Peptide-DNA interactions: a study using protein models and the furan-oxidation cross-link methodology

Lieselot Carrette

Promotor:
Prof. Dr. A. Madder

2014

ii
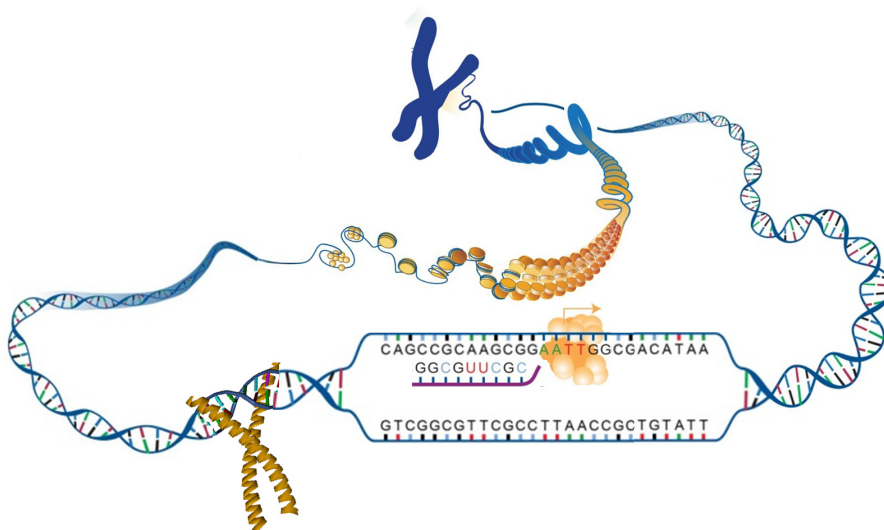
Preface

'Less then 200 years ago matter created by living organisms was regarded to be distinct from non-living matter, like salts, minerals and gases. It was with the discovery by Friedrich Wöhler in 1828 that a biological molecule could be synthesized from inorganic compounds that the inorganic and organic worlds joined. Wöhler tried to form ammonium cyanate by combining the inorganic salts silver cyanate and ammonium chloride, but found that urea was formed instead. Urea was known to be produced by animals and excreted in large amounts in the urine. In a strike it became evident the chemical reactions of organic matter must obey the same principles as the reactions of the inorganic compounds, which chemists at that time had been studying extensively. If this was true it should be possible to study the biological world in a similar way by applying a purely chemical point of view ...'

'Introduction to Pharmaceutical Bioinformatics'
by Jarl E.S. Wikberg

iv

. . . but not without

everyone that has been there in the past 5 years, who I've met due to or despite the PhD and who have contributed to the total experience and thus the person I am now.

Dat is in de eerste plaats en vooral mijn promotor Prof. Annemieke Madder, die mij een plaats in haar groep gegeven heeft (of beter 2 plaatsen, een laboruimte en later ook een heringerichte bureau toen ik allergisch werd aan de job) en steeds een aanvoer van chemicaliën en toestellen (zoals 'mijn' microgolf) geregeld heeft, voor de succesvolle uitvoering van een organisch biomimetisch doctoraat. Uw karakteriserende positiviteit was ook essentieel als motivatie toen niets leek te lukken, en ook erna toen het keerde, om positieve resultaten als zodanig te aanvaarden. Ik ben vooral dankbaar voor de vrijheid die ik kon genieten in mijn werk in Gent en ver daarbuiten, via de mogelijkheden die ik kreeg om de rest van de wereld te verkennen.

Here I would also like to thank the other members of the group for all the help, the support, the company, for challenging me and/or for continuing the work. I am especially gratefull to the DNA furan cross-linking ancestors Sami, Trinidad, Kristof and Marieke for their skillful contributions and establishment of the DNA interstrand furan cross-linking methodology as a great basis to start from. Ellen, en meer recent ook Nathalie voor het verderzetten en de uitbreiding van deze methodologie, met de geassocieerde moeilijkheden die we konden delen, bespreken en samen proberen oplossen. Ellen verdient tevens een extra bedankje voor het kritisch nalezen van deze thesis. Daarbij ook Vicky, ervaren met al-

gemene DNA kennis. Dieter, Annelies and Sara, the experienced peptide chemists who introduced me in that world. Kurt om het werk van furan incorporatie in peptiden en daaropvolgende oxidatie mee te delen, Yara and Abhishek for sharing work on the mimics. Daarbij ook Matthias, ervaren met algemene peptide kennis. Ik ben Jos diep dankbaar voor het oplossen van alle oplosbare (want ik weet dat dit veelal niet het geval was ;)) al dan niet technische problemen. Natuurlijk ook de geassocieerde PDC groep, met Prof. Pierre De Clercq, Prof. Johan Winne, Duchan en Bram was essentieel voor de organisch synthetische expertise. Ik ben ook mijn thesisstudenten Sven, Joke en Andreas dankbaar voor het werk dat ze geleverd hebben. The new generation Lars, Vincent, Eri, Margarida, Eva, Lidia, Jan, Dorien, Tim, Daan and Cecilia, I wish a lot of success.

I would also like to thank some other people of the S4 and S4bis. Prof. Richard Hoogenboom and his group that 'invaded' our corridor, are thanked for the atmosphere they brought. Prof. Du Prez en Pieter worden bedankt voor de succesvolle samenwerking voor de sequence controlled oligomers. Prof. Jose Martens en Katelijne voor de Quantas studie. Ook Tim en Katrien van de groep voor de technische ondersteuning van de NMR analyses. Jan dient uitvoerig bedankt te worden voor alle LC-MS analyses van mijn typisch niet analyseerbare stalen en de bijhorende uitleg. Mark en Pieter zijn bedankt voor de ondersteuning van ESI-MS analyses. Het admin team: Tom, Veerle, Ingrid, Freddy, Christel, Carine en in extentie ook Paul en Tom zijn bedankt voor de administratieve ondersteuning.

An important and big acknowledgement goes to Prof. Takashi Morii, who I consider as my second promotor. You gave me a place in your group in the beginning of the PhD as an unexperienced researcher. This allowed me to selfishly learn and absorb knowledge and expertise. An amazing experience to grow scientifically, but also personally, through the Japanese culture, food and drinks, that you allowed me to indulge. In the Morii group I am further endebted to Tainaka-sensee, Nakata-sensee, Shun, Matsumoto, Fong Fong, Tareq, Chisana, Li, and Yasuko for all the help.

Prof. Thomas Carell is acknowledged for allowing me to learn and try the incorporation of furan-modified amino acids in proteins. Michael Gattner explained and showed this all to me and thus also deserves credit.

Prof. Hiroshi Sugiyama is acknowledged for allowing me to learn and try to work with DNA origami. In the Sugiyama group I am endebted to Endo-sensee, Hidaka-san, Emura-san, Suzuki-san and Yangyang.

Several other people should be thanked for valuable discussions, advise and/or analysis. Prof. R. Stromberg and Malgorzata are acknowledged for ESI mass analysis and discussion. Prof. K. Gevaert en Hans worden erkent voor nano LC-MS analyse. Prof. J. Barciszewski is acknowledged gratefully for bringing kinetin under our attention. Prof. W.J. Markiewicz is thanked for synthetic discussions. Prof. Istvan Marko is acknowledged for the electrochemical oxidation of furan.

Financially, I am endebted to 'het Fonds voor Wetenschappelijk Onderzoek-Vlaanderen' for an aspirant position. UGent voor de beschikbare middelen, infrastructuur en een BOF mobility grant. The Japanese Society for the Promotion of Science [project VS.01S.10N], European Cooperation in Science and Technology [COST action TD0905], and Marie Curie ITN SO$_2$S. Also the scientific committees of several meetings for offering me a travel grant.

Uiteraard wil ik ook de mensen 'thuis' bedanken, vrienden en familie voor de steun, ontspanning en om mijn toch wel dominante wetenschapper kant te tolereren. In het bijzonder een bedankje voor mijn mama die mijn studies mogelijk gemaakt heeft door er initieel voor te zorgen dat ik ingeschreven raakte, en daarna financieel en met een overvloed aan brainfood.

In the end I would like to thank the members of my jury, with Prof. Emer. Pierre De Clercq, Prof. Daniel Summerer and Prof. Annemieke Madder for carefully reading and correcting the manuscript and further Prof. Takashi Morii, Prof. Piet Herdewijn, Prof. Kris Gevaert and Prof. Filip Du Prez as chairman.

Finally I would like to apologize to everyone who was not included here and in the reference list, in an effort to keep this thesis concise and printable.

Contents

List of compounds

List of furan-modified amino acids
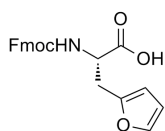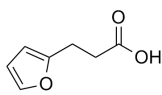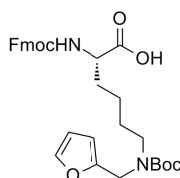
| (4.2.1) Fmoc-aa1 | (4.2.2) aa2 | (4.2.3) Fmoc-aa3 | (4.2.4) Fmoc-aa4 |

| (4.2.5) Fmoc-aa5 | (4.5.1) aa5 | (4.5.2) aa6 |
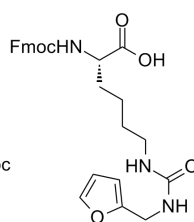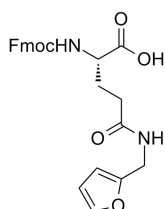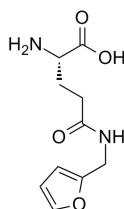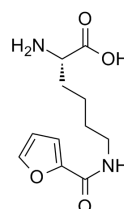
List of furan-modified nucleosides
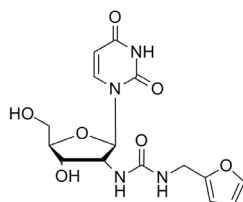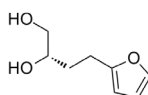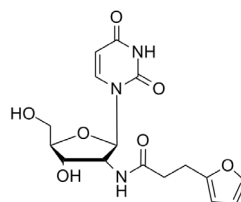
| (4.1.1) nu1 | (4.1.2) nu2 | (4.1.3) nu3 |

| (4.1.4) nu4 | (4.3.1) nu5 | (4.4.1) nu6 |
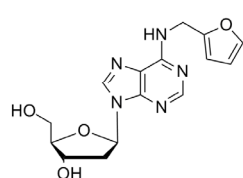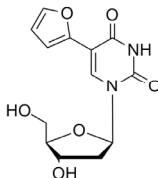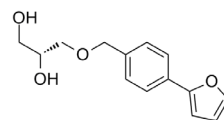
List of oligonucleotide sequences

GCN4 binding site

The GCN4 binding site is used to test the DNA binding of the Madder and the Morii mimic in sections 3.1 and 3.2 respectively. In section 4.2, it is used as the target for cross-linking from the furan-modified (**aa1**) Morii mimic. It was then modified with **nu5**, to test for cross-linking from the DNA to the binding Morii mimic, in section 4.3.

name	sequence	description
AP1	5'-CGGATGACTCATTTTTTTC-3'	unmodified
CRE	5'-CGG ATG ACG TCA TTT TTT TTC-3'	
CECRE	5'-CGG ATT GCG TCA TTT TTT TTC-3'	
CECRE mis	5'-CGG ATT GCG TCA TCT TTT TTC-3'	
CRE1	5'-CGG A**nu5**G ACG TCA TTT TTT TTC-3'	**nu5**-modified
CRE2	5'-CGG ATG ACG **nu5**CA TTT TTT TTC-3'	
CRE3	5'-CGG AT**nu5** ACG TCA TTT TTT TTC-3'	
CRE4	5'-CGG ATG **nu4**CG TCA TTT TTT TTC-3'	**nu4-modified**
AP1 comp	5'-CAA AAA AAA TGA GTC ATC CG-3'	complements
CRE comp	5'-CAA AAA AAA TGA CGT CAT CCG-3'	
CECRE comp	5'-CAA AAA AAA TGA CGC AAT CCG-3'	
CECRE flip	5'-GAA AAA AAC ATG ACG CAA TCC G-3'	
CRE3 comp	5'-CAA AAA AAA TGA CGT AAT CCG-3'	
CRE1 compT	5'-CAA AAA AAA TGA CGT CTT CCG-3'	
CRE1 compC	5'-CAA AAA AAA TGA CGT CCT CCG-3'	
CRE2 compT	5'-CAA AAA AAA TGT CGT CAT CCG-3'	
CRE2 compC	5'-CAA AAA AAA TGC CGT CAT CCG-3'	
CRE3 compT	5'-CAA AAA AAA TGA CGT TAT CCG-3'	

Testsequences for interstrand cross-linking

These testsequences were used modified with various furan-modified nucleosides for interstrand cross-link formation.

name	sequence	remark
ODN4.1.1	5'-CGG ATG **nu4**CG TCA TTT TTT TTC-3'	**nu4** modif. ODNs
ODN4.1.2	5'-CTG ACG G**nu4**G TGC-3'	
ODN4.1.3	5'-CTG ACG G**A**G TGC-3'	unmodif. reference
ODN4.1.4	5'-CTG ACG G**nu3**G TGC-3'	**nu3** modif. reference
ON4.4.1	5'-CUG ACG G**U**G UGC-3'	unmodif. 2'OMe RNA
ON4.4.2	5'-CUG ACG G**nu3**G UGC-3'	**nu3** modif. 2'OMe RNA
ON4.4.3	5'-CUG ACG G**nu6**G UGC-3'	**nu6** modif. 2'OMe RNA
ODN4.4.1	5'-CTG ACG G**T**G TGC-3'	
ODN4.4.2	5'-CTG ACG G**nu3**G TGC-3'	reference DNA
ODN4.4.3	5'-CTG ACG G**nu6**G TGC-3'	
ODN4.6.1	5'-ACG G**nu6**G T-GC TTC CGG TAC TAC GCC CAG ATG AGC TAC T-3'	**nu6** modif. origami ODN
ODN1	3'-GAC TGC CTC ACG-5'	
ODN2	3'-GAC TGC CGC ACG-5'	
ODN3	3'-GAC TGC CAC ACG-5'	DNA complements
ODN4	3'-GAC TGC CCC ACG-5'	
ODN5	3'-GAC TGC CTC ACG-5'	
ON4	3'-GAC UGC CCC ACG-5'	
ON5	3'-GAC UGC CUC ACG-5'	RNA complements
ON6	3'-CGU CAG CCC GCA-5'	

List of Abbreviations

aa	amino acid
Aba	4-acetamido benzoic acid
AcIm	1-acetylimidazole
Ac$_2$O	acetic acid anhydride
AcOH	acetic acid
Ad	adamantane
AFM	atomic force microscope
Alloc	allyloxycarbonyl (protecting group)
Boc	tert-butoxycarbonyl (protecting group)
Br	brominated species
BSA	bovine serum albumin
b-HLH-ZIP	basic helix-loop-helix zipper
bZIP	basic region leucine zipper
CD	cyclodextrin
COMU	1-Cyano-2-ethoxy-2-oxoethylidenaminooxy)dimethylamino -morpholino-carbenium hexafluorophosphate
CPG	contoled pore glass
DAPI	4',6-diamidino-2-phenylindole
DCC	dicyclohexylcarbodiimide
DCI	dicyanoimidazole
DCM	dichloromethane
DMTr	dimethoxytrityl
DHB	2,3-dihydroxybenzoic acid
DIPEA	diisopropylethylamine
DMAP	4-dimethylaminopyridine
DMF	N,N-dimethylformamide

DMT	dimethoxytrityl
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DPC	DNA-peptide/protein cross-link
dsDNA	double stranded DNA
EDT	ethanedithiol
EMSA	electrophoretic mobility shift assay
ESI	electrospray ionisation
EtOAc	ethylacetate
Exact	exact mass
FAB	fast atom bombardment
Fmoc	9-fluorenylmethoxycarbonyl
Gaba	γ-aminobutylbenzoic acid
HATU	N-[(dimethylamino-1H-1,2,3-triazolo[4,5-b]pyridin-1-ylmethylene]-N-methylmethanaminium hexafluorophosphate N-oxide
HBTU	N-[(1H-benzotrioazole-1-yl)(dimethylamino) methylene]-N-methylmethanaminium hexafluorophosphate N-oxide
HCA	α-cyano-4-hydroxycinnaminic acid
hν	UV-light or photocleavable linker
HOBt	1-hydroxybenzotriazole
HPA	hydroxypicolinic acid
HPLC	high pressure liquid chromatography
HRMS	high resolution MS
ICL	interstand cross-link
IPTG	isopropyl β-D-1-thiogalactopyranoside
LC	liquid chromatography
MALDI	matrix assisted laser desorption ionisation
MA SPPS	microwave assisted solid phase peptide synthesis
MeCN	acetonitrile
MeOH	methanol
MS	mass spectrometry
MW	molecular weight or microwave
mQ	milliQ, deionized H_2O
mRNA	messenger RNA
m/z	mass to charge ratio
NBS	N-bromosuccinimide
NMP	N-methyl-2-pyrrolidinon

NMR	nuclear magnetic resonance
NP	NovaPEG
nu	nucleoside
ODN	oligodeoxynucleic acid
oligo	oligonucleotide
ON	oligonucleic acid
PAGE	polyacrylamide gelelectrophoresis
Pbf	2,2,4,6,7-Pentamethyl-dihydrobenzofurane-5-sulfonyl (protecting group)
PDB	protein data bank
PEG	polyethyleneglycol
ppm	parts per million
PS	polystyrene
PSS	phosphorothiosulfide
PyBOP	1H-benzotriazol-1-yloxy-tris-(pyrrolidino) phosphonium hexafluorophosphate
Rf	ratio to front
RNA	ribonucleic acid
rt	room temperature
r.t.	retention time
ROS	reactive oxygen species
RP	reversed phase
SPPS	solid phase peptide synthesis
ssDNA	single stranded DNA
TBE	TrisBoratEDTA (buffer) or tris(hydroxymethyl aminomethane, $B(OH)_3$, ethylenediamine tetraacetic acid
TEA	triethylamine
TEAA	tetraethylammonium acetate
tBu	tert-butyl (protecting group)
TF	transcription factor
TFA	trifluoro acetic acid
TFO	triplex forming oligonucleotide
TG	TentaGEL
TIS	triisopropylsilane
TNBS	trinitrobenzenesulfonic acid
Tris	Trizma hydrochloride buffer solution
tRNA	transfer RNA
Trt	trityl or triphenylmethyl (protecting group)

UV   ultraviolet
XL   cross-link
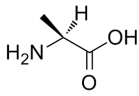
A
Ala
Alanine
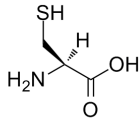
C
Cys
Cysteine
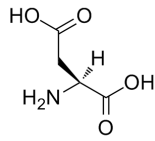
D
Asp
Aspartic Aacid
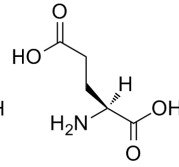
E
Glu
Glutamic Acid
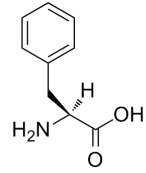
F
Phe
Phenylalanine
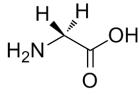
G
Gly
Glycine
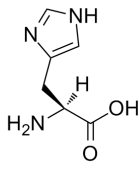
H
His
Histidine
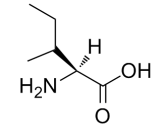
I
Ile
Isoleucne
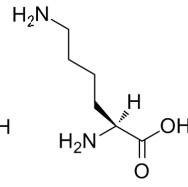
K
Lys
Lysine
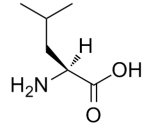
L
Leu
Leucine
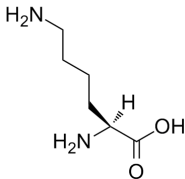
M
Met
Methionine
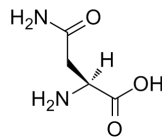
N
Asn
Asparagine
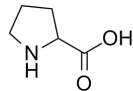
P
Pro
Proline
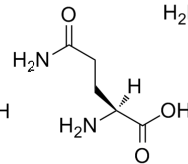
Q
Gln
Glutamine
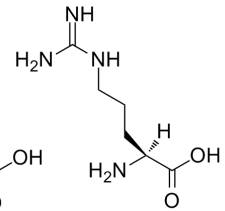
R
Arg
Arginine
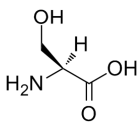
S
Ser
Serine
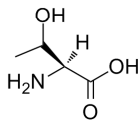
T
Thr
Threonine
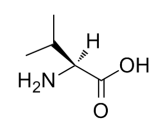
V
Val
Valine

W
Trp
Tryptophane
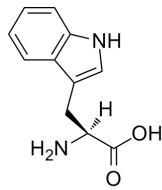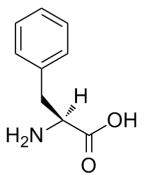
Y
Tyr
Tyrosine
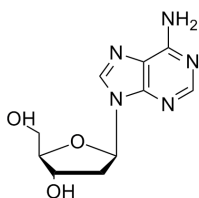
A
Deoxyadenosine
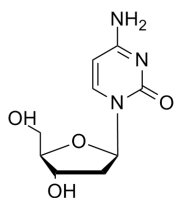
C
Deoxycytidine
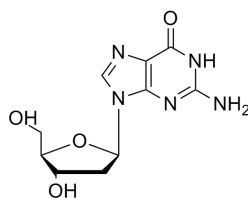
G
Deoxyguanosine
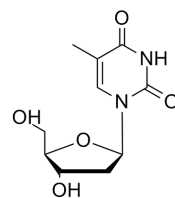
T
Deoxythymidine
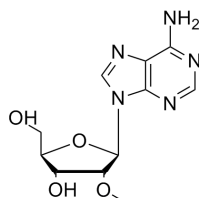
A
2'OMe-adenosine
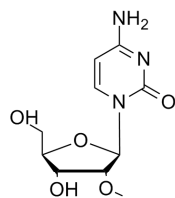
C
2'OMe-cytidine
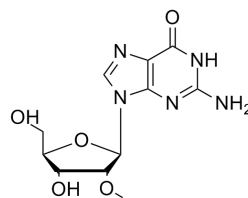
G
2'OMe-guanosine
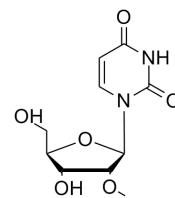
T
2'OMe-thymidine

Chapter 1

Outline of the Present Study and included Manuscripts

Part 1: Descriptive Part

chapter 1: Situation gives an overview of the available knowledge on the topic at the start of the study, based on a review written and published in the beginning of current work. Different approaches towards the miniaturization of transcription factors and cross-linking on the DNA-protein interface are described. A rationale for these efforts is provided.

- Peptide scalpels for site-specific dissection of the DNA-protein interface. Verzele D.*, Carrette L.L.G.* & Madder A., Drug Discovery Today: Technologies, 2010, 7(2), 115 (*equal contribution of both authors, upon invitation for the themed issue: Drug Discovery beyond the rule of five)

chapter 2: Transcription Factor Mimics describes the miniaturization of basic region leucine zipper transcription factor proteins. Two models are represented. First, a model developed in our group and therefore named 'the Madder mimic' is discussed. In this study, a peptidosteroid model, which was developed earlier (during the PhD thesis of D. Verzele) to improve pharmaco-kinetic properties, was revisited and optimized through the incorporation of

flexible linkers. Though this dramatically complicated the synthesis, through optimized protocols and the use of microwave assistance, the improved model could be synthesized and DNA binding for the first time shown. In the second part, 'the Morii mimic', developed by the Morii group based on a non covalent adamantane-cyclodextrin inclusion complex is described. It was selected from literature as a reliable model for DNA-peptide cross-linking in the next chapter and resynthesized during a research stay in the group of Prof. Morii.

- Peptidosteroid tweezers revisited: DNA binding through an optimized design. Carrette L.L.G., Morii T. & Madder A. (provisionally accepted in Eur. J. Org. Chem.)

chapter 3: Furan-based cross-linking   combines all results obtained with the furan oxidation cross-linking methodology. It starts with a short introduction of the methodology, developed in our group for DNA interstrand cross-linking. This is then followed by a case study on naturally occuring furan-modified DNA. The methodology was further shown to be compatible with RNA duplexes for interstrand cross-linking. It was attempted to visualise interstrand cross-link formation using DNA origami and a fast scanning atomic force microscope. The main part and goal of this section is the extension of the methodology towards its application on the DNA-protein interface. Using the transcription factor mimics from the previous chapter this was first attempted from a furan-modified peptide to the DNA, then from furan-modified DNA to the peptide and finally from a furan-modified protein to RNA.

- DNA Interstrand Cross-link Formation by using Furan as a Masked Reactive Aldehyde. Carrette L.L.G., Gyssels E. & Madder, A. Curr. Prot. Nucl. Acid. Chem. 2013, 54:5.12.1-5.12.16

- A synthetic oligonucleotide model to evaluate oxidation and cross-linking propensity of natural furan modified DNA. Carrette L.L.G.* & Madder A.* ChemBioChem 2014, 15(1), 103-107 (*co-corresponding authors), inside cover.

- Mildly inducible and selective cross-link methodology for RNA duplexes. Carrette L.L.G., Gyssels E., Loncke J. & Madder A. Org. Biomol. Chem. 2014, 12, 931-935.

- Toxicity Inspired Cross-linking for Probing DNA - Peptide Interactions. Carrette L.L.G, Morii T. & Madder A. Bioconj. Chem. 2013, 24(12), 20082014.

chapter 4: Method development   describes three results associated with solid phase synthesis, which is used to synthesize peptides and oligonucleotides: a) the use of a colour test to show the presence of functional groups on a solid support, b) an NMR based method for the concentration determination of peptides and oligonucleotides and c) a novel coupling methodology to synthesize functionalized oligomers on solid support from a single building block and without protecting groups.

- NF-31 color test uncovers 'hidden' alcohol functionalities in PEG-based resins for solid phase peptide synthesis. Carrette L.L.G., Verzele D. & Madder A. Tetrahedron Lett., 2010, 51(16), 2106-2108.

- Multi-functionalized sequence-defined oligomers from a single building block. Espeel P.*, Carrette L.L.G.*, Bury K., Capenberghs S., Martens J., Du Prez F. & Madder A. Angew. Chem. Int. Ed. 2013, 52(50), 13261-13264 (*equal contribution of both first authors).

Part 2: Experimental Part   lists the general protocols and provides additional supporting information (specific protocols and spectral data) for all sections of the descriptive part, represented in the same order.

Part I

Descriptive Part
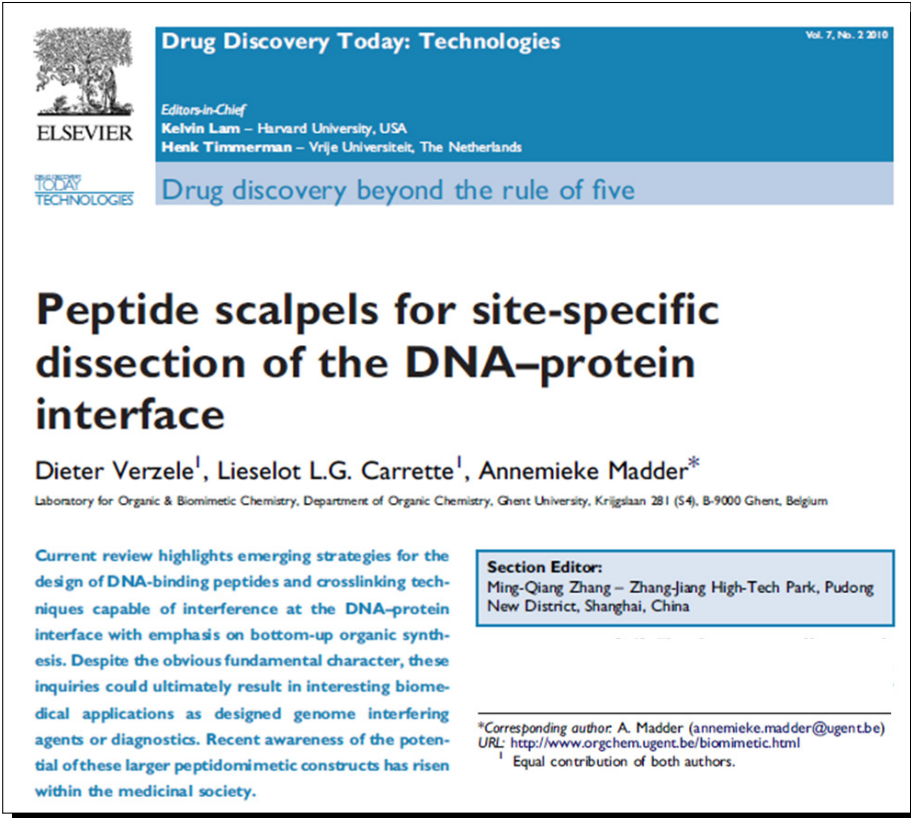
Chapter 2

Situation

To situate this work, I would like to begin with a manuscript written at the start of this PhD thesis in 2010, reviewing the state of the art in the field concerning the goals that had been set for this PhD study:

- *To prove DNA binding with our developed peptidosteroid transcription factor model*

- *To obtain cross-linking from a DNA binding peptide to its DNA target*

This review is thus ideal for the situation of the research. It gives an overview of peptidic transcription factor mimics binding duplex DNA and thus acting on the DNA-protein interface and the available methods for cross-linking on this interface with their pros and cons.

Peptide scalpels for site-specific dissection of the DNA–protein interface

Dieter Verzele[1], Lieselot L.G. Carrette[1], Annemieke Madder*

Laboratory for Organic & Biomimetic Chemistry, Department of Organic Chemistry, Ghent University, Krijgslaan 281 (S4), B-9000 Ghent, Belgium

Current review highlights emerging strategies for the design of DNA-binding peptides and crosslinking techniques capable of interference at the DNA–protein interface with emphasis on bottom-up organic synthesis. Despite the obvious fundamental character, these inquiries could ultimately result in interesting biomedical applications as designed genome interfering agents or diagnostics. Recent awareness of the potential of these larger peptidomimetic constructs has risen within the medicinal society.

Section Editor:
Ming-Qiang Zhang – Zhang-Jiang High-Tech Park, Pudong New District, Shanghai, China

*Corresponding author: A. Madder (annemieke.madder@ugent.be)
URL: http://www.orgchem.ugent.be/biomimetic.html
[1] Equal contribution of both authors.

Introduction   Aberrant modulation of gene expression at the transcriptional level is at the origin of numerous diseases. By disregulating cell growth and triggering cell proliferation, various oncoproteins carry out their biological functions as transcription factors, crucially relying on their DNA-binding capacities. Artificial transcription inhibition or modification is possible via various approaches through different interfaces (Fig. 2.1). Selective recognition between the reaction partners on the interface is essential for success. Truncated mimics that lack, for example, the activation domain while retaining the DNA-binding requirements might interfere with DNA complexation of the malignant natural counterpart, disrupt transcriptional induction and ultimately suppress the oncogenic activity.[1, 2] This short review will treat such peptidic transcription factor mimics, acting on the DNA-protein interface by binding the duplex, ultimately towards irreversible

blocking via interface cross-linking. Besides the simple desire to mimic Nature, these miniature probes hold potential for disentanglement of the molecular basis of DNA-targeting by natural proteins.

Artificial transcription factor mimics: gradual expansion of the chemical space Fifty years after the milestone discovery of the Watson and Crick double-helical structure of DNA, the official completion of the Human Genome Project (HGP) in 2003 was another landmark event in the history of natural sciences.[10, 11] Multidisciplinary research at the interface of various disciplines became prerequisite to distill the flood of new information and reap maximum reward from the genome data. The nascent discipline of chemical biology approaches biological problems from the chemist's point of view and employs its ability to proficiently design, prepare and characterize tailor-made compounds, which can be subsequently applied to probe biological systems in highly-controlled, well-defined experiments (reviewed in [12]). Organic synthesis occupies an indispensable position in the discovery process, enabling the craftful preparation of tailor-made conjugates, not limited by the genetic code or the structures of the ribosomal machinery. Through sophisticated strategies, the rigorous control of molecular structure and function provides access to the vast expanse of chemical space.[13] The targets of chemistry research in these days are no longer restricted to small drug-like molecules. Much of the intellectual adventure and challenge has gradually converged with biology and the size of aimed compounds has steadily increased. The activity, selectivity and stability of biopolymers become controllable, and they can be tailor-made like small organic compounds. Furthermore, chemical protein synthesis has enabled the systematic development of proteins with enhanced potency and specificity as candidate therapeutic agents.[14] Consequently eliciting a recent awareness within the pharmaceutical society for the unexplored opportunities of the generally ignored chemical space (GICS), medicinal chemists start thinking outside their Lipinski box.[15]

Despite the promising relevance of probing simplified peptide versions derived from natural protein counterparts, progress has been remarkably slow, failing to keep pace with the plethora of resolved complexes. The fragile balance between structural minimization and biophysical outcome impedes rational design of downsized mimics, demanding devoted trial-and-error tuning of empirical constraints.
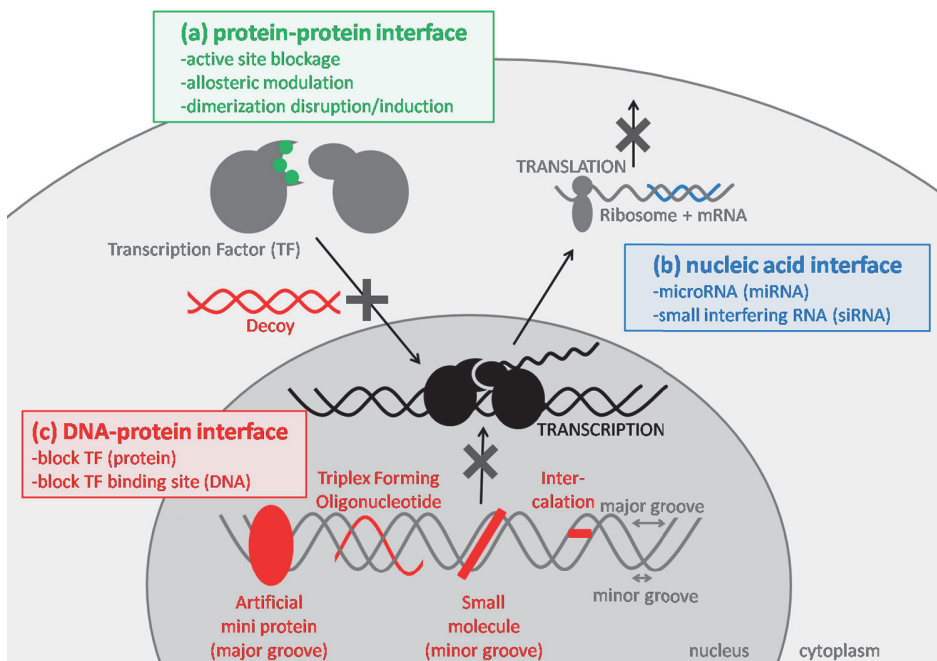
Figure 2.1: Different biological interfaces are involved in the processes of transcription and translation. a) Typically several proteins have to associate to obtain an active transcription factor and/or to form the transcription initiation complex on the DNA. Disruption of these interactions can be modulated by small molecules, which block the interaction.[3–5] b) Promising strategies exploiting the nucleic acid interface aim at inhibition of translation. This can be achieved by blocking mRNA with a short complementary RNA fragment (antisense RNA). Alternatively a double stranded RNA fragment (small interfering RNA) or single stranded RNA fragment (micro RNA) together with the RNA-induced silencing complex (RISC) can degrade mRNA.[6, 7] c) The DNA-protein interface is an emerging site for new generation therapeutics.[8] Interference can be achieved either at the protein level with decoys (small synthetic DNA fragments that bind to the transcription factor by mimicking the specific DNA binding site)[9] or at the DNA level by blocking the target site. The latter can be realized by a small intercalating molecule, which slides in between DNA base pairs; a triplex forming oligonucleotide, which binds in the major groove of DNA through Hoogsteen interactions; a small molecule binding in the minor groove or a large molecule binding in the major groove. In the latter approach, initial strategies focus on DNA-binding, and at most inhibition of gene expression. By linking this DNA-recognizing domain to an activating effector domain, selective modulation of the transcription process could ultimately be envisioned. While size of the compounds will probably hamper both synthesis and permeability, fascinating examples have been disclosed for the (synthetic) minor-groove binding Dervan polyamides[18-20] or (mostly, if not all biotechnologically-engineered) major-groove targeting zinc fingers.[25] Successes are most certainly related to the exquisite control of recognition specificity, the modularity of systems and the synthetic accessibility, as detailed hereafter.

Minor Groove Recognition: synthetic polyamides Dervans polyamides, based on the minor groove binder Distamycin A (Fig 2.2a), can be regarded as peptides with each α-carbon replaced by a five membered ring (Fig 2.2d). Natural Distamycine is a crescent shaped, small polyamide with three pyrrole (Py) rings that binds AT rich sequences in the minor groove selectively. Dervans group demonstrated that substituting pyrrole by imidazole (Im) and introducing hydroxypyrrole (Hp), it became possible to recognise all bases by a specific combination of two rings.[16] The genetic code was cracked. As the Dervan peptides bind the minor groove in a 2:1 fashion, affinity could be further increased by linking two peptides.

Additional modifications introduce specific properties to the DNA binders. Fusing a polyarginine to the tripyrrole increases binding affinity and promotes nuclear uptake even though these small molecules are already relatively cell permeable.[17] Linkage of a Dervan peptide, binding in the regulating domain of a gene, with the activator domain of a transcription factor allows for selective transcriptional activation of that specific gene.[18] Conjugation with a natural alkylating moiety, like CC-1065 or duocarmycin, increases the efficiency and selectivity of the DNA alkylation and can be used for gene silencing, as alkylated coding regions cannot be read by RNA polymerase II.[19, 20] Most proteins bind in the major groove (excluding the TATA box binding protein) but except for a few cases were major groove protein binding is inhibited by minor groove binding Dervan peptides,[21] the latter are unsuitable for blocking the DNA-protein interface, as it was shown that both can bind at the same time.

Figure 2.2 *(facing page)*: **Non-exhaustive, pictural overview of synthetic peptidic groove binders, based on the natural DNA binding moieties distamycin (a), zinc fingers (b) or basic region leucine zippers (c).** Based on distamycin, Dervan developed hairpin minor groove binding polyamides of pyrrole (Py), imidazole (Im) and hydroxypyrrole (Hp) (d). Small monomeric peptides can bind the major groove as α-helix, if the associated entropic penalties are compensated via preorganization (e). Hybrid conjugates of a tripyrrole sequence and a peptide monomer represent another strategy to cope with thermodynamics. The former module assists in the major groove recognition by the latter module which mimics either the atypical zinc finger, GAGA (f) or the basic region leucine zipper GCN4 (g). Furnishing genuine mimics of natural proteins, a final approach consists in the combination of several major groove binding peptides, such as homodimeric GCN4 (h) or heterodimeric cMyc-Max (i) sequences.
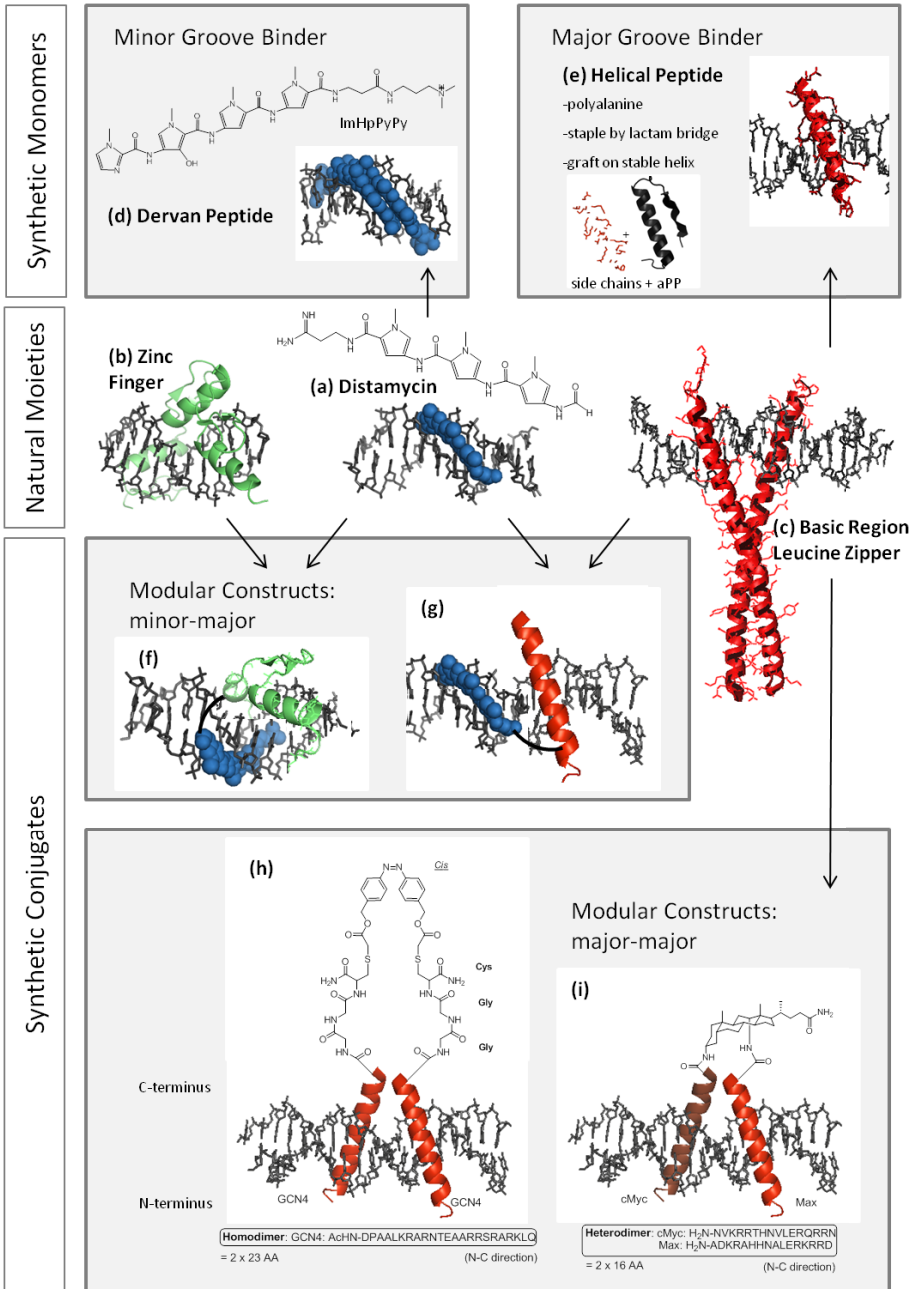
Ultimately, to uniquely distinguish a site in the three gigabase human genome, a recognition sequence of 16 base pairs is required. Although examples of polyamides that can realise this exist, mismatch sequences will probably not disrupt binding, due to the great binding affinity of such a single molecule-DNA binder.[22]

Major Groove Recognition: stabilised monomeric helices The majority of well-characterized families of native DNA-binding proteins rely on recognition α-helices to make base contacts in the major groove. The overall shape and dimensions of an α-helix allow it to fit into the major groove in a number of related, but significantly different ways. This variability explains the preference of natural proteins for targeting this wider groove over the narrower minor one.

The design of peptides mimicking the DNA binding domain of such proteins is not straightforward, as most α-helices lose the capability to fold and thus bind DNA with high affinity when removed from their natural context.

Various possible strategies have been described[23, and references cited therein] and consist in the use of peptide fragments that stabilise an α-helix (Fig 2.2e). Successful substitution of non-crucial amino acids with the helix promoting alanine can be achieved. Helical structures can also be stapled by the formation of a lactam bridge between lysine and aspartic acid, four positions further in the chain. Alternatively one can make use of stable α-helices, as in the small, well-folded avian pancreatic polypeptide (aPP), where an α-helix is stabilised by hydrophobic interactions on one side of the helix. The other side of this helix can then be used to graft residues required for DNA recognition and binding.

Exploitation of minor groove conjugation for major groove recognition It was shown that the covalent attachment of peptides to the DNA backbone through a linker can compensate for the entropic cost associated with secondary structure formation and bind DNA as an α-helix. Next to intramolecular binding to increase the affinity for DNA, cooperative binding is another option.[23, and references cited therein] This can be achieved by fusing intercalators or Dervan peptides to the major groove binding peptide (Fig 2.2f and 2.2g). Combining different smaller DNA binding units, overall binding selectivity can be significantly

improved.

Engineered Zinc Fingers gripping the major groove   In the area of
zinc finger transcription factors, an advanced level of insight and pairing control
has been achieved.[24] Whereas other DNA-binding proteins generally make use of
the 2-fold symmetry of the double helix, zinc fingers are linearly linked in tandem
to recognize DNA stretches of varying lengths, with high discrimination fidelity.
The ability to bind specifically to virtually any given DNA sequence, combined
with the potential of fusing polydactyl zinc finger peptides with effector domains,
has led to the technology of engineering of chimeric DNA modifying enzymes
(equipped with e.g. methylase, nuclease, recombinase modules) and transcription
factors (simple blockage without effector, repression vs. activation domains, and
chromatin remodeling factors as examples).[25] Opening the possibility of using
the engineered zinc finger-based factors as novel human therapeutics, ongoing suc-
cesses stimulated the creation of the first biotech companies to exploit this new
platform (Sangamo BioSciences, Inc., Richmond, California; and later Gendaq
Ltd., as a spin-off of the MRC Laboratory of Molecular Biology, Cambridge, UK).
Interestingly, customized zinc finger constructs for given DNA sequences can now
be acquired commercially from Sigma-Aldrich, branded as Compo-Zr technology.
Table 1 includes an example of a Zinc finger construct for therapeutic applica-
tion. Although a vital sequel in the gradual expansion of the chemical space
covering artificial transcription factors, most (if not all) zinc finger designs are
biotechnologically-engineered and as such beyond the scope of present overview,
emphasizing on miniature constructs synthesized via bottom-up organic synthesis.
Despite numerous further studies and applications, ranging from genome anno-
tation tools to (pre-)clinical implementation, the interested reader is therefore
referred to excellent reviews recently disclosed in literature.[26, 27]

Artificial zipper miniatures seizing the major groove   The α-helical
bZIP and b-HLH-ZIP motifs are among the simplest protein structures able to
bind the DNA major groove in a sequence-specific fashion. The simplicity and
tractability of the bipartite zipper-type proteins has made them attractive frame-
works for the design of artificial counterparts. The variety of designs exhibiting
native protein-like targeting of DNA-sequences by employing short basic region

peptides, dissected from naturally occurring bZIP transcription factors, demonstrates the potential of artificial miniatures as biomolecular recognition devices. However, considering above discussion, the inherently-low DNA-binding affinity of short, isolated, natural basic region peptide monomers does not surprise. As mentioned earlier, suppression of bZIP binding capacities by presenting so-called monomeric reading heads illustrates the delicate balance between entropy and enthalpy.

Research demonstrated that the basic region as such contains sufficient information for DNA recognition, provided helical structures are stabilized (vide supra) or artificial dimerization is applied. Predominantly serving as a structural scaffold to pre-organize the relatively small, N-terminal recognition α-helices, the extended C-terminal dimerization domain can be replaced by various connectors, both covalent and non-covalent.[23, and references cited therein][28] Since the original disclosure of a cystine-bridged peptide dimer mimicking the homodimeric yeast GCN4 leucine zipper protein, development of subsequent generation bZIP miniatures has been intensely pursued. Early work of Schepartz et al. focused on the chemical synthesis of (homodimeric) GCN4 mimics upon artificial dimerization via formation of a non-covalent bis(terpyridyl)-iron(II) coordination complex. Related contributions were further disclosed by Morii et al., typically employing either a non-covalent β-cyclodextrin:adamantyl host-guest inclusion complex or covalently bridged enantiomeric biphenyl derivatives as artificial dimerization interfaces in a variety of homo- and heterodimeric GCN4-based designs.

The apparent hibernation in this research area notwithstanding, Mascareñas' artificial GCN4 homodimer, represented in Fig. 2.2h, illustrates current state-of-the-art in the comprehension of design principles and control over binding properties for simultaneous DNA major groove targeting. Incorporation of a photo-isomerisable azobenzene moiety allowed for generating a light-modulated DNA-binding peptide.

Although sharing a similar mode of binding DNA target sites with the extensively scrutinized bZIP-equivalents, artificial mimics of the basic helix-loop-helix zipper (b-HLH-ZIP) transcription factors are conspiciously few. Despite the exceptional total chemical synthesis (172 residues) yielding a covalently-stitched replica of the vertebrate cMyc-Max (proto-) oncoprotein achieved by Canne et al. in 1995,[29] translation of this precedent towards miniature derivatives has

been long overdue. Related to the bZIP and b-HLH-ZIP type proteins, the singular MyoD-MyoD bHLH-type mimic, reported by the Morii group[30] before finally resorting to the GCN4 standard, confirms the interest for corresponding miniaturization efforts.

Directly involved in human tumorgenesis and cancer, the pivotal cMyc member of the Myc/Max/Mad network is a prominent example of how these proteins are promising targets towards novel, potent, transcription-targeting chemotherapeutica.[31] However, (1) the influence of loop-projecting DNA-contacts, (2) the ambiguous role of loop-associated flexibility, (3) the constraining/stabilizing interactions with essential accessory mediators of the transcriptional machinery in vivo and (4) the seeming regulatory subtleness of the helix-loop-helix transcription factor family conspire against the exploitation of their peptide counterparts in design efforts. In this respect, the heterodimeric Jun-Fos bZIP oncoprotein bears special mentioning, combining a cMyc-like cancer relevance with the GCN4-like structural tractability.

Although a decade passed since Mascareñas' triggerable GCN4-GCN4 bZIP contribution mentioned above, the contrasting void of helix-loop-helix counterparts stimulated initiatives from our own group. Recently, we achieved a first generation of down-sized b(-HLH-)ZIP models with emphasis on the therapeutically relevant cMyc-Max combination.[32] Based on covalent restriction by a steroid core with defined geometrical properties (as depicted in Fig. 2.2i), the orthogonal reactivity of the attachment points is an additional feature of the artificial linkage, permitting straightforward synthesis of not only homo- but also heterodimer peptidosteroid conjugates.

Table 2.1 highlights some of the therapeutic developments directly derived from the here discussed synthetic DNA binding peptidic constructs.

Cross-linking methods in development   In order to block the DNA-protein interface, the above described mimics need to compete with the natural DNA-binding proteins. Capitalizing on previous successes at the protein-protein interface, cross-linking moieties can be introduced to overcome the competition problem.

Table 2.1: Therapeutic developments of DNA-binding peptides

Name	Structure	Function	Stage	References
Distamycin A	Unmodified tripyrrole polyamide	No anti-tumor activity	No applications	[45]
Tallimustine	Benzoyl nitrogen mustard derivative of distamycin	Anti-tumor activity	Phase II (halted due to severe myelotoxicity)	[45]
Dervan-peptides	- Unmodified polyamide (binding in regulatory sequence)	Gene silencing of several genes (5S RNA, HIV-1 RNA, vascular endothelial growth factor, human transforming growth factor, prostate specific antigen, topoisomerase IIα, …)	Cell cultures	[19]
	- Polyamide (binding in regulatory or coding sequence) conjugated with an alkylating moiety (duocarmycin (Du86), CPI, CBI, CC-1065, distamicin)	Gene silencing of several genes	Cell cultures	[19]
	- Polyamide conjugated with chlorambucil (nitrogen mustard)	Blocking tumorigenicity of carcinomic cells	Animal model (nude mice)	[46]
	- Polyamide modified with an activation domain	Activation of gene expression	*In vitro* (yeast nuclear extract)	[18]
	- Polyamide modified with a positive patch (N-aminoalkylpyrrole, Arg-Pro-Arg or polyamine)	Inhibition of major-groove binding proteins (e.g. bZip)	*In vitro* (purified TF)	[21]
PNU-145156E	Tetrasodium salt of tetra-pyrrole naphthalenedisulfonic acid	Anti-tumor activity: believed to be anti-angiogenic and synergistic with other cytotoxic drugs	Phase I (halted due to its unique toxicity profile)	[46]
Brostallicin (PNU-166196)	Bromo-acrylamido tetra-pyrrole derivative	Anti-tumor activity	Phase III (generated significant interest in the clinic)	[47]
MVZ+426	Designer zinc finger transcription activator of the (VEGF)-A gene	Treatment of human peripheral arterial obstructive disease by stimulating vessel growth	Phase II	[48]

Table 2.2: Classification of chemical strategies for cross-linking

Technology	Small bifunctional crosslinking entities	Reactive DNA	DNA with inducible reactivity	Sulfur modified DNA
Crosslinking moieties	Cisplatin, nitrogen mustards, psoralen, aldehyde, epoxides, …	- Nucleotide modified with aldehyde, crotonaldehyde, trans-4-hydroxynonenal, cyclopropapyrroindole, substituted pyrophosphate, … - Postsynthetic regioselective modification of sugars in oligonucleotides with vicinal diols	- Photoactivatable functionalities: azide, diazirine, benzophenone - Intrinsic photoreactivity of DNA	- PSS-backbone oligonucleotide (phosphorothiosulfide) - N^2 thiol modified deoxyguanosine
Pro	- No modification required of DNA/protein/peptide - Bioavailability	- Embedded crosslink functionality - Better selectivity	- On demand activation within formed complexes - Increased selectivity for target protein	- Selective for cysteine - Very small destabilization of the DNA duplex
Con	Lack of selectivity	Nonspecific reactivity before complex formation	- Activation required - Possible destruction of compounds under activation conditions - Low yield	- Correct and accurate positioning of the modification required - Formed disulfide link metabolically unstable
Selected examples	Crosslinking of transcription factors CREB and JUN or nuclear matrix-bound transcription factors to their recognition sequence	Crosslinking of DNA to restriction-modification enzymes, HNF1 transcription factor and T7 RNA polymerase	Crosslinking of DNA to DNA polymerase, study positioning of Lex A repressor in the major groove, photoaffinity labeling	Crosslinking of NF-κB decoy to the p50 subunit or crosslinking of HIV-1 reverse transcriptase to its template
References	[33,36,49]	[37,40,50]	[34,35]	[38,39]

Exogenous bifunctional cross-linkers   For cross-linking purposes, it is
most easy to add small reagents, like platina complexes, nitrogen mustards,
psoralen and reactive aldehydes (Fig. 2.3). In this case no modified DNA or
protein is required and their size allows easy penetration in cell and nucleus.
Most current chemotherapeutics belong to these compound classes and result
in DNA intra- or interstrand cross-linking. The impact of covalent attachment
of proteins was recognised more recently. It was shown that cisplatin, the best
known anticancer drug, cross-links nuclear matrix-bound transcription factors
and cofactors to DNA.[33]
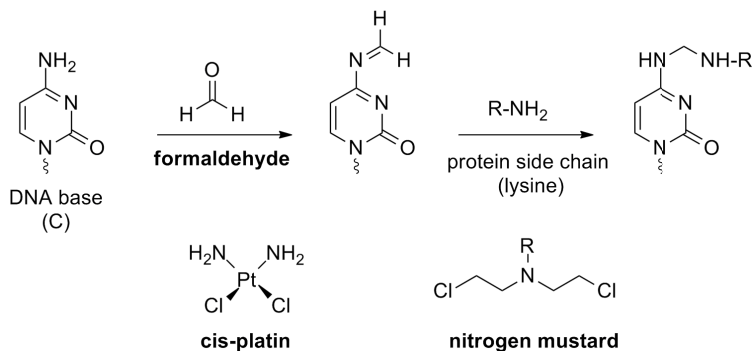
Figure 2.3: Examples of exogenous bifunctional cross-linkers formaldehyde
(top), cis-platin (left), nitrogen mustard (right)

The problem with these external reagents is their lack of selectivity. Selec-
tivity can be dramatically increased by equipping either DNA or protein with
reactive groups. Alternatively, incorporation of more selective photoactivatable
cross-linking moieties has gained popularity. Over the years, an extensive col-
lection of cross-linking techniques has been developed, mainly to study transient
complexes between proteins or between proteins and DNA. Since it is not the
intention of this paper to give an exhaustive list, the reader is referred to more
specialized literature.[34–36] Table 2.2 intends to summarize the different chemical
strategies for cross-linking.

DNA → Protein cross-linking   Many examples exist in which nucleotides
are equipped with a reactive functionality, in order to form a covalent bond to

a non-covalently bound protein (Fig. 2.4). In first instance, it is possible to modify the phosphate backbone, for example with pyrophosphate internucleotide groups [37] or disulfide groups. In the last case a PSS-backbone (phosphorothiosulfide) is formed, which can break and further react selectively with protein cysteine groups. Cells treated with a PSS-oligonucleotide of the binding site for NF-κB transcriptional factor underwent apoptosis.[38] Secondly, nucleotide building blocks equipped with a reactive functionality (aldehyde, aziridine, alkylating moiety) can be used in automated DNA synthesis for the production of reactive oligonucleotides.[39] Additionally, various approaches for postsynthetic regioselective modification to introduce reactive functionalities into DNA have been described.[40] So called decoys, as depicted in Fig 2.1, have been developed for irreversible cross-linking and inhibition of transcription proteins.[36]
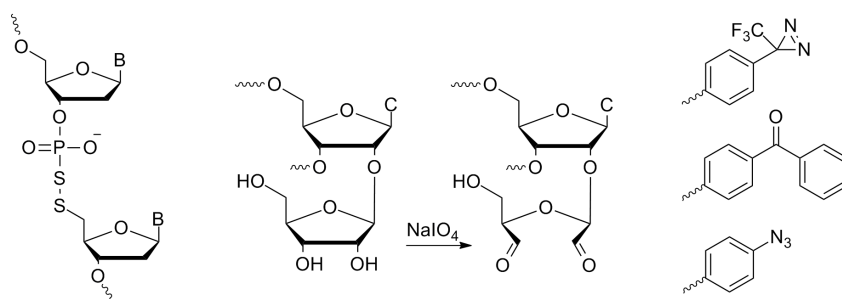
Figure 2.4: Examples of cross-linking modifications in DNA for reaction to proteins PSS-oligonucleotide (left), aldehydes (Crib) (middle), photoreactive 3-trifluoromethyl-3-phenyldiazirine, benzophenon and phenylazide (right)

Protein → DNA cross-linking As for cross-linking to DNA, strategies involving TFO (see Fig 2.1) for covalent blocking of the genome have been considered. Inhibition of transcription in cells was observed after application of a psoralen-modified oligonucleotide and irradiation.[41]

Alternative approaches involving cross-linking peptides or proteins for achievement of covalent DNA blocking have received considerably less attention. The most straightforward way to introduce a cross-linking functionality in peptides or proteins involves cysteine side chain modification. This was described for determination of the orientation of a DNA binding motif within a complex[42] and later applied for determination of other DNA-protein complexes. Furthermore,

unnatural reactive moieties (examples in Figure 2.5) can be included during so-
lution or solid phase peptide synthesis. Even in proteins unnatural amino acids
can be introduced via the natural or mutated translation system.[43, 44] Recently
a photoactivatable cross-linking protein, obtained by genetically encoding for a
benzophenone containing amino acid was cross-linked to its DNA binding site.[44]
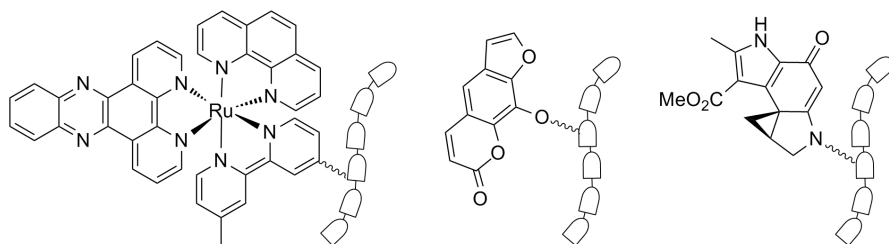
**Figure 2.5: Examples of reactive moieties incorporated in peptides for con-
jugation to DNA** Ruthenium complec [Ru(phen)(bpy')(dppz)]$^{2+}$ [51] (left), psoralen
[52] (middle), duocarmicin based (S)-CPI (cyclopropapyrroloindool) [53] (right)

Although application of specifically engineered proteins for DNA cross-linking
has thus proven feasible, the clear preference for the above described reverse ap-
proach can be explained in terms of the unpredictable influence of protein modifi-
cation on DNA binding affinity. The here described DNA binding peptides, being
potent candidates for mimicking their parent proteins, bring along new possibil-
ities in this area. As mentioned before, introduction of sequence selectivity to
DNA targeting agents does improve their efficiency as anticancer drug. Following
this trend, cross-linking transcription factor mimics might play an important role
in future cancer treatment.

Conclusions   Illumination of the molecular mechanism by which cells adapt
their phenotype in response to external stimuli is a central objective in mod-
ern life sciences. The use of exogenous chemical approaches based on simplified
probes derived from natural proteins provides a potent alternative to interrogate
and decipher cellular behavior, supplementing biological methods (e.g. knock-out
organisms or expression profiles). Chemical biology is eliciting a paradigm change
in the way scientists approach therapy-directed biomedical questions and diagnos-
tics development, improving innovation and productivity towards drug discovery
of the future. Prominent practitioners of chemical biology research recently advo-

cated the eminent role of this burgeoning field in drug discovery, towards innovative cures of the 21st century.[12] Harboring the potential to bridge fundamental academic aspirations with pharmaceutical industry's objectives, i.e. to genuinely alter the perception of the pharmacologically achievable, this integrative discipline offers invaluable incentive to elicit uncovered therapeutic targets. The ultimate goal remains the expansion of the diagnostic toolkit of biomarkers and the enrichment of the limited arsenal of contemporary pharmacopeia. Corroborating the relevance and opportunities presented in current review, a profound impact of tailor-made bioconjugates, down-sized probes and small molecule libraries is anticipated, and so-called 'undruggable' targets (and processes) such as (oncogenic) transcription factors might accordingly shift into the realm of the medicinally conceivable. Bayer's acknowledgment for the fundamental research on transcription regulation at a molecular level by Prof. Dr. Patrick Cramer (see links) illustrates this change of mentality, while the 2006 Nobel Prize of Chemistry awarded to Prof. Dr. Roger Kornberg (see links) confirms the value of continuing efforts.

Designed mimics of transcription factors are valuable to assess the implications of the exquisite recognition selectivity of natural transcription factors on the control of gene expression, towards molecular dissection of the complex regulatory networks and functional annotation of genetic information. Primarily unveiling fundamental details of the dynamic interplay between biomolecular systems, biomedicinal diagnostics and anti-gene chemotherapeutics might eventually arise from elucidation of key interactions, interrogation of aberrant transcriptional behavior, and identification of new genetic targets. Recent progress in peptide delivery systems and the prospect of developing peptidomimetic counterparts of in vitro active derivatives seem to warrant the ultimate medicinal potential of this research area. Tackling human disease in an early stage at the fundamental DNA-protein interface, the significantly rewarding ends for cancer research demonstrate the relevance and justify the means of ongoing endeavors.

Outstanding issues

- Although specific DNA recognition is claimed for most artificial transcription factor mimics, critical interpretation of assay results is needed because single target binding in a genome-wide context is yet to be verified.

- In spite of the recent progress in structural and biochemical studies, a universal code for the recognition between proteins and nucleic acids via the major groove has yet to be generalized.

- Despite the contemporary interest in compounds beyond the Lipinski rule of five, the medicinal applicability will benefit from ongoing miniaturization efforts.

Related articles [23, 54–57]

Links

- The PIN Database (Proteins Interacting in the Nucleus): `http://pin.mskcc.org/home.html`.

- The Nobel Prize in Chemistry 2006 to Roger D. Kornberg (Stanford University, USA) for his studies of the molecular basis of eukaryotic transcription: `http://nobelprize.org/nobel_prizes/chemistry/laureates/2006/`.

- Recently acknowledged at the Bayer website, Prof. Dr. Patrick Cramer (LMU Munich, Germany) occupies a prominent role at the forefront of transcription research, elucidating the chemical processes which turn the building blocks of DNA into a living, biological process: See (a) Podcast: `http://www.podcast.bayer.com/module/podcast/tv-research-en.xml`; (b) 2009 Hansen Family Award: `http://www.research.bayer.com/edition-21/science_and_education.aspx`.

- Sangamo BioSciences, Inc., biopharmaceutical company exploiting the zinc finger technology: `http://www.sangamo.com/`.

- The Sigma-Aldrich CompoZr^{TM}zinc finger technology offer: `http://www.compozrzfn.com/`.

- The Zinc Finger Database (ZiFDB): `http://bindr.gdcb.iastate.edu/ZiFDB` (Fu et al. (2009) Zinc Finger Database (ZiFDB): a repository for information on C2H2 zinc fingers and engineered zinc-finger arrays. Nucleic Acids Res. 37 (database), D279D283)

- Zinc Finger Targeter (ZiFiT), a web-based tool to facilitate design of zinc finger proteins: `http://bindr.gdcb.iastate.edu/ZiFiT` (Sander et al. (2007) Zinc Finger Targeter (ZiFiT): an engineered zinc finger/target site design tool. Nucleic Acids Res. 35 (web server), W599-W605)

Chapter 3

Transcription Factor Mimics
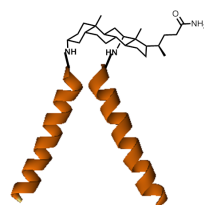
3.1 Madder Mimic: dipodal peptidosteroid

As introduced in the previous chapter there is an interest in DNA binding peptides, based on natural transcription factors for probing the DNA-protein interface and for potential biomedical applications. Our group contributed in the field with the design of a dipodal peptidosteroid tweezer, an artificial major groove zipper with a cholic acid as dimerizing scaffold.[58] The reported design did however not result in DNA binding and was therefore revisited and optimized. The following is the full paper that describes the optimized design, the associated synthetic challenges that were faced and met and the achieved selective DNA binding.

Peptidosteroid Tweezers Revisited: DNA Binding Through an Optimised Design

Lieselot L. G. Carrette,[a] Takashi Morii,[b] and Annemieke Madder*[a]

A dipodal peptidosteroid based on a deoxycholic acid scaffold was designed as a transcription factor model to study, and ultimately interfere in, selective protein–DNA interactions. In this paper, the selective binding of such a dipodal peptidosteroid to its DNA binding site is described for the first time, thus proving the concept and confirming the potential of the design. To achieve this, two 25-mer peptides were synthesized in parallel and in close proximity on a rigid amphiphilic scaffold, attached by flexible linkers. Peptide aggregation complicated the synthesis under conventional conditions to such an extent that optimization and the use of microwave assistance were required for a successful synthesis. As such, this synthesis is a good model to showcase the beneficial effects of microwave assistance in peptide synthesis.

Introduction   For few years, examples of non-Lipinski drugs from outside the traditional small-molecule box have been emerging on the market.[59, 60] To name one of over 50 approved peptide drugs, Roche's Enfuvirtide, a 36-mer peptide for HIV-fusion inhibition was approved in 2003. Clear advantages in activity, selectivity, and toxicology[60] are the basis of their utility as alternatives to low molecular weight drugs.[61] Indeed, more selective targeted interactions can result in a better treatment for a disease with fewer side effects. Such drugs can therefore be desirable alternatives to current very toxic chemotherapeutic agents, through direct selective interaction with a defective gene, or by selectively blocking the interaction of a transcription factor (TF) with its promoter. TFs play an important role in the regulation of gene expression, and it has been shown that in cancers, there is typically an overexpression of a limited number of these factors.[1, 2] This forms the basis for a new selective therapeutic platform. Transcriptional modification using artificial TF models has been demonstrated with the well-known polypyrrole amides, developed by Dervan, which show selective DNA minor-groove binding.[18–20, 62] However, a universal pairing code for recognition between natural amino acids and nucleotides for major-groove binding has not yet been established.[63–65]

In this context, TF proteins have been studied intensively in order to understand the remarkable binding selectivity and affinity that they show for the major groove of DNA in the complex cellular environment.[66] TFs typically bind DNA cooperatively as tweezers, embracing the DNA helix with two α-helical peptides that insert into the major grooves. The basic-region leucine zipper (bZIP) TF family simply consists of two uninterrupted α-helices (Fig. 3.1, left). The N-terminal basic-region peptides interact with the DNA through 20-30 mostly basic residues,[67] while the leucine-rich C-terminal part takes care of dimerization and correct positioning of the basic-region peptides through the formation of a coiled coil. By dissecting interesting areas (i.e., the DNA-binding regions) from the folded protein chain and assembling them onto a scaffold to obtain a similar geometry, protein properties can be transferred to smaller model systems.[68] Pioneering work by Kim et al. demonstrated that the entire leucine zipper part can be replaced by a smaller moiety to fulfill the dimerizing function for GCN4, a member of the bZIP family. By adding a cysteine and a spacer consisting of two glycines to the basic-region peptide, a miniature protein was obtained through disulfide-bridge formation.[69] Since then, other examples of GCN4 models have been reported, dimerizing the basic-region peptides through a branched lysine,[70] a bridged biphenyl moiety,[71] a hostguest inclusion complex (Fig. 3.2C),[72]an Fe complex,[73, 73] a photoresponsive azobenzene,[74] PNA (peptide nucleic acids),[75] etc.[76–78] The reduction of the size from proteins to peptides not only makes them synthetically accessible, allowing unnatural modifications derived from the organic chemist's imagination, but it also improves issues with stability, delivery, and availability.[79]

Previously, we designed and synthesized a dipodal peptidosteroid as a TF model of the cMyc-Max oncoprotein,[58] a transcription factor belonging to the basic helix-loop-helix zipper (b-HLH-ZIP) family, a class of TFs related to the bZIP family, containing an additional loop. Although the envisaged molecule was obtained in good purity, DNA binding could not be achieved. Visualization based on the PDB structures of the TF proteins (Fig. 3.1, middle) showed that truncation of the α-helices to the basic-region peptides and direct conjugation to the steroid scaffold does not permit adequate insertion of these peptide helices into the DNA major groove. Despite a clear parallel positioning
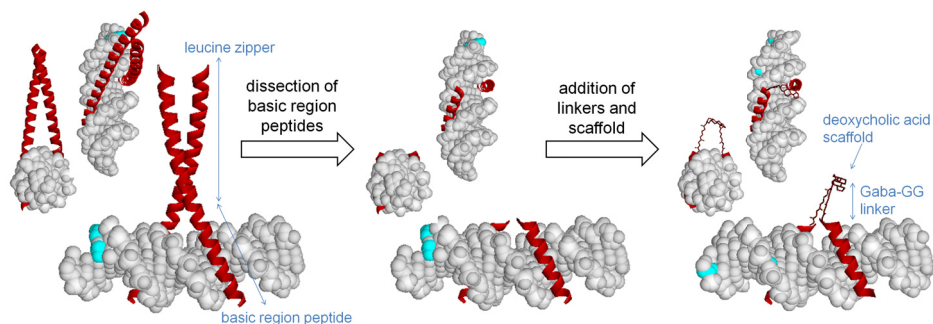
Figure 3.1: Design of a TF mimic of GCN4, a basic region leucine zipper TF (left) by dissection the basic region peptides (middle) and immobilizing them through linkers (Gaba-GG) on a smaller scaffold (deoxycholic acid) for dimerization (right) in side, front and top view. The middle panel clearly shows the required divergence for insertion in the major groove of the DNA and the steric encumbrance between the steroid and DNA that would result from direct conjugation of the basic region peptides to the steroid.

of the two peptides, imposed by immobilisation onto the steroid scaffold, the restricted divergence and steric encumbrance of such a constrained rigid model prevents the right positioning of essential side-chain contacts and the induction of α-helices, and thus DNA binding is hampered. This confirmed the need for a modified design, and prompted us to explore the potential benefits of a spacer moiety between the steroid scaffold and the basic-region peptides (Fig. 3.1, right). Although the introduction of flexible linkers into the design meant that the synthetic route had to be thoroughly optimized, we now demonstrate for the first time selective DNA binding using these peptidosteroid tweezers.

Results and Discussion

Optimized model design The use of a steroidal cholic acid moiety was inspired by the positive effects of its amphiphilic nature on the biostability[80] and bioavailability of an attached peptide,[81, 82] both of which are desirable for the further development of TF models. Cholic acid is taken up in cells both by active transport and by passive diffusion. Thus, cholic acid moieties have been added as absorption modifiers[83] or covalently attached to enhance the uptake of nanoparticles,[84] peptides, oligonucleotides, oligosaccharides, and small-molecule

drugs.[85, 86] Depending on the number of cholic acid residues and their mode of attachment, the passive diffusion is referred to as flipase activity[87] or umbrella transport.[88–93] Furthermore, cholic acid is perfectly suited to be used as a scaffold in multivalent and cooperative molecular engineering, and it has been extensively used for this purpose. In this context, these conjugates are often referred to as cholaphanes..[94] Ceragenins, cholic-acid-based mimics of antimicrobial peptides are less sensitive to degradation, and as such are more suitable for clinical development.[82] Furthermore, they are smaller and so synthetically more accessible.
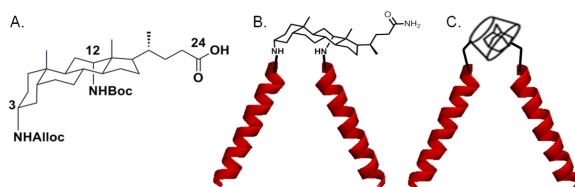
Figure 3.2: A) Orthogonally protected, C-3, C-12 diamino-modified deoxycholic acid scaffold; B) Proposed peptidosteroid model; C) Model described by Morii based on the adamantane-cyclodextrin inclusion complex.[72]

To dimerize the two GCN4 basic-region peptides, an orthogonally protected diamino steroid derived from commercially available deoxycholic acid (Fig. 3.2A) was used.[95, 96] The two orthogonally protected amino groups allow the attachment of two different DNA-binding peptides, both fixed in the α-orientation, optimal for DNA binding, by the chirality of the rigid scaffold (Fig. 3.2B). The distance of ca. 7 Å between the C-3 and C-12 anchoring points of the steroid scaffold[97] is comparable to the radius of β-cyclodextrin (7.5 Å) used in the previously reported Morii model (Fig. 3.2C).[72] This distance also lies well within the range of distances measured between anchoring points in other previously reported models (Fig. 3.3). Direct attachment of the basic-region peptides to the steroid scaffold, as previously reported,[58] did not result in a DNA-binding construct. As can be derived from Figure 3.1, this can be ascribed to the steric clash between the DNA duplex and the scaffold, which interferes with the insertion of the peptides into the grooves of the DNA.

Previous designs with other dimerizing units used two extra glycines and a cysteine as a linker to avoid this problem.[69, 71, 75] We preferred to avoid the use of oxidation-prone cysteine, and therefore a γ-aminobutyric acid (Gaba) moiety was intro duced together with two glycines as a linker to provide the necessary flexibility. Moreover, the rotational freedom of the Gaba residue should allow the peptides to more easily adopt a conformation required for DNA binding.
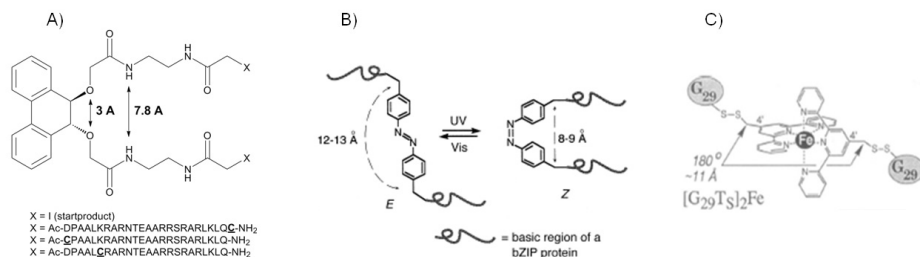
Figure 3.3: Representation of previously used models The choice for the C-3 and C-12 of deoxycholic acid as anchoring points was partly inspired by the comparable distance spanned by previously reported models. Other examples next to the Morii model (Fig. 3.2C), are given here. This illustrates that 7 Å lies well within the range of distances measured between anchoring points in previously reported models. A) Measured distance of 3 - 7.8 Å (using ChemBio3D) for the bridged biphenyl moiety reported by Morii [71]. B) Measured distance of 8-9 Å for the binding Z form of the azobenzene model reported by Mascarenas [74] (Figure taken from ref. [74]). C) Measured distance of 11 Å for the Fe-complex reported by Schepartz [73, 98] (Figure taken from ref. [73, 98]).

Additionally, an N-terminal cap, acetamido benzoic acid (Aba) was added to the N-terminus of the peptide. This aromatic moiety allows convenient UV-monitoring and concentration determination as an internal standard,[77] while efficiently blocking the N-terminus during the introduction of the second amino acid sequence (Fig. 3.4). We chose to immobilize earlier described and well-studied GCN4 basic-region peptides as DNA-binding moieties from the simple bZIP rather than b-HLH-ZIP TF family. This led to the optimized design of the target compound (**3.1.3**) shown in Figure 3.4, with two peptides consisting of a flexible linker, the DNA-binding basic-region peptide of GCN4, and an N-terminal cap.
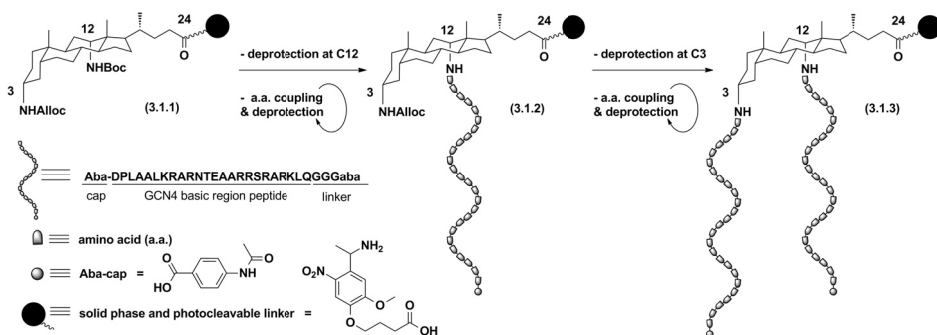
Figure 3.4: Synthesis approach for the generation of a dipodal peptidosteroid TF model (3.1.3).

Synthesis For the synthesis, several solid supports were evaluated, as it is known that the resin can have a major influence on the outcome of the synthesis of such large constructs. The most recent PEG polymers have been reported to efficiently reduce hydrophobic peptide aggregation through their increased hydrophilic character.[99–101] However, as we have shown earlier, with these more recently developed resins, monitoring of the completion of reactions can be compromised.[102] Therefore, we opted for the PEG-grafted polystyrene Tentagel support from Rapp.[103] A low loading (0.28 mmol/g) further ensures the accessibility of the growing peptide chains through a pseudodilution effect.[104]

The first step in the synthesis involved attachment of the orthogonally pro-tected C-3,C-12-diamino-modified steroid scaffold to the solid support by a Holmes photosensitive nitroveratryl-based linker.[105, 106] This linker was cho-sen as it allows semi-online monitoring of the reaction through a mild, fast, orthogonal, and non-laborious cleavage. Coupling and deprotection were car-ried out manually under standard Fmoc (9-fluorenylmethoxycarbonyl) solid-phase peptide synthesis (SPPS) conditions using PyBOP (benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate).

With the scaffold attached to the solid support, we went on to attach the GCN4 basic-region peptides (Fig. 3.4). Although the same peptide was to be introduced at both the C-3 and C-12 positions, attempted simultaneous synthesis of both peptides proved unsuccessful (see experimental part). As the C-12

position is sterically more hindered, reaction proceeds faster at the C-3 position, further inhibiting reaction at the C-12 position. Therefore the first peptide should be attached at the C-12 position before going on to functionalize the C-3 position. This differentiation is enabled by the Boc (tert-butoxycarbonyl) vs. Alloc (allyloxycarbonyl) selectivity. Furthermore, this orthogonality allows future mimicry of heterodimeric rather than homodimeric TF proteins. The amine at the C-3 position was carefully Boc-deprotected with TFA (trifluoroacetic acid; 20% solution), to ensure orthogonality with the photocleavable linker. The first residues (i.e., Gaba and Gly) were coupled manually under standard Fmoc SPPS conditions using PyBOP (4 equiv.), and the remainder of the sequence was coupled in automated fashion using HBTU (N,N,N',N'-tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate; 10 equiv.). The Aba capping was again carried out manually with PyBOP (10 equiv.) to give monopodal product (**3.1.2**) in high purity (Fig. 3.5).
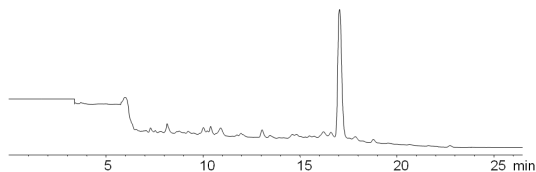
Figure 3.5: RP-HPLC chromatogram (C4, 300Å, 75-100% MeCN in 15 min) of the product (3.1.2)

The scaffold was then Alloc-deprotected at C-3 (to give **3.1.4**) for attachment of the second peptide. Previously, we reported that 1 equiv. is necessary, when using (apparently less active) commercially available catalyst. However, with freshly prepared tetrakis(triphenylphosphine)-palladium, this deprotection could be achieved using only a catalytic amount. The first amino acid coupling at the C-3 position (to give **3.1.5**) is difficult, due to steric crowding once a peptide is attached to the C-12 position (Fig. 3.6A). Acceptable conversions could only be achieved by double HATU (1-[bis(dimethylamino)-methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate) mediated coupling. A capping step was still required to fully convert the amino groups (Fig. 3.6B).

As shown in Figure 3.6C, conversions with other coupling reagents were worse. Coupling of the second amino acid again proceeded well with PyBOP. From the third residue onwards, coupling was carried out using automated synthesis equipment. The final product was obtained after a final manual capping step. The final photocleavage from the solid support had to be carried out before the acidic side-chain deprotection to avoid peptide loss, due to the observed sensitivity of the photocleavable linker to acidic conditions.
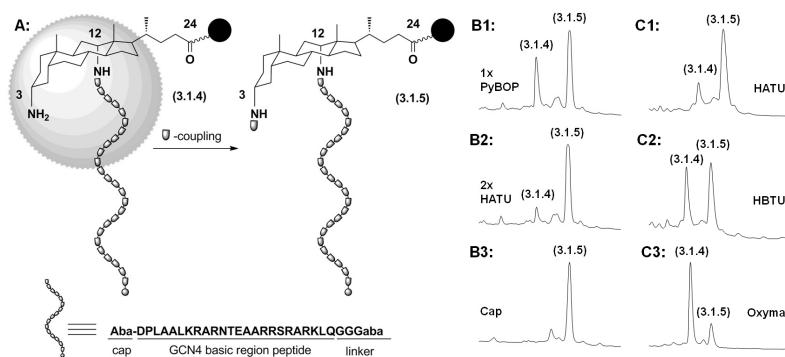
Figure 3.6: A) Coupling of the first amino acid (Gaba) of the second peptide at the C-3 position of (**3.1.4**) to obtain (**3.1.5**) is challenging. B) RP-HPLC chromatograms (C4, 300Å, 75-100% MeCN in 15 min) of the standard manual coupling conditions for the first residue of the second peptide at C-3. B1: using PyBOP: (**3.1.5**) was obtained in poor yield. B2: Subsequent double HATU coupling gave almost complete reaction. B3: The small amount of remaining (**3.1.4**) was capped. C) RP-HPLC chromatograms (C4, 300Å, 75-100% MeCN in 15 min) of double 10 min coupling with 4 equiv of different coupling reagents to obtain (**3.1.5**). C1: with HATU, C2: with HBTU, C3: with Oxyma.

The RP-HPLC chromatogram in Figure 3.7 shows that a mixture, rather than one major product, was obtained. LCMS analysis further revealed that no compound of the correct mass was present in the mixture. In view of our earlier reported successful synthesis,[58] and because linear 50-mer peptides are typically accessible by solid-phase peptide synthesis, failure of the synthesis was unexpected. The generation of two, protected 25-mer peptide chains in close proximity on a rigid steroid scaffold through flexible GabaGG linkers, thus seemed to be particularly sensitive to aggregation, which complicates the synthesis under the conditions reported earlier. Intra- and especially

intermolecular aggregation through hydrophobic forces can prevent access of
the reagents to the reactive sites, reducing the reaction rates of the coupling
and deprotection steps. The kinetic problems lead to incomplete reactions, and
thus result in mixtures of compounds with deletion sequences and prematurely
terminated sequences.[107–109] Detailed analysis of the intermediates revealed
that difficulties arose starting from the first residue of the second peptide,
which was sterically encumbered by the first peptide chain at C-12. Although
steric hindrance diminished as the peptide was extended, hydrophobic aggrega-
tion with the first peptide increased as the length of the second peptide increased.
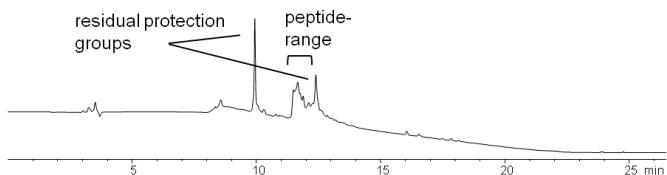
**Figure 3.7: RP-HPLC chromatogram (C18, 100Å, 0-100% MeCN in 15 min)
of the products obtained upon synthesis of (3.1.3) under conventional syn-
thesis conditions** following the reaction scheme shown in Figure 3.4.

Optimistically, the mixture obtained can be considered as a library of close
derivatives of the envisaged target model compound, and thus a preliminary
crude electrophoretic mobility shift assay (EMSA) was carried out (Fig. 3.8). As
the EMSA analysis showed a faint but clear indication of specific binding, we
were encouraged to further pursue the challenging synthesis of this new model,
optimized for DNA binding. Moreover, as failure in peptide synthesis is most
often due to steric hindrance and aggregation, like in this case, the dipodal
structure proposed here constitutes a good model system for further optimisation
of SPPS conditions. The production of high molecular weight peptidic drugs, in
which chemical synthesis plays an important role,[59] is still an important factor
that hinders their transfer into clinical use.

Lane	1	2	3	4	5	6	7	8	9	10
DNA		CREB (50 nM)					Random (50 nM)			
Peptide-concentration	/	75 nM	150 nM	225 nM	300 nM	/	75 nM	150 nM	225 nM	300 nM

mimic-DNA complex →

free DNA duplex →
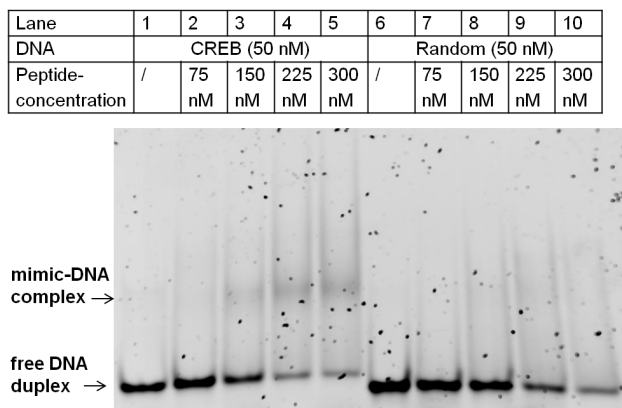
Figure 3.8: Preliminary EMSA experiment using CRE DNA in lanes 1-5 and random DNA in lanes 6-10 with the peptidosteroid (**3.1.3**) in concentrations 0 nM, 75 nM, 150 nM, 225 nM and 300 nM in respectively lanes 1 and 6, 2 and 7, 3 and 8, 4 and 9 and 5 and 10.

Completing and Improving the Synthesis In a first attempt to improve the synthetic yield and quality of the product, cleavage and deprotection conditions were carefully optimized.

For reaction monitoring and small-scale cleavage, a small lamp (1 mW/cm^2) was used with the reaction mixtures in small glass test tubes for 3 h. To maximize the yield in the final cleavage of the end product, a more powerful lamp was desired. We found that by using a 15 mW/cm^2 lamp irradiating for 1 h from both above and below, with the resin spread out as a monolayer in a Petri dish to minimize light scattering and shadowing effects, maximal cleavage yields could be obtained (optimisation described in detail in experimental part). With a stronger lamp or a longer irradiation time, side reactions started to occur.

Side-chain deprotection was then carried out using a cocktail very similar to the generally accepted but smelly reagent K (TFA, water, phenol, thioanisole, and ethanedithiol).[110, 111] Care was taken to avoid side reactions by adding the cooled cocktail to the peptide on ice. These conditions are drastically different from those described and used earlier, where cleavage mixture B (TFA, water, TIS (triisopropylsilane and phenol), including a non-smelly trialkylsilane, was

heated to 60°C to achieve complete deprotection. But under these conditions, a better result could be obtained (Fig. 3.9), and the target compound was present in the product mixture, as identified by LC-MS analysis.
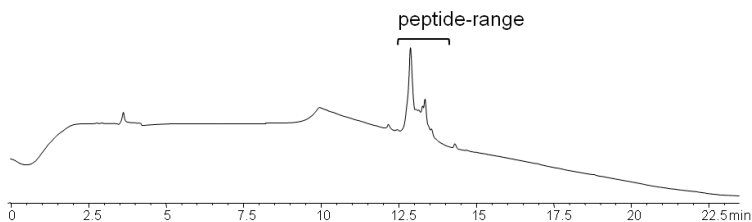
Figure 3.9: RP-HPLC chromatogram (C18, 100Å, 0-100% MeCN in 15 min) of the products obtained upon the synthesis of (3.1.3) under conventional synthesis conditions following the reaction scheme shown in Figure 3.4, but cleaved off under optimized and more mild conditions.

Further improvements were approached using microwave-assisted synthesis. Microwave irradiation was first introduced in peptide synthesis in 1992, using a modified kitchen microwave, and improved coupling yields were obtained at elevated temperatures.[112] It became an immensely popular technique, partly because of the controversial non-thermal special microwave-related phenomena.[109, 113] Today, supported by the excellent work of Kappe[108, 114, 115] through studies with accurate temperature control, it is possible to say that the improvements can be attributed to thermal effects. Still, microwave heating has favourable properties over conventional heating, including the absence of wall effects caused by an inverted temperature gradient, localized superheating, and its easy applicability to vortexed solid-phase reactions, where stirring bars should be avoided, and the processes preferably carried out in an automated fashion.[116, 117]

The current target molecule, with the challenging synthesis of the strongly aggregating second peptide, provides a good test system to explore the potential of microwave-assisted SPPS (MA SPPS), which could disrupt the aggregation in the peptide synthesis through thermal agitation. A first test consisted of the coupling of the first residue of this peptide, which required double HATU couplings under conventional conditions (Fig. 3.6). Using microwave irradiation, however, this coupling could be carried out with standard coupling reagents HBTU, PyBOP,

and DIC (N,N'-di-isopropylcarbodiimide)/HOBt (hydroxybenzotriazole) to give complete conversion, as was seen with HATU (Fig. 3.10).
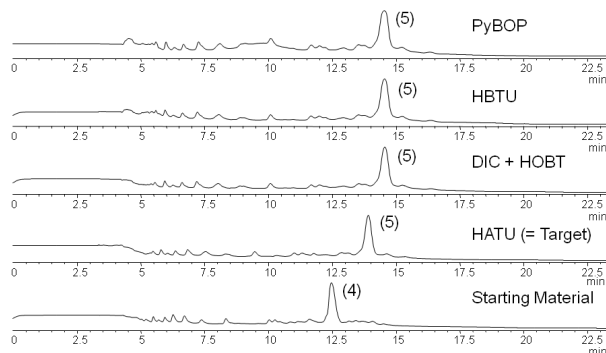
Figure 3.10: RP-HPLC chromatograms (C4, 300Å, 75-100% MeCN in 15 min) of (3.1.5) obtained by double couplings with 8 equiv of different coupling reagents at 75°C.

Encouraged by this result, the remainder of the second peptide was also synthesized under microwave-assisted conditions. The microwave-assisted synthesis gave considerable savings in time and cost, but also the final yield and purity of the product were increased (Fig. 3.11). The target peptide was shown to be the major product by LC-MS analysis. Thermal agitation was found to decrease the overall aggregation and to reduce the steric hindrance at the difficult to access reaction sites.
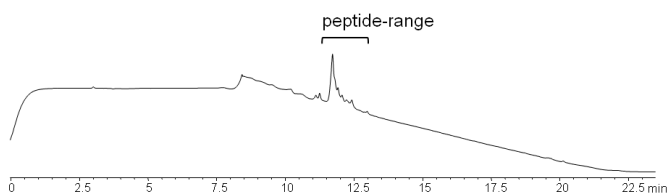
Figure 3.11: RP-HPLC chromatogram (C18, 100Å, 0-100% MeCN in 15 min) of the products obtained upon synthesis of (3.1.3) following the reaction scheme shown in Figure 3.4, with the synthesis of the second peptide under microwave-assisted conditions.

DNA binding achieved   The product obtained after purification (Fig. 3.12) was subjected to an EMSA to test for binding affinity to a DNA binding site of the original TF (CRE: 5'-ATGACGTCAT-3').[118] Figure 3.13 shows the results of the experiment. It is clear that the target molecule binds the DNA to form an upshifted complex with lower mobility. The observed smearing of the gel is probably caused by non-specific interactions with the DNA by impurities that co-eluted with the target molecule during RP-HPLC purification.[119] Indeed, with increasing peptide length, the relative differences between target and by-products decrease, while the number of variations exponentially increase. This causes strong overlap on HPLC, making purification extremely tedious. The superposition of both specific and unspecific binding modes (see experimental part) makes the present dose response between the concentration of the peptide and the peptide-DNA complex more difficult to observe, especially at high peptide concentrations. The binding capacity was further compared with the earlier reported GCN4 mimic developed by Morii [72] (Fig. 3.2C), of which an authentic sample was produced (see next section). It can be seen that our cholic-acid-based dipodal peptide shows an affinity for the target DNA that is comparable to and slightly higher than that of the earlier developed non-covalent dimer. This clearly shows the potential of our current design for synthetically accessible DNA major-groove binders.
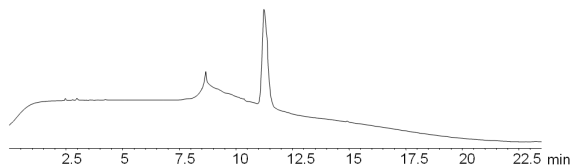
Figure 3.12: RP-HPLC chromatogram (C18, 100Å, 0-100% MeCN in 15 min) of the purified product (3.1.3)

lane	1	2	3	4	5	6	7	8	9	10
model	Peptidosteroid model (µM)					Morii model (µM)				
conc.	0	0.30	0.45	0.60	0.75	0	0.30	0.45	0.60	0.75
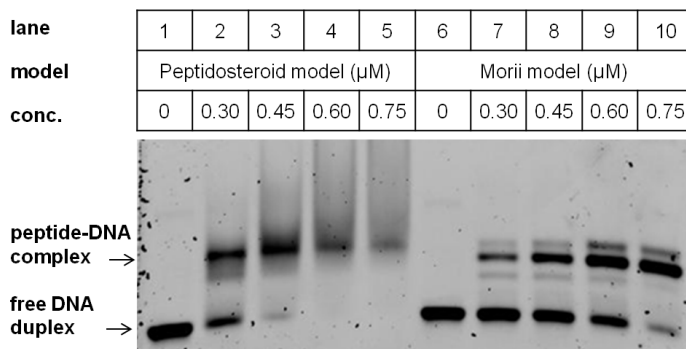
Figure 3.13: EMSA titration (8% acrylamide gel, with 0.2% TBE buffer) of the dipodal peptidosteroid (left, see Fig. 3.2B) and Morii mimic (right, see Fig. 3.2C) to the CRE binding site. Lane 1 and 6, 2 and 7, 3 and 8, 4 and 9, 5 and 10 contain respectively 0, 0.3, 0.45, 0.6 and 0.75 µM of the mimic (**3.1.3**) with 6% sucrose, 100 mM Tris pH 7.6, 20 mM KCl, 10 mM MgCl$_2$, 2.5 mM EDTA and 0.167 µM DNA. Gels were stained with SybrGold and visualized under UV light.

Conclusion   The synthesis of closely immobilized peptides is challenging, due to aggregation of the peptides during synthesis. This is particularly so in the current case, with a rigid amphiphilic steroid moiety. The flexibility required for DNA binding provided by the insertion of Gaba-GG spacers led to a dramatic increase in the steric entanglement of the protected peptides, and thus precluded successful conventional synthesis. Microwave irradiation allowed to reduce these interactions, resulting in a successful synthesis of the dipodal peptidosteroid with increased purity and yield using standard coupling reagents, and additionally in greatly reduced synthesis time. It was further shown that improvements in the yield and purity could be obtained by altering the cleavage conditions. As such, this investigation contributes to the understanding and improvement of SPPS methods for difficult sequences, where microwave irradiation can clearly be beneficial. Using these optimised SPPS conditions, we were able to produce an optimised, more flexible peptidosteroid tweezer as a TF model with demonstrated DNA-binding properties.

Our next efforts will focus on a more modular and convergent synthesis to obtained this now qualified transcription factor model. The synthesis and pu-

rification of smaller peptides is more cost-efficient and will greatly simplify the effort.[58, 120] This is the desired approach to obtain compounds in the amounts and purities required for the further development and validation of the current model system.
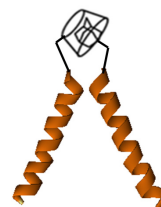
3.2 Morii mimic: Ad-CD inclusion complex

The synthetically accessible, DNA binding transcription factor model is an ideal system for modification with a cross-linking moiety as outlined earlier. However, because the optimization of synthesis and design for the peptidosteroid model was an ongoing process during the thesis, an alternative was desired for use in these experiments. The previous section already mentioned and depicted the adamantane-cyclodextrin inclusion complex based model described by Morii and coworkers. An additional advantage of this model is its non-covalent nature, which allows a reduction in size by half after cross-linking under denaturating conditions as illustrated in Figure 3.14 and simplifies its analysis and characterization. This section will give some additional and more detailed information on this model and its synthesis.
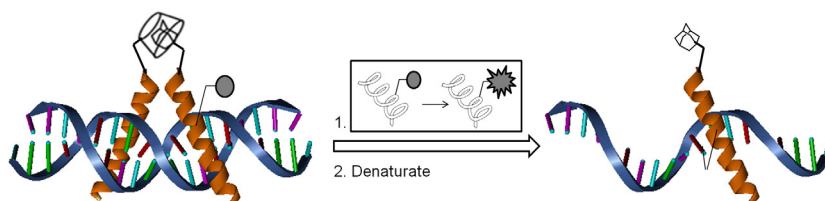
Figure 3.14: Advantage of the non-covalent nature of the Morii mimic After cross-linking the complex can be denatured to reduce the size by half for more straightforward analysis.

General The Morii group introduced the non-covalent mimic based on the adamantane (Ad)-cyclodextrin (CD) inclusion complex in 1993,[72] after the introduction of a covalent mimic based on a bridged biphenyl moiety.[121] Different members of the bZIP transcription factor family were miniaturized, GCN4[72],

EmBP1[122] and C/EBP[123]. Their DNA binding peptides could be synthesized with Ad and/or CD and mixed into Ad-CD heterodimers that preferentially bound to DNA sequences with correspondingly mixed binding half sites.[28] The synthesis of the constructs was mainly pursued to address the question of how DNA binding proteins work as dimers and what the significance of the dimer formation was. The inclusion complex allowed to test the effect of different guests (Fig. 3.15) or thus different affinities for dimerization.[124]
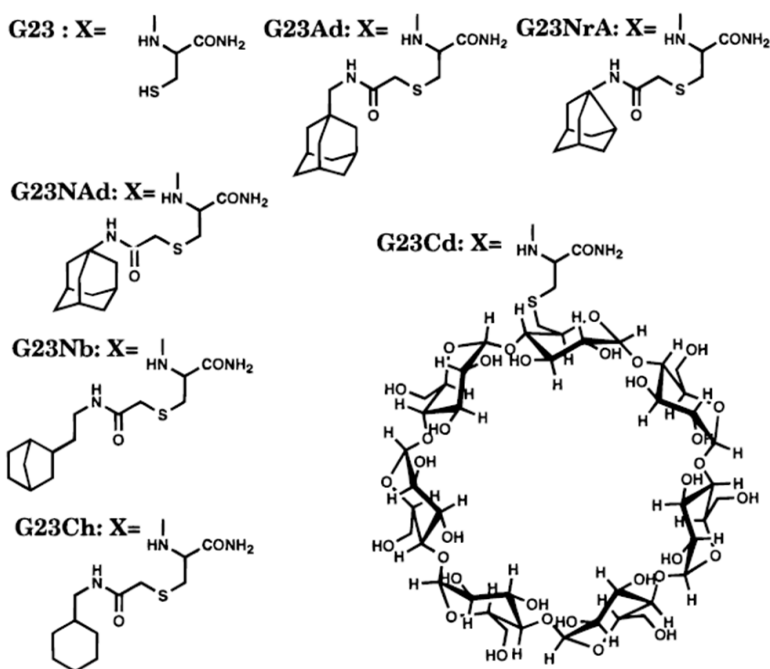
Figure 3.15: Different guests for the formation of peptide dimers through an inclusion complex Amino acid sequence for the peptide (G23) and structures of the modified C-terminal Cys residue (X). The cysteine residues were modified with 6-deoxy-6-iodo-β-cyclodextrin for G23Cd and with bromoacetyl derivatives of adamantanemethylamine, aminoadamantane, aminonoradamantane, aminoethylnorbornane and cyclohexylamine for the guest molecules G23Ad, G23NAd, G23NrA, G23Nb, and G23Ch respectively. Taken from [124]

When the host and guest molecule were each attached to a terminus of the peptide, cooperative oligomers could be formed on repeating DNA binding half sites as illustrated in Figure 3.16.[125]
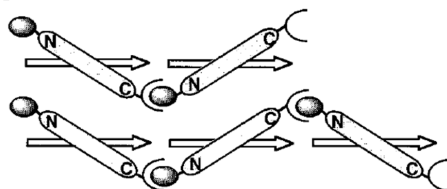
Figure 3.16: Schematic representations showing the Ad-CD homooligomer bound on the direct-repeat sequence of the DNA binding half-site. Bars represent the basic region peptide. N and C indicate the N-terminal and C-terminal of the peptide, respectively. Half-circles at the C-terminal and filled ovals at the N-terminal represent β-cyclodextrin and the adamantane group, respectively. White arrows denote the DNA binding half-site. Taken from [125]

Synthesis The synthesis of the non-covalent model based on the Ad-CD inclusion complex and from hereon now named the Morii mimic, was resynthesized under the supervision of Prof. Morii in his lab.[72]

Mono-6-deoxy-6-iodo-β-cyclodextrin (**3.2.3**) for modification of the cysteine thiol functionality in the peptides, was obtained by mono-tosylation of a primary hydroxyl functionality of the cyclodextrin followed by S_N2 substitution with NaI (Fig. 3.17).

(3.2.1)   (3.2.2)   (3.2.3)
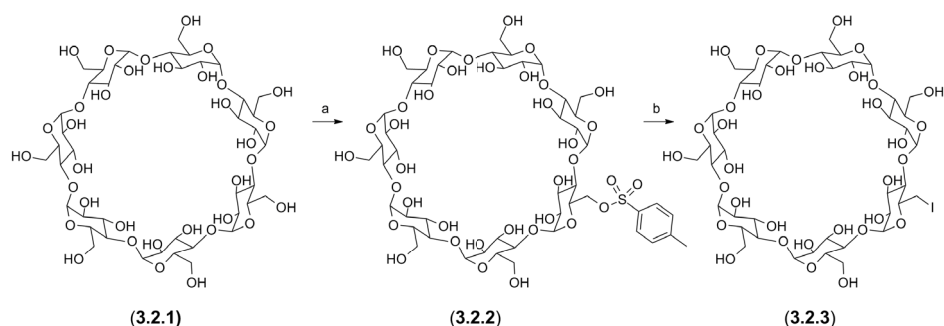
Figure 3.17: Modification of β-cyclodextrin a. 0.5 equiv TosCl in dry pyridine at rt for 2h, 45%; b. 10 equiv NaI in DMF at 70°C for 2h, 40%.

N-(bromoacetyl)-1-adamantanemethylamine (**3.2.6**) for the modification of the cysteine thiol functionality in the peptides, was obtained by activation of bromoacetic acid as succinimide ester for reaction with adamantanemethylamine (Fig. 3.18).

(3.2.4)   (3.2.5)   (3.2.6)
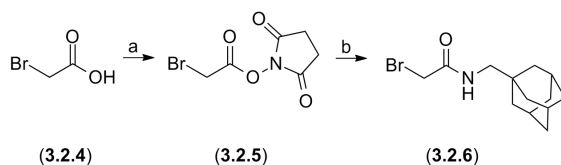
Figure 3.18: Modification of adamantane a. DCC coupling (1.1 equiv) with hydroxysuccinimide (1 equiv) in dry EtOAc at 0°C for 2h, 66%; b. 0.33 equiv adamantanemethylamine, 0.66 equiv HOBt in dry DMF at 0°C, allowed to heat up to rt overnight, 41%.

Peptide synthesis was performed under standard conditions for long peptides, using 10 equivalents of HBTU with HOBT as coupling reagent. Cleavage was performed as described in the previous section under optimized conditions. The obtained GCN4 peptides (**3.2.7**) were purified by RP-HPLC. Contrary to the peptide depicted in Figure 3.15, a double glycine linker was added between the basic region peptide and the cysteine, to also allow for dimers through a disulfide bond as described by Kim.[69]

The purified peptides where reacted with the activated adamantane to obtain Ad-GCN4 (**3.2.8**) or with the cyclodextrin moiety to obtain CD-GCN4 (**3.2.9**) (Fig. 3.19) in phosphate buffer at pH 9 under inert conditions to avoid dissulfide formation as much as possible. For modification with activated adamantane, it was found that the formation of side products could be minimized by reducing the reaction time from the reported 1 h to 5 min on ice. Modification with the less activated cyclodextrin was more difficult and required 7 h at room temperature, after which all remaining starting material had been oxidized to form disulfides. The latter were reduced and reintroduced in the reaction to increase the yield. The modified peptides were repurified on RP-HPLC. Upon combination of both Ad- and CD-GCN4, an inclusion complex gives rise to the Morii mimic (**3.2.10**).
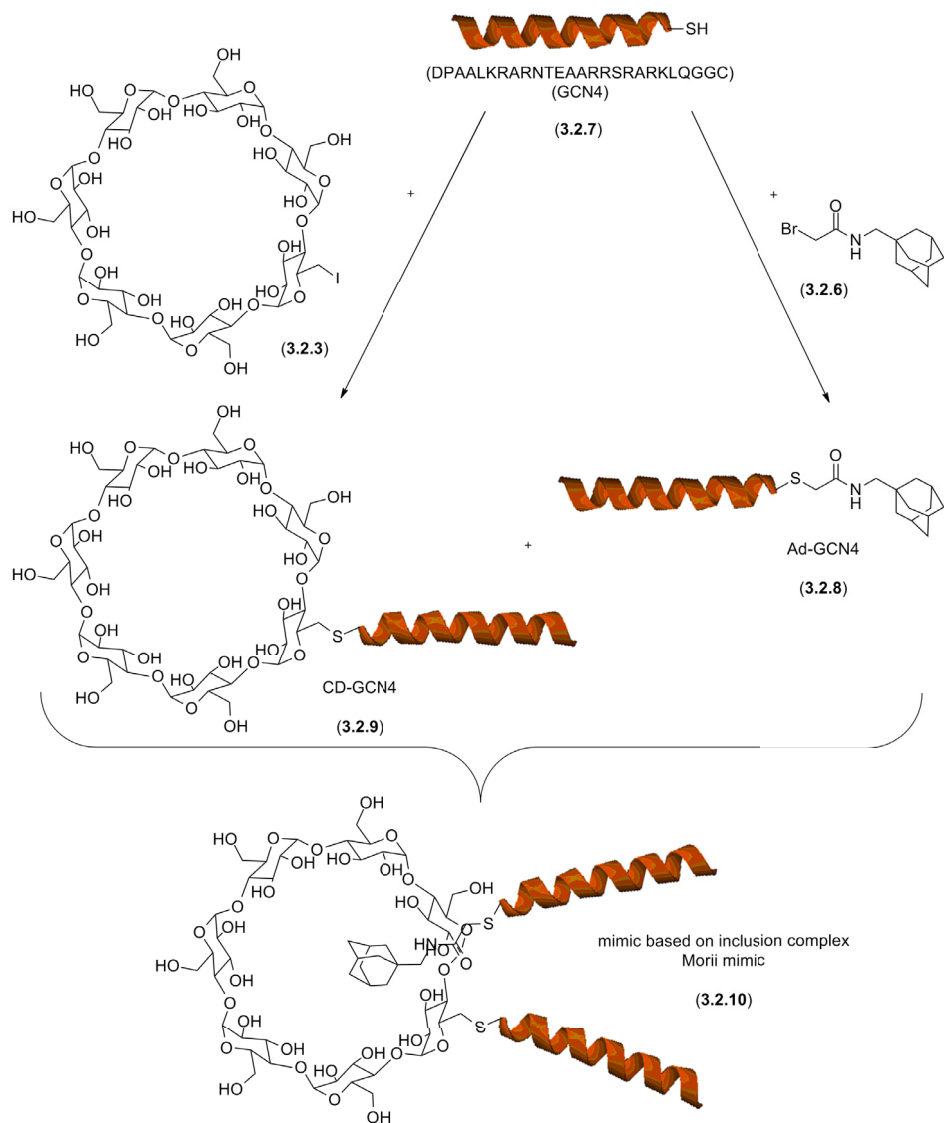
Figure 3.19: Modification of the GCN4 peptides and formation of the Morii mimic through an inclusion complex

DNA binding The DNA binding of the Morii mimic was already shown before in comparison with the Madder mimic (Fig. 3.13). Binding was also quantified by EMSA with ^{32}P-labeled AP1 DNA, which is the natural pseudopalindromic binding site of GCN4 protein. The radiolabelled DNA allows for very sensitive radiodetection. Therefore only minimal amounts of DNA are required and the equation for the dissociation constant (K_D) can be simplified, to equal the concentration at half-binding. Judging from the result in Figure 3.20 the K_D is about 150 nM. This is higher than the reported value of 30 nM for the palindromic CRE binding site,[72] due to the preference of the Morii mimic for the latter binding site.[124]

lane	1	2	3	4	5	6	7	8	9	10
c (nM)	0	1	2	5	10	20	50	100	200	500
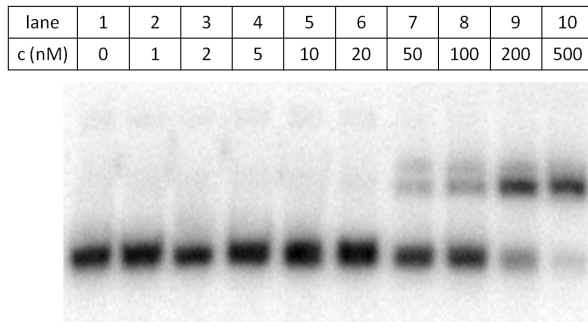
Figure 3.20: EMSA titration of the Morii mimic to ^{32}P-labeled AP1 binding site (8% acrylamide gel, with 0.2% TBE buffer). Lane 1 to 10 contain respectively 0, 1, 2, 5, 10, 20, 50, 100, 200, 500 nM of the mimic (**3.2.10**) with 6% sucrose, 100 mM Tris pH 7.6, 20 mM KCl, 10 mM MgCl$_2$, 2.5 mM EDTA and 4.88 pM DNA.

