UNIVERSITÄT DUISBURG ESSEN

SCIENCE ON THE RHINE AND THE RUHR

MOLECULAR RECOGNITION OF BIOLOGICALLY RELEVANT TARGETS

-

FROM ANION BINDING MOTIFS TO APPLICATION IN CELL BIOLOGY

Dissertation

zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften

- Dr. rer. nat. -

vorgelegt von

Diplom-Chemiker

Hannes Yacu Kuchelmeister

geboren in Illertissen

erstellt am

Institut für Organische Chemie

an der

Universität Duisburg-Essen

2011

Die vorliegende Arbeit wurde von März 2008 bis Oktober 2011 in den Instituten für Organische Chemie der Universitäten Würzburg und Duisburg-Essen unter der Anleitung von Herrn Prof. Dr. Carsten Schmuck angefertigt.

Ich erkläre hiermit des Eides statt, dass ich die vorliegende Arbeit selbst verfasst und mich dabei keiner anderen als der von mir bezeichneten Quellen und Hilfen bedient habe. Ich erkläre hiermit, dass ich an keiner anderen Stelle ein Prüfungsverfahren beantragt beziehungsweise die Dissertation in dieser oder anderer Form bereits anderweitig als Prüfungsarbeit verwendet oder einer anderen Fakultät als Dissertation vorgelegt habe.

Essen, September 2011

……

Hannes Y. Kuchelmeister

Disputation:	27. Oktober 2011
Gutachter:	Prof. Dr. Carsten Schmuck
	Prof. Dr. Shirley Knauer
Prüfungsvorsitzender:	Prof. Dr. Torsten C. Schmidt

DANKSAGUNG

An erster Stelle möchte ich natürlich meinem Doktorvater, Prof. Dr. Carsten Schmuck, für die Möglichkeit danken, auf dem interessanten Gebiet der bioorganischen Chemie zu forschen. Durch seine hilfreiche fachliche Unterstützung, eine stets offene Tür und seine umgängliche, freundliche Art hat er wesentlich zum Gelingen dieser Arbeit beigetragen. Besonders bedanken möchte ich mich für die große Gestaltungsfreiheit bei der Bearbeitung meines Themas und dem stets für neue Ideen offenen Ohr.

Für die finanzielle Unterstützung meiner Promotion möchte ich mich bei der Studienstiftung des deutschen Volkes bedanken.

Der analytischen Abteilung der Universität Duisburg-Essen gilt mein Dank: Dr. Torsten Schaller und Heinz Bandmann für die Durchführung von NMR Messungen und Werner Karow für die Aufnahme der Massenspektren.

Bei Wilhelm Sicking möchte ich mich für das DNA-Modelling bedanken.

Ursula Nüchter gilt mein Dank für das Messen zwischen den Jahren.

Allen Auszubildenden, Bachelor- und Vertiefungsstudenten möchte ich für ihren Einsatz und Beitrag zu dieser Doktorarbeit danken.

Besonderer Dank gilt meinen Kooperationspartnern ohne deren Einsatz und Arbeit diese Dissertation nur die Hälfte wert wäre:

> Magdalena Gellner, Stephan Niebling und Prof. Dr. Sebastian Schlücker möchte ich für die angenehme und unkomplizierte Zusammenarbeit danken.
>
> Bei Dr. Ivo Crnolatac, Dr. Marijana Radić Stojković, Dr. Lidija-Marija Tumir und Dr. Ivo Piantanida möchte ich mich für die fachliche Hilfe in allen praktischen und theoretischen Belangen der Welt der Nukleinsäuren, die tolle Zeit in Zagreb, das schlechte Fußballspiel und das gute Essen bedanken.
>
> Aljona Gutschmidt, Sarah Tillmann und Prof. Dr. Shirley Knauer sei gedankt für die Zuverlässigkeit und Sparsamkeit, das schnelle Zellen-Zählen, die freundliche und unkomplizierte Art und die kurzfristige und unglaublich schnelle Hilfe im Juli.

An den gesamten AK Schmuck und insbesondere an meine Würzburger und Essener Laborkollegen Dr. Volker Bickert, Dr. Nicholas Walden, Ute Michels, Pia Mereu und Tassilo Fenske geht mein Dank für die positive Atmosphäre und die gute Zusammenarbeit.

Für das Korrekturlesen und die kritische Durchsicht des Manuskripts möchte ich Barbara Geibel, Sebastian Langolf, Dr. Ivo Piantanida und Waltraud Kuchelmeister danken.

Ein herzliches Dankeschön geht an meine Eltern, Guido und Waltraud, die mich in meinem eingeschlagenen Lebensweg immer ermutigten und an meine Schwestern, Lena und Sophie, für die Ablenkung und die gute Zeit an Ostern als letztes Durchatmen vor dem Endspurt.

Der größte Dank geht an Chris, die mich in all den Jahren begleitet hat und auch in den schwierigen Momenten meiner Arbeit immer für mich da war. Danke für all deine Mühen, dein Verständnis und die selbstlose Unterstützung.

„Der Mensch hat dreierlei Wege klug zu handeln:

durch Nachdenken ist der edelste,
durch Nachahmen der einfachste,
durch Erfahrung der bitterste.“

Konfuzius

TABLE OF CONTENTS

1. INTRODUCTION

1.1 MOLECULAR RECOGNITION OF ANIONS

The recognition of anion plays a crucial role in a manifold of essential biological processes like enzyme-substrate interactions,[1] protein folding,[2] or DNA regulation.[3] The overwhelming majority of enzyme substrates for instance is of anionic nature such as carboxylates in the citric acid cycle or nucleotides for DNA polymerization (see Figure 1).[4] Hence, the importance of anion recognition processes in biological systems can hardly be overestimated. To gain a more detailed insight into these complex processes, which are up to this date not entirely understood in all their complexity, the development of synthetic model systems is worthwhile.

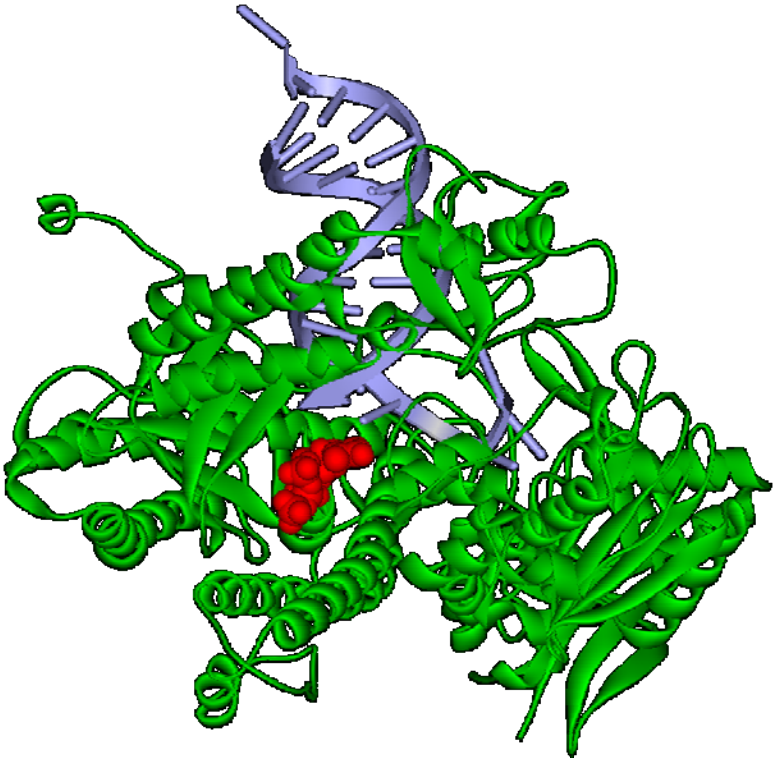


Figure 1. *X-ray crystal structure of a ternary complex between a DNA polymerase (green), a negatively charged double-stranded oligonucleotide helix (grey) and the anionic substrate thymidine triphosphate (red) (PDB access code: 3QEP[5]).*

In order to create artificial model systems, which are similarly selective and efficient as their natural archetypes, a detailled comprehension of the interactions between host and guest on the molecular level is indispensable. Such molecular recognition events are based on the combination of various attractive non-covalent interactions between receptor and substrate like electrostatic, dipole and dispersion interactions, π-stacking, and hydrogen bonding together with entropic contributions like for example the liberation of solvent molecules.[6]

While the selective recognition of cations by artificial receptor molecules was readily achieved by synthetic chemists, for example by binding alkaline metal ions with the crown ethers developed by *Pederson* as early as 1967,[7] the molecular recognition of anions has proven to be a somewhat more demanding task. Due to additional electron-electron repulsion, anions are larger in size than their corresponding neutral species and even larger in comparison to respective cations. Therefore, anion receptor systems need to be comparably larger to allow for efficient interaction with their target atom or molecule, thus increasing their complexity. Furthermore, even simple inorganic anions occur in a multitude of geometries like spherical halides, tetrahedral phosphates, planar nitrates, linear azides or in even more complex forms (e.g. oligophosphates). The substrate´s geometry has to be matched by the artificial host system for efficient and selective binding. Furthermore, anions feature a higher free energy of solution in water.[8] Consequently anion receptors have to compete more strongly with the surrounding aqueous medium for binding to their substrates and thereby replacing the tightly bound water molecules. Finally, organic and also some inorganic anions often only exist within a rather small pH window which necessitates the receptors to be in an appropriate protonation state at this pH range to allow for efficient molecular recognition.

In natural systems the amino acid arginine is very often involved in binding events between anionic substrates and proteins.[9] The energetic stabilization of carboxylates in the microenvironment of a protein by the guanidinium moiety in an arginine side chain for example outmatches the stabilization by the ammonium group of lysine by large because guanidinium is able to form a very stable bidentate hydrogen bonded salt bridge to the carboxylate.[10] Furthermore, the basicity of the guanidine group ensures protonation over a broad range of pH values due to the delocalization of the positive charge over three nitrogen atoms, thus guaranteeing for charge-charge interactions and the capability to act as efficient hydrogen bond donor.[11]

These valuable features are the reason why chemists have regularly and successfully designed guanidinium-based artificial host systems for anion recognition.[12] Likewise in our working group, a guanidiniocarbonyl pyrrole (GCP) motif was developed for efficient binding of carboxylates, which can be equipped with a linear peptidic side chain for the formation of additional interactions with the substrate.[13] Within the framework of the work presented in this thesis, novel receptors containing the GCP binding motif in one to three peptidic side chains will be developed for the recognition of biologically relevant targets like oligopeptides, nucleotides, or nucleic acids. In order to selectively and efficiently bind to these polar substrates without the help of metal ions in aqueous medium at physiological pH these new host molecules need to overcome the competing influence of the surrounding water molecules by means of multiple non-covalent interactions. From these model systems one may then deduce general information on the binding events and proceed by applying the so obtained knowledge. One potential application might for instance lie within the field of gene therapy.

1.2 GENE THERAPY

Gene therapy can be defined as the treatment of human diseases by the transfer of genetic material into specific cells of the patients.[14] Advances in biotechnology paired with the decoding of the human genome lead to the identification of disease-causing genes.[15] These findings set the road for treating inherited or acquired genetic defects by introducing the

missing genetic information, replacing corrupted ones, or by deleting deleterious genes. This way the limitations of small molecule therapies can be overcome and yet non-druggable diseases are rendered curable. Therefore, the genetic information has to be delivered into the desired cell and implemented into the cell´s own genetic code. Different from the majority of classical small molecule therapeutics,[16] nucleic acids are not able to enter the cell on their own. They are repelled from the negatively charged cell surface due to the negative charge of their ribophosphate backbone. Furthermore, they are rapidly degraded in the human body by DNAses or RNases. To surmount these obstacles transport systems, so called vectors, are necessary. Nowadays the most frequently utilized gene carriers are recombinant viruses.[17] The harmful parts of the virus´ RNA or DNA, e.g. their ability to replicate, are removed and replaced with the desired genetic information. Viruses are extremely efficient in delivering their genetic material into human cells, because this is the sole cause of their existence and ensures their survival. They were able to perfect this mechanism over millions of years. However, the gene transfer with designer viral vectors suffers from several severe drawbacks. First, the immune system recognizes the viral vectors as pathogens and reacts accordingly. This reaction is getting more severe with every treatment. Further, recombinant viruses are able to retrieve the removed harmful information by recombination with wild-type viruses and may thus become pathogenic again and spread. Another inconvenience is that the length of the therapeutic gene is limited to the size of the viral vector´s own genome. Acceptance in public is not very good. Finally, the scale-up for industrial production on large scale is very difficult to achieve for these biological systems.

Due to these inconveniences alternatives are necessary. Chemical vectors are the most promising ones because they offer the potential to overcome most of the difficulties associated with viral gene carriers.[18] These non-viral vectors are positively charged molecules which are able to condense the DNA into small, positively charged particles which are then able to cross the cell membrane. This way, the nucleic acid is also shielded from degrading enzymes and its half life in human serum is increased. Chemical vectors are not pathogenic and do not trigger an immune response. They can easily be produced in high quantities at low cost. However, there are still challenges which need to be addressed for efficient and safe delivery of genetic information by means of chemical vectors to enable application in gene therapy. Chemical vectors are not yet as efficient as viral vectors and due to their high positive charge they are often cytotoxic. Furthermore, they are not selective for a given cell type. Therefore, the development of novel chemical vectors with improved properties is necessary in order to overcome one of the major limitations of gene therapy – the transport of genetic material into the cell. Within this work novel polynucleotide-ligands containing the GCP oxoanion binding motif combined with peptidic side chains will be developed and tested for their ability to shuttle genetic information into human cells. The results of these experiments will improve our knowledge about the interactions between small molecules with nucleic acids and help to progress the highly active and still developing field of non-viral vectors for the delivery of genetic material.

2. BACKGROUND INFORMATION

2.1 CARBOXYLATE- AND OLIGOPEPTIDE RECOGNITION

2.1.1 The GCP-Oxoanion Binding Site

The binding motif **1** depicted in Figure 1 nukleotid enzyme dna ternary komplex

Figure 2 was introduced by *Schmuck* in 1999.[19] It consists of a relatively rigid, planar guanidiniocarbonyl pyrrole moiety, which allows efficient binding of planar oxoanions, e.g. carboxylates (**2**), in aqueous medium. In addition to the salt bridge between the guanidinio group and the carboxylate the complex is stabilized by several well-defined hydrogen bonds. The guanidinium group is connected to the pyrrole core by an acetyl group, thus decreasing the pK_A value of free guanidine from 13.5 to approximately 6-7.[20] As a consequence the potency to act as a hydrogen donor increases dramatically. Furthermore, the intramolecular hydrogen bond between the carboxyl and guanidinium group preorganizes the binding motif in the correct conformation for oxoanion binding. Finally, by varying R´ the receptor can be fine tuned for a given substrate.

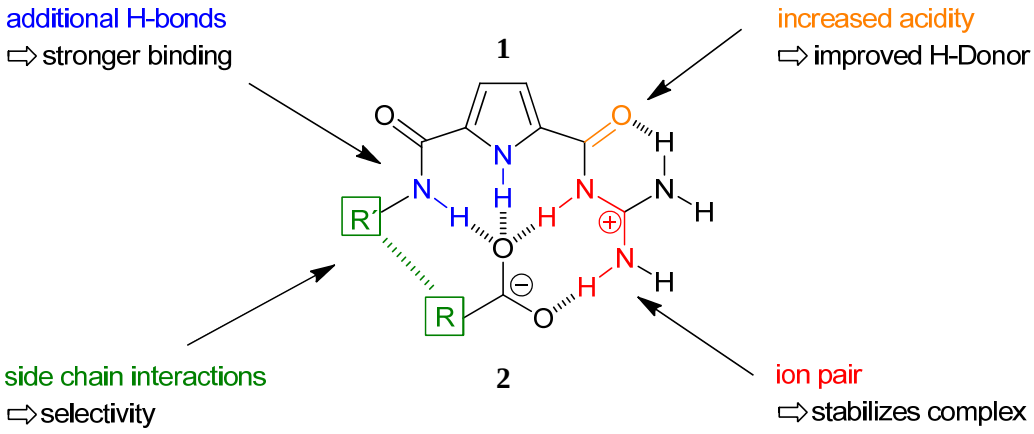


Figure 2. *Guanidiniocarbonyl pyrrole receptor **1** for the recognition of carboxylates (**2**).*

This way binding constants between $K \approx 10^3$ and 10^4 M^{-1} could be obtained in 10 % DMSO/water for N-Ac-Ala-OH.[21] This is a very good value when considering that water is an extremely competitive medium with regard to the recognition of polar substrates. Thus, by combining several weak, non-covalent interactions it was possible to devise artificial host systems for effective and selective binding of substrates in aqueous medium.

2.1.2 Recognition of Short Peptides with Linear Receptors

Following *Lehn´*s description the term "receptor" will be used throughout this work as a chemical host which binds to a given guest.[22] According to this definition, peptide receptors are synthetic organic molecules which selectively bind to a peptide sequence via multiple non-covalent interactions. The molecular recognition of short peptide sequences is a promising goal due to their importance in biochemical and medicinal processes.[23] Artificial model systems may not only contribute to gain a more substantial understanding of the yet not fully understood molecular recognition process itself, they may also serve as new starting points for drug development[24] or as sensors for diagnostics.[25] A prerequisite for potential application is always a strong and selective complexation of the target peptide sequence.

A well studied example for a short peptide sequence of biological importance is the amino acid sequence D-Ala-D-Ala-OH (**3**) and its depsipeptide analogue D-Ala-D-Lac-OH (**4**), which are both shown in Figure 3. The first one is crucial for the mode of action of the antibiotic *Vancomycin* (**5**) against *Gram* positive bacteria: *Vancomycin* binds to this sequence with high affinity ($K = 2 \times 10^5$ M^{-1}) thereby sterically blocking a transpeptidase enzyme from cross-linking the peptide side chains of peptidoglycan strands, which act as precursors during bacterial cell wall biosynthesis. This leads to a decrease of mechanical cell wall stability and ultimately to lysis upon osmotic pressure changes. The depsipeptide **4** can be found in *Vancomycin* resistant bacteria: By exchanging an amide for an ester bond the complex stability is reduced by a factor of 1,000. This dramatic decrease is due to the loss of one hydrogen bond which is instead replaced by an electrostatic repulsion between the oxygen lone pairs of the ester and the corresponding carbonyl group in the antibiotic.[26]

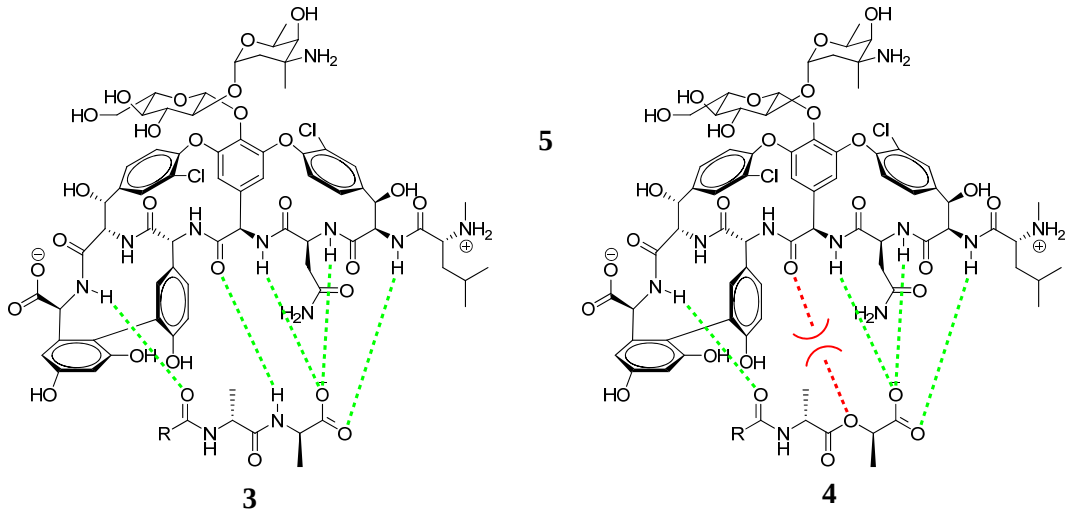
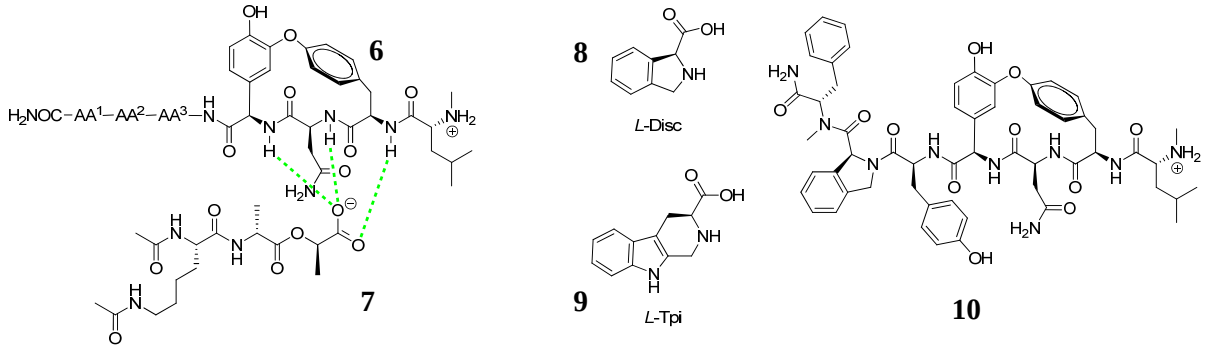


Figure 3. *Comparison of interactions between Vancomycin (**5**) with normal (**3**) and mutated (**4**) bacterial cell wall-forming peptide sequence. An attractive hydrogen bond is replaced for a repulsive electrostatic interaction thus decreasing the affinity of the antibiotic by a factor of 1,000.*

With the help of the combinatorial library **6** shown in Figure 4 comprising 39,304 members (34^3 proteinogenic and non-natural amino acids) *Ellman* was able to identify receptors which bind to the depsipeptide sequence N,N´-Ac_2-L-Lys-D-Ala-D-Lac-OH (**7**) in aqueous solution.[27] The design of *Ellman´*s receptor mimics that of *Vancomycin*: The right part corresponds to a simplified carboxylate binding pocket, which should be able to bind to both D-Ala-OH and

D-Lac-OH via three hydrogen bonds. The left side of the receptor on the other hand consists of a variable tripeptide, which is free to rotate and should therefore be able to avoid electrostatic repulsion between the antibiotic´s carbonyl group and the lactate oxygen. And indeed, an on-bead screening with fluorescent labeled substrate revealed binding constants five times as high as for *Vancomycin*. The best receptors mainly contained the amino acid sequence L-Tpi-L-His (29 %) and L-N-Me-Phe-L-Disc (20 %) in position AA^1 and AA^2 (L-Disc **8** and L-Tpi **9** are shown in Figure 4). For AA^3 no clear tendency could be observed. To determine the binding constants and to validate the results from the on-bead screening, the best performing sequences were synthesized and tested in solution for their binding affinity towards N,N´-Ac_2-L-Lys-D-Ala-D-Lac-OH (**7**) with the help of microcalorimetry. The overall best receptor **10** with the sequence L-N-Me-Phe-L-Disc-L-Tyr has a binding constant of $3 \times 10^4\ M^{-1}$ in water. For the recognition of peptides in aqueous solution this is a good result. However, the substrate is fully acetylated and therefore rather non-polar. This facilitates binding based on van-der-Waals interactions, but also is significantly different from the actual biological relevant substance.



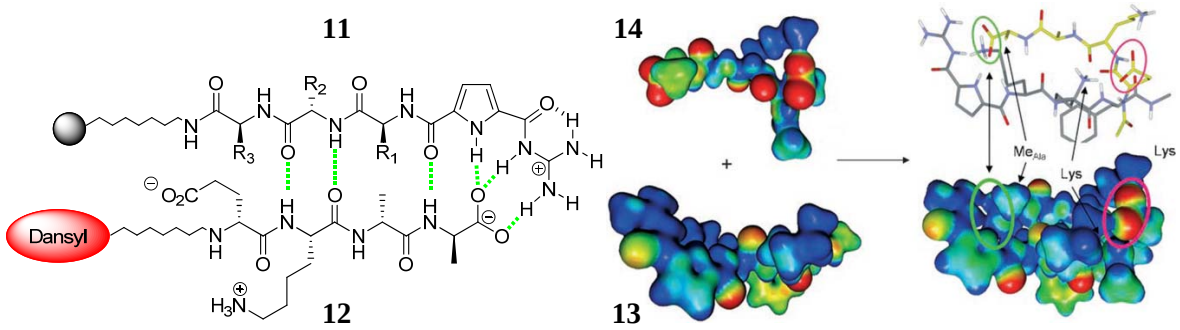
AA^{1-3} = L/D: Arg, Asn, His, Leu, N-Me-Leu, Lys, Phe, N-Me-Phe, Pro, Thr, Trp, Tyr, Tic
L: Aib, Chg, Atc, Dapa, Disc, Isn, Tpi, N-Me-Gly

$34 \times 34 \times 34 = 39{,}304$ receptors

Figure 4. *Schematic representation of receptor library **6** for the substrate N,N´-Ac_2-L-Lys-D-Ala-D-Lac-OH (**7**) (left) and receptor of highest affinity **10** with the amino acid sequence L-N-Me-Phe-L-Disc-L-Tyr (right).*

As depicted in Figure 5 *Schmuck* was able to identify receptors for the dansyl-labeled polar, anionic peptide sequence N-Ac-D-Glu-L-Lys-D-Ala-D-Ala-OH (EKAA, **12**) with binding constants of up to $K \approx 10^4\ M^{-1}$ in buffered water at pH 6 by means of screening a combinatorial library with 512 peptide-based artificial receptors (**11**) containing GCP as headgroup.[13] The best artificial receptor corresponds to the sequence $R_1 = R_2$ = lysine and R_3 = phenylalanine side chains (**13**). The C-terminus of the peptidic substrate **12** is bound by the GCP unit – this is the main driving force for the complex formation. Additionally, the attached linear tripeptide chain interacts with the substrate by forming a hydrogen bond mediated β-sheet with the peptidic substrate. An additional salt bridge is formed between glutamate and the lysine ammonium group. In order to verify the substrate selectivity the library was also screened against the inverse substrate sequence N-Ac-D-Ala-D-Ala-L-Lys-D-Glu-OH (AAKE), which resulted in binding constants that were lower by approximately one order of magnitude.[28] Thus, despite the fact that the complex stability is mainly based on undirected, long-range electrostatic interactions the binding constants are strongly dependent

on the amino acid sequence of the substrate. For binding of a polar substrate in water the so achieved binding constants are very good.



AA^{1-3} = Lys, Tyr, Ser, Glu, Phe, Val, Leu, Trp

8 × 8 × 8 = 512 receptors

Figure 5. *General scheme of the tripeptide receptor library* ***11*** *and the dansyl-labeled substrate EKAA* ***(12)*** *(left) and calculated electrostatic surface potential (right), which shows the charge complementarity between receptor* ***13*** *(GCP-KKF) and tetrapeptide* ***14*** *(EKAA). The molecular modeling images are reprinted with permission from Schmuck, C.; Heil, M. Chem. Eur. J.* ***2006****, 12, 1339-1348. Copyright 2006 John Wiley and Sons.*[28]

In another set of experiments the best receptor for the EKAA sequence (CBS-KKF) was tested against a rather small combinatorial tetrapeptide library with 320 members.[29] With only three amino acids present in the library (Ala, Lys and Glu) all substrates were closely related, only differing in the absolute configuration of one amino acid (D/L-Ala). Binding constants were determined via fluorescence in buffered water at pH 6.1 and ranged from $K < 50\ M^{-1}$ to $3 \times 10^4\ M^{-1}$. More importantly, a sequence-dependent stereoselectivity of the receptor could be observed as depicted in Figure 6 when the D/L-Ala amino acid of the substrate was fixed by strong electrostatic interactions between host and guest at both sides.

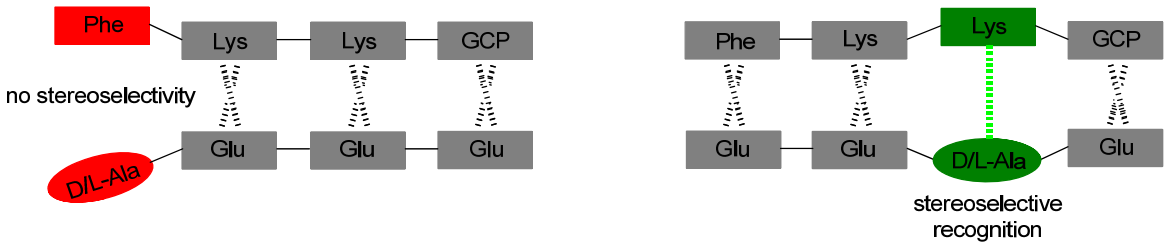


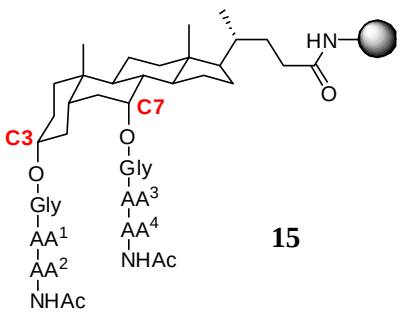
Figure 6. *Screening of a tetrapeptide library revealed that stereoselectivity is observed at positions where receptor and substrate are locked by strong electrostatic interactions at both sides of the stereogenic center (right). When this is not the case no stereoselectivity can be witnessed (left).*

These experiments clearly demonstrate the potential of small but focused libraries, which allow determining the binding properties of all members. A large and random library would not have elucidated the subtle differences of the binding properties. By characterizing all possible interactions it was possible to gain a more detailed insight into the binding event as would have been the case for a large and random library, in which only a few selected hit structures are isolated and characterized. Of course, due to the limitations of small libraries with regard to their structural and functional diversity, the library has to be carefully designed to give the correct answer for the addressed problem.[30]

2.1.3 Di- and Multivalent Receptors for Peptide Recognition

Unlike linear receptors (vide supra) so called tweezer receptors contain two arms which are bridged via a template. Depending on the linker structure this kind of molecules are able to adopt a preorganized, yet flexible cavity where the substrate is held like between molecular tweezers.[31] When correctly designed the second side chain is able to form additional attractive interactions to the substrate, thus increasing the complex stability and the substrate selectivity. The second arm may either be identical to the first one in symmetric tweezer receptors or completely different in non-symmetric tweezers. This scheme may of course be further expanded to receptors with three or more arms. In general the aim of such a multivalent approach is to strengthen the complex stability by means of simultaneous interactions between multiple complementary functionalities between host and guest. Optimally the binding affinity is then higher than the sum of the corresponding monovalent interactions. This approach is also widely made use of in nature, be it bacterium-cell-[32], antibody-antigen-interactions[33] or the interplay of transcription factors with multiple sites of DNA.[34] In the following, this chapter will be limited to di- and trivalent receptors which also corresponds to the overall focus of the dissertation.

Pioneering in the early 90s *Still* designed a library comprising 10,000 (10^4) peptidosteroidal receptors for enkephalin-like peptides (**15**, Figure 7).[35] The two arms are linked via a steroidal cheno-12-deoxycholic acid template and differ in their amino acid sequence. The non-symmetric substitution was achieved by making use of the different chemical reactivity of the two hydroxyl groups at position C3 and C7. With the help of a two-colors-two-substrates-assay receptors could be identified which were able to selectively bind to one of the two differently labeled substrates blue-dye-linker-L-Tyr-Gly-Gly-L-Phe-L-Leu-OH and red-dye-linker-L-Tyr-D-Ala-Gly-L-Phe-L-Leu-OH in chloroform, which only differ by one amino acid. Thus, by selecting beads which are stained by merely one color only those receptors which are selective for one of the two substrates could be isolated. When utilizing a flexible instead of the rigid steroidal backbone a weakening of binding strength by a factor of five and the complete loss of selectivity could be observed. In order to molecularly recognize the pentapeptidic substrates used in this work highly flexible receptors seem to be disadvantageous, probably due to a loss of preorganization.



15

$AA^{1\text{-}4}$ = Ala, Val, Leu, Phe, Pro, Ser, Thr, Lys, Asp, Glu

10 × 10 × 10 × 10 = 10,000 receptors

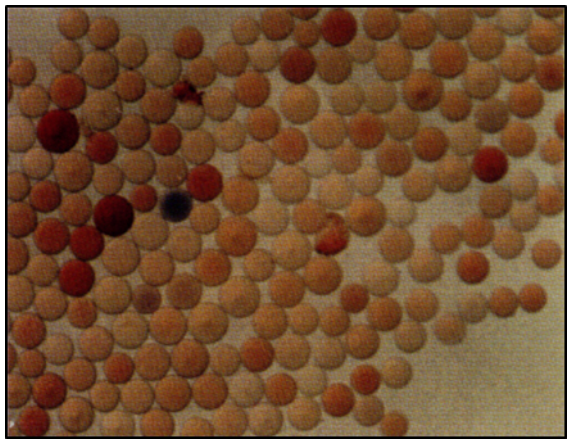


Figure 7. *The peptidosteroidal tweezer receptor library* ***15*** *with two differently substituted side chains was simultaneously screened for two substrates labeled with either a blue or a red dye. The image of the screening (right) is reprinted with permission from Boyce, R. Li, G. Nestler, H. P. Suenaga, T.; Still, W. C. J. Am. Chem. Soc.* ***1994****, 116, 7955-7956. Copyright 1994 American Chemical Society.*[35]

Wennemers synthesized five closely related, yet distinct tweezer receptors **19** which were labeled with a red dye as depicted in Figure 9. The two identical arms were linked via a rigid diamino diketopiperazine template.[36] A screening for substrate selectivity in chloroform was carried out with the help of a substrate library comprising 29^3 = 24,389 tripeptides (**18**). Binding constants of up to 10^3 M^{-1} (chloroform) and high selectivity were observed: Only one out of 5,000 sequences is bound, e.g. the receptor with the amino acid sequence L-Tyr-L-Asn(Trt)-L-Phe exclusively selects peptides containing D-His followed by two hydrophobic D-amino acids. In order to verify whether a simplified receptor design is still able to achieve effective binding, corresponding host systems with only two amino acids per side chain and one-armed analogues were screened with the same substrate library. However, as a result of the truncation the host system completely lost its binding affinity.

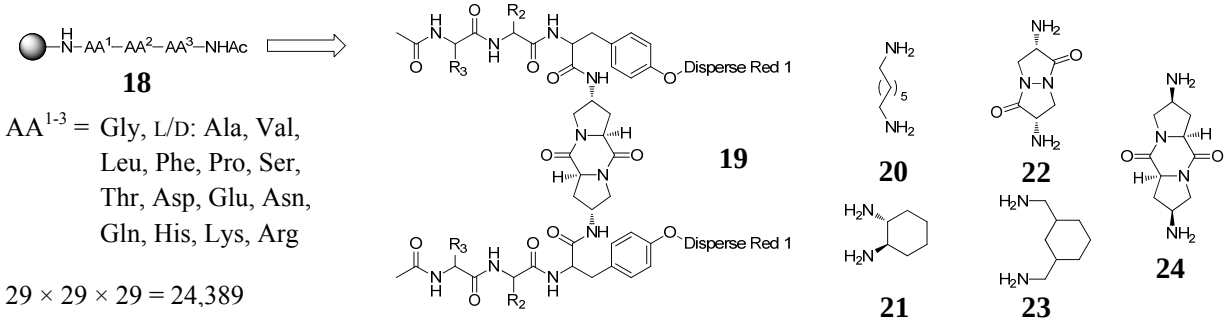


Figure 9. *Schematic design of tweezer receptors **19** containing two identical tripeptidic arms labeled with a red dye (left) and alternative templates **20-24** (right).*

Under the conditions used for the experiments **19** seems to be the minimum structure which is necessary for the recognition of the tripeptides under investigation. In order to evaluate the role of the template on the binding properties similar receptors with the same side chains but different diamine templates **20**-**24** were synthesized as well.[37] A screening against the same substrate library revealed that with only a few exceptions these analogues do not bind to the substrates at all. The U-shaped geometry of the original *trans*-*trans*-diketopiperazine template in **19** highly preorganizes the two arms in a fashion which is suitable for tripeptide binding. On the other hand side-chains which are connected via the *cis*-*cis*-diastereomer **24** adopt an almost linear geometry. Hence, a defined angle between template and side chains as well as a distance of approximately 8 Å in between (as is the case for the original diketopiperazine scaffold) seem to be superior to all other templates that were tested in this work. Accordingly, the template plays a decisive role concerning the affinity between receptor and substrate and has to be carefully chosen. However, with the help of these host systems only rather moderate affinities could be achieved. Although the selectivity is quite good, it has to be noted that the binding studies were only carried out in organic solvent and would probably suffer in more demanding aqueous surroundings.

Further studies concerning the role of the template were conducted by *Liskamp*.[38] As shown in Figure 10 tweezer receptors **25** consisting of two identical peptidosulfonamide side chains which were bridged by several templates **26**-**31**, which differ in both their flexibility and the distance between the amine groups. A dye-label enables screening of a tripeptide library containing 29^3 = 24,389 members (**18**). The overall highest binding affinity in chloroform was

found for the scaffold **26** with a binding constant for the peptide D-Ala-L-Asp-D-Ser of 4 × 10^3 M^{-1}. Since no binding affinity could be observed for a one-armed analogue it was concluded that the second arm is necessary for efficient tripeptide binding. The second best receptors with binding constants of up to 8 × 10^2 M^{-1} were those with template **27**, which is more rigid and offers less space between the two arms. Third ranked are the two scaffolds **28** and **29**. The distance between side chains seems to become too large for stable complexation of the peptidic substrates. When implementing **30** as linker, the corresponding receptor showed high selectivity (> 95 %) for the amino acid sequence Glu-His-X. However, the binding affinity was rather poor – this was again probably due to the large distance between the two arms. No affinity at all could be observed for **31**. The cause for this effect can probably be found in the repulsion between the free electron pairs of the sulfonamide oxygen groups, which would be in too close proximity in the depicted chair-chair-conformation of the template. Instead, it adopts a boat-chair-conformation which results in the loss of tweezer structure. A common feature of all examined host systems in this work is the high guest selectivity: the amino acid AA^1 almost always is an acid; AA^2 is a polar amino acid like asparagine or histidine. At position AA^3, which is bound directly to the resin, no tendency could be observed.

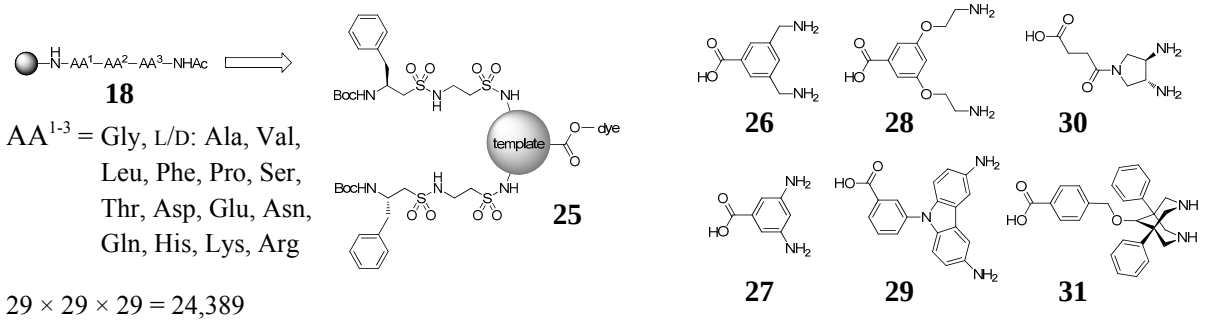


Figure 10. *Schematic representation of tweezer receptors **25** and templates **26-31**.*

The binding studies of library **25** were again conducted in organic solvents. For further studies in aqueous solution two more combinatorial tweezer libraries (**32**, Figure 11) consisting of 1,728 members each were synthesized based on the best scaffold **26**.[39] As before the two arms were identical but instead of sulfonamides they now consisted of tripeptides to facilitate the combinatorial approach. One of the two libraries contained the template without change, for the preparation of the other one the template was derivatized with a hydrocarbon chain as illustrated in Figure 11 to reduce its flexibility. Both libraries were then tested for their binding affinity towards dansyl-labeled biologically relevant substrates Gly-D-Ala-D-Ala-OH (**33**) and Gly-D-Ala-D-Lac-OH (**34**) in aqueous phosphate buffer. Both substrates were linked to the fluorescent dansyl group via a glycine spacer. Characterization of hit structures revealed a highly consistent distribution of amino acids at position 3. For the dipeptide **33** lysine was reported with 40 % and proline and alanine each with 20 %. The results from the screening for the corresponding depsipeptide **34** show that position 3 was mainly occupied by proline (50 %), followed by lysine (25 %) and alanine (20 %). The conformity was less well pronounced for positions 1 and 2. However, in both cases mainly aromatic or aliphatic amino acids were present, which might be due to an interaction of the non-polar amino acid side chains with the dansyl label. The glycine linker is probably too

short to prevent this unwanted hydrophobic interaction with the fluorescence dye. With a binding constant of 5×10^2 M^{-1} the highest affinity for the depsipeptide in aqueous phosphate buffer was found for the amino acid sequence AA^{1-3} = L-Leu-L-Phe-L-Lys. In chloroform binding constants were significantly higher with values of up to 10^4 M^{-1}. Due to the competitive influence of water the achieved binding constants are rather low.

The comparison between both libraries reveals that the hydrocarbon chain has no significant influence on the complex stability. The addition of a third arm, which was indirectly attached to the scaffold via the hydrocarbon chain (**35**), did neither improve binding affinity nor selectivity.

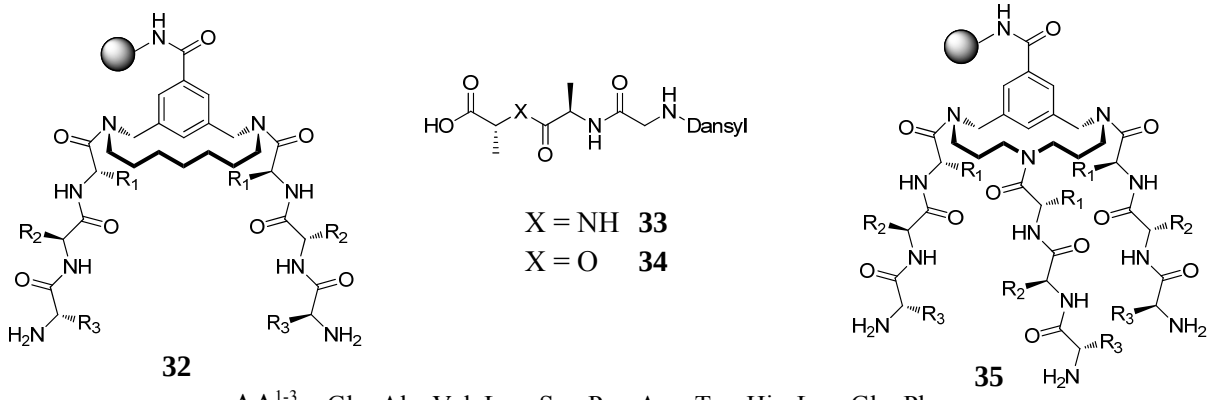


Figure 11. *Schematic illustration of receptor library **32** with an additional hydrocarbon chain. Side chains consist of three variable amino acids AA^{1-3}. The substrates D-Ala-D-Ala **33** and D-Ala-D-Lac **34** were labeled with a dansyl fluorescence marker. An additional arm attached to the hydrocarbon chain gave library **35**.*

Kilburn prepared the combinatorial tweezer receptor library **36** depicted in Figure 12 with 15,625 (5^6) members.[40] The two non-symmetrical side chains consisted of varying tripeptide sequences. A guanidinium headgroup served as both, the scaffold and as a carboxylate binding site in order to increase the binding affinity to the C-terminus of the substrate via a hydrogen bond enforced salt bridge. The positive influence of the cationic headgroup was confirmed by comparing receptors that were bound to the solid support directly via the guanidine.[41] These model systems lost all of their affinity to the substrate because the guanidine could not be protonated any more under these circumstances. An on-bead screening of **36** with the dye labeled peptide N-Ac-Lys(dye)-D-Ala-D-Ala-OH (**37**) in buffered water at pH 8.75 revealed that less than 2 % of the library members did bind to the substrate. The selected hit structure sequences were highly consistent as shown in Table 2. Additional on-bead binding studies with the best receptor of the sequence AA^{1-6} = Gly-Val-Val-Met-His-Ser showed a binding constant of 10^3 M^{-1}. The corresponding diastereomeric substrate N-Ac-Lys(dye)-Ala-Ala-OH was bound less efficiently with a binding constant of 3×10^2 M^{-1}. Binding studies in solution did not lead to an analyzable result. This indicates clearly that the resin influenced the binding event on solid support. It is therefore crucial to verify results from measurements that involve resin bound substances in solution.

Table 2. *Occurrence of amino acids of selected hit structures at each of the six positions of the receptor.*

Position	AA^1	AA^2	AA^3	AA^4	AA^5	AA^6
Amino acid	Gly (30 %) Ala (30 %)	Val (40 %) Ala (40 %)	Val (40 %) Ala (40 %)	Met (50 %)	Met (40 %) His (30 %)	Ser (70 %)

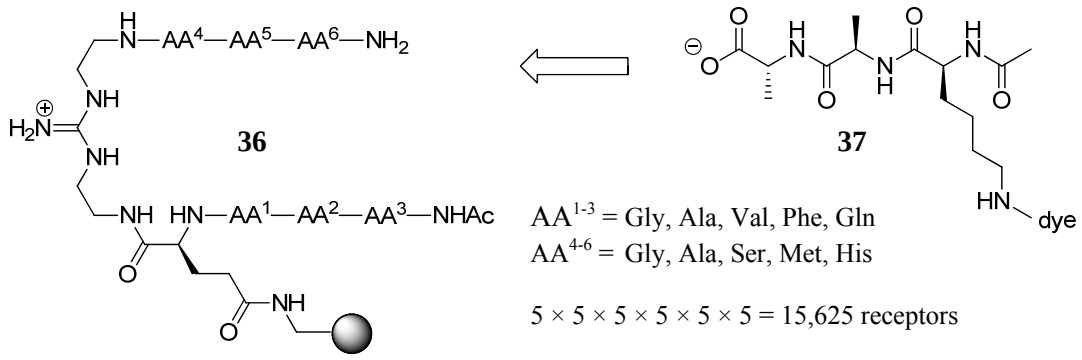


Figure 12. *Non-symmetric tweezer receptor library* ***36*** *with a guanidinium template for carboxylate recognition and dye labeled substrate N-Ac-Lys(dye)-D-Ala-D-Ala-OH* ***37****.*

In conclusion, multivalent receptors have proven to have the potential to increase both the selectivity and the affinity to a defined substrate in comparison to their one-armed analogues. By increasing the functional and structural diversity and summing up multiple non-covalent interactions the competing influence of the surrounding medium can be surmounted. This is of special importance when working in water, which is necessary when biological targets are involved. When designing such molecules several things have to be taken into account. First, the template that bridges the arms plays a decisive role for the complexation. Optimally it preorganizes the arms in a suitable fashion for substrate binding and is still flexible enough to allow for its accommodation. For peptide recognition by tweezer receptors the distance and angle between the side chains seems to be of utmost importance. Amongst the here presented scaffolds the aromatic template **26** has led to the best results. Concerning the side chains, building blocks, amino acids offer valuable resources especially when preparing combinatorial receptor libraries due to their diversity and the commercial availability. Regularly, the two arms are identical, because this way the synthesis is facilitated. Less well studied are host systems with two different arms, although this approach offers the possibility to add additional functionality. The introduction of a carboxylate binding site, like a guanidinium group, into the receptors has proven to be worthwhile for improving binding properties. Other things that should be kept in mind are that linkers and labels may also take part in the binding event as well as the solid support. Because of that it is advisable to always validate on-bead results with additional measurements in solution.

Although the presented model systems have led to a better understanding on the molecular recognition of short peptide sequences and many receptors have promising binding properties in organic solvents like chloroform most of them struggle with binding polar peptide sequences in aqueous medium at physiological pH. In order to overcome the competitive influence of water the development of tweezer receptors with non-symmetric arms and a carboxylate binding site, like e.g. *Schmuck´*s binding motif **1**, could be a promising approach.

2.2 NUCLEOTIDE RECOGNITION IN WATER

Nucleotides are amongst the most targeted anionic species for artificial host systems, because they are ubiquitously present in biological systems where they exercise key roles in many cellular functions and thus offer interesting structural features for the benchmarking of designed host systems in supramolecular chemistry.[42] Nucleotides are for example involved in DNA synthesis, energy and electron transfer events,[43] cell signaling[44] or membrane transport.[45] These are complex events which require the molecular recognition of a specific nucleotide. In order to increase our insight into these processes the development of model systems for nucleotide recognition is desirable.

Due to their structural features, nucleotides are quite interesting targets from a chemist´s point of view. As depicted in Figure 13 they consist of three main parts – nucleobase, phosphate hinge and sugar backbone. The sugar scaffold is a five-membered ribose for RNA and a 2´-deoxyribose ring for DNA building blocks. It connects the phosphate, which is attached to the 5´-hydroxy group with the nucleobase at position 1´, this way clearly defining the nucleotide´s geometry. The ribose ring itself is rather difficult to target. The hydroxy groups offer potential for hydrogen bonding. However, they are hard to differentiate from the surrounding water molecules. The hydrocarbon chain can be targeted by means of hydrophobic interactions. These are however, rather weak and unspecific. The best artificial carbohydrate receptors known today complex their targets with binding constants of merely $5 \times 10^3\ M^{-1}$ (pH 7.4, buffered water).[46] The phosphate hinge consisting of one to three phosphates is an easier target. Due to its multiple negative charges and the presence of several hydrogen bond acceptors it may be addressed via hydrogen bonds and salt bridges. In comparison to afore mentioned carboxylates, phosphate anions are of tetrahedral shape and therefore require a different kind of binding motif. The nucleobases are either purine- (A, G) or pyrimidine-based (C, U, T). Their aromatic nature allows for π-stacking while their heterocyclic nature and the exocyclic carbonyl and amine functions offer potential hydrogen bonding acceptor and donor sites.

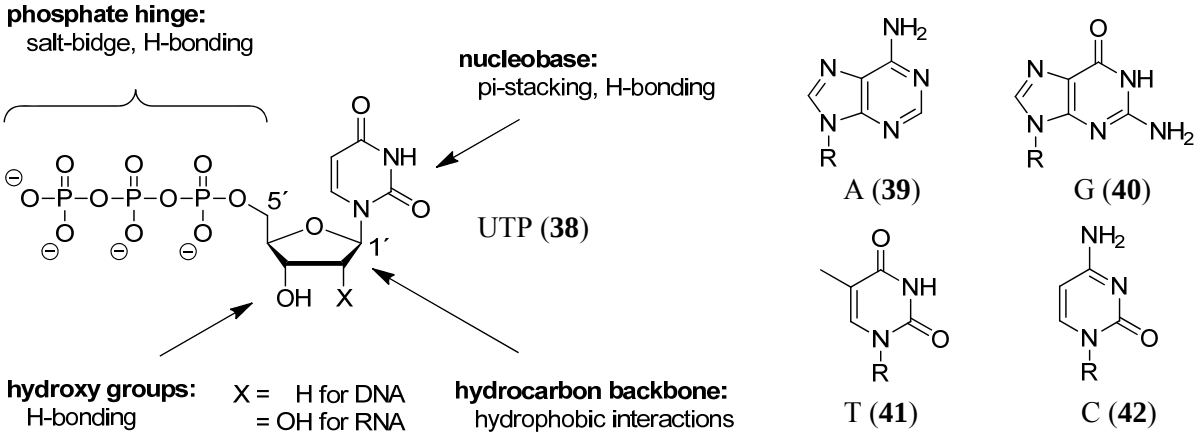


Figure 13. *Uridine triphosphate **38**, adenine (**39**), guanine (**40**), thymine (**41**) and cytosine (**42**).*

Due to the huge number of nucleotide receptors reported in literature this chapter will only focus on systems which have been used in aqueous medium.

2.2.1 Polyamine-Based Receptors

In 1977 *Nakai* could show that the naturally occurring polyamines shown in Figure 14 (top) putrescine (**43**), spermidine (**44**), and spermine (**45**) are able to bind to nucleotides.[47] Binding constants were determined by means of an anion-exchange resin method in buffered water at pH 7.5. While putrescine only forms a weak 1:1 complex with AMP (K = 82 M^{-1}), spermidine is already able to form somewhat stronger 1:1 complexes with ADP (K = 10^3 M^{-1}). Finally, spermine forms even stronger 1:1 complexes to ATP with a binding constant of 10^4 M^{-1}. Obviously the complex formation is becoming stronger the more charge-charge interactions are established.

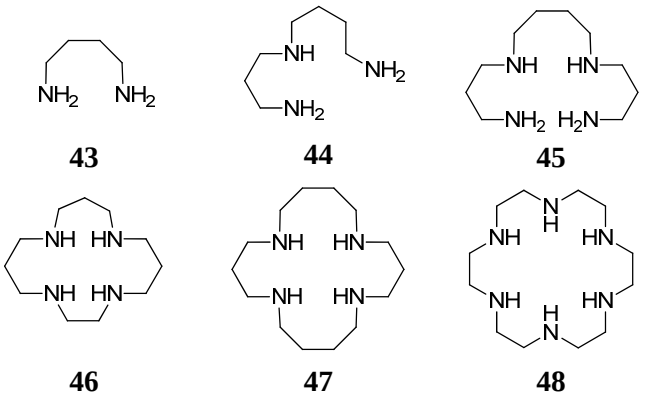
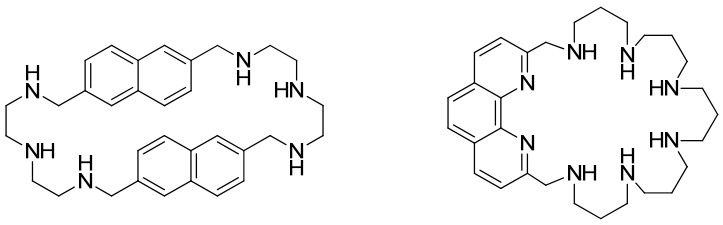


Figure 14. *Linear polyamines putrescine **43**, spermidine **44** and spermine **45** (top) and exemplary macrocyclic polyamine receptors **46-48** (bottom).*

Not surprisingly, the first artificial nucleotide receptors developed by *Kimura* were 15- to 18-members macrocyclic polyamines – some exemplary representatives are shown in Figure 14 (**46**-**48**).[48] By studying their binding properties towards phosphate, AMP, ADP and ATP via polarographic and NMR measurements in buffered water at neutral pH, binding constants of up to 4 × 10^6 M^{-1} could be observed. Again, ATP was bound strongest followed by ADP and AMP. Phosphate on its own was bound weaker than AMP by a factor 10-100, although their charge is the same. This effect was attributed to additional hydrogen bonds between the nucleobase and the receptor as indicated by NMR. By comparing spermine with its cyclic analogue **47** it could also be shown that the macrocycle bound stronger to nucleotides by approximately two orders of magnitude. Furthermore, only polyamines which are capable to incorporate at least three charges at neutral pH formed stable complexes.

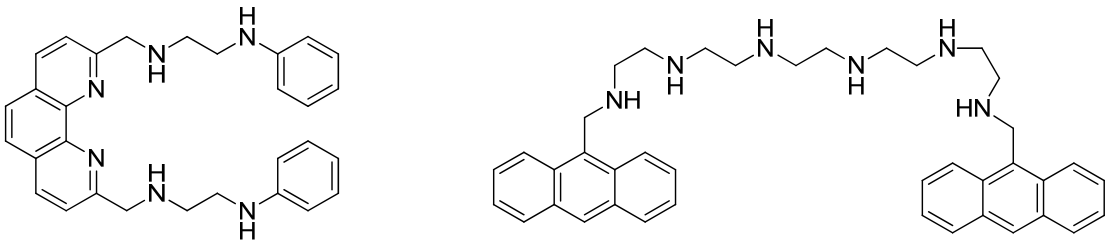
The next generation of receptors incorporated aromatic rings into the macrocycle. *Lehn* developed the nucleotide host **49** depicted in Figure 15, which is able to bind to nucleotides in buffered water at pH 6 with binding constants of up to 2 × 10^5 M^{-1} (ATP, NMR titration).[49] The lower charged AMP is bound more weakly by one order of magnitude. Interestingly the purine-derived nucleotides AMP and GMP are bound stronger than the smaller pyrimidine analogues CMP and UMP by a factor of 10. Lacking the phosphate hinge, the corresponding nucleosides are bound weaker by 2-3 orders of magnitude. Thus it can be concluded that both the stacking between nucleobase and aromatic scaffold as well as charge-charge interactions between the polyammonium chains and the phosphate part of the nucleotide are essential for efficient binding.



49 **50**

Figure 15. *Macrocyclic polyamine receptor **49** with two naphthalene bridges (left) and phenanthroline linked receptor **50** (right).*

Bencini developed the macrocyclic receptor **50** (Figure 15) with a phenanthroline unit inside the polyamine macrocycle.[50] With the help of potentiometric measurements the highest affinity was found for ATP with an extremely high binding constant of 3×10^9 M^{-1}. The corresponding nucleotides TTP, CTP and GTP were bound less well by 1-2 orders of magnitude. Furthermore, the artificial host could selectively sense ATP thanks to fluorescence quenching caused by a photoinduced electron transfer from an amine group of the receptor to the excited phenanthroline. Concerning the complex geometry it could be shown by means of NMR experiments, molecular modeling and a crystal structure of one of the complexes that the polyamine chain forms a cavity for the phosphate enforced by ion pairs and hydrogen bonds while the nucleobase stacks to the phenanthroline moiety. A macrocycle analogue with reduced ring size and one amine group less lost a huge part of its affinity, i.e. the binding constant to ATP was only 10^6 M^{-1} – lower by a factor of 3,000. These results demonstrate the importance of an appropriate mix of attractive interactions with the correct geometric prerequisites needed for efficient recognition of a given substrate molecule.



51 **52**

Figure 16. *Tweezer receptor **51** with two polyamine side chains linked via a phenanthroline template (left) and receptor **52** with two anthracene units bridged by a polyamine chain.*

Lin designed the tweezer receptor **51** (Figure 16) with two polyamine side chains linked via a phenanthroline scaffold; each arm incorporates a phenyl headgroup.[51] The binding constant for ATP was measured via potentiometry to an excellent 7.9×10^{10} M^{-1}. As expected, NMR and molecular modeling revealed charge-charge interactions between phosphate and polyammonium side chains and π-stacking between phenanthroline and phenyl groups with the nucleobase. These results prove that in order to achieve high binding affinity it is not necessary to make use of cyclic systems: Essentially more important for the molecular recognition event is the combination of several attractive non-covalent interactions in the correct molecular topology for binding.

A similar system (**52**), also depicted in Figure 16 was developed by *Garcia-España.* Two anthracene moieties are linked by a polyamine chain.[52] As determined by potentiometry the

binding constant for ATP was also similarly high with 10^9 M^{-1}. ADP and ATP were bound consecutively worse by 2 and respectively 3 orders of magnitude. Furthermore, in this work also the effective binding constant for ATP at neutral pH was determined to approximately 5×10^5 M^{-1}. This value was confirmed with the help of a fluorescence titration. It is noteworthy that the effective binding constant is lower by more than three orders of magnitude compared to the maximal possible one reported by potentiometric measurements.

At this point it is important to state that great care has to be taken when comparing binding constants. Potentiometric measurements merely represent the binding constant of host and guest in a clearly defined protonation state. However, at a given pH there is always an equilibrium of many overlapping states. Consequently these *K* values only refer to single theoretical protonation states and are therefore not valid for the macroscopic ensemble but represent rather the maximum possible binding constant between two single species in the mixture. Without the calculation of an effective binding constant these data cannot be used to compare binding constants of different host systems. In above mentioned examples which made use of potentiometry the effective binding constant was never calculated with the exception of the work conducted by *Garcia-España*. This is also why many of the extremely high binding constants must be put into perspective – the effective binding constants are likely to be lower by several orders of magnitude. A more appropriate way to obtain comparable binding constants is to utilize methods like NMR, UV/Vis, fluorescence or ITC. Whatever the choice may be, it is always strongly advisable to use at least one alternative method to verify ones results.

2.2.2 Guanidinium-Based Receptors

A statistical evaluation of 3,003 X-ray crystal structures of phosphate binding proteins revealed that two third of them do not make use of metal ions for phosphate binding.[53] More than half of the metal-free proteins use lysine or arginine instead, especially those with phosphate binding sites located on their surface. Consequently, ammonium and guanidinium moieties seem to be well suited to bind to phosphate in natural systems, even when directly opposed to the competitive influence of the surrounding water molecules and counter ions. While most of the receptors reported in literature so far make use of ammonium groups in analogy to lysine, only very few host systems have been designed yet which utilize the corresponding arginine analogues – guanidinium moieties – for phosphate recognition.

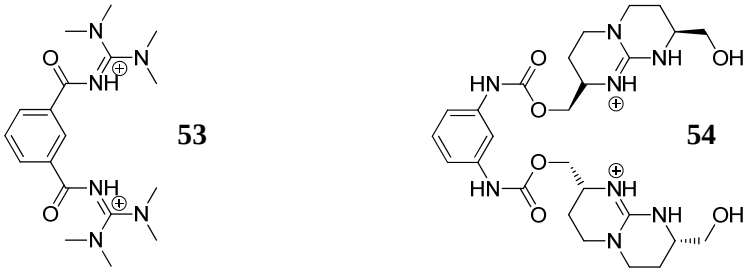


Figure 17. *Tweezer receptor **53** for AMP with two tetramethylguanidinio groups (left) and ditopic receptor **54** for TMP with two bicyclic, chiral guanidinium moieties (right).*

Wang reported the ditopic receptor **53** with two tetramethylguanidinio moieties linked via an aromatic scaffold (Figure 17).[54] With the help of an NMR titration in D_2O he could show that already this rather simple molecule is able to bind AMP with $K = 7 \times 10^4\ M^{-1}$.

In a similar approach *Schmidtchen* developed the tweezer **54** depicted in Figure 17.[55] Two chiral, bicyclic guanidinium moieties were connected via an aromatic template. With the help of NMR titrations in water a very high binding constant of $10^6\ M^{-1}$ to TMP could be observed, which was explained by a perfect preorganization of the two guanidine units perpendicular to each other. This way the geometry is perfectly matched to bind to the tetrahedral phosphate anion via two hydrogen bonded salt bridges. Once more, preorganization and structural complementarity proofed to be the key to high affinity. For phosphate binding especially ditopic receptors with two anion binding groups seem to be perfectly suitable.

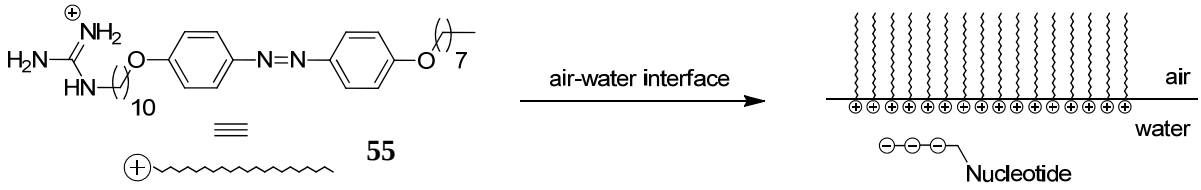


Figure 18. *The amphiphilic receptor **55** combines a long hydrophobic chain with a positively charged guanidinium headgroup. At the air-water interface the positively charged guanidinium moieties point towards the aqueous phase while the hydrophobic part points towards the air.*

The most successful approach towards guanidinium-based nucleotide recognition so far was *Kunitake´*s receptor **55** shown in Figure 18.[56] By attaching a long, lipophilic chain to a

positively charged guanidinium head group amphiphilic host systems were obtained, which were able to form micelles or bilayers in aqueous systems. The binding properties towards adenine-based nucleotides were characterized via ultrafiltration. With 2×10^7 M^{-1} the highest binding constant was observed at an air-water interface in buffered water at pH 7 for ATP; AMP was bound more weakly by one order of magnitude. The results of this approach stress the importance of the microenvironment on binding events. *Kunitake´*s system is actually a simple guanidinium moiety which is set into a defined microenvironment with the help of a long amphiphilic side chain, yet very good binding constants could be achieved this way.

2.2.3 Calixarene- and Cyclodextrin-Based Receptors

The focus of most of above mentioned nucleotide receptors clearly lies within charge-charge interactions to their substrate, sometimes accompanied by stacking between nucleobase and one or more aromatic moieties of the host systems. While π-stacking is a powerful tool for molecular recognition, especially in aqueous medium, it is not the only possibility to make use of hydrophobic interactions in water. The following examples put more focus on this kind of interaction, which may contribute substantial amounts to the Gibbs free energy in water.

Diederich for example made use of a bowl-shaped calixarene with a resorcin[4]arene scaffold and four phenylamidinium groups at the upper rim for his nucleotide receptor **56** (Figure 19).[57] Calixarenes form hydrophobic cavities in aqueous medium for the accommodation of non-polar guest molecules. NMR titrations in buffered D_2O at pH 8.3 revealed affinities corresponding to the nucleotide guest´s charge (ATP > ADP > AMP > cAMP) and binding constants for the 1:1 host-guest complexes of up to 7×10^5 M^{-1}. AMP was slightly preferred over the G, C, T, and U derivatives by a factor of 2-3. As derived from NMR and molecular modeling the nucleobase is bound within the hydrophobic bowl and the phosphate interacts with two phenylamidinium groups via hydrogen bonded salt bridges. The preference for A-nucleotides can be ascribed to a superior stereoelectronic complementarity of this nucleobase with regard to the inclusion into the bowl-type cavity.

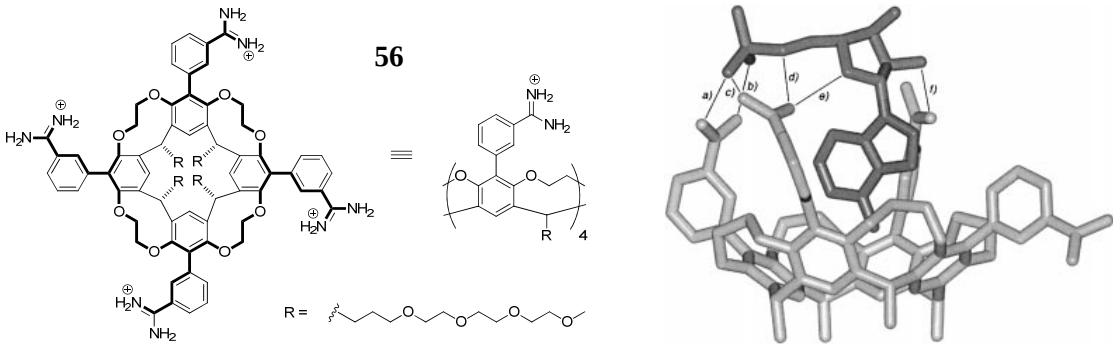


Figure 19. *Calixarene-based receptor **56** with four phenylamidinium groups (left) and calculated energy-minimized structure of the complex with AMP (right). The hydrogen bonding pattern is marked a)-f). For simplification PEG-chains R were replaced by methyl groups. The modeling image is reprinted with permission from Sebo, L.; Diederich, F. Helv. Chim. Acta **2000**, 93-113. Copyright 2000 John Wiley and Sons.*[57]

Another sort of cavity-forming molecules was used as scaffold by *Schneider*.[58] By exchanging two or respectively all seven 6-hydroxy groups in β-cyclodextrin (CyD) for aminomethyl moieties the two nucleotide receptors **57** and **58** shown in Figure 20 were obtained. CyDs are highly water soluble, cyclic, 1,4-linked oligo-α-D-glucopyranoside units. Their topology is toroidal, with a bigger and a smaller opening which present their hydroxy groups to the solvent. Although the interior of the cavity is not hydrophobic, it is less hydrophilic than water. **58** forms strong nucleotide-complexes with binding constants of up to 3×10^6 M^{-1} with ATP (effective binding constant obtained by potentiometry, verified via NMR titration). This is stronger by more than one order of magnitude compared to AMP. However, this host system is not able to differentiate between nucleobases or ribose and deoxyribose sugar moieties. **57** on the other hand forms weaker complexes (e.g. 10^5 M^{-1} for ATP) but is able to differentiate both between nucleobase and sugar.

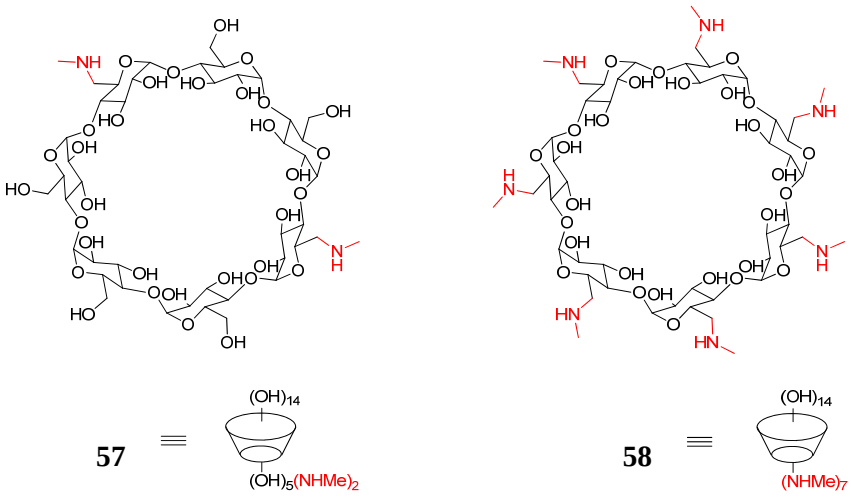


Figure 20. *Cyclodextrin-based receptors* ***57*** *and* ***58*** *with two and respectively three methylammonium groups.*

With the help of knock-out analogues (e.g. ribose, phosphoribose, phosphate) the different contributions to the overall binding constant could be assigned. The main driving force stems from ionic interactions between phosphate and methylammonium groups. Surprisingly, the nucleobase does not contribute in an advantageous way to the binding. Instead, the second most important contribution to the binding strength stems from interactions with the sugar moiety. With the help of NMR and molecular modeling it could be shown that the ribose is inside the CyD cavity and interacts with it via hydrogen bonds. This is also why the deoxyribose is bound weaker than ribose although it is more hydrophobic. The nucleobase-selectivity of **57** was explained by secondary, rather weak hydrogen bonds between nucleobase and upper rim hydroxyl groups. The higher charged **58** pulls the nucleotide deeper into the cavity and thus prevents these secondary interactions from forming. Again, these experiments are excellent examples for the difficult task to design receptors which have high selectivity as well as high affinity to a given substrate. All too often these two desirable properties are antipodal. Furthermore *Schneider´*s work is an instructive example on how model systems help to improve our understanding of anion recognition.

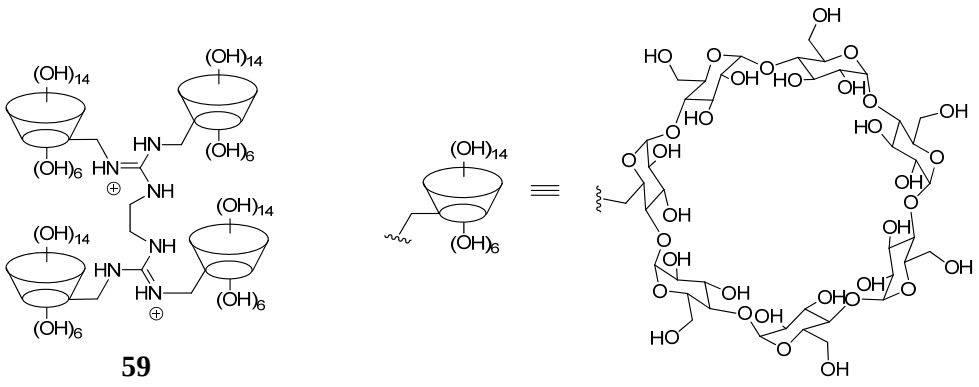


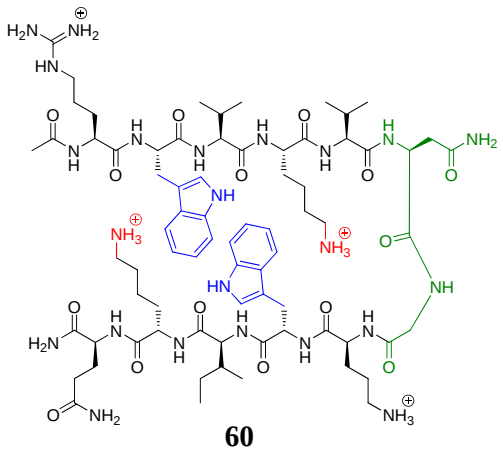
Figure 21. *Receptor* ***59*** *combines four cyclodextrin units with two guanidinium scaffolds.*

Marsura makes use of the same principle by bridging four β-cyclodextrin cavities via two guanidinium groups to receptor **59** as depicted in Figure 21.[59] The cationic scaffolds allow for

hydrogen bonded ion pairing to nucleotide phosphates while the CyDs display pockets for hydrophobic inclusion. NMR experiments in D_2O at pH 6.5 revealed that 1:2 host-guest complexes are formed to AMP, ADP and ATP with similar binding constants of 2×10^6 M^{-2}. Two substrate molecules are bound with similar binding constants representing a non-cooperative binding mode with two independent complexation steps. NMR also revealed that the phosphate hinge interacts with the guanidinium moieties via hydrogen bond enforced salt bridges. For ATP and ADP the nucleobase as well as the sugar moiety are inside the CyD cavity; for AMP on the other hand only the nucleobase is included. This is probably due to the distance between cavity and cationic template, which is becoming too large to be bridged by AMP´s monophosphate hinge when sugar and nucleobase are inside the CyD.

2.2.4 Peptide-Based Receptors

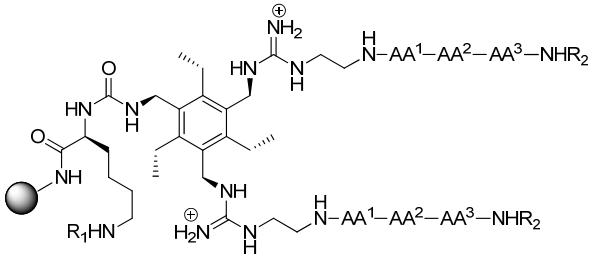
As shown for carboxylate and oligopetide recognition (vide supra) peptides have proven to be excellent building blocks for the synthesis of artificial anion receptors. Furthermore, they are excellent model systems for proteins to investigate e.g. ATP binding sites, which are commonly present in many enzymes.



60

Figure 22. *β-Hairpin receptor **60** with an Asn-Gly turn (green), a hydrophobic binding pocket composed of two Trp amino acids (blue) and two Lys moieties in close proximity for charge-charge interactions (red).*

Waters designed the 12 residue peptide-based receptor **60** depicted in Figure 22.[60] It consists of the amino acid sequence N-Ac-Arg-**Trp**-Val-**Lys**-Val-Asn-Gly-Orn-**Trp**-Ile-**Lys**-Gln-NH_2 and adopts a β-hairpin secondary structure – the turn is composed of an Asn-Gly motif. This way the aromatic Trp side chains incorporated diagonally into opposing strands form a hydrophobic pocket for the nucleobase. Adjacent to both aromatic amino acids are placed two Lys moieties in order to help develop charge-charge interactions with the phosphate. A fluorescence titration with ATP revealed 1:1 complex formation and a binding constant of 6×10^4 M^{-1} in buffered water at pH 5. ADP and AMP were bound significantly weaker by one order of magnitude. NMR experiments indicated that the nucleobase is stacked between both indole units and that the lysine ammonium groups interact with the phosphate chain.



61: R_1, R_2 = H
AA^{1-3} = Ala, Asp, Glu, Phe, Gly, His, Ile, Leu, Met, Asn, Pro, Gln, Ser, Thr, Val, Trp, Tyr

17 × 17 × 17 = 4,913 receptors

62: R_1 = coumarin
R_2 = fluoresceine
AA^{1-3} = Ser-Tyr-Ser

Figure 23. *Receptor library **61** with 4913 members and fluorescence labeled hit structure **62** for ATP binding with amino acid sequence AA^{1-3} = Ser-Tyr-Ser.*

Anslyn used the combinatorial library **61** (R_1, R_2 = H) shown in Figure 23 with 4,913 members (17^3) to obtain ATP selective receptors.[61] Two combinatorial tripeptide side chains were linked by an aromatic scaffold which preorganizes adjacent groups alternately up and down with respect to the benzene ring. As a result, a cavity is formed for substrate

accommodation. Each arm was attached to the resin via a guanidinium group in order to act as phosphate binding site. The third arm was used as connection to the solid phase and features a lysine group. The library was screened in buffered water at pH 7.1 with fluorescence labeled ATP revealing that approximately 15 % of the library members bind to ATP. Several hit structures were re-synthesized and labeled with two different dyes, fluoresceine at the N-termini of the two arms (R_2) and coumarin in the third arm (R_1, as internal standard). Binding studies in solution verified the binding properties in all but one example. This result again stresses the importance to confirm on-bead results in solution. The best receptor with the amino acid sequence AA^{1-3} = Ser-Tyr-Ser (**62**) binds to ATP with $K = 3 \times 10^3$ M^{-1} (pH 7.1, buffered water). The competing analytes AMP and GTP were bound significantly less effective. The presence of Tyr and Ser in the best receptor hints at hydrogen bonding and π-stacking between receptor and adenine and/or ribose. By attaching the fluorescence labeled receptors onto a solid support they could be used as selective sensors in arrays.[62]

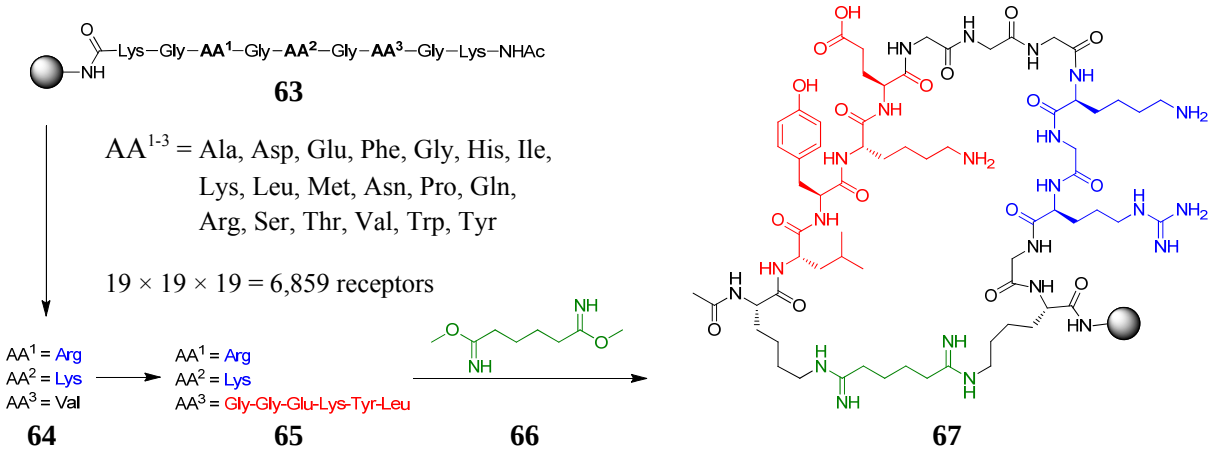


Figure 23. *Combinatorial library of 6859 (19^3) ATP receptors **63** and hit structure **64**. Rationally improved compound **65** and crosslinked ATP selective receptor **67** with phosphate (blue) and nucleobase (red) binding site. Cyclization was achieved by crosslinking the first and last Lys residue with dimethyl adipimidate (green).*

Matsui derived his ATP selective receptor **67** (Figure 24) in a three step procedure.[63] First, the nonapeptide library **63** with three combinatorial amino acid positions comprising 6,859 (19^3) members was screened with the help of fluorescence labeled ATP. The binding constant of the best receptor **64** with AA^{1-3} = Arg-Lys-Val was determined on-bead in buffered water at pH 7 to 6×10^3 M^{-1}. Molecular modeling revealed that Arg and Lys bound to the phosphate hinge, while there is no interaction with the nucleobase. Furthermore, Val does not seem to take part in the recognition process. Thus Val was replaced by Gly-Gly-Glu-Lys-Tyr-Leu, a sequence derived from the adenine binding site of biotin carboxylase. The according receptor **65** had an increased affinity to ATP ($K = 10^4$ M^{-1}). In the last step the so derived peptide was crosslinked with dimethyl adipimidate (**66**) in the presence of ATP. The cyclic peptide **67** now had a binding constant of 5×10^4 M^{-1} to ATP. Additionally, in contrary to its precursors **67** was now able to distinguish between ATP and ADP, AMP, GDP, GTP which are all bound significantly weaker by one order of magnitude. When the cyclization was carried out without the presence of ATP the affinity of the corresponding receptors was decreased, which suggests the formation of alternative cyclization products. This effect was attributed to the preoccupation of the two middle Lys moieties with binding to the phosphate hinge. Only the

first and the last lysine side chains are then freely available for crosslinking. Unfortunately, the authors did not provide additional data in solution to support this hypothesis. However, they were able to proof the general concept of combining combinatorial chemistry with rational design and molecular imprinting in order to increase the affinity of a library-derived receptor – in this case by one order of magnitude.

In conclusion, nucleotide recognition in water has received considerable attention. Many different artificial host systems have been developed and high binding constants of up to 10^7 M^{-1} could be achieved in aqueous medium. The majority of nucleotide receptors are based on polyamines – either cyclic or linear ones. Another, less frequently used binding motif for ion pairing to the phosphate hinge is the guanidinium group. Very often aromatic moieties are implemented into the hosts in order to induce π-stacking. Following a similar approach calixarenes or cyclodextrins are used in order to present hydrophobic pockets for nucleobase and/or sugar. Another successful principle is the utilization of tweezer receptors which form a preorganized cavity for the accommodation of the substrate. Very often two anion binding motifs are incorporated – one into each arm. This way the receptors may adopt a geometry that matches the tetrahedral shape of the phosphate anion. Hydrogen bonding patterns play a crucial role for selective recognition of nucleotides. While strong binding is mainly achieved by increasing charge-charge interactions, introducing selectivity is more subtle, because purine and respectively pyrimidine nucleobases are quite similar in shape and hydrophobicity. Another kind of selectivity in all of the above mentioned systems is the preference – if there is any – for tri- over di- and monophosphates, which can easily be explained by additional ionic interactions. Peptide-based systems have also been used for the recognition of nucleotides. Although the results are not yet as good as for polycationic host systems, the utilization of amino acids as building blocks offers the intriguing potential to generate artificial receptors by means of rational approaches, combinatorial chemistry or combinations thereof. Therefore the development of peptide-based tweezer systems for nucleotide recognition making use of aromatic and cationic amino acids and the implementation of a powerful anion binding site, like e.g. **1**, into each arm might present a promising new approach towards selective and efficient nucleotide recognition in water.

2.3 NUCLEIC ACIDS

Selective binding to nucleic acids, i.e. the molecular recognition of double- (ds) and single stranded (ss) RNA and DNA, has received considerable attention during the last decades. Many functions of life, like storage, readout and transmission of genetic information or regulatory processes are linked to nucleic acids and their non-covalent interplay with other molecules.[42] In order to gain more insight into these complex processes the development of small artificial molecules which are able to specifically interact with RNA or DNA is of great interest for the development of diagnostic probes and for the treatment of genetic or acquired diseases with the help of polynucleotide (PN) sensors,[64] RNA and DNA targeted drug design,[65,66] and delivery agents (see chapter 2.4). While nature has optimized complex protein structures which are able to bind to the desired sequence, chemists are trying to develop simpler, small artificial polynucleotide binding molecules which retain nature´s ability to molecularly recognize DNA and RNA with high affinity and selectivity.

2.3.1 DNA/RNA Structure

In order to develop such artificial host systems we have to understand the polynucleotide structure first. DNA consists of the nucleotides A, G, C and T depicted in Figure 13 (see chapter 2.2). Two such building blocks are linked via their 3´ and 5´ hydroxyl groups by a phosphate. RNA incorporates U instead of T. Consequently there are two ends of the strand, the 3´- and the 5´-terminus. Polymerases add new nucleotides to the hydroxyl group at 3´ – the chain grows into this direction. In the following, when describing polynucleotide sequences an annotation from 5´ to 3´ will be utilized. The sugar moiety is a five-membered ribose ring in RNA and a 2´-deoxyribose for DNA.

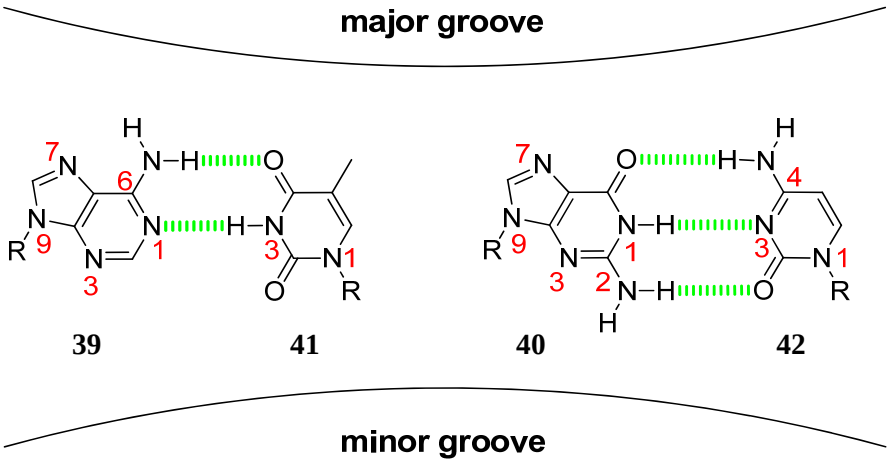


Figure 25. *Watson-Crick base pairing and schematic presentation of minor and major groove. The AT pair (left) forms two and the GC pair (right) three H-bonds, respectively.*

In 1953 *Watson* and *Crick* published their now famous paper about the "Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid",[67] in which they could show for the first time that DNA forms a double helix. The nucleobases are paired as shown in Figure 25 in the center of the helix with adenine (**39**) pairing with thymine (**41**) and guanine (**40**) with cytosine (**42**), respectively. The rather hydrophobic base pairs (bps) are stacked above each

other while the ribophosphate chain acts as backbone with the charged phosphate groups pointing outwards towards the surrounding water. The stability of the helix mainly stems from π-stacking between the base pairs and to a lower extent from the hydrogen bonds between the nucleobases. GC pairs contribute more to the thermal helix stability than AT pairs. Again, this is mainly due to superior stacking properties and only partly due to the additional hydrogen bond. The hydrogen bonding pattern between the nucleobases however is responsible for the molecular recognition between the *Watson-Crick* base pairs.[68] The two strands of the double helix are antiparallel (the 5′-end of one strand opposes the 3′-end of the other) and complementary to each other. Since purine bases are always paired with pyrimidines the helix has a regular diameter. Depending on the sequence and the surrounding medium there are different kinds of DNA helices – the main representatives are depicted in Figure 26.

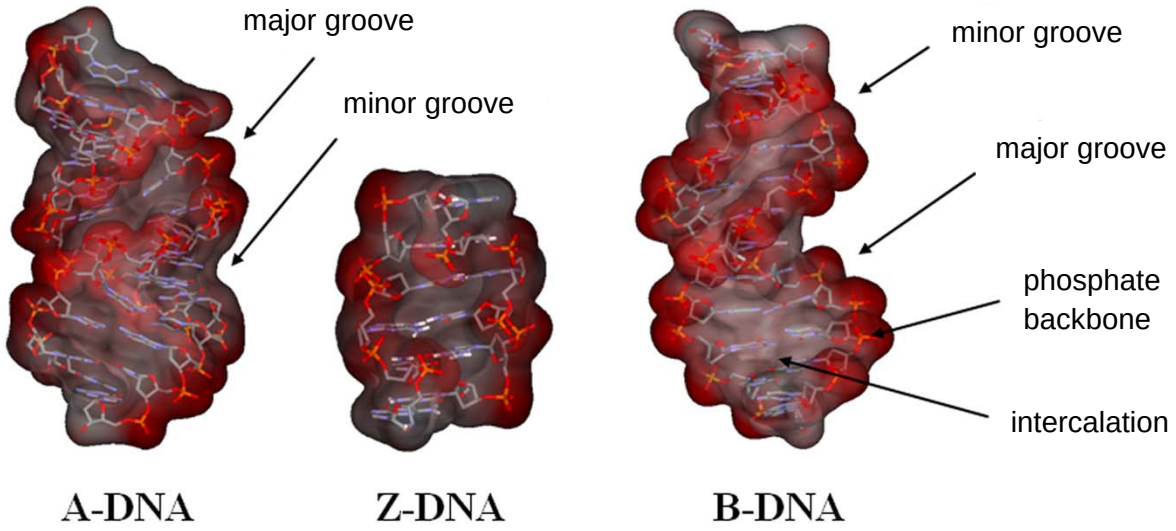


Figure 26. *From left to right: X-ray crystal structures of A-, Z- and B-DNA (PDB access code: 3QK4,[69] 2ELG,[70] and 436D[71]). In general, there are four possibilities to interact with DNA: binding in the minor or major groove, to the negatively charged phosphate backbone (red) or intercalation between two base pairs.*

B-DNA is the most commonly found form of DNA in cells. It adopts a right-handed helix with 2.0 nm diameter and 10 base pairs and a pitch of 3.4 nm per helical turn or in other words with each nucleotide the helix turns by 36° and is prolonged by 0.34 nm. The base pairs are tilted by 6° with regard to the helical axis. As a result of the regular helical winding the double strand features two openings where the nucleobases are directly exposed to the solvent, a narrow and deep minor groove and a wide and deep major groove. While proteins mainly bind to DNA via the major groove (e.g. transcription factors),[72] most artificial DNA-binders interact with the minor groove (vide infra). Other possibilities for the interaction with nucleic acids are intercalation between two base pairs or unspecific charge-mediated complexation to the phosphates.

A-DNA does not normally occur under physiological conditions but rather under partially dehydrated conditions or in RNA-DNA hybrids and DNA-protein complexes.[73] While B-DNA is usually formed at high humidity (> 85 %), A-DNA forms between 75 and 80 % humidity. At even lower values DNA becomes disordered (55-75 %). A-DNA is a right handed helix with 2.6 nm diameter and 11 base pairs and a pitch of 2.8 nm per helical turn.

This translates to a twist of 33° and a prolongation of 0.26 nm per nucleotide. The base pairs are tilted by 20°. As a result the minor groove is wide and shallow while the major groove is narrow and deep.[73]

Z-DNA is mainly found in GC-rich parts of the nucleic acid especially when cytidine is methylated.[74] Its structure is very different from the other two types presented here. It is a left-handed helix with 1.8 nm diameter, 12 base pairs and a pitch of 4.5 nm (twist of -30° and prolongation of 0.37 nm per nucleotide). The base pairs are tilted by 7°. In comparison to the B-form the guanosines are flipped upside down leading to a zigzag of the backbone. Consequently the minor groove is narrow and deep while the major groove in contrary is very broad and convex.[73]

Hud and *Plavec* introduced the concept of **A- and G-tracts** to correlate the DNA-sequence with preferred metal-ion accommodation sites and helical structure.[75] A-tracts are defined as sequences of at least four consecutive base pairs with only AA or AT steps (e.g. the synthetic homopolymer pdA × pdT). They adopt a peculiar B-type structure (B*) with an unusually narrow minor groove and a high propeller twist (angle between the planes of the two members of a base pair). Monovalent cations are predominantly bound in the minor groove, either deep inside or on top. Divalent cations are also bound in the minor groove bridging two opposing phosphate groups. The minor groove features the electronegative adenine N3 and thymine O2 and no electropositive amino protons like in GC sequences. However, they do feature the electropositive adenine amino group near the center of the major groove, which is why cations are almost never found in the major groove in this kind of sequence.

G-tracts on the other hand are defined by sequences which are rich in GG base steps and only contain GC base pairs (e.g. synthetic pdG × pdC). These sequences tend to undergo a transition from B- to A-DNA, especially with increasing cation association, which are preferentially bound in the major groove. The minor groove though is not very attractive for cations due to the electropositive protons of the N2 guanine amino group.

Most sequences are neither A- nor G-tract and thus defined as generic DNA. Their sequence is a mixture of AT and GC base pairs and the consequent helix adopts a canonical B-form with cations binding to both minor (inside and on top, preferentially at AT bps) and major groove (near N7 of guanine).

RNA differs from DNA by the 2´-OH group and the replacement of thymidine for uridine. Contrary to DNA it is not necessarily a double-strand but also exists in single-stranded forms. Therefore it may adopt a broader variety of structures: There exist for example hairpin, stemloop, loop and helical structures. As double-strand it either adopts an A or an A´-type *Watson-Crick* right-handed helix with two antiparallel and complementary strands. A´-helices are found at high salt concentrations. They accommodate 12 nucleotides per helix turn with a pitch of 3.0 nm and a base pair twist of 16-19° from the helical axis. Both structures form a very deep major and a rather shallow minor groove.

2.3.2 CD-Spectroscopy of Polynucleotides

Circular Dichroism Spectroscopy

Circular dichroism (CD) is a specific attribute of an optically active molecule: It interacts differently with right- (RCP) and left-circular polarized light (LCP).[76] Electromagnetic radiation consists of an electric (*E*) and a magnetic field (*B*) which oscillate perpendicular to each other and to the propagation direction. As depicted in Figure 27 the electric field vector´s magnitude in linear polarized light only oscillates in one plane containing the propagation direction (*k*). In circular polarized light the electric field vector rotates around the propagation direction with constant magnitude. Per definition a counterclockwise rotation towards the observer is left-circular polarized light and the other way round. When passing through an optically active medium RCP and LCP are not absorbed equally – one enantiomer absorbs more strongly RCP and the other one LCP. CD is the difference in light absorbance.

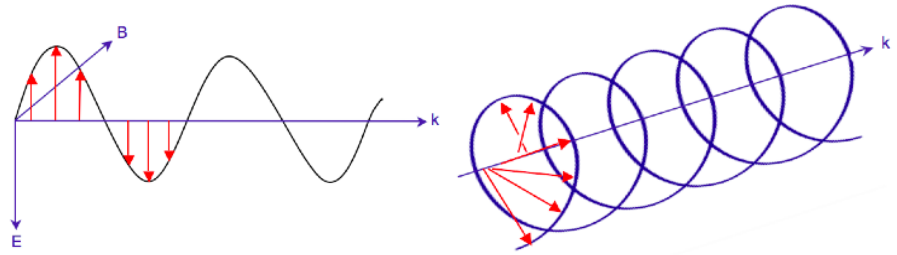


Figure 27. *The electric field vector E of linear polarized light only oscillates in one plain containing the propagation direction k (left). In circular polarized light the electric field vector rotates around k (right).*

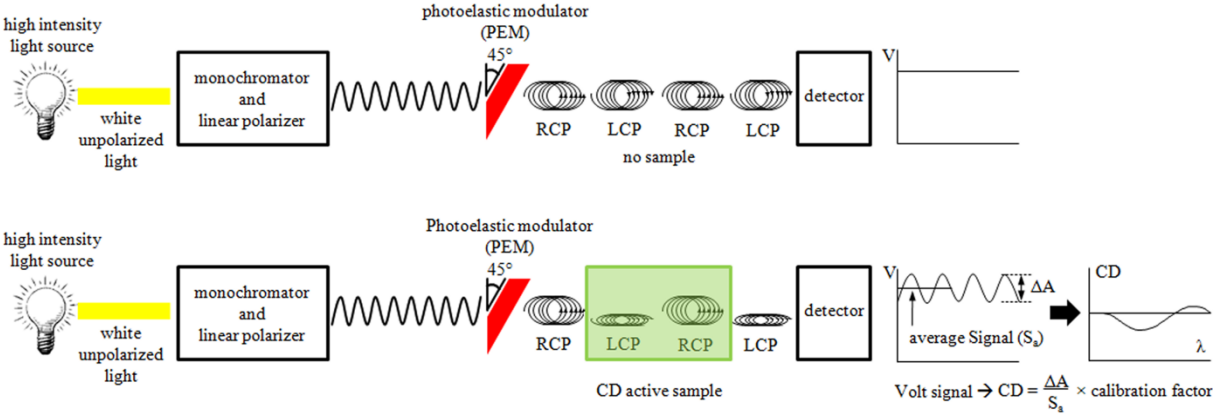


Figure 28. *Scheme of a CD-spectrometer without (top) and with a CD active sample (bottom). White light from a high intensity light source is linear polarized and then converted to circular polarized light in a photoelastic modulator (PEM). Right- (RCP) and left-circular polarized (LCP) light are emitted alternately. The detector records a voltage signal which is converted to the CD signal.*

With the help of CD-based spectroscopic methods the structure of optically active molecules can be studied. Figure 28 depicts the scheme of a CD spectrometer. A strong light source emitting white light is needed because CD values are very low in comparison to for example simple absorption spectroscopy (10^{-4} to 10^{-1} extinction values). In order to achieve results above the detection limit and standard deviation, most commonly a strong high-pressure

Xenon lamp is used. For biological systems wavelengths between 160 and 400 nm are of interest. The white light is passed through a monochromator and linear polarizer to obtain linear polarized light of a specific wavelength. A photoelastic modulator (PEM) at an angle of 45° converts the linear polarized light alternately to RCP and LCP (e.g. with a frequency of 50 kHz). Without sample the detector records a constant voltage signal. An optically active sample on the other hand will absorb more strongly either RCP or LCP. The detector will now record a signal oscillating in the frequency generated by the PEM. By dividing the difference in absorption (ΔA) through the average signal (S_a) and taking into account an internal calibration factor the voltage signal can be converted into the CD signal. With the help of the monochromator this is done for a whole spectral range resulting in the CD spectrum of the optically active probe.

Secondary Structure Determination

The absorption of nucleic acids between 240 and 280 nm stems from π-π*-transitions of the aromatic purine and pyrimidine bases. These are planar and optically inactive and consequently do not have an inherent CD signal. When incorporated into nucleosides or nucleotides however, the adjacent ribose leads to a chiral perturbation of the absorption and to a CD signal. The more nucleotides are linked together, the stronger is the CD signal. The CD signal of oligonucleotides is even stronger than the sum of the single signals from the nucleotide monomers. This is due to the coupling of bases stacked above each other. Finally, double-stranded polynucleotides additionally adopt a chiral secondary structure further increasing the strength of the chiroptical signal. Each type of secondary structure leads to a different CD spectrum which can thus be used to determine the structure of a given nucleic acid. Figure 29 shows typical CD signatures of A-, B-, and Z-form helices: A B- to A-helix transition was induced by going from high to low relative humidity. The B- to Z-transition of the synthetic alternating polynucleotide $p(dGdC)_2$ was triggered by increasing the ionic strength of the solvent.[77]

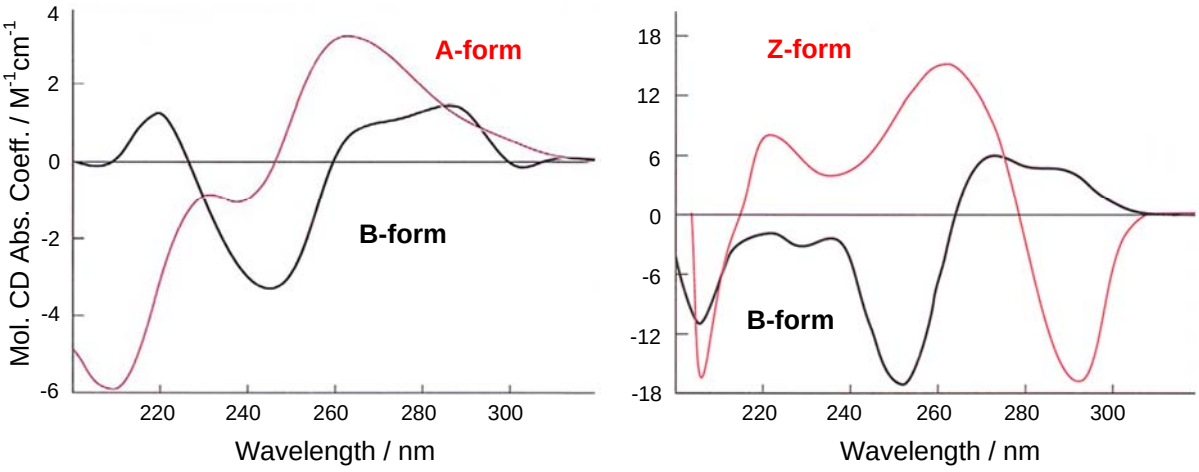


Figure 29. *Left: CD spectrum of A-form DNA at low relative humidity and B-form at high relative humidity. Right: B- to Z-transition of $p(dGdC)_2$ occurs when going from low ($[Na^+]$ = 10 mM) to high ionic strength ($[Na^+]$ = 3.9 M).*[77]

The B-form spectrum has a positive band centered at 275 nm, crosses zero around 260 nm to a negative band at 240 nm. This curvature is not due to exciton coupling but rather the sum of

all the couplings of transitions in all bases. At 220 nm the CD signal is either less negative than the band at 240 nm or slightly positive. A-helices show a strong positive band at 260 nm which is larger than the corresponding B-DNA band. Between 250 and 230 nm the curvature is rather flat, but not necessarily zero. Around 210 nm there is a big, negative band. The Z-spectrum features a negative band at 290 nm and a positive one at 260 nm and near 200 nm again a negative band.

Determination of Binding Mode

CD spectroscopy is a useful tool for studying the binding mode of polynucleotide binders, also commonly referred to as polynucleotide ligands.[78,79] Even when the ligand itself is not chiral and thus has no CD signal, upon binding to a nucleic acid it is brought into a chiral surrounding which gives rise to an induced CD (ICD) signal. The induced signal originates from coupling of electronic transition moments of the ligand and the DNA bases. Therefore the observation of an ICD is indicative of a polynucleotide-ligand interaction. The sign and intensity of the signal are dependent on the exact location and orientation of the nucleic acid-binder and the oscillator strengths of the ligand transition.

Intercalation between base pairs generally leads to small ICDs. When the transition moment of the intercalator is oriented along the short axis of the intercalation site (see Figure 30) the corresponding ICD is positive. Accordingly, it is weakly negative when oriented along the long axis. If the ligand is not placed in the middle of the intercalation site a more complicated signal is generated which depends on the lateral displacement and the surrounding base pairs.

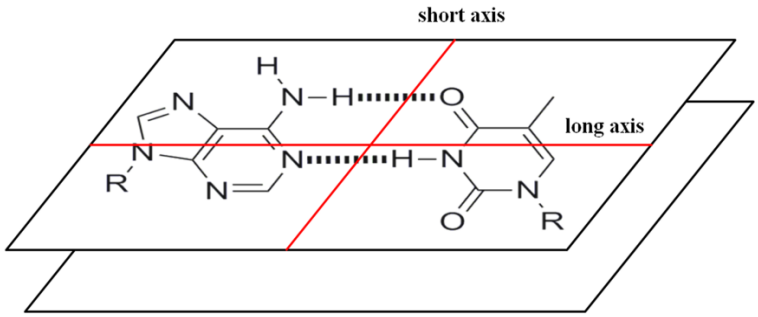
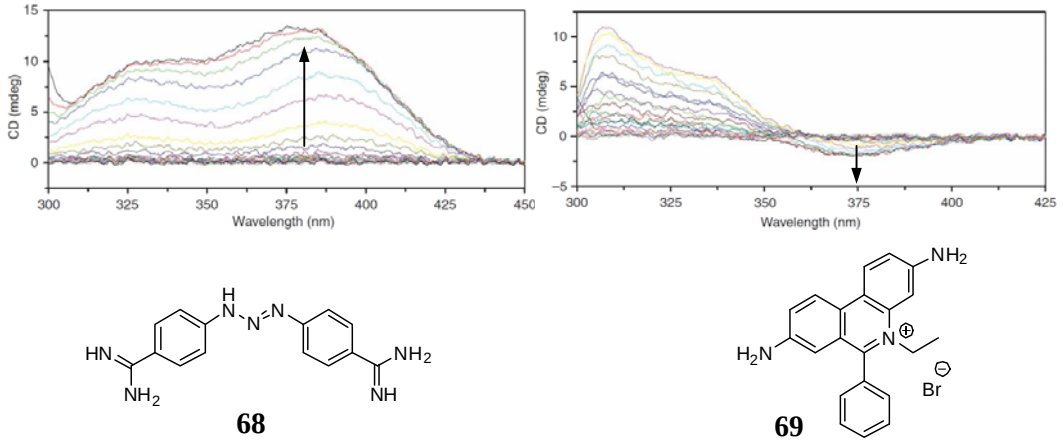


Figure 30. *Schematic representation of an intercalation site between two planes defined by two base pairs. The plain itself has a long and a short axis.*

Stronger ICD signals are indicators for groove binding. These signals can be more intensive by up to two orders of magnitude compared to an intercalator signal. When the ligand is bound in the minor groove its transition moment is oriented alongside the groove with approximately 45° with respect to the helical axis. The resulting ICD signal is strong and positive. ICDs caused by major groove binding are more variable due to the larger amount of potential orientations in the broad groove. Two exemplary examples for titration experiments with the B-type helix forming calf thymus DNA (ctDNA) with the minor groove binder berenil **68** and respectively with the intercalator ethidium bromide **69** are depicted in Figure 31. For the first one a strong positive ICD at approximately 380 nm is observed. The latter one gives rise to a weak, negative ICD band at 375 nm.

Generally, if only the intensity and not the shape of an ICD band changes during a titration experiment this is a strong indicator for a single binding mode. Alterations in the shape may be a hint for multiple binding modes, but might also originate from conformational changes of the polynucleotide or stem from ligand-ligand interactions. Ligand dimer formation or higher ordered complexes which bind to DNA often lead to a bisignate band shape with a positive and negative part localized on either side of the ligand´s absorption maximum. This kind of signal is due to exciton coupling caused by the close proximity of two identical transition moments of the same energy. Typically this is only the case for groove binders.



68 **69**

Figure 31. *A typical CD titration experiment of the minor groove binder berenil* ***68*** *with the B-helix forming ctDNA gives rise to a strong positive ICD signal (left). The CD titration with the intercalator ethidium bromide* ***69*** *leads to a weak, negative ICD signal (right). The graphs of the CD experiments are reprinted with permission from Macmillan Publishers Ltd: Nature protocols, Garbett, N. C. Ragazzon, P. A.; Chaires, J. B.* ***2007****, 3166-3172. Copyright 2007.*[79]

2.3.3 Thermal Melting Experiments

The stability of secondary nucleic acid structures can be determined with the help of thermal melting experiments.[80] When paired in a double helix the molar UV absorptivity of the nucleobases is lower than would be expected from the sum of all nucleobases. This hypochromicity is due to the coupling of transition dipoles between adjacent, stacked nucleobases and can be measured by heating a polynucleotide and observing the absorbance at 260 nm. Upon denaturation the UV absorption normally increases by 20-30 %.

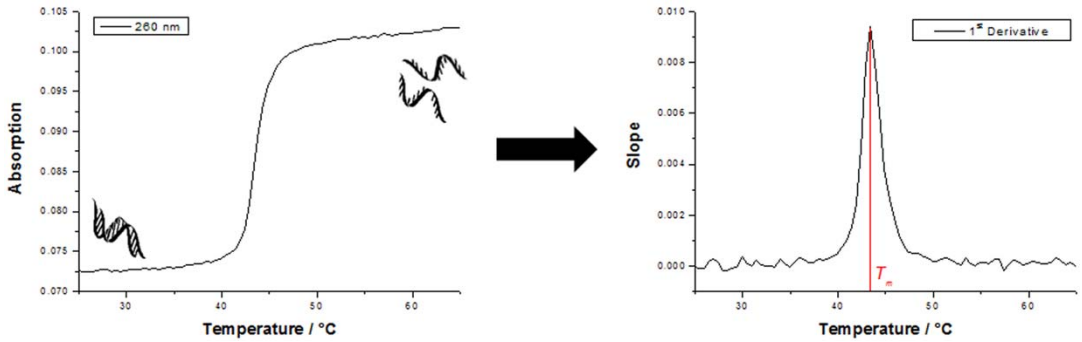


Figure 32. *Thermal melting experiment of duplex RNA recorded at 260 nm (left) and calculation of the 1st derivative (right) allow the determination of the melting point T_m.*

Upon heating the helix starts unwinding in a cooperative process which leads to a sigmoidal UV curve (see Figure 32). The point of highest slope, i.e. the point where 50 % of the duplex is unwound, is called melting point T_m. It can be calculated by plotting the first derivative of the absorbance against the temperature. AT base pairs contribute less to the stability than GC base pairs. Consequently sequences with high GC content have higher melting temperatures. Homopolymers (e.g. pdA × pdT) generate steeper melting curves than mixed sequences (e.g. ctDNA). In mixed polynucleotides AT regions start melting first, followed by mixed and finally by GC-rich regions. Hence, the melting curve is broader. The melting temperature can also be utilized to estimate the affinity of a nucleic acid ligand.[81] Generally high thermal stabilizations correlate with high binding constants.

2.3.4 *Isothermal Titration Calorimetry*

Another method to characterize binding events between nucleic acids and their ligands is the isothermal titration calorimetry (ITC), which allows direct determination of the thermodynamic parameters ΔH, K and n.[82]

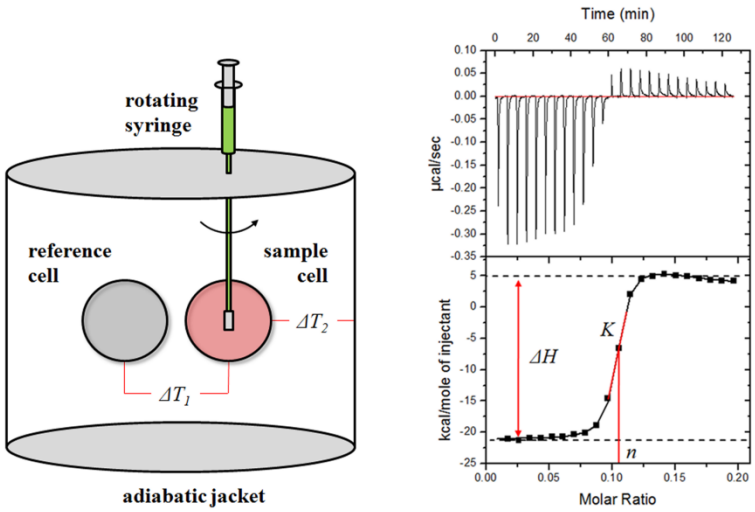


Figure 33. *Scheme of an ITC: Reference and sample cell are isolated in an adiabatic jacket. Sample is injected via a rotating syringe (left). Heat vs. time is recorded during experiment, integrated and plotted against molar ratio. Binding enthalpy (ΔH), binding constant (K), and stoichiometry (n) are obtained by curve fitting (right).*

As shown in Figure 33 (left) a calorimeter consists of two cells built from thermal highly conductive material (e.g. gold) placed in an adiabatic jacket. One is a reference and the other a sample cell of known volume. The sample cell contains one binding partner (receptor or ligand), the reference cell pure solvent/buffer. The other binding partner (ligand or receptor) is injected by a high precision, computer-controlled rotating syringe in precisely known aliquots. The experiment is carried out at constant temperature which is constantly controlled by an external reference (ΔT_2). Upon injection the association between receptor and substrate generates (exothermal) or consumes (endothermal) heat. The heat difference between sample and reference cell (ΔT_1) is measured with the help of half-conducting thermopiles and kept constant. The amount of thermal power which has to be applied in order to compensate for the heat of reaction is applied or withdrawn and recorded. Thus the measurable quantity in an ITC experiment is the thermal power as a function of time. An exothermal experiment results in negative and an endothermal one correspondingly in positive heat spikes which are in the range of µcal. Figure 33 (right) depicts the results of a typical ITC experiment. During the first few injections there is a large excess of binding partner in the sample cell. Consequently, the entire amount of titrant is bound – a large heat signal is the outcome. At a certain point, the titrate in the sample cell starts to saturate and not all of the injected substance is bound - the heat becomes smaller. At the end of the experiment no binding occurs anymore and only the heat of dilution is measured. Each injection leads to a spike which is integrated over time and plotted against the molar ratio of receptor/substrate to give a sigmoidal curve. Fitting these data with a non-linear least squares method gives the reaction enthalpy (ΔH) as the difference between the baseline at the beginning of the binding event and after saturation. The binding constant (K) can be calculated from the slope and the stoichiometry (n) can be read

out at the point of highest slope. With the help of equation (1) the reaction´s Gibbs free energy (ΔG) and entropy (ΔS) can be calculated. R is the universal gas constant and T the temperature used for the experiment.

$$\Delta G = -RT\ln(K) = \Delta H - T\Delta S \qquad (1)$$

A prerequisite for ITC measurements is that the binding event has a heat signature and that the resulting data points generate sufficient curvature of the integrated heat signals for fitting. The degree of curvature can be estimated by the C-value which should be between 10 and 100 for an optimal binding experiment and can be calculated with equation (2).

$$C = n\,K\,[\text{binding sites}] \qquad (2)$$

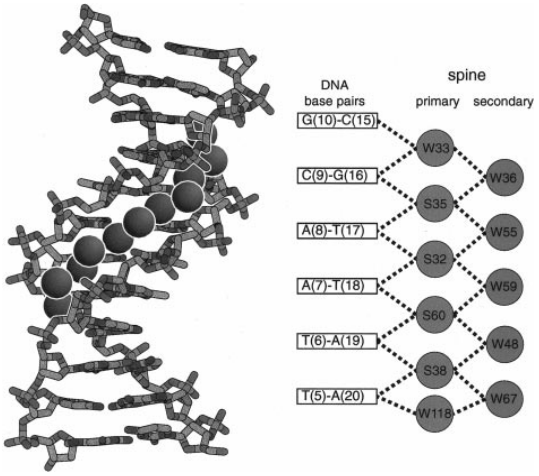
The stoichiometry n multiplied with the binding constant K and the concentration of binding sites gives C. The number of binding sites can be approximated for nucleic acids as the concentration of phosphates. Too small values (< 1) lead to shallow, non-descriptive isotherms. With too big values (> 1,000) the transition region is devoid of data points. Neither extreme can be fitted unambiguously. Generally, for high binding constants relatively low concentrations must be used and vice versa. It has to be assured however that the heat of binding is higher than the detection limit of the instrument. Binding constants in the range 10^2 - 10^9 M^{-1} are directly accessible, binding constants of up to 10^{12} M^{-1} with the help of competitive binding assays. ITC is a very sensitive method and great care has to be taken during sample preparation. All binding data have to be corrected for the heat of dilution of receptor and ligand which have to be determined in separate control experiments. If carried out correctly all important thermodynamic quantities of a binding event can be determined in just one experiment with the help of microcalorimetry. Concerning DNA-ligand interactions the energetic profile can give insight into the binding event. Several energetic barriers have to be overcome for binding, like conformational changes and the loss of translational and rotational freedom upon complexation. Advantageous contributions stem from liberation of bound water and counterions to the solvent and from molecular interactions between host and guest, like hydrogen bonds, ion pairing, and van-der-Waals forces. The exact knowledge of enthalpic and entropic contributions helps to elucidate which of these molecular interactions are involved in a given binding event.[83]

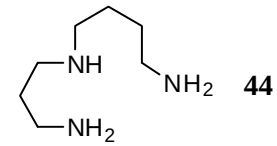
2.3.5 Specific DNA/RNA Recognition

With the help of above mentioned methods together with NMR and X-ray crystal structure determination an impressive number of nucleic acid binding molecules have been characterized. In the following an excerpt of instructive examples of natural[84] and synthetic systems[85,86] will be presented. As already mentioned, in general four possibilities exist to interact with nucleic acids: unspecific interactions via charge interactions to the phosphates, intercalation between two base pairs or minor and major groove-binding.

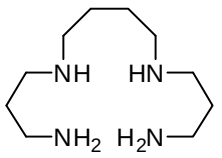
Ribophosphate-Binding and Counter Ion/Hydration Shell

Double-stranded nucleic acid helices are linear polyanions. The pK_A value for the hydroxy group of the bridging phosphates is < 1. Consequently, at physiological pH each of these phosphates carries one negative charge. In order to stabilize their structure in aqueous medium ds-polynucleotides accommodate metal cations for charge compensation. *Manning* described the metal ion binding to DNA as "counter ion condensation".[87] The metal ions are bound within a few angstrom of the surface keeping their primary hydration shell but are free to move along the helix thereby forming a hydrated counterion cloud around the negative charge density of the ribophosphate backbone. A typical B-type helix has 0.76-0.88 monovalent counter ions per phosphate group. The associated metal ions strongly influence the physical nucleic acid properties. Binding events for example are strongly dependent on the salt concentration of the solvent. Furthermore a nucleic acid accommodates per average three molecules of water per phosphate. Extensive crystallographic studies with the *Dickerson-Drew* dodecamer d(CGCGAATTCGCG)$_2$ depicted in Figure 26 as typical B-type DNA representative revealed a complex hydration shell at low temperature.[88] In the major groove there are two hydration layers. The first one is composed of 19 water molecules oriented around the N and O atoms of the nucleobases and on top of that is a second hydrate layer with 36 water molecules. As shown in Figure 34 in the minor groove there is a highly ordered, zigzag patterned shell of water partially consisting of sodium cations, the so called "spine of hydration".[89] Around the first shell there is a second spine of hydration, again highly ordered. Therefore especially upon groove binding cations and water molecules are liberated to the solvent – an entropically favored process.[90]





44



45

Figure 34. *Crystal structure depicting the primary and secondary "spine of hydration" in the minor groove of B-DNA and schematic presentation thereof (left). The crystal structure is reprinted with permission from Shui, X. McFail-Isom, L. Hu, G. G.; Williams, L. D. Biochemistry* ***1998****, 37, 8341-8355. Copyright 1998 American Chemical Society.*[89] *Non-specific ribophosphate backbone binder spermidine* ***44*** *and spermine* ***45*** *(right).*

Other molecules which compete with the metal counter ions for the unspecific electrostatic binding sites at the ribophosphate backbone are mainly polyamines like spermidine (**44**) and spermine (**45**). These are known to bind to phosphate (see chapter 2.2) and are present ubiquitously in eukaryotic cells, where they are thought to play multiple roles in the cell by interacting with nucleic acids, e.g. the modulation of protein-DNA interactions during cell proliferation.[91]

Intercalators

The concept of intercalation was first observed by *Lerman* in his studies concerning the interaction of aminoacridines like proflavine (**70**, Figure 35) with B-DNA.[92] The polycyclic aromatic amine is inserted between two base pairs, their aromatic faces stacked above each other. This leads to an extension of the double helix by approximately the van-der-Waals thickness of a phenyl ring (3.4 Å) and partial unwinding of the helix at the binding site. These physical changes have come to be known as the most important characteristics of DNA intercalation. The major driving forces for intercalation are π-stacking and electrostatics. All intercalators feature an aromatic polycyclic and electron deficient ring structure, which is in general positively charged at neutral pH and of more or less the same size as a base pair. Alternating pyrimidine-purine base-steps are preferred because they are easier to separate than their pyrimidine-pyrimidine and purine-purine analogues. Preferences apart from that are mainly due to additional interactions of non-intercalating parts in the grooves. From an energetic point of view there are many different interactions contributing to the overall binding energy. Hydrophobic transfer of the intercalator from the solvent to the binding site releases highly ordered water molecules and counterions – an entropically favored process. An enthalpic penalty has to be paid for the conformation change of nucleic acid. Enthalpy on the other hand is gained by π-stacking, van-der-Waals interactions, hydrogen bonding and the reduction of coulomb repulsion between phosphates in the unwound ribophosphate backbone.

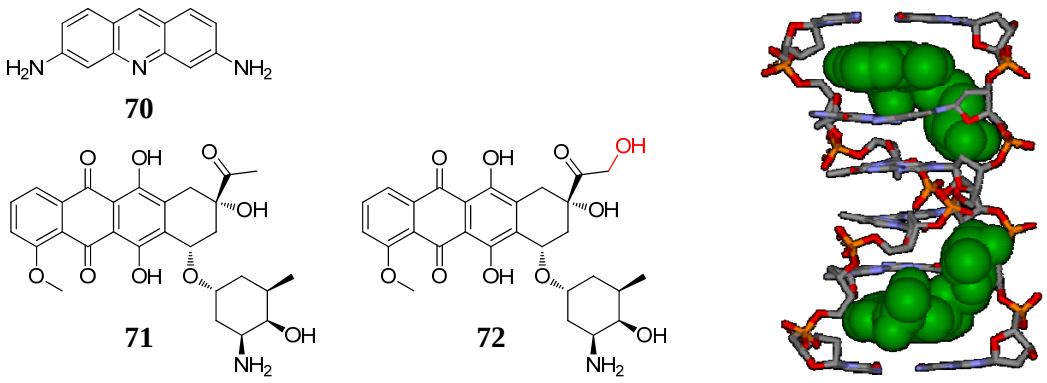


Figure 35. *Intercalators proflavine **70**, daunomycin **71**, and adriamycin **72** (left). Crystal structure of daunomycin-d(CGTACG)$_2$ complex (right) (PDB access code: 308D[93]).*

Many natural products bind to DNA via intercalation. The natural glycoside daunomycin (**71**) for instance intercalates into DNA as depicted in Figure 35. The antibiotic active compound is produced by S. *peuceticus* and S. *coeruleorubidus* and is used as cytostatic drug in cancer therapy.[94] The crystal structure with the hexamer duplex d(CGTACG)$_2$ shown in Figure 35 reveals that the amino carbohydrate lies within the minor groove while the methoxy containing ring is in the major groove. The stacking of the two adjacent rings perpendicular to

the long axis of the base pairs separates them by 3.4 Å but does not lead to unwinding at the intercalation site. The neighboring base pairs however are unwound by 8°. The tertiary hydroxy group donates hydrogen to the guanine N3 and at the same time acts as an acceptor for the amine group of respective guanine. The binding constant at neutral pH was determined as $6 \times 10^6\ M^{-1}$.[95] The binding energy originates from large favorable enthalpy minus unfavorable entropic contributions. This energetic signature is typical for intercalators. However, alternative energy profiles exist, too. The structurally very similar adriamycin **72** which only differs by one hydroxyl group intercalates stronger by a factor of 10. Here the entropy contributes a large favorable amount of energy to the binding event while enthalpy only contributes a little. Furthermore these two molecules are a good example that a similar structure does not automatically lead to a similar energetic profile.

Due to their aromatic structure intercalators often absorb light in the visible range. As depicted in Figure 36 ethidium bromide (EB, **69**) for instance is a red phenanthridine dye which intercalates into nucleic acids.[96] Upon binding its fluorescence intensity at 600 nm increases by a factor of 50-100.[97] This is why it is commonly used to stain nucleic acids, for example in gel electrophoresis experiments.[98] Ethidium bromide´s binding constant for generic DNA was determined via fluorescence experiments in buffered water at pH 7.5 to $K = 3 \times 10^6\ M^{-1}$ with saturation taking place at 0.33 equivalents EB per base pair.[99] Thus not all potential intercalation sites are occupied. It is true for all intercalators known so far that the maximum number of guest molecules which can be introduced into a double helix is 0.5 eq. Intercalators maximally bind to every second base pair. For more complex intercalators the maximum is 0.33 eq. This empiric observation is known as the "nearest neighbor principle".[73]

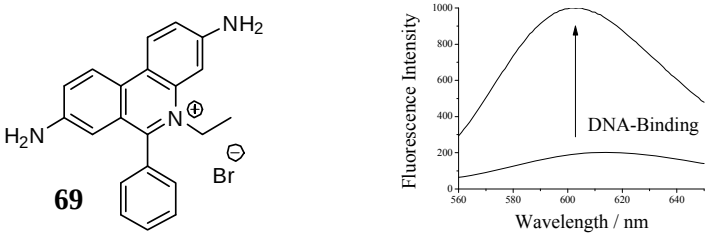


Figure 36. *Upon binding to nucleic acids the fluorescence intensity of the ethidium bromide (**69**) signal at 600 nm increases by a factor of 50-100.*

Schmuck designed the bifunctional intercalator **73** depicted in Figure 37, which is able to differentiate between ds-DNA (B-helix) and ds-RNA (A-helix) at pH 5 in buffered water independently of the base pair sequence.[100] At this pH the intercalator (0.3 eq) is able to thermally stabilize ctDNA by 7.3° while destabilizing the RNA pA × pU by 1.5°. Furthermore different spectroscopic answers are reported for RNA and DNA. At pH 7 neither stabilization nor spectroscopic differentiation could be observed. Determination of the complex stability revealed binding constants of up to $10^6\ M^{-1}$ at pH 7 and of up to $6 \times 10^6\ M^{-1}$ at pH 5. The difference of binding mode and affinity can be attributed to the protonation of the GCP moiety at pH 5 ($6 < pK_A < 7$). With the help of CD spectroscopy and molecular modeling defined binding modes could be determined at pH 5: **73** binds into the major groove of ds-RNA by formation of a dimer with stacked pyrene moieties thus lacking intercalation. For ds-DNA however intercalation of the pyrene moiety was observed while the GCP interacted

with the phosphate in the minor groove additionally forming five to six hydrogen bonds to the nucleobases.

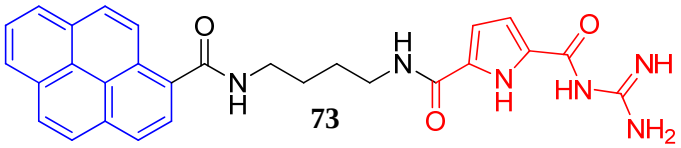


Figure 37. *Bifunctional **73** combines an intercalating pyrene (blue) and an oxoanion binding motif (red).*

Schmuck also conducted an extensive structure-activity relationship study.[101] Some exemplary potential bifunctional aryl-GCP intercalators (**74-77**) are depicted in Figure 38. It could be shown that the linker must be long enough to allow for simultaneous intercalation and minor groove binding. Ligand **77** only weakly interacted with nucleic acids because it was equipped with a too short linker. Acridine (**74**) seems to be the minimal size for intercalation – the according naphthalene derivatized ligand **75** did not intercalate into polynucleotides. The implementation of additional charges as in **76** further helped to improve the affinity, e.g. in comparison to the original ligand **73** the binding constant for ctDNA in buffered water at pH 7 was increased from 10^6 to 6×10^6 M^{-1}. Lastly, it could also be shown in a cancer cell assay that high binding correlated with antiproliferative activity most likely due to more efficient intercalation of the pyrene chromophore.

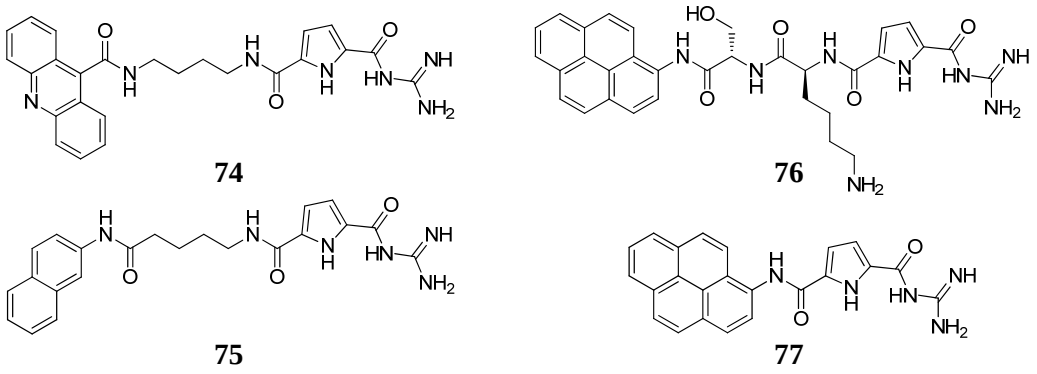


Figure 38. *Bifunctional aryl-GCP ligands **74-77** were used for a structure-activity relationship study.*

Another principle derived from nature is bisintercalation, for instance by the naturally occurring antitumor antibiotic (-)-sandramycin **78** (Figure 39) as characterized by *Boger*.[102] The binding constant to ctDNA has been determined via fluorescence titration in buffered water at pH 7.4 to 3×10^7 M^{-1}. The crystal structure of a complex with $d(GCATGC)_2$ illustrates some general principles for bisintercalation. Implementation of the two aromatic moieties most commonly takes place between next neighboring sites. In this case the depsipeptide linker rearranges its amide bonds to the more unfavorable *cis* configuration in order to span two instead of three base pairs. Next, the linker binds in the more accessible AT rich minor groove – also the biggest part of complex stability originates from non-covalent interactions of the depsipeptide with the groove. Finally, similar to mono intercalation, pyrimidine-purine sites are preferred for bisintercalation, too.

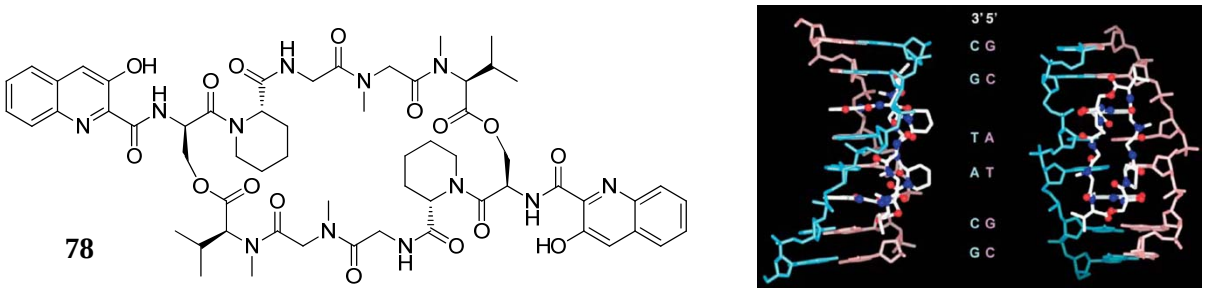


Figure 39. *(-)-Sandramycin **78** and NMR structure of its complex with d(GCATGC)$_2$. Reprinted from Tse, W.; Boger, D. Chem. Biol., 11, 1607-1617. Copyright 2004, with permission from Elsevier.* [84]

The concept of bisintercalation was very successfully utilized by *Chaires*.[103] By covalently linking two daunomycin molecules via an aromatic linker of 7 Å lengths the bisintercalator **79** depicted in Figure 40 was obtained. An extremely high binding constant of 3×10^{11} M^{-1} to herring sperm DNA could thus be achieved in buffered water. This affinity is four orders of magnitude higher than that of the monomer and almost in the region of specific protein-DNA interactions. The linker length corresponds exactly to the spacing between two daunomycin monomers as derived from the crystal structure depicted in Figure 35. By solving the crystal structure of artificial **79** with d(CDATCG)$_2$ it could be shown that much of the original binding interactions stay intact for the bisintercalator.[104] Intercalation takes place at CG steps and hydrogen bonding between the hydroxyl group and guanine is retained. The sugar moieties are still located in the minor groove but in contrast to the monomers they are lifted somewhat from the groove floor thereby forsaking the hydrogen bonding between the amines to the floor. The unwinding of DNA now is taking place more to the middle of the two intercalation sites and not at the neighboring base pairs. The energetic profile shows that binding affinity is dominated by a huge favorable enthalpy opposed by unfavorable entropy. The entropy penalty is higher than for the monomer which is probably due to the loss of rotational freedom of the linker upon binding.

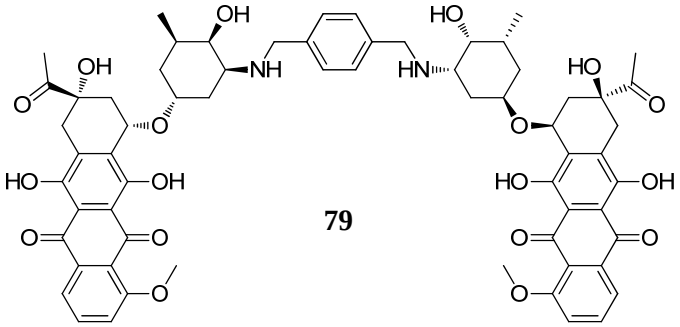


Figure 40. *Two daunomycin moieties are linked via an aromatic spacer in the bisintercalator **79**.*

Intercalation into base pairs is a very powerful method to achieve high binding constants to nucleic acids. Intercalators are polycyclic, often positively charged aromatic molecules which interact with the aromatic faces of their adjacent base pairs. Very high binding constants can be achieved by linking two of these molecules with an appropriate linker. There has been reasonable success in utilizing intercalators for the treatment of e.g. cancer or as antibiotics.

Unfortunately intercalation almost always comes at the cost of very high toxicity. Thus despite their usefulness the development of DNA-binders with alternative mode of interaction is desirable.

Minor Groove Binders

In contrary to proteins most natural and artificial groove-binding molecules bind to the minor groove due to its narrower, deeper structure which is better suited for interactions with small molecules, where they are able to interact with the deoxyribose groove walls and the groove floor which presents the edges of the base pairs as depicted in Figure 41.[105] Free electron pairs are schematically displayed as circles with dots and amine groups as circles with an H. GC base pairs offer the free guanine N3 electron pair as hydrogen acceptor (Acc) and the guanine NH_2 group as hydrogen donor (Do). With some spacing cytosine presents its O3 as acceptor. CG base pairs do possess the same Acc/Do/Acc pattern with reversed spacing. AT presents three hydrogen acceptor sites to the minor groove: the adenine N3 and in some distance the two free electron pairs of the thymine O3. For TA base pairs it is the other way round.

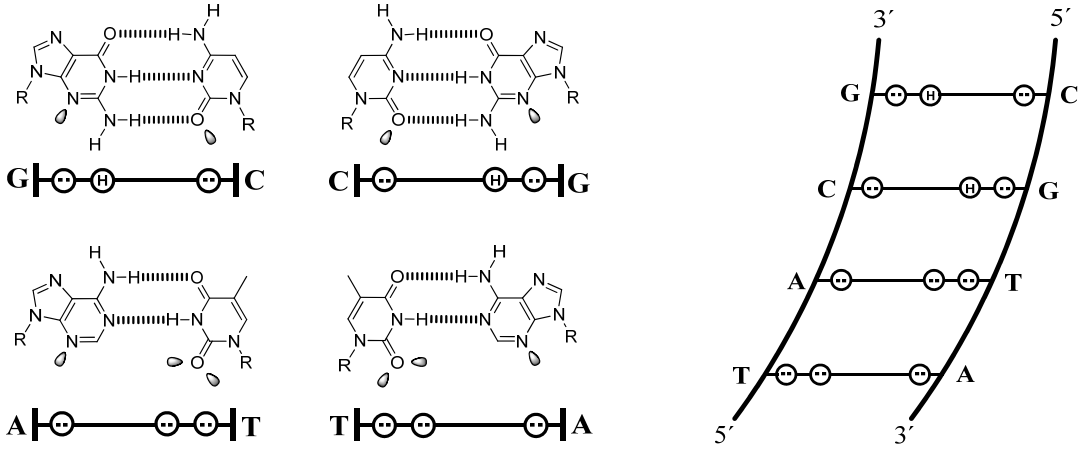
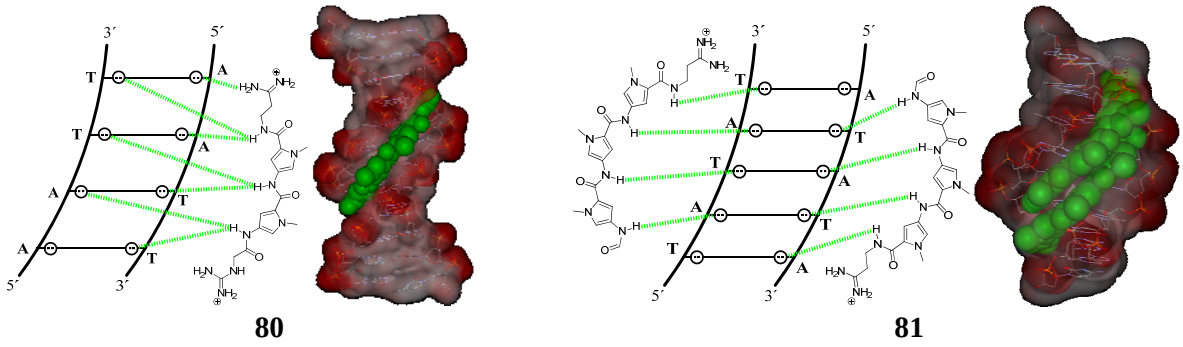


Figure 41. *Scheme of the hydrogen acceptor (circles with dots) and donor sites (circles with H) as they are presented on the floor of the minor groove.*

A very well studied group of minor groove binders is that of the naturally occurring antiviral antitumor antibiotics netropsin (**80**, from S. *netropsis*) and distamycin (**81**, from S. *distallicus*) shown in Figure 42. By binding to B-DNA they are able to interfere with both transcription and replication. Both molecules consist of repeated pyrrole (Py) subunits bridged by amide bonds and feature positively charged nitrogen-based headgroups at physiological pH. Netropsin incorporates a guanidine group at both ends and distamycin a charged amidine group at one end and an uncharged formyl moiety at the other one. This is a very typical design for minor groove binders. Very often simple aromatic rings like pyrrole, imidazole or benzene are linked via freely rotatable bonds. This way these molecules are able to adapt to the winding of the minor groove. Recognition of the groove mainly is a recognition of shape and curvature. This is also why in contrary to intercalators most minor groove binders prefer AT rich sites with narrower and deeper grooves where they can interact better with the groove walls via electrostatic and van-der-Waals contacts as well as hydrogen bonds. GC rich sites possess broader and shallower (due to the guanine amine) grooves and are less well suited for

accommodation of a ligand. Furthermore the negative electrostatic potential is greater in AT rich grooves making them better suited for positively charged ligands.[106] Upon binding there is no distortion or unwinding of the nucleic acid as would be the case for intercalation. It is rather the ligand which has to change its conformation. The crystal structure of the 1:1 complex of B-DNA with netropsin in Figure 42 was solved by *Dickerson* and clearly illustrates the shape-selective recognition in the narrow AT minor groove of the *Dickerson-Drew* dodecamer.[107] Netropsin spans approximately four AT base pairs with the carboxamide NH-groups pointing towards the minor groove floor where they form hydrogen bonds to the edges of AT and TA base pairs (adenine N3 and thymine O2). Upon binding several water molecules are displaced from the minor groove. Electrostatic contacts are formed between the guanidinium groups and the phosphates. Furthermore **80** forms van-der-Waals contacts with the ribophosphate groove walls. Thereby, the helical structure is only slightly altered. Netropsin does neither bind to ss-DNA nor to RNA or DNA/RNA hybrids. It actually promotes A to B and Z to B transformations, thus clearly underlining its propensity to bind to B-type DNA. While analogue 1:1 complexes are also known for distamycin *Wemmer* could show with the help of two dimensional NMR spectroscopy that **81** is also able to form 2:1 complexes within the minor groove of sequences containing more than four consecutive AT bps.[108] Two distamycin molecules are accommodated antiparallelly within the groove which consequently has to broaden. By accommodating two instead of one ligand the increased hydrophobic interactions overcompensate the enthalpic penalty which has to be paid for the changes of secondary DNA structure. Netropsin is probably prevented from forming a 2:1 binding mode due to its charged headgroups at both ends.



80 **81**

Figure 42. *X-ray crystal structure of the 1:1 complex of netropsin (**80**) with the Dickerson-Drew dodecamer B-DNA (left, PDB access code: 101D)[109], scheme of a 2:1 complex of distamycin (**81**) and d(GTATATAC)$_2$ and X-ray crystal structure thereof (right, PDB access code: 378D).[110] The hydrogen bonding pattern is schematically presented for both conjugates.*

More detailed calorimetric titration studies with the sequence d(GCCAAATTGGC)$_2$ in buffered water at pH 7 revealed similar binding constants for the 1:1 complex of both molecules.[111] Netropsin binds with K = 4 × 10^7 M^{-1} and distamycin with 3 × 10^7 M^{-1}. Both processes are driven by favorable enthalpy. The enthalpic stabilization for distamycin is larger than that for netropsin (ΔH = -12.3 versus -7.5 kcal mol^{-1}). This is probably due to the increased number of interactions of the additional pyrrole ring within the groove. Entropy on the other hand is only favorable for netropsin and not for distamycin (ΔS = 9.3 cal K^{-1} mol^{-1} versus -8.0 cal K^{-1} mol^{-1}) which was attributed to differences in the release of counter ions and water molecules from the groove. The second distamycin molecule binds more weakly than

the first one by one order of magnitude with $K = 3 \times 10^6\ M^{-1}$ despite a more favorable enthalpy (-18.8 kcal mol^{-1}) originating from improved van-der-Waals contacts of the stacked monomers to each other and to the groove wall. The reason for that is that the formation of a 2:1 complex leads to a huge entropic penalty (-34.0 K^{-1} mol^{-1}) preventing higher affinity.

