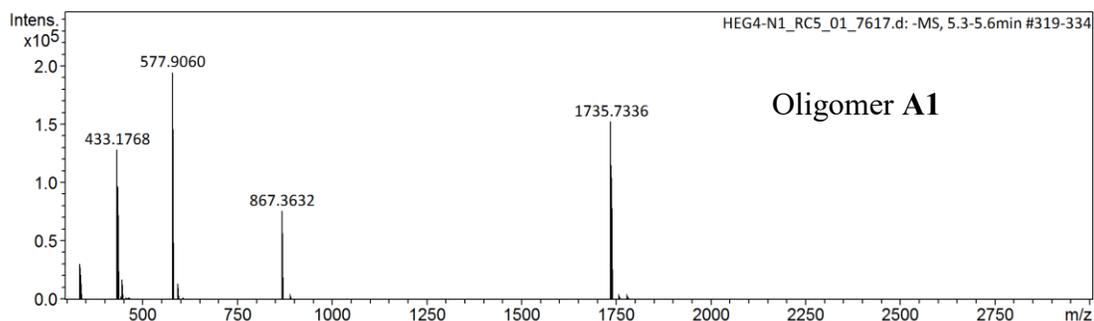
Figure 3.11. Reverse-phase HPLC traces (UV detection, 260 nm) from oligomers. Crude mixtures were injected, gradient used is from 3 to 80 % ACN in 40 minutes for oligomers A1 to A7 except A4. Other oligomers were purified using a gradient from 3 to 80 % ACN in 30 minutes.

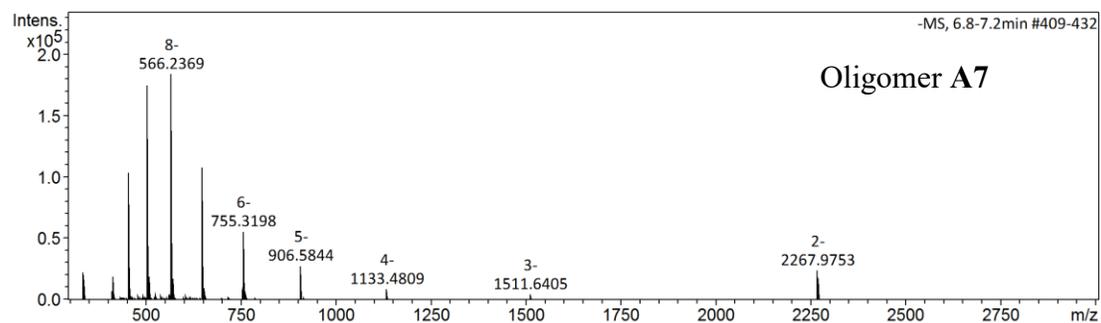
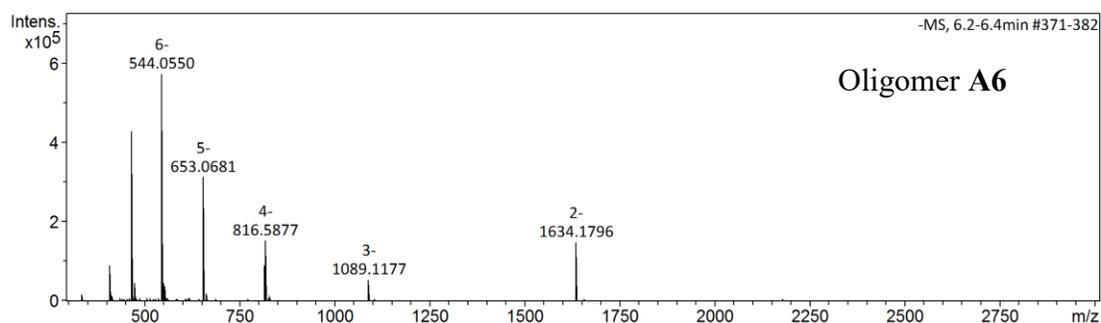
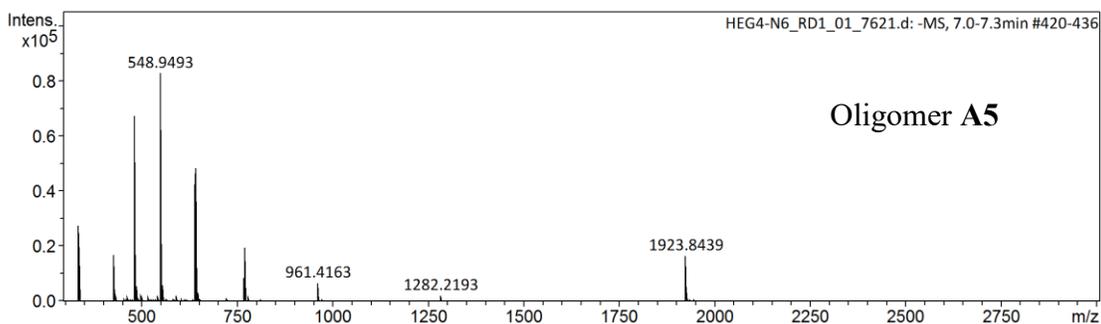
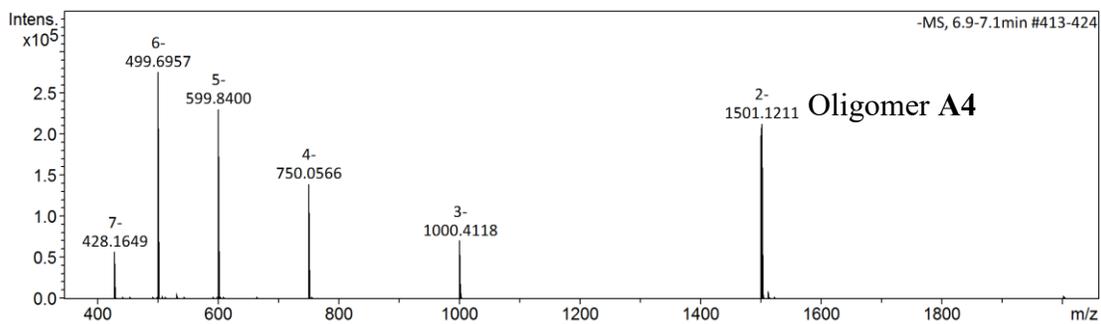
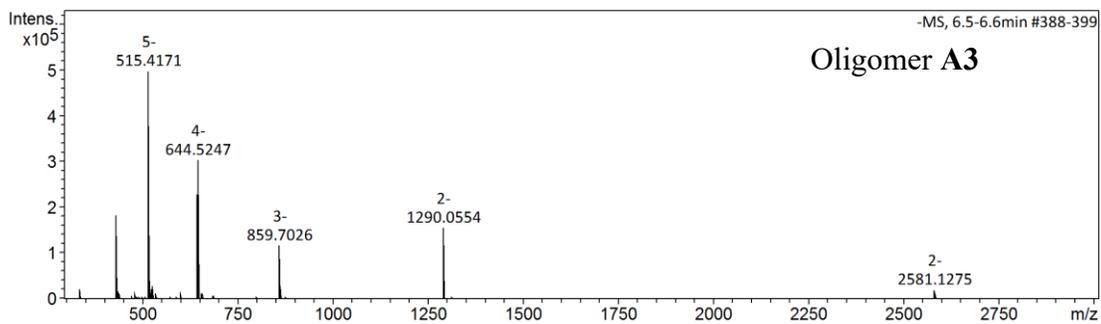
3.7.6 Gel electrophoresis purification

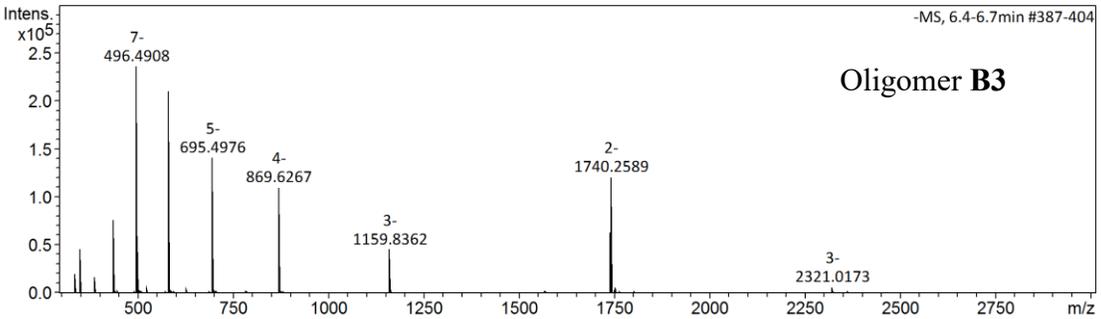
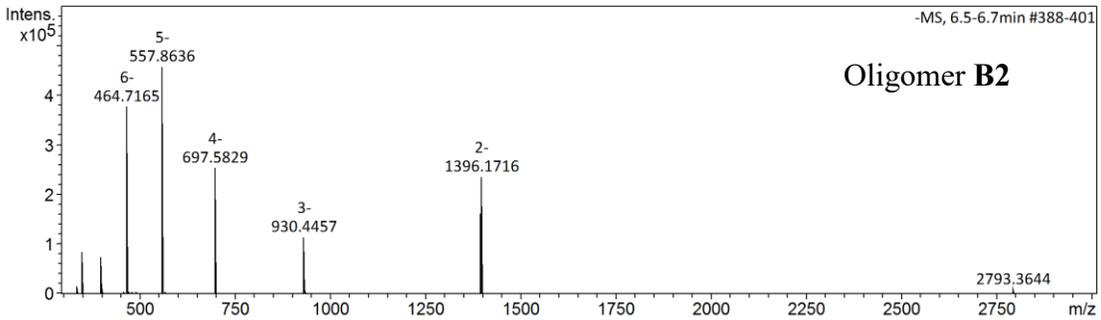
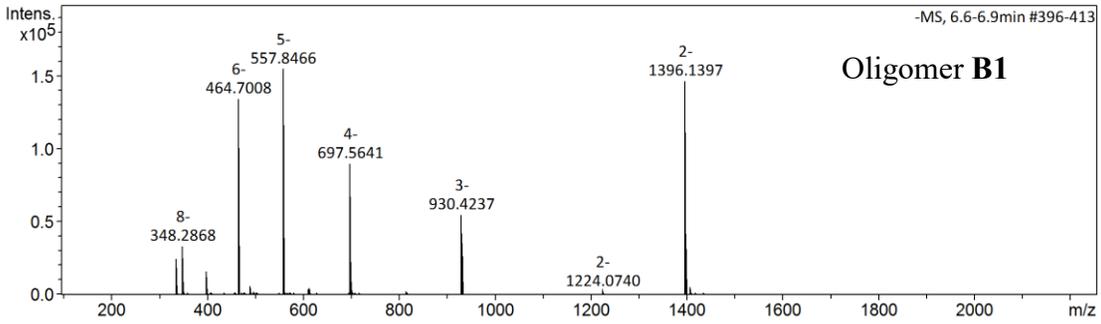
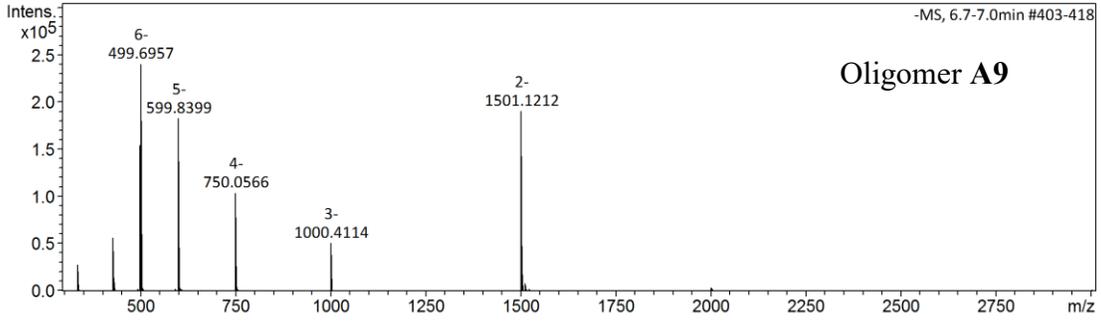
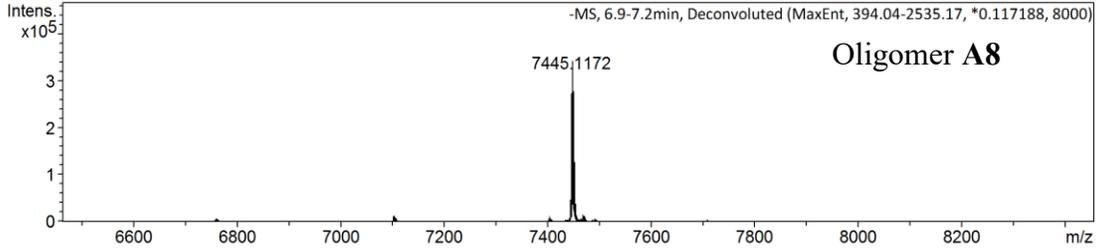
For **NAP₈-HEG₁₂**, purification could not be carried out using RP-HPLC, **HEG** being the last monomer added. In that case as well as in the cases of **HEG₆-NAP₃** and **HEG₆-NAP₆**, crude products were purified on 20 % polyacrylamide gels, supplemented with 8M urea (loading of 0.5 to 2 OD of crude oligomer per gel). Electrophoresis was run at lower voltage (250V) for the first 30 minutes followed by 1h at 500V. Following electrophoresis, the gel was wrapped in plastic and visualized by UV shadowing over a fluorescent TLC plate. The main band was quickly excised and the “crush and soak” procedure was applied: It was crushed, and incubated in ~10 mL of autoclaved water, chilled in liquid N₂ for a few minutes and left at 65 °C overnight. The supernatant was then concentrated to 1.0 mL and desalted through size exclusion chromatography (Glen Gel-Pak™ 2.5 Desalting Column from GlenResearch).

3.7.7 LC-ESI-MS characterization

The oligomers were analyzed by LC-ESI-MS in negative ESI mode. Samples (10 to 25 μM, 12 μl) were run through an Acclaim RSLC 120 C18 column (2.2 μm, 120Å 2.1 x 50 mm) using a gradient of mobile phase A (100 mM 1,1,1,3,3,3-hexafluoro-2-propanol and 5 mM triethylamine in water) and mobile phase B (Methanol) in 8 minutes (2 % to 100 % B). ~250 pmols of oligomers were injected.







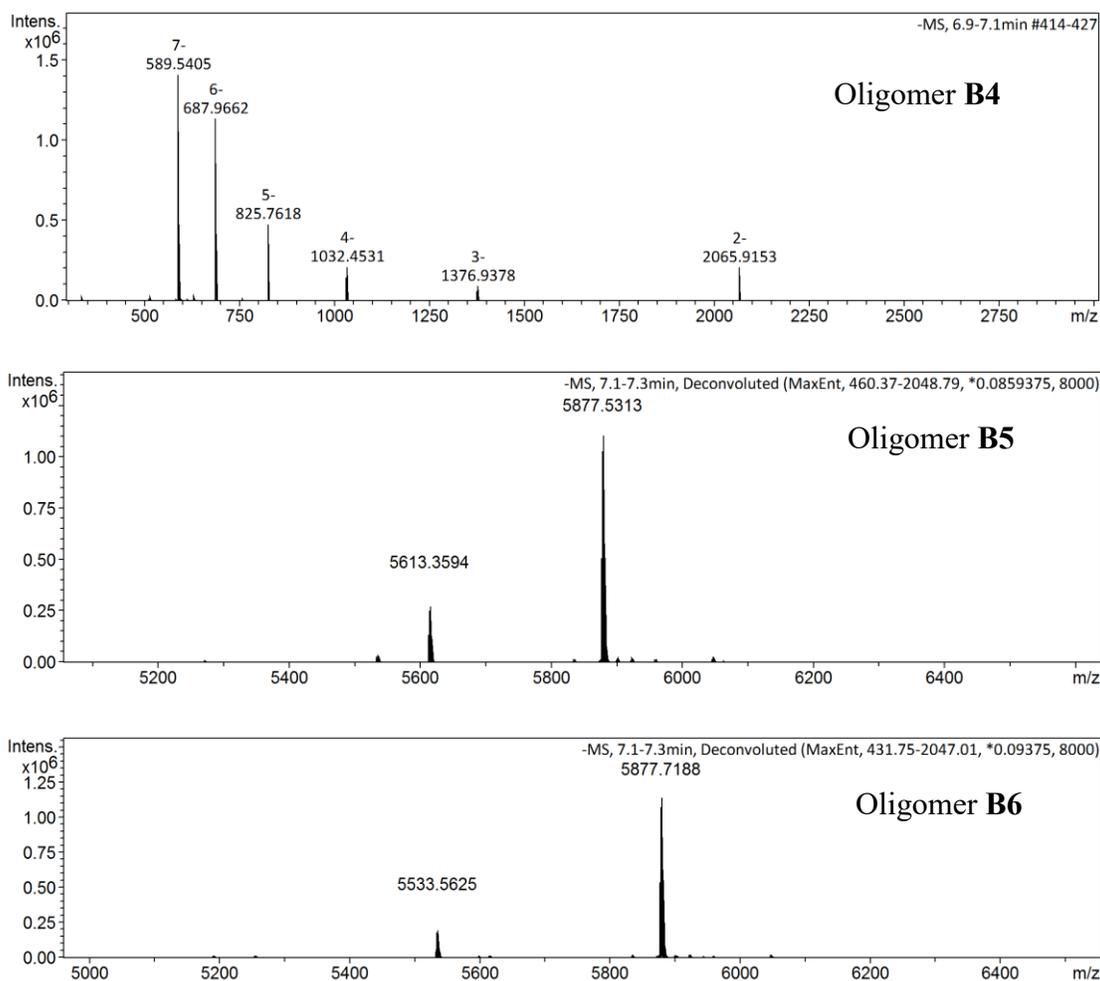


Figure 3.12. MS data for sequence-controlled oligomers, negative mode. Almost all peaks can be associated with a (M-x)/x anion. For oligomers **A8**, **B5**, **B6**, the data was processed and deconvoluted using the Bruker DataAnalysis software version 4.1. Masses reported are exact masses.

3.7.8 Atomic force microscopy

3.7.8.1 Dry conditions

Unless otherwise noticed, samples were annealed in TA5Mg from 95 °C to 4 °C at 250 μM during 1h. Samples were diluted to the concentration wanted right before deposition. NiCl₂ was added right before deposition on mica to reach 1 μM NiCl₂. 4 μL of this solution was deposited on a freshly cleaved mica surface (ca. 7 x 7 mm) and allowed to adsorb for 2-3 minutes. Samples were not washed and were dried under a flow of argon followed by vacuum for 3-4 hours prior to

imaging. Cantilevers used were Scan-Asyst Fluid +. Imaging was performed in ScanAsyst mode. Images were captured at scan rates between 0.5 and 1.5 Hz at a resolution of 512 x 512 pixels or 256 x 256 pixels for a few images. The gain was set at about 20. The peak force setpoint, z-limit and peak force amplitude were set automatically by the software. Images were processed using Nanoscope Analysis 1.5 software (Bruker).

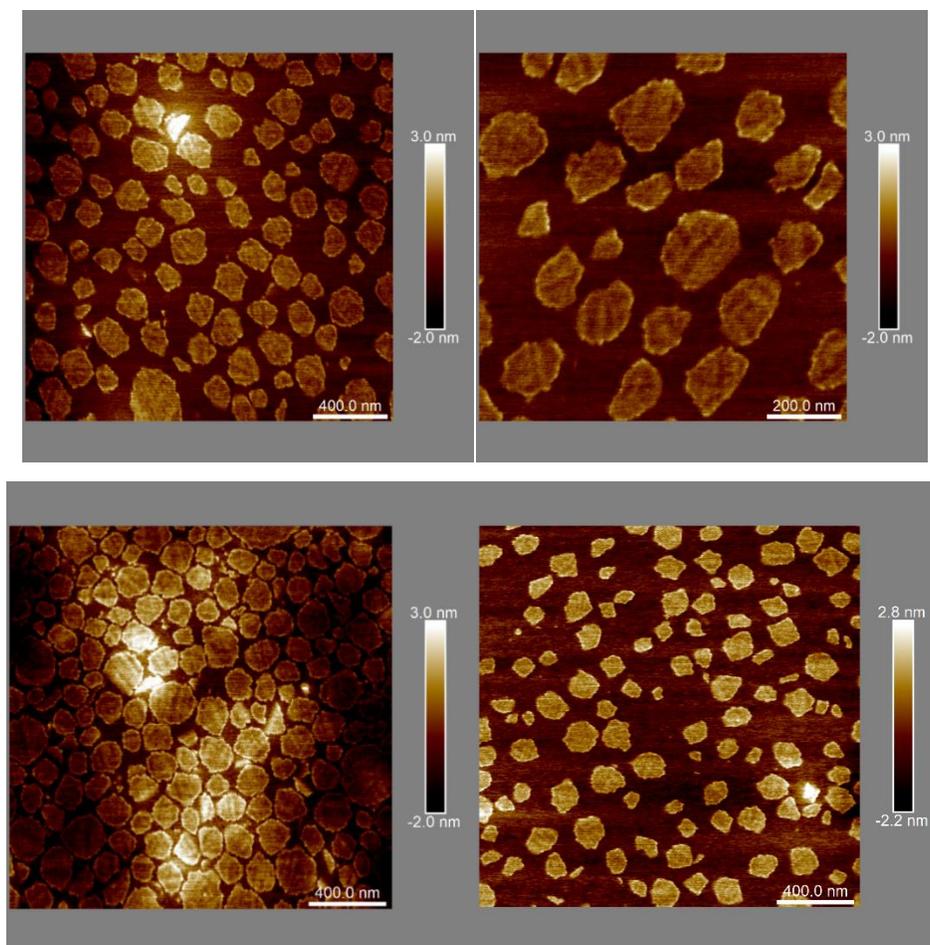


Figure 3.13. Supplementary dry AFM images of oligomer A4, 50μM.

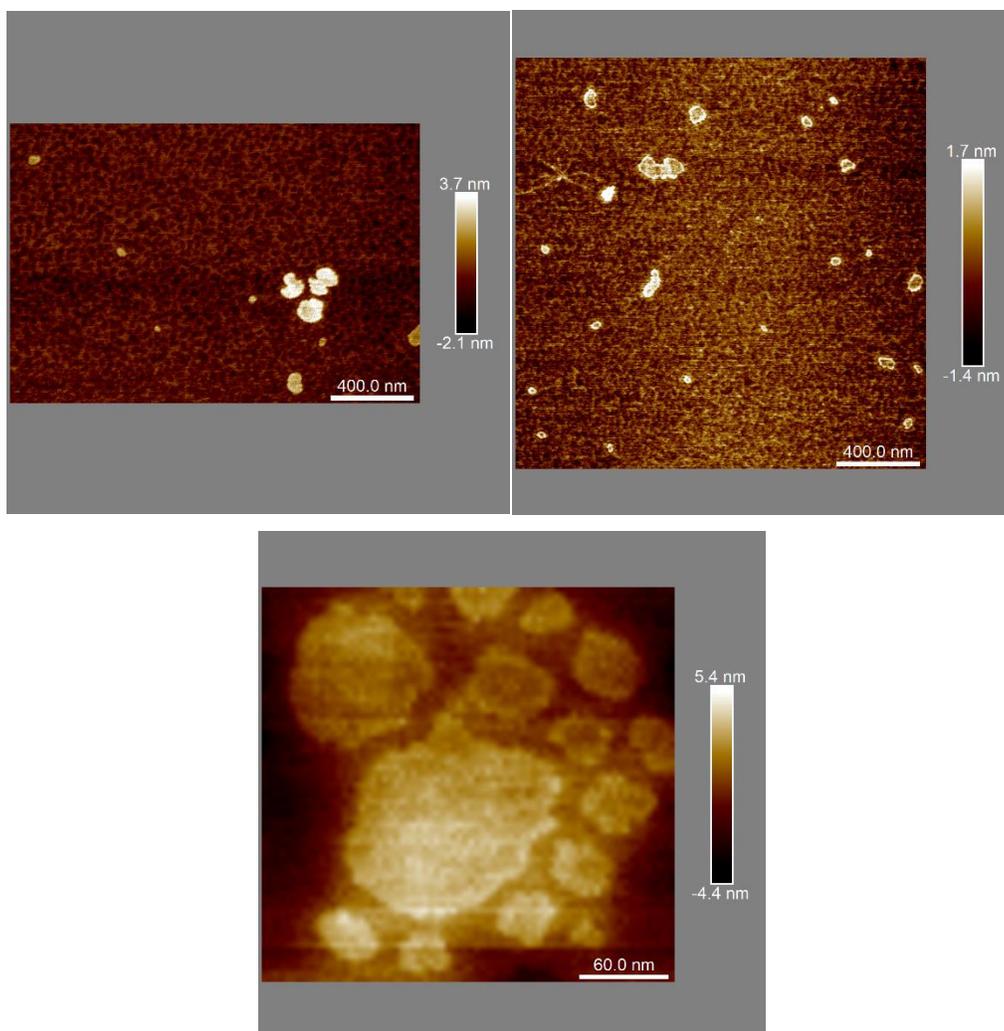


Figure 3.14. Supplementary dry AFM images of oligomer A7, 50 μ M

Imaging of these structures was also tried on highly-ordered pyrolytic graphite. However, with or without washing steps and without nickel chloride, negative controls and samples were found to be too similar to be appropriately analyzed.

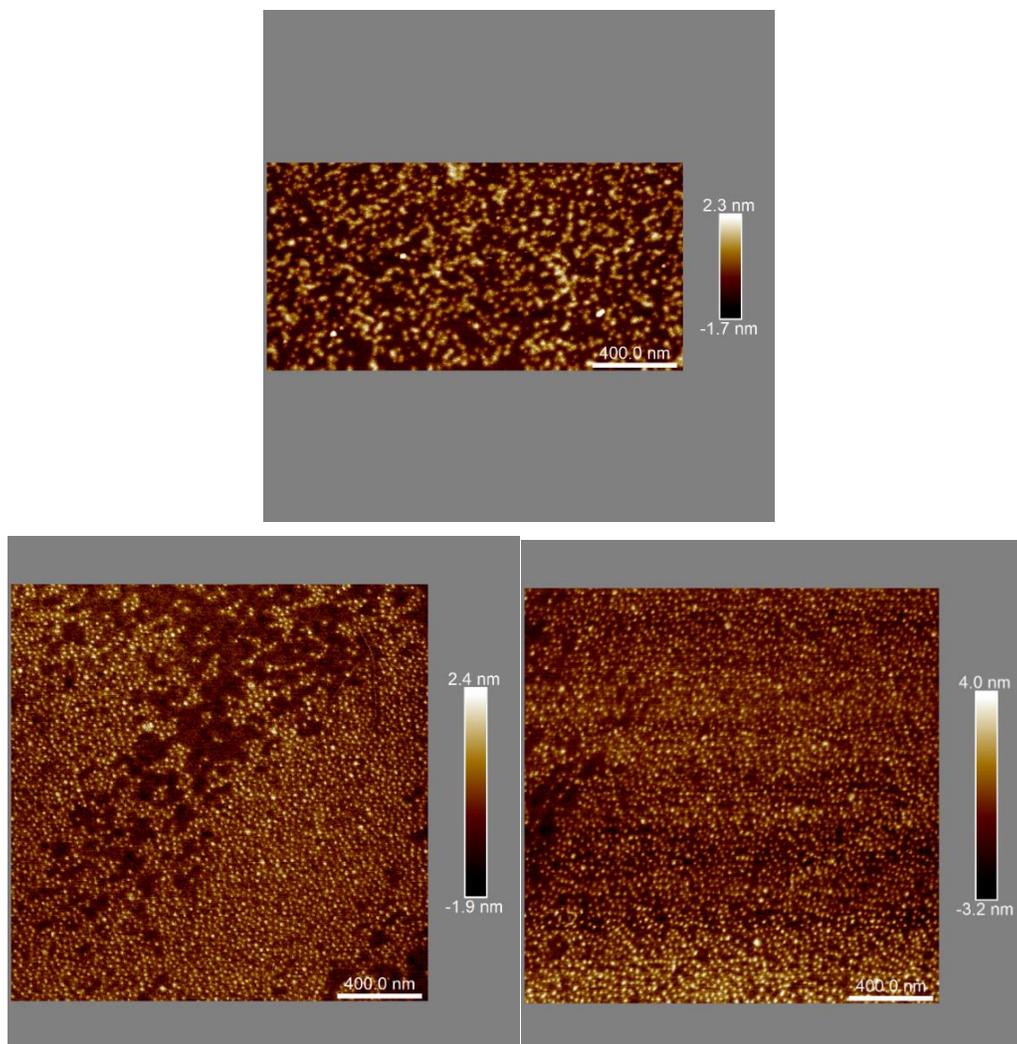


Figure 3.15. Dry AFM images of spherical micelles forming oligomers. Oligomer **A3** (top image), 50 μM and oligomer **A8** (bottom images), 10 μM . For these oligomers, mica was pretreated with a 20 mM NiCl_2 solution instead of adding NiCl_2 in the sample before deposition.

3.7.8.2 Fluid conditions

Prior to imaging, mica was treated with a 5 mM NiCl_2 solution: 100 μL were deposited on the freshly cleaved mica surface (ca. 12 x 12 mm) and left for 5 minutes. The solution was then dried using a nitrogen flow. Unless otherwise noticed, samples were annealed in TA5Mg from 95 $^\circ\text{C}$ to 4 $^\circ\text{C}$ at 250 μM during 1h. Right after, samples were diluted to the concentration wanted. 5 μL of this solution was deposited on the pretreated mica surface, allowed to adsorb for 2-3 minutes and 70 μL of TA5Mg was added. Excess was immediately removed using filter paper and dried under a nitrogen flow. After placing the sample on the AFM sample stage, fluid cell was positioned on

the sample (MTF ML fluid cell, Bruker) and 40 μL of 1xTA5Mg was injected into the cell. Cantilevers used were Scan-Asyst Fluid +. Imaging was performed in ScanAsyst mode. Images were captured at scan rates between 0.5 and 1.5 Hz at a resolution of 512 x 512 pixels or 256x256 pixels for a few images. The gain was set at about 20 and the z-limit at 1 μm . The peak force setpoint and peak force amplitude were set automatically by the software. Images were processed using Nanoscope Analysis 1.5 software (Bruker).

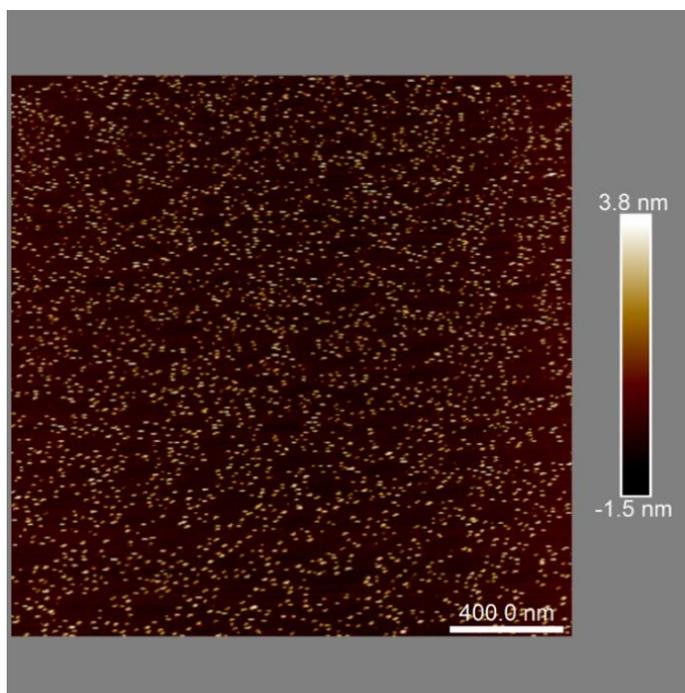


Figure 3.16. Liquid AFM image of oligomer **A8**, 10 μM .

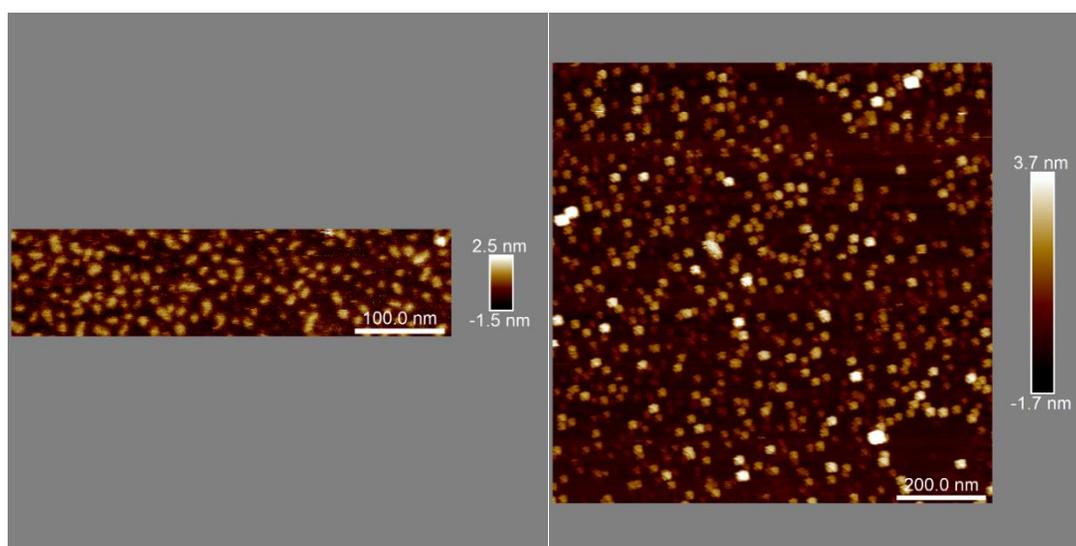


Figure 3.17. Supplementary liquid AFM images of oligomer **B3**, 50 μM .

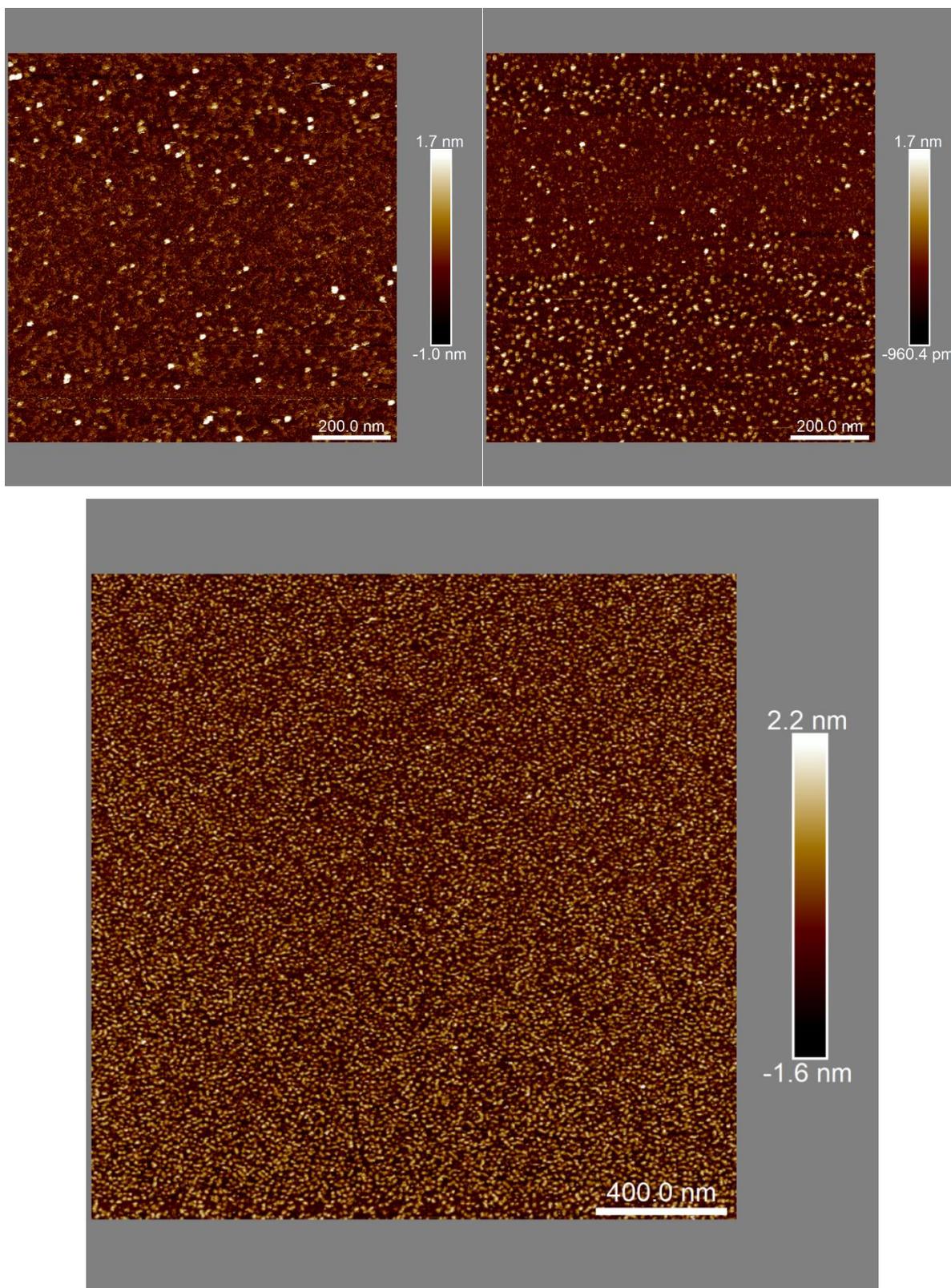


Figure 3.18. Liquid AFM images of oligomer **B2**. 1 μM (top) and 10 μM (bottom).

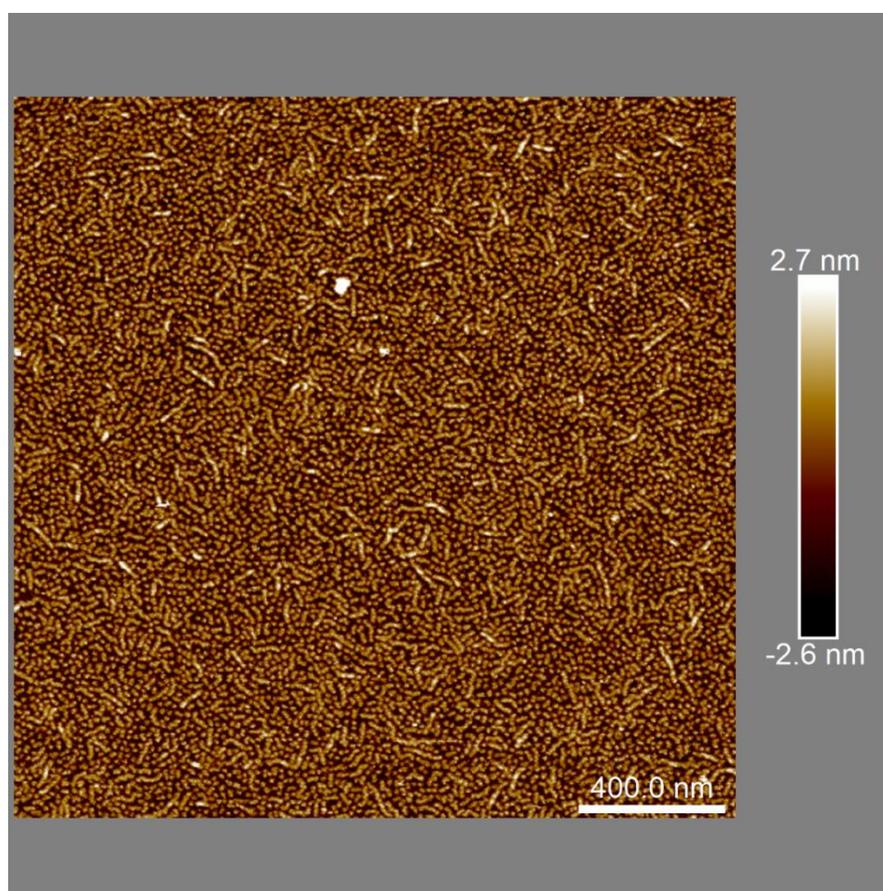
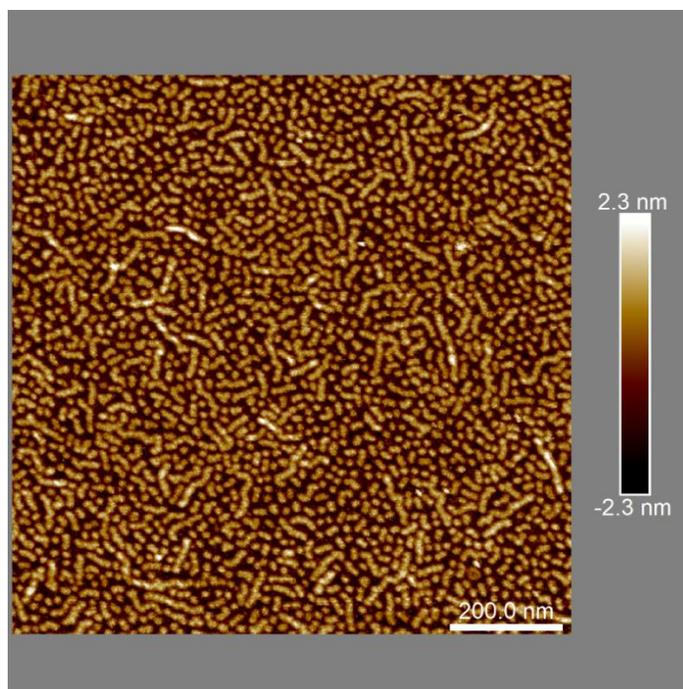


Figure 3.19. Liquid AFM images of oligomer **B6**, 10 μM . It seems that the micelles fuse during deposition on mica.

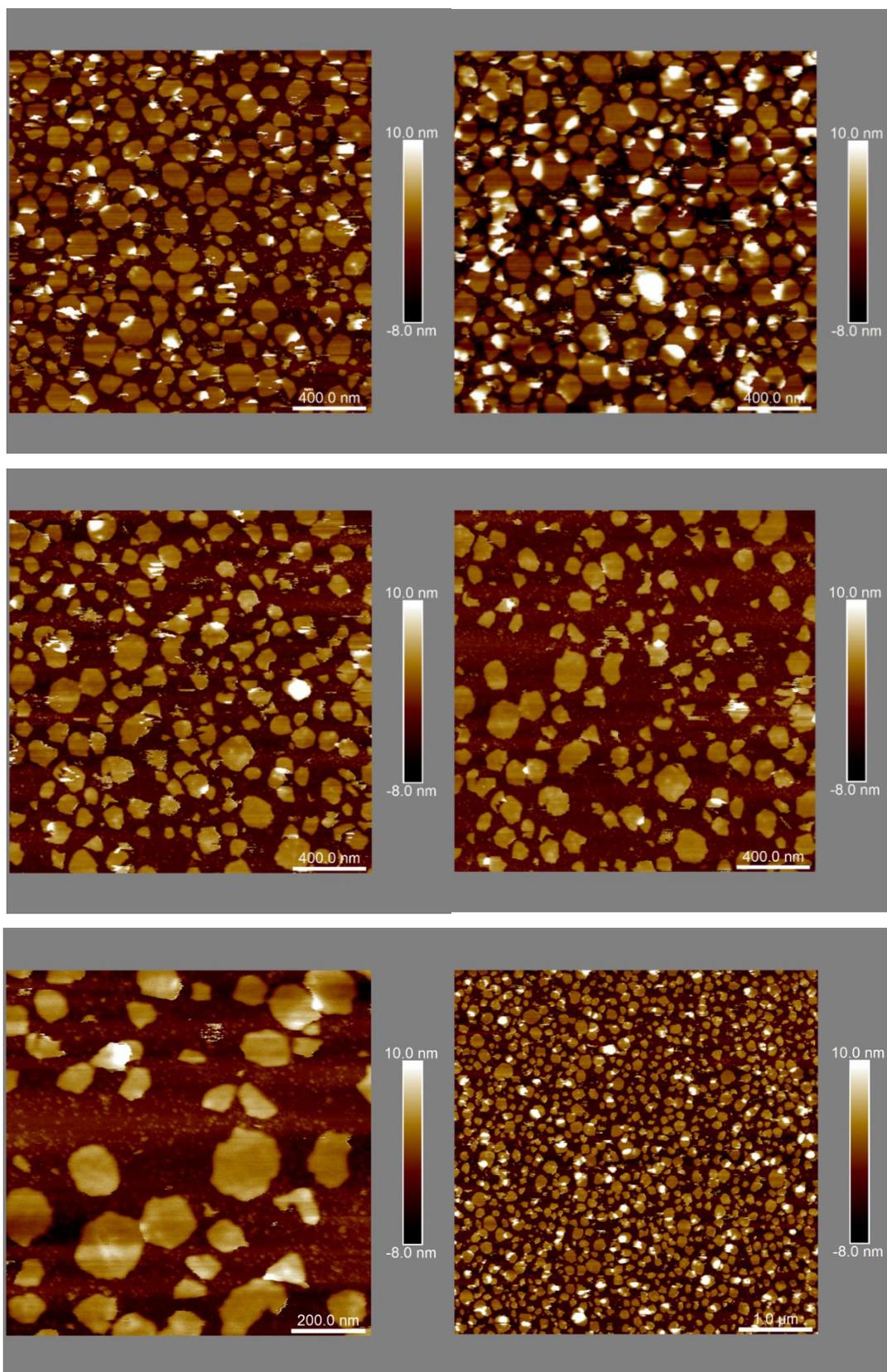


Figure 3.20. Supplementary liquid AFM images of oligomer A4, 50 μM.

3.7.9 Transmission Electron Microscopy

First trials were done using a high-voltage TEM but structures revealed to be hard to image, noticeably due to the fact that the HEG moieties are known to “burn” under a high energy electron beam. For TEM, the high amount of salts was less problematic because most samples were washed with water before imaging. However, images were taken of negative controls (without nanostructure).

Unless otherwise noticed, samples were annealed in TA5Mg from 95 °C to 4 °C at 250 μM during 1h. Right after, samples were diluted to the concentration wanted (25μM). 4μL of this solution was deposited on a TEM carbon film coated copper EM grid (300 mesh Cu grid with holey/dbl carbon films), allowed to adsorb for 2-3minutes and washed twice with MilliQ water to remove salt excess. The grids were then kept under vacuum for at least 3 hours.

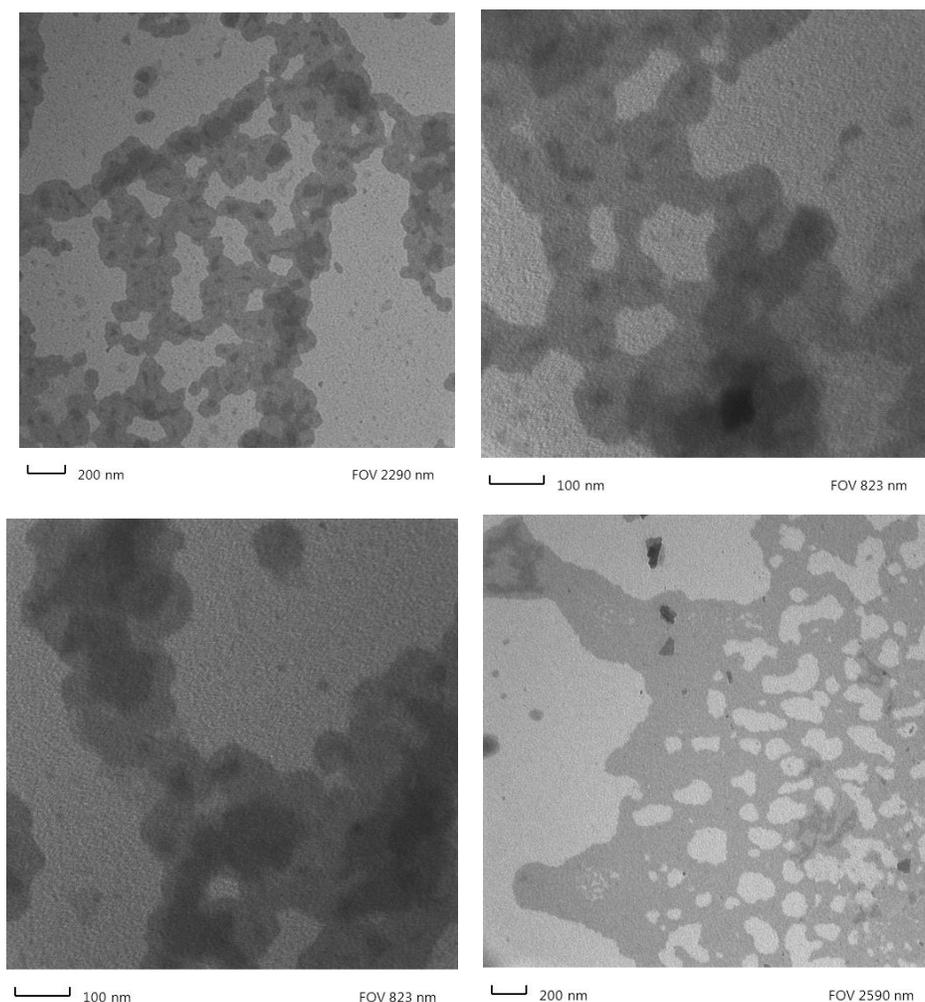


Figure 3.21. Supplementary TEM images of oligomer A4. No wash was performed for the bottom right image.

3.7.10 Dynamic Light Scattering

Dynamics V6 was used for data collection and analysis. A cumulants fit model was used to confirm/infirm the presence and determine the diffusion coefficients of a monomodal population of micellar aggregates. Hydrodynamic diameters reported in the main text were calculated with Dynamics V6 globular protein model. Sterile water and TA5Mg were filtered using a 0.2 μm nylon syringe filter before use for DLS sample preparation. All measurements were carried out at 20 $^{\circ}\text{C}$ after annealing of the constructs in TA5Mg from 95 $^{\circ}\text{C}$ to 4 $^{\circ}\text{C}$ at 250 μM during 1 h. Concentration of the sample is lowered to 25 μM after annealing. All the measurements were done at least in duplicates.

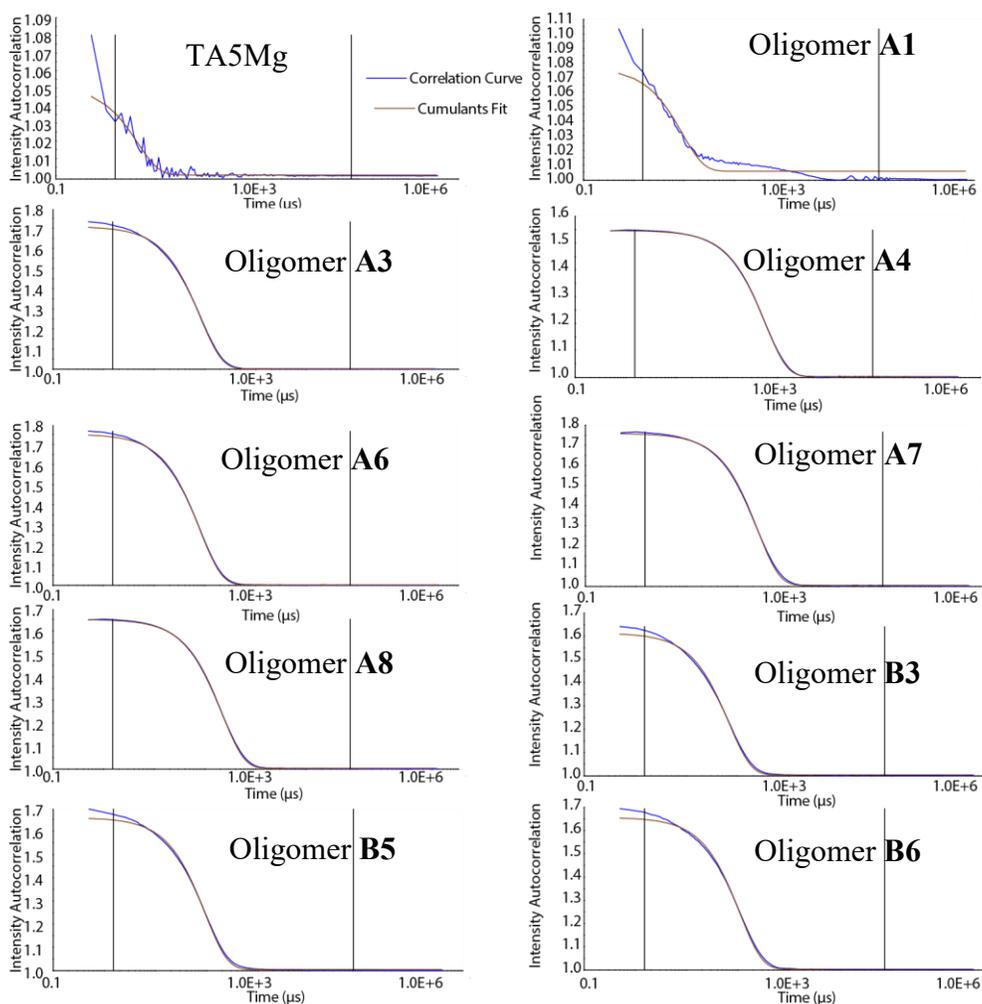


Figure 3.22. Representative DLS intensity correlation curves for 25 μM solutions of oligomers. Control sample and oligomer **A1** show poor fit because the particles are too small to be accurately analyzed. In contrast the data for self-assembling oligomers reveals excellent correlation even for oligomers **A4** and **A7** which do not assemble into spherical objects when deposited on a substrate.

Due to its perfluorocarbon core, we thought **B5** and **B6** could self-assemble without requiring the magnesium cations. However, DLS measurements seem to indicate that no self-assembly occurs after annealing in Milli-Q water. Buffering the same solution with a magnesium containing solution triggered instantaneous self-assembly of **B6**. Oligomers **B5** and **B6** were annealed with the same protocol in MilliQ water and analyzed similarly.

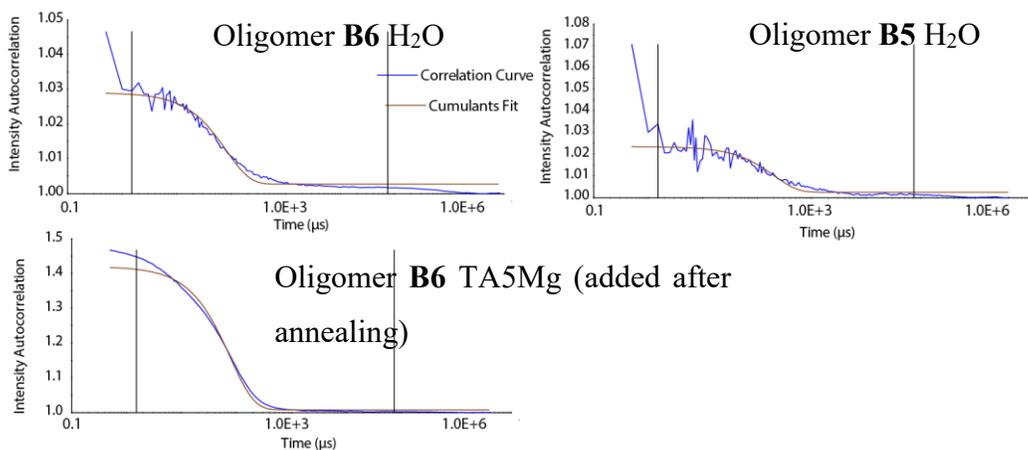


Figure 3.23. Dynamic light scattering on oligomers **B5** and **B6** in water after annealing. Representative DLS intensity correlation functions for 25 μ M solutions. In water, it looks like **B5** and **B6** are not self-assembling whereas addition of buffer to reach TA5Mg concentrations (bottom curve) triggers self-assembly. Fit is not as good as on Figure 3.22, probably because the sample has not been annealed in TA5Mg.

3.7.11 Gel electrophoresis analysis

20 % denaturing Polyacrylamide Gel Electrophoresis (PAGE) was carried out at room temperature for 30 minutes at 250V followed by 1 hour at 500V. TBE buffer (1X) was used and the concentration of urea in the gel was 7M. For each lane 5 μ L of sample in water was added to 5 μ L of 8M urea. Amount of the oligomer to load was done according to the principle that the more NAP the oligo contains, the more it will have affinity for GelRedTM. Therefore, respectively 0.8, 0.6, 0.4, 0.3, 0.6 and 0.3 nmols of **A2**, **A3**, **A4**, **A5**, **A6** and **A7** were loaded. The oligomers were visualized by incubation with GelRedTM for 10 minutes.

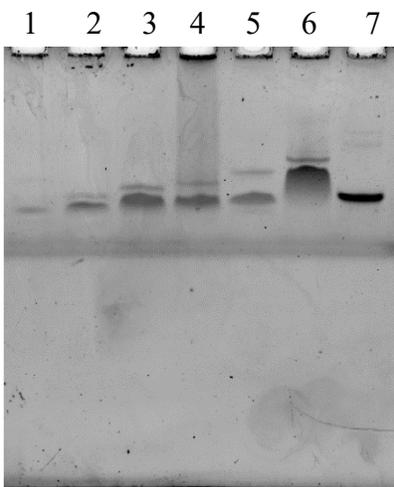


Figure 3.24. Denaturing gel (PAGE 20 %) with oligomers. Oligomers **A2**, **A3**, **A4**, **A5**, **A6**, **A7** and a DNA 22mer in respectively lanes 1 to 7.

2.5 % agarose gel electrophoresis (AGE) was carried out in TAMg at 4 °C for 2 h at 80 V. Gel was cast in TAMg and the samples were annealed with the same method detailed in section SI-VI. Respectively 1.1, 0.9, 0.7, 0.5, 0.3, 0.6, 0.3 and 0.6 nmols of **A1**, **A2**, **A3**, **A4**, **A5**, **A6**, **A7** and **A9** in 10 μ l of TA5Mg. and 2 nmols of B oligomers were loaded. 2 μ l of glycerol were added to the samples before loading. The bands for all gels were visualized by incubation with GelRed™. We hypothesize that self-assembled nanostructures are more likely to recruit GelRed® molecules than free oligomers leading to better sensitivity.

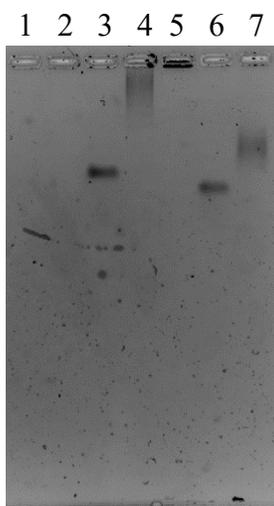


Figure 3.25. Agarose gel (2.5 % AGE) with oligomers **A1** to **A7**. Oligomers **A1**, **A2**, **A3**, **A4**, **A5**, **A6**, **A7** in respectively lanes 1 to 7.

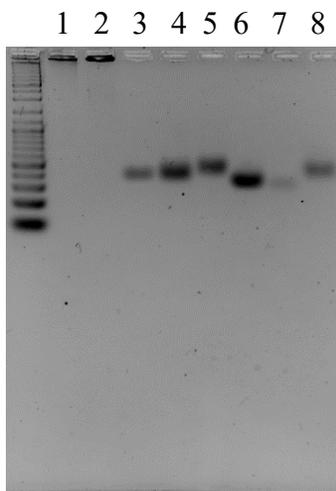


Figure 3.26. Agarose gel (2.5 % AGE) with oligomers from the **A** and **B** series. Oligomers **A4, A9, B2, B6, B5, B3, B1, B4** in respectively lanes 1 to 8. The DNA ladder used is the O'Gene Ruler Mix from ThermoScientific.

3.8 References

- (1) Edwardson, T. G. W.; Carneiro, K. M. M.; Serpell, C. J.; Sleiman, H. F. An Efficient and Modular Route to Sequence-Defined Polymers Appended to {DNA}. *Angew. Chem. Int. Ed. Engl.* **2014**, *53* (18), 4567–4571.
- (2) Dill, K. A.; MacCallum, J. L. The Protein-Folding Problem, 50 Years On. *Science* **2012**, *338* (6110), 1042–1046.
- (3) Solleder, S. C.; Schneider, R. V.; Wetzel, K. S.; Boukis, A. C.; Meier, M. A. R. Recent Progress in the Design of Monodisperse, Sequence-Defined Macromolecules. *Macromol. Rapid Commun.* **2017**, *38* (9), 1600711.
- (4) Szymański, J. K.; Abul-Haija, Y. M.; Cronin, L. Exploring Strategies to Bias Sequence in Natural and Synthetic Oligomers and Polymers. *Acc. Chem. Res.* **2018**, *51* (3), 649–658.
- (5) Gutekunst, W. R.; Hawker, C. J. A General Approach to Sequence-Controlled Polymers Using Macrocyclic Ring Opening Metathesis Polymerization. *J. Am. Chem. Soc.* **2015**, *137* (25), 8038–8041.
- (6) Lawrence, J.; Lee, S.-H.; Abdilla, A.; Nothling, M. D.; Ren, J. M.; Knight, A. S.; Fleischmann, C.; Li, Y.; Abrams, A. S.; Schmidt, B. V. K. J.; et al. A Versatile and Scalable Strategy to Discrete Oligomers. *J. Am. Chem. Soc.* **2016**, *138* (19), 6306–6310.
- (7) Ogura, Y.; Artar, M.; Palmans, A. R. A.; Sawamoto, M.; Meijer, E. W.; Terashima, T. Self-Assembly of Hydrogen-Bonding Gradient Copolymers: Sequence Control via Tandem Living Radical Polymerization with Transesterification. *Macromolecules* **2017**, *50* (8), 3215–3223.

- (8) Kleiner, R. E.; Brudno, Y.; Birnbaum, M. E.; Liu, D. R. DNA-Templated Polymerization of Side-Chain-Functionalized Peptide Nucleic Acid Aldehydes. *J. Am. Chem. Soc.* **2008**, *130* (14), 4646–4659.
- (9) Niu, J.; Hili, R.; Liu, D. R. Enzyme-Free Translation of DNA into Sequence-Defined Synthetic Polymers Structurally Unrelated to Nucleic Acids. *Nat. Chem.* **2013**, *5* (4), 282–292.
- (10) Milnes, P. J.; McKee, M. L.; Bath, J.; Song, L.; Stulz, E.; Turberfield, A. J.; O'Reilly, R. K. Sequence-Specific Synthesis of Macromolecules Using DNA-Templated Chemistry. *Chem. Commun. (Camb)*. **2012**, *48* (45), 5614–5616.
- (11) Barnes, J. C.; Ehrlich, D. J. C.; Gao, A. X.; Leibfarth, F. A.; Jiang, Y.; Zhou, E.; Jamison, T. F.; Johnson, J. A. Iterative Exponential Growth of Stereo- and Sequence-Controlled Polymers. *Nat. Chem.* **2015**, *7* (10), 810–815.
- (12) Lewis, J. E. M.; Winn, J.; Cera, L.; Goldup, S. M. Iterative Synthesis of Oligo[*n*]Rotaxanes in Excellent Yield. *J. Am. Chem. Soc.* **2016**, *138* (50), 16329–16336.
- (13) Grate, J. W.; Mo, K.-F.; Daily, M. D. Triazine-Based Sequence-Defined Polymers with Side-Chain Diversity and Backbone-Backbone Interaction Motifs. *Angew. Chem. Int. Ed. Engl.* **2016**, *55* (12), 3925–3930.
- (14) Martens, S.; Van den Begin, J.; Madder, A.; Du Prez, F. E.; Espeel, P. Automated Synthesis of Monodisperse Oligomers, Featuring Sequence Control and Tailored Functionalization. *J. Am. Chem. Soc.* **2016**, *138* (43), 14182–14185.
- (15) Solleder, S. C.; Zengel, D.; Wetzel, K. S.; Meier, M. A. R. A Scalable and High-Yield Strategy for the Synthesis of Sequence-Defined Macromolecules. *Angew. Chem. Int. Ed. Engl.* **2016**, *55* (3), 1204–1207.
- (16) Knight, A. S.; Zhou, E. Y.; Francis, M. B.; Zuckermann, R. N. Sequence Programmable Peptoid Polymers for Diverse Materials Applications. *Adv. Mater.* **2015**, *27* (38), 5665–5691.
- (17) Al Ouahabi, A.; Charles, L.; Lutz, J.-F. Synthesis of Non-Natural Sequence-Encoded Polymers Using Phosphoramidite Chemistry. *J. Am. Chem. Soc.* **2015**, *137* (16), 5629–5635.
- (18) Kanasty, R. L.; Vegas, A. J.; Ceo, L. M.; Maier, M.; Charisse, K.; Nair, J. K.; Langer, R.; Anderson, D. G. Sequence-Defined Oligomers from Hydroxyproline Building Blocks for Parallel Synthesis Applications. *Angew. Chemie Int. Ed.* **2016**, *55* (33), 9529–9533.
- (19) Chidchob, P.; Edwardson, T. G. W.; Serpell, C. J.; Sleiman, H. F. Synergy of Two Assembly Languages in DNA Nanostructures: Self-Assembly of Sequence-Defined Polymers on DNA Cages. *J. Am. Chem. Soc.* **2016**, *138* (13), 4416–4425.
- (20) de Rochambeau, D.; Barłóg, M.; Edwardson, T. G. W.; Fakhoury, J. J.; Stein, R. S.; Bazzi, H. S.; Sleiman, H. F. “DNA–Teflon” Sequence-Controlled Polymers. *Polym. Chem.* **2016**, *7* (31), 4998–5003.
- (21) König, N. F.; Al Ouahabi, A.; Poyer, S.; Charles, L.; Lutz, J.-F. A Simple Post-Polymerization Modification Method for Controlling Side-Chain Information in Digital Polymers. *Angew. Chemie Int. Ed.* **2017**, *56* (25), 7297–7301.

- (22) Ouahabi, A. Al; Kotera, M.; Charles, L.; Lutz, J.-F. Synthesis of Monodisperse Sequence-Coded Polymers with Chain Lengths above DP100. *ACS Macro Lett.* **2015**, *4* (10), 1077–1080.
- (23) Appukkutti, N.; Serpell, C. J. High Definition Polyphosphoesters: Between Nucleic Acids and Plastics. *Polym. Chem.* **2018**, *9* (17), 2210–2226.
- (24) Kosuri, S.; Church, G. M. Large-Scale de Novo DNA Synthesis: Technologies and Applications. *Nat Meth* **2014**, *11* (5), 499–507.
- (25) Serpell, C. J.; Edwardson, T. G. W.; Chidchob, P.; Carneiro, K. M. M.; Sleiman, H. F. Precision Polymers and 3D DNA Nanostructures: Emergent Assemblies from New Parameter Space. *J. Am. Chem. Soc.* **2014**, *136* (44), 15767–15774.
- (26) Vybornyi, M.; Bur-Cecilio Hechevarria, Y.; Glauser, M.; Rudnev, A. V.; Häner, R. Tubes or Sheets: Divergent Aggregation Pathways of an Amphiphilic 2,7-Substituted Pyrene Trimer. *Chem. Commun.* **2015**, *51* (90), 16191–16193.
- (27) Bösch, C. D.; Langenegger, S. M.; Häner, R. Light-Harvesting Nanotubes Formed by Supramolecular Assembly of Aromatic Oligophosphates. *Angew. Chemie Int. Ed.* **2016**, *55* (34), 9961–9964.
- (28) Kownacki, M.; Langenegger, S. M.; Liu, S.-X.; Häner, R. Integrating DNA Photonic Wires into Light-Harvesting Supramolecular Polymers. *Angew. Chemie* **2019**, *131* (3), 761–765.
- (29) Lutz, J.-F.; Lehn, J.-M.; Meijer, E. W.; Matyjaszewski, K. From Precision Polymers to Complex Materials and Systems. *Nat. Rev. Mater.* **2016**, *1*, 16024.
- (30) Schacher, F. H.; Rupa, P. A.; Manners, I. Functional Block Copolymers: Nanostructured Materials with Emerging Applications. *Angew. Chemie Int. Ed.* **2012**, *51* (32), 7898–7921.
- (31) Chen, M.; Gothelf, K. V. Labeling of DNA via Rearrangement of S-2-Aminoethyl Phosphorothioates to N-(2-Mercaptoethyl)Phosphoramidates. *Org. Biomol. Chem.* **2008**, *6* (5), 908–911.
- (32) Figg, C. A.; Carmean, R. N.; Bentz, K. C.; Mukherjee, S.; Savin, D. A.; Sumerlin, B. S. Tuning Hydrophobicity To Program Block Copolymer Assemblies from the Inside Out. *Macromolecules* **2017**, *50* (3), 935–943.
- (33) Chavis, M. A.; Smilgies, D.-M.; Wiesner, U. B.; Ober, C. K. Widely Tunable Morphologies in Block Copolymer Thin Films Through Solvent Vapor Annealing Using Mixtures of Selective Solvents. *Adv. Funct. Mater.* **2015**, *25* (20), 3057–3065.
- (34) Liu, N.; He, Q.; Wang, Y.; Bu, W. Stepwise Self-Assembly of a Block Copolymer–platinum(II) Complex Hybrid in Solvents of Variable Quality: From Worm-like Micelles to Free-Standing Sheets to Vesicle-like Nanostructures. *Soft Matter* **2017**, *13* (27), 4791–4798.
- (35) Brendel, J. C.; Schacher, F. H. Block Copolymer Self-Assembly in Solution-Quo Vadis? *Chem. - An Asian J.* **2018**, *13* (3), 230–239.
- (36) Govindaraju, T.; Avinash, M. B. Two-Dimensional Nanoarchitectonics: Organic and Hybrid Materials. *Nanoscale* **2012**, *4* (20), 6102.

- (37) Mai, Y.; Eisenberg, A. Self-Assembly of Block Copolymers. *Chem. Soc. Rev.* **2012**, *41* (18), 5969.
- (38) van Genabeek, B.; de Waal, B. F. M.; Gosens, M. M. J.; Pitet, L. M.; Palmans, A. R. A.; Meijer, E. W. Synthesis and Self-Assembly of Discrete Dimethylsiloxane–Lactic Acid Diblock Co-Oligomers: The Dononacontamer and Its Shorter Homologues. *J. Am. Chem. Soc.* **2016**, *138* (12), 4210–4218.
- (39) Berrocal, J. A.; Zha, R. H.; de Waal, B. F. M.; Lugger, J. A. M.; Lutz, M.; Meijer, E. W. Unraveling the Driving Forces in the Self-Assembly of Monodisperse Naphthalenediimide-Oligodimethylsiloxane Block Molecules. *ACS Nano* **2017**, *11* (4), 3733–3741.
- (40) Oschmann, B.; Lawrence, J.; Schulze, M. W.; Ren, J. M.; Anastasaki, A.; Luo, Y.; Nothling, M. D.; Pester, C. W.; Delaney, K. T.; Connal, L. A.; et al. Effects of Tailored Dispersity on the Self-Assembly of Dimethylsiloxane–Methyl Methacrylate Block Co-Oligomers. *ACS Macro Lett.* **2017**, *6* (7), 668–673.
- (41) van Genabeek, B.; de Waal, B. F. M.; Ligt, B.; Palmans, A. R. A.; Meijer, E. W. Dispersity under Scrutiny: Phase Behavior Differences between Disperse and Discrete Low Molecular Weight Block Co-Oligomers. *ACS Macro Lett.* **2017**, *6* (7), 674–678.
- (42) Sutton, A. T.; Arrua, R. D.; Gaborieau, M.; Castignolles, P.; Hilder, E. F. Characterization of Oligo(Acrylic Acid)s and Their Block Co-Oligomers. *Anal. Chim. Acta* **2018**, *1032*, 163–177.
- (43) Pastré, D.; Piétrement, O.; Fusil, S.; Landousy, F.; Jeusset, J.; David, M.-O.; Hamon, L.; Le Cam, E.; Zozime, A. Adsorption of DNA to Mica Mediated by Divalent Counterions: A Theoretical and Experimental Study. *Biophys. J.* **2003**, *85* (4), 2507–2518.
- (44) Robertson, E. J.; Battigelli, A.; Proulx, C.; Mannige, R. V.; Haxton, T. K.; Yun, L.; Whitelam, S.; Zuckermann, R. N. Design, Synthesis, Assembly, and Engineering of Peptoid Nanosheets. *Acc. Chem. Res.* **2016**, *49* (3), 379–389.
- (45) Vybornyi, M.; Rudnev, A.; Häner, R. Assembly of Extra-Large Nanosheets by Supramolecular Polymerization of Amphiphilic Pyrene Oligomers in Aqueous Solution. *Chem. Mater.* **2015**, *27* (4), 1426–1431.
- (46) Holder, S. J.; Sommerdijk, N. A. J. M. New Micellar Morphologies from Amphiphilic Block Copolymers: Disks, Toroids and Bicontinuous Micelles. *Polym. Chem.* **2011**, *2* (5), 1018.
- (47) Laing, B. M.; Barrow-Laing, L.; Harrington, M.; Long, E. C.; Bergstrom, D. E. Properties of Double-Stranded Oligonucleotides Modified with Lipophilic Substituents. *Bioconjug. Chem.* **2010**, *21* (8), 1537–1544.
- (48) Hisamatsu, Y.; Takami, H.; Shirai, N.; Ikeda, S.; Odashima, K. Highly Selective Adenine Recognition by a Macrocyclic Host Molecule Employing Multiple Hydrogen Bonding and Π – π Stacking Interactions. *Tetrahedron Lett.* **2007**, *48* (4), 617–621.

|4|

**Modular Strategy to Expand the Chemical Diversity
of Sequence-Defined Oligo(Phosphodiester)s**

4.1 Preface

In Chapter 2, a new perfluorocarbon containing monomer that is compatible with the synthesis of sequence-defined oligo(phosphodiester)s was developed. This new monomer imparted new properties to DNA strands. In Chapter 3, we added a monomer containing a naphthalene unit capable of π - π stacking. This last phosphoramidite, combined with the PFC monomer and two others offered the possibility to synthesize and characterize a variety of well-defined nanostructures ranging from nanosheets to spherical micelles of different sizes.

These results clearly demonstrate the potential of expanding the monomer library for sequence-defined oligo(phosphodiester)s. However, the synthetic route reported in Chapter 3 requires the use of diols as starting material and the DMT-protection is low yielding. More generally, in polymer chemistry, monomer synthesis is one of the most time-consuming steps. This issue becomes more challenging when dealing with sequence-defined polymers made of multiple types of monomers. A strategy to easily access a variety of phosphoramidite monomers that are compatible with an automated DNA synthesizer may palliate this issue and would tremendously augment the potential applications of sequence-defined oligo(phosphodiester)s.

*This chapter is mostly composed of work published as « Modular Strategy to Expand the Chemical Diversity of DNA and Sequence-Controlled Polymers » by D de Rochambeau, Y Sun, M Barlóg, HS Bazzi & HF Sleiman; J. Org. Chem, **83** (17), 9774 (2018). Some phosphoramidite monomers developed in this chapter have been the subject of a US provisional patent application, « Reagents based on a tertiary amine backbone to introduce chemical functionality in nucleic acids and sequence-controlled polymers » by D de Rochambeau, Y Sun & HF Sleiman, 08/2018.*

4.2 Contribution of authors

Donatien de Rochambeau codesigned the project and performed all experiments unless listed below, analyzed the results and cowrote the manuscript. **Yuanye Sun** synthesized compound **5'** and **5''** and prepared the first batch of **6** and its precursors. **Dr. Maciej Barłóg** synthesized compounds **1'**, **2'**, **4'**, **9'** and their precursors. **Dr. Hanadi Sleiman** codesigned the project, guided interpretation of data and discussion of results and cowrote the manuscript.

4.3 Abstract

Sequence-defined polymers with customizable sequences, monodispersity, substantial length and large chemical diversity are of great interest to mimic the efficiency and selectivity of biopolymers. We report an efficient, facile and scalable synthetic route to introduce many chemical functionalities, such as amino acids and sugars in nucleic acids and sequence-defined oligo(phosphodiester)s. Through achiral tertiary amine molecules that are perfectly compatible with automated DNA synthesis, readily available amines or azides can be turned into phosphoramidites in two steps only. Individual attachment yields on nucleic acids and artificial oligo(phosphodiester)s using automated solid-phase synthesis (SPS) were greater than 90 % in almost all cases. Using this method, multiple water-soluble sequence-defined oligomers bearing a range of functional groups in precise sequences could be synthesized and purified in high yields. The method described herein significantly expands the library of available functionalities for nucleic acids and sequence-controlled polymers.

4.4 Introduction

The precise molecular recognition and folding properties of nucleic acids and proteins derived from their primary sequence give them central roles in biological systems. Driven by the goal of mimicking the structures and functions of these biopolymers,¹ interest in sequence-controlled polymers (SCP) has grown tremendously in the past few years.² Digital storage molecules,^{3,4} foldamers – artificial oligomers with highly-ordered secondary structures,⁵⁻⁸ and single-chain nanoparticles that fold into protein-like structures⁹⁻¹⁵ are some of the applications envisaged by these studies. For these purposes, a strategy allowing the synthesis of monodisperse oligomers with a large number of chemical monomer variants is of great interest.

Solid-phase synthesis remains the method of choice for the most precise sequence control.^{16–20} Automated solid-phase phosphoramidite chemistry, in particular, has shown exceptional coupling yields for the synthesis of DNA and RNA.^{21,22} Decades of optimization have allowed it to attain the highest degrees of polymerization (DP) for a solid-phase synthesis. Up to 200 monomer-long oligonucleotides can be made in good yields and with simple purification methods.²³ Importantly, the cost of phosphoramidite synthesis has been significantly and steadily declining, making it a practical as well as powerful strategy.²⁴

In previous work, Chapter 2 and Chapter 3 of this thesis, our group developed an efficient strategy to make sequence-defined polymers based on automated phosphoramidite solid-phase synthesis.^{25–29} This method allowed the formation of poly(phosphodiester)s³⁰ that are stable and highly soluble in water due to their anionic backbone. As outlined in the introduction, other groups showed the synthesis of long artificial sequence-defined poly(phosphodiester)s (DP of 100)³¹ and reported a method for post synthesis dual functionalization.³² Phosphoramidite chemistry has also been used to make “oligopyrenotides” with novel supramolecular assembly properties.^{33,34} Thus, monodispersity, cost-efficiency, length of the oligomers and perfect sequence-control have been achieved using automated solid-phase synthesis with phosphoramidite chemistry. The next goal for this method is to expand the library of available compatible phosphoramidites in a straightforward and scalable manner.