Chapter 4

Furan based Cross-linking
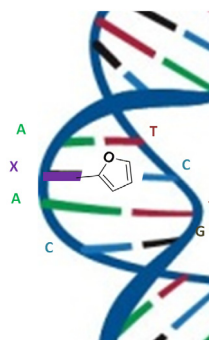
4.1 DNA interstrand cross-linking

4.1.1 The methodology

The cross-linking methodology described here is inspired by the natural toxicity of furan. Furan has been classified by the International Agency for Research on Cancer (IARC) as a possible human carcinogen (class 2B), based on several reports of toxicity in laboratory rats and mice. Furan is present in the environment, as a contaminant from incomplete combustion, as well as in food, especially after heat treatment.[126] Once in the body, furan is metabolized by cytochrome P450 enzymes to butenedial, which reacts quickly with nucleophiles in proteins and nucleic acids,[127] resulting in the observed toxicity.

To study DNA-damaging processes such as furan oxidation, and especially the subsequent endogenous DNA repair mechanism, well-defined model systems are required. As extraction from biological systems is often troublesome, much effort has been devoted to synthesizing model substrates.[128] DNA ICLs are an important form of DNA damage, not because of their abundance (they account for less than 10% of all DNA damage), but because they represent the most toxic lesions.[129] Furthermore, ICLs and their repair are of interest in the context of chemotherapeutics and increasing resistance to them.

ICLs can be generated by addition of bifunctional cross-linkers (N-mustards, mitomycin C, platinum) to double-stranded DNA[130], but selectivity in this approach can be problematic. Generation of ICLs by incorporation of a linked dinucleoside during solid-phase synthesis[131] or by postsynthetic orthogonal reaction between two modified ODNs through disulfide formation[132] and Huisgen click reaction[133] has also been described. Because the latter requires modification of both strands in the desired duplex, use with natural DNAs for therapeutic applications (e.g., antisense, antigene, plasmids, aptamers, RNA interference, alternative splicing, decoy strategies), diagnostics (SNP detection), or structural studies is not possible.

To allow reaction with unmodified DNA strands, a reactive functionality is required. However, because the incorporation and handling of such modified reactive ODNs are tricky, and premature reactions can occur, latent functionalities are desirable. In this approach, furan was used for this purpose. Reminiscent to what happens during in vivo furan metabolization, selective oxidation transforms the furan into a reactive enal functionality, which allows it to react with nearby nucleophiles (e.g., the exocyclic amine of cytidine) and form a cross-link. Figure 4.1 illustrates the reaction mechanism for the formation of the stable cross-linked ODN species.
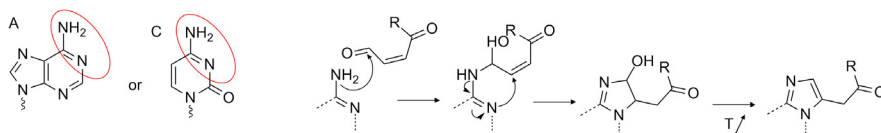
Figure 4.1: Reaction mechanism for formation of a stable cross-linked ODN.

In view of the general stability of the furan moiety before oxidation, handling during synthesis and binding events is straightforward and no special protection strategies or precautions need to be considered. This has led to the synthesis and evaluation of several alternative furan-modified ODNs (Fig. 4.2), allowing the finetuning of yield and selectivity.[134–139]
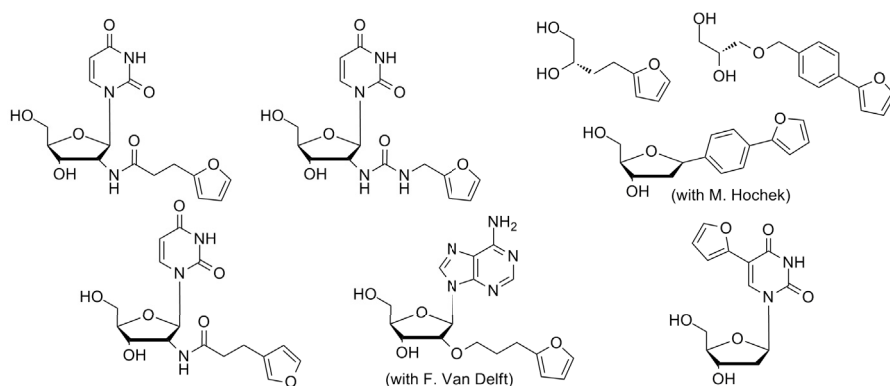
Figure 4.2: Summary of all the furan-modified nucleosides synthesized and tested for ICL formation in our group

*This was the background information from a protocol we wrote on our established furan based method for DNA interstrand cross-linking and forms the basis of the further investigations, which are the topic of this thesis. The introduction of the protocol, describing the methodology is furthermore given below. The protocols themself give a detailed description on the synthesis and use of two furan-modified nucleoside analogs. Basic Protocol 1 describes the synthesis of furan-modified 2'-amino-2'-deoxyuridine (**4.1.1 or nu1**), which provides selective cross-linking with complementary C bases (up to 30% isolated yield). The Alternate Protocol describes synthesis of a more simple and flexible acyclic furan-containing nucleoside analog (**4.1.2 or nu2**), which is less selective, but generates high quantities of ICL duplex species (up to 70% isolated yield). Incorporation of both modified phosphoramidites in oligonucleotides is described in Basic Protocol 2, and cross-link formation to an unmodified target DNA strand is described in Basic Protocol 3. As these protocols are lengthy and technical, they are not given*

here but are available at the end of the manuscript as Appendix.

DNA Interstrand Cross-Link Formation Using Furan as a Masked Reactive Aldehyde

UNIT 5.12

Lieselot L.G. Carrette,[1] Ellen Gyssels,[1] and Annemieke Madder[1]

[1]Organic and Biomimetic Chemistry Research Group, Department of Organic Chemistry, Ghent University, Krijgslaan, Ghent, Belgium

ABSTRACT

This unit describes a method for interstrand cross-linking between a furan-modified oligonucleotide and its unmodified complement. The synthesis of two furan-modified phosphoramidites, selected based on high cross-linking yield versus improved cross-linking selectivity, is described. The methods allow gram-scale synthesis starting from stable and readily available furan derivatives. Cross-linking requires selective oxidation of the furan moiety to an aldehyde. The masked nature of the latter avoids undesired and off-target reactions, resulting in clean and high-yield cross-link formation. *Curr. Protoc. Nucleic Acid Chem.* 54:5.12.1-5.12.16. © 2013 by John Wiley & Sons, Inc.

Keywords: cross-linking • oligonucleotide • nucleoside • furan

Methods for Cross-Linking Nucleic Acids

5.12.1

Supplement 54

This unit describes the use of furan-modified oligodeoxynucleotides (ODNs) for cross-linking with the complementary nucleic acid strand to achieve site-selective interstrand cross-links (ICL). Once incorporated into an ODN, the furan moiety is selectively oxidized to a reactive 4-oxobutenal. If positioned on the Watson-Crick base pairing interface in the DNA duplex, it will quickly react with the exocyclic amine functionality of a complementary base on the opposite unmodified strand, generating an ICL (Fig. 4.3). As the reactive aldehyde functionality is only uncovered on triggered demand (for example, after hybridization), nonspecific and off-target reactions are quasi excluded. Essential in the cross-linking protocol is that the obtained aldehyde reacts with an unmodified complement, allowing ICL formation with natural DNA target sequences. The cross-linked species can be isolated in good yields and are very stable.

The furan moiety is readily incorporated into the necessary phosphoramidite building blocks and does not interfere with solid-phase ODN synthesis. Given the stability and commercial availability of a series of furan derivatives, the possibilities for generating furan-modified analogs are numerous, and a range of different furan-modified nucleosides have already been synthesized and described.[134–139] Cross-link yield and selectivity are strongly dependent on the exact structure of the furan-modified nucleoside. In addition, there is a trade-off between duplex stability and nucleobase availability on the one hand and cross-link yield and selectivity on the other. Using rather simple but flexible (and thus duplex destabilizing) furan-modified nucleoside analogs like **nu1**, high yields of ICL duplex can be obtained, with cross-linking to both complementary A and C bases. On the other hand, when the furan moiety is introduced onto the more classical nucleoside building block like **nu2**, cross-linking is completely selective for complementary C bases.
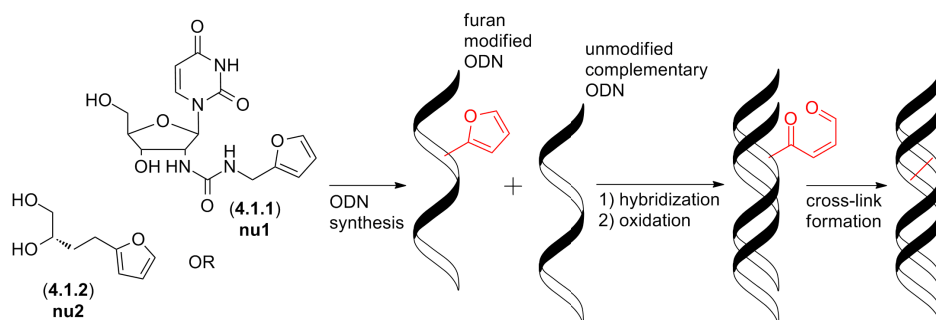
Figure 4.3: Overview of furan-based interstrand cross-linking methodology. Furan-modified nucleosides are converted to the corresponding phosphoramidites and incorporated into ODNs under standard conditions. After hybridization with a suitable complement, the furan-modified ODN is oxidized to generate a reactive 4-oxobutenal, which reacts efficiently with the nucleophilic exocyclic amine of the opposite base.

4.1.2   Case study: Kinetin

Kinetin (N^6-furfuryladenine) is obtained after base excision repair of furfuryladenine containing DNA. It has been associated with beneficial anti-aging effects and is therefore being studied for the treatment of cancer and degenerative diseases. In the context of our previous work on furan-modified oligonucleotides for the formation of interstrand cross-links, known to be very toxic lesions, we were intrigued by the absence of negative effects of the kinetin modification in DNA. Therefore, kinetin-modified oligonucleotides were synthesized, through a new, easy, post-synthetic approach and tested for duplex cross-linking under our previously reported conditions. We here show that indeed oxidation to a reactive intermediate takes place, which however degrades without cross-link formation. A tentative explanation for this observed difference between the naturally occurring kinetin-containing DNA and our previously synthesized furan-modified DNA is presented.
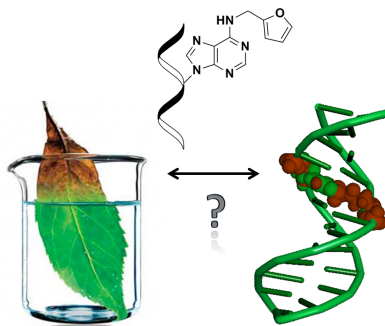
Figure 4.4: Kinetin originates from furan-modified DNA, which in principle could form ICLs, but has beneficial anti-aging effects

Kinetin modifications are formed under oxidative stress and are here shown to protect against this stress through their rapid oxidation. This finding contributes to the general understanding of the mechanism by which kinetin can exert its beneficial effects and can be of interest to a large community, in view of the importance and application potential for human health. Furthermore the fact that furan oxidation in kinetin-modified DNA does not lead to toxic DNA inter-

strand cross-links, underscores the importance of the positioning of the furan moiety in our previously described methodology. On the other hand, it reveals the potential of our furan based cross-linking to target cells under high oxidative pressure (for example cancerous cells), as these conditions can trigger furan oxidation. Such experiments will be the subject of our continued investigation.

A Synthetic Oligonucleotide Model for Evaluating the Oxidation and Crosslinking Propensies of Natural Furan-Modified DNA

Lieselot L. G. Carrette* and Annemieke Madder*[a]

We have previously developed a crosslinking methodology for oligonucleotides based on the incorporation of furan moieties, which can be selectively oxidised to reactive intermediates that will quickly react with the opposite bases in DNA, forming toxic interstrand crosslinks (ICLs). Furan moieties also occur in natural DNA, as a result of oxidative stress. Moreover, the furan-containing degradation product of this modified DNA—kinetin—has been found to display beneficial anti-ageing effects. To investigate the apparent discrepancy between the effects of the synthetic and the natural furan modifications in DNA, a quick and easy postsynthetic method providing access to the natural modification in short synthetic oligonucleotides was developed. On checking for potential crosslinking propensity, we found that the furan moiety does indeed undergo oxidation, in this way functioning as an important scavenger for oxidative stress. The reactive intermediate, however, was shown to degrade without producing toxic crosslinked products.

Introduction Furan moieties are formed upon degradation of natural DNA and can further react by addition to nucleic acid bases. The presence of furfuryl-modified adenine in cellular DNA[140, 141] has been detected in almost all organisms, including humans.[142] As illustrated in Figure 4.5, these modifications are formed under oxidative stress, characterized by the presence of increased levels of reactive oxygen species (ROS) and responsible for aging and a number of degenerative diseases including cancer.[143] Hydroxyl radicals attack the DNA deoxyribose components at C5', which leads to the spontaneous cleavage of the DNA backbone and β-elimination, giving rise to furfural as a marker of oxidative damage. Furfural can react with the exocyclic amine group of adenine in DNA to form a Shiff base, which is further reduced by an unknown mechanism to yield a stable furan modification in the DNA. The modified base can then be cut from the DNA by repair enzymes; this generates kinetin (N[6]-furfuryladenine) and an

abasic site, which is further repaired.[144]
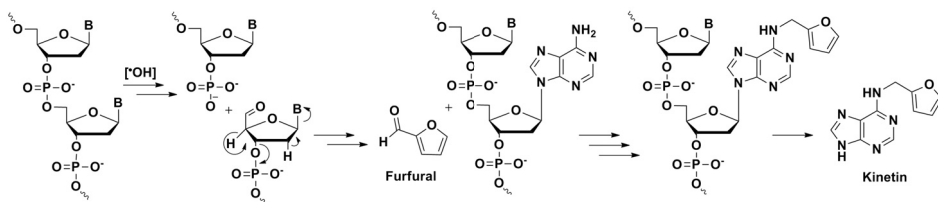
Figure 4.5: Kinetin biosynthesis

Kinetin was the first discovered cytokine plant hormone. Cytokines are products of oxidative metabolism of the cell and thus biomarkers of oxidative stress. They typically are N^6-substituted adenine derivatives that control many processes in plants.[145] Kinetin promotes cell division in plants and has been associated with several anti-aging effects such as retardation of leaf senescence and modulation of the environmental stress response.[146] In insects, kinetin can delay development and prolong the lifespan[147] and also in human fibroblast cells, kinetin was found to delay aging characteristics.[148] A detailed discussion of all effects of kinetin is beyond the scope of current report, but has been extensively reviewed by others.[145, 149]

As a result of kinetin's observed beneficial effects, it is currently being used in the treatment of dermatological problems[150] and tested in relation to neurodegenerative diseases.[151] It is further also being investigated as a chemotherapeutic.[152, 153] Importantly, unlike several other anti-ageing factors, kinetin does not promote carcinogenesis. The increased lifespan in cell cultures is not associated with proliferative activity. Kinetin's actual mechanism of action is still unknown, but it is clear that kinetin functions as antioxidant and is able to protect proteins[154] and DNA[155] from oxidative damage. Different hypotheses for its action exist; they include the prevention of radical formation by participation in complex formation with iron[155] and scavenging properties by for example free radical mechanisms.[156] It further induces maintenance and repair mechanisms by forming a bulky lesion in the DNA.[145]

We have previously reported the use of singlet oxygen, an ROS, for the oxidation of a furan moiety in synthetic, modified oligonucleotides to generate DNA interstrand cross-links (ICL) in a high-yielding and selective manner.[139] Furan oxidation by NBS[134–138, 157] or ROS such as singlet oxygen[139] results in a reactive 4-oxobutenal[127] intermediate that will quickly react with the exocyclic amines of the opposite base, resulting in an ICL as shown schematically in Figure 4.6. Hydroxyl radicals, responsible for the formation of kinetin have also been reported to oxidize furan.[158] ROS are thus scavenged both during the formation of the kinetin moiety as well as during the furan oxidation. Such antioxidant behavior has been described for naturally occurring antioxidant furan fatty acids (F acids), which were shown to scavenge hydroxyl radicals by oxidation to diketones.[159] The 4-oxobutenal obtained in case of 2-modified furans like kinetin, however is more reactive and could, similar to our synthetic furan-modified oligonucleotides cause ICL, which are considered very toxic DNA lesions.
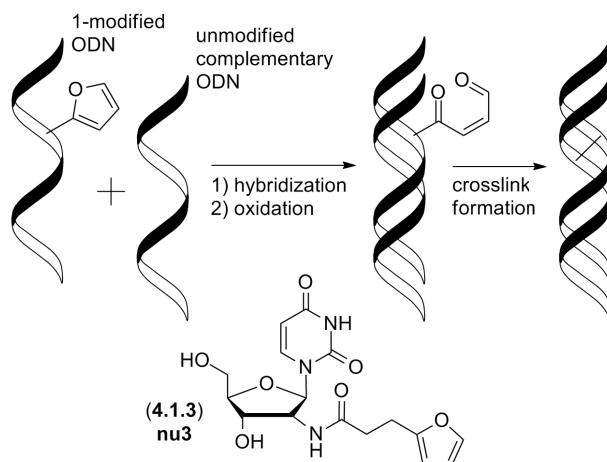
Figure 4.6: Outline of the earlier described furan oxidation cross-link methodology Furan-modified nucleosides (for example **1**) are synthesized and upon conversion into the corresponding phosphoramidites incorporated into ODNs, using standard ODN synthesis conditions. After hybridization with a suitable complement, the furan-modified ODN is oxidized to generate a reactive 4-oxobutenal, which will efficiently react with the nucleophilic exocyclic amine functionality of the opposite base.

To check the oxidation and subsequent cross-linking propensity of the naturally occurring furan-modified oligonucleotide precursors of kinetin, they were synthesized through a new fast synthesis protocol and oxidized according to our standardized procedures.[157]

Results and Discussion

Synthesis The synthesis of kinetin-modified oligonucleotides has been described before, but clear protocols remain vague. As the kinetin modification is not very stable under the acidic conditions of dimethoxytrityl removal, a post-synthetic oligonucleotide modification approach is desirable.[160] The modification of inosine with furfurylalcohol has been mentioned,[143] as well as the reaction of furfurylamine with a 1,2,4-triazol-4-yl-modified adenosine,[160] which needs to be synthesized.[161]

To avoid this need for synthesis, the commercially available activated building block O^6-phenyl-deoxyinosine-cyanoethyl phosphoramidite was used in this study (Fig. 4.7). It fulfills the role of convertible dA and can be handled and coupled in exactly the same way as standard commercial phosphoramidites. It was first described by Verdine in his study for disulfide cross-linked oligonucleotides.[162, 163] After synthesis, the oligonucleotide was cleaved from the controlled pore glass by 4 h treatment with concentrated NH_3 solution. This further led to a partial deprotection (benzoyl groups of adenosine and cytidine, isobutyryl group of guanosine) of the DNA (see Fig. 4.8). The obtained mixture was DMTr-deprotected, purified over Sep-Pak into 5 fractions (see Fig. 4.9) of which the first fraction was finally treated with 1 M aqueous furfurylamine at 55°C, to obtain complete deprotection and furfurylamine modification of the activated inosine (**4.1.4** or **nu4**). There is an optimal treatment time with furfurylamine for the partially deprotected oligonucleotide. If treatment is too short, residual protection groups remain and/or there is incomplete conversion of the inosine. Upon prolonged reaction times formation of a side product occurs, as can be derived from the peak splitting in Figure 4.10. The optimal time is between 30 and 40 h (Fig. 4.11 and Fig. 4.12). The procedure should thus best be monitored by HPLC. The identity of the obtained ODNs was confirmed by mass analysis.
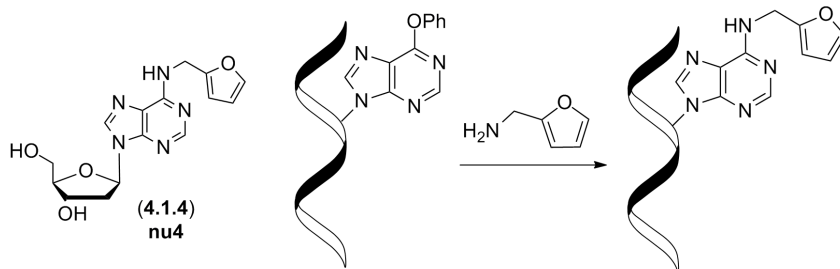
Figure 4.7: Post-synthetic modification to obtain a furfuryl-modified synthetic oligonucleotide
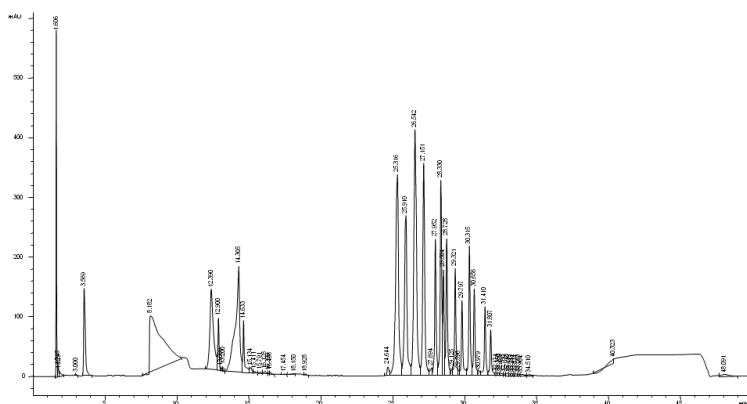
Figure 4.8: HPLC chromatogram after aqueous NH$_3$ treatment, before Sep-Pak of ODN4.1.1, a test oligonucleotide
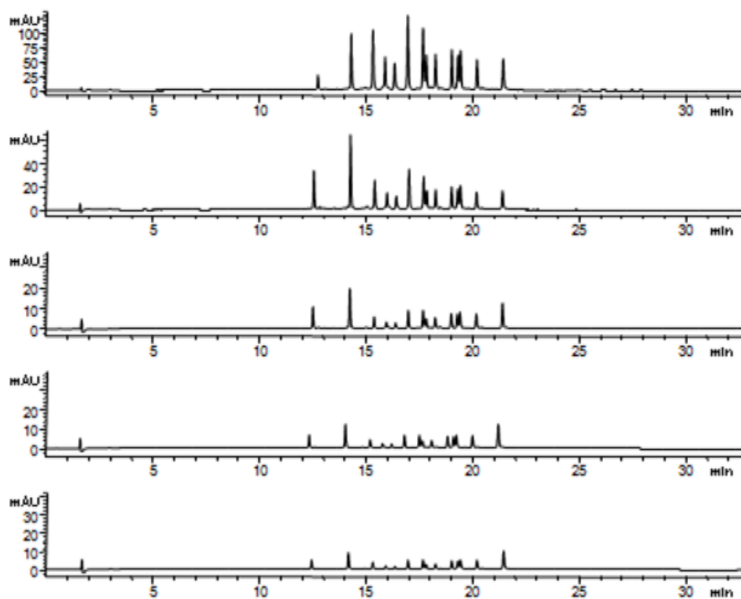
Figure 4.9: HPLC chromatograms of the 5 Sep-Pak fractions of a test oligonucleotide, ODN4.1.1,
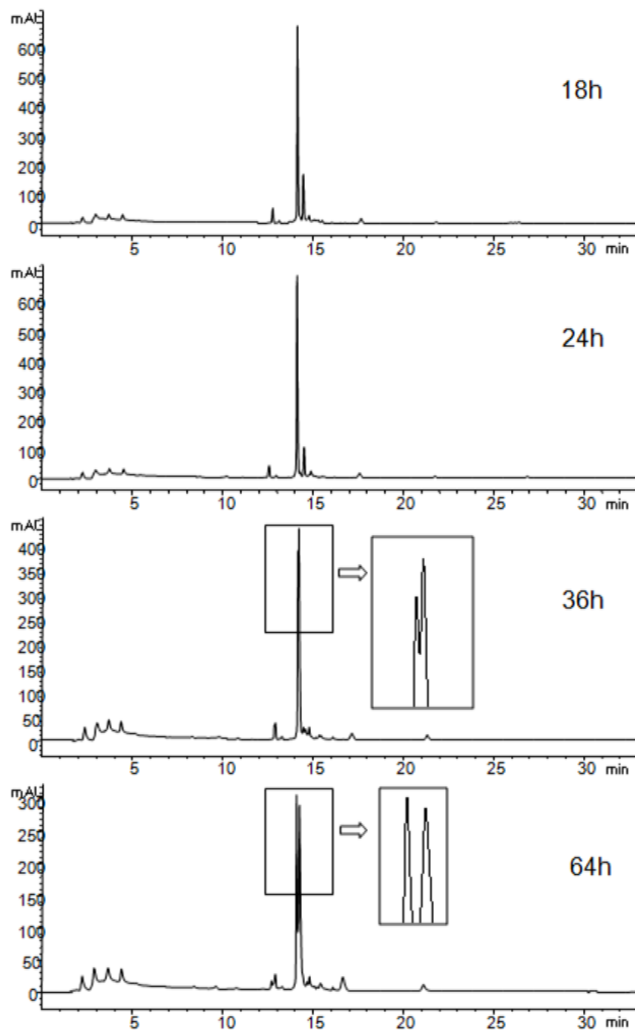
Figure 4.10: HPLC chromatogram after 18h/24h/36h/64h of furfuryl-amine treatment, with zooms on the peaks in the last 2 chromatograms of ODN4.1.1, a test oligonucleotide
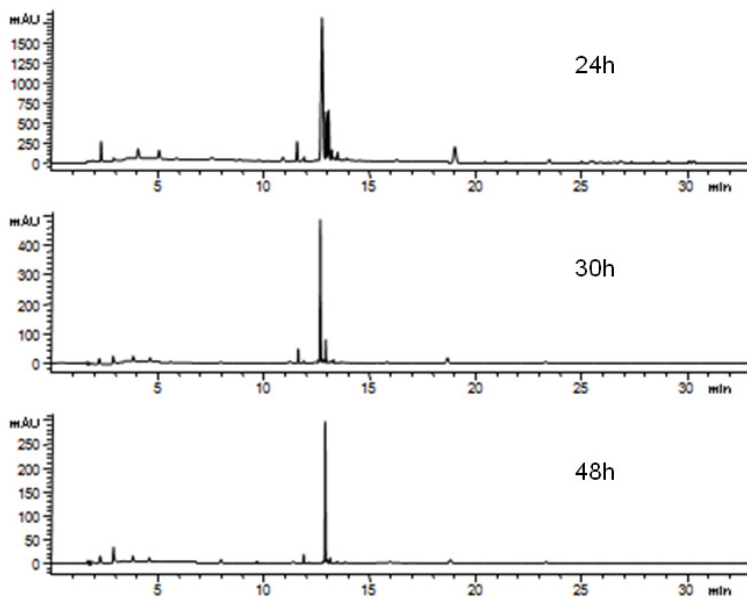
Figure 4.11: HPLC monitoring of the furfurylamine modification reaction (24h/30h/38h) of ODN4.1.1, a test oligonucleotide
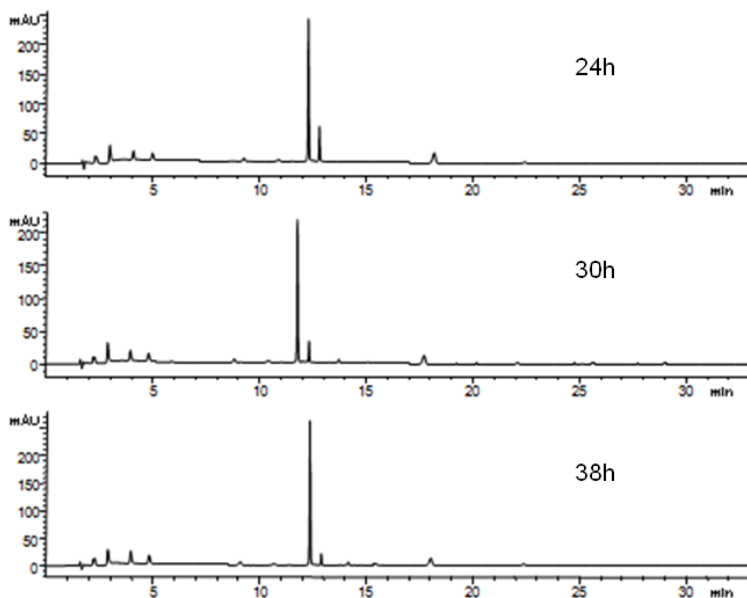
Figure 4.12: HPLC monitoring of the furfurylamine modification reaction (24h/30h/38h) of ODN4.1.2

Analysis of kinetin-modified duplexes To test for oxidation and ICL formation, the previously described sequence used to develop our furan oxidation ICL methodology was synthesized, but with replacement of nucleoside **nu3** (ODN4.1.4) with the kinetin modification (**nu4**) (ODN4.1.2, see Table 4.1). ODN4.1.2 was combined with its complement, varying the base opposite to the kinetin modification (ODN1-4). The unmodified analog of ODN4.1.2 (ODN4.1.3) was used for comparison. The effect of the kinetin modification on duplex stability was examined by measuring the melting profiles, summarized in Table 4.1. Clearly, introduction of the kinetin modification destabilizes the A-T basepair. This can be ascribed to the furan substituent on N^6 preventing the formation of a Watson-Crick hydrogen bond.[164] On the other hand, A-A and A-C mismatches are stabilized with 4 and 5 degrees respectively upon introduction of the furfuryl substituent, possibly because of a favourable stacking interaction with the furan moiety or the formation of Hoogsteen bonds.

Table 4.1: Oligonucleotide duplexes for ICL formation with their T_m values

Duplex	Base pair	T_m (°C)	Modified Duplex	Base pair	T_m (°C)	ΔT_m (°C)
ODN4.1.3-ODN1	A-T	52	ODN4.1.2-ODN1	**nu4**-T	45	-7
ODN4.1.3-ODN2	A-G	43	ODN4.1.2-ODN3	**nu4**-G	40	-3
ODN4.1.3-ODN3	A-A	37	ODN4.1.2-ODN3	**nu4**-A	41	+4
ODN4.1.3-ODN4	A-C	36	ODN4.1.2-ODN4	**nu4**-C	41	+5

With ODN sequences:
ODN4.1.1: 5'-CGG ATG **nu4**CG TCA TTT TTT TTC-3'
ODN4.1.2: 5'-CTG ACG G**nu4**G TGC-3'
ODN4.1.3: 5'-CTG ACG GA**G** TGC-3'
ODN4.1.4: 5'-CTG ACG G**nu3**G TGC-3'
ODN1: 3'-CAG TGC C**T**C ACG-5'
ODN2: 3'-CAG TGC C**G**C ACG-5'
ODN3: 3'-CAG TGC C**A**C ACG-5'
ODN4: 3'-CAG TGC C**C**C ACG-5'

Interstrand cross-linking   Oxidation of the furan moiety was performed
with NBS for all duplexes as described previously.[157] Adduct formation
would be expected with ODN3 and ODN4, in which an adenosine and cytidine
unit, respectively, are opposite the kinetin modification in ODN4.1.2, through
reaction between the exocyclic amine and the oxidized furan moiety, both of
which are positioned in the major groove. This is illustrated for cytidine in
Figure 4.13.[138] Because thymidine does not possess an exocyclic amine and the
exocyclic amine of guanosine is positioned in the minor groove, no ICL formation
is expected in those cases. Nucleosides containing furan on the base have not
been used for ICL previously, so a selection of previously obtained data with
nu3-modified ODNs[138] is represented for comparison in Figure 4.14C. With all
furan-modified nucleosides previously used for ICL in our group, the highest ICL
formation yields were obtained upon cross-linking to cytidine.[134–138, 157]
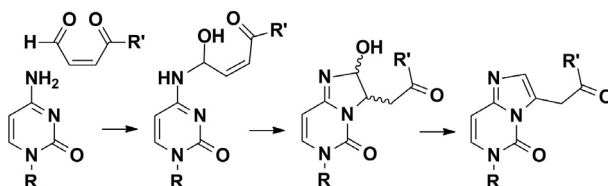
Figure 4.13: Mechanism of adduct formation with dC.[138]

Figure 4.14 shows the results of the cross-link reaction. The PAGE experi-
ment (Fig. 4.14A) shows a small amount of cross-link reaction in case of ODN4
containing a complementary C, but much less than expected and typically ob-
tained with our methodology. Moreover, and in contrast to previously obtained
ICL products, the lower-mobility compound was too unstable to be isolated and
characterized upon HPLC analysis. The HPLC chromatograms of the cross-link
reaction between ODN4.1.2 and ODN4 with increasing equivalents of added NBS
(Fig. 4.14B) show that modified ODN4.1.2 is consumed by the oxidation reaction,
thus behaving as a scavenger, but this clearly does not lead to the formation of
a new main compound or conjugates. Neither was ICL formation observed for
the other duplexes (HPLC chromatograms in the experimental part) Stable toxic
ICLs are thus hardly being formed upon oxidation, and this could explain the
observed positive effects of kinetin modifications in cells.
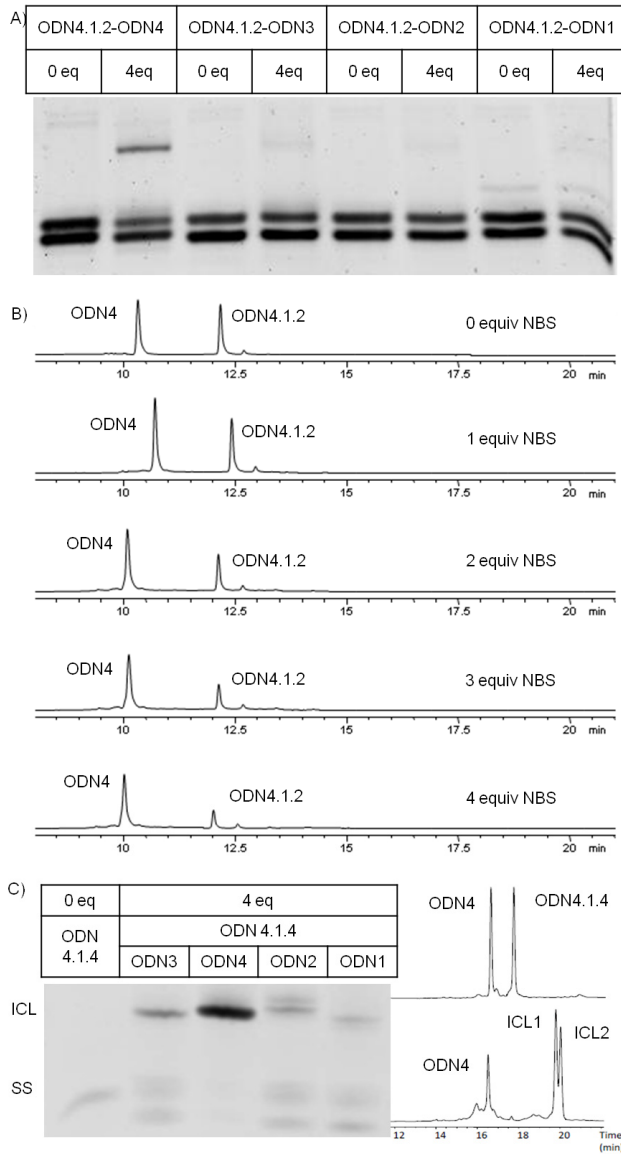
Figure 4.14: Results of the ICL formation reaction A) Denaturing PAGE of reaction mixture before and after the addition of 4 equiv NBS with ODN4.1.2-ODN4 in lanes 1 and 2, ODN4.1.2-ODN3 in lanes 3 and 4, ODN4.1.2-ODN2 in lanes 5 and 6, ODN4.1.2-ODN1 in lanes 7 and 8 respectively. B) HPLC of the cross-link reaction mixtures with ODN4.1.2-ODN4 following the reaction after addition of increasing amounts of NBS. C) For comparison, typically obtained results with **nu3**-modified ODN (ODN4.1.4) as example: denaturing PAGE (left) and HPLC (right).[138]

Visualisation   A plausible explanation for the difference in cross-linking behavior between both furan-modified oligonucleotides (**nu4** incorporation versus **nu3** incorporation) can be found in the positioning of the furan moiety of a kinetin modification in the major groove, away from the base pairing interface. A visualization of this scenario, with local optimizations is represented in Figure 4.15A. The furan moieties in the modified nucleosides that are used for cross-link formation can access the base pairing interface for reaction. The furan moiety of **nu3**, for example, can position itself via its linker from the sugar in between the bases (Fig. 4.15B), where it is possibly stabilized by stacking interactions.
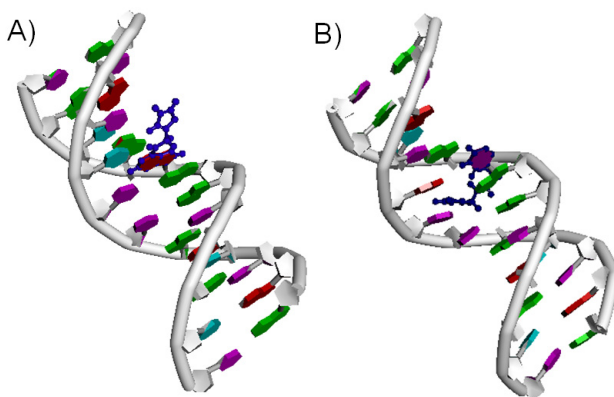
Figure 4.15: Visualisation of the positioning of the furan moiety (in ball and stick representation) in the duplex A) Visualisation of the duplex ODN4.1.2-ODN1. The furan moiety points away from the Watson-Crick interface into the major groove. B) Visualisation of the same duplex but modified with the **nu3**-A basepair for cmparison. The furan moiety is located on the Watson-Crick interface.

Conclusion   We here report on a new, fast and efficient synthetic protocol for the synthesis of naturally occurring kinetin-modified ODNs. The oxidation and cross-linking propensity of these ODNs was evaluated. We found that the furan moiety does indeed undergo oxidation, but in this case cross-linking is observed only in a very small amount, if at all. We thus propose that the antioxidant properties of kinetin are possibly linked to the oxidation potential of the furan moiety, which scavenges ROS, without causing more damage in the form of ICLs.

4.1.3 Towards extension of current methodology

The furan oxidation based cross-linking that was developed in our group works very well. However, the practical applicability is somewhat limited to stabilization of DNA duplexes as dumbbells[165, 166], the use of cross-linked duplexes for the study of repair mechanisms and the use of an inducible probe for antigene approaches. The latter however is strongly complicated by the localization of the genome in the nucleus and its highly compact form, wrapped around histone proteins as chromatin. Attempts have therefore been undertaken to expand the furan oxidation cross-link methodology and these will be discussed in the next sections of this chapter.

The first and main goal was to apply the furan oxidation cross-link methodology to DNA-protein interactions. Like discussed in the introduction, two approaches (illustrated in Fig. 4.16) can be followed for this purpose. Modifying proteins with a furan for cross-linking to their DNA target such as a transcription factor could potentially block or activate transcription. Modifying the DNA could be used to create cross-linking decoys, regulating gene expression by covalent capture and blockage of transcription factors in the cytoplasm. Furthermore, this technology could be applied for use in the study of transient DNA-protein interactions by covalent stabilization. For this latter approach, the DNA is modified with a furan containing nucleoside like for ICL formation. However because of this similarity, competitive ICL can occur, while the cross-link reaction will proceed through a different reaction mechanism targeting amino acids. The alternative approach starts by incorporating furan-modified amino acids in a protein or a more synthetically accessible peptidic protein model and proceeds through a similar cross-linking mechanism targeting the exocyclic amines of the bases of the target nucleic acid. This second approach was therefore initially followed.

To further broaden the possibilities in this context, the selective incorporation of furan moieties in proteins, rather than peptides, through amber suppression was also attempted. The obtained furan-modified proteins were tested in furan based labeling reactions.
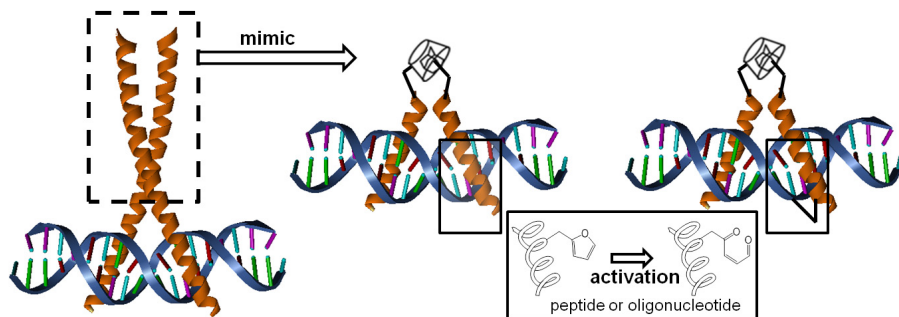
Figure 4.16: Illustration of the approaches to study furan oxidation based DNA-protein cross-linking, modifying the DNA or a protein mimic

Secondly, the furan interstrand cross-linking was applied on RNA duplexes, which is highly relevant in view of the recently found important role of RNA in regulation[167, 168] and for antisense approaches, targeting RNA in the cytoplasm. Though counterintuitive because of the completely different structure, stiffness, stability and reactivity, both NBS and singlet oxygen based furan oxidation cross-linking proceeded similar as before, apart from a different nucleoside preference in RNA context.

Lastly, an attempt was undertaken to visualize DNA interstrand cross-linking using fast scanning AFM on DNA origami. These experiments help for the application of the cross-linking approach on origami, which could be used in the assembly of complex architectures (e.g. with proteins) on such DNA chips.
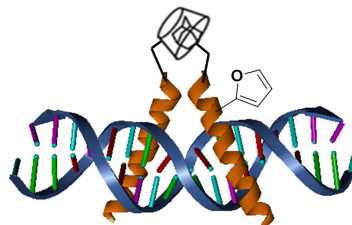
4.2 DNA ← protein mimic cross-linking

The idea to expand the furan based cross-link methodology to peptides, was not new in our group and has been extensively explored by Dr. Annelies Deceuninck.[169] Different peptides with known DNA binding behavior, have been synthesized with a furan modification and tested for cross-linking. Initially Dervan peptides *(see Fig.(2.1),* synthetic polyamides based on the natural minor groove binding distamycine A, were explored for the purpose, because of their easy synthesis and their well studied programmable DNA recognition. While selective oxidation of the furan moiety attached to a polypyrrole was difficult, no cross-link products could be observed. Therefore peptidic alternatives were searched. A synthetic peptide (SPRKSPRK), based on the SPPK motif *SPK(R)K(R)* from the histones of sea urchin sperm, was reported to interact with DNA probably through β-turns, stabilized by two hydrogen bonds. When no cross-linking was obtained, the HMG-I/Y peptide, *a fragment (TPKRPRGRPKK) of the eukaryotic, non-histon chromosomal HMG-I/Y protein,* was tested as an alternative model. More structural data was available for rational design, but still no cross-linked product was observed. However, as these three model peptides bind in the minor groove of A,T-rich DNA sequences, where no exocyclic amine for reaction is available, the search for a good model system continued. A helix (VEDRTVDVHIRRL-RKALEPG) from the PhoB transcription factor that inserts in the major groove and was reported to bind DNA on its own, was modified with furan for the purpose. However in view of the discussion in the previous chapter, it can be understood that this simplistic ap-

proach did not result in a good DNA binder. Because other major groove binders were not easily synthetically accessible, the attention was finally turned to threading intercalator peptides. A naphthalene diimide (NDI)-bisintercalator was synthesized, but also in this case no cross-linking was observed. The transcription factor mimics discussed and developed in the previous chapter were now considered for obtaining proof of principle for DNA-protein cross-link formation.

4.2.1 Furan-modified amino acid

In order to extend our furan based cross-linking methodology to peptides, furan-modified amino acids are required (Fig. 4.17). Fmoc protected furfurylamine (**Fmoc-aa1**) is commercially available. In previous work,[169] furan was also introduced N-terminally on the peptide by capping with 3-(2-furyl)propanoic acid (**aa2**), obtained from 3-(2-furyl)acrylic acid by hydrogenation. To allow for more diversification and evaluation of a different linker, additional furan-modified amino acids were synthesized (Fig. 4.17): two furan-modified lysines and a furan-modified glutamic acid.

| (4.2.1) | (4.2.2) | (4.2.3) | (4.2.4) | (4.2.5) |
| Fmoc- aa1 | aa2 | Fmoc-aa3 | Fmoc-aa4 | Fmoc-aa5 |
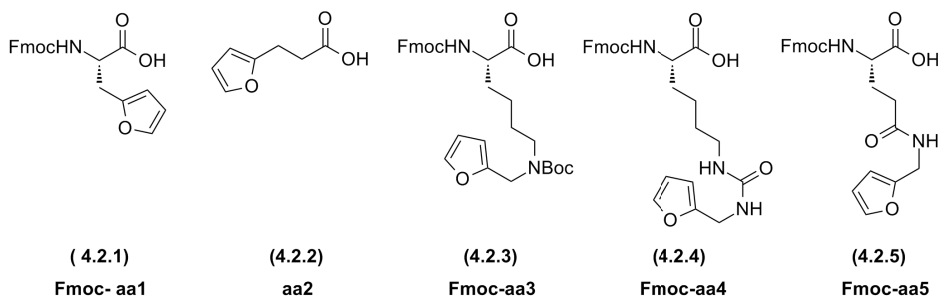
Figure 4.17: Overview of furan-modified amino acids protected for use in peptide synthesis

Initially **Fmoc-aa3** was synthesized, from the standard available Fmoc-Lys(Boc)-OH for peptide synthesis, by acid Boc deprotection, reductive amination with furfural and Boc reprotection, as illustrated in Figure 4.18. Although the Fmoc protecting group is not completely acid stable, orthogonal deprotection was possible using a 50/50 mixture of DCM and TFA for 15 to 60 min at rt. A convenient method for direct reductive alkylation of Fmoc pro-

tected amino acids with pyridine, thiazole, imidazole was described by Levadala and co-workers,[170] using the fast and high yielding alkylation conditions with NaBH(OAc)$_3$ of Abdel-Magid.[171] By appropriate stoichiometric manipulation of the amino acids and aldehydes and allowing imine formation before addition of NaBH(OAc)$_3$, monoalkylated products (**4.2.8**) could be obtained in good yield. The described reflux conditions for imine formation however were both found to be destructive and unnecessary and were thus slightly adapted. The reaction was most easily followed on TLC after a small scale reduction. Protection of the secondary amine is necessary for its use in peptide synthesis. Boc deprotection was chosen as it is an easy high yielding reaction, which in the end will result in a deprotectable furan-modified lysine, best resembling the native lysine. The dialkylated product is present in the reaction mixture as side product (10 − 15%). Though it does not require a reprotection step, the double modification may result in a lower DNA affinity due to sterical hindrance and will most probably result in a very complex cross-linking mixture. Therefore it was not further tested.
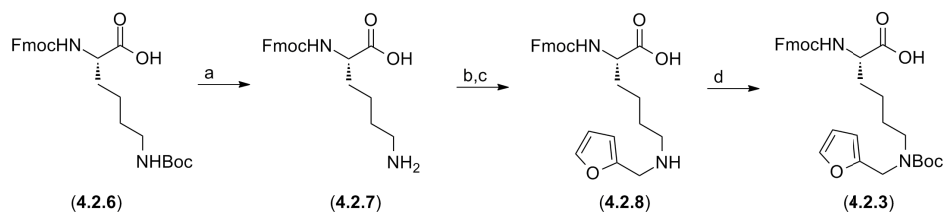
Figure 4.18: Syntesis of Fmoc-aa3 (4.2.3) a. TFA/DCM 1/1 for 1h at rt; b. addition of 0.95 equiv furfural in MeOH for 5h at rt; c. Addition of 1.4 equiv NaBH(OAc)$_3$ for another 4h; d. Quench the reaction with 1 equiv Boc$_2$O and 0.1 equiv DMAP for 1.5h (unoptimized overall yield 46%).

The monoalkylated amino acid (**4.2.3**) was also never used, as the obtained peptide modified with this furfuryllysine was thought to be able to react intramolecularly after oxidation as illustrated in Figure 4.19, rather than forming cross-links.
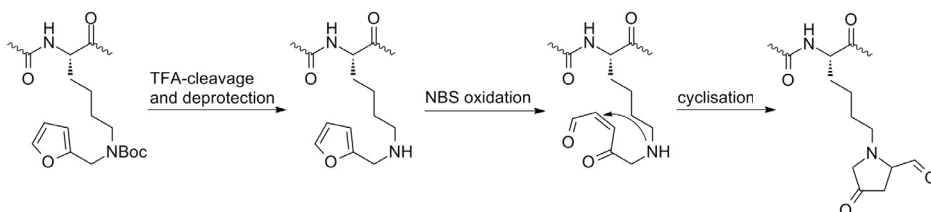
Figure 4.19: Intramolecular side reaction of a peptide modified with the synthesized furfuryllysine (4.2.3)

Alternatively, lysine can be modified with furfurylisocyanate to obtain (**4.2.4**) (Fig. 4.17), starting with the Boc deprotection as illustrated in Figure 4.20.
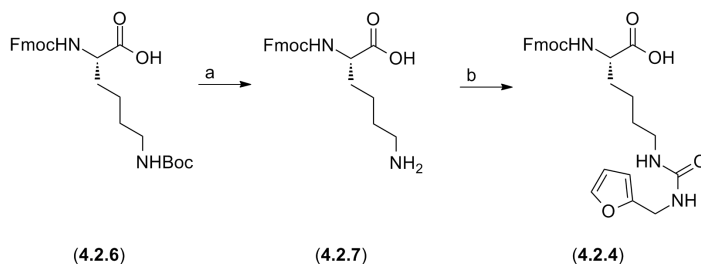
Figure 4.20: Syntesis of Fmoc-aa4 (4.2.4) a. TFA/DCM 1/1 for 1h at rt; b. 2.1 equiv furfurylisocyanate, 2.1 equiv DIPEA in DMF at rt for 2h (unoptimized overall yield 60%).

However when peptides were synthesized with this amino acid, the furan moiety was cleaved off by nucleophilic scavangers during the acid cleavage and side chain deprotection as illustrated in Figure 4.21. The ureum linker was indeed already used before in the design of a furan-modified nucleoside (Fig. 4.21B), but peptide and oligonucleotide synthesis conditions are drastically different, obviously influencing the outcome of the reactions.
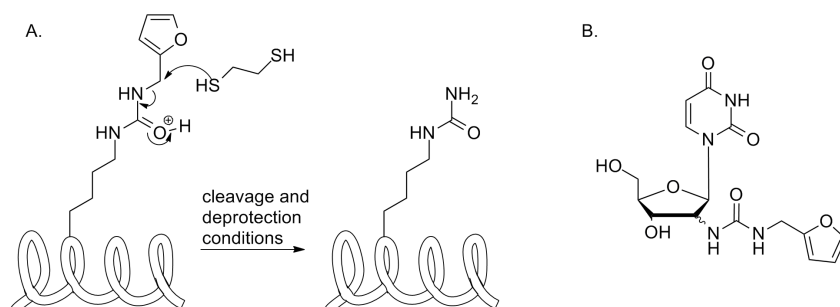
Figure 4.21: Unsuitability of the ureum linker for furan modification of an amino acid A) During the strongly acidic conditions for cleavage of the peptide from the solid support and the deprotection of its side chains, nucleophilic scavangers can cause cleavage of the furan moiety; B) Designed furan-modified nucleoside containing a ureum linker, which is stable under the phosphoramidite and DNA synthesis conditions

Because the introduction of a longer linker between the amino acid and the furan moiety might really enhance the cross-linking yield and furthermore because the stability of the furan moiety under acidic cleavage and side chain deprotection conditions should be investigated,[172] yet another furan amino acid (**4.2.5**) was synthesized. This time starting from glutamic acid rather than lysine, combination with furfurylamine and typical peptide coupling reagents can generate a stable peptide bond. Fmoc-Glu-OtBu (**5.2.10**) was bought, modified and tBu deprotected, similar like the earlier described Boc deprotection, as shown in Figure 4.22.
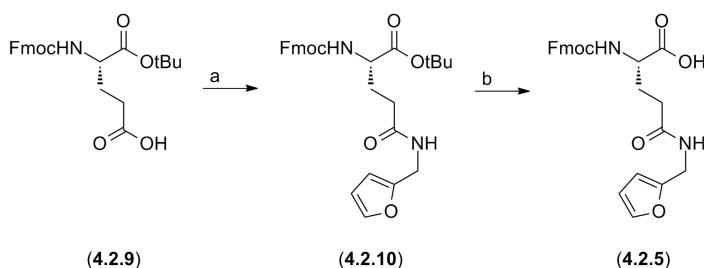
Figure 4.22: Syntesis of Fmoc-aa5 (4.2.5) a. 1 equiv furfurylamine, 1.2 equiv HBTU and 2 equiv DIPEA in DMF at rt for 2h; b. TFA/DCM 1/1 for 1h at rt (unoptimized overall yield 56%)

4.2.2   Positioning of the modifications in the mimic

To create the best possible scenario to achieve cross-linking, the position for incorporation of the furan-modified amino acid was carefully selected.

As described before, basic region leucine zippers, like the GCN4 transcription factor, bind DNA as a dimer of 2 uninterupted α-helices that dimerize through the formation of a coiled coil (Fig. 3.1). An α-helix and more specifically a coiled coil is prescribed by a characteristic 7 amino acid repeat, the 3-4 heptad repeat denoted as **a-b-c-d-e-f-g** (represented as the typical helical wheel in Figure 4.23.[173] Positions **a** and **d**, positioned 1 turn apart at the same side of the helix are typically occupied by hydrophobic amino acids such as Leu (hence the name leucine zipper), Ile, Val, and Ala. These residues cause the helix to have a hydrophobic site. Two such helices dimerize through hydrophobic interactions forming a shallow coiled coil. The residues at positions **e** and **g** are frequently polar or charged and can strengthen the coiled coil by electrostatic interactions. The basic regions extend as α-helices from the coiled coil and though the residues are now mostly basic for interaction with the DNA, their position can still be represented by the helical wheel. At the DNA interface, residues at position **g** are facing directly towards the major groove, at positions **c** and **d** the residues are just aside, one turn higher in the helix. The residues at the N and C termini of the basic region are positioned below and above the DNA.

However, the best positioned residue N235, cannot be replaced as it plays an essential role in DNA binding. Replacing it might thus compromise DNA binding. Therefore the important interactions were identified based on the crystal structures and literature.[118] Because the Morii mimic has higher affinity for the unnatural palindromic binding site CRE, experiments will be performed with DNA containing this binding site rather than the natural pseudopalindromic AP1 site. The two binding sites only differ in the center, where the two binding half sites (ATGAC) for the 2 peptides in AP1 overlap at the central C/G basepair (5'-ATGACTCAT-3'), they are separated in CRE (5'-ATGACGTCAT-3'). The interactions between the peptide and the DNA are thus merely the same, except in the center where the DNA is bent to accommodate the extra

		a	b	c	d	e	f	g
						D 289	P 290	A 291
Basic region		A 292	L 230	K 231	R 232	A 233	R 234	N 235
		T 236	E 237	A 238	A 239	R 240	R 241	S 242
		R 243	A 245	R 246	K 247	L 248	Q 249	R 250
		M 251	K 252	Q 253	L 254	E 255	D 256	K 257
Leucine zipper		V 258	E 259	E 260	L 261	L 262	S 263	K 264
		N 265	Y 266	H 267	L 268	E 269	N 270	E 271
		V 272	A 273	R 274	L 275	K 276	K 278	L 279
		V 280	E 281	G 282	R 283		(COOH)	

A. ionic interactions / hydrophobic interactions

B.

Figure 4.23: Helical wheel representation of GCN4 A) Illustration on the formation of a coiled coil; B) residues of GCN4 and their corresponding positions in the helix. The residues that are at the DNA interface are on a gray background, the dashed line accentuates position **g**, the position directly in the major groove, the dotted line accentuates position **c** and **d** that are just aside, one turn higher in the helix.

basepair. Transcription factor mimics with a large dimerizing moiety, like an adamantine-cyclodextrin inclusion complex, have a higher affinity for the CRE binding site as there is more space.

CRE:	5'-	A	T	G	A	C	G	T	C	A	T	-3'
	3'-	T	A	C	T	G	C	A	G	T	A	-5'
AP1:	5'-	A	T	G	A	C	T	C	A	T		-3'
	3'-	T	A	C	T	C	A	G	T	A		-5'
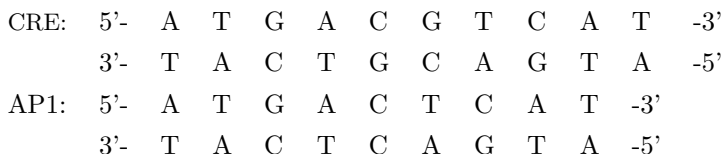
All of the residues, which make specific contacts with DNA can be found in the residues 231-245: KRARNTEAARRSRAR, which are of course on the DNA interface. Five amino acid residues, Asn235, Thr236, Ala238, Ala239 and Arg243 make direct base contacts, seven amino acid residues, Arg232, Arg234, Arg240, Arg241, Ser242, Arg243 and Arg245 further make hydrogen bonds to the phosphate backbone. Other interactions occur through an extensive network of solvent molecules (Fig. 4.24).

Figure 4.24: The protein, DNA and the interactions of the GCN4 bZIP CRE Complex, taken from Keller et. al. [118] A) GCN4-bZIP domain numbered according to the full-length sequence. Residues that make one or more DNA contacts are indicated: b, direct to base; w, via water to base; p, direct to phosphate; x, via water to phosphate. B) Self-complementary, 18 bp palindromic CRE DNA with terminal thymidine. A molecular and crystallographic 2-fold axis passes through the C-1G1 base step. The specific recognition sequence is marked with the bar. C) DNA-protein interaction summary for one basic domain and DNA half-site. The location of the 2-fold axis is indicated. Direct and water (W) mediated interactions with bases are shown as continuous lines, interactions with the phosphate backbone as broken lines and hydrophobic interactions as dotted lines. Arrows indicate the direction of hydrogen bonds.

After combination of both criteria, as symmerised in Figure 4.25, only K231 and R232 remained as possible positions for modification (Fig. 4.26). With visualization software it was possible to measure the distance of these residues to the most proximate nucleophile of the DNA. This revealed K231 to be the better option. This residue was replaced in a first modified peptide bGCN4(K231aa1). The interacting residues logically showed closer proximity. A239 interacts through a hydrophibic interaction with a thymine methyl group. As the methylene of furfurylalanine (**aa1**) could replace this interaction, this residue was replaced in a second modified peptide bGCN4(A239aa1).

residue	D	P	A	A	L	K	R	A	R	N	T	E	A	A	R	R	S	R	A	R	K	L	Q
number	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248
binding							x		x	x	x	x	x	x	x		x	x			x		
position	x	x	x	x	x			x	x		x	x			x	x		x	x	x	x	x	x
result	x	x	x	x	x			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Figure 4.25: Summary of the considerations for the incorporation of the furan-modified residue based on both the involvement in essential interactions for binding and the positioning. Non-ideal residues are crossed out with an 'x'
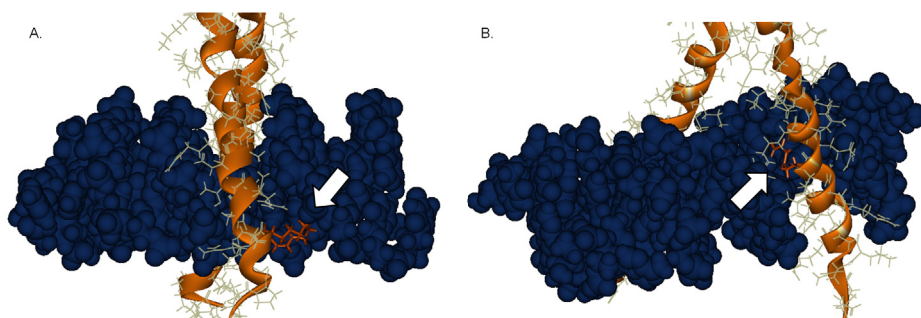
Figure 4.26: Visualisation of DNA binding of GCN4 modified with oxidized aa1 at position K231 (A) and A239 (B)

4.2.3   Synthesis of furan-modified peptides

Two different modified peptides were thus synthesized:

- bGCN4(K231aa1): DPAAL**aa1**RARNTEAARRSRARKLQ

- bGCN4(A239aa1): DPAALKRARNTEA**aa1**RRSRARKLQ

The **aa1**-modified peptides were synthesized in a similar way as the unmodified peptides before. Though oxidation and degradation of the furan moiety has been described under acidic peptide cleavage and deprotection conditions,[174] and observed when using **aa1** or **aa2** at the N terminus of a peptide,[172] cleavage and deprotection of peptides internally modified with **aa1** in our hands does not present any problems. It was reasoned that the peptide protects an internal furan moiety. Still, to minimize possible side reactions, the acidic cleavage and deprotection coctails were added cooled on ice.

It was chosen to equip all peptides as described before with Ad, as this is the most easy reaction and Ad-peptides are easier to analyze. To form dimers, these peptides were then combined with the unmodified CD-peptide.

4.2.4   DNA binding

Unexpectedly, the furan modification nearly completely abolished DNA binding, as can be seen in Figure 4.27. Compared to the unmodified mimic it is clear that the mimic modified with **aa1** at K231, hardly binds to the DNA, while the mimic modified with **aa1** at A239 binds as a complex with lower electrophoretic mobility. The same shifted band and thus complex can be observed in lane 3, which contains only CD-peptide. CD-peptides can form CD-CD dimers on their own,[72] giving rise to a complex with a slightly lower electrophoretic mobility than Ad-CD dimers, due to the size. Typically, formation of this complex is far less efficient than the formation of the Ad-CD dimer. However, in this case with furan-modified Ad-peptides that have a lower affinity for the DNA, dimerization with two unmodified CD-peptides was found to be dominating. This is problematic for cross-linking as the resulting CD-CD dimer-DNA complex does not contain a furan moiety. This is schematically represented in Figure 4.28.
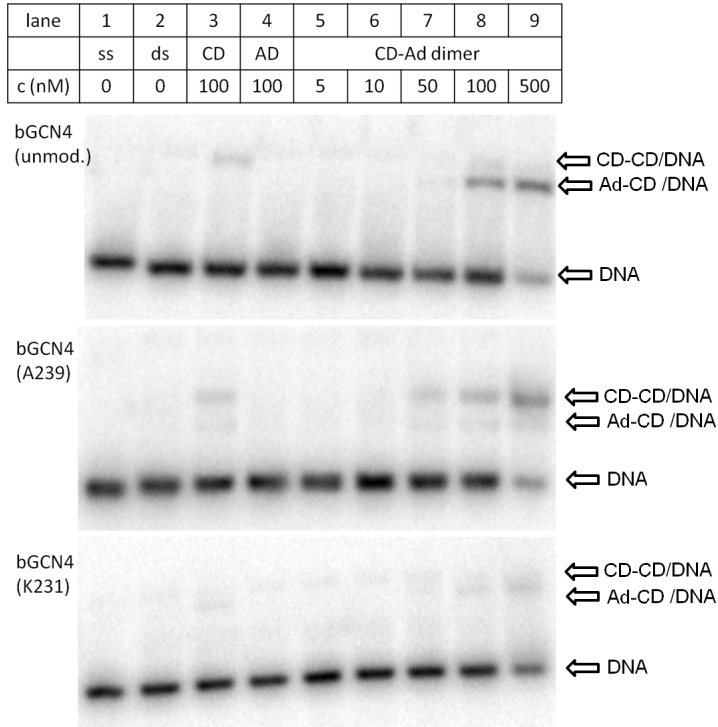
Figure 4.27: EMSA of the unmodified (top), A239 furan-modified (middle) and K231 furan-modified (bottom) Morii mimic to ^{32}P-labeled AP1 binding site (native 8% acrylamide gel, with 0.2% TBE buffer). Lane 1 contains single stranded AP1 and 2 the AP1 duplex without added peptides. Lane 3 and 4 contain 100 nM of the CD and the AD peptide respectively. Lanes 5 to 9 contain 5, 10, 50, 100, 500 nM of both AD and Cd peptides with 6% sucrose, 100 mM Tris pH 7.6, 20 mM KCl, 10 mM MgCl2, 2.5 mM EDTA and 4.88 pM DNA.
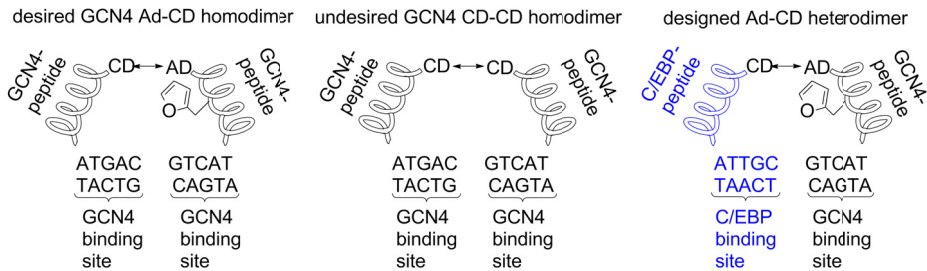
Figure 4.28: DNA binding of the furan-modified mimic From left to right the desired furan-modified GCN4 Ad-CD homodimer, the undesired GCN4 CD-CD homodimer and the designed furan-modified Ad-CD heterodimer.

To overcome the problem, heterodimers were formed through the inclusion complex of Ad and CD on different peptides, derived from different bZIP proteins, which thus bind different DNA sequences. If an oligonucleotide composed of the binding half sites for both proteins is used, then homodimerization between two cyclodextrine-modified peptides is not possible. Practically this was achieved by combining the synthesized GCN4 furan-modified Ad-peptides with a CD-peptide derived from EmpB1, another bZIP protein with a DNA duplex containing one binding halfsite for both peptides, as illustrated in Figure 4.28. This proved to be a very efficient way to disseminate the undesired CD-CD dimerization and thus reveal the real DNA binding properties (Fig. 4.29). Both modified dimers bound to DNA, but compared to the modification of the non-essential lysine bGCN4(K231aa1), modification of the central, essential alanine (A239aa1) strongly reduced the affinity.

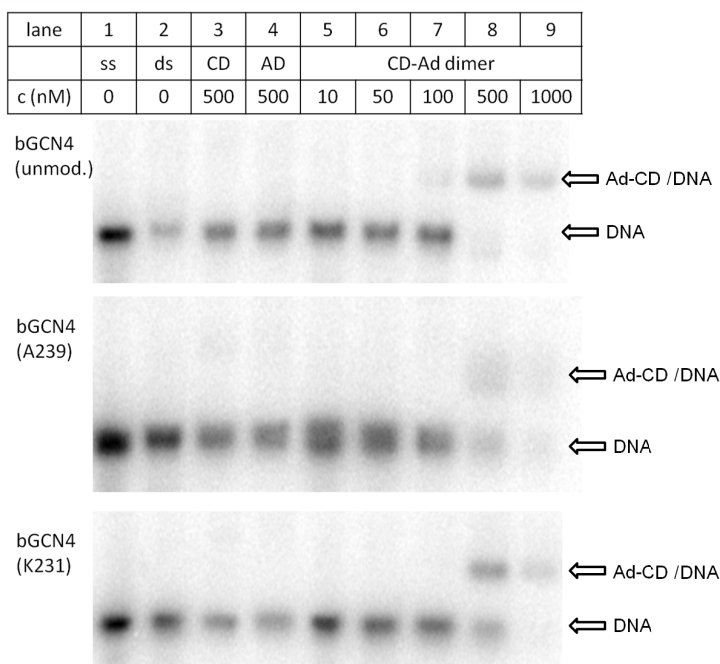| lane	1	2	3	4	5	6	7	8	9
	ss	ds	CD	AD	CD-Ad dimer				
c (nM)	0	0	500	500	10	50	100	500	1000

Figure 4.29: EMSA of the unmodified (top), A239 furan-modified (middle) and K231 furan-modified (bottom) Morii mimic to ^{32}P-labeled CECRE DNA (conditions see Fig. 4.27).

4.2.5 Cross-link tests
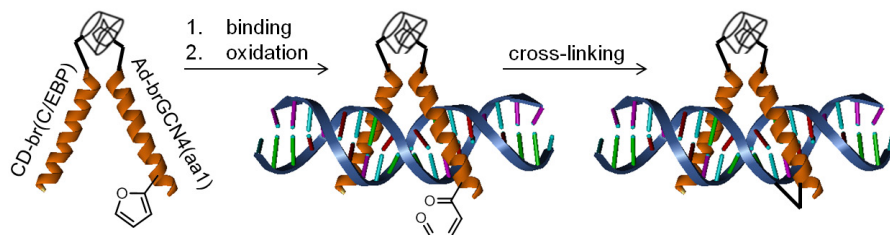
Figure 4.30: Illustration of the aim of DNA ← protein mimic cross-linking

The oxidation of the furan moiety in the Ad-peptides with NBS was monitored by mass analysis of the reaction mixture (without buffer to improve ionisation) after 5, 15, 30, 45 and 60 min. It was found (Fig. 4.31) that 1 equivalent of NBS led to complete oxidation (+16 Dalton) of the peptide (mass: 3108 Dalton) to the keto-enal (mass: 3124 Dalton) of the peptide after 5 minutes. After two hours another equivalent of NBS was added, which led to over-oxidation to the acid (+32 Dalton) (mass: 3140 Dalton).

Cross-link formation was then monitored by denaturing PAGE with ^{32}P-labelled DNA. Under these conditions, only the formation of a covalent bond can give rise to a shifted band. Many different conditions were applied (different concentration, buffer, reaction times, equivalents NBS, $NaCNBH_3$ reduction or not, ...), but a shifted band was never found.

The absence of cross-linked product in the gel shift experiments was double-checked with HPLC on both peptide (C18 column with TFA buffer) and DNA level (different oligo columns with TEAA buffer). Oxidation of the peptide in water with one equivalent of NBS indeed yielded the oxidized product, a little bit of overoxidation and a little bit of remaining start product as can be seen in Figure 4.32, exactly the same as determined before with MS. But to obtain such a clean chromatogram from the reactive 4-oxobutenal is suspicious. The oxidized compound remained stable for over 10 hours in buffer and did not react with different added nucleophiles like propylamine, cytidine mono phosphate, adenosine
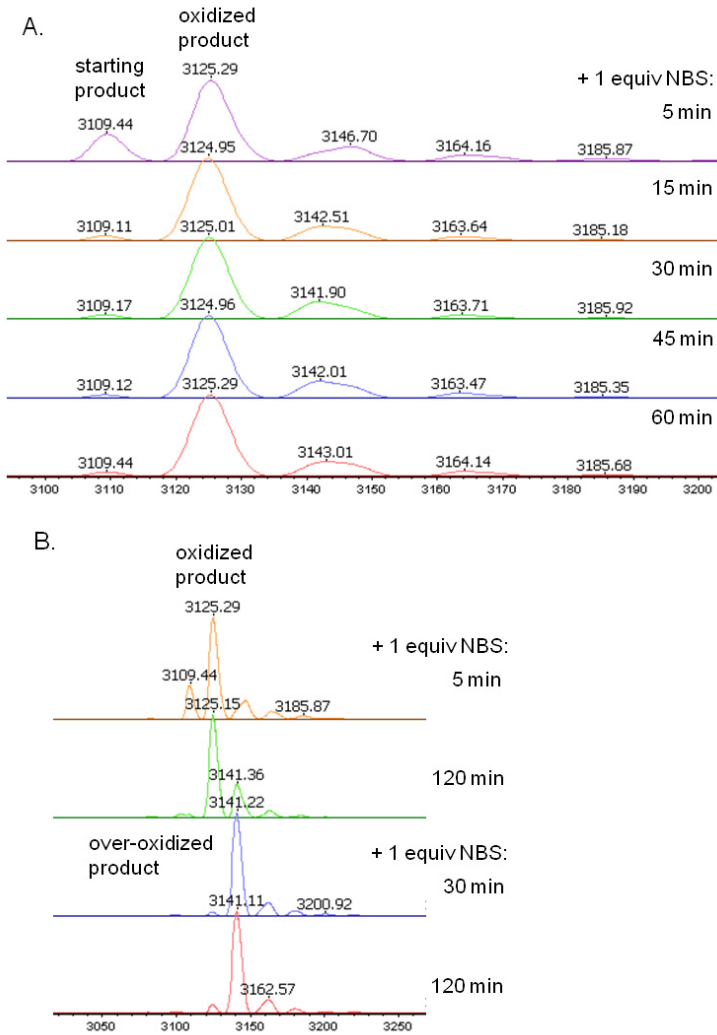
Figure 4.31: Maldi-MS spectra of the oxidation of Ad-bGCN4(K231aa1) (mass 3108 Dalton). A) From top to bottom after 5, 15, 30, 45 and 60 min with 1 equiv of NBS in mQ water; B) From top to bottom after 5 and 120 min with 1 equiv of NBS, then 1 extra equiv of NBS was added and this was monitored after 30 and 120 min.

triphosphate, glutathione, . . . with different amounts of equivalents, during different reaction times, in different buffers, at different pH and temperature. This led to the conclusion that the NBS treatment had not really oxidized the furan. In the construct two very oxidation sensitive thioether bonds, connecting Ad and CD to the peptides are present.[175, 176] The oxidation reaction had thus been optimized to generate a sulfoxide derivative, the till then assumed overoxidized product had an oxidized furan present next to the sulfoxide, rather than the originally assumed acid.
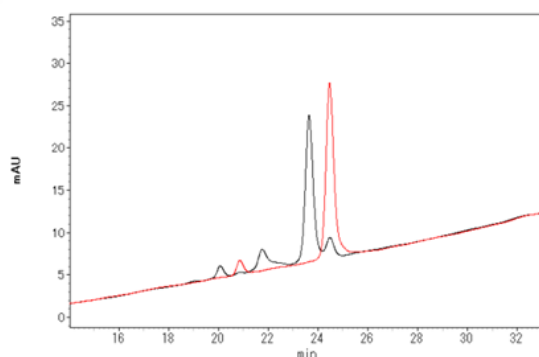
Figure 4.32: HPLC Chromatogram of the oxidation of Ad-bGCN4(K231aa1) (C18, 20-80% MeCN from 0.5% TFA buffer) before (red) and after (black) oxidation.

Optimization of the oxidation conditions thus had to be repeated. To make sure a reactive 4-oxobutenal moiety was generated, it was trapped with hydrazine forming an aromatic product that absorbs at 280 nm. Figure 4.33 shows the reaction progress as monitored by HPLC at 214 and 280 nm. Before reaction there is a single peak at 214 nm that does not absorb at 280 nm, as there are no aromatic moieties in the peptide. Upon oxidation, the starting material disappears and new products are formed. After quenching with hydrazine, there is clear formation of a product that absorbes at 280 nm. Moreover, if the peptide was treated by only hydrazine, no reaction occurs (Fig.4.34A). If NBS and hydrazine are added at the same time, again no reaction occurs (Fig.4.34B) as the reagents react with each other. The aromatic product is only formed if NBS is added before hydrazine to oxidize the furan. The reaction was repeated in this order (Fig.4.34C) and shows excellent reproducibility. The formed product could

also be identified by mass analysis.
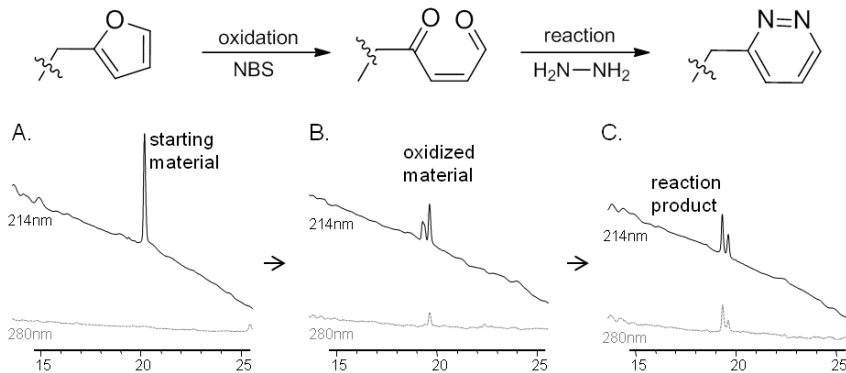
Figure 4.33: Verification of the presence of 4-oxobutenal by reaction with hydrazine reaction scheme (top), HPLC chromatograms (C18, 100Å, 0-50% MeCN from 0.1% TFA buffer in 25 min) A) peptide Ad-bGCN4(A239aa1) before oxidation; B) peptide Ad-bGCN4(A239aa1) after oxidation; C) peptide Ad-bGCN4(A239aa1) after oxidation and reaction with hydrazine (bottom). This clearly shows the formation of an aromatic product (asorbing at 280 nm) from an unstable oxidized intermediate.

However, despite the now proven oxidation, cross-link reactions still did not result in any observable covalent DNA-peptide adduct formation as shown in Figure 4.35. Only some degradation occured upon oxidation, the exact same traces, with broadening and additional peaks were observed in a blanc reaction without mimic.

Though unfavorable positioning and/or linking cannot be completely excluded, the results strongly suggest that the involvement of the targeted exocyclic amines in Watson-Crick base pairing in the highly organised duplex, strongly reduces their reactivity towards cross-linking. This rationale is further supported by the reported use of aldehydes for the detection of single stranded regions of DNA [177] and their use to elucidate RNA folding pathways.[178] Figure 4.36 depicts the difference between ICL and the targeting of double stranded DNA with a furan-modified peptide. While the reactive moiety on a peptide has to intrude in and interfere with the highly organized duplex structure for DNA ← protein cross-
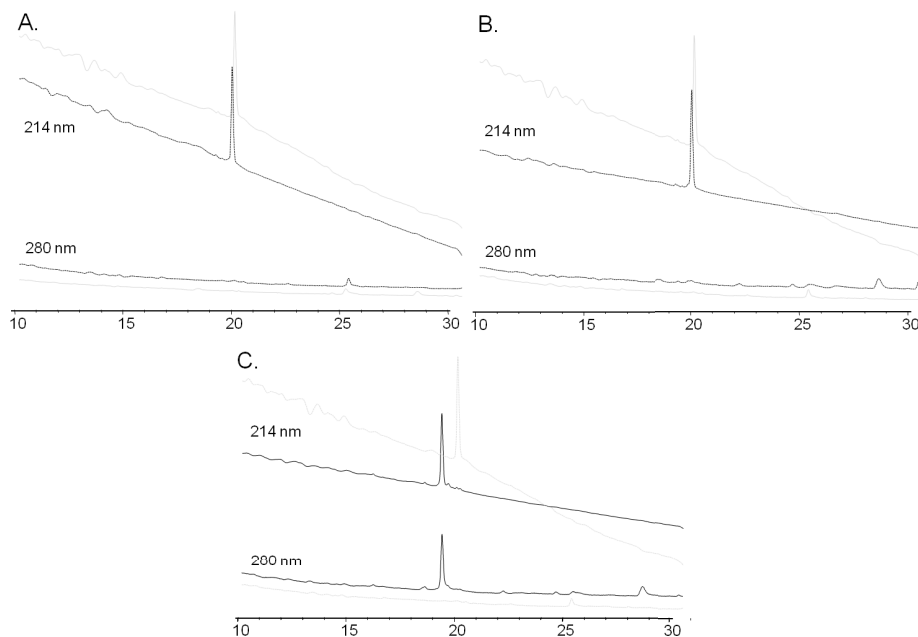
Figure 4.34: Additional experiments for the verification of the presence of 4-oxobutenal by reaction with hydrazine A) treatment with only hydrazine; B) treatment with NBS and hydrazine at the same time; C) treatment with NBS before treatment with hydrazine (check reproducibility of the reaction). The trace of the reaction mixture is presented in black, the trace of the unreacted Ad-bGCN4(A239aa1) is superposed in light gray.
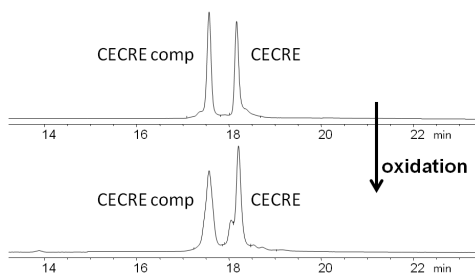
Figure 4.35: Cross-linking experiments. The top HPLC trace is before and the bottom trace after furan oxidation of K231-modified Morii mimic (which is not visible in the HPLC chromatograms) bound to the CECRE duplex. No cross-link formation can be observed.

linking, the nucleophile of the base opposite to the furan-modified ODN is readily available for ICL. Moreover, in current research with triplex cross-linking (cfr. unpublished results of Ellen Gyssels) it was found that furan cross-linking is only established from the triplex forming oligonucleotide to a duplex, when the furan modification is incorporated at the extremity of the TFO and thus positioned at the distorted transition from triplex to duplex. It was further shown that through introduction of distortions in the triplex region by mismatches and increased temperature, triplex cross-linking yield could be increased. The introduction of distortions in the target duplex for the cross-link reaction with the peptide was attempted, by introducing a mismatch in CECRE (CECRE mis: 5'-CGG ATT GCG TCA TCT TTT TTC-3') or a flipout base in CECRE comp (CECRE flip: 5'-GAA AAA AAC ATG ACG CAA TCC G-3') or by increasing the reaction temperature. However these distortions to improve the cross-link formation, also disturb the binding of the mimic and thus causing the adverse effect. It had to be concluded that cross-linking with our furan-based approach to dsDNA was not efficient and that the mimic was not a good model system for further investigation with forced cross-link formation.
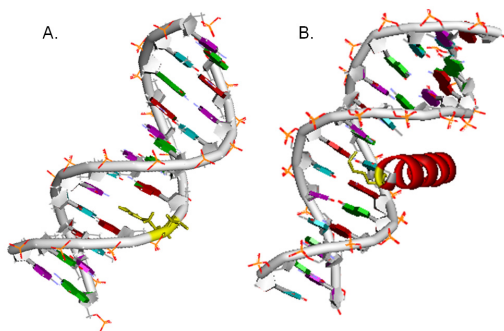
Figure 4.36: ICL vs. cross-linking from a modified peptide to double stranded DNA (modified residues indicated in yellow). In A the oxidized furan moiety faces an unpaired C base, while in B the oxidized furan moiety faces a Watson-Crick base paired C base.
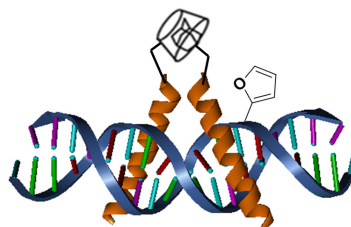
4.3   DNA → protein mimic cross-linking

The cross-link formation should thus be approached from the other side. Indeed, to form DNA-peptide cross-links, there are two options: modification of the peptide or modification of the DNA. The first approach was initially followed to keep the same cross-linking mechanism, targeting the exocyclic amines of the DNA bases. However, this proved difficult. To succeed in the other way around, the amine functionalities of lysine side chains now have to be targeted. Moreover and more importantly, competitive interstrand cross-linking should be avoided. A furan moiety facing away from the Watson Crick interface, into the major groove where the peptide binds is thus required.

Around this time in my project, thesistudent E. De Wilde, was attempting ICL with a nucleoside modified with furan on the base . . .

Bioconjugate Chemistry

Toxicity Inspired Cross-Linking for Probing DNA–Peptide Interactions

Lieselot L.G. Carrette,[†] Takashi Morii,[‡] and Annemieke Madder*,[†]

[†]Organic and Biomimetic Chemistry Research Group, Department of Organic Chemistry, Ghent University, Krijgslaan 281-S4, 9000 Gent, Belgium

[‡]Kyoto University, Uji, Kyoto 611-0011, Japan

S *Supporting Information*

ABSTRACT: A cross-linking methodology for the study of DNA–protein interactions is described. The method is inspired by the metabolic activation of furans causing toxic DNA damage, including DNA–protein cross-links (DPC). The furan moiety, representing a latent functionality, is easily incorporated into oligonucleotides, and can be activated on demand to release a reactive aldehyde. Reaction with nucleophilic lysine side chains is shown to be distance-sensitive and allows for site-selective DPC formation.
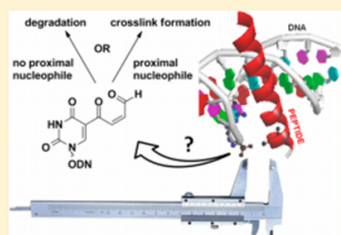
Introduction The interaction of proteins with nucleic acids is essential for cellular functioning, as they are responsible for the regulation and execution of many biological processes, including folding, replication, transcription, and degradation. Understanding the remarkable selectivity and affinity of DNA binding proteins for their targets in the complex cellular environment is a major challenge, driven by the ambition to design drugs that can compete in such interactions.[68, and references therein] In an attempt to decipher the DNA-protein recognition code, analysis of crystal structures and modified proteins has been greatly aided by the use of bifunctional cross-linking moieties, causing covalent locking of the interactions. Photoactivatable functionalities[179–182] such as phenylazides, diazirines, and benzophenones have been exploited for DNA-protein cross-link (DPC) generation, but lack selectivity for identification of the amino acid residues responsible for selective recognition at the DNA-protein binding interface, as the generated reactive species can target any amino acid. Further, due to the occurrence of high-energy intermediates, the resulting structures are not always representative of the actual structure in the absence of the cross-link. Finally, obtained DPC yields are in general rather low (10-20%).[183]

In order to probe the DNA-protein interface use of a cross-linking mechanism requiring the presence of proximal nucleophilic side chain functionalities can be considered as a viable alternative.

Aldehydes, representing sensitive functionalities reactive toward nucleophiles, common occur in the environment and are endogenously generated by metabolic activation and lipid peroxidation. As a result of random adduct formation with DNA and proteins, they are mutagenic and cytotoxic, respectively.[184, 185] Aldehydes are also known to covalently trap proteins onto DNA exploiting the nucleophilic side chain functionalities of proteins.[186–189] Consequently, externally added formaldehyde is routinely used for random DPC formation in chromatin immunoprecipitation experiments.[190] However, for site-selective formation of a DPC, oligonucleotides (ODNs) containing a suitably positioned aldehyde moiety are required. Abasic sites, which are toxic DNA lesions resulting from oxidative damage that ring-open to aldehydes[191] and can as such subsequently form DPCs, can be locally generated in DNA by action of uracil-DNA glycosylase on a correctly positioned deoxyuracil or by photolysis of a protected precursor.[192, 193] Introduction of other aldehyde functionalities into ODNs has been achieved through oxidation of a diol with $NaIO_4$. DPC formation after oxidative conversion of ODNs with diols replacing the sugar[194] or diols attached to the 2-position[195, 196] or the base[197] have been reported. In the latter case 5-formyldeoxyuridine (fdU) is generated, an abundant naturally occurring form of thymidine damage in DNA,[198, 199] proposed to be involved in the formation of DPCs.[200, 201] Unfortunately the synthesis of these diol building blocks[194–197] with extended protection schemes is lengthy. Aromatic aldehydes can be introduced without protection[202] or through postsynthetic Suzuki-Miyaura coupling with 5-iodocytosine[203] as they are less reactive. However, the lower reactivity is a disadvantage in subsequent cross-linking reactions with nonactivated nucleophiles.
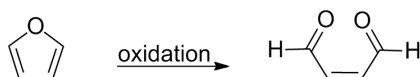
Figure 4.37: Furan oxidation

Inspired by the toxic metabolic activation of furan into 4-oxobutenal (Fig. 4.37),[127] we report here on the use of easily accessible furan-modified ODNs for site-selective DPC formation. The furan moiety represents a latent functionality, which can be oxidatively activated on demand to release the reactive aldehyde. Furan incorporation in ODNs is straightforward as the small aromatic moiety is stable under both nucleotide and ODN synthesis conditions. Moreover, various furan derivatives are commercially available, allowing modular construction of different furan-substituted nucleoside derivatives. We have previously shown that in a nucleic acid context, selective activation by oxidation to generate a reactive aldehyde is possible and can be carried out after duplex formation, thus minimizing nonspecific off-target reactions and causing site selective cross-link formation.[134, 135, 138, 139] Due to the obtained high yield and selectivity, the furan-oxidation methodology has recently been accepted as a general procedure to achieve DNA interstrand cross-linking.[157] Additionally, very recently, the usefulness and wider applicability of the furan-oxidation methodology has been illustrated by Summerer and co-workers in the cross-linking of genetically modified proteins incorporating a furan-amino acid toward RNA targets.[204] Here, using nucleoside **nu5** (**4.3.1**) (Fig. 4.38) rather than our previously developed furan-modified building blocks, we describe for the first time cross-linking to non-nucleic acid targets, considerably expanding the scope of the furan-oxidation cross-linking methodology and opening the way to a much broader application of furan-based probes.
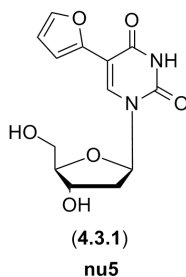

(4.3.1)

nu5

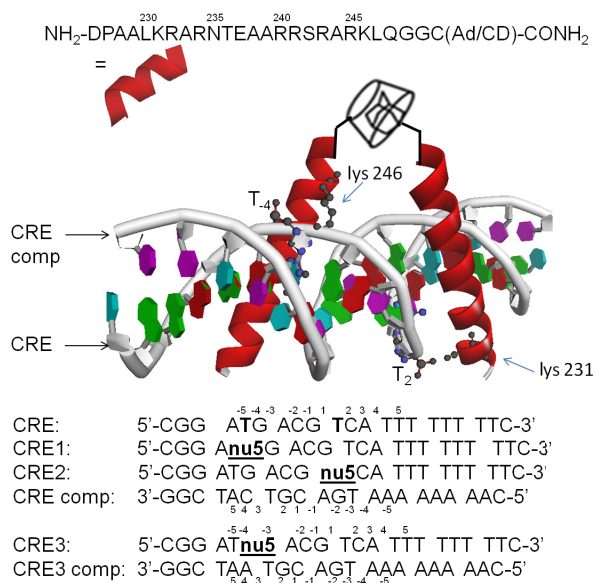**Figure 4.38: Furan-modified nucleoside nu5**

Results & Discussion   Previously we developed several furan-modified nucleosides for efficient DNA interstrand cross-linking (ICL).[134, 135, 138, 139, 157] For the formation of DPCs we decided to study a nucleoside (**nu5**) modified on the base, which positions the furan moiety away from the Watson Crick interface into the major groove where the recognition helices of proteins bind. In view of this different specific positioning of the furan moiety, it was first verified that no competitive ICL formation took place upon oxidation of the furan-modified ODN in a duplex with its complementary strand (see experimental section for spectral data).

To apply furan-oxidation based cross-linking in a DNA-protein context, one must further take into account that oxidation or modification of methionine, tryptophan, tyrosine, histidine, and cysteine residues can occur.[176, 205] However, proteins bind DNA through recognition helices, protein fragments especially rich in alanine, leucine, and arginine that have a high helix propensity[206] complemented with asparagine, glutamine, aspartate, glutamate, and arginine for specific interactions with the DNA.[65] Arginine, lysine, and histidine as charged residues further interact with the DNA backbone. These residues are highly conserved across DNA binding proteins,[207, 208] illustrating their relevance. Oxidation of the earlier mentioned oxidation sensitive residues that are not conserved and do not interact with the DNA bases should thus not interfere to a major extent with DNA binding selectivity and affinity and are not expected to complicate efficient DPC formation. In the absence of proximal nucleophiles, aldehyde-modified ODNs have earlier been shown to undergo self-destruction through several degradation processes within the aqueous buffer medium, further reducing the chances for collateral damage.

To study and deliver proof of concept for the possibility of cross-linking at the DNA-protein interface, a previously developed GCN4 transcription factor mimic was used as test system.[72] GCN4 is a transcription factor from yeast belonging to the basic region leucine zipper family. This transcription factor family features 2 uninterrupted α-helices that dimerize through a coiled coil and N-terminally interact with the DNA through 2 basic region peptides, or recognition helices.[173] The model system chosen here consists of these two

DNA binding basic region peptides modified with either a cyclodextrin (CD) or adamantane (Ad) moiety to form a dimer through an inclusion complex (as shown in Figure 4.39), which retains DNA binding selectivity and affinity of the original protein. This miniaturization allows easy synthetic access, through solid phase peptide synthesis. It further allows a more straightforward analysis of conjugation products.

NH$_2$-DPAALKRARNTEAARRSRARKLQGGC(Ad/CD)-CONH$_2$

CRE: 5'-CGG ATG ACG TCA TTT TTT TTC-3'
CRE1:   5'-CGG A**nu5**G ACG TCA TTT TTT TTC-3'
CRE2:   5'-CGG ATG ACG **nu5**CA TTT TTT TTC-3'
CRE comp: 3'-GGC TAC TGC AGT AAA AAA AAC-5'

CRE3:   5'-CGG AT**nu5** ACG TCA TTT TTT TTC-3'
CRE3 comp: 3'-GGC TAA TGC AGT AAA AAA AAC-5'

Figure 4.39: Illustration of the DNA binding protein mimic of GCN4 showing two lysines (lys 231 and 246) of the Ad-CD homodimer in close contact with two thymines (T) in the DNA binding site CRE:CRE comp (adapted from pbd structure 2DGC).[118]

The synthesis of **nu5** and its incorporation into ODNs using phosphoramidite chemistry was described by Greco and co-workers for its fluorescent properties.[209] Wicke and co-workers obtained **nu5**-modified ODNs by a postsynthetic modification.[210] It was further demonstrated that **nu5** is amenable to enzymatic incorporation into DNA.[211] For the purpose of current investigation, furan-modified nucleoside **nu5** was incorporated using standard phosphoramidite protocols in the palindromic CRE DNA binding site of GCN4,

consisting of two interaction half sites 5'-ATGAC-3' (numbered -5 to -1 in Figure 4.39) for the two DNA binding peptides. To induce minimal duplex destabilization and thus avoid loss in binding affinity, initially thymidine residues were considered for substitution by **nu5**. The base pair **nu5**-A was found to be as stable as the Watson-Crick T-A base pair, as judged from melting temperature measurements. The observed small changes in melting temperatures upon the introduction of **nu5** (T_m(CRE1) = -1°C, T_m(CRE2) = 0°C, T_m(CRE3) = -3.1°C) indicate minimal destabilization of the duplex by the modification. Visualization (using PDB structure 2DGC)[118] shows thymidines T-4 and T2 in proximity to the two lysines (246 and 231, respectively) of the major groove binding peptide (Fig. 4.39). T-4 and T2 were thus replaced by **nu5**, in CRE1 en CRE2, respectively. From this visualization it was, however, also clear that guanosine G-3 is closest in proximity to lys 246. These moieties are involved in a water mediated amino acid-base contact.[118] Therefore, an additional sequence, CRE3, was synthesized and combined with a similarly mutated complementary strand CRE3 comp to equally obtain a **nu5**-A base pair. Binding of the GCN4 mimic to the furan-modified CRE binding sites was verified using electrophoretic mobility shift assay (EMSA) binding studies (Fig. 4.40).
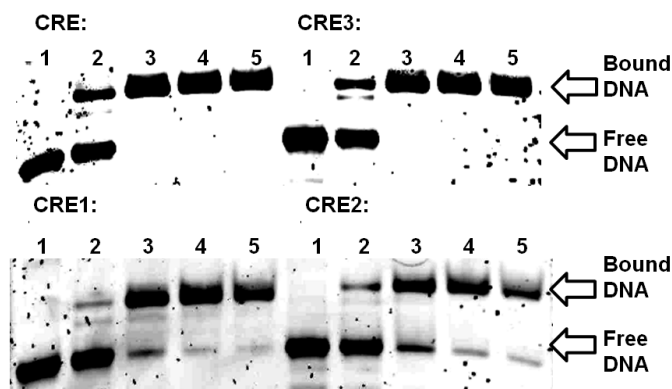

Figure 4.40: EMSA of the binding of the GCN4 mimic to CRE and furan-modified CRE variants CRE1, CRE2, and CRE3. Lanes 1 to 5 contained increasing amount of protein mimic, respectively, 0, 0.42, 0.84, 1.25, and 1.67 μM peptide versus 0.167 μM DNA duplex in lanes 1, 2, 3, 4, and 5.

The result of the furan oxidation cross-link reaction for the DNA-protein mimic complexes as analyzed on RP-HPLC is shown in Figure 4.41. The top chromatogram shows the reaction mixture containing duplex CRE:CRE comp and 5 equivalents of GCN4 mimic. Under the analysis conditions, only ODNs show up in the chromatogram. After addition of 4 equivalents of NBS to chemically cause furan oxidation, followed by NaCNBH$_3$ reduction to shift the equilibrium and stabilize the formed Schiff base,[212] the bottom chromatograms were obtained. In all cases, the furan-modified ODN disappears, while complements remain unmodified. For CRE1 no sharp new signal can be observed, only some impurities originating from degradation of the furan-modified ODN. This is in accord with the earlier observations of Ebright and co-workers, who showed a non-nucleophilic alanine residue to be closest to T-4, based on photo-cross-linking with bromouracil mutations.[213] However, for CRE2, a new small but clear signal appears. The new signal is clearly most important for CRE3, featuring closest proximity between the furan-modified nucleotide and lysine. A clear relation between cross-link yield and nucleophile proximity is observed. Although in all cases an A base is located opposite modified nucleoside **nu5**, different product yields are obtained for the three duplexes, with a maximum yield (HPLC based) of 50% in the case of close interaction. This clearly indicates the formation of a distance dependent DPC. The complementary ODN was not consumed during the reaction in contrast to what is observed during ICL formation with other studied furan-modified nucleosides (see Appendix).[134, 135, 138, 139, 157] Furthermore, mutation of the complementary ODN to include **nu5**-C and even **nu5**-T (which has no exocyclic amine for ICL formation) base pairs, has a minor influence on the product yield (for CRE3 see Fig. 4.42, for CRE1 and CRE2 see experimental part). Observed small variations are believed to be due to small

---

Figure 4.41 *(facing page)*: **Cross-linking experiments.** The top part of the figure illustrates the positioning of oxidized **nu5** in the DNA with respect to the lysine residues protruding from the protein mimic into the major groove. Images obtained from the pdb structure 2DGC using the Discovery Studio visualization software from Accelrys, the color code for the bases is as follows A = red, T = blue, G = green, and C = magenta. The bottom part shows the HPLC traces before and after furan oxidation for the duplexes of CRE1, CRE2, and CRE3 bound by the GCN4 protein mimic (which is not visible in the HPLC chromatograms).
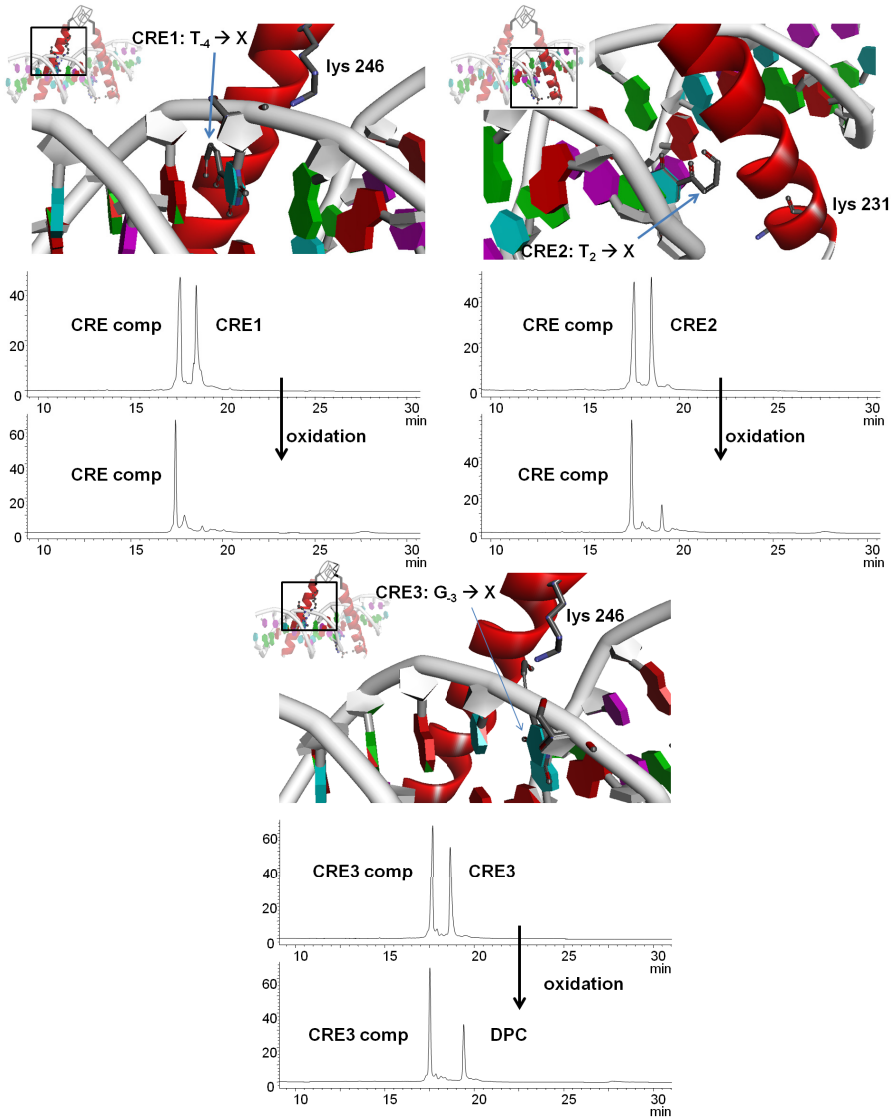
changes in the binding stability and positioning of the protein mimic to the slightly different target DNA duplex. To confirm the formation of a cross-link between the modified oligonucleotide and the target peptide, the product was isolated by purification on HPLC (Fig. 4.43).
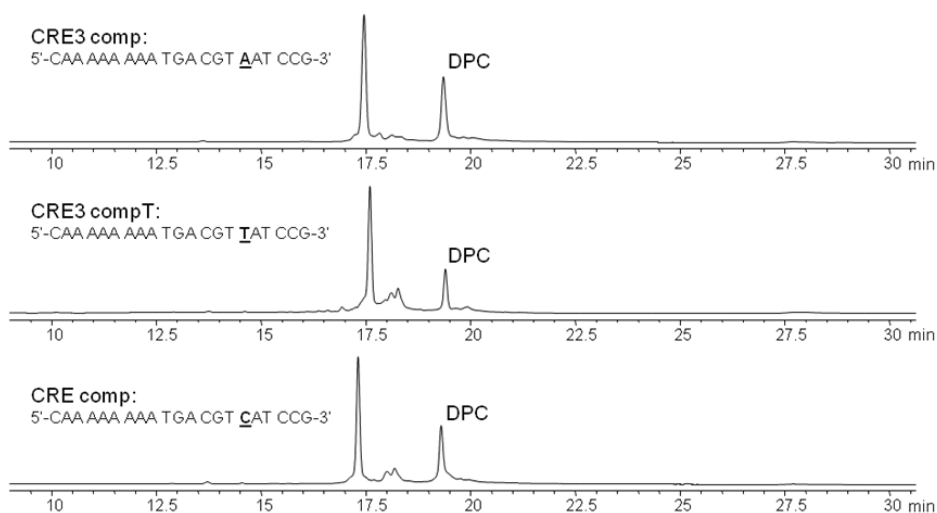
Figure 4.42: RP-HPLC chromatograms of the DPC reaction mixture with the duplex of CRE3 with the complementary ODN containing A, T and C opposite to the modified nucleotide, from top to bottom
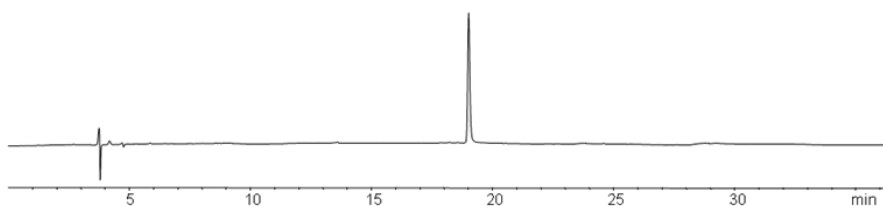
Figure 4.43: RP-HPLC chromatogram of the purified DPC of CRE3

The cross-linked product is clearly present in the mass spectrum (Fig. 4.44, DPC). Still, despite the HPLC purification, additional compounds can be observed in the mass spectrum. These are minor impurities eluting at the same retention or result from partial degradation of the sample during mass analysis. These impurities are strongly overrepresented in the mass spectra because

the ODN-peptide conjugate is very difficult to ionize. ODNs and peptides have completely opposite ionizing characteristics, which mutually impede their ionization. This is especially important in cases where both peptide and ODN are equally contributing to the total mass, like in this case.[214] Therefore, every contaminant that is present in the sample will ionize better than the target, contributing to and even dominating the mass spectrum. The sample was therefore subjected to nano LC-MS, separating the sample just before mass analysis. Deconvolution of the spectra obtained during elution of the reaction product gave the spectrum shown in Figure 4.45, with the target mass (10431 Da). The observed mass of the formed cross-linked product corresponds to the calculated molecular weight of the reduced conjugate of CD-modified peptide and furan-modified ODN (10431 Da) (Fig. 4.46).
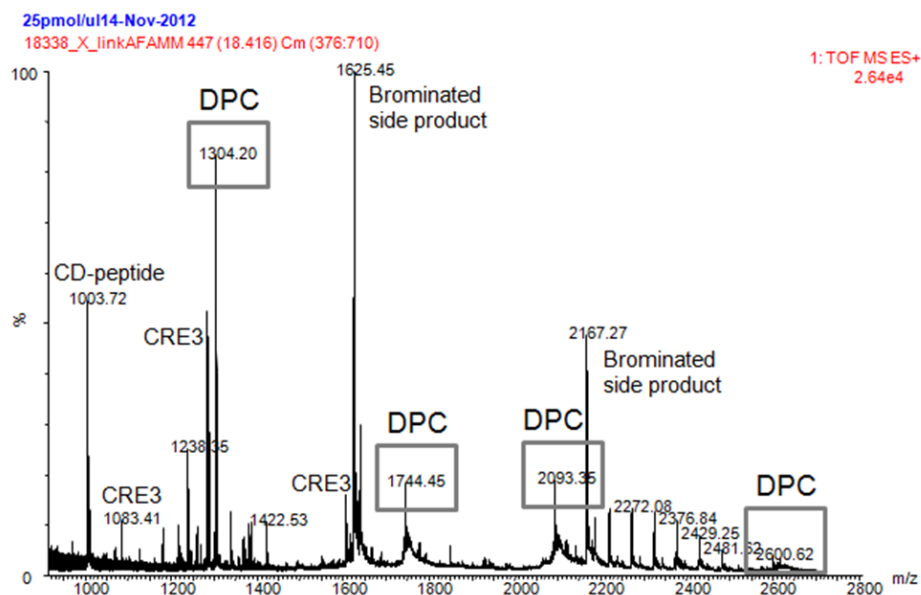

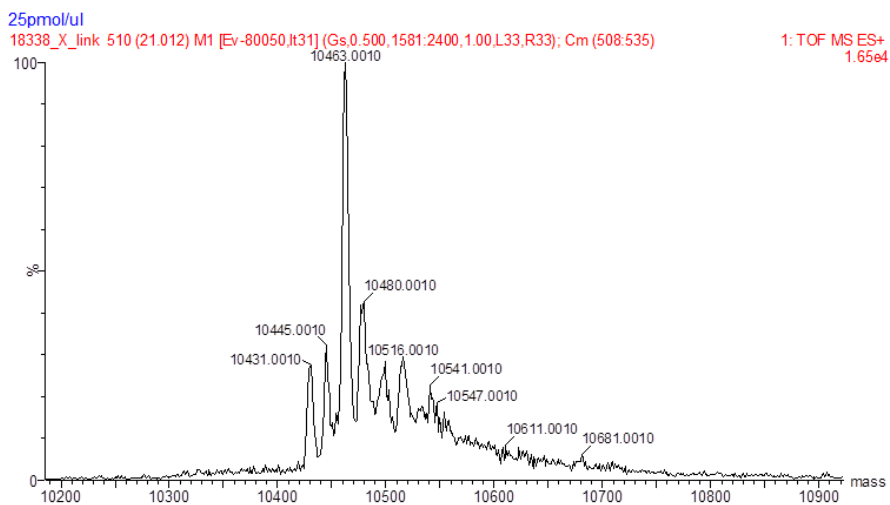**Figure 4.44: MS spectrum of the purified DPC of CRE3**

Figure 4.45: Spectrum obtained by deconvolution of the spectra during elution of the product
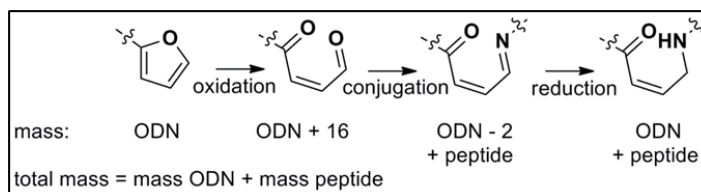

Figure 4.46: Illustration of mass changes during DPC formation

To verify the necessity of proximity between the generated enal functionality and the target lysine residue, the cross-linking experiment with CRE3 was repeated at high salt concentration. While duplex stability is generally enhanced under such conditions,[215] DNA-protein complex formation will be compromised.[216] The experiment under these conditions indeed resulted in hardly any cross-link formation, owing to the destabilization of the DNA-protein mimic complex (Fig. 4.47B). Similarly, a control experiment in the absence of protein mimic did not yield cross-linked product (Fig. 4.47C). This also again confirms the absense of ICL formation with **nu5**. (This experiment was also performed with CRE2, see experimental part for data.)
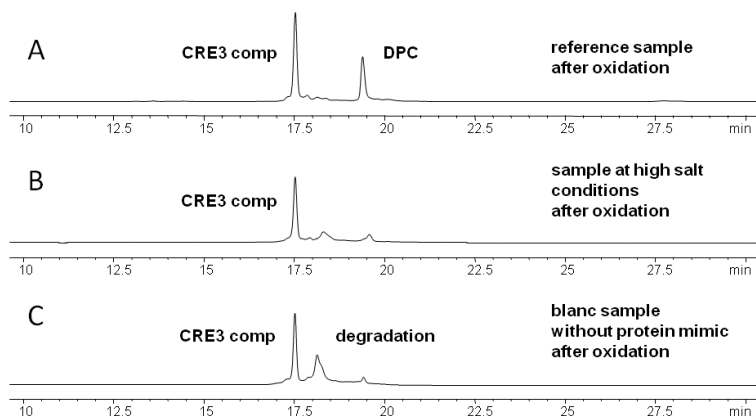
Figure 4.47: HPLC-traces of the cross-link reaction mixture of CRE3:CRE3 comp after NBS addition A) under standard conditions; B) under high salt conditions; and C) in the absence of protein mimic.