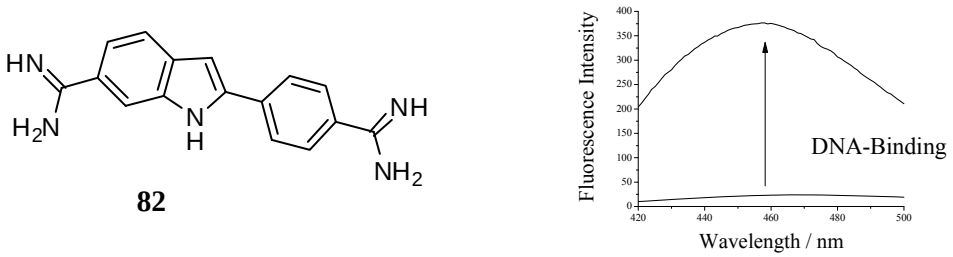


Figure 44. *The fluorescence intensity of DAPI (**82**) increases upon binding to DNA minor groove.*

Another prominent minor groove binder is 4′,6-diamidino-2-phanylindole **82** (Figure 44), or short DAPI, which was synthesized by *Dann* in 1971. Although it was not, as usually planned, a successful drug against trypanosomiasis it is nowadays one of the most commonly used dyes for staining DNA. It is a minor groove binder with selectivity for AT rich sequences. Its blue fluorescence signal at 460 nm increases strongly upon binding to DNA. When binding to RNA the fluorescence intensity is not increased as strongly and the maximum shifts to 500 nm – it is thus able to discriminate between DNA and RNA. Since DAPI easily crosses cell membranes it is extensively used for staining cell nuclei or for the visualization of nucleic acids in both live and fixed cells.[112]

In theory, sequence selectivity and the ability to span several base pairs make minor groove binders the ideal molecules for the sequence-specific recognition of nucleic acid sequences which are long enough to be unique in a cell. This would clear the road for very specific polynucleotide targeting drugs with little side effects or toxicity (cf. intercalators). While minor groove recognition itself is mainly driven by size complementarity, sequence selectivity was derived from selective recognition via hydrogen bonds of donor and acceptor sites presented at the minor groove floor by *Dervan*.[113] His ligand design is similar to the naturally occurring distamycin. By replacing one pyrrole (Py) for an imidazole (Im) ring, the polyamide ImPyPy (**83**) depicted in Figure 46 was obtained. This compound forms antiparallel 2:1 complexes in the minor groove selectively to sequences of the type XGXCX (X = A or T). The stacked Im/Py pair selectively recognizes GC over CG base pairs and vice versa Py/Im prefers to bind to CG. Both combinations do not bind well to AT or TA base pairs. The imidazole interacts with the guanine N3 amine but unlike pyrrole is not able to form hydrogen bonds to cytosine, adenine or thymine. Additional complex stability stems from several hydrogen bonds between the polyamide NH groups with the edges of GC pairs. The Py/Py pair on the other hand is able to recognize AT/TA over GC/CG bps. In combination with the specific selectivity of Im/Py pairs this was the first step towards sequence selective minor groove binding. 2:1 complex forming ligands are able to recognize a given sequence with each monomer being responsible for one “side” of the base pairing system. This scheme can in theory be applied to many ds-nucleic acid sequences, but one can easily imagine sequences where two non-identical monomers would be necessary. This would then lead to the formation of two homodimers besides the desired heterodimer and thus to the loss of a

clearly defined complex structure. Consequently, such systems would lose their sequence specificity.

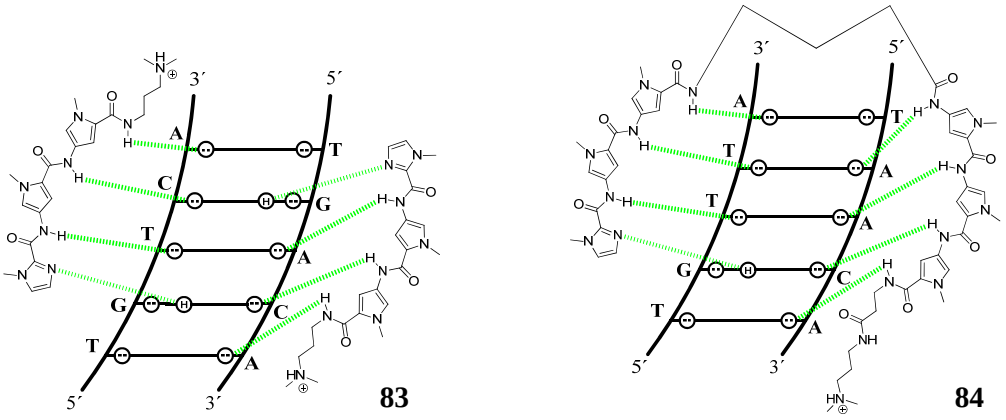


Figure 46. *Ligand **83** forms 2:1 complexes to DNA recognizing GC bps with its Im/Py pair (left). By linking two monomers via a covalent linker (**84**) high affinities without loss of selectivity can be achieved.*[105]

To circumvent this problem *Dervan* linked the ImPyPy monomer head to tail with a distamycin monomer via a γ-aminobutyric acid (GABA) resulting in the hairpin motif (**84**) depicted in Figure 46.[114] It could be shown that very selective and strong DNA complexes were formed. The sequence TGTTA was bound with a binding constant of $8 \times 10^7\ M^{-1}$. Furthermore, it could be shown that selective binding by hairpin ligands is a generally applicable concept. A manifold of different sequences could be bound selectively with high binding constants with tailored ligands, the DNA binder ImPyPyPy-GABA-ImPyPyPy for example binds selectively to AGTACT with $K = 4 \times 10^{10}\ M^{-1}$.[105] However, at this point it was only possible to design ligands which were able to differentiate between CG, GC and AT/TA, but not between TA and AT base pairs.

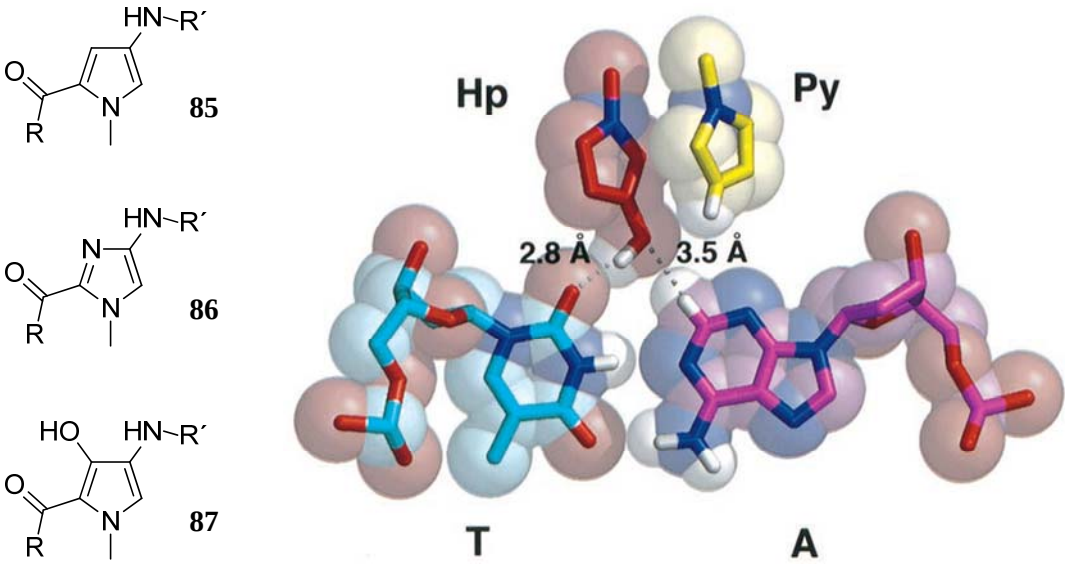


Figure 47. *The three building blocks for sequence selective double-stranded DNA recognition: Py (**85**), Im (**86**), and Hp (**87**) (left) and space-filling model of Hp/Py pair recognizing a TA base pair (right). From Kielkopf, C. L. White, S. Szewczyk, J. W. Turner, J. M. Baird, E. E. Dervan, P. B.; Rees, D. C. Science **1998**, 282, 111-115. The calculated image is reprinted with permission from AAAS.*[115]

In order to differentiate between AT and TA base pairs *Dervan* and *Rees* introduced hydroxypyrrole (**87**, Hp) as additional, rationally designed aromatic building block.[115] A pair of Py/Hp is able to bind to AT but not to TA and the other way round. As shown in Figure 47 the Hp ring forms hydrogen bonds to the thymine O7. Furthermore, the hydroxyl group perfectly fits into the cleft between A and T but sterically clashes with adenine in TA base pairs. With this finding the complete set for sequence selective minor groove binding ligands was complete.[116] For each sequence a tailor-made minor groove binder can now be designed and applied. Table 3 describes the recognition code consisting of pairs of aromatic rings able to selectively bind to one of the four possible base pairs.

Table 3. *Code for sequence selective minor groove binding. Favored (+), disfavored (-).*

Pair	GC	CG	TA	AT
Im/Py	+	-	-	-
Py/Im	-	+	-	-
Hp/Py	-	-	+	-
Py/Hp	-	-	-	+

A crystal structure solving the complex between ImImPyPy with the GGCC core sequence did not only verify that Im/Py forms three specific hydrogen bonds to a GC bp, it also revealed why it was not possible to target sequences longer than four to five base steps with this kind of molecules, although the rise per aromatic ring unit does perfectly match a base pair step.[117] The problem was that the curvature of the ligands is too strong to match the winding of the minor groove. *Dervan* could show that the implementation of β-Ala as substitute for Py rings at certain positions allowed the ligand to adapt to the nucleic acid´s curvature while retaining the specificity for AT base pairs by an Ala/Ala pair. This way it was now possible to target 11-16 base pairs.[118]

With sequence selective recognition motifs at hand *Dervan* then tested the possibility to interfere with regulatory mechanisms in the cell.[119] The hairpin motif ImPyPyPy-GABA-ImPyPyPy binds to AGTACT with a binding constant of $3 \times 10^9\ M^{-1}$. This sequence lies within the 50 bp recognition site of the gene-specific transcription factor TFIIIA, a zinc-finger protein. Binding affinity of the artificial groove binder is higher than for the protein which leads to a reduction of transcription of a specific gene encoded in the region. Thus it could be shown that the hairpin ligands are able to specifically interact with gene regulation in cell culture.

In conclusion, minor groove binding has proven to be an excellent alternative to intercalation, in many cases surpassing their affinities. Very high binding constants, at times as high as protein-DNA interactions, have been reported. Recognition of the minor groove often takes place via shape complementarity between groove curvature and ligand. Furthermore, through a multitude of binding contacts to the nucleobase groove floor and the deoxyribose groove walls excellent specificity and selectivity could be achieved – a feature not present in intercalators. With these prerequisites fulfilled interference with cellular processes like gene regulation became possible. However, not all nucleic acid sequences are accessible that way due to the base pair dependent microstructuring of the helix and packing effects inside the nucleus.

Major Groove Binders

Most sequence-selective DNA binding proteins in nature bind to the major groove where they fulfill their role in for example gene regulation.[120] The broad major groove displays more hydrogen donor and acceptor sites than the minor groove and therefore offer more potential for selective binding. Due to their size proteins are able to make use of this multitude of interaction sites. Affinity is achieved by utilizing electrostatic interactions, van-der-Waals contacts or hydrogen bonding. Normally the secondary polynucleotide structure is not altered upon protein binding. It is rather the proteins that adapt to the helical structure by forming secondary and tertiary structures on their own, e.g. α-helices or antiparallel β-sheets, which are ideally formed for DNA binding. There is a broad variety of protein architectures, like helix-turn-helix, zinc-finger, basic region-helix-loop-helix, ribbon-helix-helix or basic region-leucine zipper (bZIP) motifs. The latter one is exemplarily shown in Figure 48. The homodimeric protein consists of two monomers of approximately 60 residues which each form a continuous α-helix. Their C-termini are associated by forming a coiled-coil with an interface consisting of leucine residues. The N-termini are separated, each one binding to the major groove of B-DNA, but from opposing sides. Amazingly this highly ordered structure only forms upon binding to DNA. Apart from that the protein is unstructured.

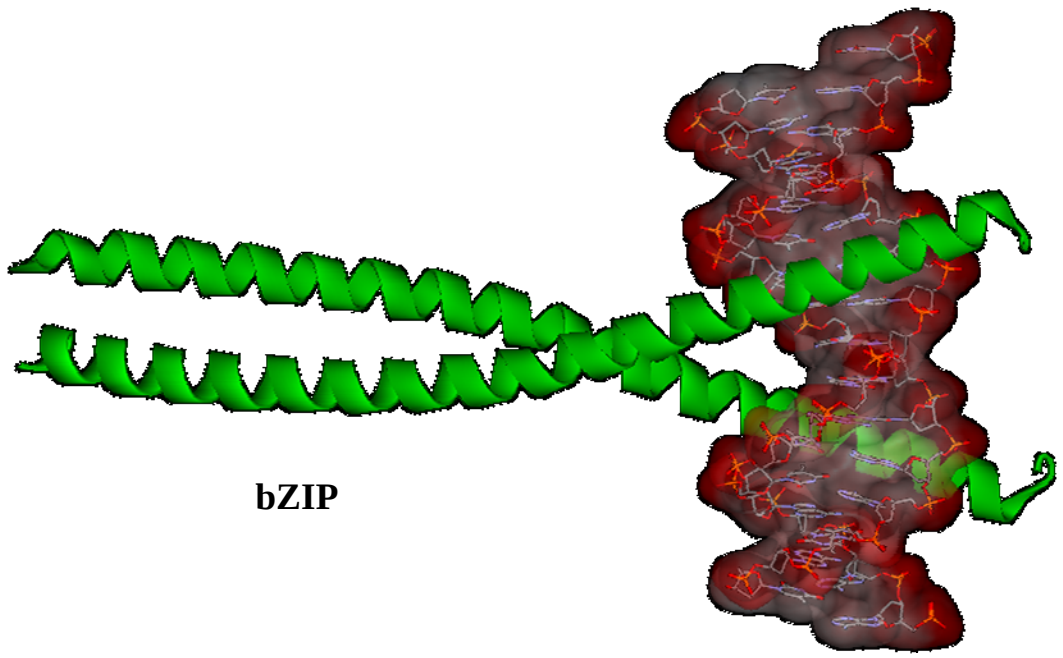


Figure 48. *A basic region-leucine zipper protein (bZIP) is binding to the major groove of B-DNA from both sides via the N-terminus. The C-termini are linked by a leucine interface (PDB access code: 2E42).*

In terms of small molecule binding to the major groove only few examples have been reported so far. Although the major groove is attractive with respect to the sheer number of potential interaction sites, it is also very broad and often shallow and more exposed to the solvent. Small ligands would have to span the widths of the groove in order to bind efficiently to it. Apart from proteins, nature, too, only seldom makes use of major groove binders. One example is for instance the natural product azinomycin A (**88**) depicted in Figure 49, which indeed forms hydrogen bonds in the major groove. However, the main

driving force of DNA binding however stems from intercalation of the naphthalene moiety.[84] Very often such natural products are alkylation agents for DNA adding most commonly to the nucleophilic guanine amine via an epoxide or aziridine.

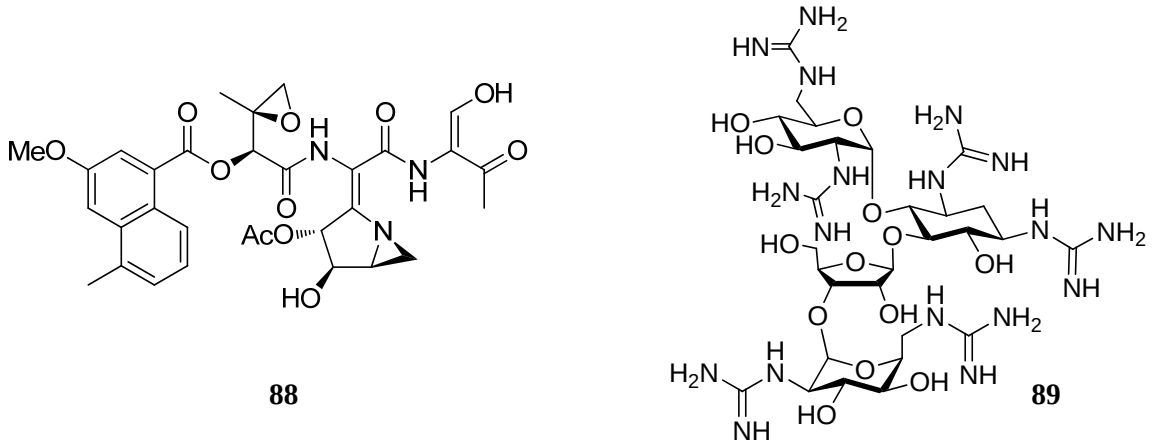


Figure 49. *Azinomycin B (**88**) is an intercalator which forms additional contacts to the major groove (left). Synthetic guanidinoneomycin B (**89**) binds to the major groove of RNA.*

The synthetic guanidinoneomycin B (**89**) synthesized by *Butcher* is shown in Figure 49.[121] It was derived from a natural aminoglycoside antibiotic analogue by replacing all amine for guanidine groups. NMR studies revealed 1:1 binding to RNA major groove with binding constants of approximately 10^8 M^{-1}. The rather large molecule spans the whole width of the major groove. At the edges of the groove the guanidinium groups interact with the phosphate backbone of the nucleic acid.

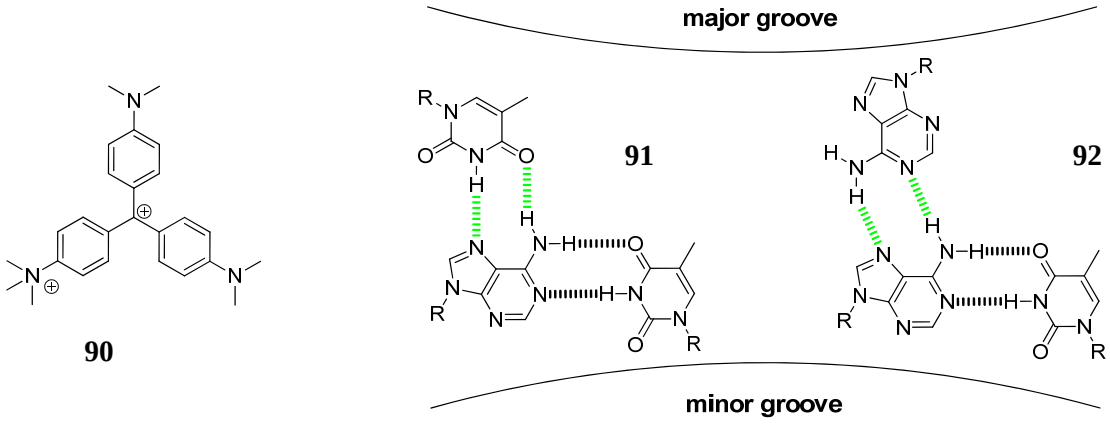


Figure 50. *Left: The dye methyl green (**90**) binds to the major groove of B-DNA. Right: In triple-stranded pdA × $(pdT)_2$ one pdT strand binds to the B-type helix pdA × pdT via Hoogsteen base pairs in the major groove. Depending on the direction of the third arm either TAT (**91**) or AAT (**92**) base-triplets are formed.*

Nordén could show that the dye methyl green (**90**, Figure 50) binds to the major groove of B-DNA.[122] Due to steric hindrances methyl green is not planar and therefore not fit for intercalation. Upon binding its CD signal clearly indicated groove binding. The triple helix pdA × $(pdT)_2$ forms a B-type helix (corresponding to pdA × pdT) with the third strand (pdT) winding alongside the major groove forming *Hoogsteen*-type hydrogen bonds to the

nucleobases. Methyl green binds to pdA × pdT but not to the triple-helix – unlike the minor groove binder DAPI which was tested as control. Thus it could be shown that methyl green binds to the major groove.

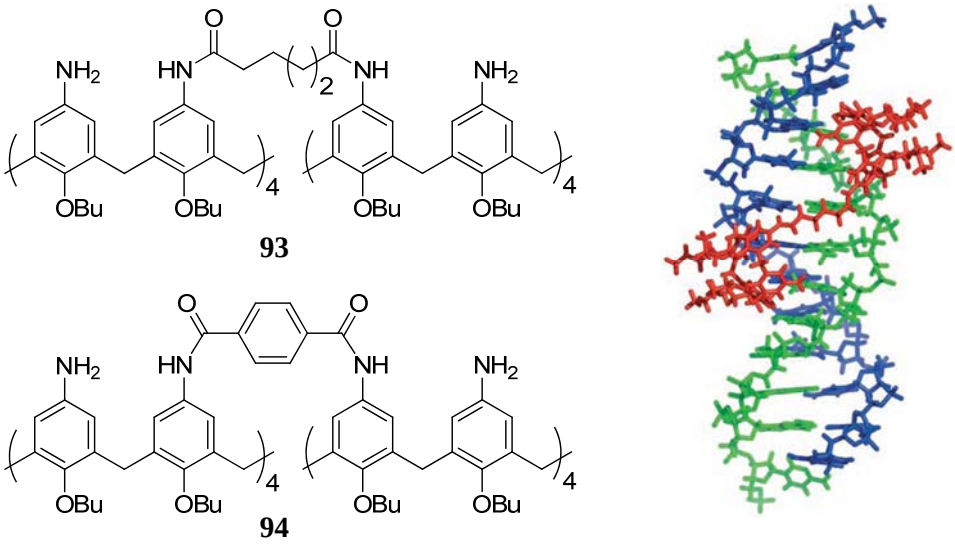


Figure 51. *The flexible anilinocalixarene dimer* ***93*** *(red) binds into the major groove of B-DNA (green, blue), the analogue* ***94*** *is too rigid for efficient binding.*[123]

Schrader reported the major groove binders shown in Figure 51.[123] Two anilinocalixarene moieties were connected with a flexible (**93**) or a rigid linker (**94**) and tested for their binding ability towards B-DNA. Binding constants were as high as 6×10^7 M^{-1} for the ligand with flexible linker, the rigidly linked ligand bound less strongly by two orders of magnitude. Obviously this setup is not flexible enough to allow for adaption to the helix curvature. The rather large size of the dimer perfectly matches the DNA major groove – there is no conformational change of the nucleic acid needed. The ammonium groups form salt bridges to the phosphate backbone and hydrogen bonds to the nucleobases at the groove floor.

In conclusion, major groove binding is a concept often used in nature by proteins for the sequence-selective recognition of DNA. However, it is rather difficult to achieve with small molecules. This is why there are only few reported examples of small molecule major groove binding by natural or synthetic ligands.

Combinatorial Approaches

Artificial DNA binding molecules are most commonly based on modified natural products or derived by rational design. With the help of combinatorial chemistry a much bigger quantity of compounds should be accessible within shorter time.

Nielsen for instance was one of the first to utilize a combinatorial approach for the development of minor groove DNA binders.[124] As shown in Figure 52 the one-bead-one-compound library **95** comprising 5,832 (18^3) members was prepared by means of split and mix standard Boc solid phase peptide synthesis (SPPS). Two consecutive N-methylpyrroles (Py) were linked to three more adjacent Py units via a combinatorial tripeptide consisting of natural and unnatural amino acids (e.g. **96-104**). The N-terminus was derivatized with γ-aminobutyric acid, thus bearing a positive charge at physiological conditions. Screening with the fluorescence labeled ds-DNA GCGTTTGCTTTCGC in neutral phosphate buffer and

consequent synthesis of the hit structure AA^{1-3} = Pip-β-Ala-Tyr and determination of the binding constant in solution revealed a moderate binding constant of 10^5 M^{-1} for the sequence TCTTTGCTTTGG. Not very surprisingly AT rich sequences were preferred due to the presence of at least seven Py units.

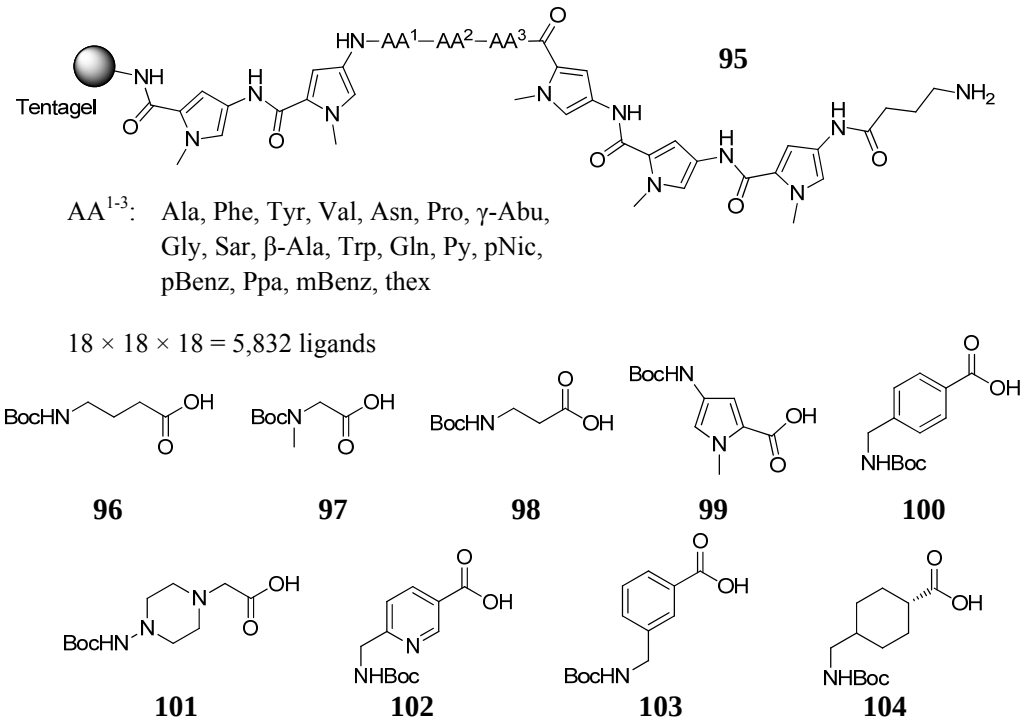


Figure 52. *Combinatorial minor groove ligand library **95**. Boc-protected building blocks γ-Abu (**96**), Sar (**97**), β-Ala (**98**) , Py (**99**), pBenz (**100**), Ppa (**101**), pNic (**102**), mBenz (**103**), and Thex (**104**).*

As shown in Figure 53 *Sasaki* screened the large pentapeptide library **105** with 3,200,000 (20^5) members for its binding affinity towards the target DNA duplex TTTTAAAG.[125] The pentapeptide library was prepared on *Merrifield* resin with two 6-aminocaproic acid units as spacer with the help of standard Fmoc SPPS in a split and mix approach. 20 proteinogenic amino acids were used in the combinatorial positions AA^{1-5}. The DNA was attached to magnetic beads and a mixture of library and nucleic acid was incubated in buffered water at pH 5. Beads with hit structures aggregated with the magnetic DNA beads and could be isolated with the help of a magnet and were then sequenced. The 71 hit structures were mainly built from hydrophobic amino acids (AA^1 = Tyr, Ile, Leu; AA^2 = Ile; AA^3 = Tyr, Gly, Trp, Phe; AA^4 = Glu; AA^5 = Phe, Asn). Several consent sequences were re-synthesized and their binding affinity for several ds-DNA sequences was determined by an EB displacement titration. The best IC_{50} value was found for the amino acid sequence AA^{1-5} = Ile-Ile-Gly-Glu-Phe to the 12mer duplex DNA CGTAGCGCTACG, corresponding to a binding constant of approximately 10^5-10^6 M^{-1}. The affinity of the artificial peptide-based ligand to this sequence is higher than that of distamycin by one order of magnitude.

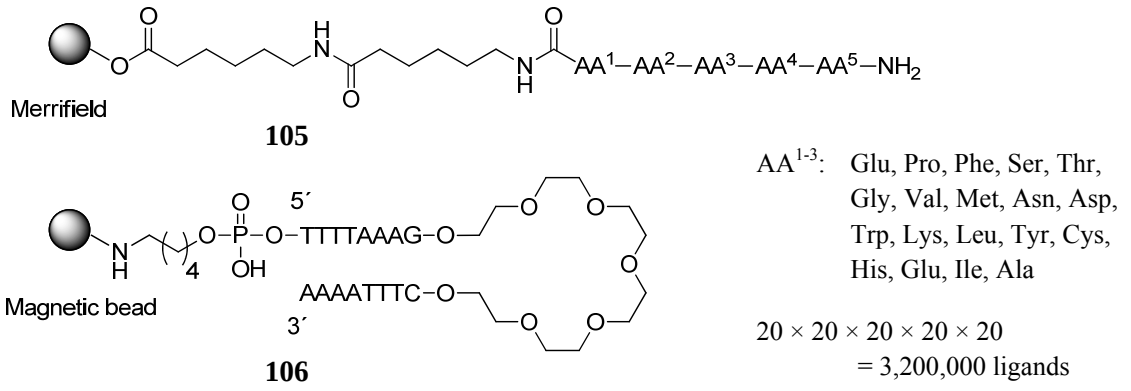


Figure 53. *Pentapeptide library **105** and target DNA sequence attached to a magnetic bead (**106**).*

Lokey designed the bisintercalator library **107** with 360 members depicted in Figure 54.[126] Two lysine-1,4,5,8-naphthalene tetracarboxylic diimide intercalator units were linked via a combinatorial tripeptide. The synthesis was carried out on solid phase by standard Fmoc SPPS preparing 12 mixtures with 30 different sequences. In position AA^2 ten different amino acids were implemented; positions AA^1 and AA^3 either consisted of five different amino acids or were skipped. After cleavage from the resin the library was screened in buffered water at pH 7.5 with a 231 bp duplex DNA with the help of DNase I footprinting. The members of the most promising sublibrary were synthesized and individually tested for their binding ability. Two hit structures could be identified: AA^{1-3} = Gly-Gly-Gly and β-Ala-β-Ala-β-Ala. While the bisintercalator with shorter linker preferred GC rich sequences like for instance TGCCGGTACTGCCGG the second compound preferred non-GC rich sites like AGCTTATCATCGATAAGCT. For the latter combination the overall highest binding constant with $K = 5 \times 10^6\ M^{-1}$ was observed. The different sequence specificity was explained by the different conformational flexibility of the linkers leading to differently oriented and spaced intercalator moieties. Furthermore NMR spectroscopic studies could show that the amino acids in the linkers form hydrogen bonds to the nucleobases in the groove.

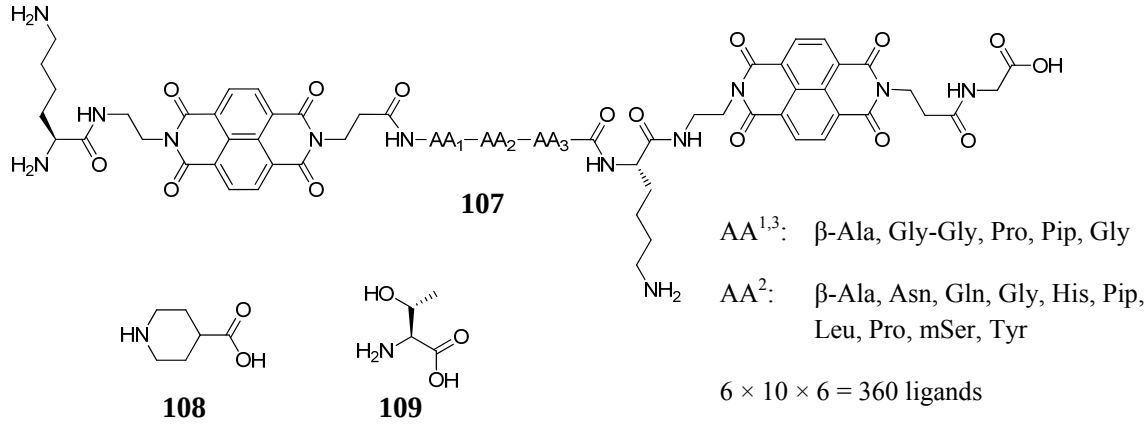


Figure 54. *Bisintercalator library **107** and two of the building blocks: Pip (**108**) and mSer (**109**).*

Boger developed a total synthesis for distamycin (**81**, Figure 43) which allowed him to individually exchange the N-methylpyrrole units of the natural product for other amino acids.[127] By implementing the building blocks **110**-**121** shown in Figure 55 two solution phase combinatorial libraries **122** and **123** were prepared, each one containing the same 1,320 (12 ×

10 × 11) tripeptide members. The two sublibraries only differed in their N-terminus which was either Boc-protected or derivatized with a dimethylaminobutyric acid (DMABA). In contrast to the natural product distamycin, the positive charge was implemented at the N- and not at the C-terminus of the artificial ligands. Both libraries were screened for their binding affinity towards various nucleic acids with the help of a fluorescence-based high-throughput ethidium bromide displacement assay. After incubation of the polynucleotides with EB in 96-well plates the combinatorial compounds were added to the fluorescent solution thus displacing the intercalator. The consequent decrease of fluorescence intensity is directly proportional to the binding affinity of the library members. This way qualitative data for all compounds were accessible. Individual hit structures were selected and quantitative EB displacement titrations were carried out to evaluate the binding constants. In general the Boc-protected, uncharged library performed significantly worse than its positively charged DMABA analogue. The highest affinity for B-type DNA pdA × pdT was achieved for structures containing large subunits at AA^2 (**110-115**) and a small subunits at AA^3 (**116-121**). The overall highest binding affinity was found for compounds **124** and the distamycin analogue **125** with $K = 6 \times 10^6\ M^{-1}$, which is as high as for the natural product. As expected, the affinity towards pdG × pdC was significantly lower compared to pdA × pdT for **124** and even lower for **125**. To fully characterize its sequence specificity, the new binding motif **124** was then screened against a library of 512 hairpin 21-mer oligonucleotides. It could be shown that purine-pyrimidine-pyrimidine sequences were preferentially bound. The distamycin analogue on the other hand prefers AT base pairs just like the natural product. Thus it could be demonstrated that with the help of solution phase combinatorial chemistry the sequence selectivity of a given nucleic acid ligand can be altered while retaining the high binding affinity of the corresponding natural product.

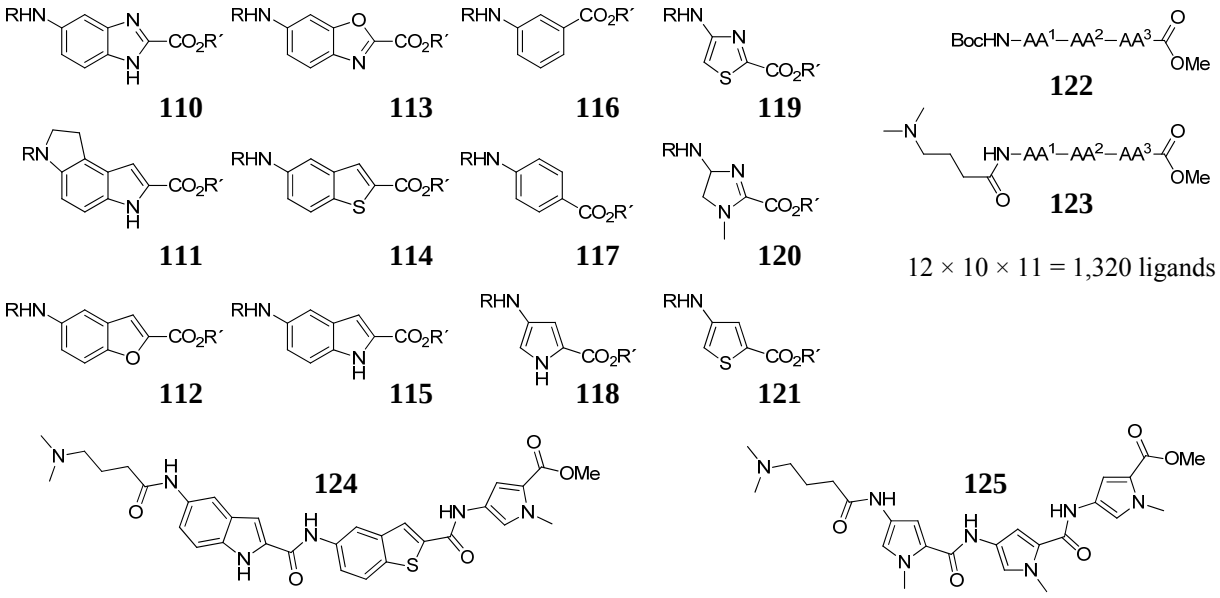


Figure 55. *The two libraries **122** and **123** each contain 1,320 members derived from the building blocks **110-121** and only differ in their N-terminus. The best results were achieved with **124** and the distamycin analogue **125**.*

This approach was further exploited by *Boger* by screening a library of > 9,000 synthetic DNA binding molecules for their ability to inhibit the transcription factor lef-1.[128] Figure 56 shows the two most efficient sequences **126** and **127** which bound to the transcription factor´s target sequence CCTTTGATC with a binding constant of $K = 9 \times 10^6$ M^{-1}. The IC_{50} values for the inhibition of lef-1 were 6.2 and 0.15 µM, respectively. Hence, with the help of combinatorial chemistry it is possible to find sequence specific DNA ligands with the ability to efficiently down-regulate gene expression.

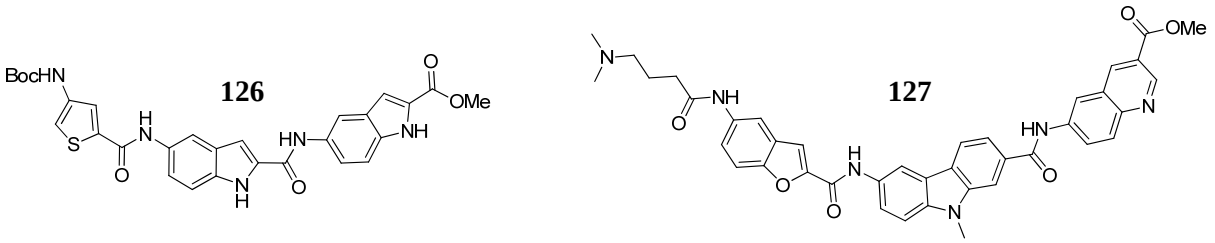


Figure 56. *The two lead compounds **126** and **127** were derived by a combinatorial library comprising > 9,000 members. They are efficient inhibitors of the transcription factor lef-1.*

Combinatorial chemistry has only rarely been utilized for the development of DNA-binding molecules. The few examples reported in literature for library-derived intercalators or groove binders were only moderately successful in finding more advanced ligand systems. These libraries however, are mainly large and random – their screening only gives positive answers for hit structures but neglects the majority of library compounds. Therefore much of the information is lost. By generating a small but focused ligand library with biased building blocks which can then be screened in a fashion that gives quantitative or qualitative information about all possible compound-DNA binding affinities a more successful ligand optimization could be possible.[30]

2.4 TRANSFECTION

2.4.1 Gene Therapy

The human genome is encoded in a long ds-DNA in the cell nucleus consisting of approximately 3 billion bps which are arranged in ca. 25,000 genes – the smallest functional units containing genetic information.[42] The structure of a eukaryotic gene and the read-out mechanism of the genetic information (transcription) are depicted in Figure 57. The gene itself consists of regions which are transcribed to RNA and regulatory elements. The promoter region at the 5′-end of the gene in the DNA antisense strand marks the starting point for transcription and is recognized by a protein – the transcription factor. Enhancer and silencer regions either increase or decrease gene expression. These elements are found near the promoter region, but are not necessarily in the same strand. Regulatory elements far away in the strand may still come within close proximity of a given promoter due to supercoiling of the DNA helix. Once the starting point for the read-out is set, the DNA unwinds by ca. one winding and an RNA polymerase enzyme generates an RNA from 5′- to 3′-end, which is complementary to the antisense-strand. This premature messenger RNA (pre-mRNA) undergoes a maturation process upon which introns are cut off the strand and only extrons remain (splicing). During this processing a 5′-cap and a 3′-poly(A) tail are attached to the RNA for protection against RNases. The processed mRNA now consists of three parts: the open reading frame (ORF) contains the information for the translation of the nucleic acid into the corresponding protein at the ribosome. Three nucleobases form one codon – the code for a specific amino acid. The remaining parts of the mRNA, the 3′- and 5′-untranslated region (UTR), serve for the regulation of translation.

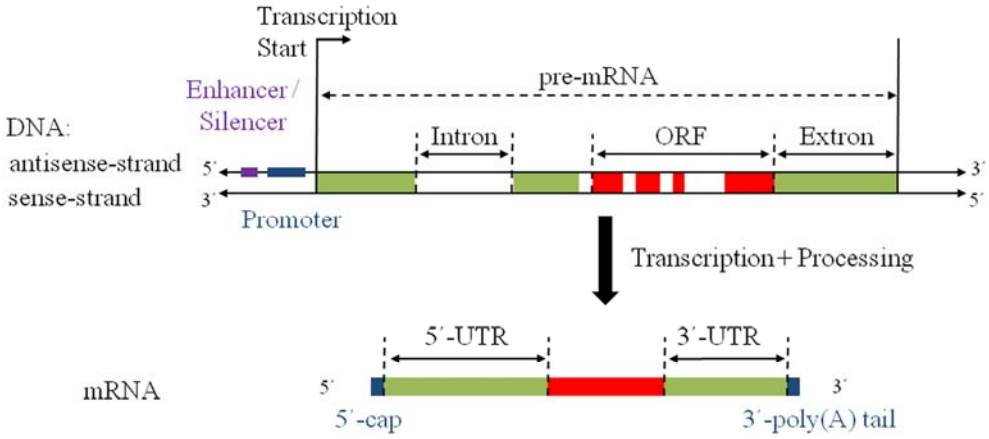


Figure 57. *Schematic illustration of the transcription process in the nucleus.*

Since the human genome was fully decoded in 2003 more and more functions of genes have come to be known.[15] The DNA´s nucleobase sequence consists of parts which are conserved and others which can be different for different persons. Differing genes which are carried by > 5 % of the population are called polymorphisms, those carried by less than 5 % are mutations. There are different variations, e.g. single nucleotide exchange, duplications, insertions, translocations, inversions, deletions or different copy numbers, which may either be part of the heritable information or arise from environmental influences during life.

Mutations in extrons coding for proteins normally lead to a drastic decrease of gene activity. In introns or promoter regions it is the other way round. These changes may either be silent, advantageous or disadvantageous for the cell functionality. In the last case dysfunctional cells and diseases can be the result.

While classical medicine introduces drugs from the outside into the cell, gene therapy strives to modify the genetic information inside the cell and thus to alter the biosynthesis as a drug from inside.[129] There are different approaches to do so: replacement therapy exchanges or corrects defective / missing genes, e.g. in monogenic hereditary diseases like the severe combined immunodeficiency (SCID).[130] Enhancement therapy amplifies or introduces genetic function (e.g. for cancer treatment) and suppression therapy turns off or prevents pathogen gene activities (like a vaccine).[131] In principal it is possible to introduce a gene individually adjusted to a person and/or to a disease.[132] However, this approach is nowadays limited to diseases which are based on a single gene defect. Unfortunately, most diseases are due to a combination of many polymorphisms and/or mutations. Up to this day it is not possible to treat multigene-derived diseases. Nevertheless, gene therapy is a powerful emerging field for the treatment of maladies which cannot be cured with classical small molecule drugs and thus opens a whole new field of medical treatment. One of the most important prerequisites for gene therapy is an effective gene transfer into the cell because it is difficult to introduce DNA directly into the cell due to its physical properties: The overall negative charge of the nucleic acid is repulsed by the negatively charged surface of the cell. Furthermore it rapidly degrades in biological surroundings due to the activity of RNases and DNases. Therefore appropriate transportation vehicles (= vectors) are necessary. These should be easy and cheap to produce in high quantities in a reproducible manner. They must enable uptake into the cell, ideally in a specific fashion. Inside the cell the genetic information has to be firmly integrated into the genome without interfering with the genetic information already inside. Last but not least vectors must be safe for application with regard to pathogenicity, cytotoxicity, adverse effects and immune responses. Nowadays, three main classes for the delivery of nucleic acids exist: physical, viral and chemical vectors.

2.4.2 Viral Vectors

First tries to shuttle genes into cells were carried out with the help of viruses in the 1970s.[133] Viruses are naturally very efficient in infecting cells and in implementing their genetic material into the host cell for the sake of self-replication. For the use as vectors their genome has to be modified:[17] The desirable functions like cell recognition, packaging and delivery of genes into the host genome are kept. Troublesome parts of the virus´ genetic information like replication, pathogenicity, or expression of immunogenic viral antigens are deleted and replaced by the gene of interest together with the necessary transcriptional regulatory elements (= transgene). These recombinant viruses are now able to transduce (= delivery + expression of a gene) the cells they would normally infect. Most commonly adeno- and retroviruses are utilized as viral vectors nowadays, and are until now the only way to permanently heal monogenetic diseases. Because viruses practically evolved as gene-delivery vehicles, they are extremely efficient with a good cell type selectivity, which may even be altered by modifying their surface glycoproteins. Therefore they are broadly applied and are the most advanced vector system available. Unfortunately, viral vectors suffer from several severe inconveniences: the immune system is still able to recognize the viral surface and thus an immune response is triggered which becomes worse with every treatment. Further, although the dangerous parts of their genome are removed before usage there is potentially the threat of a recombination with a wild type virus leading to pathogenicity and mobilization of the genetically altered virus. The new information is now able to move to other cells or persons. Insertion of the delivered genes into the genome is not site-directed and may activate for example oncogenes. The overall length of the gene which is to be delivered into the cell is restricted by the original length of the viral genome. Acceptance for the treatment with viral vectors in public is generally low. Lastly, the up-scaling for mass production on an industrial scale is very difficult.

2.4.3 Physical Vectors

An alternative for viral gene delivery are physical vectors.[134] Genes are directly delivered into the cells without the need for a vector. In order to cross the cell membrane several techniques have been applied. Electrophoresis and ultrasound for example use electrical or sonic pulses to increase the permeability of the cell membrane. Jet injection locally injects a DNA solution with high pressure so that the microdroplets penetrate the skin and the cell wall. Similarly, a gene gun shoots DNA coated heavy metal nanoparticles which directly penetrate through the cell membrane. A clear disadvantage of this method is the deposition of heavy metal in the body. Still, in general physical vectors are safer, not cytotoxic and allow for repeated administration. There is no limitation with regard to the length of DNA. Unfortunately there are several inconveniences greatly limiting these methods. They are not specific for cells and there is no inherent cell penetration mechanism. The expression level therefore is generally only low and the area of application is very limited. Only tissues at the surface like the skin are accessible without surgery.

2.4.4 Chemical Vectors

The third alternative for gene delivery is the application of chemical vectors.[135] Chemical delivery systems are positively charged molecules which form complexes with the negatively charged cargo thus neutralizing its charge and sterically shielding the ribophosphate backbone from degradation by nucleases. Unprotected plasmid for example is degraded within minutes by DNases, complexed with chemical vectors it is stable for hours. The now positively charged particles interact with the negatively charged cell surface. Unlike small molecules like amino acids or ions which can cross the cell membrane via integral protein pumps or channels the uptake of nucleic acid-chemical vector complexes usually takes place via endocytosis.[136] During this process membrane bound vesicles form around the cargo by invagination and pinching off of pieces of the plasma membrane. These cell own vesicles (endosomes) are then released into the cytosol. There are various endosomal pathways which lead to the formation of vesicles with diameters between 60 and 120 nm thus restricting the maximum size of the cargo. The next step is the liberation of the nucleic acids from the endosomes to the cytosol. The endosomal escape is a major obstacle for chemical vectors. If it fails the polynucleotides are rapidly degraded by nucleases at low pH in the late endosomes/lysosomes.[137] Once in the cytosol the gene has to be transported to the nucleus and integrated into the DNA. In general, chemical vectors are simple to prepare reproducibly in large quantities and offer improved safety compared to viral vectors. The main classes of chemical vectors are polyplexes, lipoplexes, and cell penetrating peptides.

Polyplexes

Cationic polymers have shown to form strong complexes to nucleic acids and hence to contract them to multimolecular aggregates called polyplexes. The overall size of these structures is then rather linked to the polymer size and not to that of DNA. All polyplexes form toroidal or spherical structures of similar diameter. The exact aggregation depends on the polymer structure.[18] In general, polycations are potential gene transfer agents.[138]

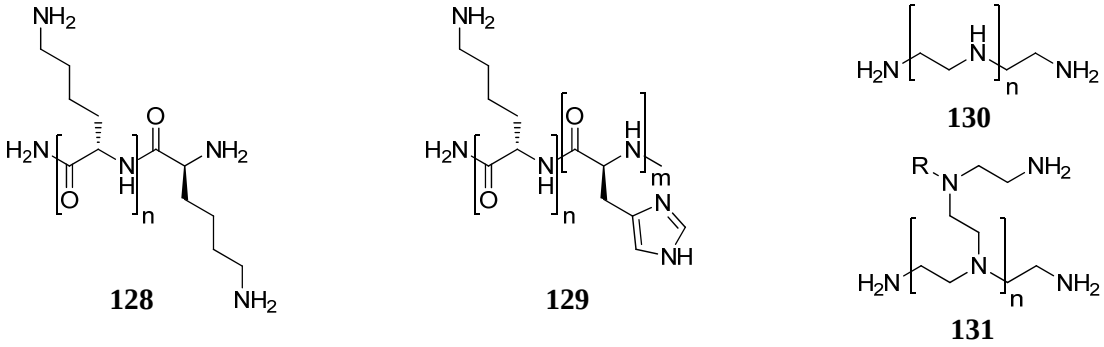


Figure 58. *Poly(L-lysine)* ***128****, PLL derivatized with L-histidine (****129****), and linear (****130****) or branched (****131****) polyethyleneimine (PEI).*

In 1975 *Laemmli* could demonstrate that poly(L-lysine) (**128**, PLL, see Figure 58) is able to strongly condense DNA.[139] The size of the polyplexes correlates with molecular weight. Large polymers (224 kDa) form aggregates of 100-300 nm diameter and small PLLs lead to the formation of 20-30 nm sized condensates. *Wu* could show later on that PLL is able to transfect cells in vitro as well as in vivo.[140] A minimum molecular weight of 3 kDa is necessary for transfection. However, most of the PLL-polyplexes are not able to escape the endosome after cellular uptake and are thus degraded in the lysosome. This is also why PLL

is a rather poor transfection agent and efficiencies are only high for cells undergoing mitosis. During the division process cells are in general more susceptible for transfection due to the general redistribution and partial breakdown of the nucleus wall. By partly replacing lysine groups for histidine (pK_A = 6) the transfection efficiency could be increased.[141] This difference can be explained by the proton sponge effect as shown in Figure 59. At physiological pH all primary lysine ε-amino groups but only very few histidine side chains are protonated. After endocytosis the pH within the endosomes is constantly reduced from 7.4 to 4-5 when going from endosome to the lysosome by a specific ATPase.[142] The weakly basic imidazole rings of the histidine amino acids are getting protonated at this pH and thus act as a buffer within the endosomes. Consequently more protons are pumped into the endosome accompanied by counterions and water. The increased osmotic pressure leads to vesicle swelling and finally to their rupture. The enclosed nucleic acid content is released into the cytosol before it can be degraded by nucleases in the lysosome. Another big disadvantage of PLL is its high cytotoxicity. This is a general problem of cationic species and especially of cationic polymers. They tend to aggregate and precipitate under the high salt concentrations (150 mM) found at physiological conditions. Furthermore cationic complexes bind to serum proteins like serum albumin, which hinders cellular uptake and further promotes aggregation. The higher the molecular weight and cationic charge, the higher is the toxicity. The cytotoxicity can be reduced by integrating polyethyleneglycol (PEG) chains into the polymer.[143] However, due to their rather poor efficiency it is rather unlikely that PLL-derived vectors will be used for clinical applications in the future.

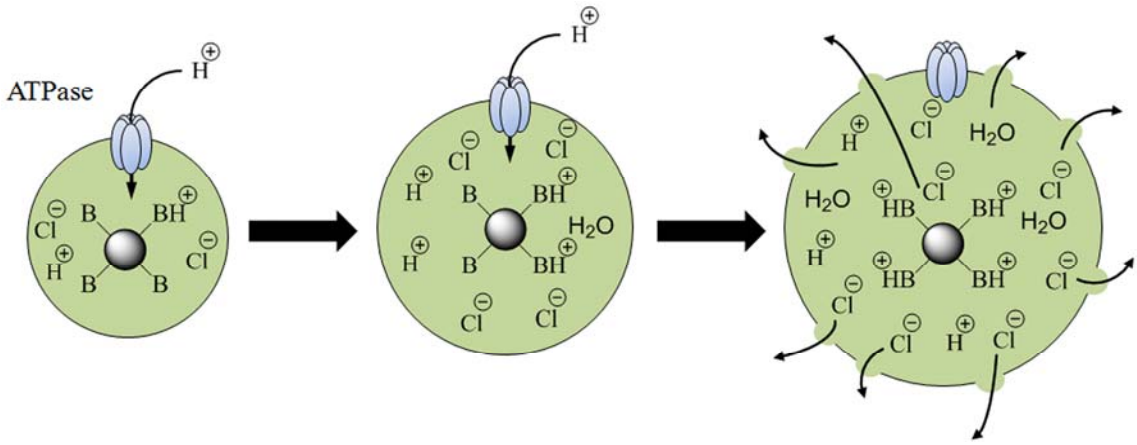


Figure 59. *Schematic illustration of the proton sponge effect. Protonation of the polyplex leads to increased influx of proton accompanied by counterions and water into the endocytic vesicles. Ultimately the vesicle bursts.*

Nowadays the most commonly used agent for cell transfection is polyethyleneimine (PEI) either as linear (**130**) or branched (**131**) polymer as shown in Figure 58. It was first utilized by *Behr* in 1995.[144] It is very efficient due to its great buffer capacity. In branched PEI for example every third atom is a nitrogen atom. At physiological pH only approximately 20 % of the amine residues are protonated. In the endosome at pH 5 more than 50 % are protonated.[145] Hence, PEI is an excellent proton sponge and allows for efficient endosomal escape.[146] Amongst polymeric substances PEI is nowadays the gold standard for in vitro and in vivo cell transfection. Linear as well as branched PEI is commercially available. Linear PEI seems to

be less efficient in condensing DNA and affords less stable polyplexes. The efficacy of transfection also correlates with its molecular weight (M_W). The higher the M_W, the higher the percentage of transfected cells. Unfortunately, cytotoxicity similarly correlates with weight. The larger the polymers the more they adhere at the cell surface, which then leads to necrosis.[147] Like PLL, highly cationic PEI easily aggregates under physiological salt concentrations and binds to serum proteins. Thus the optimal molecular weight for transfection experiments is a compromise between these two trends – a M_W between 5 and 25 kDa is regularly chosen for transfection experiments.

In conclusion cationic polymers have shown to be capable of transfecting cells in vitro and in vivo. There are still several problems like e.g. cytotoxicity or the lack of inherent cell-targeting that need to be solved. Further, polyplexes are still several orders of magnitude less efficient than viral vectors. Therefore there is still need to develop novel materials to overcome the intra- and extracellular obstacles which prevent higher efficiencies.

Lipoplexes

Another major class of gene carriers is that of cationic lipids, which are able to fuse with the cell membrane.[148] In 1987 *Felgner* was the first to utilize cationic lipids for gene delivery.[149] In general cationic lipids consist of three structural domains: a cationic headgroup, a hydrophobic hydrocarbon and a linker in between.

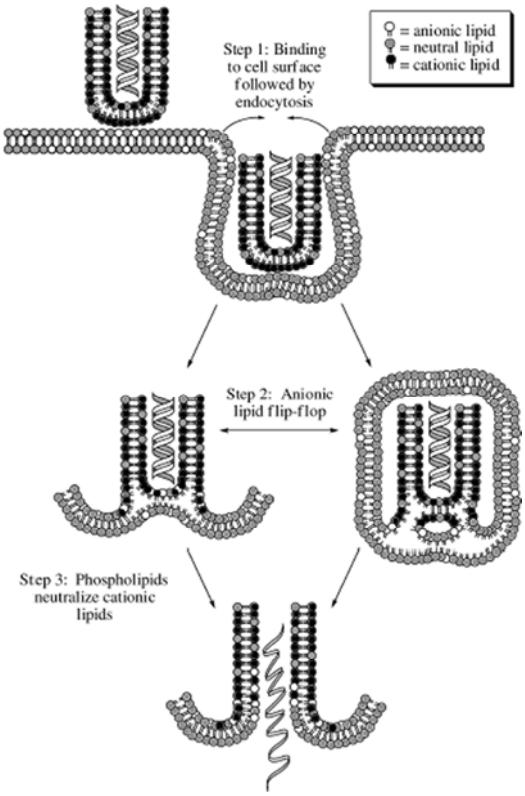


Figure 60. *Scheme of cellular uptake mediated by cationic lipids. Reprinted with permission from Mintzer, M. A.; Simanek, E. E. Chem. Rev.* ***2009****, 109, 259-302. Copyright 2009 American Chemical Society.*[18]

The first step of the lipofection mechanism is the association of the cationic headgroup with the ribophosphate backbone of the nucleic acid thus forming a supramolecular lammelarly

structured lipoplex.[18] During this process the nucleic acid is condensed. As depicted in Figure 60 the positively charged complexes recognize the negatively charged cell surface and cellular uptake then primarily takes place via endocytosis. Once inside the endosomes the positively charged lipoplexes interact with the negatively charged phospholipid-based vesicle wall, which is destabilized as a consequence. The result is a flip-flop like reorganization of the vesicle upon which the lipoplex is released into the cytoplasm. Since lipofection is very efficient there are several commercial compounds available for application in vitro and in vivo – some exemplary structures (**132-135**) are depicted in Figure 61. Especially **135**, which is better known under its commercial name lipofactamine™ is today often referred to as the gold standard for transfection with chemical vectors. Lipid-based vectors are the most utilized and best studies chemical vectors and already were applied in many human gene-therapeutic clinical trials.

Many studies concerning the structure-activity relationship have been undertaken.[18,150] Two general trends could be observed: the nature and density of the cationic headgroup has a strong, predictable influence on the efficiency, but for a given headgroup the outcome of replacing the hydrocarbon chain is not predictable. For instance, by replacing the ammonium groups by more biocompatible groups like pyridinium groups, which are able to better delocalize the charge, the transfection and cytotoxic properties could be improved. Multivalent headgroups are more efficient if the charges are well spaced. Small distances between ammonium groups lead to lower efficiencies. This is probably due to a decrease of pK_A for high charge density headgroups and thus an overall reduced charge. The better the charges are distributed in multivalent headgroups the better they are able to transfect cells. Generally the trend goes: the more charges, the higher the efficiency. Too many charges on the other hand may lead to stronger folding of the lipid conformation and a decrease of gene delivery. Variations of the hydrophobic group on the other hand did lead to different results for different head groups. Sometimes longer chains gave improved and sometimes poorer results. Finally, the replacement of the linkers for biodegradable moieties which are degraded after endosomal escape (e.g. carbamates) and hence facilitate the liberation of the nucleic acid from the lipoplex have demonstrated to improve gene delivery.

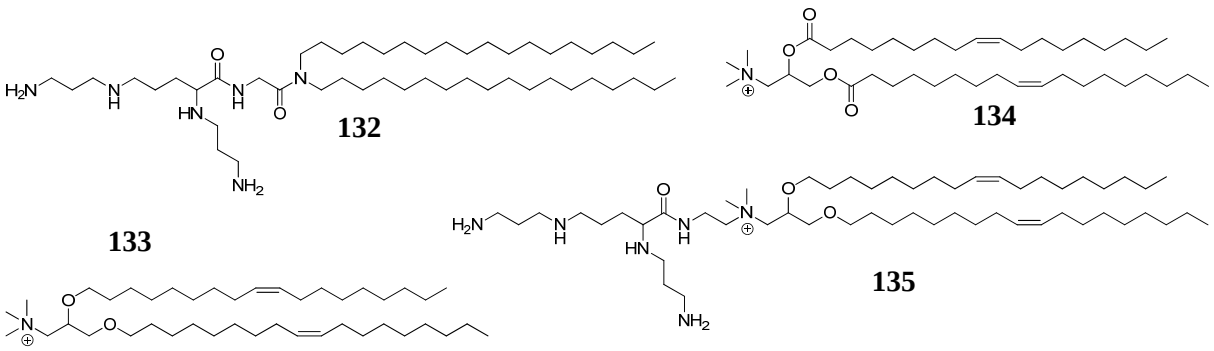


Figure 61. *Commercial cationic lipids DOGS (**132**), DOTMA (**133**), DOTAP (**134**), and DOSPA (**135**).*

Cationic lipids with two headgroups and two hydrophobic moieties which are bridged via a linker are called gemini surfactants. *Engberts* carbohydrate-based gemini surfactant **136** (Figure 62) for example is more efficient than lipofectamine.[151] At low pH these molecules

undergo a morphological change from lamellar to an inverted hexagonal structure. This change leads to endosomal fusion and very efficient liberation of the nucleic acid into the cytosol.[152]

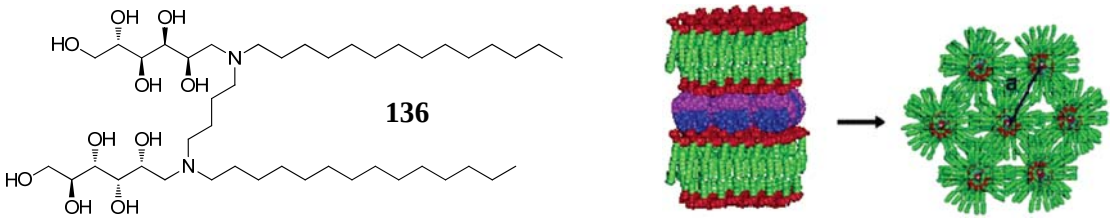


Figure 62. *Gemini-surfactants like the sugar-derived* ***136*** *undergo a morphological change from lamellar to hexagonal structure at low pH thus triggering endosomal escape. The modeled image is reprinted with permission from Bell, P. C. Bergsma, M. Dolbnya, I. P. Bras, W. Stuart, M. C. A. Rowan, A. E. Feiters, M. C.; Engberts, J. B. F. N. J. Am. Chem. Soc.* ***2003****, 125, 1551-1558. Copyright 2003 American Chemical Society.*[152]

Cationic lipids have shown to be very efficient delivery agents for the transfection of cells with genetic material. Even though they are not as effective as recombinant viral vectors they are the best known chemical vectors to this day. Together with the polyplex-forming PEI the cationic lipid lipofectamine is nowadays the most utilized chemical vector for in vitro and in vivo application. However, there are still some problems connected to cationic lipids. First, the reproducible fabrication of the liposomes and of the lipoplexes is not without trouble and strongly depends on the method of preparation and the order of preparation-steps as well as the surrounding medium. Further, the stability of the colloidal mixture is not unproblematic. Finally, cationic lipids are often cytotoxic due to their cationic nature.[153] Still, they offer a very good platform for the development of safe and efficient chemical vectors and further development of their features is desirable.

Cell Penetrating Peptides

A more recently developed class of non-viral vectors is that of cell penetrating peptides (CPPs).[154] These protein transduction domains are short (10-30 residues), basic amino acid-rich sequences originally derived from proteins which are able to cross the cell membrane. In 1988 *Frankel* and *Pabo* found that the transcription-transactivating (Tat) protein of HIV-1 is able to enter cells and translocate to the nucleus (Figure 63, bottom left),[155] and in 1991 *Prochiantz* could demonstrate that the *Drosophila* antennapedia homeobox peptide internalizes into neuronal cell (Figure 63, top left).[156] A few years later, in 1994 the same group utilized the 16mer peptide penetratin (**137**) derived from the third helix of the antennapedia peptide (residues 86-102: Arg-Gln-Ile-Lys-Ile-Tyr-Phe-Gln-Asn-Arg-Arg-Met-Lys-Trp-Lys-Lys) as first cell penetrating peptide.[157] In 1998 *Lebleu* could show that the minimal Tat-derived peptide sequence for cellular uptake is the 11mer **138** (residues 47-57: Tyr-Gly-Arg-Lys-Lys-Arg-Arg-Gln-Arg-Arg-Arg).[158] When attached covalently to a nucleic acid these CPPs were able to efficiently transfect cells. These findings led to an exploding search for other CPPs by protein screenings and rational design. Today, a broad variety of naturally derived and synthetic cell penetrating peptides is known.[160] In 1997 the first non-covalent CPP for the delivery of polynucleotides was introduced by *Divita*.[161] This chimeric peptide of 27 amino acids length contains a hydrophilic, lysine-rich C-terminal domain derived from the nuclear localization sequence of the SV40 large T-antigen (Pro-Lys-Ser-Lys-

Arg-Lys-Val). Via a tripeptide linker a variable, N-terminal hydrophobic group derived from the fusion sequence of the HIV protein gp41 (Gly-Ala-Leu-Phe-Leu-Gly-Phe-Leu-Gly-Ala-Ala-Gly-Ser-Thr-Met-Gly-Ala) is attached. The peptide forms strong complexes with polynucleotides via electrostatic interactions between its positively charged domain and the phosphates. Similar to poly- and lipoplexes the DNA is compacted. As a result of this general feature of non-covalent CPPs the stability of the nucleic acid against nucleolytic degradation is increased. The first contact to the cell is formed upon electrostatic interaction with the negatively charged proteoglycans at the cell surface. Although the exact mechanism of the cellular uptake is somewhat controversial it follows an endocytotic pathway in most cases. Transfection efficiencies are usually high, similar to lipoplexes. A big advantage is the lack of cytotoxicity of these rather small, naturally derived gene carriers.

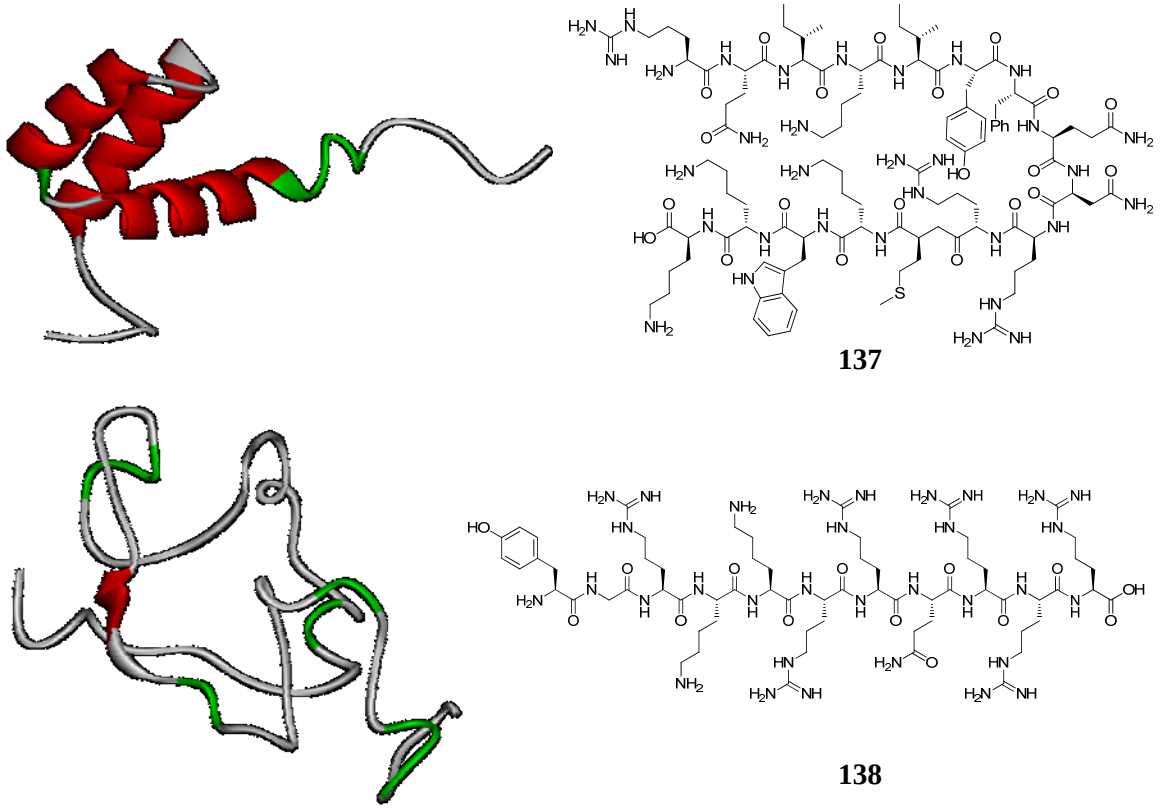
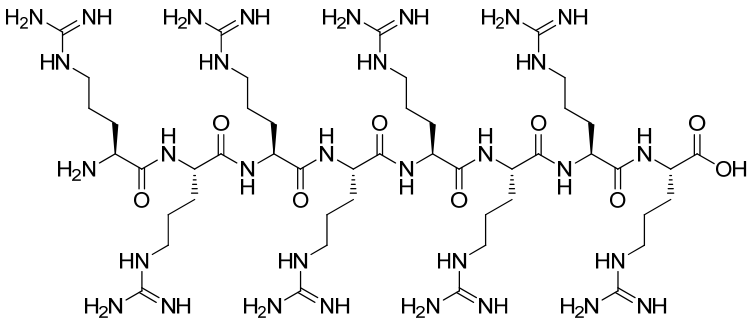
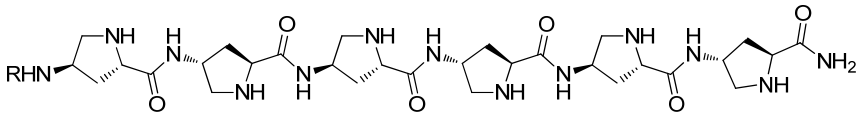


Figure 63. *NMR structure of the cell-penetrating homeobox protein from Drosophila antennapedia and the minimum sequence for cell penetration* ***137*** *(top; PDB access code: 1HOM).*[159] *NMR structure of the HIV-1 Tat protein and the corresponding minimum sequence for cell penetration* ***138*** *(bottom; PDB access code: 1TBC).*

A prominent example for artificial CPPs are oligoarginine compounds like Arg_8 (**139**, Figure 64) developed by *Futaki*.[162] Efficient internalization proceeds via endocytosis, but only if a minimum of six arginine residues is present. The transfection efficiency peaks between six and eight residues to decrease again for longer oligoarginines. The structure of the gene carrier has no influence on the efficiency: both linear and branched oligoarginines have the same efficacy. Uptake takes place via endocytosis and endosomal escape is the limiting factor for more successful efficiencies. The guanidinium group seems to play a special role prior to cellular uptake by recognizing the phosphate in the cell membrane via a bidentate hydrogen bonded salt bridge.



139



140

Figure 64. *Cell penetrating peptide Arg_8 (**139**) and octameric cis- γ -aminoproline **140**.*

A general problem of CPPs is the rather low stability for in vivo applications. Under physiological conditions they are rapidly degraded by proteases due to their peptidic nature. By implementing non-natural amino acids their proteolytic stability was increased while their ability to transduce cells was retained. *Albericio* and *Royo* for instance devised oligo-*cis*-γ-amino-proline CPPs like **140** as depicted in Figure 64.[163] Although these compounds are slightly cytotoxic they proved to be capable of cell transfection and were not degraded in human serum.

In conclusion cell penetrating peptides are a new and innovative class of gene carriers which offer great potential for efficient delivery of nucleic acids into cells. Hence, they have been frequently utilized for in vitro and in vivo applications with several ongoing clinical trials. CPPs are 10 to 30 residues long, naturally derived or rationally planned and rich in lysine and/or arginine amino acids and can be easily prepared by standard SPPS. So far, a minimum of six charges is necessary for efficient gene delivery. The major limitation is endosomal escape. In general, they are less toxic than cationic lipids or polymers. However, when not derivatized with non-natural moieties they are metabolically instable. Due to their peptidic nature a combinatorial approach for the development of even more efficient vectors should also be feasible.[164-166]

.

3. PROJECT AND OBJECTIVES

The focus of this dissertation lies within the field of molecular recognition of biologically relevant targets. The scope of the work ranges from the development of new methods and the design of novel model systems for studying supramolecular binding events to application in cell biology.

Method Development: Raman Spectroscopy

The first project is based on previous results of our working group concerning the molecular recognition between the artificial receptor GCP-KKF (**13**) and tetrapeptides (**141**) as shown in Figure 65. This host-guest system forms stable complexes in aqueous medium and is an ideal model system for the development of new Raman spectroscopy-based techniques for studying the interactions at the molecular level. This vibration-based method is in theory able to determine binding constants and to analyze at the same time which parts of the molecules interact with each other, but the method suffers from the weak intensity of the Raman signals. To circumvent these problems new approaches will be developed together with our cooperation partner, the working group of *Prof. S. Schlücker* at the Institute for Physics at the University of Osnabrück. In the first approach we will test whether it is possible to utilize Resonance Raman spectroscopy for the determination of binding constants. Therefore, the interaction between **13** and several anionic tetrapeptides will be studied. In previous work it could already be shown that Resonance Raman is able to give qualitative information for this binding event.[167] Furthermore, the binding constants are already known thus rendering this host-guest pair the ideal model system.[29]

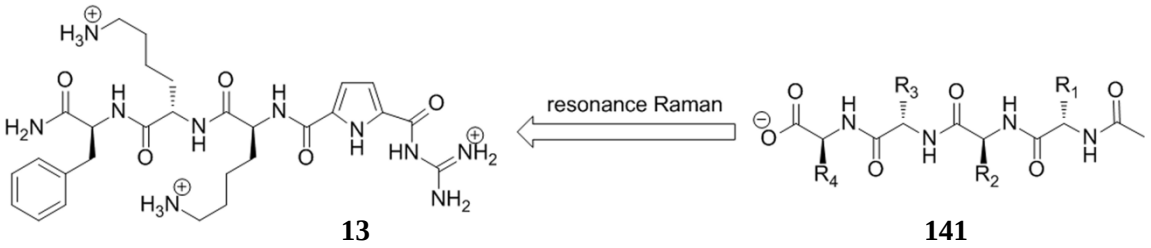


Figure 65. *Resonance Raman studies will be applied for the determination of the binding constant between receptor **13** and tetrapeptidic stubstrates **141**.*

The second Raman project, which will be attended to together with our cooperation partner, will follow an entirely different approach. In previous work it could be shown that the above mentioned receptor **13** can be visualized on bead with the help of surface enhanced Raman spectroscopy by incubating the resin with silver nanoparticles.[168] We will now try to develop a more sophisticated solid-support with integrated noble metal nanoparticles which would allow for direct and rapid detection of solid-bound substances on bead – a task which is to this day extremely difficult to achieve.

Model Receptors for Oligopeptides and Nucleotides

The second project of this thesis is based on the results of my diploma thesis, in which several divalent model receptors could be synthesized either with identical or with different arms.[195] During this doctoral work the design of these systems will be optimized in order to improve their physicochemical properties for substrate binding – especially their pK_A values, which were too low for efficient binding in water at physiological pH. The non-symmetric tweezer receptors **142** shown in Figure 66 will consist of two differently substituted peptidic side chains. One of these arms will carry a guanidiniocarbonyl pyrrole (red) unit as oxoanion binding motif. These model systems will then be tested for their binding properties towards the tripeptide **143**, which plays a decisive role in the biosynthesis of bacterial cell wall. The GCP unit is responsible for the recognition of the C-terminus of the peptide and the amino acids (AA^1-AA^4) in the two side chains allow for additional interactions with the substrate.

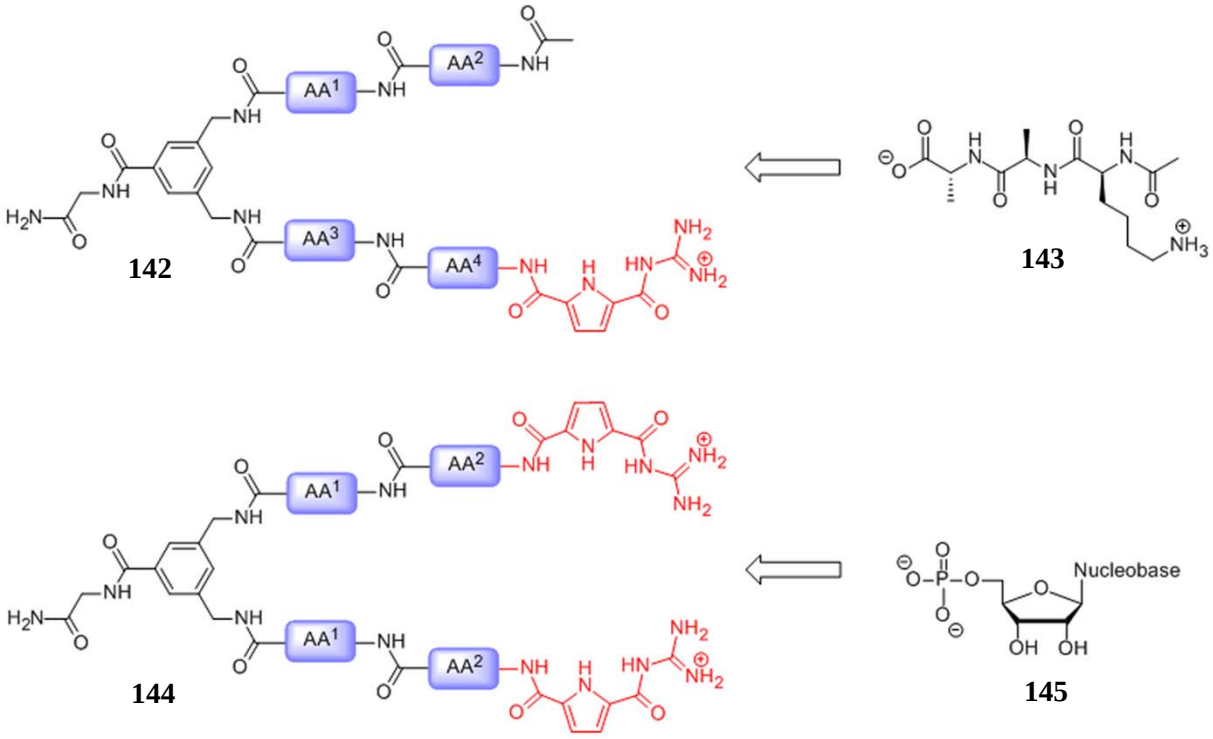


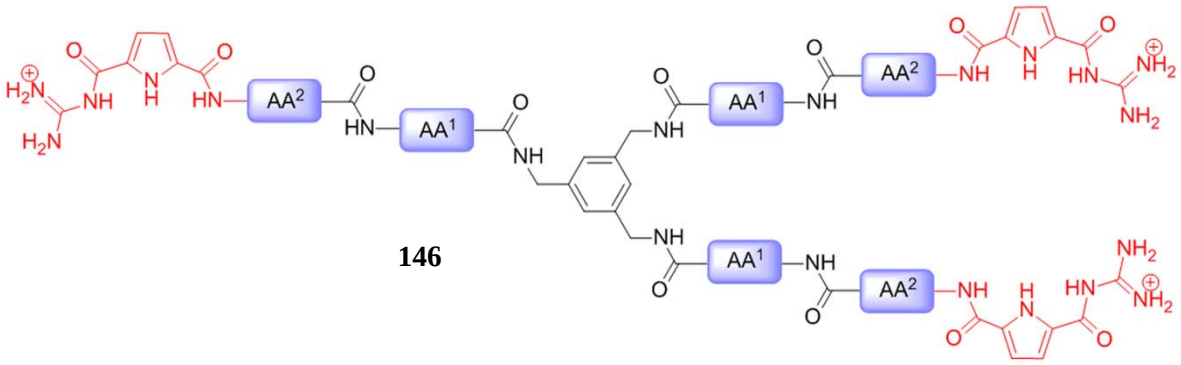
Figure 66. *Symmetric (**143**) and non-symmetric tweezer receptors (**144**) for the molecular recognition of oligopeptides (**143**) and nucleotides (**145**) in aqueous medium at physiological pH.*

The symmetric tweezer receptor **144** will correspondingly carry two identical amino acids AA^1 and AA^2, as well as two oxoanion binding motifs. In comparison to the non-symmetric receptor, the additional GCP unit should allow for the recognition of anions apart from planar ones, like for instance tetrahedral phosphates (**145**). Therefore the binding characteristics towards nucleotides will be studied for this host system.

Nucleic Acid Binding and Cell Transfection

In the next step the molecular recognition of more complex biological systems, namely nucleic acids, will be studied. Therefore the symmetric tweezer system **144** will be tested for its ability to interact with various artificial and generic polynucleotides. Additionally the trivalent ligand **146** depicted in Figure 67 of similar design will be developed and compared

to its two-armed analogue. It will feature three side chains with GCP headgroups of the same amino acid composition as **144**.



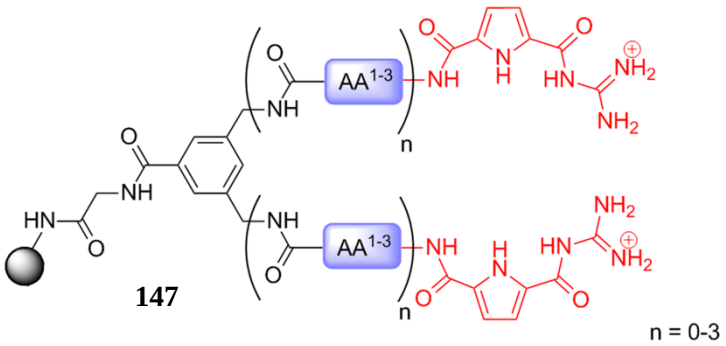
146

Figure 67. *Trivalent polynucleotide-ligand **146** with three identical side-chains and three GCP units.*

The affinity and selectivity of these systems will be studied as well as the structure of the complex. Finally, together with our cooperation partner, the working group of *Prof. S. Knauer* at the Institute for Biology at the Universtiy of Duisburg-Essen we will test the ability of these systems to transfect human cells with genetic information. Such experiments should give new insight into the interactions between small molecular ligands with nucleic acids and hence advance the applicability of non-viral vectors as an essential prerequisite for gene therapy.

Combinatorial DNA-Ligand Library

The combinatorial DNA-ligand library **147** shown in Figure 68 will be prepared next in order to improve the binding affinity of the symmetric two-armed ligands towards polynucleotides. The peptidic side chains will consist of either three, two, one or no amino acids per arm (AA^1-AA^3). This way the influence of the amino acid composition on the nucleic acid recognition event can be studied as well as the influence of the size of the ligands. Finally, the most promising DNA-binders will be tested for their transfection efficiency by our cooperation partner to study whether there is a correlation between strong binding and the ability to shuttle genetic material into human cells.

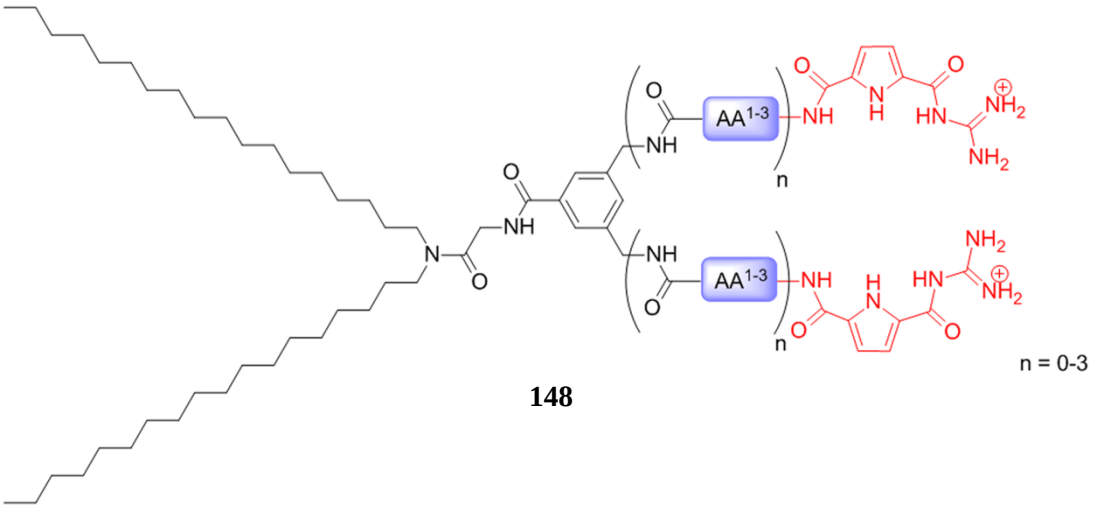


147

Figure 68. *Combinatorial DNA-binder library **147** for studying the influence of the amino acid composition and size of the ligands on the affinity towards nucleic acids.*

Amphiphilic DNA-Ligands for Cell Transfection

In the last part of this thesis the optimized lead-structures for high affinity DNA binding identified by means of the combinatorial library **147** will be derivatized with lipophilic octadecyl chains as depicted in Figure 69. These lipoplex-like structures (**148**) should then ensure very efficient gene transport. This hypothesis will also be tested by our cooperation partner.



148

Figure 69. *Combining high affinity DNA-ligands derived from the combinatorial library **147** with lipophilic hydrocarbon chains generates lipoplexes (**148**) which should be highly efficient gene carriers.*

By synthesizing the artificial host and ligand systems described above novel insights into the complex recognition events in natural systems should be revealed. Furthermore the goal of this work is to directly utilize the so obtained knowledge for the generation of efficient transfecting agents which can be directly applied in cell biology.

4. RESULTS AND DISCUSSION

4.1 RAMAN SPECTROSCOPY AND MOLECULAR RECOGNITION

The most common methods for studying supramolecular interactions between host and guest systems are NMR-,[169] calorimetry- (see chapter 2.3.4), or spectroscopy-based methods like UV/Vis and fluorescence spectroscopy.[170] A less well deployed but forthcoming spectroscopic method for the label-free studying of non-covalent receptor-ligand interactions is Raman spectroscopy which is named after the Indian physicist *C. V. Raman* who was awarded the 1930 Nobel Prize for the experimental discovery of Raman scattering in 1928.[171] This method does not only allow for the determination of the strength of a non-covalent complex, it also enables the elucidation of the dynamics and structure of the complex.[172,173] The Raman signals are characteristic features of a given molecule and are defined by its functional groups. By monitoring changes of these signals upon a binding event one may therefore deduce which functional groups are involved in the complexation process event – and to which extent.

A limiting factor for a more widespread usage of this method however, is the low intensity of the Raman signals. Hence, for the measurement of Raman spectra high concentrations, efficient detectors and a strong energy source (e.g. lasers) are necessary. However, the increase of concentration is not always feasible, and a strong energy source may lead to overheating of the sample. Therefore two methods, which are able to increase the signal strength by several orders of magnitude and thus render Raman spectroscopy an excellent tool for the studying of host-guest interactions will be presented in the following: Resonance Raman (RR) and surface enhanced Raman spectroscopy (SERS).

4.1.1 Resonance Raman Spectroscopy

When irradiating a molecule at a wavelength close to or at an electronic transition, the intensity of the total symmetric vibrations involved in the electronic transition are magnified by a factor of 10^2 to 10^6.[174,175] . An additional advantage of this method is that different parts of the molecule can be selected by tuning the wavelength to a specific absorption which corresponds to a defined moiety of the molecule. By doing so, the resulting Raman spectrum can be simplified tremendously.

With the help of RR *Schlücker* and *Schmuck* could already qualitatively study the complex formation between the receptor GCP-KKF (**13**) and a tetrapeptidic substrate by preparing solutions of different mixing ratios of host and guest and recording the corresponding Raman spectra at a very low concentration for Raman spectroscopy (0.5 mM).[167] In order to test the general feasibility for such a RR approach and to also quantitatively determine the binding constant of the host-guest system, RR titration studies were conducted within this work. Therefore, the receptor and several ligands had to be synthesized, first.

Synthesis of the Guanidinocarbonyl Pyrrole (GCP) Oxoanion Binding Motif

The synthetic work started with the preparation of the guanidiniocarbonyl pyrrole (GCP).[176] The oxoanion binding motif originally prepared by *Schmuck* will be implemented into all receptor molecules presented within this work. Since the synthesis has already been reported in literature and is frequently used in our working group, it will only be explained briefly in a short overview. As depicted in Figure 70 the synthesis started with pyrrole (**149**) which was derivatized in an electrophilic aromatic substitution reaction with trichloroacetyl chloride (**150**) to trichloroacetyl pyrrole (**151**). This compound was then converted to the corresponding benzyl ester **152** via a haloform reaction. By means of a *Vilsmeier-Haak* formylation an aldehyde group was introduced at position 5 (**153**), which was then oxidized to the free carboxylic acid with potassium permanganate. In the next step N-Boc-guanidine (**155**) was attached to the acid **156** with PyBOP as coupling agent to give **157**. Finally, the benzyl ester was selectively hydrogenated palladium-catalyzed with H_2 in the presence of triethylamine to give the Boc-protected guanidinocarbonyl pyrrole **158**, which in this form can be attached to N-termini of peptide sequences via solid phase synthesis.

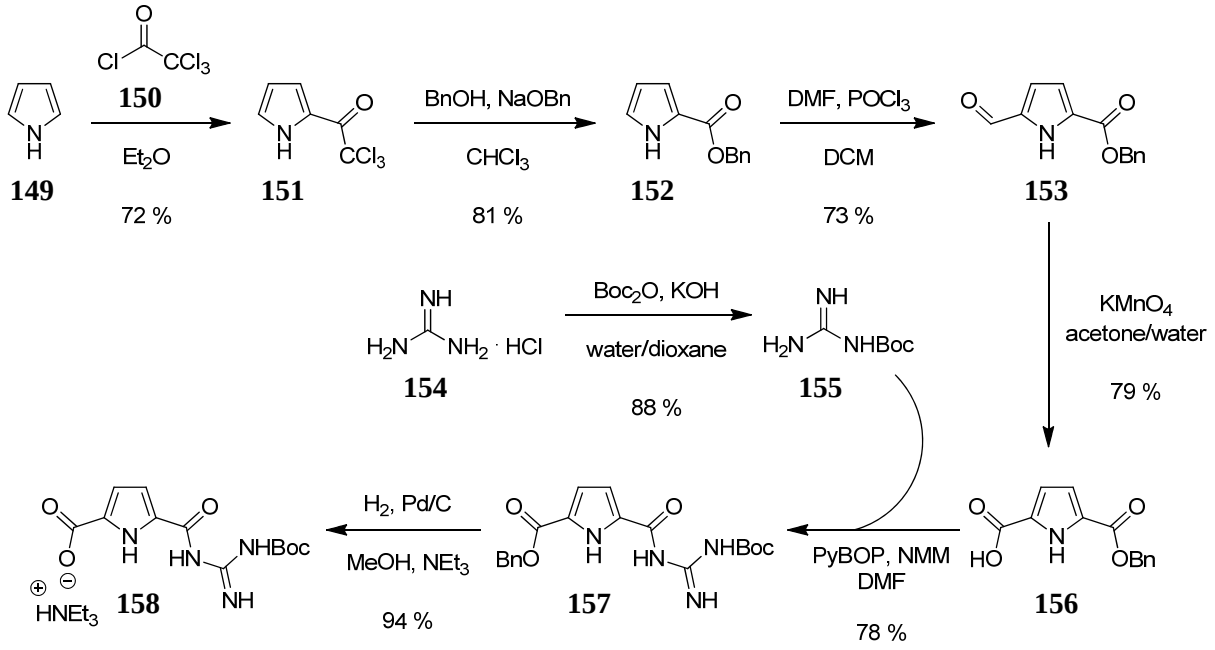


Figure 70. *Synthesis of the guanidinocarbonyl pyrrole oxoanion binding motif* ***158****.*

Solid Phase Peptide Synthesis of Receptor and Substrates

The preparation of receptor **13** has already been described by *Schmuck* and *Heil*.[177] Its synthesis was carried out with the help of standard Fmoc solid phase peptide synthesis (SPPS).[178,179] The advantage of this method is, that peptide sequences can easily be prepared on solid support with high yields and purities. The amino acid building blocks are protected with a base-labile Fmoc protecting group at their N-termini and attached to the free N-terminus of the growing peptide chain via activation of their C-terminus with a coupling agent. To prevent cross-reactions, all amino acid side chains have to be protected as well – most commonly with *tert*-butoxycarbonyl (Boc) and *tert*-butyl (*t*Bu) protecting groups.[180] By carrying out the coupling steps with an excess of reactants with respect to the resin-bound

functional groups, complete conversions can be ensured. No purification is necessary in between synthetic steps, because the reactants can easily be removed from the mixture by filtration and washing steps. Only after cleaving the resin-bound compound under acidic conditions and thereby simultaneously removing all protecting groups, the product has to be purified. This is why SPPS provides a convenient method for the preparation of oligopeptides with high yields and purities.

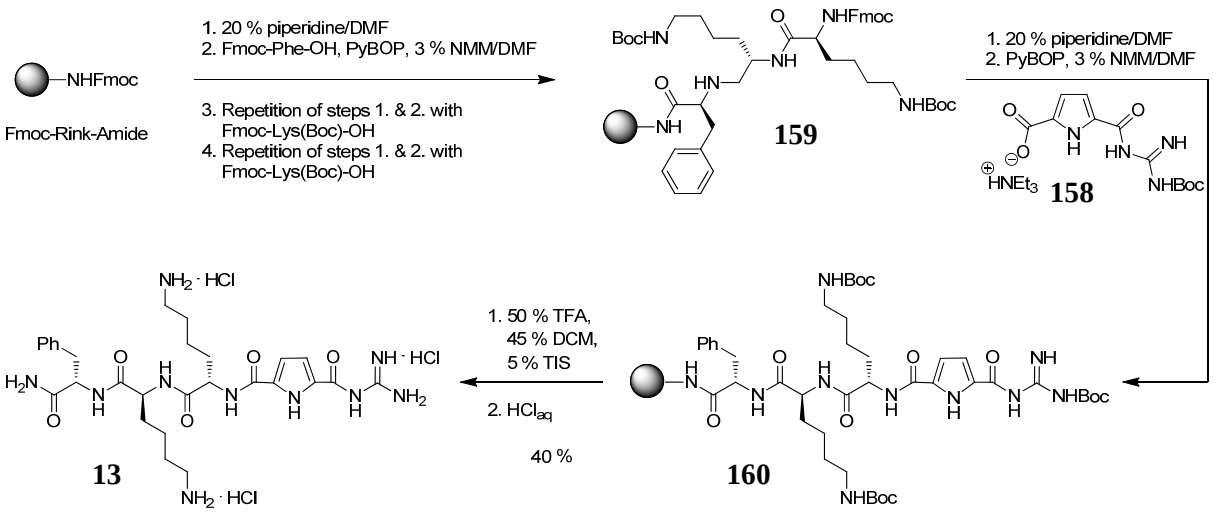


Figure 71. *Solid phase peptide synthesis of the receptor* ***13*** *on Fmoc-Rink-Amide resin.*

The synthesis of **13** was carried out as depicted in Figure 71 on Fmoc-Rink-Amide resin (loading: 0.94 mmol/g). After swelling the solid support in dimethylformamide (DMF) to increase the accessibility of the functional Fmoc-amine groups inside the resin, the Fmoc protecting group was removed by treating the polymer with 20 % piperidine/DMF (2 × 20 min). Successful liberation of the amine functionalities on the resin was confirmed with the help of the Kaiser test.[181] This color test gives a positive answer, when free amines are present on the solid-support and is negative when all amines have been converted to amides. Between all synthetic steps the resin was thoroughly washed with DMF. The first amino acid Fmoc-Phe-OH (3 eq) was activated with the coupling agent benzotriazol-1-yloxytri-pyrrolidinophoshonium hexafluorophosphate (PyBOP, 3 eq) in 3 % N-methylmorpholine (NMM)/DMF under argon atmosphere and attached to the resin by shaking it for 8-12 h. Since the first step of a solid phase synthesis is regularly rather difficult, the coupling step was repeated to ensure complete conversion of the resin´s amine groups. In order to cap eventually still remaining free amines, the resin was then treated with acetic anhydride in 3 % NMM/DMF. A negative Kaiser test indicated complete conversion. After another Fmoc deprotection the next two amino acids, two times Fmoc-Lys(Boc)-OH and finally the Boc protected GCP group were attached accordingly. If necessary, coupling steps were repeated as indicated by the Kaiser tests. After completion of the SPPS, the resin was extensively washed with DMF, dichloromethane (DCM) and methanol (MeOH) and dried in vacuum. To cleave the resin-bound product **160** from the solid support and to simultaneously remove all protecting groups, the polymeric support was treated with a mixture of 50 % trifluoroacetic acid (TFA) in DCM containing 5 % triisopropylsilane (TIS) for 4 h.[182] The latter was necessary as nucleophilic scavenger to intercept reactive cation species which are formed

under these harsh acidic conditions, e.g. *t*-butyl cations from the Boc deprotection.[183] These cationic species could else irreversibly add to the liberated product thus reducing the yield and complicating the purification. After precipitation from diethyl ether the crude product was purified with the help of reversed phase (RP) MPLC (RP18, H_2O → MeOH, 0.1 % TFA) to obtain the product with 40 % yield. Finally, the trifluoroacetate counterions were exchanged for chlorides by dissolving the product several times in aqueous HCl (0.1 N) and consequent lyophilization to obtain **13** as the hydrochloride salt.

The synthesis of the tetrapeptidic substrates N-Ac-Ala-(D-Glu)$_3$-OH (**161**), N-Ac-Ala-D-Ala-(D-Glu)$_2$-OH (**162**), N-Ac-(D-Glu)$_3$-D-Ala-OH (**163**), and N-Ac-(D-Glu)$_3$-Ala-OH (**164**) (see Figure 72) was carried out via SPPS according to a procedure reported for **161** by *Schmuck* and *Wich*.[29] Wang resin (1.63 mmol/g) was used as solid support – a polymer which features hydroxy instead of amine groups for the attachment of the first amino acid via its C-terminus. After cleavage from the resin, the products are consequently liberated as carboxylic acids instead of amides. To account for the reduced reactivity of the hydroxy group, the first coupling step was conducted with N,N´-diisopropylcarbodiimide (DIC, 2 eq) and 1-hydroxy-benzotriazole (HOBt, 2 eq) plus catalytic 4-dimethylaminopyridine (DMAP) in a mixture of DCM/DMF (7/3) under argon atmosphere (8-12 h) and repeated four times. The following coupling and deprotection steps were carried out as described for **13** (vide supra) with PyBOP as coupling agent and the building blocks Fmoc-D/L-Ala-OH and Fmoc-D-Glu(O*t*Bu)-OH. After removal of the last N-terminal Fmoc group, the resin was treated with acetic anhydride as persistent N-cap. After MPLC purification the substrates **161**-**164** were obtained with good to excellent yield (**161**: 90 %; **162**: 83 %; **163**: 53 %; and **164**: 76 %).

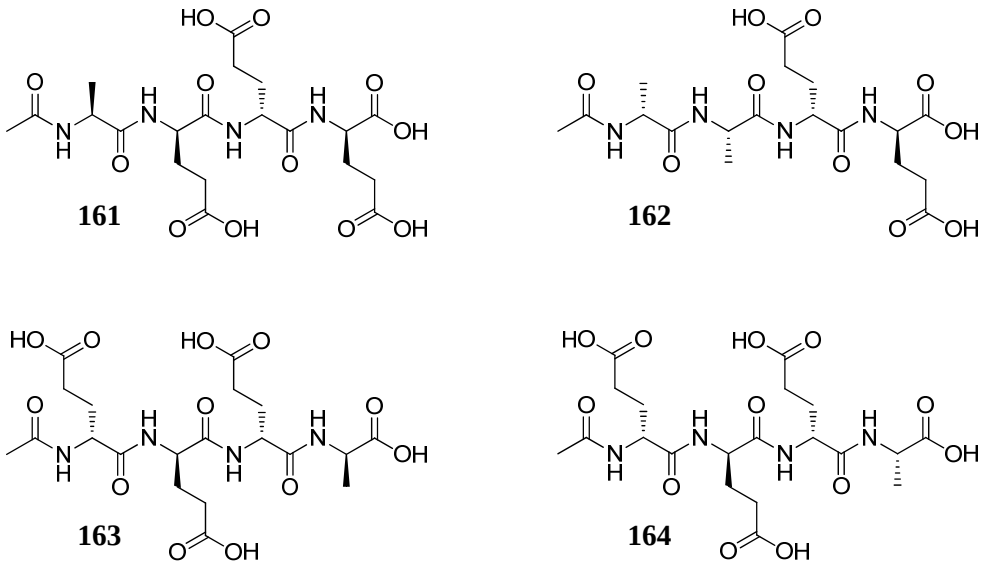


Figure 72. *The four tetrapeptides* ***161-164*** *were synthesized by SPPS on Wang resin.*

In order to ensure the high purity needed for RR studies, the substrates were further subjected to purification via preparative HPLC. After identifying the ideal isocratic conditions on analytical HPLC (acetonitrile/water + 0.5 % TFA) the pure products were isolated via preparative HPLC on multimilligram scale. An exemplary purification of the substrate **161** is depicted in Figure 73.

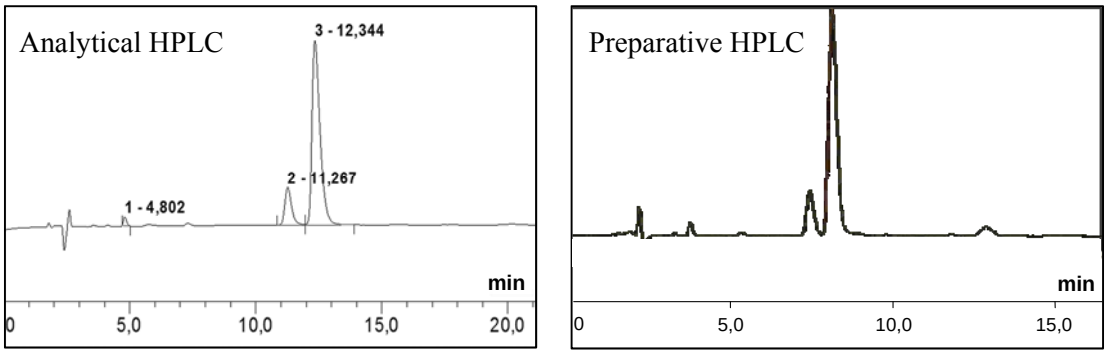


Figure 73. *Purification conditions for* ***161*** *were optimized via analytical HPLC to isocratic elution at 3 % acetonitrile/water (0.05 % TFA) as monitored by the UV/Vis absorption at 220 nm (left) and successfully transferred to purify the compound via preparative HPLC (right).*

Determination of the Binding Constants

The so obtained substances were now ready for Raman measurements, which were conducted by *Stephan Niebling* in the framework of his Ph.D. thesis in the working group of *Professor Sebastian Schlücker* at the Institute of Physics at the University of Osnabrück.[184] With the experience of the qualitative binding studies at hand, the pyrrole moiety of the receptor was excited at 275 nm, which is 25 nm away from its absorption maximum at 300 nm, to prevent fluorescence which would otherwise overlay the Raman signals.[167] The observable signals are then exclusively due to the GCP unit. In order to acquire quantitative binding information, eleven solutions were prepared at a receptor concentration of 1 mM containing increasing amounts of substrate (0 to 10 eq). The resulting RR spectra were recorded. A representative example of three mixtures of **13** with **161** is illustrated in Figure 74 (top, left). The measurements were conducted at pH 6 without buffer. Therefore the pH was checked before and after the measurement to guarantee a constant value ($\Delta pH \leq \pm 0.05$). In general, the observed spectral changes upon addition of the substrates are only weak: There are minor intensity changes at 940, 955, and 1630 cm^{-1}. The most pronounced change could be witnessed for the two bands at 1440 and 1470 cm^{-1} (red). For the latter one also a shift to higher wavenumber could be observed upon addition of the substrate. Both bands can be ascribed to bending modes of the guanidinio part of the GCP group.

The next step was to analyze the data in more detail for the calculation of the binding constant. Therefore, the experimental spectra were subjected to a non-negative matrix factorization.[185] A matrix containing the experimental spectrum was factorized into matrices containing the spectra of the individual components and into matrices containing the contributions of the components to each experimental spectrum. In this case, the experimental spectrum only contains two components due to the lack of substrate RR signals: the spectrum of the free and of the substrate-bound receptor. While the first one corresponds to the known experimental spectrum for the receptor, the latter one can be approximated by the spectrum obtained with 10 eq of substrate. Thus, the degree of complexation can be extracted from each spectrum and plotted against the amount of ligand (Figure 74, right). With the already known 1:1 host-guest stoichiometry,[29] a multivariate non-linear least square fitting procedure was utilized to determine the binding constant.

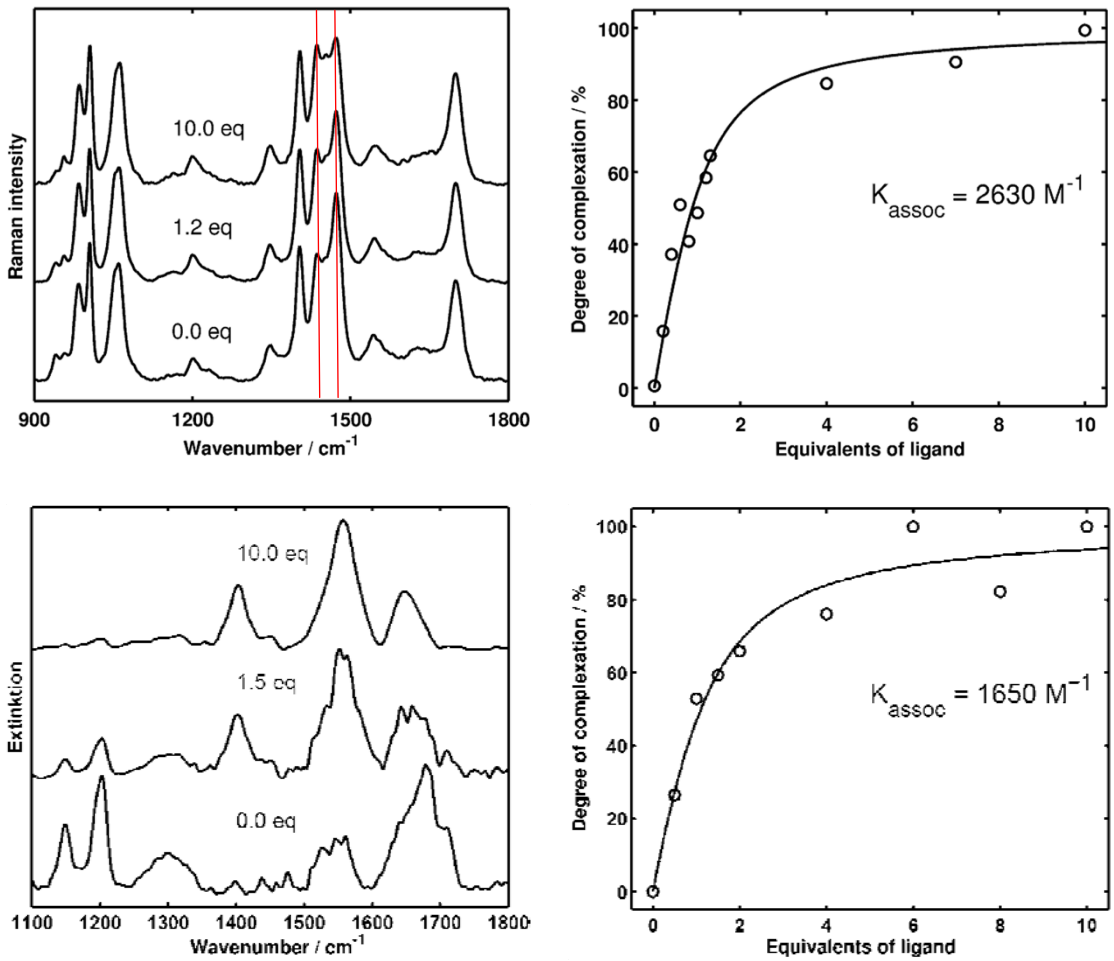


Figure 74. *Top: Resonance Raman spectra upon irradiation at 275 nm of GCP-KKF* ***13*** *(1 mM) and mixtures thereof with* ***161****. The degree of complexation was determined by non-negative matrix factorization and fitted with a non-linear regression for a 1:1 complex stoichiometry. Bottom: FT-IR spectra of* ***13*** *and mixtures with* ***161*** *and data processing thereof similar to the RR study (from: Niebling Ph.D.).*[184]

Table 1. *Comparison of binding constants K [M*$^{-1}$*] obtained by various methods.*

Substrate	On-bead fluorescence[29]	UV[29]	RR	IR
161	8500	n.d.	1910	2210
162	2600	n.d.	1260	420
163	2100	4500	n.d.	n.d.
164	200	1400	n.d.	n.d.

Table 1 lists the so obtained values and compares them to the data from the on-bead fluorescence screening and solution phase UV/Vis titrations. Additionally, FT-IR binding studies were conducted by *Niebling* and analyzed by the same procedure (Figure 74, bottom). In contrast to RR, the IR spectra contain signals from all functional groups of the receptor and not just from the GCP moiety, and additionally substrate-derived bands. The non-negative matrix factorization was thus conducted with three components instead of two. The resulting degrees of complexation were again plotted against the amount of substrate and a non-linear least square fit gave the binding constant. The data listed in Table 1 was still work in progress at the time this dissertation was written. However, it was possible to reproduce the order of

affinity **161** > **162** with this RR approach and hence this work can be regarded as proof-of-principle that quantitative binding affinity studies are accessible with the help of Resonance Raman spectroscopy. Especially when taking into account that the original data stems from on-bead assays and was carried out with completely different conditions (concentration, buffer vs. no buffer) the results are in very good agreement. In future experiments also the binding affinities of **163** and **164** will have to be analyzed.

4.1.2 Surface Enhanced Raman Spectroscopy

Another excellent method for the amplification of Raman signals is the surface enhanced Raman spectroscopy (SERS).[186] In close proximity to metal surfaces Raman signals are amplified even stronger than it is the case for Resonance Raman by a factor of 10^6 to 10^8. The effect was first observed by *Fleischmann* in 1974 for pyridine molecules which were adsorbed at the surface of a silver electrode.[187] In the best case SERS amplification by 14 orders of magnitude is possible. Most commonly noble metal nano-structures derived from gold or silver are utilized.[188] The reasons for the massive increase of signal strength are high local electromagnetic fields (due to surface plasmon Resonances at the metal surface) and a dynamic charge-transfer electronic transition between metal and molecule. The SERS effect is extremely dependent on the distance between analyte and surface: It decreases with r^{-10}.

Schmuck and *Schlücker* utilized this SERS effect for the direct and label-free characterization of solid phase bound GCP-KKF (**13**).[168] Due to the low loading of approximately 100 pmol per bead it is normally very difficult to analyze solid phase bound substances. The complexation properties of bound compounds are regularly screened with the help of UV/Vis or fluorescence tagged substrates. By doing so it is however not possible to gather structural information. It is not even possible to determine whether the expected compound is indeed present on the beads under inspection. This last point is especially troubling when working with huge, random libraries. To an extent mass spectrometry is able to characterize potential solid-bound hit structures in hindsight, but only if the compound is cleavable from the solid-support.[189] Therefore chemical tags have been developed which allow for identification of the compound after screening.[190] This approach however, increases the complexity of the synthesis and may potentially lead to unwanted interactions with the receptor or the substrate during the screening process and thus limit its application.[40] Solid-state NMR and even UV/Vis and fluorescence spectroscopy are not sensible enough to detect molecules that dilute in the presence of a high excess of solid support. Vibration-based spectroscopic techniques like IR and Raman neither allow for differentiation between actual compound and the matrix background of the bead. These inconveniences were circumvented by *Schmuck* and *Schlücker* by first preparing the receptor on TentaGel resin and subsequently adsorbing silver nanoparticles (NPs) onto the surface of the solid support. Consequently, the Raman signal could be increased by a factor of 10^6 to 10^7. Due to the distance-dependency of the SERS amplification (r^{-10}), the receptor could unambiguously be distinguished from the resin despite the low resin loading and the huge excess of polystyrene matrix. However, for general application of this method two inconveniences remain: The first is potential interactions between the silver colloids and the resin-bound substances which may alter its original electronic, vibrational and conformational properties. Second, the strong interaction between metal and solid phase bound compound is often not reversible. Therefore, this approach restricts the further usage of the resin-bound substances. For instance, the control of a synthetic procedure at the intermediate state by Raman spectroscopy followed by continuation of the synthesis will be rather difficult to achieve, because the NPs are very difficult to remove from the surface of the resin. This is why in the following a new approach will be presented. As shown in Figure 75 beads for solid phase synthesis and silica-protected plasmonic nanoparticles for SERS analysis will be integrated into a bifunctional unit. Gold nanoparticles will be adsorbed onto the resin in a way that they almost completely cover its

surface (a, b) and will then be further derivatized with e.g. amine groups to allow for SPPS (c). The following work was carried out together with *Magdalena Gellner* from the working group of *Professor Sebastian Schlücker* at the Institute of Physics at the University of Osnabrück.[191]

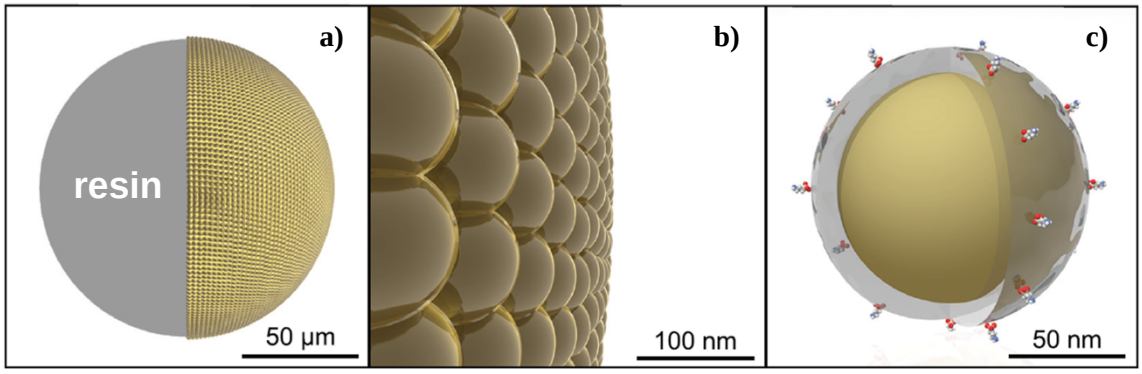


Figure 75. *Design of the plasmonic microspheres. a) Self-organized, silica-encapsulated, 80 nm sized Au nanoparticles on a 130 µm TentaGel bead; b) Zoom onto the NP-loaded surface; c) Derivatization of the silica-encapsulation with amino silane results in terminal amine groups ready for SPPS (from: Gellner Ph.D.).*[191]

Preparation of SPPS-Beads Modified with Silica-Encapsulated Au Nanoparticles

Commercially available monodisperse gold nanoparticles of 80 nm size were encapsulated by *Gellner* with an ultrathin layer (2-3 nm) of silica by treating the nanoparticles with a dilute $Na_2Si_3O_7$ solution according to a procedure described in literature.[192] The glass-encapsulation prevents adsorption of unwanted molecules onto the gold surface and improves the mechanical stability of the Au NPs.[193] Due to its thinness it still allows for plasmonic coupling and efficient SERS amplification.

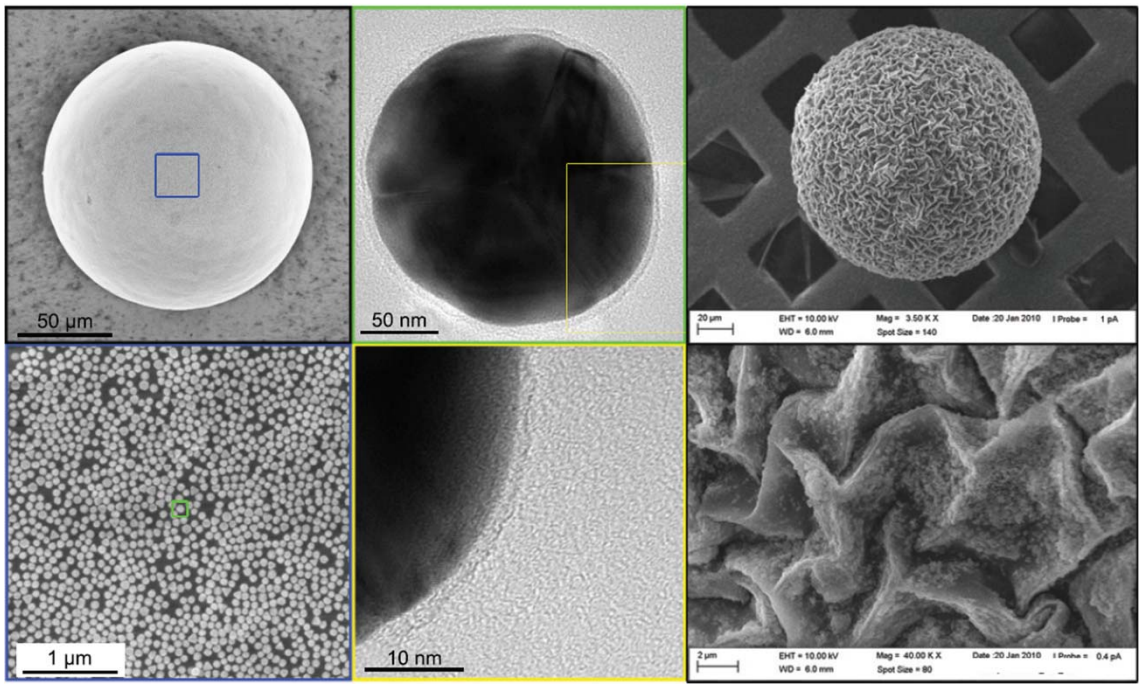


Figure 76. *Left: SEM image of a TentaGel bead, which is completely covered with silica-encapsulated Au NPs (top) and zoom onto the surface (bottom). Middle: HR-TEM image of a gold NP with 2-3 nm glass-shell (top) and higher resolution image thereof (bottom). Right: Earlier try to silica-encapsulate the bead after adsorption of the Au NPs onto the resin surface lead to the formation of a wool-like structuring (top). At high resolution the single Au NPs are not visible (bottom). All pictures from Gellner Ph.D.* [191]

In the next step the solid support TentaGel-NH_2 (130 µm) was incubated with an excess of Au colloid, which was adsorbed onto its surface via electrostatics. The large excess of NPs ensured a maximal coverage of the surface with nanoparticles almost side by side, which is a prerequisite for efficient SERS amplification.[194] This could be proved with the help of scanning electron microscopy (SEM) as shown in Figure 76 (left, bottom). The successful ultrathin silica-encapsulation of the entire surface of the gold NPs could also be demonstrated via high resolution transmission electron microscopy (HR-TEM, Figure 76 middle). It is crucial to encapsulate the Au NPs prior to adsorption onto the TentaGel surface. Earlier tries to create the ultrathin glass layer after adsorption of the NPs led to complete capsulation of the bead with a thick silica-layer (see Figure 76, right). As a consequence no Raman signals could be observed after attaching substrates to these beads caused by the increased distance between Au surface and the compounds.

SPPS of Amino Acids and a Dipeptide on the Modified Beads

The further derivatization of the so obtained Au NP modified beads was carried out in our laboratory. Due to the complete coverage of the TentaGel surface with gold nanoparticles, it was not accessible anymore for SPPS. Therefore, an amine functionality was introduced onto the silica-surface by treating the beads with a solution of 0.5 % 3-amino-N-propyltrimethoxy-silane (APTMS) in ethanol (EtOH) (50 µL) and 25 % aqueous ammonia (50 µL) in ethanol (1 mL) for 30 min. The trimethoxysilane part of this reactant adsorbs onto the glass surface and thus a free amine function is generated, which could be used for the attachment of various amino acids or oligopeptide substrates. The so obtained beads were washed with DMF and then subjected to SPPS. Approximately 20-30 beads were treated with a 30 mM solution of PyBOP and either Fmoc-Gly-OH, Fmoc-Phe-OH or Fmoc-Ser(O*t*Bu)-OH in 3 % NMM/DMF (1 ml) for 8-12 h under argon atmosphere to attach the three different compounds to the surface-bound amine. The coupling step was repeated. Then the Fmoc groups were removed with 20 % piperidine/DMF (2 × 20 min). Fmoc-Phe-OH was then attached to a fraction (10-15 beads) of the glycine derivatized beads and its Fmoc group removed according to the same procedure. Finally, all beads were treated with 33 % TFA/DCM for 4 h to remove the serine *t*Bu protecting group and to generate the TFA salts from the free amine N-termini of all four substrates for means of better comparability of the resulting spectra. Finally, the beads were washed with MeOH and water.

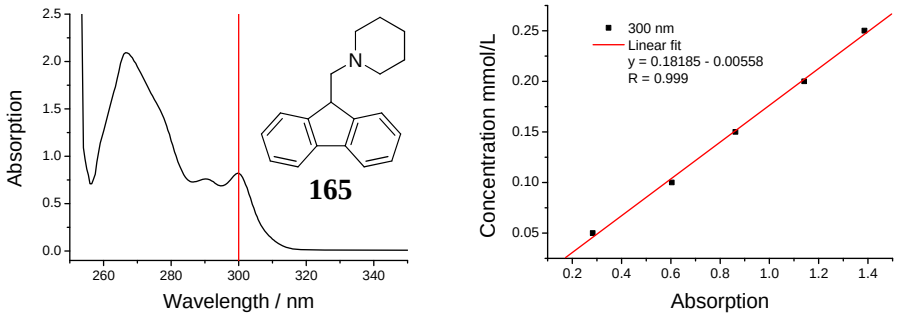


Figure 77. *The UV/Vis absorption spectrum of the Fmoc-piperidine adduct* ***165*** *(left) and the corresponding linear fit of the concentration dependent absorption at 300 nm (right).*

In order to determine the loading of the modified beads, the piperidine/DMF solutions and the following washing solutions were collected, combined and their UV/Vis spectrum was measured between 250 and 350 nm. As shown in Figure 77 the concentration of the Fmoc-piperidine adduct (**165**) can be determined with the help of its UV absorption at 300 nm and a concentration dependent calibration curve as reported in my diploma thesis.[195] By counting the number of beads with the help of a microscope, the loading per bead was then calculated as 17.14 (Gly), 22.55 (Phe), 31.69 (Ser) and 12.53 nm/bead (Gly-Phe). This loading is in the range of commercially available resin loadings (0.25-3 mmol/g).

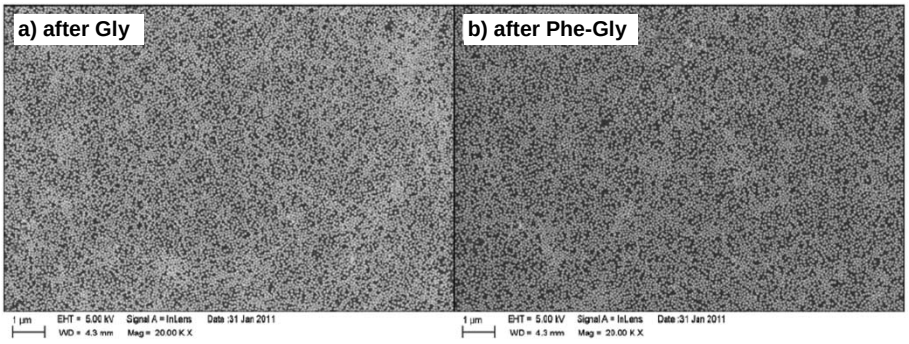


Figure 78. *SEM images of the surface of the microspheres after one (a) and respectively two SPPS steps (b) and treatment with 33 % TFA/DCM for 4 h demonstrating that the surface is still completely covered with silica-encapsulated gold nanoparticles (from Gellner Ph.D.).*[191]

Finally, for the use in surface enhanced Raman spectroscopy, it had to be ensured, that the packing density of the Au NPs on the TentaGel surface is still intact after the synthesis. As shown in Figure 78 SEM images were recorded by *Gellner* after glycine (a) or the dipeptide Gly-Phe (b) were attached to the beads via SPPS. The images clearly demonstrate that the packing of the Au NPs was unharmed even after two couplings, two basic deprotection steps, treatment with organic solvents like DMF or MeOH, and 4 h under very harsh acidic conditions. Thus the TentaGel-Au NP assembly has proven to be extremely stable under the conditions which are necessary for standard Fmoc SPPS.

SERS of the Derivatized Beads

In the next step, the so derivatized beads were analyzed via SERS measurements by *Gellner*. They were irradiated at an excitation wavelength of 785 nm emitted by a diode laser. Figure 79 shows the spectra obtained from differently derivatized beads: Serine (magenta, **166**), glycine (red, **167**), phenylalanine (green, **168**) and from beads after (cyan, **169**) and prior to derivatization with APTMS (black, **170**). Every set of spectra contains information about the mean value (colored line) and the standard deviation (upper and lower limit).

It is clearly visible that the derivatization leads to the formation of additional bands and to changes of band intensities which are different from the spectra of the untreated beads (**169**, **170**) and therefore characteristic for the amino acids. Furthermore, serine and glycine could be differentiated from each other by the ratio of the intensity of the Raman signals at 1570 and 1602 cm^{-1}. The quotient is 0.63 ± 0.35 for Gly and 1.37 ± 0.32 for Ser. Phenylalanine could be discriminated from the other two amino acids by its overall lower Raman intensity in this region of the spectrum. Comparison of these spectra with those obtained from beads prior to

and after APTMS functionalization revealed that all spectra are dominated by vibrational contributions of the polystyrene matrix and the silica network. The phenylalanine signals of the aromatic ring and the CH_2 group are overlaid by PS phenyl ring modes and APTMS CH_2 deformation modes. Therefore, it was rather difficult to detect Phe.

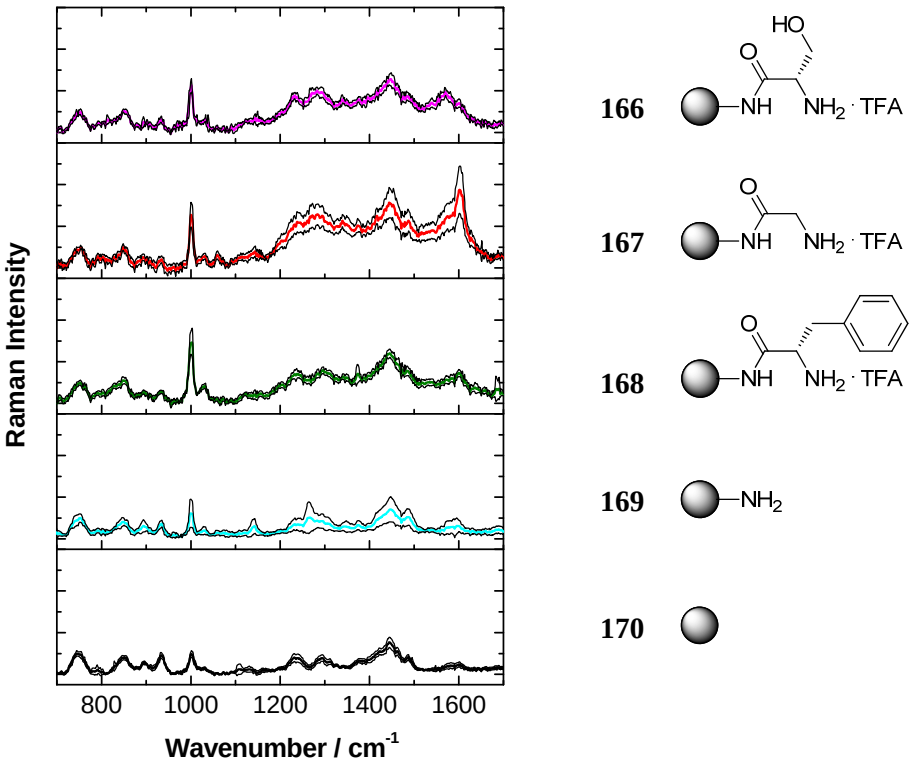


Figure 79. *SERS spectra of serine (magenta, **166**), glycine (red, **167**) and phenylalanine (green, **168**) as obtained from the Au-NP modified TentaGel as well as SERS spectra of beads after (cyan, **169**) and prior to APTMS treatment (black, **170**). Spectra from Gellner Ph.D.*[191]

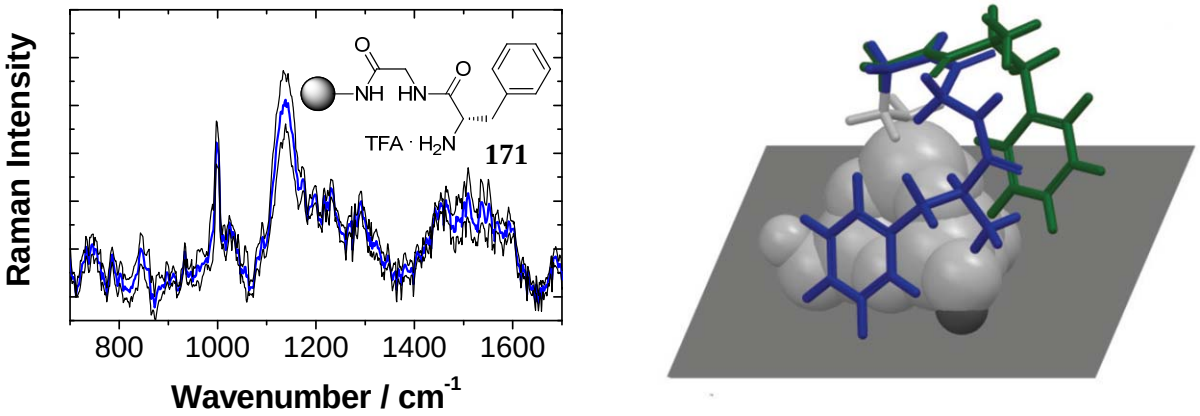


Figure 80. *Left: SERS spectrum of the dipeptide Gly-Phe (**171**) as obtained from the Au-NP modified TentaGel. Right: The calculated molecular model consisting of a small silica network, the C_3 alkyl chain of the amino silane and Phe (green) and respectively Gly-Phe (blue). Spectrum and image from Gellner Ph.D.*[191]

To test whether the dipeptide Gly-Phe (**171**) can be discriminated from its parent amino acids Gly and Phe, the so derivatized beads were analyzed next. As illustrated in Figure 80

(left) an additional Raman band at 1144 cm^{-1} could be observed which was neither present in the glycine nor in the phenylalanine spectrum. To identify the origin of this new band, calculations of model systems comprised of a small silica network, the APTMS propyl chain and either the single amino acids Gly, Phe or the dipeptide Gly-Phe were carried out by *Niebling*. First a conformational analysis was carried out with the help of classical molecular mechanics. Then a DFT calculation for the conformer of lowest energy was performed. The results of these calculations indicated that Phe features several ring mode contributions with high Raman intensity at 1140 cm^{-1}. Three normal modes which contribute to the Raman signal at this region are depicted in Figure 81.

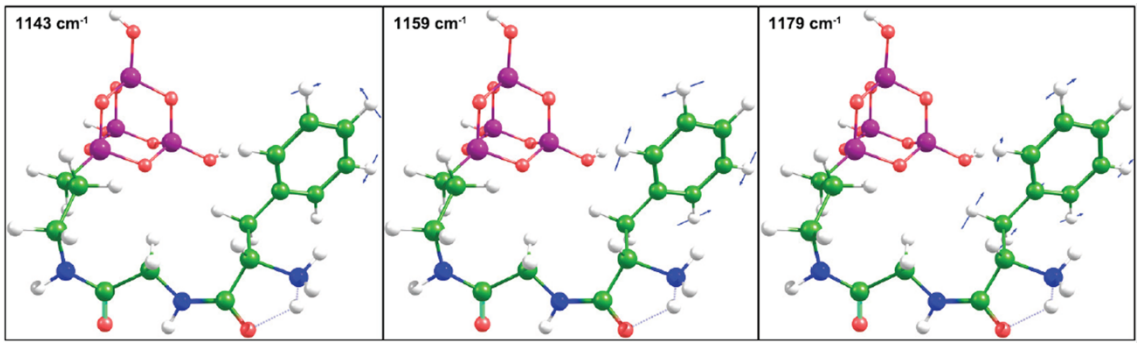


Figure 81. *Eigenvectors of selected phenyl ring modes. A scale factor of 0.96 was utilized for the DFT-derived harmonic wavenumbers.*[196] *The three normal modes stem from vibrational contributions in the region around 1144 cm^{-1} (from Gellner Ph.D.).*[191]

As indicated in Figure 80 (right) a possible explanation of the lack of this band in the Phe spectrum is that in the dipeptide (blue) the phenyl ring is able to come within close proximity of the silica network due to its increased flexibility in comparison with the simple amino acid. The proximity to the silica surface also brings the phenyl ring within closer contact to the gold surface of the nanoparticles and thus leads to an increased amplification of the ring mode-derived Raman signal. For the single amino acid Phe (green) this is not possible.

In conclusion, in cooperation with the working group of *Schlücker* it was possible to develop a bifunctional system for the preparation of short peptides via SPPS and their direct observation via SERS. Therefore, commercial gold nanoparticles were encapsulated by an ultrathin silica layer, aggregated with commercial TentaGel resin and derivatized with an amine group. In the framework of these first proof-of-principle studies three amino acids (Gly, Phe, Ser) could be attached to the modified resin and discriminated from each other and also from an analogue dipeptide (Gly-Phe). The modified solid support is stable under harsh acidic and basic conditions as well as in various organic solvents and can be utilized with great ease for the preparation of short oligopeptide sequences via SPPS. In future studies, such integrated SERS/SPPS systems might find potential application for the studying of non-covalent interactions, which can be directly observed via changes in their Raman signature. Hence, this bisignate system offers great potential in supramolecular chemistry, for instance as sensor for the rapid detection of guest molecules.

4.2 NON-SYMMETRIC TWEEZER RECEPTORS FOR PEPTIDE RECOGNITION

Although many host systems have been proposed for the molecular recognition of short peptides sequences, the development of high affinity receptors for polar oligopeptides in water remains a demanding task due to the competitive influence of the medium. Within the framework of the continuation of the work I have conducted during my diploma thesis, non-symmetric tweezer receptors, i.e. host systems with two different arms, which are linked via an aromatic template, will be prepared for the molecular recognition of the biologically relevant, polar tripeptide N-Ac-Lys-D-Ala-D-Ala-OH in water.[195]

4.2.1 Synthesis of the Orthogonally Protected Template

For the preparation of tweezer receptors with two different side chains on solid-support an orthogonally protected template is necessary. The synthesis of such a scaffold was already described in my diploma thesis and will therefore only be described in short, focusing on improvements of the original synthetic protocoll.[195] As shown in Figure 82 commercial 3,5-dimethylbenzoic acid (**172**) was converted with oxalyl chloride and catalytic amounts of DMF to the acyl chloride. In situ reaction in methanol gave the corresponding methyl ester **173**.

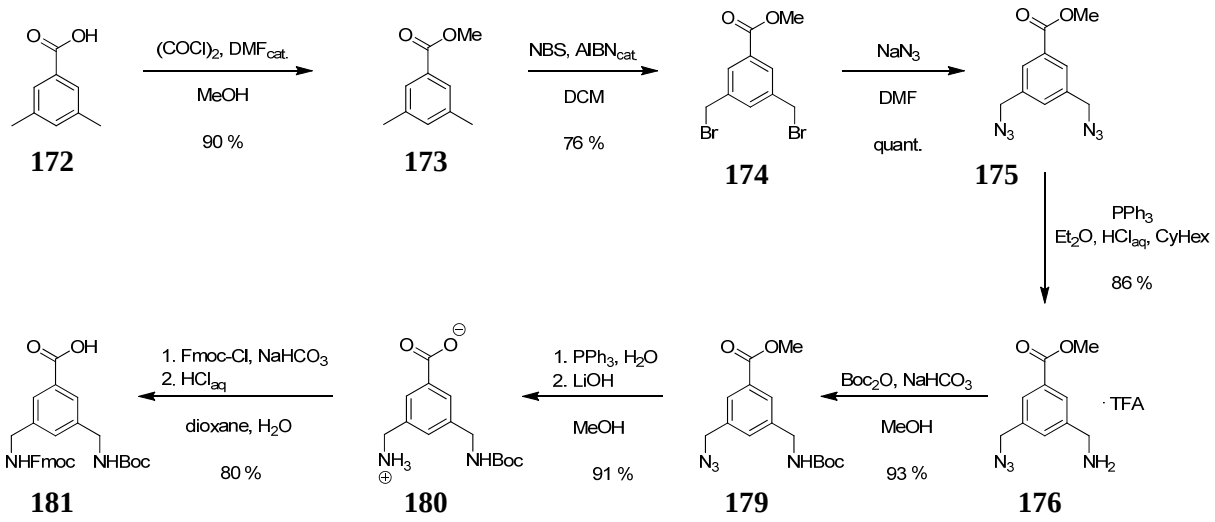


Figure 82. *Synthesis of the orthogonally protected aromatic scaffold **181**.*

A *Wohl-Ziegler* bromination with N-bromosuccinimide (NBS) and azobis(isobutyronitrile) (AIBN) afforded the dibromide **174**. Instead of carbon tetrachloride as reported in the diploma thesis dichloromethane was utilized as solvent. By doing so, the yield could be increased from 67 to 76 % thus facilitating the tedious isolation of the dibromide from various other brominated byproducts by column chromatography. Furthermore, this way the highly toxic and carcinogenic solvent CCl_4 could be replaced by a less hazardous one. A classical substitution reaction generates the diazide **175** in quantitative yield. The following key-step of the reaction procedure, the mono *Staudinger* reduction to the azido-amido derivative **176**, was carried out with triphenylphosphane in a two-phase system composed of an acidic aqueous phase and an organic phase as depicted in Figure 83. The iminophosphorane intermediate **177** is formed in the organic phase. As soon as it is reduced to the corresponding amine **178** at the interphase it gets protonated by HCl and is transferred into the aqueous phase as ammonium

salt where it cannot be further reduced. By adding 25 % cyclohexane (CyHex) to the diethyl ether phase the yield could be improved from 61 %, as reported in the diploma thesis, to 86 %. When adding 25 % ethyl acetate (EA) instead, the yield was reduced to 33 %. Thus it can be concluded that a decreased polarity of the organic phase favors the transfer of the protonated amine to the aqueous phase. The only side products were the diamine and hydrolyzed methyl ester derivatives with a free carboxylic acid. Hence, this kind of mono reduction in a biphasic system provides a facile and efficient approach for the generation of two differently derivatized amine functionalities of identical reactivity, which is otherwise often accompanied by low yields and tedious work-up procedures. In the next step the amine group was protected with a Boc group (**179**). Reduction of the second azide and saponification of the methyl ester afforded **180**, which was Fmoc-protected in the last step to give the desired template **181**. The yield of the last step could be improved from 71 to 80 % by recrystallizing the Fmoc-chloride from toluene/CyHex prior to use and by purifying the product via reversed phase (RP) MPLC with dioxane/water instead of normal phase (NP) with EA/MeOH.

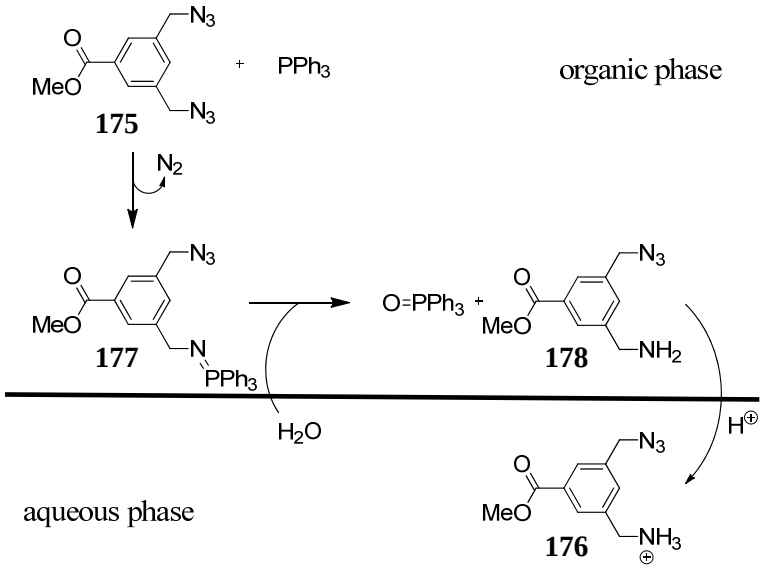


Figure 83. *Selective mono reduction of the diazide **175** to the mono amine **176** in a two-phase system.*

The overall yield of the seven step synthetic procedure could be improved from 24 to 39 %. The reaction can be carried out in multigram quantities starting from commercially available compounds. The aromatic scaffold is now ready for the solid phase synthesis of tweezer receptors with two different side chains.