Solid-phase synthesis on a DNA synthesizer requires the phosphoramidite monomers to stay unaltered to a great number of chemical conditions (repetitive treatment with oxidant and mild acid and final deprotection in aqueous base). For the purpose of versatile and multiple functionalization of synthetic oligo(phosphodiester)s, even more restrictive conditions apply, making the task arduous. (i) Due to the need for numerous monomers, their synthesis must be fast, cost effective and scalable. This notably prevents the use of nucleoside derived phosphoramidites.³⁵ (ii) Very high coupling yields, leading to high DPs are essential. As an example, a 22mer would be obtained in less than 10 % yields if coupling yields do not exceed 90 %. (iii) The monomer should preferably not contain a chiral center, or it must be enantiomerically pure. A mixture of enantiomers should be avoided as the stereochemical complexity of the oligomer increases exponentially with each monomer addition. Previous reports of non-nucleosidic phosphoramidites only partially satisfy these criteria. For example, propanediol^{36,37},

threoninol³⁸ and serinol^{39,40}-based phosphoramidites can be attached on a DNA strand but their synthesis starts either with racemic mixtures or costly enantiopure diols. Other reports show the elegant synthesis and use of hydroxyprolinol^{41,42} and oxamide^{43,44}-based phosphoramidites, but these may lead to stability issues during the final deprotection under basic conditions.

In this chapter, we explore the use of a tertiary amine scaffold to make a range of multifunctional DNA strands and sequence-defined polymers via phosphoramidite synthesis. Two monomers bearing two convenient chemical handles for the attachment of a variety of side-chains in two steps were designed: an alkyne-containing unit would allow the introduction of azides (e.g., sugar molecules), while a carboxylate-containing molecule would allow functionalization with amine containing monomers, including amino acids (**Figure 4.1**). A primary amine-containing monomer was also successfully synthesized and incorporated in DNA strands. Applicability of this strategy was tested with the synthesis of a variety of oligo(phosphodiester)s made of artificial building blocks. This tertiary amine backbone is a cost-effective, efficient, robust and achiral alternative to commercially available DNA modifications, and will open the door to significantly greater chemical diversity in both oligonucleotide and sequence-controlled polymer synthesis.

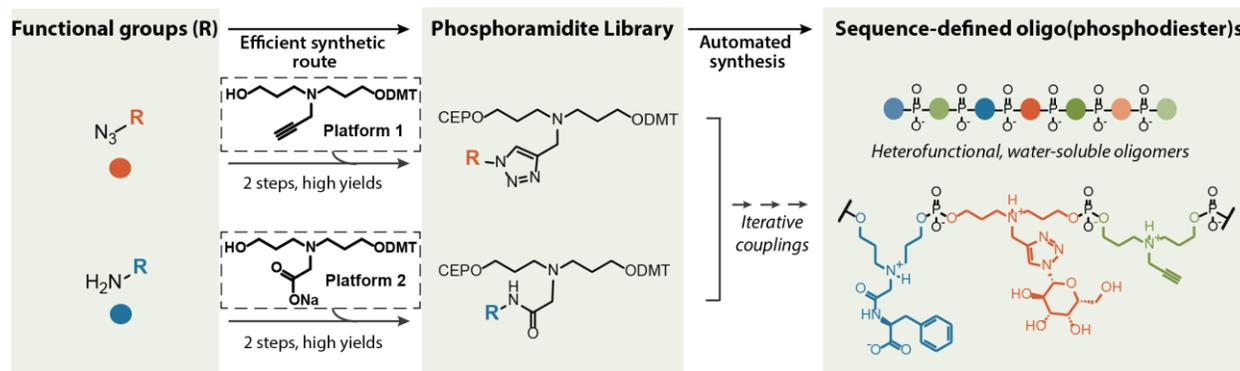


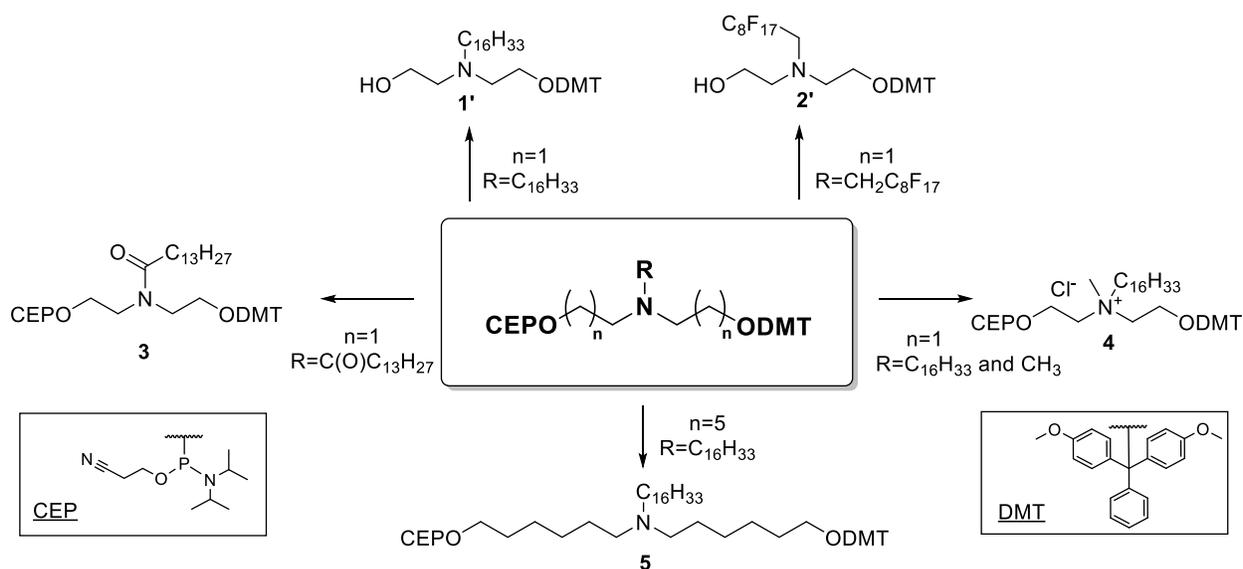
Figure 4.1. Synthesis of versatile monomers for sequence-defined oligo(phosphodiester)s. CEP stands for cyanoethylphosphoramidite and DMT for dimethoxytrityl.

4.5 Results and discussion

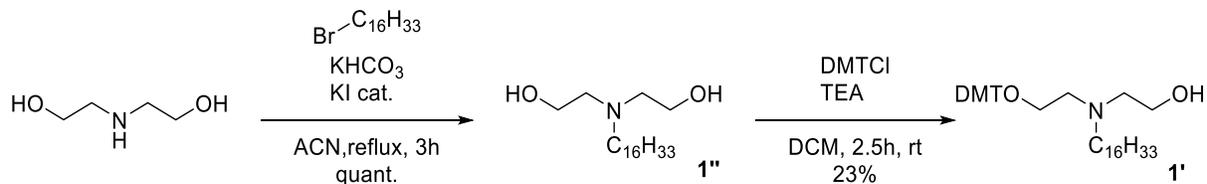
4.5.1 Tertiary amine backbones

The starting point for this study is the observation that tertiary amines (such as diisopropylethylamine and triethylamine) are commonly used as bases and thus compatible with phosphoramidite synthesis.⁴⁵ We reasoned that an achiral tertiary amine-containing backbone may be a good candidate for the introduction of functional groups. Initially, diethanolamine was chosen

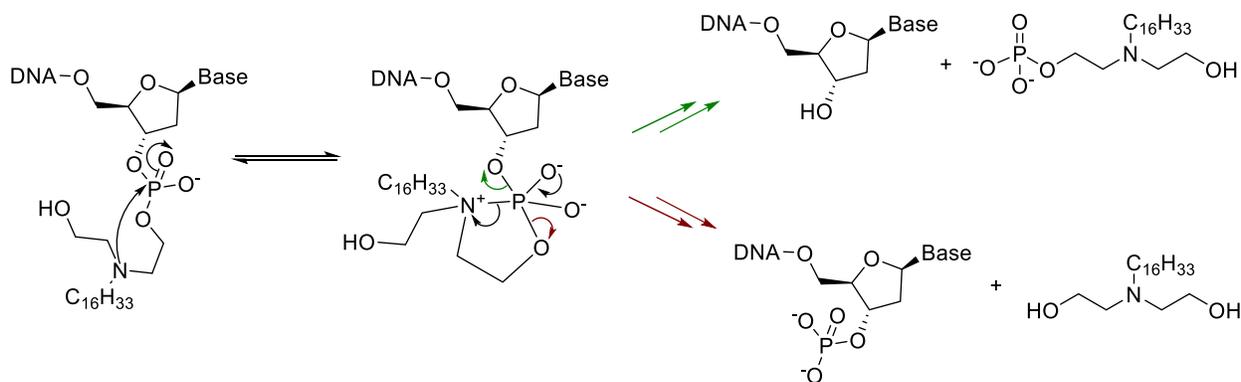
as starting material due to its ready availability.⁴⁶ Simple nucleophilic substitution with the moiety of interest on the secondary amine is then required to functionalize the monomer. As a preliminary proof of concept, molecules **1'** (**Scheme 4.1** and **Scheme 4.2**) and **2'** (precursor for N[PFC] in Chapter 2) were synthesized and attached to a DNA strand following a standard procedure that does not require isolation of the phosphoramidite (see Section 4.7.4.1). Excellent coupling yields were observed for molecule **2'** (> 90 %) as shown in Chapter 2 whereas molecule **1'** consistently led to unsatisfactory coupling yields followed by spontaneous degradation of the DNA strand in water at 4 °C after a few days. Similarly to the degradation pathway of RNA in basic conditions, we hypothesized that the nucleophilic nitrogen atom in the case of molecule **1'** can attack an electrophilic phosphorus atom forming a favorable 5-membered ring (**Scheme 4.3**). On the other hand, the fluorine atoms in **2'** exert an inductive effect that reduces the nitrogen lone pair nucleophilicity, thus possibly explaining the good yields obtained with this molecule. As further evidence for this mechanism, a phosphoramidite bearing an anthraquinone moiety directly attached to the diethanolamine nitrogen atom (thus also reducing its nucleophilicity) was reported elsewhere and led to good yields.⁴⁷



Scheme 4.1. Different backbones explored.

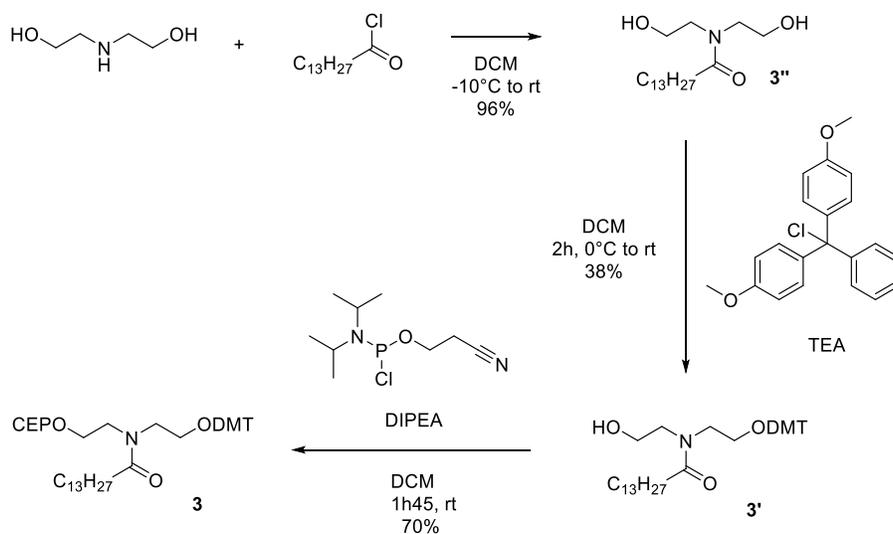


Scheme 4.2. Synthetic route for **1'**.

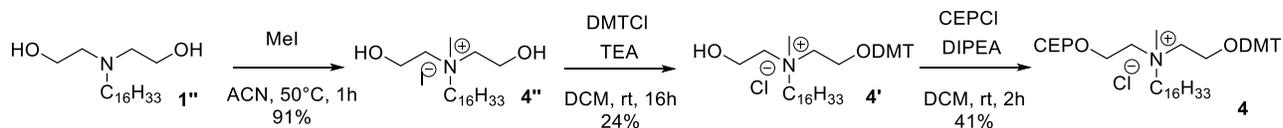


Scheme 4.3. Potential degradation mechanism of diethanolamine based monomers.

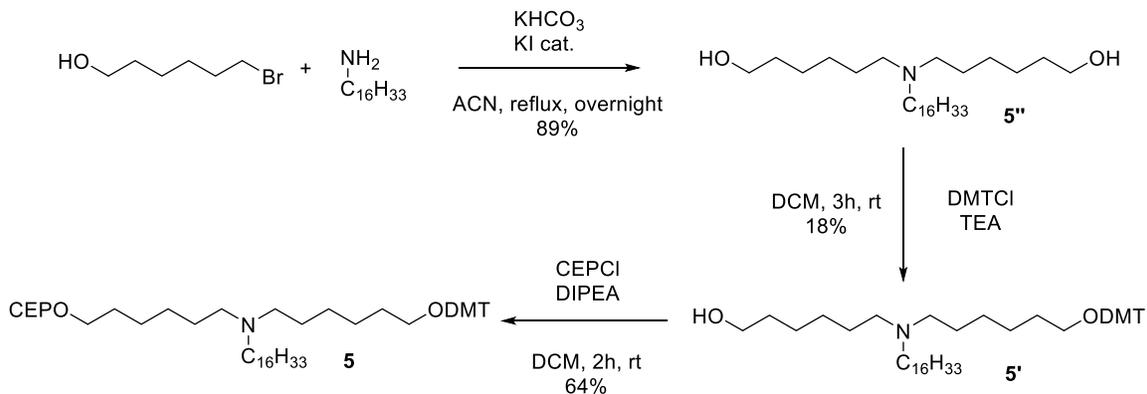
This observation led to further monomer optimization. Molecules **3**, **4** and **5** (**Scheme 4.1**) were used to check backbone suitability. For **3** and **4**, the previous degradation mechanism is less likely to happen because they do not have a nucleophilic nitrogen, while **5** has a longer spacer separating the nucleophilic nitrogen from the phosphorus, thus avoiding the 5-membered ring intermediate in this degradation mechanism. **3** was made following a straightforward synthesis: myristoyl chloride was added onto diethanolamine leading to an amide in quantitative yields, which was transformed to phosphoramidite **3** following standard procedures (**Scheme 4.4**). **4** was made from the precursor for **1'** using methyl iodide, followed by DMT protection and phosphoramidation using standard procedures (**Scheme 4.5**). Finally, bromohexanol was substituted onto hexadecylamine, leading to compound **5'** further transformed to **5** (**Scheme 4.6**).



Scheme 4.4. Synthetic route for **3**



Scheme 4.5. Synthetic route for **4**.



Scheme 4.6. Synthetic route for **5**.

These monomers were coupled to the 5' end of a DNA strand as a first step. Molecule **3** did not lead to good yields (< 30 %, **Figure 4.2**, **Table 4.1**) and the synthesis of T-HEG-HEG-3-3-T (using commercially available thymidine (T) and hexaethyleneglycol (HEG) phosphoramidites) gave low coupling yields (**Table 4.1**). On the other hand, molecule **4** led to high yields when

inserting it at the 5' end of a DNA strand and on sequence-defined oligomers (**Table 4.1**). Molecule **4** is of interest because its amphiphilic character can increase its affinity towards membrane lipids. The positively charged backbone in **4** can lead to novel sequence-defined zwitterionic polymers with positive quaternary ammonium and negative phosphate groups. The quaternary ammonium center is chiral in this case but could be made achiral by condensing twice the same iodide onto diethanolamine. Molecule **5** was incorporated in an oligomer and on a DNA strand with the best yields obtained so far (**Figure 4.2**, **Table 4.1**). This result indirectly confirms the potential degradation mechanism that occurred with **1'**. Yields obtained are high enough to consider using such a backbone for making more complex sequence-defined oligomers with a higher number of units. Moreover, **5** is pH responsive (predicted pK_b is 3.2 from online prediction tool)⁴⁸ and most probably protonated at neutral pH, which allows better solubility and possible zwitterionic character, as with oligomers of **4**.

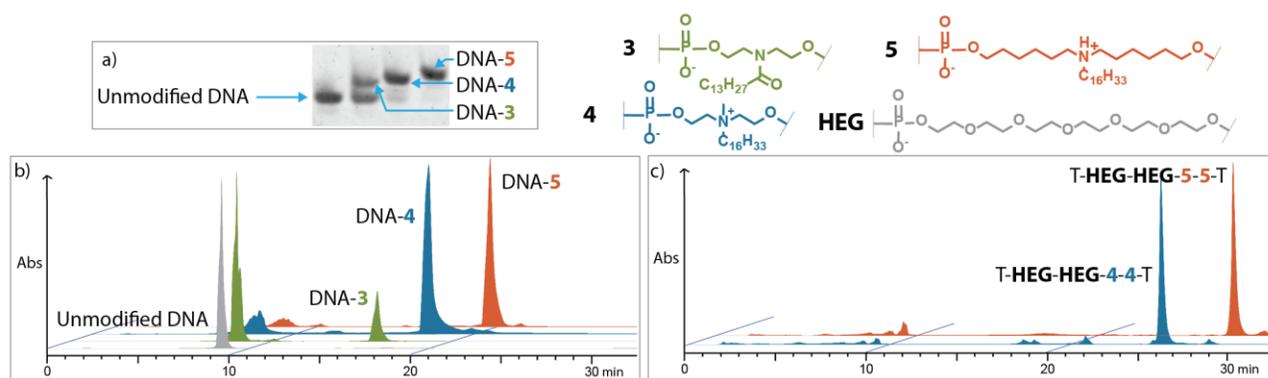


Figure 4.2. Attachment of **3**, **4** and **5** to DNA and oligomers. (a) 18 % polyacrylamide gel electrophoresis (PAGE) for DNA 19mer singly modified with **3**, **4** and **5** at the 5' end and (b) Reverse-phase high performance liquid chromatography (RP-HPLC) traces (UV detection, 260 nm) from crude mixtures. (c) RP-HPLC traces (UV detection, 260 nm) of sequence-defined oligomers containing hexaethylene glycol **HEG** and thymidines (T). Gradients are available in **Figure 4.10** and **Figure 4.11**.

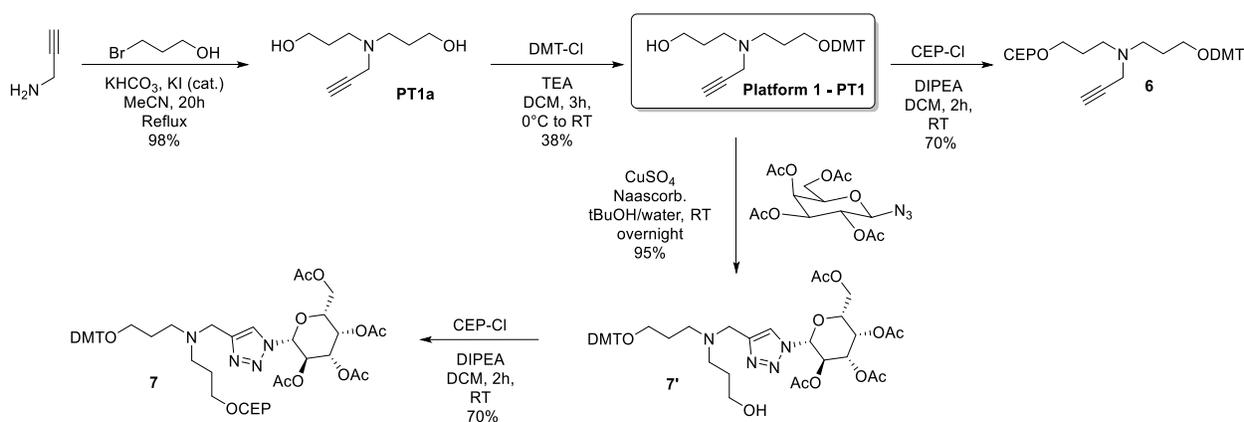
Table 4.1. Yields and ESI-MS characterization of oligomers containing monomers **3**, **4** and **5**.

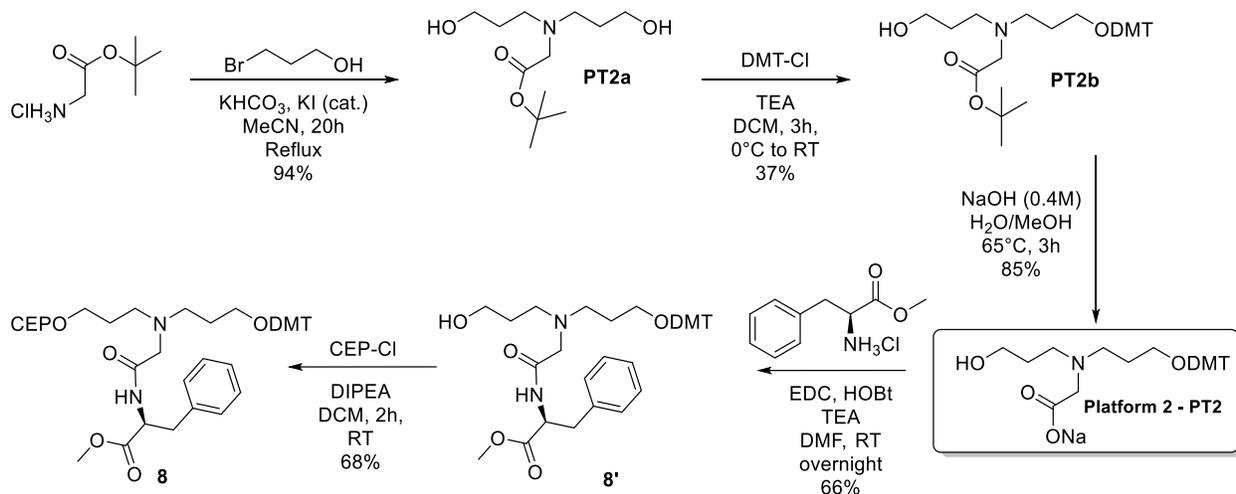
Strand	Unique incorporation yields ^a (%)	Global yields ^b (%)	Calculated exact mass (g mol ⁻¹)	Found exact mass ^c (g mol ⁻¹)
DNA- 3	26-44	-	6142.23	6142.28
DNA- 4	90-92	-	6171.29	6171.31
DNA- 5	90-96	-	6268.41	6268.47
T- HEG-HEG-3-3-T	NA	<20	1988.84	1988.69
T- HEG-HEG-4-4-T	90	79	2047.00	2046.82
T- HEG-HEG-5-5-T	97	82	2241.20	2241.03

^a Yields were calculated through image analysis of electrophoretic mobility shift assays and through product peak area by RP-HPLC (260 nm detection) vs unmodified DNA. For non-DNA oligomers, HPLC yields of the second amidite (**3**, **4** or **5**) introduction are reported. ^b Product peak area by RP-HPLC over total area. ^c Mass was found using electrospray ionization-mass spectrometry (ESI-MS) technique, detecting multiply charged species (**Figure 4.12**, **Figure 4.13**).

4.5.2 Synthesis of the optimized monomer platforms

Based on the conclusions obtained from the backbone investigations, we synthesized two molecules to be used as versatile platforms to attach a variety of moieties (**Scheme 4.7**, **Scheme 4.8**).

**Scheme 4.7.** Synthetic route for platform **PT1**, alkyne phosphoramidite **6** and a β -D-glucose-containing phosphoramidite **7**.



Scheme 4.8. Synthetic route for platform **PT2** and a phenylalanine-functionalized phosphoramidite **8**.

Molecules **PT1** and **PT2** have many advantages: they are easy to synthesize in two steps for **PT1** and three for **PT2** from readily available starting materials, they need only one chromatography purification and they have a reactive function for simple and efficient attachment of the moiety of interest. Furthermore, the DMT monoprotection being the lowest yielding step, we can limit material waste by performing it as the second step and attach the moiety of interest afterwards. Compound **PT1** has an alkyne moiety for copper-catalyzed azide-alkyne cycloaddition (CuAAC) chemistry. Many bio-related compounds are now available as azides making this platform highly versatile for possible functionalization with carbohydrates, dyes or cyclodextrin containing molecules for example. As with **5**, oligomers from **PT1** can possibly have zwitterionic character, due to protonation of their tertiary amine moiety. Compound **PT2** is of interest as well since its carboxylate allows common amide coupling with readily available amines such as amino acids or peptides. These two platforms were synthesized with good yields after double alkylation of propargylamine (**Scheme 4.7**) and glycine t-butyl ester (**Scheme 4.8**) respectively with bromopropanol, followed by DMT protection of one hydroxyl group. As before, **PT1** and **PT2b** were attached to a DNA strand without isolation of the phosphoramidite. These results confirmed that **PT1** and **PT2** backbones are suitable for phosphoramidite-based SPS (**Figure 4.3**, **Table 4.2**).

In addition, **PT1** and **PT2** were functionalized with model molecules. We chose a protected sugar azide, as well as a methyl ester amino acid, phenylalanine. Standard CuAAC procedure was performed to obtain molecule **7'** in 95 % yields which was then turned into phosphoramidite **7**. Only one chromatography purification is required here. Phenylalanine could be coupled

successfully to **PT2** and the obtained molecule was turned into monomer **8**. Successful incorporation of **7**' and **8**' on a DNA strand was achieved and high incorporation yields were obtained in both cases (**Figure 4.3**, **Table 4.2**). The deprotection conditions used are suitable for classic DNA synthesis and are described in the supporting information. Intriguingly, methylamine/ammonium hydroxide deprotection conditions allowed the quantitative conversion of the methyl ester into the secondary methyl amide derivative, which can be of interest. As a conclusion, we show the simple and cost-effective incorporation of an azide molecule or a primary amine bearing compound in DNA strands. Due to the non-chiral character of the platforms used and the good to excellent incorporation yields obtained, one could consider these molecules as good candidates for the efficient preparation of sequence-defined oligomers with a variety of side chain functions. Finally, the ease of preparation of **PT1**, **PT2** and their functionalized counterparts illustrates the great practicality and ready accessibility of this method. Sequence-specific on-bead functionalization of **PT1** in the synthesizer could also be an efficient method. Such a method has been reported with an alkyne containing phosphonate reagent elsewhere.⁴⁹ It was shown to be very modular with individual coupling yields between 63 and 89 %.

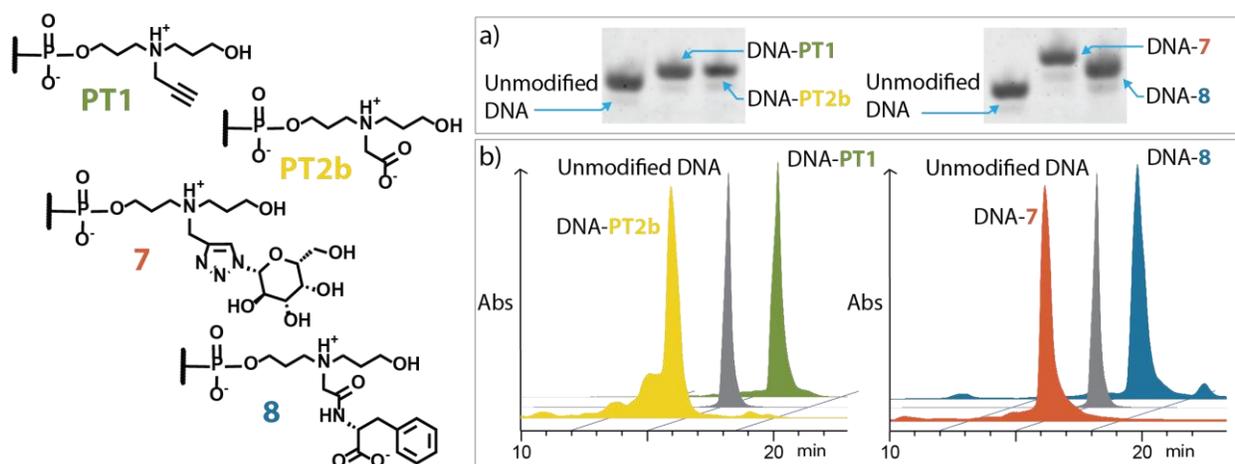


Figure 4.3. Attachment of **PT1**, **PT2b**, **7** and **8** to DNA. (a). 18 % PAGE for DNA 19mers modified with **PT1**, **PT2b**, **7** and **8** at the 5'end. (b). RP-HPLC traces (UV detection, 260 nm) from crude mixtures. Gradients are available in **Figure 4.10**.

Table 4.2. Yields and ESI-MS characterization of DNA strands modified with **PT1**, **PT2b**, **7** and **8**.

Modification	Incorporation yields ^a (%)	Calculated exact mass (g mol ⁻¹)	Found exact mass ^b (g mol ⁻¹)
PT1	93-99	5998.06	5998.00
PT2b	90-92	6018.05	6018.01
7	88-94	6203.13	6203.13
8	90-94	6165.13	6165.01

a Yields were calculated through image analysis of electrophoretic mobility shift assays and through product peak area by RP-HPLC (260 nm detection) vs unmodified DNA. b Product peak area by RP-HPLC over total UV absorption during the run. c Mass was found using ESI-MS technique, detecting multiply charged species (**Figure 4.12**).

4.5.3 Sequence-defined oligo(phosphodiester)s

The new tertiary amine platforms allowed the modular and highly efficient synthesis of sequence-defined oligomers featuring different moieties. The phosphoramidites were used in a standard automated DNA synthesizer along with commercially available DNA synthesis reagents. Coupling cycles were the same as for DNA synthesis except that the coupling time was increased to 10 minutes and phosphoramidite concentration was kept about 0.1 M in dichloromethane. After synthesis and deprotection, isolation could be performed using RP-HPLC and the oligomers were further characterized using LC-MS techniques. Due to the ionic nature of our polymers, they are perfectly soluble in water leading to great ease of manipulation. As with nucleic acids, they can be quantified easily using spectrophotometry with the appropriate absorption coefficients (see Section 4.7.4.3). Yields were measured on the HPLC traces considering the relative absorbance of each byproduct. Most sequences were designed to show high functional diversity on one chain. Oligomers **A** to **D** show the possibility of using up to 6 different phosphoramidites and highlight single coupling yields over 90 % in all cases (**Figure 4.4**, **Table 4.3**). We were able to work with the new monomers based on the tertiary amine backbone (**6**, **7** and **8**) combined with commercially available monomers (nucleotides and triethyleneglycol, **TEG**), hydrophobic 12 carbon long chain phosphoramidite (**10**) and a novel naphthalene-containing amidite (**9**). Each monomer introduces a new orthogonal interaction to DNA and sequence-defined polymers, in a similar manner to functionalities imparted to peptides and proteins by the 20 amino acids. For example, **10** and **TEG** can be used to tune the hydrophobicity and hydrophilicity of the oligomers,²⁵ **9** was synthesized to

allow the introduction of π - π stacking interactions. More interestingly, **PT1** and **PT2** can be substituted with a wide range of functionalities leading to sugar-**7** and amino acid-**8** for example. One could imagine using **7** and **8** to impart host-guest interactions, hydrogen-bonding and new biological activity. With such a palette of interactions, one could imagine making functional polymers with protein-like selectivity and efficiency.

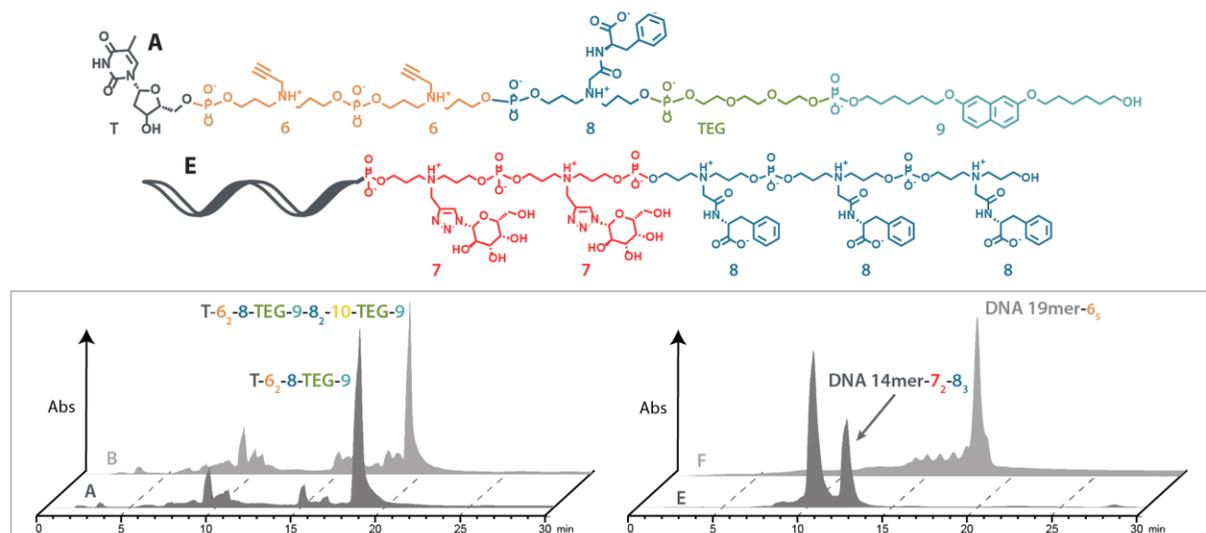


Figure 4.4. RP-HPLC chromatograms of sequence-defined oligomers with novel monomers. UV 260 nm detection, gradients are available in **Figure 4.11**.

Table 4.3. Yields and ESI-MS characterization of sequence-defined oligo(phosphodiester)s.

Oligo-mer	Sequence (iterative synthesis from left monomer)	Global yields ^a (%)	Average individual coupling yields ^b (%)	Calculated exact mass (g mol ⁻¹)	Found exact mass ^c (g mol ⁻¹)
A	T- 6 ₂ - 8 -TEG- 9	64	91	1742.62	1742.63
B	T- 6 ₂ - 8 -TEG- 9 - 8 ₂ - 10 -TEG- 9	36	90	3442.28	3442.30
C	T ₃ -HEG ₄ - 7 - 6 ₂	60	93	3130.99	3130.97
D	T- 6 - 8 - 6 - 8 - 6 - 10	52	90	2005.76	2005.74
E	DNA 14mer- 7 ₂ - 8 ₃	34	81	6321.47	6321.44
F	DNA 19mer- 6 ₅	82	96	6930.39	6930.16

^a Yields were calculated through product peak area by RP-HPLC (260 nm detection) considering the relative absorbance of byproducts. ^b These numbers were obtained by taking the global yields at the power 1/n, n being the number of couplings after the closest nucleotide. ^c Mass was found using ESI-MS technique, detecting multiply charged species except for oligomers **A** and **D**. See **Figure 4.13**.

To further show the versatility of monomer **6**, we carried out post-SPS functionalization on two DNA strands. We synthesized two DNA 19mers containing respectively 1 (**Figure 4.3**) and 5 alkyne moieties (96 % individual coupling yields, oligomer **F**, **Figure 4.4**). Following a straightforward protocol for copper catalyzed cycloaddition on oligonucleotides,⁵⁰ high yielding functionalization of these strands with β -D-glycosyl azide was achieved (> 80 %, **Figure 4.5**). Interestingly we also report the synthesis of novel hybrid oligomers containing DNA, glycosyl modifications as well as amino acids (oligomer **E**). Indeed, a DNA 14mer was functionalized twice with phosphoramidite **7** followed by three times with **8**. Yields are high for such a complex hybrid (**Table 4.3**). Functional diversity imparted by the strategy described here in DNA strands could significantly expand the toolbox of DNA nanotechnology.⁵¹

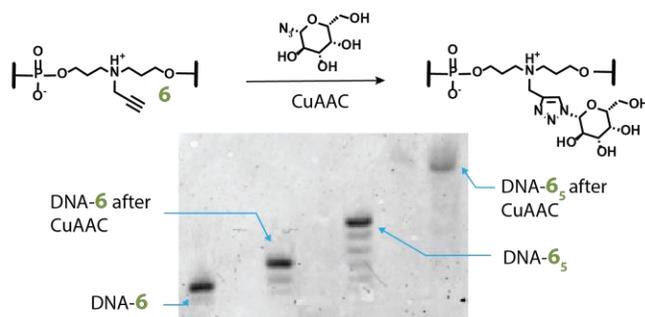
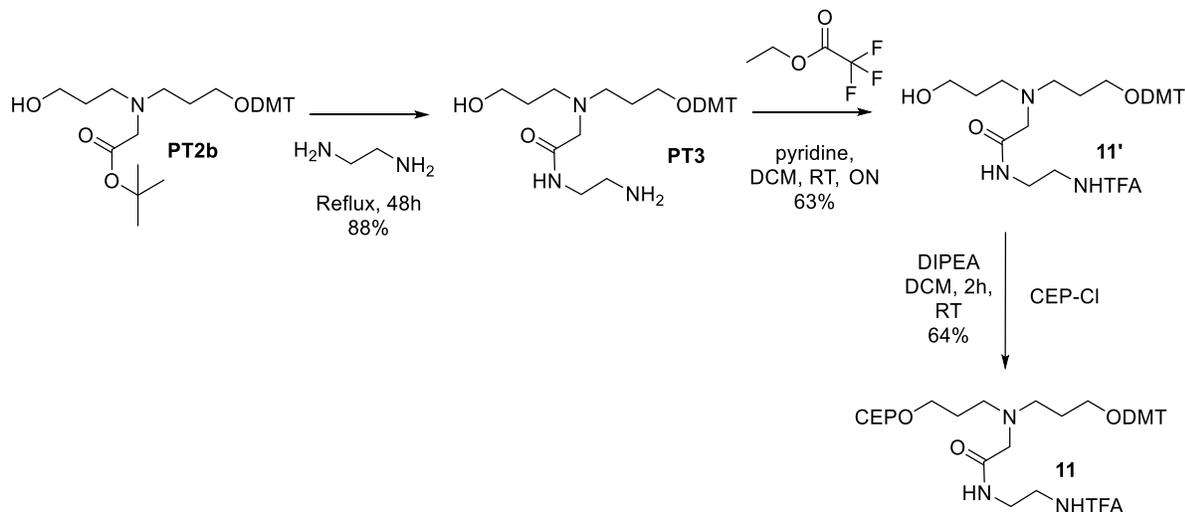


Figure 4.5. Post SPS functionalization of alkyne monomer **6** with β -D-glycosyl azide. 18 % PAGE in denaturing conditions.

4.5.4 Design of an additional primary amine-containing platform

So far, we showed the possibility to easily convert amines and azides into phosphoramidites in two steps. Biotin, fluorescein and folic acid are three moieties of interest to attach to sequence-defined oligo(phosphodiester)s. The two first ones are broadly used for immobilization of oligomers through binding to streptavidin⁵² and for fluorescent assays respectively.⁵³ Folic acid is used to target folate receptors that is overexpressed in many tumor types.⁵⁴ These moieties are mostly available as carboxylic acids. To be able to turn these moieties into phosphoramidites in two steps, we developed a third platform (**PT3**) that bears a primary amine moiety. This platform was made through reaction of **PT2b** in ethylene diamine that did not require a chromatography purification (**Scheme 4.9**). This new molecule is thought to be suitable for coupling to carboxylic acid. In order to allow post-SPS functionalization, we also protected the amine function with a trifluoroacetamide group and made the associated phosphoramidite (**11**).



Scheme 4.9. Synthesis of platform 3 (**PT3**) and phosphoramidite **11**.

Phosphoramidite **11** was successfully coupled at an internal position of a DNA 21mer (yields >94 % through gel image analysis, **Figure 4.6**) and characterized through LC-ESI-MS. It brings two positive charges and one negative charge. Therefore, it may be used similarly than in “zip nucleic acids” (oligonucleotide-oligospermine conjugates)⁵⁵ for which positively charged amino groups induce self-transfection of oligomers into cells.



Figure 4.6. Attachment of amine-functionalized containing monomer in a DNA 21mer. 15 % PAGE in denaturing conditions. Yields were found to be over 94 % by gel image analysis.

4.6 Conclusion

We highlighted here a strategy to significantly increase the chemical diversity of phosphoramidite monomers. These can be used in an automated synthesizer for the synthesis of sequence-defined oligomers and modification of nucleic acids. We investigated several tertiary amine backbones to be used as platforms for making a large variety of phosphoramidite monomers. This study led to the design of three achiral molecules that could be obtained in a few simple synthetic steps from inexpensive starting materials and that bear a reactive moiety. Their attachment on DNA strands

was shown to be very efficient. As a model study, the first platform was functionalized with a sugar azide derivative and turned into a phosphoramidite. This new monomer could be attached to a DNA strand and several sequence-defined oligomers with good yields. Phenylalanine was attached to platform 2 through standard amine coupling conditions and was further attached to DNA with very good yields. Noticeably, this strategy allowed to synthesize a new class of zwitterionic sequence-defined oligo(phosphodiester)s with average individual coupling yields rarely under 90 %. Thus, the goal of increasing side-chain diversity in DNA and sequence-defined oligomers could be achieved thanks to achiral, easily synthesized, inexpensive phosphoramidites that are compatible with automated solid-phase synthesis. This strategy paves the way to the synthesis of an extensive library of monomers. This would be a promising step towards the synthesis of functional sequence-defined polymers and ultimately fully artificial enzyme mimics.

4.7 Experimental section

4.7.1 Chemicals

All starting materials were obtained from commercial suppliers and used without further purification unless otherwise noted. Dimethoxytrityl chloride (DMT-Cl) was purchased from GenScript. (3-Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDC-Cl), glycine tert-butyl ester hydrochloride were purchased from AK Scientific. 3-bromo propanol was purchased from Alfa Aesar. Acetic acid, Boric acid, solvents were purchased from Fisher Scientific. Chloroform-*d*1 was purchased from Cambridge Isotope Laboratories. Importantly, it was stored on molecular sieves in order to keep it neutral. If used as sold, hydrolysis of phosphoramidite (fast) as well as DMT deprotection (slow) may be observed. GelRed™ nucleic acid stain was purchased from Biotium Inc. Concentrated ammonium hydroxide, ammonium persulfate, acrylamide/Bis-acrylamide (40 % 19:1 solution) and TEMED were obtained from Bioshop Canada Inc. and used as supplied. 1 μmol Universal 1000Å LCAA-CPG supports and standard reagents used for automated DNA and RNA synthesis were purchased through Bioautomation. N,N-diisopropylamino Cyanoethyl phosphonamidic-chloride (CEP-Cl) and DMT-hexaethyloxy glycol (cat.# CLP-9765) phosphoramidites were purchased from Chemgenes. Sephadex G-25 (super fine, DNA grade) was purchased from Glen Research. All other reagents were obtained from Sigma-Aldrich. TEAA (triethylammonium acetate) buffer is composed of 50 mM TEA with pH adjusted

to 8.0 using glacial acetic acid. TBE buffer is 90 mM Tris, 90 mM boric acid and 1.1 mM EDTA with a pH of 8.0.

4.7.2 Instrumentation

Standard automated solid-phase synthesis was performed on a Mermade MM6 synthesizer from Bioautomation. HPLC purification was carried out on an Agilent Infinity 1260. DNA and oligomers quantification measurements were performed by UV absorbance with a NanoDrop Lite spectrophotometer from Thermo Scientific. A Varian Cary 300 Bio spectrophotometer was used for UV absorbance studies. Polyacrylamide Gel Electrophoresis (PAGE) experiments were carried out on a 20 X 20 cm vertical Hoefer 600 electrophoresis unit. Gel images were captured using a ChemiDoc™ MP System from Bio-Rad Laboratories. Dry solvents were taken from an Innovation Technology device. Low Resolution Mass determination was carried out using Electron-Spray Ionization – Ion Trap - Mass Spectrometry (MS) on a Finnigan LCQ Duo device. High Resolution mass determination was achieved using a Bruker Maxis API (Atmospheric pressure ionization) QTOF or a THERMO Exactive Plus Orbitrap-API. Liquid Chromatography Electrospray Ionization Mass Spectrometry (LC-ESI-MS) of oligomers was carried out using Dionex Ultimate 3000 coupled to a Bruker MaXis Impact™ QTOF. Some oxygen and moisture sensitive experiments were carried out in a Vacuum Atmospheres Co. glove box. Column chromatography was performed using a CombiFlash Rf system from Teledyne Isco. The NMR spectra were recorded on Bruker 400 MHz, 500 MHz, Varian 300 MHz or 400 MHz for ^1H , ^{13}C , ^{19}F and ^{31}P with chloroform- d_1 (δ 7.26, ^1H ; δ 77.16, ^{13}C), acetone- d_6 (δ 2.04, ^1H ; δ 29.8, ^{13}C) or DMSO- d_6 (δ 2.50, ^1H ; δ 39.5, ^{13}C) as internal lock solvents and chemical shift standard.

4.7.3 Small molecule synthesis

Syntheses of **2'** was reported in Chapter 2 and syntheses of **9** and **10** in Chapter 3.

4.7.3.1 General procedure for dimethoxytrityl monoprotection.

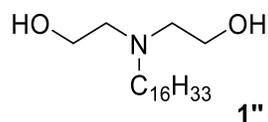
3.63 mmol of the diol starting material and 1.47 mL of TEA (10.9 mmol, 3 equiv.) were dissolved in 30 mL of dry DCM in a 100 mL dry round bottom flask. DMT chloride (1.23 g, 3.63 mmol, 1 equiv.) was dissolved in 20 mL of dry DCM and added dropwise to the reaction mixture at 0 °C under vigorous stirring. After the chloride addition, the reaction was allowed to warm up to room

temperature and left under stirring for 2h30min. Purification conditions are described for each molecule individually. Silica columns have been pretreated with a solution containing 0.1 % to 1 % of triethylamine.

4.7.3.2 General procedure for converting the monoprotected alcohols into phosphoramidites.

Monoprotected diol was suspended in toluene or acetonitrile and solvent was evaporated under reduced pressure (60 °C). The operation was repeated once and the dried compound was kept under high vacuum for at least 5 hours. In an oven-dried flask, 0.64 mmol of alcohol starting material were then dissolved in anhydrous DCM (5 mL) and 0.56 mL of dry DIPEA (3.2 mmol, 5 equiv.) was added under stirring. CEP-Cl (0.14 mL, 0.64 mmol, 1 equiv.) was added slowly and the reaction was allowed to stir under inert gas at room temperature for 2 hours. Two fast extractions with DCM from a 10 % Na₂CO₃ aqueous solution were performed. Organic fractions were combined, dried over MgSO₄, filtered and the solvent was evaporated under reduced pressure (40 °C). The crude product was loaded on celite and purified using the combiFlash system. Silica columns have been pretreated with a solution containing 1 % of triethylamine. N.B.: Care should be taken to avoid any trace of acid, or acidic pH in the presence of phosphoramidites, as this results in extensive degradation. For example, CDCl₃ needs to be stored on molecular sieves in order to keep it neutral.

4.7.3.3 2,2'-(Hexadecylazanediyl)diethanol, 1''.

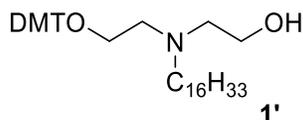


Diethanolamine (3.8 g, 36 mmol) and 1-Hexadecylbromide (9.19 g, 30 mmol) were added to dried round bottomed flask (RBF) charged with KHCO₃ (6.02 g, 60 mmol) and KI (0.5 g, 3 mmol) then mixed in dry acetonitrile (80 mL). The reaction mixture was heated under reflux for 3 hours, cooled down and the solvent removed under vacuum. Residue was taken in dichloromethane (100 mL) and washed with water (3 x 100 mL). Organic layer was dried with MgSO₄ and solvent removed under vacuum to obtain a yellow oil solidifying fast on standing, forming slightly orange waxy solid as a pure product in quantitative yield (9.82 g). Characterization data matched those reported elsewhere.⁵⁶

^1H NMR (400 MHz, CDCl_3) δ 0.87 (t, $J = 7.0$ Hz, 3H), 1.15-1.36 (m, 26H), 1.36-1.54 (br. m, 2H), 2.55 (t, $J = 7.6$ Hz, 2H), 2.69 (t, $J = 5.4$ Hz, 4H), 2.74 (br. s, 2H), 3.63 (t, $J = 5.4$ Hz, 4H).

^{13}C NMR (100 MHz, CDCl_3) δ 14.1, 22.7 – 31.9, 54.8, 56.1, 59.4.

4.7.3.4 2-((2-(Bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)(hexadecyl)amino)ethanol, **1'**



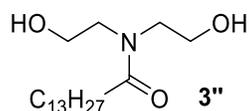
Reaction was performed from 9.1 mmol of **1''** as in Section 4.7.3.1. After the reaction solvent was evaporated and resulting yellow oily residue was purified by chromatography on triethylamine pre-treated silica gel with slow gradient of EtOAc/Hexane (0-15 %) to obtain the product as a sticky yellow oil: 1.32 g, 23 % (pure fractions).

HRMS (ESI-QTOF) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{41}\text{H}_{62}\text{NO}_4$ 632.4673; Found 632.4671.

^1H NMR (400 MHz, Acetone- d_6) δ 0.88 (t, $J = 7.0$ Hz, 3H), 1.29 (br. s, 26H), 1.43-1.48 (br. m, 2H), 2.48 (t, $J = 7.4$ Hz, 2H), 2.58 (t, $J = 6.0$ Hz, 2H), 2.71 (t, $J = 5.9$ Hz, 2H), 3.15 (t, $J = 5.9$ Hz, 2H), 3.19 (br. s, 1H), 3.51 (t, $J = 6$ Hz, 2H), 3.78 (s, 6H), 6.86-6.88 (m, 4H), 7.41-7.38 (m, 7H), 7.48 - 7.51 (m, 2H).

^{13}C NMR (100 MHz, Acetone- d_6) δ 14.4, 23.3–32.6, 55.1, 55.5, 55.9, 57.8, 60.1, 63.2, 86.8, 113.8, 127.4, 128.5, 129.0, 130.9, 137.4, 146.5, 159.5.

4.7.3.5 *N,N*-Bis(2-hydroxyethyl)tetradecanamide, **3''**

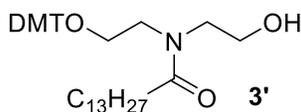


Diethanolamine (2 g, 9.51 mmol, 4 equiv.) was mixed with 10 mL of dry dichloromethane and cooled to -10 °C in a dry round bottom flask. 1.17 g of myristoyl chloride (4.78 mmol) were diluted in 6 mL of dry DCM and added dropwise in the reaction medium under stirring with temperature kept under 0 °C. After the chloride addition, reaction mixture was allowed to slowly come to RT and left under stirring overnight. The mixture was transferred in a separatory funnel. Upper layer was removed and occurred to be the unreacted diethanolamine. Water and DCM were added in the

separatory funnel and the product was extracted three times. The organic fractions were combined, washed with brine, dried over MgSO₄, filtered and the solvent was evaporated under reduced pressure (40 °C). The reaction yielded 1.44 g of a shiny white solid (96 %). Characterization data matched those reported earlier.⁵⁷

¹H NMR (500 MHz, CDCl₃): δ (ppm) = 3.90 (t, *J*=5Hz, 2H), 3.82 (t, *J*=5Hz, 2H), 3.60 (t, *J*=5Hz, 2H), 3.53 (t, *J*=5Hz, 2H), 3.90 (t, *J*=5Hz, 2H), 3.03 (br. s, 1H), 2.79 (br. s, 1H), 2.41 (t, *J*=8Hz, 2H), 1.66 (quin, *J*=7Hz, 2H) 1.28-1.33 (m, 20H), 0.90 (t, *J*=7Hz, 3H).

4.7.3.6 *N*-(2-(Bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)-*N*-(2-hydroxyethyl)tetradecanamide, **3'**.



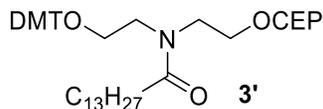
Reaction was performed from 3.63 mmol of **3** as in Section 4.7.3.1. At the end of the reaction, product was extracted twice with DCM from a 10 %Na₂CO₃ aqueous solution, washed once with a 10 %Na₂CO₃ aqueous solution, dried over MgSO₄, filtered and solvent was evaporated under reduced pressure (40 °C). The crude product was loaded on celite and purified using the combiFlash system with a 40g SiO₂ “Gold” column. The solvents used were hexanes/TEA (10:1) and ethyl acetate in a gradient from 0 to 30 % EtOAc in ~17 column volumes (CV). A pale yellow oil was isolated (843 mg, 38 %).

LRMS: Calc.exact mass: 617.41 g/mol. Measured (positive mode): 640.2 (M+23), 303.3 (DMT⁺).

¹H NMR (500 MHz, CDCl₃): δ (ppm) = 7.39-7.37 (m, 2H), 7.30-7.21 (m, 7H), 6.83 (d, *J*=9Hz, 4H), 3.79 (s, 6H), 3.69 (q, *J*=5Hz, 2H), 3.54-3.46 (m, 4H), 3.26 (t, *J*=6Hz, 2H), 2.40 (t, *J*=8Hz, 2H), 1.65-1.58 (m, 2H), 1.30-1.25 (m, 20H), 0.88 (t, *J*=7Hz, 3H).

¹³C NMR (126 MHz, CDCl₃): δ (ppm) = 176.2, 158.8, 135.8, 130.1, 128.2, 128.1, 127.1, 113.3, 113.3, 86.9, 62.9, 61.7, 55.4, 50.6, 49.4, 33.5, 32.1, 29.8, 29.7, 29.7, 29.6, 29.5, 25.4, 22.8, 14.3.

4.7.3.7 2-(*N*-(2-(Bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)tetradecanamido)ethyl (2-cyanoethyl) diisopropylphosphoramidite, **3**'.



Reaction was performed from 0.64 mmol of **3**' as in Section 4.7.3.2. Purification was achieved on a 12g SiO₂ "Gold" column. The solvents used were hexanes/TEA (10:1) and ethyl acetate in a gradient from 0 to 20 % EtOAc in ~10 CV. A clear transparent oil was isolated (366 mg, 70 %).

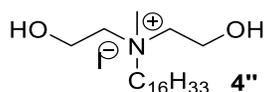
HRMS (ESI-QTOF) *m/z*: [M + Na]⁺ Calcd for C₄₈H₇₂N₃O₆PNa 840.5051; Found 840.5040.

¹H NMR (500 MHz, CDCl₃): δ (ppm) = 7.39-7.37 (m, 2H), 7.29-7.19 (m, 7H), 6.83-6.79 (m, 4H), 3.85-3.68 (m, 11H), 3.60-3.52 (m, 5H), 3.26 (t, *J*=5Hz, 1H), 3.22 (t, *J*=5Hz, 1H), 2.59 (t, *J*=6Hz, 2H), 2.37-2.33 (m, 2H), 1.63-1.57 (m, 2H), 1.32-1.23 (m, 20H), 1.19-1.14 (m, 12H), 0.88 (t, *J*=7Hz, 3H).

¹³C NMR (126 MHz, CDCl₃): δ (ppm) = 173.7, 173.5, 158.7, 158.6, 145.1, 144.8, 136.3, 136.0, 135.9, 130.1, 128.2, 128.2, 128.0, 127.9, 127.0, 126.9, 117.7, 117.7, 113.3, 113.2, 86.8, 86.7, 62.6, 61.8, 61.7, 58.6, 58.5, 58.4, 58.4, 55.4, 55.3, 53.6, 49.1, 47.3, 46.3, 43.3, 43.2, 43.1, 33.5, 33.4, 32.1, 29.8, 29.8, 29.7, 29.5, 25.6, 25.5, 24.8, 24.8, 24.7, 24.7, 22.8, 22.7, 20.5, 20.5, 14.3.

³¹P NMR (203 MHz, CDCl₃): δ (ppm) = 148.4, 147.9.

4.7.3.8 *N,N*-Bis(2-hydroxyethyl)-*N*-methylhexadecan-1-aminium iodide, **4**'.

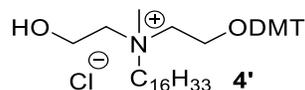


Anhydrous acetonitrile (15 mL) was added to RBF charged with **1**' (1.5 g, 4.55 mmol) and heated to 50°C until material dissolved. CH₃I (840 mg, 5.91 mmol) was added slowly and the reaction mixture was stirred at this temperature for another 1 hour. Hot reaction mixture was added to EtOAc (100 mL) at RT and left to cool down. Shiny off-white crystalline solid was filtered off and dried to obtain the pure product: 1.94g (91 %). Characterization data matched those reported elsewhere.⁵⁸

^1H NMR (600 MHz, CDCl_3) δ (ppm) 0.85-0.88 (m, 3H), 1.24(br. s, 22H), 1.35 (br. s, 4H), 1.76 (br. s, 2H), 3.33 (br. s, 3H), 3.53-3.56(m, 2H), 3.72-3.80 (m, 4H), 4.07 (br. s, 2H), 4.14 (br. s, 4H).

^{13}C NMR (125 MHz, CDCl_3) δ 14.1, 22.6-31.9, 50.8, 55.7, 64.0, 64.5.

4.7.3.9 *N*-(2-(bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)-*N*-(2-hydroxyethyl)-*N*-methylhexadecan-1-aminium chloride.



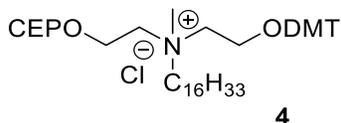
Reaction was performed from 4.2 mmol of **4''** as in Section 4.7.3.1. After 16 hours of reaction, solvents were removed under vacuum to form yellow dense oil which was purified by chromatography on triethylamine pre-treated silica gel with $\text{CH}_2\text{Cl}_2/\text{MeOH}$ with 0.1 % of NEt_3 . Desired material was collected at 3 % MeOH. It was found to contain triethylammonium salts that were removed after a short DCM extraction from 2M NaOH followed by brine wash. Organic fractions were combined, dried over MgSO_4 , filtered and the solvent was evaporated under reduced pressure (40 °C). An orange semi-solid was isolated (24 %).

HRMS (ESI-QTOF) m/z : $[\text{M}]^+$ Calcd for $\text{C}_{42}\text{H}_{64}\text{NO}_4$ 646.4830; Found 646.4845.

^1H NMR (500 MHz, CDCl_3): δ (ppm) = 7.37-7.36 (m, 2H), 7.33-7.22 (m, 7H), 6.86 (d, $J=9\text{Hz}$, 4H), 4.23 (br. s, 1H), 4.09 (br. s, 2H), 3.88-3.77 (m, 8H), 3.63-3.52 (m, 4H), 3.35-3.32 (m, 2H), 3.20 (s, 3H), 1.70-1.63 (m, 2H), 1.31-1.19 (m, 26H), 0.88 (t, $J=7\text{Hz}$, 3H).

^{13}C NMR (126 MHz, CDCl_3): δ (ppm) = 159.0, 143.6, 134.5, 129.9, 128.5, 127.8, 127.5, 113.8, 88.0, 64.5, 64.3, 62.5, 57.3, 55.8, 55.5, 50.8, 32.1, 29.8, 29.8, 29.8, 29.6, 29.5, 29.4, 26.3, 22.9, 22.8, 14.3.

4.7.3.10 *N*-(2-(Bis(4-methoxyphenyl)(phenyl)methoxy)ethyl)-*N*-(2-(((2-cyanoethoxy)(diisopropylamino)phosphanyl)oxy)ethyl)-*N*-methylhexadecan-1-aminium chloride, **4**.



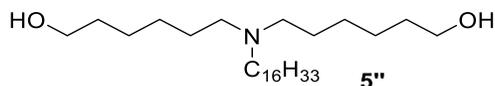
Reaction was performed from 0.57 mmol of **4'** as in Section 4.7.3.2. Purification was achieved with a 24 g neutral Al₂O₃ column. The solvents used were DCM and methanol in a gradient from 0 to 2 % MeOH in ~10 CV. A clear transparent oil was isolated (207 mg, 41 %). NB: silica based chromatography was tried several times without success.

HRMS (ESI-QTOF) *m/z*: [M]⁺ Calcd for C₅₁H₈₁N₃O₅P 846.5908; Found 846.5902.

¹H NMR (500 MHz, CDCl₃): δ (ppm) = 7.37-7.36 (m, 2H), 7.32-7.23 (m, 7H), 6.86 (d, *J*=9Hz, 4H), 4.14-4.05 (m, 2H), 4.01-3.85 (m, 5H), 3.80 (s, 6H), 3.76-3.70 (m, 1H), 3.62 (br, 2H), 3.58-3.52 (m, 2H), 3.46-3.39 (m, 2H), 3.26 (s, 3H), 2.67-2.64 (m, 2H), 1.68 (br, 2H), 1.31-1.22 (m, 26H), 1.17 (dd, *J*=7, 13Hz, 12H), 0.88 (t, *J*=7Hz, 3H).

¹³C NMR (126 MHz, CDCl₃): δ (ppm) = 159.0, 143.7, 134.6, 130.0, 128.3, 127.8, 127.4, 117.8, 113.7, 87.9, 63.5, 63.3, 63.1, 62.3, 62.2, 58.5, 58.3, 58.2, 57.6, 55.4, 49.7, 43.5, 43.4, 32.0, 29.8, 29.8, 29.7, 29.6, 29.6, 29.5, 29.4, 26.4, 24.8, 24.8, 24.7, 24.7, 23.0, 22.8, 20.6, 20.6, 14.2. ³¹P NMR (203 MHz, CDCl₃): δ (ppm) = 149.9, 149.9.

4.7.3.11 **6,6'-(Hexadecylazanediy)bis(hexan-1-ol), 5''**.

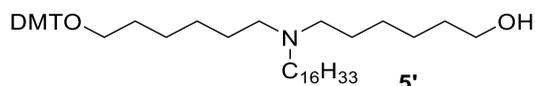


To a solution of hexadecylamine (121 mg, 0.5 mmol) in acetonitrile were added KHCO₃ (200 mg, 2 mmol, 3 equiv.), KI (17 mg, 0.1 mmol, 0.2 equiv.) and 6-bromo-1-hexanol (0.131 mL, 1.0 mmol, 2 equiv.) under stirring. Temperature was raised until a reflux of acetonitrile was reached and left under stirring overnight. The reaction was followed by TLC. After completion, the reaction was worked up with saturated NaHCO₃ solution and DCM. The organic layer was collected, dried over MgSO₄, filtered and solvent was evaporated under reduced pressure (40 °C). 197 mg of a yellow

oil were obtained (89 %). Compound obtained was found to be pure up to 90 % (NMR) and was used as is for the next step.

¹H NMR (500 MHz, CDCl₃): δ (ppm) 3.67 (m, 4H), 2.40 (m, 5H), 1.63-1.23 (m, 45H), 0.91 (t, *J*=5 Hz, 3H).

4.7.3.12 6-((6-(Bis(4-methoxyphenyl)(phenyl)methoxy)hexyl)(hexadecyl)amino)hexan-1-ol, 5'.

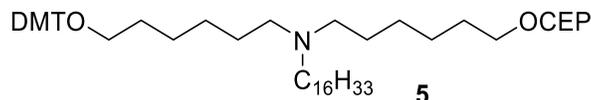


Reaction was performed from 1.54 mmol of 5' as in Section 4.7.3.1. Product was extracted twice with DCM from a 10 % Na₂CO₃ aqueous solution, washed once with a 10 % Na₂CO₃ aqueous solution, dried over MgSO₄, filtered and solvent was evaporated under reduced pressure (40 °C). The crude product was loaded on celite and purified using the combiFlash system with a 4 g SiO₂ Gold column. Hexanes/TEA (10:1) and ethyl acetate were used in a gradient from 0 to 50 % EtOAc. A clear yellow oil was isolated (204 mg, 18 %).

LRMS: Calc.exact mass: 743.6 g/mol. Measured (positive mode): 744.4 (M+1), 303.3 (DMT⁺).

¹H NMR (400 MHz, CDCl₃): δ (ppm) 7.45 (d, *J*=7Hz, 2H), 7.35-7.29 (m, 6H), 7.23 (t, *J*=7Hz, 1H), 6.85 (d, *J*=9Hz, 4H), 3.80 (s, 6H), 3.65 (t, *J*=4Hz, 2H), 3.05 (t, *J*=4Hz, 2H), 2.43-2.35 (m, 6H), 1.67-1.55 (m, 4H), 1.49-1.22 (m, 41H), 0.91 (t, *J*=4Hz, 3H).

4.7.3.13 6-((6-(Bis(4-methoxyphenyl)(phenyl)methoxy)hexyl)(hexadecyl)amino)hexyl (2-cyanoethyl) diisopropylphosphoramidite, 5.



Reaction was performed from 0.20 mmol of 5' as in Section 4.7.3.2. Purification was achieved with a 12 g SiO₂ "Gold" column. The solvents used were hexanes/TEA (10:1) and ethyl acetate in a gradient from 0 to 10 % EtOAc in ~10 CV. A clear transparent oil was isolated (121 mg, 64 %).

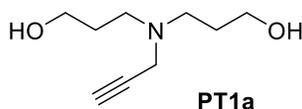
HRMS (ESI-QTOF) *m/z*: [M + H]⁺ Calcd for C₅₈H₉₅O₅N₃P 944.7004; Found 944.6990.

^1H NMR (500 MHz, CDCl_3): δ (ppm) = 7.43-7.44 (m, 2H), 7.33-7.26 (m, 6H), 7.21-7.8 (m, 1H), 6.82 (d, $J=9\text{Hz}$, 4H), 3.87-3.76 (m, 8H), 3.69-3.54 (m, 4H), 3.03 (t, $J=7\text{Hz}$, 2H), 2.62 (t, $J=7\text{Hz}$, 2H), 2.39-2.35 (m, 6H), 1.61 (quin., $J=7\text{Hz}$, 2H), 1.45-1.34 (m, 10H), 1.31-1.22 (m, 32H), 1.18 (t, $J=6\text{Hz}$, 12H), 0.88 (t, $J=7\text{Hz}$, 3H).

^{13}C NMR (126 MHz, CDCl_3): δ (ppm) = 158.4, 145.6, 136.9, 130.1, 128.3, 127.8, 126.6, 117.8, 113.0, 85.7, 63.9, 63.7, 63.6, 58.5, 58.3, 55.3, 54.4, 54.3, 54.3, 43.1, 43.0, 32.0, 31.4, 30.3, 29.8, 29.8, 29.5, 27.8, 27.7, 27.5, 27.2, 27.1, 26.5, 26.1, 24.8, 24.7, 24.7, 24.7, 22.8, 20.5, 20.4, 14.2.

^{31}P NMR (203 MHz, CDCl_3): δ (ppm) = 147.3.

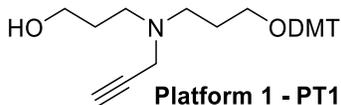
4.7.3.14 3,3'-(Prop-2-yn-1-ylazanediyl)bis(propan-1-ol), PT1a.



To a solution of propargylamine (551 mg, 10.0 mmol) in acetonitrile were added KHCO_3 (3.00 g, 30.0 mmol, 3 equiv.), KI (166 mg, 1 mmol, 0.1 equiv.) and 3-bromo-1-propanol (1.82 mL, 20.00 mmol, 2 equiv.) under stirring. Temperature was raised until a reflux of acetonitrile was reached and left under stirring overnight. The reaction was followed by TLC. After completion, solid was filtered out and solvent was evaporated under reduced pressure (40 °C). Crude obtained was resuspended in DCM, filtered, dried over MgSO_4 , filtered and solvent was evaporated under reduced pressure (40 °C). 1.68 g of a dark brown red oil were obtained. (98 %). Compound obtained was found to be pure up to 90 % (NMR) and was used as is for the next step. NB: in order to get good characterization data, some compound could be further purified using the CombiFlash with a SiO_2 Gold column, (80:20:2 DCM:EtOH: NH_4OH).

^1H NMR (500 MHz, CDCl_3): δ (ppm) 3.76 (t, $J=6\text{Hz}$, 4H), 3.50 (d, $J=2\text{Hz}$, 2H), 2.76 (t, $J=6\text{ Hz}$, 4H), 2.23 (t, $J=2.4\text{ Hz}$, 1H), 1.79 - 1.69 (m, 4H).

4.7.3.15 3-((3-(Bis(4-methoxyphenyl)(phenyl)methoxy)propyl)(prop-2-yn-1-yl)amino)propan-1-ol, PT1.



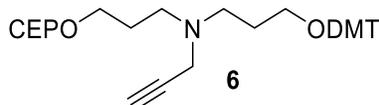
Reaction was performed from 9.8 mmol of **PT1a** as in Section 4.7.3.1. Product was extracted twice with DCM from a 10 % Na₂CO₃ aqueous solution, washed once with a 10 % Na₂CO₃ aqueous solution, dried over MgSO₄, filtered and solvent was evaporated under reduced pressure (40 °C). The crude product was loaded on celite and purified using the combiFlash system with a 80 g SiO₂ Gold column. Hexanes/TEA (100:1) and ethyl acetate were used in a gradient from 0 to 50 % EtOAc. A clear greenish oil was isolated (1.77 g, 38 %).

HRMS (ESI-QTOF) m/z: [M + Na]⁺ Calcd for C₃₀H₃₅O₄NNa 496.2458; Found 496.2448.

¹H NMR (500 MHz, CDCl₃): δ (ppm) 7.45 (d, *J*=7Hz, 2H), 7.35-7.29 (m, 6H), 7.23 (t, *J*=7Hz, 1H), 6.85 (d, *J*=9Hz, 4H), 4.33 (br. s, 1H), 3.82 (s, 6H), 3.77 (t, *J*=5Hz, 2H), 3.47 (s, 2H), 3.13 (t, *J*=6Hz, 2H), 2.75 (t, *J*=6Hz, 2H), 2.64 (t, *J*=7Hz, 2H), 2.23 (s, 1H), 1.80 (quin, *J*=7Hz, 2H), 1.69 (quin, *J*=5Hz, 2H).

¹³C NMR (126 MHz, CDCl₃): δ (ppm) 158.4, 145.2, 136.5, 130.0, 128.2, 127.7, 126.6, 113.0, 85.9, 77.9, 73.3, 64.2, 61.5, 55.2, 53.6, 51.1, 41.7, 28.2, 28.0.

4.7.3.16 3-((3-(Bis(4-methoxyphenyl)(phenyl)methoxy)propyl)(prop-2-yn-1-yl)amino)propyl (2-cyanoethyl) diisopropylphosphoramidite, 6.



Reaction was performed from 0.48 mmol of **PT1** as in Section 4.7.3.2 and purification was achieved with a 12 g SiO₂ Gold column. Hexanes/TEA (10:1) and ethyl acetate were used in a gradient from 0 to 40 % EtOAc in ~10 CV. A clear transparent oil was isolated (249 mg, 70 %).

HRMS (ESI-QTOF) m/z: [M + K]⁺ Calcd for C₃₉H₅₂N₃O₅PK 712.3276; Found 712.3296.

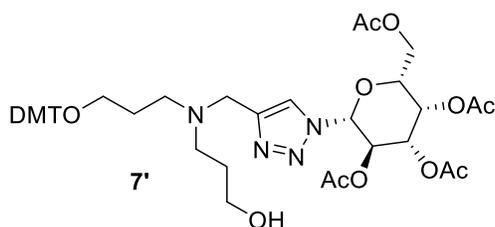
¹H NMR (500 MHz, CDCl₃): δ (ppm) 7.43 (d, *J*=8Hz, 2H), 7.33-7.26 (m, 6H), 7.20 (t, *J*=9Hz, 1H), 6.82 (d, *J*=9Hz, 4H), 3.87-3.74 (m, 2H), 3.79 (s, 6H), 3.71-3.55 (m, 4H), 3.37 (d, *J*=2Hz, 2H),

3.09 (t, $J=6\text{Hz}$, 2H), 2.61-2.54 (m, 6H), 2.14 (t, $J=2\text{Hz}$, 1H), 1.74 (sex, $J=7\text{Hz}$, 4H), 1.17 (dd, $J=7$, 9Hz, 12H).

^{13}C NMR (126 MHz, CDCl_3): δ (ppm) 158.3, 145.3, 136.6, 130.0, 128.2, 127.7, 126.6, 113.0, 85.9, 78.7, 72.7, 61.9, 61.8, 61.6, 58.4, 58.2, 55.2, 50.8, 50.2, 43.1, 43.0, 41.9, 29.3, 29.3, 28.3, 24.7, 24.6, 24.6, 20.4, 20.3.

^{31}P NMR (203 MHz, CDCl_3): δ (ppm) 147.5.

4.7.3.17 (2S,3R,4R,5S,6S)-2-(Acetoxymethyl)-6-(4-(((3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)(3-hydroxypropyl)amino)methyl)-1H-1,2,3-triazol-1-yl)tetrahydro-2H-pyran-3,4,5-triyl triacetate, 7'.



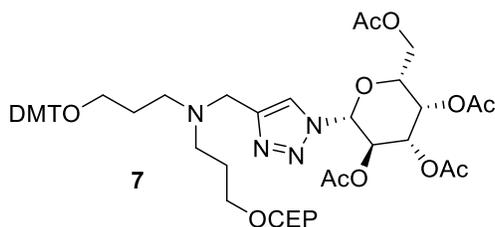
Adapted from an earlier report.⁵⁹ 1-Azido-1-deoxy- β -D-glucopyranoside tetraacetate (379 mg, 0.80 mmol) and platform 1 (300 mg, 0.80 mmol, 1 equiv.) were suspended in 1 mL CHCl_3 and a tBuOH/water (1:1) mixture (12 mL) was added. Freshly prepared solution of sodium ascorbate (12.1 mg, 0.061 mmol, 0.2equiv.) in 0.2 mL of water and another of copper sulfate (7.6 mg, 0.030 mmol, 0.1 equiv.) in 0.1 mL of water were sequentially added. Reaction was allowed to stir overnight at room temperature. Product was extracted twice with DCM from 10 % Na_2CO_3 aqueous solution, washed once with a 10 % Na_2CO_3 aqueous solution, dried over MgSO_4 , filtered and solvent was evaporated under reduced pressure (40 °C). The crude obtained was resuspended in toluene to evaporate remaining tBuOH. 643 mg of a pale golden powder were obtained (95 %).

LRMS: Calc.exact mass: 846.4 g/mol. Measured (positive mode): 869.3 (M+23), 303.3 (DMT⁺).

^1H NMR (500 MHz, CDCl_3): δ (ppm) 7.69 (s, 1H), 7.41 (d, $J=7\text{Hz}$, 2H), 7.32-7.26 (m, 6H), 7.21-7.19 (m, 1H), 6.82 (d, $J=9\text{Hz}$, 4H), 5.83-5.79 (m, 1H), 5.43-5.37 (m, 2H), 5.28-5.22 (m, 1H), 4.30 (dd, $J=5$, 13Hz, 1H), 4.14 (dd, $J=2$, 13Hz, 1H), 3.99-3.95 (m, 1H), 3.82-3.75 (m, 2H), 3.79 (s, 6H), 3.71 (t, $J=5\text{Hz}$, 2H), 3.09 (t, $J=6\text{Hz}$, 2H), 2.68-2.60 (m, 2H), 2.60-2.51 (m, 2H), 2.07 (s, 6H), 2.03 (s, 3H), 1.83 (quin, $J=6\text{Hz}$, 2H), 1.79 (s, 3H), 1.75-1.63 (m, 2H).

^{13}C NMR (126 MHz, CDCl_3): δ (ppm) 170.7, 170.0, 169.5, 169.0, 158.5, 145.6, 145.4, 136.6, 130.2, 128.3, 127.9, 126.8, 121.1, 113.2, 86.0, 75.4, 72.7, 70.5, 67.9, 64.0, 61.6, 55.3, 53.6, 51.1, 48.7, 28.3, 27.8, 20.8, 20.7, 20.7, 20.2.

4.7.3.18 (2S,3R,4R,5S,6S)-2-(Acetoxymethyl)-6-(4-(((3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)(3-(((2-cyanoethoxy)(diisopropylamino)phosphanyl)oxy)propyl)amino)methyl)-1H-1,2,3-triazol-1-yl)tetrahydro-2H-pyran-3,4,5-triyl triacetate, 7.



Reaction was performed from 0.22 mmol of **7'** as in Section 4.7.3.2. Purification was achieved with a 12 g SiO_2 "Gold" column. Hexanes/TEA (10:1) and ethyl acetate were used in a gradient from 0 to 50 % EtOAc in ~10 CV. A white crystalline powder was isolated (162 mg, 70 %).

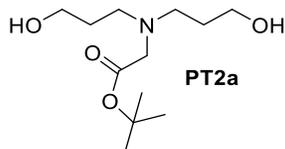
HRMS (ESI-QTOF) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{53}\text{H}_{72}\text{N}_6\text{O}_{14}\text{P}$ 1047.4839; Found 1047.4866.

^1H NMR (500 MHz, CDCl_3): δ (ppm) 7.60 (s, 1H), 7.42 (d, $J=8\text{Hz}$, 2H), 7.32-7.26 (m, 6H), 7.22-7.18 (m, 1H), 6.82 (d, $J=9\text{Hz}$, 4H), 5.82-5.77 (m, 1H), 5.44-5.37 (m, 2H), 5.27-5.23 (m, 1H), 4.30 (dd, $J=5, 13\text{Hz}$, 1H), 4.12 (dd, $J=2, 13\text{Hz}$, 1H), 3.99-3.95 (m, 1H), 3.85-3.72 (m, 10H), 3.70-3.63 (m, 1H), 3.62-6.54 (m, 2H), 3.08 (t, $J=6\text{Hz}$, 2H), 2.59 (t, $J=7\text{Hz}$, 2H), 2.53-2.48 (m, 4H), 2.07 (s, 3H), 2.06 (s, 3H), 2.03 (s, 3H), 1.82-1.73 (m, 7H), 1.16 (dd, $J=7, 10\text{Hz}$, 12H).

