Synthesis and Characterisation of Amphiphilic 3'-Peptidyl-RNA Conjugates

Inauguraldissertation
zur
Erlangung der Würde eines
Doktors der Philosophie

vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Silvia Terenzi

aus Arcevia (Italien)

Basel, 2003

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät der Universität Basel auf Antrag der Herren Professoren:

Prof. Dr. Peter Strazewski
Prof. Dr. Andreas Pfaltz
Prof. Dr. Thomas Kaden

Basel, den 28. Februar 2003

Prof. Dr. Marcel Tanner
Dekan

The work presented was carried out under the guidance of Prof. Dr. Peter Strazewski at the Institute of Organic Chemistry of the University of Basel between June 1999 and February 2003.

I would like to thank my Ph.D. supervisor, Prof. Dr. Peter Strazewski, for his guidance and help during the preparation of this work and for his counsel regarding matters of a chemical and also a non-chemical nature.

I would like to acknowledge my collegue of laboratory Quang for his help and suggestions during the work and also all the people of the Institute for rendering us available many of their facilities without which this work would have not been possible to achieve.

I would like to thank Dr. Anthony Coleman and Dr. Patrick Shahgaldian (Institut de Biologie et Chimie des Protéines, Lyon) for their collaboration and help with AFM imaging, and Profs. Paul Jenö and Thomas Kiefhaber (Biocentre, University of Basel) for making available their, respectively, MALDI-ToF and CD spectrometer facilities.

Table of Contents

List of Abbreviations

A	alanine
a.a.	amino acid
aa-tRNA	aminoacyl-transfer RNA
AFM	atomic force microscopy
Ala	alanine
AMP	aminomethyl polystyrene
Arg	arginine
BOC	*tert*-butyloxycarbonyl
C-	carboxy-
calcd.	calculated
CBMIT	[1,1'-carbonyl-*bis*(3-methylimidazolium) triflate]
CD	circular dichroism
CPG	controlled-pore-glass
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCE	1,2-dichloroethane
DEPBT	[3-(diethoxyphosphoryloxy)-1,2,3-benzotriazin-4(3*H*)-one]
DIEA	diisopropylethylamine
DLS	dynamic light scattering
DMAP	*N,N*-dimethylaminopyridine
DMF	*N,N*-dimethylformamide
DMT	4,4'-dimethoxytrityl
eq.	molar equivalent
ESI-MS	electrospray ionisation mass spectrometry
FMOC	9-fluorenylmethoxycarbonyl
Glu	glutamic acid
h	hour(s)
HATU	*O*-(7-azabenzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate
HBTU	*O*-(benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate
MALDI-ToF MS	matrix-assisted laser desorption-ionisation time-of-flight mass spectrometry

min	minutes
N-	amino
mRNA	messenger RNA
NMI	*N*-methylimidazole
NMM	*N*-methylmorpholine
O.D.	optical density
PEG	polyethyleneglycol
Pmc	2,2,5,7,8-pentamethylchroman-6-sulphonyl
PS	polystyrene
Pyr	pyridine
quant.	quantitative
R	arginine
R_f	retention factor
RP-HPLC	reversed-phase high performance liquid chromatography
rRNA	ribosomal RNA
r.t.	room temperature
SAX-HPLC	strong anion exchange high performance liquid chromatography
sat.	saturated
SPPS	solid phase peptide synthesis
TBAF	tetrabutylammonium fluoride
TBDMS	*tert*-butyldimethylsilyl
TEA	triethylamine
TFA	trifluoroacetic acid
THF	tetrahydrofurane
t_R	retention time
tRNA	transfer RNA
TOM	2'-*O*-{[(triisopropylsilyl)oxy]methyl}
UV	ultraviolet

1. INTRODUCTION

1.1 The Ribosome: structure and function

1.1.1 A brief overview

Ribosomes play a central role in biological protein synthesis. They translate the genomic information encoded in messenger RNAs into proteins. They are large ribonucleoproteins and conserve the same architecture in all branches of the tree of life. They invariably consist of a large and a small subunit, the former being roughly twice the molecular mass of the latter. The two subunits, which can be reconstituted in vitro from their components,[1] have distinct functions. The small subunit, which sediments at 30S in prokaryotes, mediates the interactions between messenger RNA codons and the transfer RNA anticodons (decoding site). The large subunit, which sediments at 50S in prokaryotes, is the site of peptide bond formation (peptidyl transferase centre). The substrates for this reaction are an aminoacyl-tRNA (aa-tRNA) located in the ribosome's A-site and a peptidyl-tRNA (located in the P-site). In the course of the reaction, the α-amino group of the aa-tRNA attacks the carbonyl group of the ester-bound peptide, forming an amide bond and transferring the peptide to the A-site bound tRNA (Fig. 1a). At every cycle, this reaction lengthens the nascent peptide chain by one residue. The association of the 30S and the 50S subunits forms the complete ribosome (70S) (Fig. 1b).

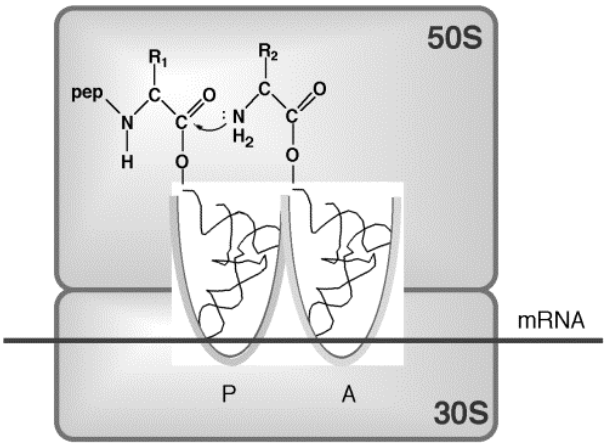


Fig. 1a: Schematic representation of protein synthesis on the ribosome.

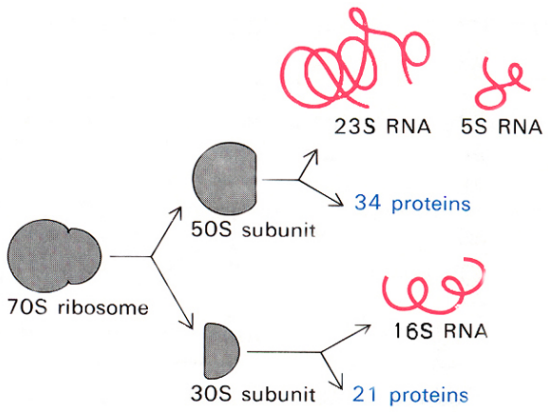


Fig. 1b: Elements forming the ribosome.

1.1.2 Determination of the ribosome structure

For many years, it was presumed that ribosomal proteins orchestrate protein synthesis and that ribosomal RNAs (rRNAs) serve primarily as a structural scaffold. Little by little, biochemical experiments produced data contrasting with this hypothesis,[2] in particular, the fact that ribosomes almost depleted of proteins can still catalyse the formation of peptide bonds.[3] Despite many findings supporting the possibility that rRNA catalyses the peptide bond formation, no direct proof was available because the complete removal of proteins resulted in loss of tertiary folding of rRNA leading to inactivation of peptidyl transferase. It was clear that a three dimensional structure of the ribosome was necessary to elucidate the mechanism of protein biosynthesis at a molecular level.

Electron microscopy techniques allowed to identify the morphological features of ribosomes but the low resolution of these structures could not provide insights into the ribosome function.[4] Recently, the X-ray crystal structure of the large ribosomal subunit at nearly atomic resolution has been published.[5] This structure, at 2.4 Å resolution, contributes to the elucidation of many features of the ribosome. However, to elucidate the catalytic mechanism of enzymes, it is necessary to resolve the structure of them complexed with substrates or substrates analogs. This is what has been done by Steitz and coworkers; they determined the structure of *Haloarcula marismortui* 50S subunit complexed either with the so-called 'Yarus analog' (CCdA-p-Puromycin) or with a mini-helix analog of a puromycyl-tRNA.[6] The Yarus molecule (Fig. 2),[7] an analog of the anionic tetrahedral intermediate in amide bond formation, can bind to the P-site so its position allowed to locate precisely the peptidyl transferase centre.

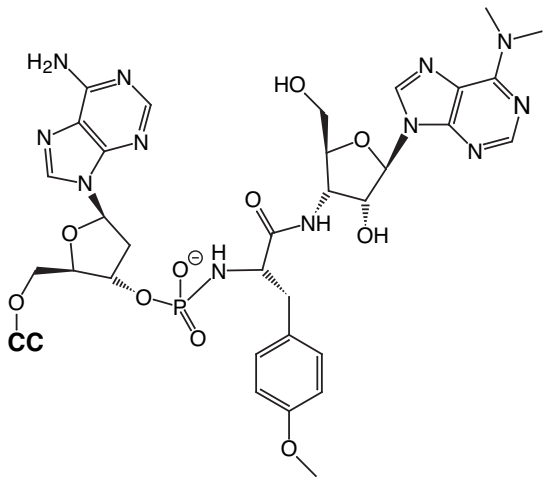


Fig. 2: Yarus analog.

The most important information obtained from these structures is that proteins are largely absent from the regions of the subunit that are of primarily functional importance for protein synthesis: the subunit interface (where the 50S interacts with the 30S subunit) and the peptidyl tranferase active site.

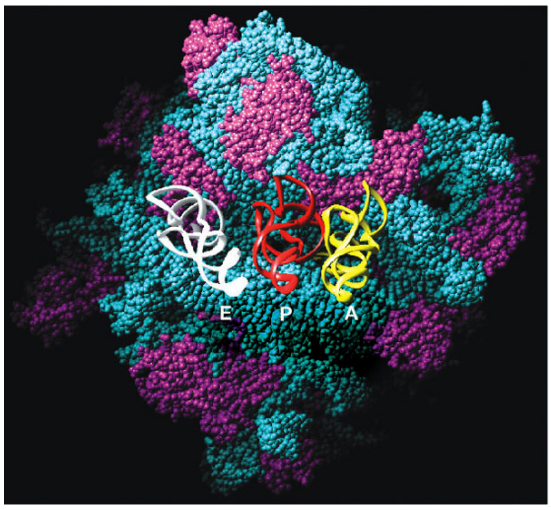


Fig. 3: 50S subunit interface with the three docked tRNAs molecules (ribbon format).The proteins are in pink and the rRNA in blue.

Only RNA lies in the vicinity of the reaction centre (Fig. 3), the nearest protein being at more than 18 Å from it. These data prove that peptide bond synthesis is performed by an RNA catalyst, without the direct involvement of proteins. In other words, the ribosome is a ribozyme. The authors suggested also a possible catalytic mechanism for peptide bond formation that has already been disproved by mutagenesis experiments.[8] A question may be raised concerning the postulated mechanism based on the structure. Does this crystallographic structure describe the active, catalytic state, or an inactive ground state? The ribosome is probably the most sophisticated molecular machine ever made by Nature. All of its components are in movements and the models produced by electron microscopy and X-ray crystallographic studies might be too static for a deep comprehension of this machinery.

1.2 The RNA world

The finding that the ribosome is essentially a ribozyme reinforced the hypothesis of an RNA world, a world that would have predated the protein and DNA one (Fig. 4).
This hypothesis was first formulated by Gilbert in 1986 after the discovery of RNA molecules which can performe catalytic activities.[9] These molecules were called ribozymes, RNA molecules with enzyme-like activity. Despite containing only four different chemical subunits, RNA can fold into a variety of complex tertiary structures rendering it able to catalyse a broad range of chemical reactions.[10]

In the hypothesised RNA world, RNA served both as the genetic material and the metabolic enzymes, probably assisted by metal ions, amino acids and other small molecular cofactors. First, the RNA molecules catalysed their own replication. In the next stage, RNA began to synthesise proteins, which emerged as superior enzymes because, having a larger repertoire of functional groups, are more versatile. Finally, DNA appeared and replaced RNA as the genetic material because its double-stranded structure renders it more stable and reliable as storing material than single-stranded RNA. In addition, the absence of 2'-hydroxyl groups renders DNA less prone to fold into complex tertiary structures which would complicate the read-out process of genetic information.

The ribosome may be a kind of molecular fossil in our world and the first ribosomes were, likely, composed entirely of RNA. The evidence for this is that the functional core of the modern ribosomes consists primarily of RNA.

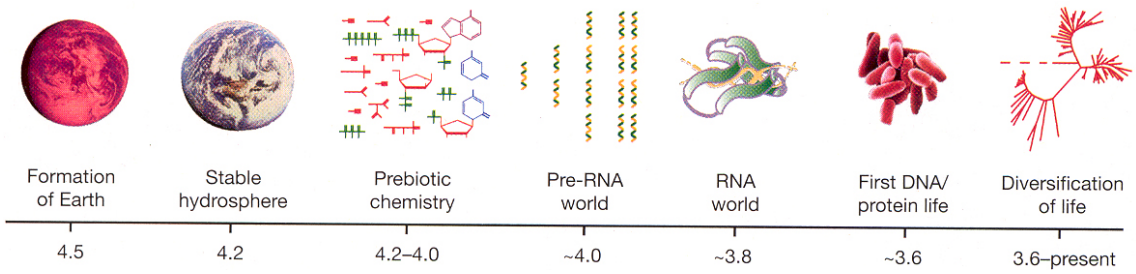


Fig. 4: Timeline of the early history of life on Earth (Ref. 9c).

In the RNA world hypothesis only RNA plays a central role excluding other molecules, such as amino acids or lipids, from the early steps of life. In our view, this is a strongly simplified theory because it does not take into account all the possible molecules that could have played also an important role. Prebiotic experiments and carbonaceous material extracted from meteorites gave evidence for the presence of amino acids in larger quantities than nucleobases, not to speak about nucleosides. Furthermore, it has been demonstrated by Harada that amino acids can spontaneously condense by simple heating through a dehydration process forming a series of random polymers, polyamides called 'proteinoids'.[11] It cannot be completely excluded that some of these polymer sequences had some kind of function, although they could not replicate and their synthesis was not controlled by RNA, meaning that they did not contain any hereditary information in their sequence. In other words, the proteins that come into play, according to Figure 4, after the emergence of an RNA world are strictly only those that have been synthesised under the control of RNA. All proteinoids have been ignored for simplification.

1.3 Self-assembling properties of amphiphilic compounds and their implication in the origin of life

The RNA world hypothesis, even if fascinating, cannot explain alone the origin of life. The hypothesis of an acellular RNA world where the RNA molecules performed both the metabolic and the genetic functions, does not explain how, in free bulk solution, these molecules could reach a concentration high enough to be in physical contact to react. Life, as we know it today, is based on a living unit called "cell". A physical boundary, the membrane, separates the living organism from the outside world. The origin of life cannot be understood without taking into account the origin of membranes.[12] The formation of a closed structure (compartmentalisation) is a key step in the evolution of life because it prevents dilution of the reacting molecules through encapsulation into vesicles. All known cells have membranes composed of amphiphilic lipids.

Amphiphilic molecules, with a hydrophilic 'head' and a hydrophobic 'tail', can self-assemble in water into a closed bilayer structure. The chain length of the hydrophobic part determines whether a given amphiphile can self-assemble into a stable membrane or not. The longer the chain, the stronger the hydrophobic interaction, forming thus more stable supramolecular structures. How did the first membranes originate? Prebiotic lipid-like molecules capable of forming membrane-bounded vesicles have been found in carbonaceous meteorites like the Murchinson meteorite.[13] It is possible to assume that the mixture of organic compounds present in meteorites, resembles components available on the primitive Earth, either through extraterrestrial infall or by synthetic processes occuring at the Earth's surface. Since it is quite sure that membrane-forming molecule were available on the early Earth, a problem arise: How could the first replicating molecules be captured and encapsulated?

The self-assembly of amphiphilic molecules leading to the formation of membranes, gives rise to three distinct regions: the inside and outside aqueous environment and the membrane itself, which is hydrophobic. This hydrophobic environment could serve to capture and concentrate the least polar molecules. A hypothesis on how the first replicating molecules could have been recruted by membranes, has been formulated by Blobel and Cavalier-Smith.[14] The probably early prebiotic amino acids are mostly hydrophobic ones. Thus, peptides formed by these amino acids, would have been capable of inserting into a bilayer membrane. The peptides we mention here, are not spontaneously polymerised, but their synthesis is RNA controlled: a sort of genetic inheritance is already coded in their sequence.

Amphipathic proto peptidyl-tRNAs formed by a hydrophobic peptide and a hydrophilic nucleotide strand, could anchor to the outer surface of membranes through the peptidic part. In this way, RNA replicases could become physically associated. In this hypothesis, the first vesicles functioned as capturing devices for metabolic molecules. These first vesicles, called also 'obcells' or 'inside-out cells', could then evolve through invagination and fission of the membrane giving rise to the first proto-cells with an internal cytosol within which a water-soluble metabolism developed.

A cell, as any living system, must be able to self-sustain, reproduce and transmit its characters. Consequently, the membrane cannot be completely impermeable, but it must allow access to the nutrients so that the cell can survive and, at the same time, it must provide a barrier sufficient to encapsulate the macromolecules responsible for its metabolism. The first living cellular systems were unlikely to have evolved specialised membrane transport systems like modern cells. Typical nutrients like amino acids and phosphate are charged molecules and they cannot pass across a phospholipid membrane through passive transport.[15] A solution to this problem could be to shorten the chain length of the lipids composing the bilayer, in this way the permeability of the membrane to ions can be increased due to a major number of transmembrane transient defects.[16] Another way to circumvent the problem of membrane permeability, is to hypothesise the formation of peptides capable of forming pores into the membranes. It has been demonstrated that very simple peptides of only 21-residues formed by leucine and serine (early prebiotic amino acids) are able to form an amphipathic α-helix spanning a lipid bilayer and functioning as an ion channel.[17]

If the macromolecules encapsulated into a vesicle are supplied with nutrients, they can replicate so that the content of the vesicle grows; this implies that the vesicle also must be able to grow. In the absence of the complex machinery that controls the growth and division of modern cells, the first proto-cells had to be able to grow and divide only relying on the physico-chemical properties of vesicles. Luisi and coworkers demonstrated that vesicles can grow by spontaneous incorporation of new lipid present in solution.[18] When a vesicle has reached a certain size, it becomes thermodynamically unstable and shear forces cause it to divide. The two new entities are structurally similar to the progenitor vesicle so that we can call this process a self-reproduction mechanism. Such a process, spontaneous growth and division, could lead to a primitive cell cycle controlled only by the biophysical properties of the membrane.

Compartmentalisation is a key process in the evolution of life and understanding all the physico-chemical laws that drive these kinds of processes would be of great utility to the

comprehension of the origin of cellular life and to try to synthesise an artificial cell. At the same time, trying to find new kinds of molecules, other than lipids, able to self-assemble, could suggest other pathways for the origin of the first proto-cells. In this scenario, peptidyl-RNA are the least understood players.

1.4 tRNA: a key molecule

1.4.1 Structure and function

Transfer RNA (tRNA) is a central molecule in protein biosynthesis. It is the adaptor molecule for the twenty proteinogenic α-amino acids. The primary structure of a tRNA molecule was first determined by Holley in 1965.[19] He provided the sequence of yeast alanine tRNA. tRNAs are small molecules containing between 76 and 93 ribonucleotides base-paired into a secondary structure called the cloverleaf structure.

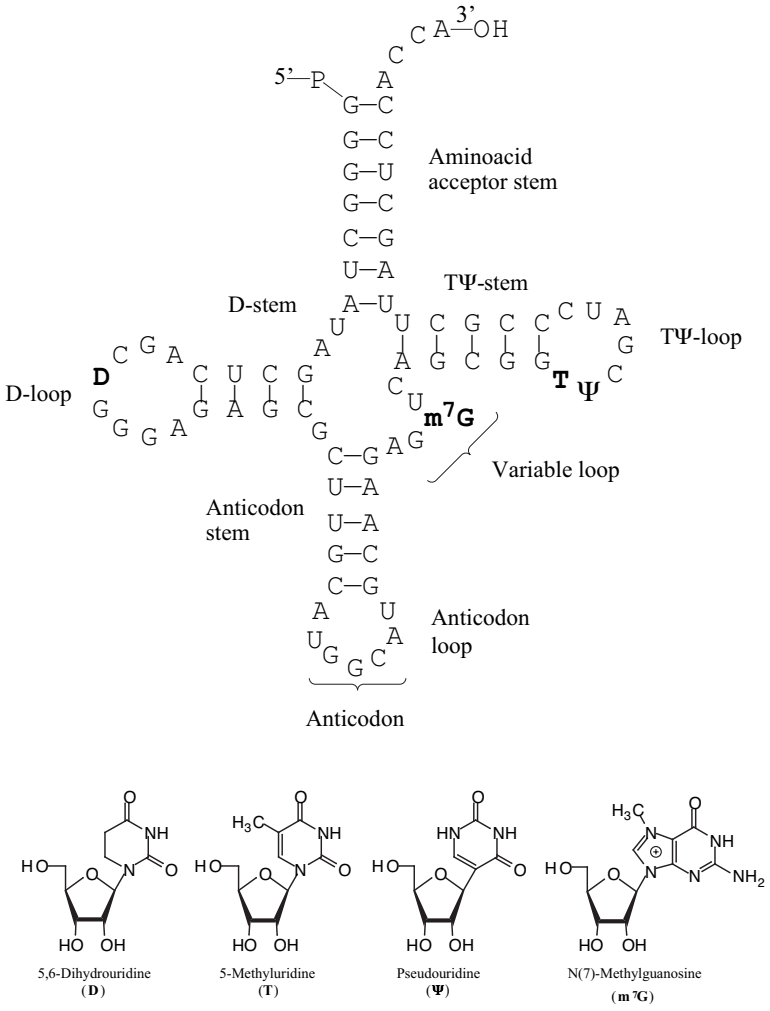


Fig. 5: Secondary structure of *E.coli* $tRNA^{Ala}$.

The secondary structure shows four base-paired stems and four loops (Fig. 5). Two sets of unpaired nucleotide residues at each end of the structure are especially important for the function of tRNA. One is the anticodon loop containing the anticodon triplet that base-pairs with the complementary triplet (codon) of the mRNA molecule. The other is the 3'-terminal end consisting of four single-stranded nucleotides (N-C-C-A: nucleotide-cytidine-cytidine-adenosine), the so-called CCA terminus. The 3'-terminus is the site of attachment of a specific amino acid corresponding to the anticodon. All tRNAs are phosphorylated at the 5'-end and contain many modified nucleotides.

The 3D structure of a tRNA molecule was first solved in 1973 for a yeast phenylalanine tRNA.[20] It folds into a tertiary structure that resembles a 3-dimensional capital L (Fig. 6). Two segments of double helix perpendicular to each other give the molecule its L shape. The CCA-terminus is at one end and the anticodon loop is located at the other end of the L shaped molecule. Today we know that all tRNAs fold into a similar tertiary structure.

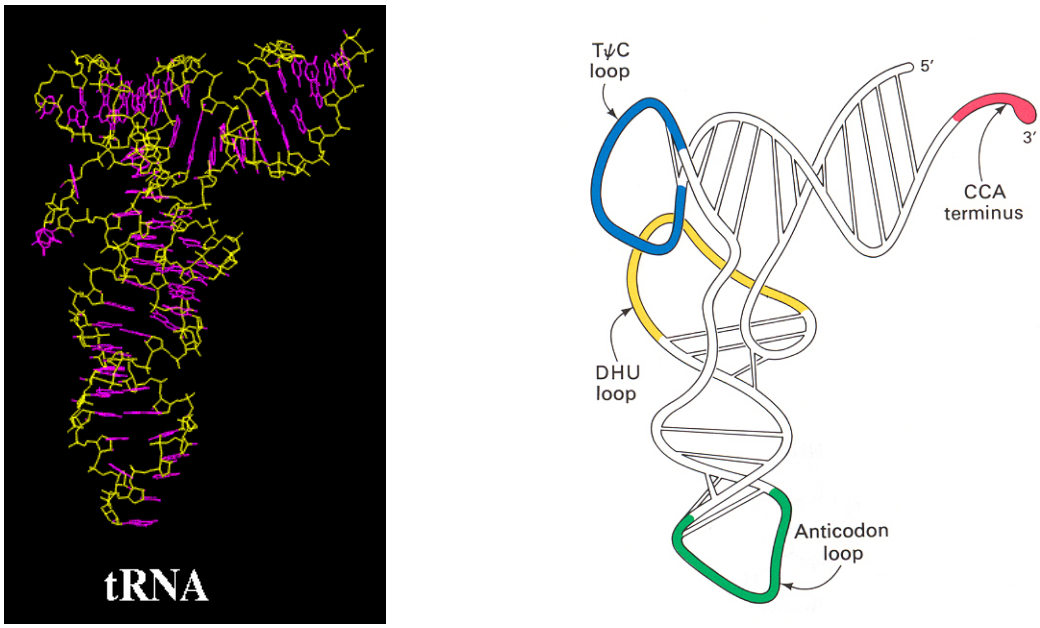
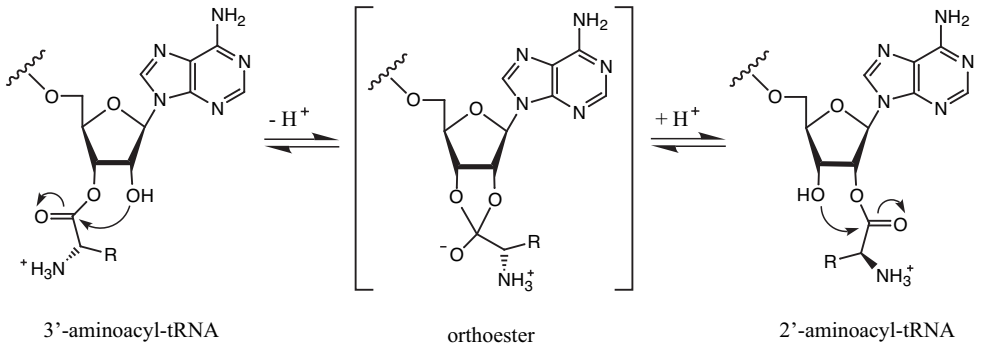


Fig. 6: Tertiary structure of tRNA.

The biologically active form of tRNA is aminoacylated at its 3'-end. The amino acid is linked to the 3'-terminal adenosine through an ester bond with either the 2' or the 3'-hydroxyl group of the ribose.[21] This ester bond is an activated, energy-rich bond with a free energy of hydrolysis comparable to ATP hydrolysis.[22] It can hydrolyse within minutes to hours in a neutral aqueous solution at room temperature. In addition, aminoacyl-tRNA undergoes fast

2'-3' transesterification due to the presence of the vicinal *cis*-hydroxyl groups on the terminal adenosine (Scheme 1).



Scheme 1: 2'-3' transesterification of aminoacyl-tRNA.

Both *N*-acylation and the absence of a neighbouring 2'-OH group can stabilise the ester bond of aminoacylated nucleoside towards hydrolysis.[23] Simultaneous *N*- and 2'-*O*-protection results in a 70-fold increase of stabilisation at 25 °C and pH 7.4.[24] The lability of the ester bond linking the amino acid and the tRNA renders very difficult to perform structural studies (NMR, X-ray crystallography) on activated aa-tRNA and peptidyl-tRNA. However, it is of the most importance for understanding the translational mechanism, to know how the ribosome interacts with its substrates, the tRNAs and mRNA. Only the knowledge of all the molecular contacts between tRNA and the ribosome, can provide a structural framework for the elucidation of the protein synthesis mechanism.

1.4.2 Cocrystallisation of tRNA analogs with the ribosome

A cocrystallisation of all the functional elements with the ribosome is necessary to achieve a deeper knowledge on the mechanism of protein synthesis. Efforts in this direction have been done but, so far, not one structure has been solved with all the elements present at the same time. Although much progress has been made with the crystal structure of the ribosome solved by Steitz and coworkers[5], many questions remain to be addressed. In particular, the determination of the structure of a peptidyl-tRNA or an analog bound to the P-site would be of great interest.

Most of the times, only analogs of aa-tRNAs have been used by crystallographers and microscopists[6,25]. Only small molecules have been used as peptidyl-tRNA analogs for the cocrystallisation, but low molecular weight P-site substrates need high concentrations of

alcohol to bind, which is decidely non physiological.[26] Recently, a longer fragment of a peptidyl-tRNA analog, which does not require alcohol for the binding, has been cocrystallised with the 50S subunit.[27] This molecule is formed by a short RNA fragment (–CCA) attached via an ester linkage to a phenylalanine whose α-amino group is linked to biotin via a capronic acid moiety (Fig. 7).

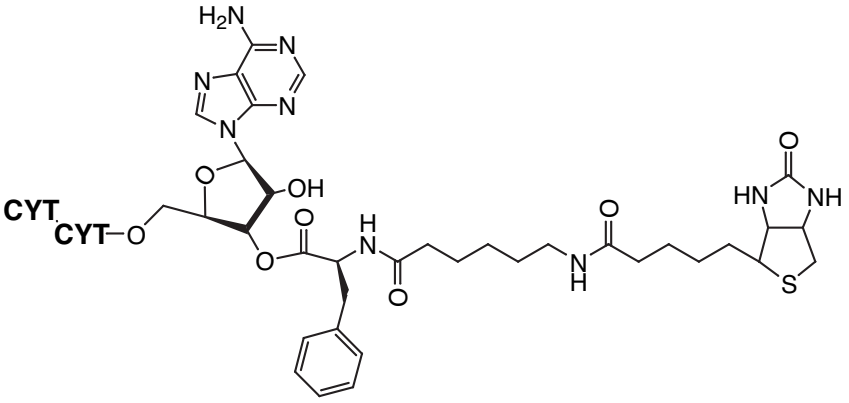


Fig. 7: Peptidyl-tRNA analog.

The capronic acid chain should mimic the growing peptide chain. However, this molecule is still far away from a native peptidyl-tRNA. A peptidyl-tRNA or an analog with a longer peptidic chain would be necessary to identify the path of the nascent peptide through the large ribosomal subunit. To date, peptidyl-tRNAs synthesised *in situ* on *E. coli* ribosomes, have been used in cross-linking studies.[28] In these experiments, the α-amino group of the terminal residue is derivatised with a photoreactive group which, after irradiation, crosslinks to RNA. Depending on the length of the peptide, the *N*-terminus will be linked to different RNA residues, rendering possible the identification of the pathway of the nascent peptide.

The only example using a peptidyl-tRNA for structural studies was published in 2001 by Blobel.[29] In this study, a long peptidyl-tRNA was assembled *in vitro* and reconstituted with the ribosome elements; subsequently, its structure was determined by cryo-EM. Even if this work is a considerable advancement for the elucidation of the peptide pathway through the exit tunnel, it has two major drawbacks: first, low resolution (15.4 Å) and, second, low yield of material (the process, scaled up, allowed to have only 0.25 OD_{260} (~5 pmol) of product). From these examples, it emerges the necessity to use peptidyl-tRNA analogs that are more stable (to allow X-ray analysis) and achievable in larger amounts (via chemical synthesis), see Section 1.8: «Aim of the thesis: synthesis and characterisation of 3'-peptidyl-RNAs».

1.5 Peptide secondary structure

The amino acid sequence defines the primary structure of a peptide. Any sequence of amino acids adopts a three dimensional structure, more or less ordered. The local conformation a polypeptide chain adopts is called secondary structure and is stabilised through backbone hydrogen bonds. Local secondary structures interact together to form a tertiary structure. Only long peptide chains can adopt a tertiary structure that is stabilised by hydrophobic forces, disulfide or salt bridges between amino acid side chains.

From the analysis of a portion of a polypeptide backbone (Fig. 8), it emerges that the bond between the carbonyl group and the nitrogen is not free to rotate because it has a partial double-bond character. A high-energy barrier must be overcome to allow a rotation about this bond (~25 kcal/mol). Conversely, the groups linked to C^{α} are free to rotate. Conventionally, rotation around the N-C^{α} bond is described by the tortion angle ϕ and that around the CO-C^{α} by φ. Normally, at room temperature, peptide bonds are in *trans* conformation (ω=180 °).

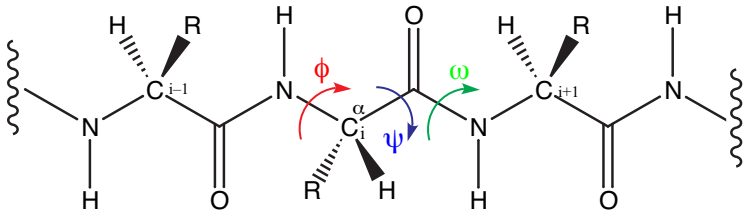


Fig. 8: Peptide backbone.

Ramachandran was the first to plot the permitted values of ϕ and φ using a hard-sphere model of the atoms.[30] The two dimentional map of these values is known as Ramachandran plot.

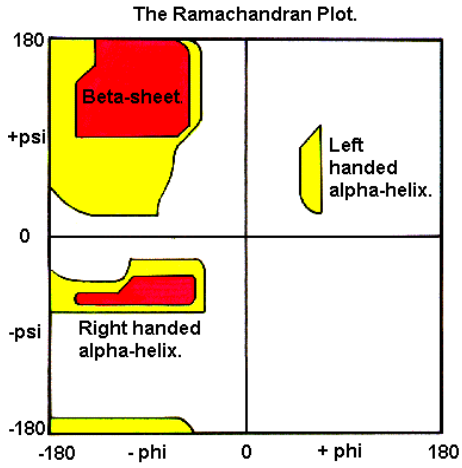


Fig. 9: The Ramachandran plot.

Certain combinations of angles are inachievable because of steric hindrance. For all amino acids, except glycine and proline, such plot shows three separate, allowed, low energy, regions. How do peptides look like in these conformations? A polypeptide in solution can adopt a completely disordered structure called random coil or it can fold into a regularly repeating structure.

1.5.1 α-Helix and β-Sheet structures

The most common secondary structures peptides adopt are α-helix and β-sheet conformations. The helix conformation (called alpha because it was the first one to be modeled), was proposed by Pauling in 1951.[31] In solution several types of helical conformations can exist, α-helix being the most common motif found in proteins. It has 3.6 residues per turn and a specific pattern of hydrogen bonds stabilising the helix. The hydrogen bonds are between the CO of residue n and the NH of residue n+4 (Fig. 10).

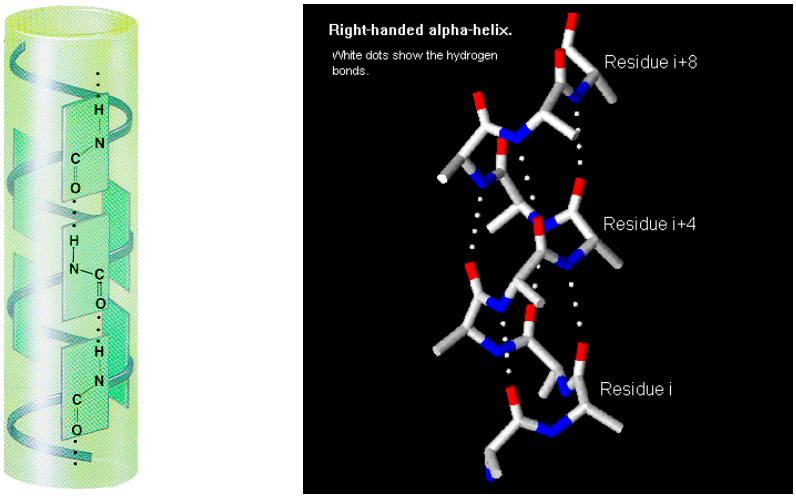


Fig. 10: Schematic representation of α-helix secondary structure.

The hydrogen bonds are almost parallel with the helix axis and have the same direction; consequently the dipole of each peptide bond cumulates, giving the helix a macrodipole moment: the amino end corresponds to the positive part of the dipole, the carbonyl end to the negative. Proline residues are incompatible with α-helical conformation and can be found only in very long, distorted helices.

β-sheet conformation is formed by parallel or antiparallel β-strands which consist of polypeptide chains almost fully extended (Fig. 11).

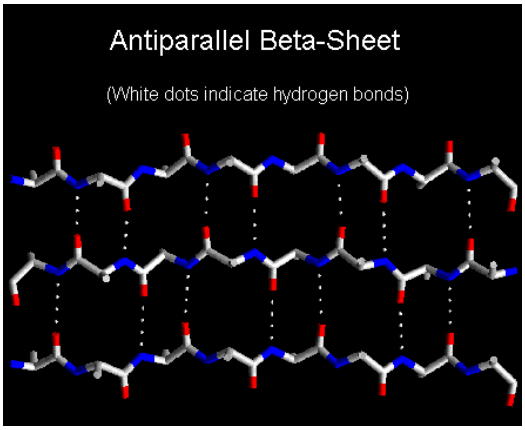


Fig. 11: Schematic representation of β-sheet secondary structure.

Like the α-helix, the β-sheet conformation is mantained by backbone hydrogen bonds. These interactions take place among residues that are distant in their primary structure whereas in the α-helix the hydrogen bonds are established between residues of the same strand that are close in the primary structure. In a β-strand the side chain groups are alternatively above and below the plane of the structure, in a minimum hindered conformation while in the α-helix the R groups point outward of the helix, tilted toward the amino end. Another secondary structure motif is the turn. Different types of turn exist and they are very abundant in globular proteins. They allow a polypeptide chain to reverse its direction.

Many studies have been done on the secondary structure motifs of polypeptides. This is a fundamental subject to understand protein folding. The aim of these studies, often, is to construct models able to predict the secondary structure of any polypeptide chain. Of course, the ultimate goal is to be able to predict the tertiary structure of any protein, starting from its primary structure. This has been done successfully with polypeptides up to ~150 residues.[32]

β-sheet conformation is more difficult to study in solution, because it tends to grow indefinitely forming insoluble aggregates. However, understanding the factors that stabilise one conformation over the other, is not only critical to study protein folding, but it is also important for material science and medicine. For example, spider silk, one of the strongest material existing in Nature, consists of β-sheet crystalline segments of polyalanine residues (4-9 amino acids) alternating with amorphous segments composed of varying amino acids, mostly glycine. Scientists have tried to reproduce, not completely successfully, this material both chemically[33] and biologically through recombinant techniques.[34]

In medicine, more and more diseases related to protein and peptide misfolding are emerging, among them Alzheimer's disease, type-II diabetes, Creutzfeldt-Jacobs disease and

many others. Here, a protein undergoes a structural change from its normal secondary structure into a β-sheet rich conformation leading to the formation of irreversible aggregates.[35] It is clear that understanding which forces are responsable for this misfolding, is crucial to propose a clinical treatment of these diseases.

Many studies have focused on the individual preference of certain amino acids to form α-helices. Any residue has a different propensity for α-helices or β-sheets and many model peptides have been used to numerically quantify this propensity. There is an ongoing debate concerning the helix propensity of certain amino acids, in particular alanine.

Theory assumes that the coil-helix transition is a two-step process. It starts from the nucleation step when the first helical hydrogen bond is generated and it is followed by the fast addition of helical residues to the ends of the existing helix (propagation). The nucleation step is the most energetically unfavorable (rate-limiting step) because three residues lose their conformational freedom (entropy reduction) balanced by only one hydrogen bond formation. Zimm and Bragg developed a model to interpret helix-coil transition using two thermodynamic parameters: σ, which describes the nucleation step and s which describes helix propagation.[36] The equilibrium constant between the two conformations for n residues is expressed by Equation 1:

$$K_{(c\rightarrow h)}=\sigma \cdot s^{n} \qquad (1)$$

σ has a value of about $2\cdot10^{-3}$ and s has a value around 1. These values imply that short peptide chains (<100 a.a.) should not form a stable helix. In reality, even very short peptides (<20 a.a.) can form stable helices. Much of the debate is around the value of s for every single amino acid.[37] However, recent molecular dynamic simulation experiments suggested that the helix formation involves other dynamical events such as the accumulation of the 3_{10}-helix or turn;[38] it seems to be a more complex process than the one hypothesised by Zimm and Bragg. Independently from the correct numerical value of the intrinsic tendency of every amino acid to form a helix, it is clear that alanine, valine, leucine and glutamic acid are prone to form α-helices. Algorithms have also been developed to predict the helix content of a sequence of amino acids with good reliability, and, through the web accessible software AGADIR, everyone can estimate the degree of helical content for a peptide of interest.[39]

1.5.2 Analysis of peptide secondary structure

Circular dichroism (CD) spectroscopy is one of the most widely used techniques for studying peptide conformation in solution.[40] The amide chromophore is responsible for the dichroism of peptides and proteins in the far-UV region (180-240 nm), therefore this dichroism reflects the overall structure of the backbone. Even though it does not provide as detailed information as NMR, it can be measured at much lower concentrations, avoiding problems of aggregation.

Figure 12 shows CD spectra of peptides that are presumed to assume a single type of secondary structure in solution.

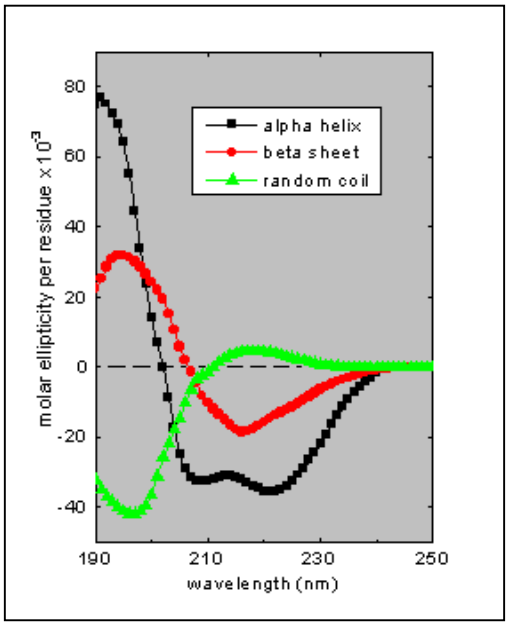


Fig. 12: CD spectra of different polypeptide conformations.

The α-helix gives rise to two negative bands at 222 and 208 nm of almost equal intensities and a strong positive band at about 192 nm. The spectrum of β-sheet conformation shows a negative band near 216 nm and a positive one at 198 nm. The random coil has an intense negative band at 198 nm.

CD spectra of α-helices are the most widely studied. Figure 12 shows an ideal α-helix but the reality can be much more complex because in solution other types of helices can exist (3_{10} helix, α_{II} helix, α_L helix).[41] Theoretical models have also been developed to predict the CD spectra of distorted α-helices and twisted β-sheets.[42] Even though it is difficult to deduce the exact content of every secondary structural element in a polypeptide chain, CD spectroscopy is still the most practical and rapid way to study the overall conformation of peptide and proteins in solution.

1.6 Solid phase synthesis: general features

The principle of all solid phase synthesis is simple. The growing chain, be it peptide, oligonucleotide or other oligomer, is built while it is attached to a solid support. It remains attached to the solid support throughout all the synthetic steps and is separated from soluble reagents and solvents by simple filtration and washing avoiding purification after each step. All the monomers are joined, one by one, to the fixed terminus. During the synthesis, at least two levels of protecting groups are required: the permanent and the temporary protecting groups. The former are used to prevent branching or other problems on the side chains and must withstand repeated applications of the conditions used for the quantitative removal of the temporary protecting groups. The latter mask the site of the chain growth and are necessary to avoid self-condensation of the incoming monomer.

Normally, orthogonal protection schemes are used. This strategy involves using protecting groups that are removed by different chemical mechanisms rendering, therefore, possible the removal of one group in the presence of the other. At the end of the synthesis, the product is detached from the solid support, purified and characterised. A disadvantage of solid phase synthesis, is the need for very high reaction efficiencies if purifiable products are desired. The presence of by-products due to the incompleteness of reactions renders the target product very difficult to purify. This problem is reduced by using a large excess of reagents which helps to drive bimolecular or higher order reactions towards completion.

The solid support, normally an insoluble polymeric matrix, must contain reactive sites at which the chains can be attached and must provide enough points of attachment to give a useful yield of final product. Moreover, it must be chemically and physically stable to the conditions of the synthesis.

A major advantage of solid phase chemistry is that the process can be completely automated. Today, automatic synthesisers are able to perfom routinely complete syntheses of peptides or oligonucleotides in a few hours. This big development of automated methods for solid support synthesis has had a major impact in many fields of biology.

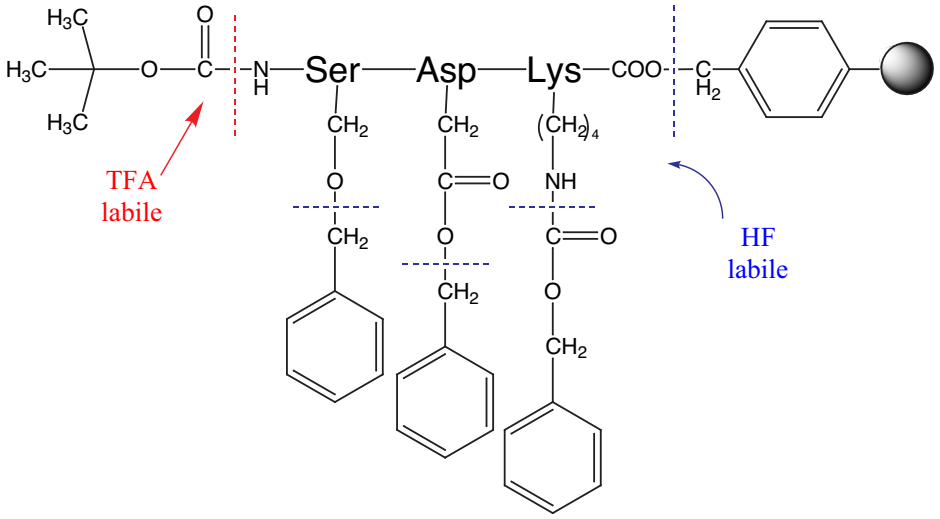
1.6.1 Solid Phase Peptide Synthesis (SPPS)

The method of solid phase peptide synthesis was introduced by Merrifield in 1963.[43] The basic principles of his method have been described in the previous paragraph. The impact and

importance of Merrifield's work was such that he was awarded the Nobel Prize for Chemistry in 1984.