4.2.2 Microwave-Assisted Solid Phase Synthesis of Two Non-Symmetric Tweezer Receptors

The so obtained orthogonally protected template has already proven to allow for the preparation of tweezer receptors with non-symmetric arms during the work which was conducted for my diploma thesis.[195] For example, such a system with the amino acid sequence Ser-Tyr-Gly in the first and Asp-Phe-GCP in the second arm could be synthesized on MBHA resin with 45 % raw product yield. However, the purity of the raw product was rather poor (15 %) and tedious purification steps via RP MPLC were the consequence. In the following an improved receptor design will be presented which should allow for facilitated synthesis and improved binding properties to peptidic substrates (**182** versus **183** in Figure 84).

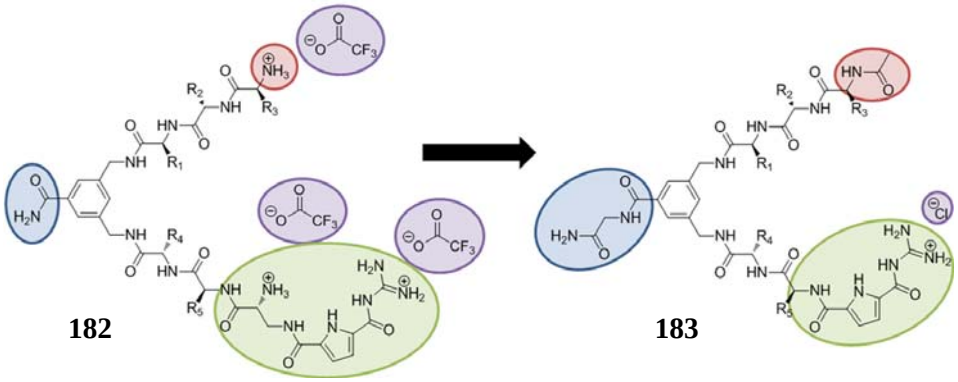


Figure 84. *Former (**182**) and improved design (**183**) for the non-symmetric tweezer receptors.*

The general design of the receptors however will stay the same. Two arms with different amino acid sequence will be linked via the aromatic template, which has already proven to feature the correct spacing and the ability to preorganize the two arms in a fashion which is favorable for the binding of peptidic substrates. One arm will feature the GCP headgroup as carboxylate binding site for the substrate´s C-terminus. The receptor is supposed to be soluble in water at pH 7 and bind to a tripeptide substrate with good affinity. Several improvements will be introduced: First of all, the resin will be preloaded with glycine instead of directly attaching the aromatic template to the resin (Figure 84, blue). This spacer will be introduced because the first step of a solid phase synthesis is regularly the critical one. Very often the first attachment is difficult, resulting in incomplete conversions and thus in low yields and purities. The coupling of the aromatic scaffold **181** is not trivial due to steric constraints and a rather low reactivity of the aromatic carboxyl group. It is therefore not very practical to use it as first residue. Glycine on the other hand is a sterically undemanding amino acid with good reactivity and therefore convenient for the first coupling step. Furthermore, the first arm will be capped with an acetyl group (red) and the GCP unit will be attached to the second arm without an amine-containing linker (green). This way the overall charge of the receptor will be reduced by two positive charges. This should lead to a higher pK_A value of the GCP unit, which is normally between 6 and 7 but was as low as 5-6 with the original design, which is inconvenient for binding studies at physiological pH. The higher the guanidine pK_A value, the more it will be protonated at neutral pH and the higher should be the binding affinity to the carboxylate group of the substrate. Lastly, the TFA counterions will be exchanged for chlorides (red). This is necessary because the trifluoroacetate anions are able to interact with the GCP unit which is known to bind efficiently to carboxylates in aqueous medium.[19] Hence,

TFA would compete with the actual substrate for the carboxylate binding site thus reducing the affinity of the receptor.

Additionally, the receptors will be synthesized with the help of microwave-assisted Fmoc-SPPS, which has shown to be a viable tool for improving the yield and purity of difficult peptide sequences.[197] With the help of a single-mode microwave reproducible results can be achieved under controlled conditions. Generally, irradiation of the reaction mixture by microwaves is a very efficient way to heat the solvent in a very homogenous fashion. This way all the reactants have the same energy and different reaction pathways caused by higher or lower energy are avoided. The heat is generated due to the excitation of rotational modes of polar molecules which are able to interact with the microwave irradiation via their dipoles. This way, the solvent molecules as well as the reactants are heated directly. In comparison, the standard heating method utilized in most chemical laboratories is the immersion of a flask into an oil bath which is heated on a heating plate. Thus, the heat is transferred from the oil bath to the outside of the flask and from there distributed by stirring the solvent with a stirring bar. This method leads to a very unbalanced distribution of heat. The consequence is spots of higher as well as lower temperature inside the reaction mixture. This may then lead to the formation of unwanted side products. Because standard SPPS is carried out at room temperature and microwave-assisted SPPS on the other hand at elevated temperature, another general advantage of the microwave is shorter reaction times for coupling and deprotection steps.

Before synthesizing the non-symmetric tweezer receptors with the help of irradiation, the microwave reaction conditions had to be optimized: Time, temperature and microwave power can be altered for a given coupling step. Microwave-assisted coupling reactions have been reported to proceed without racemization at temperatures as high as 80° C.[198] Therefore, a temperature of 60° C was chosen for the coupling steps during the synthesis of the tweezer receptor, which should ensure fast reaction without racemization. Test reactions were carried out by coupling Fmoc-Val-OH with the help of PyBOP (each 3 eq) in 5 % DIPEA/DMF. The Fmoc group was deprotected with 20 % piperidine/DMF and the resin thoroughly washed with DMF. The solutions were collected, joined, and a UV/Vis spectrum was recorded. The yield of each coupling step was determined by calculating the concentration of the Fmoc-piperidine adduct as described in chapter 4.1.2. Variation of the Power (10-180 W) did not lead to significant differences of the coupling yield. Variation of the time (5-20 min) had a more pronounced effect. The best results were obtained for the longest coupling time. After 20 min almost quantitative conversion could be observed. Hence, reaction conditions for the coupling steps with irradiation to a maximum temperature of 60° C for 20 min at 20 W were chosen.

With these settings the first receptor was synthesized on MBHA resin according to standard Fmoc-procedure as shown in Figure 85.[178] All coupling steps were conducted under argon atmosphere. The resin was at first swollen in DCM for one hour, because polystyrene-based polymers like the MBHA resin feature excellent swelling properties in this solvent. After washing with DMF a glycine spacer was attached to the solid phase as Fmoc-Gly-OH (3 eq) making use of the optimized microwave conditions with the help of PyBOP (3 eq) as coupling agent and 5 % diisopropylethyl amine (DIPEA) in DMF. To ensure complete conversion of all free amine groups on the resin, the first step was repeated. In order to cap eventually

remaining amine groups the resin was then treated with acetic anhydride in 5 % DIPEA/DMF. After each coupling step the resin was washed with DMF. Successful, quantitative coupling steps were verified by a negative Kaiser test.[181] Fmoc deprotection was carried out by treating the resin twice with 20 % piperidine/DMF under microwave irradiation (1 + 5 min, max. 60° C, 20 W). The liberation of free amine groups was verified by a positive Kaiser test. In the next step, the orthogonally protected scaffold **181** was coupled to the glycine spacer as described above. The next three amino acids, Fmoc-Ser(Bn)-OH, Fmoc-Tyr(Bn)-OH, and Ac-Gly-OH were attached accordingly. If necessary, coupling steps resulting in a positive Kaiser test were repeated.

Since the next step of the synthesis is the deprotection of the Boc group at the aromatic template, side chains in the first arm may not feature Boc or *t*Bu groups. However, the according Cbz- or benzyl (Bn) protected amino acids, which are stable under the conditions necessary for Boc deprotection are also commercially available. The last amino acid can directly be coupled as the acetyl (Ac) derivative as persistent cap for the N-terminus of the first arm. The subsequent deprotection was then carried out with 50 % TFA/DCM without microwave irradiation. The resin was thoroughly washed with DCM and 5 % DIPEA/DMF to get rid of all TFA molecules, which could potentially react with the free amine group of the template under coupling conditions. The two following amino acids, Fmoc-Asp(O*t*Bu)-OH and Fmoc-Phe-OH, as well as the GCP unit (**158**) were then again coupled under standard microwave conditions with the exception of the Fmoc deprotection step after the aspartic acid coupling, which was carried out with 5 % piperazine and 0.1 M HOBt instead of 20 % piperidine. This alteration was necessary to prevent aspartimide formation and consequent racemization during microwave-assisted Fmoc deprotection.[198]

The cleavage of the resin-bound receptor was achieved by treating the solid support with 8 % trifluoromethanesulfonic acid (TFMSA) in 80 % TFA for 4 h. Under these strongly acidic conditions the receptor is released from the resin and all side-chain protecting groups are cleaved. The deprotection mixture also contained 8 % thioanisole and 4 % 1,2-ethanedithiol (EDT) as nucleophilic scavengers for cationic species which are generated under these conditions and could otherwise add irreversibly to nucleophilic groups of the product (e.g. unprotected hydroxy groups). After precipitation from the cleavage mixture with excess Et_2O, the weight of the crude product was higher than expected (> 100 % yield). The analysis of the corresponding mass spectrum revealed compounds which still feature one or both benzyl protecting groups. Obviously these groups are not removed from the hydroxyl groups even under the harsh acidic cleavage conditions. Therefore the raw product was hydrogenated with H_2 and 10 % Pd/C in acetone for 3 h and after purification via reversed phase MPLC (RP18, H_2O → MeOH, 0.1 % TFA) the product was isolated with 53 % yield. Analytical HPLC however revealed a purity of only 24 %. MS analysis could show that the main byproduct was a methylated species, probably the methyl ester of the aspartic acid which was formed when distilling the acidic water/methanol mixture obtained from the liquid chromatography under reduced pressure and elevated temperature. After one more purification step via RP MPLC the desired product **187** could be obtained as TFA salt with 99 % purity (see Appendix C.1, Figure C. 1). Still, when comparing the results of the microwave-assisted synthesis with those from the standard procedure it becomes clear that both the yield (45 → 53 %) and the purity (15 → 24 %) of the raw product could be

improved. Furthermore, instead of two weeks the synthesis can now be carried out in merely two days.

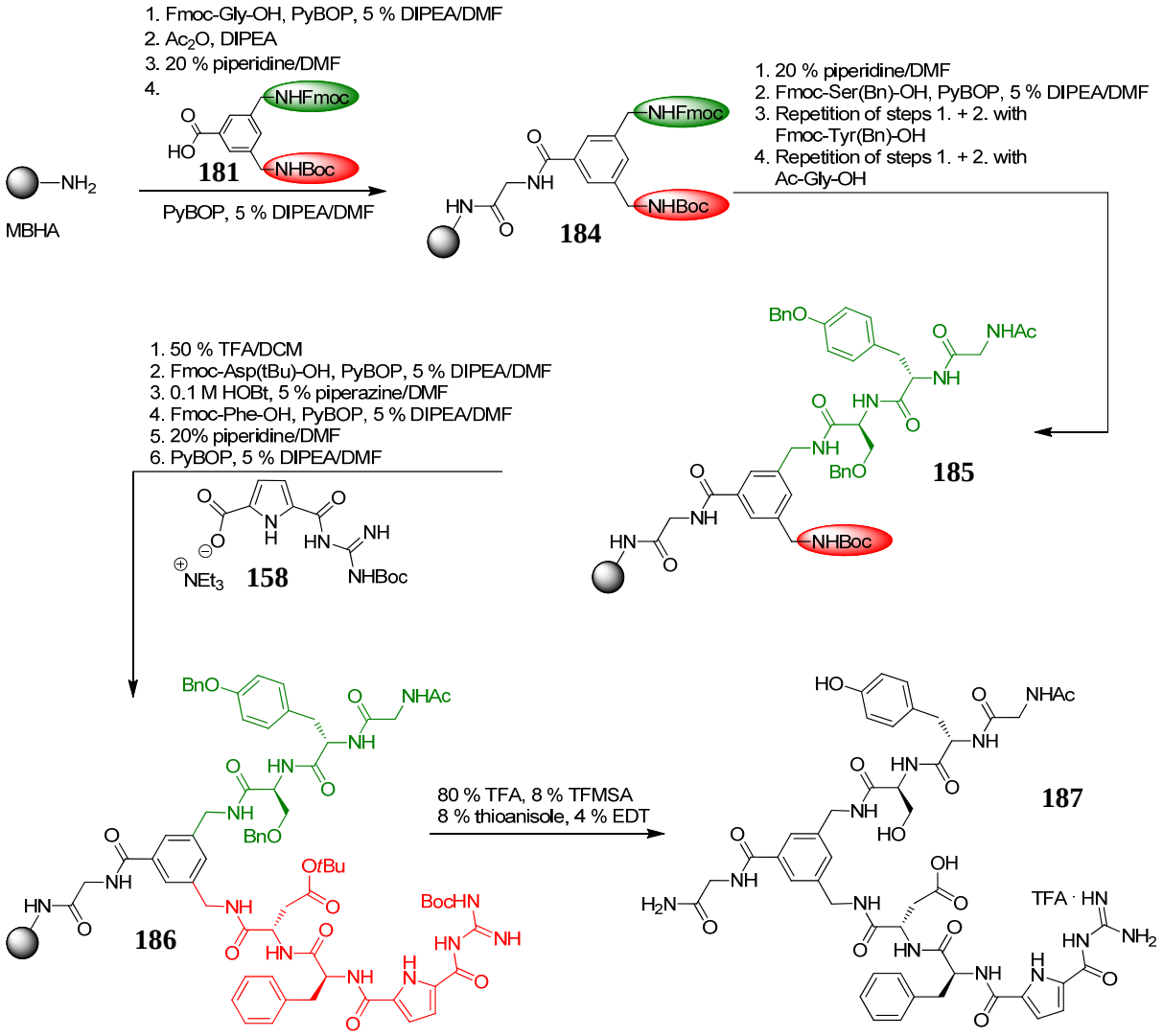


Figure 85. *The Fmoc SPPS for the non-symmetric tweezer receptor* ***187*** *was carried out with a microwave. Resin-attachment and coupling: 20 min, 20 W, max. 60° C. Fmoc deprotection: 1 + 5 min, 20 W, max. 60° C.*

According to this synthetic route a second non-symmetric tweezer receptor (**188**, Figure 86) was prepared with the amino acid sequence (from C- to N-terminus) Leu-Val-Ala in the first and Phe-Lys-GCP in the second arm. In this sequence no protecting groups are present in the first side chain. After a first MPLC (RP18, $H_2O \rightarrow$ MeOH, 0.1 % TFA) the crude product could be isolated with 71 % yield and 52 % purity. After two more purification steps the desired product was 90 % pure according to analytical HPLC (see Figure C. 2). The so achieved yield and purity are excellent when considering the complexity of the receptor, and the implementation of a rather unreactive aromatic amino acid (**181**) and a non-peptidic building block (**158**).

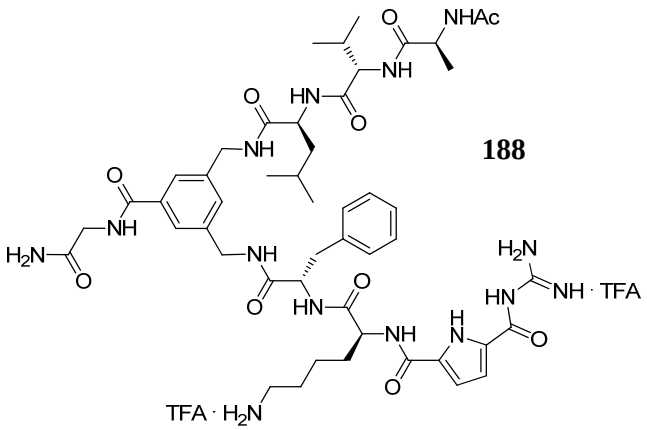


Figure 86. *The second non-symmetric tweezer receptor* ***188*** *was also prepared via microwave-assisted SPPS.*

Before carrying out the binding studies, the TFA counterions of the two receptors **187** and **188** were exchanged for chlorides by dissolving the compounds several times in dilute hydrochloric acid with subsequent lyophyilization. The receptors were now ready for binding studies.

4.2.3 Preliminary Binding Studies with N-Ac-Lys-D-Ala-D-Ala-OH

After having successfully obtained two non-symmetric tweezer receptors their physicochemical properties were characterized first. In order to determine their pK_A values the receptors were dissolved in water (0.1 mM) and half an equivalent of sodium hydroxide (0.1 N) was added and the pH value was determined. According to the *Henderson-Hasselbalch* equation (1) under these conditions the concentration of the protonated (HA) and of the deprotonated species (A^-) is equal and therefore pH = pK_A is valid.

$$pH = pK_A - \log \frac{[HA]}{[A^-]} \qquad (1)$$

The second pK_A value of each receptor was then determined by adding another equivalent of base and consequently measuring the resulting pH value. The results are listed in Table 4. Both receptors feature two different protonation sites. The carboxylic acid of the aspartic acid side chain in **187** has a pK_A of 4.82 (pK_A^I), and the GCP unit´s pK_A is 7.06 (pK_A^{II}). In comparison, the GCP moiety of **188** has a pK_A of 6.73 (pK_A^I) and its lysine ammonium group 9.82 (pK_A^{II}). The acidity of the guanidinio group in the receptors is increased tremendously in comparison to non-derivatized guanidinium (pK_A = 13.71)[199] due to the electron withdrawing carbonyl group, and to a lesser extent due to additional charges in the molecule. The hydrogen donor capability is thus greatly improved. However, a too low pK_A value, like in the receptors developed during my diploma (pK_A = 5-6), leads to the loss of the salt bridge. When comparing the pK_A value of **187** with the analogue receptor of former design, it becomes immediately clear, that the changes indeed led to a decrease of the acidity of the guanidinio group by approximately one order of magnitude (pK_A^{II}: 5.98 → 7.06).[195] The reduced acidity is mainly a result of the removal of positive charges in the receptor and should lead to improved binding properties towards carboxylate containing substrates. Hence, with the improved design of the receptors, the carboxylate binding site is now protonated to ca. 50 % at neutral pH instead of only ca. 10 %. At pH 6 already approx. 90 % of the guanidine groups are protonated.

In order to minimize sources of errors caused by weighing in of the samples, concentrated stock solutions of all compounds were prepared. These can then be diluted with higher precision to the desired concentration during the experiments. For the preparation of these solutions it was necessary to test whether the substances are water soluble at 1 mM concentration. This was only the case for **187**, but not for the less polar **188**. Therefore, the stock solution for the latter receptor was prepared in DMSO.

Table 4. *Physicochemical characterization of receptor **187** and **188**.*

Compound	pK_A^I	pK_A^{II}	Water solubility at 1 mM	ε [L/mol · cm]		
				pH 5	pH 6	pH 7
187	4.82	7.06	+	14920	15950	16440
188	6.73	9.82	-	15440	16370	17700

The determination of the binding constant was carried out with the help of a UV/Vis titration. Upon complexation, the intermolecular interactions between receptor and substrate lead to a redistribution of electron density which can be observed in the UV/Vis absorption of

the supramolecular system. The binding of carboxylates will take place via the GCP unit. Among other interactions a hydrogen bond will be formed between the pyrrole and the carboxylate.[200] This leads to a decrease of absorption at the pyrrole band at 300 nm, which can then be used for the determination of the binding constant. In order to prevent any effects derived from changes in the pH value during the titration all experiments were carried out in buffered solution (bis-tris buffer, 20 mM) at constant pH. This precaution was necessary because slight pH changes already influence the outcome of the experiments very strongly. Therefore, the concentration of the buffer must at least be 100 times the concentration of substrate and receptor. The titration itself was carried out by titrating aliquots of substrate into a solution of receptor and subsequent measurement of the UV/Vis spectrum around 300 nm. A successful complexation is characterized by a change of absorption which differs from the reduction of signal due to dilution. In order to differentiate between these two effects, dilution series of all compounds were carried out and the extinction coefficient ε was calculated according to the *Lambert-Beer* law (2). Furthermore, this way it could be shown that the absorption is directly proportional to the concentration in the range under investigation and therefore follows equation (2). This is an essential requirement for the straightforward determination of binding constants by means of spectroscopic methods.

$$A = \varepsilon \cdot c \cdot d \qquad (2)$$

Since the pyrrole´s ε depends on pH, three different values were determined for the two compounds: for pH 5, 6, and 7 (Table 4). With increasing pH the molar absorption coefficients increase as well. Within the pH range from 5 to 7 this increase is linear. The degree of protonation of the adjacent guanidinio group, which acts as a buffer in this concentration range and is therefore steadily protonated, affects the electronic surroundings of the pyrrole ring and therefore also its absorption properties. At pH 7 **187** and **188** absorb light at 300 nm more efficiently by approximately 10 %. The commercially available tripeptidic substrate N-Ac-Lys-D-Ala-D-Ala-OH (**189**, Figure 87), which plays a decisive role in the formation of the cell wall of *Gram* positive bacteria (see chapter 2.1.2), does not absorb at 300 nm.

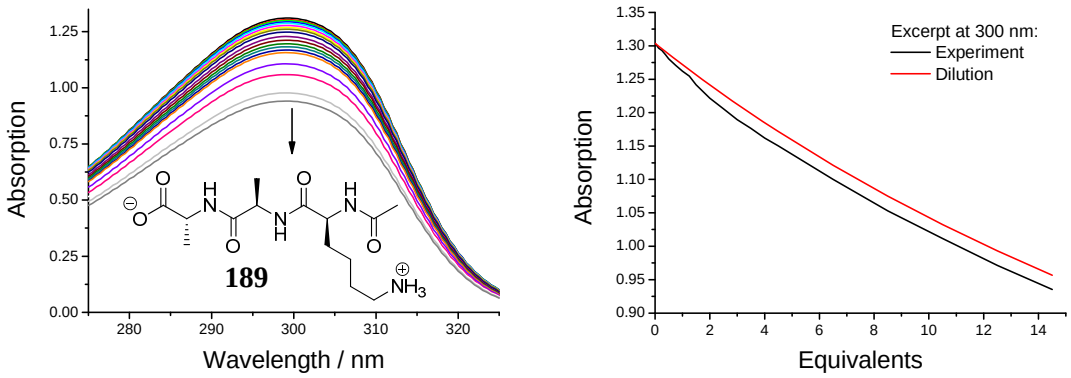


Figure 87. *Titration of aliquots of substrate **189** into a solution of **187** leads to a decrease of UV absorption (left). Comparison with the dilution shows that there is an additional effect: binding to the substrate (right).*

The titration was carried out in bis-tris buffer (0.02 M) with a receptor concentration of 5×10^{-5} M in a microcuvette (750 µL) equipped with a stopper at 25° C. Substrate of the

concentration 2×10^{-3} M was titrated into the solution in aliquots of increasing volume (e.g. 4 × 2.5, 4 × 5, 6 × 10, and 5 × 40 µL) up to ca. 15 equivalents. After one minute incubation time the resulting UV spectrum was recorded after each addition. All experiments were carried out in duplicate; the reported binding constants are average values. Figure 87 depicts a typical titration experiment (**187** at pH 5) and the comparison of the excerpt at 300 nm with the calculated dilution. As expected, the UV absorption decreases during the experiment. The comparison between experimental titration curve (black) and calculated dilution curve (red) reveals that the decrease of absorption is non-linear and thus, that a binding event is taking place between **187** and the substrate involving the GCP unit. Assuming a 1:1 receptor-substrate complex formation the binding constant can be calculated with the help of the following equation[28] which can again be solved through a non-linear regression analysis for instance with the program *Origin*.[201]

$$A = \epsilon_R[R_0] + \epsilon_S[S_0] + \frac{(\epsilon_{RS} - \epsilon_R - \epsilon_S) \cdot K \cdot [R_0] \cdot \left\{\frac{K \cdot [R_0] \cdot [S_0]}{(1 + K \cdot [S_0])^2 + K \cdot [R_0]}\right\}}{1 + K \cdot \left\{\frac{K \cdot [R_0] \cdot [S_0]}{(1 + K \cdot [S_0])^2 + K \cdot [R_0]}\right\}}$$

A	=	Absorption at 300 nm
ε_R	=	Extinction coefficient of receptor
ε_S	=	Extinction coefficient of substrate
ε_{RS}	=	Extinction coefficient of receptor-substrate complex
$[R_0]$	=	Starting concentration of receptor
$[S_0]$	=	Starting concentration of substrate
K	=	Binding constant

All variables but two, the binding constant K and the extinction coefficient of the complex ε_{RS}, are known. Thus, the equation has to be solved iteratively in order to obtain the binding constant. Figure 88 shows the experimental data from the exemplary titration experiment by plotting the equivalents against the absorption at 300 nm (black) and the corresponding iteratively obtained least-squares fit (red).

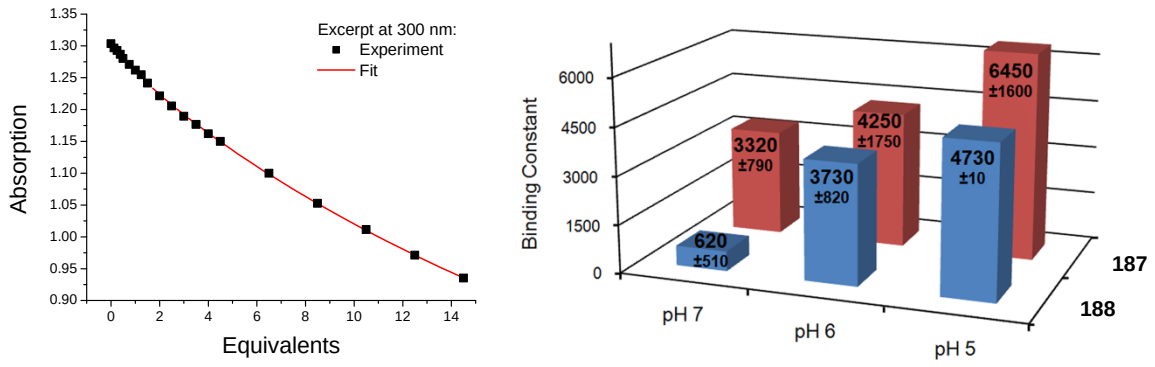


Figure 88. *Non-linear least-squares fit for the titration of the tripeptidic substrate KAA with* ***169*** *at pH 5 (left). Graphic representation of the measured binding constants for* ***187*** *and* ***188*** *with KAA at three pH values (right).*

According to this procedure the complexation of receptors **187** and **188** with the substrate KAA (**189**) was measured at three different pH values (5, 6, and 7). The results of these measurements and the corresponding standard deviations are also depicted in Figure 88

(right). The red bars represent the binding constants for **187** (Ser-Tyr-Gly; Asp-Phe-GCP), the blue ones respectively for **188** (Leu-Val-Ala; Phe-Lys-GCP). A huge difference of the binding constant was observed at pH 7. While **188** only binds with 6.2×10^2 M^{-1}, **187** binds 5 times as efficient to the substrate ($K = 3.3 \times 10^3$ M^{-1}). Only at pH 6 **188** ($K = 3.7 \times 10^3$ M^{-1}) does also bind to the substrate with similar binding constant as **187** ($K = 4.3 \times 10^3$ M^{-1}), although the latter one is still somewhat more efficient. At pH 5, again **187** ($K = 6 \times 10^3$ M^{-1}) binds more efficiently than **188** ($K = 5 \times 10^3$ M^{-1}). Generally, the binding constant is higher at lower pH. This effect can be linked to the increased protonation of the guanidine unit. At pH 7 for instance only ca. 54 % of **187**´s guanidine groups are protonated and even less for **188** (35 %). At pH 6 on the other hand, the majority of guanidine moieties is protonated (**187**: 92 % and **188**: 84 %). At pH 5 finally, almost all guanidines are in their protonated state (> 98 %). Since protonation is essential for the formation of the hydrogen bonded salt bridge to the C-terminus of the substrate, efficient binding only takes place when the majority of the GCP units are protonated. Another general trend is, that **187** has a higher affinity to KAA at all three pH values. In part, this can be explained by the differing pK_A values (**187**: 7.05 and **188**: 6.73). As shown above the consequence is, that **187** features on average a higher charge at its carboxylate binding site. This can however not entirely explain the huge difference at pH 7 and the difference at pH 5, when practically all groups are protonated. A possible explanation is an additional salt bridge between the aspartic acid, which is negatively charged at this pH, and the ammonium group of the substrate´s lysine side chain. Furthermore, in contrast to receptor **188**, the hydroxy groups in the serine and tyrosine side chains in the first arm of **187** are potential hydrogen donor and acceptor sites and might thus slightly stabilize the complex via hydrogen bonding. The huge increase of affinity between pH 7 and pH 6 for **188** demonstrates, that the binding strength of this system almost exclusively stems from the interaction between carboxylate and the GCP group, because there are no other groups in the molecule which can be protonated or deprotonated in this pH range. This is further evidence that there must form additional interactions between the substrate and **187**.

However, although these experiments concerning the interaction between the two non-symmetric receptors and the tripeptide substrate are an illustrative example on how subtle changes may affect the molecular recognition of a given host-guest system, the overall binding constants are rather moderate. In comparison with the one-armed analogue GCP-KKF (**13**, Figure 5), which has a similar amino acid composition as **187** and is able to bind to the tetrapeptide sequence Glu-Lys-D-Ala-D-Ala (**14**) with a binding constant of ca. 10^4 M^{-1}, these values are lower by a factor of 2-3.[13] Hence, the addition of the second arm, at least with the amino acid sequences utilized in this approach, does not seem to be beneficial for the recognition of this biologically relevant target. Therefore, no advanced additional studies for the detailed analysis of molecular recognition events based on these non-symmetric tweezers have been carried out. The focus of this thesis will from here onward mainly be turned on the recognition of nucleotides and nucleic acids, as will be shown in the following chapters.

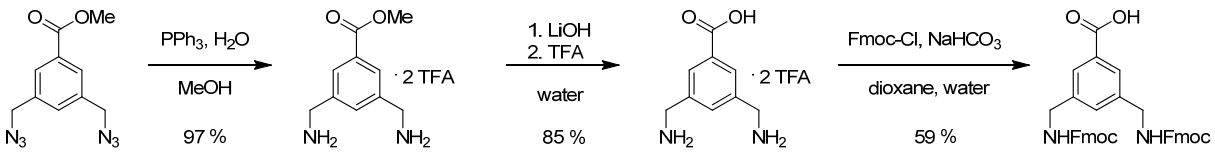
4.3 SYMMETRIC TWEEZER RECEPTOR FOR NUCLEOTIDE RECOGNITION

The recognition of phosphate-based anions like nucleotides requires a different approach in comparison to carboxylate binding. Unlike carboxylates, phosphates are not planar, but of tetrahedral shape. Furthermore, depending on the pH alkylated phosphates can carry up to two negative charges instead of only one. Hence, an approach with only one oxoanion binding site will probably not yield the desired result. Rather, phosphate and nucleotide recognition should be feasible when implementing two such units into a receptor in a fashion which allows them to adopt a geometry which matches that of the phosphate anion. In principle, the implementation of two GCP units into tweezer receptors similar to the non-symmetric ones described above, but with two identical arms and thus two oxoanion binding sites should be ideally fit for phosphate binding.

4.3.1 Synthesis of the Template

For such symmetric receptors an adequate template is necessary. Instead of two different protecting groups, only one is necessary. Already in the course of my diploma such an aromatic template with two Fmoc protecting groups was synthesized. Therefore, the reaction scheme will only be described in short, focusing on improvements in the synthetic procedure, which follows the exact same route as described for the orthogonally protected template **181** in chapter 4.2.1 up to the diazide **175**. Instead of reducing only one azide group, both are now reduced with triphenylphosphane and water to give **190** as shown in Figure 89. Instead of removing the methyl ester via saponification to **191** and attaching the Fmoc protecting group as in the former synthetic scheme, now the Fmoc group is introduced first (**192**) and the methyl ester is then cleaved under acidic condition with hydrochloric acid at 100° C with very good yield to give the template **193**, which is ready for the SPPS of symmetric tweezer receptors.

a) Original Synthetic Route



b) Novel Synthetic Route

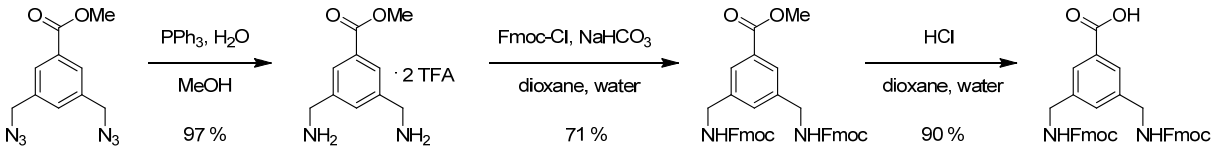


Figure 89. *Synthesis of the aromatic template **193** for the SPPS of symmetric tweezer receptors: a) according to the original; and b) according to the novel synthetic route .*

By switching the last two steps and replacing the basic with an acidic cleavage of the methyl ester, the yield over three steps could be improved from 49 to 62 %. The overall yield over six

steps starting with 3,5-dimethylbenzoic acid is 42 %. Thus, the problematic Fmoc protection step of the polar diamine benzoic acid derivative **191** with non-polar Fmoc-Cl with its crucible dioxane/water solvent ratio is circumvented: Small changes did dramatically alter the result. As a consequence, this reaction only yielded max. 59 % and was not very reproducible: It often failed totally or only gave low yields. Another problem was the purification by column chromatography. **193** forms aggregates, which hardly dissolve and only elute slowly, even at 10 % MeOH/EA. **192** on the other hand is easier to purify and the methyl ester **190** is less polar than the free acid and therefore better compatible with the solubility of Fmoc-Cl. Consequently, the yield for the Fmoc protection was now reproducible and higher, facilitating the isolation of **193**. It can be used for SPPS without chromatographic purification. The only side product is the starting material **192**, which does not affect coupling to the resin.

4.3.2 Synthesis of the Symmetric Tweezer Receptor

With the so obtained template it was now possible to prepare a tweezer receptor with two identical peptidic arms according to the optimized microwave-assisted SPPS strategy (see chapter 4.2.2). Since the receptor is supposed to bind to nucleotides in aqueous medium, it is important to choose appropriate amino acids to improve its binding properties. Therefore, lysine and phenylalanine were chosen for potential additional salt bridges between the lysine ammonium groups and the phosphate hinge of the nucleotides as well as π-stacking between the aromatic ring of phenylalanine and the nucleobases. Finally, two GCP headgroups were integrated into each arm for the recognition of a tetrahedral phosphate. As shown in Figure 90 the synthesis starts again by attaching a glycine spacer to MBHA resin in the form of Fmoc-Gly-OH (3 eq) with the help of PyBOP (3 eq) in 5 % DIPEA/DMF under microwave irradiation (max. 60° C, 20 min, 20 W) and argon atmosphere. In order to assure complete conversion of all amine groups of the resin, the first coupling was repeated twice and the solid support was afterwards treated with acetic anhydride to cap eventually leftover free amine groups. Fmoc deprotection was achieved by treating the resin with 20 % piperidine/DMF under microwave irradiation (max. 60° C, 1 + 5 min, 20 W). The attachment of the aromatic template **193** was carried out accordingly. After another Fmoc deprotection the next two amino acids Fmoc-Phe-OH and Fmoc-Lys(Boc)-OH, as well as the oxoanion binding site **158** were coupled to the growing peptide chain. Instead of utilizing 3 eq of each reactant, now the double amount was necessary, because of the doubling of amine groups. If necessary, the coupling steps were repeated as indicated by the results of the Kaiser test.

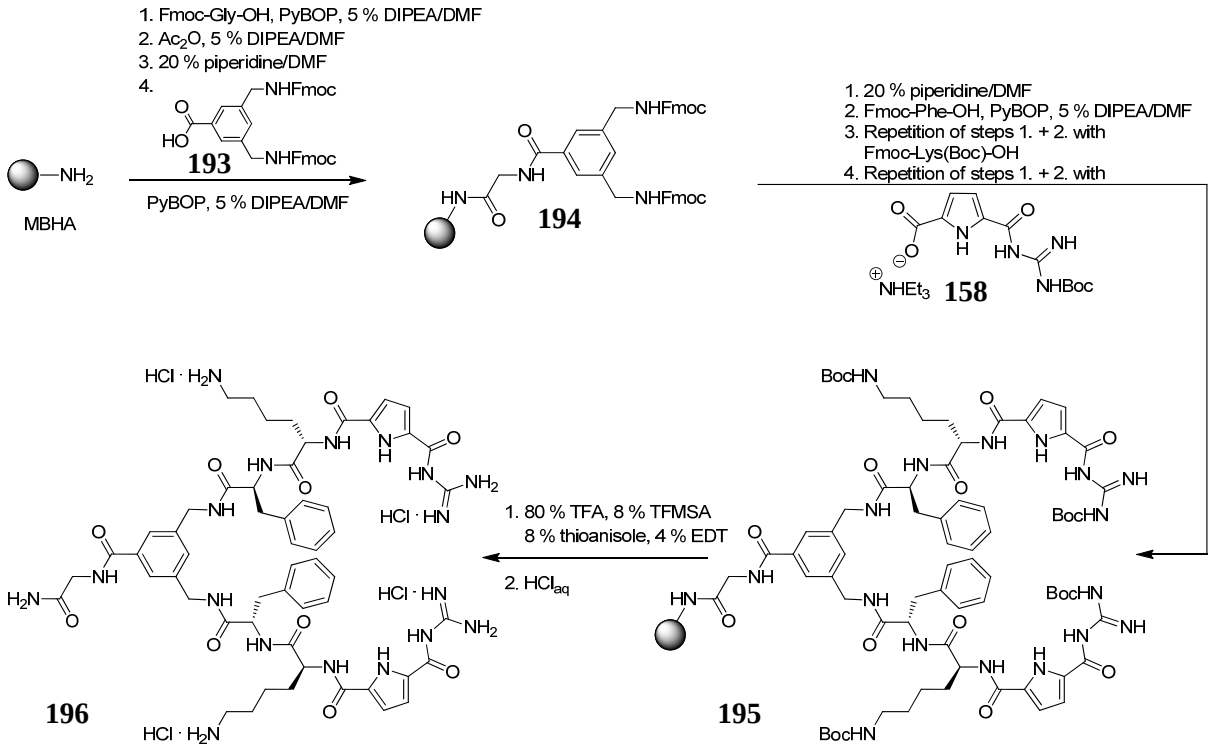


Figure 90. *The Fmoc SPPS for the non-symmetric tweezer receptor **196** was carried out with a microwave. Attachment and coupling: 20 min, 20 W, max. 60° C. Fmoc deprotection: 1 + 5 min, 20 W, max. 60° C.*

Finally, the resin bound receptor **195** was cleaved from the solid-support, at the same time removing all Boc protecting groups, by treating the resin with 80 % TFA and 8 % TFMSA together with the nucleophilic scavengers thioanisole (8 %) and EDT (4 %). After precipitation in Et_2O the crude product was obtained after a first MPLC (RP18, H_2O → MeOH, 0.1 % TFA) with 53 % yield and 74 % purity. After one more purification step via RP MPCL the desired symmetric tweezer receptor was obtained with 94 % purity (see Figure C. 3). Finally the trifluoroacetate counterions were replaced for chlorides by dissolving the receptor several times in hydrochloric acid (0.1 N) and consequent lyophilization to obtain **196** as the hydrochloride salt which can thus be used for binding studies. In comparison to the standard SPPS at ambient temperature without microwave irradiation, as carried out during my diploma thesis,[195] both the yield (20 → 53 %) and the purity (60 → 74 %) could be increased. Furthermore, the reaction time could be reduced from one week to a single day. Hence, microwave-assisted SPPS has proven to be an excellent method for the rapid and convenient synthesis of oligopeptidic tweezer structures, especially if low reactive non-proteinogenic amino acids like **193** or non-peptidic compounds (**158**) are to be implemented into the receptor.

4.3.3 Binding Studies

Prior to the binding studies a physicochemical characterization of the compound was carried out. It revealed a good solubility up to 1 mM in water. The pK_A values of the guanidinio and ammonium groups could be determined with the help of the half-equivalent point method to 7.00 (pK_A^I) and 9.90 (pK_A^{II}), respectively. The values are average values for each group in both arms and were determined by dissolving 0.1 mM of the compound in water and adding one equivalent (not just a half) of sodium hydroxide (0.1 N) to the solution with subsequent determination of the pH. The second average pK_A value was determined by adding another two equivalents of the base. Here again, the pK_A value of the guanidinio group could be increased by more than one order of magnitude (5.52 → 7.00) in comparison to the analogue host system developed during my diploma studies, by removing charges from the receptor, thus ensuring an increased protonation at neutral pH. At pH 7 exactly 50 % of the guanidinio groups are protonated and **196** carries a total of three charges.

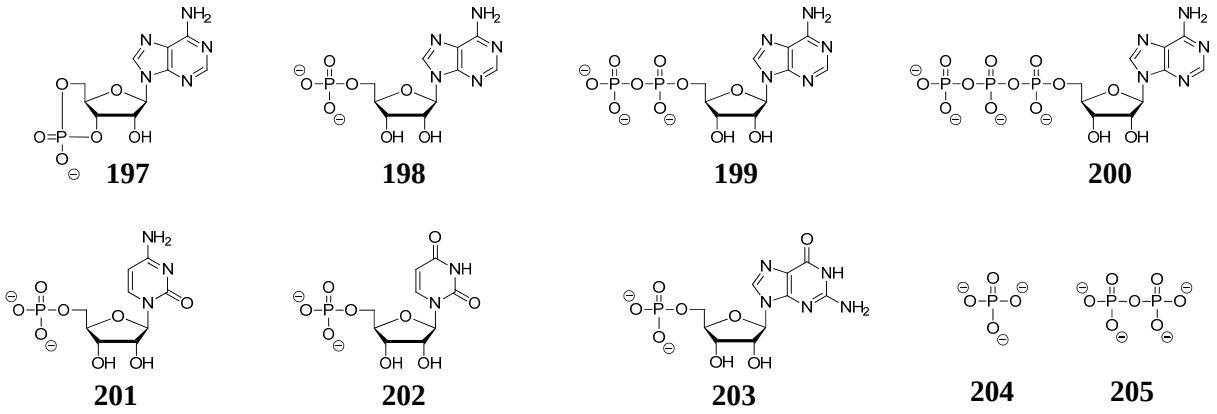


Figure 91. *Substrates for binding studies with the symmetric tweezer receptor **196**: cAMP (**197**), AMP (**198**), ADP (**199**), ATP (**200**), CMP (**201**), UMP (**202**), GMP (**203**), phosphate (**204**), and diphosphate (**205**).*

Binding studies were carried out in buffered water (bis-tris buffer, 20 mM) at 25° C and neutral pH with the substrates cyclic adenosine monophosphate (cAMP), adenosine monophosphate (AMP), adenosine diphosphate (ADP), adenosine triphosphate (ATP), cytidine monophosphate (CMP), uridine monophosphate (UMP), guanosine monophosphate (GMP), phosphate (P), and diphosphate (PPi) depicted in Figure 91 (**197**-**205**).

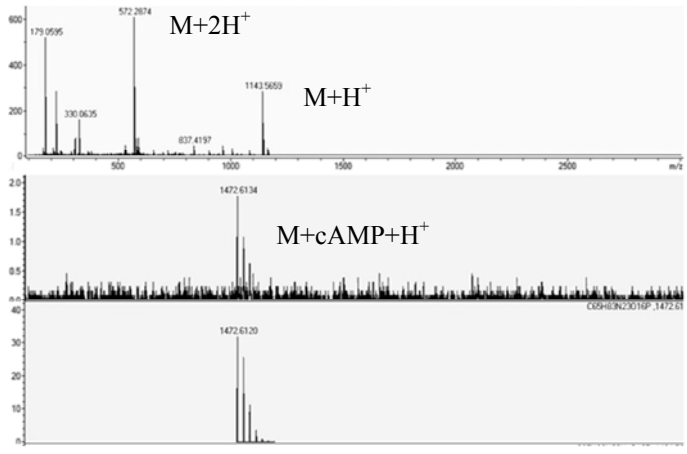


Figure 92. *Mass spectrum obtained from a solution of **196** (100 µM) and cAMP (200 µM) in water by ESI-ToF in positive operation mode. Top: Complete spectrum; middle: Excerpt; bottom: Calculated Mass.*

As depicted in Figure 92, the 1:1 complex of **196** with cAMP in water could be observed by means of ESI-ToF measurements in positive operation mode ([**196**] = 100 μM, [cAMP] = 200 μM). For AMP and UMP similar spectra could be obtained. These are clear indications, that complex formation between the symmetric tweezer receptor and nucleotides is possible. After these MS experiments and some preliminary UV/Vis titrations which showed that **196** indeed binds to these substrates, it was necessary to determine the receptor-substrate stoichiometry first, before conducting more advanced measurements for the exact determination of the binding constants. The knowledge of the stoichiometry is crucial for the calculation of the binding constants and for the selection of appropriate conditions for the titrations. By making use of the method of continuous variation known as *Job* plot, the stoichiometry can be determined.[202,203] Several solutions containing different molar fractions of receptor (R) and substrate (S) but with constant total concentration ([R] + [S]) had to be prepared. Complex formation should be maximal, when the ideal ratio of receptor and substrate is present in the solution, corresponding to the complex stoichiometry. By measuring an observable, which is directly proportional to the complex concentration a bell-shaped curve is the result. The maximum of this curve corresponds to the stoichiometry of the complex which can thus be determined. A maximum at a molar ratio of 0.5 indicates a 1:1 complex formation. A maximum at 0.33 or 0.66 corresponds to a 1:2 or respectively a 2:1 complex. Such a *Job* plot can also be carried out with the help of UV/Vis spectroscopy if the absorption of the non-covalent complex differs from that of the two reaction partners. Equation (4) then allows for directly observing the complex formation with the help of the change in absorption.

$$\Delta A = A_{exp} - A_R - A_S \qquad (4)$$

By subtracting the absorption of receptor (A_R) and substrate (A_S) of the concentration corresponding to the molar fraction from the experimentally observed absorption (A_{exp}), the difference of absorption (ΔA) is then directly proportional to the complex concentration. Generally, two common approaches exist to carry out *Job* plots: One may either prepare separate solutions with different molar fractions but of the same total concentration of R + S, and measure these separately, or one may extract the necessary data for the *Job* plot from the binding constant titrations. While the first approach is very substance-consuming, the latter one generates an inconvenient data point spreading. In a titration for the determination of binding constants, normally many data points are collected at low molar fractions and fewer at high ratios. For *Job* plots on the other hand, data points at the critical molar ratios (e.g. 0.5) or an even spreading of data points is desirable for the exact determination of the stoichiometry. Extracted data may therefore lead to inaccurate determination of the stoichiometry. This is why in this work a totally different approach was developed. First, a UV spectrum of pure receptor (60 μM) was recorded in bis-tris buffer (0.02 M) and the absorption at 300 nm was extracted. Then, a defined amount of the solution was removed, replaced by the same amount of substrate solution (60 μM), and the resulting absorption was measured. By repeating this procedure spectrums of varying molar fractions could be measured at constant total concentration. Finally, the spectrum of pure substrate (60 μM) was collected. This way, valuable amounts of receptor and substrate could be saved and a *Job* plot with the desired spacing of data points was derived. In Figure 93 the molar fractions of the substrates x_S are plotted against $\Delta A = |A_{exp} - A_R - A_S|$ for all possible combinations with the symmetric tweezer

receptor **196**. All experiments were carried out at least in duplicate. For all substrates the maximum of ΔA is at a molar fraction of 0.5 – the receptor forms 1:1 complexes to all substrates.

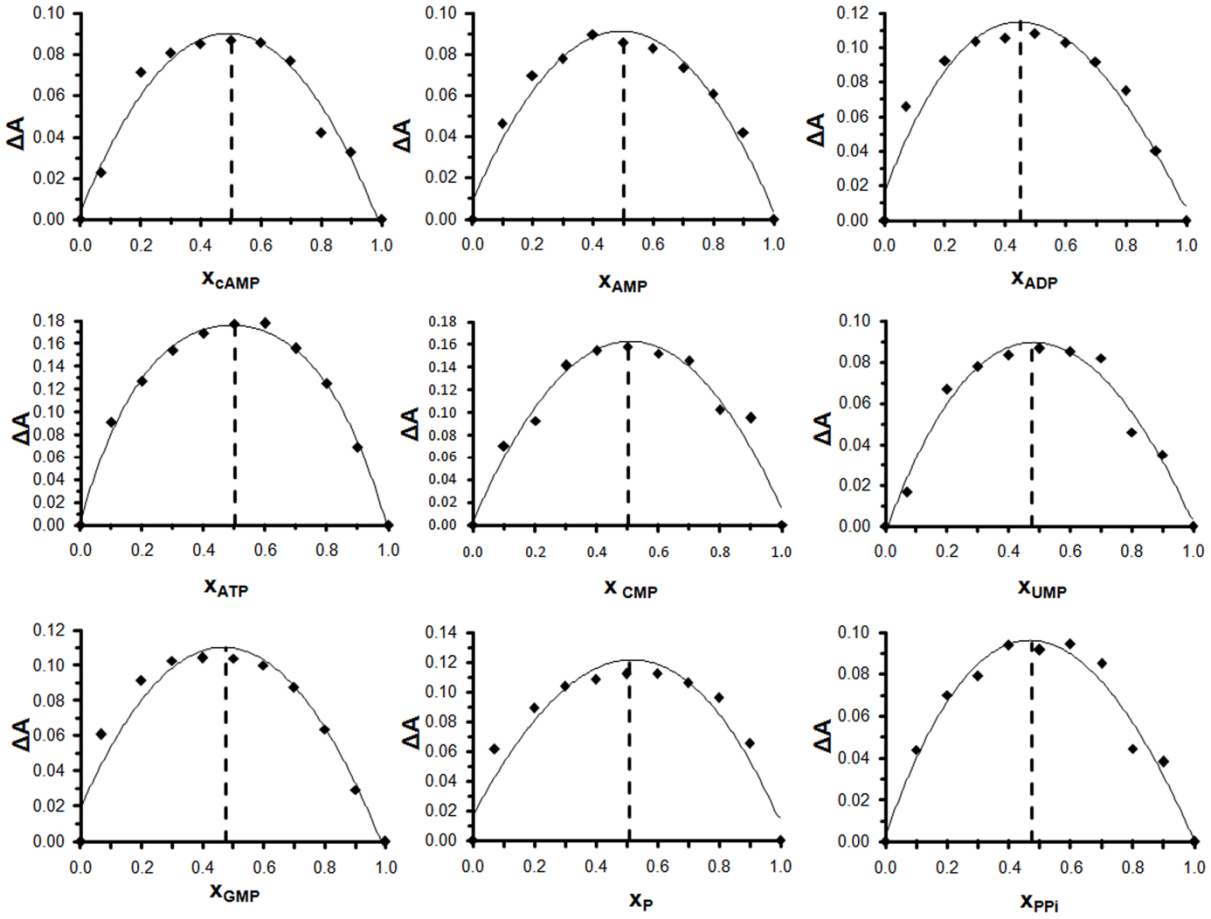


Figure 93. *Job plots for substrates **197-205** reveal 1:1 complex stoichiometry. Spectra were recorded at 300 nm in aqueous bis-tris buffer (20 mM) at pH 7 and a total concentration of [R] + [S] = 60 µM.*

With the knowledge of the complex stoichiometry, measurements concerning the binding affinity were conducted in bis-tris buffered water (20 mM) at neutral pH and 25° C. Experiments were carried out by adding aliquots of increasing volume (e.g. 8 × 10, 8 × 20, 8 × 40 µL) of substrates to a solution of **196** in a microcuvette equipped with a stopper (800 µL, 2-5 × 10^{-5} M) and subsequent recording of the UV/Vis spectrum from 250 to 350 nm after 1 min incubation time. In addition, similar titrations were carried out with the help of fluorescence spectroscopy. When irradiating the GCP unit at a fixed wavelength at its absorption maximum at 300 nm, a fluorescence signal around 355 nm can be observed. This signal´s intensity is decreasing upon complex formation, similar to the absorption signal. Again, substrate was added in aliquots of increasing volume (e.g. 8 × 5, 8 × 10, 4 × 20 µL) to a solution of **196** in a florescence microcuvette equipped with a stopper (900 µL, 0.5-1 × 10^{-5} M) and after 1 min incubation time the fluorescence emission was recorded from 335 to 400 nm. Figure 94 depicts representative examples of such UV and fluorescence titrations with AMP as substrate (see Figure C. 14-Figure C. 28 for all other titrations). The UV/Vis titration features a shifting isosbestic point – a clear indication for the formation of a host-

guest complex. All experiments were carried out at least two to three times. In the concentration range of the experiments both the fluorescence and the absorption decrease linearly with dilution and hence follow the *Lambert-Beer* law.

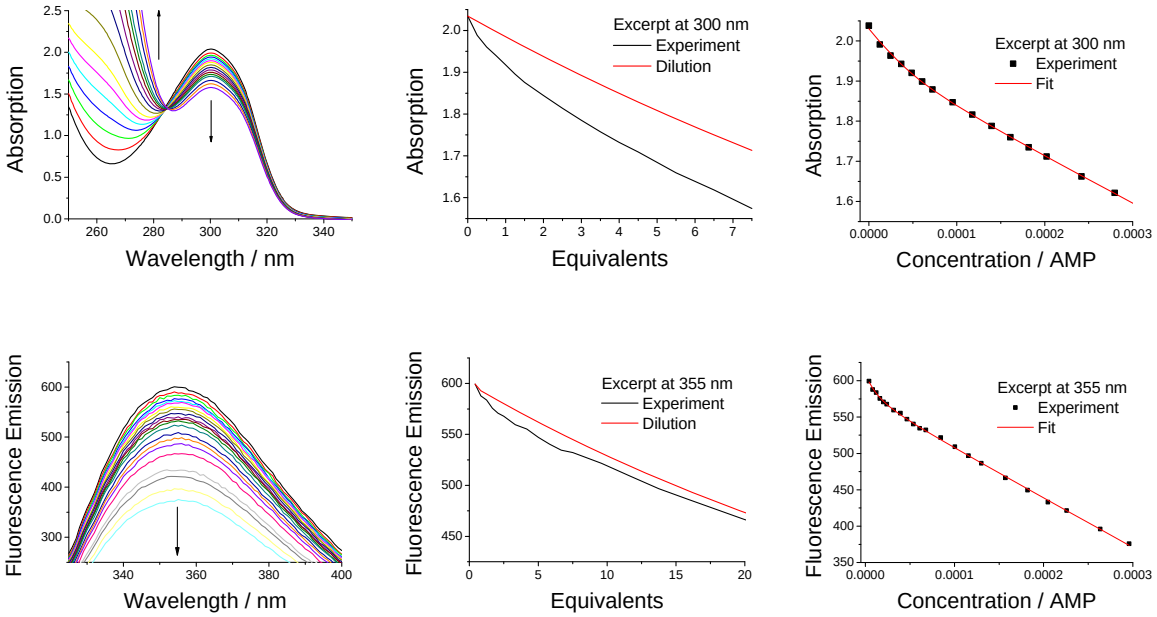


Figure 94. *Left: Absorption spectra of* ***196*** *upon incremental addition of AMP (2 mM) in aqueous bis-tris buffer (60 mM) at pH 7: UV/Vis ([**196**] = 5 × 10^{-5} M, top) and fluorescence spectrum ([**196**] = 10^{-5} M, bottom) (left). Middle: Comparison of dilution against excerpt of the experimental data at 300 nm. Right: Least-squares fit for the 1:1 complex formation as obtained by the program SpecFit.*

The plot comparing the theoretical dilution curve (red) of absorption or fluorescence versus the experimentally observed data (black) at 300 or 355 nm clearly indicates a complex formation. The data was fitted with the help of the program SpecFit developed by *Zuberbühler*.[204,205] The program makes use of a least-squares fitting procedure over the entire spectral range (instead of a single wavelength). In general, multivariate approaches generate a more accurate data-fitting than procedures which concentrate on a single wavelength. However, for accurate spectral fitting several prerequisites have to be fulfilled. First, the level of complex formation should reach approximately 80 %. This was achieved by titrating varying amounts of substrate into the receptor solution (e.g. 10 eq for AMP and 45 eq for P). Further, the spectral range for the fitting procedure must neither contain regions where the detection limit is exceeded nor overlapping absorption signals and isosbestic points. This was assured by only taking the spectral range between 290 and 325 nm for UV/Vis and between 345 and 370 nm for fluorescence titrations into account for the fitting procedure. As expected, fluorescence data was scattered stronger due to the increased sensitivity of the spectrometer. Still, excellent fitting curves could be obtained for all receptor-substrate combinations. The resulting binding constants are listed in Table 5 and visualized in Figure 95. The results obtained by the two independent methods are in excellent agreement – deviations are within the standard deviation. Generally, the standard deviation of the individual values lies within the error of the instrument, which is higher for fluorescence than for UV/Vis, as expected.

Table 5. *Binding constants (K and log K) and standard deviations (σ) as calculated by SpecFit for the formation of 1:1 complexes of* ***196*** *in bis-tris-buffered water (20 mM) at neutral pH.*

Substrate	*K* (UV) [M^{-1}]	log *K*	*σ* (UV) [M^{-1}]	*K* (Fluo) [M^{-1}]	log *K*	*σ* (Fluo) [M^{-1}]
AMP	76,400	4.88	6,200	73,000	4.86	14,000
ADP	32,500	4.51	3,300	37,100	4.57	15,700
ATP	8,300	3.92	1,400	11,400	4.06	2,700
PPi	17,000	4.23	2,100	10,800	4.03	3,400
P	13,400	4.13	2,400	17,900	4.25	3,400
cAMP	64,400	4.81	5,300	52,300	4.72	9,400
CMP	61,300	4.79	4,700	56,600	4.75	5,600
UMP	93,500	4.97	4,000	112,300	5.05	27,000
GMP	84,400	4.93	10,900	78,600	4.90	18,300

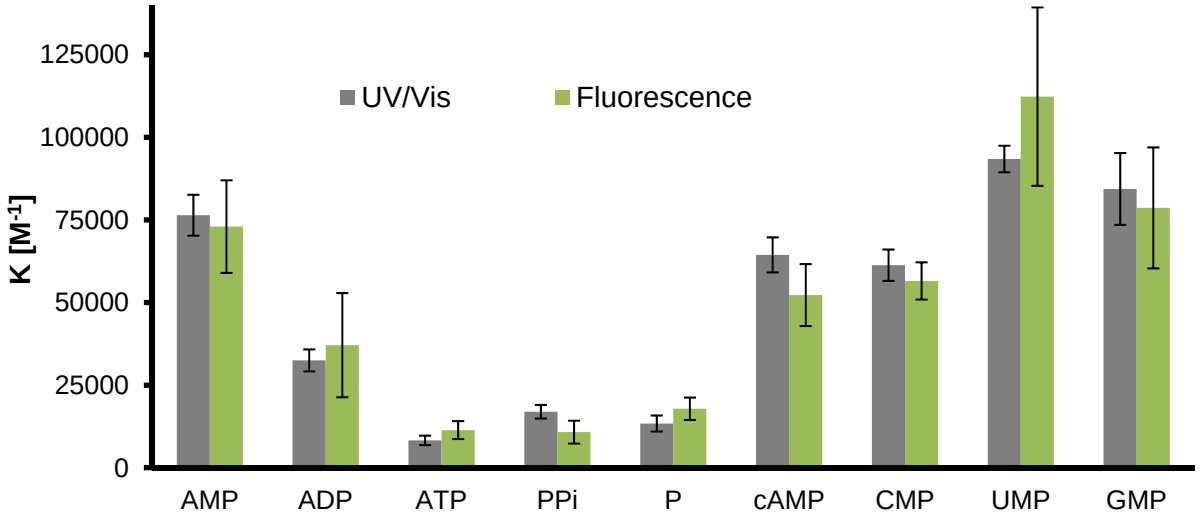


Figure 95. *The graphical representation of the binding constants as obtained by UV/Vis and fluorescence titration illustrates the excellent agreement between the two individual and independent methods.*

The general trend for binding affinity is UMP > GMP > AMP > cAMP ≈ CMP > ADP > ATP ≈ PPi ≈ P. Phosphate and diphosphate are both bound significantly weaker than nucleotides, with the exception of ATP. Therefore one can conclude, that **196** does not only interact via charge-charge interactions between the GCP guanidinio and/or lysine ammonium groups with the phosphate hinge, but that additional interactions are formed to the sugar and/or the nucleobases. The highest binding constant was found for UMP with $K = 9 \times 10^4\ M^{-1}$. However, there is no significant selectivity for any of the nucleobases. The weakest bound nucleotide CMP is still interacting with a binding constant of $6 \times 10^4\ M^{-1}$. According to that, the additional interactions are most likely non-directed π-π-interactions between the nucleobase and an aromatic moiety of the receptor. A more striking characteristic of the host **196** is, that it prefers to bind to AMP ($8 \times 10^4\ M^{-1}$) over ADP ($3 \times 10^4\ M^{-1}$) and ATP ($8 \times 10^3\ M^{-1}$) – the monophosphate is bound stronger than the triphosphate by one order of magnitude. This binding behavior is unusual and so far unprecedented in literature. Normally the higher charged nucleotides are preferred due to the formation of additional charge-charge interactions between receptor and substrate (see chapter 2.2). To investigate in more detail the

binding mode of **196** towards the substrates and to elucidate the exceptional preference for monophosphates, NMR spectroscopic measurements and molecular modeling of the complexes were conducted.

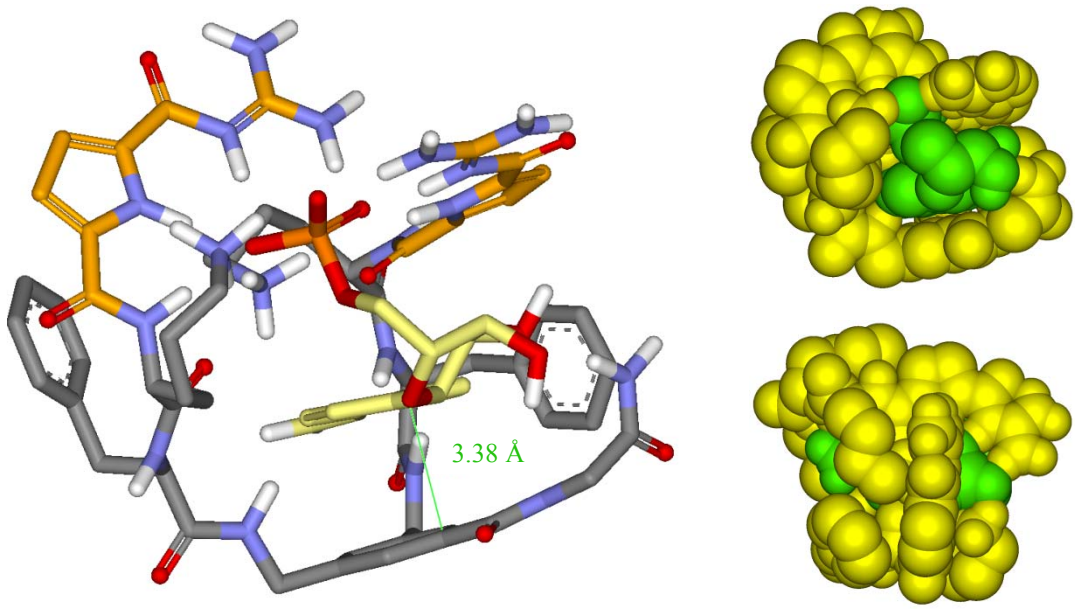


Figure 96. *Left: Side view of the calculated energy-minimized structure of the complex between* ***196*** *and AMP as obtained from a Monte Carlo conformational search. Non-polar hydrogen atoms have been removed for clarity. Color code: substrate (yellow), GCP (orange), and receptor backbone (gray). Right: Side (top) and top view (bottom) of the complex.* ***196*** *is colored yellow, AMP green. Atoms are displayed as spheres with their corresponding Van-der-Waals radius. Reprinted with permission from Kuchelmeister, H. Y.; Schmuck, C. Chem. Eur. J.* ***2011****, 17, 5311-5318. Copyright 2011 John Wiley and Sons.*[206]

The molecular modeling (MM) of the 1:1 complex between **196** and AMP was carried out with the program Macromodel v9.6 making use of the Amber* force field and a generalized Born/surface area (GBSA) water solvation.[207] The energy-minimized structure of a Monte Carlo conformational search of 50,000 steps without conformational restraints was subjected to a molecular dynamics (MD) simulation at 300 K for 500 ps. The so obtained structure is depicted in Figure 96 (left): Two GCP units (orange) are oriented perpendicular to each other thus binding the tetrahedral phosphate (orange-red) of the substrate (yellow). This is in agreement with the hypochromism of the UV absorption at the pyrrole maximum at 300 nm.

Furthermore, ^{1}H and ^{31}P NMR spectra of a 1:1 mixture of AMP and **196** were recorded at a concentration of 10 mM in DMSO-d_6 on a 500 MHz spectrometer.[208] The most important results, which are listed in Table 6, are in good agreement with the calculated structure. Excerpts of the ^{1}H spectra of the monomers as well as the 1:1 complex are depicted in Figure 97. The ^{31}P signal of the phosphate is shifted significantly upfield by 1.05 ppm with respect to pure substrate. On the side of the receptor, its guanidinium group $(NH_2)_2^+$ is shifted the most (-0.29 ppm). The lysine NH (-0.04 ppm), pyrrole CH (-0.10 ppm), pyrrole NH (-0.02 ppm) and the guanidinium NH (-0.10 ppm) peaks all shift as well. Further, the calculation indicates that the lysine ammonium groups are also involved in phosphate binding. This finding is supported by the NMR measurements: the lysine ammonium group (-0.12 ppm) and the neighboring δ- (-0.02 ppm) and γ-CH_2 groups (+0.01 ppm) all feature a complexation-induced

chemical shift (CIS). Another important structural detail is the π-stacking between the nucleobase and the aromatic template of the receptor. The calculated distance between these two rings is 3.38 Å (Figure 96, green line) – which is in excellent agreement to the ideal π-stacking distance of 3.4 to 4.5 Å as determined by a representative set of high resolution x-ray crystal structures.[209] Again, the experimental data supports this result: all involved proton signals are shifted: the aromatic protons of the nucleobase A-H2 (+0.01 ppm) and A-H8 (-0.18 ppm) as well as the template CH_{ar} signals (-0.02 and +0.03 ppm) undergo a CIS. Finally, as indicated in the space-filling representation of the calculated complex in Figure 96 (right), the receptor (yellow) forms a half-closed cavity for the accommodation of the substrate (green).

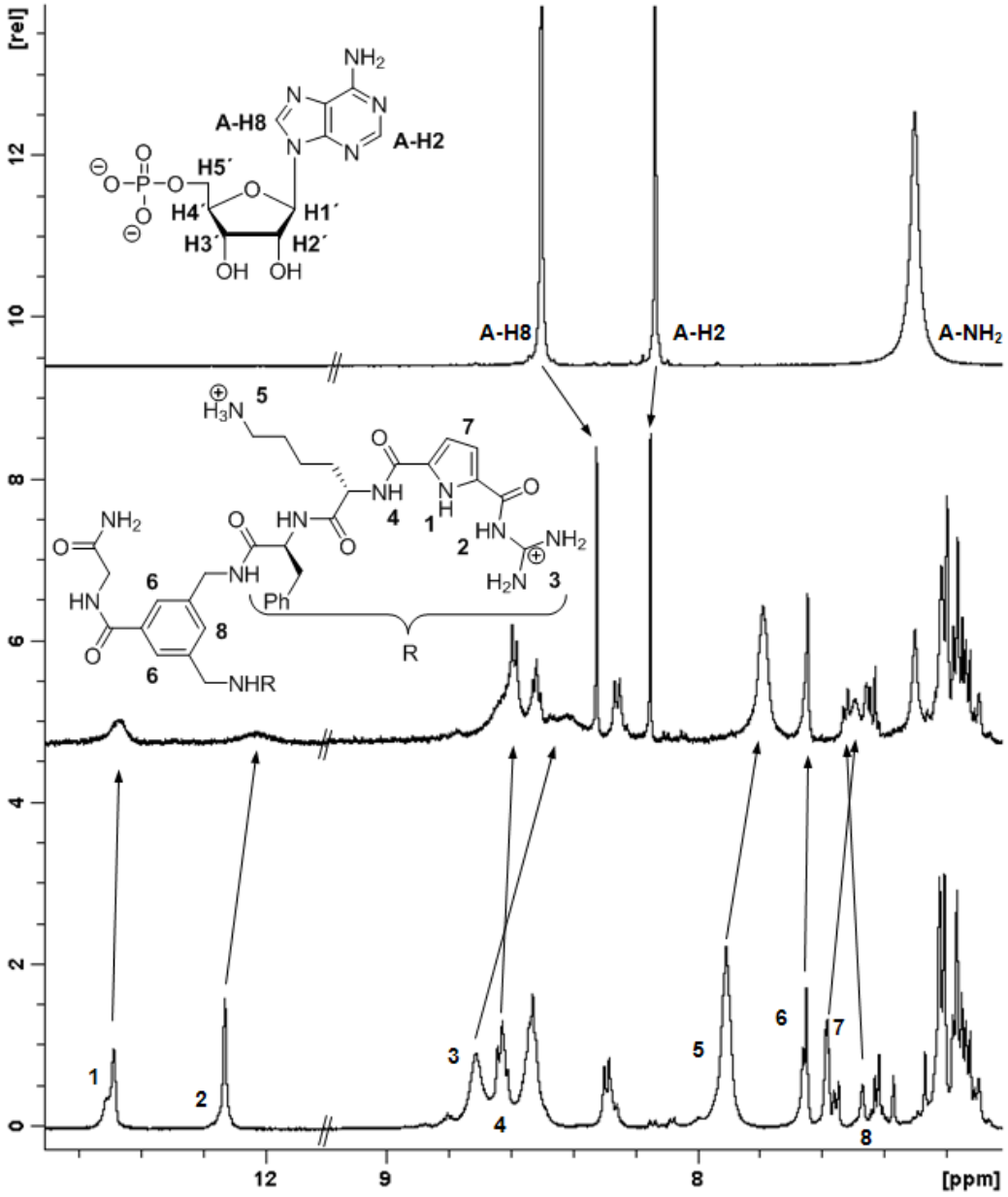


Figure 97. *Figure 98 ^{1}H NMR spectra of AMP (top), receptor **196** (bottom) and a 1:1 mixture thereof. Upon complexation, several proton signals shift. The spectra were recorded at 500 MHz in DMSO-d_6. Reprinted with permission from Kuchelmeister, H. Y.; Schmuck, C. Chem. Eur. J. **2011**, 17, 5311-5318. Copyright 2011 John Wiley and Sons.*[206]

Table 6. *Chemical shifts of selected ^{1}H and ^{31}P signals of **196** and AMP and the complexation-induced chemical shifts (CIS) of the corresponding 1:1 mixture. All measurements were carried out at a concentration of 10 mM in DMSO-d_6 at 25° C and 500 MHz.*

Signal (**196**)	δ [ppm]	$\delta_{1:1}$ [ppm]	CIS [ppm]	Signal (AMP)	δ [ppm]	$\delta_{1:1}$ [ppm]	CIS [ppm]
1: Pyr-NH	12.49	12.47	–0.02	A-H8	8.51	8.33	–0.18
2: Gua-NH	12.13	12.03	–0.10	A-H2	8.14	8.15	+0.01
3: Gua-$(NH_2)_2^+$	8.72	8.43	–0.29	A-NH_2	7.30	7.30	-
4:Lys-NH	8.63	8.59	–0.04	H1´	5.92	5.92	-
5: Lys-NH_3^+	7.91	7.79	–0.12	H2´	4.68	4.59	–0.09
6: 2 × CH_{ar}	7.66	7.64	–0.02	H3´	4.28	4.18	–0.10
7: Pyr-CH	7.58	7.48	–0.10	H4´	4.06	4.07	+0.01
8: CH_{ar}	7.47	7.50	+0.03	2 × H5´	3.85	3.96	+0.11
9: γ-Lys-CH	2.69	2.70	+0.01				
10: δ-Lys-CH	1.53	1.51	–0.02	PO_4^{2-} [a]	1.23	0.18	–1.05

a) ^{31}P NMR spectroscopy.

Due to the excellent agreement between experimental NMR data and MM/MD, additional calculations for the 1:1 complexes of **196** with ADP and ATP were carried out according to the procedure described for the calculation of the AMP complex (vide supra). In Figure 98 the solvent-accessible surfaces of receptor (yellow) and of the phosphate residues of the nucleotides (green) are displayed. From left to right: AMP, ADP, and ATP. It is clearly visible, that the solvent-accessible surface is getting larger within the series AMP < ADP < ATP. The strong charge-charge interactions between the positively charged groups of the receptor with the highest negatively charged α-phosphate of the nucleotide plus the hydrophobic interactions of the aromatic template of the receptor with the nucleobase tightly lock the complex geometry. The larger the phosphate hinge is now, the more it is located outside of the cavity, because the receptor is less well able to wrap around it. This way, the competitive influence of the surrounding water is getting stronger and consequently the binding strength is reduced. Hence, the explanation for the unique selectivity of **196**, preferring the lower charged monophosphate over its higher charged analogues, can be explained through the formation of a distinct binding mode between host and guest and the competitive influence of the aqueous medium on the binding strength.

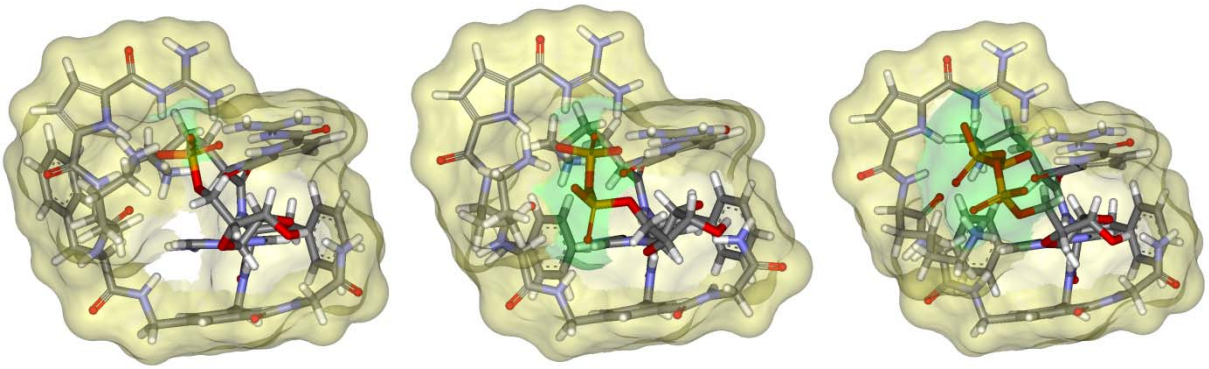


Figure 98. *Representation of the solvent-accessible surfaces of the calculated 1:1 complexes of receptor **196** (yellow) and the phosphate residue of the nucleotides (green). From left to right: AMP, ADP, and ATP. Reprinted with permission from Kuchelmeister, H. Y.; Schmuck, C. Chem. Eur. J. **2011**, 17, 5311-5318. Copyright 2011 John Wiley and Sons.*[206]

In conclusion, with the help of UV/Vis, fluorescence and NMR experiments combined with molecular modeling and molecular dynamic calculations it could be shown that the symmetric tweezer receptor **196**, with the sequence Phe-Lys (from C- to N-terminus) and a GCP headgroup in both arms, is able to form stable 1:1 complexes with nucleotides (max. $K \approx 10^5\ M^{-1}$) in buffered water at neutral pH. The host-guest relationship is based on non-covalent electrostatic and hydrophobic and/or π-π-stacking interactions. The latter one is responsible for the differentiation of the nucleotides from phosphates. The formation of a well-defined complex with specific interactions is the explanation of the unusual preference for AMP over its higher charged analogues ADP and ATP, which is yet unprecedented. All nucleotide hosts reported in literature so far form stronger complexes to the higher charged nucleotides due to unspecific charge-charge interactions. Hence, the complex formation between **196** and the phosphate-based substrates reported herewith stresses the importance to combine different types of non-covalent interactions with a defined structure in order to achieve selective molecular recognition of biologically relevant targets in aqueous solution.

4.4 TWO- AND THREE-ARMED LIGANDS FOR DNA/RNA-RECOGNITION

After having successfully utilized the symmetric two-armed receptor **196** (side-chains: Phe-Lys-GCP) for the molecular recognition of mononucleotides in buffered water at neutral pH, the next logical step was to study its ability to interact with nucleic acids. In the following, molecules which interact with DNA/RNA will be called ligand instead of receptor to do justice to the proportions of polynucleotide binder and the polynucleotide itself. Besides the ligand **196**, two more ligands will be synthesized and tested for their capability to interact with nucleic acids: Another two-armed polynucleotide (PN)-binder with inverse amino acid sequence **206** (Lys-Phe-GCP) and a three-armed ligand of the original sequence **207** (Phe-Lys-GCP).

4.4.1 Synthesis of the Two-Armed Ligand with an Inverse Amino Acid Sequence

The synthesis of **206** was carried out according to the microwave-assisted SPPS procedure on MBHA resin with the coupling agent PyBOP as described in chapter 4.3.2. Instead of the amino acid sequence Phe-Lys as in the original tweezer receptor **196**, the novel one featured the inverse sequence Lys-Phe as depicted in Figure 98. The GCP unit was again implemented as headgroup. The product was isolated as TFA salt after two purification steps via MPLC (RP18, H_2O → MeOH, 0.1 % TFA) with 22 % yield and a purity of 94 % as determined by analytical HPLC (see Figure C. 4). The trifluoroacetate counterions were then exchanged for chlorides by dissolving the TFA salt several times in dilute hydrochloric acid and consequent lyophilization to give **206** as hydrochloride salt. With the so obtained ligand the influence of the amino acid sequence on polynucleotide binding could now be studied by comparing the interaction of the two tweezer ligands **196** and **206** with nucleic acids.

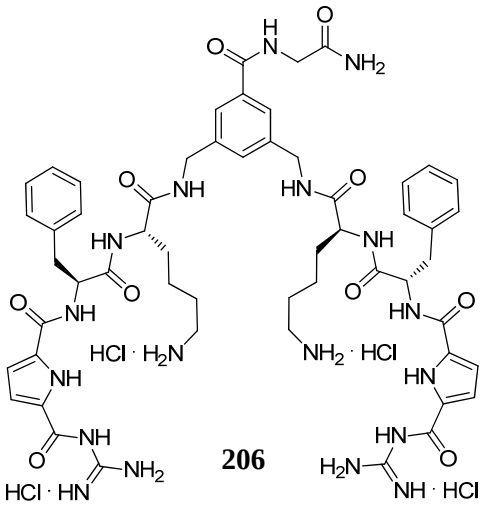


Figure 98. *The symmetric tweezer ligand **206** (Lys-Phe-GCP) was obtained by microwave-assisted SPPS.*

4.4.2 Synthesis of the Three-Armed Ligand

In order to study the influence of the valence of the nucleic acid-ligands, a three-armed analogue of the symmetric tweezer **196** with the side-chain sequence Phe-Lys-GCP was prepared. An aromatic triamino scaffold similar to the one for the divalent analogues was utilized in order to conserve the spacing and angle between the arms. However, this kind of template does not feature an additional functional group for the attachment to a solid-support. Hence, the three-armed ligand was synthesized following a convergent synthetic procedure in solution. First, the triamino template and the fully protected arms were synthesized independently and then coupled and deprotected to give the trivalent ligand **207**.

The first attempt to synthesize the template is depicted in Figure 99. The procedure was adapted from a known synthesis and started with commercial 1,3,5-tribromo benzene **208**.[210] Conversion to the tricyano derivate **209** via nucleophilic aromatic substitution according to an SN_2Ar mechanism and subsequent reduction with diisobutylaluminium hydride (DIBAL-H) at elevated temperature should afford the triamine scaffold **210**. The first step was carried out with the help of copper (I) cyanide in N-methylpyrrolidone (NMP) at high temperature. However, no product formation could be observed. Probably this was due to the lack of an electron withdrawing substituent in *ortho*- or *para*-position which would stabilize the *Meisenheimer* complex analogue and/or because bromine was used instead of the better leaving group iodine as reported in literature. For a reaction mechanism via an arine intermediate the conditions were too mild. Furthermore, this reaction would have resulted in two regioisomers. A second attempt via a modified *Heck* reaction according to *Herrmann* and *Beller* was carried out at high temperature with the help of potassium cyanide, catalytic palladium(II) acetate and 1,5-bis(diphenylphosphino)pentane (dpppe) in N,N,N′,N′-tetra-methylethylene diamine (TMEDA).[211] The reaction proceeds via in situ formation of a Pd^0 species, oxidative addition to the aryl halogenide, ligand exchange ($Br^- \rightarrow CN^-$), and reductive elimination of the aryl cyanide and regeneration of the Pd^0 catalyst.[212] The failure of this reaction might have been due to the bad solubility of the ligand[213] or due to inactivation of the palladium component caused by a too high cyanide concentration via formation of polycyanides of differing oxidation states.[214]

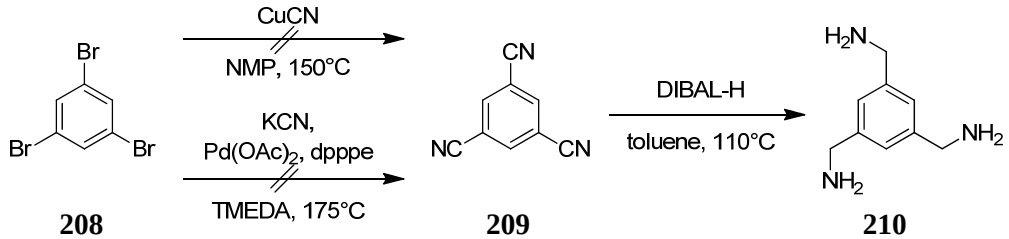


Figure 99. *The first attempt to synthesize **210** starting from 1,3,5-tribromo benzene **208** failed.*

Due to the difficulties with the approaches starting from **208** a completely different synthetic route was tested next, starting from 1,3,5-trimethyl benzene **211** as shown in Figure 100. The first step of the synthesis is a benzylic bromination to **212** with NBS and AIBN as radical starter at elevated temperature in carbon tetrachloride in analogy to the synthesis of the symmetric tweezer receptors.[215] The reaction only proceeded with moderate yield. Several differently brominated species could be isolated as byproducts. Purification was carried out

with the help of flash column chromatography and was extremely tedious due to the similar retention factors of the brominated species. Replacing CCl_4 for DCM did not improve the yield as was the case for the tweezer template. Instead, additional bromination at the ring could be observed. This was probably due to the reduced reaction temperature caused by the lower boiling point of DCM with regard to CCl_4. Furthermore, **211** is more electron rich than the benzoic acid derivative **173** utilized for the synthesis of the template for the tweezer receptors (chapter 4.2.1). Therefore, bromination at the ring should be facilitated. The next two steps, a *Gabriel* synthesis according to *Ing-Manske* with potassium phthalimide (**213**) and subsequent hydrazinolysis afforded the triamine scaffold **210** in very good yield.[216,217] No column chromatography is necessary in the last step, because the product is already highly pure after extraction from the reaction mixture.

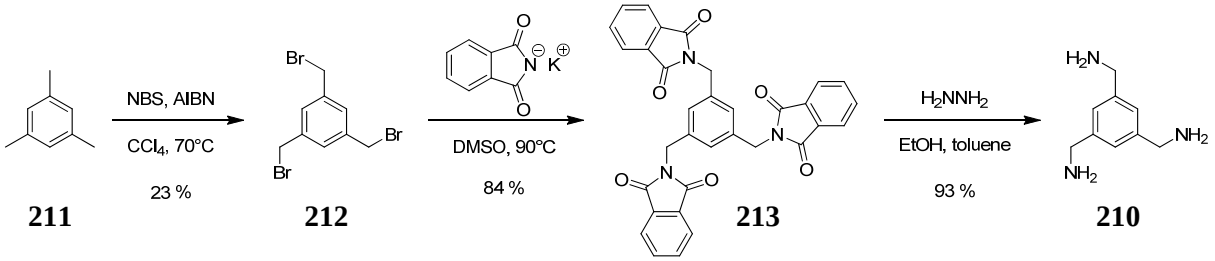


Figure 100. *The synthesis of the triamine scaffold **210** was carried out in a three step procedure.*

With the linker at hand, the next step was to synthesize the side chains. This was done with the help of standard Fmoc SPPS on SASRIN resin.[178] This solid-support is equipped with a linker that allows cleavage from the beads under very mild conditions thus leaving the Boc protecting groups of the amine functionalities intact. This is necessary to avoid unwanted coupling processes with the amino acid side chains. The synthesis was carried out manually at room temperature as depicted in Figure 101. Consequently, longer reaction times for the attachment to the resin and the coupling steps of at least 8 h were necessary. All coupling steps were conducted under argon atmosphere.

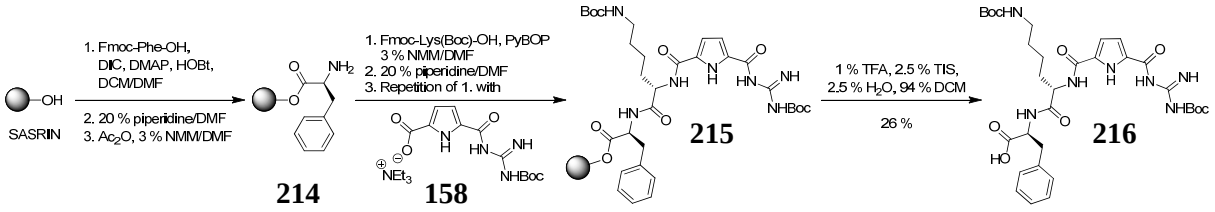
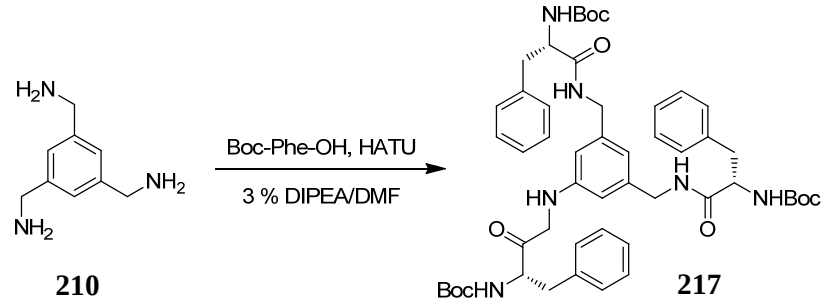


Figure 101. *Fmoc SPPS on SASRIN resin yielded the fully protected arm **216**.*

After swelling of the resin in DMF/DCM (3/7), the first amino acid, Fmoc-Phe-OH was attached to the solid support with DIC and HOBt as coupling agents with catalytic DMAP in a mixture of DMF/DCM (3/7). The first step was repeated three times and still remaining free hydroxyl groups at the resin were capped with acetic anhydride in 3 % NMM/DMF. The following Fmoc deprotection was carried out by treating the resin with 20 % piperidine/DMF twice for 20 min. Conversion to the free amine as well as the following coupling steps were monitored with the help of the Kaiser test.[181] The next amino acid Fmoc-Lys(Boc)-OH and

the oxoanion binding site Boc-GCP were coupled with PyBOP (3 eq each) in 3 % NMM/DMF. If necessary, coupling steps were repeated according to the results of the Kaiser tests. Finally, the resin-bound product **215** was cleaved from the solid-support by treating the resin with 1 % TFA in DCM containing 2.5 % TIS and 2.5 % water as nucleophilic scavengers (5 × 1 h). After purification via MPLC (RP18, H_2O → MeOH) **216** could be isolated with 28 % yield. For further improving the reaction yield in future approaches the deprotection time should be increased to five times 2 h and the cleavage mixture should immediately be neutralized with for example triethylamine to prevent partial deprotection of the Boc groups as was observed with the procedure described above.

Before linking the two building blocks **210** and **216**, a test reaction was conducted to find appropriate reaction conditions (Figure 102). In order to simulate the attachment of the fully protected arms, one equivalent of the triamine template was reacted with three equivalents Boc-Phe-OH with the coupling reagent HATU (3 eq) in 3 % DIPEA/DMF under argon atmosphere to give **217**.[218] Control via TLC indicated almost complete conversion of the starting materials. After removal of the solvent, the product could be identified via MS. It was concluded that these reaction conditions seem to be appropriate. They were thus utilized for the coupling of the scaffold **210** with the fully Boc-protected side chains **216**.



210 **217**

Figure 102. *Test reaction for the coupling of **210** with **216**. Boc-Phe-OH was utilized to simulate the fully protected side chain. The formation of the product **217** was verified by MS.*

As shown in Figure 103 the reaction of **210** with **216** was conducted accordingly. Since the triamine is not stable for long time, it was synthesized from the phthalimide precursor **213** in the amount necessary for the following reaction step and directly reacted. The resulting fully protected three-armed compound **218** was isolated via MPLC (NP, CyHex → EA), its formation verified with the help of MS, and the still crude product directly employed for the next step. Deprotection of all Boc protecting groups was achieved by treating the compound with 50 % TFA/DCM. The final product was purified via MPLC (RP18, H_2O → MeOH, 0.1 % TFA) and after exchange of the trifluoroacetate anion for chlorides the trivalent ligand **207** was obtained with 26 % yield over two steps. When testing the purity via analytical HPLC three distinct peaks with identical UV/Vis spectra could be observed (see Figure 103). An MS analysis thereof revealed that all three peaks give the same m/z signals, corresponding to the desired product. Therefore, the separate signals may either stem from different diastereomers or from varying protonation states of the product (max. 6 positive charges). The potential formation of diastereomers would most probably have taken place during the coupling step of the fully protected **216** with the template as a result of racemization of the

phenylalanine stereogenic center. This however, would have afforded four diastereomers (L-L-L, D-L-L, D-D-L, and D-D-D). Additional racemization of the lysine stereocenter during the SPPS would have afforded even more different diastereomers. The HPLC on the other hand only showed three peaks. Furthermore, as could be shown in the meantime in the work conducted by *S. Junghänel* during her master thesis,[219] the number and retention time of the HPLC peaks are strongly depending on the TFA concentration during the HPLC run. Therefore, the three peaks were attributed to different protonation states and consequently, the purity was determined to 95 %. To further clarify in future work whether these three fractions are indeed different protonation states, the fractions should be isolated via preparative HPLC, treated with base, and be subjected to analysis via analytical HPLC.

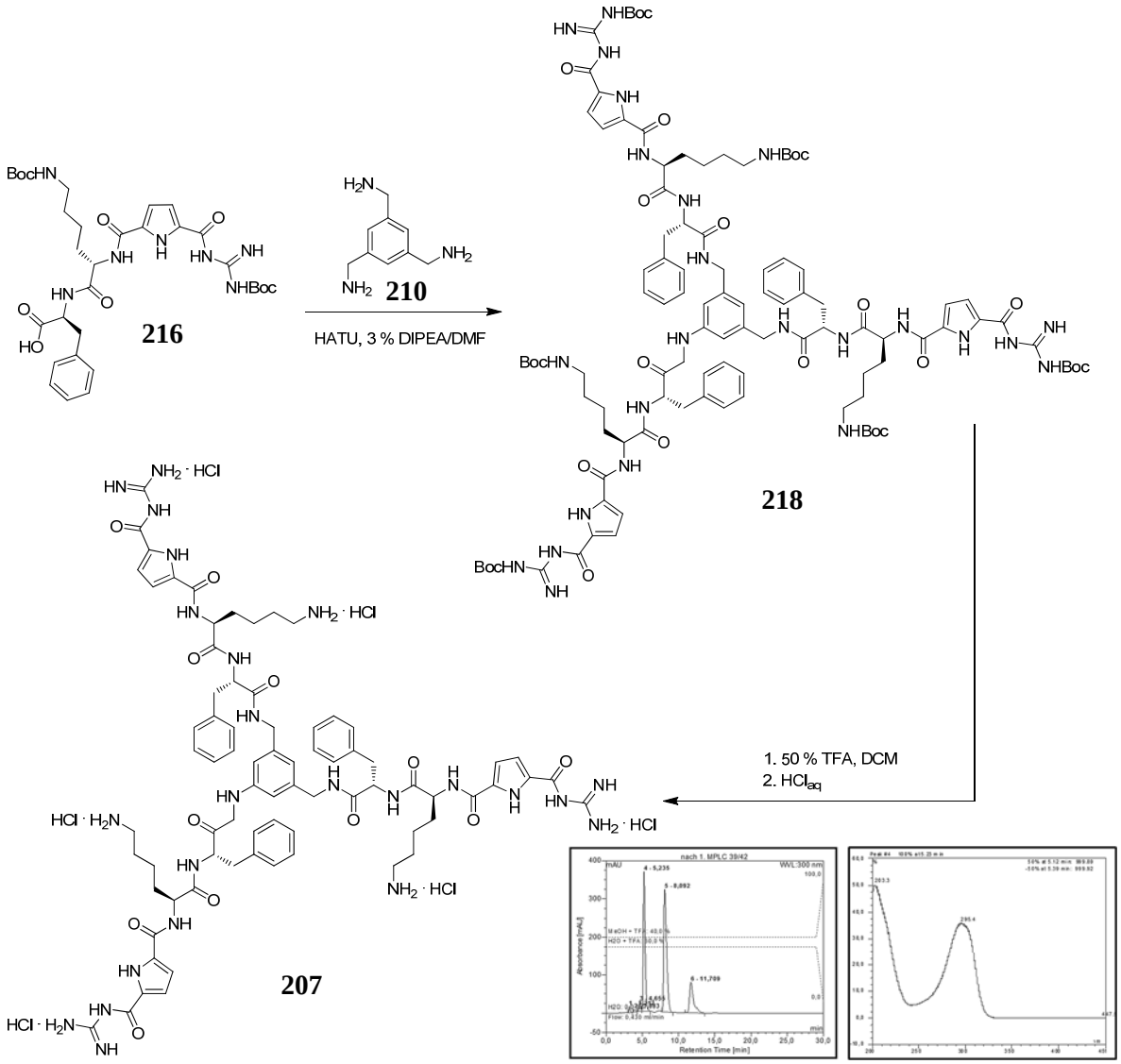


Figure 103. *Coupling of the triamine scaffold **210** and the fully protected side chains **216** and subsequent deprotection of the Boc groups gave **207**. The purity was determined via analytical HPLC.*

Thus, with the already available tweezer ligand **196** (Phe-Lys-GCP), the completion of the syntheses of the trivalent analogue **207** (Phe-Lys-GCP) and the novel tweezer **206** with inverse amino acid sequence (Lys-Phe-GCP), three compounds were now available for nucleic acid binding studies. In the following chapter their capability to bind to different polynucleotides will be elucidated.

4.4.3 Nucleic Acid Binding Studies

Synthetic and Natural Polynucleotides for Binding Studies

The three polynucleotide ligands **196**, **206**, and **207** were tested for their capability to interact with various synthetic and natural double-stranded nucleic acids consisting of both RNA and DNA. All ds-polynucleotides are commercially available and were used as obtained (except for ctDNA) from the commercial supplier (*Sigma-Aldrich*). Stock solutions were prepared in cacodylate buffer (0.05 M) at neutral pH and concentrations between 10^{-2} and 10^{-3} M. ctDNA was ultrasonicated (8 × 4 sec) and filtered prior to use to break the long double stranded nucleic acid into smaller fragments. The exact concentrations of all nucleic acids described within this work are reported with respect to the number of the polynucleotide´s phosphate groups and were determined by means of UV measurements (three point titrations). Table 7 lists the literature-reported molar extinction coefficients ε at the maxima of the nucleic acids´ absorption bands λ_{max}.[220]

Table 7. *Molar extinction coefficients ε at the maximum of UV absorption (λ_{max}) of various nucleic acids.*

Nucleic acid	pdA × pdT	$p(dAdT)_2$	pdG × pdC	$p(dGdC)_2$	pA × pU	pG × pC	ctDNA
ε [L/mol · cm]	6000	6600	7400	8400	6000	7400	6600
λ_{max} [nm]	260	262	253	254	260	253	260

The synthetic ds-PNs were either homopolymers (pdA × pdT, pA × pU, pdG × pdC, and pG × pC) or alternating heteropolymers ($p(dAdT)_2$ and $p(dGdC)_2$) of approximately 8,000-9,000 base pairs length. The naturally occurring ctDNA features ca. 20,000 bps (untreated) according to the supplier. The ultrasound treatment of the ctDNA breaks it into fragments of varying lengths. Therefore, the exact compliance with the preparation procedure has to be assured for each sample for the sake of comparability. All measurements conducted with ctDNA within this work were carried out with the same stock solution. The molecular weight of the nucleic acids can be determined by assuming an average molecular weight of 650 Da per base pair.[77] Thus the M_W of synthetic polynucleotides ranges from 5,200 to 5,850 kDa and that of untreated ctDNA is ca. 13,000 kDa.

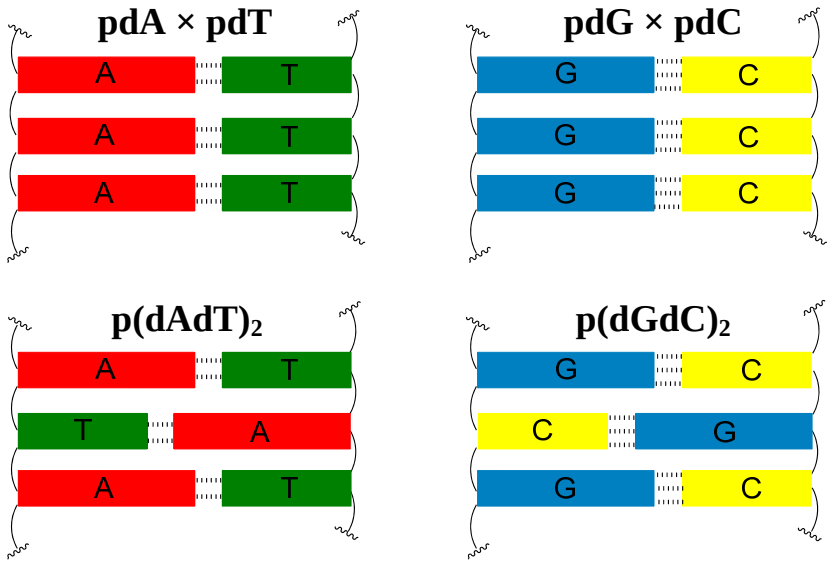


Figure 104. *Schematic representation of synthetic alternating (bottom) and homopolymeric (top) nucleic acids.*

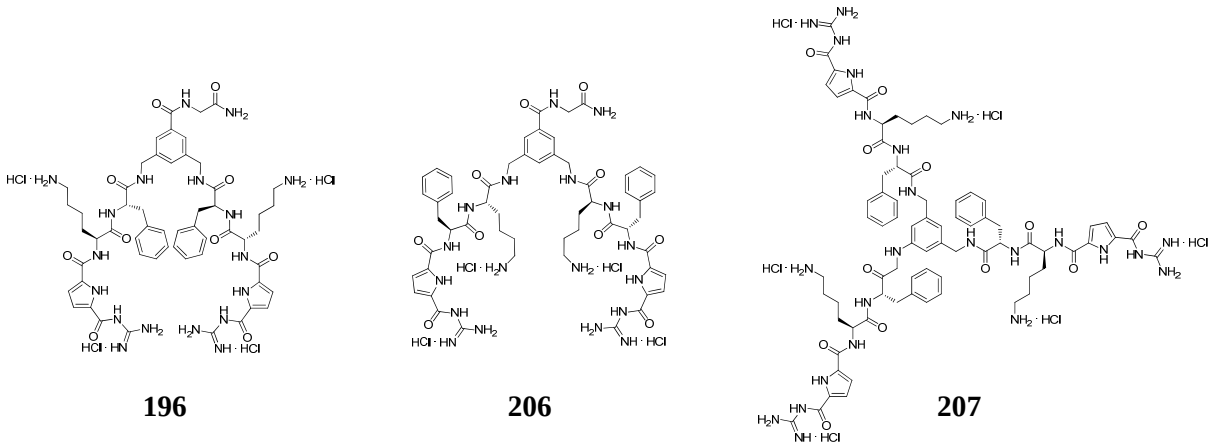
Figure 104 schematically depicts some of the polynucleotides. Depending on their sequence they feature different secondary structures, which are summarized in Table 8.[75] $p(dAdT)_2$ for instance is a classical A-tract. It forms a B-helix with a well-defined, narrow groove, a deep, well-accessible minor groove, and a rather broad and shallow major groove. pdA × pdT adopts a B*-helix with larger propeller twist and forms the deepest and narrowest minor groove of all nucleic acids utilized within this work. The major groove is broad and shallow. Both AT sequences feature the electronegative thymine O2 and adenine N2 in the minor groove (see chapter 2.3.1 for numeration of atoms). Cations are normally bound inside or on top of the minor groove. The heteropolymer $p(dGdC)_2$ also adopts a B-helix, but its minor groove features the electropositive guanine NH_2, and is therefore less well suited for minor groove binding. pdG × pdC on the other hand adopts an A-type helix with a broad minor groove which also features the guanine NH_2 and is therefore not very well suited for the accommodation of ligands at all. Its major groove is unusually narrow and could thus be an interesting target for small molecule binding. When cations are integrated into the major groove a B- to A-form transition can often be observed. ctDNA features approximately 40 % GC and 60 % AT base pairs and adopts a B-helix similar to that of $p(dAdT)_2$ with an accessible but not that well-defined minor groove. The major groove is rather broad. The two RNA polymers pA × pU and pG × pC form typical A-helices with broad minor and narrow major groove. Therefore, these polynucleotides offer a broad spectrum of different secondary structures for the testing of the nucleic acid binding properties of the three artificial ligands.

Table 8. *Characteristic features of the studied polynucleotides´ helix type and their minor and major groove.*

PN	Helix	Minor Groove	Major Groove
ctDNA	B	rather narrow	rather broad
$p(dAdT)_2$	B	narrow, deep, electronegative thymine O2, adenine N2	broad, shallow
pdA × pdT	B*	very deep, very narrow, electronegative thymine O2, adenine N2	broad, shallow
$p(dGdC)_2$	B	shallow, electropositive guanine NH_2	rather narrow
pdG × pdC	A	broad, shallow, electropositive guanine NH_2	narrow
pA × pT	A	broad	narrow
pG × pC	A	broad	narrow

Physicochemical Characterization

Before starting with the binding studies, the physicochemical properties of **196**, **206**, and **207** (see Figure 105) were determined and the results are summarized inTable 9. All three compounds are well soluble in water at 1 mM – accordingly, stock solutions in sodium cacodylate buffered water (0.01 M) were prepared at this concentration. The pK_A values were determined, again as described in chapter 4.2.3 by means of the half-equivalent method. For the two-armed ligands average values of the guanidinio groups (pK_A^{I}) and the lysine ammonium groups (pK_A^{II}) were obtained by adding 1 and 2 additional equivalents of sodium hydroxide and subsequent pH measurement after each titration. The respective average pK_A values of the three-armed ligand were obtained by adding 1.5 and then another 3 equivalents. The molar extinction coefficients ε of the absorption maxima at 300 nm were determined with the help of four point titrations in cacodylate buffer (0.01 M) at pH 7. All systems follow the *Lambert-Beer* law in the concentration range used for the experiments.



196 **206** **207**

Figure 105. *Two-armed **196** and **206**, and trivalent **207** were tested for their binding affinity to nucleic acids.*

Table 9. *Physicochemical characterization of **196**, **206**, and **207** in sodium cacodylate buffered water (0.01 M).*

Compound	pK_A^I	pK_A^{II}	Water solubility at 1 mM	ε [L/mol · cm]
196	7.00	9.90	+	55,200
206	7.20	9.94	+	53,450
207	6.50	9.93	+	80,260

Thermal Melting Experiments

In thermal melting experiments defined amounts of ligand are heated to approx. 100° C together with the nucleic acid and the melting point (T_m) is determined by generating the first derivative of the resulting melting curve as explained in chapter 2.2.3.[80] Before doing so, it has to be assured that the compounds in consideration are stable at high temperature and over time. Furthermore, it has to be tested whether the ligands´ UV/Vis absorption changes with temperature – such an effect would have to be taken into account when interpreting the thermal melting plots. Therefore, solutions of the compounds **196**, **206**, and **207** were prepared in cacodylate buffer at neutral pH (0.01 M) at a concentration of 2×10^{-5} M and the resulting UV spectra were recorded between 220 and 350 nm at 25° C and higher temperatures (40, 60, 70, 80, and 90° C). The illustrations in Figure 106 demonstrate that the changes of the absorption as a result of the rising temperature were negligible. Furthermore, after cooling down to 25° C the spectra were identical to the one recorded prior to heating within the error of the instrument. An additional spectrum recorded after 24 h did not feature any changes either. Therefore it was concluded that the ligands are stable over time and temperature and are thus suitable for the planned thermal melting experiments with polynucleotides.

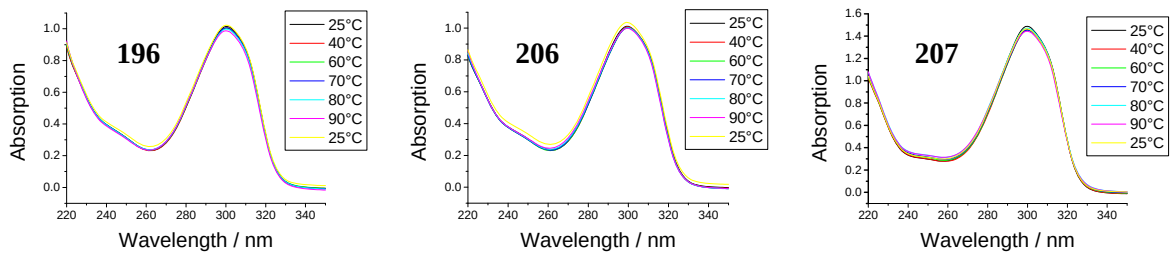


Figure 106. *Temperature-dependent UV/Vis spectra of the three nucleic acid ligands **196**, **206**, and **207**.*

The next step was to optimize the conditions for the melting experiments by varying the ionic strength and the pH of the solution. Melting experiments were carried out at an ionic strength of 0.05 M (buffer) at pH 7 and at a buffer concentration of 0.01 M at pH 5 and 7. ctDNA solutions were prepared at a concentration of approximately 1.5-2.0 × 10^{-5} M in microcuvettes which were equipped with stoppers (1 mL) resulting in absorption values around 0.15. By measuring the absorption at the polynucleotides´ maximum at 260 nm the exact concentration of the DNA was calculated with the help of the molar extinction coefficient (ε = 6600). Then 0.3 eq (0.05 M buffer) or 0.2 eq of **196** (0.01 M buffer) with respect to the DNA´s phosphate groups (r) were added. Furthermore, samples containing only ctDNA were prepared. All mixtures were then slowly heated by 0.5° C per minute to 99° C while measuring the absorption at 260 nm after 0.5 min incubation time at each 0.5° temperature step. The resulting thermal melting curves are shown in Figure 107. All measurements were carried out in duplicate.

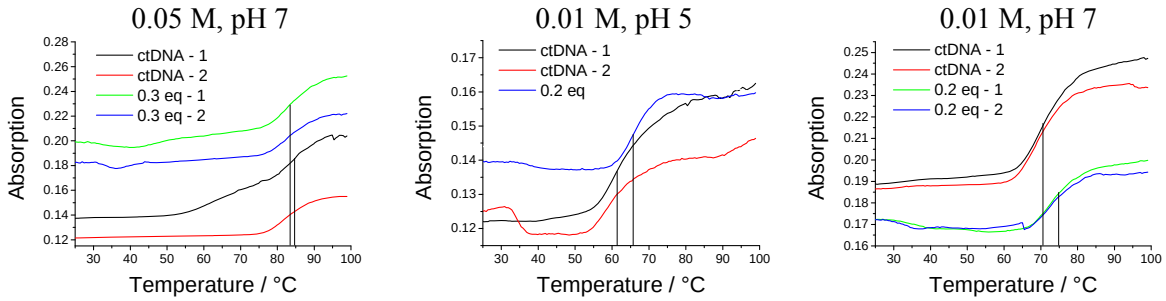


Figure 107. *Thermal melting curves of ctDNA with **196** with 0.05 molar cacodylate buffer at pH 7 (left), and with 0.01 molar buffer at pH 5 (middle) and pH 7 (right) as determined by the UV absorption at 260 nm.*

At high ionic strength (0.05 M) the melting point of ctDNA was determined to an average of 84.9° C. After addition of 0.3 eq **196** the melting point was reduced by -1.6° to 83.3° C. When taking into account the general error of the method (ΔT_m = ± 1° C), this result signifies that the two-armed ligand is not able to stabilize the helix at this high buffer concentration – if at all, there is a slight destabilization. Obviously, the competitive influence of the surrounding medium is too strong for effective binding of **196** to the B-DNA helix. At lower buffer concentration (0.01 M) the influence of the solvent should be reduced. Furthermore, when carrying out the experiments at pH 5, all guanidinio groups of the ligand should be protonated and thus the complex formation should be favored due to increased charge-charge interactions. And indeed, at pH 5 the ctDNA helix is thermally stabilized by 4.6° from 61.2° C (ctDNA) to 65.8° C (complex) already upon addition of only 0.2 eq of the ligand (the bulge around 30° C in the red line stems from bubbles in the cuvette, which are formed during the heating process). **196** is now able to bind to the ds-helix and stabilize it to a significant degree. Surprisingly, at the same ionic strength but at higher pH of 7 a similarly high stabilization of the ctDNA could be witnessed. The helix was stabilized by 4.4° from T_M = 70.6° C to 75° C. Consequently, the impact of full protonation of the guanidinio groups is only minor, while the influence of the ionic strength of the medium controls the efficiency of the stabilization of the double strand. Therefore, the more detailed additional measurements, which will be presented in the following, were conducted at a sodium cacodylate buffer concentration of 0.01 M at neutral pH. Denaturation experiments were carried out with different amounts of the three ligands (r = 0.1, 0.2, 0.3 eq) with the DNAs pdA × pdT,

p(dAdT)$_2$, and ctDNA, as well as with the RNA pA × pU. The GC-based polynucleotides feature melting points, which are too high to be measured in water. The results of the denaturation experiments are listed in Table 10. All values are averaged from at least two measurements. A representative example of a thermal melting experiment with ctDNA and increasing amounts of **196** is depicted in Figure 108 (left), all other combinations can be found in appendix C.3 (Figure C. 29-Figure C. 33 Figure C. 39). Additionally, Figure 108 (right) also depicts the 1st derivative of a ctDNA melting curve recorded with 0.2 eq of **196**.

Table 10. *Stabilization of the polynucleotides (PN) upon addition of varying equivalents (r) of the three ligands (L)* ***196***, ***206***, *and* ***207*** *as determined by thermal denaturation in 0.01 M cacodylate buffer at pH 7.*

Ligand	r = [L] / [PN]	Stabilization [°] ctDNA[a]	p(dAdT)$_2$[b]	pdA × pdT[c]	pA × pU[d]
196	0.1 eq	3.6	4.9	2.4	n.d.
	0.2 eq	4.4	6.0	3.7	0.9
	0.3 eq	4.7	6.0	5.0	3.6
206	0.1 eq	2.3	4.5	2.3	n.d.
	0.2 eq	*	6.9	3.3	0
	0.3 eq	*	5.4	8.4	(-2.6)*
207	0.1 eq	*	5.8	3.8	5.7

a) T_m = 70.6° C; b) T_m = 42.7° C; c) T_m = 51.5° C; d) T_m = 42.6° C; *) precipitation.

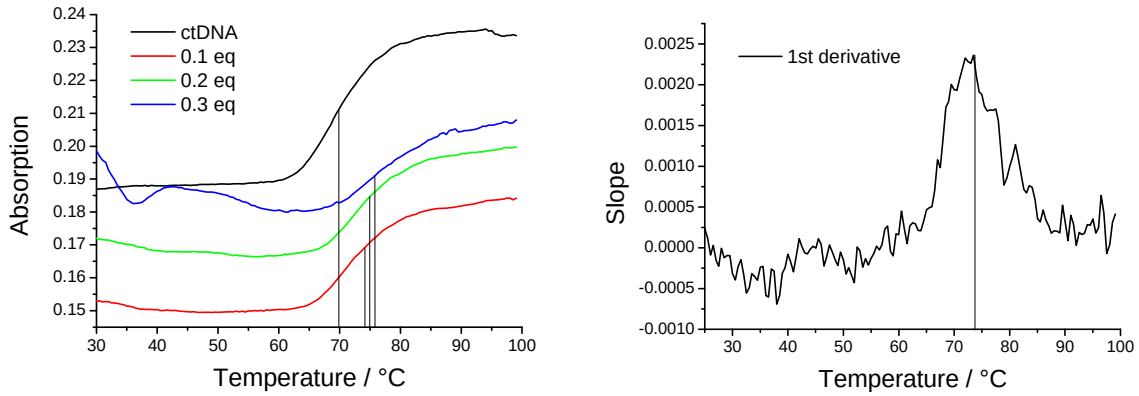


Figure 108. *Melting curves of ctDNA were recorded at 260 nm with varying amounts of* ***196*** *in cacodylate buffer (0.01 M) at neutral pH (left) and the 1st derivative of a melting experiment with ctDNA and 0.2 eq of* ***196****.*

In general, all three ligands are able to increase the thermal stability of the nucleic acids. That signifies that they prefer to bind to the double-stranded helical nucleic acids relative to their corresponding unfolded forms. The correlation between binding constant and increase of T_m is quite complex, because also the number of binding sites, potential cooperativity and the affinity for the unfolded polynucleotide have to be taken into account.[80] However, at this point it may already be stated that the experiments proof that all ligands do interact with the tested nucleic acids. A general trend which was observed was that with increasing amount of PN-binder the stabilization increases until saturation is achieved, e.g. for the interaction between the two armed **196** and ctDNA (ΔT_m = +4.4°) and the similarly shaped B-helical p(dAdT)$_2$ (ΔT_m = +6.0°) saturation occurs at 0.2 eq. For the inverse sequenced **206** this is also the case for p(dAdT)$_2$ (ΔT_m = +6.9°) – for ctDNA precipitation was observed at ratios higher

than $r = 0.1$ at increased temperature. The T_m of pdA × pdT on the other hand, which adopts a B-type helix with a narrower minor groove, is still increased by both ligands when going from 0.2 to 0.3 eq (**196**: 3.7 → 5.0°; **206**: 3.3 → 8.4°). Higher amounts of the divalent binders (0.5 eq) led to precipitation at higher temperature. Thus, the thermal stabilization properties of the two divalent ligands are rather similar for the different DNAs. For the RNA pA × pU on the other hand, only **196** is able to stabilize the A-helix, but not **206**. Reasonable stabilization of RNA by **196** though is only achieved at $r = 0.3$ ($\Delta T_m = +3.6°$). However, for both molecules a broadening of the RNA melting transition can be observed as depicted in Figure 109 (left). This phenomenon could previously be linked to ligand binding.[221] Thus, it may be concluded that both two-armed ligands do interact with RNA, but only **196** thermally stabilizes it.

For the three-armed **207**, the addition of more than 0.1 eq led to precipitation at elevated temperature. At a ratio of $r = 0.1$ eq however, **207** affords the highest stabilizations of the three compounds for both DNA and RNA, e.g. it increases the melting temperature of $p(dAdT)_2$ by 5.8° or that of pA × pU by 5.7°. For comparison, **196** and **206** stabilize $p(dAdT)_2$ for +4.9 and +4.5° at this ratio and pA × pU not at all. It has been shown, that larger ligands, which interact via the identical type of interactions with a nucleic acid, are less efficient in stabilizing a helix than corresponding smaller ligands.[222] Therefore, one may conclude that the larger trivalent **207** binds with higher affinity to the nucleic acids than its dipodal analogues **196** and **206**. Another interesting feature was observed for the interaction of **207** with pdA × pdT, which features a second transition (+20.2°) after the first one (+3.8°) as shown in Figure 109 (right). This unusual, possibly biphasic melting curve was only noted for pdA × pdT. Because there are several possible reasons for such a melting curve other methods have to be taken into account before concluding on the origin of this effect.[223]

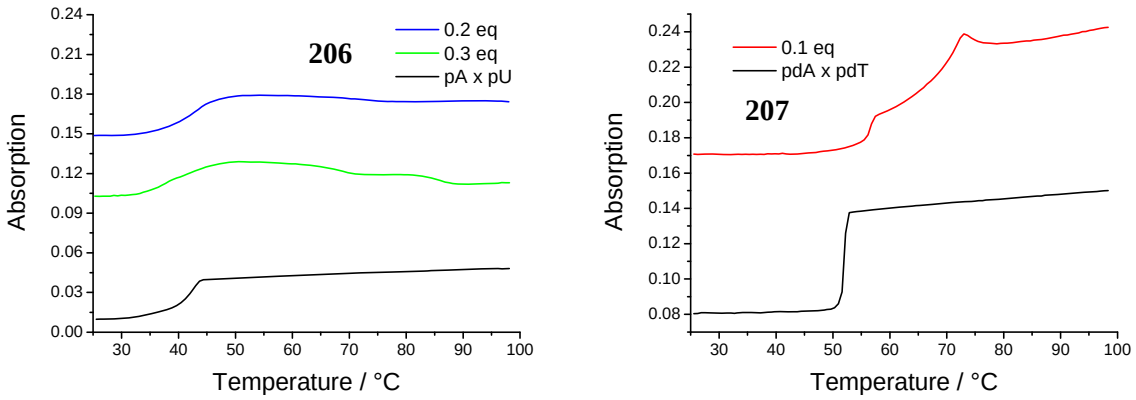


Figure 109. *Interaction of **206** with pA × pU does not lead to an increase of T_m, but to a broadening of the transition at r = 0.2. At higher ratio precipitation is observed at higher temperature (left). The interaction of **207** with pdA × pdT leads to a biphasic transition with ΔT_m values of +3.8° and +20.2° (right).*

In conclusion, the thermal melting experiments demonstrated that the three artificial ligands **196**, **206**, and **207** are able to interact with both DNA and RNA and are further able to stabilize the helical form by up to 8.4° (pdA × pdT + 0.3 eq **206**). Trivalent **207** seems to be especially apt to interact with nucleic acids, presumably with higher binding constants than its

two-armed analogues. To investigate in more detail these non-covalent interactions, experiments for the determination of the binding affinity were conducted in the following.

Spectroscopic Measurements

Similar to molecular recognition events between small host and guest systems (cf. chapter 4.3.3), the binding affinity of small artificial ligands to macromolecules can be determined by means of titrations. Due to the large number of potential binding sites in nucleic acids, specially developed equations have to be utilized for the characterization of such binding events. Already in 1949 *Scatchard* proposed an equation for the fitting of protein-ligand titrations, which can also be utilized for polynucleotide-ligand binding events.[224] This equation can be solved with the help of a statistical program (e.g. *Statistica*)[225] by generating a non-linear regression for the spectroscopic data following the formalism of *McGhee* and *von Hippel*,[226] which will give the binding constant and the stoichiometry as output.

Thus, the first attempt to qualitatively determine the binding constants of the nucleic acid-binders was conducted with the help of UV/Vis titrations. To a buffered aqueous solution (cacodylate buffer 0.01 M) of either ligand **196**, **206**, or **207** in a microcuvette equipped with a stopper (800 µL, c = 20 µM) a stock solution of ctDNA (0.75 mM) was titrated in aliquots (5 × 5, 5 × 10, 3 × 25, 3 × 50, 4 × 100 µL) to a maximum of 33 eq of nucleic acid at 25° C and neutral pH. After each addition and one minute incubation time the resulting UV/Vis spectrum was recorded from 250 to 350 nm. Figure 110 depicts excerpts of the absorption at 300 nm, which are corrected for the dilution. At this wavelength DNA only absorbs negligibly and the absorption can thus be attributed to the pyrrole unit. Changes are almost proportional to the concentration of the DNA until a clear breakpoint at approximately 0.1 mM ctDNA concentration, which corresponds to ca. 5 equivalents of ctDNA or 0.2 equivalents of ligand with regard to ctDNA could be observed. This curvature did not allow accurate non-linear analysis by means of the *Scatchard* equation, which was developed for hyperbolic function fitting. Hence, at the micromolar concentrations used for the experiments the binding constants can only be estimated to a lower limit of $K \geq 10^6$ M^{-1}.

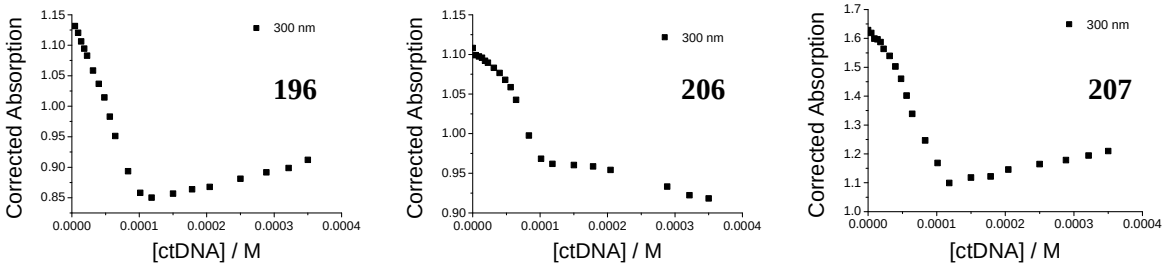


Figure 110. *The excerpts at 300 nm of the UV/Vis absorption during titrations of aliquots of ctDNA (c = 0.75 mM) to* ***196***, ***206***, *and* ***207*** *(20 µM) in cacodylate buffer (0.01 M) were too steep for accurate fitting.*

The more sensitive fluorescence spectroscopy allows carrying out such titrations at lower concentration and therefore the determination of higher binding constants. Unfortunately, attempts to accomplish such binding studies with ctDNA and the three synthetic ligands failed due to overlapping excitation wavelengths of the pyrrole (270-330 nm) and the nucleobases (230-290 nm). Therefore ethidium bromide displacement assays were conducted. The

fluorescence of the intercalator EB (see chapter 2.3.5) increases by a factor of 10 to 100 upon intercalation between two base pairs of nucleic acids. The association process to generic DNA takes place with a binding constant of 3×10^6 M^{-1} at the experimental conditions used here (ionic strength = 0.01 M). Addition of aliquots of DNA-binding molecules leads to a displacement of ethidium bromide and consequently to reduction of the EB emission band around 600 nm. The amount of ligand which is necessary to displace EB (IC_{50}) corresponds approximately to the affinity of the titrant.[227] However, due to the approximative character of the affinities derived from the IC_{50} values, only differences of ca. one order of magnitude or greater may be regarded as significant. Furthermore, IC_{50} values are only comparable for the same polynucleotide due to differing binding constants of both EB and the artificial ligands to different nucleic acid sequences. Figure 111 depicts a typical displacement experiment of an EB/ctDNA complex with **207** and the corresponding excerpt of the fluorescence emission at 600 nm, which is corrected for EB´s own emission. All other combinations can be found in appendix C.4 (Figure C. 34-Figure C. 47). First, a solution of ethidium bromide (0.75 µM, 1 eq) in aqueous cacodylate buffer (0.01 M) was prepared in a microcuvette equipped with a stopper (900 µL) and the fluorescence spectrum resulting upon irradiation at 520 nm was recorded from 560 to 650 nm. Then, four equivalents of nucleic acid (3 µM) were added and after 5 min incubation time the fluorescence signal was measured again. Further, solutions of either **196**, **206**, or **207** (50 µM) were titrated in aliquots of increasing volume (e.g. 5 × 4, 4 × 8, 16, 32 µL) and the emission was measured after each addition after 1 min incubation time. EB´s own emission was subtracted from the spectrum, the excerpt at 600 nm was plotted against the ratio EB/ligand, and the resulting data were then fitted with a first order exponential decay function. The amount of ligand which was necessary to result in 50 % of the original fluorescence emission, i.e. the point where half of the originally bound ethidium bromide is displaced, was determined (IC_{50}). The resulting listed in Table 11 are average values from at least two measurements. All data were corrected for dilution and are average values obtained from at least two individual measurements. Finally, titrations with EB and the three ligands without polynucleotide were carried out. It could thus be shown that there is no interaction between the intercalator and the synthetic nucleic acid binders. Displacement of EB is only due to the affinity of **196**, **206**, or **207** towards the nucleic acid and not due to secondary interactions with EB.

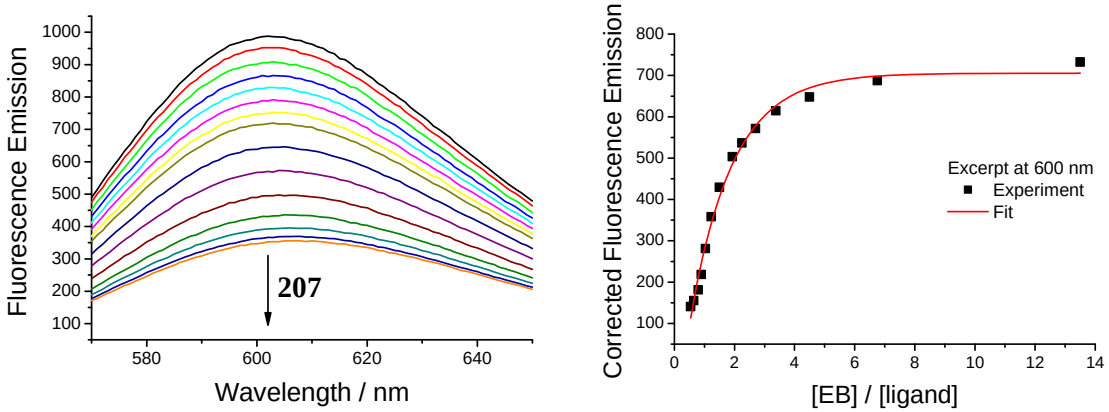


Figure 111. *EB displacement experiment with ctDNA and **207** (left) and excerpt of the corrected fluorescence emission at 600 nm plotted against the ratio of EB/**207** with the corresponding fit (right).*

The IC_{50} values for pdG × pdC and pG × pC could not be determined (n.d.) this way because mixing of EB with the GC-RNA did not produce an increase of emission signal, and respectively because the interaction of EB with the homopolymeric GC-DNA did not result in reproducible fluorescence spectra. Obviously mixed binding modes are present at this low ionic strength of the medium hampering accurate determination of the IC_{50} values for these nucleic acids.

From the results of the other experiments it is clearly visible that all three compounds are able to displace EB from all nucleic acids under investigation, thus underlining their ability to efficiently bind to both DNA and RNA. The obtained values are another clue that the binding constants range at least around $K \approx 10^6\ M^{-1}$. The three armed receptor **207** is more efficient in displacing EB than its two-armed analogues **196** and **206** by up to one order of magnitude, which is in agreement with the thermal melting experiments. The IC_{50} values of the divalent ligands are identical within the error of the method – if anything, the IC_{50} values for **207** are slightly lower than for **196** in all cases, which would translate into slightly higher binding constants. To gain a more detailed insight into the binding events and to determine the exact binding constants, ITC measurements were carried out next.

Table 11. *IC_{50} values and standard deviations as obtained from EB displacement assays.*

Polynucleotide	**196**	**206**	**207**
ctDNA	4.36 (± 0.31)	3.17 (± 0.30)	0.79 (± 0.11)
$p(dAdT)_2$	5.77 (± 0.26)	5.34 (± 0.23)	0.46 (± 0.08)
pdA × pdT	1.25 (± 0.05)	1.01 (± 0.20)	0.35 (± 0.07)
$p(dGdC)_2$	5.90 (± 0.35)	4.96 (± 0.39)	0.72 (± 0.36)
pdG × pdC	n.d.	n.d.	n.d.
pA × pU	3.03 (± 0.01)	2.96 (± 0.28)	0.69 (± 0.14)
pG × pC	n.d.	n.d.	n.d.

Isothermal Titration Calorimetry

Isothermal calorimetry allows the direct determination of the reaction enthalpy ΔH, the stoichiometry n, and the binding constant K of a given binding event with a single experiment.[82] With these values at hand the reaction entropy ΔS can also be calculated. For successful data processing it is important to collect experimental data points with sufficient but not too steep curvature for accurate fitting. The correct experimental conditions can be estimated with the help of the C value, which is the product of binding constant K, stoichiometry n and the concentration c of the nucleic acid in the cell. C values between 1 and 1000 lead to sigmoidal data, which can then be fitted with the help of the *Origin* software add-on provided by the manufacturer of the instrument (*MicroCal*). For the titration experiments with **196**, **206**, and **207** the concentration of polynucleotides was determined by setting the C value to 200. The binding constant was approximated with $K \approx 10^6\ M^{-1}$ as derived from the UV/Vis titration and EB displacement data. The stoichiometry was estimated to $n \approx 0.2$ as judged by the denaturation and UV/Vis experiments. Thus, the concentration of the nucleic acid in the cell was calculated to 0.1 mM ($= 200 : 0.2 : 10^6$). All measurements were carried out in cacodylate buffered water (0.01 M) at pH 7. The reference cell was filled with water. All solutions were degassed prior to usage by ultrasonication for at least one hour and subsequent vacuum treatment. The experiments were carried out at 25° C by titrating

aliquots of the ligand (25-35 × 5 µL) of a concentration of 0.3-0.6 mM to the polynucleotide (1.41 ml). This way 0.4 to 0.6 equivalents of the ligands had been added at the end of the titration, thus generating a data set with sufficient data points, sigmoidal curvature and clear starting and endpoints of the binding event. The delay between injections was set to 300 sec to allow for return of the signal to the baseline prior to the next injection.

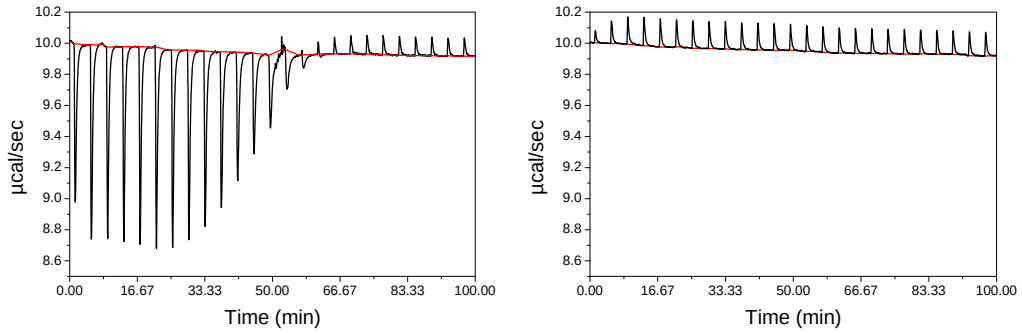


Figure 112. *Left: Raw ITC data for the titration of* ***206*** *(0.6 mM) in aliquots of 5µL to p(dGdC)$_2$ (0.1 mM). Right: Titration of* ***206*** *to pure buffer (aqueous sodium cacodylate, 0.01 M, pH 7) under identical conditions.*

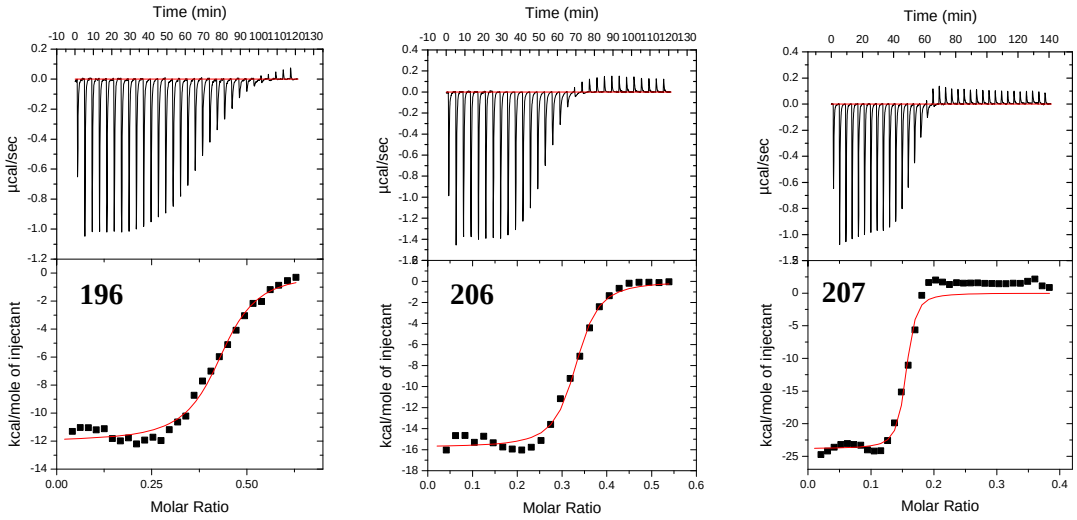


Figure 113. *Top: ITC experiments were carried out in sodium cacodylate buffer (0.01 M) at neutral pH by titrating aliquots of 5 µL of* ***196***, ***206*** *(c = 0.6 mM), or* ***207*** *(c = 0.3 mM) to ctDNA (c = 0.1 mM). Bottom: the titrations were corrected for dilution and the first data point, and then fitted with a one-site model.*

A representative experiment is depicted in Figure 112. It shows the raw data of the titration as well as a dilution experiment of ligand into buffer. Titration of the ligands into buffer led to substantial endothermal heat signals, which were thence subtracted from the titration data. On the other hand when buffer was titrated into the nucleic acid only negligible heat changes were observed. The so obtained data for ctDNA was fitted according to a one-site model for all three ligands as shown in Figure 113. All other titrations can be found in appendix C.5 (Figure C. 56-Figure C. 61). The results are listed in Table 12. Alternative curve fittings with two-site models were possible for all combinations and even necessary for fitting the titrations of p(dAdT)$_2$ with the divalent ligands **196** and **206** (see Figure 114). However, the thermodynamic data obtained for the second process with these fits was very often physically unrealistic. This second process takes place at very low ligand ratios (< 0.1 eq) and thus the concentration range and the amount of added ligand during these titrations were probably not

appropriate to resolve the second process in the presence of the dominating primary binding mode. Therefore only the data for the main binding mode can be considered to be accurate.

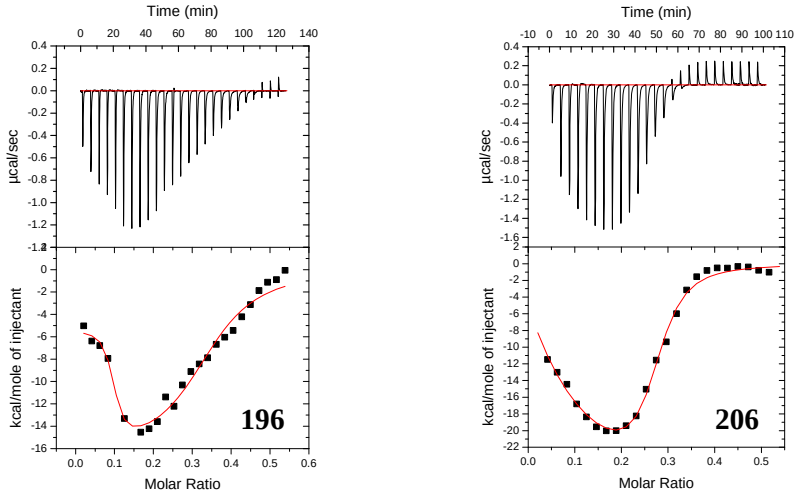


Figure 114. *Top: ITC experiments were carried out in sodium cacodylate buffer (0.01 M) at neutral pH by titrating aliquots of 5 µL of* ***196*** *or* ***206*** *(c = 0.6 mM) to p(dAdT)$_2$ (c = 0.1 mM). Bottom: the titrations were corrected for dilution and fitted with a two-site model. The titration with* ***196*** *is corrected for the first data point.*

Table 12. *Results from the ITC experiments with in aqueous cacodylate buffer (0.01 M) at neutral pH and 25° C. All experimental data was fitted according to a one-site model.*

PN	L	K [10^6 M^{-1}]	log$K$	n	ΔH [kcal/mol]	$T\Delta S$[a] [kcal/mol]	ΔG [kcal/mol]
ctDNA	**196**	1.70	6.23	0.43	-12.0	-3.5	-8.5
	206	4.89	6.69	0.32	-15.8	-6.6	-9.1
	207	27.50	7.44	0.15	-23.8	-13.7	-10.1
p(dAdT)$_2$	**196**[b]	(59.80) 0.48	(7.78) 5.68	(0.09) 0.25	(-5.4) -16.5	(5.2) -8.8	(-10.6) -7.7
	206[b]	(5.37) 2.73	(6.73) 6.79	(0.01) 0.26	(279.6) -32.5	(288.7) -23.7	(-9.1) -8.8
	207	32.70	7.51	0.20	-20.9	-10.7	-10.3
pdA × pdT	**196**	1.32	6.12	0.24	-13.3	-4.9	-8.3
	206	3.86	6.59	0.17	-17.5	-8.5	-9.0
	207	41.0	7.61	0.06	-24.6	-14.3	-10.4
p(dGdC)$_2$	**196**	4.02	6.60	0.36	-10.0	-1.0	-9.0
	206	6.02	6.78	0.28	-15.2	-5.9	-9.3
	207	23.30	7.37	0.15	-21.4	-11.4	-10.1
pdG × pdC	**196**	2.54	6.40	0.24	-12.7	-4.0	-8.7
	206	14.90	7.17	0.20	-19.9	-10.1	-9.8
	207	27.70	7.44	0.04	-30.0	-19.9	-10.2
pA × pU	**196**	6.38	6.80	0.39	-9.8	-0.6	-9.3
	206	6.34	6.80	0.34	-13.9	-4.6	-9.3
	207	23.90	7.38	0.18	-18.7	-8.6	-10.2
pG × pC	**196**	5.14	6.71	0.37	-12.7	-3.5	-9.1
	206	7.93	6.90	0.24	-19.6	-10.2	-9.4
	207	28.80	7.42	0.12	-27.5	-17.3	-10.2

a) Calculated for 25° C; b) Two-site model.

As predicted, the binding constants for the divalent ligands **196** (Phe-Lys-GCP) and **206** (Lys-Phe-GCP) are in the range of $\geq 10^6$ M^{-1}. For the three-armed **207** (Phe-Lys-GCP)

binding constants are even higher by approximately one order of magnitude (> 10^7 M^{-1}). This is in excellent agreement with the denaturation and EB displacement experiments. The two divalent receptors do not differ significantly in their binding behavior with the exception of $p(dAdT)_2$, which is bound stronger by **206** by a factor of 5 and pdG × pdC, which again is bound stronger by **206** by a factor of 6. In all other combinations the difference is rather small, although there is a general trend for stronger binding by **206** in comparison to **196** similar to the observation made during the EB displacement assays. All nucleic acids are bound with similar affinity by the divalent ligands. **206** has a slight preference, by a factor of 2-4 for pdG × pdC, which is bound with K = 1.5 × 10^7 M^{-1}. All other polynucleotides are bound weaker with binding constants in the range between 3.9 × 10^6 M^{-1} and 7.9 × 10^6 M^{-1}. **196** features similar binding constants ranging from 1.3 × 10^6 M^{-1} to 6.4 × 10^6 M^{-1} with RNA being slightly preferred by a factor of 2-3. As an exception, $p(dAdT)_2$ is bound weaker with a binding constant of only 4.8 × 10^5 M^{-1}. For **207** no preference could be witnessed at all. Binding constants closely range around 2-4 × 10^7 M^{-1}. Obviously the third arm improves the binding affinity by a great deal due to the increased number of non-covalent interactions between nucleic acid and ligand, e.g. additional charge-charge interactions.