On reviewers request, additional control experiments were performed in the absence of a peptide-oligonucleotide duplex complex, using a random peptide (Fig. 4.48) or just the single stranded furan-modified oligonucleotide (Fig. 4.49). Complex formation and thus proximity is clearly prerequisite for efficient DPC formation, as no major product peak can be observed in the HPLC chromatogram with random Rev peptide. The small amount of product formed can be explained as originating from the proximity that results from unspecific electrostatic interactions between the positively charged peptide and the negatively charged DNA.

In case there is only the single stranded oligonucleotide modified with furan, no interaction with the mimic occurs and the furan-modified oligonucleotide degrades.
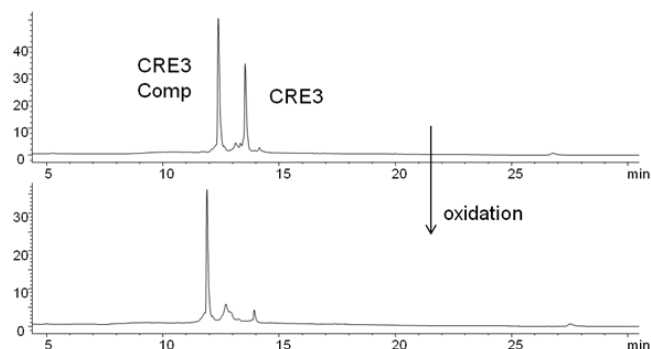
Figure 4.48: RP-HPLC chromatograms of the cross-link experiments of CRE3 and its complement with a random Rev peptide, before (top) and after (bottom) addition of NBS.
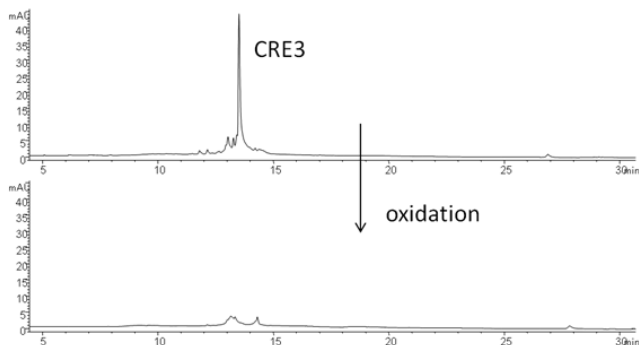
Figure 4.49: RP-HPLC chromatograms of the cross-link experiments of CRE3 without its complement with the GCN4 protein mimic, before (top) and after (bottom) addition of NBS.

Clearly, if no proximal nucleophile for selective reaction is present, random degradation of the modified ODN takes place, deactivating the cross-linking moiety. Furthermore, omitting the reduction step, the cross-linked complex was shown not to be stable (Fig. 4.50).
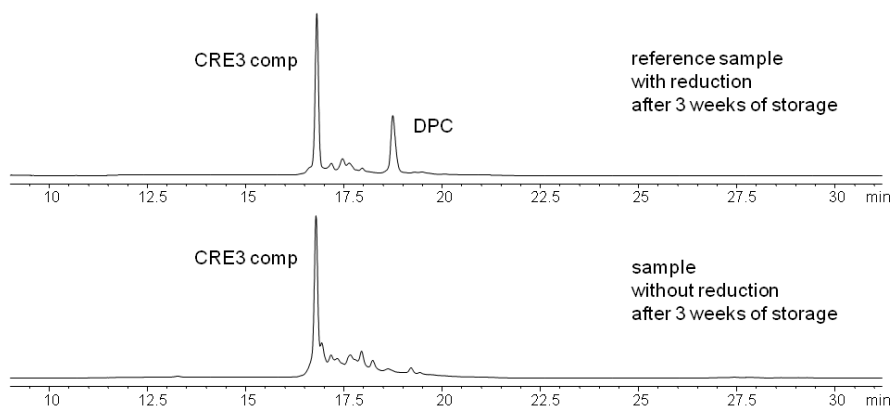
Figure 4.50: HPLC-traces of the cross-link reaction mixture of CRE3:CRE3 comp after NBS and NaCNBH$_3$ addition and 3 weeks of storage (top) without NaCNBH$_3$ addition and 3 weeks of storage (bottom).

Current findings can be rationalized mechanistically as illustrated in Figure 4.51. The major product of furan oxidation with NBS is the highly reactive oxobutenal, which degrades in the absence of proximal nucleophiles. The degradation mechanism is unclear and difficult to study due to the reactivity of the formed intermediate. We believe that this reactivity allows the aldehyde to react in many different ways. Besides that, non-reacted aldehyde-modified DNA can be retained on RP-silica, explaining the absence of clear product formation during HPLC analysis. A furan brominated species, resulting from competitive electrophilic aromatic substitution, is only formed as a minor side product. In the presence of proximal primary amines from amino acid residues, the reactive aldehyde functionality forms a Schiff base that can be selectively reduced with NaCNBH$_3$ to generate a DPC.[183, 188, 194–197, 203] Alternatively, in the case of a differently positioned furan as shown in earlier work focusing on interstrand cross-linking,[134, 135, 138, 139, 157] stable adducts can be formed with proximal nucleic acid bases, generating ICLs. It is clear that carefully planned differential positioning of the furan moiety on the nucleoside allows switching from ICL to DPC formation.
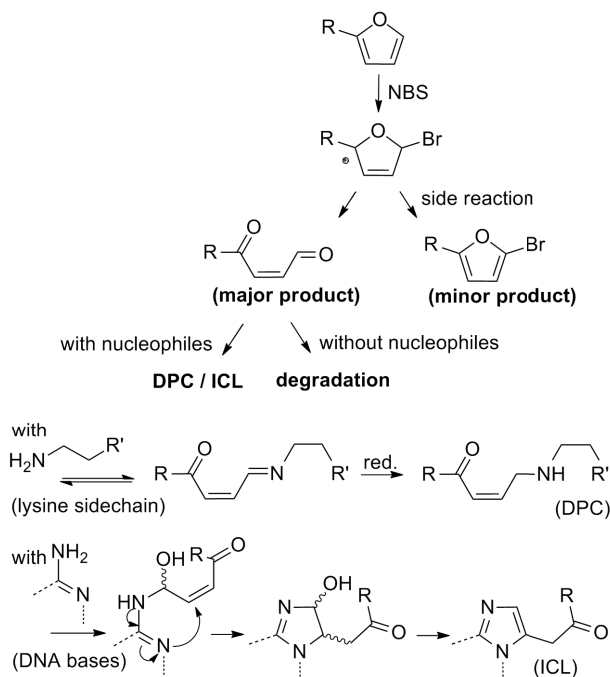
Figure 4.51: Mechanism of the furan-oxidation methodology for the formation of DPCs

Conclusion In conclusion, we have illustrated the use of furan as inducible DNA to protein cross-link forming moiety, inspired by the toxicity of furan. The stability and general availability of the small aromatic heterocycle allows for easy incorporation into ODNs. Selective oxidation then generates a very reactive aldehyde that engages in a specific reaction with proximal nucleophiles or degrades in their absence. This renders this type of furan containing ODNs particularly useful as interaction or distance scanning probes for the study of DNA-protein interactions. We've shown that positioning of the furan moiety on the base, allows **nu5** to selectively produce DPCs in high yield. As the applicability of furan-based cross-linking has recently been clearly established in a biological context by us[139] and others[204] for nucleic acid targets, current expansion toward amino acid side chains in protein targets now considerably broadens the application area of the methodology for cross-linking and bioconjugation.

Additional experiments A small amount of full lenth GCN4 dimer protein was offered by Prof. Morii. This sample was used for a real DNA-protein cross-linking reaction. However with current analysis methods the reaction mixture could not be analyzed. The protein blocked the DNA column, and also its analysis on typical PAGE gels did not yield results from which conclusions could be drawn. The addition of sodiumdodecylsulphate (SDS) helped to resolve the gels, but not enough sample was available to optimize the analysis conditions for reportable results.

DPC formation was also attempted with **nu4**, which was introduced earlier as the other base modified nucleoside, not causing ICL formation due to the positioning of furan in the major groove. As the base, position and linker is different, visualisation identified a different residue to be most suited for substitution (CRE4 = ODN4.1.1). The substitution of this adenosine did not disturb duplex formation or peptide binding (Fig. 4.52).
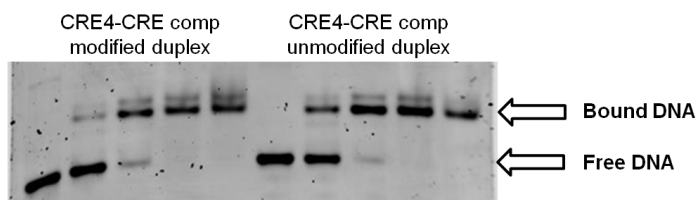
Figure 4.52: EMSA of the binding of the GCN4 mimic to CRE and furan-modified CRE variant CRE4 Lanes 1 to 5 contained increasing amount of protein mimic, respectively, 0, 0.42, 0.84, 1.25, and 1.67 μM peptide versus 0.167 μM DNA duplex in lanes 1, 2, 3, 4, and 5.

As judged from the HPLC traces (Fig. 4.53A), the furan oxidizes without the formation of a new species. PAGE experiments showed a very small amount of cross-linked products (Fig. 4.53B). This behaviour is very similar to the behaviour seen earlier with **nu4** in ICL formation. Unfortunately, it thus looks like DPC formation is not a general mechanism with base modified nucleosides, though additional sequences should be tested to prove this.
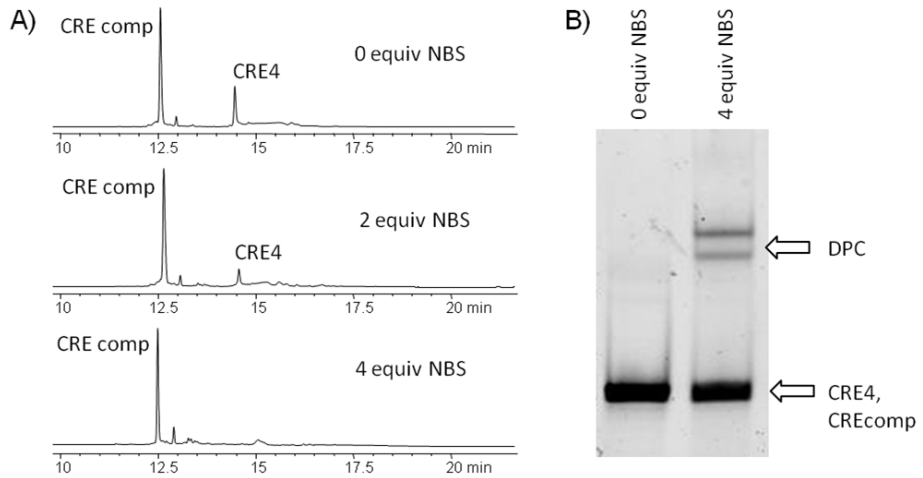
Figure 4.53: Results of the DPC formation reaction with nu4. A) HPLC traces after increasing addition of NBS and NaCNBH$_3$ for duplex of CRE4 bound by the GCN4 protein mimic (which is not visible in the HPLC chromatogram). B) Denaturating PAGE of the same reaction mixture before and after the addition of 4 equiv NBS and reduction with NaCNBH3 in lanes 1 and 2 respectively.
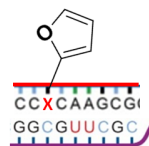
4.4 RNA interstrand cross-linking

Though the inability to cross-link to double stranded DNA was initially considered a negative result and a disadvantage, over time this has been more and more considered as a feature that could be used for structural elucidation in complex systems and to distinguish single from double stranded DNA e.g in topoisomerases. But therefore the hypothesis has to be confirmed. Currently, thesis student Andreas Baudot is analyzing DPC formation with a furan-modified Rev peptide to its binding site RRE (Rev Responsive Element) RNA. This model system is based on the strong interaction between a short synthetically accessible peptide and an RNA hairpin,[217, 218] that is highly distorted with flipout bases, a loop and a mismatch as potential cross-link targets (Fig. 4.54). This is also in interest for our collaboration with Prof. Morii, as the RRE Rev interaction is central in his ribonucleoprotein sensors. Their stabilization through a covalent bond is desirable.[219]
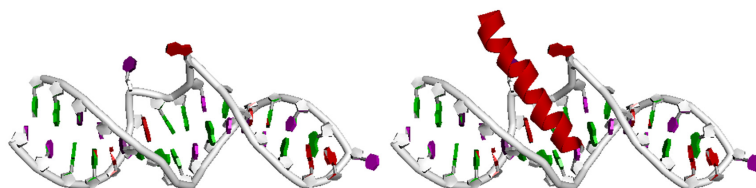
Figure 4.54: RRE-Rev interaction The distorted RRE hairpin, with a G-G mismatch, 2 flipout bases (bulges) and a loop (left) and with the Rev peptide binding in close proximity to these distortions (right)

As RNA had never been used in our group before, nor was it thus ever used in furan cross-linking reactions, an exploratory study was conducted on RNA interstrand cross-linking, which is actually very relevant on its own in view of a series of potential applications.

We here report on the furan oxidation methodology for interstrand cross-linking of RNA duplexes, which have a different structure and are more stiff, reactive and labile than their DNA counterparts. Through this mildly inducible approach, natural unmodified RNA can be selectively cross-linked in high yield. The method therefore has direct applications in the increasing number of RNA based technologies.

Introduction   While protein targets have dominated the pharmaceutical industry for over 100 years, targeting RNA is a promising alternative. With the recent FDA approval of Mipomersen[220], the second antisense drug of ISIS to reach the market after Fomivirsen,[221] the possibilities in this new area of drug discovery have been validated.[222] It was previously shown that the antisense effect obtained through a steric blockage can be enhanced by a covalent linkage, transforming the transient inhibitory effect to a permanent one.[223] Moreover, it is becoming increasingly clear that genetic expression is regulated by a large number of RNAs next to mRNA (lncRNA, miRNA, siRNA, snoRNA) which are the subject of current investigations and constitute new therapeutic targets.[167, 168] Cross-linking is further a well-established and widely used method to obtain secondary and tertiary structural information.[224] To study new RNA interactions, novel methods are continuously being developed, like CHART (capture hybridization analysis of RNA targets) with purification based on hybridization.[225] Such methods could benefit from covalently reinforced hybridization for the fish out purposes. Finally also RNA aptamers (nucleic acid

based selective binders obtained through in vitro selection) like Pegaptinap,[226] are emerging as therapeutics. Currently, these go through a long development process of post-selection modification, to increase metabolic stability, without compromising binding affinity. Resistance to nucleases could be increased by cross-linking,[165, 166] next to stabilization of the secondary structure.

Different methodologies for cross-linking of RNA duplexes have been reported. Cross-linking through selective click reactions[227–229] or through disulfide formation[230] is less relevant for current applications, as these methods require modification of both strands of the duplex. To attack a natural and thus unmodified target, a reactive moiety e.g. halocarbonyl[231], aziridine[232] or other strained rings[233] is required. These modifications however, typically lack sufficient stability resulting in low yields and off-target reactions.[234] Better results are obtained when a reactive cross-linking moiety is induced from a stable precursor. Psoralens have for this reason been extensively used. They react with thymine residues upon UV irradiation,[235] but UV light is highly energetic and can generate collateral damage. An alternative cross-linking strategy, which is induced by oxidative deprotection of a reactive vinyl group, was described by Sasaki.[236]

Our group developed a cross-linking methodology for DNA-DNA duplex interstrand cross-linking, inspired by the metabolic activation of furan (Fig. 4.55).[134–136, 138, 157] We've shown that furan moieties can be easily incorporated in oligodeoxynucleotides (ODNs) in view of the stability and wide commercial availability of furan derivatives. Furan-modified ODNs can be activated by oxidation to a very reactive oxobutenal that quickly reacts with the exocyclic amine of the opposite base resulting in selective and high yielding interstrand cross-link (ICL) formation. To increase its applicability and usefulness, the furan oxidation cross-linking approach should be applicable to RNA duplexes. However in view of the different structure and reactivity, this is counter intuitive. N-Bromosuccinimide (NBS), used to oxidize the furan moieties, has indeed been reported to react with the electron-rich C5 position of uracil nucleosides.[237] Furthermore, the extra 2'-OH, renders RNA considerably more nucleolyticaly labile and additionally influences the nucleotide

puckering. Nucleotides assume the north (C3'-endo) pucker, rather than the south (C2'-endo) pucker of deoxynucleotides. As a result an RNA-RNA duplex adopts an A-type duplex shape, which is different from the typical B-type duplex for DNA-DNA duplex.[238, 239] This different shape causes substituents to be positioned differently and thus possibly behave differently.[226] Lastly, RNA and especially 2'OMe RNA-RNA duplexes[238] are thermally more stable than DNA duplexes.[239, 240] Perturbation of such a more stiff duplex to achieve ICL is expected to be challenging.

Results and discussion For the experiments a 12-mer 2'OMe RNA probe (ON4.4.1) was used (Fig. 4.55). The different behavior of DNA versus RNA is already illustrated by the stability of the respective duplexes (Table 4.2): DNA-DNA ($T_m = 41°C$), DNA-RNA ($T_m = 39°C$) and 2'OMe RNA-RNA ($T_m = 82°C$) of the unmodified ON4.4.1 or ODN4.4.1 and its complement, ON4 or ODN4.
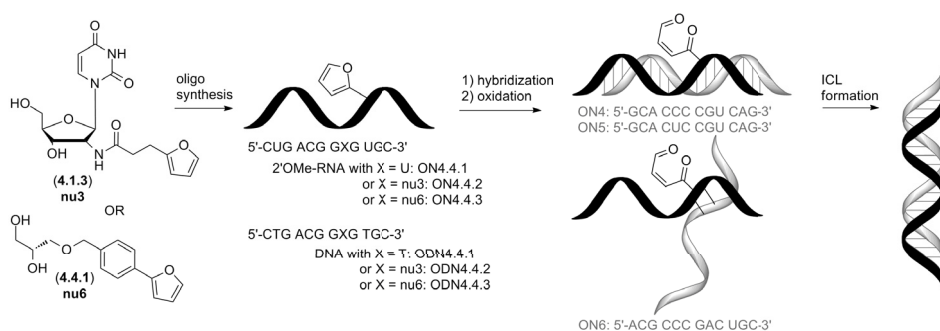
Figure 4.55: Principle of furan oxidation interstrand cross-linking. Furan-modified nucleosides **nu3** (**4.1.3**) or **nu6** (**4.4.1**) are incorporated in an ON or ODN during synthesis. After hybridization with natural RNA complements, the furan is oxidized to a very reactive ketobutenal that will quickly react with the exocyclic amine of the opposite base forming an ICL.

Table 4.2: Duplexes used for cross-linking

duplex type	furan-modified single strand	complement single strand
2'OMe RNA/ RNA	ON4.4.1*/ON4.4.2/ON4.4.3	ON4/ON5/ON6
DNA/RNA	ODN4.4.1*/ODN4.4.2/ODN4.4.3	ON4/ON5/ON6
DNA/DNA	ODN4.4.1*/ODN4.4.2/ODN4.4.3	ODN4/ODN5/ODN6

*these sequences are the unmodified analogs, with U/T
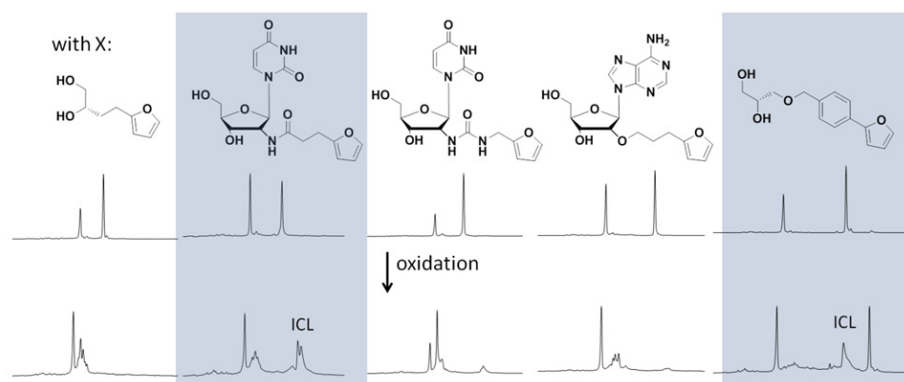
Figure 4.56: Illustration of the initial screen to identify the furan-modified nucleosides most suited for use in RNA ICL formation.

Furan-modified nucleosides **nu3** and **nu6** were selected from a range of furan modifications available in our group after an initial screen (Fig. 4.56). Both were incorporated in the 2'OMe RNA sequence resulting in ON4.4.2 and ON4.4.3 respectively and in the corresponding DNA sequence resulting in ODN4.4.2 and ODN4.4.3 respectively. Incorporation of the furan-modified nucleoside does not compromise DNA-DNA duplex formation. ΔT_m for **nu3** = -4°C and ΔT_m for **nu6** = +8°C). In a mixed DNA-RNA duplex, both modifications are destabilizing (ΔT_m for **nu3** = -8°C and ΔT_m for **nu6** = -2°C). Surprisingly, in the 2'OMe RNA-RNA duplex, severe destabilizations (ΔT_m for **nu3** = -31°C and ΔT_m for **nu6** = -22°C) were observed (a summary of all T_m values is available in Table 4.3). This local destabilization is presumably beneficial for cross-link formation in the otherwise very stiff duplex. In all cases, **nu6** is less destabilizing than **nu3**, probably due to stacking interactions of the phenyl moiety, though it is unable to form Watson-Crick base pairs.[136] To allow

evaluation of the cross-link selectivity, two additional targets were tested. The first, ON5 is identical to ON4 thus complementary to the probe, but with a U opposite to the furan-modified nucleotide or thus without an exocyclic amine to react in an ICL reaction. The other target, ON6 was completely randomized apart from the central CCC opposite to the furan-modified nucleotide.

Table 4.3: Melting temperatures of the different duplexes

	target									
	ON4		ODN4		ON5		ODN5		ON6	
probe	T_m (°C)	ΔT_m (°C)	T_m (°C)	ΔT_m (°C)	T_m (°C)	ΔT_m (°C)	T_m (°C)	ΔT_m (°C)	T_m (°C)	ΔT_m (°C)
ON 4.4.1	82		n.d.		73		n.d.		39	
ON 4.4.2	51	-31	n.d.		53	-20	n.d.		20	-19
ON 4.4.3	60	-22	n.d.		59	-14	n.d.		20	-19
ODN 4.4.1	39		41		40		44		n.d.	
ODN 4.4.2	31	-8	37	-4	34	-6	37	-7	x	
ODN 4.4.3	37	-2	49	+8	38	-2	45	+1	x	

n.d. = not determined

x = no duplex formation

Figure 4.57 shows the result of the cross-link reactions through a denaturing PAGE experiment. It is clear that the desired cross-linking from ON4.4.2 to ON4 occurs in good yield (HPLC based: 42% for ON4.4.2 and 24% for ON4.4.3). Despite the different duplex structure and positioning of the substituent, cytidine is targeted like in DNA. The south pucker of **nu3** that was determined by NMR, is thus likely maintained in the RNA duplex, like it was in the DNA duplex. The cross-link reaction with **nu3** proceeds in much higher yield than with **nu6**. This can be explained by the competitive bromination reaction on the furan

moiety, that occurs in case of an extended aromatic system like **nu6**[136] and clearly generates a significant side product (Fig. 4.58). Other, brominated side reactions, like expected on the C5-uridine were not detected in the cross-linked species or in residual DNA and RNA.
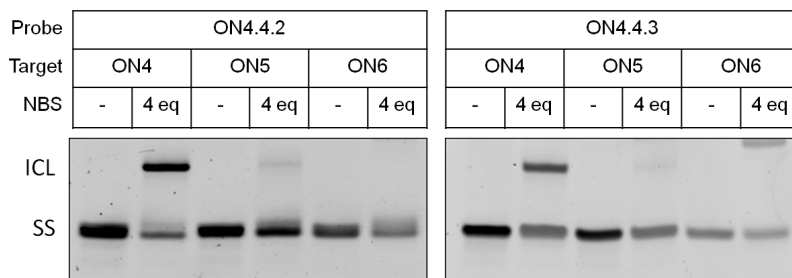
Figure 4.57: Denaturing PAGE results of the cross-link reaction. 2'OMe RNA probes containing furan-modified nucleoside **nu3** or **nu6** (ON4.4.2 and ON4.4.3) were annealed to the targets ON 4-6, after which NBS was added in portions. For each duplex the reaction mixture before and after addition of 4 equiv NBS was analyzed.
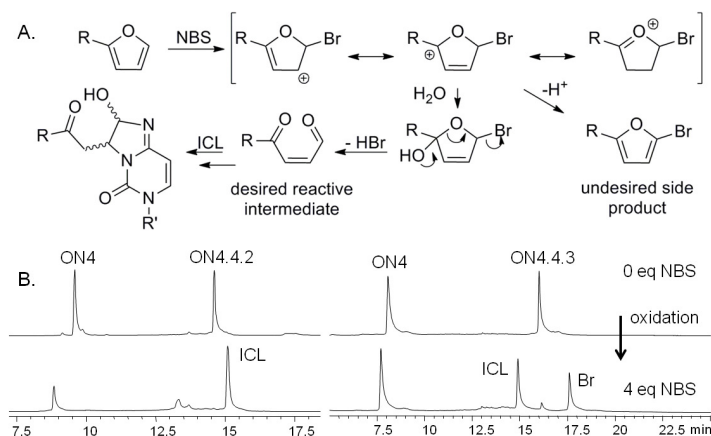
Figure 4.58: A) Oxidation mechanism of nu6-modified furan with NBS and the undesired side reaction leading to bromination. **B) Comparison of the RP-HPLC traces** of the cross-link reaction with **nu3** en **nu6** before and after NBS oxidation. (Br = brominated side product)

The selectivity of the cross-link reaction is remarkable. In case of ON5, when no exocyclic amine is available directly opposite to the modified nucleotide, no cross-linking is observed. This implies that the ICL formation is highly specific for the opposite base and that no reaction occurs with any present neighboring cytidines. Further, as expected also in case of the randomized ON6 that does not form a stable duplex, no cross-linking is observed. (PAGE and RP-HPLC data for RNA-RNA, DNA-RNA and DNA-DNA ICL formation in supporting information). These examples clearly illustrate the requirements for ICL formation: a stable duplex and a proper nucleophile in the base opposite the furan moiety. If not, the probe undergoes self-destruction, thus avoiding off-target reactions. Interestingly, with **nu3** both the yield and selectivity are better than previously observed in DNA-DNA ICL formation (HPLC yield 26% versus 42% in a 2'OMe RNA-RNA duplex), making this a very interesting modification for application in RNA context.

Though NBS oxidation proceeds unexpectedly without damage and can be easily performed, it has limited applicability in vivo. The use of singlet oxygen, generated by low energetic red light irradiation of methylene blue, for oxidation of furan was recently demonstrated by us for DNA ICL formation.[139] It allows biocompatible furan based cross-linking under spatiotemporal control. During the course of the here described experiments our red light activation methodology was used by Summerer for protein-RNA cross-linking,[204] further illustrating its broad applicability. The possibility of red light activation for RNA ICL formation is demonstrated in Figure 4.59. It can be observed that after 15 min of irradiation with **nu3** a considerable amount of ICL is formed (yield based on HPLC: 39%). Increased irradiation times can further increase the cross-link yield, though it also allows for side reactions to occur on the product, visible by the appearance of a smear on the gel. Depending on the application, the irradiation time can be kept short for high selectivity or increased for higher yield. Though there is no NBS bromination side reaction under these oxidation conditions, **nu6** was again found to be lower yielding than **nu3** (yield based on HPLC: 6% for **nu6**). As opposite results were obtained in a DNA duplex context,[240] the influence of different duplex structure on ICL formation and the importance of a proper choice of furan building block for the context is evidenced.
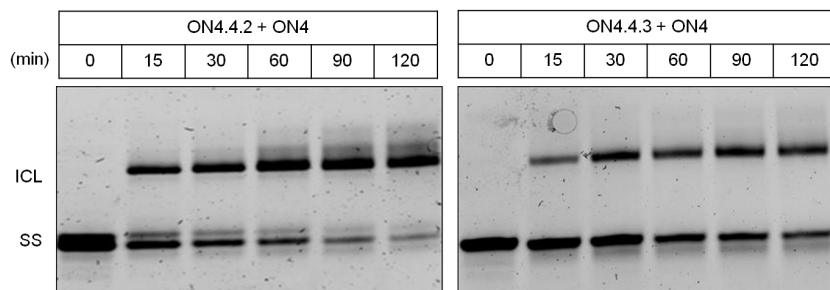
Figure 4.59: Denaturing PAGE results after different irradiation times from the cross-link reaction mixture of ON4.4.2 (left) and ON4.4.3 (right) to ON4 using singlet oxygen.

Conclusion   We here report on the ICL formation between a furan-modified 2'OMe RNA-probe and an unmodified target RNA, using 2 different furan-modified nucleoside analogues. In the current RNA context, both NBS and biocompatible red light irradiation are fast and efficient triggers to induce cross-linking. The furan oxidation methodology can thus be used as an inducible tool for selective cross-linking in stiff and reactive RNA duplexes. Though cytidine is targeted in both DNA and RNA duplex structures, the efficiency of the furan-modified nucleosides was found to be reversed (Table 4.4). Current study illustrates the superiority of building block **nu3** for selective and high yielding cross-linking in an RNA context.

Table 4.4: Cross-link yields (%)

sequences	DNA-DNA		DNA-RNA		2'OMe RNA-RNA	
	ODN4.4.2 ODN4	ODN4.4.3 ODN4	ODN4.4.2 ON4	ODN4.4.3 ON4	ON4.4.2 ON4	ON4.4.3 ON4
NBS, 25°C	17	36*	35	18	34	19
NBS, 37°C	26	28	n.d.	n.d.	42	24
$^1O_2$	15*	57*	n.d.	n.d.	39	6

n.d. = not determined

* Determined based on HPLC analysis at 50°C

Note   Interestingly, by putting all the cross-link yields together, we found that the cross-link area and thus the yield as determined by HPLC analysis is dependant on the column temperature of the RP-HPLC analysis (Fig. 4.60). The area decreases with increasing temperature, probably due to instability of the formed cross-link. All HPLC chromatograms and yields used throughout the thesis are based on analysis at 60°C, unless mentioned otherwise.
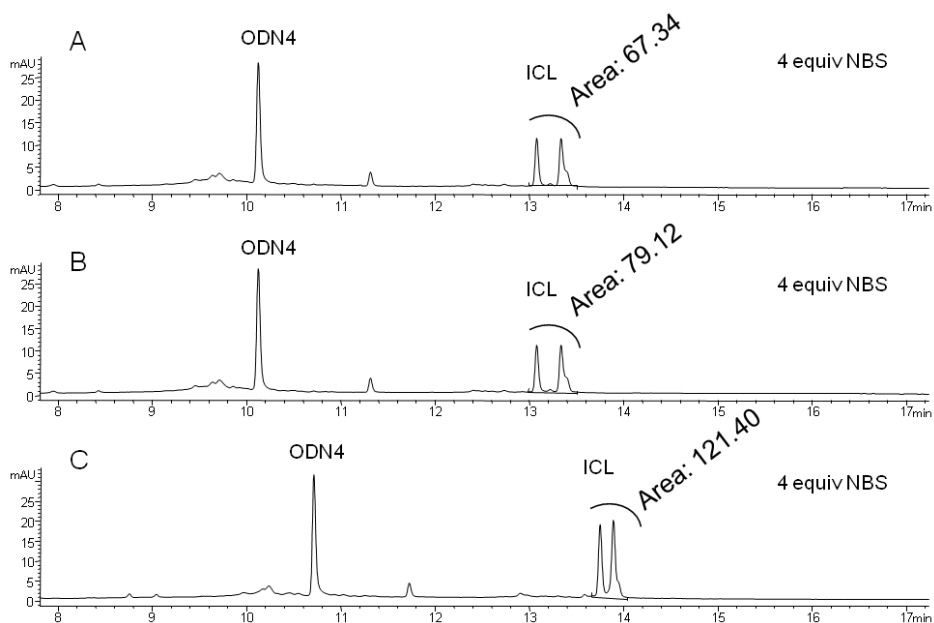
Figure 4.60: RP-HPLC chromatograms of the cross-link reaction mixture of ODN4.4.2 and ODN4 as target with 4 equiv NBS at 25°C, analyzed at A) 60°C; B) 50°C and C) 40°C.
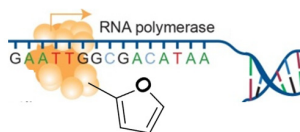
4.5 Furan-modified protein for cross-linking

For a chemist synthesizing peptides, incorporation of unnatural amino acids is quite simple and straightforward: synthesizing a modified amino acid analogue and incorporating it during the stepwise solid phase peptide synthesis. The chemist is however limited by the size of the protein that can be synthesized and has to be creative in protein mimicry with peptide fragments. As can be concluded from the discussion before, the synthetic access to a good model system is not always that easy and the model is inferior to the protein in binding affinity and selectivity, which limits the possibilities for achieving cross-linking. Ideally furan-modified amino acids could also be introduced in proteins, for example in polymerases like illustrated in Figure 4.61 that grip around a single stranded DNA template to assist in the synthesis of a new DNA strand. This system could thus serve to proof that cross-linking to single stranded DNA is possible and to show that our methodology can be extended to proteins.
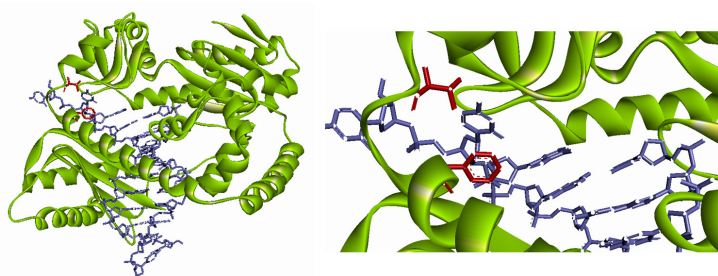
Figure 4.61: Polymerase κ Structure of polymerase κ complexed with DNA, inserting dATP opposite an 8-oxoG DNA lesion (left). Zoom on the structure showing the interaction of the protein with the single stranded DNA template, with 2 residues threonine 44 and phenylalanine 49 (red) in close proximity and suited for replacement with a furan-modified amino acid to achieve cross-linking (right)

*Chemical modification of amino acid side chains in proteins is possible, but can lead to nonselective and nonquantitative derivatization. It therefore often requires extensive mutagenesis to remove redundant labeling sites and/or purification to eliminate labeled contaminants. Incorporation of modified amino acids during protein synthesis is desired. To explore this possibility, I visited the group of Professor Thomas Carell in Munich from November 21st till December 15th 2011.*

4.5.1   Biochemical incorporaton of modified amino acid analogs in proteins

The biological synthesis of proteins is performed by the ribosomes, which translate messenger RNA (mRNA) into proteins using the genetic code. Each triplet of nucleotides (or codon) codes for the incorporation of one of the twenty natural amino acids or the termination of the sequence. There are 3 so-called stopcodons: UAG ("amber"), UAA ("ochre") and UGA ("opal"). Amino acids, the building blocks for the protein synthesis are recognized (with the help of aminoacyl tRNA synthetases) and transported to the mRNA by their transfer RNA (tRNA) that can bind to the codon for the amino acid through an anticodon. The ribosome regulates this process (illustrated in Figure 4.62), as it reads through the mRNA and chains together the arriving amino acids into the target protein.

To incorporate unnatural amino acids, it is possible to chemically misacylate nonsense tRNA with the desired amino acid and perform an in vitro translation. This is however not easy and the protein yield is typically rather low. Working in cell media, is often more easy and higher yielding. One option is to grow bacteria on media in which the encoded amino acid is replaced by a structural analogue, for example an amino acid with its ^{15}N-labeled analogue [241] or methionine by selenomethionine [242]. But, this leads to non-quantitative incorporation throughout multiple sites in the protein and requires the use of auxotrophic strains, which cannot synthesize the amino acid.
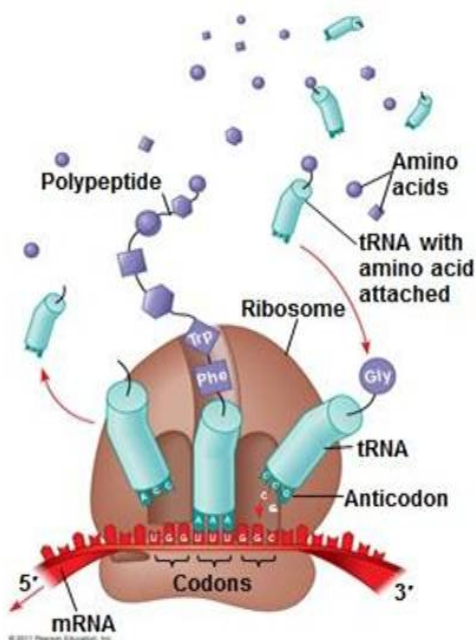
Figure 4.62: Illustration of the biological synthesis of proteins from Campbell and Reese's Biology, copyright ©2005 Pearson Education, Inc.

To really expand the genetic code, there should be a unique tRNA-codon pair, with its own aminoacyl tRNA-synthetase for the modified amino acid. If the tRNA is not recognized by endogeneous synthetases and the synthetase does not recognize endogeneous tRNA, the modified amino acid will only be incorporated at the site with the unique codon.[243] A system this clever has been developed in nature to incorporate the 21st and 22nd amino acids, respectively selenocysteine and pyrrolysine (shown in Fig. 4.63). These are not generated through post-translational modifications like many other unnatural residues, but are directly incorporated into the mRNA on the stopcodons UGA and UAG respectively.[244–246] The pyrrolysyl-tRNA synthetase and its $tRNA_{CUA}$ can be hijacked by chemists for the same purpose.

4.5.2   Amber suppression by pyrrolysine analogs

Amino acids that are similar enough to pyrrolysine, to be recognised by the
pyrrolysyl tRNA synthetase and pyrrolysyl tRNA of Methanosarcina mazei can
thus be incorporated on an amber stop codon in proteins.[247, 248] Comparison
of the successful analogs (Fig. 4.63) with the not so successful ones (Fig. 4.64),
allowes postulating of some typical features for success:

- clearly a suitable spacer (butylene) is required, with a properly sized and
  shaped hydrophobic substituent (not necessarily cyclic) that can be accom-
  modated in the hydrophobic pocket of the enzyme

- a hydrogen bond acceptor (N or O) is also required in the vicinity of the
  Nε-carbonyl group. It is not required to have it in a ring, as carbamates are
  well incorporated, but if it is in a ring the placement is important

- the stereocenter of the side chain is clearly important as the proline and
  THF derivatives are well incorporated, but their isomers are not

- aromatic moieties (thiophene, phenylalanine) are not incorporated

If the similarity is not good enough for substrate recognition by the synthetase,
the latter can be mutated to allow more diverse pyrrolysine analogs, with typically
a smaller or bigger hydrophobic moiety (Fig. 4.65). This work was pioneered
by Yokoyama and Chan. Yokoyama mutated Tyr384 of the synthetase to Phe
and Tyr306 to Ala for efficient incorporation of Nε-(o-azidobenzyloxycarbonyl)-
L-lysine [249], Emine from the Carell group mutated Tyr384 to Phe too, but also
Tyr306 to Gly and Ile401 to Arg for the incorporation of a norbornene carbamate-
L-lysine [250].
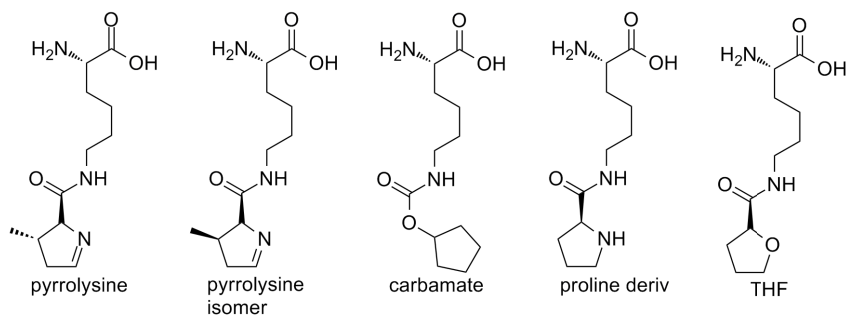
Figure 4.63: Residues well accepted by the pyrrolysine synthethase and tRNA and thus successfully incorporated in proteins
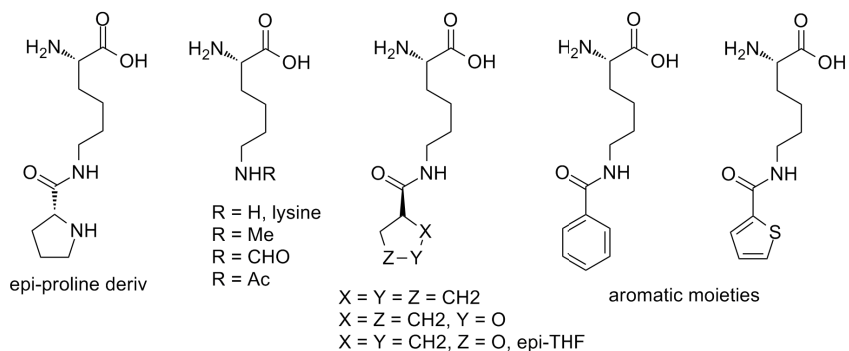
Figure 4.64: Residues not well accepted by the pyrrolysine synthethase and tRNA and thus not successfully incorporated in proteins
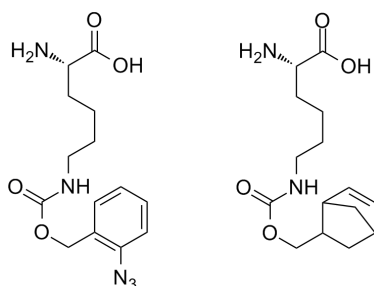
Figure 4.65: Residues accepted by mutated pyrrolysine synthethase and thus successfully incorporated in proteins

4.5.3 Synthesis of pyrrolysine-like furan-modified amino acids

Two furan-modified lysines (**aa4** and **aa5**, depicted in Fig. 4.17) were synthesized for their incorporation during peptide synthesis. **aa4** (4.5.1) seems conform with the required features for recognition by the pyrrolysine synthetase. Morover this amino acid was unsuited for the harsh conditions of SPPS, but could be suited for more mild incorporation in culture. For this purpose, removal of the Fmoc protecting group was required. This was achieved by dissolving the product in DMF with 10% piperidine at room temperature, as illustrated in Figure 4.66. Immediately the product falls out of the solution as a white powder, which was filtered after 30 min with yields close to 100%.
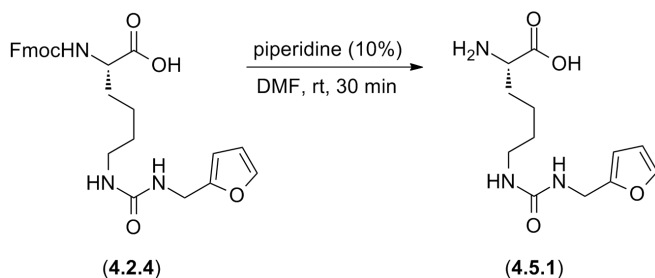
Figure 4.66: Fmoc deprotection to obtain aa4 (4.5.1) 10% piperidine in DMF for 30 min at rt (yield 95%).

Additionally a new furan-modified amino acid **aa6** (4.5.2) was synthesized as illustrated in Figure 4.67, showing stronger structural resemblance to pyrrolysine (shown in Fig. 4.63). Like before, the synthesis started from the commercial Fmoc-Lys(Boc)-OH (4.2.6) commonly used for Fmoc solid phase peptide synthesis, by acid Boc deprotection. The furan moiety was introduced by amide formation with furoic acid, activated as pentafluorophenyl ester (4.5.3). Also similar as in the previous synthesis, the reaction sequence was completed by a Fmoc deprotection.
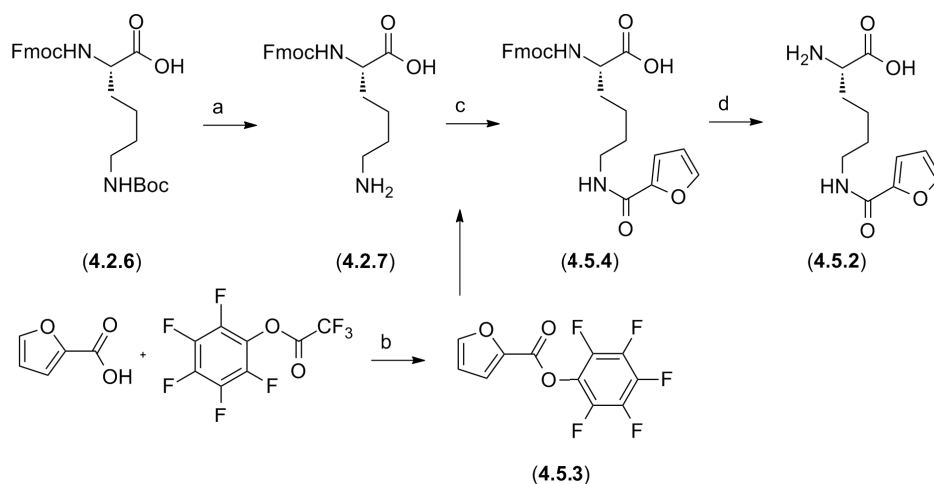
Figure 4.67: Synthesis of aa6 (4.5.2) a. TFA/DCM 1/1 for 1 h at rt; b. 1 equiv furoic acid, 1.2 equiv pentafluorophenyltrifluoroacetate and 1.2 equiv pyridine in DMF at rt for 1 h; c. 1.2 equiv furoic acid, 1.2 equiv HOBt and 1.2 equiv DIPEA in DMF at rt for 2 h; d. 10% piperidine in DMF for 30 min at rt (overall unoptimized yield 83%

The two furan-modified pyrrolysine analogues with different linker length (**aa4** and **aa6**) were both synthesized in an easy 3 step synthesis that is high yielding and easily scalable. Such short synthesis is required for amino acids to be incorporated in proteins as they have to be added to cell cultures in high concentrations of 1-10 mM.

4.5.4 Initial evaluation with Yellow Fluorescent Protein

The incorporation of unnatural amino acids by the pyrrolysyl-tRNA synthetase and its tRNA$_{CUA}$ in Yellow Fluorescent Protein (YFP) can be easily evaluated by monitoring its fluorescence with confocal microscopy. Without incorporation of the amino acid at the amber codon position engineered in the middle of the protein, this stop codon causes termination of the protein. This results in a truncated, non-fluorescent protein. The obtained fluorescence intensity thus parallels the incorporation efficiency. The incorporation of both amino acids was tested against the wild type synthetase and 2 of its mutated variants, by Yokoyama (Tyr384Phe-Tyr306Ala) and one by Emine of the Carell group (Tyr384Phe-Tyr306Gly-Ile401Arg).

Cultures with 0, 2, 5 and 10 mM of the furan-modified amino acid were ex-
amined. The results are shown in Figure 4.68. Unfortunately the wild type
synthetase could not incorporate **aa4**, but somewhat unexpected especially in
view of the described bad results with the thiophene-modified amino acid,[248]
aa6 was incorporated by the wild type synthetase in good yield. The mutated
synthetases gave lower yields. The Yokoyama mutant incorporated both **aa6** and
aa4, and again **aa6** was better than **aa4**, but in lower yield than the wild type.
Emine's mutant had very low yield with **aa4**, **aa6** was not incorporated. For more
confocal fluorescent images, see the experimental part.
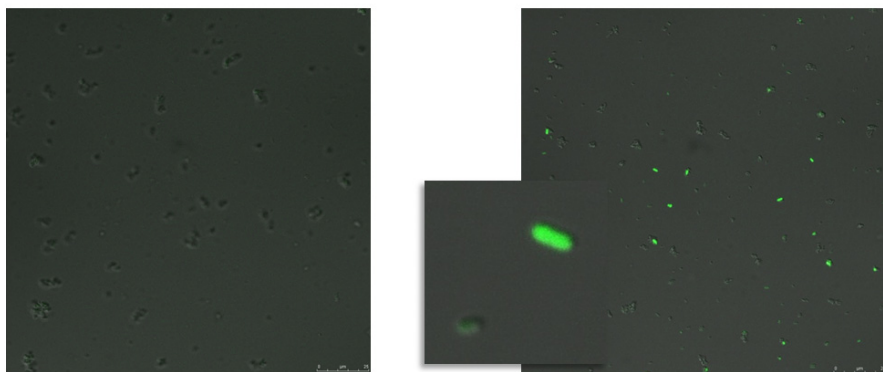
Figure 4.68: Confocal microscope images The incorporation efficiency of **aa4** (left)
and **aa6** (right, with zoom) by the wild type synthetase in YFP at 5 mM.

Based on these results, it was decided to continue with the wild type syn-
thetase and **aa6**. This was next used with a dual plasmid system: one plasmid
constitutively expresses the synthetase and tRNA, the second plasmid contains
the protein engineered with an amber codon at the desired position controlled
by a lac repressor. This increases the flexibility in protein choice and allows for
higher yields. The expression was again examined at 0, 2 and 5 mM of amino
acid concentration with confocal microscopy. The incorporation was better at 5
mM. Therefore this concentration was selected for further use.

4.5.5   Proof of concept with Thioredoxin A as testprotein

Because YFP is expressed only in low yield, its isolation is troublesome. For characterization, it is more rewarding to express another protein. With the dual plasmid system, this only requires to change the second plasmid with the engineered protein. Thioredoxin A (TrxA), a small (13 kD) and stable electron transfer protein that is expressed in better yield, was chosen for this purpose. It was mutated with an amber stop codon to replace Asn65, which is located at the surface and therefore suitable to proof the furan incorporation and also test subsequent modification reactions (Fig. 4.69).
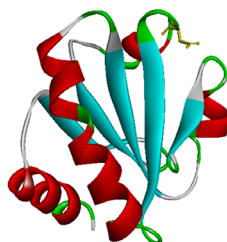
Figure 4.69: Thioredoxin A, with Asn65 indicated in yellow

Due to the limited amount of furan-modified amino acid available in this test phase, the protein expression was limited to a 1 l culture. This yielded 87 μg of protein (an additional 87 μg of protein were obtained in a less pure fraction) after purification over a Strep column. Visualisation by SDS-PAGE revealed that two main products were obtained with masses corresponding to the target protein and its dimer. For a more accurate mass determination a full protein LC-MS analysis was performed. The obtained mass proved the incorporation of the furan-modified amino acid in the protein, as the calculated average mass of the furan-modified oxidized TrxA without the starting methionine is 13039.95 Dalton, and a mass of 13039.7 Dalton was found (Fig. 4.70). The average mass of unmodified TrxA without the starting methionine is 12933.8 Dalton. The correct position of the furan modification in the protein was verified by analysis of a tryptic digest of the product.
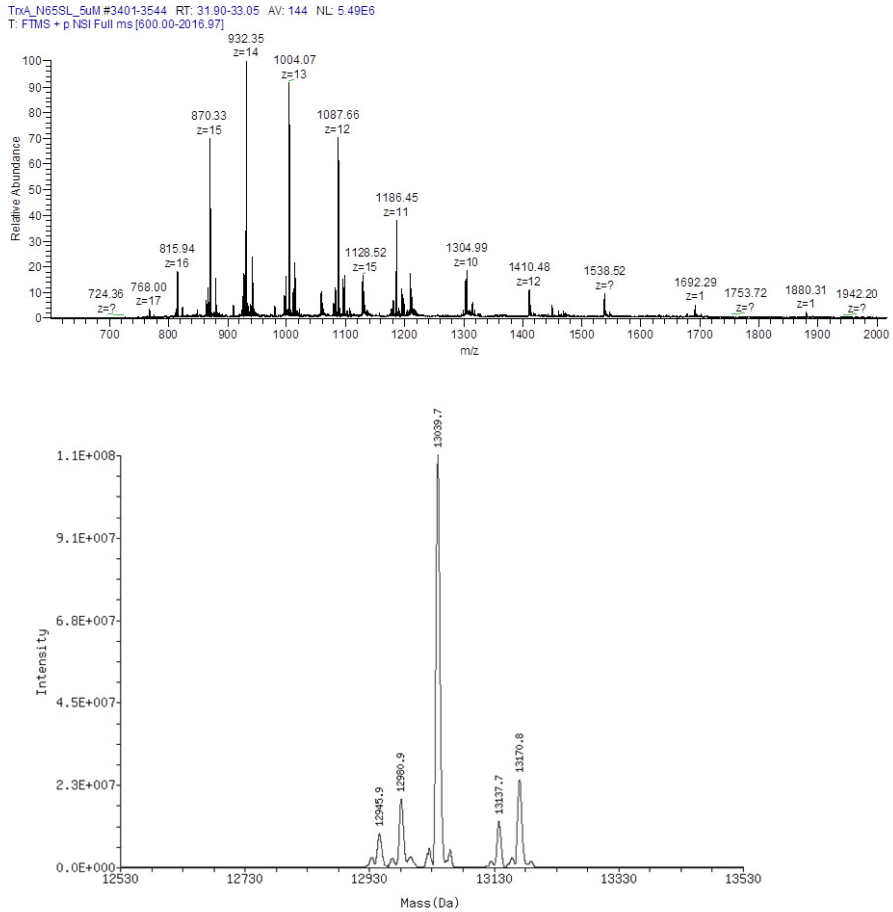
Figure 4.70: Mass spectrum of the furan-modified TrxA (top) and corresponding deconvoluted mass spectrum (bottom)

4.5.6 Labeling reactions

Before starting the cross-linking reactions, it is desired to test the reactivity of the furan moiety and the compatibility of the reaction conditions on proteins. Therefore, labeling reactions as illustrated in Figure 4.71 were performed on the TrxA test protein.
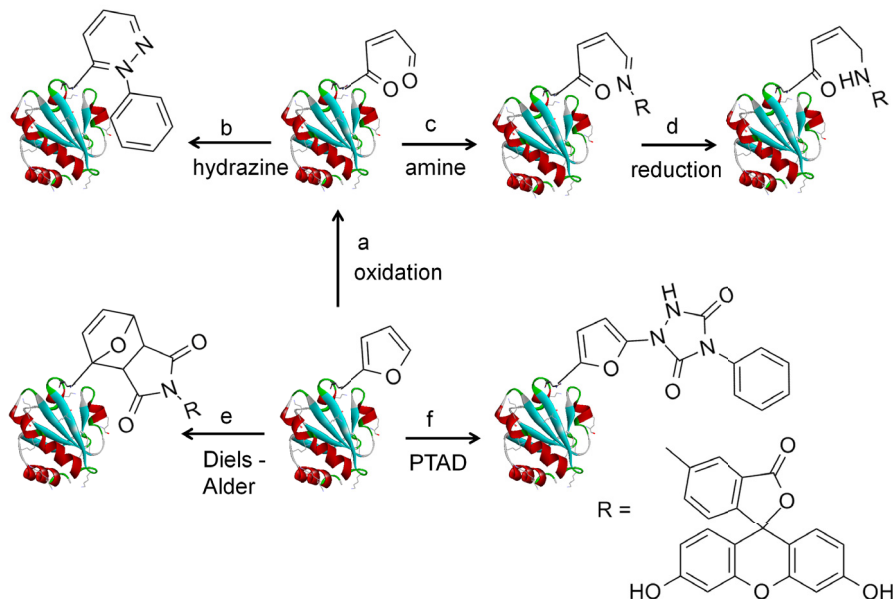
Figure 4.71: Labeling of TrxA Through oxidation of the furan (a) with 10 or 5 equiv NBS and reaction with 50 equiv phenylhydrazine (b) or fluoresceinamine (c) with subsequent NaCNBH$_3$ reduction (d) or through Diels-Alder reaction with 50 equiv fluoresceinylmaleimide (e) or through reaction with 50 equiv PTAD (f)

The furan was oxidized with 10 equivalents NBS as if for a cross-linking reaction in 0.1 M phosphate buffer pH 7, but then quenched with an excess of added phenylhydrazine to form a stabilized aromatic system. Oxidation with 10 equivalents is a large excess, and was initially chosen to guarantee furan oxidation in addition to the oxidation of a number of sensitive residues in the protein (2 tryptophans and 1 methionine). Though most of the protein had reacted, LC-MS analysis was not able to clearly characterize a reaction

product.   Probably oxidation-degradation rather than a labeling reaction had proceeded. The reaction was therefore repeated with only 5 equivalents of NBS and quenched by phenylhydrazine or fluoresceinamine. The latter forms a Shiff base that is stabilized by NaCNBH$_3$ reduction to an amine. Also in this case, no labeled products were detected by LC-MS, but now a lot of starting material was remaining.

To avoid the protein degradation issue caused by the oxidation, a Diels-Alder reaction with 50 equivalents fluorescein-5-maleimide was attempted.   These reactions were performed in 0.1 M phosphate buffer of pH 6, as it has been reported that these reactions proceed more efficiently and chemoselectively at slightly acidic pH. [251] The reaction was analysed by SDS-PAGE and checked for fluorescence (Fig. 4.72). One strongly fluorescent band was observed at the level of the product. This band was cut from the gel and subjected to in gel trypsin digest to identify the formed fluorescent product. The modification could however not be identified. The reaction was repeated to duplicate the result, but this time no fluorescent products were formed. Finally the recently developed furan labeling using 4-phenyl-1,2,4-triazole-3,5-dione (PDAT)[252] was tested. Also in this case, full protein LC-MS analysis was performed, but labeled products could not be identified.

The expression of proteins with furan-modified amino acids was thus, for the first time shown possible. With the observed strongly fluorescent product, promising results were further obtained in terms of protein labeling. Due to strong time constraints, as a result of the short research stay, this result could not be repeated, confirmed or optimized. On the other hand, the electronic properties of the furan in the pyrrolysine analog are strongly altered due to the electron withdrawing effect of the directly attached carbonyl group. This alters the reactivity of the furan so that proposed labeling reactions for electron rich furans are unlikely to proceed. Therefore it would be best to find another furan pyrrolysine analog and repeat these experiments.

1	2	3	4	5
ladder	wild	wild + label	furan	furan + label

1	2	3	4	5
ladder	wild	wild + label	furan	furan + label
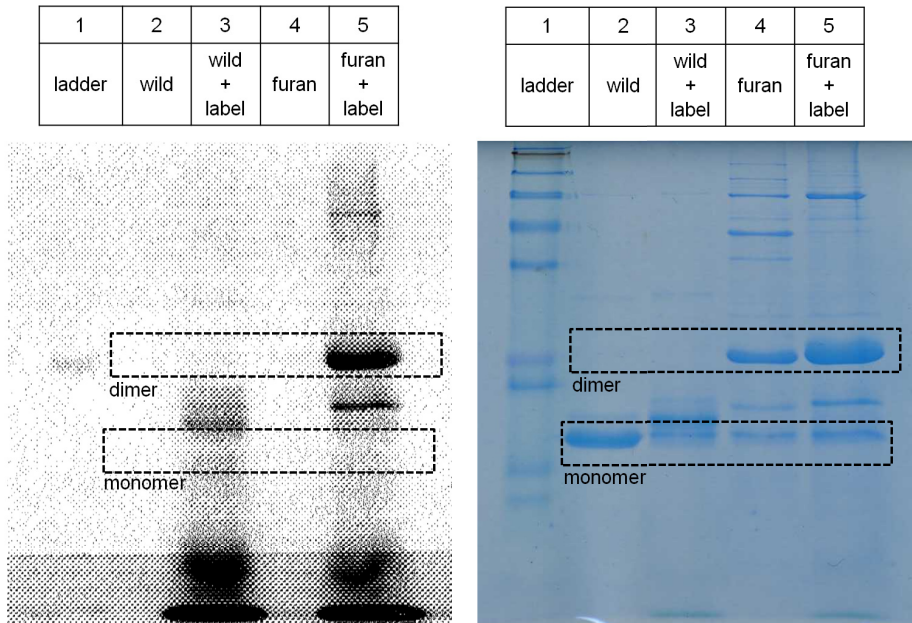
Figure 4.72: SDS-PAGE of labelling reaction with fluoresceinylmaleimide
15% SDS-Laemmli-gels, ran for 1h at 150 V and visualized using blue LED incident light source of 460 nm (left) and with protein staining solution (right) lane 1 contains the mass ladder, lane 2 the wild type TrxA, lane 3 the wild type TrxA after reaction with the maleimide, lane 4 the furan-modified TrxA and lane 5 the furan-modified TrxA after reaction with the maleimide. The TrxA monomer and dimer are indicated. The ladder represents 250, 148, 98, 64, 50, 36, 22, 16, 6 and 4 kDa.

The potential of our furan oxidation methodology was picked up by Prof. Summerer, specialized in genetic encoding of unnatural amino acids. Inspired by one of our furan ICL papers, he carried out similar experiments with an alternative furan-modified lysine derivative (depicted in Fig. 4.73) that could be incorporated into proteins by mutation of the pyrrolysine synthetase. The furan-modified protein was then cross-linked to binding RNA.[204] This further strengthens our previously formulated hypothesis concerning the importance of nucleobase located nucleophile availability, as with these results cross-linking to distorted RNA has been proven possible.
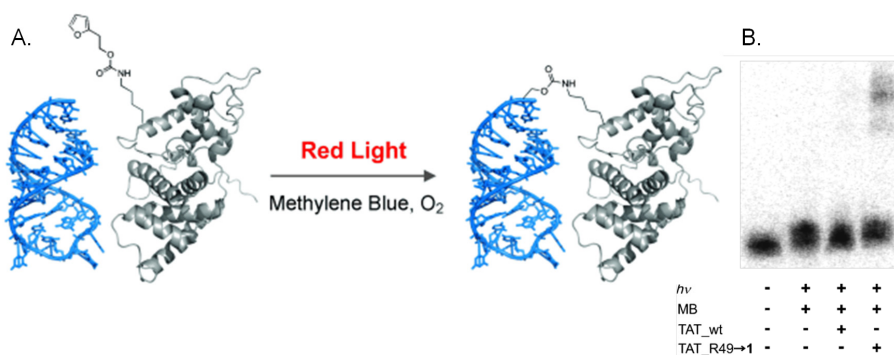
Figure 4.73: Results by Prof. Summerer A) Design of the experiment: incorporation of a furan-modified lysine into Tat-peptide for cross-linking with singlet oxygen by red light irradiation to TAR RNA; B) PAGE results of the cross-link experiments
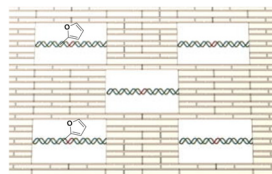
4.6 Visualisation
of furan based DNA interstrand cross-linking

*The highly organised nature of double stranded DNA, has made its targeting by our furan-oxidation cross-linking approach very tricky in the course of this study. However, the programmability of DNA, through its sequence and the Watson-Crick base pairing, is a very interesting property. It allows for the rational design of very selective antisense drugs, based on the genomic information as discussed before. It also allows for the construction of nanostructures. DNA nanotechnology was pioneered by Seeman in 1982.[253] Since then all kinds of DNA scaffolds, origami, devices, motors, [254] walkers and spiders [255] have been developed. A clever property of these assemblies is that each position of the structure has a precise address by means of its sequence code. The Morii group used this to site specifically position proteins on a nanostructure, using sequence-specific DNA-binding zinc finger proteins as illustrated in Figure 4.74.[256] To stabilize this interaction, our furan based cross-link methodology is currently being tested in the Morii group. In collaboration with the Sugiyama group, DNA-origami was used to visualise and optimize the formation of a DNA ICL.*
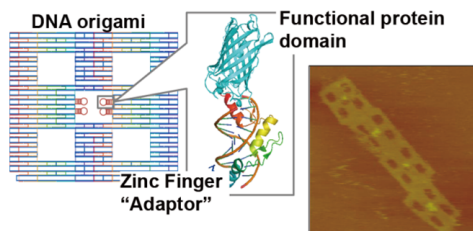
Figure 4.74: Immobilisation of zinc finger proteins on a DNA nanostructure by the Morii group.[256] Cartoon representation of the DNA origami (left); AFM image of the DNA origami: open squares can be clearly observed, small white dots indicate immobilised Zinc Finger protein (right).

4.6.1   DNA Origami

Scaffolded DNA origami was introduced in 2006 by Rothemund [257], using DNA and its well known properties, to fabricate two-dimensional DNA nanostructures of arbitrary shape, like the example in Figure 4.75, by self-assembly following a staightforward and versatile design scheme.
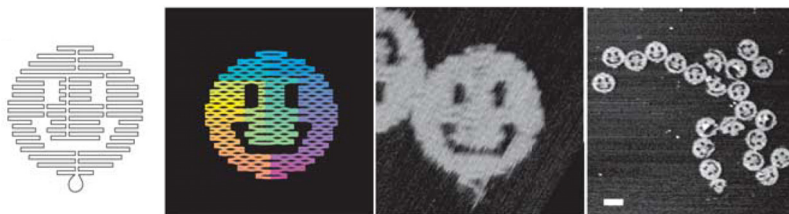
Figure 4.75: DNA origami smiley shape, taken from Rothemund.[257] From left to right: A) Folding path with dangling curves and loops represent unfolded sequence. B) Diagrams showing the bend of helices at crossovers (where helices touch) and away from crossovers (where helices bend apart). Color indicates the base-pair index along the folding path (red is the 1st base, purple the 7,000th). C) AFM image of 165 nm x 165 nm. D) AFM image with scale bar 100 nm

The design process consists of five steps, illustrated in Figure 4.76, which can be strongly simplified with computer assistance, for example with the open-source software package caDNAno.[258] In the first step the desired shape is approximated by a geometrical model, build from an even number of parallel double helices. Secondly, a single, long scaffold strand (M13mp18, a single stranded circular viral genome) is folded along all the double helices, so introducing periodic crossovers between the different helices. In the next step hundreds of staple strands' are generated. These are complementary to the DNA scaffold strand (generating the helices) and create extra crossovers between strands for stability. In the two final steps, the design is examined and refined to minimize strain and to strengthen the structure.[259]
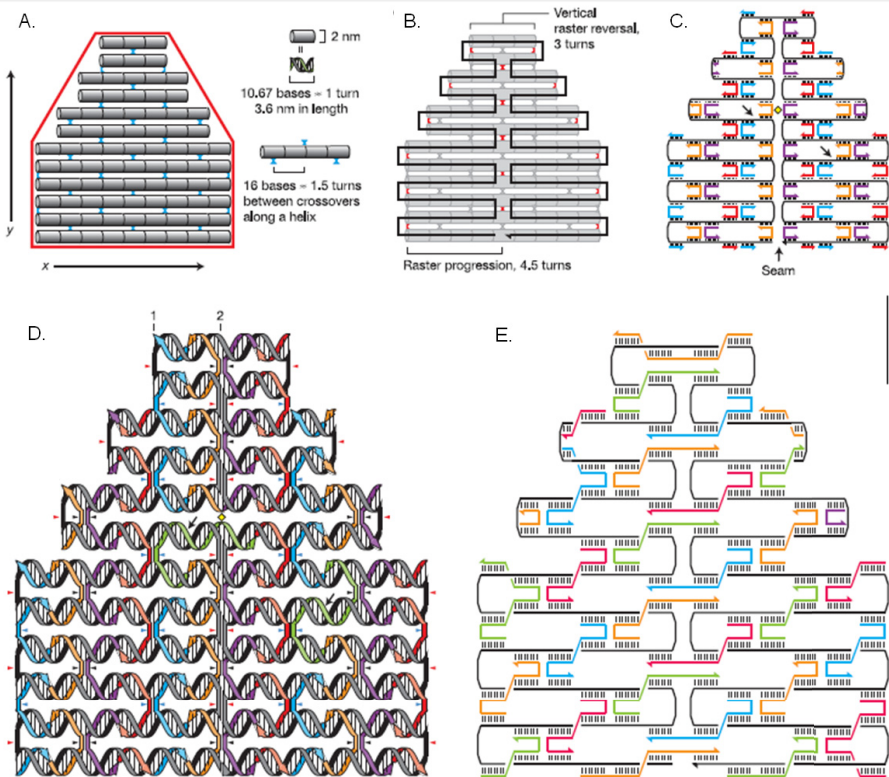
Figure 4.76: DNA origami design, taken from Rothemund.[257] A) A shape (red) approximated by parallel double helices (idealized by cylinders) held together by periodic crossovers (blue). B) A scaffold (black) runs through every helix and forms more crossovers (red). C) As first designed, most staples (coloured DNA strands) bind two helices and are 16-mers. D) Similar to C with strands drawn as helices. Red triangles point to scaffold crossovers, black triangles to periodic crossovers with minor grooves on the top face of the shape, blue triangles to periodic crossovers with minor grooves on bottom. Crosssections of crossovers (1, 2, viewed from left) indicate backbone positions with coloured lines, and major/minor grooves by large/small angles between them. Arrows in C point to nicks sealed to create green strands in D. Yellow diamonds in C and D indicate a position at which staples may be cut and resealed to bridge the seam. E) A finished design after merges and rearrangements along the seam. Most staples are 32-mers spanning three helices.

4.6.2 Single-molecule analysis using DNA Origami

With the addressability through the unique sequences, DNA origami is thus a great platform for the precise placement of a variety of functional molecules. In combination with atomic force microscope (AFM) with nanoscale resolution, this allows for single-molecule analysis. Contrary to commonly used analysis methods, which measure the bulk average, single molecule analysis can focus on the properties, behavior and trajectory in real time (with high speed AFM that allows recording of 15 images per second) of a single molecule in complex biological processes. Moreover, contrary to chip technology, these assays can be performed in solution. This has been used to elucidate the recognition and folding of biomolecules, like the puzzling human telomeric G-quadruplex and to study and visualize DNA robots and machines, label-free biomolecular recognition, chemical and enzymatic reactions, as reviewed in [260]. This is just a starting point, as this field is still in its infancy, but it is showing rapid progress.

4.6.3 The project
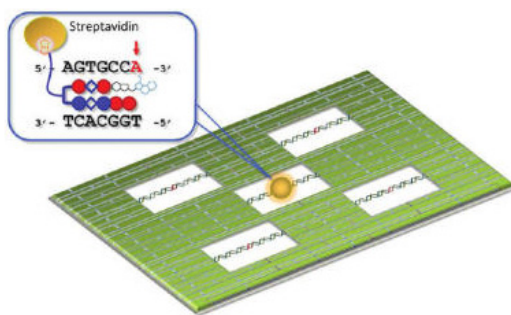
Figure 4.77: Cartoon representation of the prior art in the Sugiyama group.[261] The 'five-well DNA frame' consists of an origami tile with five cavities, carrying five different double stranded DNA sequences in these cavities. These DNA sequences can be used for detailed analysis of sequence selectivity, like alkylation by a Dervan peptide modified with an alkylating moiety and biotin. The biotin is used for streptavidin labeling, which can be observed by AFM.

Outline Based on a report of the visualization of single molecule alkylation that used a frame-like DNA origami scaffold incorporating various dsDNA substrates (bridges) into the cavities (illustrated in Fig. 4.77),[261] an attempt was undertaken to visualize the formation of ICLs through our furan methodology as illustrated in Figure 4.78A. For this purpose the bridges were designed using nicked dsDNA with sticky ends. At low temperatures the sticky overhang of 9 base pairs in length, is stable enough to form a closed bridge over the cavity. In this duplex furan interstrand cross-linking can be achieved, stabilizing the bridge, which can thus be maintained also at higher temperatures. A non-modified control nicked dsDNA bridge, will not form an ICL and will thus open at higher temperatures as illustrated in Figure 4.78B. A shorter overhang of 7 base pairs is less stable and can thus also open at lower temperatures, this should thus result in a lower cross-linking yield. In the central cavity a dsDNA bridge without nicks was annealed, as control to ensure a closed bridge. Exact sequences of the bridges can be found in the experimental part.

For the modification, the acyclic, phenyl stabilized furan-modified nucleoside (**nu6**) (Fig. 4.78C), introduced by Kristof Stevens [136] and shown to be most suited for singlet oxygen oxidation by Marieke Op de Beeck [139] was chosen. This will allow for cross-link tests triggered by both NBS and visible light. The latter has a broad potential with sensitizers for singlet oxygen production immobilized on the origami.[262]

Synthesis The acyclic, phenyl stabilized furan-modified nucleoside was converted over weekend to its phosphoramidite with the mild 2-cyanoethyl N,N,N',N'-tetraisopropylphosphoramidite in the presence of diisopropylammonium tetrazolide as activating species, rather than the standard β-cyanoethyl chlorophosphoramidite, which did not result in the formation of the phosphoramidite.[136]

The obtained phosphoramidite was manually incorporated in the designed 37-mer oligonucleotide I-2 (ODN4.6.1: 5'-ACG **Gnu6**G T-GC TTC CGG TAC TAC GCC CAG ATG AGC TAC T-3'). Probably due to the length of the ODN (37-mer), additional RP-HPLC purification was required to obtain the desired purity (Fig. 4.79).
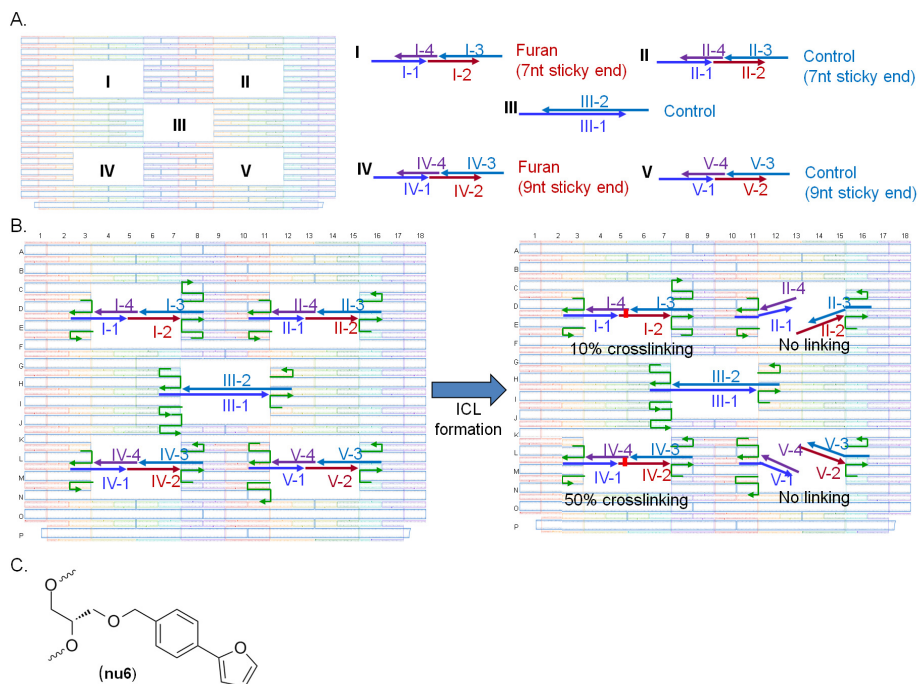
Figure 4.78: Design of the origami A) This design uses the '5-well frame', with 5 different dsDNA sequences annealed in its cavities. Cavity I and III carry the furan-modified oligonucleotide (I-2 and IV-2 respectively, which are identical), with a 7 base pair overlap in I and a more stable 9 base pair overlap in III. Well II and V are the identical control cavities, with an oligonucleotide (II-2 and V-2 respectively, which are identical) with a T instead of the furan-modified residue. In cavity III there is no furan modification and no nick in the DNA. The sequence of the overhang of oligonucleotides 1 and 3 in each cavity is unique, in order to anneal at the desired position in the origami frame. The sticky end is formed between 2 and 4 in each cavity. Oligonucleotide 4 is the same in the four cavities. The length of the different oligomers in number of bases is 88 for oligonucleotides III-1 and III-2, 51 for I-1, II-1, IV-1 and V-1, 46 for I-3 and II-3, 44 for IV-3 and V-3, 42 for I-4 (= II-4, IV-4 and V-4), 37 for I-2 (= IV-2) and II-2 (= V-2). B) When the origami tile is kept on ice, all dsDNA bridges should be annealed or closed, like the control in cavity III. Under these conditions the cross-linking reaction is performed. During analysis the tiles are not cooled anymore, which causes the non cross-linked nicked bridges (cavity II and V) to open, while the cross-linked bridges (cavity I and IV) stay closed. Due to the shorter sticky end in cavity I and II compared to cavity III and IV, the DNA bridge will partly open even on ice, which will result in a lower cross-link yield. C) The acyclic, phenyl stabilized furan-modified nucleoside (**nu6**).
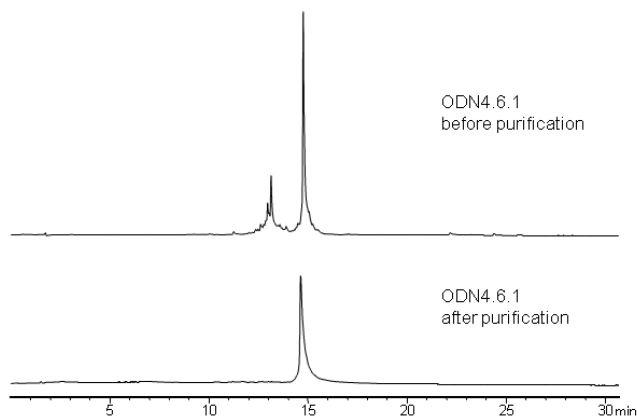
Figure 4.79: I-2, the designed furan-modified 37-mer ODN4.6.1 RP-HPLC chromatogram from Clarity column with a gradient from 0.1 M TEAA-buffer to 30% MeCN in 30 min at 60°C, before (top) and after purification (bottom).

The modified oligonucleotide was annealed together with the other bridge forming oligonucleotides in the pre-assembled five-well frame (Fig. 4.80). This two-step method (Fig. 4.81) is preferred over the one-step assembly of both frame and bridges from an enormous mixture, to reduce the risk of mismatches. The five-well frame is thus first assembled by the combination of the M13mp18 long single stranded DNA and all the designed staple strands in the annealing buffer and purified by gelfiltration. The bridges were all assembled separately following the same protocol. The frame was finally combined with 4 equivalents of the bridges, which can insert in the frame by the 5'-overhangs complementary to both edges of the cavity. This mixture was mildly heated to allow insertion of the bridges in the frame and then cooled very slowly, overnight. This very slow process was shown to be necessary, as faster cooling give inferior results without inserted bridges, as shown in Figure 4.82. The origami was again purified by gelfiltration.
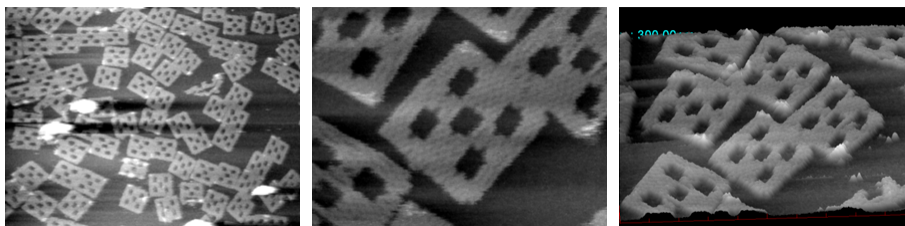
Figure 4.80: Five-well frame. AFM images from the assembled five-well frame
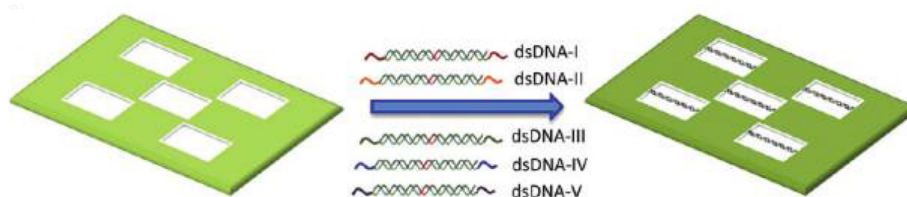
Figure 4.81: Principle of the construction of the bridged five-well origami.
In a first step the five-well frame and all the bridges are all seperately annealed, then
the duplex bridges are incorporated in the five-well frame by self-assemmbly.[261]
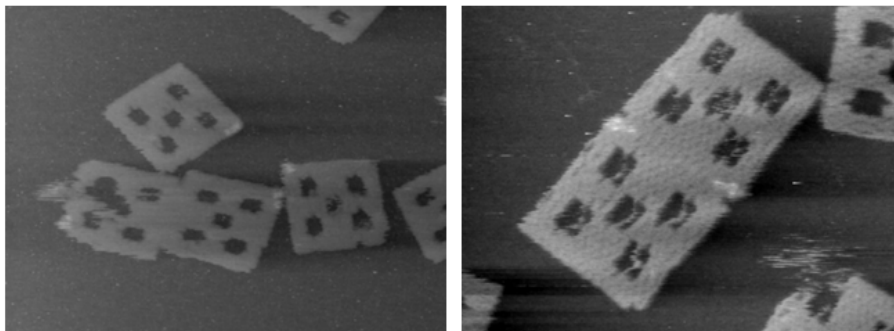
Figure 4.82: Five-well frame with inserted bridges. AFM images from the as-
sembled five-well frame with bridged, after annealing in 35 min, leaving most of the
cavities empty (left) and overnight, filling most of the cavities with a bridge (right)

Cross-linking A fresh origami stock was prepared before every experiment. Both NBS and singlet oxygen were used to oxidize the furan. The latter was generated by irradiation of the sample with methylene blue with monochromatic red light for 30 or 60 minutes. The oxidation was performed on ice, to ensure closed bridges or thus duplex sticky ends.

Looking at the AFM images, shown in Figure 4.83, it is very difficult, if not impossible to see the difference between the open and closed bridges as even the non-nicked control in the middle often looks open. In an attempt to make visualization easier, because perhaps the bridges do not open under room temperature, the Mg^{2+} concentration of the washing solution and/or the AFM solution was lowered. The lower salt concentration should cause destabilization of the duplexes and thus opening of the bridges. However under such conditions the origami did not attach to the mica surface anymore. Magnesium ions are required to neutralize the negative surface charge of the mica, and avoid repulsion of the similarly charged origami.

As an alternative method to check for cross-linking, denaturating PAGE was performed. The samples (blanc, NBS, $^1O_2$ 30 min and 60 min) were denatured by heating with formamide and loaded on a gel. The obtained result is shown in Figure 4.84. The M13mp18 long template strand could not penetrate in the gel and remained on the bottom of the wells. Due to the enormous amount of staple strands (160, varying in length from 32 to 71 bases) and bridge strands (20, varying in length from 37 to 51 bases), it was however impossible to clearly follow the reaction (between a single 37 mer and 42-mer, forming a single 79-mer).
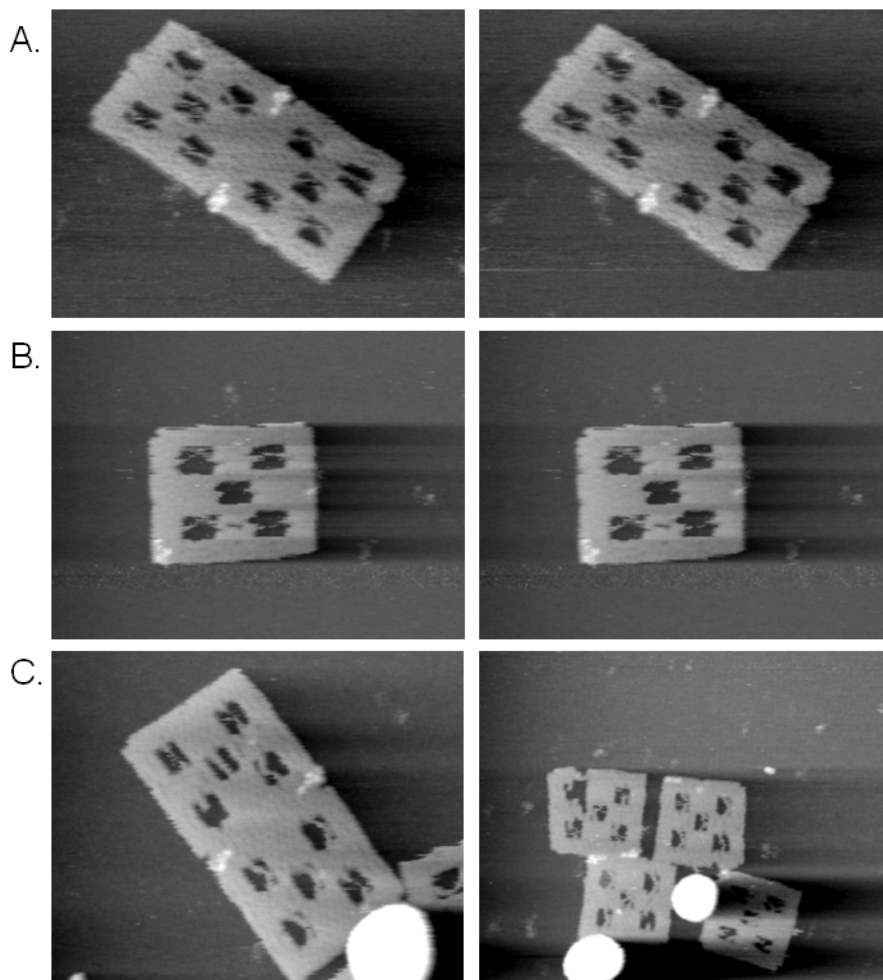
Figure 4.83: AFM images of the cross-link reactions on DNA origami A) The origami without cross-linking as reference; B) The origami after the cross-linking with NBS; C) The origami after the cross-linking with $^1O_2$: 30 min (left) and 60 min (right)
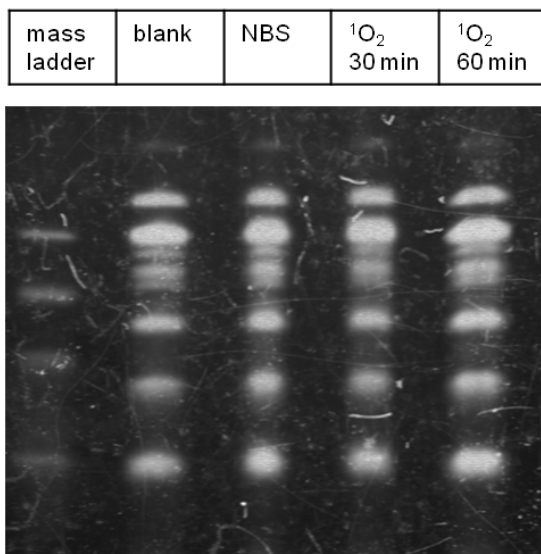
Figure 4.84: PAGE of the cross-link reactions on DNA origami The obtained bands on the polyacrylamide gel from the electrophoresis experiment, show the overwhelming presence of bridges and staples. The mass ladder consists of a mixture of oligonucleotides with 64, 56, 44, and 32 basepairs

Conclusion   The furan-modified DNA was prepared and could be incorporated in the origami. It was however impossible to visualize the cross-linking reaction using this specific design. An improved design could solve this problem and should thus be investigated. Moreover, optimization of the cross-link conditions is required, due to the very small amount and especially the very low concentration of the furan-modified DNA in the origami (0.01 μM) compared to the used concentrations in our standard protocol (20 μM). Transfer of the amount of equivalents of sensitizer or NBS, results in very low concentrations that could disturb the kinetics of the reaction, while transfer of the concentration of sensitizer or NBS, will result in DNA degradation due to the very high excess. For applications on origami, like the covalent immobilization of proteins on origami by the Morii group or for example in vivo, this needs to be further investigated.

Chapter 5

Solid Phase Methodology Development

This final chapter combines some results that are not directly related to the scope of this thesis, the binding and cross-linking of peptides and oligonucleotides, but focus on the used methodology to synthesize them. Oligomers are typically synthesized in repetitive steps from protected monomers, as represented in Figures 5.1 and 5.2.

After deprotection of a coupled monomer, the next monomer can be added for a new synthesis cycle. This process can be dramatically simplified, as first noted by Merrifield[263], by attachment of the growing oligomers to a solid support, which allows for easy purification by filtration and washing and is the basis for automation. Both peptides and oligonucleotides are now routinely synthesized on solid support, polystyrene or polyethyleneglycol and controlled pore glass respectively. The last step of the synthesis thus consists of cleaving the oligomer from the support and deprotecting it, removing semi-permanent protecting goups.

A = carboxy activating group
Y = semi-permanent side chain protecting group
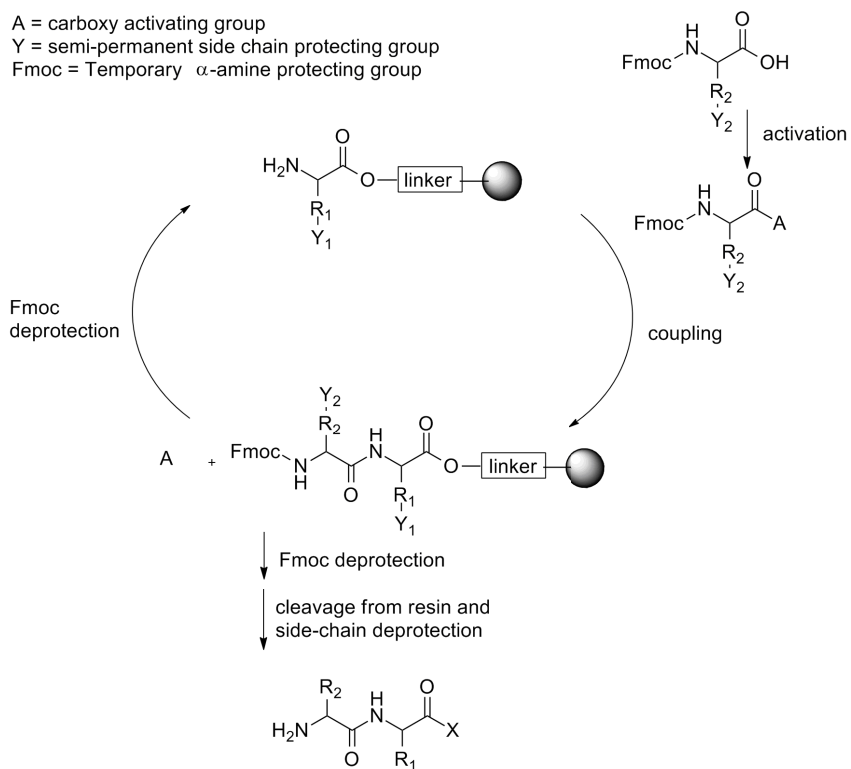Fmoc = Temporary α-amine protecting group

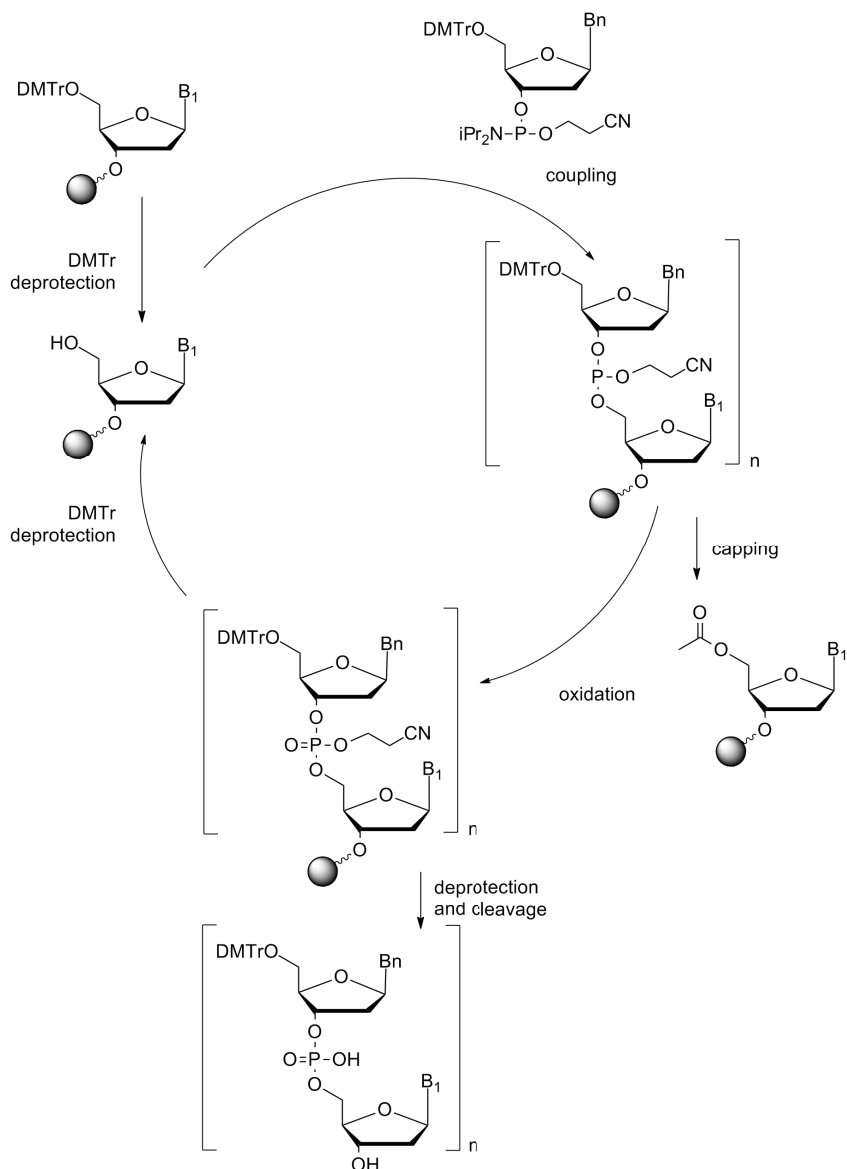**Figure 5.1: General representation of Solid Phase Peptide Synthesis**

Figure 5.2: General representation of Solid Phase Oligonucleotide Synthesis

In solid phase synthesis it is very important that the reactions are quantitative, as remaining starting material is like the product, attached to the support and thus transferred to the following reaction. By adding reagents in large excess, the reaction is driven to completion, while the excess is easily removed by filtration. To monitor the progress of the reaction, it is not possible to use TLC or MS, without intermediate cleavage from the resin. Colour tests were developed to partially circumvent this problem. Indeed, if the coupling step was complete, no reactive functionalities remain for coupling with a dye. Completion of the reaction can thus be visualized. Unless, if undeclared functionalities of the support are able to interfere like shown in the first section of this chapter.

The second section presents an NMR based method and its advantages for quantification of very small quantities of synthesized macromolecules, which is available through collaboration with the NMR Structure Analysis unit of Prof. Martins.

Though solid phase synthesis is a highly streamlined automated process, the protection-deprotection scheme makes the process inefficient. Furthermore a different monomer has to be used (and sometimes synthesized) for every different, desired functionality. In collaboration with the Polymer Chemistry Research group of Prof. Duprez, we developed a new coupling methodology for the preparation of functionalized, sequence-defined oligomers with a novel backbone from a single readily available thiolactone-containing building block, without protecting groups on solid phase (section 5.3).
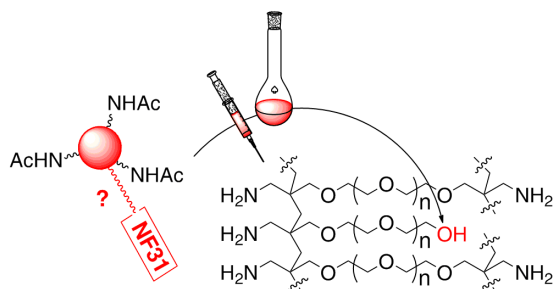
5.1 Color test

NF-31 color test uncovers 'hidden' alcohol functionalities in PEG-based resins for solid phase peptide synthesis

Lieselot L. G. Carrette, Dieter Verzele, Annemieke Madder *

Laboratory for Organic and Biomimetic Chemistry, Department of Organic Chemistry, Faculty of Sciences, Ghent University, Krijgslaan 281 S4, B-9000 Gent, Belgium

ABSTRACT

Recently developed PEG-based resins have been shown to markedly improve the quality of the synthesis in case of the so-called 'difficult' peptide sequences. Difficult coupling reactions further require sensitive color tests for amines in the assessment of the completeness of coupling. We here describe how the use of PEG-based resins in combination with one of the more sensitive color tests in SPPS can lead to severe misinterpretation and unnecessary delays during solid phase peptide synthesis due to the presence of 'non-declared' free hydroxyl functionalities in aminomethyl-PEG-based resins.

Introduction   Solid phase synthesis of highly hydrophobic peptide sequences is often problematic due to intramolecular aggregation of the peptide and hydrophobic interactions involving the solid support. Aggregation phenomena can have a severe negative influence on the reactivity of the free N-terminus, leading in subsequent reactions to deletion sequences and eventually premature termination.[108] The exact chemical composition of the resin material plays a major role and interaction with the solid support depends strongly on its hydrophobicity. With this in mind, the development of new and more performant solid supports has been and still is an active field of research. People have actively sought for solu-

tions by the development of new advanced solid supports which consist of 100% polyethyleneglycol (PEG), like the example depicted in Figure 5.3. This type of resin is less hydrophobic than the traditional polystyrene (PS) and PS based resins. Therefore the PEG-resins allow a better solvation of the growing peptide chains, making them perfectly suited for high purity synthesis of difficult hydrophobic, highly structured and also poly-arginine sequences.[101, 108, 264] These resins have good swelling properties in about any solvent. This springs from the vicinal arrangement of carbon oxygen bonds throughout the chain in the unique structure of PEG. It assumes helical structures with gauche interactions between the polarized bonds. Three helical arrangements of PEG featuring a low energy conformation exist. The first exposes the oxygen atoms to the environment and is thus quite hydrophilic. In the second low energy conformation, the oxygen atoms are oriented towards the interior of the helix, which is therefore largely hydrophobic. The third possesses an intermediate polarity. For this reason PEG displays an amphipathic nature and swells in both polar and non polar solvents.[265] PEG-resins are also thermally, mechanically and chemically very stable as they are exclusively built up from primary ether bonds. Moreover, they are easy to handle in the dry state, thus facilitating weighing and transfer.[264]
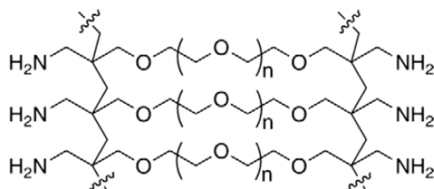
Figure 5.3: General representation of NovaPEG or ChemMatrix®resin structure: 100% cross-linked PEG.[264, 266, 267]

The synthesis of peptides on solid phase could only be conveniently and generally applied with the development of colorimetric tests. These allow monitoring the coupling efficiency of amino acid coupling, without the need for cleavage from the solid phase.[268, 269] Next to the widely used Kaiser[270] and TNBS[271] tests our group made a contribution in this area with the development of p-nitrophenylester NF31 (Fig. 5.4), which allows reliable monitoring of coupling reactions even with very sterically hindered primary and secondary amines.[272, 273]
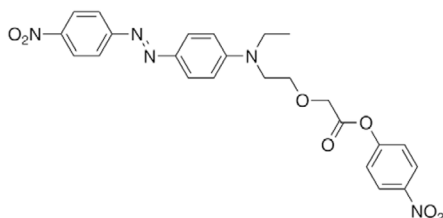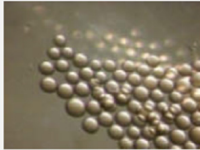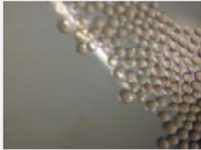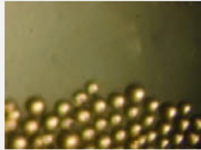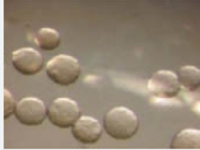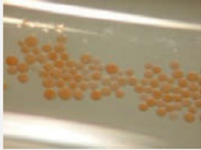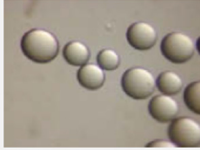
Figure 5.4: Structure of p-nitrophenylester NF31.

Combining the accuracy of sensitive color tests with the improved properties of new types of resins offers a way to improve the quality of difficult and not well-behaved peptide sequences. Clearly, next to the importance of the specific and optimized structure of the base polymer for improved properties of SPPS beads, the presence of a specific functionality on the resin is important for attachment of the desired sequence. Amine groups are generally used for this purpose, leading to stable amide bonds, but allowing for cleavage at the end of the synthesis by the introduction of suitable linker moieties. However it is important to be aware of residual functionality present on the beads, not deliberately introduced for attachment purposes but rather present as a result of the specific polymerization strategy used for construction of the resin. We here report on the presence of 'non-declared', residual, free hydroxyl groups in solid supports consisting of 100% PEG leading to background coloration with the sensitive NF-31 test.

Results and discussion   In the course of our recent endeavors in the synthesis of large dipodal peptide constructs featuring long chains of more than 20 amino acids, we were faced with severe problems in achieving the desired purity. As described above, recent papers report on the exclusive capacities of PEG-based resins in 'difficult' peptide sequences.[101, 108, 264] This prompted us to turn our attention to this novel type of resins. Before embarking on the synthesis of the actually desired large peptide constructs, a series of test reactions was run to verify and establish suitable basic peptide synthesis conditions on this new type of resin. During the simple synthesis of a small peptide, when using NovaPEG, a 100% PEG-based resin from Novabiochem, TNBS tests were negative. But to our great surprise, our regularly applied NF-color test never indicated complete coupling. Despite prolonged reaction times or multiple couplings, red coloration

Table 5.1: Color tests on resin after coupling reaction with N-Fmoc-glycine[a]

Resin	TNBS-test	NF-31-test	NF-31 + DMAP
Tenta-Gel			
Nova-PEG			
Wang[b]			
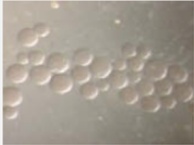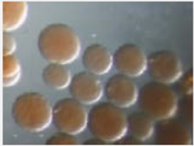
[a] N-Fmoc-glycine (4 equiv) was coupled with PyBOP (4 equiv) and DIPEA (8 equiv) for two times 2 h in DMF.
[b] No coupling reaction was performed on the Wang resin, the color tests were performed on the unmodified resin.

of beads was continuously observed, pointing towards the presence of unreacted functionalities. Using identical conditions on TentaGel, a PS-based resin, color tests indicated complete couplings (see Table 5.1).

In view of the positive reports on the performance of these PEG based resins, it was considered highly unlikely that it was impossible to obtain 100% efficiency on this novel resin during simple coupling reactions. Based on the structure of the resin, the presence of free hydroxyl groups was considered to offer a possible explanation for the unexpected positive color tests. Containing a reactive p-nitrophenyl ester functionality, NF-31 has shown to also allow for the detection of alcohols provided slightly different test conditions are applied.[274] Using these conditions, more intense red beads were obtained (see Table 5.1), as would be expected if hydroxyl groups were present. To further test this hypothesis, TNBS and NF-31 tests were carried out on the different types of resin, after applying a capping procedure with acetic acid anhydride, as illustrated in Table 5.2. To

Table 5.2: Color tests after capping with acetic acid anhydride

Entry		TNBS-test	NF-31-test
1	30 min capping		
2	4 h capping		
3	After NaOH wash		

For the capping procedure: 6 equiv acetic acid anhydride and 6 equiv DIPEA were added to the solid support in DMF. For the washing procedure: the solid support was washed four times with 0.1 M NaOH.

obtain a complete reaction of all present functionalities and associated disappearance of the red color, a longer reaction time was required than used in the standard protocol for the capping of amines, hinting toward the presence of another less nucleophilic functionality. In turn, after treatment of the resin with an aqueous 0.1 M NaOH solution, the originally observed red color reappeared. In view of the inherent difference in base sensitivity of an amide versus an ester bond, this experiment confirms the presence of free hydroxyl groups on the PEG-based resin.

The specific and beneficial characteristics of the NF 31 test, especially its increased performance and sensitivity for secondary, sterically hindered and aromatic amine functionalities (as compared to more classical TNBS or Kaiser test) have initiated its use in a variety of cases where 'difficult' peptide sequences are desired.[275–278] Furthermore, in view of the growing importance of the incorporation of unnatural (and often difficult to couple) building blocks in peptide chains and the often improved results obtained when using the novel PEG-based resins, one should be aware that the combination of this novel well-performing resins with OH-sensitive color tests can give rise to false positive results, im-

peding a correct interpretation of coupling progress and leading to the entirely unnecessary application of double couplings and concurrent waste of precious building blocks. The presence of free hydroxyl groups can further also interfere in other reactions. When determining coupling efficiencies by UV absorbance of the piperidine benzofulvene adduct [279] for example, amino acids coupled on the free hydroxyls and their subsequent Fmoc deprotection with piperidine, will give rise to wrong results.

Since schematic and graphical representations of aminomethylated NovaPEG and ChemMatrix resins in catalogs [266]and papers [108, 267] describing the benefits of their use in difficult peptide synthesis, often omit to depict residual free hydroxyl functionalities (Fig. 5.3), it is important to realize that they are present and might interfere with some of the processes carried out on the resin.

Conclusion   In conclusion, we have been able to understand and show why the application of the highly sensitive NF-31 test during SPPS on novel PEG-based resins (sold under popular brand names as Nova-PEG or ChemMatrix®) gives rise to difficult interpretation of coupling progress. We have demonstrated that the presence of remaining free hydroxyl functionality leads to misleading background coloration of the resin when checking reaction progress.

Although the presence of free hydroxyl groups does not necessarily lead to associated problems during SPPS, one should be careful when following reactions using sensitive color tests as to avoid unnecessary application of multiple coupling procedures or use of increased excesses of precious reagents.

5.2 NMR based quantification as a valuable solution for quantification of non-UV active compounds and mixture analysis

with Katelijne Gheysen and Prof. Jose Martens

Introduction   Upon synthesis of biomolecules like DNA and peptides, quantification of the acquired amount of material is indispensible to further test and determine activity, binding affinity and biological properties. Typically the obtained amounts are too low to be measured accurately on the standard laboratory analytical balance; moreover possible salt formation during synthesis makes these measurements unreliable. For the concentration determination of proteins, different indirect biochemical methods have been developed, mostly based on measurement of an optical signal resulting from binding to or hydrolysis of a reactive dye.[280, 281] Unfortunately this often requires specialized instruments and intensive manual handling as careful sample preparation and calibration have to be performed. These methods are also mostly destructive for the sample and are not always applicable to smaller synthesized peptides. Therefore UV spectrometry is routinely applied for the determination of concentrations.[282] It is cheap and easy provided the extinction coefficient of the compound is known at representative conditions of solvent, pH, ionic strength and concentration.[283] However, this is not possible for molecules without chromophoric groups, such as peptides without aromatic amino acids[284] or that are extensively modified with additional aromatic groups.