In the method he proposed, the peptide is stepwise synthesised starting from the carboxy-terminus. The C-terminal amino acid is attached to the solid support via its α-carboxylic group. The solid support used by Merrifield was a polystyrene support cross-linked with 1-2% of divinylbenzene. This solid support is still the most widely used in peptide chemistry.

The α-amino group of every amino acid is protected with the acid-labile *tert*-butyloxycarbonyl (BOC) group (temporary protection). The BOC group is easily removed with anhydrous mineral or strong organic acids (usually trifluoroacetic acid). Permanent side chain protecting groups are ether, ester and urethane derivatives based on benzyl alchohol. The peptide is attached to the solid support via a benzyl ester linkage. Treatment with anhydrous hydrogen fluoride (HF) allows the simultaneous cleavage of the peptide from the resin and the deprotection of the side chain protecting groups (Scheme 2).

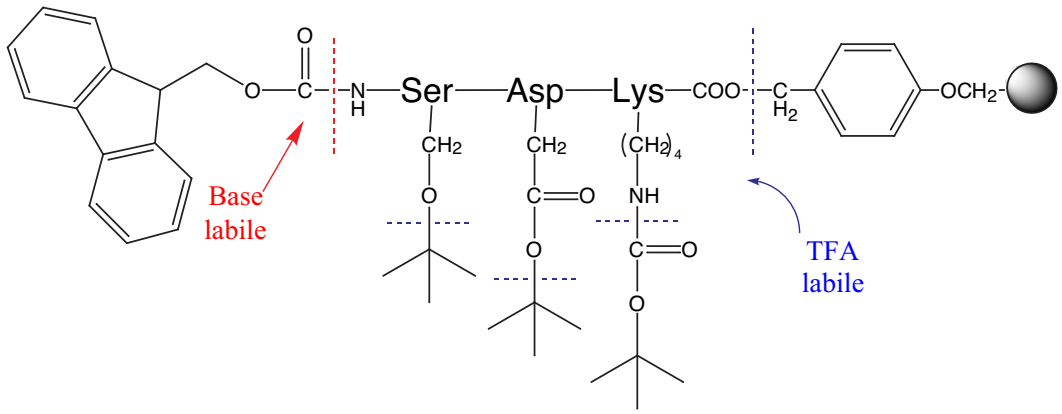


Scheme 2: General protecting group and cleavage strategy for BOC chemistry.

Since cleavage and deprotection in BOC chemistry require the use of dangerous HF, special laboratory apparatus are needed. These harsh conditions used in the procedure introduced by Merrifield were a major drawback that stimulated the development of new synthesis strategies.

The introduction of the base-labile *N*-9-fluorenylmethyloxycarbonyl (FMOC) group as temporary protection for the α-amino group represented a turning-point in solid phase peptide synthesis.[44] The FMOC/*tert*-But strategy was introduced by Atherton and Sheppard.[45] The FMOC group is cleaved with piperidine in DMF. The side chain protecting groups are

prevalently ether, ester and urethane derivatives based on *tert*-butanol. These derivatives are cleaved at the same time as the peptide from the resin by use of TFA (Scheme 3). The standard solid support for the batch synthesis of peptides in FMOC based chemistry, is the Wang resin.[46] It consists of polystyrene beads (1% cross-linked) onto which an acid-labile *p*-hydroxybenzylalchohol linker has been attached.



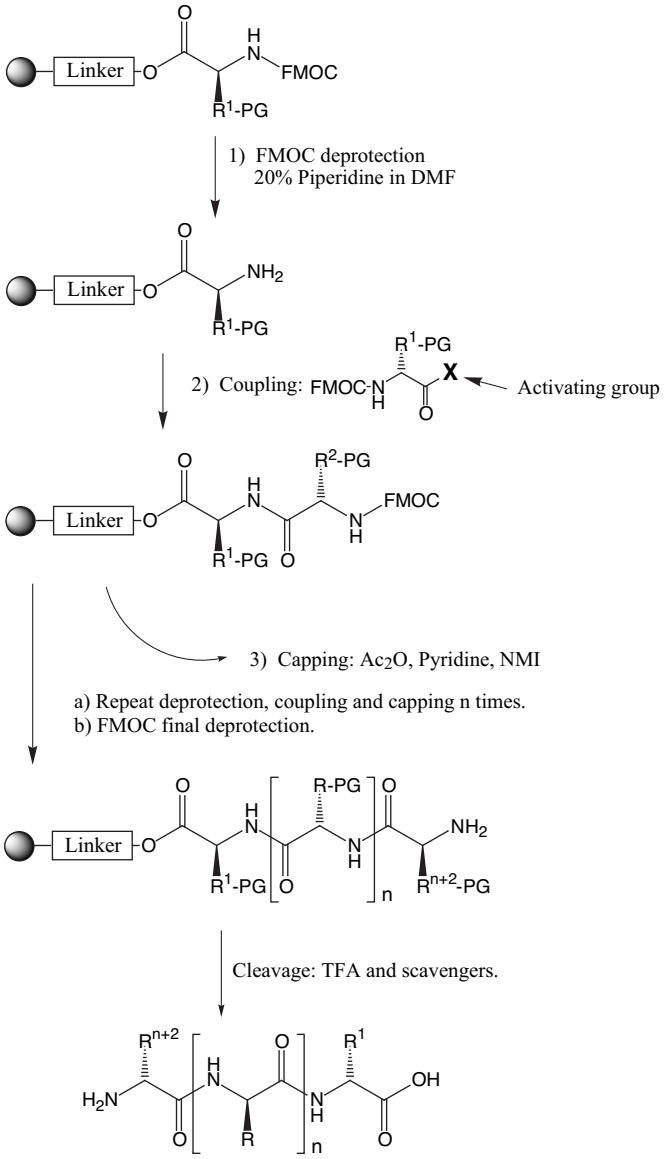
Scheme 3: General protecting group and cleavage strategy for FMOC chemistry.

For our syntheses we used the FMOC chemistry. The standard procedure for peptide synthesis is shown in Scheme 3. The first step of every cycle is the FMOC deprotection of the solid-support linked peptide or amino acid to provide a free amino-terminus. Coupling works as follows: an N^{α}-protected amino acid is contacted with an activating reagent and is converted into a highly reactive species (usually an active ester) which reacts with the support-bound terminal NH_2 group. After the amide bond formation, the chain is one residue longer. Normally, coupling time is between 30 min and 1 hour, depending on the activator and on the sequence of the peptide to synthesise. Unreacted amino groups are acetylated to avoid any further reaction of these groups in the continuation of the synthesis. Deprotection of the FMOC group of the newly coupled amino acid begins another cycle.

When the chain elongation is complete, the FMOC group of the last amino acid is removed and the peptide can be cleaved from the resin. A TFA treatment (1-2 hours) permits the deprotection of all the side chain protecting groups and the detachment of the peptide from the solid support.

The cleavage is a critical step in peptide synthesis. When removed, the linker and the protecting groups form reactive carbocations that can irreversibly modify sensitive amino acids (methionine, cysteine, tyrosine, tryptophane). The use of scavengers (H_2O, 1,2-ethandithiol, triisopropylsilane, thiophenol) during the cleavage greatly reduce these

undesirable side reactions.[47] However, the scavenger efficacy depends also on the amino acid sequence of the peptide. It is therefore recommended to perform preliminary small-scale cleavages to find the best conditions.



Scheme 4: General scheme of SPPS.

1.6.2 Synthesis monitoring

The difficulty of analytical control of BOC SPPS is one of its major drawbacks. Until the introduction of FMOC-chemistry, the most important monitoring test was the so-called "Kaiser test".[48] This test uses ninhydrin to detect the presence of unreacted primary amino groups. The appereance of a deep purple colour indicates the presence of unreacted groups.

The product formed during the reaction absorbs light in the visible region (λ_{max}=570 nm) so that the test can be performed qualitatively or quantitatively. It requires to be carried out manually, so it is not routinely performed to check the completeness of a coupling.

The advent of FMOC-chemistry introduced the possibility of direct spectrophotometric monitoring of the synthesis because the fluorenyl group has a strong absorbance in the ultraviolet region. The deprotection reaction results in the release of the amine and dibenzofulvene; the latter can further react with excess piperidine to form a piperidine adduct.

Measuring the absorbance of the deprotection solution at λ=300 nm, gives indication of the completeness of the coupling taking the first deprotection as a reference value. Atherton and Sheppard were the first to develop a method to monitor automatically a synthesis. They built a prototype machine attached to the synthesiser.[49] Even though automatic spectrophotometrical monitoring of FMOC deprotection is not so accurate as a quantitative measurement, it is a rapid way to check the quality of a synthesis.

1.6.3 Solid phase oligonucleotide synthesis

Nowadays, oligonucleotide synthesis is totally automatised, like peptide synthesis. Phosphoramidite chemistry is the standard method and it is based on protected 3'-nucleoside phosphoramidites, as illustrated in Figure 13.

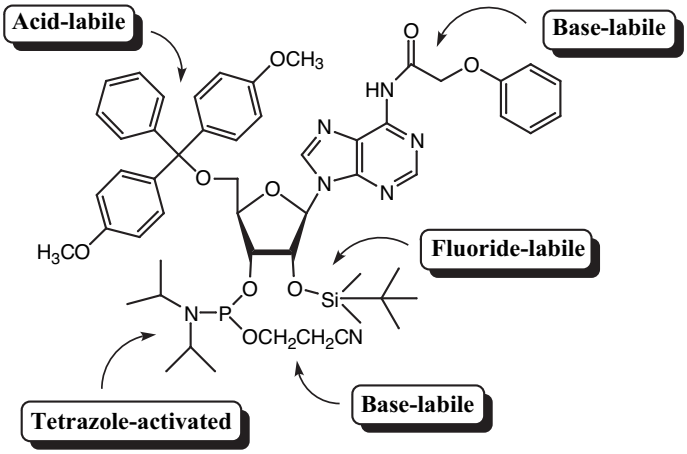


Fig. 13: Building block for RNA solid phase synthesis.

The chemical synthesis is normally carried out in the 3'- to 5'-direction to take advantage of the higher reactivity of the primary 5'-hydroxyl group which is protected with

the 4,4'-dimethoxytriphenylmethyl group (DMT or dimethoxytrityl). The DMT group is easily removable under mild acidic conditions (3% Cl_3CCOOH in CH_2Cl_2). The DMT, as temporary protecting group, is removed at every cycle to allow the coupling of a new monomer.

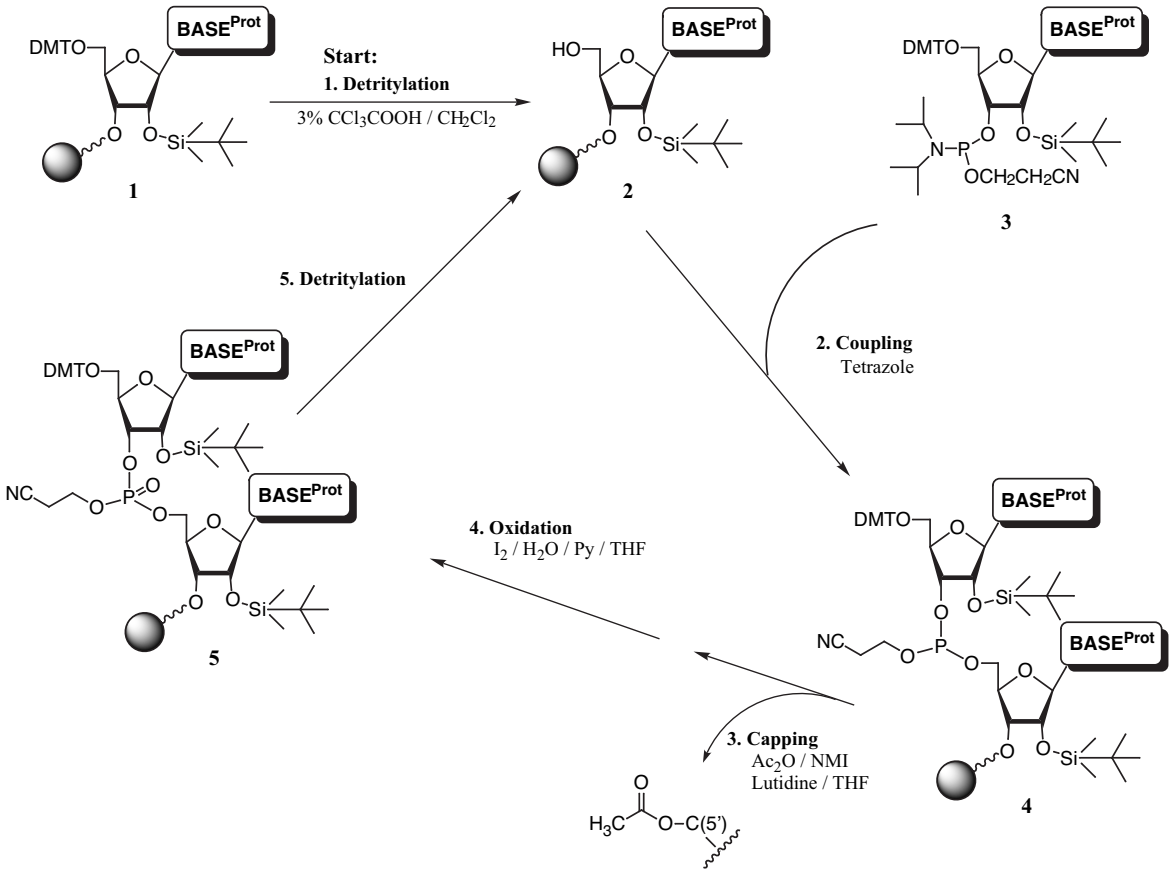
The exocyclic amine function of the bases needs to be permanently protected throughout the synthesis. The standard method developed by Khorana for the synthesis of oligodeoxyribonucleotides, requires the use of acyl protecting groups such as *N*-acetyl, *N*-benzoyl, *N*-isobutyryl.[50] However, a problem seemed to be the relatively low stability of the glycosidic linkage of *N*-benzoyl protected 2'-deoxyadenosine under acidic conditions. This has caused the development of alternative protecting groups, as amidine, in order to stabilise the glycosidic bond thereby reducing the rate of depurination during the cleavage of DMT.[51]

The power methodology developed for DNA synthesis can, in principle, be also applied to the synthesis of RNA. Compared to DNA, however, RNA contains an additional 2'-hydroxyl group, which is responsible for the instability of RNA under basic conditions (pH>12, 25 °C) and which, of course, has to be protected. This protecting group must remain intact throughout the synthesis and its removal must be the last deprotection step in such conditions that will not promote the cleavage or migration of the internucleotide linkage. Moreover, the 2'-*O*-protecting group sterically interferes with the coupling process rendering the synthesis of long RNA strands less efficient than DNA synthesis. Many 2'-OH protecting groups have been reported.[52] The fluoride-labile *tert*-butyldimethylsilyl (TBDMS) group, introduced in 1974 by Ogilvie and coworkers has found the widest application.[53] Recently, Pitsch introduced a new fluoride-labile 2'-protecting group, the 2'-*O*-{[(triisopropylsilyl)oxy]methyl} (2'-*O*-TOM).[54] This group, less steric hindered than the TBDMS group, allows coupling yields comparable to that of DNA synthesis (~99.4%). In this way, it is possible to synthesise longer and purer RNA strands. Due to their features, TOM-protected-ribonucleoside phosphoramidites are superceeding the standard TBDMS-protected monomers. In the phosphoramidite method,[55] the 3'-hydroxyl group carries a diisopropyl-phosphoramidite group protected by the base-labile cyanoethyl group.

1.6.4 The condensation cycle for RNA synthesis

A condensation cycle is illustrated in Scheme 5, which shows all the stages of solid phase synthesis of a dimer. The first nucleoside is attached by means of a linker to a solid support. In the first step of the synthesis, the support-bound nucleoside is detritylated to provide a free

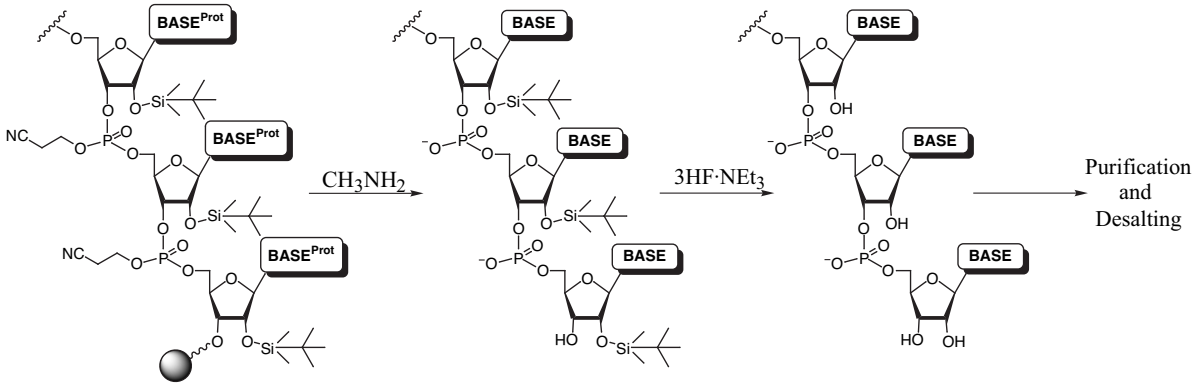
5'-hydroxyl group for the attachment of the second nucleoside. The incoming phosphoramidite is activated using large excess of tetrazole. The free 5'-OH group attacks the highly reactive intermediate to form the dinucleosidyl triphosphite **4**. The coupling time used for RNA synthesis varies between 5 and 15 min., while for DNA synthesis is only 1-2 min. A capping step follows the coupling. In this step any unreacted 5'-hydroxyl group is acetylated and therefore rendered inert to any futher reaction. After the capping, the dimer must be oxidised with aqueous iodine to convert the phosphite triester into the more stable phosphate triester. After oxidation, detritylation of the phosphotriester **5** starts another cycle.



Scheme 5: Condensation cycle in RNA synthesis.

At the end of the oligoribonucleotide synthesis, the first deprotection step is the removal of the 5'-terminal DMT-group (DMT-off). The final deprotection steps are shown in Scheme 6. Treatment of the resin-bound protected RNA with 40% methylamine in water or ethanol, allows the simultaneous removal of the protecting groups of the bases, of the phosphates and the cleavage of the molecule from the solid support. The persilylated RNA is then treated with $3HF \cdot NEt_3$ to cleave the 2'-hydroxyl protecting groups.[56] The crude product, precipitated from *n*-butanol, is purified by ion exchange and reverse phase HPLC. A final

extensive desalting by reverse-phase adsorbtion/desorbtion provides material that can be safely stored at 4 °C for a long time.



Scheme 6: Deprotection steps in RNA synthesis.

Compared to DNA, RNA strands are extremely sensitive to degradation by ribonuclease contamination. Extracellular ribonucleases are ubiquitous and the slightest trace can give rise to degradation of the oligoribonucleotides. The importance of good handling techniques and care in preparation of buffers, reagents, and apparatus cannot be overstressed. The stability of RNA fragments against ribonucleases increases when the 2'-hydroxyl groups are silylated. Therefore, it is better to perform the last deprotection shortly before the HPLC purification.

1.7 Peptide-oligonucleotide conjugates in the literature

Oligonucleotides are widely used as sequence specific reagents to block gene expression. These compounds, called antisense oligonucleotides, bind specifically to a complementary target mRNA strand and inhibit its translation into the corresponding protein.[57] Despite their big potentiality as therapeutic tools and more than one decade of efforts, the first generation of RNA therapeutics is being evaluated only now in clinical trials.[58]

One of their main limitation is their poor cellular uptake. Co-administration of certain additives such as cationic liposomes can enhance the uptake. Yet, use of these reagents is often limited by cellular toxicity.[59] An alternative approach to enhance cellular uptake is to conjugate the oligonucleotide to another molecule known for its ability to penetrate or interact with cell membranes. One of the first ideas in this area was the conjugation of cholesterol, a hydrophobic component of membranes.[60] However, no real improvement of

cellular uptake emerged from the use of cholesterol or other small molecules. Instead, some promising improvements resulted from the conjugation of a variety of peptides to oligonucleotides.[61] Peptide attachment has also other beneficial effects: conjugation at the 3'-end results in significant protection of the oligonucleotide against attack by 3'-exonucleases, which are the major source of nucleases within cells and serum.[62] Moreover, conjugated cationic peptides may increase the stability of hybrids with both single-stranded DNA and RNA targets.[63]

Usually, conjugations to oligonucleotides are carried out at either the 5' or 3' ends to prevent interference with the hybridisation to the complementary target sequence. To date, methods of peptide-oligonucleotide conjugate synthesis are not well established. There have been many routes proposed,[64] but very few have so far proved practical for routine synthesis. The two most important synthetic approaches are the stepwise solid-phase synthesis (total solid-phase conjugate synthesis) and fragment coupling (post-assembly conjugation).

In the fragment coupling approach the two moieties are independently synthesised. After deprotection (and purification if necessary), the two moieties, carrying a reactive functionality on each component, are linked together. The chemoselective reaction can be carried out with the two fragments in solution or with one fragment still bound to the solid support.

Some of the post-synthetic conjugation methods include: (1) the formation of a disulfide bond between a cysteine-containing peptide and a thiol-functionalised oligonucleotide,[65] (2) reaction of a cysteine peptide with a maleimido functionalised oligonucleotide,[66] (3) reaction of a maleimido functionalised peptide with a 5'-thiol-functionalised oligonucleotide,[67] (4) conjugation involving the "native chemical ligation"[68] of a cysteine-containing peptide with a 5'-thioester-functionalised oligonucleotide,[69] (5) "native chemical ligation" of a 5'-cysteine-functionalised oligonucleotide with a thioester peptide,[70] (6) reaction of solid-support bound oligonucleotide, functionalised at its 5'-end with a primary amino group, with a carboxy terminus of a peptide,[71] (7) reaction of an oligonucleotide, amino functionalised at its 3'-end, with the carboxy-end of a peptide,[72] (8) reaction of an incorporated 2'-amino-2'-deoxyribonucleoside with the carboxy-end of a peptide.[73] These methods are schematically illustrated in Figure 14.

The preparation of conjugates following a fragment-coupling methodology suffers from certain disadvantages. For example, a disulfide bond is unstable to reducing agents that may be present under many assay conditions or within the cells. Moreover, a cysteine is required in the peptide sequence. Post-assembly conjugation sometimes suffers from

inefficient conjugation due to the secondary structure or poor solubility of certain peptide components in aqueous solution.

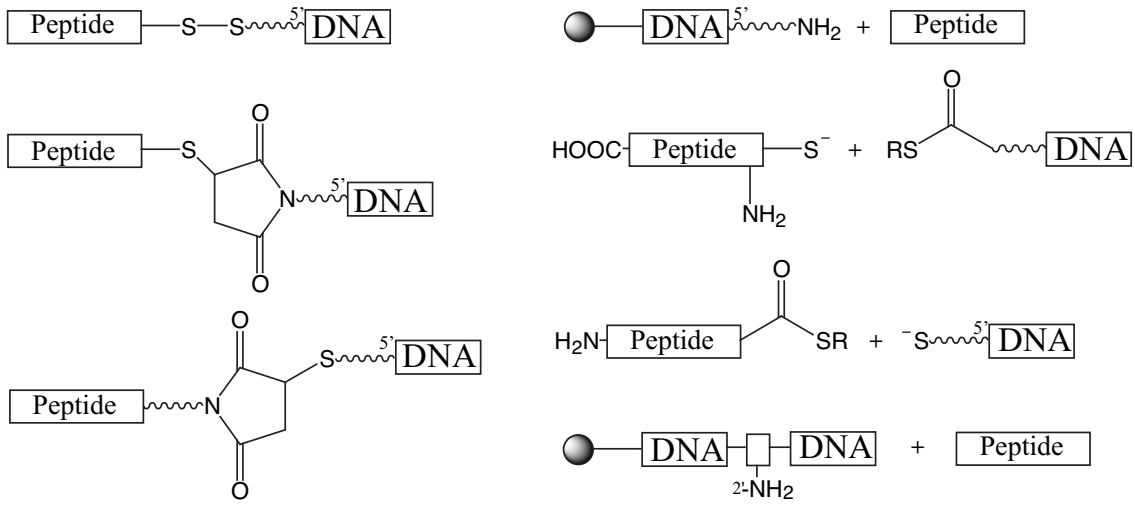


Fig. 14: Some examples of fragment-coupling synthesis.

Alternative to this methodology is the total stepwise solid-phase synthesis. This strategy consists in synthesising the two moieties on the same solid support, in most of the cases the peptidic part is synthesised before the oligonucleotidic part because it requires harsher condition of synthesis. This route is the most direct but it encounters serious difficulties due to the incompatibility of the standard peptide and oligonucleotide protecting groups and assembly chemistries. For example, at the end of the peptide synthesis, a treatment with acid is usually required and can provoke partial depurination of the oligonucleotidic strand. This problem could be solved using N^{α}-BOC-protected amino acids with fluorenyl-based protecting groups for the side chains;[74] in this case the peptide must be synthesised before the oligonucleotide. Unfortunately, these amino acids are not commercial, limiting the application of this method. Despite the difficulties of the total stepwise synthesis, many efforts have been done to develop applicable methods.

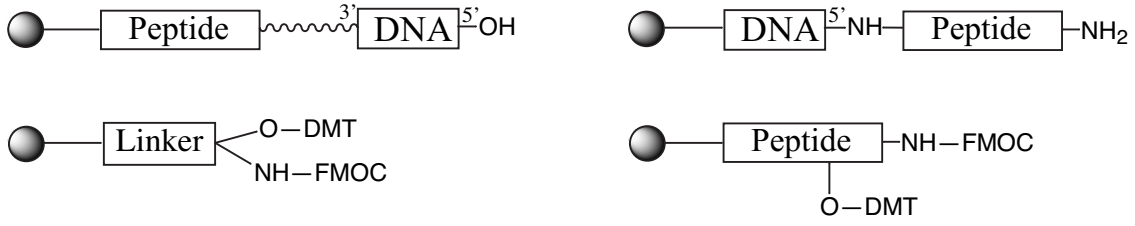
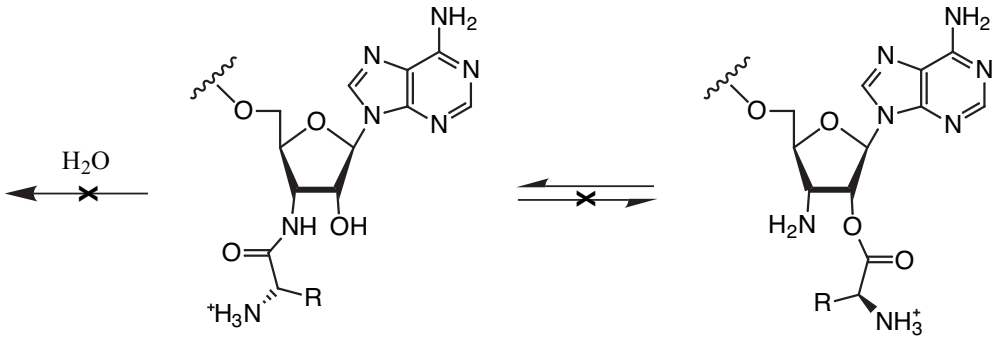


Fig. 15: Some examples of total stepwise synthesis.

Many of these methods envisage the synthesis of the peptide on solid support followed by derivatisation of the N-terminus with a suitable linker bearing an OH-group where the oligonucleotide chain can be grown.[75] Another methodology consists in synthesising the oligonucleotide strand incorporating as the last monomer a modified oligonucleoside: 5'-amino-5'-deoxythymidine. This building block allows a linker-free construction of 5'-peptide-DNA hybrids.[76] Serine or homoserine can be used as attachment site for the oligonucleotide strand. The assembly of the peptide moiety is followed by elongation of the oligonucleotide at the side chain of a hydroxylated amino acid.[77] A more flexible strategy is provided by the use of a double functional linker attached to the solid support. A molecule having a DMT-protected-hydroxyl and a FMOC-protected-amino group, allows the synthesis of the two fragments of the conjugate in any order. This method is becoming very important and, probably, in the future, it will be one of the most widely used because of its flexibility.[78]

1.8 Aim of the thesis: synthesis and characterisation of 3'-peptidyl-RNAs

The aim of this work is to synthesise amphiphilic peptidyl-RNA conjugates and study their physico-chemical properties and their interactions with membranes. These constructs mimic the peptidyl-tRNA molecule so they could be used for a cocrystallisation with the ribosome or for other structural studies (NMR). Replacing the native ester bond linking the tRNA to the nascent peptide with an amide bond renders the molecule stable against hydrolysis, thus suitable for structural experiments.

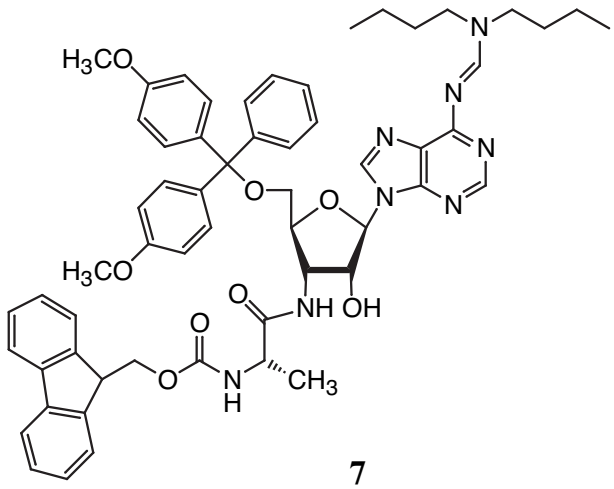


6

Scheme 7

3'-Aminoacyl-3'-deoxyadenosine **6** was used as starting molecule to build our analogs of peptidyl-tRNA. This analog has already been successfully incorporated into the 3'-terminus of tRNA by a combination of synthetic and enzymatic methods.[23,79] However, no chemical method has been developed to synthesise 3'-aminoacyl- and 3'-peptidyl-tRNA analogs using an automated procedure (RNA synthesiser). We tried, successfully, to develop such a procedure.

The starting molecule to synthesise our peptidyl-tRNA analogs is an orthogonally protected derivative of 3'-amino-3'-deoxyadenosine (Fig. 16).

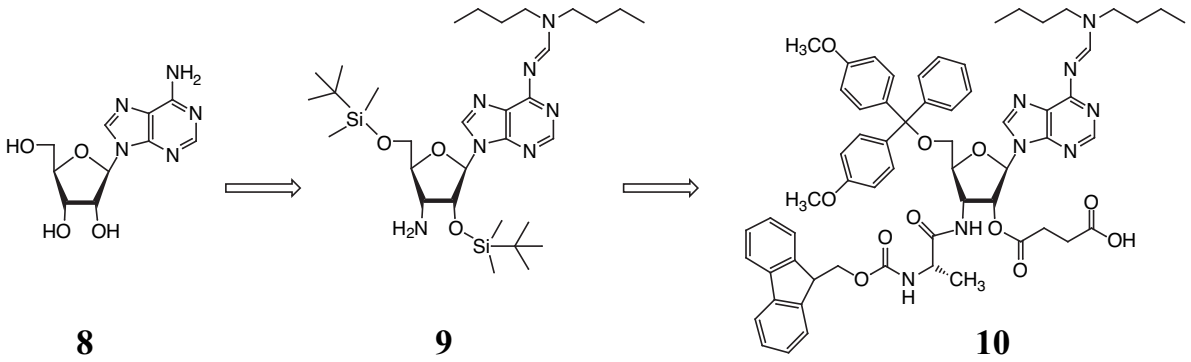


7

Fig. 16: Orthogonally protected derivative of 3'-amino-3'-deoxyadenosine.

This molecule, linked to a solid support through a 2'-succinate linkage, allows the synthesis of a peptide after the deprotection of the base-labile FMOC group. Subsequently, after the synthesis of the peptide, an oligonucleotide strand can be synthesised following the deprotection of the acid-labile trityl group. A careful choice of the protecting groups allows the automatised synthesis of both peptide and oligonucleotide strands.

The synthesis of this molecule was developed in our group by Oliver Botta[80] and Nhat Quang Nguyen-Trung (Scheme 8).[81] The optimisation of some steps of this synthesis is also presented in this work (Section 2.1).

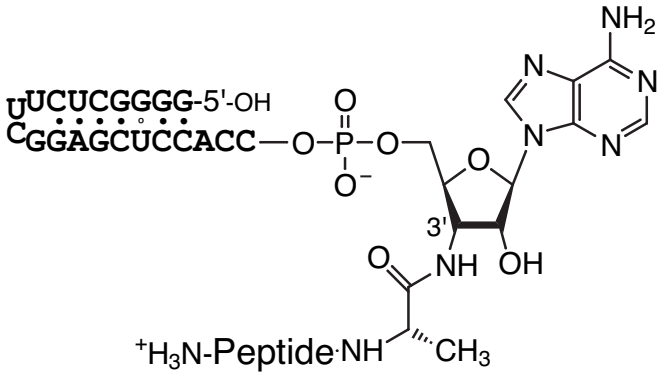


8 **9** **10**

Scheme 8

Adenosine was chosen because is the ubiquitous 3'-terminal nucleotide of all tRNAs. Alanine is the simplest chiral amino acid and no protecting group is needed for its side chain. In principle, any amino acid, with the suitable protecting group, could be incorporated into **9** replacing alanine.

Molecules constituted by an oligonucleotidic strand covalently linked to a peptide are called hybrids or, more commonly, conjugates. The conjugates we planned to synthesise consist of a 22-mer hairpin mimicking the acceptor stem of *E.coli* $tRNA^{Ala}$ closed by the stable tetraloop UUCG and peptidylated at the 3'-end (Fig. 17).



11

Fig. 17: General structure of 3'-peptidyl-RNA conjugates mimicking peptidyl-tRNA.

The peptidyl moieties consist of lipophilic peptides whose sequences are based on polyalanine and polyleucine residues. The peptide sequences were chosen on their propencity to form α-helix secondary structures and, due to the amphiphilic character they give to the molecule, to the possibility of forming higher-order structures.

2. *Results and Discussion*

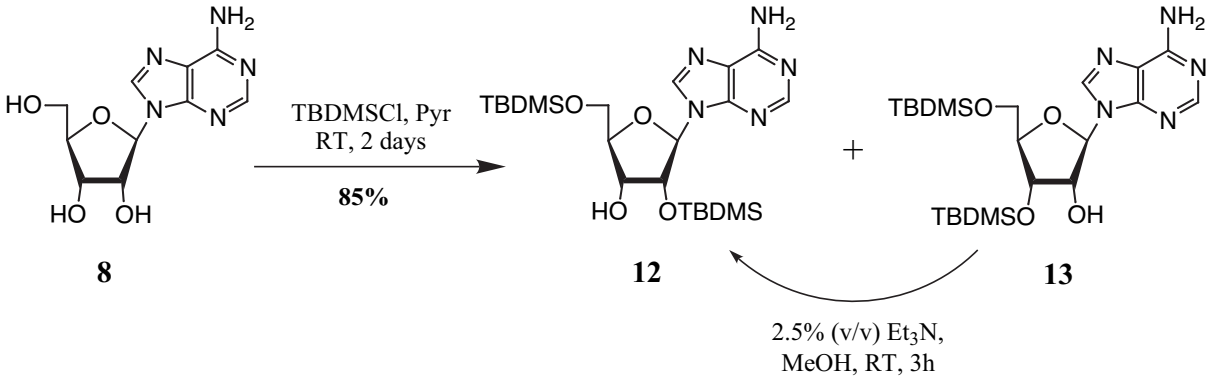
2.1 Synthesis of fully protected 3'-alanylamino-3'-deoxyadenosine (10)

The central molecule for the synthesis of our conjugates is the building block **10**. Its synthesis, starting from adenosine, is achieved through the intermediate compound 3'-amino-3'-deoxyadenosine derivative **9** (Scheme 8).

The first synthesis of 3'-amino-3'-deoxyadenosine was published in 1955 and allowed to obtain the target molecule in less than 2% overall yield.[82] Successively, other groups proposed other synthetic, more efficient pathways.[83] Our group, also, developed a multistep synthesis of 3'-amino-3'-deoxyadenosine analogs through which the building block **10** can be synthesised.[80a, 81a]

2.1.1 Synthesis of 3'-azido-3'- deoxyadenosine

The first step of the synthesis, the protection of the 2'- and 5'- hydroxyl groups of adenosine was performed following the silylation method developed by Ogilvie and coworkers.[84] Adenosine **8** (Scheme 9) was reacted with three equivalents of *tert*-butyldimethylsilyl chloride (TBDMSCl) in pyridine at room temperature giving a mixture of 2',5'-bis-*O*- (60%), 3',5'-bis-*O*- (35%) and 2',3',5'-tris-*O*-silylated adenosine (~5%). Only traces of 5'-*O*-silylated product were found.

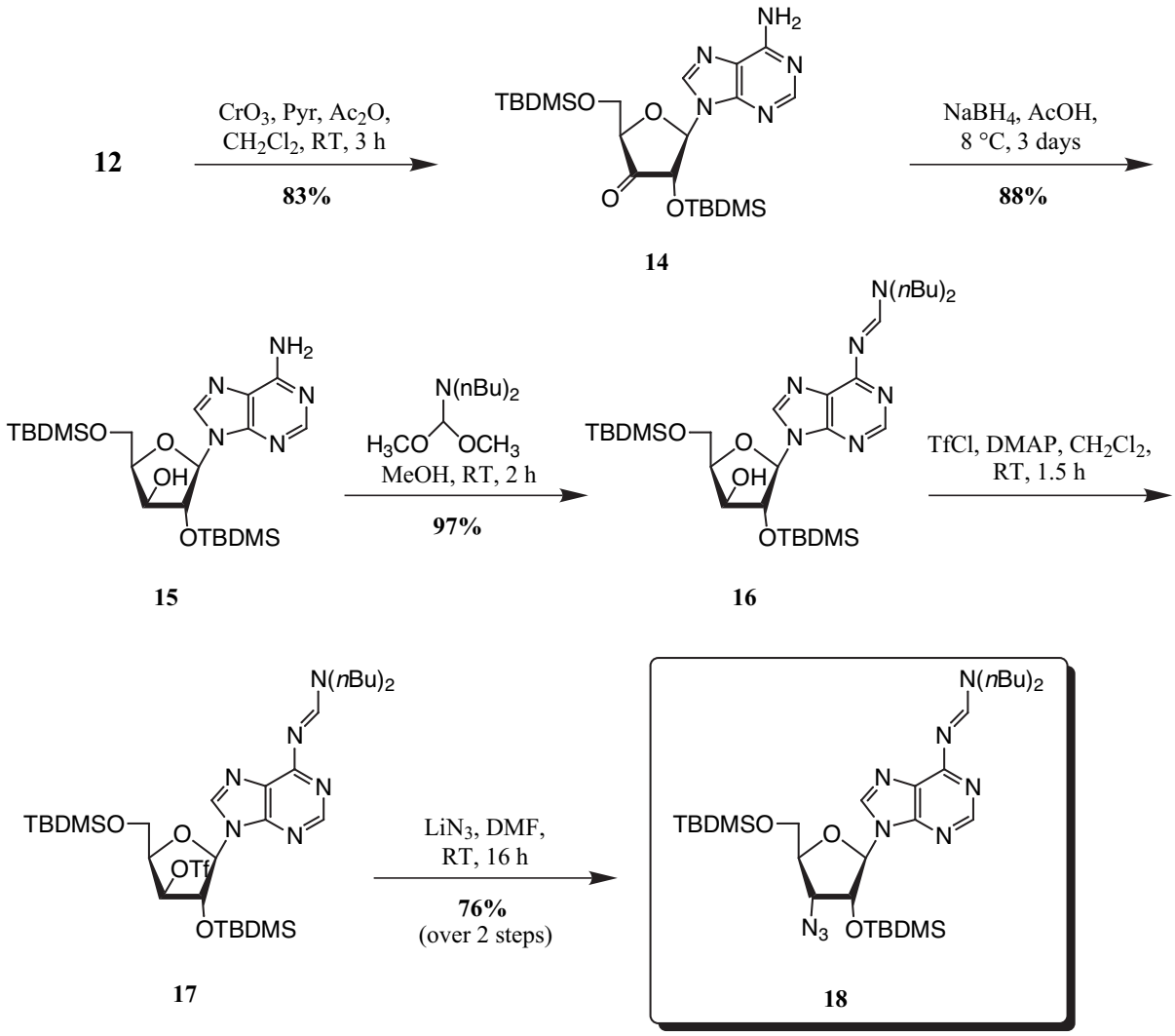


8 **12** **13**

Scheme 9

An isomerisation reaction with 2.5% Et_3N in MeOH can be performed to improve the yield of **12**.[85] This reaction gives an almost equimolar amount of **12** and **13**. Repeating this reaction two or three times, it is possible to reach a high yield of **12** (85%).

The inversion of configuration at the C(3') was achieved through an oxidation/reduction strategy (Scheme 10).



Scheme 10

Ketone **14** was obtained treating **12** with CrO_3/pyridine/Ac_2O complex in CH_2Cl_2.[86] The oxidation reaction was the most troublesome. First, the reaction should be taken to completeness otherwise the separation of the ketone-derivative **14** from the unreacted 3'-ribo-adenosine **12** is problematic due to the very similar polarity of the two compounds. Second, scaling-up the reaction to more than 4-4.5 g is not advisable due to the incompleteness of the reaction using large quantities. Third, the volume of the solvent (CH_2Cl_2) must be carefully adjusted to the quantity of product to oxidise; using too concentrated solutions always led to incomplete reaction after 3 hours necessitating the addition of more CrO_3/pyridine/Ac_2O complex.

The purification of the oxidised compound revealed to be very difficult due to the presence of chromium impurities having a polarity very similar to the target product **14**. This

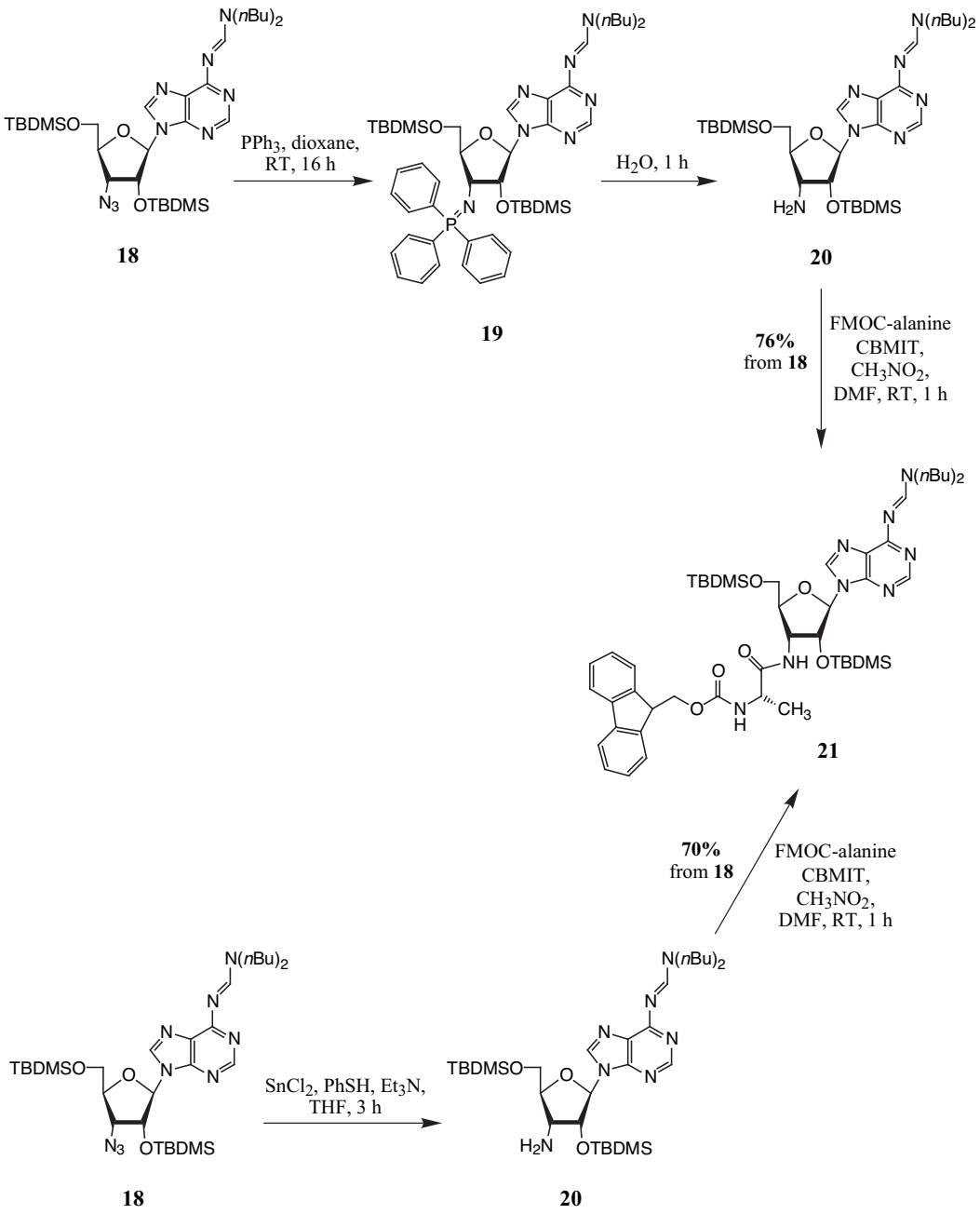
contamination renders the product very unstable and sensitive to light causing rapid degradation after brief exposure to light. These impurities have a negative effect not only on the oxidation reaction yield but also in the following reactions rendering necessary a MPLC purification of the azide derivative **18** before continuing the synthesis. It was clear that an alternative purification procedure to that proposed in literature had to be applied.[86c] The crude material, instead of being precipitated from EtOAc, was directly loaded on top of a column equilibrated with CH_2Cl_2. Isocratic elution with CH_2Cl_2 100% allowed the separation of the chromium-containing impurities (slightly less polar) from **14**. Only after the impurities have been eluted from the column, a step gradient elution with EtOAc/hexane (1:1 to 1:0) allowed the recovery of the target compound as a colorless, stable product.