From an energetic point of view, all complex formations are exothermic with unfavorable entropic contributions. Although this general energetic scheme is never changed, the complexation of different polynucleotides leads to different contributions of enthalpy and entropy to the binding event. For example, the binding of **207** with ctDNA leads to an enthalpy ΔH of -23.8 kcal/mol. Binding of pdG × pdC by the same ligand is more exothermic (ΔH = -30.0 kcal/mol). This would translate into a binding constant, which is higher by a factor of 4 × 10^4. However, the enthalpic contribution is compensated for by a less favorable entropy $T\Delta S$ (ctDNA: -13.7 → pdG × pdC: -19.9 kcal/mol). As a consequence the binding constants towards the two different DNAs are the same within the error of the instrument (K = 2.8 × 10^7 M^{-1}). Such pronounced enthalpy-entropy compensation is common for binding to macromolecules and a clear indication for the formation of well-defined complexes in contrary to unspecific long-range charge-charge interations.[228] The difference between the two-armed and the trivalent ligand stems solely from a more favorable enthalpy. For instance binding to pdA × pdT by **207** generates a ΔH of -24.6 in comparison to -13.3 kcal/mol for its two-armed analogue **196**, which features identical side chains (Phe-Lys-GCP). This would translate into an increase of binding constant by a factor of 3 × 10^8. However, this effect is by large compensated for by a less favorable entropy $T\Delta S$ (**207**: -14.3 and **196**: -4.9 kcal/mol). Once more, this enthalpy-entropy compensation hints at the formation of a well-defined complex. The stronger enthalpic interaction of **207** with the nucleic acid leads to the formation of a more tightly bound and more rigid complex leading to a significant decrease in entropy. The reason for the improved binding of the three armed ligand is probably due to its higher charge. From the pK_A values of the three ligands (Table 9) it is possible to calculate the protonation state at neutral pH. All lysine amine groups are fully protonated at this pH ($pK_A \approx$ 9.9). The guanidinio groups on the other hand are only partially charged. The two-armed ligands feature two and the three-armed one respectively three of these moieties. With the help of the rearranged *Henderson-Hasselbalch* equation (5) the percentaged amount of charged species can easily be calculated.

$$\frac{HA^+}{A} = 10^{pK_A - pH} \qquad (5)$$

Hence, 50 % of the guanidinio groups are protonated for **196** (pK_A = 7.00), 61 % for **206** (pK_A = 7.20), and 24 % for **207** (pK_A = 6.50). Consequently the three molecules feature an average of 3.0 (**196**), 3.2 (**206**), and 3.7 (**207**) positive charges at pH 7. Thus, the additional charges might be the explanation, why the three-armed compound binds significantly stronger to the nucleic acids and why **206** always binds slightly better than **196**.

With regard to the structure of the nucleic acid-ligand complexes the energetic pattern determined by these experiments does not allow for a clear conclusion. In the following it will be compared to experimentally determined binding enthalpies and entropies of literature-known systems. For instance, charge-driven, unspecific backbone binding by e.g. spermidine leads to an endothermic process which is driven by entropy.[229] Entropy is generated, when the monovalent metal counterions are displaced from the phosphate backbone together with their water shell and liberated into the solvent. Binding constants of these charge-mediated binding events are normally lower than the ones for the three artificial ligands presented herewith. The binding constant for the complexation of spermidine with pUC118 plasmid DNA for instance is 1.4×10^5 M^{-1}. Since almost all polynucleotide-binding molecules are positively charged, the entropic contributions due to liberation of territorially bound cations are general for all modes of interaction according to *Mannings* polyelectrolyte model, be it intercalation, unspecific or groove binding.[230,231] The degree to which DNA-binders add favorable entropy depends on the charge of the ligand and ranges from 1 to 4 kcal/mol.[90] Groove binders for example normally feature modest endothermal or exothermal contributions of ΔH and their complexation is mainly driven by entropy due to the displacement of the highly ordered water and counterion molecules, the spine of hydration, from the grooves.[232] Intercalation on the other hand is mainly driven by enthalpy and often features unfavorable entropic contributions. This is due to the partial unwinding of the DNA which leads to lengthening and stiffening of the helix. The more rigid nucleic acid looses degrees of motional freedom, which leads to unfavorable entropy and thus the favorable contributions to entropy by water/ion displacement are overcompensated.[233]

The energetic pattern generated by nucleic acid complex formation with **196**, **206**, and **207** resembles most the intercalation mode. However, based on the structure of the three ligands intercalation is rather unlikely. The aromatic moieties present in the three molecules, namely the phenyl and pyrrole rings, are too small for efficient stacking into base pairs. The lack of entropic driving force must thus originate from somewhere else. One possible explanation is a binding mode which does not liberate many water molecules and counterions, e.g. groove binding without deep penetration thereof. This would leave the inner water shells intact and only remove the less ordered outer ones. Still, even such a process would liberate enough water/counterions to generate a favorable entropic contribution. Hence, the liberation must be overcompensated by other processes. Possible reasons for unfavorable entropy may either be due to conformational changes of the polynucleotide or stem from the loss of rotational and vibrational degrees of freedom of the ligand, especially if the complex is very tight and well-defined. The pronounced entropy-enthalpy compensation which could also be observed speaks in the favor for the latter aspect.

Binding events of groove-binding proteins, which bind with little distortion of the DNA, for instance feature favorable enthalpic and unfavorable entropic contributions similar to the results obtained for **196**, **206**, and **207**.[234] In contrast to small molecule groove binders, which are regularly rather rigid and pre-shaped to adapt to the groove curvature (cf. DAPI or distamycin in chapter 2.3.5), proteins are large molecules. Upon binding to DNA large conformational changes can be observed (cf. bZIP in chapter 2.3.5), which may either take place at the binding site of the protein or at regions far away from it. By changing their conformation their functional nucleic acid recognizing groups are reoriented in order to favor hydrogen bonding or electrostatic and hydrophobic contacts, thus optimizing the enthalpic gain. On the other hand entropy has to be paid for these conformational changes. Probably something similar happens upon binding of **196**, **206**, and **207**. Although these molecules are still rather small, and may therefore not be treated as protein-like, they are not typical groove binders either, which are already well-shaped for binding. Hence, for the complexation of the di- and trivalent ligands a binding mode can be proposed which proceeds with favorable enthalpy due to charge-charge and hydrophobic interactions paired with hydrogen bonding but leads to a large entropic penalty due to the rigidification and subsequent loss of rotational and vibrational degrees of freedom.

In conclusion, with the help of thermal melting studies, spectroscopic methods and isothermal titration calorimetry it could be shown that the synthesized ligands are able to bind to nucleic acids with high binding constants. The divalent **196** and **206** bind with binding constants $\geq 10^6\ M^{-1}$ and the trivalent **207** binds even stronger by one order of magnitude with $K \geq 10^7\ M^{-1}$. There is no pronounced preference for any type of polynucleotide due to enthalpy-entropy compensation. The binding events are characterized by strong exothermal but entropically unfavorable contributions. Due to their structure, intercalation is not likely and a groove binding mode with severe restraints of degrees of freedom of the ligand upon the formation of a tightly bound complex can be proposed. In order to test this hypothesis additional experiments were undertaken in order to elucidate the binding mode of the three PN-binders.

DAPI Displacement Assay

In order to test the hypothesis of groove binding, displacement assays with the minor groove binder DAPI and the two nucleic acids $p(dAdT)_2$ and pdA × pdT, which form very well-defined minor grooves, were carried out similar to the ones conducted with EB in the last chapter.[235] This way it could be tested whether **196**, **206**, and **207** are able to compete with DAPI for binding in the minor groove. First, a solution of DAPI was prepared in aqueous cacodylate buffer (0.01 M) in a fluorescence microcuvette equipped with a stopper (900 µL) at a concentration of 0.5 µM (1 eq). Upon irradiation at 358 nm the fluorescence spectrum between 420 and 500 nm was recorded at 25° C. Then the polynucleotide was added (2.5 µM, 5 eq) and after 5 min incubation time the spectrum was measured again – now with increased fluorescence intensity upon minor groove binding of DAPI. Stock solutions of **196** (35 µM), **206**, and **207** (each 50 µM) were added to the mixture in aliquots of increasing volume (e.g. 10 × 1, 2, 4, 8 µL) and the spectrum was measured after 1 min incubation time after each addition. Figure 115 depicts a representative titration with pdA × pdT and **207** and the excerpt at 460 nm, which was corrected for dilution and DAPI´s own emission before fitting with a first order exponential decay function (see appendix C.6, Figure C. 66-Figure C. 70 for all

other combinations). All measurements were carried out at least in duplicate. The average amounts of artificial ligand which were necessary to displace 50 % of DAPI were calculated and are listed as the IC_{50} values in Table 13. By carrying out additional titrations of the three ligands into a solution of DAPI it was ensured that there is no interaction between these molecules. Therefore, the displacement of the minor groove binder is only due to the attraction between the nucleic acid and **196**, **206**, or **207**.

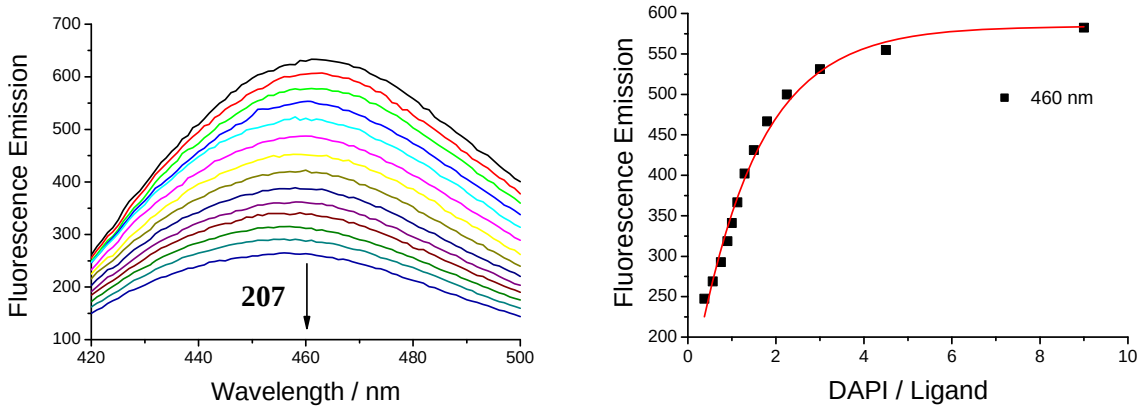


Figure 115. *DAPI displacement experiment with pdA × pdT and **207** (left) and excerpt of the fluorescence emission at 460 nm plotted against the ratio of EB/**207** plus the corresponding fit (right).*

Table 13. *IC_{50} values and standard deviations as obtained from DAPI displacement assays.*

Polynucleotide	**196**	**206**	**207**
$p(dAdT)_2$	52.15 (± 16.98)	15.93 (± 3.68)	1.72 (± 0.34)
pdA × pdT	7.73 (± 2.69)	20.79 (± 5.87)	1.28 (± 0.15)

The results of the displacement assay clearly show that especially trivalent **207** is able to efficiently displace DAPI from both nucleic acids. The divalent ligands **196** and **206** do also displace the minor groove binder but less efficient by one order of magnitude, which is in perfect agreement with the ITC data. Furthermore, **196** is less well able to displace DAPI from $p(dAdT)_2$ than **206**, which is also in good agreement with the calorimetric data. The ability to displace DAPI from pdA × pdT (**196** > **206**) on the other hand does not correspond to the ITC data. There the opposite effect was observed (**196** < **206**). Still, these measurements demonstrate that all three systems are able to displace DAPI from very well-defined minor grooves where this minor groove binder is bound very tightly (pdA × pdT: $K = 6.3 \times 10^7$ M^{-1}, and $p(dAdT)_2$: $K = 8.4 \times 10^6$ M^{-1}).[236,237] These observations are in agreement with the hypothesis of groove binding. Especially for **207** a minor groove binding mode is probable.

CD-Spectroscopy

Circular dichroism spectroscopy is an excellent tool for studying the secondary structure of nucleic acids and their interaction with ligands. Depending on the ligand´s mode of interaction, especially concerning their relative orientation and distance towards the helical axis, different changes in the CD spectra are obtained. Since **196**, **206**, and **207** are chiral molecules, they feature a CD signal on their own. Therefore, the concentration dependent CD signature of these three molecules was determined first. Solutions of increasing amount of the

ligands were prepared in a fluorescence cuvette equipped with a stopper (2 mL) in cacodylate buffer (0.01 M) at neutral pH and the corresponding CD spectra were recorded between 220 and 350 nm at 25° C. The background-corrected spectra (cuvette + buffer) of **196**, **206**, and **207** are depicted in Figure 116. As demonstrated by the excerpts of exemplary absorption maxima, the changes of the CD signals are directly proportional to the concentration.

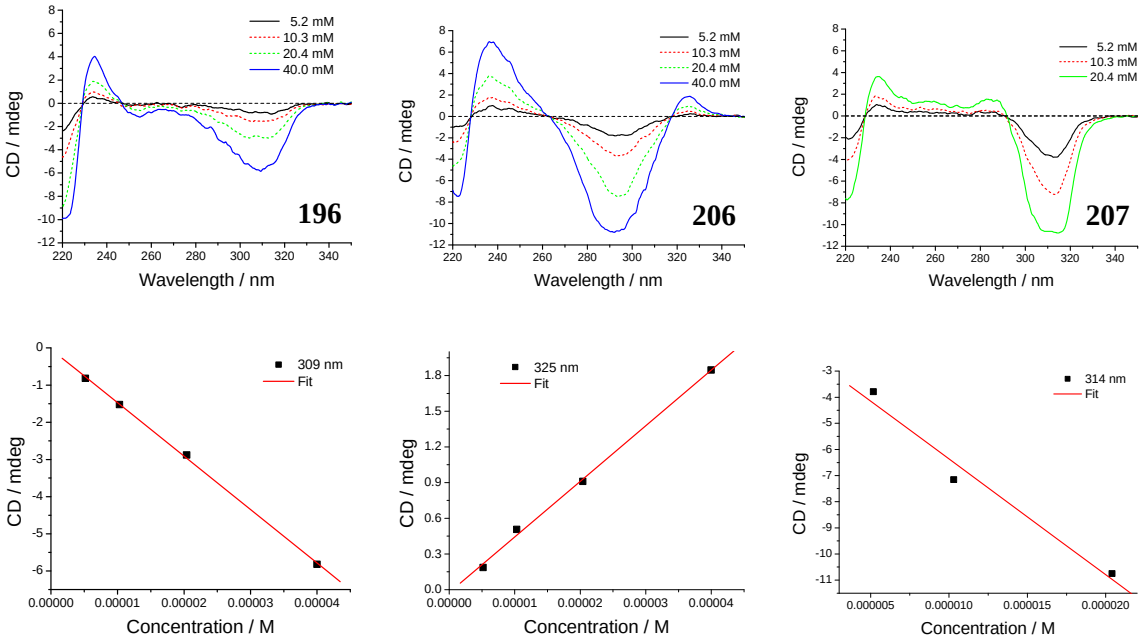


Figure 116. *Concentration dependent CD signature of the three ligands* ***196***, ***206***, *and* ***207*** *revealed linear increase/decrease of the CD signals.*

Due to this linear relationship between concentration and CD signal, the CD spectra of polynucleotide/ligand mixtures could easily be corrected for the ligand´s own CD signature in the following titration experiments. All spectra were furthermore corrected for the background CD signal (buffer + cuvette). For better comparability, the wavelengths of the CD spectra were plotted against the molar CD absorption coefficient $\Delta\varepsilon$, which takes into account the concentration of the nucleic acid and can be calculated from the observable □ after correction for the ligand´s CD signal with the help of the following equation:

$$\Delta\varepsilon = \frac{\Theta}{33{,}000 \cdot c \cdot d}$$

$\Delta\varepsilon$ = molar CD absorption coefficient

□ = ellipticity

c = concentration

d = path length = 1 cm

The factor 33,000 originates from the conversion of the ellipticity unit (mdeg).

First, the binding properties of **196** towards various double-stranded DNAs and RNAs were tested. Solutions of the polynucleotides were prepared in neutral cacodylate buffer (0.01 M) in fluorescence cuvettes equipped with stoppers (1.5-2 mL, 8-40 µM) and their CD spectra were measured between 220 and 350 nm at 25° C. Aliquots of a 1 mM stock solution of the ligand

were added (e.g. 5 × 0.75, 5 × 1.50 µL) and after 1 min incubation time the resulting CD spectra were measured again. Figure 117 illustrates the spectrum between 260 and 340 nm obtained from titrating **196** into a solution of ctDNA. Upon addition of the ligand a symmetric bisignate induced CD (ICD) signal could be observed at the UV/Vis absorption maximum of the pyrrole unit (300 nm). This kind of signal is an indication for exciton coupling between two pyrrole chromophores. Obviously two pyrrole units come within close proximity upon binding to ctDNA. Such a process is plausible when considering that only one of the two GCP units is positively charged at this pH while the other one is neutral. Exciton coupling is not possible for intercalators due to the nearest neighbor exclusion principle and is therefore an indication for groove binding.[78] Furthermore, the isoelliptic point which is present from 0.1 eq onwards suggests a predominant and well-defined binding mode of the ligand for ratios higher than 0.1.

The excerpt at 284 nm (and also at 319 nm) is of sigmoidal character. At ratios higher than 0.1 eq the slope is becoming steeper to flatten again at higher ratios until a plateau is reached at 0.35 eq. At this ratio the ICD inducing groove binding mode is saturated. This value is in good agreement with the stoichiometry n = 0.43 obtained by ITC. In general, the point of saturation correlates very well with the ITC-derived stoichiometry for all ligand-nucleic acid combinations as shown in Table 14. Hence, the process witnessed by ITC can be linked to groove binding. The sigmoidal shape of changes at both, positive and negative ICD bands, suggests that for efficient binding some definite concentration of compounds per DNA binding site has to be achieved. Above this ligand/DNA basepair ratio pyrroles are forced into dimers due to the tight packing of molecules within the DNA minor groove, whereby these pyrrole dimers are in one preferred orientation with respect to the DNA helical axis. The high saturation ratio r = 0.3-0.4 eq suggests that pyrroles from two different molecules are combined into a stacked structure which yields excitonic ICD bands. Otherwise saturation should appear at ratios r < 0.2 eq due to the size of molecules.

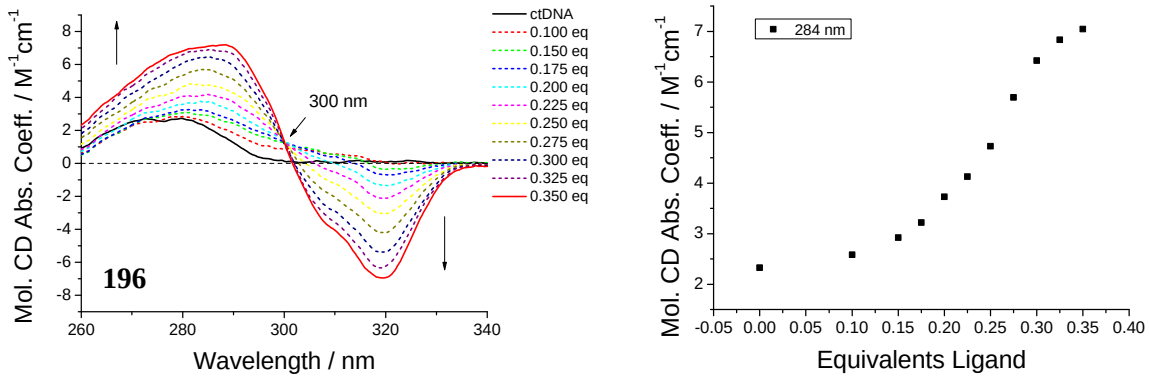


Figure 117. *CD titration experiment with ctDNA and* ***196****. The spectra are corrected for the ligand's CD signal.*

A similar binding mode is adopted by **196** towards all tested B-type helix nucleic acids. Figure 118 shows corresponding CD titration experiments for pdA × pdT, $p(dAdT)_2$, and $p(dGdC)_2$ as well as excerpts of their CD maxima around 320 nm. All spectra feature a bisignate ICD and the excerpts are of sigmoidal shape. The titration with $p(dAdT)_2$ resembles

the one with ctDNA most, which was to be expected due to the similar secondary structure of these two nucleic acids.

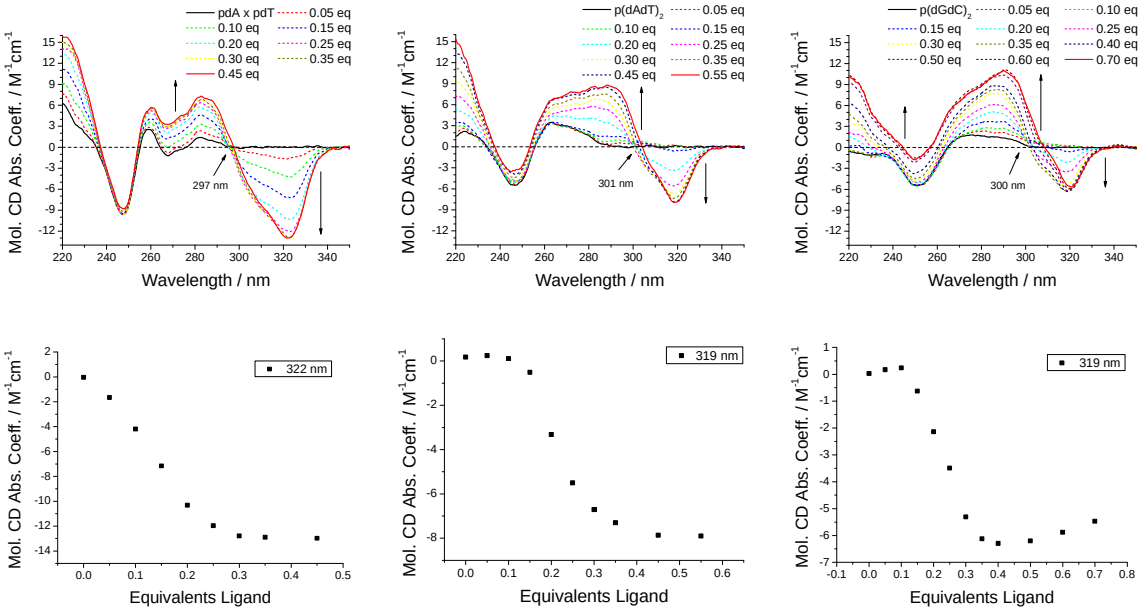


Figure 118. *Top: CD titrations of pdA × pdT (left), p(dAdT)$_2$ (middle), or p(dGdC)$_2$ (right) with* ***196****. The spectra are corrected for the ligand´s own CD signal. Bottom: Excerpts of the CD maximum around 320 nm.*

The titrations differ from the one for ctDNA in the way that the bidentate ICD signal is not symmetric any longer. For all three B-type DNAs different CD signals were obtained. The negative band around 320 nm is most pronounced for pdA × pdT, followed by p(dAdT)$_2$ and p(dGdC)$_2$. The positive signal at 290 nm behaves the other way round. The titration with this nucleic acid also features an isoelliptic point, which starts shifting after saturation (at 0.3 eq). The same was observed for the alternating GC polymer: From 0.05 eq ligand to the point of saturation (0.35 eq) there is an isoelliptic point which starts shifting upon addition of more ligand. p(dAdT)$_2$ behaves similarly, but its isoelliptic point starts shifting before saturation takes place (0.05-0.25 eq). Probably additional unspecific binding modes exist, which do either not result in ICD signals or as in the last case only give rise to a positive ICD which could be an indication for minor groove binding without exciton coupling.

Table 14. *Comparison of the stoichiometry (n) as obtained by ITC measurements with the point of saturation observed during CD titrations.*

L	PN	ctDNA	p(dAdT)$_2$	pdA × pdT	p(dGdC)$_2$	pdG × pdC	pA × pU	pG × pC
196	Saturation	0.35	0.35	0.25	0.35	0.40	0.30	0.35
	n[a]	0.43	0.25	0.24	0.36	0.24	0.39	0.37
206	Saturation	0.50	0.25	0.30	0.30	0.20	0.40	0.30
	n[a]	0.32	0.26	0.17	0.28	0.20	0.34	0.24
207	Saturation	0.20	0.20	0.15	0.10	0.10	0.20	0.13
	n[a]	0.15	0.20	0.06	0.15	0.04	0.18	0.12

a) Obtained from ITC.

As depicted in Figure 119 a totally different result was observed for the GC containing A-type helices pdG × pdC (DNA) and pG × pC (RNA). These titrations also feature a bidentate

signal, but with opposite sign. The band around 285 nm now is negative instead of positive and the one at 310 nm positive instead of negative. The bands are also somewhat shifted to lower wavelengths (5-10 nm). Both spectra feature sigmoidal excerpts at 285 nm (also at 310 nm), and isoelliptic points at the pyrrole´s UV/Vis absorption maximum at ca. 300 nm indicating exciton coupling. *Lewis* reported that the intensity of the positive and negative bands of bidentate CD signals caused by exciton coupling for two identical chromophores with parallel oriented planes depends on the angle between the transition dipoles □ following the function sin(2□).[238,239] Hence, it can be concluded from the existence of the bisignate signal that the two pyrrole units are not aligned parallelly (sin 0° = 0) but are rather twisted helically. Further, *Wagenknecht* could show that a mirror imaged CD-signature of such excition bands (+/- vs. -/+) can be attributed to left- vs. right-handed helical twisting of two chromophores caused by differing features of the site where the two chromophores are implemented into the DNA.[240] According to that the pyrrole units within the pdG × pdC and pG × pC grooves are aligned towards each other with opposite helical chirality in comparison to all other polynucleotides under investigation. For the RNA an isoelliptic point is present throughout the entire titration and the ICD is pronounced strongest. The analogue DNA only features an isoelliptic point starting from 0.15 eq and the ICD is somewhat weaker. Obviously, **196** adopts an alternative binding mode for A- and B-type helices featuring opposite chirality of the alignment of two stacked pyrroles, but is able to bind to both types as already shown by the ITC experiments. Such a differentiation between different secondary nucleic acid structures/sequences by non-covalently binding small molecules is very rare. Probably the ligand is able to bind to both, the major groove for A-helices and the minor groove for B-helices due to its flexibility. To both types it binds via unspecific electrostatic interactions paired with a groove binding mode. Additional binding modes which do not give rise to ICD signals cannot be excluded, however.

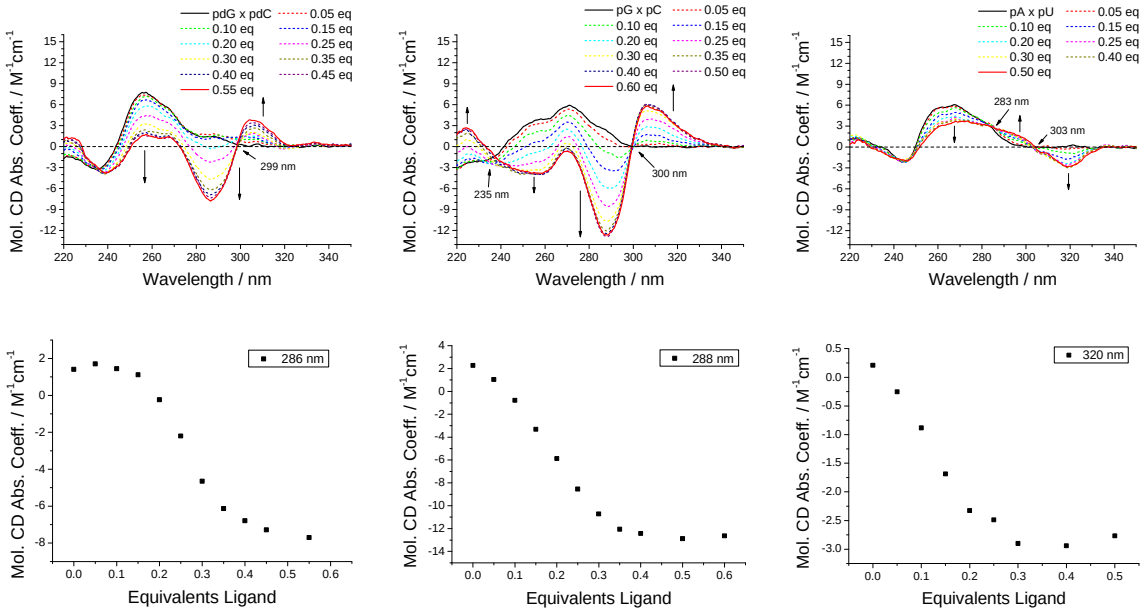


Figure 119. *Top: CD titrations of pdG × pdC (left), pG × pC (middle), or pA × pU (right) with* ***196****. The spectra are corrected for the ligand´s own CD signal. Bottom: Excerpts of the CD maxima around 280 or 320 nm.*

The A-type RNA pA × pU on the other hand is bound differently. Its CD signature resembles more the one observed for B-DNA. A weak, bisignate signal around 300 nm is formed and the spectra feature an isoelliptic point between 0.05 and 0.30 eq. The excerpt at 320 nm does not feature a pronounced sigmoidal shape. Hence, besides the secondary structure of the polynucleotide also the base sequence seems to be crucial for the binding mode of the ligand. Nevertheless, in all cases at least one well-defined groove binding mode could be observed.

Next, the analogue two-armed ligand **206** with inversed amino acid sequence (Lys-Phe-GCP) was tested. Figure 120 shows the results of the CD titrations with the three B-DNAs pdA × pdT, $p(dAdT)_2$, and ctDNA. Only the first one features a clear bisignate CD signal and and isoelliptic point at 300 nm and thus a similar binding mode as for **196**. For $p(dAdT)_2$ at least two different groove binding modes could be observed. First, a broad positive ICD around 280 nm appears, which starts to decrease again when more than 0.25 eq of ligand are present to result in a weak bisignate signal around 300 nm. The negative lobe of this signal is already saturated at 0.25 eq. For ctDNA a similar behavior was witnessed. First, a positive ICD band arises around 280 nm and decreases again with increasing amount of ligand. Unlike the alternating AT-polymer, no bisignate signal was observed at higher ratios. Obviously, it is more difficult for **206** to bring the two pyrrole moieties within close proximity than it is for **196**. A possible explanation for this is the difference in the amino acid sequence. The phenylalanine in **206** is directly attached to the pyrrole moiety, which probably makes it more difficult to stack two pyrrole units above each other. Furthermore, the strongest ICD was observed for pdA × pdT with its very well-defined minor groove. The less well the minor groove is developed (pdA × pdT > $p(dAdT)_2$ > ctDNA > $p(dGdC)_2$), the less likely it is that the ligand binds to it. Hence, the observed signals are another strong indication for minor groove binding in B-type DNA.

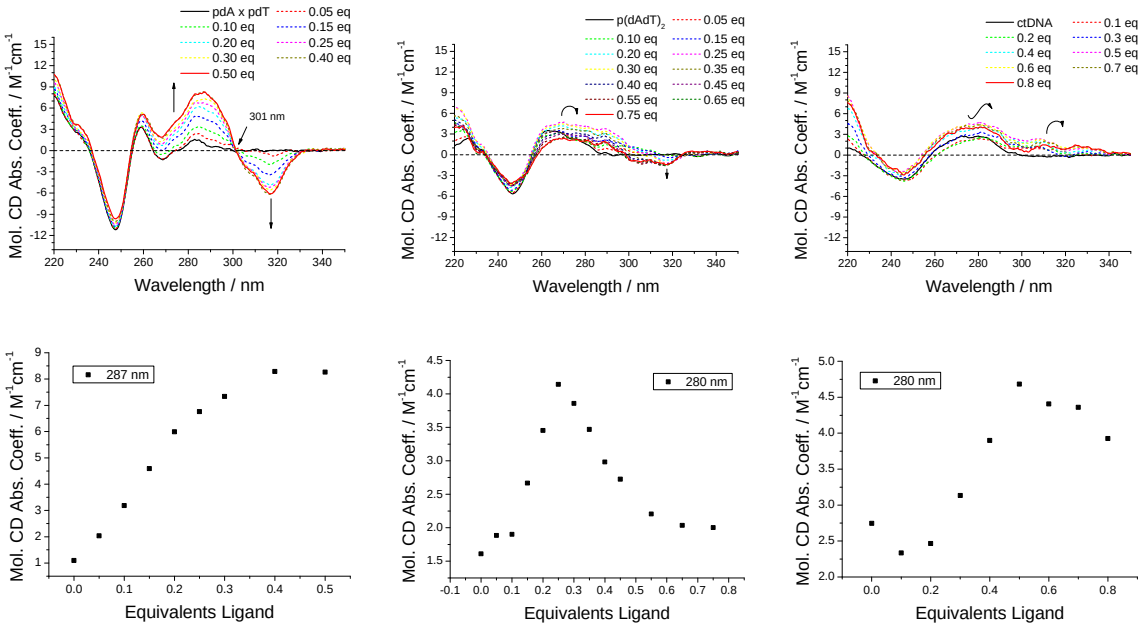


Figure 120. *Top: CD titrations of pdA × pdT (left), $p(dAdT)_2$ (middle), or ctDNA (right) with* ***206****. The spectra are corrected for the ligand's own CD signal. Bottom: Excerpts of the CD maxima around 280 nm.*

This observation is backed up by the result of the titration with $p(dGdC)_2$ depicted in Figure 121. During the titration a kinetic effect could be observed (left). Even after 4 min incubation time the equilibrium state was not yet reached. Therefore the experiment was repeated by *Dr. Lidija-Marija Tumir* from the working group of our cooperation partner *Dr. Ivo Piantanida* at the Ruđer Bošković Institute (Zagreb, Croatia). Individual mixtures of varying DNA/ligand ratios were prepared and incubating for 24 h (middle). The resulting spectra feature a complicated pattern of different binding modes without isoelliptic point. First, two positive ICD bands form at 280 and 310 nm with a maximal intensity at 0.3 eq. Then these bands decrease again and a new bisignate signal is formed around 320 nm with a negative lobe around 305 nm and a positive lobe around 330 nm. When even more ligand is added, this signal vanishes again, too. Hence, in agreement with the other CD signatures for B-DNA, **206** seems to have the greatest trouble in forming a defined complex to $p(dGdC)_2$ with its poorly accessible minor groove (shallow, featuring the electropositive guanine NH_2).

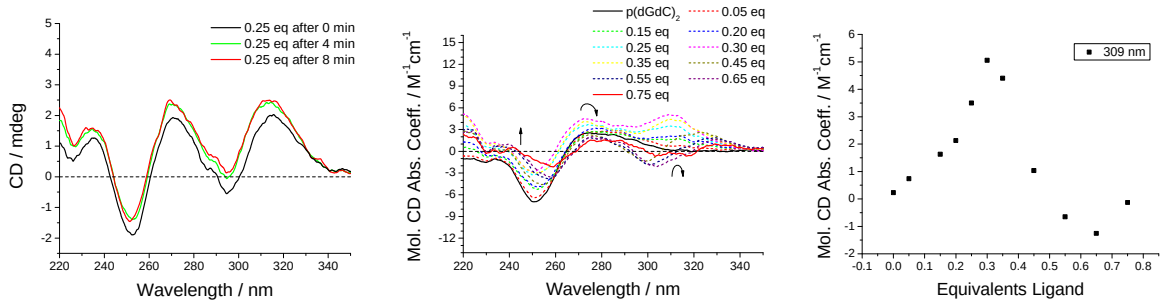


Figure 121. *Uncorrected CD spectrum of $p(dGdC)_2$ and 0.25 eq **206** (left). The CD titration with 24 h incubation time (middle) and the excerpt at 309 nm are corrected for the ligand's CD signal.*

With regard to A-type polynucleotides the experiments gave similar results as for **196**. A bisignate signal with a clear isoelliptic point at 297 nm could be observed for the homopolymeric GC RNA (Figure 122, left). The ICD signals were very strong, especially the negative lobe at 285 nm. Also for the corresponding DNA pdG × pdC a bisignate signal with an isoelliptic point at 295 nm was observed. This point however started shifting after saturation of the groove at 0.25 eq. The most striking feature of this titration was the extremely strong ICD signal with $\Delta\Delta\varepsilon$ values as strong as -34 $M^{-1}cm^{-1}$. Obviously this nucleic acid is suited best for groove binding of **206** and stacking of two pyrrole units within the groove. This finding is also backed up by the slight preference (factor 2-4) of this ligand for pdG × pdC, which was revealed by the ITC experiments. Hence, for GC-containing A-type RNA the divalent ligand most likely binds into the well accessible major groove. For pA × pU a related signal could be observed this time. A very weak bisignate signal occurred, which was shifted towards higher wavelength compared to the other ones (10 nm). Again, the type of base pairs present in the RNA seems to play a decisive role for the binding mode of the divalent ligand. Thus, just like **196** (vide supra) the ligand **206** is also able to differentiate between the polynucleotides secondary structure and sequence by non-covalently binding to the groove and thereby aligning the pyrrole moieties towards each other with opposite helical chirality (left- vs. right-handed) and thus generating an excition CD signal of opposite sign.

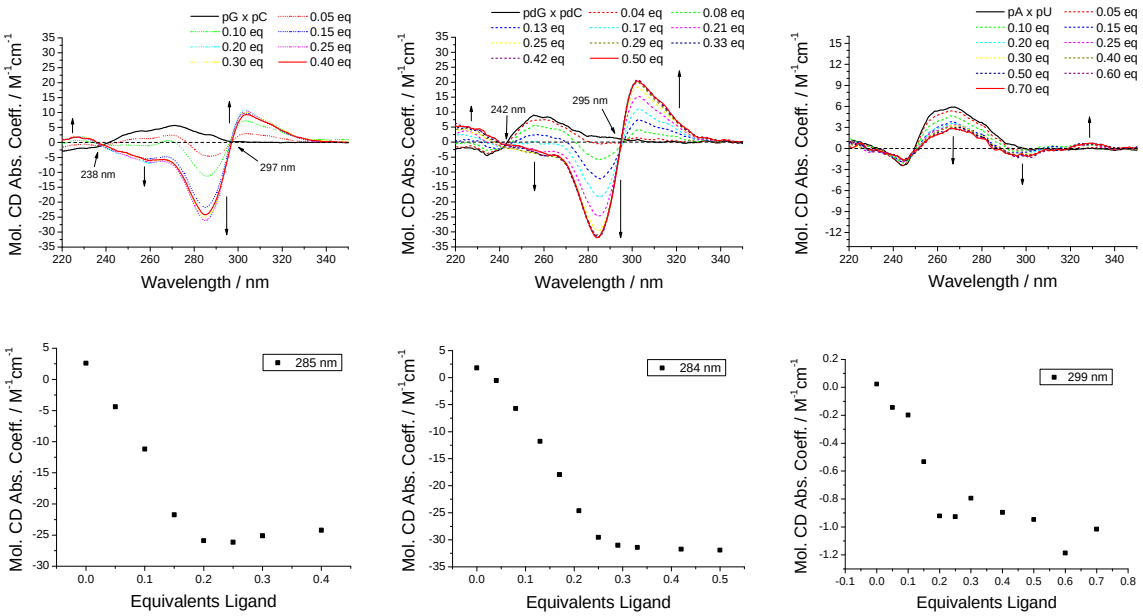


Figure 122. *Top: CD titrations of pG × pC (left), pdG × pdC (middle), or pA × pU (right) with* ***206****. The spectra are corrected for the ligand's own CD signal. Bottom: Excerpts of the CD maxima around 280 or 320 nm.*

Finally, the binding properties of the trivalent ligand **207** were characterized. This nucleic acid binder has the same amino acid sequence as its divalent analogue **196** (Phe-Lys-GCP). Accordingly, the results are quite similar to this ligand as illustrated for the four tested B-DNAs in Figure 123. For all combinations a bidentate ICD signal around 300 nm could be observed with a positive lobe at 290 nm and a negative one at 320 nm. Again, the intensity of the ICD signal correlates with the accessibility of the minor groove. The strongest signal occurred for pdA × pdT, followed by ctDNA and $p(dAdT)_2$. For the not very well accessible $p(dGdC)_2$ the signal was weakest and multiple binding modes were present. In all cases the sigmoidal character of the excerpts of the ICD signals was not very pronounced. Further, saturation takes place at lower equivalents, which is in agreement with the results from the ITC experiments. Hence, in this experimental setup there were not enough data points in the very beginning of the titration to determine the threshold value for groove binding. All polynucleotide titrations featured a shifting isoelliptic point at 310 nm, an indication for at least two binding modes. The overall signal intensity was lower than for the divalent analogue **196**. A possible explanation for this observation are additional unspecific binding modes which do not give rise to an induced CD. These modes might be due to the third arm which does not fit into the minor groove together with the other two. An alternative explanation is an orientation of the third arm which leads to a counterbalancing ICD signal, e.g. if it shows into the opposite direction of the groove as do the other two arms. This way, the electronic transition dipole moment would point ca. 180° into the other direction with regard to the helical axis. The energetic signature derived from ITC titrations is in agreement with such an explanation. The higher binding affinity of the tripodal ligand mainly stems from increased enthalpic contributions to the binding constant, which are in part compensated for by unfavorable entropic contributions. This signature can be explained by a more tight and rigid binding mode which is accompanied by a loss of rotational and vibrational degrees of freedom.

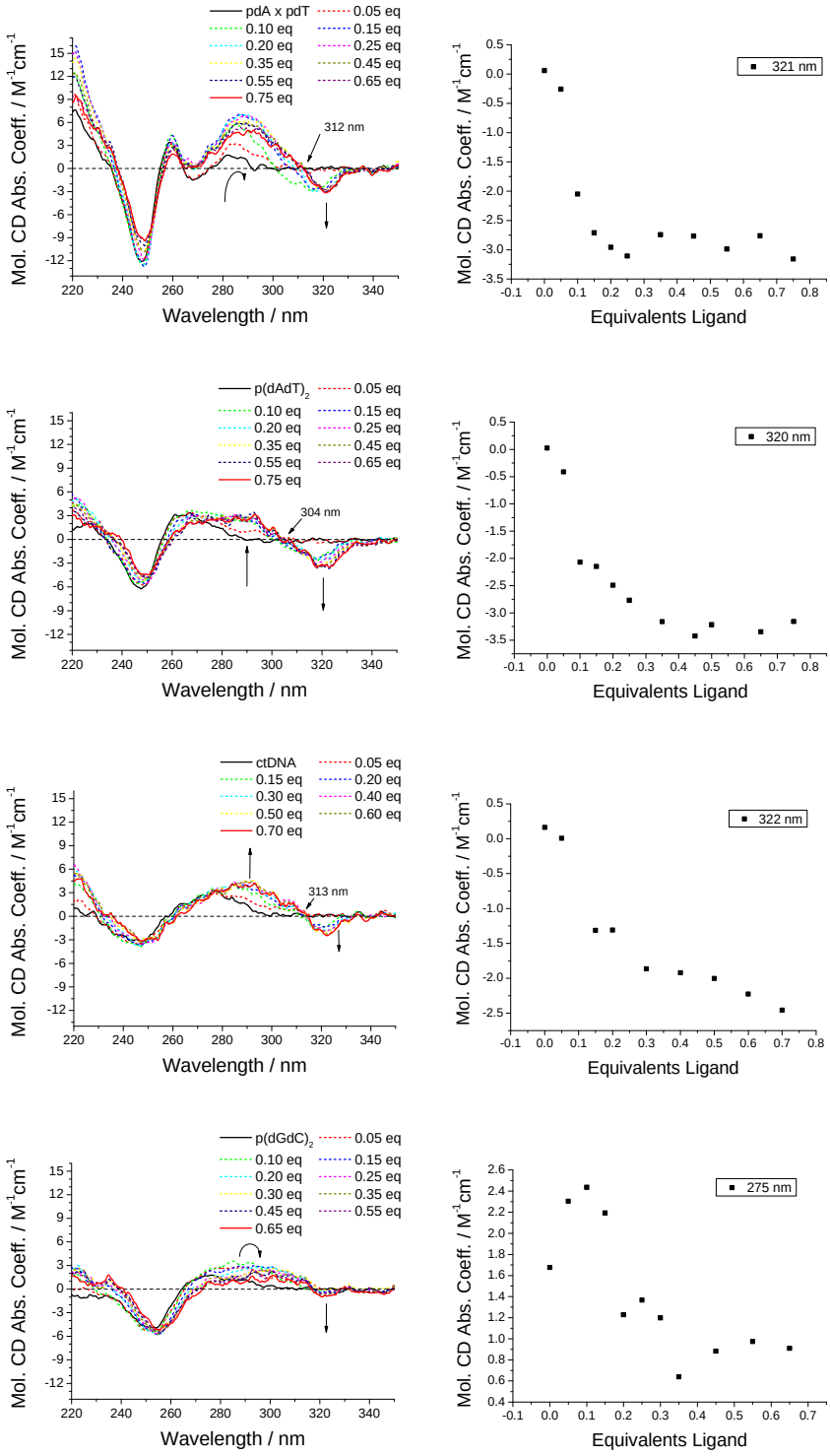


Figure 123. *Left; from top to bottom: CD titrations of pdA × pdT, $p(dAdT)_2$, ctDNA or $p(dGdC)_2$ with* ***207****. The spectra are corrected for the ligand's own CD signal. Right: Excerpts of the CD maxima around 275 or 320 nm.*

Figure 124 illustrates a calculated structure of the B-type *Dickerson-Drew* dodecamer with **207**.[71] The molecular modeling was carried out with Macromodel v9.6 making use of the OPLC_2005 force field and GBSA water solvation. First, the X-ray crystal structure-derived DNA was minimized without restraints together with the ligand **207** in the desired geometry. Then a conformational analysis with frozen DNA and 10,000 steps was conducted until the depicted minimum was found multiple times. According to the calculation the ligand adopts a

stable geometry with two parallel arms in the minor groove and the third arm directed into the opposite direction of the same groove. At neutral pH only one of three GCP units carries a positive charge. The calculation illustrates the case where two neutral GCP units are stacked above each other. The pyrroles are oriented towards the groove floor, while the lysine ammonium groups bind to the ribophosphate backbone and all phenyl rings are outside the groove. However, this binding mode is probably not the only one but represents rather one of the modes which lead to the observed CD signal.

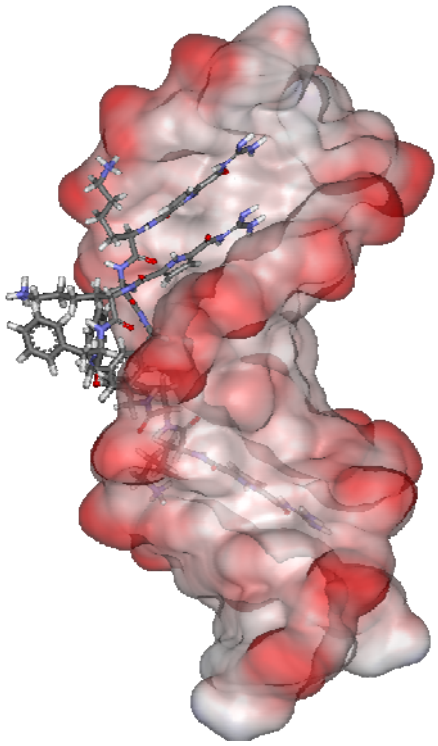


Figure 124. *Molecular modeling of* ***207*** *in the minor groove of the Dickerson-Drew dodecamer. Two arms are oriented to the same side while the third arm points into the other direction.*

The titrations with A-type nucleic acids are depicted in Figure 125. Similar to the divalent ligands a bidentate signal was observed for GC containing DNA and RNA featuring a clear isoelliptic points at ca. 300 nm. The sign of the lobes is inverted in comparison to B-type DNA and stronger. Obviously, a well-defined complex can be formed in the major groove of this kind of polynucleotides, which is well-formed for ligand accommodation and big enough to allow for binding of the large trivalent ligand. For these two nucleic acids the energetic signature derived from ITC features the largest enthalpy gain together with the largest unfavorable entropy. This signature speaks in favor of a distinct binding mode. On the other hand, a totally different signature could once more be observed for the RNA pA × pU. A negative ICD was observed at 320 nm and isoelliptic point at 297 nm. By ruling out an intercalating binding mode such a signal can only stem from major groove binding without stacking of the pyrrole units. At this point it has to be stated, that a well-defined binding mode does not necessarily correspond to a higher binding constant. For instance, according to the displacement assays and ITC experiments **206** binds slightly better to all nucleic acids than **196** although the latter one forms more defined complexes according to the CD titrations.

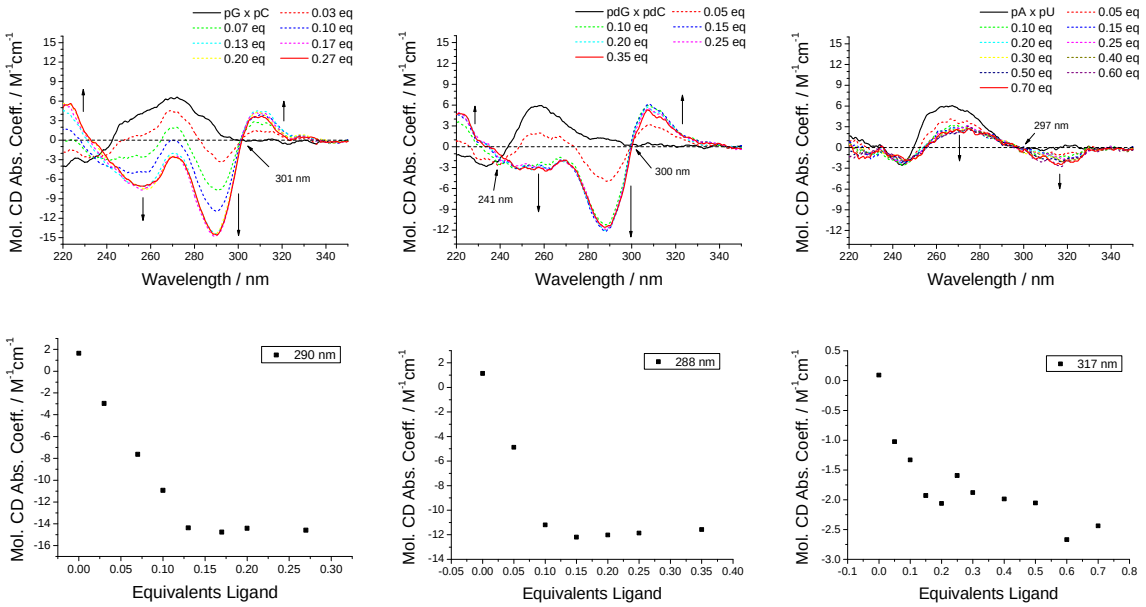


Figure 125. *Top: CD titrations of pG × pC (left), pdG × pdC (middle), or pA × pU (right) with* ***207****. The spectra are corrected for the ligand's own CD signal. Bottom: Excerpts of the CD maxima around 290 or 320 nm.*

In conclusion, the binding studies revealed that all three ligands are able to strongly interact with nucleic acids. The two-armed ligands **196** and **206** bind without clear preference for any type of polynucleotide with binding constants $K > 10^6$ M^{-1}. The results from the displacement assays and the ITC experiments show that **206** binds with slightly higher affinity than **196**. The trivalent analogue **207** also binds without preference for a specific polynucleotide but more efficiently with $K > 10^7$ M^{-1}. The energetic profiles of the binding events of all three ligands, which were determined via ITC, resemble most the one of groove binding proteins or intercalators (exothermal with unfavorable entropy). Due to the chemical structure of the ligands with their rather small aromatic moieties intercalation can be ruled out, however. Strong enthalpy-entropy compensation and the occurrence of isoelliptic points during the CD titrations suggest the formation of well-defined polynucleotide/ligand complexes. The exothermal binding event is most likely due to charge-charge and hydrophobic interactions paired with hydrogen bonding. Upon formation of the well-defined complex an enthalpic penalty has to be paid due to the loss of degrees of freedom within the ligand. According to the energetic signature obtained via ITC (more exothermic, less favorable enthalpy) the addition of the third arm of **207** allows for additional non-covalent interactions and consequently leads to the formation of an even more tighly bound complex.

However, in most cases two binding modes could be observed. One of these modes was attributed to unspecific electrostatic binding while another one could be linked to groove binding with the help of ITC experiments and CD spectroscopy. The results from the DAPI displacement assays are also in agreement with a groove binding mode. Furthermore, the CD spectra of PN/ligand complexes featured in most cases a bisignate signal at the UV/Vis absorption maximum of the pyrrole group, which was ascribed to the stacking of two GCP units with subsequent exciton coupling. Due to the nearest neighbor principle the existence of the exciton signal clearly rules out an intercalating binding mode. According to the saturation ratios (r) of the ICD signals derived from groove binding the stacking is most likely

intermolecular for the divalent ligands **196** and **206** ($r \approx 0.3$-0.4) while it is intramolecular for the tripodal analogue **207** ($r \approx 0.1$).

For B-helical DNA the ICD signal intensities correlate well with the accessibility of the minor groove. Stacking of the pyrrole units of **196** (Phe-Lys-GCP) is more efficient than for **206** (Lys-Phe-GCP) due to sterical hindrance within the groove caused by the close proximity of the rather bulky phenyl group next to the pyrrole moiety. For the other B-type polynucleotides unspecific electrostatic interactions and minor groove binding without stacking are preferred. The trivalent ligand **207** also binds to the minor groove, but it adopts a conformation in which it forms intramolecularly stacked pyrroles. The additional third arm is also bound tightly according to ITC and probably pointed into the opposite direction as indicated by the rather low intensity of the observed ICD signals. In general, the results from the CD and ITC experiments strongly suggest a minor groove binding mode for B-type nucleic acids for all three ligands together with additional unspecific alternative electrostatic binding modes.

For A-type nucleic acids on the other hand a major groove binding mode together with unspecific electrostatic binding is preferred. All nucleic acid/ligand complexes feature bisignate CD signals similar to B-helical DNA. Especially for GC-containing polynucleotides very strong ICD signals could be observed. Most strikingly, the lobes of the exciton signals feature inverted signs (-/+ vs. +/-) in comparison to B-type DNA. This inversion is probably caused by a helical twist of opposite chirality (M vs. P) of the faces of the two stacked pyrrole moieties. Furthermore, for RNA also the type of base pairs seems to be important: pA × pU always features a CD signature which is distinct from pG × pC. Thus, although the overall binding affinity of the ligands towards different nucleic acids is rather similar, they are able to differentiate between GC- and AU-containing RNAs as well as between A- and B-type secondary structures by producing distinct CD signatures.

For a more detailed analysis of the structure of the polynucleotide/ligand complexes the exact knowledge of the electronic dipole transition moment of the GCP unit would be necessary.[241,242] Alternatively, additional experimental methods, like NMR- or X-ray crystal structure analyses, would be required for the determination of the exact binding mode of the ligands. Nevertheless, as evidenced by CD spectroscopy the three ligands are able to differentiate between secondary structure and sequence of polynucleotides by means of excition coupling and chiral twisting of two stacked pyrrole units. Such differentiation by non-covalently bound small molecules is very rare and might be of interest for the development of nucleic acid sensors or for DNA analysis. In the next chapter it will be tested whether the strong interaction towards nucleic acids can also be exploited for an application in another field: cell biology.

4.4.4 Transfection Experiments

After having shown, that the three artificial nucleic acid ligands are able to strongly bind to polynucleotides, their capability to shuttle genetic information into human cells was tested. For transfection experiments, three main classes of DNA carriers are utilized nowadays: cationic lipids,[149] cell penetrating peptides,[160] and cationic polymers.[138] At the moment, there are two gold standards, which are widely being applied for gene shuttling, both in vitro and in vivo: the cationic polymer polyethyleneimine (PEI)[144] and the cationic lipid lipofectamine.[18]

The transfection experiments which will be presented in the following were carried out by our cooperation partner, the working group of *Prof. Shirley Knauer* from the Institute of Biology at the University of Duisburg-Essen by *Aljona Gutschmidt* in the course of her master thesis.[243] The human embryonic kidney cell line HEK293T was transfected with a plasmid encoding for a green fluorescence protein (pF143GFP). The plasmid was derived as reported in literature.[244] Ca. 10^4 cells were seeded in 96-well cell culture plates and grown for 24 h. Then, a mixture of the plasmid (1, 2, or 4 µg) and varying amounts of the artificial gene carriers **196**, **206**, and **207** (0.12, 0.18, 0.24, 0.30, or 0.36 mM) were incubated in phosphate buffered saline (PBS) at pH 7.2 for 10 to 15 min and added to the cells (total volume = 200 µL). The commercial transfecting agent PEI was used as a control. Successfully transfected cells were measured 24 or 48 h after treatment as the expression of GFP with the help of an inverted fluorescence microscope. The data were processed and analyzed using *MetaMorph* and *Adobe Photoshop*. All experiments were carried out at least in triplicate. The average values of the percentaged quotient of transfected cells to all plated cells as well as the corresponding standard deviations are listed in Table 15.

Table 15. *Transfection efficiency determined 24 or 48 h after treatment of HEK293T cells with varying amounts of pF143-GFP plasmid and **196**, **206**, **207**, or polyethyleneimine (PEI).*

pF143-GFP / [µg]	[Ligand] / [mM]	Time [h]	Transfection Efficiency and Standard Deviation / [%]			
			196	**206**	**207**	PEI
4	0.12	24	-	-	9.2 (± 1.5)	27.4 (± 6.0)
	0.18	24	< 1	-	21.9 (± 4.9)	n.d.
	0.24	24	< 1	-	32.5 (± 2.3)	n.d.
	0.30	24	-	3.3 (± 1.3)	12.4 (± 5.7)	n.d.
2	0.12	24	n.d.	n.d.	n.d.	21.5 (± 6.1)
	0.12	48	n.d.	n.d.	n.d.	50.7 (± 9.4)
	0.18	48	-	< 1	32.7 (± 10.7)	n.d.
	0.24	24	n.d.	n.d.	26.8 (± 10.6)	n.d.
	0.24	48	< 1	2.8 (± 0.6)	54.1 (± 12.5)	n.d.
	0.30	48	< 1	n.d.	n.d.	47.3 (± 10.0)
	0.36	48	< 1	n.d.	n.d.	33.6 (± 8.1)
1	0.12	24	n.d.	n.d.	n.d.	10.3 (± 2.9)
	0.12	48	n.d.	n.d.	n.d.	52.7 (± 14.8)
	0.24	24	n.d.	n.d.	17.3 (± 2.1)	n.d.
	0.24	48	n.d.	n.d.	30.7 (± 5.5)	n.d.

The first experiments were carried out with 4 µg plasmid DNA and 0.12 mM PEI as described by the supplier, as well as with increasing amounts of the three synthetic ligands

(0.12-0.30 mM). After 24 h there was no substantial amount of fluorescent cells visible for any concentration of the two-armed ligands **196** and **206** (< 1 to 3.3 %). For PEI (27.4 %) and trivalent gene carrier **207** at a concentration of 0.24 mM (32.5 %) on the other hand a significant amount of cells did express the fluorescent protein. At lower concentrations of the three-armed ligand the transfection efficiency dropped (e.g. 21.9 % at 0.18 mM). At higher concentrations precipitation of the DNA/ligand complex was observed and consequently the efficiency dropped again (12.4 % at 0.30 mM). To test whether the amount of plasmid influences the efficiency, additional experiments were conducted with the optimized concentration of the artificial ligands (0.24 mM) and 2 µg of pF143-GFP. The resulting fluorescence was measured 48 h afterwards. **196** was still not able to efficiently transfect the cells, and the ligand with inverse sequence **206** only with very low efficiency (2.8 %). The trivalent ligand **207** however now was able to transfect 54.1 % of the cells, thus slightly outperforming the transfection gold standard PEI (50.7 %). At an even lower amount of DNA (1 µg), the efficiency dropped for **207** (30.7 %), but not for PEI (52.7 %). Thus, it could be shown that the transfection efficiency of **207** is optimal at a concentration of 0.24 mM with 2 µg of plasmid DNA. The plasmid comprises 6,238 bps corresponding to a molecular weight of ca. 4,050 kDa (calculated with an average weight of 650 Da per bp). Consequently, 2 µg DNA dissolved in 200 µL equals a concentration of DNA phosphates of approx. 30 µM. Hence, at the optimized transfection conditions eight equivalents of **196**, **206**, and **207** or six equivalents of PEI were utilized. The results of the cell assays at these conditions are illustrated in Figure 126.

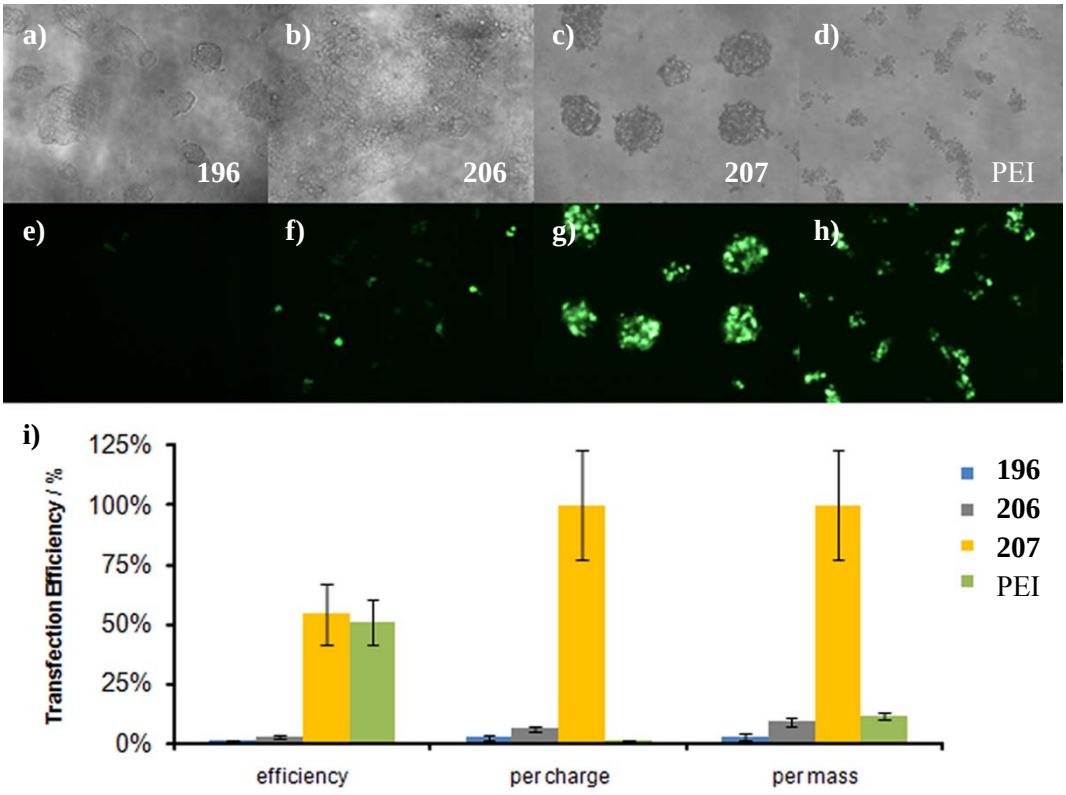


***Figure** 126. Brightfield (a-d) and fluorescence images (e-f) of HEK293T cells 48 h after transfection with 2 µg pF143-GFP plasmid and either 0.24 mM **196** (a, e), **206** (b, f), **207** (c, g), or 0.18 mM PEI (d, h); i) transfection efficiencies of **196**, **206**, **207**, and PEI: Absolute (left), per charge at pH 7.2 (middle), and per mass (right). The latter two are normalized to the efficiency of **207** as 100 %.*

Since PEI is a polymer (M_W = 25 kDa) and the synthetic gene carriers are rather small in comparison (M_W (**196**, **206**) = 1.3; M_W (**207**) = 1.7 kDa) also mass-normalized efficiencies were calculated. Furthermore, charge-normalized efficiencies were calculated. The average number of positive charges at pH 7.2 was calculated with the help of the *Henderson-Hasselbalch* equation, the reported pK_A value for PEI (pK_A = 8.4; per average 580 repetition units per polymer),[245] and the measured pK_A values for **196**, **206**, and **207** as listed in Table 9 one molecule of **196** carries on average 2.8, **206** 3.0, **207** 3.5, and PEI 546 positive charges. As illustrated in the chart at the bottom of Figure 126 the trivalent gene carrier **207** is ten times as efficient per mass as PEI (normalized efficiency of PEI relative to the mass of **207** = 9 %) and even 100 times as efficient per charge (normalized efficiency of PEI relative to the charge of **207** = 1 %). These findings demonstrate that **207** is a significant better transfecting agent than PEI.

Besides efficient transfection, the second crucial characteristic of a gene carrier is whether it is cytotoxicity or not. Therefore *Gutschmidt* determined the cytotoxicity of the non-viral vectors with the help of a trypan blue assay.[246] This dye is only able to penetrate cell walls of dead cells, and can therefore be used for the monitoring of cell toxicity. Thus, the cell vitality 48 h after treatment with the gene carriers was determined to 103 ± 3 % for **196**, 80 ± 2 % for **206**, 92 ± 4 % for **207**, and 81 ± 10 % for PEI (normalized to untreated cells = 100 %). Hence, **207** does not only feature superior transfection properties, it is also less cytotoxic than PEI.

The two-armed ligands on the other hand are not able to efficiently transfect cells. This large difference is surprising, especially when comparing the two compounds **196** and **207**, which both feature the same side chains (Phe-Lys-GCP). Transfection experiments with 0.36 mM of ligand **196** were carried out. At this molarity the number of functional groups of divalent **196** is the same as for its trivalent analogue **207** at 0.24 mM. Still, even with the 1.5 times increased concentration **196** was not able to transfect cells. It was therefore concluded that the difference between two- and three-armed gene carriers does not depend on the absolute number of functional groups but must originate from somewhere else. To get some more insight into the binding properties towards the plasmid, binding studies with the pF143-GFP plasmid were carried out. Illustrations of the EB displacement and ITC experiments can be found in appendix C.4 (Figure C. 48-Figure C. 50) and C.5 (Figure C. 62). The results of the binding studies for the plasmid are listed in Table 16. For comparison also the results for ctDNA are reported.

Table 16. *Binding data for the interaction of **196**, **206**, and **207** with ctDNA and pF143-GFP plasmid at pH 7.0 in sodium cacodylate buffer (0.01 M) at 25° C.*

Ligand	log K^a		ΔH^a		$T\Delta S^{a,b}$		n^a		$IC_{50}{}^c$	
	DNA	GFP	DNA	GFP	DNA	GFP	DNA	GFP	DNA	GFP
196	6.23	6.30	-12.0	-9.6	-3.5	-3.2	0.43	0.24	4.36	1.85
206	6.69	6.96	-15.7	-17.5	-6.6	-8.0	0.23	0.21	3.17	2.08
207	7.44	7.42	-23.8	-22.7	-13.7	-12.7	0.15	0.09	0.79	0.51

a) Determined via ITC titration of ***196***, ***206*** (each 0.6 mM) and ***207*** (0.3 mM) to DNA solutions (0.09 mM for ***196;*** 0.1 mM for ***206*** and ***207***); b) Calculated for 25° C; c) IC_{50} equals the amount of ***196***, ***206***, or ***207***, which is necessary to displace half of the ethidium bromide from a DNA/EB complex in a fluorescence displacement assay, [DNA] = 3.00 µM, [EB] = 0.75 µM.

The results for pF143-GFP are similar to those obtained for ctDNA. The binding constant for the trivalent carrier **207** ($3 \times 10^7\ M^{-1}$) is higher than for divalent **206** ($9 \times 10^6\ M^{-1}$) and respectively **196** ($2 \times 10^6\ M^{-1}$). A possible explanation for these differences is the different charge of the ligands. To an extent, the differences in affinity surely affect the transfection efficiency (**207** > **206** > **196**). However, the binding constants alone are not a satisfying explanation for the dramatic differences in transfection efficiency between di- and trivalent gene carriers (**207** >> **206** > **196**), especially between three-armed **207** and two-armed **196** which share the same side chain amino acid sequence. A potential source for this difference is dissimilar structural changes of the nucleic acid upon complexation with the ligands. In order to elucidate these changes atomic force microscopy (AFM)[247] and dynamic light scattering (DLS)[248] measurements were carried out with **196** and **207** next. Because the binding properties of the ligands towards pF143-GFP plasmid and ctDNA are similar, the latter one was utilized as model system.

In order to verify these data and to get additional information of the formed complexes, DLS measurements of mixtures of **196** and **207** with ctDNA were conducted. This technique makes use of the different *Rayleigh* light scattering properties of differently sized particles irradiated with monochromatic, coherent laser light (e.g. at 633 nm).[249] Due to the *Brownian* motion, and resulting interference of the scattered light the signal intensity of the scattered light is not constant. The *Brownian* motion depends on the size of the particles. As a measure for the mobility the translational diffusion coefficient D is characteristic for a given particle. It depends on the hydrodynamic radius R_H of the particle, i.e. the shell of solvent molecules moving with the same rate as the molecule itself, and can be calculated with the help of the *Stokes-Einstein* equation (5) assuming spherical particles.

$$R_H = \frac{kT}{3\pi\eta D} \quad (5)$$

k = *Boltzmann´*s constant
T = temperature
η = viscosity

The diffusion is faster for molecules with small hydrodynamic radius and as a consequence the fluctuation of the scattered light happens with higher frequency for small molecules. The correlation between the signal at a certain time point t with the signals within short intervals of time Δt gives a decaying, sigmoidal curve which is a characteristic for the size of a particle (rapid decay for small molecules and vice versa). Fitting of the so obtained curve with a single exponential decay function gives the mean size of the sample (Z-average diameter) and an estimation of the width of the size distribution, the polydispersity index (PDI). The higher the PDI ($0 < PDI < 1$), the less monodisperse is the sample. By fitting the correlation curve with multiple exponential decay functions the distribution of particle size for polydisperse samples is accessible. Three possibilities exist to calculate the so obtained size distribution: The size of the particle can either be plotted against the number or the volume of the differently sized particles or against the intensity of the scattered light. The scattering intensity is proportional to the particle diameter to the power of six (D^6) and thus a plot against intensity weights large particles most strongly. Plotting against the volume and respectively the number of particles weights smaller particles more and more strongly. Due to the weighting, the calculated

diameter will in general decrease in the following order: *D*(intensity) > *D*(volume) > *D*(number). For studying the influence of **196** and **207** on the size of DNA upon complexation, a solution of ctDNA (50 µM) was prepared in aqueous sodium cacodylate buffer (0.01 M) at neutral pH in a fluorescence microcuvette equipped with a stopper (1 mL) and a DLS measurement was conducted. Aliquots of a stock solution (1 mM) of **196** or **207** were then titrated into the DNA solution and after 1 min incubation time the scattering experiment was repeated. Figure 127 illustrates the volume- and number-weighted size distribution of ligand/ctDNA mixtures at different ratios *r*. The results are summarized in Table 17. All measurements were carried out at least in duplicate. The size was calculated for spherical particles and is therefore an approximation. The viscosity of the buffered aqueous solution was assumed to be equal to pure water.

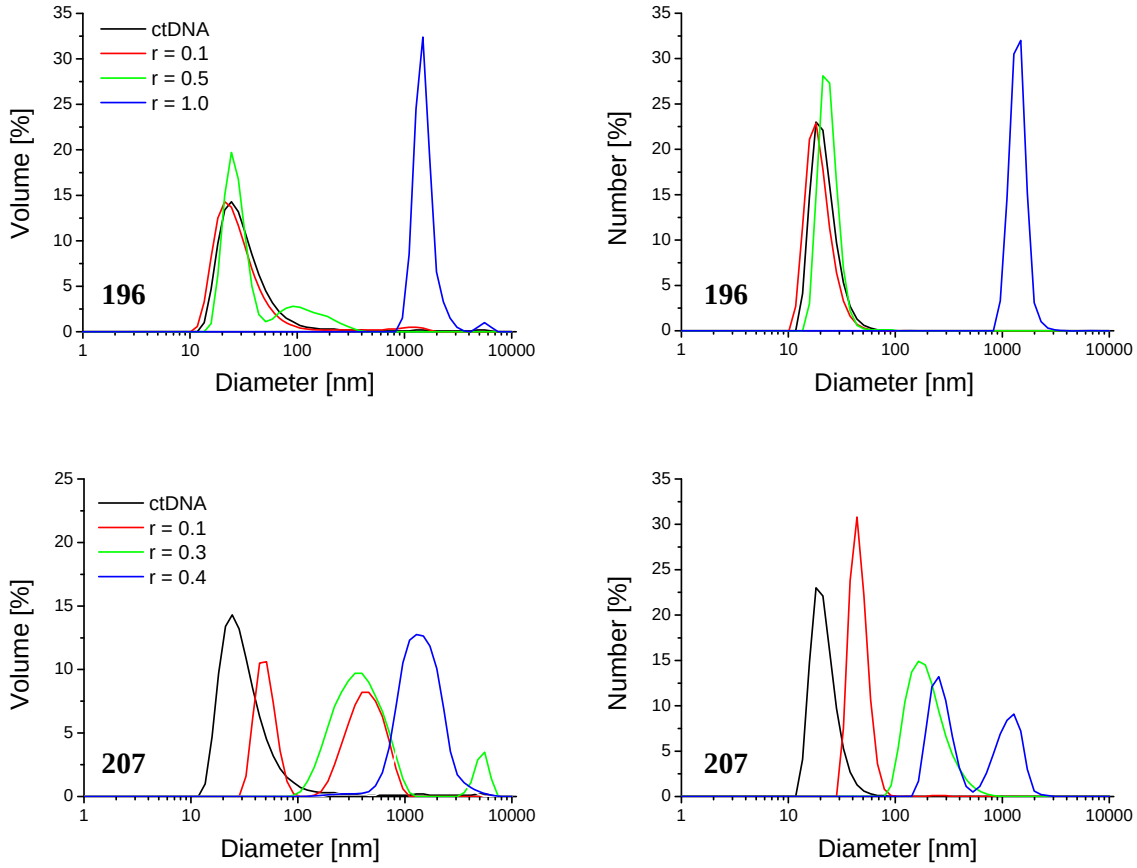


Figure 127. *Top: Volume- (left) and number-weighted (right) size distribution of complexes between ctDNA and* ***196*** *obtained from DLS measurements at pH 7.0 in 0.01 M sodium cacodylate buffer; [ctDNA] = 50 µM; r = [compound] / [ctDNA]. Bottom: Analogue results of the measurements conducted with* ***207****.*

Pure ctDNA forms particles with an average diameter (volume-weighted) of 40 nm. The addition of 0.1 eq of **207** leads to the formation of larger aggregates of 450 nm diameter besides the original ones. At a ratio of 0.3 the larger aggregates are predominant. Finally, at *r* = 0.4 even larger aggregates with diameters above 1 µm are formed. Addition of more ligand leads to precipitation under the conditions used for these experiments. The addition of **196**

leads to a similar concentration dependent aggregation, however higher ratios (r = 1.0) are needed to result in the large aggregates (> 1 µm).

Table 17. *Results of AFM and DLS measurements for the complex formation between* ***196*** *and* ***207*** *with ctDNA.*

Ligand	r[a]	AFM[b] [nm]	DLS Size Distribution by Volume[c] [nm]	DLS Size Distribution by Numbers[c] [nm]	DLS Z-average [nm] / PDI
196	0	H[d] = 0.5, 1.0, 1.5 W[d] = 31.0 ± 5.4 L[d] = 460 ± 175	39.4 (98 %)	22.3 (100 %)	73.4 / 0.608
	0.1	-	31.5 (96 %)	20.00 (100 %)	77.0 / 0.479
	0.5	H = 0.5, 1.0, 1.5 D[d] = 150 ± 30	26.8 (76 %) / 128.7 (24 %)	24.2 (100 %)	113.5 / 0.247
	1.0	H = 1.1 ± 0.3	1,529 (98 %)	1,414 (100 %)	2,222 / 0.392
	1.5	-	-[e]	-[e]	-[e]
	3.1	H = 16.2 ± 4.9 D = 90 ± 30	-	-	-
207	0.1	-	50.26 (41 %) / 451.2 (59 %)	45.9 (99 %) / 307.4 (1 %)	278.0 / 0.333
	0.3	H = 1.5 ± 0.4 D = 500 ± 100	397.6 (91 %)	211.1 (100 %)	284.0 / 0.303
	0.4	H_1 = 14.3 ± 4.2 W_1 = 700 ± 350; H_2 = 1.3 ± 0.2 D_2 = 4,000-9,000	322.2 (2 %) / 1,465 (98 %)	269.1 (56 %) / 1,174 (44 %)	1,243 / 0.293
	0.5	H = 17.5 ± 5.2 D = 350 ± 100	-[e]	-[e]	-[e]
	1.5	H = 17.0 ± 6.0 D = 150 ± 25	-	-	-

a) r = [compound] / [ctDNA]; b) Distances were measured at half height; c) Hydrodynamic diameter (2 × R_H) and percentage of total volume/numbers; d) H = height, W = width, L = length, D = diameter; e) Precipitation.

The results obtained by DLS were verified with the help of atomic force microscopy. AFM can be used to study the structure of materials deposited on a surface with resolutions up to 1000 times above the optical diffraction limit. The surface is scanned with a cantilever featuring a sharp tip with a radius of curvature on the order of nanometers, which is oscillated near its fundamental resonance frequency with an amplitude of 100-200 nm in tapping mode.[250] When tapping onto the surface the oscillation is decreased due to repulsion according to *Hooke´*s law. This deflection is measured by laser spot reflection from the top surface of the cantilever into the detector (e.g. an array of photodiodes). A feedback mechanism constantly adjusts the distance between tip and surface. For studying the structural changes upon complexation with **196** and **207**, untreated ctDNA (16 µM) was dissolved in water containing 263 µM magnesium chloride with different ratios of the gene carriers. The $MgCl_2$ facilitates the deposition of the negatively charged DNA onto the negatively charged mica surface. Furthermore, this way it can be assured that any observed structural change can be attributed to the addition of the ligands and not to changes of the surface potential. The so obtained solutions were spin coated onto freshly cleaved mica surface and analyzed in tapping mode. The resulting images are depicted in Figure 128 and the results are listed in Table 17.

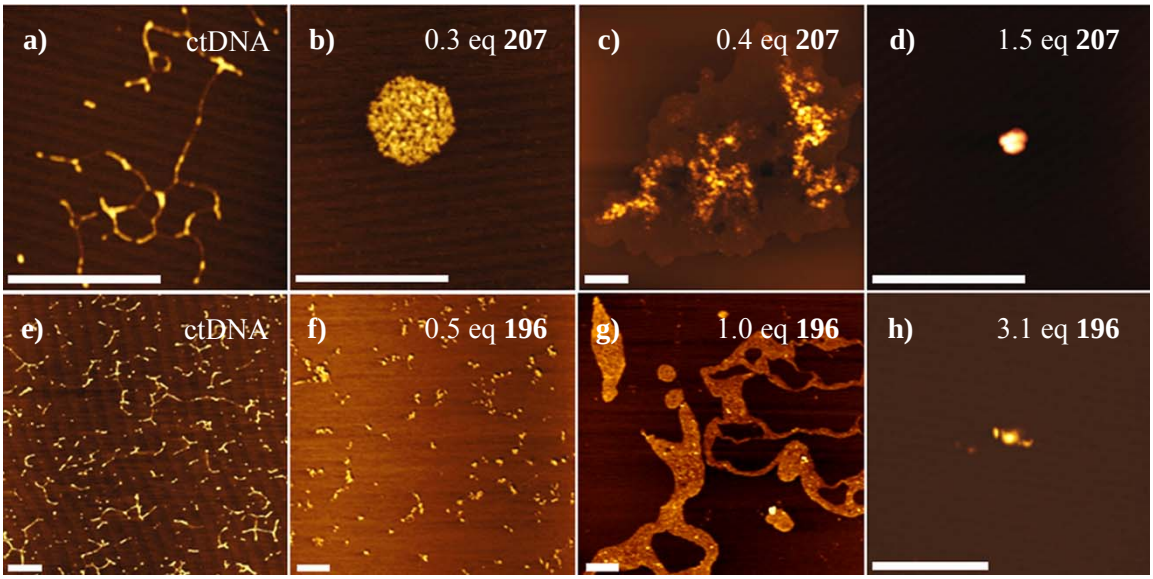


Figure 128. *AFM images were recorded in tapping mode. Mixtures of* ***207*** *(a-d) and* ***196*** *(e-h), ctDNA (16 µM) and $MgCl_2$ (263 µM) were spin coated onto freshly cleaved mica surface. a) Pure ctDNA; b) [**207**] = 5 µM (0.3 eq); c) [**207**] = 6.5 µM (0.4 eq); d) [**207**] = 25 µM (1.5 eq); e) pure ctDNA; f) [**196**] = 8 µM (0.5 eq); g) [**196**] = 16 µM (1.0 eq); h) [**196**] = 50 µM (3.1 eq). Scale bar equates to 1 µm in each picture.*

Pure ctDNA (a, e) forms strand-like, elongated structures with some crosslinking. Heights range from 0.5 over 1.0 to 1.5 nm and the average width is 31.0 ± 5.4 nm, indicating a mixture of DNA double-strands and aggregates thereof (double-strand height: 0.55 ± 0.15 nm, width: 12 ± 3 nm).[251] Addition of 0.3 equivalents of **207** led to the formation of circular aggregates of similar height (1.5 ± 0.4 nm) but with a diameter of 500 ± 100 nm (b). An increase of the ratio to 0.4 led to the formation of even larger aggregates now also with a significantly increased height of 14.3 ± 4.2 nm and a width of 700 ± 350 nm (c). Beneath these high aggregates a very large layer of up to 9 µm width and a height of 1.5 nm was formed. Thus, the large aggregate is probably an assembly of the former circular aggregates. When further increasing the ratio to r = 1.5, the aggregates become much more compact (d). Very tight, circular aggregates with a diameter of ca. 150 ± 25 nm are formed (height = 17.0 ± 6.0 nm). For **196** a similar concentration dependent condensation can be observed, although at higher ratios. At r = 0.5 the elongated strands are attracted to each other, forming loose aggregates of 0.5 to 1.5 nm height and an average diameter of 150 ± 30 nm. The addition of another 0.5 eq leads to the formation of large structures of several micrometer lengths and varying width. Finally, at a ratio of r = 3.1 compact and tight aggregates of 16.2 ± 4.9 nm and a diameter of 90 ± 30 nm are formed. The addition of higher amounts of both **196** and **207** led to precipitation under the conditions used for AFM measurements. All results obtained from the surface-based atomic force microscopy experiments are in excellent agreement with the solution phase DLS measurements. From this finding and from the ligand/DNA ratio of r = 8, which has been utilized for the transfection experiments it can be concluded that most likely even more compact and denser multimolecular complexes are present under transfection conditions resembling the polyplexes which form for example upon complex formation between nucleic acids and the cationic polymer PEI. In analogy to polyplexes and with regard to the large excess of positively charged gene carriers the net charge of these aggregates should now be positive. Hence, the compacted DNA is attracted to the negatively charged

surface of the cell instead of being repulsed as would be the case for pure DNA. Cellular uptake of particles of this size most likely takes place via an endocytotic pathway.[136]

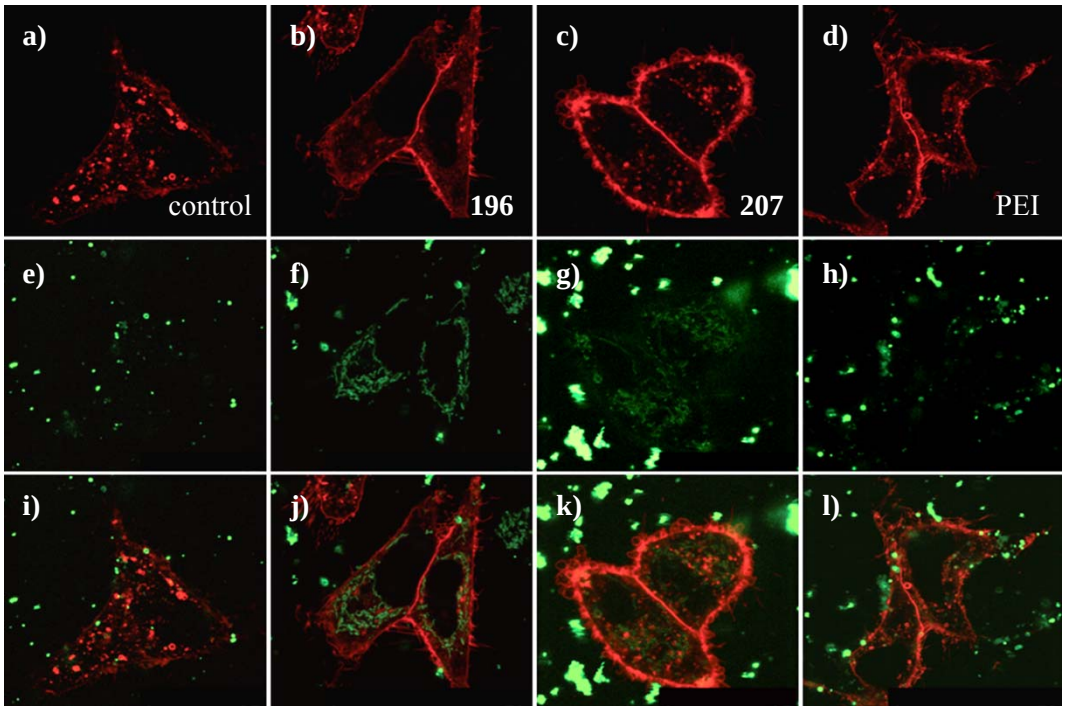


Figure 129. *Confocal microscopy images from live cell co-localization experiments for DNA uptake in the absence (control) or presence of* ***196***, ***207****, or PEI as non-viral vector. a-d) Membrane and endosomal structures stained with a RFP; e-h) DNA labeled with a green fluorescence dye; i-l) overlay of pictures above.*

In order to study in more detail the uptake mechanism into the cell co-localization studies with HeLa cells were carried out by *Gutschmidt* with the help of confocal fluorescence microscopy. First, membranes and endosomes were specifically marked by transfecting the cells with Lamp1-RFP plasmid DNA encoding for the lysosome associated membrane protein 1 (Lamp1) tagged with a red fluorescence protein (RFP). 24 h later, the same cells were transfected with fluorescence marked, fragmented DNA (2 µg) with **196**, **207**, or PEI (0.24 mM) and without carrier as control experiment. The DNA was obtained by treating pF143-GFP plasmid with a labeling kit as described by the manufacturer (*PromoKine*) and consequent purification. After another 4 h the live cells were monitored with the help of confocal microscopy as depicted in Figure 129 (from left to right: control, **196**, **207**, and PEI). Endosomes and cell membranes fluoresce in red (a-d) and the DNA in green (e-h). By overlaying the images random pattering can be observed for the control experiment, because no DNA is taken up into the cell (i). In the presence of gene carriers the labeled DNA is clearly localized inside the cell. All three ligands are able to shuttle ds-DNA into the cell. For PEI (l) co-localization mainly takes place in membranes. For **196** (j) and **207** (k) the labeled polynucleotide is co-localized within endosomes to a degree, the DNA is however mainly localized within continuous, membranous structures within the cytosol, which resemble the endoplasmic reticulum-transgolgi network. This structuring is especially pronounced for the two armed **196**, less for trivalent **207** and not at all for PEI. In conclusion, although all three carriers are able to transport the DNA into the cell only **207** and PEI enable transfection, but

not **196**. Thus, the difference between **196** and **207** does not arise from their DNA-shuttling ability but from differences in liberating the DNA once inside of the cell.

To investigate in more detail the differences in endosomal release between **196** and **207**, *Gutschmidt* carried out co-transfection experiments with chloroquine (**219**, 25 µM), a substance which is well known to facilitate endosomal escape.[252] After endocytosis the pH inside the endosomal vesicles is constantly reduced from pH 7.2 to 4-5 by accumulation of protons by the specific ATPase, which ultimately leads to degradation of the enclosed material in the lysosomes.[142] Weakly basic substances like chloroquine are protonated in the process thus acting as a buffer inside the endosome, counterbalancing the lowering of pH (cf. proton sponge effect in chapter 2.4.4). Consequently more protons, accompanying counter-ions and water are pumped into the vesicle further increasing the osmotic pressure. Ultimately the vesicle ruptures and its cargo is released into the cytosol before degradation may occur.

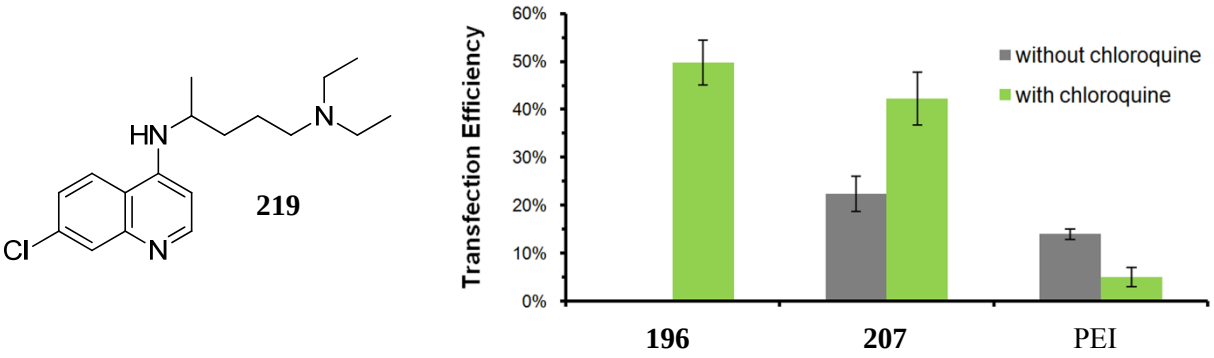


Figure 130. *Transfection efficiency of **196**, **207**, and PEI with and without addition of chloroquine (**219**).*

As illustrated in Figure 130 transfection efficiencies for **196** and **207** were increased upon co-transfection with chloroquine. This is another strong indication for endosomal uptake. PEI on the other hand does not profit from chloroquine, because it is an efficient proton sponge on its own.[144] Most important, the effect of **219** on the two-armed carrier **196** is huge (from < 1 to 50 %) but only modest for tripodal **207** (from 22 to 42 %). This result suggests an almost complete lack of buffering capacity for **196**. In contrast, **207** seems to be a capable proton sponge. The explanation for this difference most likely originates from the number of weakly basic sites in the ligands. The pK_A value for the GCP moiety is higher in **196** than in **207** (7.00 vs. 6.50). As a consequence these groups are only partially protonated at physiological pH. At the pH of the transfection experiments (7.2) **196** carries 3 and **207** between 3 and 4 positive charges. Inside the endosome full protonation of both gene carriers takes place resulting in 4 positive charges for **196** and 6 for **207**. Hence the three-armed ligand has more than double the buffer capacity than its divalent analogue (charge difference: +1.2 vs. +2.5) and is therefore able to act as proton sponge and thus allows for an early endosomal escape and efficient transfection, while **196** is not.

Figure 131 summarizes the uptake mechanism mediated by the tripodal gene carrier **207**. First, the ligand binds to the plasmid DNA with high binding constant ($K = 3 \times 10^7$ M^{-1}) and thereby the DNA is condensed into very tight, positively aggregates which are attracted to the negatively charged cell surface and of ideal size for endosomal uptake into the cell. Once inside, the GCP units of **207** are protonated and due to the proton sponge effect the cargo is

released into the cytosol, transported to the nucleus, integrated into the cell own machinery for protein biosynthesis, which ultimately leads to expression of GFP. It could be shown that **207** is a very efficient, new type of small molecule non-viral vector. It combines the advantages of cell penetrating peptides (CPPs)[160] like low cytotoxicity and ease of accessibility by means of SPPS, but does not suffer from their inconveniences, like difficulties with endosomal escape, because it is able to act as proton sponge similar to the polymer PEI.[138] In comparison to CPPs which normally comprise 10 to 30 amino acids it is very small and merely features three to four charges at physiological pH. In contrast to that CPPs like oligoarginine need at least six charges for efficient cellular uptake.[162] Therefore, this first generation of novel chemical vectors offers great potential for the delivery of nucleic acids.

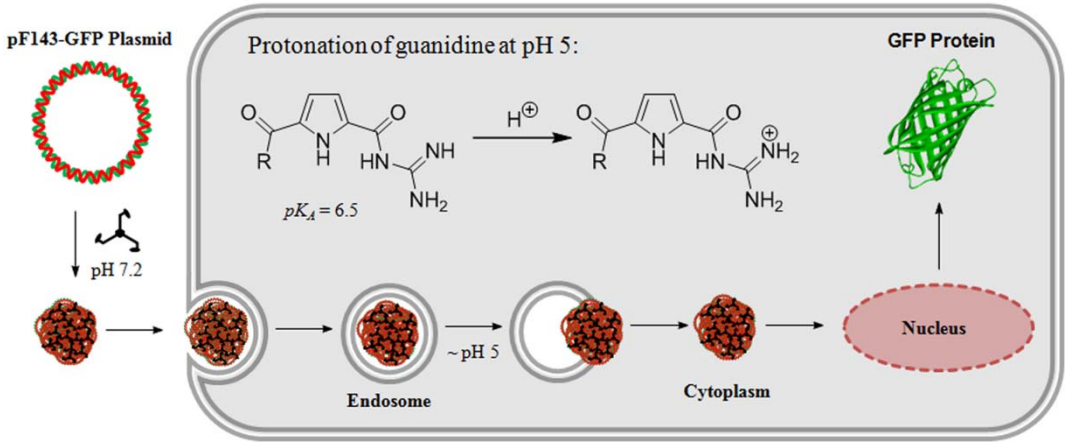
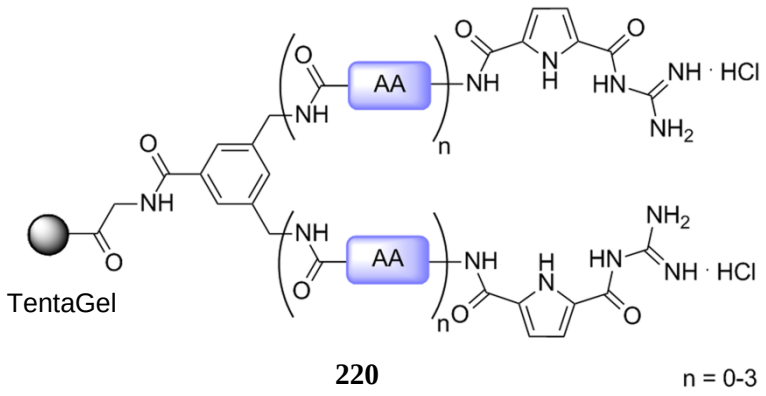


Figure 131. *The **207**-mediated uptake mechanism of plasmid DNA into the cell starts by condensing the DNA into tightly packed multimolecular aggregates, attraction to the cell surface via charge-charge interaction, uptake via endocytosis and subsequent endosomal escape with the help of the proton sponge effect. Ultimately, the plasmid encoded GFP is expressed by the cell after integration into the genome located in the nucleus.*

In conclusion, the two- (**196**, **206**) and three-armed (**207**) were successfully synthesized by means of microwave-assisted SPPS and convergent synthesis in solution. It could be shown that these substances are very efficient in binding to nucleic acids. The divalent ligands bind with binding constants higher than 10^6 M^{-1}, and the tripodal analogue binds even stronger by one order of magnitude (> 10^7 M^{-1}). Although there is no pronounced selectivity for a specific type of polynucleotide, these molecules form well-defined complexes. For B-type nucleic acids minor groove binding and for A-type polynucleotides major groove binding was observed. Furthermore, it could be shown that the ligands are able to condense linear and circular DNA in a concentration dependent manner to very densely packed aggregates and to shuttle these through the membrane of human cells via an endocytotic pathway. Only the trivalent carrier is then able to enable escape from the endosome with the help of the proton sponge effect leading to efficient transfection even surpassing the commercially available gold standard for transfection, PEI, in both the efficiency as well as the compatibility with biological systems due to a lower cytotoxicity. Because the arms of the di- and trivalent gene carriers are of peptidic nature further improving the properties of these ligands with regard to their affinity and ability to transfect cells should easily be possible with the help of combinatorial chemistry. In the next chapter such a combinatorial approach will be presented.

4.5 COMBINATORIAL DNA-LIGAND LIBRARY

Peptide-based combinatorial chemistry has until now only rarely been used for the development of small molecule DNA binders, which are not derived from proteins. Often, known natural products[127] or synthetic intercalators[126] and groove binders[124] were modified, but these examples were only moderately successful in finding more advanced ligand systems which bind with higher affinity than their parent compounds (cf. combinatorial approaches in chapter 2.3.5). For all these systems the reported binding constants range from 10^6 to 10^7 M^{-1}. These libraries were mainly of large and random nature and only a fraction of their encoded information was thus accessible: Only hit structures were characterized during the screening – the majority of library members was ignored, which led to the loss of information. This is why in the framework of this dissertation a rather small, but focused library based on divalent ligands like **196** and **206** was prepared instead, which allowed for the characterization of all possible compound-DNA interactions. The parent systems are already able to bind to nucleic acids with binding constants > 10^6 M^{-1} although their amino acid sequence was chosen for the targeting of nucleotides and not for polynucleotide binding (see chapter 4.3). Therefore, it was assumed that there is great potential for further optimizing the DNA-binding properties of these molecules by varying their amino acid sequence. Figure 132 depicts the design of the combinatorial library **220**. There are three combinatorial positions within the two symmetric arms. For the combinatorial variation aromatic (Phe, Tyr, Trp) and basic (Lys, Arg) amino acids were utilized, just like in the original ligands **196** and **206**. As a control an acidic amino acid (Glu) was also implemented. All library members feature a GCP unit as headgroup of each arm. Furthermore, sublibraries with only two, one or respectively no combinatorial position were synthesized. This way the influence of the length of the side arms on DNA binding could be studied. Furthermore, due to the choice of combinatorial amino acid building blocks the sublibrary with two amino acid positions also contained the already characterized ligands **196** and **206**, which could thus be directly compared. In total the library comprised 259 potential nucleic acid ligands ($6^3 + 6^2 + 6 + 1$) of varying length and amino acid composition.



AA = Phe, Tyr, Trp, Lys, Arg, Glu

216	= 6 × 6 × 6	(n = 3)
36	= 6 × 6	(n = 2)
6		(n = 1)
1		(n = 0)
259 ligands		

Figure 132. *The library **220** consists of 259 symmetric, divalent DNA ligands with GCP headgroups in each side arm, which contain three, two, one, or no combinatorial positions (Phe, Tyr, Trp, Lys, Arg, and Glu).*

Since the screening of the library was supposed to be carried out on-bead, a test ligand (**221**, Figure 133) with the same sequence as **196** (Phe-Lys-GCP) was synthesized first to establish

the screening method. In contrary to the polystyrene (PS)-based solid supports already described within this work like MBHA, Rink or Wang resin, which were used for the synthesis of compounds which were cleaved from the resin, the preparation of the test ligand was carried out on TentaGel.[253] This resin also consists of a PS core but is modified with polyethylene glycol (PEG) chains, which in total comprise 70 % of each bead. The advantage of this resin is that it has on the one hand good swelling properties in organic solvents like DCM and DMF, where it swells to 3 and respectively 3.2 times its original size. This swelling is necessary to ensure accessibility of the reactive sites. On the other hand it also features excellent swelling properties in aqueous medium (swelling factor: 2.5), in contrast to PS resins, and is thus compatible with binding studies of biological macromolecules in aqueous solution. The synthesis was carried out according to standard Fmoc SPPS as described for **196** in chapter 4.3.2.[178] The only difference was that the coupling and deprotection steps were not carried out under microwave irradiation but at room temperature. Therefore longer reaction times (8-12 h for coupling and 2 × 20 min for deprotection steps) were applied. All coupling steps were repeated at least twice to ensure complete conversion, or more often if indicated by the Kaiser test.[181] The Boc protecting groups were cleaved by treating the beads with 50 % TFA in DCM containing TIS and water (each 2.5 %) as nucleophilic scavengers.[182] The TentaGel linker is not cleavable and therefore **221** remains on the resin under deprotection conditions. After thoroughly washing and drying of the beads, the test ligand was then used for the optimization of the on-bead binding assay.

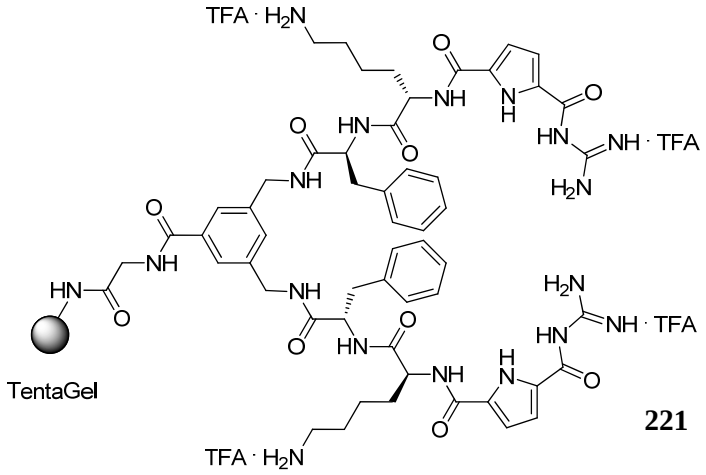


Figure 133. *Resin-bound analogue (**221**) of the symmetric divalent DNA ligand **196** .*

Boger reported the screening of a solution-based nucleic acid binder library with the help of an ethidium bromide displacement assay.[127] As described in chapter 4.4.3, the incubation of EB with a nucleic acid increases its inherent fluorescence by a factor of 10-100. Addition of a competing ligand partially displaces the intercalator thus reducing the overall fluorescence. By incubating the EB/DNA complex with resin-bound ligands, it should be possible to compare the individual library members for their ability to displace EB from the DNA. In order to optimize the conditions for such experiments and to ensure that there is no unspecific interaction between resin and nucleic acid various tests were conducted. Ethidium bromide (0.75 µM, 1 eq) was incubated in sodium cacodylate buffered water (0.01 M) at neutral pH with ctDNA (3 µM, 4 eq) in a 96-well plate (300 µL) and after 5 min incubation time the

fluorescence emission upon irradiation at 520 nm between 560 and 650 nm was recorded. Different amounts of resin bound ligand and acetylated TentaGel resin (25, 50, 100 eq based on the resin loading of 0.25 mmol/g) were then added with the help of an isopycnic solution.[254] To obtain such a solution of equal density of the swollen resin and solvent, 5 mg beads were dissolved in 400 µL of a 4:1 mixture of 1,2-dichloroethane (DCE) and isopropanol (*i*Prop). The advantage of an isopycnic solution is, that small amounts of resin do not need to be weighed in, but can be titrated thus minimizing the error. Since the so obtained solution was not perfectly isopycnic the suspension was vortexed prior to each addition. The solutions were incubated by gently shaking the well plates on a shaker and the fluorescence emission was measured again after 44 h. Each experiment was carried out in duplicate. The biggest changes in fluorescence intensity were observed for the wells containing 100 equivalents of resin bound ligand. After 44 h the fluorescence intensity was at 76 % of the original one. The acetylated TentaGel on the other hand did hardly influence the EB/DNA complex at all tested concentrations. At 100 eq for instance the fluorescence intensity after 44 h was at 94 %. This was regarded as proof-of-principle for the feasibility of an EB displacement-based on-bead screening without the need for labeling of nucleic acid or ligand. Similar tests with $p(dAdT)_2$ and the minor groove binder DAPI revealed ideal screening conditions for [DAPI] = 0.6 µM (1 eq), [$p(dAdT)_2$] = 3 µM (5 eq) and 100 eq of solid-bound ligand. In comparison to the solution-based EB-based binding assays a large excess of DNA ligand and long incubation times are necessary for displacing the intercalator from its binding sites. This is an effect which can commonly be observed for on-bead screenings and which can be attributed to the hyperentropic efficacy: Resin-bound substances have to be regarded as highly dilute even at concentrations similar to analogue solution phase conditions.[255] Therefore, higher amounts of resin-bound ligands and longer reaction times are necessary for the formation of non-covalent ligand/DNA complexes. Furthermore, ctDNA (> 13,000 kDa) and $p(dAdT)_2$ (> 5,000 kDa) are very large macromolecule and it is not very likely that they are able to fully penetrate into the resin. Therefore, not all resin-bound ligands are able to interact with the nucleic acid. Still, with both the synthesis and a screening method at hand, the preparation of the ligand library could now be accomplished.

4.5.1 Synthesis of the Library

The synthesis of the library was carried out on TentaGel S NH_2 (loading: 0.32 mmol/g) by means of split & mix synthesis.[256] During the split-step the beads are divided into parts of equal amount according to the number of combinatorial building blocks at a given position of the growing peptide chain. In Figure 134 (left) the three building blocks are visualized by colored circles. After the reaction, the solid-supports are then combined again (mix-step) and divided again during the next split step. When this procedure is repeated for example three times with three building blocks in each step, the resulting number of library members is 27 already. Thus, by synthesizing the building blocks in parallel, the split & mix method allows rapid generation of libraries.

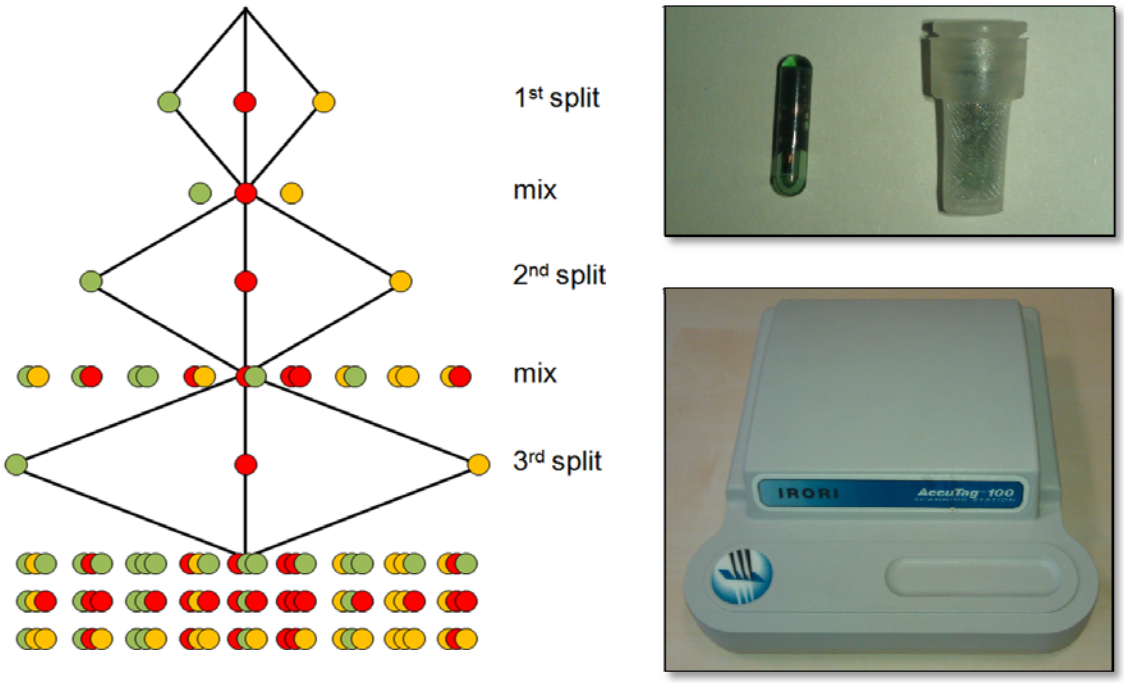


Figure 134. *Split & mix synthesis allows rapid preparation of combinatorial libraries (left). The IRORI radio-frequency technology (top: microchip and MicroKan, bottom: scanning station) helps to plan the synthesis and the identify each member of one-bead-one-compound libraries (right)*

Additionally, the synthesis of library **220** was carried out with the help of the IRORI™ radiofrequency technology.[257] Therefore an isopycnic solution of TentaGel (35 mg per 400 µL hexane/DCM = 1/8) was prepared and approximately 30 mg of the solid-support was transferred into MikroKans. These sealable polyethylene compartments contain the resin but are penetrable for the solvent and dissolved reactants through their synthetic membranes. To each Kan a radiofrequency microchip was added before sealing. Each compartment will contain exactly one library member, which can be characterized at each point of the synthesis and after preparation of the library with the help of its unique radiofrequency signal (a binary eleven-digit code), a scanning station and the corresponding IRORI Accu Tag Synthesis software, which also allows planning of the split & mix steps. After having filled the MikroKans and having identified each microchip signature at the scanning station, the library was prepared according to the scheme in Figure 135.

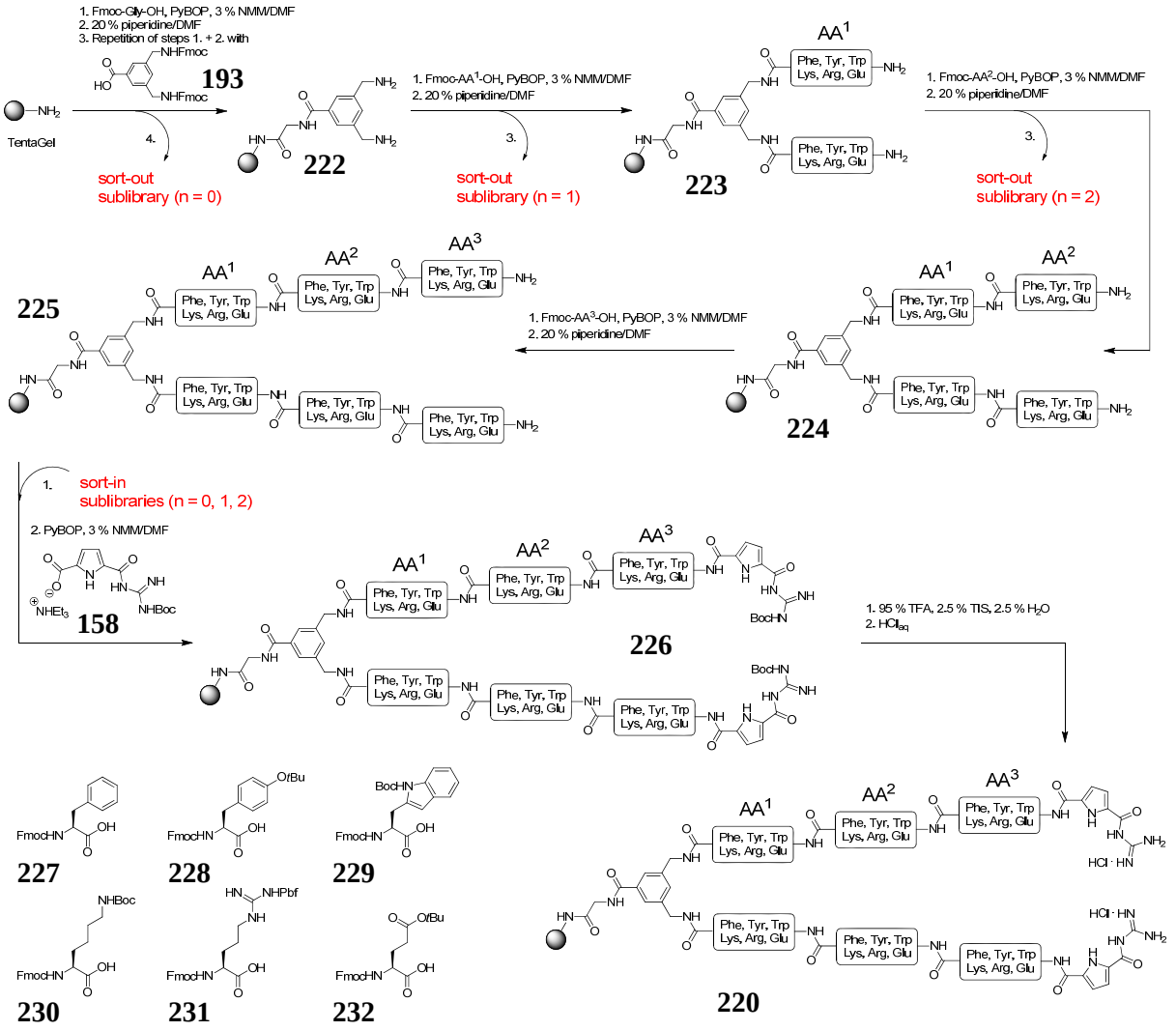


Figure 135. *The library **220** was prepared according to standard Fmoc SPPS in three split & mix synthetic steps with the six building blocks **227-232** to give a total of $6^3 + 6^2 + 6 + 1 = 259$ library members.*

The synthesis was carried out according to standard Fmoc synthesis with PyBOP (3 eq) as coupling agent in 3 % NMM/DMF under argon atmosphere.[178] In comparison to simple SPPS of singular compounds all reaction times were prolongated. The coupling of Fmoc-Gly-OH (3 eq) was repeated three times (20 h) to ensure complete conversion of the TentaGel amine groups. Fmoc deprotection was achieved by treating the solid-support with 20 % piperidine (2 × 1 h). After each coupling and deprotection step the resin was thoroughly washed and two MikroKans per compartment were opened and tested for complete conversion by means of a Kaiser test.[181] All coupling steps were carried out twice and repeated if necessary. The aromatic template **193** was reacted accordingly and afterwards the three split & mix steps were carried out by dividing the resin into six compartments according to the instructions of the IRORI software, which were then individually reacted with 6 eq of PyBOP and Fmoc-Phe-OH (**227**), Fmoc-Tyr(*t*Bu)-OH (**228**), Fmoc-Trp(Boc)-OH (**229**), Fmoc-Lys(Boc)-OH (**230**), Fmoc-Arg(Pbf)-OH (**231**), or Fmoc-Glu(O*t*Bu)-OH (**232**). Especially tryptophan and arginine are known to be problematic during SPPS. Arginine is prone to addition of electrophiles to its nucleophilic guanidinium group, e.g. acylation during coupling steps which subsequently leads to ornithine formation through intramolecular decomposition during

deprotection.[258] Because simple Boc protecting groups are not sufficiently blocking this unwanted side reaction,[259] arene sulfonyl-based protecting groups like 2,2,4,6,7-pentamethyl-dihydrobenzofuran-5-sulfonyl (Pbf) have to be used instead.[260] In order to prevent the transfer of these groups to Trp during TFA mediated deprotection and to prevent oxidation of the labile indole ring,[261] the Trp nitrogen must also be protected with a Boc group and nucleophilic scavengers have to be used during the acidic deprotection step.[262] For the preparation of the sublibraries with only two, one or no combinatorial positions, the corresponding MicroKans were sorted out according to the IRORI software before the first (n = 0), after the first (n = 1) or after the second (n = 2) split & mix step. After completion of all split & mix steps, the Kans were reunited and reacted with Boc protected GCP (**158**). This last coupling step was repeated three times. Due to the rather poor solubility of Boc-GCP-OH it was heated in 3 % NMM/DMF for 20 min together with PyBOP prior to addition to the MikroKans before the last repetition of the coupling thus ensuring the formation of the active ester which could be witnessed by the formation of a red solution. The resin was thoroughly washed (DCM, MeOH) and dried in high vacuum. The protecting groups were removed by treating the resin with 95 % TFA containing 2.5 % TIS and 2.5 % water as nucleophilic scavengers (2 × 2 h).[182] The resin was washed with organic solvents (DCM, MeOH) to remove excess TFA, and then with 0.1 N hydrochloric acid and water to exchange all trifluoroacetate counterions for chlorides and again with MeOH and DCM before drying in high vacuum for 24 h. The library (**220**) was now ready for screening with nucleic acids.

4.5.2 On-Bead Screening with p(dAdT)$_2$

The library was then first screened against the nucleic acid p(dAdT)$_2$, which forms a very well-defined and narrow minor groove by means of a DAPI displacement assay. All measurements were carried out in cacodylate buffer (0.01 M) at neutral pH. First, the fluorescence emission of DAPI (0.6 µM, 1 eq) upon irradiation at 358 nm was recorded between 455 and 465 nm in 96-well plates (300 µL). Then the polynucleotide was added (3 µM, 5 eq) and after shaking the mixture for 30 min, the resulting increased fluorescence emission was measured. Finally, isopycnic solutions (DCE/*i*Prop = 4/1, 5 mg resin per 400 µL) of the individual library members were added (300 µM, 100 eq). The suspensions were vortexed prior to each addition. The screening was carried out with three wells per library member/DNA combination and three blank samples per plate containing only p(dAdT)$_2$ and DAPI without resin for normalization of different plates. The well-plates were then incubated on a shaker and the resulting fluorescence emission was measured after 20, 40, 60, and 80 h. Surprisingly, after 20 h of incubation with the resin, the fluorescence intensity was higher instead of lower. This effect can be attributed to insufficient mixing of the nucleic acid-DAPI mixture prior to addition of the resin. Obviously shaking the plates is not enough to guarantee for homogenous solutions. In future attempts, the solutions should additionally be mixed with the help of a pipette. Nevertheless, the blank samples allowed normalization of the highest fluorescence emission to 100 % and therefore accurate processing and analysis of the library. Figure 136 depicts the normalized, relative fluorescence emissions of all library members after correction for DAPI´s own emission. Low intensities signify efficient displacement of DAPI and therefore high binding affinities.

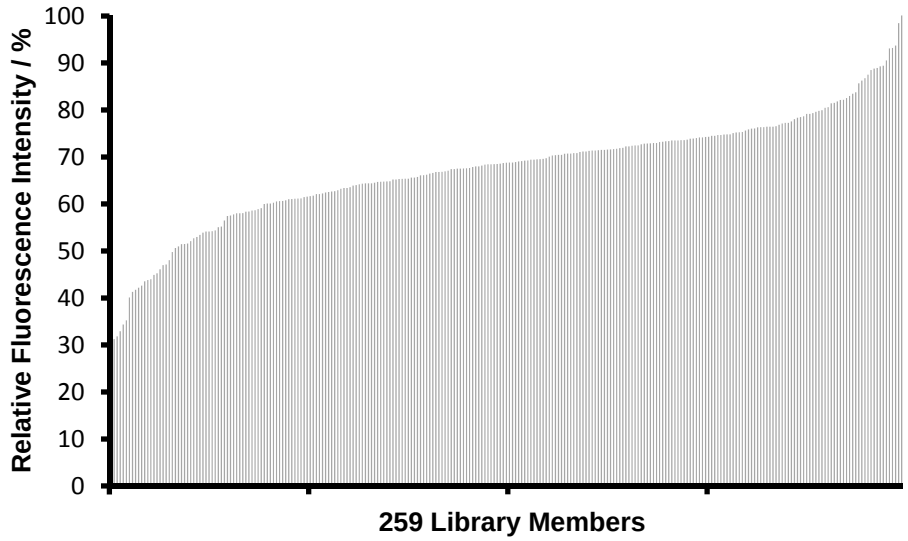


Figure 136. *Relative, normalized fluorescence emissions as obtained by a DAPI displacement assay from the B-DNA p(dAdT)$_2$ with the members of library **221** after 80 h incubation in cacodylate buffer (0.01 M) at pH 7.*

As shown in Figure 136 there is a good distribution of the fluorescence intensities between 26 and 100 % with a broad middle part and only few very good (left) and very bad (Figure 136, right) performing library members. Since the two divalent ligands **196** and **206** are also part of the library already at this time a preliminary comparison to the results in solution could be drawn. **196** (Phe-Lys-GCP) is ranked #194 with 74 % remaining DAPI and **206** (Lys-Phe-

GCP) ranks #72 with 63 %. This is in excellent agreement with the IC_{50} values obtained from the DAPI displacement in solution with IC_{50} = 52.2 for **196** and 15.9 for **206** (cf. chapter 4.4.4). Furthermore, it immediately becomes clear that the two divalent ligands are only representatives of the broad middle part of the library although their binding constants are already in the range > 10^6 M^{-1}. Hence, the top ranked members of the library **221** should feature even higher binding constants and thus excellent affinity for $p(dAdT)_2$. Table 18 lists the Top10 compounds. A complete list can be found in appendix C.7.

Table 18. *Top10 ranked members of library **220** as obtained from a DAPI displacement assay with $p(dAdT)_2$. The charges are estimates for pH 7: full protonation of Arg, Lys and 50 % protonation of the GCP group.*

#	AA^1	AA^2	AA^3	Charge	%
1	Trp	Lys	Arg	+5	26
2		Arg	Lys	+5	31
3	Arg	Arg	Lys	+7	32
4	Trp	Arg	Lys	+5	33
5	Arg	Lys	Lys	+7	34
6	Arg	Arg	Arg	+7	35
7	Tyr	Arg	Arg	+5	40
8	Arg	Arg	Tyr	+5	41
9	Phe	Tyr	Lys	+3	42
10		Arg	Arg	+5	42

The positions of the amino acids are numbered AA^1-AA^3 from C- to N-terminus with AA^3 always being the amino acid next to the GCP unit. Accordingly, the combinatorial positions of ligands which only feature two amino acids are numbered AA^2 and AA^3. The Top10 overall best performing DNA binder features the sequence Trp-Lys-Arg-CBS. All sequences feature a high amount of positive charges (+3 to +7). The charges were thereby estimated for pH 7 by assuming full protonation of amine and guanine groups, complete deprotonation of carboxylic acids and 50 % protonation of the GCP group. However, although undoubtedly charge plays a decisive role for high affinity to the negatively charged nucleic acid, there seem to be additional effects concerning the binding properties of the library members. For instance, at position AA^1 half of the compounds within the Top10 feature an aromatic amino acid, the other half arginine. Lysine on the other hand is not present at this position. In general, arginine is the most prominent amino acid in the Top10 with 15 occurrences out of a total of 28, followed by lysine (7) and aromatic amino acids (6). As expected, no glutamic acid is present, which would be negatively charged at this pH and therefore hampering the affinity due to charge-charge repulsion with the ribophosphate DNA backbone. Eight compounds feature three amino acids per arm, but there are also two ligands with just two amino acids. Both of them comprise two positively charged amino acids within each arm. Highly charged ligands with just two amino acids seem to be able to compete with their larger and higher charged analogues.

To get a more general impression of the amino acid distribution, the library was split into four compartments according to the results of the screening: Top 25 %, Top 25-50 %, Top 50-75 %, and Top 75-100 %. Then an analysis of the percentaged occurrence of amino acids within each compartment was carried out. The result is illustrated in Figure 137. As expected,

the best ligands (Top 25 %) are dominated by the basic amino acids lysine and arginine, which in total comprise over 50 % of all amino acids within this compartment. On the runner-up positions follow the aromatic amino acids and on the last spot there is glutamic acid. In the middle part of the library (Top 25-75 %) there is no clear correlation. DNA-binders within this region are of mixed sequence. The worst performing ligands (Top 75-100 %), feature an amino acid distribution antipodal to the best ones. The negatively charged Glu is most present, followed by aromatic amino acids. The positively charged basic amino acids are least present within this compartment. These results underline the general trend of higher DNA-affinity of higher charged compounds.

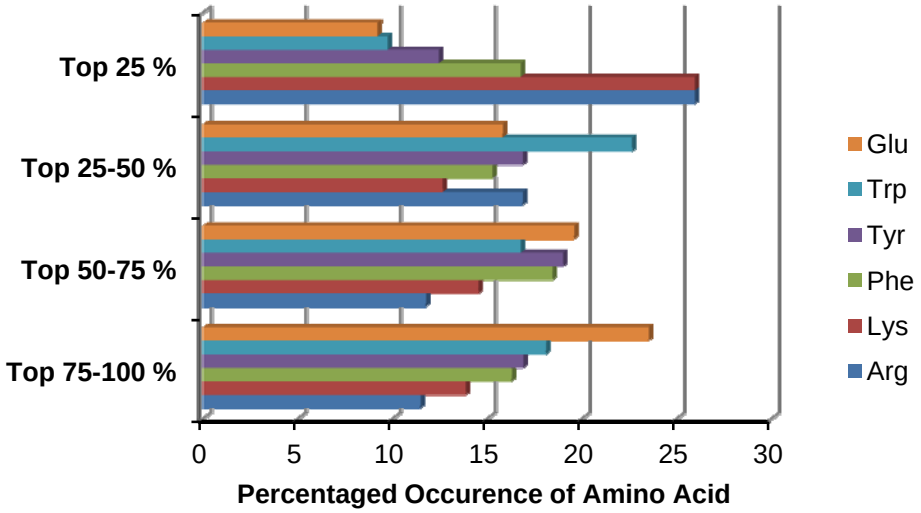


Figure 137. *Amino acid distribution in per cent within each of the four compartments of the library.*

A more detailed insight into the role of the amino acids can be obtained when analyzing the three individual combinatorial positions AA^1-AA^3 within the four compartments. The distribution of amino acids in per cent is depicted in Figure 138. At position AA^1 phenylalanine (26 %) and arginine (20 %) are preferred in the best performing quarter of the library. As already indicated by the Top10 compounds, the second positively charged building block, lysine, is underrepresented at this position with only 12 %, and is closely followed by the aromatic amino acids tyrosine (11 %) and tryptophan (9 %). Glutamic acid is present least (6 %) within the Top 25 % and does not seem to be tolerated at all in this position as can be seen in the huge amount of bad performing ligands (Top 75-100 %) with Glu in position AA^1. At the second combinatorial position, the situation is different. Here, Arg and Lys (28 and 35 %) are clearly favored. All other amino acids are not performing well. With respect to the other two positions, aromatic amino acids are tolerated worst at AA^2. The distribution of amino acids in the last position AA^3 is similar to AA^2, although aromatic amino acids seem to be better suited here than at AA^2. Hence, the positional analysis of the amino acid distribution has offered insight into some general trends. Arginine and lysine are overall preferred in the best performing ligands with the exception of sequences of the type Lys-AA^2-AA^3. Furthermore, sequences of the type AA^1-aromatic-AA^3 are not performing as good as comparable sequences with aromatic amino acids at other positions. At AA^1 (aromatic-AA^2-

AA^3) the implementation of aromatic amino acids even seems to be advantageous. This is in excellent agreement with the sequence of the first ranked compound, Trp-Lys-Arg-GCP.

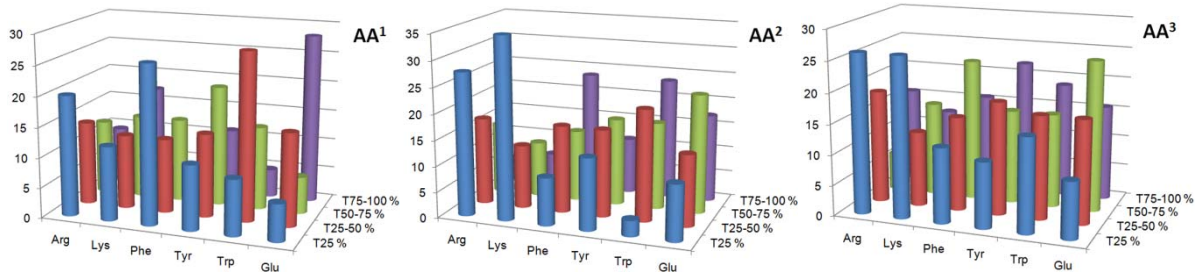


Figure 138. *Percentaged amino acid distribution at the three combinatorial positions AA^1-AA^3 within each of the four compartments of the library.*

Next, a closer look at the charge distribution was taken. By assuming pK_A values for arginine (12.5), lysine (10.5), glutamic acid (4.3), and the GCP unit (7.0),[11] the charges of the library members at neutral pH can be estimated. All Glu side chains are deprotonated, all Lys and Arg side chains protonated as well as half of the GCP groups. Therefore, there is a charge distribution in steps of two charges between the two extremes +7 and -5.

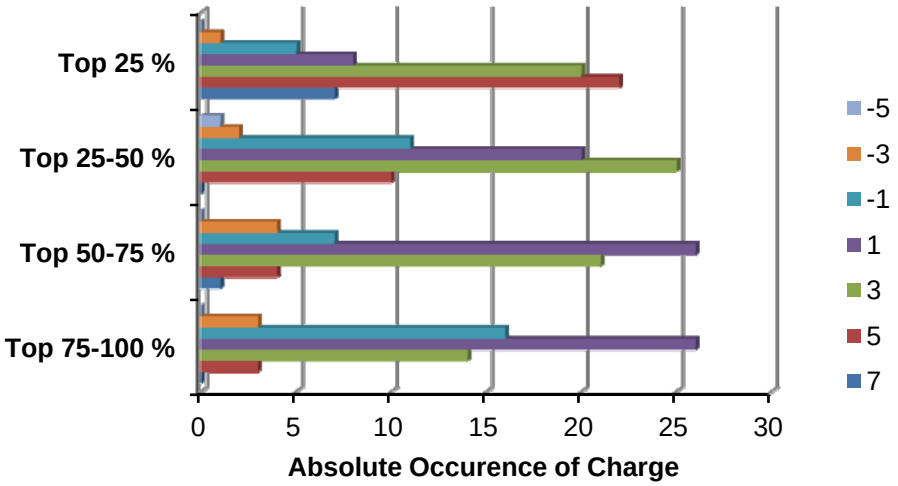


Figure 139. *Absolute charge distribution within each of the four compartments of the library.*

Figure 139 shows the absolute number of the different charges within the four compartments of the library. The highest positively charged compounds (+7, +5) are clustered within the Top 25 %. Lower positive charged compounds can be found in the broad middle region of the library. The worst performing compounds are those with net charges around zero (1, -1). Surprisingly, the higher negatively charged ligands (-3, -5) are not clustered in the low performing region. There are even some of them in the Top 25 %. These are all compounds with one or more glutamic acids combined with aromatic building blocks. Probably these compounds do interact directly with the positively charged DAPI and liberate the minor groove binder from the nucleic acid by binding to it rather than displacing it by binding to the nucleic acid. The second unexpected result is the presence of some highly charged species (+7, +5) within the low performing region (Top 50-100 %). A closer look at these substances

revealed common sequences of the type Lys-AA^2-AA^3 and AA^1-aromatic-AA^3. This is in good agreement with the observations of the position-dependent amino acid distribution, which also showed that these type of sequences are less well suited for displacing DAPI from $p(dAdT)_2$. Hence, besides the charge-driven interaction with the nucleic acid, the amino acid sequence does also seem to play a decisive role in binding to the B-DNA and displacing DAPI from the narrow minor groove. While the discrimination of Lys-AA^2-AA^3 sequences remains unclear, a possible explanation for the bad performance of AA^1-aromatic-AA^3 sequences might be that the ligands bind to the minor groove of the nucleic acid via a predominantly positive charged amino acid at position AA^3 and/or additional charge-charge or hydrophobic interactions of basic/aromatic amino acids at position AA^1. An aromatic moiety at position AA^2 would then also have to bind into the minor groove, which is difficult due to the bulkiness of the aromatic side chains in combination with the narrowness of the $p(dAdT)_2$ minor groove.

Finally, the influence of the length of the ligands on their binding properties towards the nucleic acid was analyzed. To compare ligands of increasing length, the percentage value obtained from the library screening for each ligand with n amino acids was corrected for the corresponding value of the ligand with the same amino acid at position AA^3 (or AA^2-AA^3). For example, the value for the ligand of the sequence Arg-Arg was subtracted from the percentage obtained for the corresponding DNA-binder Lys-Arg-Arg to get information on the effect of the additional amino acid, Lys. Finally, an average value for the addition of each amino acid building block was calculated for each sublibrary, e.g. for the step from sublibrary n = 1 to sublibrary n = 2.

The library member which features only two GCP units directly attached to the aromatic template without any amino acid building block (n = 0) is ranked #123 with 68 % remaining DAPI fluorescence. Addition of any one amino acid to the side chains (n = 1) leads to worse results (69-81 %). In contrast to that, addition of a second amino acid improves the performance of the ligands in all cases. Especially when adding Arg and Lys, on average 21 and respectively 16 % more DAPI are replaced in comparison to their corresponding DNA-binders with just one amino acid per arm. Addition of all other building blocks also led to modest improvement of the performance – on average 2 to 4 % more DAPI was displaced. It is noteworthy that even the addition of the negatively charged glutamic acid led to an improvement of nucleic acid binding. Obviously, two amino acids seem to be the minimum for efficient interaction with this kind of polynucleotide. A certain degree of flexibility seems to be important for the interaction. When further going from two to three amino acids (n = 3), the performance is further improved, although the overall effect is smaller. When adding Arg (6 %), Phe (5 %) or Trp (4 %) in position AA^1, more DAPI is replaced in comparison to their two amino acid analogues. Addition of Lys (0 %) and Tyr (1 %) does neither improve nor deteriorate the affinity by large. When adding glutamic acid however, 6 % less DAPI is replaced. This findings are again in good agreement with the positional analysis of the amino acid distribution, which resulted in the observation, that Glu is tolerated worst at position AA^1 and Lys performs surprisingly bad at this position.

In conclusion, the on-bead screening of the library for its binding affinity towards the B-type DNA $p(dAdT)_2$, which forms a very well-defined, narrow minor groove, could successfully be carried out with the help of a DAPI displacement assay. Besides the expected charge-

driven capability to bind to the nucleic acid, also several sequence-dependent influences on the binding affinity could be observed. For instance, sequences of the type Lys-AA^2-AA^3 and AA^1-aromatic-AA^2 seem to be less well suited for binding. In the latter case this can be attributed to the sterically difficult binding of the bulky aromatic side chains into the minor groove of $p(dAdT)_2$. An alternative explanation is the positive influence of a positively charged amino acid at position AA^2, which was revealed by comparing the sublibraries of different lengths. Presumably, the sequence AA^1-cationic-cationic is suited best to the task of binding to the nucleic acid. This would explain the difficulty of incorporating anything but a basic amino acid at position AA^2 and AA^3, the huge beneficial effect when going from n = 1 to n = 2 building blocks and the lower influence of position AA^1 on the binding efficiency as well as the compatibility of this position with aromatic amino acids. Furthermore, this also explains why DNA-binders with just two positive amino acids are able to compete with the best ligands (see Top10). In comparison to the already characterized ligands **196** and **206**, which are only medium performers, the best sequences, like Trp-Lys-Arg-GCP (#1) should bind more strongly to $p(dAdT)_2$ by at least one order of magnitude. In order to verify these results in solution, five library members were selected and synthesized.

4.5.3 Synthesis of Additional Divalent Ligands

Based on the results of the screening of the library five ligands were identified for synthesis on larger scale and characterization by binding studies in solution as depicted in Figure 140. They were chosen from different regions of the library, starting with the best performing sequence, Trp-Lys-Arg-GCP (#1, **233**), followed by Trp-Arg-Lys-GCP (#4, **234**), the highly charged Lys-Lys-Arg-GCP (#27, **235**), Lys-Trp-Arg (#94, **236**) and as a negative control Glu-Trp-Arg-GCP (#234, **237**).

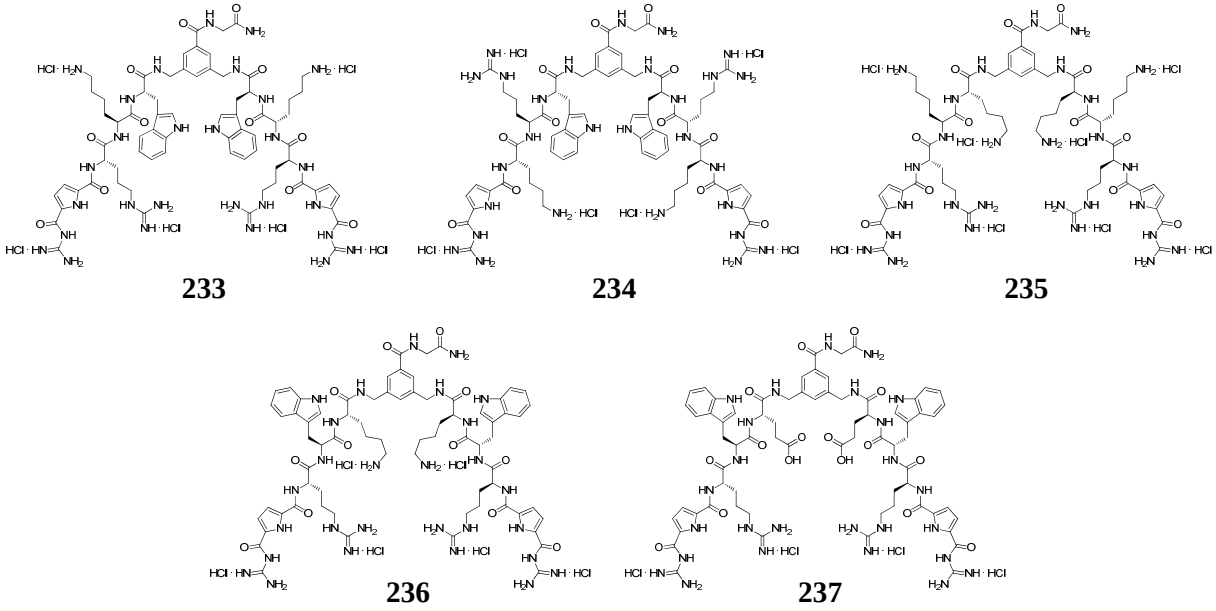


Figure 140. *Five selected members of the library (**233-237**) were synthesized for characterization in solution.*

The synthesis was carried out according to the optimized microwave-assisted procedure described in chapter 4.3.2. The first attempt to synthesize the novel DNA-ligands was therefore conducted on MBHA resin (1.94 mmol/g). The attachment and coupling steps were carried out with the help of PyBOP in 5 % DIPEA/DMF under argon atmosphere and irradiation to max. 60° C for 20 min at 20 W. Fmoc deprotection was achieved by treating the resin with 20 % piperidine/DMF. The same side-chain protection scheme as applied for the synthesis of the library was applied (see chapter 4.5.1). All coupling reactions were carried out twice or more often if indicated by the Kaiser test. Due to limitations of the available amount of Boc-GCP the coupling steps were conducted with 3 instead of 6 eq. Syntheses were carried out for the preparation of **233**, **234**, and **235**. Cleavage from the resin and removal of the protecting groups was conducted at room temperature by treating the resin with 8 % TFMSA in 80 % TFA containing 8 % TIS and 4 % EDT as nucleophilic scavengers. HPLC analysis of the crude product revealed the presence of > 40 compounds. Separation via MPLC in fractions containing < 5 compounds and consequent analysis with the help of MS revealed that no product was obtained for **233** and **234** and only minor amounts for **235**. Obviously, the cleavage conditions were too harsh for the more susceptible amino acids arginine and especially tryptophan. Therefore, the synthesis was repeated as described above, but on Fmoc-Rink-Amide resin (0.94 mmol/g) which allowed to cleave the resin-bound ligands under milder conditions with 95 % TFA containing 2.5 % TIS and 2.5 % water as nucleophilic scavengers. As summarized in Table 19 all compounds were obtained this way in

sufficient amounts with moderate yield and could be isolated after several rounds of purification via MPLC (RP18, H_2O → MeOH, 0.1 % TFA) with high purity as determined with the help of analytical HPLC (see Figure C. 5-Figure C. 9). After anion exchange of the trifluoroacetate counterions for chlorides the compounds were ready for solution phase binding studies.

Table 19. *Yield and purity of the microwave-assisted SPPS of the five selected library members.*

	233	**234**	**235**	**236**	**237**
Yield [%]	6	19	9	15	21
Purity [%]	93	93	91	95	94
Sequence	Trp-Lys-Arg	Trp-Arg-Lys	Lys-Lys-Arg	Lys-Trp-Arg	Glu-Trp-Arg
Library #	1	4	27	94	234

4.5.4 Binding Studies in Solution

To verify the results obtained from the on-bead screening of the library and to rule out potential effects on the binding event caused by the solid support, binding studies were carried out in solution with the five novel divalent DNA-ligands **233-237**. First, a DAPI displacement was conducted in cacodylate buffer (0.01 M) at neutral pH and 25° C as depicted in Figure 141 and explained in chapter 4.4.4. To a mixture of p(dAdT)$_2$ (2.5 µM, 5 eq) and DAPI (0.5 µM, 1 eq), which was prepared in a fluorescence microcuvette equipped with a stopper (900 µL), stock solutions of the DNA ligands (50 µM) were titrated in aliquots of increasing volume (e.g. 10 × 1, 2, 4 µL) and the resulting fluorescence spectrum upon irradiation at 358 nm was recorded between 420 and 500 nm. Excerpts of the fluorescence emission at 460 nm were corrected for dilution and after subtraction of DAPI´s own emission, they were plotted against [DAPI] / [Ligand] and fitted with a first order exponential decay function to obtain the amount of DNA-ligand, which is necessary to displace half of the DAPI from the nucleic acid as IC_{50} values. All measurements were carried out at least in duplicate and are depicted in appendix C.6 (Figure C. 66-Figure C. 70). The so obtained results as well as the ones for the divalent ligands **196** and **206** and for comparison the three-armed **207** are listed in Table 20.