It is common knowledge that NMR is perfectly suited for the determination of concentrations, in view of the direct proportionality of signal intensity to the number of resonating nuclei.[285] As this method does not require chromophores or other specific structural entities on the target molecule it is thus broadly applicable. Moreover, the technique is not destructive and additional information about sample purity and identity is generated. However, its implementation has been hampered for a long time by the need of good internal standards, which

have to be carefully selected taking in account solubility, chemical interaction, spectral overlap and relaxation time. Recently, this tedious search could be eliminated by the development of alternative ways for referencing, by the generation of an electronic NMR-like signal, ERETIC (Electronic Reference To access In vivo Concentrations)[286] and by use of an external standard, PULCON (PUlse Length -based CONcentration measurements).[287] For the latter technique, which is more practical in routine applications than the former, automatic post-acquisition QUANTAS (QUANTification by Artificial Signal) software was developed and implemented for Bruker TOPSPIN software, facilitating its practical implementation.[288]

In spite of the described advances, rendering the technique perfectly applicable to standard NMR apparatus, this method still has not found introduction into the synthesis laboratory. We therefore here report on a comparative study where the use of PULCON with the QUANTAS software was validated against UV as a straightforward means for concentration determination for a series of representative samples, ultimately illustrating its usefulness for non-UV active compounds and mixtures.

Discussion In order to validate the possibilities of NMR for concentration determination in our laboratories a series of test experiments was set up with stock solutions of three standards, DAPI and tryptophan in water and estradiol in methanol. The envisaged concentrations were respectively 7.50, 15.00 and 3.00 mM. Measurements were performed using UV and NMR on the same day and with minimal differential manipulations. For each compound a quantitative 1D-^{1}H NMR spectrum with full relaxation was recorded. Subsequently all signals with adequate resolution were integrated and scaled upon the artificial QUANTAS signal. As illustrated in Table 5.3, the obtained concentrations with both methods were in good correlation for the latter two standards. With QUANTAS concentrations of 15.16 $\pm$ 0.09 mM for tryptophan and 2.67 $\pm$ 0.04 mM for estradiol were obtained, whereas the UV measurements resulted in concentrations of respectively 15.10 $\pm$ 0.05 mM and 2.69 $\pm$ 0.04 mM. The concentration for the first standard, DAPI, on the other hand was lower when measured with NMR than with UV (respectively 6.45 $\pm$ 0.21 mM and 8.57 $\pm$ 0,03 mM). This inconsistency illustrates a risk associated with UV based

concentration measurements. The use of reported extinction coefficients requires data from the exact molecule, in the same solvent and at the same concentration. If not, higher or lower concentrations can result in different association behaviour, resulting in different absorption values. Not taking into account the concentration dependency of the extinction coefficient in this case, clearly results in deriving a higher UV concentration value for the DAPI standard. As such, this initial validation already illustrates the benefits of using QUANTAS instead of UV.

Table 5.3: Summary of the concentrations, obtained with NMR and UV
[a]concentrations represented in mM, concentration of the oligonucleotide taking into account length,[b] length and composition[c] or the length, composition and sequence[d]

Sample	Concentration[a]		
	envisaged	determined by NMR	determined by UV
Standards			
- DAPI	7.50	6.45 ± 0.21	8.57 ± 0.03
- tryptophan	15.00	15.16 ± 0.09	15.10 ± 0.05
- estradiol	3.00	2.67 ± 0.04	2.69 ± 0.04
oligonucleotide	2.00	1.83 ± 0.03	1.58 ± 0.02^{b}
			1.69 ± 0.03^{c}
			1.66 ± 0.02^{d}
peptides			
- AcNH-ILPEI-COOH	5.00	4.42 ± 0.25	X
H_2N-RSSVGSQS-COOH	5.00	2.98 ± 0.17	X
mixture			
- oligonucleotide	0.90	0.80 ± 0.01	0.83 ± 0.01
- tryptophan	1.50	1.52 ± 0.05	13.12 ± 0.20

The methodology was then tested on representative samples from our groups that are typically more difficult to quantify. Initially the concentration of a commercial oligonucleotide, 5'-GCA CAC CGT CAG-3', was determined. For DNA, UV concentration determination represents the standard procedure. It is routinely performed in many laboratories, since oligonucleotides are used as probes, templates and binding sites both in biological and chemical setups. It has been noted that dissolving commercial lyophilised DNA and determining its concentration by UV hardly ever matches the predetermined concentration, thus illustrating the need for accurate concentration measurement before any further experiments. Three different methods have been described to calculate extinction

coefficients of DNA based on the nucleotides it consists of; taking into account only the length,[289] the length and composition[290] or the length, composition and sequence.[291] The latter so-called nearest neighbour method, is considered the most accurate and is thus the most popular. It was however earlier shown that experimentally determined extinction coefficients are better suited for exact concentration measurement as the calculated extinction coefficients introduce an additional uncertainty of 20%.[292] This uncertainty is even more pronounced in the case of DNA with synthetic modifications that are not considered in the extinction coefficient calculation. This approximation is often necessary if the contribution of the modification and its influence on the behaviour of the oligonucleotide as a whole is not known. Using the extinction coefficients calculated with the oligonucleotides length, length and composition and length, composition and sequence, UV measurements found a concentration of respectively 1.58 ± 0.02 mM, 1.69 ± 0.03 mM and 1.66 ± 0.02 mM. NMR determined a concentration of 1.83 ± 0.03 mM. Considering the earlier mentioned additional 20% uncertainty, UV measurements and NMR measurements fall within the same range. Considering that quantitative NMR integration is independent of molecular features, except for the number of protons counting for a signal, it can be considered more accurate.

In a next series of experiments, we turned our attention to the concentration determination of peptides. Extinction coefficients can be derived from the constituting amino acids, provided aromatic moieties are present. As this is not the case for many of the peptides synthesized in our laboratory, we hitherto often faced problems when trying to quantify the obtained amounts of material. In an attempt to test the described methodology also in this peptide context, we submitted two representative non-UV active peptides to concentration determination with QUANTAS. In first instance, the pentapeptide AcNH-ILPEI-COOH, previously synthesized for a study concerning caspase activity, was selected.[293] As at the time of the experiments its concentration could not be determined using UV, estimation of concentration was based on the weight of the sample and thus determined to be 5.00 mM. The concentration obtained with QUANTAS was found to be 4.42 ± 0.25 mM which is in rather good agreement with the previous estimated value. Secondly, an octapeptide (H_2N-RSSVGSQS-COOH)

synthesized during the development of an artificial estrogen receptor mimic was tested.[294] In this case, the determined concentration was 2.98 ± 0.17 mM, which is much lower than the earlier estaminated value of 5.00 mM based on the weight. However, examination of the purity of both peptide samples (ca. 96% and 74% as determined from HPLC analysis) showed that whereas the ILPEI peptide was very pure (Fig. 5.5a), the octapeptide, though appearing as one main signal in the HPLC chromatogram (Fig. 5.5b), does contain traces of various impurities. These impurities contribute to the total weight of the sample, therefore leading to overrated reaction yields. This further illustrates another significant advantage of the QUANTAS protocol namely that impurities can be readily observed when recording the quantitative 1D spectrum and avoided in the quantification.

As is clearly illustrated here for the octapeptide case, a great merit of QUANTAS compared to UV measurements is the possibility to determine the concentration of active or desired compounds in contaminated samples. It can be further deduced that QUANTAS can thus be used in the straightforward determination of the concentration of different compounds constituting a mixture. To illustrate this added value of the QUANTAS technique, a mixture was prepared of the tryptophan standard and the commercially available oligonucleotide (with calculated concentrations of respectively 1.50 and 0.90 mM). Former results have shown that NMR as well as UV measurements resulted in the correct concentrations when the compounds were separately dissolved. Using UV, the oligonucleotide concentration within the mixture was determined to be 0.83 ± 0.01 mM with the nearest neighbour method. The tryptophan concentration on the other hand could not be determined, as the oligonucleotide dominates the absorption found (13.12 ± 0.20 mM). Indeed, inspection of the absorption in the range 500 to 200 nm reveals only one absorption signal corresponding to the absorption of the oligonucleotide with a maximum at 260 nm (Fig. 5.6). No absorption signal for tryptophan can be distinguished separately and its concentration can thus not be determined, though it slightly contributes undetected to the absorption of the oligonucleotide. Using QUANTAS NMR, no problems were encountered and the concentrations of both substances could easily be determined (0.80 ± 0.01 mM for the oligonucleotide and 1.52 ± 0.05 mM for
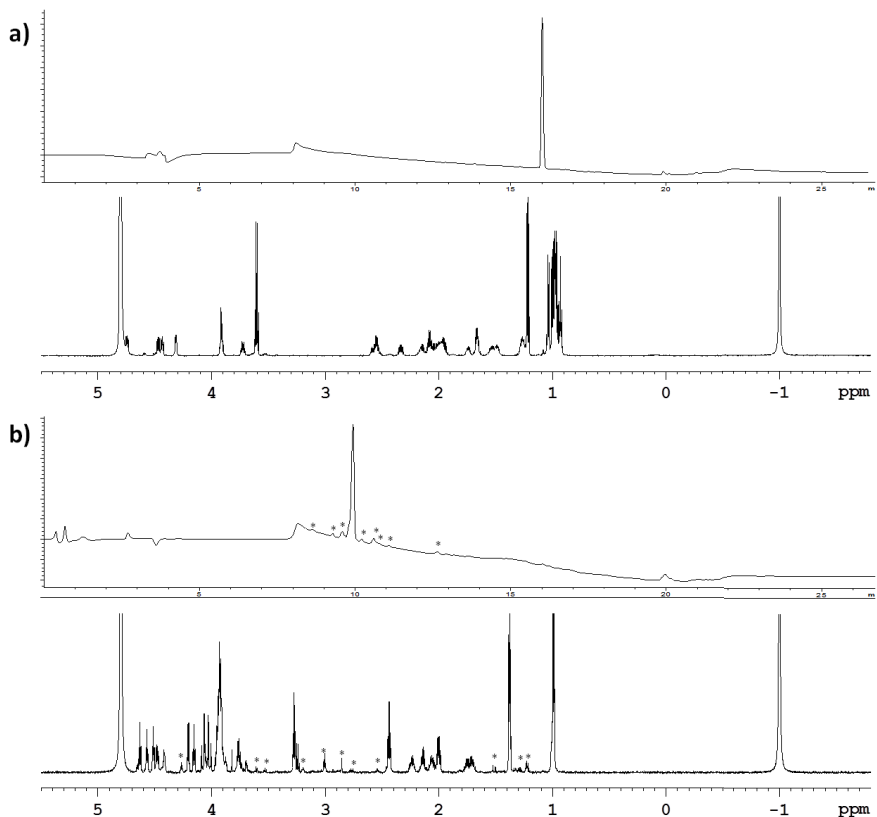
Figure 5.5: Chromatogram (top) and NMR spectrum (bottom) of a) AcNH-ILPEI-COOH and b) the octapeptide. The chromatograms were recorded at 214 nm on a Agilent 1200 Series HPLC, Luna column C18, 100 Å(250 x 4.6 mm, 5μm) from 100% 0.1 TFA in water to 100% MeCN in 15 minutes to examine the purity of the peptides. In the 1D-^{1}H NMR spectrum from the octapeptide, the signals belonging to impurities can easily be identified according to their lower concentration (ranging from 0.24 mM to 1.19 mM but the number of protons counting for the signals is unknown). The NMR signal at -1 ppm is the artificial QUANTAS signal. The impurities are denoted with *.

tryptophan). This illustrates that QUANTAS is not just a reliable fast and convenient method for concentration determination, but also further allows the accurate determination of the concentration of a compound in a contaminated sample or of compounds in mixtures.
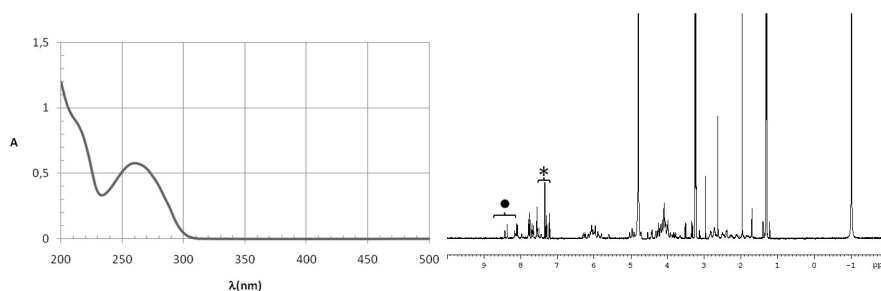
Figure 5.6: UV-spectrum (left) and 1D-^{1}H NMR spectrum (right) of the Trp/oligonucleotide mixture. The UV spectrum shows the typical absorption of an oligonucleotide with maximum at 260 nm. The tryptophan contamination is not distinguishable from the oligonucleotide signal. In the NMR spectrum on the other hand, clearly separated resonances for both compounds can be identified: ∗ are signals belong to Trp and • represent aromatic signals from the oligonucleotide. The NMR signal at -1 ppm is the artificial QUANTAS signal and the sample was contaminated with ethanol.

Conclusion In conclusion, we have tested and implemented QUANTAS in our laboratory as a viable alternative for UV concentration determination of synthetic biomolecules on standard NMR equipment. This allows us to obtain accurate data for standard and synthetically modified oligonucleotides, UV and non-UV-absorbent peptides, contaminated samples and mixtures. Furthermore, it is a practical method that requires less manipulation and is less prone to mistakes and misinterpretations than UV based quantification.
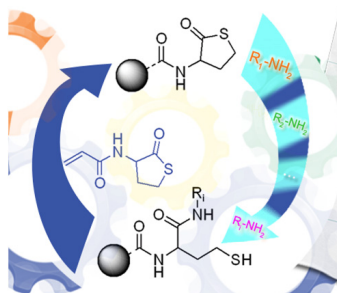
5.3   Coupling Method

Sequence-Controlled Oligomers

Multifunctionalized Sequence-Defined Oligomers from a Single Building Block**

Pieter Espeel, Lieselot L. G. Carrette, Katarzyna Bury, Sven Capenberghs, José C. Martins, Filip E. Du Prez,* and Annemieke Madder*

[*] Dr. P. Espeel,[+] K. Bury, S. Capenberghs, Prof. F. E. Du Prez
Polymer Chemistry Research Group
Department of Organic Chemistry, Ghent University
Krijgslaan 281, S4, 9000 Gent (Belgium)
E-mail: Filip.DuPrez@UGent.be

L. L. G. Carrette,[+] Prof. A. Madder
Organic and Biomimetic Chemistry Research Group
Department of Organic Chemistry, Ghent University (Belgium)
E-mail: Annemieke.Madder@UGent.be

Prof. J. C. Martins
NMR and Structure Analysis Research Group
Department of Organic Chemistry, Ghent University (Belgium)

[+] These authors contributed equally to this work.

Two decades of progress in the field of living and controlled polymerizations, combined with the elaboration of efficient conjugation reactions, have greatly contributed to the elegant preparation of functionalized macromolecular architectures.[295–298] However, these state-of-the-art methodologies, while providing a high degree of structural and topological control, are inadequate tools for controlling the polymer microstructure. Crucial parameters like primary structure (i.e. monomer sequence) and tacticity largely remain unmastered by current man-made approaches. Expectations for the next generation of synthetic polymers include performance as single chains, ability to fold and self-regulate,

and ability to sense specific molecules and/or catalyze reactions.[299–306] These precisely functionalized linear polymers should exhibit sharply defined and tailored structureactivity relationships analogous to Nature's delicately engineered macromolecules. Therefore, progress towards reliable sequence-controlled polymerization, enabling the preprogrammed distribution of multiple functional groups along the backbone, is drawing attention in a growing number of research groups worldwide.[307–311]

Pioneering efforts to control the primary structure of functionalized polymers have been based on several approaches, using different reactivity ratios of vinyl monomers,[312, 313] spatial prearrangement of monomers on a (macromolecular) template[314–318] or, as recently demonstrated, the action of a small-molecule machine.[319] Other attempts use (automated) sequential addition of building blocks on a solid[320, 321] or liquid[322–324] support, leading to sequence control as a result of iterative coupling steps, thereby omitting the need for preorganization. These methods, established for peptide[263, 325] and oligonucleotide[326] synthesis, present considerable drawbacks for sequence controlled polymerization: they generally require the use of protecting groups, and the restricted number of readily available building blocks ('monomer alphabet') equipped with the appropriate functional handle can further hamper the preparation of tailor-made functionalized sequences. The development of new chemical methods for chain elongation, often on a solid support, that result in sequence-defined (macro)molecular structures with unique backbones and side chain functionalities, or fragments thereof that could be combined to obtain sequence-controlled polymers,[311] is therefore highly desirable.

Herein, we report on a new coupling strategy for the controlled generation of sequence-defined multifunctionalized oligomers on solid support by using thiolactone-based chemistry. This protecting group free approach was inspired by the 'submonomer' synthetic method for the preparation of functionalized peptoids.[327–329] Although the generated oligomers are small in size, reconstitution approaches could further allow the synthesis of larger chains featuring the designed and repetitive display of carefully selected and well-positioned functional entities.
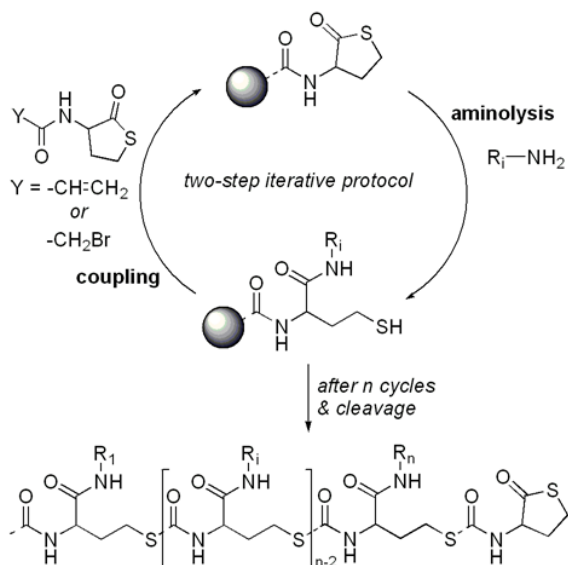
Figure 5.7: Two-step iterative method for the synthesis of functionalized oligomers on solid support: aminolysis of the resin-bound thiolactone followed by coupling of a thiolactone-containing building block that selectively reacts with the generated thiol.

The synthetic scope of thiolactones as reactive precursors for thiols in various polymeric systems was recently reported by the group of Prof. Du Prez[330–334] and other groups.[335, 336] These cyclic thioesters are susceptible to selective aminolysis, releasing a thiol and thus providing a functional handle for subsequent efficient conjugation reactions.[337, 338] From this, we reasoned that immobilization of a thiolactone unit on a solid support should enable chain extension after on-resin aminolysis. Indeed, using a judiciously selected thiolactone building block to reinstate the thiolactone functionality would allow the start of a next iterative reaction sequence (Fig. 5.7). This two-step aminolysis/chain extension method does not make use of any protecting groups and relies on a single thiolactone-containing building block for chain extension. Most importantly, a myriad of functionalities can be introduced by using the corresponding readily available amines.

The main requirement for this approach is the preparation of a suitable thiolactone-based building block for chain extension. Since such a reactant will typically be used in high excess when applying solid-support methods, its synthesis has to be high yielding on a multigram scale. Different building blocks (Fig. 5.8) were designed and could be synthesized in a one-step procedure, with a straightforward purification, from the readily available racemic DL-homocysteine thiolactone (performed by Dr. P. Espeel, see experimental part). Together with compound **5.3.1**, which is used for the initial loading of a thiolactone unit onto a suitable solid support, building blocks **5.3.2** and **5.3.3** were evaluated as chain extension moieties.

(5.3.1)   **(5.3.2)**   **(5.3.3)**
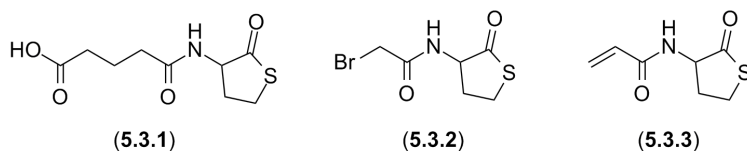
Figure 5.8: Thiolactone-containing building blocks for loading (5.3.1) and chain extension (5.3.2 and 5.3.3).

As the repetitive aminolysis and chain extension steps occur in basic medium, an acid-labile linkage was foreseen for final cleavage from the solid support. Consequently, the carboxylic acid functionalized thiolactone linker (**5.3.1**) was coupled to a 2-chlorotrityl resin using standard conditions (see experimental part).[339] The thiolactone building block for chain extension should be equipped with a functional handle to allow reaction with the on-resin generated thiol preferentially in a nucleophilic fashion. Therefore, we have chosen N-(2-bromoacetyl) homocysteine-γ-thiolactone (**5.3.2**) and N-(acryloyl) homocysteine-γ-thiolactone (**5.3.3**), which are susceptible to thio-bromo substitution and thiol-Michael addition, respectively.

Building blocks **5.3.2** and **5.3.3** were subsequently tested in the aminolysis/chain extension method. Aminolysis of the resinbound thiolactone was performed by overnight treatment of the swollen resin with an excess of benzylamine in CH_2Cl_2, thus guaranteeing full thiolactone conversion. LCMS analysis of the sample after acidic cleavage revealed quantitative consumption of the

thiolactone, but only the corresponding disulfide of the expected thiol could be identified (Fig. 5.9a). This inevitable disulfide formation during aminolysis is independent of the identity of the amine because the corresponding disulfide was detected for a variety of (functionalized) amines (see below). As a free thiol is required for the subsequent coupling step, a first attempt to avoid disulfide formation involved lowering the amine concentration. However, the thiolactone conversion was negatively influenced while the disulfide remained the dominant reaction product. The disulfides formed could be partially cleaved in a postaminolysis treatment by using dimethylphenylphosphine (Me_2PhP) as a reducing agent (Fig. 5.9b). However, their complete reduction and the subsequent conservation of the obtained resin-bound thiols appeared challenging. The synthetic strategy was thus adapted to consider the resin-bound disulfides generated upon aminolysis as stable intermediates. With respect to the two-step iterative method (Fig. 5.7), reduction of the disulfide by phosphine treatment followed by immediate in situ reaction of the generated thiol[340] with the next monomer building block is indeed an alternative for the proposed chain extension. However, these conditions were found not to be applicable when using building block **5.3.2** owing to the incompatibility of a bromide leaving group with a trialkylphosphine. On the other hand, considering chain extension by using building block **5.3.3**, the phosphine reagent can fulfill a dual role: both as a reducing agent for the disulfide generated upon the aminolysis and as an efficient catalyst for the thiol-Michael addition.[341, 342] Cleavage of the resin-bound disulfide and Michael addition of the thiol to acrylamide **5.3.3** could indeed be successfully performed in the presence of an excess of Me_2PhP and building block **5.3.3**, as demonstrated by the HPLC trace of the obtained conjugated product (Fig. 5.9c).

To demonstrate the general potential of this methodology in terms of versatility and tolerance towards functional groups, a small library of functionalized sequences was prepared by two iterations of the elaborated two-step method. A random selection of nonfunctionalized primary or secondary amines (n-propylamine, benzylamine, and pyrrolidine) for the first ring-opening step and functionalized primary amines (allylamine, propargylamine, furfurylamine, and N-(3-aminopropyl) morpholine) for the consecutive aminolysis allowed
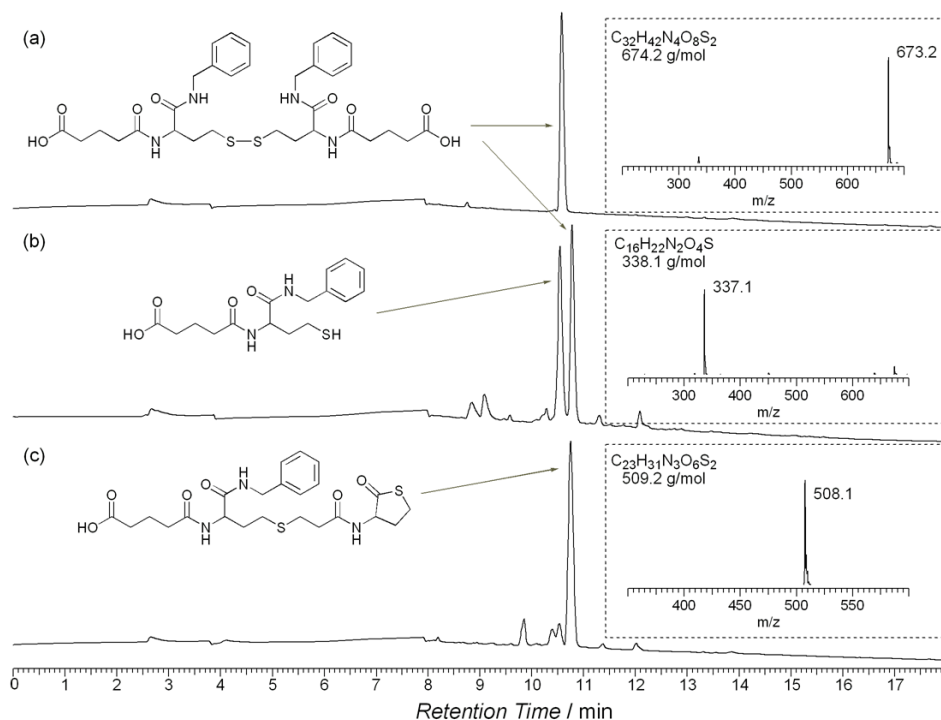
**Figure 5.9: LC-MS analysis of cleaved samples** after a) aminolysis of the resin-bound thiolactone with benzylamine, generating the corresponding disulfide, b) on-resin reduction of the formed disulfide by phosphine treatment, and c) on-resin reduction of the formed disulfide and sequential one-pot chain extension through thiolacrylamide conjugation. Inserts: ESI-MS spectra of the dominant species (negative mode).

for the synthesis of six functionalized dimers (Figure 5.10a and experimental part). By reversing the order in which the amines were employed, the isomeric sequences for each combination were obtained in good purity, thus illustrating the possibility to control the sequence without significant interference from the introduced functional groups. (Figure 5.10b and experimental part). Different functional handles could consequently be incorporated into a single, longer oligomeric motif through the application of an extra aminolysis/chain extension cycle. This was demonstrated by extension of the benzylaminepropargylamine sequence (Figure 5.10c) through a third iteration using furfurylamine during the aminolysis, thus yielding a multifunctionalized sequence-defined oligomer suitable for orthogonal modification.[138, 252, 334, 343, 344]

To further define the scope and limitations of the described method, the maximum number of repetitions of the synthetic cycle to generate longer sequences was determined. While the consecutive overnight reactions in the current method render the overall process time-consuming, it was demonstrated that heating through microwave irradiation significantly reduces the reaction times of both steps (Fig. 5.11). This accelerated microwave-assisted method was applied to the preparation of two pairs of structurally related trimer and tetramer sequences in good purity. Attempts to extend the tetramer to a multifunctionalized pentamer were only partially successful because side reactions tend to become predominant (Fig. 5.12).

The trimers and tetramer were purified and subjected to HRMS (Figures in experimental part) and 2D NMR (500 and 700 MHz) analysis, thus enabling full characterization of the obtained oligomeric species (performed by Dr. Pieter Espeel, see Appendix B). The appearance of amide NH resonances in the ^{1}H-NMR spectra (d=7.88.5 ppm) is highly specific (like a fingerprint) for each oligomer as a result of the significant influence of the amine residue on the chemical shifts of the amide protons in both the side chain and the backbone (Fig. 5.13).

In summary, a thiolactone-based approach for the solid supported preparation of multifunctionalized sequence defined oligomers was successfully developed. The elaborated two-step iterative method, consisting of an on-resin aminolysis and
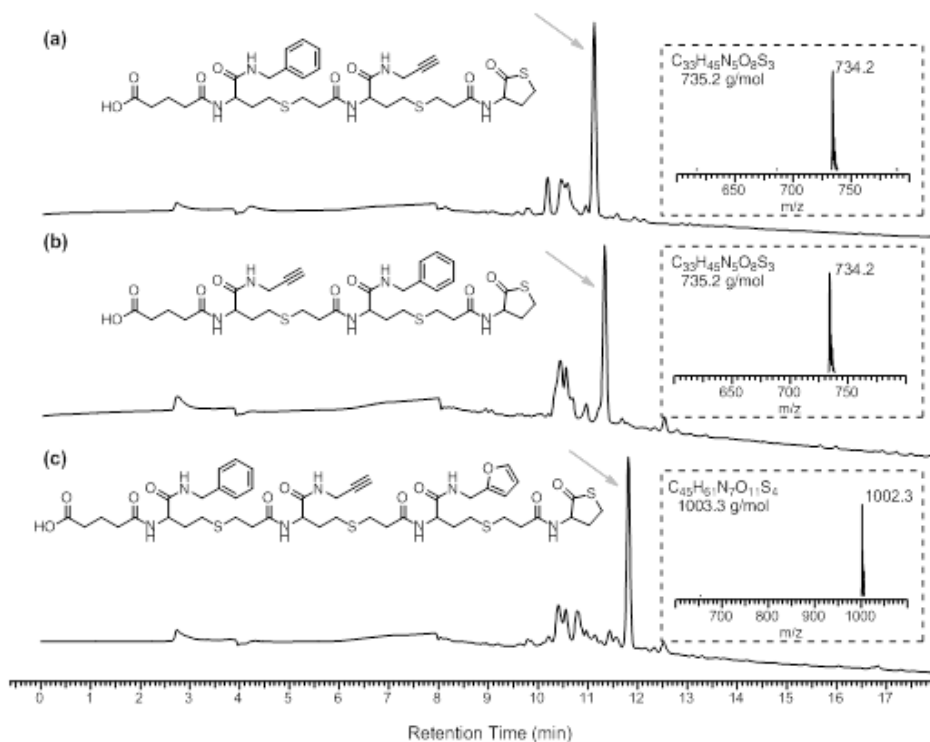
Figure 5.10: LC-MS analysis of the functionalized oligomers originating from repetitive on-resin aminolyses, using different (functional) amines, and chain extensions: a) benzylamine and propargylamine, b) propargylamine and benzylamine, and c) benzylamine, propargylamine and furfurylamine. Inserts: ESI-MS-spectra of dominant species (negative mode).
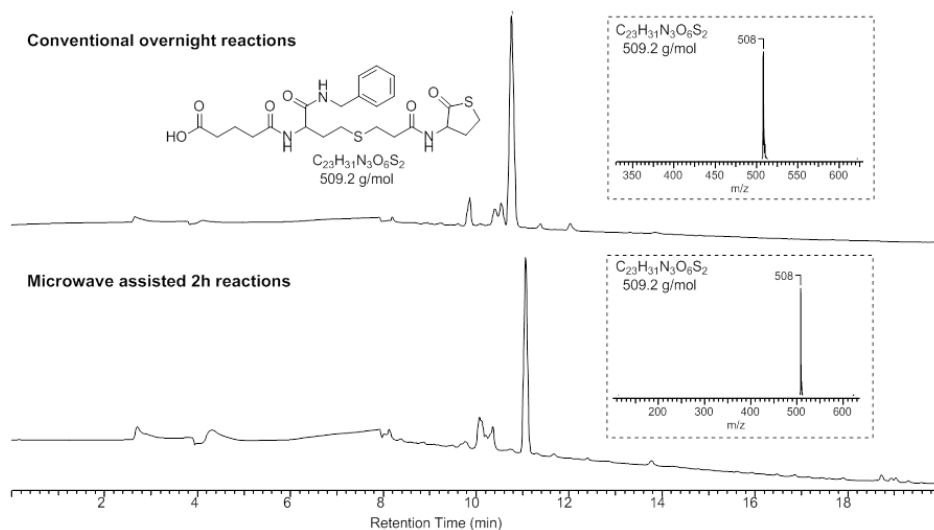
Figure 5.11: LC-MS analysis of two reaction sequences (aminolysis with benzylamine and subsequent chain extension), using conventional overnight reactions at ambient conditions (top) or a microwave-assisted accelerated protocol (bottom).

subsequent chain extension through Michael addition, does not require protecting groups and relies on the use of a single, readily available thiolactone acrylamide building block. Several functionalized short (up to tetrameric) sequences were obtained by consecutive use of a variety of commercially available amines. The obtained sequence defined motifs featuring a unique backbone and a preprogrammed organization of functional side chains were fully characterized by NMR spectroscopy and HRMS. These species could potentially exhibit particular folding and self-association properties, predominantly determined by the nature of the functionalized amines. In this regard, high-throughput automation and control

Figure 5.12 *(facing page)*: **LC-MS analysis of the functionalized oligomers originating from repetitive on-resin aminolyses, using different amines, and chain extensions**: a) trimer: *3-morpholinopropylamine - 2-methoxyethylamine -benzylamine*; b) tetramer: *3-morpholinopropylamine - 2-methoxyethylamine - benzylamine - furfurylamine*; c) pentamer: *3- morpholinopropylamine - 2-methoxyethylamine - benzylamine - furfurylamine - benzylamine*; d) pentamer: *3- morpholinopropylamine - 2-methoxyethylamine - benzylamine - furfurylamine - n-octylamine*. All samples were prepared via the microwave-assisted protocol and were analyzed as crude mixtures after cleavage from the resin.
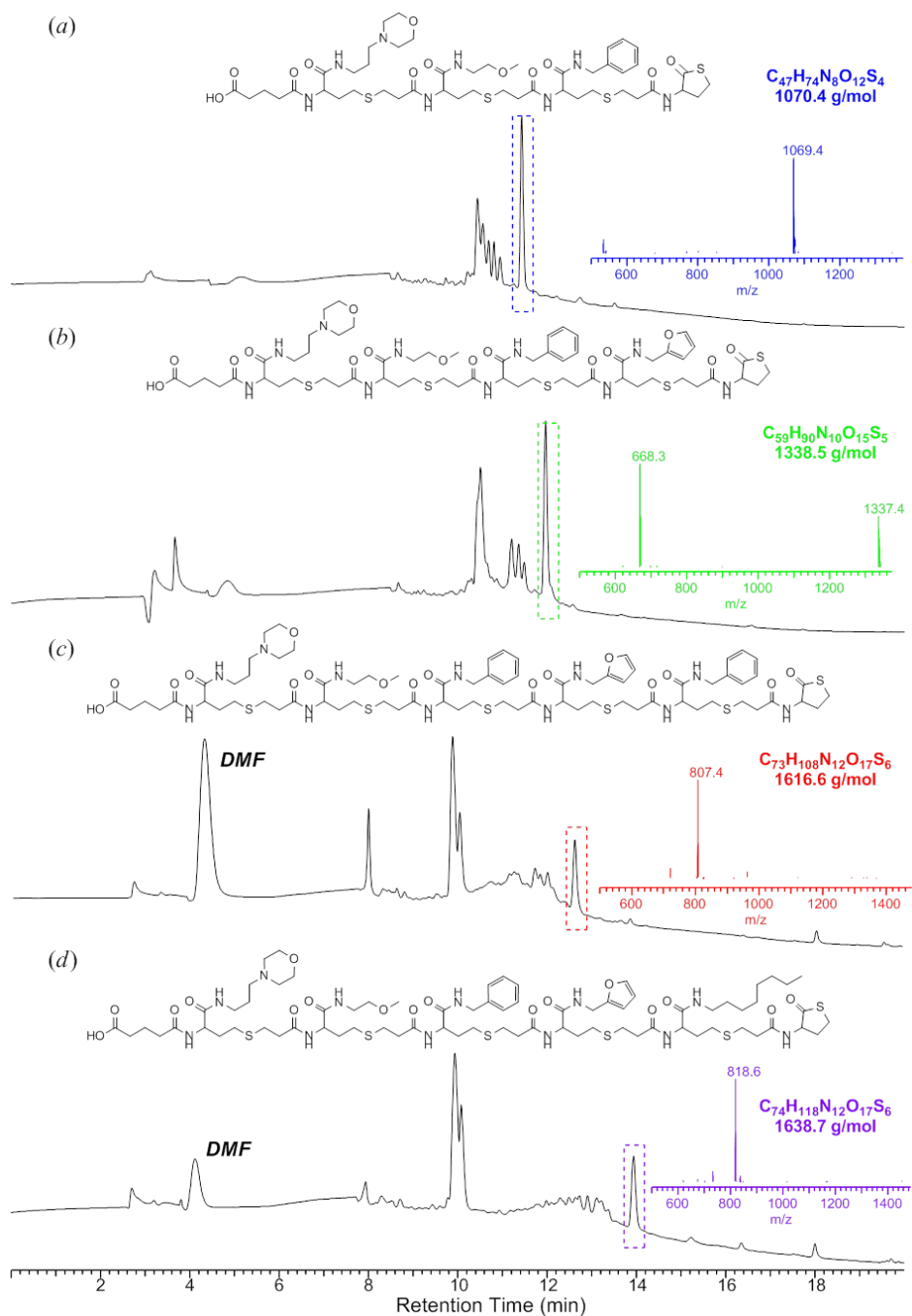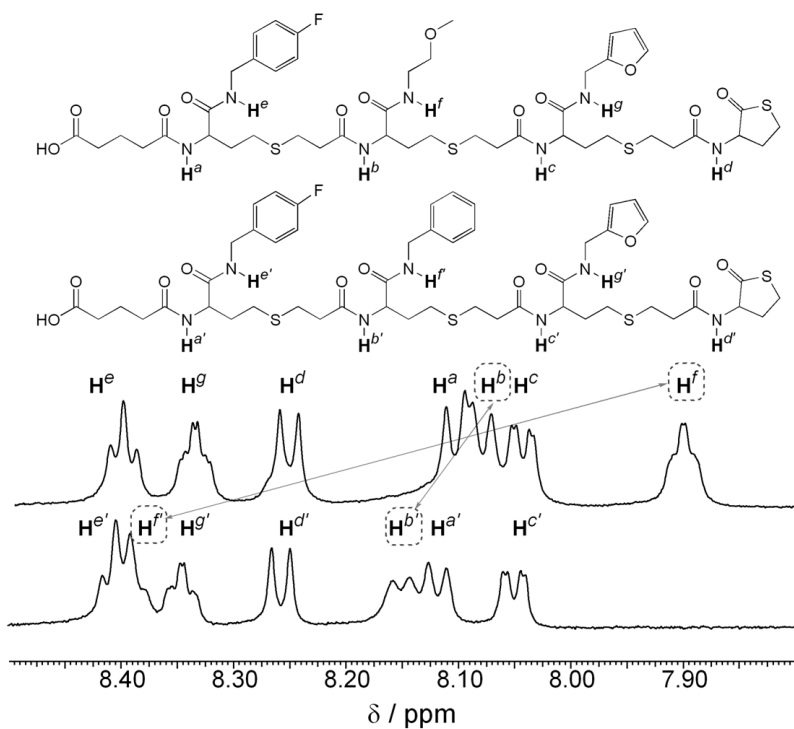
Figure 5.13: The amide signals in the ^{1}H-NMR spectra (500 MHz, [D6]DMSO) of two structurally related trimeric species.

over the stereochemistry, induced by the use of enantiomerically pure building blocks, will be further explored. Moreover, through further application of thiolactones as latent thiol functionalities and installation of appropriate functional end groups, these short heterotelechelic fragments are susceptible to periodic polymerization, a process expected to afford larger sequence-controlled polymers. This promising reconstitution approach[311] will be the focus of our future research efforts.

Chapter 6

Conclusion & Perspectives (English Summary)

DNA-protein interactions were the focal point of the work in this PhD project. These interactions are essential for cellular life and are at the basis of several biological processes like folding, replication, transcription, and degradation. The selectivity and affinity of these interactions is remarkable, as they allow DNA-binding proteins to find their target in the complex cellular environment. It is a major challenge to improve our understanding of these interactions, which is driven by the ambition to design drugs that can compete in such interactions. Here, DNA-protein interactions were studied using protein models and furan oxidation-based cross-linking.

6.1   Protein Models

Protein models consist of the minimal DNA binding peptide fragments taken from a protein. By assembling relevant peptide fragments on a scaffold, a desired geometry can be obtained (Fig. 6.1). If this geometry corresponds to that of the protein, its function can be mimicked. As such, these models allow to isolate and study important interactions in detail. These models are also easy to use in a variety of applications because of their synthetic and analytical simplicity.
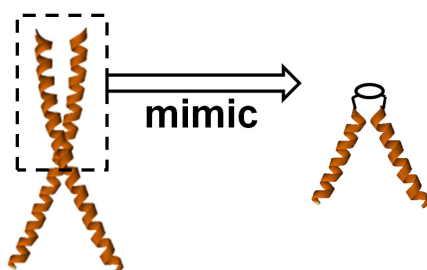
Figure 6.1: Protein models are obtained by assembling relevant peptide fragments of the protein on a scaffold.

6.1.1   Madder Model

A dipodal peptidosteroid model, developed in our group for enhanced pharmacokinetic properties, was revised and optimized. Through the incorporation of flexible linkers in the design, the geometry of the natural transcription factor protein could be better approached and for the first time with a peptidosteroid, selective DNA binding could be achieved. The increased flexibility did promote aggregation, which drastically complicated the synthesis. To obtain the peptidosteroid, optimization of the synthesis conditions with microwave irradiation was required. The synthesis on its own was thus of interest to evaluate the potential of microwave irradiation in solid phase peptide synthesis. Further research will focuss on a more convergent synthesis approach and the analysis of the pharmacokinetic properties.
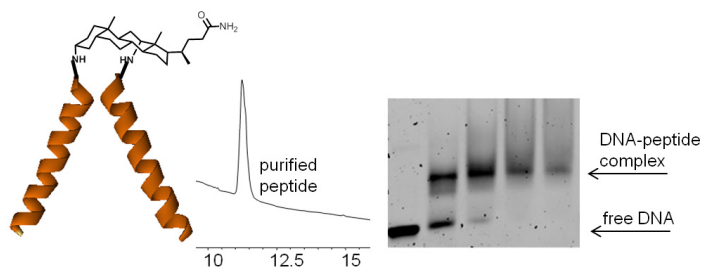
Figure 6.2: Madder Model Schematic representation of the dipodal peptidosteroid (left), HPLC trace of the purified construct after successful synthesis using microwave assistance (middle) and the polyacrylamide gel showing a supershift due to binding with the DNA target (right).

6.1.2 Morii Model

The Morii model, based on an adamantane-cyclodextrin inclusion complex, was chosen from several model systems reported in literature because of its non-covalent nature. For use in cross-link reactions, this is an advantage as it allows simplification through the use of smaller fragments. Therefore, this model was used in the next chapter.
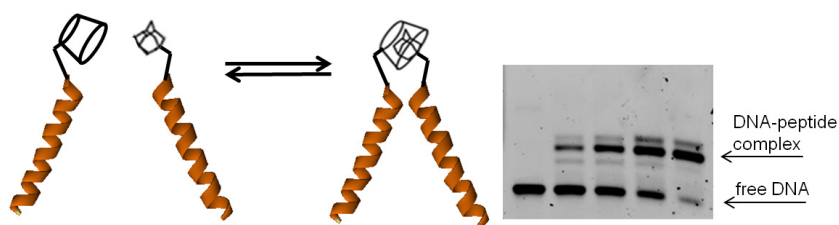
Figure 6.3: Morii Model Schematic representation of the non-covalent adamantane-cyclodextrin dipodal peptide (left) and the polyacrylamide gel showing a supershift due to binding with the DNA target (right).

6.2   Furan Oxidation Based Cross-Linking

The DNA-protein interaction was further studied by furan oxidation based cross-linking. This methodology is inspired by the toxicity of furan. In the liver, furan is oxidized into a very reactive butenedial, which covalently links biomolecules to each other, with for example formation of DNA ICLs. By synthesizing furan-modified oligonucleotides (ODNs), which is easy due to the relative stability of the furan and its commercial availability, this principle was successfully developed in our group for the controlled formation of DNA ICLs in a selective and high yielding manner (Fig. 6.4). Over the years, a whole range of furan-modified nucleosides has been synthesized and tested, resulting in a toolbox with different selectivity and yield (Fig. 6.5). In general however, the use of synthetic DNA ICLs is limited and has few applications besides their role in the study of DNA repair enzymes that cause resistance to chemotherapeutics. Therefore our group was interested to further explore the possibilities. The covalent enforcement of the transient DNA-protein interaction allows for more easy study and analysis. Moreover, a covalent and thus permanent blockage could have therapeutic applications.
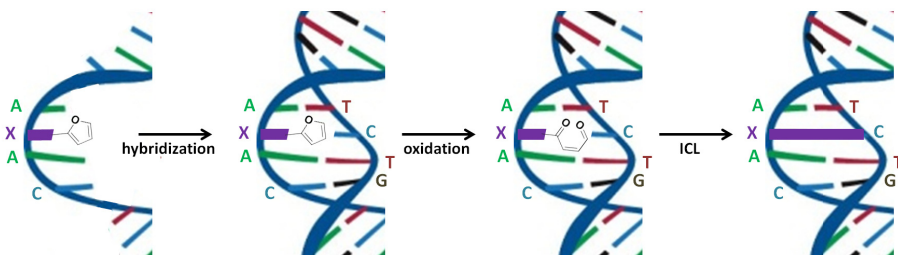
Figure 6.4: Furan oxidation based DNA interstrand cross-linking The Furan methodology was developed and successfully applied for DNA interstrand cross-linking.
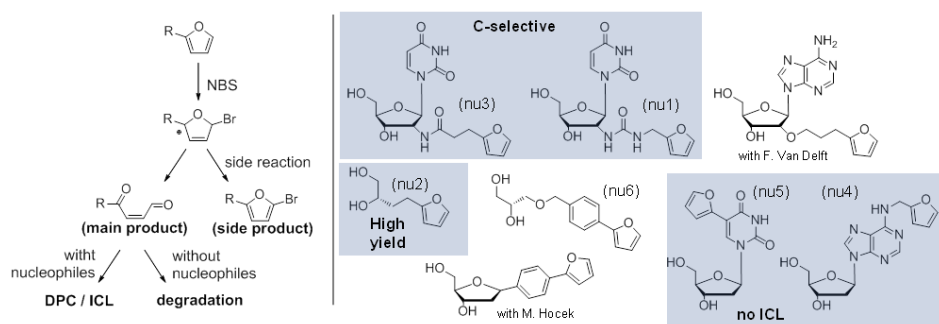
Figure 6.5: Furan oxidation Mechanism of cross-link formation (left). Furan-modified nucleosides for ICL formation, a toolbox for different selectivity and yield (right).

To apply the furan cross-linking methodology on this interface, two different approaches can be followed as illustrated in Figure 6.6: modify the DNA like before to target proteins or modify proteins to target DNA like before. Modifying the DNA on the one hand, could be used to create cross-linking decoys regulating gene expression by covalent capture and blockage of transcription factors in the cytoplasm. Modifying DNA binding proteins on the other hand, like transcription factors, with a furan for cross-linking to their natural DNA target could block the DNA an thus also block transcription.
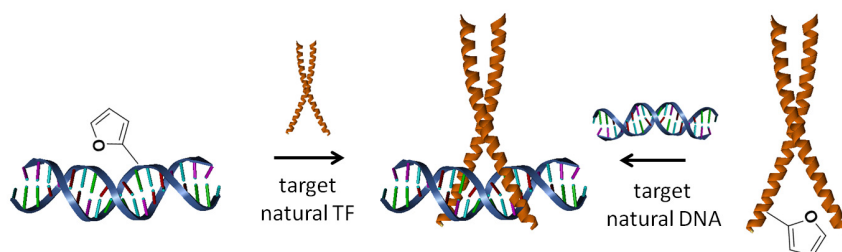
Figure 6.6: Different approaches for furan based cross-linking on the DNA-protein interface.

6.2.1   DNA → protein cross-linking

In case of reaction from furan-modified DNA to a bound protein, competitive ICL should be avoided. This was possible by using **nu5**, which features a furan on the base (Fig. 6.7). This furan is positioned away from the Watson-Crick base pairing interface, into the major groove of the DNA where proteins bind. Therefore, **nu5** did not give significant levels of ICL formation. It was rationalized that in this case, DPCs (DNA-protein cross-links) are formed through a Schiff base, which can be stabilized by $NaCNBH_3$ reduction (Fig. 6.8 top).
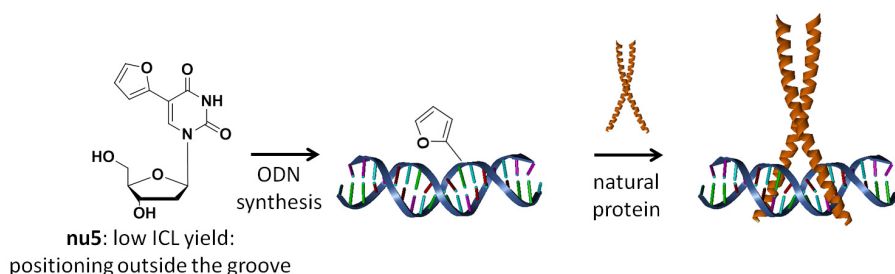
Figure 6.7: Furan-modified nucleoside for use in the synthesis of furan-modified ODNs, which were tested for cross-linking to their natural binding protein.

DPC formation could be observed after combination of the transcription factor model with its **nu5** modified DNA binding site, oxidation and reduction (Fig. 6.8 bottom). This was also confirmed by mass analysis. Interestingly, differential positioning of **nu5** in the duplex resulted in different cross-link yields, which decreased with increasing distance. As such furan-modified ODNs can be used as distance scanning probes.

Kinetin-modified DNA, that occurs in natural DNA as a result of oxidative stress is also base modified with a furan (**nu4**). It was shown that this furan-modification reacts under oxidative stress, functioning as antioxidant, but does not lead to the formation of toxic DNA ICLs. The formation of DPCs could not be shown in this case.
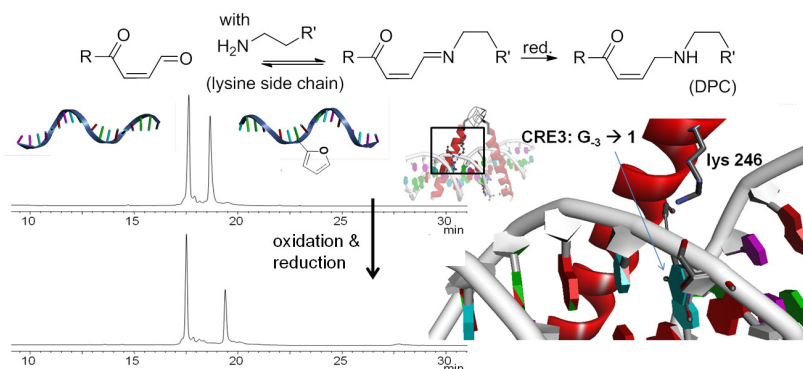
Figure 6.8: DNA → protein cross-linking. Mechanism (top), HPLC trace before and after furan oxidation and reduction (bottom left), detail of the DNA-protein interaction with the oxidized furan (bottom right).

6.2.2 DNA ← protein cross-linking

Alternatively, the protein model can be modified with furan for cross-linking to the DNA bases. For the furan-modification of the mimic, several furan-modified amino acids were synthesized next to commercial furfurylalanine, to assess the influence of linker length and type (Fig. 6.9). Finally, tests were performed with **aa1**, as **aa3** has a risk for intramolecular reactions, **aa4** is unstable under the applied peptide synthesis conditions and **aa6** has an electron poor deactivated furan. **aa5** remains to be tested.
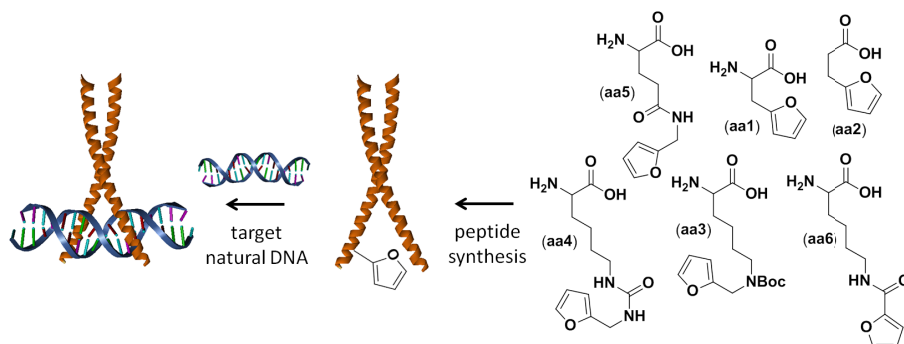
Figure 6.9: Furan-modified amino acids for the synthesis of furan-modified peptides, which were tested for cross-linking to their natural DNA binding site.

Furfurylalanine was incorporated at 2 different positions in the peptide, both directed towards the nucleophilic functionalities in the DNA with which it should react. Care was taken to conserve essential residues, to avoid loss of binding affinity. The peptides could be synthesized under standard synthesis conditions. Although oxidation and degradation of the furan moiety had been described under acidic peptide cleavage and deprotection conditions, no problems were encountered. Oxidation of the furan was checked by trapping the reactve 4-oxobutenal with hydrazine. DPC formation (under a range of different conditions) however, was not observed. Only some degradation occured upon oxidation, like shown in Figure 6.10.
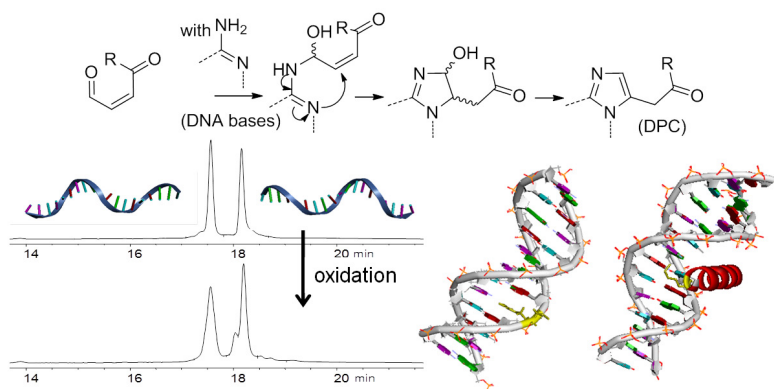
Figure 6.10: DNA ← protein cross-linking. Mechanism (top), HPLC trace before and after furan oxidation (bottom left), difference between ICL where reaction takes place with the opposite ODN and DPC where reaction has to occure with a duplex (bottom right).

Though unfavorable positioning and/or linking cannot be completely excluded, the results led to the hypothesis that it is rather difficult to cross-link to duplex DNA. The involvement of the targeted exocyclic amines in Watson-Crick base pairing in the highly organized duplex, strongly reduces their reactivity towards cross-linking. This hypothesis is further supported by the reported use of aldehydes for the detection of single stranded regions in DNA and their use to elucidate RNA folding pathways. The difference between ICL and DPC is depicted in Figure 6.10. While the reactive moiety from the peptide has to intrude in and interfere with the highly organized duplex structure for DPC formation, for ICL formation the reactive moiety is part of the duplex. Therefore the base opposite the furan-modification is unpaired and thus readily available for reaction. The introduction of distortions in the target duplex (by means of a mismatch, a flipout base or a temperature increase) to facilitate cross-linking was not possible as this disturbed binding of the model, which is dentrimental for the cross-link reaction. It could thus be concluded that cross-linking with the furan oxidation-based methodology is not efficient towards duplex DNA and that the use of a protein model is not ideal for further investigation with facilitated cross-link formation.

Though initially considered a negative result and a disadvantage, this is more and more considered as a feature that could be used for structural elucidation in complex systems and to distinguish single from double stranded DNA e.g in topoisomerases. To prove this hypothesis, cross-linking from a furan-modified protein to a single stranded ODN is required. Currently, 2 possibilities are under investigation. Cross-linking from a protein to an RNA hairpin, which is more distorted by mismatches, flipout bases and a loop (Fig. 6.11, left) should be possible. Also cross-linking from a protein that binds single stranded DNA, like polymerases should work. After initial experiments in our group and the successful experiments by Prof. Summerer, it is now also clear that furan can be incorporated in proteins by amber supression and recombinant expression. This paves the way for a good many new possibilities.
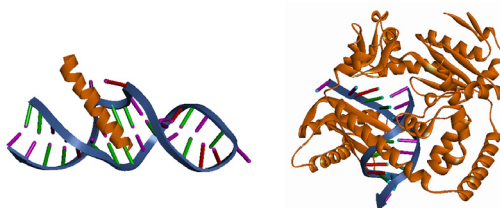
Figure 6.11: Model systems for proving the hypothesis that furan based cross-linking can differentiate between single and double stranded ODN/ON RRE RNA-Rev peptide interaction (left), polymerase associated with the single stranded template DNA (right).

6.2.3   RNA ICL

To test the mildness and generality of the methodology, to prepare for protein-RNA cross-linking and because RNA is becomming an increasingly important therapeutic target, RNA ICL formation was examined. Rather suprisingly, it was found that RNA sustained the oxidation conditions well and that high cross-link yields were achieved. The different duplex shape was reflected in a different furan-modified nucleoside preference. In this case, the best results in yield and selectivity were obtained with **nu3** (cfr. for DNA ICL better yields were obtained with **nu6** than with **nu3**).
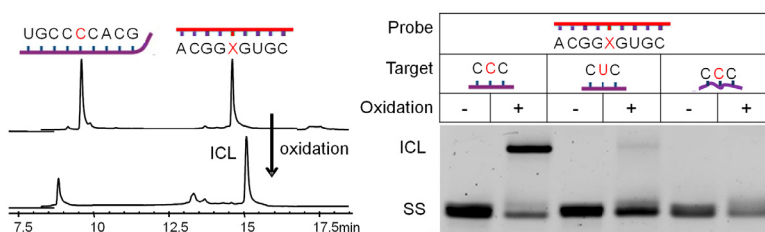
Figure 6.12: RNA ICL met X = **nu3**. RP-HPLC trace of the RNA ICL reaction before and after NBS oxidation (left). Denaturing PAGE gel with three different targets: complementary RNA with a C or a U opposite **nu3** and a randomized non-complementary RNA with a C opposite **nu3** (right).

6.3 Methodology Optimalisation

Next to the original goals in DNA-protein context, which constitute the main theme of this work, some results were accomplished in the optimalization of solid phase methodology.

- The previously developed NF31 color test was shown to detect residual OH groups in novel solid phase resins.

- The potential of NMR for the concentration determination of biomacromolecules was tested in collaboration with the NMRSTR group.

- In collaboration with the PCR group, specialized in thiolacton chemistry, a new and versatile solid phase coupling methodology for the synthesis of sequence controlled oligomers was developped (Fig. 6.13).
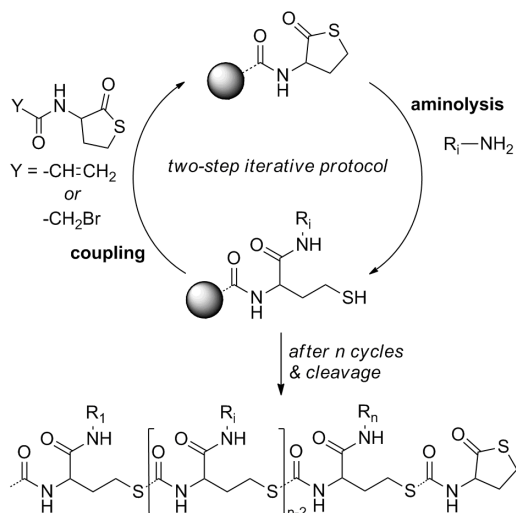
Figure 6.13: Two-step iterative method for the synthesis of functionalized oligomers on solid support: aminolysis of the resin-bound thiolactone followed by coupling of a thiolactone-containing building block that selectively reacts with the generated thiol.

Conclusion

The initial goal of this PhD work, to design a cross-linking transcription factor mimic with possible therapeutic applications, was not achieved. Cross-linking to a dubble stranded DNA helix appeared to be particularly challenging. The synthesis of the dipodal peptidosteroid mimic was however optimized and DNA binding achieved for the first time. Furthermore, the work described in this thesis has contributed to a better understanding of and insight in the furan oxidation methodology and significantly extended its applicability. Next to furan incorporation in ODNs for DNA interstrand cross-link formation, RNA interstrand cross-link formation was shown to be also possible in ON duplexes. Furan-modified duplexes can, with the choice of an appropriate modified nucleoside, also be used for the formation of DPCs. Furan has also been incorporated in peptides. DPC formation could not be shown in this case, but there are high expectations for the cross-linking in not-duplex complexes. Finally, it has been demonstrated that furan modifications can be site selectively incorporated in proteins, which opens up enormous possibilities.
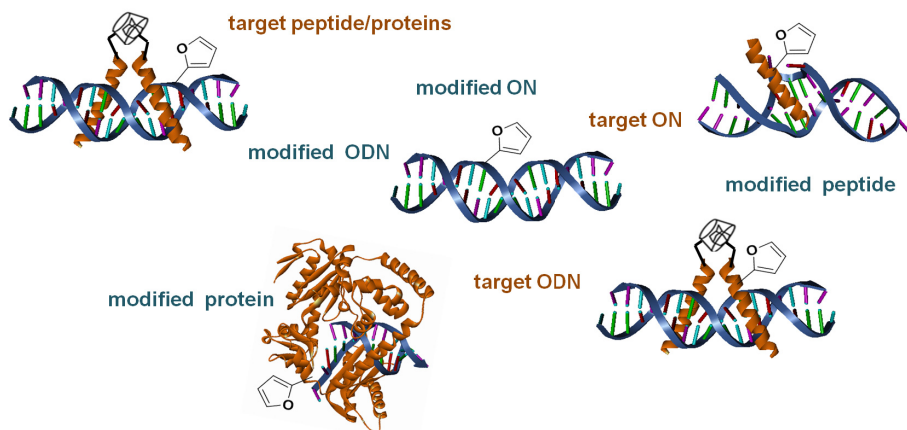
Figure 6.14: Summary of the extended applicability of the furan cross-linking approach showing the different available modifications (text in blue) and targets (text in orange).

Perspectives

Now, the following steps should persue the use of the methodology in relevant applications and in a biological context. With this purpose in mind, it is important to deal with concentration related issues and use of a biocompatible oxidation method. The production of singlet oxygen by red light irradiation of sensitizers has been developed for DNA ICL formation and was shown to be compatible with RNA ICL formation. Moreover, tests for DPC formation have been conducted and this was the method used for the reported RNA-protein cross-linking. Singlet oxygen can however still cause a lot of collateral damage. Theoretically reasonable and more selective alternatives, for which initial tests have been performed are the oxidation by cytochrome P450 that is responsible for the natural oxidation of furan in the liver [345–347] and electrochemical oxidation.[348–350] Especially the latter method holds a lot of potential. Because furan has a lower oxidation potential than the DNA/RNA bases,[351–353], a current will only flow (at the correct potential) in the presence of unoxidized furan. Moreover, the use of electrodes allows local applications.[354] The use of biological triggers for furan oxidation like oxidative stress (e. g. in cancerous cells) is an other interesting possibility.

Chapter 7

Nederlandstalige Samenvatting

De interactie tussen DNA en proteïnes stond centraal in het uitgevoerde doctoraatswerk. Deze interactie is essentieel voor cellulair leven, aangezien ze aan de basis ligt van verschillende biologische processen, zoals opvouwing, replicatie, transcriptie en degradatie. De selectiviteit en affiniteit die DNA bindende proteïnes toelaat hun target te vinden in de complexe cellulaire omgeving is uiterst opmerkelijk. Inzicht hierin verkrijgen is een interessante uitdaging, gedreven door de ambitie om geneesmiddelen te ontwikkelen die kunnen interfereren in deze interactie. De DNA-proteïne interacties werden hier bestudeerd aan de hand van proteïne modellen en furan-oxidatie gebaseerde cross-linking.

7.1   Proteïne Modellen

Proteïne modellen zijn opgebouwd uit de minimale DNA bindende peptide fragmenten van een proteïne. Door relevante peptide fragmenten te combineren op een scaffold, kan een gewenste geometrie bekomen worden (Fig. 7.1). Indien deze geometrie gelijkaardig is aan die van het proteïne, kan de functie nagebootst worden. Deze modellen laten zo toe de belangrijke interacties te isoleren en in detail te bestuderen. Deze modellen zijn ook gemakkelijk voor gebruik in allerlei toepassingen door hun synthetische en analytische eenvoud.
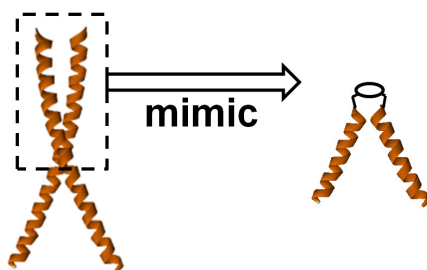
Figure 7.1: Proteïne modellen worden bekomen door relevante peptide fragmenten correct te combineren op een scaffold.

7.1.1   Madder Model

Een dipodaal peptidosteroide model, ontwikkeld in onze groep voor verbeterde farmacokinetische eigenschappen, werd herzien en geoptimaliseerd. Door het toevoegen van flexibele linkers aan het ontwerp, kon de geometrie van het natuurlijke transcriptiefactor proteïne beter benaderd worden en werd selectieve DNA binding met een peptidosteroïde voor het eerst gerealiseerd. De toegenomen flexibiliteit bemoeilijkte de synthese wel zodanig dat de condities geoptimaliseerd moesten worden. Microgolf assistentie bleek een oplossing te bieden en daarom was de synthese op zich ook interessant om de mogelijkheden van microgolf geassisteerde vaste fase peptide synthese te evalueren. Verder onderzoek zal zich richten op een meer convergente synthesestrategie en de analyse van de farmacokinetische eigenschappen.
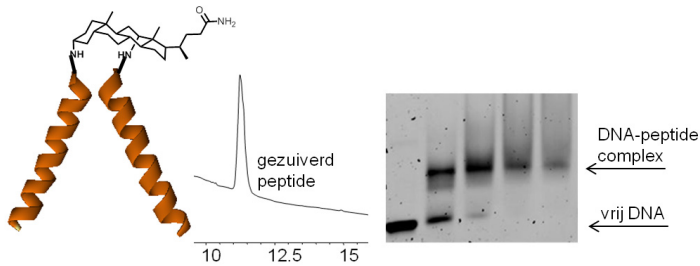
Figure 7.2: Madder Model Schematische voorstelling van het dipodaal peptidosteroide (links), HPLC trace van het opgezuiverde construct na successvolle synthese met microgolf assistentie (midden) en de polyacrylamide gel die een supershift vertoont door binding met de DNA bindingssite (rechts).

7.1.2 Morii Model

Het Morii model, gebaseerd op een cyclodextrine-adamantaan inclusie complex, werd geselecteerd uit de verschillende modelsystemen beschreven in de literatuur omwille van zijn niet-covalente natuur. Voor het gebruik in cross-link reacties is dit een voordeel omdat dit een vereenvoudiging toelaat, er kan immers met kleinere fragmenten gewerkt worden. Om die reden werd dit model dan ook gebruikt in het volgende hoofdstuk.
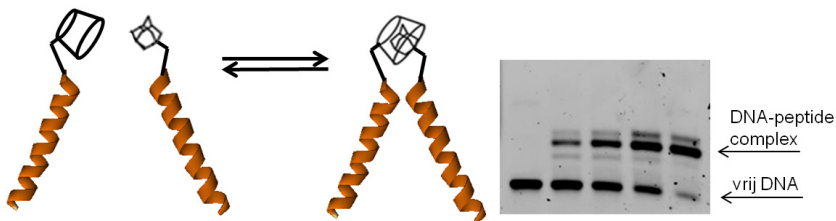
Figure 7.3: Morii Model Schematische voorstelling van het niet covalente adamantaan-cyclodextrine dipodaal peptide (links) en de polyacrylamide gel die een supershift vertoont door binding met de DNA bindingssite (rechts).

7.2   Furan Oxidatie Gebaseerde Cross-Linking

De DNA-proteïne interactie werd verder bestudeerd door middel van furan oxidatie gebaseerde cross-linking.   Deze cross-link methode is geïnspireerd op de toxiciteit van furan. Furan wordt in de lever door oxidatie omgezet tot het heel erg reactieve buteendial, dat biomoleculen covalent aan elkaar linkt door de vorming van bijvoorbeeld DNA interstrand cross-links (ICL). Aan de hand van de synthese van furan gemodificeerde oligonucleotiden (ODNs), wat eenvoudig is door de relatieve stabiliteit en commerciële beschikbaarheid van furan, kon dit principe in onze groep gebruikt worden voor de gecontroleerde en selectieve vorming van DNA ICLs met hoge opbrengst (Fig. 7.4).   Doorheen de jaren, werd een hele reeks furan gemodificeerde nucleosiden gesynthetiseerd en getest, wat geleid heeft tot een reeks aan varianten met verschillende selectiviteit en opbrengst (Fig. 7.5).   In het algemeen heeft het gebruik van artificiële DNA ICLs echter weinig toepassingen buiten hun gebruik in de studie naar DNA herstelenzymen, die voor resistentie tegen chemotherapeutica zorgen.   Daarom was er interesse om de verdere mogelijkheden van de methode te verkennen en onderzoeken. De covalente stabilisatie van de transiente DNA-proteïne interactie, laat een gemakkelijkere analyse en studie toe. Een covalente en dus permanente blokkage zou tevens therapeutische toepassingen kunnen hebben.
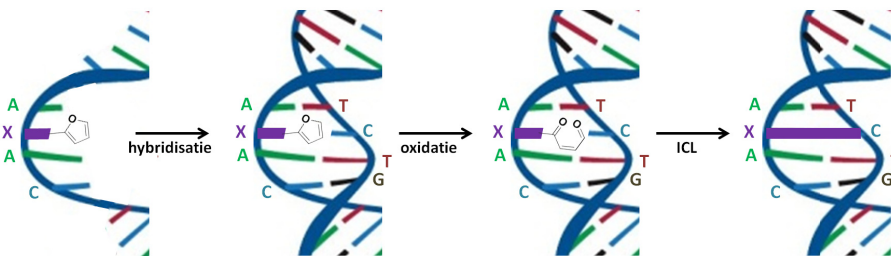
Figure 7.4: DNA interstrand cross-linking gebaseerd op furan oxidatie De furan methode werd ontwikkeld en succesvol toegepast voor DNA interstrand cross-linking.
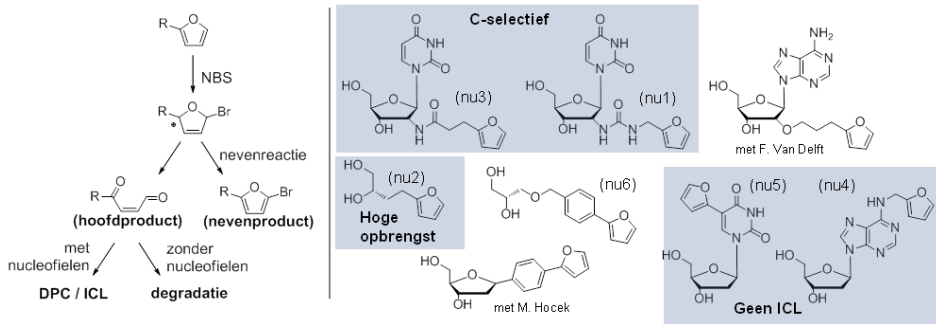
Figure 7.5: Furan oxidatie Mechanisme voor de vorming van cross-links (links). Furan gemodificeerde nucleosiden voor ICL vorming, varianten voor verschillende selectiviteit en opbrengst (rechts).

Om de furan cross-link methode op deze interface toe te passen kunnen 2 verschillende benaderingen gevolgd worden, zoals voorgesteld in Figuur 7.6: een ODN wordt gemodificeerd met furan zoals voor ICL vorming en reageert met een natuurlijk DNA-bindend proteïne. Zo wordt een cross-linkende decoy verkregen die bv. transcriptiefactoren kan blokkeren in het cytoplasma. Anderzijds, kan een DNA-bindend proteïne gemodificeerd worden met furan, voor reactie naar een natuurlijke DNA bindingssite zoals bij ICL. Door het DNA zo te blokkeren kan transcriptie ook verhinderd worden.
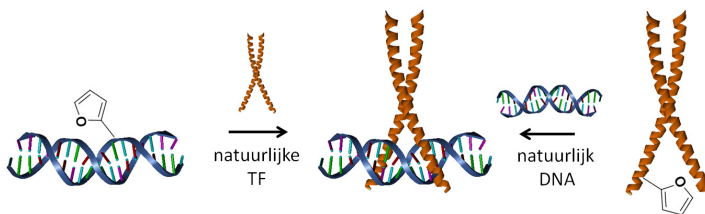
Figure 7.6:   Verschillende benaderingen die kunnen gevolgd worden voor de toepassing van furan gebaseerde cross-linking op de DNA-proteïne interface

7.2.1   DNA → proteïne cross-linking

Voor reactie van het DNA naar een gebonden proteïne, moet competitieve ICL vermeden worden. Dit was mogelijk door gebruik te maken van **nu5**, gemodificeerd met een furan op de base (Fig. 7.7). Dit furan bevindt zich in de grote groeve van het DNA, waar de proteïnen binden en weg van de Watson-Crick base paring interface. Daarom veroorzaakt **nu5** geen significante ICLs. Er werd gerationaliseerd dat DPCs (DNA-proteïne cross-links) in dit geval gevormd worden via een Schiffse base, die kan gestabiliseerd worden door NaCNBH$_3$ reductie (Fig. 7.8 boven).
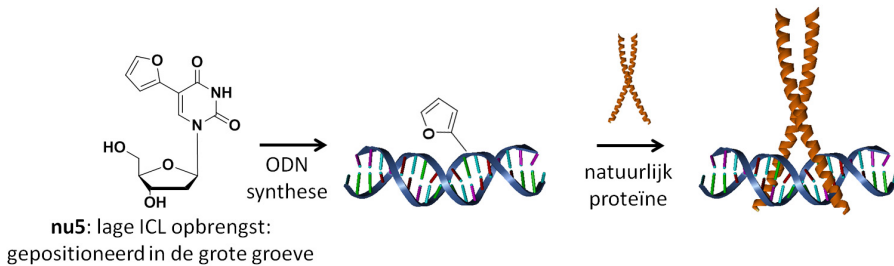
Figure 7.7: Furan gemodificeerd nucleoside voor gebruik in de synthese van furan gemodificeerde ODNs, die dan getest werden voor cross-linking naar hun natuurlijk bindend proteïne.

DPC vorming kon vastgesteld worden na combinatie van het transcriptie-factor model met zijn **nu5** gemodificeerde DNA bindingssite, oxidatie en reductie (Fig. 7.8 onder). Dit werd ook bevestigd door massa analyse. Verschillende positionering van **nu5** in de duplex resulteerde in een duidelijk verschillend cross-link rendement, dat afnam met toenemende afstand. Furan gemodificeerde ODNs kunnen dus gebruikt worden als afstand scannende probes.

Kinetine gemodificeerd DNA, wat voorkomt in natuurlijk DNA ten gevolge van oxidatieve stress is eveneens op de base gemodificeerd met een furan (**nu4**). Er werd aangetoond dat deze furan modificatie reageert onder oxidatieve stress en dus werkt als antioxidant, maar niet leidt tot de vorming van toxische DNA ICLs. De vorming van DPCs kon in dit geval niet aangetoond worden.
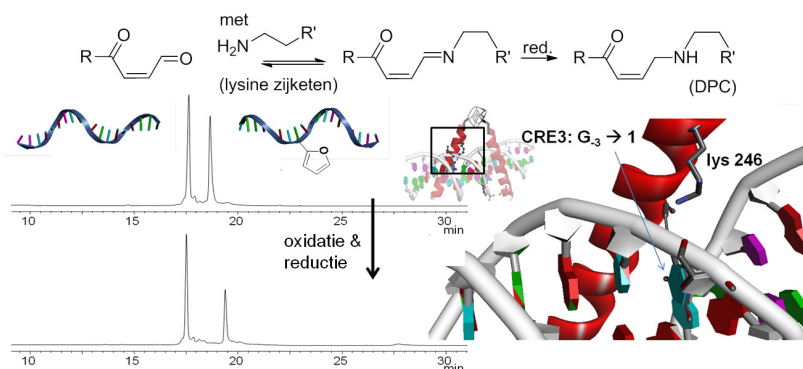
Figure 7.8: DNA → proteïne cross-linking. Mechanisme (boven), HPLC chromatogram voor en na furan oxidatie en reductie (linksonder), detail van de DNA-proteïne interactie met het geoxideerde furan (rechtsonder).

7.2.2 DNA ← proteïne cross-linking

Alternatief kan het proteïne model gemodificeerd worden met furan voor crosslinking naar de DNA basen (analoog als bij ICL). Voor de furan modificatie van het model werden, naast het commercieel beschikbare furfurylalanine, verschillende furan gemodificeerde aminozuren gesynthetiseerd om de invloed van linkerlengte en -type te kunnen inschatten (Fig. 7.9). De testen werden uiteindelijk uitgevoerd met **aa1**, aangezien **aa3** een risico inhield tot intramoleculaire reacties, **aa4** niet stabiel was onder de gebruikte peptide synthese condities en **aa6** te elektronenarm en dus gedeactiveerd was voor reactie. **aa5** moet nog getest worden.

Furfurylalanine werd geïncorporeerd op 2 verschillende posities in het peptide, beide gericht naar de nucleofiele functionaliteiten in DNA waarmee het moet reageren. Daarboven werden essentiële residuen behouden om de bindingsaffiniteit niet te reduceren. De peptiden konden gesynthetiseerd worden onder standaard condities. Hoewel furan oxidatie en degradatie onder de zure condities voor peptide afsplitsing en ontscherming beschreven is, werden geen problemen ondervonden. Oxidatie van het furan werd gecontroleerd door het 4-oxobutenal af te vangen met hydrazine. Cross-link vorming (onder een hele reeks aan verschillende condities) kon echter uiteindelijk niet worden geobserveerd. Oxidatie leidde enkel

Figure 7.9: Furan gemodificeerde aminozuren voor gebruik in de synthese van furan gemodificeerde peptiden, die dan getest werden voor cross-linking naar hun natuurlijke DNA bindingssite.

tot wat degradatie, zoals weergegeven in Figuur 7.10.

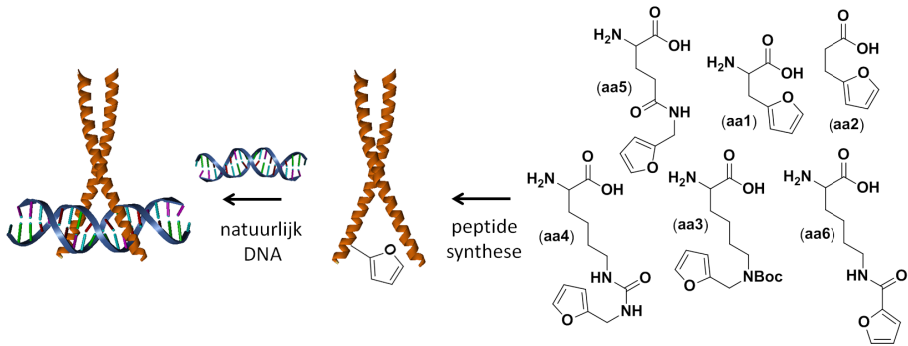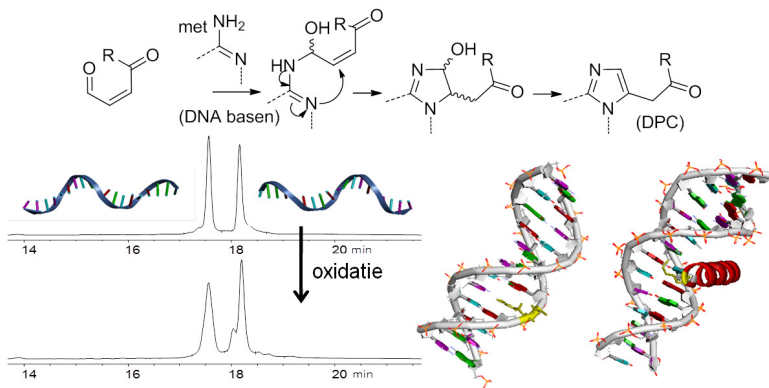**Figure 7.10: DNA ← proteïne cross-linking.** Mechanisme (boven), HPLC chromatogram voor en na furan oxidatie (linksonder), principieel verschil tussen ICL, waarbij gereageerd wordt met een oligonucleotide en DPC, waarbij er met een duplex gereageerd moet worden (rechtsonder).

Hoewel een ongunstige positionering en/of linker niet uitgesloten kunnen worden, leidden deze resultaten tot de hypothese dat het erg moeilijk is om te cross-linken naar duplex DNA. De exocyclische amines waarmee in een cross-link reactie gereageerd moet worden, zijn immers betrokken in Watson-Crick basepar-ing en daardoor minder geneigd tot verdere reactie. Deze hypothese wordt verder ondersteund door het gerapporteerde gebruik van aldehyden voor de detectie van enkelstrengige regio's in DNA en het ophelderen van RNA opvouwingspatronen. Het verschil tussen ICL en DPC vorming is weergegeven in Figuur 7.10. Terwijl de reactieve groep op het peptide moet indringen en de sterk georganiseerde duplex moet verstoren voor DPC vorming, maakt bij ICL vorming het furan deel uit van de duplex waardoor de base tegenover de furan modificatie ongepaard en dus vrij beschikbaar is. Het verstoren van de duplex (door middel van een mismatch, een flipout base of een temperatuursverhoging) om de cross-linking te faciliteren is niet mogelijk omdat de binding van het model dan wordt verstoord en dit dan weer nefast is voor cross-linken. Er kon dus geconcludeerd worden dat cross-linking met onze furan oxidatie gebaseerde methode niet efficient is naar duplex DNA en dat het gebruik van een modelsysteem niet ideaal is voor verder onderzoek met verstoring van de duplex.

Hoewel dit oorspronkelijk als een negatief resultaat en een nadeel werd beschouwd, wordt het meer en meer duidelijk dat dit een interessante eigenschap van de technologie kan zijn die kan helpen in de opheldering van complexe structuren of het onderscheiden van enkel- en dubbelstrengig DNA bv. in topoisomerasen. Om deze hypothese te bewijzen, zou er een cross-link moeten gevormd worden van een furan gemodificeerd proteïne naar een enkelstrengig ODN. Hiervoor worden momenteel 2 mogelijkheden onderzocht. Enerzijds zou de cross-linking van een proteïne naar een RNA hairpin, dat meer verstoord is met mismatches, flipout basen en een loop (Fig. 7.11, links) kunnnen lukken, anderzijds zou ook cross-linking van een proteïne dat enkelstrengig DNA bindt, zoals polymerase (Fig. 7.11, rechts) mogelijk moeten zijn. Deze experimenten vormen het onderwerp van huidige studies. Na initiële testen in onze groep en de succesvolle experimenten van Prof. Summerer, is het nu ook duidelijk dat furan ook in proteïnen geïncorporeerd kan worden door middel van amber supressie en recombinante expressie. Dit opent de weg naar heel wat nieuwe mogelijkheden.
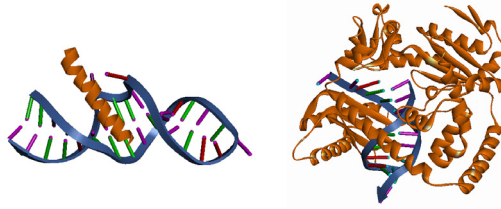
Figure 7.11: Modelsystemen voor het bewijzen van de hypothese dat furan gebaseerde cross-linking het onderscheid kan maken tussen enkel- en dubbelstrengig ODN/ON RRE RNA-Rev peptide interactie (links), polymerase geassocieerd met het enkelstrengige templaat DNA (rechts)

7.2.3   RNA ICL

Om de algemeenheid en mildheid van de methodologie te testen, proteïne-RNA cross-linking voor te bereiden en omdat RNA een steeds belangrijkere therapeutische target wordt, werd RNA ICL vorming onderzocht. Eerder onverwacht vonden we dat het RNA de oxidatiecondities goed doorstond en dat hoge cross-link rendementen bereikt konden worden. De verschillende duplex vorm werd wel weerspiegeld in een andere furan gemodificeerde nucleoside voorkeur. In dit geval gaf **nu3** immers de beste resultaten qua rendement en selectiviteit (cfr. bij DNA ICL werden betere opbrengsten verkregen met **nu6** dan met **nu3**).
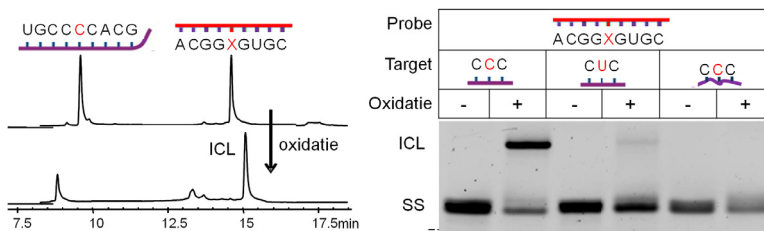
Figure 7.12: RNA ICL met X = **nu3**. RP-HPLC trace van de RNA ICL reactie voor en na NBS oxidatie (links). Denaturerende PAGE gel met drie verschillende targets: complementair RNA met een C of een U tov **nu3** en een gerandomiseerd niet complementair RNA met een C tov **nu3** (rechts).

7.3 Optimalisatie van Methodologie

Los van de oorspronkelijke doelstellingen in een DNA-proteïne context, die het hoofdthema van dit werk uitmaken, werden een aantal resultaten behaald in de optimalisatie van vaste fase methodologie.

- De eerder ontwikkelde NF31 kleurtest bleek residuele OH groepen in nieuwe vaste fase resins te kunnen detecteren.

- Het potentieel van NMR voor de concentratiebepaling van biomacro-moleculen werd getest samen met de NMRSTR groep.

- In samenwerking met de PCR groep, gespecialiseerd in thiolacton chemie, werd een nieuwe en veelzijdige vaste fase koppelingsmethode voor de synthese van sequentie gecontroleerde oligomeren ontwikkeld (Fig. 7.13).
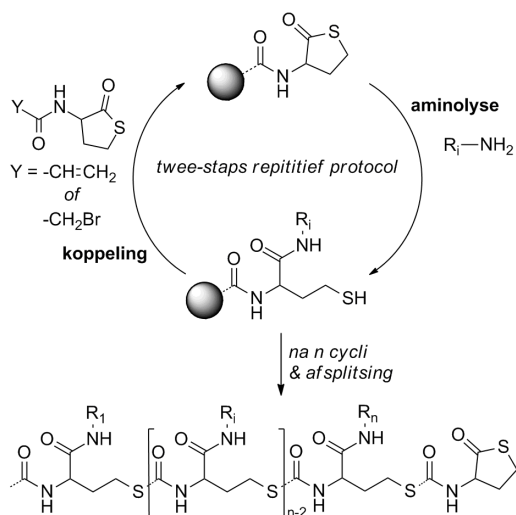
Figure 7.13: Twee-staps repititief protocol voor de synthese van gefunction-aliseerde oligomeren op vaste fase: aminolyse van het thiolacton op de vaste fase, gevolgd door de koppeling van een thiolacton bouwsteen dat selectief reageert met het vrijgestelde thiol.

Conclusie

Het oorspronkelijke doel van dit doctoraatswerk, het ontwerpen van een cross-linkend transcriptiefactor model met mogelijke therapeutische toepassingen, werd niet bereikt. Het bleek immers bijzonder moeilijk om naar een dubbelstrengige helix te cross-linken. De synthese van het dipodale peptidosteroid model werd echter geoptimaliseerd en DNA binding werd voor de eerste maal aangetoond. Verder heeft het werk beschreven in dit manuscript bijgedragen tot een beter begrip en inzicht in de furan oxidatie methode en de toepasbaarheid ervan drastisch uitgebreid. Behalve de incorporatie in ODNs voor de vorming van DNA interstrand cross-links, werd de mogelijkheid om RNA interstrand cross-links te vormen aangetoond in ON duplexen. Furan gemodificeerde duplexen kunnen, met de keuze van een geschikt gemodificeerd nucleoside, ook gebruikt worden voor de vorming van DPCs. Furan werd ook geïncorporeerd in peptiden. DPC vorming kon niet worden aangetoond in dit geval, maar er zijn hoge verwachtingen voor cross-linking in niet-duplex complexen. Tenslotte werd aangetoond dat furan modificaties ook selectief geïncorporeerd kunnen worden in proteïnes, wat veel nieuwe mogelijkheden creëert.
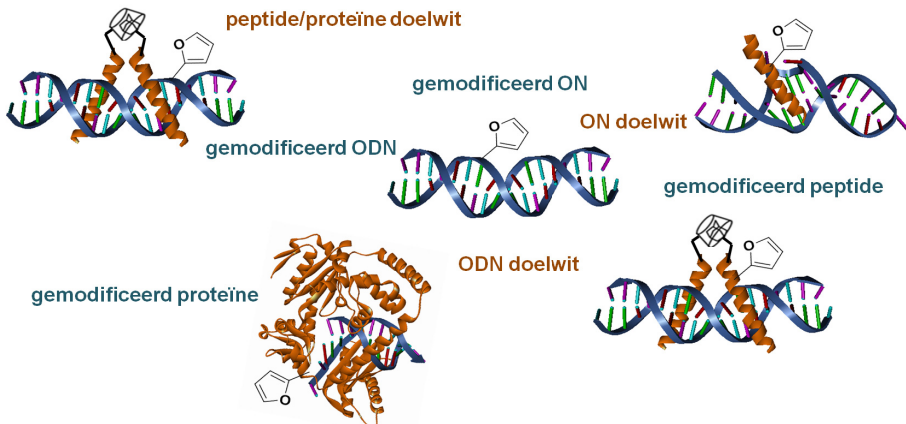
Figure 7.14: Samenvatting van de uitgebreide toepasbaarheid van de furan cross-linking technologie toont de verschillende mogelijkheden voor modificaties (blauwe tekst) en doelwitten (oranje tekst).

Perspectieven

De volgende stappen die nu gezet dienen te worden, moeten toelaten de methode te gebruiken in relevante toepassingen en een biologische context. Daarvoor is het belangrijk om concentratie gerelateerde problemen te optimaliseren en om een biologisch compatibele oxidatie methode te hanteren. De productie van singlet oxygen door sensitizers met rood licht te bestralen werd ontwikkeld voor DNA ICL en is ook compatibel met RNA ICL. Daarenboven zijn testen uitgevoerd voor DPC vorming en was dit ook de methode die gebruikt werd voor de RNA-proteïne cross-linking. Toch kan ook singlet zuurstof nog heel wat schade veroorzaken. Meer selectieve alternatieven die theoretisch haalbaar zouden moeten zijn en waarvoor initiële testen zijn uitgevoerd, zijn oxidatie door cytochroom P450, dat instaat voor de natuurlijke oxidatie van furan in de lever [345–347] of electrochemische oxidatie.[348–350] Vooral deze laatste methode heeft veel potentieel. Aangezien furan een lagere oxidatiepotentiaal heeft dan de DNA/RNA basen,[351–353] zal bij een correcte potentiaal enkel stroom vloeien zolang er ongeoxideerd furan aanwezig is. Het gebruik van elektroden maakt tevens lokale toepasingen mogelijk.[354] Het gebruik van biologische triggers, zoals oxidatieve stress (bv. in kanker cellen) is een andere interessante mogelijkheid.

Part II

Experimental Part

Chapter 8

General Material and Methods

Common solvents and chemical reagents were purchased from Sigma-Aldrich or Acros Organics. Deuterated solvents were purchased from Euriso-top. Suppliers for more specialized chemicals, are specified when used. The water used in experiments was always milliQ grade (Sartorius Arium 611 DI).

8.1   Small molecules

NMR   NMR ^{1}H-, ^{13}C- and ^{31}P-NMR (Attached Proton Test, APT) spectra were recorded on a Bruker Avance 300 at 300 MHz, a Bruker Avance 500 at 500 MHz or a Bruker Avance II 700 at 700 MHz. Chemical shifts are presented in parts per million (ppm) relative to CDCl$_3$ (7.26 ppm in ^{1}H- and 77.2 ppm in ^{13}C-NMR respectively), CD$_3$CN (1.94 ppm in ^{1}H- and 118.26 / 1.32 ppm in ^{13}C-NMR respectively) DMSO-d6 (2.50 ppm in ^{1}H- and 39.51 ppm in ^{13}C-NMR respectively) or D$_2$O (4.79 ppm in ^{1}H-NMR) as internal standard. Coupling constants (J) in ^{1}H-NMR are given in Hz. The resonance multiplicities are described as s (singlet), d (doublet), dd (doublet of doublets), t (triplet) or m (multiplet).

ESI-MS   ESI-MS spectra were recorded with an electrospray ionization source G1946C ES-MSD-single quadrupole type mass detector in positive or negative mode for small molecules using 5 mM NH$_4$OAc in H$_2$O/MeCN (1/1) as eluens.

LC-MS   Sometimes LC-MS was recorded rather than just MS. LC-MS was performed on an Agilent 1100 Series with diode array detection, using a Phenomenex Luna 100 Å, C18(2) column (5 μm, 4.6 x 250 mm) or Kinetex C18 column (5 μm, 4.6 x 250 mm) hyphenated to an electrospray ionizing, single quadrupole MS detector type VL. Linear gradient elution was applied from 0 to 100% buffer B in 15 minutes or 6 min respectively (buffer A: 5 mM NH$_4$OAc in H$_2$O; buffer B: MeCN) with a flow of 1.0 ml/min or 1.5 ml/min respectively at 35°C.

8.2 Peptides

Chemicals and reagents typical for peptide synthesis were purchased from Novabiochem or Iris Biotech GmbH. DMF and NMP peptide grade were from Biosolve.

Because SPPS of peptides is repetitive, the following section will describe the different steps performed, which are repeated throughout the project. The experimental description will refer to these standard operating procedures.

8.2.1 Synthesis

Manual synthesis on small scale was carried out in polypropylene reactors (2, 5 or 10 ml MultiSyntech GmbH) with a Teflon frit to allow easy filtration with a vacuum pump (KNF lab, Laboport). The vessel was closed with septa. Reactions on bigger scale were performed in glass reactors with a sintered glass funnel with tap again for easy filtration and washing. The reactions at room temperature were performed on a shaker (Vibromatic, Selecta). Reactions with microwave assistance were performed in semi-automated fashion with the Biotage Initiator$^+$ SP WaveTM. Automated synthesis at room temperature was performed by a robot equipped with vortexing unit Syro II (MultiSynTech GmbH) or with N_2 bubbling PSSM-8 (Shimadzu).

In those cases the peptide is attached to the resin through a photocleavable linker, all reaction steps are performed shielded from direct sunlight.

Swelling Before every reaction the resin has to be swollen. Dry resin was allowed to swell for minimum 15 min by suspending it in an adequate volume (10 ml/g) of DMF or DCM with intermediate shaking. The solvent was removed by filtration. If the resin was still swollen from a previous reaction, this step can be omitted.

Washing   After each reaction, excess reagents were filtered off using vacuum. The resin was then thoroughly washed to avoid contamination. In manual SPPS the resin was washed with DMF (4x), MeCN or MeOH (4x) and DCM (4x). About 10 ml/ 1 g resin was added to the resin, which was shaken for 20 s and filtered. In (semi-)automated SPPS, washing steps consisted of 9 cycles in which the vessels were filled with DMF or NMP, shaken for 30-45 s and then emptied under vacuum for 30-45 s.

Coupling   Coupling   reagents   (PyBOP   (benzotriazol-1-yl-oxytripyrrolidinophosphonium   hexafluorophosphate),   HBTU   (O-Benzotriazole-N,N,N,Ntetramethyl-uronium-hexauoro-phosphate),   HATU (1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium   3-oxid hexafluorophosphate), DIC (diisopropylcarbodiimide), COMU ((1-Cyano-2-ethoxy-2-oxoethylidenaminooxy)dimethylamino-morpholino-carbenium   hex-afluorophosphate)) and Fmoc (9-Fuorenylmethoxycarbonyl) amino acids were purchased from Iris Biotech GmbH or NovaBiochem: Fmoc-Gly-OH, Fmoc-Thr(tBu)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Glu(tBu)-OH, Fmoc-Asp(tBu)-OH, Fmoc-Leu-OH, Fmoc-Arg (Pbf)-OH, Fmoc-Pro-OH, Fmoc-Lys(Boc)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Ala-OH, Fmoc-Phe-OH.

In manual SPPS the protected amino acid and DIPEA were added to the resin which was suspended in half of the total volume of DMF. After 15 min the coupling reagent was added together with the other half of the DMF. The excess of reagents typically is 4 equiv amino acid, 8 equiv DIPEA and 4 equiv coupling reagent, but depends on the synthesis and will be mentioned for each case. The mixture was flushed with argon and shaken (400 U/min) at room temperature. This procedure was repeated twice for about 2 h. In automated SPPS, 10 equiv protected amino acid, 20 equiv DIPEA and 10 equiv coupling reagent HBTU were all added at the same time from stock solutions of respectively 0.5 M, 1 M and 0.5 M in NMP (As HBTU is not soluble at high concentrations in NMP, DMF was added till a concentration of 0.1%.) The mixture was allowed to react for 1 or 3 h in open vessel, during which it was shaken every 15 min for 20 s. With microwave assistance, the coupling reaction was performed with 10 equiv

amino acid, 20 equiv DIPEA and 10 equiv of the coupling reagent in a minimal amount of DMF at 75° for 5 min.

Fmoc deprotection In manual SPPS, a 20% piperidine/DMF solution was added to the resin (10 ml/g) and the suspension was shaken for 2 min after which it was filtered and washed with DMF (4x). This step was repeated for 5 min and once more for 15 min. The same procedure was followed in automated SPPS, using 40% piperidine/DMF solution. During the reaction the suspension was shaken every minute for 20 s. In between the cycles the resin was drained for 45 s and washed (6 cycles). Also Fmoc deprotection can be assisted by microwave irradiation, using 30% piperidine/DMF for 1 min at rt followed by 3 min at 65°C.

Capping For free amine functions, three different capping procedures were used throughout the work.

Acetic anhydride Acetic anhydride (6 equiv) and DIPEA (6 equiv) were added in DMF and the reaction mixture was shaken for 30 min.

1-Acetylimidazole (AcIm) 1-acetylimidazole (AcIm, 10 equiv) was added to the resin in DMF. The reaction was performed overnight, shaken at 400 U/min.

4-Acetamido benzoic acid (Aba) This capping is very similar to the coupling of an amino acid. The reaction was performed overnight and with a higher number of equivalents of Aba (10 equiv), PyBOP (10 equiv) and DIPEA (20 equiv).

Cleavage from resin

With rink linker Ice cold cleavage cocktail (TFA:water:phenol: thioanisole:EDT (40:4:2.8:2:1) (1 ml/10 mg) was added dropwise and while stirring to the sample cooled on ice. After 15 min, the icebath was removed and the sample allowed to heat up to room temperature and left for 3 h. The sample was then filtered over a sintered glass filter into ice cold diethylether, resulting in precipitation of the peptide. After centrifugation, the ether was decanted and the residue dissolved in mQ water, which was extracted twice with ether and then filtrated over a membrane filter (optionally also filtered over a sephadex column) and purified by HPLC before lyophilisation.

With photocleavable linker The UV-cleavage from the resin has to be performed before acid side chain deprotection as the photocleavable linker is not stable in strong acidic conditions. For fast on-line monitoring, it is also possible to analyze the side chain protected peptide by only performing the photocleavage.

The resin was suspended in MeCN or EtOH, a spatula tip in a small glass vial (small scale) or a monolayer spread in a petridish (big scale). This was illuminated with UV-light from the top 1 mW/cm^2 at 365 nm (Bioblock Scientific) for 3 h, or from both top and bottom 15 mW/cm^2 at 365 nm for 1 h respectively. After evaporation of the sample under reduced pressure, side chain deprotection is performed by acid treatment as described for cleavage with the rink linker.

Colour tests

TNBS Few resin-beads were transferred into a small glass tube. 3 drops of 1% TNBS in DMF and 3 drops of 10% DIPEA in DMF were added. The colour of the beads was observed immediately, after 5 min and after one hour. Red beads indicate the presence of free amine functions while colourless beads indicate their absence.

NF-31   Few resin-beads were transferred into a small glass tube. 100 μl of a 2 mM solution of NF-31 (homemade) in MeCN was added. The tube was heated at 70°C for 10 min. Afterwards the excess dye was washed away with DMF, MeOH and DCM. After washing, red beads indicate the presence of free amine functions, colourless beads the absence.

If next to the 100 μl of a 2 mM solution of NF-31 in MeCN, also 100 μl of a 0.2 M DMAP solution in MeCN is added, the presence of free hydroxyl groups can be detected after 10 min at rt.

Kaiser   Few resin-beads were transferred into a small glass tube. 1 or 2 drops of ninhydrin (500 mg in 10 ml EtOH), phenol (80 mg in 20 ml EtOH) and KCN (2 ml 0.001 M solution deluted to 100 ml with pyridine) were added. The tube was then heated in a boiling water bath for 2 min. The generation of blue colour indicates the presence of free amine functions, no colour the absence.

UV loading determination   Between 5 and 7 mg of resin-beads were weighed in a volumetric flask. Depending on the transferred amount of resin, a volume of 20% piperidine/DMF solution was added, to obtain a concentration between 0.1 and 0.15 mM. The deprotection occurs in 30 min, during which the flask was occasionally swirled. From the absorbance of the piperidinebenzovulvene adduct measured on a photospectrometer at 300 nm, the exact concentration could be determined with the formula obtained from a calibration curve: [concentration = (absorbance + 0.047) / 8.0618]. The loading of the resin in mmol/g is than equal to the product of the concentration and volume divided by the weighed mass.

8.2.2   Analysis

Mass Analysis   Mass analysis was performed using ESI-MS, with a quadrupole ion trap LC mass spectrometer, equipped with electrospray ionization (LCQ ion trap ThermoFinnigan MAT) with MeOH/H_2O 1/1 + 0.1% formic acid, in positive mode at 250°C. Further MALDI-TOF-MS spectra were acquired on a Kratos analytical, AXIMA-LNR (Shimadzu) or a Voyager-DE STR Biospectrometry Workstation (Applied Biosystems) with a high-performance nitrogen laser (337 nm) in positive mode using DHB matrix (2,5-dihydroxybenzoic acid

(98.0% pure, 10 mg) + MeCN (500 μl) + H$_2$O (470 μl) + TFAaq. (30 μl, 3%)) for protected peptides and α-HCA matrix (α-cyano-4-hydroxycinnamic acid (99% pure, 10 mg) + MeCN (500 μl) + H$_2$O (400 μl) + TFAaq. (100 μl, 3%)) for deprotected peptides, with internal or external mass standards (Applied Biosystems or AnaSpec).

RP-HPLC Analysis   Analytical HPLC was performed on an Agilent 1100 or 1200 Series with diode array detection. Protected and therefore hydrophobic peptides were analysed using a Phenomenex Jupiter 300 Å, C4 column (5 μm, 4.6 x 250 mm). After deprotection, the peptides were analysed using an Shinwa Ultron VX C18 ODS (5 μm, 4.6 x 150 mm) Phenomenex Luna 100 Å, C18 column (5 μm, 4.6 x 250 mm) or Jupiter 300 Å, C18 column (5 μm, 4.6 x 250 mm). The optimal concentration for the samples is 1 mg/ml, of which 15 μl was injected. Linear gradient elution was applied from 0 to 100% buffer B in 15 minutes (buffer A: 0.1% TFA in H$_2$O; buffer B: MeCN) with a flow of 1.0 ml/min at 35°C. Long protected peptides were eluted with a linear gradient from 75 to 100% buffer B in 15 min. Unless otherwise stated, all chromatograms shown were detected at 214 nm and obtained as described above. Fractions could be collected both manually and automatically. They were lyophilized in a Speedvac concentrator (Thermoelectron Corporation Savant SPD111V).

LC-MS Analysis   LC-MC was performed on an Agilent 1100 Series with diode array detection, using a Phenomenex Luna 100 Å, C18(2) column (5 μm, 4.6 x 250 mm) hyphenated to an electrospray ionizing, single quadrupole MS detector type VL. Linear gradient elution was applied from 0 or 75% to 100% buffer B in 15 minutes (buffer A: 0.1% formic acid in H$_2$O; buffer B: MeCN) with a flow of 1.0 ml/min at 35°C.

Concentration Analysis   Concentrations were determined by UV-Vis measurements with a Varian Cary 3E UV/Vis spectrophotometer.

8.3 Oligonucleotides (ODNs)

Commercial unmodified oligonucleotides were purchased purified over a Gold Cartidge from Eurogentec, IDT DNA Technologies or Gene Design, Inc.

8.3.1 Synthesis

Oligonucleotide were synthesized using standard phosphoramidite chemistry (reagents from Glen research or Proligo), in an automated fashion, DMT-on using an ABI 394 DNA-synthesizer at 1 μmol scale. A standard synthesis protocol was used except for coupling of the modified residues. For the manual introduction of these modified residues, the synthesis column was removed from the DNA-synthesizer and small portions (0.05 ml/min) of the phosphoramidite (0.06 M or 0.1 M if crude, 0.4 ml) and 4,5-dicyanoimidazole (1 M, 0.5 ml) solution in dry MeCN, on molecular sieves for max. 20 min, were then alternately pushed over the reaction column, during a total reaction time of 18 min. The column was flushed with acetonitrile (1 ml), and a mixture of Cap A (0.5 ml) and Cap B (0.5 ml) was pushed over the column, after which it was flushed again with acetonitrile (1 ml). The column was reinstalled on the DNA synthesizer, and automated synthesis was resumed.

Upon completion of the synthesis, the ODN was deprotected and cleaved from the controlled pore glass by incubation at 55°C overnight in concentrated aqueous ammonia (28% NH_4OH). Postsynthetic workup and purification was performed using Sep-pak (C18 cartidges from Waters). The cartridge was activated by rinsing it with MeCN (10 ml) and an aqueous TEAA solution (5 mM, 10 ml). Next the oligonucleotide solution obtained from deprotection was loaded onto the column, the eluent was collected into the vial containing the ODN solution and brought back onto the column (3x). The column was then rinsed with a diluted aqueous NH_4OH solution (2.8%, 15 ml) and H_2O (10 ml), during which truncated sequences, not containing a DMTr-group are eluted from the column. The DMTr-groups of the ODN remaining on the column were then deprotected by rinsing the column with an aqueous TFA solution (1.5%, 10 ml).

The column was again rinsed with H_2O (10 ml), after which the ODN were eluted with MeCN:H_2O (2:8, 5 ml).

The purity of the ODN was checked by RP-HPLC and the identity checked by ESI-MS or Maldi-MS.

^{32}P-labeling of ODNs for radiodetection. The ODN to be labeled (10 pmol, 10 μM, 1 μl), was combined with adenosine 5'-triphosphate,$[\gamma^{32}P]$ (Perkin Elmer) (10 pmol, 1.66 μM, 1 μl), T4 Polymer Kinase (New England Biolabs) (4 units, 0.4 μl), NE buffer for T4 polynucleotide kinase (New England Biolabs) (10x concentrate, 1 μl) and milliQ grade water (1.6 μl) in a low binding centrifugal tube (BioScience, Inc). These tubes should be used throughout experiments with very small amounts of radiolabeled ODNs. The sample was mixed by gentle tapping and incubated at 37°C for 30 min. After reaction, the labeled ODN was purified by phenol extraction (40 μl 1x TE buffer and 50 μl phenol/$CHCl_3$) and size exclusion (with a Bio-Spin 6 column (Bio Rad)).

8.3.2 Analysis

Mass Analysis   ESI-MS spectra were recorded with an electrospray ionization source G1946C ES-MSD-single quadrupole type mass detector in negative mode using H_2O/MeCN (1/1) with 1% TEA as eluens. The mass spectra were deconvoluted, i.e. the molecular weights of the compounds were reconstructed from the spectra using the Agilent LC/MSD ChemStation software (version A.08014). All MALDI-TOF-spectra were measured with an Applied Biosystems Voyager DE STR Biospectrometer, with 3-hydroxypicolinic acid/ammoniumcitrate (9/1) as matrix and dowex beads to desalt the sample.

RP-HPLC Analysis   Oligonucleotides were analyzed through ion pairing chromatography methods. Ion Pair Chromatography (IPC) is used to separate ionic analytes on a reversed phase column. An Ion Pair (IP) reagent, like triethylammonium acetate (TEAA), is added to modulate retention of the ionic analytes, in this case oligonucleotides. The ion-pairing additive in the mobile

phase is adsorbed on the stationary phase, and thus provides for charge-to-charge interactions with negative charges contained on the oligonucleotide backbone. As a result, an efficient charge-based (length-based) oligonucleotide separation is achieved. Gradient elution using an MeCN or MeOH eluent displaces both ion-pairing agent and the oligonucleotides from the sorbent surface.

RP-HPLC analysis was recorded on an Agilent 1100 or 1200 system equipped with a UV diode array detector set at 260 nm, using Phenomenex Clarity (250 x 4.6 mm, 5 μm) or Aeris Widepore (150 x 4.6 mm, 3.6 μm) columns at 60°C with eluent: 0.1 M TEAA (with 5% MeCN) and MeCN as mobile phase (linear gradient: 0-30% MeCN in 15 or 30 min, 30-100% MeCN in 3 min).