The reduction of **14** was performed with an excess of sodium triacetoxy borohydride (generated *in situ* from $NaBH_4$ and AcOH at T <15 °C). The reaction is done at 8 °C (external temperature). Attention must be paid not to go below this temperature otherwise the solution freezes (T_{int}<12-13 °C) leading to incomplete reduction of the ketone. This reaction is highly stereoselective due to the attack of the hydride from the α-face (xylo **15** / ribo **12** = ~49/1).

Before introducing the azido group, protection of the exocyclic amino group of the base was necessary to avoid it react with TfCl. A base labile amidine protecting group was chosed because of its higher stability towards acid catalysed depurination (~20-fold more stable compared to N^6-benzoyl protecting group),[87] and because its higher hydrophobicity diminishes solubility problems. The formamidine protecting group was introduced reacting **15** with *N,N*-di-(*n*-butyl)-formamide dimethyl acetal prepared as described in literature.[88]

The introduction of the 3'-azido function was carried out in two steps.[89] First, the 3'-xylo-hydroxyl group was activated by conversion into a sulfonate ester using trifluoromethanesulfonyl chloride (TfCl). The triflate intermediate was not purified because of its low stability and the crude product **17** (after NMR analysis confirmed a good quality of the triflate compound) was used for the successive reaction, the nucleophilic displacement of the triflate by an azido group with inversion of the configuration at the C(3') stereocentre. The 3'-azido derivative **18** has, thus, a ribo-configuration. LiN_3 was used instead of NaN_3 because it is soluble in DMF so that the reaction, conducted in homogeneous phase, is complete within some hours.

2.1.2 Azide reduction and amino acid coupling



Scheme 11

The azido functionality was reduced to amine using a modified Staudinger reaction (Scheme 11).[90] Treating the azide **18** with triphenylphosphine in anhydrous THF, led to the formation of the iminophosphorane **19** which was hydrolysed with H_2O producing the free amine **20** and triphenylphosphine oxide.

The amine derivative **20** was isolated as an oil; in such form degradation of part of the product was observed after having stored it for some time at low temperature. This molecule

contains a base-labile amidine protecting group that can be cleaved intermolecularly by the amino group present in the same molecule. This side reaction can be particularly deleterious when, as in an oily state, a high concentration of the molecule is reached. To avoid this problem, after column chromatography purification, the amine adenosine **20** was not dried under high vacuum but immediately acylated with an activated amino acid.

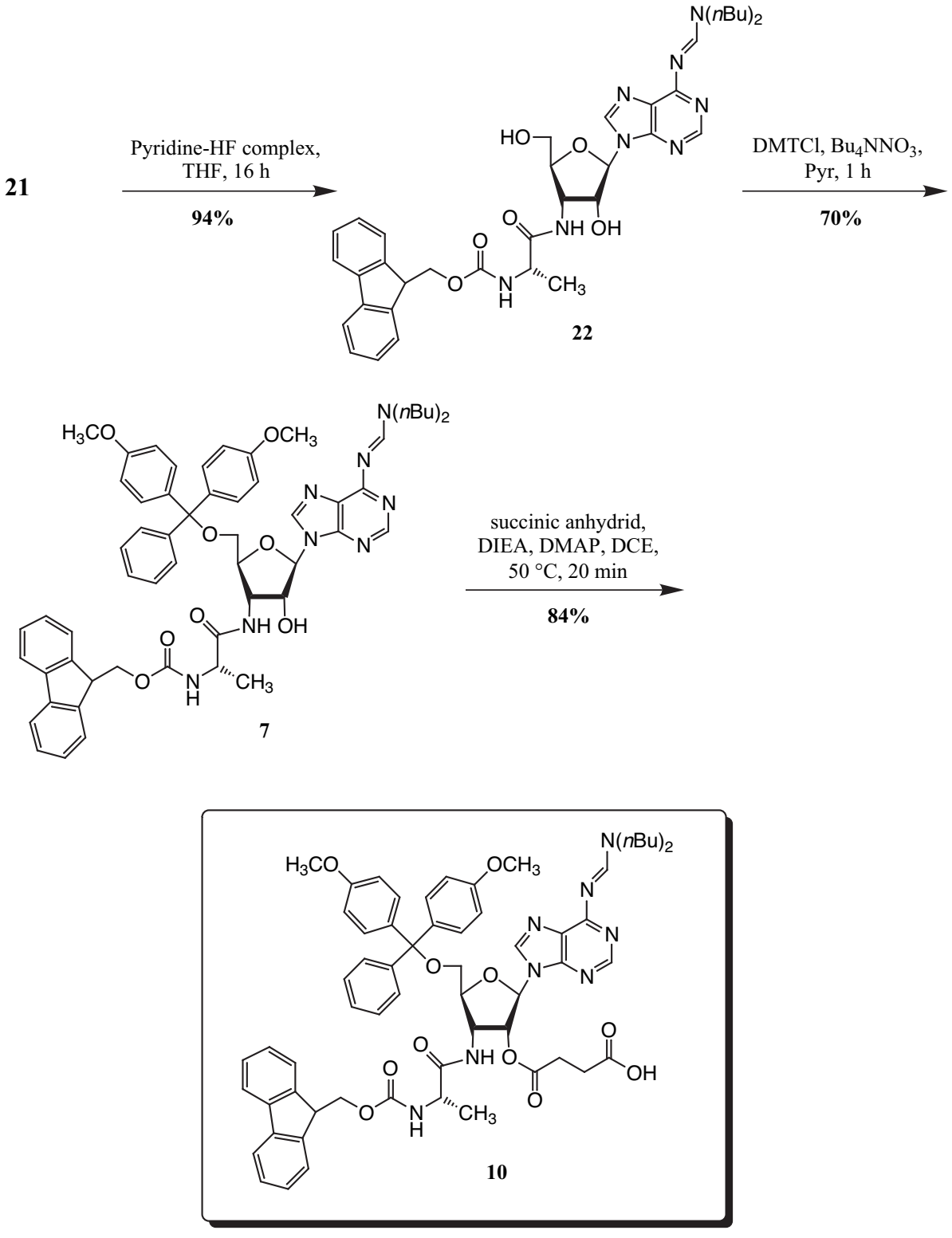
FMOC-L-alanine was coupled using as activating reagent CBMIT [1,1'-carbonylbis(3-methylimidazolium) triflate][91] which, prepared *in situ* (Section 2.3.2), allows the formation of highly reactive acyl imidazolium species which can easily acylate also hindered, secondary amines. Compound **21** was isolated with an overall yield of 75% over two steps.

To improve the yield of **21**, an alternative reducing procedure of the azide **18** was tried. According to literature,[92] tin(II) complexes prepared by treatment of $SnCl_2$ with appropriate amount of PhSH and Et_3N appear to be the best reducing agent for azides. Azide **18** was reduced following the procedure described and, after the FMOC-alanine coupling, compound **21** was isolated in 70% overall yield over two steps. Since this method did not prove to be superior to the Staudinger reaction and because of the use of the unpleasant PhSH, this reaction was abandoned.

2.1.3 Desilylation, dimethoxytritylation and succinylation

The deprotection of both TBDMS groups was achieved using pyridine-HF complex in THF (Scheme 12). This reagent was chosen because it is considerably less basic than the usual desilylating reagent TBAF which provokes the FMOC deprotection. Using only a slight excess (1.2 equivalent per TBDMS group) of pyridine-HF complex, deprotection yields around 70% was obtained. Using a higher excess (3.6 equivalents) of complex allowed almost quantitative yields (94%).

After desilylation, the 5'-hydroxyl group was protected with dimethoxytrityl chloride and Bu_4NNO_3 as catalyst.[93] The primary hydroxy group can be selectively protected even using an excess of DMTCl (2 equivalents). TLC monitoring of the reaction showed that after 1 hour the reaction was complete without any side product. Neverthless, compound **7** could be recovered only in 70% yield. A possible reason for this discrepancy, could be a partial deprotection during the column chromatography testified by the appearance of several yellow, orange bands on the column during elution.



Scheme 12

The final step in the synthesis of the building block **10**, was the succinylation of the 2'-hydroxyl group. The reaction was executed following the Kumar procedure.[94] DIEA was used instead of the suggested Et_3N to avoid any FMOC deprotection. Compound **10** contains

an acidic functionality that can deprotect the acid-labile DMT group. It is therefore important to avoid storing it as an oil. Thus, compound **10** was precipitated from a mixture of EtOAc/hexane (3/7) and obtained in good yield (83%).

Compound **18** was submitted to basic treatment to check the stability of the formamidine group during the peptide synthesis. Treatment with 20% piperidine in DMF and in another batch with 2% DBU, 2% piperidine in DMF was prolonged to three days. UV-detection of the TLC plate, revealed only a small amount of deprotected product (8-10%) with the DBU containing solution; a slightly higher amount of deprotection was detected after treatment with 20% piperidine solution. These basic treatment conditions (3 days) are, clearly, much longer than the time of contact during the synthesis of a peptide of 20-25 amino acids (5-6 hours). These data showed that the formamidine group is compatible with peptide synthesis.

In conclusion, 3'-alanylamino-3'-deoxyadenosine derivative **10** was obtained in 19% overall yield over 11 steps. The protecting groups on the molecule have been demonstrated to be compatible with peptide and oligonucleotide synthesis. After succinylation, the molecule was ready for the immobilisation on solid support.

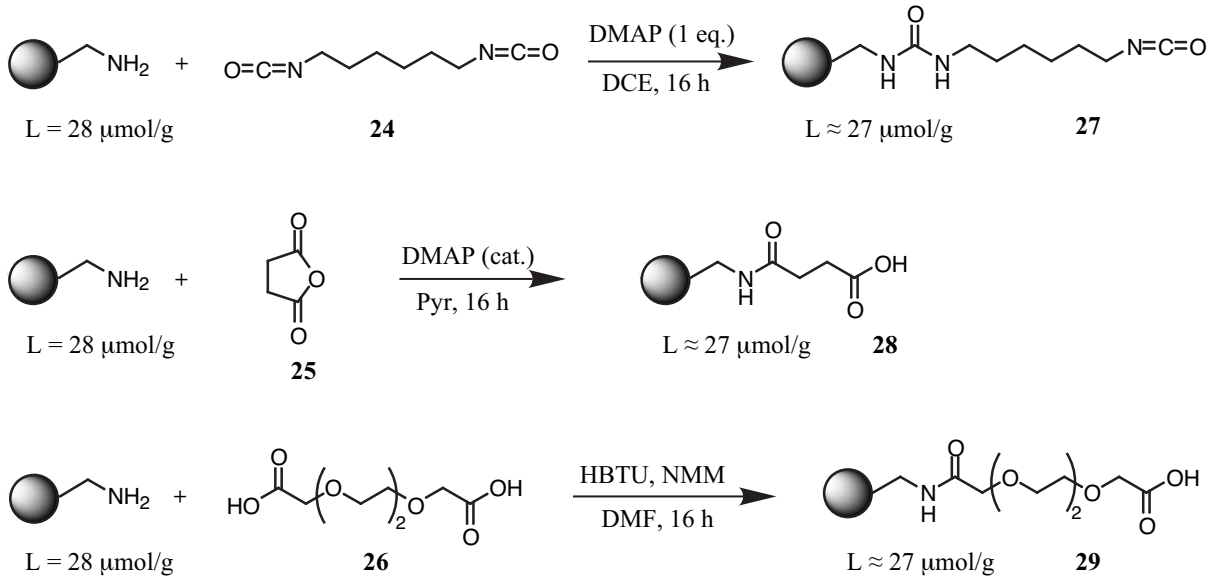
2.2 Immobilisation of the building block 10 on the solid support

2.2.1 Derivatisation of the solid support

The solid support used for the synthesis of our conjugates is an aminomethyl polystyrene 50%-crosslinked with divinylbenzene resin with an initial loading of 28 μmol NH_2 groups/g. Usually, oligonucleotide synthesis is performed using controlled pore glass (CPG) as solid support.[95] Neverthless, the superiority of highly crosslinked polymer-based solid support over CPG has been demonstrated.[96]

The immobilisation of the building block **10** on the solid support was preceded by the derivatisation of the resin with some spacers differing in length and hydrophobicity. The necessity to prepare a solid support bearing different spacer molecules emerged from the results published by Katzhendler and Richert affirming the superiority of long spacers over shorter ones.[76c, 97] A long spacer increases the steric accessibility of the reactive site during the coupling reaction, thus favoring high yields. We wanted also to confirm the superiority of hydrophilic spacers (polyethylene-glycol derived molecules) over hydrophobic ones since it has already been demonstrated that for the synthesis of oligonucleotide strands PEG-based spacers give better results.[97,98]

The first step of the derivatisation of the solid support was the attachment of three different molecules (hexamethylenediisocyanate **24**, succinic anhydride **25** and 3,6,9-trioxaundecanoic diacid **26**) to link the spacer to the solid support.



Scheme 13

In all three cases, the derivatisation was almost complete as determined by the quantitative ninhydrin test which gave a number of residual NH_2 groups (1 μmol/g) corresponding to the background (0.5-0.8 μmol/g).

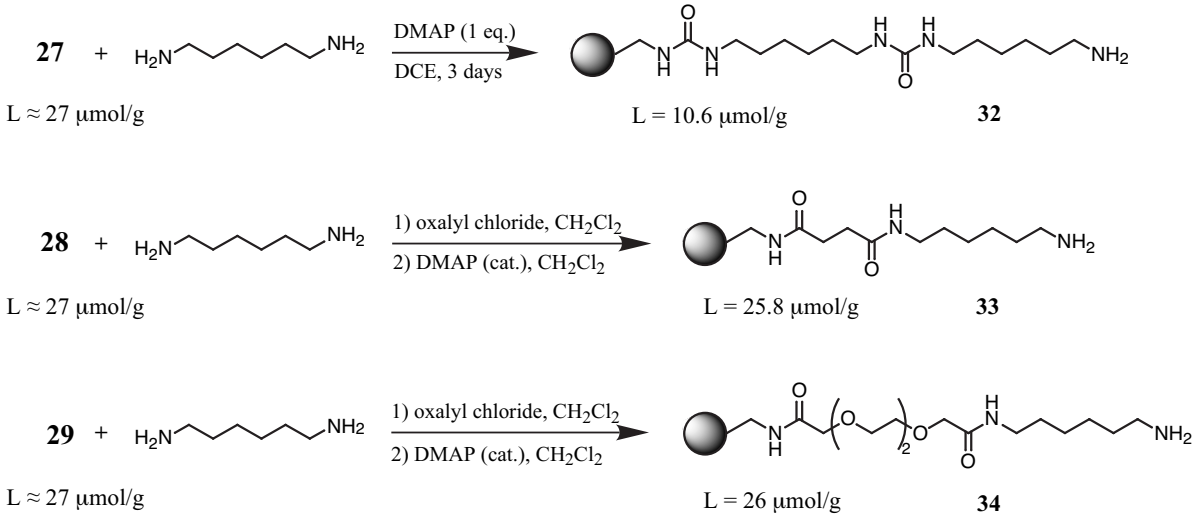
2.2.2 Attachment of different-length spacers

Two spacers were chosen: 1,6-diaminohexane **30** and a PEG-based diamino molecule **31** (Fig. 18).

30 **31**

Fig. 18: Spacer molecules attached on solid support.

1,6-Diaminohexane is short and hydrophobic whereas the PEG spacer is very long and more hydrophilic. The different physical characteristics of the two molecules could have a large influence on the overall yield of the conjugate synthesis since they create a different microenvironment which could change the swelling properties of the solid support and, therefore, influence the accessibility to the active sites and the delivery of the reagents.

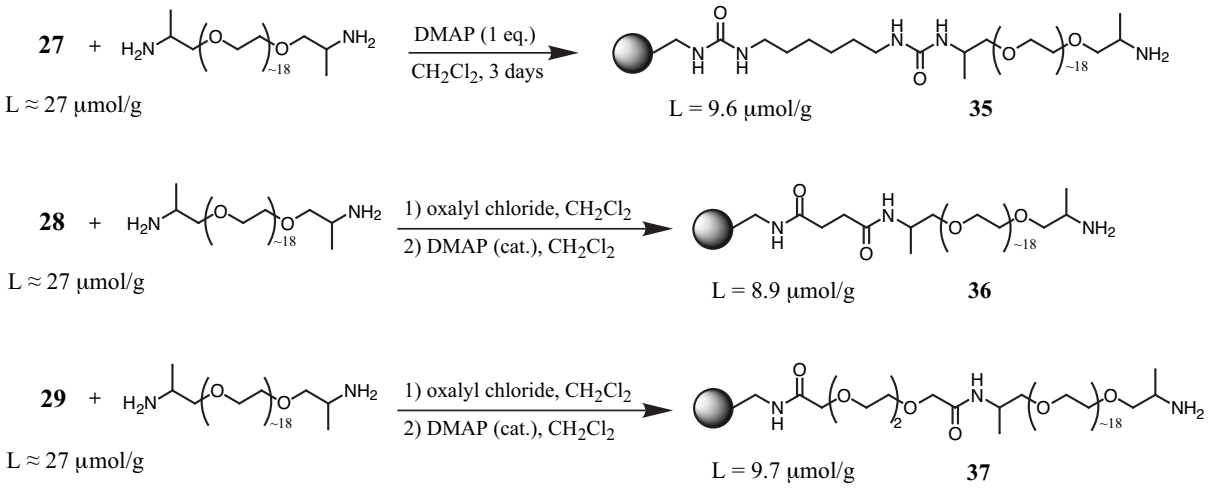


Scheme 14

1,6-Diamino spacer **30** was linked to the three derivatised solid supports **27**, **28** and **29** with different levels of substitution (Scheme 14). The carboxylic functions of **28** and **29** were activated with HBTU. The loadings obtained (16.3 μmol/g for **33** and 8.6 μmol/g for **34**) were

not satisfactory. The activation procedure was thus changed using oxalyl chloride which gave almost quantitative substitution levels.

In contrast, 1,6-diaminohexane could not be quantitatively attached to **27**; even prolonging the reaction time to three days, a loading of only 10.6 μmol/g could be obtained.



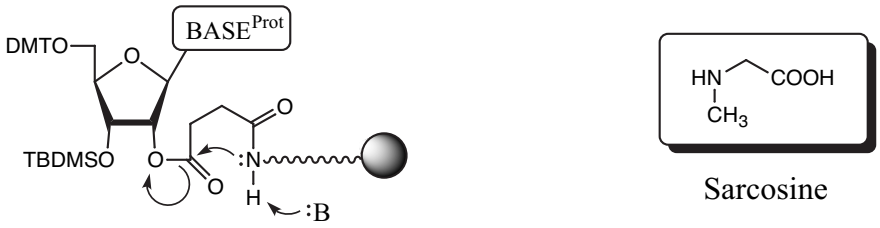
Scheme 15

The coupling of the PEG-based spacer **31** to the solid support revealed to be more difficult than that of **30** and only a loading around 10 μmol/g could be obtained. Using a huge excess of spacer **31** (up to 500 equivalents) and prolonging the reaction time to three days, did not improve the substitution level. The derivatised solid support **28** was activated with HBTU or with oxalyl chloride obtaining, in this case, only a slight difference of loading (7 μmol and 8.9 μmol respectively). Since PEG molecules are known to absorb water (also evidenced by NMR spectra), we tried to dehydrate the spacer using several methods but none of them proved to be efficient.

The difficulties encountered to attach this spacer on the solid support may be due to the presence of water or to the characteristics of the spacer itself because independentely from the solid support and coupling conditions, a value around 10 μmol/g was always obtained. One hypothesis is that the molecule, due to its length, could react with a second activated carboxylic group, forming a bridged structure, rendering thus the second amino function inactive for any further reaction.

The next step was the coupling of BOC-sarcosine to the spacer-derivatised solid supports. Sarcosine was introduced as base-stable linker between the succinylated building block **10** and the solid support.[99]

The synthesis of the peptidic moiety of the conjugate requires a basic treatment to deprotect the FMOC group before every amino acid coupling. The conventional succinate linker used to attach the nucleoside to the solid support is not stable to repeated basic treatments (Scheme 16).

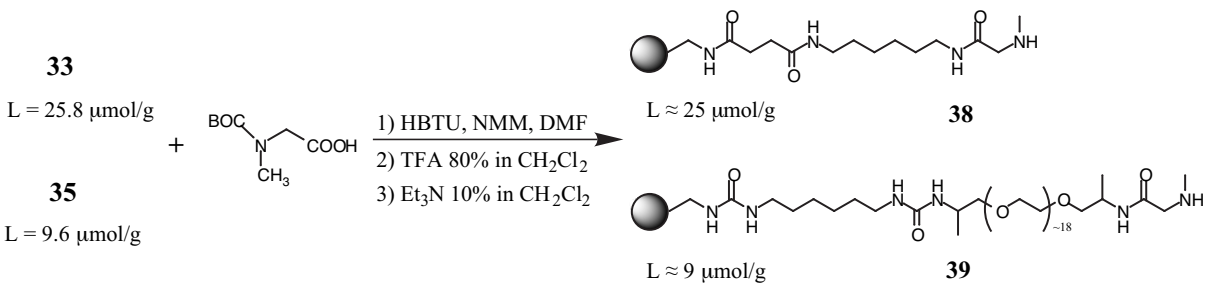


Scheme 16

Abstraction of the proton from the amide group leads to the formation of a stable 5-membered ring causing the cleavage of the nucleoside (or the nucleotidic strand) from the solid support. Replacing the proton of the amide with an alkyl group, eliminates the risk of basic cleavage.

We tested the stability of a conventional succinyl linker to repeated treatments with a solution of 2% DBU, 2% piperidine (v/v) in DMF. A commercial resin (Long Chain Alkylamino-Controlled Pore Glass) loaded with DMT-thymidine was subjected to the same basic treatments used during the peptide synthesis (10 minutes repeated 20 times). The loading of thymidine on the solid support was measured before and after the basic treatments. Measuring the trityl cation (DMT^+) released after acidic treatment of a weighed amount of resin, allows the precise quantification of nucleoside attached to the solid support. Our experiments confirmed a loss of material from the resin to a degree of about 1% for every basic treatment. We, therefore, considered necessary adding a sarcosine to the solid support before the attachment of the building block.

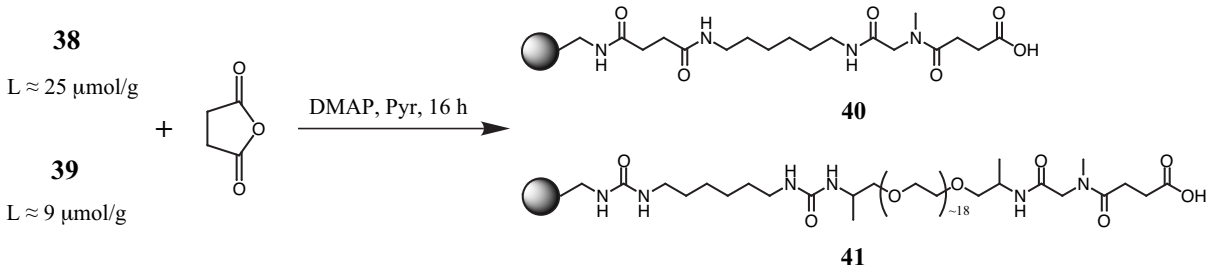
Two derivatised solid supports were selected for the further immobilisation of the building block: **33** (containing the shortest and more hydrophobic spacer) and **35**. This latter was preferred to **37** (similar loading but **37** is slightly more hydrophilic and longer) because of the urea-type bond between the PEG spacer and the hexamethylene linker which is more stable to strong basic treatment than an amide-type bond.



Scheme 17

BOC-sarcosine was preactivated with HBTU (15 minutes) and coupled to **33** and **35**. After the BOC deprotection and neutralisation of the free amine, a quantitative ninhydrin test indicated a residual amount of NH_2 groups (detection of only primary amines with ninhydrin) of ~0.8 μmol/g (coupling almost quantitative).

Portions of derivatised solid support **38** and **39** were further reacted with succinic anhydride, giving **40** and **41** (Scheme 18) which were successively used for immobilisation trials.

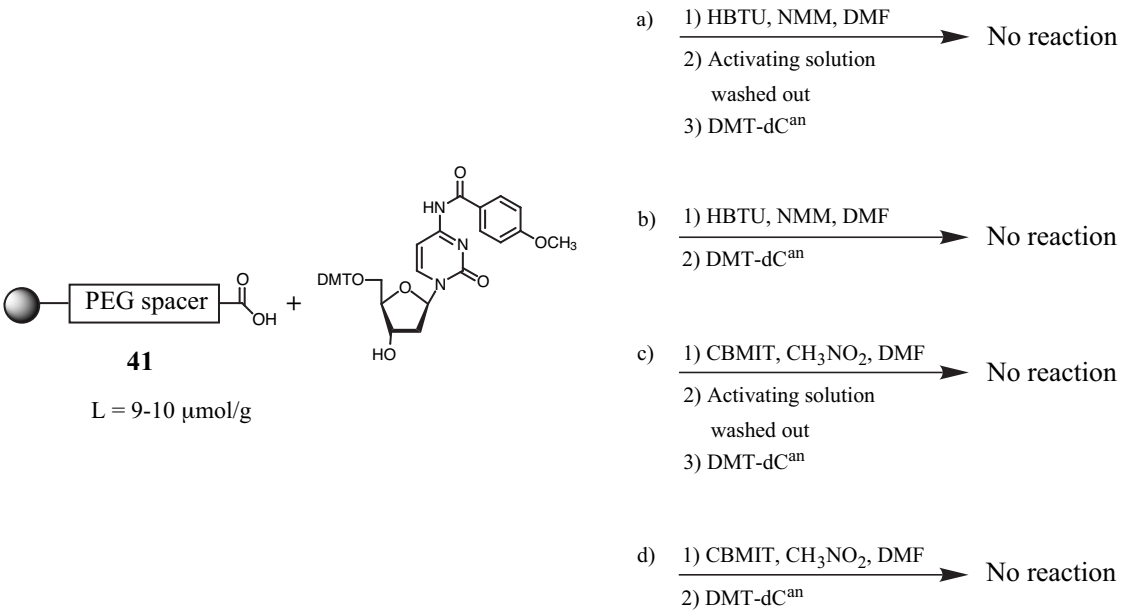


Scheme 18

2.2.3 Immobilisation trials using DMT-dC^{an} and DMT-T as model compounds

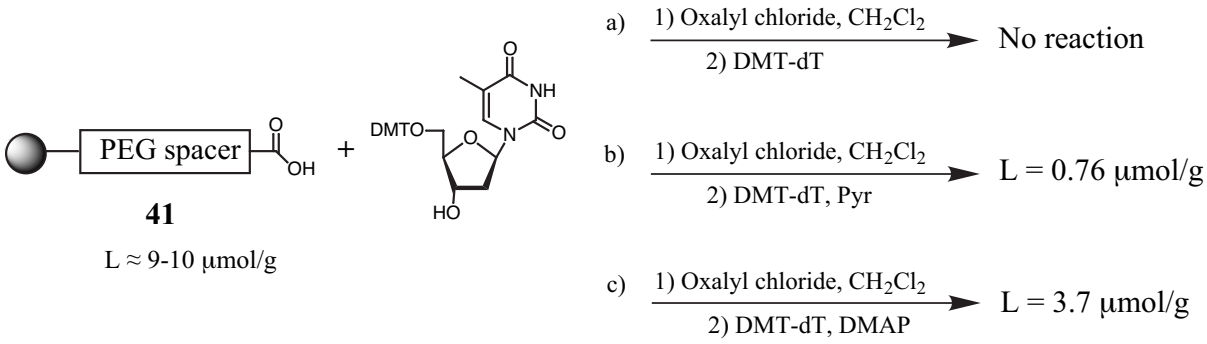
The solid supports **40** and **41** were splitted into several portions and reacted with N^4-anisoyl-5'-*O*-DMT-2'-deoxycytidine and 5'-*O*-DMT-thymidine.

The first attempts to attach DMT-dC^{an} on **41** were all unsuccessful. The following conditions were used:



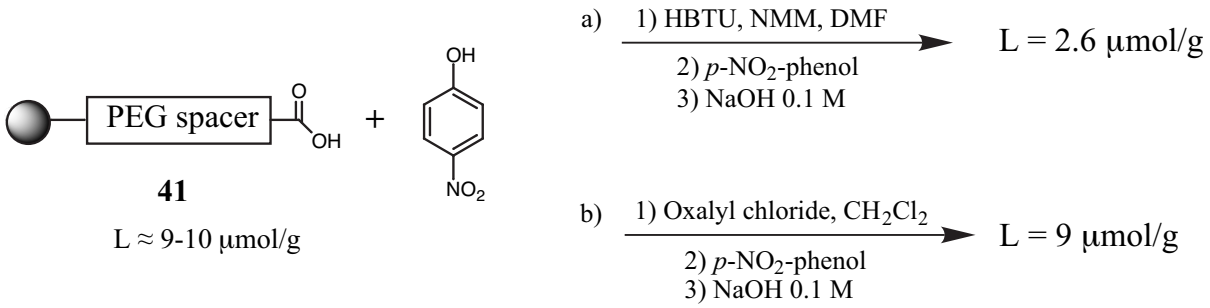
Scheme 19

No attachment of the monomer on the solid support was obtained using HBTU and CBMIT to activate the carboxylic group, so we decided to replace them with oxalyl chloride. DMT-dT was coupled on the activated solid support:



Scheme 20

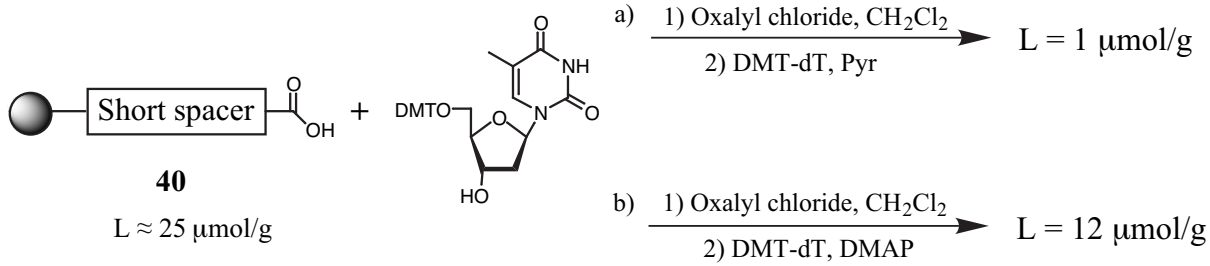
These results indicated that the activation was only partly the cause for the reaction failure. A possible reason could be that succinylation of the solid support **41** had failed. Consequently, **41** was derivatised with *p*-NO_2-phenol which, after basic hydrolysis, releases the *p*-NO_2-phenolate ion (yellow). Quantitative measurements of the ion released (absorption at λ= 405 nm, ε = 24651), allowed the determination of the level of succinylation of **41**.



Scheme 21

This latter reaction showed that succinylation of the solid support was successful. This result suggested that the procedure used for the immobilisation of the nucleosides had to be changed.

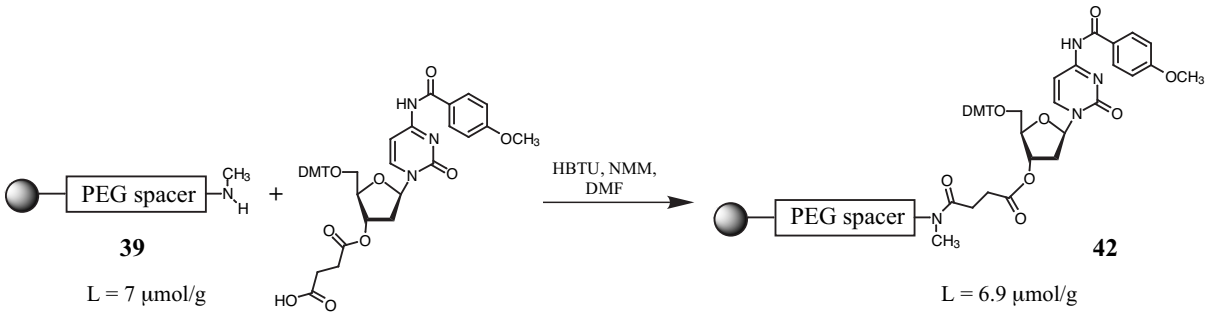
Further immobilisation experiments were carried out using the derivatised solid support **40** (short spacer) to evaluate the influence of the spacer on this reaction.



Scheme 22

These experiments evidenced, again, the importance of adding DMAP as catalyst but, although a good level of derivatisation was obtained (12 μmol/g), it represented only a 50% yield.

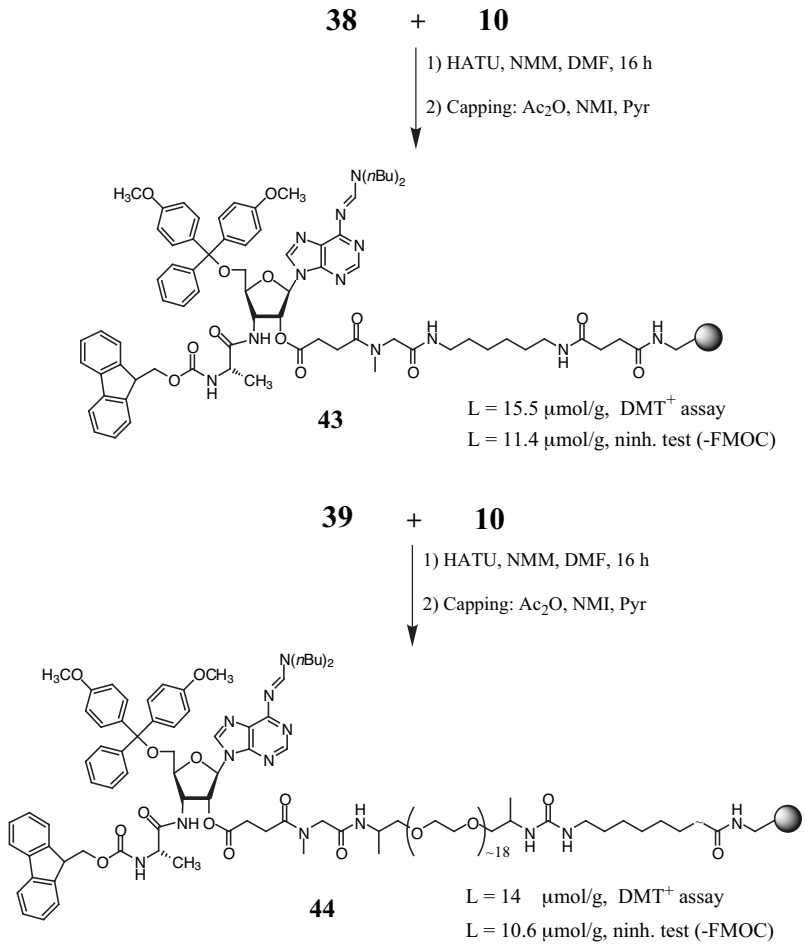
At this point an alternative procedure was explored. DMT-C^{an} was succinylated and, after activation of its carboxylic function using HBTU, it was attached to the derivatised solid support **39** (using a batch of resin with a loading of 7 μmol/g):



Scheme 23

This immobilisation procedure was successful (substitution level 98%), consequently this methodology was used to immobilise the building block **10** on the solid support.

2.2.4 Immobilisation of 10 on the solid support



Scheme 24

The immobilisation of **10** on the solid supports **38** and **39** was carried out activating the carboxylic group with HATU. Amazingly, derivatisation of **39** (estimated loading ~10 μmol/g) gave a higher final loading (substitution level 106%!). These anomalous values were obtained in three distinct experiments excluding a measuring error of the loading of **44**. It is, indeed, more probable that the loading of **35** was affected by an error because this value was not completely reproducible.

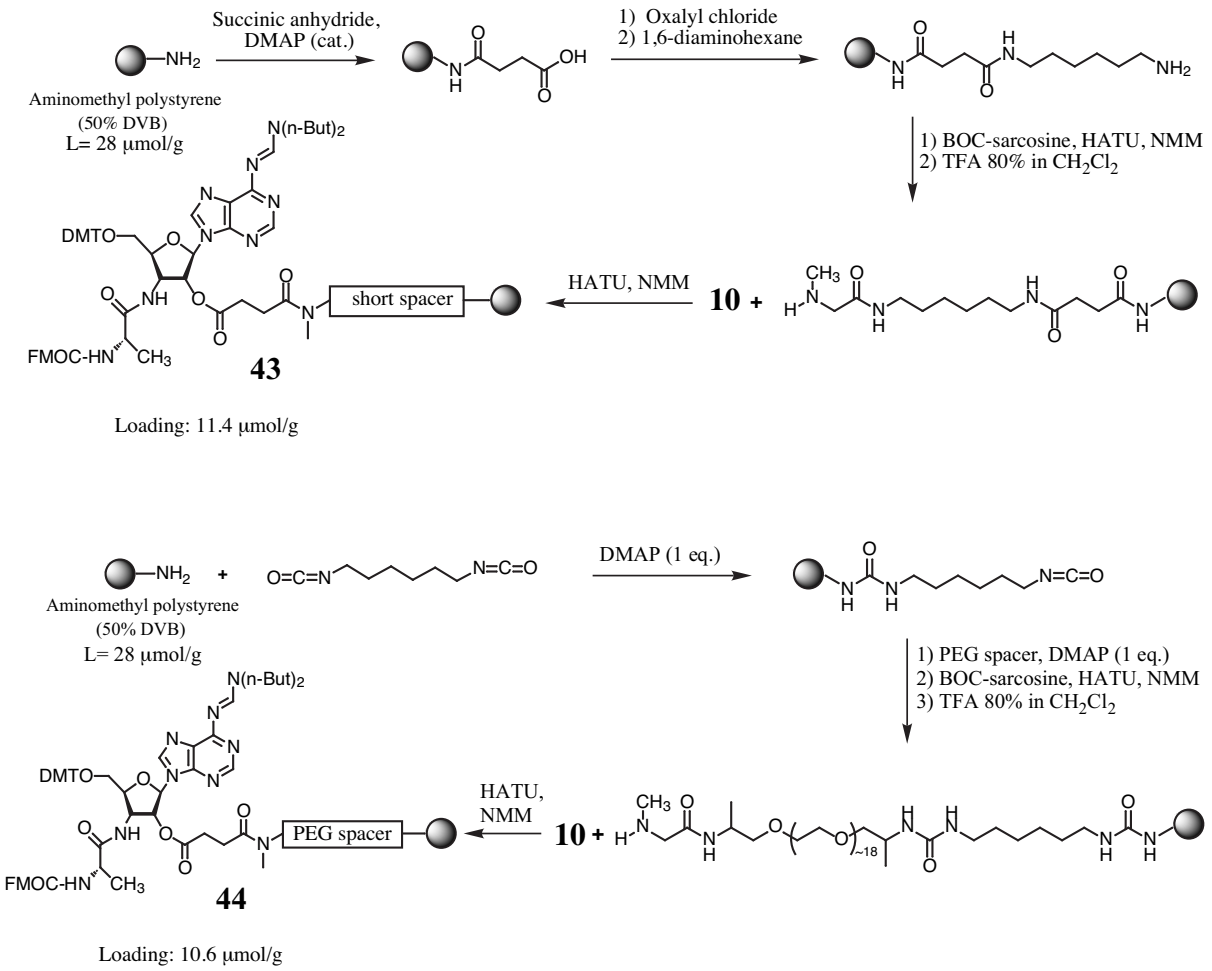
A difference in the level of loading was always observed using the DMT^+ cation release assay (+ 20-25%) compared to the ninhydrin test (after FMOC deprotection). The value obtained with the ninhydrin test was considered as the starting loading. It can be noted that, independently of the spacers, comparable loadings were obtained when the nucleoside was immobilised on the two solid supports.

A small portion of solid support **43** was treated with an ethanolic solution (33%) of CH_3NH_2 for 2 hours at room temperature. This is the deprotection procedure used at the end of the conjugate synthesis. We wanted to test the stability of the amide peptide bond under the strong basic conditions used in the final deprotection and, at the same time, we wanted to be sure that the molecule could be recovered without any side product. HPLC of the deprotected molecule showed the presence of only a major peak corresponding (ESI-MS analysis) to the target molecule and traces (2-3%) of the target product still containing the formamidine group.

2.2.5 Conclusions

The building block for the synthesis of the conjugates, **10**, was immobilised on two derivatised solid supports with a final loading of 10-11 μmol/g.

Summarising, the following strategies were followed for the immobilisation:



Scheme 25

The immobilisation on solid support incorporating two very different spacers, should give the possibility to analyse the influence of them on the synthesis of the conjugates, in particular of the peptidic moiety which, due to the steric hindrance at the 3'-position, could revealed very difficult to synthesise.

2.3 Optimisation of the peptide synthesis

The first conjugate synthesis attempt, carried out by Ewa Biala, was not completely successful. The conjugate, whose sequence is:

5'-p-GGGGCUCU**UUCG**GAGCUCCACCA-3'-NH-$(Ala)_8$ "22Ala8"

(bold: loop region, underlined: single stranded part), was synthesised using CBMIT as coupling reagent for the assembly of the peptidic moiety (10 minutes coupling time). MALDI-ToF analysis of the HPLC-SAX purified product showed the presence of conjugates having shorter peptide sequences than the target one (5,6 and 7 alanines) due to incomplete coupling reactions (Fig. 19).

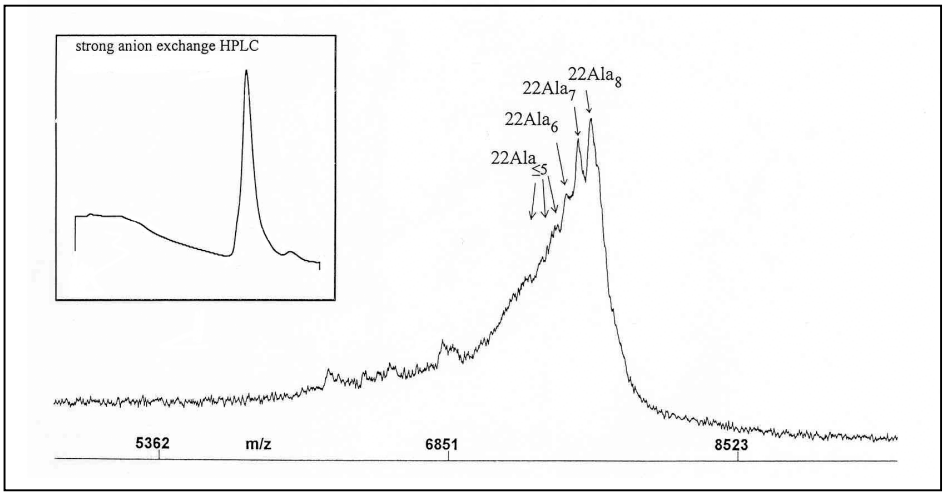


Fig. 19: MALDI-ToF and SAX-HPLC of 22Ala8

Polyalanine peptide sequences are notoriously difficult in solid phase peptide synthesis. They have often been used as model system to investigate the different factors that influence the coupling yields in peptide synthesis because of the well-known difficulties that arise when one wants to couple more than five or six subsequent alanine residues.[100] Polyalanine based peptides are also called "difficult sequences" in SPPS. The syntheses of these peptides are characterised by repetitive incomplete aminoacylation leading to the formation of peptide side products lacking one or more amino acids but with properties similar to the target sequence thus difficult to separate.

The main cause for the existence of difficult peptide sequences has been identified in the aggregation of some of the growing peptide chains by amide hydrogen bonds leading to the formation of intermolecular β-sheet structures. This aggregation renders the peptide

chains involved sterically resistant to aminoacylation and deprotection thus leading to incomplete reactions. Since we planned the synthesis of conjugates whose peptidic moieties are based on polyalanine stretches, we tried to find a good activating reagent and good coupling conditions in order to minimise the appereance of shorter peptide sequences (also called deleted sequences in SPPS).

Our first attempt of synthesis indicated that this problem is present (Fig. 19) and constitutes an important handicap for the synthesis of homogeneous products. One method to minimise the aggregation in the difficult sequences, is to use a solid support with a low substitution level of amino acid in order to dilute the growing chains and hamper the physical contact among them.