^{13}C NMR (126 MHz, CDCl_3): δ (ppm) 170.7, 170.1, 169.5, 168.8, 158.5, 146.3, 145.5, 136.7, 130.2, 128.3, 127.9, 126.7, 120.9, 117.9, 113.1, 85.9, 85.9, 75.3, 72.9, 70.4, 67.9, 62.1, 62.0, 61.7, 58.5, 58.4, 55.4, 50.8, 50.4, 48.7, 43.2, 43.1, 29.1, 28.0, 24.8, 24.8, 24.8, 24.7, 20.8, 20.7, 20.7, 20.5, 20.5, 20.2.

^{31}P NMR (203 MHz, CDCl_3): δ (ppm) 147.3, 147.3 (two diastereoisomers).

4.7.3.19 Tert-butyl bis(3-hydroxypropyl)glycinate, PT2a.

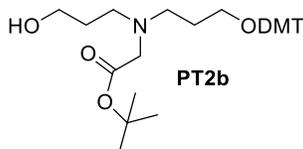


To a solution of glycine tert-butyl ester hydrochloride (2.57 g, 15.3 mmol) in acetonitrile were added KHCO_3 (4.60 g, 45.9 mmol, 3 equiv.), KI (254mg, 1.53 mmol, 0.1equiv.) and 3-bromo-1-propanol (2.77 mL, 30.6 mmol, 2 equiv.) under stirring. Temperature was raised until a reflux of acetonitrile was reached and left under stirring overnight. The reaction was followed by TLC. After completion, solvent was evaporated under reduced pressure (40 °C). Product was extracted twice with ethyl acetate from 10 % Na_2CO_3 aqueous solution, washed once with a 10 % Na_2CO_3 aqueous solution, dried over MgSO_4 , filtered and solvent was evaporated under reduced pressure (40 °C). Compound obtained was a pale yellow oil found to be pure up to 90 % (NMR) and was used as is for the next step (3.56 g, 94 %).

LRMS Calc.exact mass: 247.2 g/mol. Measured (positive mode):248.1 (M+1), 270.1 (M+23).

^1H NMR (500 MHz, CDCl_3): δ (ppm) 3.77 (t, $J=6\text{Hz}$, 4H), 3.18 (s, 2H), 2.36 (t, $J=6\text{Hz}$, 4H), 1.72 (quin., $J=6\text{Hz}$, 4H), 1.47 (s, 9H).

4.7.3.20 Tert-butyl *N*-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-*N*-(3-hydroxypropyl)glycinate, PT2b.



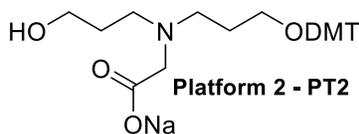
Reaction was performed from 14.4 mmol of **PT2a** as in Section 4.7.3.1. Product was extracted twice with DCM from a 10 % Na_2CO_3 aqueous solution, washed once with a 10 % Na_2CO_3 aqueous solution, dried over MgSO_4 , filtered and solvent was evaporated under reduced pressure (40 °C). The crude product was loaded on celite and purified using the combiFlash system with a 220 g SiO_2 Gold column. Hexanes/TEA (10:1) and ethyl acetate were used in a gradient from 0 to 35 % EtOAc(~25 CV). A clear yellow oil was isolated (2.96 g, 37 %).

HRMS (ESI-QTOF) m/z : $[\text{M} + \text{Na}]^+$ Calcd for $\text{C}_{33}\text{H}_{43}\text{NO}_6\text{Na}$ 572.2983; Found 572.2978.

^1H NMR (500 MHz, CDCl_3): δ (ppm) 7.42 (d, $J=7\text{Hz}$, 2H), 7.32-7.26 (m, 6H), 7.21-7.18 (m 1H), 6.82 (d, $J=9\text{Hz}$, 4H), 4.54 (br. s, 1H), 3.79 (s, 6H), 3.74 (t, $J=5\text{Hz}$, 2H), 3.19 (s, 2H), 3.09 (t, $J=6\text{Hz}$, 2H), 2.70 (t, $J=6\text{Hz}$, 2H), 2.64-2.61 (m, 2H), 1.79 (quin, $J=8\text{Hz}$, 2H), 1.65 (quin, $J=5\text{Hz}$, 2H), 1.46 (s, 9H).

^{13}C NMR (126 MHz, CDCl_3): δ (ppm) 170.7, 158.5, 145.3, 136.6, 130.1, 128.3, 127.9, 126.8, 113.2, 86.0, 81.5, 63.5, 61.8, 56.6, 55.3, 54.0, 51.8, 28.6, 28.2, 27.9.

4.7.3.21 Sodium *N*-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-*N*-(3-hydroxypropyl)glycinate, **PT2**.



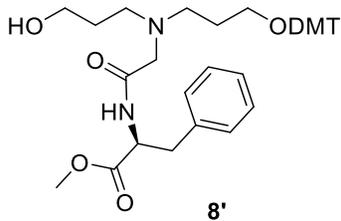
Compound **PT2b** (260 mg, 0.47 mmol) was dissolved in about 2 mL of methanol. To the mixture was added 25 mL of 0.4M NaOH in MeOH/water 4:1. The reaction mixture was left under stirring for 3h at 65 °C. Reaction was monitored by TLC. When higher mobility spot disappeared, methanol was evaporated under reduced pressure (60 °C) until a precipitate appears but a small amount of water remains. Precipitate was filtered and quickly washed four times with cold water. The solid was suspended in MeOH, solvent was evaporated under reduced pressure (40 °C). The obtained solid was resuspended in DCM, dried over MgSO_4 , solution was filtered and solvent was evaporated under reduced pressure (40 °C) to obtain a white to pale yellow solid (206 mg, 85 %). NB: An alternate protocol involved a work-up with 2 equivalents of tetrabutylammonium chloride and sodium carbonate solutions. The obtained tetrabutylammonium salt started to degrade after a few days under high vacuum.

HRMS (ESI-QTOF) m/z : $[\text{M}-\text{Na}]^-$ Calcd for $\text{C}_{29}\text{H}_{34}\text{O}_6\text{N}$ 492.2392; Found 492.2405.

^1H NMR (400 MHz, $\text{DMSO}-d_6$): δ (ppm) 7.36-7.28 (m, 4H), 7.24-7.19 (m, 5H), 6.88 (d, $J=9\text{Hz}$, 4H), 3.73 (s, 6H), 3.37 (t, $J=6\text{Hz}$, 2H), 2.96 (t, $J=6\text{Hz}$, 2H), 2.71 (s, 2H), 2.43-2.37 (m, 4H), 1.61 (quin, $J=7\text{Hz}$, 2H), 1.38 (quin, $J=6\text{Hz}$, 2H).

^{13}C NMR (126 MHz, $\text{DMSO}-d_6$): δ (ppm) 174.4, 157.9, 145.3, 136.1, 129.6, 127.8, 127.6, 126.5, 113.1, 85.2, 61.6, 59.3, 58.2, 55.0, 50.5, 50.1, 29.5, 27.3.

4.7.3.22 Methyl *N*-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-*N*-(3-hydroxypropyl)glycyl-L-phenylalaninate, 8'.



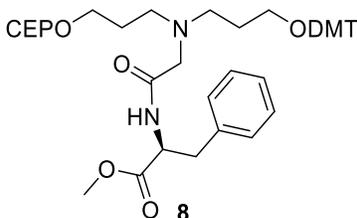
To a solution of **PT2** (206 mg, 0.40 mmol) in DMF (5 mL) were successively added anhydrous 1-hydroxybenzotriazole (HOBt) (81 mg, 0.60 mmol, 1.5 equiv.) and *N*-(3-Dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC-Cl) (115 mg, 0.60 mmol, 1.5 equiv.). The pale yellow solution was left under stirring for 5 minutes until complete dissolution of EDC-Cl. Phenylalanine methyl ester hydrochloride (95 mg, 0.44 mmol, 1.1 equiv.) and triethylamine (0.15 mL, 1.04 mmol, 2.6 equiv.) were then successively added. The reaction mixture slowly turned cloudy and was left under vigorous stirring overnight. Solvent was then evaporated under reduced pressure (60 °C). The crude product was loaded on celite and purified using the combiFlash system with a 12 g SiO₂ “Gold” column pre-treated with 1 % TEA in hexanes. Hexanes/TEA (10:1) and ethyl acetate were used in a gradient from 0 to 70 % EtOAc in ~12 CV. A clear pale yellow oil was isolated (173mg, 66 %). NB: Interestingly, when the crude in DMF is analyzed using RP-HPLC, conversion yields reached 90 % but isolated yields never exceeded 66 %. This may be due to issues with column chromatography or to the presence of undesired salts at the end of the previous step.

LRMS: Calc. exact mass: 654.33 g/mol. Measured (positive mode): 303.3 (DMT⁺), 677.3 (M+23).

¹H NMR (500 MHz, CDCl₃): δ (ppm) 7.60 (d, *J*=8Hz, 1H), 7.41 (d, *J*=7Hz, 2H), 7.31-7.18 (m, 10H), 7.08 (d, *J*=7Hz, 2H), 6.83 (d, *J*=9Hz, 4H), 4.85 (q, *J*=8Hz, 1H), 3.79 (s, 6H), 3.71 (s, 3H), 3.53 (t, *J*=6Hz, 2H), 3.15-2.94 (m, 6H), 2.61-2.53 (m, 2H), 2.49-2.43 (m, 2H), 1.68-1.61 (m, 2H), 1.60-1.55 (m, 2H).

¹³C NMR (126 MHz, CDCl₃): δ (ppm) 172.5, 171.5, 158.5, 145.3, 136.6, 136.2, 130.1, 129.3, 128.7, 128.3, 127.9, 127.2, 126.8, 113.2, 86.0, 61.5, 60.8, 58.8, 55.3, 52.7, 52.5, 52.4, 52.0, 37.9, 30.2, 28.03.

4.7.3.23 Methyl *N*-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-*N*-(3-(((2-cyanoethoxy)(diisopropylamino)phosphanyl)oxy)propyl)glycyl-L-phenylalaninate, **8.**



Reaction was performed from 0.333 mol of **8'** as in Section 4.7.3.2. Purification was achieved with a 12 g SiO₂ "Gold" column. Hexanes/TEA (10:1) and ethyl acetate were used in a gradient from 0 to 40 % EtOAc in ~12 CV. A clear transparent oil was isolated (195 mg, 68 %).

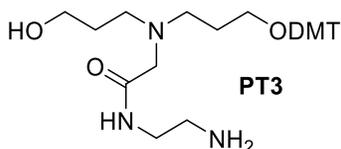
HRMS (ESI-QTOF) *m/z*: [M + Na]⁺ Calcd for C₄₈H₆₃N₄O₈PNa 877.4276; Found 877.4257.

¹H NMR (500 MHz, CDCl₃): δ (ppm) 7.64 (d, *J*=8Hz, 1H), 7.40 (d, *J*=7Hz, 2H), 7.31-7.15 (m, 10H), 7.07 (d, *J*=7Hz, 2H), 6.82 (d, *J*=9Hz, 4H), 4.83 (q, *J*=8Hz, 1H), 3.84-3.76 (m, 8H), 3.68 (s, 3H), 3.60-3.50 (m, 4H), 3.16-3.12 (m, 1H), 3.05-2.97 (m, 5H), 2.58 (td, *J*=2, 7Hz, 2H), 2.55-2.47 (m, 4H), 1.67-1.59 (m, 4H), 1.17 (dd, *J*=7, 9Hz, 12H).

¹³C NMR (126 MHz, CDCl₃): δ (ppm) 171.9, 11.5, 158.2, 145.3, 136.6, 136.2, 130.1, 129.3, 128.7, 128.3, 127.9, 127.2, 126.8, 113.2, 86.0, 61.5, 58.5, 55.3, 53.0, 52.7, 52.5, 52.3, 43.2, 43.1, 38.0, 29.3, 28.2, 24.8, 24.7, 20.5, 20.5.

³¹P NMR (203 MHz, CDCl₃): δ (ppm) 147.6.

4.7.3.24 *N*-(2-Aminoethyl)-2-(((3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)(3-hydroxypropyl)amino)acetamide (PT3)



In a round bottom flask, 350 mg (0.64 mmol) of **PT2b** were dissolved in 5 mL of ethylenediamine. Solution was left under vigorous stirring under reflux for 48h. Product was extracted twice with DCM from 10 % Na₂CO₃ aqueous solution, washed once with a saturated Na₂CO₃ aqueous

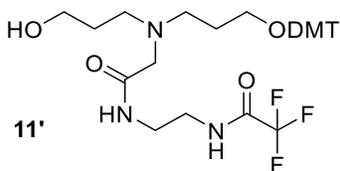
solution, dried over MgSO₄, filtered and solvent was evaporated under reduced pressure (40 °C) to obtain a yellow oil (299 mg, 88 %).

HRMS (ESI-QTOF) *m/z*: [M + Na]⁺ Calcd for C₃₁H₄₁N₃O₅Na 558.29384; Found 558.29354.

¹H NMR (500 MHz, CDCl₃): δ (ppm) 7.54 (t, *J*=6Hz, 1H), 7.42-7.40 (m, 2H), 7.31-7.26 (m, 6H), 7.22-7.20 (m, 1H), 6.82 (d, *J*=9Hz, 4H), 3.79 (s, 6H), 3.62 (t, *J*=6Hz, 2H), 3.25 (q, *J*=5Hz, 2H), 3.09 (t, *J*=6Hz, 2H), 3.04 (s, 2H), 2.78 (t, *J*=6Hz, 2H), 2.57 (q, *J*=6Hz, 4H), 1.79-1.71 (m, 2H), 1.67-1.62 (m, 2H).

¹³C NMR (126 MHz, CDCl₃): δ (ppm) 171.9, 158.5, 145.3, 136.5, 130.1, 128.2, 127.9, 126.8, 113.2, 86.1, 61.6, 59.5, 59.2, 55.4, 52.3, 51.1, 41.5, 41.1, 59.7, 27.8.

4.7.3.25 *N*-(2-(2-((3-(Bis(4-methoxyphenyl)(phenyl)methoxy)propyl)(3-hydroxypropyl)amino)acetamido)ethyl)-2,2,2-trifluoroacetamide (11')



In a round bottom flask, 648 mg (1.21 mmol, 1 equiv.) of **PT3** were dissolved in 10 mL of DCM and 1 mL of pyridine. Ethyl trifluoroacetate (0.15 mL, 1.3 mmol, 1.05 equiv.) was slowly added under vigorous stirring. Reaction was left under stirring overnight. Product was extracted twice with DCM from 10 % Na₂CO₃ aqueous solution, washed once with a 10 % Na₂CO₃ aqueous solution. Organic fractions were combined, dried over MgSO₄, filtered and the solvent was evaporated under reduced pressure (40 °C). The crude product was loaded on celite and purified using the combiFlash system with a SiO₂ “Gold” column pretreated with 0.1 % TEA in DCM. DCM/TEA (100:0.1) and DCM/Methanol/TEA (90:10:0.1) were used in a gradient. A clear yellow wax was isolated (485 mg, 0.768 mmol, 63 %).

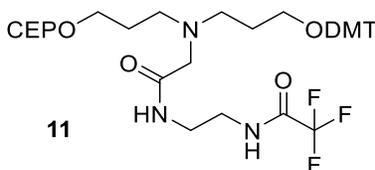
LRMS: Calc.exact mass: 631.29 g/mol. Measured (negative mode): 630.24 (M-1),

¹H NMR (500 MHz, DMSO-*d*₆): δ (ppm) 9.38 (s, 1H), 7.76 (s, 1H), 7.37-7.35 (m, 2H), 7.31-7.28 (m, 2H), 7.24-7.19 (m, 5H), 6.87 (d, *J*=9Hz, 4H), 4.43 (s, 1H), 3.73 (s, 6H), 3.40 (t, *J*=6Hz, 2H), 3.22 (s, 4H), 2.96 (t, *J*=6Hz, 2H), 2.91 (s, 2H), 2.47-2.44 (m, 4H), 1.69-1.64 (m, 2H), 1.54-1.48 (m, 2H).

^{13}C NMR (126 MHz, DMSO-*d*6): δ (ppm) 171.2, 158.0, 156.6, 156.3, 145.2, 136.0, 129.6, 127.8, 127.6, 126.6, 117.0, 114.7, 113.1, 85.2, 61.2, 58.9, 57.8, 55.0, 51.5, 51.5, 37.2, 29.8, 27.1.

^{19}F NMR (471 MHz, DMSO-*d*6): δ (ppm) -74.5.

4.7.3.26 3-((3-(Bis(4-methoxyphenyl)(phenyl)methoxy)propyl)(2-oxo-2-((2-(2,2,2-trifluoroacetamido)ethyl)amino)ethyl)amino)propyl (2-cyanoethyl) diisopropylphosphoramidite (11)



Reaction was performed from 0.744 mol of **11'** as in Section 4.7.3.2. The crude product was loaded on celite and purified using the combiFlash system with a SiO₂ "Gold" column pretreated with 1 % TEA in hexanes. Hexanes/TEA (100:1) and ethyl acetate/TEA (100/1) were used in a gradient. A white waxy solid was isolated (396 mg, 0.476 mmol, 64 %).

HRMS (APCI-QTOF) *m/z*: [M + H]⁺ Calcd for C₄₂H₅₈N₅O₇PF₃ 832.40205; Found 832.40168.

^1H NMR (500 MHz, CDCl₃): δ (ppm) 8.18 (s, 1H), 7.63-7.61 (m, 1H), 7.41-7.39 (m, 2H), 7.29-7.26 (m, 6H), 7.22-7.19 (m, 1H), 6.82 (d, *J*=9Hz, 4H), 3.85-3.81 (m, 1H), 3.79 (s, 6H), 3.75-3.55 (m, 5H), 3.37-3.34 (m, 4H), 3.08 (t, *J*=6Hz, 2H), 3.05 (s, 2H), 2.61-2.58 (m, 6H), 1.75-1.71 (m, 4H), 1.17 (t, *J*=7Hz, 12H).

^{13}C NMR (126 MHz, CDCl₃): δ (ppm) 174.6, 158.6, 157.8, 157.5, 145.2, 136.4, 130.1, 138.2, 127.9, 126.9, 117.9, 117.1, 114.8, 113.2, 86.1, 61.8, 61.7, 61.4, 58.4, 58.2, 58.1, 55.4, 52.4, 52.3, 43.2, 43.1, 42.2, 38.3, 28.9, 28.9, 27.8, 24.8, 24.8, 24.7, 24.7, 20.6, 20.6.

^{19}F NMR (471 MHz, CDCl₃): δ (ppm) -76.0.

^{31}P NMR (203 MHz, CDCl₃): δ (ppm) 147.6.

4.7.4 Solid-phase Synthesis

Standard DNA synthesis was performed on a 1 μmol scale, starting from a universal 1000 Å LCAA-CPG solid support. Amidites **HEG** and **TEG** were dissolved in dry acetonitrile and all

other amidites were dissolved in dry DCM to obtain 0.1M solutions. Extended coupling times of 10 minutes were used except for standard DNA amidites and for oligomer **A** synthesis (4min couplings). Removal of the DMT protecting group was carried out using 3 % dichloroacetic acid in dichloromethane on the DNA synthesizer. The DNA sequence used along this study is called AT and is: 5'- TTTTTCAGTTGACCATATA – 3' except for the oligomer **E** for which the sequence used was 5'- CAGTTGACCATATA – 3' and the oligomer bearing **11** for which it is 5'- TA-**11**-TTTTTCAGTTGACCATATA – 3'.

4.7.4.1 General procedure for attaching moiety without phosphoramidite isolation.

Under a nitrogen atmosphere in a glove box (<2.5 ppm trace moisture), in a 20 mL oven-dried round bottom flask, monoprotected alcohol **1'** (31.6 mg, 0.050 mmol) is dissolved in dry DCM (500 μ l). Diisopropylethylamine (8.7 μ l, 0.050 mmol, 1 equiv.) and N,N-Diisopropylamino cyanoethyl phosphoramidic-Cl (10.0 μ l, 0.045 mmol, 0.9 equiv.) are added. Reaction is allowed to stir at room temperature during 45 minutes. Coupling was done using the 'syringe' technique: the crude solution containing the phosphoramidite (200 μ l, 0.1 M) is mixed with the standard activator solution (200 μ l, 0.25 M) in presence of the CPG using syringes. After ten minutes, the solution was removed from the columns and the strands underwent capping, oxidation and deblocking steps in the synthesizer.

4.7.4.2 General deprotection procedure.

Sequences without moiety **PT2b** or **8** underwent classical deprotection procedures: completed syntheses were cleaved from the solid support and deprotected in 28 % aqueous ammonium hydroxide solution for 16-18 hours at 65 °C or in 28 % aqueous ammonium hydroxide/30 % methylamine solution 1:1 mixture for 3h hours at 65 °C. With moiety **8**, 1:3 tert-butylamine/water solution during 6h at 65 °C (recommended with dmf-protected guanosines) or 0.4 M NaOH 1:4 water/methanol solution at room temperature during 16h, followed by quenching with 2.0 M TEAA buffer (recommended with isobutyryl-protected guanosines) were used. With moiety **PT2b**, a solution of 0.4 M NaOH 1:4 water/methanol solution at 65 °C overnight, followed by quenching with 2.0 M TEAA buffer could fully deprotect the tBu ester. The crude product solution was separated from the solid support and concentrated under reduced pressure at 60 °C. This crude solid was re-suspended in 1 mL Millipore water. Filtration with 0.22 μ m centrifugal filter was then performed prior to HPLC purification. The resulting solution was quantified by absorbance at 260

nm. NB: if 1:3 tert-butylamine/water solution during 6 h at 65 °C does not lead to complete unylinker deprotection, concentrated ammonium hydroxide at room temperature overnight could be performed as a second step.

4.7.4.3 General quantification procedure.

To quantify the oligomers in solution, absorbance at 260 nm was measured using a Nanodrop Lite spectrophotometer. Extinction coefficients were calculated according to the following formula: $\epsilon = \epsilon_{\text{DNA}} + \epsilon_{\text{monomer}} \times \text{number of } \mathbf{9} \text{ in the sequence} + \epsilon_{\text{isolated thymidine}} \times \text{number of thymidine in the sequence}$. With ϵ_{DNA} being the extinction coefficient of the DNA sequence in the oligomer (if applicable, obtained from IDT oligoanalyzer tool), $\epsilon_{\text{monomer } \mathbf{9}} = 2.80 \text{ mM}^{-1} \cdot \text{cm}^{-1}$ as reported in Chapter 3, $\epsilon_{\text{isolated thymidine}} = 8.56 \text{ mM}^{-1} \cdot \text{cm}^{-1}$.⁶⁰

4.7.5 Post solid-phase synthesis functionalization of alkyne containing DNA strands

Two strands containing one and 5 times phosphoramidite **6** have been modified after solid-phase synthesis following a standard protocol reported elsewhere.⁵⁰

With DNA-6: To 20 μL of a DNA solution (directly from post-deprotection crude 6.6 nmol) in water, 10 μL of an azide solution (164 nmol, 25 equiv.) and 10 μL of a freshly prepared solution containing CuBr and tris(benzyltriazolylmethyl)-amine in a 1:1 ration in 4:3:1 water : DMSO : tBuOH was added (66 nmol, 10 equiv.). The mixture was vortexed and shaken at room temperature for 2 hours.

With DNA-65 (oligomer F): To 20 μL of a DNA solution (directly from post-deprotection crude, 9.5 nmol) in water 10 μL of an azide solution (475 nmol, 50 equiv.) and 10 μL of a freshly prepared solution containing CuBr and tris(benzyltriazolylmethyl)-amine in a 1:1 ration in 4:3:1 water : DMSO : tBuOH was added (180 nmol, 20 equiv.). The mixture was vortexed and shaken at room temperature for 2 hours.

In the case of post-SPS click chemistry compounds, purification was carried out through gel electrophoresis instead of RP-HPLC. Crude products were purified on 20 % polyacrylamide gels, supplemented with 7M urea (loading up to 10 OD260 of crude mixture per gel, 500 V field applied). Electrophoresis was run at lower voltage for the first 30 minutes. Following electrophoresis, the gel was wrapped in plastic and visualized by UV shadowing over a fluorescent TLC plate. The full-length product was quickly excised, then crushed and incubated in ~10 mL of

Millipore water. The solution was frozen in liquid N₂ for 3 minutes and left at 65 °C overnight. The supernatant was then concentrated to 1.0 mL, and desalted using size exclusion chromatography (Sephadex G-25). The DNA strand was then quantified (OD260) and converted to micromolar concentrations using the calculated extinction coefficient.

4.7.6 Gel electrophoresis

18 % denaturing Polyacrylamide Gel Electrophoresis (PAGE) was carried out at room temperature for 30 minutes at 250V followed by 1 hour at 500V. TBE buffer (1X) was used and the concentration of urea in the gel was 7 M. For each lane 5 µL of crude mixture (1.6 µM) in water was added to 5ul of 8 M urea. Only unmodified AT strand was purified beforehand. The DNA bands for all gels were visualized by incubation with GelRed™. Yields of conjugation were measured using the software Image Lab and the formula: yields =lowest mobility band intensity/ n-1 band intensity.

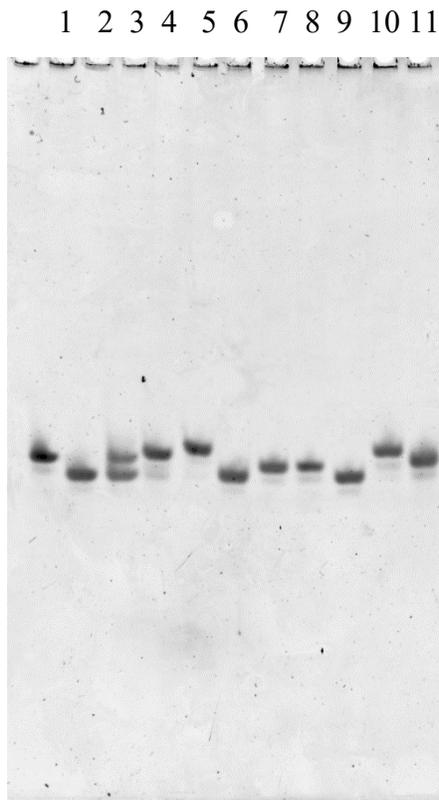


Figure 4.7. 18 % PAGE in denaturing conditions, of modified DNA strands. Lanes 1 to 11: DNA-9, DNA, DNA-3, DNA-4, DNA-5, DNA, DNA-6, DNA-PT2b, DNA, DNA-7, DNA-8.

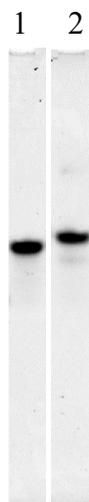


Figure 4.8. 15 % PAGE in denaturing conditions of 21mer modified with **11**. (Lane 2) and control: 21mer without **11** (Lane 1).

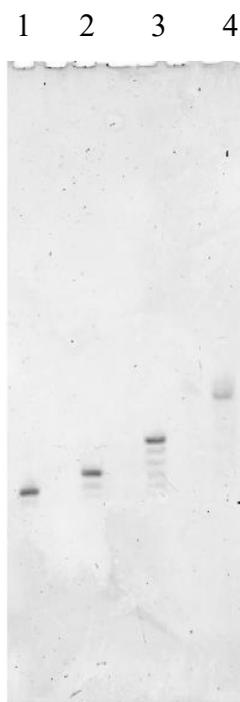


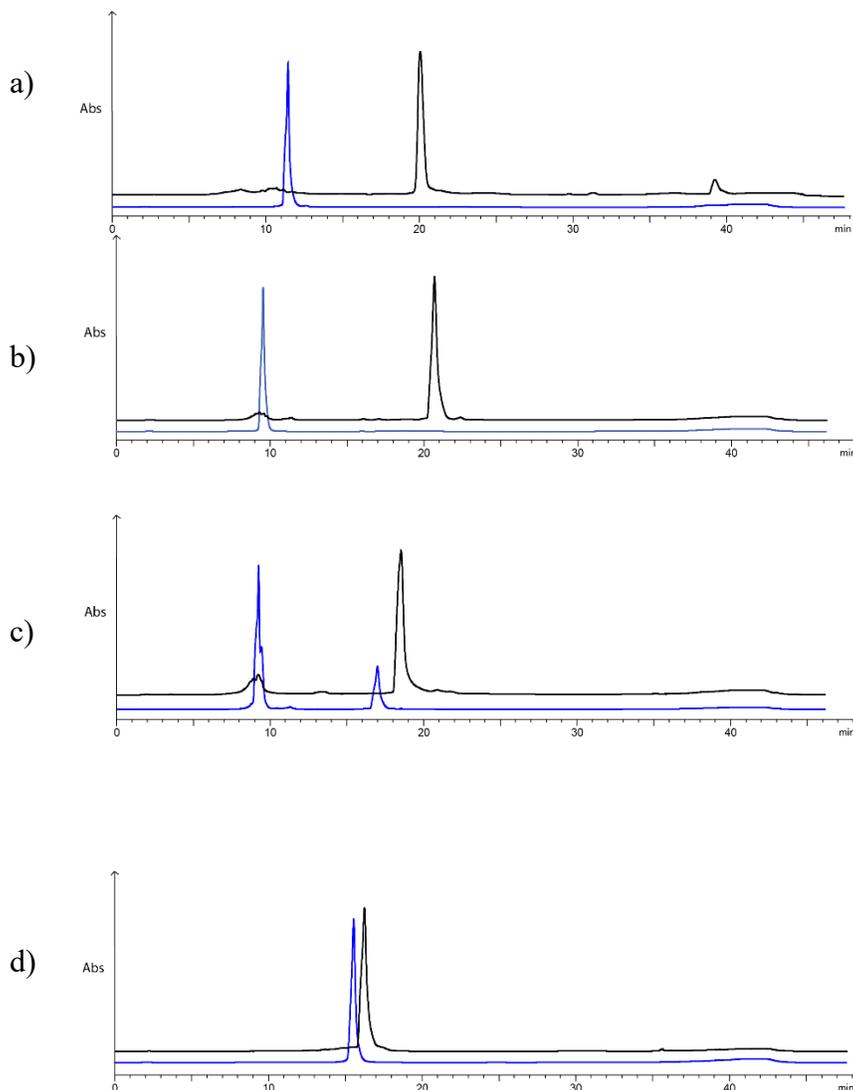
Figure 4.9. 18 % PAGE in denaturing conditions of modified DNA strands with post SPS click chemistry. Strand modified with 1 or 5 times phosphoramidite **6** (lanes 1 and 3 respectively) and associated crude mixtures after click chemistry (lanes 2 and 4 respectively).

For post-SPS click reactions, yields were calculated with the formula: yields = lowest mobility band intensity/ intensity of higher mobility bands. They were found to be higher than 80 % in both cases. The lowest mobility bands could be isolated following standard gel purification procedure. Identity of the products were identified by ESI-MS (**Figure 4.12**).

NB: Oligomer F after click chemistry and PAGE purification needed to be further purified by RP-HPLC to obtain good data by MS.

4.7.7 RP-HPLC purification and analysis.

Solvents (0.22 μm filtered): 50 mM triethylammonium acetate (TEAA) buffer (pH 8.0) and HPLC grade acetonitrile. All gradients were followed by a short column wash in 95 % acetonitrile. Column: Hamilton PRP-C18 5 μm 100 \AA 2.1 x 150 mm. For each analytical separation approximately 0.5 OD₂₆₀ of crude DNA or 0.3 OD₂₆₀ of crude oligomers was injected as a 20-50 μL solution in Millipore water. Detection was carried out using a diode-array detector, monitoring absorbance at 260 nm.



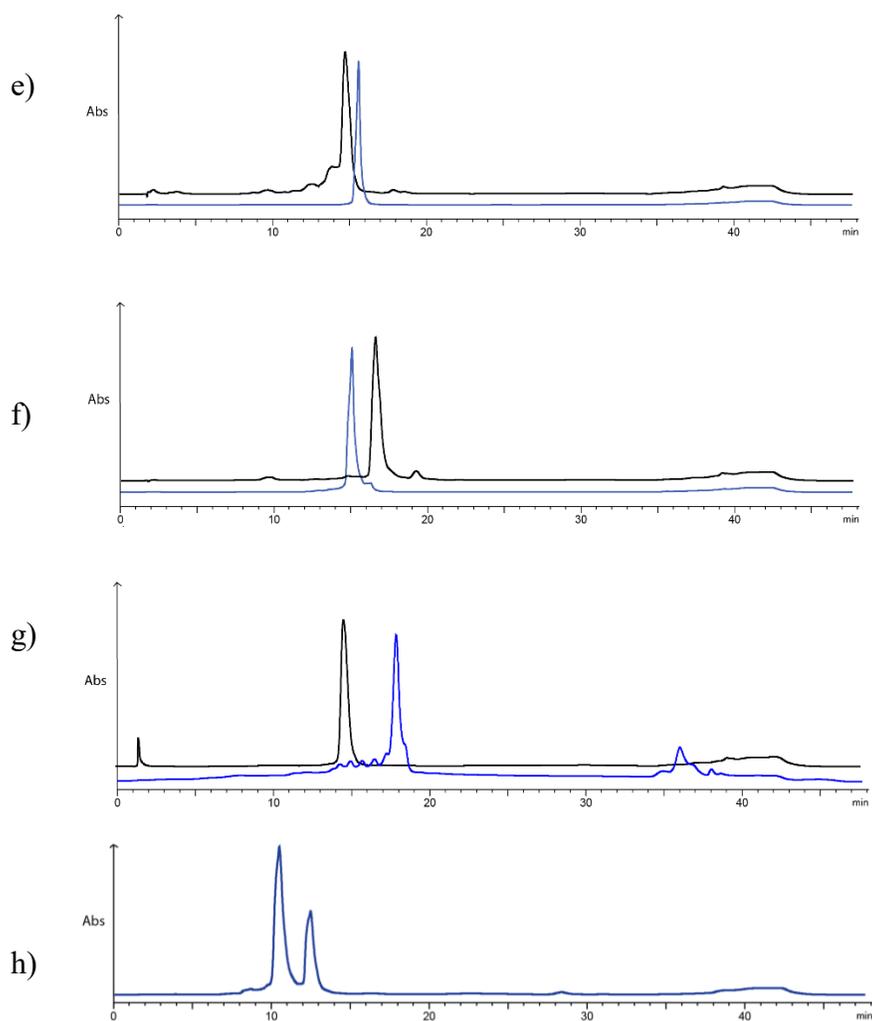


Figure 4.10. RP-HPLC traces from crude mixtures of 5' modified DNA 19mers. (UV 260 nm) a) top: DNA-9; bottom, unmodified DNA, 3 to 50 % ACN in 30 minutes. b) top: DNA-5; bottom, unmodified DNA, 3 to 80 % ACN in 30 minutes c) top: DNA-3; bottom, DNA-4, 3 to 80 % ACN in 30 minutes. d) top: DNA-6; bottom, unmodified DNA, 3 to 30 % ACN in 30 minutes e) top: DNA-PT2b; bottom, unmodified DNA, 3 to 30 % ACN in 30 minutes f) top: DNA-7 bottom DNA-8, 3 to 30 % ACN in 30 minutes g) top: DNA-6₅ after post-SPS click chemistry and gel purification. Bottom: crude from DNA-6₅ (oligomer F), overall yields: 82 %, 3 to 30 % ACN in 30 minutes, h) Oligomer E, DNA 14mer-7₂-8₃, 3 to 50 % in 30 minutes.

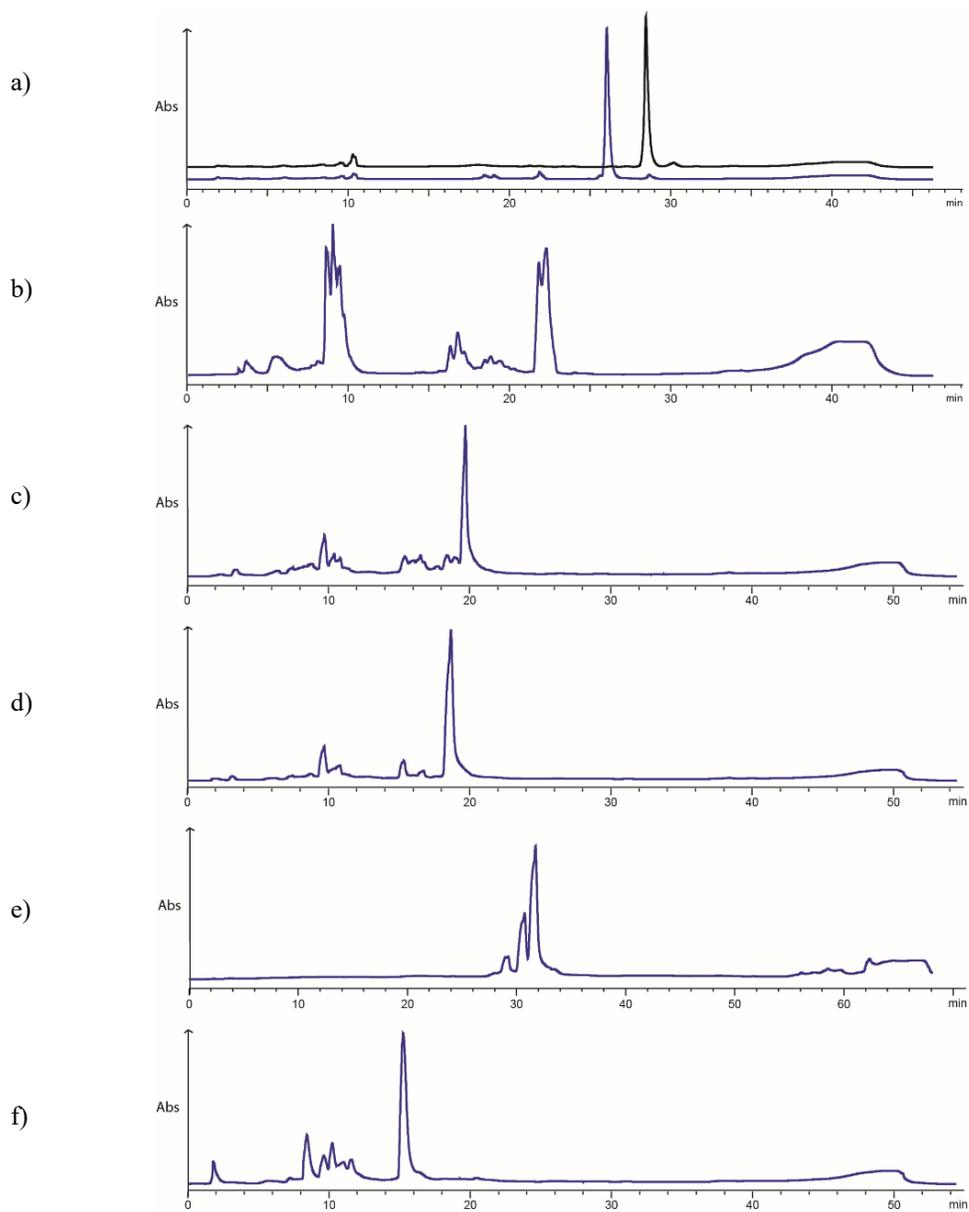
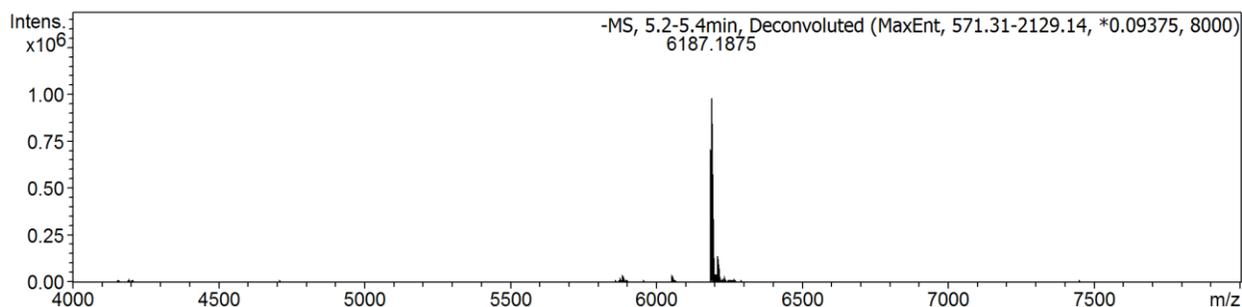


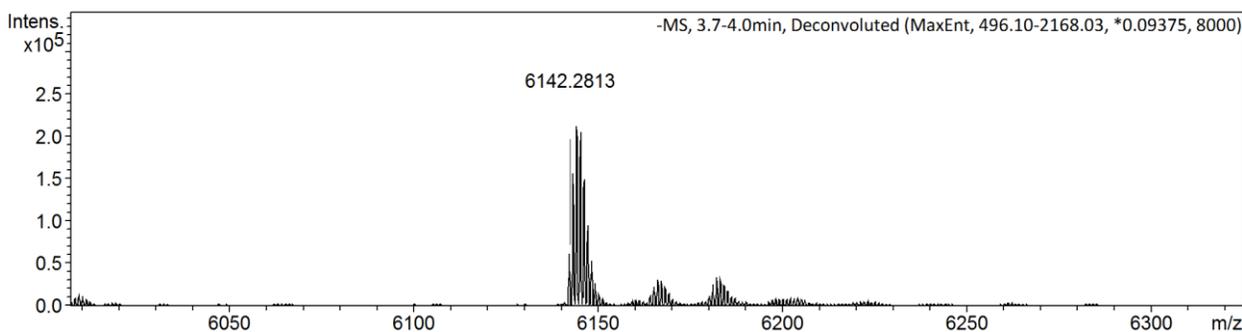
Figure 4.11. RP-HPLC traces from crude mixtures of sequence-defined oligomers. (UV 260 nm). a) top: T-**HEG-HEG-5-5-T**; bottom, T-**HEG-HEG-4-4-T**. 3 to 80 % ACN in 30 minutes. b) T-**HEG-HEG-3-3-T**, 3 to 80 % ACN in 30 minutes c) Oligomer **B** T-**6₂-8-TEG-9-8₂-10-TEG-9**, 3 to 80 % ACN in 40 minutes d) Oligomer **A** T-**6₂-8-TEG-9**, 3 to 80 % ACN in 40 minutes, e) Oligomer **C** T₃-**HEG₄-7-6₂**, 3 to 30 % ACN in 50 min., f) Oligomer **D** T-**6-8-6-8-6-10**, 3 to 80 % ACN in 40 minutes.

4.7.8 LC-ESI-MS characterization

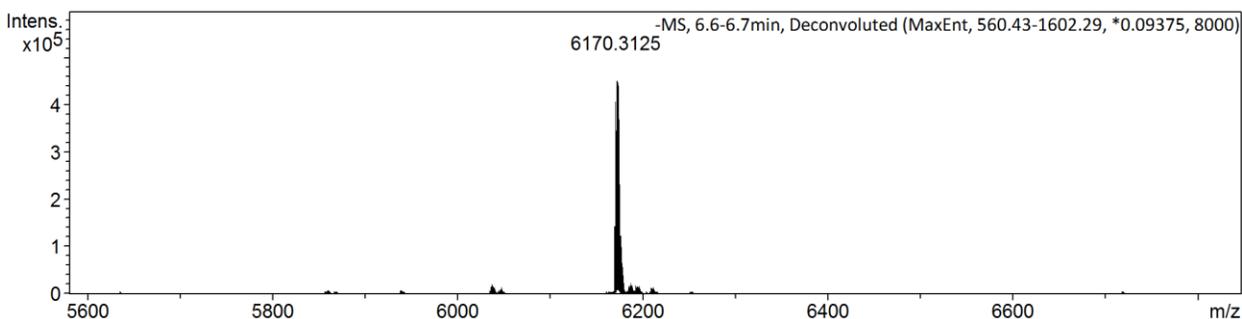
The oligomers were analyzed by LC-ESI-MS in negative ESI mode. Samples were run through an Acclaim RSLC 120 C18 column (2.2 μm , 120 \AA 2.1 x 50 mm) using a gradient of mobile phase A (100 mM 1,1,1,3,3,3-hexafluoro-2-propanol and 5 mM triethylamine in water) and mobile phase B (Methanol) in 8 minutes (2 % to 100 % B). Liquid chromatography was performed as a control for strand purity which was found to be superior to 90 % in all cases. Each time, ~ 250 pmol of artificial oligomer or ~ 60 pmol of oligonucleotide was injected.



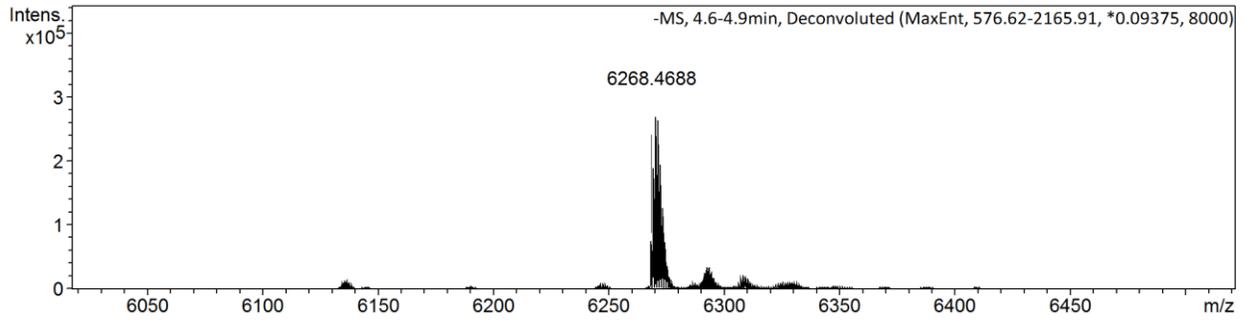
a) DNA-9



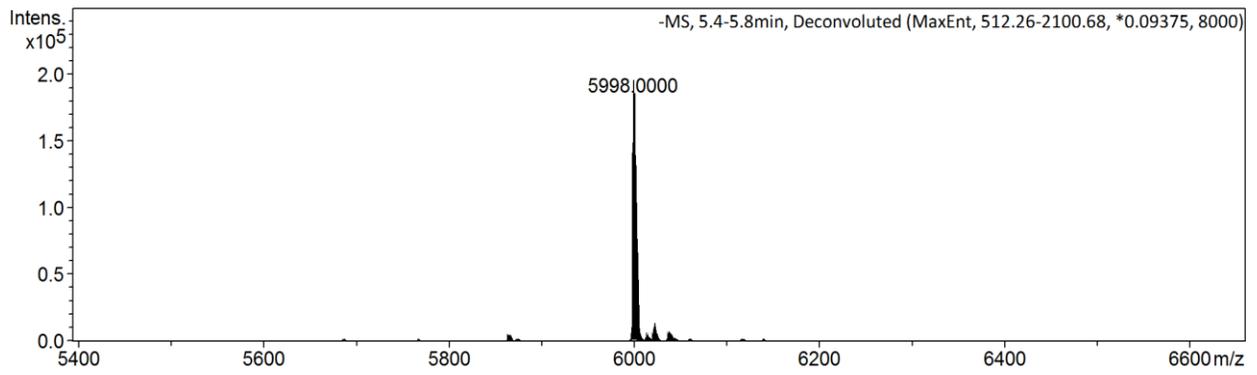
b) DNA-3



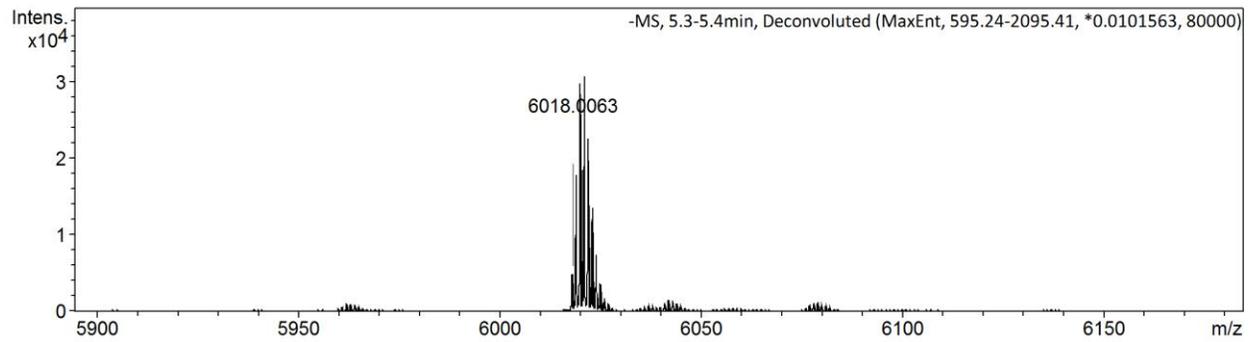
c) DNA-4. Mass measured is lower than expected (-1) due to the positive charge on 4 not taken into account by the deconvolution software.



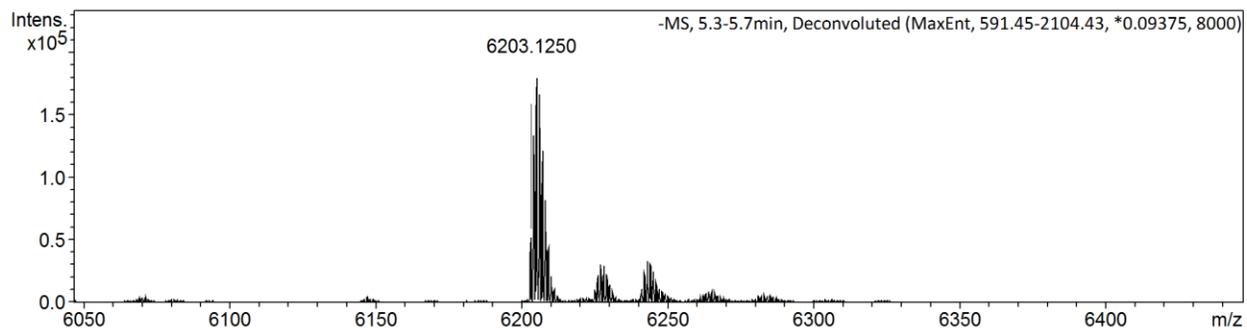
d) DNA-5



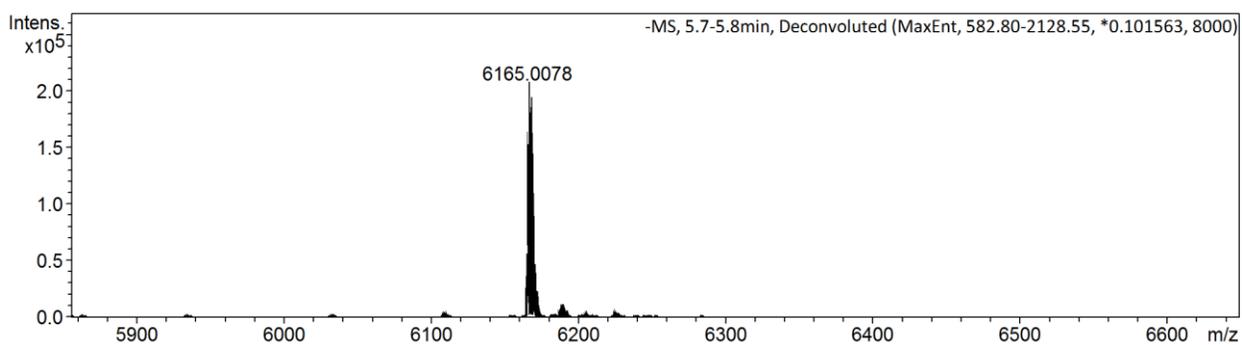
e) DNA-6



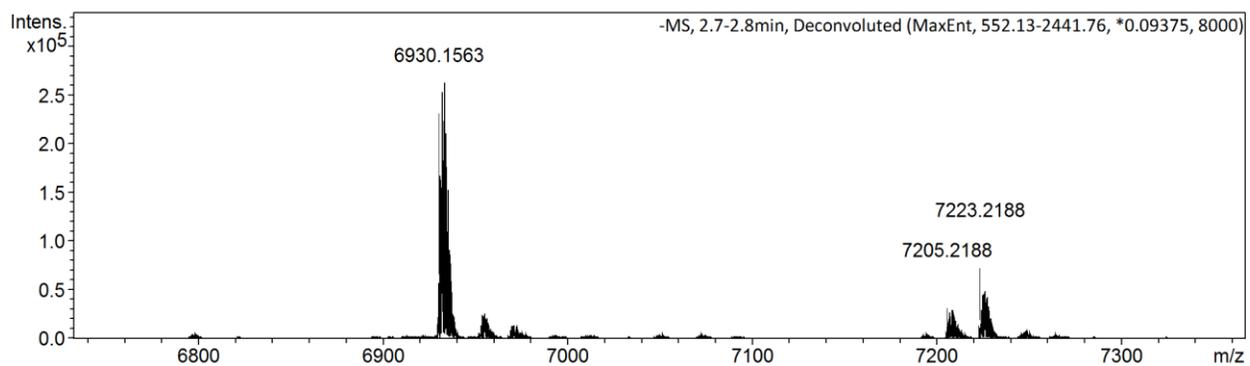
f) DNA-PT2b



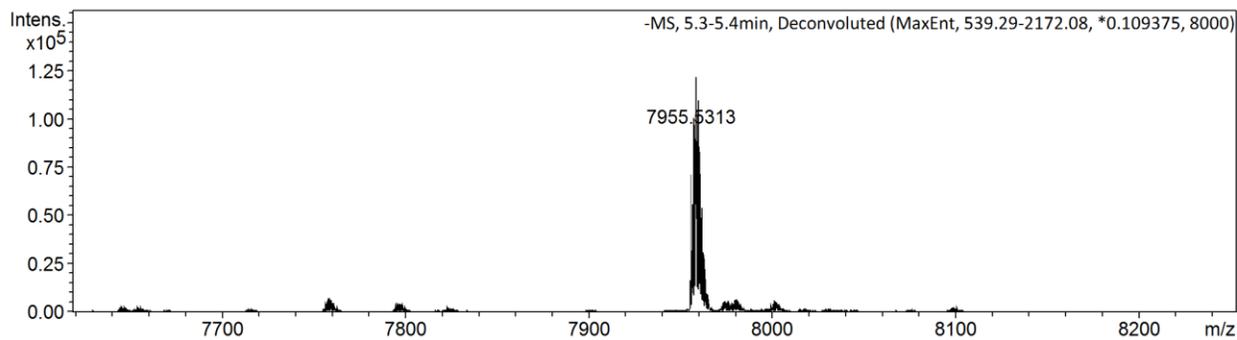
g) DNA-7



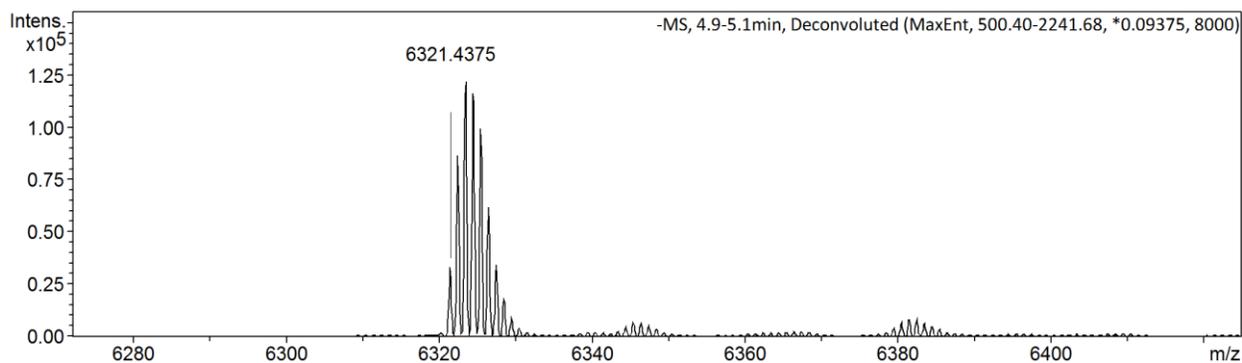
h) DNA-8



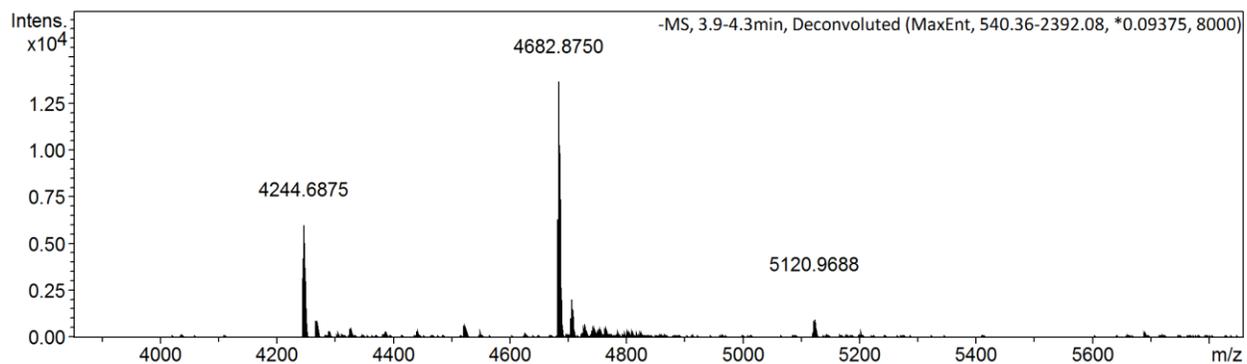
i) Oligomer **F**: DNA-6₅ (peaks at 7205 and 7223 most probably come from unylinker remaining on the strand 3' end. This issue can be solved with a longer deprotection).



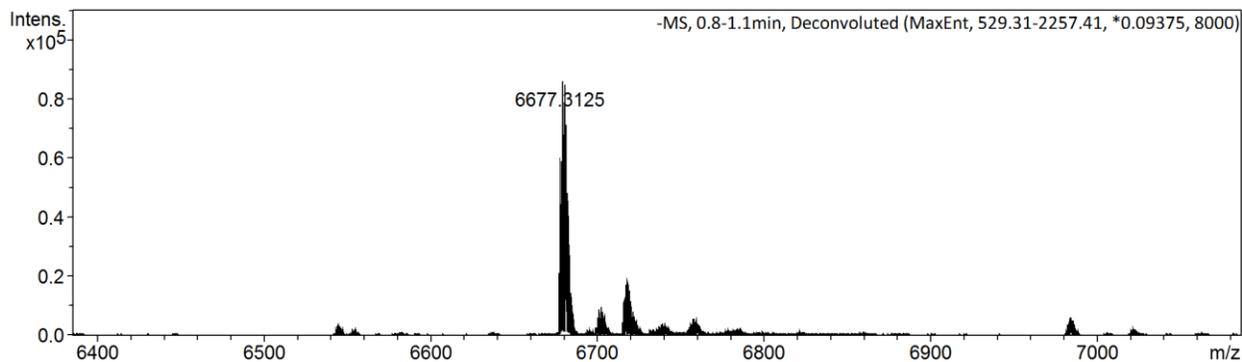
j) Oligomer **F** after click chemistry post SPS. Expected mass: 7955.75 g/mol.



k) Oligomer **E**, DNA 14mer-7₂-8₃

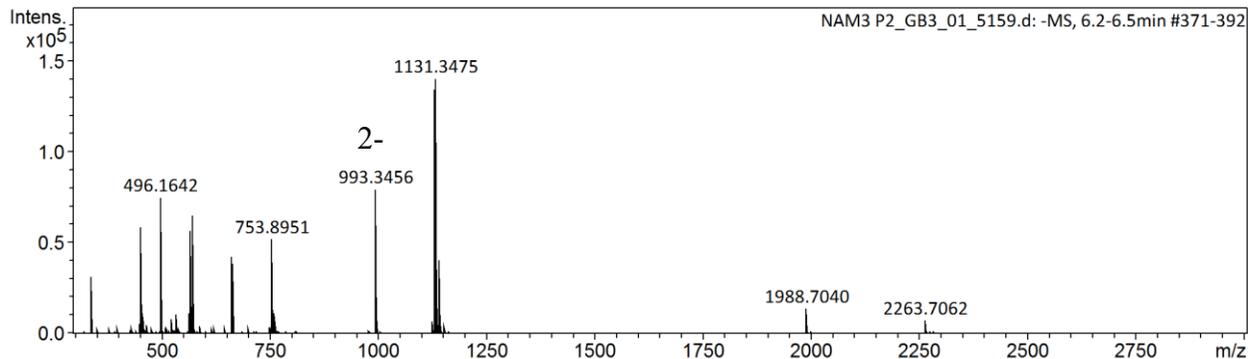


l) Byproducts from oligomer **E** synthesis (first peak from HPLC purification): DNA 14mer, DNA 14mer-7 and DNA 14mer-7₂. It seems that coupling of 7 was less efficient in this case.

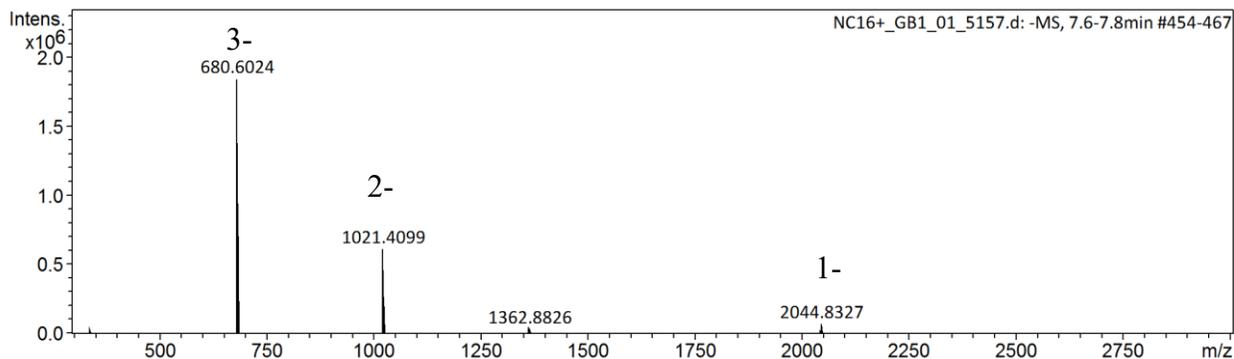


m) DNA-11-DNA. 5'-TA-11-TTTTTTCAGTTGACCATATA-3' Expected : 6677.35g/mol.

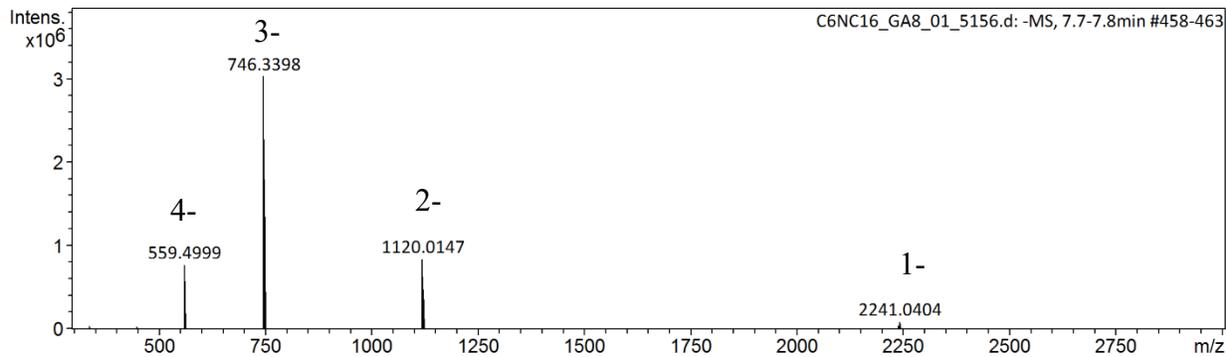
Figure 4.12. MS data for modified DNA strands. The data was processed and deconvoluted using the Bruker DataAnalysis software version 4.1. Masses reported are exact masses.



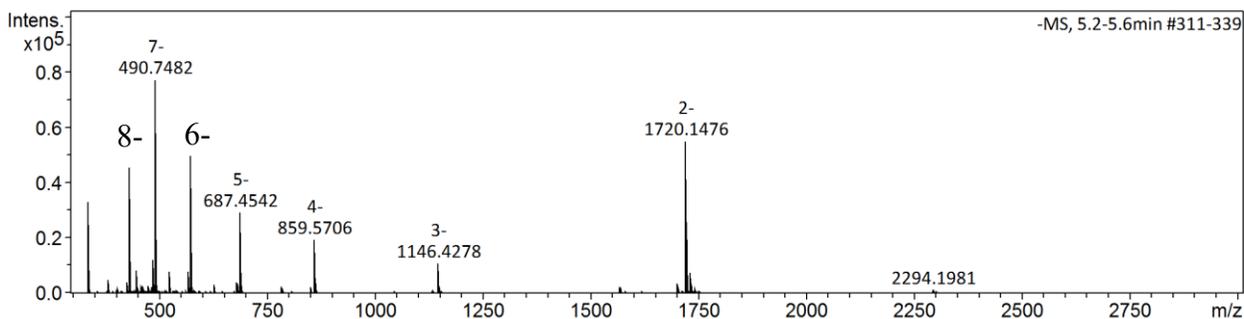
a) T-HEG-HEG-3-3-T



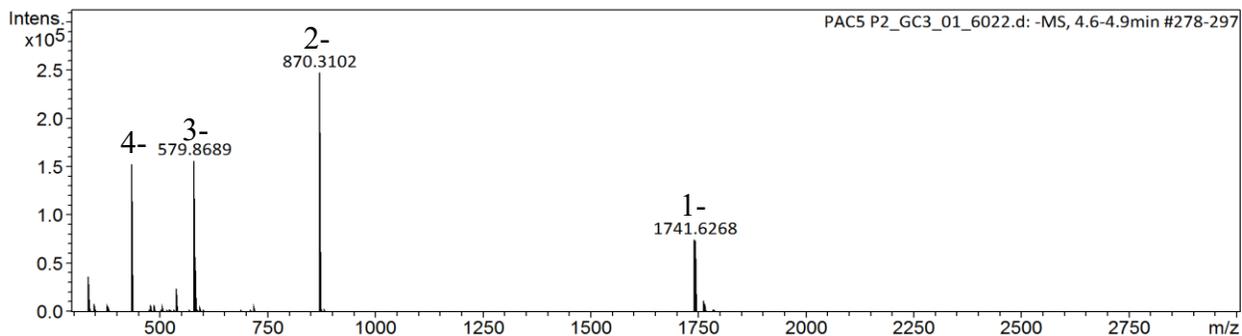
b) T-HEG-HEG-4-4-T



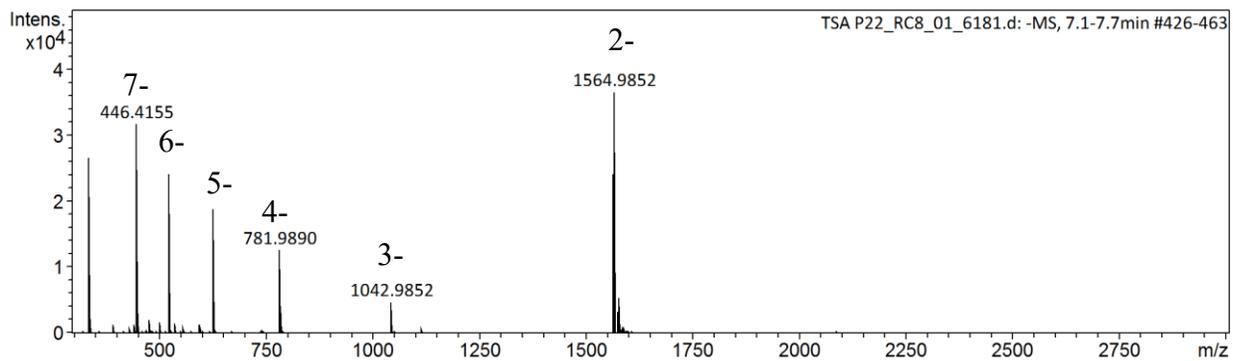
c) T-HEG-HEG-5-5-T



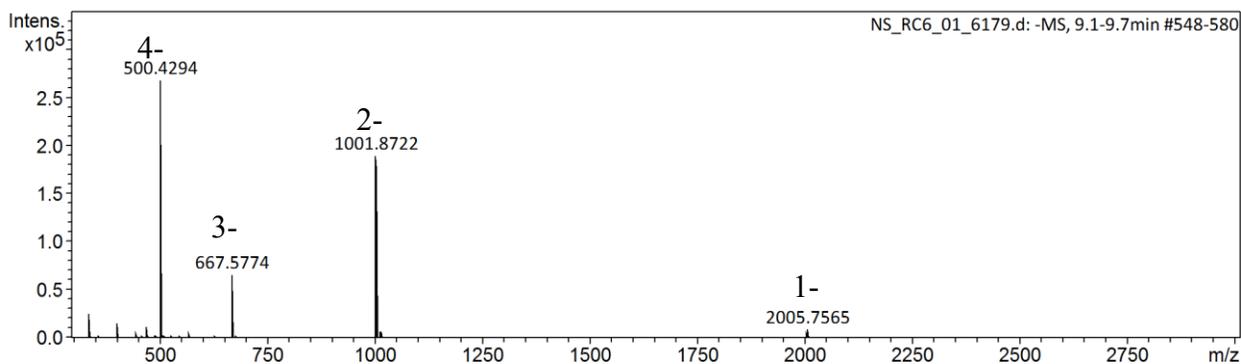
d) Oligomer B, T-6₂-8-TEG-9-8₂-10-TEG-9.



e) Oligomer A, T-6₂-8-TEG-9.



f) Oligomer C, T₃-HEG₄-7-6₂



g) Oligomer D, T-6-8-6-8-6-10

Figure 4.13. MS data for sequence-defined oligomers. Negative mode. Almost all peaks can be associated with a $(M-x)/x$ anion except for polymers containing 4 which has an intrinsic positive charge and for which the visible species are $(M - x - \text{number of } 4)/x$.

4.8 References

- (1) Cole, J. P.; Hanlon, A. M.; Rodriguez, K. J.; Berda, E. B. Protein-like Structure and Activity in Synthetic Polymers. *J. Polym. Sci. Part A Polym. Chem.* **2017**, *55* (2), 191–206.
- (2) Lutz, J.-F.; Ouchi, M.; Liu, D. R.; Sawamoto, M. Sequence-Controlled Polymers. *Science* **2013**, *341* (6146), 1238149.
- (3) Roy, R. K.; Meszynska, A.; Laure, C.; Charles, L.; Verchin, C.; Lutz, J.-F. Design and Synthesis of Digitally Encoded Polymers That Can Be Decoded and Erased. *Nat. Commun.* **2015**, *6*, 7237.
- (4) Al Ouahabi, A.; Charles, L.; Lutz, J.-F. Synthesis of Non-Natural Sequence-Encoded Polymers Using Phosphoramidite Chemistry. *J. Am. Chem. Soc.* **2015**, *137* (16), 5629–5635.
- (5) Le Bailly, B. A. F.; Clayden, J.; Patchornik, A.; Leito, I.; Morris, G. A.; Webb, S. J.; Clayden, J.; Beadle, J. D.; Clayden, J.; Rossi, F.; et al. Dynamic Foldamer Chemistry. *Chem. Commun.* **2016**, *52* (27), 4852–4863.
- (6) Guichard, G.; Huc, I. Synthetic Foldamers. *Chem. Commun. (Camb)*. **2011**, *47* (21), 5933–5941.
- (7) Delsuc, N.; Massip, S.; Léger, J.-M.; Kauffmann, B.; Huc, I. Relative Helix–Helix Conformations in Branched Aromatic Oligoamide Foldamers. *J. Am. Chem. Soc.* **2011**, *133* (9), 3165–3172.
- (8) Lombardo, C. M.; Collie, G. W.; Pulka-Ziach, K.; Rosu, F.; Gabelica, V.; Mackereth, C. D.; Guichard, G. Anatomy of an Oligourea Six-Helix Bundle. *J. Am. Chem. Soc.* **2016**, *138* (33), 10522–10530.
- (9) Harth, E.; Brooke Van Horn; Victor Y. Lee; David S. Germack; Chad P. Gonzales; Robert D. Miller, A.; Hawker, C. J. A Facile Approach to Architecturally Defined Nanoparticles via Intramolecular Chain Collapse. *J. Am. Chem. Soc.* **2002**, *124* (29), 8653–8660.

- (10) Latorre-Sánchez, A.; Pomposo, J. A. Recent Bioinspired Applications of Single-Chain Nanoparticles. *Polym. Int.* **2016**, *65* (8), 855–860.
- (11) Lyon, C. K.; Prasher, A.; Hanlon, A. M.; Tuten, B. T.; Tooley, C. A.; Frank, P. G.; Berda, E. B. A Brief User's Guide to Single-Chain Nanoparticles. *Polym. Chem.* **2015**, *6* (2), 181–197.
- (12) Terashima, T.; Mes, T.; De Greef, T. F. A.; Gillissen, M. A. J.; Besenius, P.; Palmans, A. R. A.; Meijer, E. W. Single-Chain Folding of Polymers for Catalytic Systems in Water. *J. Am. Chem. Soc.* **2011**, *133* (13), 4742–4745.
- (13) Perez-Baena, I.; Barroso-Bujans, F.; Gasser, U.; Arbe, A.; Moreno, A. J.; Colmenero, J.; Pomposo, J. A. Endowing Single-Chain Polymer Nanoparticles with Enzyme-Mimetic Activity. *ACS Macro Lett.* **2013**, *2* (9), 775–779.
- (14) Liu, Y.; Pauloehrl, T.; Presolski, S. I.; Albertazzi, L.; Palmans, A. R. A.; Meijer, E. W. Modular Synthetic Platform for the Construction of Functional Single-Chain Polymeric Nanoparticles: From Aqueous Catalysis to Photosensitization. *J. Am. Chem. Soc.* **2015**, *137* (40), 13096–13105.
- (15) Moreno, A. J.; Lo Verso, F.; Arbe, A.; Pomposo, J. A.; Colmenero, J. Concentrated Solutions of Single-Chain Nanoparticles: A Simple Model for Intrinsically Disordered Proteins under Crowding Conditions. *J. Phys. Chem. Lett.* **2016**, *7* (5), 838–844.
- (16) Grate, J. W.; Mo, K.-F.; Daily, M. D. Triazine-Based Sequence-Defined Polymers with Side-Chain Diversity and Backbone-Backbone Interaction Motifs. *Angew. Chem. Int. Ed. Engl.* **2016**, *55* (12), 3925–3930.
- (17) Martens, S.; Van den Begin, J.; Madder, A.; Du Prez, F. E.; Espeel, P. Automated Synthesis of Monodisperse Oligomers, Featuring Sequence Control and Tailored Functionalization. *J. Am. Chem. Soc.* **2016**, *138* (43), 14182–14185.
- (18) Zuckermann, R. N.; Kerr, J. M.; Kent, S. B. H.; Moos, W. H. Efficient Method for the Preparation of Peptoids [Oligo(N-Substituted Glycines)] by Submonomer Solid-Phase Synthesis. *J. Am. Chem. Soc.* **1992**, *114* (26), 10646–10647.
- (19) Knight, A. S.; Zhou, E. Y.; Francis, M. B.; Zuckermann, R. N. Sequence Programmable Peptoid Polymers for Diverse Materials Applications. *Adv. Mater.* **2015**, *27* (38), 5665–5691.
- (20) Murphy, J. E.; Uno, T.; Hamer, J. D.; Cohen, F. E.; Dwarki, V.; Zuckermann, R. N.; Schultz, P. G. A Combinatorial Approach to the Discovery of Efficient Cationic Peptoid Reagents for Gene Delivery. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 1517–1522.
- (21) Beaucage, S. L.; Caruthers, M. H. Deoxynucleoside Phosphoramidites—A New Class of Key Intermediates for Deoxypolynucleotide Synthesis. *Tetrahedron Lett.* **1981**, *22* (20), 1859–1862.
- (22) Beaucage, S. L. A Simple and Efficient Preparation of Deoxynucleoside Phosphoramidites in Situ. *Tetrahedron Lett.* **1984**, *25* (4), 375–378.
- (23) *Current Protocols in Nucleic Acid Chemistry*; Egli, M., Herdewijn, P., Matusda, A., Sanghvi, Y. S., Eds.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2001.