Concentration Analysis Concentration of ODNs was determined with a Trinean dropsense multichannel spectrophotometer. Extinction coefficients were calculated by the online IDT tool, replacing a modified nucleotide by a thymidine.

Thermal denaturation Melting temperature (T_m) studies were performed using a Varian Cary 300 Bio equipped with a six-cell thermostated cell holder. The curves were monitored at 260 nm with a heating rate of 0.3°C/min. A 1 μM sample of 2 complementary ODNs was buffered in 10 mM phosphate buffer of pH 7, with 100 mM NaCl. The oligonucleotide concentration was 1 μM for each strand. Melting temperatures were calculated from the first derivative of the heating curves using the Cary 300 Bio software. The measurements were performed in triplicate, upon which an average T_m value was calculated.

8.4   Electrophoretic shift mobility assay (EMSA)

Binding of peptides to ODN duplexes were assessed by EMSA. DNA duplexes
for binding were prepared in 10 mM Tris buffer pH 8, with 100 mM NaCl and
1 mM EDTA by annealing (heating to 95° for 24 min and slowly cool down
over min. 2 h). For fluorescent detection, equimolar amounts of both strands
was used, with radiodetection, twofold excess of the cold complementary strand
was used. The DNA concentration was dependant on the detection method,
0.167 μM in the loading sample was sufficient for fluorescent detection with
Sybr Gold, for radiodetection an amount of DNA giving 4000 counts per minute
(cpm) in the loaded sample is desired. Increasing equivalents of peptide were
added to the DNA solution with a loading buffer to obtain solutions with 6%
sucrose, 100 mM Tris pH 7.6, 20 mM KCl, 10 mM $MgCl_2$, 2.5 mM EDTA. The
binding mixtures were incubated on ice for 1 h, before 5 μl was loaded on the
gel. Polyacrylamide gelelectrophoresis experiments were performed on 10 x 10
cm 8% acrylamide:bisacrylamide (29:1) gels, (portions of 20 ml were prepared
with ammonium persulphate (200 μl, 10%, Merck), sonificated, cooled down
and initiated with tetramethylethylenediamine (20 μl, Merck)) The experiment
ran with 0.2% TBE buffer (Merck) at 150 V and 4°C for 45 min, after a prerun
under the same conditions for 30 min. DNA was visualized by fluorescent
detection: 10 min of Sybr Gold (Life Technologies) staining and illumination in
an Autochemi imaging system (UVP) or radiodetection: after drying the gel at
80°C for 30 min by 10 h irradiating of a phosphor screen (Molecular Dynamics,
Amersham Pharmacia Biotech) in a closed cassette on a Storm 820 phosphor
imager (Molecular Dynamics).

8.5 Cross-linking experiments

8.5.1 DNA interstrand cross-linking

Formation The modified strands were mixed with their complements in equimolar (1 nmol) amounts at 20 μM concentration in 10 mM phosphate buffer (pH 7) and 100 mM NaCl. Duplexes were obtained by annealing (heating to 95°C for 10 min, and then allowing very slow cool down to room temperature). Temperatures during the whole reaction were kept constant in an Eppendorf thermomixer comfort at 25°C, while shaking at 900 rpm. To start the reaction, 1 equiv NBS (1 equiv, 1 nmol, 2 μl) of a freshly prepared stock solution was added. This was repeated every 15 min until complete disappearance of the modified oligonucleotide (typically 4 equiv over 1 h). Cross-linking with singlet oxygen was carried out by adding 2.5 μM methylene blue and shining light with a Euromex 100 W cold light source, equipped with a red light filter at maximal irradiation intensity at 25°C.[139] The reactions were monitored by RP-HPLC and denaturing polyacrylamide gel electrophoresis (PAGE).

RP-HPLC Analysis For the RP-HPLC analysis of the cross-link mixtures, the same conditions were used as for the analysis of the ODNs seperately.
Cross-linking yields based on HPLC were calculated by the ratio of the peak area from the spectra after and before reaction, taking into account the extinction coefficients. These were calculated by the nearest neighbor method, replacing the modified nucleoside with a thymidine. The extinction coefficient of the cross-linked duplex is taken to be 0.9 times the sum of the extinction coefficients of its composing sequences.

Denaturing PAGE Analysis A 20% polyacrylamidegel (acrylamide:bisacrylamide 19:1) with 1x TBE buffer (Merck) and 7 M urea was used for all analyses. The gel was prerunned with 1x TBE as running buffer at 225 V, 25°C for 30 min. A portion of the reaction mixture (2 μl) was mixed with formamide (18 μl) to obtain 90% and loaded on the gel (5 μl). Gels were run with 1x TBE as running buffer at 225 V, 25°C for 1.5 h and stained with SybrGold

(1x, Life Technologies) for 10 min. Pictures were taken with an Autochemi imaging system (UVP).

8.5.2   RNA interstrand cross-linking

Experiments and analysis were performed as described for DNA interstrand cross-linking, but care was taken to avoid degradation of the more labile RNA.

- Working surface and equipment were cleaned and made RNA'se free with RNaseZAP (Sigma Aldrich)
- Gloves and a mouth mask were worn when manipulating samples
- Samples were stored in freezer or on ice, time at rt was kept to a minimum
- Annealing was performed by heating the sample for 3 min at 80°C before incubation on ice for 1 h

8.5.3   DNA-peptide cross-linking

Formation   Cross-linking reactions were performed on a 1 nmol (20 μM) DNA scale. Duplex DNA was obtained by heating 1 nmol of the 2 constituting oligonucleotides in phosphate buffer of pH 7 with 0.1 M NaCl and 0.02 M MgCl$_2$ to 95°C for 10 min, and then allowing very slow cool down to room temperature. An amount of peptide should then be added to obtain complete DNA binding and the mixture was incubated on ice for 1 h, to allow DNA-peptide complex formation. Complete furan oxidation was achieved by the addition of portions of NBS (1 equiv, 2 μl) to the complex on ice. Reaction progress was monitored by HPLC, but this typically required 4 additions over the course of 1 h. Finally the reaction was driven to completion and the product stabilized by reduction of the formed imine to an amine with an excess of NaCNBH$_3$.

RP-HPLC Analysis   For the RP-HPLC analysis of the cross-link mixtures, the same conditions were used as for the analysis of the ODNs seperately.
Cross-linking yields based on HPLC were calculated by comparing the peak area from the spectra before and after addition of NBS, taking into account the extinction coefficients and corrected for sampling. For the DNA, the extinction

coefficient was calculated by the nearest neighbor method, replacing the modified nucleoside with a thymidine. For the cross-linked complex, the extinction coefficient of the aromatic Aba cap (remainder of the GCN4 peptide is not aromatic and does not absorb at 260 nm) was additionally taken into account.

Denaturing PAGE Analysis An 8% polyacrylamidegel (acrylamide:bisacrylamide 37.5:1) with 1x TBE buffer (Merck) and 6 M urea was used for all analyses. The gel was prerunned with 1x TBE as running buffer at 225 V, 25°C for 30 min. A portion of the reaction mixture (2 μl) was mixed with formamide (18 μl) to obtain 90% and loaded on the gel (5 μl). Gels were run with 1x TBE as running buffer at 225 V, 25°C for 35 min. Visualisation was achieved by staining with SybrGold (1x, Life Technologies) for 10 min and pictures were taken with an Autochemi imaging system (UVP). Or using radiolabels, after drying the gel at 80°C for 30 min, by 10 h irradiation of a phosphor screen (Molecular Dynamics, Amersham Pharmacia Biotech) in a closed cassette and measurement on a Storm 820 phosphor imager (Molecular Dynamics).

8.6 Visualisation

The free software Discovery Studio from Accelrys was used for visualization. ODN duplexes could be drawn directly in the software. Proteins (and the duplexes they bind to) were taken from the PDB. Furan modifications could be added with the tools from the software. Geometry optimizations were performed locally around the furfurylamine modification and over the complete sturcture with a fast Dreiding-like forcefield.

Chapter 9

Specific Material and Methods

9.1 Supporting Material for the Madder Mimic

9.1.1 Additional and specific methods

Peptide synthesis was performed according to the described methods on amino TentaGel-resin (90 m, loading 0.28 mmol/g, Iris Biotech Gmbh).

For the manual couplings following specifications are additionally given: double 2.5 h PyBOP, DIPEA couplings in a minimal amount of DMF under argon atmosphere with 3 equiv of linker and PyBOP and 6 equiv of DIPEA for the photosensitive linker (homemade), with 1.2 and 2.4 equiv for the scaffold (homemade), with 4 and 8 equiv for Gaba and the 1st Gly. Capping at the end of the peptide was performed overnight with 10 equiv of Aba, PyBOP and 20 equiv of DIPEA. Intermediate capping after incomplete reactions was performed with 6 equiv of acetic anhydride and DIPEA for 30 min, or 10 equiv 1-acetylimidazole overnight.

The microwave assisted synthesis (both coupling and Fmoc deprotection) was performed as described in the general methods. For the test with different coupling reagents for the attachment of Gaba at the C3 position, 4 equiv of Gaba and the coupling reagent and 8 equiv of DIPEA were used in a double 10 min coupling reaction at 75°C.

Boc deprotection 20% TFA in DCM (1 ml/100 mg resin) was added to the resin. The suspension was shaken for 2 min, after which it was drained and washed 4x with DCM. This step was repeated but for 5 min and 15 min. The resin is subsequently washed 4x with DCM, DMF, 10% piperidine in DMF, DMF, MeOH and DCM.

Alloc deprotection $Pd(PPh_3)_4$ (1155.58 g/mol, 0.3 eq, homemade)), Bu_3SnH (291.06 g/mol, 1.092 g/ml, 10 eq) and morpholine (87.12 g/mol, 0.9955 g/ml, 90 eq) were added in dry DCM (1 ml/100 mg resin). This reaction mixture was flushed with argon, sealed and shaken (440 U/min) for 3 h. The reaction was repeated twice.

Cleavage from the resin

Previous, destructive final cleavage and deprotection was performed as described in the general procedure with UV irradiation for cleavage from the resin. After lyophilisation of the obtained sample, reagent B (88% TFA, 5% H2O, 2% TIS, 5% phenol) was added to the side-chain protected peptide. The suspension was shaken at 60°C for 3 h. The reaction was then quenched by adding 300 μL of mQ H_2O and the mixture was lyophilised. Alternatively the TFA could be evaporated under argon. The side-chain deprotected peptide was obtained as a white powder. This procedure was repeated twice. The peptide was purified by precipitation in ice cold Et_2O. After centrifugation for 10 min at 2000 rpm and 2°C, the supernatans was removed. The purification was repeated 3 times. The last Et_2O was removed by evaporation under argon.

Optimized cleavage is described in the general procedures and includes spreading the solid support as a monolayer in a pertidish, which is irradiated from top and bottom by a stronger UV lamp (15 mW/cm^2), for 1 h, the use of a different cleavage cocktail and cooling rather than heating of the mixture.

9.1.2   Spectral Data

monodipodal peptide synthesis (3.1.2)

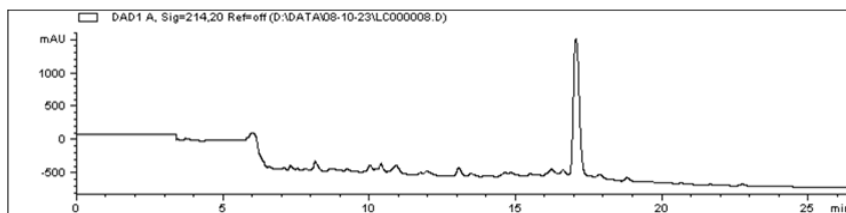Formula: $C_{295}H_{432}N_{52}O_{61}S_6$
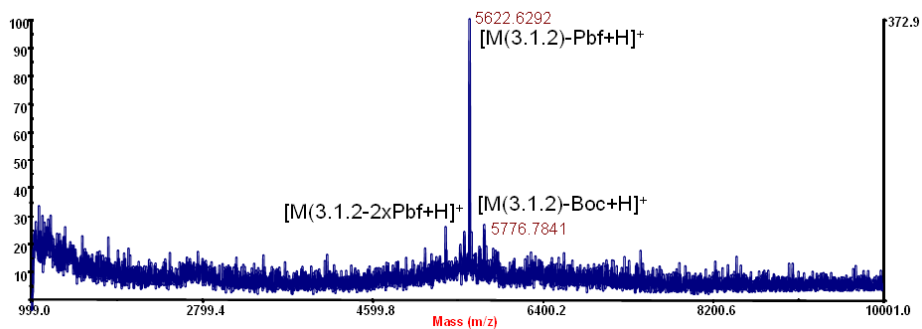Calculated Mass: 5871.1 (MW: 5875.3)
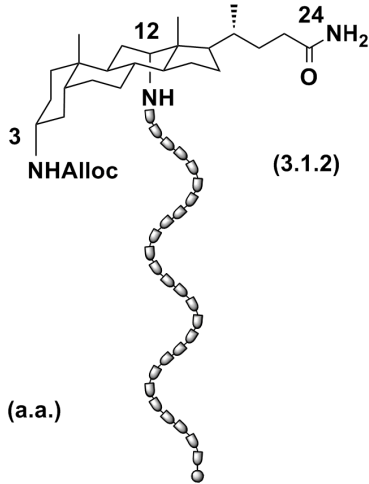
Aba- D(tBu)PLAALK(Boc)R(Pbf)AR(Pbf)N(Trt)T(tBu)E(tBu)
AAR(Pbf)R(Pbf)S(tBu)R(Pbf)AR(Pbf)K(Boc)LQ(Trt) **-GGGaba**

cap   *protected* GCN4 basic region peptide   linker

Figure 9.1: RP-HPLC (75-100% MeCN in 15 min on C4, 300 Å): r.t. $= 17$ min

Figure 9.2: MALDI of the fraction with r.t. $= 16.8$-17.5 min: 5369 [M(3.1.2)-2xPbf+H]$^+$, 5623 [M(3.1.2)-Pbf+H]$^+$, 5777 [M(3.1.2)Boc+H]$^+$
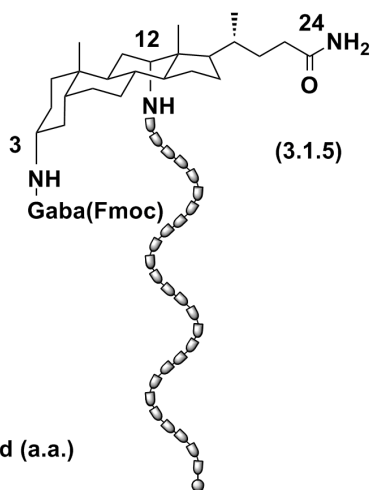
Introduction of the 1st residue of the 2nd peptide (3.1.5)

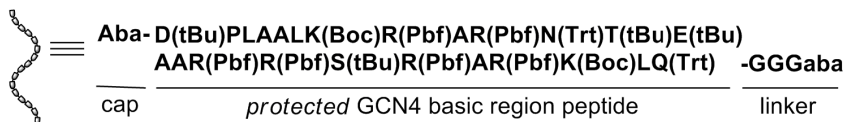Formula: $C_{310}H_{445}N_{53}O_{62}S_6$
Calculated Mass: 6094.2 (MW: 6098.6)

   (starting material (**3.1.4**):
Formula: $C_{291}H_{428}N_{52}O_{59}S_6$
Calculated Mass: 5787.0 (MW: 5791.2))

🔵 ≡	**amino acid (a.a.)**
🔵 ≡	**Aba-cap**

Aba- **D(tBu)PLAALK(Boc)R(Pbf)AR(Pbf)N(Trt)T(tBu)E(tBu)**
     AAR(Pbf)R(Pbf)S(tBu)R(Pbf)AR(Pbf)K(Boc)LQ(Trt)   **-GGGaba**

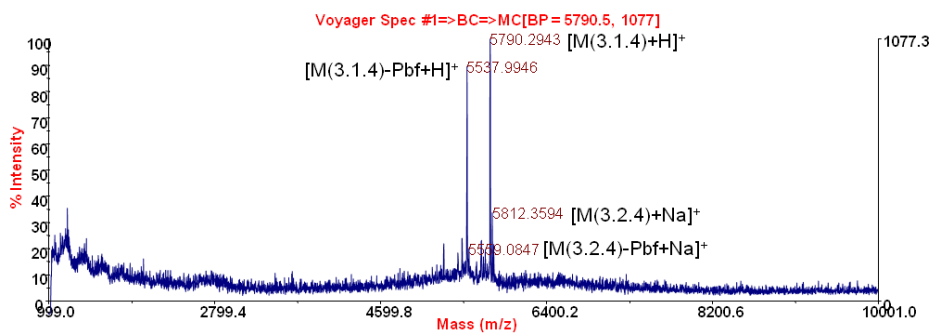<u>cap</u>   <u>*protected* GCN4 basic region peptide</u>   <u>linker</u>

After coupling with PyBOP: .
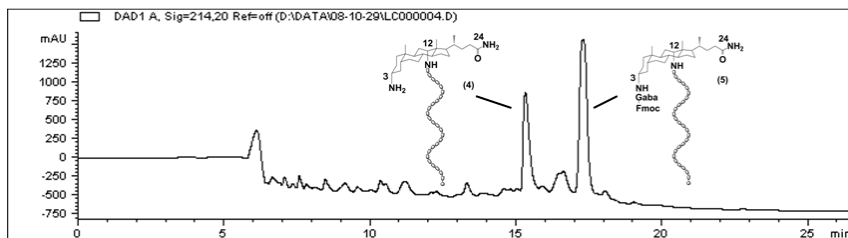
Figure 9.3: RP-HPLC (75-100% MeCN in 15 min on C4, 300 Å): r.t. = 15.5 min (**3.1.4**) and 17.4 min (**3.1.5**)

Figure 9.4: MALDI of the fraction with r.t. = 15.3-15.9 min: 5538 $[M(3.1.4)\text{-}Pbf\text{+}H]^+$, 5559 $[M(3.1.4)\text{-}Pbf\text{+}Na]^+$, 5790 $[M(3.1.4)\text{+}H]^+$, 5812 $[M(3.1.4)\text{+}Na]^+$

Figure 9.5: MALDI of the fraction with r.t. = 17.1-17.9 min: 5593 $[M(3.1.5)\text{-}2xPbf\text{+}H]^+$, 5845 $[M(3.1.5)\text{-}Pbf\text{+}H]^+$, 6120 $[M3.1.5)\text{+}Na]^+$
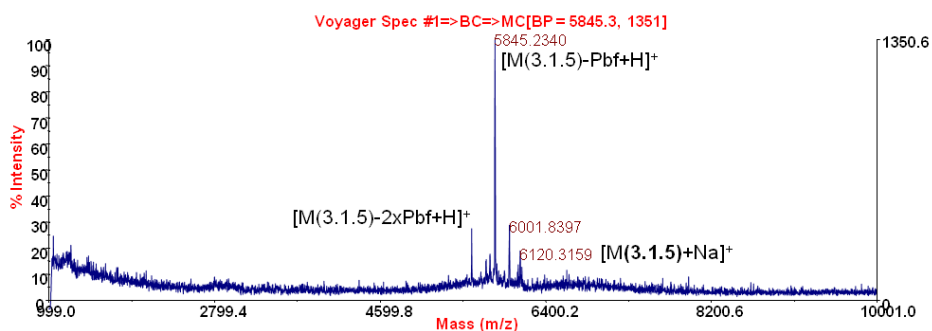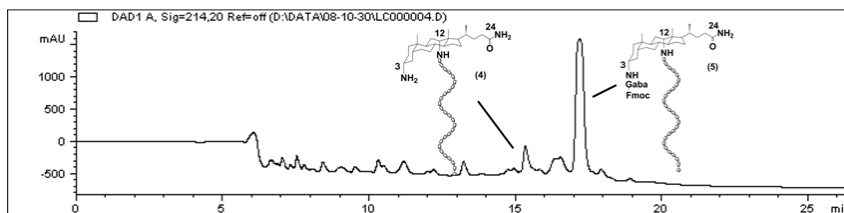
After double coupling with HATU: .

Figure 9.6: RP-HPLC (75-100% MeCN in 15 min on C4, 300 Å): r.t. = 15.5 min
(**3.1.4**) and 17.4 min (**3.1.5**)
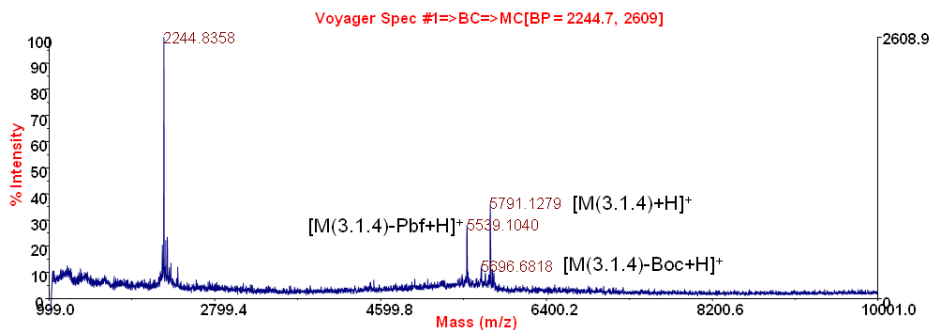
Figure 9.7: MALDI of the fraction with r.t. = 15.1-15.8 min: 5539 [M(3.1.4)-Pbf+H]$^+$,
5560 [M(3.1.4)-Pbf+Na]$^+$, 5697[M(3.1.4)-Boc+H]$^+$, 5791 [M(3.1.4)+H]$^+$
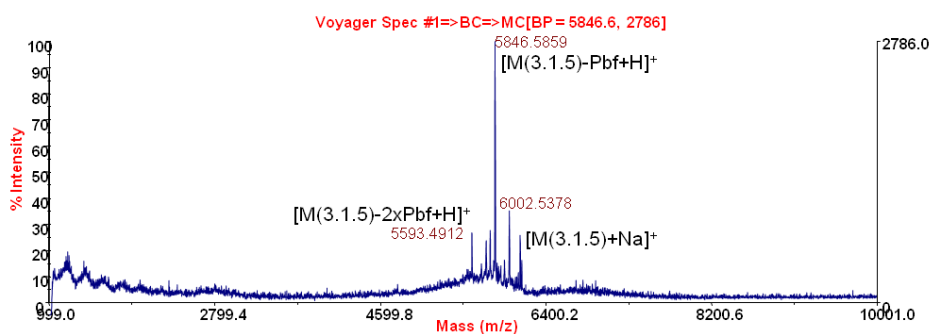
Figure 9.8: MALDI of the fraction with r.t. = 17-17.7 min: 5594 [M(3.1.5)-2xPbf+H]$^+$,
5750 [M(3.1.5)-Pbf-Boc+H]$^+$, 5847 [M(3.1.5)-Pbf+H]$^+$, 6119 [M(3.1.5)+Na]$^+$

After capping with AcIm: .

Formula Capped: $C_{293}H_{430}N_{52}O_{60}S_6$

Calculated Mass Capped: 5829.1 (MW: 5833.3)
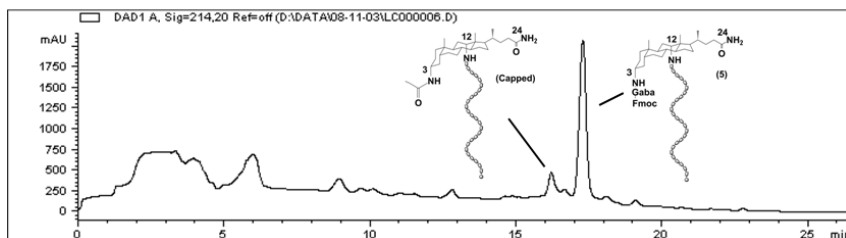
Figure 9.9: RP-HPLC (75-100% MeCN in 15 min on C4, 300 Å): r.t. = 16.2 min (**Capped**) and 17.4 min (**3.1.5**)
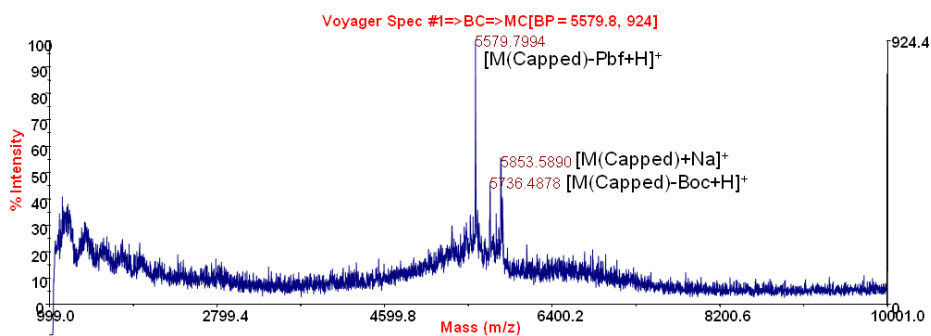
Figure 9.10: MALDI of the fraction with r.t. = 15.9-16.6 min: 5580 [M(Capped)-Pbf+H]$^+$, 5736 [M(Capped)-Boc+H]$^+$, 5854 [M(Capped)+Na]$^+$
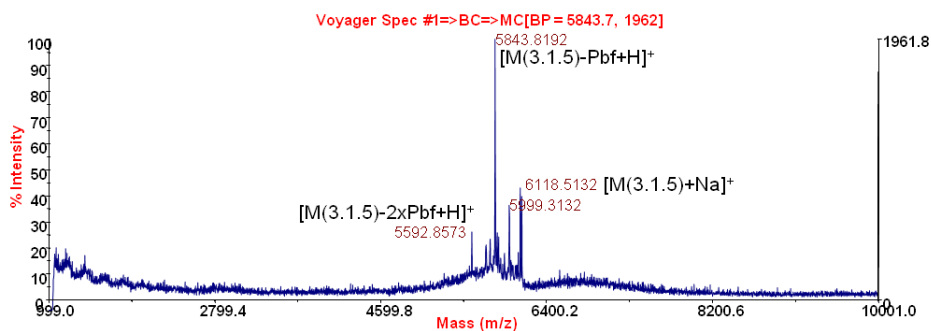
Figure 9.11: MALDI of the fraction with r.t. = 17.0-17.8 min: 5593 [M(3.1.5)-2xPbf+H]$^+$, 5844 [M(3.1.5)-Pbf+H]$^+$, 6119 [M(3.1.5)+Na]$^+$

After direct double coupling with HATU   .

Figure 9.12: RP-HPLC (75-100% MeCN in 15 min on C4, 300 Å): r.t. = 13.9 min (**3.1.5**)

Figure 9.13: MALDI of the fraction with r.t. = 13.6-14.5 min: 5862 [M(3.1.5)-Pbf+Na]$^+$, 6115 [M(3.1.5)+Na]$^+$

Introduction of the 2nd residue of the 2nd peptide (3.1.5)

Formula: $C_{312}H_{448}N_{54}O_{63}S_6$
Calculated Mass: 6151.2 (MW: 6155.6)
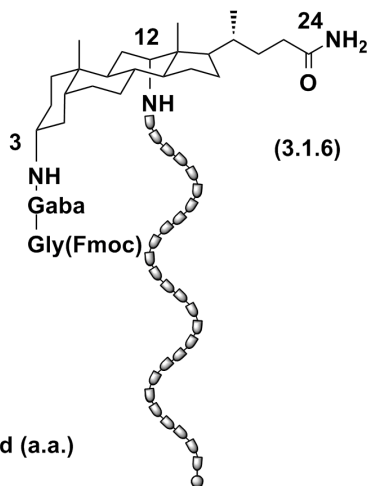
(3.1.6)

○ ≡ **amino acid (a.a.)**

● ≡ **Aba-cap**

≡ **Aba**- **D(tBu)PLAALK(Boc)R(Pbf)AR(Pbf)N(Trt)T(tBu)E(tBu)**
AAR(Pbf)R(Pbf)S(tBu)R(Pbf)AR(Pbf)K(Boc)LQ(Trt) **-GGGaba**

cap   *protected* GCN4 basic region peptide   linker
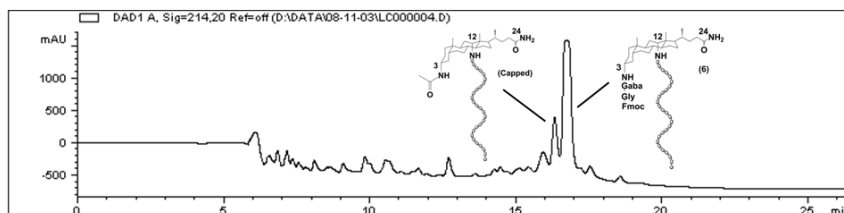
Figure 9.14: RP-HPLC (75-100% MeCN in 15 min on C4, 300 Å): r.t. = 16.2 min (**Capped**) and 16.8 min (**3.1.6**)
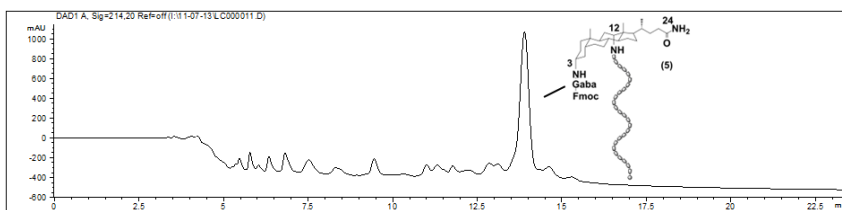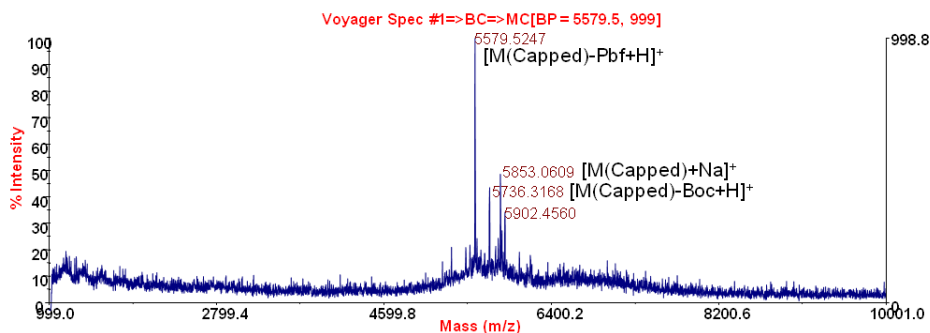

Figure 9.15: MALDI of the fraction with r.t. = 16.2-16.6 min: 5580 [M(Capped)-Pbf+H]$^+$, 5736 [M(Capped)-Boc+H]$^+$, 5853 [M(Capped)+Na]$^+$
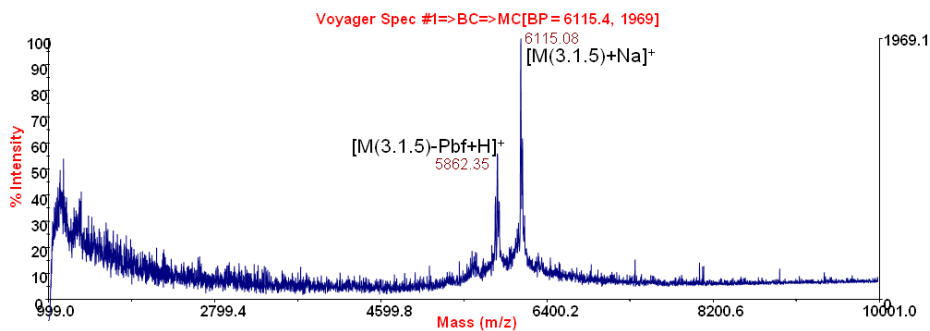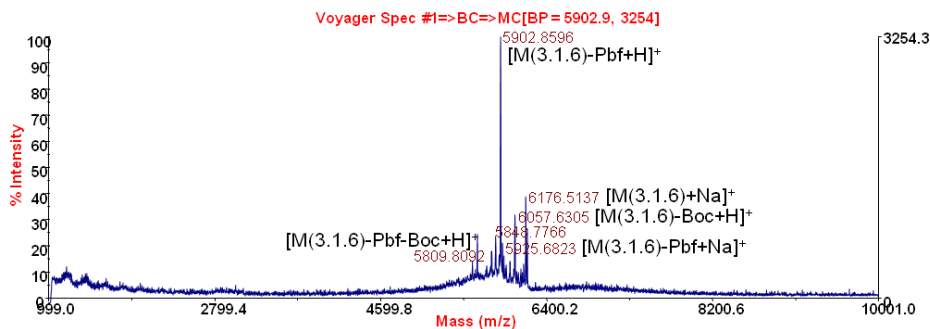

Figure 9.16: MALDI of the fraction with r.t. = 16.6-17.2 min: 5650 [M(3.1.6)-2xPbf+H]$^+$, 5809 [M(3.1.6)-Pbf-Boc+H]$^+$, 5903 [M(3.1.6)-Pbf+H]$^+$, 5926 [M(3.1.6)-Pbf+Na]$^+$, 6058 [M(3.1.6)-Boc+H]$^+$, 6177 [M(3.1.6)+Na]$^+$, 6194 [M(3.1.6)+K]$^+$

After direct double coupling with HATU .
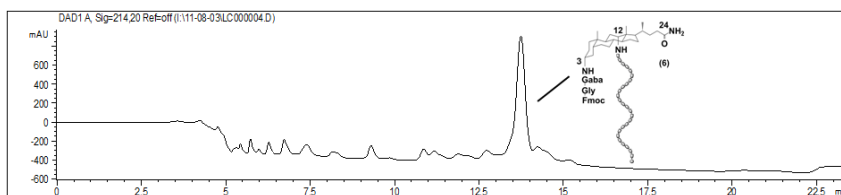
Dipodal peptide synthesis (3.1.3)

Formula: $C_{274}H_{469}N_{101}O_{73}$
Calculated Mass: 6342.6 (MW: 6346.3)

amino acid (a.a.)

Aba-cap

Aba-DPLAALKRARNTEAARRSRARKLQ-GGGaba

cap   GCN4 basic region peptide   linker
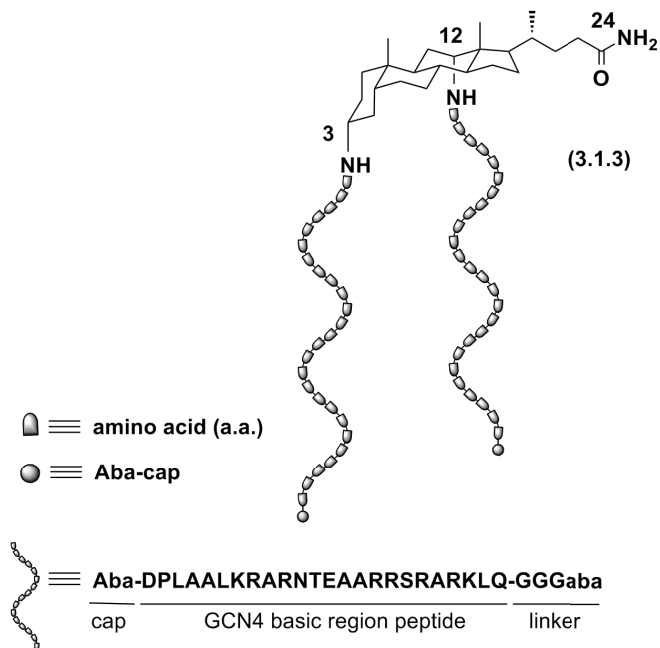
Initial synthesis conditions: .
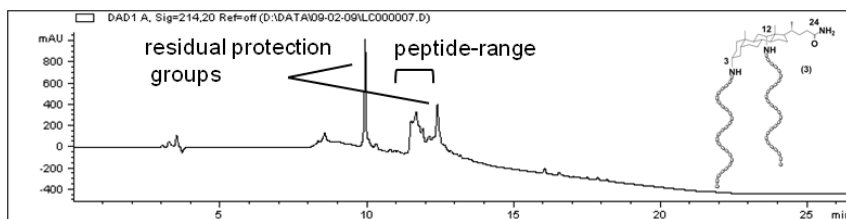
Figure 9.17: RP-HPLC (0-100% MeCN in 15 min on C18, 300 Å): mixture

MALDI of the sample and fractions does not give good MALDI spectra.

LC-MS (0-100% MeCN in 30 min on C18, 100 Å): sharp peak at r.t = 10.3 min (Fig. 9.17) is not peptidic in nature. Significant peptides with highest molecular weight are 5792 and 5890 Dalton, present at r.t. = 10.6 min. The following figures are the mass spectrum at r.t. = 10.6 min (left) and the result of the deconvolution (right) in which components A and B correspond to multiply charged ions of the peptides with a mass of 5792 and 5890 Dalton respectively.
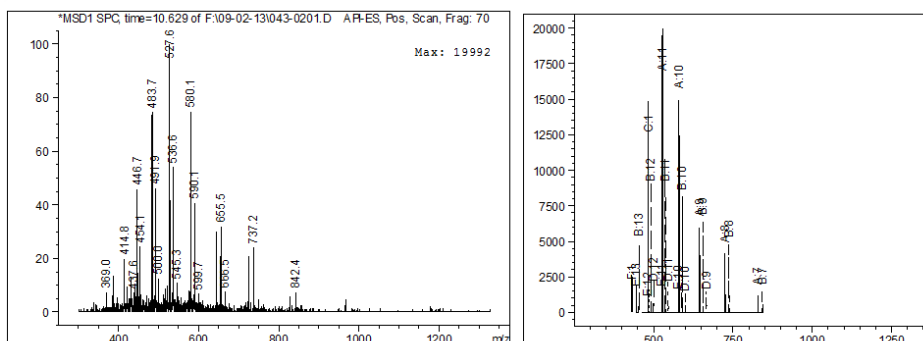
Figure 9.18: ESI-MS spectrum at r.t. = 10.6 min (left) and deconvolution to components A (5792 Dalton) and B (5890 Dalton)

The results thus revealed how the synthesis was not completely successful. The obtained chromatogram shows the impurity of the sample. The sharp peak with a retention time of 10 minutes corresponds to a protecting group. The rest of the chromatogram corresponds to peptides of different weight according to LC-MS. Between 10 and 15 minutes the most important contributing peptides are

found with a molecular weight decreasing with higher retention time. No trace of
3.1.3 could be found. If it was present, then it was only in a trace level amount.
The peptides with the highest molecular weight that were present in a significant
amount A and B, had a weight of 550 and 452 Dalton lower than **3.1.3**. Most likely
this corresponds for B to an arginine, an alanine and two leucines and for A to
the same residues and an additional proline. There are only two leucines and one
proline in the sequence, so all of them are missing. Proline has been long known
to be a difficult amino acid to couple and problems associated with leucine have
also been reported. Alanine and arginine both have six copies in the sequence,
so it is not clear which of these is missing. Possibly they are missing at different
positions, which further increases the complexity of the obtained mixture.

Parallel attachment of the peptides: .
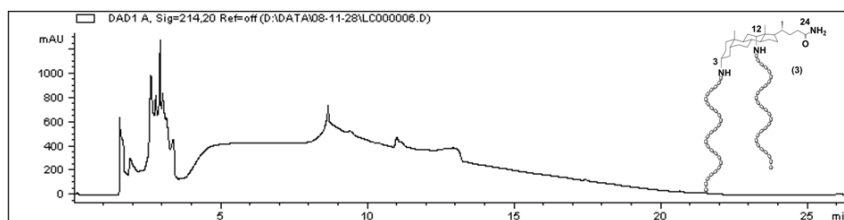
Figure 9.19: RP-HPLC (0-100% MeCN in 15 min on C18, 300 Å): mixture and low
concentration

The chromatogram (Fig. 9.19) illustrates how hardly any product had been
formed, as the purity and yield were very low. The failure of this approach is not
totally unexpected. Due to sterical hindrance, the two positions on the scaffold
have a different reactivity. The amine at position C12 is more sterically hindered
and will therefore react slower than the amine at position C3. Prolonging the
branch mainly at position C3 enforces the sterical hindrance and reduces the
reactivity even more, giving rise to extensive deletions in the branch at position
C12.

Synthesis on NovaPEG (all PEG solid support): .
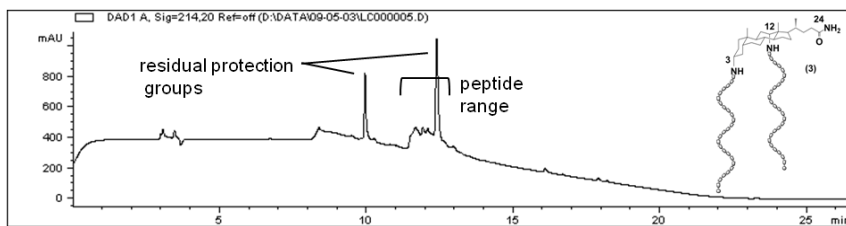
Figure 9.20: RP-HPLC (0-100% MeCN in 15 min on C18, 300 Å): mixture

LC-MS (0-100% in 30 min on C18,100): sharp peaks at r.t.= 11.5 min and 12.4 min (Fig. 9.20) are not peptidic in nature. The peptide with highest molecular weight is 5792, present at r.t. = 9.5 min. The following figures are the mass spectrum at r.t. = 9.5 min (left) and the result of the deconvolution (right) in which component A corresponds to multiply charged ions of the peptide with a mass of 5729 Dalton.
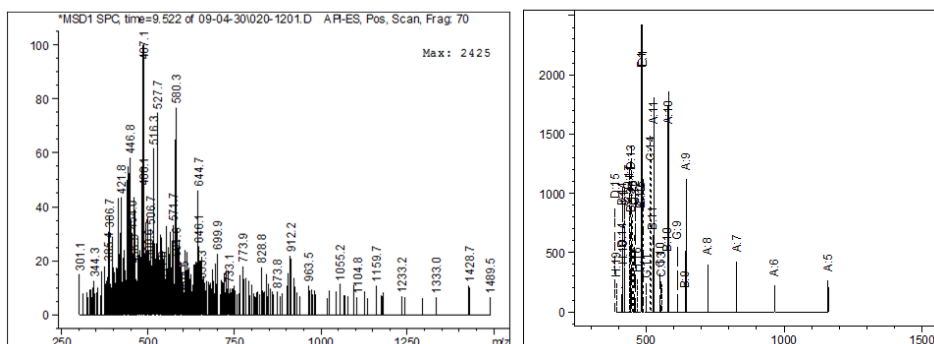
**Figure 9.21:** ESI-MS spectrum at r.t. = 9.5 min (left) and deconvolution to component A (5792 Dalton)

The first sharp peak with a retention time of 10 min in the chromatogram was already present in the previous chromatograms and corresponds to a protecting group. According to ESI-MS, the other sharp peak with a retention time of 12 min does not correspond to a peptide either. For the comparison of the synthesis on TentaGEL and NovaPEG, it can be concluded that the results are similar and not good. Once more the same compound A, with a mass 550 Dalton lower than the desired was the compound with the highest molecular weight present.

Synthesis with optimized cleavage method: .
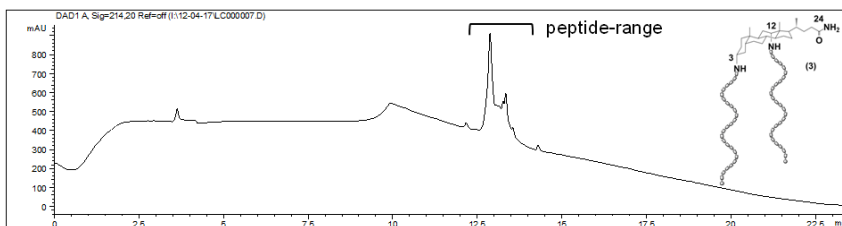
Figure 9.22: RP-HPLC (0-100% MeCN in 15 min on C18, 300 Å)

LC-MS (0-100% in 30 min on C18,100): Everything is peptidic in nature and the target mass is present as the main contributer in the major peak. The following figures are the mass spectrum at r.t. = 8.8 min (left) and the result of the deconvolution (right) in which component A corresponds to multiply charged ions of the peptide with a mass of 6345.6 Dalton or the target mass.
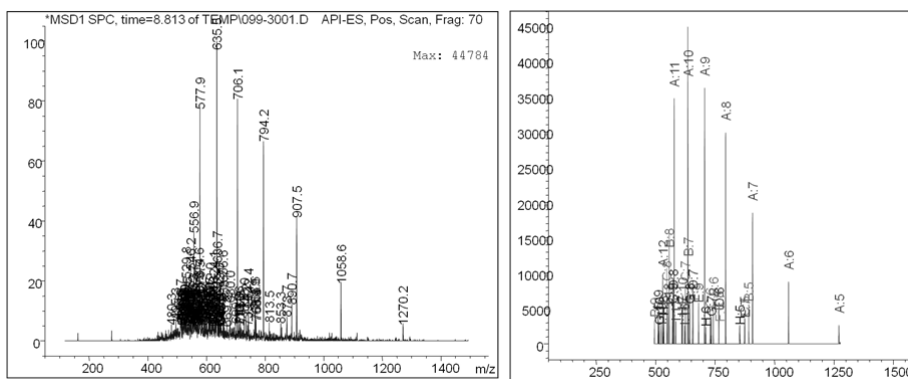
Figure 9.23: ESI-MS spectrum at r.t. = 8.8 min (left) and deconvolution to component A (6346 Dalton)

It is thus clear that the previously described and used cleavage conditions, which include 3 h heating at 60°C are destructive for the sample. More mild cleavage conditions result in the target product, which is still contaminated due to steric hindrance.

Microwave assisted synthesis: .
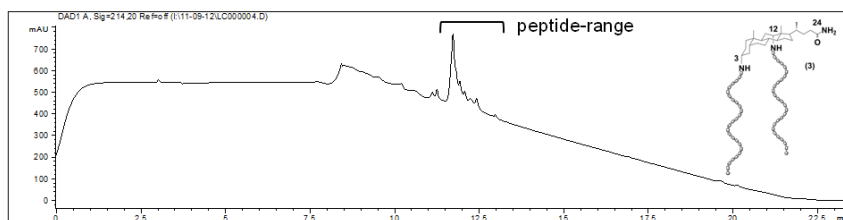
Figure 9.24: RP-HPLC (0-100% in 15 min on C18, 300 Å)

LC-MS (0-100% in 30 min on C18,100): Everything is peptidic in nature and the target mass is the main contributer in the major peak. The following figures are the mass spectrum at r.t. = 9.2 min (left) and the result of the deconvolution (right) in which component A corresponds to multiply charged ions of the peptide with a mass of 6346 Dalton or the target mass.
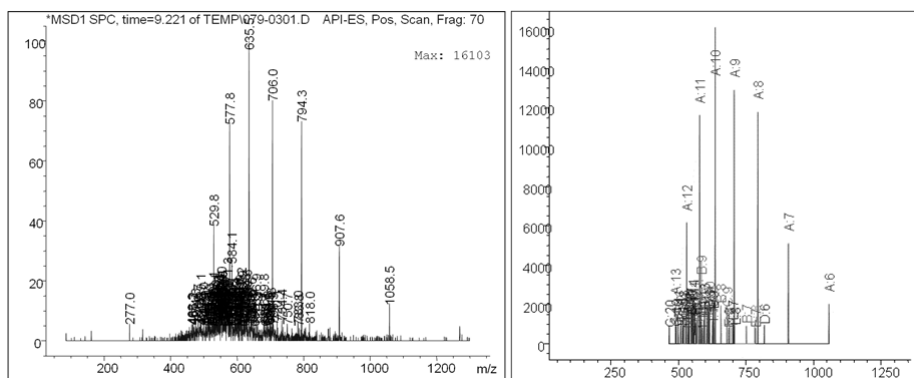
Figure 9.25: ESI-MS spectrum at r.t. = 9.2 min (left) and deconvolution to component A (6346 Dalton)

The purity of **3.1.3** (Fig. 9.24) is improved with microwave assistance (compare to Fig. 9.22)

### 9.1.3   Electrophoretic mobility shift assay data

Method for preliminary data   : DNA duplexes (CRE binding site: 5'-TGG
AGA TGA CGT CAT CTC GT-3' and its complement or Random: 5'-GAA
AAT CAC CCA ACT GCA-3' and its complement), obtained from Eurogentec
purified over gold cartridge) were combined with the mimic in buffer, to obtain
solutions for binding in 17.7 mM Tris buffer pH 7.5, 88.5 mM KCl, 1.77 mM
MgCl2, 1.77 mM EDTA, 0.09x glycerol, 0.02x nonoidet-P40 and 0.01x BSA.
These were incubated on ice for 30 min, before 17 μl was loaded on the gel.
Polyacrylamide gelelectrophoresis experiments were performed on 10 x 10 cm
10% acrylamide:bisacrylamide (37.5:1) gels, ran with 0.5% TBE buffer at 140
V and 4°C for 80 min.   Gels were stained for 10 min with SybrGold (Life
Technologies) and visualized with UV light.

Method   for   final   data :   Performed   as   described   in   the   general
procedure,   with   the   DNA   duplex   for   binding   (CRE   binding   site):   5'-
CGGATGACGTCATTTTTTTC-3' and its complement.

   There is a dose response similar to the one observed for the Morii model,
but smearing due to non-specific interactions is superposed onto it. To illustrate
this point, the intensity in the upshifted area was measured as illustrated in
Figure 9.26. The obtained values are given below. The intensity contained within
the rectangle indicated on the gel in lane 1 is the background intensity. It increases
in the second lane due to the upshifted DNA-protein complex (which contains
already more than half of the DNA). In lane 3 almost all the DNA is involved
in complex formation with the peptide. Upon addition of extra peptide, it is not
possible to form more complex because there is no free DNA left, but non-specific
interactions to the DNA-protein complex increase causing it to further smear out
over the gel.   The total amount of DNA remains the same, and therefore the
measured upshifted intensity too. Without the smearing, all this intensity would
be focused in a single band and the dose response would thus be similar like
observed with the Morii model.

- Lane1. Total intensity $3x10^5$ (= Background)
- Lane2. Total intensity $3x10^6$
- Lane3. Total intensity $4x10^6$
- Lane4. Total intensity $4x10^6$
- Lane5. Total intensity $4x10^6$
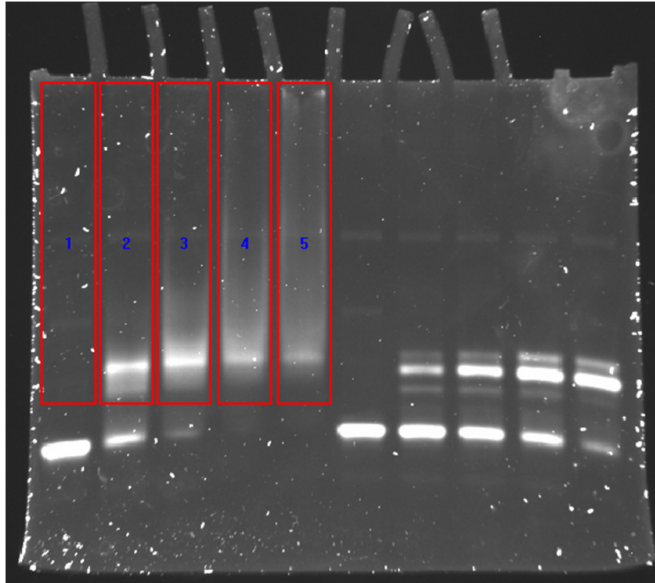
Figure 9.26: Determination of the total upshifted intensity in the EMSA experiment with the defined upshifted areas shown and numbered from left to right.

9.1.4 Optimization of the UV cleavage

A linear peptide was used to optimize the cleavage conditions:

Formula: $C_{263}H_{381}N_{51}O_{56}S_6$

Calculated Mass: 5341.7 (MW: 5345.5)
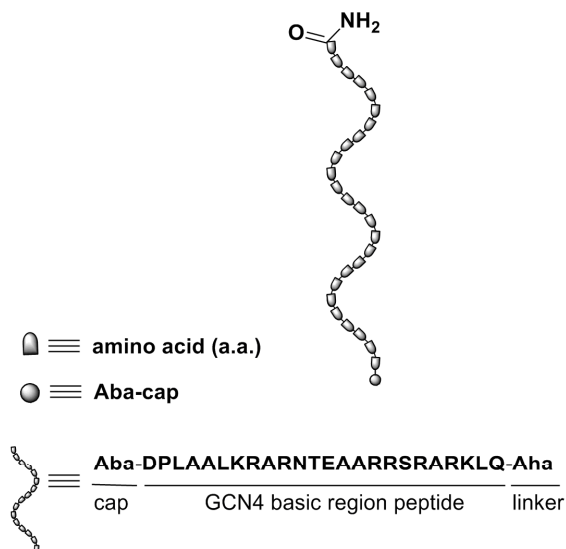
The solid support with peptide (loading 0.109 mmol/g) was split in two and suspended in EtOH. Both halfs were irradiated under different lamp setups. Then the differences in peptide yield and purity were determined:

- 49.6 mg solid support (or theoretical yield 28.9 mg peptide: 3 h irradiation under a small lamp with 1 mW/cm^2 after evaporation and filtration of the sample: 4.4 mg peptide or *15% yield*
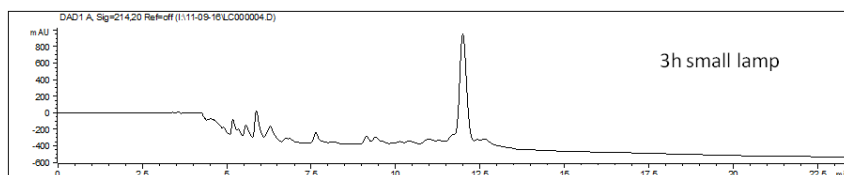
Figure 9.27: RP-HPLC (75-100% MeCN in 15 min on C4, 300 Å) of the peptide after a 3 h cleavage under a small 1 mW/cm^2 lamp

- 26 mg solid support (or theoretical yield 15.1 mg peptide): 3 h irradiation top and bottom irradiation with a more powerful lamp with 15 mW/cm^2 (cleaved sample was collected after 1 h, and after an additional 2 h.

  After evaporation and filtration of the sample: 5.8 mg peptide or *38% yield* (from 1 h) and 5.5 mg peptide or *36% yield* (from 2-3 h). That is a *total yield of 75%.*
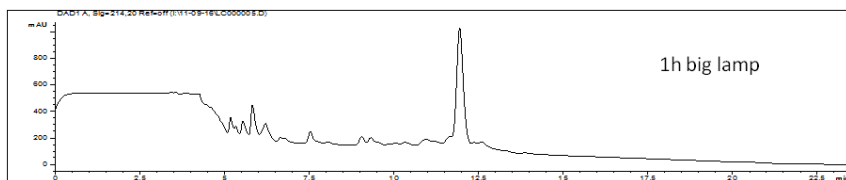
Figure 9.28: RP-HPLC (75-100% MeCN in 15 min on C4, 300 Å) of the peptide after a 1 h cleavage from top and bottom irradiated by a 15 mW/cm^2 lamp.
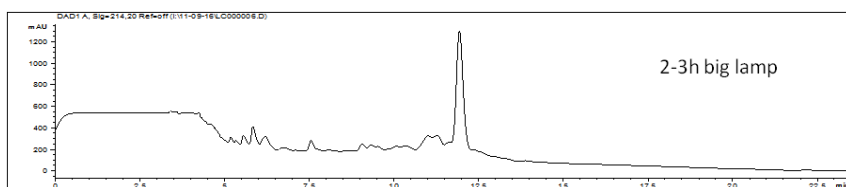
Figure 9.29: RP-HPLC (75-100% MeCN in 15 min on C4, 300 Å) of the peptide after a 2-3 h cleavage from top and bottom irradiated by a 15 mW/cm^2 lamp

The purity after 1 h irradiation by the strong lamp is exactly the same as after 3 h irradiation of the small lamp, but with more than double yield and in a third of the time. Further irradiation of the resin for an additional 2 h allows to double the yield, but with some increased impurities due to the prolonged irradiation.

9.1.5   Purification of final product

Formula: $C_{274}H_{469}N_{101}O_{73}$

Calculated Mass: 6342.6 (MW: 6346.3)
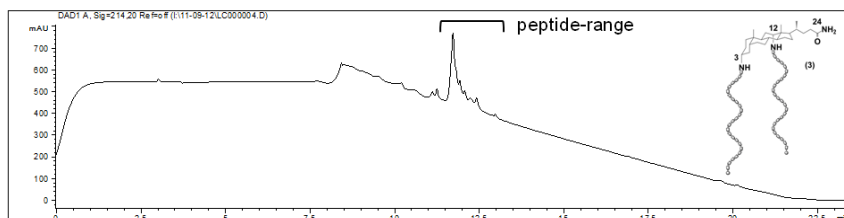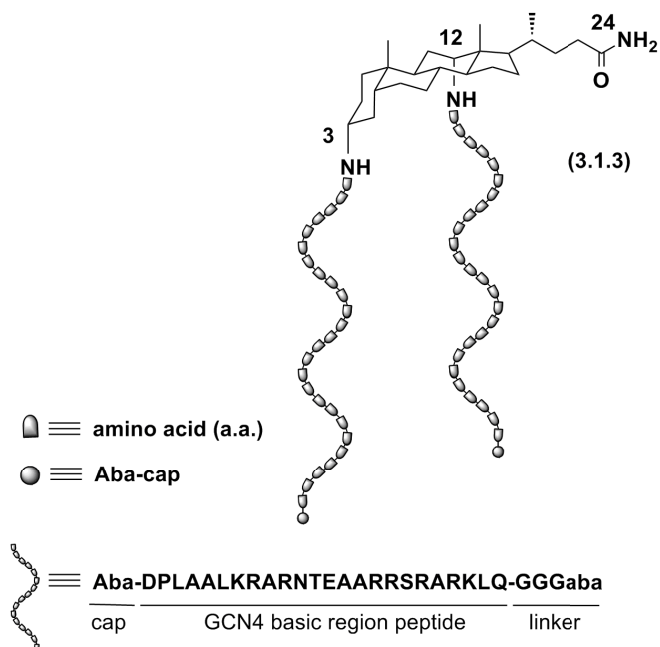
Figure 9.30: RP-HPLC (0-100% MeCN in 15 min on C18, 300 Å) of the crude dipodal peptidosteroid obtained by microwave assisted synthesis
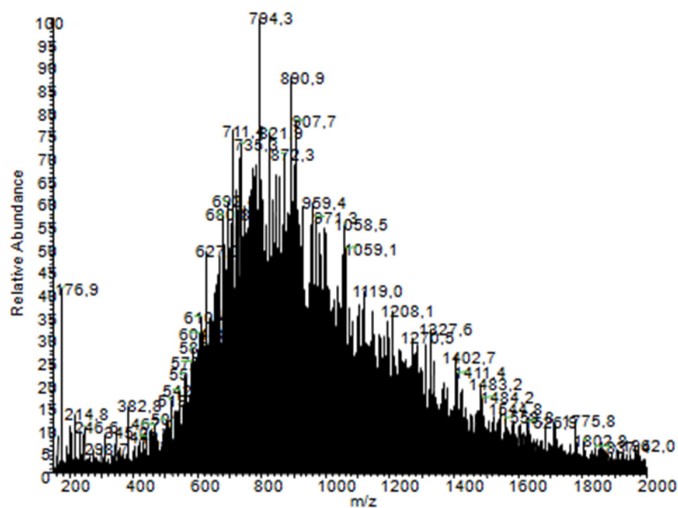
Figure 9.31: ESI-MS: 1270.5 $[M(3.1.3) + 5H]^{5+}$, 1058.5 $[M(3.1.3) + 6H]^{6+}$, 907.7 $[M(3.1.3) + 7H]^{7+}$, 794.3 $[M(3.1.3) + 8H]^{8+}$

The crude dipodal peptidosteroid obtained by microwave assisted synthesis was purified by RP-HPLC, monitored at both 214 and 280 nm. The fraction at retention time 7.5 min was collected (Fig. 9.32).
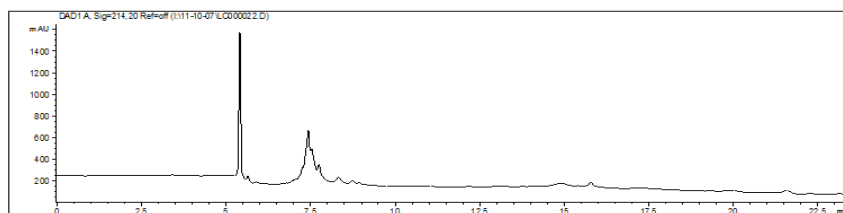
Figure 9.32: RP-HPLC (24-32% MeCN in 16 min on C18, 300 Å) of the purification process

As can be seen from the HPLC trace, purity has dramatically improved upon purification (Fig. 9.33). Purification of such large size constructs by RP-HPLC tedius, as with increasing peptide length, relative differences between target and by-products decrease while variations exponentially increase, this causes strong overlap on HPLC. This explains why the product was not obtained perfectly pure and why smearing during the EMSA occured.
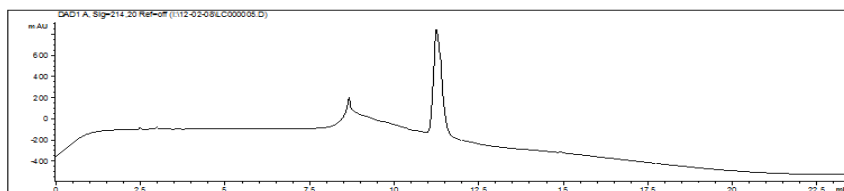
Figure 9.33: RP-HPLC (24-32% MeCNin 16 min on C18, 300 Å) of the purified product

MALDI of the obtained sample (without prior fraction collection!): $6338.4[M(3.1.3) + H]^+$
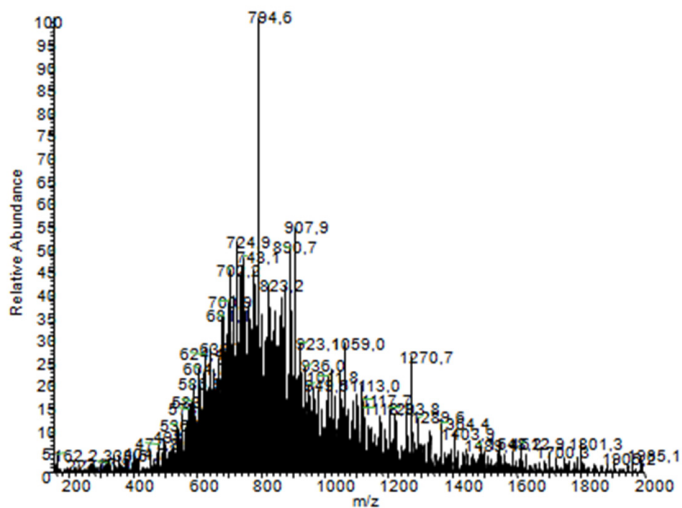
Figure 9.34: ESI: 1270.7 $[M(3.1.3) + 5H]^{5+}$, 1059.0 $[M(3.1.3) + 6H]^{6+}$, 907.9 $[M(3.1.3) + 7H]^{7+}$, 794.6 $[M(3.1.3) + 8H]^{8+}$

9.2 Supporting Material for the Morii Mimic

9.2.1 Synthesis

Because these experiments were performed in the laboratory of Prof. Morii, Kyoto University, Japan. Chemicals from different suppliers (Wako Chemical, TCI, Watanabe and Nacalai Tesque) and different equipment were used. The synthesis was performed according to the reported procedure from literature.[72]

Synthesis of mono-6-deoxy-6-iodo-beta-cyclodextrine

(3.2.1)   (3.2.2)
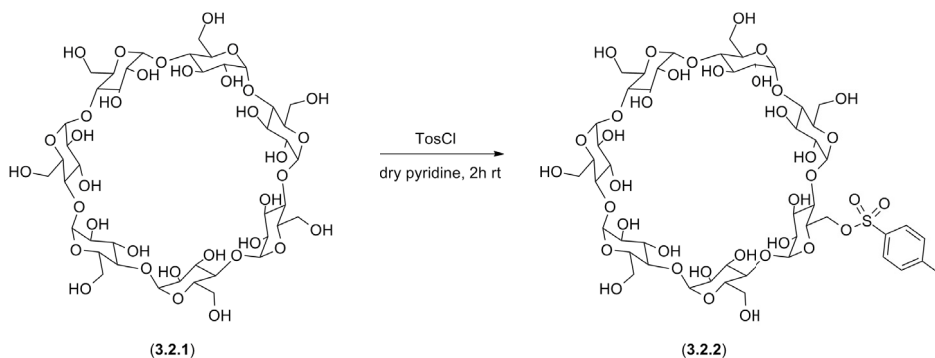
Step 1: monotosylation of beta-cyclodextrine (3.2.2) Before the start of the synthesis, it is necessary to replace crystal water present in the β-cyclodextrine (**3.2.1**) (1 equiv, 5 mmol, 5.675 g, 1135.00 g/mol, Wako Chemical) by dry pyridine (Wako Chemical). This was done, by twice dissolving the cyclodextrine in a small amount of pyridine (10 ml) and coevaporating it with the present water. The present air in the flask was also exchanged by N_2. The cyclodextrine crystals were dissolved in 100 ml dry pyridine. This is difficult and therefore was aided as much as possible by shaking, sonification, stirring and heating. The obtained suspension was cooled down and stirred in an icebath for 10 min before p-toluenesulfonylchloride (0.5 equiv, 2.5 mmol, 0.477 g, Wako Chemical) was added. After addition, the solution was allowed to heat up to room temperature. The reaction was finished after 90 min and quenched with a little water and the solvent was evaporated. Purification was based on the different solubility of modified and unmodified β-cyclodextrine. The

unmodified cyclodextrine is slightly soluble in ice-cold water, while the modified cyclodextrine is not. The majority of the cyclodextine excess could therefore be removed by dissolving it in icecold water and filtration. Because the solubility of cyclodextrine is only limited, this step had to be repeated. The evolution was monitored on TLC and showed that three filtration steps give the desired purity. The residue was recrystallised in water. After 2 days the white crystals were filtered off and dried overnight under vacuum (yield 45%).

Analysis

- Rf product: 0.53 (eluent $NH_3/H_2O/EtOAc/isopropanol$: 1/3/2/5)
- Chemical Formula: $C_{49}H_{76}O_{37}S$
- Mass: 1288.4 (Exact), 1289.2 (MW)
- Maldi: 1311 $[M + Na]^+$ (observed mass)
- ^{1}H-NMR (d6-DMSO, 300 MHz): 7.7 (2H, d, 8 Hz), 7.4 (2H, d, 8 Hz), 5.8-5.7 (14H, m), 4.9-4.8 (7H, m), 4.6-4.4 (6H, m), 3.8-3.2 (42H, m), 2.4 (3H, s)
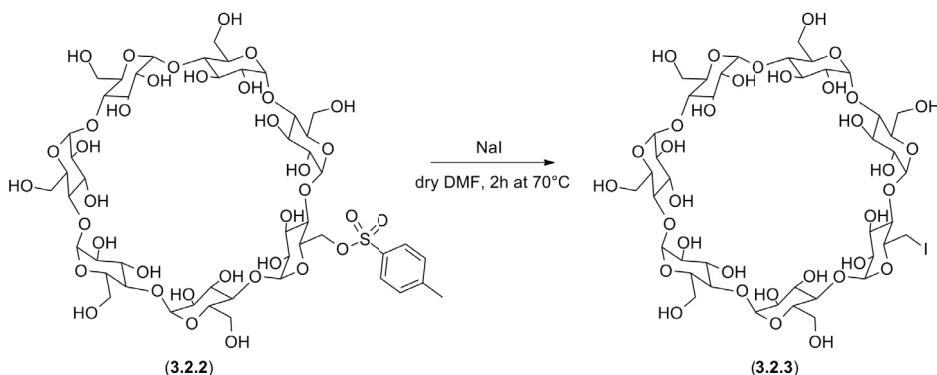
Step 2: substitution of the tosylate by iodine (3.2.3) The tosylated cyclodextrine (**3.2.2**) (1 equiv, 0.8 mmol, 1 g, 1288 g/mol obtained in the previous step) and NaI (10 equiv, 8 mmol, 1.25 g, 149.89 g/mol, Nacalai Tesque) were put together in a 25 ml 2-neck flask. Air was substituted by N_2 and 5 ml of dry DMF (Wako) was added. The mixture was stirred at 70°C for 2 h.

Excess NaI and formed NaTos were filtered off as salts, the filtrate was evaporated and redissolved in the smallest volume of eluent (5% HOAc in water)

able to dissolve the compound for size exclusion chromatography (Sephadex G-10, 7 g swollen in 21 ml of eluent for 3 h). Fractions were combined and evaporated twice, to eliminate all HOAc, before freezedrying. (Indicative yield, before recrystallization is 59%) The obtained product was reprecipitated in 3 ml H_2O/EtOH (2/1), filtered and dried under vacuum (yield 40.0%).

Analysis

- Rf product: 0.47 (eluent NH_3/H_2O/EtOAc/isopropanol: 1/3/2/5)
- Chemical Formula: $C_{42}H_{69}IO_{34}$
- Mass: 1244.3 (Exact), 1244.9 (MW)
- ESI: 1245.2 $[M+H]^+$, 1267.2 $[M+Na]^+$ (observed mass)
- ^{1}H-NMR (d6-DMSO, 300 MHz): 5.8-5.6 (14H, m), 4.9-4.8 (7H, m), 4.6-4.4 (6H, m), 3.8-3.2 (42H, m)

Synthesis of N-(bromoacetyl)-1-adamantane methylamine

(3.2.4)   (3.2.5)
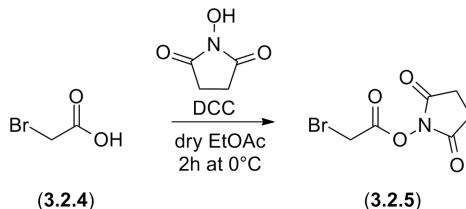
Step 1: Preperation of activated ester (3.2.5)   The reagents bromo-acetic acid (**3.2.4**) (2.0 g, 1 equiv, 14.2 mmol, 138.95 g/mol) and hydroxyl succinimide (1.656 g, 1 equiv, 14.2 mmol, 115.09 g/mol) were combined in a 100 ml 3-neck flask, put under inert atmosphere and dissolved in dry ethyl acetate (40 ml), aided by a little heating. The solution was cooled down to 0°C and DCC was added (3.268 g, 1.1 equiv, 15.9 mmol, 206.18 g/mol). After 2 h the reaction was complete.

The formed urea was filtered off with ice cold EtOAc. The filtrate was washed with 5% $NaHCO_3$, 10% citric acid and water, dried on Na_2SO_4, and concentrated under vacuum (yield 65.5%).

Analysis

- Rf product: 0-0.05 (eluent hexane/EtOAc: 2/1)or 0.75 (eluens CH_2Cl_2 with 5% MeOH)
- Chemical Formula: $C_6H_6BrNO_4$
- Mass: 234.9 (Exact), 236.0 (MW)
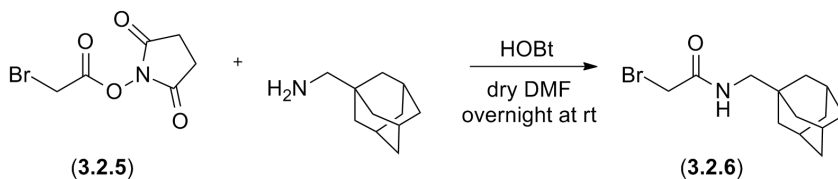
Step 2: Reaction of the activated ester with adamantanemethylamine (3.2.6)   Adamantanemethylamine (400 mg, 2.42 mmol, 1 equiv, 165.28 g/mol; Wako) was put in a small flask filled with N_2, dissolved in dry DMF (4 ml; Nacalai Tesque) and stirred in an ice bath. The activated ester (**3.2.5**) produced in the previous step (1.7 g, 7.26 mmol, 3 equiv, 234.95 g/ml) and HOBt (741 mg, 4.84 mmol, 2 equiv, 153.14 g/mol; Watanabe) were weighed in a small flask, put under inert atmosphere and dissolved in dry DMF (8 ml; Nacalai Tesque) under stirring in an ice bath. After 5 min, this solution was added to the adamantanemethylamine solution. The mixture was left for 5 more min in the icebath and was then allowed to heat till room temperature overnight.

The solvent was evaporated under vacuum, and the product redissolved in 50 ml EtOAc, extracted with 0.1% HCl and brine, dried on Na_2SO_4 and concentrated under vacuum. Purification was obtained by column chromatography (hexane/EtOAc: 2/1), after evaporating the product dissolved in 30 ml of DCM on 500 mg of silica (yield 41%).

Analysis

- Rf product: 0.44 (eluent hexane/EtOAc: 2/1)
- Chemical Formula: $C_{13}H_{20}BrNO$
- Mass: 285.1 (Exact), 286.2 (MW)

- LC-MS: 6.4 min (r.t. on kinetex), 286.1 [M+H]$^+$ (observed mass) and 327.1 [?]
- ^{1}H-NMR (CDCl$_3$, 500 MHz): 6.54 (1H, s), 3.91 (2H, s), 2.94 (2H, d, 6 Hz), 1.98 (3H, m), 1.71 (3H, d, 12 Hz), 1.62 (3H, d, 12 Hz), 1.49 (6H, d, 2 Hz)
- ^{13}C-NMR (CDCl$_3$, 125 MHz): 165 (C), 52 (CH$_2$), 40 (CH$_2$ x3), 37 (CH$_2$ x3), 34 (C), 30 (CH$_2$), 28 (CH x3)

Synthesis of 26-mer GCN4 peptides (3.2.7)

The synthesis was performed in Japan and therefor the methodology differs a little from the described general procedures.

The 26-mer GCN4 peptides (DPAALKRARNTEAARRSRARKLQGGC) were synthesized on a PSSM-8 automatic peptide synthesizer from Shimadzu, on PAL PEG PS resin with low loading (0.180 mmol/g, Applied Biosystems) using 10 fold excess of Fmoc (9-fluorenylmethoxycarbonyl) amino acids (Novabiochem), with 10 equivalents of HBTU (O-Benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate, Novabiochem) as coupling reagent and HOBt (1-hydroxybenzotriazole) as additive from a 1 M solution and 20 equivalents of diisopropylethylamine from a 2 M solution in DMF. Agitation was provided by N$_2$ bubbling. Single couplings were performed during 30 min. Fmoc deprotection was achieved in double 4 min treatments with 30% piperidine in DMF. The peptide was capped manually with acetic anhydride using N-methylimidazole as catalyst. Completion of the reaction was checked by Kaiser test. Final deprotection and cleavage from the solid support was achieved by addition of cold cleavage mixture to the product cooled on ice. The cleavage cocktail consisted of 0.375 g of phenol, 0.5 ml of water, 0.25 ml of thioanisole, 0.125 ml of ethanediol and 5 ml of TFA. After 15 min, the mixture was allowed to heat up to room temperature and was left for 3 h. The pure peptide was obtained by precipitation in cold diethylether, which was dissolved in water and extracted twice with ether, filtered over a sephadex column, purified by HPLC and lyophilised.

Modification of the peptides with Ad (3.2.8)

AcNH-DPAALKRARNTEAARRSRARKLQGGC-CONH$_2$
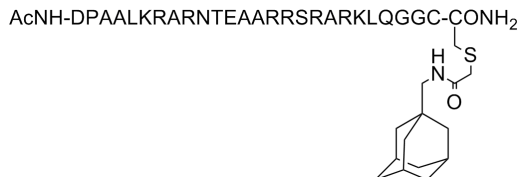
Figure 9.35: The adamantane-modified GCN4 peptide (Ad-GCN4) (**3.2.8**)

Adamantane, modified as bromoacetyl (3.2.6) was used to react with the thiol group of the C-terminal cysteine of the GCN4 peptide (**3.2.7**) to obtain Ad-GCN4 (**3.2.8**). Before reaction the peptides were reduced, by adding 2 ml 0.05 M of DTT in Tris buffer by pH 8 for 1 h at room temperature. DTT was removed by size exclusion (Sephadex G10, 10% acetic acid). The modification reaction was performed in 600 μl phosphate buffer (0.5 M Na$_2$HPO$_4$, pH 9.1). NaOH (1 M, 60 μl) was added and the solution was degassed, by stirring it under reduced pressure and exchanging air for N$_2$. The peptide (5 mg) was dissolved in 200 μl of mQ H$_2$O and added to the stirred buffer in an icebath. N-(bromoacetyl)-1-adamantanemethylamine (5 mg) was dissolved in 200 μl DMF and added too. The pH was checked with pH paper (to be 9) and the reaction mixture was stirred on ice for 5 min. The reaction was quenched with 100 μl of mercaptoethanol for 1 h, after 10 minutes the mixture was allowed to heat up to room temperature. The mixture was neutralized with 100 μl of acetic acid and purified by gelfiltration (10% acetic acid, G10). The product was purified by RP-HPLC.

   Analysis

- Chemical Formula: C$_{130}$H$_{228}$N$_{50}$O$_{36}$S
- Mass: 3097.7 (Exact), 3099.6 (MW)
- Maldi: 3100.5 $[M + H]^+$ (observed mass)
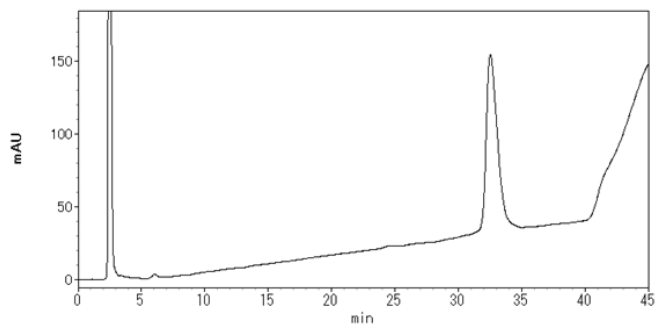
Figure 9.36: RP-HPLC chromatogram of the adamantane-modified GCN4 peptide (purification method: from 29.5 to 47% MeCN/H$_2$O 50/50 with 0.1% TFA from H$_2$O with 0.1% TFA in 35 min on C18 column (150 x 4.6 mm) at rt)
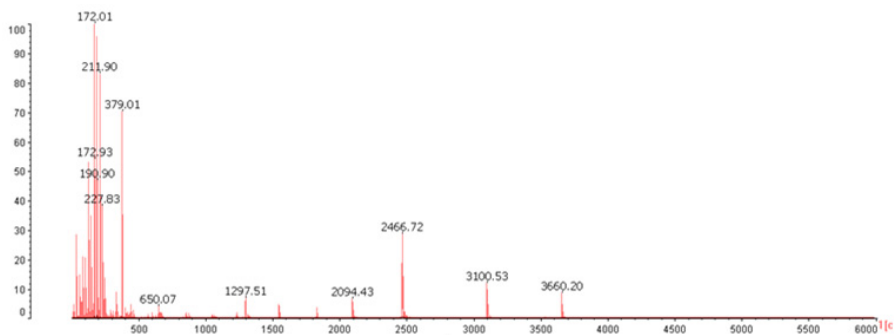
Figure 9.37: MS spectra of the adamantane-modified GCN4 peptide (with internal mass standard mix: 1297.51; 2094.46; 2466.72 and 3660.19 Da)

Modification of the peptides with CD (3.2.9)
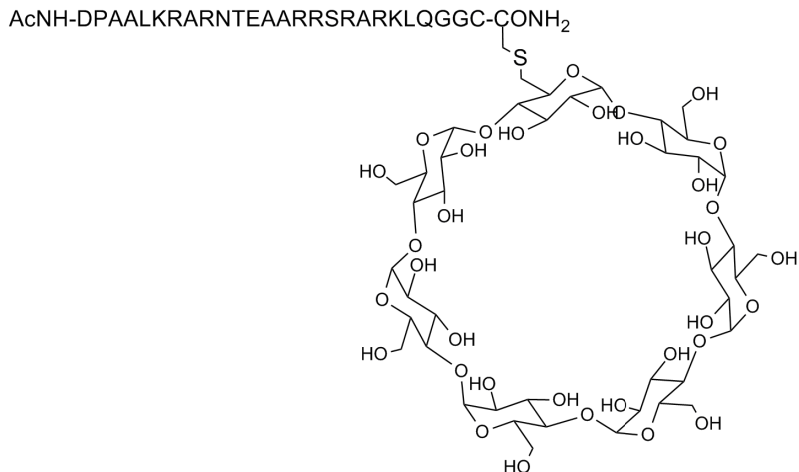
Figure 9.38: The cyclodextrin-modified GCN4 peptide (CD-GCN4) (**3.2.9**)

β-cyclodextrin was monoiodated via a tosylate (**3.2.3**) to allow reaction with the thiol group of the C-terminal cysteine of the GCN4 peptide (**3.2.7**) to obtain CD-GCN4 (**3.2.9**). The method used for this reaction is very similar to the adamantyl modification reaction, but more equivalents of the modified cyclodextrine and a longer reaction time are required. Before reaction the peptides were reduced, by adding 2 ml 0.05 M of DTT in Tris buffer by pH 8 for 1 h at room temperature. DTT was removed by size exclusion (Sephadex G10, 10% acetic acid). The modification reaction is performed in 600 μl phosphate buffer (0.5 M Na$_2$HPO$_4$, pH 9.1). NaOH (1 M, 60 μl) was added and the solution was degassed, by stirring it under reduced pressure and exchanging air for N$_2$. The pepetide (5 mg) was dissolved in 200 μl of mQ H$_2$O and added to the stirred buffer in an icebath. mono-6-deoxy-6-iodo-beta-cyclodextrine (45 mg, 35.03 μmol, 1245 g/mol, 20 equiv) was dissolved in 200 μl DMF and added too. The pH was checked with pH paper (to be 9) and the reaction mixture was stirred on ice. After 1 h the reaction was allowed to heat up to room temperature and was left overnight until all residual unmodified peptide was dimerized as disulfide. The reaction was not quenched like the Ad modification reaction, but the inert atmosphere was broken, the pH was adjusted from 9 to 8 with 10%

acetic acid and purified by gelfiltration (10% acetic acid, G10). The product was purified by RP-HPLC. The peptide disulfide was also purified and after reduction re-introduced in the modification reaction.
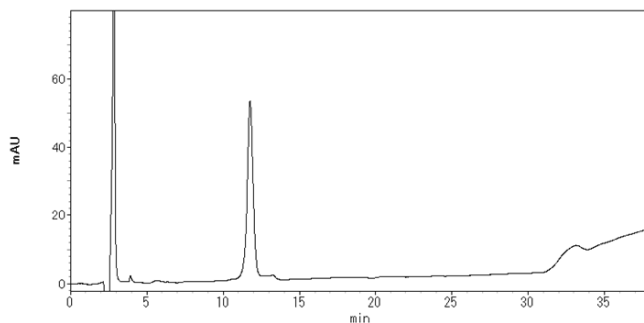
Figure 9.39: RP-HPLC chromatogram of the cyclodextrin-modified GCN4 peptide (purification method: from 26 to 39% MeCN/H$_2$O 50/50 with 0.1% TFA from H$_2$O with 0.1% TFA in 26 min on C18 column (150 x 4.6 mm) at rt)
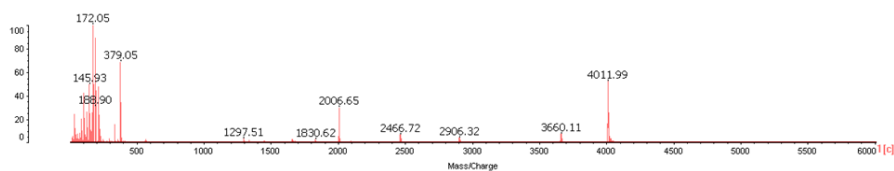
Figure 9.40: MS spectra of the cyclodextrin-modified GCN4 peptide (with internal mass standard mix: 1297.51; 2094.46; 2466.72 and 3660.19 Da)

Analysis

- Chemical Formula: C$_{159}$H$_{277}$N$_{49}$O$_{69}$S
- Mass: 4008.9 (Exact), 4011.3 (MW)
- Maldi: 4012.0 $[M + H]^+$ (observed mass)

Formation of the Ad-CD inclusion complex (3.2.10)

Upon mixing the two peptides together, a dimer was formed through an Ad-CD inclusion complex, in this way mimicking the homodimeric transcription factor GCN4.

9.3 Supporting Material for the Case Study: Kinetin

9.3.1 Synthesis of nu4-modified ODNs

O6-Phenyl dI-CE phosphoramidite (O6-Phenyl-5'-Dimethoxytrityl-2'deoxyinosine, 3'-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite) (Glen Research) was manually incorporated in an ODN during standard automated phosphoramidite synthesis, as described for modified phosphoramidites in the standard procedures. After synthesis ODNs were cleaved from the CPG by treatment with concentrated NH_3 solution (28% NH_4OH) at room temperature for 4h, rather than the typical overnight treatment at 55°C. The obtained mixture was DMTr deprotected and purified over Sep-Pak (Waters) into 5 fractions, which were analysed, lyophylised and then treated with 1 M aqueous furfurylamine at 55°C for at least 24h. The DNA was purified by removal of the furfurylamine over a 3 kDa cut-off filter (Pall, Life Science or Amicon, Millipore) by centrifugation.

Analysis

- ODN4.1.1 (5'-CGG ATG**nu4**CG TCA TTT TTT TTC-3') observed mass (deconvoluted) 6471.72 Da, calculated mass 6470.1 (Exact) and 6473.2 (MW) Da
- ODN4.1.2 (5'- CTG ACG G**nu4**G TGC-3') observed mass (deconvoluted) 3765.52 Da, calculated mass 3764.6 (Exact) and 3766.6 (MW) Da
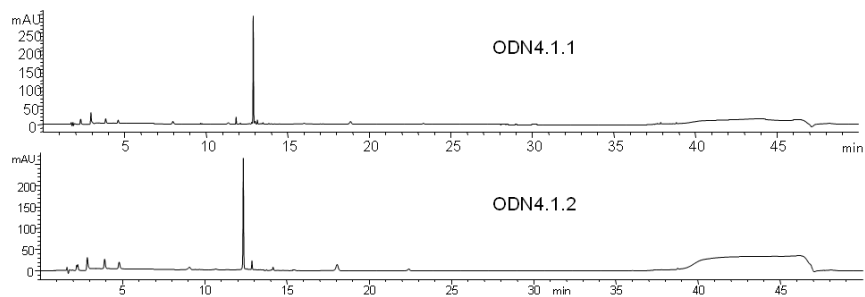
Figure 9.41: HPLC chromatograms of the nu4-modified ODNs ODN4.1.1 (top) and ODN4.1.2 (bottom). RP-HPLC chromatogram from Aeris column with a gradient from 0.1 M TEAA-buffer to 30% MeCN in 30 min at 60°C
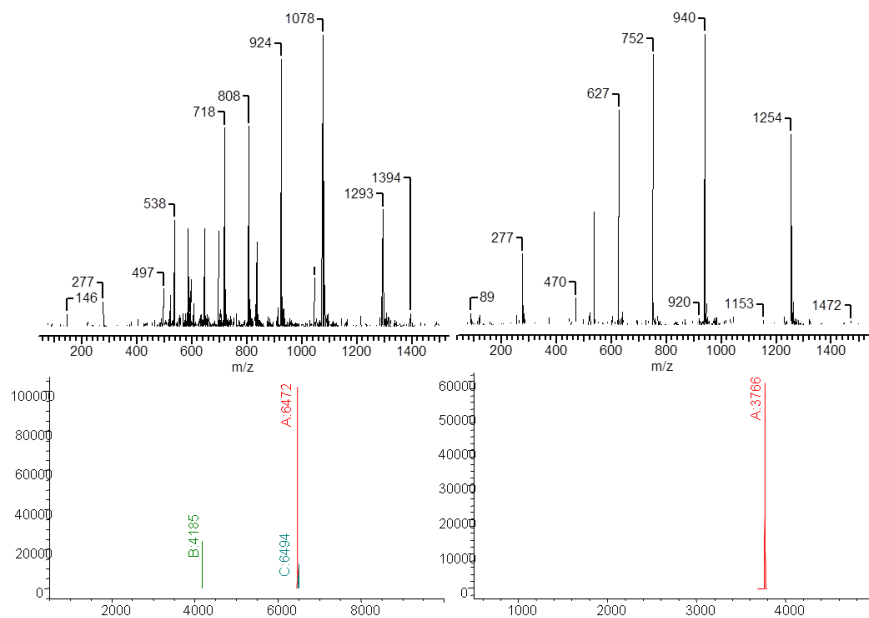
Figure 9.42: MS spectra of the nu4-modified ODNs ODN4.1.1 (left) and ODN4.1.2 (right)

9.3.2 HPLC chromatograms of the cross-linking mixtures with kinetin modification
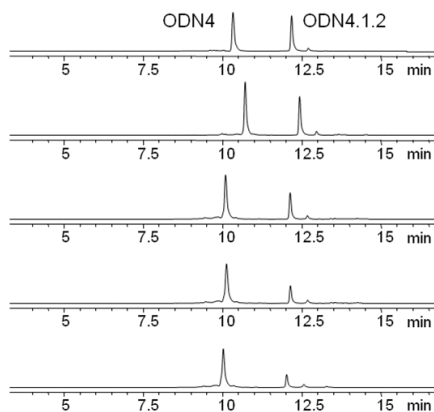
Figure 9.43: HPLC of ICL reaction mixture of ODN4.1.2-ODN4 with 0, 1, 2, 3 and 4 equiv of NBS from top to bottom. No clear ICL product can be observed.
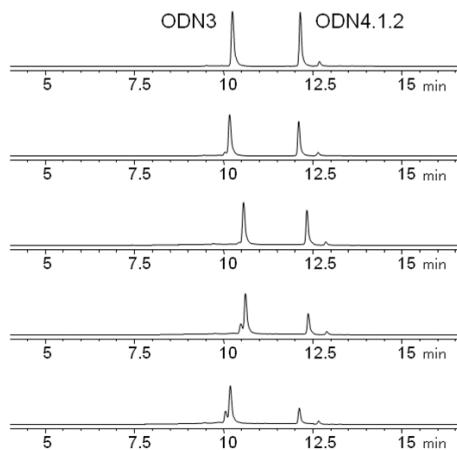
Figure 9.44: HPLC of ICL reaction mixture of ODN4.1.2-ODN3 with 0, 1, 2, 3 and 4 equiv of NBS from top to bottom. No clear ICL product can be observed.
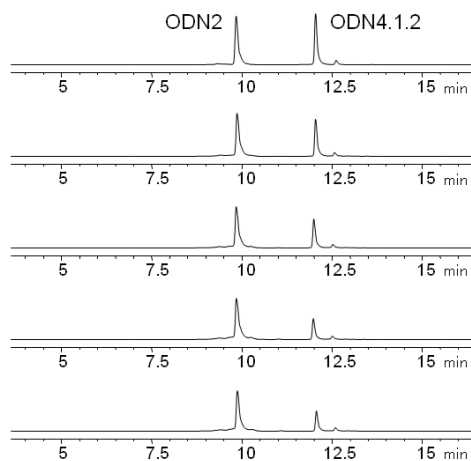
Figure 9.45: HPLC of ICL reaction mixture of ODN4.1.2-ODN2 with 0, 1, 2, 3 and 4 equiv of NBS from top to bottom. No clear ICL product can be observed.
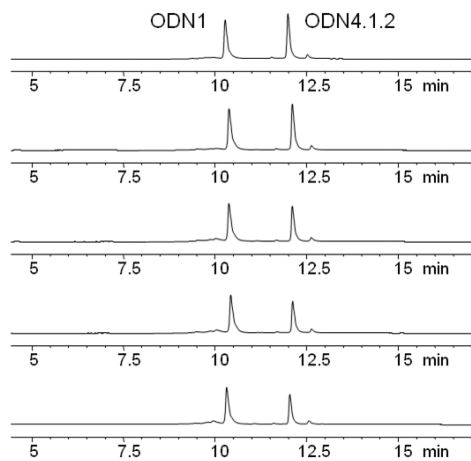
Figure 9.46: HPLC of ICL reaction mixture of ODN4.1.2-ODN1 with 0, 1, 2, 3 and 4 equiv of NBS from top to bottom. No clear ICL product can be observed.

9.4 Supporting Material for DNA ← Protein Mimic Cross-Linking

9.4.1 Synthesis of furan-modified amino acids

Synthesis of Fmoc protected aa3

(4.2.6)   (4.2.7)
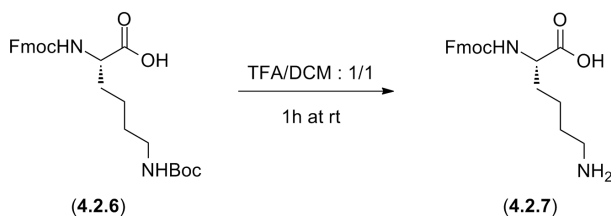
Step 1: Boc deprotection   TFA (1 ml / 100 mg) was added to a suspension of Boc-protected lysine (**4.2.6**) (468.55 g/mol) in DCM (1 ml / 100 mg). This immediately dissolves the product. After 30 min, the reaction was complete. The product (**4.2.7**) was worked up by evaporating under vacuum, and co-evaporation with MeOH. The reaction yield was calculated as 168%, this is probably due to TFA-salt formation.

Analysis

- Rf product: 0.22 (RP, eluent H_2O/MeOH : 1/1 + 1% MeCOOH)
- Chemical Formula: $C_{21}H_{24}N_2O_4$
- Mass: 368.2 (Exact), 368,4 (MW)

(4.2.7)   (4.2.8)
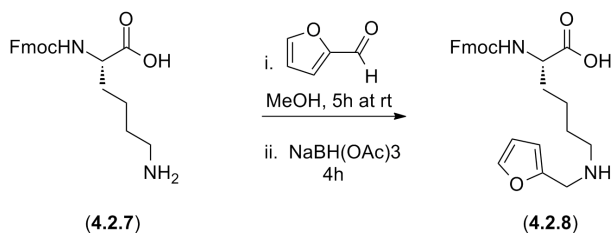
Step 2 + 3: reductive amination and Boc protection   In a multistep one-pot reaction Boc-deprotected lysine (**4.2.7**) (50 mg, 5% molar excess, 0.1357

mmol, 368.42 g/mol) was converted to the imine, reduced and Boc-reprotected. The reaction was performed in MeOH (3.5 ml) to dissolve the starting material. Furfural 99% (10.71 μl, 1 equiv, 0.12925 mmol, 96.02 g/mol, 1.159 g/mol) was added. The reaction was followed on TLC after reduction on small scale, as the formed imine decomposes on TLC. After 5 h NaBH(OAc)$_3$ (40 mg, 1.4 equiv, 1.8095 mmol, 212.05 g/ml) was added.

Analysis

- Rf product: 0.14 (normal phase, CHCl$_3$/MeOH 5/1)

(4.2.8)　　　　　　　　　　　**(4.2.3)**
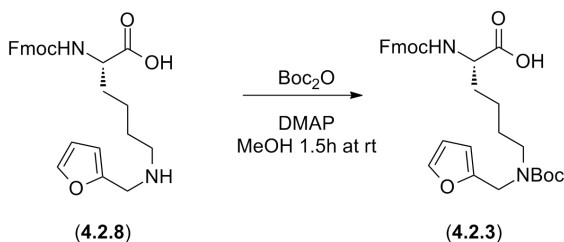
After 4 h the reaction was quenched with Boc$_2$O (32 μl, 1 equiv to starting material, 0.1357 mmol, 218.55 g/mol, 0.95 g/mol, 29.657 mg), which was activated with DMAP (2 mg, 0.1 equiv, 0.01357 mmol, 122.17 g/mol, 1.658 mg). After 45 minutes, the product is clearly formed. The reaction was finished by evaporation after 1.5 h, with some start product remaining. The residue was redissolved in 10 ml EtOAC and washed with 10% citric acid and brine, dried over Na$_2$SO$_4$ and evaporated (crude yield 32.39 mg, 46%, mainly due to the extensive sampling). The product (**4.2.3**) was purified by flash column chromatography (EtOAc/Hexane 3/1 + 0.3% acetic acid).

Analysis

- Rf product: 0.58 (normal phase, CHCl$_3$/MeOH 5/1)
- Chemical Formula: C$_{31}$H$_{36}$N$_2$O$_7$
- Mass: 548.3 (Exact), 548.6 (MW)
- FAB: 549 [M + H]$^+$; 571 [M + Na]$^+$
- ^1H-NMR (d6-DMSO): not measured, due to abandonned synthesis

Synthesis of Fmoc protected aa4

(**4.2.7**)   (**4.2.4**)
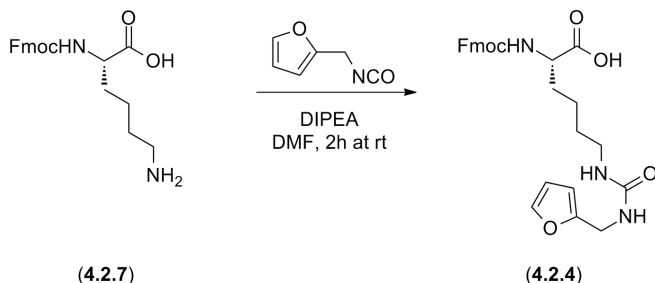
Step 2: coupling reaction   Fmoc-lysine-OH (**4.2.7**) (1.97 g, 5.34 mmol, 1
equiv, 368.4 g/mol) obtained as described before, was co-evaporated with DMF
to remove all TFA and H_2O. 12.3 ml dry DMF was added to obtain a 0.43 M
solution. DIPEA (1.94 ml, 2.1 equiv, 11.11 mmol, 1.43 g, 129.25 g/mol, 0.742
g/ml) was added and air was exchanged for N_2. The mixture was stirred in an
icebath, isocyanate (1.19 ml, 2.1 equiv, 11.11 mmol, 123.11 g/mol, 1.148 g/ml)
was added and after 10 minutes the suspension was allowed to heat up to room
temperature. The reaction was compleet after 1 h. The product (**4.2.4**) was
purified by reprecipitation (dissolve in chloroform and crash out with hexane)
and recrystalization (chloroform) and obtained in an unoptimized yield of 60%.

    Analysis

- Rf product: 0.18 (eluent $MeOH/CHCl_3$: 1/5 + 1% MeCOOH)
- Chemical Formula: $C_{27}H_{29}N_3O_6$
- Mass: 491.2 (Exact), 491.5 (MW)
- FAB: 492 $[M + H]^+$; 514 $[M + Na]^+$
- ^1H-NMR (d6-DMSO, 500 MHz): 12.5 (1H, s), 7.9 (2H, d, 7.5 Hz), 7.7 (2H,
  d, 7.5 Hz), 7.6 (1H, d, 7.9 Hz), 7.5 (1H, d, 1 Hz), 7.4 (2H, t, 7.5 Hz), 7.3
  (2H, t, 7.5 Hz), 6.4 (1H, dd, 1/3 Hz), 6.18 (1H, t, 5.7 Hz), 6.16 (1H, d, 3
  Hz) 5.9 (1H, t, 6 Hz), 4.3 (2H, d, 7 Hz), 4.2 (1H, t, 7 Hz), 4.1 (2H, d, 5.7
  Hz), 3.9 (1H, m), 3.0 (2H, q, 6/12 Hz), 1.7 (2H, m), 1.3 (4H, m)
- ^{13}C-NMR (d6-DMSO, 125 MHz): 174 (C), 158 (C), 157 (C), 154 (C), 144
  (C x2), 142 (CH), 141 (C x2), 128 (CH x2), 127 (CH x2), 126 (CH x2), 121

(CH x2), 111 (CH), 107 (CH), 66 (CH$_2$), 54 (CH), 47 (CH), 40 (CH$_2$), 37 (CH$_2$), 31 (CH$_2$), 30 (CH$_2$), 24 (CH$_2$)

Synthesis of Fmoc protected aa5

(4.2.9) **(4.2.10)**
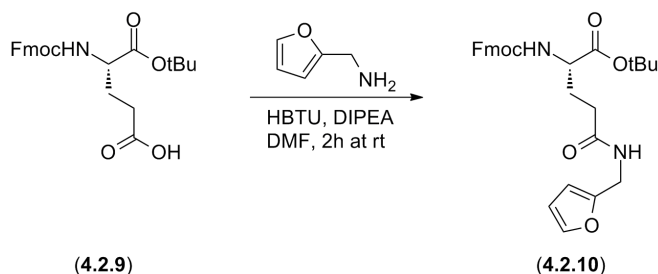
Step 1: Amide bond formation Fmoc-Glu-OtBu (**4.2.9**) (1.5 g, 1 equiv, 3.53 mmol, 425.5 g/mol) was weighed in a 2 neck flask, put under N$_2$ atmosphere and dissolved in 13 ml dry DMF. Furfurylamine (0.311 ml, 1 equiv, 3.53 mmol, 97.12 g/mol, 1.099 g/ml), 1.2 equiv of HBTU (4.23 mmol, 1.6 g, 379.25 g/mol) and 2 equiv of DIPEA (7.05 mmol, 1.23 ml, 129.25 g/mol, 0.742 g/ml) was added and the mixture was stirred at room temperature for 2 h. The reaction was finished by evaporating the solvent, the residual yellow oil was dissolved in EtOAc, washed 3x with 5% citric acid, 10% NaHCO$_3$ and brine and dried over Na$_2$SO$_4$. Pure product (**4.2.10**) was obtained in 98% yield

Analysis

- Rf product: 0.54 (eluent EtOAc/Hexane : 2/1)
- Chemical Formula: C$_{29}$H$_{32}$N$_2$O$_6$
- Mass: 504.2 (Exact), 504.6 (MW)
- FAB: 505 [M + H]$^+$
- ^1H-NMR (DMSO, 300 MHz): 8.3 (1H, t, 5.3 Hz); 7.9 (2H, d, 7.5 Hz); 7.7 (2H, d, 7.5 Hz); 7.6 (1H, d, 7 Hz); 7.5 (1H, d, 1 Hz); 7.4 (2H, t, 7.5 Hz); 7.3 (2H, t, 7.5 Hz); 6.4 (1H, dd, 1/3 Hz); 6.2 (1H, d, 3 Hz); 4.3-4.2 (5H, m); 3.9 (1H, m) 2.2 (2H, t, 7.5); 1.9 (1H, m); 1.8 (1H, m); 1.4 (9H, s) (impurity 8.3 (singlet superposed on amide signal))

- ^{13}H-NMR (d6-CDCl$_3$, 125 MHz): 172 (C), 171 (C), 156 (C), 151 (C, small but confirmed by HMBC crosspeak, 144 (2x C), 142 (CH), 141 (2x C), 128 (2x CH), 127 (2x CH), 126 (2x CH), 120 (2x CH), 110 (CH), 107 (CH), 83 (C), 67 (CH$_2$), 54 (CH), 47 (CH), 37 ((CH$_2$)), 32 (CH$_2$), 29 (CH$_2$), 28 (CH$_3$ x3)
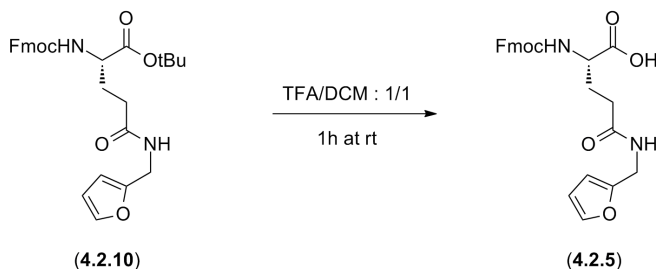
Step 2: tertiairy butyl deprotection   The obtained furan-modified glutamic acid (**4.2.10**) obtained in the previous reaction was dissolved in DCM (1 ml / 100 mg) and cooled down in an icebatch. TFA (1 ml / 100 mg) was added slowly while the mixture was stirred on ice. After 5 min the reaction was allowed to heat up to room temperature. After 3 h, the reaction was finisched by evaporation under vacuum. The product (**4.2.5**) was purified over a short silica plug (MeOH/chloroform : 1/6 + 0.1% acetic acid) and obtained in an unoptimized overall yield of 56%.

Analysis

- Rf product: 0.20 (eluent EtOAc/Hexane : 2/1)
- Chemical Formula: C$_{25}$H$_{24}$N$_2$O$_6$
- Mass: 448.2 (Exact), 448.5 (MW)
- FAB: 449 [M + H]$^+$
- ^1H-NMR (d6-DMSO, 500 MHz): 12.6 (1H, s); 8.3 (1H, t, 5.3 Hz); 7.9 (2H, d, 7.5 Hz); 7.7 (2H, d, 7.5 Hz); 7.6 (1H, d, 7 Hz); 7.5 (1H, d, 1 Hz); 7.4 (2H, t, 7.5 Hz); 7.3 (2H, t, 7.5 Hz); 6.4 (1H, dd, 1/3 Hz); 6.2 (1H, d, 3 Hz); 4.3-4.2 (5H, m); 4.0 (1H, m) 2.2 (2H, t, 6.8); 2.0 (1H, m); 1.8 (1H, m)
- ^{13}H-NMR (d6-DMSO, 125 MHz): 172 (C), 157 (C), 153 (C), 144 (2x C),

142 (CH), 141 (2xC), 128 (2x CH), 128 (2x CH), 126 (2x CH), 121 (2x CH), 111 (CH), 107 (CH), 66 (CH_2), 53 (CH, very small but confirmed by HSQC crosspeak), 47 (CH), 36 ((CH_2)), 32 (CH_2), 27 (CH_2), C_q signal is missing

9.4.2 Synthesis of furan-modified peptides

Peptide synthesis was performed as described in the general protocol, using **aa1** at positions K231 or A239 of the brGCN4 peptide. The subsequent modification with Ad, was performed as described above.

brGCN4(K231aa1)

Analysis

- Chemical Formula: $C_{118}H_{204}N_{48}O_{36}S$
- Mass: 2901.5 (Exact), 2903.3 (MW)
- Maldi: 2904.2 $[M+H]^+$ (observed mass)
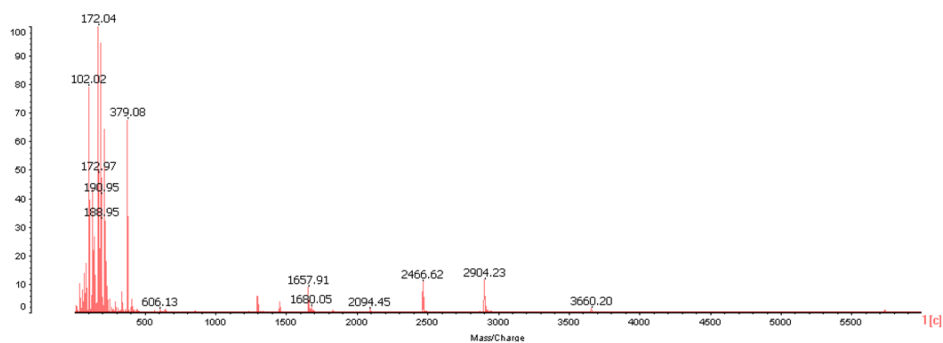
Figure 9.47: MS spectum of brGCN4(K231aa1) (with an internal mass standard mix: 1297.5, 2094.5, 2466.7 and 3660.2 Dalton, before purification)

brGCN4(A239aa1)

Analysis

- Chemical Formula: $C_{121}H_{211}N_{49}O_{36}S$
- Mass: 2958.6 (Exact), 2960.3 (MW)
- Maldi: 2961.4 $[M+H]^+$ (observed mass)
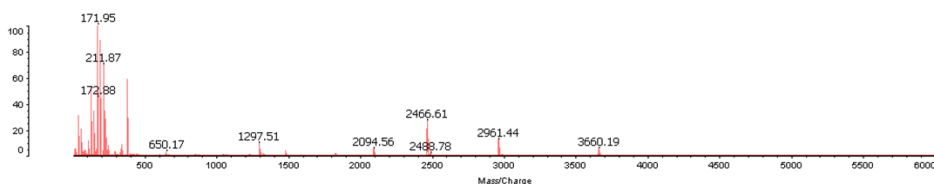
Figure 9.48: MS spectum of brGCN4(A239aa1) (with an internal mass standard mix: 1297.5, 2094.5, 2466.7 and 3660.2 Dalton)

Ad-brGCN4(K231aa1)

Analysis

- Chemical Formula: $C_{131}H_{223}N_{49}O_{37}S$
- Mass: 3106.7 (Exact), 3108.5 (MW)
- Maldi: 3109.2 $[M+H]^+$ (observed mass)
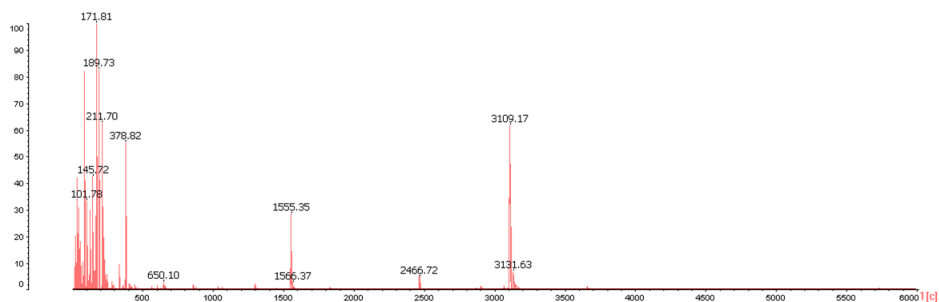
Figure 9.49: MS spectum of Ad-brGCN4(K231aa1) (with contamination of the internal mass standard at 2466 Dalton)

Ad-brGCN4(A239aa1)

Analysis

- Chemical Formula: $C_{134}H_{230}N_{50}O_{37}S$
- Mass: 3163.7 (Exact), 3165.6 (MW)
- Maldi: 3166.6 $[M+H]^+$ (observed mass)
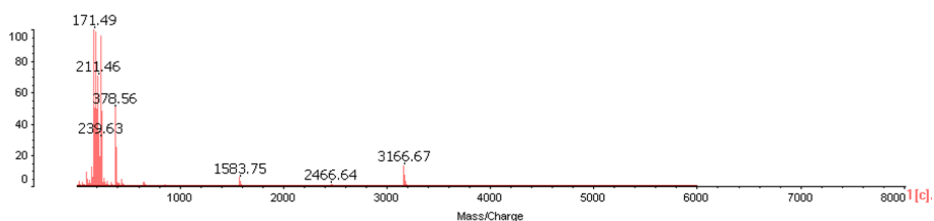
Figure 9.50: MS spectum of Ad-brGCN4(A239aa1) (with contamination of the internal mass standard at 2466 Dalton)

9.4.3   DNA binding analysis

DNA binding analysis was performed as described in the general protocol, using the radiolabeled ^{32}P-DNA sequences AP1 (5'-CGGATGACTCATTTTTTTTC-3'), CRE (5'-CGGATGACGTCATTTTTTTTC-3') and CECRE (5'-CGGATTGCGTCATTTTTTTTC-3').