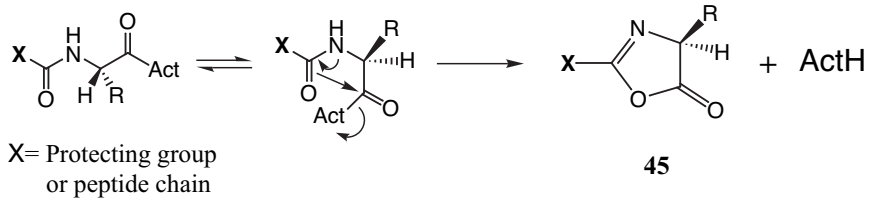
Using magic angle spinning NMR to study resin-bound polyalanine peptides, Lippens and coworkers demonstrated that using solid supports with very low loadings (~7 μmol/g), it could be possible to synthesise polyalanine stretches up to 21 amino acids without major aggregation problems.[100d]

The loading we obtained after the immobilisation of **10** on solid support (10-11 μmol/g) is very close to the value suggested by Lippens and this should constitute a positive factor for the syntheses we planned (stretches of ~20 alanines). A careful choice of the activating reagent was also necessary since this is another factor that can have a big influence on the quality of a synthesis.

2.3.1 The coupling methods and the racemisation problem in SPPS

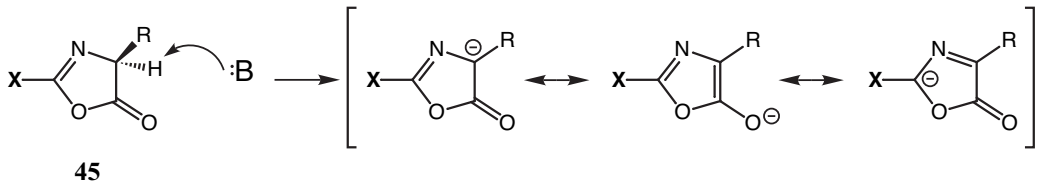
Three different coupling reagents were tested through the synthesis of model peptides based on polyalanine sequences.

An essential feature of all coupling methods is that, in addition to forming the peptide bond in very good yield (ideally in quantitative yield), they should not promote, during the activation process, the racemisation of the amino acid. It is often difficult to achieve good yields and absence of racemisation because the conversion of the carboxylic acid into a derivative bearing a good leaving group favours the formation of an oxazolone (Scheme 26, **45**).



Scheme 26

Oxazolones can still react with an amino group to form an amide bond, however, they are easily racemised under mild conditions via base-catalysed deprotonation (Scheme 27). The ease by which the proton of the chirality centre can be abstracted by bases, is due to the resonance stabilisation of the carbanion generated in the process.



Scheme 27

When the X group is a urethane-type protecting group, the carbanion is destabilised and the racemisation degree is low. Neverthless, the racemisation side reaction must be taken into account as a serious problem in peptide synthesis; in fact, although a small fraction of each residue is racemised, synthesis of long peptides can be affected by a large amount of diastereoisomers in the final product.

Taking into account this problem, we have chosen three activating reagents known to give very low racemisation but, at the same time, high yield of peptide coupling. They are: HATU [*O*-(7-azabenzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate],[101] DEPBT [3-(diethoxyphosphoryloxy)-1,2,3-benzotriazin-4(3*H*)-one],[102] CBMIT [1,1'-carbonyl-*bis*(3-methylimidazolium) triflate] [91] (Fig. 20).

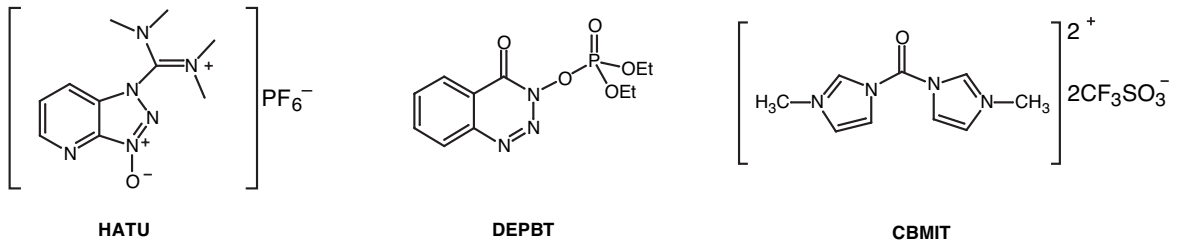


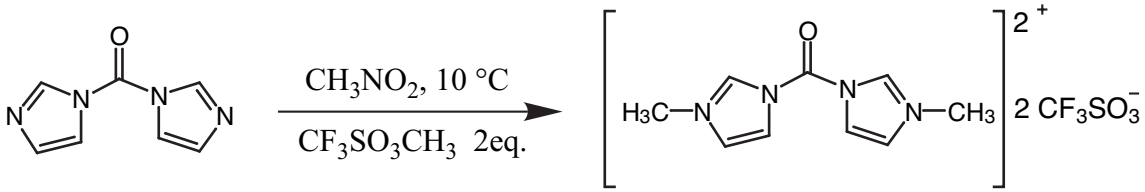
Fig. 20: Coupling reagents.

HATU is one of the best activator commercially available. It has been demonstrated to give high yields and low racemisation. However, care must be taken when using guanidinium *N*-oxide-based activators (also called uronium salts) not to use an excess relative to the carboxylic component as this can lead to capping of the amino terminus through guanidine formation.[103] In order to eliminte this side reaction, preactivation of the amino acid is advisable.

DEPBT has been recently proved to be superior to HATU and HBTU both in terms of coupling yields and racemisation levels.

CBMIT mediates the coupling of amino acids under remarkably mild conditions; the formation of a highly reactive acyl imidazolium intermediate eliminates the need of a base to deprotonate the carboxylic acid. Making the coupling reaction in the absence of a base, would minimise the risk of racemisation.

2.3.2 Synthesis of CBMIT



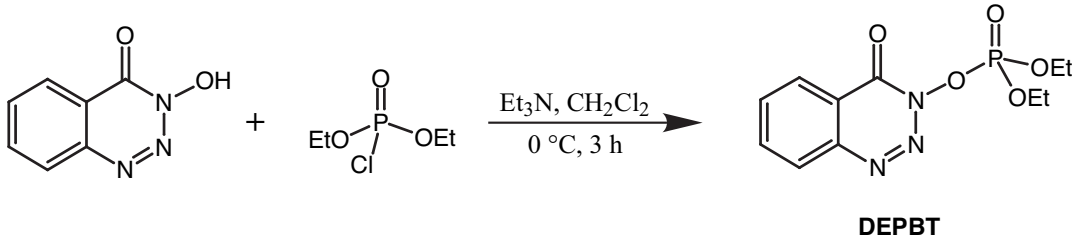
CBMIT

Scheme 28

CBMIT was prepared by alkylation of carbonyldiimidazole with methyl triflate. Trifluoromethane sulfonate anion satisfies the criteria of being non basic and non nucleophilic. CBMIT was, usually, generated in solution just prior to use. Addition of CBMIT solution to a suspension of FMOC-amino acid resulted in rapid CO_2 evolution indicating a positive reaction with the carboxylic group. After some minutes, the activated amino acid precipitated as it is insoluble in CH_3NO_2. To redissolve it, DMF was added in circa the same amount as CH_3NO_2.

Solutions of activated amino acid could be stored for 3-4 days under argon at -20 °C maintaining a high degree of activity (90-100%) when reacted with an amine. The activity was tested using a chiral amine, *R(+)*-1-(1-naphthyl)ethylamine so that also an eventually racemised product could be detected (HPLC and TLC monitored reactions).

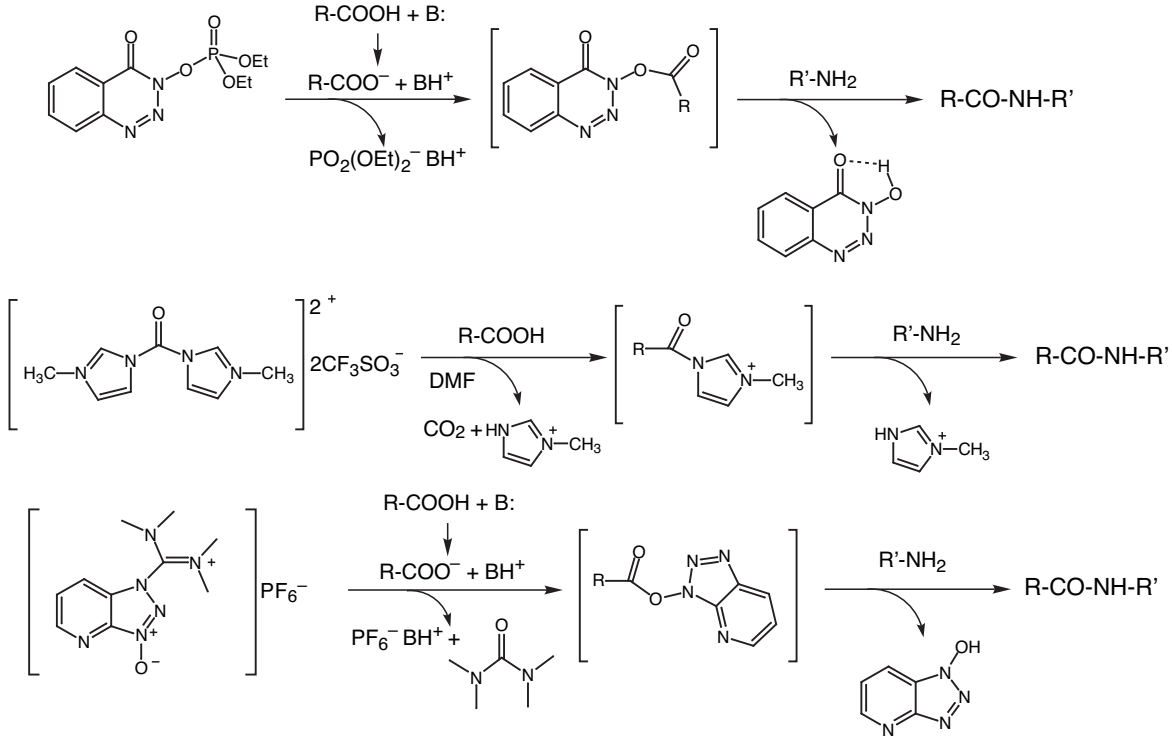
2.3.3 Synthesis of DEPBT



Scheme 29

DEPBT was prepared by reaction of 3-hydroxy-1,2,3-benzotriazin-4(3*H*)-one (HOOBT) with diethyl chlorophosphate in the presence of triethylamine in CH_2Cl_2. The compound was obtained in good yield after two crystallisations from EtOAc/hexane (72%). DEPBT is a colorless crystalline solid that can be stored for months at room temperature. Its high stability, compared to HATU and CBMIT, is a consequence of its existence as a neutral moiety, while the others are hygroscopic salts. This advantageous property of DEPBT renders it very practical to use. In order to obtain the best results, it is better to use two equivalents of DEPBT compared to the amino acid to activate and a long preactivation time (at least 30 minutes).

We can summarise the three methods in the following scheme:



Scheme 30

2.3.4 Synthesis of the model peptide $(Ala)_{12}$

The first peptide synthesised for the test was formed by twelve alanine residues:

$(Ala)_{12}$ (**46a**: CBMIT; **46b**: HATU)

As a difficult sequence to synthesise in good yield, this peptide would allow a good comparison to select the best activator.

The peptides were synthesised manually in a Teflon syringe using as solid support a Wang resin with an initial substitution of FMOC-alanine of 0.41 mmol/g. The syntheses were carried out using the following conditions:

- 10 equivalents of activated FMOC-amino acid
- double coupling
- reaction time of each coupling: 30 minutes
- FMOC deprotection: 20% piperidine in DMF, 3 times x 5 minutes
- Capping: Ac_2O, NMM in DMF, 10 minutes

Using HATU, the FMOC-amino acid was activated *in situ* (no preactivation) just prior every coupling whereas with CBMIT all the FMOC-amino acid quantity needed for the synthesis was activated in one batch and stored in small aliquots at -22 °C.

The syntheses were monitored by UV absorption of the product released after FMOC deprotection. The first value was taken as reference (100%). Normally, the first FMOC deprotection is a little bit lower compared to the second, probably due to a loss of FMOC protecting group after a long storage.

A semiquantitative method was applied to measure the FMOC release. A manual procedure similar to the method used to monitor the synthesis in an automated synthesiser, was developed. Instead of performing a quantitative measure at every cycle (very long procedure), a part of the deprotection solution (100 or 200 μl, depending on the quantity of resin and its substitution) was diluted to 10 ml and its absorbance measured at λ=300 nm. This method allowed a rapid monitoring of the synthesis.

The results obtained for the two syntheses are illustrated in Figure 21.

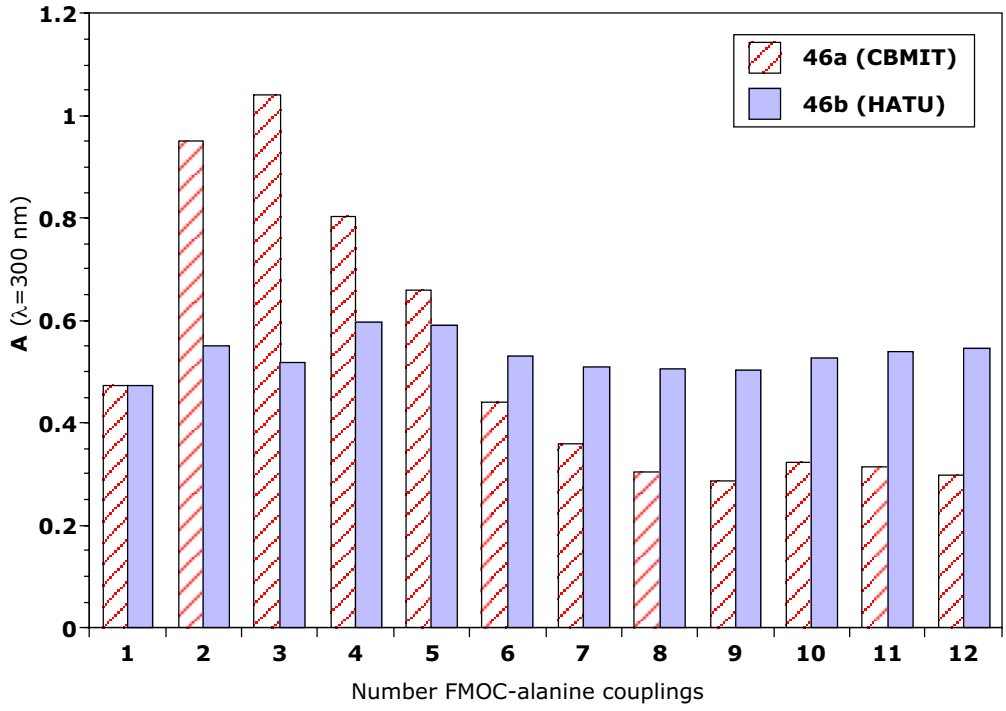
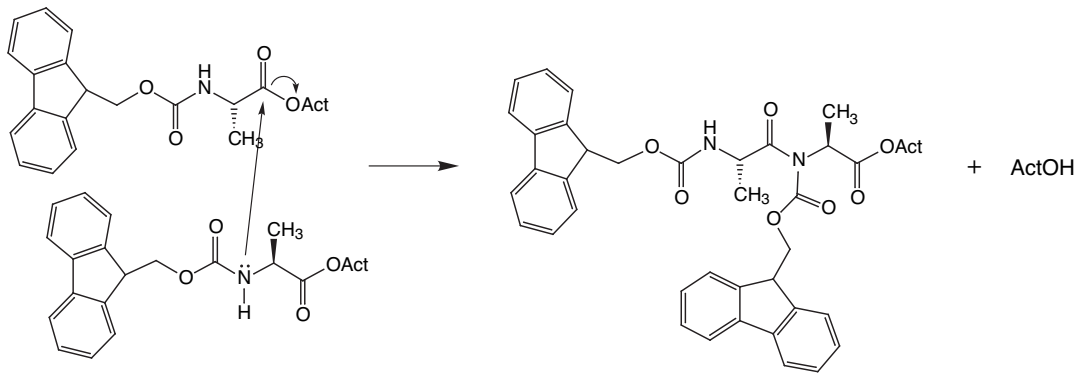


Fig. 21: FMOC-deprotection monitoring for the synthesis of **46a** and **46b**.

The analysis of these results indicated that HATU gave a peptide in very good yield (~100%) while CBMIT produced a peptide in lower yield (~60%), worse, a higher level of FMOC deprotection is detected at the beginning of the synthesis. This anomalous value is probably indicative of a side reaction called overactivation leading to the coupling of two amino acids instead of one. This side reaction is due to the FMOC-alanylation of the FMOC-carbamate nitrogen (Scheme 31).[104]



Scheme 31: Proposed mechanism of overactivation (Ref. 104).

Only mass spectrometry and HPLC analysis could confirm this hypothesis. Unfortunately, the crude products resulting from the cleavage of the solid support-bound peptides, were highly insoluble. Only neat TFA or a mixture 1:1 of AcOH and TFA could dissolve the two products. Insolubility hampered an HPLC analysis and consequently, a precise quantification of every species present in the crude mixture. MALDI-ToF spectra of **46a** and **46b** (Fig. 22 and 23) revealed the presence of many shorter sequences for both peptides and confirmed the presence of a product containing an amino acid more for **46a**.

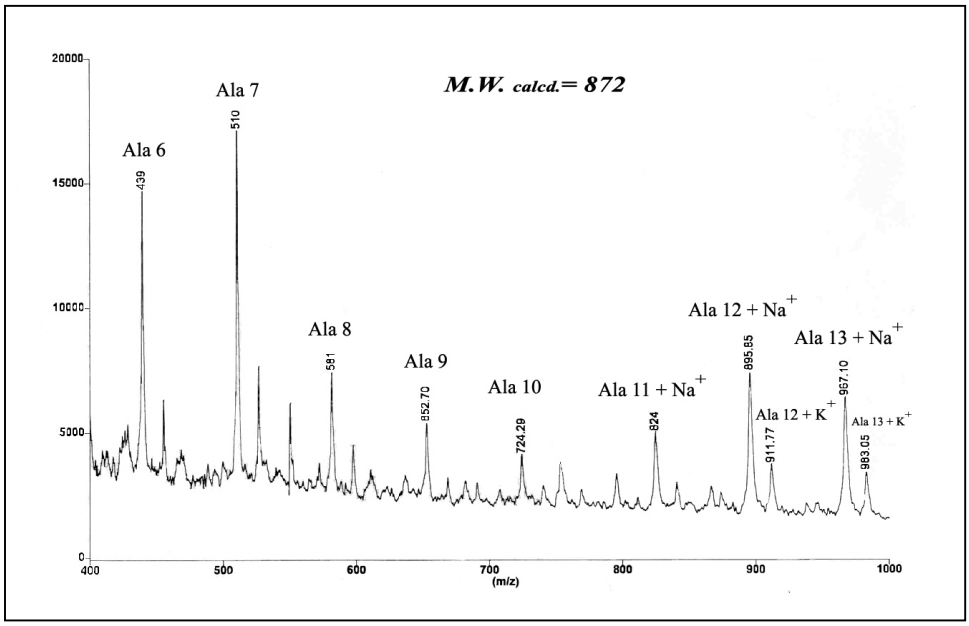


Fig. 22: MALDI-ToF spectrum of crude **46a**.

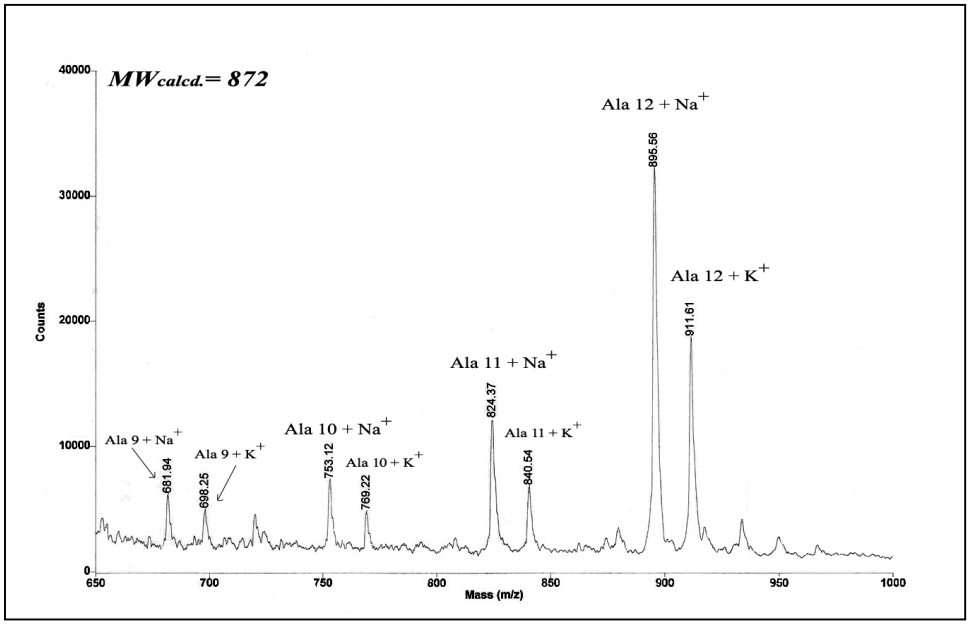


Fig. 23: MALDI-ToF spectrum of crude **46b**.

Since MALDI mass spectrometry is not a quantitative method, it is not possible to measure the quantity of every side-products. The spectra, neverthless, confirmed that the UV monitoring used is not very reliable since the overall yield (100% for **46b**) is not confirmed by the presence of several peaks in the spectrum. Anyway, it seems to be clear that HATU gave better results than CBMIT.

2.3.5 Synthesis of the model peptide $(Ala)_{12}$-Arg-Ala-Arg-Arg

In order to overcome the insolubility problems, a more hydrophilic peptide was synthesised:

$(Ala)_{12}$-Arg-Ala-Arg-Arg (**47a**: CBMIT, **47b**: HATU, **47c**: DEPBT)

Three arginines (positively charged up to pH 12) located in the C-terminal part of the peptide should render the polyalanine stretch more soluble in aqueous solution so that HPLC analysis could be feasible. The resin used had an initial substitution of FMOC-Arg of 0.5 mmol/g. The asparagine side chain was protected with the acid-labile 2,2,5,7,8-pentamethylchroman-6-sulphonyl (Pmc) group.[105]

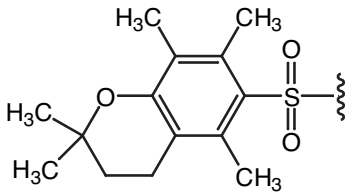


Fig. 24: Pmc protecting group.

Three peptides were synthesised using the three activators described above. In order to avoid, or at least, to minimise overactivation, a lower amount of activated FMOC-amino acid was used for each couplig. The syntheses were carried out using the following conditions:

- 5 equivalents of activated FMOC-amino acid
- double coupling
- reaction time of every coupling: 30 minutes
- FMOC-deprotection: 20% piperidine in DMF, 3 times x 5 minutes
- Capping: Ac_2O, pyridine, 1M in DMF (Capping A solution); *N*-methylimidazole 1M in DMF (Capping B solution), equal volumes of Cap. A and Cap. B; 5 minutes

When CBMIT was used as activator, all the FMOC-amino acid was activated at the beginning of the synthesis and stored at -20 °C in small aliquots (each one corresponding to the amount needed for one coupling); with HATU *in situ* activation was performed and with DEPBT the amino acid was preactivated for 30 minutes before the coupling reaction.

The synthesis carried out with CBMIT was monitored by measuring the FMOC-deprotection level. An overall yield of ~48% was obtained. The crude product **47a** was soluble in aqueous acidic solution (as expected) so it could be analysed by HPLC (Fig. 28). The MALDI-ToF spectrum showed the presence of several peaks, some of which could be identified. No peak corresponding to an overactivated product was present (Fig. 25).

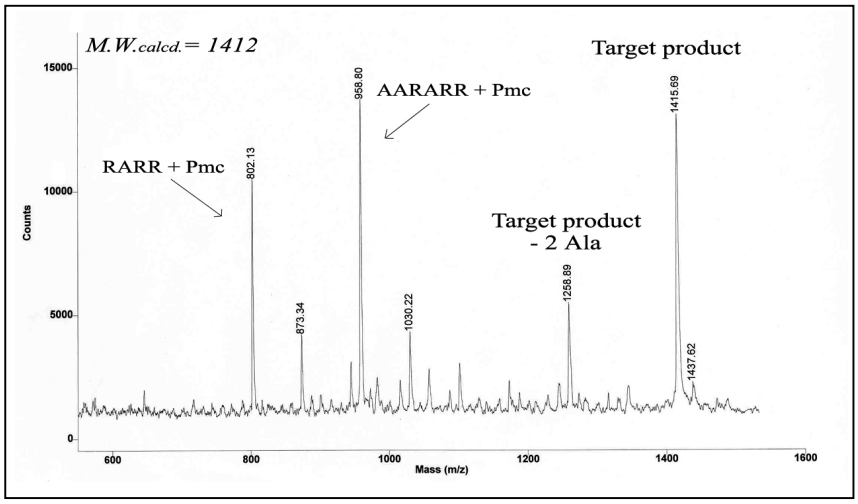


Fig. 25: MALDI-ToF spectrum of crude **47a**; R=Arg, A=Ala.

The synthesis monitoring was changed for the peptides **47b** and **47c**. Quantitative ninhydrin test was used to measure the NH_2 groups on the resin before and after every coupling. This procedure is more precise and reproducible compared to our method of detection of the FMOC-deprotection although it is destructive (the quantity of resin used for the test is lost) and a time-consuming method. The syntheses of **47b** and **47c** were achieved with a 32% and 22% overall yield respectively. MALDI spectra of **47b** and **47c** are shown in Figure 26 and 27.

The overall yield values do not fit with the average coupling yield per step obtained during the two syntheses. In fact, average values of around 97-98% were measured after every coupling, thus we should have obtained an averall yield around 60-70% over 15 steps.

An explanation to this discrepancy could be found in the paper of Warras and coworkers.[100d] They observe a collapse of the NMR signals after the FMOC deprotection when the peptide chain became long enough to allow the hydrogen bond formation between the terminal amino groups of vicinal chains. In our case, the coupling of each residue seemed to proceed nicely (low level of residual amino groups measured just after the end of each coupling, FMOC-on) but, after the FMOC deprotection, the number of available NH_2 groups for the next coupling was lower than the theorical value. Probably, when the terminal amino groups were deprotected, some growing peptide chains aggregated causing lower reaction rates and therefore low coupling yields.

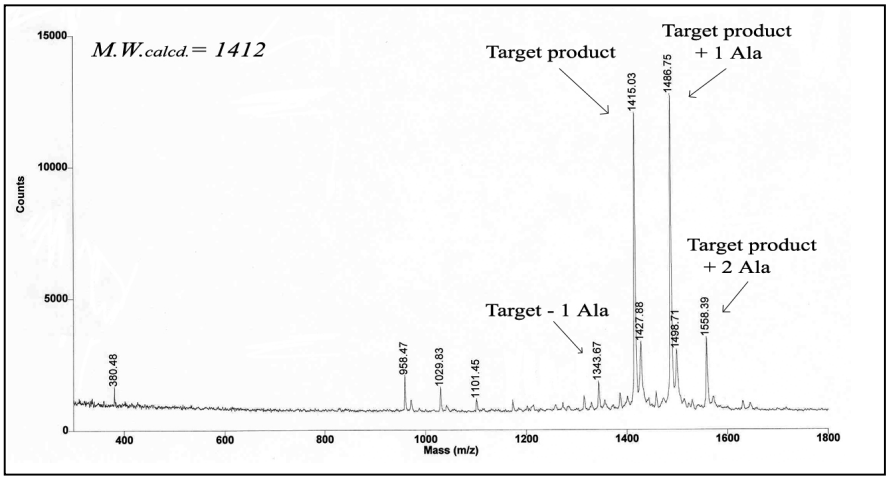


Fig. 26: MALDI-ToF spectrum of crude **47b**.

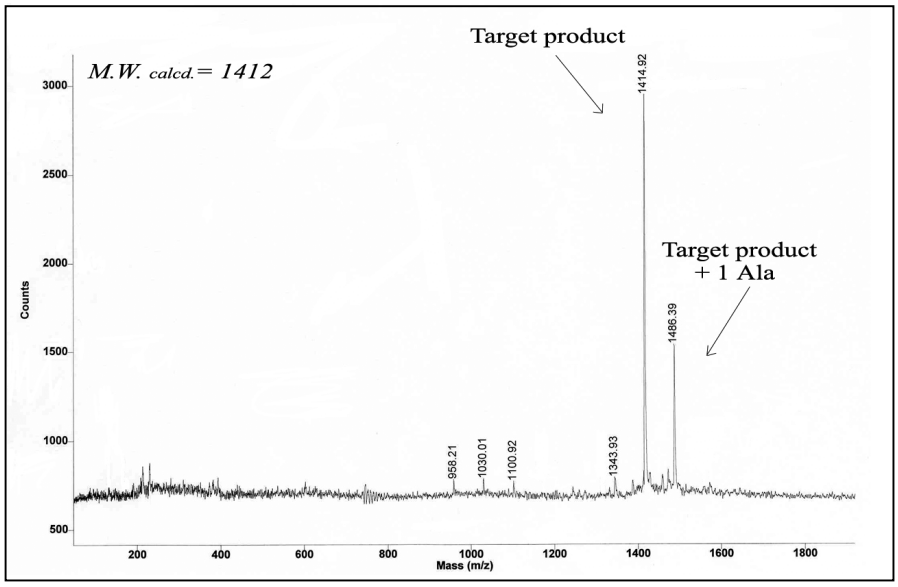


Fig. 27: MALDI-ToF spectrum of crude **47c**.

HPLC analysis of the three products indicated the presence of many peaks for all the three peptides in particular for **47a** (Fig. 28, 29 and 30). Only some peaks could be assigned through MALDI-ToF analysis. HPLC chromatogram profiles show products of lower quality when compared to the MALDI spectra of the same. This gives a further indication of the qualitative nature of the MALDI-ToF mass spectrometry technique which is more adapted to confirm the identity of a compound than to assess the quality of a synthesis.

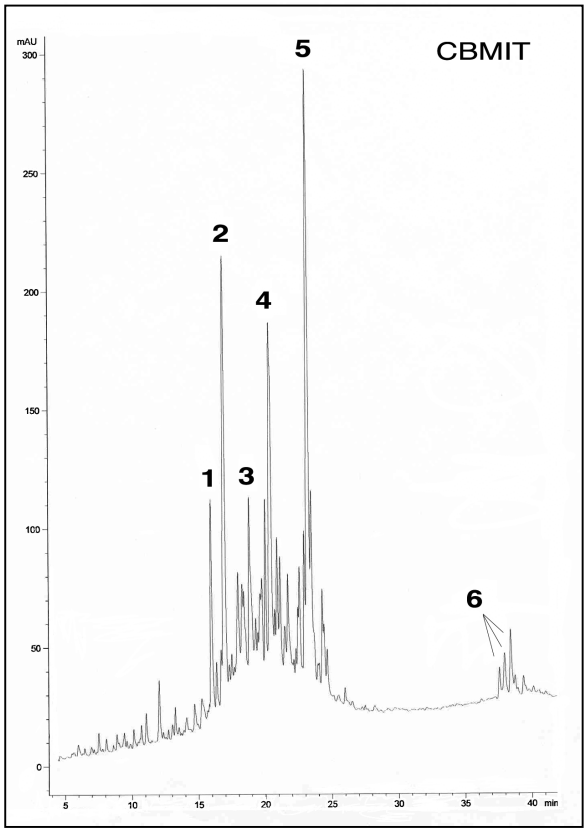


Fig. 28: RP-HPLC of crude **47a**.

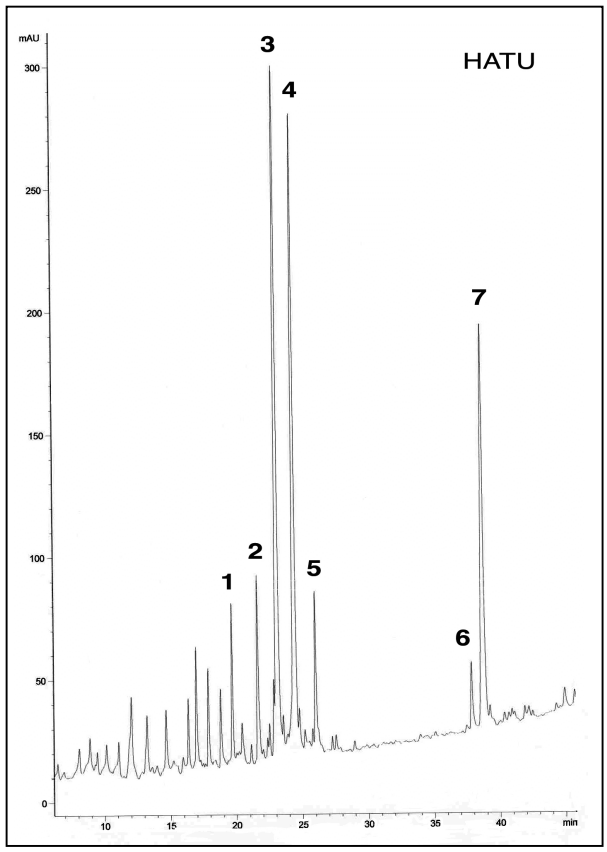


Fig. 29: RP-HPLC of crude **47b**.

Table 1: Peak analysis of the HPLC profile of **47a** (Fig. 28).

Peak	t_R (min)	Molecular Weight	Peptide fragment
1	15.7	802	RARR+Pmc
2	16.7	959	AARARR+Pmc
3	18.6	1030	AAARARR+Pmc
4	20.1	1101	AAAARARR+Pmc
5	22.9	1415	Target sequence
6	37.3-38.2	1138; 1209	A_5RARR+FMOC; A_6RARR+FMOC;

Table 2: Peak analysis of the HPLC profile of **47b** (Fig. 29).

Peak	t_R (min)	Molecular Weight	Peptide fragment
1	19.5	1272	A_{10}RARR
2	21.5	1344	Target sequence – 1 Ala
3	22.9	1415	Target sequence
4	24.3	1486	Target sequence + 1 Ala
5	25.9	1558	Target sequence + 2 Ala
6	37.9	1210	A_6RARR+FMOC
7	38.7	1419; 1489	A_9RARR +FMOC; A_{10}RARR +FMOC

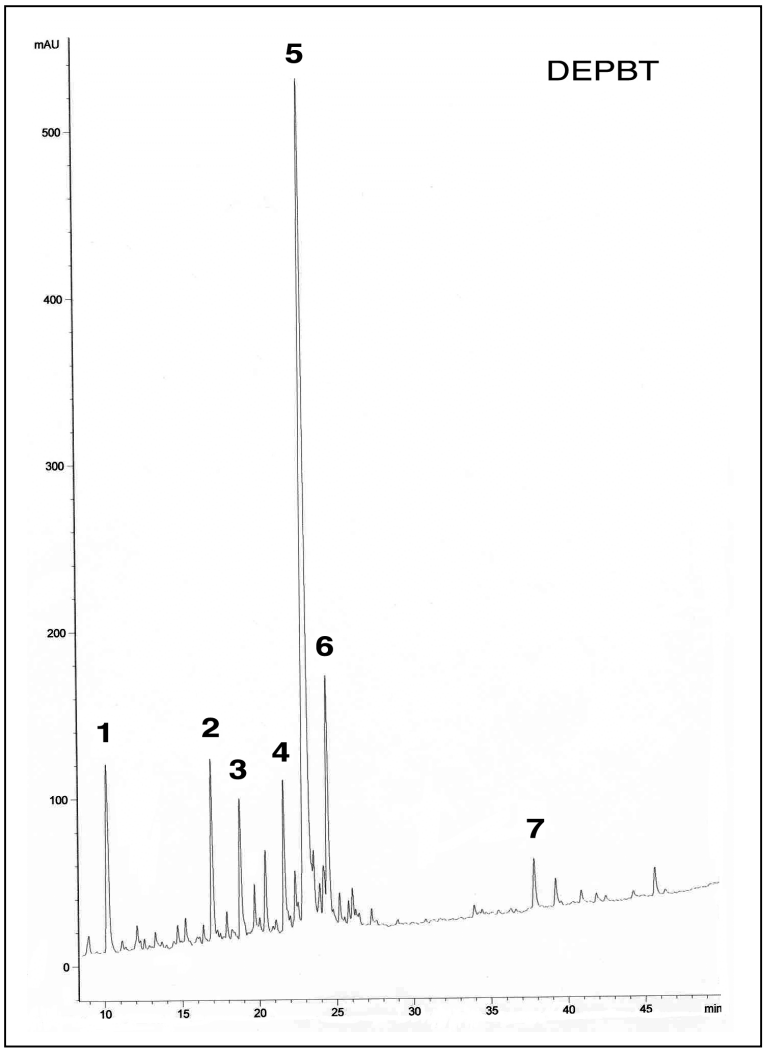


Fig. 30: RP-HPLC of crude **47c**.

Table 3: Peak analysis of the HPLC profile of **47c** (Fig. 30).

Peak	t_R (min)	Molecular Weight	Peptide fragment
1	10.2	602	ARR + FMOC
2	16.9	959	AARARR + Pmc
3	18.8	1030	AAARARR +Pmc
4	21.7	1343	Target sequence - 1 Ala
5	23.1	1415	Target sequence
6	24.5	1486	Target sequence + 1 Ala
7	37.9	1209	A_6RARR+FMOC

The three chromatograms show the presence of side products having shorter peptide sequences but also of products due to incomplete FMOC and Pmc deprotection. Due to the HPLC detection at λ=214 nm, these peaks are overestimated compared to the peaks corresponding to simple peptide sequences. Neverthless, they indicate some difficulties in the deprotection of these groups. The Pmc group can be difficult to deprotect when several arginines are present in a peptide, difficulties in the FMOC deprotection during the synthesis can only arise from aggregation of the peptide chains on the resin.

CBMIT was, by far, the worst activator but it was used in a different way compared to the others. For practical reasons, all the FMOC-amino acid was activated at the beginning of the synthesis when using CBMIT. We tried, therefore, to isolate CBMIT as a solid. This would allow the activation of the amino acid just prior the coupling, exactly in the same way used for HATU and DEPBT. Unfortunately, solid CBMIT, did not work as activator. Many syntheses were performed to try to find the reasons for its inactivity without success. It might be that the double charge renders it extremely sensitive to water traces and moisture, inactivating it very fast.

The activator selected after these trial syntheses was DEPBT which gave a cleaner product compared to the others.

2.3.6 Synthesis of $(Ala)_{12}$ and $(Ala)_{20}$ on an aminomethylpolystyrene resin

In order to evaluate the feasibility of the conjugate syntheses we planned, some model polyalanine peptides were synthesised using the aminomethylpolystyrene solid support which would be used for the construction of the conjugates.

The Wang resin, usually used in peptide synthesis, has very different swelling properties compared to our resin in polar aprotic solvents. This fact could interfere severely with the results already obtained in the comparison of the three activators.

Since the resin is functionalised with an amino group, the synthesised peptide was attached to the solid support through an amide bond and could not be cleaved and analysed. The syntheses were, therefore, monitored only by quantitative ninhydrin test every four couplings.

The activators used were DEPBT and CBMIT. In this synthesis, CBMIT was freshly prepared before every coupling in order to obtain a better comparison between the two coupling reagents. The peptide synthesised are:

$(Ala)_{12}$ (**48a**: CBMIT, **48b**: DEPBT)

$(Ala)_{20}$ (**49**: DEPBT)

The results we obtained are summarised in the following graphic:

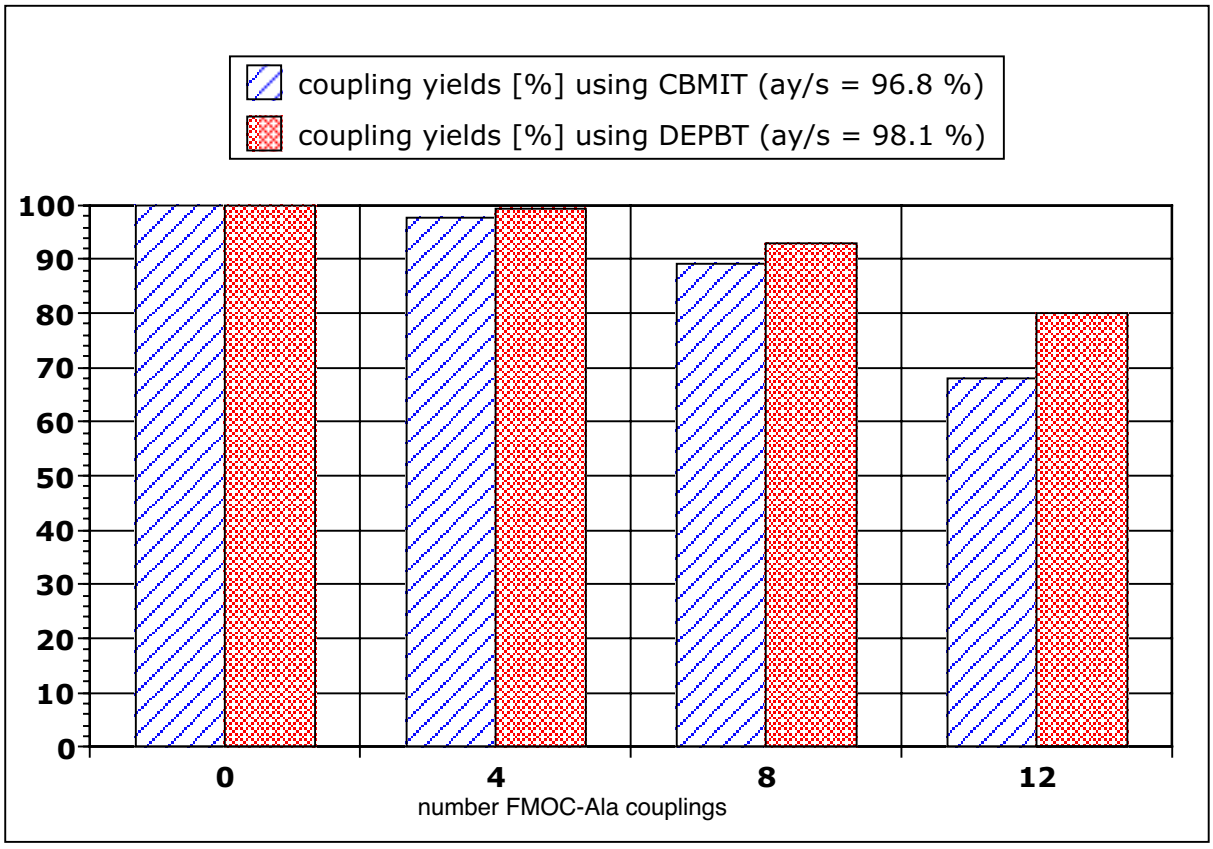


Fig. 31: Quantitative ninhydrin monitoring for the synthesis of **48a** and **48b**.

DEPBT, again, showed to be superior to CBMIT and was selected for the conjugate syntheses. The peptide **48b** was elongated to twenty amino acids. An overall yield of ~60% was obtained for **49**. This trial was necessary to verify the feasibility of a peptide containing 20 alanines in a reasonable yield.

2.3.7 Conclusions

The syntheses of polyalanine peptides revealed to be more complicated than expected, in particular the achievement of a homogeneous product seems to be problematic independently of the activating reagent. All the problems described for the synthesis of this difficult sequence were encountered (incomplete acylation, difficult deprotection, swelling difficulties). In addition, the dangerous 'overactivation' side reaction was observed in several cases. The high amount of activated FMOC-amino acid we used, on one side, helps to obtain a better acylation level, on the other hand can provoke the coupling of two residues instead of one leading to the formation of a very similar product (difference of only one residue) which results in difficult separation.

The use of the aminomethyl polystyrene resin (lower loading compared to the Wang resin) gave better yields, opening the possibility of obtaining a satisfactory result for the syntheses of the conjugates.

2.4 3'-Peptidyl-RNA conjugate synthesis

2.4.1 Introduction

The conjugate molecules we synthesised are all constituted of the same RNA strand which is modelled after the natural 3'-aminoacyl-tRNA^{Ala} (77mer) shortened by approximately two thirds (22 nucleotides). This oligonucleotide strand adopts a hairpin secondary structure stabilised by 7 base pairs and mimicks the aminoacyl acceptor stem of *Escherichia coli* tRNA^{Ala}.

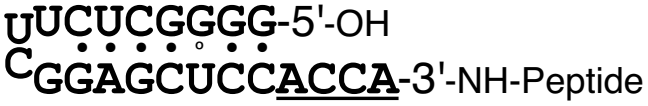


Fig. 32: RNA hairpin (bold) constituting the oligonucleotidic part of the conjugates.

The hairpin (Fig. 32) is closed by a stable UUCG tetraloop and contains a mismatched base pair at the third position. The underlined nucleotides are single-stranded and constitute the so-called CCA terminus ubiquitous in all tRNAs.

The peptidyl moieties linked to the RNA strands are designed so that they will have a strongly lipophilic character. Their sequence is based on a polyalanine or polyleucine stretch of varying length (8-22 amino acids), some with interdispersed or *N*-terminal glutamate residues to modulate their lipophilicity.

