

Synthesis of DNA-encoded screening libraries

DISSERTATION

For the requirement of the academic degree of the
doctors in natural sciences

(Dr. rer. nat.)

Submitted to

The department of chemistry and chemical biology
TU Dortmund University

By

M. Sc. **Bugain *Epse*. Kamdem, Olivia**

From Cameroon

1. **Advisor: Prof. Dr. Daniel Rauh**
2. **Advisor: Prof. Dr. Martin Engelhard**

Synthese von DNA-kodierten Screening Substanzbibliotheken

DISSERTATION

zur Erlangung des akademischen Grades

Doktors der Naturwissenschaften

(Dr. rer. nat.)

Fakultät Chemie und Chemische Biologie

der Technischen Universität Dortmund

vorgelegt von

M. Sc. **Bugain *Epse.* Kamdem, Olivia**

- 1. Gutachter: Prof. Dr. Daniel Rauh**
- 2. Gutachter: Prof. Dr. Martin Engelhard**

Eingereicht am **14.07.2017**

Tag der mündliche Prüfung: **12.09.2017**



Eidesstattliche Versicherung (Affidavit)

Bugain Epse. Kamdem, Olivia

164139

Name, Vorname

(Surname, first name)

Matrikel-Nr.

(Enrolment number)

Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden, § 63 Abs. 5 Hochschulgesetz NRW.

Die Abgabe einer falschen Versicherung an Eides statt ist strafbar.

Wer vorsätzlich eine falsche Versicherung an Eides statt abgibt, kann mit einer Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft werden, § 156 StGB. Die fahrlässige Abgabe einer falschen Versicherung an Eides statt kann mit einer Freiheitsstrafe bis zu einem Jahr oder Geldstrafe bestraft werden, § 161 StGB.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offence can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offences of this type is the chancellor of the TU Dortmund University. In the case of multiple or other serious attempts at deception, the candidate can also be unenrolled, Section 63, paragraph 5 of the Universities Act of North Rhine-Westphalia.

The submission of a false affidavit is punishable.

Any person who intentionally submits a false affidavit can be punished with a prison sentence of up to three years or a fine, Section 156 of the Criminal Code. The negligent submission of a false affidavit can be punished with a prison sentence of up to one year or a fine, Section 161 of the Criminal Code.

I have taken note of the above official notification.

14.07.2017

Ort, Datum

(Place, date)

Unterschrift

(Signature)

Titel der Dissertation:

(Title of the thesis):

Synthese von DNA-kodierten Screening Substanzbibliotheken.

Synthesis of DNA-encoded screening libraries.

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel selbstständig und ohne unzulässige fremde Hilfe angefertigt habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht.

Die Arbeit hat in gegenwärtiger oder in einer anderen Fassung weder der TU Dortmund noch einer anderen Hochschule im Zusammenhang mit einer staatlichen oder akademischen Prüfung vorgelegen.

I hereby swear that I have completed the present dissertation independently and without inadmissible external support. I have not used any sources or tools other than those indicated and have identified literal and analogous quotations.

The thesis in its current version or another version has not been presented to the TU Dortmund University or another university in connection with a state or academic examination.*

***Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the PhD thesis is the official and legally binding version.**

14.07.2017

Ort, Datum

(Place, date)

Unterschrift

(Signature)

*“Commit everything you do to the Lord.
Trust him, and he will help you.”*

Psalm 37.5

Dedication:

To the three pillars of my life: God almighty, my husband and my parents. Without you, my life will fall apart. I might not know where the life's road will take me, but walking with You, God, through this journey has given me strength. You have given me so much, thanks for your faith in me, and for teaching me that I should never surrender.

ACKNOWLEDGEMENTS

Undertaking the PhD study at the technical university of Dortmund was a life-changing experience for me, and it would not have been accomplished without the help, support and collaboration of many people. I have here the great pleasure to thank them all.

First and foremost, I must express my sincere gratitude to my supervisor **Dr. Andreas Brunschweiger** for providing me the opportunity to undertake my PhD in his research group. I thank him for initiating the thesis project and giving me the possibility to use the excellent research facilities in his lab. His insightful discussions and enthusiastic determination are extremely valuable to me. I have immensely appreciated his guidance over the years of my PhD. I am extremely grateful for all his effort, advices, interpretations and for correcting all posters, presentations and drafts of papers. He helped me not only with priceless counsel in the field of DNA encoded chemistry; but he also broadened my horizon towards the application of chemistry to pharmaceutical. I thank him for recruiting me, being also always approachable and spent a huge amount of time reading and correcting this thesis.

I greatly acknowledge **Prof. Dr. Daniel Rauh** for offering me a warm welcome in the MedChem group. I sincerely thank him for his great support, as well as his guidance and encouragement through my PhD study. I thank him for the precious advices I received from him during my progress reports as well as all insightful suggestions and comments throughout my doctorate. This thesis would not have been possible without his coordination. His intelligent project guidance, his intervention and the many fruitful discussions were crucial for the completion of this thesis.

My sincere thanks go to **Prof Dr. Martin Engelhard** for taking over the second supervisor position. More than a second supervisor, he helped me a lot by his constant support, by his presence and assistance during all my TAC meetings. I thank him for all the time he has invested to follow my project and for all the constructive conversations, we had together. I thank him for his esteem, for believing in me and for being on my side. To see him in the audience during my reports was a great relief, helping me to stay focused and concentrated. I am also grateful for his help with proof reading and revisions of this thesis. I thank him for his help and advices with the use of EndNote. For his expertise and helpful advices. I thank him from the bottom of my heart. It was an outstanding experience having him as advisor.

For the collaboration on macrocycles synthesis I would like thank **Prof. Dr. Tom Großmann** together with his student Dr. Adrian Glas from the Chemical Genomic Center in Dortmund. I acknowledge them not only for the nice collaboration, but I thank Mrs Großmann for being my second IMPRS-supervisor and for actively participating to my TAC meeting. I thank **Dr. Adrian Glas** for providing me enough amount of the amino acids for my amide syntheses as well as the Grubbs catalyst for macrocyclization experiments. I thank him for explaining more than one time, for being so patient and pedagogic with all my questions.

Christa Hornemann: Since the beginning of my stay in Dortmund till the present day, I could always count on her goodwill and helpfulness. More than a contact person relating to IMPRS issues, you were a friend and mother for me during my PhD time. I thank you very much for believing in me. Thank you for being so nice and lovely to me. Thank you very much Christa for advocating for me. Thank you very much for pushing me to achieve my goals and for reassuring me when I felt they were out of reach. A big thanks to you for assisting me in the german version of the summary of this work and for being a source of love and energy ever since. Words cannot express how grateful I am to you.

I would like to express my special appreciation and thanks to **Petra Alhorn** and **Martina Reibner** for their help with administrative questions and ordering issues.

The members of the Brunschweiger group have contributed immensely to my personal and professional time at Dortmund. The group has been a source of friendships as well as good advice and collaboration. I am especially grateful to **Mateja Klika**, **Kathrin Jung** and **Hazem Salamon**. You have made the lab an enjoyable place to work and I very much appreciated your enthusiasm, help, friendship and collaboration. Thank you for an enjoyable atmosphere, inside and outside the workplace. A special acknowledgement goes to my office mate **Mateja Klika**. I thank her for her love, respect, trust, support and for our collaborative works. A friend like you is hard to find. One that touches you deep inside. You were a true friend. Thank you Mateja.

I thank my bachelor students **Nadine Kaiser** and **Sven Brandherm** for their respect and consideration. It was a great privilege for me to work with you and be in charge of you. I had a wonderful time getting to know you, and felt privileged to have met you.

Further I would like to acknowledge the financial support by **IMPRS-CMB** (the International Max-Planck-Research School in Chemical and Molecular Biology) that supported my work for my first year. I gratefully acknowledge the **BMBF** (Bundesministerium für Bildung und Forschung) for their financial support for the second and third years.

My time at Dortmund was made enjoyable in large part due to the many friends that became a part of my life: Esther, Bärbel, Gervais, Simone. I am grateful for time spent with them. My time at Dortmund was also enriched by the support of members of my church, who never stop to pray for me: Peter Lutz, Martha Speck, Adi Meier, Hannah Ermakov, Wolfgang Dirks and David Lückhof. I thank Dr. Raymond Mathis, Deborah Benett, Nick Bradshaw, Dr. Esther Nkuipou and Thorsten Doll for reading and correcting parts of this thesis. I thank them for their prayers and precious advices.

My thanks also go to all my dear friends, as well as the other people sharing my pain and happiness in the past, the present and the future: Conrad and Carole Bobale, Landry and Peclare Kamdem, Sinkoni and Rolande Tchana, Larissa, Vulon and Gaelle Meneckdem.

I would like to deeply thank my father, my mother, as well as my father and mother in law. Words cannot describe how important their care, affection, loves, prayers and tireless dedication have been essential for me to reach this stage of my life. Thank you, from the bottom of my heart “mon Papa chéri et ma Maman chérie”, for everything that you have done and still do for me. Your prayer for me was what sustained me thus far. I Love you so much. I give a very special thanks to all my brothers and sister: Gilles, Christian, Landry, Manuela and Kevin, for the essential support and encouragement you gave me all this time. You raised me with a love of science and supported me in all my pursuits.

Last but not least, I would like to express a more than a special thanks to my loving, supportive, encouraging, and patient husband **Philippe Kamdem**. The person with whom I share every moment and detail of my life. Thank you for always supporting me, for listening with attention and affection to all my thoughts, for our dreams and for every celebrated step we make together. Thank you Honey for your love, support and understanding in the time of Ph.D. hardship. Especially the writing period required sacrifices, which have been well appreciated.

I thank you for your faithful support during the last three years. You have always encouraged me. Thank you for being there for me during the years I have needed for my thesis, and taking an interest in what I did. You are a wonderful husband. Without you, I would never be here. I love you.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	10
ABSTRACT	2
ZUSAMMENFASSUNG	6
Chapter I: INTRODUCTION	10
I-1. Background of drug discovery	11
I-2. Drug discovery today	13
I-3. The identification of small organic molecules in drug discovery	14
I-3.1. Technologies for identification of bioactive compounds	15
I-3.2. First practical applications of DNA-encoded combinatorial chemistry	17
I-3.3. Progress in DNA-encoded chemical libraries	22
I-4. Selection strategies	28
I-5. DNA sequencing for library decoding	30
I-6. Bioactive compounds from DNA-encoded small molecule libraries	31
I-7. DNA-routed libraries	35
I-8. DNA-templated libraries	36
I-9. Self-assembled libraries	37
I-10. Privileged scaffolds in medicinal chemistry	46
I-11. Limitations of DNA-encoded libraries	49
I-12. Aims and objectives	49
Chapter II: Synthesis of a DNA-encoded library based on a benzodiazepine scaffold.	58
II-1. Benzodiazepines as bioactive compounds and drugs	59
II-2. Synthesis of trifunctionalized benzodiazepines for library synthesis	60
II-2.1. Synthesis of trifunctionalized 1,4-benzodiazepine-2,5-dione scaffold (1)	61
II-2.2. Synthesis of trifunctionalized 1,4-benzodiazepine-2,5-dione scaffold (2)	63
II-2.3. Synthesis of trifunctionalized tetrahydro-benzodiazepine (3)	67

II-2.4.	Optimization of the amide synthesis for the coupling of scaffolds to 5'-aminolinker modified DNA	72
II-2.5.	Synthesis of a DNA-encoded library based on tetrahydrobenzodiazepine (3)	75
Chapter III:	Development of a modular solid phase synthesis strategy for DNA-macrocyclic conjugates	80
III-1.	Introduction	81
III-2.	Macrocyclic peptides	84
III-3.	Macrocyclization of peptides by ruthenium catalyzed ring-closing metathesis	85
Chapter IV:	Development of an acid catalysed approach to an oligo-thymidine initiated DNA-encoded β-carboline library	90
IV-1.	Introduction	91
IV-2.	β -carboline	92
IV-3.	The Pictet-Spengler reaction	93
IV-4.	Oligo-thymidine initiated DNA-encoded β -carbolines library	99
IV-4.1.	Synthesis of the hexT-5'-PEG-aminolinker conjugate	100
IV-4.2.	Synthesis of hexT-tryptophane conjugates	101
IV-4.3.	Synthesis of hexT- β -carbolines from hexT-tryptophane conjugates by Pictet-Spengler condensation	102
IV-4.4.	Investigation of different catalysts	103
IV-4.5.	Investigation of different solvents ^[a]	105
IV-4.6.	Investigation of different concentrations and reaction times for the synthesis of hexT- β -carboline conjugates	106
IV-4.7.	Control experiments with Fmoc protected alanine	108
IV-4.8.	Control experiments: Pictet-Spengler reaction with tryptophane conjugates of hexa-A, hexa-G, hexa-C, and a mixed sequence	109
IV-4.9.	Introduction of a second set of building blocks	111
IV-4.10.	Projected synthesis of a tiDEL library based on β -carbolines	115
IV-4.11.	Synthesis of a reference compound	116
Chapter V:	SUMMARY AND OUTLOOK	118
Chapter VI:	EXPERIMENTAL PART	124
VI-1.	Materials and instruments	125
VI-2.	Synthesis of DEL9588 DNA-encoded library based on one scaffold	128

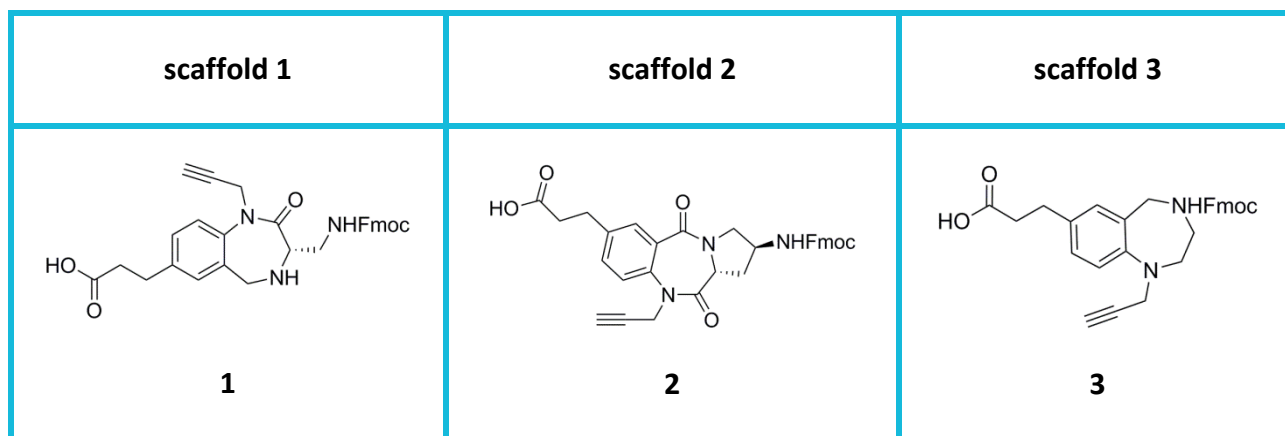
VI-2.1.	Preparation of benzodiazepine scaffolds for library synthesis	128
VI-2.2.	Synthesis of the functionalized benzo-1,4-diazepine-2,5-dione 2	136
VI-2.3.	Synthesis of the tetrahydrobenzodiazepine 3	145
VI-3.	Coupling of carboxylic acid building blocks to DNA scaffold conjugates	152
VI-3.1.	Synthesis of DNA-PEG-linker conjugate 94a	152
VI-3.2.	Synthesis of DNA-scaffold conjugate 95	153
VI-3.3.	Coupling reactions of 114 carboxylic acids to DNA-scaffold conjugate 96	154
VI-4.	Coupling of carboxylic acid building blocks to DNA scaffold conjugates	163
VI-5.	Evaluation of halides for library synthesis	164
VI-6.	Macrocyclization of peptides by ruthenium-catalyzed ring-closing metathesis	174
VI-6.1.	1 st amide coupling: Fmoc-NH-PEG(4)-COOH	175
VI-6.2.	2 nd amide coupling: Fmoc-X ₅ -OH	175
VI-6.3.	3 rd amide coupling: Fmoc-Leu-OH	176
VI-6.4.	4 th amide coupling: Fmoc-Ala-OH	176
VI-6.5.	5 th amide coupling: Fmoc-X _R -OH	177
VI-6.6.	Ring-closing metathesis on CPG by ruthenium catalysis	177
VI-7.	Development of an acid catalyzed approach to an oligo-thymidine initiated DNA-encoded β-carbolines library	178
VI-7.1.	Synthesis of the 5'-amino-PEG(4)-hexathymidine conjugate "hexT"	178
VI-7.2.	Synthesis of hexT conjugate 141	179
VI-8.	Synthesis of hexT-β-carboline conjugates 142A – 142DJ by Pictet-Spengler reaction, and building block validation for tiDEL synthesis	180
VI-8.1.	Optimization of the reaction conditions for the synthesis of hexT- β -carboline conjugates	180
VI-8.2.	Synthesis of hexT- β -carboline conjugates 142A – 142DJ	181
VI-9.	Synthesis of reference molecules and educts	191
APPENDIX		194
LIST OF ABBREVIATIONS		212
LIST OF TABLES		216
LIST OF CHROMATOGRAMS		218
REFERENCES		220

ABSTRACT

The discovery of small organic molecules capable of binding to specific biological targets represents a decisive step towards the development of novel therapies for the treatment of human diseases. However, identifying such small bioactive molecules is often a complex process involving numerous challenges, including low-affinity binding to target proteins. This necessitates the development of effective and powerful strategies to generate large sets of chemical compound libraries and novel screening methodologies to identify binding molecules. High throughput screening (HTS) has proven to be a powerful tool in facilitating the screening of large compound collections against disease-related protein targets. However, HTS has many requirements that must be fulfilled, including: miniaturized and compatible assay formats, an infrastructure with a high degree of automatization, large-scale data analysis, and trained staff. Owing to these significant technical and financial requirements, new technologies to facilitate ligand discovery that can be implemented in the academic sphere are highly desirable. In this regard, technologies based on DNA encoded chemical libraries (DELs) offer a considerable advantage over conventional HTS technologies as they are rapid, cost-effective, and employ user-friendly screening logistics. Two points in particular make the screening of DNA-encoded libraries attractive. Firstly, large numbers of compounds can be pooled and screened, so that a dedicated infrastructure for handling of compound libraries is unnecessary. Moreover, only one assay format is required for all protein targets facilitating assay development. Hits emerging from the selection assay must be validated in additional assays and may serve as lead compounds for further development towards probes or even therapeutic agents.

In the first section of this thesis, the general synthesis of three different functionalized benzodiazepines (compound **1**, **2**, **3**) known as “**privileged scaffolds**” is established. Privileged scaffolds are interesting structures that are overrepresented among bioactive small molecules and therefore constitute an attractive framework for probe and drug discovery. With this in mind, the DEL methodology was applied to construct a DNA-encoded library based on one of these benzodiazepines (compound **3**). This yielded a DNA-encoded chemical library of 9588 members (DEL9588) relying on introduction of building blocks by amide bond forming reaction and Cu (I)-catalyzed azide-alkyne cycloaddition. Although the DEL-approach increases chances to find specific binders to relevant proteins, it remains limited due to the incompatibility of DNA with many chemical transformations.

ABSTRACT



The third and final section of the thesis proposes a strategy to overcome this limitation by establishing new protocols of DNA-compatible chemistries that are robust, working in the presence of strong catalysts without degradation of the DNA. For this purpose, the “oligoThymidine initiated DNA-Encoded Chemistry” (TiDEC) strategy for DEL synthesis has been developed and applied for the successful Pictet-Spengler cyclization of tryptophan and aldehydes to form drug-like heterocyclic structures (β -Carbolines) at ambient temperature under acidic conditions using a hexT DNA. This resulted in the formation of 114 β -carboline conjugates that were subsequently encoded by DNA ligation. Besides, methodologies for the construction of a DNA-encoded chemical library based on those β -carboline were investigated, featuring the stepwise addition of two independent sets of chemical moieties onto the initial β -carboline scaffold.

Amide synthesis represents an important reaction for the generation of DNA-encoded combinatorial libraries. The investigation process leading to suitable reaction conditions for amide conjugation on solid phase on DNA was a lengthy optimization effort. Fortunately, it revealed the Fmoc chemistry to be compatible with the DNA on solid support, preserving the DNA integrity. Encouraged by this observation, a modular solid phase synthesis strategy for DNA-macrocycle conjugates was developed. The second section of this thesis presents the successful macrocyclization of a peptide composed of four amino acids by ruthenium-catalyzed ring-closing metathesis (RCM) on DNA on solid phase making use of the first generation Grubbs catalyst **39**. This experiment has proved on one hand, synthesis on solid-phase to be particularly attractive, as it allows to drive chemical reactions to completion and to expand the scope of DNA-compatible chemical reactions to water-free conditions.

ABSTRACT

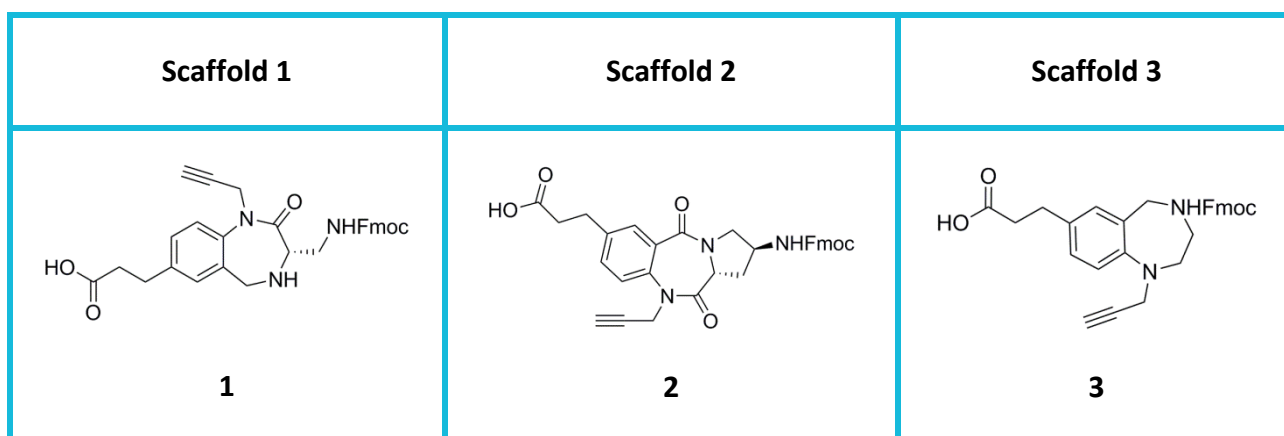
On the other hand, the macrocyclization of peptides by RCM is tolerated by a hexT adapter oligonucleotide designed for this purpose, opening the door to the synthesis of a DNA-encoded chemical library based on macrocycles that can be screened against targets involved in protein-protein interactions. Overall, the experiments performed show that an adapter oligonucleotide consisting of thymine nucleobases can enable the use of various catalysts utilized by medicinal chemists to initiate DNA-encoded library synthesis based on heterocycles, as well as on macrocycles.

ZUSAMMENFASSUNG

ZUSAMMENFASSUNG

Die Identifizierung und Isolierung neuer bioaktiver Verbindungen, die mit medizinisch relevanten Proteinen in Wechselwirkung treten können, ist eine Herausforderung nicht nur in der pharmazeutischen Industrie, sondern auch in der akademischen Forschung. Das Verfahren, kleine bioaktive Moleküle zu identifizieren, ist oft komplex. Neue Wirkstoffe werden oft in Screening-Verfahren mit dem Hochdurchsatz-Screening (HTS) identifiziert. Diese gut etablierte Strategie ist allerdings sehr zeitaufwändig und kostspielig. Aus diesem Grund besteht ein zunehmender Bedarf für die Entwicklung neuer Methoden, die in der Forschung, insbesondere der universitären, umgesetzt werden können, um die Entdeckung neuer Liganden zu erleichtern und zu beschleunigen. In diesem Zusammenhang erweisen sich besonders DNA-kodierte chemische Substanzbibliotheken (DELS, *engl.* DNA Encoded Libraries) als eine vielversprechende Methode, um sehr große Substanzbibliotheken simultan auf Bindungsaffinitäten gegen ein Zielprotein der Wahl zu testen. DNA-kodierte chemische Bibliotheken bestehen aus einer Vielzahl verschiedener organischer Verbindungen, die individuell an spezifische DNA-Sequenzen gekoppelt sind. Dies ermöglicht eine Bibliothek als Substanzgemisch in Selektionsexperimenten mit einem immobilisierten Protein in Kontakt zu bringen und Proteinbinder durch Sequenzierung ihres spezifischen DNA-Codes zu identifizieren. Das Verfahren ist auch deshalb attraktiv, da nur ein Assayformat benötigt wird.

Im ersten Teil der vorliegenden Dissertation wird die Synthese von drei mit Carbonsäure, geschütztem Amin und terminalem Alkin funktionalisierten Benzodiazepinen dargestellt (Verbindung **1**, **2**, **3**). Benzodiazepine werden in der medizinischen Chemie auch als "privilegierte Gerüste" bezeichnet.



Privilegierte Gerüste sind unter kleinen bioaktiven Molekülen überrepräsentiert und gelten daher als attraktive Verbindungen in der Wirkstoffforschung. Daher wurde eine DNA-kodierte Bibliothek basierend auf einem dieser Benzodiazepine (Verbindung **3**) dargestellt. Dies ergab eine DNA-kodierte chemische Bibliothek von 9588 Molekülen (DEL9588) beruhend auf Amidsynthese und Click-Chemie für die Einführung von Substituenten.

Obwohl die Synthese von DELs ein validiertes Verfahren darstellt, die von einer Vielzahl von akademischen und industriellen Arbeitsgruppen als Quelle für bioaktive Moleküle genutzt wird, können viel Synthesemethoden aufgrund der Inkompatibilität mit DNA nicht eingesetzt werden. Die DNA besteht aus vier Nucleobasen: die Purine Adenin (**A**) und Guanin (**G**) sowie die Pyrimidine Thymin (**T**) und Cytosin (**C**). Verschiedene Reaktionsbedingungen können die DNA-Integrität beeinflussen. Starke Oxidationsmittel können zum Beispiel den Imidazolring von Guanin oxidieren. Saure pH-Werte führen zur Depurinierung. Entsprechend katalysierte Reaktionen, werden in der Literatur als inkompatibel mit DNA-Chemie beschrieben. Die sogenannte *oligoThymidine DNA-kodierte Chemie* (TiDEC)-Strategie, basiert auf der Feststellung, dass die Inkompatibilität von vielen Katalysatorsystemen mit DNA auf der Reaktivität der Purinbasen beruht. Daher werden zur Initiierung des Aufbaus unserer DELs kurze DNA-Stränge, die nur aus Pyrimidinbasen bestehen, die wir Initiator-DNAs nennen, verwendet. Durch die Verwendung kurzer 5'-Amino-C6-Hexathymidin-Einzelstränge (hexT) ist es möglich, bekannte Arzneistoffe und Derivate in Form einer DEL zu synthetisieren.

Fmoc-Peptidchemie ist kompatibel mit an Festphase gebundenem hexT. Daher wurde eine modulare Festphasensynthese-Strategie für die Synthese eines DNA-Makrozyklus-Konjugats an der DNA entwickelt. Der zweite Teil dieser Arbeit stellt die erfolgreiche Makrozyklisierung eines Peptids dar, das aus vier Aminosäuren synthetisiert wurde, und durch Ruthenium-katalysierte Ringschlussmetathese (RCM) zyklisiert wurde. Dabei wurde der Grubbs (I) Katalysator **39** verwendet. Dieses Experiment hat gezeigt, dass die Makrozyklisierung von Peptiden durch RCM auf einem für diesen Zweck konzipierten hexT-Adapter-Oligonukleotid toleriert wird. Dieses Ergebnis kann die Tür zur Synthese einer DNA-kodierten chemischen Bibliothek basierend auf Makrozyklen öffnen, die zum Beispiel gegen Ziele getestet werden können, die in Protein-Protein-Wechselwirkungen involviert sind.

ZUSAMMENFASSUNG

Der dritte Teil dieser Arbeit schlägt eine weitere Strategie vor, um die Stabilität des hexT auszunutzen: Die TiDEC-Strategie wurde für die erfolgreiche Pictet-Spengler-Zyklisierung von Tryptamin mit verschiedenen Aldehyden unter Wasserabspaltung zur Herstellung von β -Carbolinen verwendet. Die Reaktionen wurden unter Säure-katalyse durchgeführt. Dies führte zur Bildung von 114 β -Carboline-Konjugaten, die anschließend durch DNA-Ligation kodiert wurden.

Letztendlich zeigen diese Experimente, inwiefern ein Adapter-Oligonukleotid bestehend aus Thymin-Nukleobasen die Verwendung von verschiedenen Katalysatoren erlaubt, die in der organischen Chemie häufig verwendet werden. Diese Erkenntnis ermöglicht auf der einen Seite die Synthese von DNA-kodierten Bibliotheken basierend auf Heterozyklen sowie Makrozyklen. Auf der anderen Seite erweitert sie den Katalog von DNA-kompatiblen chemischen Reaktionen, die der Konstruktion von großen, vielfältigen Molekülbibliotheken dienen.

Chapter I: INTRODUCTION

The identification of bioactive small molecules capable of binding to a target protein is one focus of modern drug discovery.¹⁻² Technologies like sequencing of the human genome and advances in proteomic as well as transcriptomic analysis have furthered understanding of living systems and have enabled the identification of various new biological targets associated with many human diseases.³ However, the synthesis of small molecules capable of addressing such target proteins still remains a challenge. Conventional high-throughput screening (HTS) campaigns are a standard technology in drug development but closed to academic research as they require dedicated infrastructure.⁴⁻⁵ There is consequently an urgent need in the field of medicinal chemistry to develop more efficient drug discovery techniques that accelerate drug identification. DNA-encoded combinatorial chemistry shows great promise as drug discovery technology. It addresses the limitations of the conventional high-throughput methodology for hit discovery.⁶ The concept of antibody phage display technology, which involves linking antibody binding properties (“phenotype”) with the genetic information coding for the antibody (“genotype”), has been translated to DNA-encoded chemical libraries based on the collections of organic molecules covalently linked to a unique DNA tag serving as an amplifiable identification bar code.⁷ DNA-encoding accelerates the drug discovery process, since it allows the rapid and inexpensive in vitro selection of ligands by affinity capture at sub-picomolar concentrations of any target protein of interest. The fact that this technology does not require screening infrastructure is also a great advantage.⁸

I-1. Background of drug discovery

Drug discovery and development has a long history. It is the process by which new drug candidates are synthesized and spans the fields of medicine, biotechnology, pharmacology and medicinal chemistry.⁹ In the past, new candidate medications were often discovered by chance.¹⁰ They were isolated by identifying the active ingredient from traditional remedies or a combination of trial and error experimentation, followed by observation of human and animal reactions to ingestion of the drug candidates.¹¹ Later, collections of small chemical compounds, natural products or extracts were used in screening assays with cells or organisms in order to identify substances or moieties possessing a desirable therapeutic effect.

Over time, drug discovery developed still further, becoming a capital-intensive process that necessitates large investments by the pharmaceutical industry, as well as national governments providing grants and loan guarantees. However, despite remarkable advances in technology and an improved understanding of biological systems, drug discovery remains a cumbersome, lengthy, and extremely costly process with a low rate of new therapeutic discovery (Figure I-1). In fact, the process takes approximately 15 years to complete, and the current costs of developing a new drug are estimated to be around a billion dollars.¹² The identification of biologically active compounds to provide an insight into patho-physiological processes and novel drugs against diseases is both an aim in academia and an important priority in many pharma companies.² However, in the last decade, there have been an increasing number of discussions about a looming crisis in pharma companies due to the so-called patent cliff.¹³ To date, there are still many diseases for which medications are currently not available. These include neurodegenerative diseases such as Alzheimer’s or Parkinson’s disease, various malignant diseases and the resurgence of infectious diseases due to the development of resistance to available therapeutics. On the other hand, decoding of the human genome has revealed several proteins that are involved in pathophysiological processes and are therefore of great interest for pharmacology.¹⁴ However, several promising targets, such as protein-protein interactions in the ubiquitin-proteasome system and epigenetic proteins are challenging to drug discovery.¹⁴

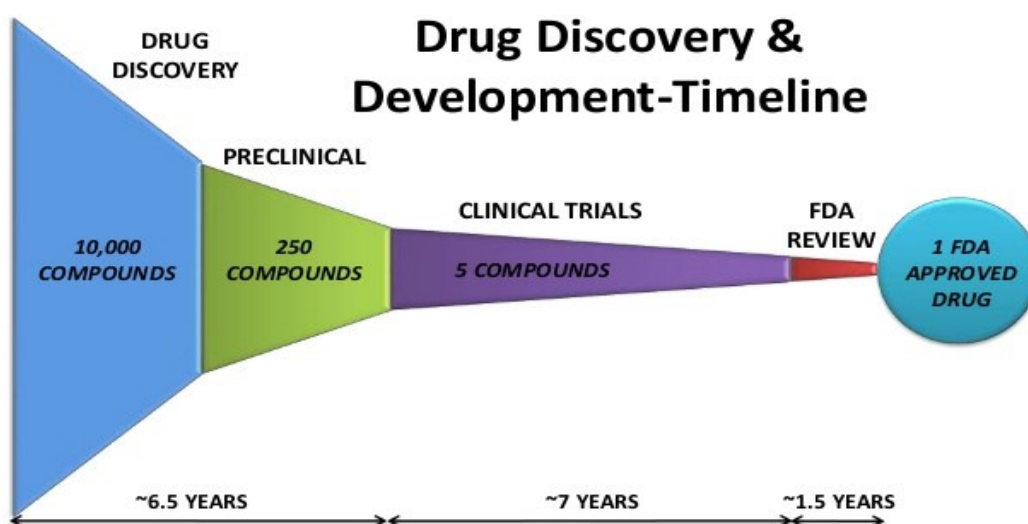


Figure I-1: Overview of drug discovery research and development timeline of the screening of a 10,000 compounds library. The process leading to a drug or therapy is lengthy and costly. The figure was taken from rahul_pharma/drug-discovery-and-development homepage.¹⁵

I-2. Drug discovery today

The identification of bioactive small molecules against medically relevant proteins constitutes a challenge in drug discovery (Figure I-2).¹ The screening of large compound collections containing up to 10^6 different molecules using automated HTS has proven to be very effective in the discovery of new bioactive compounds, and therefore appears to represent a good basis in industry for hit-to-lead development.⁴ However, traditional hit development programs by HTS are very demanding in terms of time, suitable logistics and finance.^{1, 5, 11} Accordingly, new innovative, cost-effective and broadly applicable concepts to screen small molecule libraries are urgently needed in academia.⁷

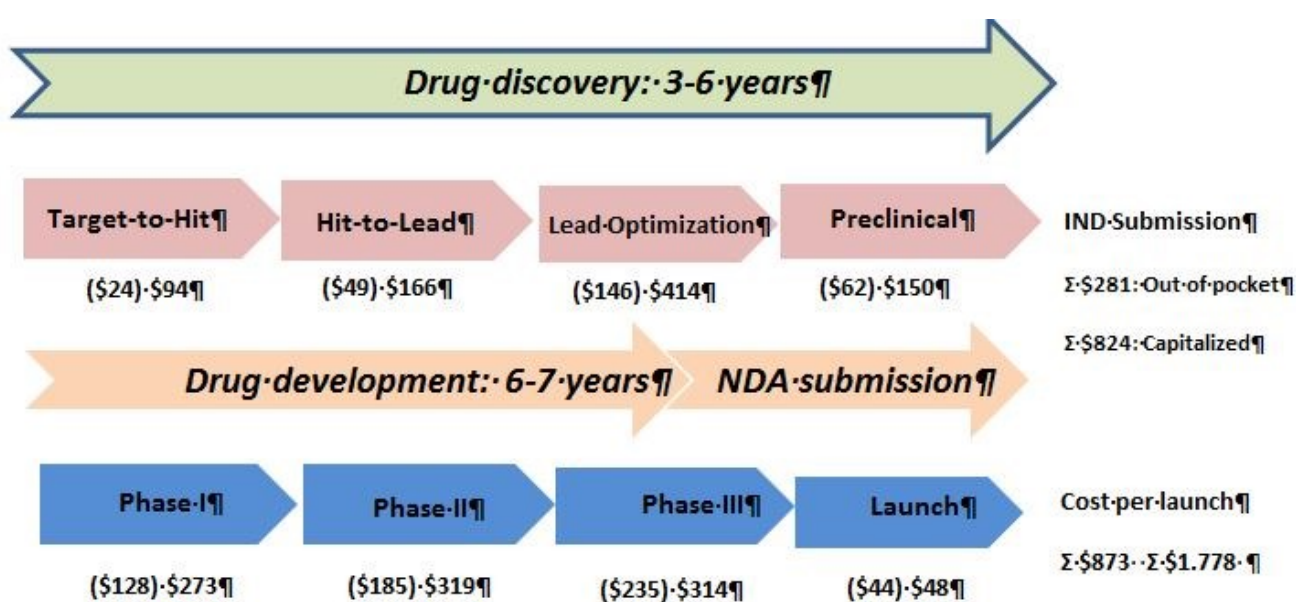


Figure I-2: The discovery of new ligand methodologies may aid in considerably reducing the cost of the early drug discovery process. It is important to note that the risks of costly late-stage failures in preclinical or clinical trials can be decreased by good-quality drug candidates. IND: Investigational New Drug. NDA: New Drug Application Capitalized costs include interest expenses until drug launch for financing activities.¹⁶ The costs during drug development and discovery process was adapted from Paul et al., 2010.¹⁶

In the last decade, there has been slow progress regarding the number of novel drugs approved by the US Food and Drug Administration (FDA): also called new molecular entities (NMEs).¹⁷⁻¹⁸ This is the case despite the fact that worldwide R&D investments rose strongly during the same period.¹⁷⁻¹⁸ This outcome can be explained by the fact that the R&D efficiency (R&D costs as a percentage of net sales) of pharmaceutical industries declined constantly until the 2000s.

This decline in R&D performance is mainly attributed to the tortuous road (up to 15 years) and the associated costs leading to the identification of novel compounds following FDA stringent drug approval processes. Given the current challenges, it is therefore of pivotal importance to find alternative measures to improve R&D efficiency and subsequently strengthen the pharmaceutical industry sector.^{14, 17-18}

The scientific progress in medicinal chemistry leading to the development of new and approved biological entities plays a significant role in the acceleration of further drug discoveries in the pharmaceutical industry.^{11, 17, 19} The example of antibody phage display emphasizes this point, as it was made possible using a selection-based strategy to generate monoclonal antibodies against any target protein of choice within a short period of time.²⁰⁻²⁴ In this context, large collections of phages exhibiting single-chain variable fragments (scFv) as fusion complexes on their surface could be simultaneously screened against a target protein of interest in affinity selection experiments. This led to the identification of several compounds, such as adalimumab (Humira®); a fully human monoclonal antibody against TNF- α that has further contributed to the management of breast cancer. The cost of this project reached the \$3 billion annual sales mark in the year 2007.²⁵⁻²⁶ The technology also possesses several analogies to DNA-encoded combinatorial chemistry, which is useful for the rapid generation and interrogation of massive collections of chemical compounds to tackle new target classes. DNA-encoding of small molecules therefore allows selection-based interrogation of small-molecule libraries against a target protein of interest.

I-3. The identification of small organic molecules in drug discovery

An important challenge in medicinal chemistry remains the development or identification of small molecules capable of specifically binding to proteins with pharmaceutical properties. Sequencing the human genome as well as the genomes of multiple pathogens, yielded hundreds to thousands of potentially novel biological targets. Functional genomic approaches were introduced into the drug discovery area, comprising RNA profiling, proteomics, antisense and RNA interference, model organisms and high-throughput, genome-wide over-expression or knockdowns in the identification of new drug targets.²⁷

Thanks to an improved understanding of the mechanisms of many diseases at a molecular level, scientists are now able to associate different proteins with the pathogenesis of diseases. However, it is challenging to use this rich source of information in a manner that can assist in the development of novel, more effective and safer drugs. Moreover, new ligands need to be isolated in order to analyze molecular pathways and study protein function.¹ Technologies that enable new ligand development are sorely needed, as they contribute to a better understanding of complex biological processes and speed up drug discovery. Straightforward development and isolation of specific binders to target proteins can be achieved with the selection of large polypeptide collections (e.g. antibody libraries).^{21-22, 24} However, these reagents are not able to interfere with intracellular biochemical processes.²⁴

I-3.1. Technologies for identification of bioactive compounds

From a historical point of view, drug discovery has been divided into three periods; (i) before, (ii) during and (iii) after the twentieth century.¹⁰ Almost all drugs discovered before the twentieth century were found by chance.¹¹ However, as a result of notable advances in the different disciplines involved, drug discovery quickly became a more rational process. Among the noteworthy techniques made available to bring this process forward by the end of the twentieth century was the development of molecular structures of drugs using different methodologies such as molecular modeling, combinatorial chemistry, high-throughput screening and advanced molecular biology methods.¹⁰⁻¹¹

I-3.1.1. High-throughput screening

Traditional technologies used in the identification of small molecule ligands against a target protein of choice use a collection of chemical compounds or compound libraries, which are screened one molecule at a time. The principle of the assay here is based on enzymatic activity of the target protein. The high-throughput screening (HTS) approach enables the screening of between 10,000 to 1,000,000 compounds. This greatly improves the chances of discovering novel and efficient drugs. However, the screening is cumbersome and expensive in terms of library synthesis, logistics, screening and management.²⁸

Moreover, the HTS of compound libraries does not always yield hits. In summary, although the advantages of the technology are apparent, its use in the academic environment is restricted due to its high costs. A frequently used strategy to circumvent this problem is the use of structural information in “virtual screening” with the aim of reducing the number of compounds that need to be evaluated.²⁸⁻³⁰

I-3.1.2. Fragment-based approach

Another routinely used alternative for the identification of ligands with convenient binding affinity and pharmacological features is the fragment-based approach.^{29, 31-32} Fragment-based approaches analyze small molecules with molecular masses of < 300 Da using NMR or X-ray crystallography, and have proved to be an efficient and complementary method for early stage drug development in recent years.²⁹ Indeed, this approach differs from HTS in almost every respect: library size, screening method, as well as the dependence on structural methods.³² Fragment-based drug discovery relies on the close interaction of structural biology and synthetic chemistry. Taking a structural and molecular view of biological targets, combined with a stepwise and logical medicinal chemistry strategy, this technology represents an attractive tool not only to chemical biologists, but also to synthetic and computational chemists.³¹

I-3.1.3. Display technologies

Among the various biological libraries that have been developed, most attention has been directed at those which enable the identification of bioactive proteins or peptides, especially as proteins are natural modulators of phenotype and therefore represent obvious drug candidates. Display technologies as combinatorial biology techniques are likely to play an increasingly more important role in the future of drug discovery and include phage display technology and ribosome display technology. Phage display technology, involving the isolation of binding polypeptides from large libraries, has also demonstrated its value in the development of new bioactive molecules (Figure I-3).³³ Millions to billions of polypeptides displayed on the surface of a filamentous phage make use of the physical linkage between their potential binding properties (“Phenotype”) and the corresponding genetic information enclosed in the phage (“Genotype”).

Libraries can be panned against target antigens that are immobilized on a solid support. Through several washings steps, non-binding molecules will be eliminated whereas the remaining binding molecules will be captured.³⁴⁻³⁵

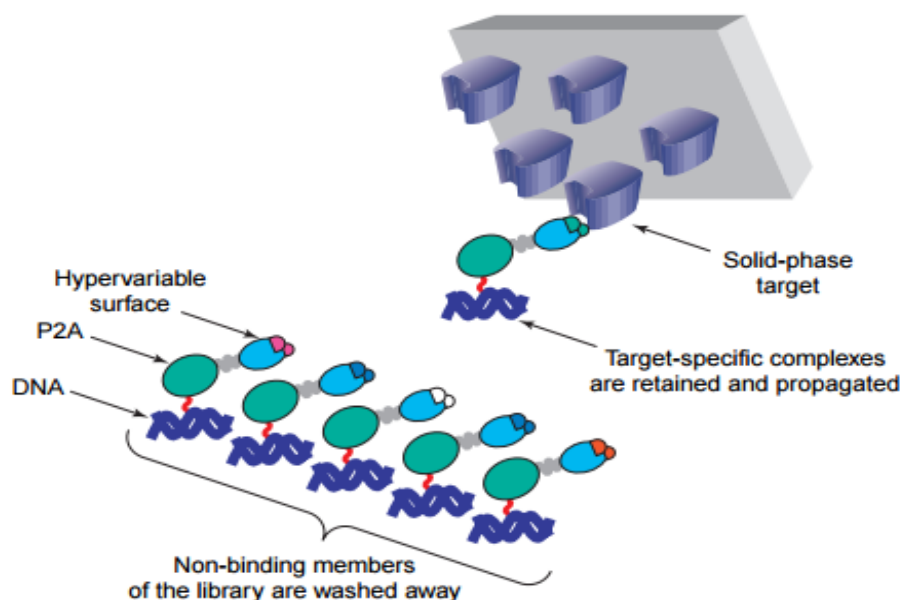


Figure I-3: The display technology exploits the properties of a replication initiator protein from the *E. coli* bacteriophage P2. The protein is the product of the viral A-gene (P2A) and is an endonuclease that initiates a rolling circle replication process by binding to the viral origin (*ori*) and introducing a single strand discontinuity (nick) in the DNA. Affinity selection consists of the incubation of pools of covalent DNA–protein complexes with an immobilized target molecule. Binding moieties are retained on the solid surface while non-specific complexes are removed by washing steps. The physical link between the gene and the gene product results in the coincident retention of both target-specific polypeptide and its encoding DNA. The retained DNA is then amplified by PCR for further rounds of selection or used to generate a clonal archive of the binding domain’s gene. The figure was taken from Mattheakis et al., 1994.³⁵

I-3.2. First practical applications of DNA-encoded combinatorial chemistry

The emergence of combinatorial chemistry, particularly the “split-and-pool concept” has opened the door to the synthesis of compound collections with unprecedented size (Figure I-5). This concept operates as follows: (i) First a set of chemical compounds comprising a members is pooled in equimolar amounts and split into b aliquots. (ii) Then, a specific reactant of a second set of chemical compounds containing b members is added to each aliquot yielding a reaction products. (iii) Finally, aliquots are pooled again resulting in a collection of $a \times b$ different reaction products.

This synthetic route can be repeated several times, generating polymeric compounds such as polypeptides (Figure I-4). Despite the great advantages offered by the split-and-pool methodology, the identification of bioactive small molecules from these compound collections remains problematic since individual members fail to retain their spatial information during split-and-pool synthesis. Brenner and Lerner were the first to suggest the usage of DNA as a tag to provide the information that would be lost during split-and-pool synthesis.³⁶ In their theoretical article published in 1992, Brenner and Lerner described how to adapt the concept of linking phenotype and genotype to organic chemistry. In comparison to polypeptide display technologies, they proposed the synthesis of amino acid sequences (“phenotype”) and encoded oligonucleotides on the same bead (Figure I-4). In this approach, synthetic chemical moieties and encoded DNA are physically connected and the encoded chemical libraries can be screened in the same manner as the affinity selections utilized in phage display technology. PCR priming sites are attached to the appropriate DNA sequences, which in turn play the role of bar codes. This facilitates the decoding of chemical structures by PCR amplification and DNA sequencing. To demonstrate that the library construction strategy does not depend on biological systems, an enlarged catalog of chemical building blocks can be used. The practical application of DNA-encoding strategies necessitates the compatibility of unprotected DNA with the covalent synthesis of the chemical moiety. Brenner and coworkers thus demonstrated the possibility of synthesizing DNA-encoded chemical libraries based on peptides, which is helpful for further investigation in this area. During the 1990s, diverse alternative encoding strategies were developed with the aim of expanding the field of compound synthesis methodologies beyond DEL-compatible chemistry. However, the selective amplification of DNA by PCR was found to be an indispensable step for hit detection and statistical analysis of DELs. Moreover, DNA molecules display additional benefits for encoded chemical libraries, including (i) commercially available oligonucleotides that can be designed by each scientist; (ii) the straightforwardness and efficiency of DNA purification methods, such as ethanol precipitation; (iii) the accessibility of DNA-processing enzymes, and (iv) the availability of high-throughput DNA sequencing technologies (e.g., Roche 454, Illumina Solexa)³⁷ that permit the extraction of DNA sequence information in a parallel fashion. However, the identification of DNA-compatible reaction conditions remains crucial to access new chemistry for library construction.

INTRODUCTION

The first implementation of DNA-encoded combinatorial chemistry was performed in 1993 by Brenner and Janda in an experiment where five alternating synthesis cycles of amide-bond were performed to form reactions and nucleoside phosphoramidite chemistry. This demonstrated the potential to orthogonally synthesize pentapeptides and their corresponding oligonucleotide tags on the same bead (Figure I-4).³⁸

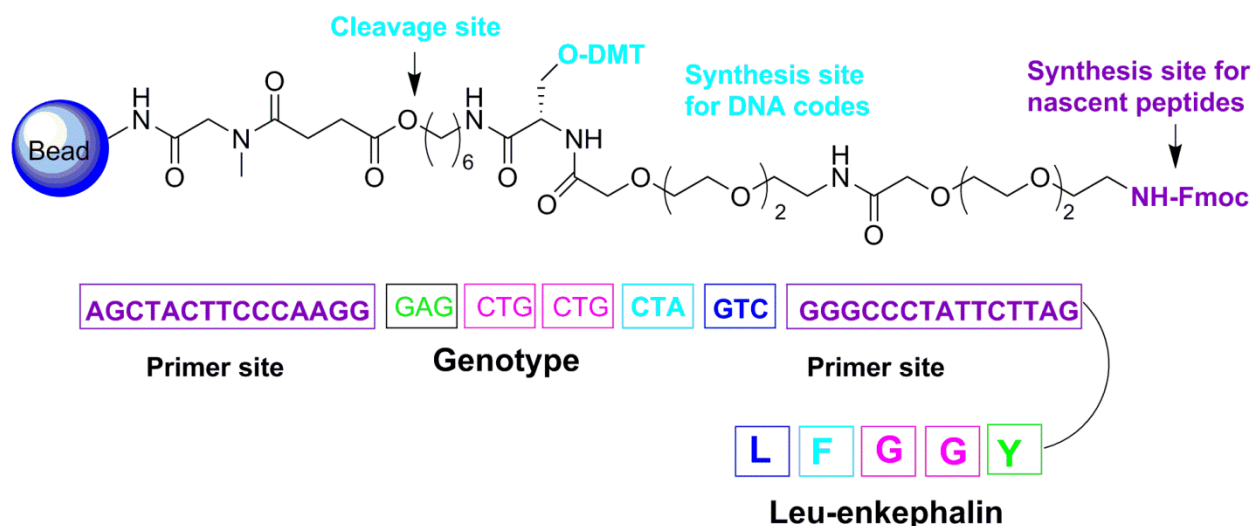


Figure I-4: A chemical scaffold with orthogonal protecting groups (O-DMT and N-Fmoc) allows the bidirectional synthesis of oligonucleotide and peptide sequences. A cleavage site allows the release of the DNA-encoded peptides from the solid support. An example of oligonucleotide-peptide conjugate after bead-release is depicted at the bottom. This picture was adapted from Leimbacher et al., 2012.³⁸

The selection of DNA-encoded Leu-enkephalin pentapeptide (LFGGY) in this manner, as the result of synthesis against the antibody 3-E7, displayed strong binding when compared with the reference peptide ($K_d = 24$ nM), whereas the bivalent Leu-enkephalin displayed a $K_d = 7$ nM.³⁹ Hence, not only were they able to use DNA-encoded combinatorial chemistry, but they were also able to show that a greater binding at the target protein was achieved. Brenner and coworkers also showed that bulky DNA tags did not interact with target binding and corresponding DNA codes were amplifiable by PCR.³⁹ This finding demonstrated the efficiency of this technology in terms of the screening of libraries against target proteins. Another independent group, Gallop and co-workers, experimented on the orthogonal synthesis of a library containing 823,543 (7^7) heptapeptides. These were bound to corresponding DNA tags, on the same bead.³⁹ Coding sequences were enclosed by PCR priming sites for DNA amplification.

An affinity selection was then performed and beads bearing the amino acid sequence RQFKVVT were isolated based on their aptitude to bind the fluorescent monoclonal antibody D32.39, known to bind the reference peptide RQFKVVT with a $K_d = 0.5$ nM.³⁹ FACS-based isolation and PCR amplification revealed the sequence of the related DNA code. In 2004, Harbury and co-workers showed that the field of chemical reactions can be extended to completely water-free conditions using diethylaminoethyl (DEAE) resin for the immobilization of DNA during the synthesis of displayed chemical moieties.⁴⁰

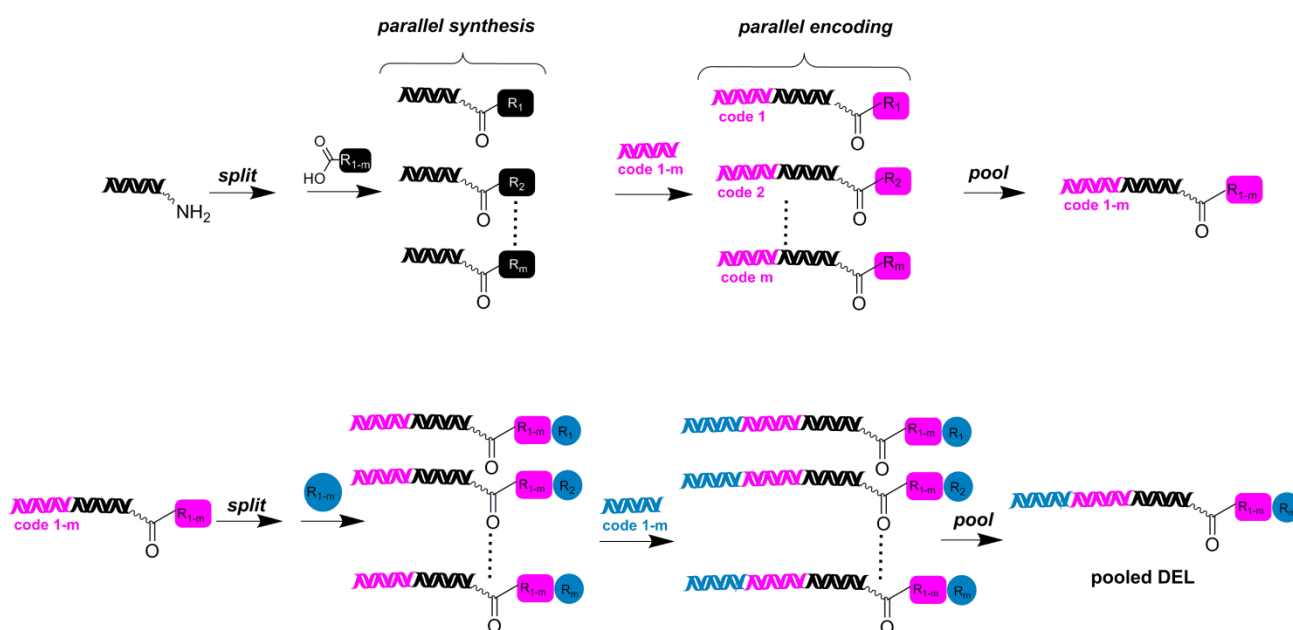


Figure I-5: Schematic representation of a DNA-encoded library containing two building blocks constructed in a split-&-pool fashion. The initial building block is conjugated to the oligonucleotide by amide synthesis and encoded with a further set of oligonucleotide by DNA ligation. Afterwards the oligonucleotide conjugate compounds are mixed, splitted into different reaction vessels and reacted again with an additional building block. Following encoding, these steps are repeated a given number of times and the final library is pooled together.

As written above, the principle of linking “Genotype” and “Phenotype” for selection assays can be transposed into the area of small organic molecules by encoding them with unique DNA fragments that are used as amplifiable identification bar codes (Figure I-6).³⁶ This concept would be useful for the identification of small organic molecules capable of binding to target proteins with high affinity and specificity, based on the association of individual chemical compounds to unique DNA-fragments serving as identification bar-codes.

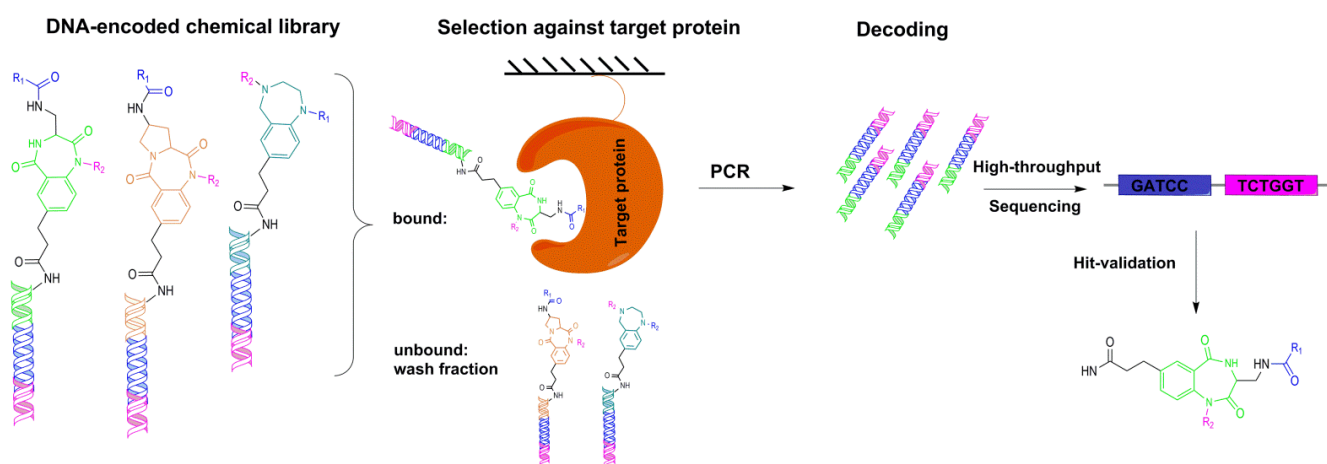


Figure I-6: A typical selection process with DNA-encoded chemical libraries comprises the following steps: (i) incubation of the library with a target protein of interest immobilized on a solid support, (ii) removal of non-binders by washing steps of the solid support, (iii) PCR amplification of the DNA codes corresponding to the binders, (iv) assessment of the relative quantification of library codes after selection using high-throughput sequencing, and (v) re-synthesis of identified compounds in the absence of a DNA-tag for biological or biochemical assays.

Like polypeptide-based display technologies, DNA-encoded small molecules can be panned against a target protein of choice. Non-binders are then removed by washing cycles and sequences of binders can subsequently be enriched and then amplified by PCR (Figure I-6).³⁶ The concept of DNA-encoded chemical libraries was firstly described in 1992 by Brenner and Lerner, proposing the alternated synthesis of peptides and oligonucleotides on the same solid support (Figure I-7C).³⁶ The authors planned to use this concept for the synthesis of peptides and oligonucleotides on the same bead. However, their concept could not be put into practice for library construction, presumably due to limitations in library size (number of beads is limiting) and the requirement for an orthogonal synthesis of peptides and oligonucleotides on the same solid support.³⁶ Since that time, a number of different approaches have been developed. In addition to the direct linkage of small molecules to an encoding DNA tag (Figure I-7E), PNA-encoded libraries have been introduced and are now used by many groups (Figure I-7D).⁴¹⁻⁴² The use of different DEL construction methodologies has extended the discovery of novel ligands and inhibitors against a variety of biomedically important targets.⁴³

I-3.3. Progress in DNA-encoded chemical libraries

The DNA encoding technology has become a formidable strategy for the synthesis of large combinatorial libraries, since the DNA tag itself allows library construction in different ways (Figure I-7).^{41, 44-48} The simplest way to obtain a DNA-encoded chemical library involves the attachment of an individual organic entity to a unique DNA strand. It is obtained when a single organic compound is coupled to each DNA strand (Figure I-7A-D).^{41, 44-48} In contrast, the so-called “*dual pharmacophore libraries*” are characterized by the presence of pairs of chemical moieties at the two extremities of complementary DNA fragments (Figure I-7E).⁴⁹ A classification is then achieved based on library synthesis categorizing single pharmacophore libraries which includes: i) libraries in which building blocks are brought together by a split-and-pool strategy in a combinatorial fashion as well as encoding DNA strands (Figure I-7A, B).^{41, 44-45}; and ii) libraries for which DNA-templated synthesis is applied in the construction step (Figure I-7C, D).^{46, 48, 50}

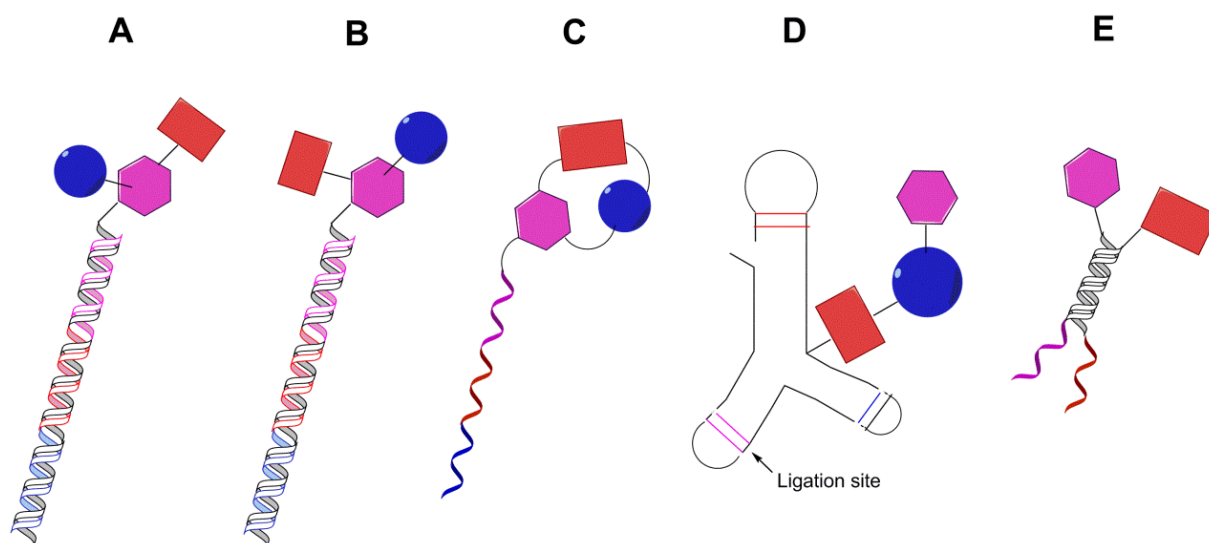


Figure I-7: “Single pharmacophore DNA-encoded chemical libraries” display a single chemical compound on one DNA-strand, in comparison to “dual pharmacophore DNA-encoded chemical libraries”. E illustrates pairs of chemical structures at the two extremities of complementary DNA strands. For the construction of single pharmacophore libraries, different chemical synthesis strategies can be considered including: A the simple attachment of small molecules to one DNA strand (e.g. to a 5'-aminomodified oligonucleotide), using a linker which connects both DNA strands B or The use of DNA templated synthesis C, D to build macrocyclic structures C or linear structures with the help of a triple hairpin structure D. The figure was adapted from Buller et al., 2009.³⁷

The split-and-pool strategy has proven to be a powerful tool for the synthesis of DNA-encoded chemical libraries^{44-45, 47} and is an alternative strategy for constructing a DNA-encoded library in analogy with the encoded “split-&-pool” technique described by Brenner and Lenner.³⁶ The principle is based on the synthesis of chemical compounds directly on the oligonucleotide, but without the use of beads. Figure I-8 describes a suitable synthetic route comprising the following: (i) a set of unique oligonucleotides, each bearing a specific sequence chemically attached to a corresponding set of small organic molecules. These represent the first set of building blocks that carry a suitable reactive group. (ii) These oligonucleotide conjugates pooled together, mixed and split into a number of groups in separate reaction vessels. (iii) A second set of building blocks is conjugated to the first one under suitable chemical conditions, (iv) the second set of building blocks is then encoded by using a further oligonucleotide, which codes the second moiety. As a result, they are hybridized to the initial oligonucleotide and enzymatically encoded either by DNA ligation or Klenow fragment-mediated DNA polymerization. (v) The encoded reactions are pooled to yield a DNA-encoded chemical library. These steps can be reiterated in a “split-&-pool” fashion for the construction of libraries containing three or more sets of building blocks as well as three or more coding regions. DNA-encoded chemical libraries are efficient in the de novo identification of binders against a target protein of interest. Frequently, lead compounds yielded from these selection assay experiments and decoding are remodeled with a view to improve their performance for drug discovery applications.⁵¹ Single pharmacophore DNA-encoded chemical libraries can be effectively used to operate “affinity maturation” of lead structures, including the lead structure during DNA-encoded library synthesis.⁵⁰⁻⁵¹ In this manner, Neri and coworkers recently reported the synthesis of a benzamidine-based single pharmacophore library which enabled the affinity maturation of benzamidine towards serine proteases and the discovery of nanomolar binders against trypsin as a model target.⁴⁹

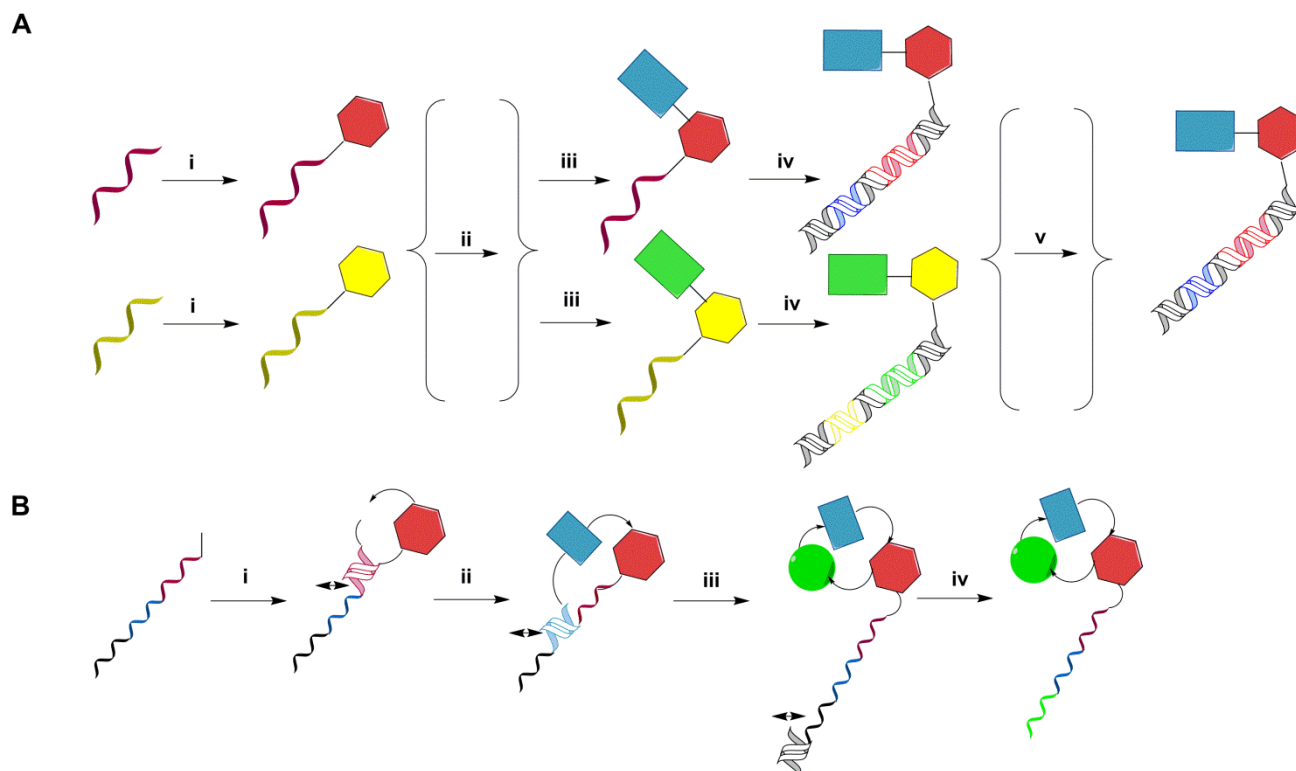


Figure I-8:.(A) Split-and-pool synthesis of DNA-encoded and templated chemical libraries relies on the alternated synthesis and encoding steps: (i) conjugation of a first set of chemical compounds representing building block 1 to distinct amino-modified synthetic oligonucleotides; (ii) pooling of the reactions, followed by a splitting step into separate reaction vessels corresponding to the reactions for the incorporation of a second building block; (iii) chemical coupling of the second building block; (iv) encoding of the second building block (e.g., by hybridization of partially complementary oligonucleotides followed by Klenow fragment-mediated DNA polymerization); (v) pooling of the encoded reactions to yield a DNA-encoded chemical library. (B) A single pharmacophore DNA-encoded chemical library can be synthesized using DNA-template synthesis, in which individual chemical moieties are connected to a short single-stranded biotinylated DNA-fragment serving as a “donor strand”. Upon hybridization to the template DNA strand, the chemical moiety is transferred to the template strand by covalent bond formation, followed by cleavage and avidin-assisted donor strand removal (i-iv). Macrocyclic structures could be obtained in the last reaction step (iv). The picture was adapted from Buller et al., 2009.³⁷

Liu and coworkers developed the so called “DNA-templated synthesis,” which is an alternative method to synthesize a DNA-encoded chemical library using complementary oligonucleotide derivatives with the objective of facilitating chemical reactions (Figure I-7C, B).^{46, 48} The DNA sequence represents the code for the organic chemistry steps performed on the same DNA strand. Each chemical fragment is attached to a short single-stranded biotinylated DNA-fragment, playing the role of a “donor strand”. The transfer of the chemical moiety to the template strand occurs by covalent bond formation as a result of hybridization to the template DNA strand. The cleavage takes place immediately, followed by avidin-assisted donor strand removal.

They also described how DNA-templated chemical reactions can be accelerated using hybridization-induced proximity of the two reactants.^{46, 48} When compared to single and dual pharmacophore DNA encoded chemical libraries, the concept of DNA-templated synthesis may allow the execution of multiple rounds of selection and amplification, not unlike the rounds of panning in phage and ribosome display technologies.⁴⁶ The library format describes the chemical re-synthesis of the enriched library members after PCR amplification (Figure I-7C, I-7B), so that preferential binders can be amplified in a repetitive manner. Furthermore, the synthesis of a templated DNA encoded library of 13,000 small molecule macrocycles via amide synthesis has been published.^{48, 52}

The use of DNA three-way junctions generated a different version of DNA-templated library synthesis (Figure I-7D). This library is characterized by the coupling of amino acids and short peptides representing the chemical building blocks to oligonucleotides using cleavable and non-cleavable linkers. The oligonucleotide sequences display a code sequence specific for the chemical building block and a hybridization domain building hairpin loop structures.⁵⁰ The stepwise self-assembly of the DNA-conjugated chemical building blocks is facilitated by the hybridization domain, enabling a chemical reaction between the different building blocks. The next step entails gel-purification of the reaction product, ligation of the DNA fragments and cleavage of the linker (Figure I-7D). The DNA-junction seems to promote DNA-templated chemical reactions due to its close proximity to the reaction centers. During the formation of the three chemical reactions, the DNA-junction is taken apart by primer extension, and a single pharmacophore library with linear, double-stranded DNA tags is built and employable for selection experiments. DNA-templated synthesis also permits the re-synthesis of library members after selection assay and PCR amplification. This methodology presents the possibility of recombining the DNA-junctions after digestion of the PCR product and removal of the non-coding DNA-strand. The resynthesis of the encoded compounds from the PCR products is resumed in a series of steps, including (i) ligation of building blocks, (ii) digestion and purification, (iii) recombination of the DNA-junction, (iv) chemical coupling, and (v) linker cleavage.⁵⁰ Further experiments have been performed in this area in order to expand the practical use of this technology. In this context, a library of 100 members was tested by Hansen and coworkers in a proof-of-principle experiment, executing two rounds of *in vitro* selection on an immobilized antibody target, leading to the isolation of an enkephalin peptide used as a positive control.⁵⁰

A different concept for the encoding of chemical reactions with DNA was developed by Halpin and Harbury.^{40, 53-55} In contrast to DNA-templated reactions, they established a strategy based on the sequence specific separation of DNA fragments followed by a chemical reaction. The use of oligonucleotide hybridization chromatography enables this separation. These oligonucleotides make use of columns with oligonucleotides complementary to coding sequences in the library that are immobilized on sepharose resin (the “anticodon column”). This methodology allows for libraries of oligonucleotides (containing constant regions and coding regions) to be directed via a series of anticodon columns. The hybridization of the specific sequence to a column in this series is reached using high salt conditions. In analogy with a DNA-encoded chemical library, this specific code separation of oligonucleotides represents a part of a split-and-pool synthesis. More specifically, the capture of the oligonucleotide on the anticodon column occurs first, followed by the chemical modification of the DNA with a small molecule. In the next step, the library is pooled and a second round of separation and modification can take place.⁵⁵

Another approach to build dual-pharmacophore chemical libraries involves the combinatorial self-assembly of two complementary sub-libraries, in which each DNA strand represents a small molecule (Figure I-7E, I-9).^{9, 49, 56-59} The first sub-library is comprised of chemical moieties linked to the 5'-end of distinct single-stranded DNA fragments, while chemical moieties are attached at the 3'-end of unique DNA fragments in the second sub-library, including a hybridization domain for the first sub-library (Figure I-9).^{9, 49, 59} Both sub-libraries can be obtained by the individual covalent coupling of chemical moieties to modified DNA fragments. An example is the coupling of a reactive moiety to an amino-modified DNA fragment (Figure I-9).⁶⁰

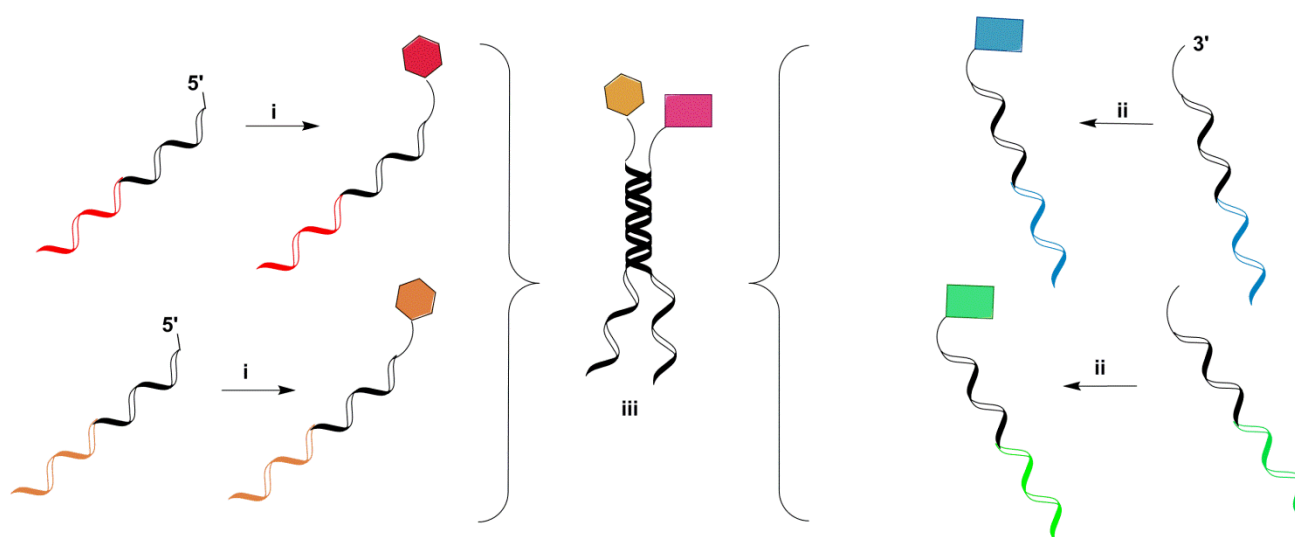


Figure I-9: Dual-pharmacophore chemical libraries can be self-assembled out of two complementary sub-libraries,⁶¹ in which each DNA strand presents a small molecule (iii). The first sub-library contains chemical moieties attached at the 5'-end of distinct single-stranded oligonucleotides (i), whereas the second sub-library displays chemical moieties at the 3'-end of unique oligonucleotides (ii), containing a hybridization domain for the first sub-library. The picture was adapted from Buller et al., 2009.³⁷

The HPLC purification of the sub-libraries after conjugation with a small molecule guarantees a high quality library. Furthermore, large combinatorial libraries can be constructed by hybridizing two smaller sub-libraries. The exposition of two neighboring chemical structures makes it possible to simultaneously bind two compounds to adjacent non-overlapping binding sites on a target protein (Figure I-10).⁵⁸ Later, akin to fragment-based drug discovery technologies, these individual binding entities must be attached and assembled into a single organic molecule.^{9, 59} In fact, the methodology of self-assembling DNA-encoded chemical libraries can be enlarged to “triple-“and “quadruple-libraries”. As was the case with DNA duplexes, the incorporation of a third and fourth sub-library will further increase library size. The process of bringing four sub-libraries together would necessitate reverse Hoogsteen base pairing,⁶² whereas it is possible to assemble four sub-libraries with oligonucleotides capable of quadruplex formation.⁶³⁻⁶⁴ However, neither “triple” nor “quadruple” DNA-encoded chemical libraries have been published to date.

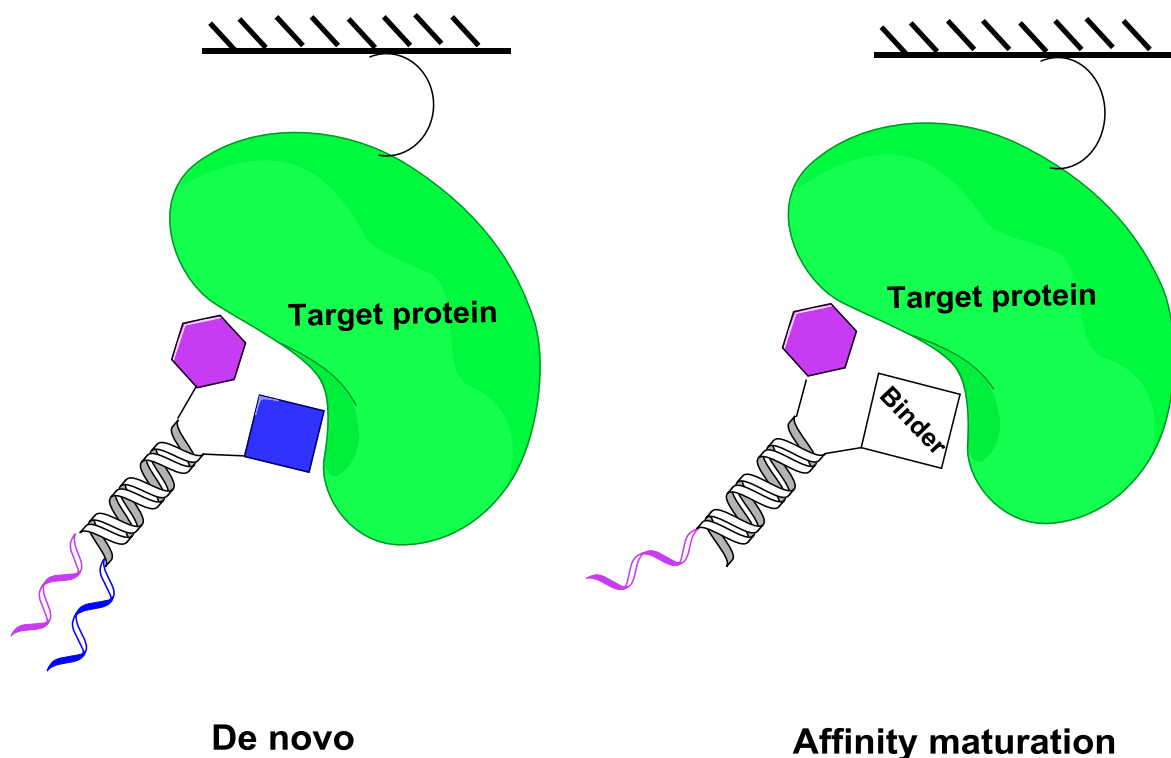


Figure I-10: De novo selection against a target protein results after hybridization of two DNA-encoded sub-libraries. In addition, a known lead structure can be affinity matured using a DNA-encoded library which is constructed by hybridization of an oligonucleotide conjugated to the lead structure and a DNA-encoded sub-library. The figure was adapted from Buller et al., 2009.³⁷

I-4. Selection strategies

Similar to the antibody phage display concept, a DNA-encoded chemical library can be selected by affinity capture against any target protein of choice (Figure I-11).^{27, 65-66} The library is incubated at low pM to fM concentrations of individual library members with the target protein. Unbound fractions are removed by washing steps, while bound library members are amplified by PCR (Figure I-13). In the next step, the obtained amplicon mixture must then be deconvoluted by quantifying the relative concentration of all library members, for example with high-throughput sequencing.^{37, 45, 47}

Finally, detected hits can be resynthesized and characterized biologically in the absence of the DNA tag.^{37, 45, 47} The separation of bound and unbound library members (known as “selection”) can be performed using multiple technologies. The choice of technology depends also on the characteristics of the protein of interest.⁶⁵ Normally, libraries with a volume of 10-100 μ L can be panned against immobilized protein.