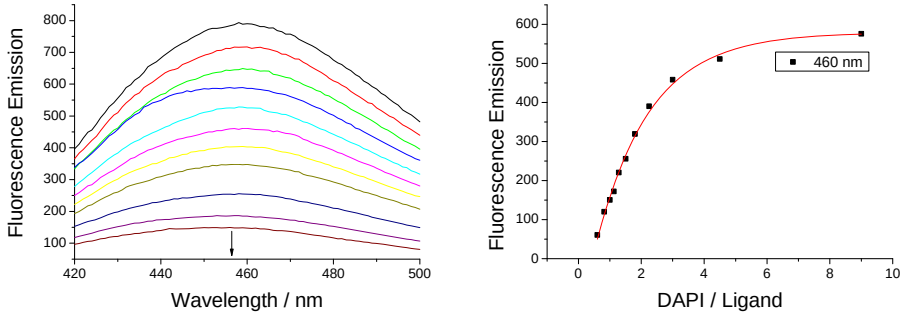


Figure 141. *DAPI displacement experiment with p(dAdT)$_2$ and **233** (left) and excerpt of the fluorescence emission at 455 nm plotted against the ratio of DAPI/**233** with the corresponding fit (right).*

Table 20. *IC_{50} values for divalent (**196**, **206**, **233-237**) and trivalent (**207**) DNA-binders as obtained by DAPI displacement assays in buffered aqueous solution (cacodylate buffer, 0.01 M) at neutral pH and the results of the screening on-bead. The binding constants obtained from ITC measurements are listed for comparison.*

	233	**234**	**235**	**206**	**236**	**196**	**237**	**207**[a]
IC_{50}[b] [eq]	0.77 ± 0.2	0.61 ± 0.1	2.23 ± 0.3	15.9 ± 3.7	0.52 ± 0.1	52.2 ± 16.9	42.3 ± 2.0	1.72 ± 0.3
On-bead[c] [%]	26	33	52	63	65	74	80	-
Library #	1	4	27	72	94	194	234	-
Solution #	3	2	4	5	1	7	6	-
logK_2[d]	9.76	9.09	8.77	6.79	9.04	5.80	6.18	7.51
Sequence	Trp-Lys-Arg-GCP	Trp-Arg-Lys-GCP	Lys-Lys-Arg-GCP	Lys-Phe-GCP	Lys-Trp-Arg-GCP	Phe-Lys-GCP	Glu-Trp-Arg-GCP	Phe-Lys-GCP

a) Trivalent ligand; b) Solution phase DAPI displacement; c) On-bead DAPI displacement; d) ITC.

In general, there is an excellent agreement between the results in solution with those obtained from the on-bead screening thus proofing their validity. Low IC_{50} values correspond to low percentaged values of non-displaced DAPI on-bead. All deviations are within the

standard deviation of the obtained data with the exception of ligand **236**, which has a very low IC_{50} value despite its mediocre performance during the on-bead screening (possibly due to corrupted beads). Together with **233** and **234** this DNA-binder had the highest affinity to $p(dAdT)_2$ according to the solution phase DAPI displacement. These three ligands all feature the same amino acids, Trp, Arg, and Lys, but in different order. The combination of these three amino acids seems to be best fit for displacing DAPI from the minor groove of the DNA. The IC_{50} values of the tested ligands range from 0.52 for the best sequence **236** to 52.2 equivalents for the original divalent ligand **196** thus spanning two orders of magnitude. Comparison with the IC_{50} value of the trivalent ligand **207** (1.72), which had the highest affinity of all synthesized nucleic acid binders until this point, clearly demonstrates the tremendous improvement of binding affinity as a result of optimization by means of the combinatorial library. Based on these values, binding constants of at least 10^8 M^{-1} were expected. After this qualitative validation of the results of the screening, ITC measurements were conducted in order to obtain quantitative binding data. Therefore, aliquots (20-30 × 5 µL) of the ligands (0.1 mM) were titrated into solutions of $p(dAdT)_2$ (0.04 mM, 1451 µL) in neutral cacodylate buffer (0.01 M) at 25° C. The raw data was corrected for the endothermic heat of dilution of the ligands and fitted according to two-site models for **233-236** and to a one-site model for **237**. In contrast to all other ligands, the latter one also featured an exothermal heat of dilution. Other binding models did not give satisfying curve fitting. Two exemplary titrations are shown in Figure 142, all others can be found in appendix C.5 (Figure C. 63). The results are listed in Table 21 together with those obtained for **196**, **206**, and **207** (cf. chapter 4.4.3). From top to bottom the divalent ligands are in order of decreasing affinity.

Table 21. *Results of ITC measurements obtained by titrating the library-derived DNA binders **233-236** (0.1 mM) to $p(dAdT)_2$ (0.04 mM, 1451 µL) in cacodylate buffer (0.01 M) at neutral pH and 25° C.*

Ligand	Sequence	K [10^7 M^{-1}]	log K	n	ΔH [kcal/mol]	$T\Delta S$[a] [kcal/mol]	ΔG [kcal/mol]
233	Trp-Lys-Arg	0.4	6.59	0.09	3.2	12.2	-9.0
		578	9.76	0.07	-18.5	-5.2	-13.3
234	Trp-Arg-Lys	0.5	6.73	0.05	4.8	14.0	-9.2
		123	9.09	0.09	-19.5	-7.1	-12.4
236	Lys-Trp-Arg	0.1	5.99	0.06	9.3	17.4	-8.2
		111	9.05	0.07	-21.9	-9.5	-12.3
235[b]	Lys-Lys-Arg	0.1	6.08	0.07	7.8	16.1	-8.3
		59	8.77	0.07	-13.4	-1.4	-12.0
206	Lys-Phe	0.6	6.79	0.27	-18.7	-9.4	-9.3
237	Glu-Trp-Arg	0.2	6.18	0.55	-4.2	4.2	-8.4
196	Phe-Lys	0.1	5.80	0.34	-16.4	-8.5	-7.9
207[d]	Phe-Lys	3.3	7.51	0.20	-21.0	-10.7	-10.3

a) Calculated for 25° C; b) [**235**] = 0.06 mM, $[p(dAdT)_2]$ = 0.03 mM; c) [**237**] = 0.2 mM, $[p(dAdT)_2]$ = 0.04 mM; d) Trivalent ligand.

The results of the ITC experiments are in excellent agreement with the IC_{50} values of the DAPI displacement and with the on-bead screening results. As predicted in the screening, **233** features the highest binding constant towards $p(dAdT)_2$ with an excellent 5.8×10^9 M^{-1}. With the help of the combinatorial optimization of the divalent ligands their affinity could be increased by a factor of 1000 (**206**: $K = 6.2 \times 10^6$ M^{-1}). In total, the binding constants of the

library members span four orders of magnitude (**196**: $K = 6.3 \times 10^5$ M^{-1}) clearly proofing the importance of the choice of appropriate amino acids for high affinity binding. The best receptors of the library are even more efficient in binding to the nucleic acid than their three armed analogue **207** ($K = 3.3 \times 10^7$ M^{-1}) by two orders of magnitude.

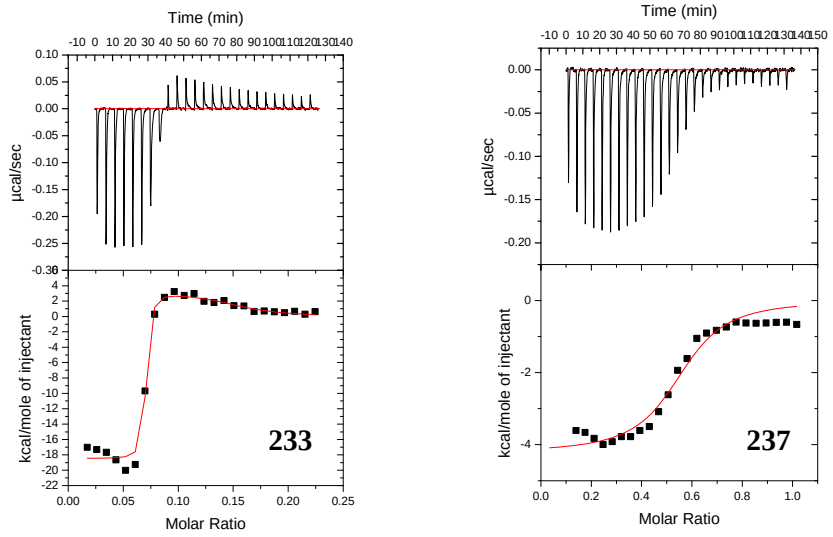


Figure 142. *Top: ITC experiments were carried out in sodium cacodylate buffer (0.01 M) at neutral pH by titrating aliquots of 5 µL of **233** (c = 0.1 mM), or **237** (c = 0.2 mM) to p(dAdT)$_2$ (0.04 mM). Bottom: the titrations were corrected for dilution and fitted with a two-site model for **233** and a one-site model for **237**. The titration with **233** is corrected for the first; and the one for **237** for the first two data points.*

The binding mode of the ligands with three amino acids per arm is different than for their shorter analogues **196** and **206** and the negative control **237**. While the binding data could only be fitted with a two-site model for the first mentioned, the latter ones had to be fitted according to a one-site model. For **233-236** two processes seem to be happening simultaneously. The first one is similar for all ligands: It takes place with a binding constant in the region 0.1-0.5 $\times 10^7$ M^{-1} and is endothermic and driven by entropy. Such an energetic pattern is typical for unspecific binding to the ribophosphate backbone, which is driven by the entropically favored liberation of counterions and water molecules into the solvent.[229] The second binding event manifests with varying affinity and has the same energetic pattern (exothermal, unfavorable entropy) as the binding of **196** and **206** and was therefore attributed to groove binding. Probably the charge-mediated backbone binding can only be resolved by an ITC experiment if the interaction is strong. Therefore, only the highly charged compounds **233-236**, but not the lower charged **196** and **206** featured this second binding mode. For the latter two, the groove binding event is predominant and overlays the backbone binding, which is without doubt taking place to some extent with these systems, too. The negative control **237** finally features yet another binding mode, which is driven by a combination of enthalpy and entropy.

Furthermore, the preferences found in the positional and charge analysis of the library were validated, too. When comparing the two ligands with identical amino acids (Trp, Lys, Arg), but in different order **233** (Trp-Lys-Arg-GCP) and **236** (Lys-Trp-Arg-GCP) it becomes clear that the receptor of type AA^1-charged-charged-GCP is favored over the one which combines

the two sequence criterions, which were identified as unfavorable for binding: Lys-AA^2-AA^3 and AA^1-aromatic-AA^3. Although this effect is less pronounced than predicted by the on-bead screening, the two compounds still differ in their binding affinity by a factor of 5. The second important finding of the library screening was validated as well: The affinity to the nucleic acid correlates with the charge of the artificial ligands, but only to some extent. A DNA binder with the correct amino acid sequence can outperform an analogue ligand with higher charge, e.g. **233** binds stronger to $p(dAdT)_2$ by one order of magnitude, although it features two charges less than **235** (Lys-Lys-Arg-GCP). Finally, the last trend of the library, the occurance of aromatic amino acids in high performing ligands at the first combinatorial position (aromatic-AA^2-AA^3) was also confirmed. The two best ligands in solution (**233** and **234**) featured tryptophan in position AA^1.

In conclusion, with the help of standard Fmoc SPPS and a split & mix strategy, a small but focused library for the identification of efficient minor groove binders was synthesized. Qualitative information about the affinity towards $p(dAdT)_2$ for all library member/nucleic acid combinations was collected by screening the library with a DAPI displacement assay. Except for one deviation, the qualitative results and general trends were verified in solution by DAPI displacement and ITC experiments revealing an improvement of the affinity by three orders of magnitude compared to the original divalent ligands. The highest binding constants were found for sequences containing the three amino acids Trp, Lys, and Arg. The overall highest binding constant was as high as an excellent $5.8 \times 10^9\ M^{-1}$. Clearly, with the help of this rather small library, it was not only possible to reveal sequence dependent binding properties, it was also possible to dramatically increase the potency of the divalent DNA binders for the recognition of B-type DNA.

4.5.5 Identification of DNA-Ligands for Transfection Experiments

On-Bead Screening with pF143-GFP plasmid DNA

After having successfully utilized the combinatorial library **220** for the identification of efficient minor groove binders, which bind to $p(dAdT)_2$ with high affinity, the idea was to utilize the library to identify efficient binders of plasmid DNA and to test whether improved binding affinity does lead to enhanced ability to act as transfecting agent. Therefore an ethidium bromide displacement on-bead screening was conducted. Since the secondary structure of the generic plasmid DNA is not regular, and therefore does not necessarily feature a well-defined minor groove for the accommodation of DAPI, the intercalator EB was utilized for the displacement assay instead. First, a solution of EB (0.75 µM, 1 eq) was prepared in 96-well plates (300 µL) in neutral cacodylate buffer (0.01 M) and its fluorescence emission upon irradiation at 520 nm was recorded between 595 and 605 nm. Then, pF143-GFP plasmid DNA (3 µM, 4 eq) was added, the solutions thoroughly mixed with the help of a pipette and after 30 min incubation time the increased fluorescence emission was determined. Then, 100 equivalents of the individual library members were added as isopycnic suspensions (DCE/*i*Prop = 4/1; 5 mg resin per 400 µL). Three wells per ligand/DNA combination were analyzed to obtain average values, and three blank samples per plate (without resin) were utilized for normalization of the different plates. The mixtures were incubated on a shaker and the resulting fluorescence emission was measured after 16, 40, 68 and 88 h. Figure 143 depicts the relative fluorescence emissions of all library members after 88 h which is normalized to 100 % for the lowest performing ligand and corrected for EB´s own emission. Low intensities signify efficient displacement of the intercalator and therefore high binding affinities to the plasmid DNA.

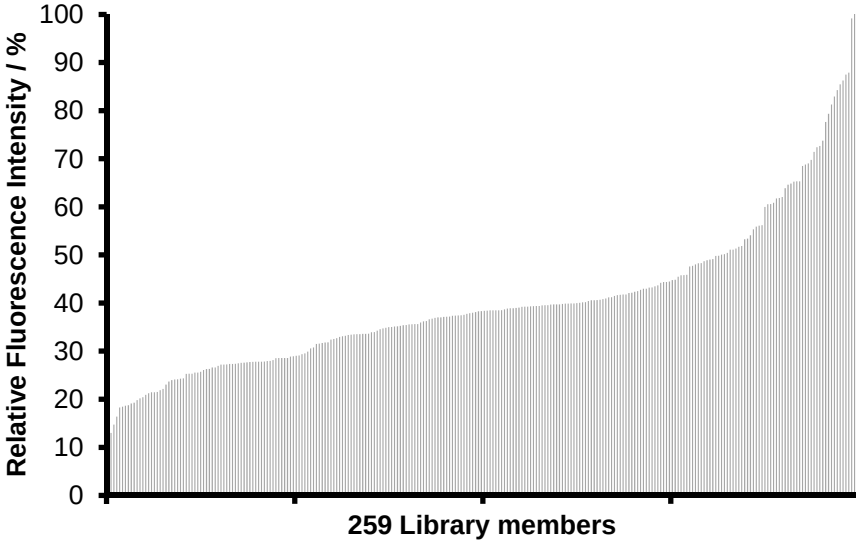


Figure 143. *Relative, normalized fluorescence emissions as obtained by a EB displacement assay from pF143-GFP plasmid DNA with the members of library* ***220*** *after 88 h incubation in cacodylate buffer (0.01 M) at pH 7.*

There is an excellent distribution of the fluorescence intensities between 9 and 100 % with a few low affinity ligands and a steadily improving, already well performing middle part, followed by a few very well performing DNA binders. The Top10 ranked library members are

listed in Table 22. All other results can be found in appendix C.7. The absolute occurrence of amino acids within the Top10 is similar to the one observed during the DAPI screening: arginine is present most (13), followed by lysine (6) and aromatic amino acids (total: 9). Glutamic acid is again not present in the best performing systems. In general, the ligands feature high positive charges (+5), but not the highest possible charge (+7). The charges were estimated for pH 7 as described in chapter 4.5.2. Similar to the results from other screening high charge seems to be advantageous for DNA binding, but not the only criterion. There is even an efficient ligand with just three positive charges and two amino acids in the Top10. Similar to the DAPI screening, there are two short ligands present, which seem to be able to compete with their longer versions. The distribution of the amino acids on the other hand is different. Although the two best ligands feature a tryptophan at position AA^1, all others do not feature an aromatic amino acid at this position in contrary to the DAPI screening with 50 % aromatic amino acids at AA^1. At AA^2 on the other hand, there are now more aromatic amino acids present. All these are indications that other prerequisites are necessary to bind to the plasmid instead to the AT-heteropolymer.

Table 22. *Top10 ranked members of library* ***220*** *as obtained from an EB displacement assay with pF143-GFP plasmid DNA and comparison to their rank as obtained from the DAPI displacement screening with* $p(dAdT)_2$*. The charges are estimates for pH 7: full protonation of Arg, Lys and 50 % protonation of GCPs.*

# plasmid	# $p(dAdT)_2$	AA^1	AA^2	AA^3	Charge	%
1	4	Trp	Arg	Lys	5	9
2	1	Trp	Lys	Arg	5	13
3	131	Lys	Phe	Arg	5	15
4	148	Lys	Lys	Phe	5	16
5	196		Arg	Tyr	3	18
6	37	Arg	Lys	Tyr	5	18
7	79	Arg	Tyr	Arg	5	19
8	204	Arg	Trp	Arg	5	19
9	31	Arg	Phe	Arg	5	19
10	10		Arg	Arg	5	19

Comparison with the two divalent ligands **196** and **206**, which have already been characterized for their binding affinity towards the plasmid ($K = 2$ and $9 \times 10^6\ M^{-1}$) reveals that these are merely ranked in the middle part of the library (**206**: #136 with 38 %; **196**: #148 with 39 %). Therefore, the compounds with the highest efficiency in displacing EB from the plasmid most likely bind to this ds-DNA with binding constants higher by 2 to 3 orders of magnitude than the original ligands.

Comparison of the Top10 ranked library members with their performance during the DAPI displacement assay further suggests that there is no direct correlation between the binding properties towards the two different nucleic acids. In Figure 144 the ranking of all library members in the EB displacement screening is plotted against the one obtained from the DAPI displacement. It now becomes clearly obvious that there is no correlation between the two screenings. Efficient $p(dAdT)_2$ binders are not necessarily good plasmid binders and vice versa.

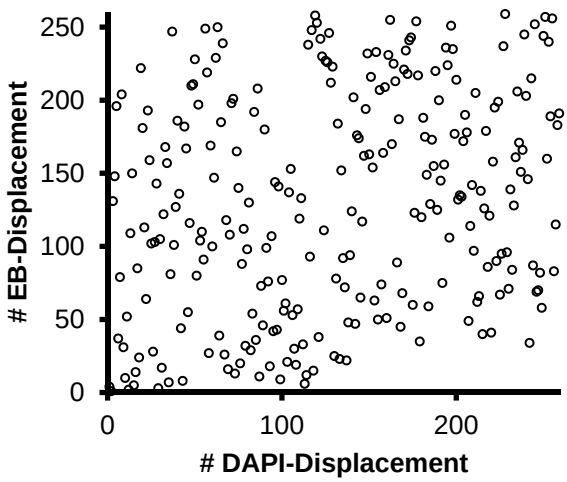


Figure 144. *Correlation between the ranking of the library members during DAPI and EB displacement assays.*

Division of the library into four compartments according to their efficiency and determination of the percentaged appearance of the individual building blocks within each compartment leads to the illustration in Figure 145. A clear trend can be observed for the best performing ligands (Top 25 %). Arginine is implemented most often (31 %), followed by lysine, tyrosine (both 20 %) and the other two aromatic amino acids (both 13 %). Almost no glutamic acid is present within this region of the library. When going to lower affinity ligands, the occurrence of arginine is steadily dwindling. Similarly, lysine peaks in the second compartment and is less often implemented in DNA-binders of lower affinity. The aromatic amino acids are distributed quite evenly over all four compartments, with slightly rising amounts for less efficient ligands. The worse the performance of the library members, the more often glutamic acid can be found.

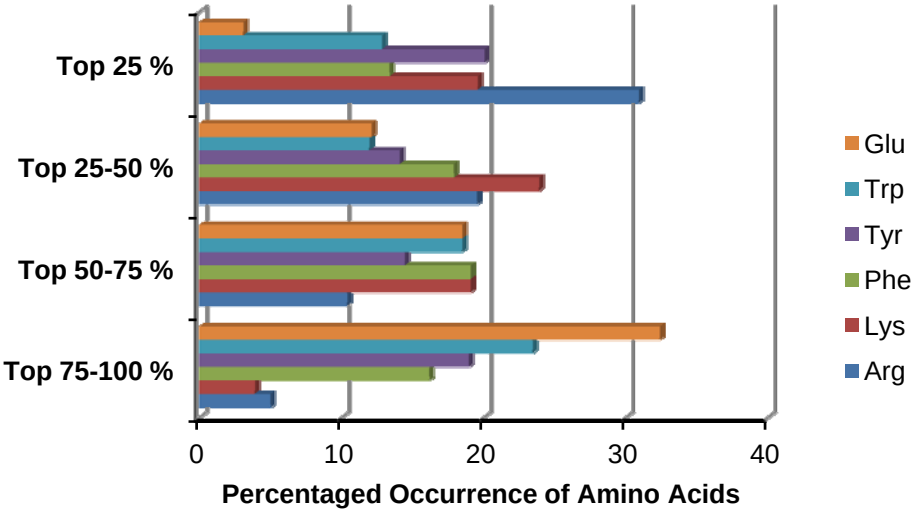


Figure 145. *Percentaged amino acid distribution within each of the four compartments of the library.*

A closer look at the amino acid distribution at the three combinatorial positions AA^1-AA^3 within the four compartments was carried out next. The results are shown in Figure 146. At the first combinatorial position the most beneficial amino acid by far is arginine, followed by

tyrosine and lysine. The other two aromatic amino acids, phenylalanine and especially tryptophan seem to prevent efficient binding when implemented at this position. Especially the latter trend is in strong contrast to the two overall best sequences, which feature Trp at position AA^1 and therefore seem to be the exception to this rule. A comparison of sequences of the type Trp-X-X in the Top 25 % with those in the Top 75-100 % compartment revealed, that all good performing ligands with Trp in position AA^1 are of the type Trp-cation-X, while sequences that only bind weakly do not feature an adjacent cationic amino acid, but another aromatic one or glutamic acid. Thus, Trp only seems to be favorable for binding in the first combinatorial position if Arg or Lys are present in AA^2. Not a single glutamic acid is implemented in AA^1 within the first compartment.

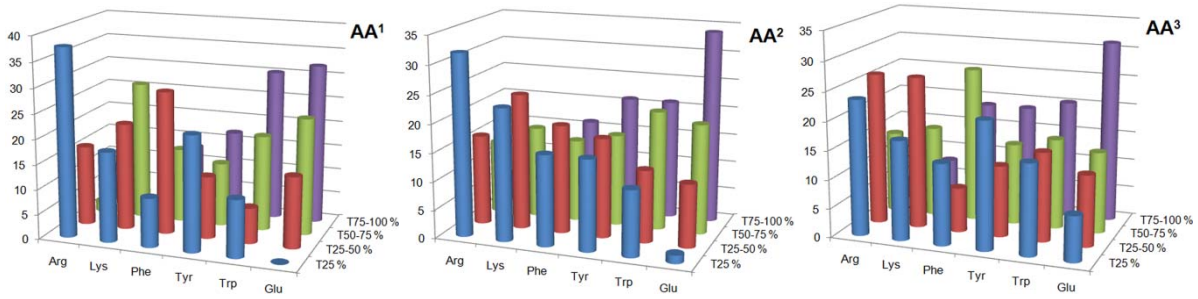


Figure 146. *Percentaged amino acid distribution at the three combinatorial positions AA^1-AA^3 within each of the four compartments of the library.*

At position AA^2 it is arginine again, which seems to improve binding the most, followed by lysine. Aromatic moieties are evenly spread over all four compartments at this position, neither benefiting nor preventing binding. This is in contrast to the results obtained for $p(dAdT)_2$, where an implementation of an aromatic amino acid at this position had a negative influence on the binding affinity. Although generic plasmid DNA is mainly of the B-type, too, it does not feature the very deep and narrow minor groove as is the case for the AT-heteropolymer. Thus, a bulky aromatic amino acid does not prevent efficient binding, because it does not have to adapt to the narrow groove. At position AA^3, a slight preference for arginine and tyrosine was observed, although this trend was not very pronounced as it was the case for the other positions. The exception is again glutamic acid, which is not tolerated at AA^3 either.

Finally, the binding affinity of the ligands with respect to their charge was studied. The absolute appearance of species with the same charge within the four compartments is depicted in Figure 147. As expected, the general trend is that highly charged compounds do interact stronger with the plasmid than lower charged or even negatively charged ligands. The latter ones are mainly clustered in the region of lowest performance. In comparison with the positively charged ligands, these molecules do interact worse with the DNA by large. However, the highest charged compounds are not necessarily the most efficient ones. For instance, half of the compounds of highest charge (+7) can be found within the Top 25-50 % region, and not within the Top 25 %.

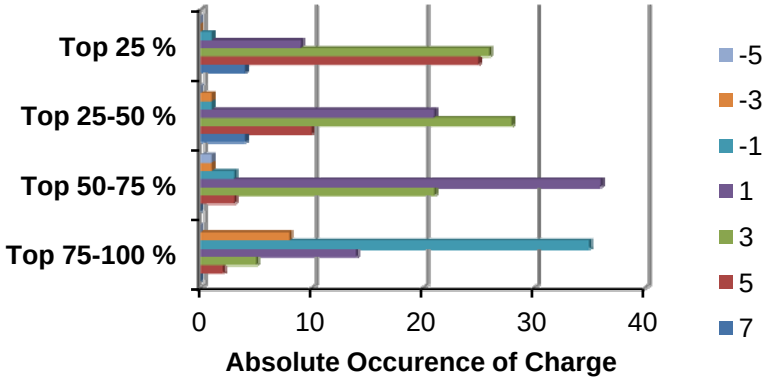


Figure 147. *Absolute charge distribution within each of the four compartments of the library.*

Although the general trends for efficient binding of the plasmid DNA are not as clear as it was the case for binding to $p(dAdT)_2$, the combinatorial library helped to identify compounds, which bind stronger to the plasmid DNA than the original divalent ligands **196** and **206** by at least 1-2 orders of magnitude.

Binding Studies in Solution

In order to verify the results from the on-bead screening, EB displacement assays were carried out in solution with the five already available novel divalent ligands **233-237** (see Figure C. 51-Figure C. 55). These ligands represent very well the different parts of the library. A solution of ethidium bromide (0.75 µM, 1 eq) was prepared in sodium cacodylate buffer (0.01 M) at neutral pH in a fluorescence microcuvette equipped with a stopper (900 µL). Upon irradiation at 520 nm the fluorescence emission between 560 and 650 nm was measured at 25° C. Then the plasmid was added to the solution (3 µM, 4 eq) and after 5 min incubation time the increased fluorescence emission was recorded. Stock solutions of the ligands (50 µM) were added in aliquots of increasing volume (e.g. 3 × 1, 5 × 2, 4 µL) and after 1 min incubation time after each addition the resulting fluorescence spectrum was measured. All experiments were carried out at least in duplicate. The averaged results of the EB displacement assays and those for **196**, **206** and **207** (cf. chapter 4.4.5) are listed in Table 23.

Table 23. *IC_{50} values for divalent (**196**, **206**, **233-237**) and trivalent (**207**) DNA-binders as obtained by EB displacement assays with pF143-GFP plasmid DNA in buffered aqueous solution (cacodylate buffer, 0.01 M) at neutral pH and on-bead screening results. Binding constants from ITC measurements are listed for comparison.*

	234	**233**	**235**	**206**	**236**	**196**	**237**
IC_{50}[a] [eq]	0.51 ± 0.03	0.55 ± 0.04	0.64 ± 0.05	2.08 ± 0.18	0.36 ± 0.04	1.85 ± 0.12	31.2 ± 0.7
On-bead[b] [%]	9	13	28	38	39	39	41
Library #	1	2	58	136	139	148	171
Solution #	2	3	4	6	1	5	7
$\log K_2$[c]	8.74	8.42	7.51	6.96	9.22	6.30	5.82
Sequence	Trp-Arg-Lys-GCP	Trp-Lys-Arg-GCP	Lys-Lys-Arg-GCP	Lys-Phe-GCP	Lys-Trp-Arg-GCP	Phe-Lys-GCP	Glu-Trp-Arg-GCP

a) Solution phase EB displacement; b) On-bead EB displacement; c) ITC.

With the exception of **236** there is an excellent agreement between the results from the on-bead screening with the solution phase measurements. Similar to the observations made for $p(dAdT)_2$ **236** binds stronger than predicted by the screening. This further underlines the hypothesis of corrupted beads for this library member. Probably the synthesis failed in this case. However, out of the seven substances tested in solution, this is the only one which does not behave as expected, indicating that the overwhelming majority of the library is intact. In the screening against the plasmid the ligand **236** seems to be the most potent binder with an IC_{50} value as low as 0.36 equivalents, followed by the two ligands **234** (0.51 eq) and **233** (0.55 eq).

Table 24. *Results of ITC measurements obtained by titrating the library-derived DNA binders **233-236** (0.1 mM) to pF143-GFP plasmid DNA (0.04 mM, 1451 µL) in cacodylate buffer (0.01 M) at neutral pH and 25° C.*

Ligand	Sequence	K [10^6 M^{-1}]	log K	n	ΔH [kcal/mol]	$T\Delta S$[a] [kcal/mol]	ΔG [kcal/mol]
236[b]	Lys-Trp-Arg	2.6	6.42	0.12	4.6	13.4	-8.8
		1670	9.22	0.08	-21.5	-8.9	-12.6
234[c]	Trp-Arg-Lys	0.55	5.74	0.12	5.4	13.2	-7.8
		553	8.74	0.10	-18.1	-6.2	-11.9
233[d]	Trp-Lys-Arg	(0.003)	(3.40)	(0.41)	(113)	(118)	(-4.7)
		266	8.42	0.09	-16.6	-5.1	-11.5
235[e]	Lys-Lys-Arg	0.25	5.40	0.03	(95)	(102)	-7.3
		32.6	7.51	0.08	-16.6	-6.4	-10.2
206	Lys-Phe	9.1	6.96	0.23	-17.5	-8.0	-9.5
196	Phe-Lys	2.0	6.30	0.24	-9.6	-1.0	-8.6
237[f]	Glu-Trp-Arg	0.7	5.82	0.55	-5.8	2.1	-7.9
207[g]	Phe-Lys	26.4	7.42	0.09	-22.7	-12.6	-10.1

a) Calculated for 25° C; b) [**236**] = 0.1 mM, [pF143-GFP] = 0.04 mM; c) [**234**] = 0.1 mM, [pF143-GFP] = 0.035 mM; d) [**233**] = 0.125 mM, [pF143-GFP] = 0.05 mM; e) [**235**] = 0.06 mM, [pF143-GFP] = 0.03 mM; f) [**237**] = 0.2 mM, [pF143-GFP] = 0.03 mM; g) Trivalent Ligand.

In order to determine quantitative data, ITC measurements were conducted next. Therefore solutions of the ligands **233-237** (0.06-0.2 mM) were titrated in aliquots (25 × 5 µL) to a solution of the plasmid (0.03-0.05 mM, 1451 µL) in cacodylate buffer (0.01 M) at pH 7 and 25° C. The raw data was corrected for the heat of dilution of the ligands and the data was fitted according to two-site models for **233-236**, and with the help of a one-site model for **237**. Other binding models did not result in satisfying fits. The titration experiments are depicted in appendix C.5 (Figure C. 64-Figure C. 65) and the results are listed in the order according to their binding affinity in Table 24. The observed binding constants were in good agreement with the EB displacement and the on-bead screening (except for **236** as denoted above). **236** featured the highest binding constant to the plasmid with $K = 1.7 \times 10^9$ M^{-1}, followed by **234** (5.5×10^8 M^{-1}) and **233** (2.7×10^8 M^{-1}). Thus the binding constants of the best performing ligands are two to three orders of magnitude higher than those for the original divalent ligands **196** and **206**. They are even higher than for the trivalent DNA binder **207** by two orders of magnitude. Again, the highest charged compound **235** does not feature the highest binding constant (3.3×10^7 M^{-1}), suggesting that the correct combination of amino acid building blocks plays a decisive role besides mere charge-charge interactions to the nucleic acid. The binding mode is similar to the one observed for $p(dAdT)_2$, although in this case the

experimental results for the first binding event are not as accurate (and therefore listed in brackets in Table 24). It occurs with binding constants around 10^6 M^{-1} and is entropy driven and endothermic. The other event is exothermic but with unfavorable entropy and takes place with different binding constants for different ligands. With these data a binding mode with two simultaneous processes, an unspecific attraction to the phosphate backbone together with groove binding, can be proposed similar to the recognition event with $p(dAdT)_2$. The glutamic acid containing **237** again features a different binding mode, which is exothermal and with favorable entropy, but takes place with lower affinity.

Interestingly, for different nucleic acids different high affinity-compounds could be identified. Hence, with the help of combinatorial chemistry, polynucleotide-selective identification of efficient DNA-binders could be achieved. With these data at hand the correlation between the affinity of the ligands towards the pF143-GFP plasmid DNA to the ability to shuttle genetic material into human cells could be studied in the next step.

Transfection Experiments

Therefore transfection experiments were conducted with the optimized conditions from chapter 4.4.5 by *Gutschmidt* in the working group of *Prof. Knauer*. HEK293T cells were seeded in 96-well cell culture plates and grown for 24 h. Then a mixture of pF143-GFP plasmid DNA (2 µg) and the novel, library-derived gene carriers **233-237** (0.24 mM) were incubated in phosphate buffered saline at pH 7.2 for 10 to 15 min and added to the cells (total volume = 200 µL). After 48 h the number of GFP expressing cells was counted with the help of an inverted fluorescence microscope. All experiments were carried out at least in triplicate. The average values of the percentage of fluorescent cells were calculated and are reported together with the corresponding standard deviations in Table 25. The results obtained for **196**, **206**, and **207** are listed for comparison.

As depicted in Figure 148 three of the novel DNA binders were able to transfect the cells. The highest transfection efficiency was achieved by **236** with 9.6 %, followed by **233** (6.5 %) and **234** (6.0 %). These three systems all contain the amino acids Trp, Lys and Arg, but in different order. The higher charged **235** is not able to transfect cells. Neither is the negative control **237**. The overall transfection efficiencies are higher than for the original divalent DNA ligands **196** and **206** (< 1 and 2.9 %), but lower by large in comparison to the three-armed **207** (54.1 %). When carrying out the experiments at higher ligand concentrations (0.30 and 0.36 mM) the efficiencies dropped.

Table 25. *Transfection efficiency determined 48 h after treatment of HEK293T cells with pF143-GFP plasmid (2 µg) and varying amounts (0.24-0.36 mM) of* ***233-237****,* ***196****,* ***206****, or* ***207****.*

[Ligand] / [mM]	Transfection Efficiency and Standard Deviation / [%]							
	233	**234**	**235**	**236**	**237**	**196**	**206**	**207**
0.24	6.5 ± 2.3	6.0 ± 2.1	< 1	9.6 ± 2.8	< 1	< 1	2.8 ± 0.6	54.1 ± 12.5
0.30	3.1 ± 1.6	5.0 ± 1.6	< 1	1.6 ± 0.9	< 1	< 1	n.d.	n.d.
0.36	1.9 ± 0.5	4.8 ± 2.8	< 1	1.8 ± 1.4	< 1	< 1	n.d.	n.d.

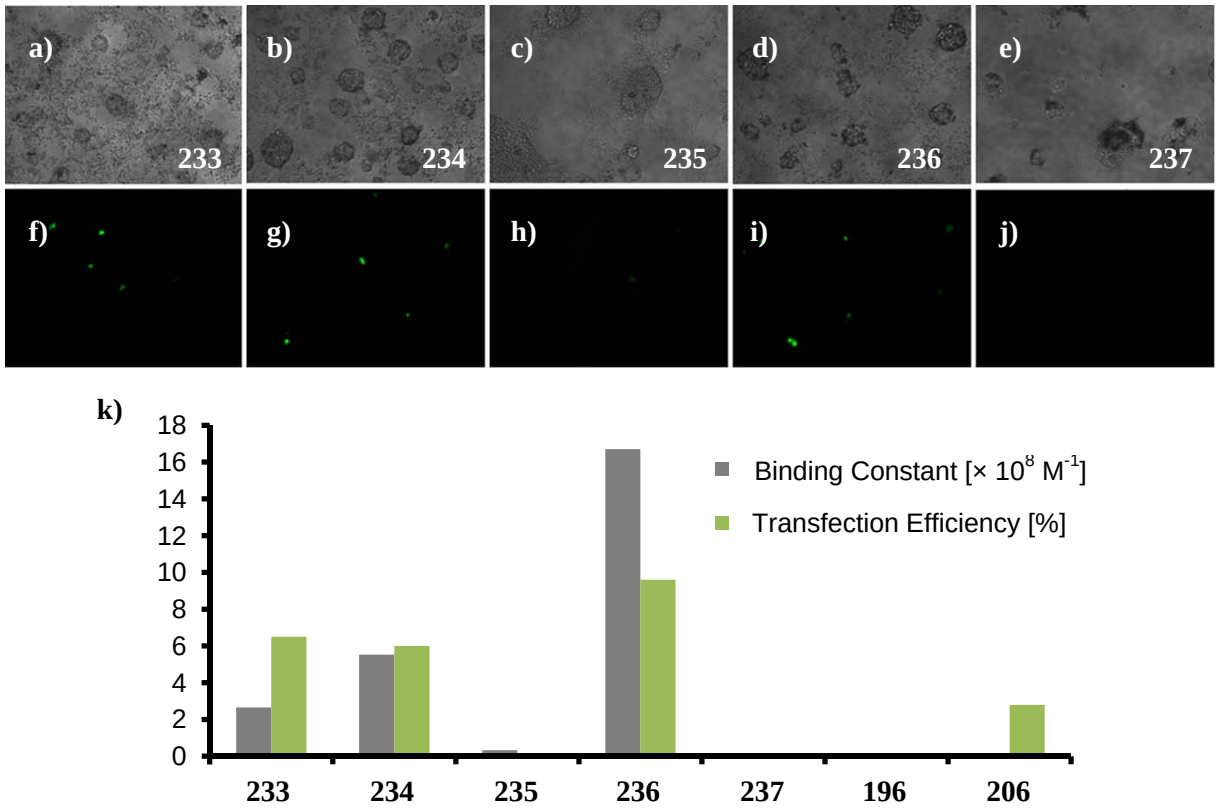


Figure 148. *Brightfield (a-e) and fluorescence images (f-j) of HEK293T cells 48 h after transfection with 2 µg pF143-GFP plasmid and* ***233-237*** *(0.24 mM); k) comparison of binding constants of* ***233-237****,* ***196****, and* ***206*** *with their transfection efficiency.*

The explanation for the lower transfection efficiencies is probably due to a lack of buffering capacity as could be shown for **196** compared to the tripodal carrier **207** in chapter 4.4.5. It is likely that the optimized divalent ligands are extremely efficient in transporting the plasmid into the cell, but are prevented from more efficient transfection by their lacking capability to escape from the endosome. Still, in comparison to the original two-armed carriers **196** and **206** the transfection efficiency could be increased. Figure 148 (bottom) illustrates that there is a direct correlation between binding constant and transfection efficiency for comparable DNA-carriers. Only the three ligands with the highest affinity led to transfection efficiencies > 5 %. The carrier with the highest affinity also featured the highest efficacy for cell transfection. The premise for such a comparison is of course that there are no additional effects, which differentiate the non-viral vectors from each other (cf. **196** vs. **207**: proton sponge effect). Hence, the optimization of binding affinity with the help of combinatorial chemistry is a viable tool for improving the transfecting properties of non-viral vector systems. The so optimized side chains can now be further modified chemically, for instance by implementing them into trivalent systems like **207** or by attaching lipophilic groups for improving their capability for endosomal escape.

In conclusion, the preparation of a small but focused library containing 259 divalent DNA ligands of varying lengths, which were built from aromatic (Trp, Tyr, Phe) and basic (Lys, Arg) amino acids (and acidic Glu as control), allowed to improve the binding affinity of their parent compounds to two different nucleic acids ($p(dAdT)_2$ and pF143-GFP plasmid) by a factor of up to 1000. The additional binding affinity does not only originate from additional charge-charge interactions of additional basic amino acids in the ligand arms with the

ribophosphate backbone, but there is also a sequence-dependent influence onto the binding properties. Sequences containing the three amino acids Trp, Lys, and Arg in varying order have proved to be excellent for binding to both types of DNA on bead and in solution. Finally, the transfection efficiency of the dipodal ligands could also be improved. Although the enhancement is not large due to their lacking ability to escape from endosomal vesicles, high affinity ligands could be identified and linked to improved transfection ability. By further modifying these sequences with for instance lipophilic groups, which would facilitate endosomal escape, their efficiency might be further improved. In the next chapter such molecules will be prepared and tested for their transfection efficiency.

4.6 AMPHIPHILIC GENE CARRIERS FOR CELL TRANSFECTION

Amphiphilic molecules are frequently being used for efficient cell transfection, for example the commercial Lipofectamine™ (cf. chapter 2.4.4).[18] These gene carriers feature a cationic headgroup which binds to the nucleic acid and contracts it. The lipophilic part of the molecule forms lamellar bilayers. Hence, the resulting lipoplexes are positively charged and attracted to the negatively charged cell surface. After uptake via endocytosis the cationic lipids interact with the anionic lipid vesicle wall which as a consequence is disrupted and releases its cargo into the cytosol. With the help of the combinatorial library presented in the last chapter cationic headgroups could be identified which are able to efficiently bind to plasmid DNA. By attaching C_{18} chains to this headgroup they become amphiphilic, similar to lipoplexes. In the following a synthetic route to this kind of gene carrier will be presented as well as the results from the corresponding transfection experiments.

4.6.1 Combined Solution Phase and Solid Phase Peptide Synthesis

The synthesis of the amphiphilic DNA carriers was carried out in a converged synthesis. First, the fully protected ligand was synthesized via microwave-assisted standard Fmoc SPPS on SASRIN resin as illustrated in Figure 149.[178]

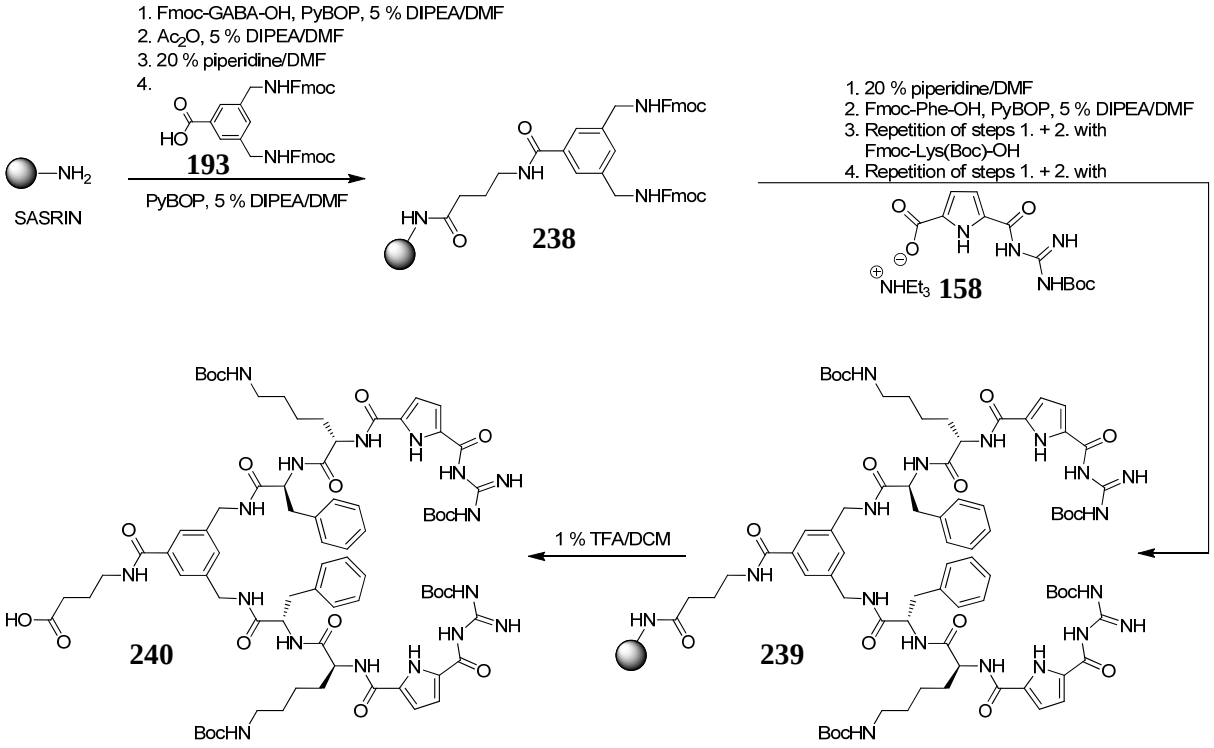


Figure 149. *Microwave-assisted SPPS of the fully protected headgroup **240**.*

After swelling the resin in DCM/DMF (7/3) the attachment of the first amino acid, Fmoc-γ-aminobutyric acid-OH (Fmoc-GABA-OH, 2 eq), was achieved with DIC (2 eq), HOBt (2 eq), and catalytic DMAP in a mixture of DCM/DMF (7/3) under argon atmosphere. The coupling step was repeated four times (8-12 h) to ensure complete conversion of the resin´s hydroxy groups. GABA was used instead of Gly in this case in order to introduce a longer

spacer between the lipophilic and the hydrophilic part of the gene carrier. Afterwards the solid support was treated with acetic anhydride to cap eventually still remaining free hydroxy groups. The Fmoc group was deprotected with 20 % piperidine/DMF and the aromatic template **193** (3 eq) was coupled to the liberated N-terminus with the help of PyBOP (3 eq) and microwave irradiation (max. 60° C, 20 min, 20 W) in 5 % DIPEA/DMF under argon atmosphere. All coupling steps were carried out twice. Coupling and Fmoc deprotection steps were monitored with the help of the Kaiser test and repeated if necessary.[181] The next two amino acids Fmoc-Phe-OH and Fmoc-Lys(Boc)-OH and the Boc protected GCP unit (**155**) were coupled accordingly, but with six instead of three equivalents of reactants. The resin-bound product **239** was cleaved from the solid support by treating the resin with 1 % TFA/DCM (12 × 15 min) followed by thorough washing steps with DCM and MeOH. The crude product was purified with the help of MPLC (RP18, H_2O → MeOH) to obtain **240** with 26 % yield.

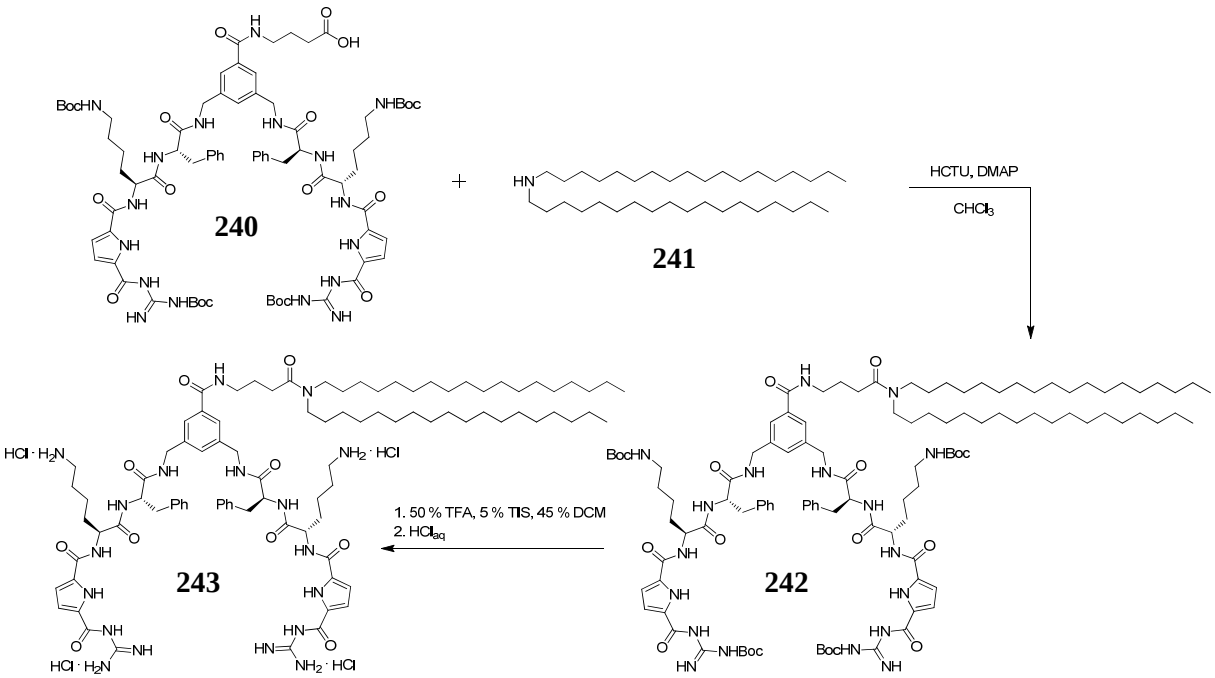


Figure 150. *Solution phase coupling of dioctadecylamine **241** with the fully protected headgroup **240** and consequent Boc deprotection with TFA gives the amphiphilic gene carrier **243**.*

In the next step, which is shown in Figure 150, the fully protected headgroup **240** (1 eq) was reacted with dioctadecylamine (**241**, 1 eq), HCTU (1.2 eq) as coupling agent and DMAP (3 eq) in chloroform adopted from a known procedure for the reaction of Boc-Gly-OH with **241**.[263] The resulting product **242** was directly liberated from its Boc protecting groups by treating it with 50 % TFA in DCM containing 5 % TIS as nucleophilic scavenger. The crude product was isolated via MPLC (RP8, H_2O → MeOH, 0.1 % TFA) with 90 % purity according to analytical HPLC (see Figure C. 10). Finally, the trifluoroacetate counterions were exchanged for chlorides by treating the product several times with 0.1 N HCl and consequent lyophilization to give the amphiphilic gene carrier **243** with 24 % yield over two steps. This rather moderate yield is probably due to a low conversion of **240** to C_{18}-

derivatized **242**. The amine nucleophile is sterically quite demanding and therefore, the reaction probably does not proceed very smoothly.

To circumvent these difficulties, dioctadecylamine (1 eq) was first reacted with Boc-Gly-OH (**244**, 1 eq) with HBTU (1.4 eq) and DMAP (3.7 eq) in chloroform under argon atmosphere with 93 % yield as illustrated in Figure 151.[263] The Boc protecting group was then removed by treating **245** with 38 % TFA in DCM to give **246** with 88 % yield. This compound could now be attached to protected headgroups more easily than the dioctadecylamine on its own.

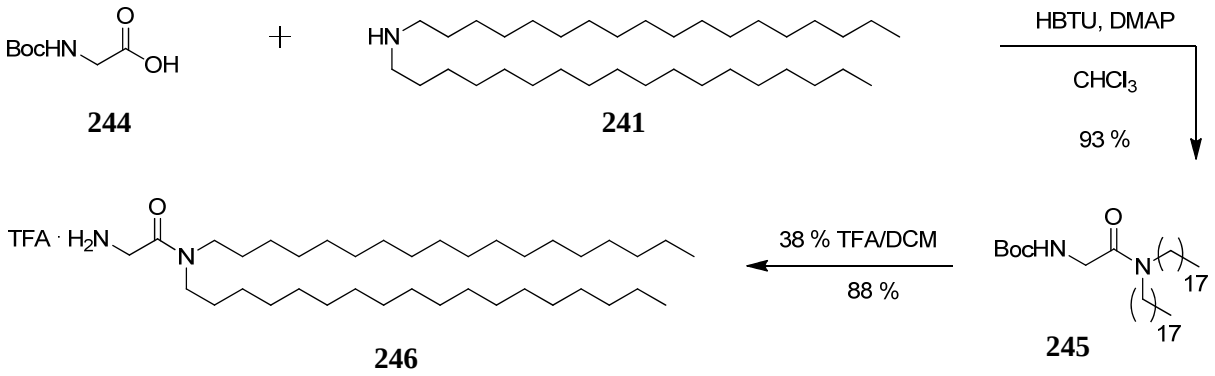


Figure 151. *Attachment of dioctadecylamine to Boc-Gly-OH and deprotection of the Boc group.*

The fully protected headgroups **247**-**249** shown in Figure 152 were synthesized via microwave-assisted SPPS on SASRIN resin as described above for **243**. The attachment of the first amino acid Fmoc-Gly-OH was achieved with DIC, HOBt in DCM/DMF (7/3) under argon atmosphere and catalytic DMAP and repeated four times. The resin was then treated with acetic anhydride in 5 % DIPEA/DMF to cap eventually remaining free hydroxy groups. Fmoc deprotections were conducted under microwave irradiation (1 + 5 min, max. 60° C, 20 W) and controlled by the Kaiser test. The following three amino acids Fmoc-Lys(Boc)-OH, Fmoc-Trp(Boc)-OH and Fmoc-Arg(Pbf)-OH were coupled in varying order with the help of PyBOP in 5 % DIPEA/DMF under argon atmosphere and microwave irradiation (20 min, max. 60° C, 20 W) and repeated at least twice or more often if indicated by the Kaiser test. Finally, the Boc protected GCP moiety was coupled accordingly and the compounds were cleaved from the solid support by treating the resin with a mixture containing 1 % TFA, 2.5 % TIS, and 2.5 % H_2O (4 × 2 h, 8 h) to obtain the fully protected headgroups with moderate yield (**249**: 12 %, **248**: 12 %, and **249**: 14 %).

The amino acid sequence of these three building blocks was chosen according to the results of the transfection experiments reported in chapter 4.5.5. They correspond to the three most efficient transfection agents **233**, **234**, and **236** derived from the combinatorial library **220**. The three headgroups (1 eq) were then coupled with C_{18}-derivatized glycine **246** (1.2 eq) with the help of PyBOP (1.4 eq) in 7.5 % DIPEA/DMF under argon atmosphere. The crude products were then treated with 95 % TFA, 2.5 % TIS, and 2.5 % H_2O to remove all protecting groups and purified via MPLC (RP8, H_2O → MeOH, 0.1 % TFA). The yield over two steps and the purity of the amphiphilic gene carriers **250**-**252** are listed in Table 26. Compared to the previous direct linking of the lipophilic amine to the fully protected head

group, it was possible to improve the yield by a factor of 2-3. The purity was determined via analytical HPLC at 300 nm and is also listed in Table 26 (see Figure C. 11-Figure C. 13). Excellent purities could thus be achieved. Finally, the trifluoroacetate counterions were exchanged for chlorides by dissolving the product several times in 0.1 N hydrochloric acid and consequent lyophilization. The so obtained compounds were now ready for further characterization and transfection experiments with human cells.

Table 26. *Yield over two steps of gene carriers **243**, and **250-252**. The purity was determined via analytical HPLC according to the absorption at 300 nm.*

Compound	**243**	**250**	**251**	**252**
Yield [%]	24	65	43	53
Purity [%]	90	94	99	99

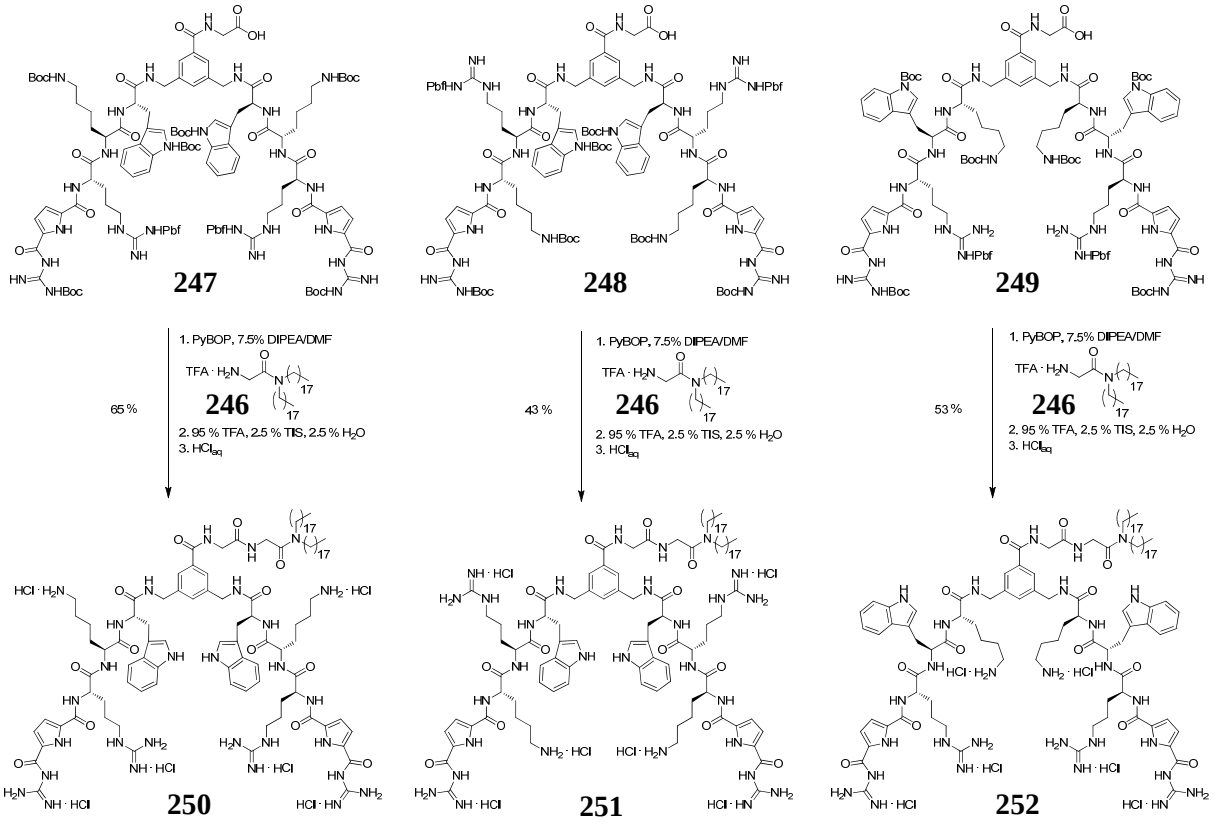


Figure 152. *The three fully protected headgroups **247-249** were synthesized via microwave-assisted SPPS. C_{18}-derivatized glycine (**246**) was coupled to their free C-termini and all protecting groups were removed to obtain the amphiphilic gene carriers **250-252**.*

4.6.2 Critical Micelle Concentration

The so obtained gene carriers **243**, and **250**-**252** are amphiphilic molecules and should therefore form micelles in aqueous medium. The critical micelle concentration (CMC) was determined with the help of a fluorescence assay.[264] Solutions of 1,6-diphenyl-1,3,5-hexatriene (DPH, **253**) were prepared at a concentration of 1 µM in in cacodylate buffer (0.01 M) at neutral pH. To these solutions varying amounts of amphiphile (1 µM - 0.1 mM) were added and the solutions incubated while shaking for 12 h. Afterwards the DPH fluorescence emission upon irradiation at 360 nm was measured between 380 and 600 nm. DPH is not soluble in aqueous medium on its own. Consequently only weak fluorescence could be observed in the absence and at low concentrations of the amphiphiles. As soon as the concentration is above a critical concentration, micelles are formed and the fluorophore is taken up within the hydrophobic core. This leads to an increase of fluorescence signal. A fluorescence experiment with **243** and the corresponding excerpt at 429 nm is depicted in Figure 153. It is clearly visible, that at a concentration of 1.9 µM the fluorescence intensity increases abruptly. This concentration corresponds to the critical micelle concentration. Table 27 features the CMC values for gene carriers **243**, and **250**-**252**.

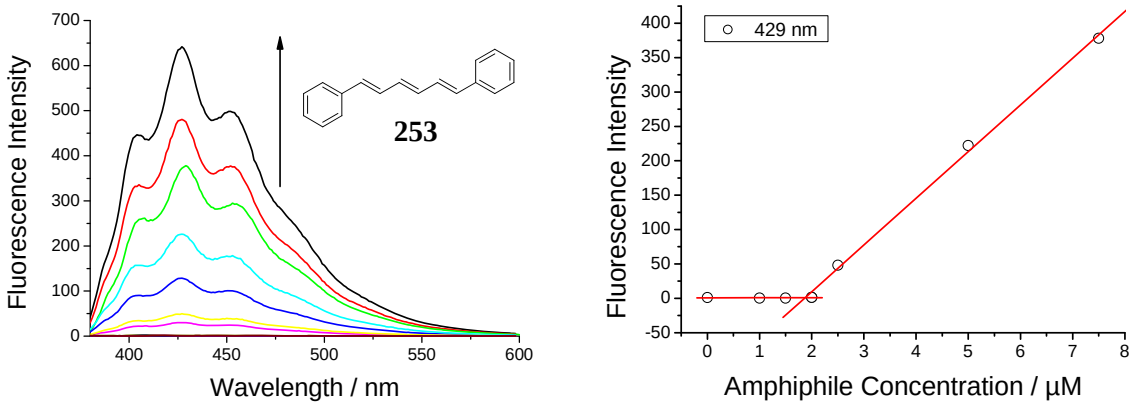


Figure 153. *Fluorescence experiment with DPH (1 µM) and increasing amounts of **243** in sodium cacodylate buffer (0.01 M) at pH 7 (left). Excerpt at 429 nm for determination of the critical micelle concentration (right).*

Table 27. *Critical micelle concentration in aqueous cacodylate buffer (0.01 M) at neutral pH as determined by a fluorescence assay with DPH.*

Compound	**243**	**250**	**251**	**252**
CMC [µM]	1.9	> 100	> 100	> 100

For the compounds **250**-**252** no micelle formation could be observed even at concentrations as high as 100 µM. Therefore, only a lower limit could be determined for the CMC of these three amphiphiles. In comparison to **243**, the higher charged headgroup of **250**-**252** improves their solubility in water and thus increases the critical concentration for micelle formation.

4.6.3 Transfection Experiments

The four amphiphilic DNA ligands were tested for their transfection efficiency by *Sarah Tillman* from the working group of *Prof. Knauer*. HEK293T cells were seeded in 96-well cell culture plates and grown for 24 h. Then a mixture of pF143-GFP plasmid DNA (2 µg) and various amounts of the gene carriers **243**, **250-252** (0.0075 mM, 0.03 mM, and 0.12 mM) were incubated in phosphate buffered saline at pH 7.2 for 10 to 15 min and added to the cells (total volume = 200 µL). After 48 h the number of GFP expressing cells was counted with the help of an inverted fluorescence microscope. All experiments were carried out at least in triplicate. The average values of the percentage of fluorescent cells were calculated and are reported together with the standard deviations in Table 28. The transfection efficiencies of the corresponding parent compounds **196**, **233**, **234**, and **236** (0.24 mM), which are lacking the C_{18} chains, as well as the one of PEI (0.12 mM) are listed for comparison, too.

Table 28. *Transfection efficiencies determined 48 h after treatment of HEK293T cells with varying amounts of pF143-GFP plasmid and **243**, **250-252**. The results from PEI (0.12 mM) and the divalent ligands without C_{18} chains **196**, **233**, **234**, and **236** (0.24 mM) are listed for comparison. For the latter ones the binding constant (K) towards the pF143-GFP plasmid is listed as well.*

Ligand	C_{18}	log K[a]	Transfection Efficiency and Standard Deviation / [%]			
			0.0075 mM	0.03 mM	0.12 mM	0.24 mM
PEI	-	-	n.d.	n.d.	50.7 (± 9.4)	n.d.
196	-	6.30	n.d.	n.d.	-	< 1
243	+	-	< 1[b]	< 1[b]	< 1[b]	n.d.
233	-	8.42	n.d.	n.d.	n.d.	6.5 (± 2.3)
250	+	-	< 1	2.8 (±1.7)	12.3 (± 2.9)	n.d.
234	-	8.74	n.d.	n.d.	n.d.	6.0 (±2.1)
251	+	-	11.0 (± 4.7)	30.0 (± 3.0)	68.2 (± 11.5)	n.d.
236	-	9.22	n.d.	n.d.	n.d.	9.6 (± 2.8)
252	+	-	14.3 (± 6.8)	49.0 (± 6.5)	21.7 (± 2.6)	n.d.

a) From ITC experiments; b) after 24 h.

As depicted in Figure 154 the C_{18}-derivatized transfecting agents (**243**, **250-252**) clearly outmatch their parent compounds (**196**, **233**, **234**, **236**) by up to one order of magnitude. The best nonviral vector **251** (68.2 %) even outmatches the gold standard PEI (50.7 %). Hence, the amphiphilic character of the novel DNA ligands dramatically improves their properties as gene carriers, most probably due to an enhanced capability for endosomal escape (cf. chapter 2.4.4). Furthermore, the minimal concentration for efficient transfection is reduced by half (0.24 mM → 0.12 mM). At even lower concentrations however, the efficiency drops, with the exception of **252**, which performs best at a concentration of 0.03 mM. At 0.12 mM precipitation was observed for the **252**/DNA complex, which reduced its efficiency. Precipitation is also the cause for an upper concentration limit of 0.12 mM for the amphiphilic gene carriers.

Another striking feature of the results is that the transfection efficiency does indeed correlate with the affinity of the headgroup of the amphiphiles towards generic DNA. While the compounds derived from high affinity sequences **250-252** (K of parent compounds ca. 10^9 M^{-1}) feature modest to excellent transfection efficiencies, the low affinity headgroup-derived **243** (K of parent compound ca. 10^6 M^{-1}) only enables transfection of minor amounts

of the cells. While the correlation is straightforward for **243** (< 1 %, log *K* = 6.30), **250** (12.3 %, log *K* = 8.42), and **251** (68.2 %, log *K* = 8.74), the gene carrier with the headgroup of highest affinity **252** (21.7 %, log *K* = 9.22) performs less well than expected at a concentration of 0.12 mM. Probably, the extremely tight complexation of the nucleic acid by **252** is the cause for the precipitation at this concentration, which prevents higher transfection efficiency. This problem can probably be solved in future experiments by choosing a concentration between 0.03 mM (49.0 %) and 0.12 mM (21.7 %) as a compromise between efficiency and solubility. Nevertheless, these results clearly confirm the direct correlation between the affinity of the DNA carrier towards its cargo DNA and the resulting transfection efficiency, which was already observed during the screening of library **220** (cf. chapter 4.5.5).

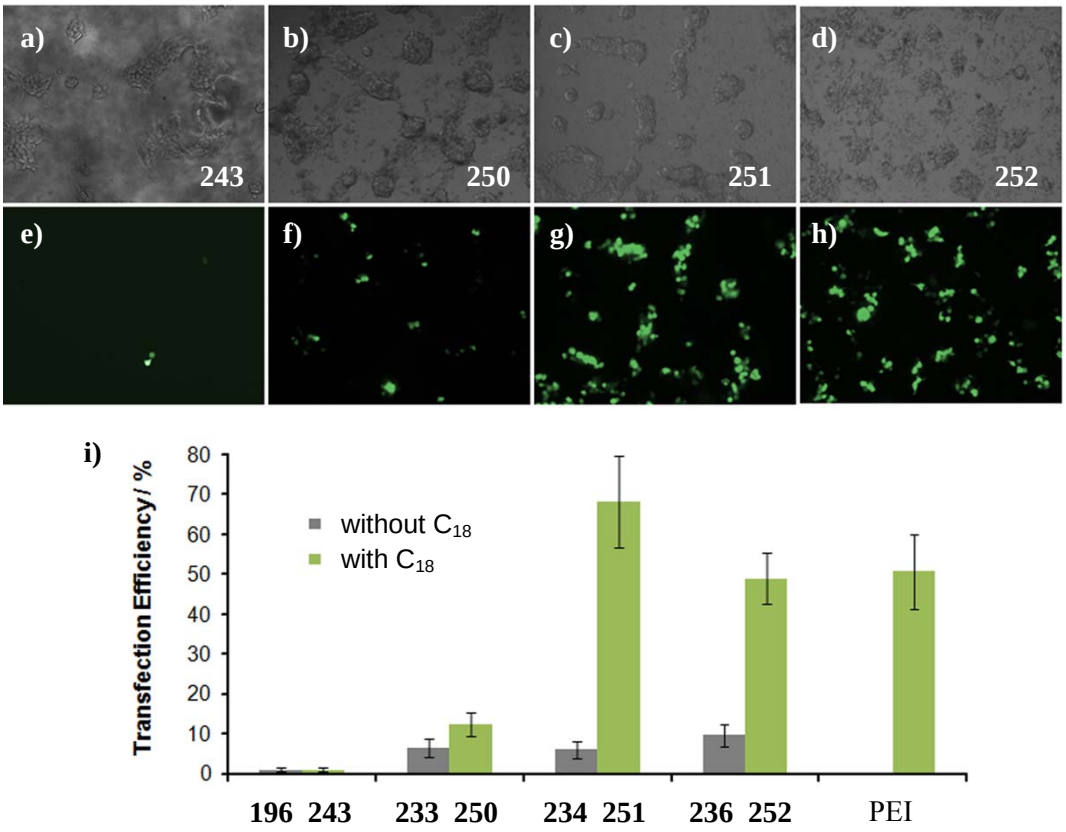


Figure 154. *Brightfield (a-d) and fluorescence images (e-h) of HEK293T cells 24 h (a, e) or 48 h (b-d, f-h) after transfection with pF143-GFP plasmid (2 µg) and* ***243*** *(0.0075 mM),* ***250***, ***251*** *(each 0.12 mM), or* ***252*** *(0.03 mM); i) Transfection efficiencies with C_{18} chains (****243***, ***250-252****, green) and without (****196***, ***233***, ***234***, ***236****, each 0.30 mM, grey).*

In conclusion, by combining combinatorial chemistry with rational design extraordinarily efficient gene carriers could be identified, which are even more efficient than the gold standard PEI. By screening a small but focused library of divalent peptide-based DNA ligands, sequences with high affinity for the pF143-GFP plasmid were found, which already featured moderate inherent gene carrier characteristics. By further derivatizing these sequences with hydrophobic alkyl chains extremely well performing amphiphilic transfecting agents were obtained and the minimum concentration, which is necessary for efficient transfection could be reduced by half. The direct correlation between the affinity of the amphiphile´s headgroup and its optimization via combinatorial chemistry is to the best of our knowledge unprecedented in literature.

5. SUMMARY

In the first part of this thesis we carried out Resonance Raman studies together with our cooperation partners from the working group of *Prof. Schlücker* concerning the binding event of a small molecular receptor with tetrapeptides. Therefore the receptor **13** and the four ligands **161**-**164** shown in Figure 155 were synthesized with the help of standard Fmoc SPPS in good to excellent yield and their binding properties were determined via RR. The data was analyzed by means of non-negative matrix factorization. Within the first preliminary studies we could thus reproduce the literature-known order of affinity **161** > **162**. Although more work is still on-going, already at this early stage these measurements can be viewed as proof-of-principle that we are not only able to qualitatively analyze a given binding event by means of RR, but that we are now also capable to quantitatively determine the binding constants of the complex formations.

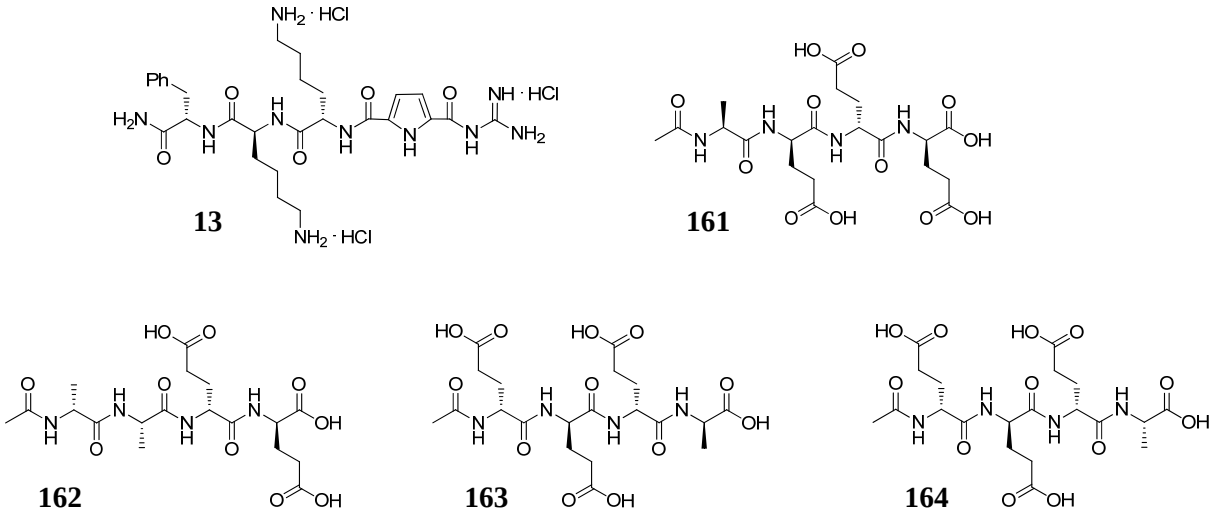
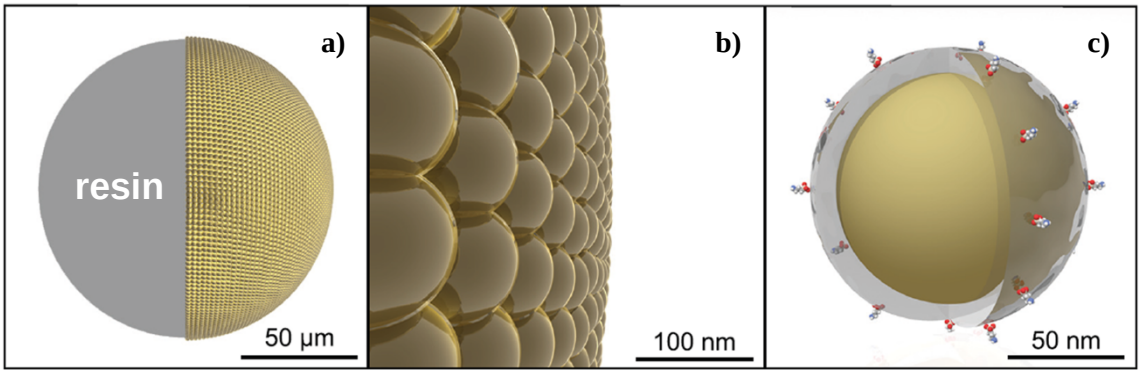
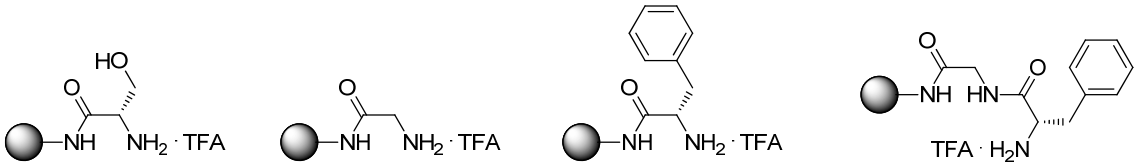


Figure 155. *Receptor **13** and substrates **161-164** for Resonance Raman binding studies.*

In the second Raman project which we carried out together with our cooperation partner we were able to prepare a novel bifunctional solid-support which combines the advantages of solid phase synthesis with rapid characterization of the bound substances via surface enhanced Raman spectroscopy. As depicted in Figure 156 gold nanoparticles were coated with an ultrathin silica layer and aggregated onto a solid support for SPPS thereby completely covering its surface. Derivatization with an amine function enabled the attachment of three different amino acids Gly, Phe, and Ser (**166**-**168**) as well as the dipeptide Gly-Phe (**171**) via standard Fmoc SPPS. The synthesis involved treatment of the designer beads with organic solvents like DMF or DCM as well as basic (20 % piperidine) or acidic conditions (33 % TFA). Afterwards it could be shown via microscopy techniques, that the gold NP covered bead surface was still intact thus ensuring that the bifunctional solid-support is suitable for

application in SPPS. The four compounds were then characterized and could be clearly distinguished from each other via SERS experiments very rapidly and reliably. The combined SPPS/SERS beads fulfill all necessary prerequisites for a manifold of applications, for instance the characterization of reaction intermediates during SPPS or the studying of molecular recognition events on-bead. This sophisticated nanoparticles-derivatized resin is an excellent addition to the fields of supramolecular chemistry and solid phase synthesis because it is one of the very few techniques that allows for direct and label-free analysis of resin-bound compounds.





166 **167** **168** **171**

Figure 156. *Design of the bifunctional beads which combine SPPS with SERS: a) Silica-encapsulated, 80 nm sized Au NPs on a 130 µm TentaGel bead; b) Zoom onto the surface; c) Amine functionalization of the silica-encapsulation with amino silane (image from: Gellner Ph.D.).*[191] *The three amino acids* ***166-168*** *and the dipeptide* ***171*** *were synthesized via SPPS and could be characterized and distinguished from each other via on-bead SERS.*

The next major part of the dissertation was concerned with the molecular recognition of biologically relevant substrates. Therefore the two non-symmetric (**187**, **188**) and the symmetric tweezer receptor **196** shown in Figure 157 were synthesized via microwave-assisted SPPS. In comparison to previous work conducted during my diploma thesis the synthesis of both templates, the orthogonally protected **181** of non-symmetric tweezers and **193** for symmetric hosts, could be improved. The overall yields over seven (**181**) and respectively six (**193**) steps could be augmented from 24 to 39 % and respectively from 29 to 42 %. Furthermore the yield and purity of the SPPS could be dramatically ameliorated for all receptors with the help of a single mode microwave. Additionally, at the optimized coupling conditions (max. 60° C, 20 min, 20 W) the time for the synthesis of e.g. **187** could be tremendously reduced from two weeks to merely two days. Furthermore, this second generation of host systems features improved physicochemical properties for binding studies in water at physiological pH. Especially the reduction of the acidity of the guanidinio group by approximately one order of magnitude in comparison to the first generation ensures a higher degree of protonation of the GCP oxoanion binding site at pH 7 and therefore improved binding properties towards anionic substrates.

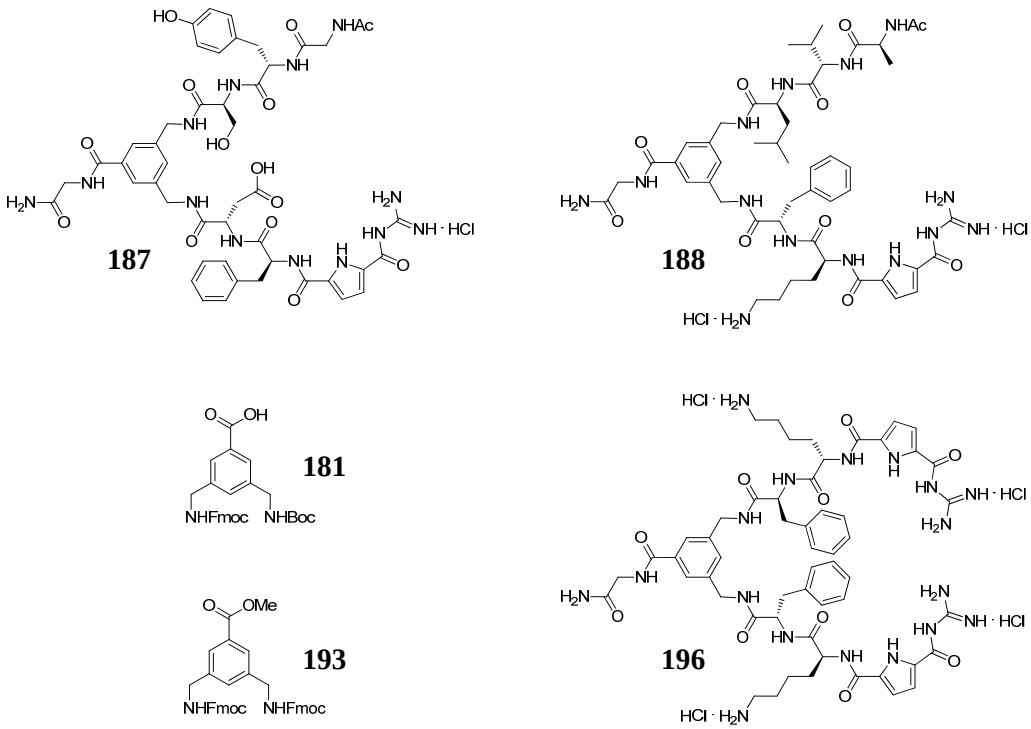


Figure 157. *Orthogonally protected template **181** for the microwave-assisted SPPS of non-symmetric tweezer receptors **187** and **188**. The scaffold **193** was utilized for the preparation of the symmetric tweezer **196**.*

UV/Vis titrations revealed that the non-symmetric tweezer receptors **187** and **188** feature binding constants of up to 6×10^3 M^{-1} in buffered water to the polar peptidic substrate N-Ac-Lys-D-Ala-D-Ala-OH, which plays a decisive role in the biosynthesis of the cell wall of *Gram* positive bacteria. It could be shown that the binding constants increase at lower pH (7 → 6 → 5). This finding could be linked to the higher degree of protonation of the GCP groups. The binding affinity mainly originates from the interaction between the GCP motif and the tripeptide´s carboxylate group. Furthermore it was observed that **187** has a higher affinity towards the substrate than **188** caused by the lower pK_A value of its GCP group and additional interactions of the receptor´s side chains with the tripeptide.

With the help of UV/Vis and fluorescence spectroscopy it could be shown that the symmetric tweezer **196** binds to nucleotides by forming 1:1 complexes in buffered water at neutral pH with good binding constants of up to 9×10^4 M^{-1}. Nucleotides are bound stronger than phosphate or diphosphate by almost one order of magnitude. Most remarkably, adenine monophosphate is preferred over its corresponding di- and triphosphate. This preference is as of yet unprecedented in literature and could be explained with the help of NMR spectroscopy and molecular modeling. The receptor forms salt-bridges to the highest charged α-phosphate of the nucleotide with its GCP units, which are oriented perpendicular to each other to allow for simultaneous binding to the tetrahedral phosphate anion. Moreover, the aromatic template stacks with the nucleobase thus locking the nucleotides in a well-defined geometry where the host is able to wrap around its substrate as a half-closed shell. As illustrated by the solvent accessible surfaces of the host-guest system depicted in Figure 158 the longer the phosphate hinge of the nucleotide (green), the more it is exposed to the solvent which is then better able to compete for binding. Consequently, the binding constants drop in the order AMP > ADP >

ATP. The specific interactions between the symmetric tweezer receptor **196** and phosphate and nucleotide anions is an excellent example for the importance to combine multiple non-covalent interactions in the correct orientation to allow for the formation of distinct host-guest complexes for the selective molecular recognition of the desired substrate.

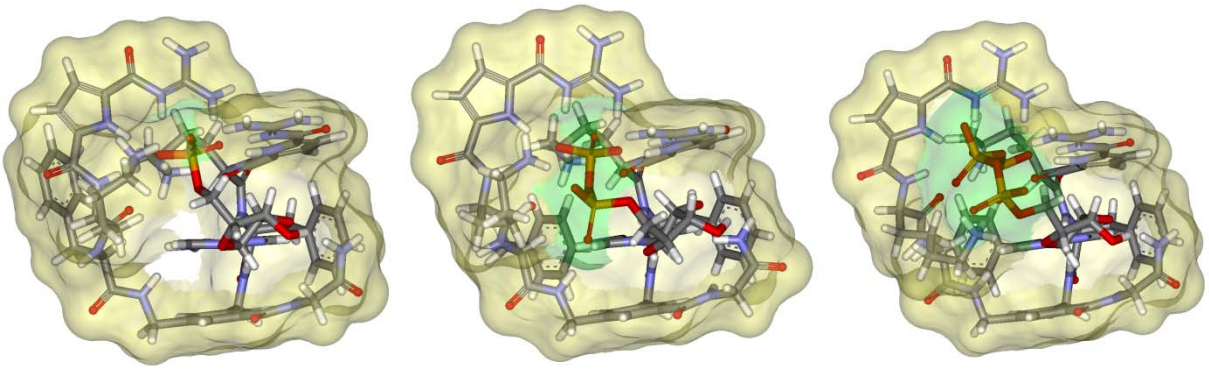
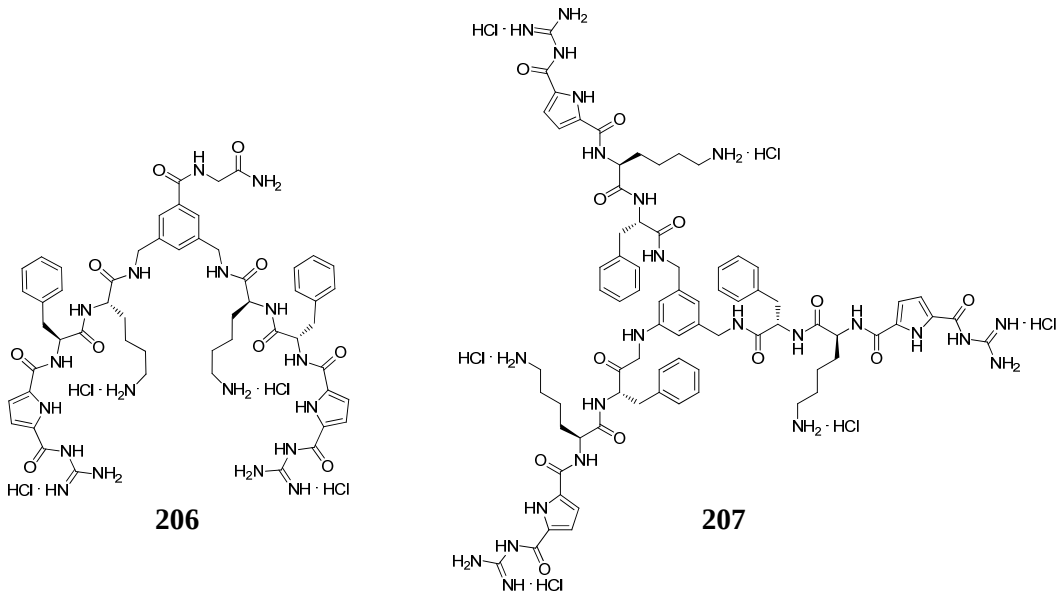


Figure 158. *Representation of the solvent-accessible surfaces of the calculated 1:1 complexes of receptor* ***196*** *(yellow) and the phosphate residue of the nucleotides (green). From left to right: AMP, ADP, and ATP.*

For the next project two more artificial systems were synthesized. As shown in Figure 159 the two-armed **206** features an inverted amino acid sequence in comparison to the parent compound **196**. It was prepared accordingly with the help of microwave-assisted SPPS. The tripodal **207** features the original amino acid sequence and was prepared by a convergent synthesis. First, the arms were prepared by standard Fmoc SPPS with all side chain protecting groups intact, coupled to the triamine scaffold and finally deprotected.



206 **207**

Figure 159. *Additional divalent ligand with inverted amino acid sequence* ***206*** *and trivalent ligand* ***207****.*

These three molecules were then tested for their binding properties towards various synthetic and generic nucleic acids with the help of thermal melting experiments, displacement assays, ITC and CD spectroscopy. The two-armed polynucleotide-ligands **196**

and **206** form highly stable complexes with binding constants of $K > 10^6$ M^{-1} and the trivalent ligand **207** binds even stronger by one order of magnitude ($K > 10^7$ M^{-1}). The binding mode involves electrostatic binding to the ribophosphate backbone paired with minor groove binding for B-type DNA and major groove binding for A-type DNA and RNA. Furthermore CD spectroscopy revealed a binding mode with two stacked pyrrole units which allowed differentiating between the secondary structure and sequence of different nucleic acids by means of the sign of the exciton CD signal. Such selective recognition of polynucleotides by non-covalently bound small molecule ligands is very rare and has potential for applications in polynucleotide sensing. Furthermore, it could be shown by means of AFM and DLS experiments that **196** and **207** are able to condense generic DNA in a concentration dependent fashion to very tight and densely packed spherical aggregates. Together with our cooperation partner, the working group of *Prof. Knauer*, we could demonstrate that these multimolecular complexes are taken up by human cells via endocytosis. However, only the trivalent ligand is able to escape from the endosome and transfect cells with genetic information. As indicated in Figure 160 this differentiation is due to **207**´s supreme buffering capacity in comparison to its two-armed analogue. The guanidinio group is getting protonated within the endosome upon specific ATPase-mediated proton influx and acts as proton sponge which ultimately leads to rupture of the vesicle caused by osmotic pressure. The three-armed gene carrier **207** is extremely efficient in transfecting cells. It outmatches the gold standard for non-viral vectors, PEI, in terms of efficiency and is at the same time less toxic. Furthermore, **207** is extraordinarily small in comparison to other cell penetrating peptides which normally feature 10-30 residues. It also carries fewer charges than comparable systems like oligo-Arg, which needs at least six positive charges for cellular uptake. Hence, this novel chemical vector offers great potential for the application in cell biology.

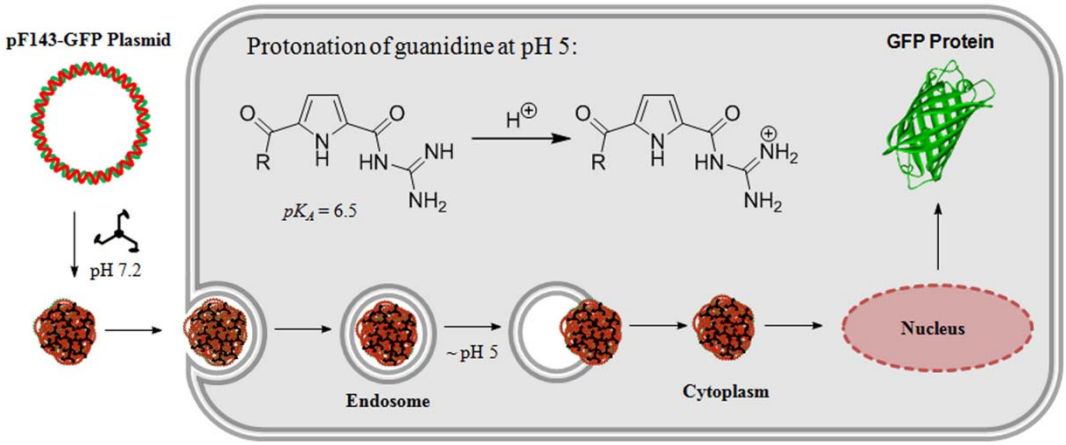
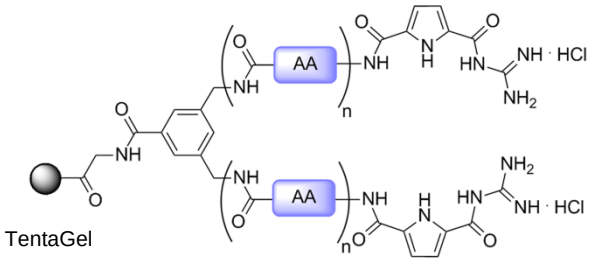


***Figure 160. 207**-mediated cellular uptake mechanism. The plasmid DNA is condensed into tight aggregates which are taken up via endocytosis. The GCP group acts as proton sponge in the endosome leading to its rupture. The genetic information is transported into the nucleus and ultimately the GFP protein is expressed.*

In order to improve the binding properties of the divalent ligands the combinatorial DNA-ligand library **220** shown in Figure 161 was generated by means of split & mix synthesis. This small but focused library comprised three combinatorial positions with six possible amino acids per position. The building blocks were of cationic (Lys, Arg) and aromatic character (Phe, Trp, Tyr) or anionic (Glu) as a control. Furthermore, the number of amino acids (n) and

thus the length of the arms was varied from 3 to 0. Accordingly, the library contained a total of 259 members ($6^3 + 6^2 + 6 + 1$). On-bead displacement assays were developed next for the screening of the library with two nucleic acids: the B-DNA $p(dAdT)_2$ with its well-defined, narrow minor groove and pF143-GFP plasmid DNA, which was utilized for above mentioned transfection experiments. With the help of the IRORI radiofrequency technology quantitative data could be obtained for all members of the library and not just for a few hit structures, which is regularly the case for large and random libraries.



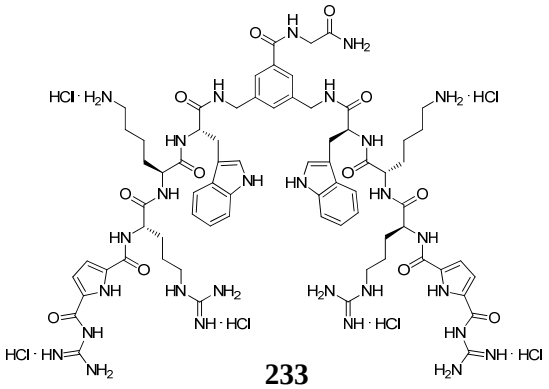
220

AA = Phe, Tyr, Trp, Lys, Arg, Glu

216	= 6 × 6 × 6	(n = 3)
36	= 6 × 6	(n = 2)
6		(n = 1)
1		(n = 0)
259 ligands		

Figure 161. *The library **220** consists of 259 symmetric, divalent DNA ligands with GCP headgroups in each side arm, which contain three, two, one, or no combinatorial positions (Phe, Tyr, Trp, Lys, Arg, and Glu).*

The screening of the library revealed high affinity sequences for both nucleic acids which clearly outmatched the original two-armed ligands. Additional charge-charge interactions were one reason for the improved binding properties, but the amino acid sequence also played a decisive role in binding to the nucleic acids. Sequences of the type AA^1-cationic-cationic-GCP were especially successful in binding to $p(dAdT)_2$, while those of the type AA^1-aromatic-AA^2-GCP were less well suited, probably due to the difficulty to arrange the bulky aromatic side chains in the minor groove of the nucleic acid. Concerning the length of the DNA binders a huge beneficial effect could be observed when going from one to two amino acids. Obviously a certain degree of flexibility is necessary within these ligands for efficient binding.



233

Figure 162. *Five library members were synthesized for characterization in solution. The ligand **233** with the amino acid sequence Trp-Lys-Arg-GCP is depicted exemplarily. The four other compounds were of the sequence Trp-Arg-Lys-GCP (**234**), Lys-Lys-Arg-GCP (**235**), Lys-Trp-Arg-GCP (**236**), and Glu-Trp-Arg-GCP (**237**).*

Five DNA-binders (**233**-**237**) were selected as cross-section of the library and synthesized in order to verify the results obtained from the on-bead screening. The ligand **233** is shown exemplarily in Figure 162. Solution-based displacement assays were in excellent agreement with the results obtained from solid phase. Furthermore, ITC measurements revealed binding

constants of $K > 10^9\ M^{-1}$. Hence, with the help of combinatorial chemistry the affinity of the divalent ligands could be increased by three orders of magnitude

The on-bead screening with the plasmid DNA revealed that there is no general correlation between efficient binding to $p(dAdT)_2$ and high affinity towards the plasmid. However, sequences containing the three amino acids Trp, Lys and Arg were very efficient in binding to both types of DNA. The highest observed binding constants were increased by a factor of 1,000 in comparison to the parent compounds ($K > 10^9\ M^{-1}$). Hence, with the help of combinatorial chemistry it was possible to optimize the binding properties of the divalent ligands by large and to fine-tune them for specific nucleic acids. Finally, we could show that the increased binding affinity to the plasmid also correlated with higher transfection efficiencies. Although the so derived, novel two-armed gene carriers were not as efficient as the trivalent ligand **207** with its superior buffering capacity, this finding was of great use at it revealed the unexpected dependency of the transfection efficiency from the binding affinity.

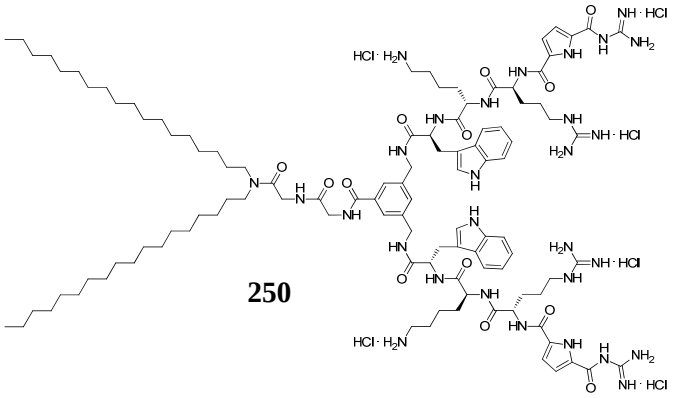


Figure 163. *Four amphiphilic gene carriers were sythesized for cell transfection. The one with the sequence Trp-Lys-Arg-GCP (**250**) is depicted exemplarily. The three others were of the sequence Trp-Arg-Lys-GCP (**251**), Lys-Trp-Arg-GCP (**252**), and Phe-Lys-GCP (**243**).*

In order to further exploit the improved ability to transfect cells, three high affinity sequences were derivatized with octadecyl chains to generate the gene carriers **250**-**252**. One of them is shown exemplarily in Figure 163. Furthermore, an amphiphilic analogue (**243**) of the less affine parent compound **196** was synthesized for comparison. The amphiphilic ligands were prepared in a convergent synthesis: First, the fully side chain protected DNA-binders were synthesized with the help of microwave-assisted SPPS, coupled with C_{18}-derivatized glycine at their free C-termini, and finally deprotected to obtain the amphiphilic carriers. The following cell transfection experiments confirmed the correlation between the affinity of the amphiphile´s headgroup towards the plasmid DNA and the efficiency as gene carrier. **251** was identified as extremely efficient transfecting agent, which even outperformes the gold standard PEI. Furthermore, the minimum concentration for efficient transfection could be reduced by half.

In conclusion, within the framework of this dissertation new insight into molecular recognition events of biologically important targets like peptides, nucleotides and nucleic acids could be gathered. These findings do not only help to improve our general understanding of the complex processes which take place in biological systems but could also be utilized for an application in cell biology by generating a novel approach for the development of highly efficient gene carriers.

6. OUTLOOK

After having shown that it is principally possible to calculate binding constants for molecular recognition events from Resonance Raman experiments, the as of yet preliminary data has to be validated and its reproducibility has to be ensured. Concerning the novel SERS/SPPS beads, the next step could be to test if the GCP oxoanion binding motif is SERS active when bound to the designer solid support (**253**). If this is the case it would then be possible to determine whether a molecular recognition event, e.g. with simple carboxylate anions like acetate (**254**), can be visualized on-bead as indicated in Figure 164. By recording the SERS answers upon binding of different substrates and comparing them with their solution-derived binding constants it would be possible to analyze whether this method is sensitive for binding affinity. The SPPS/SERS beads could then be used for the development and consequent label-free on-bead screening of receptor libraries. Furthermore, the detection of more complicated host-guest systems of increasing size, e.g. with one or two additional amino acids, or even more complex molecules could be tested to explore the scope and limitations of this novel technique.

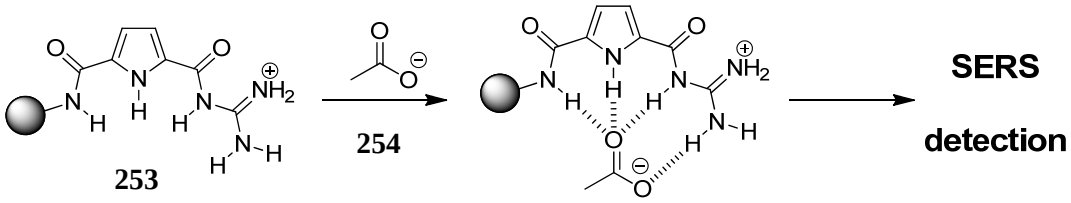
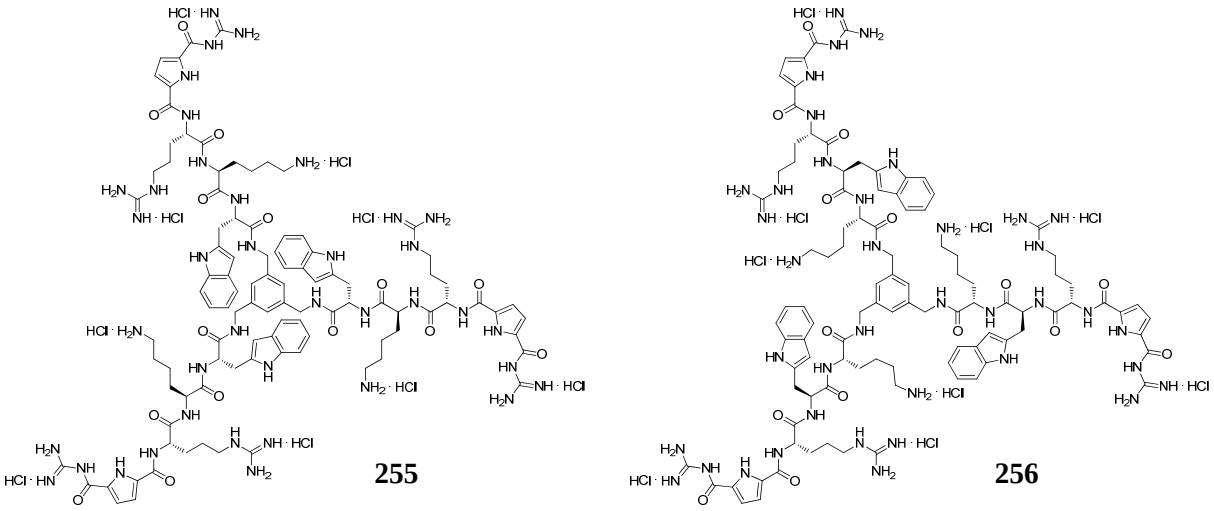


Figure 164. *By attaching the GCP oxoanion binding motif onto the novel bifunctional SPPS/SERS beads (**253**) binding events of small carboxylates like acetate (**254**) can be directly visualized on-bead via SERS.*

With regard to the combinatorial library **220** a screening for nucleotides could be conducted next. Since the parent symmetric tweezer compound **196** is already able to selectively recognize nucleotides with good binding constants, even more efficient hosts could thus be identified. Furthermore, other types of nucleic acids, for instance siRNAs could be subjected to a screening in order to identify high affinity sequences to this completely different class of polynucleotides.[265] The so identified sequences could then be tested for their efficiency in transporting siRNAs into cells and thus trigger the RNAi mechanism.

In general, the library-derived amino acid sequences, which are optimized for binding ds-DNA, could be implemented into three-armed ligands of the type **207**, which has already proven to be an excellent gene carrier for plasmid DNA. The divalent high affinity ligands already provided better transfection efficiencies than their parent compounds **196** and **206**. Thus, by attaching the improved sequences to a trivalent scaffold according to the procedure developed during this doctoral work, buffering capacity would be introduced into these systems and therefore the generation of extremely efficient transfecting agents is likely. By preparing for example the gene carriers **255** and **256** shown in Figure 165 a direct comparison

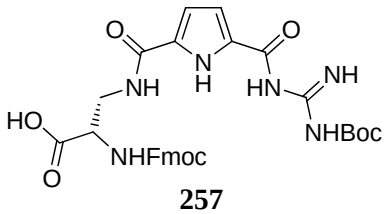
with their already synthesized two-armed analogues **233** and **236** would be possible which would allow for validating the general applicability of this approach.



255 **256**

Figure 165. *By implementing the library-derived high affinity amino acid sequences into three-armed DNA-ligands (**255**, **256**) highly efficient gene carriers should be accessible.*

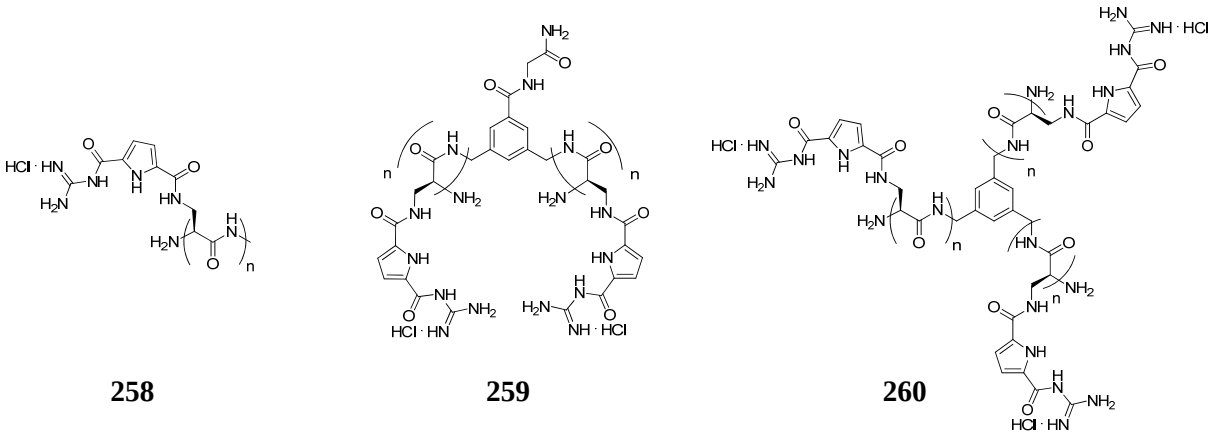
In general, the buffering capacity of the GCP unit seems to be crucial for efficient gene transfer. Therefore, the generation of a second generation DNA-ligand library with GCP units also present within the combinatorial positions could help to identify even more efficient gene carriers. The GCP unit can easily be implemented in the form of the building block **257** shown in Figure 166. The synthesis of this compound is already known and it has already been utilized for the preparation of compounds in standard Fmoc-SPPS syntheses.[266]



257

Figure 166. *The building block **257** allows for implementation of GCP units into combinatorial positions within the side chains of di- or trivalent DNA ligands. Furthermore, oligo-GCP units can thus be synthesized for the development of even more efficient gene carriers.*

Furthermore, with the help of this building block the preparation of oligo-GCP units by means of standard Fmoc SPPS is possible. Such systems could be very efficient small molecule gene carriers combining the propensity to bind to DNA and the capability to escape from the endosome. Structure-activity studies could be carried out concerning their minimum length and the effect of multivalency by preparing linear DNA-ligands of the type **258** depicted in Figure 167 and by implementing them into di- (**259**) or trivalent systems (**260**) similar to those developed during this doctoral work.



258 **259** **260**

Figure 167. *Linear (**258**), divalent (**259**) and trivalent (**260**) oligo-GCP compounds as transfecting agents.*

In conclusion, the work conducted during this doctoral thesis has laid the foundation for the development of more efficient gene carriers by developing small molecule DNA-binders and elucidating their uptake mechanism into human cells. This knowledge suggests to enable further enhancement of the characteristics of non-viral vectors in future research. This will help to render gene therapy into a broadly applicable standard method for the treatment of a wide range of diseases.

7. EXPERIMENTAL SECTION

7.1 GENERAL EXPERIMENTAL AND ANALYTICAL METHODS

Solvents and Chemicals

All solvents were dried according to literature procedures.[267] Diethyl ether was distilled from sodium with benzophenon as indicator. DCM and CCl_4 were distilled from calcium hydride. Methanol was distilled. DMF was distilled from calcium hydride in vacuum (the first 250 mL were discarded). Water was purified with a *TKA* MicroPure ultrapure water system. Unless otherwise stated, all other solvents and chemicals were utilized as supplied from the manufacturer.

Inert Gas

Reactions under inert gas were carried out with Argon (99.998 %) from *Air Liquide* which was dried with Silica Gel Orange.

Orbital Shaker for Peptide Synthesis

The solid phase peptide synthesis was carried out on a Yellowline OS 5 basic orbital shaker from *Heidolph*.

Magnetic Stirrer

Reactions were stirred and heated with the help of magnetic stirring bars and a MR 3001 K magnetic stirrer from *Heidolph*.

Single Mode Microwave

Mircowave assisted SPPS was carried out with a *CEM* Discover single mode microwave.

Equipment for Combinatorial Synthesis

The synthesis of the DNA-ligand library was performed with the help of the AccuTag-100 Combinatorial Chemistry System from *IRORI*. The individual members were synthesized in MikroKans^{TM} which were equipped with radiofrequency tags. A scanning station was used to sort and identify the library members.

Thin layer Chromatography (TLC)

Analytical TLC was carried out in order to monitor reactions on aluminium foils ALUGRAM SIL G/UV_{254} and C18 SiO_2 aluminium foils ALUGRAM RP-18W/UV_{254} from *Macherey-Nagel*. The spots were visualized by means of fluorescence quenching upon irradiation with UV light of 254 nm or with the help of a ninhydrin (0.5 g) solution in acidic ethanol (EtOH/AcOH/TFA = 90/10/1).

Column Chromatography

Flash chromatography was done on columns packed with Silica gel 60M with a spherical size of 40-63 μm from *Macherey-Nagel*. Organic solvents were distilled prior to use.

Preparative MPLC

Purification of polar substances was conducted with the help of reversed phase preparative MPLC on a Liquid Chromatography Flash apparatus from *Armen Instrument*. Glass columns were packed with *YMC* Gel ODS-A or Gel ODS octyl C8 of 50 µm diameter and 12 nm pore size. Distilled or HPLC grade solvents and ultrapure water were used as eluents. Detection of the products was achieved with the help of a UV detector.

Analytical HPLC

The purity of the compounds was determined with the help of an HPLC apparatus from *Dionex* containing the following components: P680 HPLC-pump, ASI-100 Automated Sample Injector and UVD 340U detector. A Supelcosil^{TM} LC-18 column (25 cm × 4.6 mm, 5 µm) from *Supelco* or a *YMC* ODS-A column (15 cm × 3.0 mm, 5 µm) was utilized. Ultrapure water and HPLC-grade solvents were used as eluents. Detection was achieved with the help of a UV detector.

Preparative HPLC

Separation of multimilligram amounts of the products were performed after pre-purification via MPLC on a *Merck* SepTech preparative HPLC apparatus equipped with a *Knauer* Dynamic Mixing Chamber and a *Shimadzu* SPD-10A UV detector. A *YMC* ODS-A column (15 cm × 3.0 cm, 5 µm) was utilized. The conditions were optimized via analytical HPLC on the corresponding *YMC* column (vide supra) and the amount of substance per injection and the elution speed (ml/min) were scaled-up by a factor of 100.

Lyophilization

Freeze drying of aqueous solutions was carried out with an Alpha 1-4 LD plus apparaturs from *Christ*.

Vacuum Pumps

Vacuum was generated with a *Vacuubrand* CVC3000 pump. High vacuum was generated with a RC6 *Vacuubrand* chemistry hybrid pump.

Rotary Evaporator

Distillation of solvent from the reaction mixture was carried out under reduced pressure with a *Heidolph* Hei-VAP Advantage rotary evaporator at 40° C.

Melting Point Determination

Melting points were measured with open end glass capillary tubes without correction with a *Büchi* Melting Point B-540.

pH Measurement

The pH was determined with the pH-Meter 766 Calimatic from *Knick*, which was calibrated with commercial available buffer standards (pH = 4.00 and 7.00).

Nuclear Magnetic Resonance (NMR)

700, 500, 400, and 300 MHz ^{1}H and respectively 175, 125, 100, and 75 MHz ^{13}C-NMR spectra were recorded with an Avance II+ 700, a DRX 500, an AC400 and a DMX300 from *Bruker* at 25° C. Chemical shifts are reported as δ values in ppm. The spectra were calibrated with the help of the residual peaks of the deuterated solvents as internal standard: DMSO-d_6 (δ [^{1}H] = 2.50 ppm, δ [^{13}C] = 39.52 ppm), MeOD-d_4 (δ [^{1}H] = 3.31 ppm, δ [^{13}C] = 49.05 ppm), or $CDCl_3$ (δ [^{1}H] = 7.26 ppm, δ [^{13}C] = 77.16 ppm). Coupling constants are

reported in Hz. For the signal multiplicity the following abbreviations are utilized: br. = broad, s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet. All assignments have been performed according to literature.[268]

Fourier Transform Infrared Spectroscopy (FT-IR)

Infrared spectra were recorded on a FT-IR 430 from *Jasco* with the help of a MIRacle ATR unit from *Pike Technologies* or as KBr disks on a FT-IR 1600 from *Perkin-Elmer*. The bands are reported in cm^{-1} and the following abbreviations have been utilized for the characterization of the band intensities: br. = broad, s = strong, m = middle, w = weak.

UV/Vis Spectroscopy

UV/Vis spectra were recorded with a *Jasco* V-660 spectrometer in standard quartz microcuvettes of 1 cm width which were equipped with stoppers. All spectra are corrected for the background (empty cuvette) unless otherwise stated.

Fluorescence Spectroscopy

UV/Vis spectra were recorded with a *Varian* Carey Eclipse spectrometer in standard quartz fluorescence microcuvettes of 1 cm width which were equipped with stoppers or in sterile Nunclon Surface 96-well plates from *Nunc*.

Isothermal Titration Calorimetry

ITC measurements were carried out with a VP-ITC Microcalorimeter from MicroCal. The reference cell was filled with ultrapure, degassed water. All substances were dissolved in degassed, ultrapure buffered water. The data were analyzed with the help of Origin 7 from *Origin Lab* and the corresponding ITC plugin from *MicroCal*.

Atomic Force Microscopy

AFM images were recorded in tapping mode with an Innova NanoDrive AFM controller and an Innova Atomic Force Microscope from *Veeco* with N-doped silica cantilevers (AC 160TS OLYMPUS). The solutions were spin coated onto freshly cleaved Mica surface from *Plano*. The data were analyzed with Gwiddion-2.20.

Dynamic Light Scattering

DLS measurements were conducted with a Zetasizer Nano-ZS from *Malvern Instruments GmbH* and analyzed with the software provided from the manufacturer.

Force Field Calculations

Computational force field calculations were carried out with Macromodel v9.6 from *Schrödinger*. For small host-guest systems a 50,000 step conformational analysis (Monte Carlo) was carried out first, making use of the Amber* force field and a generalized Born/surface area (GBSA) water solvation, followed by a molecular dynamics simulation at 300 K for 500 ps. Systems containing nucleic acids were calculated with the help of the OPLC_2005 force field an GBSA water solvation by minimizing the structure without restraints first, followed by a conformational analysis with frozen DNA and 10,000 steps.

Liquid Extraction

Extraction of caffeine was conducted with a PrimaDonna apparatus from *De´Longhi*.

Centrifuge

Precipitates were centrifuged with a Rotofix 32 from *Hettich*.

Mass Spectrometry

ESI mass spectra were recorded with a *Bruker* BioTOF III spectrometer. EI MS was conducted with a MAT 90 from *Finnigan*.

Vortexer

Isopycnic solutions were vortexed prior to transfer to 96-well plates with a Vortex-Genie 2 from *Scientific Industries*.

CD Spectroscopy

CD spectra were recorded with a J-815 spectrometer from *Jasco* in quartz fluorescence cuvettes of 1 cm width which were equipped with stoppers. All spectra are corrected for the background (cuvette + buffer).

Oven

All glass ware was dried in a Function Line oven from *Heraeus* before usage.

7.2 GENERAL PROCEDURES FOR SPPS

Solid phase peptide synthesis was carried out in Schlenck glass vessels equipped with a frit and a stopper. Reactions were conducted under argon atmosphere while shaking the reaction mixture on a shaker. After completion of a reaction step, the solvent was removed via filtration. The amount of solvent used for the reaction steps was chosen in a way so that the concentration of the reactants was between 0.1 and 0.2 M.

Kaiser test

Two solutions were prepared: Ninhydrin (1.00 g) and respectively phenol (40 g) were dissolved in ethanol (10 mL). A few beads of the analyte were suspended in a mixture of 0.3 mL of each of the two solutions and heated to 110° C for 1-2 min. Consequently, beads with free amine functionalities were coloured deeply blue while resin without free amino functions stayed colourless.

Fmoc Deprotection

In order to remove Fmoc protecting groups the resin was shaken with 20 % piperidine in DMF two times for 20 min and then thoroughly washed six times for 5 min with DMF.

Attachment

The attachment of the first compound to the resin differed depending on the reactive groups present on the solid support. Resins featuring hydroxy groups (Wang, SASRIN) were first swollen in a mixture of DCM/DMF (7/3). The first N-Fmoc protected amino acid (2 eq) was then attached to the solid support via its free C-terminus. The resin (1 eq) and the amino acid (2 eq) were suspended in DCM/DMF (7/3) and then DIC (2 eq) was slowly added dropwise under argon atmosphere, followed by HOBt (2 eq) and DMAP (0.2 eq). The mixture was shaken for at least 8 h and washed three times for 5 min with DCM/DMF (7/3). The coupling and washing steps were repeated four times in total. After the last repetition of the coupling, the washing step was carried out with DMF. The solid support was then treated with Ac_2O in 3-5 % NMM or DIPEA in DMF under argon atmosphere for 1 h. Finally, the resin was washed three times with DMF for 5 min.

Resins featuring amine groups (Fmoc-Rink-Amide, MBHA) were swollen in DMF for 1 h. The Rink-Amide´s Fmoc protecting group was removed as described above. The first amino acid (3 eq) was attached to the resin with PyBOP (3 eq) in 3-5 % NMM or DIPEA in DMF under argon atmosphere while shaking the reaction mixture for at least 8 h. The solid support was then washed three times for 5 min with DMF. The coupling and washing steps were repeated. The solid support was then treated with Ac_2O in 3-5 % NMM or DIPEA in DMF under argon atmosphere for 1 h. Finally, the resin was washed three times for 5 min with DMF.

Coupling

The resin (1 eq) was suspended in 3-5 % NMM or DIPEA in DMF and then amino acid (3 eq) and PyBOP (3 eq) were added. The reaction mixture was shaken for at least 8 h under argon atmosphere. The solid support was then washed three times for 5 min with DMF.

Cleavage

Before cleaving the raw product from the solid support, the beads were thoroughly washed with DCM (3 × 5 min), MeOH (3 × 5 min), and again three times with DCM (3 × 5 min). The resin was then dried in vacuum for 1 h. The cleavage was carried out as follows:

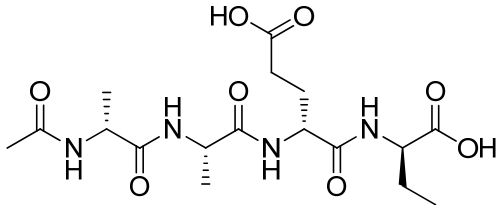
- **SASRIN**: The resin was treated with 1 % TFA containing 2.5 % TIS and 2.5 % H_2O in DCM four times for 2 h and once more for 20 h and then washed three times with DCM. The filtrates were immediately neutralized with the corresponding amount of NEt_3.
- **Wang**: The solid support was shaken for 4 h in 50 % TFA in DCM containing 5 % TIS under argon atmosphere and washed three times for 5 min with the cleavage cocktail.
- **Fmoc-Rink-Amide**: The polymer was treated with 95 % TFA containing 2.5 % TIS and 2.5 % H_2O for 4 h under argon atmosphere and washed three times with the cleavage mixture.
- **MBHA**: The beads were suspended in 80 % TFA containing 8 % thioanisole and 4 % EDT. 8 % TFMSA were slowly added under argon atmosphere, the resin was shaken for 4 h, and then washed three times with TFA for 5 min.

The filtrates were combined and the solvent reduced to a few mL by distillation under reduced pressure. The raw product was then precipitated according to its polarity by adding either 40 mL of water or Et_2O, centrifuged and the supernatant solvent decanted. The remaining solid was dissolved in little MeOH, water was added and the mixture was freeze dried under reduced pressure. The raw product was then purified by reversed phase MPLC.

7.3 SYNTHESIS OF SMALL OLIGOPEPTIDE SUBSTRATES

The SPPS of the tetrapeptides **162-164** was carried out on 250-300 mg Wang resin (1.63 mmol/g) as described above. 3 % NMM was utilized as base during the coupling steps. The raw products were precipitated with Et_2O. Purification was carried out via RP MPLC (H_2O → acetonitrile, 0.1 % TFA) and RP HPLC (H_2O/acetonitrile, 0.05 % TFA).

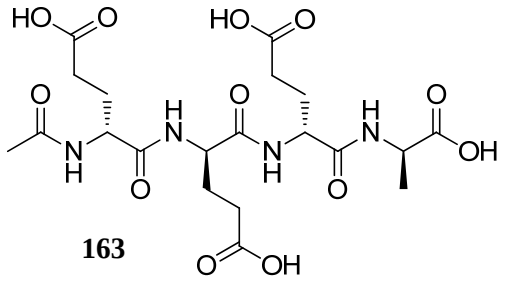
N-Ac-D-Ala-L-Ala-D-Glu-D-Glu-OH (162)



162

$C_{18}H_{28}N_4O_{10}$	460.44 g/mol
Yield	83 %
Mp.	102.7° C
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 1.16 (d, ^3J = 7.05 Hz, 3H, Ala-CH_3), 1.21 (d, ^3J = 7.14 Hz, 3H, Ala-CH_3), 1.68-2.00 (m, 7H, Ac-CH_3, 2 × CH_2), 2.14-2.32 (m, 4H, 2 × CH_2), 4.15-4.21 (m, 3H, 3 × C*H*), 4.28 (dt, ^3J = 8.72 Hz, ^3J = 5.04 Hz, Glu-C*H*), 7.86 (d, ^3J = 8.24 Hz, 1H, N*H*), 8.07 (d, ^3J = 6.47 Hz, 1H, N*H*), 8.11 (d, ^3J = 7.61 Hz, 1H, N*H*), 8.22 (d, ^3J = 7.21 Hz, 1H, N*H*), 12.22 (br. s, CO_2H) ppm.
^{13}C-NMR	**(125 MHz, DMSO-d_6):** δ = 17.54 (*C*H_3), 17.88 (*C*H_3), 22.14 (*C*H_3), 26.02 (*C*H_2), 27.06 (*C*H_2), 29.91 (*C*H_2), 48.41 (*C*H), 48.54 (*C*H), 51.12 (*C*H), 51.47 (*C*H), 169.45 (*C*O), 171.10 (*C*O), 171.98 (*C*O), 172.41 (*C*O), 173.00 (*C*O), 173.67 (*C*O), 173.90 (*C*O) ppm.
HR-MS	**(pos. ESI-ToF):** m/z = 459.1663 $[M - H^+]^-$ calculated for $C_{18}H_{27}N_4O_{10}$: 459.1733
FT-IR	**(ATR):** $\tilde{\nu}$ = 3290 (br. w), 2938 (w), 1716 (m), 1634 (s), 1539 (s), 1447 (m), 1375 (m), 1168 (s), 787 (m) cm^{-1}.

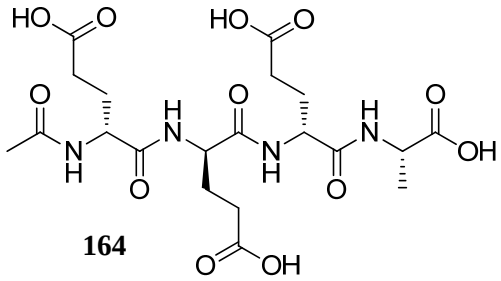
N-Ac-D-Glu-D-Glu-D-Glu-D-Ala-OH (163)



163

$C_{20}H_{30}N_4O_{12}$	518.47 g/mol
Yield	53 %
Mp.	228° C (decomposition)
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 1.28 (br. s, 3H, Ala-CH_3), 1.66-1.97 (m, 9H, Ac-CH_3, 3 × CH_2), 2.25 (br. s, 6H, 3 × CH_2), 4.09-4.34 (m, 4H, 4 × CH), 7.89-7.99 (m, 2H, 2 × NH), 8.01 (d, ^3J = 7.25 Hz, 1H, NH), 8.04 (d, 1H, ^3J = 7.23 Hz, NH), 8.17 (br. s, 1H, NH), 12.15 (br. s, 4H, 4 × CO_2H) ppm.
^{13}C-NMR	**(125 MHz, DMSO-d_6):** δ = 16.92 (Ala-CH$_3$), 22.46 (Ac-CH$_3$), 27.41 (Glu-CH$_2$), 27.50 (Glu-CH$_2$), 27.59 (Glu-CH$_2$), 30.00 (Glu-CH$_2$), 30.10 (Glu-CH$_2$), 30.27 (Glu-CH$_2$), 51.63 (Ala-CH), 51.87 (Glu-CH), 51.96 (Glu-CH), 52.12 (Glu-CH), 169.61 (CO), 170.75 (CO), 170.80 (CO), 170.91 (CO), 171.02 (CO), 171.43 (CO), 173.50 (CO), 173.92 (CO) ppm.
HR-MS	**(neg. ESI-ToF):** m/z = 258.0865 $[M - 2\ H^+]^{2-}$ calculated for $C_{20}H_{28}N_4O_{12}^{2-}$: 258.0858
FT-IR	**(ATR):** $\tilde{\nu}$ = 3328 (br. w), 3218 (w), 2945 (w), 1731 (s), 1601 (s), 1547 (s), 1237 (w), 1442 (w), 1408 (m), 1350 (w), 1295 (w), 1275 (m), 1167 (s), 1063 (w), 1018 (w), 949 (w), 892 (w), 790 (s), 621 (s) cm^{-1}.

N-Ac-D-Glu-D-Glu-D-Glu-L-Ala-OH (164)



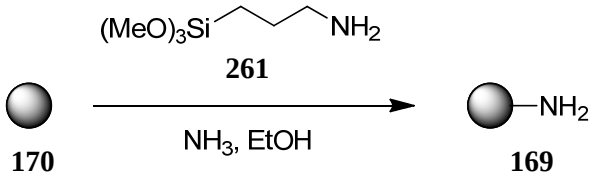
164

$C_{20}H_{30}N_4O_{12}$	518.47 g/mol
Yield	76 %
Mp.	181.6° C
^{1}H-NMR	**(300 MHz, DMSO-d_6):** δ = 1.25 (d, ^3J = 7.24 Hz, 3H, Ala-CH_3), 1.63-1.97 (m, 9H, Ac-CH_3, 3 × CH_2), 2.13-2.32 (m, 6H, 3 × CH_2), 4.13-4.34 (m, 4H, 4 × C*H*), 7.91 (d, ^3J = 8.01 Hz, 1H, N*H*), 7.98-8.16 (m, 3H, 3 × N*H*), 12.28 (br. s, 4H, 4 × CO_2H) ppm.
^{13}C-NMR	**(75 MHz, DMSO-d_6):** δ = 17.21 (Ala-*C*H_3), 22.37 (Ac-*C*H_3), 26.89 (Glu-*C*H_2), 27.08 (Glu-*C*H_2), 27.40 (Glu-*C*H_2), 29.94 (Glu-*C*H_2), 30.03 (Glu-*C*H_2), 30.10 (Glu-*C*H_2), 47.39 (Ala-*C*H), 51.80 (Glu-*C*H), 51.87 (Glu-*C*H), 52.01 (Glu-*C*H), 169.55 (*C*O), 170.60 (*C*O), 170.79 (*C*O), 171.38 (*C*O), 173.68 (*C*O), 173.80 (*C*O), 173.84 (*C*O), 173.88 (*C*O) ppm.
HR-MS	**(neg. ESI-ToF):** m/z = 258.0847 $[M - 2\ H^+]^{2-}$ calculated for $C_{20}H_{28}N_4O_{12}^{2-}$: 258.0858
FT-IR	**(ATR):** $\tilde{\nu}$ = 3338 (br. w), 1730 (s), 1601 (s), 1542 (s), 1443 (w), 1409 (m), 1295 (w), 1273 (m), 1169 (s), 1062 (w), 1016 (w), 949 (w), 893 (w), 790 (s) cm^{-1}.

7.4 DERIVATIZATION OF AU-NANOPARTICLES

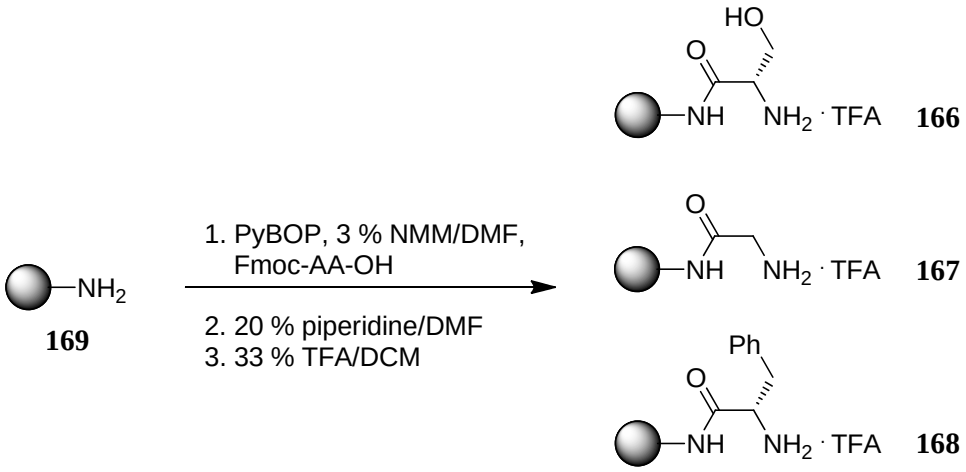
Silica-encapsulated Au nanoparticles were obtained aggregated onto TentaGel S NH_2 beads (**170**) from *Gellner* from the working group of *Prof. Schlücker* at the Institute for Physics at the University of Osnabrück.

Amino Functionalization (169)



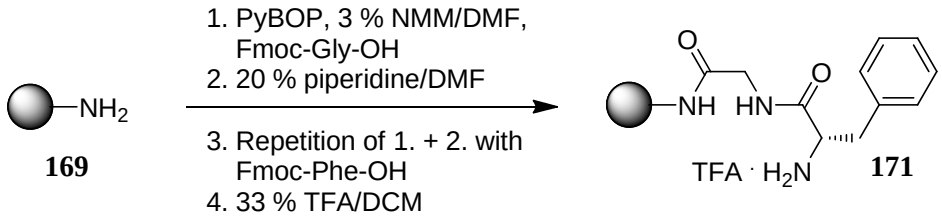
Approximately 100 beads of **170** were suspended ethanol (1 mL). 30 % aqueous ammonia (50 µL) and a solution of 0.5 % APTMS (**261**) in EtOH (50 µL) were added under argon atmosphere and the mixture was shaken for 30 min. The solid support was washed three times for 5 min with DMF (1 mL) to obtain the functionalized beads **169**.

Attachment of Amino Acids (166-168)



Approximately 20 beads of **169** were suspended in 3 % NMM/DMF (1 mL). PyBOP (16 mg, 0.03 mmol) and one of the three amino acids Fmoc-Ser-OH (12 mg, 0.03 mmol), Fmoc-Gly-OH (9 mg, 0.03 mmol), or Fmoc-Phe-OH (12 mg, 0.03 mmol) were added and the mixture was shaken for at least 8 h. The resin was washed three times for 5 min with DMF and the Fmoc group was deprotected by treating the resin with 20 % piperidine/DMF two times for 20 min (1 mL). Afterwards the beads were washed six times for 5 min with DMF (1 mL). The coupling, deprotection and washing steps were repeated. Finally, the solid support was treated with 33 % TFA/DCM and washed with DCM (3 × 5 min), MeOH (3 × 5 min), and water (3 × 5 min) to obtain the derivatized beads **166-168**.

Attachment of the Dipeptide Gly-Phe (171)



Approximately 20 beads of **169** were suspended in 3 % NMM/DMF (1 mL). PyBOP (16 mg, 0.03 mmol) and Fmoc-Gly-OH (9 mg, 0.03 mmol) were added and the mixture was shaken for at least 8 h. The resin was washed three times for 5 min with DMF and the Fmoc group was deprotected by treating the resin with 20 % piperidine/DMF two times for 20 min (1 mL). Afterwards the beads were washed six times for 5 min with DMF (1 mL). The coupling, deprotection and washing steps were repeated. The second amino acid, Fmoc-Phe-OH (12 mg, 0.03 mmol) was coupled accordingly. Finally, the solid support was treated with 33 % TFA/DCM and washed with DCM (3 × 5 min), MeOH (3 × 5 min), and water (3 × 5 min) to obtain the dipeptide-derivatized beads **171**.

7.5 GENERAL PROCEDURES FOR MICROWAVE ASSISTED SPPS

Microwave-assisted SPPS was carried out in polyethylene tubes equipped with a frit. Reactions were conducted while bubbling argon through the solution. After completion of a reaction step, the solvent was removed via filtration. The amount of solvent used for the reaction steps was chosen in a way so that the concentration of the reactants was between 0.1 and 0.2 M.

Fmoc Deprotection

In order to remove Fmoc protecting groups the resin was treated with 20 % piperidine in DMF two times for 1 and respectively 5 min under argonatmophere while irradiating the mixture with microwaves (20 W, max. 60° C). The solid support was then thoroughly washed six times for 2 min with DMF.

Attachment

For hydroxy resins (Wang, SASRIN) the attachment of the first compound to the resin was carried out as described in chapter 7.2. Resins featuring amine groups (Fmoc-Rink-Amide, MBHA) were swollen in DMF for 1 h. The Rink-Amide´s Fmoc protecting group was then removed as described above. The first amino acid (3 eq) was attached to the resin (1 eq) with PyBOP (3 eq) in 3-5 % NMM or DIPEA in DMF under argon atmosphere while irradiating the mixture with microwaves for 20 min (20 W, max. 60° C). The solid support was then washed three times for 2 min with DMF. The coupling and washing steps were repeated. The solid support was subsequently treated with Ac_2O in 3-5 % NMM or DIPEA in DMF under argon atmosphere while irradiating the mixture with microwaves for 20 min (20 W, max. 60° C). Finally, the resin was washed three times for 2 min with DMF.

Coupling

The resin (1 eq) was suspended in 3-5 % NMM or DIPEA in DMF and then amino acid (3 eq) and PyBOP (3 eq) were added. The reaction mixture was irradiated with microwaves for 20 min (20 W, max. 60° C) under argon atmosphere. The solid support was then washed three times for 2 min with DMF.

Cleavage

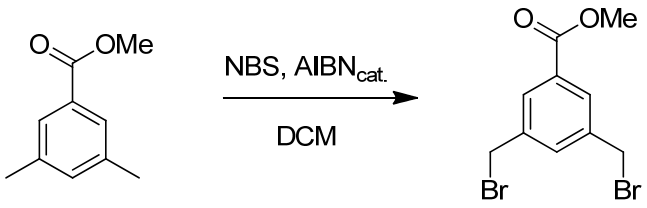
Cleavage of the raw product from the solid support was carried out without microwave irradiation. Therefore the beads were transferred into Schlenck glass vessels equipped with a frit and a stopper and treated according to the cleavage procedure described in chapter 7.2.

7.6 PREPARATION OF NON-SYMMETRIC TWEEZER RECEPTORS

Synthesis of the Orthogonally Protected Template 181

The preparation of **181** has already been described.[195] Therefore only the optimized reaction steps will be reported in the following.

Methyl-3,5-bis(bromomethyl)benzoate (174)

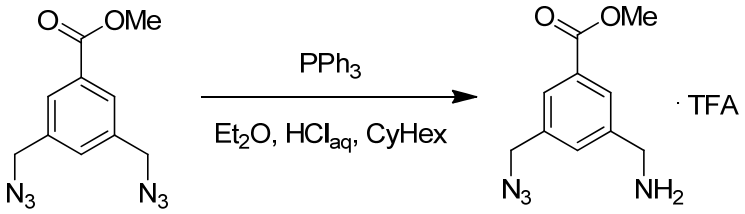


173 **174**

To a refluxing solution (55° C oil bath temperature) of **173** (5.00 g, 30.45 mmol, 1 eq) in DCM (125 mL) NBS (11.92 g, 66.99 mmol, 2.2 eq) and a catalytic amount of AIBN were added and the reaction mixture was stirred for 48 h. After cooling down to room temperature the solution was washed three times with water (75 mL) and dried with sodium sulfate. The solvent was removed via distillation under reduced pressure and the raw product was purified by flash chromatography (SiO_2, CyHex/EA = 19/1) to obtain **174** as white solid (7.49 g, 23.26 mmol, 76 %).

$C_{10}H_{10}Br_2O_2$	321.99 g/mol
Yield	76 %
Mp.	99° C
R_f	0.67 (SiO_2, hexane/EA = 1/9)
^{1}H-NMR	**(400 MHz, $CDCl_3$):** δ = 3.93 (s, 3H, CH_3), 4.50 (s, 4H, 2 × CH_2), 7.61 (dd, $^4J_1$ = $^4J_2$ = 1.7 Hz, 1H, CH_{ar}), 7.99 (d, ^4J = 1.7 Hz, 2H, 2 × CH_{ar}) ppm.
^{13}C-NMR	**(100 MHz, $CDCl_3$):** δ = 31.8 (*C*H_2), 52.4 (*C*H_3), 130.0 (*C*H_{ar}), 131.4 (C_q), 133.8 (*C*H_{ar}), 138.9 (C_q), 165.9 (*C*O_2Me) ppm.
MS	**(pos. EI):** m/z = 319.9 $[M - e^-]^{+\cdot}$ calculated for $C_{10}H_{10}Br_2O_2^{+\cdot}$: 319.9
FT-IR	**(KBr disk):** $\tilde{\nu}$ = 2117 (w), 1723 (s), 1435 (m), 1315 (m), 1212 (s), 994 (m), 898 (w), 771 (m), 697 (s) cm^{-1}.

Methyl-3-aminomethyl-5-azidomethylbenzoate, TFA-salt (176)

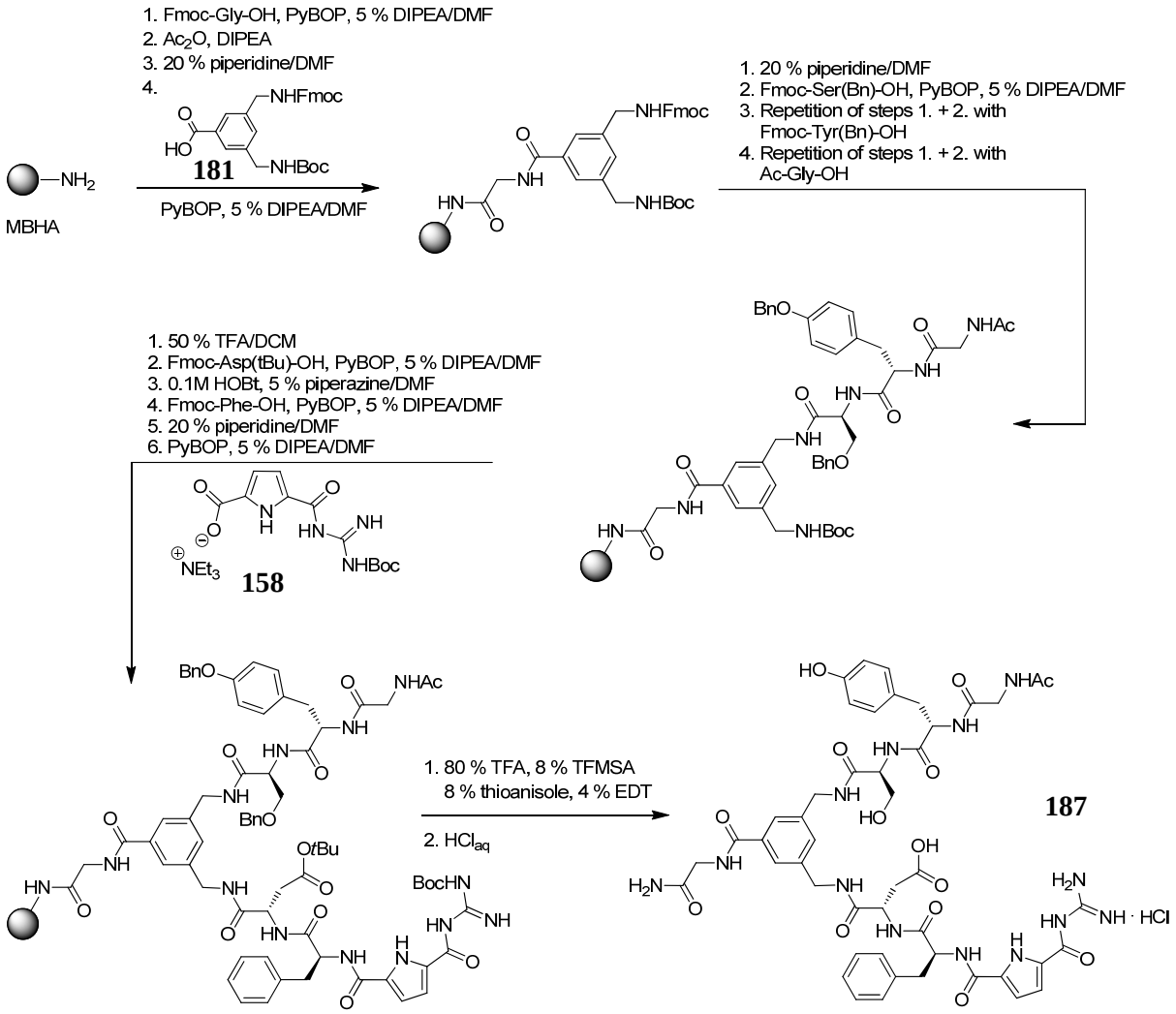


175 **176**

To a solution of **175** (4.00 g, 16.24 mmol, 1 eq) in diethyl ether (120 mL) cyclohexane (40 mL) and 5 % aqueous hydrochloric acid (160 mL) were added. The mixture was cooled to 0° C and triphenylphosphane (4.26 g, 16.24 mmol, 1 eq) was added. The reaction mixture was stirred for 8 h at 0° C and for another 40 h at room temperature. After phase separation the organic layer was extracted with 5 % aqueous hydrochloric acid (80 mL) once. The aqueous phases were combined and lyophilized under reduced pressure. The raw product was purified via MPLC (RP18, 20 % water/methanol → 40 % water/methanol, 0.1 % TFA) to give **176** as white solid (4.43 g, 13.26 mmol, 86 %).