A peptide formed by a stretch of 15-20 alanine or leucine residues should adopt a stable α-helix conformation and could interact with lipid bilayer membranes (a hydrophobic segment of ~20 amino acids, arranged in a α-helix conformation, corresponds approximately to the thickness of a lipid bilayer).[106]

Alanine, leucine and glutamic acid, all early prebiotic amino acids, were chosen in view of evolutionary biological studies. In fact, if our conjugates will reveal capable of forming a stable α-helix, they could be used to give evidence for the existence of the hypothesised prebiotic 'ob-cells' or other kinds of self-driven compartment assemblies (Section 1.3). More important, the peptidic part gives our conjugates an amphiphilic character rendering them potentially capable of forming higher-order structures (micelle, vesicles, nanostructures) with, maybe, unexpected properties.[107]

2.4.2 Synthesis strategy

The conjugates were synthesised adopting a stepwise synthesis strategy. Fragment coupling would not be feasible for the synthesis of our conjugates because of the high insolubility of hydrophobic peptides. Moreover, the construction of our 3'-peptidyl-RNAs would require the activation of the carboxylic end of the peptide fragment with the danger of a high level of epimerisation. Indeed, this is one of the major drawbacks of the peptide fragment approach synthesis limiting its use to the coupling of peptide having a glycine or a proline as the C-terminal residue.

In our strategy the peptidic moiety was synthesised before the oligonucleotidic strand, so we had to develop a robust synthetic protocol for the peptide synthesis on this solid support; this involved the monitoring of the synthesis at every step to optimise the procedure.

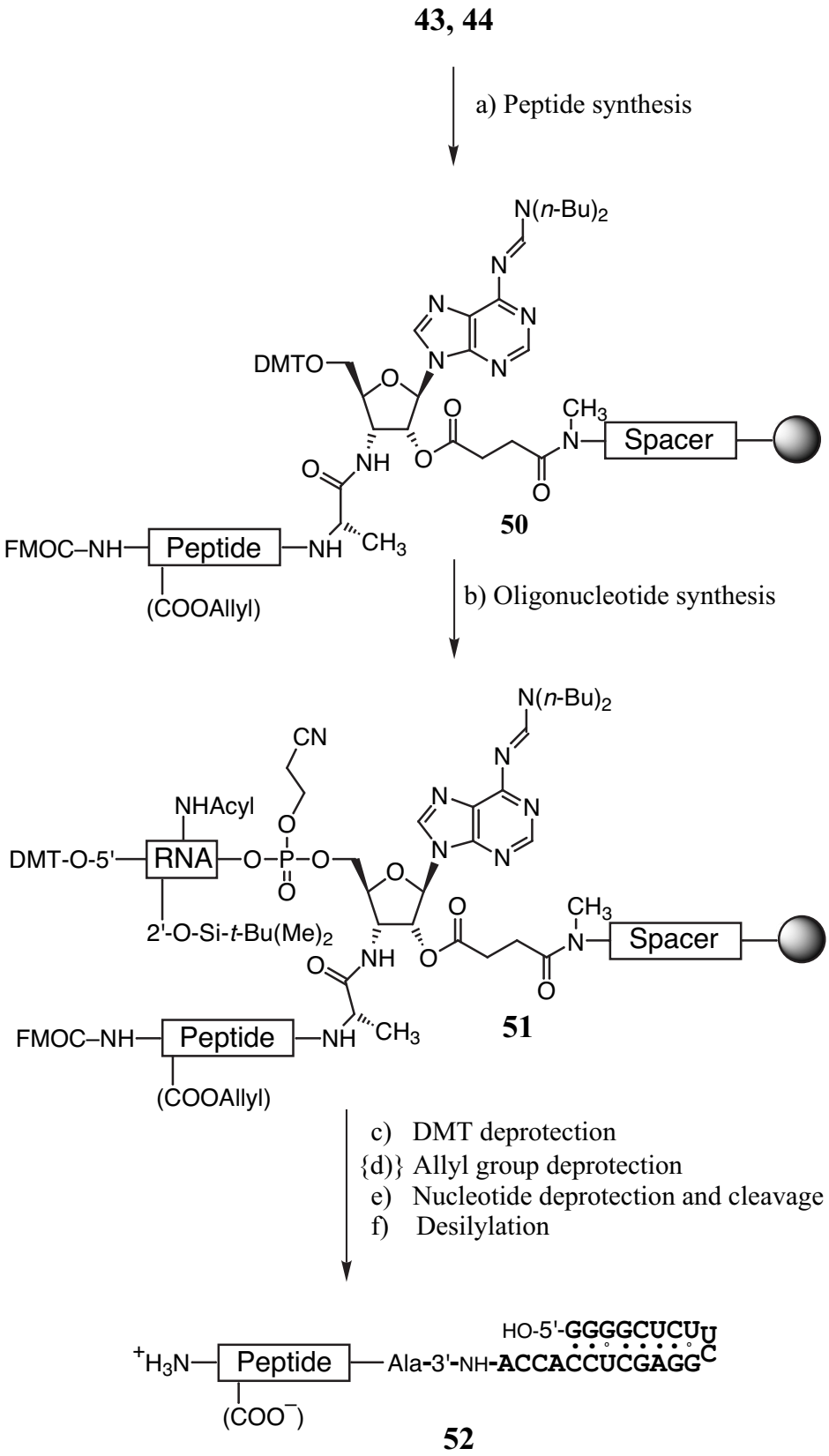
The peptide syntheses were carried out manually using a Teflon syringe as reaction vessel equipped with a polyethylene filter, while the oligoribonucleotide syntheses were carried out on an automated DNA/RNA synthesiser.

Starting from **43** or **44** a number of 3'-peptidyl-RNAs were synthesised according to the synthetic plan depicted in Scheme 32. The conjugates synthesised are listed in Table 4.

Table 4: Peptide sequences of 3'-peptidyl-RNA.

Conjugate	Sequence
52a	$^{-}O_3PO$-5'-RNA-3'-$(Ala)_8$
52b	HO-5'-RNA-3'-$(Ala)_{15,16}$
52c	HO-5'-RNA-3'-$(Ala)_{10}$Glu
52d [a]	HO-5'-RNA-3'-$(Ala)_{18}(Glu)_2$pGlu
52e	HO-5'-RNA-3'-$(Ala)_7$Glu$(Ala)_7$Glu$(Ala)_4(Glu)_2$
52f	HO-5'-RNA-3'-$(Ala)_{20\text{-}22}$
52g	HO-5'-RNA-3'-Ala$(Leu)_{18}(Glu)$
52h	HO-5'-RNA-3'-Ala$(Leu)_{18}(Glu)_3$

[a] pyroglutamate probably formed during the purification steps.


43, 44
a) Peptide synthesis
N(n-Bu)2
DMTO
CH3
Spacer
FMOC–NH
Peptide
NH
CH3
(COOAllyl)
50
b) Oligonucleotide synthesis
CN
NHAcyl
DMT-O-5'
RNA
2'-O-Si-t-Bu(Me)2
51
c) DMT deprotection
{d)} Allyl group deprotection
e) Nucleotide deprotection and cleavage
f) Desilylation
HO-5'-GGGGCUCUU
+H3N
Ala-3'-NH-ACCACCUCGAGGC
(COO−)
52

Scheme 32

The peptidic moiety was synthesised according to FMOC solid support peptide synthesis protocol using DEPBT as activating reagent. The syntheses were monitored by measuring the FMOC-release at every cycle or by quantitative ninhydrin test every four to six couplings.

The FMOC deprotection was carried out using a solution of 2% DBU, 2% piperidine in DMF instead of using 20% piperidine in DMF.[108] The addition of piperidine to DBU solution avoids dibenzofulvene precipitation and allows the utilisation of this mixture in a batchwise strategy. Incomplete deprotection can be associated with subsequent slow or incomplete amino acid coupling. In fact, a sign of the beginning of aggregation of the polypeptide chains is the difficulty of FMOC-release. So, also the deprotection conditions can influence the yield of a coupling. Dettin and coworkers obtained a substantial enhancement in the amount of the target peptide only replacing 20% piperidine with 2% DBU in the deprotection step of the synthesis of polyalanine peptides.[108] It has also been observed that the effectiveness of deprotection is more dependent on the number of deprotection cycles than on the total deprotection time.[109]

During the conjugate syntheses, two deprotection cycles of 8 and 5 minutes were perfomed, when the FMOC-deprotection was difficult, a third deprotection (3 minutes) was performed. At the end of the peptide synthesis, the N^{α}-terminal FMOC group was left on to avoid a reaction with the incoming phosphoramidite in the following ribonucleotide synthesis. Quantitative ninhydrin test evidenced a loss of the N^{α}-terminal FMOC group two or three days after the end of the peptide synthesis. In order to overcome this problem, the last amino acid was coupled just prior the oligonucleotide synthesis.

2.4.3 Cleavage, deprotection and purification of the conjugates

At the end of the oligonucleotide synthesis, the DMT group was deprotected (DMT-off) and a treatment with methyl amine allowed for the simoultaneous removal of the FMOC and all base-labile protecting groups on the nucleotides along with the detachment of the conjugate from the solid support. When glutamic acid was present in the peptide sequence, previous removal of its side-chain protecting group was accomplished (Section 2.4.7).

Finally, the 2'-OH-persilylated soluble conjugates were desilylated with neat $Et_3N{\cdot}3HF$ containing some DMF and precipitated from *n*-butanol. Some syntheses were carried out using phosphoramidite with the TOM protecting group. In this case, the

desilylation was carried out with 1M Bu_4NF in THF and the excess of fluoride was removed by gel filtration chromatography on NAP cartridges containing Sephadex G-10 resin.

The crude products were purified, first, by strong anion exchange (SAX) HPLC using a phosphate buffer at pH 7.0. After this first step of purification, the fraction collected after SAX-HPLC were desalted on RP-HPLC C_{18} column and concentrated to a volume of ~1ml on a Rotavap system. Normally, oligonucleotides are concentrated or lyophilised using a SpeedVac system. Unfortunately, this practical method caused the precipitation of a part of the products that was not possible to solubilise again.

After SAX-HPLC and desalting, RP-HPLC allowed (not in all cases) to obtain an homogeneous product. The fraction containing the target product was, finally, extensively desalted on a RP-HPLC column. The final desalting is very important for the storage of RNA which is easily degraded by RNAase and this process is favoured by the presence of salts. Complete desalting allowed the storage of the samples at 4 °C for months without degradation.

2.4.4 Analysis of the conjugates with MALDI-ToF MS (Matrix Assisted Laser Desorption/Ionisation Time-of-Flight Mass Spectrometry)

MALDI-ToF MS, introduced in the last 80s,[110] has become in the last years one of the most important method for the structural analysis of biological molecules such as proteins, carbohydrates and oligonucleotides. These method (together with ESI-MS) solved the difficult problem of generating ions from large, nonvolatile analytes such as biological probes without significant fragmentation. Ionisation is so soft that even noncovalent interactions may be maintained during the ionisation process. Due to the tendency of the method to generate singly charged ions, the spectra are simple to interpret. ToF mass detectors are most commonly used with MALDI. They are very sensitive (femtomole detection) and virtually unlimited in mass range. The ions are produced by short pulsed laser irradiation of the biomolecules embedded in a matrix which are strongly UV-absorbing organic compounds. The matrix absorbs the incident laser light and makes it possible to transfer the energy of the laser to the sample. Proton transfer from the matrix is supposted to be the major ionisation mechanism.[111] Even though the matrix is very important to obtain good spectra, it is not yet possible to select a matrix theoretically. There are some widely-spread matrixes used for some classes of compounds (for example: sinapinic acid for proteins, α-cyano-4-hydroxycinnamic acid for peptides, 2,5-dihydroxybenzoic acid for oligonucleotides), but for the analysis of non

common classes of compounds, like RNA-peptide conjugates, a suitable matrix must be found empirically. The optimum matrixes used for the single components of a hybrid molecule (the peptide and the oligonucleotide) generally clash. Unfortunately, the suitability of matrices used for peptides is inverse to that for oligonucleotides. It seems that the ionisation of a hybrid molecule roughly reflects the relative masses of the two components.[112] Our conjugates are composed by ~20-25% of peptide (in mass), so a matrix used for oligonucleotides was employed (2,4,6-trihydroxyacetophenone) for the analysis.

2.4.5 First trial of synthesis

The first conjugate to be synthesised had the following sequence:

$$HO\text{-}5'\text{-}RNA\text{-}3'\text{-}(Ala)_{20} \qquad \mathbf{53}$$

(We abbreviate here the hairpin sequence as RNA). An old batch of resin having a starting loading of 5 μmol/g was used. Since the building block **10** could be attached on the solid support almost quantitatively using a dilute solution (14 or 30 mM), the synthesis of the peptidic moiety was started using a 30 mM solution of FMOC-amino acid to activate (20 equivalents). DEPBT was used as activator for all the conjugate syntheses.

The synthesis of **53** was monitored only through the UV-detection of the FMOC-deprotection level. After the first FMOC-amino acid coupling, the monitoring indicated a coupling yield of only 32%. It was clear that the concentration of the FMOC-amino acid was too low, so the following coupling was performed using 100 equivalents of FMOC-amino acid in a 0.13 M solution. Again, monitoring indicated an incomplete coupling. All the following couplings were thus carried out using ~200 equivalents in a 0.2 M solution. The synthesis was completed but, due to the low level of absorption of the FMOC deprotection solution, it was not clear if these values corresponded to the background or they really indicated FMOC deprotection of the resin-bound peptide (Fig. 33) so that the synthesis of the RNA part was not performed.

This first trial indicated clearly that the low loading of the resin rendered the synthesis on our solid support quite different from a classical peptide synthesis were normally 4-6 equivalents of FMOC-amino acid are used. The activation process and the coupling being bimolecular reactions, a low concentration of the reagents disfavours the kinetics (unless the reaction is performed for a long time). The successive syntheses were performed using a higher concentration of FMOC-amino acid (0.6-1.6 M) corresponding to an excess of 200-500

equivalents. At the same time, using a high excess of FMOC-amino acid, we exposed to the risk of overactivation (Section 2.3.4).

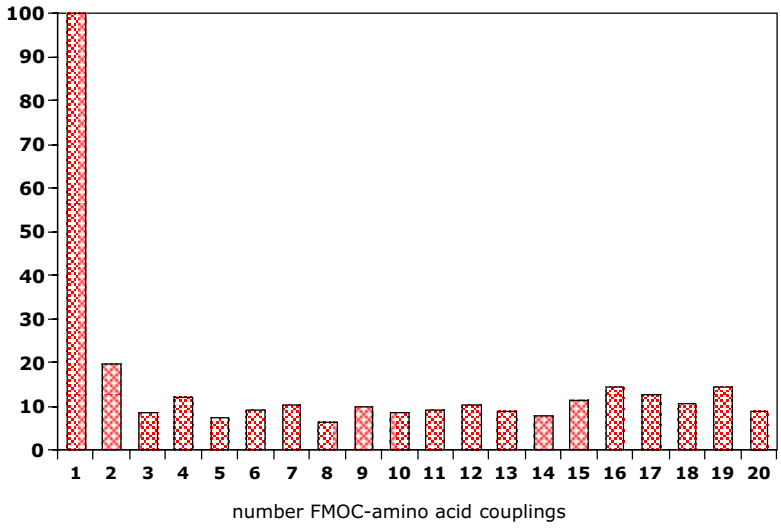


Fig. 33: FMOC-deprotection monitoring for the synthesis of **53**.

The first syntheses of the nucleotidic part revealed, as well as the peptidic part, difficulties in the couplings. Consequently, the coupling time was prolonged from 5 to 13-15 minutes. It might be that the peptide chain causes steric hindrance during the RNA synthesis, rendering more difficult the couplings.

2.4.6 Synthesis and purification of RNA-$(Ala)_{15,16}$ (52b) and RNA-$(Ala)_{20\text{-}22}$ (52f)

The syntheses of **52b** and **52f** were carried out using both solid supports **43** (short spacer) and **44** (long spacer). On the solid support **43** two peptides were planned to be synthesised: Ala_{21} **50f** and Ala_{15} **50b** starting with a concentration of FMOC-amino acid of ~0.4 M. After the 6^{th}-7^{th} amino acid, difficulties in the release of the FMOC appeared and the concentration of FMOC-amino acid was augmented to 0.8-1 M. Quantitative ninhydrin test at the end of the peptide syntheses gave an overall yield of 57% and 71% for **50f** and **50b** respectively.

The synthesis of **50f** was repeated on the solid support bearing the long PEG spacer **44** in order to analyse the influence of the solid support on the peptide synthesis. From the beginning, difficulties appeared in the coupling of FMOC-alanine on this support, so a higher concentration of FMOC-amino acid was used and the reaction time was prolonged from the 4^{th} coupling. Apparently, from the FMOC deprotection values, this procedure worked well because a higher coupling yield per step could be obtained while for the synthesis of **50a** a

constant decrease of the yield was observed (Fig. 34). The ninhydrin test, however, indicated an overall yield of 55% instead of the apparent 70% yield resulting from the value of the FMOC deprotection. After SAX-HPLC, MALDI analysis evidenced the presence of products with one or two alanines more than the synthesised sequence, indicating that probably the high FMOC-deprotection levels resulted from a small amount of overactivation reaction at several steps.

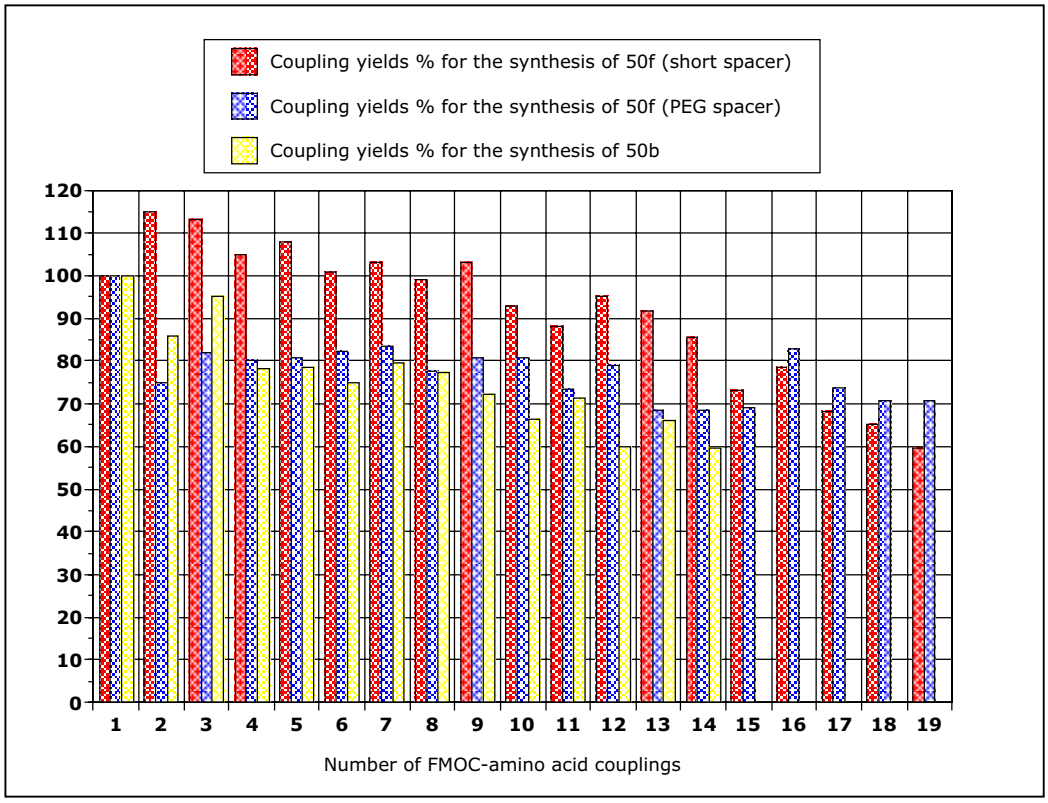


Fig. 34: FMOC-deprotection monitoring for the synthesis of **50b** and **50f**.

At the end of the peptide synthesis, the oligonucleotide strand was synthesised obtaining the following overall yields: 67.2% for **52b**, 55.8% for **52f** (short spacer) and 59.8% for **52f** (PEG spacer). These latter data seems to indicate that the influence of the spacer is minimal on the overall yield of the oligonucleotide synthesis. These results were later confirmed by the other conjugate syntheses. Also the influence of the peptide on the RNA synthesis cannot be rationalised; neither the length of the peptide nor its hydrophobicity have a tendential influence (positive or negative) on the oligonucletide synthesis yield. The values obtained seems to indicate more of a stochastic influence of the peptide and of the spacer on the RNA overall yield.

The purification of the three products revealed to be very problematic. After precipitation from *n*-butanol, part of the pellets could not be redissolved. This part revealed to be almost the total quantity for **52f** (short spacer) and no target product could be recovered from this synthesis. **52f** (PEG spacer synthesis) also was partly lost because of insolubility. Even with a shorter alanine sequence (**52b**) a part of the product was lost because of insolubilty, albeit to a lesser extent. We thought that the 22 negative charges of the RNA part could give the constructs enough hydrophilicity to diminish the insolubility problem that we had already encountered with the syntheses of **46a** and **46b**. Attempts to redissolve this precipitate with organic solvents mixtures, 6 M guanidinium hydrochloride (pH 8) or a detergent (Triton-X) were unsuccessful. Only neat TFA could redissolve these pellets meaning that no purification was possible.

The soluble part of the crude products was subjected to SAX-HPLC purification. The product eluted as two main peaks both for **52b** and **52f** (Fig. 35).

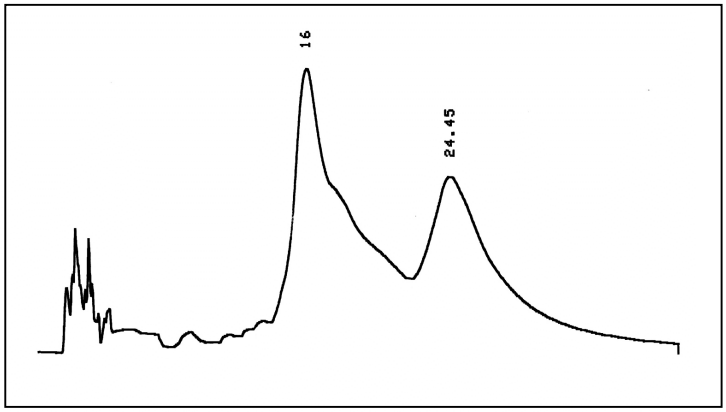


Fig. 35: SAX-HPLC profile of the crude **52f**.

Fractions from SAX-HPLC, after desalting and concentration, were subjected to RP-HPLC purification. The target product resulted to be present only in the first peak (MALDI analysis). Several attempts to obtain MALDI spectra from the second SAX peak of this and of the other conjugates were unsuccessful. Therefore, in this and all subsequent conjugate syntheses only the first SAX peak was used for further purification.

The RP-HPLC profiles revealed the presence of a very broad peak eluting in front of all the other peaks for both compounds. For **52f**, this was the main peak, so that no RP-HPLC purification was carried out.

MALDI-ToF analysis of **52f** revealed the presence of four main products corresponding to conjugates having 20, 21 (target sequence), 22 and 23 alanines. Deleted sequences composed of 9, 10 and 11 alanines were also detected from MALDI (Fig. 36).

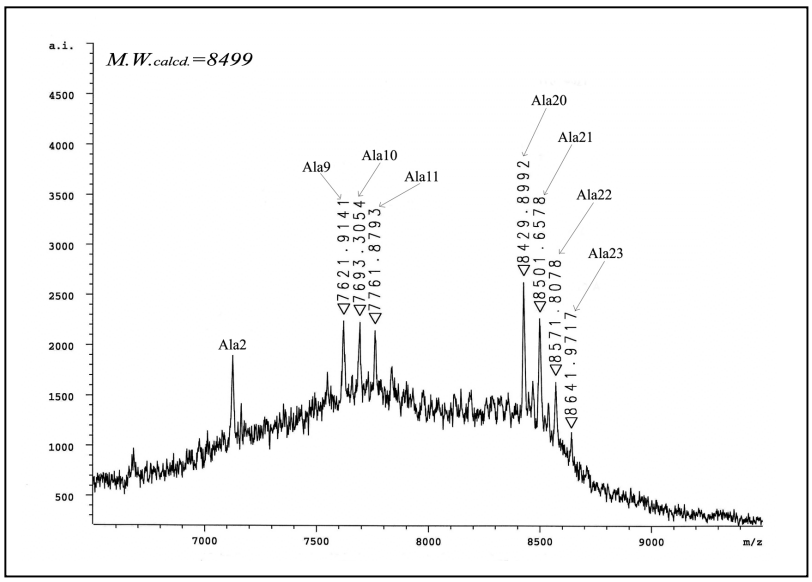


Fig. 36: MALDI-ToF spectrum of **52f** (SAX-HPLC purified).

RP-HPLC profile of **52b** (Fig. 37) showed, again, the presence of the broad peak eluting in front of the others (t_R~10-15 min). This first peak, which was present in the chromatogram of all the conjugates to a higher or lower extent (see below), resulted in poorly ionizable material that could not be analysed properly by MALDI-ToF. We hypothesise that this broad peak is formed by the aggregation of some molecules of the conjugates. Our conjugates have an amphiphilic character that might lead to the formation of micelles or other kinds of aggregates masking the hydrophobic part, hence, shortening the retention time of this material on RP-HPLC. Unfortunately, the presence of these 'aggregates' considerably lowered the yield of the final, purified material.

The RP-HPLC chromatogram of **52b** evidenced, apart from the broad peak, two other main peaks. The two peaks could be isolated and purified to homogeneity (Fig. 38 and 39). MALDI MS indicated that they corresponded to the target conjugate (Peak 2) and to the target product containing an alanine more (Peak 3). The MALDI spectra could be only obtained with big difficulties, a high laser power was needed to desorb the product causing depurination of the RNA strand as can be seen in the MS spectra.

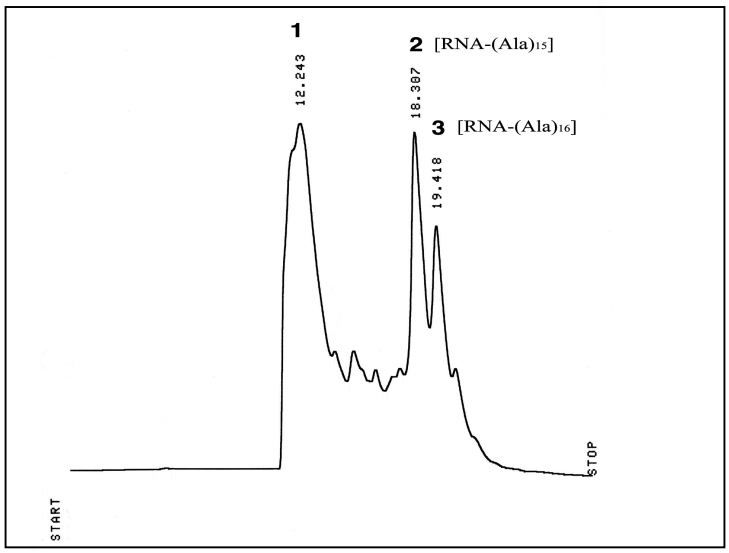


Fig. 37: RP-HPLC profile of **52b** (SAX-HPLC purified).

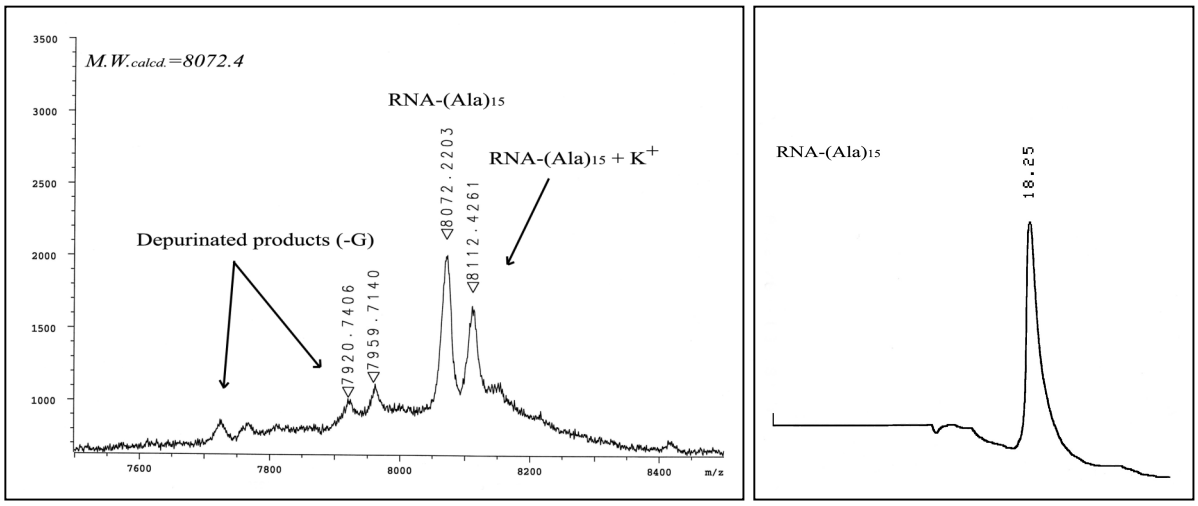


Fig. 38: MALDI and RP-HPLC profile of purified **52b** (Peak 2).

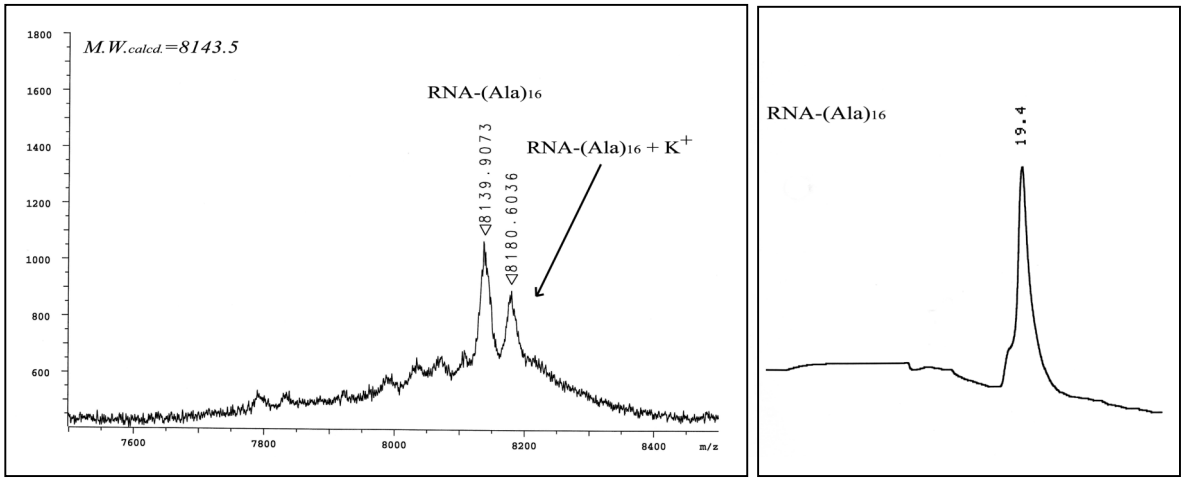
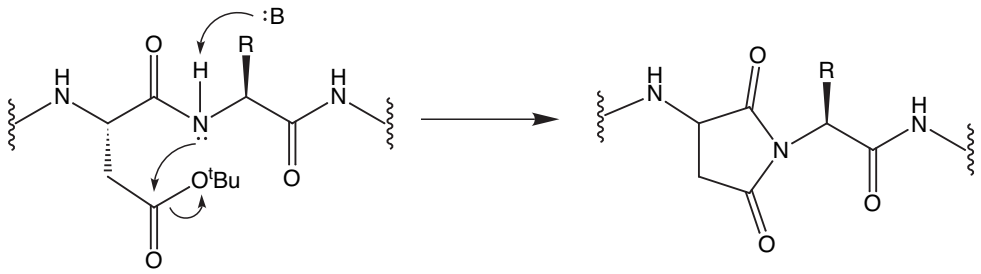


Fig. 39: MALDI and RP-HPLC profile of purified **52b** (Peak 3).

2.4.7 Introduction of glutamic acid into the peptide sequence

The insolubility problems forced us to introduce into the peptide sequence hydrophilic residues in order to diminish the hydrophobicity of this part of the molecule and, thus, eliminate the aggregation. A residue bearing a negative charge was chosen because aspartic and glutamic acid are early prebiotic amino acids while the positive charged amino acids (Lys, Arg) are probably of later biosynthetic origin. In addition, other work on DNA-conjugate has shown that arginine and lysine within a covalently attached peptide strongly interacts with the DNA backbone due to Coulomb attraction. We wished to avoid that kind of idiosyncrasy at this stage.[113]

Glutamic acid was preferred to aspartic acid because this latter can give more problems during the synthesis, in particular it can form the dangerous base-catalysed aspartimide which, after nucleophilic ring opening by piperidine, can lead to the formation of β-peptides and piperidine adducts (Scheme 33).[114]



Scheme 33: Mechanism of aspartimide formation during peptide synthesis.

The glutamic acid side chain protecting group must be deprotected before the final basic treatment (CH_3NH_2) to avoid undesired ester aminolysis generating N^{δ}-methyl glutamine. Two glutamic acid protecting groups, orthogonal to peptide and oligonucleotide synthesis, were tested through the synthesis of small model peptides: the Dmab group[115] {*N*-[1-(4,4-dimethyl-2,6-dioxocyclohexylidene)-3-methylbuthyl]-amino} benzyl ester (Fig. 40) and the allyl group.[116]

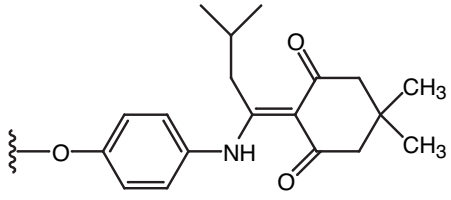


Fig. 40: Dmab protecting group.

The Dmab group is quasi-orthogonal to the FMOC chemistry since it is stable to 20% piperidine treatment and is removed with 2% hydrazine whereas the FMOC group is not stable to hydrazine treatment. The deprotection trials we executed, showed incomplete deprotection of the Dmab group after hydrazine treatment. Moreover, a recent work by Johnson showed that the Dmab group can favour the formation of truncated N^{α}-pyroglutamyl peptides via intramolecular cyclisation during the FMOC removal.[117] Deprotection difficulties and possible side-reactions made us choose the allyl group that, in our deprotection trials, could be removed without any problems. The allyl group is orthogonal to peptide and oligonucleotide chemistries and is removed under mild conditions by Pd^0-catalysed allyl transfer in the presence of a nucleophile acting as a scavenger. $Pd(PPh_3)_4$ was used as the catalyst and $PhSiH_3$ as the allyl scavenger.

After the palladium treatment, the conjugate (still bound to the solid support) was washed with a solution of *N,N*-diethyldithiocarbamate to ensure the removal of any palladium contaminant. Atomic absorption spectroscopy indicated that circa 1 Pd atom/DNA chain was present as contaminant while, after *N,N*-diethyldithiocarbamate washing only 0.003 Pd atom/DNA chain.[118]

A number of conjugates containing glutamic acid residues were synthesised (see Table 5): **52c**, **52d**, **52e**, **52g**, **52h** and, unsuccessfully, **52i** the sequence of which is: HO-5'-RNA-3'-$(Ala)_{14}Glu(Ala)_5$.

Table 5: Predicted percentage of helicity for the peptidic moiety of the conjugates. Conditions: pH 7.5, T=0 °C, ionic strength = 0.1 M.

Conjugate	Sequence	Predicted helicity [%]
52a	$^{-}O_3PO$-5'-RNA-3'-$(Ala)_8$	2.41
52b	HO-5'-RNA-3'-$(Ala)_{16}$	33.78
52c	HO-5'-RNA-3'-$(Ala)_{10}Glu$	13.15
52d	HO-5'-RNA-3'-$(Ala)_{18}(Glu)_3$	79.58
52e	HO-5'-RNA-3'-$(Ala)_7Glu(Ala)_7Glu(Ala)_4(Glu)_2$	70.00
52f	HO-5'-RNA-3'-$(Ala)_{20\text{-}22}$	68.62
52g	HO-5'-RNA-3'-$Ala(Leu)_{18}(Glu)$	96.68
52h	HO-5'-RNA-3'-$Ala(Leu)_{18}(Glu)_3$	94.11

We introduced 1, 2, 3 or 4 negative charges along the peptide sequence. The position of the glutamate residue was selected on the basis of the theoretical α-helix content of the sequence. The charges were placed in such position as to maximise the α-helix content predicted using the Web-accessible application AGADIR.[39] In Table 5, the predicted helicity (calculated with AGADIR) for the conjugates we synthesised are listed.

All the syntheses containing glutamic acid were performed using less concentrated activated FMOC-amino acids (0.3-0.4 M) compared to the previous ones, in order to minimise or avoid the pernicious overactivation reaction.

The synthesis of **52i** was a failure. The product, after SAX-HPLC chromatography was subjected to RP-HPLC purification. RP-HPLC profile revealed the presence of many peaks but no major one so that no target compund could be recovered.

The synthesis of **52c** was carried out starting with a low concentration of FMOC-amino acid (0.15-0.2 M). Since the FMOC-deprotection assay revealed a constant and rapid decrease of the coupling yields, the concentration was augmented to ~0.5 M for the second half of the peptide synthesis. Ninhydrin test indicated an overall yield ~30%. The RNA part was assembled as usual but a technical problem during the monitoring did not allow the measurement of the oligonucleotide overall yield.

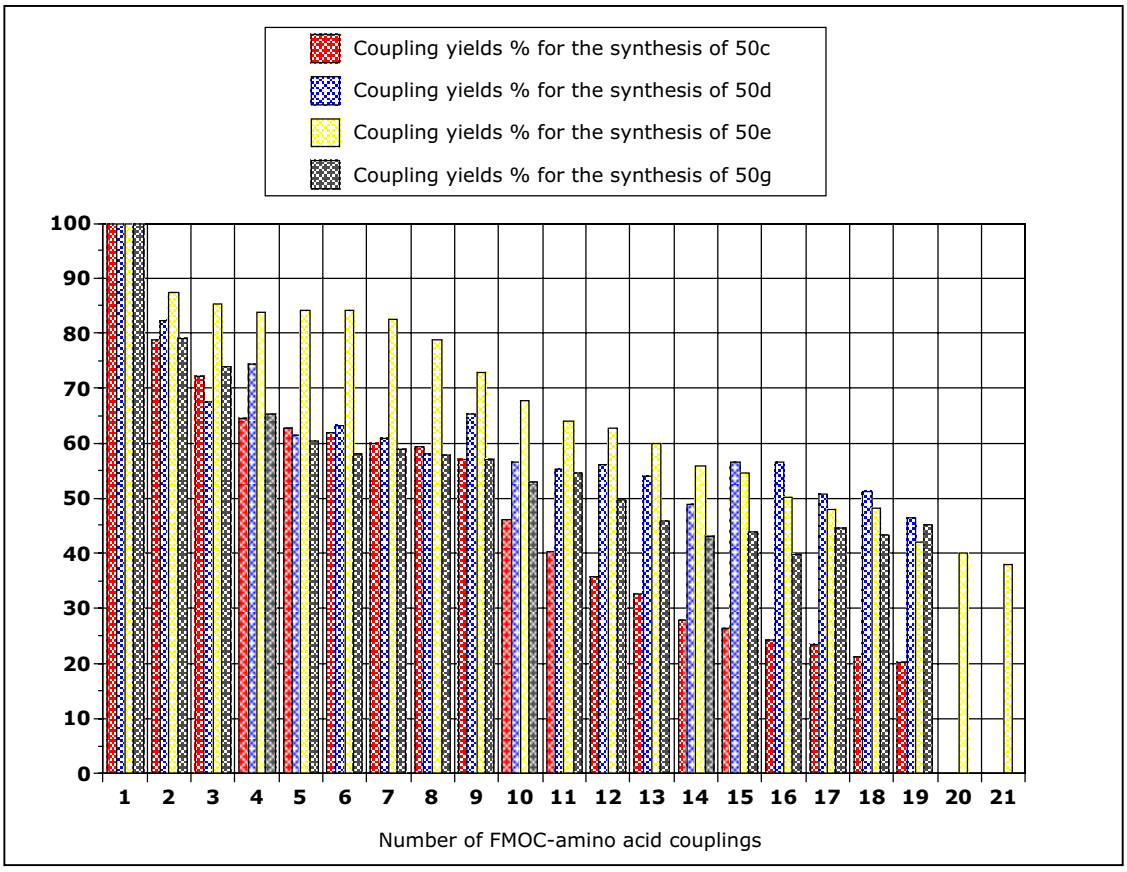


Fig. 41: FMOC-deprotection monitoring for the synthesis of **50c**, **50d**, **50e** and **50g**.

After SAX-HPLC, the RP-HPLC profile for **52c**, revealed the presence of several peaks. Also in this case the broad peak (1) was present as major product (Fig. 42).

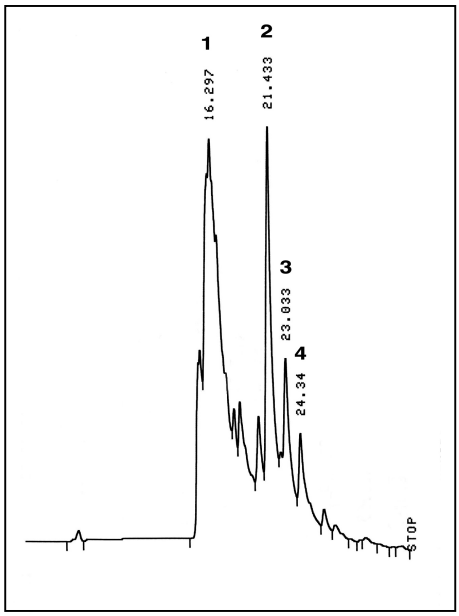


Fig. 42: RP-HPLC of **52c** (SAX purified).

MALDI-ToF of peaks 2 and 3 revealed that they do not correspond to the target sequence RNA-$(Ala)_{10}Glu(Ala)_9Glu$ but to a conjugate bearing the peptidic moiety $(Ala)_{10}Glu$ (Peak 2) and to $(Ala)_9Glu$ (Peak 3). This was really unexpected because both the ninhydrin and FMOC-monitoring indicated an overall peptide yield of around ~20-30%. The most plausible explanation is that the peptide was cleaved (probably during the last basic deprotection). Nevertheless, peak 2 was purified to homogeneity and used for further analysis. (Fig. 43).

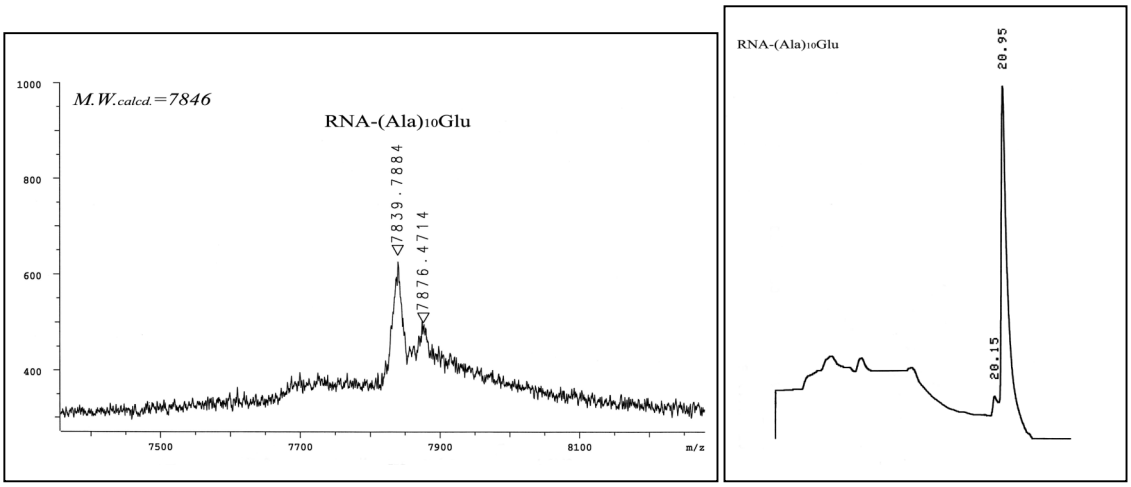


Fig: 43: MALDI and RP-HPLC profile of purified **52c** (Peak 2).

Three negative charges were then introduced in the successive synthesis (**52d**), since aggregation problems were still present for the conjugate **52c**. **52d** was synthesised using a slightly higher FMOC-amino acid concentration from the beginning on (0.4-0.5 M) since the synthesis of **52c** showed a decrease of yield at every coupling (Fig. 41). An overall yield of ~ 36% was obtained for the peptidic part and of 67% for the RNA strand.

SAX-HPLC evidenced the presence of the two peaks as for all the other conjugates but no more to a ratio of ~1:1 (as for the more hydrophobic sequences) but as a minor product. (Fig. 44).