Antigens immobilized via CNBr-activated sepharose resin^{9, 37, 47, 59} and His-tagged proteins on metal affinity chromatography resin (IMAC) can also be envisaged (Figure I-11A).³² It has been reported that the density of the protein on the resin modifies the selection pressure, with lower coating densities associated with more stringent selection conditions.^{9, 67} Another alternative is to incubate the DNA-encoded chemical library in solution with a biotinylated interaction partner, followed by capture of the binding fraction on streptavidin-coated resin or magnetic beads (Figure I-11B).⁴⁷ As with antibody phage display technology, selection efficacy can be improved by modifying the concentration of the target protein (e.g. the coating density on the resin) or the number of washing steps.^{9, 67} Unspecific binding can be prevented or avoided during the selection using blocking agents, comprising non-related DNA (e.g. herring sperm DNA) and bovine serum albumin.^{9, 37, 45, 47, 59} Selection technologies could also be investigated in which binding DNA-conjugates are substituted by known ligands with the objective of guiding library selections towards a different pocket of the protein (Figure I-11C). In the case of polypeptide display methodologies, selection experiments can be conducted via an epitope of choice on the targeted protein by the use of a competing ligand (Figure I-11D), or by executing parallel panning reactions on a set of closely related proteins.⁶⁸

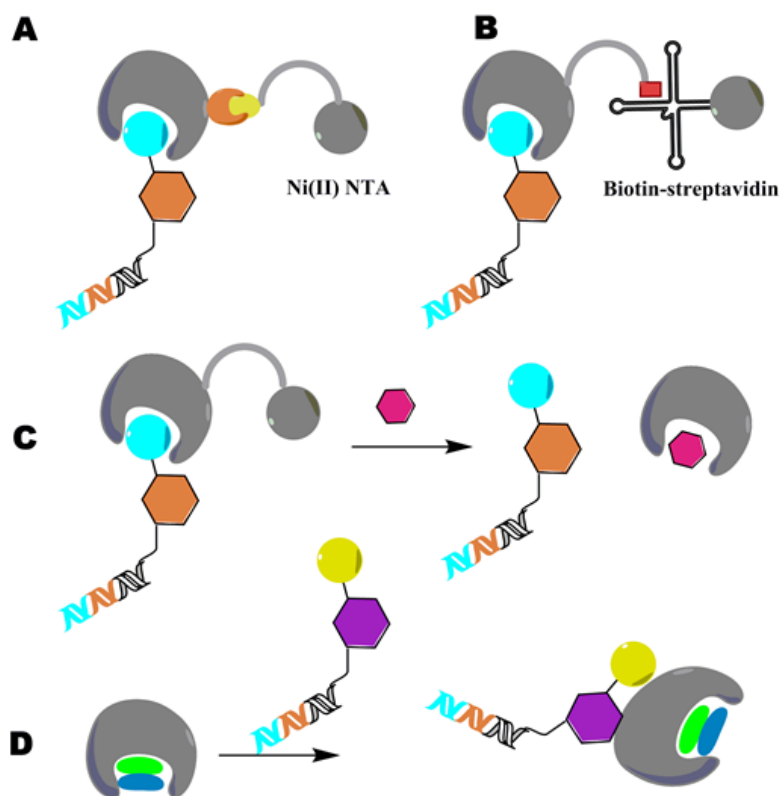


Figure I-11: Affinity-based selections can be performed by A) immobilization using common tags for protein purification (e.g. polyhistidine tag with Ni^{2+} or Co^{2+} immobilized via chelating ligands on resin or magnetic beads); by B) immobilization of biotinylated proteins on streptavidin coated resin or magnetic beads; by C) Elution of the protein from the solid support can be performed by washing of the resin or by competing ligands in order to direct library selections towards a distinct pocket of the protein; D) In principle, the blocking of epitopes or binding pockets with known binders before selection might permit the exclusion of possible binding sites for the selection of ligands (e.g., to direct ligands to allosteric binding sites). The figure was adapted from Hoogenboom et al, 2009.⁶⁸

I-5. DNA sequencing for library decoding

Following affinity selection experiments, the enriched binding molecules can be identified and quantified from a DEL by combining PCR amplification of eluted DNA codes, DNA sequencing, and computational decoding.^{37, 47} Initially, the decoding of a selection amplicon was carried out using “Sanger-sequencing”, which involves sub-cloning and sequencing of the individual colonies. This approach is based on a relatively long and accurate sequencing read length and is therefore inappropriate for decoding a library comprising millions of compounds. Sanger-sequencing is cost-effective in cases where the amplicon population is restricted to a small number of sequences, but not for a DEL containing up to hundred compounds.⁶⁷⁻⁶⁸ Microarray-based methodologies have since emerged, allowing high-resolution read-out of large DNA-encoded chemical libraries.

Collection of DNA-distribution information was fulfilled by hybridization of fluorescent DNA-amplicons to complementary DNA-microarrays. However, the physical limitation of the number of probes that can be arrayed on the surface and the low resolution for similar DNA-sequences make the use of micro-array inefficient for the deconvolution of DNA-encoded libraries possessing more than a few thousand compounds. With the emergence of ultra-high-throughput DNA sequencing technologies, selection experiments using libraries containing millions of compounds could be efficiently and inexpensively processed in less than a few days, allowing the cost-effective sequencing of more than 10^7 sequences per run (at about 1,000 \$ per selection). As described by Mannocci *et al.* in 2008, a single Roche 454 high-throughput sequencing experiment afforded over one million useful deconvolution sequences and enabled the rapid "one-pot" deconvolution of several selection experiments using a 4000 member DEL library.⁴⁷ Indeed, DNA-encoded chemical libraries are becoming larger and larger. This observation highlights the need for alternative high-throughput sequencing platforms that further facilitate the deconvolution of larger libraries comprising up to billions of chemical moieties in just a few days and at affordable costs. Technological improvements based on single-molecule sequencing (SMRT sequencing) or on sequencing by ligation (SOLiD sequencing) may be a good starting point in the process of facilitating the analysis of these large libraries.

I-6. Bioactive compounds from DNA-encoded small molecule libraries

Since 2004, several academic groups and pharmaceutical companies have successfully explored and demonstrated the efficacy and efficiency of DNA-encoded technology in the optimization and discovery or identification of small molecules capable of specific binding to medically relevant target proteins. In 2012, the synthesis and screening of a high-quality library containing 30,000 drug-like compounds yielded a specific inhibitor of interleukin-2 (IL-2, a pro-inflammatory cytokine) in the micromolar range.³⁸ The novel compound **14** (see table 1) showed inhibition of IL-2-mediated T-cell proliferation with an $IC_{50}=32 \mu\text{M}$, without any cytotoxicity to primary fibroblast cell cultures up to $128 \mu\text{M}$ concentrations.³⁸ Computational docking experiments based on the crystal structure of IL-2 yielded a 2-methyl-1H-indole moiety of 67 out of 100 structurally related enriched compounds in the same hydrophobic groove of IL-2 with a strong tendency for accommodating aromatic groups.³⁸

The same group reported in another publication a Diels-Alder based DNA-conjugate library from which they have isolated specific inhibitors against the tumor associated antigen carbonic anhydrase IX (compound **4** with an IC_{50} of 240 nM) and tumor necrosis factor (TNF alpha): compound **9**. This compound showed complete inhibition of L-M fibroblast TNF-mediated killing of L-M fibroblast occurring at concentrations as high as 300 nM).³⁷ Selections against Bcl-xL, an anti-apoptotic target protein which plays a role in cancer treatment, enabled the discovery of different related structures with affinity to the target ranging between 60 μ M and 0.93 μ M.⁵¹ The best ligand identified was able to compete with a known Bcl-xL antagonist and to induce cell-death in Raji cells at micromolar concentrations.⁵¹ Noticeably, the selected compound contained the molecular structure of indomethacin, a known anti-inflammatory compound able to induce apoptosis in multiple cancer cell lines. In addition to the selection of these new compounds, scientists from Philochem and from the Swiss Federal Institute of Technology in Zurich (ETH) used DNA-conjugate libraries to improve previously discovered protein binders. In a first publication, a DNA-conjugate "affinity maturation" library permitted a 10,000-fold improvement of a benzamidine-based trypsin inhibitor in a single round of optimization, yielding a nanomolar range compound.⁶⁷

In 2009, the selection and high throughput sequencing-based decoding of DNA-conjugate libraries comprising more than one billion members enabled the isolation of a series of structurally related triazine inhibitors against Aurora A and p38 MAP (p38 mitogen-activated protein) kinases **8** with an EC_{50} in the submicromolar range.^{45, 69} In 2012, Deng et al. described the identification of a novel ADAMTS-5 inhibitor scaffold **13** (see table 1) devoid of the classical zinc-binding functionality, obtained from a GlaxoSmithKline's (GSK) DNA-encoded library containing 4 billion of compounds.⁷⁰⁻⁷¹ ADAMTS-5 represents a metalloprotease considered as a potential target for the treatment of osteoarthritis.⁷⁰ Following affinity selection, the authors were able to isolate an exemplar compound (R)-N-((1-(4-(but-3-en-1-ylamino)-6-(((2-(thiophen-2-yl)thiazol-4-yl)methyl)amino)-1,3,5-triazin-2-yl)pyrrolidin-2-yl)methyl)-4-propylbenzenesulfonamide **13** able to block the action of ADAMTS-5 with an IC_{50} = 30 nM and excellent selectivity towards other related metalloprotease (>50-fold against ADAMTS-4 and >1000-fold against ADAMTS-1, ADAMTS-13, MMP-13, and TACE).⁷⁰

In 2013, GSK published the application of DEL technology for the identification of a novel selective inhibitor of recombinant human, rat and mouse soluble epoxide hydrolase (sEH or EPHX2).⁷² The novel compound ((1R,3S)-N-(4-cyano-2-(trifluoromethyl)benzyl)-3-((4-methyl-6-(methylamino)-1,3,5-triazin-2-yl)amino)cyclohexanecarboxamide) **16** displayed a potent cell-based activity (human IC₅₀ = 0.66 nM) as well as a concentration-dependent in vitro inhibition of the enzyme in human, rat and mouse whole blood (human IC₅₀ = 6.83 nM). Soluble epoxide hydrolase (sEH) inhibitors have shown a great efficacy as therapeutic agents in some pathological conditions, including diabetes, metabolic syndrome, cardiovascular disease, and pain. This novel isolated compound is giving rise to a new chemical class of sEH inhibitors highlighted by the authors. Unfortunately, this compound failed in the clinic phase.⁷¹ In the same year, GSK reported the application of DNA-encoding in the affinity screening of a 1.2 million DNA-encoded heterocycle library for the identification of SIRT1, SIRT2 and SIRT3 pan-inhibitors with nanomolar potency.⁶⁹ SIRT1/2/3 are NAD⁺ dependent deacetylases.⁶⁹ They are considered to be potential targets for metabolic, inflammatory, oncologic and neurodegenerative disorders.^[69b] After selection and structure-affinity relationship (SAR) analysis were applied to improve physiochemical properties, three thieno[3,2-d]pyrimidine-6-carboxamide derivatives **15** were identified, revealing IC₅₀ as high as 3.6, 2.7 and 4.0 nM for SIRT1, SIRT2 and SIRT3 respectively.⁶⁹

In early 2014, two multi-laboratory collaborations headed by GSK described in two independent reports the use of DNA-encoded conjugate libraries for the de novo identification of key compounds for the treatment of multidrug resistant tuberculosis and autoimmune inflammatory diseases respectively. The first report described the use of a 16.1 million member DNA-conjugate library generated by split-and-pool assembly of three different set of building blocks, making use of orthogonally protected di-amino acids as core structures and peptide-like chemistry for the synthesis.⁷³ Affinity selection against biotinylated InhA (Inhabin, alpha) immobilized on a streptavidin matrix allowed the identification of a new class of potent and selective *M. tuberculosis* direct InhA inhibitors.⁷³ Finally, medicinal chemistry optimization of the initial hit **18** (IC₅₀ = 34nM) led to the synthesis of a diethylpyrazole derivative with an InhA IC₅₀ as good as 4nM.⁷³ Unfortunately, despite the optimal in vitro antitubercular potency and pharmacokinetic profile, the compound proved to be inactive in vivo against a murine TB acute infection model.⁷³

A GlaxoSmithKline (GSK) team in collaboration with the laboratories of the Immune Disease Institute, Children's Hospital Boston, described in a further report the use of a previously synthesized 4.1 billion member library to identify potent inhibitors of the protein-protein interaction between lymphocyte function-associated antigen 1 (Integrin LFA-1) and its major ligand intercellular adhesion molecule 1 (ICAM-1).⁷⁴ Standard affinity selection experiments were performed and yielded an enriched small-molecule hit family **19** with submicromolar potency in both ELISA and cell adhesion assays after off-DNA resynthesis (ELISA $IC_{50} > 16\text{nM}$).⁷⁴

In addition, the resynthesis of these compounds and conjugation to DNA or a fluorophore showed them to bind to cells expressing the target protein by fluorescent flow cytometry.^[68d] This data was the first published evidence that DEL affinity selections can be performed against a target in its natural state on the cell surface.⁷⁴ In 2015, the Swiss Federal Institute of Technology in Zurich (ETH Zurich), together with the Structural Genomics Consortium (University of Toronto), the Department of Medical Biochemistry and Biophysics (Karolinska Institute, Stockholm) and Philochem, reported the design and synthesis of DAL-100K, a DNA-encoded chemical library comprising 103,200 compounds. Affinity-based selection experiments and DNA high-throughput sequencing analysis led to the identification of ligands with nanomolar or subnanomolar affinities to prostate-specific membrane antigen (PSMA; $K_D = 830\text{ nM}$ fluorescence polarization), human serum albumin (HSA; $K_D = 0.70 \pm 0.04\text{ nM}$ fluorescence polarization) and tankyrase 1 (TNKS1; $IC_{50} < 370\text{ nM}$ PARylation assay) respectively.

In December 2015, GSK Molecular Discovery Research described the first use of a cell-based selection method for identifying small-molecule inhibitors from DNA-encoded libraries. The selection campaign was conducted against tachykinin receptor 3 (NK3), a GPCR trans-membrane protein target for the treatment of central nervous system disorders as well as irritable bowel syndrome, and enabled the isolation of highly efficient (nanomolar and sub-nanomolar range) structurally diverse triazine and non-triazine small molecular antagonists which showed excellent selectivity versus the related NK1 and NK2 receptors.⁷⁵

In 2016, GSK researchers were able to isolate novel biphenyl pyrrolidine ether based inhibitors from the screening of a 34.7 million member Suzuki–Miyaura based DNA-encoded compound library against the mitochondrial isozyme (BCATm).^{71, 76}

Branched chain aminotransferases (BCAT) are defined as aminotransferase enzymes which catalyze the transamination of branched amino acids.⁷⁶ Studies in mice demonstrate the mitochondrial isozyme (BCATm) to be a possible therapeutic target for metabolic diseases, such as obesity.⁷⁶ SAR investigations showed both the methylsulfonamido and the bromothiophene moieties present on the compound structure to be essential for biological activity.⁷⁶

Finally, X-ray crystallographic analysis was performed for the most biologically active compound **20** ($IC_{50} = 2.0 \mu\text{M}$), revealing the newly isolated compound to bind to the BCATm catalytic site through multiple H-bond and van der Waals interactions.⁷⁶ It is important to mention that the library used in this work constitutes the first report of a fully disclosed DEL compound library synthesized using on-DNA Suzuki–Miyaura cross-coupling reactions.

I-7. DNA-routed libraries

In 2007, Harbury and coworkers described in a “proof-of principle” experiment the DNA-routed chemical translation of an octamer peptoid library containing 107 compounds. The structurally related peptoid structures were identified from six rounds of molecular evolution on N-CrkSH3 (N-terminal SH3 domain of Crk protein) with micromolar affinities. The best binder exhibited a K_D of 16 μM .⁵⁴ With the objective of developing a novel and potentially isoform and/or mutant selective class of PI3Kp110 α inhibitors, the group from GSK performed selection assays against both wildtype His-tagged PI3K α and the mutant H1047R, using their DNA-encoded chemical library platform. The panning of a combinatorial DEL library of about 3.5 million compounds consisted of 191 amino acids (used for peptide coupling), 95 boronates (used in Suzuki coupling), and 196 amines (for reductive amination) and, in addition to compounds interacting with the ATP binding site, yielded a 4-amino-N-(5-methyl-1,3,4-thiadiazol-2-yl)benzenesulfonamide building block preferentially enriched together with three different boronate building blocks.⁷⁷ They chose a representative set of 10 compounds from the selected molecules and synthesized them off-DNA in order to investigate their activity against PI3K α . The inhibition assay exhibited exceptional activity for most of the re-synthesized compounds with IC_{50} values as good as 6.5 nM.

Nevertheless, they displayed poor selectivity against mutant H1047R⁷⁷ and appeared to show a preference for the gamma isoform. In March 2015, the drug discovery team of GSK also reported the use of DNA-encoded chemical library technology for the identification of a novel class of compounds that bind to the HCV NS4B protein. This is a small (27 kDa), hydrophobic, membrane associated protein involved in hepatitis C viral replication.⁷⁷ Following screening of 28 different DELs, containing up to 8 billion compounds, the GSK team found GSK0109 and GSK4809 to bind to the HCV NS4B protein with an activity against the HCV replicon in tissue culture of more than 100 nM.⁷⁷ This report emphasizes the utility of DELs as a powerful tool for identifying antivirals which act directly on non-enzymatic targets.

I-8. DNA-templated libraries

In 2012, Liu and coworkers described the use of a number of methods for the de novo discovery of bioactive molecules. The selection and decoding of a 13,824-member macrocycle library against 36 different target proteins yielded new classes of kinases modulators.⁴⁸ Selected compounds included p38 α -MAPKAP2 cascade inhibitor (11 μ M), VEGFR2 activators (up to 300% activity enhancement at 100 μ M concentration) as well as ATP-competitive Src inhibitors (with IC₅₀ as potent as 680nM and 960nM).⁴⁸ “Molecular evolution” analysis of this library produced a second-generation of macrocycles with IC₅₀ \leq 4 nM and the ability to inhibit Src activity in cultured mammalian cells.⁴⁸

Ensemble Therapeutics also used DNA-directed libraries to build a 160,000 macro-cyclic peptidomimetic small molecule library to address the protein-protein interaction between XIAP (X-linked inhibitor of apoptosis protein) binding domain (BIR2 and BIR3) and caspase initiator to promote apoptosis.⁷⁸ Cancer is one of the diseases for which the apoptosis of malignant cells is desired. The process and mechanism of apoptosis is complex, with many pathways implicated. However, most of them converge upon the activation of common effector caspases. X-linked inhibitor of apoptosis protein (XIAP) is known to bind and inhibit the activation of different key initiator and effector caspases.⁷⁸

For this reason, the isolation of small molecule antagonists capable of disrupting the protein–protein interaction between the XIAP binding domains (BIR2 and BIR3) and caspase initiator is essential for the clinical treatment of multiple disorders, including neurodegeneration, cancer, and autoimmune diseases. To reach this goal, the scientists of Ensemble Therapeutics, in collaboration with BMS, used this library as well as various more focused libraries to perform multiple rounds of selection (against biotinylated BIR2 on streptavidin resin and his-tagged BIR3 on nickel resin). SAR investigations and compound re-synthesis ultimately led to the identification of three macrocycles with potent antiproliferative activity in human triple negative breast cancer and A875 melanoma cell-lines with sub-100 nM IC₅₀ values.⁷⁸ ViperGen patented the YoctoReactor technology (yR), which they applied for the synthesis of a three building block DNA-directed library containing 12.6 million compounds. The selection of the library against the p38 α kinase protein resulted in the identification of a new specific and potent small molecular inhibitor **31** with an IC₅₀ as good as 7nM for inhibition of TNF- α secretion in a human monocytic cell line.^{37, 50} X-ray crystallographic investigations proved an alternative binding mode of the newly isolated inhibitor with respect to previously reported MAP kinase small molecule modulators.³⁷

I-9. Self-assembled libraries

In 2008, the application of an ESAC single-library allowed the identification and characterization of a new family of specific portable albumin binders based on 4-(p-iodophenyl)-butyric acid moiety, with affinities to human serum albumin as high as 3.2 μ M.⁵⁷ The best of the newly identified ligands was conjugate to blood pool contrast agents (fluorescein and Gd-DTPA), showing a notable increase in the compound circulatory half-life and improving imaging performance with regard to the non-conjugate compounds. Recently, the use of a pseudo-ESAC duplex format has been described for the discovery of novel inhibitors of stromelysin-1 (MMP3), a matrix metalloproteinase over-expressed in pathological tissue remodeling processes.⁵⁹ Screening experiments on human MMP-3 were performed making use of a 550-membered ESAC sub-library, while microarray-based decoding yielded the initial lead binding structure. This compound was further investigated in an ESAC affinity maturation experiment. Decoding and covalent conjugation of binding molecules in the absence of the DNA-Tag revealed a MMP3 inhibitor with an IC₅₀ as good as 9.9 μ M.⁵⁹

Using a PNA/DNA self-assembling approach, Winssinger and co-workers at ISIS Strasbourg described for the first time in 2011 the synthesis of a 37,485 member carbohydrate DNA-encoded library containing consensus sequences for DC-SIGN (dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin), a protein implicated in the initial stages of the human immunodeficiency virus infection.⁶⁰ After selection, SPR (Surface Plasmon Resonance) of the best isolated pair revealed an improved binding to the target protein ($K_D = 0.127 \mu\text{M}$).⁶⁰ SPR represents an optical technique used to study molecular interactions.⁷⁹

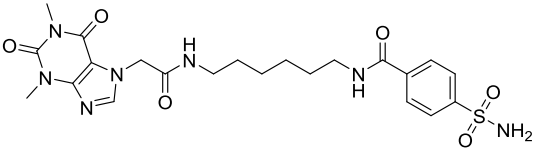
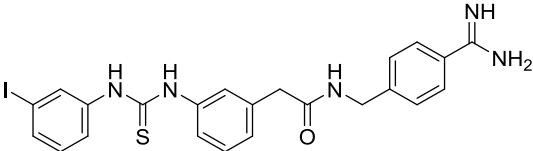
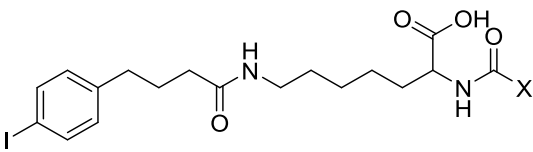
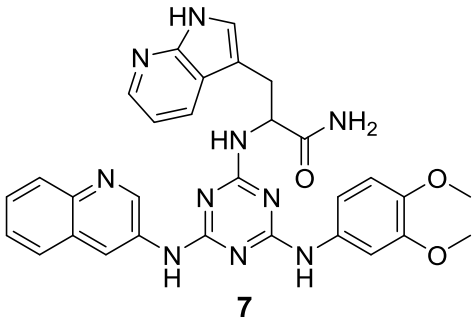
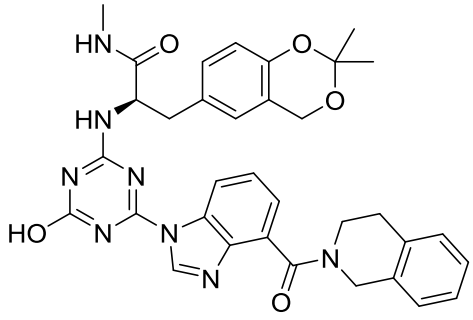
A research team at the ETH Zurich together with Philochem reported the synthesis and high-throughput sequencing decoding of an ESAC-duplex library comprising 111,100 compounds. The application of a novel encoding strategy, transferring the sequence information from one strand to the complementary one, enabled the isolation of binding pairs working in combination.⁶¹ The authors specifically described the de novo identification of an alpha-1-acid glycoprotein (AGP) binder **4** with a K_D (isothermal titration calorimetry) as good as $5.8 \pm 2.3 \mu\text{M}$ and the optimization of a potent carbonic anhydrase IX (CAIX) ligand.⁶¹ This compound displayed a subnanomolar affinity towards CAIX by fluorescence polarization ($K_D = 0.2 \pm 0.1 \text{nM}$) and slightly higher dissociation constants by SPR analysis ($K_D = 2.6 \text{nM}$).⁶¹ An in vivo experiment performed using mice with subcutaneously grafted SK-RC-52 tumors revealed the selective effect and long-lasting targeting of the tumor mass of the newly discovered high-affinity bidentate CAIX ligand. At the same time, the cognate monodentate acetazolamide analog gradually dissociated from the neoplastic mass, demonstrating the therapeutic benefits of the compounds. Quantitative biodistribution of homogenized tissues presented a superior uptake of the bidentate compound ($10 \pm 0.2\% \text{IDg}^{-1}$ at 24h) with respect to the monodentate acetazolamide derivative ($2 \pm 0.3\% \text{IDg}^{-1}$) and two other monoclonal antibodies specific to CAIX ($1-5\% \text{IDg}^{-1}$).⁶¹ The authors also reported the resynthesis and hit-validation of the selected binding pairs by fluorescence polarization, making use of a fluorescently labelled 8-mer LNA heteroduplex conjugate.⁶¹ In the same year, the group of Professor Winssinger (NCCR Chemical Biology, University of Geneva) published work on the use of self-assembling DNA/PNA-hybrid encoded libraries containing 65,000 compounds to identify binding pharmacophore to Hsp 70.⁶⁰

The description of a suitable linker to covalently pair the fragments is generally a difficult task. The information yielded from the fragment screen constituted a starting point for the synthesis of a related 10,000-member PNA-conjugate library encompassing different linkers with different lengths and geometrical constraints.⁶⁰ Ultimately, the selection campaign led to the identification of the most potent Hsp70 ligand described so far ($K_D < 12.10$ nM by SPR). The ligand displayed good selectivity for Hsp70 (Hsp70/Hsp72) across the proteome. At the same time, none of the individual fragments used in the synthesis of the bidentate compounds showed any interactions with the target under the same conditions, thus confirming a synergetic interaction.⁷⁷ The phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha (PI3K p110 α) represents the most frequently mutated kinase in human cancer.⁷⁷

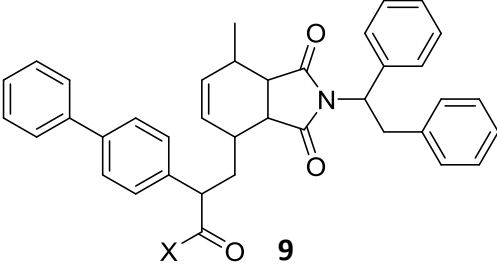
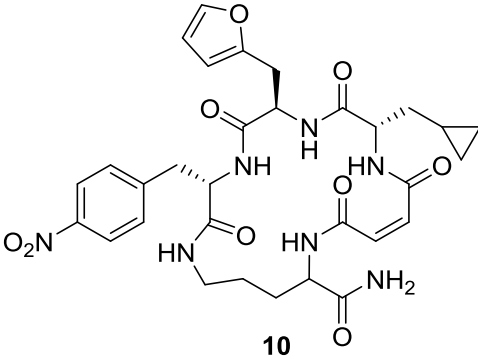
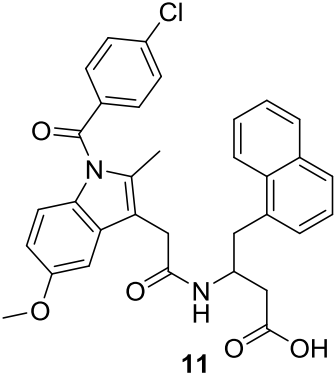
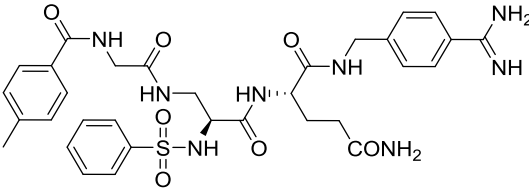
Screening against the prototypical phosphatase (PTP1B) and microarray-based decoding of a hybrid self-assembled combinatorial library, containing 62,500 fragment combinations, enabled the isolation of several fragments.⁴² Following this, a 10,000 member DNA-conjugate “affinity maturation” library, covalently bound to the newly discovered fragments with linkers of different length and geometry, was synthesized. A panning experiment of the latest lead-optimization library on PTP1B as well as on the closely related TCPTP allowed the development of orthogonal ligands with nanomolar affinity ($K_D = 50$ nM) and enzymatic inhibition ($EC_{50} = 250$ nM).⁴² In summary, DEL technology is gaining more attention leading to its implementation as a routine for their drug discovery programs in various academic groups and pharmaceutical companies. Table 1 below provides an overview of different lead compounds obtained from the synthesis and screening of various DELs with diverse size.

INTRODUCTION

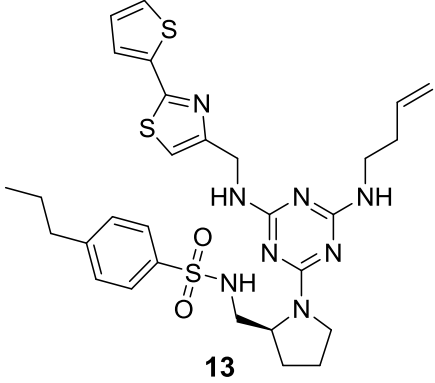
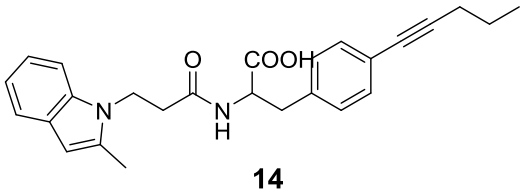
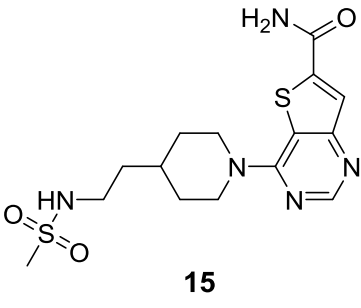
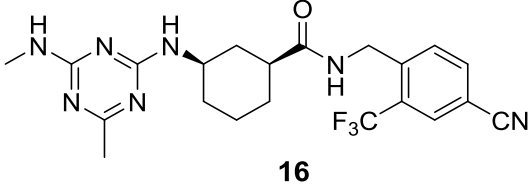
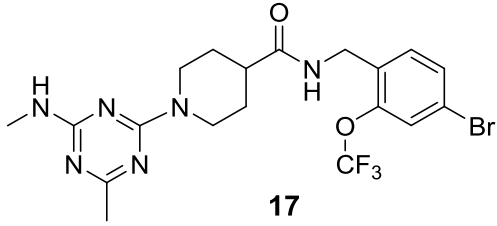
Table 1: Overview of bioactive compounds **4 -31**.

entry	library size	target (potency)	structure	organisation	reference
1	138	Carbonic Anhydrase II (12 nM)	 <p style="text-align: center;">4</p>	ETH Zurich	Melkko et al. 2004 ⁴⁹
2	550	Trypsin (98 nM)	 <p style="text-align: center;">5</p>	ETH Zurich	Melkko et al. 2007 ⁹
3	550	HSA (3.2 μM)	 <p style="text-align: center;">6</p>	Philochem	Dumelin et al. 2008 ⁵⁷
4	7e6	Aurora A kinase (0.27 μM)	 <p style="text-align: center;">7</p>	GlaxoSmithKine	Clark et al. 2009 ⁴⁵
5	8e8	P38 MAP kinase (0.25 μM)	 <p style="text-align: center;">8</p>	GlaxoSmithKine	Clark et al. 2009 ⁴⁵

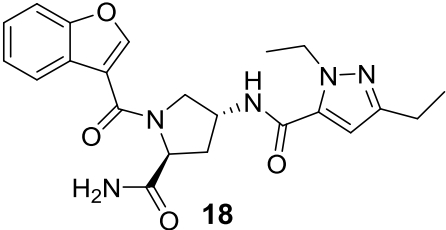
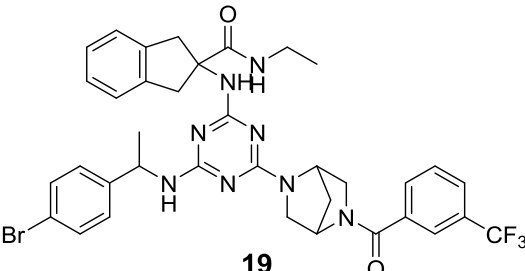
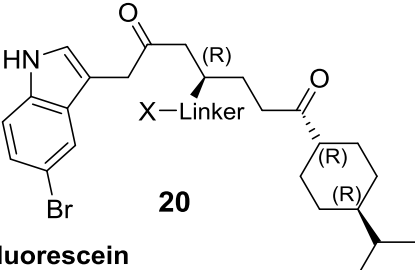
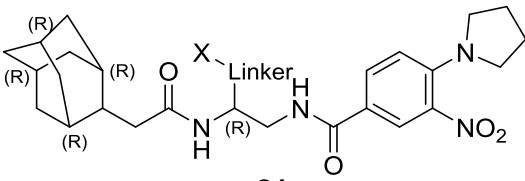
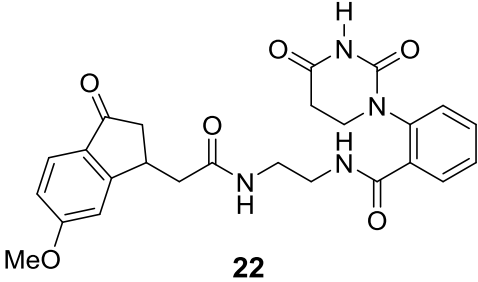
INTRODUCTION

entry	library size	target (potency)	structure	organisation	reference
6	4000	TNF (20 μ M)	 <p>9</p>	ETH Zurich	Buller et al. 2009 ³⁷
7	13824	Src kinase (680 nM)	 <p>10</p>	Harvard University	Kleiner et al. 2010 ⁸⁰
8	4000	Bcl-sL (0.93 μ M)	 <p>11</p>	Philochem	Melkko et al. 2010 ⁵¹
9	8000	Trypsin (3nM)	 <p>12</p>	Philochem	Manocci et al. 2010 ⁶⁷

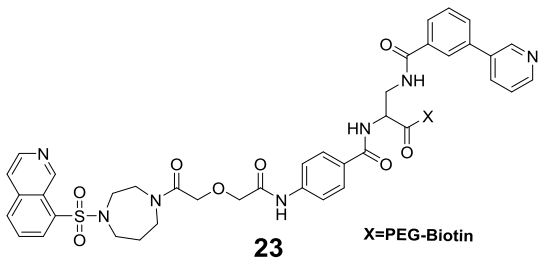
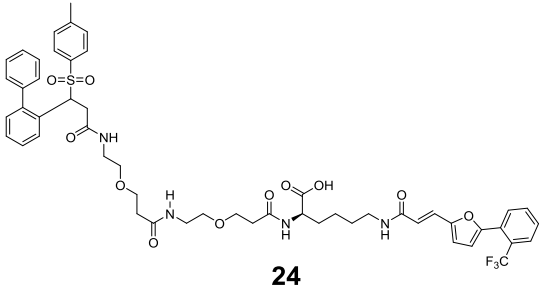
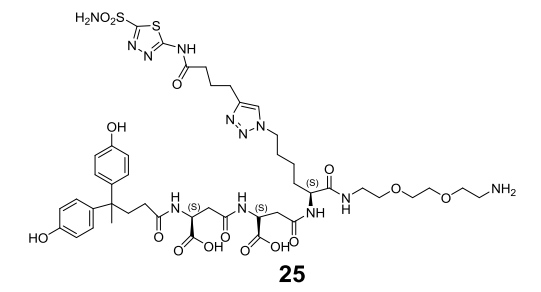
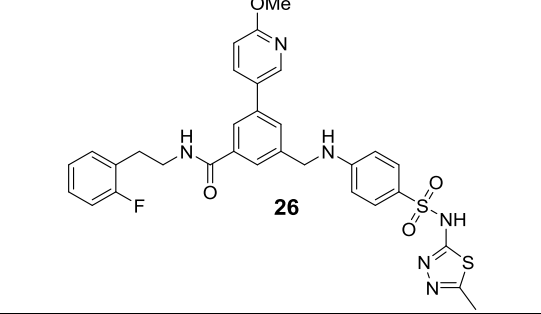
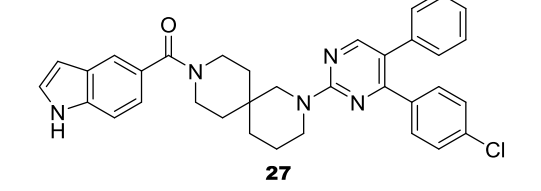
INTRODUCTION

entry	library size	target (potency)	structure	organisation	reference
10	4e9	ADAMTS-5 (30 nM)	 13	GlaxoSmithKine	Deng et al. 2012 ⁷⁰
11	30000	IL-2 (2.5 μM)	 14	ETH Zurich/Philochem	Leimbacher et al. 2012 ³⁸
12	1.2e6	SIRT 1/2/3 (ca. 3.5 nM)	 15	GlaxoSmithKine	Disch et al. 2013 ⁶⁹
13	Not disclosed	sEH (27 pM)	 16	GlaxoSmithKine	Podolin et al. 2013 ⁸¹
14	Not disclosed	sEH (0.5 nM)	 17	GlaxoSmithKine	Thalji et al. 2013 ⁷²

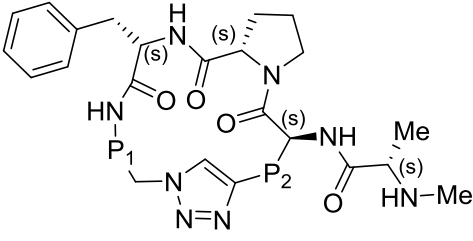
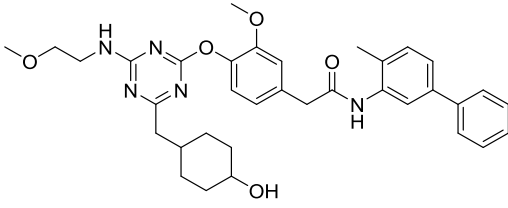
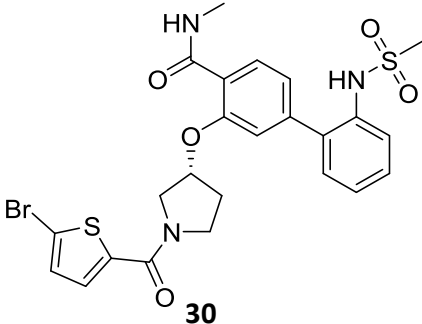
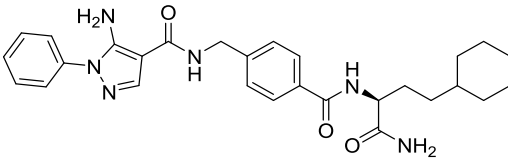
INTRODUCTION

entry	library size	target (potency)	structure	organisation	reference
15	16e6	InhA (4nM)	 18	GlaxoSmithKine	Encinas et al. 2014 ⁷³
16	4e9	Integrin LFA-1 (16 nM)	 19	GlaxoSmithKine	Kollmann et al. 2014 ⁷⁴
17	103200	HSA (0.70 nM)	 20 X= Fluorescein	ETH Zurich	Franzini et al. 2015 ⁸²
18	103200	PSMA (830 nM)	 21 X= Fluorescein	ETH Zurich	Franzini et al. 2015 ⁸²
19	103200	TSKS 1 (290 nM)	 22	ETH Zurich	Franzini et al. 2015 ⁸²

INTRODUCTION

entry	library size	target (potency)	structure	organisation	reference
20	62500	Hsp 70 (0.38 nM)	 <p>23 X=PEG-Biotin</p>	University of Geneva	Daguer et al. 2015 ⁸³
21	111100	AGP (5.8 μM)	 <p>24</p>	ETH Zurich	Wichert et al. 2015 ⁶¹
22	111100	CAIX (2.6 nM)	 <p>25</p>	ETH Zurich	Wichert et al. 2015 ⁶¹
23	Ca. 3.5e6	PI3K, p110a (> 6.5 nM)	 <p>26</p>	GSK	Yang et al. 2015 ⁷⁷
24	Up to 8 billion	HCV NS4B (< 100 nM)	 <p>27</p>	GSK	Arico-Muendel et al. 2015 ⁸⁴

INTRODUCTION

entry	library size	target (potency)	structure	organisation	reference
25	Up to 16e4 (macrocycles)	XIAP BIR2-3 (<100 nM)	 <p>28</p>	Ensemble Therapeutics Corp BMS Research & Development	Seigal et al. 2015 ⁷⁸
26	Up to 15 billion	NK3 (nM/pM)	 <p>29</p>	GSK	Wu et al. 2015 ⁷⁵
27	34.7 million	BCATm (2 μM)	 <p>30</p>	GSK	Deng et al. 2016 ⁷⁰
28	12 million	P38a kinase (7 nM)	 <p>31</p>	Vipergeren ApS	Petersen et al. 2016 ⁸⁵

I-10. Privileged scaffolds in medicinal chemistry

In 1988, Evans used the term “privileged structures” for the first time in an article,⁸⁶ in which he defined them as important molecular fragments or subunits present in the molecules capable of interacting with several different receptors, probably because of being able to present functional groups or substituents in a favorable arrangement.⁸⁶ Accordingly such compounds or sub-structures are highly represented in the population of bioactive compounds. The concept of a privileged scaffold concept can be used to design high-affinity protein ligands from core structures that can bind to more than one receptor. The intention behind the introduction of this concept in medicinal chemistry was to accelerate the drug discovery process. The identification and use of privileged scaffolds might increase the probability to identify hits after screening or selection experiments against target proteins with pharmaceutical relevance.

The first class of molecules described as “privileged” was the benzodiazepine family, a class of drugs that is efficient in the treatment of anxiety, insomnia, psychomotor agitation, epileptic seizure and many others.⁸⁶ The benzodiazepines had been discovered by chance, following an experiment on the contents of an abandoned flask found during a lab clean-up.⁸⁶ Yet, not only benzodiazepines are nowadays views as privileged scaffolds but also 1,4-dihydropyridines, indoles, quinolones, benzofuran, isoquinoline, biphenyls⁸⁷ and more. Several of these scaffolds also occur in natural products, thereby proposing an evolutionary pressure favoring such structures.

On the other hand, we classify other structures, such as pyrrolinone, as “privileged” because of their rigid framework, which enables them to direct functional groups in a well-defined space. These classes of privileged structures produce leads with improved drug-like properties and the privileged scaffold-based design has emerged from an independent concept to an integral component in various lead discovery platforms (Table 2, Table 3). The applications of the privileged scaffold concept in current lead discovery processes have been described by many authors. Most reviews have highlighted the role of privileged scaffolds in the development of compounds to feed high-throughput screening and their role in the design of ligands targeting protein-protein interactions, multiple ligands or warhead-based ligands. Table 2 and 3 provides a list of privileged identified from molecules which are largely used by the pharmaceutical industry and which are contained in drugs.⁸⁷⁻⁸⁸

Table 2: Selected molecular scaffolds and corresponding examples of privileged structures. The figure is adapted from Schneider and Baringhaus, 2008.⁸⁸


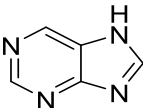
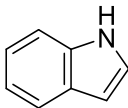
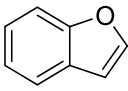

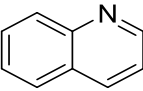
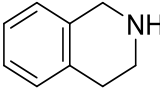
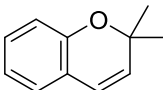

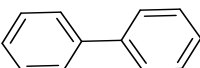
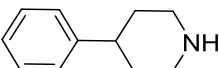
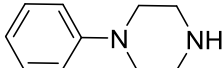
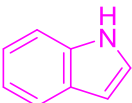
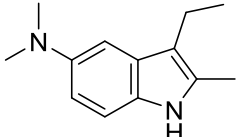
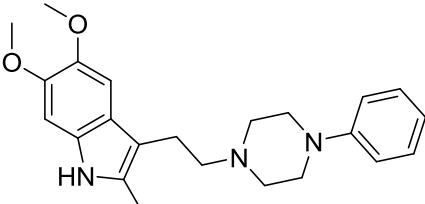
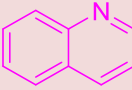
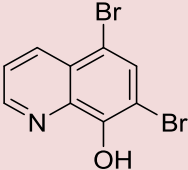
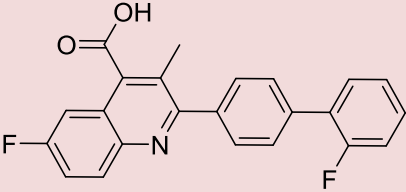
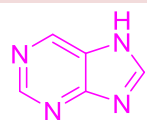
Molecular scaffold	Examples of privileged scaffolds		
	 Purine	 Indole	 Benzofuran
	 Quinoline	 Tetrahydroisoquinoline	 3,3-dimethylbenzopyran
	 Biphenyl	 Arylpiperidine	 Arylpiperazine

Table 3: Selected privileged structures and corresponding drug examples. Privileged structures are found in drugs against several classes of protein targets. The table was adapted from Welsch et al., 2010.⁸⁹

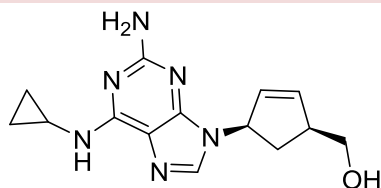
Privileged scaffolds	Drugs structures	
 Indole	 Medmain <i>Serotonin inhibitor</i>	 Oxypertine <i>Antidepressant</i>
 Quinoline	 Broxyquinoline <i>Antiseptic, disinfectant</i>	 Brequinar <i>Immunosuppressant</i>

Privileged scaffolds

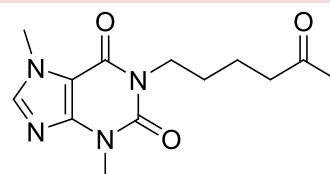
Drugs structures



Purine



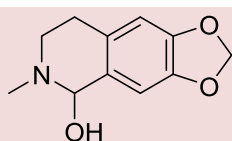
Abacavir
Antiviral (HIV)



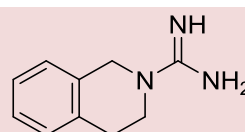
Pentoxifylline
Hemorheologic agent



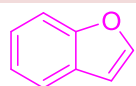
Tetrahydroisoquinoline



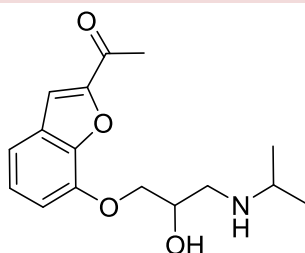
Hydrastinine
Hemostatic



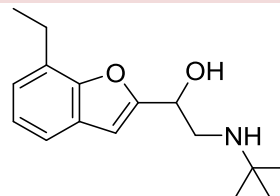
Debrisoquin
Antihypertensive



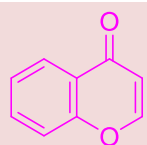
Benzofuran



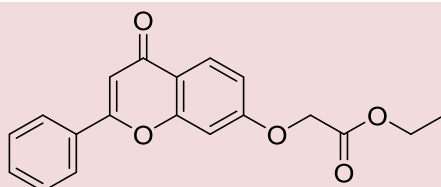
Befunolol
Antiglaucoma



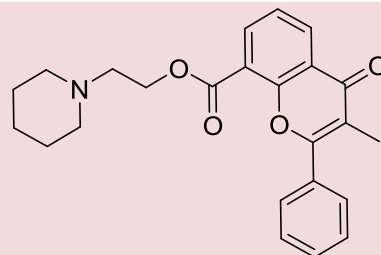
Buturalol
Antianginal; antihypertensive



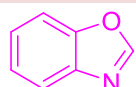
Chromone



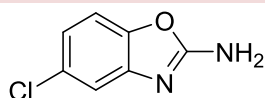
Efloxate
Vasodilator (Coronary)



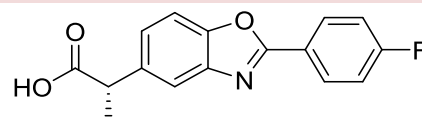
Flavoxate
Antispasmodic



Benzoxazole



Zoxazolamine
Muscle relaxant (skeletal)



Flunoxaprofen
Anti-inflammatory

I-11. Limitations of DNA-encoded libraries

DNA-encoded chemical library technology enables the construction of very large compound libraries. The synthesis and screening of DNA-encoded chemical libraries are more efficient than biochemical HTS processes. Moreover, the technology has reached a good level of maturity and has yielded ligands to several different targets. However, it does have limitations. The principal limitation in the field of DNA-encoded Chemical Libraries have been identified as (i) unsatisfactory quality controls in the synthesis of DELs (e.g., systematic verification of sufficiently high reaction yields in order to achieve equimolarity among library members)⁴³, (ii) deviation from the standard rules of library construction, such as the concepts of lead-likeness and privileged structures⁴³, and (iii) the poorly documented impact of experimental conditions in affinity selection on selection readouts.⁴³ The synthesis of DELs based on privileged scaffolds, as outlined in this work, might allow for identification of bioactive compounds with properties that enable further development.

I-12. Aims and objectives

As explained above, there is a vital need to develop drug discovery approaches capable of rapidly generating a large chemical compound libraries for identification of bioactive compounds.^{4, 8} The central objective of my PhD thesis was to establish a protocol for the synthesis of a sizable DNA-encoded library based on the benzodiazepine scaffold (Figure I-13). Benzodiazepines are important members of this family. Benzodiazepine-based compounds are well-established approved drugs (e.g. Diazepam[®]). Furthermore the scaffold has yielded, for example, inhibitors of bromodomains (**6**), E-3-ubiquitin-ligase HDM2 (**7**), RabGGTase (**8**) and the clinical candidate BMS214662 (**9**) (Figure I-14).⁹⁰⁻⁹¹ The benzodiazepine-based library synthesized may represent a good possibility to identify bioactive compounds for members from this interesting class of targets as demonstrated by the clinical candidate **9**. Based on the conceptual translation of antibody phage display technology, which relies on linkage between the antibody binding properties (phenotype) and the genetic information coding for the antibody (genotype), the DNA-encoded chemical library consisted of collections of benzodiazepines molecules, covalently linked to a unique DNA tag serving as an amplifiable identification bar code (Figure I-12).

INTRODUCTION

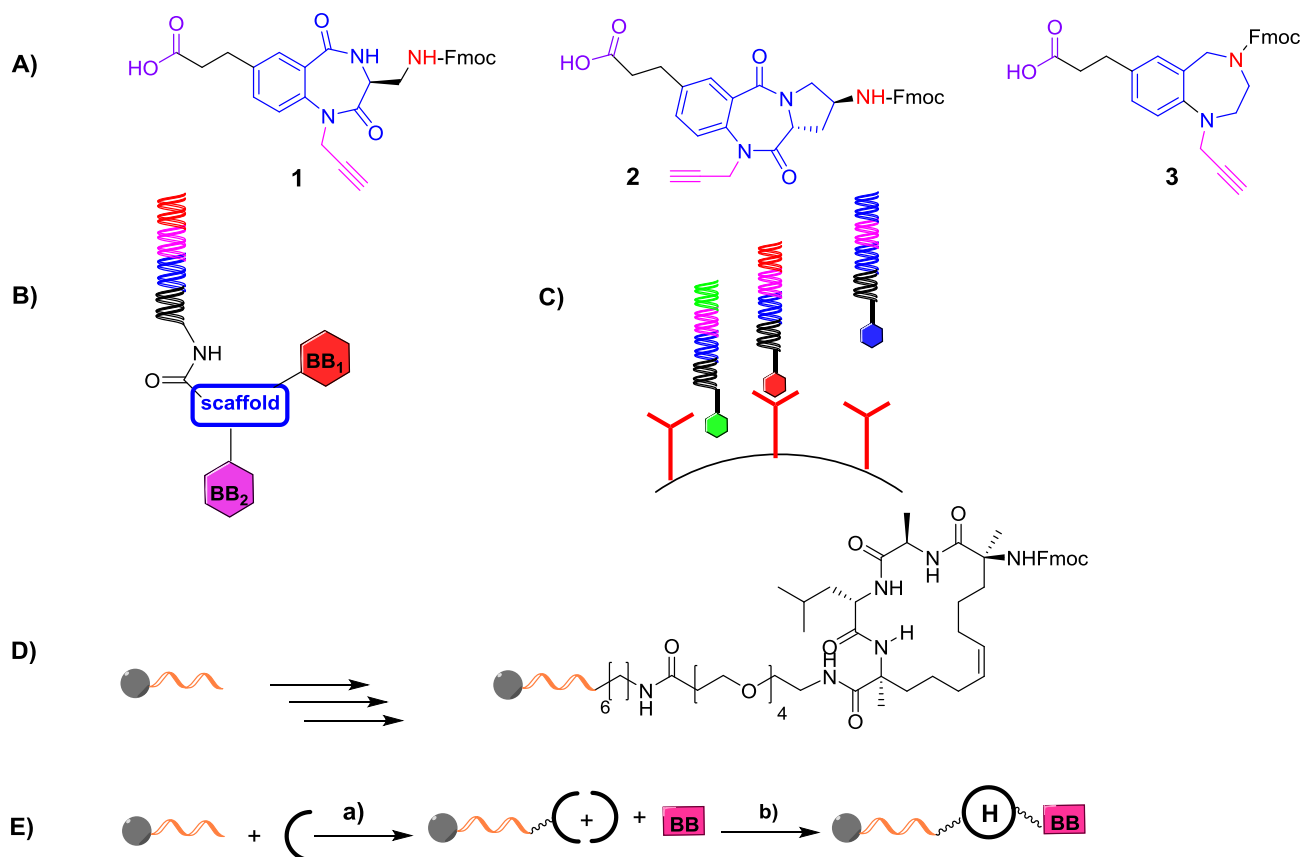


Figure I-12: Overview of the objectives in the current work. A) Synthesis of benzodiazepine functionalized scaffolds. B) Attachment of the scaffolds to DNA, combinatorial chemistry, exploration of DNA-compatible synthetic methods. C) Screening of the pooled libraries at protein targets. D) Macrocyclization of a peptide composed of four amino acids by ruthenium-catalyzed ring-closing metathesis (RCM) on DNA on solid phase making use of the first generation Grubbs catalyst. E) Development of an oligoThymidine initiated DNA-Encoded Chemistry.

In the first step, three benzodiazepine scaffolds were synthesized in sufficient amounts for test reactions and library synthesis (200-400 mg for each scaffold). **Figure II-9:** Appearance of the three synthesized benzodiazepines **1, 2, 3**.

These scaffolds were designed in a manner that allowed attachment of a DNA-code and decoration by combinatorial chemistry. They possessed a carboxylic acid to couple them to aminolinker-modified DNA, as well as an amino functional group to which a set of carboxylic acid (first set of building blocks) were coupled and a terminal alkyne that reacted with a set of azides (second set of building blocks) to furnish triazoles. The two sets of building blocks, 114 carboxylic acids and 104 azides (Figure I-15), were selected by chemoinformatic methods with the aim of controlling the physicochemical properties of the final library removing both undesired chemical fragments and PAINS-like-structures (Pan-Assay-Interference compounds),⁹² and improving library diversity.

INTRODUCTION

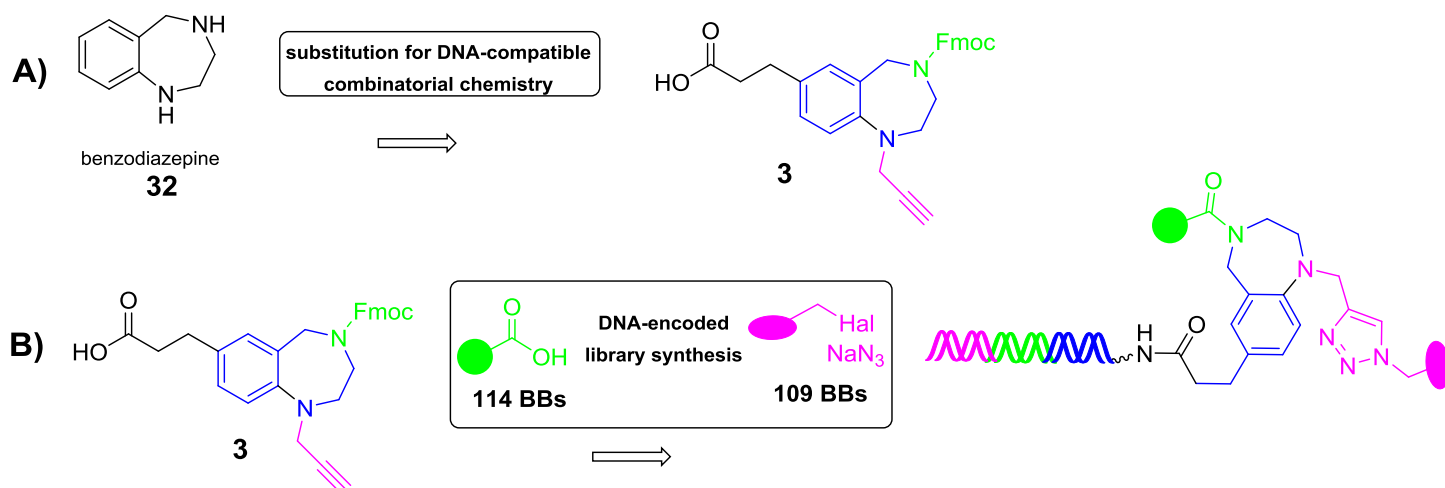


Figure I-13: A) Benzodiazepine structure functionalized for encoded library synthesis **3**; B) DNA-encoded library based on the structure **3**.

The DNA-encoded library was then synthesized with the benzodiazepine scaffold **32** by split-and-pool combinatorial chemistry, yielding a library that contains 9588 compounds. The combinatorial library was encoded by DNA-ligation, where 5'-phosphorylated double-stranded coding DNA sequences with four base overhangs were ligated with T4 ligase. Finally, the encoded library was pooled together with libraries synthesized by other members of the group and screened against disease-relevant target proteins provided by collaboration partners. Combinatorial libraries based on privileged scaffolds may hold more promise of yielding bioactive compounds, which may be improved in subsequent rounds of optimization with regard to their pharmacokinetic properties, for example.

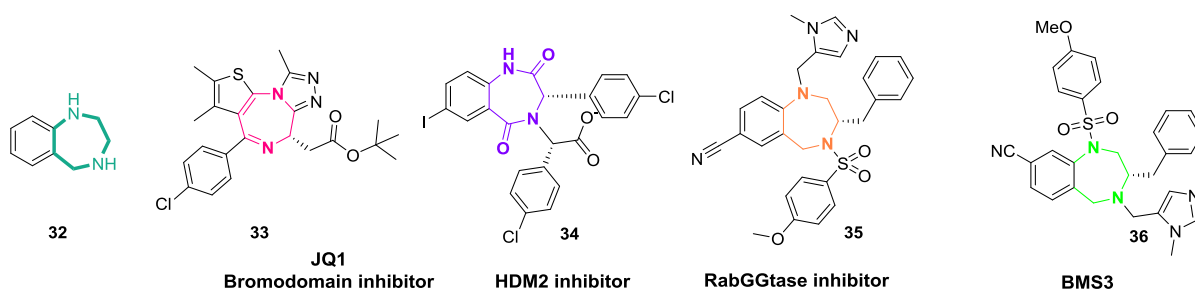


Figure I-14: Bioactive benzodiazepine derivatives.

Only a small number of reactions have so far been utilized to generate DNA-encoded combinatorial libraries: amide synthesis, Diels-Alder reaction, Wittig reaction, reductive amination and stepwise substitution of cyanuric chloride.^{37, 45, 93} A secondary project of the current thesis was to screen for conditions for chemical reactions important in medicinal chemistry, but which have not yet been applied in the construction of DEL. DEL-compatible chemistry demands methodologies that are robust, work in the presence of protic solvents, do not require strong acidic or basic conditions and will not degrade the DNA. At present, synthesis routes to substituted (hetero-) cyclic structures from readily available starting materials are lacking. Neri and coworkers have reported several bioactive compounds obtained from DELs that was synthesized by the Diels-Alder-cycloaddition reaction.³⁷ Also recently, different heterocycles have been synthesized by condensation reactions which are mild and DNA compatible.⁹⁴ In this context, an approach for the synthesis of spirocycles on DNA has been reported.⁹⁵ These reports revealed the necessity to develop new synthetic strategies based on new chemistry for the synthesis of DELs.^{53, 55} The deoxyribonucleic acid molecule (DNA) is composed of four nucleobases: the purines adenine (**A**) and guanine (**G**), as well as the pyrimidines thymine (**T**) and cytosine (**C**) (Figure I-16). Important sources of drug-like heterocyclic structures are Brønsted acid and transition metal catalyzed reactions.⁹⁵ However, the use of strong acids on DNA oligomers would cleave the glycosidic bond of purine nucleosides, and many transition metal ions oxidize the C8 of purines, and/or bind to the N7 of purines leading to depurination through a tautomerism-based mechanism.⁹⁵ The interactions of nucleobases with acids and transition metals are depicted in Figure I-16. In addition, different experiments performed on DNA sequences consisting of purines and pyrimidines bases have shown that the incompatibility of the DNA with many catalyst systems used in organic chemistry is based only on the purine bases. As depurination is a huge barrier or hindrance (Figure I-15) to method development for DELs, DNA-encoded library synthesis based on heterocycles was synthesized and enabled the subsequent ligation of coding DNA sequences based on the assumption that an adapter oligonucleotide composed of thymine nucleobases might allow the use of acid- and transition metal-catalysis. Therefore, to initiate the construction of a DEL based for instance on heterocycles, macrocycles or other scaffolds, short DNA strands consisting only of pyrimidine bases called DNA (HexT) were used.

Thus, the oligoThymidine initiated DNA-Encoded Chemistry (“TiDEC”) strategy for DEL synthesis was developed, in which a 5'-aminolinker modified hexathymidine oligonucleotide “hexT”, enables the utilization of a surprisingly broad spectrum of catalysts and reaction conditions. The hexT was bound to controlled pore glass (CPG) solid support which enabled us to use organic solvents in the synthesis of target molecules, and facilitated purification of the DNA-conjugates.

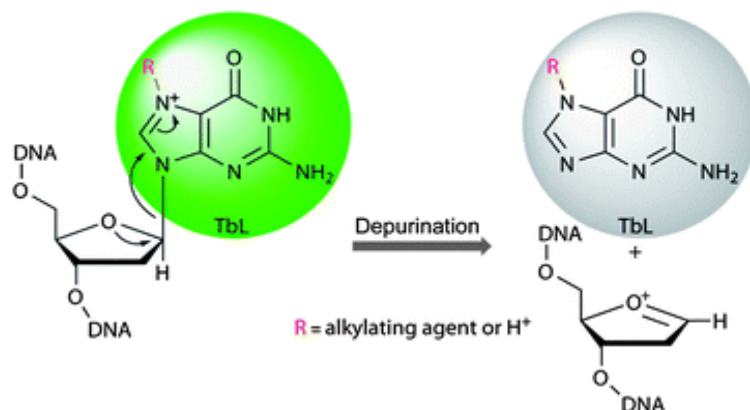


Figure I-15: Specific recognition of DNA depurination by a luminescent terbium (III) complex. The figure was taken from Xiaohui et al., 2013.⁹⁶

The DNA sequences based on pyrimidine bases tolerate the reactivity of a surprisingly wide range of catalyst systems, for instance: Transition metal ions and phosphoric acid-based organocatalysts. This therefore opens up a wide spectrum of chemical reactions, for instance heterocycles synthesis on this DNA. This thesis reported the ability to synthesize large number of interesting structures derived in a DNA-encoded format. The encoding of the synthetic steps with DNA, which contains all nucleobases, takes place only after the first synthesis step has been carried out. Finally it was confirmed that the short pyrimidine initiator DNAs used for the synthesis of DELs can be subsequently recognized by DNA ligases.

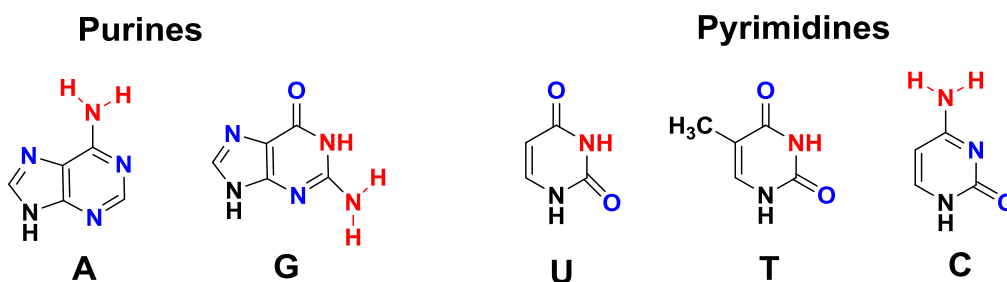


Figure I-16: The four nucleobases and their interactions sites.

INTRODUCTION

In this context, a modular solid phase synthesis strategy for DNA-macrocycle conjugates was developed. Four selected amino acids were coupled together on DNA on solid support by Fmoc-based solid phase peptide synthesis (SPPS) (Figure I-17) for the introduction of the hydrocarbon cross-link. Subsequent cross-linking via RCM used first generation Grubbs catalyst with solid-phase-bound protected peptides to provide the macrocyclic peptide **40**. For this project conditions for Fmoc-peptide chemistry to assemble peptide strands connected to the hexT were optimized. Different parameters were investigated: CPG porosity of the solid support, coupling reagent, activation and reaction time, base, solvent, as well as equivalents of reactants and reagents. The investigation process based on experiments with different Fmoc-protected amino acids yielded suitable reaction conditions for amide synthesis on DNA on solid support. The usage of 20 % of piperidine in DMF was found to be the best protocol to remove the Fmoc group on DNA conjugates without affecting the DNA integrity. Thus a modular synthesis strategy provided access to the ring-closed peptide **40**.

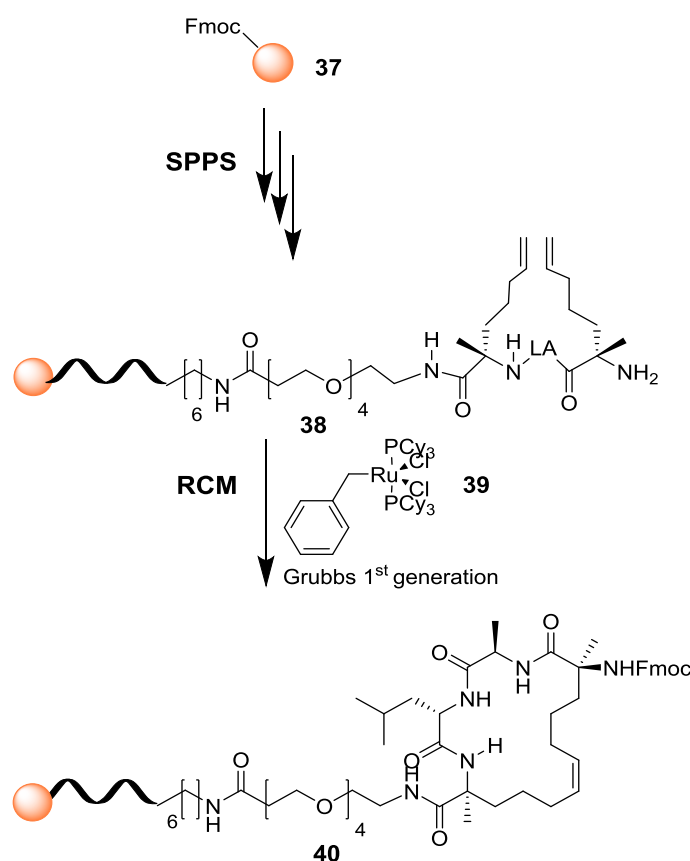


Figure I-17: Olefinic, non-natural amino acids were introduced by SPPS. Ruthenium-catalyzed RCM was performed with first generation Grubbs catalyst at room temperature in dichloromethane (DCM). Reactions were analyzed by HPLC.

Encouraged by the stability of the hexT DNA in the presence of the Grubbs catalyst, a third project was developed applying the Pictet-Spengler reaction conditions to tryptophan coupled to the hexT adapter oligonucleotide in the presence of an aldehyde or ketone to access β -carboline conjugates. This resulted in the successful synthesis of 114 hexT- β -carboline conjugates by Brønsted acid catalyzed Pictet-Spengler reaction. The hexT-heterocycle conjugates were ligated to two coding dsDNAs in a single step and used in a final step as starting material for the introduction of a second set of building blocks to build a DEL library. The construction of such libraries is currently in progress (Figure I-18).

A) DNA-encoded libraries



B) oligoThymidine initiated DNA-Encoded Chemistry

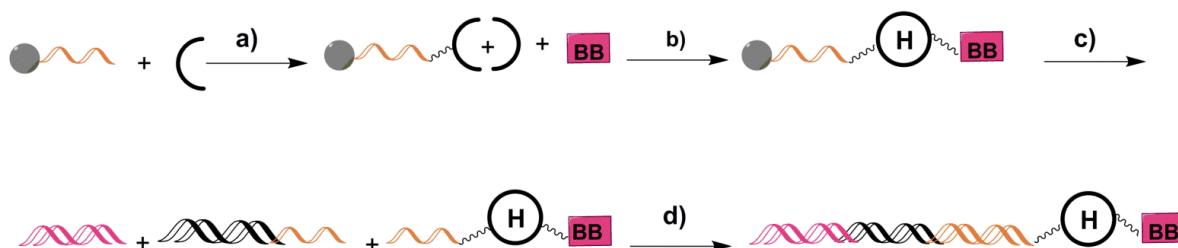
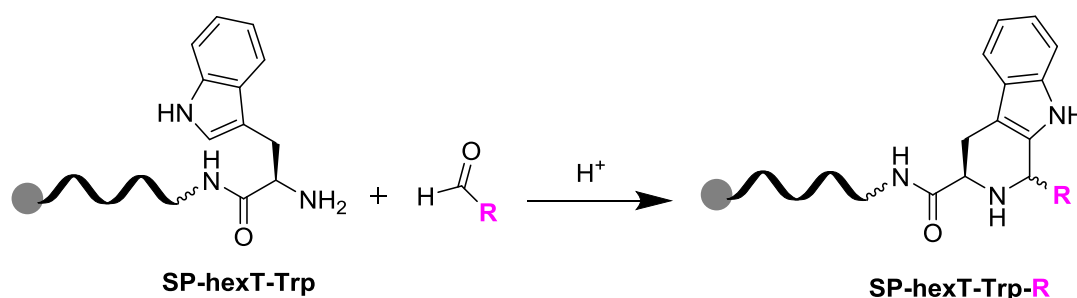


Figure I-18: Synthesis of a DNA-encoded compound by alternated synthesis and encoding steps; a) conjugation of a building block; b) encoding by DNA ligation. B) TiDEC; a) conjugation of starting materials to the hexT adapter; b) synthesis of target heterocycles by acid catalysis; c) isolation of the hexT conjugates; d) encoding by DNA ligation.



Scheme I-1: Pictet-Spengler reaction on hexT bound to controlled pore glass (CPG) solid support for the synthesis of hexT- β -carbolines.

INTRODUCTION

Overall this result will be useful for the development of novel DNA-compatible methods and their use in library design. Since library chemical diversity is often limited to DNA compatible synthetic reactions, this result will extend the field of chemical diversity displayed by DELs. For the chemical diversity displayed by a library is key to successful discovery of potent, novel and drug-like chemical matter.

Chapter II: Synthesis of a DNA-encoded library based on a benzodiazepine scaffold.

II-1. Benzodiazepines as bioactive compounds and drugs

Benzodiazepines continue to be a long-standing topic of scientific interest.⁹⁷⁻¹⁰¹ The understanding of the biological processes taking place in the brain was still limited in the mid 1950's, making the identification of new drugs that influenced the central nervous system a real challenge to medicinal chemists.¹⁰² The discovery of Librium and its clinical value motivated to find more selective and more potent compounds. 7-Chloro-1,3-dihydro-1-methyl-5-phenyl-2H-1,4 benzodiazepin-2-one **42** showed to be significantly more potent than chlordiazepoxide. This benzodiazepinone showing anxiolytic properties was introduced into clinical practice in 1963 under the generic name diazepam or Valium **42**.¹⁰³⁻¹⁰⁶ Today, benzodiazepines (BZDs) represent an important class of antianxiety, hypnotic and muscle relaxing agents.¹⁰⁶⁻¹⁰⁹ The widespread usage of eight benzodiazepine derivatives throughout the U.S. confirms their acceptance as clinical drugs.¹⁰⁴ These benzodiazepine derivatives depicted in figure II-1 belong to anxiolytics except of flurazepam **44**, which is used as hypnotic, and clonazepam **46**, which is marked as an antiepileptic.

1,4-Benzodiazepines have a broad range of biological utilities and have been employed as anxiolytic,¹⁰⁶ anticonvulsant,¹¹⁰ antitumor,¹⁰⁸ and anti-HIV agents.^{109, 111} Among the family of benzodiazepines, 1,4-benzodiazepine-2,5-dione represent a fundamental class of drugs that possess selective activities against a wide array of biological targets. They have been identified as inhibitors of platelet aggregation to mimic the arginine-glycine-aspartic acid (RGD) peptide sequence,¹¹² as precursors of benzodiazepines,¹¹²⁻¹¹³ as anxiolytic agents,¹⁰⁶ and as Hdm2 antagonists to disrupt the p53-Hdm2 protein-protein interaction and induce cell growth arrest and apoptosis.¹¹² Compounds containing the 1,4-benzodiazepin-2,5-dione skeleton have earned attention as potential therapeutics for endocrine disorders.¹⁰² The benzodiazepinedione derivative G5598 has been demonstrated as a potent inhibitor of the binding of GpIIb/IIIa to fibrinogen as well as a potent inhibitor of platelet aggregation.¹¹⁴ 1,4-benzodiazepin-2,5-diones have shown to be valuable pharmacological mimics of the tripeptide RGD (ArgGly-Asp).^{112, 115} The development of new synthetic protocols for BZDs and preparation of BZD analog libraries for biological screening are topics of continuous interest. Introducing the 1,4-benzodiazepin-2,5-dione moiety as scaffold in the synthesis of DNA-encoded library would increase our chances to find novel bioactive compounds.

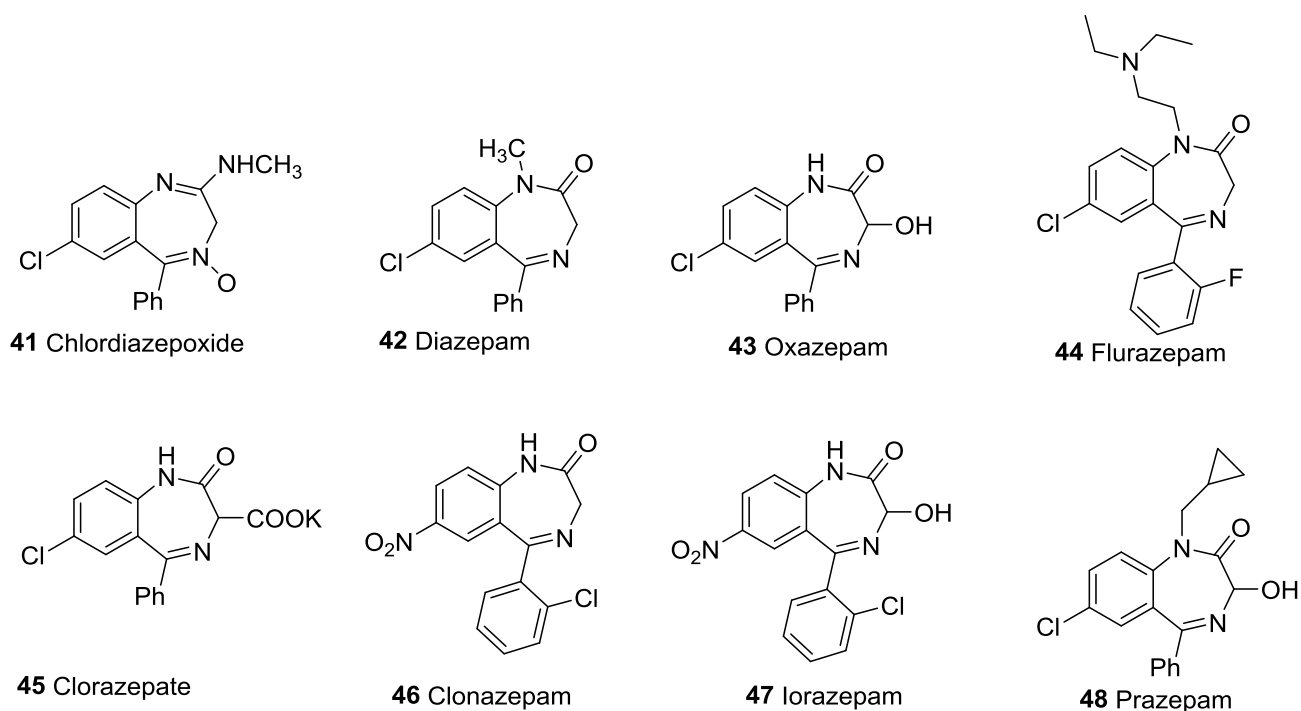


Figure II-1: Benzodiazepine drugs.¹¹⁶

II-2. Synthesis of trifunctionalized benzodiazepines for library synthesis

One focus of this thesis was the synthesis of a DNA-encoded library based on the benzodiazepine scaffold. In figure I-12 the three appropriately substituted benzodiazepines (**1**, **2** and **3**) are shown. The scaffolds carry a carboxylic acid to attach it to amino-linker-modified DNA a), an amino-function to attach diverse building blocks by amide synthesis and a terminal alkyne through which fragments can be attached by copper catalyzed alkyne-azide-cycloaddition (CuAAC). The functionalized, properly protected benzodiazepine scaffold **3** (Figure II-6) will be coupled to a commercially available synthetic DNA containing a 5'-amino group by amide synthesis on solid support (Figure II-10, step a). This strand contains both a 5'-primer sequence for the PCR-amplification following the selection assay and an identification code for the scaffold.

II-2.1. Synthesis of trifunctionalized 1,4-benzodiazepine-2,5-dione scaffold (1)

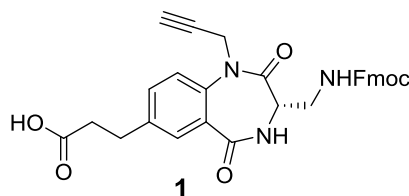
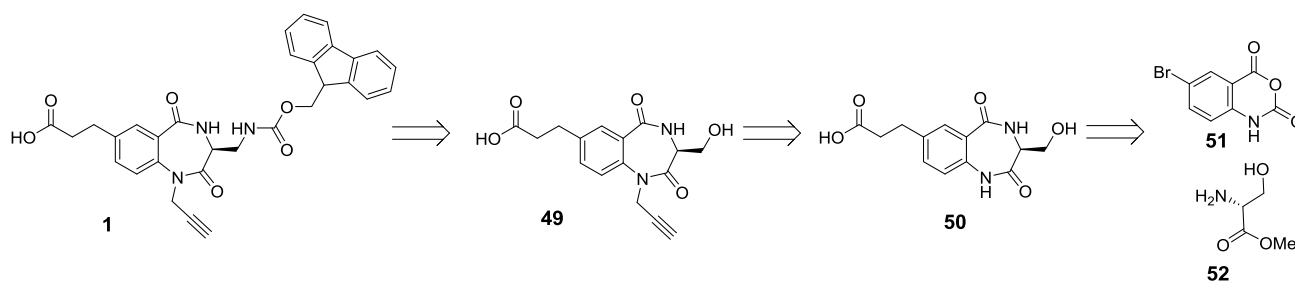


Figure II-2: Structure of the trifunctionalized 1,4-benzodiazepine-2,5-dione scaffold **1**.