$C_{12}H_{13}F_3N_4O_4$	334.25 g/mol
Yield	86 %
Mp.	138° C
R_f	0.66 (RP18, water/methanol = 1/1, 0.1 % TFA)
^{1}H-NMR	**(400 MHz, DMSO-d_6):** δ = 3.89 (s, 3 H, OCH_3), 4.15 (s, 2 H, CH_2), 4.61 (s, 2 H, CH_2), 7.73 (dd, $^4J_1$ = $^4J_2$ = 1.6 Hz, 1 H, CH_{ar}), 7.97 (d, ^4J = 1.6 Hz, 1 H, CH_{ar}), 8.26 (dd, ^4J = 1.6 Hz, 1 H, CH_{ar}), 8.39 (br. s, 3 H, NH^{3+}) ppm.
^{13}C-NMR	**(100 MHz, DMSO-d_6):** δ = 41.6 (*C*H_2), 52.2 (*C*H_3), 52.6 (*C*H_2), 128.7 (*C*H_{ar}), 129.3 (*C*H_{ar}), 130.2 (C_q), 133.5 (*C*H_{ar}), 135.2 (C_q), 137.0 (C_q), 165.5 (*C*O_2) ppm.
HR-MS	**(pos. ESI-ToF):** m/z = 221.1033 $[M + H^+ - TFA]^+$ calculated for $C_{10}H_{13}N_4O_2^+$: 221.1033
FT-IR	**(KBr disk):** $\tilde{\nu}$ = 2969 (br. w), 2098 (w), 1710 (m), 1671 (s), 1423 (w), 1313 (w), 1237 (m), 1176 (s), 889 (w), 836 (w), 779 (m), 722 (m) cm^{-1}.

Microwave Assisted SPPS of the Non-Symmetric Tweezer 187

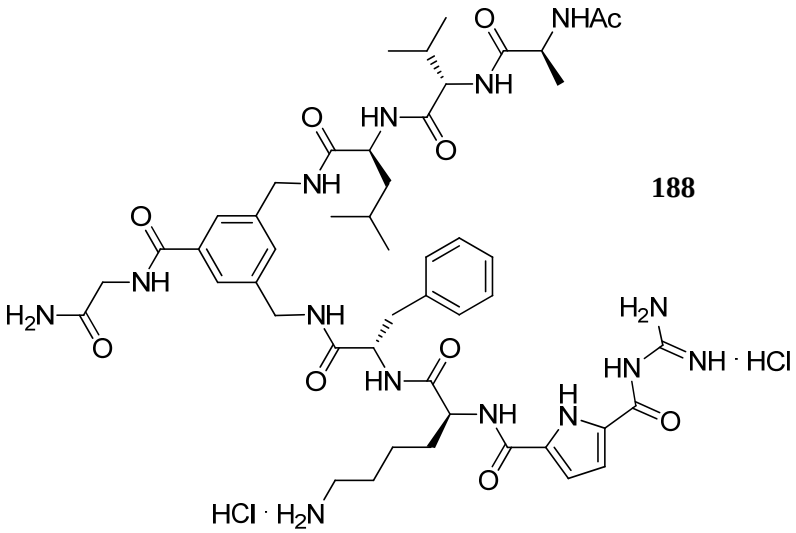


The synthesis of the non-symmetric tweezer was carried out on 100 mg MBHA resin (1.52 mmol/g) according to the standard microwave-assisted SPPS procedure described in chapter 7.5. During the coupling steps 5 % DIPEA was used as base and PyBOP (229 mg, 0.46 mmol, 3 eq) as coupling agent. After attaching Fmoc-Gly-OH (226 mg, 0.76 mmol, 5 eq), the resin was treated with Ac_2O (71 µL, 0.76 mmol, 5 eq). After the coupling step the Fmoc groop was removed with 20 % piperidine/DMF. Then the template **181** (285 mg, 0.46 mmol, 3 eq), Fmoc-Ser(Bn)-OH (190 mg, 0.46 mmol, 3 eq), Fmoc-Tyr(Bn)-OH (225 mg, 0.46 mmol, 3 eq), and Ac-Gly-OH (53 mg, 0.46 mmol, 3 eq) were coupled and the Fmoc-protecting groups were respectively removed after each coupling step. The following Boc deprotection was conducted without microwave irradiation by treating the solid support with 50 % TFA/DCM two times for 15 min. The resin was then washed three times for 5 min with DCM and another three times for 5 min with 5 % DIPEA/DMF. The following coupling steps of Fmoc-Asp(O*t*Bu)-OH (188 mg, 0.46 mmol, 3 eq), Fmoc-Phe-OH (177 mg, 0.46 mmol, 3 eq), and Boc-GCP-OH (**158**, 181 mg, 0.46 mmol, 3 eq) as well as the corresponding Fmoc deprotection steps were again achieved via the standard microwave-assisted procedure with the exception of the Fmoc deprotection following the attachment of asparagine which was achieved by treating the polymeric support with 0.1 M HOBt in 5 %

piperazine/DMF instead of 20 % piperidine/DMF. The cleavage of the product from the solid support was achieved without microwave irradiation according to the standard procedure for MBHA resin described in chapter 7.2. Then the crude product was hydrogenated with H_2 and 10 % Pd/C in acetone for 3 h. The suspension was filtrated over celite and the solvent was removed via distillation under reduced pressure. The raw product was purified via MPLC (RP18, methanol/water, 0.1 % TFA) and the trifluoroacetate counterions were exchanged by dissolving the product two times in 0.1 M aqueous HCl and one time in water followed by freeze drying under reduced pressure to obtain **187** as voluminous white solid (21 mg, 0.02 mmol, 13 %).

$C_{47}H_{56}ClN_{13}O_{14}$	1062.48 g/mol
Yield	13 %
Mp.	162° C (decomposition)
^{1}H-NMR	**(700 MHz, MeOD-d_4, TFA-salt):** δ = 1.96 (s, 3H, CH_3), 2.66 (s, 1H, CH), 2.79 (br. s, 1H, CH_2), 2.86 (dd, $^3J_1$ = 8.83 Hz, $^3J_2$ = 14.19 Hz, 1H, CH_2), 2.90 (s, 1H, CH), 2.99-3.08 (m, 3H, 2 × CH_2), 3.14 (dd, $^3J_1$ = 5.25 Hz, $^3J_2$ = 17.29 Hz, 1H, CH_2), 3.77 (dd, $^3J_1$ = $^3J_2$ = 16.62 Hz, 2H, CH_2), 3.98 (d, ^3J = 16.69 Hz, 1H, CH_2), 4.05-4.60 (m, 5H, 2 × CH, CH_2, CH_2-OH), 6.71 (d, ^3J = 8.57 Hz, 2H, CH_{ar}), 6.76 (s, 1H, CH_{ar}), 6.89-6.96 (m, 2H, 2 × NH), 6.99 (d, ^3J = 8.28, 2H, CH_{ar}), 7.01-7.09 (m, 1H, NH), 7.09-7.31 (m, 7H, 7 × CH_{ar}), 7.32-7.39 (m, 1H, NH), 7.44-7.52 (m, 1H, NH), 7.54 (s, 1H, NH), 7.62-7.76 (m, 1H, NH), 8.03 (s, 1H, NH), 8.88 (br. s, 1H, NH) ppm.
^{13}C-NMR	**(175 MHz, MeOD-d_4, TFA-salt):** δ = 9.22 (CH_3), 18.70, 19.74, 21.85, 23.45, 25.96, 26.92, 28.12, 30.45 (CH_2), 31.69 (CH_2), 33.06 (CH_2), 36.54 (CH_2), 40.47 (CH_2), 41.41 (CH_2), 43.61 (CH_2), 47.96 (CH), 50.86 (CH), 53.67 (CH), 57.50 (CH), 60.62 (CH), 114.15 (CH_{ar}), 125.52 (CH_{ar}), 125.95 (CH_{ar}), 127.79 (CH_{ar}), 129.68 (CH_{ar}), 130.27 (CH_{ar}), 130.60 (C_q), 130.99 (C_q), 135.00 (C_q), 138.73 (C_q), 140.51 (C_q), 140. 72 (C_q), 169.61 (CO), 173.50 (CO), 173.59 (CO), 173.88 (CO), 174.73 (CO), 174.87 (CO), 175.75 (CO) ppm.
HR-MS	**(pos. ESI-ToF):** m/z = 1026.4103 $[M - Cl^-]^+$ calculated for $C_{47}H_{56}N_{13}O_{14}^-$: 1026.4064
FT-IR	**(ATR):** $\tilde{\nu}$ = 3308 (br. m), 2372 (w), 1669 (s), 1542 (m), 1436 (m), 1280 (m), 1200 (s), 1134 (s), 839 (m), 800 (m), 762 (w), 722 (m) cm^{-1}.

Microwave Assisted SPPS of the Non-Symmetric Tweezer 188



188

The second non-symmetric tweezer receptor **188** was synthesized accordingly on 80 mg MBHA resin (1.30 mmol/g) following the above described microwave-assisted SPPS procedure. During coupling steps 3 % DIPEA was used as base and PyBOP as coupling agent (162 mg, 0.31 mmol, 3 eq). After attaching Fmoc-Gly-OH (93 mg, 0.31 mmol, 3 eq) and the template (160 mg, 0.31 mmol, 3 eq), the amino acids Fmoc-Leu-OH (110 mg, 0.31 mmol, 3 eq), Fmoc-Val-OH (106 mg, 0.31 mmol, 3 eq) and Ac-Ala-OH (41 mg, 0.31 mmol, 3 eq) were coupled to the first arm and Fmoc-Phe-OH (121 mg, 0.31 mmol, 3 eq), Fmoc-Lys(Boc)-OH (146 mg, 0.31 mmol, 3 eq) and Boc-GCP-OH (124 mg, 0.31 mmol, 3 eq) in the second arm. The product was obtained as voluminous white solid (42 mg, 0.04 mmol, 37 %).

$C_{49}H_{72}Cl_2N_{14}O_{10}$	1088.09 g/mol
Yield	37 %
Mp.	221° C (decomposition)
^{1}H-NMR	**(700 MHz, MeOD-d_4, TFA-salt):** δ = 0.89-0.94 (m, 6H, 2 × Val-CH_3), 0.96 (d, ^3J = 6.28 Hz, 3H, Ala-CH_3), 1.26-1.36 (m, 13H, 2 × Val-CH_3, 2 × Lys-CH_2, Ac-CH_3), 1.54-1.61 (m, 2H, Lys-CH_2), 1.67-1.73 (m, 2H, Lys-CH_2), 2.00-2.06 (m, 2H, CH_2), 2.07-2.13 (m, 1H, CH), 2.19 (t, ^3J = 7.53 Hz, 1H, CH), 2.77 (t, ^3J = 7.69 Hz, Lys-CH), 3.21 (dd, $^3J_1$= 7.33 Hz, $^3J_2$ = 14.64 Hz, 1H, CH), 3.42 (tdd, $^3J_1$ = 1.53 Hz, $^3J_2$ = 3.12 Hz, ^2J = 28.69 Hz, 2H, Leu-CH_2), 4.10 (d, ^3J = 4.28 Hz, 1H, CH), 4.13 (d, ^3J = 6.90 Hz, 1H, CH), 4.19 (d, ^3J = 16.21, 1H, Gly-CH_2), 4.25-4.35 (m, 2H, CH_2), 4.36-4.44 (m, 2H, CH_2), 4.58 (s, 2H, NH$_2$), 4.63 (d, ^3J = 15.93 Hz, 1H, Gly-CH_2), 5.35 (t, ^3J = 4.61 Hz, Phe-CH), 6.81 (d, ^3J = 4.08 Hz, 1H, Pyr-CH_{ar}), 6.87 (d, ^3J = 3.90 Hz, 1H, Pyr-CH_{ar}), 7.15-7.20 (m, 1H, Phe-CH_{ar}), 7.23 (d, ^3J = 4.34 Hz, 4H, 4 × Phe-CH_{ar}), 7.26-7.30 (m, 1H, NH), 7.32 (s, 1H, CH_{ar}), 7.35-7.40 (m, 1H, NH), 7.41-7.48 (m, 1H, NH), 7.56

(s, 1H, CH_{ar}), 7.64 (s, 1H, CH_{ar}), 7.66-7.71 (m, 1H, NH), 7.80 (br. s, 1H, *N*H), 7.90 (s, 1H, *N*H) ppm.

^{13}C-NMR **(175 MHz, MeOD-d$_4$, TFA-salt):** δ = 22.47 (*C*H$_3$), 35.80 (*C*H$_2$), 36.86 (*C*H$_2$), 37.83 (*C*H$_2$), 40.39 (*C*H), 43.26 (*C*H$_2$), 43.52 (*C*H$_2$), 44.05 (*C*H$_2$), 45.47 (*C*H), 53.50 (*C*H), 55.66 (*C*H), 63.74 (*C*H-OH), 114.06 (*C*H$_{ar}$), 116.36 (*C*H$_{ar}$), 117.38 (C_q), 119.05 (C_q), 120.73 (C_q), 124.53 (*C*H$_{ar}$), 125.14 (*C*H$_{ar}$), 128.37 (*C*H$_{ar}$), 129.69 (*C*H$_{ar}$), 130.22 (*C*H$_{ar}$), 130.44 (*C*H$_{ar}$), 131.28 (*C*H$_{ar}$), 134.55 (C_q), 137.86 (C_q), 139.71 (C_q), 140.64 (C_q), 157.54 (*C*O), 161.34 (*C*O), 162.97 (*C*O), 163.15 (*C*O), 166.95 (*C*O), 169.02 (*C*O), 171.94 (*C*O), 172.16 (*C*O), 172.82 (*C*O), 174.14 (*C*O), 175.07 (*C*O) ppm.

HR-MS **(pos. ESI-ToF):** m/z = 1015.5537 $[M - HCl - Cl^-]^+$
calculated for **$C_{49}H_{71}N_{14}O_{10}^+$**: 1015.5472

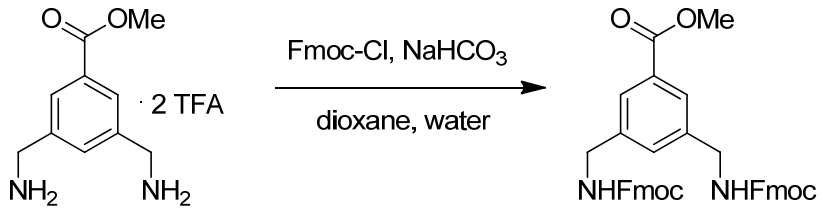
FT-IR **(ATR):** $\tilde{\nu}$ = 3268 (br. m), 2364 (w), 1638 (s), 1542 (s), 1469 (m), 1278 (m), 1202 (s), 1136 (m), 839 (w), 722 (w), 704 (w) cm^{-1}.

7.7 SYNTHESIS OF SYMMETRIC TWEEZERS

Synthesis of the Template 193

The preparation of **193** has already been described elsewhere.[195] Therefore only the optimized reaction steps will be reported in the following.

Methyl-3,5-bis(N-Fmoc-aminomethyl)benzoate (192)

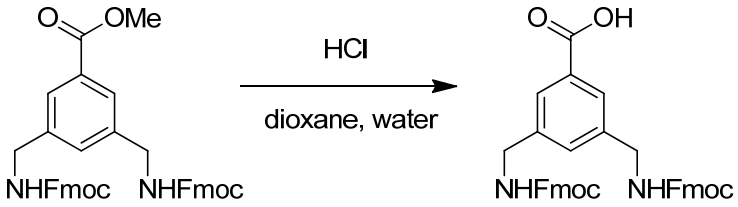


190 **192**

A solution of **190** (4.37 g 10.35 mmol, 1 eq) and sodium hydrogencarbonate (3.91 g, 46.54 mmol, 4.5 eq) in water (80 mL) was stirred for 10 min. Then dioxane (160 mL) was added, the mixture cooled to 0° C and a solution of Fmoc-Cl (5.90 g, 22.8 mmol, 2.2 eq) in dioxane (80 mL) was added dropwise during 15 min. The reaction mixture was stirred for 20 h while warming up to room temperature. The dioxane was removed via distillation under reduced pressure and 500 mL of water were added. The raw product was extracted with EA (4 × 750 mL). The organic phases were combined and the solvent was removed via distillation under reduced pressure. The resulting solid was purified by means of flash chromatography (SiO_2, first EA/CyHex = 1/2 until elution of the first substance, then MeOH/EA = 1/9) to obtain **192** as white solid (4.70 g, 7.46 mmol, 71 %).

$C_{40}H_{34}N_2O_6$	638.71 g/mol
Yield	71 %
Mp.	205° C
R_f	0.53 (SiO_2, CyHex/EA = 1/1)
^{1}H-NMR	**(300 MHz, DMSO-d_6):** δ = 3.83 (s, 3H, CH_3), 4.10-4.36 (m, 10H, 2 × CH_2, 2 × Fmoc-CH_2, 2 × Fmoc-CH), 7.27-7.46 (m, 9H, CH_{ar}, 8 × Fmoc-CH_{ar}), 7.63-7.63-7.72 (m, 4H, 4 × Fmoc-CH_{ar}), 7.78 (s, 2H, 2 × CH_{ar}), 7.87 (t, ^3J = 7.86 Hz, 4H, 4 × Fmoc-CH_{ar}), 7.98 (t, ^3J = 6.06 Hz, 2H, 2 × NH) ppm.
^{13}C-NMR	**(75 MHz, DMSO-d_6):** δ = 43.48 (*C*H_2), 46.69 (Fmoc-*C*H), 52.13 (*C*H_3), 65.54 (Fmoc-*C*H_2), 120.11 (Fmoc-*C*H_{ar}), 125.12 (*C*H_{ar}), 126.32 (*C*H_{ar}), 127.03 (*C*H_{ar}), 127.60 (*C*H_{ar}), 129.79 (C_q), 140.71 (C_q), 143.82 (C_q), 145.19 (C_q), 156.32 (Fmoc-*C*O), 166.16 (*C*O) ppm.
HR-MS	**(pos. ESI-ToF):** m/z = 661.2340 $[M + Na^+]^+$ calculated for $C_{40}H_{35}N_2O_6^+$: 661.2309
FT-IR	**(ATR):** $\tilde{\nu}$ = 3303 (br. m), 3064 (w), 3015 (w), 2951 (w), 2891 (w), 1719 (m), 1691 (s), 1542 (s), 1446 (m), 1254 (s), 1217 (s), 1146 (s), 1046 (m), 995 (m), 737 (s), 667 (m) cm^{-1}.

3,5-bis(N-Fmoc-aminomethyl)benzoic acid (193)

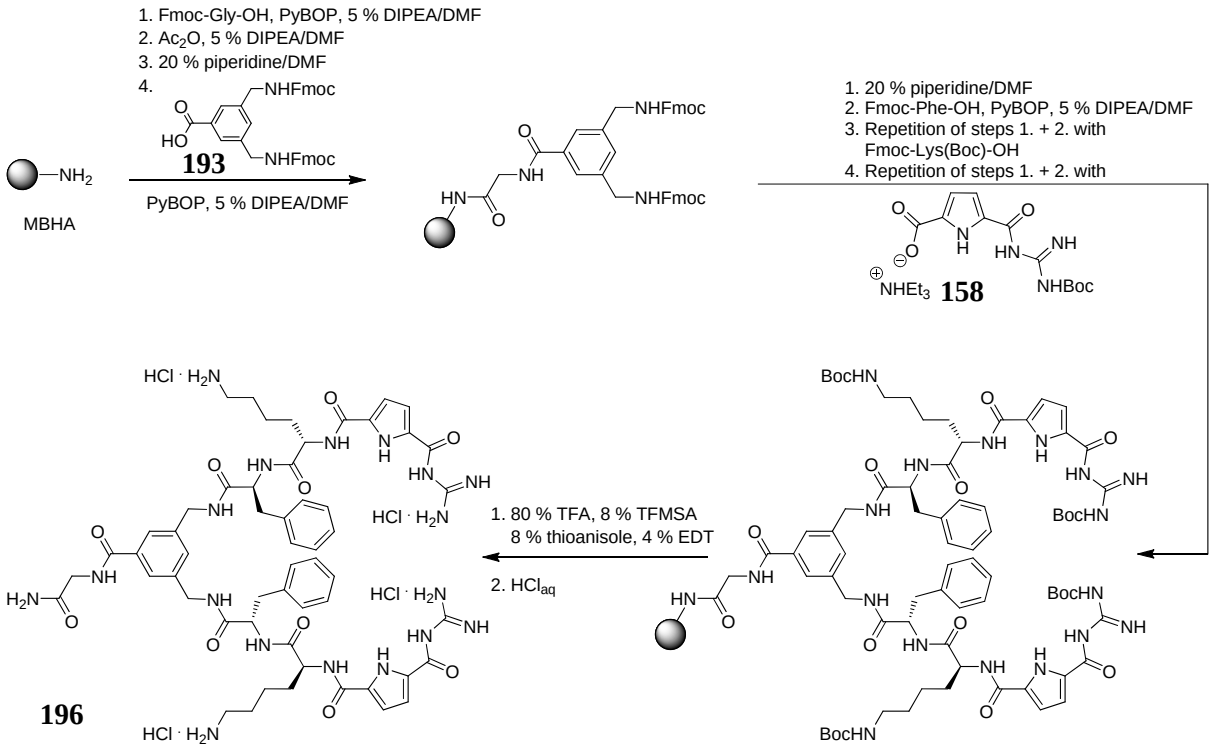


192 **193**

A suspension of **192** (2.35 g, 3.68 mmol) in dioxane/water (2/1, 150 mL) was cooled to 0° C and concentrated, aqueous HCl (10 mL) was added slowly. The reaction mixture was heated (100° C oil bath temperature) and stirred for 26 h. The resulting solution was cooled to room temperature and 400 mL of water were added. The precipitate was filtrated, washed with water and dried in an dissicator over phosphorous pentoxide to give **193** (2.08 g, 3.33 mmol, 90 %) as white solid.

$C_{39}H_{32}N_2O_6$	624.68 g/mol
Yield	90 %
Mp.	212° C (decomposition)
R_f	0.30 (SiO_2, EA)
^{1}H-NMR	**(400 MHz, $CDCl_3$):** δ = 4.18-4.33 (m, 10H, 2 × CH_2, 2 × Fmoc- CH_2, 2 × Fmoc-CH), 7.28 (m, 5H, 4 × Fmoc-CH_{ar}, CH_{ar}), 7.40 (dd, $^3J_1 = {}^3J_2$ = 7.4 Hz, 4H, 4 × Fmoc-CH_{ar}), 7.69 (d, ^3J = 7.3 Hz, 4H, 4 × Fmoc-CH_{ar}), 7.77 (s, 2H, 2 × CH_{ar}), 7.88 (d, ^3J = 7.5 Hz, 4H, 4 × Fmoc-CH_{ar}), 7.95 (dd, $^3J_1 = {}^3J_2$ = 6.1 Hz, 2H, 2 × NH) ppm.
^{13}C-NMR	**(100 MHz, $CDCl_3$):** δ = 43.5 (CH_2), 46.6 (Fmoc-CH), 65.4 (Fmoc-CH_2), 120.0 (Fmoc-CH_{ar}), 125.1 (CH_{ar}), 126.4 (CH_{ar}), 126.9 (CH_{ar}), 127.4 (CH_{ar}), 139.9 (C_q), 140.6 (C_q), 143.7 (C_q), 156.2 (Fmoc-CO) ppm.
HR-MS	**(neg. ESI-ToF):** m/z = 623.2160 $[M - H^+]^-$ calculated for $C_{39}H_{31}N_2O_6^-$: 623.2188
FT-IR	**(KBr disk):** $\tilde{\nu}$ = 3343 (m), 2921 (m), 2847 (w), 2093 (m), 1713 (s), 1691 (s), 1522 (s), 1452 (m), 1364 (w), 1254 (s), 1222 (s), 1156 (s), 1044 (w), 862 (w), 773 (w), 738 (m) cm^{-1}.

Microwave Assisted SPPS of the Tweezer 196 (Phe-Lys-GCP)



The synthesis of **196** was carried out on 68 mg MBHA resin (1.30 mmol/g) according to the standard microwave-assisted procedure described in chapter 7.5. During coupling steps 5 % DIPEA was used as base. After attaching Fmoc-Gly-OH (93 mg, 0.27 mmol, 3 eq) with PyBOP (157 mg, 0.27 mmol, 3 eq), the resin was treated with Ac_2O (48 µL, 0.44 mmol, 5 eq). After the coupling step the Fmoc groop was removed with 20 % piperidine/DMF. Then the template **193** (195 mg, 0.27 mmol, 3 eq) was coupled with 3 eq of PyBOP. Afterwards Fmoc-Phe-OH (242 mg, 0.54 mmol, 6 eq), Fmoc-Lys(Boc)-OH (292 mg, 0.54 mmol, 6 eq), and Boc-GCP-OH (248 mg, 0.54 mmol, 6 eq) were coupled with 6 eq of PyBOP (314 mg, 0.54 mmol). The cleavage of the product from the solid support was achieved without microwave irradiation according to the standard procedure for MBHA resin described in chapter 7.2. The raw product was purified via MPLC (RP18, methanol/water, 0.1 % TFA) and the trifluoroacetate counterions were exchanged by dissolving the product two times in 0.1 M aqueous HCl and one time in water followed by freeze drying under reduced pressure to obtain **187** as voluminous white solid (42 mg, 0.03 mmol, 37 %).

$C_{55}H_{74}Cl_4N_{18}O_{10}$	1289.10 g/mol
Yield	37 %
Mp.	230° C (decomposition)
^{1}H-NMR	**(500 MHz, MeOD-d_4, TFA-salt):** δ = 1.19-1.35 (m, 4H, 2 × Lys-CH_2), 1.52-1.62 (m, 4H, 2 × Lys-CH_2), 1.64-1.75 (m, 4H, 2 × Lys-CH_2), 2.80 (t, ^3J = 7.40 Hz, 4H, 2 × Lys-CH_2), 4.07 (d, ^3J = 15.49 Hz, 2H, CH_2), 4.25-4.39 (m, 4H, 2 × CH_2), 4.55 (dd, $^3J_1$ = 9.49 Hz, $^3J_2$ = 5.20 Hz, 2H, 2

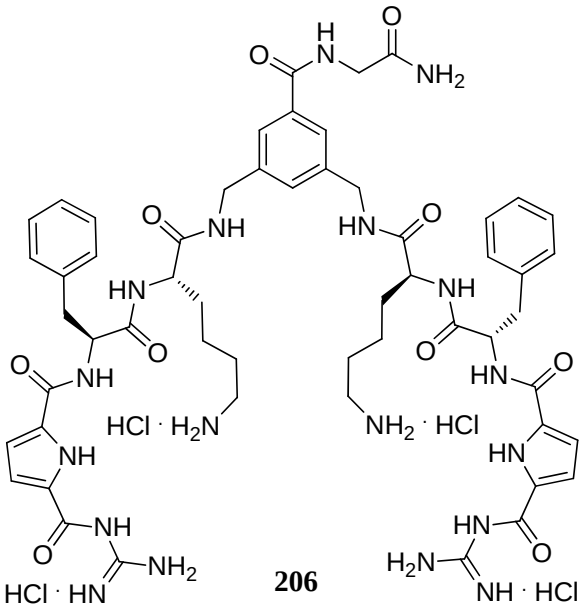
× C*H*), 4.62 (d, ^{3}J = 15.58 Hz, 2H, 2 × C*H*), 6.89 (d, ^{3}J = 3.82 Hz, 2H, 2 × Pyr-C*H*), 7.06 (d, ^{3}J = 3.87 Hz, 2H, 2 × Pyr-C*H*), 7.12 (s, 1H, CH_{ar}), 7.15-7.19 (m, 2H, 2 × Phe-CH_{ar}), 7.21-7.27 (m, 8H, 8 × Phe-CH_{ar}), 7.45 (s, 2H, 2 × CH_{ar}) ppm.

^{13}C-NMR **(125 MHz, MeOD-d$_4$, TFA-salt):** δ = 23.08 (Lys-*C*H$_2$), 28.16 (Lys-*C*H$_2$), 31.48 (Lys-*C*H$_2$), 37.27 (Phe-*C*H$_2$), 40.47 (Lys-*C*H$_2$), 43.35 (Gly-*C*H$_2$), 43.49 (*C*H$_2$), 55.97 (*C*H), 56.87 (*C*H), 114.16 (*C*H$_{ar}$), 116.25 (*C*H$_{ar}$), 125.26 (*C*H$_{ar}$), 127.10 (C_q), 127.77 (*C*H$_{ar}$), 129.55 (*C*H$_{ar}$), 130.31 (*C*H$_{ar}$), 132.86 (C_q), 134.32 (C_q), 139.01 (C_q), 140.49 (), 157.18 (C_q), 161.50 (C_q), 162.61 (C_q), 168.99 (C_q), 173.50 (C_q), 175.04 (C_q), 175.18 (C_q) ppm.

HR-MS **(pos. ESI-ToF):** m/z = 1143.5690 [M – 3 HCl – Cl$^-$]+
calculated for $C_{55}H_{71}N_{18}O_{10}^+$: 1143.5595

FT-IR **(ATR):** $\tilde{\nu}$ = 3307 (br m), 2362 (w), 1665 (s), 1541 (s), 1280 (m), 1198 (s), 1132 (s), 839 (m), 800 (m), 721 (w), 655 (m) cm^{-1}.

Microwave Assisted SPPS of the Tweezer 206 (Lys-Phe-GCP)



206

The tweezer **206** was synthesized accordingly on 200 mg MBHA resin (1.52 mmol/g) following the above described microwave-assisted SPPS procedure. During coupling steps 5 % DIPEA was used as base and PyBOP as coupling agent. After attaching Fmoc-Gly-OH (272 mg, 0.92 mmol, 3 eq) capping with Ac_2O (145 µL, 1.52 mmol, 5 eq), Fmoc deprotection and coupling of the template (379 mg, 0.92 mmol, 3 eq), the amino acids Fmoc-Lys(Boc)-OH (855 mg, 1.79 mmol, 6 eq), Fmoc-Phe-OH (707 mg, 1.79 mmol, 6 eq) and the binding motif Boc-GCP-OH (725 mg, 1.79 mmol, 6 eq) were coupled. The product was obtained as voluminous white solid (82 mg, 0.06 mmol, 21 %).

$C_{55}H_{74}Cl_4N_{18}O_{10}$	1289.10 g/mol
Yield	21 %
Mp.	> 250° C (decomposition)
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 1.28-1.39 (m, 4H, 2 × Lys-CH_2), 1.50-1.57 (m, 4H, 2 × Lys-CH_2), 1.59-1.67 (m, 2H, Lys-CH_2), 1.69-1.78 (m, 2H, Lys-CH_2), 2.70-2.78 (m, 4H, 2 × Lys-CH_2), 2.90 (dd, $^3J_1$ = 10.77 Hz, $^3J_2$ = 13.77 Hz, 2H, Phe-CH_2), 2.90 (dd, $^3J_1$ = 3.82 Hz, $^3J_2$ = 13.75 Hz, 2H, Phe-CH_2), 3.83 (ddd, $^3J_1$ = 5.87 Hz, $^3J_2$ = $^3J_3$ = 13.75 Hz, 2H, Gly-CH_2), 4.24-4.37 (m, 6H, 2 × CH, 2 × CH_2), 4.55 (ddd, $^3J_1$ = 4.34 Hz, $^3J_2$ = 8.31 Hz, $^3J_3$ = 10.83 Hz, 2H, 2 × CH), 6.85 (dd, ^3J = 2.42 Hz, ^4J = 4.01 Hz, 2H, 2 × Pyr-CH), 7.14 (dd, $^3J_1$ = 6.27 Hz, $^3J_2$ = 13.56 Hz, 2H, 2 × Phe-CH_{ar}), 7.23 (t, ^3J = 7.58 Hz, 4H, 4 × Phe-CH_{ar}), 7.29 (s, 1H, CH_{ar}), 7.33 (d, ^3J = 7.26 Hz, 4H, 4 × Phe-CH_{ar}), 7.46 (dd, ^3J = 2.35 Hz, ^4J = 4.03 Hz, 2H, 2 × Pyr-CH), 7.49 (s, 2H, NH_2), 7.66 (s, 2H, 2 × CH_{ar}), 7.79 (br. s, 6H, 2 × Lys-NH_3^+), 8.36-8.66 (m, 13H, 2 × GCP-NH_4^+, 5 × NH),

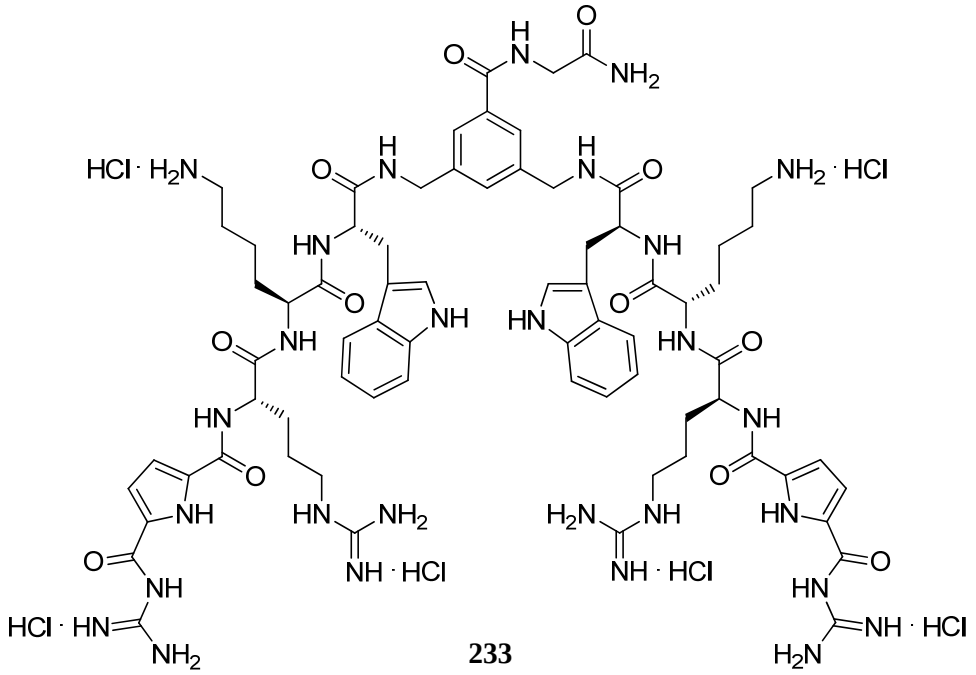
8.75 (d, ^{3}J = 8.05 Hz, 2H, 2 × N*H*), 11.95 (s, 2H, 2 × GCP-N*H*), 12.41 (s, 2H, 2 × Pyr-N*H*) ppm.

^{13}C-NMR **(125 MHz, DMSO-d$_6$):** δ = 22.21 (Lys-*C*H$_2$), 26.46 (Lys-*C*H$_2$), 31.19 (Lys-*C*H$_2$), 37.10 (Phe-*C*H$_2$), 41.81 (Lys-*C*H$_2$), 42.30 (Gly-*C*H$_2$), 52.69 (*C*H), 54.22 (*C*H), 113.21 (Pyr-*C*H$_{ar}$), 115.68 (Pyr-*C*H$_{ar}$), 124.39 (*C*H$_{ar}$), 125.54 (*C*$_q$), 126.20 (*C*H$_{ar}$), 127.99 (*C*H$_{ar}$), 129.02 (*C*H$_{ar}$), 132.02 (*C*$_q$), 134.09 (*C*$_q$), 137.84 (*C*$_q$), 139.35 (*C*$_q$), 155.28 (*C*O), 158.79 (*C*O), 159.57 (*C*O), 166.17 (*C*O), 171.18 (*C*O), 171.31 (*C*O), 171.35 (*C*O) ppm.

HR-MS **(pos. ESI-ToF):** m/z = 1143.5649 $[M + H^+]^+$
calculated for $C_{55}H_{71}N_{18}O_{10}^+$: 1143.5595

FT-IR **(ATR):** $\tilde{\nu}$ = 3091 (br. m), 2948 (w), 2908 (w), 2840 (w), 1683 (s), 1650 (s), 1558 (s), 1540 (s), 1508 (s), 1216 (s), 1116 (m), 1007 (w), 1268 (s), 998 (s), 871 (s), 763 (m) cm^{-1}.

Microwave Assisted SPPS of the Tweezer 233 (Trp-Lys-Arg-GCP)



233

The divalent DNA-ligand **233** was synthesized accordingly on 200 mg Fmoc-Rink-Amide resin (0.94 mmol/g) following the above described microwave-assisted SPPS procedure. During coupling steps 5 % DIPEA was used as base and PyBOP as coupling agent. After attaching Fmoc-Gly-OH (335 mg, 1.13 mmol, 6 eq) capping with Ac_2O (176 µL, 1.188 mmol, 10 eq), Fmoc deprotection and coupling of the template (352 mg, 0.56 mmol, 3 eq), the amino acids Fmoc-Trp(Boc)-OH (594 mg, 1.13 mmol, 6 eq), Fmoc-Lys(Boc)-OH (529 mg, 1.13 mmol, 6 eq), Fmoc-Arg(Pbf)-OH (732 mg, 1.13 mmol, 6 eq), and the binding motif Boc-GCP-OH (448 mg, 1.13 mmol, 6 eq) were coupled. The cleavage of the raw product was carried out without microwave irradiation according to the procedure described in chapter 7.2. After purification via MPLC (RP8, MeOH/water, 0.1 % TFA) and exchange of the counter ion from trifluoroacetate to chloride the product was obtained as voluminous white solid (20 mg, 0.06 mmol, 6 %).

$C_{71}H_{102}Cl_6N_{28}O_{12}$	1752.47 g/mol
Yield	6 %
Mp.	> 250° C (decomposition)
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 1.23-1.29 (m, 4H, 2 × Lys-CH_2), 1.42-1.57 (m, 12H, 2 × Lys-CH_2, 4 × Arg-CH_2), 1.59-1.69 (m, 4H, 2 × Lys-CH_2), 2.65-2.72 (m, 4H, 2 × Lys-CH_2), 2.99-3.13 (m, 8H, 2 × Trp-CH_2, 2 × Arg-CH_2), 3.80-3.93 (m, 2H, Gly-CH_2), 4.18-4.33 (m, 6H, 2 × Arg-CH, 2 × CH_2,), 4.44-4.51 (m, 2H, 2 × CH), 4.55 (dd, $^3J_1$ = 7.05 Hz, $^3J_2$ = 14.16 Hz, 2H, 2 × CH), 6.90 (s, 2H, 2 × Pyr-CH_{ar}), 6.94-7.00 (m, 2H, 2 × Trp-CH_{ar}), 7.06 (br. s, 8H, 2 × Arg-NH_4^+), 7.08-7.16 (m, 6H, NH_2, 2 × Pyr-CH_{ar}, 2 × Trp-CH_{ar}), 7.32 (d, ^3J = 8.18 Hz, 2H, 2 × Trp-CH_{ar}),

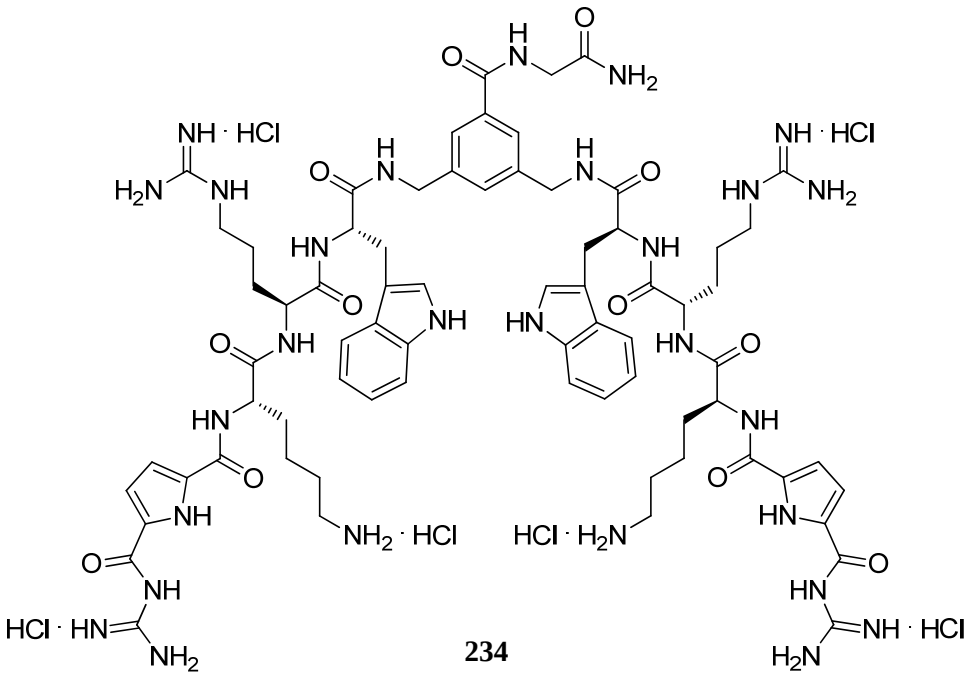
7.51 (s, 1H, C*H*$_{ar}$), 7.54-7.61 (m, 4H, 2 × N*H,* 2 × Trp-C*H*$_{ar}$), 7.62-7.74 (m, 8H, 2 × Lys-N*H*$_3^+$, 2 × C*H*$_{ar}$), 8.05 (d, ^3J = 7.56 Hz, 2H, 2 × NH), 8.17 (d, ^3J = 5.84 Hz, 2H, 2 × NH), 8.31-8.66 (m, 13H, 5 × N*H*, 2 × GCP-N*H*$_4^+$), 10.79 (s, 2H, Trp-N*H*), 11.38 (s, 2H, GCP-N*H*), 12.57 (s, 2H, Pyr-N*H*) ppm.

^{13}C-NMR **(175 MHz, DMSO-d$_6$):** δ = 22.09 (*C*H$_2$), 25.15 (*C*H$_2$), 26.56 (*C*H$_2$), 27.93 (*C*H$_2$), 29.10 (*C*H$_2$), 31.22 (*C*H$_2$), 38.65 (*C*H$_2$), 40.43 (*C*H$_2$), 42.03 (*C*H$_2$), 42.41 (*C*H$_2$), 48.52 (*C*H), 52.33 (*C*H), 53.64 (*C*H), 109.73 (*C*H$_{ar}$), 111.33 (*C*H$_{ar}$), 113.77 (*C*H$_{ar}$), 114.95 (*C*$_q$), 116.50 (*C*$_q$), 118.24 (*C*H$_{ar}$), 118.43 (*C*H$_{ar}$), 120.87 (*C*H$_{ar}$), 123.69 (*C*H$_{ar}$), 124.54 (*C*H$_{ar}$), 125.70 (*C*$_q$), 127.31 (*C*H$_{ar}$), 129.01 (*C*$_q$), 129.57 (*C*$_q$), 134.08 (*C*$_q$), 136.01 (*C*$_q$), 139.21 (*C*$_q$), 155.10 (*C*O), 156.64 (*C*O), 158.02 (*C*O), 158.20 (*C*O), 158.36 (*C*O), 166.38 (*C*O), 171.36 (*C*O), 171.41 (*C*O), 171.44 (*C*O) ppm.

HR-MS **(pos. ESI-ToF):** m/z = 767.4103 [M – 4 HCl – 2 Cl$^-$]$^{2+}$
calculated for $C_{71}H_{98}N_{28}O_{12}^{2+}$: 767.3954

FT-IR **(ATR):** $\tilde{\nu}$ = 3292 (br. w), 1699 (br. s), 1541 (br. S), 1418 (br. m), 1232 (br. s), 1166 (br. s) cm^{-1}.

Microwave Assisted SPPS of the Tweezer 234 (Trp-Arg-Lys-GCP)



234

The divalent DNA-ligand **234** was synthesized accordingly on 200 mg Fmoc-Rink-Amide resin (0.94 mmol/g) following the above described microwave-assisted SPPS procedure. During coupling steps 5 % DIPEA was used as base and PyBOP as coupling agent. After attaching Fmoc-Gly-OH (335 mg, 1.13 mmol, 6 eq), capping with Ac_2O (176 µL, 1.188 mmol, 10 eq), Fmoc deprotection and coupling of the template (352 mg, 0.56 mmol, 3 eq), the amino acids Fmoc-Trp(Boc)-OH (594 mg, 1.13 mmol, 6 eq), Fmoc-Arg(Pbf)-OH (732 mg, 1.13 mmol, 6 eq), Fmoc-Lys(Boc)-OH (529 mg, 1.13 mmol, 6 eq), and the binding motif Boc-GCP-OH (448 mg, 1.13 mmol, 6 eq) were coupled. The cleavage of the raw product was carried out without microwave irradiation according to the procedure described in chapter 7.2. After purification via MPLC (RP8, MeOH/water, 0.1 % TFA) and exchange of the counter ion from trifluoroacetate to chloride the product was obtained as voluminous white solid (63 mg, 0.06 mmol, 19 %).

$C_{71}H_{102}Cl_6N_{28}O_{12}$	1752.47 g/mol
Yield	19 %
Mp.	> 250° C (decomposition)
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 1.30-1.72 (m, 20H, 6 × Lys-CH_2, 4 × Arg-CH_2), 2.71-2.81 (m, 4H, 2 × Lys-CH_2), 2.98-3.09 (m, 6H, Trp-CH_2, 2 × Arg-CH_2), 3.14 (dd, $^3J_1$ = 5.46 Hz, $^3J_2$ = 14.41 Hz, 2H, Trp-CH_2), 3.81-3.91 (m, 2H, Gly-CH_2), 4.16-4.34 (m, 6H, 2 × CH_2, 2 × CH), 4.46 (dd, $^3J_1$ = 8.65 Hz, $^3J_2$ = 13.16 Hz, 2H, 2 × CH), 4.61 (dd, $^3J_1$ = 7.08 Hz, $^3J_2$ = 13.96 Hz, 2H, 2 × CH), 6.90 (dd, ^4J = 2.33 Hz, ^3J = 3.85 Hz, 2H, 2 × Pyr-CH_{ar}), 6.96 (t, ^3J = 7.37 Hz, 2H, 2 × Trp-CH_{ar}), 7.05 (t, ^3J = 7.48 Hz, 2H, Trp-CH_{ar}), 7.09-7.15 (m, 4H, 2 × Pyr-CH_{ar}, 2 × Trp-CH_{ar}),

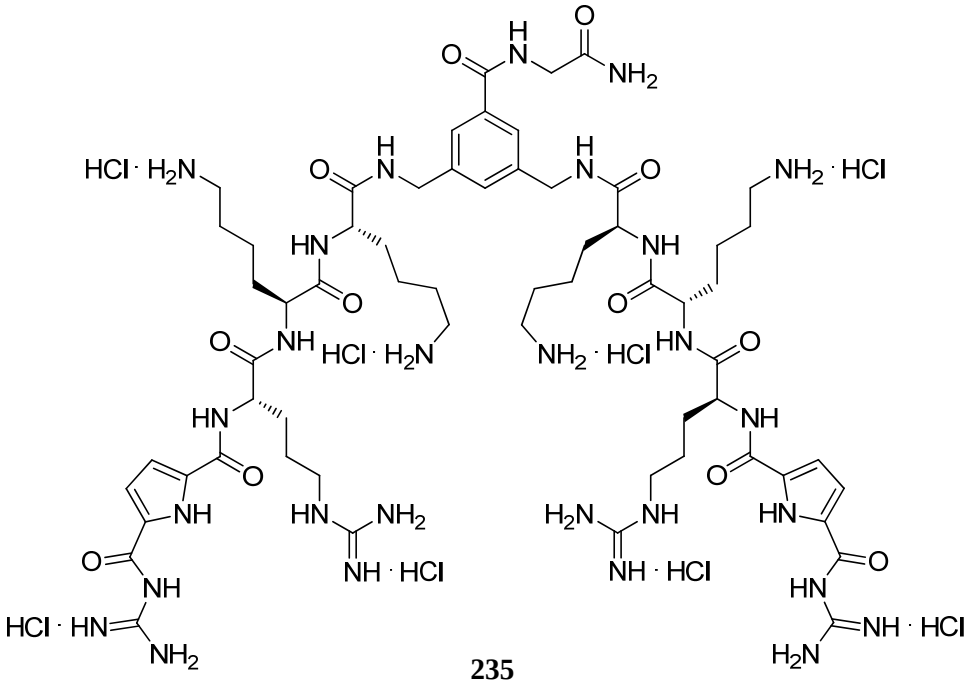
7.17 (s, 2H, N*H*$_2$), 7.24 (br. s, 8H, 2 × Arg-N*H*$_4^+$), 7.32 (d, ^3J = 8.11 Hz, 2H, 2 × Trp-C*H*$_{ar}$), 7.52 (s, 1H, C*H*$_{ar}$), 7.56-7.56 (m, 4H, 2 × N*H,* 2 × Trp-C*H*$_{ar}$), 7.64 (s, 2H, 2 × C*H*$_{ar}$), 7.74 (s, 6H, 2 × Lys-N*H*$_3^+$), 8.00 (d, ^3J = 7.44 Hz, 2H, 2 × NH), 8.25 (d, ^3J = 7.48 Hz, 2H, 2 × NH), 8.43-8.79 (m, 13H, 5 × N*H*, 2 × GCP-N*H*$_4^+$), 10.79 (s, 2H, Trp-N*H*), 11.51 (s, 2H, GCP-N*H*), 12.52 (s, 2H, Pyr-N*H*) ppm.

^{13}C-NMR **(125 MHz, DMSO-d$_6$):** δ = 22.50 (*C*H$_2$), 24.91 (*C*H$_2$), 26.56 (*C*H$_2$), 27.93 (*C*H$_2$), 28.89 (*C*H$_2$), 31.20 (*C*H$_2$), 38.65 (*C*H$_2$), 40.32 (*C*H$_2$), 41.96 (*C*H$_2$), 42.31 (*C*H$_2$), 52.19 (*C*H), 52.53 (*C*H), 53.55 (*C*H), 109.54 (*C*H$_{ar}$), 111.22 (*C*$_q$), 113.57 (*C*H$_{ar}$), 114.90 (*C*H$_{ar}$), 118.16 (*C*H$_{ar}$), 118.32 (*C*H$_{ar}$), 120.83 (*C*H$_{ar}$), 123.55 (*C*H$_{ar}$), 124.56 (*C*H$_{ar}$), 125.58 (*C*$_q$), 127.21 (*C*H$_{ar}$), 128.96 (*C*$_q$), 132.10 (*C*$_q$), 134.03 (*C*$_q$), 135.95 (*C*$_q$), 139.17 (*C*$_q$), 155.13 (*C*O), 156.63 (*C*O), 158.22 (*C*O), 158.47 (*C*O), 158.73 (*C*O), 158.93 (*C*O), 166.41 (*C*O), 171.15 (*C*O), 171.21 (*C*O), 171.63 (*C*O) ppm.

HR-MS **(pos. ESI-ToF):** m/z = 767.4194 [M – 4 HCl – 2 Cl$^-$]$^{2+}$
calculated for $C_{71}H_{98}N_{28}O_{12}^{2+}$: 767.3954

FT-IR **(ATR):** $\tilde{\nu}$ = 3280 (br. w), 1698 (m), 1636 (br. s), 1540 (br. s), 1456 (br. m), 1252 (br. s), 1264 (br. s), 746 (s), 615 (br. s) cm^{-1}.

Microwave Assisted SPPS of the Tweezer 235 (Lys-Lys-Arg-GCP)



235

The divalent DNA-ligand **235** was synthesized accordingly on 125 mg Fmoc-Rink-Amide resin (0.94 mmol/g) following the above described microwave-assisted SPPS procedure. During coupling steps 5 % DIPEA was used as base and PyBOP as coupling agent. After attaching Fmoc-Gly-OH (210 mg, 0.71 mmol, 6 eq), capping with Ac_2O (110 µL, 1.18 mmol, 10 eq), Fmoc deprotection and coupling of the template (440 mg, 0.71 mmol, 6 eq), the amino acids Fmoc-Lys(Boc)-OH (330 mg, 0.71 mmol, 6 eq), Fmoc-Lys(Boc)-OH (330 mg, 0.71 mmol, 6 eq), Fmoc-Arg(Pbf)-OH (457 mg, 0.71 mmol, 6 eq), and the binding motif Boc-GCP-OH (280 mg, 0.71 mmol, 6 eq) were coupled. The cleavage of the raw product was carried out without microwave irradiation according to the procedure described in chapter 7.2. After purification via MPLC (RP8, MeOH/water, 0.1 % TFA) and exchange of the counter ion from trifluoroacetate to chloride the product was obtained as voluminous white solid (18 mg, 0.01 mmol, 9 %).

$C_{61}H_{108}Cl_8N_{28}O_{12}$	1709.31 g/mol
Yield	9 %
Mp.	> 250° C (decomposition)
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 1.22-1.39 (m, 8H, 4 × Lys-CH_2), 1.42-1.84 (m, 24H, 4 × Arg-CH_2, 8 × Lys-CH_2), 2.69-2.78 (m, 8H, 4 × Lys-CH_2), 3.12 (td, 4H, $^3J_1$ = 6.87 Hz, $^3J_2$ = 13.80 Hz, 2 × Arg-CH_2), 3.78-3.92 (ddd, $^3J_1$ = $^3J_2$ = 16.65 Hz, $^3J_3$ = 6.56 Hz, Gly-CH_2), 4.09-4.31 (m, 6H, 2 × CH, 2 × CH_2), 4.36 (dd, $^3J_1$ = 15.41 Hz, $^3J_2$ = 5.01 Hz, 2H, 2 × CH), 4.47 (dd, $^3J_1$ = 14.11 Hz, $^3J_2$ = 6.50 Hz, 2H, 2 × CH), 6.91 (d, ^3J = 3.59 Hz, 2H, 2 × Pyr-CH_{ar}), 7.12 (d, ^3J = 4.37 Hz, 2H, 2 × Pyr-CH_{ar}),7.14 (br. s, 8H, 2 × Arg-NH_4^+), 7.17 (s, 2H, NH_2), 7.28 (s, 2H, 2

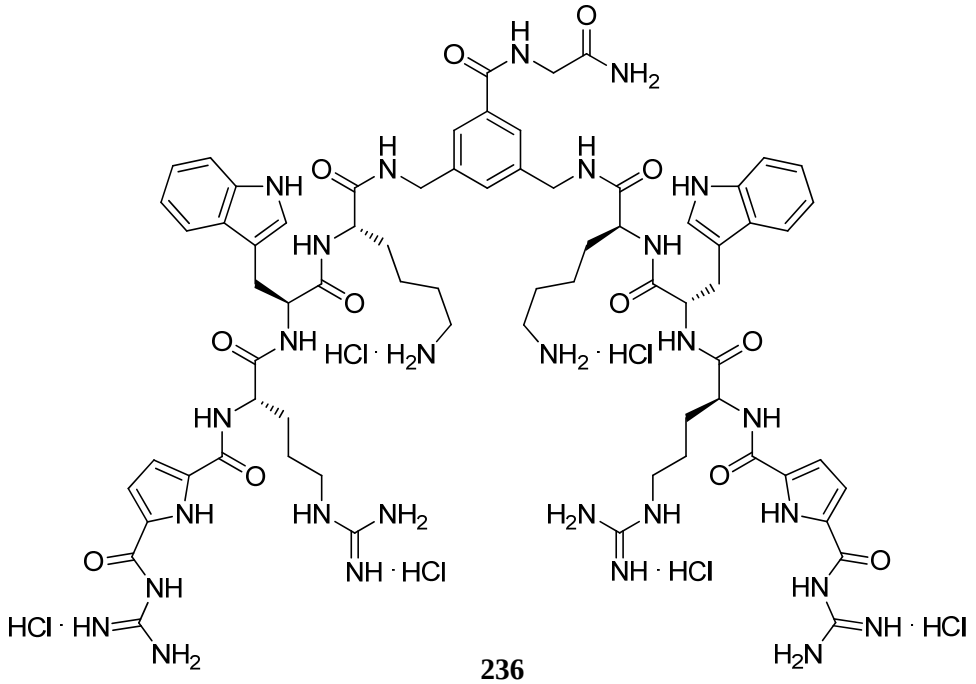
× N*H*), 7.51 (s, 1H, C*H*$_{ar}$), 7.64 (s, 2H, 2 × C*H*$_{ar}$), 7.75 (br. s, 12H, 4 × Lys-N*H*$_3^+$), 8.04 (d, ^3J = 6.42 Hz, 2H, 2 × N*H*), 8.17 (d, ^3J = 7.10 Hz, 2H, 2 × N*H*), 8.20-8.79 (m, 13H, 5 × N*H*, 2 × GCP-N*H*$_4^+$), 11.51 (s, 2H, 2 × GCP-N*H*), 12.52 (s, 2H, 2 × Pyr-N*H*) ppm.

^{13}C-NMR **(125 MHz, DMSO-d$_6$):** δ = 22.24 (*C*H$_2$), 25.18 (*C*H$_2$), 26.44 (*C*H$_2$), 26.52 (*C*H$_2$), 28.83 (*C*H$_2$), 30.87 (*C*H$_2$), 31.27 (*C*H$_2$), 38.54 (*C*H$_2$), 38.57 (*C*H$_2$), 40.34 (*C*H$_2$), 41.88 (*C*H$_2$), 42.41 (*C*H$_2$), 44.00 (*C*H), 52.51 (*C*H), 52.68 (*C*H), 113.67 (*C*H$_{ar}$), 115.65 (*C*H$_{ar}$), 124.37 (*C*H$_{ar}$), 125.79 (*C*$_q$), 128.89 (*C*H$_{ar}$), 134.15 (*C*$_q$), 139.42 (*C*$_q$), 156.85 (*C*$_q$), 159.11 (*C*$_q$), 166.35 (*C*$_q$), 171.31 (*C*$_q$), 171.47 (*C*$_q$), 171.57 (*C*$_q$),171.60 (*C*$_q$) ppm.

HR-MS **(pos. ESI-ToF):** m/z = 709.4035 $[M - 4\ HCl - 2\ Cl^-]^{2+}$
calculated for $C_{61}H_{102}N_{28}O_{122}^{2+}$: 709.4111

FT-IR **(ATR):** $\tilde{\nu}$ = 3273 (br. w), 1636 (br. s), 1541 (br. s), 1457 (w), 1280 (br. m), 1195 (br. m) cm^{-1}.

Microwave Assisted SPPS of the Tweezer 236 (Lys-Trp-Arg-GCP)



236

The divalent DNA-ligand **236** was synthesized accordingly on 200 mg Fmoc-Rink-Amide resin (0.94 mmol/g) following the above described microwave-assisted SPPS procedure. During coupling steps 5 % DIPEA was used as base and PyBOP as coupling agent. After attaching Fmoc-Gly-OH (335 mg, 1.13 mmol, 6 eq), capping with Ac_2O (176 µL, 1.88 mmol, 10 eq), Fmoc deprotection and coupling of the template (352 mg, 0.56 mmol, 3 eq), the amino acids Fmoc-Lys(Boc)-OH (529 mg, 1.13 mmol, 6 eq), Fmoc-Trp(Boc)-OH (594 mg, 1.13 mmol, 6 eq), Fmoc-Arg(Pbf)-OH (732 mg, 1.13 mmol, 6 eq), and the binding motif Boc-GCP-OH (224 mg, 0.56 mmol, 3 eq) were coupled. The cleavage of the raw product was carried out without microwave irradiation according to the procedure described in chapter 7.2. After purification via MPLC (RP8, MeOH/water, 0.1 % TFA) and exchange of the counter ion from trifluoroacetate to chloride the product was obtained as voluminous white solid (49 mg, 0.03 mmol, 15 %).

$C_{71}H_{102}Cl_6N_{28}O_{12}$	1752.47 g/mol
Yield	15 %
Mp.	> 250° C (decomposition)
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 1.22-1.32 (m, 4H, 2 × Lys-CH_2), 1.39-1.61 (m, 12H, 2 × Lys-CH_2, 4 × Arg-CH_2), 1.64-1.75 (m, 4H, 2 × Lys-CH_2), 2.68-2.79 (m, 4H, 2 × Lys-CH_2), 2.99 (dd, $^3J_1$ = 8.19 Hz, $^3J_2$ = 14.63 Hz, 2H, Trp-CH_2), 3.07 (td, $^3J_1$ = 6.61 Hz, $^3J_2$ = 12.77 Hz, 4H, 2 × Arg-CH_2), 3.16 (dd, $^3J_1$ = 4.04 Hz, $^3J_2$ = 14.44 Hz, 2H, Trp-CH_2), 3.86 (ddd, $^3J_1$ = 5.79 Hz, $^3J_2$ = 16.30 Hz, $^3J_3$ = 16.66 Hz, 2H, 2 × Gly-CH_2), 4.20-4.30 (m, 4H, 2 × CH_2), 4.34 (dd, $^3J_1$ = 5.55 Hz, $^3J_2$ = 15.57 Hz, 2H, 2 × CH), 4.46 (dd, $^3J_1$ = 7.60 Hz, $^3J_2$ = 13.72 Hz, 2H, 2 ×

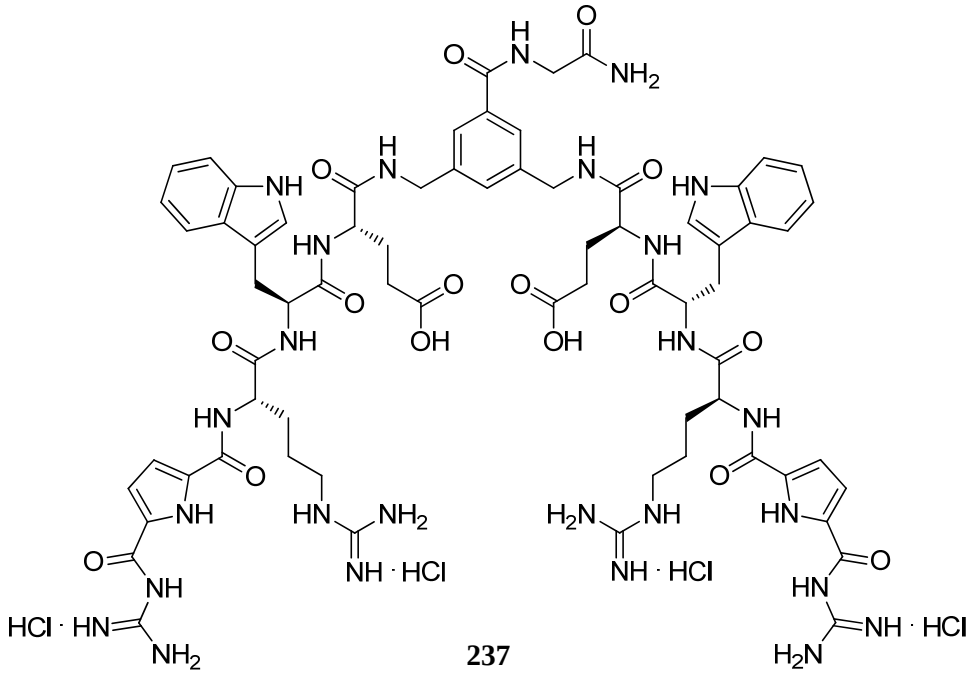
C*H*), 4.61 (dd, $^{3}J_{1}$ = 7.82 Hz, $^{3}J_{2}$ = 13.17 Hz, 2H, 2 × C*H*), 6.88-6.95 (m, 4H, 2 × Trp-CH_{ar}, 2 × Pyr-CH_{ar}), 7.01 (t, ^{3}J = 7.63 Hz, 2H, 2 × Trp-CH_{ar}), 7.07 (br. s, 8H, 2 × Arg-NH_4^+), 7.08-7.11 (m, 6H, 2 × Pyr-CH_{ar}, 2 × Trp-CH_{ar}, NH_2), 7.29 (d, ^{3}J = 8.16 Hz, 2H, 2 × Trp-CH_{ar}), 7.49 (s, 1H, CH_{ar}), 7.54 (d, ^{3}J = 7.86 Hz, 2H, 2 × Trp-CH_{ar}), 7.61 (t, ^{3}J = 5.42, 2H, 2 × N*H*), 7.65 (s, 2H, 2 × CH_{ar}), 7.73 (s, 6H, 2 × Lys-NH_3^+), 8.13 (t, ^{3}J = 8.18 Hz, 4H, 4 × N*H*), 8.36 (t, ^{3}J = 5.43 Hz, 2H, 2 × N*H*), 8.47 (br. s, 8H, 2 × GCP-NH_4^+), 8.54-8.62 m, 3H, 3 × N*H*), 10.78 (s, 2H, Trp-N*H*), 11.44 (s, 2H, GCP-N*H*), 12.52 (s, 2H, Pyr-N*H*) ppm.

^{13}C-NMR **(125 MHz, DMSO-d$_6$):** δ = 22.15 (*C*H_2), 25.01 (*C*H_2), 26.61 (*C*H_2), 27.25 (*C*H_2), 28.91 (*C*H_2), 31.29 (*C*H_2), 38.60 (*C*H_2), 40.28 (*C*H_2), 41.81 (*C*H_2), 42.32 (*C*H_2), 52.48 (*C*H), 52.68 (*C*H), 53.26 (*C*H), 109.63 (*C*H_{ar}), 111.13 (C_q), 113.60 (*C*H_{ar}), 114.89 (*C*H_{ar}), 118.10 (*C*H_{ar}), 118.30 (*C*H_{ar}), 120.74 (*C*H_{ar}), 123.53 (*C*H_{ar}), 124.29 (*C*H_{ar}), 125.58 (C_q), 127.20 (*C*H_{ar}), 132.01 (C_q), 134.07 (C_q), 135.88 (C_q), 139.35 (C_q), 155.04 (*C*O), 156.61 (*C*O), 158.28 (*C*O), 158.53 (*C*O), 158.95 (*C*O), 166.28 (*C*O), 171.17 (*C*O), 171.32 (*C*O), 171.36 (*C*O) ppm.

HR-MS **(pos. ESI-ToF):** m/z = 767.4194 $[M - 4\ HCl - 2\ Cl^-]^{2+}$
calculated for $C_{71}H_{98}N_{28}O_{12}^{2+}$: 767.3954

FT-IR **(ATR):** $\tilde{\nu}$ = 3176 (br. s), 11646 (br. s), 1541 (br. s), 1473 (br. m), 1254 (br. s), 1194 (m), 814 (w), 747 (s) cm^{-1}.

Microwave Assisted SPPS of the Tweezer 237 (Glu-Trp-Arg-GCP)



237

The divalent DNA-ligand **237** was synthesized accordingly on 200 mg Fmoc-Rink-Amide resin (0.94 mmol/g) following the above described microwave-assisted SPPS procedure. During coupling steps 5 % DIPEA was used as base and PyBOP as coupling agent. After attaching Fmoc-Gly-OH (168 mg, 0.56 mmol, 3 eq), capping with Ac_2O (176 µL, 1.88 mmol, 10 eq), Fmoc deprotection and coupling of the template (352 mg, 0.56 mmol, 3 eq), the amino acids Fmoc-Glu(O*t*Bu)-OH (480 mg, 1.13 mmol, 6 eq), Fmoc-Trp(Boc)-OH (594 mg, 1.13 mmol, 6 eq), Fmoc-Arg(Pbf)-OH (732 mg, 1.13 mmol, 6 eq), and the binding motif Boc-GCP-OH (224 mg, 0.56 mmol, 3 eq) were coupled. The cleavage of the raw product was carried out without microwave irradiation according to the procedure described in chapter 7.2. After purification via MPLC (RP8, MeOH/water, 0.1 % TFA) and exchange of the counter ion from trifluoroacetate to chloride the product was obtained as voluminous white solid (66 mg, 0.04 mmol, 21 %).

$C_{69}H_{90}Cl_4N_{26}O_{16}$	1681.43 g/mol
Yield	21 %
Mp.	> 250° C (decomposition)
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 1.41-1.52 (m, 4H, 2 × CH_2), 1.52-1.54 (m, 2H, CH_2), 1.67-1.76 (m, 2H, CH_2), 1.77-1.95 (m, 2H, CH_2), 1.92-1.99 (m, 2H, CH_2), 2.16-2.29 (m, 4H, 2 × Glu-CH_2), 2.98 (dd, $^3J_1$ = 8.11 Hz, $^3J_2$ = 14.24 Hz, 2H, Trp-CH_2), 3.07 (td, $^3J_1$ = 6.22 Hz, $^3J_2$ = 12.49 Hz, 4H, 2 × Arg-CH_2), 3.16 (dd, $^3J_1$ = 4.35 Hz, $^3J_2$ = 14.55 Hz, 2H, Trp-CH_2), 3.86 (ddd, $^3J_1$ = 5.79 Hz, $^3J_2$ = 16.30 Hz, $^3J_3$ = 16.66 Hz, 2H, Gly-CH_2), 4.22-4.35 (m, 6H, 2 × CH, 2 × CH_2), 4.45-4.53 (m, 2H, 2 × CH), 4.60 (dd, $^3J_1$ = 7.85 Hz, $^3J_2$ = 13.36 Hz, 2H, 2 × CH), 6.88-6.93

(m, 4H, 2 × Trp-C*H*$_{ar}$, 2 × Pyr-C*H*$_{ar}$), 7.01 (t, ^{3}J = 7.75 Hz, 2H, 2 × Trp-C*H*$_{ar}$), 7.07-7.16 (m, 6H, N*H*$_2$, 2 × Pyr-C*H*$_{ar}$, 2 × Trp-C*H*$_{ar}$), 7.17 (br. s, 8H, 2 × Arg-N*H*$_4^+$), 7.28 (d, ^{3}J = 8.14 Hz, 2H, 2 × Trp-C*H*$_{ar}$), 7.44 (s, 1H, C*H*$_{ar}$), 7.48 (t, ^{3}J = 5.46 Hz, 2H, 2 × N*H*), 7.55 (d, ^{3}J = 7.97 Hz, 2H, 2 × Trp-C*H*$_{ar}$), 7.66 (s, 2H, 2 × C*H*$_{ar}$), 8.16 (d, ^{3}J = 7.07 Hz, 2H, 2 × N*H*), 8.22 (d, ^{3}J = 7.43 Hz, 2H, 2 × N*H*), 8.27 (t, ^{3}J = 5.86 Hz, 2H, 2 × N*H*), 8.85 (br. s, 8H, 2 × GCP-N*H*$_4^+$), 8.54-8.61 (m, 3H, 3 × N*H*), 10.77 (s, 2H, Trp-N*H*), 11.29 (s, 2H, GCP-N*H*), 12.13 (br. s, 2H, CO_2*H*), 12.54 (s, 2H, Pyr-N*H*) ppm.

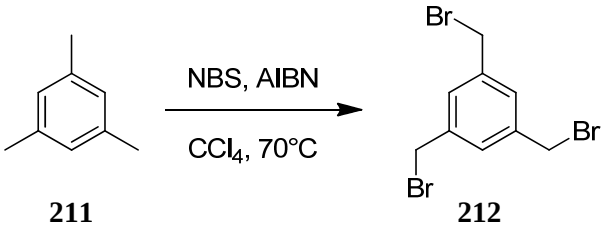
^{13}C-NMR **(125 MHz, DMSO-d$_6$):** δ = 24.96 (*C*H$_2$), 27.17 (*C*H$_2$), 26.56 (*C*H$_2$), 27.25 (*C*H$_2$), 29.10 (*C*H$_2$), 30.06 (*C*H$_2$), 40.31 (*C*H$_2$), 41.89 (*C*H$_2$), 42.38 (*C*H$_2$), 52.26 (*C*H), 53.30 (*C*H), 53.34 (*C*H), 109.69 (*C*H$_{ar}$), 111.39 (*C*H$_{ar}$), 111.83 (*C*$_q$), 113.62 (*C*H$_{ar}$), 114.90 (*C*H$_{ar}$), 115.03 (*C*$_q$), 118.11 (*C*H$_{ar}$), 118.32 (*C*H$_{ar}$), 120.74 (*C*H$_{ar}$), 123.50 (*C*H$_{ar}$), 124.48 (*C*H$_{ar}$), 127.16 (*C*$_q$), 135.88 (*C*$_q$), 139.29 (*C*$_q$), 156.53 (*C*O), 158.09 (*C*O), 158.32 (*C*O), 170.95 (*C*O), 171.10 (*C*O), 171.21 (*C*O), 171.47 (*C*O) ppm.

HR-MS **(pos. ESI-ToF):** m/z = 768.3458
calculated for $C_{69}H_{88}N_{26}O_{16}^{2+}$: 768.3431

FT-IR **(ATR):** $\tilde{\nu}$ = 3275 (br. w), 1636 (br. s), 1540 (br. s), 1436 (br. m), 1271 (br. m), 1197 (br. s), 1134 (br. s), 800 (m), 721 (s) cm^{-1}.

7.8 PREPARATION OF A TRIVALENT POLYNUCLEOTIDE LIGAND

1,3,5-Tris(bromomethyl)benzene (212)

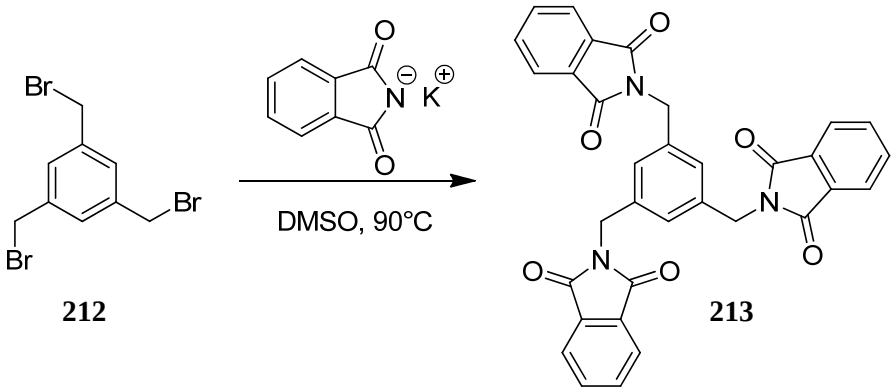


211 **212**

Mesitylene (**211**, 2 mL, 14.3 mmol, 1 eq), NBS (9.83 g, 55.2 mmol, 3.8 eq) and AIBN (catalytic) were dissolved in CCl_4 (60 mL) and refluxed (70° C oil bath temperature) for 40 h. After cooling down to room temperature, the resulting solid was filtrated and washed with CCl_4. The solvent was removed via distillation in vacuum and the resulting brown oil was purified by means of flash chromatography (SiO_2, CyHex/EA = 15/1) to obtain **212** as white solid (1.19 g, 3.35 mmol, 23 %.

$C_9H_9Br_3$	356.59 g/mol
Yield	23 %
Mp.	89-92° C
R_f	0.56 (SiO_2, CyHex/EA = 15/1)
^{1}H-NMR	**(300 MHz, $CDCl_3$):** δ = 4.46 (s, 6H, CH_2), 7.35 (s, 3H, C$H_{ar.}$) ppm.
^{13}C-NMR	**(75 MHz, $CDCl_3$):** δ = 32.2 (CH$_2$), 129.9 (CH$_{ar.}$), 139.2 ($C_{ar.q}$) ppm.
MS	**(pos. GC-EI)** m/z = 356 $[M - e^-]^{+\cdot}$
FT-IR	**(ATR):** $\tilde{\nu}$ = 3027 (w), 2998 (w), 2965 (w), 2923 (w), 1209 (s) 1145 (s), 979 (s), 701 (s) cm^{-1}.

1,3,5-Tris(N-phtalimidomethyl)benzene (213)

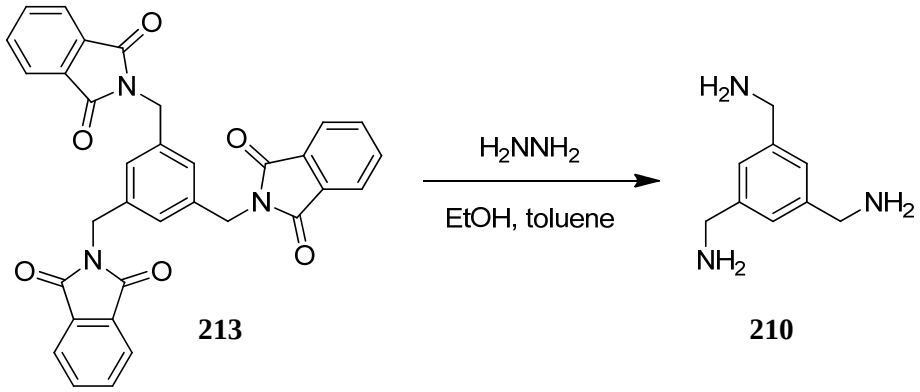


212 **213**

Potassium phthalimide (411 mg, 2.22 mmol, 3.0 eq) and **212** (261 mg, 0.73 mmol, 1 eq) were dissolved in DMSO (40 mL) and refluxed (90° C oil bath temperature) for 6 h. After cooling down to room temperature water was added (200 mL) and the suspension was extracted with DCM (5 × 100 mL). The organic phases were combined and dried with Na_2SO_4. The solvent was then removed via distillation under reduced pressure and the resulting solid purified by flash chromatography (SiO_2, DCM/acetone = 20/1) to obtain **213** as white solid (285 mg, 0.51 mmol, 84 %).

$C_{33}H_{21}N_3O_6$	555.47 g/mol
Yield	84 %
Mp.	267° C
R_f	0.55 (SiO_2, DCM/acetone = 20/1)
^{1}H-NMR	**(500 MHz, $CDCl_3$):** δ = 4.78 (s, 6H, CH_2), 7.35 (s, 3H, CH_{ar}), 7.68-7.83 (m, 12H, CH_{ar}) ppm.
^{13}C-NMR	**(125 MHz, $CDCl_3$):** δ = 41.6 (*C*H_2), 123.6 (*C*H_{ar}), 127.9 (*C*H_{ar}), 132.2 (C_q), 134.1 (*C*H_{ar}), 137.5 (C_q), 168.2 (*C*O) ppm.
HR-MS	**(pos. ESI-ToF):** m/z = 578.1343 $[M + Na^+]^+$ calculated for $C_{33}H_{21}N_3O_6Na^+$: 578.1323
FT-IR	**(ATR):** $\tilde{\nu}$ = 3027 (w), 2998 (w), 2965 (w), 2923 (w), 1209 (s) 1145 (s), 979 (s), 701 (s) cm^{-1}.

1,3,5-Triaminomethylbenzene (210)

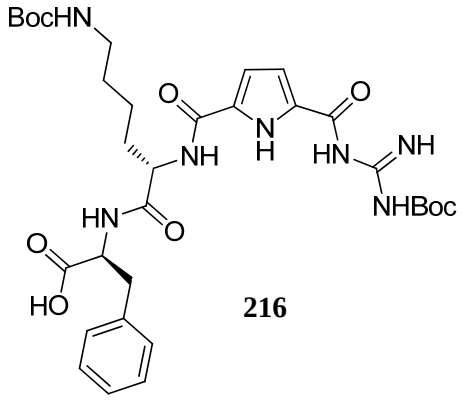


213 **210**

213 (39 mg, 0.07 mmol, 1 eq) was dissolved in EtOH/toluene (2/1, 30 mL) and heated to 90° C. Hydrazine monohydrate (105 µL, 2.12 mmol, 30 eq) were added and the mixture refluxed for 72 h. After cooling down to room temperature the resulting solid was filtrated and washed with EtOH/toluene (2/1). The solvent was then removed via distillation under reduced pressure and the raw product resuspended in EA (30 mL). After addition of cold 40 % KOH solution and phase separation the aqueous phase was was extracted with EA (4 × 30 mL). The organic phases were combined and dried with Na_2SO_4. Afterwards the solvent was removed via distillation under reduced pressure to give **210** (11 mg, 0.07 mmol, 93 %) as white solid.

$C_9H_{15}N_3$	165.24 g/mol
Yield	93 %
Mp.	109-110° C
^{1}H-NMR	**(300 MHz, MeOD-d_4):** δ = 4.78 (s, 6H, CH_2), 7.35 (s, 3H, CH_{ar}) ppm.
^{13}C-NMR	**(75 MHz, MeOD-d_4):** δ = 41.6 (CH_2), 127.9 (CH_{ar}), 132.2 (C_q) ppm.
HR-MS	**(pos. ESI-ToF):** m/z = 166.1353 $[M + H^+]^+$ calculated for $C_9H_{16}N_3^+$: 166.1339
FT-IR	**(ATR):** $\tilde{\nu}$ = 3196 (br. s), 3019 (w), 2955 (w), 1631 (m), 1539 (m), 997 (m), 882 (m), 764 (s), 628 (m), 619 (s) cm^{-1}.

Microwave Assisted SPPS of the Fully Protected Side Chain 216 (Phe-Lys-GCP)



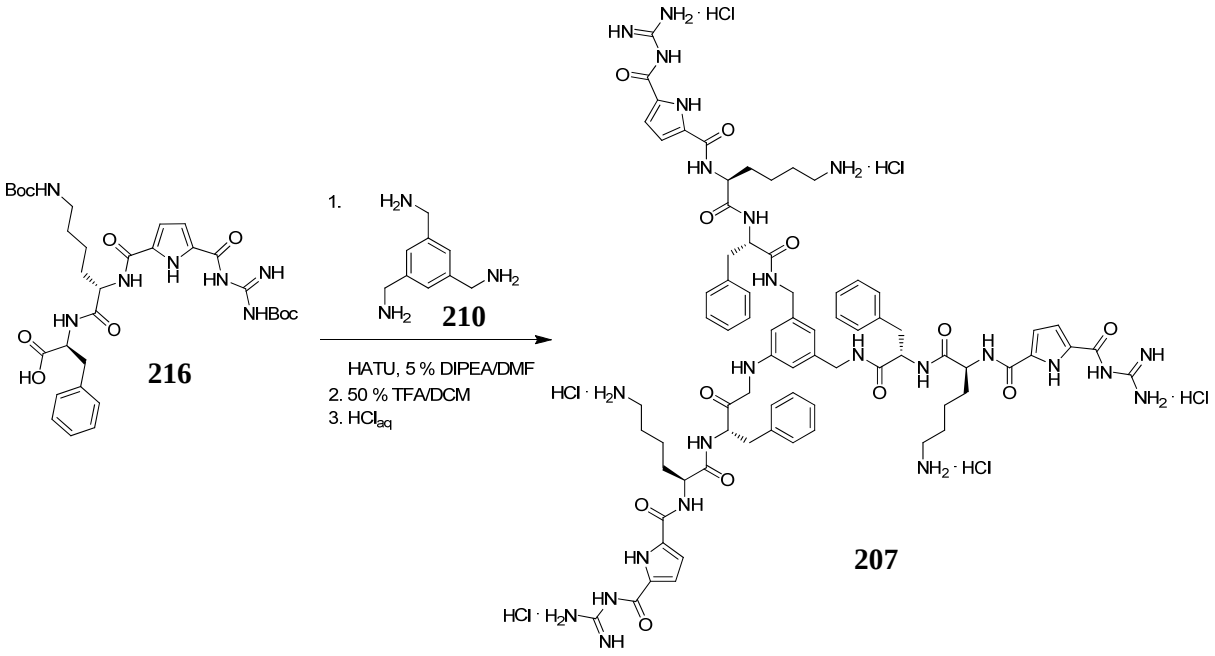
216

The synthesis of fully protected **216** was carried out on 300 mg SASRIN resin (1.3 mmol/g) according to standard Fmoc SPPS as described in chapter 7.2. The attachment of the first amino acid Fmoc-Phe-OH (471 mg, 1.17 mmol, 3 eq) was carried out with DIC (182 µl, 1.17 mmol, 3 eq), HOBt (179 mg, 1.17 mmol, 3 eq) and DMAP (catalytic) in DCM/DMF (7/3). Afterwards the resin was treated with Ac_2O (110 µl, 1.17 mmol, 3 eq) in 3 % NMM/DMF (10 ml) and the Fmoc group was removed with 20 % piperidine/DMF. The next amino acid Fmoc-Lys(Boc)-OH (547 mg, 1.17 mmol, 3 eq) and the binding motif Boc-GCP-OH (347 mg, 1.17 mmol, 3 eq) were coupled with PyBOP. The cleavage of the raw product was carried out without microwave irradiation according to the procedure described in chapter 7.2. The product was precipitated by adding water, purified via MPLC (RP18, MeOH/water) to obtain the product **216** as voluminous white solid (74 mg, 0.04 mmol, 28 %).

$C_{32}H_{45}N_7O_9$	671.74 g/mol
Yield	28 %
Mp.	177° C (decomposition)
^{1}H-NMR	**(500 MHz, $CDCl_3$):** δ = 1.36 (s, 9H, Boc-C*H*$_3$), 1.38-1.47 (m, 8H, Lys-C*H*$_2$), 1.59 (s, 9H, Boc-C*H*$_3$), 1.68-1.93 (m, 2H, Lys-C*H*$_2$), 2.95-3.08 (m, 2H, Lys-C*H*$_2$), 3.12-3.30 (m, 2H Phe-C*H*$_2$), 4.75 (s, 1H, N*H*); 4.89-5.01 (m, 1H, Phe-C*H*), 5.69-5.80 (m, 1H, Lys-C*H*), 6.56-6.65 (m, 2H, Phe-C*H*$_{ar}$, Pyr-C*H*$_{ar}$), 6.80-6.88 (m, 2H, Phe-C*H*$_{ar}$), 7.01-7.13 (m, 4H, N*H*, 2 × Phe-C*H*$_{ar}$, Pyr-C*H*$_{ar}$), 8.35-8.44 (m, 1H, N*H*), 9.73 (br. s, 1H, Gua-N*H*), 10.41 (br. s, 1H, Gua-N*H*), 11.65 (s, 1H, Pyr-N*H*), 11.82 (s, 1H, CO_2*H*), 12.01 (s, 1H, Gua-N*H*) ppm.
^{13}C-NMR	**(125 MHz, DMSO-d_6):** δ = 22.8 (Lys-*C*H$_2$), 28.0 (Boc-*C*H$_3$), 28.5 (Boc-*C*H$_3$), 29.9 (Lys-*C*H$_2$), 35.0 (Lys-*C*H$_2$), 39.0 (Phe-*C*H$_2$), 40.5 (Lys-*C*H$_2$), 52.0 (Lys-*C*H), 55.0 (Phe-*C*H), 78.9 (Boc-*C*$_q$), 84.9 (Boc-*C*$_q$), 109.8 (Pyr-*C*H), 119.6 (Pyr-*C*H), 126.5 (Phe-*C*H$_{ar}$), 127.2 (Phe-*C*H$_{ar}$), 129.2 (Phe-*C*H$_{ar}$), 131.63 (Pyr-*C*$_q$), 136.6 (Phe-*C*$_q$), 155.3 (*C*$_q$), 156.0 (*C*$_q$), 157.7 (*C*$_q$), 159.1 (*C*$_q$), 161.7 (*C*$_q$), 171.1 (*C*$_q$), 178.8 (*C*$_q$) ppm.
HR-MS	**(neg. ESI-ToF):** m/z = 670.3342 $[M - H^+]^-$

	calculated for $C_{32}H_{44}N_7O_9^-$: 670.3206
FT-IR	**(ATR):** $\tilde{\nu}$ = 3145 (br. w), 2948 (s), 1683 (s), 1623 (s), 1521 (s), 1247 (s), 1199 (s), 1143 (s), 971 (s), 833 (s), 771 (s) cm^{-1}.

Synthesis of the Trivalent DNA-Ligand 207 (Phe-Lys-GCP)

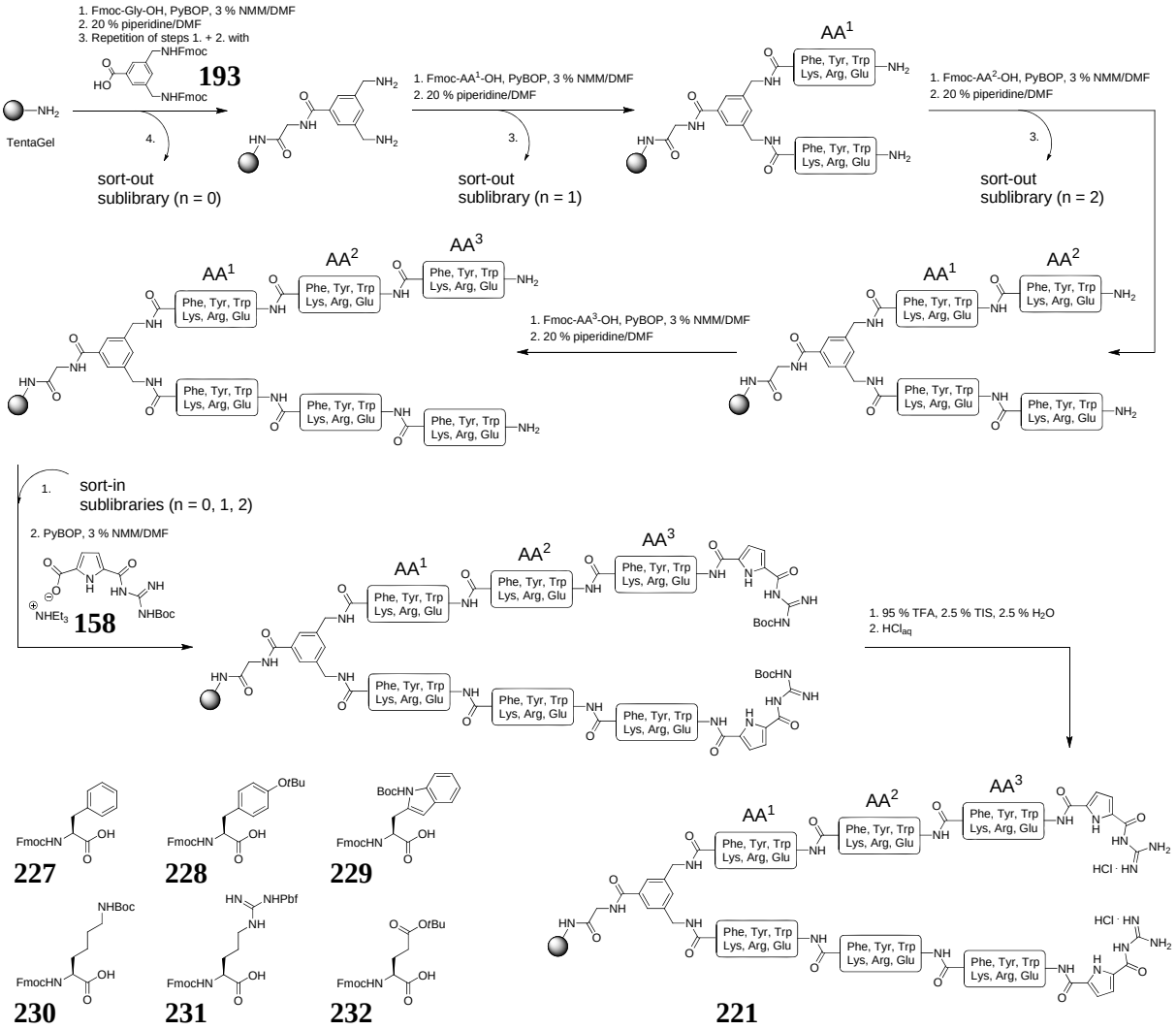


To a solution of **216** (71 mg, 0.11 mmol, 3 eq) and HATU (41 mg, 0.11 mmol, 3 eq) in DMF (5 ml) 150 µl of DIPEA were added and the mixture was stirred for 1 h under argon atmosphere. After addition of the aromatic triamine scaffold **210** (6 mg 0.04 mmol, 1 eq) the solution was stirred for another 36 h under argon atmosphere. **210** was obtained as described in literature.[216,217] Water (10 ml) was added and the resulting suspension extracted with ethyl acetate (3 × 25 ml). The combined organic phases were dried with Na_2SO_4 and the solvent was removed in vacuum. The resulting solid was purified by MPLC on normal phase silica gel (cyclohexane to ethyl acetate within 30 min), dissolved in 50 % TFA/DCM (10 ml) and stirred for 16 h under argon atmosphere. After removal of the solvent in vacuum the crude product was purified by MPLC (RP18, water to methanol within 30 min, 0.1 % TFA), re-dissolved in 1 M aqueous HCl and freeze dried. The last step was repeated twice and after an additional lyophilization in pure water **207** was obtained as a white, voluminous hydrochloride salt (16 mg, 0.01 mmol, 26 %).

$C_{75}H_{96}N_{24}O_{12}$	1744.48 g/mol
Yield	26 %
Mp.	> 250° C (decomposition)
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 0.94-1.10 (m, 6H, 3 × Lys-CH_2), 1.17-1.62 (m, 12H, 63 × Lys-CH_2), 2.63-3.11 (m, 12 H, 3 × Lys-CH_2, 3 × Phe-CH_2); 4.07-4.60 (m, 12H, 3 × Lys-C*H*, 3 × CH_2, 3 × Phe-C*H*), 6.83-6.91 (m, 3H, 3 × Pyr-CH_{ar}), 6.97 (s, 3H, 3 × CH_{ar}), 7.01 (s, 3H, 3 × CH_{ar}), 7.06-7.27 (m, 18H, 3 × Pyr-C*H*, 15 × Phe-CH_{ar}), 7.70 (br, s, 9H, 3 × Lys-NH_3), 8.30-8.67 (m, 12H, 3 × N*H*, 3 × Pyr-N*H,* 6 × Gua-N*H*), 11.38 (s, 3H, 3 × Pyr-N*H*), 12.48 (s, 3H, 3 × Gua-N*H*) ppm.

^{13}C-NMR	**(125 MHz, DMSO-d_6):** δ = 22.2 (Lys-*C*H_2), 26.6 (Lys-*C*H_2), 31.4 (Lys-*C*H_2), 37.8 (Phe-*C*H_2), 38.7 (Lys-*C*H_2), 42.2 (*C*H_2), 52.6 (Lys-*C*H), 53.9 (Phe-*C*H), 113.7 (Pyr-*C*H), 115.0 (Pyr-*C*H), 125.0 (*C*H_{ar}), 126.3 (Phe-*C*H_{ar}), 127.9 (Phe-*C*H_{ar}), 129.4 (*C*H_{ar}), 137.9 (Phe-C_q), 139.3 (C_q), 139.5 (C_q), 158.4 (C_q), 158.7 (C_q), 158.9 (C_q), 171.0 (C_q), 171.4 (C_q) ppm.
HR-MS	**(pos. ESI-ToF):** m/z = 1525.7583 $[M - 5\ HCl - Cl^-]^+$ calculated for $C_{75}H_{97}N_{24}O_{12}^+$: 1525.7712
FT-IR	**(ATR):** $\tilde{\nu}$ = 3152 (br. m), 2905 (w), 2362 (s), 2336 (s), 1701 (s), 1685 (s), 1657 (s), 1541 (s), 1517 (s), 1289 (m), 828 (w), 685 (m) cm^{-1}.

7.9 SPPS OF THE DNA-BINDER LIBRARY

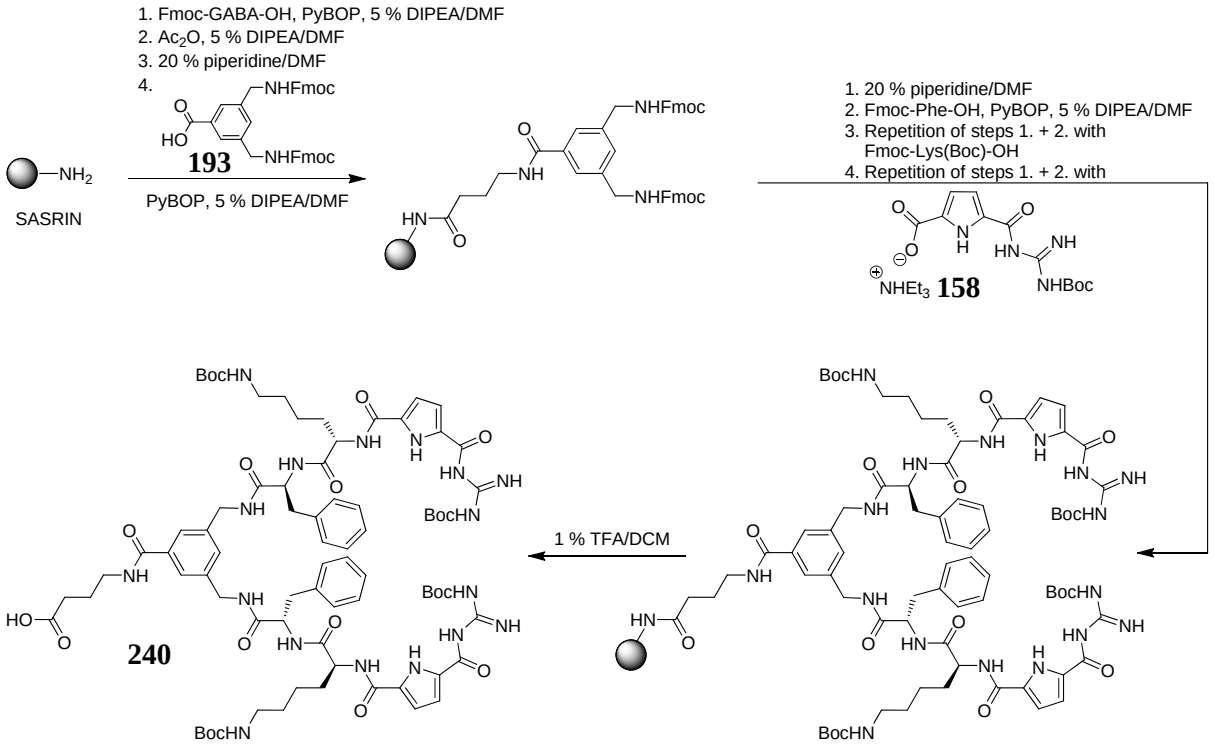


The synthesis of the library was carried out by means of standard Fmoc SPPS (chapter 7.2), which was altered slightly due to the IRORI MikroKan setting. First an isopycnic solution of Tentagel S NH_2 (9.10 g, 0.32 mmol/g) was prepared in a mixture of hexane/DCM (1/8, 104 mL) and 400 µL thereof were filled into each of the 259 MikroKans which were then equipped with a radiofrequency chip. After scanning of the MikroKans with the IRORI scanning station the resin was randomly split into three equal parts and swollen in DMF (80 mL) for 20 h in three Schlenck glass reaction vessels equipped with a frit. Then the attachment of the first amino acid Fmoc-Gly-OH (1.44 g per reaction vessel, 4.83 mmol, 5 eq) was conducted with PyBOP (2.54 g per reaction vessel, 4.83 mmol, 5 eq) in 3 % NMM/DMF (80 mL) by shaking the mixture for 20 h under argon atmosphere. After washing the beads three times for 10 min with DMF (80 mL) the coupling and washing steps were repeated another two times. 12 randomly chosen MicroKans were opened and a negative Kaiser test revealed complete conversion. This test was carried out after each coupling and deprotection step and in all cases led to the expected result. Fmoc deprotection was achieved by treating the solid support with 20 % piperidine/DMF (80 mL) two times for 1 h. The MikroKans were then thoroughly washed three times with DMF (80 mL), three times with DCM (80 mL), and another three times with DMF (80 mL) each time for 10 min. Afterwards the template **193**

(1.81 g per reaction vessel, 2.90 mmol, 3 eq) was coupled similarly with PyBOP (1.51 g per reaction vessel, 2.90 mmol, 3 eq). This and all following coupling steps were repeated twice. Before carrying out the combinatorial steps the MikroKan containing the compound without amino acids in its side arms was identified and removed after the Fmoc deprotection step. The remaining MikroKans were split into six parts according to the IRORI software and each compartment was reacted with either Fmoc-Phe-OH (**227**, 1.87 g, 4.83 mmol, 10 eq), Fmoc-Tyr(*t*Bu)-OH (**228**, 2.21 g, 4.83 mmol, 10 eq), Fmoc-Trp(Boc)-OH (**229**, 2.54 g, 4.83 mmol, 10 eq), Fmoc-Lys(Boc)-OH (**230**, 2.26 g, 4.83 mmol, 10 eq), Fmoc-Arg(Pbf)-OH (**231**, 3.12 g, 4.83 mmol, 10 eq), or Fmoc-Glu(OtBu)-OH (**232**, 2.05 g, 4.83 mmol, 10 eq) with PyBOP (2.51 g, 4.83 mmol, 10 eq) in 3 % NMM/DMF (50 mL) under argon atmosphere for 20 h. After the following Fmoc deprotection the sublibrary with just one amino acid per side chain was removed and the remaining MikroKans were again split into six parts and reacted as described above. After the next Fmoc deprotection step the next sublibrary containing members with two amino acids per arm was removed and another combinatorial step was carried out. After Fmoc deprotection all MikroKans were combined (containing the sublibraries) and randomly split into three parts. Finally, the binding motif Boc-GCP-OH (**158**, 2.39 g per reaction vessel, 5.80 mmol, 6 eq) was coupled with the help of PyBOP (3.02 g per reaction vessel, 5.80 mmol, 6 eq) in 3 % NMM/DMF (80 mL). The last coupling step was repeated three times in total. Before the last repetition Boc-GCP-OH was heated for 20 min to 50° C together with PyBOP in 3 % NMM/DMF whereupon it dissolved with red colour and then added to the resin. The MikroKans were then washed three times for 10 min with DMF (80 mL), three times with DCM (80 mL), three times with MeOH (80 mL), another three times with (80 mL) and then dried under reduced pressure for 4 h. Afterwards the beads were treated with 95 % TFA containing 2.5 % water and 2.5 % TIS two times for 2 h and washed two times for 10 min with DCM (80 mL), two times with MeOH (80 mL), three times with 0.1 N aqueous HCl (80 mL), two times with water (80 mL), two times with MeOH (80 mL), and two times with DCM (80 mL). Finally, the library **221** was dried under reduced pressure for 24 h.

7.10 SYNTHESIS OF AMPHIPHILIC GENE CARRIERS

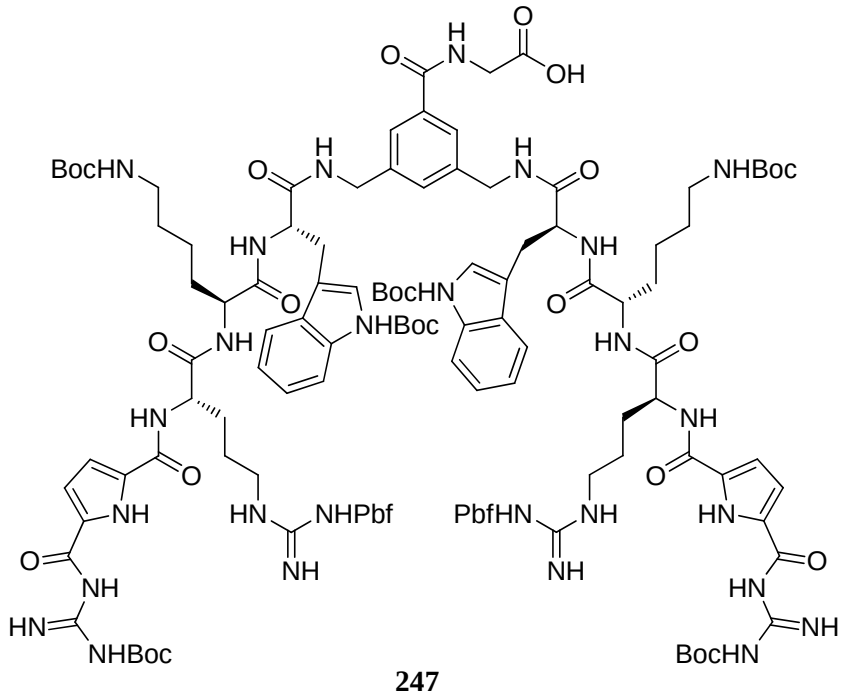
SPPS of the Fully Protected Headgroup 240 (Phe-Lys-GCP)



The synthesis of fully protected 240 was carried out on 150 mg SASRIN resin (1.3 mmol/g) according to the microwave-assisted Fmoc SPPS procedure described in chapter 7.5. After swelling of the resin in DCM/DMF (7/3, 5 mL), the attachment of the first amino acid Fmoc-GABA-OH (190 mg, 0.59 mmol, 3 eq) was carried out without microwave irradiation with DIC (91 µl, 0.59 mmol, 3 eq), HOBt (79 mg, 0.59 mmol, 3 eq) and DMAP (catalytic) in DCM/DMF (7/3, 5 mL). The coupling step was repeated four times in total. Afterwards the resin was treated with Ac_2O (55 µl, 0.59 mmol, 3 eq) in 5 % DIPEA/DMF (5 ml). The following removal of the Fmoc group and the coupling of the template **193** (365 mg, 0.59 mmol, 3 eq) with PyBOP (304 mg, 0.59 mmol, 3 eq) as well as the couplings of the following amino acids Fmoc-Phe-OH (453 mg, 1.17 mmol, 6 eq), Fmoc-Lys(Boc)-OH (548 mg, 1.17 mmol, 6 eq) and the binding motif Boc-GCP-OH (**158**, 465 mg, 1.17 mmol, 6 eq) were all carried out under argon atmosphere and microwave irradiation with PyBOP (609 mg, 1.17 mmol, 6 eq) in 5 % DIPEA/DMF (5 mL). All coupling steps were repeated twice. The cleavage of the raw product was carried out without microwave irradiation by treating the resin with 1 % TFA/DCM twelve times for 15 min followed by washing three times with DCM, three times with MeOH, and three times with DCM. The organic phases were combined and reduced via distillation in vacuum. The product was precipitated by adding water, purified via MPLC (RP8, H_2O → MeOH) to obtain the product **240** as voluminous white solid (79 mg, 0.05 mmol, 26 %).

$C_{77}H_{105}N_{17}O_{19}$	1572.76 g/mol
Yield	26 %
Mp.	218° C
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 1.07-1.20 (m, 4H, 2 × Lys-CH_2), 1.30-1.38 (m, 22H, 2 × Boc-*t*Bu, 2 × Lys-CH_2), 1.46 (s, 18H, 2 × Boc-*t*Bu), 1.50-1.65 (m, 4H, 2 × Lys-CH_2), 1.70-1.77 (m, 2H, CH_2), 2.24 (t, ^{3}J = 7.47 Hz, 2H, CH_2), 2.81-2.89 (m, 6H, 2 × Lys-CH_2, CH_2), 3.05 (dd, $^{3}J_1$ = 4.75 Hz, $^{3}J_2$ = 13.67 Hz, 2H, Phe-CH_2), 3.22-3.30 (m, 2H, Phe-CH_2), 4.22-4.35 (m, 6H, 2 × CH_2, 2 × Lys-C*H*), 4.53 (dd, $^{3}J_1$ = 8.60 Hz, $^{3}J_2$ = 13.76 Hz, 2H, 2 × Phe-C*H*), 6.73 (t, ^{3}J = 5.52 Hz, 2H, 2 × Lys-N*H*Boc), 6.78-6.82 (m, 2H, 2 × Pyr-CH_{ar}), 6.83-6.86 (m, 2H, 2 × Pyr-CH_{ar}), 7.10-7.21 (m, 11H, 10 × Phe-CH_{ar}, CH_{ar}), 7.61 (s, 2H, 2 × CH_{ar}), 8.08 (d, ^{3}J = 7.87 Hz, 2H, 2 × N*H*), 8.36-8.47 (m, 5H, 5 × N*H*), 8.57 (s, 2H, N*H*), 9.34 (s, 2H, N*H*), 11.37 (br. s, 2H, N*H*), 11.63 (s, 2H, Pyr-N*H*) ppm.
^{13}C-NMR	**(125 MHz, DMSO-d_6):** δ = 22.81 (*C*H_2), 24.49 (*C*H_2), 27.55 (Boc-*C*H_3), 27.68 (Boc-*C*H_3), 29.19 (*C*H_2), 31.04 (*C*H_2), 37.29 (*C*H_2), 38.50 (*C*H_2), 39.54 (*C*H_2), 42.04 (*C*H_2), 53.12 (CH), 53.93 (CH), 77.22 (Boc-C_q), 80.76 (Boc-C_q), 112.36 (*C*H_{ar}), 112.93 (*C*H_{ar}), 113.58 (*C*H_{ar}), 124.70 (*C*H_{ar}), 126.09 (*C*H_{ar}), 127.85 (*C*H_{ar}), 129.04 (*C*H_{ar}), 129.11 (*C*H_{ar}), 134.65 (C_q), 137.58 (C_q), 139.08 (C_q), 155.42 (*C*O), 159 .63 (*C*O), 166.02 (*C*O), 170.72 (*C*O), 171.65 (*C*O), 174.11 (*C*O) ppm.
HR-MS	**(neg. ESI-ToF):** m/z = 1570.7765 $[M - H^+]^-$ calculated for $C_{77}H_{104}N_{17}O_{19}^-$: 1570.7700
FT-IR	**(ATR):** $\tilde{\nu}$ = 3291 (br. w), 2930 (w), 1635 (br. s), 1523 (br. s), 1456 (m), 1394 (m), 1367 (m), 1240 (br. s), 1147 (br. s), 1047 (m), 841 (m), 755 (s), 700 (s) cm^{-1}.

SPPS of the Fully Protected Headgroup 247 (Trp-Lys-Arg-GCP)

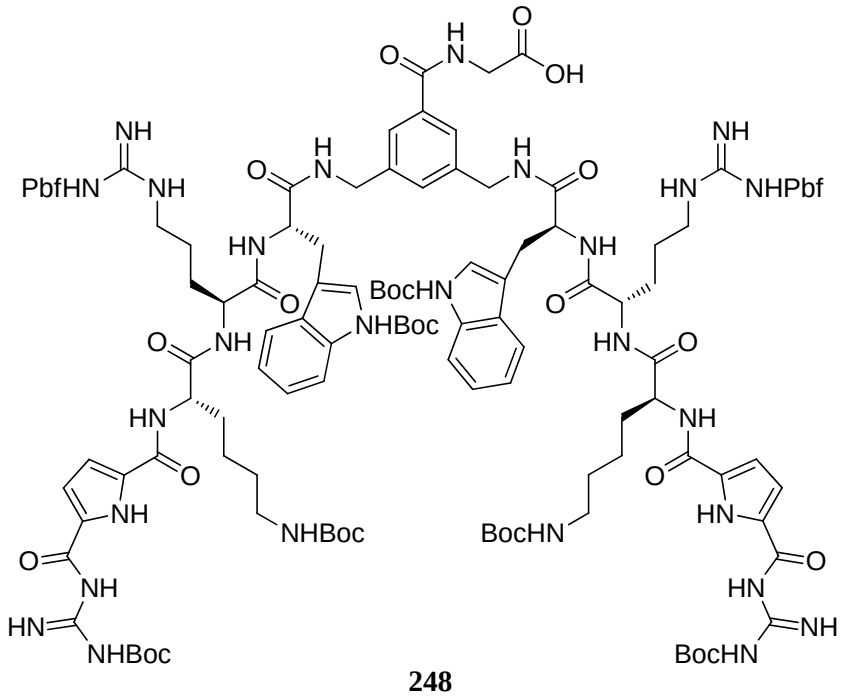


247

The synthesis of fully protected **247** was synthesized accordingly on 150 mg SASRIN resin (1.3 mmol/g). After swelling of the resin in DCM/DMF (7/3, 5 mL), the attachment of the first amino acid Fmoc-Gly-OH (174 mg, 0.59 mmol, 3 eq) was carried out without microwave irradiation with DIC (91 µl, 0.59 mmol, 3 eq), HOBt (79 mg, 0.59 mmol, 3 eq) and DMAP (catalytic) in DCM/DMF (7/3, 5 mL). The coupling step was repeated four times in total. Afterwards the resin was treated with Ac_2O (55 µl, 0.59 mmol, 3 eq) in 5 % DIPEA/DMF (5 ml). The following removal of the Fmoc group and the coupling of the template **193** (365 mg, 0.59 mmol, 3 eq) with PyBOP (304 mg, 0.59 mmol, 3 eq) as well as the couplings of the following amino acids Fmoc-Trp(Boc)-OH (616 mg, 1.17 mmol, 6 eq), Fmoc-Lys(Boc)-OH (548 mg, 1.17 mmol, 6 eq), and Fmoc-Arg(Pbf)-OH (759 mg, 1.17 mmol, 6 eq) were all carried out under argon atmosphere and microwave irradiation with PyBOP (609 mg, 1.17 mmol, 6 eq) in 5 % DIPEA/DMF (5 mL). The binding motif Boc-GCP-OH (**158**, 233 mg, 0.59 mmol, 3 eq) was coupled similarly with PyBOP (304 mg, 0.59 mmol, 3 eq). All coupling steps were repeated twice. The cleavage of the raw product was carried out without microwave irradiation by treating the resin with 1 % TFA/DCM containing 2.5 % water and 2.5 % TIS four times for 2 h and another time for 20 h followed by washing three times with DCM, three times with MeOH, and three times with DCM. The organic phases were combined and reduced via distillation in vacuum. The product was precipitated by adding water, purified via MPLC (RP8, $H_2O \rightarrow$ MeOH) to obtain the product **247** as voluminous white solid (62 mg, 0.02 mmol, 12 %).

$C_{127}H_{175}N_{27}O_{31}S_2$	2640.04 g/mol
Yield	12 %
Mp.	225-230° C (decomposition)
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 1.20-1.30 (m, 8H, 4 × CH_2), 1.33 (s, 18H, 2 × Boc-*t*Bu), 1.35-1.40 (m, 12H, 4 × Pbf-CH_3), 1.41-1.44 (m, 4H, 2 × CH_2), 1.45 (s, 18H, 2 × Boc-*t*Bu), 1.50-1.54 (m, 4H, 2 × CH_2), 1.58 (s, 18H, 2 × Boc-*t*Bu), 1.61-1.78 (m, 4H, 2 × CH_2), 1.95 (s, 6H, 2 × Pbf-CH_3), 2.38 (s, 6H, 2 × Pbf-CH_3), 2.45 (s, 6H, 2 × Pbf-CH_3), 2.78 (dd, $^3J_1$ = 6.22 Hz, $^3J_2$ = 11.91 Hz, 4H, 2 × Lys-CH_2), 2.89 (s, 4H, 2 × Pbf-CH_2), 2.96-3.11 (m, 8H, 2 × Arg-CH_2, 2 × Trp-CH_2), 3.87-3.95 (m, 2H, Gly-CH_2), 4.20 (dd, $^3J_1$ = 7.34 Hz, $^3J_2$ = 13.16 Hz, 4H, 2 × CH_2), 4.40 (s, 2H, 2 × C*H*), 4.29-4.36 (m, 2H, 2 × C*H*), 4.59 (dd, $^3J_1$ = 6.73 Hz, $^3J_2$ = 13.83 Hz, 2H, 2 × C*H*), 6.57 (s, 2H, 2 × N*H*), 6.67 (t, ^3J = 5.34 Hz, 2H, 2 × Trp-CH_{ar}), 6.81 (s, 4H, 4 × Pyr-CH_{ar}), 6.95 (s, 1H, CH_{ar}), 7.11 (s, 2H, 2 × N*H*), 7.20 (t, ^3J = 7.49 Hz, 2H, 2 × Trp-CH_{ar}), 7.25 (t, ^3J = 7.53 Hz, 2H, 2 × Trp-CH_{ar}), 7.49 (s, 2H, 2 × CH_{ar}), 7.62-7.69 (m, 4H, 4 × Trp-CH_{ar}), 8.00 (d, ^3J = 8.47 Hz, 2H, 2 × N*H*), 8.14 (s, 4H, 4 × N*H*), 8.31 (s, 3H, 3 × N*H*), 8.55-8.68 (m, 6H, 6 × N*H*), 9.32 (s, 2H, N*H*), 11.75 (s, 2H, Pyr-N*H*) ppm.
^{13}C-NMR	**(125 MHz, DMSO-d_6):** δ = 9.05 (*C*H_3), 12.12 (*C*H_3), 17.48 (*C*H_3), 18.82 (*C*H_3), 22.51 (*C*H_2), 25.41 (*C*H_2), 27.13 (*C*H_2), 27.58 (*C*H_3), 27.67 (*C*H_3), 27.95 (*C*H_3), 27.99 (*C*H_2), 28.15 (*C*H_2), 28.86 (*C*H_2), 29.07 (*C*H_2), 31.62 (*C*H_3), 40.08 (*C*H_2), 40.25 (*C*H_2), 42.32 (*C*H_2), 45.61 (*C*H_2), 52.41 (*C*H), 52.55 (*C*H), 52.65 (*C*H), 77.18 (C_q), 83.35 (C_q), 86.13 (C_q), 112.93 (*C*H_{ar}), 113.52 (*C*H_{ar}), 114.36 (*C*H_{ar}), 114.48 (C_q), 116.09 (C_q), 116.29 (C_q), 119.26 (*C*H_{ar}), 122.27 (*C*H_{ar}), 123.69 (*C*H_{ar}), 124.16 (*C*H_{ar}), 130.12 (C_q), 131.33 (C_q), 133.68 (C_q), 134.08 (C_q), 134.45 (C_q), 137.14 (C_q), 139.08 (C_q), 148.91 (*C*O), 155.36 (*C*O), 156.08 (*C*O), 157.28 (*C*O), 158.30 (*C*O), 159.53 (*C*O), 170.94 (*C*O), 171.46 (*C*O), 171.64 (*C*O), 174.40 (*C*O) ppm.
MS	**(neg. and pos. ESI-ToF):** not found
FT-IR	**(ATR):** $\tilde{\nu}$ = 3290 (br. m), 2924 (m), 2853 (m), 2161 (w), 2035 (w), 1654 (s), 1542 (s), 1445 (m), 1275 (m), 1200 (s), 1132 (s), 838 (m), 800 (m), 722 (m) cm^{-1}.

SPPS of the Fully Protected Headgroup 248 (Trp-Arg-Lys-GCP)

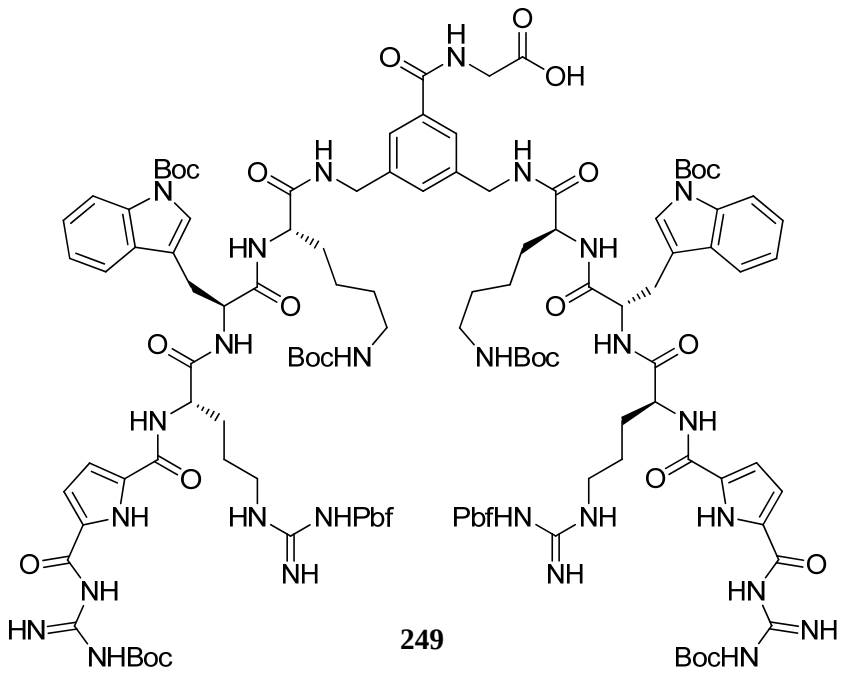


248

The synthesis of fully protected **248** was synthesized accordingly on 150 mg SASRIN resin (1.3 mmol/g). After swelling of the resin in DCM/DMF (7/3, 5 mL), the attachment of the first amino acid Fmoc-Gly-OH (174 mg, 0.59 mmol, 3 eq) was carried out without microwave irradiation with DIC (91 µl, 0.59 mmol, 3 eq), HOBt (79 mg, 0.59 mmol, 3 eq) and DMAP (catalytic) in DCM/DMF (7/3, 5 mL). The coupling step was repeated four times in total. Afterwards the resin was treated with Ac_2O (55 µl, 0.59 mmol, 3 eq) in 5 % DIPEA/DMF (5 ml). The following removal of the Fmoc group and the coupling of the template **193** (365 mg, 0.59 mmol, 3 eq) with PyBOP (304 mg, 0.59 mmol, 3 eq) as well as the couplings of the following amino acids Fmoc-Trp(Boc)-OH (616 mg, 1.17 mmol, 6 eq), Fmoc-Arg(Pbf)-OH (759 mg, 1.17 mmol, 6 eq), and Fmoc-Lys(Boc)-OH (548 mg, 1.17 mmol, 6 eq) were all carried out under argon atmosphere and microwave irradiation with PyBOP (609 mg, 1.17 mmol, 6 eq) in 5 % DIPEA/DMF (5 mL). The binding motif Boc-GCP-OH (**158**, 233 mg, 0.59 mmol, 3 eq) was coupled similarly with PyBOP (304 mg, 0.59 mmol, 3 eq). All coupling steps were repeated twice. The cleavage of the raw product was carried out without microwave irradiation by treating the resin with 1 % TFA/DCM containing 2.5 % water and 2.5 % TIS four times for 2 h and another time for 20 h followed by washing three times with DCM, three times with MeOH, and three times with DCM. The organic phases were combined and reduced via distillation in vacuum. The product was precipitated by adding water, purified via MPLC (RP8, $H_2O \rightarrow$ MeOH) to obtain the product **248** as voluminous white solid (62 mg, 0.02 mmol, 12 %).

$C_{127}H_{175}N_{27}O_{31}S_2$	2640.04 g/mol
Yield	12 %
Mp.	225-230° C (decomposition)
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 1.19-1.29 (m, 12H, 6 × Lys-CH_2), 1.32 (s, 18H, 2 × Boc-*t*Bu), 1.38 (s, 12H, 4 × Pbf-CH_3), 1.46 (s, 18H, 2 × Boc-*t*Bu), 1.47-1.53 (m, 4H, 2 × CH_2), 1.56 (s, 18H, 2 × Boc-*t*Bu), 1.59-1.67 (m, 4H, 2 × CH_2), 1.97 (s, 6H, 2 × Pbf-CH_3), 2.39 (s, 6H, 2 × Pbf-CH_3), 2.45 (s, 6H, 2 × Pbf-CH_3), 2.81-2.89 (m 4H, 2 × Lys-CH_2), 2.92 (s, 4H, 2 × Pbf-CH_2), 2.93-3.04 (m, 8H, 2 × Arg-CH_2, 2 × Trp-CH_2), 3.84-3.94 (m, 2H, Gly-CH_2), 4.18-4.30 (m, 6H, 2 × CH, 2 × CH_2), 4.40 (s, 2H, 2 × CH), 4.34 (s, 2H, 2 × CH), 4.60 (dd, $^3J_1$ = 7.12 Hz, $^3J_2$ = 13.49 Hz, 2H, 2 × CH), 6.44 (s, 4H, 4 × NH), 6.71 (t, ^3J = 5.49 Hz, 2H, 2 × Trp-CH_{ar}), 6.78-6.84 (m, 4H, 4 × Pyr-CH_{ar}), 7.14 (s, 1H, CH_{ar}), 7.19 (t, ^3J = 7.41 Hz, 2H, 2 × Trp-CH_{ar}), 7.27 (t, ^3J = 7.73 Hz, 2H, 2 × CH_{ar}), 7.62-7.67 (m, 4H, 4 × Trp-CH_{ar}), 7.99 (d, ^3J = 8.13 Hz, 2H, 2 × NH), 8.05 (d, ^3J = 5.82 Hz, 2H, 2 × NH), 8.07-8.15 (s, 3H, 3 × NH), 8.44 (s, 2H, 2 × NH), 8.57 (s, 6H, 6 × NH), 9.32 (s, 2H, 2 × NH), 11.67 (s, 2H, Pyr-NH) ppm.
^{13}C-NMR	**(125 MHz, DMSO-d_6):** δ = 8.63 (*C*H$_3$), 12.15 (*C*H$_3$), 17.48 (*C*H$_3$), 18.83 (*C*H$_3$), 22.95 (*C*H$_2$), 25.24 (*C*H$_2$), 27.36 (*C*H$_2$), 27.57 (*C*H$_3$), 27.67 (*C*H$_3$), 27.94 (*C*H$_2$), 28.14 (*C*H$_3$), 28.18 (*C*H$_3$), 29.00 (CH_2), 29.15 (*C*H$_3$), 31.25 (*C*H$_2$), 39.93 (*C*H$_2$), 40.02 (*C*H$_2$), 42.00 (*C*H$_2$), 42.34 (*C*H$_2$), 45.65 (*C*H$_2$), 52.33 (*C*H), 52.55 (*C*H), 53.00 (*C*H), 77.19 (C_q), 83.33 (C_q), 83.49 (C_q), 86.16 (C_q), 112.90 (*C*H$_{ar}$), 113.62 (*C*H$_{ar}$), 114.47 (*C*H$_{ar}$), 116.13 (C_q), 116.26 (C_q), 122.97 (*C*H$_{ar}$), 123.73 (*C*H$_{ar}$), 124.20 (*C*H$_{ar}$), 124.50 (*C*H$_{ar}$), 131.35 (C_q), 133.09 (C_q), 134.46 (C_q), 134.94 (C_q), 137.18 (C_q), 139.12 (C_q), 148.92 (*C*O), 155.42 (*C*O), 155.96 (*C*O), 157.33 (*C*O), 157.74 (*C*O), 158.36 (*C*O), 159.65 (*C*O), 170.86 (*C*O), 171.45 (*C*O), 172.09 (*C*O) ppm.
MS	**(neg. and pos. ESI-ToF):** not found
FT-IR	**(ATR):** $\tilde{\nu}$ = 3311 (br. m), 2977 (m), 2931 (m), 2159 (w), 1634 (s), 1541 (s), 1453 (m), 1368 (m), 1242 (s), 1148 (s), 1087 (s), 834 (w), 750 (m) cm^{-1}.

SPPS of the Fully Protected Headgroup 249 (Lys-Trp-Arg-GCP)

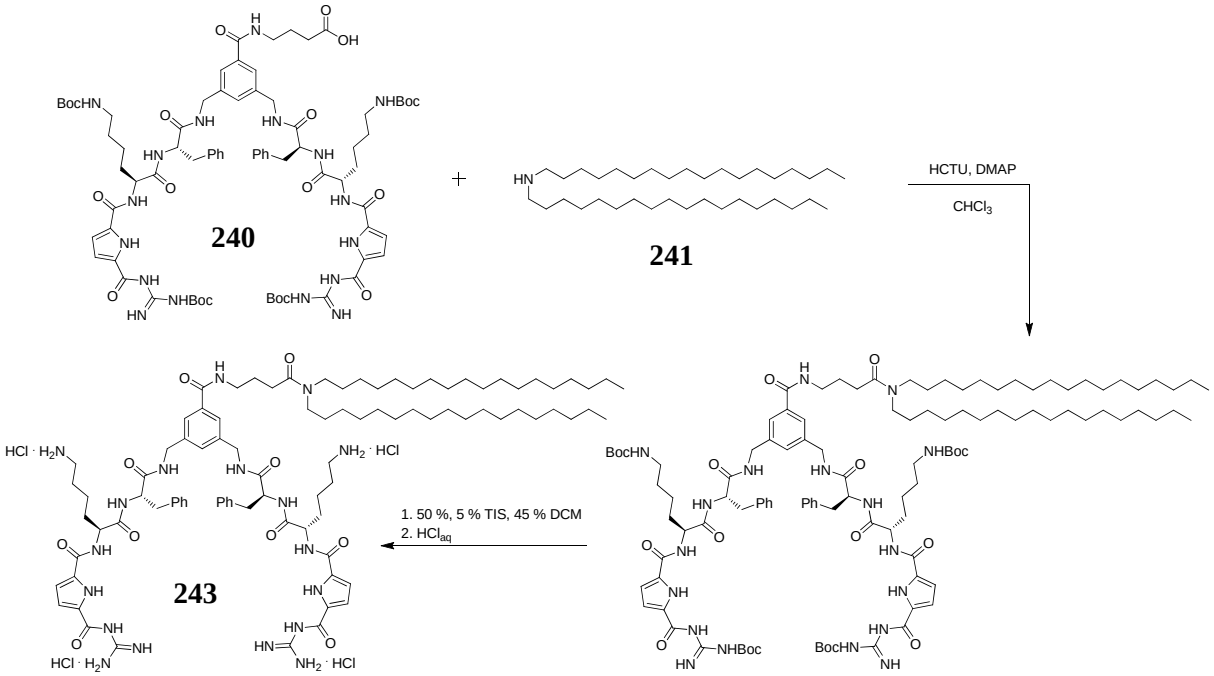


249

The synthesis of fully protected **248** was synthesized accordingly on 150 mg SASRIN resin (1.3 mmol/g). After swelling of the resin in DCM/DMF (7/3, 5 mL), the attachment of the first amino acid Fmoc-Gly-OH (174 mg, 0.59 mmol, 3 eq) was carried out without microwave irradiation with DIC (91 µl, 0.59 mmol, 3 eq), HOBt (79 mg, 0.59 mmol, 3 eq) and DMAP (catalytic) in DCM/DMF (7/3, 5 mL). The coupling step was repeated four times in total. Afterwards the resin was treated with Ac_2O (55 µl, 0.59 mmol, 3 eq) in 5 % DIPEA/DMF (5 ml). The following removal of the Fmoc group and the coupling of the template **193** (365 mg, 0.59 mmol, 3 eq) with PyBOP (304 mg, 0.59 mmol, 3 eq) as well as the couplings of the following amino acids Fmoc-Lys(Boc)-OH (548 mg, 1.17 mmol, 6 eq), Fmoc-Trp(Boc)-OH (616 mg, 1.17 mmol, 6 eq), and Fmoc-Arg(Pbf)-OH (759 mg, 1.17 mmol, 6 eq) were all carried out under argon atmosphere and microwave irradiation with PyBOP (609 mg, 1.17 mmol, 6 eq) in 5 % DIPEA/DMF (5 mL). The binding motif Boc-GCP-OH (**158**, 233 mg, 0.59 mmol, 3 eq) was coupled similarly with PyBOP (304 mg, 0.59 mmol, 3 eq). All coupling steps were repeated twice. The cleavage of the raw product was carried out without microwave irradiation by treating the resin with 1 % TFA/DCM containing 2.5 % water and 2.5 % TIS four times for 2 h and another time for 20 h followed by washing three times with DCM, three times with MeOH, and three times with DCM. The organic phases were combined and reduced via distillation in vacuum. The product was precipitated by adding water, purified via MPLC (RP8, H_2O → MeOH) to obtain the product **248** as voluminous white solid (72 mg, 0.02 mmol, 14 %).

$C_{127}H_{175}N_{27}O_{31}S_2$	2640.04 g/mol
Yield	14 %
Mp.	225-230° C (decomposition)
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 1.21-1.27 (m, 4H, 2 × Lys-CH_2), 1.28-1.32 (m, 4H, 2 × Lys-CH_2), 1.34 (s, 18H, 2 × Boc-*t*Bu), 1.37 (s, 12H, 4 × Pbf-CH_3), 1.38 (s, 18H, 2 × Boc-*t*Bu), 1.48-1.51 (m, 4H, 2 × Lys-CH_2), 1.54 (s, 18H, 2 × Boc-*t*Bu), 1.69-1.71 (m, 4H, 2 × Lys-CH_2), 1.95 (s, 6H, 2 × Pbf-CH_3), 2.38 (s, 6H, 2 × Pbf-CH_3), 2.44 (s, 6H, 2 × Pbf-CH_3), 2.84 (dd, $^3J_1$ = 6.98 Hz, $^3J_2$ = 12.72 Hz, 4H, 2 × Lys-CH_2), 2.89 (s, 4H, 2 × Pbf-CH_2), 2.93-3.05 (m, 8H, 2 × Arg-CH_2, 2 × Trp-CH_2), 3.91 (d, ^3J = 5.07 Hz, 2H, Gly-CH_2), 4.21-4.33 (m, 6H, 2 × C*H*, 2 × CH_2), 4.40 (s, 2H, 2 × C*H*), 4.67 (dd, $^3J_1$ = 7.58 Hz, $^3J_2$ = 13.64 Hz, 2H, 2 × C*H*), 6.35 (s, 2H, 2 × N*H*), 6.73 (t, ^3J = 5.14 Hz, 2H, 2 × Lys-N*H*Boc), 6.79 (s, 2H, 2 × Pyr-CH_{ar}), 6.84 (s, 2H, 2 × Pyr-CH_{ar}), 7.16 (t, ^3J = 7.42 Hz, 2H, 2 × Trp-CH_{ar}), 7.25 (t, ^3J = 7.48 Hz, 2H, 2 × Trp-CH_{ar}), 7.29 (s, 1H, CH_{ar}), 7.48 (s, 2H, 2 × CH_{ar}), 7.61-7.68 (m, 4H, 4 × Trp-CH_{ar}), 7.98 (d, ^3J = 7.85 Hz, 2H, 2 × N*H*), 8.06 (d, ^3J = 6.54 Hz, 2H, 2 × N*H*), 8.22 (d, ^3J = 7.41 Hz, 2H, 2 × N*H*), 8.41 (d, ^3J = 5.95 Hz, 4H, 4 × N*H*), 8.59 (s, 2H, 2 × NH), 8.69 (t, ^3J = 5.73 Hz, 1H, Gly-NH), 9.02 (s, 4H, 4 × N*H*), 9.36 (s, 2H, N*H*), 11.64 (s, 2H, Pyr-N*H*) ppm.
^{13}C-NMR	**(125 MHz, DMSO-d_6):** δ = 8.54 (*C*H_3), 12.13 (*C*H_3), 17.48 (*C*H_3), 18.82 (*C*H_3), 22.60 (*C*H_2), 25.49 (*C*H_2), 26.92 (*C*H_2), 27.53 (*C*H_3), 27.68 (*C*H_3), 28.17 (*C*H_3), 29.16 (*C*H_2), 29.24 (*C*H_2), 30.61 (*C*H_3), 31.69 (*C*H_2), 41.63 (*C*H_2), 41.87 (*C*H_2), 42.34 (*C*H_2), 45.67 (*C*H_2), 52.18 (*C*H), 52.40 (*C*H), 52.87 (*C*H), 77.25 (C_q), 83.25 (C_q), 86.18 (C_q), 113.00 (*C*H$_{ar}$), 113.60 (*C*H$_{ar}$), 114.30 (*C*H$_{ar}$), 116.12 (C_q), 116.34 (C_q), 119.28 (*C*H$_{ar}$), 122.27 (*C*H$_{ar}$), 123.71 (*C*H$_{ar}$), 124.11 (*C*H$_{ar}$), 124.19 (C_q), 124.39 (*C*H$_{ar}$), 128.92 (*C*H$_{ar}$), 129.39 (C_q), 130.22 (C_q), 131.37 (C_q), 133.89 (C_q), 134.02 (C_q), 134.50 (C_q), 137.17 (C_q), 139.43 (C_q), 148.92 (*C*O), 155.44 (*C*O), 155.93 (*C*O), 157.33 (*C*O), 157.60 (*C*O), 157.85 (*C*O), 158.08 (*C*O), 158.22 (*C*O), 159.37 (*C*O), 166.23 (*C*O), 170.83 (*C*O), 171.19 (*C*O), 171.37 (*C*O), 171.60 (*C*O) ppm.
MS	**(neg. and pos. ESI-ToF):** not found
FT-IR	**(ATR):** $\tilde{\nu}$ = 3318 (br. m), 2981 (m), 2506 (w), 1670 (s), 1541 (s), 1454 (m), 1369 (m), 1248 (m), 1128 (s), 1034 (m), 829 (m), 799 (m), 719 (m) cm^{-1}.

Synthesis of the Amphiphile 243 (Phe-Lys-GCP)

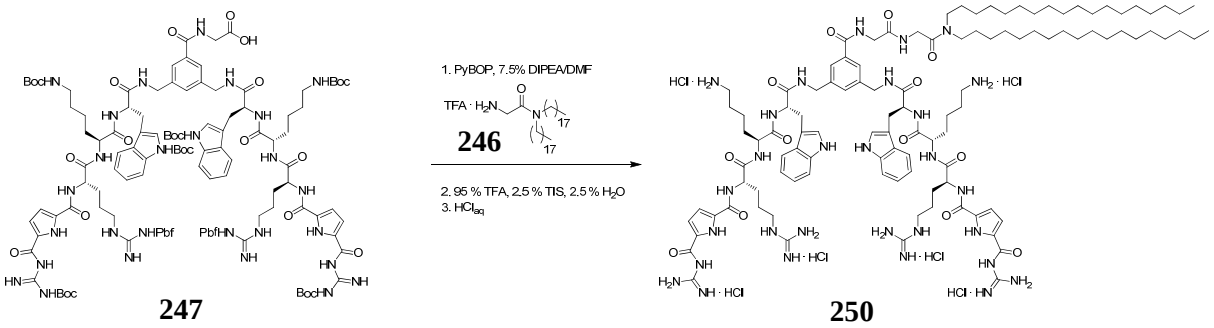


The protected headgroup **240** (65 mg, 0.04 mmol, 1 eq) and dioctadecylamine (**241**, 22 mg, 0.04 mmol, 1 eq) were suspended in $CHCl_3$. Under argon atmosphere HCTU (21 mg, 0.05 mmol, 1.2 eq) and DMAP (15 mg, 0.11 mmol, 3 eq) were added and the suspension was stirred for 20 h at room temperature. Then 10 mL DMF were added and the mixture was stirred for 4 h at 40° C (oil bath temperature). The chloroform was removed via distillation under reduced pressure and water (40 mL) was added. The resulting solid was filtrated, purified by MPLC (RP8, water/MeOH) and treated with 50 % TFA/DCM containing 5 % TIS for 9 h under argon atmosphere. The solvent was removed via distillation under reduced pressure, water was added and the solution was lyophilized. The raw product was then purified by MPLC (RP8, H_2O → MeOH, 0.01 % TFA) and dissolved two times in 0.1 N aqueous HCl and another time in water followed by lyophilization to obtain **243** (18 mg, 0.01 mmol, 24 %) as voluminous white solid.