9.4.4   Analysis of furan oxidation by hydazine trapping

Ad-bGCN4(A239aa1) (0.75 nmol, 50μM) was oxidized with NBS (2 equiv, 5 μl) on ice for 1 h. Then hydrazine was added (100 equiv, 5 μl). The mixture was left on ice for 30 min, then it was shaken 750 rpm at 25°C for another 30 min. The control experiments followed the same protocol but A) omitted the NBS oxidation step, B) added NBS and hydrazine at the sam time and C) exactly reproduced these conditions.

   Analysis

- Chemical Formula: $C_{134}H_{230}N_{52}O_{37}S$
  vs. starting material (Ad-bGCN4(A239aa1)): $C_{134}H_{230}N_{50}O_{37}S$
- Mass: 3191.7 (Exact), 3193.7 (MW)
  vs. starting material (Ad-bGCN4(A239aa1)): 3163.7 (Exact), 3165.6 (MW)
- Observed mass A: 3167 Dalton (Maldi); no reaction
- Observed mass B: 3167 Dalton (Maldi); no reaction
- Observed mass C: 3196 Dalton (Maldi); reaction

9.4.5   Analysis of cross-link formation

Analysis of cross-link formation was performed as described in the general protocol. No spectral data is shown here due to the numerous attempts, all with negative outcome.

9.5 Supporting Material for DNA → Protein Mimic Cross-Linking

9.5.1 Synthesis of nu5

Nucleoside **nu5** was synthesized through a Stille cross coupling of stannylated furan with commercially available 5-iodo-2'-deoxyuridine. DMTr protection and conversion into the phosphoramidite, was carried out by M. Op de Beeck as described in literature.[209]

9.5.2 Synthesis of nu5-modified CRE

The modified residue **nu5** (**4.3.1**) was incorporated into the CRE (5'-CGG ATG ACG TCA TTT TTT TTT TTC-3') DNA binding site of the transcription factor mimic of GCN4, replacing natural T (CRE1 and CRE2) or natural G (CRE3) residues. The synthesis was performed as described in the general procedures.

Analysis

- CRE1 (5'-CGG A**nu5**G ACG TCA TTT TTT TTC-3') observed mass (deconvoluted) 6443.79 Da, calculated mass 6442.07 (Exact) and 6445.17 (MW) Da with formula $C_{209}H_{263}N_{67}O_{132}P_{20}$
- CRE2 (5'-CGG ATG ACG **nu5**CA TTT TTT TTC-3') observed mass (deconvoluted) 6443.85 Da, calculated mass 6442.07 (Exact) and 6445.17 (MW) Da with formula $C_{209}H_{263}N_{67}O_{132}P_{20}$
- CRE3 (5'-CGG AT**nu5** ACG TCA TTT TTT TTC-3') observed mass (deconvoluted) 6418.53 Da, calculated mass 6417.06 (Exact) and 6420.16 (MW) Da with formula $C_{209}H_{264}N_{64}O_{133}P_{20}$
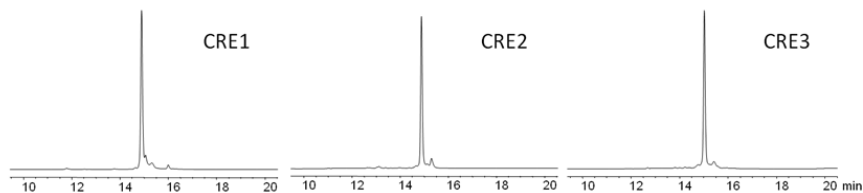
Figure 9.51: RP-HPLC chromatograms of the nu5-modified CRE ODNs
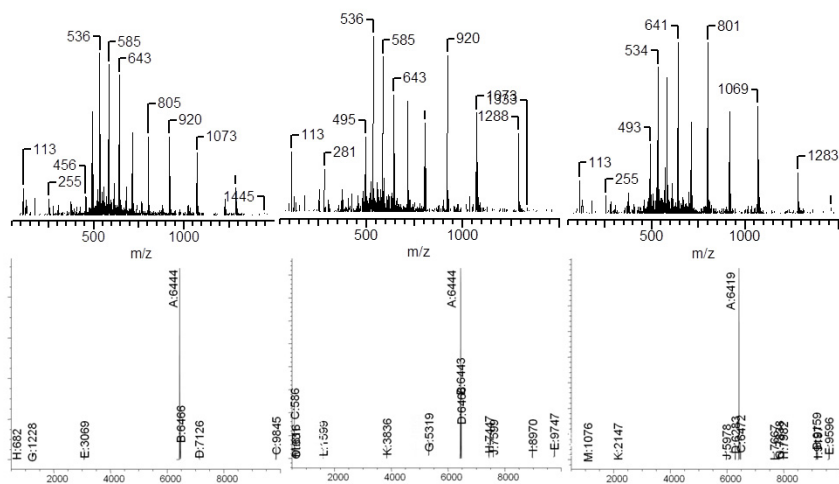
Figure 9.52: MS spectra (top) and the deconvolution (bottom) of the nu5-modified CRE ODNs (from left to right: CRE1, CRE2, CRE3)

9.5.3   Melting temperatures of the formed duplexes

Duplex	T_m (°C)	ΔT_m (°C)
CRE:CRE comp	60.7	
CRE1:CRE comp (**nu5**)	59.7	-1
CRE2:CRE comp (**nu5**)	60.7	0
CRE3:CRE3 comp (**nu5**)	57.6	-3.1
CRE4:CRE comp (**nu4**)	54.5	-6.2

9.5.4 Cross-linking

The reaction was performed as described in the general procedure. To achieve full binding, 5 equivalents of protein mimic, or thus of both peptides (5 nmol) were added to the duplex. Since NBS will quickly oxidize the thioether bonds in Ad-GCN4 and CD-GCN4, 1 equiv of NBS should equal 11 nmol. The reduction with $NaCNBH_3$ (20 nmol) restored the thioether bonds.

9.5.5 Mass analysis of the formed product with CRE3

With the purpose of structural characterization, the obtained product was isolated by purification on HPLC (first on a Clarity column (Phenomenex) and then on an Aeris column (Phenomenex)). Mass spectra were obtained by using a Nano-Acquity HPLC (Waters Corporation, Milford, MA) in-line coupled to a Q-Tof Premier mass spectrometer (Waters Corporation). Therefore the purified sample was dissolved in 5 mM ammonium acetate buffer of pH 5 and first loaded onto a trapping column (100 μm internal diameter 20 mm, packed in-house with Reprosil-Pur Basic C18-HD, 5 μm) with 5 mM ammonium acetate in water containing 1% MeCN pH 5. It was then further transferred onto the nano column (75 μm internal diameter 150 mm, packed in-house with Reprosil-Pur Basic-C18-HD, 3 μm) with 5 mM ammonium acetate in water pH 5. Material retained by the C18 column was eluted using a linear gradient of 1.69% acetonitrile increase per minute at a flow rate of 300 nl/min and analyzed in MS mode. The mass spectrometer was calibrated every 10 seconds using Leu-Enkephalin as lockmass. Molecular weight of the conjugate was reconstructed from the m/z values of the multiply charged ions using the mass deconvolution program of the instrument.

Analysis

- CRE3 - CD-GCN4 observed mass (deconvoluted) 10431.0 Da, calculated mass 10426.0 (Exact) and 10431.4 (MW) Da with formula $C_{368}H_{541}N_{113}O_{202}P_{20}S$

9.5.6   Additional Data

DPC under High Salt Conditions

Cross-linking was performed as described above using the CRE3 duplex, but at 1 M NaCl concentration.

DPC attempt in absence of GCN4 Mimic

A cross-linking experiment with CRE3 duplex in the absence of the GCN4 mimic was performed under identical conditions as described before. A volume of MilliQ water was added equal to the used volume of peptide, to ensure the same final concentration for the reaction and obtain the best comparable results.

DPC attempt without NaCNBH$_3$ reduction

A cross-linking experiment with CRE3 duplex was performed under identical conditions as described before, and afterward it was divided in two. Half of the sample was reduced with NaCNBH$_3$, the other half was left untreated. Both samples were stored at 20°C for 3 weeks and analyzed together.

ICL experiment with CRE2 in the absence of protein mimic

Additional cross-link reactions were performed with CRE2 and its complements CRE comp, CRE2 compT and CRE2 compC. These ICL experiments were performed as DPC experiments, described in the general protocol in the absence of mimic. A volume of milliQ water was added equal to the normally used volume of peptide, this assures the same final concentration for the ICL and DPC reactions and allows comparison of the obtained results. As can be seen in Figure 9.53 there is no ICL formation. The earlier observed product in presence of mimic, thus requires the presence of the mimic.
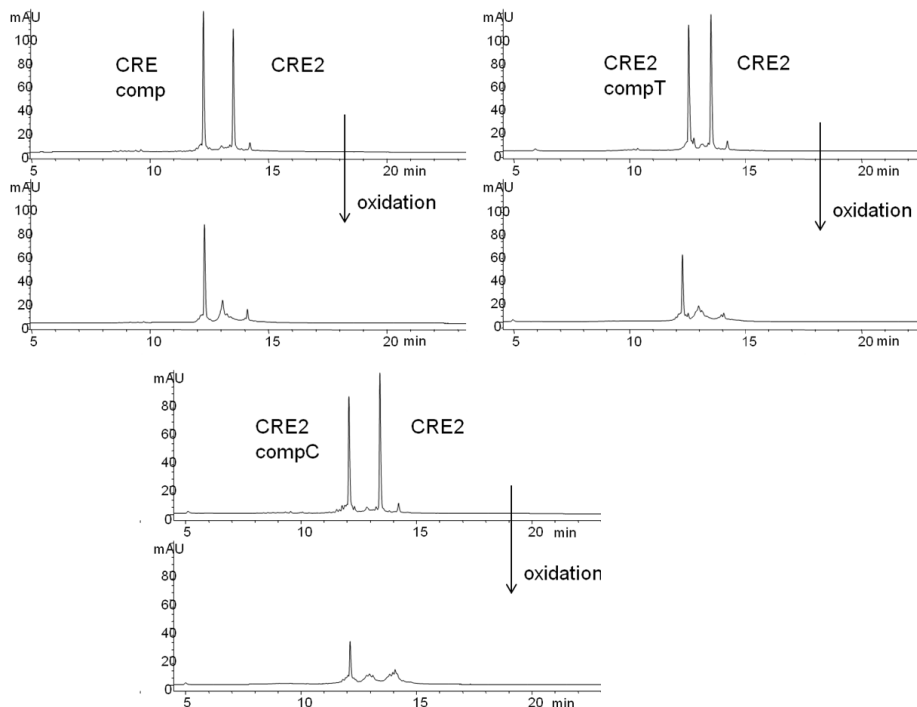
Figure 9.53: **RP-HPLC chromatograms of the cross-link experiments of CRE2** and its complements: CRE comp, CRE2 compT and CRE2 compC (left to right) in the absence of GCN4 mimic, before (top) and after (bottom) addition of NBS.

Cross-link experiment with variation of the opposite base in complementary ODN

CRE1, 2 and 3 were each combined additionally with two alternative strands, to allow positioning of a different base opposite to the furan-modified residue and obtain the base pairs **nu5**-C and **nu5**-T next to **nu5**-A. In case of ICL these variations would drastically change the cross-linking yield, as T has no nucleophilic amine to cross-link to. Cross-linking reactions and analysis were performed under identical conditions as described before. Chromatograms after oxidation are presented in Figures 9.54, 9.55 and 4.42, and show only minor changes in the yield. The observed variation in yield can be explained by the effect of the introduction of mismatches in different positions in the ODN on the

binding affinity of the GCN4 protein mimic and thus the DPC cross-linking.
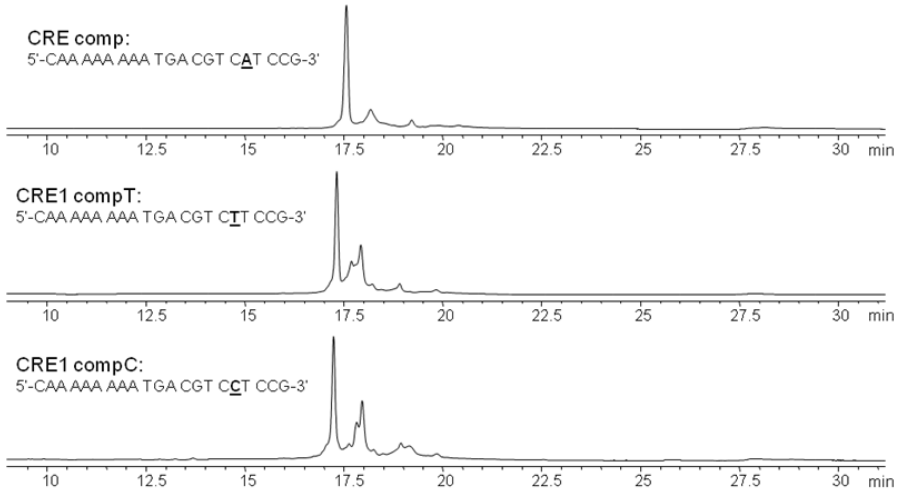
Figure 9.54: RP-HPLC chromatograms of the DPC reaction mixture with the duplex of CRE1 with the complementary ODN containing A, T and C opposite to the modified nucleotide, from top to bottom
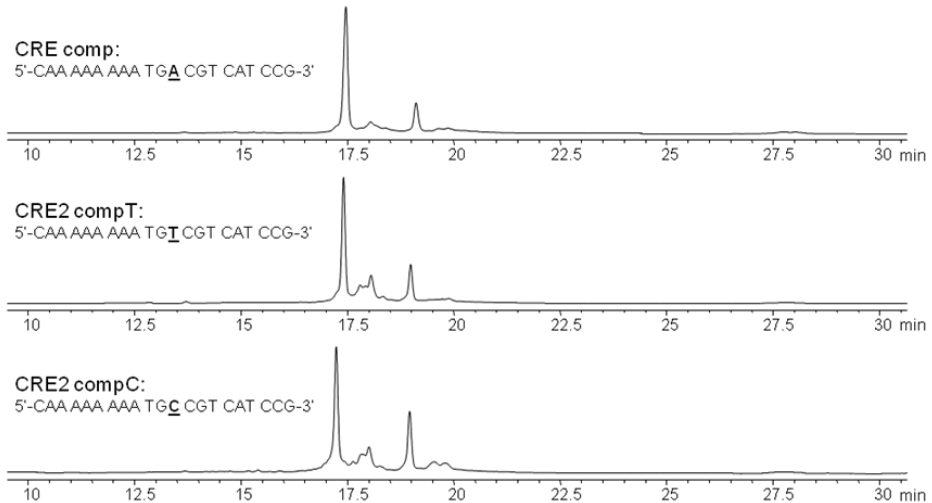
Figure 9.55: RP-HPLC chromatograms of the DPC reaction mixture with the duplex of CRE2 with the complementary ODN containing A, T and C opposite to the modified nucleotide, from top to bottom

DPC experiments in the absence of a peptide - oligonucleotide duplex complex

with a random peptide In an attempt to work with a model where complex formation is completely avoided (CD-GCN4 homodimers have been shown to bind the DNA similarly to Ad-CD heterodimers[72]), the cross-link reaction was repeated with 5 equivalents of another peptide of similar length and composition (Rev-peptide: DTRQARRNRRRRWRERQRAAAAR). The experiment was performed as before (ensuring similar reaction concentrations and obtain the best comparable results).

with the single stranded furan-modified CRE3 The experiment was performed as before, a volume of mQ water was added equal to the used volume of CRE3 comp, to assure the same final concentration for the reaction and obtain the best comparable results.

9.5.7 Control ICL with nu5-modified test ODNs

(performed by M. Op de Beeck)

The behavior of **nu5** in ICL formation was verified in a series of ICL experiments with **nu5**-modified test ODNs (Test G, Test T, Test C, Test A).

Synthesis of nu5-modified test ODNs

Similarly as in the previously reported experiments with furan-modified oligonucleotides[135, 138], the modified residue was incorporated in between two identical bases, to assess the cross-linking selectivity. Four different sequences were thus synthesized, varying the bases flanking the C5-modified uridine. All four ODN sequences (Test G, C, T and A) could easily be obtained in good yield and purity using the standard procedure for the manual incorporation of modified residues.

Analysis

- Test G (5'-CTG ACG G**nu5**G TGC-3') observed mass (deconvoluted) 3728.7 Da, calculated mass 3727.6 (Exact) and 3729.4 (MW) Da with formula $C_{120}H_{148}N_{45}O_{73}P_{11}$

- Test T (5'-CTG ACG T**nu5**T TGC-3') observed mass (deconvoluted) 3678.5 Da, calculated mass 3677.6 (Exact) and 3679.4 (MW) Da with formula $C_{120}H_{150}N_{39}O_{75}P_{11}$

- Test C (5'-CTG ACG C**nu5**C TGC-3') observed mass (deconvoluted) 3648.5 Da, calculated mass 3647.6 (Exact) and 3649.4 (MW) Da with formula $C_{118}H_{148}N_{41}O_{73}P_{11}$

- Test A (5'-CTG ACG A**nu5**A TGC-3') observed mass (deconvoluted) 3696.5 Da, calculated mass 3695.6 (Exact) and 3697.4 (MW) Da with formula $C_{120}H_{148}N_{45}O_{71}P_{11}$
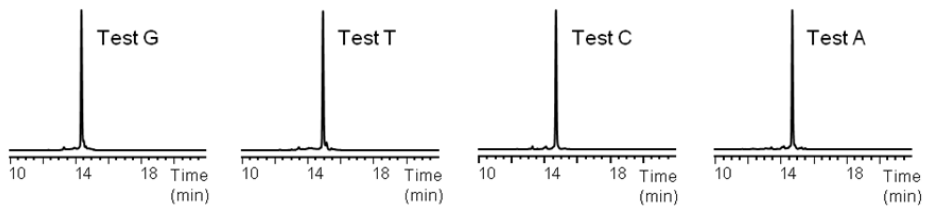
Figure 9.56: RP-HPLC chromatograms of the nu5-modified Test ODNs from left to right Test G, Test T, Test C and Test A
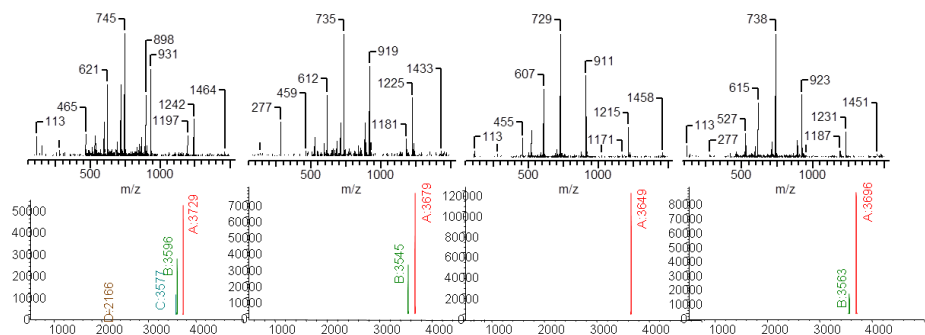
Figure 9.57: MS spectra (top) and the deconvolution (bottom) of the nu5-modified Test ODNs from left to right Test G, Test T, Test C and Test A

ICL with nu5-modified test ODNs

The obtained modified ODNs were each combined with four different commercial complementary ODNs where the base opposite to the modified residue, N was varied to be A, C, G or T:

Test G: 5'-CTG ACG G**nu5**G TGC-3' with 3'-GAC TGC C**N**C ACG-5'

Test T: 5'-CTG ACG T**nu5**T TGC-3' with 3'-GAC TGC A**N**A ACG-5'

Test C: 5'-CTG ACG C**nu5**C TGC-3' with 3'-GAC TGC G**N**G ACG-5'

Test A: 5'-CTG ACG A**nu5**A TGC-3' with 3'-GAC TGC T**N**T ACG-5'

Cross-linking reactions were performed using the standard ICL protocol. The resulting reaction mixtures were analyzed by RP-HPLC and the results shown in Figures 9.58- 9.61. Since the reaction yield in all cases is absent or low, this was not further investigated.
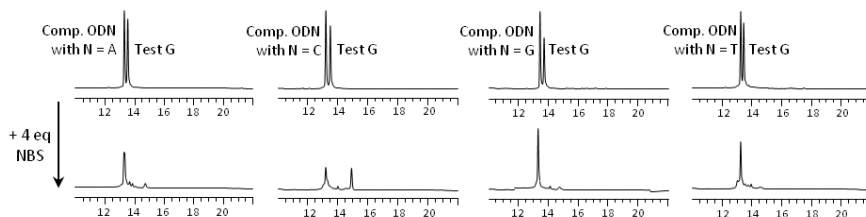
Figure 9.58: RP-HPLC chromatograms of ICL experiment of Test G and complementary ODN with N = A, C, G and T (left to right) before (top) and after (bottom) addition of NBS
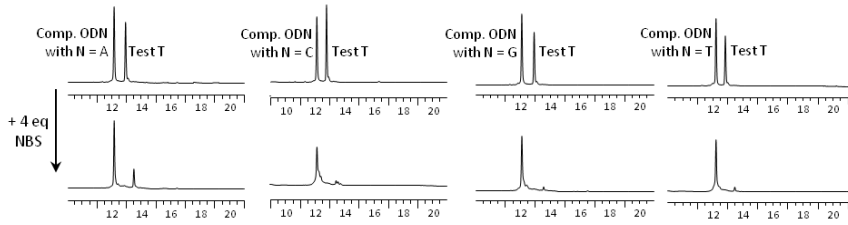
Figure 9.59: RP-HPLC chromatograms of ICL experiment of Test T and complementary ODN with N = A, C, G and T (left to right) before (top) and after (bottom) addition of NBS
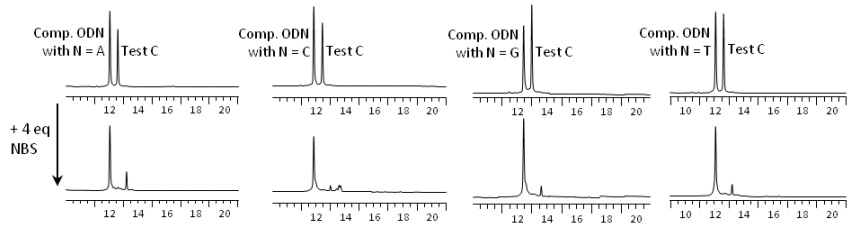
Figure 9.60: RP-HPLC chromatograms of ICL experiment of Test C and complementary ODN with N = A, C, G and T (left to right) before (top) and after (bottom) addition of NBS
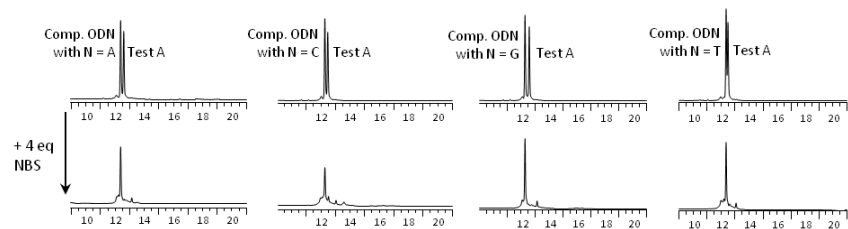
Figure 9.61: RP-HPLC chromatograms of ICL experiment of Test A and complementary ODN with N = A, C, G and T (left to right) before (top) and after (bottom) addition of NBS

9.6 Supporting Material for RNA Interstrand Cross-Linking

9.6.1 Synthesis of furan-modified nucleosides

nu3 (**4.1.3**) and **nu6** (**4.4.1**) were synthesized as described by our group before.[136, 138]

9.6.2 Synthesis of oligonucleotides (ONs) and oligodeoxynucleotides (ODNs)

ODNs were synthesized as described in the general protocol. For the synthesis of the ONs, 2'OMe phosphoramidites (Glen Research) were coupled using the same automated protocol, but with the coupling time increased to 6 min.

Analysis

- ODN4.4.2: 5'-CTG ACG G**nu3**G TGC-3' ($C_{123}H_{153}N_{46}O_{74}P_{11}$) observed mass: 3799 Da (Mass-H) (Maldi-TOF MS) and 3799.5 Da (deconvoluted) (ESI MS), calculated mass 3798.7 (exact) and 3800.5 (MW) Da
- ODN4.4.3: 5'-CTG ACG G**nu6**G TGC-3' ($C_{121}H_{150}N_{43}O_{71}P_{11}$) observed mass: 3683 Da (Mass-H) (Maldi-TOF MS) and 3682.5 Da (deconvoluted) (ESI MS), calculated mass 3681.7 (exact) and 3683.4 (MW) Da
- ON4.4.2: 5'-C*U*G* A*C*G* G***nu3**G* U*G*C*-3' (*2'OMe-RNA) ($C_{131}H_{169}N_{46}O_{84}P_{11}$) observed mass: 4071 Da (Mass-H) (Maldi-TOF MS) and 4071.8 Da (deconvoluted) (ESI MS), calculated mass 4070.7 (exact) and 4072.7 (MW) Da
- ON4.4.3: 5'-C*U*G* A*C*G* G***nu6**G* U*G*C*-3' (*2'OMe-RNA) ($C_{129}H_{166}N_{43}O_{81}P_{11}$) observed mass: 3955 Da (Mass-H) (Maldi-TOF MS) and 3954.8 Da (deconvoluted) (ESI MS), calculated mass 3953.7 (exact) and 3955.6 (MW) Da
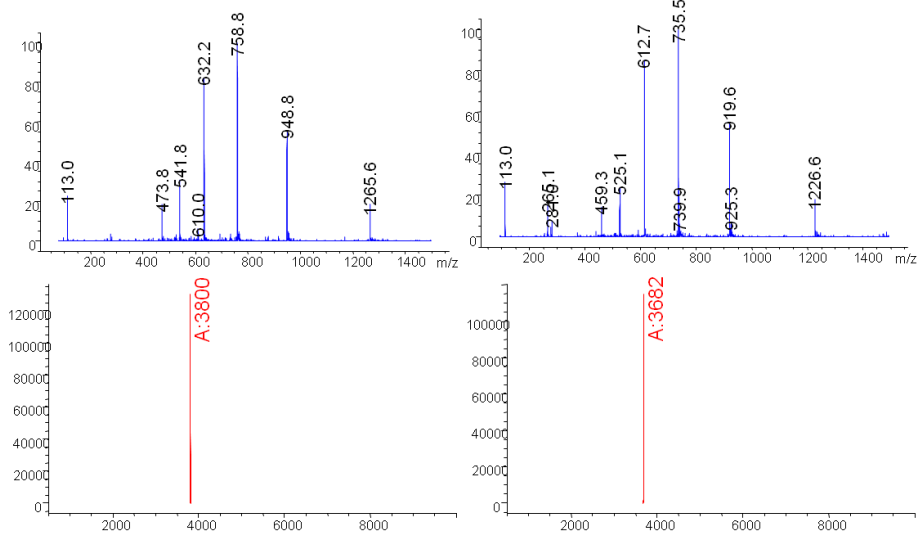
Figure 9.62: ESI-MS spectra of ODN4.2.2 (left) and ODN4.4.3 (right)
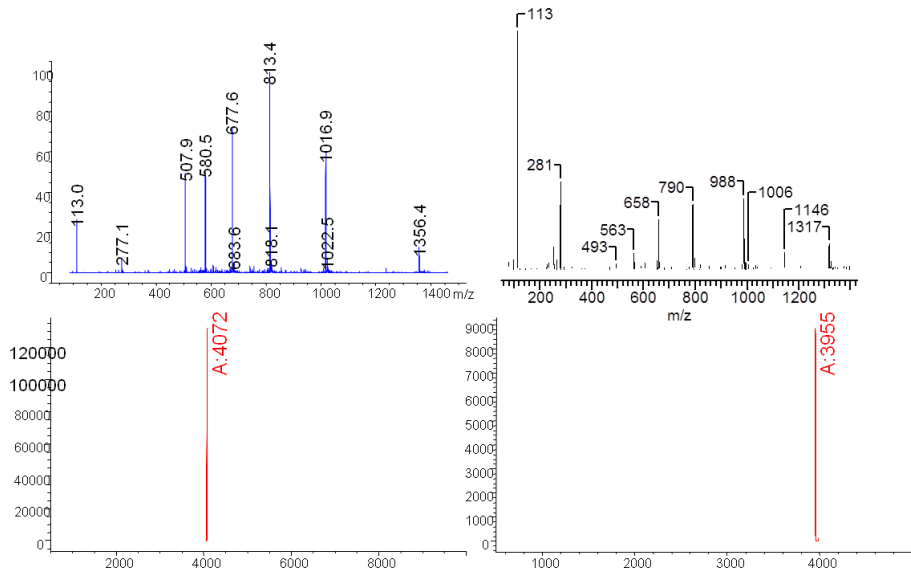
Figure 9.63: ESI-MS spectra of ON4.2.2 (left) and ON4.4.3 (right)
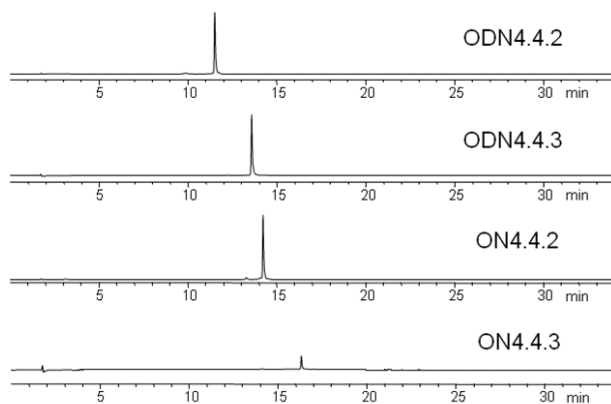
9.6.3 Cross-linking

The cross-linking reactions were performed as described in the general protocol, but both at 25°C and 37°.

Analysis

- ON4.4.2-ON4 ($C_{244}H_{313}N_{91}O_{166}P_{22}$) observed mass: 7855 Da (Mass-H) (Maldi-TOF MS); calculated mass 7858 Da
- ON4.4.3-ON4 ($C_{242}H_{310}N_{88}O_{163}P_{22}$) observed mass: 7777 Da (Mass-2H+K) (Maldi-TOF MS); calculated mass 7741 Da
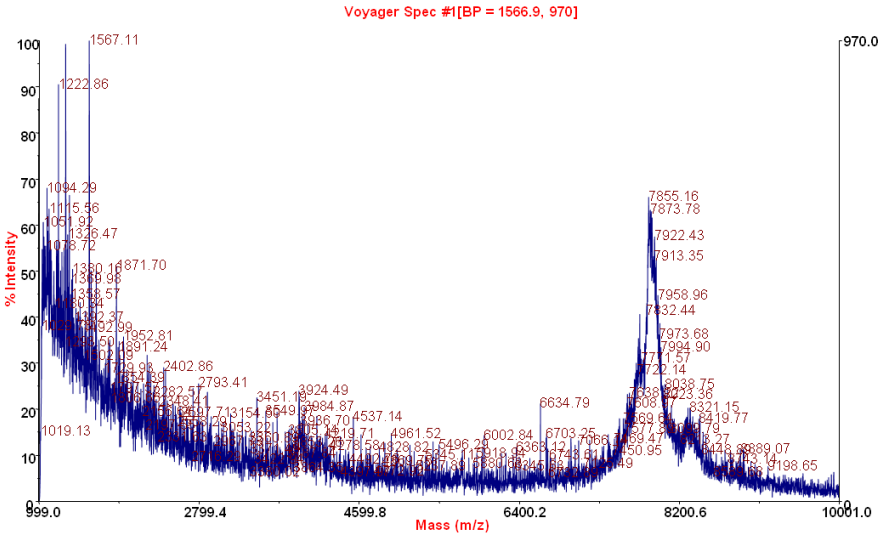
Figure 9.64: Maldi Mass spectrum of ICL: ON4.4.2-ON4
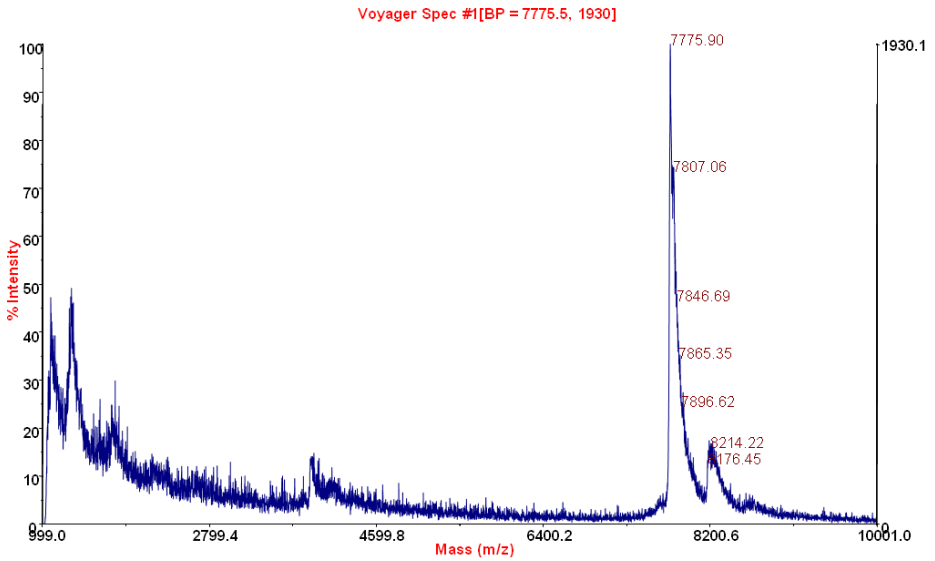
Figure 9.65: Maldi Mass spectrum of ICL: ON4.4.3-ON4

Cross-linking of ON4.4.2 (PAGE) .
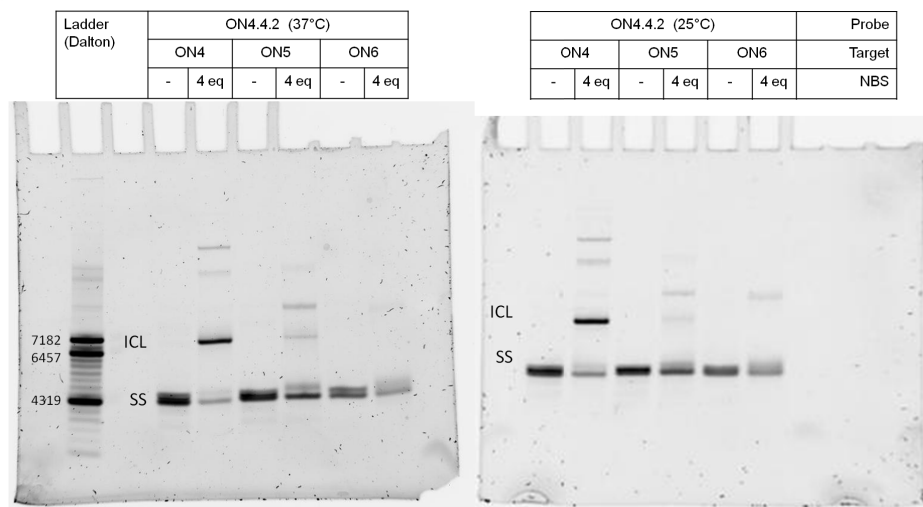
Figure 9.66: PAGE of cross-link reaction of ON4.4.2 and RNA targets at 37°C (left) and 25°C (right). Ladder consist of a mixture of 4 ODNs with masses 7182, 6457, 4319, 2096 Da

Cross-linking of ON4.4.2 to ON4 (RP-HPLC) .
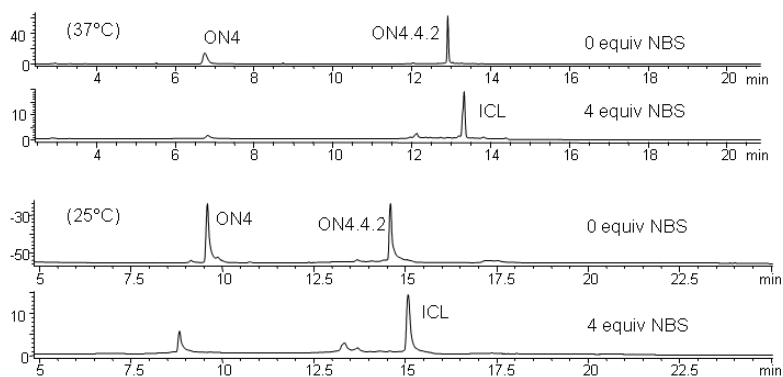
Figure 9.67: RP-HPLC chromatograms of cross-link reaction of ON4.4.2 with ON4 as target at 37°C (top) and 25°C (bottom).
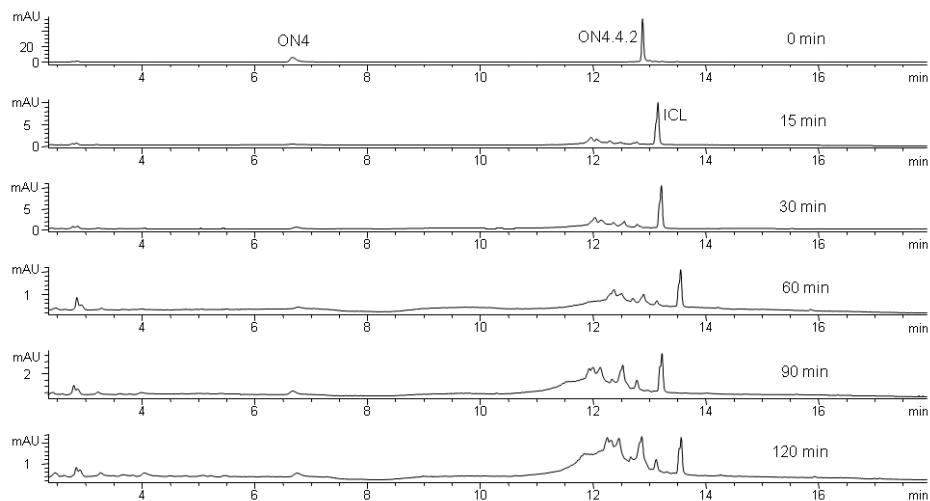
Figure 9.68: RP-HPLC chromatograms of cross-link reaction of ON4.4.2 with ON4 using singlet oxygen.

Cross-linking of ON4.4.2 to ON5 (RP-HPLC)   .
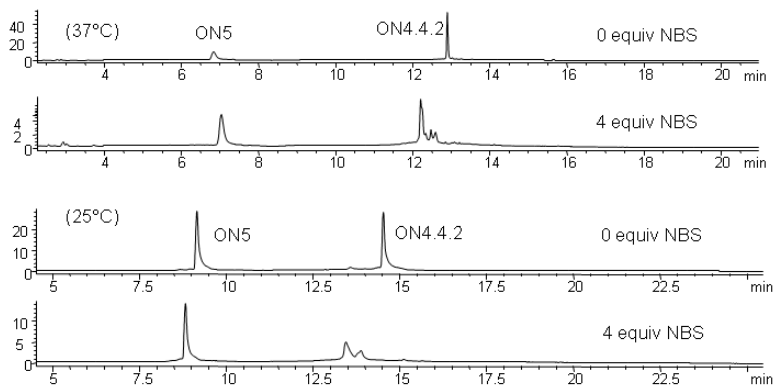
Figure 9.69: RP-HPLC chromatograms of cross-link reaction of ON4.4.2 with ON5 as target at 37°C (top) and 25°C (bottom).

Cross-linking of ON4.4.2 to ON6 (RP-HPLC)   .
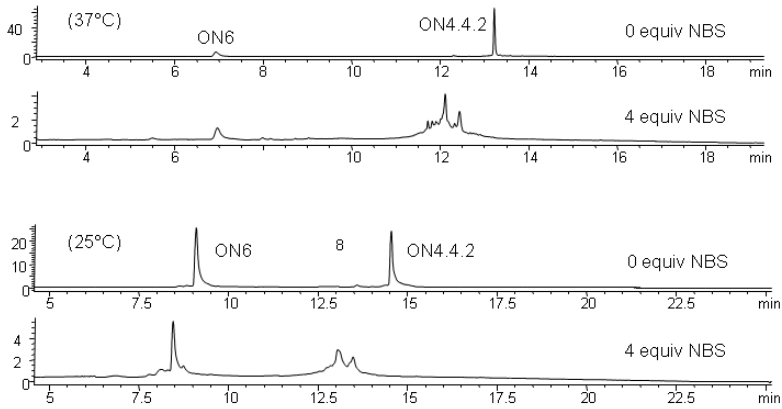
Figure 9.70: RP-HPLC chromatograms of cross-link reaction of ON1 with ON6 as target at 37°C (top) and 25°C (bottom).

Cross-linking of ON4.4.3 (PAGE)   .
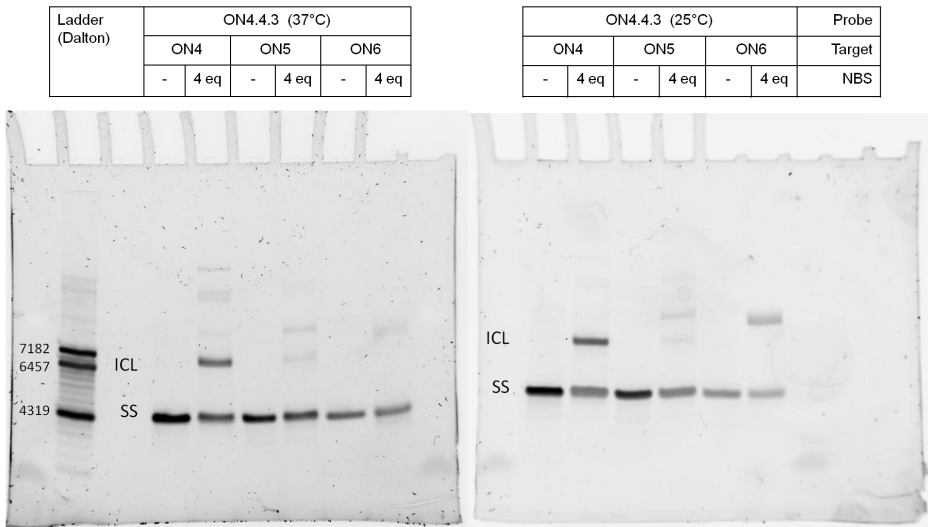
Figure 9.71: PAGE of cross-link reaction of ON4.4.3 and RNA targets at 37°C (left) and 25°C (right). Ladder consist of a mixture of 4 ODNs with masses 7182, 6457, 4319, 2096 Da

Cross-linking of ON4.4.3 to ON4 (RP-HPLC)   .
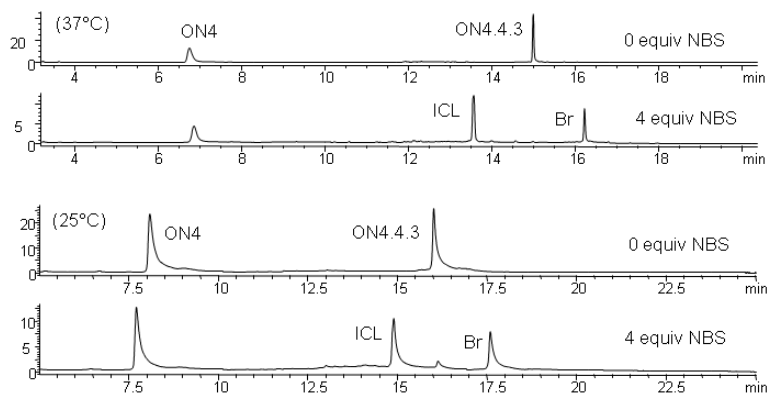
Figure 9.72: RP-HPLC chromatograms of cross-link reaction of ON4.4.3 with ON4 as target at 37°C (top) and 25°C (bottom).
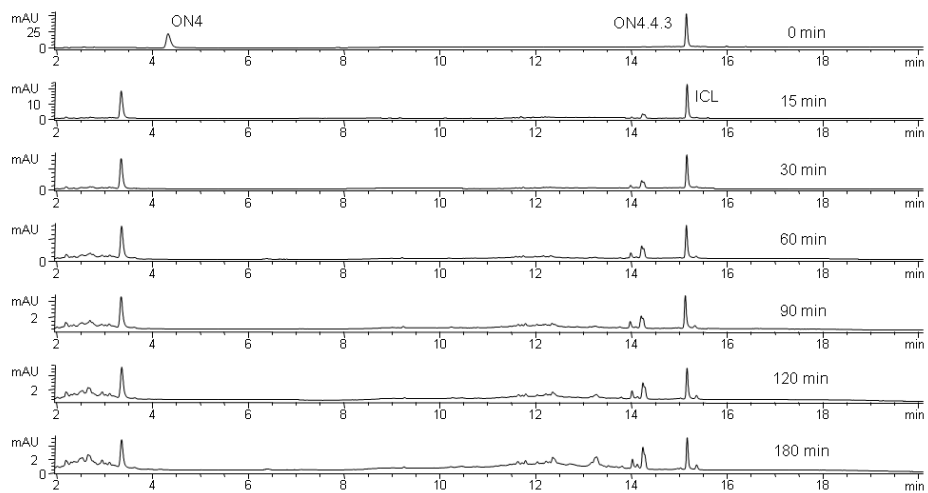
Figure 9.73: RP-HPLC chromatograms of cross-link reaction of ON4.4.3 with ON4 using singlet oxygen.

Cross-linking of ON4.4.3 to ON5 (RP-HPLC) .
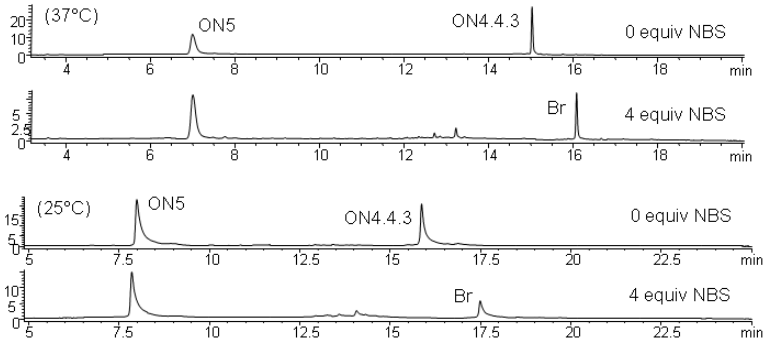
Figure 9.74: RP-HPLC chromatograms of cross-link reaction of ON4.4.3 with ON5 as target at 37°C (top) and 25°C (bottom).

Cross-linking of ON4.4.3 to ON6 (RP-HPLC) .
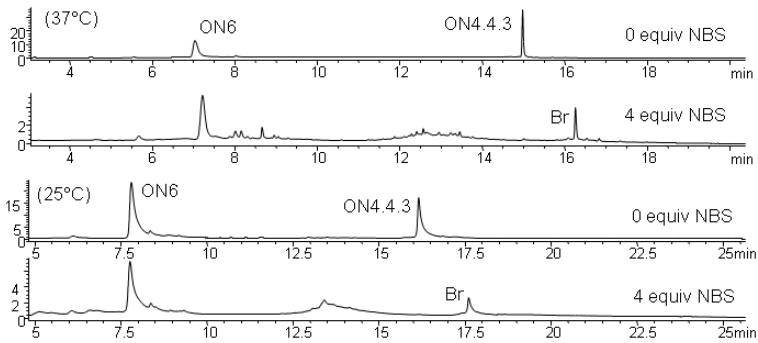
Figure 9.75: RP-HPLC chromatograms of cross-link reaction of ON4.4.3 with ON6 as target at 37°C (top) and 25°C (bottom).

9.6.4   Additional experiments

Initial screen

A range of furan-modified nucleosides has been developed and or tested by our group for DNA ICL formation. The furan moiety has been used to replace the sugar of the nucleoside or has been attached to it, on the 1' or 2' position, through different linkers (amide, urea and ether) and combined with different bases (uracil and adenine). The furan moiety can also be attached to the base.

Because an RNA-RNA duplex is structurally different from a DNA-DNA duplex, the properties of the different furan-modified building blocks cannot be assumed to be simply transferable. For the selection of the 2 most promising furan-modified nucleosides to be used in this study, an initial screen was performed with 5 furan-modified nucleosides, incorporated in the DNA sequence 5'-CTG ACG G**X**G TGC-3' (available from previous experiments) and targeting ON4 (5'-GCA CCC CGU CAG-3'). The combination of DNA and RNA in a helix forms an intermediate duplex between A and B and can give an indication of the behavior in pure RNA duplex.

The cross-link reaction was performed at 0°C (initial mild conditions) by addition of 4 equiv of NBS as described before. Figure 4.56 shows the RP-HPLC chromatograms before and after oxidation with NBS. The 2nd and 5th furan-modified nucleosides were identified as the most promising and therefore selected for further study in this context.

Cross-link tests from DNA to the RNA targets

Cross-linking of ODN4.4.2 (PAGE) .
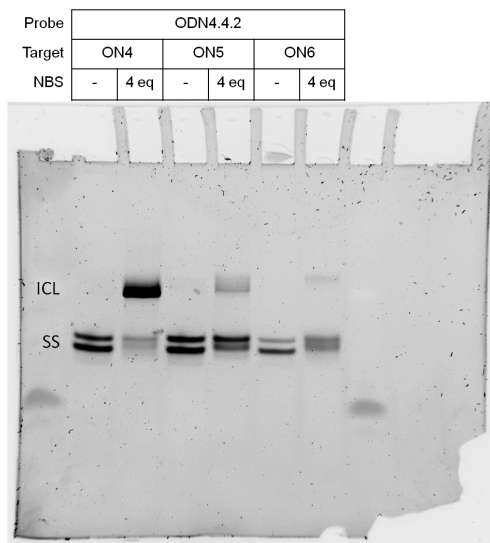
Figure 9.76: PAGE of cross-link reaction of ODN4.4.2 and RNA targets at 25°C.

Cross-linking of ODN4.4.2 to ON4 (RP-HPLC) .
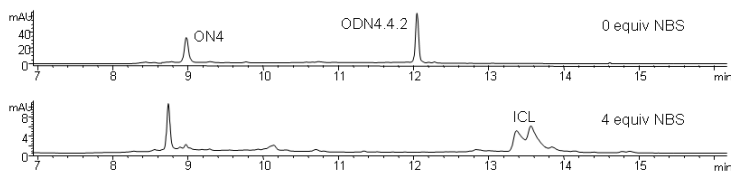
Figure 9.77: RP-HPLC chromatograms of cross-link reaction of ODN4.4.2 with ON4 as target at 25°C. ICL indicates the cross-linked species, which consists of 2 pairs of pseudo enantiomers

Cross-linking of ODN4.4.2 to ON5 (RP-HPLC)   .
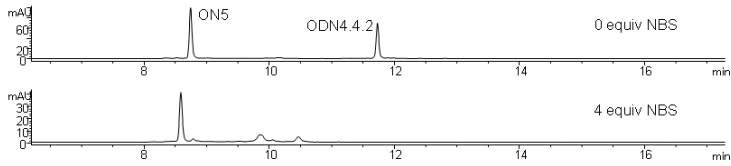
Figure 9.78: RP-HPLC chromatograms of cross-link reaction of ODN4.4.2 with ON5 as target at 25°C. Only remaining RNA2 can be observed after NBS treatment.

Cross-linking of ODN4.4.2 to ON6 (RP-HPLC)   .
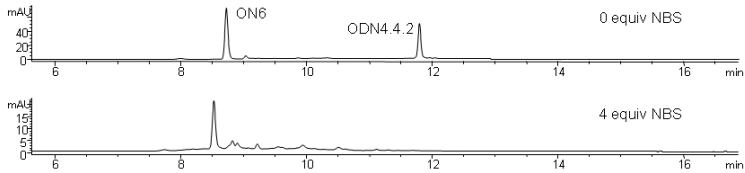
Figure 9.79: RP-HPLC chromatograms of cross-link reaction of ODN4.4.2 with ON6 as target at 25°C. Only remaining RNA2 can be observed after NBS treatment.

Cross-linking of ODN4.4.3 (PAGE)   .
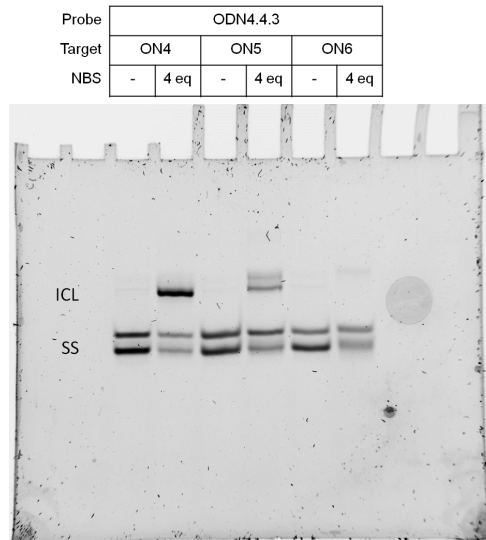
Figure 9.80: PAGE of cross-link reaction of ODN4.4.3 and RNA targets at 25°C.

Cross-linking of ODN4.4.3 to ON4 (RP-HPLC) .
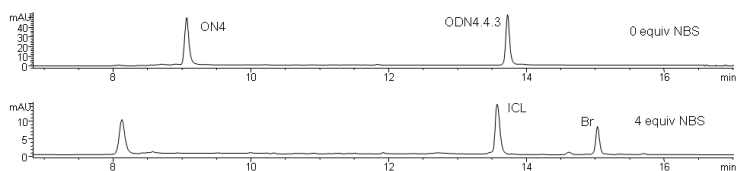
Figure 9.81: RP-HPLC chromatograms of cross-link reaction of ODN4.4.3 with ON4 as target at 25°C. ICL indicates the cross-linked species and Br indicates the brominated side product

Cross-linking of ODN4.4.3 to ON5 (RP-HPLC) .
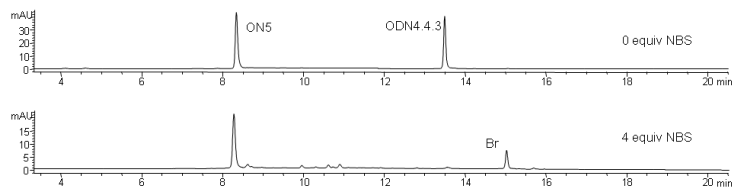
Figure 9.82: RP-HPLC chromatograms of cross-link reaction of ODN4.4.3 with ON5 as target at 25°C. Br indicates the brominated side product, none of the other visible peaks could be correlated with a cross-linked species, but correspond to degradation products.

Cross-linking of ODN4.4.3 to ON6 (RP-HPLC) .
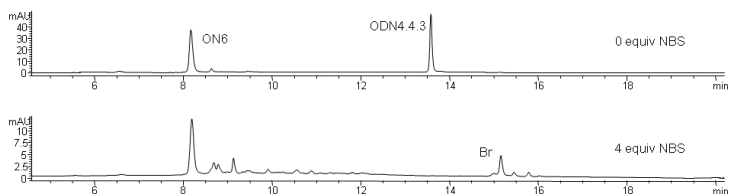
Figure 9.83: RP-HPLC chromatograms of cross-link reaction of ODN4.4.3 with ON6 as target at 25°C. Br indicates the brominated side product, none of the other visible peaks could be correlated with a cross-linked species, but correspond to degradation products.

DNA ICL tests

Cross-linking of ODN4.4.2 to ODN4 (RP-HPLC)   .
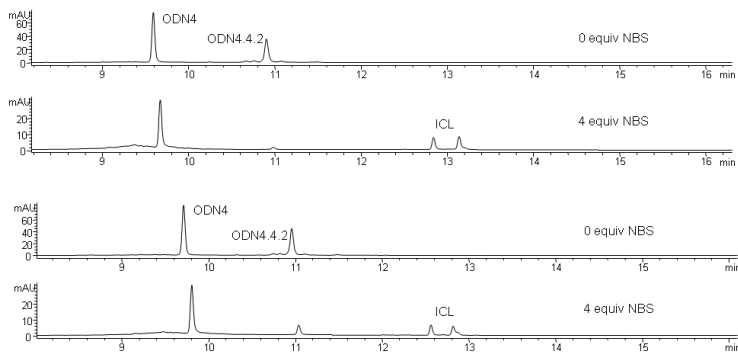
Figure 9.84: RP-HPLC chromatograms of cross-link reaction of ODN4.4.2 with ODN4 as target at 37°C (top) and 25°C (bottom). ICL indicates the cross-linked species, which consists of 2 pairs of pseudo enantiomers

Cross-linking of ODN4.4.3 to ODN4 (RP-HPLC)   .
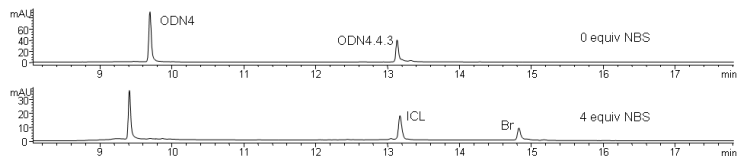
Figure 9.85: RP-HPLC chromatograms of cross-link reaction of ODN4.4.3 with ODN4 as target at 37°C. ICL indicates the cross-linked species and Br indicates the brominated side product

9.7 Supporting Material for Furan-Modified Proteins

9.7.1 Synthesis of furan-modified amino acids

Synthesis of aa4 (4.5.1)

For previous steps, see section DNA ← protein mimic cross-linking.
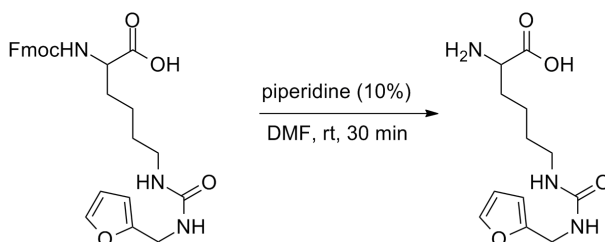
Step 3: Fmoc deprotection A solution of 10% piperidine in DMF was added to Fmoc protected **aa4** (**4.2.6**) and stirred for 30 min. The product (**4.5.1**) fell out of the reaction mixture as a white powder and was purified by fitration (Yield 95%).

Analysis

- Rf product: 0 (eluent MeOH/CHCl$_3$ 1/6 + 0.1% acetic acid)
- Chemical Formula: C$_{12}$H$_{19}$N$_3$O$_4$
- Mass: 269.1 (Exact), 269.3 (MW)
- LC-MS: 8.8 min (r.t.), 270.1 [M+H]$^+$ (observed mass)
- ^1H-NMR (300 MHz, D$_2$O): 7.42 (1H, d, 1 Hz), 6.38 (1H, dd, 1/3 Hz), 6.23 (1H, d, 3 Hz), 4.24 (2H, s), 3.68 (1H, t, 6 Hz), 3.09 (2H, t, 6.7 Hz), 1.82 (2H, m), 1.48 (2H, m), 1.35 (2H, m)
- ^{13}C-NMR (125 MHz, D$_2$O: 175 (C), 160 (C), 153 (C), 142 (CH), 110 (CH), 106 (CH), 55 (CH), 39 (CH$_2$), 37 (CH$_2$) 30 (CH$_2$), 29 (CH$_2$), 22 (CH$_2$)

Synthesis of aa6

Step 1: Boc deprotecion   Was described before, see section DNA ← protein mimic cross-linking.

Activation of furoic acid   .

(4.5.3)
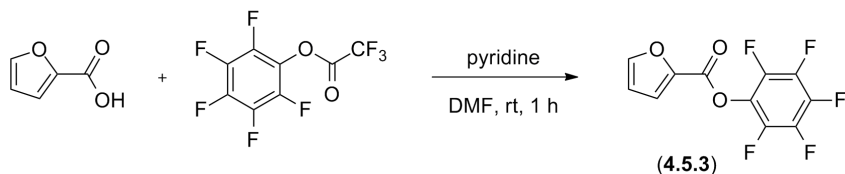
Furoic acid (1 equiv, 1.5 g, 13.4 mmol, 112.08 g/mol), was dissolved in DMF (25 ml), then pentafluoraphenyl trifluoroacetate (1.2 equiv, 2.68 ml, 15.5 mmol, 280.08 g/mol, 1.63 g/ml) and pyridine (1.1 equiv, 1.19 ml, 14.7 mmol, 79.10 g/mol, 0.98 g/ml) were added. The mixture was stirred at rt for 2 h. The reaction mixture was diluted in EtOAc (500 ml), washed with 0.05 M HCl and 5% NaHCO$_3$, dried over MgSO$_4$ and concentrated under reduced pressure. White crystals (**4.5.3**) were obtained in 99% yield.

   Analysis

- Rf product: 0.71 (eluent DCM)
- Chemical Formula: $C_{11}H_3F_5O_3$
- Mass: 278.0 (Exact), 278.1 (MW)
- LC-MS: 18.22 min (r.t.), 279.1 $[M+H]^+$ (observed mass)

Step 2: coupling reaction   .

(4.2.4)   **(4.5.1)**
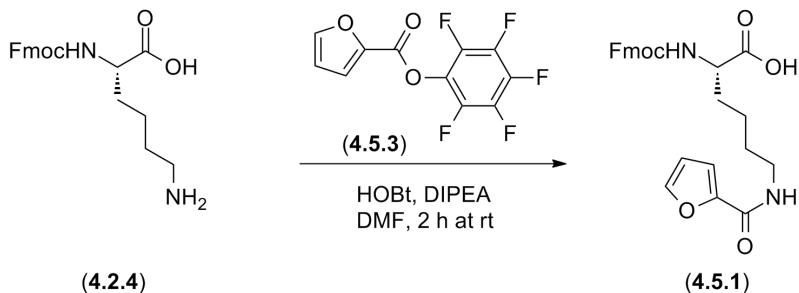
The obtained Boc deprotected lysine (**4.2.7**) (1 equiv, 10.7 mmol) was combined with the obtained activated furoic acid (**4.5.3**) (1.2 equiv, 12.8 mmol, 3.56 g, 278 g/mol), together with HOBt (1.2 equiv, 12.8 mmol, 1.73 g, 135 g/mol) and DIPEA (1.2 equiv, 12.8 mmol, 2.12 ml, 129.25 g/mol, 0.78 g/ml) in DMF (35 ml). The coupling was performed at rt for 2 h. The reaction mixture was then concentrated under reduced pressure and the product (**4.5.4**) purified using column chromatography (CHCl$_3$/MeOH 12/1 with acetic acid 0.46%).

Analysis

- Rf product: 0.45 (eluent MeOH/CHCl$_3$ 1/6 + 0.1% acetic acid)
- Chemical Formula: C$_{26}$H$_{26}$N$_2$O$_6$
- Mass: 462.2 (Exact), 462.5 (MW)
- LC-MS: 12.1-13 min (r.t.), 463.1 [M+H]$^+$ (observed mass)

Step 2: Fmoc deprotection .
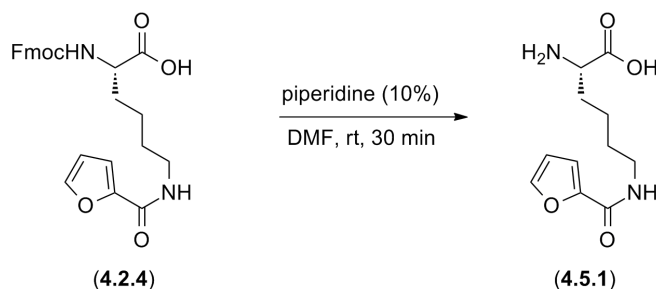
A solution of 10% piperidine in DMF was directly added to the product optained in the previous step (**4.5.4**) and stirred for 30 min. The product (**4.5.2**) falls out of the reaction mixture as a white powder and is purified by fitration (unoptimized overall yield 83%).

Analysis

- Rf product: 0 (eluent MeOH/CHCl$_3$ 1/6 + 0.1% acetic acid)
- Chemical Formula: C$_{11}$H$_{16}$N$_2$O$_4$
- Mass: 240.1 (Exact), 240.3 (MW)

- ESI MS: 239.3 [M-H]$^-$, 241.3 [M+H]$^+$ (observed mass)
- ^1H-NMR (500 MHz, D$_2$O: 7.6 (1H, d, 1 Hz), 7.1 (1H, d, 3 Hz), 6.6 (1H, dd, 3/1 Hz), 3.7 (1H, t, 6 Hz), 3.3 (2H, t, 7 Hz), 1.9-1.8 (2H, m), 1.6 (2H, m), 1.5-1.3 (2H, m)
- ^{13}C-NMR (125 MHz, D$_2$O: 175 (C), 161 (C), 146 (C), 145 (CH), 115 (CH), 113 (CH), 55 (CH), 39 (CH$_2$), 30 (CH$_2$), 28 (CH$_2$), 22 (CH$_2$), 21 (CH$_2$)

9.7.2 Biotechnology protocols

It is important to work sterile. Therefore the working space and hands are cleaned with 80% EtOH. All experiments should be performed in the airflow of a bunsen-burner flame.

LB medium

Tryptone (10 g/l, Roth), yeast extract (5 g/l, Roth), NaCl (10 g/l) were dissolved in distilled water (the pH should be 7.5, but it is typically 7 and can be used as such) and autoclaved.

Agar plates

Agar-agar (15 g/l) wass added to LB medium and dissolved by heating in a microwave. Before pouring in the petridishes it was cooled down till 56° in a water bath and the appropriate antibiotics (1/1000) were added. Then this was poured into a petridish, to just cover the bottom and cooled down for 2 h.

Transfection

1 μl of plasmid (pET DUET 1) (or 2 μl of a 2 plasmid system), was added to T7 express competent cells (E. coli, New England Biolabs) on ice and left for 30 min before a heat shock of 45 s at 42°C. LB medium (0.5 ml) was added and the suspension was shaken at 37°C for 1 h. Half of the cells was plated on agar with the appropriate antibiotics, for storage. The other half was added to LB medium (50 ml) with the appropriate antibiotics (50 μl of a 1000x stock). Both were incubated overnight at 37°C.

Inocculation and induction of YFP expression

100 μl (1/50) of an overnight culture was added to LB medium (5 ml) with (1/1000) of the appropriate antibiotics (5 μl). The amino acid can be added to obtain different concentrations. The cells are incubated at 37°C, 200 rpm till an optical density (OD) of 0.6 is reached. Then the expression is induced with isopropylthio-β-galactoside (IPTG) (1/1000) (1 mM, 5 μl). After 4 h the cells (10 μl) were monitored on a TCS SPE spectral confocal microscope (Leica) with an inverse stand and the HCX FL APO 63x/140-0.60 oil objective (excitation: 488 nm).

Expression of TrxA

An overnight culture (50 ml from a swipe of a plate) of T7 express chemical competent cells transfected with a plasmid for TrxA with an Amber mutation at N65 and carbenicillin resistance and a constitutive plasmid with the pyrrolysyl-tRNA synthetase and its tRNA$_{CUA}$ and chloramphenicol resistance (OD 2.5) was inoculated (1/50) two times in half a litre of LB medium with carbenicillin, chloramphenicol and 5 mM of **aa6** and was grown till OD 0.6 (over 9 hours). Then IPTG was added to induce the expression at 18°C overnight (16 h).

TrxA protein purification

All purification steps were carried out at 4°C. The media were centrifuged at 4°C and 8000 rpm for 15 min and decanted to harvest the cells. Cells were resuspended in StrepA buffer (IBA) (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 150 mM NaCl, 30 ml) supplemented with protease inhibitor mix (Roche). The cells were lysed with a French Press and centrifuged at 4°C and 18 000 rpm for 40 min. The supernatant was purified over a Strep-Tactin-Superflow column (IBA) on KTA (Amersham Biosciences) equilibrated with StrepA buffer (1 ml/min, Flow through, 7 colom volumes wash) in 5 ml fractions during elution with 100% Strep B buffer (0.56 g/l d-desthiobiotin in Strep A buffer) The protein in fraction 2 and fraction 3+4, was concentrated using an (Amicon Ultra 10K) centrifugal tube in 0.1 M phosphate buffer pH 8 with 0.1 M NaCl. The concentration was determined with a nanodrop (2 = 6.13 mg/ml and 3+4 = 6.48 mg/ml), but was inaccurate

due to DNA contamination (260/280 for 2 =1.49 and 3+4 = 1.7). Therefore a
Bradford Assay was performed with Protein Assay solution (Bio-Rad). This way
the concentration of 2 and of 3+4 was determined to be 1.56 g/L.

SDS-PAGE

The analysis of the proteins was performed on Laemmli gels. These consist of
the main separating gel of 12% acrylamide, 0.375 M Tris pH 8.8 and a small
collecting gel on top of 5% acrylamide 0.1 M Tris pH 6.8. The loading buffer
consisted of 0.045 M Tris-HCl, pH 6.8, 10% glycerol, 1% sodiumdodecylsulphate,
0.01% bromophenolblue a,d 0.05 M dithiotreitol. The running buffer was 25 mM
Tris, 192 mM glycin and 0.1% sodiumdodecylsulphate. The gel was run at 150 V
for about 50 min. The running time was derived with SeeBluePlus2 pre-stained
standard. The gels were washed 3 times in 100 ml distilled water before staining.
Staining was achieved with 20 ml Page Blue protein staining solution (Thermo
Scientific) for 60 min (or overnight) with gentle agitation and 5 min rinsing in
distilled water.

In Gel Trypsin digest

The desired bands were cut from the Coomassie Blue stained SDS-gel that had
been washed in water (2x, 10 min), on a sterile underground and further cut into
pieces of approximately 1 mm^2. The obtained fragments were put in a centrifugal
tube and washed by incubation in 100 μl water for 15 min and shaking at 550
rpm at rt. The water was taken off and replaced with 100 μl of a 1/1 solution of
50 μM NH_4HCO_3 (200 mg in 50 ml)/MeCN. This mixture was shaken for 5 to
10 min at 550 rpm at rt. This was then replaced by 100 μl MeCN and shaken
for 10 min. The gel was rehydrated by 100 μl 50 mM NH_4HCO_3 and shaken for
10 min. To this mixture 100 μl MeCN was added and shaken for 10 min. This
was then replaced by 100 μl MeCN to shrink the gel pieces and shaken for 10
min, before it was removed and the sample dried in the speedvac. The protein in
the gel was then reduced and alkylated by adding 100 μl of a 1/1 solution of 10
mM dithiotreitol (7.7 mg in 5 ml)/50 mM NH_4HCO_3, this was heated till 56° for
45 min. After cooling down, the solution was replaced by a freshly prepared 55
mM iodoacetamide (51 mg in 5 ml 50 mM NH_4HCO_3) solution. The mixture was

shielded from light and shaken for 30 min at rt. The solution was replaced by 100 μl of a 1/1 solution of 50 μM NH_4HCO_3/MeCN and shaken at rt for 15 min. This last step was repeated once more. Finally the solution was replaced by 100 μl MeCN and shaken for 15 min, before it was removed and the sample dried in the speedvac. The protein was digested by adding 100 μl freshly prepared enzyme (90 μl 25 mM NH_4HCO_3 with 10 μl enzyme trypsin or chymotrypsin (Promega) (10 ng/μl in 25 mM NH_4HCO_3)) and overnight incubation (without shaking) at 37°. Afterwards, the peptide fragments could be extracted from the gel. The digest solution was collected in a new centrifugal tube. 100 μl 25 mM NH_4HCO_3 was added to the gel rests and sonificated for 15 min. Then 100 ml MeCN was added and again sonificated for 15 min. The 200 μl solution was taken out and added to the centrifugal tube with the digest solution. 100 μl 5% formic acid was added to the gel rests and sonificated for 15 min. Then 100 ml MeCN was added and again sonificated for 15 min. The 200 μl solution was taken out and added to the centrifugal tube with the digest solution. The combined solution was evaporated till dryness in the speedvac and then redissolved in 25 μl of 0.1% formic acid for analysis.

Mass spectrometry

Mass spectra were recorded on the following machines: Thermo Finnigan LTQ-FT (ESI-ICR) and Thermo LTQ-Orbitrap XL. For analytical HPLC separations of protein and peptide samples with subsequent MS a Dionex Ultimate 3000 Nano HPLC was used. The purified intact protein was analyzed by HPLC-MS. A self-packed nano C4 column (ReproSil Gold 300 Å, 3μM) was used for HPLC and mass spectrometry was performed on a LTQ-Orbitrap XL mass spectrometer (Thermo). For the nano-HPLC-ESI-MS analysis the tryptic digested peptides were loaded onto a Dionex C18 Nano Trap Column (100 μm), subsequently eluted and separated by the self packed C18-AQ (Reprosil Pur, 3 μm, 120) column. The obtained peptide fragment data was analyzed with the SEQUEST algorithm implemented in the software 'Xcalibur bioworks' against the TrxA sequence.

9.7.3 Additional Data

Confocal fluorescent images with the wildtype synthetase

.
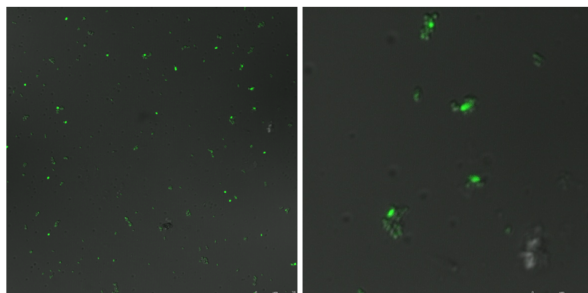
Figure 9.86: Wild type synthetase with an alkyn-modified amino acid (10 mM), for reference (with zoom on the right)
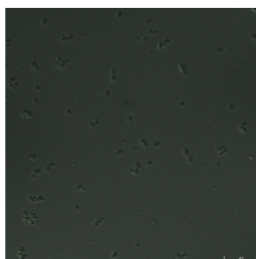
Figure 9.87: Wild type synthetase with **aa4** (5 mM)
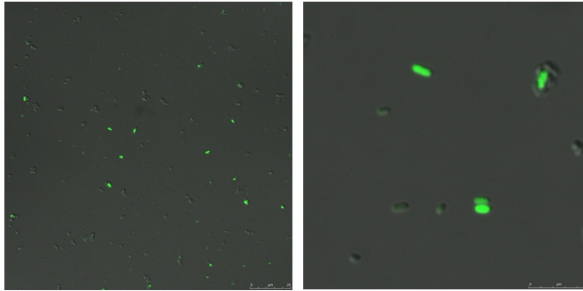
Figure 9.88: Wild type synthetase with **aa6** (5 mM) (with zoom on the right)

Confocal fluorescent images with the mutant synthetase of Yokoyama

.
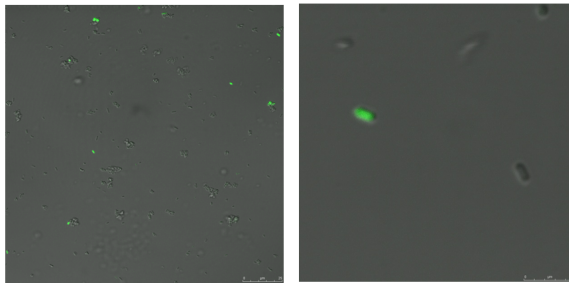
Figure 9.89: Mutant synthetase of Yokoyama with **aa4** (10 mM) (with zoom on the right)
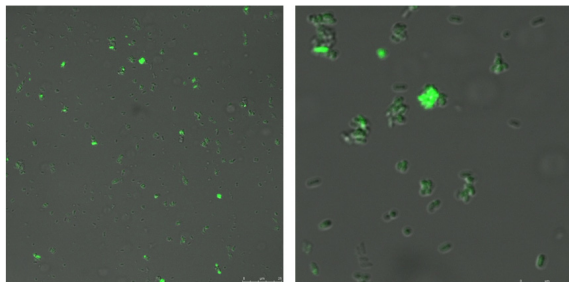
Figure 9.90: Mutant synthetase of Yokoyama with **aa6** (10 mM) (with zoom on the right)
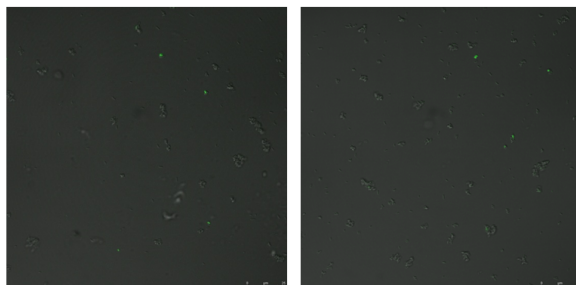
Figure 9.91: Mutant synthetase of Yokoyama with **aa4** (left) and **aa6** (right) (5 mM)

Confocal fluorescent images with the mutant synthetase of Emine

.
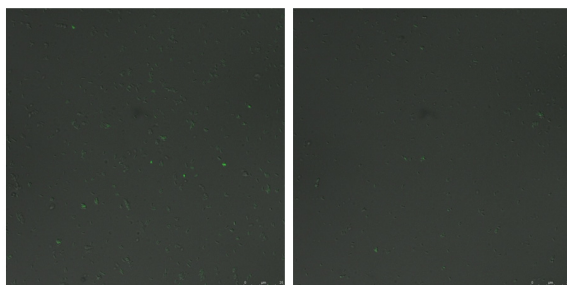
Figure 9.92: Mutant synthetase of Emine with **aa4** (left) and **aa6** (right) (10 mM)

Confocal fluorescent images with the Dual plasmid system, with wild type synthetase and aa6

.
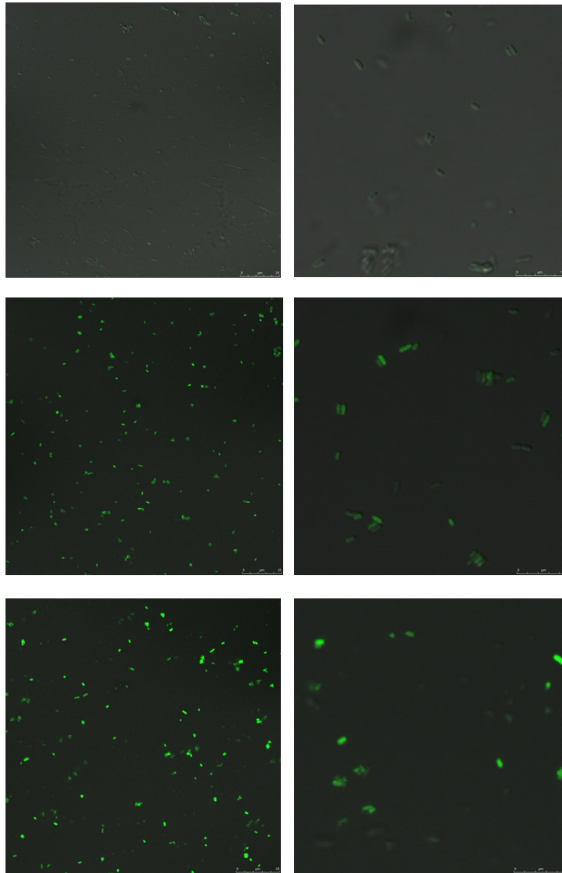
Figure 9.93: Wild type synthetase in the dual plasmid system with **aa6**, 0 mM or blanc (top), 2 mM (middle) and 5 mM (bottom) using maximum laser intensity (with zoom on the right)
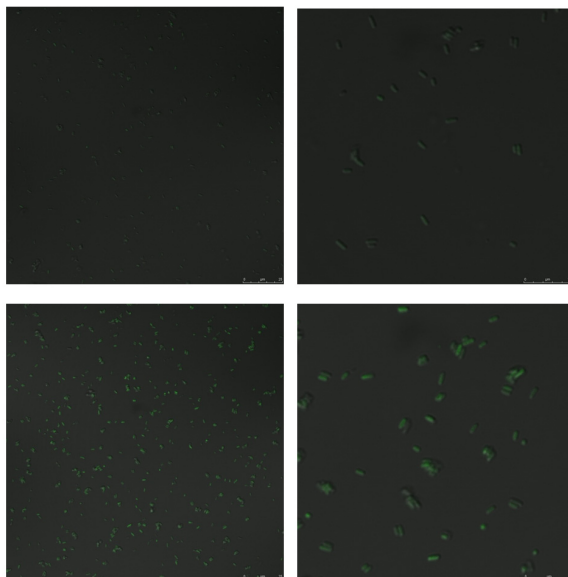
Figure 9.94: Wild type synthetase in the dual plasmid system with **aa6**, 2 mM (top) and 5 mM (bottom) using a lower laser intensity to see the difference (with zoom on the right)

SDS purification
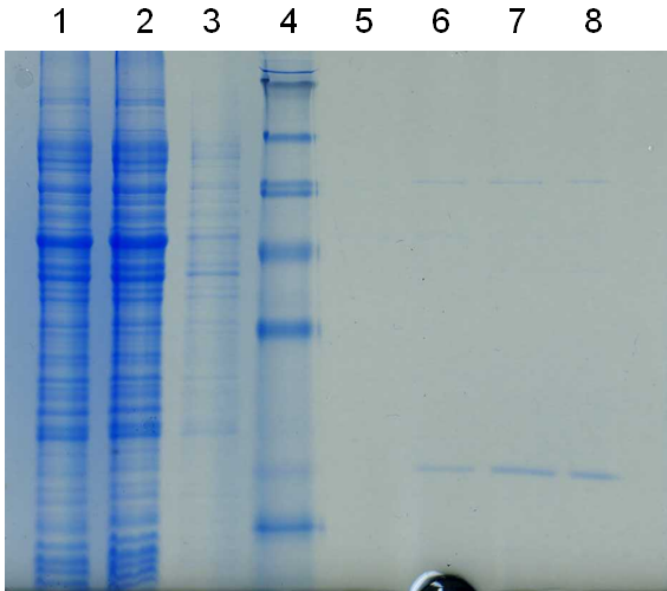
Figure 9.95: SDS-PAGE gel (12%): 1) Lysate (2 μl + 1 μl loading buffer), 2) Flow Through (2 μl + 1 μl loading buffer), 3) Wash (8 μl + 2 μl loading buffer), 4) 7.5 μl SeeBlue Plus2 Pre-stained standard, 5-8) Fractions 1, 2, 3, 4 (16 μl + 4 μl loading buffer).

9.7.4   Labeling reactions

Oxidative
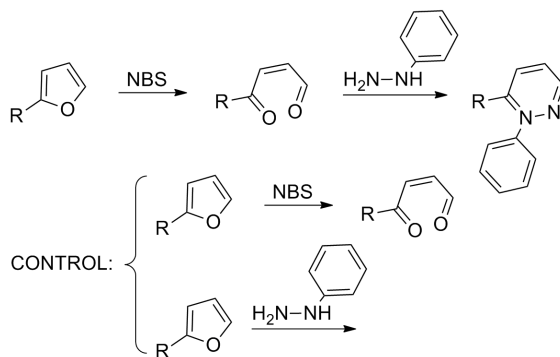
The protein (both furan-modified and blanc as control) was diluted to a 20 μM solution in phosphate buffer of pH 7. NBS solution (10 equiv, 2 nmol, 1 μl, 3.5 mg/10 ml H_2O, 177.98 g/mol) was added to the protein solution on ice. After 1 h, phenylhydrazine solution (50 equiv, 10 nmol, 1 μl, 9.85 μg/10 ml H_2O, 108.14 g/mol) was added and left to react for 1 h. As control reactions, the same procedure was followed with only oxidation and no quenching and with only addition of phenylhydrazine without oxidation.

For analysis 30 μl of water was added to the samples to obtain 5 μM samples.

In a second attempt the reaction was repeated but oxidized with less NBS (5 equiv, 1 nmol, 1 μl, 3.5 mg/20 ml H_2O, g/mol).
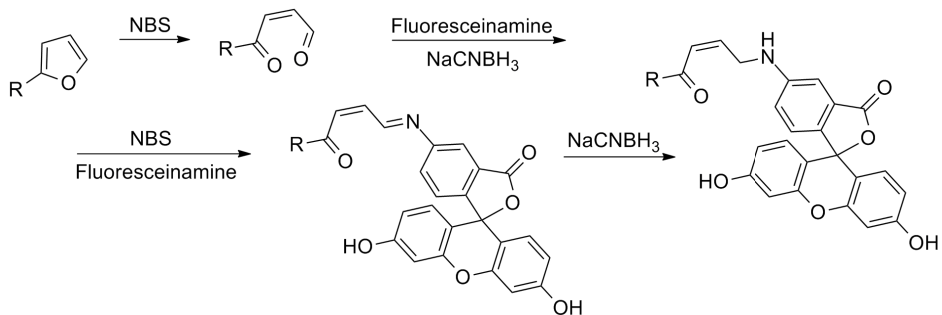
Also, in a seperate sample the reaction was quenched by labeling with fluoresceinamine (50 equiv, 1 μl, 1 mg/288 μl H$_2$O, 347 g/mol) and reduction with NaCNBH$_3$ (10 equiv, 2 nmol, 1 μl, 12.36 mg/20 ml H$_2$O, 62.8 g/mol). For this reaction the amine was added together with the NBS or together with the NaCNBH$_3$.

Non-oxidative
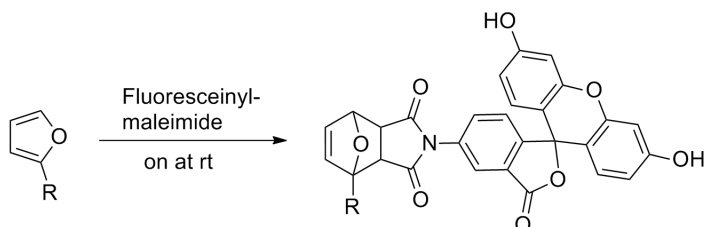
The protein (both furan-modified and blanc as control) was diluted to a 20 μM solution in phosphate buffer of pH 6. Fluoresceinylmaleimide (50 equiv, 10 nmol, 1 μl, 1 mg/234 μl H$_2$O, 427.36 g/mol) was added to 10 μl and the mixture was shaken overnight at room temperature.
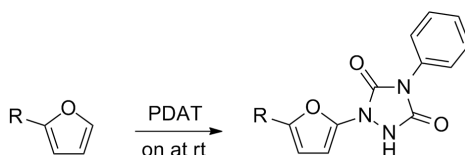
The protein (both furan-modified and blanc as control) was diluted to a 20 μM solution in phosphate buffer of pH 6. 4-phenyl-1,2,4-triazole-3,5-dione (PTAD) (50 equiv, 10 nmol, 1 μl, 1 mg/571.5 μl H$_2$O, 175.14 g/mol) was added to 10 μl and the mixture was shaken overnight at room temperature.

9.8   Supporting Material for Visualisation of Furan based Cross-Linking

9.8.1   Sequences of the bridges

- I-2 = IV-2: furan-modified ODN with **nu6**
- II-2 = V-2: control ODN with **T** instead of **nu6**
- underlined parts of the ODN insert into the five well frame
- italic parts of the ODN are the overlapping sticky ends
- I-4 = II-4 = IV-4 = V-4

7 nt sticky end (furan):

I-1: <u>CGTTAGAATCAGAGCG</u>-
CCACGGCACGGACCTGGTCTCTCGTATGAGTCCTG

I-2: *ACGGnu6GT*-GCTTCCGGTACTACGCCCAGATGAGCTACT

I-3: <u>TGGAAGGGTTAGAACC</u>-
AGTAGCTCATCTGGGCGTAGTACCGGAAGC

I-4: *ACCCCGT*-CAGGACTCATACGAGAGACCAGGTCCGTGCCGTGG

7 nt sticky end (control):

II-1: <u>TTTCAATTACCTGAGC</u>-
CCACGGCACGGACCTGGTCTCTCGTATGAGTCCTG

II-2: *ACGG**T**GT*-GCTTCCGGTACTACGCCCAGATGAGCTACT

II-3: <u>CCGACTTGCGGGAGGT</u>-
AGTAGCTCATCTGGGCGTAGTACCGGAAGC

II-4: *ACCCCGT*-CAGGACTCATACGAGAGACCAGGTCCGTGCCGTGG

control:

III-1: <u>GCTGTTTCCTGTGTGA</u>-
CCACGGCACGGACCTGGTCTCTCGTATGAGTC-CTGACGGGGTGC-
TTCCGGTACTACGCCCAGATGAGCTAC

III-2: <u>CCATTAGCAAGGCCGG</u>-
AGTAGCTCATCTGGGCGTAGTACCGGAA-GCACCCCGTCAG-

GACTCATACGAGAGACCAGGTCCGTGCCGTGG

    9 nt sticky end (furan):

IV-1: <u>TCTAGCTGATAAATTA</u>-
CCACGGCACGGACCTGGTCTCTCGTATGAGTCCTG

IV-2: *ACGG**nu6**GTGC*-TTCCGGTACTACGCCCAGATGAGCTACT

IV-3: <u>AAAGGAATTACGAGGC</u>-
AGTAGCTCATCTGGGCGTAGTACCGGAA

IV-4: *GCACCCCGT*-CAGGACTCATACGAGAGACCAGGTCCGTGCCGTGG

    9 nt sticky end (control):

V-1: <u>TAATTTCAACTTTAAT</u>-
CCACGGCACGGACCTGGTCTCTCGTATGAGTCCTG

V-2: *ACGG**T**GTGC*-TTCCGGTACTACGCCCAGATGAGCTACT

V-3: <u>GTACCGTAACACTGAG</u>-
AGTAGCTCATCTGGGCGTAGTACCGGAA

V-4: *GCACCCCGT*-CAGGACTCATACGAGAGACCAGGTCCGTGCCGTGG

9.8.2 Synthesis of the furan-modified ODN

Available DMT protected intermediate of nu6

The DMT protected intermediate (**4.6.1**) (150 mg) was available from a previous synthesis by Kristof Stevens.[136] Analysis on NMR indicated the sample to be still of good quality.

Phosphoramidite synthesis of nu6

For the synthesis of the phosphoramidite (**4.6.2**) of the nucleoside (**4.6.1**), it was shown that in case of similar glycerol nucleic acid modifications, use of the standard α-cyanoethyl chlorophosphoramidite does not result in formation of the phosphoramidite. The less reactive α-cyanoethyl N,N,N',N'-tetraisopropylphosphoramidite in the presence of diisopropylammonium tetrazolide as activating species is used because of the mildness and chemoselectivity

of this phosphitylation reaction.
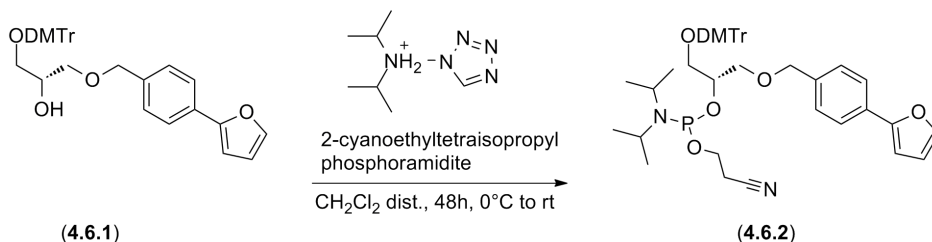
The DMT protected intermediate (4.6.1) (75 mg, 1 equiv, 0.136 mmol, g/mol) was dissolved in fresh distilled CH_2Cl_2 (4 ml) under argon atmosphere. N,N-diisopropylammoniumtetrazolide (62.6 mg, 2.7 equiv, 0.368 mmol, g/mol) was added and stirred at room temperature. 2-cyanoethyltetraisopropylphosphoramidite (82.1 mg, 2 equiv, 0.272 mmol, g/mol) was added while cooling on an ice-bath and the reaction mixture was allowed to stir 48 h at room temperature. When no further conversion occurred, the reaction mixture was quenched with 5% $NaHCO_3$ (4 ml) and then further diluted with CH_2Cl_2. The reaction mixture was extracted 4x with CH_2Cl_2 and the combined organic phases washed with 5% $NaHCO_3$ and brine. It (4.6.2) was then dried on anhydrous Na_2SO_4, and the solvent was removed under reduced pressure.

Analysis

- Rf product: 0.4 (iso-octaan/ethylacetaat : 6/4)
- Chemical Formula: $C_{44}H_{51}N_2O_7P$
- Mass: 750.9 (MW)
- ^{31}P-NMR (121 MHz, $CDCL_3$): 149,48 and 149,31 ppm (product), 14.169 ppm (pentavalent phosphorous due to quenched excess reagent)

I-2 synthesis (ODN4.6.1)

Furan-modified ODN synthesis was performed as described in the general procedure. Though in this case, the phosphoramidite and dicyanoimidazole solutions were mixed and pushed together over the column. Also, no additional washing and

capping steps were performed manually. Due to the length of the ODN (37-mer), additional RP-HPLC purification was required to obtain the desired purity.
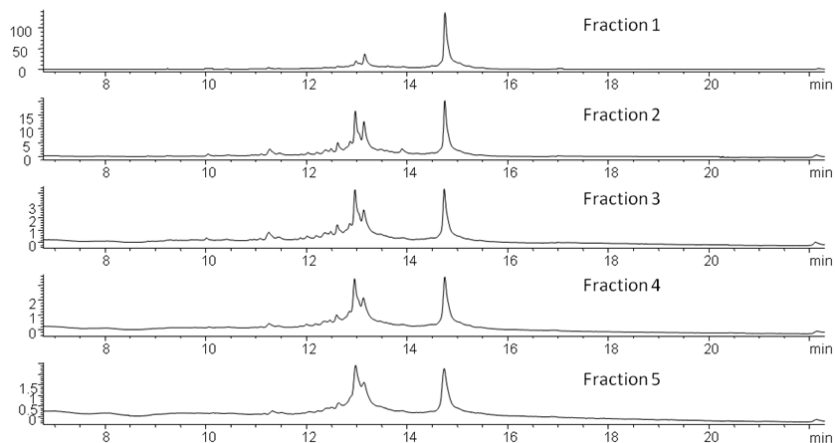
Figure 9.96: RP-HPLC chromatogram of the deprotected ODN4.6.1 fractions after Sep-pak purification, analysed on Clarity column 0-30% MeCN from 0.1M TEAA-buffer in 30 min at 60°C
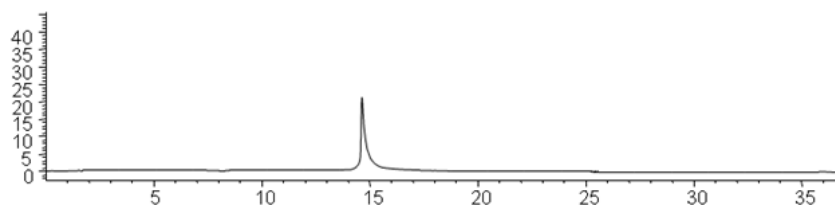
Figure 9.97: RP-HPLC chromatogram of the purified I-3 (ODN4.6.1), analysed on Clarity column 0-30% MeCN from 0.1M TEAA-buffer in 30 min at 60°C

Mass was determined in Japan with a Bruker Microflex KS-II, after desalting using ZipTip (Millipore) with 5-hydroxypicolinic acid/diammonium citrate 5/1 matrix.

Analysis

- Chemical Formula: $C_{364}H_{458}N_{134}O_{221}P_{36}$
- Mass: 11355.9 (Exact), 11361.4 (MW)
- Maldi: 11364.6 $[M\text{-}H]^-$

9.8.3   Origami related protocols

AFM measurements (performed by K. Hidaka)

AFM images were obtained using a fast-scanning AFM system (Nano Live Vision, RIBM Co. Ltd., Tsukuba, Japan) with a silicon nitride cantilever (resonant frequency = 1.0-2.0 MHz, spring constant = 0.1-0.3 N/m, EBD Tip radius <15 nm, Olympus BLAC10EGS- A2). The sample (2 μL) was adsorbed onto a freshly cleaved mica plate (1.5 mm, RIBM Co. Ltd., Tsukuba, Japan) for 5 min at rt and then washed three times using the same buffer solution. Scanning was performed by tapping mode in the same buffer solution.

Denaturing PAGE

The samples were combined with formamide containing loading dye and heated to 60°C for 5 min. Then the samples were loaded on a gel (15% acrylamide gel (37.5:1), 7% urea) preran at 150 V for 15 min and ran at 150 V for 2 h. The gel was then stained for 10 min in fresh SybrGold solution. A mass ladder (mixture of 64, 56, 44 and 32 base pairs) was added in a seperate lane.

Construction of '5-well frame'

An annealing mixture (20 μl) with 0.01 μM M13mp18 single stranded DNA (NewEngland BioLabs) and 0.04 μM of all the staples was prepared in 20 mM Tris buffer pH 7.6 with 1 mM EDTA and 10 mM MgCl$_2$. Annealing was performed with a thermocycler (ASTEC PC350) from 85°C to 15°C, decreasing 1°C per minute (70 min). The origami frame was purified with Sephacryl S-300 high resolution (from GE Healthcare) with Micro Bio-Spin Chromatography Columns (from Biorad).

Insertion of the bridges in '5-well frame'

Two possibilities were explored to insert the bridges in the 5-well frame: Annealing from a mixture of M13mp18, staples and single stranded bridges and annealing of double stranded bridges into an assembled 5-well frame.

Annealing from mixture A mixture of all the bridge strands was made, to have each bridge component in 0.2 μM concentration. 4 and 2 equivalents excess were then combined with 0.01 μM M13mp18 single stranded DNA (NewEngland BioLabs) and 0.04 μM of all the staples in 20 mM Tris buffer pH 7.6 with 1 mM EDTA and 10 mM MgCl$_2$. Annealing was performed from 85 to 15°C, decreasing 1°C per minute (70 min). Gelfiltration was performed as before.
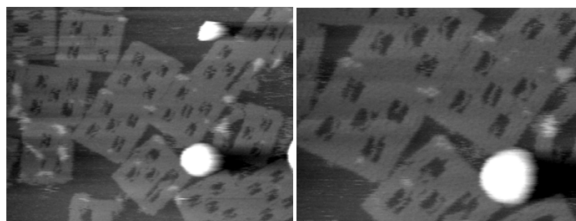
Figure 9.98: AFM images of the 5-well frame with (4 equivalents of) bridges annealed from a mixture
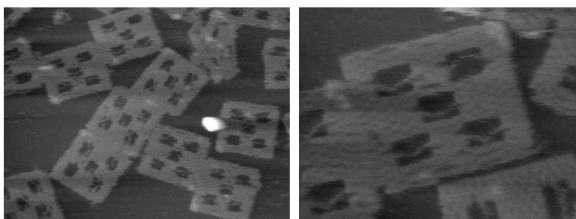
Figure 9.99: AFM images of the 5-well frame with (2 equivalents of) bridges annealed from a mixture

Annealing of bridges in frame The frame was prepared as before. The bridges 1 μM were each annealed separately in solutions of 20 mM Tris buffer pH 7.6 with 1 mM EDTA and 10 mM MgCl$_2$. Annealing was performed from 85 to 15°C, decreasing 1°C per minute (70 min). The annealed bridges were then combined to a mixture with concentration of 0.2 μM. 2 and 4 equivalents of the annealed bridges were added to the frame in 20 mM Tris buffer pH 7.6 with 1 mM EDTA and 10 mM MgCl$_2$. Annealing was performed by heating the mixture to 50°C and cooling at 1°C per minute till 15°C (35 min). Gelfiltration was performed as described before.
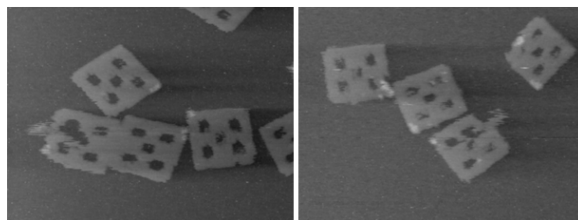
Figure 9.100: AFM images of the 5-well frame with (4 equivalents of) bridges annealed into it from 55°C in 35 minutes
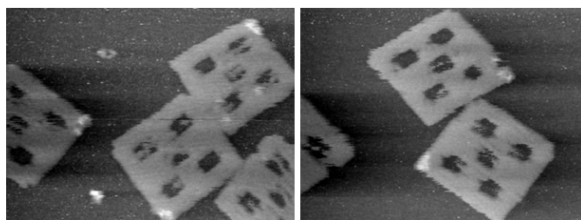
Figure 9.101: AFM images of the 5-well frame with (2 equivalents of) bridges annealed into it from 55°C in 35 minutes

As this clearly does not result in nicely inserted bridges. The procedure was optimized with an overnight annealing. This gives very nice looking origami with bridges assembled in the 5-well frame. These conditions were thus used to prepare origami for the subsequent cross-linking reactions.

Figure 9.102: AFM images of the 5-well frame with (4 equivalents of) bridges annealed into it from 55°C overnight

9.8.4 Cross-linking

A fresh origami stock was prepared for every experiment (10 μl), as described before. Both NBS (4 equiv over the course of 1 h, 4x 1 equiv in 0.375 μl) and singlet oxygen were used to oxidize the furan. The latter was generated by irradiation of the sample with 4 nM of methylene blue with monochromatic red light of 661 nm (JASCO HM-3) for 30 or 60 min. The oxidation was performed on ice, to ensure closed bridges or thus duplex sticky ends.

9.9   Supporting Material for Color Test

The reactions and subsequent color tests were performed according to the standard conditions. Prior to reaction dry resins are allowed to swell in DMF with occasional swirling. For the coupling reaction 4 equiv of N-Fmoc-glycine (297.3 g/mol) and 8 equiv of DIPEA (129.3 g/mol, 0.782 g/ml) were added to the resin which was suspended in half of the total volume of DMF (1.5 ml/mg). After 15 min, 4 equiv of PyBOP (520.4 g/mol) was added together with the remaining necessary volume of DMF. The mixture was flushed with argon and vortexed at rt. This procedure was repeated twice for 2 h. In the standard capping procedure, acetic acid anhydride (6 equiv) and DIPEA (6 equiv) were added to the resin in DMF, which was vortexed for 30 min at rt. To obtain complete capping of the hydroxyl groups, reaction time was prolonged to 4 h. To cleave formed ester bonds, the resin was washed four times for 20 s with an aqueous 0.1 M NaOH solution. After each reaction, excess reagent was filtered off using vacuum. The resin was then thoroughly washed with DMF (4x), MeOH (4x), and DCM (4x).

9.10 Supporting Material for NMR Concentration Determination

Preparation of standards and samples Tryptophan, estradiol, DAPI and menthol were purchased from Sigma Aldrich and the oligonucleotide 5'-GCACACCGTCAG-3' from IDT DNA Technologies. The peptides AcNH-ILPEI-COOH and the H_2N-RSSVGSQS-COOH were synthesized in our lab in previous projects.[293, 294] For all compounds, stock solutions of typically 3 ml were prepared in D_2O 99.9% (or CD_3OD 99.96% for estradiol) using a balance with a precision of 0.1 mg. For NMR measurements, the stock solutions could be used as such, for UV measurements on the other hand all stock solutions had to be diluted with Milli Q H_2O (or MeOH for estradiol) to be within the analysable range for UV spectrometry. This was every time repeated three times, to eliminate random errors. The volumes of the volumetric flasks were chosen as large as possible to further reduce relative errors (exact dilution conditions are given for each measurement).

UV measurements Measurements were performed on a Varian BioSpec UV-Vis spectrophotometer, using the 'Advanced Reads' software. The background was taken into account by zeroing with a blank solution and double beam mode was used to correct for changes in background absorption. Concentration c was derived from the measured absorption A following the law of Lambert-Beer, with l the optical path length (1 cm) and the molar extinction coefficient: $A = \epsilon lc$. To eliminate systematic errors, from each dilution three cuvettes of 1 ml were filled and measured 5 times.

The conditions for the measurements were as follows:

- For the DAPI sample 0.95 ml of the 7.5 mM stock solution in D_2O was diluted 3 times with H_2O to 250 ml and measured at $\lambda max = 344$ nm with $\epsilon = 27\,000$ 1/(cm M).[355]

- For the tryptophan sample 0.95 ml of the 15 mM stock solution in D_2O was diluted 3 times with H_2O to 100 ml and measured at $\lambda max = 278$ nm with $\epsilon = 5579$ 1/(cm M).[355]

- For the estradiol sample 1.00 ml of the 3 mM stock solution in MeOD was diluted 3 times with MeOH to 10 ml and measured at the λmax = 280 nm with ϵ = 2053.9 1/(cm M) (determined via calibration curve).

- For the oligonucleotide sample 0.150 ml of the 2 mM stock solution in D_2O was diluted 3 times with H_2O to 50 ml and measured at λmax = 260 nm with ϵ = 120 000, 112 140 or 114 000 1/(cm M) (calculated taking in account respectively length,[289] length and composition[290] and length, composition and sequence[291])

- For the mixture 0.15 ml of the stock solution of 0.9 mM oligonucleotide and 1.5 mM tryptophan in D_2O was diluted 3 times with H_2O to 25 ml and measured for the oligonucleotide at λmax = 260 nm with ϵ = 114 000 1/(cm M) and for the tryptophan at λmax = 278 nm with ϵ = 5579 1/(cm M).

NMR measurements (performed by K. Gheysen)   For the QUANTAS calibration, tryptophan was used as the standard for water and menthol for methanol. For both compounds, two stock solutions of 3 mM were independently prepared and from each stock solution three samples were taken which were measured in duplicate. All quantitative ^{1}H NMR measurements were performed with a Bruker Avance-II spectrometer running TopSpin 1.3 software, equipped with a 5 mm TXI ^{1}H-^{13}C/^{15}N probe z-gradient probe and operating at 700.13 MHz. The temperature was set to 298 K throughout all the measurements. All spectra were acquired in a quantitative way with a relaxation delay of 30 s and an acquisition time of 3.34 s. The receiver gain was kept constant at 128. The 1D data were accumulated from 8 to 64 scans and multiplied with an exponential function with 2.0 Hz line broadening, followed by fourier transformation and baseline correction. For the integration relative to the artificial QUANTAS signal at -1 ppm, well resolved resonances were selected.

9.11 Supporting material for coupling method

9.11.1 Synthesis of thiolactone-containing building blocks (performed by Dr. P. Espeel)

Synthesis of linker 5.3.1 .
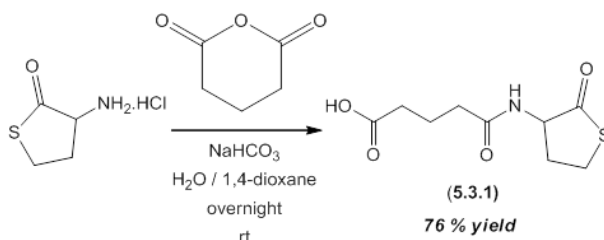
Figure 9.103: One-step synthesis of carboxyl-functionalized linker compound **5.3.1**

An ice-cooled solution of DL-homocysteine thiolactone hydrochloride (7.0 g, 45.6 mmol) in H_2O/1,4-dioxane (1/1, 100 ml) was treated with $NaHCO_3$ (19.15 g, 227.9 mmol) and stirred for 30 minutes at 0°C. Glutaric anhydride (10.4 g, 91.2 mmol) was added to the mixture in several portions. The reaction mixture was allowed to reach room temperature overnight. The pH was adjusted (pH = 1) by drop-wise addition of 12 M HCl. Brine (100 ml) was added and the mixture was extracted with EtOAc (3 x 200 ml). The pooled organic fractions were dried ($MgSO_4$) and concentrated. The crude residue was purified by column chromatography (silicagel, gradient elution: 100% CH_2Cl_2 to CH_2Cl_2/acetone : 6/4), yielding linker **5.3.1** as a white solid (8.1 g, 35.0 mmol, 76%).

Analysis

- ^{1}H-NMR (300 MHz, DMSO-d6, ppm) 12.04 (s, 1 H), 8.17 (d, 1H, 8.1 Hz), 4.58 (ddd, 1 H, 12.7, 8.0, 7.2 Hz), 3.31 (m, 2 H), 2.40 (m, 1 H), 2.31 to 1.99 (m, 5 H), 1.71 (m, 2 H).
- ^{13}C-NMR (75 MHz, DMSO-d6, ppm) 205.5 (C), 174.1 (C), 171.7 (C), 58.1 (CH), 34.2 (CH_2), 32.8 (CH_2), 30.1 (CH_2), 26.7 (CH_2), 20.5 (CH_2).
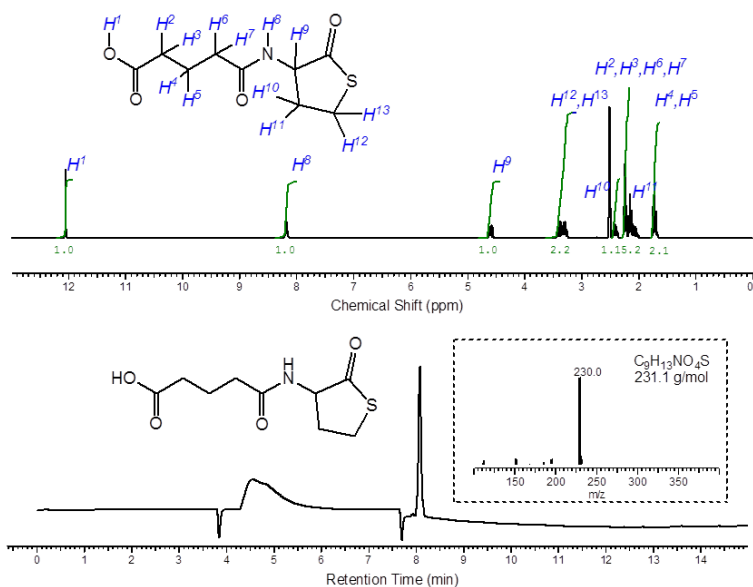
Figure 9.104: ^{1}H-NMR spectrum (300 MHz, DMSO-d6) with peak assignment and integration (top) and HPLC trace with MS analysis (negative mode) of the dominant species (bottom) of the purified linker **5.3.1**.