The first task was to synthesize the benzodiazepine **1** that is functionalized in way that it allows coupling to 5'-aminolinker modified DNA. It also possesses a carboxylic acid moiety, an Fmoc-protected amino group and a terminal alkyne for subsequent substitutions in combinatorial fashion. The chemical structure of this 1,4-benzodiazepine-2,5-dione consists of a core benzene ring fused to a seven-member 1,4-diazepine ring. It shows a semi-planar platform. The N1-position was substituted with a terminal alkyne.

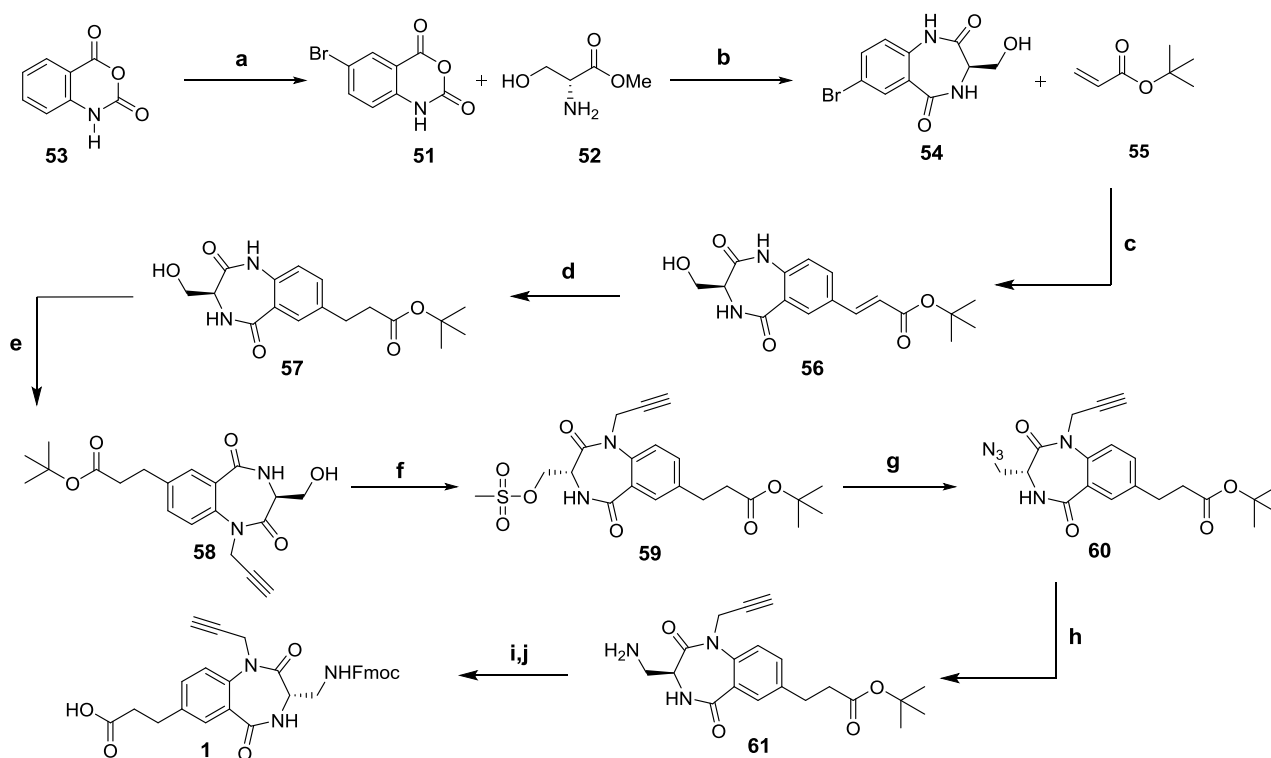
In order to obtain the benzodiazepine **1** that fulfills the requirements for library synthesis, a straightforward route needed to be developed. It was envisioned that the 1,4-benzodiazepine-2,5-dione **1** would be an excellent starting point for introduction of sets of building blocks. Compound **1** could be obtained from building block **49**, which could be constructed from 5-bromoisatoic anhydride **51** and serine methyl ester **52** (Scheme II-1). The 1,4-benzodiazepine-2,5-dione **1** was accessed through a 10 steps synthesis route that started with the bromination of isatoic anhydride **53** (Scheme II-2), the condensation of this product **62** with serine methyl ester **52** led to the core heterocycle **54**.



Scheme II-1: Retrosynthetic analysis of compound **1**.

Synthesis of a DNA-encoded library based on a benzodiazepine scaffold.

This heterocycle reacted with acrylic acid *tert*-butylester **55** by Heck reaction to introduce the carboxylic acid function. The Heck reaction was followed by hydrogenation of the double bond to remove the α,β -unsaturated Michael acceptor function. Alkylation of the aryl amine with propargyl bromide introduced the terminal alkyne (compound **58**). The hydrogenated product **58** was then converted to compound **59** after mesylation with methanesulfonyl chloride. The crude mesylate **59** was then converted without purification to the azide **60** in presence of DMF and sodium azide. The azide **60** was reduced to the amine **61** by the Staudinger reaction. The phosphine imine-forming reaction was followed by hydrolysis to produce a phosphine oxide and the desired amine **61**. Finally, protective groups were removed and the free amine of the compound **61** was protected with the Fmoc group yielding the target compound **1**.



Scheme II-2: Synthesis of 1,4-benzodiazepine-2,5-dione. Reagents and conditions: a) bromine in H₂O at 50 °C for 1 h; b) Serine methyl ester **52**, DMAP, pyridine, 120°C, 3 days; c) Pd(OAc)₂, P(O-tol)₃, *tert*-butyl acrylate **55**, Et₃N, dry CH₃CN, 100 °C, 18 h; d) Pd/C, dry MeOH, rt, 18 h; e) Cs₂CO₃, propargyl bromide, dry DMF, 60 °C, 18 h; f) CH₃SO₂Cl, Pyridine, 0° C-rt, 24 h; g) NaN₃, DMF, 50°C, 72 h; h) Staudinger reaction, PPh₃, dry THF, rt, 24 h; i) TFA in dry CH₂Cl₂, rt, 19 h; j) Fmoc-Osu, NaHCO₃, H₂O, dry 1,4-dioxane, rt, 18 h.

The functionalized 1,4-benzodiazepine-2,5-dione **1** was synthesized in straightforward manner. However, the first step consisting in the benzodiazepine formation **51** was a problematic step due to the poor yield obtained after synthesis under reactions conditions presented above. In fact, the expensive commercially available starting material 5-bromoisatoic anhydride **53** used for the reaction was impure. Thin layer chromatography (T.L.C.) analysis in the system DCM/MeOH (9:1) showed four impurities. These were probably leading to the formation of many side products.

The target benzodiazepine **54** was therefore obtained in poor yields and very complex and difficult to purify. One alternative was to prepare the starting material **51** from the isatoic anhydride **53** according to the scheme above. The synthesis was successful affording the desired 5-bromoisatoic anhydride **51** in satisfying yield which was then used for the benzodiazepine synthesis. It is also important to mention that the use of serine methyl ester hydrochloride **52** instead of serine, as well as the increasing the amount of 4-dimethylaminopyridine (DMAP) and reaction time (4 days instead of 2) has increased the yield of the reaction from 10 to 32 %. The synthetic protocol for the preparation of 1,4-benzodiazepine-2,5-dione **1** using serine methyl ester **52** was thus successfully established. The scaffold **1** was synthesized in sufficient amount (290 mg) for test reactions on DNA and for library synthesis.

II-2.2. Synthesis of trifunctionalized 1,4-benzodiazepine-2,5-dione scaffold (2)

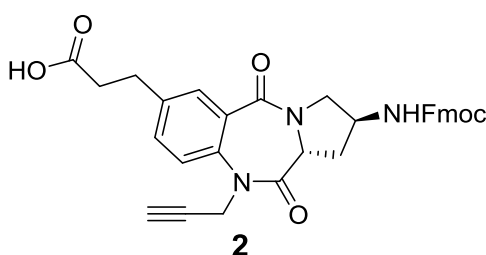


Figure II-3: Structure of the trifunctionalized 1,4-benzodiazepine-2,5-dione scaffold **2**.

Proline-derived 1,4-benzodiazepin-2,5-diones are useful scaffolds in medicinal chemistry (Figure II-4).¹¹⁷⁻¹¹⁸ They have been reported as starting materials for the synthesis of anthramycin-inspired anticancer agents **63**, more recently they have shown promise as herbicides **64**.¹¹⁸

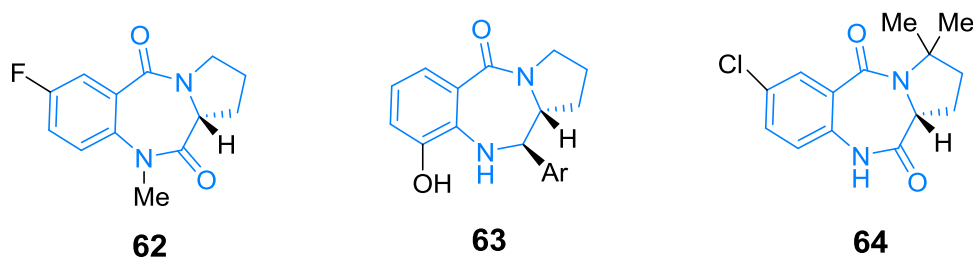


Figure II-5: Proline-derived 1,4-benzodiazepin-2,5-diones **62**, **63**, **64**.

The objective here was to synthesize a novel tricyclic benzodiazepinedione scaffold that is functionalized in a way that allows library synthesis. In order to obtain the benzodiazepine **2** that fulfills the requirements for library synthesis, a versatile route was developed. It was envisioned that the 1,4-benzodiazepine-2,5-dione **2** would be an excellent starting point for introduction of sets of building blocks. Compound **2** could be obtained from building block **66**, which could be constructed from 5-bromoisatoic anhydride **51** and trans-4-hydroxy-L-proline **67**.

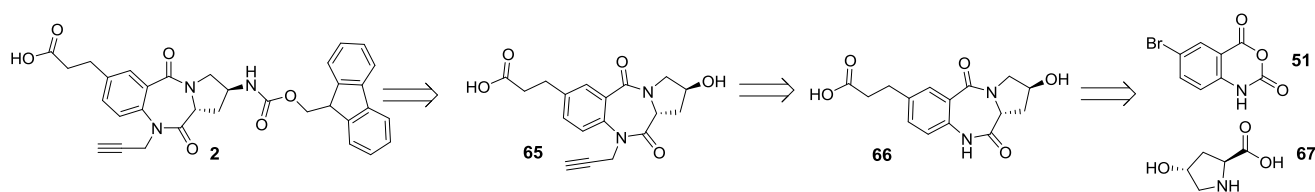


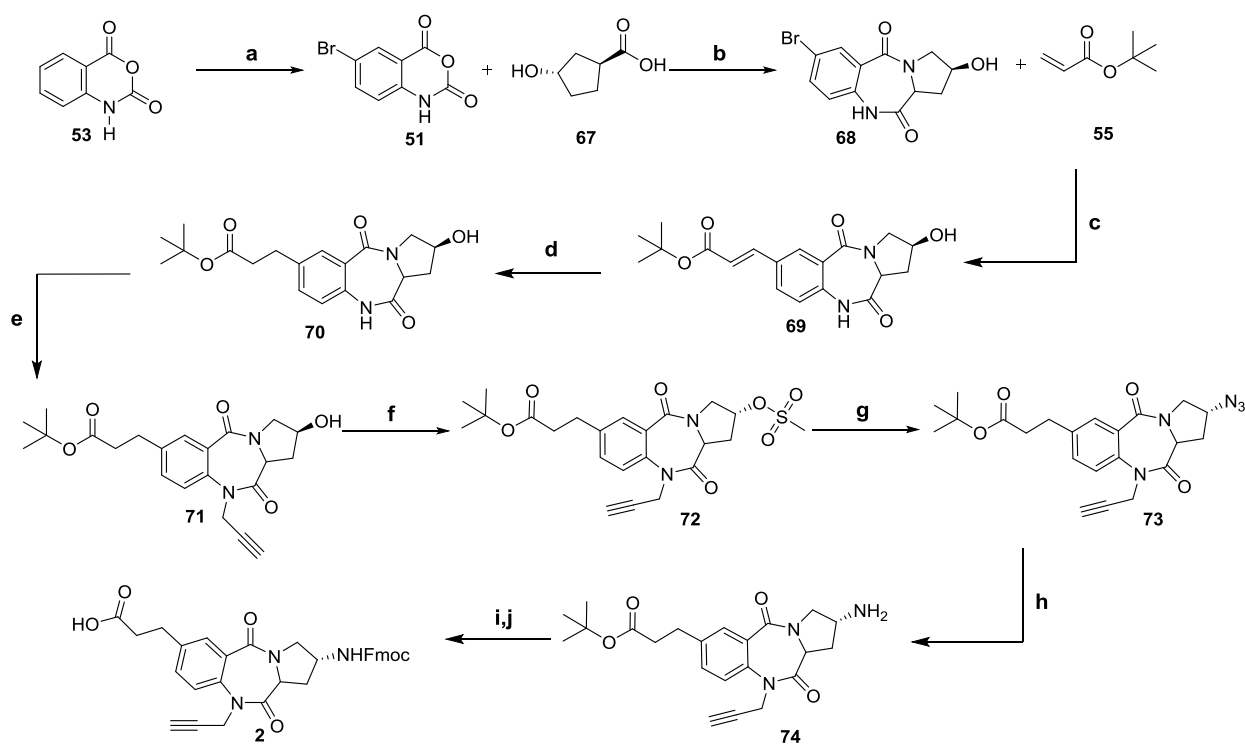
Figure II-6: Retrosynthetic analysis of compound **2**.

II-2.2.1. Synthetic route to trifunctionalized 1,4-benzodiazepine-2,5-dione scaffold (**2**)

The 1,4-benzodiazepine-2,5-dione **2** was prepared according to a procedure of 10 steps outlined in Scheme 2-. The synthesis route started with the bromination of isatoic anhydride **53** which yielded the 5-bromoisatoic anhydride **51**. 5-bromoisatoic anhydride **51** with trans-4-hydroxy-L-proline **67** in dimethyl sulfoxide (DMSO) afforded benzodiazepinedione derivative **68**, which after a Heck reaction with *tert*-butyl acrylate **55** in acetonitrile gave compound **69** as a mixture of diastereomers due to partial epimerization. The diastereoisomers were separated by flash chromatography. Subsequent steps were carried out with (*E*)-isomer **69**. The Heck reaction was followed by hydrogenation of the double bond to remove the α,β -unsaturated Michael acceptor function.

Synthesis of a DNA-encoded library based on a benzodiazepine scaffold.

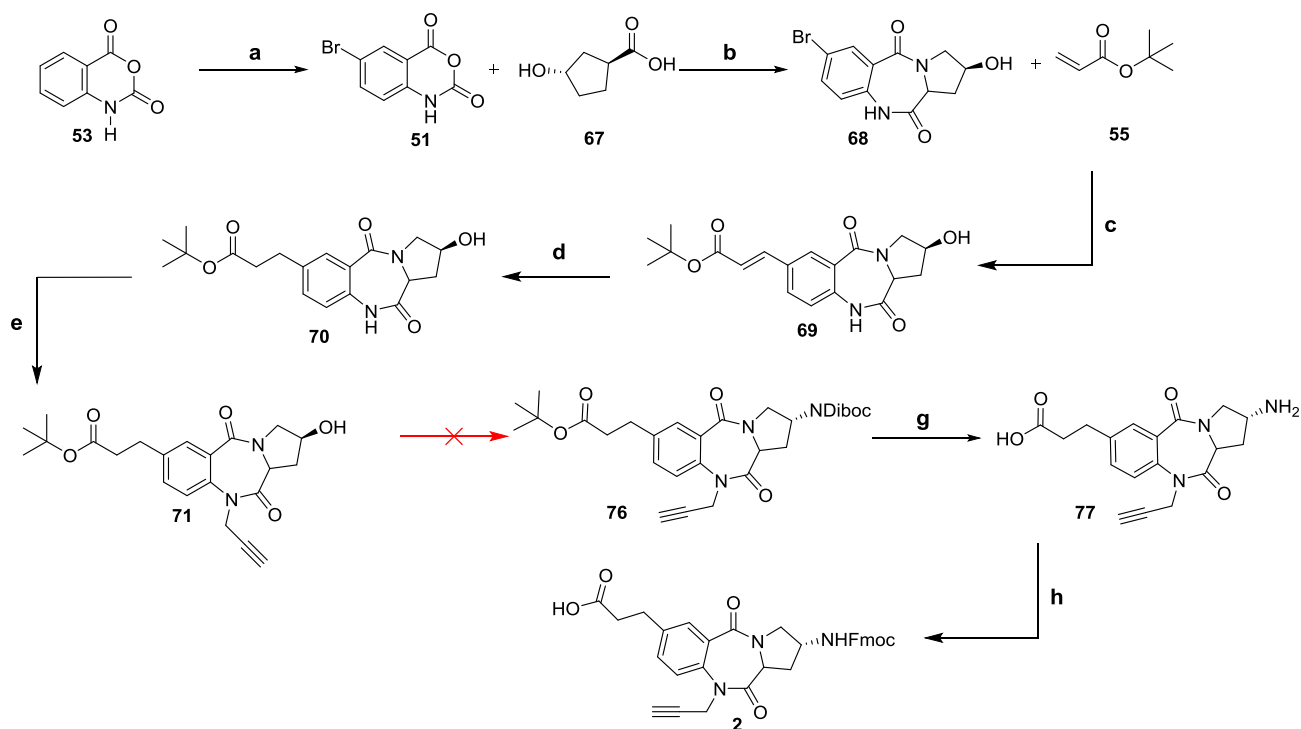
Alkylation of the aryl amine **70** with propargyl bromide introduced the terminal alkyne. Transformation of alcohol **71** into the corresponding mesylate compound **72** was performed by use of methanesulfonyl chloride in dry DCM. The crude mesylate **72** was then converted without purification to the azide **73** in presence of DMF and sodium azide. The azide **73** was reduced to the amine **74** by the Staudinger reaction. The phosphine imine-forming reaction was followed by hydrolysis to produce a phosphine oxide and the desired amine **74**. Finally, removal of the protective groups with trifluoroacetic acid (TFA) followed by Fmoc protection of the primary amine **74** afforded the target molecule **2**.



Scheme II-3: Synthesis of 1,4-benzodiazepine-2,5-dione **2**. Reagents and conditions: a) bromine in H₂O at 50 °C for 1 h; b) trans-4-hydroxy-L-proline **67**, DMSO, 120 °C for 5 h, then rt for 18 h; c) Pd(OAc)₂, P(O-tol)₃, tert-butyl acrylate **55**, Et₃N, dry CH₃CN, 100 °C, 18 h; d) Pd/C, dry MeOH, rt, 18 h; e) Cs₂CO₃, propargyl bromide, dry DMF, 60 °C, 18 h; f) CH₃SO₂Cl, Pyridine, 0° C-rt, 24 h; g) NaN₃, DMF, 50 °C, 72 h; h) Staudinger reaction, PPh₃, dry THF, rt, 24 h; i) TFA in dry CH₂Cl₂, rt, 19 h; j) Fmoc-Osu, NaHCO₃, H₂O, dry 1,4-dioxane, rt, 18 h.

In the initial synthetic route leading to compound **2** (Scheme II-4), the Mitsunobu reaction was envisaged to convert the hydroxyl group of compound **71** to the protected amine **76** with inversion of configuration. The nucleophile employed should be acidic, since one of the reagents (DIAD or DEAD) must be protonated during the reaction to prevent side reactions.

The triphenylphosphine was first combined with DEAD, then to DIAD to generate a phosphonium intermediate that was supposed to bind to the alcohol oxygen, activating it as a leaving group. The formation of the amine failed in this case. The principal mass observed after LC-MS analysis was 279 g/mol corresponding to triphenylphosphine oxide. The Mitsunobu reaction was in this case time consuming and didn't yield the target product. The increase of equivalents of DIAD/DEAD brought no significant changes, also higher temperature, as well as the increase of the amine did not lead to improvement. This leads to the investigation of a different approach for the conversion of the alcohol to the amine. The formation of an azide followed by a reduction through the two steps Staudinger reaction showed to be the best way to access the target compound. These optimizations led to the new synthetic pathway outlined in Scheme II-3.



Scheme II-4: Initial synthesis of 1,4-benzodiazepine-2,5-dione **2**. Reagents and conditions: a) bromine in H₂O at 50 °C for 1 h; b) trans-4-hydroxy-L-proline **67**, DMSO, 120°C for 5 h, then rt for 18 h; c) Pd(OAc)₂, P(O-tol)₃, *tert*-butyl acrylate **55**, Et₃N, dry CH₃CN, 100 °C, 18 h; d) Pd/C, dry MeOH, rt, 18 h; e) Cs₂CO₃, propargyl bromide, dry DMF, 60 °C, 18 h; f) Mitsunobu reaction Ph₃P/ DEAD or DIAD, THF, rt, 16 h; g) TFA in dry CH₂Cl₂, rt, 19 h; h) Fmoc-Osu, NaHCO₃, H₂O, dry 1,4-dioxane, rt, 18 h.

The synthetic protocol for the preparation of the proline-derived 1,4-benzodiazepine-2,5-dione **2** using trans-4-hydroxy-L-proline was successfully established. The scaffold **2** was synthesized in satisfying amount (420 mg) for test reactions on DNA and library synthesis.

II-2.3. Synthesis of trifunctionalized tetrahydrobenzodiazepine (3)

The objective was to synthesize the trifunctionalized tetrahydrobenzodiazepine **3**. The chemical structure of this 1,4-benzodiazepine consists of a core benzene ring fused to a seven-member 1,4-diazepine ring. It shows a semi-planar platform and distinguishes itself from the Scaffold 1 and 2 by the absence of the oxo groups which are reduced. In order to obtain the benzodiazepine **3** that mimic the core structure of the clinical trial candidates **36** and **78** (figure II-7), and at the same time fulfill the requirements for library synthesis, a versatile route needed to be developed.

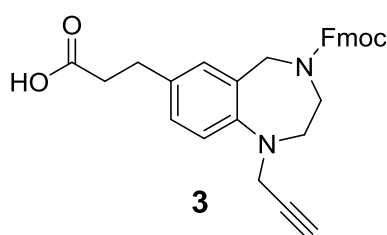


Figure II-7: Structure of the trifunctionalized tetrahydrobenzodiazepine scaffold **3**.

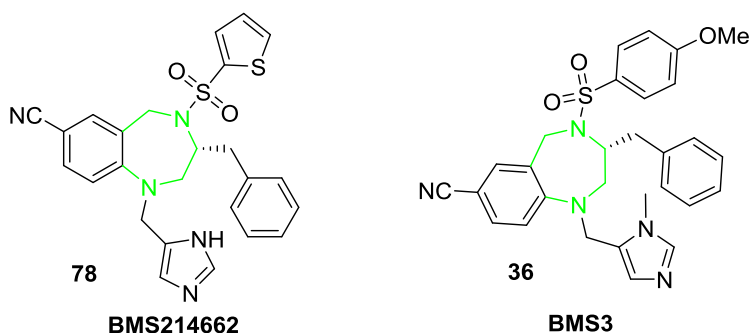
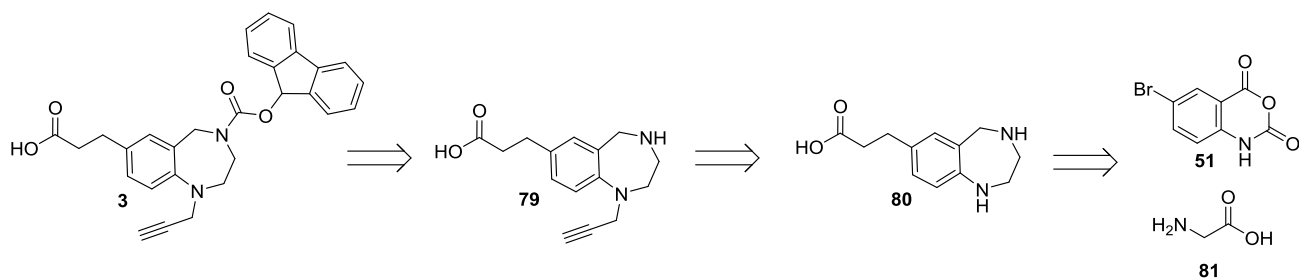


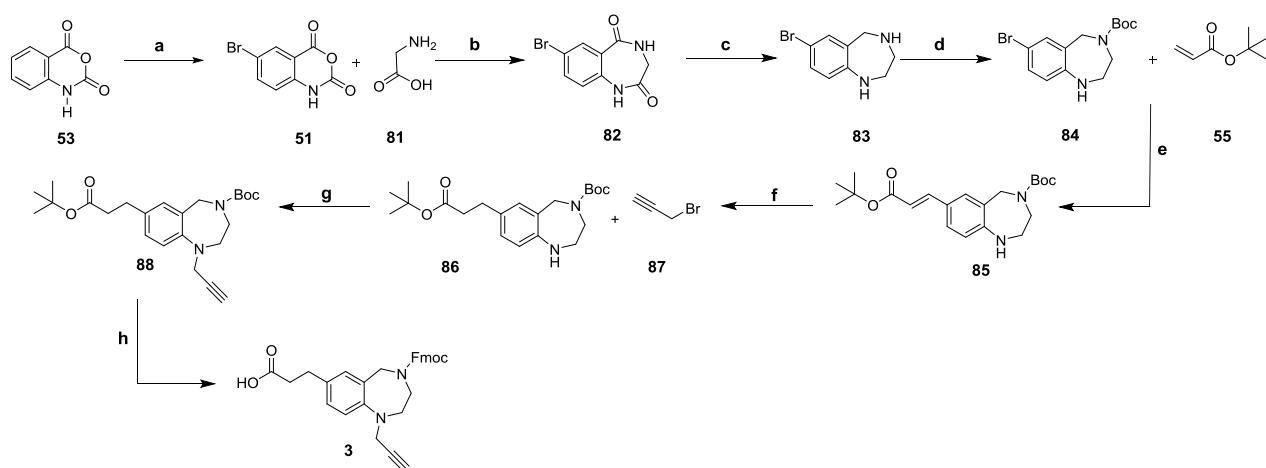
Figure II-8: Tetrahydrobenzodiazepine in clinical trials **36**, **78**.¹¹⁹

Compound (**3**) could be obtained from building block **80**, which could be constructed from 5-bromoisatoic anhydride **51** and glycine **81**.



Scheme II-5: Retrosynthetic analysis of compound **3**.

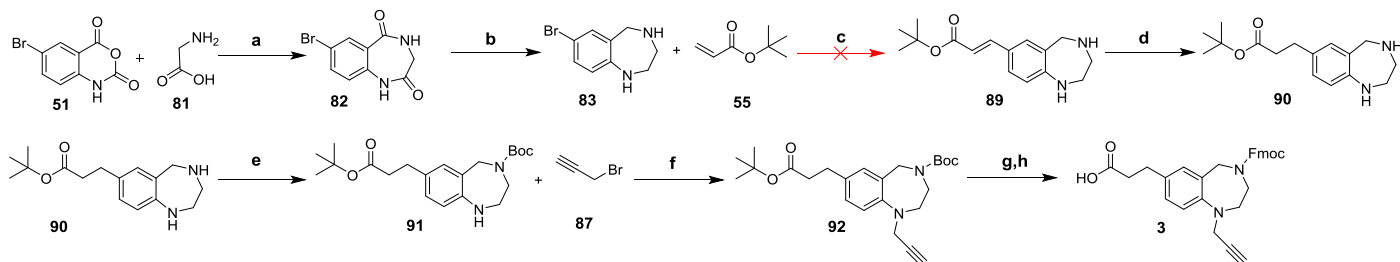
The tetrahydrobenzodiazepine **3** was synthesized through 8 steps. The first step consisted in the bromination of isatoic anhydride **53** (Scheme II-6), followed by the condensation of 5-bromoisatoic anhydride **51** with glycine **81** to access the core benzodiazepinedione **82**. The reduction of the carbonyl groups of **82** yielded the tetrahydrobenzodiazepine structure **96**. The N-Boc protection of the secondary amine of the tetrahydrobenzodiazepine structure **83** will facilitate the introduction of the carboxylic function by Heck reaction with acrylic acid *tert*-butylester. This yielded compound **85**. The prior protection of the amino function will avoid the N-alkylation by the acrylic acid ester. Compound **86** was obtained after hydrogenation of the double bond removing the α,β -unsaturated Michael acceptor function. Compound **88** was accessed through alkylation of the aryl amine with propargyl bromide **87**. Finally, Boc and Fmoc protective groups were removed yielding the target compound **3**.



Scheme II-6: Synthesis of tetrahydrobenzodiazepine **3**. Reagents and conditions: a) bromine in H₂O at 50 °C for 1 h; b) glycine **81** in H₂O, Et₃N, rt, 4 h; c) BH₃ in dry THF, reflux, 18 h; d) “diboc” in dry MeOH, rt, 18 h; e) Pd(OAc)₂, P(O-tol)₃, *tert*-butyl acrylate **55**, Et₃N, dry CH₃CN, 100 °C, 18 h; f) Pd/C, dry MeOH, rt, 18 h; g) Cs₂CO₃, propargyl bromide **87**, dry DMF, 60 °C, 18 h; h) TFA in dry CH₂Cl₂, rt, 19 h; i) Fmoc-Osu, NaHCO₃, H₂O, dry 1,4-dioxane, rt, 18 h.

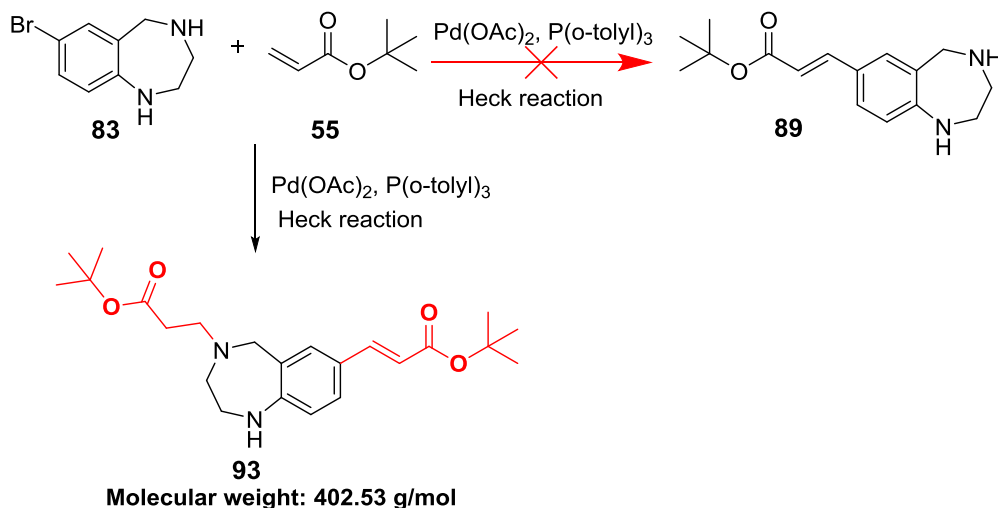
Initially, the synthesis of the tetrahydrobenzodiazepine **3** was planned according to the synthetic route depicted in scheme II-7. This synthetic route however presented two issues that prompted to modify the initial synthetic pathway leading to the target compound **3**.

Synthesis of a DNA-encoded library based on a benzodiazepine scaffold.



Scheme II-7: Initial synthesis of tetrahydrobenzodiazepine **3**. Reagents and conditions: a) glycine **81** in H₂O, Et₃N, rt, 4 h; b) BH₃ in dry THF, reflux, 18 h; c) Pd(OAc)₂, P(O-tol)₃, *tert*-butyl acrylate **55**, Et₃N, dry CH₃CN, 100 °C, 18 h; d) Pd/C, dry MeOH, rt, 18 h; e) “diboc” in dry MeOH, rt, 18 h; f) Cs₂CO₃, propargyl bromide **87**, dry DMF, 60 °C, 18 h; g) TFA in dry CH₂Cl₂, rt, 19 h; h) Fmoc-Osu, NaHCO₃, H₂O, dry 1,4-dioxane, rt, 18 h.

The first issue concerned the Heck reaction product **89** (see step 4 in the experimental section). During the Heck reaction, an addition to the secondary amine was observed, yielding the addition product as major product (Scheme II-8). The analysis of the obtained compound by LC-MS always showed a mass of 402.53 g/mol corresponding to the addition product **93**. For this reason, the synthesis pathway had to be changed in order to get the product **89**. In the new synthetic pathway, the amine at the 4th position was first protected (compound **84**) with the Boc group prior to the Heck reaction to avoid an addition at this position and the Heck reaction was performed.

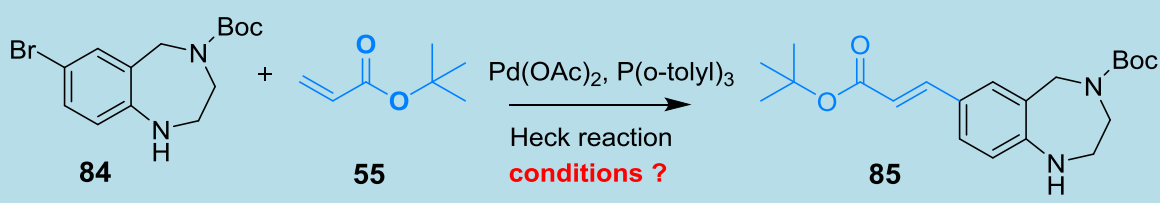


Scheme II-8: Michael addition on the secondary amine **83**.

Many test reactions were performed to find out the suitable temperature at which the Heck reaction could be performed without the deprotection of the protective Boc group.

The Heck reaction was firstly performed with the Boc protected compound **84** at 60 °C for 5 hours using 0.1 equivalent of the catalyst palladium acetate in combination with 3 equivalents *tert*-Butyl acrylate **55**. These reaction conditions yielded 2% of the desired compound **85** (Table 4, **entry 1**). Increasing the reaction time to 8 hours raised the yield of the desired compound **85** to 8 % (Table 4, **entry 2**). Longer reaction time, as well as an increase of the temperature (60°C to 100 °C), catalyst amount (0.1 to 0.2) yielded 45 % of the desired product (Table 4, **entry 3**). The use of 6 eq of the reagent instead of 3 eq raised the yield of compound **85** to 51 % (Table 4, **entry 4**). These last reaction conditions were used to perform the Heck reaction.

Table 4: Optimization of the Heck reaction (preparation of compound 98).



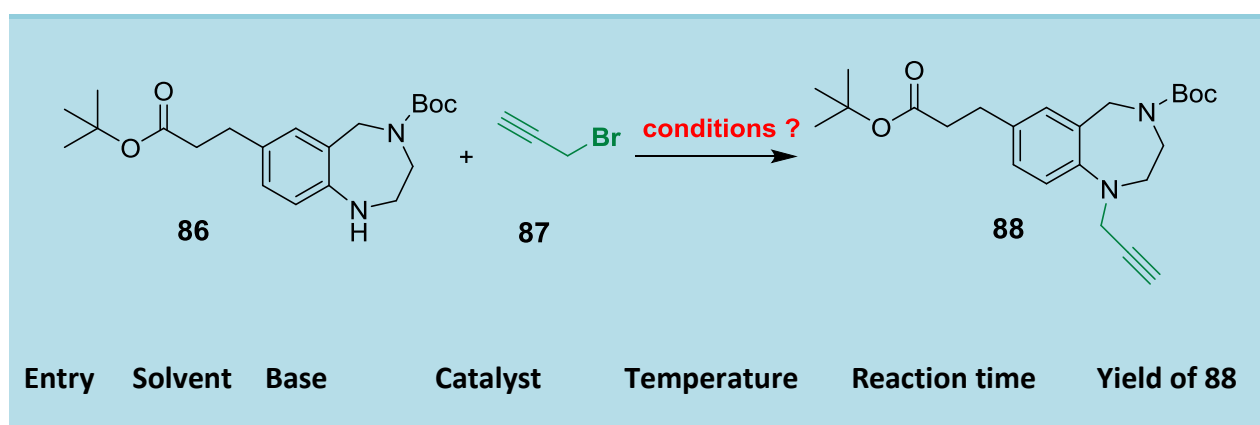
Entry	Reaction time	Temperature	Catalyst amount	TBA amount	Yield of compound 85
1	5h	60°C	0.1	3 eq	2%
2	8h	60°C	0.1	3 eq	8%
3	18h	100°C	0.2	3 eq	45%
4	18h	100°C	0.2	6 eq	51%

The second issue concerns alkylation of the secondary amine **86** in high yields for the successful synthesis of a sufficient amount of the final scaffold **3** (see step 5 in the experimental section). The literature describes alkylation reactions of secondary amines in dimethylformamide (DMF) in presence of NaH at room temperature or under slight heating.¹²⁰ However, these conditions did not lead to satisfying yields even in presence of a catalyst (Table 5, **entries 1** and **2**), and also higher temperature did not lead to improvement. Actually higher temperatures lead to degradation of starting material.

Synthesis of a DNA-encoded library based on a benzodiazepine scaffold.

In addition NaH causes many side products making the reaction purification problematic. Changing the base sodium hydride (NaH), to 2 eq of potassium carbonate (K_2CO_3) and 0,1 eq of the catalyst tetrabutylammonium iodide (TBAI) at room temperature gave no significant changes (Table 5, **entry 3**). The yield slightly increased by the use of 2 eq of cesium carbonate (Cs_2CO_3) as base at room temperature for 18 hours in the absence of any catalyst. This is probably because Cs_2CO_3 dissolves better in DMF than K_2CO_3 (Table 5, **entry 4**). Finally, by increasing temperature (Table 5, **entry 5**) the yield of compound **88** could be further optimized. All these improvements led to the new synthetic pathway depicted above in scheme II-17 that encompasses eight reaction steps.

Table 4: N-alkylation optimization of the secondary amine **86**.



Entry	Solvent	Base	Catalyst	Temperature	Reaction time	Yield of 88
1	DMF	3 eq NaH	-	25 °C	18 h	25 %
2	DMF	3 eq NaH	0,1 eq TBAI	60 °C	18 h	32 %
3	DMF	2 eq K_2CO_3	0,1 eq TBAI	25 °C	18 h	35 %
4	DMF	2 eq Cs_2CO_3	-	25 °C	18 h	42 %
5	DMF	2 eq Cs_2CO_3	-	60 °C	18 h	52 %

The synthesis of tetrahydrobenzodiazepine (**3**) was successfully established in sufficient amount (380 mg) for test reactions on DNA and for library synthesis. In summary, two approaches have been developed for the synthesis of three BDZs based on three different amino acids: **D-serine methyl ester 52** (1,4-benzodiazepine-2,5-dione **1**), **trans-4-Hydroxy-L-proline 67** (tricyclic benzodiazepinedione scaffold **2**) and **glycine 81** (tetrahydrobenzodiazepine **3**).

The presented syntheses demonstrate the variety of possibilities for functionalization of the benzodiazepine core structure and, as a consequence, present 1,4-benzodiazepine-2,5-dione and 1,4-benzodiazepine as interesting scaffolds for library synthesis. The three scaffolds were successfully synthesized and characterized by 1-dimensional proton (^1H) or carbon-13 NMR (^{13}C) spectroscopy. They were synthesized in sufficient amount for test reactions on DNA and library synthesis (scaffold amount was comprised between 100 and 500 mg) (Figure II-8).



scaffold 1: 290 mg



scaffold 2: 420 mg



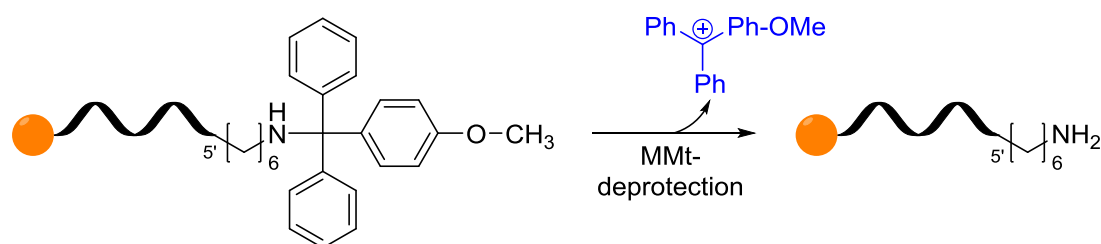
scaffold 3: 380 mg

Figure II-9: Appearance of the three synthesized benzodiazepines **1**, **2**, **3**.

II-2.4. Optimization of the amide synthesis for the coupling of scaffolds to 5'-aminolinker modified DNA

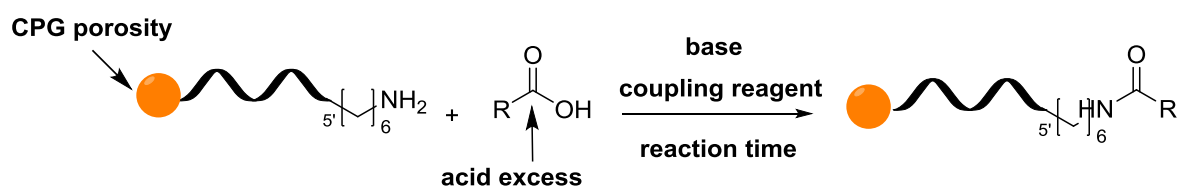
The construction of DNA-encoded chemical libraries is facilitated through the feasibility of performing chemical reactions at high yields and with good purities, by the use of synthetic protocols which work well in water and preserve the integrity of DNA. Up to now, most libraries reported in the literature have relied on amide bond forming reactions.⁷⁻⁸ The step following the synthesis of the privileged scaffolds depicted above is the combinatorial synthesis of DNA-Encoded Libraries based on these compounds. The library construction started with the attachment of a functionalized scaffold to a DNA-PEG linker conjugate synthesized by amide synthesis. Therefore, we needed to investigate reaction conditions (CPG porosity, coupling reagent, solvent, base, reaction time and carboxylic acid excess) that provide the target DNA conjugate in satisfying yield and preserve the integrity of the DNA. The different experiments were performed on solid support (CPG) carrying a 32-mer 5'-C6-amino linker DNA: **GTC TTG CCG AAT TCC GTC CTG ACG TAT GCT GG**.

II-2.4.1. Deprotection of the MMT group



Scheme II-9: Deprotection of the MMT group

For the deprotection of the MMT group, the DNA on solid support was distributed in small columns with filters and treated with 3 % trichloroacetic acid (TCA) in *dry* DCM (3 x 200 μ L) followed by washing with each 3 x 200 μ L of 1 % TEA in MeCN, DMF, MeOH, MeCN and DCM. Afterwards the oligonucleotides were dried *in vacuo* for about 15 minutes, and used for the optimization of amide coupling conditions.



Scheme II-10: Optimization of the amide synthesis.

The first investigation was the determination of the suitable CPG porosity for the attachment of small molecules to the 5'-aminolinker-DNA. For this purpose, 500 Å CPG, 1000 Å CPG, and 2000 Å CPG were tested. Different amide syntheses were performed with different Fmoc protected amino acids.^[67a] The reaction with DNA on 500 Å CPG yielded no product at all. Amide syntheses performed with 2000 Å CPG gave low yields; while syntheses performed on 1000 Å CPG gave satisfying yields (yields were comprised between 50 and 95 %). The second investigation was to find out the suitable coupling reagent for the successful amide synthesis. The different experiments performed on 1000 Å CPG DNA are summarized in table 2. Firstly, different coupling reagents were tested: HATU, PyBOP, HBTU as well as a combination of EDC with HOBT. Also different bases were combined: *N,N*-diisopropylethylamine (DIPEA) and *N*-methylmorpholine (NMM) with a variation of the equivalents of coupling reagent (0.25 eq, 0.5 eq and 1 eq). Also, different excesses of the base were investigated: 1 eq, 2.5 eq.

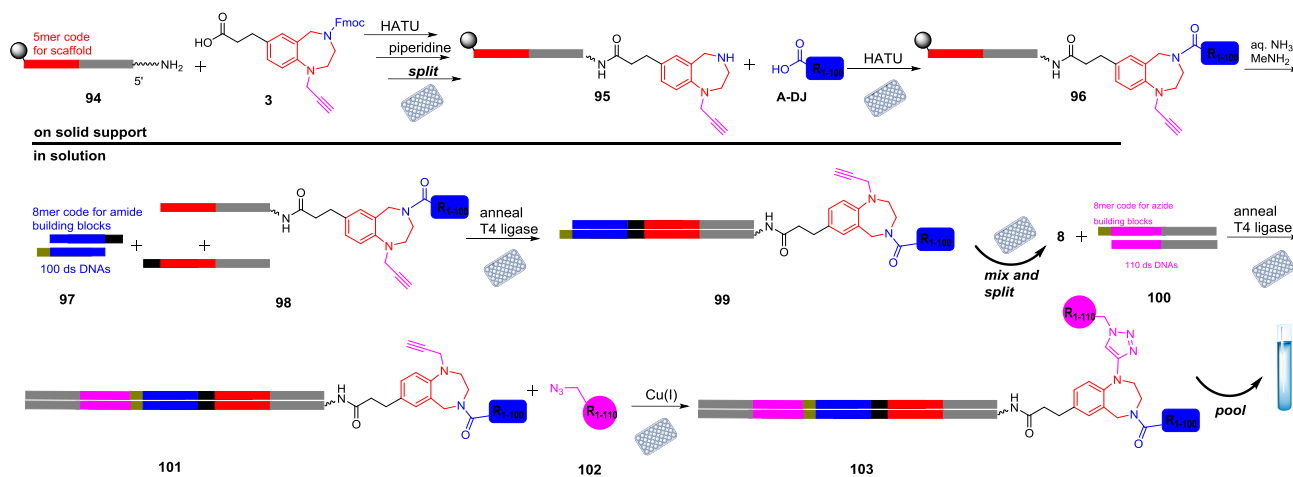
The best yields for the amides synthesis were obtained by combining HATU as coupling reagent (0.5 eq/1 eq) with DIPEA (2.5 eq) as base in *dry* DMF (Table 6, **entry 2**). Different reaction times and different temperatures were also investigated. The best results were obtained after 2-4 hours at room temperature (Table 6, **entry 2, 3**). Performing the coupling overnight led to DNA degradation (Table 6, **entry 1**). The use of 25 eq of the coupling reagent HATU led to lower yields (Table 6, **entry 4**). The same result was observed by the use of 50 eq of PyBOP (Table 6, **entry 6**). Comparing the coupling agents HATU and PyBOP revealed HATU to be the superior reagent for amide synthesis on DNA on solid support. The yields of reactions with PyBOP were slightly lower than those with HATU (Table 6, **entry 5**). In conclusion, every amide synthesis was performed using HATU and DIPEA in DMF at room temperature for 2 to 4 hours. After each coupling, the DNA conjugates were removed from CPG by AMA treatment at room temperature for 4 hours and analysed by HPLC. After each coupling step, remaining free amines were blocked with a capping solution prior to a following coupling. HPLC data of these amide syntheses showed a shift when compared to the starting material suggesting that reaction took place (see *experimental section*). The successful coupling of Fmoc protected amino acids and scaffolds to the 5'-aminolinker DNA enabled library synthesis. In parallel a protocol for the removal of the Fmoc group on the solid support was established (20% piperidine in DMF for 15 min). The successful removal of the Fmoc group on the solid support was also the starting point for Fmoc solid phase peptide synthesis described in chapter 3.

Table 5: Optimization of the amide synthesis on solid phase.

Entry	Carboxylic acid	Coupling reagent	Base	Solvent	Time
1	100 eq	HATU (100 eq)	DIPEA (2.5 eq)	DMF	Overnight
2	100 eq	HATU (100 eq)	DIPEA (2.5 eq)	DMF	2h/4h
3	100 eq	HATU (50 eq)	DIPEA (2.5 eq)	DMF	2h/4h
4	100 eq	HATU (25 eq s)	DIPEA (2.5 eq)	DMF	2h
5	100 eq	PyBOP (100 eq)	DIPEA (2.5 eq)	DMF	2h
6	100 excess	PyBOP (50 eq)	DIPEA (2.5 eq)	DMF	2h

II-2.5. Synthesis of a DNA-encoded library based on tetrahydrobenzodiazepine (3)

DNA-encoding enables the synthesis and selection of large small molecules libraries. Here, I outline the general synthetic route for the stepwise coupling of coding oligonucleotides fragments to appropriate organic molecules throughout individual reaction steps. The procedure was illustrated in the synthesis of a DNA-encoded chemical library containing 9588 compounds named “DEL9588” (DNA Encoded Library 9588). The library was yielded using a split-and-pool methodology characterized by the following steps: (i) Attachment of functionalized scaffolds to amino modified synthetic DNA fragments; (ii) deprotection of the *N*-Fmoc-amino group of the scaffolds; (iii) amide bond formation reaction with selected carboxylic acids; (iv) pool and split ; (v) encoding of the carboxylic acids used in the previous step as well as encoding of the different halides used in the next step by DNA ligation, yielding the fully encoded compounds in a double-stranded DNA format; (vi) Introduction of the second set of building blocks by click chemistry. The DNA conjugates were purified using HPLC and analyzed by mass spectrometry and MALDI-TOF.



Scheme II-11: Synthesis of the DEL by a two stage synthesis strategy. a) HATU; b) piperidine in DMF; c) aq. NH_3 /aq. MeNH_2 ; d) encoding by T4 DNA ligation; e) NaN_3 in DMF; f) DEAE sepharose, Cu (I), TBTA, Na-ascorbate.

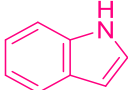
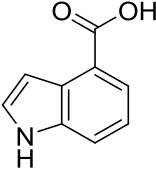
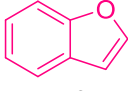
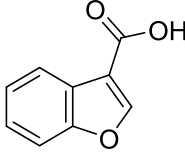
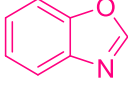
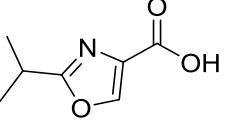
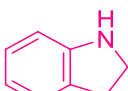
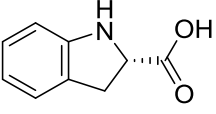
Contrary to privileged scaffolds, there are various chemical motifs with undesirable properties for drug design; for instance, chemically reactive groups or moieties,^[65] which tend to confer poor water solubility. Chemical building blocks were selected by chemoinformatics and unsuitable chemical substructures were eliminated in the selection of building blocks for the synthesis of the DNA-encoded library described in this thesis. The scheme II-13 described the synthesis of the DNA-encoded library **103**.

The benzodiazepine scaffold **3** was attached to a fully protected, solid phase-bound 5'-aminolinker modified single-strand DNA sequence **94**. Fmoc deprotection yielded the DNA conjugate **95**. The next step was the coupling of 114 carboxylic acid building blocks A-DJ (Table 15, *experimental part*) to the DNA-conjugate **95**. Afterwards, the conjugates **96A-DJ** were deprotected from the solid phase and each DNA conjugate was purified by HPLC. The DNA-encoded library has been synthesized in this way from a uniform and purified set of DNA-conjugates. The purity of DNA conjugates is crucial for library quality. The encoding of the purified conjugates **96A-DJ** was achieved by DNA ligation with T4 ligase by Kathrin Jung making use of two consecutive double-stranded DNA sequences reactions. The final library was obtained after introduction of a second set of azide building blocks **1-104** (Table 16, *experimental part*) by copper(I)-catalyzed alkyne-azide cycloaddition on DEAE sepharose.

Library synthesis started with the synthesis of a DNA-PEG linker conjugate by coupling of a protected amino-PEG-carboxylic acid to a solid phase-coupled 5'-C6-aminolinker modified 23mer DNA using HATU as coupling reagent. The amide synthesis was performed on 1 μ mol-scale. The 23mer DNA contains the primer and scaffold code (see *experimental part*). First of all, the protective group was cleaved; after amide synthesis, unreacted amines were capped with acetic acid anhydride before going in the next reaction. Two spacers MMT-amino-PEG (8)-linker and Fmoc-protected amino-PEG (4)-linker were used for library synthesis. They were found to be appropriate as spacer to avoid interactions between small molecule and target protein (Table 15, *experimental part*). The investigation of reaction conditions for the deprotection of the Fmoc group revealed the use of piperidine in DMF to be the best combination. However, long reaction time caused a transamidation reaction yielding the formation of a lipophilic side product as described by other groups.¹²¹ Decreasing the cleavage of the Fmoc group to 5 minutes annulled the formation of this side product. The next step consisted in the coupling of the scaffold **3** by amide synthesis to the DNA-PEG-linker conjugate (from the MMT-protected PEG-linker) on the 1 μ mol scale. After washing steps, unreacted amines were capped again, and the Fmoc-groups of the scaffolds were cleaved with piperidine/DMF. The carboxylic acid building blocks A-DJ were then coupled to the DNA-PEG conjugates (Table 15, *experimental part*). Structures of carboxylic acid building blocks as well as halides building blocks are diverse. They include aliphatic, cyclic aliphatic, aromatic and heteroaromatic structures and some of the substituted heteroatoms displayed functional groups. They are depicted in Table 7.

The investigation of these building blocks (Tables 15 and 16, *experimental part*) revealed that a certain number of these moieties are present in bioactive compounds. A few are enumerated here: the benzopyrazole AI (Table 15, *experimental part*) for instance is a structure binding to the kinase CDK2;¹²² Dihydrouracil BK was recognized as binding motif for members of the family of PARP enzymes after a screening of a DNA-encoded library,¹²³ and the uracil structure BN was found to bind to the activated complement factor C3d.¹²⁴ In addition, the indole structure R, the benzofuran AC and the benzimidazole 75 (Table 16, *experimental part*) occur in numerous bioactive compounds.

Table 6: Selected privileged structures and examples of chemical building blocks used in the synthesis of the DEL9588.

Privileged scaffold	Chemical building block
 Indole	 BBa = R, BF
 Benzofuran	 BBa = AC
 Benzoxazole	 BBa = CS
 Indoline	 BBa = DE

The introduction of the first set of building blocks started with the coupling of 20 carboxylic acids to the DNA conjugate **95**. The solid phase (400 nmol, ca. 16 mg) was distributed in 20 nmol aliquots into a 96 well plate by dissolving the solid phase in DMF and distributing the suspension (see *experimental part*). Afterwards, 20 carboxylic acid building blocks (Table 15, *experimental part*) were attached to the DNA scaffold conjugate by amide synthesis using HATU as coupling reagent. The solid support-bound DNA conjugates were distributed in a 96 well filter plate. Washing steps with different solvents followed. The solid support-bound DNA conjugates were then dry under vacuum, deprotected and cleaved from the CPG with a mixture of aq. ammonia/methylamine on the filter plate.

This filter plate was linked to a receiver plate and sealed so that the solid support would remain in the filter plate and the DNA conjugates were collected in the receiver plate. The conjugates were finally purified by ion pair reverse phase HPLC. The analysis of HPLC chromatograms after coupling and purification of two sets of carboxylic acid building blocks, i.e. 40 carboxylic acids, revealed that ten (**H, W, AC, AD, AE, CW, CX, CY, CZ, DA**, Table 15) of these building blocks did not yield the desired amide products. The initial 23 mer DNA sequence was changed to a shorter 14 mer, for the Fmoc-peptide chemistry has been described to be more efficient on shorter solid support-bound DNA oligonucleotides.¹²¹ With this shorter DNA-sequence, the amide coupling of the unreacted carboxylic acid building blocks mentioned above was repeated. Five of the ten carboxylic acid building blocks (**H, W, AC, AD, AE**) yielded the target amide. However, the absence of a new corresponding to the amide coupling product could either be explained by the low reactivity of some carboxylic acids or by the tendency of the amide bond to hydrolytic deprotection in the cleavage step. The amide coupling with the total set of 114 carboxylic acid building blocks A-DJ (Table 15, *experimental part*) to the DNA-conjugate **95** generated 94 products for the DNA-benzodiazepine conjugate. The so obtained DNA conjugates **96A-CV** displayed a purity of more than 95% after HPLC analysis, and the identity of the products was confirmed by MALDI MS analysis (Table 15, *experimental part*). Most building blocks caused a shift to longer retention times. However, a number of building blocks, for instance the primary amide **H** and the dihydrouracil **BL** gave a shift to shorter retention times. In this manner, HPLC purification and isolation of the DNA-conjugates **96A-CV** produced a set of 94 DNA-small molecule conjugates for encoding and combinatorial library synthesis.

The obtained DNA-small molecule conjugates was then encoded by T4 ligation of 5'-phosphorylated dsDNA containing overhangs by Kathrin Jung. The ligation of dsDNAs took place either for 4 hours or overnight with equal efficiency.¹²⁵ The next step consisted in the introduction of a second set of building blocks by Cu (I) catalyzed azide-alkyne cycloaddition (CuAAC). Previous library synthesis, the reactivity of the in situ synthesized azides with a DNA-alkyne conjugate **101** (Table 16, *experimental part*) was investigated by Mateja Klika Skopic. For this purpose, each 400 pmol of DNA-alkyne conjugate **101** was immobilized on DEAE sepharose in 96 well plates.⁵⁵ On the other hand, the azides were synthesized by substitution of the halides with NaN_3 /tetrabutylammonium iodide.³⁹ The cycloaddition reaction was then effected according to a previously established procedure with a 2500 fold excess of the azide at 45°C for 16 h.⁹²

The resin was firstly washed with different solvents to remove the excess of reactants and reagents, followed by the elution of the triazole. To avoid DNA degradation by Cu-ion contaminants and preserve the integrity of DNA conjugates during the storage, a 1 N buffered aqueous solution of EDTA was applied to remove Cu-ion contaminants. The conjugates were analyzed by MALDI MS (Table 16, *experimental part*). 104 azides were used for test experiments (Table 16) and 102 gave the triazole products. The figure VII-1 presents a statistical analysis of the results. 82 azides showed complete conversion and compound identities were confirmed by MALDI MS. In case of azide building blocks **A10**, **A16**, **A38**, **A44**, **A51**, **A53**, **A62**, **A68**, **A75**, **A76**, **A77**, **A101**, **A102** and **A104** incomplete conversion was observed, however the yield for these building blocks reached and surpassed 50% as approximated by MALDI MS analysis, which is satisfying for library synthesis.⁴⁵ The second set of building blocks is principally composed of heterocyclic structures. From the 104 tested azides, only azides **A6** and **A87** did not give the desired triazoles. Lastly, the synthesis of DEL library on solid-phase appears as a particularly interesting concept, for it allows to perform chemical reactions to completion and to enlarge the area of DNA-compatible chemical reactions to water-free conditions.

Chapter III: Development of a modular solid phase synthesis strategy for DNA-macrocycle conjugates

III-1. Introduction

Through the development of numerous new synthesis strategies and application possibilities, peptides have in recent years increasingly taken an important place in medical research. While only a few years ago they played a subordinate role in drug development, due to their low stability and poor cellular uptake, peptides are now an important component of pharmaceutical research and development. Peptides are characterized by excellent selectivity, high activity and low toxicity, as well as a predictable metabolism.¹²⁶⁻¹²⁸ This has led in recent years to more than 140 peptide-based drug candidates in clinical trials.^{127, 129} A key role in the recovery of peptides in medical research was played by progress in peptide synthesis, as well as the development of chemical modifications leading to improvement in the availability and biological activity of peptides. The well-established and modular synthesis of peptides provides rapid access to extensive libraries with structural diversity since, in addition to the 20 proteinogenic amino acids, a large number of different building blocks are accessible for peptide synthesis. Several peptides are highly selective and effective binders of receptors or ion channels and operate as hormones or neurotransmitters. Due to their flexible backbone, peptides are very adaptable and they are able to cover large-area binding sites. According to the "Lock-and-key" principle, in which both binding partners incur conformational changes during binding, the biologically active conformation is particularly crucial for high affinity binding (Figure III-1A).¹³⁰⁻¹³³ The preorganization of the structure by restriction of the conformational free space is advantageous in many ways. By stabilizing the biologically active conformation (a "match"), the affinity of the peptide to the target protein is increased, as the required entropic energy is reduced during binding (Figure III-1b). Furthermore, a stable conformation affords a higher stability against proteases and in some cases even the cellular uptake of peptides.¹³⁰⁻¹³⁵ The stabilization of a non-binding conformation (a "mismatch") instead leads to a loss of binding affinity.

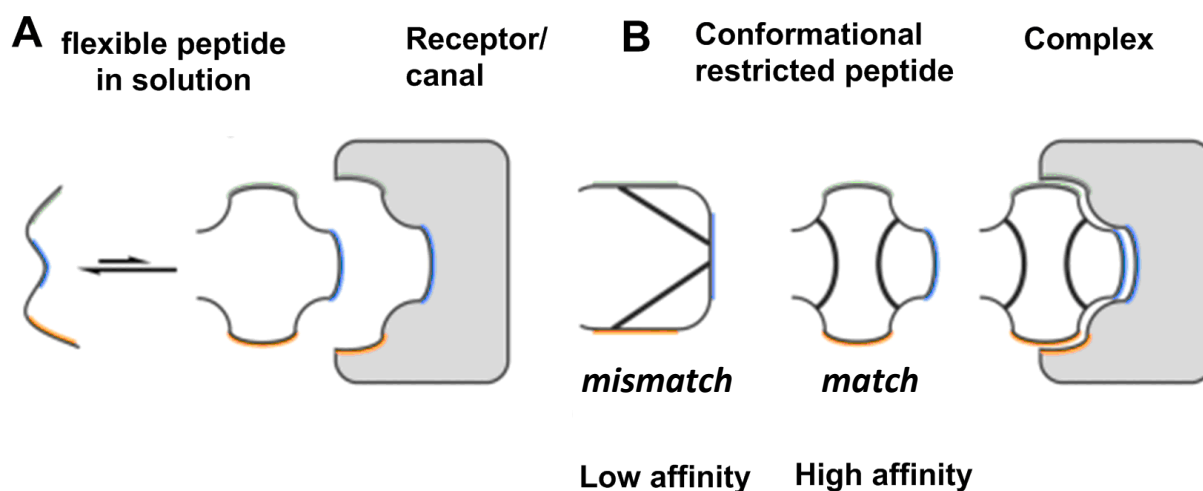


Figure III-1: (a) flexible peptides in solution; (b) conformational restricted peptide. As a result of the reinvigorated interest in peptides, many peptide-based binders for numerous receptors and target proteins have been developed in advanced preclinical or clinical trials.¹²⁷ Therefore, the current challenge in peptide-based drug discovery is to develop new strategies and methods which transcend the application spectrum of peptides. These include multifunctional peptides that can cross cell membranes or peptide-drug conjugates.¹²⁷ For this reason, peptides are increasingly used for addressing extensive and low-profile protein-protein interactions (PPIs).^{128, 134-138} In particular, mimicking of secondary structure elements enables the identification of numerous peptide-based modulators for challenging PPIs, which are not accessible with the classical methods of medical chemistry (so-called “undruggable” PPIs). The figure was adapted from Patgiri et al, 2008.¹³⁹

PPIs are more and more in the focus of drug research but represent difficult targets for low molecular-weight compounds, since they rarely possess addressable binding pockets.¹³⁹ Peptide-based approaches for addressing PPIs are mainly based on the imitation of binding epitopes, which are involved in the interaction of proteins. Additionally, peptides display some beneficial characteristics. Their flexibility allows for very useful adaptation to large surfaces and their simple modular structure increases their structural diversity through the incorporation of numerous amino acids and unnatural building blocks. In addition, peptide-based active substances often possess a high efficiency or activity as well as excellent selectivity. Furthermore, peptides and their degradation products are biocompatible and non-toxic.^[113] Peptides have been increasingly successfully used to modulate PPIs due to their ability to mimic important secondary structures motifs (Figure III-2).^{134, 138-142} However, the isolation of the peptide sequence of the interacting structural element is insufficient since short peptide sequences in free solution lose their order and embrace flexible and disorganized conformations.¹⁴²⁻¹⁴³

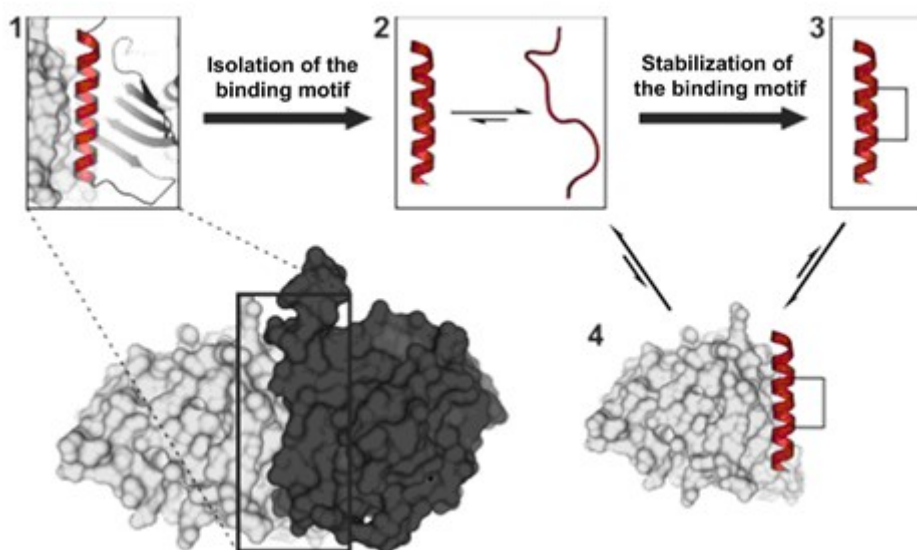


Figure III-2: (1) The major components of the binding energy of PPIs are provided by discrete binding motifs. (2) Without a stabilizing tertiary structure, the isolated peptide sequences lose their arrangement and can bind to the target protein only with reduced affinity. (3) Through chemical modifications, the disorganized peptide sequences become stabilized in their biologically active conformation. (4) Thereby, the chemical modification allows the binding of the peptide sequence to the target protein and the Inhibition of the protein-protein interaction. The figure was adapted from Cromm et al., 2015.^{141, 144}

Such peptidomimetics resulting from chemical modification can be organized into four classes (A-D) based on their similarity with the unmodified original peptide.¹³⁵ The class-A-peptidomimetics consists mainly of α -amino acids and contains only a small number of modified amino acids. The backbone and the side chains are largely similar to those of the original peptide. The peptidomimetics of class B include foldamers, in which the peptide backbone has been additional modified by the incorporation of β -amino acids or the use of peptoids. In contrast, for peptidomimetics of classes C and D, the peptide backbone is completely substituted. While class-C-mimetics orientate their side chains analog to the bioactive conformation of the peptide precursor, the structural reference to the original peptide for class D-mimetics no longer exists. They simply imitate the biological mode of action of the bioactive original peptide.

III-2. Macrocyclic peptides

In the search for inhibitors, which can cover a large area of interaction, macrocycles have proved to be increasingly successful in the inhibition of PPIs. Cyclic peptides and peptidomimetics represent a promising class of substances.¹⁴⁵⁻¹⁴⁸ Macrocyclic, peptide-based natural products like vancomycin or cyclosporin A gain attention as a result of their unique mechanisms of action and are successful and established drugs on the market. Therefore, numerous macrocycles are generated by the cyclization of linear peptide sequences. The macrocyclization of linear peptides restricts their conformational free space and leads thus to increased proteolytic stability, reduced polarity, increased affinity and selectivity, and improved overall drug-like properties.^{124, 146-147, 149-152} However, macrocyclization of peptides represents in many cases a synthetic challenge.^{146, 153} In particular, the 3D conformation of peptides represents a central determining factor of the biological activity and bioavailability of peptides.¹³² A peptide is able to exhaust its full potential only when this is stabilized in its optimal 3D structure (*match*) (Figure III-1).

The cyclization of peptides can be summarized as taking place in four different geometries (Figure III-3): (1) Head-to-tail (C-terminus to N-terminus), (2) Head-to-side chain, (3) Side-chain-to-tail, (4) Side-chain-to-side-chain. Therefore, several different chemical reactions are available for the ring closure reaction. Ring closures are widely used and are obtained by the formation of a lactam,¹⁵⁴ lactones,¹⁵⁵ or a disulfide.¹⁵⁶ Meanwhile many other reactions such as the Ugi reaction¹⁵⁷ or azide alkyne-click chemistry¹⁵⁸ are applied for the synthesis of macrocyclic peptides. Additionally, the Pd-catalyzed C-H activation of Trp for the coupling with Iodo-aryl amino acids can be used as a ring closure reaction.¹⁵⁹ In particular, ring-closing olefin metathesis (RCM) is becoming more and more attractive for the synthesis of peptides macrocycles.¹⁶⁰ RCM is principally used in the synthesis of macrocyclic α -helical peptides.^{141, 161} The analogous ring-closing alkyne metathesis (RCAM)¹⁶² has already been successfully used for the cyclization of linear peptide sequences.¹⁶³⁻¹⁶⁵

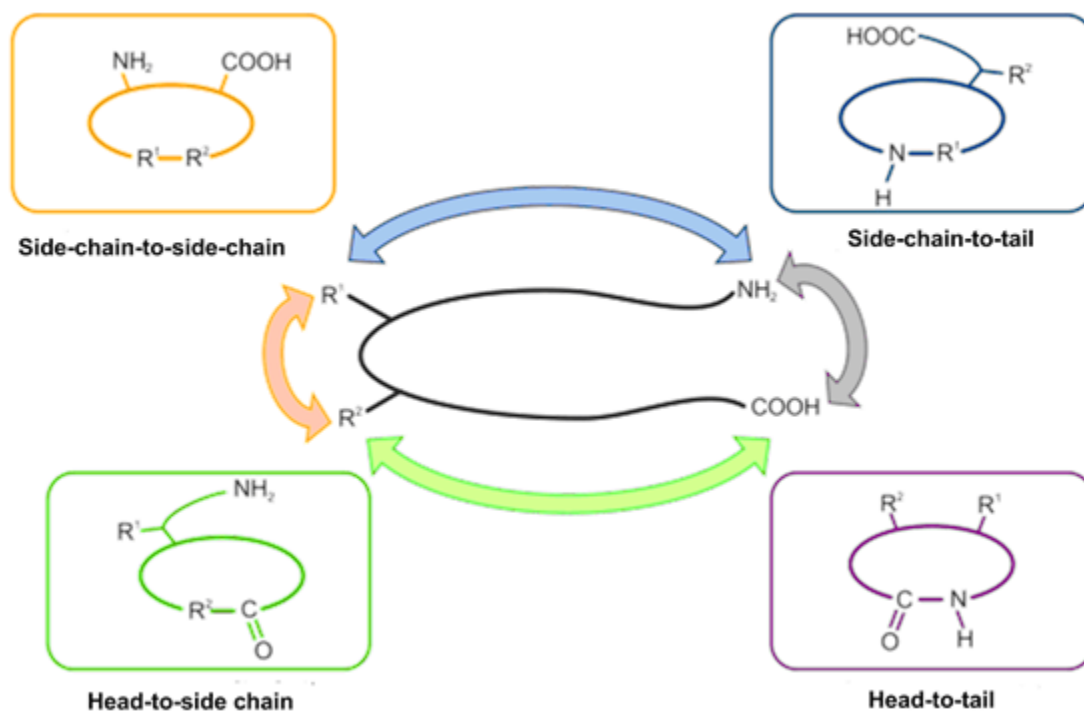


Figure III-3: Schematic representation of the four different cyclization geometries of peptide sequences. The figure was adapted from White et al., 2011.¹⁶⁶

III-3. Macrocyclization of peptides by ruthenium catalyzed ring-closing metathesis

Peptidomimetic macrocycles are considered as viable modulators due to their ability to mimic natural peptides or specific secondary structure in proteins.^{154-158, 166} Generating large numbers of peptidomimetic-based macrocycles using DNA-encoded technology can be valuable for the discovery of molecules that modulate PPI's. This chapter describes the first application of the oligoThymidine initiated DNA-Encoded Chemistry ("TiDEC") strategy for DEL synthesis developed in our group, in which the selection of a 5'-aminolinker modified hexathymidine oligonucleotide hexT, enables the utilization of a surprisingly broad spectrum of catalysts and reaction conditions.

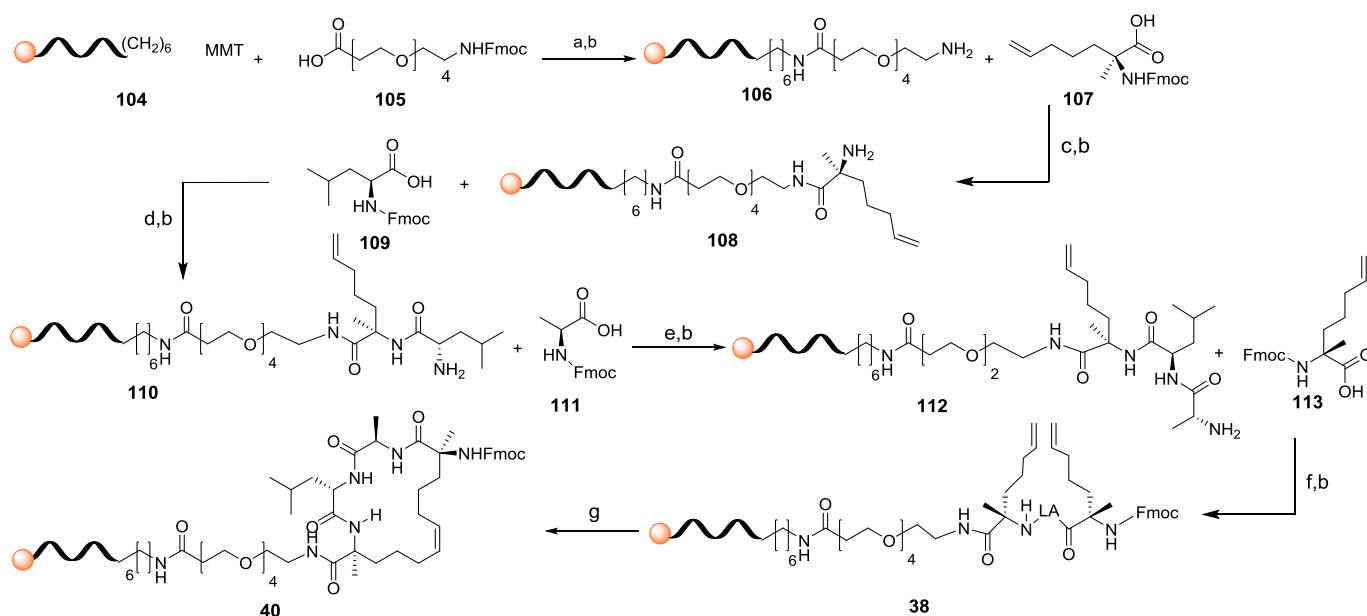
The hexT (**5'-AM TTT TTT**) was bound to controlled pore glass (CPG) solid support which enabled us to use organic solvents in the synthesis of target molecules, and facilitated purification of the DNA-conjugates. The task was the macrocyclization of a peptide sequence on DNA based on this new TiDEC strategy. For the introduction of the hydrocarbon cross-link, two selected non-natural amino acids were coupled together with two natural amino acids. Synthesis of the peptide sequence was performed by Fmoc-based SPPS (Figure III-4).

Subsequent cross-linking via RCM used first generation Grubbs catalyst **39** with solid-phase-bound protected peptides provided the macrocyclic peptide **40**. This project was initiated by the successful experiments showing the compatibility of the Fmoc chemistry with the DNA on solid support. This project was done in collaboration with the group of Prof. Dr. Tom Grossmann who provided us modified amino acids and the Grubbs catalyst for the ring-closing metathesis.

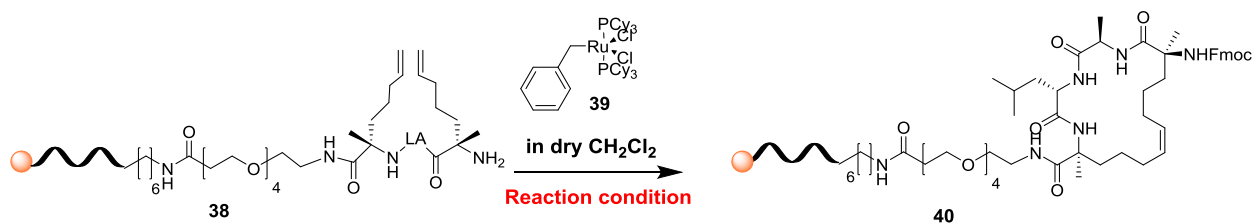
In line with our aim to synthesize peptide sequences that can be cyclized by ring-closing metathesis and the possibility in turn to introduce different sets of building blocks for the construction of a DEL, four amino acids were chosen: (*S*)-*N*-alpha-(9-Fluorenylmethyloxycarbonyl)-4-pentenylalanine (**107**), *N*-alpha-(9-Fluorenylmethyloxycarbonyl)-L-leucine (**109**), *N*-alpha-(9-Fluorenylmethyloxycarbonyl)-L-alanine (**111**) and (*R*)-*N*-alpha-(9-Fluorenylmethyloxycarbonyl)-4-pentenylalanine (**113**). Different absolute configurations of the unnatural amino acids were chosen and a polyethylene glycol spacer (Fmoc-NH-PEG (4)-COOH **105**) was introduced between DNA and the first amino acid aa₁. For an optimal spacing and installation of the hydrophobic bridges via olefin RCM, the *S*-configured amino acid (**107**) was introduced first, and the *R*-configured amino acid (**113**) was introduced at last. The metathesis reaction provided high yields under mild conditions and tolerates the functional groups present in the protected peptide. Synthesis of the peptide sequence on DNA on solid support was performed by Fmoc-based SPPS (Figure III-4). Subsequent cross-linking via RCM used first-generation Grubbs catalyst (Bis(tricyclohexylphosphine) benzylidene ruthenium (IV) dichloride **39**) with solid-phase-bound, protected peptides.

The synthesis of the peptide-DNA conjugate **38** was performed as described in Figure III-1. The polyethylene glycol spacer **105** was firstly coupled to a fully protected, solid phase-bound 5'-aminolinker modified single-strand DNA sequence 5'-AM-TTTTTT **104**, and the Fmoc-group was removed yielding **106**. The peptide-DNA conjugates were synthesized by repetitive Fmoc-solid-phase peptide synthesis (SPPS) according to the general procedure 1 described in the experimental section making use in the following order of the (*S*)-*N*-9-Fluorenylmethyloxycarbonyl)-alpha-4-pentenylalanine (**104**), *N*-alpha-(9-Fluorenylmethyloxycarbonyl)-L-leucine (**105**), *N*-alpha-(9-Fluorenylmethyloxycarbonyl)-L-alanine (**106**) and (*R*)-*N*-9-Fluorenylmethyloxycarbonyl)-alpha-4-pentenylalanine (**107**). The Fmoc-deprotection was performed according to the general procedure 2.

Aliquots of all conjugates **38**, **40**, **106**, **108**, **110**, **112** were cleaved from the solid phase, purified to a single peak by ion pair reversed phase HPLC and analyzed by analytical HPLC and MALDI-TOF. Fmoc-deprotection was performed with 20% piperidine in DMF for 15 min. After each coupling step, remaining free amines were blocked with capping solution. Crosslinking of the peptide **38** was finally performed by ring closing metathesis (RCM) in a solution of the Grubbs 1st generation catalyst (Bis (tricyclohexylphosphine) benzylidene ruthenium (IV) dichloride **39**, 4.9 mM, 4 mg) dissolved in 1 mL of dry dichloromethane (DCM) three times for 2 h at room temperature (100 μ l of the solution were used for each coupling) (Figure III-4). After each coupling, the DNA was washed with DCM (3 x 200 μ l), the CPG was then dried *in vacuo* for 5 min and the next coupling was performed. Subsequent to the three couplings, the DNA was washed successively with DMF (3 x 200 μ l), MeCN (3 x 200 μ l), MeOH (3 x 200 μ l) and DCM (3 x 200 μ l). The filter carrying CPGs was then dried *in vacuo* for 15 min. The DNA-macrocycle conjugate was deprotected from the CPG by treatment with AMA, purified by HPLC and analyzed by MALDI-TOF.