- (24) Kosuri, S.; Church, G. M. Large-Scale de Novo DNA Synthesis: Technologies and Applications. *Nat Meth* **2014**, *11* (5), 499–507.
- (25) Edwardson, T. G. W.; Carneiro, K. M. M.; Serpell, C. J.; Sleiman, H. F. An Efficient and Modular Route to Sequence-Defined Polymers Appended to {DNA}. *Angew. Chem. Int. Ed. Engl.* **2014**, *53* (18), 4567–4571.
- (26) de Rochambeau, D.; Barlóg, M.; Edwardson, T. G. W.; Fakhoury, J. J.; Stein, R. S.; Bazzi, H. S.; Sleiman, H. F. “DNA–Teflon” Sequence-Controlled Polymers. *Polym. Chem.* **2016**, *7* (31), 4998–5003.
- (27) Serpell, C. J.; Edwardson, T. G. W.; Chidchob, P.; Carneiro, K. M. M.; Sleiman, H. F. Precision Polymers and 3D DNA Nanostructures: Emergent Assemblies from New Parameter Space. *J. Am. Chem. Soc.* **2014**, *136* (44), 15767–15774.
- (28) Chidchob, P.; Edwardson, T. G. W.; Serpell, C. J.; Sleiman, H. F. Synergy of Two Assembly Languages in DNA Nanostructures: Self-Assembly of Sequence-Defined Polymers on DNA Cages. *J. Am. Chem. Soc.* **2016**, *138* (13), 4416–4425.
- (29) Lacroix, A.; Edwardson, T. G. W.; Hancock, M. A.; Dore, M. D.; Sleiman, H. F. Development of DNA Nanostructures for High-Affinity Binding to Human Serum Albumin. *J. Am. Chem. Soc.* **2017**, *139* (21), 7355–7362.
- (30) Appukutti, N.; Serpell, C. J. High Definition Polyphosphoesters: Between Nucleic Acids and Plastics. *Polym. Chem.* **2018**, *9* (17), 2210–2226.
- (31) Ouahabi, A. Al; Kotera, M.; Charles, L.; Lutz, J.-F. Synthesis of Monodisperse Sequence-Coded Polymers with Chain Lengths above DP100. *ACS Macro Lett.* **2015**, *4* (10), 1077–1080.
- (32) König, N. F.; Al Ouahabi, A.; Poyer, S.; Charles, L.; Lutz, J.-F. A Simple Post-Polymerization Modification Method for Controlling Side-Chain Information in Digital Polymers. *Angew. Chemie Int. Ed.* **2017**, *56* (25), 7297–7301.
- (33) Häner, R.; Garo, F.; Wenger, D.; Malinovskii, V. L. Oligopyrenotides: Abiotic, Polyanionic Oligomers with Nucleic Acid-like Structural Properties. *J. Am. Chem. Soc.* **2010**, *132* (21), 7466–7471.
- (34) Vyborna, Y.; Vybornyi, M.; Häner, R. Pathway Diversity in the Self-Assembly of DNA-Derived Bioconjugates. *Bioconjug. Chem.* **2016**, *27* (11), 2755–2761.
- (35) Vasilyeva, S. V.; Budilkin, B. I.; Konevets, D. A.; Silnikov, V. N. Synthesis of Novel Nucleoside 5'-Triphosphates and Phosphoramidites Containing Alkyne or Amino Groups for the Postsynthetic Functionalization of Nucleic Acids. *Nucleosides, Nucleotides and Nucleic Acids* **2011**, *30* (10), 753–767.
- (36) Nelson, P. S.; Sherman-Gold, R.; Leon, R. A New and Versatile Reagent for Incorporating Multiple Primary Aliphatic Amines into Synthetic Oligonucleotides. *Nucleic Acids Res.* **1989**, *17* (18), 7179–7186.
- (37) Nelson, P. S.; Kent, M.; Muthini, S. Oligonucleotide Labeling Methods. 3. Direct Labeling of Oligonucleotides Employing a Novel, Non-Nucleosidic, 2-Aminobutyl-1,3-Propanediol

- Backbone. *Nucleic Acids Res.* **1992**, *20* (23), 6253–6259.
- (38) Ito, H.; Liang, X.; Nishioka, H.; Asanuma, H. Construction of Photoresponsive RNA for Photoswitching RNA Hybridization. *Org. Biomol. Chem.* **2010**, *8* (24), 5519.
- (39) Putnam, W. C.; Bashkin, J. K. De Novo Synthesis of Artificial Ribonucleases with Benign Metal Catalysts. *Chem. Commun.* **2000**, *0* (9), 767–768.
- (40) Paul S. Nelson, Hugh Mackie, A. M. Reagents Utilizing a Serinol Scaffold for Labeling Synthetic Oligonucleotides. US 8394948 B2, 2010.
- (41) Reed, M. W.; Adams, A. D.; Nelson, J. S.; Meyer, R. B. Acridine- and Cholesterol-Derivatized Solid Supports for Improved Synthesis of 3'-Modified Oligonucleotides. *Bioconjug. Chem.* **1991**, *2* (4), 217–225.
- (42) Hébert, N.; Davis, P. W.; DeBaets, E. L.; Acevedo, O. L. Synthesis of N-Substituted Hydroxyprolinol Phosphoramidites for the Preparation of Combinatorial Libraries. *Tetrahedron Lett.* **1994**, *35* (51), 9509–9512.
- (43) Kupryushkin, M. S.; Pyshnyi, D. V. Oligonucleotide Derivatives in Nucleic Acid Hybridization Analysis. III. Synthesis and Investigation of Properties of Oligonucleotides, Bearing Bifunctional Non-Nucleotide Insertion. *Russ. J. Bioorganic Chem.* **2012**, *38* (6), 625–638.
- (44) Kupryushkin, M. S.; Nekrasov, M. D.; Stetsenko, D. A.; Pyshnyi, D. V. Efficient Functionalization of Oligonucleotides by New Achiral Nonnucleosidic Monomers. *Org. Lett.* **2014**, *16* (11), 2842–2845.
- (45) Beaucage, S. L. Oligodeoxyribonucleotides Synthesis: Phosphoramidite Approach. In *Methods in molecular biology (Clifton, N.J.)*; Agrawal, S., Ed.; Humana Press: New Jersey, 1993; pp 33–62.
- (46) Desaulniers, J.-P.; Hagen, G.; Anderson, J.; McKim, C.; Roberts, B. Effective Gene-Silencing of siRNAs That Contain Functionalized Spacer Linkages within the Central Region. *RSC Adv.* **2017**, *7* (6), 3450–3454.
- (47) Lin, K. Y.; Matteucci, M. Hybridization Properties of Deoxyoligonucleotides Containing Anthraquinone Pseudonucleosides. *Nucleic Acids Res.* **1991**, *19* (11), 3111–3114.
- (48) Online pKa prediction tool <http://epoch.uky.edu/ace/public/pKa.jsp>.
- (49) Jezowska, M.; Honcharenko, D.; Ghidini, A.; Strömberg, R.; Honcharenko, M. Enabling Multiple Conjugation to Oligonucleotides Using “Click Cycles.” *Bioconjug. Chem.* **2016**, *27* (11), 2620–2628.
- (50) Johannes Gierlich; Glenn A. Burley; Philipp M. E. Gramlich; David M. Hammond, A.; Carell, T. Click Chemistry as a Reliable Method for the High-Density Postsynthetic Functionalization of Alkyne-Modified DNA. *Org. Lett.* **2006**, *8* (17), 3639–3642.
- (51) Seeman, N. C.; Sleiman, H. F. DNA Nanotechnology. *Nat. Rev. Mater.* **2017**, *3*, 17068.
- (52) Hamblin, G. D.; Rahbani, J. F.; Sleiman, H. F. Sequential Growth of Long DNA Strands with User-Defined Patterns for Nanostructures and Scaffolds. *Nat. Commun.* **2015**, *6*, 7065.

- (53) Huang, P.-J. J.; Liu, J. Rational Evolution of Cd²⁺-Specific DNAzymes with Phosphorothioate Modified Cleavage Junction and Cd²⁺ Sensing. *Nucleic Acids Res.* **2015**, *43* (12), 6125–6133.
- (54) Fernández, M.; Javaid, F.; Chudasama, V. Advances in Targeting the Folate Receptor in the Treatment/Imaging of Cancers. *Chem. Sci.* **2018**, *9* (4), 790–810.
- (55) Gagnon, K. T.; Watts, J. K.; Pendergraff, H. M.; Montañier, C.; Thai, D.; Potier, P.; Corey, D. R. Antisense and Antigene Inhibition of Gene Expression by Cell-Permeable Oligonucleotide-Oligospermine Conjugates. *J. Am. Chem. Soc.* **2011**, *133* (22), 8404–8407.
- (56) Li, Y.-J.; Nakamura, N.; Chen, T.-M.; Wang, Y.-F.; Kitamura, M.; Nakaya, T. Synthesis of Comb-like Polyurethanes Containing Hydrophilic Phosphatidylcholine Analogues in the Main Chains and Hydrophobic Long Chain Alkyl Groups in the Side Chains. *Macromol. Rapid Commun.* **1996**, *17* (10), 737–744.
- (57) Sagnella, S. M.; Conn, C. E.; Krodziewska, I.; Drummond, C. J. Nonionic Diethanolamide Amphiphiles with Saturated Hydrocarbon Chains: Neat Crystalline and Lyotropic Liquid Crystalline Phase Behavior. *J. Colloid Interface Sci.* **2012**, *385* (1), 87–95.
- (58) He, J.; Söderling, E.; Lassila, L. V. J.; Vallittu, P. K. Synthesis of Antibacterial and Radio-Opaque Dimethacrylate Monomers and Their Potential Application in Dental Resin. *Dent. Mater.* **2014**, *30* (9), 968–976.
- (59) Ramakrishna, K. K. G.; Thakur, R. K.; Pasam, V. R.; Pandey, J.; Mahar, R.; Shukla, S. K.; Tamrakar, A. K.; Tripathi, R. P. Synthesis of Novel Glycosyl-1,2,3-1H-Triazolyl Methyl Quinazolin-4(3H)-Ones and Their Effect on GLUT4 Translocation. *Tetrahedron* **2017**, *73* (2), 187–203.
- (60) Cavaluzzi, M. J.; Borer, P. N. Revised UV Extinction Coefficients for Nucleoside-5'-Monophosphates and Unpaired DNA and RNA. *Nucleic Acids Res.* **2004**, *32* (1), e13.

|5|

DNA-Encoded Functionalized Aptamers

5.1 Preface

In the chapters 2 and 3, we demonstrated that some new monomers permitted the formation of novel supramolecular structures and imparted new properties to DNA strands. In Chapter 4, we showed the possibility to dramatically expand the number of monomers available for the synthesis of sequence-defined oligo(phosphodiester)s. This approach will lead to numerous applications, such as new biological properties of DNA strands, single-chain folded particles for catalysis, data storage, and new molecular recognition properties; the latter application is the subject of this chapter.

In the introduction, the difficulty of predicting a folded 3-dimensional structure *a priori* from a specific sequence was pointed out. It is a similar challenge as the protein folding problem but with oligomers made of artificial building blocks. Therefore, we can apply similar strategies as those developed for peptide sequence discovery. Noticeably, the generation of large combinatorial libraries containing thousands of different sequences has been found to be valuable in discovering new antibacterial peptides¹ and oligonucleotide aptamers² for example.

DNA-encoding is a strategy used when generating combinatorial libraries. In this method, a short DNA sequence acts as a tag for each added building block in the molecule, so that the best binding molecule from a large library can be identified by DNA sequencing. The oligomers presented in this thesis have a similar backbone as DNA and are made using a similar chemical process. Therefore, DNA-encoding technology seems to be the method of choice to make combinatorial libraries of sequence-defined oligomer libraries. It would greatly enhance the chances to find useful sequences that would take advantage of the monomer diversity achieved in Chapter 4.

5.2 Contribution of authors

Donatien de Rochambeau codesigned the project and performed all experiments unless listed below, analyzed the results and cowrote the chapter. **Michael Dore** synthesized the DMT-protected precursor of **Bal** and **Dr. Violeta Toader** performed the phosphoramidite synthesis step for **Bal** and **C12**. **Daniel Saliba** performed and optimized the PCR amplification experiments. **Dr. Hanadi Sleiman** codesigned the project, guided interpretation of data and discussion of results and cowrote the chapter.

5.3 Abstract

Sequence-defined oligomers made of unnatural building blocks can be used as biologically relevant ligands with improved properties compared to their natural counterparts such as oligonucleotide aptamers. Herein, we show the synthesis of DNA-encoded sequence-defined functionalized aptamers that will enable the selection and identification of potent target binding sequences from a large highly functionalized library. The use of levulinyl protecting groups is shown to be fully orthogonal to dimethoxytrityl and allows the parallel synthesis of two oligo(phosphodiester)s. Therefore, on a solid phase support, a DNA code made of nucleotide phosphoramidites was synthesized simultaneously with a sequence-defined oligomer made of non-nucleosidic monomers. Using the split-and-pool combinatorial strategy, a library of ~300,000 DNA-encoded unnatural aptamers based on the thrombin-binding aptamer sequence was synthesized. Monomers used were designed to improve affinity to thrombin, increase the aptamer serum half-life, and potentially help cellular transfection. The oligonucleotide of a control DNA-encoded structure was amplified and sequenced successfully from very low concentrations to show the possibility to use our library for target affinity selection. This synthetic methodology represents a potential major breakthrough in both the aptamer field, with the use of non-nucleosidic modifications; and the DNA-encoded library field, by exploring a new chemical space through phosphoramidite chemistry and solid-phase synthesis.

5.4 Introduction

Aptamers are oligonucleotide sequences that specifically bind relevant targets due to their well-defined sequence-dependent tertiary structures. Aptamers are attractive alternatives to antibodies

because of their stability and smaller-size, and are more versatile than small molecules due to their specific 3D shape.^{3,4} Complex targets that are considered “undruggable” with small molecules such as protein-protein interactions (PPIs) can be efficiently targeted.⁵ As oligonucleotides, they are easy to synthesize through solid-phase synthesis (SPS), are less immunogenic than antibodies, and potent sequences can be obtained from large libraries through *in vitro* selection, also known as systematic evolution of ligands by exponential enrichment (SELEX).^{6,7} However, the chemical space offered by aptamers (4 nucleobases: ATGC or AUGC) is much narrower than proteins (20 amino acids with very different functional groups). This has made the discovery of high affinity-binding sequences with no off-target binding difficult in many cases.⁸ In addition, due to their nucleic acid nature, aptamers are susceptible to nuclease degradation leading to short half-lives *in vivo*, and they suffer from poor cellular uptake. These major drawbacks have considerably slowed the emergence of therapeutic applications,² and currently only one aptamer drug (pegaptanib, Macugen®) has been approved by the United States Food and Drug Administration (FDA).² To improve the properties of oligonucleotide aptamers, many scientists have chemically modified DNA and RNA nucleobases. Four recent strategies highlight significant steps forward in the field: (i) Slow off-rate modified aptamers (SOMAmers) were developed by Gold and coworkers by introducing nucleotides bearing amino acid-like side-chains, and they showed superior binding affinities;⁹ (ii) click-SELEX allows the addition of an unnatural modifier to an alkyne containing base during SELEX;¹⁰ (iii) ligation of modified short nucleic acids fragments on an evolved template was demonstrated by the Liu and Hili groups;¹¹⁻¹³ (iv) finally, polymerases and reverse-transcriptases capable of reading and synthesizing backbone-modified oligonucleotides (xeno-nucleic acids, XNA) were developed by Holliger and others.^{14,15} All these strategies expanded the chemical space explored in aptamers or increased their serum stability. However, the three first methods offer the possibility to modify only few nucleotides of a given strand and the fourth only yielded few aptamers so far.¹⁶ These strategies also all rely on nucleotides with relatively conservative modifications to natural nucleic acids to allow for enzyme-aided amplification or ligation. Therefore, to completely open the chemical space available to aptamers, a strategy that allows the incorporation of non-nucleosidic functional modifications would be highly attractive.

Sequence-defined polymers made of unnatural building blocks have attracted considerable attention in the past decade due to their potential to broaden the scope of biopolymer function and to extend their performance.¹⁷ However, the emergence of applications that require reliable and

precise folding is slowed down by the difficulty to predict the final structure from an initial sequence.¹⁸ Large combinatorial libraries help overcome this issue. Initially developed for natural biopolymers and small molecules, this strategy has been expanded to artificial sequence-defined polymers such as oligocarbamates,¹⁹ peptoids²⁰ and thiolactone precision polymers²¹, that have found applications as protein ligands, antimicrobial agents and drug solubilizers respectively. These studies highlight the great potential of combinatorial libraries of unnatural sequence-defined polymers. Classic oligomer library generation methods are limited by the surface area and instrumentation needed for high throughput screening (HTS) and by the difficulty to accurately sequence oligomers in mixed libraries. In 1992, Brenner and Lerner introduced the idea of DNA-encoded libraries, where each library member is barcoded with a DNA strand (*Figure 5.1*).²² This code can be amplified and sequenced from very small concentrations ($<10^{-15}$ nM) thanks to the high reliability of the polymerase chain reaction (PCR).²³ This area has been thoroughly developed in academic labs and pharmaceutical companies, and it has led to the discovery of active small molecules that are now entering clinical trials.²⁴ Finding new reactions and strategies to make DNA-encoded libraries (DEL) is an active research area to access a greater diversity of drug candidates. A wide chemical space²⁵ was explored so far but DEL of sequence-defined oligomers, while holding great promise, is considerably underexplored.^{26,27} Liquid-phase organic reactions that are orthogonal to DNA-ligation chemistry typically used for DEL synthesis²⁵ may not provide the high coupling yields necessary for making sequence-defined polymers. To solve this issue, strategies involving DNA-templated reactions²⁸ and solid-phase synthesis (SPS) of DNA-encoded peptides²⁹ have been reported. SPS can afford high-purity DNA strands³⁰ and is even used for the production of FDA-approved oligonucleotide therapeutics.³¹ In the chapters 2 to 4, we showed that artificial sequence-defined oligo(phosphodiester)s³² can also be synthesized in high yields with this method. This therefore represents an opportunity to expand the chemical space explored with DNA-encoded libraries to a variety of sequence-defined oligomers.

Herein, we report the design and synthesis of a DNA-encoded library of aptamers made of non-nucleosidic building blocks. The constructs were designed as branched DNA strands, each composed of a DNA code and a highly functionalized sequence-defined oligo(phosphodiester) aptamer. The specific nature and position of each non-nucleosidic unit is encoded by a three-DNA base codon residing on the covalently linked DNA arm. This strategy relies on the use of the levulinyl orthogonal protecting group to be able to synthesize the DNA-code and the unnatural

oligomer simultaneously and conveniently through phosphoramidite chemistry on solid phase support. We used the split-and-pool strategy to synthesize a ~300,000 member library of DNA-encoded potential thrombin-binding aptamers. Each aptamer was designed using a semi-rational approach to resemble an already reported thrombin-binding aptamer.³³ We used the non-nucleosidic building blocks developed in Chapter 3 and Chapter 4 since they are versatile, flexible and can induce weak interactions including hydrogen-bonding, π - π stacking and the hydrophobic effect.³² Moreover, some of the modifications are positively charged at physiological pH, which may help the oligomers' cell internalization. Successful amplification and sequencing of the code of model DNA-encoded unnatural aptamers demonstrates the applicability of our method to a selection process (**Figure 5.1**) that we envision as future work.

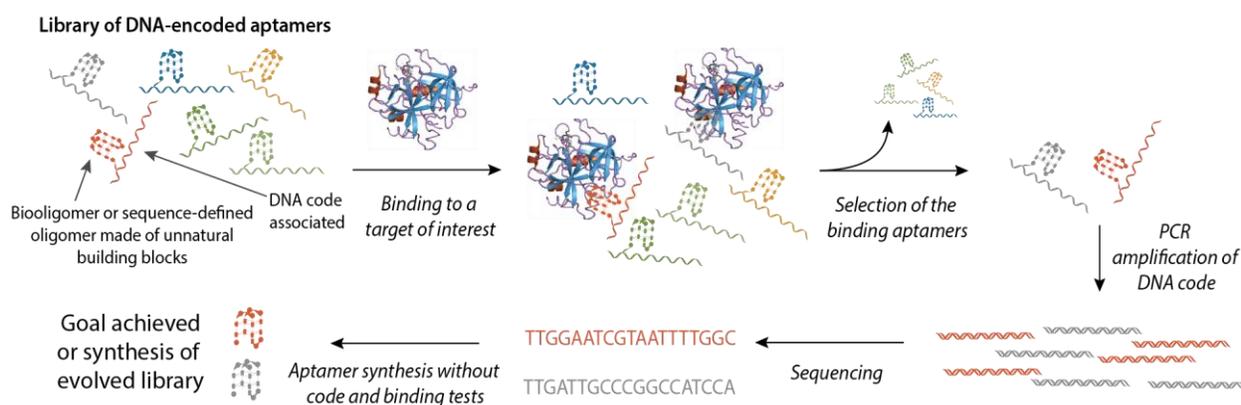


Figure 5.1. Selection principle of DNA-encoded unnatural aptamer libraries. Our idea takes inspiration from Brenner and Lerner's who initially proposed the synthesis of DNA-encoded peptides on beads. 1. A library of aptamers with a DNA tag is incubated with the target protein that is usually immobilized on a surface. 2. The unbound aptamers are eluted out. 3. The bound aptamers are recovered and their DNA tags are amplified. 4. Sequencing of the DNA amplicons obtained. 5. The DNA sequences allow to know the aptamer sequences. The latter are resynthesized without the DNA tag and tested. If the affinity for the target is not satisfactory, one can synthesize an evolved library on solid-phase.

5.5 Results and discussion

5.5.1 Finding an orthogonal protecting group

We developed a novel "genetic code" that would associate a 3-base codon to each unnatural monomer used in this study. As explained later in this chapter (**Scheme 5.65**), the split-and-pool strategy requires the parallel synthesis of the aptamer and the code associated. The requirement for a constant reverse primer region at the 3' end of the code for future PCR amplification led us to the design shown in Figure 5.2. It consists of a branched oligonucleotide in which one of the

arms is made of unnatural building blocks and eventually forms a sequence-defined oligo(phosphodiester) aptamer. This branch will be called the aptamer throughout the chapter. The other arm of the construct consists of regular DNA and codes for the sequence of the aptamer – it will be called the code.

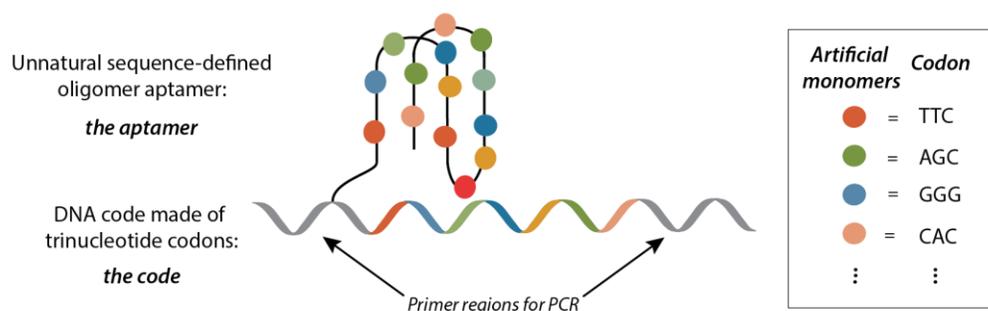
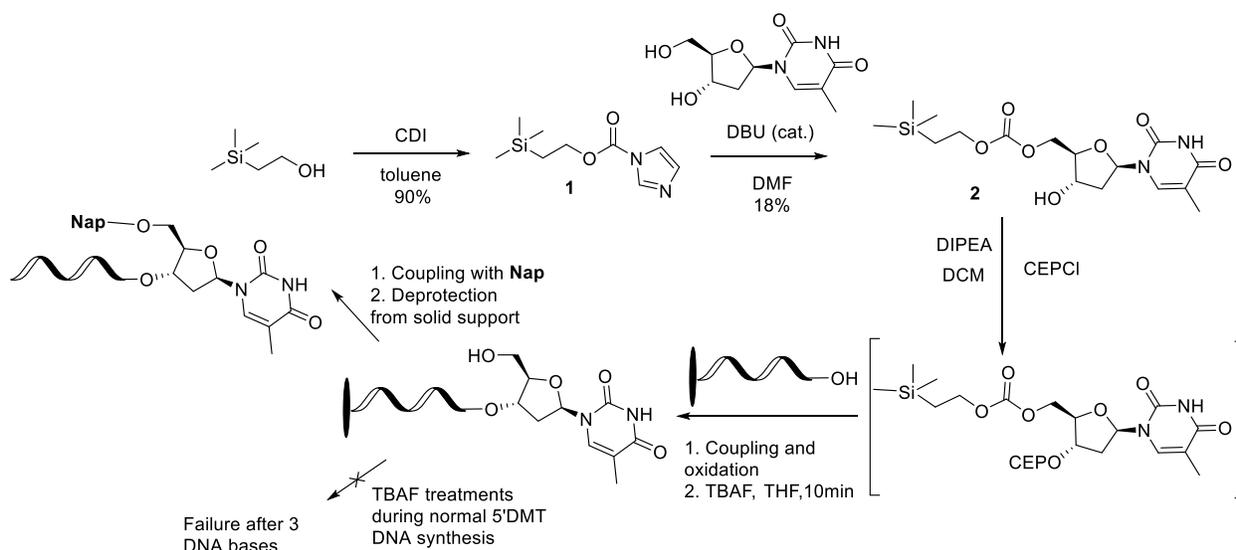


Figure 5.2. General design of DNA-encoded unnatural aptamers.

The original idea of Brenner and Lerner discussed the use of orthogonal chemistries because the aptamer and the code need to grow independently.²² In our case, phosphoramidite chemistry is used for the aptamer on the branch as well as the code. Therefore, an orthogonal protecting group to the common dimethoxytrityl (DMT) is required. This new group must fulfil several criteria: (i) it needs to be resistant to all reagents used in a DNA synthesizer, (ii) it must be quickly and quantitatively removed and (iii) classical DNA protecting groups on nucleotides and phosphates need to stay unaltered to the orthogonal deprotection conditions. Our first design strategy was to use silylated protecting groups. For example, *tert*-butyldimethylsilane (TBDMS) and triisopropylsilyloxymethyl (TOM) 2'-hydroxyl protection of RNA are resistant to synthesizer reagents.³⁴ However, we were unable to determine an efficient and orthogonal way to deprotect the TBDMS and TOM groups. 2'-acetoxy ethyl orthoester (ACE) RNA chemistry involves the use of a silylated group at the 5' end of RNA nucleotides that is readily deprotected using hydrofluoric acid (HF) in triethylamine (TEA).³⁵ However, this imposes the use of a methyl group instead of a cyanoethyl as phosphate protection, requiring harsher deprotection conditions that may not be compatible with the unnatural monomers of the aptamers. We also explored the more seldomly used trimethylsilylethoxycarbonyloxy (Teoc) protecting group. This group has notably been used in the total synthesis of verrucaric acid.³⁶ We followed a similar protocol to make a 5'-Teoc protected thymidine nucleoside (**2**, **Scheme 5.1**). Successful attachment to a DNA 19mer was achieved. The Teoc group showed resistance to synthesizer reagents and was orthogonally deprotected using

tetrabutylammonium fluoride (TBAF) for 10 minutes. However, several TBAF treatments during a common DNA strand synthesis did not lead to satisfactory yields. TBAF probably deprotects cyanoethyl groups and alters Controlled Pore Glass (CPG) solid support. Polystyrene beads as an alternative to CPG are not recommended for long and branched oligomers similar to the ones we want to make.



Scheme 5.1. Synthesis of 5'-Teoc protected thymidine and attachment to DNA. Phosphoramidite was not isolated and coupling to DNA was achieved directly with the crude mixture as described in Section 5.7.4. One coupling was accomplished successfully but several TBAF treatments prevented couplings. **Nap** is the naphthalene containing phosphoramidite monomer of Chapter 3. This hydrophobic moiety helped for reverse-phase high performance liquid chromatography (RP-HPLC) analysis and purification (**Figure 5.8**).

The levulinyl group (Lev) can be used as a protecting group in oligonucleotide synthesis.^{37,38} As shown by several reports, it can be orthogonally deprotected using hydrazine to make branched oligonucleotides.³⁹⁻⁴² We examined whether the levulinyl group could be used efficiently multiple times to grow two different strands simultaneously from a branching point using the same CPG. **HYD1** is a 47mer DNA strand synthesized according to standard DNA synthesis protocols, except that a hydrazine solution was injected 30 times during the synthesis. Yields were satisfactory but the synthesis led to a crude mixture containing multiple byproducts as observed by polyacrylamide gel electrophoresis (PAGE). We hypothesized that the main reason for the presence of byproducts was the use of acetyl protected cytidine phosphoramidite, which can be deprotected during hydrazine treatment.⁴³ We therefore performed the next tests with benzoyl-protected cytidine. Conveniently, 5'-levulinyl deoxynucleotide monomers are commercially available. We

synthesized two DNA strands: a 20mer and 14mer termed **LEV1** and **LEV2** respectively, by coupling 5' DMT-phosphoramidites followed by 4 5'-levulinyl phosphoramidite couplings (**Figure 5.3**). Steps involving trityl response allowed to visually assess the coupling efficiency during the synthesis. Both **LEV1** and **LEV2** were synthesized in high yields. Gel electrophoresis revealed that the synthesis was cleaner than **HYD1** further demonstrating that acetyl protected cytidine was the main issue encountered during initial tests. 5'-levulinyl monomers of **LEV2** were deprotected using 0.5 M hydrazine solution as reported elsewhere,⁴² while we used higher concentrations of hydrazine for **LEV1**. Mass spectrometry (MS) and gel electrophoresis analyses revealed that the crude mixture of **LEV1** contained species with higher masses than the expected product. We hypothesized that nucleobase deprotection of A, C or G may still happen during the synthesis leading to their phosphitylation followed by the formation of branched byproducts. On the contrary, no byproduct was identified in the analysis of **LEV2**. Further optimization showed that deprotection time could be lowered from 32 to 6 minutes. These results convinced us that the levulinates are good orthogonal protecting groups for the parallel synthesis of branched oligonucleotides on solid phase.

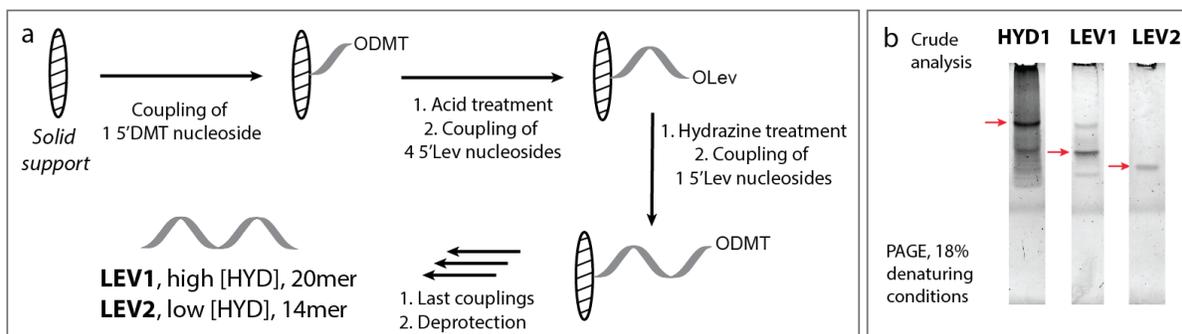
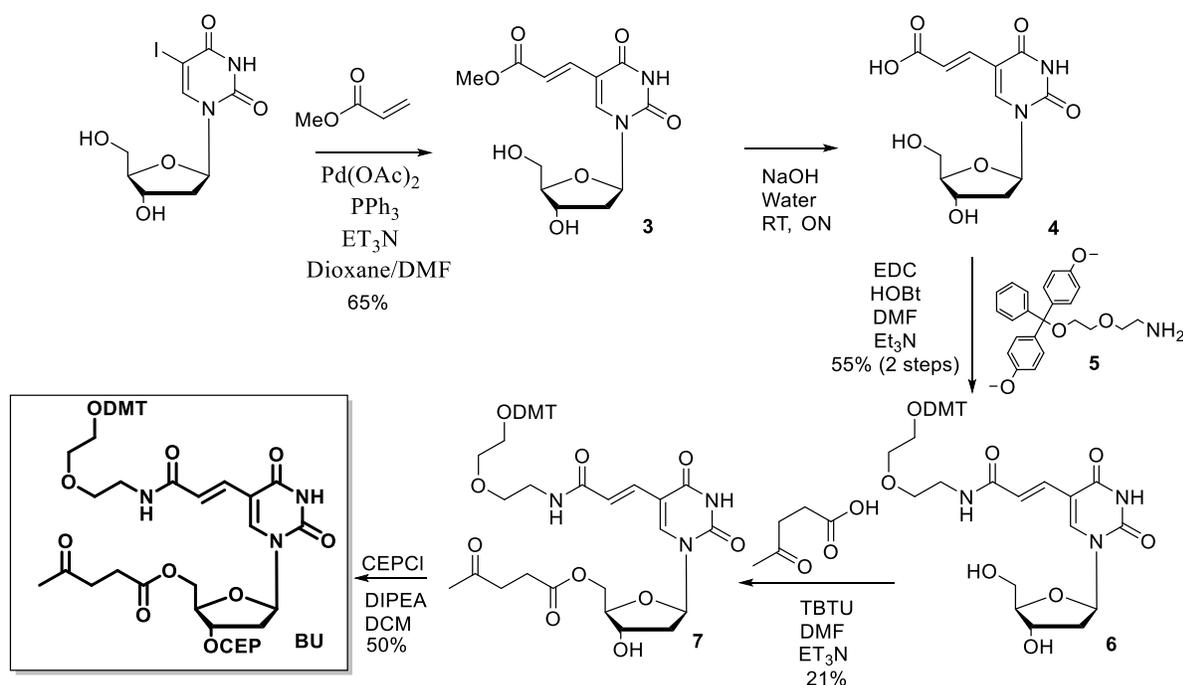


Figure 5.3. Synthesis of DNA strands with 5'-Lev and 5'-DMT phosphoramidites. (a) Synthesis of **LEV1** and **LEV2** and (b) Crude mixture analysis of **HYD1** (45mer, 30 hydrazine (0.5 M) treatments, no 5'-Lev amidites, acetyl protected cytidines), **LEV1** (20mer, 15 hydrazine (1.5 M) treatments, 15 5'-Lev amidites, benzoyl-protected cytidines), **LEV2** (14mer, 10 hydrazine (0.5 M) treatments, 10 5'-Lev amidites, benzoyl-protected cytidines). There is a faint higher mobility band visible in every lane that comes from experimental conditions (see section 5.7.6).

5.5.2 Branching unit and branched oligonucleotides

We chose to use DMT-protected monomers for the aptamer branch. This choice was motivated by the possibility to visually gauge the more challenging unnatural monomer couplings efficiency thanks to the DMT cation orange color. The other part would be 5'-levulinyl protected for growing

the code. For the branch point, we designed a deoxyuridine modified with an acrylamido side chain at the C5 position (**BU**) similar to the ones used by the Cook and Liu groups (**Scheme 5.2**).^{13,44} These groups showed that this modification still allows efficient DNA hybridization and can even be bypassed by some polymerases efficiently. Starting with the Mizoroki-Heck coupling of iodouridine with methyl acrylate, the ester obtained was deprotected and underwent an amine coupling with DMT protected 2-(2-amino)ethoxyethan-1-ol. The protection of the 5'-OH of **3** with a levulinyl group was then performed to yield compound **7** which was finally turned into phosphoramidite **BU**.

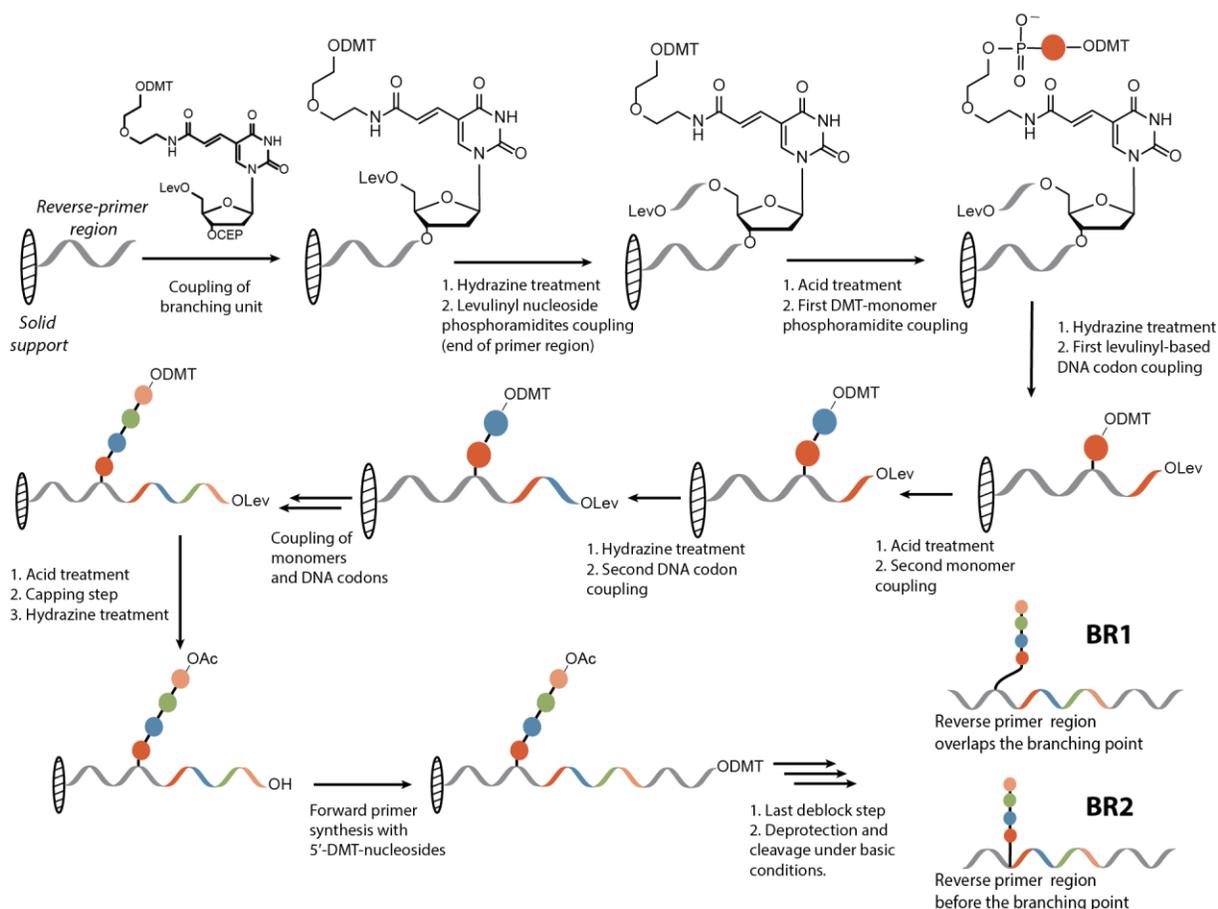


Scheme 5.2. Synthesis of branching unit **BU**. CEP=cynoethylphosphoramidite.

Synthesis of two model branched strands, **BR1** and **BR2** was first examined by exclusively using DNA nucleotides (**Scheme 5.3**). **BR1** is designed so that the reverse primer would bind over the branching unit, while for **BR2**, the whole reverse primer would hybridize at the 3' end of the branching unit. The “aptamer branch” for these strands is a DNA 4mer. Horn and coworkers recommended the attachment of the branching unit on a growing chain on CPG after the 20th nucleotide⁴⁰ to prevent steric hindrance close to the solid support. Therefore, **BR1** synthesis started with a 9-thymidine linker followed by 12 nucleotides forming the main part of the reverse primer region of **BR1**. **BU** was then incorporated efficiently. Four nucleotides belonging to the reverse

primer region of the code were added at the 5' end of **BU** using 5'-levulinyl chemistry. On the branch, the aptamer was synthesized with 5'-DMT phosphoramidites. After each 5'-DMT monomer insertion, DMT was kept on while a codon was synthesized on the code branch. We used 2-nucleotide codons at this point as a proof of concept. This resulted with a DNA 4mer as the aptamer and a DNA 8mer as the code. The last aptamer nucleotide was deprotected and intentionally capped with usual capping reagents.

Initial tests have shown the resistance of acetyl ester to hydrazine treatment. Indeed, during a regular 19mer synthesis, the 5'end was capped and underwent multiple hydrazine treatments. Further couplings failed while trityl response was optimal in the case of an uncapped control. Therefore, the 5' end of the code could be deprotected with hydrazine and elongated with the forward primer region (11 nucleotides) using DMT chemistry. Thanks to the trityl response, this strategy allowed to visually observe if the code synthesis had been efficient. **BR2** synthesis started with the entire 17-mer primer region, followed by **BU** and similar aptamer and code than **BR1**. Yields were low (<10 %) according to gel electrophoresis image analysis but high enough to isolate **BR1** and **BR2**. Subsequent syntheses detailed thereafter have shown higher yields, possibly due to improved phosphoramidite dryness. After gel purification, we confirmed the formation of the expected branched oligonucleotide by liquid chromatography coupled to mass spectrometry.



Scheme 5.3. Parallel synthesis of branched oligo(phosphodiester)s.

We next synthesized a new design, **BR3** to further show the applicability of our strategy to longer strands. **BR3** is similar to **BR1** except that the aptamer branch contains 10 nucleotides instead of 4 and the code is made with $10 \times 2 = 20$ nucleotides instead of 8. Again, we capped the aptamer branch at the end of its synthesis and continued the code with 5'-DMT phosphoramidites for the forward primer region synthesis. We left the DMT on at the 5' end of the code branch to get an insight into the byproducts formed. After deprotection from the solid support, the crude product was analyzed by reverse-phase high-performance liquid chromatography (RP-HPLC), gel electrophoresis and mass spectrometry (**Figure 5.4**). The expected product was obtained in yields between 10 (determined by gel electrophoresis image analysis) and 22 % (determined by HPLC chromatogram analysis). While this number can be further improved, it allowed the isolation of **BR3** through gel purification. The DMT-on strategy revealed that most byproducts do not have the full DNA code on. In other words, they are the result of failed couplings on the code part.

Through MS, we identified the DNA strand before **BU** coupling and the strand comprising the aptamer and the 3' DNA section but missing the code. We hypothesize the branching unit coupling as well as the first levulinyl phosphoramidite coupling are low yielding. Longer coupling and deprotection times for these phosphoramidites were therefore implemented. Some species with longer retention times than the expected DMT-on product were also observed. Mass spectrometry and gel electrophoresis revealed higher masses and lower mobility than **BR3** (**Figure 5.4c**), indicating that byproducts of larger size and probably containing another or several other DMT groups are present. Similarly to **LEV1**, we hypothesize that hydrazine deprotects small amounts of nucleobases leading to the formation of minor oligomeric byproducts with several branches.

In summary, the successful synthesis of **BR1**, **BR2** and **BR3** showed that our strategy to grow two different strands simultaneously is successful. From **BU**, the aptamer and the code could be synthesized in parallel. Moreover, careful analyses of the byproducts formed allowed to understand some synthetic drawbacks and paved the way to further improvement of yields.

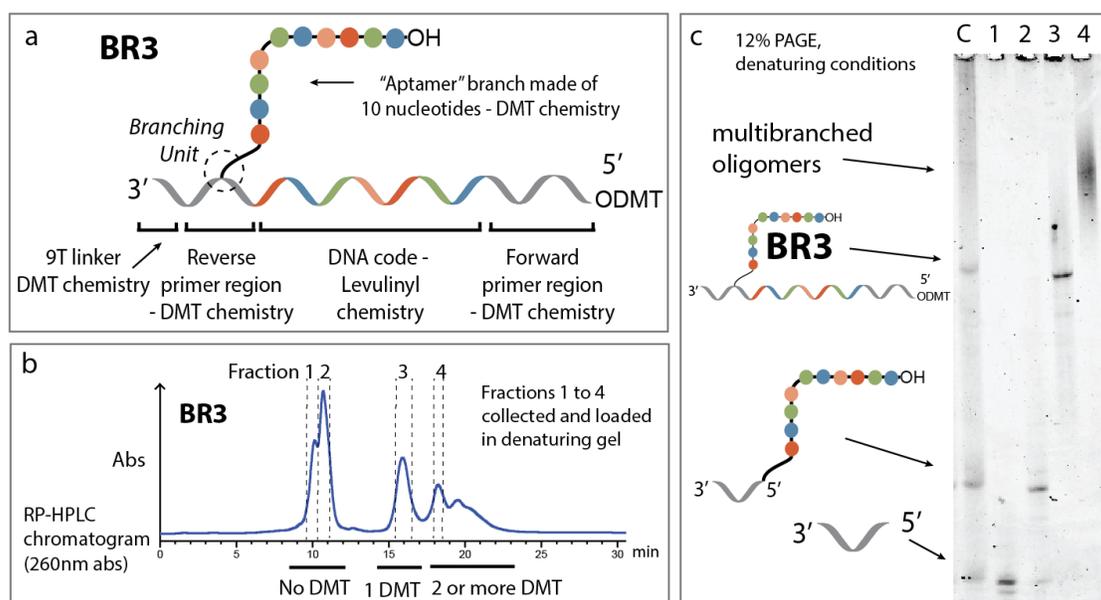


Figure 5.4. Analysis of **BR3** synthesis. (a) Schematic representation of **BR3**, (b) RP-HPLC of crude mixture obtained from synthesis of **BR3** with DMT-on, (c) gel analysis of crude mixture (lane C) and HPLC fractions (lanes 1 to 4). Coupled to MS results, this in-depth analysis revealed the nature of the synthesis' main byproducts. Detailed experimental procedures are available in the experimental section.

5.5.3 Aptamer design

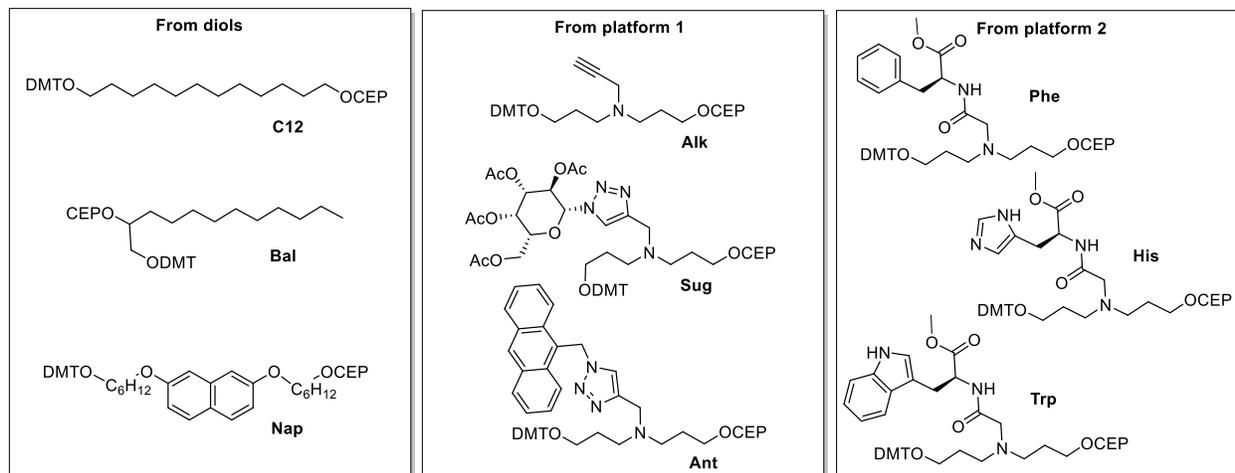
Contrary to aptamers developed through SELEX and other modified aptamer synthesis, we use non-nucleosidic modifications. These non-nucleosidic modifications prevent amplification using enzymatic reactions. As such, the selection step would need to be followed by DNA-code amplification and sequencing. At this point, binding sequences can be synthesized efficiently without the code through automated SPS and tested. If desired, another round of selection could then be set up. Indeed, sequences of interest can be “evolved” manually or using an appropriate software and a new library synthesized through automated SPS for the next selection round. However, this makes *in vitro* selection cycles potentially time-consuming to implement. This drawback can be compensated with two strategies: (i) the use of very large libraries combined with the use of stringent selection conditions and (ii) the semi-rational design of the library components. The first strategy (i) is beyond the scope of this chapter but is conceptually possible to apply; therefore, we focused on (ii) by reproducing some major features of the thrombin-binding aptamer to increase our chances of finding a non-nucleosidic sequence that binds to thrombin with high affinity.

Thrombin is a protein involved in coagulation mechanisms. The thrombin binding aptamer (TBA) with sequence 5'-GG-TT-GG-TGT-GG-TT-GG-3' (TBA) is a G-quadruplex-forming 15mer. It was first discovered using SELEX in 1992 by Toole and coworkers.³³ Several reports then described binding interactions notably through TBA-thrombin structure elucidation with nuclear magnetic resonance (NMR)^{45,46} and X-Ray Diffraction.⁴⁷⁻⁴⁹ These studies revealed that TBA indeed folds into an intramolecular G-quadruplex with one TGT and two TT loops. TBA most probably interacts with exosite I of thrombin through its two TT loops. Being the main contact regions between thrombin and TBA, many scientists have reported the chemical modification of these loops.⁵⁰ The success of this strategy was evidenced by enhancement of binding affinities by more than 10-fold compared to the original aptamer.⁵¹ It is noteworthy that studies usually use one or two types of aptamer modifications. Moreover, their choice can be motivated by *in silico* design⁵² but most often, modifications are placed at partly random locations and each modified aptamer has to be studied individually. Within our work, we show the possibility of trying very large numbers of aptamers with several modifications at once. We decided to design our modified

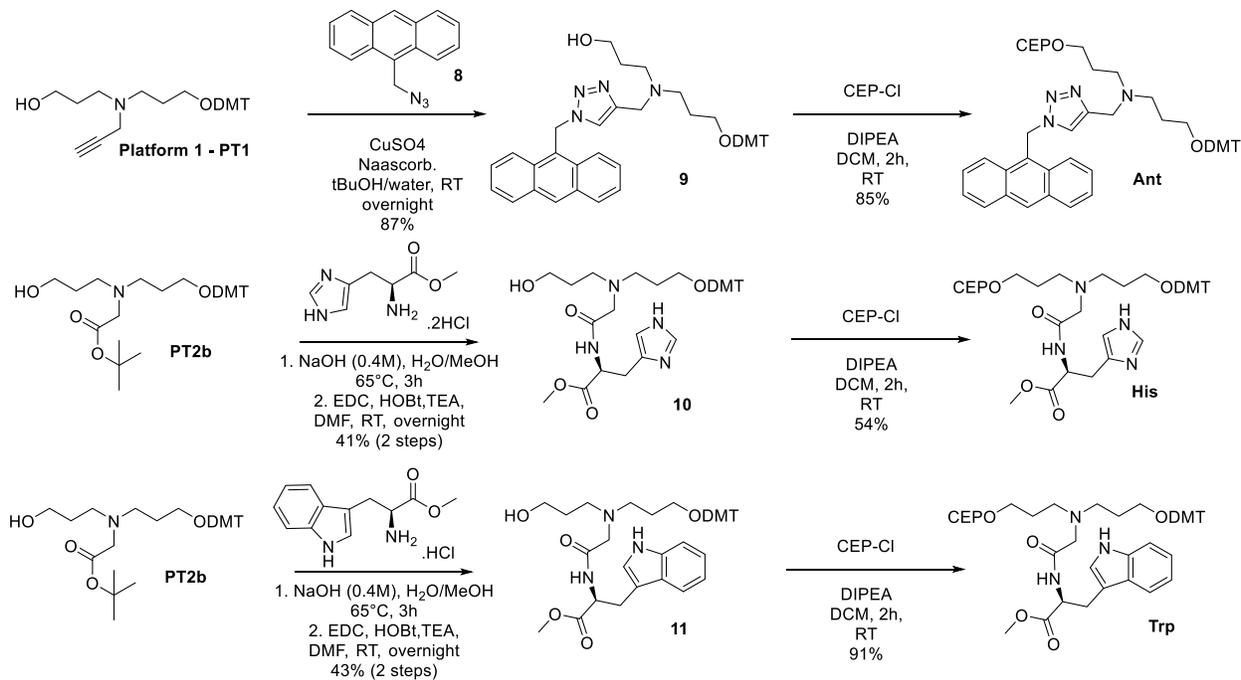
aptamer so that the G-tetrads stay intact while the loops would be modified with one or two monomers from our new library. All modifications of the loop are expected to influence the global G quadruplex stability. The TT loops are in close proximity to thrombin in TBA-thrombin complexes, hence modifications in these regions should greatly influence the binding affinity to thrombin. The 3' end of the aptamer was also modified since some reports showed a potential increase in stability at this position.⁵³ In this chapter we focus on the synthesis of a library of DNA-encoded thrombin-binding aptamer analogues. Testing of this library is part of a collaborative work in progress.

We chose the monomers shown in Scheme 5.4. Compatibility with the DNA synthesizer and the need for high coupling yields are restricting conditions for non-nucleosidic monomer design. In our case, the monomers were shown in the chapters 3 and 4 to fulfil the synthetic criteria. **C12** was chosen to improve hydrophobic interactions with thrombin. We were also initially interested in using a branched version of the **C12** monomer (**Bal**) that has been reported elsewhere to be efficiently incorporated in DNA strands (but not at the 3' and 5' ends).⁵⁴ However, **Bal** was only used in some of our initial studies due to the presence of a chiral carbon as a racemic mixture, which may make thrombin-binding studies more complex. Naphthalene has been shown to improve aptamer binding affinity to some targets.⁵⁵ Therefore, we thought **Nap** could be a good candidate for π - π stacking and hydrophobic interactions. All the other modifications are built from platform 1 and 2 described in Chapter 4. Thanks to the versatility of this method, modifications can be designed and synthesized at will following criteria dictated by the protein being studied. Moreover, the positive charge on the tertiary nitrogen at physiological pH could improve cellular uptake properties of our constructs, which is appealing for targeted therapy applications (aptamer-drug conjugates).² Again, we analyzed findings reported elsewhere^{8,9} which directed us to the phenylalanine modification (**Phe**) from Chapter 4 and to the synthesis of a histidine-like (**His**) and a tryptophan-like (**Trp**) modification (**Scheme 5.5**). These aromatic ring-containing modifications could create π - π stacking, polar and hydrogen-bonding interactions with thrombin. **Alk** was also used as a small moiety with a positive charge that can be turned into a functional group through click chemistry after SPS. Finally, the carbohydrate-containing (**Sug**) and anthracene (**Ant**) modifications were used to expand the structural diversity of aptamer modifications. The first one is of interest because carbohydrates have been shown to interact tightly with some proteins (especially lectins). **Ant** has an extended aromatic system that could interact with several amino

acids in thrombin. However, **Ant** was used at only two positions in the library to avoid higher-order self-assembly of **Ant**-containing aptamers. **His**, **Trp** and **Ant** syntheses further demonstrate the versatility of the synthetic method described in Chapter 4 (**Scheme 5.5**).



Scheme 5.4. Phosphoramidite monomers used in this chapter. CEP=Cyanoethylphosphoramidite, DMT=dimethoxytrityl.



Scheme 5.5. Synthesis of **Ant**, **His**, **Trp**.

5.5.4 Synthesis of model compounds and the library

Before starting the synthesis of the actual library, we made three strands of known sequences with different numbers of unnatural nucleotides. **TBA1** is a strand made of DNA nucleotides only and whose branch is the original TBA, **TBA2** has an aptamer branch with three unnatural monomers, one in each loop of the TBA and **TBA3** is heavily modified with seven positions bearing a different moiety than on the original TBA sequence (**Figure 5.6**). As explained before, synthesis started with a 9-thymidine linker to put the branching unit away from the solid support. 11 nucleotides, part of the 3' reverse primer region, were then attached. The branching unit uridine and the 4 following nucleotides are also part of the reverse primer region on the code. Simultaneously, a 4-thymidine linker is attached on the aptamer branch to prevent interactions between the later formed aptamer and the code. To build the aptamer, a 3-nucleotide codon was first attached to the code followed by the associated monomer on the aptamer. Each GG section on the aptamer is not accompanied with code growth because these are part of the constant regions of the aptamer. In total, seven nucleotides or unnatural monomers were encoded meaning that the aptamer branch was 4 (linker) + 7 (unnatural monomers) + 8 (GG regions) = 20 monomers long while the code was 4 (end of reverse primer region) + 21 (7x3 nucleotides coding for the aptamer sequence) = 25 nucleotides long. Again, the aptamer branch synthesis was capped allowing the forward primer region to be grown with 5' DMT phosphoramidites. **TBA1**, **TBA2** and **TBA3** were synthesized in reasonable yields allowing the isolation of a few nanomoles of each. From gel image analysis, synthetic yields are estimated to be 22 % for **TBA1**, 16 % for **TBA2** and 10 % for **TBA3**. Mass spectrometry results confirmed the identity of the products. We noticed that all three strands have similar mobility shifts on gel showing that a library of constructs can be purified by gel electrophoresis techniques. In DNA-encoded libraries, full-length products are usually not separated from the crude mixture.²⁴ The negatively charged nature of our oligomers allows for easy purification.

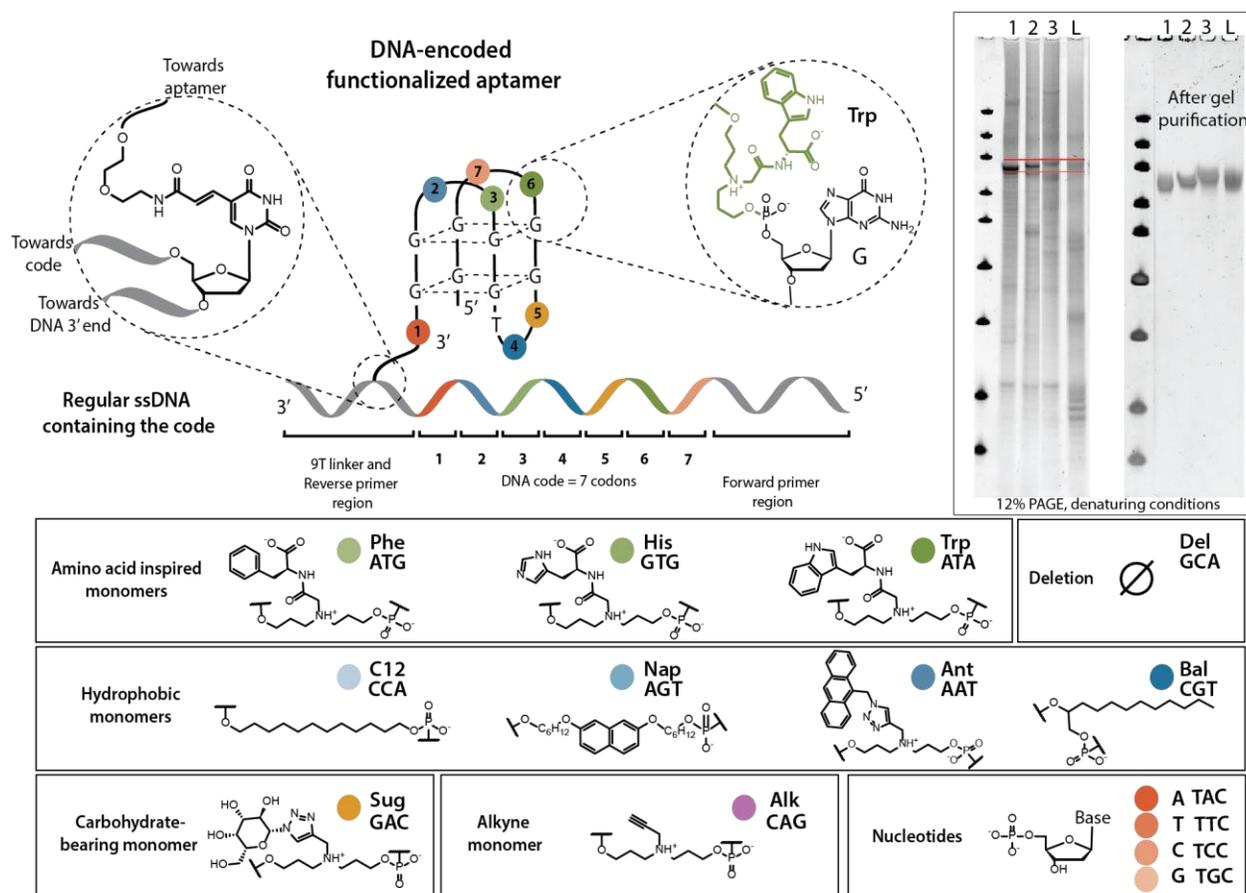
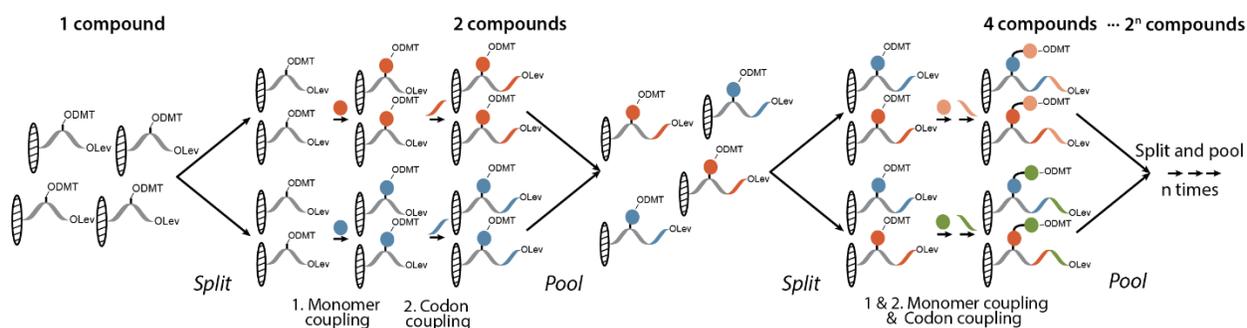


Figure 5.6. Design and synthesis of DNA-encoded thrombin-binding aptamers made of unnatural monomers. Lanes 1, 2, 3 and L in the gels were respectively loaded with crude mixtures (left) and purified strands (right) of **TBA1**, **TBA2**, **TBA3** and **LIB**. Aptamer branch: **TBA1**: 5'-GGTTGGTGTGGTTGGTTTTT-3', **TBA2**: 5'-GGHisTGGC12GTGGAlkTGGTTTTT-3', **TBA3**: 5'-GGTrpSugGGPheAntTGGBalAGGNapTTTTT-3'. **LIB** is a library of DNA-encoded aptamers with seven positions modified. Monomers name, structure and codons are shown in the bottom.

After the successful syntheses of **TBA1**, **TBA2**, and **TBA3**, we applied the split-and-pool strategy to make **LIB**, a library of DNA-encoded aptamers (**Scheme 5.6**). When reaching position n of the aptamer, the solid support is split into x aliquots. The determination of equivalent amounts of CPG in each aliquot was performed visually to avoid CPG spilling. This induced a small bias: some sequences may be slightly overrepresented, and others may be underrepresented. In our case, 6 different modifications were used at each cycle because it fits our MerMade 6 synthesizer ($x=6$). Each solid support fraction underwent an unnatural monomer or nucleotide coupling on the aptamer branch. This step is followed by the attachment of the associated codon of 3 DNA nucleotides on the code branch. The solid support is then mixed together and split again for the coupling $n+1$. Attachments of the G bases, which are indispensable to keep the G quadruplex

backbone, are executed without split and pool nor DNA codon synthesis. In the end, the library should theoretically be composed of x^n members. In our case, we expected a $6^7=279,936$ members library. The library crude mixture was loaded on a gel for analysis and purification. We observed the presence of a relatively diffuse band corresponding to the mobility of **TBA1** to **TBA3** strands (**Figure 5.6**). Gel image analysis allowed to measure yields of synthesis of about 11 %. We isolated 2.5 nmols of the library using about 1/6th of the crude product obtained (despite the spilling of some CPG during the split-and-pool steps). This amount will allow to perform affinity tests towards thrombin.



Scheme 5.6. Split and pool strategy for the synthesis of a DNA-encoded unnatural aptamer library. We show an example with 2 different monomers at each step for clarity. In this study, we couple 6 different monomers at each step yielding 6^n compounds.

5.5.5 PCR and sequencing

To show the possible sequencing of strands after selection, PCR amplification of the code of branched oligonucleotides was performed. For PCR, we used a forward primer that has the same sequence as the constant 5' region of the branched oligomers and a reverse primer that hybridizes the 3' region of our strands. We first showed that **BR1** and **BR2** were amplified efficiently using a Taq polymerase through hot-start PCR. This demonstrated that the strategies where the reverse primer binds across the branching unit (in **BR1**) and where it binds before the branch (in **BR2**) are both suitable for our study (**Figure 5.7**). We thought the design of **BR1** had better chances to be amplifiable in the case of more complex aptamer branches and decided to apply it to all other strands made. The strands **BR3**, **TBA1** and **TBA2** were shown to amplify efficiently using the same polymerase starting from $\sim 2 \times 10^{-16}$ moles of template. During the first round of PCR, we expect the polymerase to produce the complement of the templates code branch from the reverse primers (**Figure 5.7a**). From round 2, both template-like sequence and its complementary

counterpart should be copied in an exponential way. **TBA3** is the strand containing the most unnatural monomers. To show the possibility of amplifying the lowest possible outcome of selected strands from a library of compounds, the amplification of **TBA3** code was performed using $\sim 2 \cdot 10^{-20}$ moles and was shown to be successful (**Figure 5.7b**).

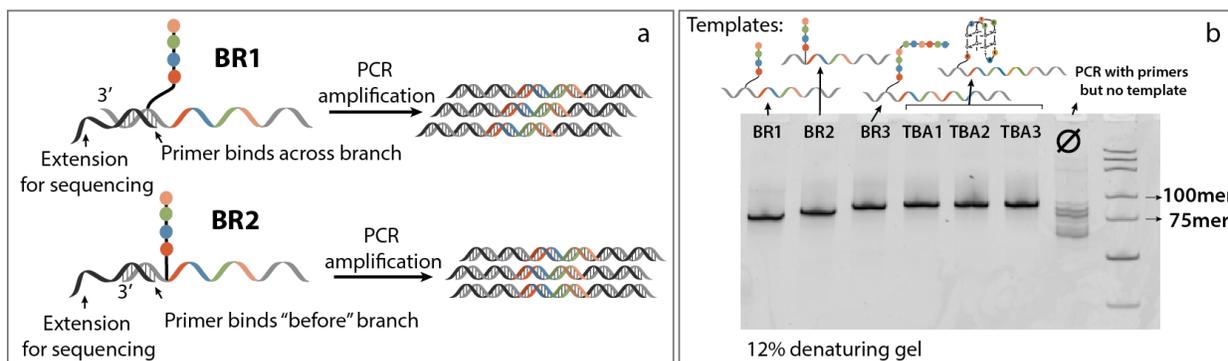


Figure 5.7. PCR with DNA-encoded aptamers. (a) Schematic representation of reverse primer binding **BR1** and **BR2**. Primers needed for sequencing have a 22mer tag on both sides. (b) Amplicons obtained after PCR from the 6 different templates and a control with no template on a denaturing gel.

The amplicons obtained were indexed and sequenced through MiSeq next-generation sequencing according to standard procedures. This technology has been used for the analysis of modified aptamers pools after selection cycles.¹² Each sequence was read between 20,000 and 60,000 times. The results are summarized in Table 5.1. For **BR1** and **BR2**, we observed that more than 94 % of the reads started with the forward primer region sequence. We deduced an average error rate per base of ~ 0.5 % which is standard for SPS-made oligomers.⁵⁶ We also counted the number of times the correct code sequence was present and deduced an average error rate between 0.5 and 0.7 % for both strands. It shows that, for such templates, having the reverse primer region before or across the branching unit does not lead to a significant difference in amplification and sequencing fidelity. For **BR3**, **TBA1**, **TBA2** and **TBA3**, the expected sequence code was found in between 69 and 87 % of the reads showing that our templates are suitable for PCR and sequencing. In particular, the error rate obtained using **TBA3** is in the same range than the other TBA strands whereas it contained 6 unnatural monomers and was amplified from very low concentrations of templates. The error rates are slightly higher in the code region than in the primer region but values are still in the range of SPS made oligonucleotides. Future work will examine the use of high-fidelity polymerases to reduce the error rates obtained in our sequencing results.

Table 5.1. Sequencing of amplicons from **BR1**, **BR2**, **BR3** and **TBA1** to **TBA3** templates.

Amplicon	Number of reads	Percentage of expected primer sequence ^a (%)	Average error rate per base for primer sequence ^b (%)	Percentage of expected code sequence (%)	Average error rate per base for code sequence ^b (%)
BR1	28,652	94	0.53	96	0.54
BR2	34,146	95	0.49	95	0.68
BR3	36,860	94	0.35	87	0.69
TBA1	23,325	93	0.45	75	1.4
TBA2	34,456	93	0.43	69	1.7
TBA3	58,347	93	0.41	75	1.4

a. For **BR1** and **BR2**, we searched for the forward primer sequence. For the other strands, we searched for the reverse primer sequence as explained in Section 5.7.9. b. Approximation obtained from $1 - (\% \text{expected sequence})^{(1/n)}$, n being the length of the sequence examined.

5.6 Conclusions

In conclusion, we reported a successful synthetic method to generate DNA-encoded unnatural sequence-defined polymers by solid-phase synthesis. This achievement allowed the synthesis of a 300,000 member library of thrombin-binding aptamers with up to seven unnatural modifications. At very low concentrations, the DNA code of an artificial aptamer could be amplified and sequenced, further confirming the validity of our design. The monomers accessible through the method described in Chapter 4 may greatly enhance the type and number of interactions that these novel aptamers would form with their target. Moreover, they are expected to induce better serum stability and cell-internalization properties compared to regular oligonucleotide aptamers. The use of our non-nucleosidic modifications in aptamers is unprecedented, and our strategy is therefore an important contribution to the discovery of aptamers made of modified bases. Future work will involve testing our library in thrombin binding tests in order to identify potential aptamers with better affinity than TBA.

In the DNA-encoded library (DEL) context, our method is quite different from most of the synthetic routes reported earlier.²⁵ Indeed, in general, the DNA strand in DEL is synthesized through enzymatic ligation, while we propose to use solid-phase synthesis. Such a synthetic route

can broaden the chemical space explored by DEL, because it is carried out in non-aqueous solvents and under an inert gas atmosphere. While we base this study on already numerous monomers developed in the other chapters of this thesis, one can use even more phosphoramidites from the plethora that are commercially available. Importantly as well, solid-phase synthesis allows the use of other non-DNA compatible organic coupling reactions, in contrast to traditional DEL.

Finally, this study highlights the potential of combinatorial strategies to help designing potent target-binding sequences with unnatural sequence-defined polymers. We showed the potential of DEL of sequence-defined oligo(phosphodiester)s in the context of potential ligand discovery. Our method may be extended to other applications, such as the discovery of ribozyme-like catalysts. New monomers would also certainly allow to explore new reactions and expand the scope and efficiency of these catalysts.

5.7 Experimental section

5.7.1 Chemicals

All starting materials were obtained from commercial suppliers and used without further purification unless otherwise noted. Acetic acid, Boric acid, solvents were purchased from Fisher Scientific. 4,4'-(chloro(phenyl)methylene)bis(methoxybenzene) (DMT-Cl) and (3-Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDC-Cl) were purchased from AK Scientific. O-(Benzotriazol-1-yl)-N,N,N',N'-tetramethylammonium tetra-fluoroborate (TBTU) was purchased from Oakwood Chemicals. Chloroform-*d*1 was purchased from Cambridge Isotope Laboratories. Importantly, it was stored on molecular sieves in order to keep it neutral. If used as sold, hydrolysis of phosphoramidite (fast) as well as DMT deprotection (slow) may be observed. GelRed™ nucleic acid stain was purchased from Biotium Inc. Concentrated ammonium hydroxide, ammonium persulfate, acrylamide/Bis-acrylamide (40 % 19:1 solution) and tetramethylethylenediamine (TEMED) were obtained from Bioshop Canada Inc. and used as supplied. 1 μmol Universal 1000Å LCAA-CPG supports, dry packs, activator solution, 5'-DMT-phosphoramidites used for automated DNA and RNA synthesis were purchased through Bioautomation. Universal 2000Å LCAA-CPG, N,N-diisopropylamino cyanoethyl phosphonamidic-chloride (CEP-Cl) and 5'-levulinyl phosphoramidites were purchased from

Chemgenes. Sephadex G-25 (super fine, DNA grade) and 5'-6-carboxyfluorescein phosphoramidite were purchased from Glen Research. Invitrogen Ultra Low Range ladder was obtained from ThermoFisher Scientific. MyTaq™ HS Red Mix was purchased from Froggabio. QIAquick PCR Purification Kit were obtained from Qiagen. All other reagents were obtained from Sigma-Aldrich. TEAA (triethylammonium acetate) buffer is composed of 50 mM TEA with pH adjusted to 8.0 using glacial acetic acid. TBE buffer is 90 mM Tris, 90 mM boric acid and 1.1 mM EDTA with a pH of 8.0. TAMg buffer is 40 mM Tris, 7.6 mM magnesium chloride and 1.4 mM acetic acid.

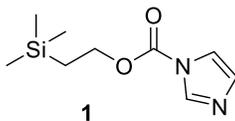
5.7.2 Instrumentation

Standard automated solid-phase synthesis was performed on a Mermade MM6 synthesizer from Bioautomation. HPLC purification was carried out on an Agilent Infinity 1260. DNA and oligomers quantification measurements were performed by UV absorbance with a NanoDrop Lite spectrophotometer from Thermo Scientific. Eppendorf Mastercycler 96-well thermocycler and Bio-Rad T100™ thermal cycler were used for PCR. PAGE experiments were carried out on a 20 X 20 cm vertical Hoefer 600 electrophoresis unit and Mini-PROTEAN electrophoresis units. Gel images were captured using a ChemiDoc™ MP System from Bio-Rad Laboratories. Dry solvents were taken from an Innovation Technology device. Low Resolution Mass determination was carried out using Electron-Spray Ionization – Ion Trap - Mass Spectrometry (MS) on a Finnigan LCQ Duo device. High Resolution mass determination was achieved using a Bruker Maxis API (Atmospheric pressure ionization) QTOF or a THERMO Exactive Plus Orbitrap-API. Liquid Chromatography Electrospray Ionization Mass Spectrometry (LC-ESI-MS) of oligomers was carried out using Dionex Ultimate 3000 coupled to a Bruker MaXis Impact™ QTOF. Some oxygen and moisture sensitive experiments were carried out in a Vacuum Atmospheres Co. glove box. Column chromatography was performed using a CombiFlash Rf system from Teledyne Isco. The NMR spectra were recorded on Bruker 400 MHz, 500 MHz, Varian 300 MHz or 400 MHz for ¹H, ¹³C and ³¹P with chloroform-*d*₁ (δ 7.26, ¹H; δ 77.16, ¹³C), acetone-*d*₆ (δ 2.04, ¹H; δ 29.8, ¹³C) or DMSO-*d*₆ (δ 2.50, ¹H; δ 39.5, ¹³C) as internal lock solvents and chemical shift standard.

5.7.3 Small molecule synthesis

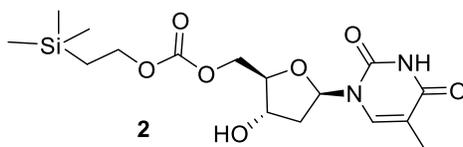
Protocols for **C12** and **Nap** are available in Chapter 3. Protocols for **Alk**, **Sug**, **Phe**, **PT2b** are available in Chapter 4.

5.7.3.1 2-(Trimethylsilyl)ethyl *1H*-imidazole-1-carboxylate (**1**)



This compound was made according to the protocol described by Blizzard and coworkers³⁶ and was fully characterized in the same report. We obtained 3.78g of compound **1** from 2.30 g of 2-(trimethylsilyl)ethan-1-ol. Yields : 91 %.

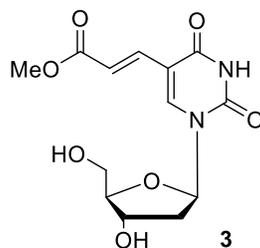
5.7.3.2 ((2*R*,3*S*,5*R*)-3-Hydroxy-5-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2*H*)-yl)tetrahydrofuran-2-yl)methyl (2-(trimethylsilyl)ethyl) carbonate (**2**)



Protocol inspired by Blizzard and coworkers.³⁶ Thymidine (1.52 g, 6.28 mmol, 1 equiv.) and 1,8-Diazabicyclo[5.4.0]undec-7-ene (DBU, 0.19 mL, 1.26 mmol, 0.2 equiv.) were loaded sequentially to a solution of **1** (1.33 g, 6.28 mmol, 1 equiv.) in 35 mL of DMF in a dry round bottom flask. The mixture was stirred for 22 h at 25 °C and then cooled in an ice bath as 10 mL of 0.1 N HCl was added. The aqueous phase was then extracted with ethyl acetate. The combined extracts were dried (MgSO₄), filtered, and evaporated. The residue was purified by chromatography on silica gel with slow gradient of DCM/Methanol (0-5 %) in about 5 column volumes (CV) to obtain the product as a white solid. 428 mg, 18 %.

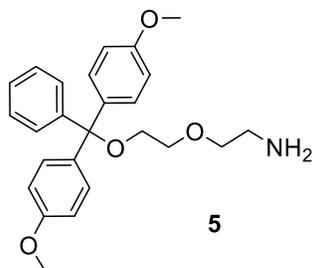
¹H NMR (500 MHz, DMSO-*d*₆): 11.31 (s, 1H), 7.45 (d, *J*=1Hz, 1H), 6.18 (t, *J*=7Hz, 1H), 5.43 (d, *J*=4Hz, 1H), 4.31-4.16 (m, 5H), 3.93-3.90 (m, 2H), 2.19-2.07 (m, 2H), 1.78 (d, *J*=1Hz, 3H), 1.01-0.97 (m, 2H), 0.02 (m, 9H).

5.7.3.3 Methyl (E)-3-(1-((2R,4S,5R)-4-Hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)acrylate (3)



Protocol inspired by Liu and coworkers.¹³ To a solution of 5'-deoxy-5-iodouridine (2.50 g, 7.06 mmol, 1 equiv.), methyl acrylate (1.22 g, 14.1 mmol, 2 equiv.), triphenylphosphine (370 mg, 1.41 mmol, 0.2 equiv.), and triethylamine (1.97 mL, 14.1 mmol, 2 equiv.) in a mixture of dry DMF (25 mL) and dry dioxane (25 mL) was added Pd(OAc)₂ (159 mg, 0.706 mmol, 0.1 equiv.) at 25 °C under Ar current. The mixture was heated to 90 °C and stirred for 16 h. The reaction mixture was evaporated under reduced pressure to remove the dioxane followed by coevaporation with toluene to dryness. The residue was purified by column chromatography (SiO₂, DCM/10 % methanol in DCM= 0 to 70 % in 7CV) to give the compound **3** (1.39 g, 4.45 mmol, 64 % yield) as a white solid. Spectroscopic data matched those for previously reported compound in the literature.⁵⁷

5.7.3.4 2-(2-(Bis(4-methoxyphenyl)(phenyl)methoxy)ethoxy)ethan-1-amine (5)

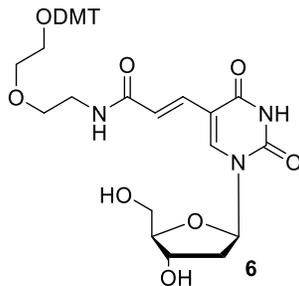


To a solution of 2-(2-aminoethoxy)ethan-1-ol (1.50 mL, 14.95 mmol, 1 equiv.) and triethylamine (6.06 mL, 44.86 mmol, 3 equiv.) in dry DCM (75 mL) was slowly added DMT-Cl (10.13 g, 29.9 mmol, 2 equiv.) on ice. The solution was left under stirring at room temperature under Ar for 2 hours. Solvent was evaporated under reduced pressure, residue was resuspended in DCM and ethylthiotetrazole (~1.4 g) was added until a pale pink/orange color appears. Solution was left under stirring for a few minutes. Product was extracted twice with DCM from 10 % Na₂CO₃ aqueous solution, washed once with a 10 % Na₂CO₃ aqueous solution, dried over MgSO₄, filtered

and solvent was evaporated under reduced pressure (40 °C). The residue was purified by column chromatography (Solid loading on celite, SiO₂ pretreated with 0.1 %TEA in DCM, DCM/TEA (100:0.1) and DCM/MeOH/TEA (90:10:0.1)) to give the compound **5** (4.16 g, 10.2 mmol, 68 % yield) as a yellow oil.