$C_{93}H_{150}Cl_4N_{18}O_{10}$	1822.11 g/mol
Yield	24 %
Mp.	> 275° C (decomposition)
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 0.75-1.86 (m, 82H, 32 × C_{18}-CH_2, 6 × Lys-CH_2, 2 × C_{18}-CH_3), 1.93-2.07 (m, 2H, CH_2), 2.29 (t, ^3J = 7.21 Hz, 2H, CH_2), 2.35-2.47 (m, 2H, C_{18}-NCH_2), 2.56-2.65 (m, 2H, C_{18}-NCH_2), 2.67-2.80 (m, 4H, 2 × Lys-CH_2), 2.85 (dd, $^3J_1$ = 9.46 Hz, $^3J_2$ = 13.27 Hz, 2H, CH_2), 3.00-3.10 (m, 2H, Phe-CH_2), 3.14-3.19 (m, 2H, Phe-CH_2), 4.20-4.41 (m, 6H, 2 × CH_2, 2 × Lys-CH), 4.51-4.60 (m, 2H, 2 × Phe-CH), 6.84 (br. s, 2H, 2 × Pyr-CH_{ar}), 7.03-7.40 (m, 13H, 10 × Phe-

	C*H*$_{ar}$, C*H*$_{ar}$, 2 × Pyr-C*H*$_{ar}$), 7.41-8.82 (m, 15H, 8 × GCP-N*H*$_4^+$, 7 × N*H*), 11.31 (br. s, 2H, 2 × N*H*), 11.94 (br. s, 2H, 2 × Pyr-N*H*) ppm.
^{13}C-NMR	**(125 MHz, DMSO-d$_6$):** δ = 14.00 (*C*H$_3$), 22.19 (*C*H$_2$), 22.45 (*C*H$_2$), 25.11 (*C*H$_2$), 26.16 (*C*H$_2$), 26.41 (*C*H$_2$), 26.70 (*C*H$_2$), 27.32 (*C*H$_2$), 28.59 (*C*H$_2$), 28.65 (*C*H$_2$), 28.74 (*C*H$_2$), 28.79 (*C*H$_2$), 28.88 (*C*H$_2$), 28.94 (*C*H$_2$), 28.96 (*C*H$_2$), 28.99 (*C*H$_2$), 31.04 (*C*H$_2$), 31.36 (*C*H$_2$), 37.59 (*C*H$_2$), 38.83 (*C*H$_2$), 38.86 (*C*H$_2$), 42.20 (*C*H$_2$), 45.28 (*C*H$_2$), 47.33 (*C*H$_2$), 49.37 (*C*H$_2$), 52.96 (*C*H), 54.11 (*C*H), 113.71 (*C*H$_{ar}$), 124.82 (*C*H$_{ar}$), 126.32 (*C*H$_{ar}$), 128.04 (*C*H$_{ar}$), 129.25 (*C*H$_{ar}$), 137.65 (*C*$_q$), 139.22 (*C*$_q$), 158.25 (*C*O), 158.48 (*C*O), 158.75 (*C*O), 166.24 (*C*O), 170.97 (*C*O), 171.27 (*C*O), 171.51 (*C*O) ppm.
HR-MS	**(pos. ESI-ToF):** m/z = 1677.1617 [M – 3 HCl – Cl$^-$]$^+$ calculated for $C_{93}H_{147}N_{18}O_{10}^+$: 1677.1572
FT-IR	**(ATR):** $\tilde{\nu}$ = 3264 (br. w), 2921 (w), 2852 (w), 1698 (m), 1636 (s), 1541 (s), 1457 (m), 1277 (m), 1200 (m), 749 (s), 700 (s), 626 (s) cm^{-1}.

Synthesis of the Amphiphile 250 (Trp-Lys-Arg-GCP)



Fully protected **247** (20.0 mg, 7.6 µmol, 1 eq), Boc-Gly-dioctadecylamide **246** (6.5 mg, 9.3 mmol, 1.2 eq) and PyBOP (5.4 mg, 11.4 mmol, 1.4 eq) were dissolved in 7.5 % DIPEA/DMF (13 mL) under argon atmosphere and stirred for 16 h. **246** was prepared as described in literature.[263] Water (100 mL) was added, the pH was adjusted to 7 by adding 1 N aqueous HCl and the mixture was extracted four times with EA (100 mL). The organic phases were combined and the solvent was removed via distillation under reduced pressure. The yellowish, viscous residue was treated with TFA containing 2.5 % water and 2.5 % TIS under argon atmosphere for 4 h. The solvent was removed via distillation in vacuum and the crude product was purified by MPLC (RP8, H_2O → MeOH, 0.1 % TFA). The resulting product was dissolved two times in 0.1 N aqueous HCl and another time in water followed by lyophilization to obtain **250** (11.3 mg, 4.9 µmol, 65 %) as hydrochloride salt.

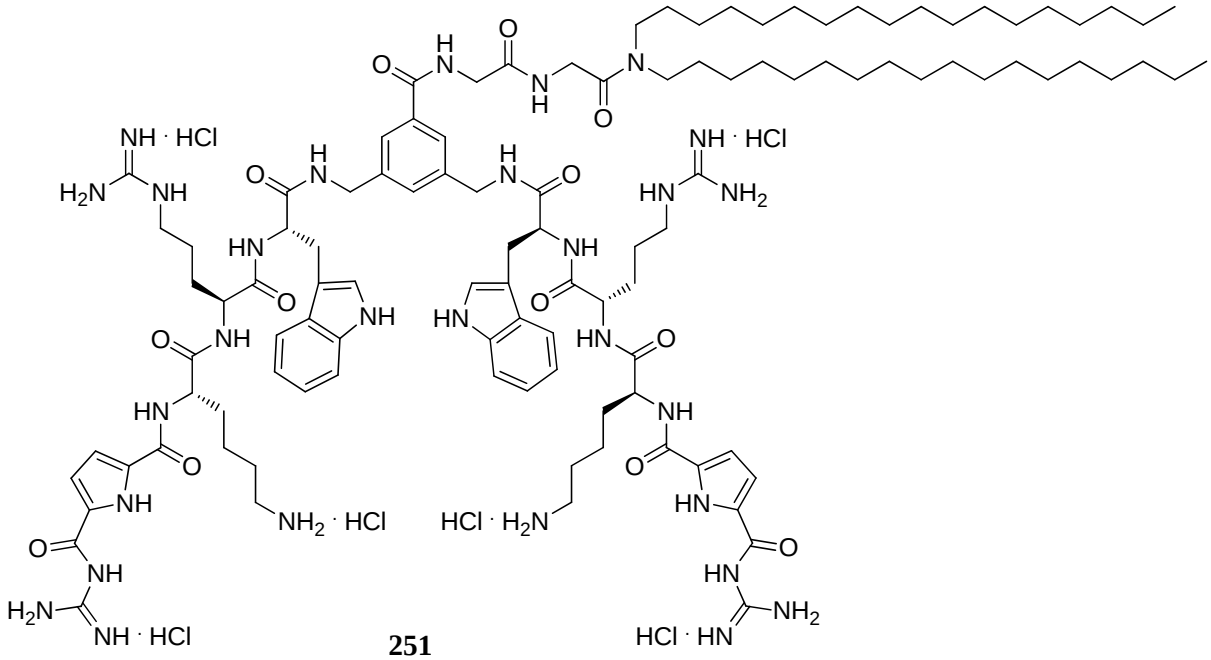
$C_{109}H_{177}Cl_6N_{29}O_{13}$	2314.48 g/mol
Yield	65 %
Mp.	> 180° C (decomposition)
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 1.20-1.30 (m, 8H, 4 × CH_2), 1.33 (s, 18H, 2 × Boc-*t*Bu), 1.35-1.40 (m, 12H, 4 × Pbf-CH_3), 1.41-1.44 (m, 4H, 2 × CH_2), 1.45 (s, 18H, 2 × Boc-*t*Bu), 1.50-1.54 (m, 4H, 2 × CH_2), 1.58 (s, 18H, 2 × Boc-*t*Bu), 1.61-1.78 (m, 4H, 2 × CH_2), 1.95 (s, 6H, 2 × Pbf-CH_3), 2.38 (s, 6H, 2 × Pbf-CH_3), 2.45 (s, 6H, 2 × Pbf-CH_3), 2.78 (dd, $^3J_1$ = 6.22 Hz, $^3J_2$ = 11.91 Hz, 4H, 2 × Lys-CH_2), 2.89 (s, 4H, 2 × Pbf-CH_2), 2.96-3.11 (m, 8H, 2 × Arg-CH_2, 2 × Trp-CH_2), 3.87-3.95 (m, 2H, Gly-CH_2), 4.20 (dd, $^3J_1$ = 7.34 Hz, $^3J_2$ = 13.16 Hz, 4H, 2 × CH_2), 4.40 (s, 2H, 2 × C*H*), 4.29-4.36 (m, 2H, 2 × C*H*), 4.59 (dd, $^3J_1$ = 6.73 Hz, $^3J_2$ = 13.83 Hz, 2H, 2 × C*H*), 6.57 (s, 2H, 2 × N*H*), 6.67 (t, ^3J = 5.34 Hz, 2H, 2 × Trp-CH_{ar}), 6.81 (s, 4H, 4 × Pyr-CH_{ar}), 6.95 (s, 1H, CH_{ar}), 7.11 (s, 2H, 2 × N*H*), 7.20 (t, ^3J = 7.49 Hz, 2H, 2 × Trp-CH_{ar}), 7.25 (t, ^3J = 7.53 Hz, 2H, 2 × Trp-CH_{ar}), 7.49 (s, 2H, 2 × CH_{ar}), 7.62-7.69 (m, 4H, 4 × Trp-CH_{ar}), 8.00 (d, ^3J = 8.47 Hz, 2H, 2 × N*H*), 8.14 (s, 4H, 4 × N*H*), 8.31 (s, 3H, 3 × N*H*), 8.55-8.68 (m, 6H, 6 × N*H*), 9.32 (s, 2H, N*H*), 11.75 (s, 2H, Pyr-N*H*) ppm.

^{13}C-NMR **(125 MHz, DMSO-d_6):** δ = 9.05 (*C*H$_3$), 12.12 (*C*H$_3$), 17.48 (*C*H$_3$), 18.82 (*C*H$_3$), 22.51 (*C*H$_2$), 25.41 (*C*H$_2$), 27.13 (*C*H$_2$), 27.58 (*C*H$_3$), 27.67 (*C*H$_3$), 27.95 (*C*H$_3$), 27.99 (*C*H$_2$), 28.15 (*C*H$_2$), 28.86 (*C*H$_2$), 29.07 (*C*H$_2$), 31.62 (*C*H$_3$), 40.08 (*C*H$_2$), 40.25 (*C*H$_2$), 42.32 (*C*H$_2$), 45.61 (*C*H$_2$), 52.41 (*C*H), 52.55 (*C*H), 52.65 (*C*H), 77.18 (C_q), 83.35 (C_q), 86.13 (C_q), 112.93 (*C*H$_{ar}$), 113.52 (*C*H$_{ar}$), 114.36 (*C*H$_{ar}$), 114.48 (C_q), 116.09 (C_q), 116.29 (C_q), 119.26 (*C*H$_{ar}$), 122.27 (*C*H$_{ar}$), 123.69 (*C*H$_{ar}$), 124.16 (*C*H$_{ar}$), 130.12 (C_q), 131.33 (C_q), 133.68 (C_q), 134.08 (C_q), 134.45 (C_q), 137.14 (C_q), 139.08 (C_q), 148.91 (*C*O), 155.36 (*C*O), 156.08 (*C*O), 157.28 (*C*O), 158.30 (*C*O), 159.53 (*C*O), 170.94 (*C*O), 171.46 (*C*O), 171.64 (*C*O), 174.40 (*C*O) ppm.

HR-MS **(pos. ESI-ToF):** m/z = 1048.6949 $[M - 2\ HCl - 2\ Cl^-]^{2+}$
calculated for $C_{109}H_{173}N_{29}O_{13}^{2+}$: 1048.6893

FT-IR **(ATR):** $\tilde{\nu}$ = 2924 (m), 2853 (m), 2358 (w), 1663 (s), 1549 (m), 1433 (m), 1278 (m), 1179 (s), 1129 (s), 836 (m), 800 (m), 720 (s) cm^{-1}.

Synthesis of the Amphiphile 251 (Trp-Arg-Lys-GCP)



251

Fully protected **248** (30.1 mg, 11.4 µmol, 1 eq), Boc-Gly-dioctadecylamide **246** (10.4 mg, 15.0 mmol, 1.3 eq) and PyBOP (8.9 mg, 17.1 mmol, 1.5 eq) were dissolved in 7.5 % DIPEA/DMF (15 mL) under argon atmosphere and stirred for 16 h. **246** was prepared as described in literature.[263] Water (100 mL) was added, the pH was adjusted to 7 by adding 1 N aqueous HCl and the mixture was extracted four times with EA (100 mL). The organic phases were combined and the solvent was removed via distillation under reduced pressure. The yellowish, viscous residue was treated with TFA containing 2.5 % water and 2.5 % TIS under argon atmosphere for 4 h. The solvent was removed via distillation in vacuum and the crude product was purified by MPLC (RP8, H_2O → MeOH, 0.1 % TFA). The resulting product was dissolved two times in 0.1 N aqueous HCl and another time in water followed by lyophilization to obtain **251** (11.4 mg, 4.9 µmol, 43 %) as hydrochloride salt.

$C_{109}H_{177}Cl_6N_{29}O_{13}$	2314.48 g/mol
Yield	43 %
Mp.	> 180° C (decomposition)
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 0.84 (t, ^3J = 6.75 Hz, 6H, 2 × CH_3), 1.13-1.29 (m, 64H, 16 × C_{18}-CH_2), 1.31-1.73 (m, 20H, 6 × Lys-CH_2, 4 × Arg-CH_2), 2.75 (s, 4H, 2 × Lys-CH_2), 2.98-3.22 (m, 12H, 2 × C_{18}-CH_2, 2 × Trp-CH_2, 2 × Arg-CH_2), 3.89-4.05 (m, 4H, 2 × Gly-CH_2), 4.19 (dd, $^3J_1$ = 4.83 Hz, $^3J_2$ = 15.70 Hz, 2H, 2 × CH), 4.25-4.35 (m, 4H, 2 × CH_2), 4.45 (dd, $^3J_1$ = 8.54 Hz, $^3J_2$ = 12.33 Hz, 2H, 2 × CH), 4.56 (dd, $^3J_1$ = 6.88 Hz, $^3J_2$ = 13.63 Hz, 2H, 2 × CH), 6.90 (s, 2H, 2 × Pyr-CH_{ar}), 6.96 (t, ^3J = 7.41 Hz, 2H, 2 × Trp-CH_{ar}), 7.04 (t, ^3J = 7.20 Hz, 2H, 2 × Trp-CH_{ar}), 7.08-7.15 (m, 3H, CH_{ar}, 2 × Trp-CH_{ar}), 7.16-7.20 (m, 4H, 2 × NH, 2 × Pyr-CH_{ar}), 7.25 (br. s, 8H, 2 × Arg-NH_4^+), 7.31 (d, ^3J =

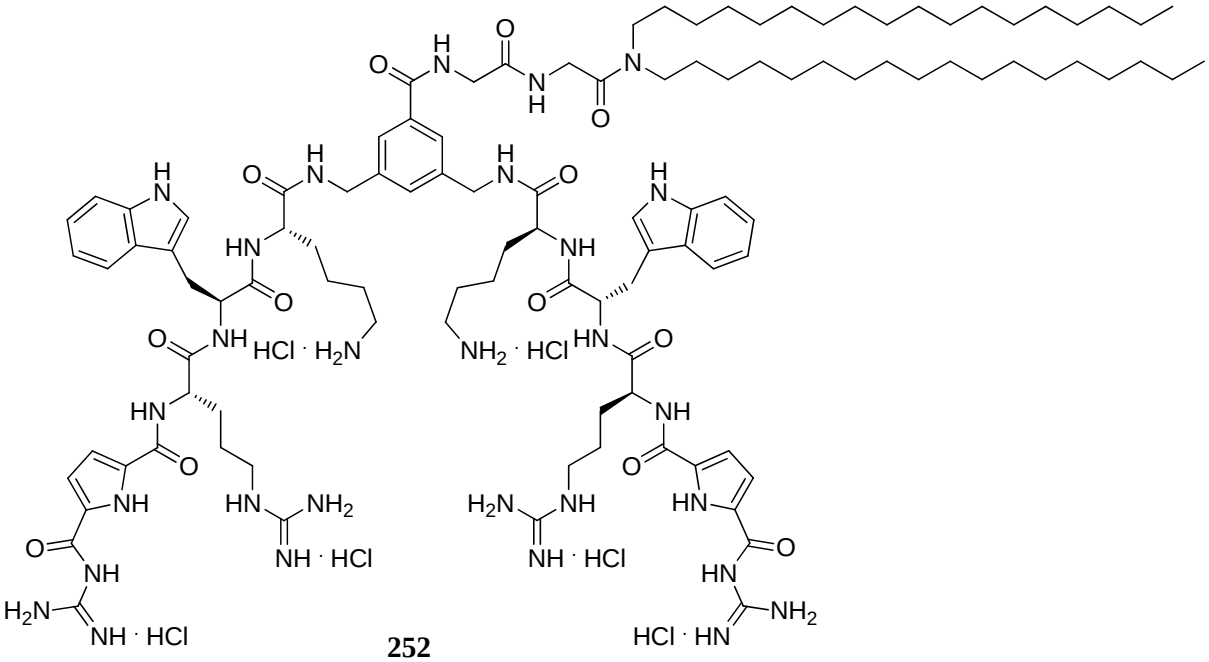
8.12 Hz, 2H, 2 × Trp-C*H*$_{ar}$), 7.53-7.61 (m, 4H, 2 × N*H*, 2 × Trp-C*H*$_{ar}$), 7.66 (s, 2H, 2 × C*H*$_{ar}$), 7.71 (br. s, 6H, 2 × Lys-N*H*$_3^+$), 8.01 (d, ^3J = 6.91 Hz, 2H, 2 × NH), 8.08 (t, ^3J = 4.82 Hz, 1H, Gly-N*H*), 8.24 (d, ^3J = 7.05 Hz, 2H, 2 × N*H*), 8.43-8.63 (m, 10H, 2 × N*H*, 2 × GCP-N*H*$_4^+$), 8.71 (t, ^3J = 5.47 Hz, 1H, Gly-N*H*), 10.77 (s, 2H, Trp-N*H*), 11.52 (s, 2H, GCP-N*H*), 12.52 (s, 2H, Pyr-N*H*) ppm.

^{13}C-NMR **(125 MHz, DMSO-d$_6$):** δ = 13.85 (*C*H$_3$), 22.00 (*C*H$_2$), 22.48 (*C*H$_2$), 24.89 (*C*H$_2$), 26.02 (*C*H$_2$), 26.22 (*C*H$_2$), 26.55 (*C*H$_2$), 27.09 (*C*H$_2$), 28.13 (*C*H$_2$), 28.53 (*C*H$_2$), 28.61 (*C*H$_2$), 28.78 (*C*H$_2$), 28.80 (*C*H$_2$), 28.83 (*C*H$_2$), 28.88 (*C*H$_2$), 28.91 (*C*H$_2$), 28.93(*C*H$_2$), 31.20 (*C*H$_2$), 34.27 (*C*H$_2$), 38.65 (*C*H$_2$), 40.24 (*C*H$_2$), 40.31 (*C*H$_2$), 41.73 (*C*H$_2$), 41.98 (*C*H$_2$), 42.44 (*C*H$_2$), 44.08 (*C*H$_2$), 46.26 (*C*H$_2$), 52.18 (*C*H), 52.53 (*C*H), 53.49 (*C*H), 109.53 (*C*H$_{ar}$), 111.21 (*C*$_q$), 113.51 (*C*H$_{ar}$), 115.89 (*C*H$_{ar}$), 118.14 (*C*$_q$), 118.27 (*C*H$_{ar}$), 118.32 (*C*H$_{ar}$), 120.75 (*C*H$_{ar}$), 123.57 (*C*H$_{ar}$), 124.58 (*C*$_q$), 124.72 (*C*$_q$), 127.20 (*C*H$_{ar}$), 133.86 (*C*$_q$), 135.95 (*C*$_q$), 139.22 (*C*$_q$), 156.59 (*C*O), 157.91 (*C*O), 158.16 (*C*O), 158.40 (*C*O), 158.65 (*C*O), 167.28 (*C*O), 169.19 (*C*O), 171.18 (*C*O), 171.62 (*C*O) ppm.

HR-MS **(pos. ESI-ToF):** m/z = 1048.6911 [M – 2 HCl – 2 Cl$^-$]$^{2+}$
calculated for $C_{109}H_{173}N_{29}O_{13}^{2+}$: 1048.6893

FT-IR **(ATR):** $\tilde{\nu}$ = 2924 (m), 2539 (m), 1639 (s), 1663 (s), 1549 (m), 1276 (m), 1200 (s), 1179 (s), 1132 (s), 837 (m), 800 (m), 721 (s) cm^{-1}.

Synthesis of the Amphiphile 252 (Lys-Trp-Arg-GCP)



252

Fully protected **248** (30.9 mg, 11.5 µmol, 1 eq), Boc-Gly-dioctadecylamide **246** (9.9 mg, 14.3 mmol, 1.2 eq) and PyBOP (7.2 mg, 13.8 mmol, 1.2 eq) were dissolved in 7.5 % DIPEA/DMF (15 mL) under argon atmosphere and stirred for 16 h. **246** was prepared as described in literature.[263] Water (100 mL) was added, the pH was adjusted to 7 by adding 1 N aqueous HCl and the mixture was extracted four times with EA (100 mL). The organic phases were combined and the solvent was removed via distillation under reduced pressure. The yellowish, viscous residue was treated with TFA containing 2.5 % water and 2.5 % TIS under argon atmosphere for 4 h. The solvent was removed via distillation in vacuum and the crude product was purified by MPLC (RP8, H_2O → MeOH, 0.1 % TFA). The resulting product was dissolved two times in 0.1 N aqueous HCl and another time in water followed by lyophilization to obtain **251** (14.2 mg, 6.1 µmol, 53 %) as hydrochloride salt.

$C_{109}H_{177}Cl_6N_{29}O_{13}$	2314.48 g/mol
Yield	53 %
Mp.	> 180° C (decomposition)
^{1}H-NMR	**(500 MHz, DMSO-d_6):** δ = 0.85 (t, ^3J = 6.87 Hz, 6H, 2 × CH_3), 1.17-1.27 (m, 64H, 16 × C_{18}-CH_2), 1.35-1.62 (m, 16H, 3 × Lys-CH_2, 4 × Arg-CH_2), 1.63-1.76 (m, 4H, 2 × Lys-CH_2), 2.70-2.76 (m, 4H, 2 × Lys-CH_2), 2.96-3.20 (m, 12H, 2 × C_{18}-CH_2, 2 × Trp-CH_2, 2 × Arg-CH_2), 3.88-4.05 (m, 4H, 2 × Gly-CH_2), 4.20-4.29 (m, 4H, 2 × CH_2), 4.34 (dd, $^3J_1$ = 5.25 Hz, $^3J_2$ = 15.81 Hz, 2H, 2 × CH), 4.44 (s, 2H, 2 × CH), 4.60 (dd, $^3J_1$ = 7.35 Hz, $^3J_2$ = 12.77 Hz, 2H, 2 × CH), 6.85-6.93 (m, 4H, 2 × Pyr-CH_{ar}, 2 × Trp-CH_{ar}), 7.00 (t, ^3J = 7.51 Hz, 2H, 2 × Trp-CH_{ar}), 7.08-7.24 (m, 7H, CH_{ar}, 2 × Trp-CH_{ar}, 2 × NH, 2 × Pyr-CH_{ar}), 7.26 (br. s, 8H, 2 × Arg-NH_4^+), 7.28 (d, ^3J = 7.95 Hz, 2H, 2 × Trp-CH_{ar}), 7.54 (d, ^3J

= 7.91 Hz, 2 × Trp-CH_{ar}), 7.61 (s, 2H, 2 × N*H,*), 7.67 (s, 2H, 2 × CH_{ar}), 7.73 (br. s, 6H, 2 × Lys-NH_3^+), 8.04 (d, ^3J = 4.85 Hz, 1H, Gly-NH), 8.14 (t, ^3J = 6.80 Hz, 2H, 2 × N*H*), 8.35 (s, 2H, 2 × N*H*), 8.41-8.70 (m, 10H, 2 × N*H*, 2 × GCP-NH_4^+), 8.74 (t, ^3J = 5.64 Hz, 1H, Gly-N*H*), 10.77 (s, 2H, Trp-N*H*), 11.49 (s, 2H, GCP-N*H*), 12.52 (s, 2H, Pyr-N*H*) ppm.

^{13}C-NMR **(125 MHz, DMSO-d_6):** δ = 13.95 (*C*H_3), 22.09 (*C*H_2), 22.26 (*C*H_2), 25.09 (*C*H_2), 26.12 (*C*H_2), 26.32 (*C*H_2), 26.70 (*C*H_2), 27.16 (*C*H_2), 28.21 (*C*H_2), 28.64 (*C*H_2), 28.70 (*C*H_2), 28.89 (*C*H_2), 28.94 (*C*H_2), 28.97 (*C*H_2), 29.00 (*C*H_2), 29.03 (*C*H_2), 31.29 (*C*H_2), 38.72 (*C*H_2), 40.34 (*C*H_2), 40.38 (*C*H_2), 41.89 (*C*H_2), 42.40 (*C*H_2), 42.61 (*C*H_2), 44.18 (*C*H), 52.38 (*C*H), 52.39 (*C*H), 109.74 (*C*H_{ar}), 111.21 (C_q), 113.67 (*C*H_{ar}), 116.02 (*C*H_{ar}), 118.17 (C_q), 118.40 (*C*H_{ar}), 120.81 (*C*H_{ar}), 123.62 (*C*H_{ar}), 135.98 (C_q), 139.51 (C_q), 156.70 (*C*O), 158.20 (*C*O), 158.44 (*C*O), 166.37 (*C*O), 167.47 (*C*O), 171.48 (*C*O), 171.53 (*C*O) ppm.

HR-MS **(pos. ESI-ToF):** m/z = 1048.6923 $[M - 2\ HCl - 2\ Cl^-]^{2+}$
calculated for $C_{109}H_{173}N_{29}O_{13}^{2+}$: 1048.6893

FT-IR **(ATR):** $\tilde{\nu}$ = 3273 (br. m), 2923 (m), 2852 (m), 2359 (w), 1638 (s), 1541 (s), 1434 (m), 1275 (m), 1200 (s), 1132 (s), 837 (m), 800 (m), 721 (s) cm^{-1}.

7.11 BIOLOGICAL TESTS

Preparation of Polynucleotides

$p(dAdT)_2$, $p(dGdC)_2$, pdA × pdT, pdG × pdC, pA × pU, and pG × pC were dissolved in sodium cacodylate buffer (pH 7, 0.05 M) as obtained from the supplier *Aldrich* (10 UN in 250 µL or 5 mg in 350 µL), cooled to 4° C for 20 h, transferred into a glass vial rinsing the original vessel with 25 or 35 µL buffer. ctDNA (10 mg) was transferred to a falcon, dissolved in sodium cacodylate buffer (1.5 ml, pH 7, 0.05 M), cooled to 4° C for 20 h, and sonicated eight times for 4 sec. Then it was filtrated through a 0.45 µm PTFE filter, which was washed with 0.5 ml of buffer. The concentration of the compounds was determined as the concentration of phosphates with the help of the epsilon value at the absorption maxima listed in Table 29. by titrating five times 1 µL into sodium cacodylate buffer (pH 7, 0.05 M) and measuring the resulting UV/Vis absorption. The mammalian expression plasmid pF143 encoding for GFP under the control of a CMV-promoter was obtained as described in literature.[244]

Table 29. *Molar extinction coefficients ε at the maximum of UV absorption (λ_{max}) of various nucleic acids.*

Nucleic acid	pdA × pdT	$p(dAdT)_2$	pdG × pdC	$p(dGdC)_2$	pA × pU	pG × pC	ctDNA
ε [L/mol · cm]	6000	6600	7400	8400	6000	7400	6600
λ_{max} [nm]	260	262	253	254	260	253	260

Cell lines

The human embryonic kidney cell line HEK293T (ATCC-No. CRL-1573) was obtained from the American Type Culture Collection and maintained as recommended in complete Dulbecco`s Modified Eagle Medium supplemented with 10 % fetal bovine serum, 1 % Glutamine and 1 % Antibiotic-Antimycotic (*Invitrogen*) at 37° C in a humidified atmosphere of 5 % CO_2.

Transfection and Microscopy

Per well, 1×10^4 cells were seeded in 96 well cell culture plates (*Greiner bio-one*) in a total medium volume of 200 µl 24 h before transfection. The cationic transfection reagent polyethylenimine (PEI, pH 6.8, *Sigma-Aldrich*) was used at concentration of 0.12 mM in phosphate buffered saline (PBS) (*Invitrogen*) with 2 µg of plasmid DNA per well. Transfection with **196**, **206**, **207**, **233-237** was carried out at concentrations ranging from 0.12 to 0.36 mM with 1, 2, or 4 µg plasmid DNA in a total volume of 30 µl in PBS buffer. Transfection efficiency was analyzed 24 and 48 h after transfection with an inverted fluorescence microscope (Axiovert 200M, *Carl Zeiss*) with a 10X air objective. Images were processed and analyzed using MetaMorph 6.3r6 (*Molecular Devices*) and Adobe Photoshop CS2 (*Adobe Systems*).

Endosomal Escape Assay

To facilitate endosomal escape, cells were incubated with complete medium containing 25 µM chloroquine 30 min before transfection. Transfection was carried out with 0.24 mM transfecting agent and 2 µg plasmid DNA.

Trypan Blue Cell Viability Assay

Cytotoxity of the transfection reagents was evaluated using the negatively charged azo-dye trypan blue (*Applichem*) which only penetrates cells with membrane damage. HEK293T cells were grown and transfected as described above. Equal amounts of the cell solution were mixed with trypan blue (0.2 % in 0.9 % NaCl), incubated for 5 min at RT. Cells were counted under a light microscope in counting cambers (*MEGUMED*).

Plasmid DNA Labeling and Co-localization Experiments

Plasmid DNA was labeled with the PromoFluor-500 Nick Translation Labeling Kit (*PromoKine*) according to the manufacturer's instructions. Briefly, following nick translation of the pF143-GFP vector with PromorFluor-500-dUTPs for 2 h at 15° C, cleanup of the mixture was performed using a DNA purification Kit (*Macharey-Nagel*) as advised. For transfection, HeLa cells were seeded in 8-well plates, grown for 24 h and transfected with 2 µg of Lamp1-RFP Plasmid-DNA mixed with 1 mM PEI. After an additional incubation of 24 h, cells were transfected again with 0.2 µg of labeled DNA mixed with 0.24 mM transfecting agent or with the labeled DNA alone and examined under a confocal fluorescence microscope (*Leica*) 2 and 4 h after transfection.

APPENDICES

A: ZUSAMMENFASSUNG UND AUSBLICK

Zusammenfassung

Im ersten Teil der Doktorarbeit wurden zusammen mit unseren Kooperationspartnern, der Arbeitsgruppe von *Prof. Schlücker*, Resonanz-Raman (RR) Bindungsstudien bezüglich des Erkennungsvorganges zwischen einem kleinen molekularem Rezeptor und tetrapeptidischen Substraten durchgeführt. Dafür wurden der in Abbildung 1 abgebildete Rezeptor **13** und die vier Liganden **161**-**164** mittels Fmoc-Festphasensynthese in guten bis exzellenten Ausbeuten hergestellt. Mittels RR und der mathematischen Methode der nicht-negativen Matrixfaktorisierung konnten die Bindungseigenschaften aller vier Kombinationen erfolgreich untersucht werden. So wurde die in der Literatur bereits bekannte Reihenfolge der Affinität des Rezeptors gegenüber den Substraten **161** > **162** korrekt abgebildet. Auch wenn diese Ergebnisse in den momentan durchgeführten weiterführenden Arbeiten bestätigt werden müssen, können diese ersten Resultate bereits als Beweis für die allgemeine Tauglichkeit dieser Resonanz-Raman-basierten Methode zur Analyse von supramolekularen Komplexbildungen angesehen werden. Es ist nunmehr nicht nur möglich, einen derartigen Bindungsvorgang rein qualitativ zu beschreiben, vielmehr konnte gezeigt werden, dass mit Hilfe dieser Methodik ebenfalls eine quantitative Beschreibung supramolekularer Vorgänge mittels der Bestimmung der Bindungskonstanten möglich ist.

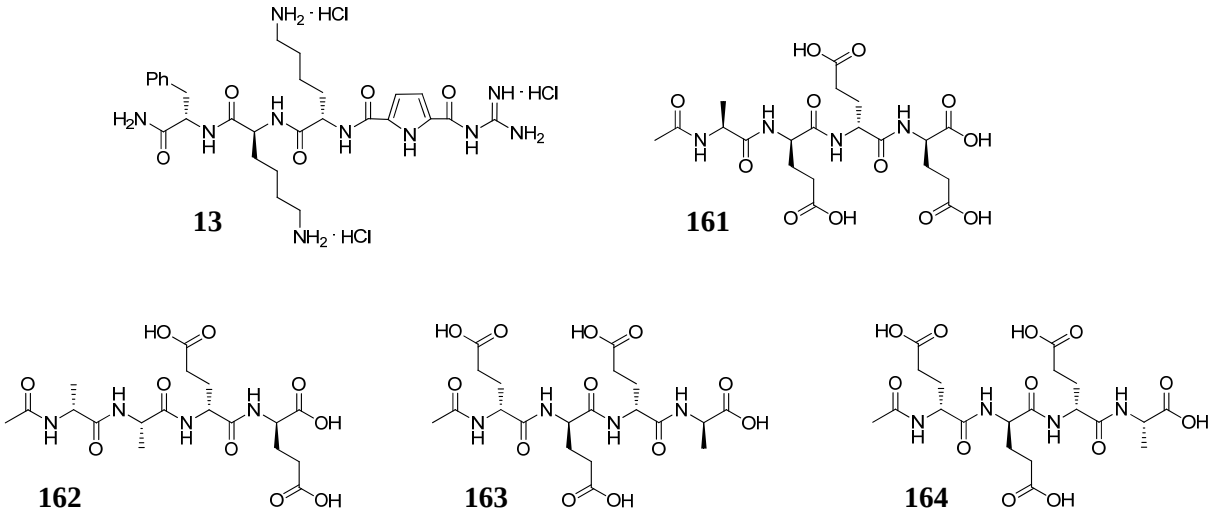


Abbildung 1. *Rezeptor **13** und vier tetrapeptidische Substrate **(161-164)** für Resonanz-Raman Bindungsstudien.*

Im zweiten Raman Projekt, das ebenfalls zusammen mit unseren Kooperationspartnern bearbeitet wurde, konnten wir erfolgreich neuartige, bifunktionelle polymere Träger herstellen, die die Vorteile der Festphasensynthese mit der Möglichkeit der schnellen Charakterisierung von festphasengebundenen Substanzen mittels oberflächenverstärkter

Raman Spektroskopie verbinden. Dafür wurden, wie in Abbildung 2 gezeigt, Gold-Nanopartikel mit einer extrem dünnen Glasschicht versehen und auf einen polymeren Träger derartig aufgebracht, dass dessen Oberfläche vollständig mit den Nanopartikeln bedeckt war. Die oberflächengebundenen Partikel wurden dann mit einer Aminfunktion derivatisiert und somit für die Peptidsynthese mittels standard Fmoc-Strategie zugänglich gemacht. Somit konnten die drei Aminosäuren Gly, Phe, und Ser (**166-168**) sowie das Dipeptid Gly-Phe (**171**) auf den Träger angebracht werden. Mittels mikroskopischer Techniken konnte gezeigt werden, dass die bifunktionellen Einheiten auch nach der Synthese und somit trotz der Behandlung mit Base (20 % Piperidin), Säure (33 % TFA) und organischen Lösemitteln (DMF, DCM) vollkommen intakt waren, und aus diesem Grund hervorragend für eine Anwendung in der Festphasensynthese geeignet sind. Die vier festphasengebundenen Substanzen konnten dann mittels Raman Experimenten schnell und zuverlässig zweifelsfrei voneinander unterschieden werden. Somit erfüllen die kombinierten Festphasen/Raman Einheiten alle notwendigen Voraussetzungen für eine Vielzahl von Anwendungen, z.B. für die Charakterisierung von Zwischenprodukten einer Festphasensynthese oder der Untersuchung von molekularen Erkennungsvorgängen. Zusammenfassend stellt dieses neuartige mit Nanopartikeln versehene bifunktionelle Harz also eine exzellente Erweiterung der Felder der supramolekularen Chemie und der Festphasensynthese dar, die eine direkte und markerfreie Analyse von festphasengebundenen Substanzen ermöglicht.

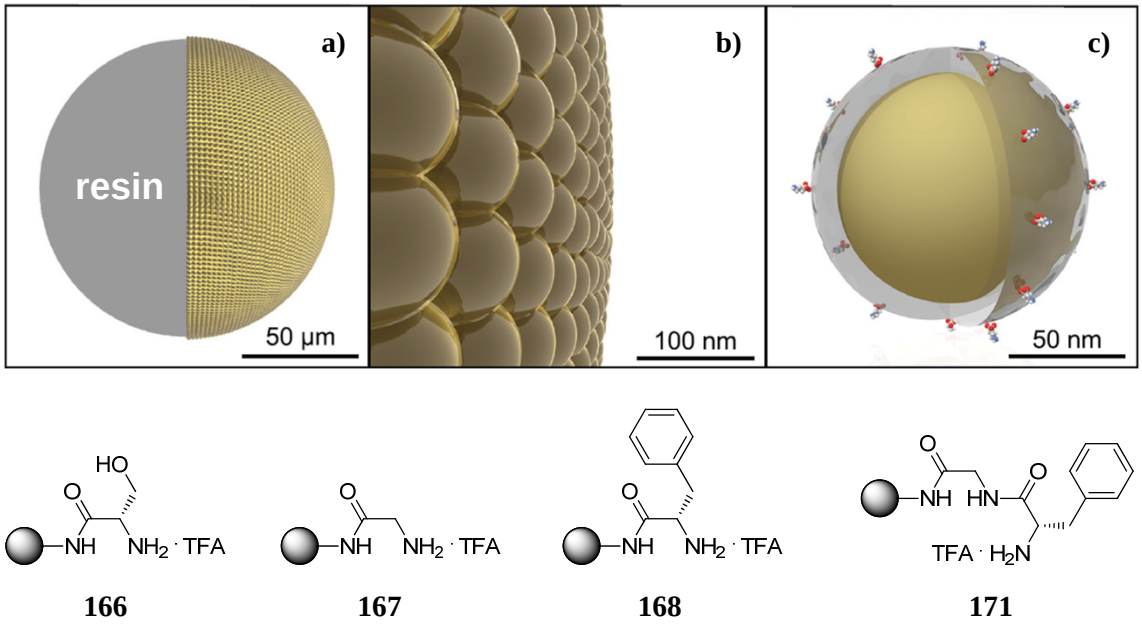


Abbildung 2. *Design von bifunktionellen Einheiten zur Kombination von Festphasensynthese mit oberflächenverstärkter Raman Spektroskopie. a) Glasverkapselte, 80 nm große Goldnanopartikel auf einem 130 µm großen TentaGel Kügelchen; b) Zoom auf die Oberfläche; c) Aminfunktionalisierung der Glasverkapselung mittels eines Aminosilans (Bild aus: Doktorarbeit Gellner).*[191] *Die drei Aminosäuren* ***166-168*** *und das Dipeptid* ***171*** *wurden mittels Festphasensynthese hergestellt und konnten mit Hilfe von verstärkter Raman Spektren direkt am Harz charakterisiert und voneinander unterschieden werden.*

Ein weiteres großes Arbeitsfeld der Dissertation befasste sich mit der molekularen Erkennung von biologisch relevanten Substraten. Dazu wurden zunächst zwei unsymmetrische (**187**, **188**) und ein symmetrischer Tweezer Rezeptor **(196**, Abbildung 3)

mittels mikrowellengestützter Festphasensynthese hergestellt. Im Vergleich zu früheren Arbeiten im Rahmen meiner Diplomarbeit konnte die Synthese für beide dafür benötigte Template, dem orthogonal geschützten Derivat der Benzoesäure **181** für die Darstellung von unsymmetrischen Rezeptoren, sowie dem analogen **193** für die Synthese von symmetrischen Tweezern, verbessert werden. So konnte die Gesamtausbeute über sieben (**181**) bzw. sechs (**193**) Reaktionsschritte von 24 auf 39 % und respektive von 29 auf 42 % erhöht werden. Des Weiteren konnten sowohl die Ausbeute als auch die Reinheit der Rezeptoren durch die Verwendung einer Mikrowelle signifikant erhöht werden. Mit den eigens für die Synthese optimierten Mikrowellenbedingungen (max. 60° C, 20 min, 20 W) war es außerdem möglich z.B. den unsymmetrischen Rezeptor **187** innerhalb von nur zwei Tagen anstatt von zwei Wochen, die für die konventionelle Synthese notwendig waren, herzustellen. Die so synthetisierte zweite Generation von zweiarmigen Rezeptoren verfügte im Vergleich zu ähnlichen während der Diplomarbeit synthetisierten Systemen über verbesserte physikochemische Eigenschaften für die Komplexierung von Substraten in Wasser bei physiologischem pH. Vor allem die Verringerung der Azidität der Guanidiniogruppe um ca. eine Größenordnung garantiert einen höheren Protonierungsgrad der GCP Oxoanionen Bindungsstelle bei pH 7, was zu stark verbesserten Bindungseigenschaften hinsichtlich der Interaktion mit anionischen Substraten führte.

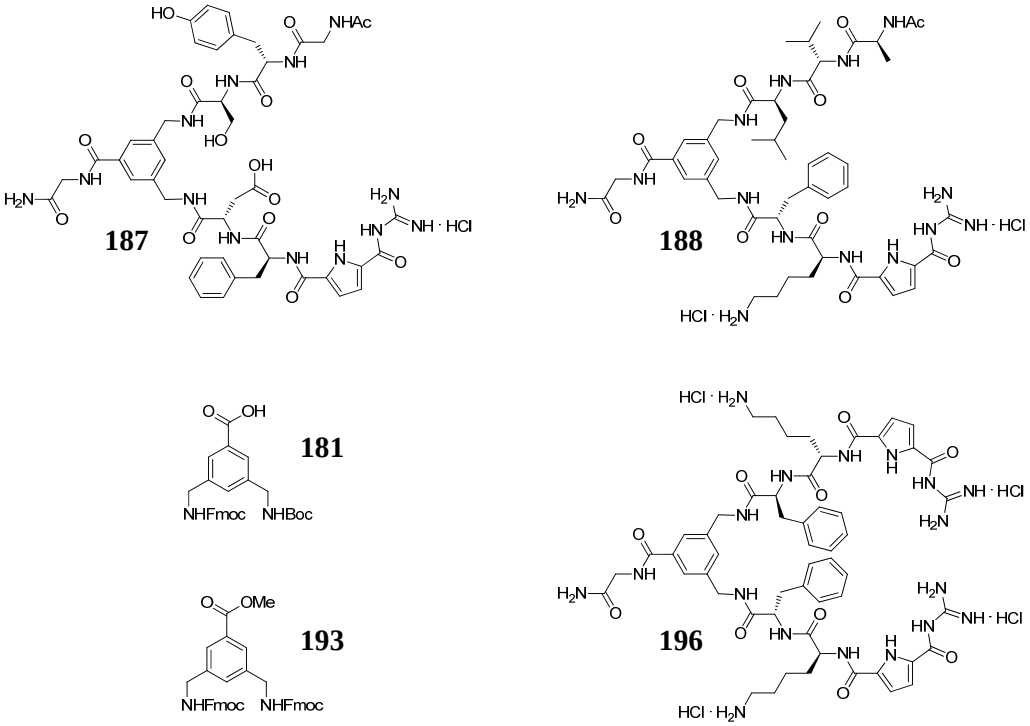


Abbildung 3. *Orthogonal geschütztes Templat **181** für die mikrowellengestützte Festphasensynthese der unsymmetrischen Tweezer Rezeptoren **187** und **188**. Das entsprechende symmetrische Gegenstück **193** wurde für die Darstellung des divalenten Rezeptors **196** mit zwei gleichartigen Seitenketten verwendet.*

Nach der erfolgreichen Synthese konnte mittels UV/Vis-Titrationen ermittelt werden, dass die beiden unsymmetrischen Tweezer Rezeptoren **187** und **188** Bindungskonstanten von bis zu 6×10^3 M^{-1} in gepuffertem Wasser für das polare, peptidische Substrat N-Ac-Lys-D-Ala-D-Ala-OH, welches eine entscheidende Rolle in der Biosynthese der Zellwand von *Gram*

positiven Bakterien spielt, aufweisen. Es konnte demonstriert werden, dass die Bindungskonstante mit sinkendem pH (7 → 6 → 5) steigt, und dass dieser Zusammenhang auf den Protonierungsgrad der GCP Gruppen zurückzuführen ist. Die Bindungsaffinität beruht hauptsächlich auf der Wechselwirkung zwischen dem GCP Bindungsmotiv des Rezeptors mit der Carboxylatgruppe des Tripeptids. Außerdem konnte gezeigt werden, dass **187** aufgrund des geringeren pK_A-Wertes seiner GCP Gruppe gepaart mit zusätzlichen Wechselwirkungen der Seitenketten des Rezeptors mit dem Substrat eine höhere Affinität zu dem getesteten Tripeptid aufweist als der analoge Tweezer **188**.

Bezüglich des symmetrischen Rezeptors **196** konnte mittels übereinstimmender UV/Vis- und fluoreszenzspektroskopischer Methoden gezeigt werden, dass dieser Tweezer in gepuffertem Wasser bei neutralem pH stabile 1:1 Komplexe mit Nukleotiden bildet, die mit guten Bindungskonstanten von bis zu 9×10^4 M^{-1} gebunden werden. Nukleotide werden dabei um annähernd eine Größenordnung stärker gebunden als Phosphate oder Diphosphate. Die bemerkenswerteste Eigenschaft des Rezeptors liegt jedoch darin, dass er Adenosinmonophosphat über das entsprechende Diphosphat und dieses wiederum über das Triphosphat präferiert. Eine derartige Bevorzugung ist bisher beisspiellos in der Literatur und konnte mit Hilfe von NMR Spektroskopie und Kraftfeldrechnungen erklärt werden. Der Rezeptor bildet mit seinen beiden GCP Einheiten eine doppelte Salzbrücke zu dem am höchsten geladenen α-Phosphat des jeweiligen Nukleotids aus, indem sich die beiden Pyrrole perpendikulär zueinander ausrichten und gleichzeitig das tetraedrische Phosphatanion binden. Gleichzeitig findet eine π-π-Wechselwirkung zwischen dem aromatischen Templat des Rezeptors und der Nukleobase statt, wodurch das Nukleotid in einer hochgeordneten Geometrie komplexiert wird. **196** umschließt dabei das Substrat als eine halbgeschlossene Schale. Die in Abbildung 4 illustrierte lösemittelzugängliche Oberfläche dieses Wirt-Gast-Systems verdeutlicht, dass der Nukleotid-Phosphatrest (grün) mit größer werdender Länge immer mehr aus der Halbschale des Rezeptors (gelb) in das umgebende Lösemittel herausragt und somit für kompetitive Einflüsse desselben exponiert wird, wodurch dieses immer mehr um die Komplexierung des Substrates konkurriert. Die Konsequenz davon ist, dass die Bindungskonstante in der Reihenfolge AMP > ADP > ATP sinkt.

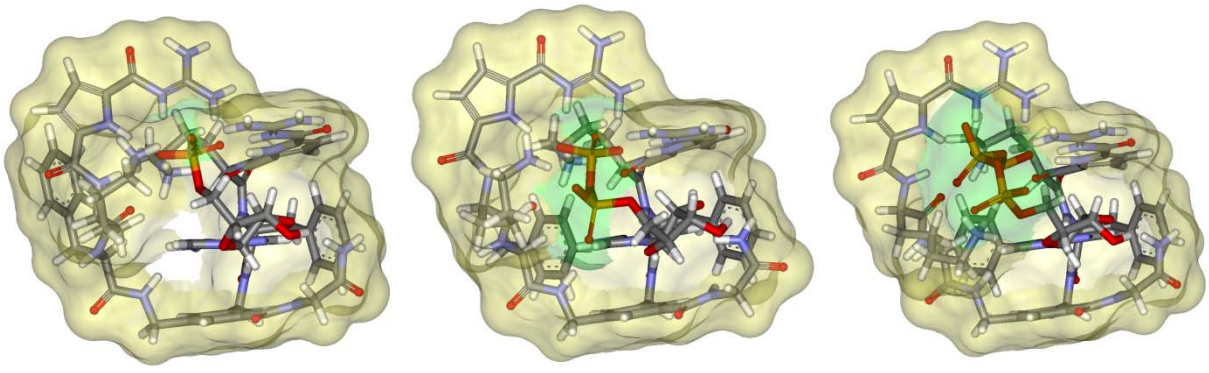
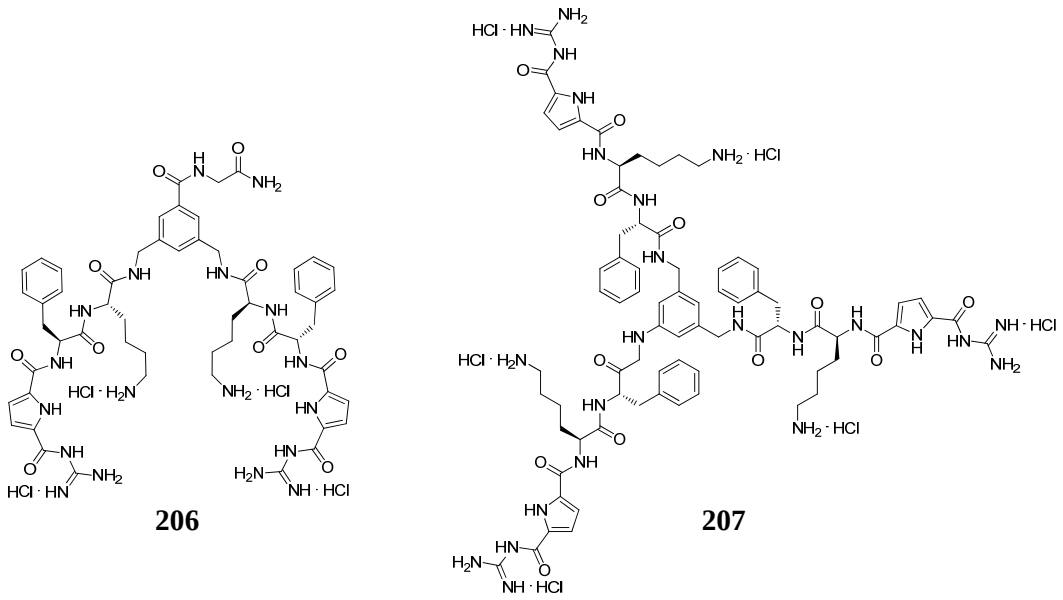


Abbildung 4. *Darstellung der lösemittelzugänglichen Oberfläche von berechneten 1:1 Komplexen des Rezeptors* ***196*** *(gelb) und der Phosphatreste der Nukleotidsubstrate (grün). Von links nach rechts: AMP, ADP und ATP.*

Mit Hilfe dieser Experimente konnte eindrucksvoll gezeigt werden, dass die spezifische Interaktion zwischen dem symmetrischen Tweezer Rezeptor **196** mit dem Phosphat- und

Nukleotidanionen ein exzellentes Beispiel für die Notwendigkeit dafür ist, mehrere nichtkovalente Wechselwirkungen in einer geeigneten Orientierung miteinander zu kombinieren, um die Ausbildung von spezifischen Wirt-Gast-Komplexen und somit eine selektive molekulare Erkennung eines Substrates zu ermöglichen.

Für das nächste bearbeitete Projekt wurden zwei weitere artifizielle Systeme synthetisiert. Wie in Abbildung 5 gezeigt, weist der ebenfalls mittels mikrowellengestützter Festphasensynthese dargestellte symmetrische Tweezer **206** im Vergleich mit dem ursprünglichen Analogon **196** (Phe-Lys-GCP) eine invertierte Aminosäuresequenz auf (Lys-Phe-GCP). Das mittels einer konvergenten Synthese hergestellte trivalente Analogon **207** verfügt wiederum über die ursprüngliche Aminosäuresequenz (Phe-Lys-GCP). Zu dessen Darstellung wurde zunächst die Seitenkette mittels Fmoc-Festphasensynthese derart synthetisiert, dass alle Schutzgruppen der Aminosäuren intakt waren um die Arme dann in Lösung über den freien C-Terminus an das Triamintemplat zu kuppeln und anschließend zu entschützen.



206 **207**

Abbildung 5. *Divalenter Ligand **206** mit invertierter Aminosäuresequenz Lys-Phe-GCP und trivalenter Ligand **207** mit ursprünglicher Sequenz Phe-Lys-GCP.*

Die drei Moleküle **196**, **206** und **207** wurden dann mittels thermaler Denaturierungsexperimente, Verdrängungsassays, ITC und CD Spektroskopie bezüglich ihrer Bindungseigenschaften gegenüber verschiedenen synthetischen und generischen Nukleinsäuren getestet. Die zweiarmigen Polynukleotid-Liganden **196** und **206** bilden äußerst stabile Komplexe mit Bindungskonstanten von $K > 10^6\ M^{-1}$. Der trivalente Ligand **207** ist sogar in der Lage um eine weitere Größenordnung stärker ($K > 10^7\ M^{-1}$) an die Biomoleküle zu binden. Der Bindungsmodus setzt sich bei diesen Vorgängen aus elektrostatischer Wechselwirkung mit dem negativ geladenen Ribophosphat Rückgrade der Nukleinsäuren mit einer Komplexierung in der kleinen Furche für B-Typ DNA bzw. in der großen Furche für A-Typ DNA und RNA zusammen. Mit Hilfe von CD-Spektroskopie konnte gezeigt werden, dass die drei Nukleinsäureliganden mittels zweier gestapelter Pyrroleinheiten binden und dass sie so in der Lage sind zwischen verschiedenen Sekundärstrukturen und Sequenzen von Polynukleotiden zu unterscheiden und Exzitonen CD-Signal mit unterschiedlichem

Vorzeichen hervorrufen. Eine derartige selektive Erkennung von Polynucleotiden durch nicht-kovalent gebundene kleine Moleküle ist bisher nur sehr selten in der Literatur beschrieben und ermöglicht eine potentielle Anwendung dieser Systeme als DNA/RNA-Sensoren.

Des Weiteren konnte mittels AFM und DLS Experimenten gezeigt werden, dass **196** und **207** in der Lage sind generische DNA in einer konzentrationsabhängigen Art und Weise zu dicht gepackten, sphärischen Aggregaten zu komprimieren. Zusammen mit unseren Kooperationspartnern, der Arbeitsgruppe von *Prof. Knauer*, konnten wir zeigen, dass diese multimolekularen Komplexe endozytotisch von humanen Zellen aufgenommen werden. Jedoch war nur der trivalente Ligand in der Lage aus den endosomalen Vesikeln zu entkommen und somit Zellen mit genetischem Material zu transfizieren. Wie in Abbildung 6 schematisch gezeigt, beruht diese Differenzierung auf der überlegenen Pufferkapazität von **207**, die in etwa doppelt so hoch ist wie beim analogen zweiarmigen Äquivalent **196**. Die Guanidino-Gruppe wird aufgrund der Einströmung von Protonen mittels der spezifischen ATPase in das Endosom sukzessive protoniert und agiert somit als „Protonenschwamm“, was letztendlich zum Bruch der endosomalen Vesikelwand aufgrund des erhöhten osmotischen Drucks führt. In vergleichenden Experimenten konnte bewiesen werden, dass der dreiarmige Ligand **207** effizienter als der standardmäßig eingesetzte nichtvirale Vektor PEI bei gleichzeitig verringerter Zytotoxizität ist. Im Vergleich zu anderen zellpenetrierenden Peptiden, die normalerweise von DNA-bindenden Peptiden abgeleitet und aus 10-30 Aminosäuren aufgebaut sind, ist **207** außerordentlich klein und weist weniger Ladungen auf und verfügt somit über großes Potential für die Anwendung in der Zellbiologie.

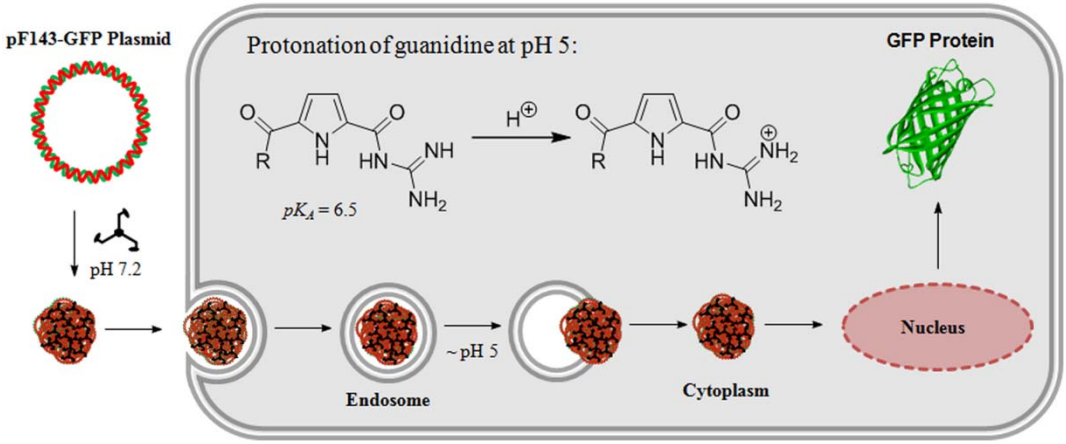
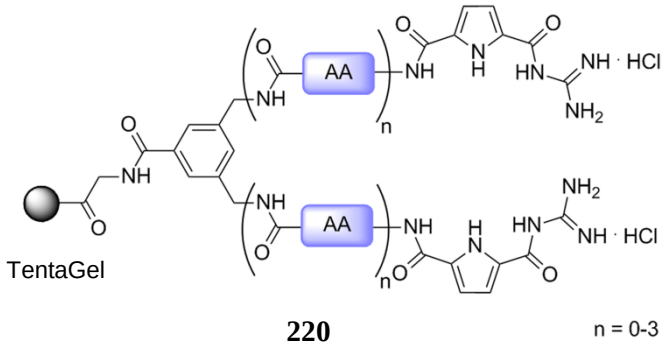


***Abbildung 6. 207**-vermittelter zellulärer Aufnahmemechanismus. Plasmid DNA wird zu dichten Aggregaten kondensiert, welche mittels Endozytose in die Zelle aufgenommen werden. Die GCP Gruppe fungiert dabei als Protonenschwamm, was zum Bruch des endosomalen Vesikels führt. Die genetische Information wird dann in den Zellkern transportiert um letztendlich eine Expression des GFP Proteins auszulösen.*

Um die Bindungseigenschaften der divalenten Liganden zu verbessern wurde im nächsten Schritte die kombinatorische DNA-Liganden Bibliothek **220**, die in Abbildung 7 schematisch dargestellt ist, mittels Split & Mix Synthese hergestellt. Die Mitglieder dieser relativ kleinen, jedoch hochgradig fokussierten Bibliothek verfügen über drei kombinatorische Positionen, die jeweils mit sechs unterschiedlichen Aminosäuren bestückt wurden. Als Bausteine wurden dabei kationissche (Lys, Arg), aromatische (Phe, Trp, Tyr) oder als negative Kontrolle

anionische (Glu) Aminosäuren verwendet. Zusätzlich wurde die Länge der beiden Arme der Liganden variabel gestaltet, indem die Anzahl der eingebauten Aminosäuren zwischen 0 und 3 variiert wurde. Dementsprechend enthielt die Substanzbibliothek insgesamt 259 Mitglieder ($6^3 + 6^2 + 6 + 1$). Um die Bibliothek direkt am Harz zu screenen wurde zunächst ein Assay entwickelt, welcher die Bestimmung der Affinitäten der Liganden zu den zwei getesteten Nukleinsäuren ermöglichte. Dabei wurde zum Einen die synthetische B-DNA $p(dAdT)_2$ mit ihrer sehr ausgeprägten, gut definierten kleinen Furche und zum Anderen die pF143-GFP Plasmid DNA, welche für die oben erwähnten Transfektionsexperimente genutzt wurde, verwendet. Im Gegensatz zu großen, randomisierten Bibliotheken, bei denen lediglich Hit-Strukturen ermittelt werden können, konnten mit Hilfe der IRORI Radiofrequenztechnologie quantitative Daten für alle Bibliotheksmitglieder ermittelt werden.



220

AA = Phe, Tyr, Trp, Lys, Arg, Glu

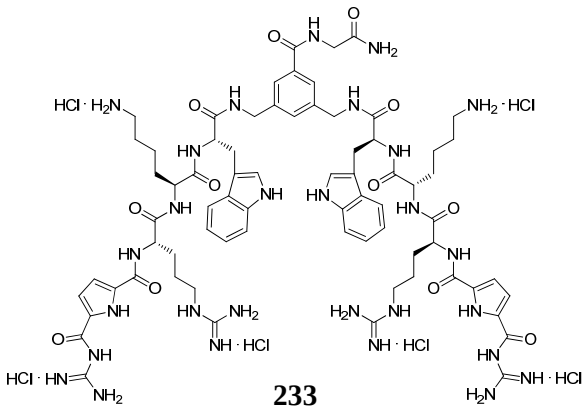
216	= 6 × 6 × 6	(n = 3)
36	= 6 × 6	(n = 2)
6		(n = 1)
1		(n = 0)
259 Liganden		

Abbildung 7. *Die Bibliothek **220** besteht aus 259 symmetrischen, divalenten DNA Liganden, die über jeweils eine GCP Kopfgruppe in den beiden Seitenketten verfügen. Diese bestehen aus entweder drei, zwei, einer oder keiner kombinatorischen Position (Phe, Tyr, Trp, Lys, Arg und Glu).*

Mit Hilfe des neu entwickelten Screenings konnten für $p(dAdT)_2$ hochaffine Sequenzen identifiziert werden, die die ursprünglichen zweiarmigen Liganden klar übertreffen. Wie erwartet führten zusätzliche Ladungswechselwirkungen von höher positiv geladenen Liganden mit der negativ geladenen DNA zu verbesserten Bindungseigenschaften. Wesentlich aufschlussreicher war jedoch, dass auch die Aminosäuresequenz eine entscheidende Rolle dabei spielte. So waren Sequenzen des Typs AA^1-kationisch-kationisch-GCP besonders effizient das alternierende Heteropolymer $p(dAdT)_2$ zu binden, wohingegen solche des Typs AA^1-aromatisch-AA^2-GCP weniger geeignet waren. Dies ist wohl darauf zurückzuführen, dass sperrige aromatische Seitenketten relativ schlecht in die schmale kleine Furche dieser DNA eingefügt werden können. Bezüglich der Rolle der Länge der zwei Arme der Liganden konnte ein immens positiver Effekt für den Übergang von einer zu zwei Aminosäuren pro Arm beobachtet werden. Offensichtlich ist ein gewisser Grad an Flexibilität notwendig um eine hohe Affinität der Liganden zu den getesteten Nukleinsäuren zu gewährleisten.

Um die Resultate des Screenings an fester Phase mittels Experimente in Lösung zu verifizieren wurden die fünf DNA-Liganden **233-237** synthetisiert, die einen Querschnitt über alle Affinitätsbereiche der Bibliothek repräsentieren. Der Ligand **233** ist exemplarisch in Abbildung 8 gezeigt. Die Wiederholung des Assays in Lösung mit diesen Substanzen lieferte Ergebnisse, die in exzellenter Übereinstimmung mit denen der festen Phase waren. Außerdem konnte mittels ITC Messungen gezeigt werden, dass extrem hohe Bindungskonstanten von bis

zu $K > 10^9$ M^{-1} erzielt werden können. Die Affinität der zweiarmigen Liganden konnte somit mit Hilfe der kombinatorischen Bibliothek um hervorragende drei Größenordnungen gesteigert werden.



233

Abbildung 8. *Fünf Bibliotheksmitglieder wurden für eine Charakterisierung in Lösung synthetisiert. Neben dem abgebildeten Ligand **233** der Sequenz Trp-Lys-Arg-GCP weisen die vier weiteren Moleküle die Sequenzen Trp-Arg-Lys-GCP (**234**), Lys-Lys-Arg-GCP (**235**), Lys-Trp-Arg-GCP (**236**) und Glu-Trp-Arg-GCP (**237**) auf.*

Ein analoges Festphasenscreening für die Plasmid DNA sowie dessen Verifizierung in Lösung konnte aufzeigen, dass im Allgemeinen keine Korrelation zwischen effizienter Komplexierung von $p(dAdT)_2$ und hoher Affinität zum Plasmid besteht. Es konnte jedoch gezeigt werden, dass Liganden, die die drei Aminosäuren Trp, Lys und Arg enthalten sehr effizient darin sind beide DNA Typen zu binden. Die höchsten so erzielten Bindungskonstanten bezüglich des Plasmids waren um einen Faktor 1000 höher als diejenige der urprünglichen divalenten Liganden ($K > 10^9$ M^{-1}). Somit konnte mittels der kombinatorischen Bibliothek **220** nicht nur die Affinität der divalenten Liganden optimiert werden, sondern die Bibliothek ermöglichte außerdem die Identifizierung von Polynukleotid-spezifischen Liganden. In Zusammenarbeit mit unseren Kooperationspartnern konnte schließlich eine direkte Korrelation zwischen erhöhten Bindungskonstanten zum Plasmid mit einer erhöhten Transfektionseffizienz ermittelt werden. Auch wenn die so abgeleiteten neuen zweiarmigen Gentransportmoleküle aufgrund fehlender Pufferkapazität nicht die Effizienz des dreiarmigen Analogons **207** erreichten, ist dieser unerwartete Zusammenhang von hohem Nutzen für die Entwicklung zukünftiger nichtviraler Vektoren.

Um die verbesserten Transfektionseigenschaften der neu entwickelten Tweezer weiter auszuschöpfen, wurden drei der hochaffinen Sequenzen mit Oktadecylketten derivatisiert um so die Substanzen **250**-**252** zu erhalten. **250** ist als Vertreter dieser amphiphilen Gentransporter in Abbildung 9 exemplarisch aufgeführt. Des Weiteren wurde zu Vergleichszwecken ein amphiphiles Analogon (**243**) des ursprünglichen Tweezers **196** synthetisiert. Die Darstellung dieser vier Substanzen erfolgte mittels einer konvergenten Syntheseroute: Zunächst wurden die vollständig geschützten DNA-Liganden mittels mikrowellengestützter Festphasensynthese hergestellt um sie dann über ihre freien C-Termini mit C_{18}-derivatisiertem Glycin zu kuppeln und anschließend zu entschützen. Die folgenden Zelltransfektionsexperimente bestätigten den Zusammenhang zwischen einer zum Plasmid hochaffinen Kopfgruppe am Amphiphil mit der einhergehenden Effizienz als Gentransporter. **251** konnte als äußerst effizienter Transfektionsvektor identifiziert werden, der sogar den

Goldstandard PEI übertrifft. Zudem konnte so die für eine effiziente Transfektion minimal benötigte Vektorkonzentration um die Hälfte reduziert werden.

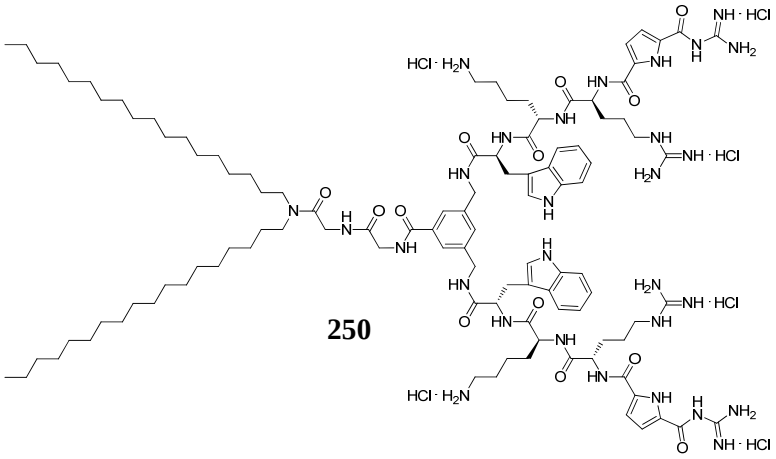


Abbildung 9. *Vier amphiphile Gentransporter wurden für die Zelltransfektion synthetisiert. Exemplarisch ist derjenige mit der Sequenz Trp-Lys-Arg-GCP (**250**) abgebildet. Die drei weiteren Systeme verfügen über die Sequenzen Trp-Arg-Lys-GCP (**251**), Lys-Trp-Arg-GCP (**252**) und Phe-Lys-GCP (**243**).*

Zusammenfassend konnten im Rahmen dieser Dissertation neue Erkenntnisse über molekulare Erkennungsvorgänge von biologisch relevanten Substraten wie Peptide, Nukleotide und Nukleinsäuren gewonnen werden. Diese Resultate tragen nicht nur dazu bei unser generelles Verständnis über die komplexen Prozesse, die in biologischen Systemen stattfinden, zu verbessern, sie konnten auch direkt für eine Anwendung in der Zellbiologie mittels der Entwicklung eines neuen Ansatzes für die Identifizierung und Darstellung neuartiger, hocheffizienter nichtviraler Vektoren verwendet werden.

Ausblick

Nachdem gezeigt werden konnte, dass es prinzipiell möglich ist, Bindungskonstanten von molekularen Erkennungsereignissen mittels Resonanz-Raman Experimenten zu bestimmen, müssen diese bisher vorläufigen Ergebnisse verifiziert werden und auf ihre Reproduzierbarkeit hin getestet werden. Bezüglich der bifunktionellen Festphasensynthese/Raman-Partikel könnte der nächste Schritt sein zu testen ob das GCP Oxoanionen Bindungsmotif durch oberflächenverstärkte Raman Spektroskopie abgebildet werden kann, wenn es wie in Abbildung 10 gezeigt auf dem modifiziertem Harz gebunden ist. Wenn dem so ist, kann im Folgenden bestimmt werden, ob es möglich ist, ein molekulares Erkennungsereignis wie z.B. die Komplexierung von Acetat (**254**) direkt an der festen Phase zu visualisieren. Durch den Vergleich der Raman-Spekren nach der Zugabe von unterschiedlichen Substraten mit den dazugehörigen experimentell bestimmten Bindungskonstanten, kann des Weiteren getestet werden ob diese Methode dafür geeignet ist unterschiedliche Bindungsaffinitäten abzubilden und somit ein rasches Substratscreening zu ermöglichen. Somit könnten die modifizierten festen Trägerpartikel für die Darstellung und das markerfreie Screening von Rezeptor- oder Substratbibliotheken verwendet werden. Schließlich könnten in zukünftigen Experimenten auch komplexere Wirt-Gast-Systeme mit z.B. ein oder zwei zusätzlichen Aminosäuren untersucht werden um die Grenzen und den potentiellen Anwendungsbereich dieser neuen Technik auszuloten.

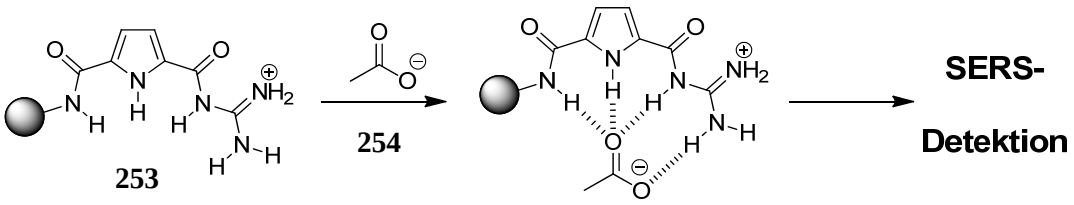


Abbildung 10. *Indem das GCP Oxoanionen Bindungsmotif auf die bifunktionellen Harzkügelchen angebracht wird (**253**) könnten Bindungsvorgänge von kleinen Carboxylaten wie z.B. Acetat (**254**) direkt an der festen Phase mittels oberflächenverstärkter Raman-Spektroskopie sichtbar gemacht werden.*

Hinsichtlich der kombinatorischen Bibliothek **220** könnte als nächstes ein Screening mit Nukleotiden durchgeführt werden. Da der entsprechende ursprüngliche Tweezer **196** schon in der Lage dazu war selektiv Nukleotide mit guten Bindungskonstanten zu erkennen, sollte es möglich sein noch effizientere Rezeptoren mit Hilfe der Bibliothek zu ermitteln. Alternativ könnten weitere Arten von Nukleinsäuren wie z.B. siRNAs einem Screening unterzogen werden, um hochaffine Sequenzen für diesen vollkommen andersartigen Typ von Oligonukleotiden zu identifizieren. Die so ermittelten Sequenzen könnten dann auf ihre Effizienz hin getestet werden, siRNAs in Zellen zu schleusen und somit den RNAi Mechanismus auszulösen.[265]

Die bereits identifizierten hochaffinen Aminosäuresequenzen, die für die Komplexierung von Plasmid DNA optimiert sind, könnten des Weiteren in dreiarmige Liganden des Typs **207** eingebaut werden, der schon mit einer weniger potenten Sequenz ein exzellenter Gentransporter ist. Da hochaffine Sequenzen bereits bei divalenten Liganden zu einer Verbesserung der Transfektionseffizienz führten, sollte eine Implementierung in ein dreiarmiges Gerüst zu einer weiteren Steigerung der bereits sehr guten Eigenschaften derartiger nichtviraler Vektoren führen, da somit auch die für den Ausbruch aus den endosomalen Vesikeln notwendige Pufferkapazität in den Transporter implementiert würde. Die Darstellung der Gentransportern **255** und **256** (Abbildung 11) würde einen direkten Vergleich mit den bereits vorhandenen zweiarmigen Analoga **233** und **236** ermöglichen und somit die allgemeine Anwendbarkeit dieser Methodik verifizieren.

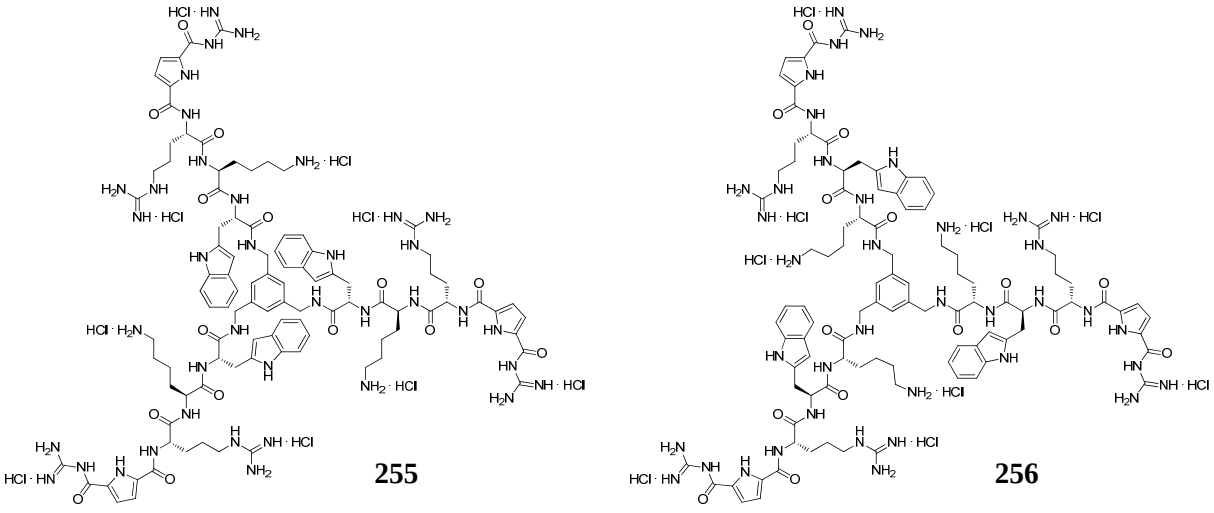
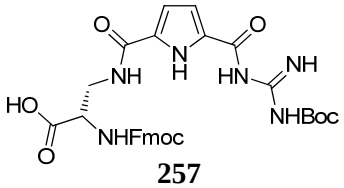


Abbildung 11. *Der Einbau von durch das Bibliotheksscreening identifizierten zu DNA hochaffinen Aminosäuresequenzen in dreiarmige Systeme wie **255** und **256** sollte hocheffiziente Gentransporter zugänglich machen.*

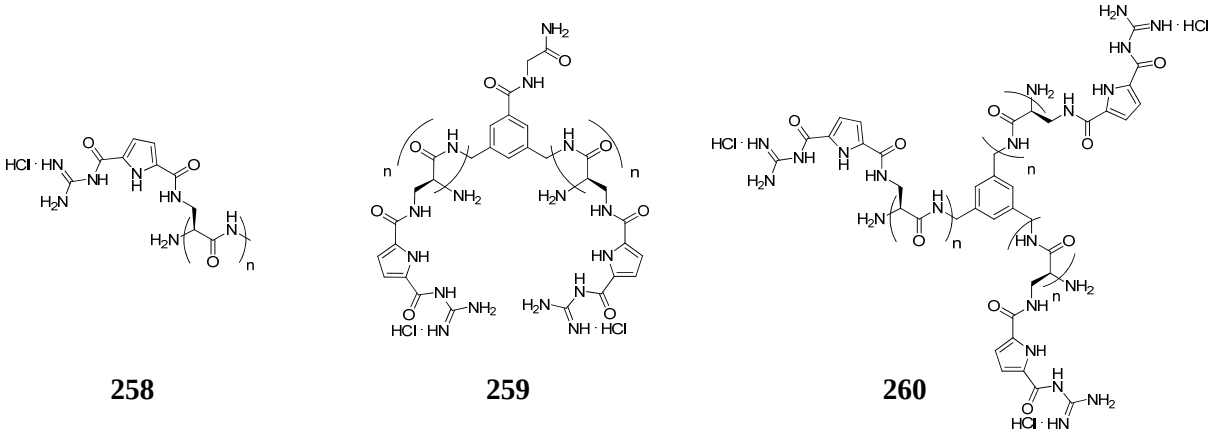
Im Allgemeinen scheint die Pufferkapazität der GCP Einheit ein entscheidendes Kriterium für die Effizienz der in dieser Arbeit entwickelten DNA-Transporter zu sein. Somit wäre die Darstellung einer zweiten Generation der Polynukleotid-Ligand-Bibliothek mit GCP Einheiten nicht nur als Kopfgruppe, sondern auch in den kombinatorischen Positionen zielführend um neue effektive Systeme zu identifizieren. Der Einbau kann sehr einfach mittels des Bausteins **257** (siehe Abbildung 12), dessen Synthese und Verwendung in der Festphasensynthese von Peptiden bereits beschrieben ist, erfolgen.[266]



257

Abbildung 12. *Der Baustein **257** ermöglicht den Einbbau von GCP Einheiten in die kombinatorischen Positionen innerhalb der Seitenketten von di- oder trivalenten DNA-Liganden. Des Weiteren können damit oligo-GCP-basierte Systeme synthetisiert werden um noch effizientere Gentransporter zu entwickeln.*

Außerdem ermöglicht dieser Baustein auch die Herstellung von oligo-GCP Einheiten mittels Fmoc-Festphasensynthese. Derartige Systeme könnten aufgrund ihrer Befähigung einerseits DNA zu binden und andererseits aufgrund ihrer Pufferkapazität den Ausbruch aus dem Endosom zu ermöglichen sehr effiziente, kleine molekulare Gentransporter darstellen. Struktur-Wirkungs-Studien bezüglich der minimalen Länge oder des Einflusses der Valenz könnten durch die Synthese von linearen Nukleinsäure-Liganden des Typs **258** (siehe Abbildung 13) oder durch den Einbau in analoge di- (**259**) oder trivalente Systeme (**260**) durchgeführt werden. Die notwendigen Synthesestrategien dazu wurden bereits innerhalb dieser Dissertation entwickelt, was eine rasche Umsetzung ermöglichen sollte.



258 **259** **260**

Abbildung 13. *Lineare (**258**), divalente (**259**) und trivalente (**260**) oligo-GCP Einheiten für die Transfektion.*

Zusammenfassend wurde mit der Arbeit, die im Rahmen dieser Dissertation durchgeführt wurde, der Grundstein für die Entwicklung von effizienten Gentransportern gelegt, indem kleine molekulare DNA-Liganden entwickelt wurden und deren Bindungseigenschaften zu verschiedenen Polynukleotiden sowie der zelluläre Aufnahmemechanismus von Plasmid/ Ligand-Komplexen in humane Zellen untersucht und erklärt werden konnte. Dieses Wissen ermöglicht es, die Charakteristika von Gentransportern in zukünftigen Forschungsprojekten weiter zu verbessern und somit dazu beizutragen, Gentherapie als eine allgemein und breit einsetzbare Methode für die Behandlung einer Vielzahl von Krankheiten zu verwirklichen.

B: LIST OF ABBREVIATIONS

°	degree
µ	micro
^{13}C	carbon isotope 13
^{1}H	proton isotope 1
A	alanine / adenine / absorption
Å	Ångstrom
AA	amino acid
abs	absolute
Abu	aminobutyric acid
Ac	acetyl
Acc	acceptor
ADP	adenosine diphosphate
AFM	atomic force microscopy
Aib	2-amino-2-methylpropionic acid
AIBN	azobisisobutyronitrile
AMP	adenosine monophosphate
analyt.	analytical
APTMS	3-amino-N-propyltrimethoxysilane
aq.	aqueous
Arg	arginine
Asn	asparagine
Atc	4-thiazolidinecarboxylic acid
ATP	adenosine triphosphate
Benz	benzene
bis-tris	2,2-bis(hydroxymethyl)-2,2´,2´´-nitrilotriethanol
Bn	benzyl
Boc	*tert*-butoxycarbonyl
bp	base pair
br.	broad
bZIP	basic-leucine zipper protein
c	centi
c	concentration
C	cytidine / Celsius
C	ITC optimization factor
cat.	catalytic
Cbz	carbobenzyloxy
CD	circular dichroism
CDP	cytidine diphosphate
cf.	compare
Chg	cyclohexyl glycine
CIS	complexation-induced chemical shift
CMC	critical micelle concentration
CMP	cytidine monophosphate
ctDNA	calf thymus DNA
CTP	cytidine triphosphate
CyD	cyclodextrine
Cys	cysteine
d	doublet / desoxy / path length
D	translational diffusion coefficient / diameter
Da	Dalton
Dapa	2,3-diaminopropionic acid
DCE	1,2-dichloroethane
DCM	dichloromethane
deg	degree
DIBAL-H	diisobutylaluminium hydride
DIPEA	diisopropylethyl amine
Disc	1,3-dihydro-2H-isoindole carboxylic acid
DLS	dynamic light scattering
DMABA	dimethylaminobutyric acid
DMF	dimethyl formamide
DMSO	dimethylsulfoxide
DNA	desoxyribonucleic acid
Do	donor
DOGS	dioctadecylamido-glycylspermine
DOSPA	2,3-dioleyloxy-N-[2-(spermine-carboxamido)ethyl]-N,N-dimethyl-1-propaneaminium trfluoroacetate
DOTAP	1,2-dioleoyl-3-trimethylammonium-propane
DOTMA	N-[1-(2,3-dioleyloxy)propyl]-N,N,N-trimethyl-ammonium chloride
DPH	1,3-diphenyl-1,3,5-hexatriene

dppe	1,5-bis(diphenylphosphino)pentane
ds	double stranded
e^-	electron
E	glutamic acid
EI	electron ionization
e.g.	for example
EA	ethyl acetate
EB	ethidium bromide
EDT	1,2-ethanedithiol
eq	equivalent
ESI	electrospray ionization
et al.	et alii
Et	ethyl
Fmoc	fluorenylmethyloxycarbonyl
FT-IR	fourier-transform infrared
g	gramm
G	guanine
GABA	γ-aminobutyric acid
GC	gas chromatography
GCP	guanidiniocarbonyl pyrrole
GDP	guanosine diphosphate
GFP	green fluorescence protein
GMP	guanosine monophosphate
GSABA	generalized Born/surface area
GTP	guanosine triphosphate
H	height
H	proton
HBTU	2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HCTU	1-(bis-dimethylamino-methylene)-5-chloro-3-oxy-1H-benzotriazol-1-ium hexafluorophosphate
HEK293T	human embryonic kidney cell line
HeLa	human cancer cell line
His	histidine
HOBt	1-hydroxybenzotriazole
Hp	hydroxypyrrole
HPLC	high pressure liquid chromatography
HR	high resolution
Hz	Hertz
i.e.	that is
IC_{50}	equivalents necessary to displace half of the bound polynucleotide-ligand
ICD	induced CD
Ile	isoleucine
Im	imidazole
IR	infrared
ITC	isothermal titration calorimetry
k	kilo / Boltzmann´s constant
K	lysine
L	length
L	liter / ligand
Lamp1	lysosome associated membrane protein
LCP	left-circular polarized light
Leu	leucine
Lys	lysine
M	mega / mol/l
m	*meta*
m	milli / medium / multiplett / meter / methyl
m/z	mass per charge
max.	maximal
MBHA	4-methylbenzhydrylamine
MD	molecular dynamics
Me	methyl
Met	methionine
min	minute(s)
MM	molecular modeling
Mp.	melting point
MPLC	medium performance liquid chromatography
mRNA	messenger RNA
MS	mass spectrometry
n	nano
n	stoichiometry
n.d.	not determined

NBS	N-bromosuccinimide
NMM	N-methylmorpholine
NMP	N-methylpyrrolidone
NMR	nuclear magnetic resonance
NP	nanoparticles / normal phase
ORF	open reading frame
p	*para*
P	phosphate
p	poly
Pbf	2,2,4,6,7-pentamethyldihydrobenzo-furan-5-sulfonyl
PBS	phosphate buffered saline
*i*Prop	isopropanol
Pbf	2,2,4,6,7-pentamethyldihydrobenzo-furan-5-sulfonyl
PBS	phosphate buffered saline
Pd/C	palladium on charcoal
PDI	polydispersity index
PEG	polyethylene glycol
PEI	polyethylenimine
PEM	photoelastic modulator
pF143-GFP	plasmid encoding for a green fluo-rescence protein
Ph	phenyl
pH	pondus hydrogenii
Phe	phenylalanine
Pip	piperidine
PLL	poly(l-lysine)
PN	polynucleotide
Ppa	piperazin
PPi	diphospahte
ppm	parts per million
pre	premature
Pro	proline
PS	polystyrene
Py	pyrrole
PyBOP	benzotriazol-1-yloxytripyrrolidino-phoshonium hexafluorophosphate
q	quadruplet
r	equivalents of ligand with respect to the nucleic acid´s phosphate groups
Ra	Raman
RCP	right-circular polarized light
R_f	retention factor
RFP	red fluorescence protein
R_H	hydrodynamic radius
RNA	ribonucleic acid
RP	reversed phase
RR	Resonance Raman
s	strong / singulett / second
Sar	sarcosine
sec	second
SEM	surface electron microscopy
Ser	serine
SERS	surface enhanced Raman spectros-copy
SPPS	solid phase peptide synthesis
ss	single stranded
t	*tert* / time
T	thymidine
t	triplet
TEM	transmission electron microscopy
TFA	trifluoroacetic acid
Thr	threonine
Tic	1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid
TLC	thin layer chromatography
T_m	melting point
TMEDA	N,N,N´,N´-tetramethylethylenedi-amine
TMP	thymidine monophosphate
ToF	time of flight
Trp	tryptophan
TTP	thymidine triphosphate
Tyr	tyrosine
U	uracil
UDP	uridine diphosphate
UMP	uridine monophosphate
UTP	uridine triphosphate

UTR	untranslated region
UV	ultraviolet
Val	valine
Vis	visible
vs.	versus
W	Watt
w	weak
W	width
δ	chemical shift / standard deviation
η	viscosity
λ	wavelength

C: SUPPLEMENTARY EXPERIMENTAL DATA

C.1 HPLC Data

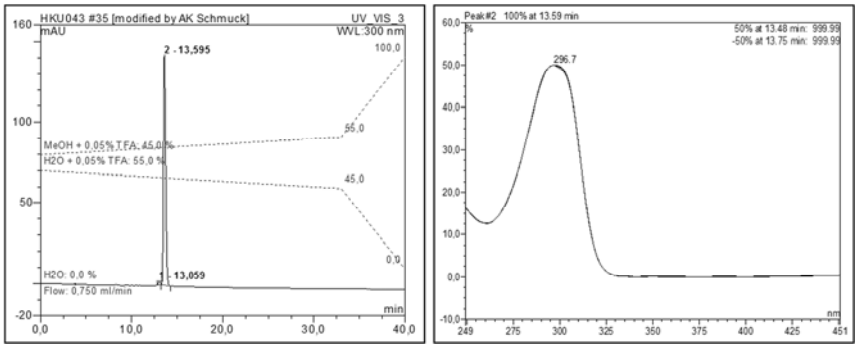


Figure C. 1. *The purity of the non-symmetric tweezer receptor* ***187*** *was determined to 99 % via analytical HPLC on a Supelcosil™ LC-18 column (25 cm × 4.6 mm, 5 µm), solvent: 45 to 55 % methanol/water in 33 min (0.05 % TFA), flow rate 0.75 mL/min, retention time t_R = 13.6 min; the peak integration is based on detection at 300 nm.*

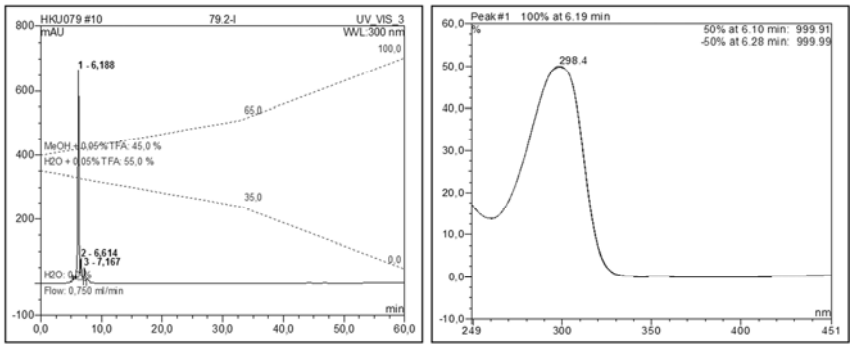


Figure C. 2. *The purity of the non-symmetric tweezer receptor* ***188*** *was determined to 90 % via analytical HPLC on a Supelcosil™ LC-18 column (25 cm × 4.6 mm, 5 µm), solvent: 45 to 65 % methanol/water in 32 min (0.05 % TFA), flow rate 0.75 mL/min, retention time t_R = 6.2 min; the peak integration is based on detection at 300 nm.*

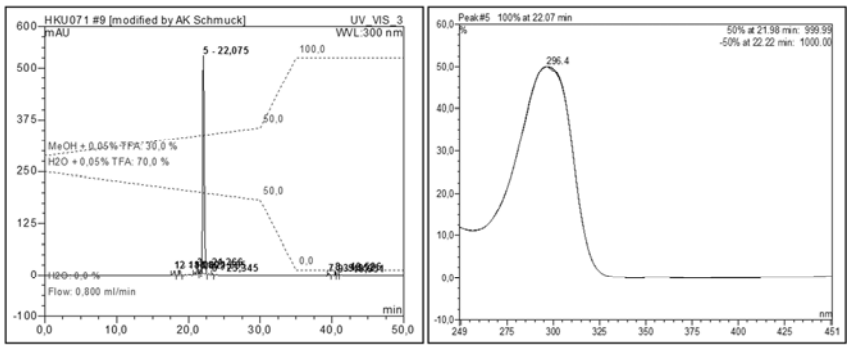


Figure C. 3. *The purity of the symmetric tweezer receptor* ***196*** *was determined to 94 % via analytical HPLC on a YMC ODS-A column (15 cm × 3.0 mm, 5 µm), solvent: 40 to 50 % methanol/water in 30 min (0.05 % TFA), flow rate 0.43 mL/min, retention time tR = 22.1 min; the peak integration is based on detection at 300 nm. Reprinted with permission from Kuchelmeister, H. Y.; Schmuck, C. Chem. Eur. J.* ***2011****, 17, 5311-5318. Copyright 2011 John Wiley and Sons.*[206]

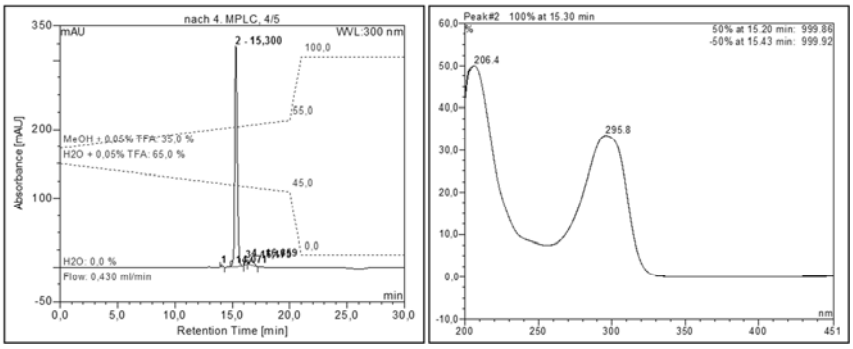


Figure C. 4. *The purity of the symmetric tweezer receptor* ***206*** *was determined to 94 % via analytical HPLC on a YMC ODS-A column (15 cm × 3.0 mm, 5 µm), solvent: 35 to 55 % methanol/water in 20 min (0.05 % TFA), flow rate 0.43 mL/min, retention time tR = 16.3 min; the peak integration is based on detection at 300 nm.*

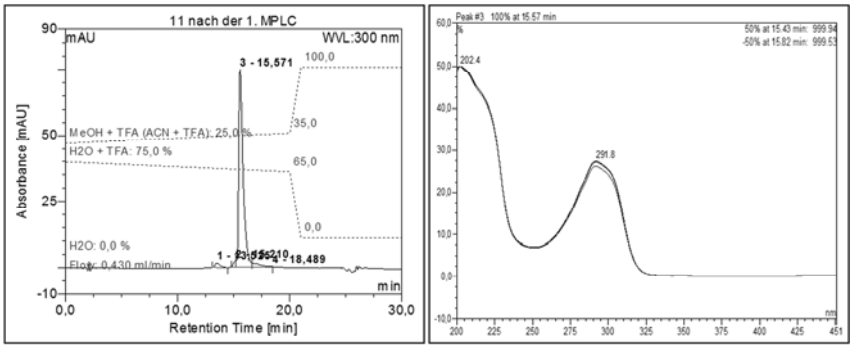


Figure C. 5. *The purity of the symmetric divalent ligand* ***233*** *was determined to 93 % via analytical HPLC on a YMC ODS-A column (15 cm × 3.0 mm, 5 µm), solvent: 25 to 35 % methanol/water in 20 min (0.05 % TFA), flow rate 0.43 mL/min, retention time t_R = 15.6 min; the peak integration is based on detection at 300 nm.*

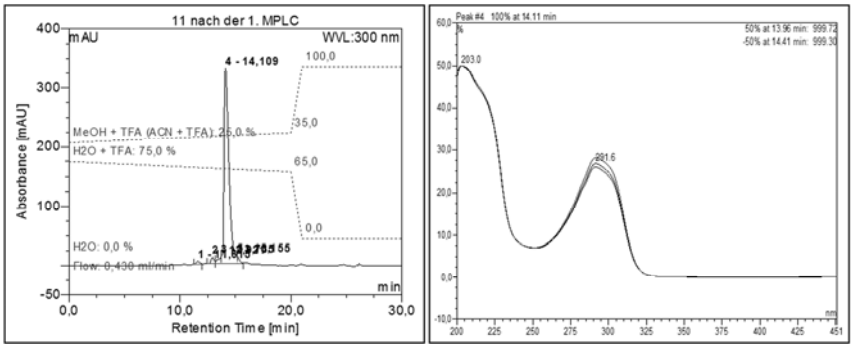


Figure C. 6. *The purity of the symmetric divalent ligand* ***234*** *was determined to 93 % via analytical HPLC on a YMC ODS-A column (15 cm × 3.0 mm, 5 µm), solvent: 25 to 35 % methanol/water in 20 min (0.05 % TFA), flow rate 0.43 mL/min, retention time t_R = 14.1 min; the peak integration is based on detection at 300 nm.*

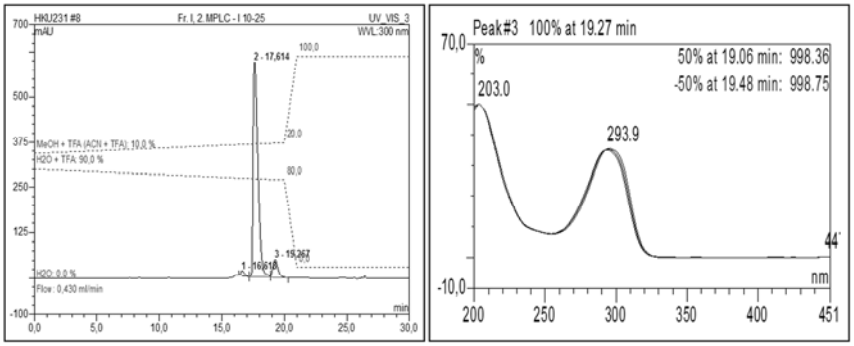


Figure C. 7. *The purity of the symmetric divalent ligand* ***235*** *was determined to 91 % via analytical HPLC on a YMC ODS-A column (15 cm × 3.0 mm, 5 µm), solvent: 10 to 10 % methanol/water in 20 min (0.05 % TFA), flow rate 0.43 mL/min, retention time t_R = 17.6 min; the peak integration is based on detection at 300 nm.*

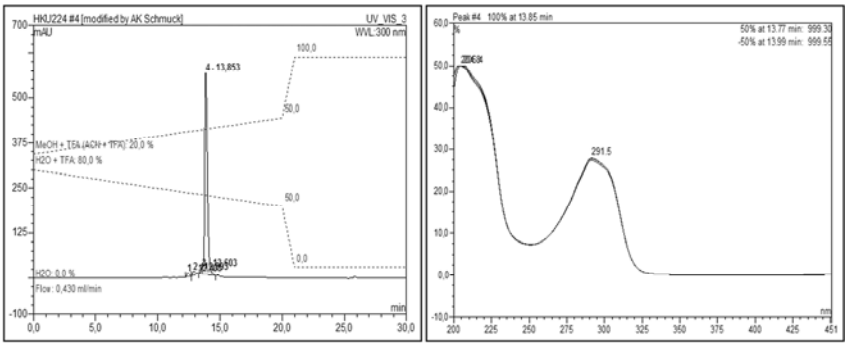


Figure C. 8. *The purity of the symmetric divalent ligand* ***236*** *was determined to 95 % via analytical HPLC on a YMC ODS-A column (15 cm × 3.0 mm, 5 µm), solvent: 20 to 50 % methanol/water in 20 min (0.05 % TFA), flow rate 0.43 mL/min, retention time t_R = 13.9 min; the peak integration is based on detection at 300 nm.*

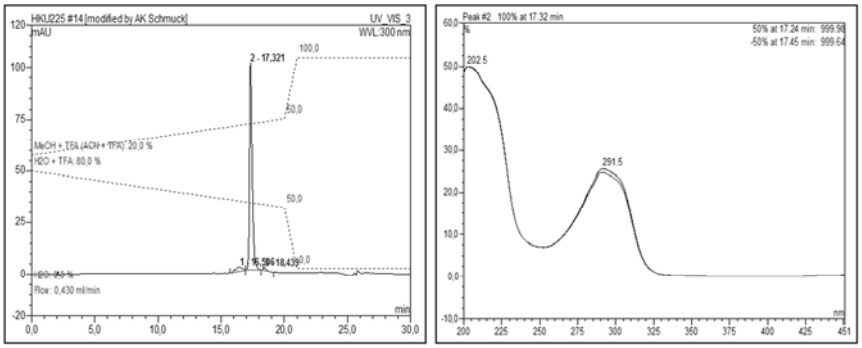


Figure C. 9. *The purity of the symmetric divalent ligand* ***237*** *was determined to 94 % via analytical HPLC on a YMC ODS-A column (15 cm × 3.0 mm, 5 µm), solvent: 20 to 50 % methanol/water in 20 min (0.05 % TFA), flow rate 0.43 mL/min, retention time t_R = 17.3 min; the peak integration is based on detection at 300 nm.*

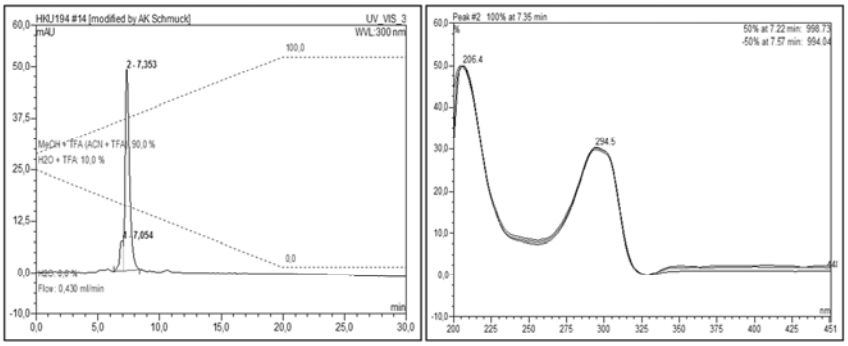


Figure C. 10. *The purity of the amphiphilic ligand* ***243*** *was determined to 90 % via analytical HPLC on a YMC ODS-A column (15 cm × 3.0 mm, 5 µm), solvent: 10 to 100 % methanol/water in 20 min (0.05 % TFA), flow rate 0.43 mL/min, retention time t_R = 7.4 min; the peak integration is based on detection at 300 nm.*

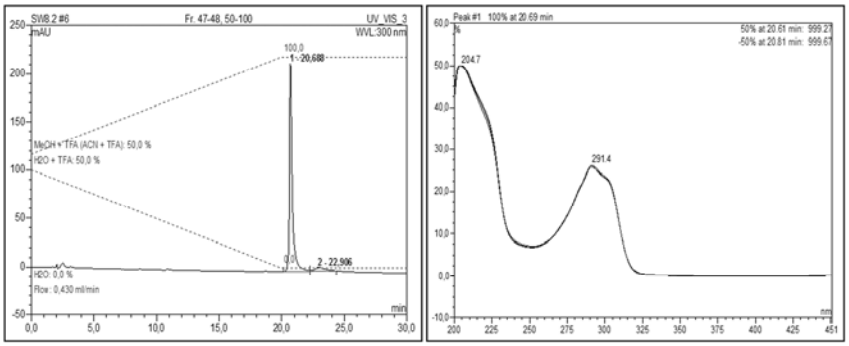


Figure C. 11. *The purity of amphiphilic ligand* ***250*** *was determined to 94 % via analytical HPLC on a YMC ODS-A column (15 cm × 3.0 mm, 5 µm), solvent: 50 to 1000 % methanol/water in 20 min (0.05 % TFA), flow rate 0.43 mL/min, retention time t_R = 20.7 min; the peak integration is based on detection at 300 nm.*

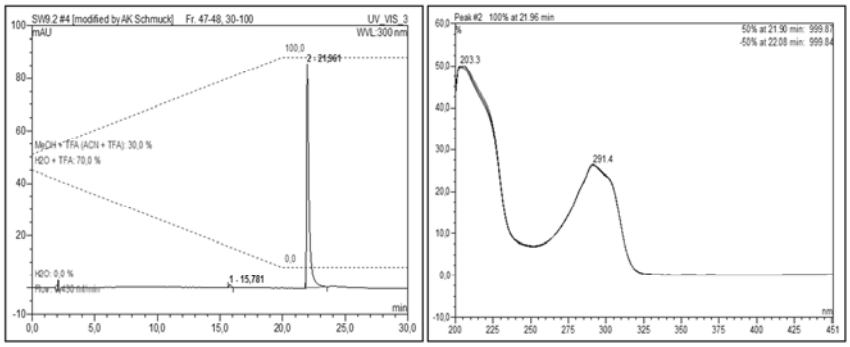


Figure C. 12. *The purity of the amphiphilic ligand* ***251*** *was determined to 99 % via analytical HPLC on a YMC ODS-A column (15 cm × 3.0 mm, 5 µm), solvent: 30 to 100 % methanol/water in 20 min (0.05 % TFA), flow rate 0.43 mL/min, retention time* t_R *= 22.0 min; the peak integration is based on detection at 300 nm.*

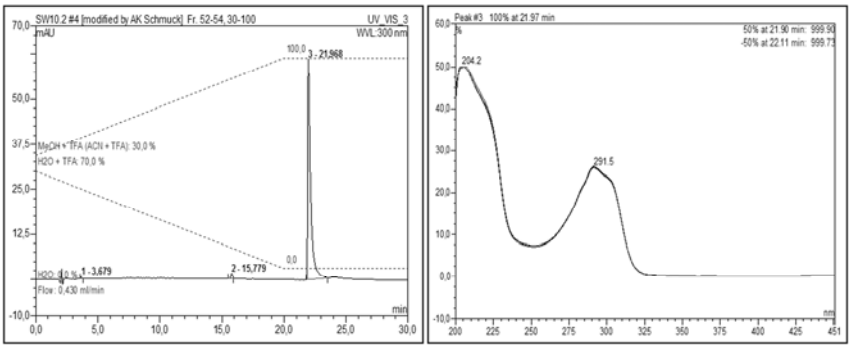


Figure C. 13. *The purity of the amphiphilic ligand* ***252*** *was determined to 99 % via analytical HPLC on a YMC ODS-A column (15 cm × 3.0 mm, 5 µm), solvent: 30 to 100 % methanol/water in 20 min (0.05 % TFA), flow rate 0.43 mL/min, retention time* t_R *= 22.0 min; the peak integration is based on detection at 300 nm.*

C.2 UV/Vis and Fluorescence Titrations

UV/Vis Titrations

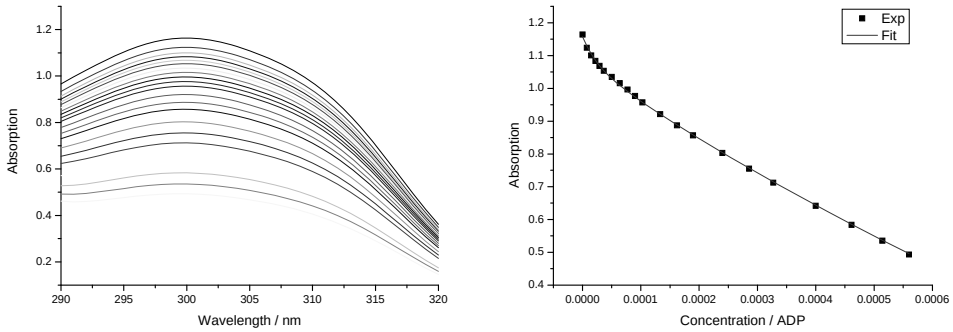


Figure C. 14. *UV Titration with ADP and excerpt at 300 nm of global fit according to SpecFit.*

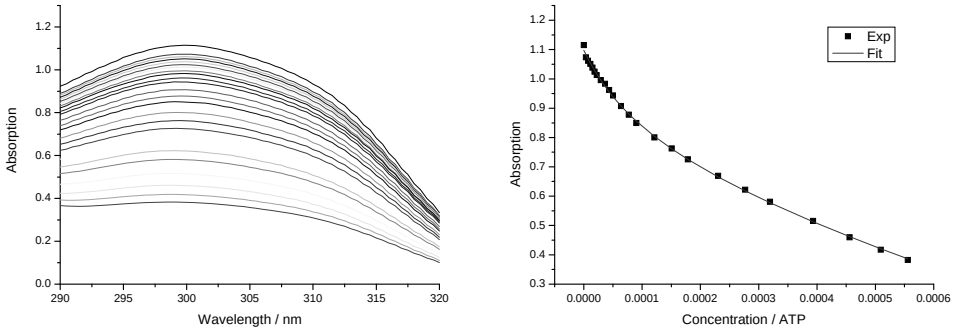


Figure C. 15. *UV Titration with ATP and excerpt at 300 nm of global fit according to SpecFit.*

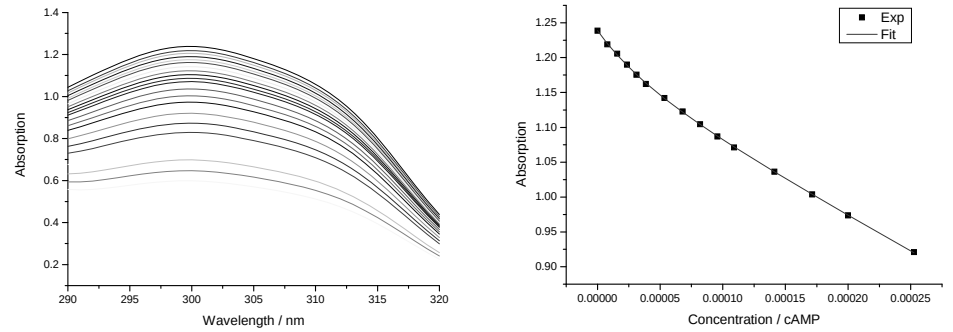


Figure C. 16. *UV Titration with cAMP and excerpt at 300 nm of global fit according to SpecFit.*

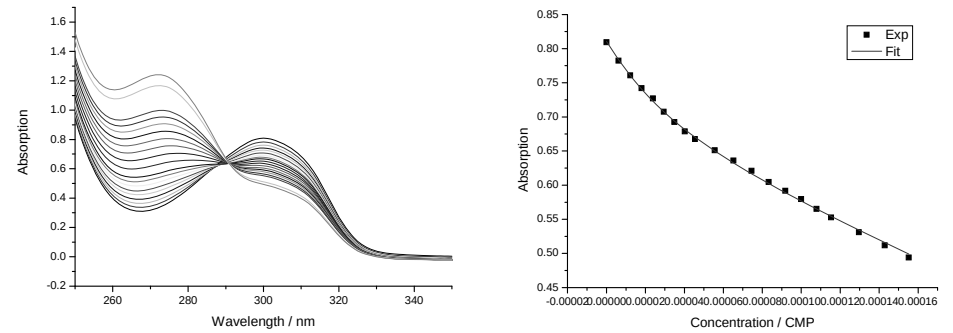


Figure C. 17. *UV Titration with CMP and excerpt at 300 nm of global fit according to SpecFit.*

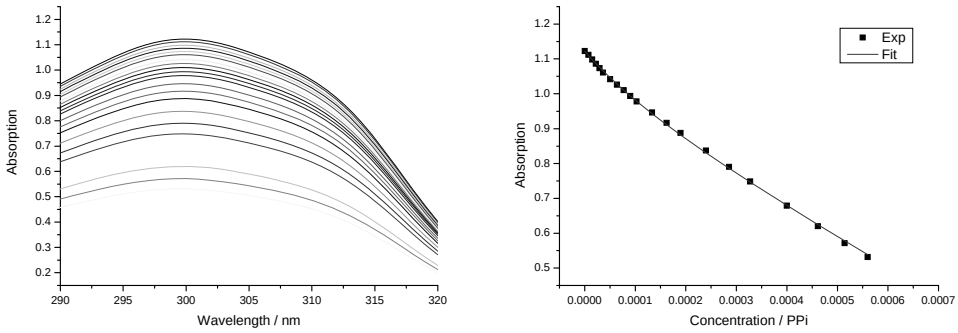


Figure C. 18. *UV Titration with PPi and excerpt at 300 nm of global fit according to SpecFit.*

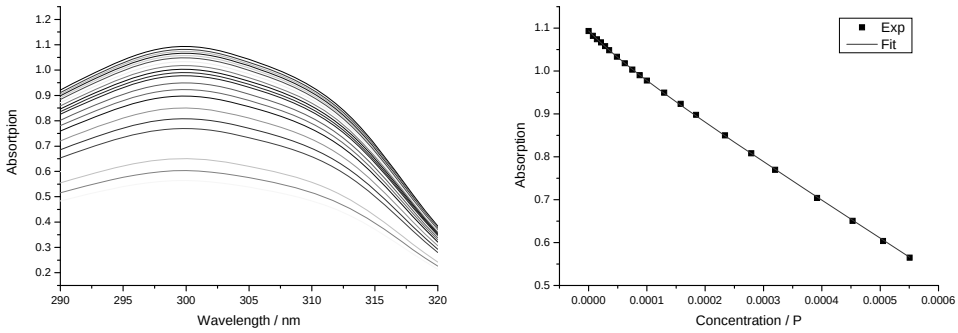


Figure C. 19. *UV Titration with P and excerpt at 300 nm of global fit according to SpecFit.*

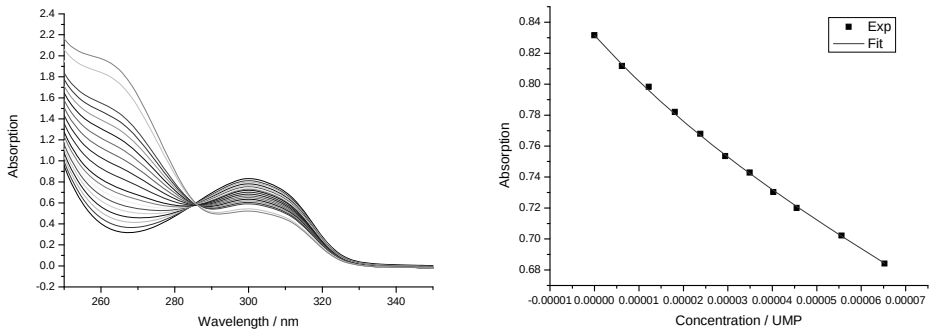


Figure C. 20. *UV Titration with UMP and excerpt at 300 nm of global fit according to SpecFit.*

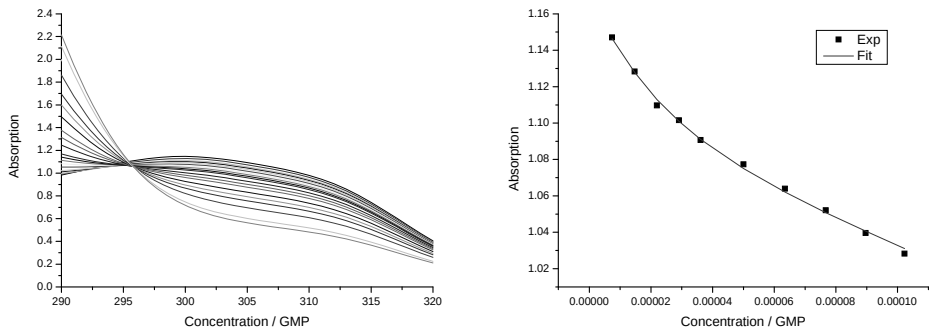


Figure C. 21. *UV Titration with GMP and excerpt at 300 nm of global fit according to SpecFit.*

Fluorescence Titrations

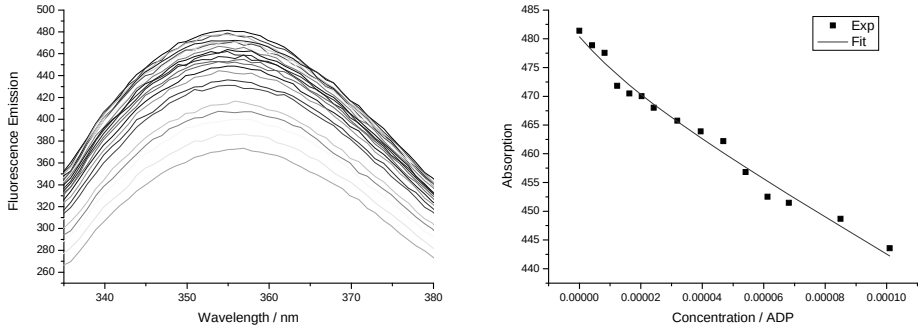


Figure C. 22. *Fluorescence Titration with ADP and excerpt at 355 nm of global fit according to SpecFit.*

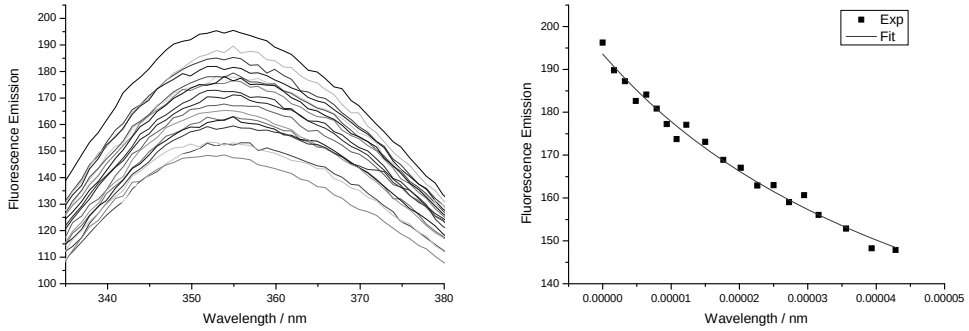


Figure C. 23. *Fluorescence Titration with ATP and excerpt at 355 nm of global fit according to SpecFit.*

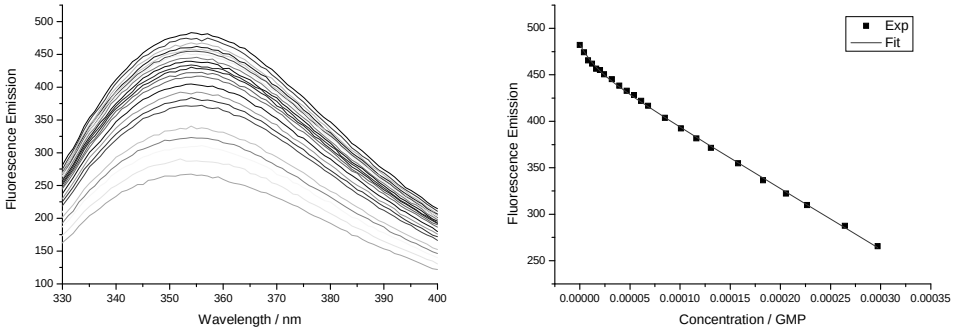


Figure C. 24. *Fluorescence Titration with GMP and excerpt at 355 nm of global fit according to SpecFit.*

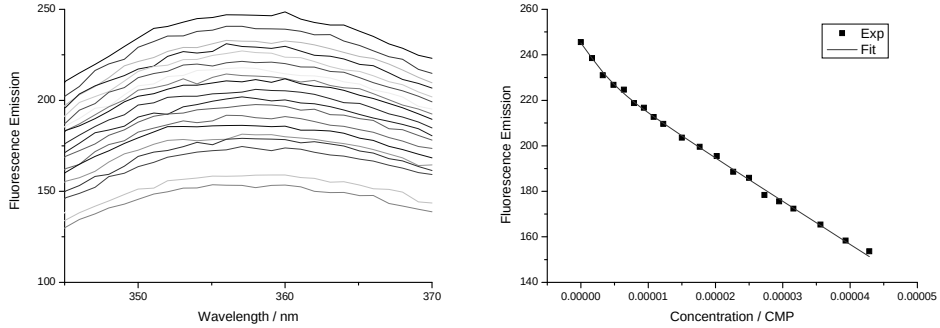


Figure C. 25. *Fluorescence Titration with CMP and excerpt at 355 nm of global fit according to SpecFit.*

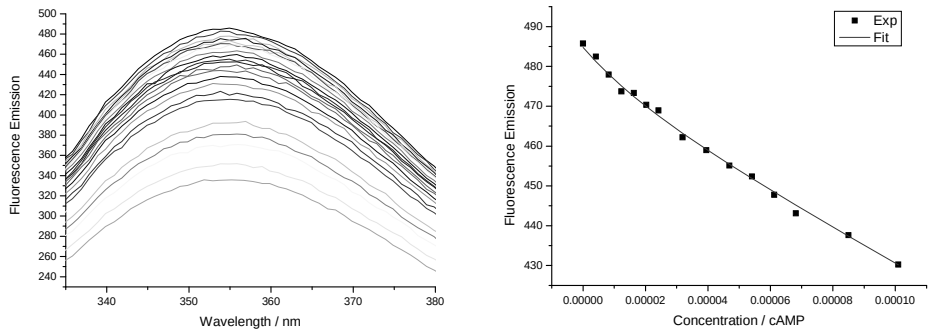


Figure C. 26. *Fluorescence Titration with cAMP and excerpt at 355 nm of global fit according to SpecFit.*

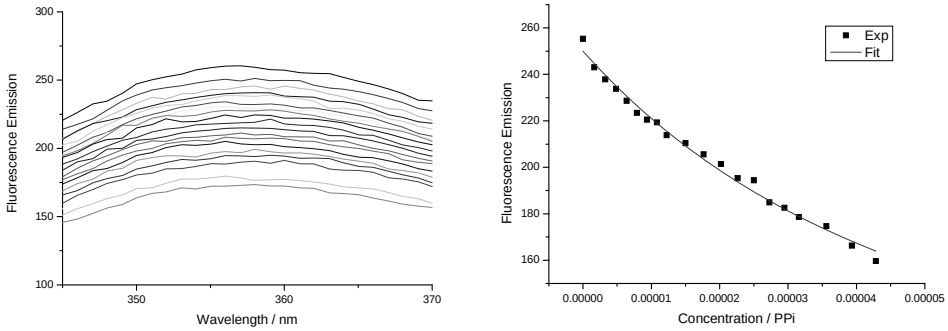


Figure C. 27. *Fluorescence Titration with PPi and excerpt at 355 nm of global fit according to SpecFit.*

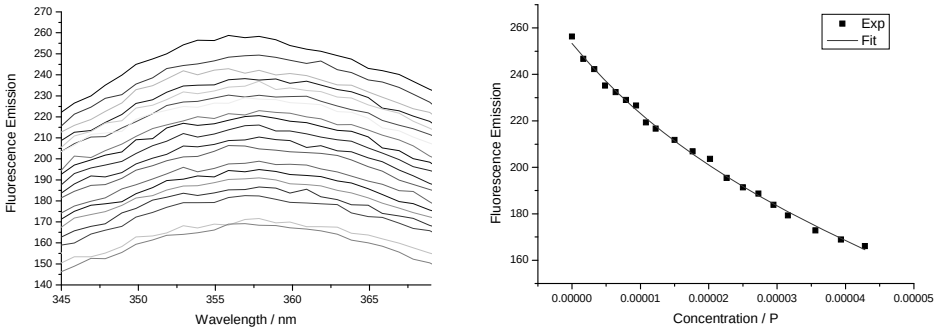


Figure C. 28. *Fluorescence Titration with P and excerpt at 355 nm of global fit according to SpecFit.*

C.3 Thermal Melting Experiments

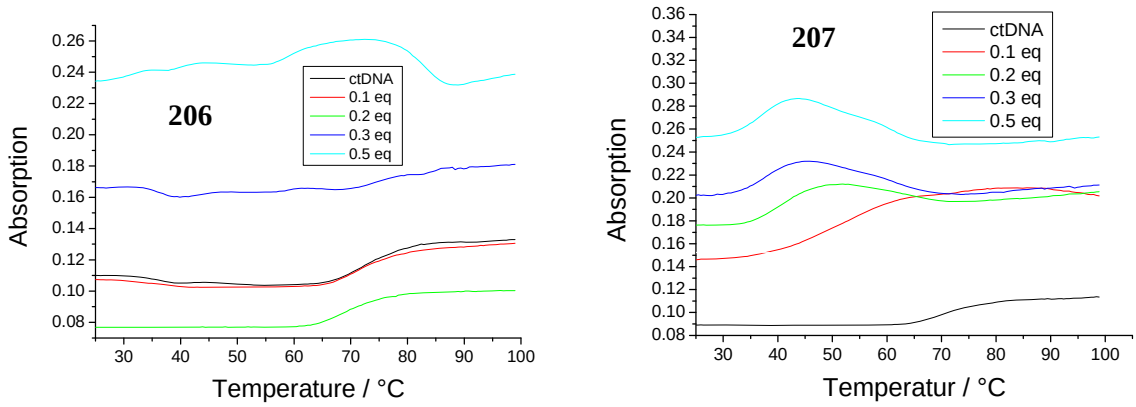


Figure C. 29. *Thermal melting experiment of ctDNA with* ***206*** *(left) and* ***207*** *(right).*

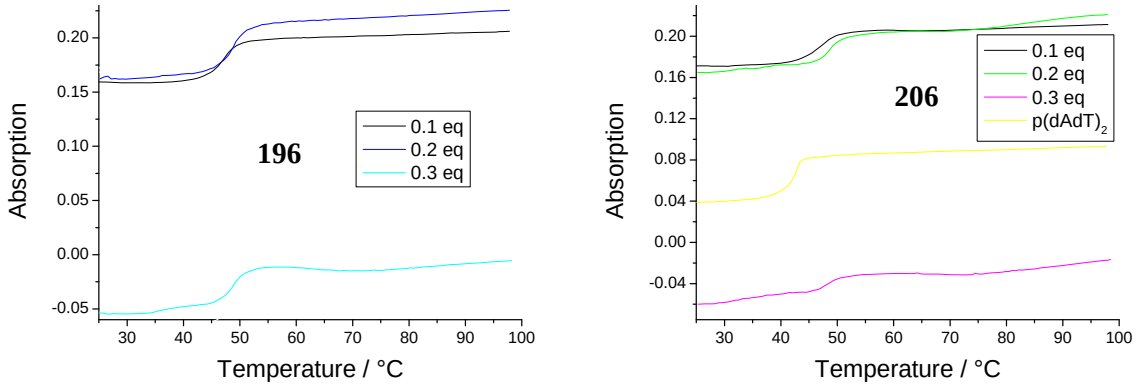


Figure C. 30. *Thermal melting experiment of p(dAdT)$_2$ with* ***196*** *(left) and* ***206*** *(right).*

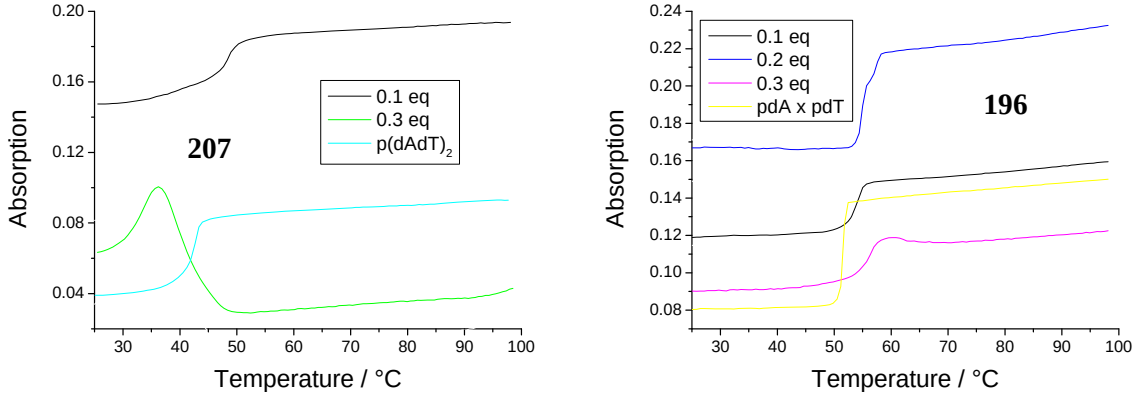


Figure C. 31. *Thermal melting experiment of p(dAdT)$_2$ with* ***207*** *(left) and of pdA × pdT with* ***196*** *(right).*

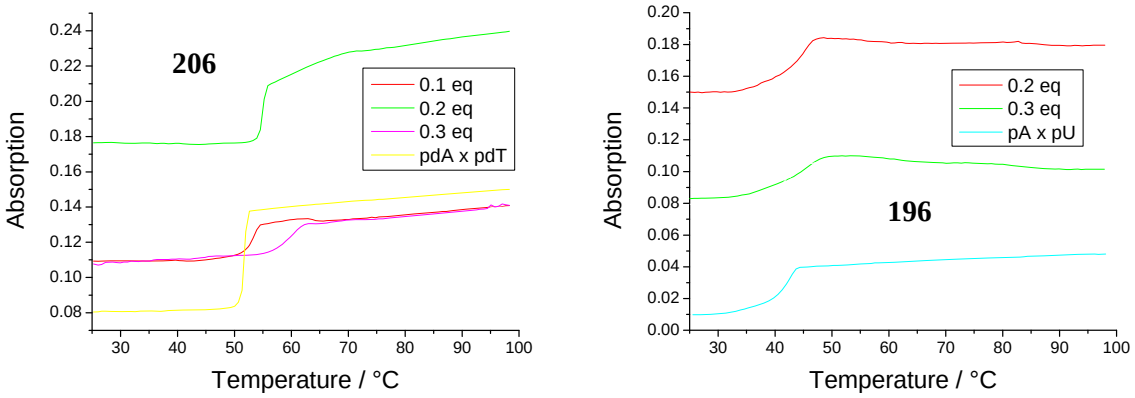


Figure C. 32. *Thermal melting experiment of pdA × pdT with* ***206*** *(left) and of pA × pU with* ***196*** *(right).*

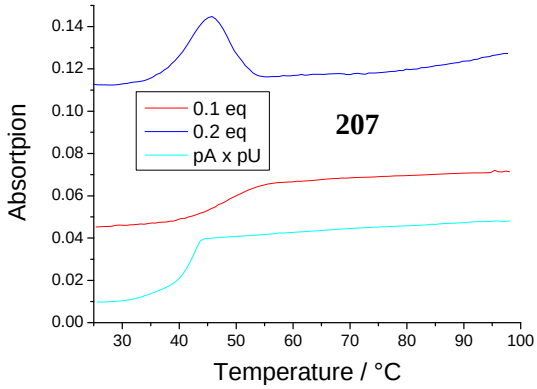


Figure C. 33. *Thermal melting experiment of pA × pU with* ***207****.*

C.4 EB Displacement Assays

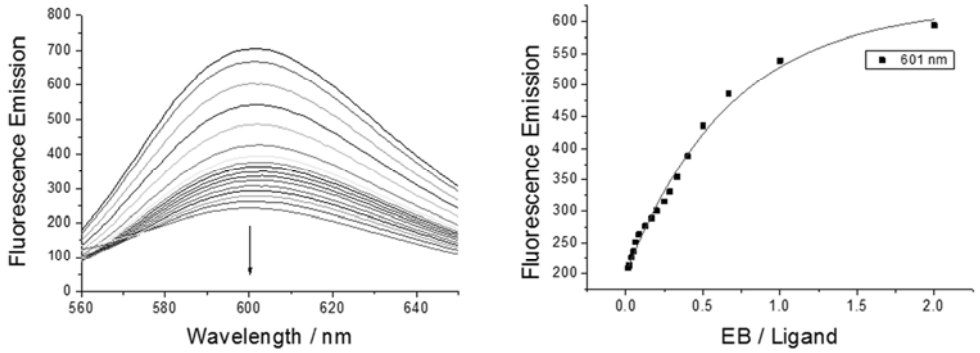


Figure C. 34. *EB displacement experiment with ctDNA and* ***196*** *(left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/****196*** *with the corresponding fit (right).*

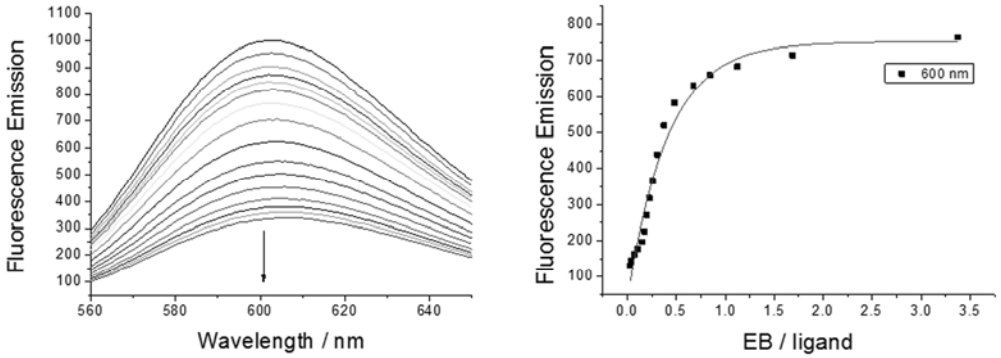


Figure C. 35. *EB displacement experiment with ctDNA and* ***206*** *(left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/****206*** *with the corresponding fit (right).*

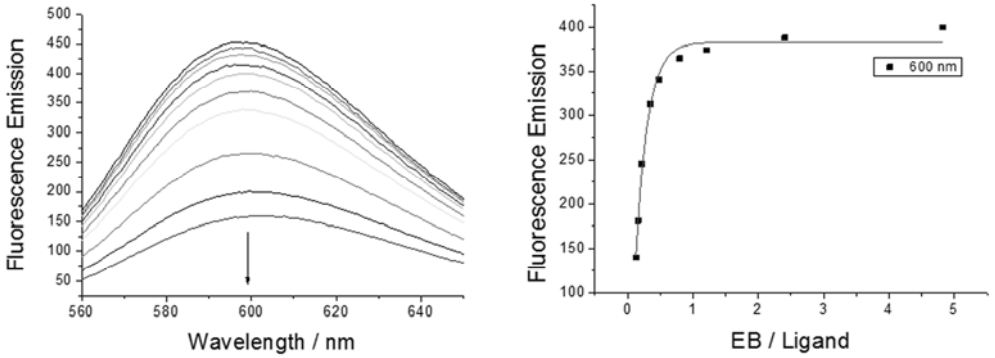


Figure C. 36. *EB displacement experiment with p(dAdT)$_2$ and* ***196*** *(left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/****196*** *with the corresponding fit (right).*

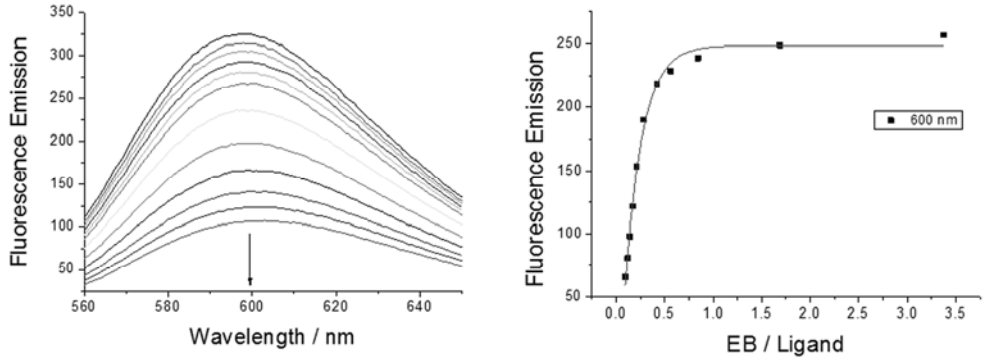


Figure C. 37. *EB displacement experiment with p(dAdT)$_2$ and* ***206*** *(left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/****206*** *with the corresponding fit (right).*

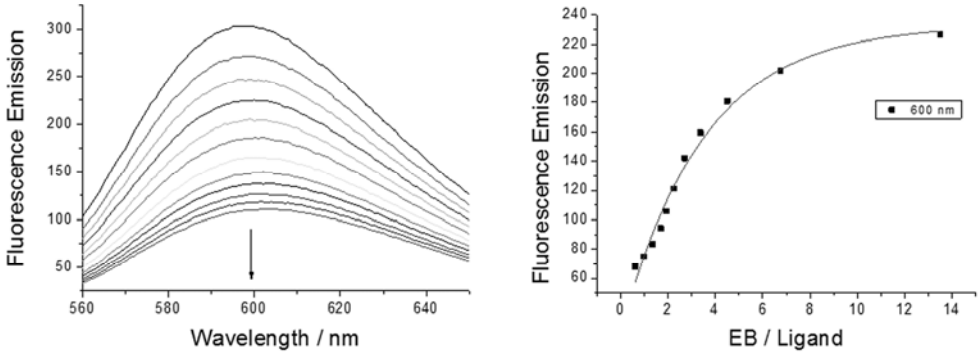


Figure C. 38. *EB displacement experiment with $p(dAdT)_2$ and **207** (left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/**207** with the corresponding fit (right).*

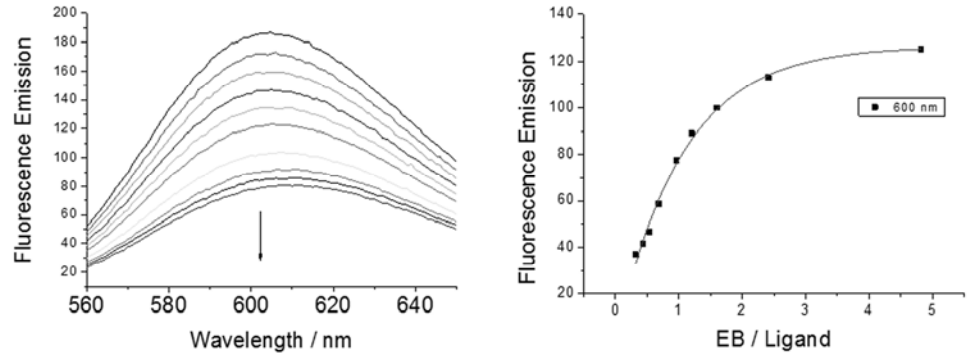


Figure C. 39. *EB displacement experiment with pdA × pdT and **196** (left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/**196** with the corresponding fit (right).*

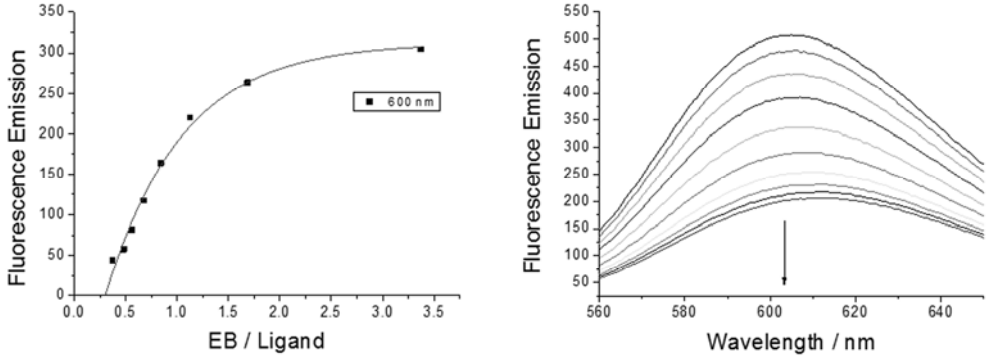


Figure C. 40. *EB displacement experiment with pdA × pdT and **206** (left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/**206** with the corresponding fit (right).*

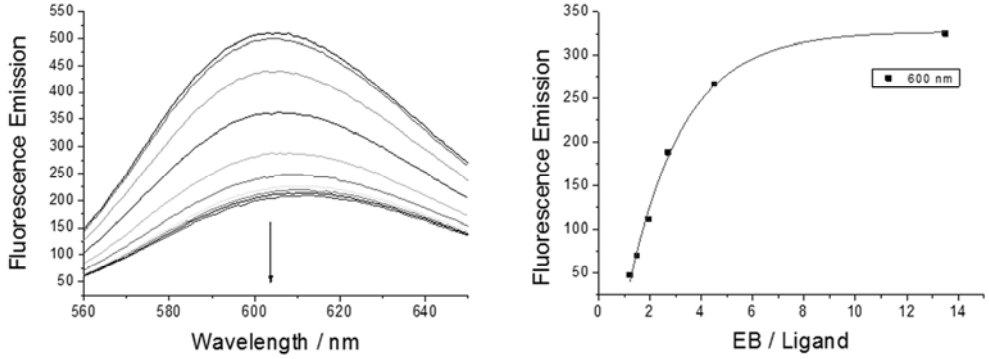


Figure C. 41. *EB displacement experiment with pdA × pdT and **207** (left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/**207** with the corresponding fit (right).*

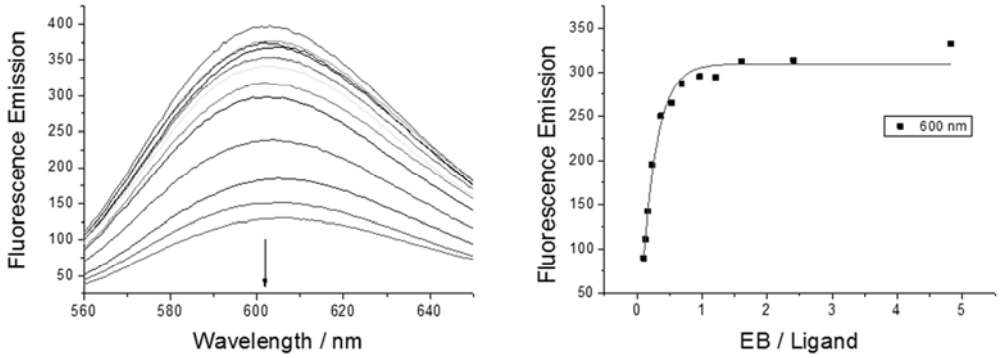


Figure C. 42. *EB displacement experiment with p(dGdC)$_2$ and **196** (left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/**196** with the corresponding fit (right).*