Synthesis of building block 5.3.2 .
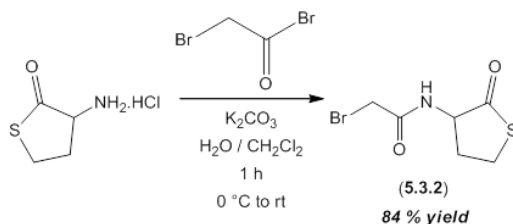
Figure 9.105: One-step synthesis of building block **5.3.2**

Procedure adapted from Rojas-Quijano, F. A.; Benyo, E. T.; Tircso, G.; Kalman, F. K.; Baranyai, Z.; Aime, S.; Sherry, A. D.; Kovacs, Z. Chem. Eur. J. 2009, 15, 13188-13200.

DL-homocysteine thiolactone hydrochloride (9.2 g, 60 mmol) and potassium carbonate (25 g, 180 mmol) were dissolved in H$_2$O (200 ml), and CH$_2$Cl$_2$

(200 ml) was added. The biphasic reaction mixture was cooled to 0°C and a solution of bromoacetyl bromide (10.5 ml, 120 mmol) in CH_2Cl_2 (40 ml) was added dropwise. The reaction was stirred at 0°C for 30 minutes, then allowed to warm to room temperature and stirred for 1 hour. The two layers were separated and the organic layer was washed twice with 5% aqueous citric acid (40 ml) and twice with H_2O (40 ml) before drying over Na_2SO_4. The solvent was then removed in vacuo to afford 12.1 g (84%) of an off-white solid.

Analysis

- ^{1}H-NMR (300 MHz, CD_3CN, ppm) 7.08 (s, 1 H), 4.57 (app dt, 1 H, 12.8, 7.5 Hz), 3.84 (s, 2 H), 3.38 (app dt, 1 H, 11.4, 5.3 Hz), 3.27 (ddd, 1 H, 11.2, 7.1, 1.4 Hz), 2.56 (m, 1 H), 2.15 (m, 1 H).

- ^{13}C-NMR (75 MHz, CD_3CN, ppm) 205.8 (C), 167.4 (C), 60.0 (CH), 31.0 ((CH$_2$), 29.4 (CH$_2$), 27.9 (CH$_2$).
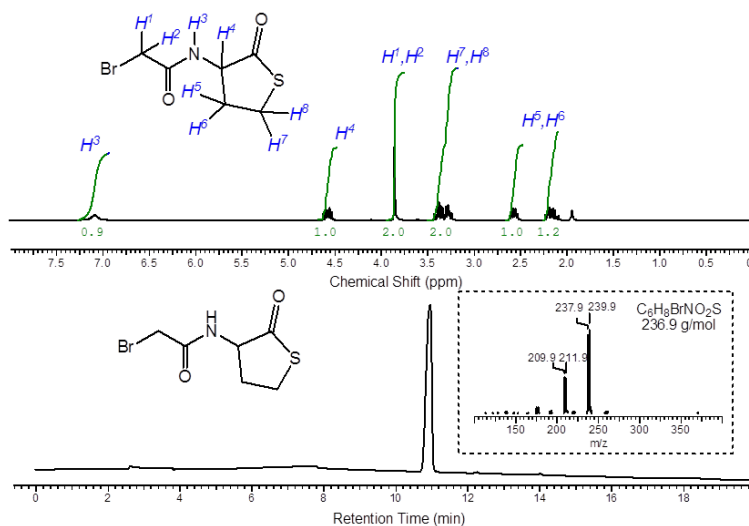
Figure 9.106: ^{1}H-NMR spectrum (300 MHz, CD_3CN) with peak assignment and integration (top) and HPLC trace with MS analysis (positive mode) of the dominant species (bottom) of the purified building block **5.3.2**.

Synthesis of building block 5.3.3 .
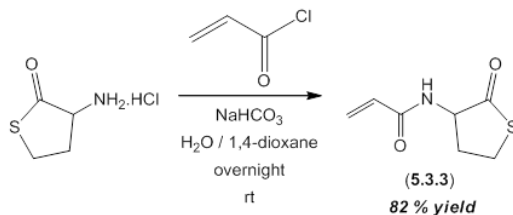
Figure 9.107: One-step synthesis of building block **5.3.3**

An ice-cooled solution of DL-homocysteine thiolactone hydrochloride (7.0 g, 45.6 mmol) in H_2O/1,4-dioxane (1/1, 100 ml) was treated with $NaHCO_3$ (19.15 g, 227.9 mmol) and stirred for 30 minutes at 0°C. Acryloyl chloride (8.3 g, 91.2 mmol) was added to the mixture in several portions. The reaction mixture was allowed to reach room temperature overnight. Brine (100 ml) was added and the mixture was extracted with EtOAc (3 x 200 ml). The pooled organic fractions were dried ($MgSO_4$) and concentrated. The crude residue was purified by column chromatography (silicagel, gradient elution: 100% CH_2Cl_2 to CH_2Cl_2/acetone: 95/5), yielding building block **5.3.3** as a white solid (6.4 g, 37.3 mmol, 82%).

Analysis

- ^{1}H-NMR (300 MHz, DMSO-d6, ppm) 8.45 (d, 1 H, 8.2 Hz), 6.23 (dd, 1 H, 17.1, 9.7 Hz), 6.12 (dd, 1 H, 17.1, 2.6 Hz), 5.65 (dd, 1 H, 9.7, 2.6 Hz), 4.70 (ddd, 1 H, 15.3, 8.2, 7.1 Hz), 3.42 (app dt, 1 H, 11.3, 5.3 Hz), 3.30 (ddd, 1 H, 10.9, 6.9, 1.4 Hz), 2.44 (m, 1 H), 2.09 (m, 1 H).

- ^{13}C-NMR (75 MHz, DMSO-d6, ppm) 205.3 (C), 164.6 (C), 131.0 (CH), 126.2 (CH_2), 58.2 (CH), 30.3 (CH_2), 26.8 (CH_2).
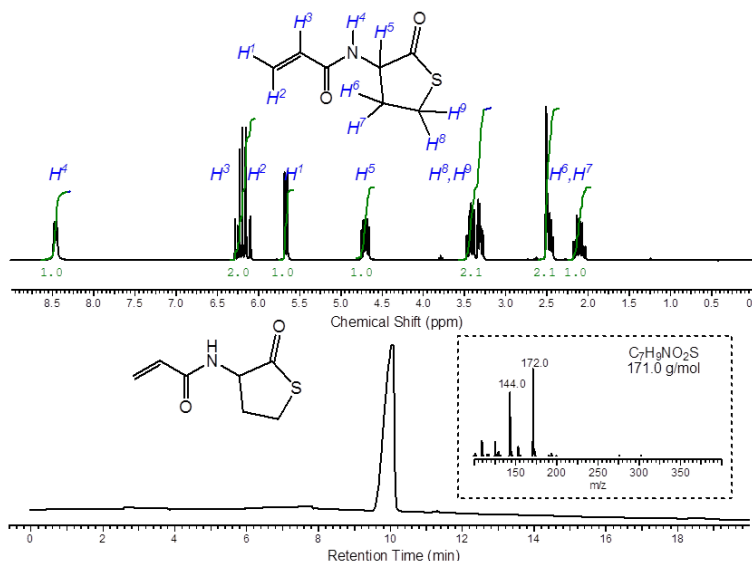
Figure 9.108: ^{1}H-NMR spectrum (300 MHz, DMSO-d6) with peak assignment and integration (top) and HPLC trace with MS analysis (positive mode) of the dominant species (bottom) of the purified building block **5.3.3**.

9.11.2 Oligomer synthesis

Methods

Swelling Prior to the reactions, the solid support was allowed to swell in the solvent, while shaking for at least 30 minutes.

Washing After each reaction, the solid support was washed with the solvent of the reaction (4x), with DMF (4x), methanol (4x), CH_2Cl_2 (4x) and diethyl ether (1x) and dried under vacuum for storage.

Loading The 2-chlorotrityl chloride resin (200 mg, 1.55 mmol/g) was loaded with linker **5.3.1** (86.03 mg, 1.2 equiv rel. to resin) in dry CH_2Cl_2 (2 ml, 10 ml/1 g of resin) with DIPEA (0.26 ml, 4 equiv rel. to acid) and 10 vol% of DMF (0.2 ml), to facilitate the dissolution of the acid. The reaction mixture was shaken for 3 h. The resin was subsequently washed with a CH_2Cl_2/MeOH/DIPEA mixture (17:2:1, 3 x 6 ml), CH_2Cl_2 (3 x 6 ml), DMF (2 x 6 ml), and CH_2Cl_2 (2 x 6 ml). The resin was dried under vacuum. To determine the loading of the resin, the

calibration curve for the carboxylic acid was prepared by UV spectrometry at 254 nm (Fig. 9.109). The efficiency of the immobilization of the thiolactone unit according to the optimized procedure (vide supra) is 75%, resulting in a loading of 1.16 mmol thiolactone units per gram resin.
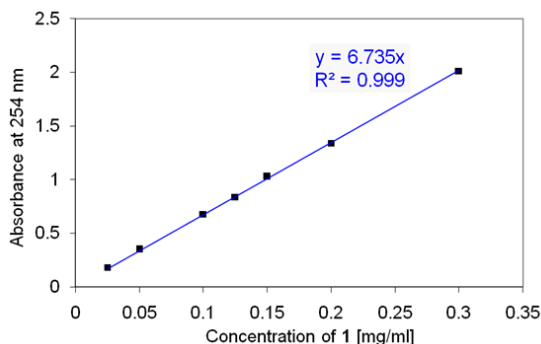
Figure 9.109: Calibration curve for the carboxyl-functionalized linker**1**.

Cleavage The compatibility of the thiolactone ring with standard cleavage conditions for 2-chlorotrityl chloride resins (1% TFA in CH_2Cl_2) was tested by adding 0.5 ml of TFA in CH_2Cl_2 (1%) to 5 mg of loaded resin and shake for 2 h. The resin was filtered, the solvent evaporated, and the residue dried under vacuum. The product was dissolved in MeCN for LC-MS analysis. The results were very similar to the LC-MS results of **1**, confirming the expected structure of the product after the cleavage reaction.

For cleavage, 2 mg of solid support was thus suspended in 0.5 ml of 1% trifluoroacetic acid (TFA) in CH_2Cl_2 and shaken for 2 h. The solid support was then filtered, the sample concentrated by evaporation and dissolved in MeCN for LC-MS analysis.

Aminolysis: general procedure The resin is suspended in CH_2Cl_2 (10 mg resin per 100 μL reaction mixture). The selected amine (50 equiv relative to the resin-bound thiolactone) was added and this mixture is shaken overnight at room temperature. The resin is washed with DMF (4x), methanol (4x), DCM

(4x) and diethyl ether (1x) and dried under vacuum.

Chain extension: general procedure The resin is suspended in DMF (10 mg resin per 100 μL reaction mixture) then the building block **5.3.3** (10 equiv rel. to the resin-bound thiolactone), PhMe$_2$P (10 eq. rel. to the resin-bound thiolactone), and 1 drop of H$_2$O are added. This mixture is shaken overnight at rt. The resin is washed with DMF (4x), MeOH (4x), CH$_2$Cl$_2$ (4x) and diethyl ether (1x) and dried under vacuum.

Microwave-assisted accelerated protocol In order to speed up the overnight aminolysis and chain extension reaction, both reactions were performed with microwave assistance, while all other reaction parameters were according to the general procedures, except for the use of DMF as solvent in both reactions and the limitation to non-volatile amines as the reactions are performed in an open vessel. The aminolysis with benzylamine and subsequent chain extension were performed on the swollen resin by heating at 70°C for 2 h. These conditions were applied for the preparation of longer sequences.

Analysis The oligomers were analysed as described in the general protocol for peptides.

Purification The crude samples were purified by reversed-phase HPLC. The freeze-dried fractions were consequently subjected to NMR and HRMS analysis

9.11.3   Additional data
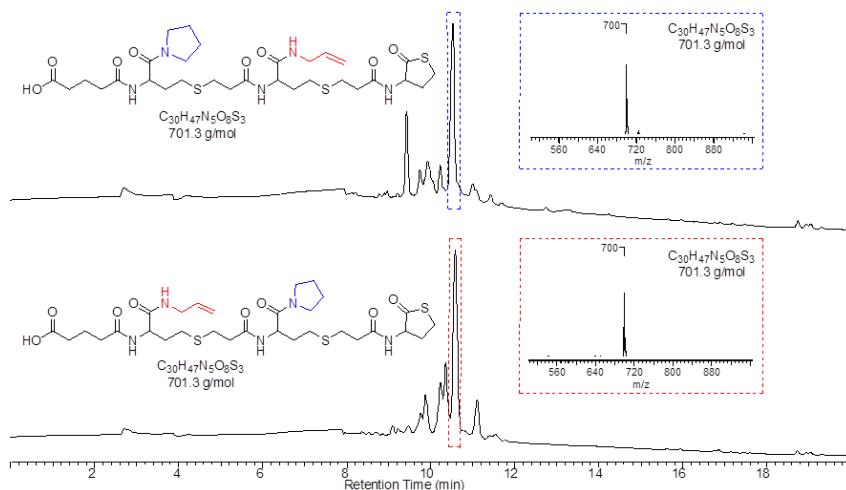
Figure 9.110: LC-MS analysis of the functionalized oligomers originating from repetitive on-resin aminolyses, using two (functional) amines, and chain extensions: *pyrrolidine - allylamine* (top) and the reverse order sequence (bottom). Inserts: ESI-MS-spectra of dominant species (negative mode).
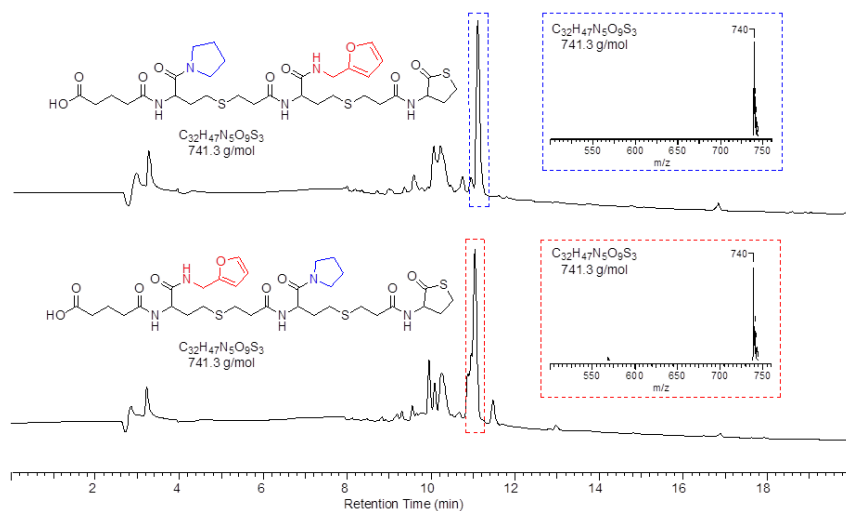
Figure 9.111: LC-MS analysis of the functionalized oligomers originating from repetitive on-resin aminolyses, using two (functional) amines, and chain extensions: *pyrrolidine - furfurylamine* (top) and the reverse order sequence (bottom). Inserts: ESI-MS-spectra of dominant species (negative mode).
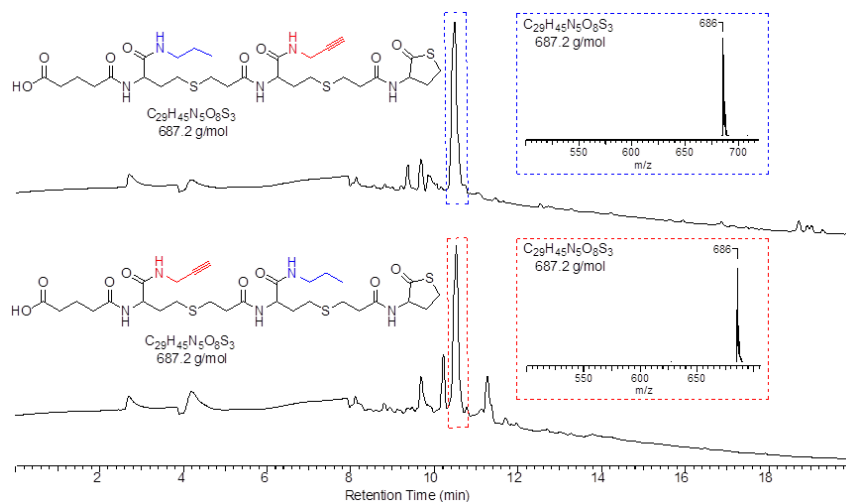
Figure 9.112: LC-MS analysis of the functionalized oligomers originating from repetitive on-resin aminolyses, using two (functional) amines, and chain extensions: *n-propylamine - propargylamine* (top) and the reverse order sequence (bottom). Inserts: ESI-MS-spectra of dominant species (negative mode).
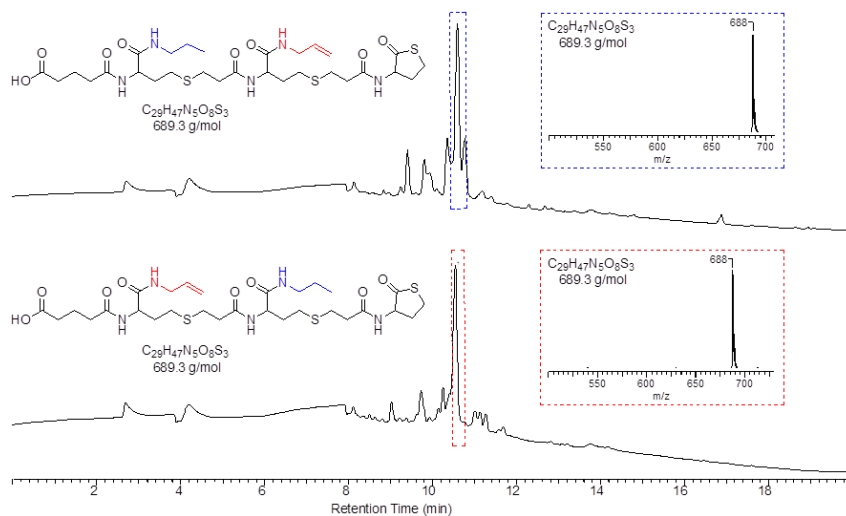
Figure 9.113: LC-MS analysis of the functionalized oligomers originating from repetitive on-resin aminolyses, using two (functional) amines, and chain extensions: *n-propylamine - allylamine* (top) and the reverse order sequence (bottom). Inserts: ESI-MS-spectra of dominant species (negative mode).
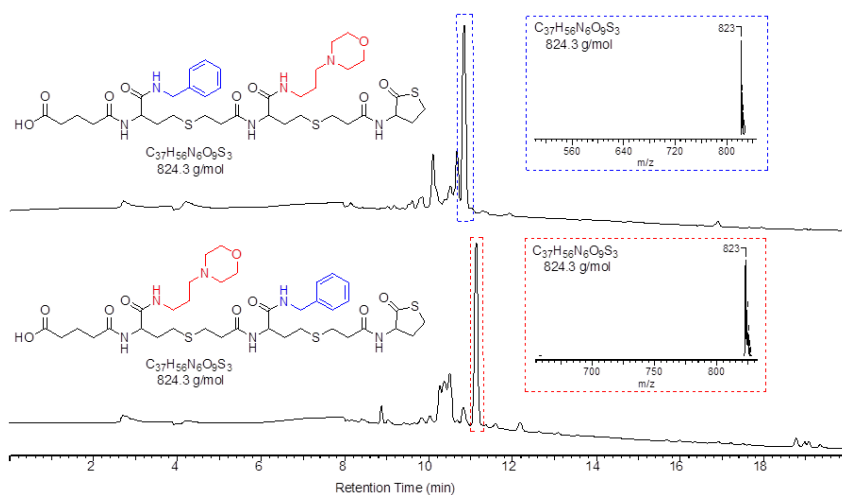
Figure 9.114: LC-MS analysis of the functionalized oligomers originating from repetitive on-resin aminolyses, using two (functional) amines, and chain extensions: *benzylamine - N-(3-aminopropyl)morpholine* (top) and the reverse order sequence (bottom). Inserts: ESI-MS-spectra of dominant species (negative mode).

HRMS data   High resolution mass spectra (HRMS) were collected using an Agilent 6220 Accurate-Mass time-of-flight (TOF) equipped with a multimode ionization (MMI) source. Monoisotopic masses were calculated using MolE - Molecular Mass Calculator v2.03.
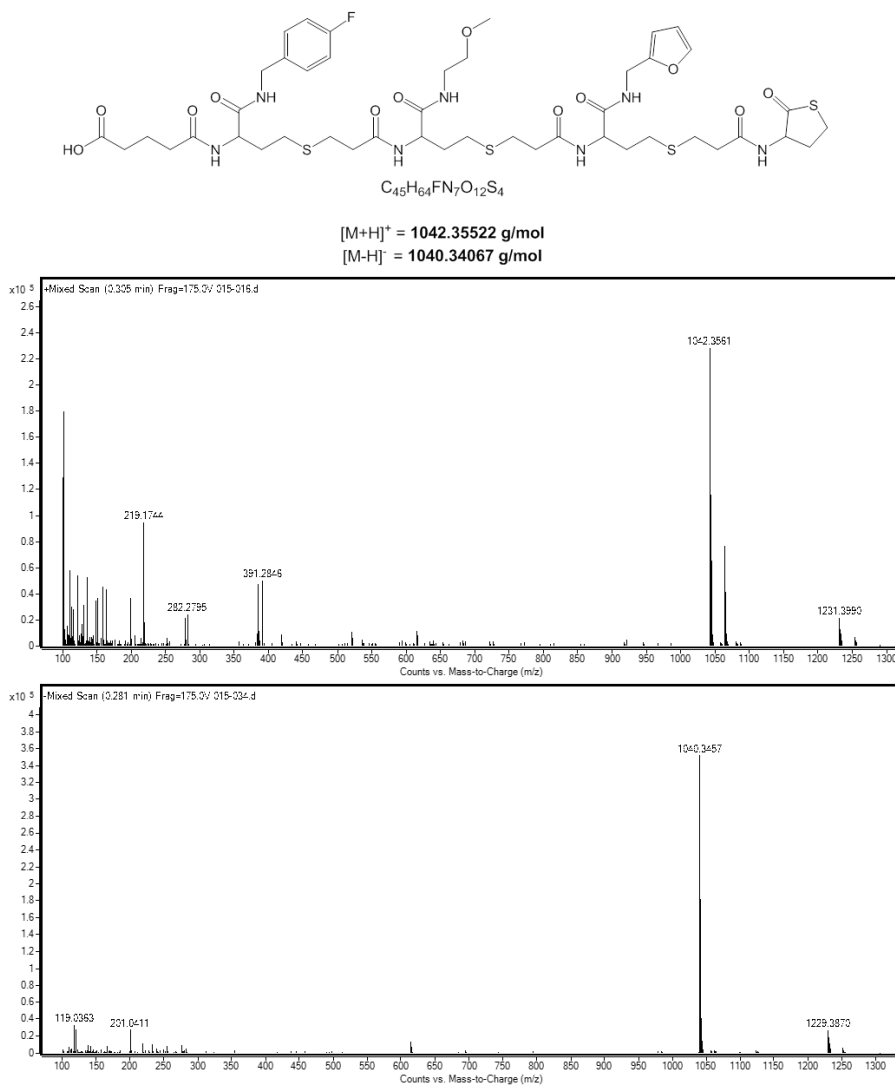
Figure 9.115: HRMS analysis of depicted trimer sequence (**A I**), in positive (top) and in negative (bottom) mode
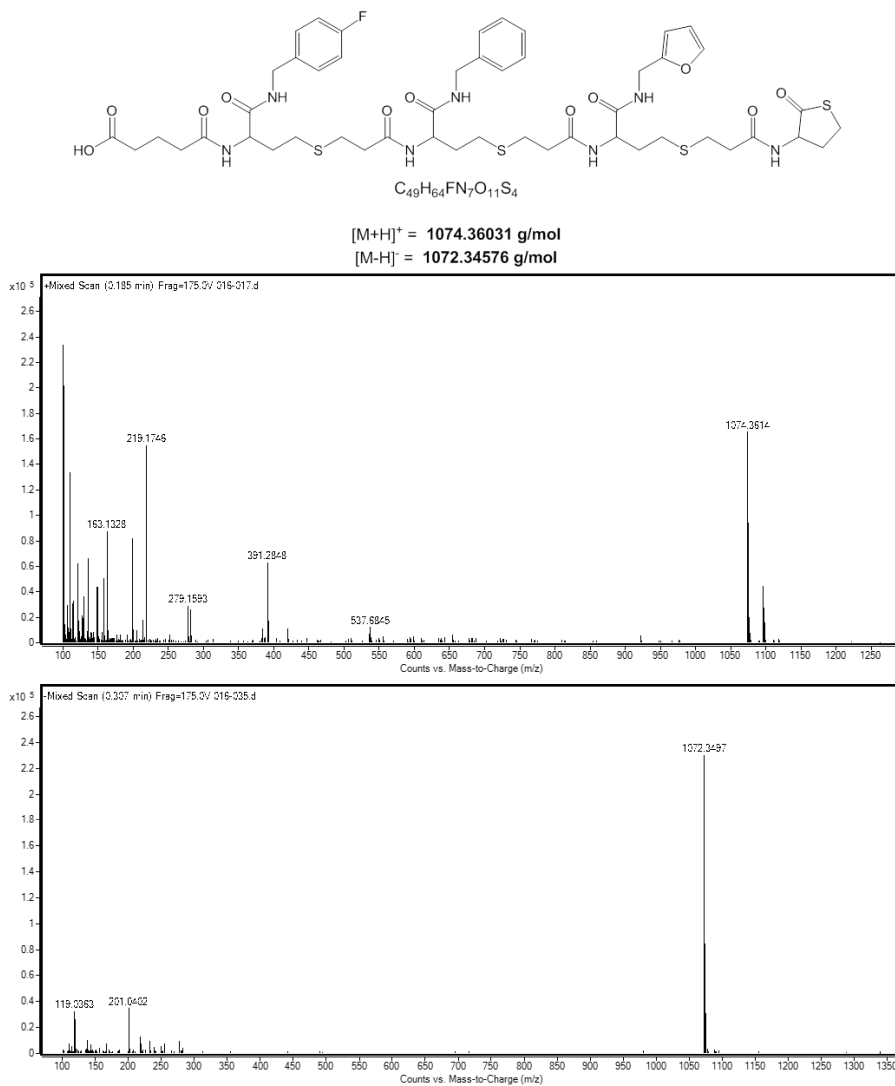
Figure 9.116: HRMS analysis of depicted trimer sequence (**A II**), in positive (top) and in negative (bottom) mode
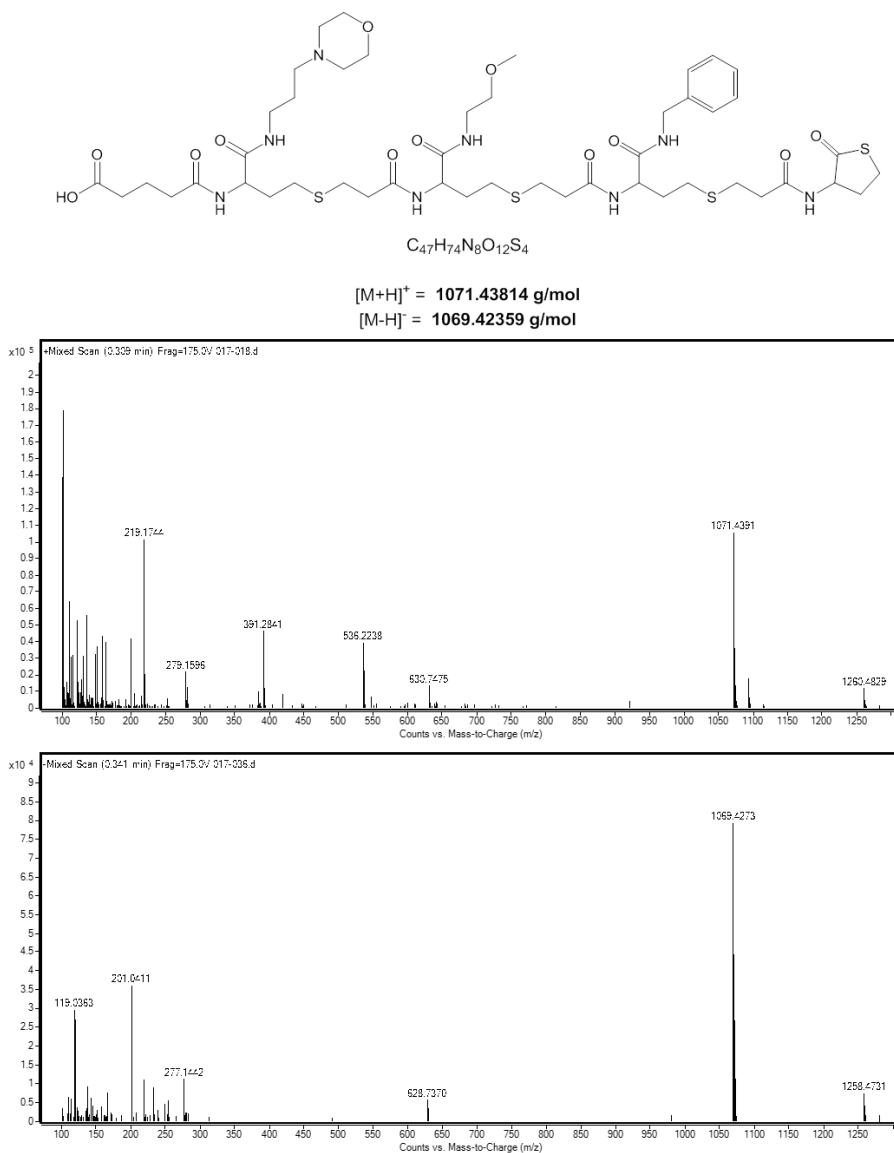
Figure 9.117: HRMS analysis of depicted trimer sequence (**B I**), in positive (top) and in negative (bottom) mode
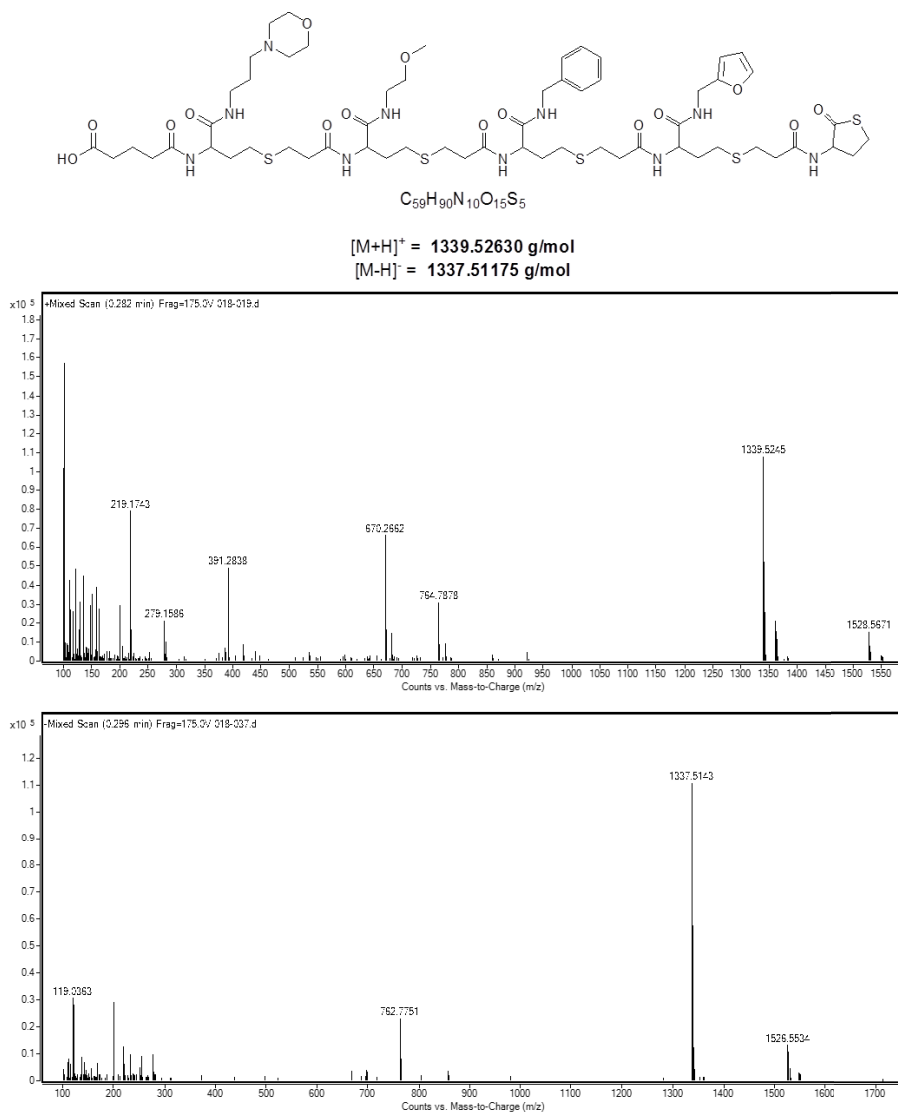
Figure 9.118: HRMS analysis of depicted trimer sequence (**B II**), in positive (top) and in negative (bottom) mode

Appendix A

Current Protocols in Nucleic Acid Chemistry

A.1 BASIC PROTOCOL 1: Synthesis of selective 2'-furan-modified uridine nucleoside phosphoramidite

The sequence of reactions in this protocol was reported by Op de Beeck and Madder[138] and is shown in Figure A.1. First, uridine is converted to 2,2'-O-anhydro-1-(β-D-arabinofuranosyl)-uracil (**2**) by cyclization as described by McGee et al.[356] This is converted to 2'-azido-uridine (**3**) by substitution [357], which yields the key intermediate 2'-amino-2'-deoxyuridine (**4**) after Staudinger reduction.[358] The furan modification is then introduced by treatment with furfurylisocyanate to give **5**. Finally, 5'-O-tritylation and 3'-O-phosphitylation yield the desired phosphoramidite **7**, which is highly selective for ICL formation with C bases in the complementary strand.

All intermediate compounds up to the trityl-protected **6** are stable enough for prolonged storage (under argon at -20°C) and may be prepared in larger quantities. In contrast, because the final phosphoramidite **7** is very unstable, it is advisable to prepare only the required quantity just prior to oligonucleotide synthesis. In addition, phosphitylation should be performed under dry and inert conditions.
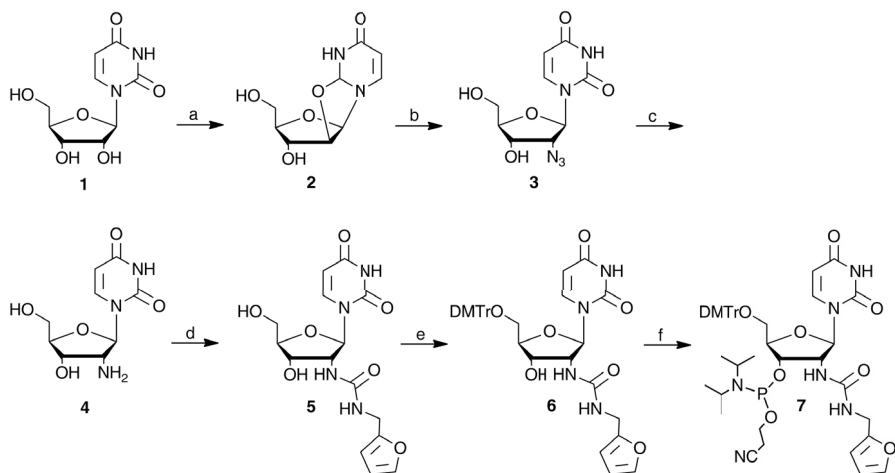
Figure A.1: Synthesis of 2'-furan-modified nucleoside phosphoramidite 7.
a. (PhO)$_2$CO, NaHCO$_3$, DMF, 16 h, 115°C, 74%; b. LiF, Me$_3$SiN$_3$, DMF/TMEDA,
63 h, 110°C, 72%; c. PPh$_3$, MeOH, H$_2$O, 3 h, 50°C, 89%; d. furfurylisocyanate,
DMF, 1 h, 96%; e. DMTrCl, pyridine, 17 h, rt, 79%; f. 2-cyanoethyl-N,N-
diisopropylchlorophosphine, DIPEA, 2.5 h, 0°C, 50%.

Materials

- Uridine
- Diphenylcarbonate ((PhO)$_2$CO)
- N,N-Dimethylformamide (DMF), extra dry, on molecular sieves
- Sodium bicarbonate (NaHCO$_3$)
- Dichloromethane (CH$_2$Cl$_2$)
- Methanol (MeOH)
- Lithium fluoride (LiF)
- Azidotrimethylsilane (Me$_3$SiN$_3$)
- Tetramethylethylenediamine (TMEDA)
- Ethyl acetate (EtOAc)
- Triphenylphosphine (PPh$_3$)
- Toluene
- Furfurylisocyanate (Sigma-Aldrich)
- Pyridine, extra dry, on molecular sieves
- 4,4'-Dimethoxytritylchloride (DMTrCl)
- Saturated aqueous sodium bicarbonate (NaHCO$_3$)
- Brine (saturated aqueous NaCl)
- Sodium sulfate (Na$_2$SO$_4$)

- Isooctane
- Triethylamine (TEA)
- Diisopropylethylamine (DIPEA)
- 2-Cyanoethyl N,N-diisopropylchlorophosphoramidite
- 25-, 50-, and 100-ml round-bottom flasks
- 10- and 25-ml oven-dried round-bottom flask
- Liebig condensers
- No. 4 sintered glass filter
- Rotary evaporator
- Vacuum pump (dry ice + isopropanol)
- Distillation apparatus for CH2Cl2
- Additional reagents and equipment for TLC (Curr. Protoc. Nucl. Acid Chem. APPENDIX 3D) and flash chromatography (Curr. Protoc. Nucl. Acid Chem. APPENDIX 3E)

Prepare 2,2'-O-anhydro-1-(β-D-arabinofuranosyl)-uracil (2)

1. Add 4.99 g uridine (**1**; 25.6 mmol) to a solution of 5.49 g $(PhO)_2CO$ (1 eq, 25.6 mmol) in 5.5 ml dry DMF in a 50-ml round-bottom flask with a Liebig condenser. Heat to 80°C until a white turbid solution is obtained ($\sim$2 h).

2. Add 0.046 g $NaHCO_3$ (0.02 eq, 0.55 mmol) and stir at 115°C until uridine is completely consumed (5 h to overnight). Monitor the reaction by TLC (Curr. Protoc. Nucl. Acid Chem. APPENDIX 3D) using 8:2 (v/v) CH_2Cl_2/MeOH (Rf 1 = 0.2; Rf 2 = 0.12).
   After $\sim$4 h, gas evolution subsides and a white precipitate forms.

3. Allow mixture to cool to room temperature, then isolate the white precipitate by filtration through a no. 4 sintered glass filter. Wash with 10 to 20 ml MeOH.

4. Partially evaporate the filtrate on a rotary evaporator, then isolate the remaining solid again by filtration. Repeat the concentration/precipitation process twice more.
   2,2'-O-Anhydro-1-(β-D-arabinofuranosyl)-uracil (2) is obtained as a white powder (yield 74%).

Prepare 2'-azido-uridine (3)

5. Add 0.62 g LiF (1.8 eq, 23.9 mmol) in dry DMF (13.3 ml) to a 50-ml round-bottom flask with a Liebig condenser and heat to 105°C.

6. Add 3.2ml Me$_3$SiN$_3$ (1.8 eq, 23.9 mmol) and 16.6 ml TMEDA (8.3 eq, 110.7 mmol) and stir for 30 min.

7. Add 3.00 g **2** (13.3 mmol) and continue stirring for another 64 h. Monitor the reaction by TLC using 8:2 (v/v) CH$_2$Cl$_2$/MeOH (Rf = 0.59).

8. Allow the reaction mixture to cool to room temperature, then concentrate by rotary evaporation.

9. Dissolve the residue in 10 ml MeOH in a 100-mL round-bottom flask.

10. Add 40 ml EtOAc to precipitate the majority of salts and starting material.

11. Remove solid residue by filtration and wash with 40 ml of 1:4 (v/v) MeOH/EtOAc. Collect and concentrate the filtrate.

12. Perform a first rudimentary purification by flash chromatography (Curr. Protoc. Nucl. Acid Chem. APPENDIX 3E) using 4:1 (v/v) EtOAc/MeOH. Monitor fractions by TLC as above (Rf = 0.59).
    If the compound is not soluble in the eluent, evaporate the filtrate on silica (10 ml) before loading on the column.

13. Combine fractions containing product and remove solvent under reduced pressure on a rotary evaporator to give a brown foam.

14. Repurify by flash chromatography using 9:1 (v/v) CH$_2$Cl$_2$/MeOH, monitoring fractions by TLC as above (Rf = 0.59).

15. Combine fractions containing product and remove the solvent under reduced pressure on a rotary evaporator and then vacuum pump.
    *2'-Azido-uridine (**3**) is obtained as an off-white foam (yield 72%).*

Prepare 2'-amino-2'-deoxyuridine (4) by Staudinger reduction

16. Prepare a solution of 0.406 g **3** (1.51 mmol) in 16.5 ml of 9:1 (v/v) MeOH/H$_2$O in a 50-ml round-bottom flask with a Liebig condenser.

17. Add 0.434 g PPh$_3$ (1.1 eq, 1.65 mmol) and stir at 50°C for 3 h. Monitor by TLC in 8:2 (v/v) CH$_2$Cl$_2$/MeOH (Rf = 0.09).

18. Combine fractions that contain product and concentrate by rotary evaporation.

19. Purify the product by recrystallizing overnight in 15 ml toluene under reflux conditions (110°C).
    *2'-Azido-2'-deoxyuridine (**4**) is obtained as a yellowish powder (yield 89%). Best results are obtained with overnight crystallization.*

Prepare furan-modified nucleoside 5

20. Dissolve 0.200 g **4** (0.82 mmol) in 8.2 ml dry DMF in a 25-ml round-bottom flask, and then cool in an ice bath.

21. Add 100 μl furfurylisocyanate (1.1 eq, 0.90 mmol) dropwise and then stir the cold reaction for 1 h. Monitor by TLC using 9:1 (v/v) CH_2Cl_2/MeOH (Rf = 0.24).

22. Quench the reaction with 1 ml MeOH, then concentrate the product under vacuum.

23. Purify the residue by flash chromatography using 95:5 (v/v) CH_2Cl_2/MeOH and monitor fractions by TLC as above (Rf = 0.24).

24. Combine fractions containing the product and remove solvent under reduced pressure on a rotary evaporator and then vacuum pump.
    2'-Deoxy-2'-[3-(2-furyl)-2-azapropanamido)]uridine (5) is obtained as an off-white foam (yield 96%).

Perform dimethoxytritylation to give 6

25. Dry 76 mg **5** (0.207 mmol) by co-evaporating three times with 1 ml dry pyridine in an oven-dried 25-ml round-bottom flask.

26. Dissolve in 2 ml dry pyridine and cool to 0°C on ice.

27. Add 99 mg DMTrCl (1.4 eq, 0.29 mmol) and stir overnight at room temperature.

28. Monitor by TLC using 9:1 (v/v) CH_2Cl_2/MeOH (Rf = 0.43). If the reaction is not complete, add another 35 mg DMTrCl (0.5 eq, 0.10 mmol) at 0°C, and continue stirring at room temperature for an additional 4 h.

29. When the reaction is complete, cool in an ice bath and quench with 100 μl MeOH.

30. Dilute with 10 ml EtOAc, then wash successively with 10 ml water, 10 ml saturated aqueous $NaHCO_3$, and 10 ml brine.

31. Extract the combined aqueous layers with 30 ml EtOAc.

32. Dry the combined organic layers on Na_2SO_4 and then concentrate by rotary evaporation.

33. Purify the residue by flash chromatography using 70:29:1 (v/v/v) EtOAc/isooctane/TEA and monitor fractions by TLC using 9:1 (v/v) CH_2Cl_2/MeOH (Rf = 0.43).
    The eluent should contain 1% TEA because the DMTr protecting group is acid-sensitive. This eluent should also be used to pretreat the silica resin by eluting 2 column volumes before the product is loaded on the column.

34. Combine fractions containing product and remove the solvent under reduced pressure on a rotary evaporator and then vacuum pump.
    2'-Deoxy-2'-[3-(2-furyl)-2-azapropanamido)]-5'-O-(4,4'-dimethoxytrityl)uridine
    (6) is obtained as an off-white foam (yield 79%).

Perform phosphitylation to give phosphoramidite 7

35. Dry 79 mg **6** (0.118 mmol) by co-evaporating twice with 1 ml toluene and once with 1 ml freshly distilled CH_2Cl_2 in an oven-dried 10-ml round-bottom flask.

36. Dissolve in 2 ml dry CH_2Cl_2 and cool to 0°C on ice.

37. Add 65 μl DIPEA (3 eq, 0.364 mmol) followed by 40 μl 2-cyanoethyl N,N-diisopropylchlorophosphoramidite (1.5 eq, 0.179 mmol) and stir for 1.5 h at 0°C. Monitor by TLC using 8:2 (v/v) isoochane/EtOAc (Rf = 0.45).

38. Quench with 100 μl MeOH.

39. Dilute with 20 μl EtOAc and wash the organic phase once with 20 ml saturated aqueous $NaHCO_3$ and three times with 20 ml brine.

40. Dry the organic phase on Na_2SO_4 and concentrate by rotary evaporation.

41. Purify the residue by flash chromatography using 55:44:1 (v/v/v) EtOAc/isooctane/TEA and monitor by TLC as above (Rf = 0.45).
    As the phosphoramidite is very sensitive, the purification should be performed swiftly, using 1% TEA in the eluent and pretreating the silica as in step 33.

42. Combine fractions containing product and remove solvent under reduced pressure on a rotary evaporator and then vacuum pump.
    2'-Deoxy-2'-[3-(2-furyl)-2-azapropanamido)]-5'-O-(4,4'-dimethoxytrityl)uridine
    3'-(2-cyanoethyl N,N-diisopropylphosphoramidite) (7) is obtained as an off-white
    foam (yield 50%).

A.2   ALTERNATE PROTOCOL: Synthesis of 2'-furan-modified acyclic nucleoside phosphoramidite

The sequence of reactions in this protocol was reported by Stevens and Madder[135] and is illustrated in Figure A.2. First, (4S)-(+)-4-(2-hydroxyethyl)-2,2-dimethyl-1,3-dioxolane (**8**) is activated by tosylation to **9**. This is converted by $S_N 2$ to the bromine derivative **10**. The furan modification is then introduced by addition of lithiated furan to give **11**. Acid deprotection results in (S)-4-furan-2-yl-butane-1,2-diol (**12**). Finally, 5'-O-tritylation and 3'-O-phosphitylation yield the desired phosphoramidite **14**, which

is useful for high yield ICL reactions with complementary A or C bases.

Intermediates **9** and **1113** are stable enough for prolonged storage (under argon at -20°C). We are not sure about the stability of **10** and advise to use it immediately. Phosphoramidite **14** is unstable and should be prepared only in the quantity needed just before oligonucleotide synthesis. Phosphitylation should be performed under dry and inert conditions.
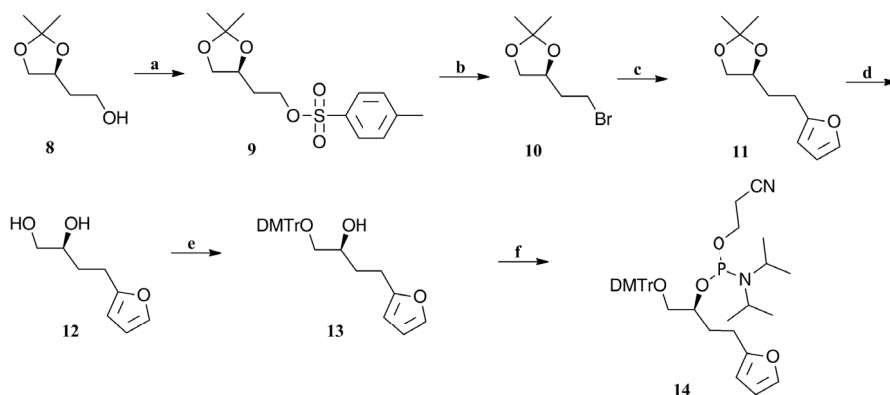
Figure A.2: Synthesis of acyclic furan-modified phosphoramidite 14. a. pTsCl (1.15 eq.), pyridine, 0°C, 2.5 h, 91%; b. LiBr (5 eq.), DMF, 60°C, 1 h, 76%; c. furan (2 eq.), BuLi (1.8 eq.), THF, -78° to 0°C, 16 h, 62% over two steps; d. 4 M HCl, THF, 3 h, 0°C, 89%; e. DMTrCl (1 eq.), pyridine, 0°C, 1.25 h, 66%; f. (iPr$_2$N)(NCCH$_2$CH$_2$O)PCl (2.5 eq.), DIPEA, CH$_2$Cl$_2$, 1.5h, rt, 79%.

Materials

- (4S)-(+)-4-(2-Hydroxyethyl)-2,2-dimethyl-1,3-dioxolane (**8**; Sigma-Aldrich)
- Pyridine, extra dry, on molecular sieves
- p-Toluenesulfonyl chloride (pTsCl)
- Argon gas
- Ethyl acetate (EtOAc)
- Pentane
- 1 M and 4 M aqueous HCl
- Diethyl ether (Et$_2$O)
- Saturated aqueous sodium bicarbonate (NaHCO$_3$)
- Anhydrous sodium sulfate (¡eNa2SO4)
- N,N-Dimethylformamide (DMF), extra dry, on molecular sieves

- Lithium bromide (LiBr),
  made anhydrous by heating under vacuum for 1 min
- Isooctane
- Tetrahydrofuran (THF), extra dry, on molecular sieves
- Isopropanol
- Dry ice
- n-Butyllithium (n-BuLi)
- Hexane
- Furan
- 4,4'-Dimethoxytritylchloride (DMTrCl)
- Cyclohexane
- Methanol (MeOH)
- Triethylamine (TEA)
- Toluene
- Dichloromethane (CH_2Cl_2)
- Diisopropylethylamine (DIPEA)
- 2-Cyanoethyl N,N-diisopropylchlorophosphoramidite
- 25- and 50-ml two-neck flasks
- 25- and 50-ml oven-dried round-bottom flasks
- Rotary evaporator
- Vacuum pump (dry ice + isopropanol)
- Distillation apparatus for CH_2Cl_2
- Additional reagents and equipment for TLC (Curr. Protoc. Nucl. Acid Chem. APPENDIX 3D) and flash chromatography (Curr. Protoc. Nucl. Acid Chem. APPENDIX 3E)

Prepare tosylate 9

1. Weigh 1.73 g (4S)-(+)-4-(2-hydroxyethyl)-2,2-dimethyl-1,3-dioxolane (**8**; 11.85 mmol) in a 25-ml two-neck flask.

2. Add 1.5 ml dry pyridine and 2.59 g pTsCl (1.15 eq, 13.6 mmol), flush the flask with argon, and stir at room temperature for 2.5 h. Monitor the reaction by TLC (Curr. Protoc. Nucl. Acid Chem. APPENDIX 3D) using 1:1 (v/v) EtOAc/pentane (Rf = 0.71).
   A white suspension is formed after a couple of minutes.

3. Quench the reaction by pouring the mixture into 3 ml of 1 M aqueous HCl.

4. Extract product with 5 ml $Et_2O\cdot$

5. Wash the organic phase three times with 5 ml of 1 M aqueous HCl and three times with 5 ml saturated aqueous $NaHCO_3$.

6. Dry with anhydrous Na_2SO_4 and evaporate under reduced pressure on a rotary evaporator and then vacuum pump.
   *(S)-Toluene-4-sulfonic acid 2-(2,2-dimethyl-[1,3]dioxolan-4-yl)-ethyl ester (**9**) is obtained as a colorless oil (yield 91%) and can be used for further reactions without purification.*

Prepare (S)-4-(2-Bromo-ethyl)-2,2-dimethyl-[1,3]dioxolane (10)

7. Dissolve 1.99 g **9** (6.63 mmol) in 10 ml extra dry DMF in a 50-ml oven-dried round-bottom flask.

8. Add 2.54 g anhydrous LiBr (5 eq, 33 mmol). Slowly heat the mixture to 60°C with stirring, and then let stir another 1 h at 60°C. Monitor by TLC using 2:3 (v/v) EtOAc/isooctane (Rf = 0.53).

9. Quench reaction by pouring the mixture into 25 ml ice water.

10. Extract product three times with 25 ml Et_2O. Combine the organic phases and wash with 75 ml water and then 75 ml saturated aqueous $NaHCO_3$.

11. Dry the organic phase with anhydrous Na_2SO_4 and evaporate under reduced pressure.
    *Product **10** is volatile (b.p. 90° C/19 Torr). To prevent evaporation of the compound, never turn on the heating water bath of the rotary evaporator.*

12. Purify residue by flash chromatography (Curr. Protoc. Nucl. Acid Chem. APPENDIX 3E) using 1:5 (v/v) EtOAc/pentane and monitor fractions by TLC as above (Rf = 0.53).

13. Combine fractions containing product and remove solvent under reduced pressure on a rotary evaporator and then vacuum pump.
    *(S)-4-(2-Bromo-ethyl)-2,2-dimethyl-[1,3]dioxolane (**10**) should be used soon after preparation.*

Prepare (S)-4-(2-Furan-2-yl-ethyl)-2,2-dimethyl-[1,3]dioxolane (11)

14. Add 6 ml dry THF to a 50-ml two-neck flask and cool to -78°C in an isopropanol/dry ice bath.

15. While stirring, add 7.5 ml of 1.6 M n-BuLi in hexane and 0.9 g furan (2 eq, 13.25 mmol). Leave at -78°C for another 15 min.

16. Transfer flask to a normal ice bath ($0°C$) and leave for 2 h.
    A white precipitate is formed.

17. During precipitation, co-evaporate the entire amount of bromide **10** three times with 10 ml dry THF to ensure that it is dry for the next reaction.

18. Add dry **10** to the precipitate in the flask and stir the reaction overnight at room temperature. Monitor by TLC using 2:3 (v/v) EtOAc/isooctane (Rf = 0.56).

19. Quench the reaction by pouring into 25 ml ice water.

20. Extract product with 25 ml Et_2O. Dry the organic phase with anhydrous Na_2SO_4 and evaporate under reduced pressure.

21. Purify residue by flash chromatography using 1:10 (v/v) EtOAc/pentane and monitor fractions by TLC as above (Rf = 0.56).

22. Combine fractions containing product and remove solvent under reduced pressure on a rotary evaporator and then vacuum pump.
    *(S)-4-(2-Furan-2-yl-ethyl)-2,2-dimethyl-[1,3]dioxolane (**11**) is obtained as a dark yellow oil (yield 62% over two steps).*

Prepare (S)-4-furan-2-yl-butane-1,2-diol (12)

23. Dissolve the 0.309 g diol **11** (1.57 mmol) in 3 ml THF in a 25-ml round-bottom flask.

24. Add 1.5 ml of 4 M HCl while stirring in an ice bath ($0°C$), then continue stirring at $0°C$ for 3 h. Monitor by TLC using 2:3 (v/v) EtOAc/pentane (Rf = 0.14).

25. Extract product with 10 ml EtOAc and then evaporate under reduced pressure.

26. Purify residue by flash chromatography using 1:1 (v/v) EtOAc/pentane and monitor fractions by TLC as above (Rf = 0.14).

27. Combine fractions containing product and remove solvent under reduced pressure on a rotary evaporator and then vacuum pump.
    *(S)-4-Furan-2-yl-butane-1,2-diol (**12**) is obtained as a red oil (yield 89%).*

Prepare (S)-1-[bis-(4-methoxy-phenyl)-phenyl-methoxy]-4-furan-2-yl-butane-2-ol (13)

28. Dry 0.212 g diol **12** (1.36 mmol) by co-evaporating three times with 5 ml extra dry pyridine in an oven-dried 25-ml two-neck flask.

29. Dissolve in 2 ml extra dry pyridine and cool to $0°C$ on ice.

30. Add 0.46 g DMTrCl (1 eq, 1.36 mmol) and stir for 1.25 h at 0°C. Monitor by TLC using 1:5 (v/v) EtOAc/cyclohexane (Rf = 0.4).

31. Add 1 ml MeOH and stir for 10 min at 0°C.

32. Add 15 ml EtOAc and wash the organic phase three times with 20 ml saturated aqueous NaHCO$_3$.

33. Dry the organic phase with anhydrous Na$_2$SO$_4$ and evaporate under reduced pressure.

34. Purify product by flash chromatography using 12:87:1 (v/v/v) EtOAc/cyclohexane/TEA and monitor fractions by TLC as above (Rf = 0.4).

    The eluent should contain 1% TEA because the DMTr protecting group is acid-sensitive. This eluent should also be used to pretreat the silica resin by eluting 2 column volumes before the product is loaded on the column.

35. Combine fractions containing product and remove solvent under reduced pressure on a rotary evaporator and then vacuum pump.

    (S)-1-[Bis-(4-methoxy-phenyl)-phenyl-methoxy]-4-furan-2-yl-butane-2-ol (**13**) *is obtained as an oil (yield 66%).*

Perform phosphitylation to give phosphoramidite 14

36. Dry 220 mg **13** (0.48 mmol) by co-evaporating twice with 2 ml toluene and once with 2 mL freshly distilled CH$_2$Cl$_2$ in an oven-dried 25-ml round-bottom flask.

37. Dissolve in 10 mL dry CH$_2$Cl$_2$.

38. Add 0.9 ml DIPEA (11.25 eq, 5.4 mmol) and 269 μl 2-cyanoethyl N,N diisopropylchlorophosphoramidite (2.5 eq, 1.2 mmol) and stir for 2 h. Monitor by TLC using 1:7 (v/v) EtOAc/cyclohexane (Rf = 0.25).

39. Quench reaction by adding 20 ml saturated aqueous NaHCO$_3$.

40. Extract product with 20 ml CH$_2$Cl$_2$.

41. Dry the organic phase with anhydrous Na$_2$SO$_4$ and evaporate under reduced pressure.

42. Purify compound by flash chromatography using 1:7 (v/v) EtOAc/cyclohexane and monitor fractions by TLC as above (Rf = 0.25).

43. Combine fractions containing product and remove solvent under reduced pressure on a rotary evaporator and then vacuum pump.

    As the phosphoramidite is very sensitive, the purification should be performed

swiftly, using 1% TEA in the eluent and pretreating the silica as in step 34.
Diisopropyl-phosphoramidous   acid-1-[bis-(4-methoxy-phenyl)-phenyl-methoxy]-4-
*furan-2-yl-propyl ester 2-cyano-ethyl ester (**14**) is obtained as a colorless oil (yield*
79%).

A.3   BASIC PROTOCOL 2: Synthesis of furan-containing oligodeoxynucleotides

Phosphoramidites **7** and **14** are incorporated as a manual step in an otherwise standard automated ODN synthesis. Preparation of coupling solution, manual coupling, and postsynthetic workup steps compatible with the incorporated furan moiety are described here.

Materials

- Furan-modified phosphoramidite **7** (see Basic Protocol 1) or **14** (see Alternate Protocol)
- 4,5-Dicyanoimidazole (DCI)
- Acetonitrile (MeCN), extra dry, on molecular sieves
- Nitrogen and argon gas
- Controlled-pore glass (CPG) derivatized with desired 3' nucleoside (Proligo)
- Unmodified phosphoramidites, configured for DNA synthesizer (Proligo): DMTr-dA^{Bz}, DMTr-dC^{Bz}, DMTr-dGi^{Bu}, and DMTr-T
- Activated molecular sieves
- Reagents for automated ODN synthesis, configured for DNA synthesizer (Proligo): activator, oxidizer, cap A and B solutions, TCA deblock
- 50 mM aqueous ammonia (NH$_4$OH)
- DMTr-selective cartridge (Sep-pak from Waters)
- Rotary evaporator
- Vacuum pump (dry ice + isopropanol)
- Microcentrifuge tube
- Additional reagents and equipment for automated DNA synthesis (Curr. Protoc. Nucl. Acid Chem. APPENDIX 3C) and ODN purification using a DMTr-selective cartridge (UNIT 10.7)

NOTE: ODN synthesis should be performed under anhydrous conditions using oven-dried glassware, needles, and syringes, or new plastic syringes and needles that have been dried in a desiccator for 2 days.

Synthesize ODN

1. Evaporate crude phosphoramidite **7** or **14** with dry MeCN on a rotary evaporator and then dry overnight using a vacuum pump.

2. Prepare 0.5 ml of 1 M DCI activator solution (59 mg) per sequence in dry MeCN with molecular sieves and flush with argon.

3. Weigh the proper amount of CPG to obtain 1 μmol of the 3' nucleoside in a column and place the column in the synthesizer.

4. Follow standard setup procedures (diluting commercial phosphoramidites to 0.1 M in MeCN, priming lines, programming) for automated ODN synthesis (Curr. Protoc. Nucl. Acid Chem. APPENDIX 3C) on a 1-μmol scale, using the furan-modified phosphoramidite as the extra (fifth) phosphoramidite. Install a clean, dry, empty flask in the fifth phosphoramidite position, and prime this extra line with the other lines. Program synthesis to leave the trityl group attached to the last unmodified nucleoside.

5. Perform synthesis cycles using the unmodified phosphoramidites.

6. During automated synthesis, dissolve modified phosphoramidite at ∼0.06 M in dry MeCN. Transfer 0.4 ml (per sequence) to a dried flask and flush well with argon. No more than 30 min before incorporation, add 3 molecular sieves and flush again with argon.

7. When coupling of the modified phosphoramidite begins, interrupt the synthesizer and remove the column.

8. Take 0.4 ml modified phosphoramidite in a dry syringe and 0.5 ml DCI activator solution in another dry syringe. To perform manual coupling by the double-syringe technique, place an empty dry syringe at one end of the column and, from the other end, alternately push portions of amidite solution and activator solution over the column. Never push more than 1 column volume at once.
   Total coupling time is ∼20 min.
   *If desired, phosphoramidite **14** and the activator solution can be mixed in a single syringe and added simultaneously over the column, and manual capping (step 9) can be omitted. This is not recommended for **7**, because the coupling efficiency is very low. Aside from a lower reactivity caused by the sterical hindrance of the 2'-modification of **7**, the nucleophile present in the modified building block can attack the activated phosphoramidite in a side reaction, deactivating it. Coupling efficiency could be improved by adding the activator to the phosphoramidite in the*

presence of the growing ODN chain, rather than mixing activator and phosphor-
*amidite beforehand. An extra manual capping step is performed for **7**, because of*
the coupling yield, to avoid a deletion sequence. Regular capping steps are per-
formed for both on the synthesizer.

9. Using the same double-syringe method, rinse the column with 1 ml MeCN, then push 1 mL of 1:1 (v/v) cap A/cap B solutions over the column, and rinse again with 1 ml MeCN.

10. Return the column to the synthesizer and resume automated synthesis.

Deprotect, release, and purify ODN

11. Remove the column containing the completed ODN from the synthesizer and the CPG under a stream of nitrogen.

12. Pour the CPG into a microcentrifuge tube.

13. Deprotect and cleave the ODN from the CPG by adding 1 ml of 50 mM NH_4OH and incubating overnight at 55°C.

14. Purify the ODN over a DMTr-selective cartridge (UNIT 10.7).
    ODNs are best stored lyophilized at -20° C, but they can be stored for years in
    aqueous solution at -20° C.

A.4   BASIC PROTOCOL 3: Selective formation of DNA interstrand cross-links

Cross-link formation occurs upon oxidation of furan to the reactive 4-oxobutenal with NBS in water. Under these specific and mild oxidation conditions furan is quickly oxidized, while the DNA remains unaffected. The in-situ oxidation after hybridization, strongly reduces unspecific site reactions. This protocol describes the formation and analysis of ICL by RP-HPLC and denaturing PAGE.

Materials

- Furan-modified ODN (see Basic Protocol 2)
- Complementary ODN
- 100 mM (10x) phosphate buffer, pH 7
- 1 M (10x) NaCl
- Milli-Q-purified water
- N-Bromosuccinimide (NBS)

- Acetonitrile (MeCN), HPLC grade, >99.9
- 0.1 M triethylammonium acetate (TEAA) buffer, pH 7
- Formamide
- GelRed stock solution (Biotium): 10,000x in dimethyl sulfoxide (DMSO)
- Microcentrifuge tube
- Eppendorf thermomixer
- Additional reagents and equipment for RP-HPLC (UNIT 10.5) and PAGE (Curr. Protoc. Nucl. Acid Chem. UNIT 10.4 & Curr. Protoc. Nucl. Acid Chem. APPENDIX 3B)

Activate furan-mediated cross-link formation

1. In a microcentrifuge tube, prepare a 50-μl ODN solution as follows: 1 nmol furan-modified ODN, 1 nmol complementary ODN, 5 μl 100 mM phosphate buffer, pH 7 (final 10 mM), 5 μl 1 M NaCl (final 100 mM), Milli-Q-purified water to 50 μl.

2. Place in a thermomixer and heat from room temperature to 95°C, thenmaintain 95°C for an additional 10 min.

3. Turn off the thermomixer and allow sample to cool to room temperature over at least 2 h.
   Better annealing is achieved by turning off the thermomixer rather than removing the sample, as this allows the sample to cool as slowly as possible.

4. If desired, remove several μl as reference material for later analysis by RP-HPLC or PAGE.

5. Prepare a fresh solution of 500 μM NBS in Milli-Q-purified water.

6. While holding the annealed sample at 25°C in the thermomixer, add 2 μl freshly prepared NBS (1 eq) with shaking at 950 rpm. Repeat for a total of four additions at 15-min intervals.
   If samples were removed for analysis, the volume of NBS should be adjusted to account for the reduced amount of furan-modified ODN.

Analyze cross-link formation by HPLC

7. Take 2 μl reaction mixture and dilute to 20 μl with Milli-Q-purified water in an HPLC vial.

8. Analyze by RP-HPLC (UNIT 10.5) on a C18 column with a gradient of 5% to 30% MeCN in 100 mM TEAA buffer over 15 min at 50°C and UV detection at 260 nm.

Analyze cross-link formation by denaturing PAGE

9. Mix 2 μl reaction mixture with 2 μl formamide as loading buffer and analyze by denaturing PAGE (Curr. Protoc. Nucl. Acid Chem. UNIT 10.4 & Curr. Protoc. Nucl. Acid Chem. APPENDIX 3B).

   Specific conditions depend on the length of ODN. For 12-mer ODNs, use a 20 x 20-cm 20% polyacrylamide gel (19:1 acrylamide/bisacrylamide) with 7 M urea and run at 120 V overnight at 25° C.

10. Stain the gel for 30 min with shaking in 100 ml water containing 10 μl GelRed stock solution. Visualize ODNs by illuminating the gel with UV light in a black box.

    CAUTION: GelRed is an intercalator and is therefore toxic. Careful handling with appropriate gloves is required.

Appendix B

NMR study of the Sequence Controlled Oligomers

Samples Two pairs of structurally related oligomeric sequences (Table B.1) were prepared using the microwave-assisted protocol (vide supra). The first pair of trimers A I & A II only differ in the second amine residue as a different amine (2-methoxyethylamine in A I vs benzylamine in A II) was used in the second iterative cycle. The second pair consists of a trimer B I and a tetramer B II; an additional aminolysis (furfurylamine)/chain extension cycle starting from B I resulted in the formation of B II. The crude samples were purified by preparative reversed-phase HPLC. The freeze-dried fractions were consequently subjected to NMR and HRMS analysis (vide infra). Oligomers A I & A II (1.3 mg each) were dissolved in 650 μL DMSO-d6 in a 5 mm-NMR tube, prior to analysis using the Bruker 500 MHz Avance II spectrometer. Oligomers B I & B II (1.5 mg each) were dissolved in 17 μL DMSO-d6 an transferred to a 1 mm-NMR tube, prior to analysis using the Bruker 700 MHz Avance II spectrometer.

NMR experiments (performed by Dr. P. Espeel) Each sample was analyzed by 1D-^{1}H- and 2D-NMR (gCOSY, TOCSY, ^{1}H-^{13}C multiplicity-edited gHSQC, ^{1}H-^{13}C gHMBC and NOESY). In the following, the prefix g, signifying gradient-enhanced sequences, will be dropped. NMR experiments were performed either at 500 MHz or 700 MHz, both using a Bruker Avance II spectrometer equipped with a 5 mm TXI-Z ^{1}H,^{13}C,^{31}P probe, respectively a 1 mm TXI-Z 1H,13C,15N probe. The latter affords a 5-fold higher mass sensitivity compared to a 5 mm

343

Table B.1: Structural information of two pairs of related oligomeric sequences, trimers **A I, A II, B I** & tetramer **B II**, subjected to NMR and HRMS analysis.

Code	Structure	Formula	MW (g/mol)
A I		$C_{45}H_{54}FN_7O_{12}S_4$	1041.3
A II		$C_{49}H_{54}FN_7O_{11}S_4$	1073.4
B I		$C_{47}H_{74}N_8O_{12}S_4$	1070.4
B II		$C_{59}H_{50}N_{10}O_{15}S_5$	1338.5

probe at the same field. Used at 700 MHz, this affords a 15-fold increase in signal to noise compared to the 5 mm probe at 500 MHz, in addition to the obvious improvement in resolution. The sensitivity advantage proved most interesting to obtain 2D NOESY spectra of sufficient quality to support the assignment process, given that limited quantities of material were available and that the nOe cross-peaks generally consist at most of about 5% of the diagonal peak intensity. All measurements were performed at 25°C and used TopSpin 3.1 for acquisition and processing of all data. All spectra including 1D-^{1}H, 2D gCOSY, TOCSY (80 ms), NOESY (300 ms), ^{1}H-^{13}C gHSQC and ^{1}H-^{13}C gHMBC (optimized for nJCH of 8 Hz) were recorded using in a standard fashion with pulse programs available in the Bruker library.

Data analysis and peak assignment (performed by Dr. P. Espeel) The 1D-^{1}H NMR spectra of **A I, A II, B I** & **B II** are shown respectively in Figure B.1, B.2, B.3 and B.4.

The analysis of the obtained 2D-NMR data is illustrated by a selection of plots (zoom) of the relevant, well-resolved spectra.

A detail of the 1D-^{1}H NMR spectra of **A I, A II, B I** & **B II** (8.5 to 6.0 ppm) with assignment of the amide N-H and aromatic signals of the amine residues is shown in Figure B.5 and B.6.
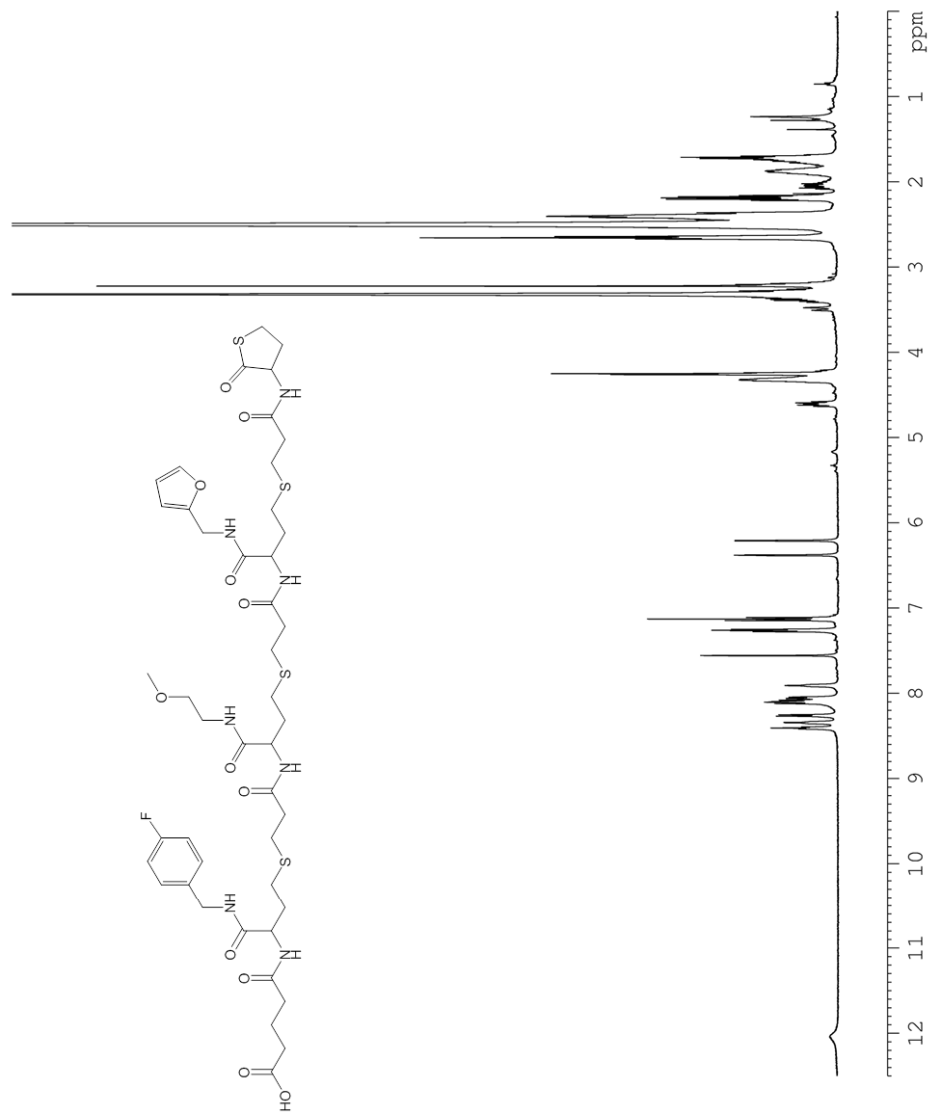
Figure B.1: 1D-^{1}H-NMR spectrum (500 MHz, DMSO-d6) of trimer sequence **A I**.
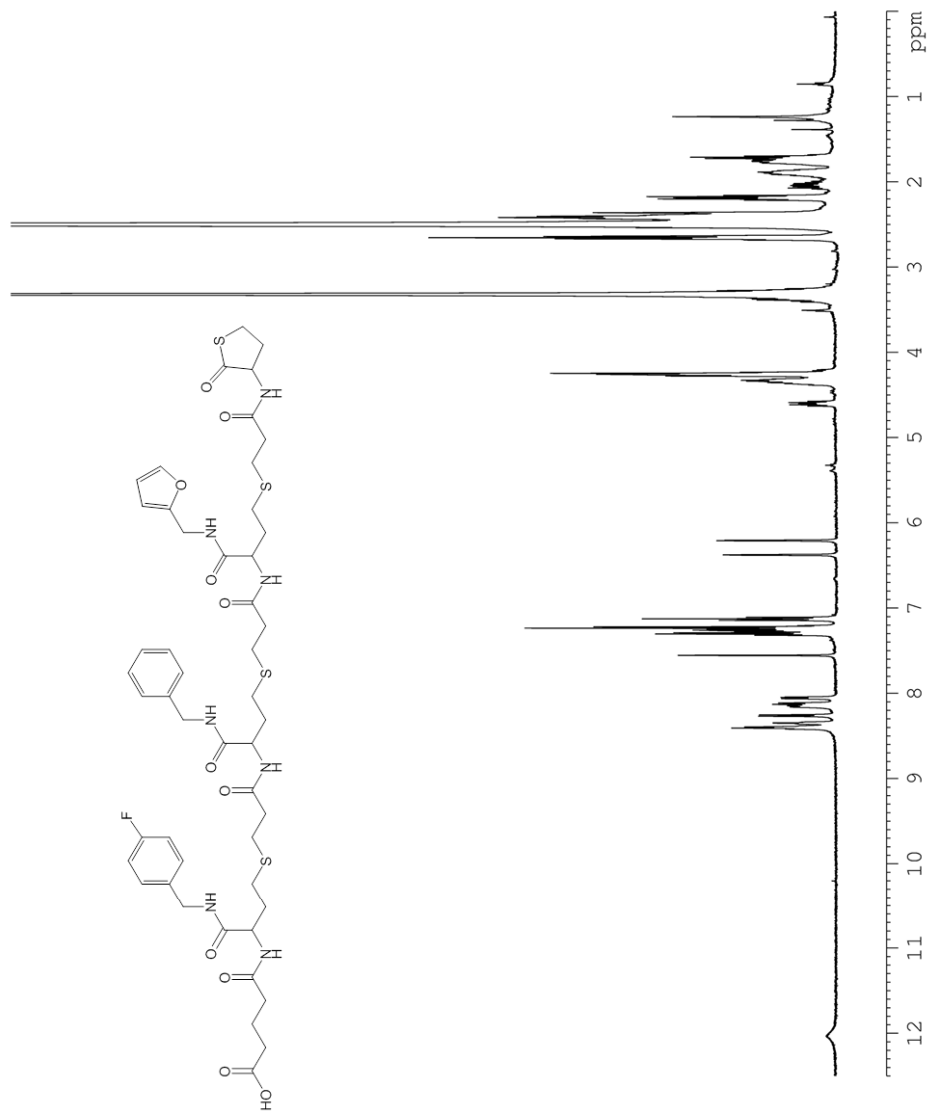
Figure B.2: 1D-^{1}H-NMR spectrum (500 MHz, DMSO-d6) of trimer sequence **A II**.
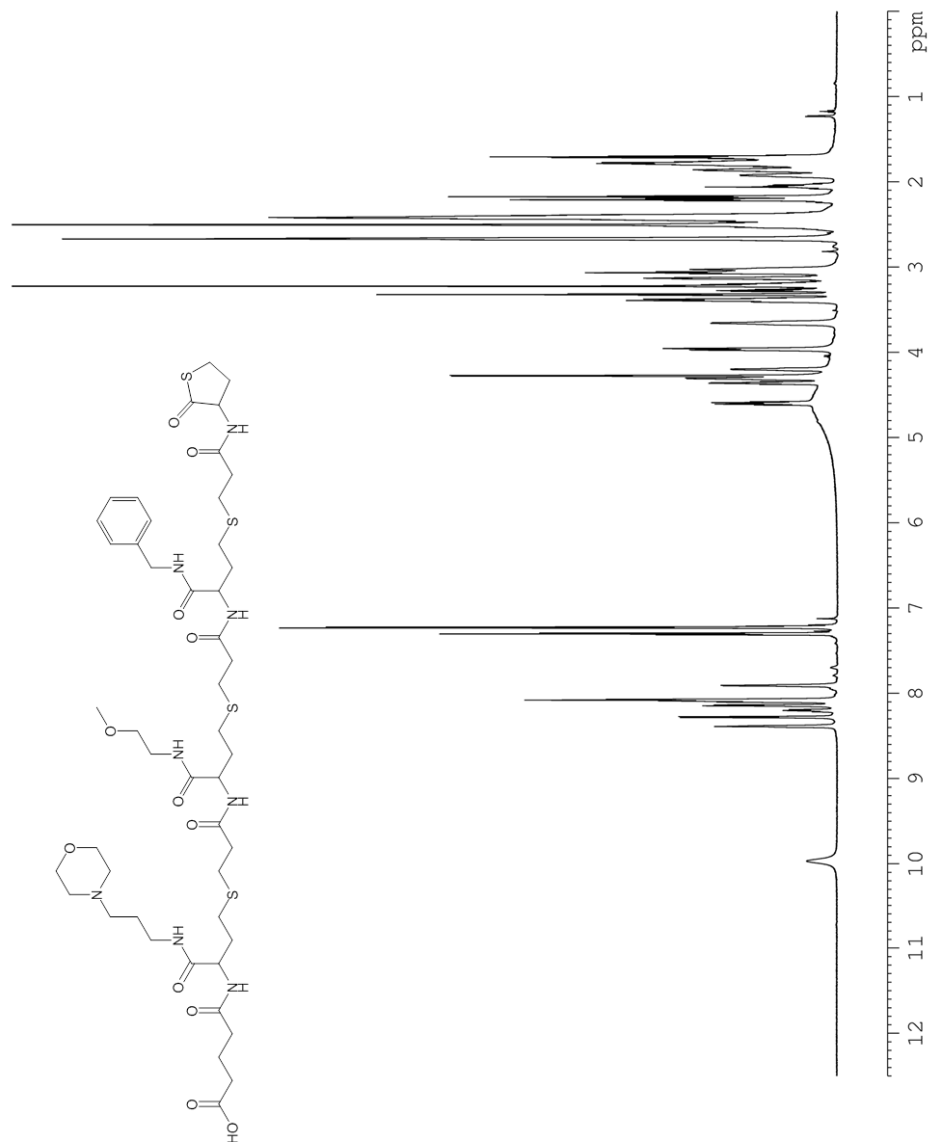
Figure B.3: 1D-^{1}H-NMR spectrum (500 MHz, DMSO-d6) of trimer sequence **B I**.
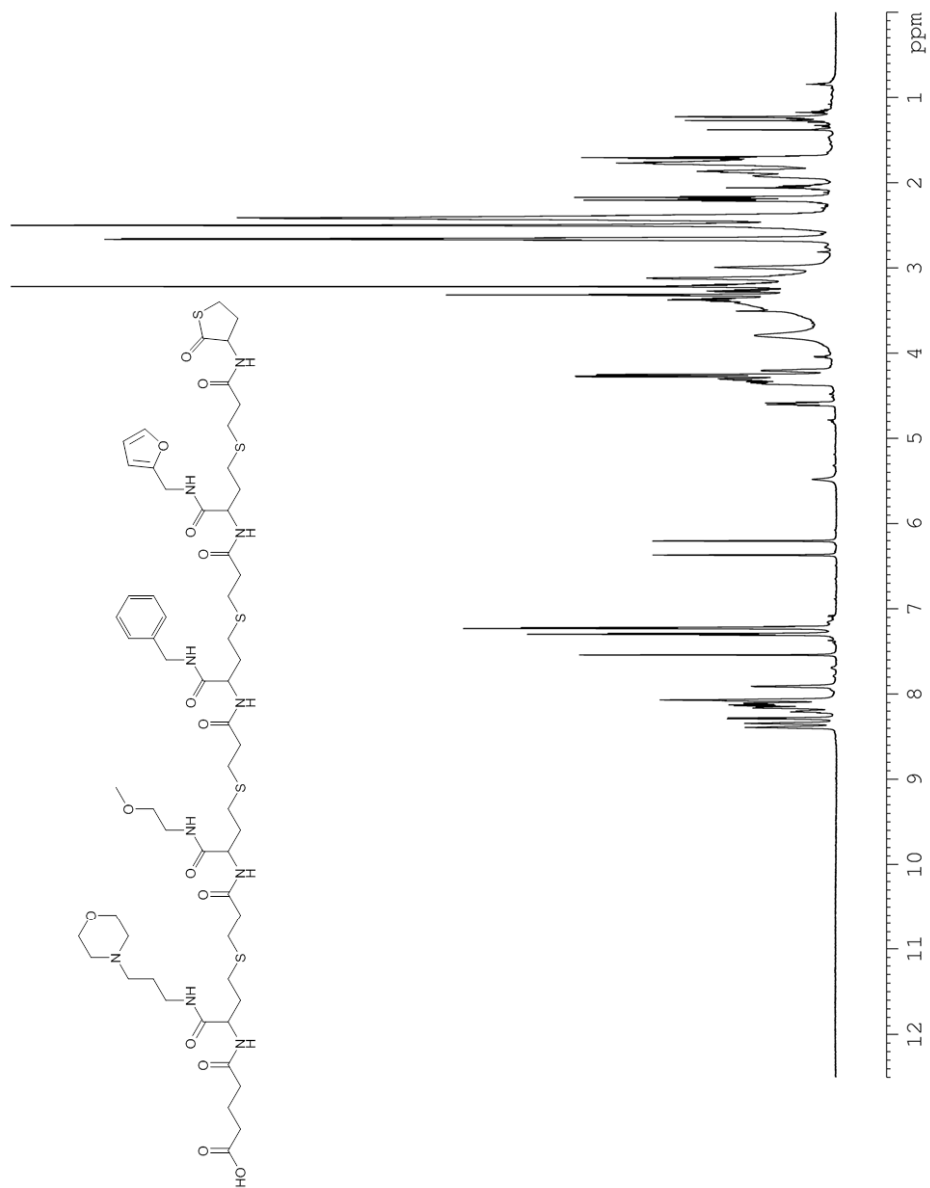
Figure B.4: 1D-^{1}H-NMR spectrum (500 MHz, DMSO-d6) of tetramer sequence **B II**.
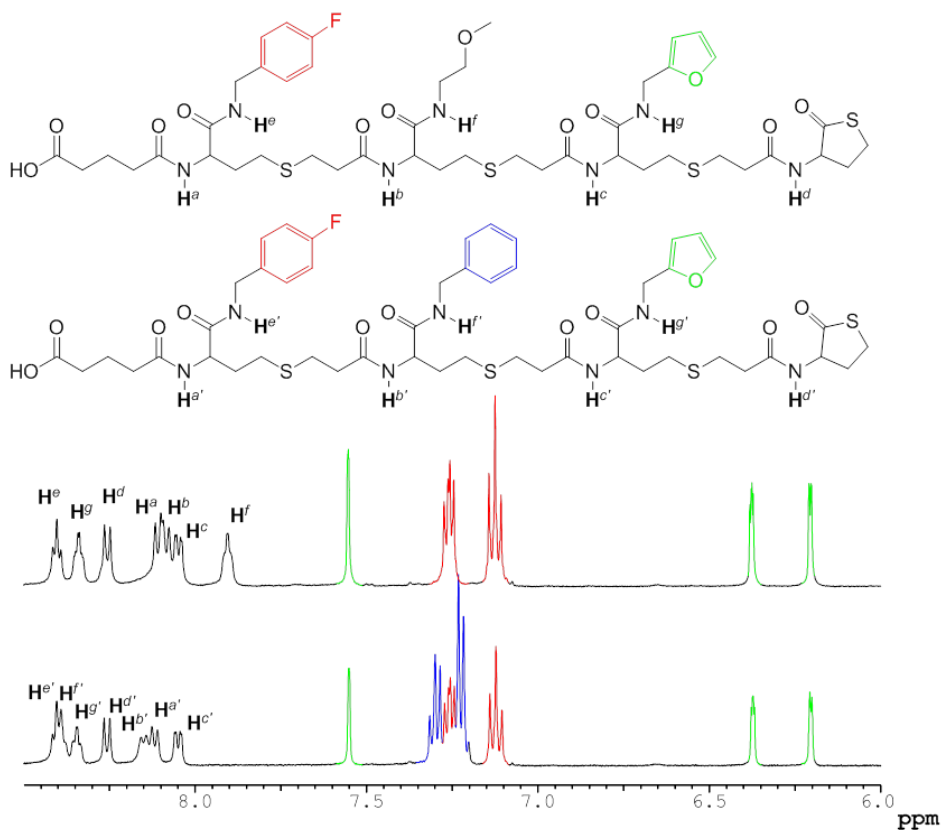
Figure B.5: Details (8.5 to 6 ppm) of 1D-^{1}H-NMR spectrum (500 MHz, DMSO-d6) of **A I** (top) and **A II** (bottom).
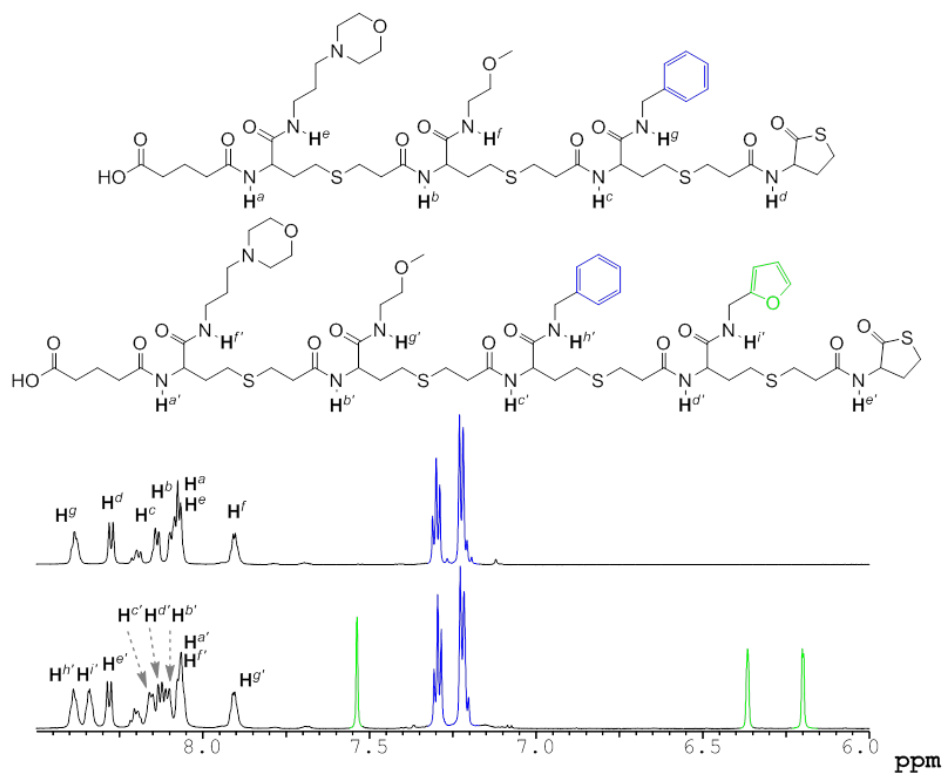
Figure B.6: Details (8.5 to 6 ppm) of 1D-^{1}H-NMR spectrum (500 MHz, DMSO-d6) of **B I** (top) and **B II** (bottom).

At first glance, the aromatic protons (Figure B.5 and B.6) can be assigned by comparison of the obtained spectra. More importantly, each sequence is characterized by a set of typical amide N-H signals (H^i in Figure B.5 and B.6), for which only tentative assignment is possible by comparison of the 1D-^{1}H NMR spectra. Additionally, stacked plots of the respective gCOSY and TOCSY spectra (zoom) of both **A I** & **A II** (Fig. B.7 and B.8) reveal a clear distinction between the amide N-H signals in the backbone (H^a, H^b and H^c), at the thiolactone chain end (H^d) and those in the side chains (H^e, H^f and H^g). Moreover, side chain amide N-H signals, originating from aromatic amines, have a higher chemical shift than those from aliphatic amines. Similar conclusions can be made from the corresponding COSY/TOCSY spectra of **B I** & **B II** (not shown here).

Despite the high degree of spectral crowding in the high-field region, stacking the COSY and TOCSY spectra of **B I** (Fig. B.9) facilitates further assignment of overlapping signals. Furthermore, 5 well-resolved nOe contacts, involving amide protons, in **B I** (Figure S21), allow for the connection of the specific amine residues to the backbone (nOe 4 and 5, Fig. B.10). Additionally, nOe contacts, bridging the amide C=O bonds in the backbone, even enable partial elucidation of the oligomeric sequence, as nOe 1 connects the amide NH, Ha, to a CH2 signal of the linker (Fig. B.10). Unfortunately, nOe 2 and 3 cannot be unambiguously attributed to fully confirm the sequential build-up via NMR.

Due to the poor signal-to-noise ratio in the APT spectra, the ^{13}C-chemical shift values were extracted from the corresponding HSQC and HMBC spectra, the latter for quaternary carbons. The carbons in aromatic residues can be readily assigned from an overlay of the HSQC and HMBC spectra, as demonstrated for the sequence A I (Fig. B.11, left). The amide C=O signals (Ī71 ppm) are linked to the specific amide N-H signals in the well-resolved HMBC spectrum (Fig. B.11, right).

The remaining C=O carbons at both ends of the sequence (COOH and thiolactone unit) can be assigned in a similar manner (B I, Figure S23). Furthermore, combining the COSY/TOCSY spectra (Figure S20) and the gHSQC spectrum (Figure S24) of B I allows for the assignment of all (CH)-, (CH$_2$)- and (CH$_3$)- carbons in the high-field region.

This in-depth study of the obtained 2D-NMR spectra enables full characterization of all obtained samples, as for each sequence, the chemical shift of all ^1H- and ^{13}C-atoms was determined (Table B.3 to B.10), despite the spectral crowding in the high-field region.
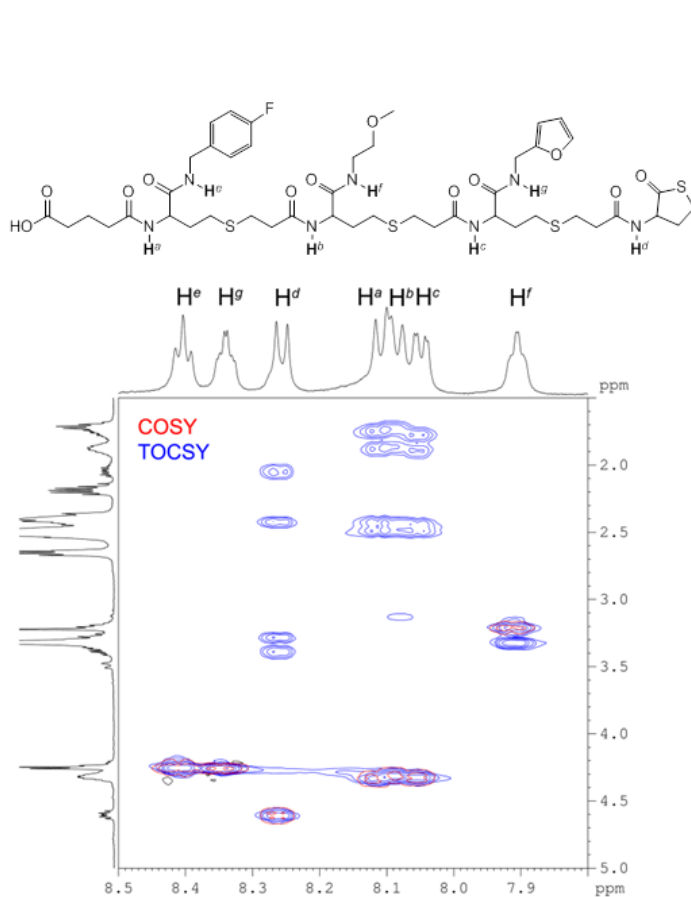
Figure B.7: Stacked plots of the COSY (red) and TOCSY (blue) spectra (500 MHz, DMSO-d6) of **A I** reveal a clear distinction between the amide N-H signals in the backbone (H^a, H^b and H^c), at the thiolactone chain end (H^d) and those in the side chains (H^e, H^f and H^g).
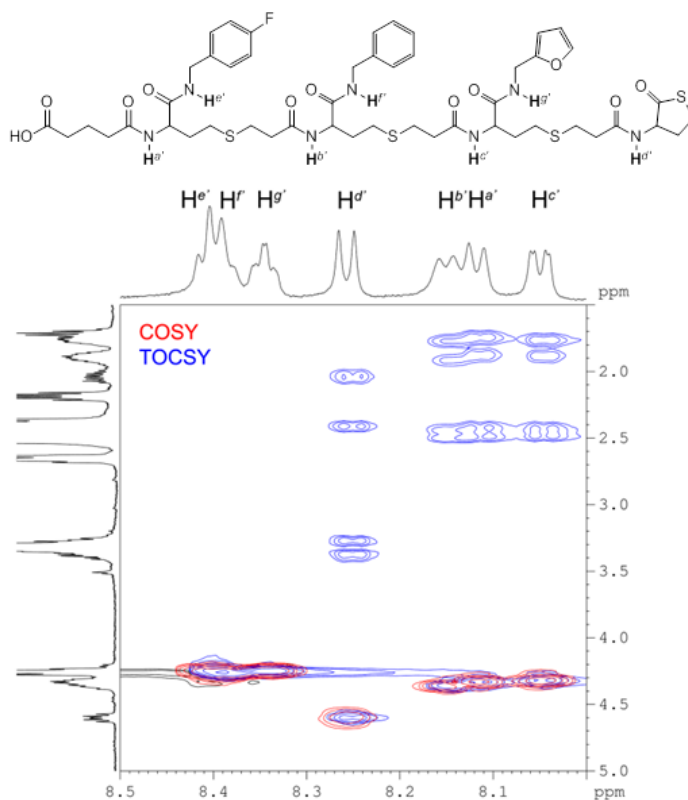
Figure B.8: Stacked plots of the COSY (red) and TOCSY (blue) spectra (500 MHz, DMSO-d6) of **A II** (right) reveal a clear distinction between the amide N-H signals in the backbone (H^a, H^b and H^c), at the thiolactone chain end (H^d) and those in the side chains (H^e, H^f and H^g).
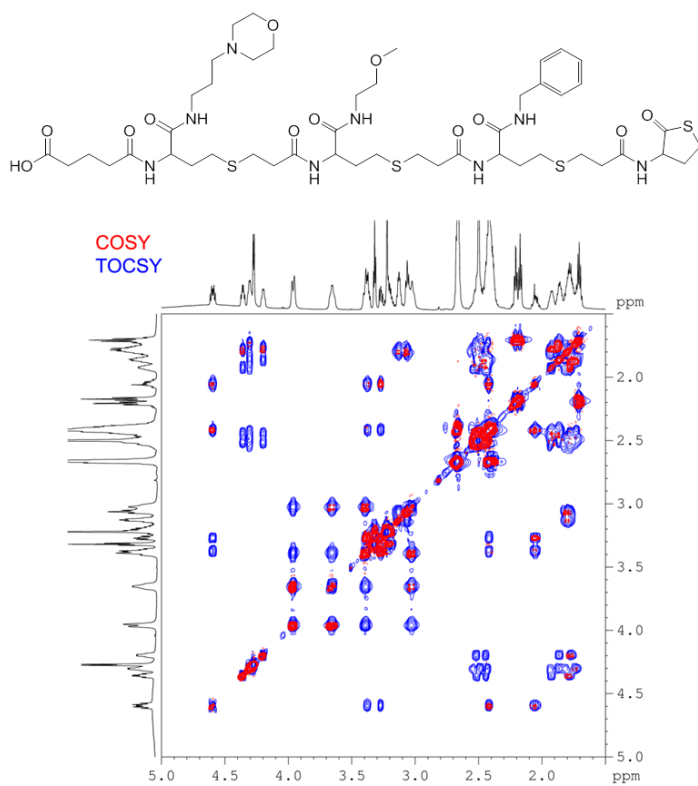
Figure B.9: Stacked plot of the COSY (red) and TOCSY (blue) spectra (700 MHz, DMSO-d6) of **B I** to facilitate the assignment of high-field protons. (5 to 1.5 ppm).
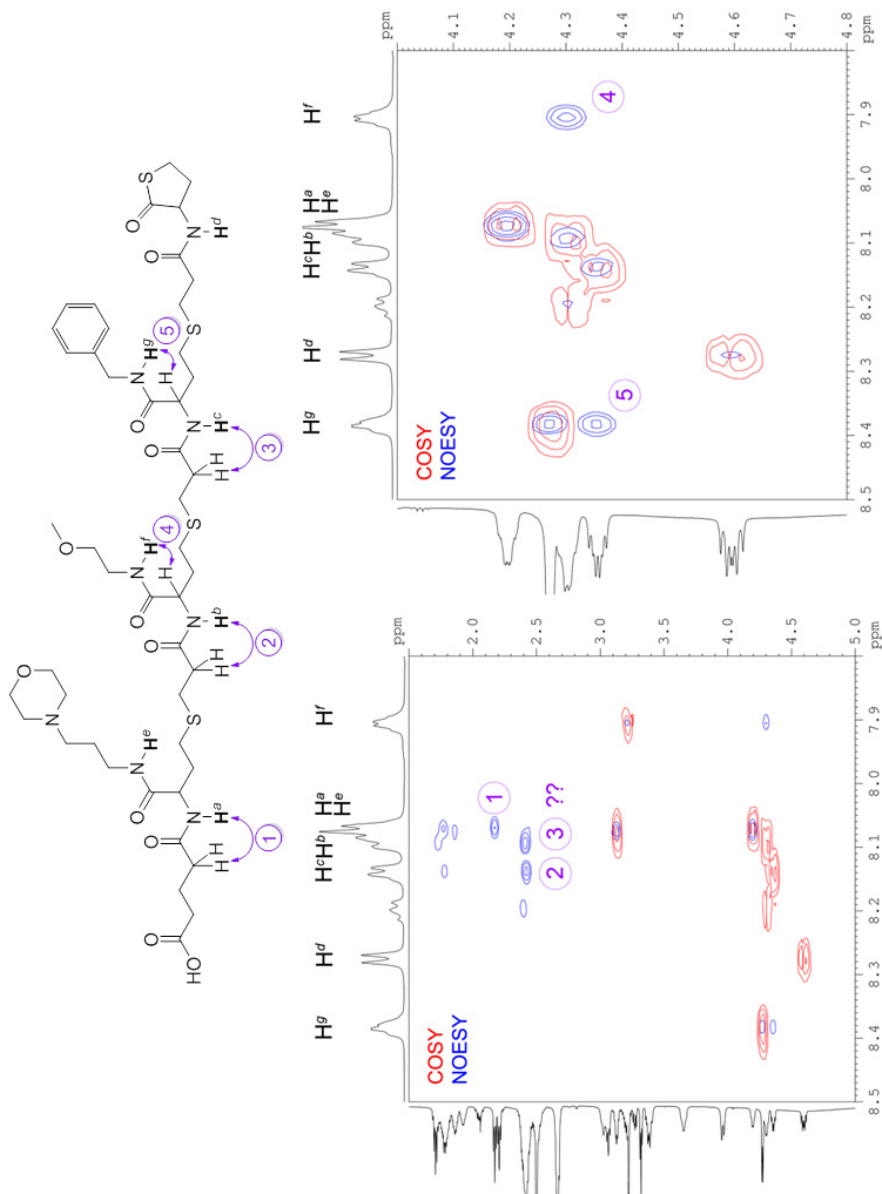
Figure B.10: Stacked plot of the COSY (red) and the NOESY (blue) spectra (700 MHz, DMSO-d6) of **B I**. 5 nOe Contacts allow for the connection between side chain residues and the backbone and the partial elucidation of the sequential build-up.
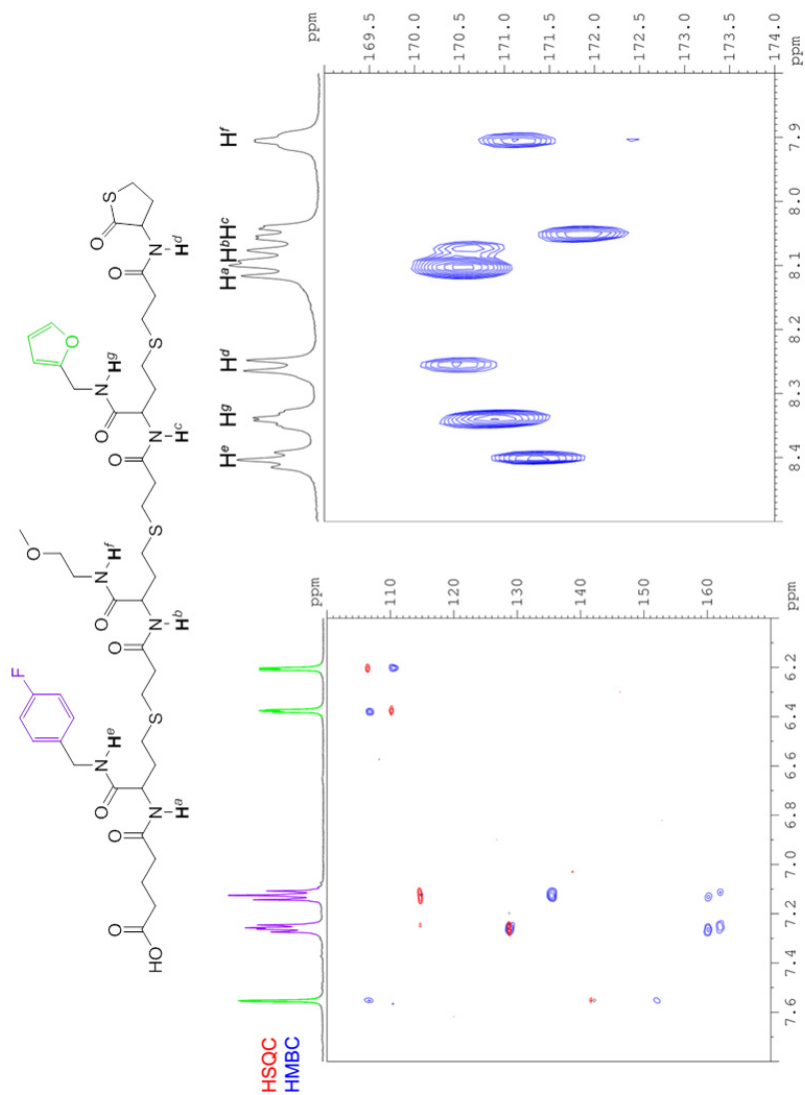
Figure B.11: HSQC (red) and HMBC (blue) spectra (500 MHz, DMSO-d6) of **A I**. The carbons in aromatic residues can be assigned from an overlay of the HSQC and HMBC spectra (left) and the amide C=O signals are linked to the specific amide N-H signals (right).
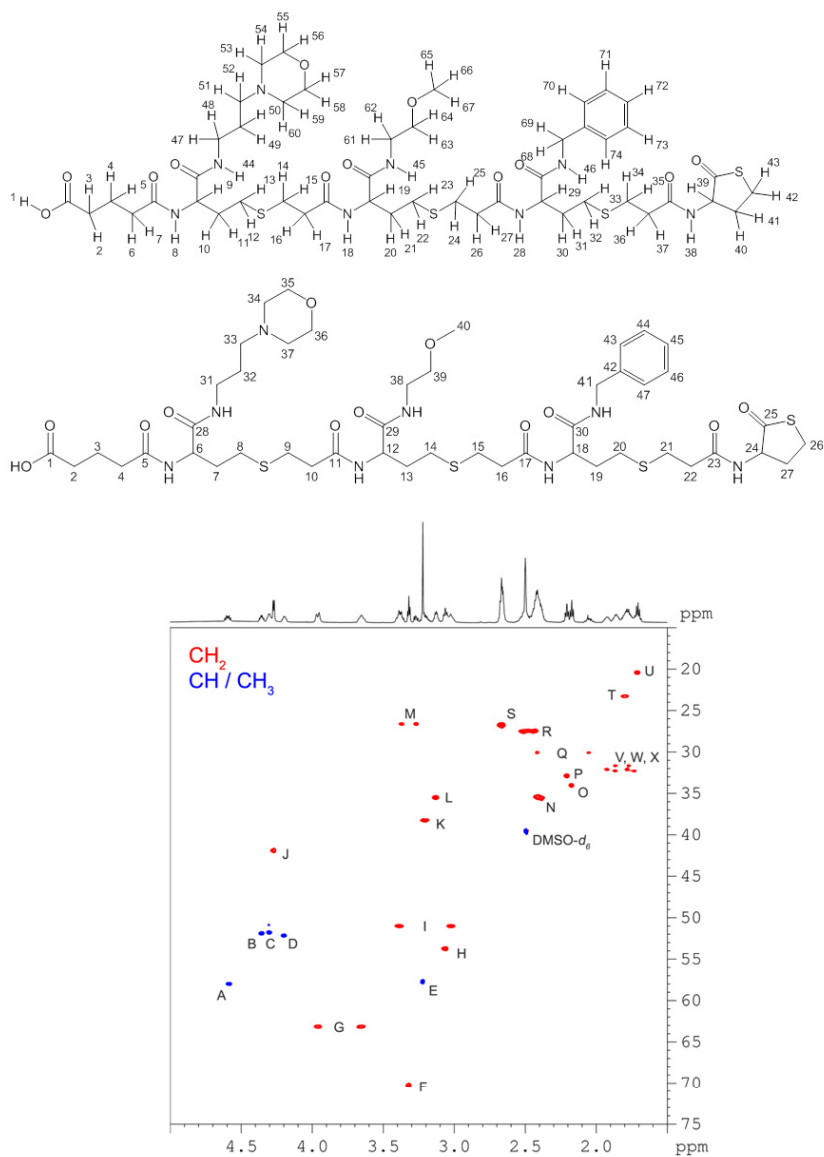
Figure B.12: High-field region of the HSQC spectrum showing multiplicity editing (700 MHz, DMSO-d6) of **B I**

Table B.2: Determination of the ^{1}H- and ^{13}C- chemical shift values of **B I**

Peak	δ (^{1}H, ppm)	δ (^{13}C, ppm)	Assignment	
			H	C
A	4.60	58.0	*39*	*24*
B	4.36	51.8	*29*	*18*
C	4.30	51.8	*19*	*12*
D	4.19	52.1	*9*	*6*
E	3.22	57.6	*55/56/57*	*40*
F	3.33	70.3	*63/64*	*39*
G	3.96 3.65	63.1	*55/56 57/58*	*35/36*
H	3.06	53.7	*51/52*	*33*
I	3.39 3.02	51.0	*53/54 59/60*	*34/37*
J	4.27	41.9	*68/69*	*41*
K	3.21	38.2	*61/62*	*38*
L	3.13	35.5	*47/48*	*31*
M	3.37 3.27	26.5	*42/43*	*26*
N	2.41	35.5	*16/17 26/27 36/37*	*10/16/22*
O	2.17	34.0	*6/7*	*4*
P	2.20	32.9	*2/3*	*2*
Q	2.41 2.05	30.0	*40/41*	*27*
R	2.50 2.44	27.5	*12/13 22/23 32/33*	*8/14/20*
S	2.67	26.6	*14/15 24/25 34/35*	*9/15/21*
T	1.80	23.3	*49/50*	*32*
U	1.71	20.4	*4/5*	*3*
V	1.86 1.76	31.7	*10/11*	*7*
W	1.92 1.77	32.3	*30/31*	*19*
X	1.85 1.72	32.3	*20/21*	*13*

Table B.3: [1]H - chemical shifts of trimer **A I** (500 MHz, DMSO-d6).