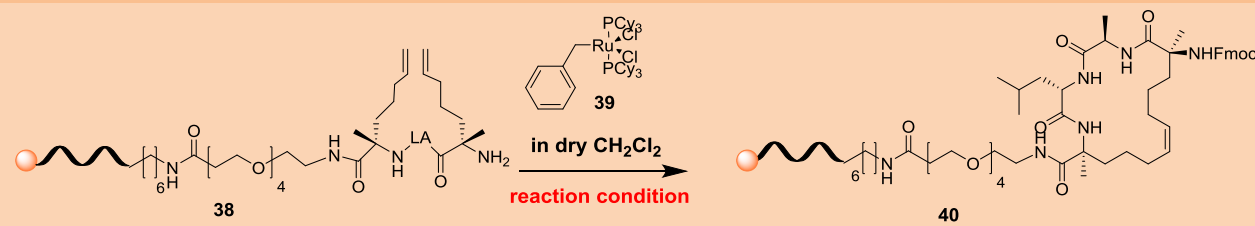


Scheme III-1: Synthetic route leading to the macrocycle **40**. a) Fmoc-NH-PEG(4)-COOH **105**, HATU, DIPEA in DMF, 2 h, rt; b) Fmoc deprotection, 20 % piperidine in DMF, 15 Min; c) Fmoc-S₅-OH **107**, HATU, DIPEA in DMF, 2 h, rt; d) Fmoc-Leu-OH **109**, HATU, DIPEA in DMF, 2 h, rt; e) Fmoc-Ala-OH **111**, HATU, DIPEA in DMF, 2 h, rt; f) Fmoc-R₅-OH **113**, HATU, DIPEA in DMF, 2 h, rt; g) RCM: Grubbs-(I)-catalyst **39**, DCM, three coupling for 2 h.



Scheme III-2: Ring closing metathesis using the Grubbs-(I)-catalyst **39**.

Different conditions were investigated for the ring closing metathesis (Figure III-5) in order to find out the suitable concentration of the catalyst and the appropriate reaction time. Therefore, the ring closing was performed in a solution of the catalyst **39** in a concentration of (1.2 mM) dissolved in dry dichloromethane (DCM) three times for 2 h at room temperature (100 μl of the solution were used for each coupling). Analysis of the crude compound by HPLC showed that the cyclization hasn't reached completion (Table 8, **entry 1**). Raising the concentration of the catalyst from 1.2 mM to 2.4 mM slightly increased the yield of the desired macrocyclic conjugate (Table 8, **entry 2**). An additional increase of the yield was observed when the ring closing was performed under the conditions described in Table 8, **entry 3**. Interestingly, raising the concentration of the catalyst from 3.6 mM to 4.9 mM and performing the macrocyclization three times for 30 minutes led to 50 % of the macrocycle **40** (Table 8, **entry 4**). Keeping the concentration to 4.9 mM, the reaction time was investigated. Longer reaction times gave better yield (Table 8, **entries 5 and 6**). The metathesis reaction provided high yields under mild conditions and tolerates the functional groups present in the protected peptide **38**.

Table 7: Optimization of the ring closing metathesis.


entry	concentration of the Grubbs catalyst 39	reaction time (for each coupling)	yield of the macrocyclic conjugate 40
1	1 mg/mL (1.2 mM)	2 hours	20%
2	2 mg/mL (2.4 mM)	2 hours	30 %
3	3 mg/mL (3.6 mM)	2 hours	40 %
4	4 mg/mL (4.9 mM)	0.5 hours	50 %
5	4 mg/mL (4.9 mM)	1 hours	65 %
6	4 mg/mL (4.9 mM)	2 hours	95 %

This chapter has proved the synthesis on solid-phase to be particularly attractive, as it allows to drive chemical reactions to completion and to expand the scope of DNA-compatible chemical reactions to water-free conditions. It has also shown that the macrocyclization of peptides by RCM is tolerated by a hexT adapter oligonucleotide, opening the door to the synthesis of a DNA-encoded chemical library based on macrocycles that can be screened against targets involved in protein-protein interactions.

Chapter IV: Development of an acid catalysed approach to an oligo-thymidine initiated DNA-encoded β -carboline library

IV-1. Introduction

The indole moiety **114** (Figure IV-1) is an aromatic heterocyclic component of biologically active natural products, and its exploration has been a major focus of several research groups for generations.¹⁶⁷⁻¹⁷⁰

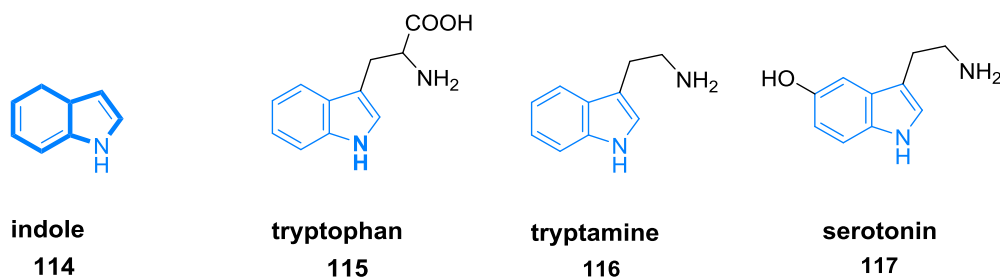


Figure IV-1: Naturally occurring indole structures **114-117**.

The indole scaffold has become an important core structure or intermediate in the synthesis of a broad number of biologically active natural and synthetic products including simple to complex indole derived scaffolds, serving as ligands for a diverse array of therapeutic targets. To these belong anti-inflammatories, phosphodiesterase inhibitors, 5-hydroxytryptamine receptor agonists and antagonists, cannabinoid receptors agonists, HMG-CoA reductase inhibitors and many more.¹⁷¹⁻

¹⁷³ An important class of indole derived scaffolds is the tetrahydro- β -carboline ring system contained in harmicine, indoloquinoline and related analogues as depicted in figure IV-2: yohimbine **118**, vallesiachotamine **119** and 10-hydroxyauginine **120**, the harmicine alkaloid¹⁷⁴⁻¹⁷⁵ **121** has the tetrahydro- β -carboline ring fused to a 5-membered ring as the core scaffold as depicted in figure IV-2. Compound collections built upon these complex scaffolds might afford diversely bioactive small molecules as drug and probe candidates.

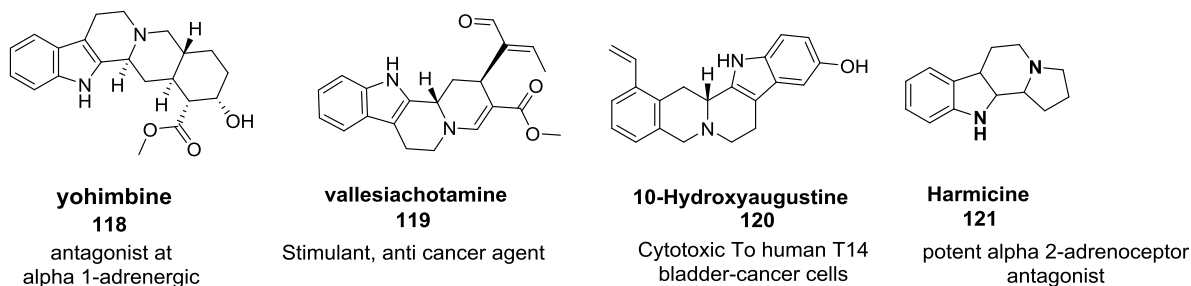


Figure IV-2: Structures of representative polycyclic indole alkaloids.

IV-2. β -carboline

Carbolines are pyridoindoles, which depending on the position of the nitrogen in the pyridine ring are divided in α -, β -, γ or δ -carbolines (Figure IV-3). α -, γ - and δ -carbolines are less studied than β -Carboline because of their minor occurrence in the nature.¹⁷⁶ It is known that some α -carboline derivatives possess an antitumor effect, whereas antihistamine and nootropic properties are attributed to the γ -carboline.¹⁷⁷⁻¹⁷⁸ β -Carbolines are formed endogenously e.g. from tryptophan or serotonin by condensation with aldehydes or α -keto acids. This reaction takes place in human tissues predominantly in the central nervous system, in plant and fungal tissues, and in food processing by heating or fermentation processes. In addition, β -carbolines are found in tobacco smoke.¹⁷⁹⁻¹⁸¹

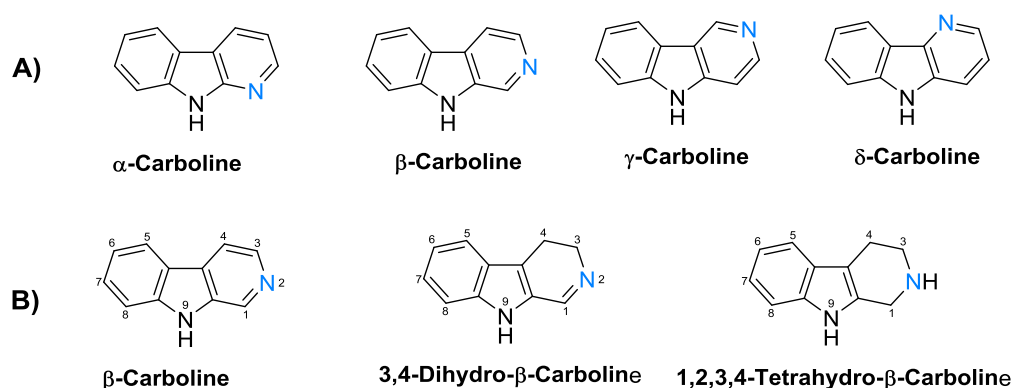


Figure IV-3: carbolines isomere.

A diversity of β -carboline derivatives has been isolated from the plant and animal kingdom with diverse biological activities ranging from antineoplastic (tubulin binding),^{170, 182} anticonvulsive, hypnotic and anxiolytic (benzodiazepine receptor ligands)¹⁸³⁻¹⁸⁴ antiviral,¹⁸⁵ antimicrobial¹⁸⁶ as well as topoisomerase-II inhibition¹⁸³ to inhibition of cGMP-dependent processes.¹⁸⁶ The relevance of β -carboline derivatives as potential therapeutics is underpinned by the β -carboline based compounds Tadalafil and Abecarnil that are used clinically for the treatment of erectile dysfunction and CNS disorders.

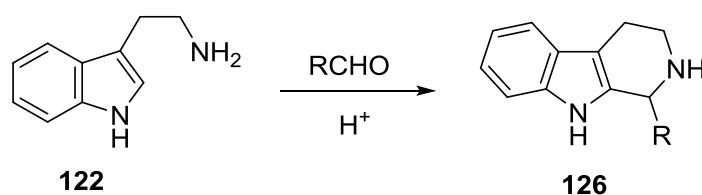


Figure IV-4: General scheme of PSR.

The 1-methyl- β -carboline-containing maracuja is for instance used in Brazil as drug against spasms and as a sedative.¹⁸⁵ Besides, various plants containing β -carbolines such as the liana (*Banisteria caapi*) are used as hallucinogenic substances by South American Indians.¹⁸⁶

IV-3. The Pictet-Spengler reaction

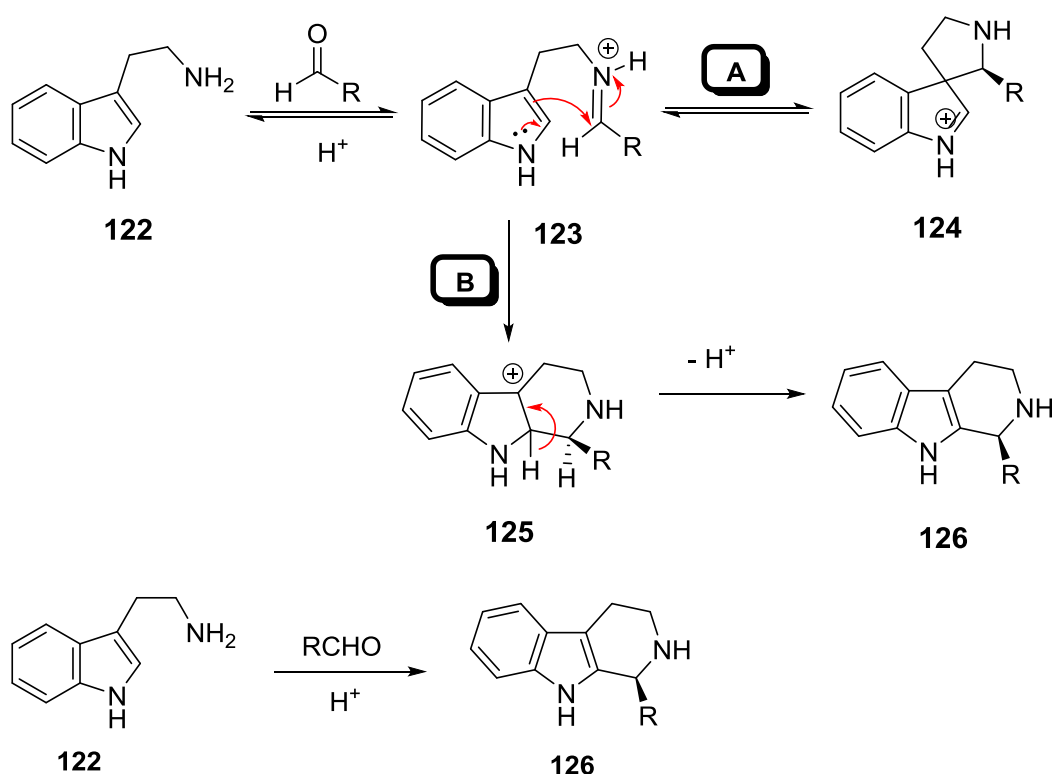
During the last ~100 years, the Pictet-Spengler reaction has represented one of the most widely used reactions in heterocyclic chemistry.¹⁸⁷⁻¹⁸⁹ The reaction was discovered in 1911 by Ame Pictet and Theodor Spengler. The Pictet-Spengler (PS) reaction, in its simplest form, consists of the condensation of a beta-arylethylamine with a ketone or an aldehyde to yield a tetrahydroisoquinoline or tetrahydro- β -carboline. This reaction is applicable to phenylethylamine, pyrrole-ethylamines,¹⁹⁰ imidazole-ethylamines¹⁹¹ and tryptamines¹⁹², where the latter gives rise to formation of tetrahydro- β -carbolines.^{188, 193} The mechanism consists in a two-step method including Brønsted acid catalysed condensation of an aliphatic amine attached to a sufficiently reactive aromatic nucleus with aldehyde.^{187, 189}

A linkage occurs at the carbon *ortho* between an activated arene ring and an aliphatic amine group to the activated carbon nucleophile in order to enable ring closure via 6-*endo* cyclization. The first step represents the formation of the imine obtained by condensation of an aldehyde and an amine. This imine is then activated under acidic conditions yielding an iminium ion in situ (Scheme IV-1). The iminium ion that generates the C-C bond is composed of two components: a nucleophilic partner coming from the activated heterocyclic ring that contributes one of the carbons as nucleophile, whereas the second carbon is contributed from the aldehydic carbon implicated in the iminium ion formation and playing the role of an electrophilic partner.

The mechanism of the Pictet-Spengler condensation leading to β -carbolines is still subject of discussion, but it is well established to follow either of two likely pathways A or B.^{190, 194} The condensation of tryptamine **122** with an aldehyde yields the iminium salt **123**. At this step attack of either the indole 2 or 3-position has been postulated. In pathway **A** initial attack of the more electrophilic 3-position gives the intermediate spiro-indolenine **124**, which rearranges to **123** and subsequent protonation affords the tetrahydro- β -carboline intermediate **125**. This intermediate can also be obtained by direct attack of the indole 2-position to the imine carbon atom via pathway **B**.

Development of an acid catalysed approach to an oligo-thymidine initiated DNA-encoded β -carboline library

According to Baldwin's rules for cyclisations,¹⁹⁵⁻¹⁹⁶ the attack of the indole 3-position would involve the disfavoured 5-endo-trig pathway, whereas attack of the 2-position proceeds through the favoured 6-endo-trig pathway.^{195, 197} The mechanism of the reaction is initiated by the protonation of the carbonyl oxygen by the acid which is consequently attacked by the amine of the tryptamine **122**. Proton transfer steps and loss of water molecule emanates in a protonated imine intermediate **123**, which then undergoes a 6-endo-trig cyclization reaction followed by a final deprotonation restoring the aromaticity of the indole ring and resulting in the tetrahydro- β -carboline product **126** (Scheme IV-1).

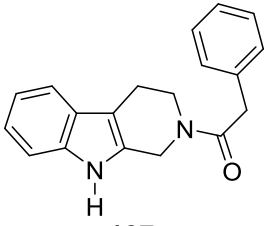
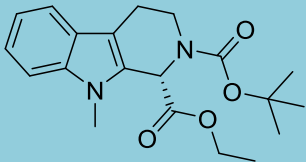
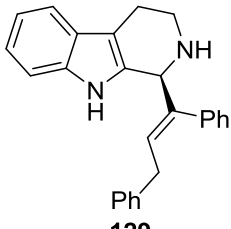
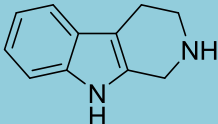
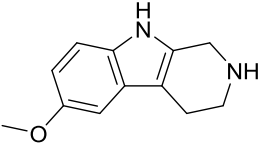


Scheme IV-1: Typical acid-catalyzed Pictet-Spengler reaction.¹⁸⁹

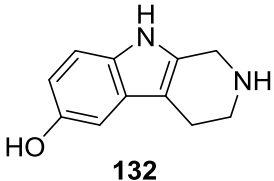
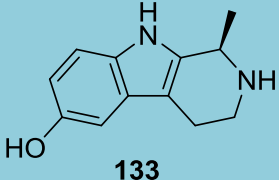
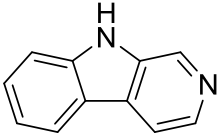
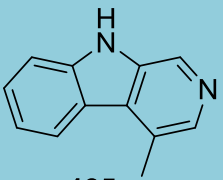
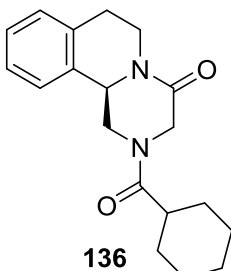
Many reviews described how extensively this reaction has been modified to different variants and promoted by many catalysts. Among the different classes of catalyst used over the years in the PS reaction, we count protic acids like: TFA,¹⁹⁸ HCl,¹⁹² H₂SO₄,¹⁹³ Lewis acids like boron trifluoride diethyl etherate (BF₃·Et₂O),¹⁹⁹ AuCl₃/AgOTf²⁰⁰ and recently lanthanide triflates²⁰¹⁻²⁰³ emerged also as efficient Lewis acid catalysts; halosilanes like chlorotrimethylsilane²⁰⁴ and molecular iodine²⁰⁵ have been also employed as suitable condensation agents for the PS reaction.

The PS condensation has also been subjected to different conditions from standard room temperature and heating conditions, to being subjected to microwave¹⁸² and ultrasound treatment in pursuance of better conversions and higher yields. Conforming to the retrosynthetic analysis, the Pictet-Spengler cyclization was found to be a powerful method for the synthesis of the tetrahydro- β -carboline core.¹⁸⁸ In recent years, a large number of targets of β -carbolines, particularly in the central nervous system have been identified. Some examples are shown in table 9. Tetrahydro- and dihydro-derivatives can be converted carbolines by enzymatic oxidation into full aromatized β - and methylated. They can be methylated in position 2 and /or 9 by N-methyltransferases.²¹¹ Non-N-methylated β -carbolines, such as norharman **134** and harman **135**, exhibit strong inhibitory activity on the monoamine oxidase, which has a favorable effect on the survival of cells. On the other hand, Norharman promotes apoptotic processes in the cell by stimulating or inhibiting various enzymes and proteins (phosphoinositol phospholipase C, glucose-regulating protein, triose phosphate isomerase).²⁰⁹ Quaternary salts, compounds methylated in position 2 inhibit the mitochondrial respiratory chain. They lead to reduced ATP synthesis and as a consequence to cell death.²¹⁸ The permanent charge appears to be essential, since the 9-methyl derivative shows only slight cytotoxic effects.²¹⁹ The in position 3 carboxylated derivatives, for instance β -carboline carboxylic acid (β CCM) are inverse agonists on the benzodiazepine receptor and show benzodiazepine antagonistic effects.

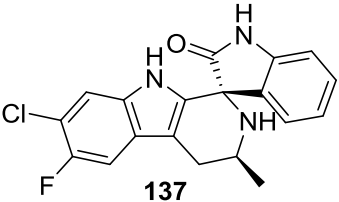
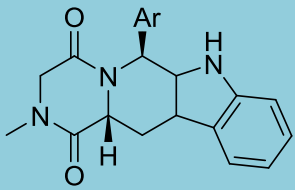
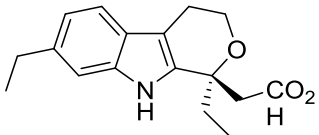
Table 8: Characteristics of β -carbolines **127-139**.

β -carboline	target	classification	properties
 <p>127</p>	TGFβ	Inhibitor	<ul style="list-style-type: none"> - Inhibition of cell proliferation in lung cancer (<i>in vitro</i>).²⁰⁶ - Growth suppression of lung and breast cancer (<i>in vivo</i>).²⁰⁶
 <p>128</p>	RAS-Oncogene -	Inhibitor	<ul style="list-style-type: none"> - Selective against oncogenes RAS (BJ-TERT/LT/ST/RAS^{V12}) - Induces non-apoptotic cell death (no activation of caspases).²⁰⁷
 <p>129</p>	Colon carcinoma	Selective cytotoxic compound	Specific, cytotoxic activity in colorectal cancer in COLO 205 and HCC-2998 cell lines. ²⁰⁸
 <p>130</p>	5-HTT, NET	Inhibitor	Inhibits the resorption of 5-HT and noradrenaline into synaptosome-rich regions of the hippocampus / hypothalamus. ²⁰⁹
 <p>131</p>	monoamine oxidase A	Inhibitor	<ul style="list-style-type: none"> - Inhibits the monoamine oxidase A.²⁰⁹ - Stimulates insulin secretion by binding imidazoline I₃ and ryanodine receptor.²¹⁰

Development of an acid catalysed approach to an oligo-thymidine initiated DNA-encoded β -carboline library

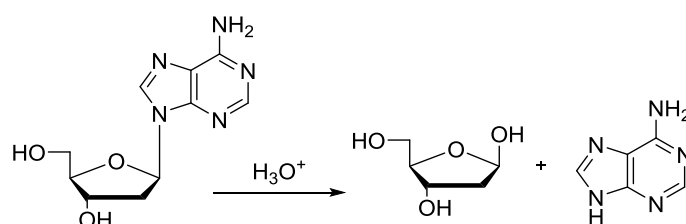
β -carboline	target	classification	properties
 <p>132</p>	G-protein coupled receptor	Inhibitor	Inhibits strongly and reversibly the dopamine, Noradrenaline and serotonin reuptake. ²¹¹
 <p>133</p>	G-protein coupled receptor	Inhibitor	High affinity to opioid receptors, ²¹² Possibly associated with dependency and neurons death by alcohol abuse. ²¹²
 <p>134</p>	monoamine oxidase A	Inhibitor	- Inhibits the monoamine oxidase A. ^{181, 213} - Stimulates the Phosphoinositol-Phospholipase C. ^{181, 213} - Inhibits Glucose-regulating protein. ^{181, 213}
 <p>135</p>	monoamine oxidase A	Inhibitor	- Inhibits the monoamine oxidase A. ^{181, 213} - Stimulates insulin secretion by binding to Imidazoline I ₃ - und ryanodine receptor. ²¹⁰
 <p>136</p>	Parasitic worms	Agonist	-Increases the permeability of trematode tegument to calcium. ²¹⁴ -Provokes the contraction of the parasite muscle. ²¹⁴

Development of an acid catalysed approach to an oligo-thymidine initiated DNA-encoded β -carboline library

β -carboline	target	classification	properties
 <p>137</p>	<p>P-type ATPase PfATP4</p>	<p>Inhibitor</p>	<ul style="list-style-type: none"> - Inhibits PfATP4, a P-type Na(+)-ATPase in the plasma membrane of the <i>plasmodium</i> parasite.²¹⁵ - Causes a disruption of its sodium homeostasis.²¹⁵
 <p>138</p>	<p>phosphodiesterase type 5</p>	<p>Inhibitor</p>	<ul style="list-style-type: none"> - Inhibits the phosphodiesterase type 5.²¹⁶
 <p>139</p>	<p>cyclooxygenase COX-2</p>	<p>Inhibitor</p>	<ul style="list-style-type: none"> - Inhibits the cyclooxygenase COX-2.²¹⁷

IV-4. Oligo-thymidine initiated DNA-encoded β -carbolines library

The deoxyribonucleic acid molecule (DNA) is composed of four nucleobases: the purines adenine (**A**) and guanine (**G**), as well as the pyrimidines thymine (**T**) and cytosine (**C**). Different experiments performed on DNA sequences consisting of purines and pyrimidines bases have shown that the incompatibility of the DNA with many catalyst systems used for heterocycles synthesis is based only on the purine bases. Strong acids usually used for the PSR protonate the N7-position of purine bases leading to depurination (Scheme IV-2).²²⁰⁻²²²

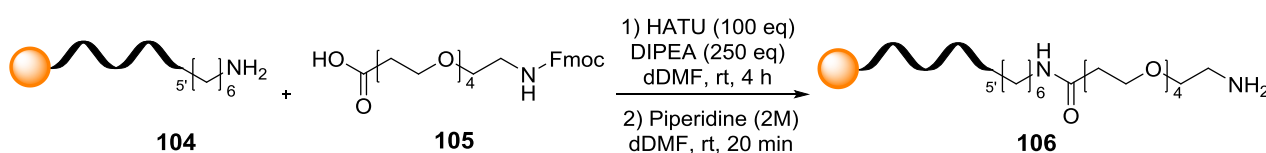


Scheme IV-2: The effect of strong acid on purine bases.

Thus, it was postulated that oligonucleotides that are deprived of purine bases should have a much broader tolerance towards harsh reaction conditions as such sequences will not depurinate. Therefore we designed for the construction of DEL libraries a hexathymidine-adapter oligonucleotide “hexT”. This hexT was bound to CPG solid support in order to broaden the range of applicable solvents for synthesis and contains a 5'-aminolinker to append functionalized building blocks. Subsequent to the synthetic transformation, the hexathymidine will be annealed and ligated to a duplex DNA containing a hexa A overhang.

The PSR gives access to several drug-like heterocycles molecules. The table 9 presents some widely-used pharmaceuticals accessible with the PSR and the general scheme of this reaction.¹⁸⁹ Fmoc-tryptophane was conjugated to hexT under the conditions described below (see *synthesis of the hexT-5'-PEG-aminolinker conjugate*). HPLC data of this amide synthesis show a distinct shift when compared to the starting material suggesting that reaction took place. The product identity was confirmed by MALDI-TOF-MS and analyzed by LC (Figure IV-8). The cyclization is then carried out after Fmoc deprotection and capping of the free amine.

IV-4.1. Synthesis of the hexT-5'-PEG-aminolinker conjugate



Scheme IV-3: Synthesis of DNA (hexa T) +PEG 4.conjugate **106**.

The amide coupling of the linker Fmoc-NH-PEG₄-COOH **105** to the hexT was performed making use of the hexT on 1000 Å CPG **104** (200 nM, ca. 7.2 mg), Fmoc-NH-PEG₄-COOH **105** (20 μmol, 9.6 mg), HATU (20 μmol, 7.6 mg) and 8.7 μL DIPEA. The removal of the Fmoc-protective group yielding compound **106** was done using 20% of piperidine in dry DMF for 20 minutes. For analysis, an aliquot of the DNA (20nM, ca 0.72 mg) was deprotected with 500 μL of 1:1 solution of conc. aq. NH₃ and conc. aq. MeNH₂ for 30 minutes at room temperature in order to remove it from CPG. 20 μL of 1M Tris-buffer (pH=7.5) was added and the DNA was dried in a SpeedVac and re-dissolved in 100 μL of distilled water and analyzed by HPLC with a gradient of aqueous triethylammonium acetate buffer (100 mM, pH= 8) and methanol (20:80 – 70:30 of methanol over 19 min).

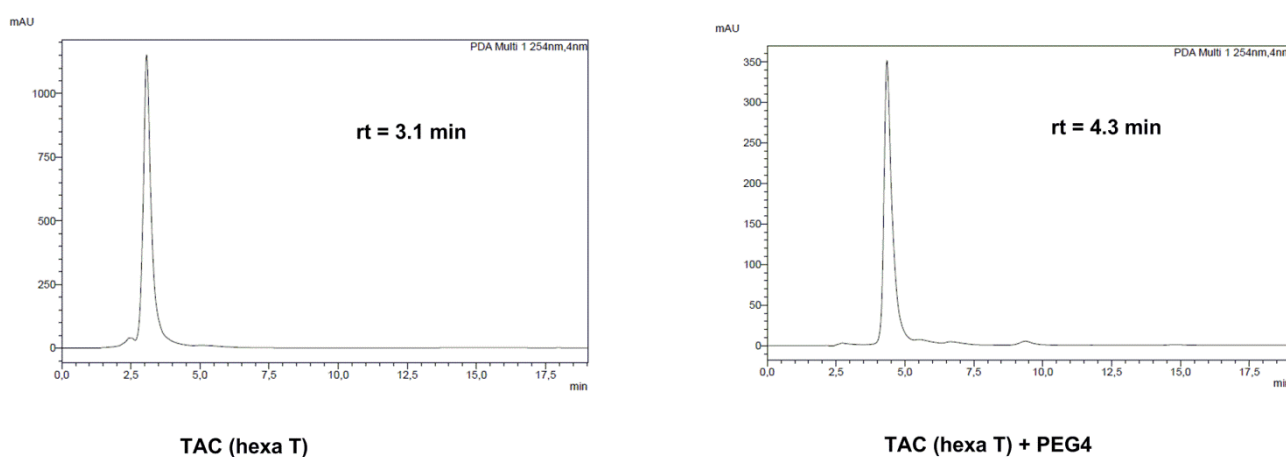
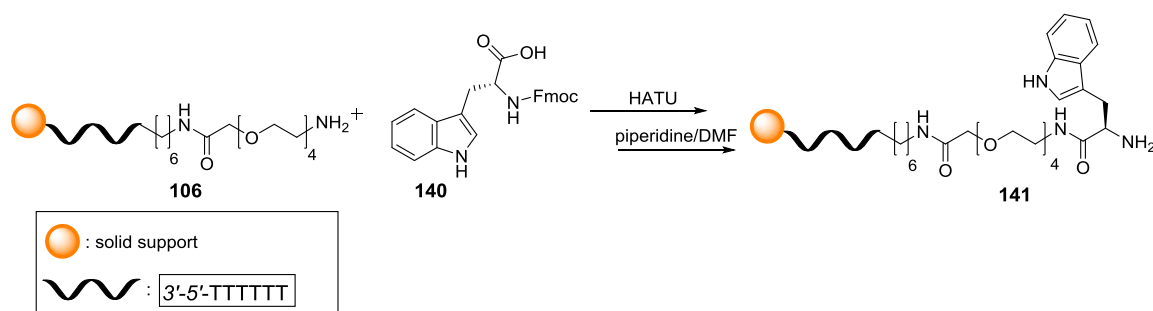


Figure IV-5: HPLC analysis of DNA (hexa T) +PEG 4.conjugate **106**.

IV-4.2. Synthesis of hexT-tryptophane conjugates



Scheme IV-4: Synthesis of the DNA tryptamine conjugate **141**.

The amide coupling of Fmoc-tryptophane **140** to the hexT-PEG conjugate **106** was performed as described above using the hexT on 1000 Å CPG, Fmoc-Trp-OH (100 eq), HATU (100 eq), and 250 eq of DIPEA. The removal of the Fmoc-protective group was performed with 20 % piperidine in DMF (100 μL) for 15 min at room temperature. The CPG containing the deprotected conjugate was filtered off and washed with each 3 x 200 μL of DMF, MeOH, MeCN and DCM. The CPG was then dried *in vacuo* for 15 min and a further cycle of coupling, capping, and Fmoc deprotection was performed as described above.

For analysis, an aliquot of the DNA (20nM, ca 0.72 mg) was deprotected with 500 μL of 1:1 solution of conc. aq. NH_3 and conc. aq. MeNH_2 for 30 minutes at room temperature in order to remove it from CPG. 20 μL of 1M Tris-buffer (pH=7.5) was added and the DNA was dried in a SpeedVac and re-dissolved in 100 μL of distilled water and analyzed by HPLC with a gradient of aqueous triethylammonium acetate buffer (100 mM, pH= 8) and methanol (20:80 – 70:30 of methanol over 19 min).

Development of an acid catalysed approach to an oligo-thymidine initiated DNA-encoded β -carboline library

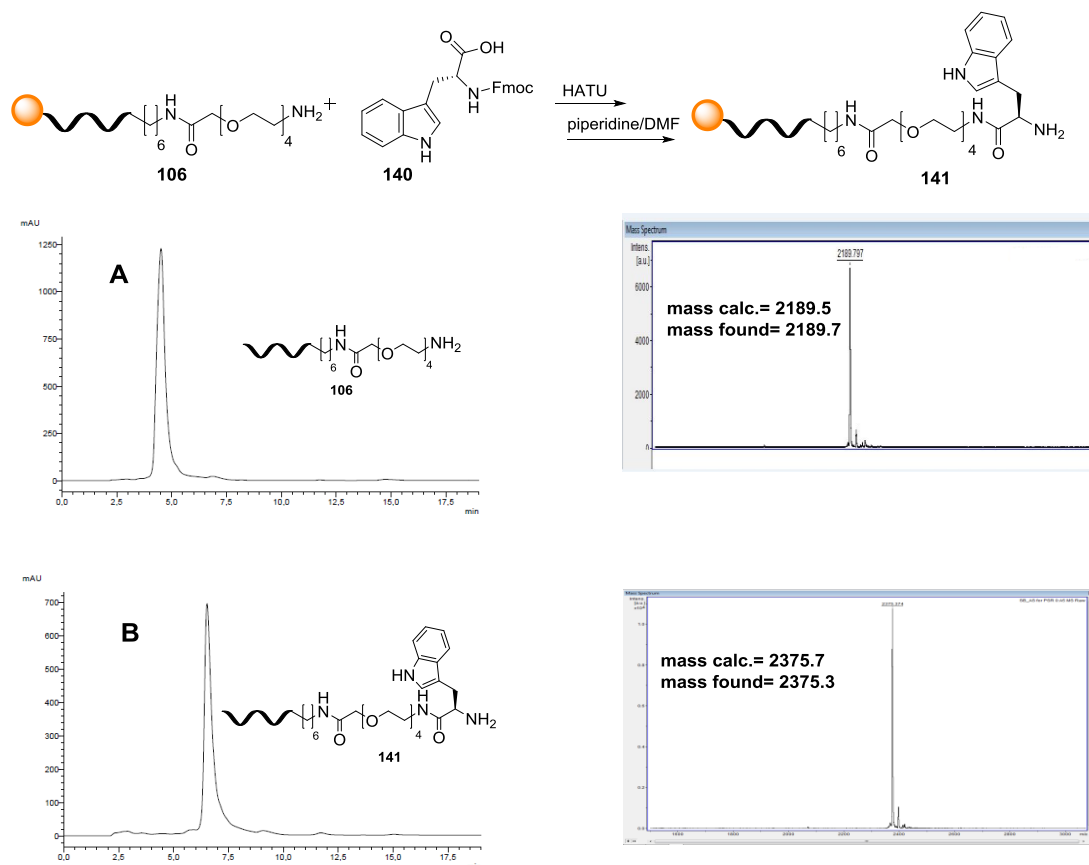
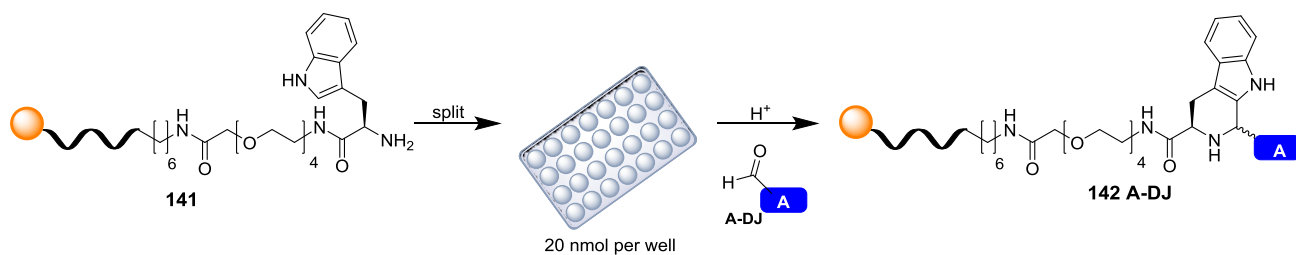


Figure IV-6: Synthesis of hexT- β -Carboline by Pictet-Spengler reaction. A) HPLC analysis of the DNA+PEG 4 conjugate **106**. B) HPLC analysis of the DNA + PEG 4+Tryp conjugate **141**.

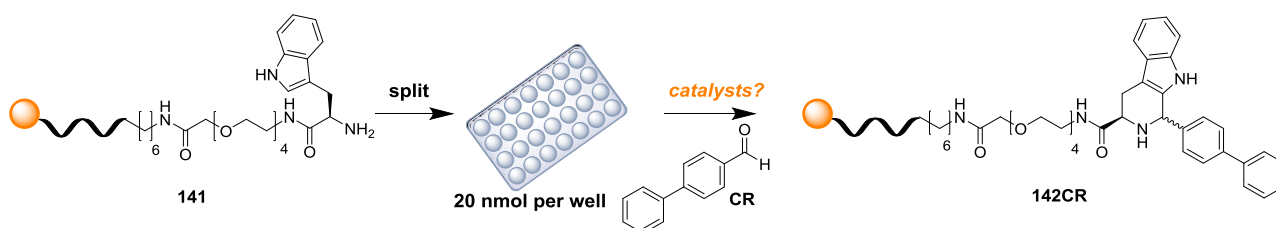
IV-4.3. Synthesis of hexT- β -carbolines from hexT-tryptophane conjugates by Pictet-Spengler condensation



Scheme IV-5: Synthesis of β -carboline conjugates from hexT tryptophane conjugate **141**.

To a suspension of the hexT-Tryp-NH₂ **141** conjugate bound to CPG in a solution of a suitable catalyst (*trichloroacetic acid (TCA)*, 1%; *trifluoroacetic acid (TFA)*, 1%; *diphenyl hydrogen phosphate (DHP)* 100 eq, 0.75 mg) in 45 μ l of a suitable dry solvent (*DCM*, MeCN, *DCE*, *PhMe*) was added 100 eq of an aldehyde or ketone of choice. The reaction was shaken at room temperature for 16 h, then, the DNA was filtered off, washed successively with DMF (3 x 200 μ l), MeCN (3 x 200 μ l), MeOH (3 x 200 μ l) and DCM (3 x 200 μ l), and dried *in vacuo* for 15 min. The DNA was deprotected and removed from the CPG by treatment with AMA (AMA= aqueous ammonia (30 %)/ aqueous methylamine (40 %) 1:1 vol/vol) at room temperature for 30 min. Then, 20 μ l of 1M Tris-buffer (pH=7.5) was added, the mixture was dried in a SpeedVac, re-dissolved in 100 μ l of distilled water and purified by HPLC on a Gemini 5u C₁₈ 110A column; 100*10.0 mm with a gradient of aqueous triethylammonium acetate buffer (100 mM, pH= 8) and methanol (20:80 – 70:30 of methanol over 19 min). (See experimental section).

IV-4.4. Investigation of different catalysts



Scheme IV-6: Investigation of different catalysts for the synthesis of hexT- β -Carboline conjugate **142CR**.

The mild and efficient protocol for the Pictet-Spengler tetrahydro- β -carboline synthesis in DCM using TFA as the acid catalyst is reported.¹⁹⁸ In order to investigate, if other catalysts in combination to other solvents can be used, different catalysts such as Lewis acid, Brønsted acid and organocatalysts have been tested for C-C bond forming reactions of tryptamine derivatives and various aldehydes. For this purpose, 20 nmol of hexT-tryptophane conjugate **141** was used for each reaction in combination with 100 eq of the 4-biphenylbenzaldehyde **CR**. We estimated conversion rates based on integration of HPLC peak of the product **142CR** versus starting material in the HPLC. The best conversion was achieved using 0.03 M of TCA in dry DCM for 18 hours (Table 10, **entry 3**). We noticed that lower loadings of catalyst lead to lower conversion of the desired product. 100 eq of *diphenyl* hydrogen phosphate (DHP) under the same reactions conditions lead to 80 % conversion (Table 10, **entry 1**).

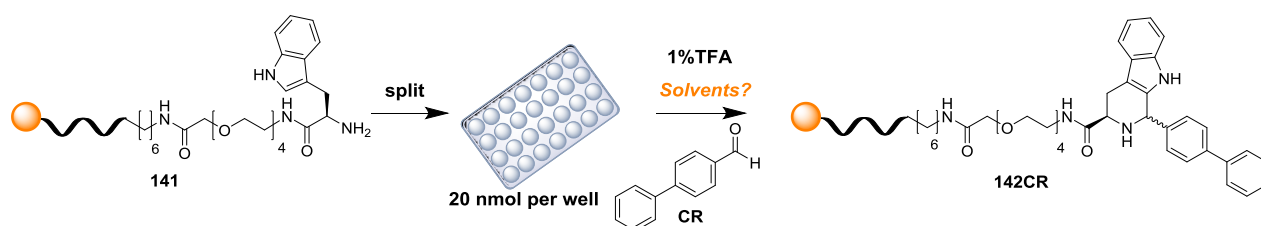
Development of an acid catalysed approach to an oligo-thymidine initiated DNA-encoded β -carboline library

The use of 0.1 M of BCl_3 and 1% of BF_3 gave no conversion of the starting material at all after 18 hours. Instead DNA degradation was observed in both cases (Table 10, **entries 2 and 6**). No product was formed by the use of 100 eq of 4-nitrobenzoic acid. An explanation could be the difficulty to dissolve this catalyst in dry DCM (Table 10, **entry 4**). The solubility problem was solved using DMF as solvent; however no reaction took place in presence of DMF. 50% of formic acid in dry DCM led to a 70 % conversion of the tryptophane conjugate after 4 hours. Finally, the use of 0.1 eq of silver trifluoroacetate in DCM with 2,6 lutidine as base for 4 hours, as well as the use of 1% of thioglycolic acid over 18 hours gave 10 % of the desired product (Table 10, **entries 5 and 7**). Increasing the reaction time in (Table 10, **entry 7**) should increase the yield. Yet, 1% of thioglycolic acid overnight led to DNA degradation. In conclusion, the catalyst screen at room temperature for overnight revealed product formation and satisfying product yields with 1% TFA for 18 hours, 10% TFA for 4 hours, 0.03 M TCA for 18 hours, 100 eq DHP for 18 hours and 50% formic acid for 4 hours for aldehydes, and 10% TFA for 6 hours for ketones.

Table 9: [a] The reactions were carried out on 20 nmol of hexT -tryptophane conjugate and purified by preparative HPLC. [b] Conversions of the tryptamine conjugate **141** to the β -carbolines after reaction using 4-biphenylcarboxaldehyde **CR**.

entry ^[a]	catalyst (conc.)	solvent	T (°C)	t (h)	conversion (%) ^[b]
1	$(\text{C}_6\text{H}_5\text{O})_2\text{P}(\text{O})\text{OH}$ (100 eq)	DCM	rt	18	80
2	BCl_3 (0.1 M)	DCM	rt	18	0
3	CCl_3COOH (0.03 M)	DCM	rt	18	100
4	$\text{C}_6\text{H}_4\text{NO}_2\text{CO}_2\text{H}$ (100 eq)	THF	rt	18	0
5	HSCH_2COOH (1%)	DCM	rt	18	10
6	BF_3 (1%)	DCM	rt	18	0
7	CF_3COOAg with 2,6 lutidine	DCM	rt	4	10
8	HCOOH (50%)	DCM	rt	4	70

IV-4.5. Investigation of different solvents [a]



Scheme IV-7: Investigation of different solvents for the synthesis of hexT- β -Carboline conjugate **142CR**.

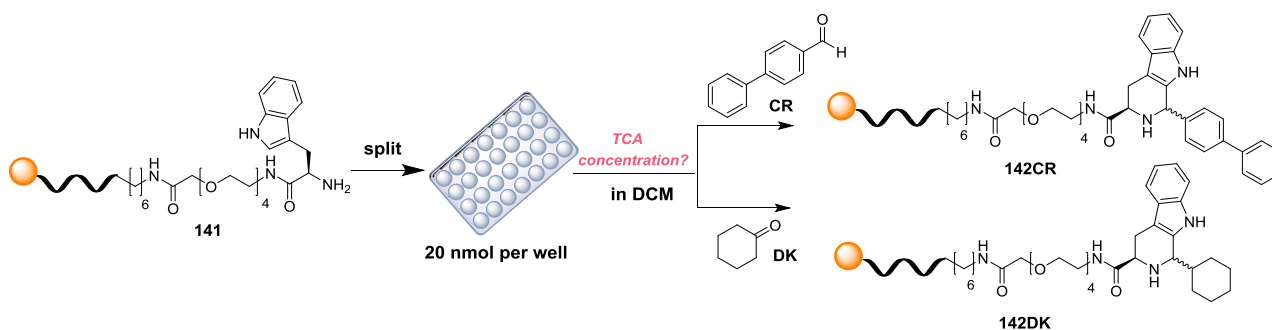
A solvent screen at room temperature for overnight revealed product formation in different solvents. For this purpose, eight different solvents were investigated. The results are summarized in the table 11 depicted below. The highest conversion was observed in DCM, ACN, 1,2-DCE and toluene with 100 % conversion of the tryptamine **141** (Table 11, **entries 1, 3, 5 and 9**). No reaction took place at room temperature with THF, DMSO and 1,4-dioxane (Table 11, **entries 4, 6 and 8**). Some aldehydes like isatin, substituted imidazole carbaldehydes, non-substituted and substituted indole carboxaldehydes are insoluble in DCM. For those aldehydes, we have tried to find alternative solvents. Anhydrous DMF (with or without DCM as co-solvent) gave better solubility but the reaction led just to 50-60% conversion (Table 11, **entry 2**). In general, we observed that reactions did not take place with polar solvents that might be contaminated with water. The diagram depicted below summarizes the solubility of the 114 aldehydes used in the first step of library synthesis.

Development of an acid catalysed approach to an oligo-thymidine initiated DNA-encoded β -carboline library

Table 10: [a] The reactions were carried out on 20 nmol of hexT -tryptophane conjugate and purified by preparative HPLC. [b] Conversions of the tryptamine conjugate 141 to the β -carbolines after reaction using 4-biphenylcarboxaldehyde CR.

entry ^[a]	catalyst used	solvent	T (°C)	t (h)	conversion (%) ^[b]
1	1% TFA	DCM	rt	18	100
2	1% TFA	DMF/anhydrous DMF	rt	18	0/ 60
3	1% TFA	MeCN	rt	18	100
4	1% TFA	THF	rt	18	0
5	1% TFA	1,2-DCE	rt	18	100
6	1% TFA	DMSO	rt	18	0
7	1% TFA	MeOH/anhydrous MeOH	50°C-rt	18	0/ 55
8	1% TFA	1,4-dioxane	rt	18	0
9	1% TFA	toluene	rt	18	100

IV-4.6. Investigation of different concentrations and reaction times for the synthesis of hexT- β -carboline conjugates



Scheme IV-8: Investigation of different concentrations and reaction times for the synthesis of hexT- β -carboline conjugates.

For 0.03 M TCA showed complete conversion (almost 100 %) in the formation of β -carbolines, the investigation of different concentrations of this catalyst at different reaction times with two different substrates (an aldehyde and a ketone) were one of our focus. For this purpose, the PSR was performed with two reference compounds (4-Biphenylcarboxaldehyde **CR** and cyclohexanone **DK**) at room temperature in TCA dissolved in DCM modifying the reaction time. Reaction conditions tested are summarized in table 12. Better conversion was noticed with compound **CR** than with **DK**. The Pictet-Spengler cyclisation of the tryptamine to the β -carboline using 0.05 M of TCA in dry DCM at room temperature for 4 hours gave 35% yield (Table 12, **entry 1**). No conversion of the tryptamine was observed with the ketone after 6 hours (Table 12, **entry 2**). Increasing the reaction time brought no change in the reaction with the cyclohexanone (Table 12, **entry 3**). We observed a better conversion when raising the concentration of the catalyst from 0.05 M to 0.1 M after 4 hours, as well as from 0.1 M to 0.3 M (Table 12, **entry 4**). Indeed 0.03 M TCA showed 100 % conversion of the tryptamine, but the reaction should run at least for 16 hours, otherwise we observed an uncompleted conversion (Table 12, **entry 6**). However, this catalyst was not robust enough to allow cyclisation of ketones also after long reaction times (Table 12, **entry 7**). Aldehydes show a better conversion than ketones. Neither the increase of the concentration of the catalyst, nor the longer reaction times show an improvement in the conversion of ketones by PSR.

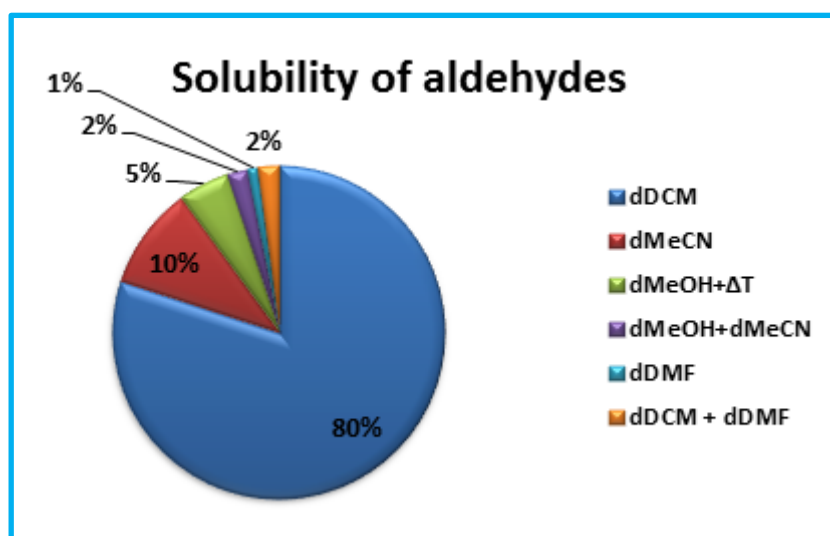


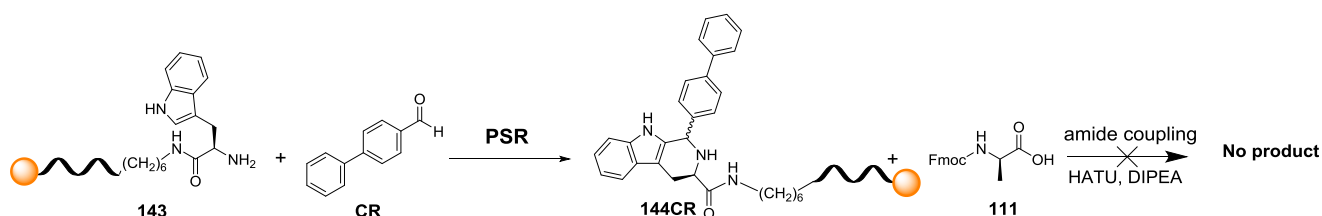
Figure IV-7: Diagram showing the solubility of 114 aldehydes used for the PSR.

Development of an acid catalysed approach to an oligo-thymidine initiated DNA-encoded β -carboline library

Table 11: [a] The reactions were carried out on 20 nmol of hexT-tryptophane conjugate and purified by preparative HPLC. [b] Conversions of the tryptamine conjugate to the β -carbolines after reaction using 4-biphenylcarboxaldehyde CR. The reaction was also repeated using cyclohexanone DK.

entry ^[a]	catalyst conc.	solvent	T (°C)	t (h)	conversion (%) ^[b]
1	0.05 M	DCM	rt	4	35
2*	0.05 M	DCM	rt	6	0
3*	0.05 M	DCM	rt	18	0
4	0.1 M	DCM	rt	4	55
5*	0.1 M	DCM	rt	6	0
6	0.03 M	DCM	rt	4	60-70
7*	0.03 M	DCM	rt	18	0

IV-4.7. Control experiments with Fmoc protected alanine



Scheme IV-9: Amide synthesis with Fmoc protected alanine 111.

β -carbolines are synthesized from tryptamine **143** as described above. One important control experiment was to investigate whether the cyclization would also take place using an alanine derivative, what of course would mean that the PSR could also work with a broader range of conjugates. Therefore, we have performed the PSR after attachment of alanine to the hexT adapter via amide synthesis. The HPLC chromatograms (Figure IV-13) show that no reaction took place.

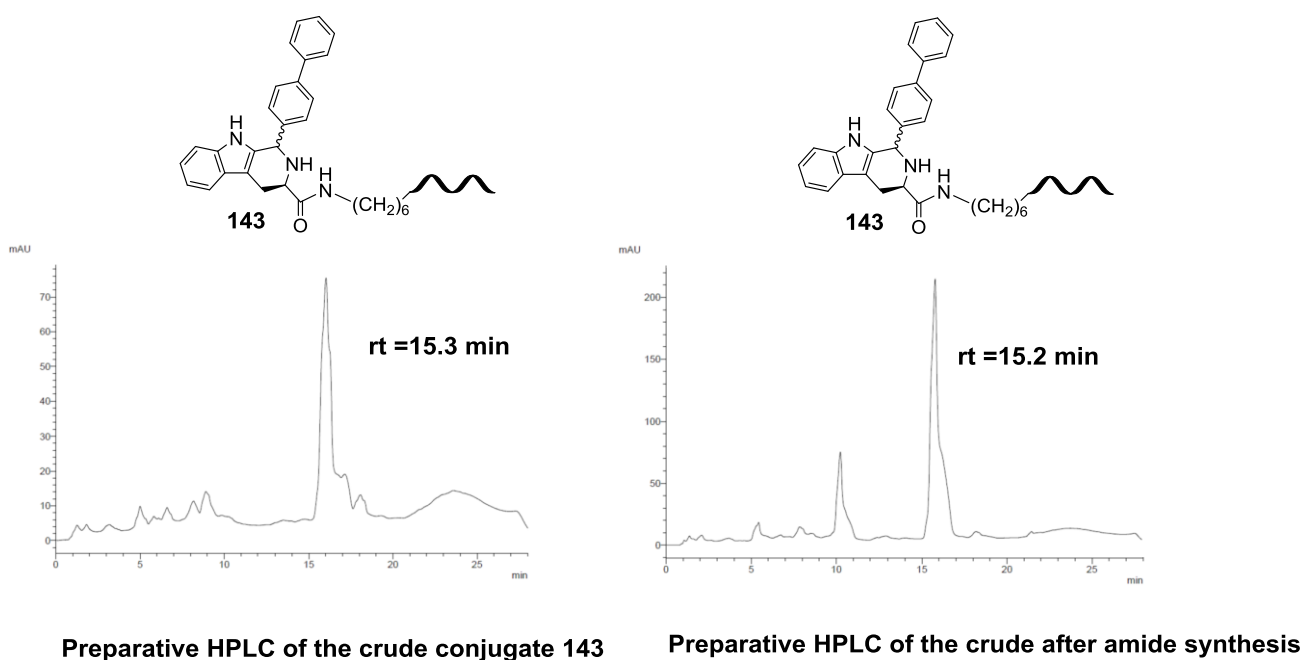
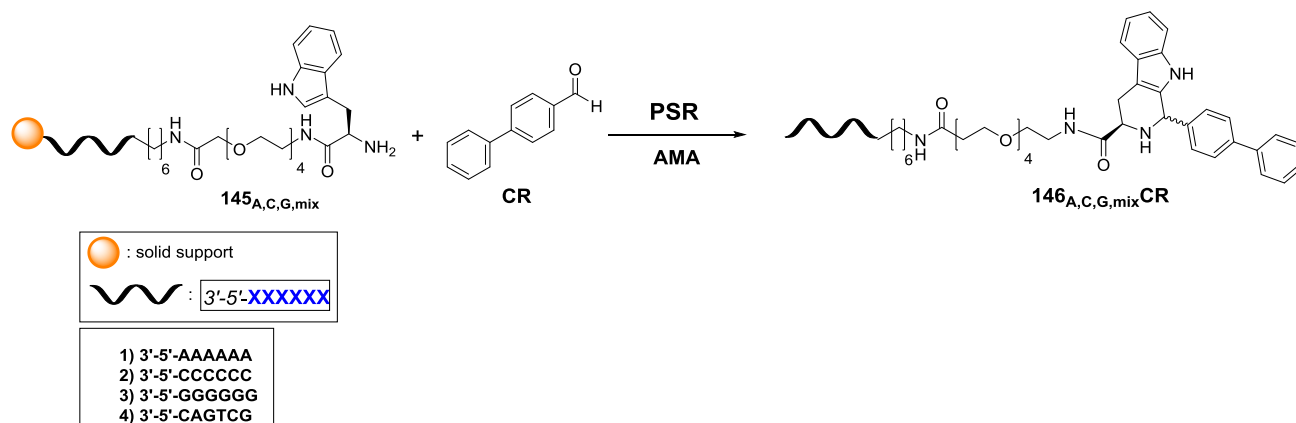


Figure IV-8: Left: Preparative HPLC of the crude conjugate **143**. Right: Preparative HPLC of the crude after amide synthesis. It corresponds to compound **143**.

IV-4.8. Control experiments: Pictet-Spengler reaction with tryptophane conjugates of hexa-A, hexa-G, hexa-C, and a mixed sequence

The Pictet-Spengler reaction experiments described above were repeated with tryptophane conjugates of hex-C and homopurine sequences hexa-A and hexa-G, as well as with a mixed sequence DNA: **CAG TCG**. The preparative HPLC chromatograms showed that all purine-containing DNA sequences were degraded by these catalysts. The hexa-C sequence in contrary tolerated the PSR conditions mentioned above and HPLC chromatograms show the successful formation of the β -carboline **146_CCR**.

Development of an acid catalysed approach to an oligo-thymidine initiated DNA-encoded β -carboline library



Scheme IV-10: Pictet-Spengler reaction with tryptophan conjugates **145_A** (on hexa-A), **145_G** (on hexa G), **145_C** (on hexa-C), and **145_{mix}** (on a mixed sequence CAGTCG).

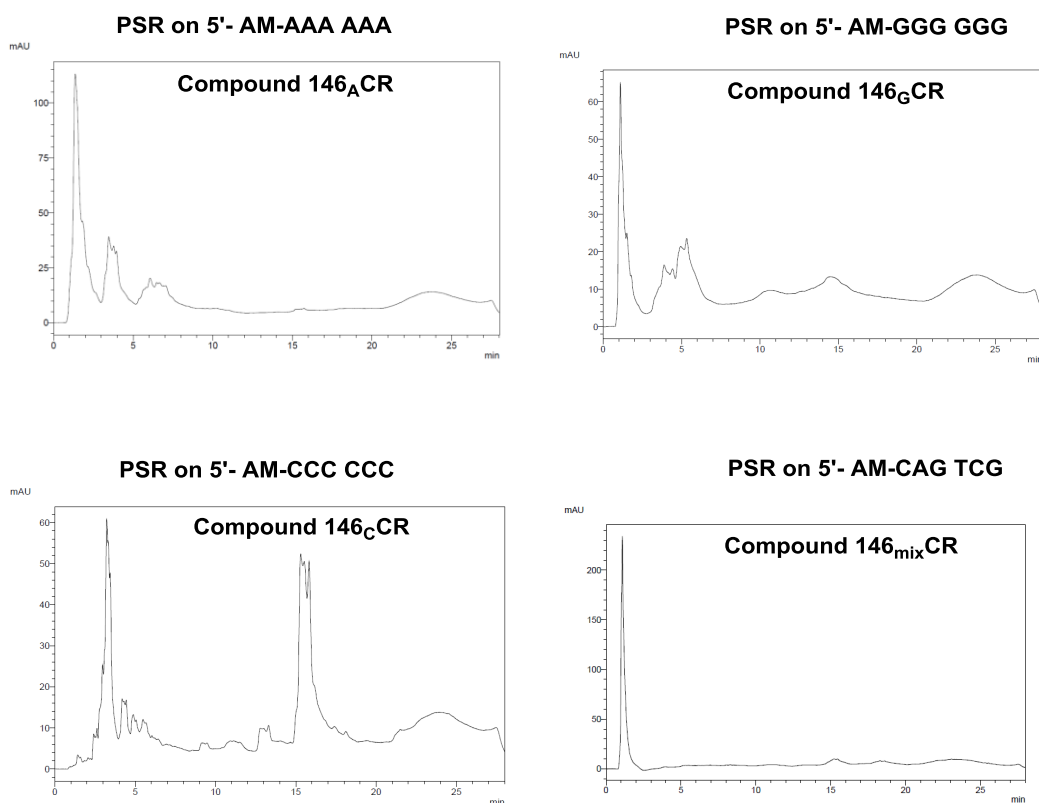
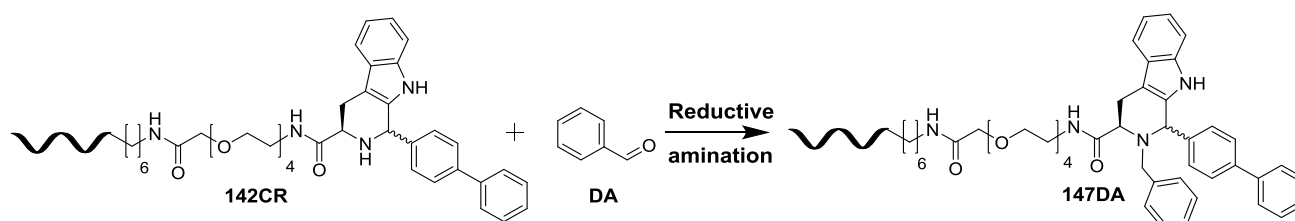


Figure IV-9: Preparative HPLC chromatograms of crude conjugate **146_{A,C,G,mix}CR** synthesized with 1% TFA in dry DCM at room temperature for 18 hours on vier differents DNA sequence containing purines.

IV-4.9. Introduction of a second set of building blocks

We next explored different strategies for library synthesis. As the β -carboline attached to the hexT adapter contains a free amine, we intended to introduce a second set of building blocks at this position via amide synthesis. The reaction was performed under standard reaction conditions for amide synthesis already described.¹²⁵ The HPLC chromatogram showed unfortunately no shift of the starting material suggesting that the reaction did not take place. We have tried to force the coupling of carboxylic acids to the hexT- β -carboline conjugate using different coupling reagents (HBTU, PYBOP and EDC). These experiments gave unfortunately no product. One alternative was the introduction of aldehydes at the same position rather via reductive amination. As the reductive amination is described on DNA on solid support,²²³ an alternative was to investigate this reaction. In addition, it is known that reductive amination yields higher conversion with secondary amines than with primary amines on solid support.²²³



Scheme IV-11: Reductive amination on hexT- β -carboline in solution. Reductive amination in solution with benzaldehyde **DA** on the hexT- β -carboline **142CR** synthesized by PSR with 4-biphenylcarboxaldehyde **CR**.

Reductive amination comprises a one- or two-step procedure in which an amine and a carbonyl compound condense to afford an imine or iminium ion that is reduced *in situ* to form an amine product. We tested reductive amination as a route for the synthesis of DNA-conjugates prepared from a hexT- β -carboline **142CR** and aldehydes. The reductive amination was performed in aqueous solution on DEAE sepharose.²²³ Different reaction conditions were tested and we estimated conversion rates based on integration of HPLC peak of the product versus starting material in the HPLC. Incubating the substrate **142CR** in buffer (300 mM MOPS, pH 7.4) with benzaldehyde **DA** and sodium cyanoborohydride (NaBH_3CN) as the reducing agent for overnight at 37°C provided the desired iminium ion conjugated on DNA. Unfortunately, HPLC chromatograms of the reaction show that the imine was not reduced, when compared to the reference compound (See Figure IV-10). Increasing the concentration of NaBH_3CN did not enhance the product formation in the reduction step (Table 13, **entry 1**).

Development of an acid catalysed approach to an oligo-thymidine initiated DNA-encoded β -carboline library

The use of borate buffer (pH 9.4, 250 mM) with 50 eq of aldehyde at 60°C for 4 hours showed a complete condensation but no reduction of the imine was observed (Table 13, **entry 2**). A successful condensation and reduction was observed making use of reactions conditions described in table 13, **entry 3**. The best conversion was achieved when 100 eq of aldehyde and 40 eq of the reducing agent NaBH_4 were used. We noticed that higher loadings of the reducing agent lead to DNA degradation already after 1 hour. We studied the influence of different buffers in combination with NaCNBH_3 as reducing agent and observed no successful reduction of the imine also overnight (Table 13, **entries 4, 5 and 6**). Reductive amination of DNA-NH-R conjugates generates under conditions described in table 13, **entry 7** the desired conjugates in excellent yields for the majority of aldehydes, when performed on solid support.

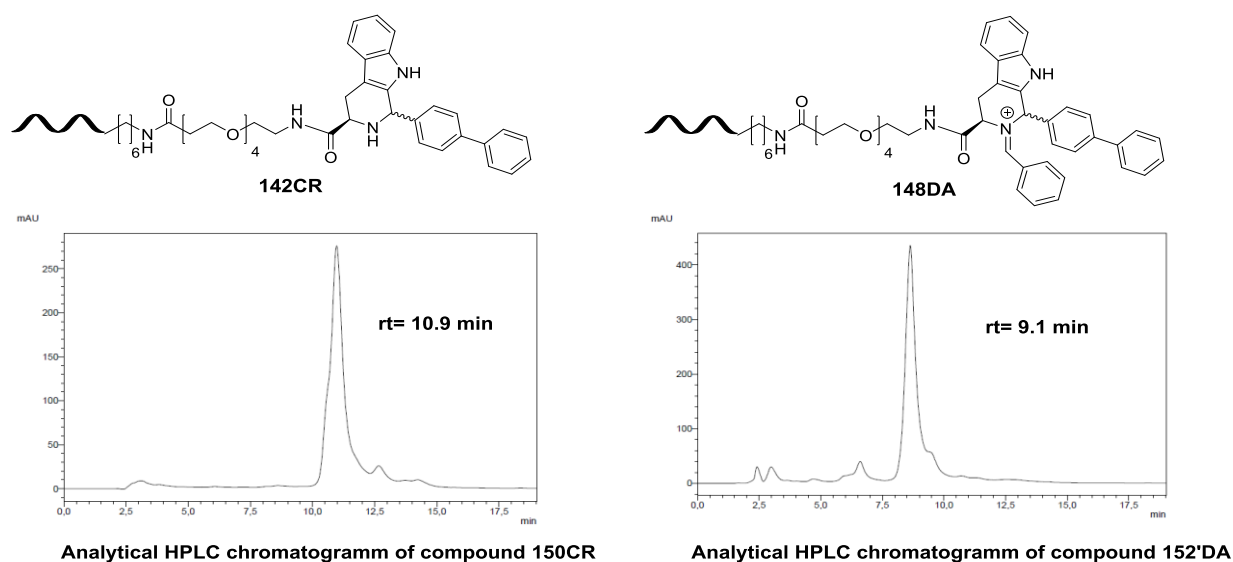


Figure IV-10: Left: Analytical HPLC Chromatogram of the hexT- β -carboline **142CR** synthesized by PSR with 4-biphenylcarboxaldehyde **CR**. Right: analytical HPLC chromatogram of the condensed product **148DA** obtained after reductive amination with benzaldehyde **DA**.

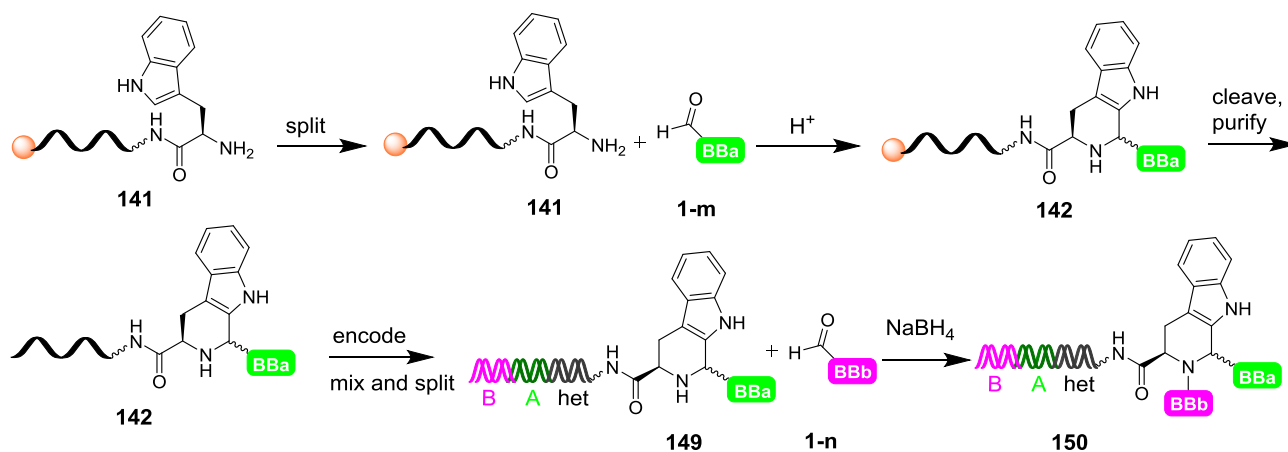
Development of an acid catalysed approach to an oligo-thymidine initiated DNA-encoded β -carboline library

Table 12: [a] The reactions were carried out on 500 pmol of hexT - β -carboline conjugate and analysed by analytical HPLC. [b] The formation of the iminium ion was evaluated by analytical HPLC. We estimated conversion rates based on integration of HPLC peak of the product versus starting material in the HPLC.

entry ^[a]	reaction conditions	iminium formation ^[b]	reduction Step
1	MOPS Buffer (300 mM, pH 7.4), 50 mM of aldehyde Acetic acid (50 mM) NaBH ₃ CN (50 mM) at 37°C overnight	Condensation was observed	No reduction was observed
2	1 mM DNA in borate buffer (pH 9.4, 250 mM) 50 equiv aldehyde (200 mM DMA), 60°C, 4 h.	Condensation was observed	No reduction was observed
3	Step 1: 1mM DNA in borate buffer, 100 equiv aldehyde (200 mM DMA), rt, 1h. Step 2: Then add 40 equiv NaBH ₄ (200 mM ACN stock), rt, 1 h.	Condensation was observed	80% of the iminium ion was reduced
4	DNA in sodium acetate buffer 2500 equiv. Aldehyde (from a 200 mM DMSO stock) ~50 mM final conc. 1400 equiv. NaCNBH ₃ ~28 mM	Condensation was observed	No reduction was observed
5	DNA in sodium phosphate buffer 100 equiv. of the aldehyde (13 μ l 200 mM in DMF) followed by 100 equiv. NaCNBH ₃ solution (13 μ L 200 mM in CH ₃ CN) and incubated at RT overnight.	Condensation was observed	No reduction was observed
6	2500 equiv. aldehyde (from a 200 mM DMSO stock), ~50 mM final conc. 1400 equiv. NaCNBH ₃ , ~28 mM pH 5 buffer (200 mM NaOAc), ~20 μ M DNA at 30°C overnight.	Condensation was observed	No reduction was observed
7	MOPS Buffer (300 mM, pH 7.4) 50 mM of aldehyde Acetic acid (50 mM) NaBH ₄ (50 mM) at 37°C overnight	Condensation was observed	No reduction was observed
8	Step 1: 1mM DNA in borate buffer, 40 equiv aldehyde (200 mM DMA), rt, 1h. Step 2: Then add 40 equiv NaBH ₃ CN (200 mM ACN stock), rt, 1 h.	Condensation was observed	No reduction was observed

In the set of 45 different aldehydes that have been tested for their reactivity by reductive amination, we obtained under optimized reaction conditions a successful condensation in the first reaction step and satisfying reduction in the second step with 28 aldehydes. 2 aldehydes **BD** and **DG** (Table 18) showed no conversion at all and the remaining 15 aldehydes showed formation of the imine or condensed product after 1 hour, but no reduction was observed, also after higher loadings of the reducing agent to 60 or 100 eq. Increasing the reaction time did not bring any significant changes. In contrary, 100 eq of the reducing agent in the second reaction step led to hexT degradation already after 1 hour. This observation motivated us to keep the catalyst amount by 40 eq per reaction. From the 28 aldehydes showing a successful conversion of the starting material to the target product, 27 aldehydes yielded the desired product with more than 50 % (Table 18). We observed that branched and straight aliphatic aldehydes were partially converted to the imine, but they did not give rise to the desired product at all. Substituted pyrazole carbaldehyde showed under optimized reaction conditions no formation of the condensed product. Substituted and non-substituted pyridine carboxaldehydes as well as indole carboxaldehydes gave good conversion of the secondary amine to the imine but no reduction was observed under optimized reaction conditions. *Ortho*-substituted aldehydes show the same behavior leading to the condensed product. Olefinic aldehydes like citral gave similar result. They also showed no reduction of the imine to the desired product, but a satisfying condensation was observed (Table 18). The most aromatic aldehydes that have been used gave medium to high conversion into desired β -carbolines, whereby the *para*-substituted and the most *meta*-substituted aldehydes gave rise to the hexT- β -carboline conjugate in higher conversion (Table 18).

IV-4.10. Projected synthesis of a tiDEL library based on β -carboline



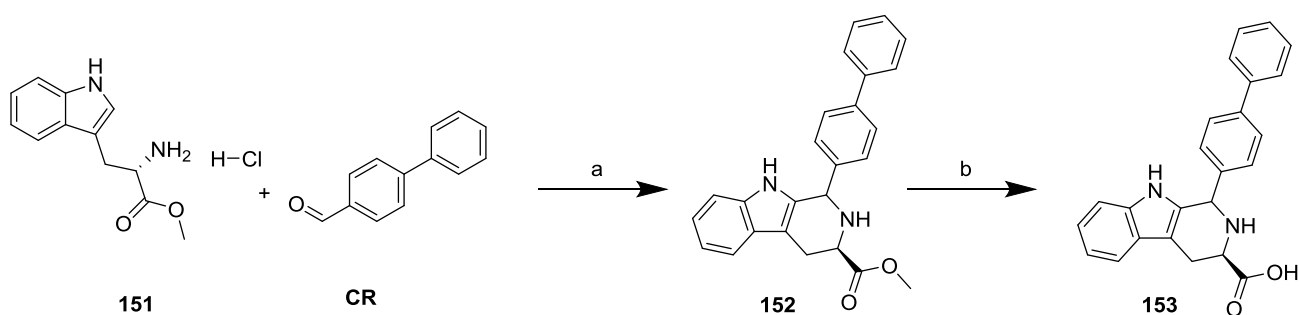
Scheme IV-12: Projected synthesis of a tiDEL library based on β -carboline.

The synthesis of the tiDEL library will be performed as described in figure 3. The first step of library synthesis takes place on controlled pore glass (CPG) solid support. The next step after the encoding by DNA ligation will be done on DEAE ion-exchange resin. Library synthesis was initiated with coupling of a protected PEG (4) linker **105** to solid phase-coupled 5'-aminolinker modified hexT adapter. For encoding and screening purpose, the PSR was performed with a PEG linker playing the role of a spacer between hexT adapter and small molecule to avoid interactions between DNA and target proteins. The tryptamine conjugate will be formed in the next step by coupling of Fmoc-Tryp-OH to the hexT-PEG (4) linker conjugate **106**. The solid phase containing hexT-PEG(4)-tryptamine conjugate **141** is then splitted in 20 nmol aliquots in two 96well plates and the Pictet-Spengler reaction was performed using 114 aldehydes. The most aldehydes were coupled using 1%TFA as catalytic system in DCM at 25°C for overnight, while insoluble aldehydes in DCM like isatin and substituted imidazole carbaldehyde were coupled to tryptamine conjugate using 10% TFA as catalytic system in DMF at 25°C for 4 hours. Non substituted and substituted indole carboxaldehydes show to be completely insoluble in DMC, ACN and DMF. These aldehydes were coupled to the tryptamine using 10% TFA as catalytic system in a mixture of MeOH/DCM at 50°C for 4 hours. The Pictet-Spengler cyclization yielded hexT-PEG(4)- β -carboline conjugates that were washed after the reaction on a 96 well filter plate, deprotected and removed from the CPG with aq. ammonia/methylamine for 30 min at room temperature on the filter plate which was sealed for this purpose.

The hexT-PEG (4)- β -carbolines conjugates were then purified by ion pair reverse phase HPLC. HPLC analysis of the products indicated a purity of the conjugates and MALDI-TOF/TOF-MS analysis confirmed the identity of the products (Table 18). 114 aldehydes were tested for their reactivity. After deprotection and removal from the CPG, and HPLC purification of hexT conjugates will be submitted to encoding by T4 DNA ligation yielding the primary library. Library will be finalized by introducing of a second set of building blocks by reductive amination on DEAE sepharose that will lead to complete tiDEL.

IV-4.11. Synthesis of a reference compound

In order to confirm our hypothesis that TiDEC can yield hexT- β -carboline conjugates and to validate the reaction conditions used “on-DNA”, the reference molecule **153** was synthesized and coupled to the hexT oligonucleotide yielding a hexT- β -carboline conjugate. The β -carboline reference compound **153** was synthesized in a straightforward manner. The Pictet-Spengler reaction was firstly performed with tryptophan methyl ester hydrochloride and biphenyl-4-carboxaldehyde **CR** using trifluoroacetic acid as catalyst. The carboxylic ester **152** was then hydrolyzed in the presence of lithium hydroxide yielding the compound **153**.



Scheme IV-13: Synthesis of the β -carboline reference compound **157**. Reagents and conditions: a) L-Tryptophan methyl ester hydrochloride **155**, biphenyl-4-carboxaldehyde **CR**, 3-5% TFA in dry CH_2Cl_2 , $T=0-25^\circ\text{C}$, 18 hours. b) 2.4 eq LiOH; water in THF in ratio 1:1, $0-25^\circ\text{C}$, 18 hours.

The reference conjugate **153** was attached to DNA-PEG 4 conjugate **106** by amide synthesis yielding the β -carboline conjugate **154CR**. Compound **154CR** matched the conjugate **142CR** synthesized “on-DNA” in HPLC and MALDI analysis when both were injected separately and also when they were co-injected (Figure IV-9). In conclusion, this chapter has presented the successful formation of β -carbolines conjugates on a hexT oligonucleotide designed for this purpose. The β -carbolines conjugates were subsequently purified by HPLC and encoded by DNA ligation.