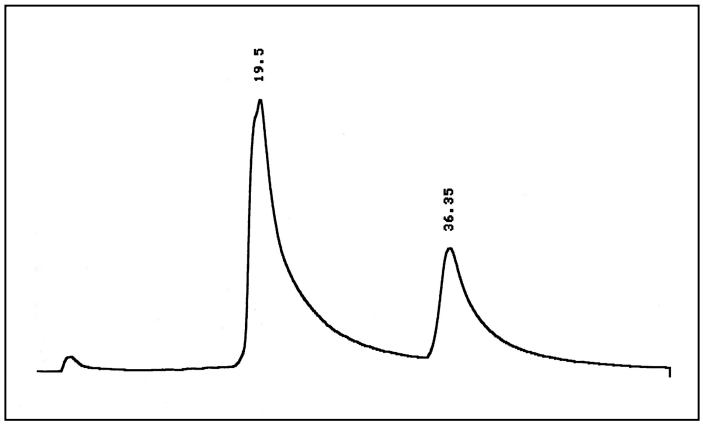


Fig. 44: SAX-HPLC of crude **52d**.

RP-HPLC of the SAX purified product showed again the presence of many peaks (Fig. 45).

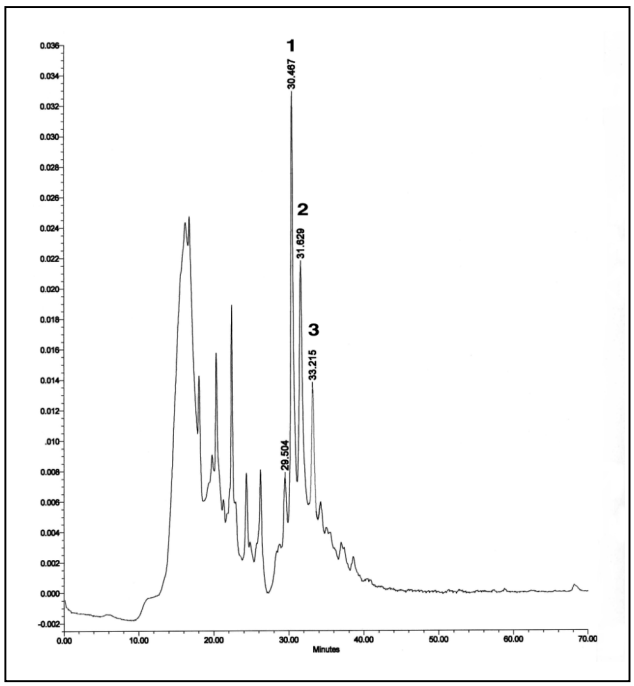


Fig. 45: RP-HPLC of **52d** (SAX-purified).

Not even three negative charges could eliminate the aggregated product. Peak 1, 2 and 3 were collected and analysed by MALDI which revealed peak 1 being the target sequence with the terminal glutamate cyclised as pyroglutamate. Peak 2 and 3 correspond to conjugates with one and two alanines more, respectively. Again, overactivation had been produced during the synthesis. Consequently the following syntheses were carried out using ~0.3 M or less of activated animo acid solution.

In the synthesis of **52e**, four negative charges were interdispersed in the peptide chain. Their relative position should avoid the formation af an amphipatic helix because this would not solve the aggregation problems. If the charges are all situated on the same side of the helix, lateral aggregation is favoured. In order to avoid this phenomenon, the glutamic acid residues were positioned as fas as possible as shown in the Figure 46 (always in compatibility with a high α-helix content).

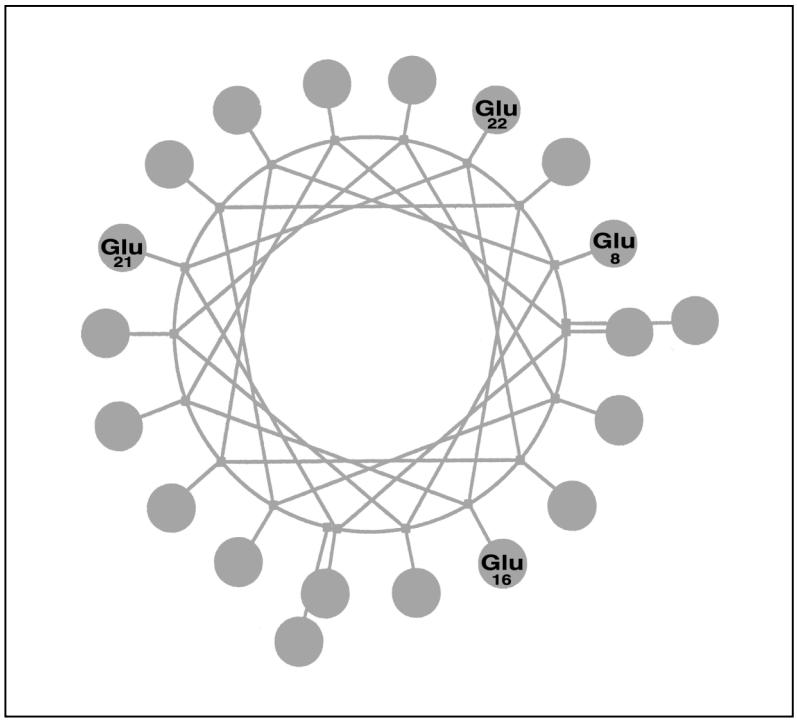


Fig. 46: Helical wheel projection of the peptidic moiety of **52e**.

The synthesis of **52e** was achieved with an overall yield of 24% for the peptide and 55% for the RNA part. After SAX purification (two peaks still present), the RP-HPLC profile showed a quite good product with no overactivation peaks. Unfortunately, the first peak eluting at t_R~12-15 minutes, is still present, indicating that not even four interdispersed charges could eliminate its presence (Fig. 47). The conjugate was purified to homogeneity and MALDI MS confirmed it to be the target sequence.

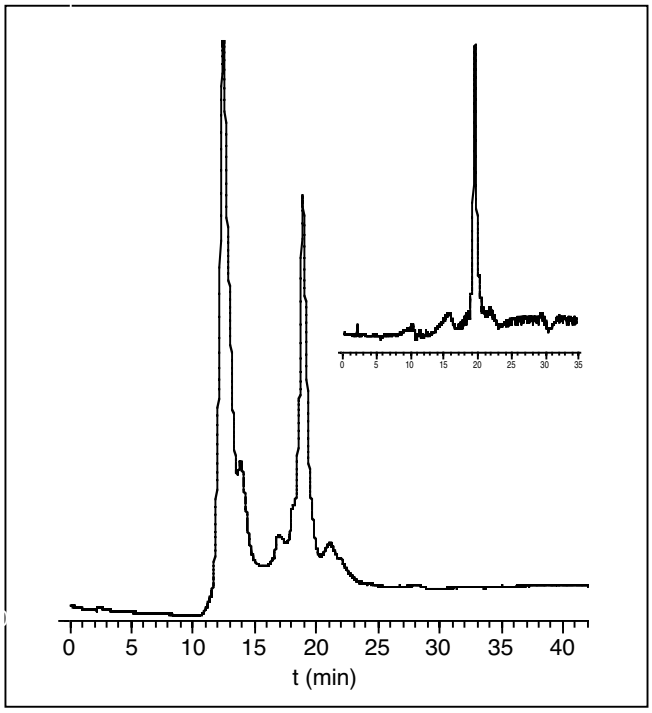


Fig. 47: RP-HPLC of **52e** (SAX-purified) and final purified conjugate (frame).

52g, containing a long stretch of 18 leucines, was synthesised with an overall yield of ~40% for the peptidic part and 55% for the RNA. SAX-purification showed again, the presence of two main peaks to a ratio of ~1:1 (Fig. 48).

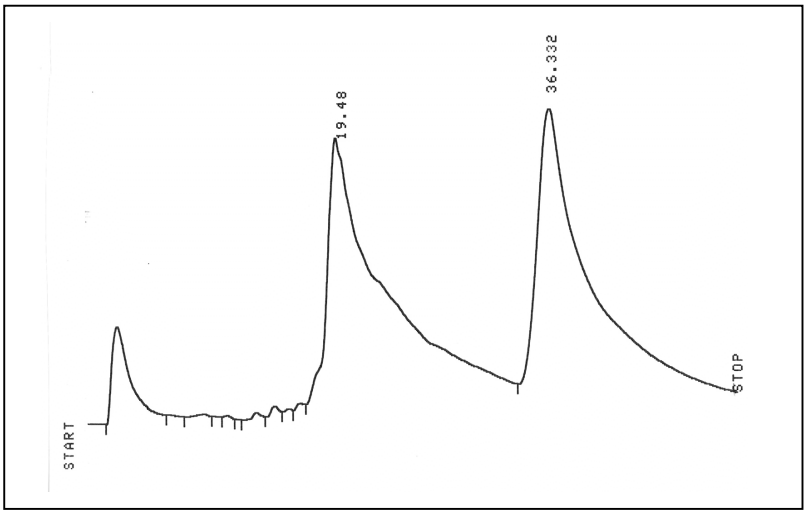


Fig. 48: SAX-HPLC of crude **52g**.

RP-HPLC chromatography revealed that only one broad early peak was present (‘aggregates’) and no peaks corresponding to monomeric conjugates were present. (Fig. 49).

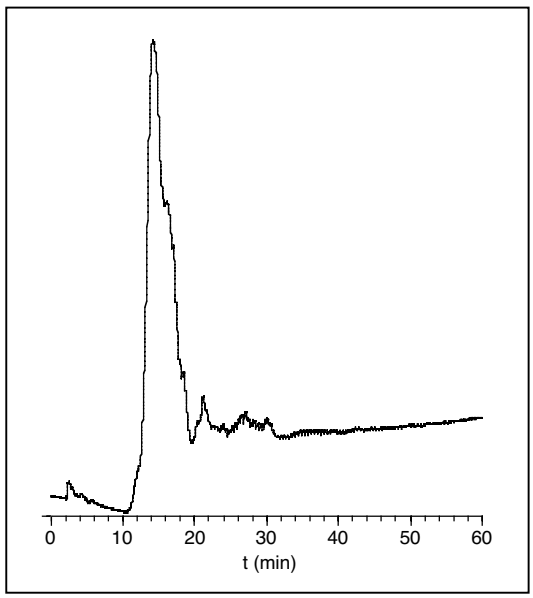


Fig. 49: RP-HPLC of **52g** (SAX purified).

Leucine is more hydrophobic than alanine, more aggregation problems could be possible but we did not expect it to occur to this extent. Already after the methylamine treatment, the solution was milky, indicating the formation of an emulsion. MALDI analysis showed only some broad peaks corresponding to short peptide sequences. This behaviour in the MALDI could originate from strong aggregation of the molecules that prevents the ionisation and desorption of the sample. If this is the case, only the short truncated sequences are visible in the spectrum.

Although the MALDI spectrum does not correspond to the correct molecular weight such that we do not have a proof of identity of the target conjugate, this product was used for further analysis. The same can be written for **52h**. The presence of two more negative charges did not change the situation.

2.4.8 Conclusions

Our syntheses showed that the synthetic strategy we developed, allows the synthesis of 3'-peptidyl-RNA conjugates. However, two main problems should be avoided: the overactivation and the aggregation.

The overactivation side reaction can be limited using a low concentration of activated FMOC-amino acid even though this causes a lower overall yield. A compromise should be found between an acceptable yield and a low level or absence of overactivation. After the

results, we can deduce that it is better to have lower amounts of product but more homogeneous material, thus, easier to purify.

Aggregation, in our case, depends on the peptide sequence since the unpeptidylated hairpin does not exhibit any aggregation problem. Since we planned to synthesise amphiphilic structures, this problem cannot be avoided. The more hydrophobic the peptidic part is, the more severe aggregation problems were encountered leading to a lower amount of monomeric, homogeneous material recovered.

Analysis of the synthesis monitoring (Fig. 34), indicates that the second coupling is the most difficult, probably due to steric hindrance. It can also been evidenced that starting from the 10^{th} amino acid, FMOC-deprotection difficulties arise, indicating aggregation of the peptide chains. A high FMOC-deprotection detected also in the second half of the peptide synthesis may be an indication of overactivation, as could be observed for the synthesis of **52d**. As already written, the FMOC-release monitoring should be regarded only as a qualitative indication of the course of synthesis.

2.5 Conjugate secondary and quaternary structure analysis

2.5.1 CD spectroscopy

Some of the conjugates were analysed by CD spectroscopy to study the secondary structure conformation adopted by the peptidic moieties and its influence on the hairpin secondary structure.

In the literature, examples of CD spectra of peptide-oligonucleotide conjugates are very rare.[72c,119] In contrast, CD spectroscopy is widely used to study protein-nucleic acid interactions.[120] In this kind of CD spectra, peptide and protein secondary structures generally dominate the CD at wavelength below 250 nm, while the region from about 250 nm to 320 nm provides contributions of the secondary structure of nucleic acids. Although aromatic amino acids can contribute to the CD in the 250-320 nm region, it is usually small relative to the CD of polymeric nucleic acids.

RNA molecules adopt an A-type helix conformation. Characteristic of the A-form CD is a positive band centered around 260 nm, a fairly intense negative band around 210 nm and a very intense positive band at about 190 nm.[121]

The following conjugates were analysed:

RNA (the unpeptidylated haipin as reference)	**X**
RNA-$(Ala)_8$	**52a**
RNA-$(Ala)_{16}$	**52b**
RNA-$(Ala)_{18}(Glu)_2$pGlu	**52d**
RNA-$(Ala)_7$Glu$(Ala)_7$Glu$(Ala)_4(Glu)_2$	**52e**

The spectra were measured at 0, 25 and 60 °C, in phosphate buffer at pH 7.5 and ~0.1 M ionic strength. The region scanned was between 200 and 350 nm. The strong absorbance of the buffer below 200 nm precluded searching for the α-helical characteristic maximun between 190 and 200 nm. Consequently, attention was focused on the region between 208 and 222 nm, where peptide α-helices exhibit a minimum and in the region around 260 nm where the RNA part exhibit a maximum.

The RNA hairpin alone exhibits a T_m value of 87.6 °C under the conditions used for CD spectroscopy and is completely folded (stable A_{260} baseline) up to ~60 °C.[122]

Figures 50, 51 and 52 show the CD spectra of the conjugates at 0, 25 and 60 °C respectively.

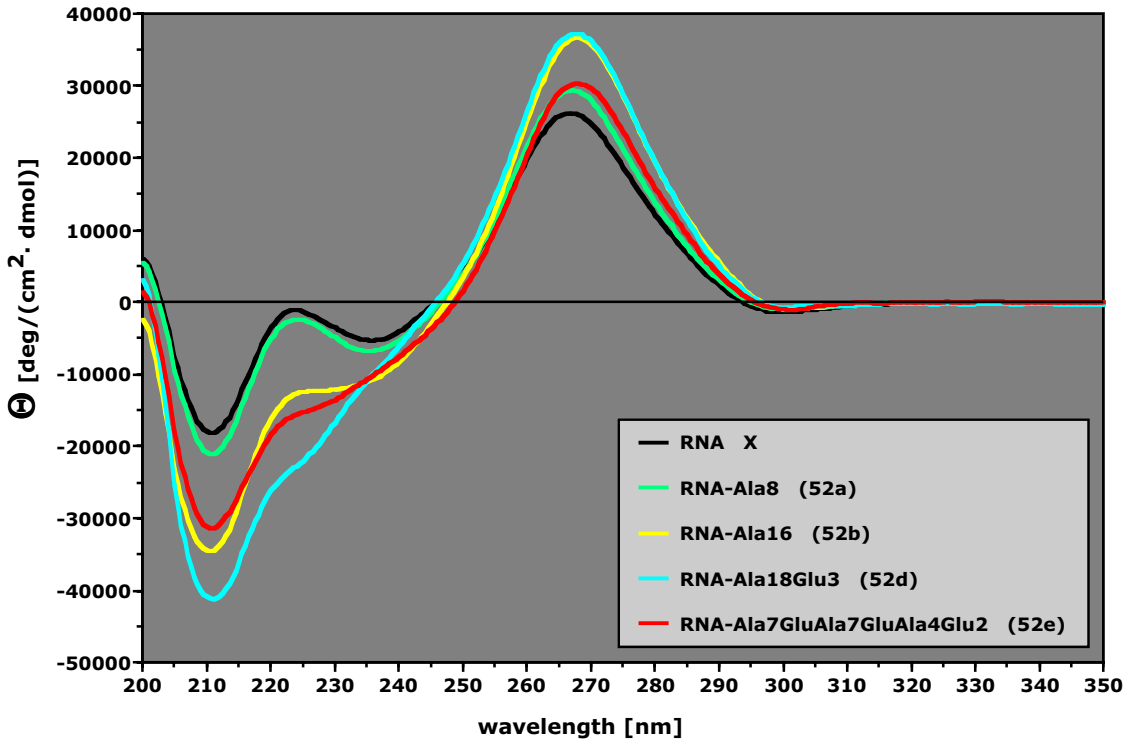


Fig. 50: CD spectra at 0 °C of a ~2.4 µM solution (A_{260nm}, 25 °C ≈ 0.35, $\varepsilon_{260,calcd}$ = 145'100 M^{-1} cm^{-1}, 100 mM NaCl, 10 mM $Na_xH_yPO_4$, pH 7.5) of **52a**, **52b**, **52d**, **52e** and **X**.

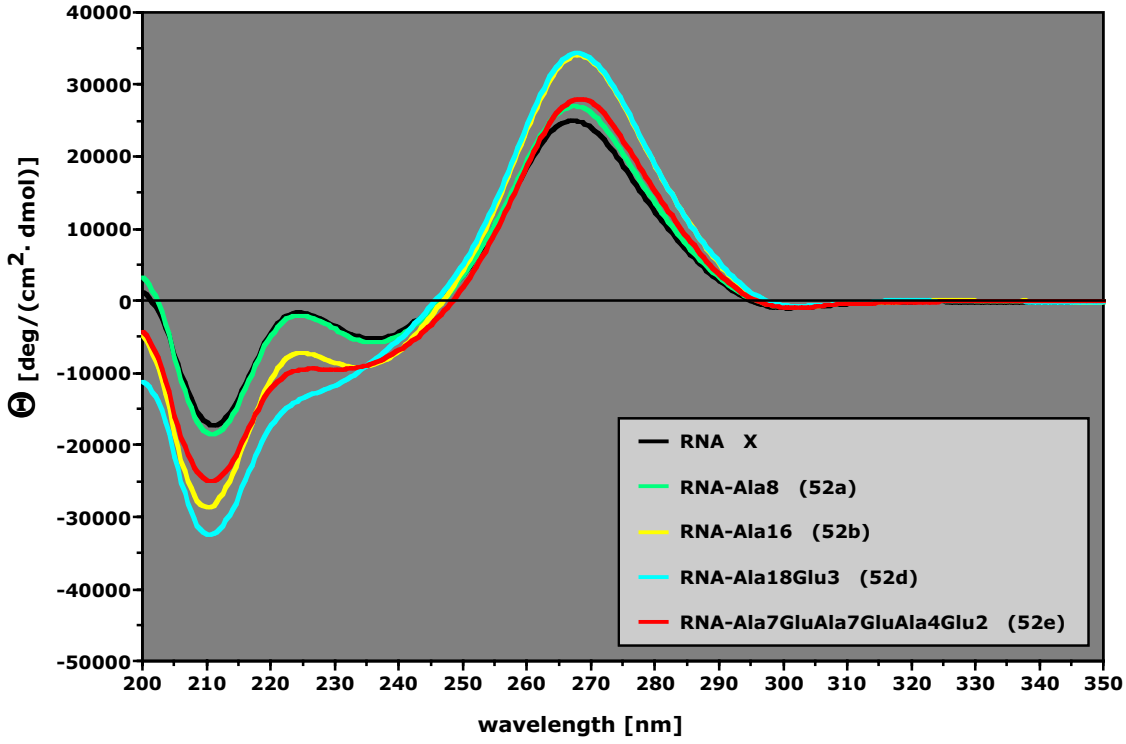


Fig. 51: CD spectra at 25 °C of **52a**, **52b**, **52d**, **52e** and **X**.

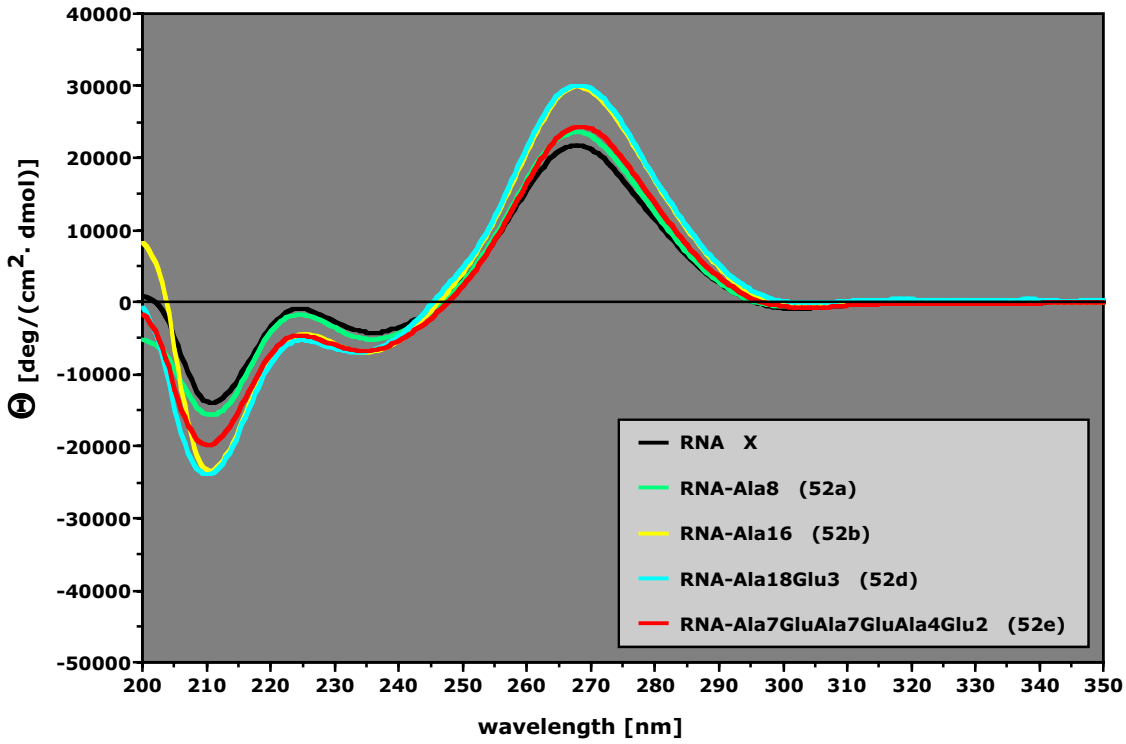


Fig. 52: CD spectra at 60 °C of **52a**, **52b**, **52d**, **52e** and **X**.

Analysis of the CD at 0 °C evidenced that the spectrum of the hairpin **X** shows the typical bands for an A-RNA and, in fact, the spectra of all the conjugates at 0 °C show a strong positive band at 267 nm owing only to the nucleotidic part. At this wavelength, the conjugates analysed have a stronger signal as compared to the hairpin alone. The differences span over +13% (**52a**) to 42% (**52d**). The higher Cotton Effect indicates that the single-stranded CCA-terminus of the hairpin, quite free to move when unpeptidylated, is rigidified by the presence of a peptide attached to it. Moreover, the different secondary structure of the peptide seems to correlate with the degree of rigidification of the single stranded region of the hairpin: the higher the calculated helicity, the stronger the effect. **52e**, in contrast, does not correlate with its predicted helicity (Table 5). We can hypothesise that the negative charges distributed along the peptide chain produce some electrostatic repulsion with the phosphate backbone reducing its 3'-terminal rigidity. **52b** and **52d** excert the same degree of hairpin rigidification although **52d** has a higher content of helicity. It might be that the hairpin rigidification has reached a maximun value so that the peptidic part cannot have a higher influence on it.

The peptidic part shows a Cotton Effect in the region 200-230 nm. The difference spectra between the unpeptidylated RNA and the conjugates at 0 °C show the change of the

Cotton Effect caused by the presence of the peptide and give evidence of the conformation adopted by the RNA-bound peptides in solution (Fig. 53).

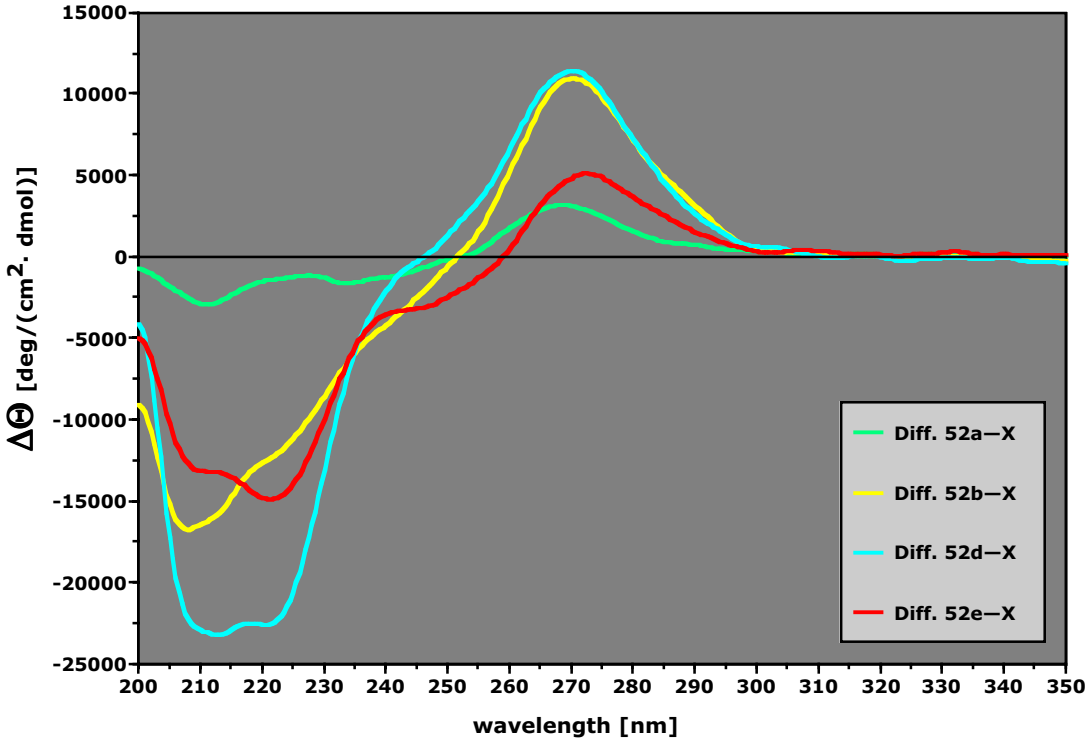


Fig. 53: Difference CD spectra between the conjugates and the hairpin at 0 °C normalised for the number of nucleotide residues.

The negative bands around 208 and 210 nm, characteristic of α-helical peptides, are present in the difference spectra of **52b**, **52d** and **52e** which, clearly, adopt a helical secondary structure. By contrast, **52a** which, having only eight alanines, should be predominantly random coil, does not show the two bands. Besides, this latter conjugate does not exert almost any effect on the hairpin Cotton Effect meaning that the peptide must likely adopt a stable secondary structure on its own in order to influence the hairpin structure. **52a** bears a 5'-phosphate group. CD analysis of the RNA hairpin with and without a 5'-phosphate do not show any difference excluding the phosphate being the cause for the different behaviour of **52a**.

The CD spectra are normalised for the number of nucleotide residues. They have not been normalised for the number of amino acid residues because the peptide length is not the same. One could thus argue that the stronger bands for the peptidic part of conjugates **52b**, **52d** and **52e** derive from the longer peptide chain. Normalising the difference spectra for the number of amino acid residues composing each peptide, always at 0 °C, gives the intrinsic amount of helicity of every amino acid residue in each conjugate (Fig. 54).

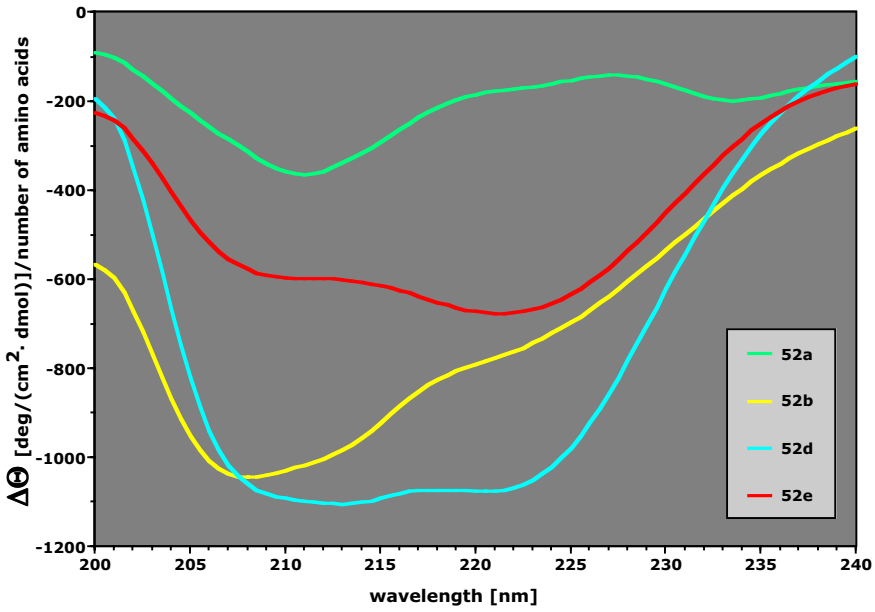


Fig. 54: Difference CD spectra normalised for the number of amino acids at 0 °C.

The analysis of the normalised spectra indicates clearly that the intrinsic content of helicity per residue is higher for **52b**, **52d** and **52e** as compared to **52a**, as expected. **52b** (containing only alanines) has a high content of helicity per residue even though this is the shortest peptide of the three, meaning that the presence of glutamate residues tends somehow to lower the propencity to form a helix. Figure 55 shows all the CD at all temperatures overlayed.

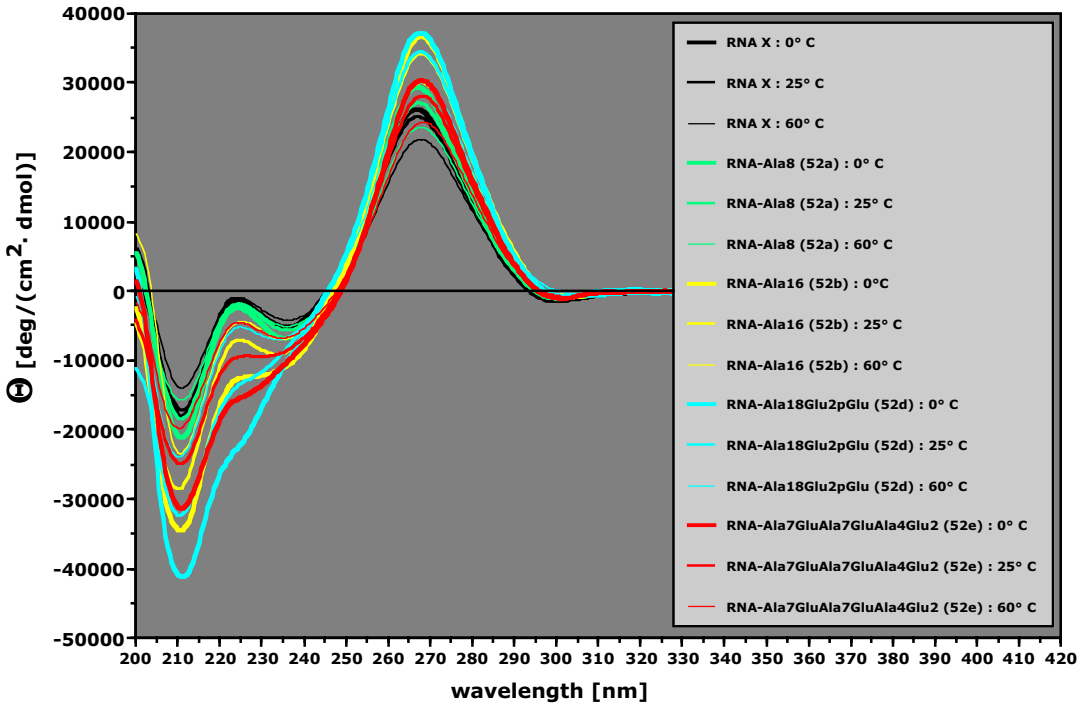


Fig. 55: All CD spectra at all temperatures overlayed.

The analysis of the spectra at the three temperatures gives evidence for the melting of the secondary structure elements of the conjugates. A better analysis can be achieved plotting the difference spectra of the conjugates with the hairpin at all temperatures (Fig. 56).

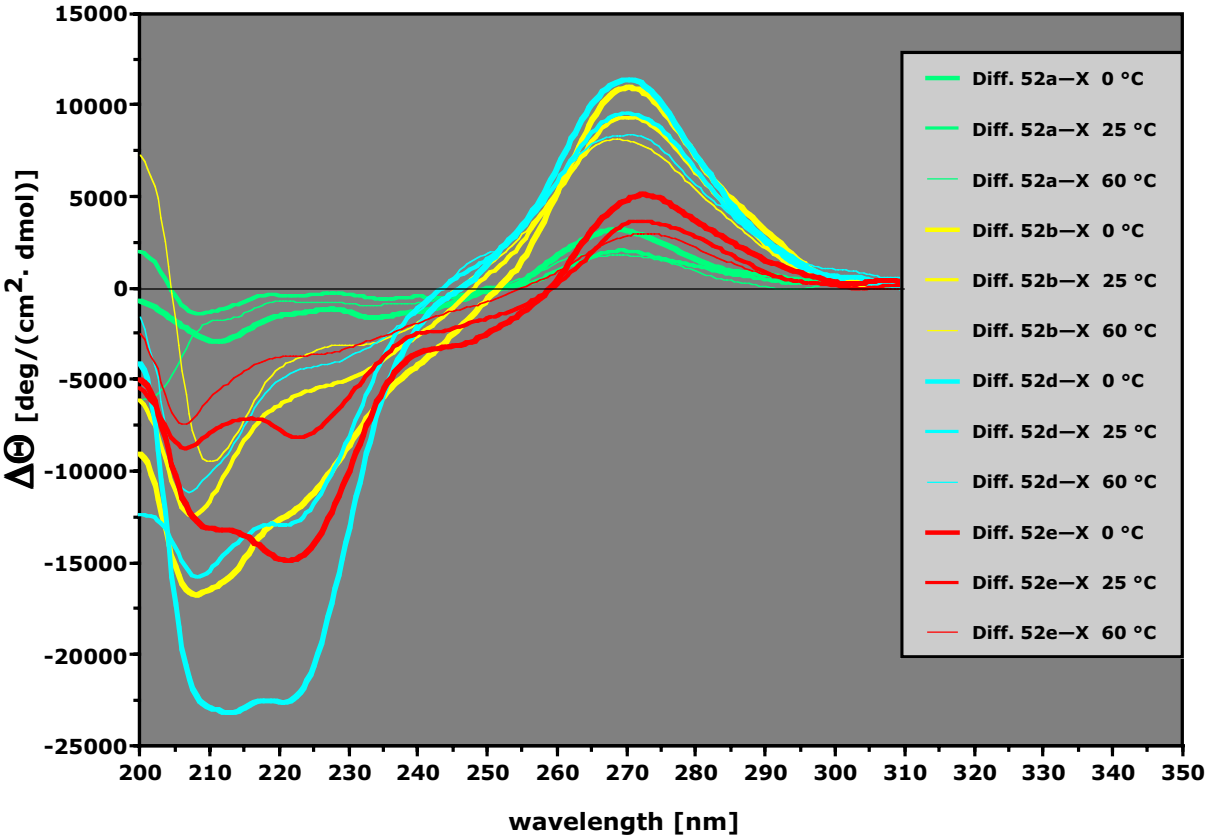


Fig. 56: All CD difference spectra at all temperatures overlayed.

It can be noted that the nucleobase region (245-295 nm) withstands heat denaturation more than the peptide region. Already at 25 °C, a major part of the helices seems to be melted. Glutamate seems to favour a rapid denaturation evidenced by the CD profiles of **52b** and **52e**. At 0 °C they have a similar content of α-helix whereas at 25 °C **52e** has a lower one. An important element that can be noticed is that the band at 222 nm denatures with temperature more readily than that at 208 nm.

This phenomenon has been analysed, recently, by D. S. Kemp.[123] They noticed that for polylanine or alanine-rich peptides $[\Theta]_{222}$ and $[\Theta]_{208}$ change with temperature in a disproportional way, in particular as the temperature is lowered, $[\Theta]_{222}$ becomes much stronger than $[\Theta]_{208}$. This can be visualised easily plotting the ratio of $[\Theta]_{222}/[\Theta]_{208}$ against temperature, which for most of the peptides is linear. They attribute this phenomenon, typical

of polyalanines, to the presence in solution of different conformers af α-helices (differing in the dihedral angles ϕ and φ) which have a temperature-dependent equilibration.

The oligoleucine conjugate **52g** was also analysed by CD even though its identity could not be confirmed by MALDI-ToF. The spectra show a strange phenomenon (Fig. 57); only RNA bands are visible as if the peptidic moiety was absent but a comparison with the spectra of **X** shows that **52g** melts much more rapidly. If the peptidic part is not present, why should the conjugate **52g** behave differently from **X**? This data are a further evidence that this conjugate behaves differently as compared to the others.

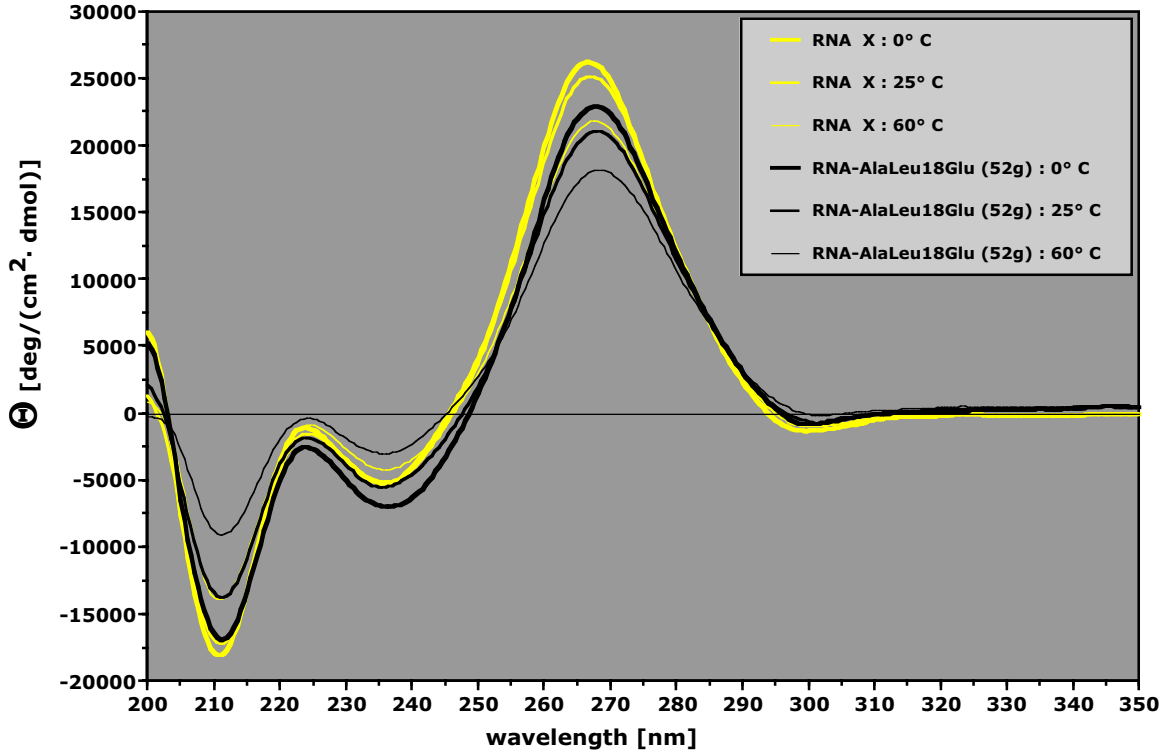


Fig. 57: CD spectra of **52g** and **X** at all the temperatures.

CD spectroscopy secondary structure analysis of the conjugates evidenced an influence of the peptidyl part on the overall structure of the molecule rendering it more rigid but, to exert this effect the peptide should have a rigid and stable secondary structure on its own. Moreover, from the spectra emerges that the peptidic moieties of conjugates **52b**, **52d** and **52e** adopt an α-helical conformation in solution.

2.5.2 Denaturation profiles and thermodynamic analysis

Thermal denaturation experiments were carried out to evaluate the influence of the peptide fragment conjugated at the 3'-end of the 22mer hairpin. The melting curves were further analysed to obtain thermodynamic parameters.

Thermal denaturation profiles of the following compounds were measured at 260 nm (only RNA denaturation observed): the parent RNA hairpin (**X**) and the conjugates **52a**, **52b**, **52c**, **52d**, **52e**, **52h**, all at pH 7.5 and 0.1 M ionic strength (NaCl).

The van't Hoff thermodynamic data can be abstracted from the denaturation profile by fitting the data to a unimolecular two-state equilibrium between A and A' (folded and denatured respectively). The equilibrium can be described as:

(2) $$K_{(A \Leftrightarrow A')} = \alpha/(1-\alpha) = e^{-(\Delta H° - T\Delta S°)/RT}$$

where α is the fraction of folded molecule A (α = 0-1). Solving the quadratic equation for α produces one solution between 0 and 1:

(3) $$\alpha(T) = 1/[1 + e^{(\Delta H° - T\Delta S°)/RT}]$$

For a UV denaturing profile with non-drifting baselines the upper (1) and lower (0) limits of α have to be adjusted to the low- and high-temperature baselines of $1/A_{260}$. This is achieved by fitting the function 4 to the datapoints.

(4) $$1/A_{260} = a + b \cdot \alpha(T)$$

Substitution of 3 in 4 gives the fitting formula:

(5) $$1/A_{260} = a + b/[1 + e^{(\Delta H° - T\Delta S°)/RT}]$$

The parameters to be optimised are *a*, *b*, $\Delta H°$ [cal/mol] and $\Delta S°$ [cal/(mol·K)]. From the values of $\Delta H°$ and $\Delta S°$, $\Delta G°$ and T_m are calculated:

(6) $$\Delta G° = \Delta H° - T\,\Delta S°$$

(7) $$T(\alpha = 0.5) = \Delta H° / \Delta S° = T_m$$

The 'optimised temperature-window fitting' method developed in our group, which takes into account only the main transition of the RNA hairpin above 85 °C, was used to abstract the thermodynamic data.[122b] This method is particularly useful when additional secondary minor transitions (a few percent of the main transition) at lower temperatures are present and produce 'tilted' baselines (this was the case for the conjugate profiles). In these cases the fitting is performed using the datapoints of two temperature windows, the first close to 20 °C and the second containing most of the main transition datapoints. Fitting the function over all datapoints, produces a fitting curve visibly deviating from the datapoints of the main transition (black curve in Figure 58).

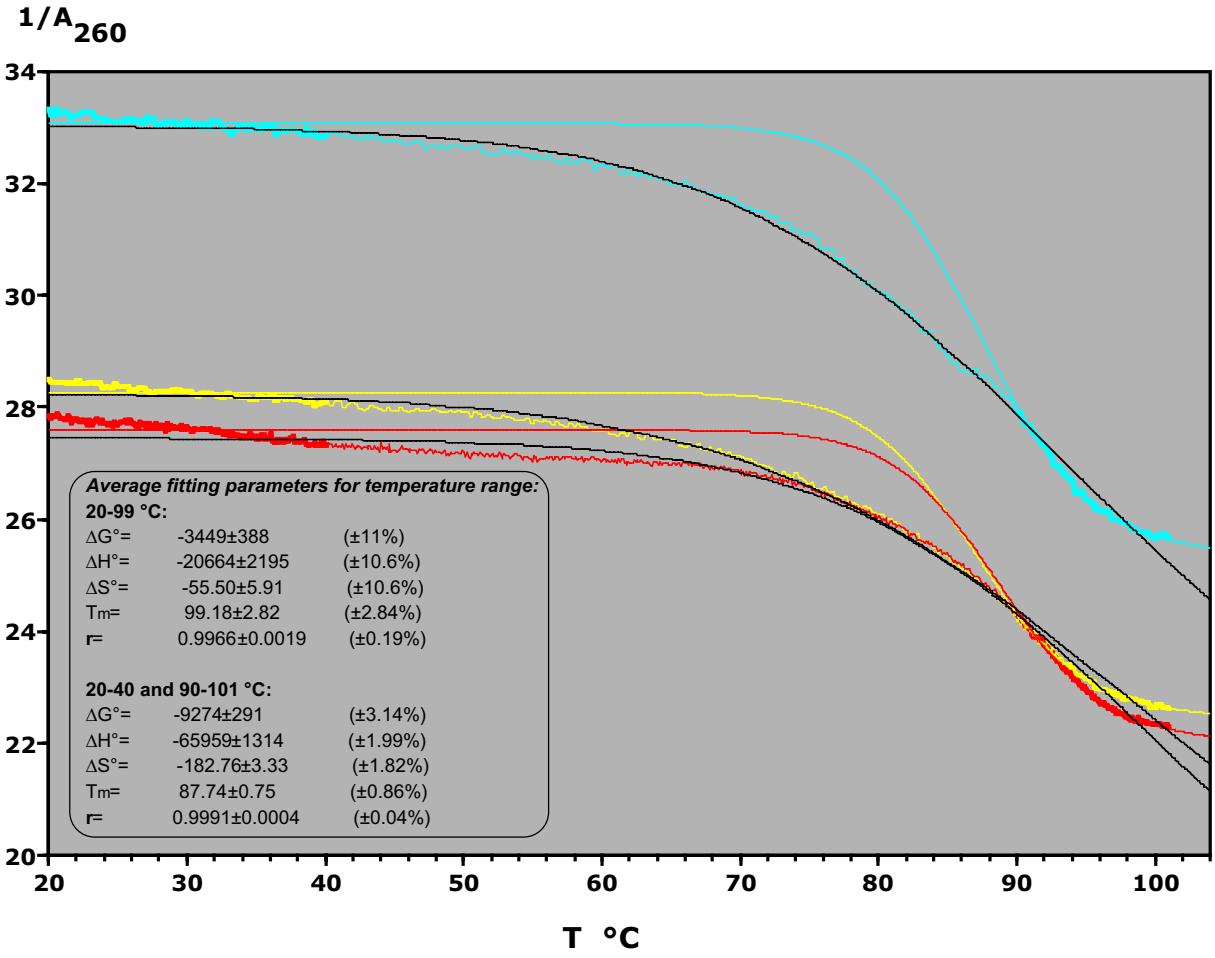


Fig. 58: Example of fitting procedure using all datapoints or two selected windows.