¹H NMR (500 MHz, CDCl₃): δ (ppm) 7.46 (d, *J*=8Hz, 2H), 7.36-7.18 (m, 7H), 6.82 (d, *J*=9Hz, 4H), 3.78 (s, 6H), 3.64 (t, *J*=5Hz, 2H), 3.55 (t, *J*=5Hz, 2H), 3.23 (t, *J*=5Hz, 2H), 2.90 (t, *J*=5Hz, 2H), 2.30 (bs, 2H).

5.7.3.5 (E)-N-(2-(2-(Bis(4-methoxyphenyl)(phenyl)methoxy)ethoxy)ethyl)-3-(1-((2R,4S,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)acrylamide (6**)**



To a 1M NaOH aqueous solution was added **3** (1.39 g, 4.45 mmol, 1 equiv.). The solution was left under stirring overnight, pH was adjusted to 7 with 1M HCl aqueous solution and water was evaporated under reduced pressure (60 °C). Residue was resuspended in acetonitrile and solvent was evaporated under reduced pressure (60 °C). This step was repeated once. Residue was suspended in dry DMF (20 mL). Anhydrous 1-hydroxybenzotriazole (HOBt) (782 mg, 5.79 mmol, 1.3 equiv.) and N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDC-Cl) (1.11 g, 5.79 mmol, 1.3 equiv.) were successively added. The pale yellow solution was left under stirring for 5 minutes until complete dissolution of EDC-Cl. Compound **5** (1.90 g, 4.67 mmol, 1.05 equiv.) and triethylamine (2.48 mL, 17.8 mmol, 4 equiv.) were then successively added. The reaction mixture slowly turned cloudy and was left under vigorous stirring overnight. Product was extracted twice with EtOAc from 10 % Na₂CO₃ aqueous solution, washed once with a 10 % Na₂CO₃ aqueous solution, dried over MgSO₄, filtered and solvent was evaporated under reduced pressure (40 °C). Toluene was used as cosolvent for complete evaporation. The crude product was loaded on celite and purified using the combiFlash system with a “Gold” column pretreated with 0.1 % TEA in

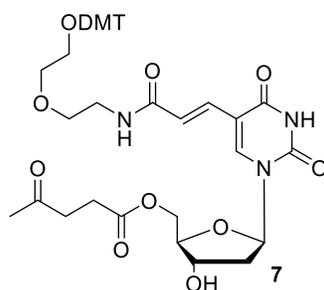
DCM. DCM/TEA (100:0.1) and DCM/MeOH/TEA (90:10:0.1) were used in a gradient from 0 to 60 % in ~8 CV. A pale yellow solid was isolated (1.74 g, 2.53 mmol, 56 %).

HRMS (ESI-QTOF) m/z : $[M + Na]^+$. Calcd for $C_{37}H_{41}N_3O_{10}Na$ 710.26842; Found 710.26991.

1H NMR (500 MHz, DMSO-*d*₆): δ (ppm) 11.55 (s, 1H), 8.29 (s, 1H), 7.40-7.38 (m, 2H), 7.30 (t, $J=8$ Hz, 2H), 7.27-7.19 (m, 5H), 7.16-7.02 (m, 2H), 6.88 (d, $J=9$ Hz, 4H), 6.16 (t, $J=7$ Hz, 1H), 5.26 (d, $J=4$ Hz, 1H), 5.16 (t, $J=5$ Hz, 1H), 4.28-4.26 (m, 1H), 3.82-3.80 (m, 1H), 3.72 (s, 6H), 3.68-3.64 (m, 1H), 3.61-3.56 (m, 3H), 3.49 (t, $J=6$ Hz, 2H), 3.35-3.32 (m, 2H), 3.06 (t, $J=5$ Hz, 2H), 2.20-2.12 (m, 2H).

^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) 165.8, 161.8, 158.0, 149.3, 145.0, 142.5, 135.8, 132.3, 129.7, 127.8, 127.7, 126.6, 121.4, 113.2, 109.1, 87.6, 85.3, 84.6, 70.0, 69.5, 69.2, 62.8, 61.0, 55.0.

5.7.3.6 ((2R,3S,5R)-5-(5-((E)-3-((2-(2-(Bis(4-methoxyphenyl)(phenyl)methoxy)ethoxy)ethyl)amino)-3-oxoprop-1-en-1-yl)-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-3-hydroxytetrahydrofuran-2-yl)methyl 4-oxopentanoate (7)



Protocol inspired from Damha and coworkers.⁵⁸ O-(Benzotriazol-1-yl)-N,N,N',N'-tetramethylammonium tetra-fluoroborate (TBTU, 532 mg, 1.66 mmol, 1 equiv.) was dissolved in DMF (2 mL) and N,N-diisopropylethylamine (1.15 mL, 6.63 mmol, 4 equiv.) and stirred at room temperature under Ar in a dry round bottom flask. Freshly distilled levulinic acid (0.17 mL, 1.7 mmol, 1 equiv.) was added, and the reaction was stirred at room temperature for 30 min. Compound **6** (1.14 g, 1.66 mmol, 1 equiv.) was dried under vacuum in a round-bottom flask with a magnetic stirrer. The TBTU/levulinic acid reaction mixture was cannulated to the nucleoside, and the reaction was stirred at room temperature. After 10 h, product was extracted twice with DCM from a saturated $NaHCO_3$ aqueous solution, washed twice with brine, dried over $MgSO_4$,

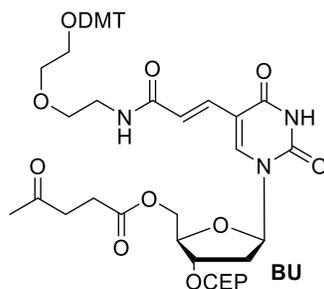
filtered and solvent was evaporated under reduced pressure (40 °C). Toluene was used as cosolvent for complete evaporation. The crude product was loaded on celite and purified using the combiFlash system with a SiO₂ “Gold” column pretreated with 0.1 % TEA in DCM. DCM and DCM/MeOH (9:1) were used in a gradient from 0 to 50 % in ~8 CV. A pale yellow solid was isolated (271 mg, 0.345 mmol, 21 %). NB: Addition of TEA in the solvents resulted in poor peak resolution. Nucleoside deprotected with levulinyl and 3’Lev protected were found to have similar shorter retention times than the product.

HRMS (APCI-QTOF) *m/z*: [M + Cl]⁻ Calcd for C₄₂H₄₇N₃O₁₂Cl 820.28537; Found 820.28403.

¹H NMR (500 MHz, CDCl₃): δ (ppm) 9.63 (s, 1H), 7.84 (s, 1H), 7.41-7.39 (m, 2H), 7.30-7.14 (m, 10H), 6.79 (d, *J*=9Hz, 4H), 6.23 (t, *J*=6Hz, 1H), 4.56-4.55 (m, 1H), 4.47-4.44 (m, 1H), 4.26-4.23 (m, 1H), 4.11-4.09 (m, 2H), 3.74 (s, 6H), 3.73-3.53 (m, 6H), 3.27 (t, *J*=5Hz, 2H), 2.89-2.83 (m, 1H), 2.79-2.73 (m, 1H), 2.64-2.58 (m, 1H), 2.55-2.45 (m, 2H), 2.25-2.20 (m, 1H), 2.14 (s, 3H).

¹³C NMR (126 MHz, CDCl₃): δ (ppm) 207.8, 173.2, 166.8, 162.1, 158.5, 149.1, 145.0, 142.0, 136.3, 132.7, 130.1, 128.3, 127.9, 126.9, 122.8, 113.2, 110.3, 86.3, 85.5, 84.7, 70.8, 70.7, 70.50, 63.1, 55.3, 40.6, 39.6, 38.2, 29.8, 28.2.

5.7.3.7 Branching Unit (BU)



Compound **7** was suspended in acetonitrile and solvent was evaporated under reduced pressure (60 °C). The operation was repeated once and the dried compound was kept under high vacuum for at least 5 hours. In an oven-dried flask, compound **7** (271 mg, 0.345 mmol, 1 equiv.) was then dissolved in anhydrous DCM (5 mL) and 0.30 mL of dry DIPEA (1.7 mmol, 5 equiv.) were added under stirring. CEP-Cl (0.12 mL, 0.52 mmol, 1.5 equiv.) was added slowly and the reaction was allowed to stir under inert gas at room temperature for 2 hours. Two fast extractions with DCM from a NaHCO₃ (sat.) aqueous solution were performed. Organic fractions were combined, dried

over MgSO₄, filtered and the solvent was evaporated under reduced pressure (40 °C). The crude product was loaded on celite and purified using the combiFlash system with a SiO₂ “Gold” column pretreated with 1 % TEA in DCM. DCM/TEA (100:1) and DCM/MeOH/TEA (90:10:1) were used in a very slow gradient from 0 to 10 % in ~10 CV. A white solid was isolated (170 mg, 0.172 mmol, 50 %). NB: an impurity thought to be 2-cyanoethyl *N,N*-diisopropylphosphoramidate (³¹P NMR signal at 14 ppm) was found to coelute with the product. Purification conditions involving toluene or hexanes and ethyl acetate were not helpful. In the very slow gradient described, the impurity elutes slightly before the product.

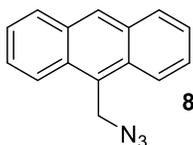
HRMS (ESI-QTOF) *m/z*: [M + Na]⁺ Calcd for C₅₁H₆₄N₅O₁₃PNa 1008.4130; Found 1008.4155.

¹H NMR (500 MHz, CDCl₃): δ (ppm) 7.84-7.82 (m, 1H), 7.41-7.39 (m, 2H), 7.35-7.13 (m, 10H), 6.79 (d, *J*=9Hz, 4H), 6.26-6.22 (m, 1H), 4.52-4.33 (m, 2H), 4.32-4.24 (m, 2H), 3.89-3.83 (m, 1H), 3.78-3.68 (m, 11H), 3.68-3.55 (m, 5H), 3.28 (t, *J*=5Hz, 2H), 2.91-2.83 (m, 1H), 2.80-2.73 (m, 1H), 2.69-2.51 (m, 5H), 2.25-2.20 (m, 1H), 2.15, (s, 3H), 1.19-1.17 (m, 12H).

¹³C NMR (126 MHz, CDCl₃): δ (ppm) 207.1, 207.0, 172.7, 172.6, 166.6, 162.0, 158.5, 148.9, 145.0, 141.8, 136.3, 132.4, 130.1, 128.3, 127.9, 126.8, 123.1, 117.7, 117.7, 113.2, 110.5, 110.4, 86.3, 85.7, 84.5, 73.0, 72.8, 72.6, 71.0, 70.7, 63.4, 63.3, 63.1, 58.4, 58.4, 58.3, 58.2, 55.3, 46.1, 43.5, 43.4, 40.3, 40.3, 39.5, 38.1, 29.8, 29.8, 28.2, 28.2, 24.8, 24.7, 24.7, 24.6, 20.5, 20.4.

³¹P NMR (203 MHz, CDCl₃): δ (ppm) 149.4, 149.3.

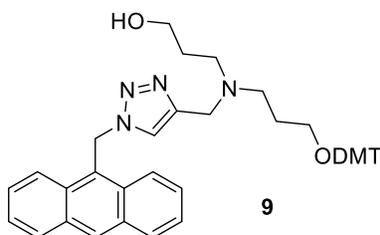
5.7.3.8 9-(Azidomethyl)anthracene (**8**)



In a dry flask covered with aluminum foil, under an argon atmosphere, to a solution of dimethylformamide (10 mL) were introduced with vigorous stirring, 9-(chloromethyl)anthracene (500 mg, 2.21 mmol, 1 equiv.) and sodium azide (430 mg, 6.62 mmol, 3 equiv.). The reaction was left under vigorous stirring overnight. Distilled water and DCM were added to the reaction medium while pH was not allowed to be under 8.0 with a diluted sodium hydroxide solution. The phases were separated and the aqueous phase was extracted twice with DCM. The combined organic phases were washed once with distilled water and once with brine. Organic phase was dried over

magnesium sulfate, filtered under vacuum and the solvent was evaporated (40 °C, 50 mbar, 1h); yielding 400 mg (1.71 mmol, 78 %) of a pure yellow solid that was stored away from light. Spectroscopic data matched those for previously reported compound in the literature.⁵⁹

5.7.3.9 3-(((1-(Anthracen-9-ylmethyl)-1H-1,2,3-triazol-4-yl)methyl)(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)amino)propan-1-ol (9)



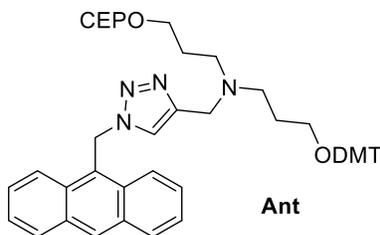
In a 10 mL round-bottom flask covered with an aluminum foil, azide **8** (257 mg, 0.80 mmol, 0.95 equiv.) and platform 1 (550 mg, 1.16 mmol, 1 equiv.) were suspended in 1.5 mL CHCl₃ and a tBuOH/water (1:1) mixture (12 mL) was added. Freshly prepared solution of sodium ascorbate (46.0 mg, 0.232 mmol, 0.20 equiv.) in 0.3 mL of water and another of copper sulfate (29.0 mg, 0.116 mmol, 0.1 equiv.) in 0.15 mL of water were sequentially added. Reaction was allowed to stir for 6h at room temperature. Product was extracted twice with DCM from 10 % Na₂CO₃ aqueous solution, washed once with a 10 % Na₂CO₃ aqueous solution, dried over MgSO₄, filtered and solvent was evaporated under reduced pressure (40 °C). The crude obtained was resuspended in toluene to evaporate remaining tBuOH. 643mg of a dark yellow/brown solid were obtained (714 mg, 1.01 mmol, 87 %). The compound was kept away from light.

HRMS (APCI-QTOF) m/z: [M + Cl]⁻ Calcd for C₄₅H₄₆N₄O₄Cl 741.32131; Found 741.32122.

¹H NMR (500 MHz, CDCl₃): δ (ppm) 8.56 (s, 1H), 8.26 (d, *J*=9Hz, 2H), 8.06 (d, *J*=8Hz, 2H), 7.54-7.49 (m, 4H), 7.38-7.37 (m, 2H), 7.27-7.24 (m, 7H), 7.20-7.17 (m, 1H), 7.06 (s, 1H), 6.80 (d, *J*=9Hz, 4H), 6.50 (s, 2H), 3.78 (s, 6H), 3.60 (s, 2H), 3.49 (t, *J*=5Hz, 2H), 2.99 (t, *J*=6Hz, 2H), 2.54 (t, *J*=6Hz, 2H), 2.48-2.45 (m, 2H), 1.72-1.66 (m, 2H), 1.55 (quint, *J*=5Hz, 2H).

¹³C NMR (126 MHz, CDCl₃): δ (ppm) 158.4, 145.3, 144.6, 136.6, 131.5, 130.9, 130.1, 129.9, 129.6, 128.2, 127.8, 127.7, 126.7, 125.5, 124.0, 123.1, 122.1, 113.1, 85.9, 63.9, 61.6, 55.3, 53.6, 51.0, 48.8, 46.6, 28.0, 27.6.

5.7.3.10 3-(((1-(Anthracen-9-ylmethyl)-1H-1,2,3-triazol-4-yl)methyl)(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)amino)propyl (2-cyanoethyl) diisopropylphosphoramidite (Ant)



Compound **9** was suspended in toluene and solvent was evaporated under reduced pressure (60 °C). The operation was repeated once and the dried compound was kept under high vacuum for at least 5 hours. In an oven-dried flask covered with aluminum foil, compound **9** (714 mg, 1.01 mmol, 1 equiv.) was then dissolved in anhydrous DCM (10 mL) and 0.88 mL of dry DIPEA (5.1 mmol, 5 equiv.) were added under stirring. CEP-Cl (0.29 mL, 1.3 mmol, 1.3 equiv.) was added slowly and the reaction was allowed to stir under inert gas at room temperature for 2 hours. Two fast extractions with DCM from a 10 % Na₂CO₃ aqueous solution were performed. Organic fractions were combined, dried over MgSO₄, filtered and the solvent was evaporated under reduced pressure (40 °C). The crude product was loaded on celite and purified using the combiFlash system with a SiO₂ “Gold” column pretreated with 1 % TEA in hexanes. Hexanes/TEA (100:1) and ethyl acetate were used in a gradient. A white solid was isolated (690 mg, 0.855 mmol, 85 %). This compound was stored away from light.

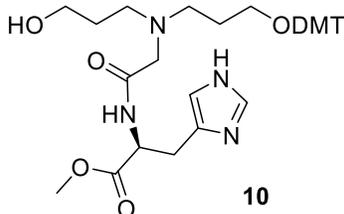
HRMS (APCI-QTOF) m/z: [M + Cl]⁻ Calcd for C₅₄H₆₃N₆O₅PCl 941.42916; Found 941.42839.

¹H NMR (500 MHz, CDCl₃): δ (ppm) 8.57 (s, 1H), 8.26-8.23 (m, 2H), 8.07-8.05 (m, 2H), 7.53-7.47 (m, 4H), 7.37-7.35 (m, 2H), 7.21-7.15 (m, 7H), 7.06 (s, 1H), 6.78 (d, *J*=9Hz, 4H), 6.49 (s, 2H), 3.77-3.66 (m, 8H), 3.57 (s, 2H), 3.56-3.44 (m, 4H), 2.94 (t, *J*=6Hz, 2H), 2.53 (t, *J*=6Hz, 2H), 2.45-2.39 (m, 4H), 1.65-1.58 (m, 4H), 1.14 (d, *J*=6Hz, 6H), 1.06 (d, *J*=7Hz, 6H).

¹³C NMR (126 MHz, CDCl₃): δ (ppm) 158.4, 145.7, 145.4, 136.7, 131.6, 130.9, 130.1, 129.9, 129.5, 128.3, 127.8, 127.7, 126.7, 125.5, 124.1, 123.2, 121.9, 117.8, 113.1, 85.8, 62.0, 61.9, 61.6, 58.5, 58.3, 55.3, 50.9, 50.4, 49.1, 46.6, 43.1, 43.0, 28.8, 28.7, 27.7, 24.8, 24.7, 24.7, 24.6, 20.5, 20.4.

^{31}P NMR (203 MHz, CDCl_3): δ (ppm) 147.2.

5.7.3.11 Methyl N-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-N-(3-hydroxypropyl)glycyl-L-histidinate (10)



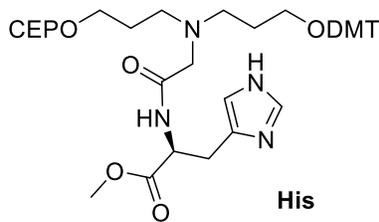
Compound **PT2b** (484 mg, 0.88 mmol, 1 equiv.) was dissolved in about 2 mL of methanol. To the mixture was added 25 mL of 0.4M NaOH in MeOH/water 4:1. The reaction mixture was left under stirring for 3h at 65 °C. Reaction was monitored by TLC. When the higher mobility spot disappeared, methanol was evaporated under reduced pressure (60 °C) until a precipitate appears but a small amount of water remains. DCM is added to the obtained solution and 2 equiv. of tetrabutylammonium chloride (489 mg, 1.76 mmol) were added. Two extractions with DCM from sat. Na_2CO_3 solution were performed followed by one washing with 10 % Na_2CO_3 solution. Organic fractions were combined, dried over MgSO_4 , filtered and the solvent was evaporated under reduced pressure (40 °C). The tetrabutylammonium carboxylate salt obtained was suspended in anhydrous DMF (5 mL). Anhydrous HOBt (154 mg, 1.14 mmol, 1.3 equiv.) and EDC-Cl (219 mg, 1.14 mmol, 1.3 equiv.) were successively added under argon. The solution was left under stirring for 5 minutes until complete dissolution of EDC-Cl. Histidine methyl ester dihydrochloride (224 mg, 0.92 mmol, 1.05 equiv.) and triethylamine (0.61 mL, 4.4 mmol, 5 equiv.) were then successively added. The reaction mixture was left under vigorous stirring overnight. Product was extracted twice with DCM from 10 % Na_2CO_3 aqueous solution, washed once with a 10 % Na_2CO_3 aqueous solution, dried over MgSO_4 , filtered and solvent was evaporated under reduced pressure (60 °C). The crude product was loaded on celite and purified using the combiFlash system with a SiO_2 “Gold” column pretreated with 0.1 % TEA in DCM. DCM/TEA (100:0.1) and DCM/Methanol/TEA (90:10:0.1) were used in a gradient. Product was extracted again to remove triethylammonium salts, twice with DCM from 10 % Na_2CO_3 aqueous solution, washed once with a 10 % Na_2CO_3 aqueous solution, dried over MgSO_4 , filtered and solvent was evaporated under reduced pressure (60 °C). A white solid was isolated (233 mg, 0.361 mmol, 41 %).

LRMS: Calc.exact mass: 644.32 g/mol. Measured (positive mode): 667.2 (M+23).

^1H NMR (500 MHz, CDCl_3): δ (ppm) 9.17 (bs, 1H), 8.16 (bs, 1H), 7.47 (d, $J=1\text{Hz}$, 1H), 7.40-7.39 (m, 2H), 7.29-7.26 (m, 6H), 7.22-7.18 (m, 1H), 6.81 (d, $J=9\text{Hz}$, 4H), 6.77 (s, 1H), 4.86-4.82 (m, 1H), 3.79 (s, 6H), 3.73-3.69 (m, 5H), 3.14-3.04 (m, 6H), 2.62-2.54 (m, 4H), 1.73-1.65 (m, 4H).

^{13}C NMR (126 MHz, DMSO): δ (ppm) 171.6, 170.7, 157.9, 145.2, 136.0, 135.0, 129.6, 127.8, 127.6, 126.6, 113.1, 85.2, 61.2, 58.9, 57.9, 55.0, 51.8, 51.6, 51.4, 45.7, 30.0, 27.2.

5.7.3.12 Methyl N-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-N-(3-(((2-cyanoethoxy)(diisopropylamino)phosphanyl)oxy)propyl)glycyl-L-histidinate (His)



Compound **10** was suspended in acetonitrile and solvent was evaporated under reduced pressure (60 °C). The operation was repeated once and the dried compound was kept under high vacuum for at least 5 hours. In an oven-dried flask, compound **10** (154 mg, 0.239 mmol, 1 equiv.) was then dissolved in anhydrous DCM (5 mL) and 0.21 mL of dry DIPEA (1.2 mmol, 5equiv.) were added under stirring. CEP-Cl (0.06 mL, 0.2 mmol, 1.1 equiv.) was added slowly and the reaction was allowed to stir under inert gas at room temperature for 2 hours. Two fast extractions with DCM from a 10 % Na_2CO_3 aqueous solution were performed. Organic fractions were combined, dried over MgSO_4 , filtered and the solvent was evaporated under reduced pressure (40 °C). The crude product was loaded on celite and purified using the combiFlash system with a SiO_2 “Gold” column pretreated with 1 % TEA in hexanes. Hexanes/TEA (100:1) and ethyl acetate were used in a gradient. A white solid was isolated (110 mg, 0.130 mmol, 54 %).

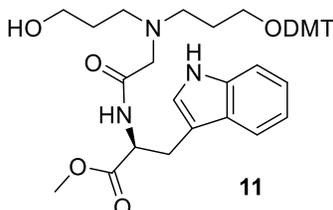
HRMS (ESI-QTOF) m/z : $[\text{M} + \text{Na}]^+$ Calcd for $\text{C}_{45}\text{H}_{61}\text{N}_6\text{O}_8\text{PNa}$ 867.4181; Found 867.4164.

^1H NMR (500 MHz, CDCl_3): δ (ppm) 9.41 (bs, 1H), 8.04 (bs, 1H), 7.47 (s, 1H), 7.40-7.39 (m, 2H), 7.29-7.26 (m, 6H), 7.21-7.18 (m, 1H), 6.81 (d, $J=9\text{Hz}$, 4H), 6.73 (s, 1H), 4.80-4.77 (m, 1H), 3.88-3.55 (m, 15H), 3.11-3.03 (m, 6H), 2.63-2.52 (m, 6H), 1.72-1.63 (m, 4H), 1.17 (t, $J=6\text{Hz}$, 12H).

^{13}C NMR (126 MHz, CDCl_3): δ (ppm) 158.5, 145.3, 136.6, 135.2, 130.1, 128.3, 127.9, 126.8, 113.2, 86.0, 62.0, 61.7, 58.7, 58.2, 58.1, 58.0, 55.3, 53.0, 52.9, 52.6, 52.4, 52.2, 51.8, 51.1, 45.6, 43.2, 43.1, 29.8, 29.4, 28.1, 24.8, 24.8, 24.8, 24.7, 20.6, 20.6.

^{31}P NMR (203 MHz, CDCl_3): δ (ppm) 147.6 147.4.

5.7.3.13 Methyl N-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-N-(3-hydroxypropyl)glycyl-L-tryptophanate (11)



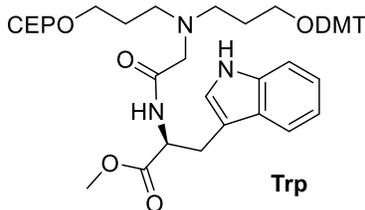
Compound **PT2b** (484 mg, 0.88 mmol, 1 equiv.) was dissolved in about 2 mL of methanol. To the mixture was added 25 mL of 0.4M NaOH in MeOH/water 4:1. The reaction mixture was left under stirring for 3h at 65 °C. Reaction was monitored by TLC. When higher mobility spot disappeared, methanol was evaporated under reduced pressure (60 °C) until a precipitate appears but a small amount of water remains. DCM is added to the obtained solution and 2 equiv. of tetrabutylammonium chloride (489 mg, 1.76 mmol) were added. Two extractions with DCM from sat. Na_2CO_3 solution were performed followed by one washing with 10 % Na_2CO_3 solution. Organic fractions were combined, dried over MgSO_4 , filtered and the solvent was evaporated under reduced pressure (40 °C). The tetrabutylammonium carboxylate salt obtained was suspended in DMF (5 mL). Anhydrous HOBt (154 mg, 1.14 mmol, 1.3 equiv.) and EDC-Cl (219 mg, 1.14 mmol, 1.3 equiv.) were successively added. The solution was left under stirring for 5 minutes until complete dissolution of EDC-Cl. Tryptophan methyl ester hydrochloride (235 mg, 0.92 mmol, 1.05 equiv.) and triethylamine (0.61 mL, 4.4 mmol, 5 equiv.) were then successively added. The reaction mixture was left under vigorous stirring overnight. Solvent was evaporated under reduced pressure (60 °C). The crude product was loaded on celite and purified using the combiFlash system with a SiO_2 “Gold” column pretreated with 0.1 % TEA in DCM. DCM/TEA (100:0.1) and DCM/Methanol/TEA (90:10:0.1) were used in a gradient. A pale yellow solid was isolated (260 mg, 0.374 mmol, 43 %).

HRMS (ESI-QTOF) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{41}\text{H}_{48}\text{N}_3\text{O}_7$ 694.34868; Found 694.34783.

^1H NMR (500 MHz, CDCl_3): δ (ppm) 8.36 (s, 1H), 7.66 (d, $J=8\text{Hz}$, 1H), 7.52 (d, $J=8\text{Hz}$, 1H), 7.43-7.41 (m, 2H), 7.31-7.21 (m, 8H), 7.13 (t, $J=7\text{Hz}$, 1H), 7.08 (t, $J=7\text{Hz}$, 1H), 6.92 (d, $J=2\text{Hz}$, 1H), 6.83 (d, $J=9\text{Hz}$, 4H), 4.92-4.88 (m, 1H), 3.78 (s, 6H), 3.68 (s, 3H), 3.38-3.26 (m, 4H), 3.09-2.96 (m, 4H), 2.34-2.52 (m, 4H), 1.63-1.52 (m, 2H), 1.47-1.39 (m, 2H).

^{13}C NMR (126 MHz, CDCl_3): δ (ppm) 172.7, 171.7, 158.4, 145.3, 136.5, 136.2, 130.1, 128.2, 127.9, 127.7, 126.8, 122.9, 122.2, 119.6, 118.6, 113.1, 111.3, 109.9, 86.0, 61.5, 60.5, 58.7, 55.3, 52.5, 52.5, 52.1, 52.0, 30.1, 27.8, 27.4.

5.7.3.14 Methyl N-(3-(bis(4-methoxyphenyl)(phenyl)methoxy)propyl)-N-(3-(((2-cyanoethoxy)(diisopropylamino)phosphanyl)oxy)propyl)glycyl-L-tryptophanate (Trp)



Compound **11** was suspended in acetonitrile and solvent was evaporated under reduced pressure (60 °C). The operation was repeated once and the dried compound was kept under high vacuum for at least 5 hours. In an oven-dried flask, compound **11** (254 mg, 0.353 mmol, 1 equiv.) was then dissolved in anhydrous DCM (6 mL) and 0.31 mL of dry DIPEA (1.8 mmol, 5 equiv.) were added under stirring. CEP-Cl (0.10 mL, 0.46 mmol, 1.3 equiv.) was added slowly and the reaction was allowed to stir under inert gas at room temperature for 2 hours. Two fast extractions with DCM from a 10 % Na_2CO_3 aqueous solution were performed. Organic fractions were combined, dried over MgSO_4 , filtered and the solvent was evaporated under reduced pressure (40 °C). The crude product was loaded on celite and purified using the combiFlash system with a SiO_2 “Gold” column pretreated with 1 % TEA in hexanes. Hexanes/TEA (100:1) and ethyl acetate were used in a gradient. A white solid was isolated (286 mg, 0.320 mmol, 91 %).

HRMS (ESI-QTOF) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{50}\text{H}_{65}\text{N}_5\text{O}_8\text{P}$ 894.45653; Found 894.45527.

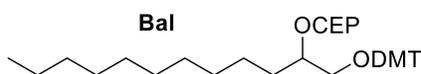
^1H NMR (500 MHz, CDCl_3): δ (ppm) 8.54 (s, 1H), 7.72 (d, $J=8\text{Hz}$, 1H), 7.50 (d, $J=8\text{Hz}$, 1H), 7.41-7.40 (m, 2H), 7.29-7.19 (m, 8H), 7.11 (t, $J=7\text{Hz}$, 1H), 7.05 (t, $J=7\text{Hz}$, 1H), 6.92 (m, 1H), 6.81 (d, $J=9\text{Hz}$, 4H), 4.88-4.85 (m, 1H), 3.83-3.69 (m, 8H), 3.64 (s, 3H), 3.62-3.42 (m, 4H), 3.35-3.24

(m, 2H), 3.06-2.95 (m, 4H), 2.55 (t, $J=5\text{Hz}$, 2H), 2.49-2.44 (m, 4H), 1.55-1.47 (m, 4H), 1.20-1.15 (m, 12H).

^{13}C NMR (126 MHz, CDCl_3): δ (ppm) 172.1, 171.5, 158.3, 145.2, 136.4, 136.4, 136.1, 129.9, 128.1, 127.7, 127.5, 126.6, 122.8, 121.9, 119.3, 118.4, 117.8, 117.8, 113.0, 111.3, 109.7, 109.7, 85.8, 61.8, 61.7, 61.6, 61.6, 61.3, 58.5, 58.5, 58.2, 58.1, 58.0, 58.0, 55.1, 52.6, 52.4, 52.2, 52.0, 43.0, 42.9, 29.0, 28.9, 27.9, 27.8, 27.4, 24.6, 24.6, 24.6, 24.6, 20.3, 20.3.

^{31}P NMR (203 MHz, CDCl_3): δ (ppm) 147.5 147.4.

5.7.3.15 1-(Bis(4-methoxyphenyl)(phenyl)methoxy)dodecan-2-yl (2-cyanoethyl) diisopropylphosphoramidite (**Bal**)



2 steps protocol to obtain **Bal** from 1,2-dodecanediol and characterization data are detailed elsewhere.⁵⁴

5.7.4 Solid-Phase Synthesis

For unbranched oligomers, synthesis was performed on a 1 μmol scale, starting from a universal 1000 Å LCAA-CPG solid support. 5'-DMT and 5'-Lev nucleoside phosphoramidites (benzoyl protected adenosine, benzoyl protected cytidine, isobutyryl protected guanosine and thymidine) were dissolved in dry acetonitrile and coupling times of 3 minutes were used. Molecular trap packs were used to maintain the acetonitrile, the activator and the phosphoramidite solutions dry. Removal of the DMT protecting group was carried out using 3 % dichloroacetic acid in dichloromethane on the DNA synthesizer. Removal of the Levulinyl protecting group was carried out using hydrazine hydrate (50-60 %) diluted to 0.5M in a 3:2 (v:v) pyridine/acetic acid solution. Hydrazine treatment lasts 6 minutes with three injections of 2 minutes each unless otherwise noted.

For branched oligomers, synthesis was performed similarly starting from a universal 2000 Å LCAA-CPG solid support and using coupling times of 10 minutes. The codons were synthesized after the unnatural monomer couplings in the case of the library but before otherwise.

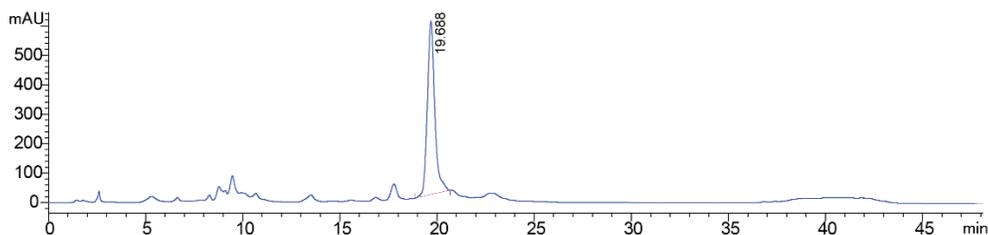
LEV2	<i>TTTTTTTGTCCG</i>	Hydrazine 32min
T10	<i>GGTCCTAATGCCGATCGA</i>	Hydrazine 10min
T30	<i>GGTCCTAATGCCGATCGA</i>	Hydrazine 30min
DNA 21mer	<i>TA-X-TTTTTTCAGTTGACCATATA</i>	From X, 20 hydrazine treatments of 10 min
BR1	<i>CGTCGAGGCCCGGGAAGAAAGGA-BU- ACACGTCACGCCTTTTTTTTTT</i> Branch: TCGA	Hydrazine 20 min
BR2	<i>CGTCGAGGCCCGGGAAGAAA-BU- GCATAGGATACACGTCACGCC</i> Branch: TCGA	Hydrazine 20 min
BR3	DMT- <i>CGTCGAGGCCCTTAATTGGTTCCTTAACCTTAGGA- BU-ACACGTCACGCCTTTTTTTTTT</i> Branch: TATGTCTACT	Hydrazine 15 min
TBA1	<i>CGTCGAGGCCCTTCTTCTTCGTCTTCTTCTTAGGA- BU-ACACGTCACGCCTTTTTTTTTT</i> Branch: GGTTGGTGTGGTTGGTTTTT	Hydrazine 10 min
TBA2	<i>CGTCGAGGCCCGTGCTTACCCGTGACCTTCTTAGGA- BU-ACACGTCACGCCTTTTTTTTTT</i> Branch: GGHisTGGC12GTGGAlkTGGTTTTT	Hydrazine 10 min
TBA3	<i>CGTCGAGGCCCATACAGGTATAATGCCATTGAAGGA- BU-ACACGTCACGCCTTTTTTTTTT</i> Branch: GGTrpSugGGPheAntTGGBalAGGNapTTTT	Hydrazine 10 min
Forward Primer	<i>ACACTGACGACATGGTTCTACACGTCGAGGCC</i>	
Reverse Primer	<i>TACGGTAGCAGAGACTTGGTCTGGCGTGACGTGTAT CCT</i>	
LIB	<i>CGTCGAGGCC NNN NNN NNN NNN NNN NNN NNN AGGA-BU- ACACGTCACGCCTTTTTTTTTT</i> Branch:GG 7 6 GG 5 4 TGG 3 2 GG 1 TTTT	

Table 5.3. Monomers used during library synthesis.

Position modified	Aliquot A	Aliquot B	Aliquot C	Aliquot D	Aliquot E	Aliquot F
1	G	T	Alk	C12	Nap	Phe
2	A	T	Alk	C12	Sug	Trp
3	C	T	Ant	Nap	Phe	Trp
4	A	G	Alk	C12	Nap	Sug
5	G	T	C12	Phe	Sug	Trp
6	A	T	Alk	His	Nap	Phe
7	G	T	Ant	C12	Sug	Trp

5.7.5 Reverse-phase HPLC

Analysis and purification with the HPLC were performed as follows. Solvents (0.22 μm filtered): 50 mM triethylammonium acetate (TEAA) buffer (pH 8.0) and HPLC grade acetonitrile. Elution gradient is described on **Figure 5.9**, **Figure 5.10** and **Figure 5.11**. All gradients were followed by a short column wash in 95 % acetonitrile. Column: Hamilton PRP-C18 5 μm 100 \AA 2.1 x 150 mm. For each analytical separation approximately 0.5 OD₂₆₀ of crude DNA was injected as a 20-50 μL solution in Millipore water. Detection was carried out using a diode-array detector, monitoring absorbance at 260 nm.

**Figure 5.8.** RP-HPLC chromatogram of **AT-2-Nap**. Gradient : 3 to 50 % in 30 minutes.

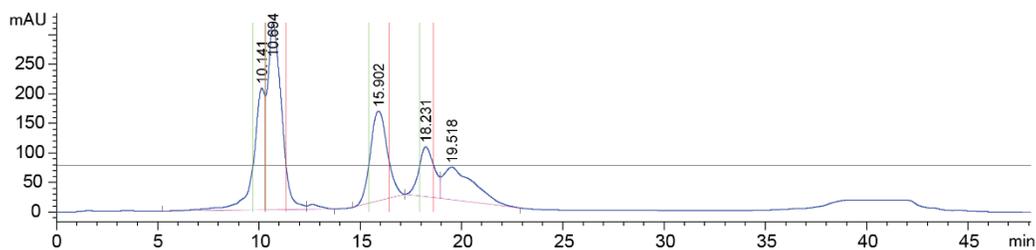


Figure 5.9. RP-HPLC chromatogram of **BR3**. Gradient : 3 to 50 % in 30 minutes.

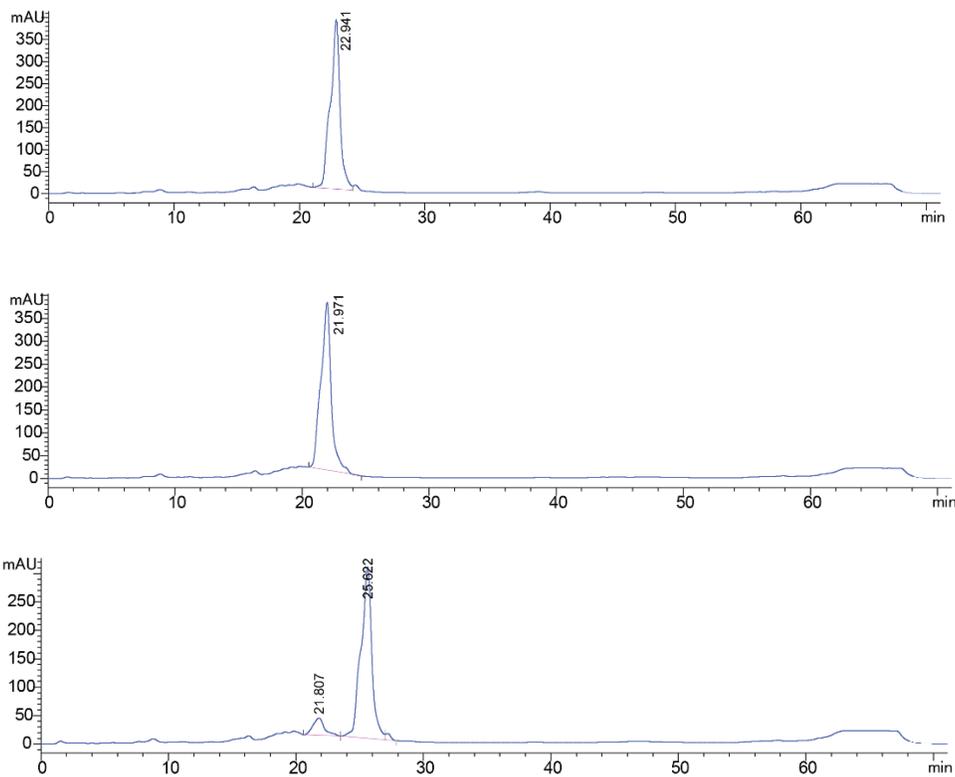


Figure 5.10. RP-HPLC chromatogram of **AT**, **AT-His**, **AT-Trp**. Respectively top, middle and bottom traces. Gradient : 3 to 30 % in 50 minutes. These strands were made and HPLC-purified in order to check for their purity by LC-MS.

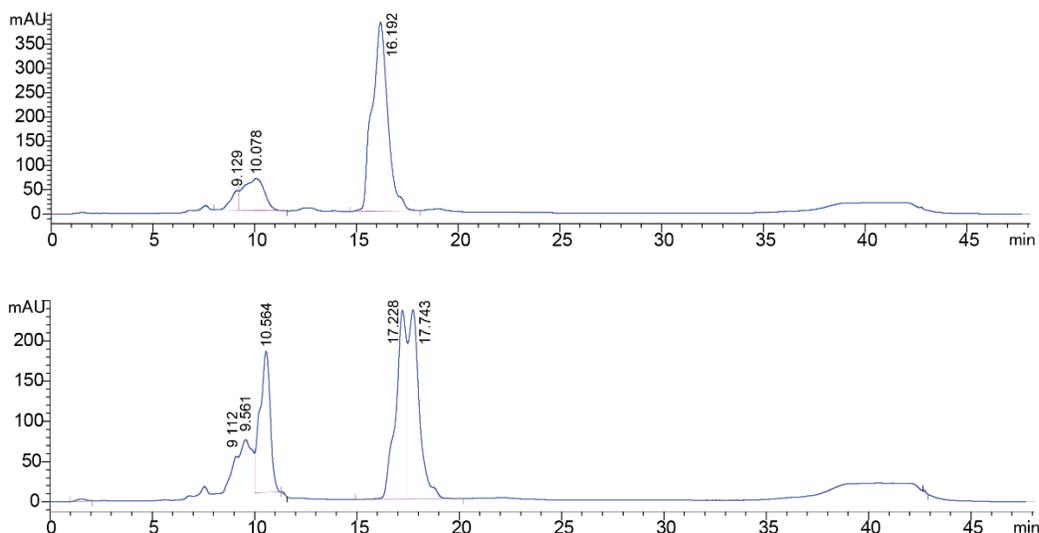


Figure 5.11. RP-HPLC chromatogram of **AT-Ant** (top), **AT-Bal** (bottom). Gradient : 3 to 50 % in 30 minutes. Peak at 10 minutes correspond to unmodified DNA strands while other peak is the expected modified strand. Yields of **Bal** are lower than for the other modifications since this monomer is not suitable for 5' end modification. At internal positions, yields are similar to the other monomers as shown on Figure 5.5. We hypothesize double peak is due to the presence of two diastereomers. These strands were made and HPLC-purified in order to check for their purity by LC-MS.

5.7.6 Gel electrophoresis

Denaturing Polyacrylamide Gel Electrophoresis (PAGE) was carried out at room temperature for 30 minutes at 250V followed by 1 hour at 500V with big plates and only for 1h at 100V with small plates. TBE buffer (1X) was used and the concentration of urea in the gel was 7M. For each lane 5 μ L of DNA (0.1 to 2 μ M) in water was added to 5 μ L of 8 M urea. The DNA bands for all gels were visualized by incubation with GelRed™. In all gels, the ladder used is the Invitrogen Ultra Low Range ladder.

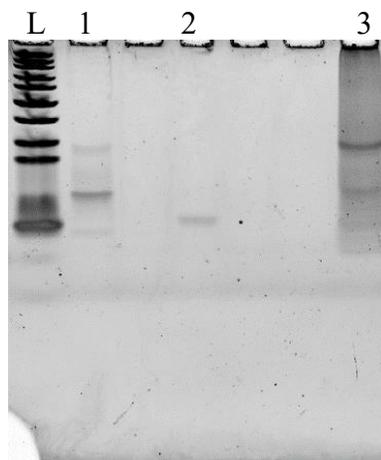


Figure 5.12. 15 % PAGE, denaturing conditions, **HYD1, LEV1, LEV2**. Crude mixture was loaded. Lanes 1, 2,3: **LEV1, LEV2, HYD1**. Strands are discussed in Section 5.5.1. The highest mobility band is visible on the whole gel width.

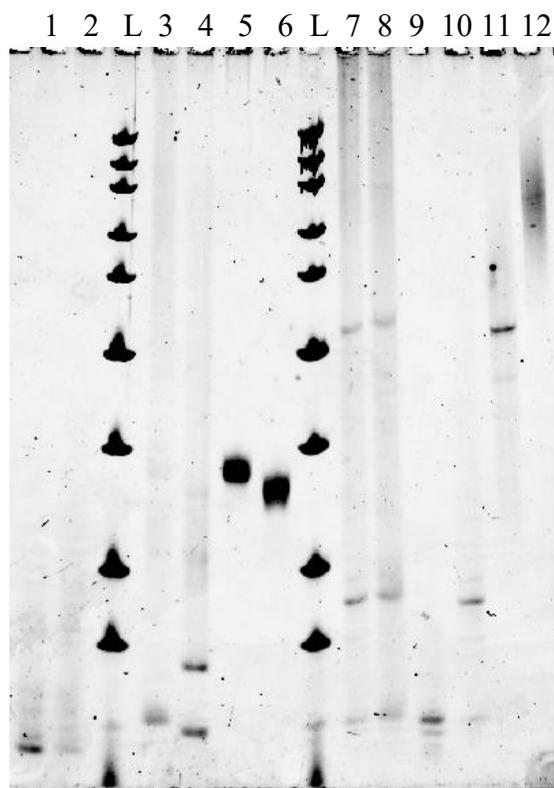


Figure 5.13. 12 % PAGE, denaturing conditions, **T10, T30, BR1, BR2, BR3**. Lanes : 1 : **T10** (crude), 2 : **T30** (crude), 3 : **BR1** (crude), 4 : **BR2** (crude), 5 : **BR1** (gel purified), 6 : **BR2** (gel purified), 7 : **BR3-DMT off** (crude), 8 : **BR3** (crude), 9 : **BR3-f1** (fraction 1 from HPLC purification), 10: **BR3-f2** (fraction 2 from HPLC purification), 11: **BR3-f3** (fraction 3 from HPLC purification), 12: **BR3-f4** (fraction 4 from HPLC purification). **T10** and **T30** were synthesized to compare 10 min vs 30 min hydrazine deprotection times (see **Table 5.2**). Other strands are discussed in Section 5.5.2.



Figure 5.14. 15 % PAGE denaturing conditions, **AT-Ant**, **AT**, **AT-Bal**, **AT-His**, **AT-Trp**. Lanes 1 to 5: Crude mixtures from synthesis of **AT-Ant**, **AT**, **AT-Bal**, **AT-His**, **AT-Trp**. **Bal** is found to have lower attachment yields in this case because it is located at the 5' end. These strands were made and HPLC-purified in order to characterize them by LC-MS.

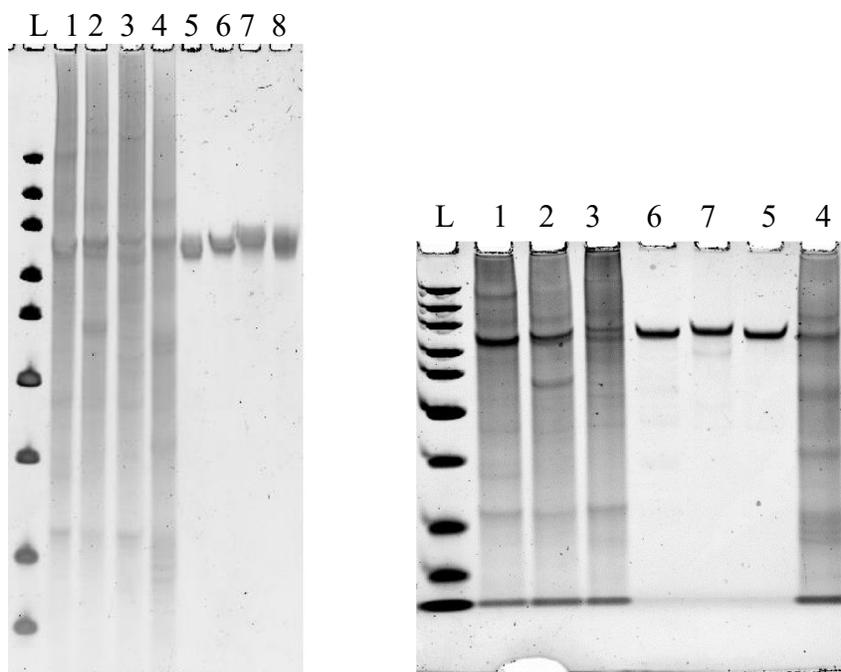


Figure 5.15. 12 % PAGE, denaturing conditions, **TBA1**, **TBA2**, **TBA3** and **LIB**. Lanes 1 to 4: crude mixtures from synthesis of **TBA1**, **TBA2**, **TBA3** and **LIB**, Lanes 5 to 8: gel purified products: **TBA1**, **TBA2**, **TBA3** and **LIB**. These strands are discussed in Section 5.5.4.

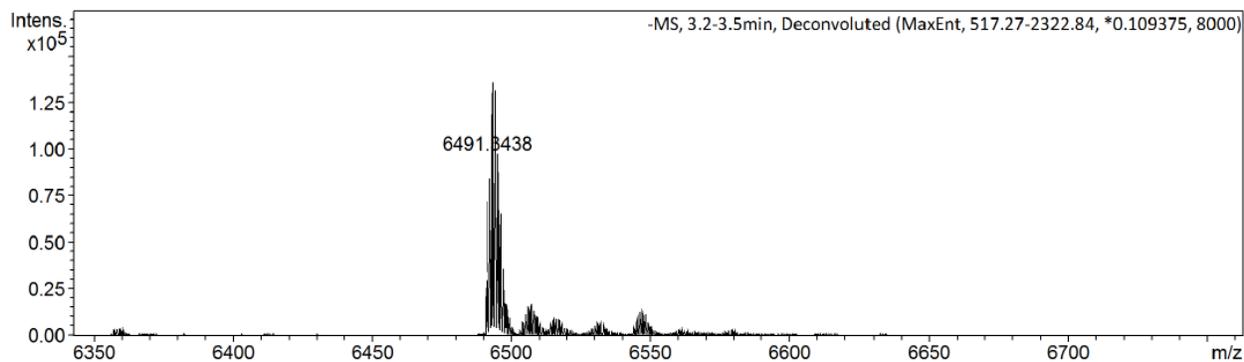
5.7.7 LC-MS

The oligomers were analyzed by LC-ESI-MS in negative ESI mode. Samples (~60 pmols in H₂O) were run through an Acclaim RSLC 120 C18 column (2.2 μ m, 120 Å 2.1 x 50 mm) using a gradient of mobile phase A (100 mM 1,1,1,3,3,3-hexafluoro-2-propanol and 5 mM triethylamine in water) and mobile phase B (Methanol) in 8 minutes (2 % to 100 % B). Liquid chromatography was performed as a control for strand purity which was found to be superior to 90 % in all cases.

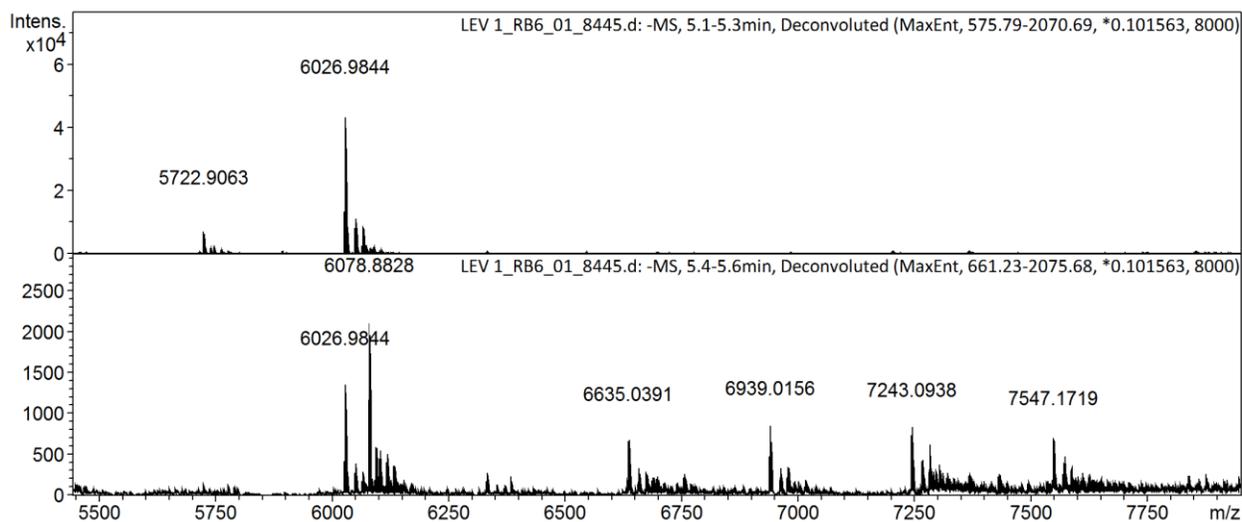
Table 5.4. ESI-MS characterization of the strands synthesized.

Strand	Expected exact mass (g/mol)	Found exact mass (g/mol)
AT-2-Nap	6491.23	6491.34
LEV1	6027.00	6026.98
LEV2	4214.70	4215.76
BR1	15230.90 (MW)	15230.20 (MW)
BR2	14346.39 (MW)	14345.68 (MW)
BR3-DMT_{off}	20590.40 (MW)	20589.71 (MW)
BR3 , bottom band ^a	6310.06	6310.00
BR3 , middle band ^a	9810.66	9810.47
BR3 , HPLC fraction 4 ^a	>20590	Mostly >27000
AT-Ant	6231.17	6231.25
AT-Bal	6029.14	6029.22
AT-His	6155.12	6155.22
AT-Trp	6204.14	6204.25
TBA1	24042.59 (MW)	24041.20 (MW)
TBA2	24110.81 (MW)	24109.73 (MW)
TBA3	24839.97 (MW)	24839.96 (MW)

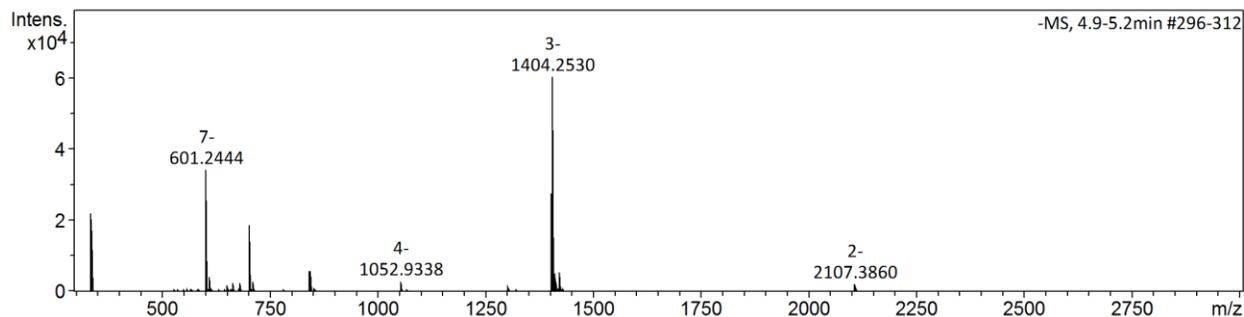
MW stands for molecular weight. a. These strands are byproducts from **BR3** synthesis. They are discussed in Figure 5.4.



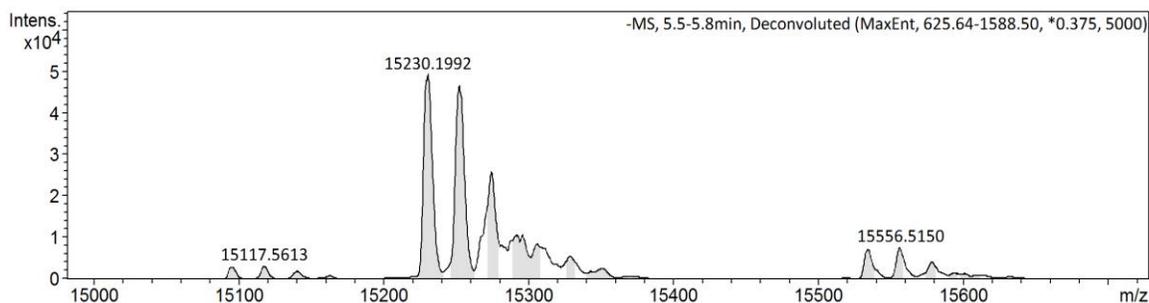
a) **AT-2-Nap** after HPLC purification.



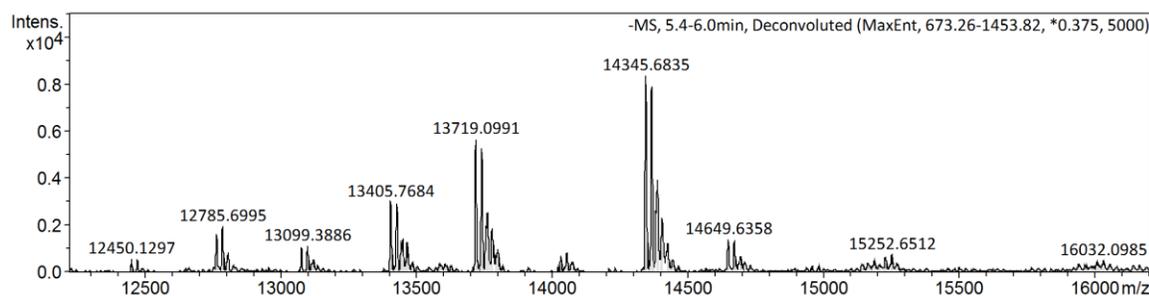
b) **LEV1**, analyzed before purification. Difference between peaks of the bottom spectrum is about 304 g/mol, which is the mass of a thymidine nucleotide. It shows that the byproducts shown on this spectrum may be branched oligomers.



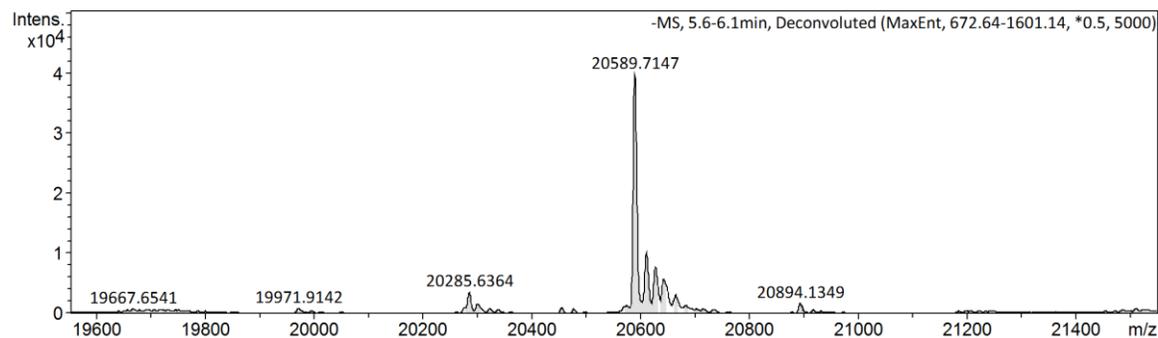
c) **LEV2**, analyzed before purification.



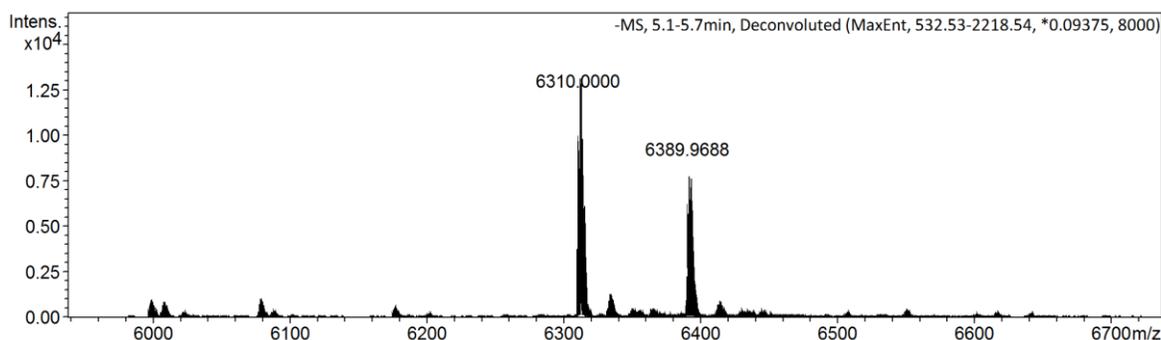
d) **BR1**, analyzed after gel purification



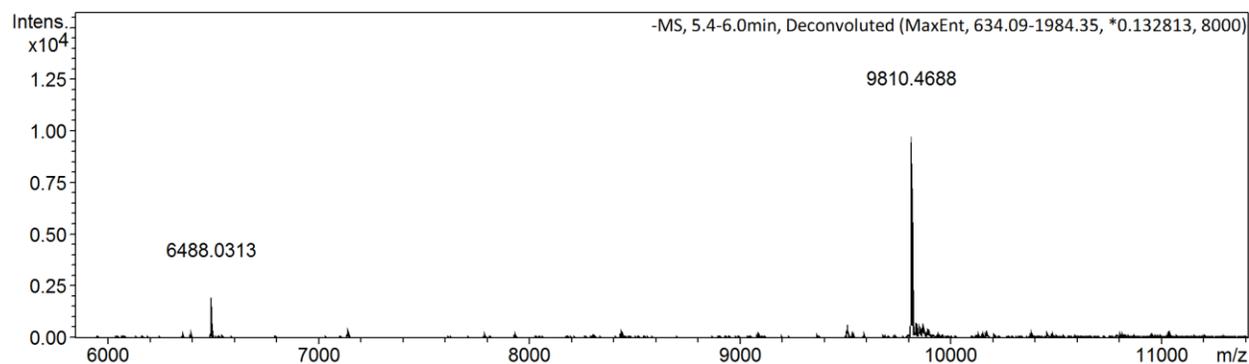
e) **BR2**, analyzed after gel purification. Peaks of lower masses probably come from truncated products due to mispurification or strand degradation after purification.



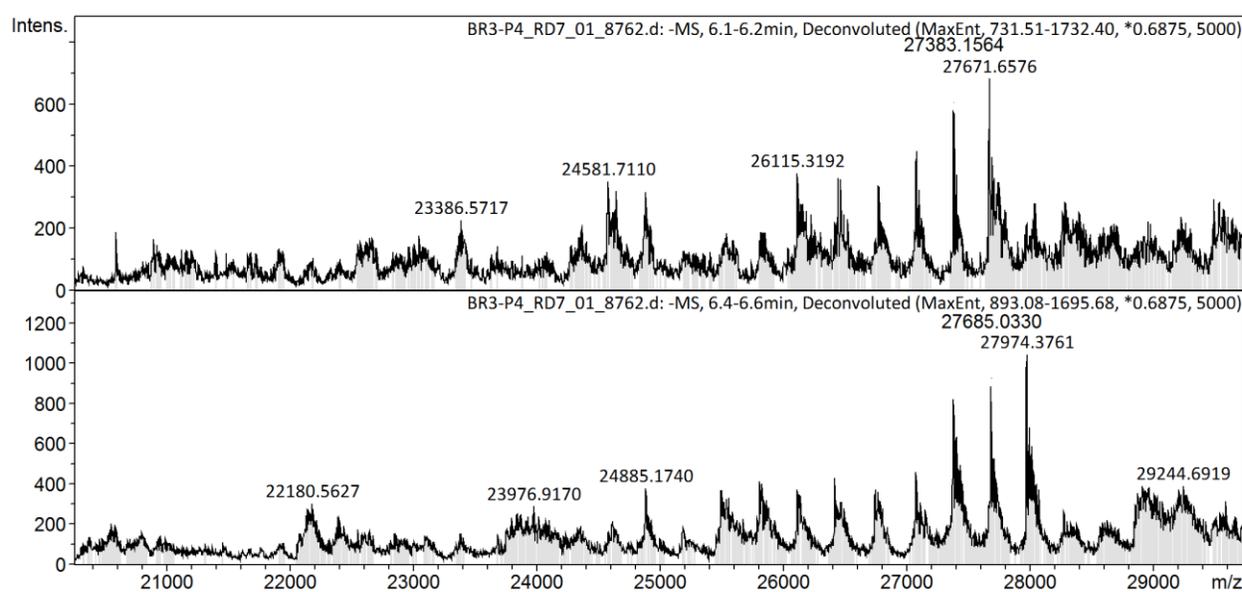
f) **BR3-DMT off**, after gel purification.



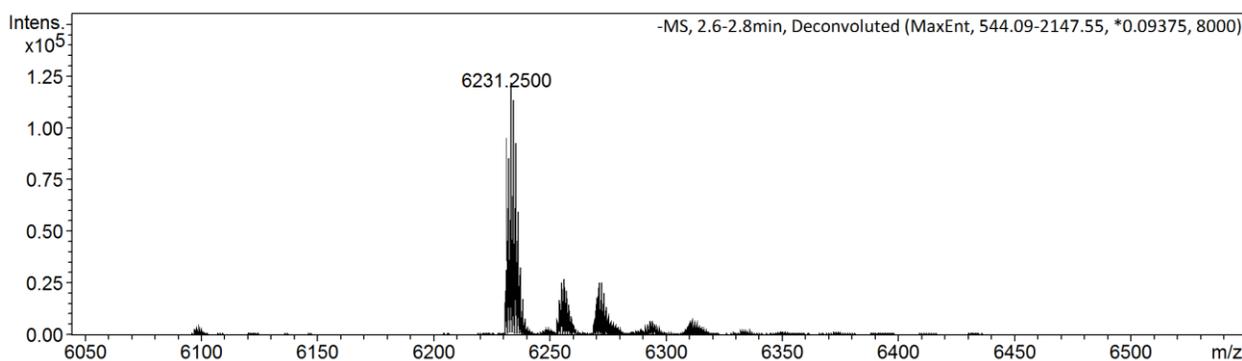
g) **BR3**, bottom band, after gel purification. Truncated monomer before the insertion of **BU** as explained on Figure 5.4. Second peak has a +80 g/mol difference corresponding to an extra phosphate group present at the 5' end.



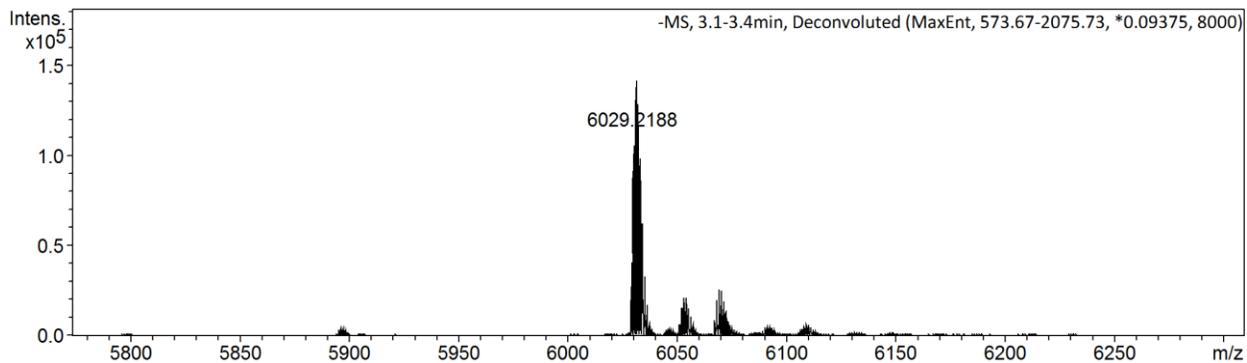
h) **BR3**, middle band, after gel purification. Truncated monomer for which the code did not grow; as explained on Figure 5.4.



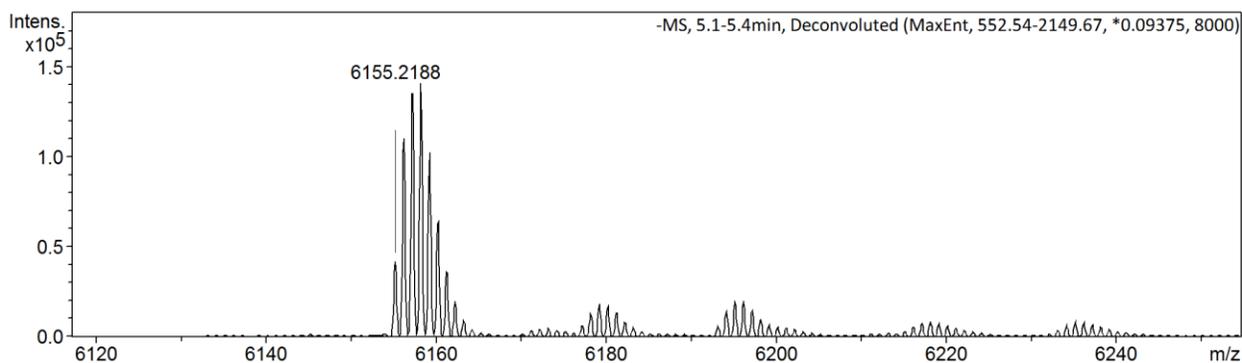
i) **BR3**, fraction 4 after HPLC purification. High masses may be due to the presence of multi-branched oligomers as explained on Figure 5.4.



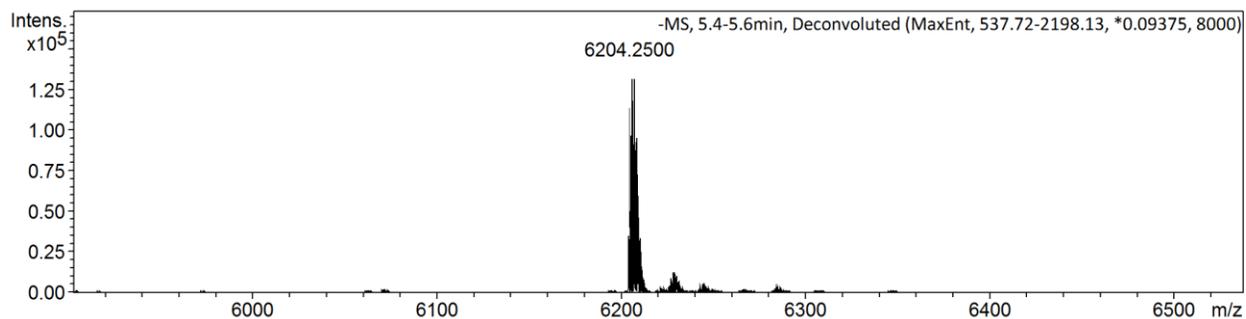
j) **AT-Ant**, after HPLC purification.



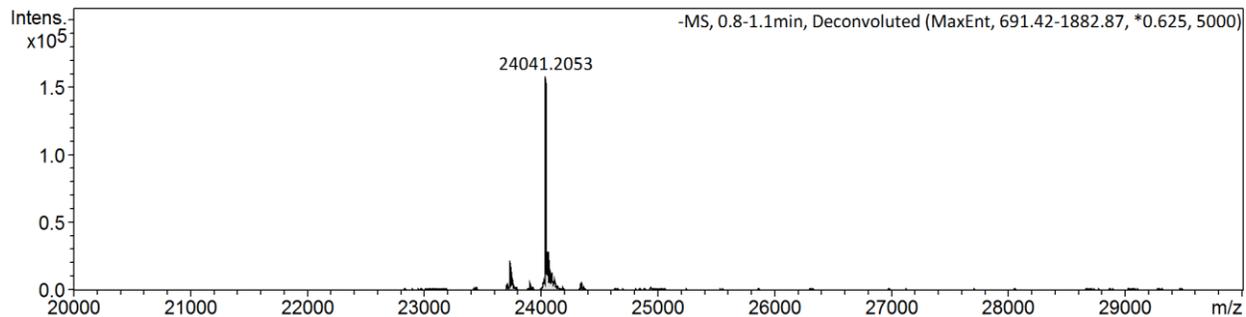
k) **AT-Bal**, after HPLC purification.



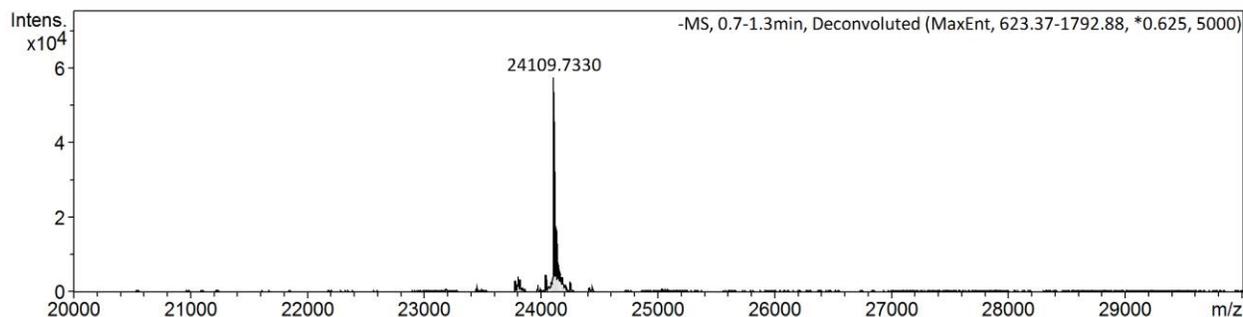
l) **AT-His**, after HPLC purification.



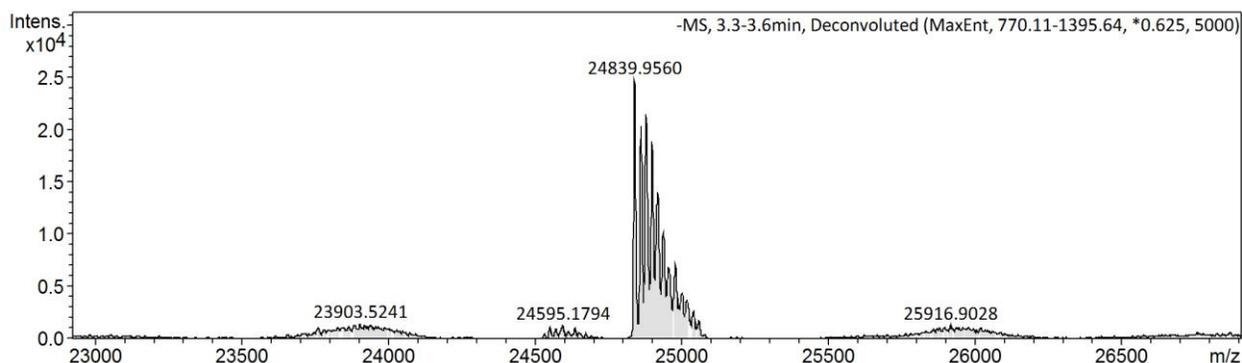
m) **AT-Trp**, after HPLC purification.



n) **TBA1**, after gel purification.



o) **TBA2**, after gel purification.



p) **TBA3**, after gel purification.

Figure 5.16. MS data for modified DNA strands. The data was processed and deconvoluted using the Bruker Data Analysis software version 4.1. Masses reported are exact masses except for **BR1**, **BR2**, **BR3-DMToff**, **TBA1**, **TBA2**, **TBA3** for which molecular weight is reported.

5.7.8 PCR

Polymerase chain reaction (PCR) was carried out using the MyTaq™ HS Red Mix PCR kit. The reaction was performed in a batch of 60 μL , using $0.1 \text{ ng} \cdot \mu\text{L}^{-1}$ of template (except for $1 \times 10^{-5} \text{ ng} \cdot \mu\text{L}^{-1}$ of **TBA3**), $0.625 \mu\text{M}$ of each of the forward and reverse primers, and a final concentration of 1x MyTaq™ HS Red Mix. The mixture was then heated at $95 \text{ }^\circ\text{C}$ for 1 minute and was followed by 30 cycles of: 1) $95 \text{ }^\circ\text{C}$ for 15 seconds, 2) $60\text{-}67 \text{ }^\circ\text{C}$ (temperature was optimized depending on the sample) for 15 seconds, and 3) $72 \text{ }^\circ\text{C}$ for 15 seconds. After PCR, the samples were purified using a QIAquick PCR Purification Kit (manufacturer protocol was followed). Electrophoresis gel experiments were performed in native and denaturing conditions.

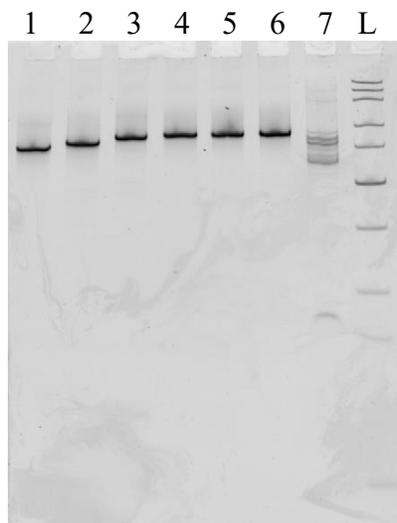


Figure 5.17. 12 % PAGE, denaturing conditions of amplicons after PCR with **BR1, BR2, BR3, TBA1, TBA2, TBA3**. Strands are respectively in lanes 1, 2, 3, 4, 5, 6. Lane 7 contains a control with no template added. This gel is discussed on section 5.5.5. L contains the Invitrogen Ultra Low Range ladder.