Development of an acid catalysed approach to an oligo-thymidine initiated DNA-encoded β -carboline library

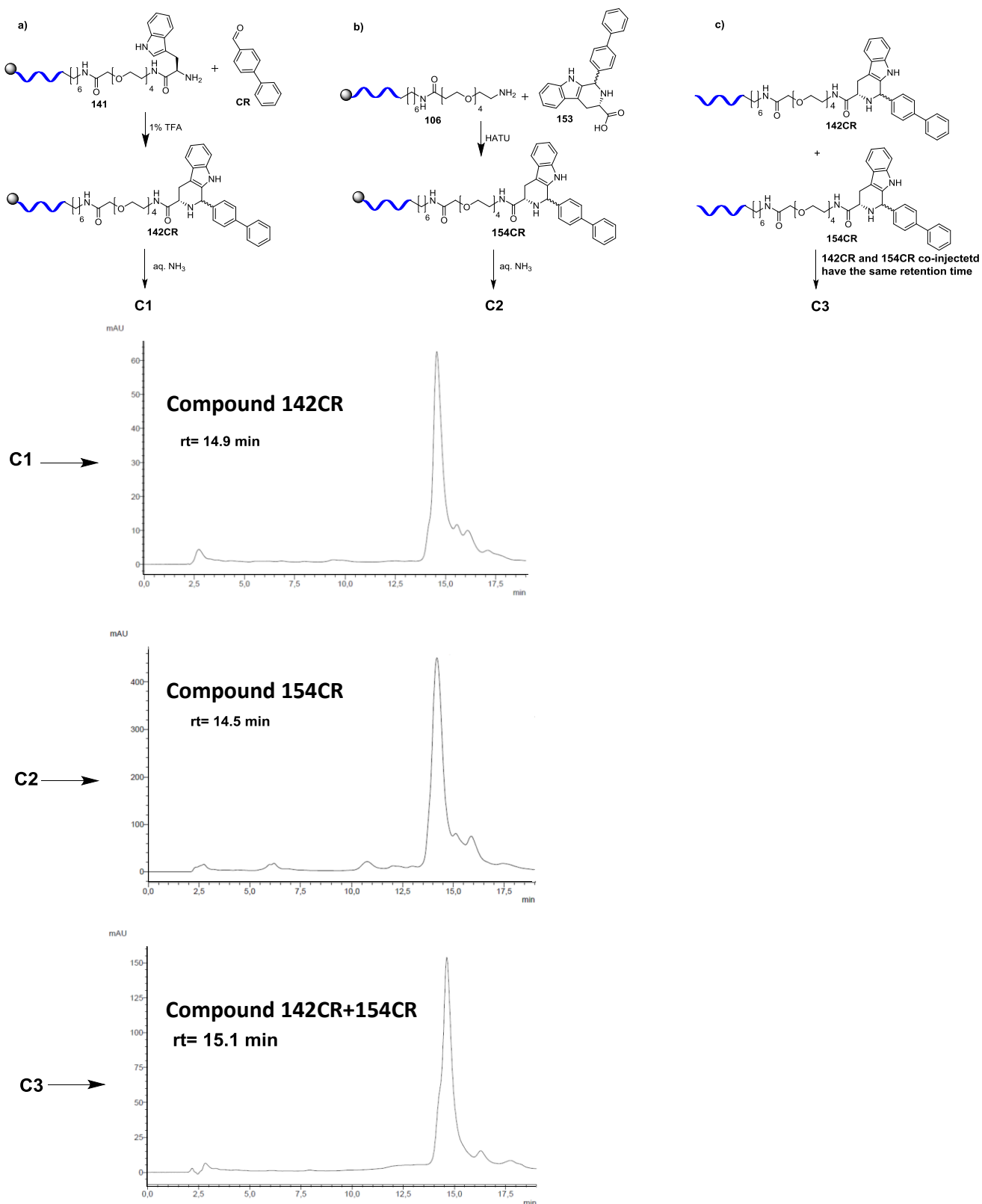


Figure IV-11: Comparison between DNA-conjugate **154CR** synthesized off-DNA with DNA-conjugate **142CR** synthesized on-DNA.

Chapter V:SUMMARY AND OUTLOOK

The identification of novel small-molecule ligands against medically relevant proteins is a central challenge in drug research. However, the identification of binding compounds by conventional HTS campaigns is highly demanding in terms of time and resources. The technology based on the synthesis of DNA-encoded chemical libraries presented in this thesis represents a selection-based ligand-identification methodology. This methodology addresses the need for technologies substantially more efficient than HTS for hit discovery programs. This technology allows affinity selections to be performed using DNA-encoded combinatorial libraries (DELs) composed of millions of DNA-encoded compounds.^{41, 44, 59} The identification of high-quality ligands from DNA-encoded chemical libraries against target proteins of interest is conditioned by diverse factors, including library design, the quality of the library and of target proteins, experimental conditions in affinity selections, and the amount of sample sizes after decoding of DNA sequences. In recent years, several academic and industrial laboratories have described the identification of novel biologically active molecules against pharmaceutically relevant proteins, including tumor-associated antigens,^{41, 50, 54, 59} kinases⁸⁵ and immunocytokines.⁴¹

DNA-encoded chemical libraries (DELs) represent a new and powerful tool for the identification of small organic molecules. Briefly, a collection of chemical compounds are individually coupled to unique DNA fragments which serve as identification barcodes. The first part of this thesis reports on the design and synthesis of three novel scaffolds, including benzodiazepines **1**, **2** and **3**. The benzodiazepine **3** was used for the synthesis of a DNA-encoded library (**103**) containing 9588 molecules. The scaffold was substituted with functional groups (Figure I-19) allowing for combinatorial DNA-encoded library synthesis using high-yielding DNA-compatible reactions with broad reactant scope: amide coupling and Cu (I)-catalyzed azide–alkyne cycloaddition. The substituents for library synthesis were selected with the help of chemoinformatic tools to avoid unwanted structural motifs, such as PAINS,⁹² and to control the physicochemical properties of the library members. The structural diversity of the substituents used for the library synthesis was refined following the filtering steps. The chemical space of the DNA-encoded library showed only low similarity to three existing databases of bioactive compounds, indicating that this library covers novel chemical space. The synthesis of the DEL was initiated with the coupling of the scaffold **3** to a 5'-aminolinker-modified DNA on solid support. The first set of 114 carboxylic acid building blocks was appended to the DNA-scaffold conjugates by amide synthesis on the solid phase. These DNA conjugates were purified by ion-pair chromatography, and characterized by HPLC and LC-ESI-MS.

Approximately 80% of the carboxylic acids yielded the target products, though with variable yields justifying the effort of purifying this first set of DNA conjugates. Library synthesis began with combinatorial ligation of coding dsDNA sequences with four-nucleotide overhangs by an optimized T4 DNA ligation protocol. The synthesis of the library was achieved with the introduction of a set of 102 validated azide building blocks. Finally, one library member containing desthiobiotin was used to validate the synthesis and encoding strategy through successful selection of its target protein, streptavidin.⁴⁴ Screening of the synthesized DNA-encoded libraries to identify novel binders for target proteins is currently in progress.

Modulating protein–protein interactions (PPIs) has emerged as a major point of research in the drug discovery process due to its prominence in molecular recognition and signaling pathways. The identification of molecules that can adequately disrupt these interactions with sufficient affinity is a challenge given the limited ability of small molecules to block PPIs.^{124, 127} One appealing strategy is the generation of macrocycles, given that irregular structures participate widely in protein-protein interactions (PPIs). This thesis also presents the successful synthesis of a macrocycle directly on DNA using the synthetic potential of the metathesis reaction. This synthesis is based on the fact that Fmoc chemistry is compatible with DNA on solid support, preserving the integrity of the DNA. Experiments were performed on a 5'-aminolinker modified hexathymidine oligonucleotide, named "hexT", allowing the use of organic solvents for the synthesis of a peptide with unnatural amino acids, which was subsequently cyclized by RCM. This investigation proves that the macrocyclization of peptides by RCM is tolerated by a hexT adapter oligonucleotide designed for this purpose, opening the door for the synthesis of a DNA-encoded chemical library based on macrocycles that can be screened against targets involved in protein-protein interactions.

The main point of criticism of the current DEL-technology is that the chemical space covered by DELs is very limited due to the narrow repertory of chemical reactions that can be used at present. The low diversity of DELs reduces the probability of finding a bioactive substance. In addition, many drugs present so-called heterocyclic structures. This means that many drugs contain, as a central structural motif, a ring-shaped structure with at least one atom which is not a carbon. Many of these structures cannot be synthesized using the DNA-encoded format with the state-of-the-art of this technique. Therefore, several interesting scaffolds capable to generate novel chemical libraries remain unusable due to the incompatibility of DNA with many chemical transformations.

This thesis proposes a strategy to overcome the limitation on the incompatibility of DNA to protein transformations through the development of the “oligoThymidine initiated DNA-Encoded Chemistry” (TiDEC) strategy. This thesis reports a successful synthesis strategy for β -carbolines on DNA, based on a 5'-aminolinker modified hexathymidine adapter oligonucleotide which allows the initiation of DEL-synthesis with acid-catalyzed heterocycle-forming reactions: TiDEC, oligothymidine initiated DNA-encoded chemistry. The possibility of performing acid catalysis in a way which can lead to a library of DNA-encodable compounds breaks the limitations of chemistry methodology that were present in the field. An important heterocyclic structure in medicinal chemistry and drug research is the β -carboline scaffold. It can be accessed through the Brønsted acid catalyzed Pictet-Spengler reaction.¹⁸⁹ It is the core scaffold of many natural products including the pharmacologically active alkaloids ajmalicine and reserpine,^{169, 224} and also of both synthetic drugs and drug candidates such as tadalafil,²¹⁶ AZD9496,²²⁵ and the antimalarial candidate NITD609.²¹⁵ There are plenty of interesting reactions, e.g. the Povarov reaction, which could be applied in a similar manner using acid-catalysis to yield different heterocycles core structures. The TiDEC strategy aims at overcoming key technical limitations in the field of DNA-encoded chemical libraries and opens a door for acid catalysis covering a huge area of chemical space, which now becomes part of DEL chemical space as well. The hexT- β -carboline conjugates obtained by Brønsted acid catalyzed Pictet-Spengler reaction were successfully encoded by DNA ligation and will be used for the synthesis of a thymidine initiated DNA-encoded library based on β -carboline scaffold. The resulting library will subsequently be used to identify binders for several target proteins.

In conclusion, this thesis shows the successful application of DNA-encoded chemical library based on benzodiazepine and suggests this technology to be a rapid and cost-effective tool for small-molecule ligand discovery. However, there are plenty of opportunities for further improvements in the performance of DELs for hit discovery purposes. The synthesis of functionally and chemically diverse libraries will require the exploration of more robust DNA-compatible chemical reactions. For this purpose, library synthesis on solid-phase appears to be particularly attractive, as it allows the driving of chemical reactions to completion and expanding the scope of DNA-compatible chemical reactions to water-free conditions. In addition, the utilization of (i) more diverse privileged scaffolds and new DNA-compatible reactions for library construction, (ii) the synthesis of high-quality DELs with increased complexity and diversity, as well as (iii) the application of sophisticated affinity selection protocols, will help to exhaust the full potential of DNA-encoded combinatorial chemistry, which provides attractive starting points for hit-to-lead development.

Chapter VI: EXPERIMENTAL PART

VI-1. Materials and instruments

Unless otherwise noted, all chemicals were purchased from Sigma-Aldrich (Taufkirchen, Germany), Bachem (Bubendorf, Switzerland), Thermo Fisher Scientific (Karlsruhe, Germany), AppliChem (Darmstadt, Germany), and VWR (Langenfeld, Germany). The building blocks shown in tables 15 and 16 were selected and purchased from the Aldrich Market Select building blocks catalogue (version 2013/3). Enzymes were purchased from Thermo Fisher Scientific (PNK), Biozym (T4 DNA ligase rapid), NEB (E.coRI), Roche (FastStart Universal Probe Master (Rox) for qPCR containing SYBR GREEN and Taq polymerase and dNTPs) and Eurogentec (Takyon™ No Rox Probe MasterMix dTTP). 5'-Aminolinker-modified DNA oligonucleotides attached to controlled pore glass solid phase (CPG, 1000 Å) were synthesized by IBA (Goettingen, Germany); unmodified DNA-oligonucleotides were purchased from Integrated DNA Technologies (Integrated DNA Technologies, Leuven, Belgium). Controlled pore glass solid phase was dried on a synthesis column plugged onto a vacuum manifold (Vac-Man®, Promega). Oligonucleotide-small molecule conjugates were purified by ion pair reverse-phase high-pressure liquid chromatography (HPLC, Shimadzu Prominence) using a C18 stationary phase (Phenomenex, Gemini; 5 µm, C18, 110 Å, 100*10.0 mm) and a gradient of 100 mM aqueous triethylammonium acetate/MeOH. The triethylammonium acetate buffer was set to pH 8. Oligonucleotide-small molecule conjugates were analyzed by ion pair reverse phase ultra-pressure liquid chromatography (UPLC, Agilent Technologies 1100) using a C18 stationary phase (Waters, Acquity UPLC® BEH, 1.7 µm, C18, 50*2.1 mm) and a gradient of 100 mM aqueous triethylammonium acetate/MeOH. Oligonucleotide concentrations were quantitated by UV spectroscopy using a spectrophotometer (NanoDrop 2000, Thermo Fisher Scientific). Oligonucleotides were analyzed by MALDI-TOF/TOF-MS (Bruker Daltonics) using THAP or 3-HPA matrix (Dichrom). ¹H-NMR-spectra were measured at 400 or 500 MHz on a Bruker DRX400 or Inova 500 spectrometer, respectively. ¹³C-NMR-spectra were measured at 101 MHz or 126 MHz on a Bruker DRX400 or Inova 500 spectrometer, respectively. The pure substance was dissolved in deuterated chloroform (CDCl₃, 99.8 %, VWR) or dimethyl sulfoxide-d₆ (DMSO-d₆, 99.8 %, VWR, Langenfeld, Germany). Chemical shifts are listed relative to the deuterated solvent. Each proton signal was analyzed regarding its multiplicity, coupling constant J [Hz] and the amount of protons. The multiplicity was abbreviated as follows: s = sigulett, d = duplet, t = triplet, q = quartet, quint = quintet, m = multiplet and br = broad signal.

Silica gel chromatography was performed on NORMASIL 60 silica gel 40-63 μm (VWR, Langenfeld, Germany); thin layer chromatography was performed on aluminium-backed silica gel 60 F₂₅₄ plates provided by (Merck Millipore, Darmstadt, Germany). LC-MS analysis of low-molecular weight compounds was performed on reverse-phase high-pressure liquid chromatography (HPLC, Shimadzu Prominence) using a C18 column stationary phase (Phenomenex, Luna; 5 μm , C18, 100 Å) and MeOH/1% aq. formic acid, 50:50 to 100:0 over 13 min. The following chemicals and materials were used for the experiments performed on DNA.

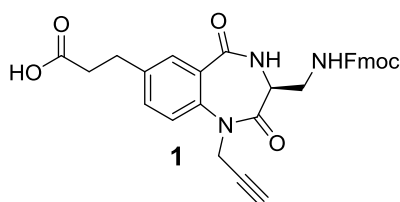
Table 13: List of chemicals and materials.

Chemicals and materials	Manufacturer
Acetic acid	J. T. Baker
Acetic anhydride	VWR
Ammonia	VWR
Eppendorf tubes (0.5 ml, 1.5 ml, 2 ml)	Eppendorf/Sarstedt
Ethanol (abs.)	Sigma-Aldrich
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich
Hydrochloric acid (HCl)	J. T. Baker
Isopropanol	J. T. Baker
Methylamine	VWR
N-Methylimidazole	VWR
Syringe needle 100 Sterican®	B. Braun Melsungen AG
Sodium chlorid (NaCl)	AnalaR® NORMAPUR® VWR
Sodium hydroxide	J. T. Baker
Oligonucleotides	IBA life sciences/ Intregrated DNA Technologies
Pipette tips	Nerbe Plus
Pipettes (2.5, 10, 20, 100, 200, 1000 µl)	Eppendorf
Syringes	B. Braun Melsungen AG
Tris (tris(hydroxymethyl)aminomethane	Roth

VI-2. Synthesis of DEL9588 DNA-encoded library based on one scaffold

VI-2.1. Preparation of benzodiazepine scaffolds for library synthesis

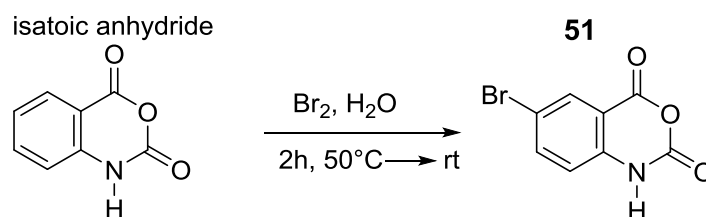
VI-2.1.1. Synthesis of the trifunctionalized benzo-1,4-diazepine-2,5-dione **1**



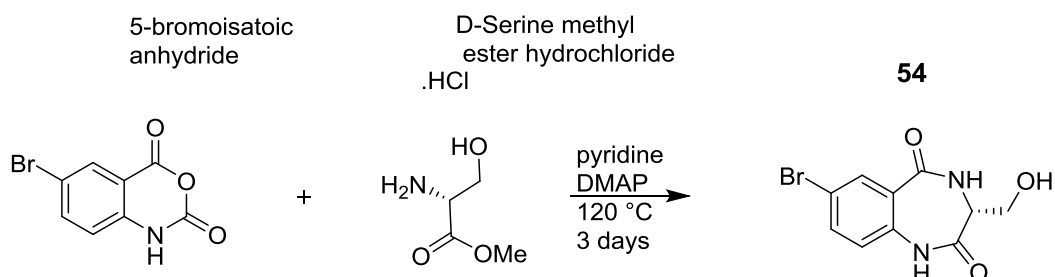
Step 1: Synthesis of the starting material

Structure name: 6-bromo-2*H*-benzo[*d*][1,3]oxazine-2,4(1*H*)-dione **51**

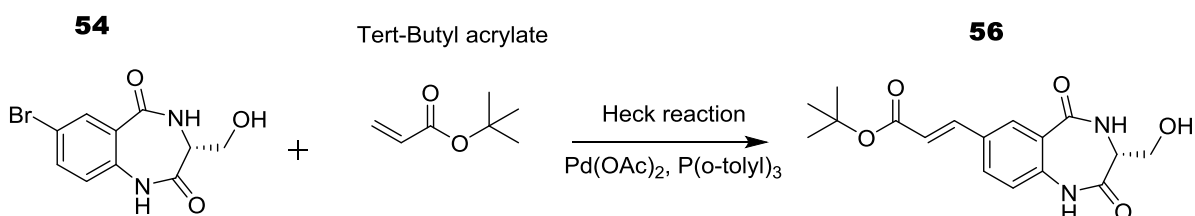
Scheme:



Procedure: Bromine (10.8 g, 1.1 eq, 67.5 mmol) was added dropwise at 50 °C to a suspension of isatoic anhydride (10 g, 1 eq, 61.3 mmol) in water (300 mL). The mixture was stirred for 1 hour at 50 °C then it was allowed to cool to room temperature. The solid filtered off, washed with water, and with acetone giving 5-bromoisatoic anhydride as an off-white solid (6.9 g, 46.5 % yield) MS (ESI): calc. 242.03, found ($[M+H]^+$) 243.03. Purity (HPLC): 97%. **¹H-NMR:** (500 MHz, DMSO-*d*₆) δ ppm: 8.04 (br. s, 1H), 7.99 (d, ⁴*J* = 2.4 Hz, 1H), 7.86 (dd, ³*J* = 8.6 Hz and ⁴*J* = 2.4 Hz, 1H), 7.11 (d, ³*J* = 8.6 Hz, 1H). **¹³C-NMR:** (126 MHz, DMSO-*d*₆) δ ppm: 158.9, 146.9, 140.7, 139.4, 130.7, 117.8, 114.7, 112.4.

Step 2: Benzodiazepine formation**Structure name:** 6-bromo-2*H*-benzo[*d*][1,3]oxazine-2,4(1*H*)-dione **54****Scheme:**

Procedure: A stirred solution of 5-bromoisatoic anhydride (5 g, 1 eq, 20.7 mmol), D-Serine methyl ester hydrochloride (3.2 g, 1 eq, 20.7 mmol) and DMAP (126.4 mg, 0.05 eq, 1.03 mmol) in anhydrous pyridine (5 mL) was heated under reflux for 4 days. The reaction mixture was allowed to cool to room temperature and the solvents were evaporated. The residue was dissolved in EtOAc, washed with 10% HCl, brine, dried over Na₂SO₄ and concentrated *in vacuo*, giving crude OB 001 as a dark brown solid. This crude material was purified by column chromatography (CH₂Cl₂/MeOH), affording OB 001 as a beige powder (1.5 g, 25 % yield) MS (ESI): calc. 285.09, found ([M+H]⁺) 284.93. Purity (HPLC): 96%. ¹H-NMR: (500 MHz, DMSO-*d*₆) δ ppm: 8.48 (d, ⁴*J*= 5.02 Hz, 1 H), 7.84 (d, ³*J*= 2.26 Hz, 1 H), 7.68 - 7.75 (m, 1 H), 6.99 - 7.12 (m, 1 H), 4.65 - 4.86 (m, 1 H), 3.71 - 3.81 (m, 2 H), 3.61 - 3.68 (m, 1 H), 3.18 (d, ³*J*= 3.76 Hz, 2 H).

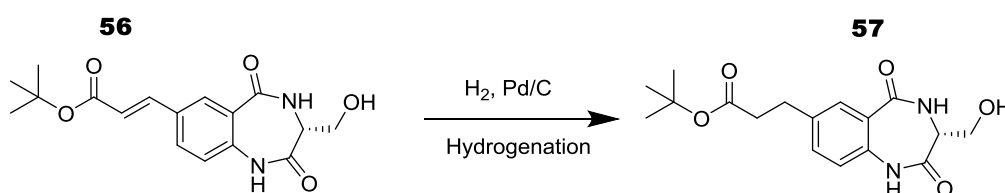
Step 3: Heck reaction**Structure name:** *tert*-butyl (*R,E*)-3-(3-(hydroxymethyl)-2,5-dioxo-2,3,4,5-tetrahydro-1*H*-benzo[*e*][1,4]diazepin-7-yl)acrylate **56****Scheme:**

Procedure: Palladium (II) acetate (47 mg, 0.05 eq, 0.21 mmol) was added to a degassed solution of OB 001 (1.4 g, 1 eq, 4.2 mmol) tri-*o*-tolylphosphine (127.8 g, 0.1 eq, 0.42 mmol), triethylamine (1.76 mL, 3 eq, 12.6 mmol) and *tert*-butyl acrylate (1.84 mL, 3 eq, 12,6 mmol) in dry acetonitrile (10 mL). The mixture was stirred for 18 h at 100 °C. The solvent was evaporated under reduced pressure and the residue was purified by flash chromatography (CH₂Cl₂/MeOH) to give **56** as a light white solid (893 mg, 64 % yield) MS (ESI): calc. 332.35, found ([M+H]⁺) 332.95. Purity (HPLC): 96%. ¹H-NMR: (500 MHz, CDCl₃-*d*1) δ ppm: 8.21 (d, ⁴J= 4.15 Hz, 1 H), 8.15 (br. s., 1 H), 7.67 (d, ³J= 8.41 Hz, 1 H), 7.59 (d, ³J= 8.51 Hz, 1 H), 7.06 (d, ³J= 8.03 Hz, 1 H), 6.95 (d, ³J= 7.65 Hz, 1 H), 6.44 (d, ²J=15.68 Hz, 1 H), 5.33 (s, 1 H), 4.26 - 4.37 (m, 1 H), 3.81 - 3.96 (m, 2 H), 3.26 - 3.34 (m, 1 H), 3.19 (dd, ²J=11.66 Hz and ³J= 6.69 Hz, 1 H), 3.12 (dd, ²J= 13.00 Hz and ³J= 5.74 Hz, 1 H), 3.00 (t, ³J= 7.46 Hz, 1 H), 2.13 (s, 1 H), 1.42 - 1.50 (s, 9 H).

Step 4: Hydrogenation of the alkene

Structure name: *tert*-butyl (*R*)-3-(3-(hydroxymethyl)-2,5-dioxo-2,3,4,5-tetrahydro-1*H*-benzo[*e*][1,4]diazepin-7-yl)propanoate **57**

Scheme:



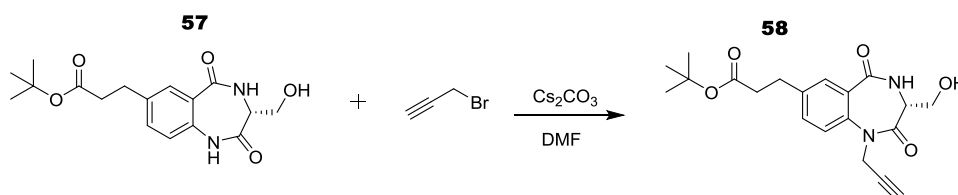
Procedure: To a stirred solution of compound **56** (570 mg, 1 eq, 1.7 mmol) in dry methanol (5 mL) at room temperature under hydrogen was added palladium on carbon (181 mg, 1 eq, 1.7 mmol). The reaction mixture was stirred under hydrogen for 18 hours and then filtered. The filtrate was concentrated and purified by flash column chromatography (CH₂Cl₂/MeOH 2:1 →5:1 →10:1), affording pure **57** as a white off solid. (443.4 g, 78 % yield) MS (ESI): calc. 334.37, found ([M+H]⁺) 334.76. Purity (HPLC): 96%. ¹H-NMR: (500 MHz, DMSO-*d*6) δ ppm: 7.59 (d, ⁴J= 1.76 Hz, 1 H), 7.38 (dd, ³J= 8.28 Hz and ⁴J= 2.01 Hz, 1 H), 7.02 (d, ³J= 8.03 Hz, 1 H), 3.77 (d, ³J= 5.52 Hz, 1 H), 3.63 (dd, ³J= 4.27 Hz, 1 H), 3.56 - 3.67 (m, 2 H), 2.52 (t, J=7.1, 2H), 2.83 (t, ³J= 7.40 Hz, 2 H), 1.36 (s, 9 H).

Optimisation: A Pd/C-catalyzed hydrogenation using Palladium on carbon as a catalyst selectively reduces the alkene functionality without hydrogenolysis of aromatic carbonyls. The amount of palladium on carbon should be not below 1eq. Increasing the reaction time also increases the yield.

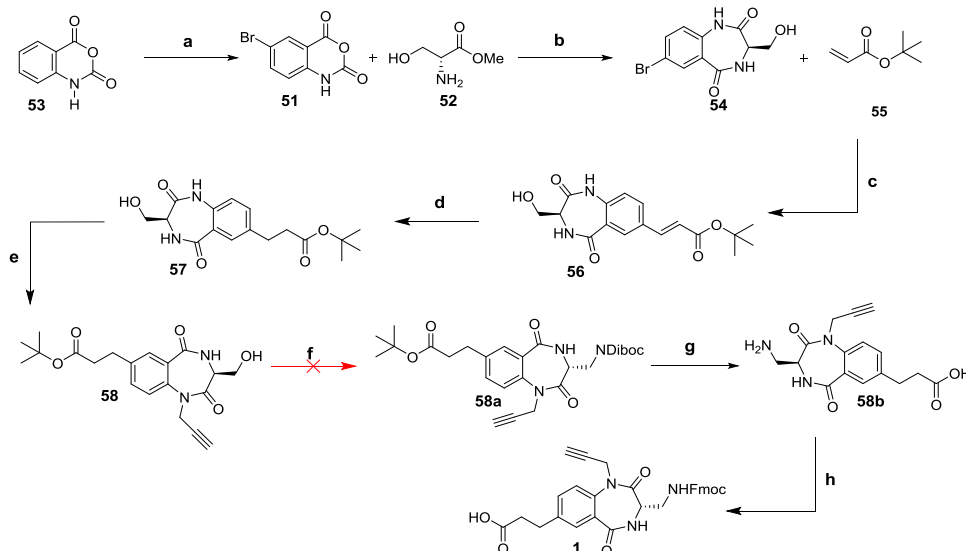
Step 5: Alkylation of the N-amine

Structure name: *tert*-butyl (*R*)-3-(3-(hydroxymethyl)-2,5-dioxo-1-(prop-2-yn-1-yl)-2,3,4,5-tetrahydro-1*H*-benzo[*e*][1,4]diazepin-7-yl)propanoate **58**

Scheme:



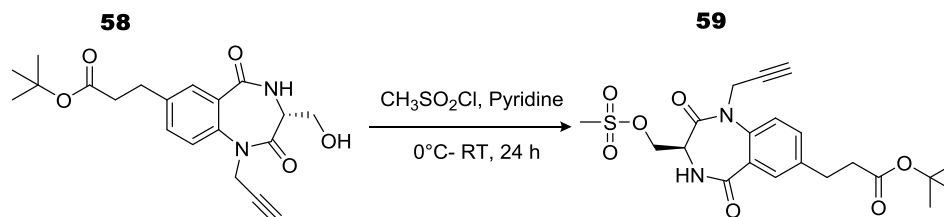
Procedure: To a degassed solution of **57** (500 mg, 1 eq, 1.5 mmol) in DMF was added Cs_2CO_3 (977 mg, 2 eq, 3 mmol). The mixture was stirred at room temperature. After 10 min propargyl bromide (200 mL, 1.5 eq, 2.25 mmol) was added slowly and the stirring continued at an ambient temperature for 18 h at room temperature. After completion of the reaction, the solvent was evaporated to afford an oil, which was poured into ice-cooled water (25 mL) and the organic layer extracted with ethyl acetate. The organic layer was washed with water, brine, dried over anhydrous sodium sulphate and concentrated on rotary evaporator. The residue was purified by flash chromatography to give **58** as yellow oil (282 mg, 50.5 % yield) MS (ESI): calc. 372.41, found ($[\text{M}+\text{H}]^+$) 372.39. Purity (HPLC): 96%. **¹H-NMR:** (500 MHz, CDCl_3 -*d*1) δ ppm: 7.66 (d, $^4J = 2.01$ Hz, 1 H), 7.47 (d, $^3J = 8.39$ Hz, 1H), 7.34 (dd, $^3J = 8.39$ Hz and $^4J = 2.18$ Hz, 1 H), 4.71 (dd, $^2J = 17.37$ Hz and $^3J = 2.43$ Hz, 1 H), 4.58 (br. s., 1 H), 4.14 - 4.21 (m, 2 H), 3.78 (d, $^3J = 1.68$ Hz, 1 H), 3.58 - 3.62 (m, 1 H), 2.84 - 2.89 (m, 2 H), 2.47 - 2.51 (m, 2 H), 2.26 (t, $^3J = 2.35$ Hz, 1 H), 1.36 (s, 9 H).



Step 6: Mesylate formation

Structure name: *tert*-butyl (S)-3-(3-(((methylsulfonyl)oxy)methyl)-2,5-dioxo-1-(prop-2-yn-1-yl)-2,3,4,5-tetrahydro-1*H*-benzo[*e*][1,4]diazepin-7-yl)propanoate **59**

Scheme:

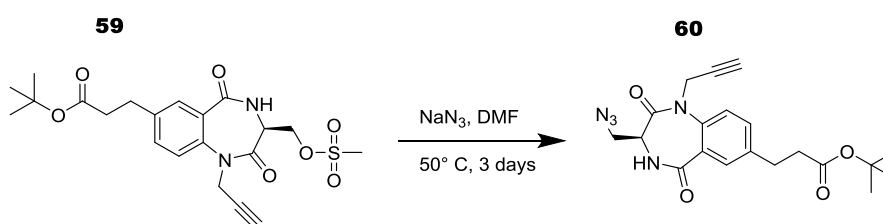


Procedure: Compound **58** (200 mg, 1 eq, 0.54 mmol) was dissolved in dichloromethane (3 mL). After addition of pyridine (3 mL), the solution was cooled to -10°C . Methanesulfonyl chloride (62.7 μL , 1.5 eq, 0.81 mmol) was slowly added with stirring. The mixture was then stirred for 1 h at -10°C and for 10 h at room temperature. The solution was washed with sat. NaHCO_3 solution and the aqueous layer was extracted with dichloromethane. The combined organic layers were washed with 10% citric acid and water and dried over MgSO_4 . The solvent was removed under reduced pressure. (197 mg, 81 % yield) MS (ESI): calc. 450.51, found ($[\text{M}+\text{H}]^+$) 451.39. Purity (HPLC): 96%.

Step 7: Azide formation

Structure name: *tert*-butyl (S)-3-(3-(azidomethyl)-2,5-dioxo-1-(prop-2-yn-1-yl)-2,3,4,5-tetrahydro-1*H*-benzo[*e*][1,4]diazepin-7-yl)propanoate **60**

Scheme:



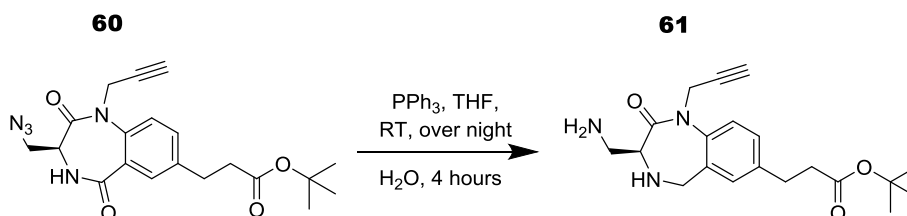
Procedure: The crude mesylate **59** (195 mg, 1 eq, 0.5 mmol) was dissolved in *N,N*-dimethylformamide (3 mL). Sodium azide (617.6 mg, 19 eq, 9.5 mmol) was added, followed by stirring at 50°C for 72 h. Afterwards, the mixture was poured into ice-cooled water and the aqueous phase was extracted with dichloromethane. After the organic layer had been dried over Na_2SO_4 , the solvent was evaporated.

Purification by flash chromatography (CH₂Cl₂/CH₃OH 10:1) afforded product **60**. (143 mg, 72 % yield) MS (ESI): calc. 397.43, found ([M+H]⁺) 398.57. Purity (HPLC): 96%. ¹H NMR (500 MHz, CDCl₃-d₁) δ ppm: 9.10 (br, s., 1H), 8.13 (d, ⁴J= 1.99 Hz, 1H), 7.62 (dd, ³J= 8.34 Hz and ⁴J= 1.99 Hz, 1H), 7.55 (d, ³J= 5.89 Hz, 1H), 7.10 (m, 1H), 6.39 (d, ³J= 15.90 Hz, 1H), 4.37 (m, 1H; CH), 4.20 (dd, ³J= 1.99 Hz and ³J= 9.14 Hz, 1H), 3.82 (dd, ²J= 12.98 Hz and ³J= 5.14 Hz, 1H), 3.73 (m, 1H), 3.10 (dd, ²J= 13.91 Hz and ³J= 1.99 Hz, 1H), 2.38 (m, 1H), 1.52 (s, 9H).

Step 8: Staudinger azide reduction

Structure name: *tert*-butyl (S)-3-(3-(aminomethyl)-2-oxo-1-(prop-2-yn-1-yl)-2,3,4,5-tetrahydro-1H-benzo[e][1,4]diazepin-7-yl)propanoate **61**

Scheme:

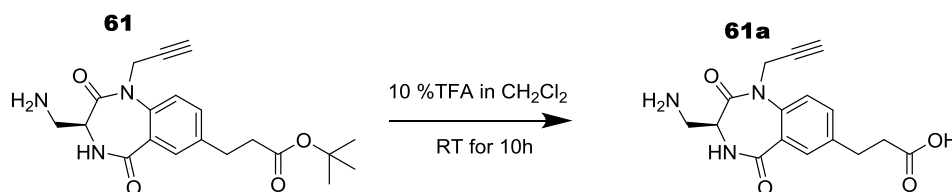


Procedure: The compound **60** (130 mg, 1 eq, 0.35 mmol) was dissolved in tetrahydrofuran (5 mL) and triphenylphosphine (110.2 mg, 1.2 eq, 0.42 mmol) was added. The reaction mixture was stirred at room temperature for 24 h and then water (2 mL) was added. Stirring was continued overnight and the solvents were removed in vacuo. The residue was purified by chromatography on silica using dichloromethane/methanol (98/2) followed by dichloromethane/methanol (95/5) to afford **61** as an off white solid. (90 mg, 69 % yield) MS (ESI): calc. 371.43, found ([M+H]⁺) 372.33. Purity (HPLC): 96%. ¹H-NMR: (500 MHz, CDCl₃-d₁): 8.03 (d, ⁴J= 1.99 Hz, 1H), 7.74 (dd, ³J= 8.34 Hz and ⁴J= 1.99 Hz, 1H), 7.55 (d, ³J= 15.89 Hz, 1H), 7.15 (d, ³J= 8.74 Hz, 1H), 6.41 (d, ³J= 16.29 Hz, 1H), 4.34 (dd, ³J= 3.18 Hz and ³J= 8.35 Hz, 1H), 3.88 (dd, ²J= 11.92 Hz and ³J= 6.75 Hz, 1H), 3.76 (m, 1H), 3.53 (dd, ²J= 11.92 Hz and ³J= 7.15 Hz, 1H), 2.93 (m, 1H), 2.04 (m, 1H), 1.52 (s, 9H).

Step 9: Cleavage of protective groups

Structure name: (S)-3-(3-(aminomethyl)-2,5-dioxo-1-(prop-2-yn-1-yl)-2,3,4,5-tetrahydro-1H-benzo[e][1,4]diazepin-7-yl)propanoic acid **61a**

Scheme:

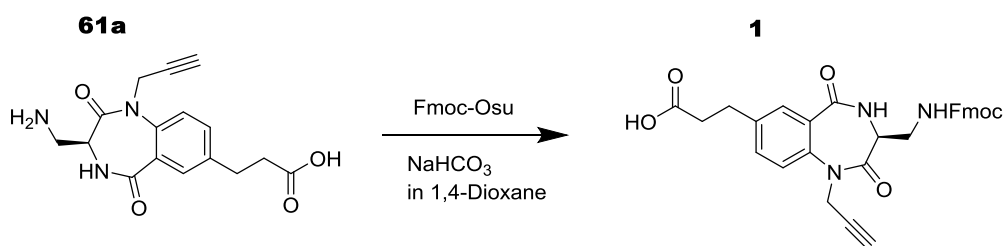


Procedure: 300 μ L TFA was added to a solution of compound **61** (80 mg, 1 eq, 0.25 mmol) in dry dichloromethane (3 mL), followed by stirring overnight at room temperature. The solvent was removed in vacuo and the residue was triturated with diethyl ether to afford **61a** (70 mg, 88 % yield) MS (ESI): calc. 315.32, found ($[M+H]^+$) 316.12. Purity (HPLC): 96%. $^1\text{H-NMR}$: (500 MHz, DMSO): 10.79 (br., s., 1H), 7.97 (d, $^4J= 1.77$ Hz, 1H), 7.89 (dd, $^4J= 1.77$ Hz and $^3J= 8.48$ Hz, 1H), 7.79 (br., s., 1H), 7.55 (d, $^3J= 15.89$ Hz, 1H), 7.19 (d, $^3J= 8.47$ Hz, 1H), 6.48 (d, $^3J= 15.89$ Hz, 1H), 4.29 (dd, $^3J= 3.53$ Hz and $^3J= 8.83$ Hz, 1H), 4.16 (m, 1H; CH), 3.96 (dd, $^2J(\text{H,H})= 12.36$ Hz and $^3J= 5.65$ Hz, 1 H), 3.43 (d, $^2J= 12.36$ Hz, 1H), 2.62 (m, 1H), 2.35 (m, 1H).

Step 10: Fmoc protection of the secondary amine

Structure name: (S)-3-(3-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)methyl)-2,5-dioxo-1-(prop-2-yn-1-yl)-2,3,4,5-tetrahydro-1H-benzo[e][1,4]diazepin-7-yl)propanoic acid **1**

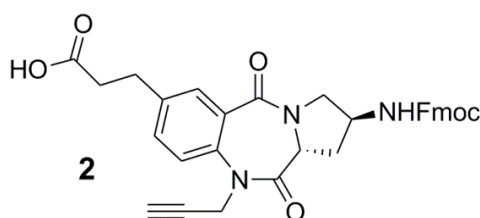
Scheme:



Procedure: The amine **61a** (60 mg, 1 eq, 0.2 mmol) is dissolved in water (2 mL) and sodium bicarbonate (34 mg, 2 eq, 0.4 mmol) is added with stirring. The resulting solution is cooled to 5°C and Fmoc-OSu (101.2 mg, 1.5 eq, 0.3 mmol) is added slowly as a solution in *para*-dioxane also cooled (3 mL). The resulting mixture is stirred at 0° for 1 h and allowed to warm to room temperature overnight.

Water is then added and the aqueous layer is extracted two times with EtOAc. The organic layer is back extracted twice with saturated sodium bicarbonate solution. The combined aqueous layers are acidified to a pH of 1 with 10% HCl, and then extracted 3 times with EtOAc. The combined organic layers are dried (sodium sulfate) and concentrated *in vacuo*. The resulting residue was then be purified by column chromatography (69.9 mg, 65 % yield) MS (ESI): calc. 537.56, found ($[M+H]^+$) 538.2. Purity (HPLC): 96%. **¹H-NMR:** (500 MHz, CDCl₃-d₁): 8.03 (d, ⁴J= 1.99 Hz, 1H), 7.74 (dd, ³J= 8.34 Hz and ⁴J= 1.99 Hz, 1H), 7.55 (d, ³J= 15.89 Hz, 1H), 7.15 (d, ³J= 8.74 Hz, 1H), 6.41 (d, ³J= 16.29 Hz, 1H), 4.34 (dd, ³J= 3.18 Hz and ³J= 8.35 Hz, 1H), 3.88 (dd, ²J= 11.92 Hz and ³J= 6.75 Hz, 1H), 3.76 (m, 1H), 3.53 (dd, ²J= 11.92 Hz and ³J= 7.15 Hz, 1H), 2.93 (m, 1H), 2.04 (m, 1H), 1.52 (s, 9H). **¹³C-NMR:** (126 MHz, DMSO-d₆) δ ppm: 174.4, 172.3, 171.1, 157.8, 143.6, 142.6, 131.3, 135.1, 136.2, 126.7, 126.6, 126.2, 125.2, 123.1, 120.5, 116.1, 73.2, 72.1, 67.3, 58.1, 47.1, 43.9, 34.4, 34.2, 29.8.

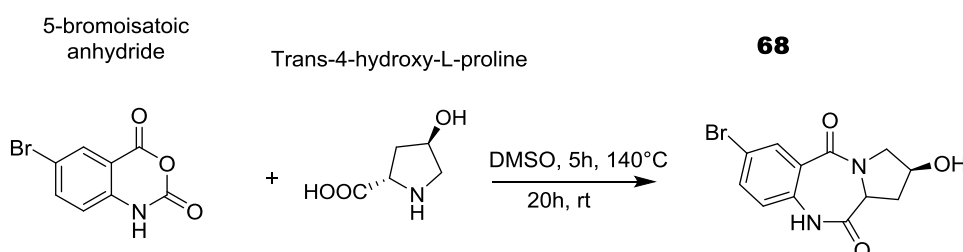
VI-2.2. Synthesis of the functionalized benzo-1,4-diazepine-2,5-dione **2**



Step 1: Benzodiazepine formation

Structure name: (2*S*)-7-bromo-2-hydroxy-2,3-dihydro-1*H*-benzo[*e*]pyrrolo[1,2-*a*][1,4]diazepine-5,11(10*H*,11*aH*)-dione **68**

Scheme:



Procedure: A stirred solution of 5-bromoisatoic anhydride (4 g, 1 eq, 16.53 mmol) and (2*S*, 4*R*)-4-hydroxypyrrolidine-2-carboxylic acid (3.25 g, 1.5 eq, 24.8 mmol) in dry DMSO (25 ml) was heated for 5 h at 140°C. Stirring was continued at room temperature for 16 h. The solution was added to 300 mL ice-cooled water and the aqueous phase was extracted with ethyl acetate (4 x 100 mL). The organic layer was washed with water (2 x 50 mL), dried over Na₂SO₄, and evaporated to give **68** as a slightly yellow solid (3.8 g, 74% yield). MS (ESI): calc. 311.13, found 311.01. Purity (HPLC): 98%.

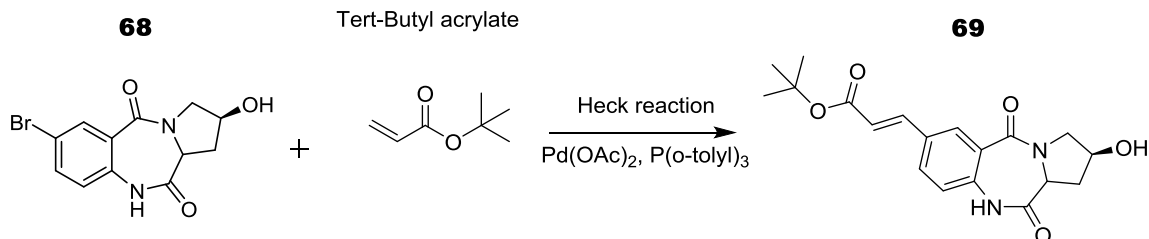
¹H-NMR: (500 MHz, CDCl₃-*d*1) δ ppm: 8.11 (d, ⁴J= 2.35 Hz, 1 H), 8.03 (br. s., 1 H), 7.61 (dd, ³J= 8.48 Hz and ⁴J= 2.43 Hz, 1 H), 7.61 (dd, ³J=8.48 Hz and ⁴J= 2.43 Hz, 1 H), 6.91 (d, ³J= 8.56 Hz, 1 H), 4.61 - 4.69 (m, 1 H), 4.29 - 4.35 (m, 1 H), 3.96 - 4.04 (m, 1 H), 3.68 (dd, ²J= 12.76 Hz and ³J= 4.36 Hz, 1 H), 2.86 - 2.98 (m, 1 H), 2.23 (dddd, ³J= 9.84 Hz, ³J= 7.95 Hz, ⁴J= 4.11 Hz and ⁴J= 1.93 Hz, 1 H). **¹³C-NMR:** (126 MHz, DMSO-*d*6) δ ppm: 169.6, 164.5, 135.6, 130.9, 126.4, 118.5, 117.7, 68.2, 62.2, 56.4, 34.3.

Optimization: Amount of (2*S*,4*R*)-4-hydroxypyrrolidine-2-carboxylic acid not below 1.5 eq. Heating longer than 3 h. Heating 5 hours under reflux increases the yield.

Step 2: Heck reaction

Structure name: *tert*-butyl (*E*)-3-((2*S*)-2-hydroxy-5,11-dioxo-2,3,5,10,11,11a-hexahydro-1*H*-benzo[*e*]pyrrolo[1,2-*a*][1,4]diazepin-7-yl)acrylate **69**

Scheme:



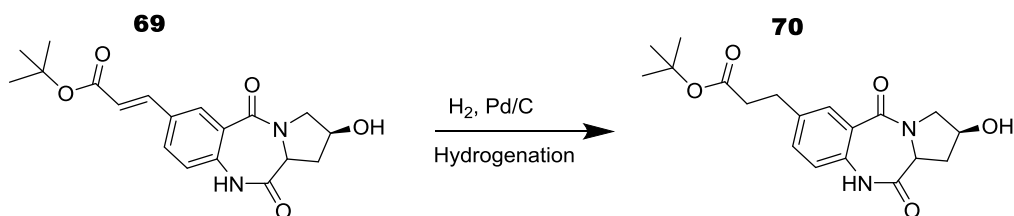
Procedure: Palladium (II) acetate (71.8 mg, 0.05 eq, 0.32 mmol) was added to a degassed solution of **68** (2 g, 1 eq, 6.4 mmol) tri-*o*-tolylphosphine (195 mg, 0.1 eq, 0.64 mmol), triethylamine (2.7 mL, 3 eq, 19.2 mmol) and *tert*-butyl acrylate (2.8 mL, 3 eq, 19.2 mmol) in dry acetonitrile (10 mL). The mixture was stirred for 18 h at 100 °C. The solvent was evaporated under reduced pressure and the residue was purified by flash chromatography (CH₂Cl₂/MeOH) to give **69** as a light white solid (1.8 g, 78.6 % yield) MS (ESI): calc. 358.39, found 358.93. Purity (HPLC): 96%. ¹H-NMR: (500 MHz, CDCl₃-*d*1) δ ppm: 8.98 (s, 1 H), 8.04 (s, 1 H), 7.56 - 7.65 (m, 1 H), 7.51 (d, ²*J*= 15.81 Hz, 1 H), 7.09 (d, ³*J*= 8.53 Hz, 1 H), 6.35 (d, ²*J*= 15.81 Hz, 1 H), 4.63 (br. s., 1 H), 4.34 (t, ³*J*= 7.28 Hz, 1 H), 4.04 (d, ²*J*= 13.05 Hz, 1 H), 3.66 (dd, ²*J*=12.55 Hz and ³*J*= 4.27 Hz, 1 H), 3.15 (s, 1 H), 2.90 (s, 1 H), 1.54 (s, 9 H). ¹³C-NMR: (126 MHz, DMSO-*d*6) δ ppm: 169.6, 166.5, 164.5, 145.1, 135.9, 131.8, 130.0, 123.8, 121.3, 116.2, 81.2, 68.2, 62.2, 56.4, 34.3, 28.8.

Optimization: It is important to keep the reaction under dry conditions; Using Tri-*o*-tolylphosphine at the place of PPh₃ increased pushes the yield of the reaction from 20 to more than 70 %. The reaction need to be monitored by TLC. Dependent to the starting material's amount, the reaction time can be increased.

Step 3: Hydrogenation of the alkene

Structure name: *tert*-butyl 3-((2*S*)-2-hydroxy-5,11-dioxo-2,3,5,10,11,11*a*-hexahydro-1*H*-benzo[*e*]pyrrolo[1,2-*a*][1,4]diazepin-7-yl)propanoate **70**

Scheme:



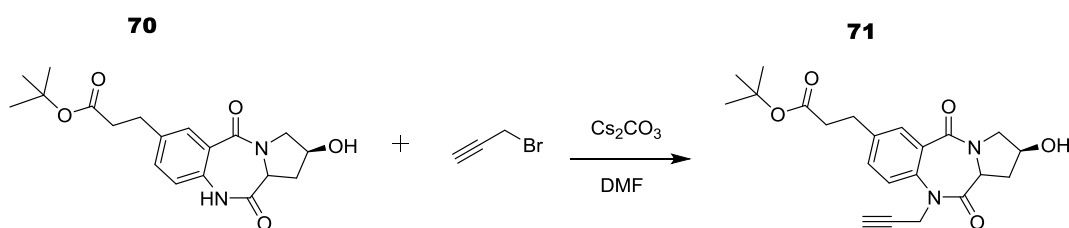
Procedure: To a stirred solution of **69** (1.5 g, 1 eq, 4.2 mmol) in methanol (8 mL) at room temperature under hydrogen was added palladium on carbon (446.9 mg, 1 eq, 4.2 mmol). The reaction mixture was stirred under hydrogen for 18 hours and then filtered. The filtrate was concentrated and purified by flash column chromatography (CH₂Cl₂/MeOH 10:1) affording pure **70** as a white off solid. (1.26 g, 83.2 % yield) MS (ESI): calc. 360.4, found ([M+H]⁺) 360.83. Purity (HPLC): 96%. ¹H-NMR: (500 MHz, CDCl₃-*d*1) δ ppm: 7.75 - 7.82 (m, 1 H), 7.35 (dd, ³J=8.03 Hz and ⁴J=1.76 Hz, 1 H), 6.96 (d, ³J= 8.03 Hz, 1 H), 4.58 - 4.66 (m, 1 H), 4.27 (t, ³J=7.15 Hz, 1 H), 3.98 (br. s., 1 H), 3.67 (dd, ²J= 12.80 Hz and ³J= 4.52 Hz, 1 H), 2.90 - 2.95 (m, 3 H), 2.56 (t, ³J= 7.53 Hz, 2 H), 1.44 (s, 10 H). ¹³C-NMR: (126 MHz, DMSO-*d*6) δ ppm: 171.7, 169.6, 164.5, 134.8, 132.9, 132.6, 128.9, 126.3, 120.3, 82.1, 68.2, 62.2, 56.4, 34.7, 34.3, 30.1, 28.7.

Optimization: Amount of palladium on carbon not below 1eq. Increasing the reaction time also increases the yield. It is important to keep the reaction under hydrogen.

Step 4: Alkylation of the N-amine

Structure name: *tert*-butyl 3-((2*S*)-2-hydroxy-5,11-dioxo-10-(prop-2-yn-1-yl)-2,3,5,10,11,11*a*-hexahydro-1*H*-benzo[*e*]pyrrolo[1,2-*a*][1,4]diazepin-7-yl)propanoate **71**

Scheme:



Procedure: To a degassed solution of **70** (1 g, 1 eq, 2.8 mmol) in DMF (5 mL) was added Cs₂CO₃ (1.82 g, 2 eq, 5.6 mmol). The mixture was stirred at room temperature. After 10 min propargyl bromide (0.5 mL, 2 eq, 5.6 mmol) was added slowly and the stirring continued at an ambient temperature for 18 h at room temperature. After completion of the reaction, the mixture was poured into ice-cooled water (25 mL) and the organic layer extracted with ethyl acetate. The organic layer was washed with water, brine, dried over anhydrous sodium sulphate and concentrated on rotary evaporator. The residue was purified by flash chromatography to give **71** as pale yellow powder. (688 g, 62 % yield) MS (ESI): calc. 398.45, found 398.91. Purity (HPLC): 96%. ¹H-NMR: (500 MHz, CDCl₃-d₁) δ ppm: 7.66 (d, ⁴J= 2.01 Hz, 1 H), 7.47 (d, ³J= 8.39 Hz, 1H), 7.34 (dd, ³J= 8.39 Hz and ⁴J= 2.18 Hz, 1 H), 4.71 (dd, ²J= 17.37 Hz and ³J= 2.43 Hz, 1 H), 4.58 (br. s., 1 H), 4.14 - 4.21 (m, 2 H), 3.78 (d, ⁴J= 1.68 Hz, 1 H), 3.58 - 3.62 (m, 1 H), 2.84 - 2.89 (m, 2 H), 2.73 (br. s., 2 H), 2.47 - 2.51 (m, 2 H), 2.26 (t, ⁴J= 2.35 Hz, 1 H), 1.36 (s, 9 H). ¹³C-NMR: (126 MHz, DMSO-d₆) δ ppm: 173.7, 171.6, 164.5, 135.7, 134.8, 128.9, 126.3, 124.1, 115.8, 73.2, 72.1, 68.2, 59.1, 56.4, 37.1, 34.7, 34.4, 30.1, 28.7.

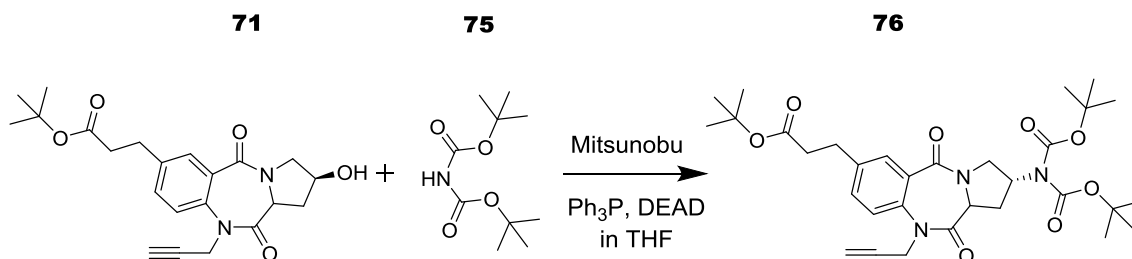
Optimization: Use of 1.2 eq NaH in THF.

In a 3-neck round-bottomed flask fitted with argon gas and septum was placed THF and sodium hydride (1.2 eq). The suspension was cooled to 0 °C and a solution of **70** in THF was added slowly. Ice bath was removed and the stirring continued at an ambient temperature for 1 hour. The clear solution was cooled down to 0 °C again and propargyl bromide added slowly. The reaction mixture was allowed to warm to an ambient temperature and stirring was continued overnight. After completion of reaction, reaction mixture was diluted in water and organics extracted with ethyl acetate. Organic layer was washed with water, brine, dried over anhydrous sodium sulfate and concentrated on rotary evaporator. The residue was purified by flash chromatography with a gradient of CH₂Cl₂/MeOH 5 → 10%.

Step 5: Mitsunobu reaction

Structure name: *tert*-butyl 3-((2*R*)-2-amino-5,11-dioxo-10-(prop-2-yn-1-yl)-2,3,5,10,11,11a-hexahydro-1*H*-benzo[*e*]pyrrolo[1,2-*a*][1,4]diazepin-7-yl)propanoate **76**

Scheme:



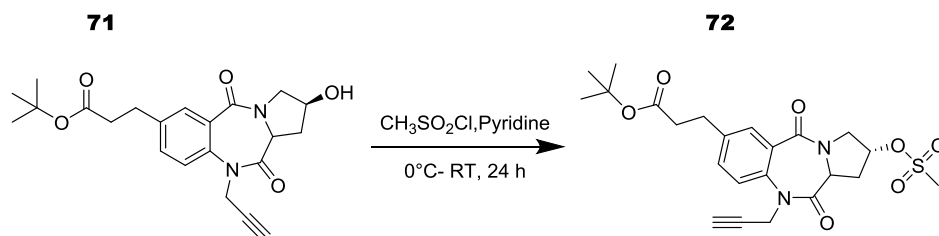
Procedure: Under argon diisopropyl azodicarboxylate (1.1 eq) was added to a degassed solution of **71** (1 eq), di-*tert*-butyl iminodicyclohexanecarboxylate **75** (1.2 eq) and triphenylphosphine (1.1 eq) in tetrahydrofuran and the resulting mixture was stirred at room temperature for 18 hours. Water was added to the reaction mixture, followed by extraction with ethyl acetate. The organic layer was then washed successively with water and brine. The organic layer thus obtained was dried over sodium sulfate and concentrated under reduced pressure. The residue thus obtained was subjected to flash chromatography on a silica gel column.

Observation: The Mitsunobu reaction was performed to convert the hydroxyl group OH of compound **71** to the amine **76** with inversion of configuration. The nucleophile employed should be acidic, since one of the reagents (DIAD or DEAD) must be protonated during the reaction to prevent from side reactions. The triphenylphosphine combines with DEAD should generate a phosphonium intermediate that binds to the alcohol oxygen, activating it as a leaving group. The formation of the amine failed in this case. The principal mass observed on the LC-MS was 279 corresponding to triphenylphosphine oxide.

Optimisation: The preparation of the amine definitively failed by the use of triphenylphosphine and DIAD/DEAD. Therefore, an alternative way to convert the alcohol to the amine using a different synthetic route was investigated. The increase of DIAD/DEAD amount brought no significant changes, also higher temperature, as well as the increase of the amine did not lead to improvement. This led to a new synthetic pathway to the final molecule.

New synthetic route:**Step 5:** Mesylate formation

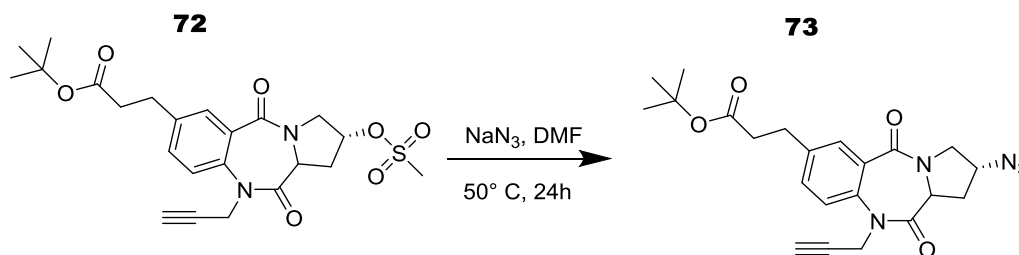
Structure name: *tert*-butyl 3-((2*R*)-2-((methylsulfonyl)oxy)-5,11-dioxo-10-(prop-2-yn-1-yl)-2,3,5,10,11,11a-hexahydro-1*H*-benzo[*e*]pyrrolo[1,2-*a*][1,4]diazepin-7-yl)propanoate **72**

Scheme:

Procedure: Compound **71** (200 mg, 1 eq, 0.55 mmol) was dissolved in dichloromethane (3 mL). After addition of pyridine (3 mL) the solution was cooled to -10°C . Methanesulfonyl chloride (64 μL , 1.5 eq, 0.82 mmol) was slowly added with stirring. The mixture was then stirred for 1 h at -10°C and for 10 h at room temperature. The solution was washed with sat. NaHCO_3 solution and the aqueous layer were extracted with dichloromethane. The combined organic layers were washed with 10% citric acid and water and dried over MgSO_4 . The solvent was removed under reduced pressure. (220 mg, 84 % yield) MS (ESI): calc. 476.54, found ($[\text{M}+\text{H}]^+$) 477.38. Purity (HPLC): 96%. $^1\text{H-NMR}$: (500 MHz, CDCl_3 -*d*1) δ ppm: 9.10 (*br, s,* 1H), 8.13 (*d,* $^4J= 1.99$ Hz, 1H), 7.62 (*dd,* $^3J= 8.34$ Hz and $^4J= 1.99$ Hz, 1H), 7.55 (*d,* $^3J= 15.89$ Hz, 1H), 7.10 (*m,* 1H), 6.39 (*d,* $^3J= 15.90$ Hz, 1H), 4.37 (*m,* 1H), 4.20 (*dd,* $^3J= 9.14$ Hz and $^3J= 1.99$ Hz, 1H), 3.82 (*dd,* $^2J= 12.98$ Hz and $^3J= 5.14$ Hz, 1H), 3.73 (*m,* 1H), 3.10 (*dd,* $^2J= 13.91$ Hz and $^3J= 1.99$ Hz, 1H), 2.38 (*m,* 1H), 1.52 (*s,* 9H).

Step 6: Azide formation

Structure name: *tert*-butyl 3-((2*R*)-2-azido-5,11-dioxo-10-(prop-2-yn-1-yl)-2,3,5,10,11,11a-hexahydro-1*H*-benzo[*e*]pyrrolo[1,2-*a*][1,4]diazepin-7-yl)propanoate **73**

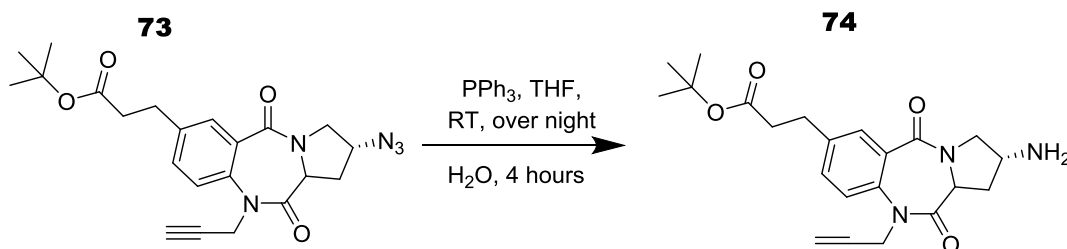
Scheme:

Procedure: The crude mesylate **72** (220 mg, 1 eq, 0.46 mmol) was dissolved in N,N-dimethylformamide (4 mL). Sodium azide (570.2 mg, 19 eq, 8.77 mmol) was added, followed by stirring at 50 °C for 72 h. Afterwards, the mixture was poured into ice-cooled water (50 mL) and the aqueous phase was extracted with dichloromethane (3 x 100 mL). After the organic layer had been dried over Na₂SO₄, the solvent was evaporated. Purification by flash chromatography (CH₂Cl₂/CH₃OH 10:1) afforded product **73** (148 mg, 76 % yield) MS (ESI): calc. 423.47, found ([M+H]⁺) 424.56. Purity (HPLC): 96%. ¹H-NMR: (500 MHz, CDCl₃-d₁) δ ppm: 9.10 (br, s, 1H), 8.13 (d, ⁴J= 1.99 Hz, 1H), 7.62 (dd, ³J= 8.34 Hz and ⁴J= 1.99 Hz, 1H), 7.55 (d, ³J= 15.89 Hz, 1H), 7.10 (m, 1H), 6.39 (d, ³J= 15.90 Hz, 1H), 4.37 (m, 1H; CH), 4.20 (dd, ³J= 9.14 Hz and ³J= 1.99 Hz, 1H), 3.82 (dd, ²J= 12,98 Hz and ³J= 5.14 Hz, 1H), 3.73 (m, 1H; CH₂), 3.10 (dd, ²J= 13.91 Hz and ³J=1.99 Hz, 1H), .38 (m, 1H; CH₂), 1.52 (s, 9H). ¹³C-NMR: (126 MHz, DMSO-d₆) δ ppm: 171.8, 171.6, 164.5, 135.7, 134.8, 134.4, 128.9, 126.3, 124.1, 115.8, 82.1, 73.2, 60.2, 52.2, 49.2, 34.7, 30.1, 28.7, 28.1.

Step 7: Staudinger azide reduction

Name: *tert*-butyl 3-((2*R*)-2-azido-5,11-dioxo-10-(prop-2-yn-1-yl)-2,3,5,10,11,11a-hexahydro-1*H*-benzo[*e*]pyrrolo[1,2-*a*][1,4]diazepin-7-yl)propanoate **74**

Scheme:



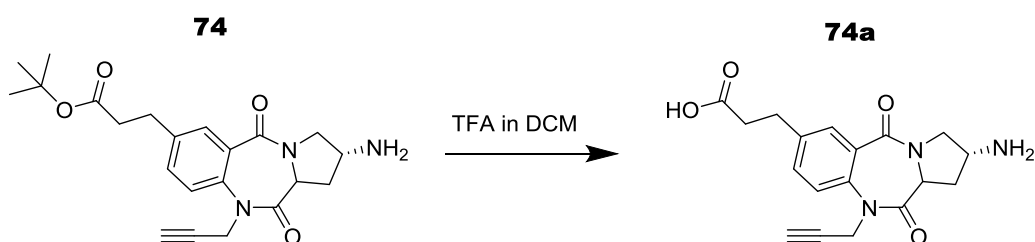
Procedure: The compound **73** (181 mg, 1 eq, 0.43 mmol) was dissolved in dry tetrahydrofuran (3 mL) and triphenylphosphine (135.3 mg, 1.2 eq, 0.52 mmol) was added. The reaction mixture was stirred at room temperature for 24 h and then water (2 mL) was added. Stirring was continued overnight and the solvents were removed in vacuo.

The residue was purified by chromatography on silica using dichloromethane/methanol (98/2) followed by dichloromethane/methanol (95/5) to afford the title compound **74** as a white off solid. (133.3 mg, 78 % yield) MS (ESI): calc. 397.47, found ($[M+H]^+$) 398.68. Purity (HPLC): 96%. **¹H-NMR**: ¹H NMR (500 MHz, CDCl₃-d₁): 8.03 (d, ⁴J= 1.99 Hz, 1H), 7.74 (dd, ⁴J= 1.99 Hz and ³J= 8.34 Hz, 1H), 7.55 (d, ³J= 15.89 Hz, 1H), 7.15 (d, ³J= 8.74 Hz, 1H), 6.41 (d, ³J= 16.29 Hz, 1H), 4.34 (dd, ³J= 3.18 Hz and ³J= 8.35 Hz, 1H), 3.88 (dd, ³J= 6.75 Hz and ²J= 11.92 Hz, 1H), 3.76 (m, 1H), 3.53 (dd, ²J= 11.92 Hz and ³J= 7.15 Hz, 1H), 2.93 (m, 1H), 2.04 (m, 1H), 1.52 (s, 9H). **¹³C-NMR**: (126 MHz, DMSO-d₆) δ ppm: 171.8, 171.6, 164.5, 135.7, 134.8, 128.9, 126.3, 124.3, 115.8, 82.1, 73.2, 72.1, 60.2, 56.7, 44.1, 34.8, 34.5, 30.2, 28.7.

Step 8: Deprotection of the protective groups

Structure name: 3-((2*R*)-2-amino-5,11-dioxo-10-(prop-2-yn-1-yl)-2,3,5,10,11,11a-hexahydro-1*H*-benzo[*e*]pyrrolo[1,2-*a*][1,4]diazepin-7-yl)propanoic acid **74a**

Scheme:

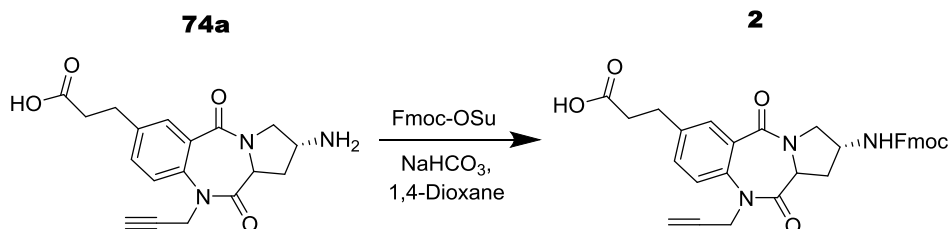


Procedure: TFA (500 μL) was added to a solution of compound **74** (150 mg, 1 eq, 0.38 mmol) in dichloromethane (5 mL), followed by stirring overnight at room temperature. The solvent was removed in vacuo and the residue was triturated with diethyl ether to afford the desired compound **74a** (110.2 mg, 85 % yield) MS (ESI): calc. 341.36, found ($[M+H]^+$) 342.46. Purity (HPLC): 98 %. **¹H-NMR**: (500 MHz, DMSO): 10.79 (br., s., 1H), 7.97 (d, ⁴J= 1.77 Hz, 1H), 7.89 (dd, ⁴J= 1.77 Hz and ³J= 8.48 Hz, 1H), 7.79 (br., s., 1H), 7.55 (d, ³J= 15.89 Hz, 1H), 7.19 (d, ³J= 8.47 Hz, 1H), 6.48 (d, ³J= 15.89 Hz, 1H), 4.29 (dd, ³J= 3.53 Hz and ³J= 8.83 Hz, 1H), 4.16 (m, 1H), 3.96 (dd, ²J= 12.36 Hz and ³J= 5.65 Hz, 1H), 3.43 (d, ²J= 12.36 Hz, 1H), 2.62 (m, 1H), 2.35 (m, 1H). **¹³C-NMR**: (126 MHz, DMSO-d₆) δ ppm: 171.8, 171.6, 164.5, 135.7, 134.8, 128.9, 126.3, 124.1, 115.8, 82.1, 73.2, 72.1, 60.2, 56.7, 44.1, 34.8, 34.5, 32.8, 30.2, 28.7.

Step 9: Fmoc protection of the free amine

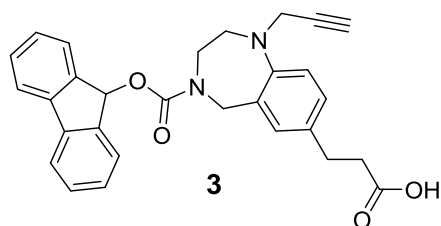
Structure name: 3-((2*R*)-2-(((9*H*-fluoren-9-yl)methoxy)carbonyl)amino)-5,11-dioxo-10-(prop-2-yn-1-yl)-2,3,5,10,11,11*a*-hexahydro-1*H*-benzo[*e*]pyrrolo[1,2-*a*][1,4]diazepin-7-yl)propanoic acid **2**

Scheme:



Procedure: Compound **74a** (120 mg, 1 eq, 0.30 mmol) was dissolved in water (2 mL) and sodium bicarbonate (50.4 mg, 2 eq, 0.60 mmol) was added with stirring. The resulting solution was cooled to 5°C and Fmoc-OSu (152 mg, 1.5 eq, 0.45 mmol) was added slowly as a solution in *para*-dioxane (also cooled, 4 mL). The resulting mixture was stirred at 0° for 1 h and allowed to warm to room temperature overnight. Water was then added and the aqueous layer was extracted 2 times with EtOAc. The organic layer was extracted back twice with saturated sodium bicarbonate solution. The combined aqueous layers were acidified to a pH of 1 with 10 % aq. HCl, and extracted 3 times with EtOAc. The combined organic layers were dried (sodium sulfate) and concentrated *in vacuo*. The residue was purified by column chromatography using DCM/MeOH (90:10). (135 mg, 79.8 % yield) MS (ESI): calc. 563.60, found ([M+H]⁺) 564.58. Purity (HPLC): 98%. ¹H-NMR: (500 MHz, CDCl₃-*d*1): 10.79 (s, 1H), 7.88 (d, ³J= 6.8 Hz, 2H), 7.53 (d, ³J= 7.4 Hz, 2H), 7.38 (dd, ³J= 7.1 Hz and ³J= 5.1 Hz, 2H), 7.27 (dd, ³J= 7.9 Hz and ⁴J= 2.1 Hz, 2H), 7.16 (d, ⁴J= 2.1 Hz, 1H), 7.03 (dd, ³J= 8.4 Hz and ⁴J= 2.1 Hz, 1H), 6.59 (d, ³J= 8.4 Hz, 1H), 4.73 (dd, ²J= 15.4 Hz and ³J= 4.3 Hz, 2H), 4.48 (d, ³J= 5.1 Hz, 1H), 4.07 (d, ²J= 15.8 Hz, 2H), 3.71 (m, 1H), 3.69 (d, ³J= 7.5 Hz, 1H), 3.44 (d, ³J= 7.5 Hz, 1H), 2.27 (d, ³J= 8.6 Hz, 1H), 2.02 (d, ³J= 8.6 Hz, 1H), 2.81 (dd, ²J= 12.6 Hz and ³J= 4.5 Hz, 2H), 2.63 (s, 1H), 2.49 (dd, ²J= 12.6 Hz and ³J= 8.1 Hz, 2H). ¹³C-NMR: (126 MHz, DMSO-*d*6) δ ppm: 174.4, 171.6, 164.5, 155.6, 143.6, 142.6, 135.7, 134.8, 128.9, 126.7, 126.3, 126.2, 125.2, 124.1, 120.5, 115.8, 73.2, 71.1, 67.3, 60.2, 53.9, 47.1, 46.4, 34.4, 34.2, 29.8.

VI-2.3. Synthesis of the tetrahydrobenzodiazepine 3



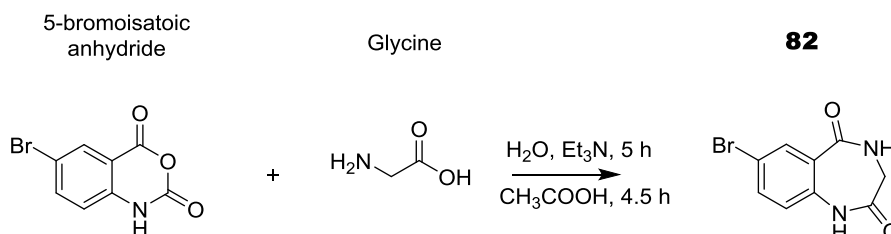
The objective of this part was the synthesis of scaffold **3** that is functionalized in way that it allows coupling to 5'-aminolinker modified DNA. It also possesses a carboxylic acid moiety, an Fmoc-protected amino group and a terminal alkyne for subsequent substitutions in combinatorial fashion. The chemical structure of this 1,4-benzodiazepine consists of a core benzene ring fused to a seven-member 1,4-diazepine ring. It shows a semi-planar platform and distinguishes itself from the Scaffold 1 and 2 by the absence of the oxo groups which are reduced in the second step of the synthesis. The N1-position will be implanted with the desired terminal alkyne.

Synthesis of tetrahydrobenzodiazepine 3. Reagents and conditions: a) bromine in H₂O at 50 °C for 1 h; b) glycine in H₂O, Et₃N, rt, 4 h; c) BH₃ in dry THF, reflux, 18 h; d) "diboc" in dry MeOH, rt, 18 h; e) Pd(OAc)₂, P(O-tol)₃, *tert*-butyl acrylate, Et₃N, dry CH₃CN, 100 °C, 18 h; f) Pd/C, dry MeOH, rt, 18 h; g) Cs₂CO₃, propargyl bromide, dry DMF, 60 °C, 18 h; h) TFA in dry CH₂Cl₂, rt, 19 h; i) Fmoc-Osu, NaHCO₃, H₂O, dry 1,4-dioxane, rt, 18 h.

Step 1: Benzodiazepine formation

Structure name: 7-bromo-3,4-dihydro-1*H*-benzo[*e*][1,4]diazepine-2,5-dione **82**

Scheme:

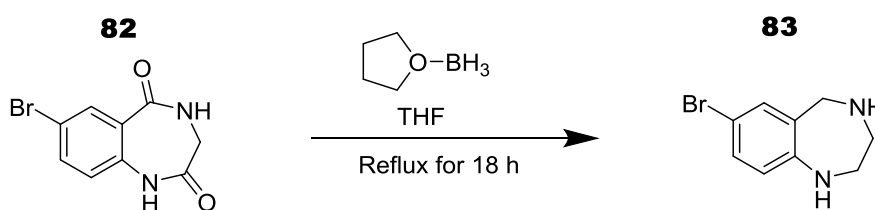


Procedure: To a solution of 5-bromoisatoic anhydride (10 g, 1 eq, 41.3 mmol) in water (40 mL) was added glycine (4.4 g, 1.4 eq, 58 mmol) at room temperature. The reaction mixture was stirred at room temperature for 4 hours to give a cloudy solution. The reaction mixture was concentrated *in vacuo*. Glacial acetic acid was added and the reaction mixture was refluxed for 4.5 hours. The reaction mixture was cooled down slowly to room temperature. A precipitate formed. The reaction mixture was diluted with diethyl ether, then filtered through a sintered funnel to yield the title product **82** (7.6 g, 72.5 % yield). ¹H-NMR (500 MHz, DMSO-*d*₆) δ ppm: 8.02 (t, 1H), 8.04 (s, 1H), 7.84 (d, ⁴J= 2.2 Hz, 1H), 7.70 (dd, ³J= 8.6 Hz and ⁴J= 2.2 Hz, 1H), 7.07 (d, ³J= 8.7 Hz, 1H), 3.63 (d, ²J= 13.3 Hz, 2H). ¹³C-NMR (126 MHz, DMSO-*d*₆) δ ppm: 168.9, 167.5, 137.4, 135.7, 130.2, 128.2, 124.1, 116.5, 45.1. MS (ESI): calc. 253.97, found 254.4 ([M+H]⁺). Purity (HPLC): 98 %.

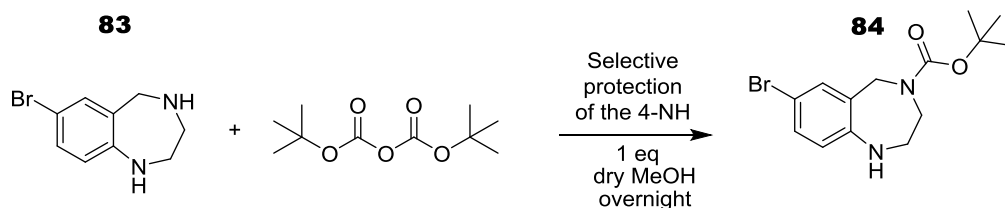
Step 2: Reduction of the oxo groups

Structure name: 7-bromo-2,3,4,5-tetrahydro-1*H*-benzo[*e*][1,4]diazepine **83**

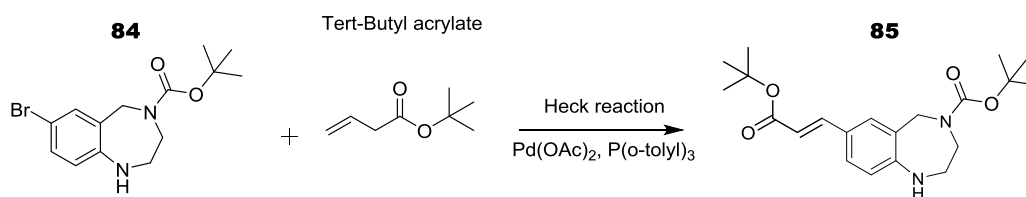
Scheme:



Procedure: A solution of borane in THF (1M, 15 eq, 28 mL) was slowly added to a stirred solution of **82** (5 g, 1 eq, 19.6 mmol) in dry THF. The reaction mixture was refluxed for 18h. After cooling to 0°C, MeOH (9 mL) was added carefully, the solvent was evaporated and the residue was dissolved in MeOH (20 mL). To this solution, 7N aq. HCl (5 mL) was added and the mixture was heated to dryness. The resulting solid was suspended in NaHCO₃ (saturated aqueous solution, 100 mL) and the suspension was brought to pH 9 with 5N aq. NaOH. The product was extracted with 3 x 100 mL of CH₂Cl₂ and the combined organic layers were washed with brine, dried over Na₂SO₄ and concentrated *in vacuo*. Purification by column chromatography (CH₂Cl₂/ MeOH 10:1) afforded **83** as a light yellow solid (2.55 g, 57.6 % yield). ¹H-NMR (500 MHz, DMSO-*d*₆) δ ppm: 7.21 (dd, ³J= 8.4 Hz and ⁴J= 2.4 Hz, 1H), 7.18 (d, ⁴J= 2.4 Hz, 1H), 6.46 (d, ³J= 8.4 Hz, 1H), 5.61 (br. s, 1H), 3.65 (s, 2H), 2.89 - 3.00 (m, 2H), 2.75 - 2.82 (m, 1H), 2.12 (m, 1H). ¹³C-NMR (126 MHz, DMSO-*d*₆) δ ppm: 151.2, 135.5, 132.2, 130.1, 121.2, 110.7, 54.5, 52.3, 50.6. MS (ESI): calc. 226.01, found 227.95 ([M+H]⁺). Purity (HPLC): 98 %.