The thermodynamic parameters for the conjugates are listed in Table 6.

Table 6: Thermodynamic parameters of RNA hairpin formation. Buffer solutions measured at ~0.3 mM strand concentration (A_{260nm}≈0.04, however the profile of the RNA alone was shown to be concentration-indipendent.[122b]

Conjugate	T_m [°C]	ΔH° [kcal/mol]	ΔS° [kcal/mol]	$\Delta G°_{25°C}$ [kcal/mol]
(X)	87.6	–64.0 ± 2.7	–177.4 ± 7.8	–11.1 ± 0.8
52a	89.3	–69.2 ± 1.8	–190.9 ± 4.8	–12.3 ± 0.3
52b	87.7	–65.9 ± 1.3	–182.8 ± 3.3	–11.5 ± 0.3
52c	88.8	–56.0 ± 1.4	–154.8 ± 3.9	–9.9 ± 0.2
52d	88.3	–58.6 ± 1.9	–162.0 ± 5.4	–10.2 ± 0.3
52e	88.7	–56.3 ± 0.5	–155.6 ± 1.4	–9.9 ± 0.1
52h	88.7	–64.5 ± 0.7	–178.3 ± 1.8	–11.3 ± 0.1

Apart from these data, the melting temperatures of the following modified hairpins were determined: the RNA hairpin bearing a 5'-phosphate (T_m = 89.9 °C), a 3'-alanine (T_m = 87.9 °C) or both (T_m = 91.2 °C). These data indicate that a positive charge (3'-terminal α-ammonium) does not influence the stability of the hairpin (ΔT_m = +0.3 °C) whereas a 5'-terminal negative charge stabilises the structure somewhat (ΔT_m = +2.3 °C). This effect can also be noted from the melting temperature of **52a** which bears a 5'-phosphate and has a higher T_m compared to the other conjugates. When a 3'-positive and a 5'-negative charges are both present, a larger stabilisation is observed (ΔT_m = 3.6 °C), perhaps due to an interaction between them. The formation of a salt bridge between the two terminal charges was already hypothesised by Sprinzl on the basis of biochemical experiments.[124] He noticed that the aminoacylation renders the 3'-terminus inaccessible to binding complementary oligonucleotides. Based on this finding, he proposed that the α-amino group of the $tRNA^{Phe}$-A-C-C-A-3'-NH-Phe folds back to the 5'-terminal phosphate to form a salt bridge rendering the 3'-terminus inaccessible for binding.

Analysis of the conjugate T_m suggests that no major influence on the stability of the hairpin is observed in any conjugate (ΔT_{max} = 1.2 °C). However, the denaturation profiles do show differences dependent on the abscence or presence and, to some degree, identity of the peptidyl moiety (Fig. 59).

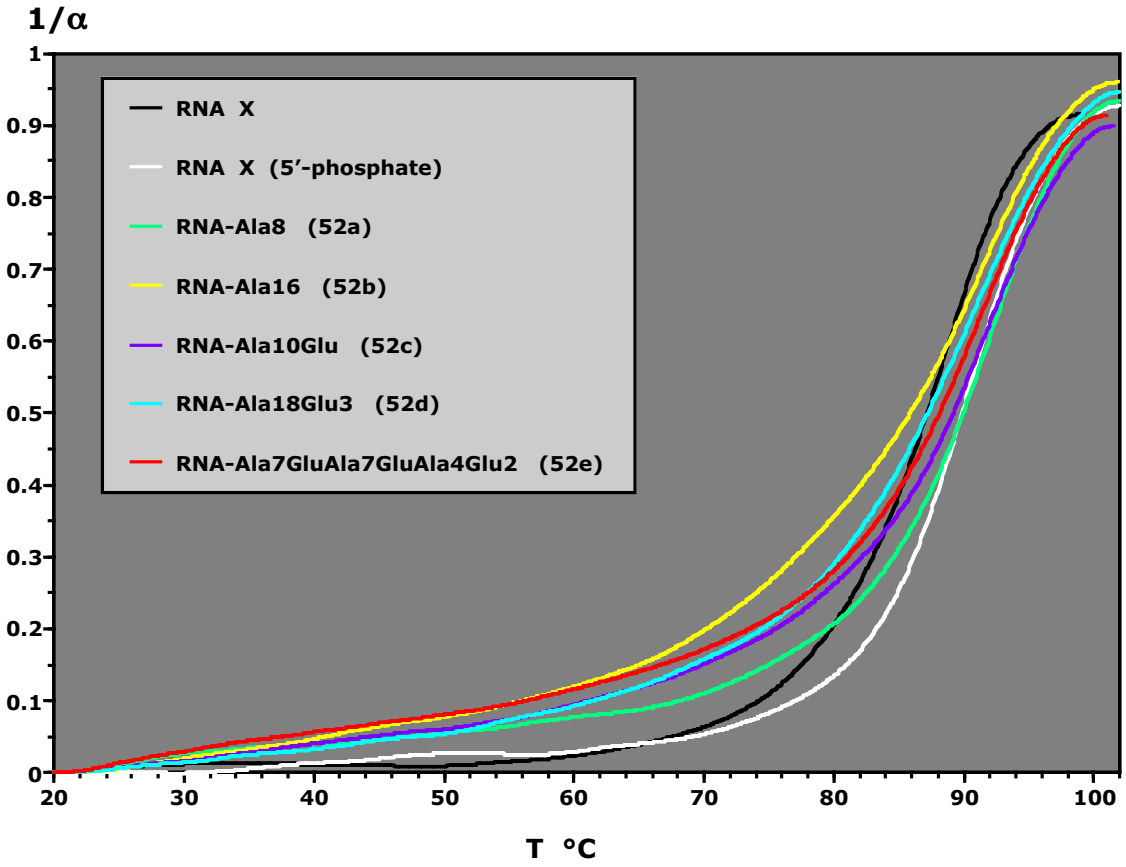


Fig 59: Normalised denaturation profiles at 260 nm.

The presence of the peptide renders the transitions somewhat shallower over the same temperature range, a strong indication for changes in the molecularity of the melting process.[125] The close to exclusively unimolecular equilibria are observed in the RNA-only profiles (including 5'-terminal phosphate). The tendency of aggregation of the conjugates, which translates into molecularities higher than one ($n > 1$), can be qualitatively recognised in the differential shallowness of the profiles at intermediate temperatures (below 87 °C), albeit without a visibly systematic relationship to the peptide sequence.

2.5.3 AFM analysis

The atomic force microscope (AFM) is one of the scanned-proximity probe microscopes. These types of microscope work by measuring a local property, such as height, optical absorption or magnetism, with a probe or tip placed very close to the sample. Unlike traditional microscopes, scanned-probe systems do not use lenses, so the size of the probe rather than diffraction limits their resolution. AFM measures the attractive or repulsive forces (at the atomic level) between a sharp probing tip (which is attached to a cantilever spring) and a sample surface.[126] In the 'contact mode' the tip touches the sample surface and hard-sphere repulsion forces are measured. As a raster-scan drags the tip over the sample, some sort of detection apparatus measures the vertical deflection of the cantilever, which indicates the local sample height. In the 'non-contact' mode (distance greater than 10 Å between the tip and the surface), the AFM derives topographic images from measuring attractive forces between the tip and the sample.[127] AFM can achieve a resolution of 10 pm and can image samples in air and under liquids.

In collaboration with Dr. Anthony Coleman and Patrick Shahgaldian, Institut de Biologie et Chimie des Protéines (IBCP), Lyon, non-contact AFM images were generated after deposition of an aqueous solution of some conjugates on a glass surface and mild drying. Some conjugates, in particular **52f** and **52e** were able to form supramolecular structures. **52f** forms polydisperse nanovesicles (Fig. 60).

These nanovesicles were treated with spermidine in order to evaluate an eventual interaction of this molecule with the RNA part. It emerged that the same nanovesicles, after treatment with spermidine, became smaller and higher as if this interaction had rigidified them (Fig. 61). The analysis of the dimension of the nanovesicles (Fig. 62), reveals that without spermidine the particles have an average height of ~13 nm and an average diameter of 204 nm whereas after treatment with spermidine the average height is 16.7 nm and the diameter 102 nm. The distribution of the diameter dimensions is broader before spermidine while the distribution of the height values becomes broader with spermidine. The plotting of the ratio height/diameter reveals that before spermidine treatment the ratio is much more uniform meaning that, without spermidine, the blobs have a quite uniform shape, albeit not size, while with spermidine they shrink, but not proportionally, so their shapes become much less uniform.

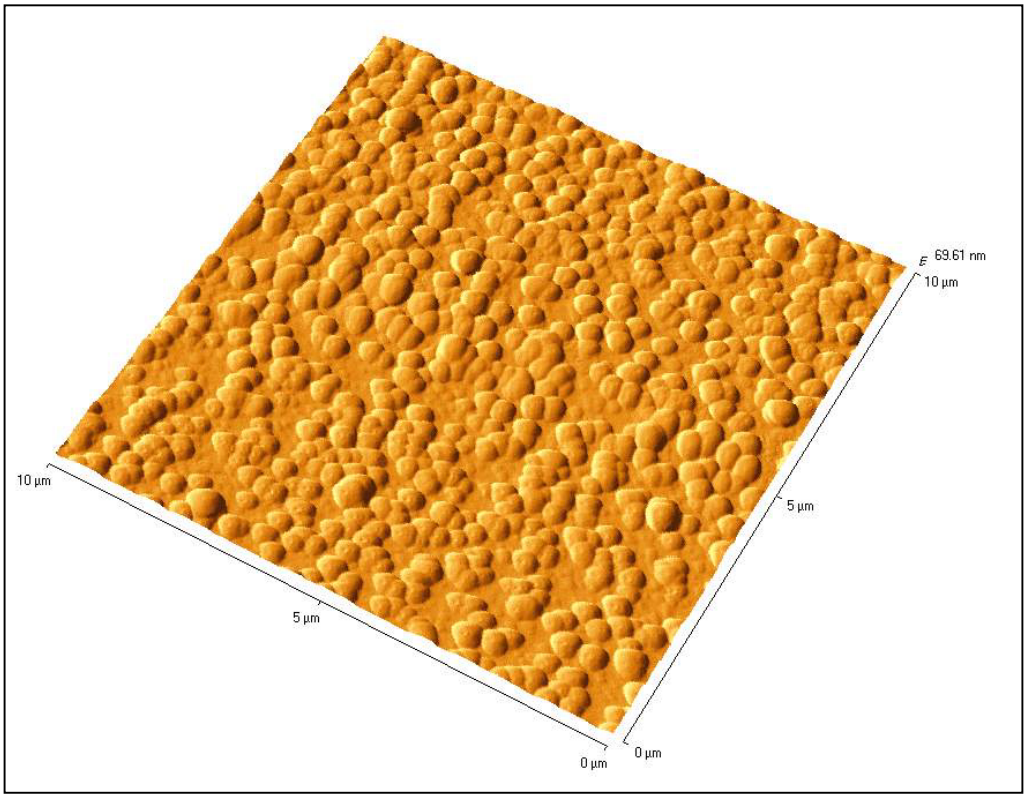


Fig. 60: AFM scan of conjugate **52f** on glass plate.

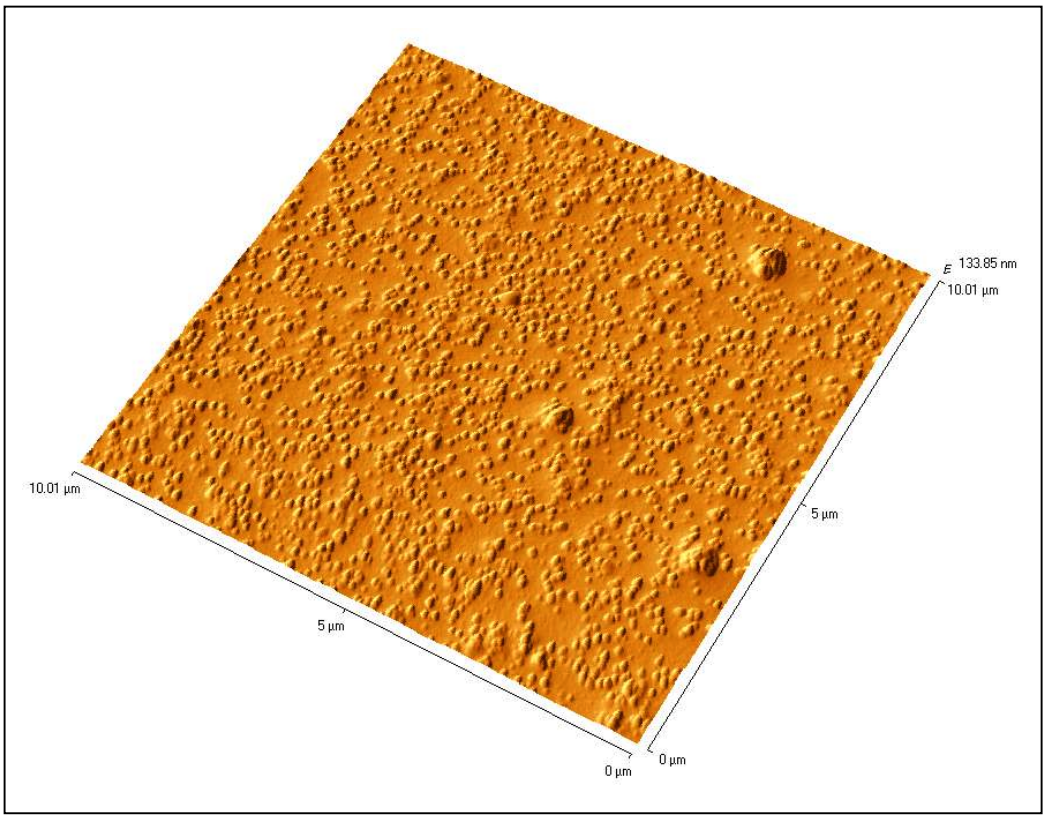


Fig. 61: AFM scan of **52f** after mixing with a spermidine solution.

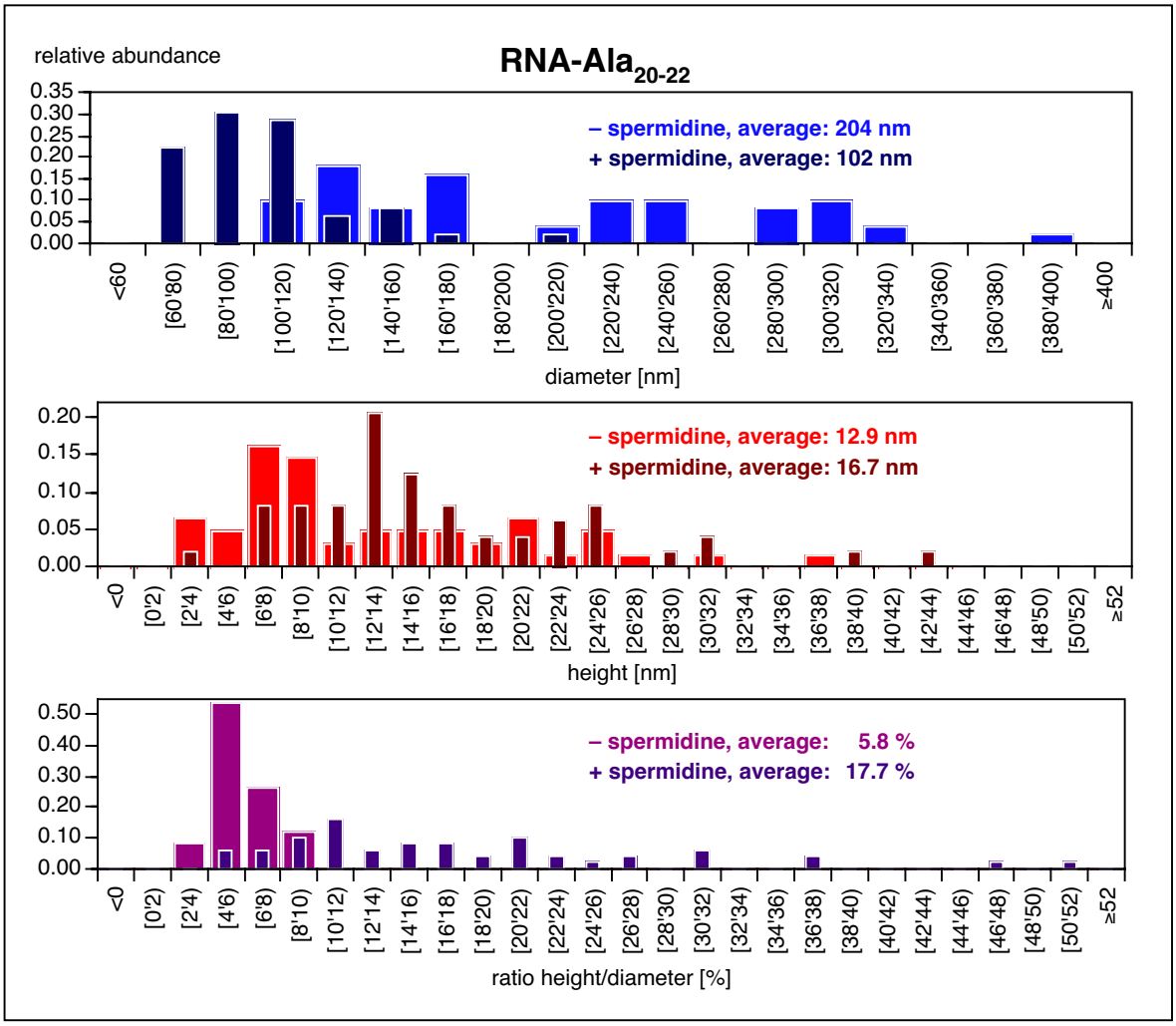


Fig. 62: Analysis of vesicle sizes distribution. Number of vesicles analysed: 49.

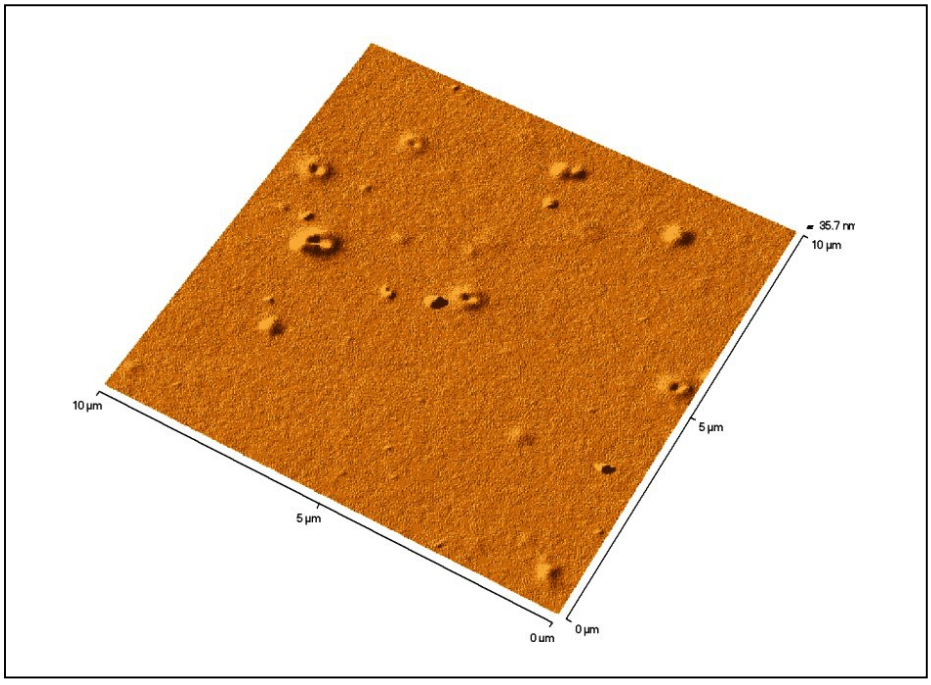


Fig. 63: AFM scan of **52e** on glass plate.

52e was able to form a kind of nanostructure with a shape resembling craters (Fig. 63) of 500 nm external diameter, 300 nm internal diameter and 5-6 nm height.

Not all conjugates were able to form nanostructures, **52g** and **52d** formed small crystals after deposition and drying on the glass surface (Fig. 64 and 65).

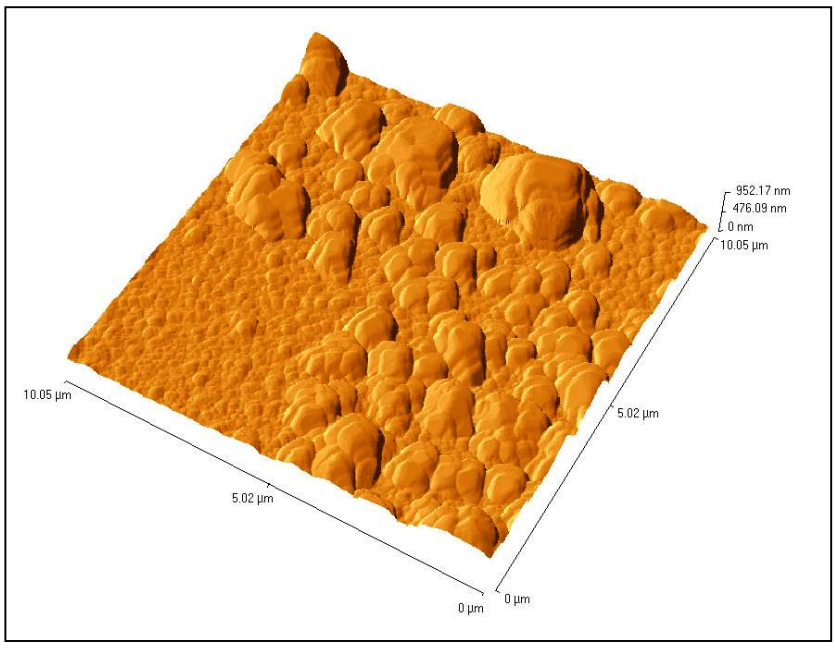


Fig. 64: AFM scan of **52g** on glass plate.

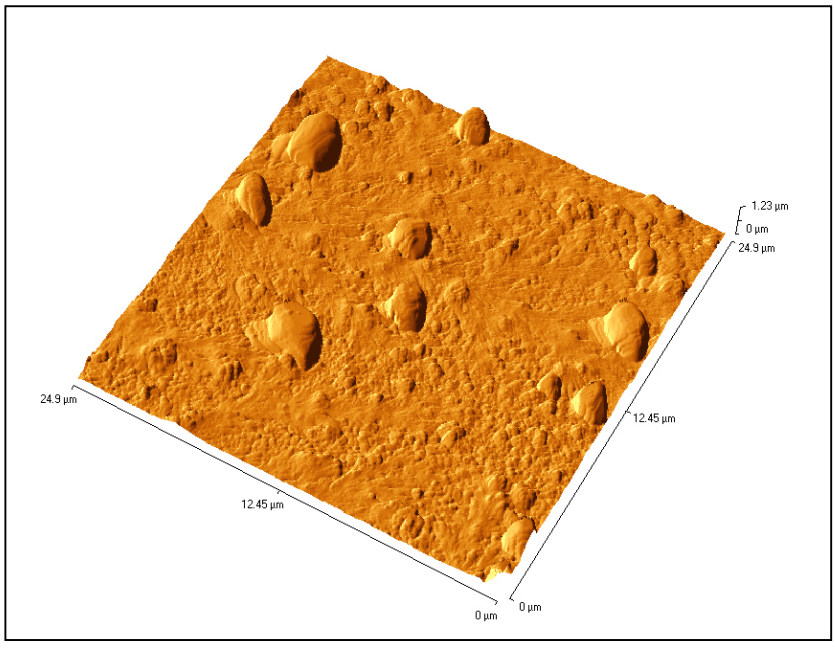


Fig. 65: AFM scan of **52d** on glass plate.

The conjugate **52b**, bearing a peptide of 16 alanines, was able to form some kind of nanostructure but less ordered and more polydisperse than **52f** (Fig. 66).

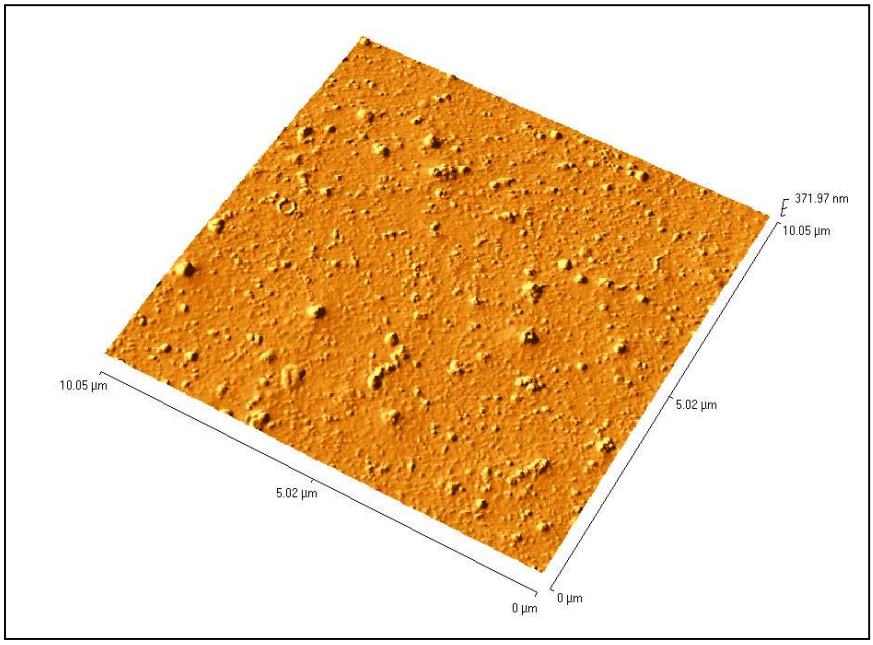


Fig. 66: AFM scan of **52b** on glass plate.

The RNA hairpin alone **X** was also analysed (as a control) and it does not seem to form any kind of supramolecular assembly. Instead, at rather high concentation, it forms microcrystals (Fig. 67).

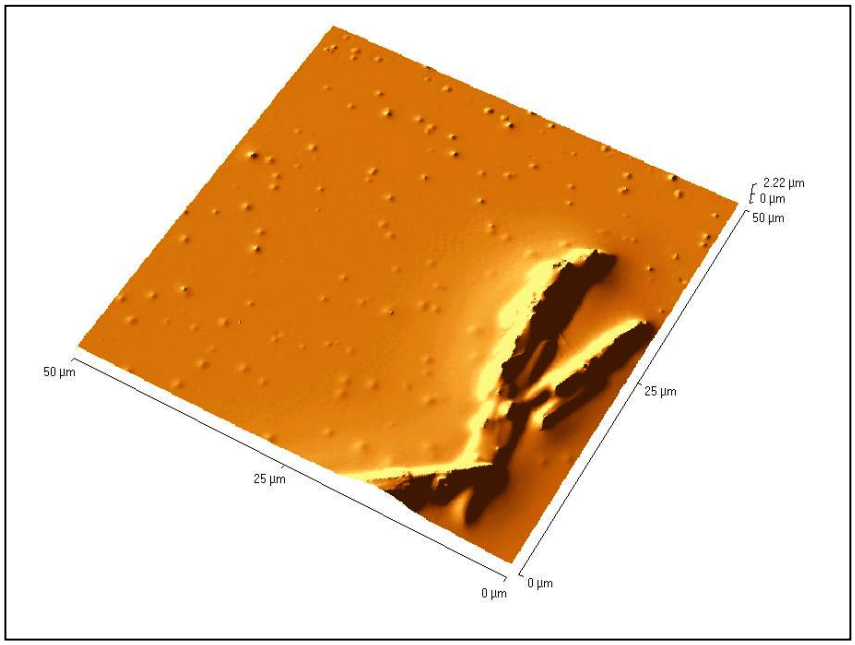


Fig. 67: AFM scan of **X** on glass plate.

2.5.4 DLS analysis

Preliminary dynamic light scattering (DLS) experiments confirmed the formation of supramolecular structures for **52f** - without spermidine- also in solution where the vesicles are even more polydisperse (60-600 nm, as compared to 100-400 nm diameter on the surface). The correlation analysis of the DLS measurement (Fig. 68) shows the presence of two families of aggregates, a larger and a smaller one (around 250 and 500 nm). This could arise from the inhomogeneity of **52f** which is formed prevalently from conjugates containing a peptide with 20-23 alanines and a second group bearing a peptide with 9-11 alanines (see MALDI-ToF of **52f**, Fig. 36).

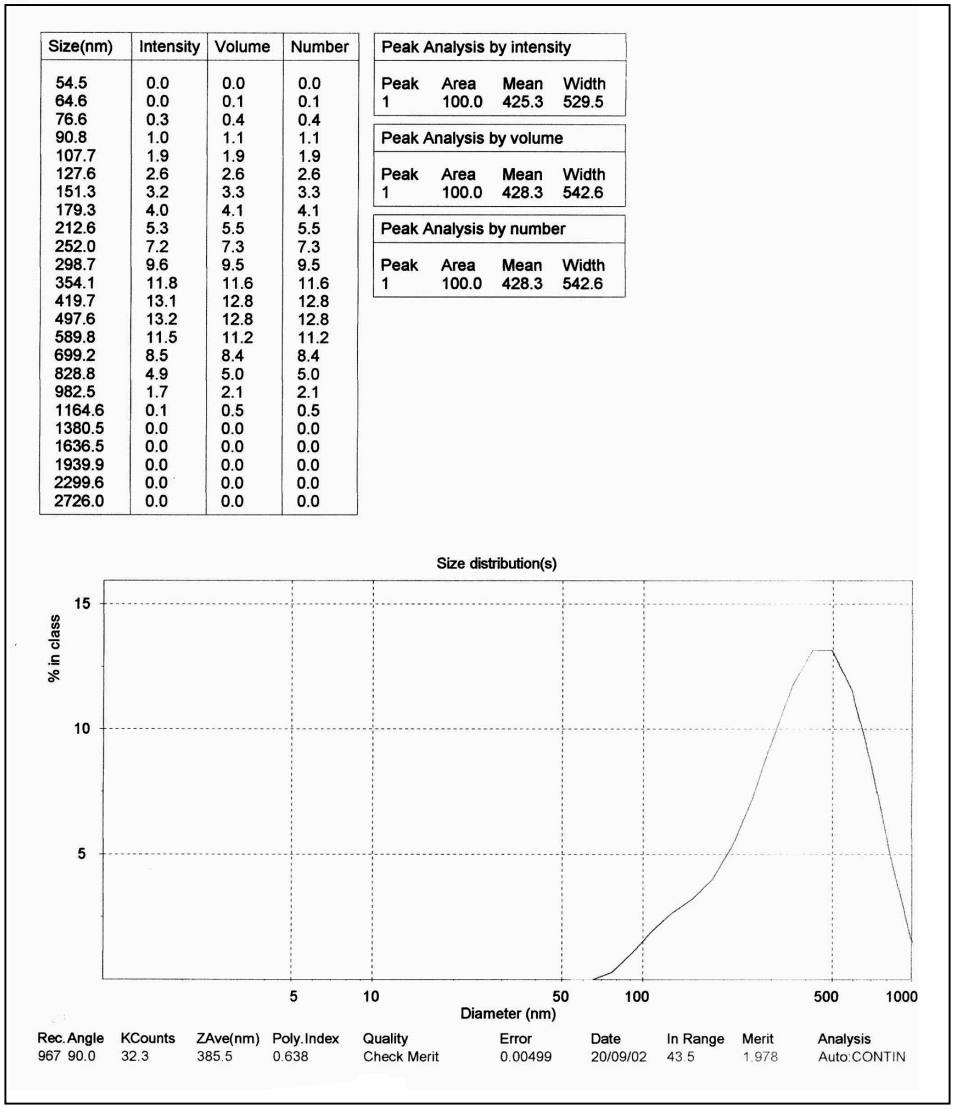


Fig. 68: DLS measurement of **52f**.

Self-assembling and compartmentalisation is a key process for the evolution of life. In this regard, our conjugates, in particular **52b**, **52f** and **52e** are examples of new constructs able to form higher-order structures and, at the same time, interesting molecules for further prebiotic studies. These are the first examples of self-asssemblig molecules formed by RNA and peptides, and these results are very important from a prebiotic prospective. In fact, proto peptidyl-tRNAs (similar to our constructs) have been hypothesised to be protagonist for the formation of the first cells.[14] The self-assembling of our conjugates could mimic a process leading to the formation of some sort of vesicles in the 'early cellular' world.

2.6 Summary and outlook

A new solid phase stepwise procedure for the synthesis of amphiphilic 3'-peptidyl-RNA conjugates was developed. Our constructs are the first examples of oligopeptides directly attached in a biomimetic (and spacer-free) way to the 3'-terminus of an oligoribonucleotide. They mimic the peptidyl-tRNA and could be used for cocrystallisation with the ribosome. The central molecule for the synthesis of our conjugates is the protected derivative **10** of 3'-amino-3'-adenosine. This molecule was immobilised on aminomethyl-derivatised polystyrene with a loading of ~10-11 μmol/g. Two different spacers, **30** (short and hydrophobic) one **31** (long and more hydrophilic), were used to attach the building block to the solid support in order to evaluate the influence of the spacer on the overall yield of the synthesis.

The syntheses of the conjugates were carried out using DEPBT as activating reagent, which after several synthesis trials, revealed to be the best activator compared to HATU and CBMIT (Fig. 20). The syntheses were quite problematic because of the 'overactivation' side reaction and aggregation problems. Overactivation, which leads to the incorporation of two amino acid residues instead of one, can be limited using a lower concentration of activated amino acid but, at the same time this causes a lower overall yield. A compromise should be found between a higher yield of material of poor quality or less material but more homogeneous.

Aggregation is probably the major problem for our syntheses owing to the high hydrophobicity of the peptidic moiety. As a matter of fact, even the incorporation of four negatively charged amino acids along the peptide chain could not avoid what we assume to be an aggregation problem. In the future, a still more hydrophilic peptide should be synthesised on our solid support to ascertain that the hydrophobic character of the peptide is the cause of aggregation and not the overall structure of the conjugates. The nature of the spacer did not prove important to improve the yield of neither the peptide nor the oligonucleotide synthesis. Overactivation and aggregation problems lowered considerably the final yield of purified monomeric conjugates which, nevertheless, could be recovered in sufficient amounts for structural characterisation.

The conjugates were analysed by CD, UV-detected thermal denaturation, AFM and one by DLS. CD spectroscopy revealed that the oligoalanine moieties adopt an α-helix secondary structure as we hoped for, apart from **52a** which, having only eight alanines, cannot form a stable helix in solution. A helical conformation is a prerequisite for hydrophobic peptides to span a lipid membrane, so this result is important in view of future experiments in

this direction. Moreover, the peptides influence the hairpin secondary structure rigidifying its 3'-single stranded part.

Thermal denaturation did not reveal a significant influence of the peptide on the stability of the hairpin. Probably, the hairpin being already very stable (T_m = 87.6 °C), cannot be influenced too much by the presence of a peptide attached to it. Even when the peptide bears four negative charges, which could interact in a repulsive way with the polyanionic phosphate backbone, the influence is minimal. However, the shallow denaturation pofiles show that the conjugate's higher-order structure departs from unimolecularity, a strong indication for the precence of aggregates in diluted solution.

Preliminary AFM analysis of the conjugates revealed that some of them are able to self-assemble and form nanovesicles. DLS measurements confirmed that also in solution there is a formation of nanovesicles, albeit more polydisperse than after deposition of the product on a glass plate and mild drying. This is very important, not only because our conjugates could suggest a new way for the formation of the first proto-cells, but also because they form a new class of molecules and could have unexpected properties. In this regard, the antibacterial agents based on a peptide architecture presented by Ghadiri and coworkers are a remarkable and challenging example.[128]

In the future, more systematic studies should be carried out in order to analyse the three-dimensional arrangement of the conjugates within the nanovesicles. In this regard, freeze-cut cryo electron transmission microscopy imaging will be performed to study the internal structure of the vesicles. Moreover, AFM imaging using hydrophilic or hydrophobic derivatised silicon wafers should clarify if the hydrophilic RNA strand is at the outer surface as we expect, since the vesicles were formed in water. The addition of proteases and RNAases during AFM imaging could also shed light on the arrangement of the conjugate molecules; in fact, rapid degradation should be observed with these enzymes depending on which part of the molecule is exposed. Finally, studying vesicle properties like their encapsulation capacity could be interesting.

3. EXPERIMENTAL PART

3.1 General

Chemicals were obtained from *Fluka Chemie AG* and *Sigma Aldrich* unless otherwise mentioned. BOC-sarcosine and FMOC-alanine were obtained from *Fluka Chemie AG*. FMOC-Arg(Pmc) and HBTU were obtained from *Novabiochem*, HATU from *ABI* and $Pd(PPh_3)_4$ from *Strem Chemicals*. The solid support, aminomethyl polystyrene 50% crosslinked with divinylbenzene was obtained from *ABI* (360865 C, Lot: 9609225 GB). TBDMS-monomers and the reagents for the oligoribonucleotide synthesis were obtained from *Glen Research*. TOM-monomers were a gift from Stephan Pitsch (EPFL). All the reagents were used as purchased. Solvents were dried under standard conditions and freshly distilled prior to use.

Column chromatography was performed with flash silica gel (35-70 μm) from *Uetikon* and with *Merck* 60 (40-63μm). The solvent ratios are given in % volume.

Thin Layer Chromatography (TLC) was performed on pre-coated silica gel F_{254} plates with fluorescent indicator from *Merck*. The detection of compounds was done with UV-light (254 nm). Nucleosides were visualised on TLC plates by subsequent spraying with conc. H_2SO_4 and 2% naphtoresorcin solutions in ethanol, followed by heating. For the detection of the amino groups, spraying with an ethanolic solution of ninhydrin, followed by heating was used.

NMR spectroscopy was performed on the instruments *Varian Gemini 300* (^{1}H-NMR 300 MHz), *Bruker VXR-400* (^{1}H-NMR 400 MHz) and *Bruker DRX-500* (^{1}H-NMR 500 MHz, ^{13}C-NMR 125 MHz, ^{31}P-NMR 202 MHz). Chemical shifts (δ) are reported in ppm relative to residual solvent peaks, and coupling constants *J* are in Hertz (Hz). NMR solvents were obtained from *Dr. Glaser AG* (Basel, Switzerland).

Melting points were determined on a *Kofler* block and are not corrected.

Ultra Violet-Visible (UV/Vis) spectra were recorded on a *Perkin-Elmer Lambda Bio 40* spectrophotometer equipped with a deuterium and a tungsten-halogen lamp.

Solid phase peptide syntheses were carried out manually using a Teflon syringe equipped with a polyethylene filter purchased from *MultiSynTech* (Germany).

Oligoribonucleotide syntheses were carried out on an *Applied Biosystems* DNA/RNA Synthesizer, model 392.

The buffer solutions were prepared with water purified through the Nanopure Ultrapure D4742 water system of *Barnstead*. The salts (biochemical quality) were obtained from *Fluka*.

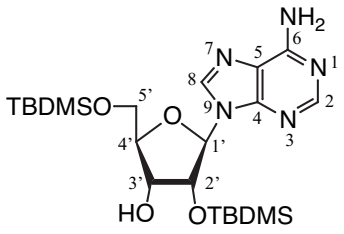
HPLC were performed using the following instruments: *Shimadzu LC-7A* HPLC system (high pressure gradient mixing, total flow rate max. 20 ml/min, 2 ml loop) coupled with a *Kontron-Instruments* UV spectrophotometer, model 432; the chromatograms were recorded on a *Shimadzu Integrator*, model C-R3A; *Waters 600* system equipped with a pump model 511 and a photodiode array detector 991; the chromatograms were processed using the software *Millennium* V. 2.01 from Waters; *Hewlett Packard Series 1090* system equipped with a diode array detector; the chromatograms were processed using the software *HP Chemstation Rev. A.06.03* from Hewlett Packard.

Tubes and tips for RNA handling are RNAase free and were purchased from *AMBION*.

MALDI-ToF mass spectra were measured on a *Applied Biosystems Voyager Elite* mass spectrometer for the analysis of the peptides which were analysed in positive ion mode using as a matrix α-cyano-4-hydroxycinnamic acid. The conjugates were analysed on a *Bruker Riflex III* mass spectrometer in negative ion mode using as a matrix 2,4,6-trihydroxy acetophenone.

3.2 Synthesis of the building block 10

3.2.1 2',5'-*Bis*-*O*-(*tert*-butyldimethylsilyl)-β-D-adenosine (12)



TBDMS-Cl (24.7 g, 163.9 mmol) was added to a solution of adenosine (14.6 g, 54.6 mmol) in anhydrous pyridine (110 ml). The mixture was stirred at r.t. for 48 h, then evaporated to dryness. The crude product was redissolved in CH_2Cl_2 and washed with 5% aq. HCl (v/v 3:1). The organic layer was then washed with sat. $NaHCO_3$-solution and brine, dried (Na_2SO_4) and evaporated. Purification by chromatography (EtOAc/hexane gradient from 70/30 to 100/0)

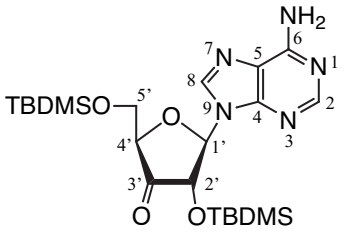
yielded **12** (16.2 g, 60%), **13** (9.4 g, 35%) and 2,3,5-*tris*-*O*-silylated adenosine (1.4 g, 5%) as white solids.

Isomerization of 13: To a suspension of **13** (4.5 g, 9 mmol) in MeOH (40 ml), triethylamine (1 ml, ~2.5% v/v) was added and the solution was stirred at r.t. for 3 h. The solution was evaporated to dryness and purification by chromatography (EtOAc/hexane gradient from 60/40 to 80/20) furnished more pure **12**. After some isomerisation cycles, the yield of **12** for the silylation reaction increased to 85%.

TLC R_f (EtOAc/hexane 7/3) = 0.25 (**12**) and 0.15 (**13**).
R_f (EtOAc 100%) = 0.40 (**12).**

^{1}H-NMR ($CDCl_3$) δ: –0.14, –0.08, 0.11, 0.12 (4*s*, 12H, Si-CH_3); 0.81, 0.93 (2*s*, 18H, Si-$C(CH_3)_3$); 2.81 (*d*, 1H, ^3J=4.2, OH(3')); 3.82 (*dd*, 1H, ^3J=2.5, ^2J=11.5, H_A(5')); 3.99 (*dd*, 1H, ^3J=2.7, ^2J=11.5, H_B(5')); 4.18 (*m*, 1H, H(4')); 4.26 (*m*, 1H, H(3')); 4.62 (*t*, 1H, ^3J=5.1, H(2')); 5.90 (*br. s*, 2H, NH_2(6)); 6.08 (*d*, 1H, ^3J=5.1, H(1')); 8.19 (*s*, 1H, H(8)); 8.32 (*s*, 1H, H(2)).

3.2.2 9-(2',5'-*Bis*-*O*-(*tert*-butyldimethylsilyl)-β-D-*erythro*-pentofuran-3-ulosyl)-9*H*-adenine (**14**)



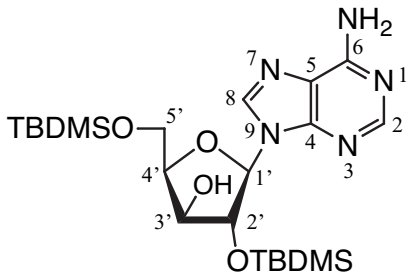
To a stirred suspension of chromium(VI) oxide (CrO_3, 2.42 g, 24 mmol) in CH_2Cl_2 (40 ml) at 0 °C was added, subsequently and slowly, acetic anhydride (2.3 mL, 24 mmol) and pyridine (3.9 ml, 48 mmol). The solution was stirred at 0 °C for 5-8 min. and then allowed to reach room temperature. Stirring was continued at r.t. until a black homogeneous solution was observed. Compound **12** (4 g, 8.07 mmol) and CH_2Cl_2 (12 ml) were added to the oxidation mixture. The solution was stirred at r.t. for 3 h. and then directly poured onto a chromatography column (800 g flash silica conditioned with CH_2Cl_2 in a 8 cm ∅ column). Isocratic elution with CH_2Cl_2 allowed the separation of the impurities (yellow band) from the target product. Only after the impurities have been eluted from the column, a step gradient

elution with EtOAc/hexane (from 50/50 to 100/0) afforded **14** (3.32 g, 83%) as a colourless powder.