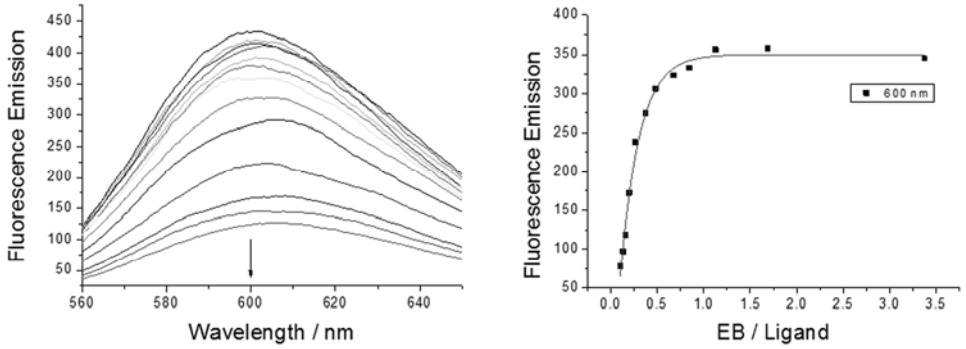


Figure C. 43. *EB displacement experiment with p(dGdC)$_2$ and **206** (left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/**206** with the corresponding fit (right).*

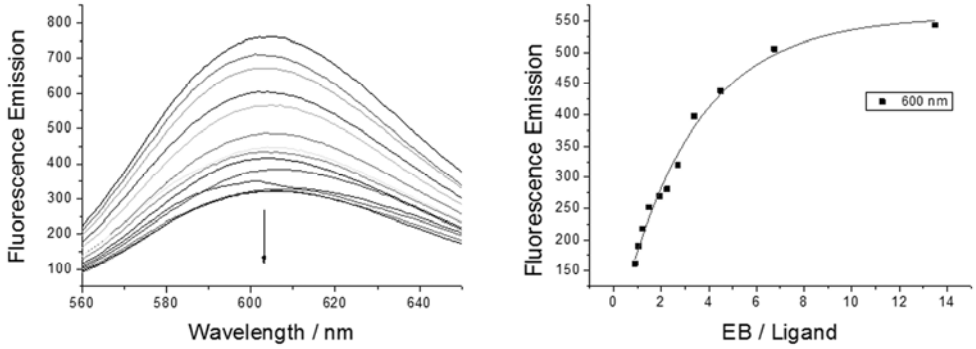


Figure C. 44. *EB displacement experiment with p(dGdC)$_2$ and **207** (left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/**207** with the corresponding fit (right).*

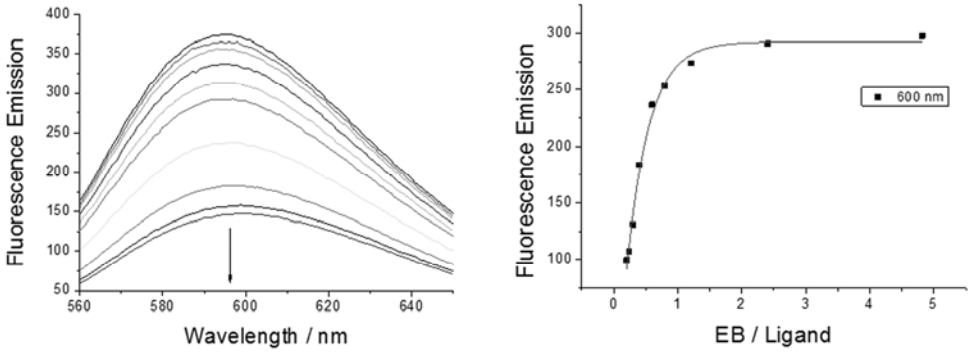


Figure C. 45. *EB displacement experiment with pA × pU and **196** (left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/**196** with the corresponding fit (right).*

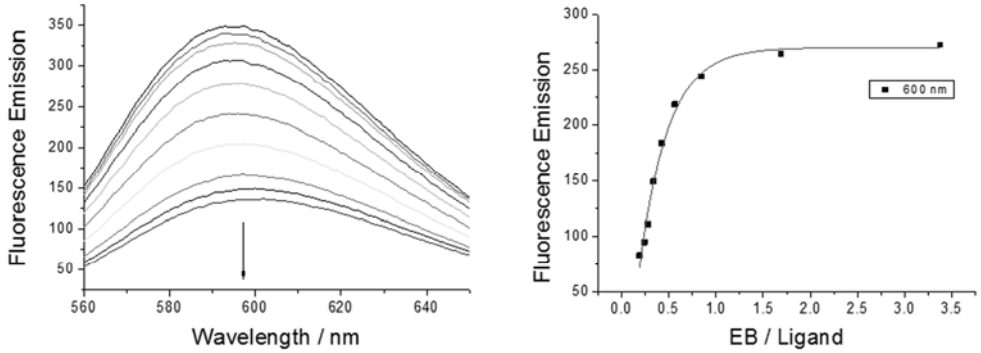


Figure C. 46. *EB displacement experiment with pA × pU and* ***206*** *(left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/****206*** *with the corresponding fit (right).*

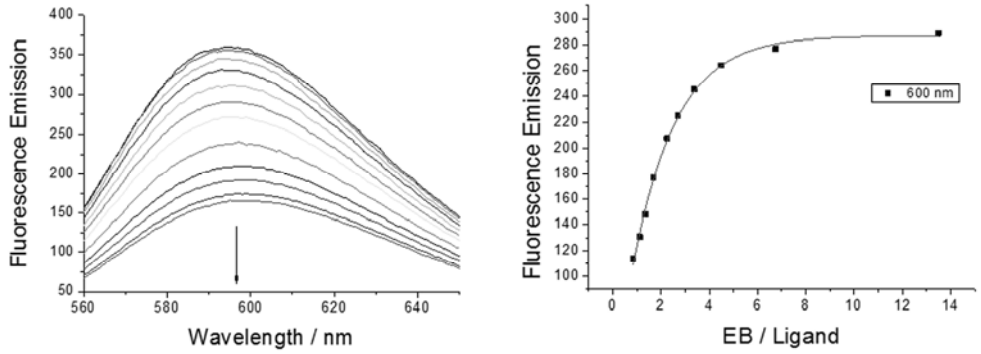


Figure C. 47. *EB displacement experiment with pA × pU and* ***207*** *(left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/****207*** *with the corresponding fit (right).*

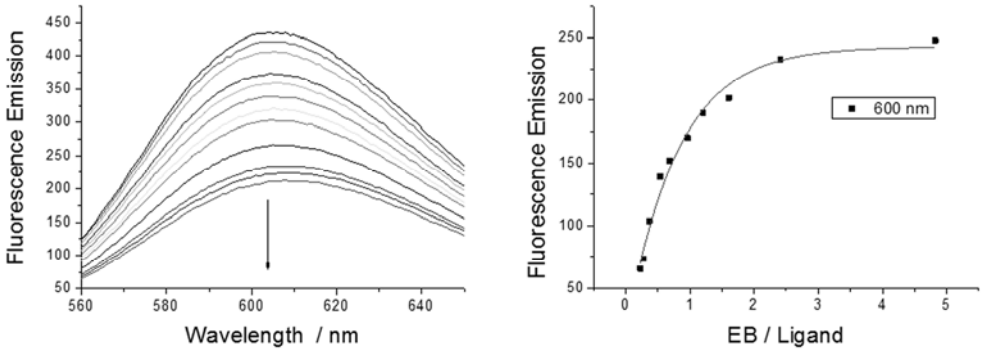


Figure C. 47. *EB displacement experiment with pF143-GFP and* ***196*** *(left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/****196*** *with the corresponding fit (right).*

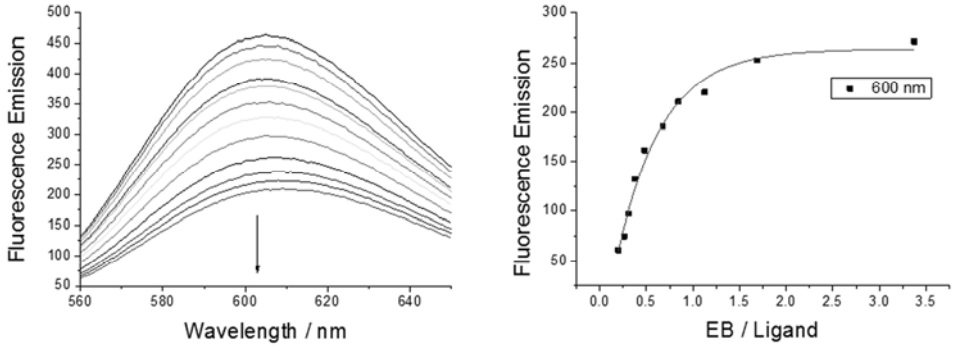


Figure C. 47. *EB displacement experiment with pF143-GFP and* ***206*** *(left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/****206*** *with the corresponding fit (right).*

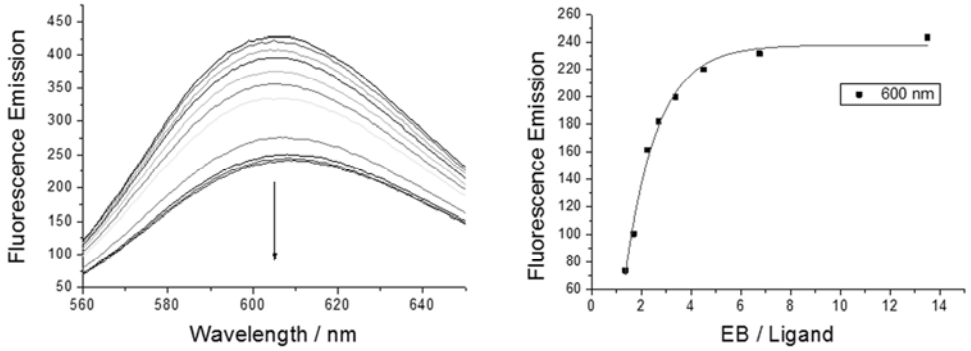


Figure C. 47. *EB displacement experiment with pF143-GFP and* ***207*** *(left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/****207*** *with the corresponding fit (right).*

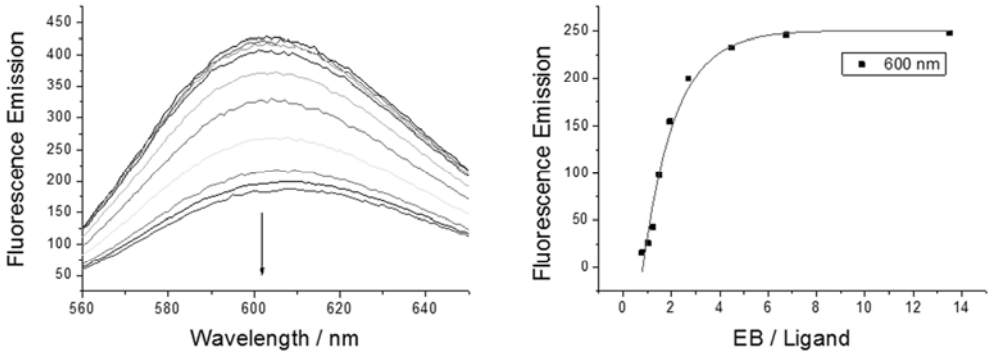


Figure C. 51. *EB displacement experiment with pF143-GFP and* ***233*** *(left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/****233*** *with the corresponding fit (right).*

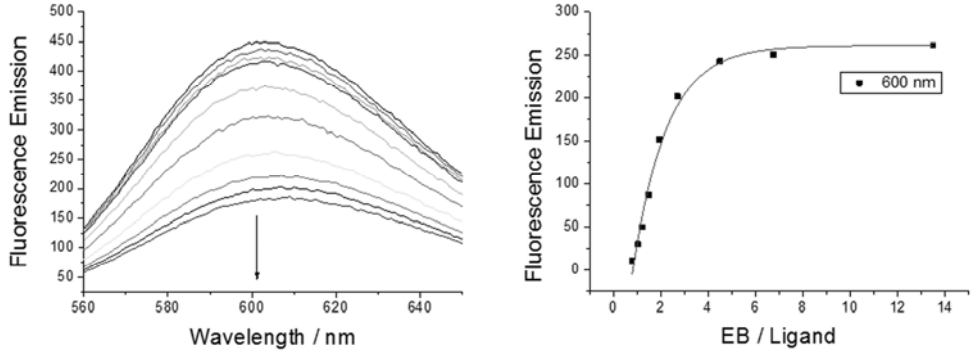


Figure C. 52. *EB displacement experiment with pF143-GFP and* ***234*** *(left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/****234*** *with the corresponding fit (right).*

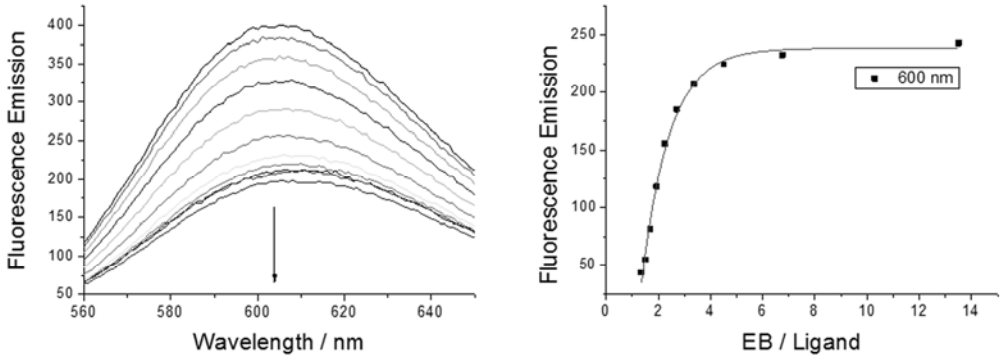


Figure C. 53. *EB displacement experiment with pF143-GFP and* ***235*** *(left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/****235*** *with the corresponding fit (right).*

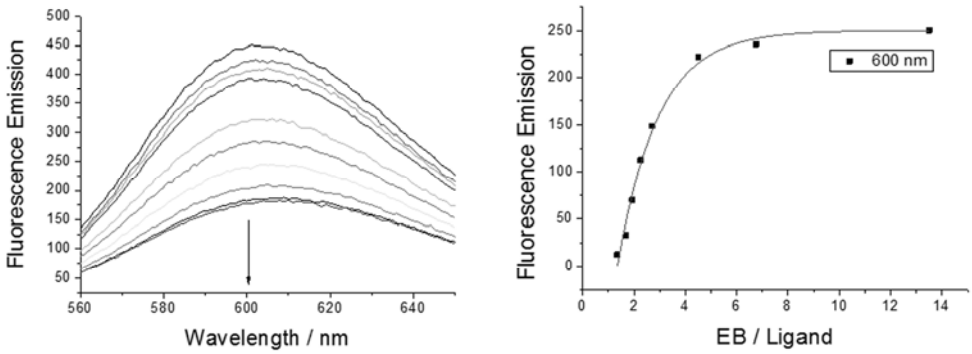


Figure C. 54. *EB displacement experiment with pF143-GFP and* ***236*** *(left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/****236*** *with the corresponding fit (right).*

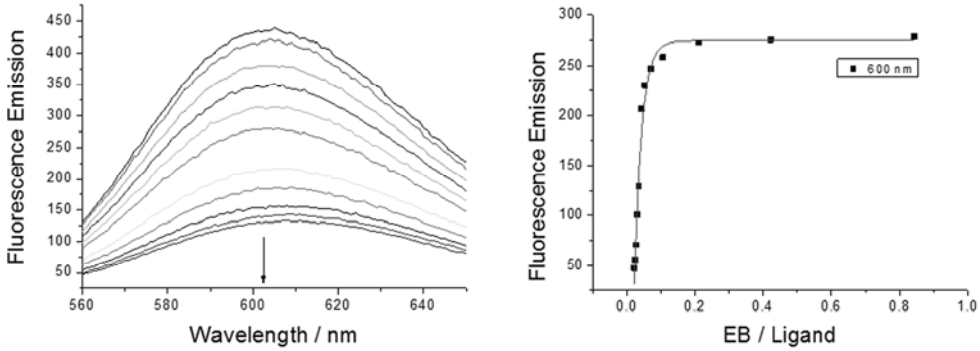


Figure C. 55. *EB displacement experiment with pF143-GFP and* ***237*** *(left) and excerpt of the fluorescence emission at 600 nm plotted against the ratio of EB/****237*** *with the corresponding fit (right).*

C.5 ITC Experiments

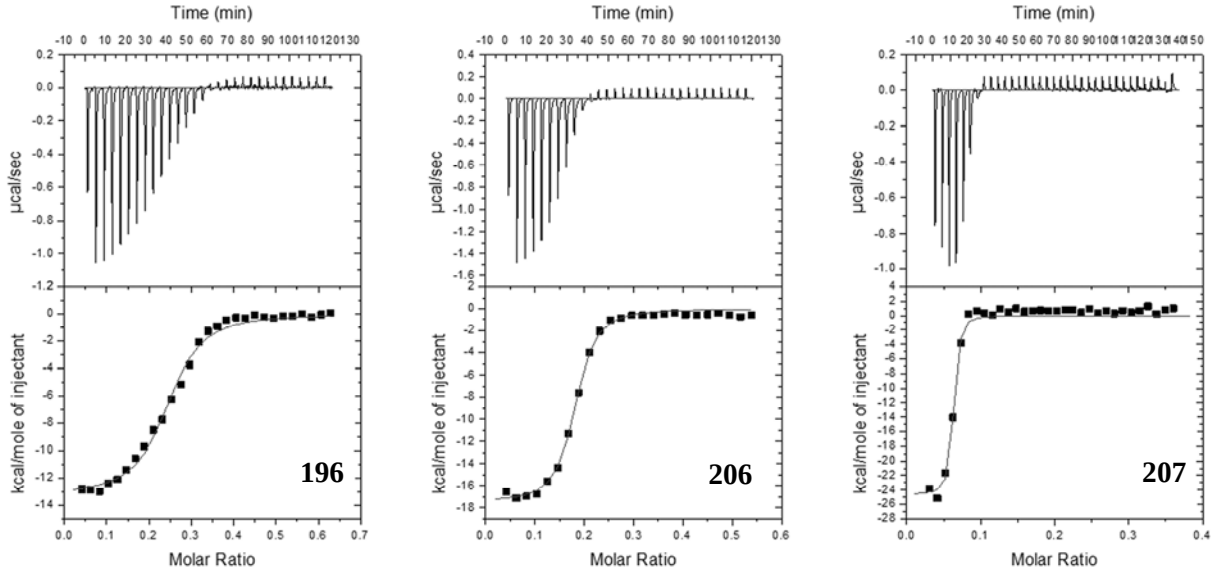


Figure C. 56. *Top: ITC experiments were carried out in sodium cacodylate buffer (0.01 M) at neutral pH by titrating aliquots of 5 µL of* ***196***, ***206*** *(c = 0.6 mM), or* ***207*** *(c = 0.3 mM) to pdA × pdT (c = 0.1 mM). Bottom: the titrations were corrected for dilution and fitted with a one-site model. The titrations with* ***196*** *and* ***206*** *are corrected for the first, and the one for* ***207*** *for the first two data points.*

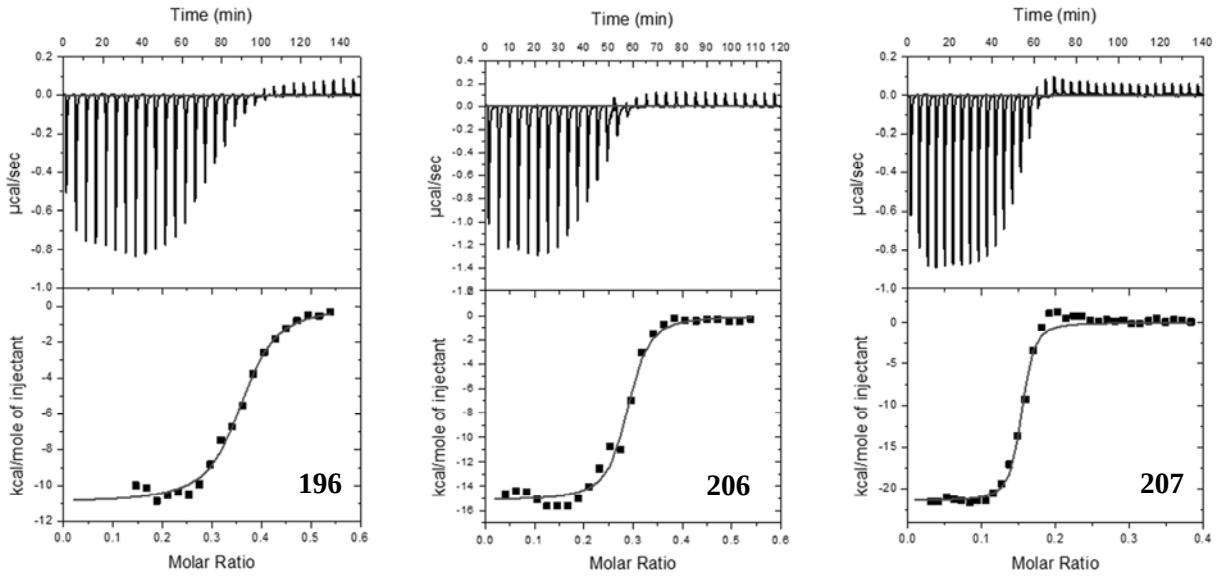


Figure C. 57. *Top: ITC experiments were carried out in sodium cacodylate buffer (0.01 M) at neutral pH by titrating aliquots of 5 µL of* ***196***, ***206*** *(c = 0.6 mM), or* ***207*** *(c = 0.3 mM) to $p(dGdC)_2$ (c = 0.1 mM). Bottom: the titrations were corrected for dilution and fitted with a one-site model. The titration with* ***196*** *is corrected for the first five, and the ones for* ***206*** *and* ***207*** *for the first data point.*

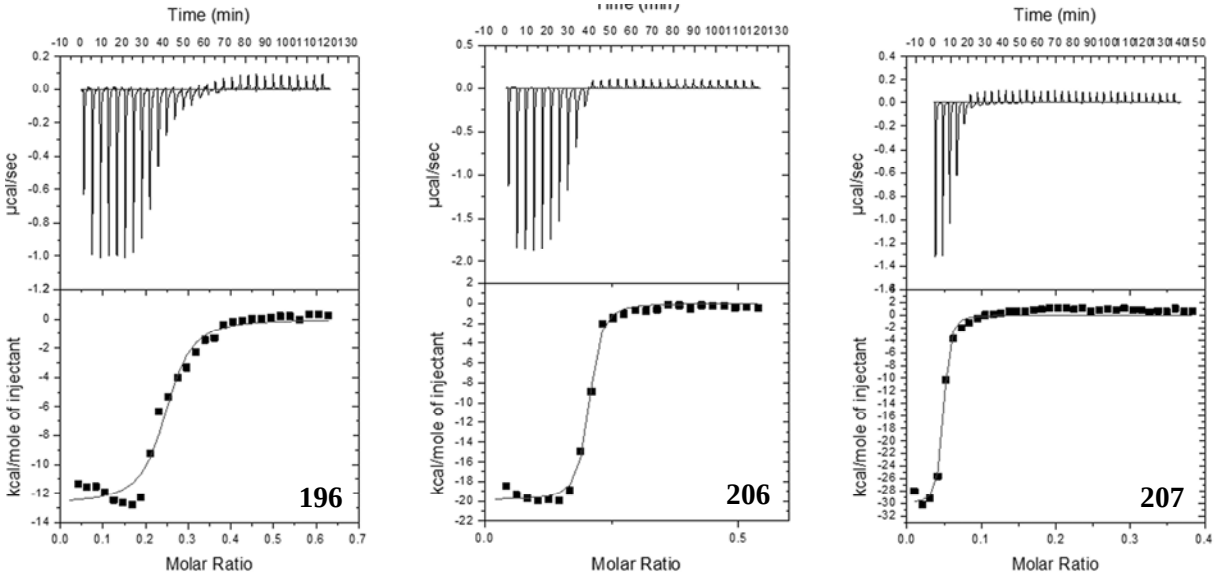


Figure C. 58. *Top: ITC experiments were carried out in sodium cacodylate buffer (0.01 M) at neutral pH by titrating aliquots of 5 µL of* ***196***, ***206*** *(c = 0.6 mM), or* ***207*** *(c = 0.3 mM) to pdG × pdC (c = 0.1 mM). Bottom: the titrations were corrected for dilution and fitted with a one-site model. The titrations with* ***196***, ***206***, *and* ***207*** *are corrected for the first data point.*

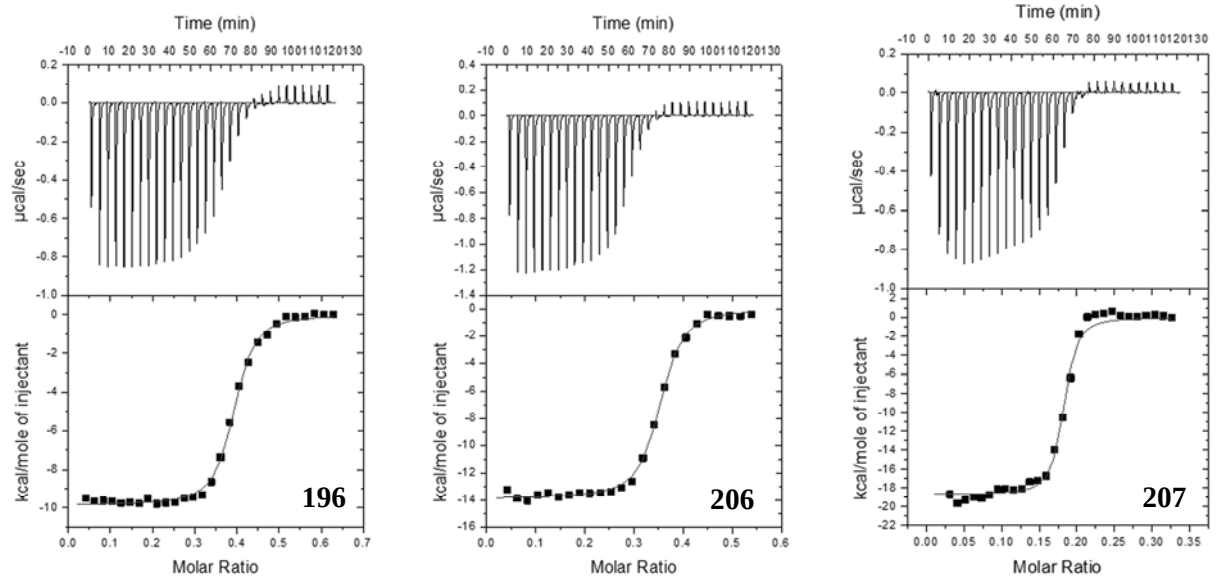


Figure C. 59. *Top: ITC experiments were carried out in sodium cacodylate buffer (0.01 M) at neutral pH by titrating aliquots of 5 µL of* ***196***, ***206*** *(c = 0.6 mM), or* ***207*** *(c = 0.3 mM) to pA × pU (c = 0.1 mM). Bottom: the titrations were corrected for dilution and fitted with a one-site model. The titrations with* ***196*** *and* ***206*** *are corrected for the first, and the one for* ***207*** *for the first two data points.*

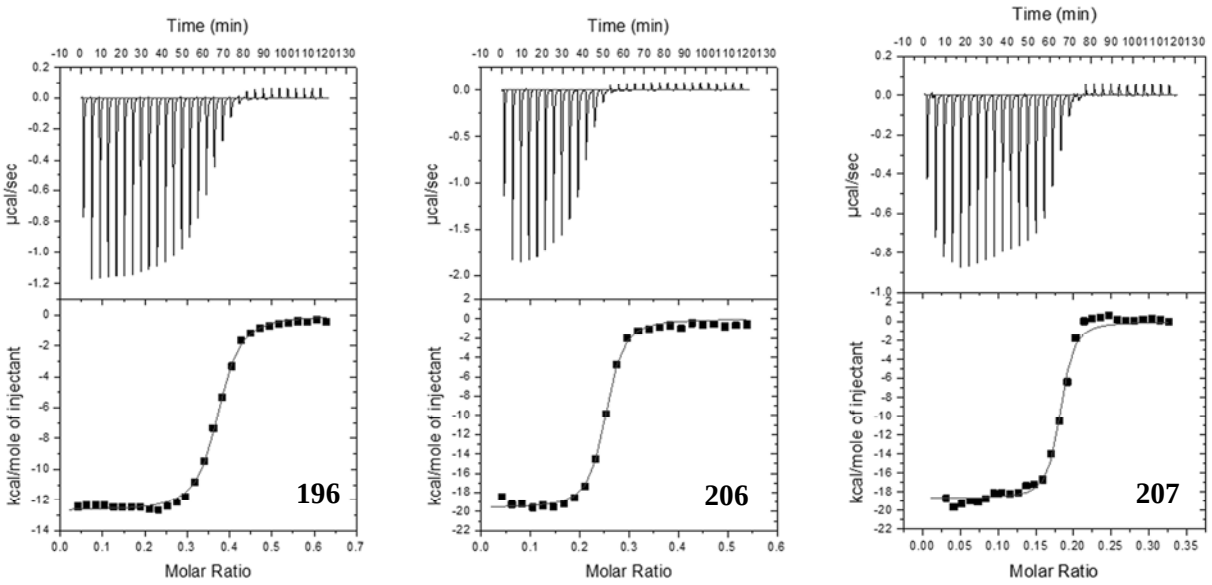


Figure C. 60. *Top: ITC experiments were carried out in sodium cacodylate buffer (0.01 M) at neutral pH by titrating aliquots of 5 µL of* ***196***, ***206*** *(c = 0.6 mM), or* ***207*** *(c = 0.3 mM) to pG × pC (c = 0.1 mM). Bottom: the titrations were corrected for dilution and fitted with a one-site model. The titration with* ***196*** *is corrected for the first five, and the ones for* ***206*** *and* ***207*** *for the first data point.*

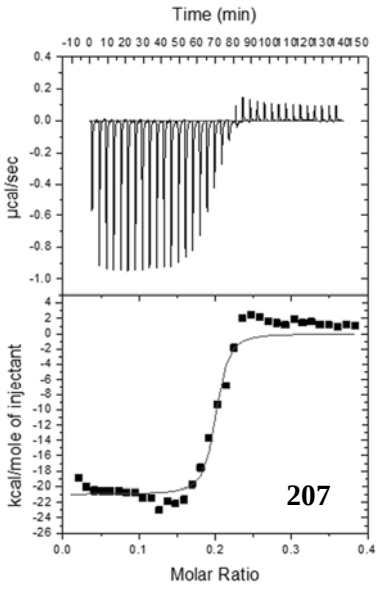


Figure C. 61. *Top: ITC experiments were carried out in sodium cacodylate buffer (0.01 M) at neutral pH by titrating aliquots of 5 µL of* ***196***, ***206*** *(c = 0.6 mM), or* ***207*** *(c = 0.3 mM) to p(dAdT)$_2$ (c = 0.1 mM). Bottom: the titrations were corrected for dilution and fitted with a one-site model. The titrations with* ***196*** *and* ***206*** *are corrected for the first six, and the one for* ***207*** *for the first data point.*

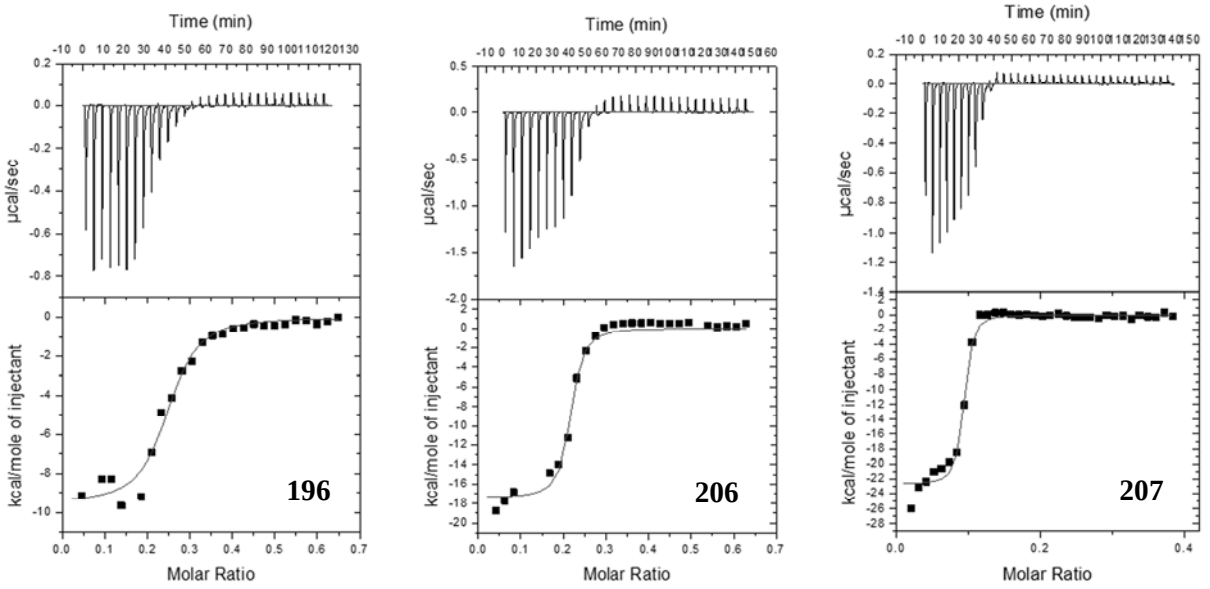


Figure C. 62. *Top: ITC experiments were carried out in sodium cacodylate buffer (0.01 M) at neutral pH by titrating aliquots of 5 µL of* ***196***, ***206*** *(c = 0.6 mM), or* ***207*** *(c = 0.3 mM) to pF143-GFP (****196****: 0.09 mM;* ***206***, ***207****: 0.1 mM). Bottom: the titrations were corrected for dilution and fitted with a one-site model. The titration with* ***196*** *is corrected for the first, third, and seventh; the one for* ***206*** *for the first, fifth, sixth, and seventh; and the one for* ***207*** *for the first data point.*

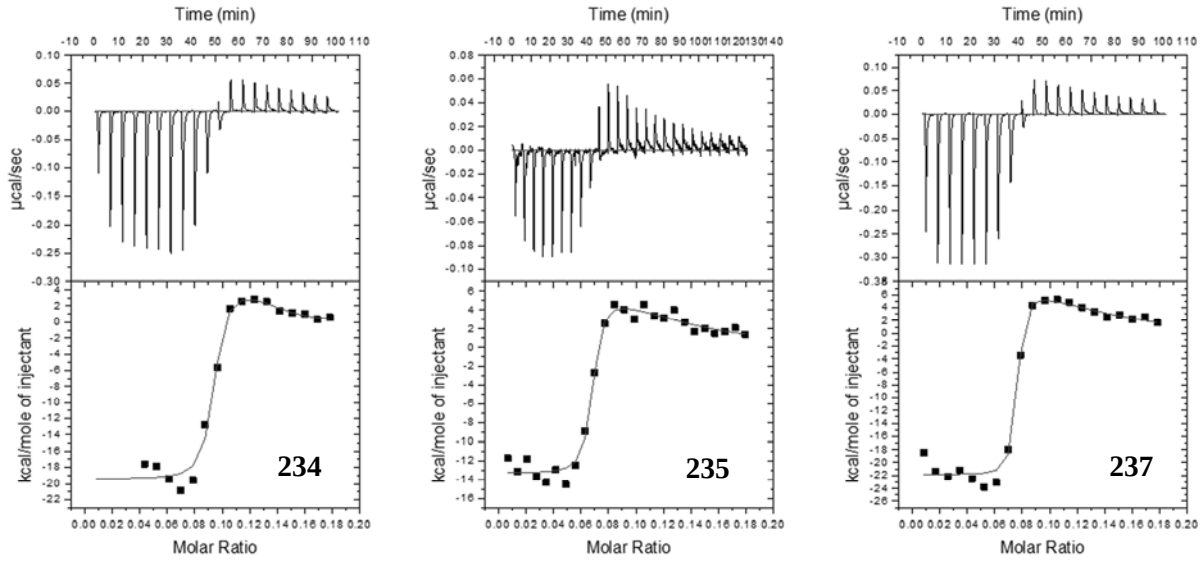


Figure C. 63. *Top: ITC experiments were carried out in sodium cacodylate buffer (0.01 M) at neutral pH by titrating aliquots of 5 µL of* ***234***, ***235*** *(c = 0.1 mM), or* ***237*** *(c = 0.2 mM) to p(dAdT)$_2$ (0.04 mM). Bottom: the titrations were corrected for dilution and fitted with a two-site model for* ***234***, ***235*** *and a one-site model for* ***237****. The titration with* ***234*** *is corrected for the first three; and the one for* ***237*** *for the first two data points.*

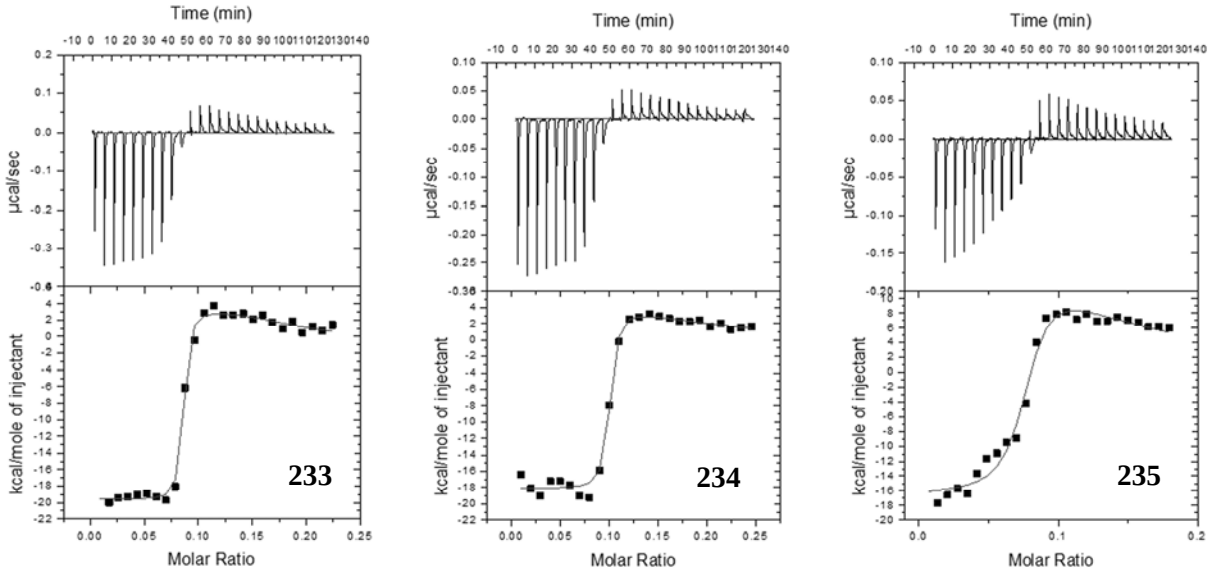


Figure C. 64. *Top: ITC experiments were carried out in sodium cacodylate buffer (0.01 M) at neutral pH by titrating aliquots of 5 µL of* ***233*** *(c = 0.125 mM),* ***234*** *(c = 0.1 mM), or* ***235*** *(c = 0.06 mM) to pF143-GFP (****233****: 0.05 mM;* ***234****: 0.035 mM;* ***235****: 0.03 mM). Bottom: the titrations were corrected for dilution and fitted with a two-site model. The titrations are corrected for the first data point.*

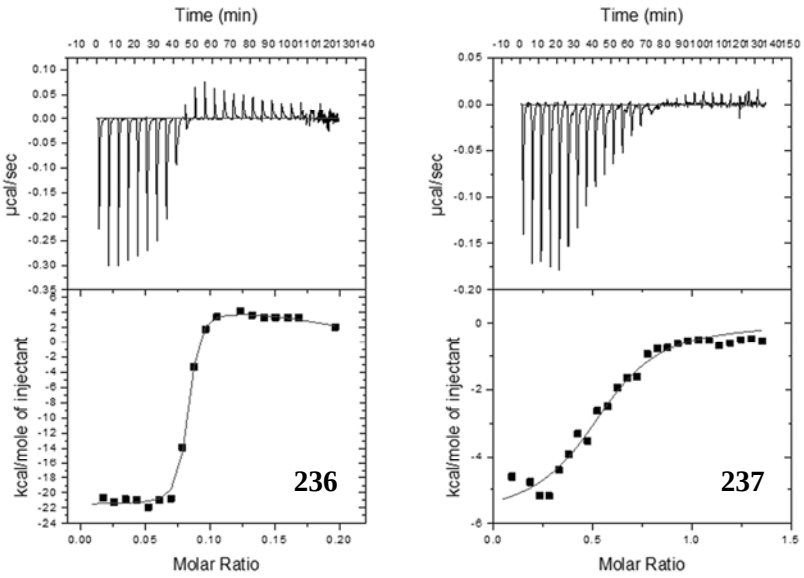


Figure C. 65. *Top: ITC experiments were carried out in sodium cacodylate buffer (0.01 M) at neutral pH by titrating aliquots of 5 µL of* ***236****(c = 0.1 mM) or* ***237*** *(c = 0.2 mM) to pF143-GFP (****236****: 0.04 mM;* ***237****: 0.03 mM). Bottom: the titrations were corrected for dilution and fitted with a two-site model for* ***236*** *and a one-site model for* ***237****. The titration of* ***236*** *is corrected for the first, 13th, 20th, 21st, 23rd, 24th, 25th data points.*

C.6 DAPI Displacement Assays

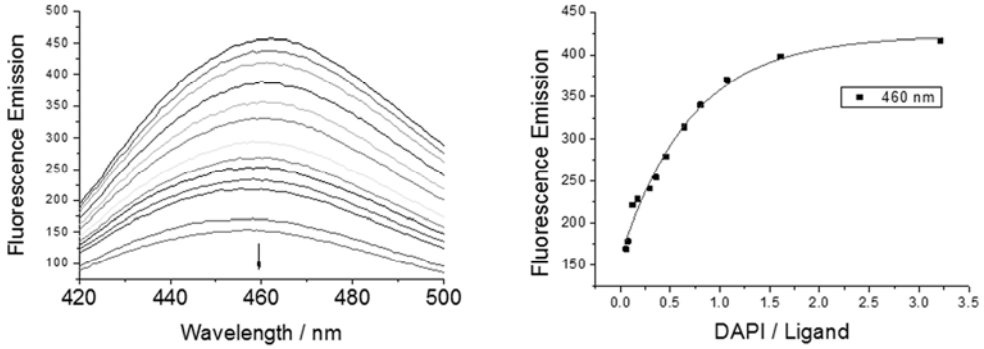


Figure C. 66. *DAPI displacement experiment with pdA × pdT and* ***196*** *(left) and excerpt of the fluorescence emission at 460 nm plotted against the ratio of DAPI /****196*** *with the corresponding fit (right).*

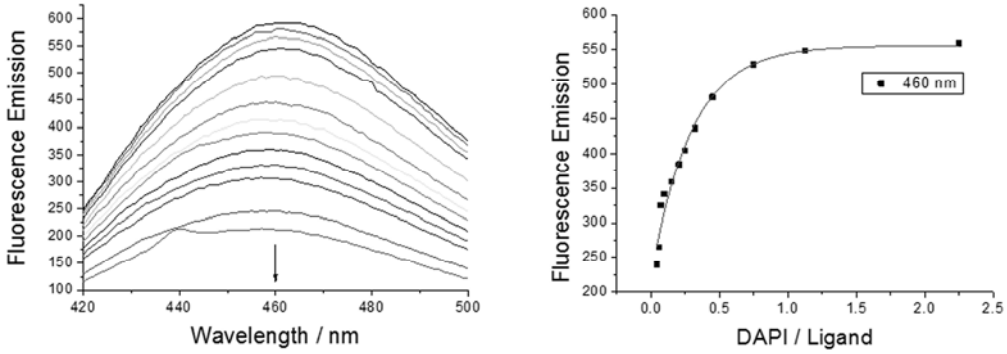


Figure C. 67. *DAPI displacement experiment with pdA × pdT and* ***206*** *(left) and excerpt of the fluorescence emission at 460 nm plotted against the ratio of DAPI /****206*** *with the corresponding fit (right).*

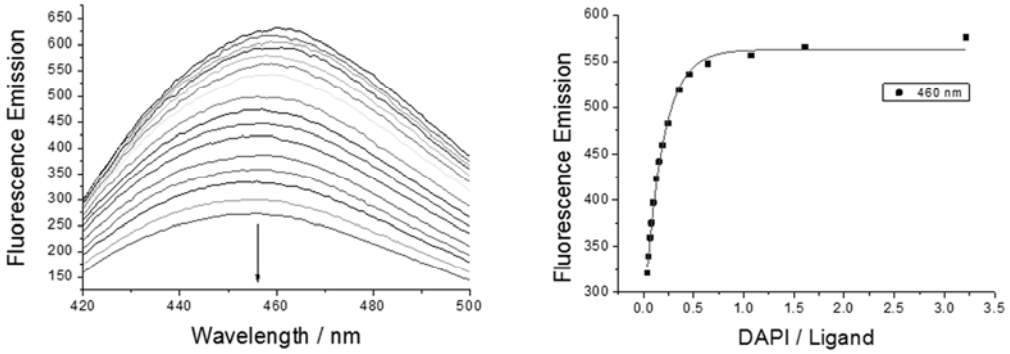


Figure C. 68. *DAPI displacement experiment with p(dAdT)$_2$ and* ***196*** *(left) and excerpt of the fluorescence emission at 460 nm plotted against the ratio of DAPI /****196*** *with the corresponding fit (right).*

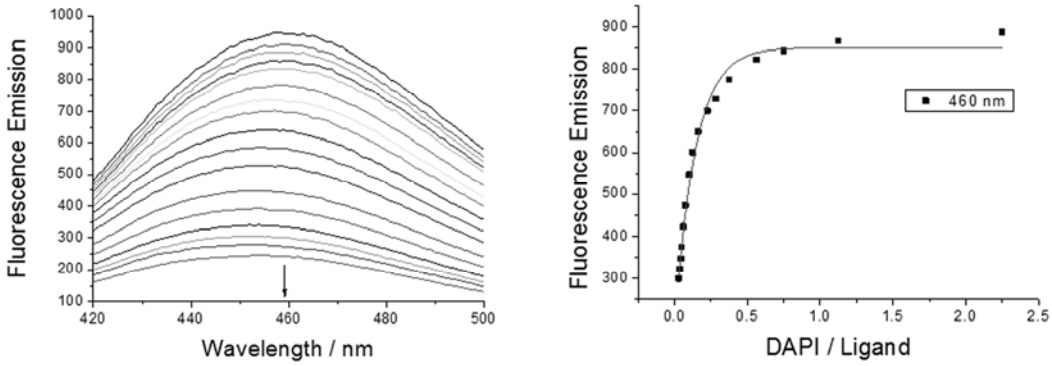


Figure C. 69. *DAPI displacement experiment with p(dAdT)$_2$ and* ***206*** *(left) and excerpt of the fluorescence emission at 460 nm plotted against the ratio of DAPI /****206*** *with the corresponding fit (right).*

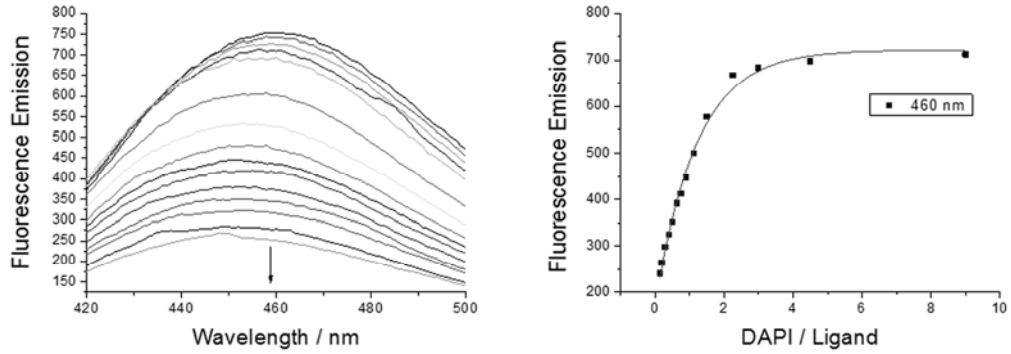


Figure C. 70. *DAPI displacement experiment with p(dAdT)$_2$ and **207** (left) and excerpt of the fluorescence emission at 455 nm plotted against the ratio of DAPI /**207** with the corresponding fit (right).*

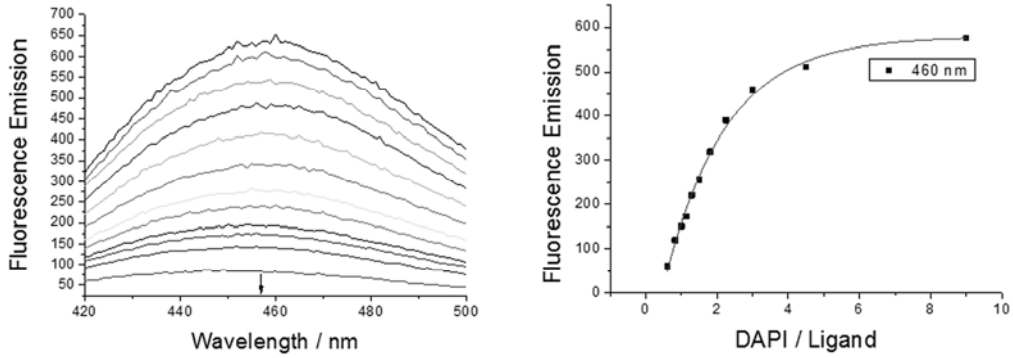


Figure C. 71. *DAPI displacement experiment with p(dAdT)$_2$ and **234** (left) and excerpt of the fluorescence emission at 455 nm plotted against the ratio of DAPI /**234** with the corresponding fit (right).*

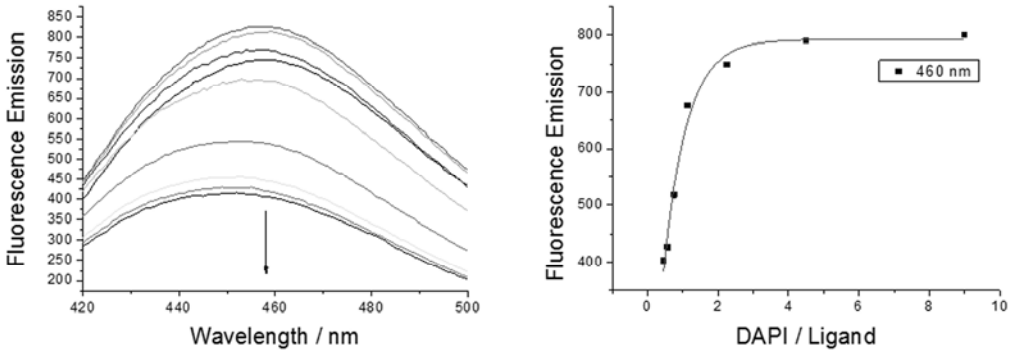


Figure C. 72. *DAPI displacement experiment with p(dAdT)$_2$ and **235** (left) and excerpt of the fluorescence emission at 455 nm plotted against the ratio of DAPI /**235** with the corresponding fit (right).*

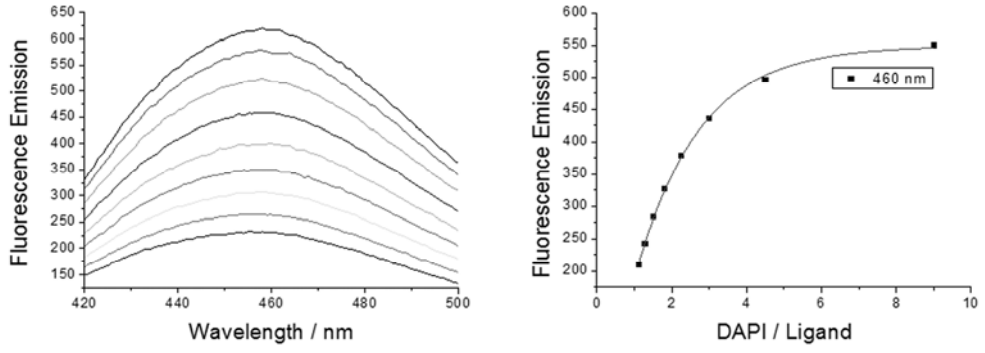


Figure C. 73. *DAPI displacement experiment with p(dAdT)$_2$ and **236** (left) and excerpt of the fluorescence emission at 455 nm plotted against the ratio of DAPI /**236** with the corresponding fit (right).*

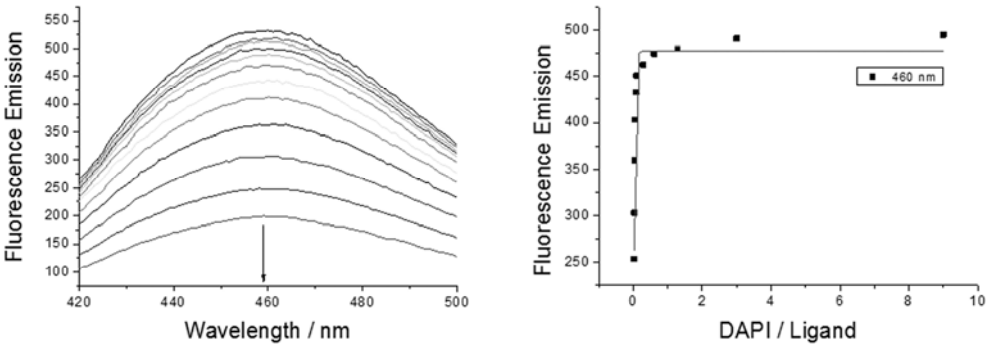


Figure C. 74. *DAPI displacement experiment with p(dAdT)$_2$ and* ***237*** *(left) and excerpt of the fluorescence emission at 455 nm plotted against the ratio of DAPI /****237*** *with the corresponding fit (right).*

C.7 On-Bead Screening of the Library

Screening Results for $p(dAdT)_2$

Table C. 1. *Screening results of the combinatorial library* ***220*** *normalized to 100 %. The DNA ligands are listed in descending order of their ability to displace DAPI from* $p(dAdT)_2$.

#	AA[1]	AA[2]	AA[3]	Remaining DAPI [%]	Archive
1	Trp	Lys	Arg	26	2-E07
2		Arg	Lys	31	3-C02
3	Arg	Arg	Lys	32	1-A02
4	Trp	Arg	Lys	33	2-E02
5	Arg	Lys	Lys	34	1-A08
6	Arg	Arg	Arg	35	1-A01
7	Tyr	Arg	Arg	40	2-B01
8	Arg	Arg	Tyr	41	1-A04
9	Phe	Tyr	Lys	42	1-H08
10		Arg	Arg	42	3-C01
11	Arg	Arg	Trp	43	1-A05
12	Lys	Arg	Trp	44	1-D05
13	Phe	Tyr	Arg	44	1-H07
14	Trp	Lys	Lys	44	2-E08
15	Phe	Phe	Lys	45	1-H02
16	Lys	Lys	Lys	45	1-D08
17	Phe	Lys	Arg	46	1-G07
18	Arg	Lys	Phe	47	1-A09
19	Arg	Trp	Lys	47	1-C02
20	Glu	Lys	Lys	48	2-H08
21	Tyr	Arg	Lys	50	2-B02
22	Tyr	Lys	Lys	51	2-B08
23	Tyr	Tyr	Phe	51	2-C09
24	Lys	Arg	Phe	51	1-D03
25		Phe	Arg	51	3-D01
26	Glu	Arg	Lys	52	2-H02
27	Lys	Lys	Arg	52	1-D07
28	Lys	Tyr	Arg	53	1-E07
29	Tyr	Lys	Tyr	53	2-B10
30	Arg	Lys	Arg	53	1-A07
31	Arg	Phe	Arg	54	1-B01
32		Lys	Lys	54	3-C08
33	Phe	Phe	Trp	54	1-H05
34	Phe	Phe	Glu	54	1-H06
35	Phe	Tyr	Phe	54	1-H09
36		Lys	Arg	55	3-C07
37	Arg	Lys	Tyr	55	1-A10
38	Tyr	Tyr	Arg	56	2-C07
39		Arg	Trp	57	3-C05
40	Phe	Glu	Tyr	58	2-A10
41	Glu	Glu	Trp	58	3-B11
42	Arg	Glu	Lys	58	1-C08
43	Lys	Arg	Arg	58	1-D01
44	Phe	Phe	Arg	58	1-H01
45		Arg	Phe	58	3-C03
46	Arg	Glu	Arg	58	1-C07
47	Glu	Arg	Tyr	59	2-H04
48		Trp	Arg	59	3-E01
49	Phe	Glu	Phe	59	2-A09
50		Lys	Trp	59	3-C11
51	Trp	Lys	Glu	60	2-E12

52	Arg	Lys	Trp	60	1-A11
53	Phe	Lys	Trp	60	1-G11
54	Phe	Lys	Lys	60	1-G08
55	Lys	Arg	Glu	61	1-D06
56	Phe	Lys	Tyr	61	1-G10
57	Lys	Tyr	Glu	61	1-E12
58	Trp	Glu	Trp	61	2-G11
59	Phe	Glu	Lys	61	2-A08
60	Trp	Lys	Phe	61	2-E09
61	Phe	Lys	Phe	61	1-G09
62	Tyr	Tyr	Glu	61	2-C12
63	Phe	Tyr	Trp	61	1-H11
64		Lys	Tyr	61	3-C10
65	Phe	Arg	Glu	62	1-G06
66	Glu	Arg	Glu	62	2-H06
67	Arg	Glu	Tyr	62	1-C10
68	Tyr	Phe	Glu	62	2-C06
69	Trp	Glu	Lys	62	2-G08
70	Trp	Glu	Tyr	62	2-G10
71	Trp	Tyr	Glu	62	2-F12
72		Lys	Phe	63	3-C09
73	Lys	Trp	Lys	63	1-F02
74	Glu	Glu	Glu	63	3-B12
75	Trp	Phe	Arg	63	2-F01
76	Tyr	Trp	Lys	63	2-D02
77	Glu	Lys	Arg	63	2-H07
78	Glu	Lys	Phe	63	2-H09
79	Arg	Tyr	Arg	64	1-B07
80	Arg	Tyr	Lys	64	1-B08
81	Arg	Arg	Phe	64	1-A03
82	Trp	Trp	Glu	64	2-G06
83	Trp	Glu	Phe	64	2-G09
84	Glu	Glu	Phe	64	3-B09
85	Trp	Arg	Tyr	64	2-E04
86	Tyr	Trp	Phe	64	2-D03
87	Glu	Trp	Phe	65	3-B03
88	Phe	Tyr	Glu	65	1-H12
89	Trp	Trp	Phe	65	2-G03
90	Trp	Phe	Phe	65	2-F03
91	Trp	Arg	Trp	65	2-E05
92	Trp	Lys	Trp	65	2-E11
93	Glu	Arg	Arg	65	2-H01
94	Lys	Trp	Arg	65	1-F01
95	Glu	Trp	Glu	65	3-B06
96	Trp	Phe	Lys	65	2-F02
97	Trp	Tyr	Arg	65	2-F07
98	Trp	Lys	Tyr	65	2-E10
99	Phe	Phe	Tyr	66	1-H04
100	Tyr	Arg	Tyr	66	2-B04
101	Lys	Lys	Trp	66	1-D11
102	Trp	Arg	Arg	66	2-E01
103	Lys	Lys	Glu	66	1-D12
104	Tyr	Phe	Trp	66	2-C05
105	Arg	Tyr	Trp	66	1-B11
106	Arg	Lys	Glu	67	1-A12
107	Phe	Arg	Tyr	67	1-G04
108	Phe	Tyr	Tyr	67	1-H10
109	Lys	Arg	Tyr	67	1-D04
110	Tyr	Tyr	Lys	67	2-C08

111	Lys	Phe	Tyr	67	1-E04
112	Phe	Trp	Trp	67	2-A05
113	Tyr	Tyr	Trp	67	2-C11
114	Arg	Phe	Glu	68	1-B06
115	Tyr	Glu	Tyr	68	2-D10
116	Phe	Trp	Arg	68	2-A01
117	Glu	Arg	Trp	68	2-H05
118	Arg	Trp	Trp	68	1-C05
119	Phe	Glu	Arg	68	2-A07
120	Lys	Trp	Phe	68	1-F03
121	Phe	Arg	Arg	68	1-G01
122	Tyr	Phe	Lys	68	2-C02
123				68	3-F07
124	Trp	Glu	Arg	68	2-G07
125	Lys	Phe	Lys	68	1-E02
126	Glu	Tyr	Glu	68	3-A12
127	Arg	Trp	Tyr	69	1-C04
128	Trp	Phe	Trp	69	2-F05
129	Trp	Trp	Trp	69	2-G05
130		Tyr	Arg	69	3-D07
131	Lys	Phe	Arg	69	1-E01
132	Phe	Arg	Lys	69	1-G02
133	Trp	Arg	Glu	69	2-E06
134			Glu	69	3-F06
135	Phe	Trp	Glu	69	2-A06
136	Arg	Trp	Glu	69	1-C06
137		Tyr	Lys	69	3-D08
138		Tyr	Tyr	69	3-D10
139	Trp	Tyr	Tyr	69	2-F10
140	Tyr	Arg	Glu	70	2-B06
141	Trp	Trp	Arg	70	2-G01
142	Trp	Glu	Glu	70	2-G12
143	Lys	Arg	Lys	70	1-D02
144	Phe	Lys	Glu	70	1-G12
145		Glu	Lys	70	3-E08
146		Trp	Glu	70	3-E06
147	Tyr	Tyr	Tyr	70	2-C10
148	Lys	Lys	Phe	70	1-D09
149		Glu	Arg	71	3-E07
150	Arg	Phe	Tyr	71	1-B04
151	Trp	Tyr	Trp	71	2-F11
152		Lys	Glu	71	3-C12
153		Phe	Trp	71	3-D05
154	Phe	Phe	Phe	71	1-H03
155	Arg	Tyr	Phe	71	1-B09
156	Tyr	Glu	Lys	71	2-D08
157	Arg	Tyr	Glu	71	1-B12
158		Glu	Phe	71	3-E09
159	Phe	Trp	Lys	71	2-A02
160	Tyr	Glu	Trp	71	2-D11
161	Trp	Tyr	Lys	71	2-F08
162	Phe	Arg	Trp	72	1-G05
163		Arg	Glu	72	3-C06
164	Lys	Tyr	Tyr	72	1-E10
165	Tyr	Phe	Phe	72	2-C03
166	Tyr	Glu	Phe	72	2-D09
167	Arg	Phe	Trp	72	1-B05
168	Tyr	Lys	Phe	72	2-B09
169	Arg	Glu	Phe	72	1-C09

170	Lys	Glu	Glu	72	1-F12
171		Glu	Trp	72	3-E11
172	Tyr	Trp	Trp	72	2-D05
173	Tyr	Glu	Arg	72	2-D07
174	Lys	Trp	Tyr	73	1-F04
175		Trp	Lys	73	3-E02
176	Glu	Arg	Phe	73	2-H03
177	Glu	Lys	Glu	73	2-H12
178	Tyr	Trp	Tyr	73	2-D04
179	Arg	Trp	Phe	73	1-C03
180	Phe	Glu	Glu	73	2-A12
181	Trp	Arg	Phe	73	2-E03
182	Tyr	Arg	Phe	73	2-B03
183	Trp	Phe	Glu	73	2-F06
184	Lys	Glu	Phe	73	1-F09
185	Tyr	Lys	Trp	74	2-B11
186	Arg	Tyr	Tyr	74	1-B10
187	Phe	Arg	Phe	74	1-G03
188	Trp	Phe	Tyr	74	2-F04
189	Glu	Phe	Trp	74	3-A05
190	Glu	Glu	Lys	74	3-B08
191	Tyr	Trp	Glu	74	2-D06
192	Lys	Tyr	Phe	74	1-E09
193	Lys	Lys	Tyr	74	1-D10
194		Phe	Lys	74	3-D02
195	Phe	Glu	Trp	74	2-A11
196		Arg	Tyr	74	3-C04
197	Tyr	Phe	Tyr	74	2-C04
198	Tyr	Lys	Glu	75	2-B12
199	Glu	Phe	Phe	75	3-A03
200	Tyr	Trp	Arg	75	2-D01
201			Arg	75	3-F05
202	Glu	Lys	Trp	75	2-H11
203	Tyr	Glu	Glu	75	2-D12
204	Arg	Trp	Arg	75	1-C01
205	Trp	Tyr	Phe	75	2-F09
206	Glu	Tyr	Tyr	75	3-A10
207	Lys	Tyr	Lys	75	1-E08
208			Lys	76	3-F04
209		Trp	Tyr	76	3-E04
210	Tyr	Phe	Arg	76	2-C01
211	Arg	Glu	Glu	76	1-C12
212	Glu	Phe	Lys	76	3-A02
213	Lys	Glu	Lys	76	1-F08
214		Glu	Glu	76	3-E12
215		Tyr	Glu	76	3-D12
216	Lys	Glu	Tyr	76	1-F10
217		Tyr	Phe	76	3-D09
218	Trp	Trp	Lys	77	2-G02
219	Arg	Phe	Phe	77	1-B03
220	Lys	Trp	Glu	77	1-F06
221	Lys	Trp	Trp	77	1-F05
222	Arg	Phe	Lys	77	1-B02
223		Tyr	Trp	78	3-D11
224		Trp	Phe	78	3-E03
225		Phe	Phe	78	3-D03
226	Lys	Phe	Glu	78	1-E06
227	Lys	Phe	Phe	79	1-E03
228	Tyr	Arg	Trp	79	2-B05

229	Phe	Trp	Phe	79	2-A03
230	Trp	Trp	Tyr	79	2-G04
231	Lys	Phe	Trp	80	1-E05
232			Phe	80	3-F03
233	Glu	Lys	Tyr	80	2-H10
234	Glu	Trp	Arg	80	3-B01
235			Trp	81	3-F01
236			Tyr	81	3-F02
237	Glu	Glu	Tyr	82	3-B10
238	Lys	Glu	Trp	82	1-F11
239	Arg	Glu	Trp	82	1-C11
240		Glu	Tyr	82	3-E10
241	Glu	Tyr	Lys	83	3-A08
242	Glu	Tyr	Arg	83	3-A07
243	Lys	Tyr	Trp	83	1-E11
244	Glu	Trp	Trp	84	3-B05
245		Phe	Glu	86	3-D06
246	Glu	Glu	Arg	86	3-B07
247	Tyr	Lys	Arg	87	2-B07
248	Glu	Phe	Arg	87	3-A01
249	Phe	Trp	Tyr	88	2-A04
250	Arg	Arg	Glu	89	1-A06
251		Trp	Trp	89	3-E05
252	Glu	Tyr	Phe	89	3-A09
253	Lys	Glu	Arg	89	1-F07
254		Phe	Tyr	91	3-D04
255	Glu	Phe	Glu	93	3-A06
256	Glu	Tyr	Trp	93	3-A11
257	Glu	Phe	Tyr	94	3-A04
258	Glu	Trp	Lys	98	3-B02
259	Glu	Trp	Tyr	100	3-B04

Screening Results for pF143-GFP plasmid DNA

Table C. 2. *Screening results of the combinatorial library* ***220*** *normalized to 100 %. The DNA ligands are listed in descending order of their ability to displace EB from pF143-GFP plasmid DNA.*

#	AA[1]	AA[2]	AA[3]	Remaining EB [%]	Archive
1	Trp	Arg	Lys	9	2-E02
2	Trp	Lys	Arg	13	2-E07
3	Lys	Phe	Arg	15	1-E01
4	Lys	Lys	Phe	16	1-D09
5		Arg	Tyr	18	3-C04
6	Arg	Lys	Tyr	18	1-A10
7	Arg	Tyr	Arg	19	1-B07
8	Arg	Trp	Arg	19	1-C01
9	Arg	Phe	Arg	19	1-B01
10		Arg	Arg	19	3-C01
11	Arg	Lys	Trp	20	1-A11
12		Arg	Lys	20	3-C02
13	Lys	Arg	Tyr	20	1-D04
14	Arg	Phe	Tyr	21	1-B04
15	Arg	Lys	Lys	21	1-A08
16	Trp	Lys	Lys	21	2-E08
17	Trp	Arg	Tyr	21	2-E04
18	Lys	Arg	Phe	22	1-D03
19	Arg	Phe	Lys	22	1-B02
20	Trp	Arg	Phe	22	2-E03
21	Tyr	Tyr	Trp	23	2-C11
22		Lys	Tyr	24	3-D10
23	Lys	Lys	Tyr	24	1-D10
24	Phe	Trp	Lys	24	2-A02
25	Trp	Arg	Arg	24	2-B01
26	Lys	Tyr	Arg	24	1-E07
27	Lys	Lys	Glu	24	1-D12
28	Lys	Arg	Lys	25	1-D02
29	Arg	Arg	Lys	25	1-A02
30	Arg	Tyr	Trp	25	1-B11
31	Phe	Lys	Arg	26	1-G07
32	Tyr	Phe	Lys	26	2-C02
33	Tyr	Lys	Phe	26	2-B09
34	Arg	Tyr	Glu	26	1-B12
35	Tyr	Arg	Arg	26	2-B01
36	Arg	Arg	Phe	26	1-A03
37	Tyr	Lys	Arg	27	2-B07
38	Lys	Lys	Trp	27	1-D11
39	Arg	Trp	Tyr	27	1-C04
40	Arg	Tyr	Tyr	27	1-B10
41	Arg	Trp	Glu	27	1-C06
42	Phe	Phe	Arg	27	1-H01
43	Arg	Arg	Tyr	27	1-A04
44	Tyr	Arg	Phe	27	2-B03
45	Arg	Phe	Trp	27	1-B05
46	Lys	Arg	Glu	27	1-D06
47	Phe	Trp	Arg	28	2-A01
48	Tyr	Phe	Arg	28	2-C01
49	Arg	Glu	Glu	28	1-C12
50	Tyr	Arg	Trp	28	2-B05
51	Arg	Tyr	Lys	28	1-B08
52	Tyr	Phe	Tyr	28	2-C04
53	Tyr	Phe	Trp	28	2-C05

54	Tyr	Tyr	Lys	28	2-C08
55	Trp	Arg	Trp	28	2-E05
56	Phe	Trp	Tyr	28	2-A04
57	Arg	Phe	Phe	28	1-B03
58	Lys	Lys	Arg	28	1-D07
59	Arg	Glu	Phe	29	1-C09
60	Tyr	Arg	Tyr	29	2-B04
61	Tyr	Tyr	Tyr	29	2-C10
62	Phe	Trp	Phe	29	2-A03
63	Arg	Arg	Glu	29	1-A06
64		Arg	Trp	29	3-C05
65	Tyr	Lys	Trp	29	2-B11
66	Arg	Glu	Trp	29	1-C11
67	Glu	Arg	Lys	29	2-H02
68	Arg	Trp	Trp	29	1-C05
69	Lys	Lys	Lys	30	1-D08
70	Phe	Tyr	Tyr	30	1-H10
71	Tyr	Lys	Glu	31	2-B12
72			Arg	31	3-F05
73	Phe	Tyr	Arg	31	1-H07
74	Tyr	Phe	Phe	32	2-C03
75	Tyr	Arg	Glu	32	2-B06
76	Glu	Lys	Lys	32	2-H08
77	Phe	Tyr	Glu	32	1-H12
78	Phe	Trp	Trp	32	2-A05
79		Lys	Lys	33	3-C08
80	Trp	Lys	Tyr	33	2-E10
81	Lys	Phe	Lys	33	1-E02
82		Tyr	Arg	33	3-D07
83	Tyr	Lys	Tyr	33	2-B10
84	Phe	Lys	Lys	33	1-G08
85	Lys	Tyr	Phe	33	1-E09
86		Lys	Arg	33	3-C07
87			Lys	34	3-F04
88	Arg	Arg	Trp	34	1-A05
89	Lys	Trp	Lys	34	1-F02
90	Arg	Glu	Arg	34	1-C07
91	Phe	Glu	Glu	34	2-A12
92	Phe	Phe	Tyr	34	1-H04
93	Tyr	Trp	Lys	34	2-D02
94	Arg	Lys	Phe	34	1-A09
95	Phe	Arg	Tyr	35	1-G04
96	Arg	Glu	Lys	35	1-C08
97	Phe	Lys	Glu	35	1-G12
98	Lys	Arg	Arg	35	1-D01
99	Trp	Trp	Arg	35	2-G01
100	Phe	Tyr	Lys	35	1-H08
101	Glu	Lys	Arg	35	2-H07
102	Phe	Lys	Tyr	35	1-G10
103	Phe	Lys	Phe	35	1-G09
104	Tyr	Arg	Lys	35	2-B02
105		Tyr	Lys	36	3-D08
106		Phe	Trp	36	3-D05
107	Phe	Lys	Trp	36	1-G11
108	Arg	Lys	Arg	36	1-A07
109	Arg	Trp	Lys	36	1-C02
110	Lys	Tyr	Glu	36	1-E12
111	Phe	Glu	Arg	36	2-A07
112	Trp	Arg	Glu	37	2-E06

113	Phe	Phe	Trp	37	1-H05
114	Arg	Arg	Arg	37	1-A01
115	Lys	Arg	Trp	37	1-D05
116	Lys	Glu	Trp	37	1-F11
117	Glu	Arg	Arg	37	2-H01
118	Glu	Phe	Arg	37	3-A01
119	Phe	Phe	Lys	37	1-H02
120	Glu	Trp	Lys	37	3-B02
121	Lys	Glu	Arg	37	1-F07
122	Tyr	Tyr	Arg	37	2-C07
123	Glu	Tyr	Arg	37	3-A07
124	Trp	Trp	Tyr	38	2-G04
125	Lys	Phe	Tyr	38	1-E04
126	Lys	Phe	Phe	38	1-E03
127	Lys	Phe	Glu	38	1-E06
128	Glu	Phe	Lys	38	3-A02
129		Tyr	Trp	38	3-D11
130		Phe	Arg	38	3-D01
131	Glu	Lys	Phe	38	2-H09
132	Lys	Glu	Phe	38	1-F09
133	Tyr	Tyr	Phe	38	2-C09
134		Lys	Glu	38	3-C12
135	Trp	Lys	Trp	38	2-E11
136		Lys	Phe	38	3-C09
137	Tyr	Lys	Lys	39	2-B08
138		Trp	Arg	39	3-E01
139	Lys	Trp	Arg	39	1-F01
140	Trp	Glu	Arg	39	2-G07
141	Glu	Lys	Trp	39	2-H11
142	Glu	Arg	Tyr	39	2-H04
143	Glu	Arg	Phe	39	2-H03
144	Lys	Trp	Tyr	39	1-F04
145	Phe	Arg	Glu	39	1-G06
146	Glu	Arg	Trp	39	2-H05
147	Phe	Arg	Trp	39	1-G05
148		Phe	Lys	39	3-D02
149			Phe	39	3-F03
150		Arg	Glu	39	3-C06
151	Lys	Glu	Tyr	40	1-F10
152	Phe	Phe	Phe	40	1-H03
153	Phe	Tyr	Trp	40	1-H11
154	Glu	Lys	Tyr	40	2-H10
155		Lys	Trp	40	3-C11
156	Lys	Tyr	Lys	40	1-E08
157	Glu	Glu	Glu	40	3-B12
158	Lys	Tyr	Tyr	40	1-E10
159		Trp	Tyr	40	3-E04
160	Trp	Lys	Glu	40	2-E12
161	Lys	Phe	Trp	40	1-E05
162	Glu	Phe	Glu	40	3-A06
163	Lys	Glu	Glu	40	1-F12
164		Phe	Phe	40	3-D03
165	Lys	Glu	Lys	40	1-F08
166	Trp	Trp	Phe	40	2-G03
167	Phe	Arg	Phe	40	1-G03
168		Arg	Phe	41	3-C03
169	Tyr	Phe	Glu	41	2-C06
170	Lys	Trp	Trp	41	1-F05
171	Glu	Trp	Arg	41	3-B01

172	Trp	Trp	Lys	41	2-G02
173	Glu	Tyr	Lys	41	3-A08
174	Lys	Tyr	Trp	41	1-E11
175	Trp	Lys	Phe	41	2-E09
176			CBS	41	3-F07
177		Phe	Tyr	42	3-D04
178		Tyr	Phe	42	3-D09
179	Glu	Glu	Arg	42	3-B07
180	Phe	Tyr	Phe	42	1-H09
181	Lys	Trp	Phe	42	1-F03
182	Trp	Phe	Tyr	42	2-F04
183		Trp	Lys	42	3-E02
184		Glu	Arg	42	3-E07
185	Phe	Glu	Lys	43	2-A08
186	Trp	Trp	Trp	43	2-G05
187	Tyr	Glu	Arg	43	2-D07
188	Arg	Tyr	Phe	43	1-B09
189	Lys	Trp	Glu	43	1-F06
190	Tyr	Trp	Arg	43	2-D01
191		Glu	Lys	44	3-E08
192	Trp	Phe	Arg	44	2-F01
193	Tyr	Glu	Lys	44	2-D08
194			Tyr	44	3-F02
195		Trp	Phe	44	3-E03
196	Arg	Lys	Glu	45	1-A12
197		Trp	Trp	45	3-E05
198			Trp	45	3-F01
199	Glu	Lys	Glu	46	2-H12
200		Glu	Glu	46	3-E12
201	Phe	Arg	Lys	46	1-G02
202	Phe	Trp	Glu	48	2-A06
203			Glu	48	3-F06
204	Tyr	Trp	Trp	48	2-D05
205	Glu	Glu	Lys	48	3-B08
206	Tyr	Trp	Tyr	48	2-D04
207	Phe	Glu	Phe	49	2-A09
208	Arg	Phe	Glu	49	1-B06
209	Trp	Glu	Glu	49	2-G12
210	Trp	Tyr	Arg	49	2-F07
211	Trp	Tyr	Phe	50	2-F09
212	Tyr	Tyr	Glu	50	2-C12
213	Glu	Arg	Glu	50	2-H06
214		Tyr	Tyr	50	3-D10
215	Phe	Glu	Tyr	50	2-A10
216	Glu	Tyr	Glu	51	3-A12
217	Arg	Trp	Phe	51	1-C03
218	Tyr	Trp	Phe	51	2-D03
219	Phe	Arg	Arg	52	1G01
220	Glu	Glu	Trp	52	3-B11
221		Glu	Phe	53	3-E09
222	Phe	Glu	Trp	53	2-A11
223	Trp	Phe	Phe	54	2-F03
224	Glu	Phe	Phe	55	3-A03
225	Arg	Glu	Tyr	56	1-C10
226	Glu	Trp	Glu	56	3-B06
227	Glu	Glu	Tyr	56	3-B10
228	Glu	Trp	Tyr	60	3-B04
229	Trp	Phe	Lys	61	2-F02
230	Trp	Tyr	Glu	61	2-F12

231	Trp	Tyr	Tyr	61	2-F10
232	Glu	Glu	Phe	62	3-B09
233	Trp	Phe	Trp	62	2-F05
234	Trp	Tyr	Lys	62	2-F08
235	Glu	Tyr	Tyr	64	3-A10
236		Glu	Trp	65	3-E11
237	Trp	Tyr	Trp	65	2-F11
238	Tyr	Glu	Phe	65	2-D09
239		Phe	Glu	65	3-D06
240	Tyr	Glu	Glu	65	2-D12
241		Trp	Glu	68	3-E06
242	Phe	Phe	Glu	69	1-H06
243		Tyr	Glu	69	3-D12
244	Glu	Trp	Phe	70	3-B03
245	Glu	Tyr	Phe	71	3-A09
246	Trp	Glu	Lys	72	2-G08
247	Trp	Glu	Tyr	73	2-G10
248	Trp	Trp	Glu	74	2-G06
249	Trp	Glu	Trp	78	2-G11
250	Glu	Trp	Trp	79	3-B05
251	Glu	Phe	Tyr	81	3-A04
252	Tyr	Glu	Trp	83	2-D11
253		Glu	Tyr	84	3-E10
254	Glu	Phe	Trp	85	3-A05
255	Glu	Tyr	Trp	86	3-A11
256	Trp	Glu	Phe	87	2-G09
257	Tyr	Glu	Tyr	88	2-D10
258	Trp	Phe	Glu	99	2-F06
259	Tyr	Trp	Glu	100	2-D06

D: CURRICULUM VITAE

The CV is not part of this online version of the dissertation.

E: LIST OF PUBLICATIONS

Journal (Peer-Reviewed)

1. H. Y. Kuchelmeister, A. Gutschmidt, S. Knauer, C. Schmuck; A surprisingly small trivalent dipeptide ligand allows efficient gene delivery into cells; **2011**, *submitted*.
2. M. Gellner, S. Niebling, H. Y. Kuchelmeister, C. Schmuck, S. Schlücker; Plasmonically active micron-sized beads for integrated solid phase synthesis and label-free SERS analysis; **2011**, *under revision*.
3. H. Y. Kuchelmeister, C. Schmuck; Nucleotide recognition in water by a guanidinium-based artificial tweezer receptor; *Chem. Eur. J.* **2011***, 17*, 5311-5318.
4. S. Niebling, H. Y. Kuchelmeister, C. Schmuck, S. Schlücker; Quantitative, label-free and site-specific monitoring of molecular recognition: A multivariate Resonance Raman approach; *Chem. Commun.* **2011**, *47*, 568-570.
5. H. Y. Kuchelmeister, C. Schmuck; An efficient synthesis of an orthogonally protected aromatic diamine as scaffold for tweezer receptors with two different arms; *Eur. J. Org. Chem.* **2009**, 4480-4485.

Book Chapter

1. C. Schmuck, H. Y. Kuchelmeister; Guanidinium based anion receptors; *Artificial receptors for chemical sensors*; Eds. V. M. Mirsky, A. Yatsimirsky, Wiley-VCH, **2011**, 273-317.

Poster Presentation

1. H. Y. Kuchelmeister, I. Piantanida, C. Schmuck; Molecular recognition of (poly)-nucleotides; *7th ERA Bioinspired Chemistry Conference*, Santiago de Compostella, Spain, **Jun. 2010**; *Industry Meets Science*, Essen, **Mar. 2011**.
2. M. Gellner, H. Y. Kuchelmeister, C. Schmuck, S. Schlücker; SERS and Solid-Phase Synthesis; *22nd International Conference on Raman Spectroscopy*, Boston, USA, **Aug. 2010**.
3. H. Y. Kuchelmeister; Comparison of different drugs for the treatment of retroviral diseases; *GDCh Project Management Training*, Münster, **Jun. 2010**.
4. H. Y. Kuchelmeister, C. Schmuck; Design and synthesis of peptide receptors; *JCF Spring Symposium,* Essen, **May 2009**; *1st Microwave Assisted Organic and Peptide Synthesis Conference,* Montpellier, France, **Jun. 2009**.
5. H. Y. Kuchelmeister, C. Schmuck; Combinatorial synthesis of peptide receptors; *2nd EuCheMS Chemistry Congress*, Turin, Italy, **Sep. 2008**; *International Summer School for Supramolecular Systems in Chemistry and Biology*, Tuapse, Russia, **Oct. 2008**.
6. O. Dopfer, H.-S. Andrei, H. Y. Kuchelmeister, N. Solca; IR spectrum of the ethyl cation: Nonclassical versus classical structure; *105th General Meeting of the German Bunsen-Society*

Oral Presentation

1. H. Y. Kuchelmeister; Molecular recognition of nucleotides and nucleic acids; *Symposium in Supramolecular Chemistry*, Essen, **Feb. 2011**.
2. H. Y. Kuchelmeister; Molecular recognition of biologically relevant substrates; *Ph.D. Symposium of the Studienstiftung des deutschen Volkes*, Düsseldorf, **Apr. 2010**.

F: BIBLIOGRAPHY

(1) Kim, D. H.; Park, J. The nature of interaction between the carboxylate of substrates and the guanidinium moiety of Arg-145 in carboxypeptidase A probed by inhibitors of the enzyme. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2967-2970.

(2) Luo, R.; David L.; Hung, H.; Devaney, J.; Gilson, M. K. Strength of solvent-exposed salt-bridges. *J. Phys. Chem. B* **1999**, *103*, 727-736.

(3) García-Pérez, M.; Pinto, M.; Subirana, J. A. Nonsequence-specific arginine interactions in the nucleosome core particle. *Biopolymers* **2003**, *69*, 432-439.

(4) Schmidtchen, F. P. Molekulare Wirte für Anionen. *Nachr. Chem. Tech. Lab.* **1988**, *36*, 8-17.

(5) Xia, S.; Konigsberg, W. H.; Wang, J. Hydrogen-bonding capability of a templating difluorotoluene nucleotide residue in an RB69 DNA polymerase ternary complex. *J. Am. Chem. Soc.* **2011**, *133*, 10003-10005.

(6) Oshovsky, G. V.; Reinhoudt, D. N.; Verboom, W. Supramolecular chemistry in water. *Angew. Chem. Int. Ed.* **2007**, *46*, 2366-2393.

(7) Pedersen, C. J. Cyclic polyethers and their complexes with metal salts. *J. Am. Chem. Soc.* **1967**, *89*, 7017-7036.

(8) Goldman, S.; Bates, R. G. Calculation of thermodynamic functions for ionic hydration. *J. Am. Chem. Soc.* **1972**, *94*, 1476-1484.

(9) Anslyn, E. V.; Hannon, C. L. The guanidinium group: Its biological role and synthetic analogs. In *Bioorganic chemistry frontiers*; Springer: Berlin, 1993; Bd. 3, 193-256.

(10) Inoue, Y.; Kuramitsu, S.; Inoue, K.; Kagamiyama, H.; Hiromi, K.; Tanase, S.; Morino, Y. Substitution of a lysyl residue for arginine 386 of Escherichia coli aspartate aminotransferase. *J. Biol. Chem.* **1989**, *264*, 9673-9681.

(11) Dawson, R. M. C.; Elliott, D. C.; Elliott, W. H.; Jones, K. M. *Data for biochemical research*; Clarendon Press: Oxford, 1959.

(12) Schmuck, C.; Kuchelmeister, H. Y. Guanidinium based anion receptors. In *Artificial receptors for chemical sensors*; Wiley-VCH: Weinheim, 2011, 273-317.

(13) Schmuck, C.; Heil, M.; Scheiber, J.; Baumann, K. Charge interactions do the job: A combined statistical and combinatorial approach to finding artificial receptors for binding tetrapeptides in water. *Angew. Chem. Int. Ed.* **2005**, *44*, 7208-7212.

(14) Mulligan, R. C. The basic science of gene therapy. *Science* **1993**, *260*, 926-932.

(15) Richards, J. E.; Hawley, S. *The human genome: a user´s guide*; Academic Press Inc.: London, 2010.

(16) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3-26.

(17) Bouard, D.; Alazard-Dany, N.; Cosset, F.-L. Viral vectors: From virology to transgene expression. *Brit. J. Pharmacol.* **2009**, *157*, 153-165.

(18) Mintzer, M. A.; Simanek, E. E. Nonviral vectors for gene delivery. *Chem. Rev.* **2009**, *109*, 259-302.

(19) Schmuck, C. Side chain selective binding of N-acetyl-α-amino acid carboxylates by a 2-(guanidiniocarbonyl) pyrrole receptor in aqueous solvents. *Chem. Commun.* **1999**, 843-844.

(20) Perrin, D. D. *Dissociation constants of organic bases in aqueous solution*; Supplement; Butterworths: London, 1972.

(21) Schmuck, C. Carboxylate binding by 2-(guanidiniocarbonyl)pyrrole receptors in aqueous solvents: Improving the binding properties of guanidinium cations through additional hydrogen bonds. *Chem. Eur. J.* **2000**, *6*, 709-718.

(22) Lehn, J.-M. *Supramolecular chemistry: Concepts and perspectives*; VCH: Weinheim,

1995.
(23) Sewald, N.; Jakubke, H.-D. *Peptides: Chemistry and biology*; Wiley-VCH: Weinheim, 2002.
(24) Schmuck, C. Von der molekularen Erkennung zum Design neuer Wirkstoffe. *Chem. uns. Zeit* **2001**, 356-366.
(25) Mirsky, V. M.; Yatsimirsky, A. *Artificial receptors for chemical sensors*; Wiley-VCH: Weinheim, 2010.
(26) Williams, D. H.; Bardsley, B. Die Vancomycin-Antibiotica und der Kampf gegen resistente Bakterien. *Angew. Chem.* **1999**, 1264-1286.
(27) Xu, R.; Greiveldinger, G.; Marenus, L. E.; Cooper, A.; Ellman, J. A. Combinatorial library approach for the identification of synthetic receptors targeting Vancomycin-resistant bacteria. *J. Am. Chem. Soc.* **1999**, *121*, 4898-4899.
(28) Schmuck, C.; Heil, M. One-armed artificial receptors for the binding of polar tetrapeptides in water: Probing the substrate selectivity of a combinatorial receptor library. *Chem. Eur. J.* **2006**, *12*, 1339-1348.
(29) Schmuck, C.; Wich, P. Sequence-dependent stereoselectivity in the binding of tetrapeptides in water by a flexible artificial receptor. *Angew. Chem. Int. Ed.* **2006**, *45*, 4277-4281.
(30) Schmuck, C.; Wich, P. Combinatorial receptor finding: Large and random vs. small and focused libraries. *New J. Chem.* **2006**, *30*, 1377-1385.
(31) Chen, C. W.; Whitlock Jr, H. W. Molecular tweezers: A simple model of bifunctional intercalation. *J. Am. Chem. Soc.* **1978**, *100*, 4921-4922.
(32) Westerlund, B.; Korhonen, T. K. Bacterial proteins binding to the mammalianextracellular matrix. *Mol. Biol.* **1993**, *9*, 687-694.
(33) Dower, S. K.; DeLisi, C.; Titus, J. A.; Segal, D. M. Mechanism of binding of multivalent immune complexes to Fc receptors. 1. Equilibrium binding. *Biochemistry* **1981**, *20*, 6326-6334.
(34) Chen, H.; Privalsky, M. L. Cooperative formation of high-order oligomers by retinoid X receptors: An unexpected mode of DNA recognition. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 422.
(35) Boyce, R.; Li, G.; Nestler, H. P.; Suenaga, T.; Still, W. C.; Peptidosteroidal receptors for opioid peptides. Sequence-selective binding using a synthetic receptor library. *J. Am. Chem. Soc.* **1994**, *116*, 7955-7956.
(36) Wennemers, H.; Conza, M.; Nold, M.; Krattiger, P. Diketopiperazine receptors: A novel class of highly selective receptors for binding small peptides. *Chem. Eur. J.* **2001**, *7*, 3342-3347.
(37) Wennemers, H.; Nold, M. C.; Conza, M. M.; Kulicke, K. J.; Neuburger, M. Flexible but with a defined turn - influence of the template on the binding properties of two-armed receptors. *Chem. Eur. J.* **2003**, *9*, 442-448.
(38) Löwik, D. W. P. M.; Weingarten, M. D.; Broekema, M.; Brouwer, A. J.; Still, W. C.; Liskamp, R. M. J. Tweezers with different bite: Increasing the affinity of synthetic receptors by varying the hinge part. *Angew. Chem. Int. Ed.* **1998**, *37*, 1846-1850.
(39) Monnee, M. C. F.; Brouwer, A. J.; Liskamp, R. M. J. Synthesis, screening and evaluation of a combined library of tweezer-and tripodal synthetic receptors. *QSAR Comb. Sci.* **2004**, *23*, 546-559.
(40) Shepherd, J.; Gale, T.; Jensen, K. B.; Kilburn, J. D. Synthesis of unsymmetrical tweezer receptor libraries and identification of receptors for Lys-D-Ala-D-Ala in aqueous solution. *Chem. Eur. J.* **2006**, *12*, 713-720.
(41) Davies, M.; Bonnat, M.; Guillier, F.; Kilburn, J. D.; Bradley, M. Screening an inverted peptide library in water with a guanidinium-based tweezer receptor. *J. Org. Chem.* **1998**, *63*, 8696-8703.

(42) Voet, D.; Voet, J. G.; Pratt, C. W. *Fundamentals of biochemistry*; Wiley-VCH: Weinheim, 2002.
(43) Nath, S. Beyond the chemiosmotic theory: Analysis of key fundamental aspects of energy coupling in oxidative phosphorylation in the light of a torsional mechanism of energy transduction and ATP synthesis - invited review part 2. *J. Bioenerg. Biomembr.* **2010**, *42*, 301-309.
(44) Riedl, S. J.; Salvesen, G. S. The apoptosome: Signalling platform of cell death. *Nature Rev. Mol. Cell Biol.* **2007**, *8*, 405-413.
(45) Tojima, T.; Hines, J. H.; Henley, J. R.; Kamiguchi, H. Second messengers and membrane trafficking direct and organize growth cone steering. *Nature Rev. Neurosci.* **2011**, 191-203.
(46) Rauschenberg, M.; Bomke, S.; Karst, U.; Ravoo, B. J. Dynamic peptides as biomimetic carbohydrate receptors. *Angew. Chem. Int. Ed.* **2010**, *49*, 7340-7345.
(47) Nakai, C.; Glinsmann, W. Interactions between polyamines and nucleotides. *Biochemistry* **1977**, *16*, 5636-5614.
(48) Kimura, E.; Kodama, M.; Yatsunami, T. Macromonocyclic polyamines as biological polyanion complexons. 2. Ion-pair association with phosphate and nucleotides. *J. Am. Chem. Soc.* **1982**, *104*, 3182-3187.
(49) Dhaenens, M.; Lehn, J.-M.; Vigneron, J.-P. Molecular recognition of nucleosides, nucleotides and anionic planar substrates by a water-soluble bis-intercaland-type receptor molecule. *J. Chem. Soc., Perkin Trans. 2* **1993**, 1379-1381.
(50) Bazzicalupi, C.; Bencini, A.; Biagini, S.; Faggi, E.; Meini, S.; Giorgi, C.; Spepi, A.; Valtancoli, B. Exploring the binding ability of phenanthroline-based polyammonium receptors for anions: Hints for design of selective chemosensors for nucleotides. *J. Org. Chem.* **2009**, 7349-7363.
(51) Guo, Y.; Ge, Q.; Lin, H.; Lin, H.; Zhu, S.; Zhou, C. Recognition promoted by Zn^{2+} between phenanthroline bridging polyaza ligands and nucleotides – Zn^{2+} acts as ´messenger´ between the receptor and substrate. *J. Mol. Recognit.* **2003**, *16*, 102-111.
(52) Albelda, M. T.; Bernardo, M. A.; Garcia-España, E.; Godino-Salido, M. L.; Luis, S. V.; Melo, M. J.; Pina, F.; Soriano, C. Thermodynamics and fluorescence emission studies on potential molecular chemosensors for ATP recognition in aqueous solution. *J. Chem. Soc., Perkin Trans. 2* **1999**, 2545-2549.
(53) Hirsch, A. K. H.; Fischer, F. R.; Diederich, F. Molekulare Erkennung von Phosphaten in der Strukturbiologie. *Angew. Chem.* **2007**, *119*, 342-357.
(54) Gao, H.; Cai, L.; Qi, Y.; Wang, H. Synthesis of 1,3-{Di-[N-bis(dimethylamino) methane]}-benzyl-diamide and its molecular recognition of nucleotides in aqueous solution. *Supramol. Chem.* **2003**, *15*, 323-325.
(55) Schmidtchen, F. P. A non-macrocyclic host for binding organic phosphates in protic solvents. *Tetrahedron Lett.* **1989**, *30*, 4493-4496.
(56) Onda, M.; Yoshihara, K.; Koyano, H.; Ariga, K.; Kunitake, T. Molecular recognition of nucleotides by the guanidinium unit at the surface of aqueous micelles and bilayers. A comparison of microscopic and macroscopic interfaces. *J. Am. Chem. Soc.* **1996**, *118*, 8524-8530.
(57) Sebo, L.; Diederich, F. Tetrakis(phenylamidinium)-substituted resorcin[4]arene receptors for the complexation of dicarboxylates and phosphates in protic solvents. *Helv. Chim. Acta* **2000**, 93-113.
(58) Eliseev, A. V.; Schneider, H. J. Molecular recognition of nucleotides, nucleosides, and sugars by aminocyclodextrins. *J. Am. Chem. Soc.* **1994**, *116*, 6081-6088.
(59) Menuel, S.; Duval, R. E.; Cuc, D.; Mutzenhardt, P.; Marsura, A. Molecular recognition of nucleotides by a new bis(guanidinium)tetrakis(β-cyclodextrin) tetrapod. *New J. Chem.* **2007**, *31*, 995-1000.

(60) Butterfield, S. M.; Waters, M. L. A designed β-hairpin peptide for molecular recognition of ATP in water. *J. Am. Chem. Soc.* **2003**, *125*, 9580-9581.
(61) Schneider, S. E.; O'Nei, S. N.; Anslyn, E. V. Coupling rational design with libraries leads to the production of an ATP selective chemosensor. *J. Am. Chem. Soc.* **2000**, *122*, 542-543.
(62) McCleskey, S. C.; Griffin, M. J.; Schneider, S. E.; McDevitt, J. T.; Anslyn, E. V. Differential receptors create patterns diagnostic for ATP and GTP. *J. Am. Chem. Soc.* **2003**, *125*, 1114-1115.
(63) Matsui, J.; Nagano, J.; Miyoshi, D.; Tamaki, K.; Sugimoto, N. An approach to peptide-based ATP receptors by a combination of random selection, rational design, and molecular imprinting. *Biosens. Bioelectron.* **2009**, 563-567.
(64) Gröger, K.; Baretić, D.; Piantanida, I.; Marjanović, M.; Kralj, M.; Grabar, M.; Tomić, S.; Schmuck, C. Guanidiniocarbonyl-pyrrole-aryl conjugates as nucleic acid sensors: Switch of binding mode and spectroscopic responses by introducing additional binding sites into the linker. *Org. Biomol. Chem.* **2011**, *9*, 198-209.
(65) Hermann, T. Strategies for the design of drugs targeting RNA and RNA-protein complexes. *Angew. Chem. Int. Ed* **2000**, *39*, 1890-1904.
(66) Turner, P. R.; Denny, W. A. The genome as a drug target sequence specific minor groove binding ligands. *Current Drug Targets* **2000**, *1*, 1-14.
(67) Watson, J. D.; Crick, F. H. C. Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature* **1953**, 737-738.
(68) Egli, M.; Saenger, W. *Principles of nucleic acid structure*; Springer: New York, 1984.
(69) Venkadesh, S.; Mandal, P. K.; Gautham, N. The structure of a full turn of an A DNA duplex $d(CGCGGGTACCCGCG)_2$. *Biochem. Biophys. Res. Commun.* **2011**, 307-312.
(70) Ohishi, H.; Tozuka, Y.; Da-Yang, Z.; Ishida, T.; Nakatani, K. The rare crystallographic structure of $d(CGCGCG)_2$: The natural spermidine molecule bound to the minor groove of left-handed Z-DNA $d(CGCGCG)_2$ at 10 °C. *Biochem. Biophys. Res. Commun.* **2007**, *358*, 24-28.
(71) Tereshko, V.; Minasov, G.; Egli, M. The Dickerson-Drew B-DNA dodecamer revisited at atomic resolution. *J. Am. Chem. Soc.* **1999**, *121*, 470-471.
(72) Harrison, S. C. Three-dimensional intricacies in protein-DNA recognition and transcriptional control. *Nat. Struc. Mol. Biol.* **2007**, *14*, 1118-1119.
(73) Blackburn, G. M.; Gait, M. J.; Loakes, D.; Williams, D. M. *Nucleic acids in chemistry and biology*; RSC Publishing: Cambridge, 2006.
(74) Herbert, A.; Rich, A. The biology of left-handed Z-DNA. *J. Biol. Chem.* **1996**, *271*, 11595-11598.
(75) Hud, N. V.; Plavec, J. A unified model for the origin of DNA sequence-directed curvature. *Biopolymers* **2003**, *69*, 144-158.
(76) Rodger, A.; Nordén, B. *Circular dichroism and linear dichroism*; Oxford University Press: New York, 1997.
(77) Blackburn, G. M.; Gait, M. J. *Nucleic acids in chemistry and biology*; Oxford University Press: New York, 1997.
(78) Nordén, B.; Kurucsev, T. Analysing DNA complexes by circular and linear dichroism. *J. Mol. Recognit.* **1994**, *7*, 141-155.
(79) Garbett, N. C.; Ragazzon, P. A.; Chaires, J. B. Circular dichroism to determine binding mode and affinity of ligand-DNA interactions. *Nature Prot.* **2007**, 3166-3172.
(80) Mergny, J. L.; Lacroix, L. Analysis of thermal melting curves. *Oligonucleotides* **2003**, *13*, 515-537.
(81) Chaires, J. B.; Waring, M. J. *Drug-nucleic acid interactions*. In *Methods in enzymology*; Academic Press Inc., 2001; Bd. 340.
(82) Jelesarov, I.; Bosshard, H. R. Isothermal titration calorimetry and differential scanning

calorimetry as complementary tools to investigate the energetics of biomolecular recognition. *J. Mol. Recognit.* **1999**, *12*, 3-18.
(83) Chaires, J. B. Energetics of drug-DNA interactions. *Biopolymers* **1997**, *44*, 201-205.
(84) Tse, W.; Boger, D. Sequence-selective DNA recognition: Natural products and nature´s lessons. *Chem. Biol.* **2004**, *11*, 1607-1617.
(85) Berman, H. M.; Young, P. R. The interaction of intercalating drugs with nucleic acids. *Annu. Rev. Biophys. Bioeng.* **1981**, 87-114.
(86) Neidle, S. DNA minor-groove recognition by small molecules. *Nat. Prod. Rep.* **2001**, *18*, 291-309.
(87) Manning, G. S. Counterion binding in polyelectrolyte theory. *Acc. Chem. Res.* **1979**, *12*, 443-449.
(88) Drew, H. R.; Samson, S.; Dickerson, R. E. Structure of a B-DNA dodecamer at 16 K. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4040-4044.
(89) Shui, X.; McFail-Isom, L.; Hu, G. G.; Williams, L. D. The B-DNA dodecamer at high resolution reveals a spine of water on sodium. *Biochemistry* **1998**, *37*, 8341-8355.
(90) Chaires, J. B. Dissecting the free energy of drug binding to DNA. *Anticancer Drug Des.* **1996**, *11*, 569-80.
(91) Westhof, E.; Beveridge, D. L. Hydration of nucleic acids. *Water Sci. Rev.* **1990**, *5*, 24-136.
(92) Lerman, L. S. Structural considerations in the interaction of DNA and acridines. *J. Mol. Biol.* **1961**, *3*, 18-30.
(93) Gao, Y.; Robinson, H.; Wijsman, E. R.; van der Marel, G. A.; van Boom, J. H.; Andrew, H. J. W. Binding of daunorubicin to β-D-glucosylated DNA found in protozoa Tryponosoma brucei studied by X-ray crystallography. *J. Am. Chem. Soc.* **1997**, *119*, 1496-1497.
(94) Wang, A. H.-J.; Ughetto, G.; Quigley, G. J.; Rich, A. Interactions between an anthracycline antibiotic and DNA: Molecular structure of daunomycin complexed to d(DpGpTpApCpG) at 1.2-Å resolution. *Biochemistry* **1987**, 1152-1163.
(95) Demeunynck, M.; Bailly, C.; Wilson, W. D. *Small molecule DNA and RNA binders: From synthesis to nucleic acid complexes*; Wiley-VCH: Weinheim, 2003.
(96) Waring, M. J. Complex formation between ethidium bromide and nucleic acids. *J. Mol. Biol.* **1965**, *13*, 269-282.
(97) Le Pecq, J.-B.; Paoletti, C. A new fluorometric method for RNA and DNA determination. *Analyt. Biochem.* **1966**, *17*, 100-107.
(98) Borst, P. Ethidium DNA agarose gel electrophoresis: How it started. *IUBMB Life* **2005**, *57*, 745-747.
(99) Alonso, A.; Almendral, M. J.; Curto, Y.; Criado, J. J.; Rodríguez, E.; Manzano, J. L. Determination of the DNA-binding characteristics of ethidium bromide, proflavine, and cisplatin by flow injection analysis: Usefulness in studies on antitumor drugs. *Analyt. Biochem.* **2006**, *335*, 157-164.
(100) Hernandez-Folgado, L.; Schmuck, C.; Tomić, S.; Piantanida, I. A novel pyrene-guanidiniocarbonyl-pyrrole cation efficiently differentiates between ds-DNA and ds-RNA by two independent, sensitive spectroscopic methods. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2977-2981.
(101) Hernandez-Folgado, L.; Baretić, D.; Piantanida, I.; Marjanović, M.; Kralj, M.; Rehm, T.; Schmuck, C. Guanidiniocarbonylpyrrole-aryl derivatives: Structure tuning for spectrophotometric recognition of specific DNA and RNA sequences and for antiproliferative activity. *Chem. Eur. J.* **2010**, *16*, 3036-3056.
(102) Boger, D. L.; Chen, J. H.; Saionz, K. W. (-)-Sandramycin: Total synthesis and characterization of DNA binding properties. *J. Am. Chem. Soc.* **1996**, *118*, 1629-1644.
(103) Chaires, J. B.; Leng, F.; Przewloka, T.; Fokt, I.; Ling, Y. H.; Perez-Soler, R.; Priebe,

W. Structure-based design of a new bisintercalating anthracycline antibiotic. *J. Med. Chem.* **1997**, *40*, 261-266.
(104) Hu, G. G.; Shui, X.; Leng, F.; Priebe, W.; Chaires, J. B.; Williams, L. D. Structure of a DNA-bisdaunomycin complex. *Biochemistry* **1997**, *36*, 5940-5946.
(105) Dervan, P. B. Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.* **2001**, *9*, 2215-2236.
(106) Perahia, D.; Pullman, A.; The molecular electrostatic potential of B-DNA helix. *Theor. Chem. Acc.* **1979**, *50*, 351-354.
(107) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. The molecular origin of DNA-drug specificity in netropsin and distamycin. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 1376-1380.
(108) Pelton, J. G.; Wemmer, D. E. Structural characterization of a 2:1 distamycin A d(CGCAAATTGGC) complex by two-dimensional NMR. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5723-5727.
(109) Goodsell, D. S.; Kopka, M. L.; Dickerson, R. E. Refinement of netropsin bound to DNA: Bias and feedback in electron density map interpretation. *Biochemistry* **1995**, *34*, 4983-4993.
(110) Mitra, S. N.; Wahl, M. C.; Sundaralingam, M. Structure of the side-by-side binding of distamycin to d(GTATATAC)$_2$. *Acta Crystallogr. Sect. D* **1999**, 602-609.
(111) Rentzeperis, D.; Marky, L. A.; Dwyer, T. J.; Geierstanger, B. H.; Pelton, J. G.; Wemmer, D. E. Interaction of minor groove ligands to an AAATT/AATTT site: Correlation of thermodynamic characterization and solution structure. *Biochemistry* **1995**, *34*, 2937-2945.
(112) Kapuscinski, J. DAPI: A DNA-specific fluorescent probe. *Biotech. Histochem.* **1995**, *70*, 220-233.
(113) Wade, W. S.; Mrksich, M.; Dervan, P. B. Design of peptides that bind in the minor groove of DNA at 5'-(A, T)G(A, T)C(A, T)-3' sequences by a dimeric side-by-side motif. *J. Am. Chem. Soc.* **1992**, *114*, 8783-8794.
(114) Mrksich, M.; Parks, M. E.; Dervan, P. B. Hairpin peptide motif. A new class of oligopeptides for sequence-specific recognition in the minor groove of double-helical DNA. *J. Am. Chem. Soc.* **1994**, *116*, 7983-7988.
(115) Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. A structural basis for recognition of AT and TA base pairs in the minor groove of B-DNA. *Science* **1998**, *282*, 111-115.
(116) White, S.; Szewcyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. Recognition of the four Waston-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* **1998**, 468-471.
(117) Kielkopf, C. L.; Baird, E. E.; Dervan, P. B.; Rees, D. C. Structural basis for GC recognition in the DNA minor groove. *Nat. Struc. Biol.* **1998**, 104-109.
(118) Trauger, J. W.; Baird, E. E.; Dervan, P. B. Recognition of 16 base pairs in the minor groove of DNA by a pyrrole-imidazole polyamide dimer. *J. Am. Chem. Soc.* **1998**, *120*, 3534-3535.
(119) Gottesfeld, J. M.; Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B. Regulation of gene expression by small molecules. *Nature* **1997**, 202-205.
(120) Garvie, C. W.; Wolberger, C. Recognition of specific DNA sequences. *Mol. Cell.* **2001**, *8*, 937-946.
(121) Staple, D. W.; Venditti, V.; Niccolai, N.; Elson-Schwab, L.; Tor, Y.; Butcher, S. E. Guanidinoneomycin B recognition of an HIV-1 RNA helix. *ChemBioChem* **2008**, *9*, 93-102.
(122) Kim, S. K.; Nordén, B. Methyl green: A DNA major-groove binding drug. *FEBS Lett.* **1993**, *315*, 61-64.

(123) Zadmard, R.; Schrader, T. DNA recognition with large calixarene dimers. *Angew. Chem. Int. Ed.* **2006**, *45*, 2703-2706.
(124) Behrens, C.; Nielsen, P. E. A combinatorial approach to new DNA minor groove binders. *Comb. Chem. High Throughput Screen.* **1998**, 127-134.
(125) Alam, R.; Maeda, M.; Sasaki, S. DNA-binding peptides searched from the solid-phase combinatorial library with the use of the magnetic beads attaching the target duplex DNA. *Bioorg. Med. Chem.* **2000**, 465-473.
(126) Guelev, V. M.; Harting, M. T.; Lokey, R. S.; Iverson, B. L. Altered sequence specifity identified from a library of DNA-binding small molecules. *Chem. Biol.* **2000**, 1-8.
(127) Boger, D. L.; Fink, B. E.; Hedrick, M. P. Total synthesis of distamycin A and 2640 analogues: A solution-phase combinatorial approach to the discovery of new, bioactive DNA binding agents and development of a rapid, high-throughput screen for determining relative DNA binding affinity or DNA binding sequence selectivity. *J. Am. Chem. Soc.* **2000**, *122*, 6382-6394.
(128) Stover, J. S.; Shi, J.; Jin, W.; Vogt, P. K.; Boger, D. L. Discovery of inhibitors of aberrant gene transcription from libraries of DNA binding molecules: Inhibition of LEF-1-mediated gene transcription and oncogenic transformation. *J. Am. Chem. Soc.* **2009**, *131*, 3342-3348.
(129) Verma, I. M.; Weitzman, M. D. Gene therapy: Twenty-first century medicine. *Annu. Rev. Biochem.* **2005**, *74*, 711-738.
(130) Cavazzana-Calvo, M.; Hacein-Bey, S.; de Saint Basile, G.; Gross, F.; Yvon, E.; Nusbaum, P.; Selz, F.; Hue, C.; Certain, S.; Casanova, J.-L.; Bousso, P.; Le Deist, F.; Fischer, A. Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* **2000**, *288*, 669-672.
(131) Nguyen, D. M.; Spitz, F. R.; Yen, N.; Cristiano, R. J.; Roth, J. A. Gene therapy for lung cancer: Enhancement of tumor suppression by a combination of sequential systemic cisplatin and adenovirus-mediated p53 gene transfer. *J. Thorac. Cardiovasc. Surg.* **1996**, *112*, 1372-1377.
(132) Müller, G. Personalisierte Medizin – Vision oder Fiktion? Beispiel: Altersdiabetes. *Biol. uns. Zeit* **2010**, *40*, 46-54.
(133) Rogers, S.; Lowenthal, A.; Terheggen, H. G.; Columbo, J. P. Induction of arginase activity with the Shope papilloma virus in tissue culture cells from argininemic patients. *J. Exp. Med.* **1973**, 1091-1096.
(134) Villemejane, J.; Mir, L. M. Physical methods of nucleic acid transfer: General concepts and applications. *Brit. J. Pharmacol.* **2009**, *157*, 207-219.
(135) Midoux, P.; Pichon, C.; Yaouanc, J.-J.; Jaffrès, P.-A. Chemical vectors for gene delivery: A current review on polymers, peptides and lipids containing histidine or imidazole as nucleic acids carriers. *Brit. J. Pharmacol.* **2009**, *157*, 166-178.
(136) Conner, S. D.; Schmid, S. L. Regulated portals of entry into the cell. *Nature* **2003**, *422*, 37-44.
(137) Mukherjee, S.; Ghosh, R. N.; Maxfield, F. R. Endocytosis. *Physiol. Rev.* **2007**, *77*, 759-799.
(138) Pack, D. W.; Hoffman, A. S.; Pun, S.; Stayton, P. S. Design and development of polymers for gene delivery. *Nat. Rev. Drug Disc.* **2005**, *4*, 581-593.
(139) Laemmli, U. K. Characterization of DNA condensates induced by poly(ethylene oxide) and polylysine. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 4288-4292.
(140) Wu, G. Y.; Wu, C. H. Receptor-mediated gene delivery and expression in vivo. *J. Biol. Chem.* **1988**, *263*, 14621-14624.
(141) Benns, J. M.; Choi, J.-S.; Mahato, R. I.; Park, J.-S.; Kim, S. W. pH-sensitive cationic polymer gene delivery vehicle: N-Ac-poly(L-histidine)-graft-poly(L-lysine) comb shaped polymer. *Bioconjugate Chem.* **2000**, *11*, 637-645.

(142) Futai, M.; Oka, T.; Moriyama, Y.; Wada, Y. Diverse roles of single membrane organelles: Factors establishing the acid lumenal pH. *J. Biochem.* **1998**, *124*, 259-267.
(143) Choi, Y. H.; Liu, F.; Kim, J. S.; Choi, Y. K.; Park, J. S.; Kim, S. W. Polyethylene glycol-grafted poly-L-lysine as polymeric gene carrier. *J. Cont. Rel.* **1998**, *54*, 39-48.
(144) Boussif, O.; Lezoualc'h, F.; Zanta, M. A.; Mergny, M. D.; Scherman, D.; Demeneix, B.; Behr, J. P. A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: Polyethylenimine. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 7297-7301.
(145) Suh, J.; Paik, H.-J.; Hwang, B. K. Ionization of poly(ethylenimine) and poly(allysamine) at various pH´s. *Bioorg. Chem.* **1994**, *22*, 318-327.
(146) Sonawane, N. D.; Szoka, F. C.; Verkman, A. S. Chloride accumulation and swelling in endosomes enhances DNA transfer by polyamine-DNA polyplexes. *J. Biol. Chem.* **2003**, *278*, 44826-44831.
(147) Fischer, D.; Li, Y.; Ahlemeyer, B.; Krieglstein, J.; Kissel, T. In vitro cytotoxicity testing of polycations: Influence of polymer structure on cell viability and hemolysis. *Biomaterials* **2003**, *24*, 1121-1131.
(148) Tranchant, I.; Thompson, B.; Nicolazzi, C.; Mignet, N.; Scherman, D. Physicochemical optimisation of plasmid delivery by cationic lipids. *J. Gene Med.* **2004**, *6*, 24-35.
(149) Felgner, P. L.; Gadek, T. R.; Holm, M.; Roman, R.; Chan, H. W.; Wenz, M.; Northrop, J. P.; Ringold, G. M.; Danielsen, M. Lipofection: A highly efficient, lipid-mediated DNA-transfection procedure. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7413-7417.
(150) Zhi, D.; Zhang, S.; Wang, B.; Zhao, Y.; Yang, B.; Yu, S. Transfection efficiency of cationic lipids with different hydrophobic domains in gene delivery. *Bioconjugate Chem.* **2010**, *21*, 563-577.
(151) Fielden, M. L.; Perrin, C.; Kremer, A.; Bergsma, M.; Stuart, M. C.; Camilleri, P.; Engberts, J. B. F. N. Sugar-based tertiary amino gemini surfactants with a vesicle-to-micelle transition in the endosomal pH range mediate efficient transfection in vitro. *Eu. J. Biochem.* **2001**, *268*, 1269-1279.
(152) Bell, P. C.; Bergsma, M.; Dolbnya, I. P.; Bras, W.; Stuart, M. C. A.; Rowan, A. E.; Feiters, M. C.; Engberts, J. B. F. N. Transfection mediated by gemini surfactants: Engineered escape from the endosomal compartment. *J. Am. Chem. Soc.* **2003**, *125*, 1551-1558.
(153) Filion, M. C.; Phillips, N. C. Major limitations in the use of cationic liposomes for DNA delivery. *Int. J. Pharmacol.* **1998**, *162*, 159-170.
(154) Heitz, F.; Morris, M. C.; Divita, G. Twenty years of cell-penetrating peptides: From molecular mechanisms to therapeutics. *Brit. J. Pharmacol.* **2009**, *157*, 195-206.
(155) Frankel, A. D.; Pabo, C. O. Cellular uptake of the tat protein from human immunodeficiency virus. *Cell* **1988**, *55*, 1189-1193.
(156) Joliot, A.; Pernelle, C.; Deagostini-Bazin, H.; Prochiantz, A. Antennapedia homeobox peptide regulates neural morphogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 1864-1868.
(157) Derossi, D.; Joliot, A. H.; Chassaing, G.; Prochiantz, A. The third helix of the Antennapedia homeodomain translocates through biological membranes. *J. Biol. Chem.* **1994**, *269*, 10444-10450.
(158) Vivès, E.; Brodin, P.; Lebleu, B. A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. *J. Biol. Chem.* **1997**, *272*, 16010-16017.
(159) Billeter, M.; Qian, Y.-Q.; Otting, G.; Müller, M.; Gehring, W. J.; Wüthrich, K. Determination of the three-dimensional structure of the Antennapedia homeodomain

from Drosophila in solution by 1H nuclear magnetic resonance spectroscopy. *J. Mol. Biol.* **1990**, 183-197.

(160) Hassane, S. F.; Saleh, A. F.; Abes, R.; Gait, M. J.; Lebleu, B. Cell penetrating peptides: Overview and applications to the delivery of oligonucleotides. *Cell. Mol. Life Sci.* **2009**, *67*, 715-726.

(161) Morris, M. C.; Vidal, P.; Chaloin, L.; Heitz, F.; Divita, G. A new peptide vector for efficient delivery of oligonucleotides into mammalian cells. *Nucl. Acids Res.* **1997**, *25*, 2730-2736.

(162) Futaki, S. Arginine-rich peptides: Potential for intracellular delivery of macromolecules and the mystery of the translocation mechanism. *Int. J. Pharm.* **2002**, 1-7.

(163) Farrera-Sinfreu, J.; Giralt, E.; Castel, S.; Albericio, F.; Royo, M. Cell-penetrating cis-γ-amino-L-proline-derived peptides. *J. Am. Chem. Soc.* **2005**, *127*, 9459-9468.

(164) Gao, S.; Simon, M. J.; Hue, C. D.; Morrison, B.; Banta, S. An unusual cell penetrating peptide identified using a plasmid display-based functional selection platform. *ACS Chem. Biol.* **2011**, 484-491.

(165) Marks, J. R.; Placone, J.; Hristova, K.; Wimley, W. C. Spontaneous membrane-translocating peptides by orthogonal high-throughput screening. *J. Am. Chem. Soc.* **2011**, 8995-9004.

(166) Kamide, K.; Nakakubo, H.; Uno, S.; Fukamizu, A. Isolation of novel cell-penetrating peptides from a random peptide library using in vitro virus and their modifications. *Int. J. Mol. Med.* **2009**, *25*, 41-51.

(167) Küstner, B.; Schmuck, C.; Wich, P.; Jehn, C.; Srivastava, S. K.; Schlücker, S. UV Resonance Raman spectroscopic monitoring of supramolecular complex formation: Peptide recognition in aqueous solution. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4598-4603.

(168) Schmuck, C.; Wich, P.; Küstner, B.; Kiefer, W.; Schlücker, S. Direct and label-free detection of solid-phase-bound compounds by using Surface-Enhanced Raman Scattering microspectroscopy. *Angew. Chem. Int. Ed.* **2007**, 4786-4789.

(169) Funasaki, N.; Nomura, M.; Ishikawa, S.; Neya, S. NMR chemical shift references for binding constant determination in aqueous solutions. *J. Phys. Chem. B* **2001**, *105*, 7361-7365.

(170) Schneider, H. J.; Dürr, H. *Frontiers in supramolecular organic chemistry and photochemistry*; Wiley-VCH: New York, 1991.

(171) Raman, C. V.; Krishnan, K. S. Polarisation of scattered light-quanta. *Nature* **1928**, 169.

(172) Petry, R.; Schmitt, M.; Popp, J. Raman spectroscopy – a prospective tool in the life sciences. *ChemPhysChem 4*, 14-30.

(173) Spiro, T. G. *Biological applications of Raman spectroscopy*; Wiley-VCH: New York, 1988.

(174) Asher, S. A. UV Resonance Raman spectroscopy for analytical, physical, and biophysical chemistry part 1. *Anal. Chem.* **1993**, 59-66.

(175) Asher, S. A. UV Resonance Raman spectroscopy for analytical, physical, and biophysical chemistry part 2. *Anal. Chem.* **1993**, 201-210.

(176) Schmuck, C.; Bickert, V.; Merschky, M.; Geiger, L.; Rupprecht, D.; Dudaczek, J.; Wich, P.; Rehm, T.; Machon, U. A facile and efficient multi-gram synthesis of N-protected 5-(guanidinocarbonyl)-1H-pyrrole-2-carboxylic acids. *Eur. J. Org. Chem.* **2008**, *2008*, 324-329.

(177) Heil, M. *Doktorarbeit: Synthese und Screening einer kombinatorischen Rezeptorbibliothek für biologisch relevante Tetrapeptide*; Bayerische Julius-Maximilians-Universität: Würzburg, 2004.

(178) Chan, W. C.; White, P. D. *Fmoc solid phase peptide synthesis – a practical approach*; Oxford University Press: New York, 1999.
(179) Coin, I.; Beyermann, M.; Bienert, M. Solid-phase peptide synthesis: From standard procedures to the synthesis of difficult sequences. *Nat. Protoc.* **2007**, *2*, 3247-3256.
(180) Isidro-Llobet, A.; Álvarez, M.; Albericio, F. Amino acid-protecting groups. *Chem. Rev.* **2009**, *109*, 2455-2504.
(181) Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal. Biochem.* 595-598.
(182) Drijfhout, J. W.; Perdijk, E. W.; Weijer, W. J.; Bloemhoff, W. Controlled peptide-protein conjugation by means of 3-nitro-2-pyridinesulfenyl protection-activation. *Int. J. Pept. Protein Res.* **1988**, *32*, 161-166.
(183) Han, Y.; Albericio, F.; Barany, G. Occurrence and minimization of cysteine racemization during stepwise solid-phase peptide synthesis. *J. Org. Chem.* **1997**, *62*, 4307-4312.
(184) Niebling, S. *Doktorarbeit: Schwingungsspektroskopische und computerchemische Untersuchungen zur molekularen Erkennung von Tetrapeptiden durch künstliche Rezeptoren*; Universität Osnabrück: Osnabrück.
(185) Lee, D. D.; Seung, S. Learning the parts of objects by non-negative matrix factorization. *Nature* **1999**, 788-791.
(186) Aroca, R. *Surface enhanced vibrational spectroscopy*; Wiley-VCH: New York, 2006.
(187) Fleischmann, M.; Hendra, P. J.; McQuillan, A. J. Raman spectra of pyridine adsorbed at a silver electrode. *Chem. Phys. Lett.* **1974**, 163-166.
(188) Kneipp, K.; Kneipp, H.; Itzkan, I.; Dasari, R. R.; Feld, M. S. Ultrasensitive chemical analysis by Raman spectroscopy. *Chem. Rev.* **1999**, *99*, 2957-2976.
(189) Kassel, D. B. Combinatorial chemistry and mass spectrometry in the 21st century drug discovery laboratory. *Chem. Rev.* **2001**, *101*, 255-268.
(190) Affleck, R. L. Solutions for library encoding to create collections of discrete compounds. *Curr. Opin. Chem. Biol.* **2001**, *5*, 257-263.
(191) Gellner, M. *Doktorarbeit: Plasmonisch aktive Kern/Schale-Nanopartikel in der oberflächenverstärkten Raman-Spektroskopie*; Universität Osnabrück: Osnabrück, 2011.
(192) Li, J. F.; Huang, Y. F.; Ding, Y.; Yang, Z. L.; Li, S. B.; Zhou, X. S.; Fan, F. R.; Zhang, W.; Zhou, Z. Y.; Wu, D. Y.; Ren, B.; Wang, Z. L.; Tian, Z. Q. Shell-isolated nanoparticle-enhanced Raman spectroscopy. *Nature* **2010**, *464*, 392-395.
(193) Küstner, B.; Gellner, M.; Schütz, M.; Schöppler, F.; Marx, A.; Ströbel, P.; Adam, P.; Schmuck, C.; Schlücker, S. SERS labels for red laser excitation: Silica-encapsulated SAMs on tunable gold/silver nanoshells. *Angew. Chem. Int. Ed.* **2009**, *48*, 1950-1953.
(194) McMahon, J. M.; Henry, A.-I.; Wustholz, K. L.; Natan, M. J.; Freeman, R. G.; Duyne, R. P.; Schatz, G. C. Gold nanoparticle dimer plasmonics: Finite element method calculations of the electromagnetic enhancement to surface-enhanced Raman spectroscopy. *Anal Bioanal. Chem.* **2009**, *394*, 1819-1825.
(195) Kuchelmeister, H. Y. *Diplomarbeit: Synthese von unsymmetrischen Tweezer-Rezeptoren für das biologisch relevante Depsipeptid L-Lys-D-Ala-D-Lac*; Bayerische Julius-Maximilians-Universität: Würzburg, 2007.
(196) Merrick, J. P.; Moran, D.; Radom, L. An evaluation of harmonic vibrational frequency scale factors. *J. Phys. Chem. A* **2007**, *111*, 11683-11700.
(197) Bacsa, B.; Horváti, K.; Bõsze, S.; Andreae, F.; Kappe, C. O. Solid-phase synthesis of difficult peptide sequences at elevated temperatures: A critical comparison of microwave and conventional heating technologies. *J. Org. Chem.* **2008**, *73*, 7532-7542.

(198) Palasek, S. A.; Cox, Z. J.; Collins, J. M. Limiting racemization and aspartimide formation in microwave-enhanced Fmoc solid phase peptide synthesis. *J. Pept. Sci.* **2007**, *13*, 143-148.
(199) Albert, A.; Goldacre, R.; Phillips, J. The strength of heterocyclic bases. *J. Chem. Soc.* **1948**, 2240.
(200) Schmuck, C. How to improve guanidinium cations for oxoanion binding in aqueous solution? The design of artificial peptide receptors. *Coord. Chem. Rev.* **2006**, *250*, 3053-3067.
(201) *Origin 8 user guide*; OriginLab Corporation: Massachusetts, 2007.
(202) Job, P. Spectrographic study of the formation of complexes in solutions and of their stability. *Compt. Rend.* **1925**, 928-930.
(203) MacCarthy, P. Simplified experimental route for obtaining Job´s curves. *Anal. Chem.* **1978**, *50*, 2165.
(204) Gampp, H.; Maeder, M. Calculation of equilibrium constants from multiwavelength spectroscopic data – II. SPECFIT: Two user-friendly programs in BASIC and standard FORTRAN 77. *Talanta* 257-264.
(205) Maeder, M.; Zuberbühler, A. D. Nonlinear least-squares fitting of multivariate absorption data. *Anal. Chem.* **1990**, *62*, 2220-2224.
(206) Kuchelmeister, H. Y.; Schmuck, C. Nucleotide recognition in water by a guanidinium-based artificial tweezer receptor. *Chem. Eur. J.* **2011**, *17*, 5311-5318.
(207) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. Macromodel – an integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J. Comput. Chem.* **1990**, *11*, 440-467.
(208) González-Gaitano, G.; Tardajos, G. Chemical equilibrium in supramolecular systems as studied by NMR spectrometry. *J. Chem. Edu.* **2004**, *81*, 270-274.
(209) McGaughey, G. B.; Gagné, M.; Rappé, A. K. π-Stacking interactions. *J. Biol. Chem.* **1998**, *273*, 15458-15463.
(210) Kato, S.-I.; Matsumoto, T.; Ideta, K.; Shimasaki, T.; Goto, K.; Shinmyozu, T. Supramolecular assemblies and redox modulation of pyromellitic diimide-based cyclophane via noncovalent interactions with naphthol. *J. Org. Chem.* **2006**, *71*, 4723-4733.
(211) Veauthier, J. M.; Carlson, C. N.; Collis, G. E.; Kiplinger, J. L.; John, K. D. The synthesis of poly-nitrile aromatic and oligopyridine ligands via palladium-catalyzed cyanation of aryl halides. *Synthesis* **2005**, 2683-2686.
(212) Sundermeier, M.; Zapf, A.; Mutyala, S.; Baumann, W.; Sans, J.; Weiss, S.; Beller, M. Progress in the palladium-catalyzed cyanation of aryl chlorides. *Chem. Eur. J.* **2003**, *9*, 1828-1836.
(213) Brückner, R. *Reaktionsmechanismen*; Spektrum: Heidelberg, 2004.
(214) Sundermeier, M.; Zapf, A.; Beller, M. A convenient procedure for the palladium-catalyzed cyanation of aryl halides. *Angew. Chem. Int. Ed.* **2003**, *42*, 1661-1664.
(215) Siva, A.; Murugan, E. A new trimeric cinchona alkaloid as a chiral phase-transfer catalyst for the synthesis of asymmetric α-amino acids. *Synthesis* **2005**, 2927-2933.
(216) Gibson, M. S.; Bradshaw, R. W. The Gabriel synthesis of primary amines. *Angew. Chem. Int. Ed.* **1968**, *7*, 919-930.
(217) Grawe, T.; Schrader, T.; Zadmard, R.; Kraft, A. Self-assembly of ball-shaped molecular complexes in water. *J. Org. Chem.* **2002**, *67*, 3755-3763.
(218) Carpino, L. A. 1-Hydroxy-7-azabenzotriazole. An efficient peptide coupling additive. *J. Am. Chem. Soc.* **1993**, *115*, 4397-4398.
(219) Junghänel, S. *Masterarbeit: Synthese neuer Guanidiniocarbonylpyrrol-Rezeptoren zur DNA-Erkennung*; Universität Duisburg-Essen: Essen, 2011.

(220) Stojković, M. R.; Piantanida, I. Tuning urea-phenanthridinium conjugates for DNA/RNA and base pair recognition. *Tetrahedron* **2008**, *64*, 7807-7814.
(221) Plum, G. E.; Bloomfield, V. A. Structural and electrostatic effects on binding of trivalent cations to double-stranded and single-stranded poly [d (AT)]. *Biopolymers* **1990**, *29*, 13-27.
(222) McGhee, J. D. Theoretical calculations of the helix-coil transition of DNA in the presence of large, cooperatively binding ligands. *Biopolymers* **1976**, *15*, 1345-1375.
(223) Crothers, D. M. Statistical thermodynamics of nucleic acid melting transitions with coupled binding equilibria. *Biopolymers* **1971**, *10*, 2147-2160.
(224) Scatchard, G. The attractions of proteins for small molecules and ions. *Acad. Sci.* **1949**, 660-672.
(225) Weiß, C. *Datenanalyse und Modellierung mit Statistica*; Wissenschaftsverlag: Oldenburg, 2006.
(226) McGhee, J. D.; von Hippel, P. H. Erratum: Theoretic aspects of DNA-protein interactions: Co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice. *J. Mol. Biol.* **1976**, 679.
(227) Cain, B. F.; Baguley, B. C.; Denny, W. A. Potential antitumor agents. 28. Deoxyribonucleic acid polyintercalating agents. *J. Med. Chem.* **1978**, *21*, 658–668.
(228) Liu, L.; Guo, Q.-X. Isokinetic Relationship, Isoequilibrium Relationship, and Enthalpy-Entropy Compensation. *Chem. Rev.* **2001**, *101*, 673-696.
(229) Matulis, D.; Rouzina, I.; Bloomfield, V. A. Thermodynamics of DNA binding and condensation: Isothermal titration calorimetry and electrostatic mechanism. *J. Mol. Biol.* **2000**, *296*, 1053-1063.
(230) Manning, G. S. The molecular theory of polyelectrolyte solutions with applications to the lectrostatic properties of polynucleotides. *Q. Rev. Biophys.* **1978**, *11*, 179-246.
(231) Record, M. T. J.; Anderson, C. F.; Lohman, T. M. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: The roles of ion association or release, screening, and ion effects on water activity. *Q. Rev. Biophys.* **1978**, *11*, 103-178.
(232) Chaires, J. B. A thermodynamic signature for drug-DNA binding mode. *Arch. Biochem. Biophys.* **2006**, *453*, 26-31.
(233) BloomWeld, V. A.; Crothers, D. M.; Ignacio Tinoco, J. *Nucleic acids: Structures, properties and functions*; University Science Books: Sausalito, 2000.
(234) Jen-Jacobson, L.; Engler, L. E.; Jacobson, L. A. Structural and thermodynamic strategies for site-specific DNA binding proteins. *Structure* **2000**, *8*, 1015-1023.
(235) Spillane, C. B.; Smith, J. A.; Morgan, J. L.; Keene, F. R. DNA affinity binding studies using a fluorescent dye displacement technique: The dichotomy of the binding site. *J. Biol. Inorg. Chem.* **2007**, *12*, 819-824.
(236) Tanious, F. A.; Veal, J. M.; Buczak, H.; Ratmeyer, L. S.; Wilson, W. D. DAPI (4', 6-diamidino-2-phenylindole) binds differently to DNA and RNA: Minor-groove binding at AT sites and intercalation at AU sites. *Biochemistry* **1992**, *31*, 3103-3112.
(237) Wilson, W. D.; Tanious, F. A.; Barton, H. J.; Jones, R. L.; Fox, K.; Wydra, R. L.; Strekowski, L. DNA sequence dependent binding modes of 4', 6-diamidino-2-phenylindole (DAPI). *Biochemistry* **1990**, *29*, 8452-8461.
(238) Lewis, F. D.; Zhang, L.; Kelley, R. F.; McCamant, D.; Wasielewski, M. R. A perylenedicarboxamide linker for DNA hairpins. *Tetrahedron* **2007**, *63*, 3457-3464.
(239) Zheng, Y.; Long, H.; Schatz, G. C.; Lewis, F. D. Duplex and hairpin dimer structures for perylene diimide-oligonucleotide conjugates. *Chem. Commun.* **2005**, 4795-4797.
(240) Baumstark, D.; Wagenknecht, H.-A. Fluorescent hydrophobic zippers inside duplex DNA: Interstrand stacking of perylene-3,4:9,10-tetracarboxylic acid bisimides as artificial DNA base dyes. *Chem. Eur. J.* **2008**, *14*, 6640-6645.

(241) Lyng, R.; Rodger, A.; Nordén, B. The CD of ligand-DNA systems. I. Poly (dG-dC) B-DNA. *Biopolymers* **1991**, *31*, 1709-1720.
(242) Lyng, R.; Rodger, A.; Nordén, B. The CD of ligand-DNA systems. 2. Poly (dA-dT) B-DNA. *Biopolymers* **1992**, *32*, 1201-1214.
(243) Gutschmidt, A. *Masterarbeit: Molekulare Bedeutung von Survivin für die DNA-Reparatur*; Universität Duisburg-Essen: Essen, 2011.
(244) Rosorius, O.; Heger, P.; Stelz, G.; Hirschmann, N.; Hauber, J.; Stauber, R. H. Direct observation of nucleocytoplasmic transport by microinjection of GFP-tagged proteins in living cells. *Biotechniques* **1999**, *27*, 350–355.
(245) Clamme, J. P.; Azoulay, J.; Mély, Y. Monitoring of the formation and dissociation of polyethylenimine/DNA complexes by two photon fluorescence correlation spectroscopy. *Biophys. J.* **2003**, *84*, 1960-1968.
(246) Correa, G. T. B.; Veranio, G. A. C.; Silva, L. E.; Hirata Junior, R.; Coil, J. M.; Scelza, M. F. Z. Cytotoxicity evaluation of two root canal sealers and a commercial calcium hydroxide paste on THP1 cell line by Trypan Blue assay. *J. App. Oral Sci.* **2009**, *17*, 457-461.
(247) Hansma, H. G. Surface Biology of DNA by atomic force microscopy. *Annu. Rev. Phys. Chem.* **2001**, *52*, 71-92.
(248) Dong, X.; Wang, X.; He, Y.; Yu, Z.; Lin, M.; Zhang, C.; Wang, J.; Song, Y.; Zhang, Y.; Liu, Z.; Li, Y.; Guo, Z. Reversible DNA condensation induced by a tetranuclear nickel(II) complex. *Chem. Eur. J.* **2010**, *16*, 14181-14189.
(249) Berne, B. J.; Pecora, R. *Dynamic light scattering*; Wiley-VCH: New York, 1976.
(250) Zhong, Q.; Inniss, D. Fractured polymer / silica fiber surface studied by tapping mode atomic force microscopy. *Surface Sci. Lett.* **1993**, 688-692.
(251) Rivetti, C.; Guthold, M.; Bustamante, C. Scanning force microscopy of DNA deposited onto mica: Equilibration versus kinetic trapping studied by statistical polymer chain analysis. *J. Mol. Biol.* **1996**, 919-932.
(252) Dean, R. T.; Jessup, W.; Roberts, C. R. Effects of exogenous amines on mammalian cells, with particular reference to membrane flow. *Biochem. J.* **1984**, 27-40.
(253) Bayer, E. Towards the chemical synthesis of proteins. *Angew. Chem. Int. Ed.* **1991**, *30*, 113-129.
(254) Thakkar, A.; Cohen, A. S.; Connolly, M. D.; Zuckermann, R. N.; Pei, D. High-throughput sequencing of peptoids and peptide-peptoid hybrids by partial Edman degradation and mass spectrometry. *J. Comb. Chem.* **2009**, *11*, 294-302.
(255) Rapoport, H.; Crowley, J. I. Cyclization via solid phase synthesis. Unidirectional Dieckmann products from solid phase and benzyl triethylcarbinyl pimelates. *J. Am. Chem. Soc.* **1970**, *92*, 6363-6365.
(256) Jung, G.; Beck-Sickinger, G. Multiple peptide synthesis methods and their applications. *Angew. Chem. Int. Ed.* **1992**, *31*, 367-383.
(257) Nicolaou, K. C.; Pfefferkorn, J. A.; Mitchell, H. J.; Roecker, A. J.; Barluenga, S.; Cao, G.-Q.; Affleck, R. L.; Lillig, J. E. Natural product-like combinatorial libraries based on privileged structures. 2. Construction of a 10 000-membered benzopyran library by directed split-and-pool chemistry using NanoKans and optical encoding. *J. Am. Chem. Soc.* **2000**, *122*, 9954-9967.
(258) Rink, H.; Sieber, P.; Raschdorf, F. Conversion of NG-urethane protected arginine to ornithine in peptide solid phase synthesis. *Tetrahedron Lett.* **1984**, *25*, 621-624.
(259) Brunfeldt, K. *Peptides 1980: Proceedings of the sixtheenth European Peptide Symposium: Helsingor Denmark August 31 - September 6, 1980*; Scriptor: Kopenhagen, 1981.
(260) Carpino, L. A.; Shroff, H.; Triolo, S. A.; Mansour, E.-S. M. E.; Wenschuh, H.; Albericio, F. The 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl group (Pbf) as

arginine side chain protectant. *Tetrahedron Lett.* **1993**, *34*, 7829-7832.
(261) Steward, J. M.; Young, J. D. *Solid-phase peptide synthesis*; Pierce Chemical Company: Rockford, 1984.
(262) Fields, C. G.; Fields, G. B. Minimization of tryptophan alkylation following 9-fluorenylmethoxycarbonyl solid-phase peptide synthesis. *Tetrahedron Lett.* **1993**, *34*, 6661-6664.
(263) Kamaly, N.; Kalber, T.; Ahmad, A.; Oliver, M. H.; So, P.-W.; Herlihy, A. H.; Bell, J. D.; Jorgensen, M. R.; Miller, A. D. Bimodal paramagnetic and fluorescent liposomes for cellular and tumor magnetic resonance imaging. *Bioconjugate Chem.* **2008**, *19*, 118-129.
(264) Chattopadhyay, A.; London, E. Fluorimetric determination of critical micelle concentration avoiding interference from detergent charge. *Anal. Biochem.* **1984**, *139*, 408-412.
(265) Kurreck, J. RNA interference: From basic research to therapeutic applications. *Angew. Chem. Int. Ed.* **2009**, *48*, 1378-1398.
(266) Schmuck, C.; Geiger, L. Design and synthesis of a new class of arginine analogues with an improved anion binding site in the side chain. *Chem. Commun.* **2005**, 772-774.
(267) Leonard, J.; Lygor, B.; Procter, G. *Praxis der Organischen Chemie*; Wiley-VCH: Weinheim, 1996.
(268) Hesse, M.; Meier, H.; Zeeh, B. *Spektroskopische Methoden in der Organischen Chemie*; Georg Thieme Verlag Stuttgart: New York, 1991.