Step 3: Boc protection of the N-4 secondary amine**Structure name:** *tert*-butyl-7-bromo-2,3-dihydro-1*H*-benzo[*e*][1,4]diazepine-4(5*H*)-carboxylate **84****Scheme:**

Procedure: To a stirred solution of **83** (2.73 g, 1 eq, 12.1 mmol) in dry methanol (10 mL) was added di-*tert*-butyl dicarboxylate (2.64 g, 1 eq, 12.1 mmol) at 0°C under argon. The mixture was stirred at room temperature for 18 hours. The solvent was removed and the residue was purified by column chromatography (petrolether/EtOAc 80:20), affording the pure compound **84** as a beige solid (2.29 g, 58 % yield). ¹H-NMR (500 MHz, DMSO-*d*₆) δ ppm: 7.42 (dd, ³*J*= 8.0 Hz and ⁴*J*= 2.6 Hz, 1H), 7.39 (d, ⁴*J*= 2.6 Hz, 1H), 6.42 (d, ³*J*= 8.0 Hz, 1H), 4.25 (d, ²*J*= 16.0 Hz, 2H), 4.11 (s, br., 1H), 3.33 (dd, ²*J*= 13.6 Hz and ³*J*= 4.8 Hz, 2H), 3.19 (dd, ²*J*= 13.6 Hz and ³*J*= 8.3 Hz, 2H), 1.39 (s, 9H). ¹³C-NMR (126 MHz, DMSO-*d*₆) δ ppm: 154.3, 147.3, 133.5, 130.7, 123.4, 113.8, 111.3, 79.8, 56.3, 55.5, 43.6, 28.4. MS (ESI): calc. 326.06, found 327.22 ([M+H]⁺). Purity (HPLC): 96 %.

Step 5: Heck reaction**Structure name:** *tert*-butyl (*E*)-7-(3-(*tert*-butoxy)-3-oxoprop-1-en-1-yl)-1,2,3,5-tetrahydro-4*H*-benzo[*e*][1,4]diazepine-4-carboxylate **85****Scheme:**

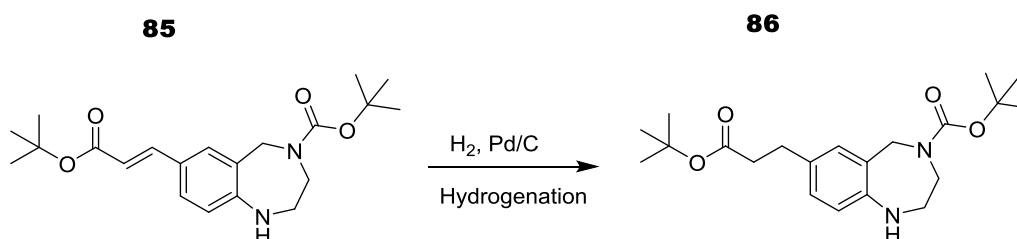
Procedure: Palladium (II) acetate (200 mg, 0.2 eq, 0.92 mmol) was added to a degassed solution of **84** (1.5 g, 1 eq, 4.6 mmol), tri-*o*-tolylphosphine (700 mg, 0.5 eq, 2.3 mmol), triethylamine (2.6 mL, 3 eq, 13.8 mmol) and *tert*-butyl acrylate (4.1 mL, 6 eq, 27.6 mmol) in dry acetonitrile (8 mL). The mixture was refluxed for 18 h.

The solvent was evaporated under reduced pressure and the residue was purified by column chromatography (CH₂Cl₂/MeOH 90:10) giving the compound **85** (870 mg, 50.5 % yield). **¹H-NMR**: (500 MHz, CDCl₃) δ ppm: 7.95 (dd, ³J= 8.4 Hz and ⁴J= 2.1 Hz, 1H), 7.44 (d, ²J= 15.8 Hz, 1H), 7.38 (d, ⁴J= 2.1 Hz, 1H), 7.06 (d, ³J= 8.4 Hz, 1H), 6.29 (d, ²J= 16.0 Hz, 1H), 4.25 (d, ²J= 15.2 Hz, 2H), 4.05 (s, br., 1H), 3.34 (dd, ²J= 13.6 Hz and ³J= 4.8 Hz, 2H), 3.20 (dd, ²J= 13.6 Hz and ³J= 8.3 Hz, 2H), 1.48 (s, 9H), 1.39 (s, 9H). **¹³C-NMR**: (126 MHz, CDCl₃) δ ppm: 166.5, 154.3, 147.5, 145.1, 127.9, 126.3, 123.4, 121.1, 116.2, 107.8, 81.2, 79.8, 57.4, 55.5, 43.6, 28.8, 28.4. MS (ESI): calc. 374.22, found 375.47 ([M+H]⁺). Purity (HPLC): 98 %.

Step 6: Hydrogenation of the alkene

Structure name: *tert*-butyl-7-(3-(*tert*-butoxy)-3-oxopropyl)-1,2,3,5-tetrahydro-4*H*-benzo[*e*][1,4]diazepine-4-carboxylate **86**

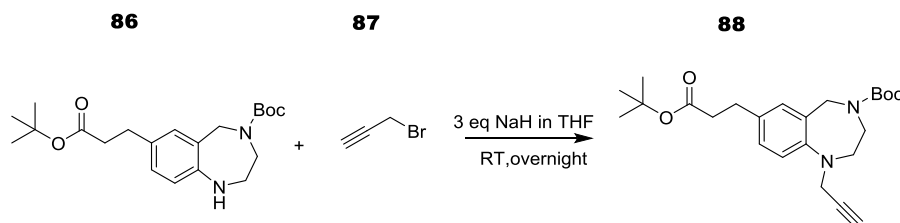
Scheme:



Procedure: To a stirred solution of **85** (700 mg, 1 eq, 1.87 mmol) in dry methanol (5 mL) at room temperature under hydrogen was added palladium on carbon (50 mg, 0.25 eq, 0.5 mmol). The reaction mixture was stirred under hydrogen for 18 hours, filtered and the residue was additionally washed with methanol. The filtrate was then concentrated and the product was purified by column chromatography (CH₂Cl₂/MeOH 90:10), affording pure compound **86** (476 mg, 67.6 % yield). **¹H-NMR**: (500 MHz, DMSO-*d*₆) δ ppm: 7.12 (d, ⁴J= 2.1 Hz, 1H), 6.99 (dd, ³J= 8.4 Hz and ⁴J= 2.1 Hz, 1H), 6.56 (d, ³J= 8.4 Hz, 1H), 4.25 (d, ²J= 15.2 Hz, 2H), 4.05 (s, br., 1H), 3.34 (dd, ²J= 13.6 Hz and ³J= 4.8 Hz, 2H), 3.20 (dd, ²J= 13.6 Hz and ³J= 8.3 Hz, 2H), 2.92 (dd, ²J= 12.6 Hz and ³J= 4.5 Hz, 2H), 2.55 (dd, ²J= 12.6 Hz and ³J= 8.1 Hz, 2H), 1.48 (s, 9H), 1.39 (s, 9H). **¹³C-NMR**: (126 MHz, DMSO-*d*₆) δ ppm: 171.7, 154.3, 145.5, 129.3, 127.6, 126.8, 121.1, 113.2, 82.1, 79.8, 57.3, 55.5, 43.6, 34.7, 28.7, 28.4. MS (ESI): calc. 376.24, found 377.49 ([M+H]⁺). Purity (HPLC): 97 %.

Step 7: Alkylation of the N-1 secondary amine

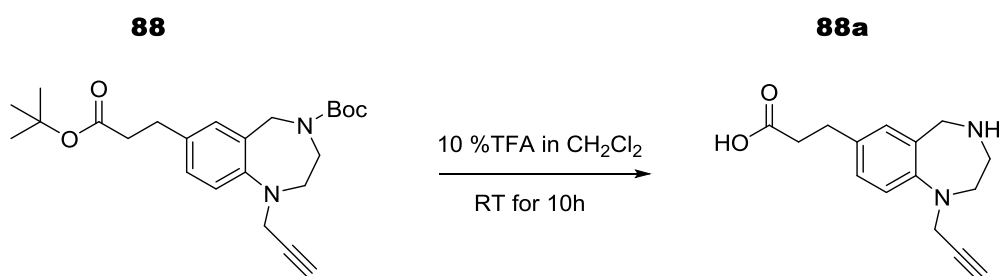
Structure name: *tert*-butyl 7-(3-(*tert*-butoxy)-3-oxopropyl)-1-(prop-2-yn-1-yl)-1,2,3,5-tetrahydro-4*H*-benzo[*e*][1,4]diazepine-4-carboxylate **88**

Scheme:

Procedure: In a 2-neck round-bottomed flask fitted with argon gas and septum was placed compound **86** (535 mg, 1 eq, 1.42 mmol) dissolved in dry DMF (3 mL). To the suspension was added (925 mg, 2 eq, 2.84 mmol) of Cs₂CO₃ at room temperature. Stirring was continued for 20 minutes and propargyl bromide **87** (245 μL, 2 eq, 2.84 mmol) was slowly added. Stirring was then continued at 60 °C overnight. Then, the solvent was evaporated, and the residue was purified by column chromatography with a gradient of CH₂Cl₂/MeOH 95:5 to 90:10 (the solvent contained 0.5 % of TEA) yielding pure compound **88** (278.2 mg, 47.3 % yield). ¹H-NMR: (400 MHz, DMSO-*d*₆) δ ppm: 7.16 (d, ⁴*J* = 2.1 Hz, 1H), 7.03 (dd, ³*J* = 8.4 Hz and ⁴*J* = 2.1 Hz, 1H), 6.59 (d, ³*J* = 8.4 Hz, 1H), 4.25 (d, ²*J* = 15.2 Hz, 2H), 4.07 (d, ²*J* = 15.8 Hz, 2H), 3.62 (dd, ²*J* = 13.6 Hz and ³*J* = 8.3 Hz, 2H), 3.18 (dd, ²*J* = 13.6 Hz and ³*J* = 4.8 Hz, 2H), 2.92 (dd, ²*J* = 12.6 Hz and ³*J* = 4.5 Hz, 2H), 2.63 (s, 1H), 2.55 (dd, ²*J* = 12.6 Hz and ³*J* = 8.1 Hz, 2H), 1.48 (s, 9H), 1.39 (s, 9H). ¹³C-NMR: (101 MHz, DMSO-*d*₆) δ ppm: 171.7, 154.3, 145.5, 129.3, 127.6, 126.8, 121.1, 113.2, 82.1, 79.8, 78.1, 73.2, 57.3, 55.5, 46.7, 43.6, 34.7, 28.7, 28.4. MS (ESI): calc. 414.54, found 415.34 ([M+H]⁺). Purity (HPLC): 95 %.

Step 8: Deprotection of the Boc groups

Structure name: *tert*-butyl-7-(3-(*tert*-butoxy)-3-oxopropyl)-1-(prop-2-yn-1-yl)-2,3-dihydro-1*H*-benzo[*e*][1,4]diazepine-4(5*H*)-carboxylate **88a**

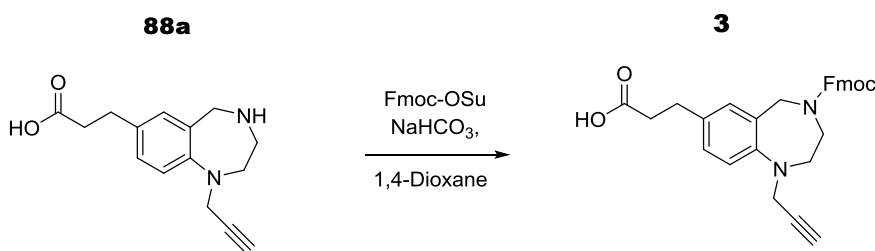
Scheme:

Procedure: TFA (4.4 mL) was added to a solution of compound **88** (440 mg, 1.1 mmol) in dry dichloromethane (12 mL). The reaction was stirred overnight at room temperature for 19 hours. The solvent was removed *in vacuo* and the residue was triturated with diethyl ether to afford compound **88a** (276 mg, 97.2 % yield). **¹H-NMR** (500 MHz, DMSO-*d*₆) δ ppm: 10.79 (s, 1H), 7.16 (d, ⁴*J*= 2.1 Hz, 1H), 7.03 (dd, ³*J*= 8.4 Hz and ⁴*J*= 2.1 Hz, 1H), 6.59 (d, ³*J*= 8.4 Hz, 1H), 4.07 (d, ²*J*= 15.8 Hz, 2H), 3.89 (d, ²*J*= 15.2 Hz, 2H), 3.42 (dd, ²*J*= 13.6 Hz and ³*J*= 8.3 Hz, 2H), 2.75 (dd, ²*J*= 13.6 Hz and ³*J*= 4.8 Hz, 2H), 2.81 (dd, ²*J*= 12.6 Hz and ³*J*= 4.5 Hz, 2H), 2.63 (s, 1H), 2.49 (dd, ²*J*= 12.6 Hz and ³*J*= 8.1 Hz, 2H), 1.99 (m, 1H). **¹³C-NMR** (126 MHz, DMSO-*d*₆) δ ppm: 174.4, 143.4, 129.4, 128.7, 126.9, 121.9, 120.2, 78.1, 73.2, 59.2, 53.3, 50.1, 46.7, 34.2, 30.5. MS (ESI): calc. 258.14, found 259.32 ([M+H]⁺). Purity (HPLC): 98 %.

Step 9: Fmoc protection of the secondary amine

Structure name: 3-(4-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-1-(prop-2-yn-1-yl)-2,3,4,5-tetrahydro-1*H*-benzo[*e*][1,4]diazepin-7-yl)propanoic acid **3**.

Scheme:



Procedure: Compound **88a** (350 mg, 1eq, 1.35 mmol) was dissolved in water (2 mL) and sodium bicarbonate (227 mg, 2 eq, 2.7 mmol) was added with stirring. The resulting solution was cooled to 5°C and Fmoc-OSu (685 mg, 1.5 eq, 2.03 mmol) was added slowly as a solution in *para*-dioxane (also cooled, 4 mL). The resulting mixture was stirred at 0° for 1 h and allowed to warm to room temperature overnight. Water was then added and the aqueous layer was extracted 2 times with EtOAc. The organic layer was extracted back twice with saturated sodium bicarbonate solution. The combined aqueous layers were acidified to a pH of 1 with 10 % aq. HCl, and extracted 3 times with EtOAc.

The combined organic layers were dried (sodium sulfate) and concentrated *in vacuo*. The residue was purified by column chromatography using DCM/MeOH (90:10). (457.4 mg, 70.5 % yield). **¹H-NMR** (500 MHz, DMSO-*d*₆) δ ppm: (500 MHz, CDCl₃-*d*₁): 10.79 (s, 1H), 7.88 (d, ³*J*= 6.8 Hz, 2H), 7.53 (d, ³*J*= 7.4 Hz, 2H), 7.38 (dd, ³*J*= 7.1 Hz and ³*J*= 5.1 Hz, 2H), 7.27 (dd, ³*J*= 7.9 Hz and ⁴*J*= 2.1 Hz, 2H), 7.16 (d, ⁴*J*= 2.1 Hz, 1H), 7.03 (dd, ³*J*= 8.4 Hz and ⁴*J*= 2.1 Hz, 1H), 6.59 (d, ³*J*= 8.4 Hz, 1H), 4.73 (dd, ²*J*= 15.4 Hz and ³*J*= 4.3 Hz, 2H), 4.48 (d, ³*J*= 5.1 Hz, 1H), 4.27 (d, ²*J*= 15.2 Hz, 2H), 4.07 (d, ²*J*= 15.8 Hz, 2H), 3.66 (dd, ²*J*= 13.6 Hz and ³*J*= 8.3 Hz, 2H), 3.22 (dd, ²*J*= 13.6 Hz and ³*J*= 4.8 Hz, 2H), 2.81 (dd, ²*J*= 12.6 Hz and ³*J*= 4.5 Hz, 2H), 2.63 (s, 1H), 2.49 (dd, ²*J*= 12.6 Hz and ³*J*= 8.1 Hz, 2H). **¹³C-NMR** (126 MHz, DMSO-*d*₆) δ ppm: 174.4, 162.8, 143.6, 143.4, 142.6, 129.4, 128.7, 126.9, 126.7, 126.2, 125.2, 121.9, 120.5, 120.2, 78.1, 73.2, 67.6, 57.6, 56.4, 52.4, 47.0, 46.7, 34.2, 30.5. MS (ESI): calc. 480.55, found 481.57 ([M+H]⁺), and 503.01 ([M+Na]⁺). Purity (HPLC): 98 %.

VI-3. Coupling of carboxylic acid building blocks to DNA scaffold conjugates

VI-3.1. Synthesis of DNA-PEG-linker conjugate **94a**

The MMT-protective group of the 5'-aminolinker-modified DNA bound to 1000 Å controlled pore glass (CPG) solid support **94** (500 nmol, ca. 20 mg) was removed by addition of 3 % trichloroacetic acid in dry DCM (3 x 200 µL) for 3 x 1 min. A yellow color indicated successful removal of the protective group. The CPG containing the deprotected DNA was then washed three times with each 200 µL of 1 % TEA in MeCN, then DMF, MeOH, MeCN and DCM. The CPG, the MMT-NH-PEG(8)-COOH linker **155** or Fmoc-NH-PEG(4)-COOH linker **105**, and HATU as a coupling reagent were dried *in vacuo* for 15 min. Stock solutions of all reactants in dry DMF were prepared immediately before reaction was started. To 150 µL of a solution of the MMT-NH-PEG(8)-COOH linker **155** (35.7 mg, 50 µM, 100 eq.) or the Fmoc-NH-PEG(4)-COOH linker **105** (26.9 mg, 50 µmol, 100 eq.) in dry DMF were added 150 µL of HATU (19 mg, 50 µM, 100 eq.) dissolved in dry DMF and DIPEA (21.3 µL, 250 eq.). This reaction mixture was shaken for 5 min and added to the solid support-bound DNA suspended in dry DMF (150 µL). The amide coupling reaction was shaken at room temperature for 2 hours. Then, the CPG containing the DNA-PEG linker conjugate was filtered off using a filter column and washed subsequently with each 3 x 200 µL of DMF, MeOH, MeCN and DCM. Unreacted amines were capped with acetic acid anhydride, and the CPG was again washed subsequently with each 3 x 200 µL of DMF, MeOH, MeCN and DCM, and dried *in vacuo* for 15 min. For analysis, an aliquot ca. 10 nmol of the DNA-PEG conjugate **94a** was deprotected and cleaved from the CPG by treatment with 500 µL of AMA (AMA= aqueous ammonia (30 %)/ aqueous methylamine (40 %)= 1:1, vol/vol) for 4 hours at room temperature. Then, 20 µL of 1 M Tris buffer (pH= 7.5) were added, the mixture was dried in a SpeedVac, re-dissolved in 100 µL of distilled water, and the product was analyzed by RP-HPLC (Gemini 5u C18 110A column; 100*10.0 mm) with a gradient of aqueous triethylammonium acetate buffer (100 mM, pH= 8) and methanol (20% – 70% of methanol over 19 min).

VI-3.2. Synthesis of DNA-scaffold conjugate **95**

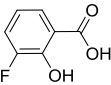
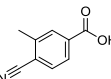
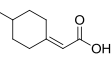
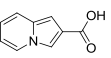
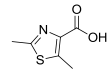
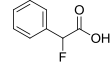
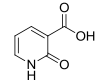
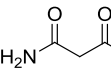
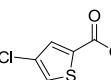
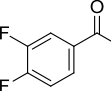
Prior coupling of the scaffold to the PEG-linker, the protective group of the linker was removed: The MMT-protective group of the DNA-PEG(8) linker conjugate bound to 1000 Å controlled pore glass (CPG) solid support (1 μmol, ca. 40 mg of the solid support) was removed by addition of 3 % trichloroacetic acid in dry DCM (3 x 200 μL) for 3 x 1 min. A yellow color indicated successful removal of the protective group. The Fmoc-group of the DNA-PEG(4) linker conjugate bound to 1000 Å controlled pore glass (CPG) solid support (1 μmol, ca. 40 mg of the solid support) was removed by addition of 20 % piperidine in dry DMF (1 mL). The reaction mixture was shaken for 5 min at room temperature. The CPG containing the deprotected DNA-PEG linker conjugate was washed three times with each 200 μL of 1 % TEA in MeCN, then DMF, MeOH, MeCN and DCM. The solid support-bound DNA-PEG linker conjugate, compounds **3** and HATU as a coupling reagent were dried *in vacuo* for 5 min. To a 250 μL of a solution of the compound **3** (100 μmol, 100 eq.) in dry DMF were added HATU (38 mg, 100 μmol, 100 eq.) dissolved in 250 μL of dry DMF, and DIPEA (43 μL, 250 μmol, 250 eq.). Compound **3** were activated for 5 min at room temperature in a shaker. Then, the activated carboxylic acids were added to the solid support-bound DNA-PEG linker conjugate **105a** suspended in dry DMF (250 μL). The amide coupling reactions were shaken at room temperature for 4 hours. The CPGs containing the amide coupling products **95** were filtered off using a filter column and washed subsequently with each 3 x 200 μL of DMF, MeOH, MeCN and DCM, they were then capped with acetic acid anhydride (a 1:1 mixture of THF/methylimidazole, 9:1, vol/vol, and THF/pyridine/acetic acid anhydride, 8:1:1, vol/vol), washed again with each 3 x 200 μL of DMF, MeOH, MeCN and DCM, and dried *in vacuo* for 15 min. For analysis, an aliquot of 20 nmol of the each DNA-small molecule conjugate **95** was deprotected and cleaved from the CPG by treatment with 500 μL of AMA (AMA= aqueous ammonia (30 %)/ aqueous methylamine (40 %)= 1:1, vol/vol) for 4 hours at room temperature. Then, 20 μL of 1 M Tris buffer (pH= 7.5) were added, the mixture was dried in a SpeedVac, re-dissolved in 100 μL of distilled water, and the product was analyzed by RP-HPLC (Gemini 5u C18 110A column; 100*10.0 mm) with a gradient of aqueous triethylammonium acetate buffer (100 mM, pH= 8) and methanol (20% – 70% of methanol over 19 min).

VI-3.3. Coupling reactions of 114 carboxylic acids to DNA-scaffold conjugate 96

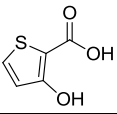
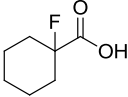
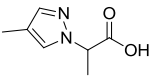
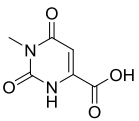
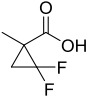
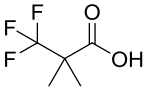
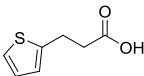

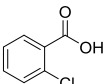

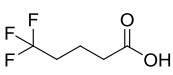
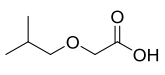
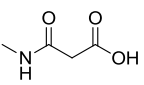
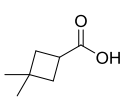
Part of this work was done by **Mateja Klika Skopic**, **Kathrin Jung** and **Sven Brandherm**. For library synthesis, sets of 20 carboxylic acids (see Table 15, building blocks A-CV) were coupled to the DNA-scaffold conjugate **95** in parallel (a total of 60 reactions were performed in parallel). For Fmoc-deprotection of DNA-small molecule conjugate **95** 400 nmol (ca. 16 mg) of each DNA-small molecule conjugate bound to 1000 Å controlled pore glass (CPG) solid support was treated with 20 % piperidine in dry DMF (400 µL) for 5 min at room temperature. The solid support containing the Fmoc-deprotected conjugate **95** was filtered off and washed with each 3 x 200 µL of DMF, MeOH, MeCN and DCM. Then, the batch of solid support with the DNA conjugates **95**, a set of 20 carboxylic acids (each 2 µmol per coupling reaction, a total of 6 µmol), and HATU as a coupling reagent were dried *in vacuo* for 15 min. Then, each of the batch of solid support containing one of the DNA conjugate **95** (each 16 mg) was split in 4 Eppendorf tubes (each 4 mg) on a balance, giving a total of 12 sub-batches (3 x 4). Each of these 12 sub-batches was suspended in 100 µL of dry DMF. These 12 suspensions were distributed to a total of 60 reaction vessels of a 96well plate, so that each reaction vessel contained ca. 20 nmol of a DNA-conjugate **95** suspended in 20 µL of dry DMF. Then, each of the 20 carboxylic acids (6 µmol;) was dissolved in 60 µL of dry DMF. To these solutions were added each HATU (6 µmol), dissolved in 60 µL of DMF, (taken from a stock solution: 120 µmol, 76 mg in 2 mL of dry DMF), and 2.6 µL (250 eq.) of DIPEA giving a total volume of ca. 120 µL. The 20 carboxylic acids were activated for 5 min with shaking. Then, from each solution of activated carboxylic acids 40 µL were added to each of the three DNA conjugate **95**, so that the coupling reaction was performed with 100 eq. of the carboxylic acid. The reaction mixtures were shaken at room temperature for 4 hours. The solid phase containing the amide coupling products were filtered off using a filter plate and washed subsequently with each 3 x 200 µL of DMF, MeOH, MeCN and DCM and dried *in vacuo* for 15 min. Then, the DNA-conjugates were deprotected and cleaved from the CPG by treatment with 500 µL of AMA (AMA= aqueous ammonia (30 %)/ aqueous methylamine (40 %) = 1:1, vol/vol) for 5 hours at room temperature. Samples were filtered directly from a filter plate into a 96-deep well plate. Then, 20 µL of 1 M Tris buffer (pH= 7.5) was added, the mixtures were dried in a SpeedVac, re-dissolved in 100 µL of distilled water, and all coupling products were purified by RP-HPLC with a gradient of aqueous triethylammonium acetate buffer (100 mM, pH= 8) and methanol (20% – 70% of methanol over 19 min).

Fractions containing the product were collected, evaporated in a SpeedVac, co-evaporated with 3 x 200 μ L of ethanol/distilled water (1:1) and analyzed by MALDI-TOF-MS analysis and by analytical HPLC to assert purity and identity. This procedure was repeated until all 114 carboxylic acids were coupled to DNA-conjugate **95**.


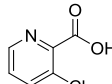
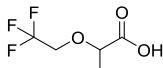
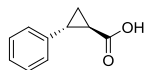

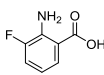
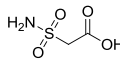
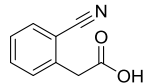
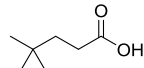
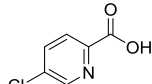
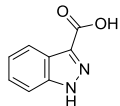
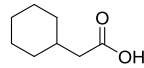
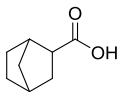
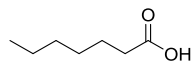
Table 14: Yield, conversion rates and MALDI MS data of compounds **96A-CZ**.

No.	structure	DNA conjugate 96		
		yield ² [nmol]	con. (%) ³	mass calc. mass found ⁴
A		0.10	50	7971.3 7974.0
B		0.91	65	7976.3 7977.3
C		0.41	60	7969.4 7971.0
D		0.10	80	7976.3 7977.5
E		4.66	95	7972.3 7976.6
F		0.10	30	7969.3 7967.8
G		1.95	100	7954.3 7954.6
H		1.33	95	5157.7 5154.6
I		3.16	95	7977.7 7979.8
J		2.85	90	7973.2 7974.5

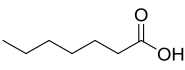
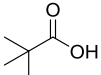
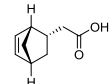
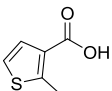
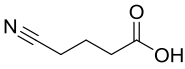
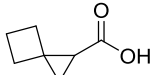
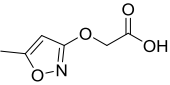
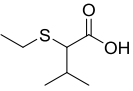
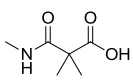
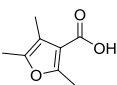
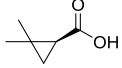
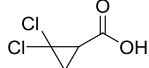
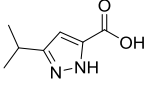
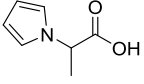
EXPERIMENTAL PART

No.	Structure	DNA conjugate 96		
		Yield ² [nmol]	Con. (%) ³	Mass calc. Mass found ⁴
K		0.58	30	7959.2 7957.1
L		0.10	60	7961.3 7960.4
M		0.19	95	7969.3 7974.6
N		1.19	95	7985.3 7988.1
O		0.74	15	7951.2 7949.8
P		5.73	90	7971.2 7969.2
Q		4.88	85	7971.3 7974.4
R		2.29	85	7976.3 7977.6
S		1.20	80	7971.3 7974.6
T		0.75	75	7945.3 7943.7
U		3.39	90	7971.2 7971.5
V		0.82	90	7947.3 7949.5
W		1.68	80	5171.7 5179.0
X		1.63	90	7943.3 7945.1

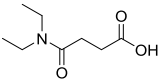
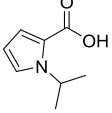
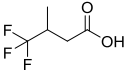
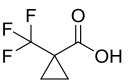
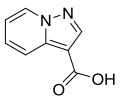
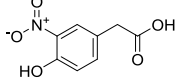
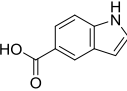
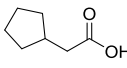
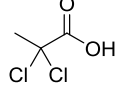
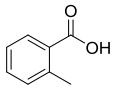
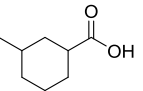
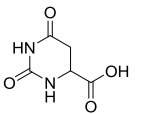
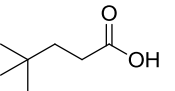
EXPERIMENTAL PART

No.	Structure	DNA conjugate 96		
		Yield ² [nmol]	Con. (%) ³	Mass calc. Mass found ⁴
Y		1.49	85	7978.3 7981.2
Z		2.17	10	7972.7 7978.2
AA		0.59	80	7987.3 7988.2
AB		2.74	55	7977.3 7979.3
AC		0.60	80	7977.3 7978.3
AD		1.14	80	5209.7 5212.0
AE		3.69	80	7954.3 7954.7
AF		2.82	40	5215.8 5215.3
AG		1.41	100	5182.8 5181.2
AH		4.59	100	5212.2 5212.0
AI		3.70	70	5216.8 5217.0
AJ		2.75	85	5196.8 5195.0
AK		2.75	100	5194.8 5194.0
AL		4.42	100	5184.8 5184.0

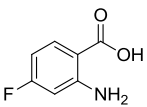
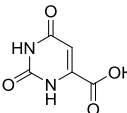
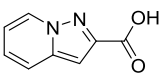
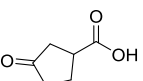
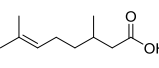
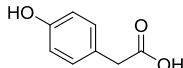
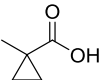
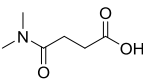
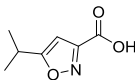
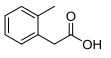
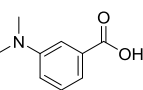
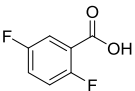
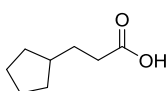
EXPERIMENTAL PART

No.	Structure	DNA conjugate 96		
		Yield ² [nmol]	Con. (%) ³	Mass calc. Mass found ⁴
AL		4.42	100	5184.8 5184.0
AM		1.76	100	5156.7 5157.0
AN		1.39	100	5206.8 5206.0
AO		1.05	100	5196.8 5197.0
AP		1.74	90	4991.5 4992.0
AQ		2.65	100	5180.8 5179.0
AR		4.51	55	5211.7 5207.1
AS		/	/	/
AT		3.89	100	5199.8 5198.0
AU		2.67	65	5208.8 5211.6
AV		0.10	90	4992.5 4992.0
AW		0.90	55	5209.6 5210.0
AX		1.60	100	5208.8 5208.0
AY		3.30	100	5193.8 5194.0

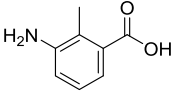
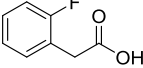
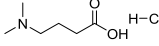
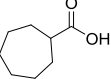
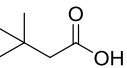
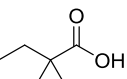
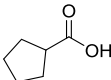
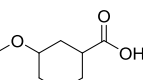
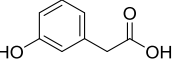
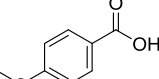
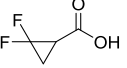
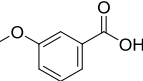
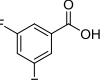
EXPERIMENTAL PART

No.	Structure	DNA conjugate 96		
		Yield ² [nmol]	Con. (%) ³	Mass calc. Mass found ⁴
AZ		5.12	100	5227.8 5226.0
BA		7.18	95	5207.8 5208.0
BB		4.88	100	5210.7 5210.0
BC		5.46	100	5208.7 5208.7
BD		17.87	100	5216.7 5215.0
BE		13.46	100	5251.8 5251.8
BF		6.76	100	5215.8 5215.0
BG		5.23	100	5182.8 5184.5
BH		4.36	100	5197.6 5197.0
BI		28.80	90	5190.8 5190.0
BJ		15.80	100	5196.8 5195.0
BK		4.13	100	5212.7 5212.0
BL		8.40	100	5184.8 5184.0

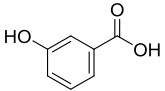
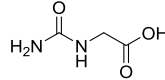
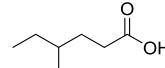
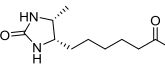
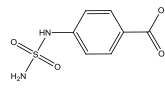
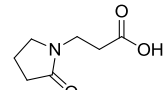
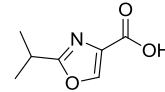
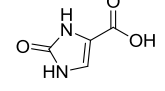
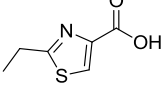
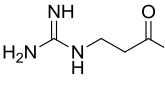
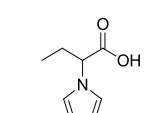
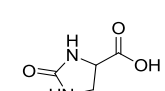
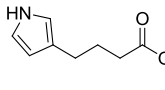
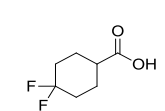
EXPERIMENTAL PART

No.	Structure	DNA conjugate 96		
		Yield ² [nmol]	Con. (%) ³	Mass calc. Mass found ⁴
BM		3.71	75	5209.7 5210.0
BN		10.51	85	5210.7 5210.0
BO		11.50	95	5216.8 5215.0
BP		9.91	90	5182.7 5182.0
BQ		0.15	60	5224.9 5225.0
BR		13.16	95	5206.8 5209.7
BS		0.47	100	4978.5 4978.8
BT		3.36	95	5199.8 5198.0
BU		1.66	55	5209.8 5211.8
BV		0.10	100	5204.8 5207.5
BW		0.10	100	5219.8 5221.0
BX		0.10	100	5212.7 5215.8
BY		2.44	100	5196.8 5197.0

EXPERIMENTAL PART

No.	Structure	DNA conjugate 96		
		Yield ² [nmol]	Con. (%) ³	Mass calc. Mass found ⁴
BZ		1.49	90	5205.8 5205.0
CA		/	/	/
CB		1.11	35	5222.2 5222.0
CC		0.10	90	5196.8 5195.0
CD		/	/	/
CE		0.91	95	5170.8 5170.0
CF		1.92	95	5168.8 5168.0
CG		3.74	90	5212.8 5212.0
CH		5.68	85	5206.8 5205.0
CI		3.59	100	5206.8 5206.0
CJ		0.10	100	5176.7 5171.4
CK		/	/	/
CL		/	/	/

EXPERIMENTAL PART

No.	Structure	DNA conjugate 96		
		Yield ² [nmol]	Con. (%) ³	Mass calc. Mass found ⁴
CM		0.34	100	5192.7 5192.0
CN		/	/	/
CO		0.68	100	5184.8 5185.0
CP		2.09	95	5270.8 5270.0
CQ		3.23	90	5268.9 5269.0
CR		3.39	100	7983.6 7987.3
CS		0.82	100	7981.6 7982.6
CT		2.74	100	7943.5 7944.5
CU		0.60	100	7972.6 7974.0
CV		1.68	100	5185.8 5186.2
CW		/		
CX		/		
CY		/		
CZ		/		

¹ **96A-G, 96I-AB** and **96CR-CU** were coupled to a 23mer DNA, all other compounds to a 14mer DNA

² measured by nanodrop

³ % conversion estimated based on the area under the curve of the product versus the educt in the HPLC-chromatograms, 100 %: no starting material detectable.

⁴ measured by MALDI MS

VI-4. Coupling of carboxylic acid building blocks to DNA scaffold conjugates

All carboxylic acids **A-CV** were coupled to DNA-scaffold conjugates containing a PEG(8) linker serving as spacer between DNA and scaffold, except for carboxylic acids **AJ, AO, AS, AT, AU**, Carboxylic acids **AP, AV** and **BS** that were coupled to the DNA-scaffold conjugate **95** containing a PEG(4) linker.

<i>DNA conjugates 96A-CZ</i>	
<i>Conversion</i>	<i>acids</i>
0%	20
1-25%	2
26-50%	5
51-75%	12
76-100%	75
Total	94

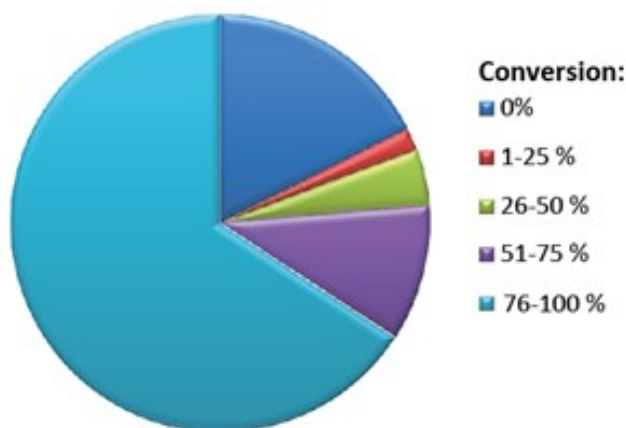
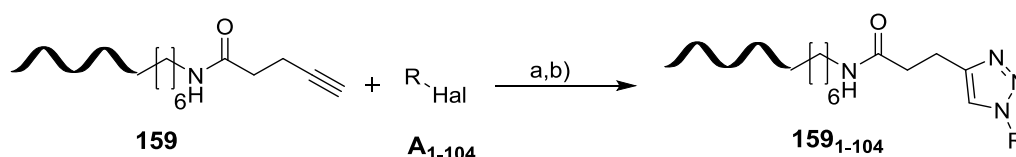


Figure VI-1: Statistical analysis of the coupling efficiency of carboxylic acids A-DJ to the DNA-scaffold conjugate **95**.

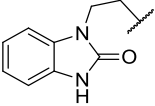
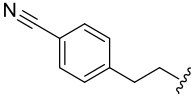
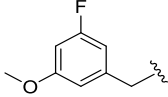
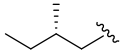
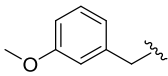
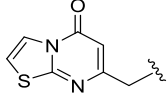
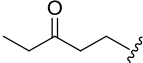
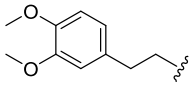
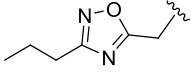
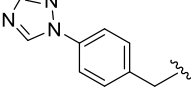
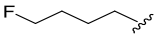
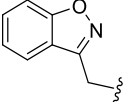
VI-5. Evaluation of halides for library synthesis

The evaluation of halides for library synthesis was done by **Mateja Klika Skopic**. For this purpose, DEAE sepharose (100 μL) was washed with 10 mM aq. NaAc buffer ($2 \times 350 \mu\text{L}$) and water ($2 \times 350 \mu\text{L}$). Then, 400 pmol of the DNA alkyne conjugate **101** (Table 16) was immobilized on DEAE anion exchange resin by incubation of an aqueous solution of the oligonucleotide conjugate (2 μL) with the resin for 15 min, followed by washing of the resin with 10 mM aq. NaAc buffer ($2 \times 350 \mu\text{L}$), distilled water ($2 \times 350 \mu\text{L}$) and MeOH/H₂O/DMF (2:2:1) mixture ($2 \times 350 \mu\text{L}$). For the synthesis of the azide for *in situ*-CuAAC, 10 μmol of a halide 1-104 (Table 16) was dissolved in 800 μL of DMF in an Eppendorf tube, to this were added 100 μL of an aqueous Na-azide solution (123 μmol , 8 mg/100 μL) and 100 μL of TBAI in DMF (20 μmol , 7.5 mg/100 μL). Both Na-azide and TBAI solutions were prepared as stock solutions immediately before the substitution reaction. In order to generate the azide from halide, the reaction was shaken for 8 hour at RT. In case of aliphatic bromides and all chlorides, azide formation was performed for 4 h at 70 $^{\circ}\text{C}$. For parallel CuAAC, the DEAE sepharose carrying the oligonucleotide conjugate was suspended in 100 μL of DMF. Subsequently, 380 μL of H₂O/MeOH (1:1), the azide (1 μmol in 100 μL of DMF/H₂O (9:1), 2500 eq.), TBTA (0.053 mg in 20 μL of DMF, 0.05 μmol , 125 eq.), Na-ascorbate (0.02 mg in 10 μL of H₂O, 0.05 μmol , 125 eq.) and CuSO₄·5H₂O (0.00125 mg in 10 μL of H₂O, 0.0005 μmol , 1.25 eq.) were added to the suspension in this order. All solutions had been prepared as stock solutions immediately before the CuAAC reaction was started. The reaction mixtures were shaken at 45 $^{\circ}\text{C}$ for overnight. Then, DEAE sepharose was filtered over receiver plate (20 μm) and washed subsequently with each 3x200 μL of DMF, 0.1 N aqueous EDTA to remove the copper-ion contaminants, MeOH/H₂O/DMF (2:2:1)mixture, water and 10 mM aq. NaAc buffer. After that, the oligonucleotide conjugates were eluted from DEAE sepharose by incubation with 60 μL of 3 M NaAc buffer (pH= 4.75) for 30 min. The products were analyzed by MALDI-TOF-MS analysis.

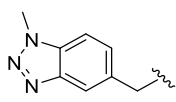
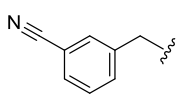
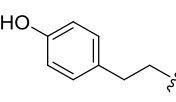
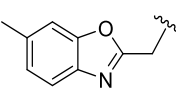
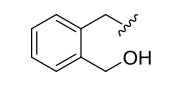
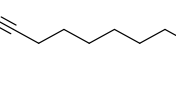
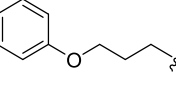
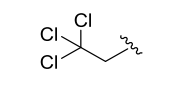
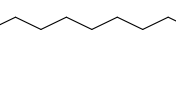
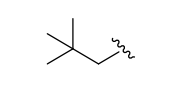
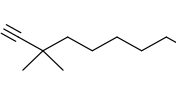
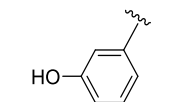


Scheme VI-1: Evaluation of halides **A1-A104**. a) NaN₃, TBAI, DMF/H₂O; b) CuSO₄, TBTA, Na-ascorbate, DMF, MeOH, H₂O, DEAE-sepharose.

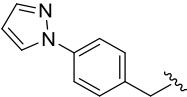
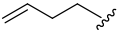
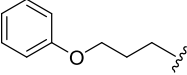
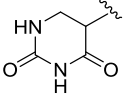
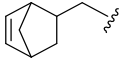

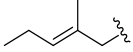
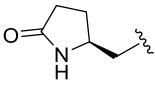
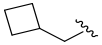
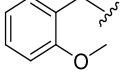
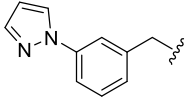

Table 15: Evaluation of halides A1-A104.

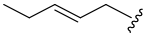
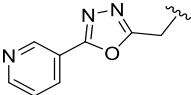
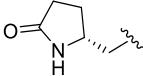
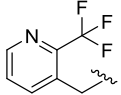
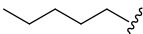
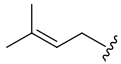
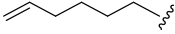
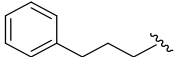
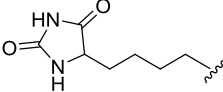
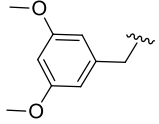
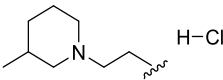
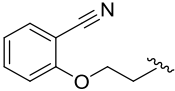
halides A1-A104			
No. ¹	R	halide	mass calc. mass found ²
A1		Cl	4692.1 4689.3
A2		Br	4661.1 4669.0
A3		Br	4670.1 4669.0
A4		Br	4602.1 4600.0
A5		Br	4652.1 4653.0
A6		Cl	4696.1 Not found
A7		Cl	4616.0 4618.3
A8		Br	4696.2 4693.0
A9		Cl	4656.1 4661.0
A10 ³		Br	4689.1 4698.0
A11		Br	4606.0 4610.0
A12		Br	4663.1 4664.0

EXPERIMENTAL PART

halides A1-A104			
No. ¹	R	halide	mass calc. mass found ²
A13		Br	4677.1 4672.1
A14		Br	4647.1 4653.0
A15		Br	4652.1 4656.7
A16 ³		Cl	4677.1 4668.5
A17		Br	4652.1 4651.3
A18		Br	4641.1 4650.0
A19		Br	4680.2 4682.5
A20		Cl	4663.4 4670.6
A21		Br	4660.2 4663.0
A22		Br	4602.1 4603.5
A23		Br	4655.2 4648.0
A24		Cl	4652.1 4653.9

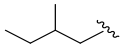
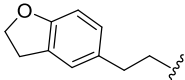
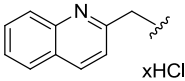
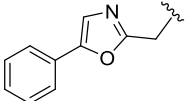
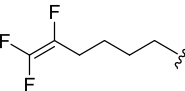
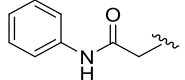
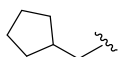
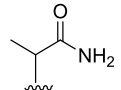
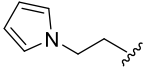
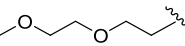
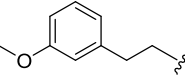
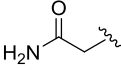
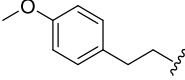
EXPERIMENTAL PART

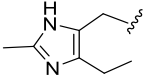
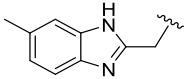
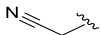
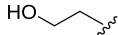
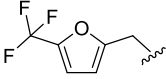
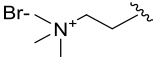
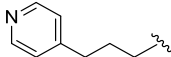
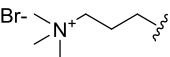
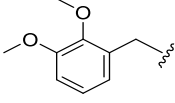
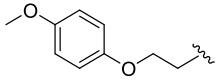
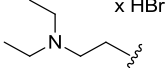
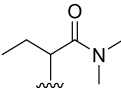
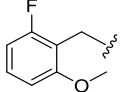
halides A1-A104			
No. ¹	R	halide	mass calc. mass found ²
A25		Br	4688.1 4693.6
A26		Br	4586.0 4583.1
A27		Br	4666.1 4654.3
A28		Br	4644.0 4647.6
A29		Cl	4638.1 4639.4
A30		Br	4588.0 4599.4
A31		Br	4614.1 4619.3
A32		Br	4629.1 4629.9
A33		Br	4600.1 4581.6
A34		Cl	4651.8 4657.8
A35		Cl	4688.1 4686.9
A36		Br	4616.1 4626.4

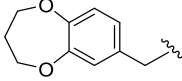
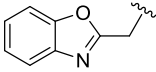
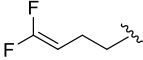
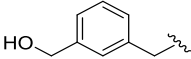
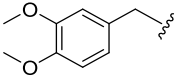
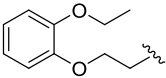
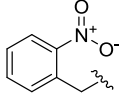
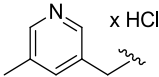
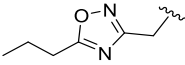
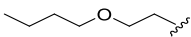
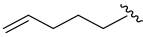
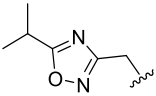
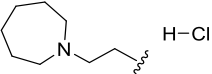
halides A1-A104			
No. ¹	R	halide	mass calc. mass found ²
A37		Br	4600.1 4600.2
A38 ³		Cl	4691.1 4691.9
A39		Br	4629.1 4637.4
A40		Cl	4691.1 4691.5
A41		Br	4602.1 4602.8
A42		Br	4600.1 4599.8
A43		Br	4599.1 4600.4
A44 ³		Cl	4650.1 4649.3
A45		Br	4686.1 4679.0
A46		Br	4682.1 4683.1
A47		Cl	4693.6 4691.5
A48		Br	4677.1 4675.2

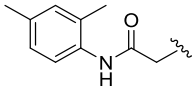
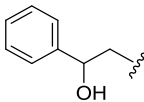
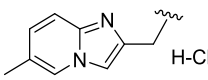
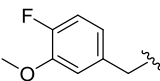
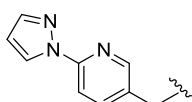
EXPERIMENTAL PART

halides A1-A104			
No. ¹	R	halide	mass calc. mass found ²
A49		Cl	4656.1 4644.2
A50		Br	4670.5 4670.5
A51 ³		Cl	4731.0 4731.8
A52		Cl	4670.1 4670.2
A53 ³		Cl	4710.6 4710.9
A54		Br	4676.1 4676.9
A55		Br	4677.4 4677.1
A56		Br	4598.1 4598.1
A57		Br	4622.1 4622.4
A58		Br	4689.1 4689.1
A59		Cl	4671.2 4671.3
A60		Cl	4654.1 4654.6

halides A1-A104			
No. ¹	R	halide	mass calc. mass found ²
A61		Br	4602.1 4602.4
A62 ³		Br	4678.1 4672.0
A63		Cl	4673.1 4673.1
A64		Cl	4689.1 4694.5
A65		Br	4668.0 4667.8
A66		Br	4665.1 4666.2
A67		Br	4614.1 4626.7
A68 ³		Br	4603.0 4602.6
A69		Br	4625.1 4625.1
A70		Br	4634.1 4622.8
A71		Br	4666.1 4666.5
A72		Br	4589.0 4585.2
A73		Cl	4666.1 4664.7

halides A1-A104			
No. ¹	R	halide	mass calc. mass found ²
A74		Cl	4654.1 4655.1
A75 ³		Cl	4676.1 4665.1
A76 ³		Br	4571.0 4571.1
A77 ³		Br	4576.0 4574.1
A78		Br	4680.0 4680.2
A79		Br	4603.1 4592.1
A80		Cl	4651.1 4650.2
A81		Br	4632.1 4630.5
A82		Cl	4682.1 4686.2
A83		Br	4682.1 4683.3
A84		Br	4631.1 4637.1
A85		Br	4645.1 4646.1
A86		Br	4670.1 4671.0

halides A1-A104			
No. ¹	R	halide	mass calc. mass found ²
A87		Cl	4694.1 not found
A88		Cl	4663.0 4660.1
A89		Br	4622.0 4613.2
A90		Br	4652.1 4655.2
A91		Cl	4682.1 4682.7
A92		Br	4696.1 4698.2
A93		Br	4667.1 4670.3
A94		Cl	4637.1 4638.0
A95		Cl	4656.1 4634.7
A96		Cl	4632.1 4630.8
A97		Br	4600.0 4609.3
A98		Cl	4656.1 4655.1
A99		Cl	4657.1 4655.3

halides A1-A104			
No. ¹	R	halide	mass calc. mass found ²
A100		Br	4693.1 4665.0
A101 ³		Br	4652.1 4658.0
A102 ³		Cl	4676.1 4675.0
A103		Br	4670.1 4665.4
A104 ³		Cl	4689.1 4690.0

¹ building blocks that did not yield products are marked in red

² measured by MALDI MS

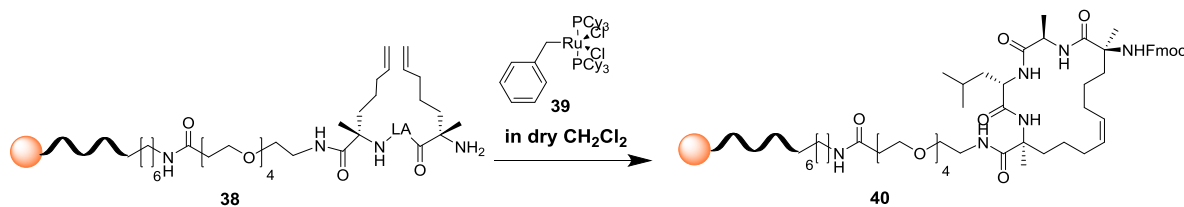
³ incomplete conversion

Table 16: Conversions of the DNA alkyne conjugate **101** to triazole.

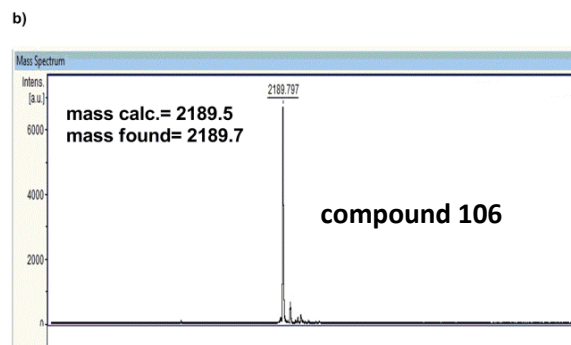
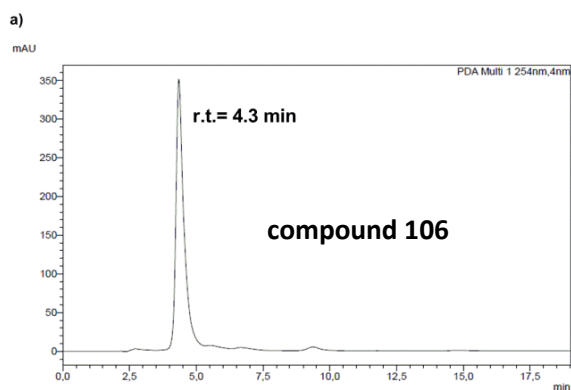
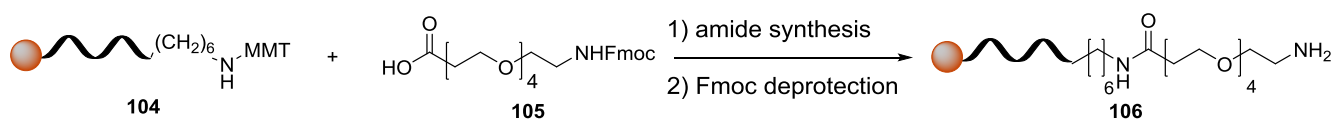
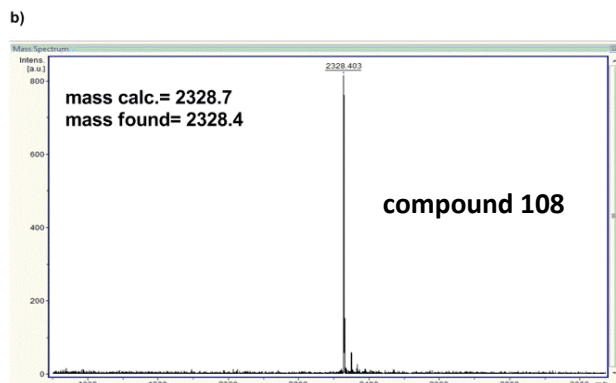
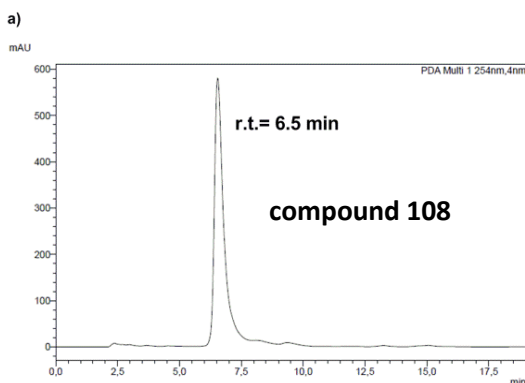
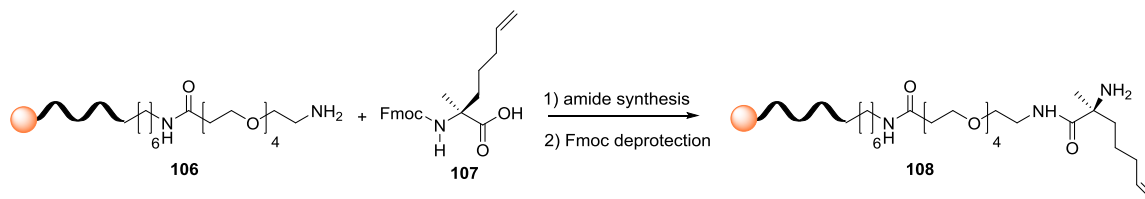
DNA alkyne conjugate	
	<i>Azides</i>
fully converted	82
incomplete conversion	14
no product observed	2
oxidation product observed	6
Total	104

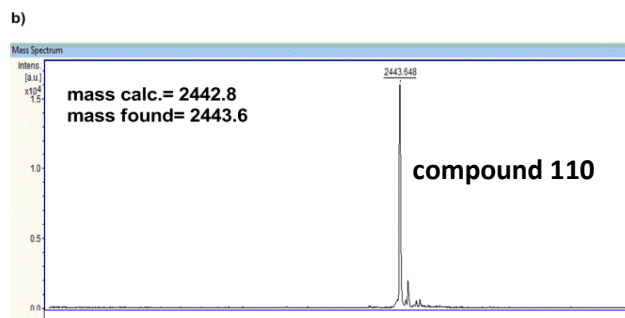
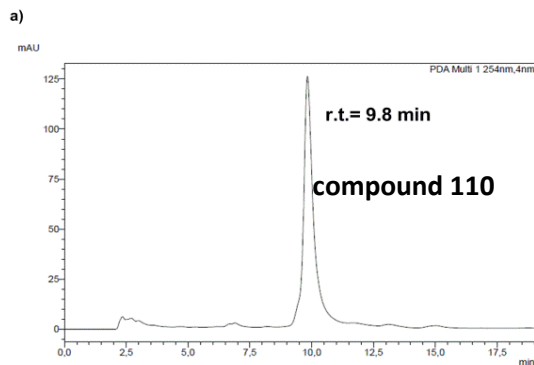
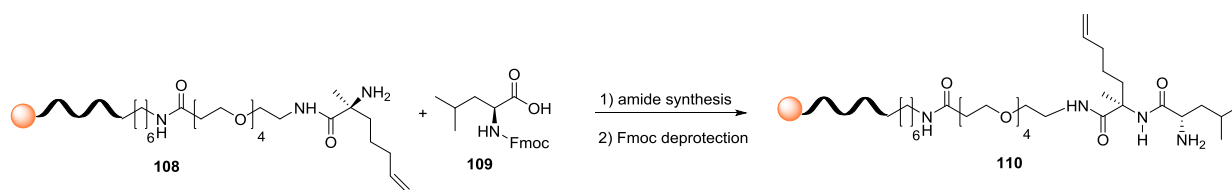
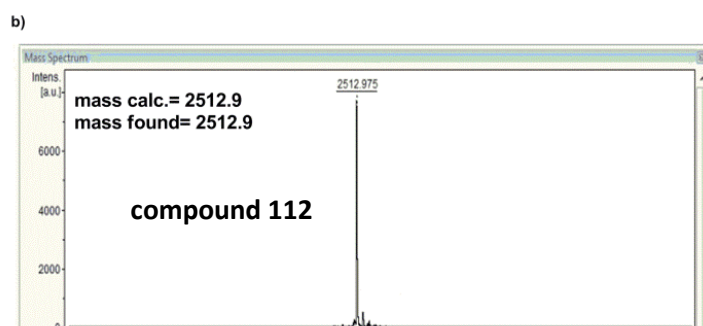
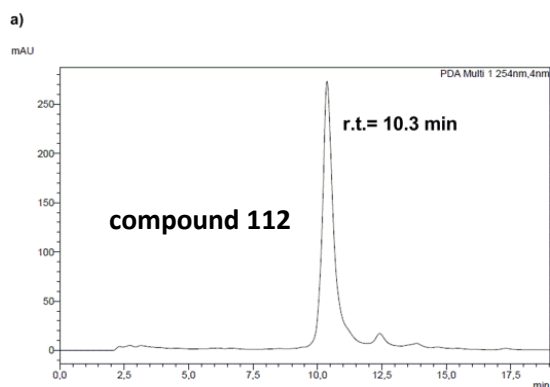
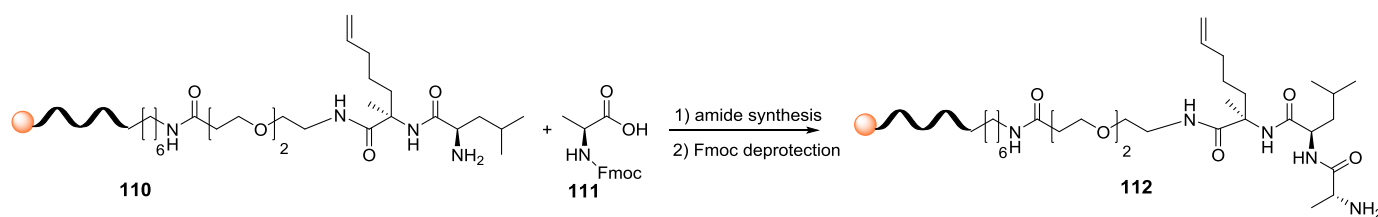
VI-6. Macrocyclization of peptides by ruthenium-catalyzed ring-closing metathesis

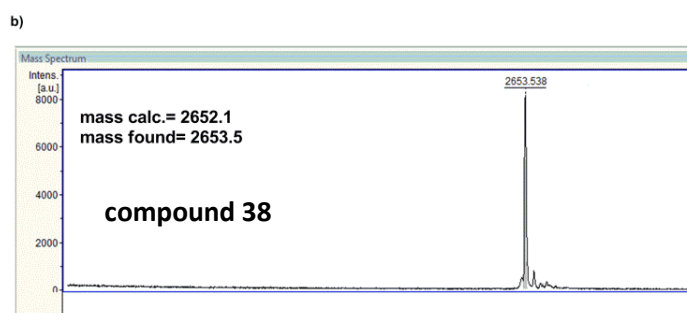
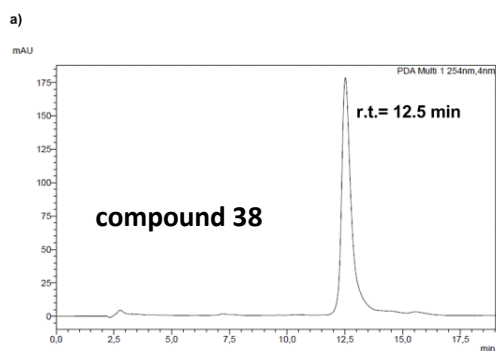
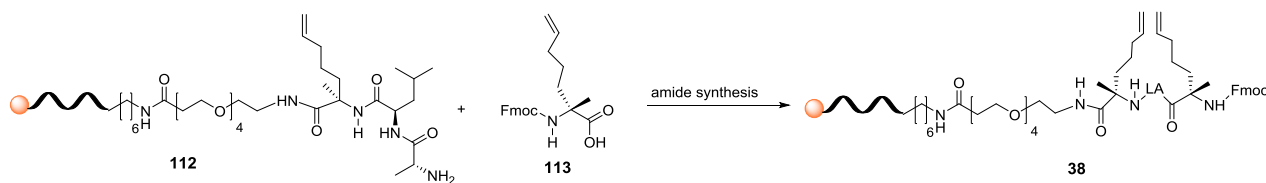
The synthesis of peptide-DNA conjugates was performed by repetitive Fmoc-solid-phase peptide



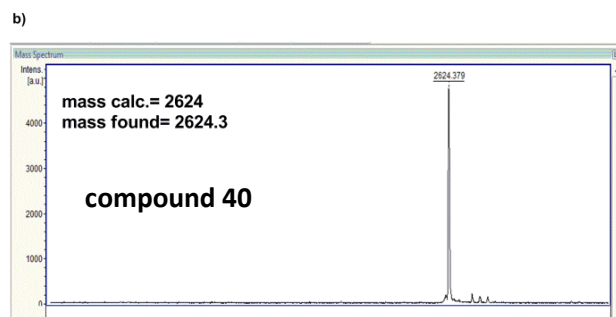
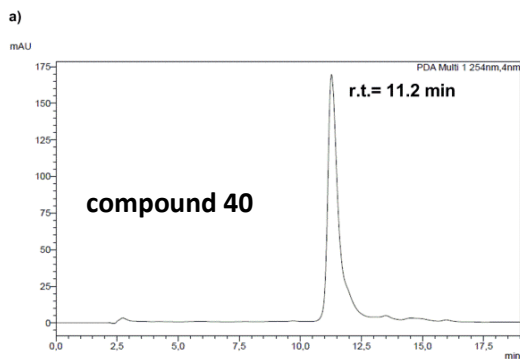
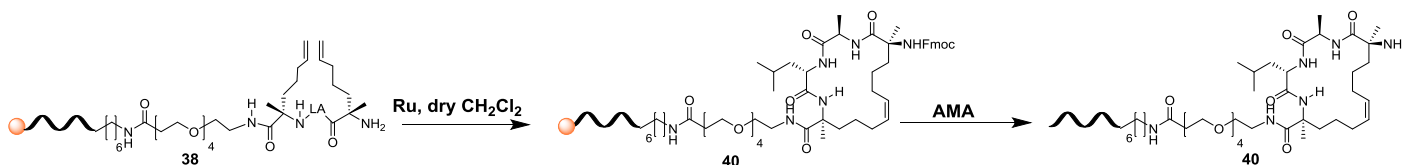
synthesis (SPPS) according to the general procedure 1 making use in the following order of the (1. *S*)-*N*-Fmoc- α -4-pentenylalanine **107** (25 μmol , 9.5 mg); 2. *N*- α -(9-Fluorenylmethyloxycarbonyl)-L-leucine **109** (22 μmol , 7.8 mg); 3. *N*- α -(9-Fluorenylmethyloxycarbonyl)-L-alanine **111** (19 μmol , 5.9 mg); 4. (*R*)-*N*-Fmoc- α -4-pentenylalanine **113** (16 μmol , 6.1 mg)). The Fmoc-deprotection was performed according to the general procedure 2. Crosslinking of unnatural olefinic amino acids was performed by ring closing metathesis (RCM) in a solution of the Grubbs 1st generation catalyst (Bis(tricyclohexylphosphine)benzylidene ruthenium(IV) dichloride, 4.9 mM, 4 mg) **39** dissolved in 1 mL of dry dichloromethane (DCM) three times for 2 h at room temperature (100 μl of the solution were used for each coupling). After each coupling, the DNA was washed with DCM (3 x 200 μl), the CPG was then dried *in vacuo* for 5 min and the next coupling was performed. Subsequent to the three couplings, the DNA was washed successively with DMF (3 x 200 μl), MeCN (3 x 200 μl), MeOH (3 x 200 μl) and DCM (3 x 200 μl). The filter carrying CPGs was then dried *in vacuo* for 15 min. The DNA-macrocyclic conjugate was deprotected and removed from the CPG by treatment with AMA (AMA= aqueous ammonia (30 %)/ aqueous methylamine (40 %) 1:1 vol/vol). Then, 20 μL of 1M Tris buffer was added, the mixture was dried in a SpeedVac, re-dissolved in 100 μL of distilled water and purified by HPLC on a Gemini 5u C_{18} 110A column; 100*10.0 mm with a gradient of aqueous triethylammonium acetate buffer (100 mM, pH= 8) and methanol (20:80 – 70:30 of methanol over 20 min).

VI-6.1. 1st amide coupling: Fmoc-NH-PEG(4)-COOHVI-6.2. 2nd amide coupling: Fmoc-X₅-OH

VI-6.3. 3rd amide coupling: Fmoc-Leu-OHVI-6.4. 4th amide coupling: Fmoc-Ala-OH

VI-6.5. 5th amide coupling: Fmoc-X_R-OH

VI-6.6. Ring-closing metathesis on CPG by ruthenium catalysis



VI-7. Development of an acid catalyzed approach to an oligo-thymidine initiated DNA-encoded β -carbolines library

VI-7.1. Synthesis of the 5'-amino-PEG(4)-hexathymidine conjugate "hexT"

The MMT-protective group of the 5'-aminolinker-modified hexathymidine bound to 1000 Å controlled pore glass (CPG) solid support (1 μ mol, ca. 40 mg) was removed by addition of 3 % trichloroacetic acid in dry DCM (3 x 200 μ L) for 3 x 1 min. A yellow color indicated successful removal of the protective group. The CPG containing the deprotected DNA was then washed three times with each 200 μ L of 1 % TEA in MeCN, DMF, MeOH, MeCN and DCM. The CPG, Fmoc-NH-PEG(4)-COOH, and HATU were dried *in vacuo* for 15 min. Stock solutions of all reactants in dry DMF were prepared immediately before the reaction was started. To 150 μ L of a solution of the Fmoc-NH-PEG(4)-COOH linker (54 mg, 100 μ mol, 100 eq.) in dry DMF were added HATU (19 mg, 50 μ M, 100 eq.) dissolved in 150 μ L of dry DMF and DIPEA (42 μ L, 250 eq.). This reaction mixture was shaken for 5 min and added to the solid support-bound DNA suspended in dry DMF (150 μ L). The amide coupling reaction was shaken at room temperature for 2 hours. Then, the CPG containing the DNA-PEG linker conjugate was filtered off using a filter column and washed subsequently with each 3 x 200 μ L of DMF, MeOH, MeCN and DCM. Unreacted amines were capped with acetic acid anhydride (a 1:1 mixture of THF/methylimidazole, 9:1, vol/vol, and THF/pyridine/acetic acid anhydride, 8:1:1, vol/vol was used), and the CPG was again washed subsequently with each 3 x 200 μ L of DMF, MeOH, MeCN and DCM, and dried *in vacuo* for 15 min. For analysis, an aliquot of ca. 10 μ mol of the DNA-PEG conjugate was deprotected and cleaved from the CPG by treatment with 500 μ L of AMA (AMA= aqueous ammonia (30 %)/ aqueous methylamine (40 %), 1:1, vol/vol) for 30 min at room temperature. Then, 20 μ L of 1 M Tris buffer (pH= 7.5) were added, the mixture was dried in a SpeedVac, re-dissolved in 100 μ L of distilled water, and the product was analyzed by RP-HPLC (Gemini, 5u, C18, 110A column; 100*10.0 mm) with a gradient of aqueous triethylammonium acetate buffer (100 mM, pH= 8) and methanol (20% – 70% of methanol over 19 min). For further coupling of carboxylic acids to the linker, the Fmoc-group (1 μ mol, ca. 40 mg of the solid support) was removed with 20 % piperidine in dry DMF (0.4 mL). The reaction mixture was shaken for 5 min at room temperature. The CPG containing the deprotected hexT was washed three times with each 200 μ L of DMF, MeOH, MeCN and DCM.

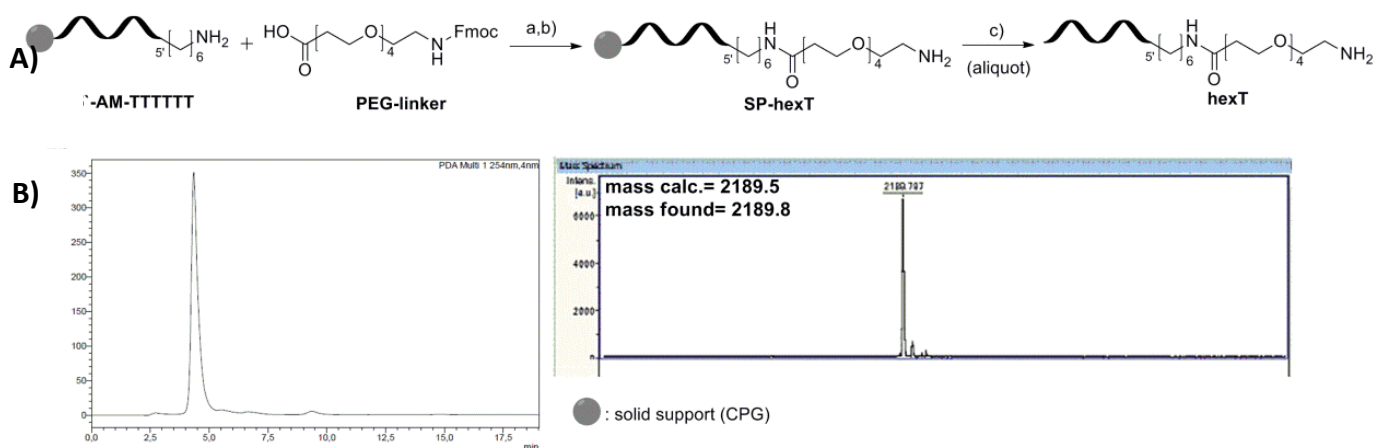


Figure VI-2: Synthesis of the 5'-amino-PEG (4)-linker modified hexathymidine oligonucleotide ("hexT"). Reagents and conditions: a) HATU, DIPEA, dry DMF, room temperature, 2 hours; b) 20 % piperidine in dry DMF; c) cleavage of an aliquot of the hexT from the solid support: AMA (aqueous ammonia (30 %)/ aqueous methylamine (40 %) = 1:1, vol/vol), 30 min, room temperature. b) HPLC trace and MALDI-MS spectrum of the hexT. Filled circle denotes solid support (CPG).

VI-7.2. Synthesis of hexT conjugate 141

The hexT **104**, either Fmoc-L-tryptophane **140** (100 μ mol, 100 eq.) and HATU (100 μ mol, 100 eq.) were dried *in vacuo* for 5 min. To a 250 μ L of a solution of either Fmoc-L-tryptophane (100 μ mol, 100 eq) in dry DMF were added HATU (38 mg, 100 μ mol, 100 eq.) dissolved in 250 μ L of dry DMF, and DIPEA (43 μ L, 250 μ mol, 250 eq.). The acids were activated for 5 min at room temperature. They were added to the solid support-bound HexT suspended in dry DMF (250 μ L). The amide coupling reactions were shaken at room temperature for 4 hours. The CPG containing the amide coupling product **141** was filtered off, and washed subsequently with each 3 x 200 μ L of DMF, MeOH, MeCN and DCM. Unreacted amines were capped with acetic acid anhydride (a 1:1 mixture of THF/methylimidazole, 9:1, vol/vol, and THF/pyridine/acetic acid anhydride, 8:1:1, vol/vol was used), the CPG was washed again with each 3 x 200 μ L of DMF, MeOH, MeCN and DCM, and dried *in vacuo* for 15 min. For analysis, an aliquot of ca. 10 nmol of conjugate **141** was deprotected and cleaved from the CPG with 500 μ L of AMA (AMA= aqueous ammonia (30 %)/ aqueous methylamine (40 %), 1:1, vol/vol) for 30 min at room temperature. Then, 20 μ L of 1 M Tris buffer (pH= 7.5) were added, the mixture was dried in a SpeedVac, redissolved in 100 μ L of distilled water, and the product was analyzed by RP-HPLC (Gemini, 5u, C18, 110A column; 100*10.0 mm) with a gradient of aqueous triethylammonium acetate buffer (100 mM, pH= 8) and methanol (20% – 70% of methanol over 19 min).

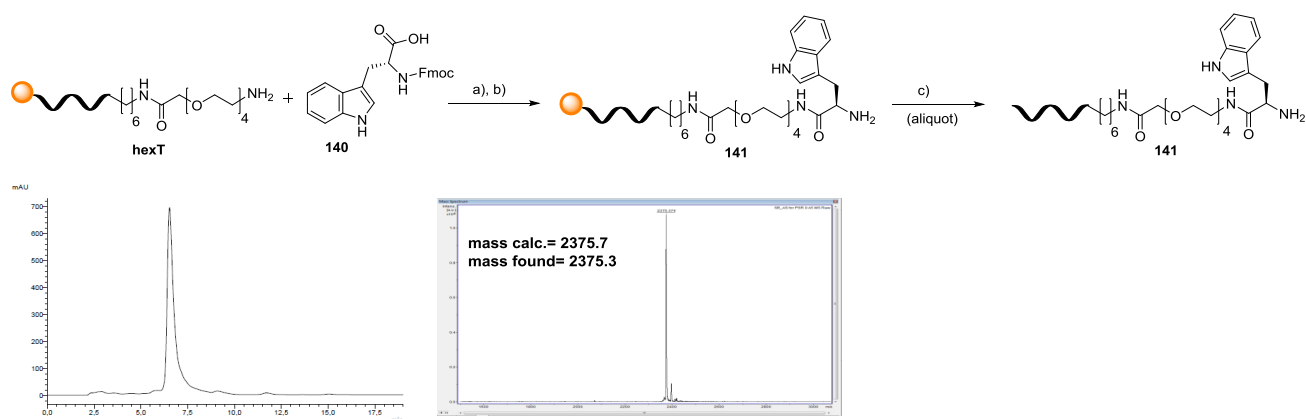


Figure VI-3: Synthesis of hexT-tryptophane conjugate **141**. Reagents and conditions: a) HATU, DIPEA, dry DMF, room temperature, 4 hours; b) 20 % piperidine in dry DMF; c) AMA (aqueous ammonia (30 %)/ aqueous methylamine (40 %), 1:1, vol/vol), 30 min, room temperature (aliquot); b) HPLC trace and MALDI-MS spectrum of **141**.

VI-8. Synthesis of hexT- β -carboline conjugates **142A – 142DJ** by Pictet-Spengler reaction, and building block validation for tiDEL synthesis

VI-8.1. Optimization of the reaction conditions for the synthesis of hexT- β -carboline conjugates

A solution of a catalyst in a solvent (**table 1**) was freshly prepared as stock. Aldehyde **CR** (20 μ mol, **table S3**) was dissolved in this solution to a final concentration of 0.44 M. This solution (45 μ L, 20 μ mol of **CR**) was added to 20 nmol (ca. 0.8 mg) of CPG-bound hexT-conjugate **141**. The reaction was run for different times as given in **table 1**. The CPG was filtered off, washed successively with DMF (3 x 200 μ L), MeCN (3 x 200 μ L), MeOH (3 x 200 μ L) and DCM (3 x 200 μ L), and dried *in vacuo* for 15 min. The DNA was removed from the CPG by treatment with 500 μ L of AMA (AMA= aqueous ammonia (30 %)/ aqueous methylamine (40 %) 1:1 vol/vol) at room temperature for 30 min. Then, 20 μ L of 1M Tris-buffer (pH=7.5) was added, the mixture was dried in a SpeedVac, re-dissolved in 100 μ L of distilled water and the product was purified by HPLC on a Gemini 5u C₁₈ 110A column; 100*10.0 mm with a gradient of aqueous triethylammonium acetate buffer (100 mM, pH= 8) and methanol (20:80 – 70:30 of methanol over 19 min).

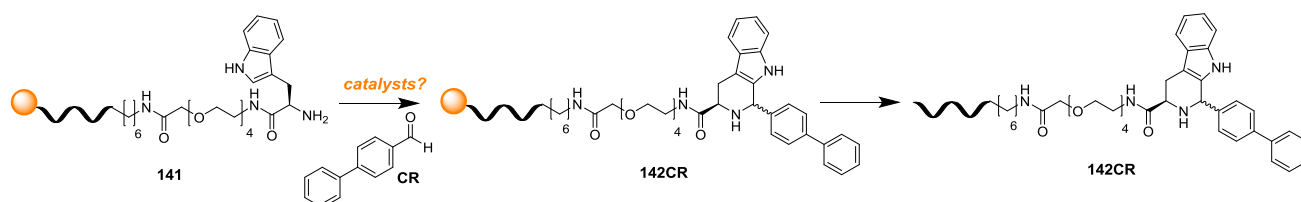


Figure VI-4: Optimization of reaction conditions for the synthesis of hexT-β-carbolines. a) AMA (aqueous ammonia (30 %)/ aqueous methylamine (40 %)= 1:1, vol/vol), 30 min, room temperature.