Native Polyacrylamide Gel Electrophoresis (PAGE) was carried out at 4 °C for 1 hour at 100V using small gel plates. TAMg buffer (1X) was used. For each lane ~10 µL of DNA (0.1 to 1 µM) from PCR crude mixture was added to 2 µL of a glycerol solution. The DNA bands for all gels were visualized by incubation with GelRed™.

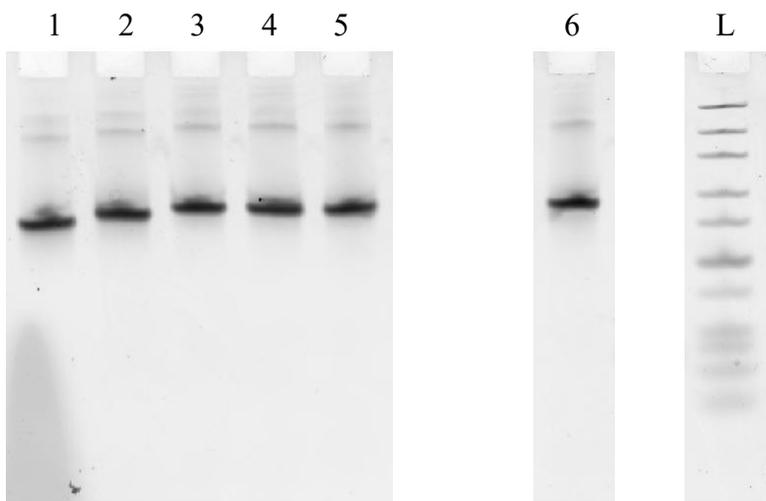


Figure 5.18. 8 % PAGE, native conditions of amplicons after PCR with **BR1, BR2, BR3, TBA1, TBA2, TBA3**. Strands are respectively in lanes 1, 2, 3, 4, 5, 6. L contains the Invitrogen Ultra Low Range ladder.

5.7.9 Sequencing

Sequencing was performed by the McGill University and Génome Québec Innovation Center. Methodology for MiSeq Illumina next generation sequencing sample preparation is indicated below.

Barcoding step: The barcoding step adds an index (or barcode) to each sample and the sequence of Illumina adapters required for DNA binds to flow cell (i5 and i7).

Table 5.5. Master Mix components for barcoding step.

Master Mix Components	1X	8	Final Concentration
Roche PCR 10X Buffer without MgCl ₂	2.000	16.0	1 X
Roche MgCl ₂ 25 mM	1.438	11.5	1.8 mM
Roche DMSO	1.000	8.0	5 %
dNTP mix 10 mM_FroggaBio	0.400	3.2	0.2 mM
TAQ 5U-ul Roche FastStart High Fi	0.100	0.8	0.025 U/ul
H ₂ O	12.063	96.5	

Polymerase chain reaction (PCR) was carried out using 17 μL of the Master Mix for a 20 μL final volume. The reaction was performed using $\sim 0.1 \text{ ng} \cdot \mu\text{L}^{-1}$ of template (except for $1 \times 10^{-5} \text{ ng} \cdot \mu\text{L}^{-1}$ of **TBA3**), 0.2 μM of each of the forward and reverse primers with the barcode corresponding to each sample. The mixture was then heated at 95 °C for 10 minutes and was followed by 15 cycles of: 1) 95 °C for 15 seconds, 2) 60 °C for 15 seconds, and 3) 72 °C for 15 seconds. The cycles were followed by 3 minutes at 72 °C.

Verification of barcode incorporation for each sample on 2 % agarose gel: All amplicons had a similar profile on agarose gel so no quantification was necessary to generate the pool (library). The library was then generated by pooling 5ul of each sample except for sample **TBA3** (7ul) in order to get more reads for this sample.

Final steps:

- Cleaning-up of the pool (or library) with a ratio of 1.5 of sparQ PureMag Beads (from Quantabio).

- QC of the library as follows: library was quantified by fluorescence using Qubit dsDNA high-sensitivity (HS) kit (ThermoFisher). Average fragment size was estimated on 2 % agarose gel with 100bp DNA ladder.
- The library was added to the “principal” library at a ratio of 1 % of the MiSeq lane. 10 % of Phix control library was spiked into the final amplicon pools (loaded at a final concentration of 10pM) to improve the unbalanced base composition.
- Sequencing with the MiSeq Reagent Kit v2 500 cycles from Illumina. Sequencing was done with LNATM modified custom primers (Exiqon):

Primer read1: LNA-CS1: ACACTGACGACATGGTTCTACA

Primer read2: LNA-CS2: TACGGTAGCAGAGACTTGGTCT

Primer index read: LNA-CS2rc: AGACCAAGTCTCTGCTACCGTA

Results were obtained as 2 FASTQ files per sample containing the forward strand read and the reverse strand read. The forward strand reads were used for **BR1** and **BR2** to count the number of times code sequences were found as well as the number of times the forward primer region sequence (5'-CGTCGAGGCC-3') was found.

For **BR3**, **TBA1**, **TBA2**, **TBA3**, we privileged the reverse strand read in order to count the number of times the reverse primer region was found (5'-GGCGTGACGTGTATCCT-3'). It is a 17mer which is closer in length and therefore more comparable to the code sequences' length (20mer for **BR3** and 21mer for **TBA1**, **TBA2** and **TBA3**). We also counted the number of times the complementary code sequence was found. Ratio of reads found/total number of reads are reported in Table 4.3.

Using the reads of the complementary strand led to similar results.

5.8 References

- (1) ten Brummelhuis, N.; Wilke, P.; Börner, H. G. Identification of Functional Peptide Sequences to Lead the Design of Precision Polymers. *Macromol. Rapid Commun.* **2017**, *38* (24), 1700632.

- (2) Zhou, J.; Rossi, J. Aptamers as Targeted Therapeutics: Current Potential and Challenges. *Nat. Rev. Drug Discov.* **2017**, *16* (3), 181–202.
- (3) Lau, J. L.; Dunn, M. K. Therapeutic Peptides: Historical Perspectives, Current Development Trends, and Future Directions. *Bioorg. Med. Chem.* **2018**, *26* (10), 2700–2707.
- (4) Dunn, M. R.; Jimenez, R. M.; Chaput, J. C. Analysis of Aptamer Discovery and Technology. *Nat. Rev. Chem.* **2017**, *1* (10), 0076.
- (5) Parekh, P.; Martin, J.; Chen, Y.; Colon, D.; Wang, H.; Tan, W. Using Aptamers to Study Protein–Protein Interactions. In *Advances in biochemical engineering/biotechnology*; 2008; Vol. 110, pp 177–194.
- (6) Ellington, A. D.; Szostak, J. W. In Vitro Selection of RNA Molecules That Bind Specific Ligands. *Nature* **1990**, *346* (6287), 818–822.
- (7) Tuerk, C.; Gold, L. Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase. *Science* **1990**, *249* (4968), 505–510.
- (8) Rohloff, J. C.; Gelinas, A. D.; Jarvis, T. C.; Ochsner, U. A.; Schneider, D. J.; Gold, L.; Janjic, N. Nucleic Acid Ligands With Protein-like Side Chains: Modified Aptamers and Their Use as Diagnostic and Therapeutic Agents. *Mol. Ther. - Nucleic Acids* **2014**, *3*, e201.
- (9) Gold, L.; Ayers, D.; Bertino, J.; Bock, C.; Bock, A.; Brody, E. N.; Carter, J.; Dalby, A. B.; Eaton, B. E.; Fitzwater, T.; et al. Aptamer-Based Multiplexed Proteomic Technology for Biomarker Discovery. *PLoS One* **2010**, *5* (12), e15004.
- (10) Pfeiffer, F.; Tolle, F.; Rosenthal, M.; Brändle, G. M.; Ewers, J.; Mayer, G. Identification and Characterization of Nucleobase-Modified Aptamers by Click-SELEX. *Nat. Protoc.* **2018**, *13* (5), 1153–1180.
- (11) Hili, R.; Niu, J.; Liu, D. R. DNA Ligase-Mediated Translation of DNA Into Densely Functionalized Nucleic Acid Polymers. *J. Am. Chem. Soc.* **2013**, *135* (1), 98–101.
- (12) Kong, D.; Yeung, W.; Hili, R. In Vitro Selection of Diversely Functionalized Aptamers. *J. Am. Chem. Soc.* **2017**, *139* (40), 13977–13980.
- (13) Chen, Z.; Lichtor, P. A.; Berliner, A. P.; Chen, J. C.; Liu, D. R. Evolution of Sequence-Defined Highly Functionalized Nucleic Acid Polymers. *Nat. Chem.* **2018**, *10* (4), 420–427.
- (14) Loakes, D.; Holliger, P. Polymerase Engineering: Towards the Encoded Synthesis of Unnatural Biopolymers. *Chem. Commun.* **2009**, *0* (31), 4619.
- (15) Pinheiro, V. B.; Taylor, A. I.; Cozens, C.; Abramov, M.; Renders, M.; Zhang, S.; Chaput, J. C.; Wengel, J.; Peak-Chew, S.-Y.; McLaughlin, S. H.; et al. Synthetic Genetic Polymers Capable of Heredity and Evolution. *Science* **2012**, *336* (6079), 341–344.
- (16) Mei, H.; Liao, J.-Y.; Jimenez, R. M.; Wang, Y.; Bala, S.; McCloskey, C.; Switzer, C.; Chaput, J. C. Synthesis and Evolution of a Threose Nucleic Acid Aptamer Bearing 7-Deaza-7-Substituted Guanosine Residues. *J. Am. Chem. Soc.* **2018**, *140* (17), 5706–5713.

- (17) Lutz, J.-F. Defining the Field of Sequence-Controlled Polymers. *Macromol. Rapid Commun.* **2017**, *38* (24), 1700582.
- (18) Austin, M. J.; Rosales, A. M. Tunable Biomaterials from Synthetic, Sequence-Controlled Polymers. *Biomater. Sci.* **2019**, *7* (2), 490–505.
- (19) Cho, C. Y.; Youngquist, R. S.; Paikoff, S. J.; Beresini, M. H.; Hebert, A. R.; Berleau, L. T.; Liu, C. W.; Wemmer, D. E.; Keough, T.; Schultz, P. G. Synthesis and Screening of Linear and Cyclic Oligocarbamate Libraries. Discovery of High Affinity Ligands for GPIIb/IIIa. *J. Am. Chem. Soc.* **1998**, *120* (31), 7706–7718.
- (20) Schneider, A. C.; Fritz, D.; Vasquez, J. K.; Vollrath, S. B. L.; Blackwell, H. E.; Bräse, S. Microwave-Facilitated SPOT-Synthesis of Antibacterial Dipeptoids. *ACS Comb. Sci.* **2017**, *19* (12), 715–737.
- (21) Celasun, S.; Remmler, D.; Schwaar, T.; Weller, M. G.; Du Prez, F.; Börner, H. G. Digging into the Sequential Space of Thiolactone Precision Polymers: A Combinatorial Strategy to Identify Functional Domains. *Angew. Chemie Int. Ed.* **2019**, *58* (7), 1960–1964.
- (22) Brenner, S.; Lerner, R. A. Encoded Combinatorial Chemistry. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (12), 5381–5383.
- (23) Clark, D. P.; Pazdernik, N. J.; Clark, D. P.; Pazdernik, N. J. *Polymerase Chain Reaction*; Academic Press, 2013.
- (24) Favalli, N.; Bassi, G.; Scheuermann, J.; Neri, D. DNA-Encoded Chemical Libraries - Achievements and Remaining Challenges. *FEBS Lett.* **2018**, *592* (12), 2168–2180.
- (25) Goodnow, R. A.; Dumelin, C. E.; Keefe, A. D. DNA-Encoded Chemistry: Enabling the Deeper Sampling of Chemical Space. *Nat. Rev. Drug Discov.* **2017**, *16* (2), 131–147.
- (26) Zhu, Z.; Shaginian, A.; Grady, L. C.; O’Keeffe, T.; Shi, X. E.; Davie, C. P.; Simpson, G. L.; Messer, J. A.; Evindar, G.; Bream, R. N.; et al. Design and Application of a DNA-Encoded Macrocyclic Peptide Library. *ACS Chem. Biol.* **2018**, *13* (1), 53–59.
- (27) Usanov, D. L.; Chan, A. I.; Maianti, J. P.; Liu, D. R. Second-Generation DNA-Templated Macrocyclic Libraries for the Discovery of Bioactive Small Molecules. *Nat. Chem.* **2018**, *10* (7), 704–714.
- (28) O’Reilly, R. K.; Turberfield, A. J.; Wilks, T. R. The Evolution of DNA-Templated Synthesis as a Tool for Materials Discovery. *Acc. Chem. Res.* **2017**, *50* (10), 2496–2509.
- (29) Needels, M. C.; Jones, D. G.; Tate, E. H.; Heinkel, G. L.; Kochersperger, L. M.; Dower, W. J.; Barrett, R. W.; Gallop, M. A. Generation and Screening of an Oligonucleotide-Encoded Synthetic Peptide Library. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (22), 10700–10704.
- (30) Beaucage, S. L.; Caruthers, M. H. Deoxynucleoside Phosphoramidites—A New Class of Key Intermediates for Deoxypolynucleotide Synthesis. *Tetrahedron Lett.* **1981**, *22* (20), 1859–1862.
- (31) Shen, X.; Corey, D. R. Chemistry, Mechanism and Clinical Status of Antisense Oligonucleotides and Duplex RNAs. *Nucleic Acids Res.* **2018**, *46* (4), 1584–1600.

- (32) De Rochambeau, D.; Sun, Y.; Barlog, M.; Bazzi, H. S.; Sleiman, H. F. Modular Strategy to Expand the Chemical Diversity of DNA and Sequence-Controlled Polymers. *J. Org. Chem.* **2018**, *83* (17), 9774–9786.
- (33) Bock, L. C.; Griffin, L. C.; Latham, J. A.; Vermaas, E. H.; Toole, J. J. Selection of Single-Stranded DNA Molecules That Bind and Inhibit Human Thrombin. *Nature* **1992**, *355* (6360), 564–566.
- (34) Wincott, F. E. Strategies for Oligoribonucleotide Synthesis According to the Phosphoramidite Method. *Curr. Protoc. Nucleic Acid Chem.* **2000**, *00* (1), 3.5.1-3.5.12.
- (35) Stephen A. Scaringe; Francine E. Wincott, and; Caruthers, M. H. Novel RNA Synthesis Method Using 5'-O-Silyl-2'-O-Orthoester Protecting Groups. *J. Am. Chem. Soc.* **1998**, *120* (45), 11820–11821.
- (36) Roush, W. R.; Blizzard, T. A. Synthesis of Verrucarin B. *J. Org. Chem.* **1984**, *49* (23), 4332–4339.
- (37) van Boom, J. H.; Burgers, P. M. J. Use of Levulinic Acid in the Protection of Oligonucleotides via the Modified Phosphotriester Method: Synthesis of Decaribonucleotide u-a-u-a-u-a-u-a-u-A. *Tetrahedron Lett.* **1976**, *17* (52), 4875–4878.
- (38) Lackey, J. G.; Sabatino, D.; Damha, M. J. Solid-Phase Synthesis and on-Column Deprotection of RNA from 2'- (and 3'-) O-Levulinated (Lv) Ribonucleoside Monomers. *Org. Lett.* **2007**, *9* (5), 789–792.
- (39) Horn, T.; Urdea, M. S. Forks and Combs and DNA: The Synthesis of Branched Oligodeoxyribonucleotides. *Nucleic Acids Res.* **1989**, *17* (17), 6959–6967.
- (40) Horn, T.; Chang, C. A.; Urdea, M. S. An Improved Divergent Synthesis of Comb-Type Branched Oligodeoxyribonucleotides (BDNA) Containing Multiple Secondary Sequences. *Nucleic Acids Res.* **1997**, *25* (23), 4835–4841.
- (41) Lyttle, M. H.; Walton, T. A.; Dick, D. J.; Carter, T. G.; Beckman, J. H.; Cook, R. M. New Reagents and Methods for the Synthesis of Internal and 3'-Labeled DNA. *Bioconjug. Chem.* **2002**, *13* (5), 1146–1154.
- (42) Katolik, A.; Johnsson, R.; Montemayor, E.; Lackey, J. G.; Hart, P. J.; Damha, M. J. Regiospecific Solid-Phase Synthesis of Branched Oligoribonucleotides That Mimic Intronic Lariat RNA Intermediates. *J. Org. Chem.* **2014**, *79* (3), 963–975.
- (43) Glen Research. *Glen Report 24.1: Technical Brief – Compatibility of UltraMild Monomers or Ac-DC with Hydrazine Reagent*; 2012.
- (44) Walton, T. A.; Lyttle, M. H.; Dick, D. J.; Cook, R. M. Evaluation of New Linkers and Synthetic Methods for Internal Modified Oligonucleotides. *Bioconjug. Chem.* **2002**, *13* (5), 1155–1158.
- (45) Macaya, R. F.; Schultze, P.; Smith, F. W.; Roe, J. A.; Feigon, J. Thrombin-Binding DNA Aptamer Forms a Unimolecular Quadruplex Structure in Solution. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (8), 3745–3749.

- (46) Wang, K. Y.; McCurdy, S.; Shea, R. G.; Swaminathan, S.; Bolton, P. H. A DNA Aptamer Which Binds to and Inhibits Thrombin Exhibits a New Structural Motif for DNA. *Biochemistry* **1993**, *32* (8), 1899–1904.
- (47) K Padmanabhan; K P Padmanabhan; J D Ferrara; J E Sadler; A Tulinsky. The Structure of Alpha-Thrombin Inhibited by a 15-Mer Single-Stranded DNA Aptamer. *J. Biol. Chem.* **1993**, *268*, 17651–17654.
- (48) Padmanabhan, K.; Tulinsky, A.; IUCr. An Ambiguous Structure of a DNA 15-Mer Thrombin Complex. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1996**, *52* (2), 272–282.
- (49) Kelly, J. A.; Feigon, J.; Yeates, T. O. Reconciliation of the X-Ray and NMR Structures of the Thrombin-Binding Aptamer d(GGTTGGTGTGGTTGG). *J. Mol. Biol.* **1996**, *256* (3), 417–422.
- (50) Avino, A.; Fabrega, C.; Tintore, M.; Eritja, R. Thrombin Binding Aptamer, More than a Simple Aptamer: Chemically Modified Derivatives and Biomedical Applications. *Curr. Pharm. Des.* **2012**, *18* (14), 2036–2047.
- (51) Dolot, R.; Lam, C. H.; Sierant, M.; Zhao, Q.; Liu, F.-W.; Nawrot, B.; Egli, M.; Yang, X. Crystal Structures of Thrombin in Complex with Chemically Modified Thrombin DNA Aptamers Reveal the Origins of Enhanced Affinity. *Nucleic Acids Res.* **2018**, *46* (9), 4819–4830.
- (52) Ikebukuro, K.; Okumura, Y.; Sumikura, K.; Karube, I. A Novel Method of Screening Thrombin-Inhibiting DNA Aptamers Using an Evolution-Mimicking Algorithm. *Nucleic Acids Res.* **2005**, *33* (12), e108.
- (53) Smirnov, I.; Shafer, R. H. Effect of Loop Sequence and Size on DNA Aptamer Stability. *Biochemistry* **2000**, *39* (6), 1462–1468.
- (54) Laing, B. M.; Juliano, R. L. DNA Three-Way Junctions Stabilized by Hydrophobic Interactions for Creation of Functional Nanostructures. *Chembiochem* **2015**, *16* (9), 1284–1287.
- (55) Gawande, B. N.; Rohloff, J. C.; Carter, J. D.; von Carlowitz, I.; Zhang, C.; Schneider, D. J.; Janjic, N. Selection of DNA Aptamers with Two Modified Bases. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (11), 2898–2903.
- (56) Kosuri, S.; Church, G. M. Large-Scale de Novo DNA Synthesis: Technologies and Applications. *Nat Meth* **2014**, *11* (5), 499–507.
- (57) Hervé, G.; Len, C. First Ligand-Free, Microwave-Assisted, Heck Cross-Coupling Reaction in Pure Water on a Nucleoside – Application to the Synthesis of Antiviral BVDU. *RSC Adv.* **2014**, *4* (87), 46926–46929.
- (58) Katolik, A.; Johnsson, R.; Montemayor, E.; Lackey, J. G.; Hart, P. J.; Damha, M. J. Regiospecific Solid-Phase Synthesis of Branched Oligoribonucleotides That Mimic Intronic Lariat RNA Intermediates. *J. Org. Chem.* **2014**, *79* (3), 963–975.
- (59) Kai-Chi Chang; In-Hao Su; Annamalai Senthilvelan, and; Chung, W.-S. Triazole-Modified Calix[4]Crown as a Novel Fluorescent On–Off Switchable Chemosensor. *Org. Lett.*, **2007**, *17* (9), 3363–3366.

|6|

Conclusions and Future Work

6.1 Contributions to original knowledge

Throughout this thesis, we have significantly expanded the types of monomers that can be used to build sequence-defined oligo(phosphodiester)s, and to functionalize DNA and RNA. This allowed the exploration of novel self-assembly behaviors and applications for such precision polymers in DNA-encoded combinatorial libraries.

In Chapter 2, we specifically addressed the use of a monomer capable of fluorophilic interactions that are not found in natural compounds. The strategy reported allowed the synthesis of an oligomer made of perfluorocarbon chains (PFC) to a DNA strand. This represents a breakthrough in the context of a hydrophobic polymer conjugation to DNA. Resulting “DNA-Teflon” oligomers were shown to self-assemble into almost monodisperse micelles responsive to the concentration of divalent cations in solution. More importantly, the micelles were shown to be detectable and even quantifiable through ^{19}F NMR, while fluorine-rich aggregates are typically invisible with such a technique, underlining their potential for *in vivo* imaging by MRI. When incorporated into DNA duplexes, the PFC chains interacted together through the “fluorous” effect, significantly influencing hybridization thermodynamics. The van’t Hoff plot-based study of entropy and enthalpy of hybridization was successful to determine locations at which PFC chains conferred a stability increase upon double-stranded DNA. PFC chains were also demonstrated to increase the nuclease resistance of DNA duplexes. Finally, the gene silencing ability of a silencing RNA duplex remained unaffected by its modification with a PFC chain. Thus, PFC were shown to impart valuable properties to DNA strands demonstrating the benefits of introducing new supramolecular interactions in oligo(phosphodiester)s.

With the synthesis of a novel naphthalene-containing monomer used along a hydrophilic, a hydrophobic and a fluorophilic monomer, many different interactions could be rationally studied in Chapter 3. A fine balance of these interactions was investigated along with systematic variation of oligomer length and sequence through the synthesis of a library of sequence-defined block oligo(phosphodiester)s. Sequence-control enabled by iterative automated synthesis was a unique tool to probe subtle differences in monomer sequence that could not be achieved with regular block copolymerization techniques. For example, increasing the hydrophobic content by one monomer triggered a drastic morphology switch from spherical micelles to well-defined nanosheets. This

chapter sets out elementary rational rules to predict the self-assembly outcomes for sequence-defined oligomers.

In polymer synthesis and especially in the sequence-controlled polymers field, monomers design and synthesis is a time-consuming step. Chapter 4 describes the development of a novel strategy to make a large variety of phosphoramidite monomers in two steps only. Three readily accessible molecular platforms were designed to bear chemical handles (alkyne, carboxylate, amine) for further functionalization with a moiety of interest. Contrary to most non-nucleosidic commercially available phosphoramidites, these platforms do not contain a chiral center and showed very high coupling efficiency and stability, which allows their use multiple times in the same oligomer or DNA strand. To demonstrate the versatility of this approach, phenylalanine and carbohydrate-containing phosphoramidites were successfully synthesized and introduced into fully artificial oligomer in high yields. The alkyne platform already found applications in the synthesis of DNA trimers with a specific strand pattern¹ and of sulfonated dye-DNA conjugates.² The histidine-containing moiety developed using the strategy reported in Chapter 4 and presented in Chapter 5 was used in a collaboration with the group of Prof. Juewen Liu (U. Waterloo) in the generation of DNA strands with high affinity and specificity for Zn²⁺.³ Thus, a large scope of applications was already made possible following Chapter 4 methodology.

Rational design of sequence-defined polymers is an arduous task especially in the case of artificial monomers. Combinatorial strategies are good alternatives, especially in the case of ligand discovery through DNA-encoded libraries. In Chapter 5, a valuable tool to discover aptamers made of non-nucleosidic monomers is described. The successful parallel synthesis of a sequence-defined oligomer made of the artificial monomers that were described in the other chapters, and a DNA strand coding for the oligomer identity and sequence was implemented. It led to the synthesis of a 300,000 sequence-defined oligomer aptamers library through the split-and-pool method. Amplification and sequencing were both accomplished from small amounts of these constructs to show the possibility for target affinity selection. In the DNA-encoded library field, this finding represents an access door to the new chemical space of phosphodiesteres. Moreover, the use of a solid-phase synthesis (SPS) strategy allows to envision the synthesis of much longer oligomers than what was possible through common strategies. As aptamers, our constructs are also of

significant interest since they are made of non-nucleosidic modifications that have only been rarely explored.

Overall, by increasing the number of monomers available for sequence-defined synthesis of oligo(phosphodiester)s, we have discovered new applications in drug delivery, self-assembly and ligand discovery. More importantly, this work sets up novel tools for the rational study of sequence-defined oligomers as outlined in Chapter 3, the development of more phosphoramidite monomers and a powerful synthetic method to make DNA-encoded libraries of precision poly(phosphodiester)s.

6.2 Future work

Due to their great stability, low CMC (critical micelle concentration) and ability to be detected and quantified through ^{19}F NMR, “DNA-Teflon” micelles hold great promise as potential drug carriers. Therefore, it would be of great interest to pursue research in this direction. For example, drug encapsulation tests could be performed. Once *in vivo*, micelles disassemble due to an important dilution factor and interactions with numerous biological components. Micelle crosslinking has been investigated in our group for example through the use of *N*-hydroxysuccinimide (NHS) esters in micelles with a core made of alkyl chains.⁴ This strategy could be implemented with “DNA-Teflon” micelles. For any drug delivery applications, biodistribution and bioaccumulation of PFC chains need to be tested. Indeed, the biggest limitation of “DNA-Teflon” micelles might be that the PFC chains do not get cleared and accumulate. The design and use of shorter PFC chains would be a good alternative to palliate this issue. Besides, PFC chains could be of interest to combine the “fluorous” effect with other supramolecular interactions in DNA nanostructures. For example, on the DNA cube scaffold our group developed,⁵ it would be possible to precisely position alkyl and PFC chains and study the new structures this may generate.

Fundamental studies on self-assembly and folding of sequence-defined polymers are valuable for the *de novo* design of complex nanostructures. Chapter 3 reveals the potential to make *many* sequences and study their self-assembly, at the same time by combining fast “screening” techniques such as agarose gel electrophoresis and accurate characterization methods such as atomic force microscopy. Cryo-electron microscopy would probably be a valuable tool to analyze self-assembled structures in solution in more details. Also, many parameters can be further

explored to increase our understanding of sequence-defined oligomers. We mainly limited our study to oligomers with a degree of polymerization (DP) <15 . The yields we obtained demonstrate the possibility to make longer oligomers. With a DP >50 , one could think of block co-oligomers with 10 blocks or more, giving rise to novel nanostructures. More importantly, new monomers derived from the platforms studied in Chapter 4 may be used as well. For example, the creation of sequence-defined oligo(phosphodiester)s that contain large aromatic units is an on-going project in our group. These may have applications in photonics and molecular wires.

Chapter 4 describes a methodology for making novel phosphoramidite monomers. Therefore, it can lead to many opportunities. So far, the amine-containing platform has not been used to turn a carboxylic acid into a suitable phosphoramidite. Such an experiment could be done with promising candidates such as folates. Many studies already reported the synthesis of oligonucleotide-folate conjugates for targeted delivery applications.⁶ However, folate has two carboxylates that are usually reacted non-specifically for bioconjugation. Moreover, very few studies offer a method to analyze the impact of folate positioning on a DNA strand or on a sequence-defined oligo(phosphodiester). An on-going project with Dr. Violeta Toader aims at developing a specific folate phosphoramidite that could be introduced at different positions on a DNA strand. More generally, two avenues would be worth exploring in order to make artificial proteins mimics made of poly(phosphodiester)s. The optimization of synthetic conditions to reach higher synthetic yields and therefore higher DPs of multifunctional oligomers is of interest. The generation of a library of monomers containing amino acids and short peptides would also greatly enhance the capacity of oligomers to fold similarly to proteins.

In Chapter 5, a strategy to encode sequence-defined oligo(phosphodiester)s with a DNA tag is presented. It probably holds the greatest promises of this thesis. We started a collaboration with the McKeague group to test the design of our DNA-encoded thrombin-binding aptamers and to perform selection on the library we have obtained. The presence of a hit could confirm the power of our strategy to find protein ligands and could lead to potential high-impact discoveries. For example, it could be adapted to other targets such as proteins involved in important diseases and receptors of specific cell lines for targeted therapies. The scope of DNA-encoded sequence-defined oligomers library is not limited to ligand discovery. Indeed, motivated by the discovery of potent DNazymes using a histidine modification we developed,³ we envision that sequence-defined

oligo(phosphodiester)s may be good candidates to discover novel catalysts. Ideally, these would be close in selectivity and efficiency to natural enzymes but adapted to new reactions and substrates. Such discoveries would epitomize the long-term goals of this thesis – outperforming biopolymers selectivity and efficiency with sequence-defined oligo(phosphodiester)s made of artificial building blocks.

6.3 Publications

de Rochambeau, D.; Barlóg, M.; Edwardson, T. G. W.; Fakhoury, J. J.; Stein, R. S.; Bazzi, H. S.; Sleiman, H. F. “DNA–Teflon” Sequence-Controlled Polymers. *Polym. Chem.* **2016**, *7* (31), 4998–5003.

de Rochambeau, D.; Sun, Y.; Barlog, M.; Bazzi, H. S.; Sleiman, H. F. Modular Strategy to Expand the Chemical Diversity of DNA and Sequence-Controlled Polymers. *J. Org. Chem.* **2018**, *83* (17), 9774–9786.

Trinh, T.; Saliba, D.; Liao, C.; de Rochambeau, D.; Prinzen, A. L.; Li, J.; Sleiman, H. F. “Printing” DNA Strand Patterns on Small Molecules with Control of Valency, Directionality, and Sequence. *Angew. Chemie Int. Ed.* **2019**, *58* (10), 3042–3047

de Rochambeau, D.; Barlog, M.; Bazzi, H. S.; Sleiman, H. F. Self-Assembled Nanostructure Library from Monodisperse Sequence-Defined Oligo(phosphodiester)s. ChemRxiv. Preprint. **2019**.

Lacroix, A.; Vengut-Climent, E.; de Rochambeau, D.; Sleiman, H. F. Uptake and Fate of Fluorescently Labeled DNA Nanostructures in Cellular Environments: A Cautionary Tale. *ACS Cent. Sci.* **2019**, acscentsci.9b00174

Huang, P.-J. J.; de Rochambeau, D.; Sleiman, H. F.; Liu, J. Drastically Increased Specificity by Recruiting Multiple Cooperative Metal Ions in Ligand-Modified DNazymes. *In Preparation*.

6.4 References

- (1) Trinh, T.; Saliba, D.; Liao, C.; de Rochambeau, D.; Prinzen, A. L.; Li, J.; Sleiman, H. F. “Printing” DNA Strand Patterns on Small Molecules with Control of Valency, Directionality, and Sequence. *Angew. Chemie Int. Ed.* **2019**, *58* (10), 3042–3047.
- (2) Lacroix, A.; Vengut-Climent, E.; de Rochambeau, D.; Sleiman, H. F. Uptake and Fate of Fluorescently Labeled DNA Nanostructures in Cellular Environments: A Cautionary Tale. *ACS Cent. Sci.* **2019**, acscentsci.9b00174.
- (3) Huang, P.-J. J.; de Rochambeau, D.; Sleiman, H. F.; Liu, J. Drastically Increased Specificity by Recruiting Multiple Cooperative Metal Ions in Ligand-Modified DNAzymes. *In Preparation*.
- (4) Trinh, T.; Liao, C.; Toader, V.; Barlóg, M.; Bazzi, H. S.; Li, J.; Sleiman, H. F. DNA-Imprinted Polymer Nanoparticles with Monodispersity and Prescribed DNA-Strand Patterns. *Nat. Chem.* **2018**, *10* (2), 184–192.
- (5) McLaughlin, C. K.; Hamblin, G. D.; Hänni, K. D.; Conway, J. W.; Nayak, M. K.; Carneiro, K. M. M.; Bazzi, H. S.; Sleiman, H. F. Three-Dimensional Organization of Block Copolymers on “{DNA}-Minimal” Scaffolds. *J. Am. Chem. Soc.* **2012**, *134* (9), 4280–4286.
- (6) Fernández, M.; Javaid, F.; Chudasama, V. Advances in Targeting the Folate Receptor in the Treatment/Imaging of Cancers. *Chem. Sci.* **2018**, *9* (4), 790–810.

[7]

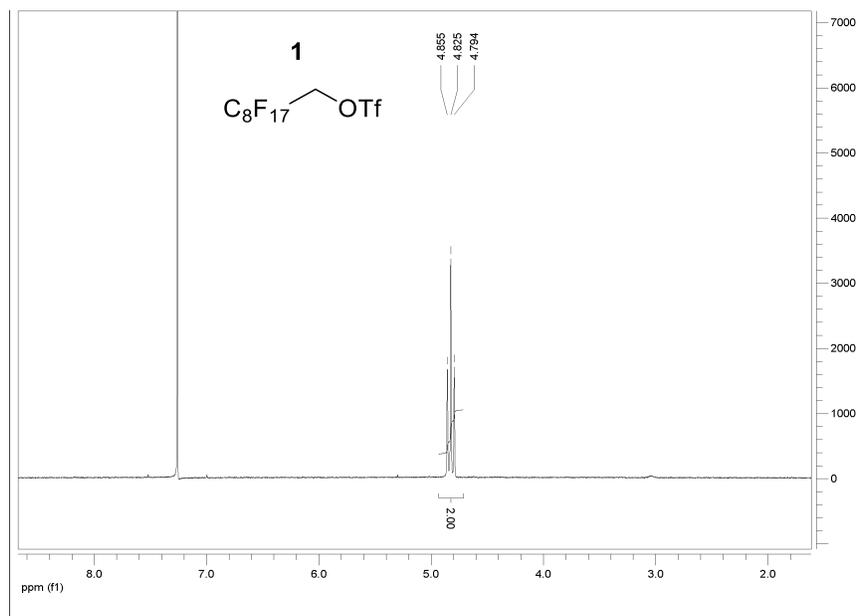
NMR Spectra

A ^1H NMR spectrum is provided for each small molecule synthesized. A ^{13}C NMR spectrum is included for each unreported compound.

7.1 Molecules from Chapter 2

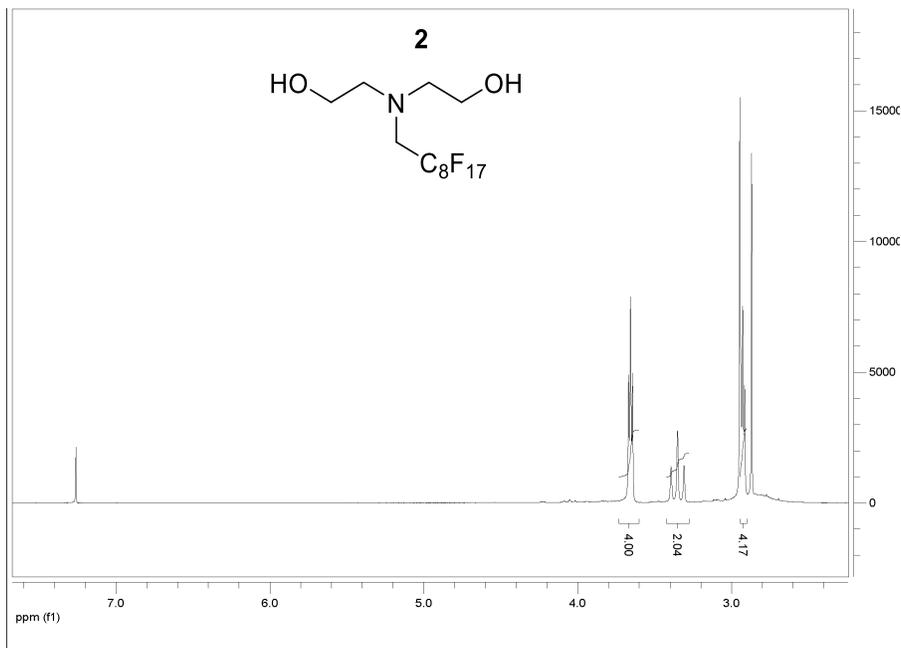
7.1.1 Compound 1

^1H NMR (400 MHz), CDCl_3 CDCl_3



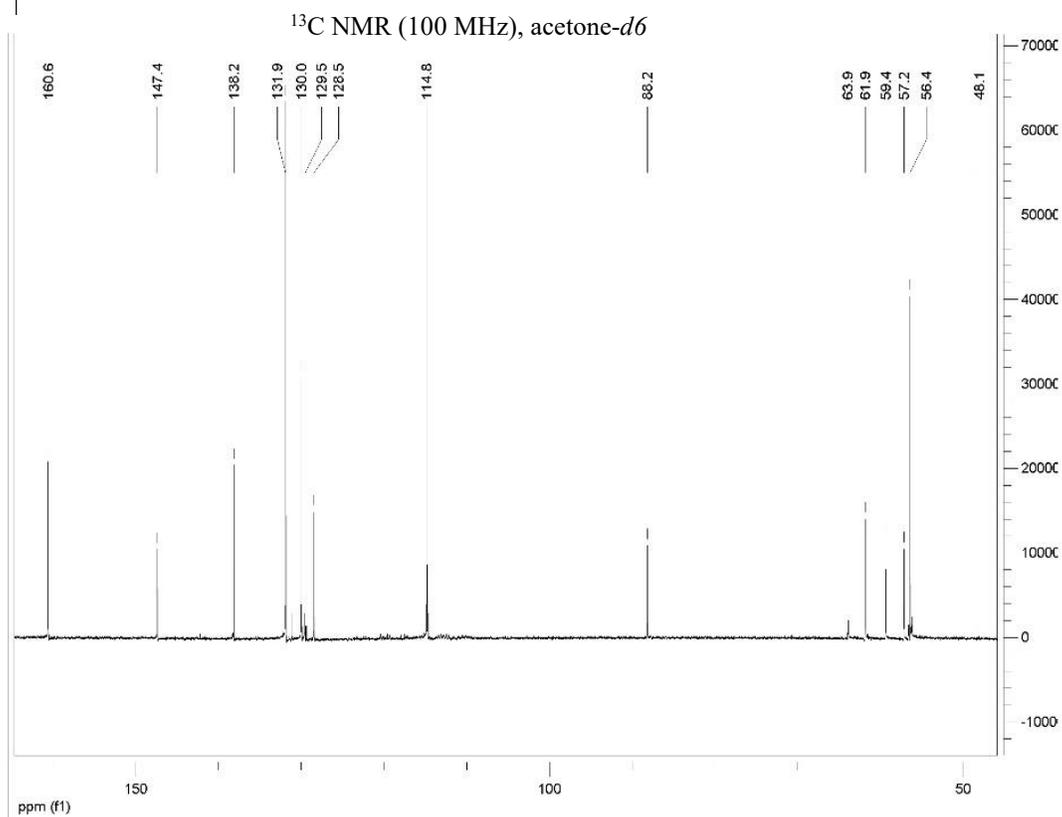
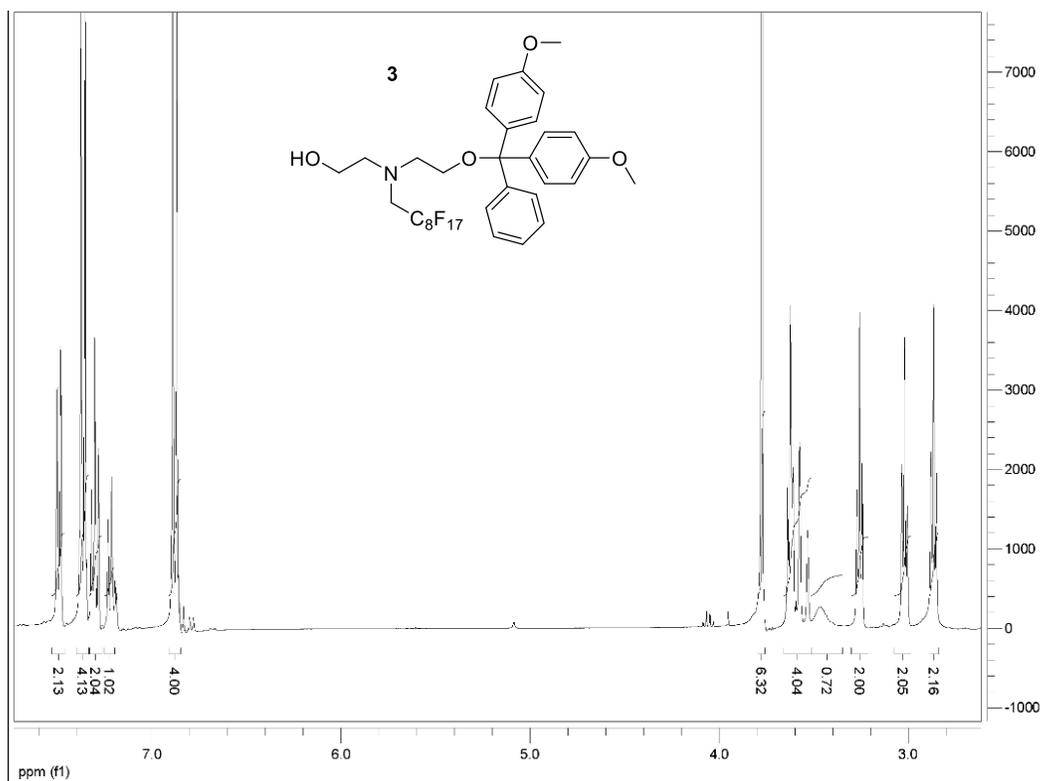
7.1.2 Compound 2

^1H NMR (400 MHz), CDCl_3

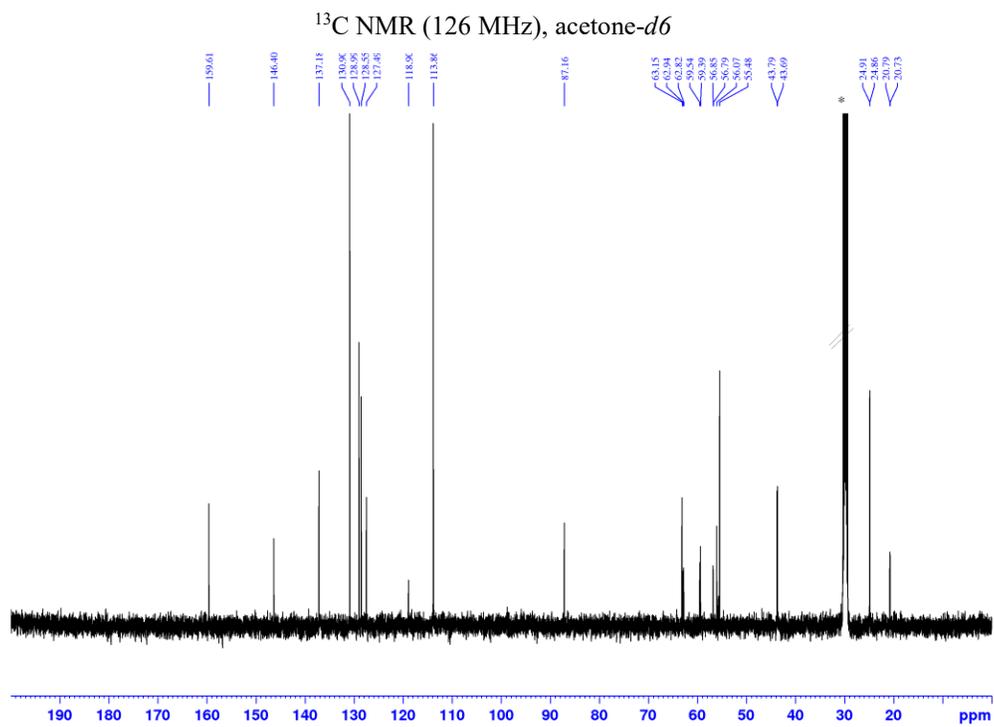
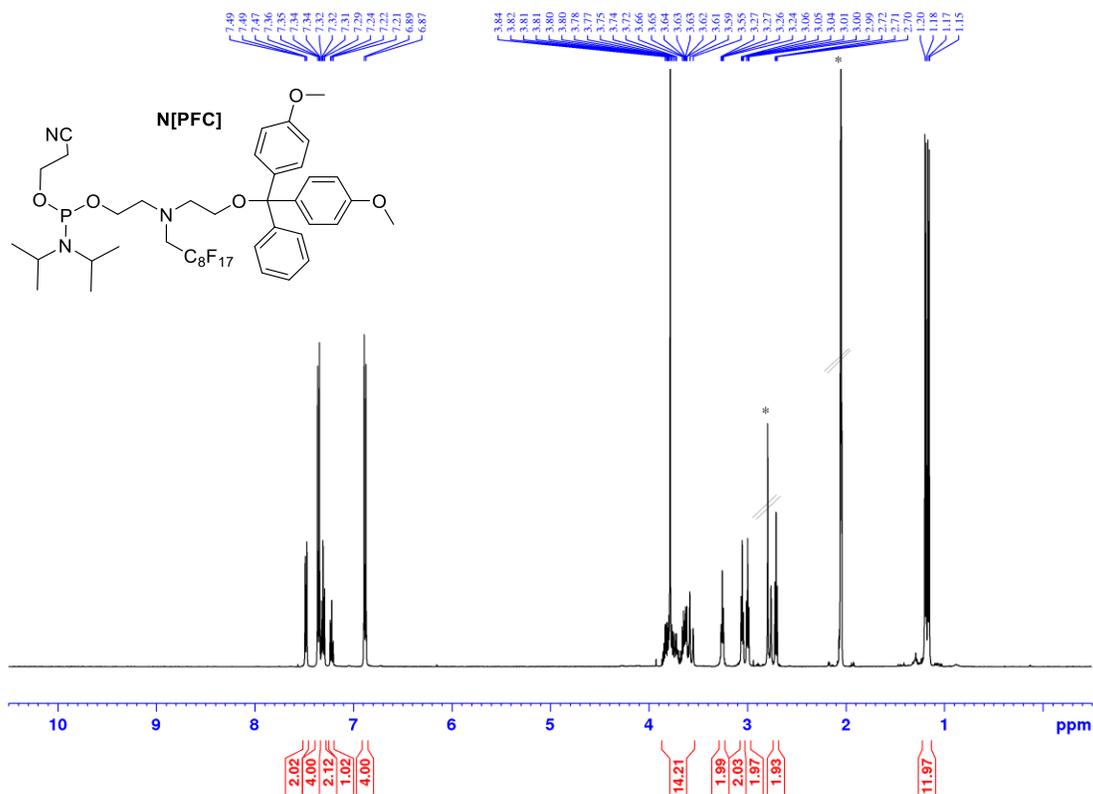


7.1.3 Compound 3

^1H NMR (400 MHz), acetone- d_6

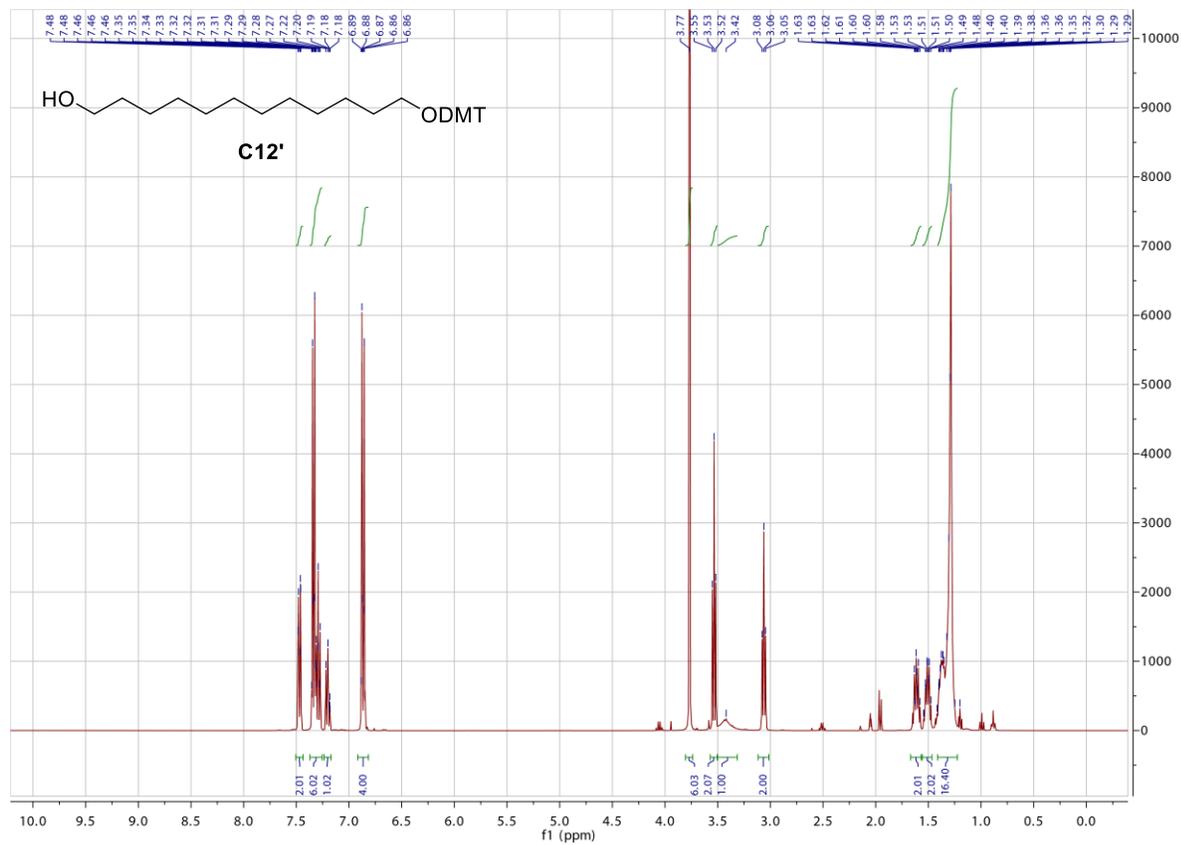


7.1.4 Monomer N[PFC] ¹H NMR (500 MHz), acetone-*d*₆

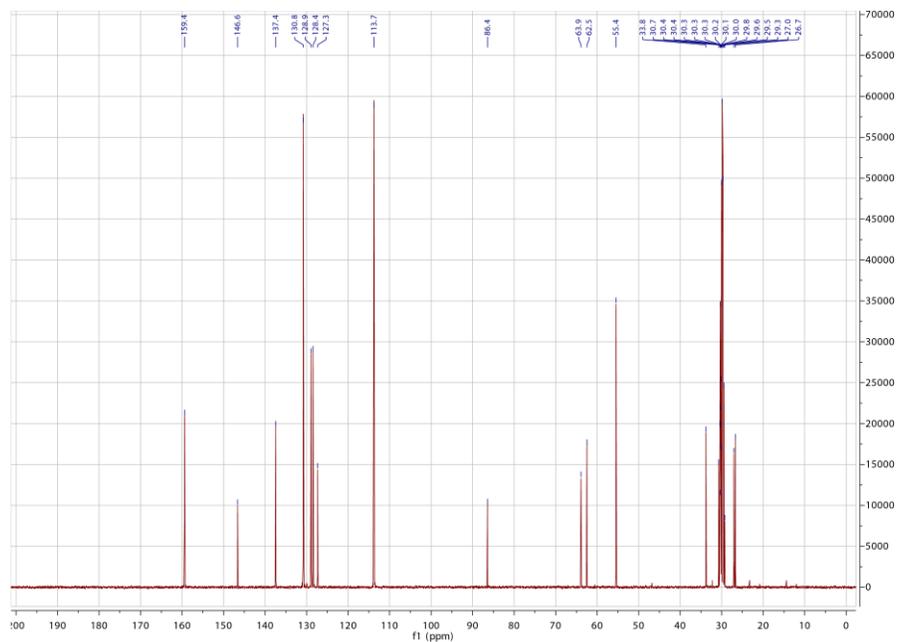


7.2 Molecules from Chapter 3

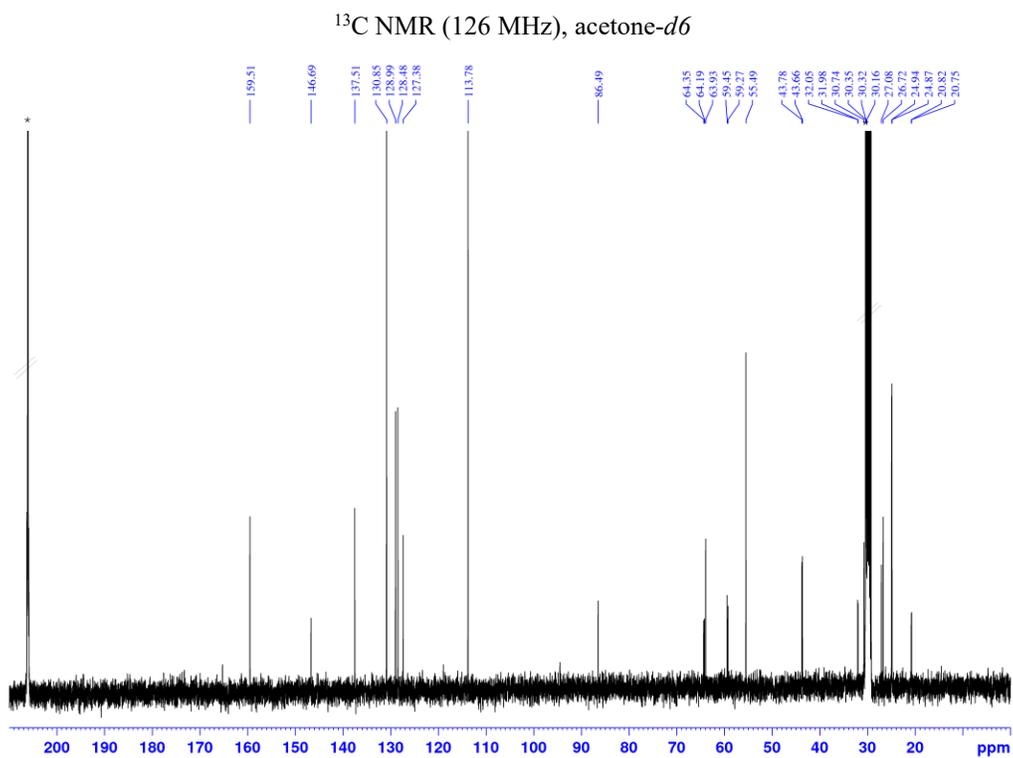
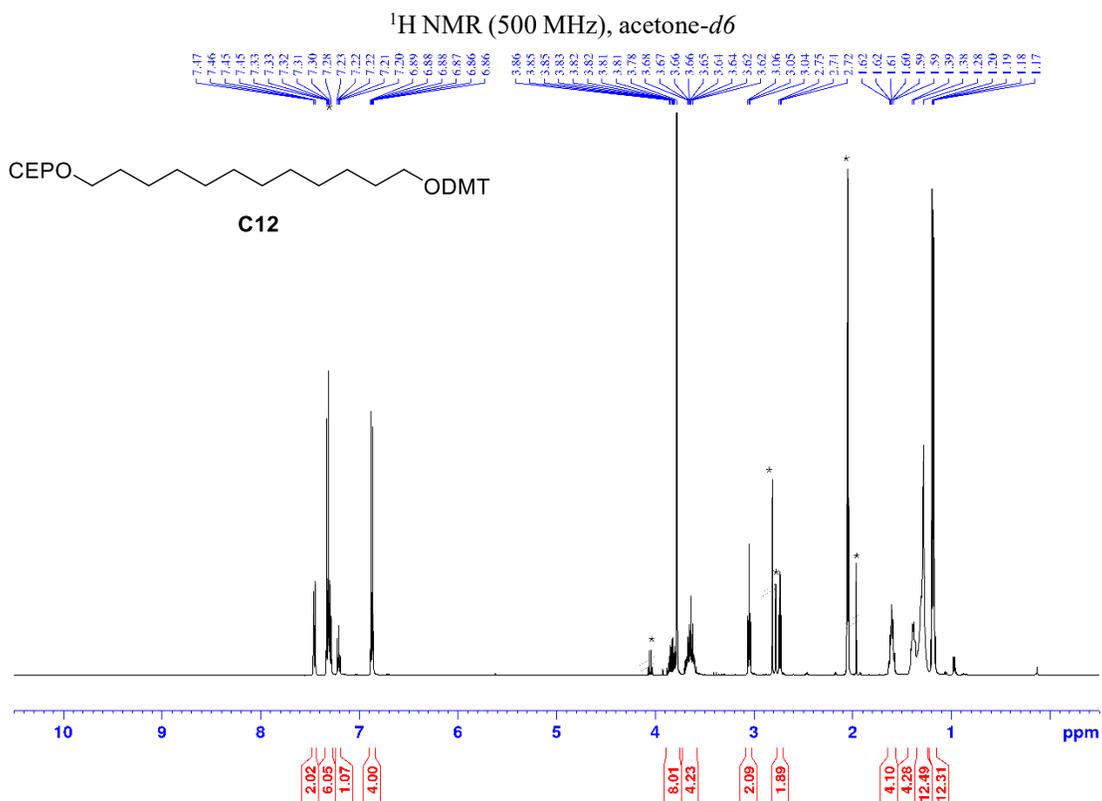
7.2.1 Compound C12' ¹H NMR (400 MHz), acetone-*d*₆



¹³C NMR (100 MHz), acetone-*d*₆

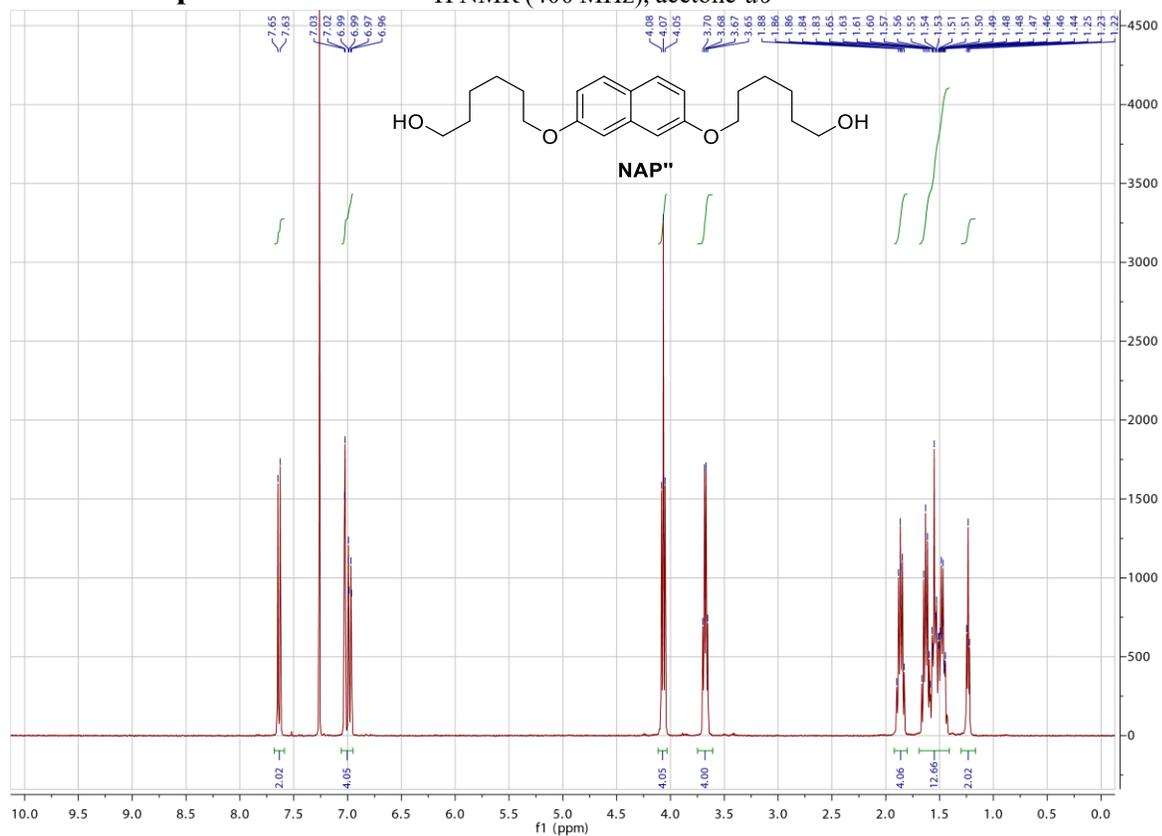


7.2.2 Monomer C12

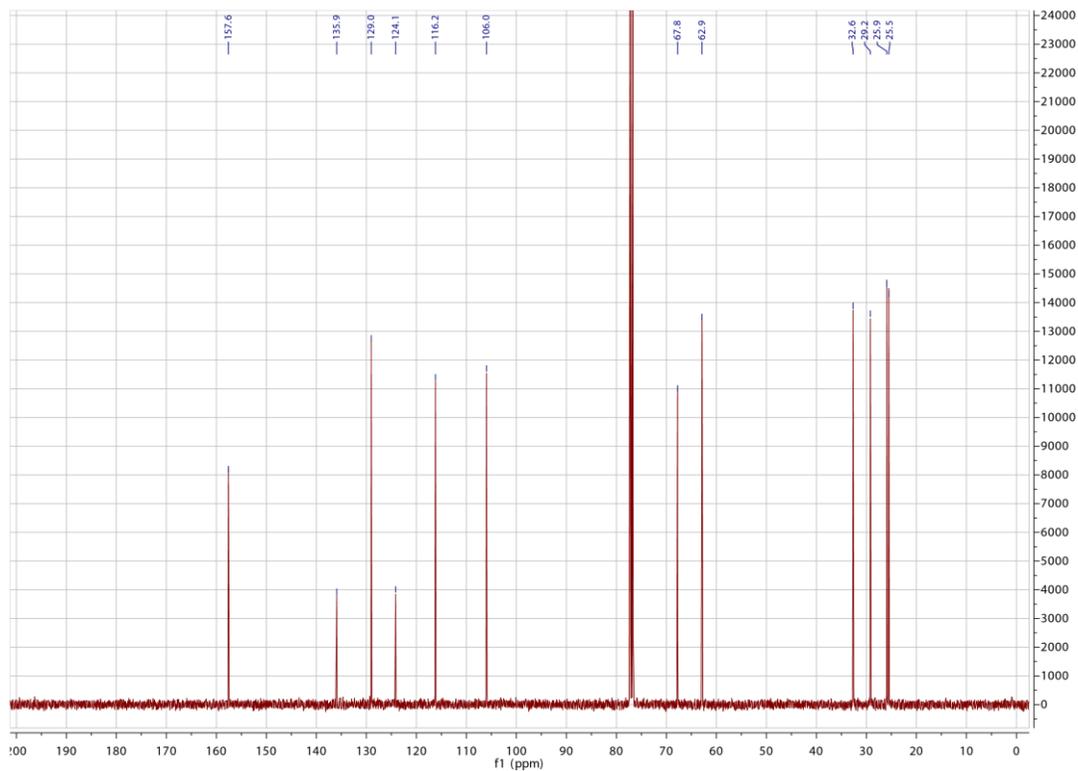


7.2.3 Compound NAP''

¹H NMR (400 MHz), acetone-*d*₆

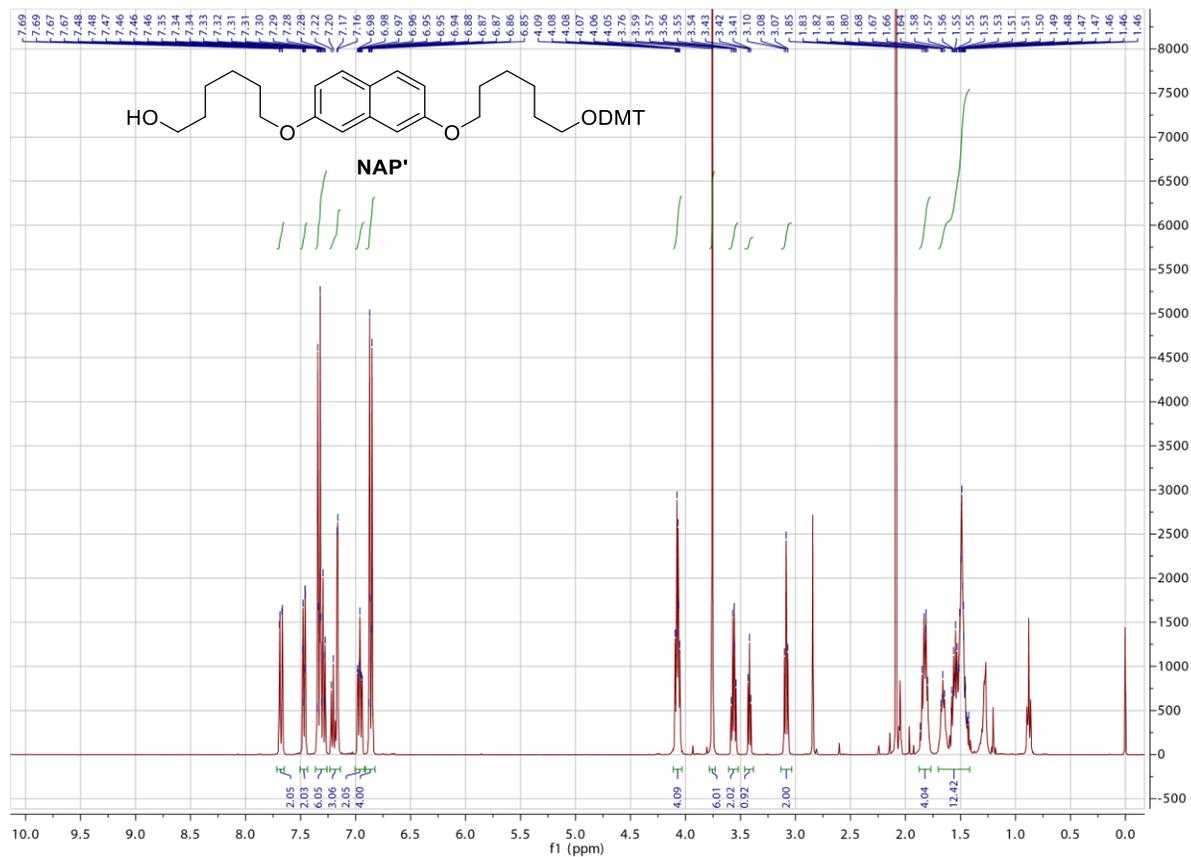


¹³C NMR (100 MHz), acetone-*d*₆

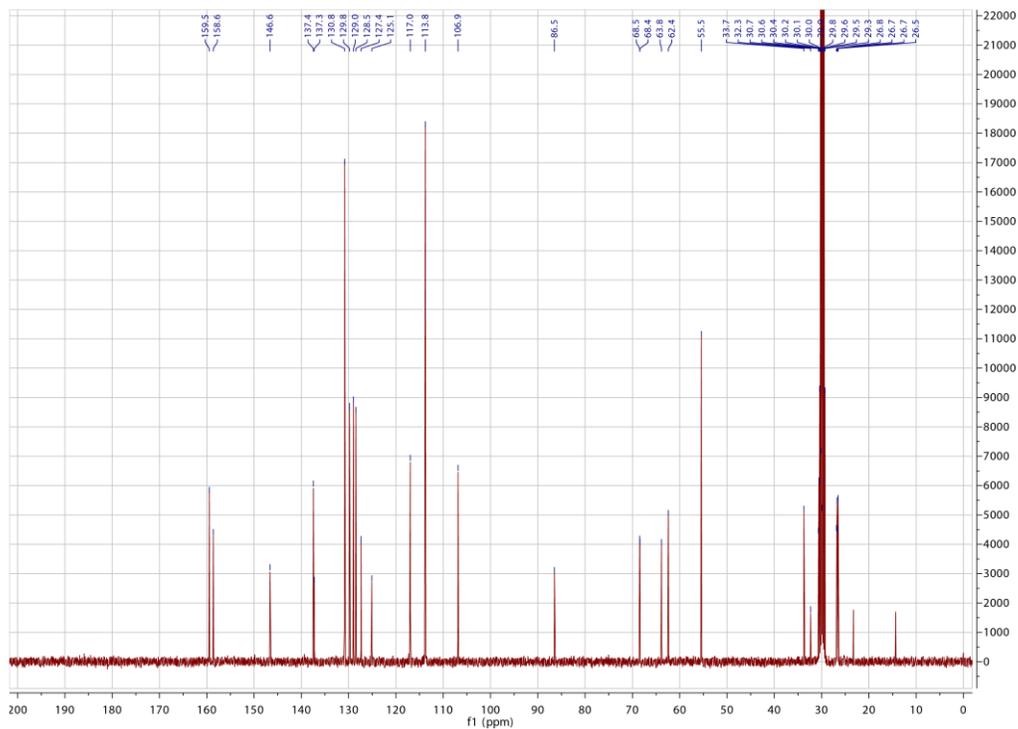


7.2.4 Compound NAP'

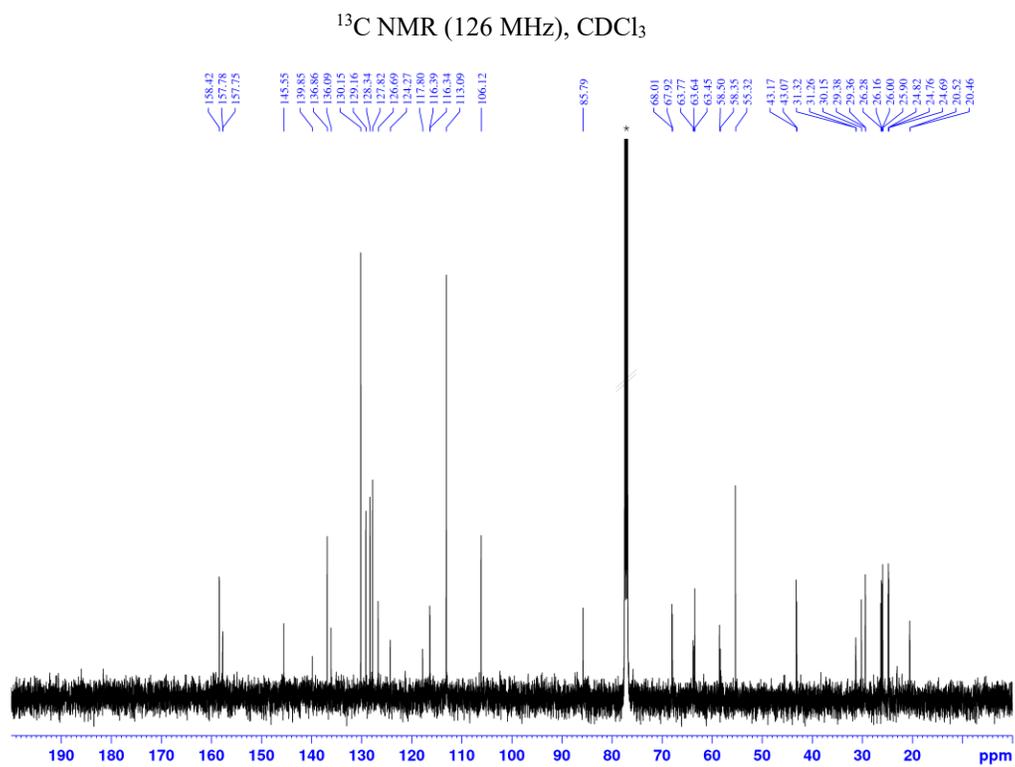
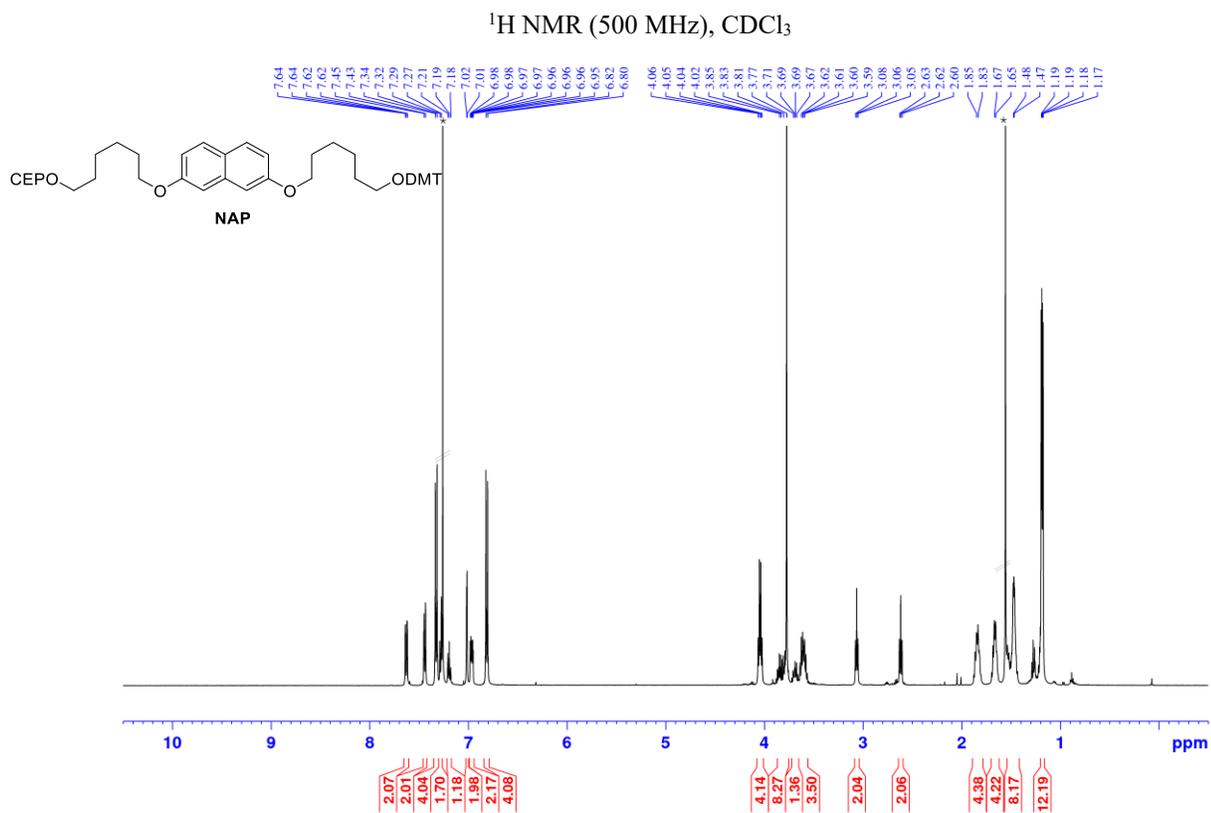
¹H NMR (400 MHz), acetone-*d*₆