H	δ (ppm)		H	δ (ppm)		H	δ (ppm)
1	12.04		24/25	2.65		46	8.34
2/3	2.20		26/27	2.40		47/48	4.25
4/5	1.71		28	8.05		49/52	7.26
6/7	2.18		30/31	1.87/1.74		50/51	7.12
8	8.10		32/33	2.46		53/54	3.20
9/19/29	4.32		34/35	2.65		55/56	3.32
10/11	1.87/1.74		36/37	2.40		57/58/59	3.22
12/13	2.46		38	8.25		60/61	4.25
14/15	2.65		39	4.60		62	6.20
16/17	2.40		40/41	2.41/2.04		63	6.37
18	8.08		42/43	3.37/3.27		64	7.55
20/21	1.87/1.74		44	8.40			
22/23	2.46		45	7.90			

Table B.4: [13]C - chemical shifts of trimer **A I** (125 MHz, DMSO-d6).

C	d (ppm)		C	d (ppm)		C	d (ppm)
1	174.0		24	57.9		39	70.2
2	32.8		25	205.2		40	57.7
3	20.4		26	26.5		41	35.4
4	34.1		27	30.0		42	152.1
5	172.1		28	171.4		43	106.5
6/12/18	51.7		29	171.1		44	110.1
7/13/19	32.1		30	170.9		45	141.5
8/14/20	27.4		31	41.1			
9/15/21	26.6		32	135.5			
10/16/22	35.3		33/37	128.7			
11	170.6		34/36	114.7			
17	171.9		35	160.1/162.0			
23	170.5		38	38.2			

Table B.5: ^{1}H - chemical shifts of trimer **A II** (500 MHz, DMSO-d6).

H	δ (ppm)
1	12.03
2/3	2.20
4/5	1.71
6/7	2.18
8	8.11
9/19/29	4.34
10/11	1.89/1.76
12/13	2.46
14/15	2.66
16/17	2.41
18	8.15
20/21	1.89/1.76
22/23	2.46

H	δ (ppm)
24/25	2.66
26/27	2.41
28	8.05
30/31	1.89/1.76
32/33	2.46
34/35	2.66
36/37	2.41
38	8.26
39	4.60
40/41	2.42/2.05
42/43	3.38/3.28
44	8.41
45	8.39

H	δ (ppm)
46	8.34
47/48	4.25
49/52	7.26
50/51	7.12
53/54	4.27
55/57/59	7.22
56/58	7.30
60/61	4.26
62	6.20
63	6.38
64	7.55

Table B.6: ^{13}C - chemical shifts of trimer **A II** (125 MHz, DMSO-d6).

C	δ (ppm)
1	174.3
2	32.8
3	20.4
4	34.0
5	172.0
6/12/18	51.7
7/13/19	31.9
8/14/20	27.4
9/15/21	26.6
10/16/22	35.3
11/17	170.7
23	170.8
24	57.8

C	δ (ppm)
25	???
26	26.5
27	30.0
28/29	171.1
30	171.0
31	41.0
32	135.6
33/37	128.6
34/36	114.6
35	160.2/162.2
38	41.8
39	139.3
40/42/44	126.7

C	δ (ppm)
41/43	127.9
45	35.3
46	152.2
47	106.4
48	110.2
49	141.6

Table B.7: ^{1}H - chemical shifts of trimer **B I** (700 MHz, DMSO-d6).

H	δ (ppm)	H	δ (ppm)	H	δ (ppm)
1	9.96	22/23	2.48/2.43	45	7.91
2/3	2.20	24/25	2.67	46	8.38
4/5	1.71	26/27	2.41	47/48	3.13
6/7	2.17	28	8.14	49/50	1.80
8/44	8.07	29	4.36	51/52	3.06
9	4.19	30/31	1.92/1.77	53/54/59/60	3.39/3.02
10/11	1.86/1.76	32/33	2.50/2.44	55/56/57/58	3.96/3.65
12/13	2.52/2.44	34/35	2.67	61/62	3.21
14/15	2.67	36/37	2.41	63/64	3.33
16/17	2.41	38	8.27	65/66/67	3.22
18	8.09	39	4.60	68/69	4.27
19	4.30	40/41	2.41/2.05	70/72/74	7.22
20/21	1.85/1.72	42/43	3.37/3.27	71/73	7.30

Table B.8: ^{13}C - chemical shifts of trimer **B I** (175 MHz, DMSO-d6).

C	δ (ppm)	C	δ (ppm)	C	δ (ppm)
1	174.1	17	170.7	34/37	51.0
2	32.9	18	51.8	35/36	63.1
3	20.4	23	170.7	38	38.2
4	34.0	24	58.0	39	70.3
5	172.1	25	205.4	40	57.6
6	52.1	26	26.5	41	41.9
7	31.7	27	30.0	42	139.3
8/14/20	27.5	28	172.1	43/45/47	126.8
9/15/21	26.6	29	171.2	44/46	128.0
10/16/22	35.5	30	171.2		
11	170.6	31	35.5		
12	51.8	32	23.3		
13/19	32.3	33	53.7		

Table B.9: ¹H - chemical shifts of tetramer **B II** (700 MHz, DMSO-d6).

H	δ (ppm)		H	δ (ppm)		H	δ (ppm)		H	δ (ppm)
1	-		*22/23*	2.48/2.43		*46/47*	2.42		*68/69*	3.79
2/3	2.20		*24/25*	2.67		*48*	8.28		*70/71*	3.12
4/5	1.72		*26/27*	2.42		*49*	4.59		*72/73*	3.20
6/7	2.17		*28*	8.16		*50/51*	2.41/2.04		*74/75*	3.32
8/54	8.07		*29/39*	4.35		*52/53*	3.37/3.27		*76/77/78*	3.22
9	4.20		*30/31*	1.92/1.77		*55*	7.91		*79/80*	4.27
10/11	1.86/1.77		*32/33*	2.51/2.41		*56*	8.39		*81/83/85*	7.21
12/13	2.50/2.44		*34/35*	2.67		*57*	8.34		*82/84*	7.30
14/15	2.67		*36/37*	2.42		*58/59*	3.12		*86/87*	4.25
16/17	2.42		*38*	8.13		*60/61*	1.79		*88*	6.20
18	8.11		*40/41*	1.88/1.74		*62/63*	2.99		*89*	6.37
19	4.30		*42/43*	2.49/2.42		*64/65*	3.12		*90*	7.55
20/21	1.86/1.73		*44/45*	2.67		*66/67*	3.79			

Table B.10: ¹³C - chemical shifts of tetramer **B II** (175 MHz, DMSO-d6).

C	δ (ppm)		C	δ (ppm)		C	δ (ppm)		C	δ (ppm)
1	174.2		*17*	170.8		*38*	35.5		*56*	152.2
2	32.8		*18/24*	51.8		*39*	23.3		*57*	106.4
3	20.3		*23*	170.8		*40*	53.8		*58*	110.1
4	34.0		*25*	31.1		*41/44*	51.1		*59*	142.0
5	172.1		*29*	170.7		*42/43*	63.3			
6	52.0		*30*	57.9		*45*	38.2			
7	31.7		*31*	205.4		*46*	70.1			
8/14/20/26	27.4		*32*	26.5		*47*	57.7			
9/15/21/27	26.7		*33*	30.0		*48*	41.8			
10/16/22/28	35.4		*34*	172.1		*49*	139.3			
11	170.8		*35*	171.3		*50/52/54*	126.7			
12	51.7		*36*	171.3		*51/53*	128.0			
13/19	32.1		*37*	171.1		*55*	35.4			

References

[1] Darnell, J. E. (2002) Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer 2*, 740–749.

[2] Brennan, P., Donev, R., and Hewamana, S. (2008) Targeting transcription factors for therapeutic benefit. *Mol. Biosyst. 4*, 909–919.

[3] Strausberg, R. L., and Schreiber, S. L. (2003) From knowing to controlling: A path from Genomics to drugs using small molecule probes. *Science 300*, 294–295.

[4] Reayi, A., and Arya, P. (2005) Natural product-like chemical space: search for chemical dissectors of macromolecular interactions. *Curr. Opin. Chem. Biol. 9*, 240–247.

[5] Arndt, H. D. (2006) Small molecule modulators of transcription. *Angew. Chem. Int. Ed. 45*, 4552–4560.

[6] Gewirtz, A. M. (2007) On future's doorstep: RNA interference and the pharmacopeia of tomorrow. *J. Clin. Invest. 117*, 3612–3614.

[7] Karkare, S., and Bhatnagar, D. (2006) Promising nucleic acid analogs and mimics: characteristic features and applications of PNA, LNA, and morpholino. *Appl. Microbiol. Biotechnol. 71*, 575–586.

[8] Boger, D. L., Desharnais, J., and Capps, K. (2003) Solution-phase combinatorial libraries: Modulating cellular signaling by targeting protein-protein or protein-DNA interactions. *Angew. Chem. Int. Ed. 42*, 4138–4176.

[9] Sinha, N. D., and Michaud, D. P. (2007) Recent developments in the chemistry, analysis and control for the manufacture of therapeutic-grade synthetic oligonucleotides. *Curr. Opin. Drug Discovery Dev. 10*, 807–818.

[10] Collins, F. S., Lander, E. S., Rogers, J., Waterston, R. H., and Conso, I. H. G. S. (2004) Finishing the euchromatic sequence of the human genome. *Nature 431*, 931–945.

[11] Reiss, T. (2001) Drug discovery of the future: the implications of the human genome project. *Trends Biotechnol. 19*, 496–499.

[12] Altmann, K. H. et al. (2009) The State of the Art of Chemical Biology. *Chembiochem 10*, 16–29.

[13] Dobson, C. M. (2004) Chemical space and biology. *Nature 432*, 824–828.

[14] Rademann, J. (2004) Organic protein chemistry: Drug discovery through the chemical modification of proteins. *Angew. Chem. Int. Ed. 43*, 4554–4556.

[15] Drews, J. (2006) What's in a number? *Nat. Rev. Drug Discovery 5*, 975–975.

[16] White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., and Dervan, P. B. (1998) Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature 391*, 468–471.

[17] Vazquez, O., Blanco-Canosa, J. B., Vazquez, M. E., Martinez-Costas, J., Castedo, L., and Mascarenas, J. L. (2008) Efficient DNA Binding and Nuclear Uptake by Distamycin Derivatives Conjugated to Octa-arginine Sequences. *Chembiochem 9*, 2822–2829.

[18] Mapp, A. K., Ansari, A. Z., Ptashne, M., and Dervan, P. B. (2000) Activation of gene expression by small molecule transcription factors. *Proc. Natl. Acad. Sci. U. S. A. 97*, 3930–3935.

[19] Bando, T., and Sugiyama, H. (2006) Synthesis and biological properties of sequence-specific DNA-alkylating pyrrole-imidazole polyamides. *Acc. Chem. Res. 39*, 935–944.

[20] Gniazdowski, M., Denny, W. A., Nelson, S. M., and Czyz, M. (2005) Effects of anticancer drugs on transcription factor-DNA interactions. *Expert Opin. Ther. Targets 9*, 471–489.

[21] Bremer, R. E., Baird, E. E., and Dervan, P. B. (1998) Inhibition of major-groove-binding proteins by pyrrole-imidazole polyamides with an Arg-Pro-Arg positive patch. *Chem. Biol. 5*, 119–133.

[22] Moretti, R., and Ansari, A. Z. (2008) Expanding the specificity of DNA targeting by harnessing cooperative assembly. *Biochimie 90*, 1015–1025.

[23] Vazquez, M. E., Caamano, A. M., and Mascarenas, J. L. (2003) From transcription factors to designed sequence-specific DNA-binding peptides. *Chem. Soc. Rev. 32*, 338–349.

[24] Sera, T., and Uranga, C. (2002) Rational design of artificial zinc-finger proteins using a nondegenerate recognition code table. *Biochem. (Mosc.) 41*, 7074–7081.

[25] Gommans, W. M., Haisma, H. J., and Rots, M. G. (2005) Engineering zinc finger protein transcription factors: The therapeutic relevance of switching endogenous gene expression on or off at command. *J. Mol. Biol. 354*, 507–519.

[26] Klug, A. (2010) The Discovery of Zinc Fingers and Their Applications in Gene Regulation and Genome Manipulation. *Ann. Rev. Biochem., Vol 79 79*, 213–231.

[27] Jamieson, A. C., Miller, J. C., and Pabo, C. O. (2003) Drug discovery with engineered zinc-finger proteins. *Nat. Rev. Drug Discovery 2*, 361–368.

[28] Sato, S., Hagihara, M., Sugimoto, K., and Morii, T. (2002) Chemical approaches untangling sequence-specific DNA binding by proteins. *Chem. Eur. J. 8*, 5066–5071.

[29] Canne, L. E., Ferredamare, A. R., Burley, S. K., and Kent, S. B. H. (1995) Total Chemical Synthesis of a Unique Transcription Factor-Related Protein - Cmyc-Max. *J. Am. Chem. Soc. 117*, 2998–3007.

[30] Morii, T., Simomura, M., Morimoto, S., and Saito, I. (1993) Sequence-Specific DNA-Binding by a Geometrically Constrained Peptide Dimer. *J. Am. Chem. Soc. 115*, 1150–1151.

[31] Ponzielli, R., Katz, S., Barsyte-Lovejoy, D., and Penn, L. Z. (2005) Cancer therapeutics: Targeting the dark side of Myc. *Eur. J. Cancer 41*, 2485–2501.

[32] Verzele, D. Development of miniature b(-HLH-)ZIP peptidosteroid conjugates as new transcription factor models. PhD, 2009.

[33] Samuel, S. K., Spencer, V. A., Bajno, L., Sun, J. M., Holth, L. T., Oesterreich, S., and Davie, J. R. (1998) In situ cross-linking by cisplatin of nuclear matrix-bound transcription factors to nuclear DNA of human breast cancer cells. *Cancer Res. 58*, 3004–3008.

[34] Vodovozova, E. L. (2007) Photoaffinity labeling and its application in structural biology. *Biochem. (Mosc.) 72*, 1–20.

[35] Meisenheimer, K. M., and Koch, T. H. (1997) Photocross-linking of nucleic acids to associated proteins. *Crit. Rev. Biochem. Mol. Biol. 32*, 101–140.

[36] Rajski, S. R., and Williams, R. M. (1998) DNA cross-linking agents as antitumor drugs. *Chem. Rev. 98*, 2723–2795.

[37] Kuznetsova, S. A., Clusel, C., Ugarte, E., Elias, I., Vasseur, M., Blumenfeld, M., and Shabarova, Z. A. (1996) Crosslinking of double-stranded oligonucleotides containing O-methyl-substituted pyrophosphate groups to the HNF1 transcription factor in nuclear cell extract. *Nucleic Acids Res. 24*, 4783–4790.

[38] Lesage, D., Metelev, V., Borisova, O., Dolinnaya, N., Oretskaya, T., Baran-Marszak, F., Taillandier, E., Raphael, M., and Fagard, R. (2003) Specific covalent binding of a NF-kappa B decoy hairpin oligonucleotide targeted to the p50 subunit and induction of apoptosis. *FEBS Lett. 547*, 115–118.

[39] Peletskaya, E. N., Boyer, P. L., Kogon, A. A., Clark, P., Kroth, H., Sayer, J. M., Jerina, D. M., and Hughes, S. H. (2001) Cross-linking of the fingers subdomain of human immunodeficiency virus type 1 reverse transcriptase to template-primer. *J. Virol. 75*, 9435–9445.

[40] Gritsenko, O. M., Koudan, E. V., Mikhailov, S. N., Ermolinsky, B. S., Van Aerschot, A., Herdewijn, P., and Gromova, E. S. (2002) Affinity modification of EcoRII DNA methyl-transferase by the dialdehyde-substituted DNA duplexes: Mapping the enzyme region that interacts with DNA. *Nucleos. Nucleot. Nucl. Acids 21*, 753–764.

[41] Guieysse, A. L., Praseuth, D., Giovannangeli, C., Asseline, U., and Helene, C. (2000) Psoralen adducts induced by triplex-forming oligonucleotides are refractory to repair in HeLa cells. *J. Mol. Biol. 296*, 373–383.

[42] Pendergrast, P. S., Chen, Y., Ebright, Y. W., and Ebright, R. H. (1992) Determination of the Orientation of a DNA-Binding Motif in a Protein DNA Complex by Photo-Cross-Linking. *Proc. Natl. Acad. Sci. U. S. A. 89*, 10287–10291.

[43] Suchanek, M., Radzikowska, A., and Thiele, C. (2005) Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. *Nat. Methods 2*, 261–267.

[44] Lee, H. S., Dimla, R. D., and Schultz, P. G. (2009) Protein-DNA photo-crosslinking with a genetically encoded benzophenone-containing amino acid. *Bioorg. Med. Chem. Lett. 19*, 5222–5224.

[45] Cai, X. M., Gray, P. J., and Von Hoff, D. D. (2009) DNA minor groove binders: Back in the groove. *Cancer Treat. Rev. 35*, 437–450.

[46] Dickinson, L. A., Burnett, R., Melander, C., Edelson, B. S., Arora, P. S., Dervan, P. B., and Gottesfeld, J. M. (2004) Arresting cancer proliferation by small-molecule gene regulation. *Chem. Biol. 11*, 1583–1594.

[47] Lorusso, D., Mainenti, S., Pietragalla, A., Ferrandina, G., Foco, G., Masciullo, V., and Scambia, G. (2009) Brostallicin (PNU-166196), a new minor groove DNA binder: preclinical and clinical activity. *Expert Opin. Invest. Drugs 18*, 1939–1946.

[48] Rebar, E. J., Huang, Y., Hickey, R., Nath, A. K., Meoli, D., Nath, S., Chen, B. L., Xu, L., Liang, Y. X., Jamieson, A. C., Zhang, L., Spratt, S. K., Case, C. C., Wolffe, A., and Giordano, F. J. (2002) Induction of angiogenesis in a mouse model using engineered transcription factors. *Nat. Med. 8*, 1427–1432.

[49] Chu, B. C. F., and Orgel, L. E. (1992) Cross-Linking Transcription Factors to Their Recognition Sequences with Pt(Ii) Complexes. *Nucleic Acids Res. 20*, 2497–2502.

[50] Lukhtanov, E. A., Kutyavin, I. V., Gorn, V. V., Reed, M. W., Adams, A. D., Lucas, D. D., and Meyer, R. B. (1997) Sequence and structure dependence of the hybridization-triggered reaction of oligonucleotides bearing conjugated cyclopropapyrroloindole. *J. Am. Chem. Soc. 119*, 6214–6225.

[51] Copeland, K. D., Lueras, A. M. K., Stemp, E. D. A., and Barton, J. K. (2002) DNA Cross-Linking with Metallointercalator-Peptide Conjugates. *Biochem. (Mosc.) 41*, 12785–12797.

[52] Wang, Z., Wang, X., and Rana, T. M. (1996) Protein Orientation in the Tat-TAR Complex Determined by Psoralen Photocross-linking. *J. Biol. Chem. 271*, 16995–16998.

[53] Tao, Z.-F., Fujiwara, T., Saito, I., and Sugiyama, H. (1999) Sequence-Specific DNA Alkylation by Hybrid Molecules between Segment A of Duocarmycin A and Pyrrole/Imidazole Diamide. *Angew. Chem. Int. Ed. 38*, 650–653.

[54] Hurley, L. H. (2002) DNA and its associated processes as targets for cancer therapy. *Nat. Rev. Cancer 2*, 188–200.

[55] Uil, T. G., Haisma, H. J., and Rots, M. G. (2003) Therapeutic modulation of endogenous gene function by agents with designed DNA-sequence specificities. *Nucleic Acids Res. 31*, 6064–6078.

[56] Redell, M. S., and Tweardy, D. J. (2006) Targeting transcription factors in cancer: Challenges and evolving strategies. *Drug Discovery Today: Technologies 3*, 261–267.

[57] Schmitz, K., and Schepers, U. (2004) Polyamides as artificial transcription factors: Novel tools for molecular medicine? *Angew. Chem. Int. Ed. 43*, 2472–2475.

[58] Verzele, D., and Madder, A. (2013) Synthetic Progress in cMyc-Max Oncoprotein Miniaturization: Semi-Online Monitoring Gives Solid-Phase Access to Hydrophobic b(-HLH-)ZIP Peptidosteroid Tweezers. *Eur. J. Org. Chem.* 673–687.

[59] Otvos, J., Laszlo In *Peptide-Based Drug Design*; Otvos, L., Ed.; Methods In Molecular Biology; Humana Press, 2008; Vol. 494; pp 1–8.

[60] Goodwin, D., Simerska, P., and Toth, I. (2012) Peptides As Therapeutics with Enhanced Bioactivity. *Curr. Med. Chem. 19*, 4451–4461.

[61] Nissen, F., Kraft, T. E., Ruppert, T., Eisenhut, M., Haberkorn, U., and Mier, W. (2010) Hot or not - the influence of elevated temperature and microwave irradiation on the solid phase synthesis of an affibody. *Tetrahedron Lett. 51*, 6216–6219.

[62] Xiao, X., Yu, P., Lim, H.-S., Sikder, D., and Kodadek, T. (2007) A Cell-Permeable Synthetic Transcription Factor Mimic. *Angew. Chem. Int. Ed. 46*, 2865–2868.

[63] Pabo, C. O., and Nekludova, L. (2000) Geometric analysis and comparison of protein-DNA interfaces: why is there no simple code for recognition? *J.Mol. Biol. 301*, 597–624.

[64] Luscombe, N. M., Laskowski, R. A., and Thornton, J. M. (2001) Amino acid–base interactions: a three-dimensional analysis of protein–DNA interactions at an atomic level. *Nucleic Acids Res. 29*, 2860–2874.

[65] Kondo, J., and Westhof, E. (2011) Classification of pseudo pairs between nucleotide bases and amino acids by analysis of nucleotide–protein complexes. *Nucleic Acids Res. 39*, 8628–8637.

[66] Shultzaberger, R. K., Maerkl, S. J., Kirsch, J. F., and Eisen, M. B. (2012) Probing the Informational and regulatory plasticity of a transcription factor DNA–binding domain. *PLoS genetics 8*, e1002614.

[67] Talanian, R. V., Mcknight, C. J., Rutkowski, R., and Kim, P. S. (1992) Minimum Length of a Sequence-Specific DNA-Binding Peptide. *Biochem. (Mosc.) 31*, 6871–6875.

[68] Verzele, D., Carrette, L. L. G., and Madder, A. (2010) Peptide scalpels for site-specific dissection of the DNA-protein interface. *Drug Discov. Today Technol. 7*, e115–e123.

[69] Talanian, R. V., Mcknight, C. J., and Kim, P. S. (1990) Sequence-Specific DNA-Binding by a Short Peptide Dimer. *Science 249*, 769–771.

[70] Pellegrini, M., and Ebright, R. H. (1996) Artificial sequence-specific DNA binding peptides: Branched-chain basic regions. *J. Am. Chem. Soc. 118*, 5831–5835.

[71] Okagami, M., Ueno, M., Makino, K., Shimomura, M., Saito, I., Morii, T., and Sugiura, Y. (1995) Sequence-Specific DNA-Binding by Covalently Constrained Peptide Dimers of the Basic Leucine-Zipper Protein Gcn4. *Bioorg. Med. Chem. 3*, 777–784.

[72] Ueno, M., Murakami, A., Makino, K., and Morii, T. (1993) Arranging quaternary structure of peptides by cyclodextrin-guest inclusion complex: sequence-specific DNA binding by a peptide dimer with artificial dimerization module. *J. Am. Chem. Soc. 115*, 12575–12576.

[73] Cuenoud, B., and Schepartz, A. (1993) Design of a Metallo Bzip-Protein That Discriminates between Cre and Ap1 Target Sites - Selection against Ap1. *Proc. Natl. Acad. Sci. U. S. A. 90*, 1154–1159.

[74] Caamano, A. M., Vazquez, M. E., Martinez-Costas, J., Castedo, L., and Mascarenas, J. L. (2000) A light-modulated sequence-specific DNA-binding peptide. *Angew. Chem. Int. Ed. 39*, 3104–3107.

[75] Pensato, S., Renda, M., Leccia, F., Saviano, M., D'Andrea, L. D., Pedone, C., Pedone, P. V., and Romanelli, A. (2010) PNA zipper as a dimerization tool: Development of a bZip mimic. *Biopolymers 93*, 434–441.

[76] Jimenez-Balsa, A., Pazos, E., Martinez-Albardonedo, B., Mascarenas, J. L., and Vazquez, M. E. (2012) Temporary Electrostatic Impairment of DNA Recognition: Light-Driven DNA Binding of Peptide Dimers. *Angew. Chem. Int. Ed. 51*, 8825–8829.

[77] Mosquera, J., Jimenez-Balsa, A., Dodero, V. I., Vazquez, M. E., and Mascarenas, J. L. (2013) Stimuli-responsive selection of target DNA sequences by synthetic bZIP peptides. *Nature Commun. 4*.

[78] Pazos, E., Mosquera, J., Vazquez, M. E., and Mascarenas, J. L. (2011) DNA Recognition by Synthetic Constructs. *Chembiochem 12*, 1958–1973.

[79] Vorherr, T. (2008) Peptdies as drugs Molecular evolvement in recent years. *Pharmachem 7*, 13–16.

[80] Bode, C. A., Muller, C. P., and Madder, A. (2007) Validation of a solid-phase-bound steroid scaffold for the synthesis of novel cyclic peptidosteroids. *J. Pept. Sci. 13*, 702–708.

[81] Virtanen, E., and Kolehmainen, E. (2004) Use of Bile Acids in Pharmacological and Supramolecular Applications. *Eur. J. Org. Chem. 2004*, 3385–3399.

[82] Lai, X.-Z., Feng, Y., Pollard, J., Chin, J. N., Rybak, M. J., Bucki, R., Epand, R. F., Epand, R. M., and Savage, P. B. (2008) Ceragenins: Cholic Acid-Based Mimics of Antimicrobial Peptides. *Acc. Chem. Res. 41*, 1233–1240.

[83] Mrozek, L., Dvorakova, L., Mandelova, Z., Rarova, L., Rezacova, A., Placek, L., Opatrilova, R., Dohnal, J., Paleta, O., Kral, V., Drasar, P., and Jampilek, J. (2011) Investigation of new acyloxy derivatives of cholic acid and their esters as drug absorption modifiers. *Steroids 76*, 1082–1097.

[84] Samstein, R. M., Perica, K., Balderrama, F., Look, M., and Fahmy, T. M. (2008) The use of deoxycholic acid to enhance the oral bioavailability of biodegradable nanoparticles. *Biomaterials 29*, 703–708.

[85] Lee, Y.-k., Kim, S. K., Lee, D. Y., Lee, S., Kim, C.-Y., Shin, H.-C., Moon, H. T., and Byun, Y. (2006) Efficacy of orally active chemical conjugate of low molecular weight heparin and deoxycholic acid in rats, mice and monkeys. *J. Control. Rel. 111*, 290–298.

[86] Kramer, W. Transporters, Trojan horses and therapeutics: suitability of bile acid and peptide transporters for drug delivery. 2011.

[87] Lambert, T. N., Boon, J. M., Smith, B. D., Prez-Payn, M. N., and Davis, A. P. (2002) Facilitated Phospholipid Flip-Flop Using Synthetic Steroid-Derived Translocases. *J. Am. Chem. Soc. 124*, 5276–5277.

[88] Mehiri, M., Chen, W.-H., Janout, V., and Regen, S. L. (2009) Molecular Umbrella Transport: Exceptions to the Classic Size/Lipophilicity Rule. *J. Am. Chem. Soc. 131*, 1338–1339.

[89] Ge, D., Wu, D., Wang, Z., Shi, W., Wu, T., Zhang, A., Hong, S., Wang, J., Zhang, Y., and Ren, L. (2009) Cellular Uptake Mechanism of Molecular Umbrella. *Bioconjug. Chem. 20*, 2311–2316.

[90] Cline, L. L., Janout, V., Fisher, M., Juliano, R. L., and Regen, S. L. (2011) A Molecular Umbrella Approach to the Intracellular Delivery of Small Interfering RNA. *Bioconjug. Chem. 22*, 2210–2216.

[91] Shiraishi, T., and Nielsen, P. E. (2012) Nanomolar Cellular Antisense Activity of Peptide Nucleic Acid (PNA) Cholic Acid (Umbrella?) and Cholesterol Conjugates Delivered by Cationic Lipids. *Bioconjug. Chem. 23*, 196–202.

[92] Janout, V., Jing, B., and Regen, S. L. (2005) Molecular Umbrella-Assisted Transport of an Oligonucleotide across Cholesterol-Rich Phospholipid Bilayers. *J. Am. Chem. Soc. 127*, 15862–15870.

[93] Jing, B., Janout, V., and Regen, S. L. (2003) Fully Detachable Molecular Umbrellas as Peptide Delivery Agents. *Bioconjug. Chem. 14*, 1191–1196.

[94] Davis, A. P. (1993) Cholaphanes et al.; steroids as structural components in molecular engineering. *Chem. Soc. Rev. 22*, 243–253.

[95] Verzele, D., Goeman, J. L., and Madder, A. (2007) LC-(TIC/EIC)-MS as tool in the analysis of diastereomeric 3,12-aza-analogues of deoxycholic acid. *Arkivoc* 325–336.

[96] Verzele, D., and Madder, A. (2007) Short synthesis of orthogonally protected 3 alpha,12 alpha-diamino-5 beta-cholan-24-oic acid, a dipodal steroid scaffold for combinatorial chemistry. *Eur. J. Org. Chem.* 1793–1797.

[97] D'Souz, L. J., and Maitra, U. (1996) Design, Synthesis, and Evaluation of Bile Acid-Based Molecular Tweezers. *J. Org. Chem. 61*, 9494–9502.

[98] Cuenoud, B., and Schepartz, A. (1993) Altered Specificity of DNA-Binding Proteins with Transition-Metal Dimerization Domains. *Science 259*, 510–513.

[99] Indrevoll, B., Bjerke, R., Hvattum, E., Mantzilas, D., and Rogstad, A. (2010) Solid phase synthesis and characterisation of a platelet-derived growth factor receptor (PDGFR) specific Affibody molecule. *J. Pept. Sci. 16*, 51–52.

[100] Garcia-Martin, F., Quintanar-Audelo, M., Garcia-Ramos, Y., Cruz, L. J., Gravel, C., Furic, R., Ct, S., Tulla-Puche, J., and Albericio, F. (2006) ChemMatrix, a Poly(ethylene glycol)-Based Support for the Solid-Phase Synthesis of Complex Peptides. *J. Comb. Chem. 8*, 213–220.

[101] de la Torre, B. G., Jakab, A., and Andreu, D. (2007) Polyethyleneglycol-based resins as solid supports for the synthesis of difficult or long peptides. *Int. J. Pept. Res. Ther. 13*, 265–270.

[102] Carrette, L. L. G., Verzele, D., and Madder, A. (2010) NF-31 color test uncovers 'hidden' alcohol functionalities in PEG-based resins for solid phase peptide synthesis. *Tetrahedron Lett. 51*, 2106–2108.

[103] Bayer, E. (1991) Towards the Chemical Synthesis of Proteins. *Angew. Chem. Int. Ed. 30*, 113–129.

[104] Coantic, S., Subra, G., and Martinez, J. (2008) Microwave-assisted Solid Phase Peptide Synthesis on High Loaded Resins. *Int. J. Pept. Res. Ther. 14*, 143–147.

[105] Holmes, C. P., and Jones, D. G. (1995) Reagents for Combinatorial Organic Synthesis: Development of a New o-Nitrobenzyl Photolabile Linker for Solid Phase Synthesis. *J. Org. Chem. 60*, 2318–2319.

[106] Holmes, C. P. (1997) Model studies for new o-nitrobenzyl photolabile linkers: Substituent effects on the rates of photochemical cleavage. *J. Org. Chem. 62*, 2370–2380.

[107] Grillo-Bosch, D., Rabanal, F., and Giralt, E. (2011) Improved Fmoc-based solid-phase synthesis of homologous peptide fragments of human and mouse prion proteins. *J. Pept. Sci. 17*, 32–38.

[108] Bacsa, B., Horva?ti, K., Bosze, S., Andreae, F., and Kappe, C. O. (2008) Solid-Phase Synthesis of Difficult Peptide Sequences at Elevated Temperatures: A Critical Comparison of Microwave and Conventional Heating Technologies. *J. Org. Chem. 73*, 7532–7542.

[109] Pedersen, S. L., Bhatia, V. K., Jurt, S., Paulsson, J. F., Pedersen, M. H., Jorgensen, R., Holst, B., Stamou, D., Vrang, N., Zerbe, O., and Jensen, K. J. (2012) Improving membrane

binding as a design strategy for amphipathic peptide hormones: 2-helix variants of PYY3-36. *J. Pept. Sci. 18*, 579–587.

[110] King, D. S., Fields, C. G., and Fields, G. B. (1990) A cleavage method which minimizes side reactions following Fmoc solid phase peptide synthesis. *Int. J. Pept. Protein Res. 36*, 255–266.

[111] Guillier, F., Orain, D., and Bradley, M. (2000) Linkers and cleavage strategies in solid-phase organic synthesis and combinatorial chemistry (vol 100, pg 2091, 2000). *Chem. Rev. 100*, 3859–3859.

[112] Yu, H. M., Chen, S. T., and Wang, K. T. (1992) Enhanced coupling efficiency in solid-phase peptide synthesis by microwave irradiation. *J. Org. Chem. 57*, 4781–4784.

[113] Collins, J. M., and Leadbeater, N. E. (2007) Microwave energy: a versatile tool for the biosciences. *Org. Biomol. Chem. 5*, 1141–1150.

[114] Kappe, C. O., Pieber, B., and Dallinger, D. (2013) Microwave Effects in Organic Synthesis: Myth or Reality? *Angew. Chem. Int. Ed. 52*, 1088–1094.

[115] Herrero, M. A., Kremsner, J. M., and Kappe, C. O. (2007) Nonthermal Microwave Effects Revisited: On the Importance of Internal Temperature Monitoring and Agitation in Microwave Chemistry. *J. Org. Chem. 73*, 36–47.

[116] Malik, L., Tofteng, A. P., Pedersen, S. L., Sorensen, K. K., and Jensen, K. J. (2010) Automated 'X-Y' robot for peptide synthesis with microwave heating: application to difficult peptide sequences and protein domains. *J. Pept. Sci. 16*, 506–512.

[117] Pedersen, S. L., Srensen, K. K., and Jensen, K. J. (2010) Semi-automated microwave-assisted SPPS: Optimization of protocols and synthesis of difficult sequences. *Pept. Sci. 94*, 206–212.

[118] Keller, W., Knig, P., and Richmond, T. J. (1995) Crystal Structure of a bZIP/DNA Complex at 2.2 : Determinants of DNA Specific Recognition. *J. Mol. Biol. 254*, 657 – 667.

[119] Verzele, D., and Madder, A. (2013) Patchwork Protein Chemistry: A Practitioner's Treatise on the Advances in Synthetic Peptide Stitchery. *Chembiochem 14*, 1032–1048.

[120] Bray, B. L. (2003) Large-scale manufacture of peptide therapeutics by chemical synthesis. *Nat. Rev. Drug Discovery 2*, 587–593.

[121] Morii, T., Morimoto, S., and Saito, I. (1991) DNA recognition by synthetic peptides with dyad symmetry. *J. Inorg. Biochem. 43*, 468–468.

[122] Hirata, A., Ueno, M., Aizawa, Y., Ohkubo, K., Morii, T., and Yoshikawa, S. (2005) Dual DNA recognition codes of a short peptide derived from the basic leucine zipper protein EmBP1. *Bioorg. Med. Chem. 13*, 3107–3116.

[123] Ueno, M., Sawada, M., Makino, K., and Morii, T. (1994) Recognition of nonpalindromic DNA sequence by a peptide heterodimer with artificial dimerization module. *J. Am. Chem. Soc. 116*, 11137–11138.

[124] Aizawa, Y., Sugiura, Y., Ueno, M., Mori, Y., Imoto, K., Makino, K., and Morii, T. (1999) Stability of the dimerization domain effects the cooperative DNA binding of short peptides. *Biochem. (Mosc.) 38*, 4008–4017.

[125] Morii, T., Yamane, J., Aizawa, Y., Makino, K., and Sugiura, Y. (1996) Cooperative

oligomerization enhances sequence-selective DNA binding by a short peptide. *J. Am. Chem. Soc. 118*, 10011–10017.

[126] Bakhiya, N., and Appel, K. (2010) Toxicity and carcinogenicity of furan in human diet. *Arch. Toxicol. 84*, 563–578.

[127] Peterson, L. A. (2006) Electrophilic Intermediates Produced by Bioactivation of Furan*. *Drug Metab. Rev. 38*, 615–626.

[128] Guainazzi, A., and Scharer, O. (2010) Using synthetic DNA interstrand crosslinks to elucidate repair pathways and identify new therapeutic targets for cancer chemotherapy. *Cell. Mol. Life Sci. 67*, 3683–3697.

[129] Grillari, J., Katinger, H., and Voglauer, R. (2007) Contributions of DNA interstrand cross-links to aging of cells and organisms. *Nucleic Acids Res. 35*, 7566–7576.

[130] Noll, D. M., Mason, T. M., and Miller, P. S. (2005) Formation and Repair of Interstrand Cross-Links in DNA. *Chem. Rev. 106*, 277–301.

[131] Wilds, C. J., Booth, J. D., and Noronha, A. M. (2011) Synthesis of building blocks and oligonucleotides with GO(6)-alkyl-O(6)G cross-links. *Curr. Protoc. Nucleic Acid Chem. 44*, 5.9.1–5.9.19.

[132] Cohen, S. B., and Cech, T. R. (2001) Engineering disulfide cross-links in RNA using thiol-disulfide interchange chemistry. *Curr. Protoc. Nucleic Acid Chem.* 5.1.1–5.1.10.

[133] Singh, I., Vyle, J. S., and Heaney, F. (2009) Fast, copper-free click chemistry: a convenient solid-phase approach to oligonucleotide conjugation. *Chem. Commun.* 3276–3278.

[134] Halila, S., Velasco, T., De Clercq, P., and Madder, A. (2005) Fine-tuning furan toxicity: fast and quantitative DNA interchain crosslink formation upon selective oxidation of a furan containing oligonucleotide. *Chem. Commun.* 936–938.

[135] Stevens, K., and Madder, A. (2009) Furan-modified oligonucleotides for fast, high-yielding and site-selective DNA inter-strand cross-linking with non-modified complements. *Nucleic Acids Res. 37*, 1555–1565.

[136] Stevens, K., Claeys, D. D., Catak, S., Figaroli, S., Hocek, M., Tromp, J. M., Schuerch, S., Van Speybroeck, V., and Madder, A. (2011) Furan-Oxidation-Triggered Inducible DNA Cross-Linking: Acyclic Versus Cyclic Furan-Containing Building Blocks-On the Benefit of Restoring the Cyclic Sugar Backbone. *Chem. Eur. J. 17*, 6940–6953.

[137] Jawalekar, A. M., de Beeck, M. O., van Delft, F. L., and Madder, A. (2011) Synthesis and incorporation of a furan-modified adenosine building block for DNA interstrand crosslinking. *Chem. Commun. 47*, 2796–2798.

[138] Op de Beeck, M., and Madder, A. (2011) Unprecedented C-Selective Interstrand Cross-Linking through in Situ Oxidation of Furan-Modified Oligodeoxynucleotides. *J. Am. Chem. Soc. 133*, 796–807.

[139] Op de Beeck, M., and Madder, A. (2012) Sequence Specific DNA Cross-Linking Triggered by Visible Light. *J. Am. Chem. Soc. 134*, 10737–10740.

[140] Barciszewski, J., Siboska, G. E., Pedersen, B. O., Clark, B. F., and Rattan, S. I. (1996) Evidence for the presence of kinetin in DNA and cell extracts. *FEBS Lett. 393*, 197–200.

[141] Barciszewski, J., Siboska, G. E., Pedersen, B. O., Clark, B. F., and Rattan, S. I. (1997)

Furfural, a precursor of the cytokinin hormone kinetin, and base propenals are formed by hydroxyl radical damage of DNA. *Biochem. Biophys. Res. Commun. 238*, 317–319.

[142] Barciszewski, J., Mielcarek, M., Stobiecki, M., Siboska, G., and Clark, B. F. (2000) Identification of 6-furfuryladenine (kinetin) in human urine. *Biochem. Biophys. Res. Commun. 279*, 69–73.

[143] Wyszko, E., Barciszewska, M. Z., Markiewicz, M., Szymański, M., Markiewicz, W. T., Clark, B. F., and Barciszewski, J. (2003) Action-at-a distance of a new DNA oxidative damage product 6-furfuryl-adenine (kinetin) on template properties of modified DNA. *Biochim. Biophys. Acta 1625*, 239–245.

[144] Barciszewski, J., Siboska, G. E., Pedersen, B. O., Clark, B. F., and Rattan, S. I. (1997) A mechanism for the in vivo formation of N6-furfuryladenine, kinetin, as a secondary oxidative damage product of DNA. *FEBS Lett. 414*, 457–460.

[145] Barciszewski, J., Massino, F., and Clark, B. F. (2007) Kinetina multiactive molecule. *Int. J. Biol. Macromol. 40*, 182–192.

[146] Gan, S., and Amasino, R. M. (1995) Inhibition of leaf senescence by autoregulated production of cytokinin. *Science 270*, 1986–1988.

[147] Sharma, S. P., Kaur, J., and Rattan, S. I. (1997) Increased longevity of kinetin-fed Zaprionus fruitflies is accompanied by their reduced fecundity and enhanced catalase activity. *IUBMB Life 41*, 869–875.

[148] Rattan, S. I., and Clark, B. F. (1994) Kinetin delays the onset of aging characteristics in human fibroblasts. *Biochem. Biophys. Res. Commun. 201*, 665–672.

[149] Frebort, I., Kowalska, M., Hluska, T., Frébortová, J., and Galuszka, P. (2011) Evolution of cytokinin biosynthesis and degradation. *J. Exp. Bot. 62*, 2431–2452.

[150] Wu, J., Weinstein, G., Kricorian, G., Kormeili, T., and McCullough, J. (2007) Topical kinetin 0.1% lotion for improving the signs and symptoms of rosacea. *Clin. Exp. Dermatol. 32*, 693–695.

[151] Liu, Y., Zhang, Z., and Yang, X. (2011) Kinetin protects against lipid peroxidation and improves antioxidant status in cultured astrocytes and mouse brain exposed to D-galactose. *Afric. J. Biotech. 10*, 11721–11727.

[152] Cabello, C. M., Bair III, W. B., Ley, S., Lamore, S. D., Azimian, S., and Wondrak, G. T. (2009) The experimental chemotherapeutic N6-furfuryladenosine (kinetin-riboside) induces rapid ATP depletion, genotoxic stress, and CDKN1A (p21) upregulation in human cancer cell lines. *Biochem. Pharmacol. 77*, 1125–1138.

[153] Mehrzad, J., and Rajabi, M. (2011) Kinetin (N 6-furfuryladenine): Cytotoxicity against MCF-7 breast cancer cell line and interaction with bovine serum albumin. *Afric. J. Biotech. 10*, 6304–6309.

[154] Verbeke, P., Siboska, G. E., Clark, B. F., and Rattan, S. I. (2000) Kinetin Inhibits Protein Oxidation and Glycoxidation in Vitro. *Biochem. Biophys. Res. Commun. 276*, 1265–1270.

[155] Olsen, A., Siboska, G. E., Clark, B. F., and Rattan, S. I. (1999) N¡ sup¿ 6¡/sup¿-Furfuryladenine, Kinetin, Protects against Fenton Reaction-Mediated Oxidative Damage to DNA. *Biochem. Biophys. Res. Commun. 265*, 499–502.

[156] Kurchii, B. A. Possible free radical mechanisms of action of auxin and kinetin. 2000.

[157] Carrette, L. L., Gyssels, E., and Madder, A. (2013) DNA Interstrand Cross-Link Formation Using Furan as a Masked Reactive Aldehyde. *Curr. Protoc. Nucleic Acid Chem.* 5–12.

[158] Takayama, K., Noguchi, T., Nakano, M., and Migita, T. (1977) Reactivities of diphenylfuran (a singlet oxygen trap) with singlet oxygen and hydroxyl radical in aqueous systems. *Biochem. Biophys. Res. Commun. 75*, 1052–1058.

[159] Okada, Y., Kaneko, M., and Okajima, H. (1996) Hydroxyl radical scavenging activity of naturally occurring furan fatty acids. *Biol. Pharm. Bull. 19*, 1607–1613.

[160] Markiewicz, M., Adrych-Rozek, K., and Markiewicz, W. T. Synthesis of Kinetin 2'-Deoxynucleoside and Its Introduction to Oligonucleotides. 2002.

[161] Godzina, P., Adrych-Rozek, K., and Markiewicz, W. T. (1999) Synthetic Oligonucleotide Combinatorial Libraries. 3. Synthesis of Polyamevonucleosides. *Nucleos. Nucleot. 18*, 2397–2414.

[162] Ferentz, A. E., and Verdine, G. L. (1991) Disulfide-crosslinked oligonucleotides. *J. Am. Chem. Soc. 113*, 4000–4002.

[163] Larson, C. J., and Verdine, G. L. (1992) A high-capacity column for affinity purification of sequence-specific DNA-binding proteins. *Nucleic Acids Res. 20*, 3525.

[164] Soriano-Garcia, M., and Parthasarathy, R. (1977) Stereochemistry and hydrogen bonding of cytokinins: 6-furfurylaminopurine (kinetin). *Acta Crystallogr. Sect. B-Struct. Sci. 33*, 2674–2677.

[165] Lee, I. K., Ahn, J. D., Kim, H. S., Park, J. Y., and Lee, K. U. (2003) Advantages of the Circular Dumbbell Decoy in Gene Therapy and Studies of Gene Regulation. *Curr. Drug Targets 4*, 619–623.

[166] Murakami, A., Yamamoto, Y., Namba, M., Iwase, R., and Yamaoka, T. (2001) Photo-Cross-Linked Oligonucleotide Duplex as a Decoy-DNA for Inhibition of Restriction Endonuclease Activity. *Bioorg. Chem. 29*, 223 – 233.

[167] Lee, J. T. (2012) Epigenetic Regulation by Long Noncoding RNAs. *Science 338*, 1435–1439.

[168] Turner, A.-M. W., and Morris, K. V. (2010) Controlling transcription with noncoding RNAs in mammalian cells. *Biotechniques 48*, ixxvi.

[169] Deceuninck, A. Evaluation of a furan moiety as a masked reactive enal functionality for peptide-DNA crosslinking and peptide labeling purposes. Ph.D. thesis, Ghent University, 2010.

[170] Levadala, M. K., Banerjee, S. R., Maresca, K. P., Babich, J. W., and Zubieta, J. (2004) Direct Reductive Alkylation of Amino Acids: Synthesis of Bifunctional Chelates for Nuclear Imaging. *Synthesis 2004*, 1759–1766.

[171] Abdel-Magid, A. F., and Mehrman, S. J. (2006) A Review on the Use of Sodium Triacetoxyborohydride in the Reductive Amination of Ketones and Aldehydes. *Org. Process Res. Dev. 10*, 971–1031.

[172] Hoogewijs, K., Deceuninck, A., and Madder, A. (2012) Aromatic capping surprisingly stabilizes furan moieties in peptides against acidic degradation. *Org. Biomol. Chem. 10*, 3999–4002.

[173] Hurst, H. C. (1995) Transcription factors .1. bZIP proteins. *Protein Profile 2*, 105–168.

[174] Schulz, A., Busmann, A., Kluver, E., Schnebel, M., and Adermann, K. (2004) Stability and Cleavage Conditions of (2-Furyl)-L-Alanine-Containing Peptides. *Prot. Pept. Lett. 11*, 601–606.

[175] Stellwagen, E., and Van Rooyan, S. (1967) The Structural Environment of the Tryptophanyl Residue of Horse Heart Ferricytochrome c. *J. Biol. Chem. 242*, 4801–4805.

[176] Williams, M. (1975) Effect of N-bromosuccinimide modification on dihydrofolate reductase from a methotrexate-resistant strain of Escherichia coli. Activity, spectrophotometric, fluorescence and circular dichroism studies. *J. Biol. Chem. 250*, 322–330.

[177] Murray, V. (1999) A survey of the sequence-specific interaction of damaging agents with DNA: emphasis on antitumor agents. *Progr. nucleic acid res. mol. biol. 63*, 367–415.

[178] Mathews, D. H., and Turner, D. H. (2002) Use of chemical modification to elucidate RNA folding pathways. *Curr. Protoc. Nucleic Acid Chem.* 11.9.

[179] Tate, J. J., Persinger, J., and Bartholomew, B. (1998) Survey of four different photoreactive moieties for DNA photoaffinity labeling of yeast RNA polymerase III transcription complexes. *Nucleic Acids Res. 26*, 1421–1426.

[180] Winnacker, M., Breeger, S., Strasser, R., and Carell, T. (2009) Novel Diazirine-Containing DNA Photoaffinity Probes for the Investigation of DNA-Protein-Interactions. *Chembiochem 10*, 109–118.

[181] Shigdel, U. K., Zhang, J., and He, C. (2008) Diazirine-Based DNA Photo-Cross-Linking Probes for the Study of Protein–DNA Interactions. *Angew. Chem. 120*, 96–99.

[182] Zhu, G., and Lippard, S. J. (2009) Photoaffinity Labeling Reveals Nuclear Proteins That Uniquely Recognize Cisplatin- DNA Interstrand Cross-Links. *Biochem. (Mosc.) 48*, 4916–4925.

[183] Dolinnaya, N. G., Zubin, E. M., Kubareva, E. A., Zatsepin, T. S., and Oretskaya, T. S. (2009) Design and Synthesis of 2-Functionalised Oligonucleotides. Their Application for Covalent Trapping the Protein-DNA Complexes. *Curr. Org. Chem. 13*, 1029–1049.

[184] Joenje, H. (2011) Metabolism: Alcohol, DNA and disease. *Nature 475*, 45–46.

[185] O'Brien, P. J., Siraki, A. G., and Shangari, N. (2005) Aldehyde sources, metabolism, molecular toxicity mechanisms, and possible effects on human health. *Crit. Rev. Toxicol. 35*, 609–662.

[186] Ide, H., Shoulkamy, M. I., Nakano, T., Miyamoto-Matsubara, M., and Salem, A. M. H. (2011) Repair and biochemical effects of DNA-protein crosslinks. *Mutat. Res. 711*, 113–122.

[187] Barker, S., Weinfeld, M., and Murray, D. (2005) DNA–protein crosslinks: their induction, repair, and biological consequences. *Mutat. Res. 589*, 111–135.

[188] Szekely, J., Rizzo, C. J., and Marnett, L. J. (2008) Chemical properties of oxopropenyl adducts of purine and pyrimidine nucleosides and their reactivity toward amino acid cross-link formation. *J. Am. Chem. Soc. 130*, 2195–2201.

[189] Voitkun, V., and Zhitkovich, A. (1999) Analysis of DNA–protein crosslinking activity of malondialdehyde in vitro. *Mutat. Res. 424*, 97–106.

[190] Nelson, J. D., Denisenko, O., and Bomsztyk, K. (2006) Protocol for the fast chromatin immunoprecipitation (ChIP) method. *Nat. Protocols 1*, 179–185.

[191] Wilde, J. A., Bolton, P. H., Mazumder, A., Manoharan, M., and Gerlt, J. A. (1989) Characterization of the equilibrating forms of the aldehydic abasic site in duplex DNA by oxygen-17 NMR. *J. Am. Chem. Soc. 111*, 1894–1896.

[192] Sczepanski, J. T., Wong, R. S., McKnight, J. N., Bowman, G. D., and Greenberg, M. M. (2010) Rapid DNA–protein cross-linking and strand scission by an abasic site in a nucleosome core particle. *Proc. Nat. Acad. Sci. U.S.A. 107*, 22475–22480.

[193] Guan, L., and Greenberg, M. M. (2010) Irreversible inhibition of DNA polymerase β by an oxidized abasic lesion. *J. Am. Chem. Soc. 132*, 5004–5005.

[194] Brevnov, M. G., Gritsenko, O. M., Mikhailov, S. N., Efimtseva, E. V., Ermolinsky, B. S., Van Aerschot, A., Herdewijn, P., Repyk, A. V., and Gromova, E. S. (1997) DNA duplexes with reactive dialdehyde groups as novel reagents for cross-linking to restriction-modification enzymes. *Nucleic Acids Res. 25*, 3302–3309.

[195] Romanenkov, A. S., Ustyugov, A. A., Zatsepin, T. S., Nikulova, A. A., Kolesnikov, I. V., Metelev, V. G., Oretskaya, T. S., and Kubareva, E. A. (2005) Analysis of DNA-protein interactions in complexes of transcription factor NF-kappa B with DNA. *Biochem. (Mosc.) 70*, 1212–1222.

[196] Tunitskaya, V. L., Rusakova, E. E., Memelova, L. V., Kochetkov, S. N., Van Aerschot, A., Herdewijn, P., Efimtseva, E. V., Ermolinsky, B. S., and Mikhailov, S. N. (1999) Mapping of T7 RNA polymerase active site with novel reagents - oligonucleotides with reactive dialdehyde groups. *FEBS Lett. 442*, 20–24.

[197] Sugiyama, T., Kittaka, A., Takayama, H., Tomioka, M., Ida, Y., and Kuroda, R. (2003) Aggregation of RecA-derived peptides on single-stranded oligonucleotides triggered by schiff base-mediated crosslinking. *Bioorg. Med. Chem. Lett. 13*, 2847–2851.

[198] Klungland, A., Paulsen, R., Rolseth, V., Yamada, Y., Ueno, Y., Wiik, P., Matsuda, A., Seeberg, E., and Bjelland, S. (2001) 5-Formyluracil and its nucleoside derivatives confer toxicity and mutagenicity to mammalian cells by interfering with normal RNA and DNA metabolism. *Toxicol. Lett. 119*, 71–78.

[199] Kellinger, M. W., Song, C.-X., Chong, J., Lu, X.-Y., He, C., and Wang, D. (2012) 5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription. *Nature Struct. Molec. Biol. 19*, 831–833.

[200] Mee, L. K., and Adelstein, S. J. (1981) Predominance of core histones in formation of DNA–protein crosslinks in gamma-irradiated chromatin. *Proc. Nat. Acad. Sci. U.S.A. 78*, 2194–2198.

[201] Kasai, H., Iida, A., Yamaizumi, Z., Nishimura, S., and Tanooka, H. (1990) 5-Formyldeoxyuridine: a new type of DNA damage induced by ionizing radiation and its mutagenicity to Salmonella strain TA102. *Mutat. Res. Lett. 243*, 249–253.

[202] Okamoto, A., Tainaka, K., and Saito, I. (2002) A facile incorporation of the aldehyde function into DNA: 3-formylindole nucleoside as an aldehyde-containing universal nucleoside. *Tetrahedron Lett. 43*, 4581–4583.

[203] Raindlova, V., Pohl, R., and Hocek, M. (2012) Synthesis of Aldehyde-Linked Nucleotides

and DNA and Their Bioconjugations with Lysine and Peptides through Reductive Amination. *Chem. Eur. J. 18*, 4080–4087.

[204] Schmidt, M. J., and Summerer, D. (2013) Red-Light-Controlled Protein–RNA Crosslinking with a Genetically Encoded Furan. *Angew. Chem. Int. Ed. 52*, 46904693.

[205] Spande, T. t., and Witkop, B. (1967) [58] Determination of the tryptophan content of proteins with N-bromosuccinimide. *Methods Enzymol. 11*, 498–506.

[206] Pace, N. C., and Scholtz, M. J. (1998) A helix propensity scale based on experimental studies of peptides and proteins. *Biophys. J. 75*, 422–427.

[207] Suzuki, M. (1993) Common features in DNA recognition helices of eukaryotic transcription factors. *The EMBO journal 12*, 3221–3226.

[208] Luscombe, N. M., and Thornton, J. M. (2002) Protein–DNA interactions: amino acid conservation and the effects of mutations on binding specificity. *J. Mol. Biol. 320*, 991–1009.

[209] Greco, N. J., and Tor, Y. (2005) Simple fluorescent pyrimidine analogues detect the presence of DNA abasic sites. *J. Am. Chem. Soc. 127*, 10784–10785.

[210] Wicke, L., and Engels, J. W. (2012) Postsynthetic on Column RNA Labeling via Stille Coupling. *Bioconjug. Chem. 23*, 627–642.

[211] Srivatsan, S. G., and Tor, Y. (2007) Fluorescent pyrimidine ribonucleotide: synthesis, enzymatic incorporation, and utilization. *J. Am. Chem. Soc. 129*, 2044–2053.

[212] Zatsepin, T. S., Stetsenko, D. A., Gait, M. J., and Oretskaya, T. S. (2005) Use of Carbonyl Group Addition-Elimination Reactions for Synthesis of Nucleic Acid Conjugates. *Bioconjug. Chem. 16*, 471–489.

[213] Blatter, E. E., Ebright, Y. W., and Ebright, R. H. (1992) Identification of an amino acid-base contact in the GCN4-DNA complex by bromouracil-mediated photocrosslinking. *Nature 359*, 650–652.

[214] Steen, H., and Jensen, O. N. (2002) Analysis of protein–nucleic acid interactions by photochemical cross-linking and mass spectrometry. *Mass Spectrom. Rev. 21*, 163–182.

[215] Doktycz, M. J. (1997) Nucleic Acids: Thermal Stability and Denaturation. *eLS* 3123?3140.

[216] Misra, V. K., Hecht, J. L., Sharp, K. A., Friedman, R. A., and Honig, B. (1994) Salt effects on protein-DNA interactions: the λcI repressor and EcoRI endonuclease. *J. Mol. Biol. 238*, 264–280.

[217] Battiste, J. L., Mao, H., Rao, N. S., Tan, R., Muhandiram, D., Kay, L. E., Frankel, A. D., and Williamson, J. R. (1996) α helix-RNA major groove recognition in an HIV-1 Rev peptide-RRE RNA complex. *Science 273*, 1547–1551.

[218] Ping, Y.-H., Liu, Y., Wang, X., Neenhold, H. R., and Rana, T. M. (1997) Dynamics of RNA-protein interactions in the HIV-1 Rev-RRE complex visualized by 6-thioguanosine-mediated photocrosslinking. *RNA 3*, 850.

[219] Nakano, S., Fukuda, M., Tamura, T., Sakaguchi, R., Nakata, E., and Morii, T. (2013) Simultaneous Detection of ATP and GTP by Covalently Linked Fluorescent Ribonucleopeptide Sensors. *J. Am. Chem. Soc. 135*, 3465–3473.

[220] Hair, P., Cameron, F., and McKeage, K. (2013) Mipomersen Sodium: First Global Approval. *Drugs* 1–7.

[221] CROOKE, S. T. (1998) An overview of progress in antisense therapeutics. *Antisense Nucleic Acid Drug Dev. 8*, 115–122.

[222] Bennett, C. F., and Swayze, E. E. (2010) RNA targeting therapeutics: molecular mechanisms of antisense oligonucleotides as a therapeutic platform. *Annu. Rev. Pharmacol. Toxicol. 50*, 259–293.

[223] Imoto, S., Hori, T., Hagihara, S., Taniguchi, Y., Sasaki, S., and Nagatsugi, F. (2010) Alteration of cross-linking selectivity with the 2 -OMe analogue of 2-amino-6-vinylpurine and evaluation of antisense effects. *Bioorg. Med. Chem. Lett. 20*, 6121–6124.

[224] Harris, M. E., and Christian, E. L. (2009) RNA crosslinking methods. *Methods Enzymol. 468*, 127–146.

[225] Simon, M. D., Wang, C. I., Kharchenko, P. V., West, J. A., Chapman, B. A., Alekseyenko, A. A., Borowsky, M. L., Kuroda, M. I., and Kingston, R. E. (2011) The genomic binding sites of a noncoding RNA. *Proc. Nat. Acad. Sci. U.S.A. 108*, 20497–20502.

[226] Ng, E. W., Shima, D. T., Calias, P., Cunningham, E. T., Guyer, D. R., and Adamis, A. P. (2006) Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nat. Rev. Drug Discovery 5*, 123–132.

[227] Kocalka, P., El-Sagheer, A. H., and Brown, T. (2008) Rapid and Efficient DNA Strand Cross-Linking by Click Chemistry. *Chembiochem 9*, 1280–1285.

[228] Pujari, S. S., and Seela, F. (2012) Cross-linked DNA: propargylated ribonucleosides as click ligation sites for bifunctional azides. *J. Org. Chem. 77*, 4460–4465.

[229] Xiong, H., and Seela, F. (2012) Cross-linked DNA: Site-selective click ligation in duplexes with bis-azides and stability changes caused by internal cross-links. *Bioconjug. Chem. 23*, 1230–1243.

[230] Cohen, S. B., and Cech, T. R. (1997) Dynamics of thermal motions within a large catalytic RNA investigated by cross-linking with thiol-disulfide interchange. *J. Am. Chem. Soc. 119*, 6259–6268.

[231] Iverson, B. L., and Dervan, P. B. (1987) Nonenzymatic sequence-specific cleavage of single-stranded DNA to nucleotide resolution. DNA methyl thioether probes. *J. Am. Chem. Soc. 109*, 1241–1243.

[232] Webb, T. R., and Matteucci, M. D. (1986) Sequence-specific cross-linking of deoxyoligonucleotides via hybridization-triggered alkylation. *J. Am. Chem. Soc. 108*, 2764–2765.

[233] Lukhtanov, E. A., Podyminogin, M. A., Kutyavin, I. V., Meyer, R. B., and Gamper, H. B. (1996) Rapid and efficient hybridization-triggered crosslinking within a DNA duplex by an oligodeoxyribonucleotide bearing a conjugated cyclopropapyrroloindole. *Nucleic Acids Res. 24*, 683–687.

[234] Sasaki, S. (2001) Active oligonucleotides incorporating alkylating an agent as potential sequence-and base selective modifier of gene expression. *Eur. J. Pharm. Sci. 13*, 43–51.

[235] Kean, J. M., Murakami, A., Blake, K. R., Cushman, C. D., and Miller, P. S. (1988) Photochemical cross-linking of psoralen-derivatized oligonucleoside methylphosphonates to rabbit globin messenger RNA. *Biochem. (Mosc.) 27*, 9113–9121.

[236] Ali, M., Oishi, M., Nagatsugi, F., Mori, K., Nagasaki, Y., Kataoka, K., and Sasaki, S. (2006) Intracellular inducible alkylation system that exhibits antisense effects with greater

potency and selectivity than the natural oligonucleotide. *Angew. Chem. Int. Ed. 45*, 3136–3140.

[237] Kumar, R., Wiebe, L. I., and Knaus, E. E. (1994) A mild and efficient methodology for the synthesis of 5-halogeno uracil nucleosides that occurs via a 5-halogeno-6-azido-5, 6-dihydro intermediate. *Can. J. Chem. 72*, 2005–2010.

[238] Kvaerno, L., and Wengel, J. (2001) Antisense molecules and furanose conformationsis it really that simple? *Chem. Commun.* 1419–1424.

[239] Shaw, N. N., and Arya, D. P. (2008) Recognition of the unique structure of DNA: RNA hybrids. *Biochimie 90*, 1026–1039.

[240] Wang, S., and Kool, E. T. (1995) Origins of the large differences in stability of DNA and RNA helixes: C-5 methyl and 2'-hydroxyl effects. *Biochem. (Mosc.) 34*, 4125–4132.

[241] McIntosh, L. P., and Dahlquist, F. W. (1990) Biosynthetic Incorporation of 15N and 13C for Assignment and Interpretation of Nuclear Magnetic Resonance Spectra of Proteins. *Q. Rev. Biophys. 23*, 1–38.

[242] Hendrickson, W. A., Horton, J. R., and LeMaster, D. M. (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J. 9*, 1665–1672.

[243] Wang, L., Xie, J., and Schultz, P. G. (2006) Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct. 35*, 225–249.

[244] Atkins, J. F., and Gesteland, R. (2002) The 22nd Amino Acid. *Science 296*, 1409–1410.

[245] Srinivasan, G., James, C. M., and Krzycki, J. A. (2002) Pyrrolysine Encoded by UAG in Archaea: Charging of a UAG-Decoding Specialized tRNA. *Science 296*, 1459–1462.

[246] Hao, B., Gong, W., Ferguson, T. K., James, C. M., Krzycki, J. A., and Chan, M. K. (2002) A New UAG-Encoded Residue in the Structure of a Methanogen Methyltransferase. *Science 296*, 1462–1466.

[247] Fekner, T., Li, X., and Chan, M. K. (2010) Pyrrolysine analogs for translational incorporation into proteins. *Eur. J. Org. Chem. 2010*, 4171–4179.

[248] Li, W.-T., Mahapatra, A., Longstaff, D. G., Bechtel, J., Zhao, G., Kang, P. T., Chan, M. K., and Krzycki, J. A. (2009) Specificity of pyrrolysyl-tRNA synthetase for pyrrolysine and pyrrolysine analogs. *J. Mol. Biol. 385*, 1156–1164.

[249] Yanagisawa, T., Ishii, R., Fukunaga, R., Kobayashi, T., Sakamoto, K., and Yokoyama, S. (2008) Multistep Engineering of Pyrrolysyl-tRNA Synthetase to Genetically Encode N-(o-Azidobenzyloxycarbonyl) lysine for Site-Specific Protein Modification. *Chem. Biol. 15*, 1187 – 1197.

[250] Kaya, E., Vrabel, M., Deiml, C., Prill, S., Fluxa, V. S., and Carell, T. (2012) A Genetically Encoded Norbornene Amino Acid for the Mild and Selective Modification of Proteins in a Copper-Free Click Reaction. *Angew. Chem. Int. Ed. 51*, 4466–4469.

[251] Palomo, J. M. (2010) Diels–Alder Cycloaddition in Protein Chemistry. *Eur. J. Org. Chem. 2010*, 6303–6314.

[252] Hoogewijs, K., Buyst, D., Winne, J., Martins, J. C., and Madder, A. (2013) Exploiting furans versatile reactivity in reversible and irreversible orthogonal peptide labeling. *Chem. Commun. 49*, 2927–2929.

[253] Seeman, N. C. (1982) Nucleic acid junctions and lattices. *J. Theor. Biol. 99*, 237–247.

[254] Wickham, S. F., Bath, J., Katsuda, Y., Endo, M., Hidaka, K., Sugiyama, H., and Turberfield, A. J. (2012) A DNA-based molecular motor that can navigate a network of tracks. *Nature nanotech. 7*, 169–173.

[255] Lund, K., Manzo, A. J., Dabby, N., Michelotti, N., Johnson-Buck, A., Nangreave, J., Taylor, S., Pei, R., Stojanovic, M. N., Walter, N. G., Winfree, E., and Yan, H. (2010) Molecular robots guided by prescriptive landscapes. *Nature 465*, 206–210.

[256] Nakata, E., Liew, F. F., Uwatoko, C., Kiyonaka, S., Mori, Y., Katsuda, Y., Endo, M., Sugiyama, H., and Morii, T. (2012) Zinc-Finger Proteins for Site-Specific Protein Positioning on DNA-Origami Structures. *Angew. Chem. 124*, 2471–2474.

[257] Rothemund, P. W. (2006) Folding DNA to create nanoscale shapes and patterns. *Nature 440*, 297–302.

[258] Douglas, S. M., Marblestone, A. H., Teerapittayanon, S., Vazquez, A., Church, G. M., and Shih, W. M. (2009) Rapid prototyping of 3D DNA-origami shapes with caDNAno. *Nucleic Acids Res. 37*, 5001–5006.

[259] Rajendran, A., Endo, M., and Sugiyama, H. (2012) DNA Origami: Synthesis and Self-Assembly. *Curr. Protoc. Nucleic Acid Chem.* 12–9.

[260] Rajendran, A., Endo, M., and Sugiyama, H. (2012) Single-Molecule Analysis Using DNA Origami. *Angew. Chem. Int. Ed. 51*, 874–890.

[261] Yoshidome, T., Endo, M., Kashiwazaki, G., Hidaka, K., Bando, T., and Sugiyama, H. (2012) Sequence-Selective Single-Molecule Alkylation with a Pyrrole–Imidazole Polyamide Visualized in a DNA Nanoscaffold. *J. Am. Chem. Soc. 134*, 4654–4660.

[262] Helmig, S., Rotaru, A., Arian, D., Kovbasyuk, L., Arnbjerg, J., Ogilby, P. R., Kjems, J., Mokhir, A., Besenbacher, F., and Gothelf, K. V. (2010) Single molecule atomic force microscopy studies of photosensitized singlet oxygen behavior on a DNA origami template. *ACS nano 4*, 7475–7480.

[263] Merrifield, R. B. (1963) Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc. 85*, 2149–2154.

[264] Garcia-Martin, F., Quintanar-Audelo, M., Garcia-Ramos, Y., Cruz, L. J., Gravel, C., Furic, R., Cote, S., Tulla-Puche, J., and Albericio, F. (2006) ChemMatrix, a poly (ethylene glycol)-based support for the solid-phase synthesis of complex peptides. *J. Comb. Chem. 8*, 213–220.

[265] Meldal, M. (1997) [6] Properties of solid supports. *Methods Enzymol. 289*, 83–104.

[266] Information from suppliers: Merck Novabiochem Peptide Synthesis catalogue 2008/2009, pp 2.4-2.5, fig 2-6: Structure of NovaPEG resin. IRIS Biotech catalogue Edition 2006, p. 203. Matrix Innovation homepage: Chemical structure of Aminomethyl-ChemMatrix.

[267] Lu, J., and Toy, P. H. (2009) Organic polymer supports for synthesis and for reagent and catalyst immobilization. *Chem. Rev. 109*, 815–838.

[268] Gaggini, F., Porcheddu, A., Reginato, G., Rodriquez, M., and Taddei, M. (2004) Colorimetric tools for solid-phase organic synthesis. *J. Comb. Chem. 6*, 805–810.

[269] Vazquez, J., Qushair, G., and Albericio, F. (2003) Qualitative colorimetric tests for solid phase synthesis. *Methods Enzymol. 369*, 21–35.

[270] Kaiser, E., Colescott, R., Bossinger, C., and Cook, P. (1970) Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal. Biochem. 34*, 595–598.

[271] Hancock, W., and Battersby, J. (1976) A new micro-test for the detection of incomplete coupling reactions in solid-phase peptide synthesis using 2, 4, 6-trinitrobenzene-sulphonic acid. *Anal. Biochem. 71*, 260–264.

[272] Madder, A., Farcy, N., Hosten, N. G., De Muynck, H., De Clercq, P. J., Barry, J., and Davis, A. P. (1999) A Novel Sensitive Colorimetric Assay for Visual Detection of Solid-Phase Bound Amines. *Eur. J. Org. Chem. 1999*, 2787–2791.

[273] Van der Plas, S. E., De Clercq, P. J., and Madder, A. (2007) Fast and easy detection of aromatic amines on solid support. *Tetrahedron Lett. 48*, 2587–2589.

[274] Caroen, J., and Van der Eycken, J. (2009) A sensitive and practical colorimetric test for polymer-supported hydroxyl and thiol groups. *Tetrahedron Lett. 50*, 41–44.

[275] Botana, E., Ongeri, S., Arienzo, R., Demarcus, M., Frey, J. G., Piarulli, U., Potenza, D., Kilburn, J. D., and Gennari, C. (2001) Synthesis, Conformational Studies and Binding Properties of Acyclic Receptors for N-Protected Amino Acids and Dipeptides. *Eur. J. Org. Chem. 2001*, 4625–4634.

[276] Sanclimens, G., Crespo, L., Giralt, E., Albericio, F., and Royo, M. (2005) Preparation of de novo globular proteins based on proline dendrimers. *J. Org. Chem. 70*, 6274–6281.

[277] Bayo-Puxan, N., N.-Puxan, Tulla-Puche, J., and Albericio, F. (2009) Oxathiocoraline: Lessons to be Learned from the Synthesis of Complex N-Methylated Depsipeptides. *Eur. J. Org. Chem. 2009*, 2957–2974.

[278] Bayo-Puxan, N., Fernando Albericiondez, A., Tulla-Puche, J., Riego, E., Cuevas, C., Alvarez, M., and Albericio, F. (2006) Total Solid-Phase Synthesis of the Azathiocoraline Class of Symmetric Bicyclic Peptides. *Chem. Eur. J. 12*, 9001–9009.

[279] Fields, G., Tian, Z., Barany, G., and Grant, G. (1992) Synthetic Peptides: A User's Guide. *WH Freeman, Salt Lake City, UT* 77–183.

[280] Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem. 72*, 248–254.

[281] Smith, P., Krohn, R. I., Hermanson, G., Mallia, A., Gartner, F., Provenzano, M., Fujimoto, E., Goeke, N., Olson, B., and Klenk, D. (1985) Measurement of protein using bicinchoninic acid. *Anal. Biochem. 150*, 76–85.

[282] Layne, E. (1957) [73] Spectrophotometric and turbidimetric methods for measuring proteins. *Methods Enzymol. 3*, 447–454.

[283] Lenevich, S., and Distefano, M. D. (2011) Nuclear magnetic resonance-based quantification of organic diphosphates. *Anal. Biochem. 408*, 316–320.

[284] Holzgrabe, U. (2010) Quantitative NMR spectroscopy in pharmaceutical applications. *Prog. Nucl. Magn. Reson. Spectrosc. 57*, 229–240.

[285] Pauli, G. F., Jaki, B. U., and Lankin, D. C. (2005) Quantitative 1H NMR: Development and Potential of a Method for Natural Products Analysis. *J. Nat. Prod. 68*, 133–149.

[286] Akoka, S., Barantin, L., and Trierweiler, M. (1999) Concentration measurement by proton NMR using the ERETIC method. *Anal. Chem. 71*, 2554–2557.

[287] Wider, G., and Dreier, L. (2006) Measuring protein concentrations by NMR spectroscopy. *J. Am. Chem. Soc. 128*, 2571–2576.

[288] Farrant, R. D., Hollerton, J. C., Lynn, S. M., Provera, S., Sidebottom, P. J., and Upton, R. J. (2010) NMR quantification using an artificial signal. *Magn. Reson. Chem. 48*, 753–762.

[289] Sambrook, J., Fritsch, E. F., and Maniatis, T. *Molecular cloning*; Cold spring harbor laboratory press New York, 1989; Vol. 2.

[290] Wallace, B. R., and Miyada, G. C. (1987) [47] Oligonucleotide probes for the screening of recombinant DNA libraries. *Methods Enzymol. 152*, 432–442.

[291] Borer, P. In *Handbook of Biochemistry and Molecular Biology: Nucleic Acids*; Fasman, G., Ed.; CRC Press, Cleveland, OH, 1975.

[292] Kallansrud, G., and Ward, B. (1996) A comparison of measured and calculated single-and double-stranded oligodeoxynucleotide extinction coefficients. *Anal. Biochem. 236*, 134–138.

[293] Demon, D. et al. (2009) Proteome-wide substrate analysis indicates substrate exclusion as a mechanism to generate caspase-7 versus caspase-3 specificity. *Molec. Cell. Proteomics 8*, 2700–2714.

[294] Figaroli, S., and Madder, A. (2010) Design and automated generation of artificial estrogen receptor as potential endocrine disruptor chemical binders. *Tetrahedron 66*, 6912–6918.

[295] Hawker, C. J., and Wooley, K. L. (2005) The convergence of synthetic organic and polymer chemistries. *Science 309*, 1200–1205.

[296] Barner-Kowollik, C., and Inglis, A. J. (2009) Has Click Chemistry Lead to a Paradigm Shift in Polymer Material Design? *Macromol. Chem. Phys. 210*, 987–992.

[297] Mansfeld, U., Pietsch, C., Hoogenboom, R., Becer, C. R., and Schubert, U. S. (2010) Clickable initiators, monomers and polymers in controlled radical polymerizations - a prospective combination in polymer science. *Polym. Chem. 1*, 1560–1598.

[298] Kempe, K., Krieg, A., Becer, C. R., and Schubert, U. S. (2012) "Clicking" on/with polymers: a rapidly expanding field for the straightforward preparation of novel macromolecular architectures. *Chem. Soc. Rev. 41*, 176–191.

[299] Foster, E. J., Berda, E. B., and Meijer, E. (2009) Metastable supramolecular polymer nanoparticles via intramolecular collapse of single polymer chains. *J. Am. Chem. Soc. 131*, 6964–6966.

[300] Brudno, Y., and Liu, D. R. (2009) Recent Progress Toward the Templated Synthesis and Directed Evolution of Sequence-Defined Synthetic Polymers. *Chem. Biol. 16*, 265–276.

[301] Schmidt, B. V. K. J., Fechler, N., Falkenhagen, J., and Lutz, J.-F. (2011) Controlled folding of synthetic polymer chains through the formation of positionable covalent bridges. *Nature Chem. 3*, 234–238.

[302] Ouchi, M., Badi, N., Lutz, J.-F., and Sawamoto, M. (2011) Single-chain technology using discrete synthetic macromolecules. *Nature Chem. 3*, 917–924.

[303] Terashima, T., Mes, T., De Greef, T. F., Gillissen, M. A., Besenius, P., Palmans, A. R., and Meijer, E. (2011) Single-chain folding of polymers for catalytic systems in water. *J. Am. Chem. Soc. 133*, 4742–4745.

[304] Altintas, O., and Barner-Kowollik, C. (2012) Single Chain Folding of Synthetic Polymers by Covalent and Non-Covalent Interactions: Current Status and Future Perspectives. *Macromol. Rapid Commun. 33*, 958–971.

[305] Baradel, N., Fort, S., Halila, S., Badi, N., and Lutz, J.-F. (2013) Synthesis of Single-Chain Sugar Arrays. *Angew. Chem. Int. Ed. 125*, 2391–2395.

[306] Zhang, Q., Collins, J., Anastasaki, A., Wallis, R., Mitchell, D. A., Becer, C. R., and Haddleton, D. M. (2013) Sequence-Controlled Multi-Block Glycopolymers to Inhibit DC-SIGN-gp120 Binding. *Angew. Chem. 125*, 4531–4535.

[307] Badi, N., and Lutz, J.-F. (2009) Sequence control in polymer synthesis. *Chem. Soc. Rev. 38*, 3383–3390.

[308] Hartmann, L., and Borner, H. G. (2009) Precision Polymers: Monodisperse, Monomer-Sequence-Defined Segments to Target Future Demands of Polymers in Medicine. *Advanced Materials 21*, 3425–3431.

[309] Borner, H. G. (2011) Precision Polymers - Modern Tools to Understand and Program Macromolecular Interactions. *Macromol. Rapid Commun. 32*, 115–126.

[310] Lutz, J.-F. (2013) Writing on Polymer Chains. *Acc. Chem. Res. doi:10.1021/ar400097a.*

[311] Lutz, J. F. (2010) Sequence-controlled polymerizations: the next Holy Grail in polymer science? *Polym. Chem. 1*, 55–62.

[312] Pfeifer, S., and Lutz, J.-F. (2007) A facile procedure for controlling monomer sequence distribution in radical chain polymerizations. *J. Am. Chem. Soc. 129*, 9542–9543.

[313] Pfeifer, S., and Lutz, J.-F. (2008) Development of a Library of N-Substituted Maleimides for the Local Functionalization of Linear Polymer Chains. *Chem. Eur. J. 14*, 10949–10957.

[314] McKee, M. L., Milnes, P. J., Bath, J., Stulz, E., Turberfield, A. J., and O'Reilly, R. K. (2010) Multistep DNA-Templated Reactions for the Synthesis of Functional Sequence Controlled Oligomers. *Angew. Chem. Int. Ed. 49*, 7948–7951.

[315] Ida, S., Ouchi, M., and Sawamoto, M. (2010) Template-Assisted Selective Radical Addition toward Sequence-Regulated Polymerization: Lariat Capture of Target Monomer by Template Initiator. *J. Am. Chem. Soc. 132*, 14748–14750.

[316] Hibi, Y., Ouchi, M., and Sawamoto, M. (2011) Sequence-Regulated Radical Polymerization with a Metal-Templated Monomer: Repetitive ABA Sequence by Double Cyclopolymerization. *Angew. Chem. Int. Ed. 50*, 7434–7437.

[317] Ida, S., Ouchi, M., and Sawamoto, M. (2011) Designer Template Initiator for Sequence Regulated Polymerization: Systems Design for Substrate-Selective Metal-Catalyzed Radical Addition and Living Radical Polymerization. *Macromol. Rapid Commun. 32*, 209–214.

[318] Bertran-Vicente, J., and Hackenberger, C. P. (2013) A Supramolecular Peptide Synthesizer. *Angew. Chem. Int. Ed. 52*, 6140–6142.

[319] Lewandowski, B., De Bo, G., Ward, J. W., Papmeyer, M., Kuschel, S., Aldegunde, M. J., Gramlich, P. M., Heckmann, D., Goldup, S. M., DSouza, D. M., Fernandes, A. E., and

Leigh, D. A. (2013) Sequence-specific peptide synthesis by an artificial small-molecule machine. *Science 339*, 189–193.

[320] Hartmann, L. (2011) Polymers for Control Freaks: Sequence-Defined Poly(amidoamine)s and Their Biomedical Applications. *Macromol. Chem. Phys. 212*, 8–13.

[321] Mosca, S., Wojcik, F., and Hartmann, L. (2011) Precise Positioning of Chiral Building Blocks in Monodisperse, Sequence-Defined Polyamides. *Macromol. Rapid Commun. 32*, 197–202.

[322] Pfeifer, S., Zarafshani, Z., Badi, N., and Lutz, J.-F. (2009) Liquid-Phase Synthesis of Block Copolymers Containing Sequence-Ordered Segments. *J. Am. Chem. Soc. 131*, 9195–9197.

[323] Pfeifer, S., and Lutz, J.-F. (2010) Tailor-Made Soluble Polymer Supports: Synthesis of a Series of ATRP Initiators Containing Labile Wang Linkers. *Macromol. Chem. Phys. 211*, 940–947.

[324] Meszynska, A., Badi, N., Boerner, H. G., and Lutz, J.-F. (2012) "Inverse" synthesis of polymer bioconjugates using soluble supports. *Chem. Commun. 48*, 3887–3889.

[325] Fields, G. B., and Noble, R. L. (1990) Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int. J. Pept. Protein Res. 35*, 161–214.

[326] Beaucage, S. L., and Iyer, R. P. (1992) Advances in the synthesis of oligonucleotides by the phosphoramidite approach. *Tetrahedron 48*, 2223–2311.

[327] Holub, J. M., Jang, H. J., and Kirshenbaum, K. (2006) Clickity-click: highly functionalized peptoid oligomers generated by sequential conjugation reactions on solid-phase support. *Org. Biomol. Chem. 4*, 1497–1502.

[328] Caumes, C., Hjelmgaard, T., Remuson, R., Faure, S., and Taillefumier, C. (2011) Highly Convenient Gram-Scale Solution-Phase Peptoid Synthesis and Orthogonal Side-Chain Post-Modification. *Synthesis-Stuttgart* 257–264.

[329] Liskamp, R. M. J., Rijkers, D. T. S., Kruijtzer, J. A. W., and Kemmink, J. (2011) Peptides and Proteins as a Continuing Exciting Source of Inspiration for Peptidomimetics. *Chembiochem 12*, 1626–1653.

[330] Espeel, P., Goethals, F., and Du Prez, F. E. (2011) One-Pot Multistep Reactions Based on Thiolactones: Extending the Realm of Thiol-Ene Chemistry in Polymer Synthesis. *J. Am. Chem. Soc. 133*, 1678–1681.

[331] Espeel, P., Goethals, F., Stamenovic, M. M., Petton, L., and Du Prez, F. E. (2012) Double modular modification of thiolactone-containing polymers: towards polythiols and derived structures. *Polym. Chem. 3*, 1007–1015.

[332] Stamenovic, M. M., Espeel, P., Baba, E., Yamamoto, T., Tezuka, Y., and Du Prez, F. E. (2013) Straightforward synthesis of functionalized cyclic polymers in high yield via RAFT and thiolactone-disulfide chemistry. *Polym. Chem. 4*, 184–193.

[333] Reinicke, S., Espeel, P., Stamenovic, M. M., and Du Prez, F. E. (2013) One-Pot Double Modification of p (NIPAAm): A Tool for Designing Tailor-Made Multiresponsive Polymers. *ACS Macro Letters 2*, 539–543.

[334] Espeel, P., Goethals, F., Driessen, F., Nguyen, L.-T. T., and Du Prez, F. E. (2013) One-pot, additive-free preparation of functionalized polyurethanes via amine–thiol–ene conjugation. *Polym. Chem. 4*, 2449–2456.

[335] Mommer, S., Lamberts, K., Keul, H., and Möller, M. (2013) A novel multifunctional coupler; Concept of coupling and proof of principle. *Chem. Commun. 49*, 3288–3290.

[336] Yan, J.-J., Wang, D., Wu, D.-C., and You, Y.-Z. (2013) Synthesis of sequence-ordered polymers via sequential addition of monomers in one pot. *Chem. Commun. 49*, 6057– 6059.

[337] Hoyle, C. E. (2010) Thiol-click chemistry: a multifaceted toolbox for small molecule and polymer synthesis. *Chem. Soc. Rev. 39*, 1355–1387.

[338] Gokmen, M. T., Brassinne, J., Prasath, R. A., and Du Prez, F. E. (2011) Revealing the nature of thio-click reactions on the solid phase. *Chem. Commun. 47*, 4652–4654.

[339] Barlos, K., Chatzi, O., Gatos, D., and Stavropoulos, G. (1991) 2-Chlorotrityl chloride resin - studies on anchoring of Fmoc amino acids and peptide cleavage. *Int. J. Pept. Protein Res. 37*, 513–520.

[340] Bartolozzi, A., Foudoulakis, H. M., and Cole, B. M. (2008) Development of a tandem base-catalyzed, triphenylphosphine-mediated disulfide reduction-Michael addition. *Synthesis-Stuttgart* 2023–2032.

[341] Li, G.-Z., Randev, R. K., Soeriyadi, A. H., Rees, G., Boyer, C., Tong, Z., Davis, T. P., Becer, C. R., and Haddleton, D. M. (2010) Investigation into thiol-(meth)acrylate Michael addition reactions using amine and phosphine catalysts. *Polym. Chem. 1*, 1196–1204.

[342] Chan, J. W., Hoyle, C. E., Lowe, A. B., and Bowman, M. (2010) Nucleophile-Initiated Thiol-Michael Reactions: Effect of Organocatalyst, Thiol, and Ene. *Macromolecules 43*, 6381–6388.

[343] Gauthier, M. A., Gibson, M. I., and Klok, H. A. (2009) Synthesis of Functional Polymers by Post-Polymerization Modification. *Angew. Chem. Int. Ed. 48*, 48–58.

[344] Gunay, K. A., Theato, P., and Klok, H.-A. (2013) Standing on the shoulders of hermann staudinger: Post-polymerization modification from past to present. *J. Polym. Sci. A Polym. Chem. 51*, 1–28.

[345] Phillips, M. B., Sullivan, M. M., Villalta, P. W., and Peterson, L. A. (2013) Covalent Modification of Cytochrome c by Reactive Metabolites of Furan. *Chem. Res. Toxicol.* DOI: 10.1021/tx400368r.

[346] Peterson, L. A., Cummings, M. E., Vu, C. C., and Matter, B. A. (2005) Glutathione trapping to measure microsomal oxidation of furan to cis-2-butene-1,4-dial. *Drug Metab. Disp. 33*, 1453–1458.

[347] Khojasteh, S. C., Oishi, S., and Nelson, S. D. (2010) Metabolism and Toxicity of Mentho-furan in Rat Liver Slices and in Rats. *Chem. Res. Toxicol. 23*, 1824–1832.

[348] Weinberg, N. L., and Weinberg, H. R. (1968) Electrochemical oxidation of organic compounds. *Chem. Rev. 68*, 449–523.

[349] Nessakh, B., Kotkowska-Machnik, Z., and Tedjar, F. (1990) Electrochemical behaviour of furan, 2-methylfuran and 2,5-dimethylfuran in acetonitrile. *J. Electroanal. Chem. Interf. Electrochem. 296*, 263 – 268.

[350] Malode, S. J., Abbar, J. C., Shetti, N. P., and Nandibewoor, S. T. (2012) Voltammetric oxidation and determination of loop diuretic furosemide at a multi-walled carbon nanotubes paste electrode. *Electrochim. Acta 60*, 95 – 101.

[351] Oliveira Brett, A. M., and Serrano, S. H. P. (1995) The Electrochemical Oxidation of DNA. *J. Braz. Chem. Soc. 6*, 97–100.

[352] Ivandini, T. A., Sarada, B. V., Rao, T. N., and Fujishima, A. (2003) Electrochemical oxidation of underivatized-nucleic acids at highly boron-doped diamond electrodes. *Analyst. 128*, 924–929.

[353] Wang, Q., Zheng, M., Shi, J., Gao, F., and Gao, F. (2011) Electrochemical Oxidation of Native Double-Stranded DNA on a Graphene-Modified Glassy Carbon Electrode. *Electroanal. 23*, 915–920.

[354] Kim, S., Hirose, K., Uematsu, J., Mikami, Y., and Chiba, K. (2012) Electrochemically Active Cross-Linking Reaction for Fluorescent Labeling of Aliphatic Alkenes. *Chem. Eur. J. 18*, 6284–6288.

[355] Du, H., Fuh, R.-C. A., Li, J., Corkan, L. A., and Lindsey, J. S. (1998) PhotochemCAD: A Computer-Aided Design and Research Tool in Photochemistry. *Photochem. Photobiol. 68*, 141–142.

[356] McGee, D. P., Vaughn-Settle, A., Vargeese, C., and Zhai, Y. (1996) 2'-Amino-2'-deoxyuridine via an intramolecular cyclization of a trichloroacetimidate. *J. Org. Chem. 61*, 781–785.

[357] Kirschenheuter, G. P., Zhai, Y., and Pieken, W. A. (1994) An improved synthesis of 2 -azido-2 -deoxyuridine. *Tetrahedron Lett. 35*, 8517–8520.

[358] Lin, F. L., Hoyt, H. M., Van Halbeek, H., Bergman, R. G., and Bertozzi, C. R. (2005) Mechanistic investigation of the Staudinger ligation. *J. Am. Chem. Soc. 127*, 2686–2695.

Curriculum Vitae

Personal Data

Lieselot *Leen Georges* Carrette
born on the August 13th, 1986 in Tielt, Belgium.

Education and Research Experience

Education

- PhD student in Sciences: Chemistry at Ghent University, Organic and Biomimetic Chemistry Group of Prof. Dr. A. Madder

    Scientific Assistant and Marie Curie Project Manager (01-10-2013 - 31-03-2014)

    Aspirant of the Fund for Scientific Research of Flanders (01-10-2009 - 30-09-2013)

- Master in Chemistry with the highest distinction at Ghent University, 01-07-2009, Ghent.

    Masterthesis in Organic and Biomimetic Chemistry Group of Prof. Dr. A. Madder on the topic: Evaluation of the Linear Synthesis Approach Towards a Dipodal Peptidosteroid as a Miniature DNA Major Groove Binding Protein

- Bachelor in Chemistry with the highest distinction at Ghent University, 04-07-2007, Ghent.

Additional courses

- COST Epigenetic Training School, May 21-24th 2013, Rome, Italy.

- Uppsala University regular course: 'Pharmaceutical Bioinformatics' by Prof. J. Wikberg, 2nd semester of the academic year 2012-2013

- UGent regular course: 'Moleculaire Simulaties bij Biosystemen' by Dr. E. Pauwels, 1st semester of the academic year 2011-2012

- UGent regular course: 'Basic Japanese I' by Prof. Niehaus, 1st semester of the academic year 2009-2010

- Roche Continents: 'Explore the common ground of creativity and innovation in Arts & Science!', August 4-10th 2009, Salzburg, Austria.

- HRSMC Summer School: 'New Horizons in Synthetic Methodology', July 6-10th 2009, Maastricht, The Netherlands.

International Short Research Stays

- Laboratory of Prof. H. Sugiyama, Kyoto University: 'Visualisation of furan based interstrand cross-linking on DNA origami', July 9th-26th 2012, Kyoto, Japan.

- Laboratory of Prof. T. Carell, Ludwig-Maximilians-Universitt Mnchen: 'Incorporaton of furan modified amino acid analogs in proteins', November 21st December 15th 2011, Munich, Germany.

- Laboratory of Prof. T. Morii, Biofunctional Science, Institute of Advanced Energy, Kyoto University: 'Detection of peptide-DNA interactions via cross-linking: bidirectional use of the furan-oxidation methodology', April 1st - November 31st 2010, Kyoto, Japan.

- Analytical R&D Lab, Pfizer: 'SFC COMET: Orthogonal Selectivity In Packed Column Supercritical Fluid Chromatography For Pharmaceutical Analysis: Selectivity Tuning On The Stationary Phase', July 9th August 31st 2007, Sandwich, United Kingdom.

Awards & Funding

Scientific Awards

- Travel Award by the scientific committee to attend the 23rd American Peptide Symposium (APS), in conjugation with the 6th International Peptide Symposium, June 22nd - 27th 2013, Hawaii, United States.

- Poster Prize B at the Belgian-German (Macro)Molecular Meeting, December 3-4 2012 Houffalize, Belgium.

- Travel Award by the scientific committee to attend the 32nd European Peptide Symposium (EPS), September 2nd- 9th 2012, Athens, Greece.

- Travel Award by the scientific committee to attend the XX International Round Table, August 5th - 9th 2012, Montreal, Canada.

- Best Poster- Organic Chemistry, Award from the Royal Flemisch Chemical Society March 2nd 2012, Blankenberge, Belgium

- DOW Chemical Company Award 2009 (award for the best chemistry student at UGent)

- Selected for Roche Continents 2009 (100 best students of science, music and fine arts in Europe)

Funding

- Renewal Ph.D grant from the Fonds voor Wetenschappelijk Onderzoek Vlaanderen (FWO): aspirant 2011-2013

- Collaboration project from the Fonds voor Wetenschappelijk Onderzoek Vlaanderen (FWO) and the Japan Society for the Promotion of Science (JSPS): 'Design and synthesis of functional biomacromolecules with inducible reactivity: tools for chemical biology', Prof. Madder, A. and Morii, T.

- Mobility grant from the Bijzonder Onderzoeks Fonds UGent (BOF) for a research stay in Kyoto, Japan (April - December 2010)

- Ph.D grant from the Fonds voor Wetenschappelijk Onderzoek Vlaanderen (FWO): aspirant 2009-2011

- Additional Support to cover travel costs for attending meetings:

  - Trainee grant from the Management Committee of the Cost Action TD0905 to participate to the COST Epigenetic Training School, May 21-24th 2013, Rome, Italy.

  - Travel grant (part of the travel costs) from the European Science Foundation and from the Faculty Committee for Scientific Research (registration costs and remaining travel costs) to attend EST-COST High-Level Research Conference on Systems Chemistry III, October 23th-28th 2011, Crete, Greece

Scientific Output

A1 publications:

- Carrette L.L.G., Morii T. & Madder A. Peptidosteroid tweezers revisited: DNA binding through an optimized design. Eur. J. Org. Chem. (accepted).

- Carrette L.L.G., Gyssels E., Loncke J. & Madder A. Mildly inducible and selective cross-link methodology for RNA duplexes. Org. Biomol. Chem. 2014, 12,931-935.

- Carrette L.L.G.* & Madder A.* A synthetic oligonucleotide model to evaluate oxidation and cross-linking propensity of natural furan modified DNA. ChemBioChem 2014, 15(1), 103-107 (*co-corresponding authors, inside cover).

- Carrette L.L.G, Morii T. & Madder A. Toxicity Inspired Cross-linking for Probing DNA - Peptide Interactions. Bioconj. Chem. 2013, 24(12), 20082014.

- Espeel P.*, Carrette L.L.G.*, Bury K., Capenberghs S., Martens J., Du Prez F. & Madder A. Multi-functionalized sequence-defined oligomers from a single building block. Angew. Chem. Int. Ed. 2013, 52(50), 13261-13264 (*equal contribution of both first authors).

- Carrette L.L.G., Verzele D. & Madder A. NF-31 color test uncovers 'hidden' alcohol functionalities in PEG-based resins for solid phase peptide synthesis. Tetrahedron Lett. 2010, 51(16), 2106-2108.

A2 publication:

- Verzele, D.*; Carrette, L.L.G.* & Madder, A. Peptide scalpels for site-specific dissection of the DNAprotein interface. Drug Discovery Today: Technologies 2010, 7(2), 115 (*equal contribution of both authors, upon invitation for the themed issue: Drug Discovery beyond the rule of five).

Patent:

- Madder, A.; Hoogewijs, K.; Deceuninck, A. & Carrette, L. Method for Crosslinking Peptides. 2012, European Patent WO2012085279.

Book Chapter:

- Carrette L.L.G, Gyssels E. & Madder A. DNA Interstrand Cross-link Formation by using Furan as a Masked Reactive Aldehyde. Curr. Prot. Nucl. Acid. Chem. 2013, 54:5.12.1-5.12.16

Meetings, Symposia and Conferences (with associated C publications)

Oral presentations as presenter:

- 'Furan Oxidation Cross-linking: A New Approach for the Study and Targeting of Protein and Nucleic Acid Interactions' L.L.G. Carrette, E. Gyssels, N. De Laet, T. Morii, A. Madder, WE07 Departmental Scientific Workshop, UGent, March 6-7th 2014, Gent, Belgium.
- 'Furan Oxidation Cross-linking: A New Approach for the Study and Targeting of Protein and Nucleic Acid Interactions' L.L.G. Carrette, E. Gyssels, J. Loncke, T. Morii, A. Madder, ChemCys, 12th Chemistry Conference for Young Scientists, February 27-28th 2014, Blankenberge, Belgium.
- 'Metabolism inspired crosslinking: the furan oxidation strategy' Carrette L.L.G, Gyssels E., Hoogewijs, K., Op de Beeck M., Stevens K., Madder A., company visit ISIS Pharmaceuticals, July 2nd 2013, Carlsbad, United States.
- 'Study on the application of Furan Crosslinking at the Protein-DNA Interface' Carrette Lieselot L.G., Gattner Michael, Carell Thomas, Morii Takashi and Madder Annemieke; 16th Sigma-Aldrich Organic Synthesis Meeting, December 6-7th 2012, Spa, Belgium.
- 'Application of Furan Crosslinking on the DNA-Protein Interface' Carrette L.L.G., Op de Beeck M., Morii T., Madder A.; Sugiyama group visit, July 13th 2012, Kyoto, Japan.
- 'Progress in the application of furan crosslinking on the DNA-protein interface' Lieselot Carrette and Annemieke Madder; Network Meeting and Midterm Review PhosChemRec, May 22nd-24th 2012, Gent, Belgium.
- 'Furan as masked reactive aldehyde: beyond DNA recognition, towards peptide-DNA crosslinking' Lieselot Carrette and Annemieke Madder; Network Meeting PhosChemRec, July 2nd-4th 2011, Stockholm, Sweden.

- 'Evaluation of the linear synthesis of a transcription factor mimic' Lieselot Carrette, Dieter Verzele and Annemieke Madder; Network Meeting PhosChemRec, February 16th-18th 2011, Padua, Italy.

- 'Design of crosslinking transcription factor mimics' Lieselot Carrette and Annemieke Madder; Annual One-Day Meeting on Medicinal Chemistry of SRC & KVCV, Does size matter? Beyond small molecule therapeutics: challenges and success stories, November 6th 2009, Brussels, Belgium (poster selected for oral presentation)

Oral presentations of work presented by professor Madder

- 'Furan-Based Crosslinking Approaches at the DNA/DNA and DNA/Protein interface' Carrette, L.L.G.; Gyssels, E.; Op de Beeck, M. and Madder, A.;Drug Discovery & Therapy World Congress, June 3rd-6th 2013, Boston, United States.

- 'Furan-Based Oligonucleotides for Cross-linking at the DNA/DNA and DNA/ Protein Interface' Carrette, L.L.G.; Gyssels, E.; Op de Beeck, M. and Madder, A.; XXth International Round Table on Nucleosides Nucleotides and Nucleic Acids, August 5-9th 2012, Montreal, Canada (abstract selected for oral presentation)

- 'Furan-Based Crosslinking Approaches at the DNA/DNA and DNA/Protein interface' Annemieke Madder, Lieselot Carrette, Ellen Gyssels and Marieke Op de Beeck; Epigenetics Bench to Bedside, COST Action TG 0905, April 26-27th 2012, Riga, Latvia.

- 'Cholic acid based peptidosteroids as miniature proteins with defined structure' A. Madder, C. Bode, L. Carrette, Y. Ruiz and D. Verzele, 1st Belgian Peptide Group Meeting, February 9th - 10th 2012, Brussels, Belgium.

Poster presentations as presenter:

- 'Furan Oxidation Cross-linking: A New Approach for the Study and Targeting of Protein and Nucleic Acid Interactions' L.L.G. Carrette, E. Gyssels, J. Loncke, T. Morii, A. Madder, ChemCys, 12th Chemistry Conference for Young Scientists, February 27-28th 2014, Blankenberge, Belgium.

- 'Furan Oxidation Cross-linking: A New Approach for the Study and Targeting of Protein and Nucleic Acid Interactions' L.L.G. Carrette, E. Gyssels, J. Loncke, T. Morii, A. Madder, 2nd Belgian Pepetide Group Meeting, February 8-9th 2014, Blankenberge, Belgium.

- 'Furan Oxidation Cross-linking: A New Approach for the Study and Targeting of Protein and Nucleic Acid Interactions' L.L.G. Carrette, E. Gyssels, J. Loncke, T. Morii, A. Madder, Sigma Aldrich Synthesis meeting, Decemmber 5th-6th 2013, Blankenberge, Belgium.

- 'Furan Oxidation Cross-linking: A New Approach for the Study and Targeting of Protein and Nucleic Acid Interactions' L.L.G. Carrette, E. Gyssels, J. Loncke, T. Morii, A. Madder, Annual One-Day Meeting on Medicinal Chemistry of SRC & KVCV, 'Constrained

Peptides and Macrocycles New Opportunities for Drug Discovery' November 22nd 2013, Beerse, Belgium.

- 'Furan Oxidation Cross-linking: A New Approach for the Study and Targeting of Protein and Nucleic Acid Interactions' L.L.G. Carrette, E. Gyssels, J. Loncke, A. Madder, 9th Annual Meeting of the Oligonucleotide Therapeutics Society, October 6th 8th 2013, Naples, Italy.

- 'Crosslinking Transcription Factor Mimics' L.L.G. Carrette, T. Morii, A. Madder; 23th American Peptide Symposium, in conjugation with the 6th International Peptide Symposium, June 22nd - 27th 2013, Hawaii, United States.

- 'Inspired Crooslinking of DNA to its Binding Proteins' Lieselot Carrette, Marieke Op de Beeck, Takashi Morii, Annemieke Madder; Belgian-German (Macro)Molecular Meeting, December 3-4th 2012, Houffalize, Belgium.

- 'From Sequence Defined Oligomers to Sequence Controlled Polymers through a Thiolactone Coupling Strategy' Lieselot Carrette, Pieter Espeel, Sven Capenberghs, Katarzyna Bury, Andreas Baudot, Dorien Van Lysebetten, Matthias Van Zele, Annemieke Madder, Filip Du Prez; Belgian-German (Macro)Molecular Meeting, December 3-4th 2012, Houffalize, Belgium (Poster Prize B).

- 'Study on the Application of Furan Crosslinking on the Protein-DNA Interface' Carrette Lieselot L.G., Gattner Michael, Van den Begin Jos, Carell Thomas, Morii Takashi and Madder Annemieke; 32nd European Peptide Symposium, September 2nd-7th 2012, Athens, Greece (Travel Award granted to attend: registration).

- 'Toxicity Inspired Crosslinking of DNA to its Binding Proteins' Carrette, L.L.G.; Op de Beeck, M.; Morii, T. And Madder, A.; XXth International Round Table on Nucleosides Nucleotides and Nucleic Acids, August 5-9th 2012, Montreal, Canada (Travel Award granted to attend: registration, travel and accommodation costs).

- 'Study on the Application of Furan Crosslinking on the Protein-DNA Interface' Carrette Lieselot L.G., Gattner Michael, Van den Begin Jos, Carell Thomas, Morii Takashi and Madder Annemieke; Kekule-cyclus XIV 'Chemistry, part of your life', 'Doorbraken in Farmaceutische R&D' March 27th 2012, Antwerp, Belgium (Poster presentation upon invitation as a result of the award best poster Organic Chemistry on ChemCys 2012).

- 'Study on the Application of Furan Crosslinking on the Protein-DNA Interface' Carrette Lieselot L.G., Gattner Michael, Van den Begin Jos, Carell Thomas, Morii Takashi and Madder Annemieke; 11th Flemisch Youth Conference of Chemistry, March 1st-2nd 2012, Blankenberge, Belgium (Award for best poster organic chemistry).

- 'Challenging the possibilities of peptide synthesis, towards a biomimetic transcription factor' Lieselot Carrette, Yara Ruiz Garcia, Jos Van den Begin and Annemieke Madder; EST-COST High-Level Research Conference on Systems Chemistry III, 23th-28th October 2011, Crete, Greece.

- 'Design of crosslinking transcription factor mimics' Lieselot Carrette and Annemieke Madder; Annual One-Day Meeting on Medicinal Chemistry of SRC & KVCV, Does size mat-

ter? Beyond small molecule therapeutics: challenges and success stories, 6th November 2009, Brussels, Belgium (poster selected for oral presentation).

- 'Development of Crosslinking Transcription Factor Mimics' Lieselot Carrette, Dieter Verzele and Annemieke Madder; 10th Flemisch Youth Conference of Chemistry, March 1st-2nd 2010, Blankenberge, Belgium.

- 'Development of Crosslinking Transcription Factor Mimics' Lieselot Carrette, Dieter Verzele and Annemieke Madder; 13th Sigma-Aldrich Organic Synthesis Meeting, 3-4th December 2009, Spa.

Assistance with the organisation of Meetings

- Symposium 'The Art of Chemistry and the Chemistry of Art', a scientific and social meeting between the organic chemistry groups of Prof. Madder, Prof. Van der Eycken and Prof. Hoogenboom from UGent and of Prof. Feringo, Prof. Harutyunyan, Prof. Otten and Prof. Barta from Groningen university, June 5th 7th, Gent, Belgium.

- Network Meeting and Midterm Review PhosChemRec, May 22nd-24th 2012, Gent, Belgium.

Academic Responsibilities:

Teaching:

- Theoretical exercise sessions on Organic Chemistry sessions for the 1st year Pharmaceutical Sciences (academic years 2009-2010, 2010-2011, 2011-2012)

- Theoretical exercise sessions on Organic Chemistry sessions for the 1st year Medicine (academic year 2013-2014)

- Practical exercise sessions on Organic Chemistry for the 1st year Pharmaceutical Sciences (academic year 2012-2013)

- Introduction day of Pharmaceutical Sciences 2014, presentation of Organic Chemistry

Support of Master Students:

- Andreas Baudot (academic year 2013-2014) 'Furan oxidation based cross-linking from peptides to their RNA binding site'

- Joke Loncke (academic year 2012-2013) 'Synthesis and evaluation of furan modified oligonucleotides for RNA crosslinking'

- Sven Capenberghs (academic year 2010-2011) 'Solid phase synthesis of sequence controlled polymers based on thiol click chemistry'; joined master thesis with the Polymer Research Group of Prof. Du Prez

Support of Bachelor Students:

- Andreas Baudot, Dorien Van Lysebetten and Mathias Van Zele (academic year 2011-2012): 'Sequence-controlled synthesis of oligomers based on thiolactone chemistry' joined project with the Polymer Research Group of Prof. Du Prez

Students representative:

- in the Selection Commission for New Assistant Professor in the Department of Inorganic and Physical Chemistry at UGent 2009.

- in the Study Programme Committee for the study programme Chemistry 2008-2009.