VI-8.2. Synthesis of hexT-β-carboline conjugates 142A – 142DJ

The CPG (400 nmol, ca. 16 mg) containing the deprotected hexT conjugate **141** was suspended in dry DCM and split into 20 aliquots of each 20 nmol in Eppendorf tubes. These were evaporated to dryness in a speed vac. A solution of 1% TFA in dry CH₂Cl₂ was freshly prepared as stock. Aldehydes **A – DJ** were dissolved in 1% TFA in dry CH₂Cl₂ to obtain a 0.44 M solution (20 μmol of each aldehyde **A – DJ** in 45 μL of 1% TFA in dry CH₂Cl₂). For the Pictet-Spengler reaction, aldehydes **A – DJ** dissolved in 1% TFA in dry CH₂Cl₂ (45 μL, 20 μmol of the aldehyde) were added to the 20 nmol aliquots of CPG-bound **141**. Typically 20 reactions were run in parallel at room temperature for 18 h. The CPG was filtered off, washed successively with DMF (3 x 200 μl), MeCN (3 x 200 μl), MeOH (3 x 200 μl) and DCM (3 x 200 μl), and dried *in vacuo* for 15 min. The DNA was removed from the CPG by treatment with 500 μL of AMA (AMA= aqueous ammonia (30 %)/ aqueous methylamine (40 %) 1:1 vol/vol) at room temperature for 30 min. Then, 20 μL of 1M Tris-buffer (pH=7.5) was added, the mixture was dried in a SpeedVac, re-dissolved in 100 μL of distilled water and purified by HPLC on a Gemini 5u C₁₈ 110A column; 100*10.0 mm with a gradient of aqueous triethylammonium acetate buffer (100 mM, pH= 8) and methanol (20:80 – 70:30 of methanol over 19 min).

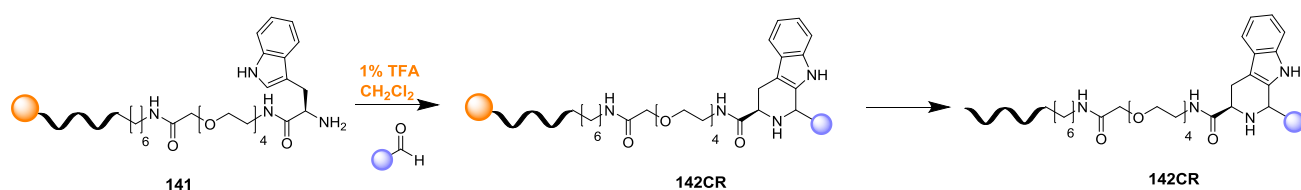
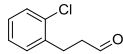
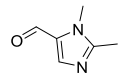
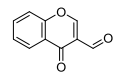
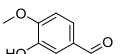
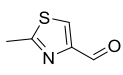
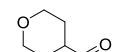
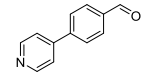
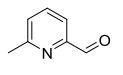
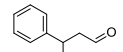
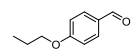
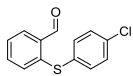

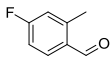


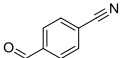
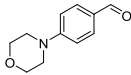
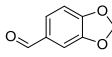


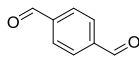
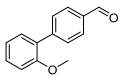
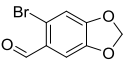
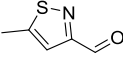
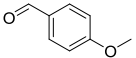
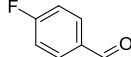


Figure VI-5: Synthesis of hexT-conjugates **142A-142DJ**.

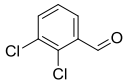
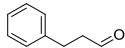
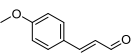
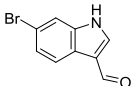
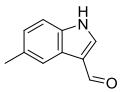


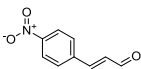
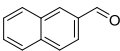

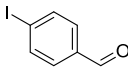

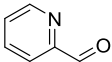
Table 17: Total yield, conversion rates, and MALDI MS data of compounds **142A – 142DJ**.

No.	structure	yield [nmol] ^[a]	conversion (%) ^[b]	mass calc. mass found ^[c]
A		2.6	>95	2526.3 2525.5
B		2.6	80	2481.8 2482.1
C		1.9	50	2531.9 2532.3
D ^[e]		1	70	2509.9 2510.5
E		2.4	85	2484.9 2482.2
F		1.6	80	2471.8 2471.5
G ^[e]		5.5	85	2540.9 2540.6
H		5.4	85	2478.8 2479.7
I		1.9	>95	2505.9 2506.0
J		1.9	80	2521.9 2521.6
K		1.8	>95	2606.4 2606.5
L ^[f]		2.8	>95	2504.8 2505.4

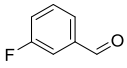
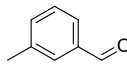
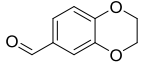
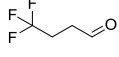
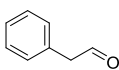
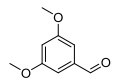
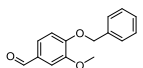
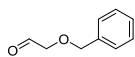
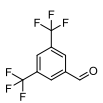
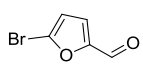
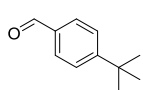
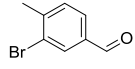
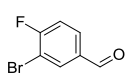
EXPERIMENTAL PART

No.	structure	yield [nmol] ^[a]	Conversion (%) ^[b]	mass calc. mass found ^[c]
M		0.7	90	2495.8 2496.4
N ^[g]		1	85	2453.8 2450.5
O		4.9	90	2467.8 2468.6
P		0.8	90	2488.8 2487.5
Q		1.5	>95	2548.9 2548.7
R ^[e]		1.7	>95	2507.8 2507.5
S		2.6	>95	2443.8 2443.6
T		1	80	2429.8 2429.5
U ^[d]		1.3	90	2491.8 2491.2
V		1.1	60	2569.9 2569.5
W ^[e]		1.7	80	2586.7 2587.7
X		1.7	85	2484.9 2483.1
Y		6.8	60	2493.9 2493.5
Z		8.7	90	2478.8 2477.7

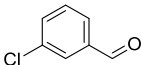

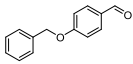
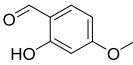
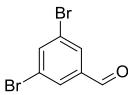
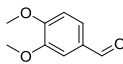
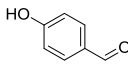
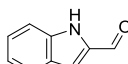
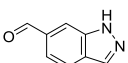
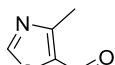
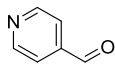
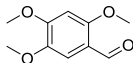
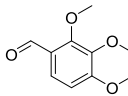
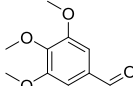
EXPERIMENTAL PART

No.	structure	yield [nmol] ^[a]	conversion (%) ^[b]	mass calc. mass found ^[c]
AA		9.2	90	2532.7 2533.3
AB		6.8	>95	2491.9 2492.4
AC		7.7	60	2519.9 2519.5
AD ^[h]		2.2	40	2581.8 2583.4
AE ^[h]		1	90	2516.9 2513.9
AF ^[g]		5.9	90	2453.8 2455.0
AG		1.4	90	2453.8 2453.5
AH ^[e]		0.4	90	2534.9 2533.3
AI		0.5	60	2513.9 2513.9
AJ		1.7	>95	2429.8 2429.7
AK		5.6	80	2589.7 2590.2
AL		2.1	90	2443.8 2444.6
AM ^[e]		0.8	55	2464.8 2468.2

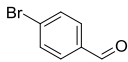
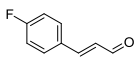
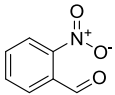
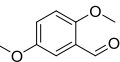
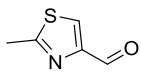
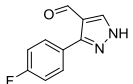
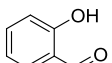
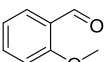
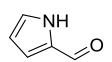
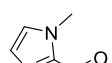

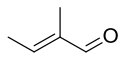
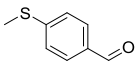
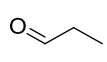
EXPERIMENTAL PART

No.	structure	yield [nmol] ^[a]	conversion (%) ^[b]	mass calc. mass found ^[c]
AN		1.7	>95	2481.8 2479.7
AO		1.8	85	2477.9 2477.7
AP ^[e]		1.3	>95	2521.9 2521.8
AQ		0.9	>95	2483.8 2484.1
AR		1.6	85	2463.8 2478.5
AS		1.3	>95	2523.9 2524.4
AT		0.7	80	2599.9 2600.4
AU		1.9	85	2507.9 2508.3
AV		0.7	90	2599.8 2600.1
AW		2.9	80	2532.7 2533.4
AX		0.5	90	2519.9 2520.5
AY		1	75	2556.7 2557.6
AZ		1.5	70	2560.7 2557.2


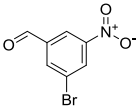
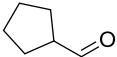
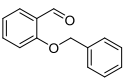

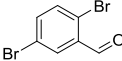
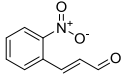
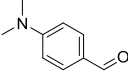
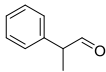

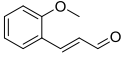
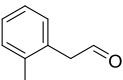
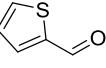
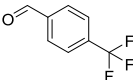
EXPERIMENTAL PART

No.	structure	yield [nmol] ^[a]	conversion (%) ^[b]	mass calc. mass found ^[c]
BA		1.3	90	2498.3 2494.5
BB		1.8	>95	2469.9 2469.1
BC		1.5	90	2569.9 2570.4
BD ^[e]		1.4	45	2523.9 2524.4
BE		1.6	75	2621.6 2622.3
BF		3.9	85	2523.9 2524.2
BG ^[i]		5.3	>95	2479.8 2479.7
BH ^[h]		5.6	70	2502.9 2505.6
BI ^[j]		4.8	70	2503.8 2505.4
BJ ^[e]		2.6	>95	2484.9 2485.5
BK		1.3	>95	2464.8 2461.3
BL		0.8	90	2553.9 2553.9
BM		1.6	85	2553.9 2554.1
BN		0.4	85	2553.9 2553.9

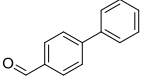

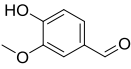
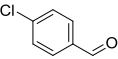
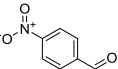
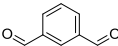
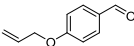
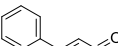
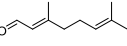

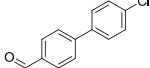
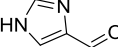
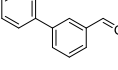
EXPERIMENTAL PART

No.	structure	yield [nmol] ^[a]	conversion (%) ^[b]	mass calc. mass found ^[c]
BO		2.7	80	2542.7 2541.1
BP		3.1	90	2507.8 2508.2
BQ		1.8	80	2508.8 2508.8
BR		1	80	2523.9 2524.4
BS		0.8	>95	2484.9 2485.8
BT ^[i]		0.6	65	2547.9 2549.8
BU		1.8	85	2481.8 2482.5
BV		3.9	90	2479.8 2480.2
BW ^[g]		1.1	90	2452.8 2454.2
BX ^[g]		1.8	85	2467.8 2468.6
BY		2.8	90	2457.8 2458.4
BZ		1.6	60	2441.8 2443.6
CA		2.1	85	2509.9 2510.4
CB		1.8	80	2415.8 2415.1


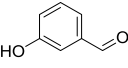
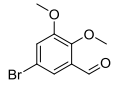
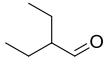
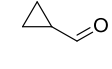
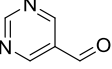
EXPERIMENTAL PART

No.	structure	yield [nmol] ^[a]	conversion (%) ^[b]	mass calc. mass found ^[c]
CC ^[g]		1.2	65	2467.8 2466.3
CD ^[i]		2.7	80	2587.7 2573.3
CE		3.5	60	2455.8 2455.1
CF		1.2	75	2569.9 2470.4
CG		3.4	90	2429.8 2429.9
CH		1.4	90	2621.6 2622.5
CI ^[j]		3.7	>95	2534.9 2536.1
CJ		4.2	70	2506.9 2507.4
CK		1.4	70	2491.9 2492.1
CL		1.9	85	2539.9 2541.6
CM		3.8	65	2519.9 2520.2
CO		2.1	90	2491.9 2491.3
CP		5.1	85	2469.9 2469.1
CQ		1.8	90	2528.8 2527.3

EXPERIMENTAL PART

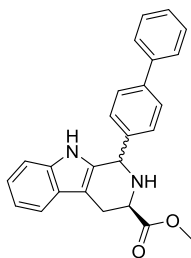
No.	structure	yield [nmol] ^[a]	conversion (%) ^[b]	mass calc. mass found ^[c]
CR		2.8	>95	2539.9 2539.5
CS		1.5	55	2401.8 2402.5
CT		2.2	85	2509.8 2509.4
CU		2.6	>95	2498.3 2493.9
CV ^[ij]		2.7	>95	2508.8 2509.4
CW		0.5	>95	2491.8 2491.6
CX		3.6	>95	2519.9 2520.2
CY		1.2	90	2489.9 2489.9
CZ		5.0	80	2519.9 2520.2
DA		1.2	50	2463.8 2462.1
DB		5.2	>95	2574.4 2574.3
DC		4.9	60	2453.8 2451.1
DD		3.9	>95	2539.9 2539.8

EXPERIMENTAL PART

No.	structure	yield [nmol] ^[a]	conversion (%) ^[b]	mass calc. mass found ^[c]
DE		3.5	90	2387.7 2387.3
DF ^[e]		2.4	90	2479.8 2478.0
DG		1.5	>95	2602.8 2603.1
DH		5.1	90	2457.9 2459.1
DI		5.4	>95	2427.8 2427.4
DJ ^[i]		2.9	90	2465.8 2465.3

VI-9. Synthesis of reference molecules and educts

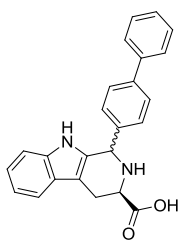
Methyl (R)-1-([1,1'-biphenyl]-4-yl)-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole-3-carboxylate (**152**)



In a 50 mL two-necked round-bottomed flask equipped with a magnetic bar and an argon-filled balloon, L-tryptophan methyl ester hydrochloride **151** (637 mg, 1 eq, 2.5 mmol) was suspended in dry CH_2Cl_2 (10 mL). The solution was cooled to 0°C in an ice bath, and biphenyl-4-carboxaldehyde **CR** (911 mg, 2 eq, 5 mmol) was added. To this solution TFA (0.5 mL) was added dropwise and the reaction was stirred at

room temperature for 18 hours under argon atmosphere. On completion, the reaction mixture was basified with dilute NH_4OH solution and extracted with CH_2Cl_2 (3×50 mL). The organic layer was washed with water, brine, dried over anhydrous Na_2SO_4 , filtered, and evaporated under reduced pressure. The residue was purified by silica gel column chromatography using a gradient of $\text{CH}_2\text{Cl}_2/\text{MeOH}$ (100:0 to 95:5) to yield the β -carboline as a mixture of two diastereomers in a ratio of 1:1 according to NMR (sticky gum, 231 mg, 24 % yield). ^1H NMR (600 MHz, CDCl_3) δ ppm: 10.34 (br. s, 1H), 7.62 (d, $^3J = 7.68$ Hz, 1H), 7.61 (d, $^3J = 8.34$ Hz, 2H), 7.57 (d, $^3J = 7.68$ Hz, 1H), 7.55 (d, $^3J = 8.34$ Hz, 2H), 7.44 (t, $^3J = 7.86$ Hz, 2H), 7.35 (t, $^3J = 7.38$ Hz, 2H), 7.31 (d, $^3J = 8.04$ Hz, 2H), 7.20 (d, $^3J = 7.86$ Hz, 2H), 7.14–7.07 (m, 4H) (overlap due to diastereomerism), 6.22 (s, 1H), 5.30 (br. s, 1H), 3.88 (dd, $^3J = 10.08$ Hz, 1H), 3.84 (s, 3H), 3.79 (dd, $^3J = 10.08$ Hz, 1H), 3.74 (s, 3H), 3.39 (dd, $^3J = 10.08$ Hz, $^2J = 2.94$ Hz, 1H), 3.08 (dd, $^3J = 14.52$ Hz, $^2J = 2.4$ Hz, 1H). ^{13}C NMR (600 MHz, CDCl_3) δ ppm: 175.28, 139.46, 137.73, 136.58, 135.19, 134.50, 127.84, 127.77, 127.60, 126.63, 126.34, 126.10, 125.98, 120.47, 117.86, 117.19, 109.66, 54.00, 51.38, 40.19, 28.42. MS (ESI): calculated for $\text{C}_{25}\text{H}_{22}\text{N}_2\text{O}_2$ 382.17, found 383.51 ($[\text{M}+\text{H}]^+$). Purity (HPLC): 97 %.

(R,S)-([1,1'-Biphenyl]-4-yl)-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole-3-(R)-carboxylic acid (**153**)



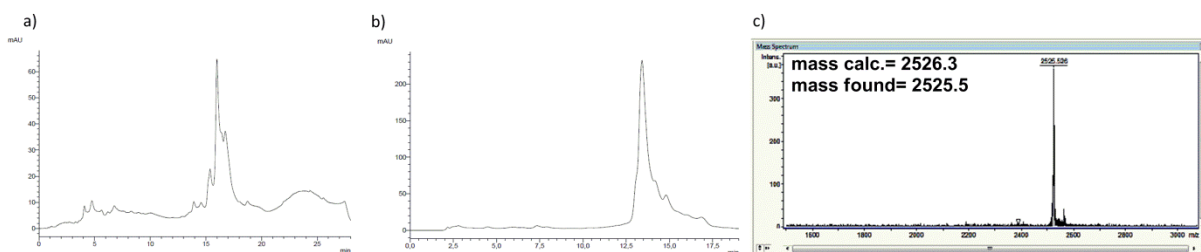
In a 50 mL round-bottomed flask equipped with a magnetic bar, a solution of ester **152** (200 mg, 1 eq, 0.52 mmol) in tetrahydrofuran (0.5 mL) and water (0.5 mL) was cooled to 0°C . To this mixture was added lithium hydroxide monohydrate (87 mg, 4 eq, 2.08 mmol), and it was allowed to slowly warm up to room temperature overnight. Then it was neutralized to pH 7 with hydrogen chloride (1 M in water, 30 mL) and cooled in an ice bath to obtain a yellow precipitate.

EXPERIMENTAL PART

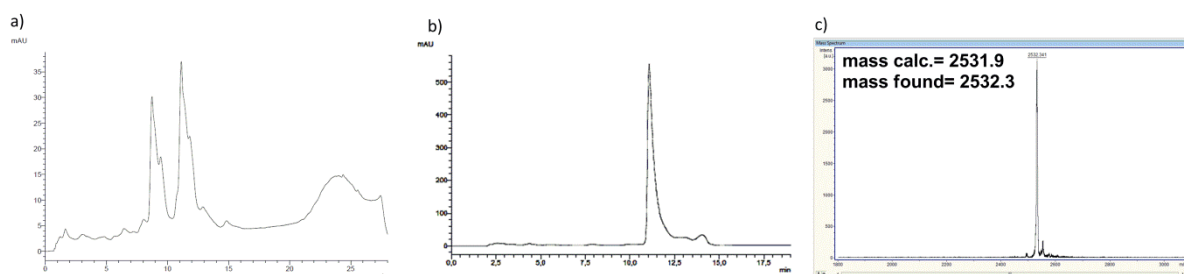
The precipitate was filtered off and the residue was resolved in methanol and then concentrated. The filtrate was extracted with methylene chloride (2 x 100 mL). The combined organic extracts were dried over Na₂SO₄, filtered, and then concentrated. The residue from filtration and extraction was purified by preparative HPLC (gradient of H₂O and MeOH) to give **153** as a yellow powder (135 mg, 85 % yield). Melting temperature: 208 °C (decomposition). ¹H NMR (500 MHz, CDCl₃) δ ppm: 10.87 (s, 1H), 7.62 (d, ³J = 7.5 Hz, 1H), 7.55 (d, ³J = 6.8 Hz, 2H), 7.51 (t, ³J = 7.4 Hz, 2H), 7.49 (d, ³J = 7.7 Hz, 1H), 7.46 (t, ³J = 6.8 Hz, 1H), 7.35 (d, ³J = 7.4 Hz, 2H), 7.28 (t, ³J = 7.6 Hz, 1H), 7.02 (dt, ³J = 7.1 Hz and ⁴J = 1.0 Hz, 1H), 7.06 (t, ³J = 6.3 Hz, 1H), 6.31 (d, ³J = 7.6 Hz, 1H), 5.58 (s, 1H), 3.96 (t, ³J = 6.9 Hz, 1H), 3.03 (dd, ²J = 10.5 Hz and ³J = 2.8 Hz, 1H), 2.75 (dd, ²J = 10.5 Hz and ³J = 4.1 Hz, 1H), 1.89 (br s, 1H). ¹³C NMR (600 MHz, CDCl₃) δ ppm: 175.28, 139.46, 137.73, 136.58, 135.19, 134.50, 127.84, 127.77, 127.60, 126.63, 126.34, 126.10, 125.98, 120.47, 117.86, 117.19, 109.66, 59.20, 56.42, 28.42. MS (ESI): calculated for C₂₄H₂₀N₂O₂ 368.43, found: 369.20. ([M+H]⁺). Purity (HPLC): 99 %.

APPENDIX

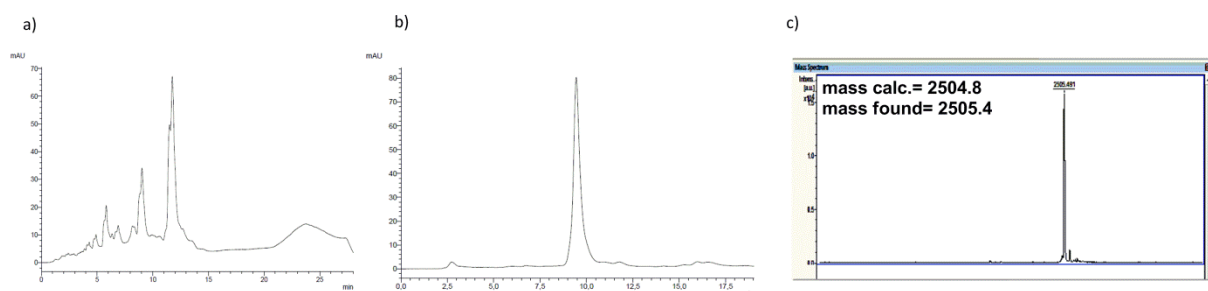
A) HPLC traces and MALDI MS spectra of hexT- β -carboline conjugates 142A-142D]



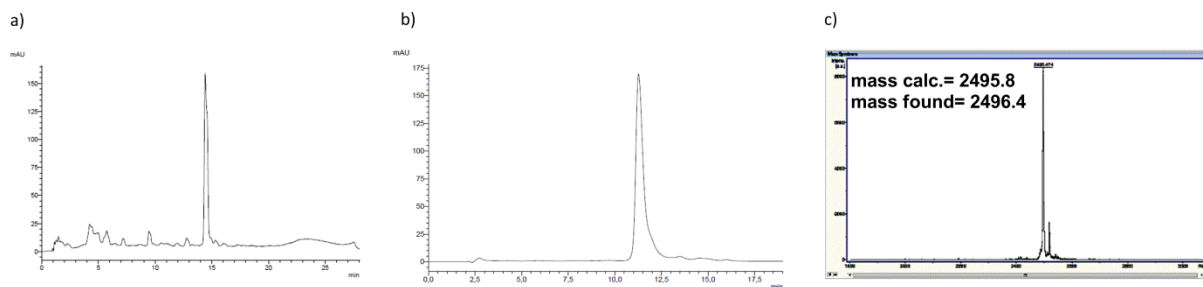
Chromatogram 1: HPLC trace of the crude hexT-conjugate **142A** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142A**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142A**.



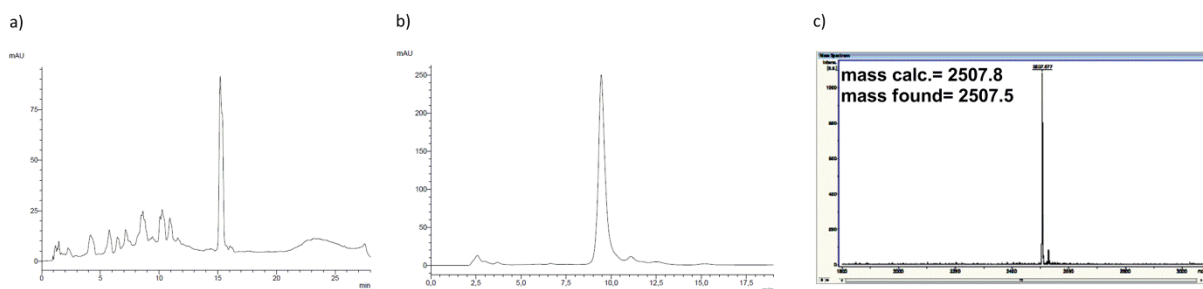
Chromatogram 2: HPLC trace of the crude hexT-conjugate **142C** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142C**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142C**.



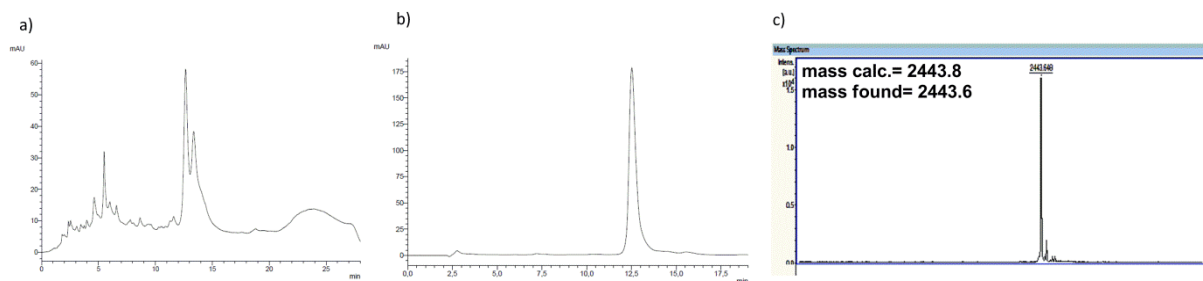
Chromatogram 3: HPLC trace of the crude hexT-conjugate **142L** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142L**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142L**.



Chromatogram 4: HPLC trace of the crude hexT-conjugate **142M** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142M**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142M**.

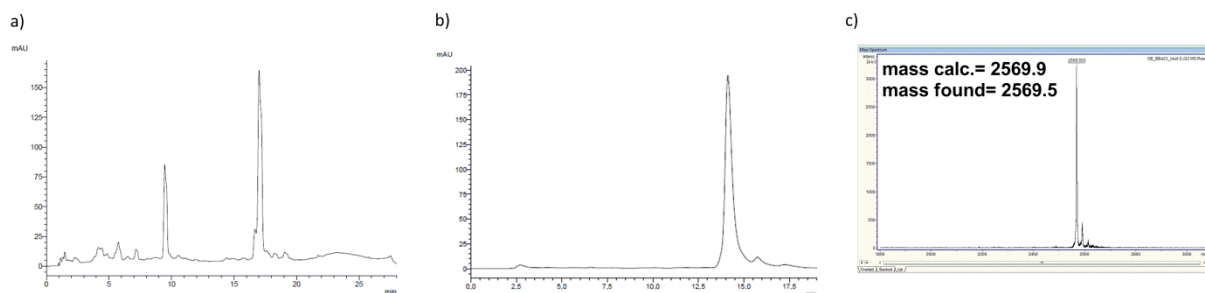


Chromatogram 5: HPLC trace of the crude hexT-conjugate **142S** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142S**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142S**.

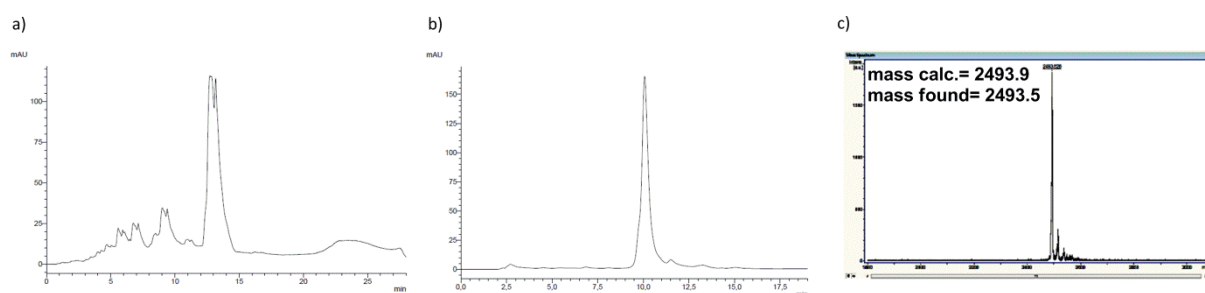


Chromatogram 6: HPLC trace of the crude hexT-conjugate **142R** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142R**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142R**.

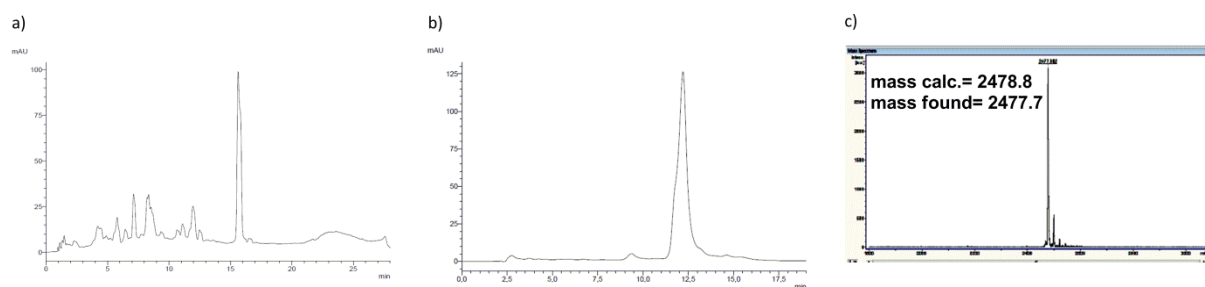
APPENDIX



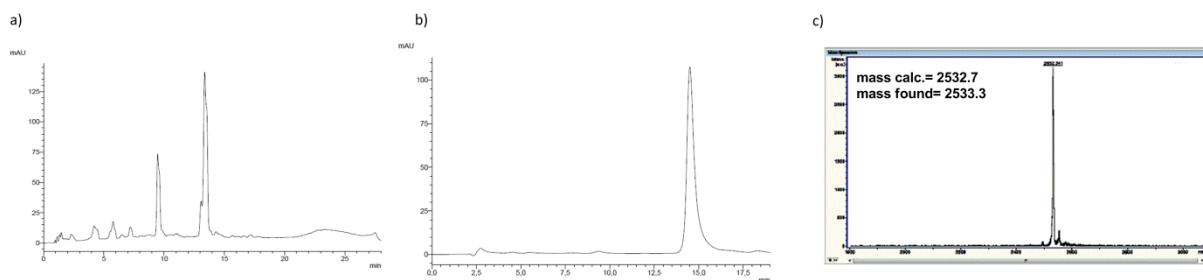
Chromatogram 7: HPLC trace of the crude hexT-conjugate **142V** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142V**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142V**.



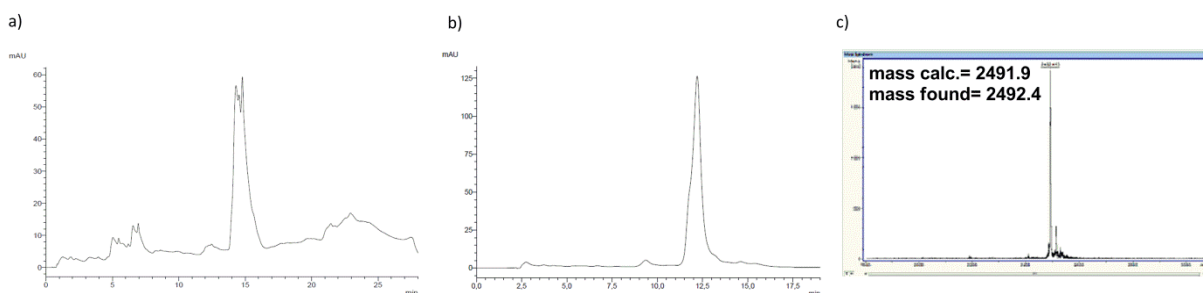
Chromatogram 8: HPLC trace of the crude hexT-conjugate **142Y** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142Y**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142Y**.



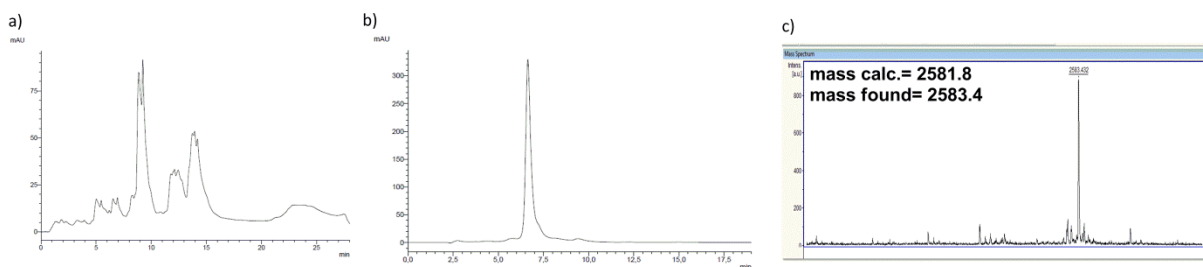
Chromatogram 9: HPLC trace of the crude hexT-conjugate **142Z** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142Z**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142Z**.



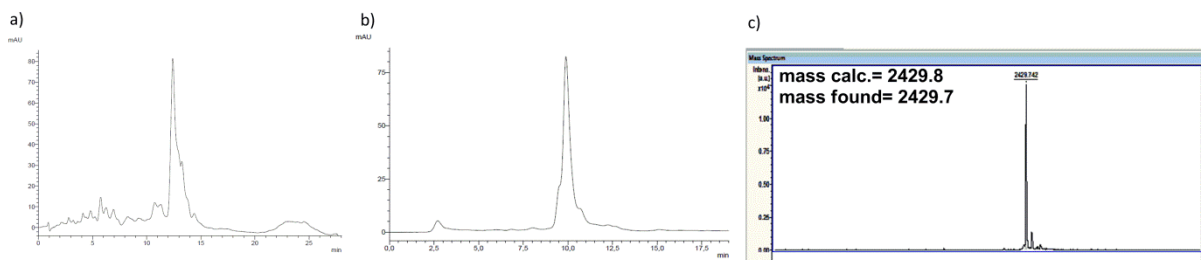
Chromatogram 10: HPLC trace of the crude hexT-conjugate **142AA** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142AA**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142AA**.



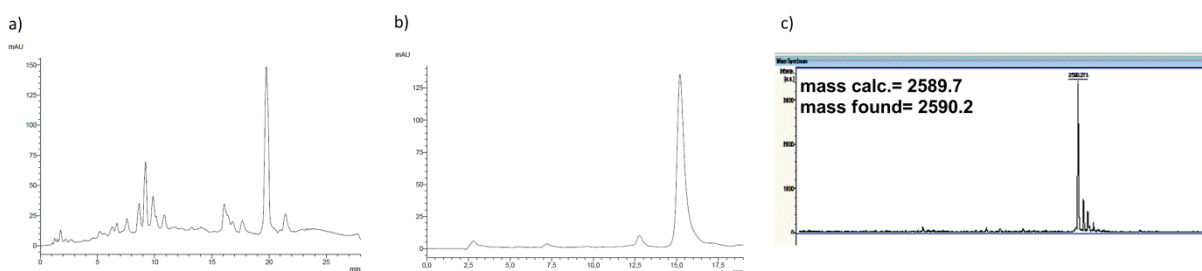
Chromatogram 11: HPLC trace of the crude hexT-conjugate **142AB** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142AB**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142AB**.



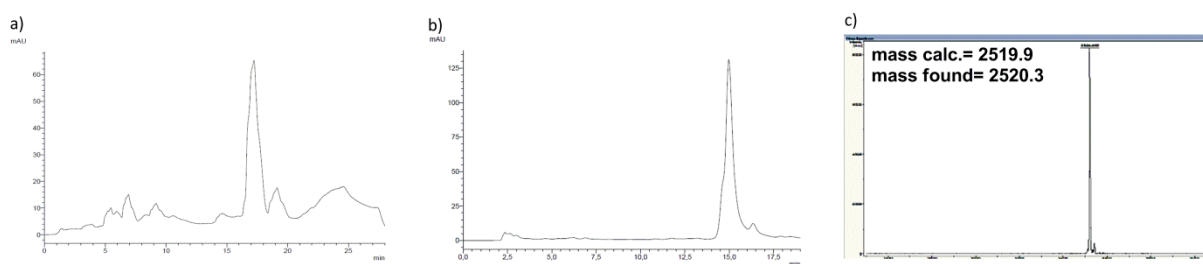
Chromatogram 12: HPLC trace of the crude hexT-conjugate **142AD** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142AD**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142AD**.



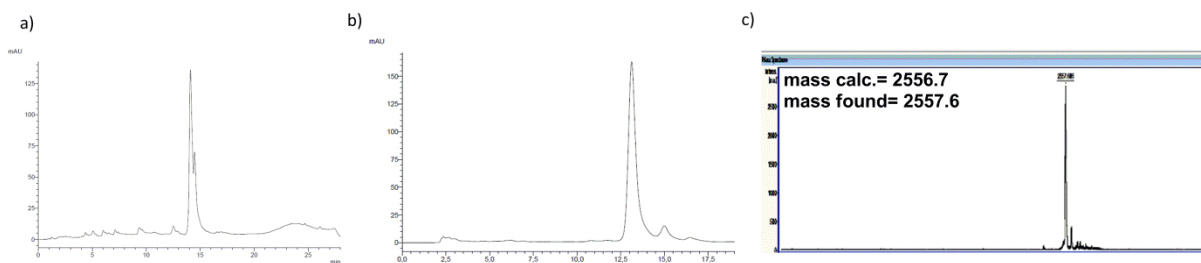
Chromatogram 13: HPLC trace of the crude hexT-conjugate **142AJ** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142AJ**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142AJ**.



Chromatogram 14: HPLC trace of the crude hexT-conjugate **142AK** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142AK**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142AK**.

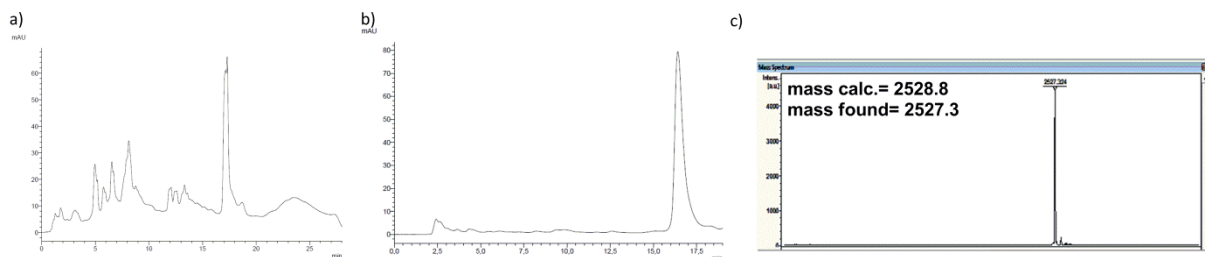


Chromatogram 15: HPLC trace of the crude hexT-conjugate **142AX** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142AX**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142AX**.

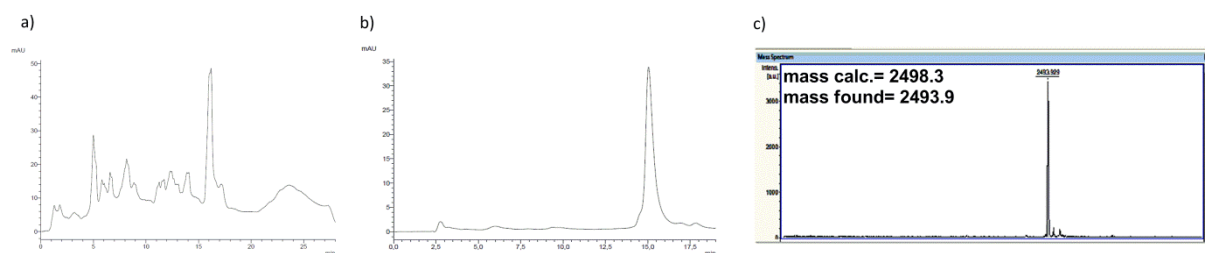


Chromatogram 16: HPLC trace of the crude hexT-conjugate **142AY** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142AY**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142AY**.

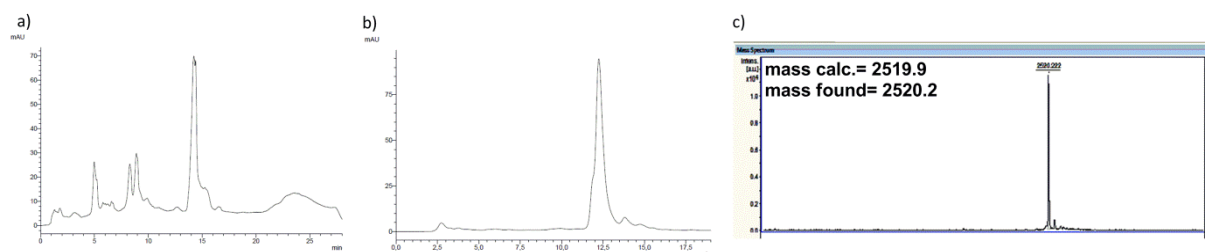
APPENDIX



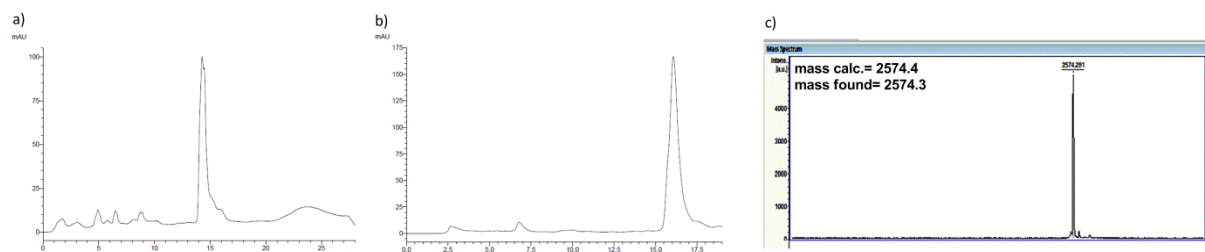
Chromatogram 17: HPLC trace of the crude hexT-conjugate **142CQ** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142CQ**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142CQ**.



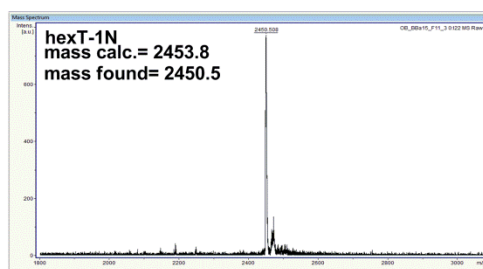
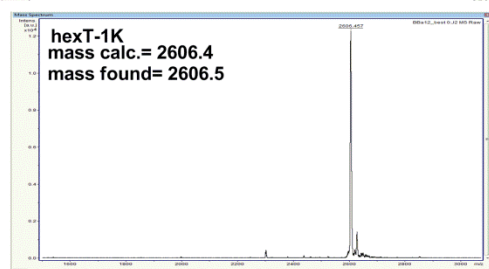
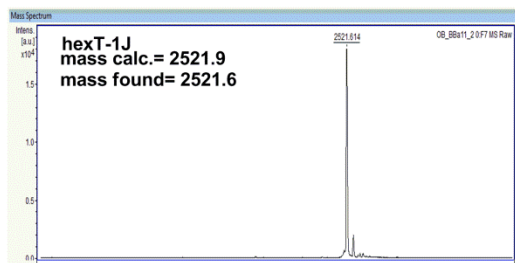
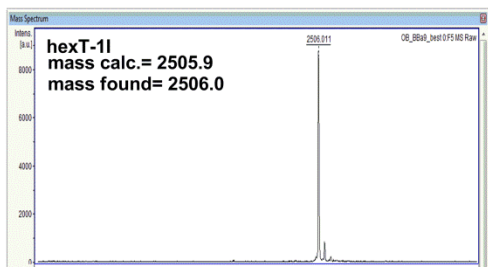
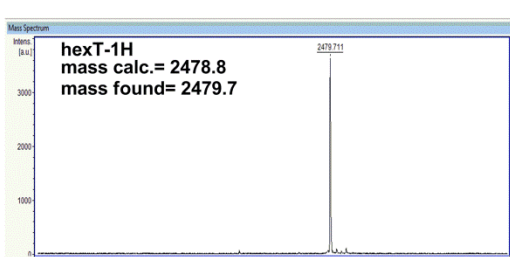
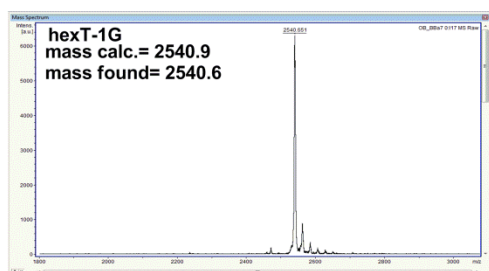
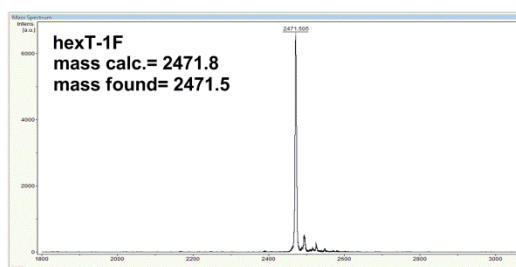
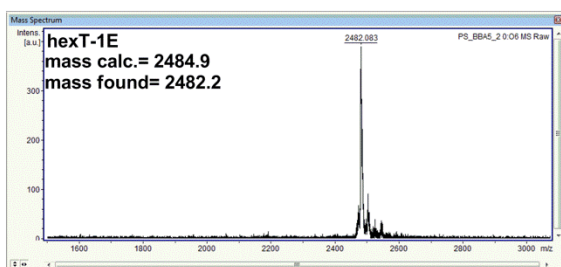
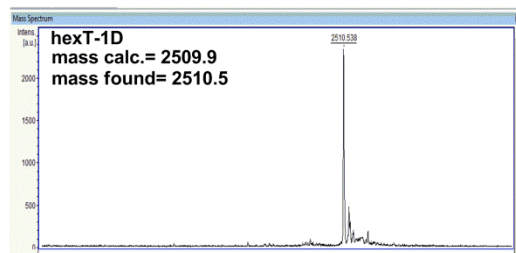
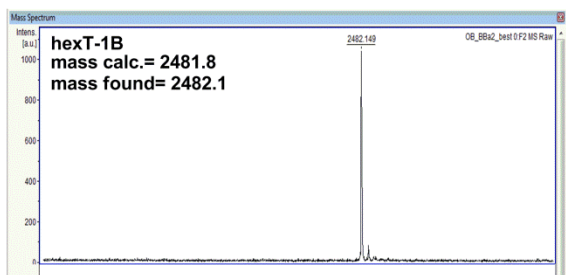
Chromatogram 18: HPLC trace of the crude hexT-conjugate **142CU** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142CU**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142CU**.



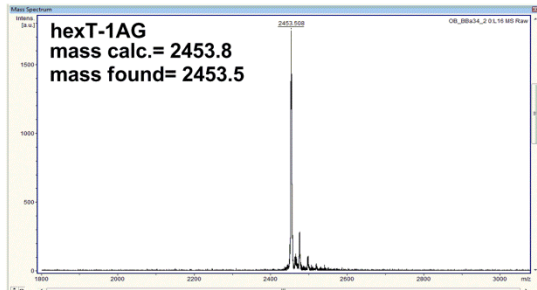
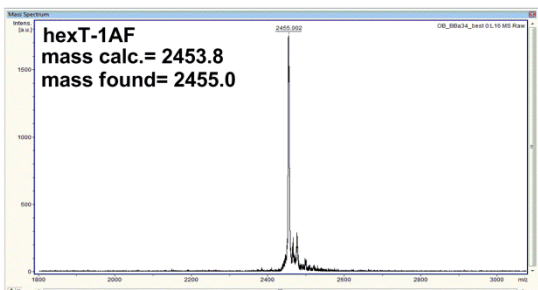
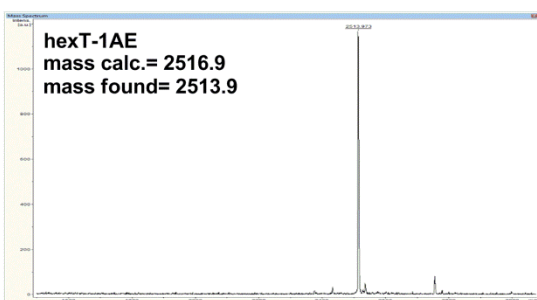
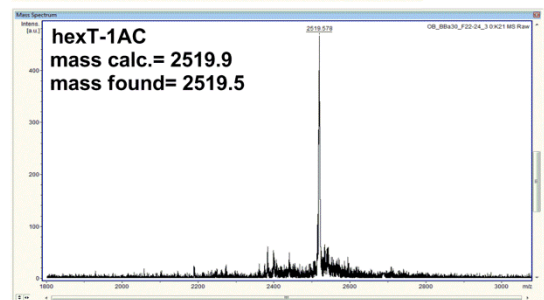
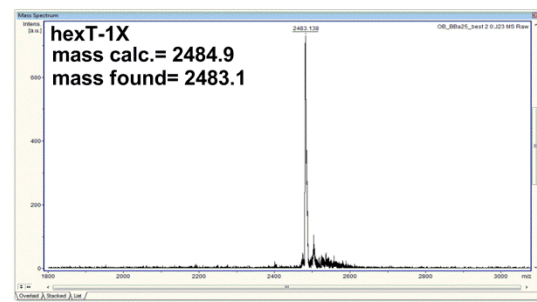
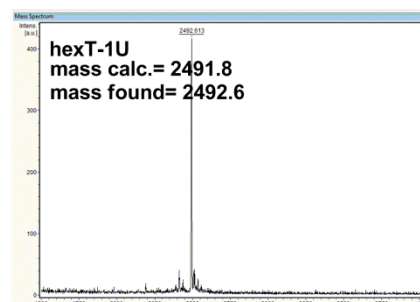
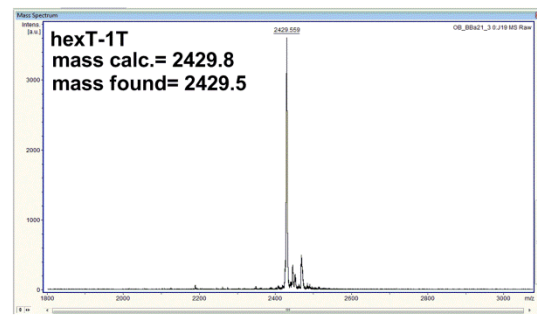
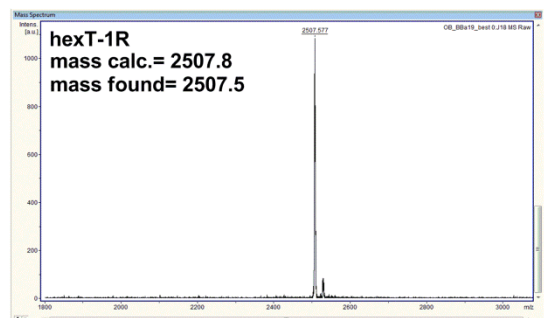
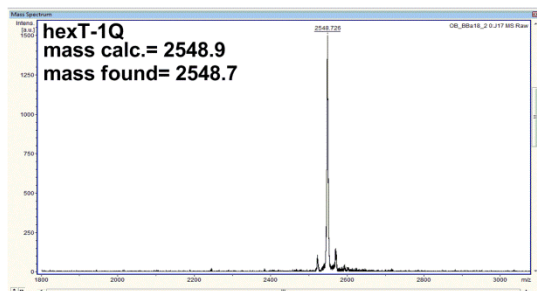
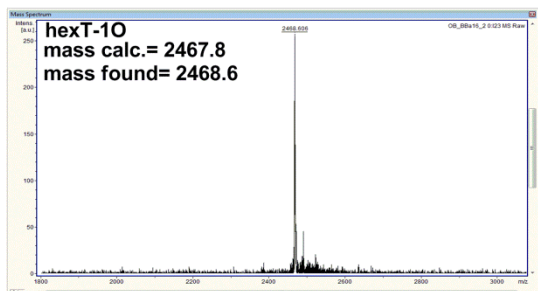
Chromatogram 19: HPLC trace of the crude hexT-conjugate **142CZ** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142CZ**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142CZ**.



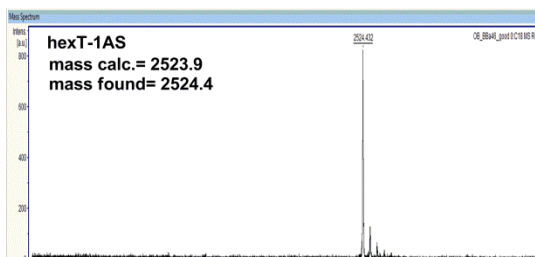
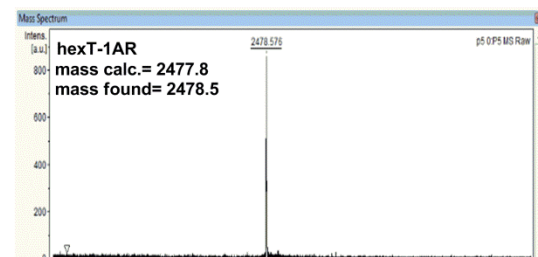
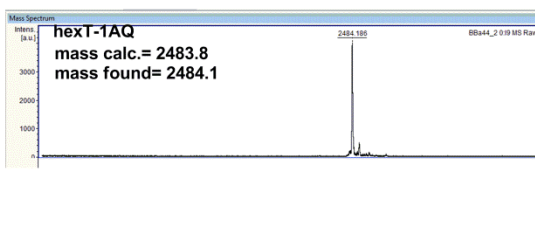
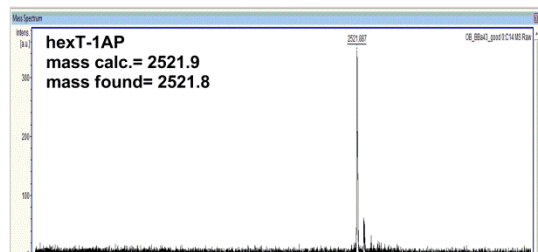
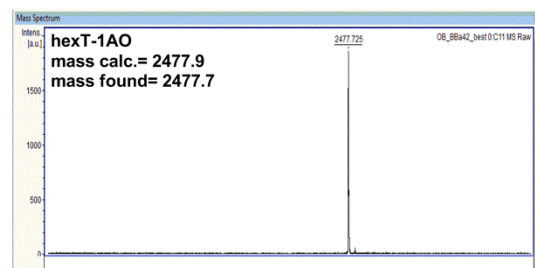
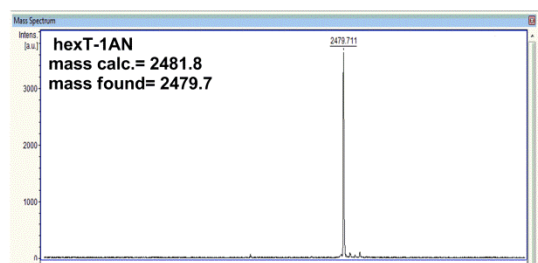
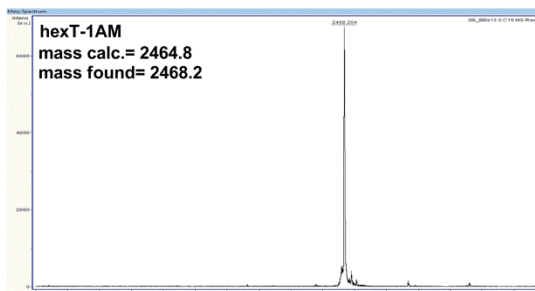
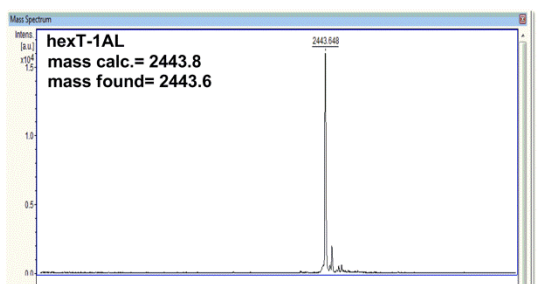
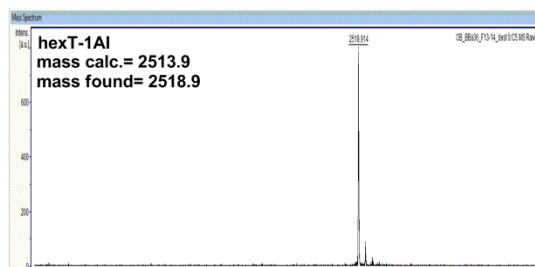
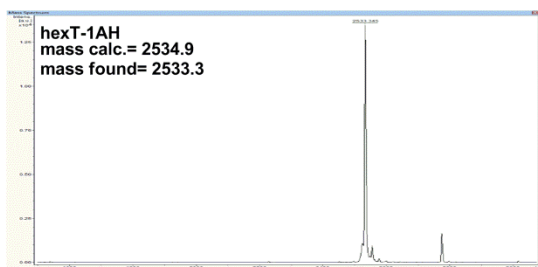
Chromatogram 20: HPLC trace of the crude hexT-conjugate **142DB** (preparative HPLC); b) HPLC trace of the purified hexT-conjugate **142DB**; c) MALDI-MS analysis of the purified purified hexT-conjugate **142DB**.

C) MALDI MS spectra of 94 hexT- β -carboline conjugates

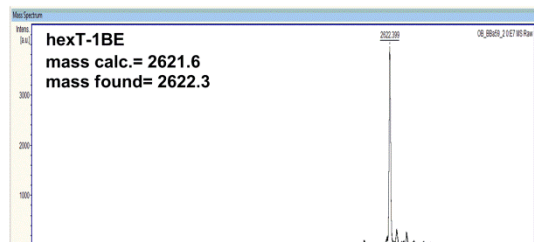
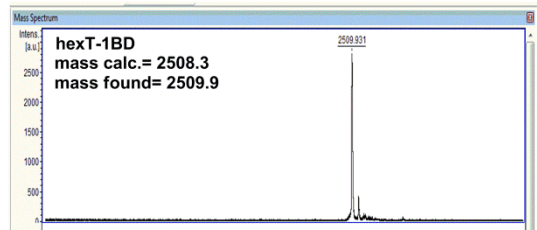
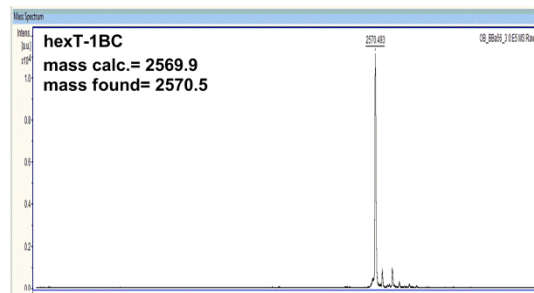
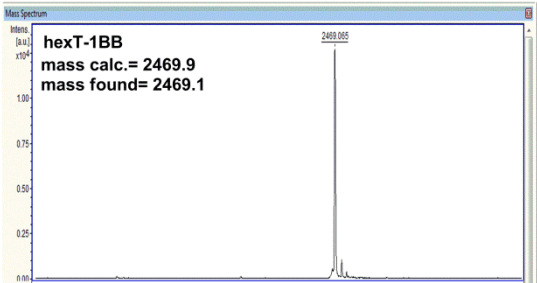
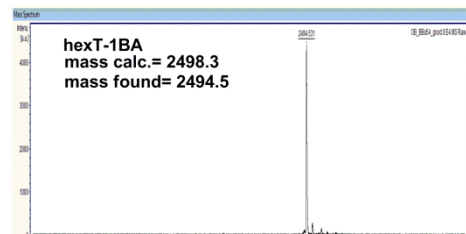
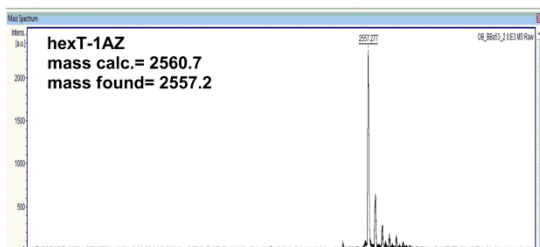
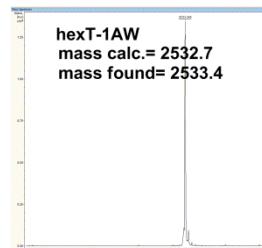
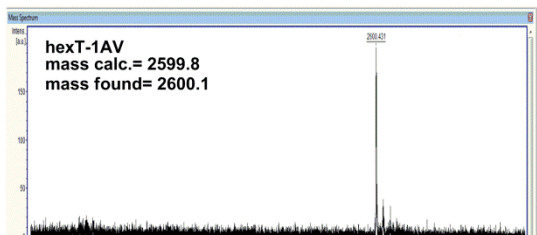
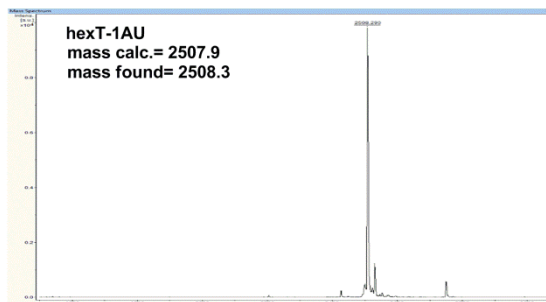
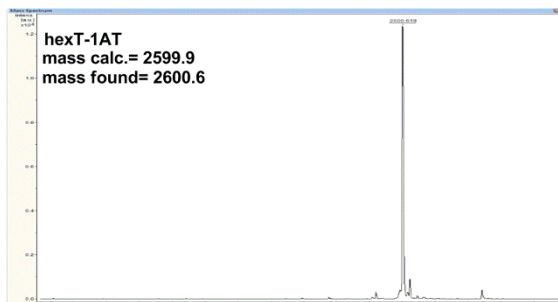
APPENDIX



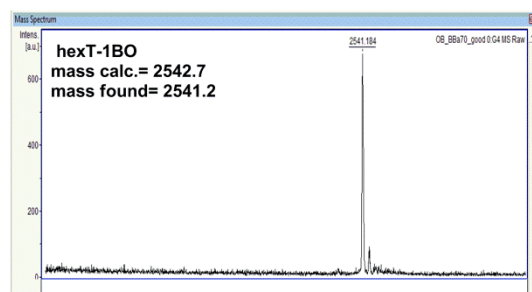
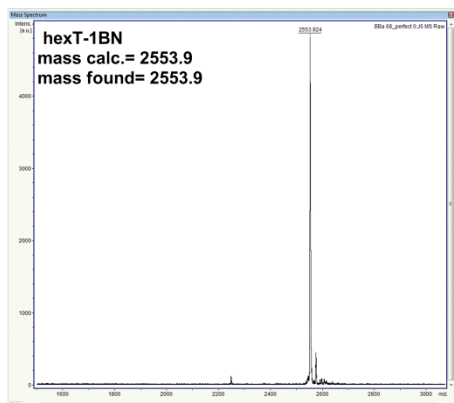
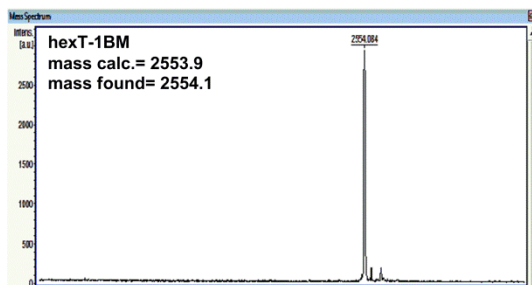
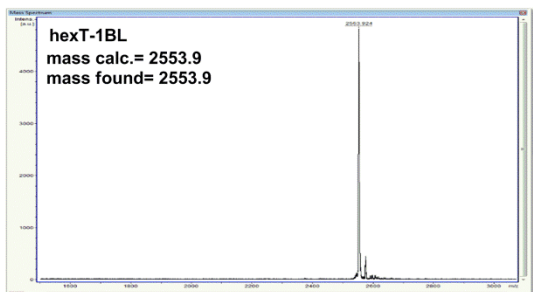
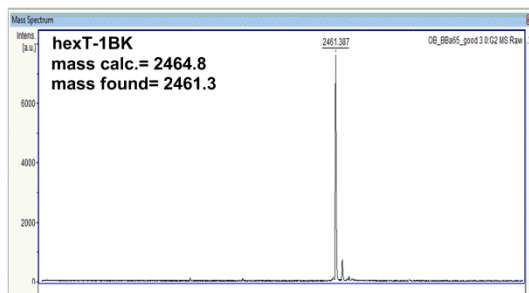
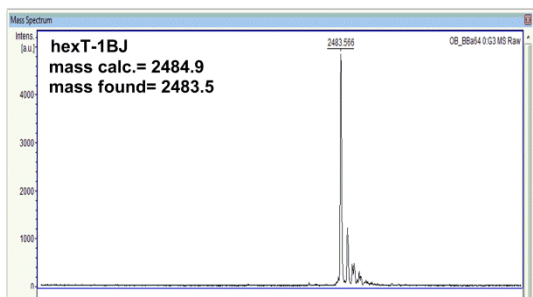
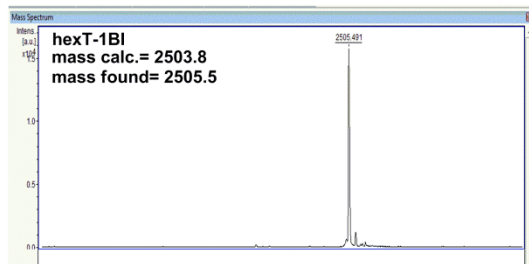
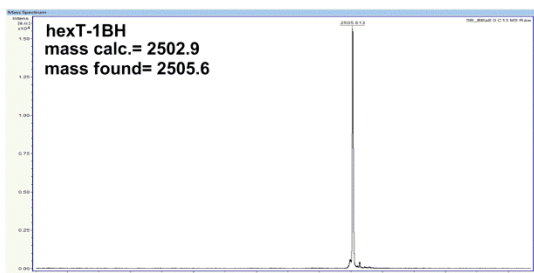
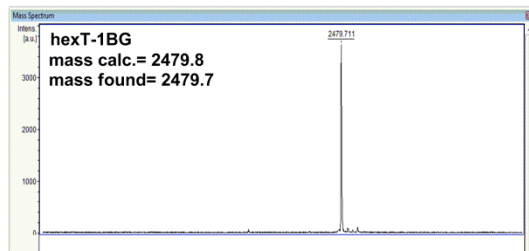
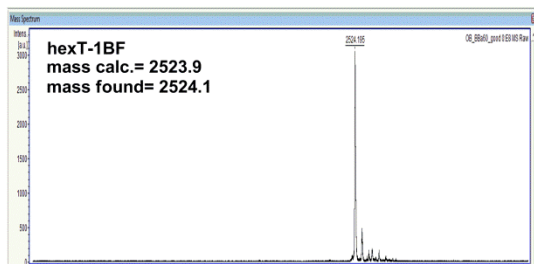
APPENDIX



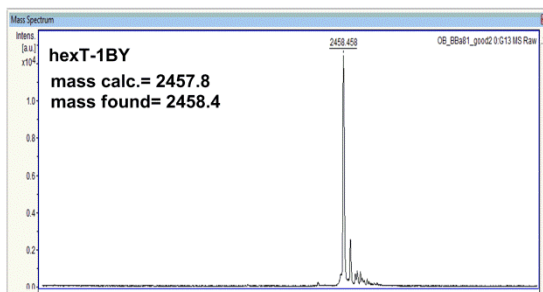
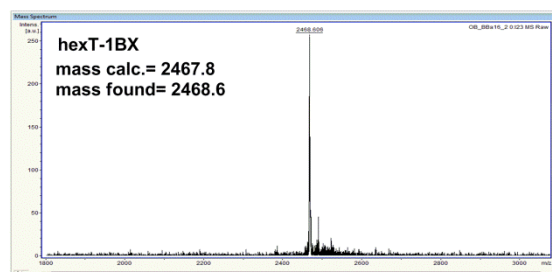
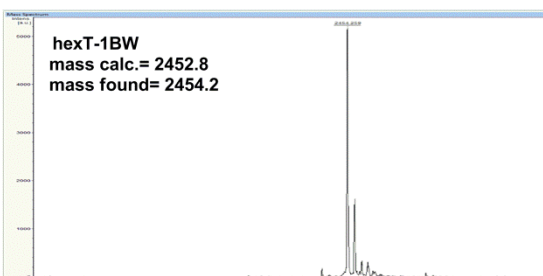
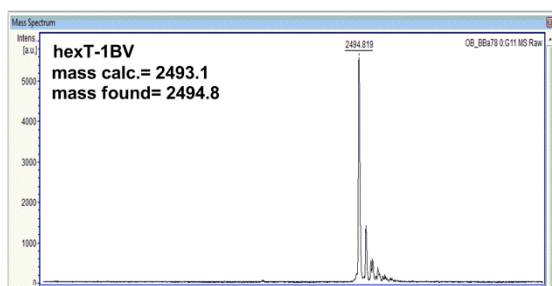
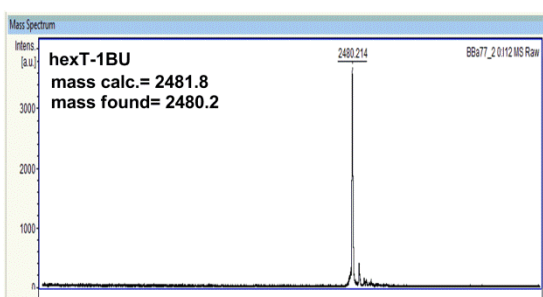
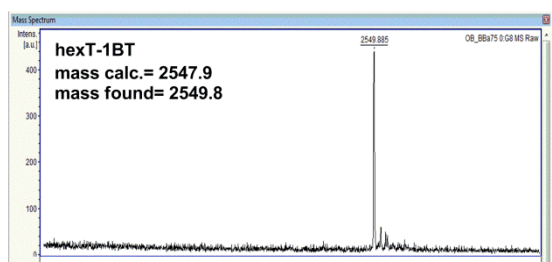
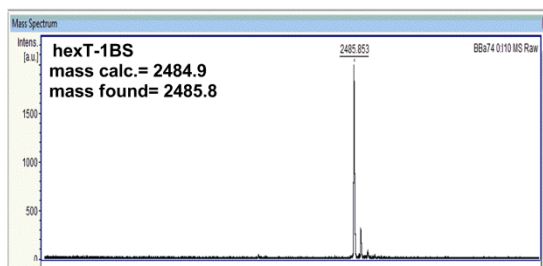
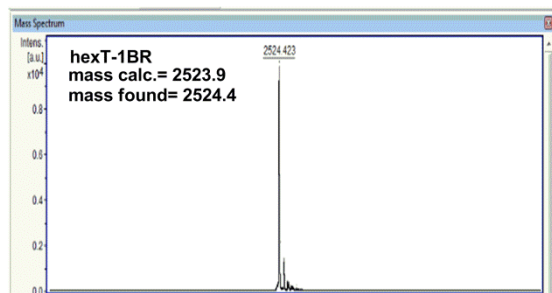
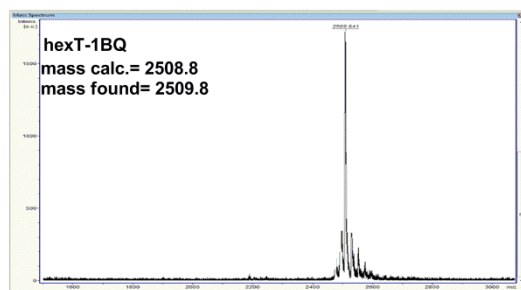
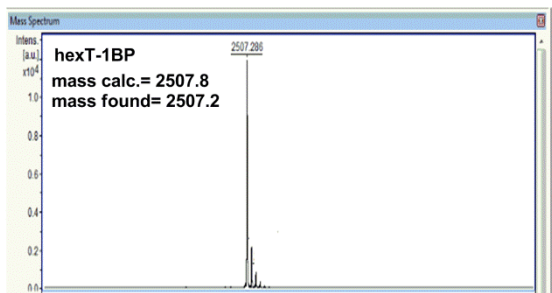
APPENDIX



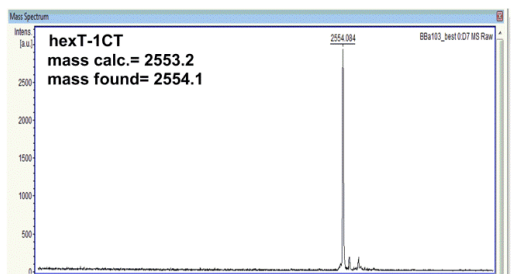
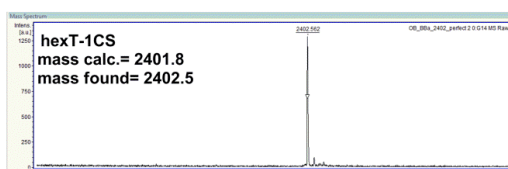
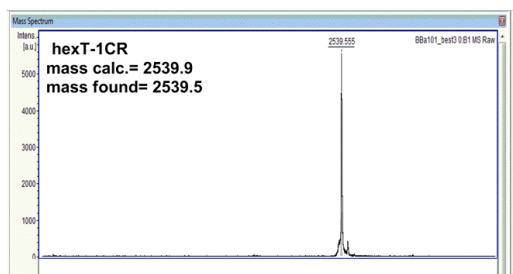
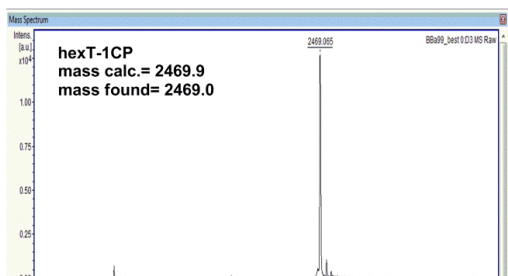
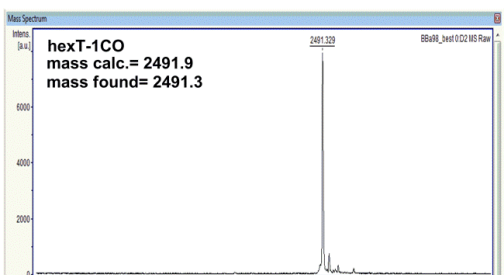
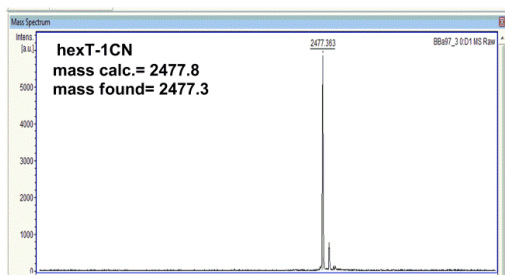
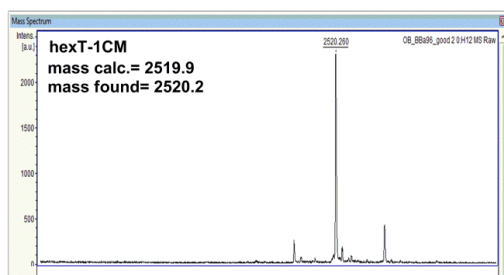
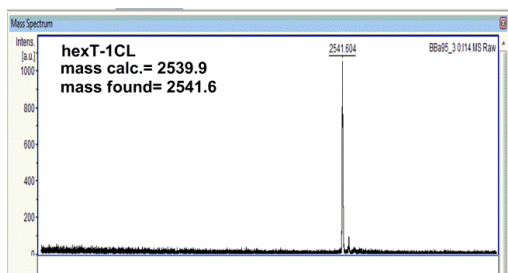
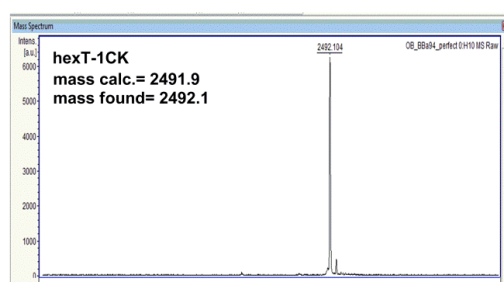
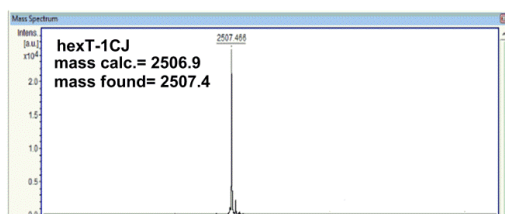
APPENDIX



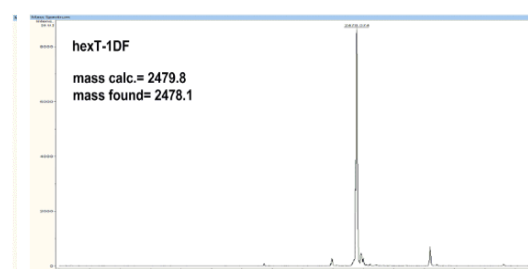
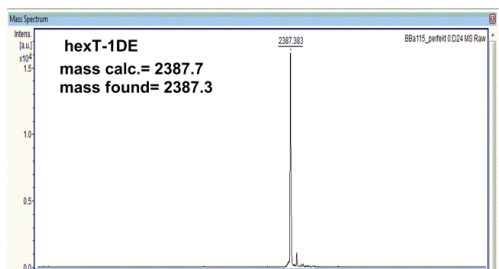
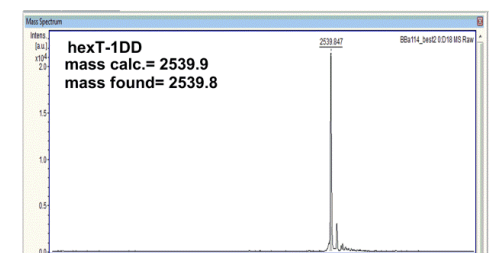
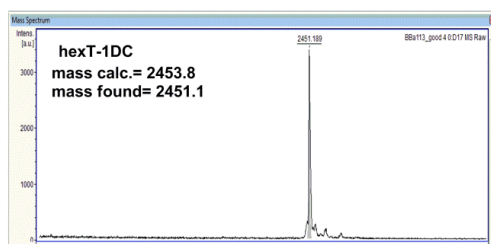
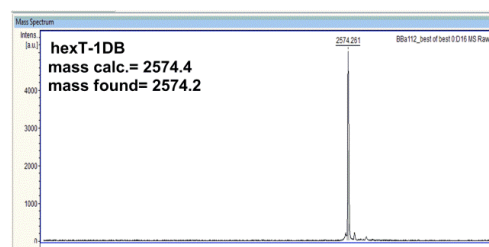
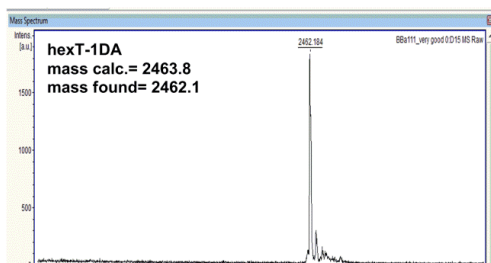
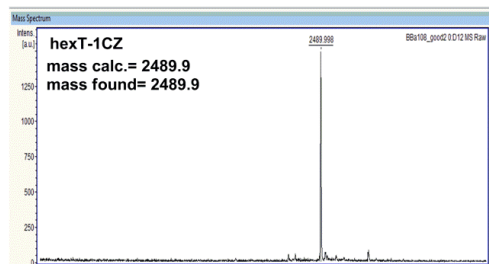
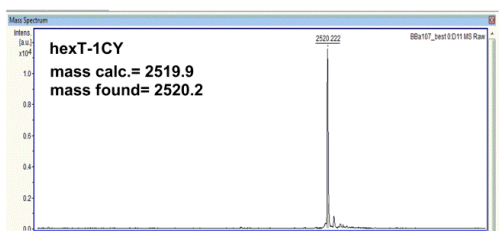
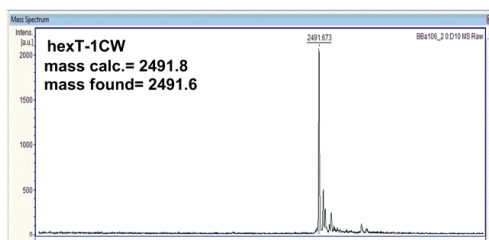
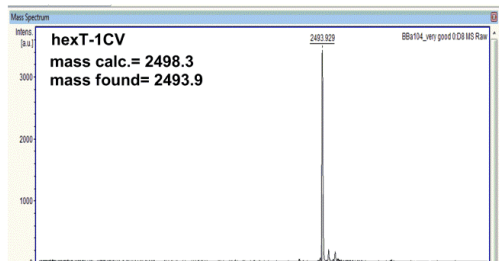
APPENDIX



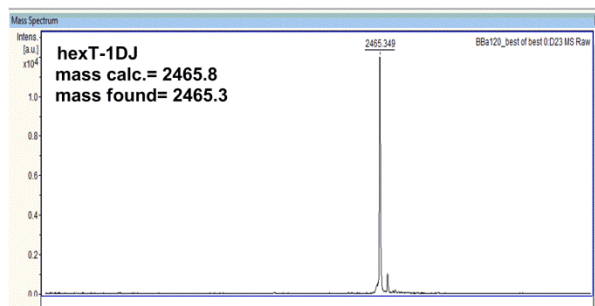
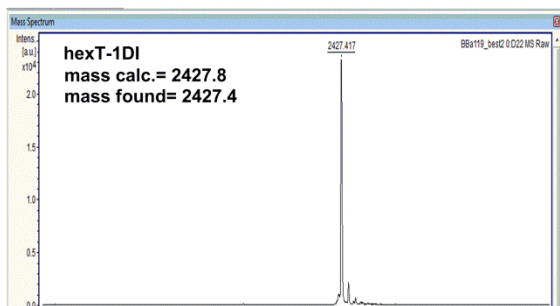
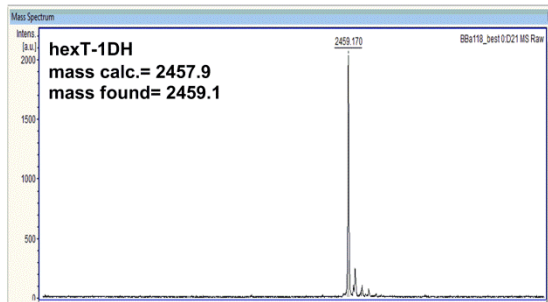
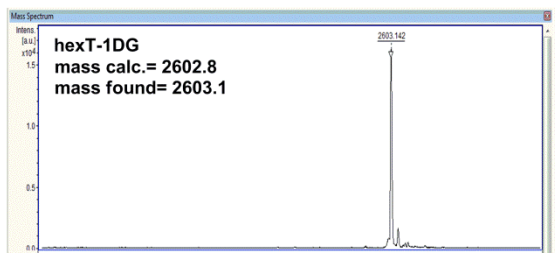
APPENDIX



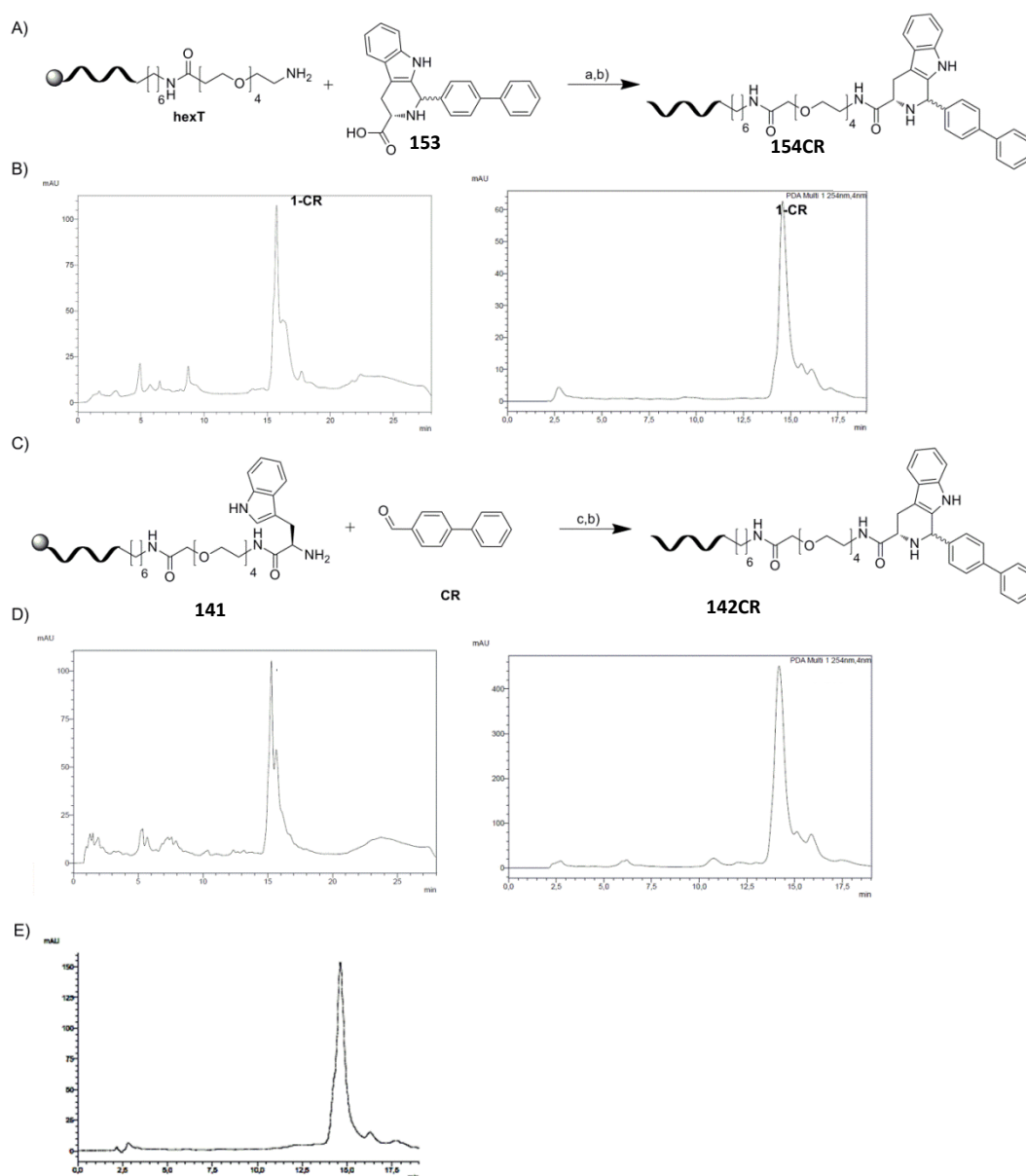
APPENDIX



APPENDIX



D) Comparison of the hexT- β -carboline conjugate **142CR** with a reference molecule



Chromatogram 21: Comparison of the product of the Pictet-Spengler reaction **142CR** with a reference compound **142CR** synthesized from the β -carboline **3** and hexT by amide coupling. A) synthesis scheme for the synthesis of the reference molecule, a) HATU, DIPEA, room temperature, 4 hours; b) AMA, room temperature, 30 min; B) HPLC trace of the crude reference product **154CR** (left hand), and HPLC trace of the purified reference product **154CR**; C) synthesis scheme for the Pictet-Spengler reaction of hexT conjugate **141** with the aldehyde CR furnishing **142CR**, c) 1 % TFA in CH_2Cl_2 , room temperature, 16 hours; D) HPLC trace of the crude product **142CR** (left hand), and HPLC trace of the purified product **142CR**; E) HPLC trace of the co-injected product **154CR** and reference **142CR**.