¹³C NMR (100 MHz), acetone-*d*₆

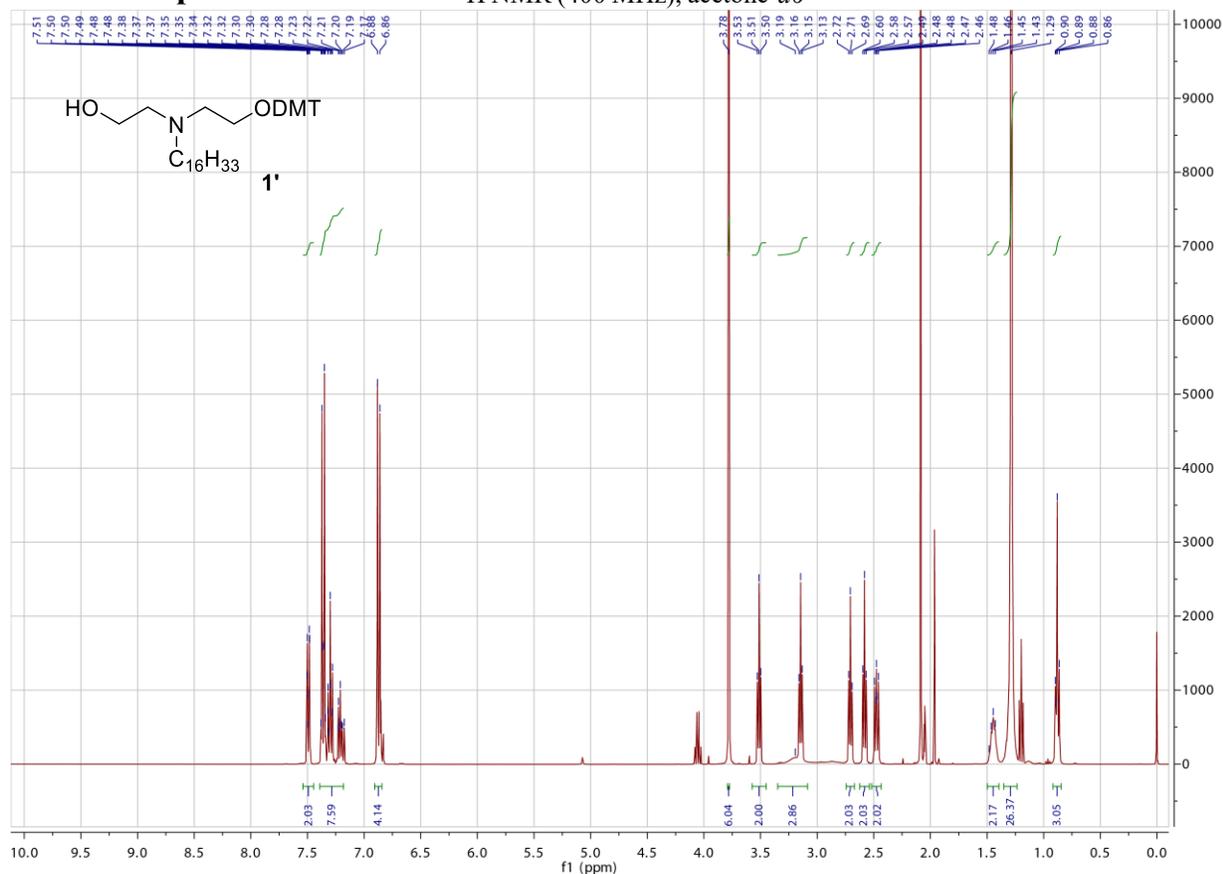


7.2.5 Monomer NAP

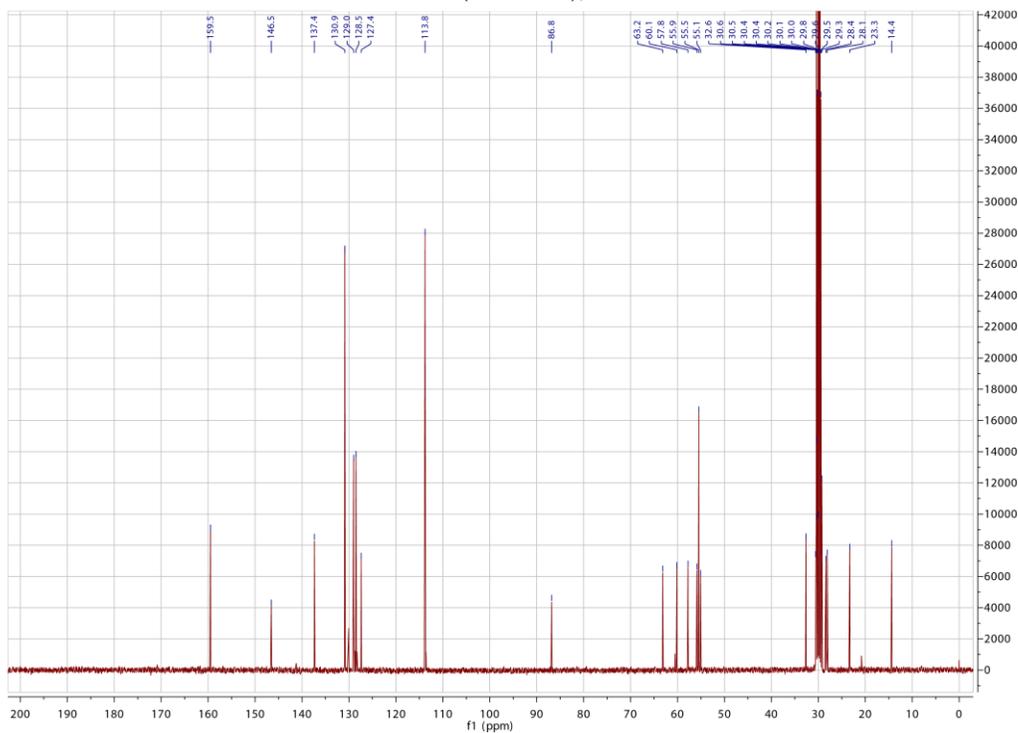


7.3.2 Compound 1'

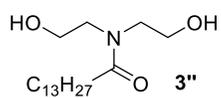
¹H NMR (400 MHz), acetone-d₆



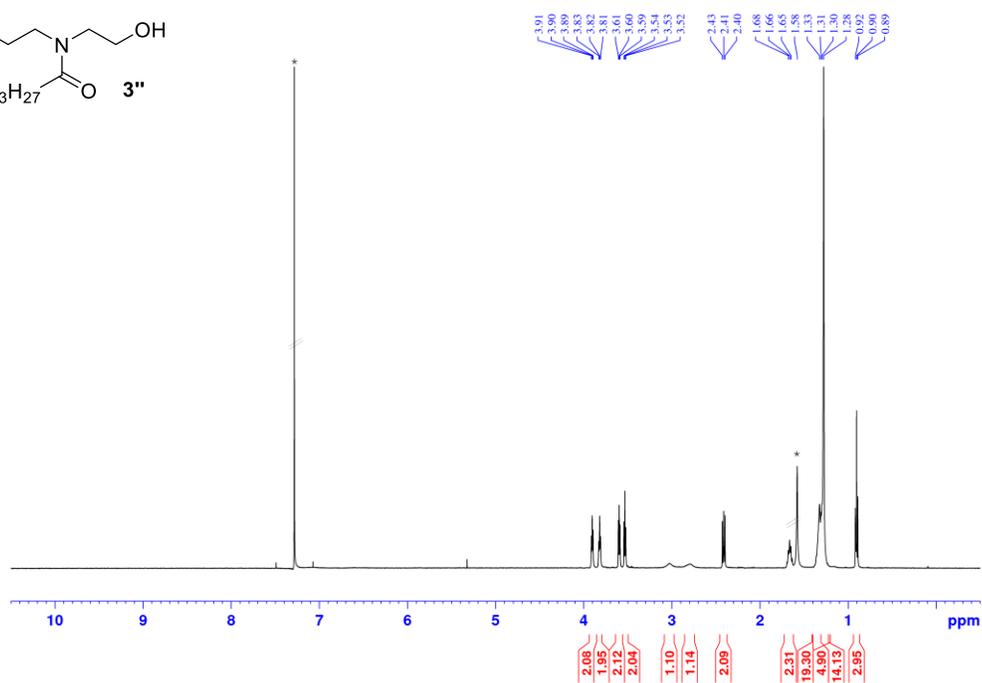
¹³C NMR (100 MHz), acetone-d₆



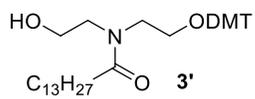
7.3.3 Compound 3''



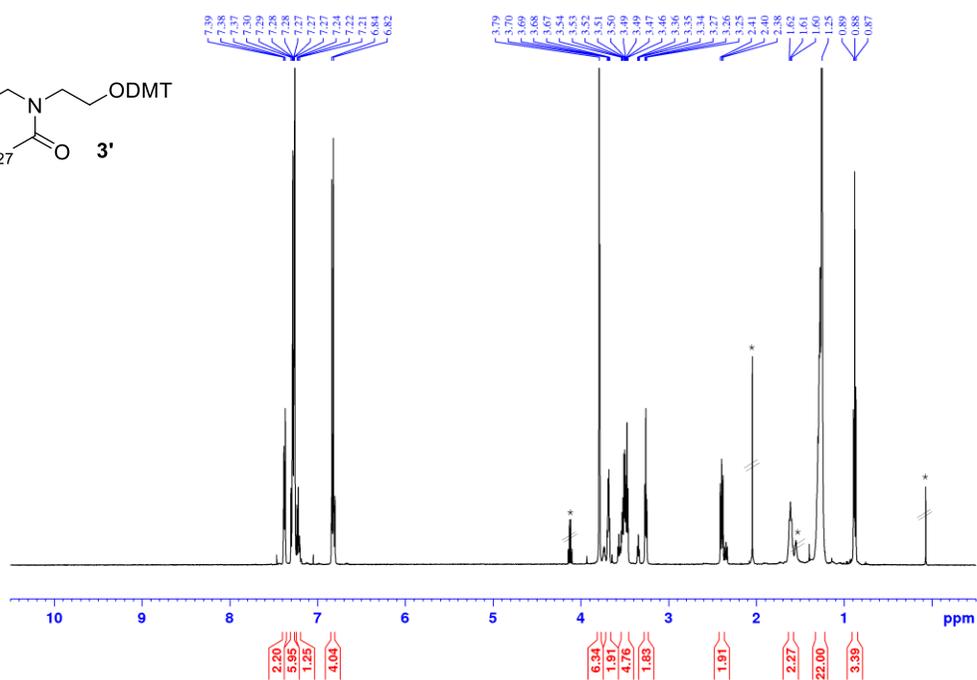
1H NMR (500 MHz), $CDCl_3$

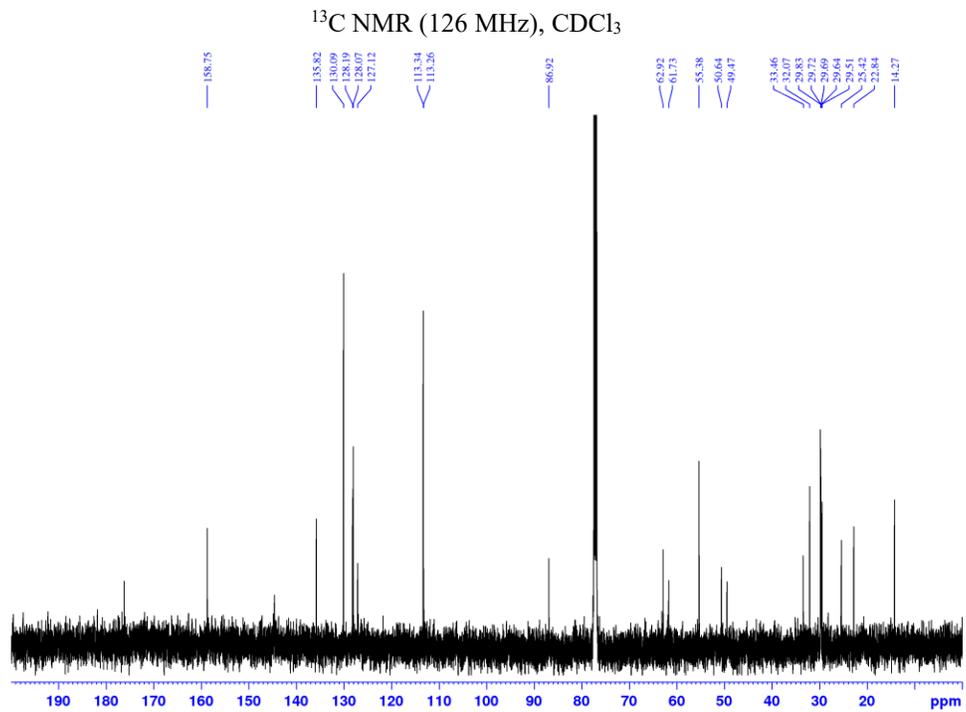


7.3.4 Compound 3'

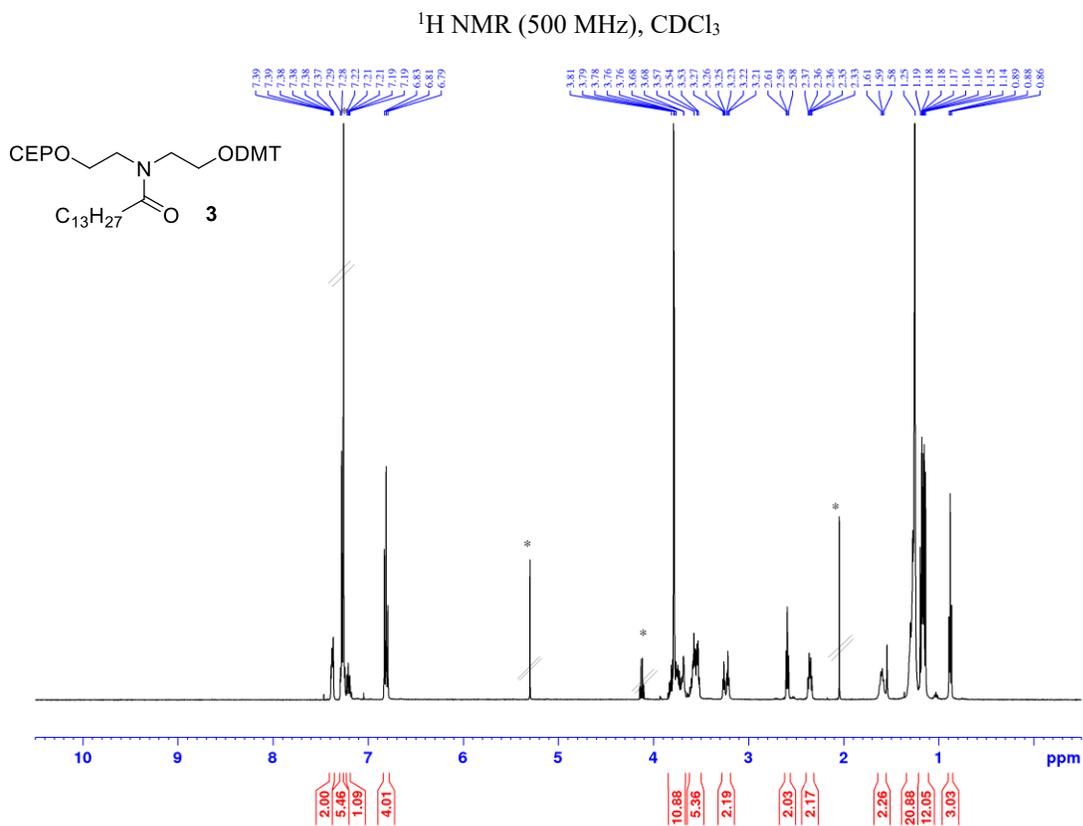


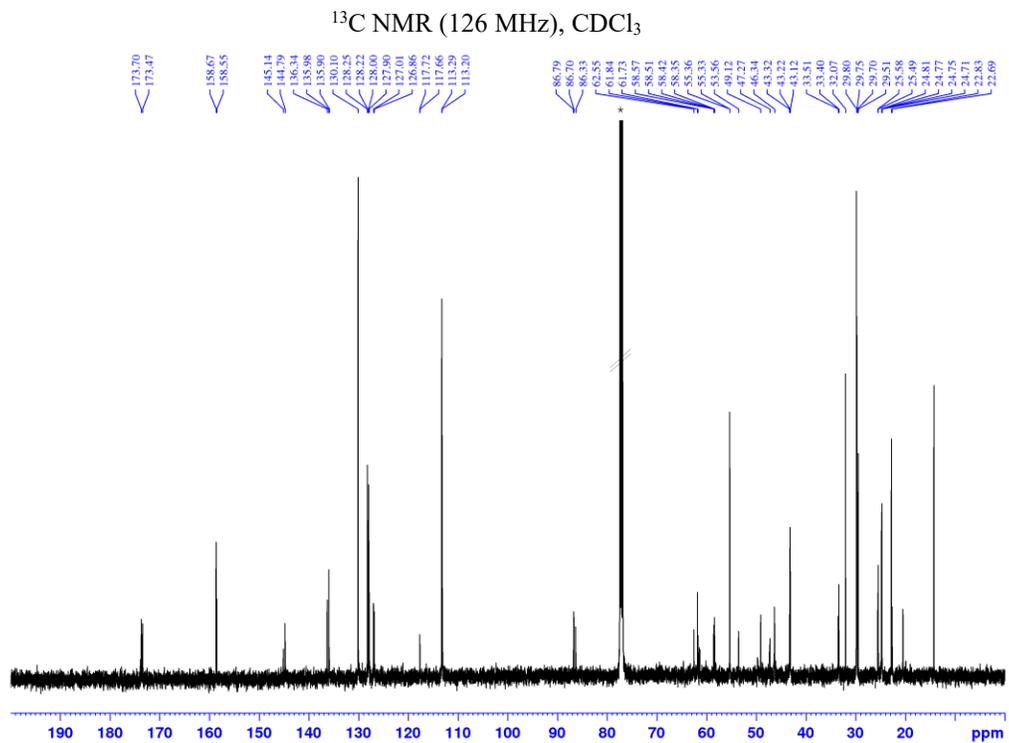
1H NMR (500 MHz), $CDCl_3$





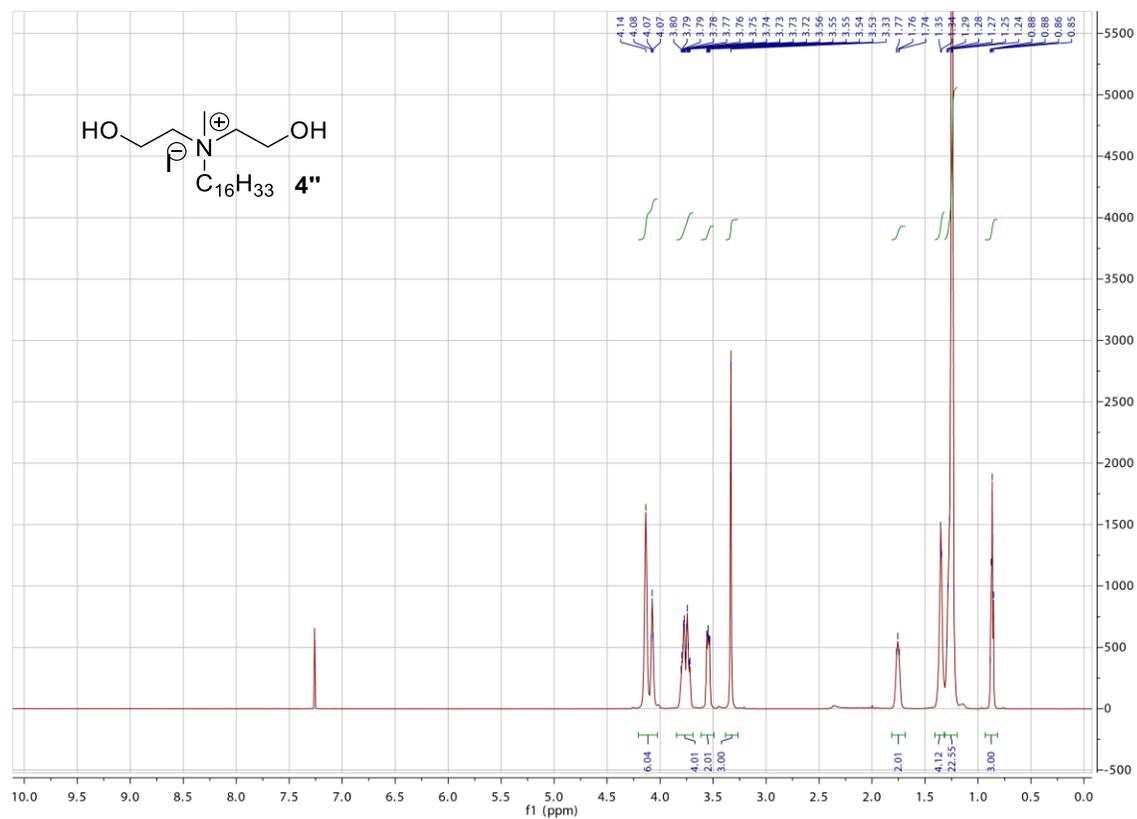
7.3.5 Monomer 3

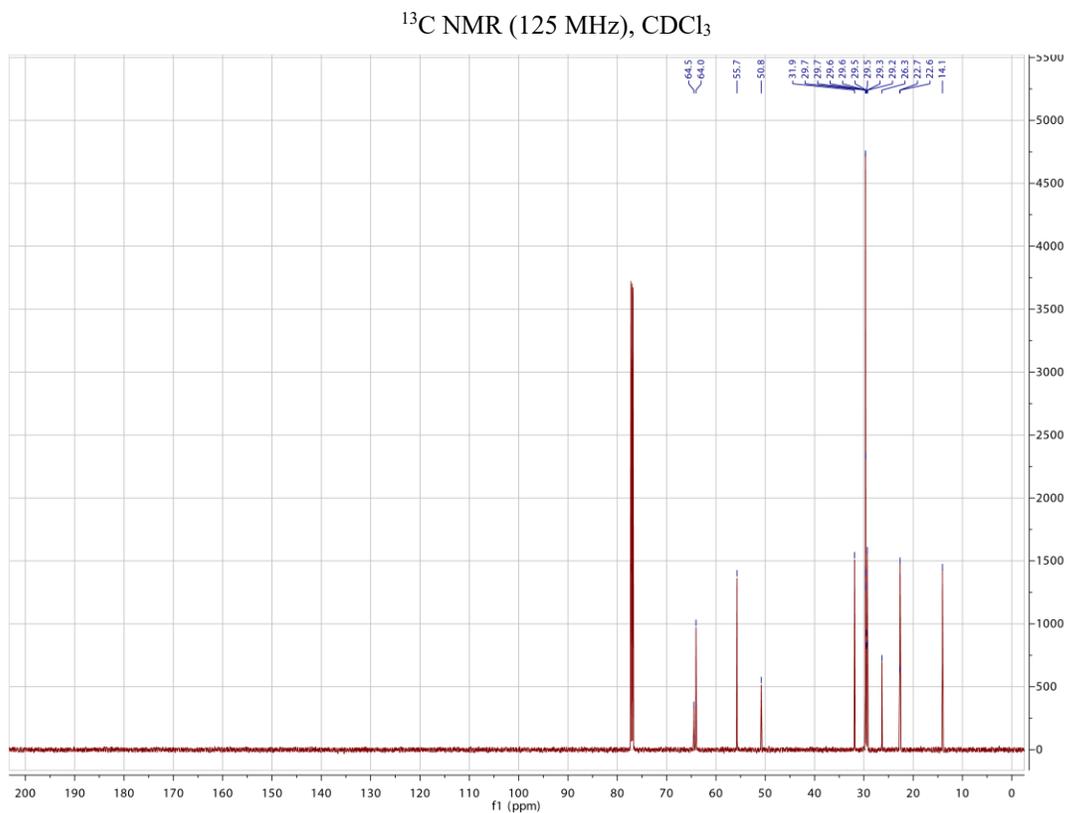




7.3.6 Compound 4''

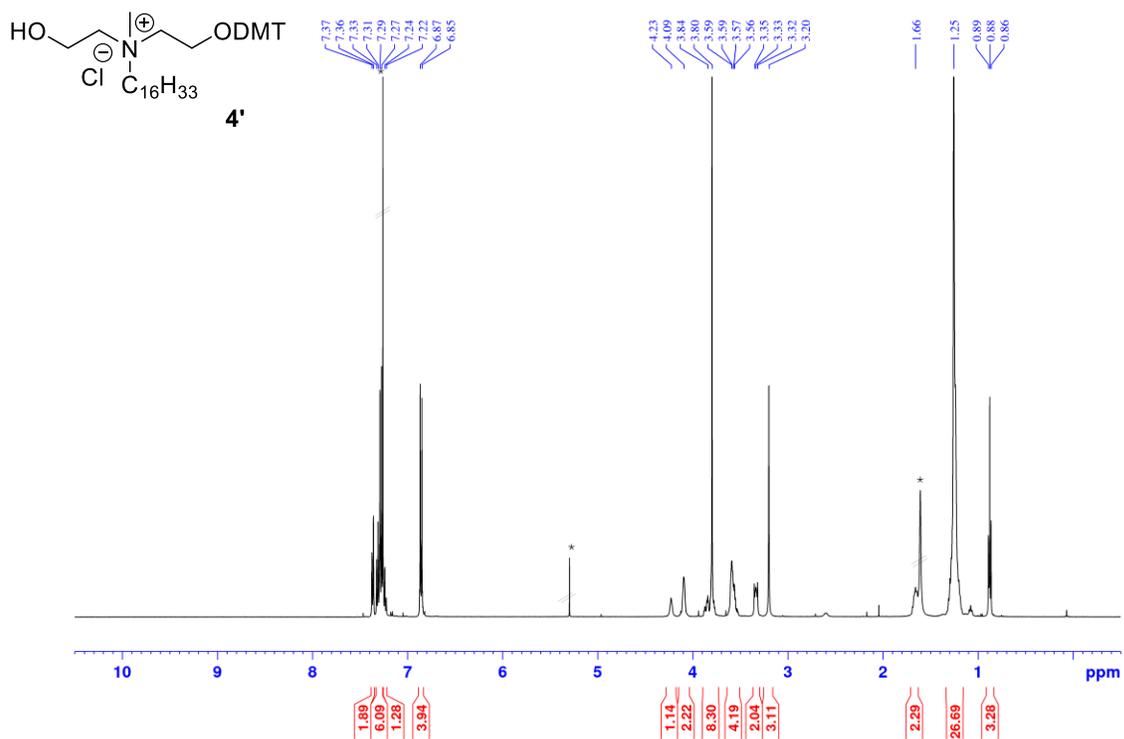
^1H NMR (600 MHz), CDCl_3



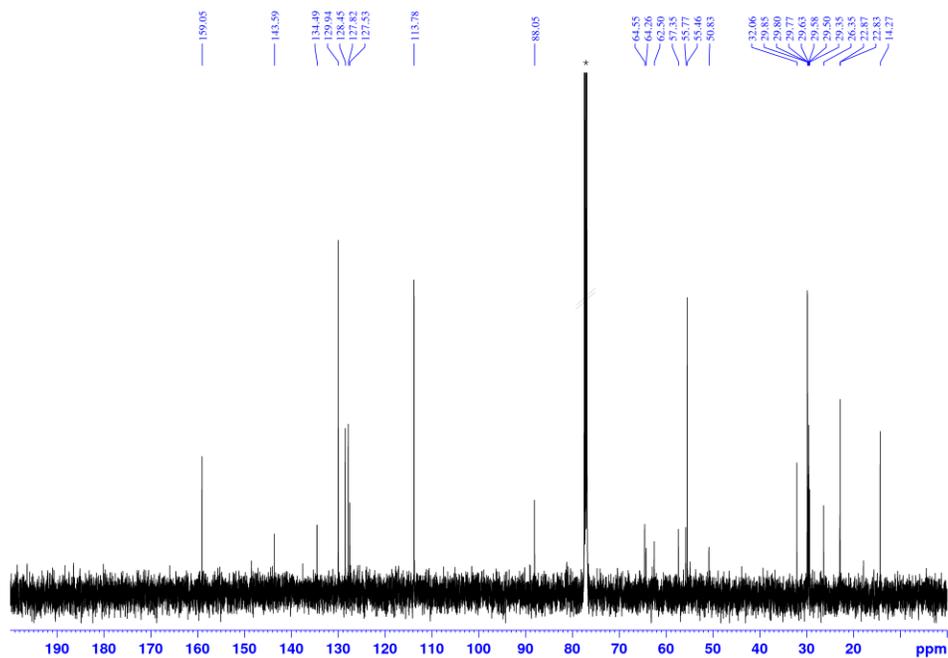


7.3.7 Compound 4'

^1H NMR (500 MHz), CDCl_3

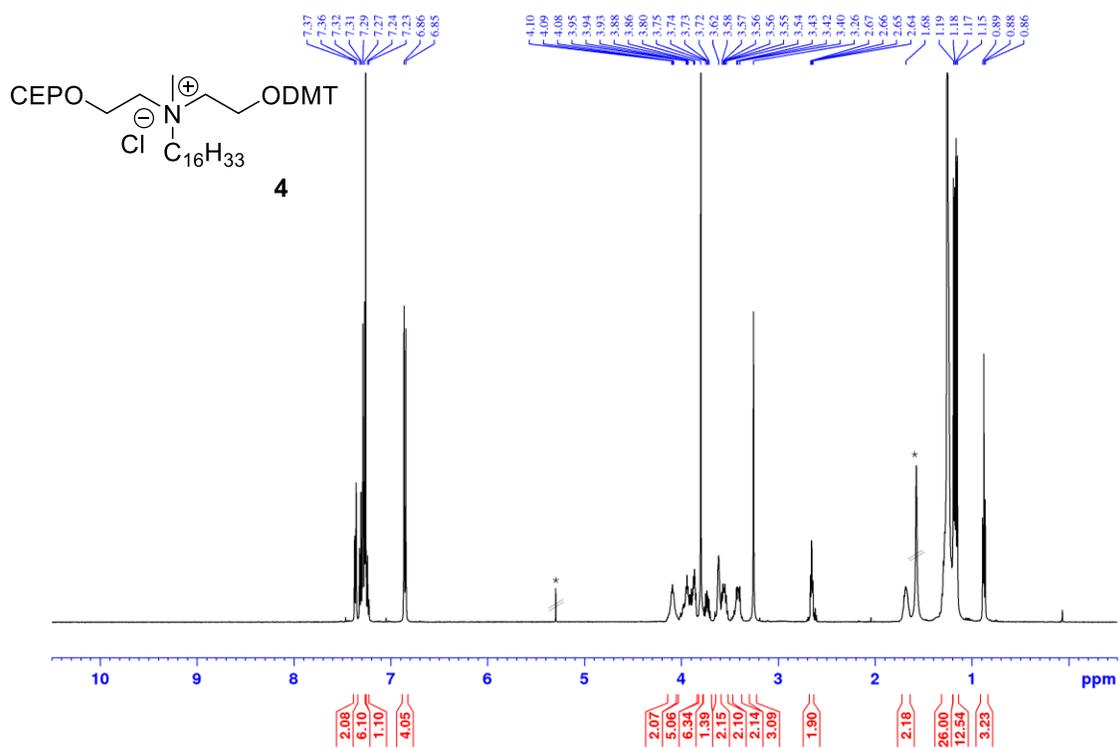


^{13}C NMR (125 MHz), CDCl_3

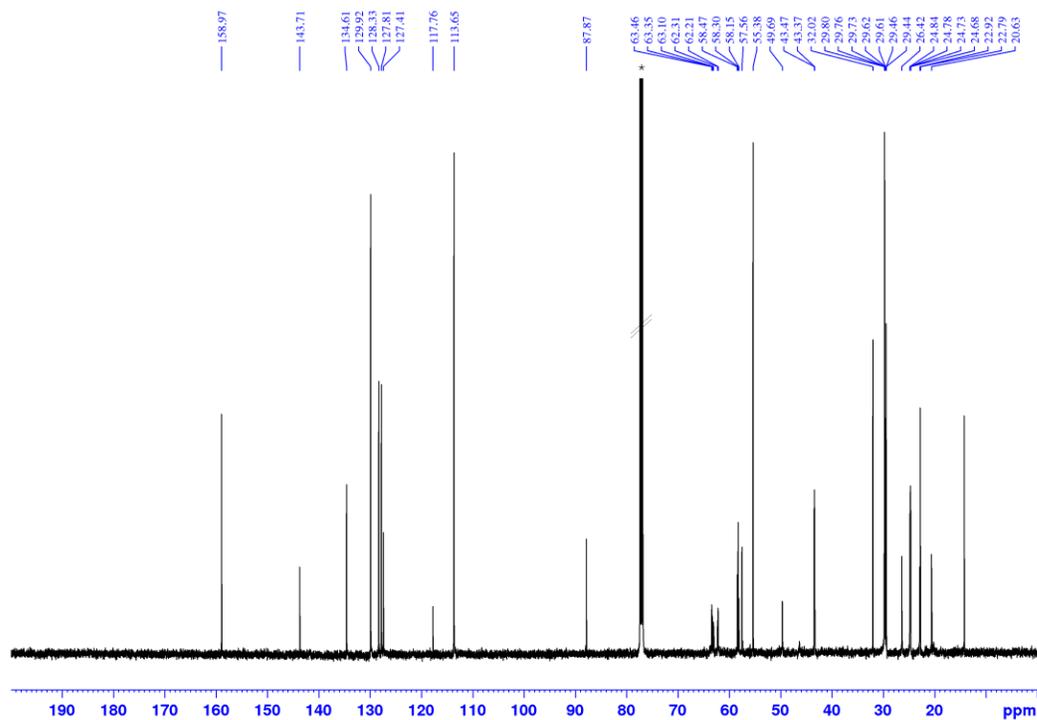


7.3.8 Monomer 4

^1H NMR (500 MHz), CDCl_3

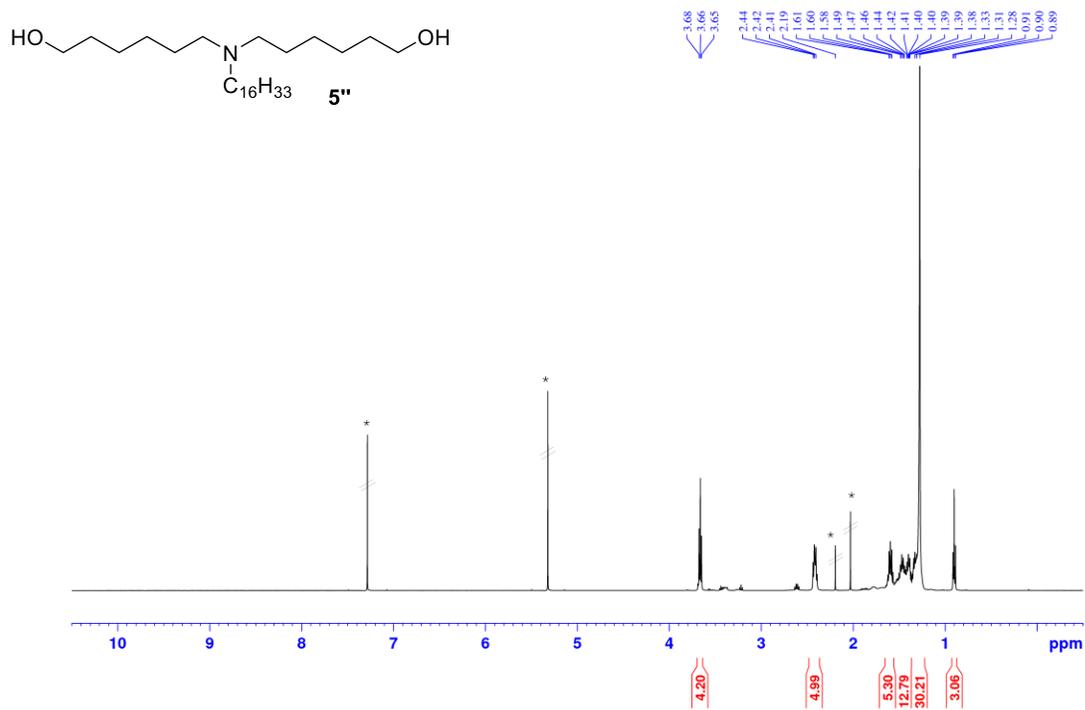


¹³C NMR (126 MHz), CDCl₃



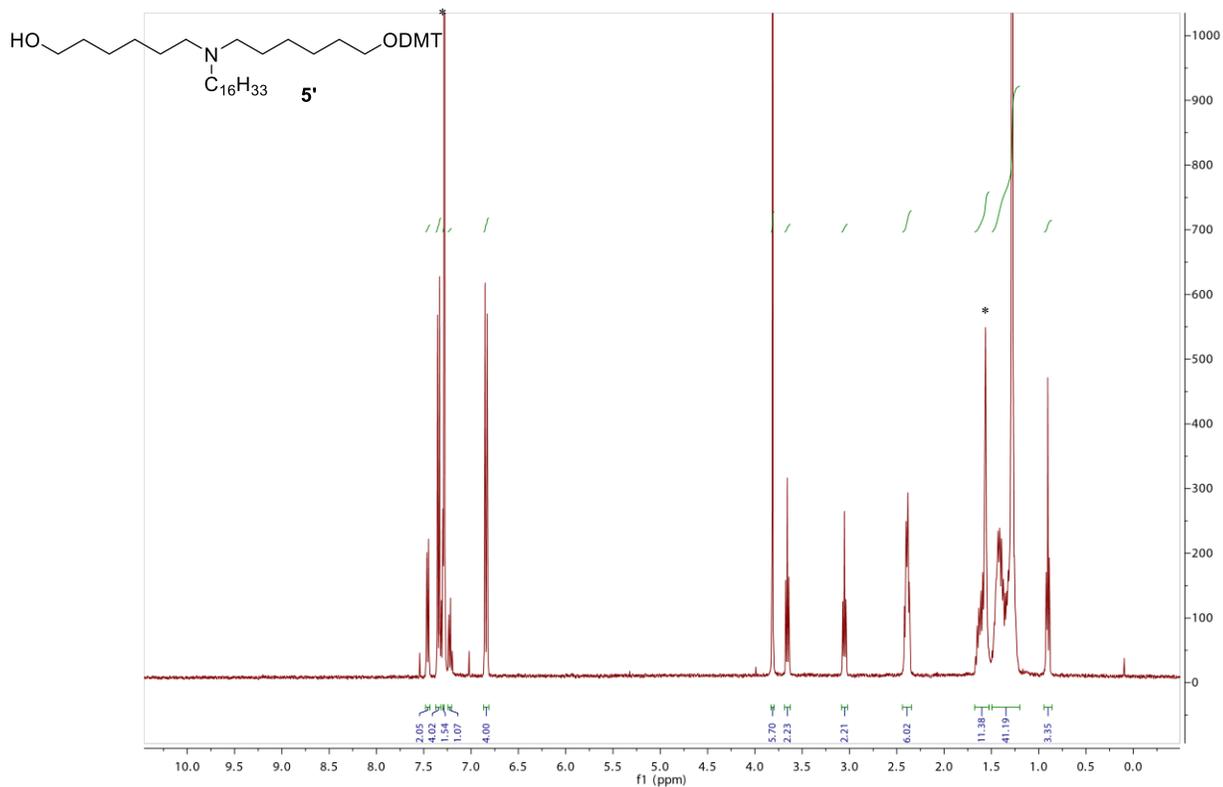
7.3.9 Compound 5''

¹H NMR (500 MHz), CDCl₃



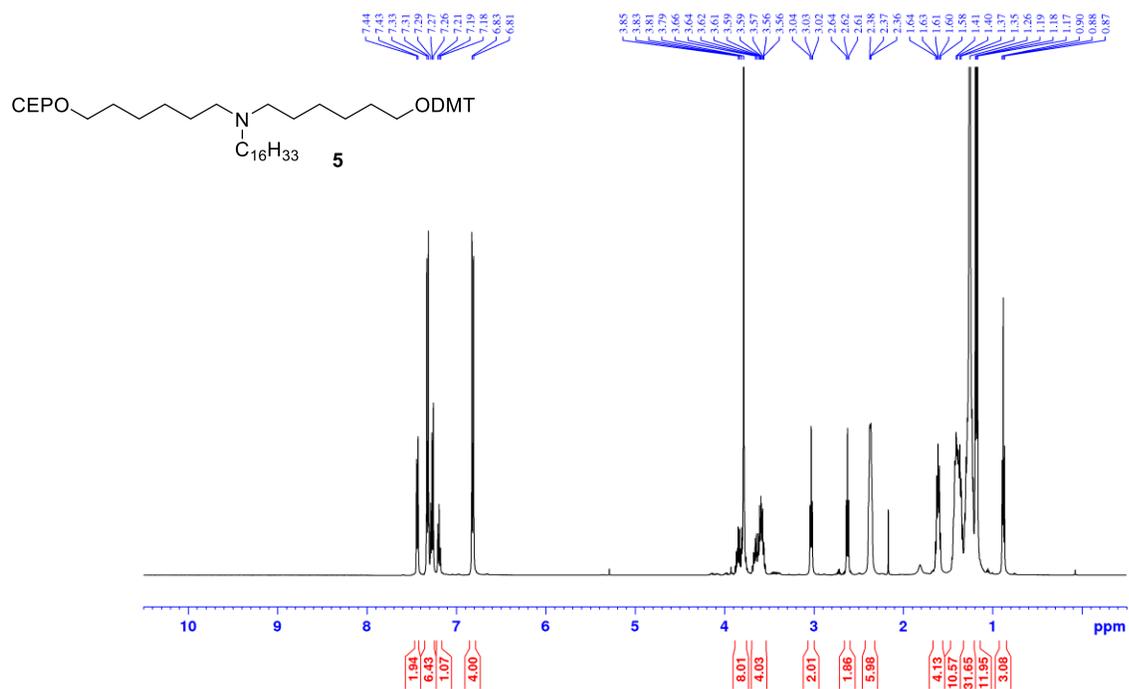
7.3.10 Compound 5'

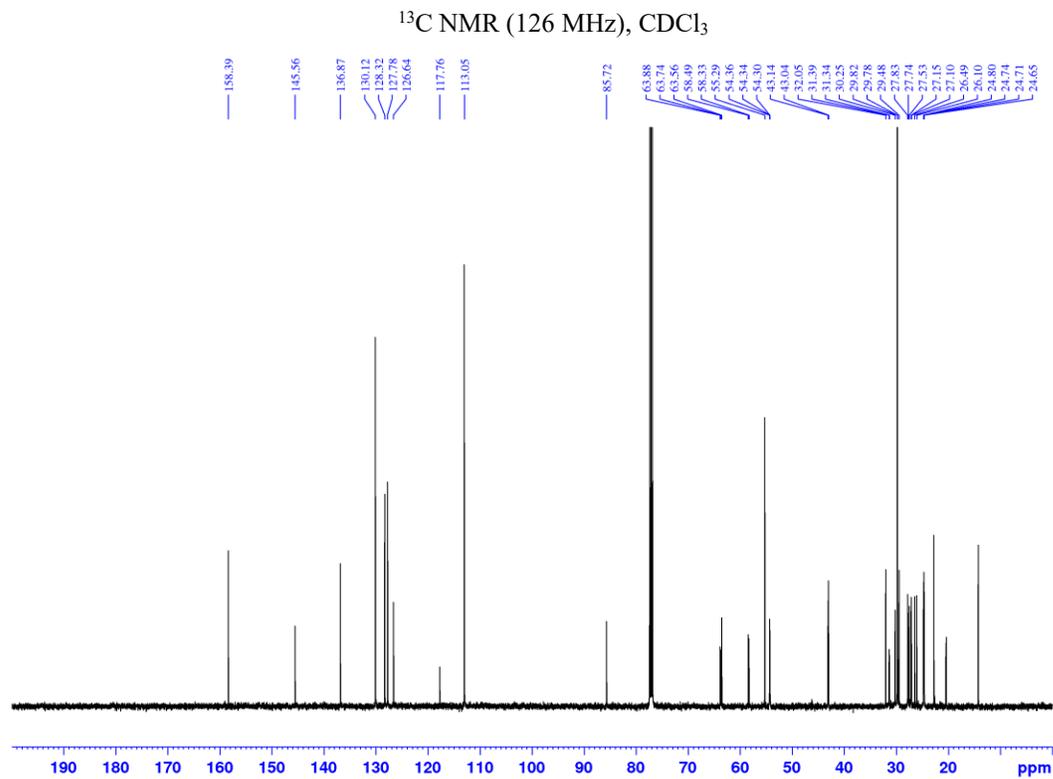
¹H NMR (400 MHz), CDCl₃



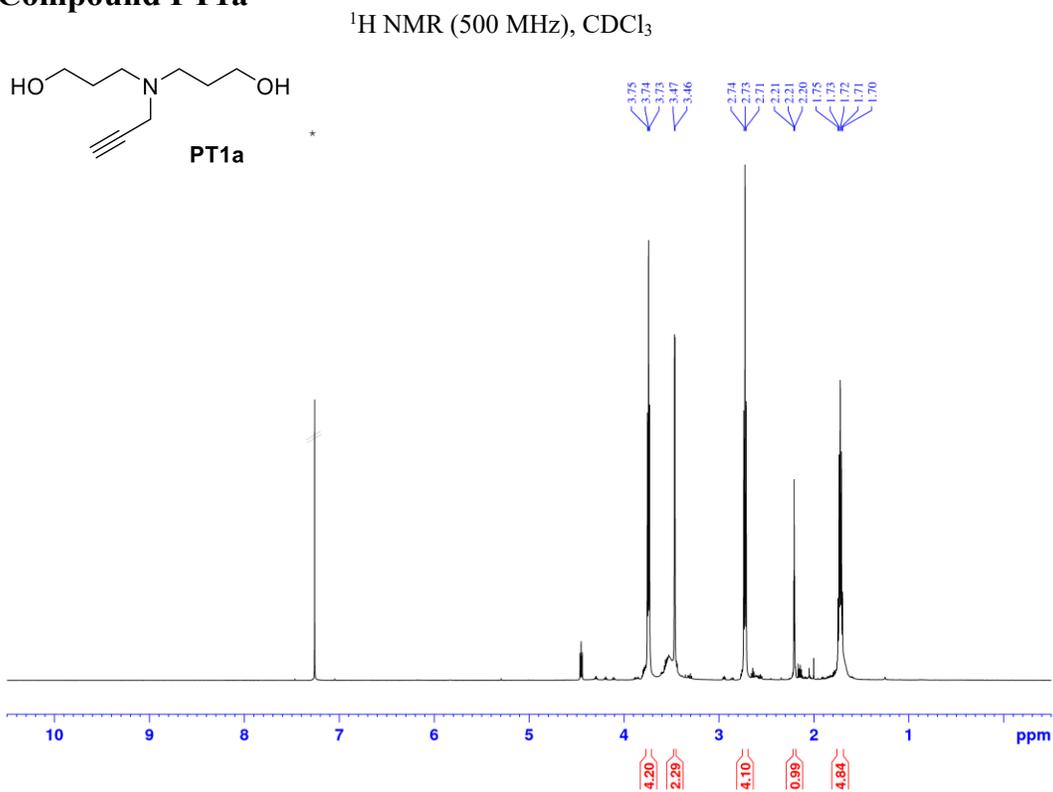
7.3.11 Monomer 5

¹H NMR (500 MHz), CDCl₃



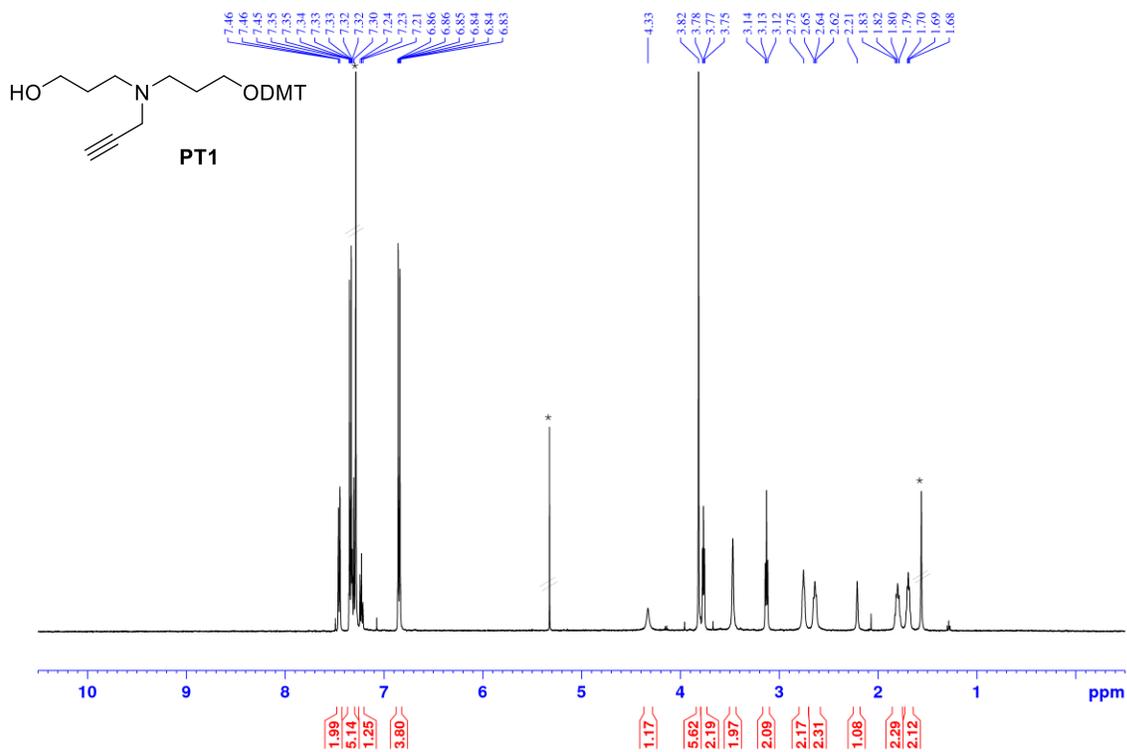


7.3.12 Compound PT1a

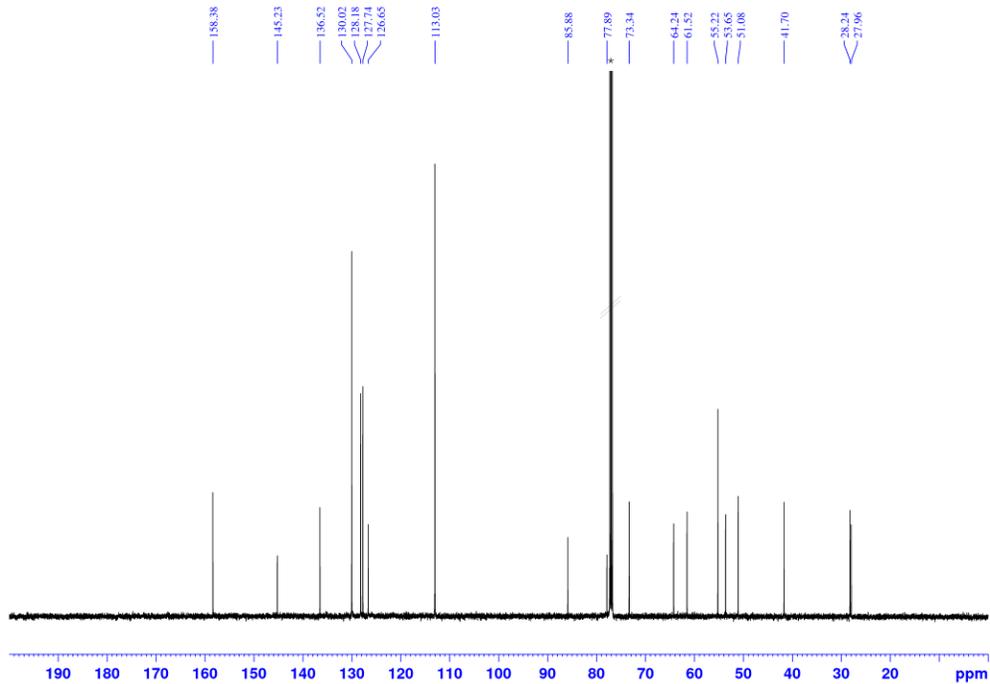


7.3.13 Platform PT1

$^1\text{H NMR}$ (500 MHz), CDCl_3

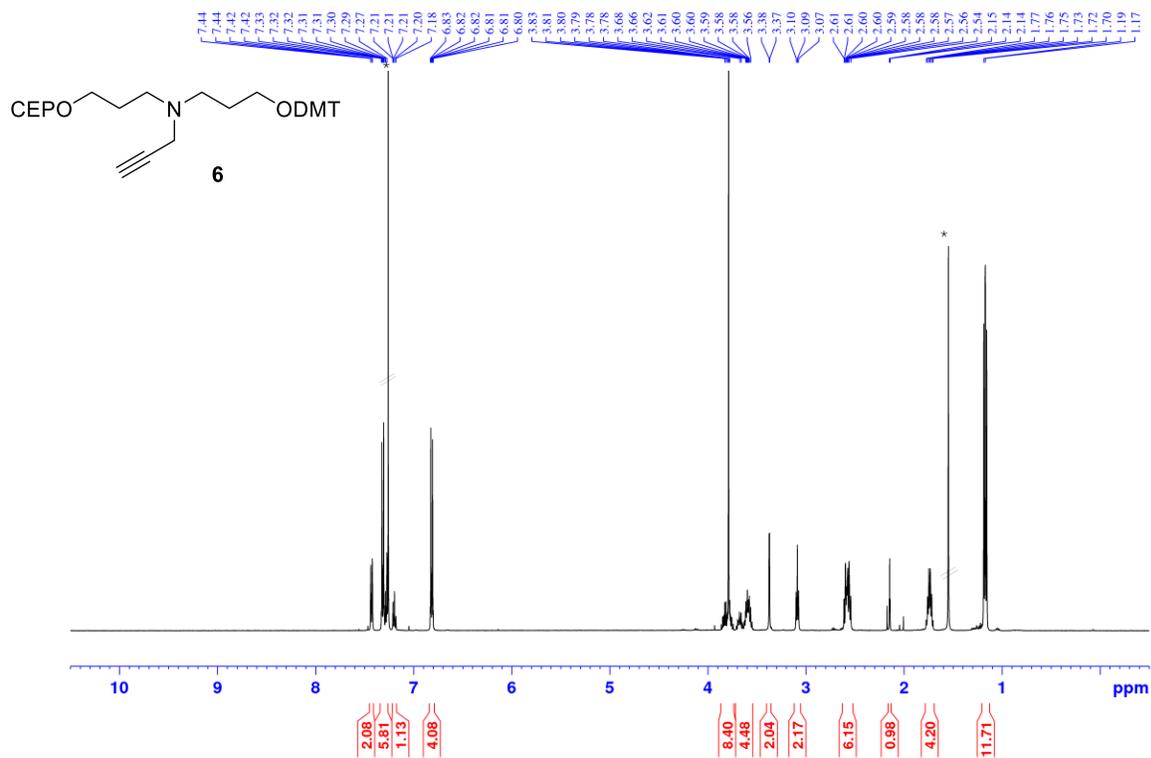


$^{13}\text{C NMR}$ (126 MHz), CDCl_3

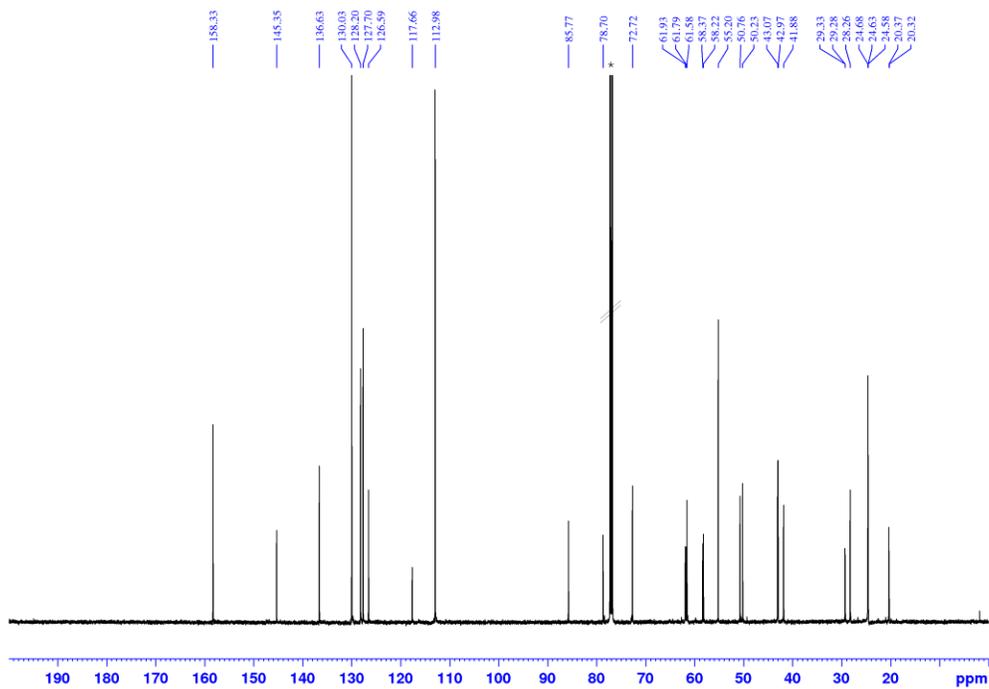


7.3.14 Monomer 6

^1H NMR (500 MHz), CDCl_3

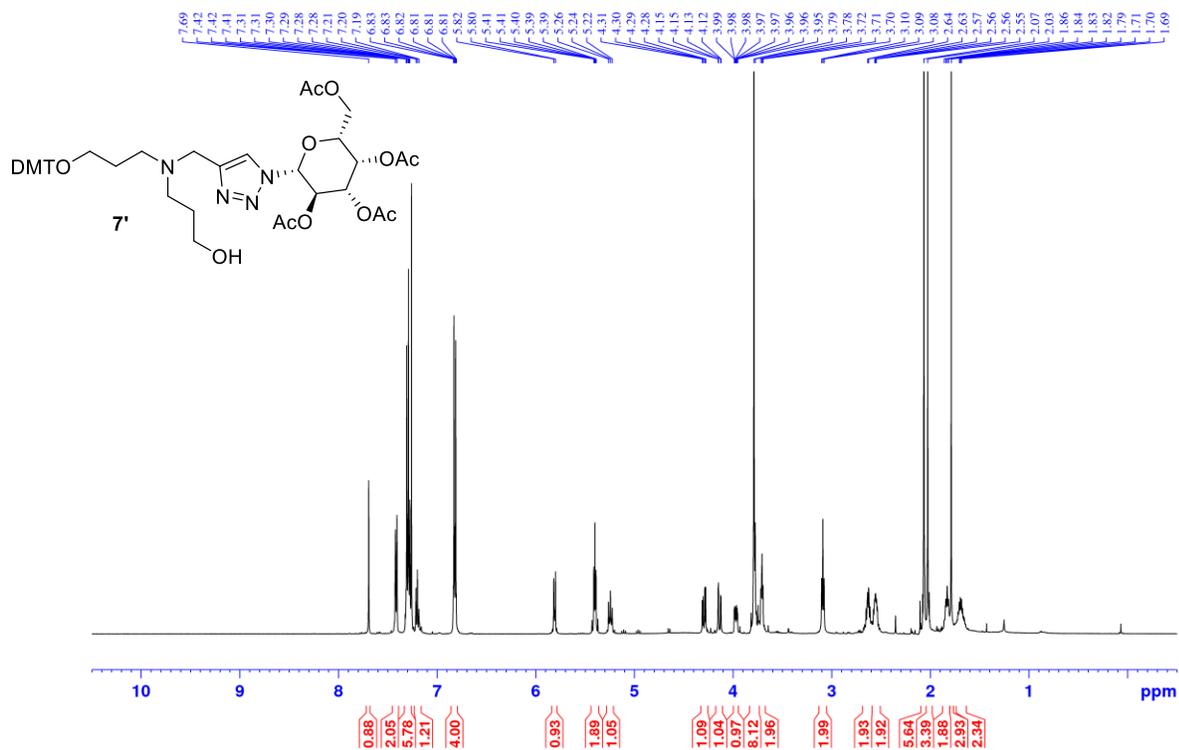


^{13}C NMR (126 MHz), CDCl_3

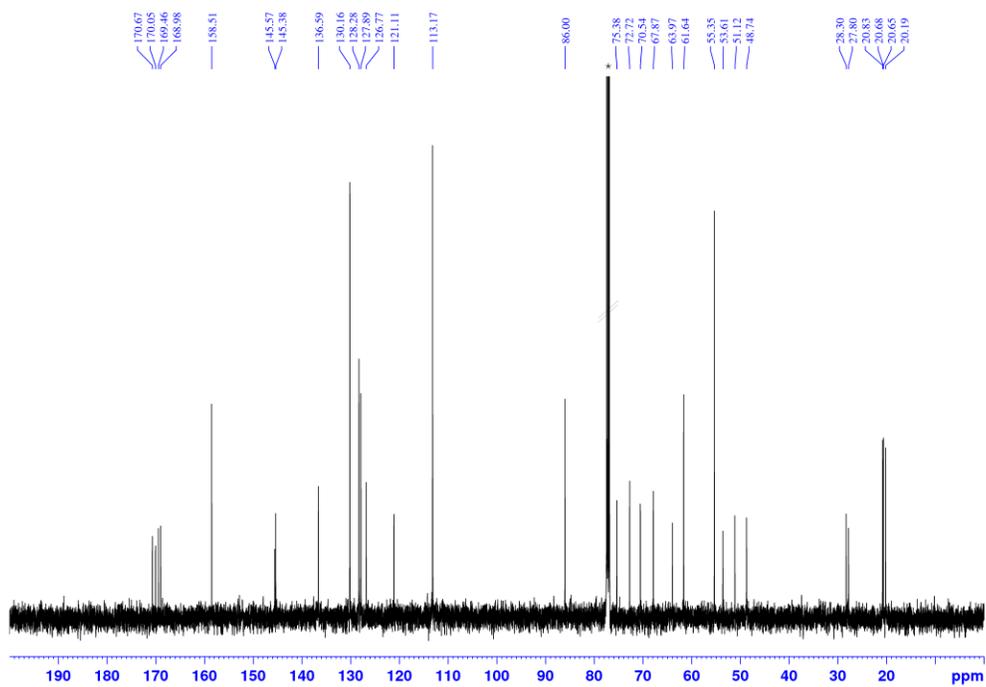


7.3.15 Compound 7'

¹H NMR (500 MHz), CDCl₃

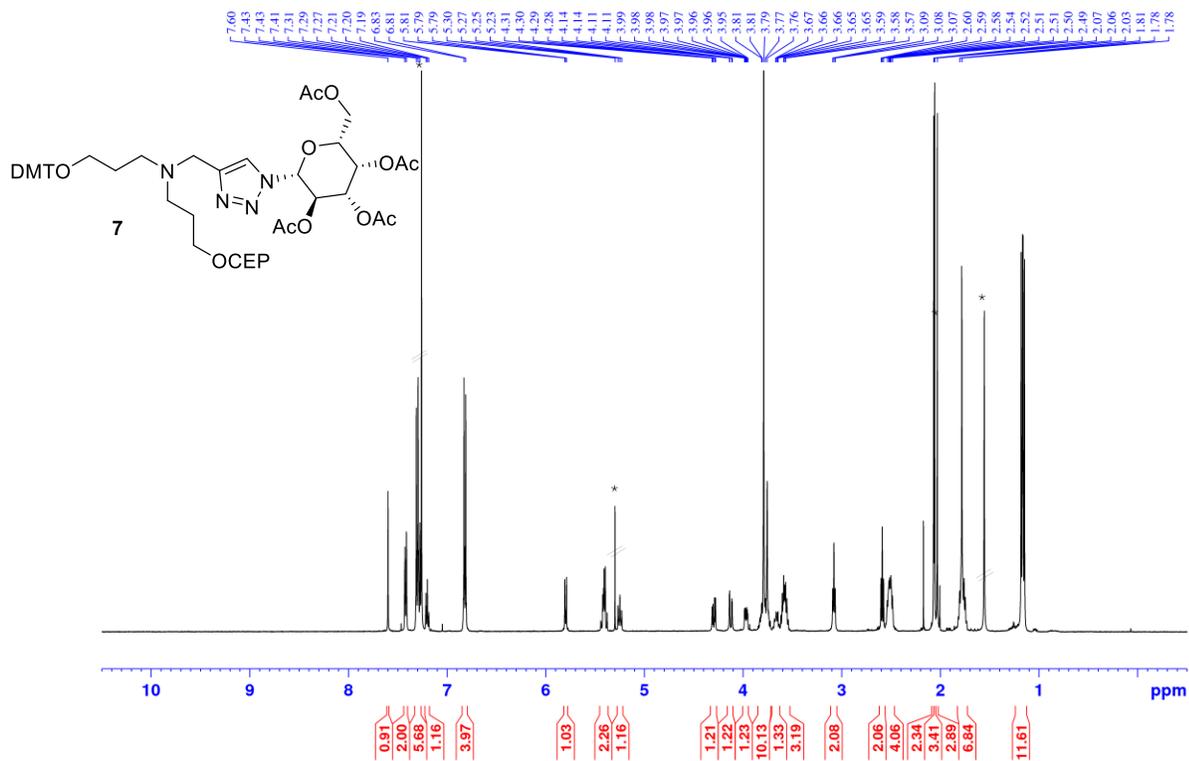


¹³C NMR (126 MHz), CDCl₃

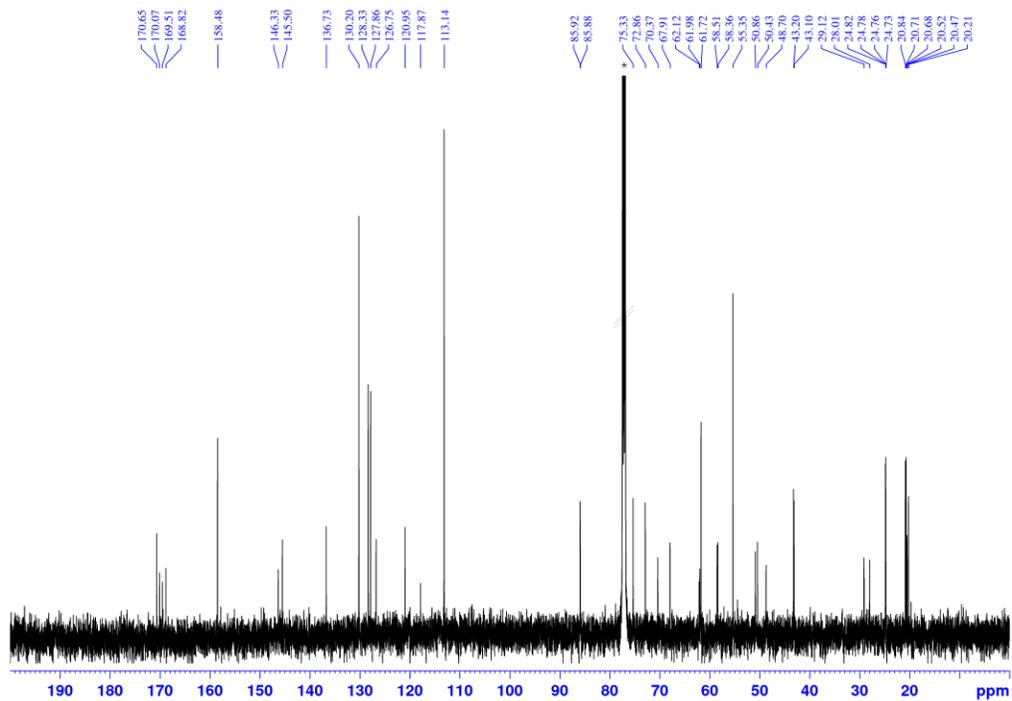


7.3.16 Monomer 7

$^1\text{H NMR}$ (500 MHz), CDCl_3

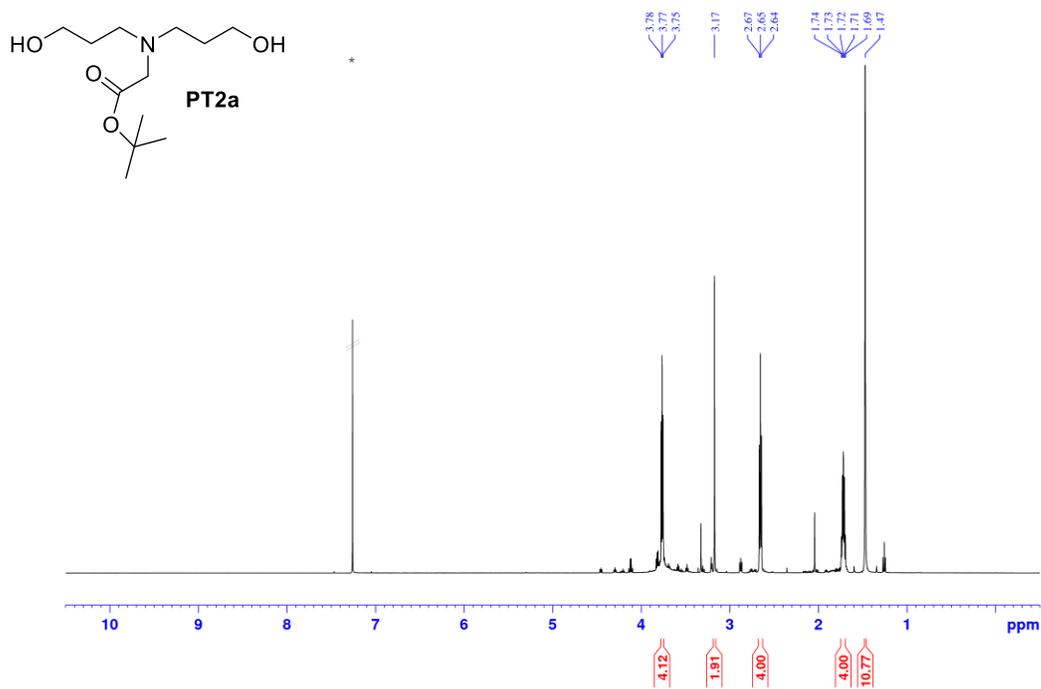


$^{13}\text{C NMR}$ (126 MHz), CDCl_3



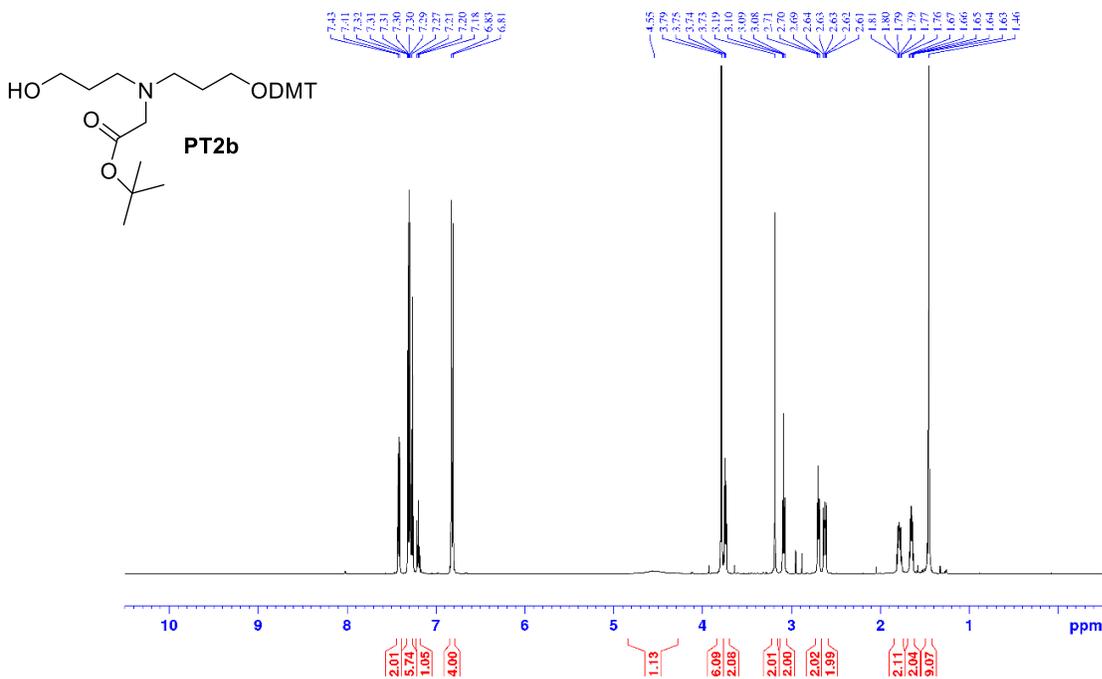
7.3.17 Compound PT2a

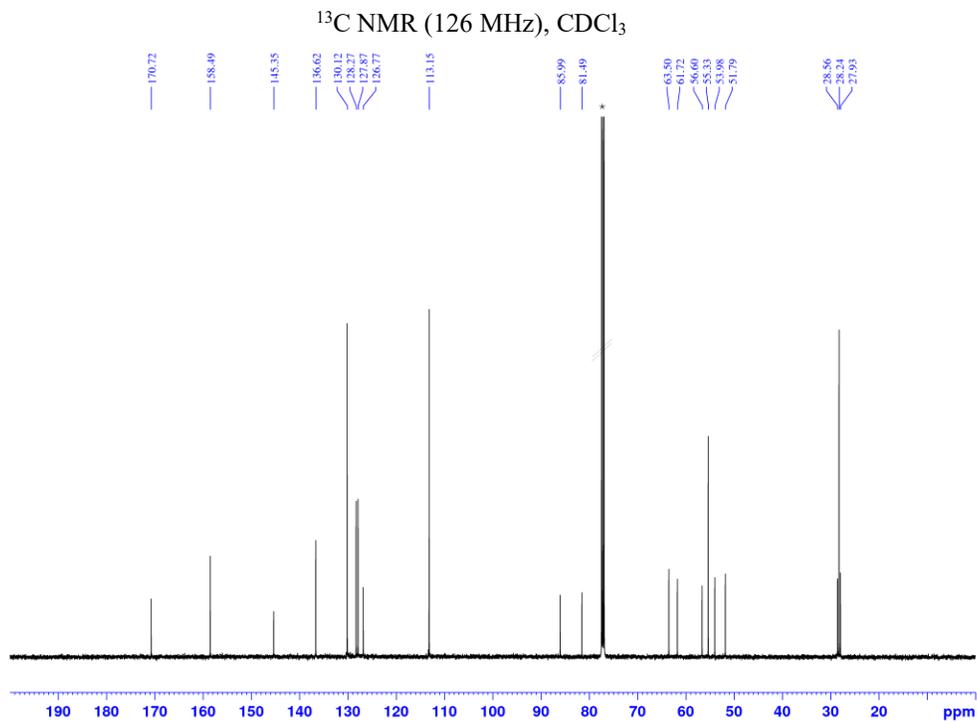
¹H NMR (500 MHz), CDCl₃



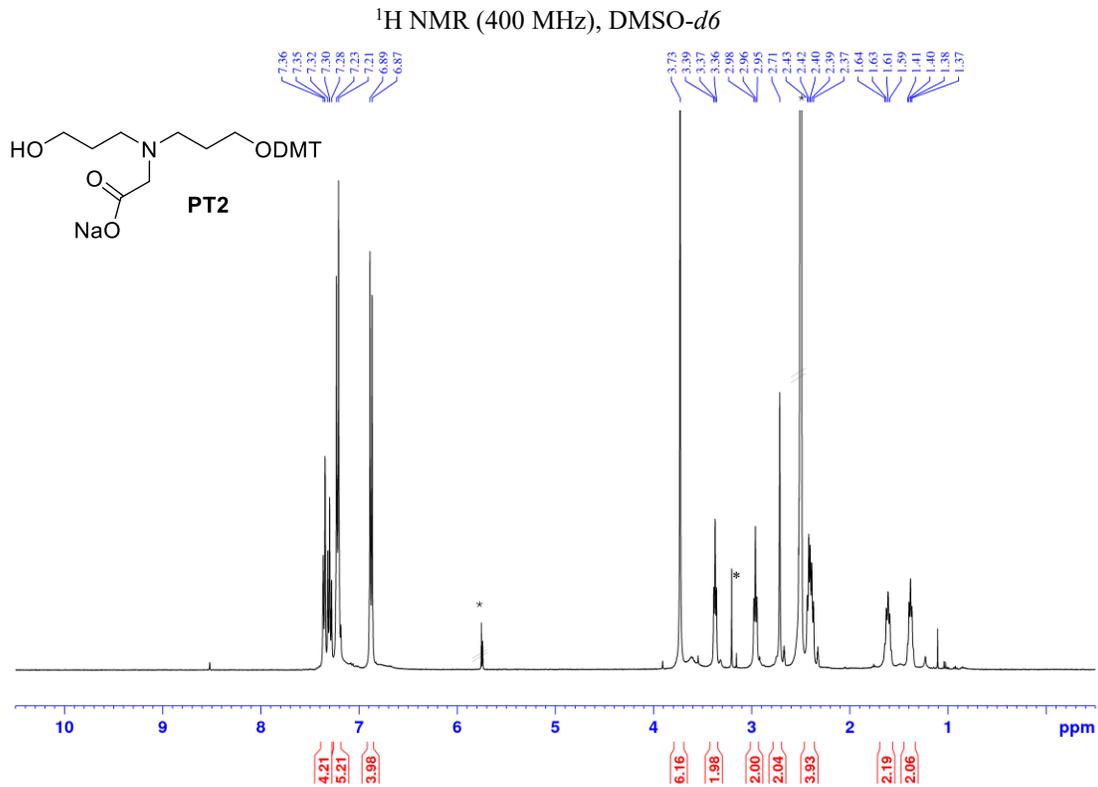
7.3.18 Compound PT2b

¹H NMR (500 MHz), CDCl₃

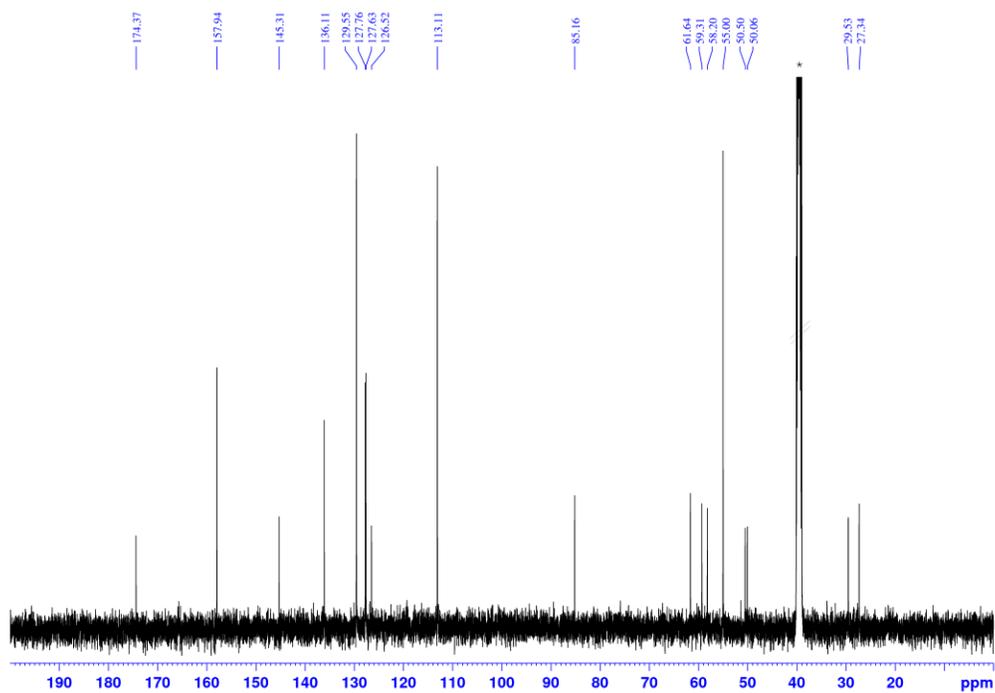




7.3.19 Platform PT2

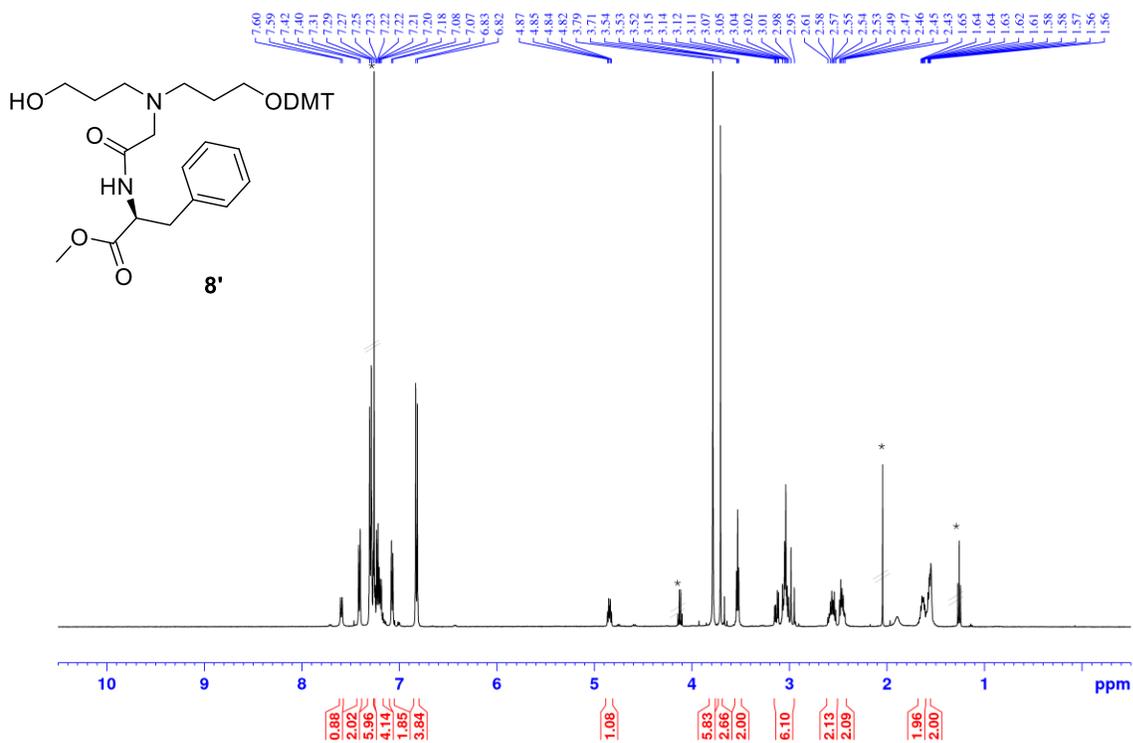


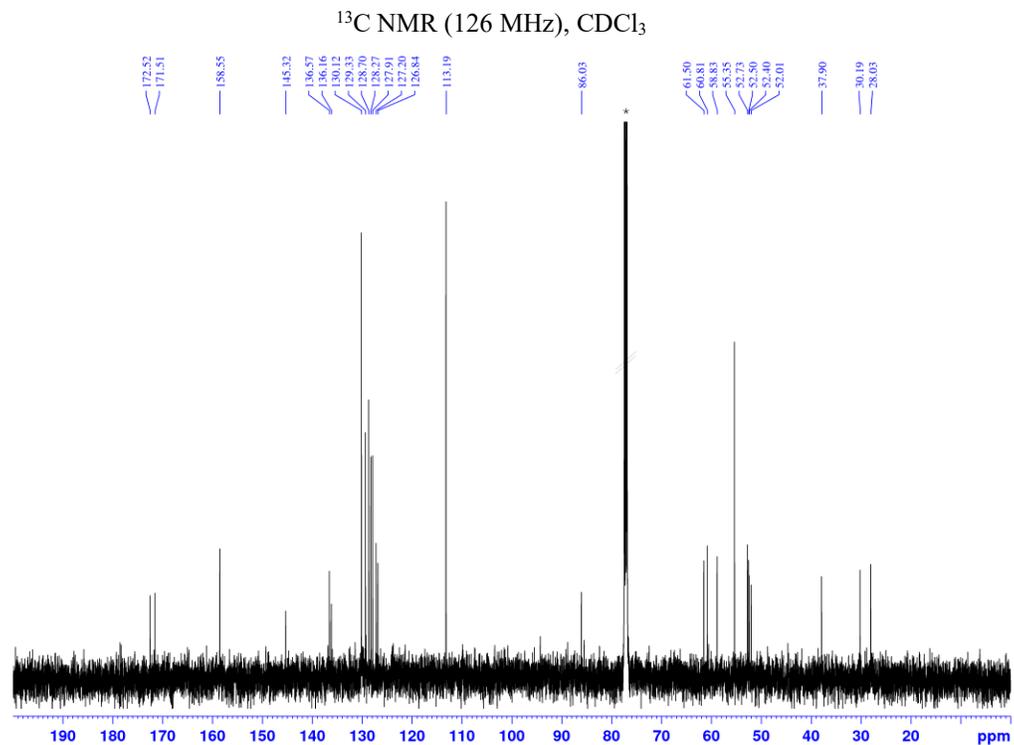
^{13}C NMR (126 MHz), DMSO- d_6



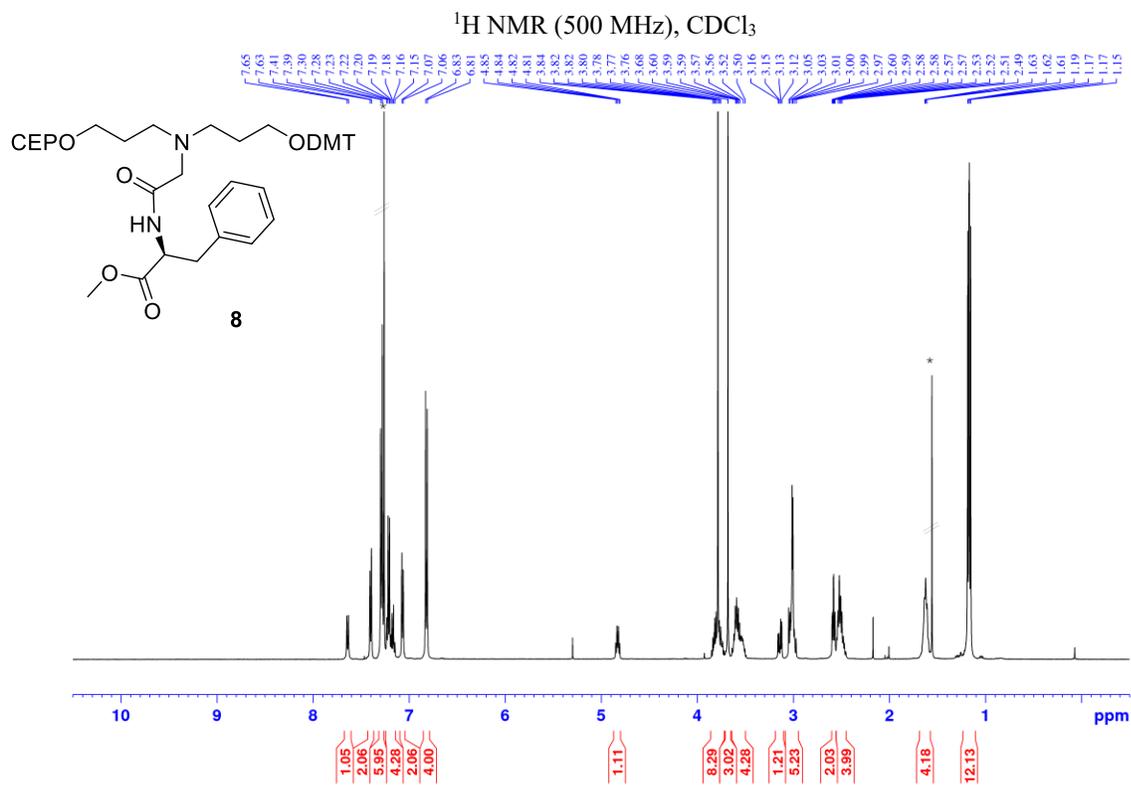
7.3.20 Compound 8'