LIST OF ABBREVIATIONS

LIST OF ABBREVIATIONS

5-HTT	5-Hydroxytryptamine transporter	mAU	Milli absorption units
Ac₂O	Acetic anhydride	m/z	Mass-to-charge ratio
AcOH	Acetic acid	MeCN	Acetonitrile
AMA	Ammonia/methylamine 1:1	MeOH	Methanol
ATP	Adenosine triphosphate	MMP-3	Matrix metalloproteinase 3
AUC	Area under the curve	MS	Mass spectrometry
BOC	<i>tert</i> -butyloxycarbonyl	MW	Molecular weight
BB	Building block	NaCl	Sodium chloride
CPG	Controlle dpore glass	DEAD	Diethyl azodicarboxylate
CA IX	Carbonic Anhydrase IX	NaN₃	Sodium azide
CuI	Copper (I) iodide	NaOAc	Sodium acetate
DCE	Dichloroethane	Na₂SO₄	Sodium sulfate
DCM	Dichloromethane	NBE	New biological entity
ddH₂O	Distilled water	NET	Norepinephrine transporter
DEAE	Diethylaminoethyl	ng	nanogramm
DEL	DNA-Encoded library	nm	nanomole
DIPEA	N,N-Diisopropylrthylamine	nM	nanomolare

LIST OF ABBREVIATIONS

DMF	Dimethylformamide	NME	New molecular entity
DMSO	Dimethylsulfoxide	NMR	Nuclear magnetic resonance
DNA	Deoxyribonucleic acid	PAK4	P21 activated kinase 4
DTS	DNA-templated synthesis	PCR	Polymerase chain reaction
EDC	N-ethyl-N'-(3-dimethylaminopropyl)-carbodiimide	PEG	Polyethylene glycol
EDTA	Ethylenediaminetetraacetic acid	PhMe	Toluol
Eq	Equivalent	PPh₃	Triphenylphosphine
ESI-MS	Electrospray ionization MS	PSR	Pictet-Spengler reaction
Et₂O	Diethylether	R&D	Research & Development
EtOAc	Ethylacetate	RT	Room temperature
EtOH	Ethanol	R_t	Retention time
Fmoc	9-Fluorenylmethoxycarbonyl	TCA	Trichloroacetic acid
HATU	2-(7-Aza-1 <i>H</i> -benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluoro-phosphate	TEA	Triethylamine
HBTU	2-(1 <i>H</i> -Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluoro-phosphate	TEAA	Triethylamine acetate
HOAt	1-hydroxy-7-aza-benzotriazole	TFA	Trifluoroacetic acid

LIST OF ABBREVIATIONS

HOBt	1-hydroxybenzotrizol	THAP	2',4',6'-Trihydroxyacetophenone monohydrate
HOAc	Acetic acid	THBC	Tetrahydro- β -carboline
HPLC	High performance liquid chromatography	THF	Tetrahydrofuran
HTS	High-throughput screening	TiDEC	oligoThymidine Initiated DNA-Encoded Chemistry
IL-2	Interleukin-2	tiDEL	oligoThymidine Initiated DNA-Encoded Library
IU	International unit	TNF-α	Tumor necrosis factor alpha
K_d	Dissociation constant	Tris	Tris(hydroxymethyl)aminomethane
K_i	Inhibitory constant	ml	microliter
LC	Liquid chromatography	μmol	micromolar

LIST OF TABLES

Table 1: Overview of bioactive compounds 4 -31	40
Table 2: Selected molecular scaffolds and corresponding examples of privileged structures. The figure is adapted from Schneider and Baringhaus, 2008. ⁸⁸	47
Table 3: Selected privileged structures and corresponding drug examples. Privileged structures are found in drugs against several classes of protein targets. The table was adapted from Welsch et al., 2010. ⁸⁹	47
Table 4: Optimization of the Heck reaction (preparation of compound 98).	70
Table 5: N-alkylation optimization of the secondary amine 86	71
Table 6: Optimization of the amide synthesis on solid phase.	74
Table 7: Selected privileged structures and examples of chemical building blocks used in the synthesis of the DEL9588	77
Table 8: Optimization of the ring closing metathesis.	89
Table 9: Characteristics of β -carbolines 127-139	96
Table 10: [a] The reactions were carried out on 20 nmol of hexT-tryptophane conjugate and purified by preparative HPLC. [b] Conversions of the tryptamine conjugate 141 to the β -carbolines after reaction using 4-biphenylcarboxaldehyde CR.	106
Table 12: [a] The reactions were carried out on 20 nmol of hexT-tryptophane conjugate and purified by preparative HPLC. [b] Conversions of the tryptamine conjugate to the β -carbolines after reaction using 4-biphenylcarboxaldehyde CR. The reaction was also repeated using cyclohexanone DK.	108
Table 13: [a] The reactions were carried out on 500 pmol of hexT- β -carboline conjugate and analysed by analytical HPLC. [b] The formation of the iminium ion was evaluated by analytical HPLC. We estimated conversion rates based on integration of HPLC peak of the product versus starting material in the HPLC.	113
Table 14: List of chemicals and materials.	127
Table 15: Yield, conversion rates and MALDI MS data of compounds 96A-CZ	155
Table 16: Evaluation of halides A1-A104.	164
Table 17: Conversions of the DNA alkyne conjugate 101 to triazole.	173
Table 18: Total yield, conversion rates, and MALDI MS data of compounds 150A – 150DJ	182

LIST OF TABLES

LIST OF CHROMATOGRAMS

<i>Chromatogram 1: HPLC trace of the crude hexT-conjugate 142A (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142A; c) MALDI-MS analysis of the purified purified hexT-conjugate 142A.</i>	<i>195</i>
<i>Chromatogram 2: HPLC trace of the crude hexT-conjugate 142C (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142C; c) MALDI-MS analysis of the purified purified hexT-conjugate 142C.</i>	<i>195</i>
<i>Chromatogram 3: HPLC trace of the crude hexT-conjugate 142L (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142L; c) MALDI-MS analysis of the purified purified hexT-conjugate 142L.</i>	<i>195</i>
<i>Chromatogram 4: HPLC trace of the crude hexT-conjugate 142M (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142M; c) MALDI-MS analysis of the purified purified hexT-conjugate 142M.</i>	<i>196</i>
<i>Chromatogram 6: HPLC trace of the crude hexT-conjugate 142R (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142R; c) MALDI-MS analysis of the purified purified hexT-conjugate 142R.</i>	<i>196</i>
<i>Chromatogram 5: HPLC trace of the crude hexT-conjugate 142S (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142S; c) MALDI-MS analysis of the purified purified hexT-conjugate 142S.</i>	<i>196</i>
<i>Chromatogram 7: HPLC trace of the crude hexT-conjugate 142V (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142V; c) MALDI-MS analysis of the purified purified hexT-conjugate 142V.</i>	<i>197</i>
<i>Chromatogram 8: HPLC trace of the crude hexT-conjugate 142Y (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142Y; c) MALDI-MS analysis of the purified purified hexT-conjugate 142Y.</i>	<i>197</i>
<i>Chromatogram 9: HPLC trace of the crude hexT-conjugate 142Z (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142Z; c) MALDI-MS analysis of the purified purified hexT-conjugate 142Z.</i>	<i>197</i>
<i>Chromatogram 10: HPLC trace of the crude hexT-conjugate 142AA (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142AA; c) MALDI-MS analysis of the purified purified hexT-conjugate 142AA.</i>	<i>198</i>
<i>Chromatogram 11: HPLC trace of the crude hexT-conjugate 142AB (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142AB; c) MALDI-MS analysis of the purified purified hexT-conjugate 142AB.</i>	<i>198</i>
<i>Chromatogram 12: HPLC trace of the crude hexT-conjugate 142AD (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142AD; c) MALDI-MS analysis of the purified purified hexT-conjugate 142AD.</i>	<i>198</i>
<i>Chromatogram 13: HPLC trace of the crude hexT-conjugate 142AJ (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142AJ; c) MALDI-MS analysis of the purified purified hexT-conjugate 142AJ.</i>	<i>199</i>
<i>Chromatogram 14: HPLC trace of the crude hexT-conjugate 142AK (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142AK; c) MALDI-MS analysis of the purified purified hexT-conjugate 142AK.</i>	<i>199</i>
<i>Chromatogram 15: HPLC trace of the crude hexT-conjugate 142AX (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142AX; c) MALDI-MS analysis of the purified purified hexT-conjugate 142AX.</i>	<i>199</i>
<i>Chromatogram 16: HPLC trace of the crude hexT-conjugate 142AY (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142AY; c) MALDI-MS analysis of the purified purified hexT-conjugate 142AY.</i>	<i>199</i>
<i>Chromatogram 17: HPLC trace of the crude hexT-conjugate 142CQ (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142CQ; c) MALDI-MS analysis of the purified purified hexT-conjugate 142CQ.</i>	<i>200</i>

LIST OF CHROMATOGRAMS

<i>Chromatogram 18: HPLC trace of the crude hexT-conjugate 142CU (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142CU; c) MALDI-MS analysis of the purified purified hexT-conjugate 142CU.</i>	<i>200</i>
<i>Chromatogram 19: HPLC trace of the crude hexT-conjugate 142CZ (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142CZ; c) MALDI-MS analysis of the purified purified hexT-conjugate 142CZ.</i>	<i>200</i>
<i>Chromatogram 20: HPLC trace of the crude hexT-conjugate 142DB (preparative HLPC); b) HPLC trace of the purified hexT-conjugate 142DB; c) MALDI-MS analysis of the purified purified hexT-conjugate 142DB.</i>	<i>200</i>
<i>D) Comparison of the hexT-β-carboline conjugate 142CR with a reference molecule.....</i>	<i>210</i>
<i>Chromatogram 21: Comparison of the product of the Pictet-Spengler reaction 142CR with a reference compound 142CR synthesized from the β-carboline 3 and hexT by amide coupling. A) synthesis scheme for the synthesis of the reference molecule, a) HATU, DIPEA, room temperature, 4 hours; b) AMA, room temperature, 30 min; B) HPLC trace of the crude reference product 154CR (left hand), and HPLC trace of the purified reference product 154CR; C) synthesis scheme for the Pictet-Spengler reaction of hexT conjugate 141 with the aldehyde CR furnishing 142CR, c) 1 % TFA in CH₂Cl₂, room temperature, 16 hours; D) HPLC trace of the crude product 142CR (left hand), and HPLC trace of the purified product 142CR; E) HPLC trace of the co-injected product 154CR and reference 142CR.</i>	<i>210</i>

REFERENCES

REFERENCES

1. Drews, J., Drug discovery: a historical perspective. *Science* **2000**, *287* (5460), 1960-4.
2. Strausberg, R. L.; Schreiber, S. L., From knowing to controlling: a path from genomics to drugs using small molecule probes. *Science* **2003**, *300* (5617), 294-5.
3. Stockwell, B. R., Exploring biology with small organic molecules. *Nature* **2004**, *432* (7019), 846-54.
4. Macarron, R.; Banks, M. N.; Bojanic, D.; Burns, D. J.; Cirovic, D. A.; Garyantes, T.; Green, D. V.; Hertzberg, R. P.; Janzen, W. P.; Paslay, J. W.; Schopfer, U.; Sittampalam, G. S., Impact of high-throughput screening in biomedical research. *Nat Rev Drug Discov* **2011**, *10* (3), 188-95.
5. Mayr, L. M.; Bojanic, D., Novel trends in high-throughput screening. *Curr Opin Pharmacol* **2009**, *9* (5), 580-8.
6. Mannocci, L.; Leimbacher, M.; Wichert, M.; Scheuermann, J.; Neri, D., 20 years of DNA-encoded chemical libraries. *Chem Commun (Camb)* **2011**, *47* (48), 12747-53.
7. Scheuermann, J.; Dumelin, C. E.; Melkko, S.; Neri, D., DNA-encoded chemical libraries. *J Biotechnol* **2006**, *126* (4), 568-81.
8. Melkko, S.; Dumelin, C. E.; Scheuermann, J.; Neri, D., Lead discovery by DNA-encoded chemical libraries. *Drug Discov Today* **2007**, *12* (11-12), 465-71.
9. Melkko, S.; Zhang, Y.; Dumelin, C. E.; Scheuermann, J.; Neri, D., Isolation of high-affinity trypsin inhibitors from a DNA-encoded chemical library. *Angew Chem Int Ed Engl* **2007**, *46* (25), 4671-4.
10. Warren, J., Drug discovery: lessons from evolution. *Br J Clin Pharmacol* **2011**, *71* (4), 497-503.
11. Swinney, D. C.; Anthony, J., How were new medicines discovered? *Nat Rev Drug Discov* **2011**, *10* (7), 507-19.
12. Posner, B. A., High-throughput screening-driven lead discovery: meeting the challenges of finding new therapeutics. *Curr Opin Drug Discov Devel* **2005**, *8* (4), 487-94.
13. Scannell, J. W.; Blanckley, A.; Boldon, H.; Warrington, B., Diagnosing the decline in pharmaceutical R&D efficiency. *Nat Rev Drug Discov* **2012**, *11* (3), 191-200.
14. Brunschweiler, A.; Hall, J., A decade of the human genome sequence--how does the medicinal chemist benefit? *ChemMedChem* **2012**, *7* (2), 194-203.
15. Strovel, J.; Sittampalam, S.; Coussens, N. P.; Hughes, M.; Inglese, J.; Kurtz, A.; Andalibi, A.; Patton, L.; Austin, C.; Baltezor, M.; Beckloff, M.; Weingarten, M.; Weir, S., Early Drug Discovery and Development Guidelines: For Academic Researchers, Collaborators, and Start-up Companies. In *Assay Guidance Manual*, Sittampalam, G. S.; Coussens, N. P.; Brimacombe, K.; Grossman, A.; Arkin, M.; Auld, D.; Austin, C.; Baell, J.; Bejcek, B.; Chung, T. D. Y.; Dahlin, J. L.; Devanaryan, V.; Foley, T. L.; Glicksman, M.; Hall, M. D.; Hass, J. V.; Inglese, J.; Iversen, P. W.; Kahl, S. D.; Kales, S. C.; Lal-Nag, M.;

REFERENCES

- Li, Z.; McGee, J.; McManus, O.; Riss, T.; Trask, O. J., Jr.; Weidner, J. R.; Xia, M.; Xu, X., Eds. Bethesda (MD), 2004.
16. Paul, S. M.; Mytelka, D. S.; Dunwiddie, C. T.; Persinger, C. C.; Munos, B. H.; Lindborg, S. R.; Schacht, A. L., How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Drug Discov* **2010**, *9* (3), 203-14.
17. Cohen, F. J., Macro trends in pharmaceutical innovation. *Discov Med* **2005**, *5* (26), 153-8.
18. Munos, B., Lessons from 60 years of pharmaceutical innovation. *Nat Rev Drug Discov* **2009**, *8* (12), 959-68.
19. Walsh, G., Biopharmaceutical benchmarks 2006. *Nat Biotechnol* **2006**, *24* (7), 769-76.
20. Kehoe, J. W.; Kay, B. K., Filamentous phage display in the new millennium. *Chem Rev* **2005**, *105* (11), 4056-72.
21. McCafferty, J.; Griffiths, A. D.; Winter, G.; Chiswell, D. J., Phage antibodies: filamentous phage displaying antibody variable domains. *Nature* **1990**, *348* (6301), 552-4.
22. Smith, G. P., Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **1985**, *228* (4705), 1315-7.
23. Smith, G. P.; Petrenko, V. A., Phage Display. *Chem Rev* **1997**, *97* (2), 391-410.
24. Winter, G.; Griffiths, A. D.; Hawkins, R. E.; Hoogenboom, H. R., Making antibodies by phage display technology. *Annu Rev Immunol* **1994**, *12*, 433-55.
25. Kempeni, J., Preliminary results of early clinical trials with the fully human anti-TNFalpha monoclonal antibody D2E7. *Ann Rheum Dis* **1999**, *58 Suppl 1*, I70-2.
26. Petrenko, V., Evolution of phage display: from bioactive peptides to bioselective nanomaterials. *Expert Opin Drug Deliv* **2008**, *5* (8), 825-36.
27. Kramer, R.; Cohen, D., Functional genomics to new drug targets. *Nat Rev Drug Discov* **2004**, *3* (11), 965-72.
28. Bajorath, J., Integration of virtual and high-throughput screening. *Nat Rev Drug Discov* **2002**, *1* (11), 882-94.
29. Erlanson, D. A.; Wells, J. A.; Braisted, A. C., Tethering: fragment-based drug discovery. *Annu Rev Biophys Biomol Struct* **2004**, *33*, 199-223.
30. Schneider, G., Virtual screening: an endless staircase? *Nat Rev Drug Discov* **2010**, *9* (4), 273-6.
31. Leach, A. R.; Hann, M. M.; Burrows, J. N.; Griffen, E. J., Fragment screening: an introduction. *Mol Biosyst* **2006**, *2* (9), 430-46.

REFERENCES

32. Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R., Fragment-based lead discovery. *Nat Rev Drug Discov* **2004**, *3* (8), 660-72.
33. Clackson, T.; Hoogenboom, H. R.; Griffiths, A. D.; Winter, G., Making antibody fragments using phage display libraries. *Nature* **1991**, *352* (6336), 624-8.
34. Hanes, J.; Pluckthun, A., In vitro selection and evolution of functional proteins by using ribosome display. *Proc Natl Acad Sci U S A* **1997**, *94* (10), 4937-42.
35. Mattheakis, L. C.; Bhatt, R. R.; Dower, W. J., An in vitro polysome display system for identifying ligands from very large peptide libraries. *Proc Natl Acad Sci U S A* **1994**, *91* (19), 9022-6.
36. Brenner, S.; Lerner, R. A., Encoded combinatorial chemistry. *Proc Natl Acad Sci U S A* **1992**, *89* (12), 5381-3.
37. Buller, F.; Zhang, Y.; Scheuermann, J.; Schafer, J.; Buhlmann, P.; Neri, D., Discovery of TNF inhibitors from a DNA-encoded chemical library based on diels-alder cycloaddition. *Chem Biol* **2009**, *16* (10), 1075-86.
38. Leimbacher, M.; Zhang, Y.; Mannocci, L.; Stravs, M.; Geppert, T.; Scheuermann, J.; Schneider, G.; Neri, D., Discovery of small-molecule interleukin-2 inhibitors from a DNA-encoded chemical library. *Chemistry* **2012**, *18* (25), 7729-37.
39. Needels, M. C.; Jones, D. G.; Tate, E. H.; Heinkel, G. L.; Kochersperger, L. M.; Dower, W. J.; Barrett, R. W.; Gallop, M. A., Generation and screening of an oligonucleotide-encoded synthetic peptide library. *Proc Natl Acad Sci U S A* **1993**, *90* (22), 10700-4.
40. Halpin, D. R.; Lee, J. A.; Wrenn, S. J.; Harbury, P. B., DNA display III. Solid-phase organic synthesis on unprotected DNA. *PLoS Biol* **2004**, *2* (7), E175.
41. Brudno, Y.; Birnbaum, M. E.; Kleiner, R. E.; Liu, D. R., An in vitro translation, selection and amplification system for peptide nucleic acids. *Nat Chem Biol* **2010**, *6* (2), 148-55.
42. Zambaldo, C.; Barluenga, S.; Winssinger, N., PNA-encoded chemical libraries. *Curr Opin Chem Biol* **2015**, *26*, 8-15.
43. Salamon, H.; Klika Skopic, M.; Jung, K.; Bugain, O.; Brunschweiler, A., Chemical Biology Probes from Advanced DNA-encoded Libraries. *ACS Chem Biol* **2016**, *11* (2), 296-307.
44. Buller, F.; Mannocci, L.; Zhang, Y.; Dumelin, C. E.; Scheuermann, J.; Neri, D., Design and synthesis of a novel DNA-encoded chemical library using Diels-Alder cycloadditions. *Bioorg Med Chem Lett* **2008**, *18* (22), 5926-31.
45. Clark, M. A.; Acharya, R. A.; Arico-Muendel, C. C.; Belyanskaya, S. L.; Benjamin, D. R.; Carlson, N. R.; Centrella, P. A.; Chiu, C. H.; Creaser, S. P.; Cuzzo, J. W.; Davie, C. P.; Ding, Y.; Franklin, G. J.; Franzen, K. D.; Gefter, M. L.; Hale, S. P.; Hansen, N. J.; Israel, D. I.; Jiang, J.; Kavarana, M. J.; Kelley, M. S.; Kollmann, C. S.; Li, F.; Lind, K.; Mataruse, S.; Medeiros, P. F.; Messer, J. A.; Myers, P.; O'Keefe, H.; Oliff, M. C.; Rise, C. E.; Satz, A. L.; Skinner, S. R.; Svendsen, J. L.; Tang, L.; van Vloten, K.; Wagner,

REFERENCES

- R. W.; Yao, G.; Zhao, B.; Morgan, B. A., Design, synthesis and selection of DNA-encoded small-molecule libraries. *Nat Chem Biol* **2009**, *5* (9), 647-54.
46. Gartner, Z. J.; Tse, B. N.; Grubina, R.; Doyon, J. B.; Snyder, T. M.; Liu, D. R., DNA-templated organic synthesis and selection of a library of macrocycles. *Science* **2004**, *305* (5690), 1601-5.
47. Mannocci, L.; Zhang, Y.; Scheuermann, J.; Leimbacher, M.; De Bellis, G.; Rizzi, E.; Dumelin, C.; Melkko, S.; Neri, D., High-throughput sequencing allows the identification of binding molecules isolated from DNA-encoded chemical libraries. *Proc Natl Acad Sci U S A* **2008**, *105* (46), 17670-5.
48. Tse, B. N.; Snyder, T. M.; Shen, Y.; Liu, D. R., Translation of DNA into a library of 13,000 synthetic small-molecule macrocycles suitable for in vitro selection. *J Am Chem Soc* **2008**, *130* (46), 15611-26.
49. Melkko, S.; Scheuermann, J.; Dumelin, C. E.; Neri, D., Encoded self-assembling chemical libraries. *Nat Biotechnol* **2004**, *22* (5), 568-74.
50. Hansen, M. H.; Blakskjaer, P.; Petersen, L. K.; Hansen, T. H.; Hojfeldt, J. W.; Gothelf, K. V.; Hansen, N. J., A yoctoliter-scale DNA reactor for small-molecule evolution. *J Am Chem Soc* **2009**, *131* (3), 1322-7.
51. Melkko, S.; Mannocci, L.; Dumelin, C. E.; Villa, A.; Sommovilla, R.; Zhang, Y.; Grutter, M. G.; Keller, N.; Jermutus, L.; Jackson, R. H.; Scheuermann, J.; Neri, D., Isolation of a small-molecule inhibitor of the antiapoptotic protein Bcl-xL from a DNA-encoded chemical library. *ChemMedChem* **2010**, *5* (4), 584-90.
52. Georghiou, G.; Kleiner, R. E.; Pulkoski-Gross, M.; Liu, D. R.; Seeliger, M. A., Highly specific, bisubstrate-competitive Src inhibitors from DNA-templated macrocycles. *Nat Chem Biol* **2012**, *8* (4), 366-74.
53. Halpin, D. R.; Harbury, P. B., DNA display II. Genetic manipulation of combinatorial chemistry libraries for small-molecule evolution. *PLoS Biol* **2004**, *2* (7), E174.
54. Wrenn, S. J.; Weisinger, R. M.; Halpin, D. R.; Harbury, P. B., Synthetic ligands discovered by in vitro selection. *J Am Chem Soc* **2007**, *129* (43), 13137-43.
55. Halpin, D. R.; Harbury, P. B., DNA display I. Sequence-encoded routing of DNA populations. *PLoS Biol* **2004**, *2* (7), E173.
56. Dumelin, C. E.; Scheuermann, J.; Melkko, S.; Neri, D., Selection of streptavidin binders from a DNA-encoded chemical library. *Bioconjug Chem* **2006**, *17* (2), 366-70.
57. Dumelin, C. E.; Trussel, S.; Buller, F.; Trachsel, E.; Bootz, F.; Zhang, Y.; Mannocci, L.; Beck, S. C.; Drumea-Mirancea, M.; Seeliger, M. W.; Baltes, C.; Muggler, T.; Kranz, F.; Rudin, M.; Melkko, S.; Scheuermann, J.; Neri, D., A portable albumin binder from a DNA-encoded chemical library. *Angew Chem Int Ed Engl* **2008**, *47* (17), 3196-201.

REFERENCES

58. Melkko, S.; Dumelin, C. E.; Scheuermann, J.; Neri, D., On the magnitude of the chelate effect for the recognition of proteins by pharmacophores scaffolded by self-assembling oligonucleotides. *Chem Biol* **2006**, *13* (2), 225-31.
59. Scheuermann, J.; Dumelin, C. E.; Melkko, S.; Zhang, Y.; Mannocci, L.; Jaggi, M.; Sobek, J.; Neri, D., DNA-encoded chemical libraries for the discovery of MMP-3 inhibitors. *Bioconjug Chem* **2008**, *19* (3), 778-85.
60. Zambaldo, C.; Daguer, J. P.; Saabach, J.; Barluenga, S.; Winssinger, N., Screening for covalent inhibitors using DNA-display of small molecule libraries functionalized with cysteine reactive moieties. *MedChemComm* **2016**, *7* (7), 1340-1351.
61. Wichert, M.; Krall, N.; Decurtins, W.; Franzini, R. M.; Pretto, F.; Schneider, P.; Neri, D.; Scheuermann, J., Dual-display of small molecules enables the discovery of ligand pairs and facilitates affinity maturation. *Nat Chem* **2015**, *7* (3), 241-9.
62. Hoogsteen, K.; Trenner, N. R., The structure and conformation of the cis and trans isomers of 1-(p-chlorobenzylidene)-2-methyl-5-methoxyindenylic acid. *J Org Chem* **1970**, *35* (2), 521-3.
63. Kang, C.; Zhang, X.; Ratliff, R.; Moyzis, R.; Rich, A., Crystal structure of four-stranded Oxytricha telomeric DNA. *Nature* **1992**, *356* (6365), 126-31.
64. Laughlan, G.; Murchie, A. I.; Norman, D. G.; Moore, M. H.; Moody, P. C.; Lilley, D. M.; Luisi, B., The high-resolution crystal structure of a parallel-stranded guanine tetraplex. *Science* **1994**, *265* (5171), 520-4.
65. Makara, G. M.; Athanasopoulos, J., Improving success rates for lead generation using affinity binding technologies. *Curr Opin Biotechnol* **2005**, *16* (6), 666-73.
66. Scheuermann, J.; Neri, D., DNA-encoded chemical libraries: a tool for drug discovery and for chemical biology. *Chembiochem* **2010**, *11* (7), 931-7.
67. Mannocci, L.; Melkko, S.; Buller, F.; Molnar, I.; Bianke, J. P.; Dumelin, C. E.; Scheuermann, J.; Neri, D., Isolation of potent and specific trypsin inhibitors from a DNA-encoded chemical library. *Bioconjug Chem* **2010**, *21* (10), 1836-41.
68. Hoogenboom, H. R., Selecting and screening recombinant antibody libraries. *Nat Biotechnol* **2005**, *23* (9), 1105-16.
69. Disch, J. S.; Evindar, G.; Chiu, C. H.; Blum, C. A.; Dai, H.; Jin, L.; Schuman, E.; Lind, K. E.; Belyanskaya, S. L.; Deng, J.; Coppo, F.; Aquilani, L.; Graybill, T. L.; Cuozzo, J. W.; Lavu, S.; Mao, C.; Vlasuk, G. P.; Perni, R. B., Discovery of thieno[3,2-d]pyrimidine-6-carboxamides as potent inhibitors of SIRT1, SIRT2, and SIRT3. *J Med Chem* **2013**, *56* (9), 3666-79.
70. Deng, H.; O'Keefe, H.; Davie, C. P.; Lind, K. E.; Acharya, R. A.; Franklin, G. J.; Larkin, J.; Matico, R.; Neeb, M.; Thompson, M. M.; Lohr, T.; Gross, J. W.; Centrella, P. A.; O'Donovan, G. K.; Bedard, K. L.; van Vloten, K.; Mataruse, S.; Skinner, S. R.; Belyanskaya, S. L.; Carpenter, T. Y.; Shearer, T. W.; Clark, M. A.; Cuozzo, J. W.; Arico-Muendel, C. C.; Morgan, B. A., Discovery of highly potent and

REFERENCES

selective small molecule ADAMTS-5 inhibitors that inhibit human cartilage degradation via encoded library technology (ELT). *J Med Chem* **2012**, *55* (16), 7061-79.

71. Deng, H.; Zhou, J.; Sundersingh, F.; Messer, J. A.; Somers, D. O.; Ajakane, M.; Arico-Muendel, C. C.; Beljean, A.; Belyanskaya, S. L.; Bingham, R.; Blazensky, E.; Boullay, A. B.; Boursier, E.; Chai, J.; Carter, P.; Chung, C. W.; Daugan, A.; Ding, Y.; Herry, K.; Hobbs, C.; Humphries, E.; Kollmann, C.; Nguyen, V. L.; Nicodeme, E.; Smith, S. E.; Dodic, N.; Ancellin, N., Discovery and Optimization of Potent, Selective, and in Vivo Efficacious 2-Aryl Benzimidazole BCATm Inhibitors. *ACS Med Chem Lett* **2016**, *7* (4), 379-84.

72. Thalji, R. K.; McAtee, J. J.; Belyanskaya, S.; Brandt, M.; Brown, G. D.; Costell, M. H.; Ding, Y.; Dodson, J. W.; Eisennagel, S. H.; Fries, R. E.; Gross, J. W.; Harpel, M. R.; Holt, D. A.; Israel, D. I.; Jolivet, L. J.; Krosky, D.; Li, H.; Lu, Q.; Mandichak, T.; Roethke, T.; Schnackenberg, C. G.; Schwartz, B.; Shewchuk, L. M.; Xie, W.; Behm, D. J.; Douglas, S. A.; Shaw, A. L.; Marino, J. P., Jr., Discovery of 1-(1,3,5-triazin-2-yl)piperidine-4-carboxamides as inhibitors of soluble epoxide hydrolase. *Bioorg Med Chem Lett* **2013**, *23* (12), 3584-8.

73. Encinas, L.; O'Keefe, H.; Neu, M.; Remuinan, M. J.; Patel, A. M.; Guardia, A.; Davie, C. P.; Perez-Macias, N.; Yang, H.; Convery, M. A.; Messer, J. A.; Perez-Herran, E.; Centrella, P. A.; Alvarez-Gomez, D.; Clark, M. A.; Huss, S.; O'Donovan, G. K.; Ortega-Muro, F.; McDowell, W.; Castaneda, P.; Arico-Muendel, C. C.; Pajk, S.; Rullas, J.; Angulo-Barturen, I.; Alvarez-Ruiz, E.; Mendoza-Losana, A.; Ballell Pages, L.; Castro-Pichel, J.; Evindar, G., Encoded library technology as a source of hits for the discovery and lead optimization of a potent and selective class of bactericidal direct inhibitors of Mycobacterium tuberculosis InhA. *J Med Chem* **2014**, *57* (4), 1276-88.

74. Kollmann, C. S.; Bai, X.; Tsai, C. H.; Yang, H.; Lind, K. E.; Skinner, S. R.; Zhu, Z.; Israel, D. I.; Cuozzo, J. W.; Morgan, B. A.; Yuki, K.; Xie, C.; Springer, T. A.; Shimaoka, M.; Evindar, G., Application of encoded library technology (ELT) to a protein-protein interaction target: discovery of a potent class of integrin lymphocyte function-associated antigen 1 (LFA-1) antagonists. *Bioorg Med Chem* **2014**, *22* (7), 2353-65.

75. Wu, Z.; Graybill, T. L.; Zeng, X.; Platchek, M.; Zhang, J.; Bodmer, V. Q.; Wisnoski, D. D.; Deng, J.; Coppo, F. T.; Yao, G.; Tamburino, A.; Scavello, G.; Franklin, G. J.; Mataruse, S.; Bedard, K. L.; Ding, Y.; Chai, J.; Summerfield, J.; Centrella, P. A.; Messer, J. A.; Pope, A. J.; Israel, D. I., Cell-Based Selection Expands the Utility of DNA-Encoded Small-Molecule Library Technology to Cell Surface Drug Targets: Identification of Novel Antagonists of the NK3 Tachykinin Receptor. *ACS Comb Sci* **2015**, *17* (12), 722-31.

76. Deng, H.; Zhou, J.; Sundersingh, F. S.; Summerfield, J.; Somers, D.; Messer, J. A.; Satz, A. L.; Ancellin, N.; Arico-Muendel, C. C.; Sargent Bedard, K. L.; Beljean, A.; Belyanskaya, S. L.; Bingham, R.; Smith, S. E.; Boursier, E.; Carter, P.; Centrella, P. A.; Clark, M. A.; Chung, C. W.; Davie, C. P.; Delorey, J. L.; Ding, Y.; Franklin, G. J.; Grady, L. C.; Herry, K.; Hobbs, C.; Kollmann, C. S.; Morgan, B. A.; Pothier Kaushansky, L. J.; Zhou, Q., Discovery, SAR, and X-ray Binding Mode Study of BCATm Inhibitors from a Novel DNA-Encoded Library. *ACS Med Chem Lett* **2015**, *6* (8), 919-24.

77. Yang, H.; Medeiros, P. F.; Raha, K.; Elkins, P.; Lind, K. E.; Lehr, R.; Adams, N. D.; Burgess, J. L.; Schmidt, S. J.; Knight, S. D.; Auger, K. R.; Schaber, M. D.; Franklin, G. J.; Ding, Y.; DeLorey, J. L.; Centrella, P. A.; Mataruse, S.; Skinner, S. R.; Clark, M. A.; Cuozzo, J. W.; Evindar, G., Discovery of a

Potent Class of PI3K α Inhibitors with Unique Binding Mode via Encoded Library Technology (ELT). *ACS Med Chem Lett* **2015**, 6 (5), 531-6.

78. Seigal, B. A.; Connors, W. H.; Fraley, A.; Borzilleri, R. M.; Carter, P. H.; Emanuel, S. L.; Fagnoli, J.; Kim, K.; Lei, M.; Naglich, J. G.; Pokross, M. E.; Posy, S. L.; Shen, H.; Surti, N.; Talbott, R.; Zhang, Y.; Terrett, N. K., The discovery of macrocyclic XIAP antagonists from a DNA-programmed chemistry library, and their optimization to give lead compounds with in vivo antitumor activity. *J Med Chem* **2015**, 58 (6), 2855-61.

79. Gassman, N. R.; Nelli, J. P.; Dutta, S.; Kuhn, A.; Bonin, K.; Pianowski, Z.; Winssinger, N.; Guthold, M.; Macosko, J. C., Selection of bead-displayed, PNA-encoded chemicals. *J Mol Recognit* **2010**, 23 (5), 414-22.

80. Kleiner, R. E.; Dumelin, C. E.; Tiu, G. C.; Sakurai, K.; Liu, D. R., In vitro selection of a DNA-templated small-molecule library reveals a class of macrocyclic kinase inhibitors. *J Am Chem Soc* **2010**, 132 (33), 11779-91.

81. Podolin, P. L.; Bolognese, B. J.; Foley, J. F.; Long, E., 3rd; Peck, B.; Umbrecht, S.; Zhang, X.; Zhu, P.; Schwartz, B.; Xie, W.; Quinn, C.; Qi, H.; Sweitzer, S.; Chen, S.; Galop, M.; Ding, Y.; Belyanskaya, S. L.; Israel, D. I.; Morgan, B. A.; Behm, D. J.; Marino, J. P., Jr.; Kurali, E.; Barnette, M. S.; Mayer, R. J.; Booth-Genthe, C. L.; Callahan, J. F., In vitro and in vivo characterization of a novel soluble epoxide hydrolase inhibitor. *Prostaglandins Other Lipid Mediat* **2013**, 104-105, 25-31.

82. Franzini, R. M.; Ekblad, T.; Zhong, N.; Wichert, M.; Decurtins, W.; Nauer, A.; Zimmermann, M.; Samain, F.; Scheuermann, J.; Brown, P. J.; Hall, J.; Graslund, S.; Schuler, H.; Neri, D., Identification of structure-activity relationships from screening a structurally compact DNA-encoded chemical library. *Angew Chem Int Ed Engl* **2015**, 54 (13), 3927-31.

83. Dagher, J. P.; Zambaldo, C.; Abegg, D.; Barluenga, S.; Tallant, C.; Muller, S.; Adibekian, A.; Winssinger, N., Identification of Covalent Bromodomain Binders through DNA Display of Small Molecules. *Angew Chem Int Ed Engl* **2015**, 54 (20), 6057-61.

84. Arico-Muendel, C.; Zhu, Z.; Dickson, H.; Parks, D.; Keicher, J.; Deng, J.; Aquilani, L.; Coppo, F.; Graybill, T.; Lind, K.; Peat, A.; Thomson, M., Encoded library technology screening of hepatitis C virus NS4B yields a small-molecule compound series with in vitro replicon activity. *Antimicrob Agents Chemother* **2015**, 59 (6), 3450-9.

85. Petersen, L., Blakskjær, P., Chaikuad, A., Christensen A., Dietvorst, J., Holmkvist, S., Knapp, M., Kořínek, L., Larsen, A., Pedersen, S., Röhm, F., Hansen, N., Novel p38 α MAP kinase inhibitors identified from yoctoReactor DNA-encoded small molecule library. *Med. Chem. Commun* **2016**, 7 (7), 1332-1339.

86. Evans, B. E.; Rittle, K. E.; Bock, M. G.; DiPardo, R. M.; Freidinger, R. M.; Whitter, W. L.; Lundell, G. F.; Veber, D. F.; Anderson, P. S.; Chang, R. S.; et al., Methods for drug discovery: development of potent, selective, orally effective cholecystokinin antagonists. *J Med Chem* **1988**, 31 (12), 2235-46.

REFERENCES

87. DeSimone, R. W.; Currie, K. S.; Mitchell, S. A.; Darrow, J. W.; Pippin, D. A., Privileged structures: applications in drug discovery. *Comb Chem High Throughput Screen* **2004**, *7* (5), 473-94.
88. Schneider, G., Baringhaus, K. ,
Molecular Design: Concepts and Applications. 2008; p 277.
89. Welsch, M. E.; Snyder, S. A.; Stockwell, B. R., Privileged scaffolds for library design and drug discovery. *Curr Opin Chem Biol* **2010**, *14* (3), 347-61.
90. Filippakopoulos, P.; Qi, J.; Picaud, S.; Shen, Y.; Smith, W. B.; Fedorov, O.; Morse, E. M.; Keates, T.; Hickman, T. T.; Felletar, I.; Philpott, M.; Munro, S.; McKeown, M. R.; Wang, Y.; Christie, A. L.; West, N.; Cameron, M. J.; Schwartz, B.; Heightman, T. D.; La Thangue, N.; French, C. A.; Wiest, O.; Kung, A. L.; Knapp, S.; Bradner, J. E., Selective inhibition of BET bromodomains. *Nature* **2010**, *468* (7327), 1067-73.
91. Wells, J. A. M., C. L, Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. . *Nature* **2007**, *450* (7172), 1001-1009.
92. Baell, J. B.; Holloway, G. A., New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J Med Chem* **2010**, *53* (7), 2719-40.
93. Li, X.; Liu, D. R., DNA-templated organic synthesis: nature's strategy for controlling chemical reactivity applied to synthetic molecules. *Angew Chem Int Ed Engl* **2004**, *43* (37), 4848-70.
94. Satz, A. L.; Cai, J.; Chen, Y.; Goodnow, R.; Gruber, F.; Kowalczyk, A.; Petersen, A.; Naderi-Oboodi, G.; Orzechowski, L.; Strebler, Q., DNA Compatible Multistep Synthesis and Applications to DNA Encoded Libraries. *Bioconjug Chem* **2015**, *26* (8), 1623-32.
95. Skopic, M. K.; Salamon, H.; Bugain, O.; Jung, K.; Gohla, A.; Doetsch, L. J.; Dos Santos, D.; Bhat, A.; Wagner, B.; Brunschweiler, A., Acid- and Au(i)-mediated synthesis of hexathymidine-DNA-heterocycle chimeras, an efficient entry to DNA-encoded libraries inspired by drug structures. *Chem Sci* **2017**, *8* (5), 3356-3361.
96. Xiaohui, W., Xiaoyong, W., Shanshan, C., Yan, W., Guan, C., Zijian G., Specific recognition of DNA depurination by a luminescent terbium(III) complex. *Chem. Sci.* **2013**, *4* (13), 3748-3752.
97. Basile, A. S.; Lippa, A. S.; Skolnick, P., Anxiolytic drugs: can less be more? *Eur J Pharmacol* **2004**, *500* (1-3), 441-51.
98. Berezhnoy, D.; Nyfeler, Y.; Gonthier, A.; Schwob, H.; Goeldner, M.; Sigel, E., On the benzodiazepine binding pocket in GABAA receptors. *J Biol Chem* **2004**, *279* (5), 3160-8.
99. Mohler, H.; Okada, T., Benzodiazepine receptor: demonstration in the central nervous system. *Science* **1977**, *198* (4319), 849-51.

REFERENCES

100. Mori, K.; Togashi, H.; Kojima, T.; Matsumoto, M.; Ohashi, S.; Ueno, K.; Yoshioka, M., Different effects of anxiolytic agents, diazepam and 5-HT(1A) agonist tandospirone, on hippocampal long-term potentiation in vivo. *Pharmacol Biochem Behav* **2001**, *69* (3-4), 367-72.
101. Sigel, E.; Buhr, A., The benzodiazepine binding site of GABAA receptors. *Trends Pharmacol Sci* **1997**, *18* (11), 425-9.
102. Geller, I.; Kulak, J. T., Jr.; Seifter, J., The effects of chlordiazepoxide and chlorpromazine on a punishment discrimination. *Psychopharmacologia* **1962**, *3*, 374-85.
103. Lopes, D. V.; Caruso, R. R.; Castro, N. G.; Costa, P. R.; da Silva, A. J.; Noel, F., Characterization of a new synthetic isoflavonoid with inverse agonist activity at the central benzodiazepine receptor. *Eur J Pharmacol* **2004**, *495* (2-3), 87-96.
104. Sternbach, L. H., The benzodiazepine story. *J Med Chem* **1979**, *22* (1), 1-7.
105. Sternbach, L., Kaiser, S., Reeder, E., , Quinazoline 3-oxide structure of compounds previously described in the literature as 3, 1, 4-benzoxadiazepines. *J. Am. Chem. Soc.* **1960**, *82* (2), 475-480.
106. Sternbach, L., Reeder, E., , Quinazolines and 1,4-benzodiazepines IV. Transformations of 7-chloro-2-methylamino-5-phenyl-3H-1,4-benzodiazepine 4- oxide. *J. Org. Chem.* **1961**, *26* (12), 4488-4497.
107. Covelli, V.; Maffione, A. B.; Nacci, C.; Tato, E.; Jirillo, E., Stress, neuropsychiatric disorders and immunological effects exerted by benzodiazepines. *Immunopharmacol Immunotoxicol* **1998**, *20* (2), 199-209.
108. Rudolph, U.; Crestani, F.; Benke, D.; Brunig, I.; Benson, J. A.; Fritschy, J. M.; Martin, J. R.; Bluethmann, H.; Mohler, H., Benzodiazepine actions mediated by specific gamma-aminobutyric acid(A) receptor subtypes. *Nature* **1999**, *401* (6755), 796-800.
109. Tecott, L. H., Designer genes and anti-anxiety drugs. *Nat Neurosci* **2000**, *3* (6), 529-30.
110. Venter, J., Harrison, L. , Benzodiazepine/GABA Receptors and Chloride Channels. *Liss. Inc.* **1986**, *5793* (87), 80785-80788.
111. Lader, M., Antianxiety drugs: clinical pharmacology and therapeutic use. *Drugs* **1976**, *12* (5), 362-373.
112. Addicks, E.; Mazitschek, R.; Giannis, A., Synthesis and biological investigation of novel tricyclic benzodiazepinedione-based RGD analogues. *Chembiochem* **2002**, *3* (11), 1078-88.
113. Woods, J. H.; Winger, G., Current benzodiazepine issues. *Psychopharmacology (Berl)* **1995**, *118* (2), 107-15; discussion 118, 120-1.
114. Webb II, R. R., Barker, P., Baier, M., Reynolds, M. E., Robarge, K. D.,; Blackburn, B. K., Tischler, M., Weese, K. J., Mono-N-alkylation of anthranilamides via quinazolinones. *Tetrahedron Letters* **1994**, (35).

REFERENCES

115. Mc Dowell, R. S., Blackburn, B. K., Gadek, T. R., McGee, L. R., Rawson, T., Reynolds, M. E., K.D., R., Somers, T. C., Thorsett, E. D., Tischler, M., Webb II, R. R., M. C., V., From peptide to non-peptide. The de novo design of potent, nonpeptidal inhibitors of the platelet aggregation based on a benzodiazepinedione scaffold. *J. Am. Chem. Soc.* **1994**, *116* (12), 5077–5083.
116. Trimble, M., Benzodiazepines Divided. *John Wiley and Sons: New York* **1983**, *15* (8), 1259-1267.
117. Carlier, P. R.; Zhao, H.; DeGuzman, J.; Lam, P. C., Enantioselective synthesis of "quaternary" 1,4-benzodiazepin-2-one scaffolds via memory of chirality. *J Am Chem Soc* **2003**, *125* (38), 11482-3.
118. MacQuarrie-Hunter, S.; Carlier, P. R., Highly enantioselective synthesis of rigid, quaternary 1,4-benzodiazepine-2,5-diones derived from proline. *Org Lett* **2005**, *7* (23), 5305-8.
119. Stigter, E. A.; Guo, Z.; Bon, R. S.; Wu, Y. W.; Choidas, A.; Wolf, A.; Menninger, S.; Waldmann, H.; Blankenfeldt, W.; Goody, R. S., Development of selective, potent RabGGTase inhibitors. *J Med Chem* **2012**, *55* (19), 8330-40.
120. Wright, W. B., Jr.; Brabander, H. J.; Greenblatt, E. N.; Day, I. P.; Hardy, R. A., Jr., Derivatives of 1,2,3,11a-tetrahydro-5H-pyrrolo[2,1-c][1,4]benzodiazepine-5,11(10H)-dione as anxiolytic agents. *J Med Chem* **1978**, *21* (10), 1087-9.
121. Tetzlaff, C., Schwöpe, I., Blecziński, J., Steinberg, A. Richert, C., A convenient synthesis of 5'-amino-5'-deoxythymidine and preparation of peptide-DNA hybrids. *Tetrahedron Lett.* **1998**, *39* (1), 4215–4218.
122. Wyatt, P. G.; Woodhead, A. J.; Berdini, V.; Boulstridge, J. A.; Carr, M. G.; Cross, D. M.; Davis, D. J.; Devine, L. A.; Early, T. R.; Feltell, R. E.; Lewis, E. J.; McMenemy, R. L.; Navarro, E. F.; O'Brien, M. A.; O'Reilly, M.; Reule, M.; Saxty, G.; Seavers, L. C.; Smith, D. M.; Squires, M. S.; Trewartha, G.; Walker, M. T.; Woolford, A. J., Identification of N-(4-piperidinyl)-4-(2,6-dichlorobenzoylamino)-1H-pyrazole-3-carboxamide (AT7519), a novel cyclin dependent kinase inhibitor using fragment-based X-ray crystallography and structure based drug design. *J Med Chem* **2008**, *51* (16), 4986-99.
123. Samain, F.; Ekblad, T.; Mikutis, G.; Zhong, N.; Zimmermann, M.; Nauer, A.; Bajic, D.; Decurtins, W.; Scheuermann, J.; Brown, P. J.; Hall, J.; Graslund, S.; Schuler, H.; Neri, D.; Franzini, R. M., Tankyrase 1 Inhibitors with Drug-like Properties Identified by Screening a DNA-Encoded Chemical Library. *J Med Chem* **2015**, *58* (12), 5143-9.
124. Gorham, R. D., Jr.; Nunez, V.; Lin, J. H.; Rooijackers, S. H.; Vullev, V. I.; Morikis, D., Discovery of Small Molecules for Fluorescent Detection of Complement Activation Product C3d. *J Med Chem* **2015**, *58* (24), 9535-45.
125. Škopić, M., Bugain, O., Jung, K., Onstein, S., Brandherm, S., Brunschweiler, A., Design and synthesis of DNA-encoded libraries based on a benzodiazepine and a pyrazolopyrimidine *Chem.Med.Comm* **2016**, (7), 1957-1965.
126. Craik, D. J.; Fairlie, D. P.; Liras, S.; Price, D., The future of peptide-based drugs. *Chem Biol Drug Des* **2013**, *81* (1), 136-47.

REFERENCES

127. Fosgerau, K.; Hoffmann, T., Peptide therapeutics: current status and future directions. *Drug Discov Today* **2015**, *20* (1), 122-8.
128. Nevola, L.; Giralt, E., Modulating protein-protein interactions: the potential of peptides. *Chem Commun (Camb)* **2015**, *51* (16), 3302-15.
129. Kaspar, A. A.; Reichert, J. M., Future directions for peptide therapeutics development. *Drug Discov Today* **2013**, *18* (17-18), 807-17.
130. Bock, J. E.; Gavenonis, J.; Kritzer, J. A., Getting in shape: controlling peptide bioactivity and bioavailability using conformational constraints. *ACS Chem Biol* **2013**, *8* (3), 488-99.
131. Fischer, E., Einfluss der Konfiguration auf die Wirkung der Enzyme. *Ber. Dtsch. Chem. Ges.* **1894**, *27* (1), 2985– 2993.
132. Gilon, C.; Halle, D.; Chorev, M.; Selinger, Z.; Byk, G., Backbone cyclization: A new method for conferring conformational constraint on peptides. *Biopolymers* **1991**, *31* (6), 745-50.
133. Koshland, D. E., Jr., Correlation of Structure and Function in Enzyme Action. *Science* **1963**, *142* (3599), 1533-41.
134. Fominaya, J.; Bravo, J.; Rebollo, A., Strategies to stabilize cell penetrating peptides for in vivo applications. *Ther Deliv* **2015**, *6* (10), 1171-94.
135. Pelay-Gimeno, M.; Glas, A.; Koch, O.; Grossmann, T. N., Structure-Based Design of Inhibitors of Protein-Protein Interactions: Mimicking Peptide Binding Epitopes. *Angew Chem Int Ed Engl* **2015**, *54* (31), 8896-927.
136. Milroy, L. G.; Grossmann, T. N.; Hennig, S.; Brunsveld, L.; Ottmann, C., Modulators of protein-protein interactions. *Chem Rev* **2014**, *114* (9), 4695-748.
137. Raj, M.; Bullock, B. N.; Arora, P. S., Plucking the high hanging fruit: a systematic approach for targeting protein-protein interactions. *Bioorg Med Chem* **2013**, *21* (14), 4051-7.
138. Watkins, A. M.; Arora, P. S., Structure-based inhibition of protein-protein interactions. *Eur J Med Chem* **2015**, *94*, 480-8.
139. Patgiri, A.; Jochim, A. L.; Arora, P. S., A hydrogen bond surrogate approach for stabilization of short peptide sequences in alpha-helical conformation. *Acc Chem Res* **2008**, *41* (10), 1289-300.
140. Arkin, M. R.; Tang, Y.; Wells, J. A., Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. *Chem Biol* **2014**, *21* (9), 1102-14.
141. Cromm, P. M.; Spiegel, J.; Grossmann, T. N., Hydrocarbon stapled peptides as modulators of biological function. *ACS Chem Biol* **2015**, *10* (6), 1362-75.
142. Wojcik, P.; Berlicki, L., Peptide-based inhibitors of protein-protein interactions. *Bioorg Med Chem Lett* **2016**, *26* (3), 707-13.

REFERENCES

143. Houk, K. N.; Leach, A. G.; Kim, S. P.; Zhang, X., Binding affinities of host-guest, protein-ligand, and protein-transition-state complexes. *Angew Chem Int Ed Engl* **2003**, *42* (40), 4872-97.
144. Cromm, P. M.; Spiegel, J.; Grossmann, T. N.; Waldmann, H., Direct Modulation of Small GTPase Activity and Function. *Angew Chem Int Ed Engl* **2015**, *54* (46), 13516-37.
145. Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K., The exploration of macrocycles for drug discovery an underexploited structural class. *Nat Rev Drug Discov* **2008**, *7* (7), 608-24.
146. Mallinson, J.; Collins, I., Macrocycles in new drug discovery. *Future Med Chem* **2012**, *4* (11), 1409-38.
147. Marsault, E.; Peterson, M. L., Macrocycles are great cycles: applications, opportunities, and challenges of synthetic macrocycles in drug discovery. *J Med Chem* **2011**, *54* (7), 1961-2004.
148. Scholtz, J. M.; Baldwin, R. L., The mechanism of alpha-helix formation by peptides. *Annu Rev Biophys Biomol Struct* **1992**, *21*, 95-118.
149. Bhat, A.; Roberts, L. R.; Dwyer, J. J., Lead discovery and optimization strategies for peptide macrocycles. *Eur J Med Chem* **2015**, *94*, 471-9.
150. Clark, R. J.; Akcan, M.; Kaas, Q.; Daly, N. L.; Craik, D. J., Cyclization of conotoxins to improve their biopharmaceutical properties. *Toxicon* **2012**, *59* (4), 446-55.
151. Dharanipragada, R., New modalities in conformationally constrained peptides for potency, selectivity and cell permeation. *Future Med Chem* **2013**, *5* (7), 831-49.
152. Katsara, M.; Tselios, T.; Deraos, S.; Deraos, G.; Matsoukas, M. T.; Lazoura, E.; Matsoukas, J.; Apostolopoulos, V., Round and round we go: cyclic peptides in disease. *Curr Med Chem* **2006**, *13* (19), 2221-32.
153. Hruby, V. J., Peptide science: exploring the use of chemical principles and interdisciplinary collaboration for understanding life processes. *J Med Chem* **2003**, *46* (20), 4215-31.
154. Montalbetti, C., Falque, V., Amide bond formation and peptide coupling. *Tetrahedron Lett.* **2005**, *61* (46), 10827– 10852.
155. Parenty, A.; Moreau, X.; Niel, G.; Campagne, J. M., Update 1 of: macrolactonizations in the total synthesis of natural products. *Chem Rev* **2013**, *113* (1), PR1-40.
156. Gongora-Benitez, M.; Tulla-Puche, J.; Albericio, F., Multifaceted roles of disulfide bonds. Peptides as therapeutics. *Chem Rev* **2014**, *114* (2), 901-26.
157. Vasco, A. V.; Perez, C. S.; Morales, F. E.; Garay, H. E.; Vasilev, D.; Gavin, J. A.; Wessjohann, L. A.; Rivera, D. G., Macrocyclization of Peptide Side Chains by the Ugi Reaction: Achieving Peptide Folding and Exocyclic N-Functionalization in One Shot. *J Org Chem* **2015**, *80* (13), 6697-707.
158. Turner, R. A.; Oliver, A. G.; Lokey, R. S., Click chemistry as a macrocyclization tool in the solid-phase synthesis of small cyclic peptides. *Org Lett* **2007**, *9* (24), 5011-4.

REFERENCES

159. Mendive-Tapia, L.; Preciado, S.; Garcia, J.; Ramon, R.; Kielland, N.; Albericio, F.; Lavilla, R., New peptide architectures through C-H activation stapling between tryptophan-phenylalanine/tyrosine residues. *Nat Commun* **2015**, *6*, 7160.
160. Perez de Vega, M. J.; Garcia-Aranda, M. I.; Gonzalez-Muniz, R., A role for ring-closing metathesis in medicinal chemistry: mimicking secondary architectures in bioactive peptides. *Med Res Rev* **2011**, *31* (5), 677-715.
161. Wang, D.; Chen, K.; Kulp Iii, J. L.; Arora, P. S., Evaluation of biologically relevant short alpha-helices stabilized by a main-chain hydrogen-bond surrogate. *J Am Chem Soc* **2006**, *128* (28), 9248-56.
162. Furstner, A., Alkyne metathesis on the rise. *Angew Chem Int Ed Engl* **2013**, *52* (10), 2794-819.
163. Aguilera, B.; Wolf, L. B.; Nieczypor, P.; Rutjes, F. P.; Overkleeft, H. S.; van Hest, J. C.; Schoemaker, H. E.; Wang, B.; Mol, J. C.; Furstner, A.; Overhand, M.; van der Marel, G. A.; van Boom, J. H., Synthesis of diaminosuberic acid derivatives via ring-closing alkyne metathesis. *J Org Chem* **2001**, *66* (10), 3584-9.
164. Ghalit, N.; Poot, A. J.; Furstner, A.; Rijkers, D. T.; Liskamp, R. M., Ring-closing alkyne metathesis approach toward the synthesis of alkyne mimics of thioether A-, B-, C-, and DE-ring systems of the lantibiotic nisin Z. *Org Lett* **2005**, *7* (14), 2961-4.
165. IJsselstijn, M., Ring-closing alkyne metathesis mediated synthesis of cyclic β -turn mimetics. *Tetrahedron Letters* **2004**, *45* (22), 4379-4382.
166. White, C. J.; Yudin, A. K., Contemporary strategies for peptide macrocyclization. *Nat Chem* **2011**, *3* (7), 509-24.
167. Baran, P. S.; Guerrero, C. A.; Corey, E. J., Short, enantioselective total synthesis of okaramine N. *J Am Chem Soc* **2003**, *125* (19), 5628-9.
168. Ekebergh, A.; Borje, A.; Martensson, J., Total synthesis of nostodione A, a cyanobacterial metabolite. *Org Lett* **2012**, *14* (24), 6274-7.
169. Ishikawa, H.; Colby, D. A.; Seto, S.; Va, P.; Tam, A.; Kakei, H.; Rayl, T. J.; Hwang, I.; Boger, D. L., Total synthesis of vinblastine, vincristine, related natural products, and key structural analogues. *J Am Chem Soc* **2009**, *131* (13), 4904-16.
170. Muscia, G., De María, L., Buldain, G., Asís, S., , Ultrasound Assisted Pictet–Spengler Synthesis of Tetrahydro- β -Carboline Derivatives. *Journal of Heterocyclic Chemistry* **2015**, *53* (2), 647–650.
171. Gribble, G., Recent developments in indole ring synthesis-methodology and applications. *J. Chem. Soc* **2000**, *1* (7), 1045-1075.
172. Zhang, M. Z.; Chen, Q.; Yang, G. F., A review on recent developments of indole-containing antiviral agents. *Eur J Med Chem* **2015**, *89*, 421-41.

REFERENCES

173. Sharath, V., Kumar, H., Naik, N. , Synthesis of novel indole based scaffolds holding pyrazole ring as anti-inflammatory and antioxidant agents. *Journal of Pharmacy Research* **2013**, (7), 785-790.
174. Kaushik, N. K.; Kaushik, N.; Attri, P.; Kumar, N.; Kim, C. H.; Verma, A. K.; Choi, E. H., Biomedical importance of indoles. *Molecules* **2013**, 18 (6), 6620-62.
175. Lood, C., Koskinen, A. , Harmicine: a Tetracyclic Tetrahydro- β -Carboline: From the First Synthetic Precedent to Isolation from Natural Sources to Target-Oriented Synthesis. *Chem. Heterocycl. Comp.* **2015**, 50 (10), 1367-1387.
176. Sennhenn, P., Mantoulidis, A., Treu, M., Tontsch-Grunt, U., Spevak, W., McConnell, D., Schoop, A., Brueckner, R., Jacobi, A., Guertler, U., Schnapp, G., Klein, C., Himmelsbach, F., Pautsch, A., Betzemeier, B., Herfurth, L., Mack, J., Wiedenmayer, D., Bader, G., Reiser, U. Preparation of α -carbolines as CDK1 inhibitors. 2006.
177. Matveeva, I. A., [Action of dimebon on histamine receptors]. *Farmakol Toksikol* **1983**, 46 (4), 27-9.
178. Shevtsova, E. F.; Kireeva, E. G.; Bachurin, S. O., [Mitochondria as the target for neuroprotectors]. *Vestn Ross Akad Med Nauk* **2005**, (9), 13-7.
179. Drucker, G.; Raikoff, K.; Neafsey, E. J.; Collins, M. A., Dopamine uptake inhibitory capacities of beta-carboline and 3,4-dihydro-beta-carboline analogs of N-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) oxidation products. *Brain Res* **1990**, 509 (1), 125-33.
180. Matsubara, K.; Gonda, T.; Sawada, H.; Uezono, T.; Kobayashi, Y.; Kawamura, T.; Ohtaki, K.; Kimura, K.; Akaike, A., Endogenously occurring beta-carboline induces parkinsonism in nonprimate animals: a possible causative protoxin in idiopathic Parkinson's disease. *J Neurochem* **1998**, 70 (2), 727-35.
181. Rommelspacher, H.; Meier-Henco, M.; Smolka, M.; Kloft, C., The levels of norharman are high enough after smoking to affect monoamineoxidase B in platelets. *Eur J Pharmacol* **2002**, 441 (1-2), 115-25.
182. Kuo, F., Tseng, M., Yen, Y., Chu, Y. , Microwave accelerated Pictet– Spengler reactions of tryptophan with ketones directed toward the preparation of 1,1- disubstituted indole alkaloids. *Tetrahedron Lett.* **2004**, 60 (52), 12075-12084.
183. Cao, R.; Chen, Q.; Hou, X.; Chen, H.; Guan, H.; Ma, Y.; Peng, W.; Xu, A., Synthesis, acute toxicities, and antitumor effects of novel 9-substituted beta-carboline derivatives. *Bioorg Med Chem* **2004**, 12 (17), 4613-23.
184. Polychronopoulos, P.; Magiatis, P.; Skaltsounis, A. L.; Myrianthopoulos, V.; Mikros, E.; Tarricone, A.; Musacchio, A.; Roe, S. M.; Pearl, L.; Leost, M.; Greengard, P.; Meijer, L., Structural basis for the synthesis of indirubins as potent and selective inhibitors of glycogen synthase kinase-3 and cyclin-dependent kinases. *J Med Chem* **2004**, 47 (4), 935-46.

REFERENCES

185. Dekhane, M., Dubois, L., Blanchet, G., Garrigue, H., Sentenacroumanou, H., Potier, P., Dodd, R., N-2 Methylated Quaternary Derivatives of Beta-Carboline-3-Carboxylates Inhibit Acetylcholinesterase In-Vitro *Bioorg. Med. Chem. Lett.* **1993**, 3 (12), 2831-2836.
186. Gearhart, D. A.; Collins, M. A.; Lee, J. M.; Neafsey, E. J., Increased beta-carboline 9N-methyltransferase activity in the frontal cortex in Parkinson's disease. *Neurobiol Dis* **2000**, 7 (3), 201-11.
187. Hutchins, S., Chapman, K., Solid phase synthesis of tetrahydroisoquinolines & tetrahydroimidazopyridines. *Tetrahedron Letters* **1996**, 37 (28), 4865-4868.
188. Stockigt, J.; Antonchick, A. P.; Wu, F.; Waldmann, H., The Pictet-Spengler reaction in nature and in organic chemistry. *Angew Chem Int Ed Engl* **2011**, 50 (37), 8538-64.
189. Wang, Z., *Pictet-Spengler Reaction. In Comprehensive Organic Name Reactions and Reagents*. 2010.
190. Sinha, M. K.; Khoury, K.; Herdtweck, E.; Domling, A., Tricycles by a new Ugi variation and Pictet-Spengler reaction in one pot. *Chemistry* **2013**, 19 (25), 8048-52.
191. Vitali, T.; Impicciatore, M.; Plazzi, P. V.; Bordi, F.; Vitto, M., [Imidazole H-2 agonists. Synthesis and activity of 2-(2-amino-4-imidazolyl)ethylamine (2-aminohistamine) dihydrochloride]. *Farmaco Sci* **1984**, 39 (1), 70-80.
192. Hahn, G., Hansel, A. *Kondensation von Tryptamin mit α -Ketonsäuren, α -Ketodicarbonsäuren und α,α' -Diketo-dicarbonsäuren*; 1938; pp 2163-2175.
193. Snyder, H. R.; Hansch, C. H.; et al., The synthesis of derivatives of beta-carboline; syntheses from dl-tryptophan and aldehydes. *J Am Chem Soc* **1948**, 70 (1), 219-21.
194. Subba Reddy, B. V.; Swain, M.; Reddy, S. M.; Yadav, J. S.; Sridhar, B., Gold-catalyzed domino cycloisomerization/Pictet-Spengler reaction of 2-(4-aminobut-1-yn-1-yl)anilines with aldehydes: synthesis of tetrahydropyrido[4,3-b]indole scaffolds. *J Org Chem* **2012**, 77 (24), 11355-61.
195. Baldwin, J., Rules for ring closure. *Ciba Found Symp* **1978**, (53), 85-99.
196. Johnson, C. D., Stereoelectronic effects in the formation of 5- and 6-membered rings: the role of Baldwin's rules. *Acc. Chem. Res* **1993**, 26 (9), 476-482.
197. Gremmen, C.; Willemse, B.; Wanner, M. J.; Koomen, G. J., Enantiopure Tetrahydro-beta-carbolines via Pictet-Spengler Reactions with N-Sulfinyl Tryptamines. *Org Lett* **2000**, 2 (13), 1955-1958.
198. Grigg, R., Gunaratne, H., McNaghten, E., Observations on the pictet-spengler synthesis of 1,2,3,4-tetrahydro-[small beta]-carbolines. *J. Chem. Soc.* **1983**, 1 (52), 185-187.
199. Silveira, C., Vieira, A., Kaufman, T., Thiophenol-mediated improvement of the Pictet-Spengler cyclization of N-tosyl- β -phenethylamines with aldehydes. *Tetrahedron Letters* **2006**, 47 (43), 7545-7549.

REFERENCES

200. Youn, S. W., Development of the Pictet-Spengler reaction catalyzed by AuCl₃/AgOTf. *J Org Chem* **2006**, *71* (6), 2521-3.
201. Manabe, K.; Nobutou, D.; Kobayashi, S., Catalytic Pictet-Spengler reactions using Yb(OTf)₃. *Bioorg Med Chem* **2005**, *13* (17), 5154-8.
202. Srinivasan, N.; Ganesan, A., Highly efficient Lewis acid-catalysed Pictet-Spengler reactions discovered by parallel screening. *Chem Commun (Camb)* **2003**, (7), 916-7.
203. Tsuji, R., Yamanaka, M., Nishida, A., Nakagawa, M. , Pictet–Spengler Reaction of Nitrones and Imines Catalyzed by Yb(OTf)₃–TMSCl. *Chemistry Letters* **2002**, *31* (4), 428-429.
204. Ryabukhin, S., Panov, D., Plaskon, A., Tolmachev, A., Smaliy, R., Application of chlorotrimethylsilane in Pictet–Spengler reaction. *Monatsh. Chem.* **2012**, *143* (11), 1507-1517.
205. Prajapati, D., Gohain, M., Iodine-Catalyzed Highly Effective Pictet–Spengler Condensation: An Efficient Synthesis of Tetrahydro-β-carbolines. *Synthetic Communications* **2008**, *38* (24), 4426-4433.
206. Zheng, C.; Fang, Y.; Tong, W.; Li, G.; Wu, H.; Zhou, W.; Lin, Q.; Yang, F.; Yang, Z.; Wang, P.; Peng, Y.; Pang, X.; Yi, Z.; Luo, J.; Liu, M.; Chen, Y., Synthesis and biological evaluation of novel tetrahydro-beta-carboline derivatives as antitumor growth and metastasis agents through inhibiting the transforming growth factor-beta signaling pathway. *J Med Chem* **2014**, *57* (3), 600-12.
207. Skouta, R.; Hayano, M.; Shimada, K.; Stockwell, B. R., Design and synthesis of Pictet-Spengler condensation products that exhibit oncogenic-RAS synthetic lethality and induce non-apoptotic cell death. *Bioorg Med Chem Lett* **2012**, *22* (17), 5707-13.
208. Liew, P., Fleming, J., Longeon, A., Mouray, E., Florent, I., Bourguet-Kondracki, M., Copp, B., Synthesis of 1-indolyl substituted β-carboline natural products and discovery of antimalarial and cytotoxic activities. *Tetrahedron Lett.* **2014**, *70* (33), 4910–4920.
209. Ho, B. T.; Taylor, D.; Walker, K. E.; Mclsaac, W. M., The mode of action of 6-methoxy-1,2,3,4-tetrahydro-β-carboline on brain serotonin. *Can J Biochem* **1973**, *51* (4), 482-5.
210. Squires, P. E.; Hills, C. E.; Rogers, G. J.; Garland, P.; Farley, S. R.; Morgan, N. G., The putative imidazoline receptor agonist, harmaline, promotes intracellular calcium mobilisation in pancreatic beta-cells. *Eur J Pharmacol* **2004**, *501* (1-3), 31-9.
211. Rommelspacher, H.; Strauss, S. M.; Rehse, K., beta-Carbolines: a tool for investigating structure-activity relationships of the high-affinity uptake of serotonin, noradrenaline, dopamine, GABA and choline into a synaptosome-rich fraction of various regions from rat brain. *J Neurochem* **1978**, *30* (6), 1573-8.
212. Wrona, M. Z.; Waskiewicz, J.; Han, Q. P.; Han, J.; Li, H.; Dryhurst, G., Putative oxidative metabolites of 1-methyl-6-hydroxy-1,2,3,4-tetrahydro-beta-carboline of potential relevance to the addictive and neurodegenerative consequences of ethanol abuse. *Alcohol* **1997**, *14* (3), 213-23.

REFERENCES

213. Herraiz, T.; Chaparro, C., Human monoamine oxidase enzyme inhibition by coffee and beta-carbolines norharman and harman isolated from coffee. *Life Sci* **2006**, *78* (8), 795-802.
214. de Jesus, M. B.; Pinto Lde, M.; Fraceto, L. F.; Magalhaes, L. A.; Zanotti-Magalhaes, E. M.; de Paula, E., Improvement of the oral praziquantel anthelmintic effect by cyclodextrin complexation. *J Drug Target* **2010**, *18* (1), 21-6.
215. Turner, H., Spiroindolone NITD609 is a novel antimalarial drug that targets the P-type ATPase PfATP4. *Future Med Chem* **2016**, *8* (2), 227-38.
216. Coward, R. M.; Carson, C. C., Tadalafil in the treatment of erectile dysfunction. *Ther Clin Risk Manag* **2008**, *4* (6), 1315-30.
217. Bellamy, N., Etodolac in the management of pain: a clinical review of a multipurpose analgesic. *Inflammopharmacology* **1997**, *5* (2), 139-52.
218. Pahkla, R.; Zilmer, M.; Kullisaar, T.; Rago, L., Comparison of the antioxidant activity of melatonin and pinoline in vitro. *J Pineal Res* **1998**, *24* (2), 96-101.
219. Komulainen, H.; Tuomisto, J.; Airaksinen, M. M.; Kari, I.; Peura, P.; Pollari, L., Tetrahydro-beta-carbolines and corresponding tryptamines: In vitro inhibition of serotonin, dopamine and noradrenaline uptake in rat brain synaptosomes. *Acta Pharmacol Toxicol (Copenh)* **1980**, *46* (4), 299-307.
220. Boerner, L. J.; Zaleski, J. M., Metal complex-DNA interactions: from transcription inhibition to photoactivated cleavage. *Curr Opin Chem Biol* **2005**, *9* (2), 135-44.
221. Hadjiliadis, N., Sletten, E. , *Metal Complex-DNA Interactions*. Wiley-Blackwell New York.: **2009**.
222. Bannwarth, W., Gene Technology: a challenge for chemist. *Chimia* **1987**, *41* (9), 302–317.
223. Franzini, R. M.; Samain, F.; Abd Elrahman, M.; Mikutis, G.; Nauer, A.; Zimmermann, M.; Scheuermann, J.; Hall, J.; Neri, D., Systematic evaluation and optimization of modification reactions of oligonucleotides with amines and carboxylic acids for the synthesis of DNA-encoded chemical libraries. *Bioconjug Chem* **2014**, *25* (8), 1453-61.
224. Majik, M. S.; Rodrigues, C.; Mascarenhas, S.; D'Souza, L., Design and synthesis of marine natural product-based 1H-indole-2,3-dione scaffold as a new antifouling/antibacterial agent against fouling bacteria. *Bioorg Chem* **2014**, *54*, 89-95.
225. De Savi, C.; Bradbury, R. H.; Rabow, A. A.; Norman, R. A.; de Almeida, C.; Andrews, D. M.; Ballard, P.; Buttar, D.; Callis, R. J.; Currie, G. S.; Curwen, J. O.; Davies, C. D.; Donald, C. S.; Feron, L. J.; Gingell, H.; Glossop, S. C.; Hayter, B. R.; Hussain, S.; Karoutchi, G.; Lamont, S. G.; MacFaul, P.; Moss, T. A.; Pearson, S. E.; Tonge, M.; Walker, G. E.; Weir, H. M.; Wilson, Z., Optimization of a Novel Binding Motif to (E)-3-(3,5-Difluoro-4-((1R,3R)-2-(2-fluoro-2-methylpropyl)-3-methyl-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indol-1-yl)phenyl)acrylic Acid (AZD9496), a Potent and Orally Bioavailable Selective Estrogen Receptor Downregulator and Antagonist. *J Med Chem* **2015**, *58* (20), 8128-40.