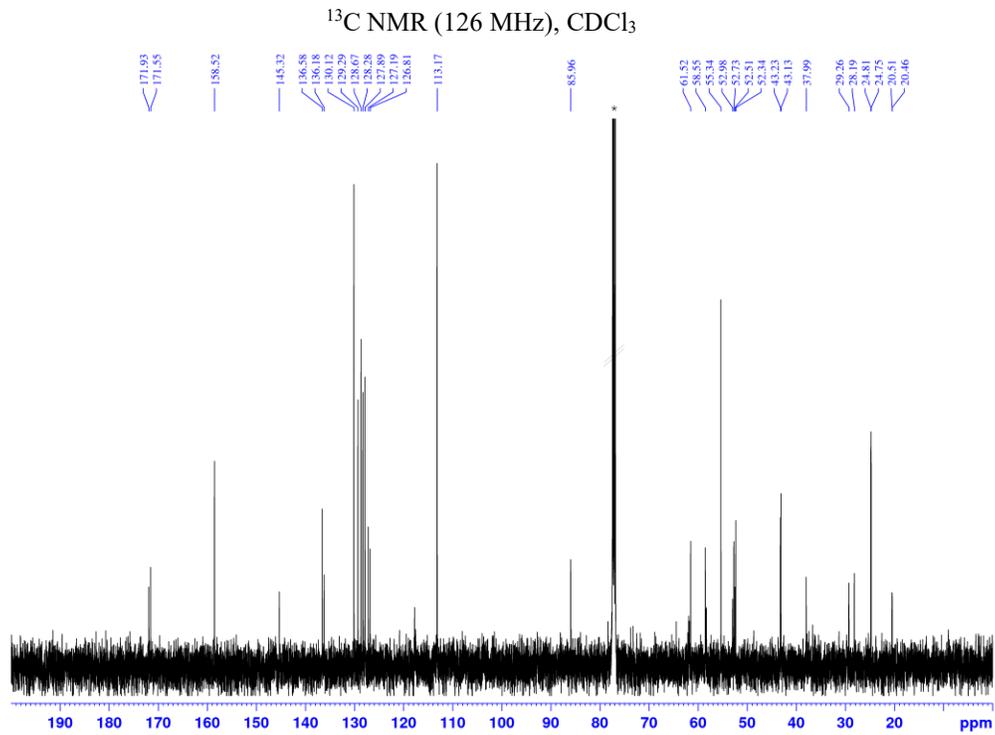
^1H NMR (500 MHz), CDCl_3



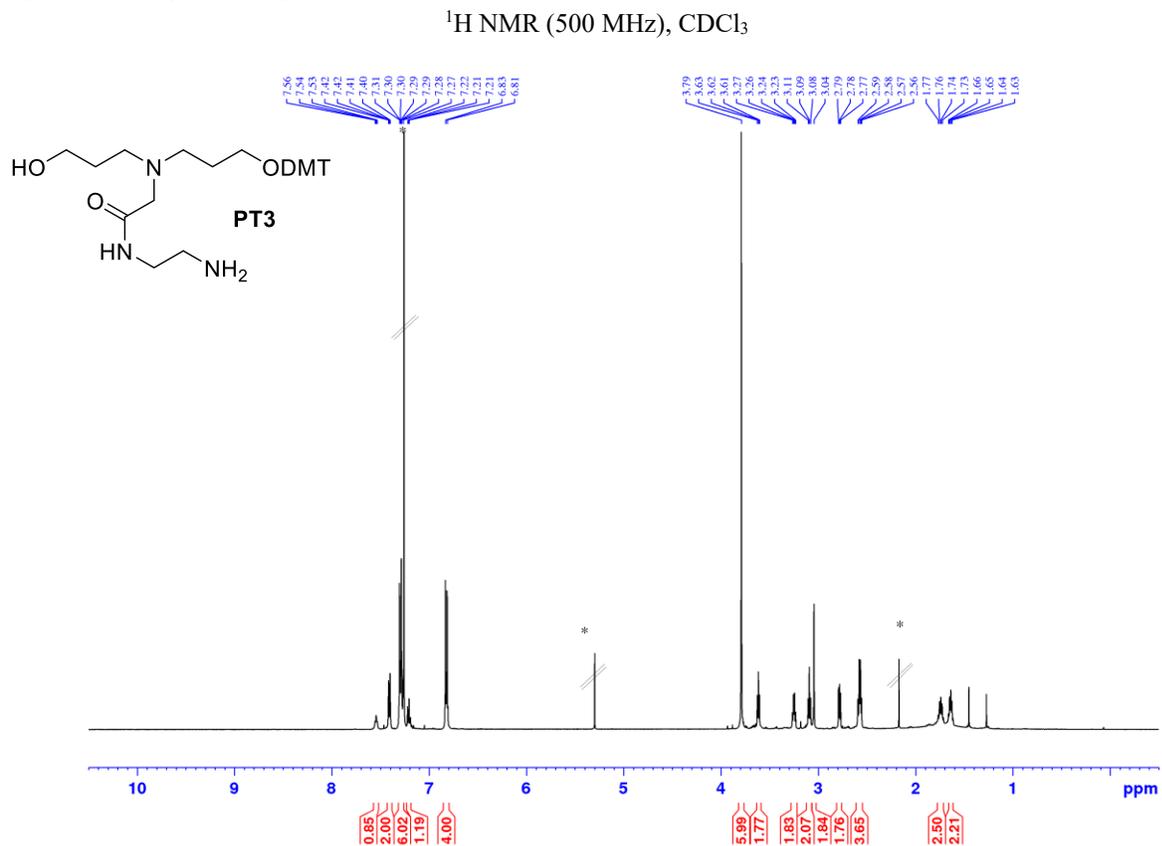


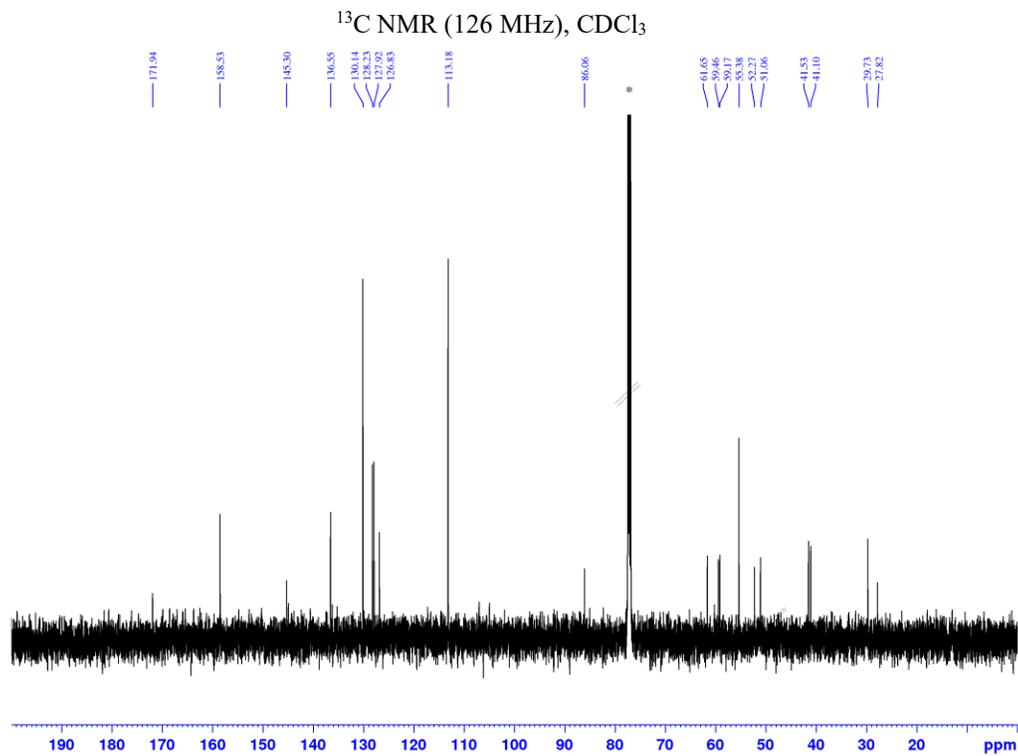
7.3.21 Monomer 8



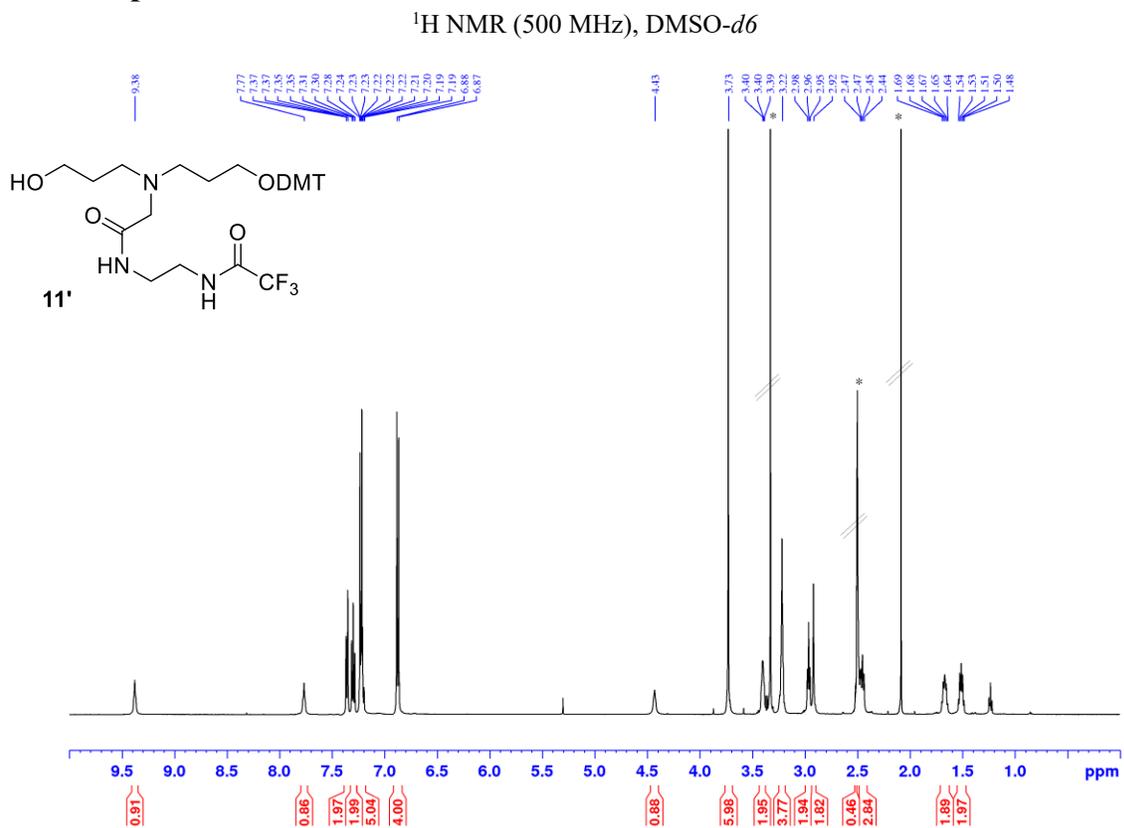


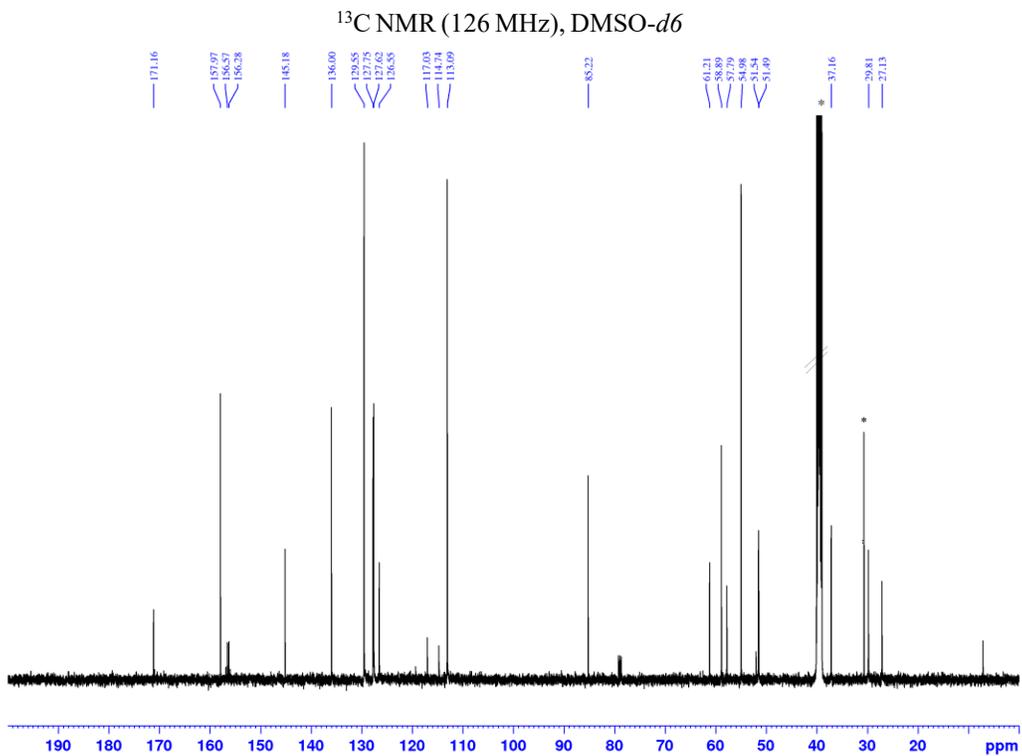
7.3.22 Platform PT3



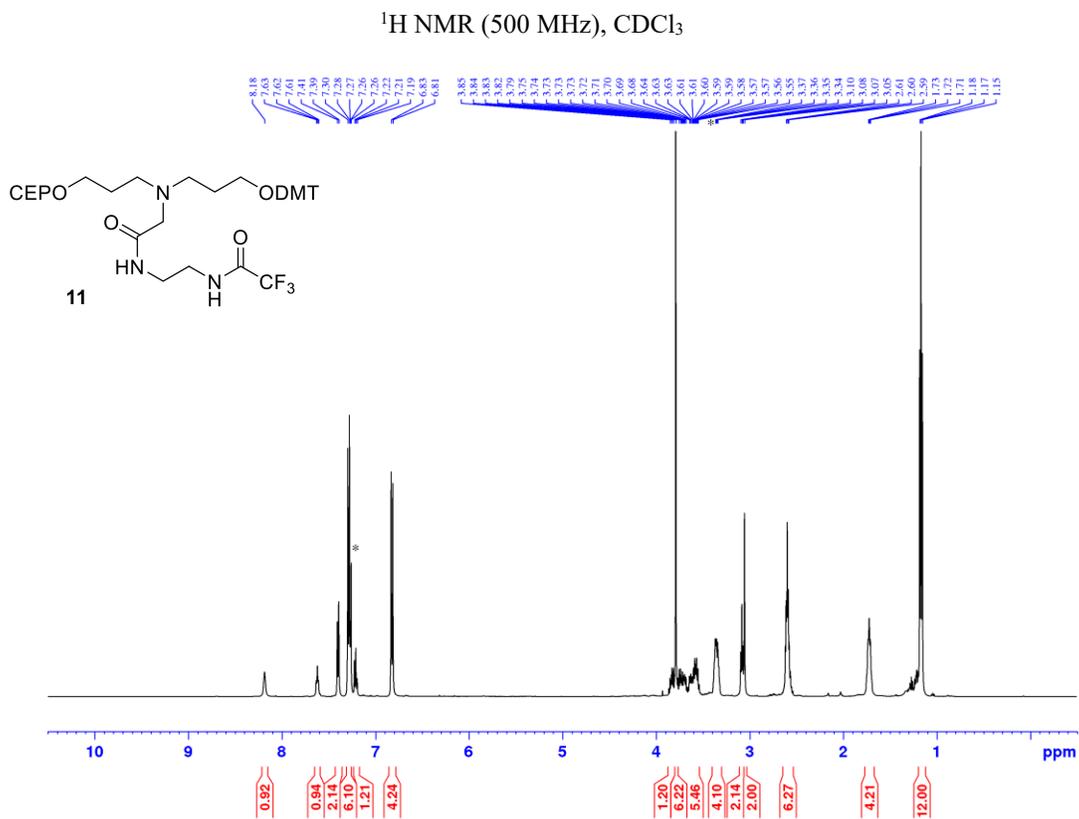


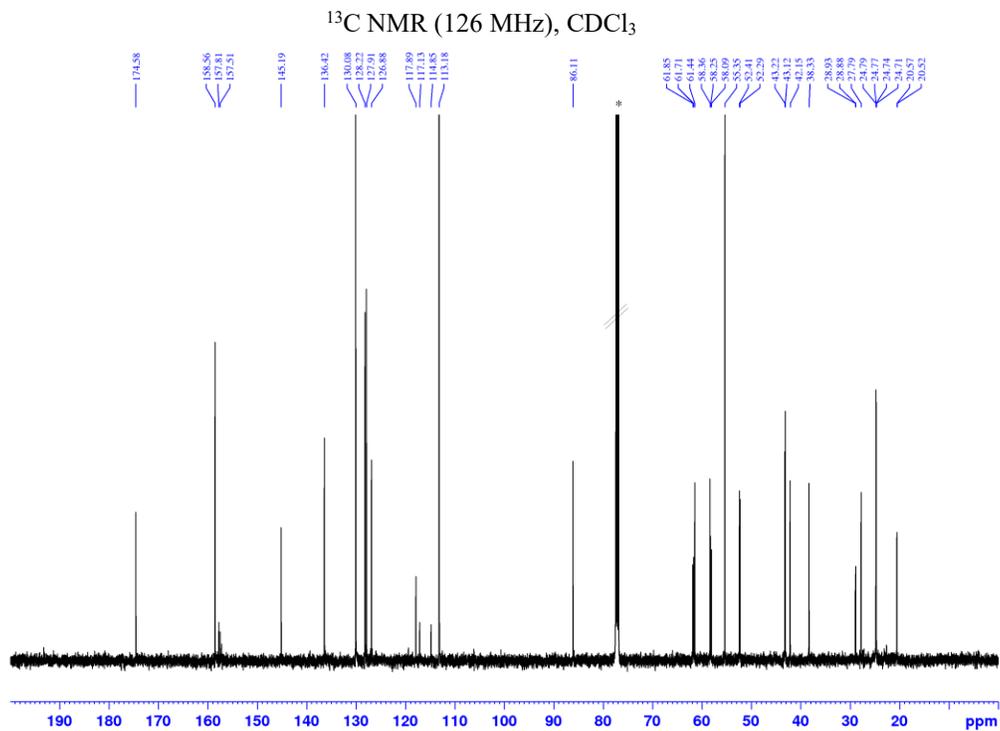
7.3.23 Compound 11'





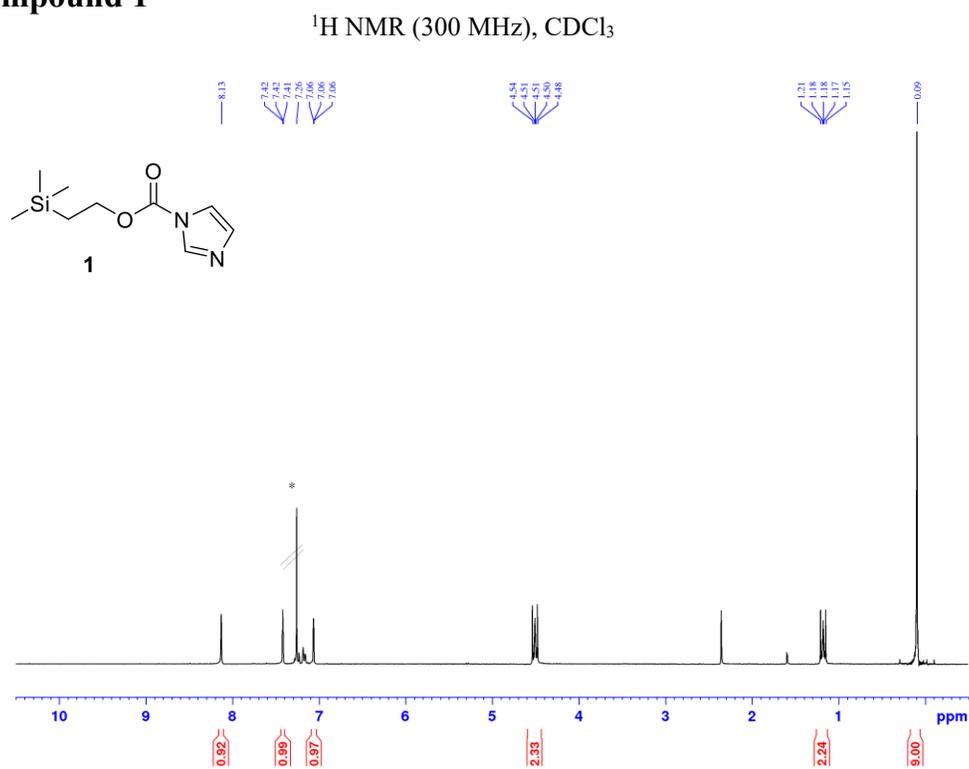
7.3.24 Monomer 11





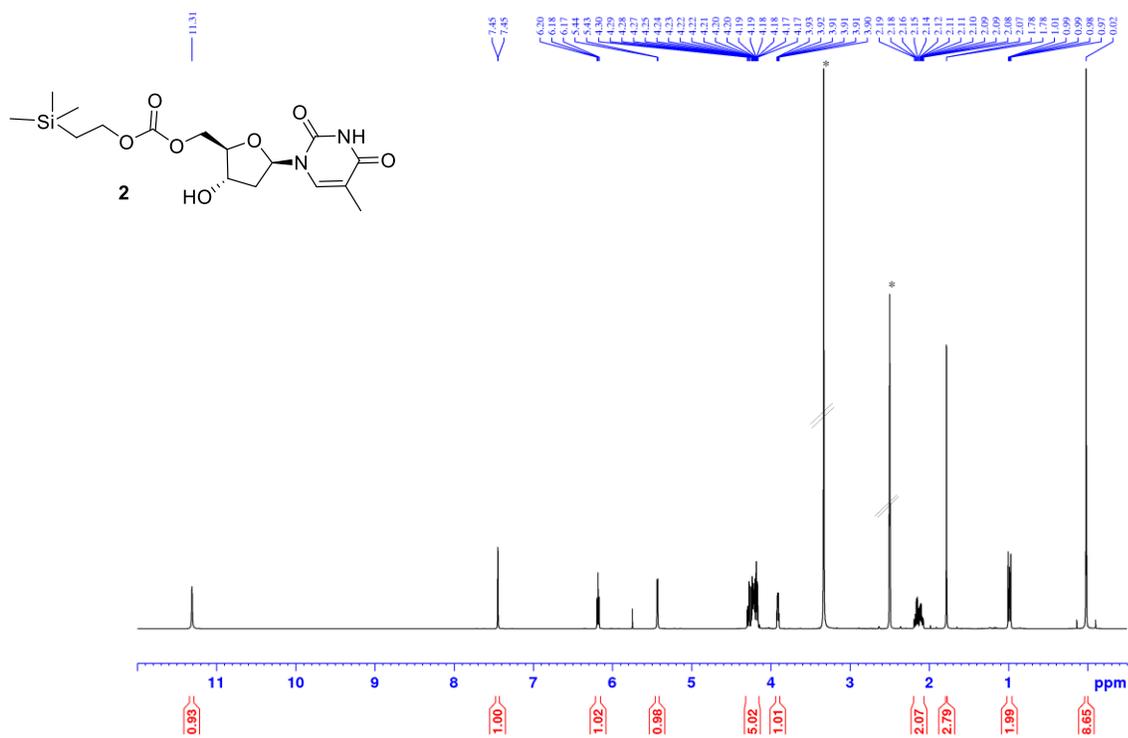
7.4 Molecules from Chapter 5

7.4.1 Compound 1



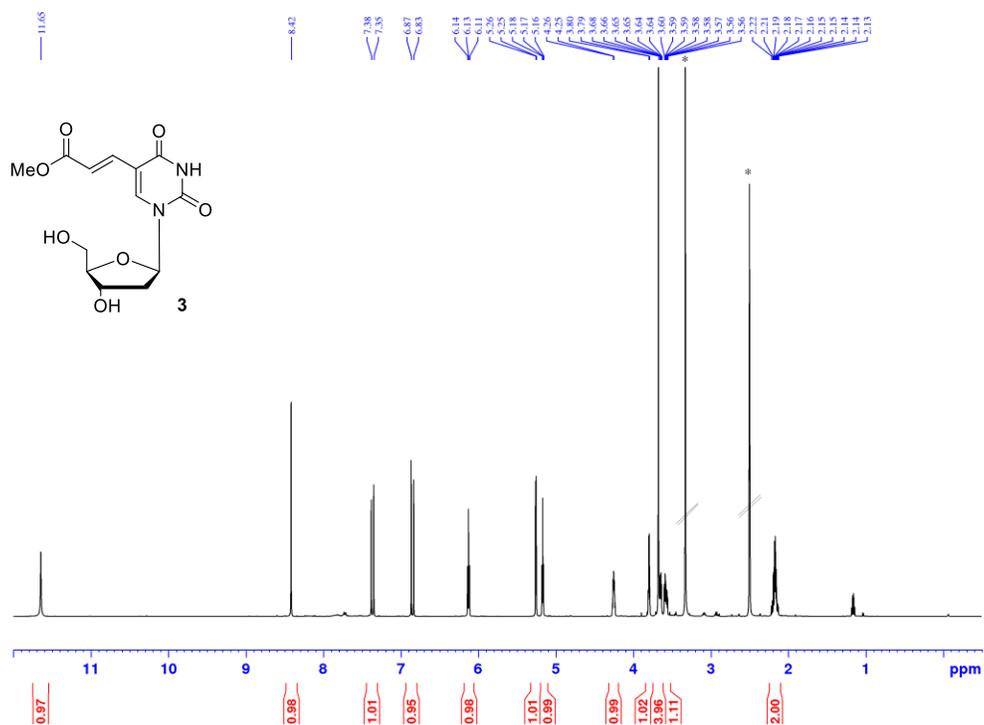
7.4.2 Compound 2

$^1\text{H NMR}$ (500 MHz), $\text{DMSO-}d_6$



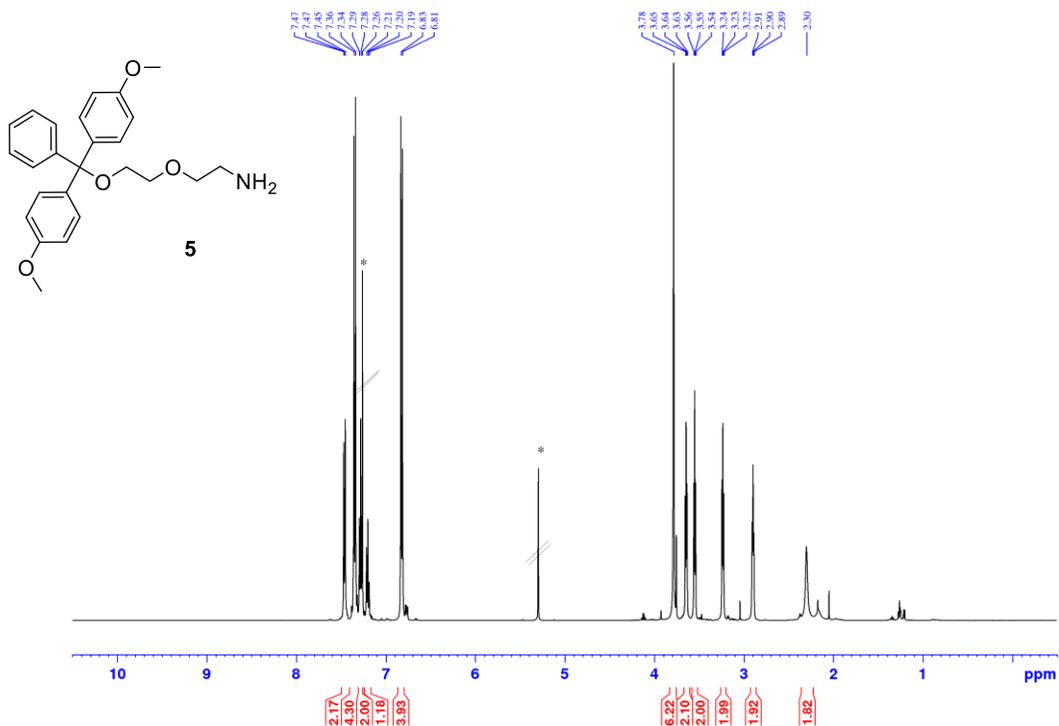
7.4.3 Compound 3

$^1\text{H NMR}$ (500 MHz), $\text{DMSO-}d_6$



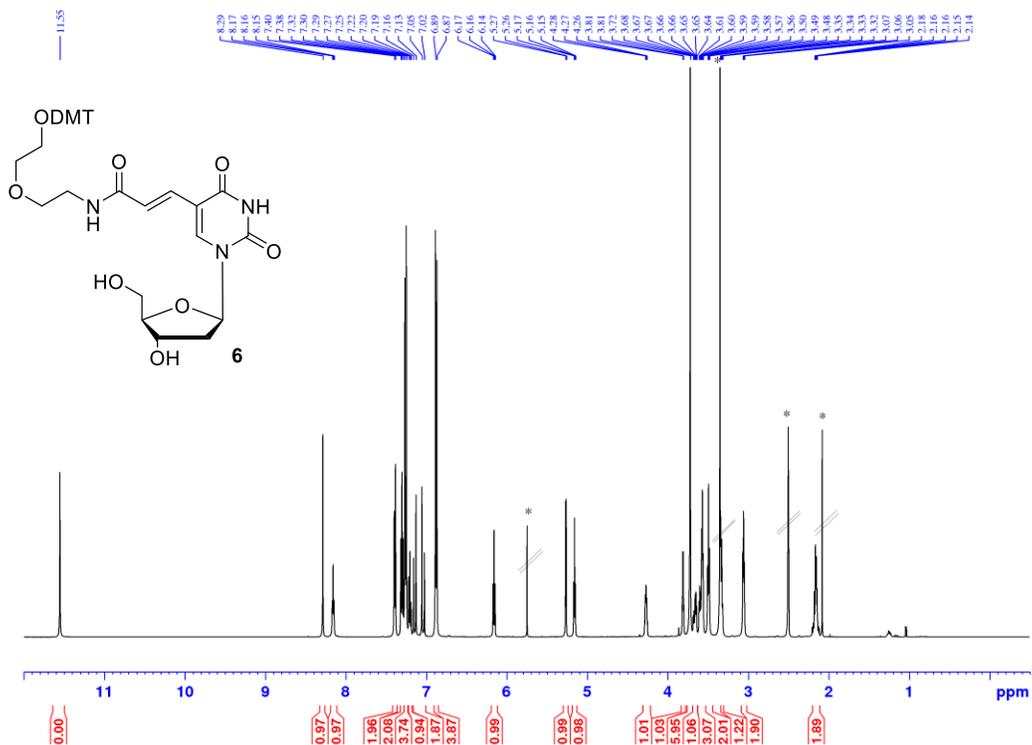
7.4.4 Compound 5

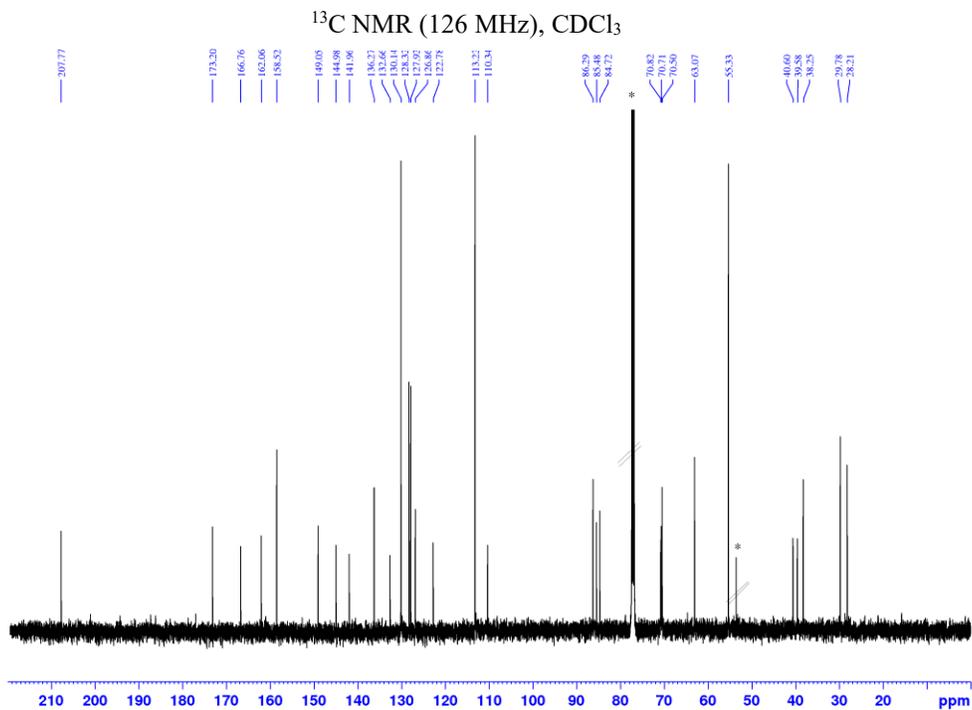
$^1\text{H NMR}$ (500 MHz), CDCl_3



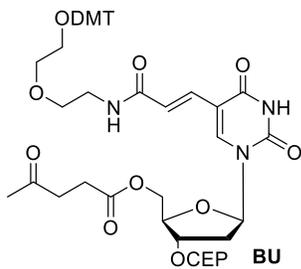
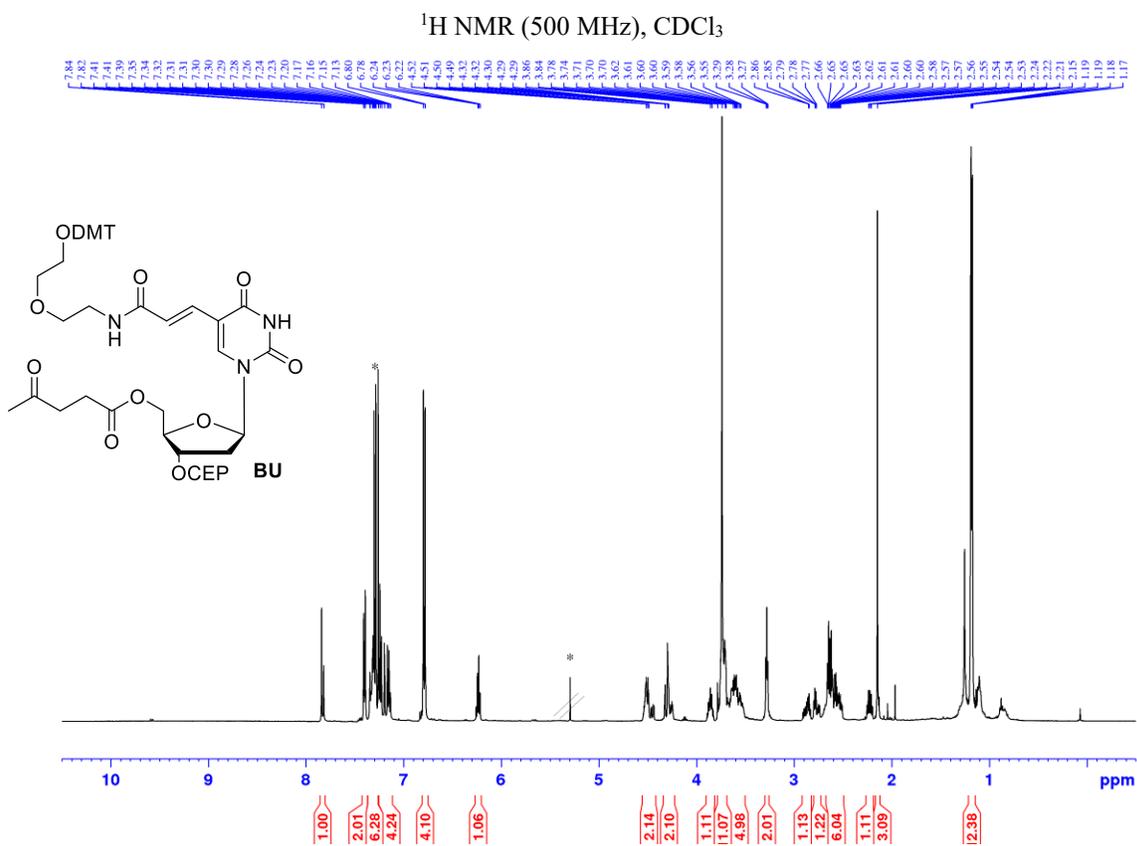
7.4.5 Compound 6

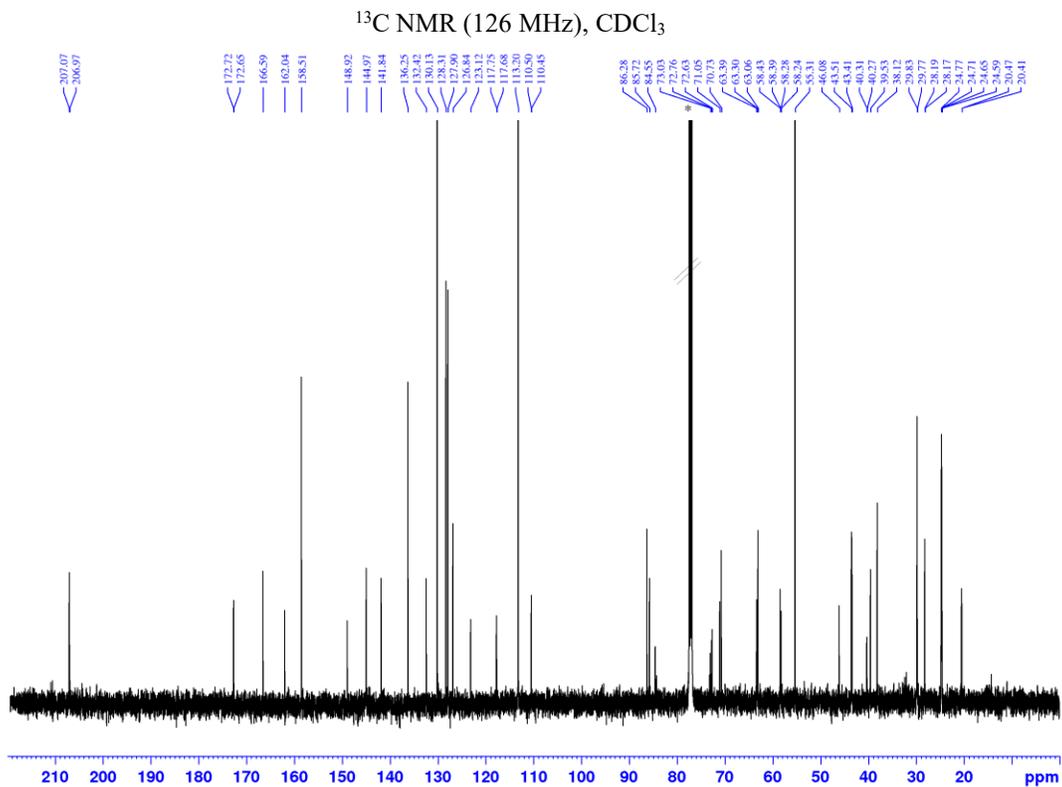
$^1\text{H NMR}$ (500 MHz), $\text{DMSO-}d_6$



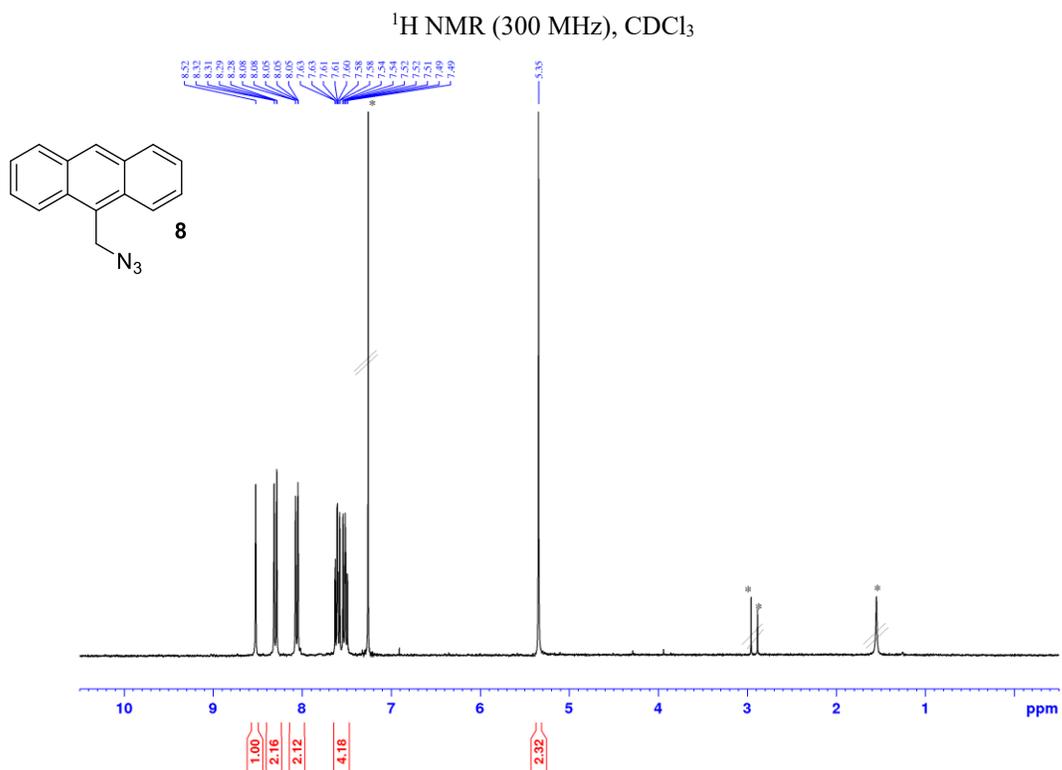


7.4.7 Branching unit BU

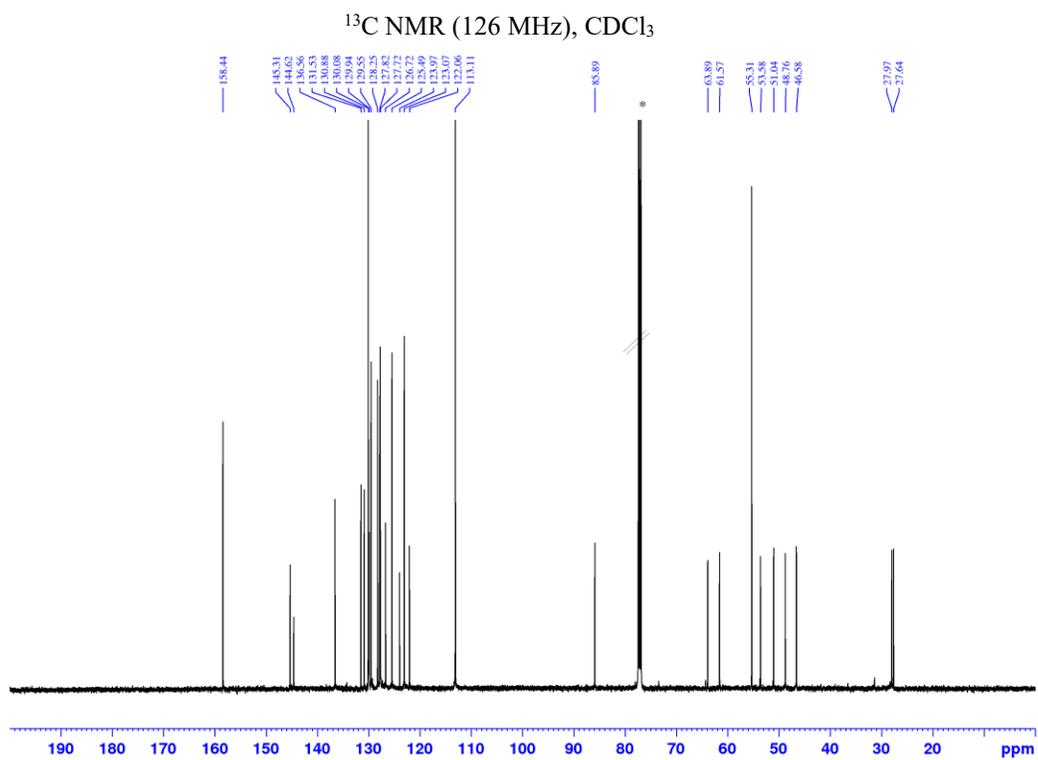
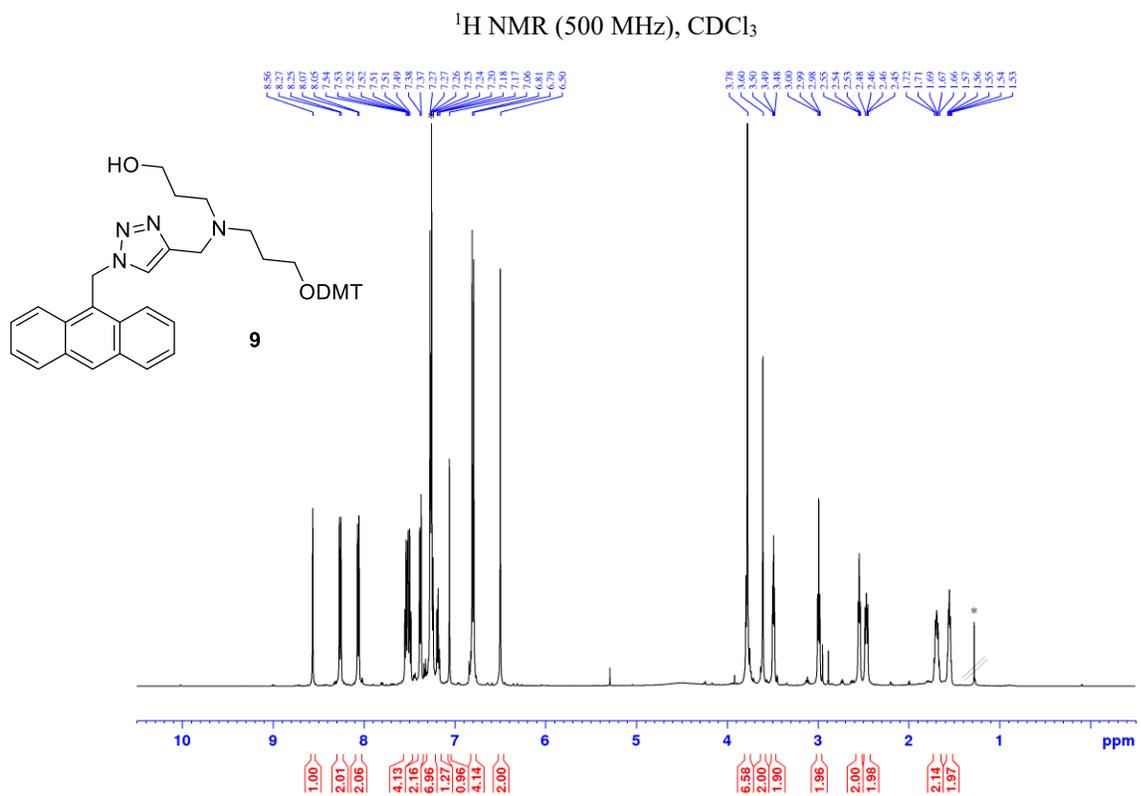




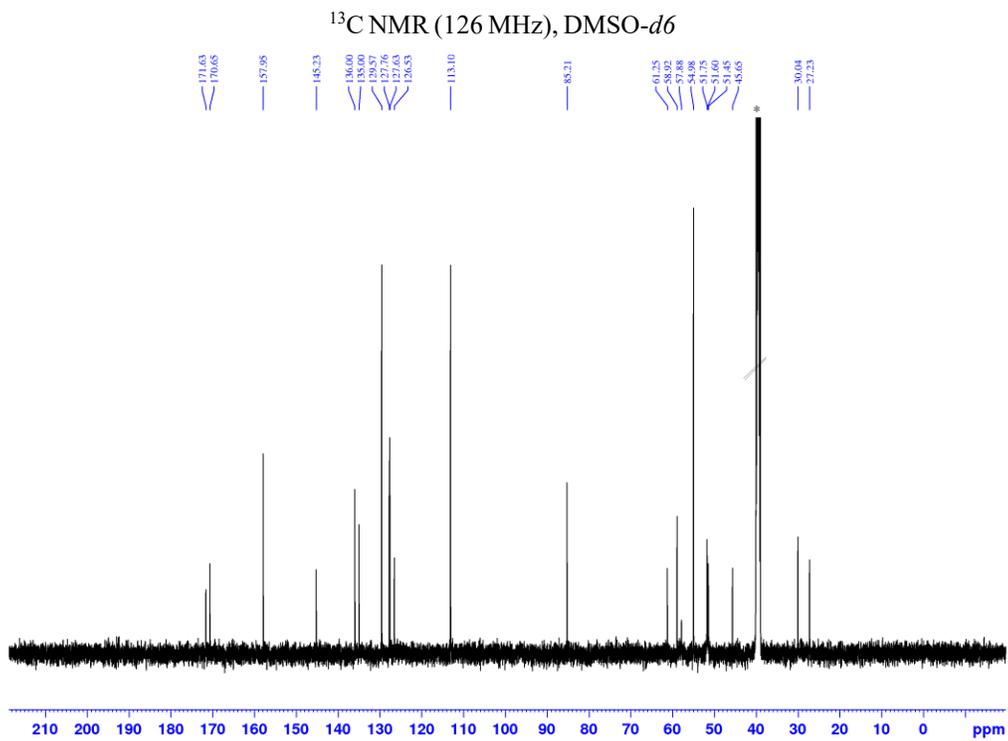
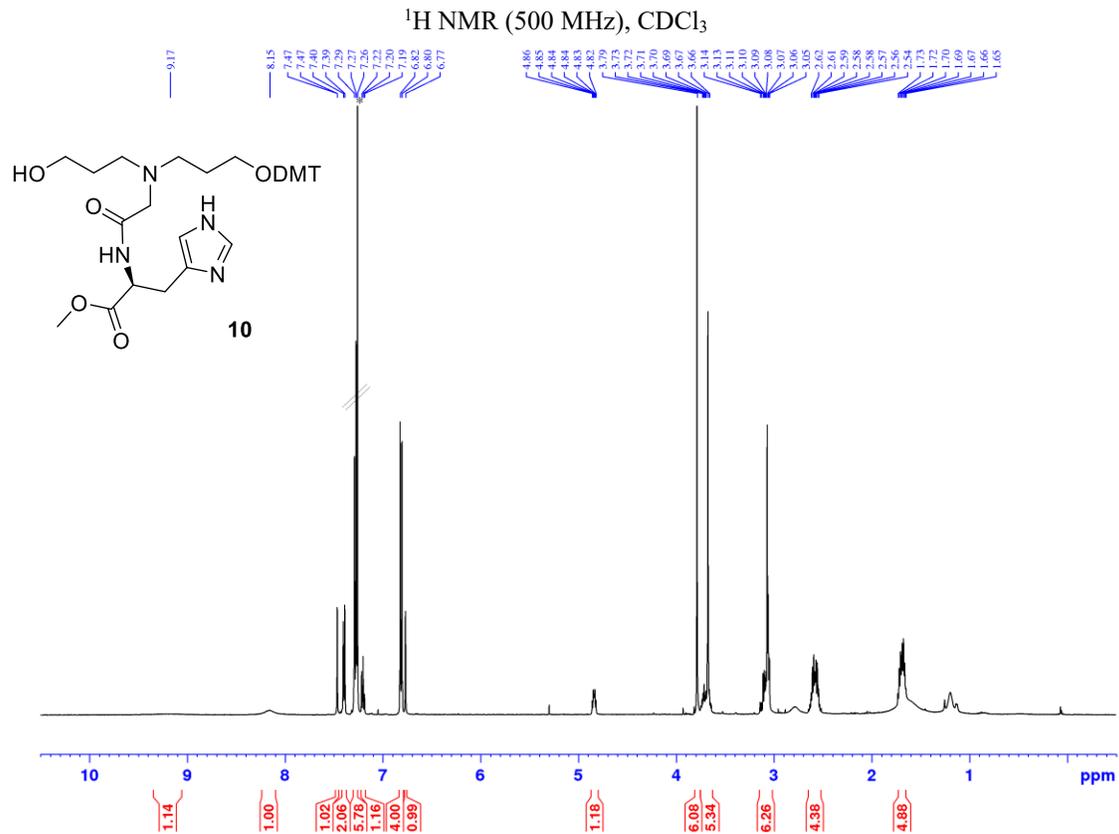
7.4.8 Compound 8



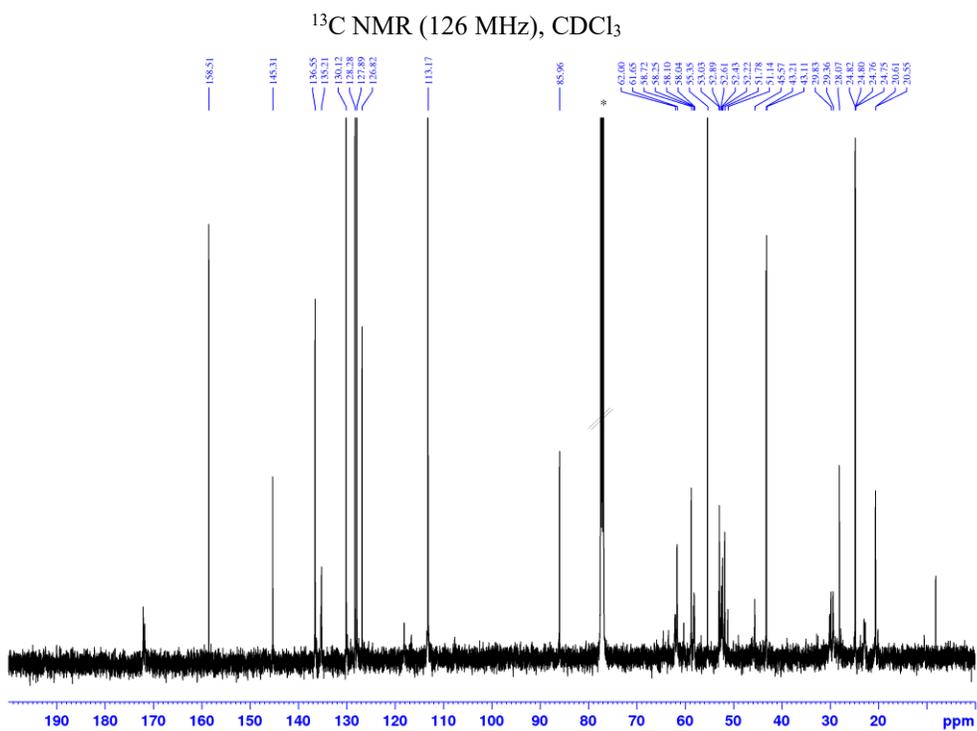
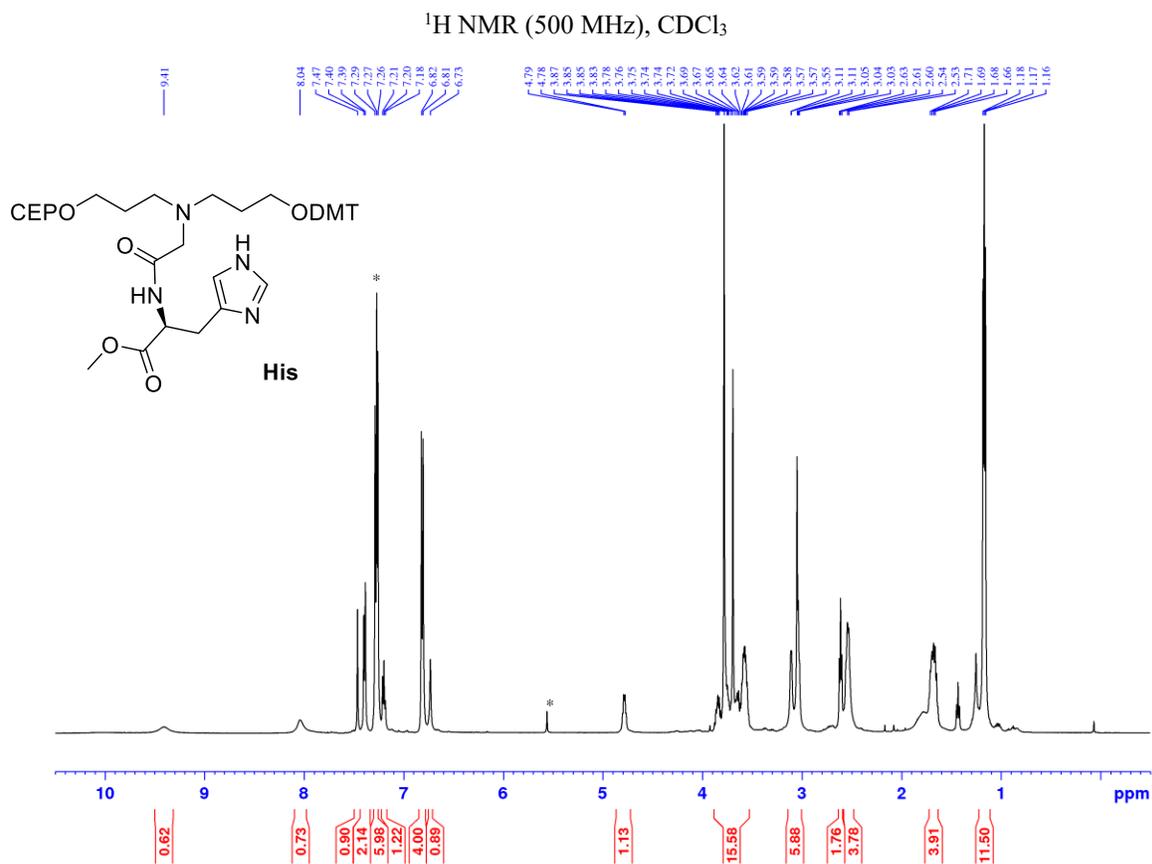
7.4.9 Compound 9



7.4.11 Compound 10

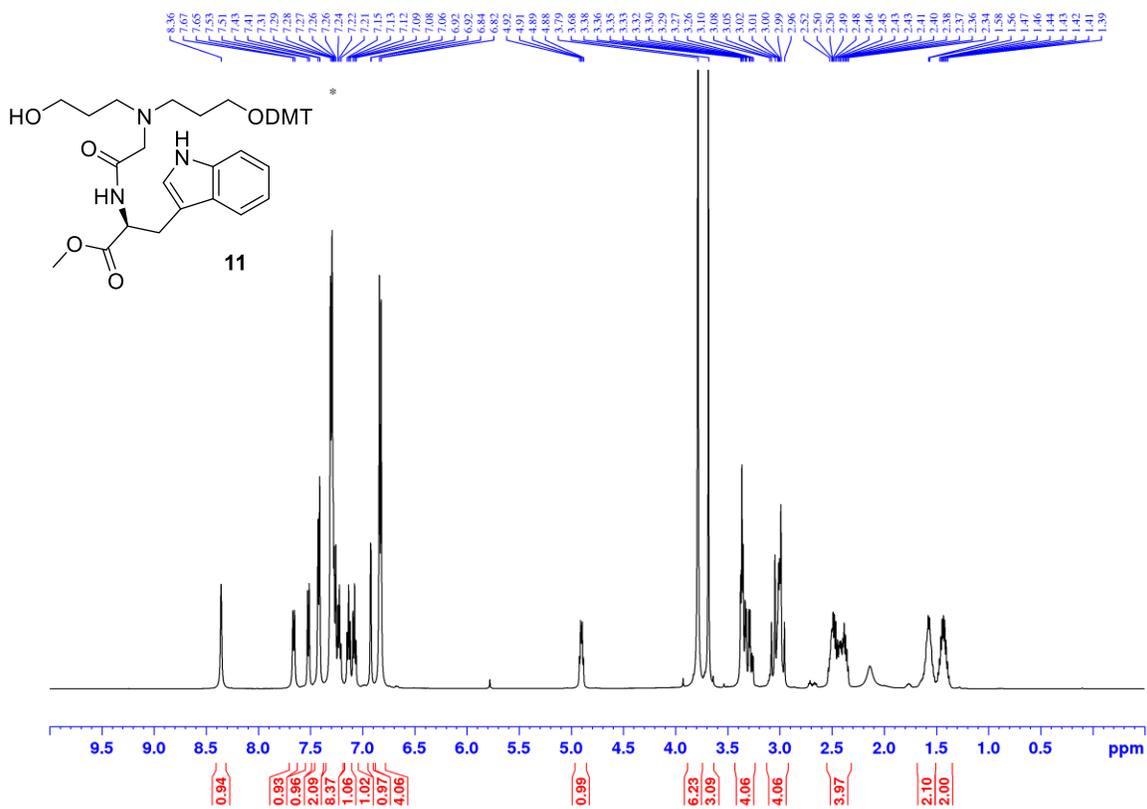


7.4.12 Monomer His

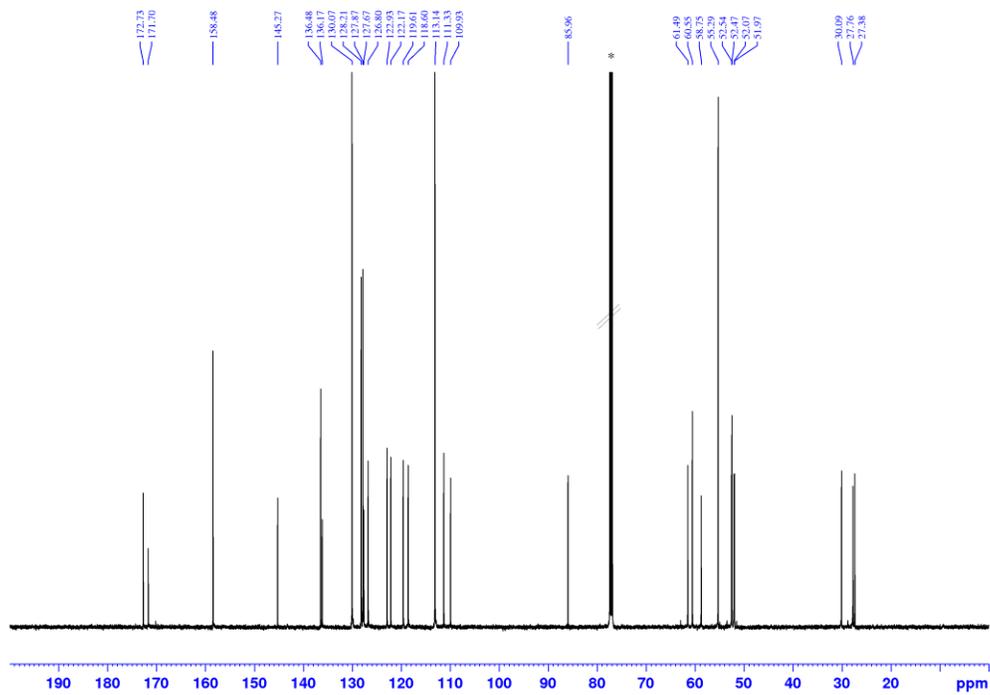


7.4.13 Compound 11

$^1\text{H NMR}$ (500 MHz), CDCl_3

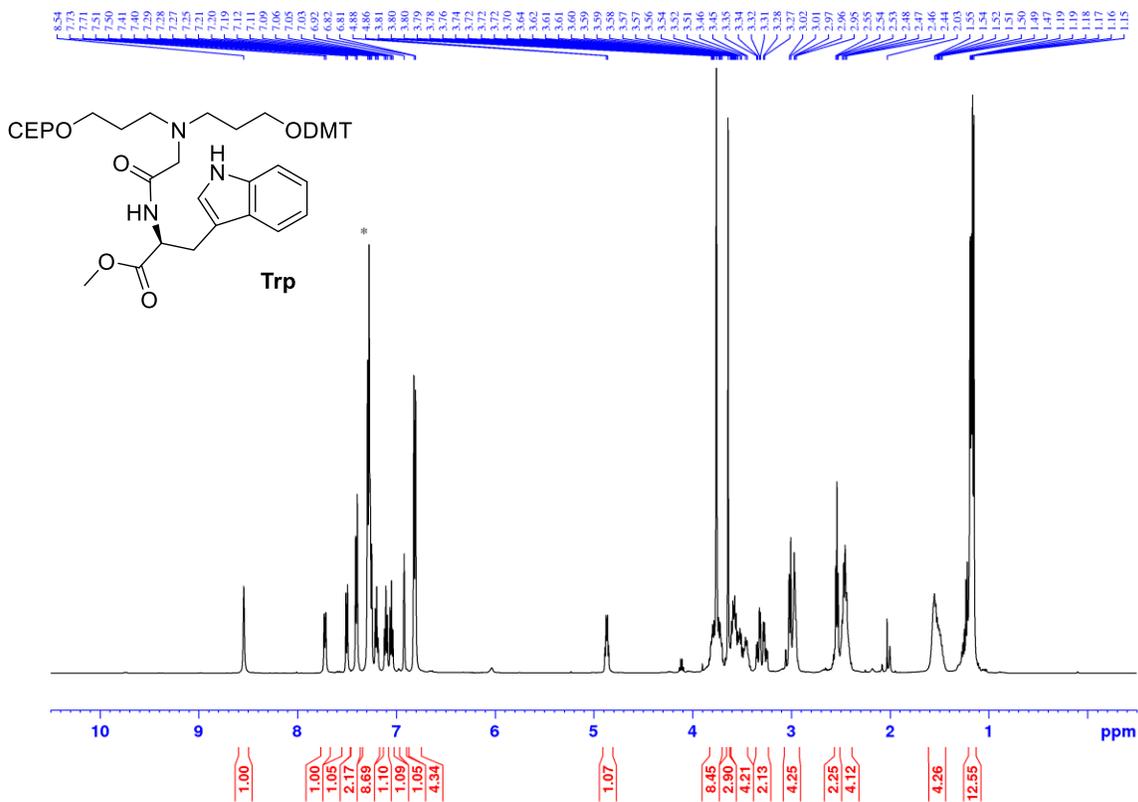


$^{13}\text{C NMR}$ (126 MHz), CDCl_3

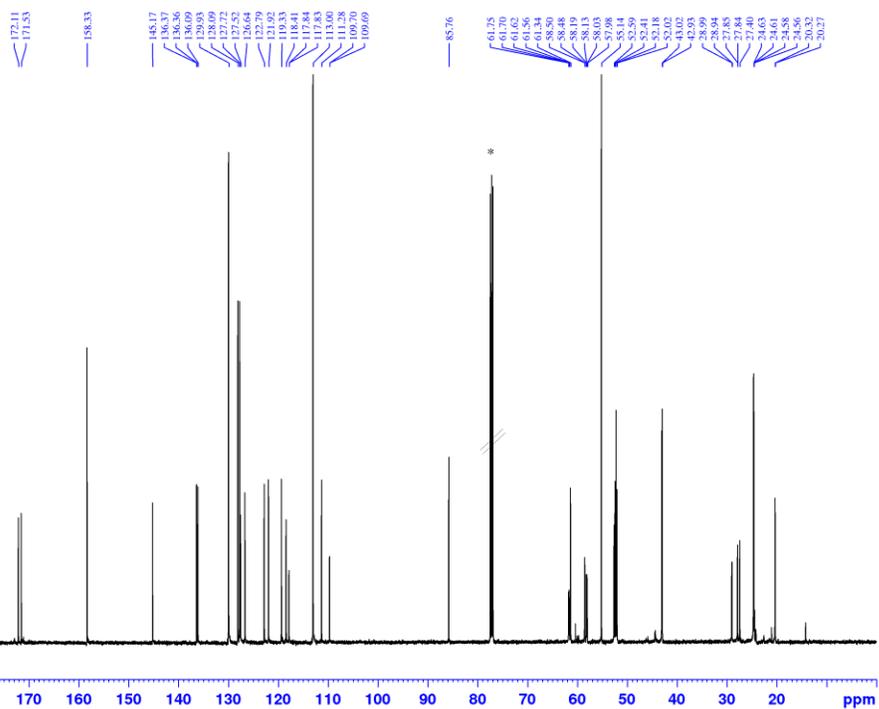


7.4.14 Monomer Trp

$^1\text{H NMR}$ (500 MHz), CDCl_3

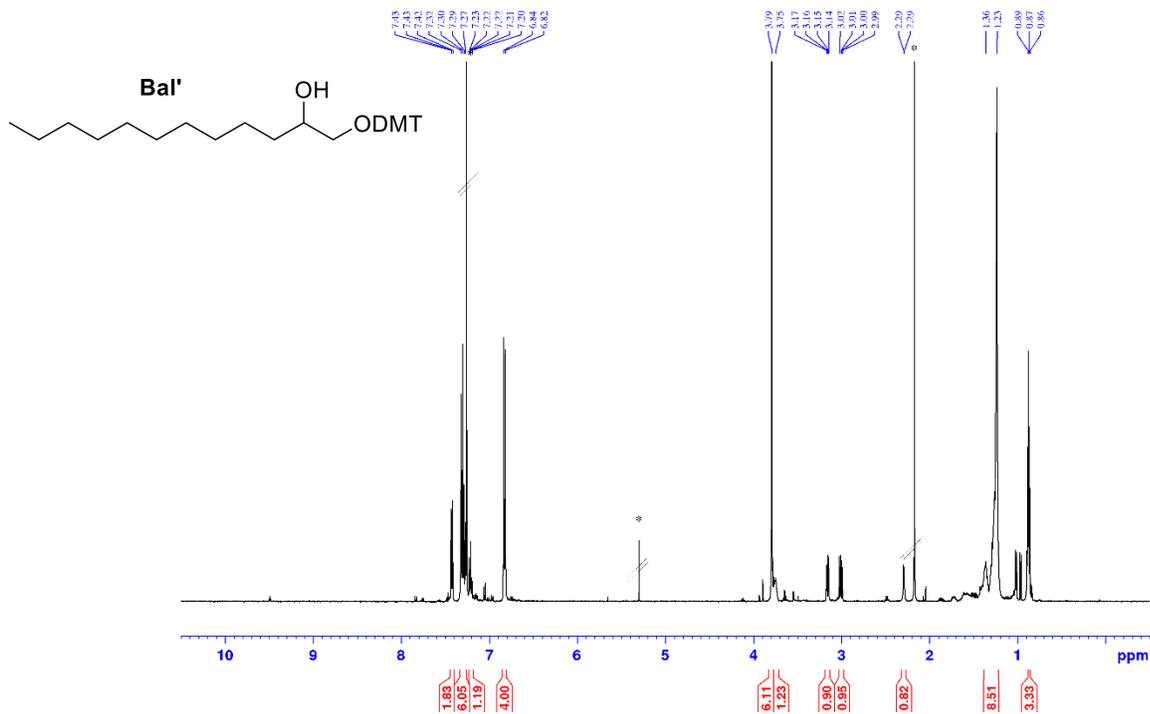


$^{13}\text{C NMR}$ (126 MHz), CDCl_3



7.4.15 Compound Bal'

¹H NMR (500 MHz), CDCl₃



7.4.16 Monomer Bal

¹H NMR (400 MHz), CDCl₃