TLC R_f (EtOAc/hexane 70/30) = 0.32; (EtOAc 100%) = 0.44.

^{1}H-NMR ($CDCl_3$) δ: –0.21, –0.02, 0.07, 0.10 (4*s*, 12H, Si-CH_3); 0.72, 0.91 (2*s*, 18H, Si-$C(CH_3)_3$); 3.97 (*s*, 2H, H_A(5'), H_B(5')); 4.30 (*s*, 1H, H(4')); 4.92 (*d*, 1H, ^3J=8.3, H(2')); 6.13 (*d*, 1H, ^3J=8.3, H(1')); 8.15 (*s*, 1H, H(8)); 8.33 (*s*, 1H, H(2)).

3.2.3 9-(2',5'-*Bis*-*O*-(*tert*-butyldimethylsilyl)-β-D-xylofuranosyl-9*H*-adenine (15)

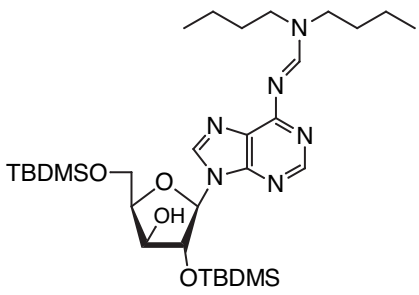


Sodium borohydride (1.5 g, 39 mmol) was slowly added to mechanically stirred glacial acetic acid (48 ml) at $T_{int.}$~13°C. After 1-1.5 h the solution became milky, indicating that the complex $HB(CH_3COO)_3$ was formed. Compound **14** (3 g, 6.07 mmol) was then added as a solid to the reaction mixture together with 10 ml of CH_3COOH. The solution was stirred for 2.5 days maintaining the $T_{ex.}$~8 °C which corresponds to a $T_{int.}$~12-13 °C (be careful not to freeze). Acetic acid was evaporated under reduced pressure and the yellow residue was taken up in an ice cold mixture of sat. $NaHCO_3$-solution/EtOAc (1:3) and extracted. The organic layer was washed with brine, dried (Na_2SO_4) and evaporated. Purification by chromatography (EtOAc/hexane 50/50) and evaporation furnished **15** (2.66 g, 88%) as a pale yellow foam.

TLC R_f (EtOAc/hexane 70/30) = 0.27; (EtOAc 100%) = 0.38.

^{1}H-NMR ($CDCl_3$) δ: –0.03, 0.06, 0.08 (3*s*, 12H, Si-CH_3); 0.89, 0.91 (2*s*, 18H, Si-$C(CH_3)_3$); 3.96-4.02 (*m*, 1H, H(3')); 4.08-4.16 (*m*, 2H, H_A(5'), H_B(5')); 4.25 (*m*, 1H, H(4')); 4.52 (*s*, 1H, H(2')); 5.76 (*d*, 1H, ^3J=1.3, H(1')); 5.98 (*br. s*, 2H, NH_2(6)); 6.69 (*d*, 1H, ^3J=9, OH(3')); 7.97 (*s*, 1H, H(8)); 8.33 (*s*, 1H, H(2)).

3.2.4 6-*N*-[(di-*n*-butylamino)methylene]-9-(2',5'-*bis*-*O*-(*tert*-butyldimethylsilyl)-β-D-xylofuranosyl)-9*H*-adenine (16)

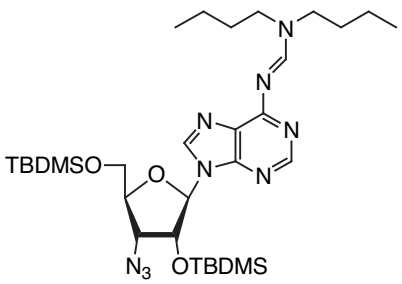


Compound **15** (3.2 g, 6.45 mmol) was coevaporated twice with absolute pyridine under reduced pressure and dissolved in dry MeOH (16 ml). After addition of *N*,*N*-di-*n*-butylformamide dimethylacetal (2.63 g, 13 mmol) the solution was stirred at r.t. for 2 h followed by evaporation. The column for the chromatography was first equilibrated with EtOAc to deactivate the silanol groups and then reequilibrated with a mixture of EtOAc/hexane 50/50. Rapid elution of the product furnished **16** (4.06 g, 97%) as a yellowish oil.

TLC R_f (EtOAc/hexane 50/50) = 0.55.

^{1}H-NMR ($CDCl_3$) δ: 0.01, 0.03, 0.06, 0.07 (*4s*, 12H, Si-CH_3); 0.88, 0.89 (*2s*, 18H, Si-$C(CH_3)_3$); 0.87-0.98 (*m*, 6H, $N(CH_2CH_2CH_2C\underline{H}_3)_2$); 1.25-1.45 (*m*, 4H, $N(CH_2CH_2C\underline{H}_2CH_3)_2$); 1.60-1.75 (*m*, 4H, $N(CH_2C\underline{H}_2CH_2CH_3)_2$); 3.40 (*dd= 't'*, 2H, ^3J=7.2, $N(C\underline{H}_2CH_2CH_2CH_3)$); 3.73 (*m*, 2H, $N(C\underline{H}_2CH_2CH_2CH_3)$); 3.96 (*m*, 1H, H(3')); 4.08-4.15 (*m*, 2H, H_A(5'), H_B(5')); 4.21-4.26 (*m*, 1H, H(4')); 4.52 (*d*, 1H, ^3J=1.4, H(2')); 5.74 (*d*, 1H, ^3J=1.4, H(1')); 6.97 (*br. s*, 1H, OH (3')); 7.99 (*s*, 1H, H(8)); 8.51 (*s*, 1H, H(2)); 9.00(*s*, 1H, N^6=CH).

3.2.5 3'-Azido-6-*N*-[(di-*n*-butylamino)methylene]-2',5'-*bis*-*O*-(*tert*-butyldimethylsilyl)-3'-deoxy-β-D-adenosine (18)

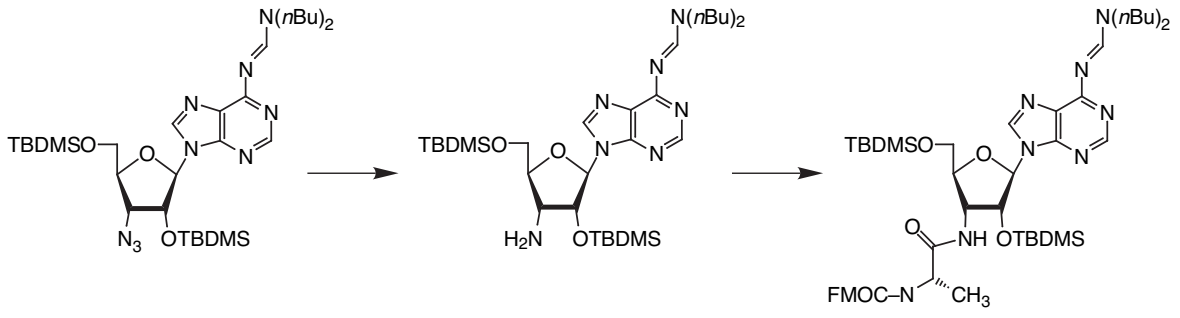


Compound **16** (4.06 g, 6.4 mmol) and DMAP (1.95 g, 16 mmol) were dissolved in dry CH_2Cl_2 (38 ml). The solution was cooled to 0°C and trifloromethanesulfonylchloride (TfCl, 1.3 ml, 12.8 mmol) was added dropwise. After stirring at 0°C for 1.5 h, the yellow solution was taken up in AcOH/H_2O(1:99)/CH_2Cl_2 and extracted twice. The organic layers were washed with ice cold sat. $NaHCO_3$-solution and brine, dried (Na_2SO_4), evaporated and dried under 'high vacuum' (oil pump) resulting in a yellow oil. The residue was dissolved in DMF (25 ml) and LiN_3 (1.25 g, 25.6 mmol) was added under argon and let it react overnight at r.t. The solution was then taken up in sat. $NaHCO_3$-solution/EtOAc. After extraction the organic layers were collected, washed with brine, dried (Na_2SO_4) and evaporated. The residue was purified (EtOAc/hexane 50/50) to give **18** (3.2 g, 76%) as a pale yellow oil.

TLC R_f (EtOAc 100%) = 0.72; (EtOAc/hexane 80/20) = 0.65.

^{1}H-NMR ($CDCl_3$) δ: –0.09, 0.05, 0.13, 0.14 (*4s*, 12H, Si-CH_3); 0.86 (*s*, 9H, Si-C(CH_3)$_3$); 0.91-0.98 (*m*, 15H, Si-C(CH_3)$_3$, N($CH_2CH_2CH_2C\underline{H}_3$)$_2$); 1.32-1.46 (*m*, 4H, N($CH_2CH_2C\underline{H}_2CH_3$)); 1.59-1.71 (*m*, 4H, N($CH_2C\underline{H}_2CH_2CH_3$)$_2$); 3.39 (*t*, 2H, ^3J=7.3, N($C\underline{H}_2CH_2CH_2CH_3$)); 3.72 (*t*, 2H, ^3J=7.3, N($C\underline{H}_2CH_2CH_2CH_3$)); 3.83 (*dd*, 1H, ^3J=3.0, ^2J=11.5, H_A(5')); 4.05 (*dd*, 1H, ^3J=3.3, ^2J=11.5, H_B(5')); 4.11 (*m*, 1H, H(3')); 4.21 (*m*, 1H, H(4')); 4.92 (*t*, 1H, ^3J=4.4, H(2')); 6.05 (*d*, 1H, ^3J=4.4, H(1')); 8.17 (*s*, 1H, H(8)); 8.53 (*s*, 1H, H(2)); 8.98 (*s*, 1H, N^6=CH).

3.2.6 6-*N*-[(di-*n*-butylamino)methylene]-2',5'-*bis*-*O*-(*tert*-butyldimethylsilyl)-3'-[*N*-(9-fluorenyl)methoxycarbonyl-L-alanylamino]-3'-deoxy-β-D-adenosine (21)



To a solution of **18** (151 mg, 0.23 mmol) in dioxane (5 ml, filtered through Al_2O_3 to remove traces of peroxide) was added PPh_3 (78 mg, 0.3 mmol) and the solution was stirred overnight at r.t. The reaction was quenched with H_2O (1 ml) under stirring for 1h at r.t. The solution was taken up in sat. $NaHCO_3$-solution/EtOAc and extracted twice. The organic layers were

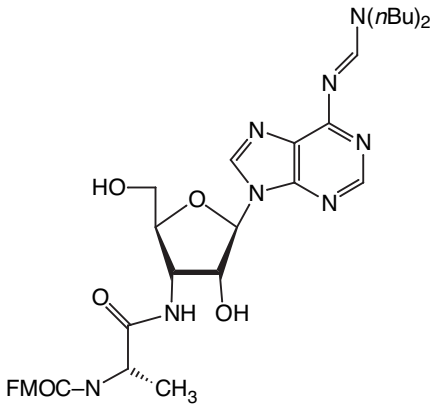
combined, washed with brine, dried (Na_2SO_4) and concentrated to 2-3 ml volume. Immediately after the purification of the amino compound, the FMOC-amino acid was activated and coupled to the amino nucleoside.

To a solution of carbonyldiimidazole (148 mg, 0.915 mmol) in dry CH_3NO_2 (1.0 ml) at 10°C methyl triflate (200 µl, 1.83 mmol) was added dropwise. After the addition of methyl triflate, the solution was allowed to warm to r.t. The solution was then added to a suspension of *N*-(9-fluorenyl)methoxycarbonyl-L-alanine (285 mg, 0.915 mmol) and *N*-methylimidazole (7 µl, 0.091 mmol) in dry CH_3NO_2 (0.7 mL) resulting in a clear solution that was stirred at r.t. until a precipitate formed (activated FMOC-amino acid). At this point, **20** was taken to dryness, redissolved in freshly distilled THF (2.5 ml) and added to the activated FMOC-amino acid solution. After stirring at r.t. for 1 h. the solution was taken up in a CH_2Cl_2/H_2O and extracted three times. The organic layers were combined, washed with sat. $NaHCO_3$-solution and H_2O, dried (Na_2SO_4) and evaporated. After flash chromatography (isocratic elution with EtOAc 100%) compound **21** was obtained as a colourless oil (162 mg, 76.4% over two steps).

TLC R_f (EtOAc/hexane 80/20) = 0.57; (EtOAc 100%) = 0.65 (**21**)
R_f (EtOAc/hexane 80/20) = 0.28; (EtOAc 100%) = 0.35 (**20**)

^{1}H-NMR ($CDCl_3$) δ: 0.04-0.14 (4*s*, 12H, Si-CH_3); 0.83 (*s*, 9H, Si-C(CH_3)$_3$,); 0.90-0.98 (*m*, 6H, N($CH_2CH_2CH_2C\underline{H}_3$)$_2$); 0.96 (*s*, 9H, Si-C($CH_3$)$_3$); 1.36-1.44 (*m*, 7H, N($CH_2CH_2C\underline{H}_2CH_3$)$_2$, CH_3(β)-Ala); 1.61-1.67(*m*, 4H, N($CH_2C\underline{H}_2CH_2CH_3$)$_2$); 3.37-3.67 (*m*, 2H, N($C\underline{H}_2CH_2CH_2CH_3$)); 3.72 (*m*, 2H, N($C\underline{H}_2CH_2CH_2CH_3$)); 3.91 (*dd*, 1H, ^3J=2.9, ^2J=11.2, H_A(5')); 3.99 (*dd*, 1H, ^3J=3.7, ^2J=11.2, H_B(5')); 4.17-4.26 (*m*, 3H, H(4'), H(9'')-Fmoc, H(α)-Ala); 4.41 (*br. t*, 2H, (O-CH_2)-Fmoc); 4.56 (*m*, 1H, H(3')); 4.66 (*dd*, 1H, ^3J=4.3, ^3J=6.0, H(2')); 5.34 (*br. d*, 1H, NH-carbamate); 6.10 (*d*, 1H, ^3J=4.3, H(1')); 6.52 (*br. s*, 1H, NH-amide); 7.30-7.77 (*m*, 8H, arom. H-Fmoc); 8.30 (*s*, 1H, H(8)); 8.53 (*s*, 1H, H(2)); 9.00 (*s*, 1H, N^6=CH).

3.2.7 6-*N*-[(di-*n*-butylamino)methylene]-3'-[*N*-(9-fluorenyl)methoxycarbonyl-L-alanylamino]-3'-deoxy-β-D-adenosine (22)

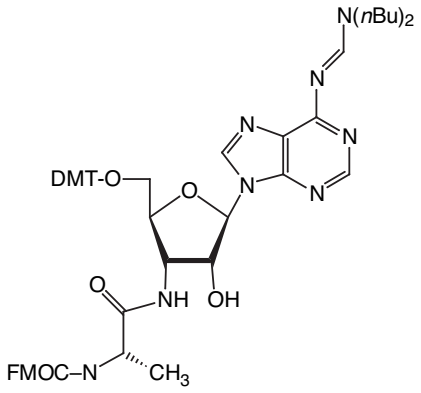


To a solution of **21** (308 mg, 0.332 mmol) in THF (3.6 ml splitted in three Eppendorf tubes), pyridine-hydrogen fluoride (~70 % HF) was added (215 µl, 2.39 mmol). The solution was stirred overnight at r.t. and then taken up in sat. $NaHCO_3$-solution/EtOAc and extracted three times. The organic layers were combined, washed with brine, dried (Na_2SO_4) and evaporated. Purification by chromatography (step gradient from EtOAc to EtOAc/MeOH 9:1) gave **22** (232 mg, 94%) as a white solid.

TLC Rf (EtOAc/MeOH 90/10) = 0.47.

^{1}H NMR ($CDCl_3$) δ: 0.91 (*t*, 3H, ^3J=7.2, ^3J=7.5, $N(CH_2CH_2CH_2C\underline{H}_3)$); 0.98 (*t*, 3H, ^3J=7.2, ^3J=7.5, $N(CH_2CH_2CH_2C\underline{H}_3)$); 1.23-1.42 (*m*, 4H, $N(CH_2CH_2C\underline{H}_2CH_3)_2$); 1.46 (*d*, 3H, ^3J=6.9, CH_3(β)-Ala); 1.59-1.71 (*m*, 4H, $N(CH_2C\underline{H}_2CH_2CH_3)_2$); 3.32-3.46 (*m*, 3H, $N(C\underline{H}_2CH_2CH_2CH_3)_2$); 3.81-3.99 (*m*, 3H, H_A(5'), H_B(5'), $N(C\underline{H}_2CH_2CH_2CH_3)$); 4.25 (*br. t*, 1H, H(9'')-Fmoc); 4.33-4.48 (*m*, 5H, H(3'), H(4'), H(α)-Ala, (O-CH_2)-Fmoc); 5.30 (*br. t*, 1H, H(2')); 5.51 (*br. d*, 1H, NH-carbamate); 5.66 (*d*, 1H, ^3J=5.4, H(1')); 7.18 (*br. d*, 1H, NH-amide); 7.32-7.76 (*m*, 8H, arom. H-Fmoc); 7.55 (*s*, 1H, H(8)); 8.16 (*s*, 1H, H(2)); 8.77 (*s*, 1H, N^6=CH).

3.2.8 6-*N*-[(di-*n*-butylamino)methylene]-3'-[*N*-(9-fluorenyl)methoxycarbonyl-L-alanylamino]-5'-*O*-(4,4'-dimethoxytrityl)-3'-deoxy-β-D-adenosine (7)

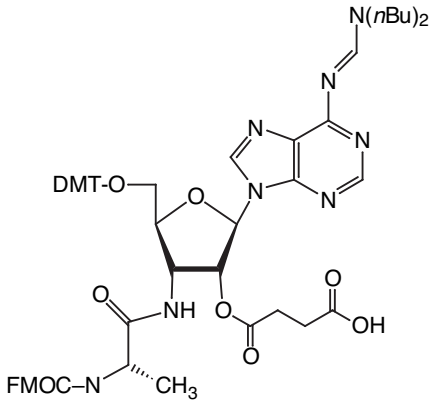


Compound **22** (40 mg, 0.057 mmol) was coevaporated three times with dry pyridine, then redissolved in pyridine (0.6 ml) under argon atmosphere. DMTCl (39 mg, 0.1144 mmol) and tetrabutylammonium nitrate (5 mg, 0.017 mmol) were added to the solution and stirred for 1 hour at r.t. MeOH (0.6 ml) was added to quench the reaction. The solution was taken up in sat. $NaHCO_3$-solution/CH_2Cl_2 and extracted twice. The organic layers were combined, washed with brine, dried (Na_2SO_4) and evaporated. After column chromatography (gradient elution from CH_2Cl_2 100% to CH_2Cl_2/MeOH 95/5, 0.1% pyridine) **7** was obtained (57 mg, 70%) as a yellowish foam.

TLC R_f (CH_2Cl_2/MeOH 93/7) = 0.50; (CH_2Cl_2/MeOH 97/3) = 0.16.

^{1}H-NMR ($CDCl_3$) δ: 0.94, 0.95 (2*t*, 6H, ^3J=7.4, $N(CH_2CH_2CH_2C\underline{H}_3)_2$); 1.33-1.41 (m, 7H, $N(CH_2CH_2C\underline{H}_2CH_3)_2$, ^3J=7.6, CH_3(β)-Ala); 1.61-1.69 (*m*, 4H, $N(CH_2C\underline{H}_2CH_2CH_3)_2$); 3.35-3.44 (*m*, 3H, $N(C\underline{H}_2CH_2CH_2CH_3)$, H_A(5')); 3.45-3.47 (*dd*, 1H, ^2J=10.7, H_B(5')); 3.65-3.77 (*m*, 2H, $N(C\underline{H}_2CH_2CH_2CH_3)$); 3.75 (2*s*, 6H, OCH_3); 4.19-4.24 (*m*, 2H, H(α)-Ala, H(9'')-Fmoc); 4.33-4.42 (*m*, 3H, H(4'), (O-CH_2)-Fmoc); 4.60 (*br. q*, 1H, ^3J=5.8, H(3')); 4.79 (*dd*, 1H, ^3J=5.8, ^3J=4.1, H(2')); 5.58 (*br. d*, 1H ^3J=7.4, NH-carbamate); 6.02 (*d*, 1H, ^3J=4.0, H(1')); 6.74-6.77 (*m*, 4H, Ph); 6.92 (*br. d*, 1H, ^3J=4.5 NH-amide); 7.15-7.74 (*m*, 17H, arom. Fmoc, Ph); 8.17 (*s*, 1H, H(2)); 8.46 (*s*, 1H, H(8)); 9.04 (*s*, 1H, N^6=CH).

3.2.9 6-*N*-[(di-*n*-butylamino)methylene]-3'-[*N*-(9-fluorenyl)methoxycarbonyl-L-alanylamino]-5'-*O*-(4,4'-dimethoxytrityl)-3'-deoxy-β-D-adenosine-2'-succinate (10)



To a solution of **7** (39 mg, 0.039 mmol) in dry 1,2-dichloroethane (350 μl) in an Eppendorf tube, DMAP (2.4 mg, 0.019 mmol), succinic anhydride (5.8 mg, 0.058 mg) and $(iPr)_2NEt$ (6.7 ml, 0.039 mmol) were added. The mixture was stirred at 50 °C in a thermoshaker and the reaction was complete within 20 min. The solution was cooled, diluted with CH_2Cl_2 (5 ml), washed with a cold aqueous solution of 10% citric acid and water. The organic layer was dried ($NaSO_4$) and concentrated (~1 ml). This concentrated solution was precipitated from a mixture of EtOAc/hexane 30/70 (~10 ml) giving **10** (36 mg, 84%) as a white solid.

TLC R_f (CH_2Cl_2/MeOH 95/5) = 0.25.

^{1}H-NMR ($CDCl_3$) δ: 0.94, 0.95 (2*t*, 6H, ^3J=7.3, $N(CH_2CH_2CH_2C\underline{H}_3)_2$); 1.15 (d, 3H, ^3J=7.6, CH_3(β)-Ala); 1.32-1.43 (m, 4H, $N(CH_2CH_2C\underline{H}_2CH_3)_2$,); 1.61-1.70 (*m*, 4H, $N(CH_2C\underline{H}_2CH_2CH_3)_2$); 2.54-2-85 (m, 4H, -CH2CH2-); 3.36-3.40 (*m*, 4H, $N(C\underline{H}_2CH_2CH_2CH_3)$, H_2(5')); 3.68-3.70 (*m*, 2H, $N(C\underline{H}_2CH_2CH_2CH_3)$); 3.76 (*s*, 6H, OCH_3); 4.16-4.20 (*m*, 2H, H(α)-Ala, H(9'')-Fmoc); 4.27-4.35 (*"d"*, 3H, H(4'), (O-CH_2)-Fmoc); 4.53 (*m*, 1H, H(3')); 5.30 (*m*, 1H, H(2')); 5.81 (*"d"*, 1H, NH-carbamate); 6.21 (*d*, 1H, ^3J=1.7, H(1')); 6.70 (*br. d*, 0.5H, NH-amide); 6.77-7.76 (*m*, 21H, arom. Fmoc, Ph); 8.13 (*s*, 1H, H(2)); 8.54 (*s*, 1H, H(8)); 8.96 (*s*, 1H, N^6=CH).

3.3 Derivatisation of the solid support

3.3.1 Quantitative ninhydrin test

The quantitative ninhydrin test requires the use of three reagent solutions: 1) Phenol (40 g, 0.425 mol) dissolved in 10 ml ethanol; 2) 0.2 mM KCN in pyridine (100 ml); Sol. 1 and 2 are combined to form Sol. A. Sol. B is prepared dissolving ninhydrin in ethanol (0.28 M).

A sample of resin (10-15 mg of aminomethyl polystyrene (AMP) resin or 1.5-2.5 mg of Wang resin) is weighed into a small test tube and incubated for 5 minutes with 200 μl of Sol. A and 50 μl of Sol. B at 100-104 °C in an oil bath. Immediately following the designated incubation time, 60% aqueous ethanol is added and the solution is filtered through a pipette pasteur equipped with a cotton plug directly into a measuring flask. The solution is taken to a final volume of 2 or 5 ml, depending on the loading. When the amount of NH_2 groups/g is high (Wang resin) the solution is taken to 10 or 20 ml in order to obtain an absorbance value in the range of linearity of the Lambert-Beer law. 1 ml of the solution is put into a cuvette and the VIS absorbance is read at 570 nm ($\varepsilon = 15000\ cm^{-1}\ M^{-1}$). The aqueous ethanolic solution is used as blank. The resin drying is a critical parameter in order to compare the tests. We tried to use always the same conditions of drying. The resin was put in a funnel equipped with a filter and washed several times with DMF and CH_2Cl_2 before, first, drying with a water pump, and then with a membrane pump for 8-10 minutes. In order to aquire a very precise measure, the resin should be dried under high vacuum for several hours. This procedure, very time costing, was used only for very important samples.

3.3.2 Derivatisation of AMP solid support with hexamethylene diisocyanate (27)

Aminomethyl polystyrene (AMP) solid support (1 g, 28 μmol) was placed in a Teflon syringe. Hexamethylene diisocyanate (448 μl, 2.8 mmol) and DMAP (342 mg, 2.8 mmol) were dissolved in dry 1,2-dichloroethane (DCE, 4 ml) and poured onto the polymer support. The reaction was allowed to proceed overnight at r.t. under shaking. At the end of the reaction, the solution was pushed out of the syringe and the solid support was carefully rinsed with DCE (5 ml X 5) and CH_2Cl_2 (5ml X 5) followed by drying under reduced pressure. Quant. ninhydr. test: ~1 μmol/g of residual NH_2 groups. A capping procedure was performed to block any unreacted amino group using equal volumes (2 ml of each) of capping solutions (Cap. A: 1 M acetic anhydride, 1 M pyridine in DMF, Cap. B: 1 M *N*-methylimidazole (NMI in DMF) for

10 minutes followed by thoroughly washing with DMF (5 ml X 5) and CH_2Cl_2 (5 ml X 5) and then dried in vacuo.

3.3.3 Derivatisation of AMP solid support with succinic anhydride (28)

Aminomethyl polystyrene (AMP) solid support (1.5 g, 42 μmol) was placed in a Teflon syringe. Succinic anhydride (210 mg, 2.1 mmol) and DMAP (15.4 mg, 0.126 mmol) were dissolved in dry pyridine (6 ml) and poured onto the polymer support. The reaction was allowed to proceed overnight at r.t. under shaking. At the end of the reaction, the solution was pushed out of the syringe and the solid support was carefully rinsed with pyridine (5 ml X 5) and CH_2Cl_2 (5 ml X 5) followed by drying under reduced pressure. Quant. ninhydr. test: ~1 μmol/g of residual NH_2 groups. A capping procedure was performed as described before using 3 ml of each capping solution and reacted for 10 minutes. Thoroughly washing with DMF (5 ml X 5) and CH_2Cl_2 (5 ml X 5) preceded the drying in vacuo.

3.3.4 Derivatisation of AMP solid support with 3,6,9-trioxaundecanoic diacid (29)

Aminomethyl polystyrene (AMP) solid support (1.9 g, 53.2 μmol) was placed in a Teflon syringe. 3,6,9-trioxaundecanoic diacid (440 mg, 1.98 mmol), HBTU (403 mg, 1.06 mmol) and NMM (120 μl, 1.06 mmol) were dissolved in dry DMF (5 ml) and the diacid molecule was activated for 15 minutes before pouring this solution onto the polymer support. The reaction was allowed to proceed overnight at r.t. under shaking. At the end of the reaction, the solution was pushed out of the syringe and the solid support was carefully rinsed with DMF (6 ml X 5) and CH_2Cl_2 (6 ml X 5) followed by drying under reduced pressure. Quant. ninhydr. test: ~1 μmol/g of residual NH_2 groups. A capping procedure was performed as described before using 3 ml of each capping solution and reacted for 10 minutes. Thoroughly washing with DMF (6 ml X 5) and CH_2Cl_2 (6 ml X 5) preceded the drying in vacuo.

3.3.5 Attachment of 1,6-diaminohexane to the solid support 27 (32)

The derivatised solid support **27** (61 mg, 1.65 μmol) was placed in a Teflon syringe. 1,6-diaminohexane (96 mg, 0.823 mmol) and DMAP (60 mg, 0.494 mmol) were dissolved in dry DCE (1 ml) and poured onto the polymer support **27**. The reaction was allowed to proceed for 3 days at r.t. under shaking. At the end of the reaction, the solution was pushed out of the

syringe and the solid support was carefully rinsed with DCE (2 ml X 5) and CH_2Cl_2 (2 ml X 5) followed by drying under reduced pressure. Quant. ninhydr. test: 10.6 μmol/g of NH_2 groups.

3.3.6 Attachment of 1,6-diaminohexane to the solid support 28 (33)

The derivatised solid support **28** (1.5 g, 40.5 μmol) was placed in a Teflon syringe. Oxalyl chloride (2 ml) was added to CH_2Cl_2 (4 ml) and the activating solution was poured onto the solid support **28** and reacted overnight. After washing of the activated solid support with dry CH_2Cl_2 under an argon atmosphere and prolonged drying under high vacuum pump, 1,6-diaminohexane (941 mg, 8.1 mmol) and DMAP (98 mg, 0.8 mmol) were dissolved in dry CH_2Cl_2 (5 ml) and poured onto the polymer support. The reaction was allowed to proceed for 1 day at r.t. under shaking. At the end of the reaction, the solution was pushed out of the syringe and the solid support was carefully rinsed with CH_2Cl_2 (6 ml X 8) followed by drying under reduced pressure. Quant. ninhydr. test: 25.8 μmol/g of NH_2 groups.

3.3.7 Attachment of 1,6-diaminohexane to the solid support 29 (34)

The derivatised solid support **29** (156 mg, 4.13 μmol) was placed in a Teflon syringe. Oxalyl chloride (172 μl) was added to CH_2Cl_2 (0.5 ml) and the activating solution was poured onto the solid support **29** and reacted overnight. After washing of the activated solid support with dry CH_2Cl_2 under an argon atmosphere and prolonged drying under high vacuum pump, 1,6-diaminohexane (96 mg, 0.826 mmol) and DMAP (10.5 mg, 0.086 mmol) were dissolved in dry CH_2Cl_2 (5 ml) and poured onto the polymer support. The reaction was allowed to proceed for 1 day at r.t. under shaking. At the end of the reaction, the solution was pushed out of the syringe and the solid support was carefully rinsed with CH_2Cl_2 (2 ml X 8) followed by drying under reduced pressure. Quant. ninhydr. test: 26 μmol/g of NH_2 groups.

3.3.8 Attachment of the diamino PEG 900 spacer to the solid support 27 (35)

The derivatised solid support **27** (1 g, 27 μmol) was placed in a Teflon syringe. The diamino PEG 900 spacer (4.5 ml, 5.4 mmol) and DMAP (165 mg, 1.35 mmol) were dissolved in dry DCE (4 ml) and poured onto the polymer support **27**. The reaction was allowed to proceed for 3 days at r.t. under shaking. At the end of the reaction, the solution was pushed out of the

syringe and the solid support was carefully rinsed with DCE (5 ml X 5) and CH_2Cl_2 (5 ml X 5) followed by drying under reduced pressure. Quant. ninhydr. test: 9.6 μmol/g of NH_2 groups.

3.3.9 Attachment of the the diamino PEG 900 spacer to the solid support 28 (36)

The derivatised solid support **28** (100 mg, 2.7 μmol) was placed in a Teflon syringe. Oxalyl chloride (112 μl, 1.3 mmol) was added to CH_2Cl_2 (0.6 ml) and the activating solution was poured onto the solid support **28** and reacted overnight. After washing of the activated solid support with dry CH_2Cl_2 under an argon atmosphere and prolonged drying under high vacuum pump, the diamino PEG 900 (656 μl, 0.78 mmol) and DMAP (6 mg, 0.05 mmol) were dissolved in dry DCE (0.6 ml) and poured onto the polymer support. The reaction was allowed to proceed for 3 days at r.t. under shaking. At the end of the reaction, the solution was pushed out of the syringe and the solid support was carefully rinsed with DCE (2 ml X 5) and CH_2Cl_2 (2 ml X 5) followed by drying under reduced pressure. Quant. ninhydr. test: 8.9 μmol/g of NH_2 groups.

3.3.10 Attachment of the the diamino PEG 900 spacer to the solid support 29 (37)

The derivatised solid support **29** (60 mg, 1.62 μmol) was placed in a Teflon syringe. Oxalyl chloride (70 μl, 0.81 mmol) was added to CH_2Cl_2 (0.4 ml) and the activating solution was poured onto the solid support **28** and reacted overnight. After washing of the activated solid support with dry CH_2Cl_2 under an argon atmosphere and prolonged drying under high vacuum pump, the diamino PEG 900 (272 μl, 0.324 mmol) and DMAP (4 mg, 0.033 mmol) were dissolved in dry DCE (0.4 ml) and poured onto the polymer support. The reaction was allowed to proceed for 3 days at r.t. under shaking. At the end of the reaction, the solution was pushed out of the syringe and the solid support was carefully rinsed with DCE (1 ml X 5) and CH_2Cl_2 (1 ml X 5) followed by drying under reduced pressure. Quant. ninhydr. test: 9.7 μmol/g of NH_2 groups.

3.3.11 Attachment of BOC-sarcosine to the solid support 33 (38)

The derivatised solid support **33** (1.3 g, 33.5 μmol) was placed in a Teflon syringe. BOC-sarcosine (639 mg, 3.38 mmol), HBTU (1.28 g, 3.38 mmol) and NMM (751 μl, 6.76 mmol) were dissolved in dry DMF (5 ml) and the BOC-amino acid was activated for 15 minutes

before pouring this solution onto the polymer support. The reaction was allowed to proceed for 2 hours at r.t. under shaking. A second coupling was performed using a lower amount of activated BOC-sarcosine (127 mg, 0.67 mmol). At the end of the reaction, the solution was pushed out of the syringe and the solid support was carefully rinsed with DMF (6 ml X 5) and CH_2Cl_2 (6 ml X 5) followed by drying under reduced pressure. Qualitative ninhydrin test indicated incomplete reaction (blue colour). A third coupling was thus performed using: BOC-sarcosine (1.45 g, 7.66 mmol), HBTU (2.9 g, 7.66 mmol) and NMM (1.7 ml, 15.3 mmol) preactivation time 15 minutes and coupling time 3 h. After extensive washing, quant. ninhydr. test indicated a residual amount of ~1 μmol/g NH_2 groups. A capping procedure was performed as described before using 3 ml of each capping solution and reacted for 10 minutes. The BOC group was then removed treating the resin twice (8 + 3 min.) with a mixture of TFA/CH_2Cl_2 (95/5, 6 ml). A neutralisation step with triethylamine/DMF (20/80) for 1 minute followed the removal of the BOC group. Thoroughly washing with DMF (6 ml X 5) and CH_2Cl_2 (6 ml X 5) preceded the drying in vacuo.

3.3.12 Attachment of BOC-sarcosine to the solid support 35 (39)

The derivatised solid support **35** (0.7 g, 6.7 μmol) was placed in a Teflon syringe. BOC-sarcosine (290 mg, 1.53 mmol), HBTU (580 mg, 1.53 mmol) and NMM (340 μl, 3.06 mmol) were dissolved in dry DMF (3 ml) and the BOC-amino acid was activated for 15 minutes before pouring this solution onto the polymer support. The reaction was allowed to proceed for 3 hours at r.t. under shaking. A second coupling was performed using a half amount of activated BOC-sarcosine (145 mg, 0.77 mmol). At the end of the reaction, the solution was pushed out of the syringe and the solid support was carefully rinsed with DMF (3 ml X 5) and CH_2Cl_2 (3 ml X 5) followed by drying under reduced pressure. After extensive washing, quant. ninhydr. test indicated a residual amount of ~1 μmol/g NH_2 groups. A capping procedure was performed as described before using 2 ml of each capping solution and reacted for 10 minutes. The BOC group was then removed treating the resin twice (8 + 3 min.) with a mixture of TFA/CH_2Cl_2 (95/5, 4 ml). A neutralisation step with triethylamine/DMF (20/80) for 1 minute followed the removal of the BOC group. Thoroughly washing with DMF (3 ml X 5) and CH_2Cl_2 (3 ml X 5) preceded the drying in vacuo.

3.3.13 Attachment of succinic anhydride to the solid support 38 (40)

The derivatised solid support **38** (190 mg, 4.75 μmol) was placed in a Teflon syringe. Succinic anhydride (58 mg, 0.57 mmol) and DMAP (10 mg, 0.074 mmol) were dissolved in dry pyridine (1.5 ml) and poured onto the polymer support. The reaction was allowed to proceed for 20 hours at r.t. under shaking. At the end of the reaction, the solution was pushed out of the syringe and the solid support was carefully rinsed with pyridine (2 ml X 5) and CH_2Cl_2 (2 ml X 5) followed by drying under reduced pressure. A capping procedure was performed as described before using 1 ml of each capping solution and reacted for 5 minutes. Thoroughly washing with DMF (2 ml X 5) and CH_2Cl_2 (2 ml X 5) preceded the drying in vacuo.

3.3.14 Attachment of succinic anhydride to the solid support 39 (41)

The derivatised solid support **39** (400 mg, 3.6 μmol) was placed in a Teflon syringe. Succinic anhydride (50 mg, 0.5 mmol) and DMAP (8 mg, 0.065 mmol) were dissolved in dry pyridine (3 ml) and poured onto the polymer support. The reaction was allowed to proceed for 20 hours at r.t. under shaking. At the end of the reaction, the solution was pushed out of the syringe and the solid support was carefully rinsed with pyridine (4 ml X 5) and CH_2Cl_2 (4 ml X 5) followed by drying under reduced pressure. A capping procedure was performed as described before using 2 ml of each capping solution and reacted for 8 minutes. Thoroughly washing with DMF (4 ml X 5) and CH_2Cl_2 (4 ml X 5) preceded the drying in vacuo.

3.3.15 Immobilisation of DMT-dC^{an} on solid support 41

The derivatised solid support **41** was splitted in several batches of ~50 mg of resin (0.45 μmol) The same amount of nucleoside was used for all the reactions (15 mg, 20 μmol). The batches a) and b) were activated with HBTU (30 mg, 80 μmol) and NMM (9 μl, 80 μmol) in 0.5 ml of dry DMF. The solid support was activated for 20 minutes before adding the nucleoside. In batch a) the activating solution was washed out and *N*^{4}-anisoyl-5'-*O*-DMT-2'-deoxycytidine (~15 mg, 20 μmol) dissolved in 0.5 ml DMF was added to the resin. In batch b) the activating solution was left into the syringe and the nucleoside added as a solid. After 3 days of reaction, trityl assay was negative. Batches c) and d) were activated with CBMIT. CBMIT (130 μmol), prepared as described in Section 3.4.1, was poured onto the batches c)

and d) and the activation was prolonged for 20 minutes before adding the nucleoside. In batch c) the activating solution was washed out. After 3 days of reaction the trityl assay was negative indicating no coupling of the nucleoside to the solid support.

3.3.16 Immobilisation of DMT-T on solid support 41

Three other batches of the derivatised solid support **41** (50 mg, 0.45 μmol) were activated with oxalyl chloride (11 μl, 0.132 mmol) in CH_2Cl_2 (0.3 ml) overnight. After washing of the activated solid support with dry CH_2Cl_2 under an argon atmosphere and prolonged drying under high vacuum pump, the nucleoside 5'-O-DMT-thymidine (~11 mg, 20 μmol) was added to the solid support. In batches a) no base was added, in batch b) pyridine was added (~1 μl, 12 μmol) and in batch c) DMAP was added (~2 mg, 16 μmol). Trityl assay : batch a) negative, batch b) 0.76 μmol/g and batch c) 3.7 μmol/g.

3.3.17 Loading measurement of solid support 41 using *p*-NO_2-phenolate assay

Two batches of the derivatised solid support **41** (42 mg, 0.38 μmol) were activated: one with oxalyl chloride (10 μl, 0.116 mmol) in CH_2Cl_2 (0.5 ml) and the other with HBTU (40 mg, 0.105 mmol) and NMM (15 μl, 0.135 mmol). *p*-NO_2-phenol (24 mg, 0.18 mmol) was added to each batch and reacted for 4 hours at r.t. Basic hydrolysis was carried out with 0.1 M NaOH mixed with dioxane (1:1). The addition of 1 ml of this solution to a weighed amount of resin produced the *p*-NO_2-phenolate ion (strong yellow). Measuring the absorbance of this solution at λ=405 nm (ε=24651) indicated an amount of derivatisation of 2.6 μmol/g for the resin activated with HBTU and 9 μmol/g for the resin activated with oxalyl chloride.

3.3.18 Immobilisation of DMT-T on solid support 40

Two batches of the derivatised solid support **40** (30 mg, 0.74 μmol) were activated with oxalyl chloride (33 μl, 0.384 mmol) in CH_2Cl_2 (0.3 ml). After the activation, washing and drying of the solid supports, the nucleoside was added (8 mg, 0.0144 mmol). In one batch (a) pyridine (~1 μl, 0.0124 mmol) was added and in the other (b) DMAP was added (~1.5 mg, 0.0123 mmol). The reaction was prolonged for 20 hours at r.t. Trityl assay indicated a loading of 1 μmol/g for (a) and 12 μmol/g for (b).

3.3.19 Immobilisation of N^4-anisoyl-5'-*O*-DMT-3'succinate-2'-deoxycytidine on the solid support 39

The derivatised solid support **39** (54 mg, 0.378 μmol, 7 μmol/g of loading) was placed in a Teflon syringe. The nucleoside (12 mg, 15.7 μmol), HBTU (6 mg, 15.7 μmol) and NMM (3 μl, 27 μmol) were dissolved in dry DMF (0.4 ml) and the nucleoside was activated for 20 min. before adding this solution to the resin. The coupling reaction was prolonged for 20 h. After extensive washing and drying, trityl assay indicated a loading of 6.9 μmol/g.

3.3.20 Immobilisation of the building block 10 on solid supports 38 and 39 (43 and 44)

The derivatised solid support **38** (578 mg, 14.45 μmol) was placed in a Teflon syringe. The building block **10** (31 mg, 28 μmol), HATU (10.6 mg, 28 μmol) and NMM (6 μl, 54 μmol) were dissolved in dry DMF (2 ml) and the nucleoside was activated for 15 min. before adding this solution to the resin. After 20 hours the solution was pushed out of the syringe and used to attached the building block to a second portion of resin (**39**, 300 mg, 3 μmol). After extensive washing and drying, trityl assay indicated a loading of 15.5 μmol/g for the first batch of resin (**43**) and 14 μmol/g for the second batch (**44**). After a capping procedure and FMOC deprotection, quantitative ninhydrin test gave the following loadings: 11.4 μmol/g for **43** and 10.6 mmol/g for **44**.

3.4 Peptide coupling optimisation

3.4.1 Synthesis of CBMIT

Methyl triflate (1.35 ml, 12.3 mmol) was added dropwise via a syringe to a solution of carbonyldiimidazole (1 g, 6.17 mmol) in dry CH_3NO_2 (8 ml) at 10 °C (ice-water bath). After the addition of methyl triflate, the solution is allowed to reach room temperature. For the *in situ* activation the CBMIT solution is added to a suspension of FMOC-amino acid. In order to isolate CBMIT as a solid, it was precipitated from cold ether. The solution was added to a large amount of cold ether under fast stirring and CBMIT precipitated as a fine white powder in 70-80% yield.

m.p. 81-83 °C

^{1}H-NMR (DMSO-d_6) δ: 3.85 (*s*, 6H); 7.61 (*m*, 2H); 7.65 (*m*, 2H); 8.93 (*s*, 2H).

3.4.2 Synthesis of DEPBT

Diethylphosphorochloridate (20 ml, 0.135 mol) in CH_2Cl_2 (100 ml) was added dropwise to a solution of 3-hydroxy-1,2,3-benzotriazin-4(3*H*)-one (HOOBt, 20 g, 0.122 mol) and Et_3N (17 ml, 0.122 mol) in CH_2Cl_2 (1500 ml) at 0 °C. The reaction mixture was stirred for 2 hours at 0 °C and then for 2 hours at r.t. After the resulting triethylamine hydrochloride salt is removed by filtration, the solvent is evaporated. The residue is dissolved in EtOAc (200 ml) and washed with 0.1 M HCl, water and brine. The organic layer is dried ($MgSO_4$) and taken to dryness. The crude product (yellow oil) is recrystallised from EtOAc/hexane (1/5) to give pale yellow crystals of DEPBT (26.3 g, 72%).

m.p. 72-74.5 °C

Anal. Calcd. (C, 44.15; H, 4.72, N, 26.749 Found (C, 44.19, H, 4.83, N, 14.10)

^{1}H-NMR ($CDCl_3$) δ: 1.45-1.62 (*m*, 6H, ($OCH_2C\underline{H}_3)_2$); 4.46-4.57 (*m*, 4H, ($OC\underline{H}_2CH_3)_2$); 7.82-8.38 (*m*, 4H, arom.).

^{13}C-NMR ($CDCl_3$) δ: 15.96, 16.02, 66.33, 66.38, 122.52, 125.74, 128.94, 132.68, 135.29, 144.05, 150.26.

^{31}P-NMR ($CDCl_3$, 85% H_3PO_4 as external standard) δ: -0.46.

3.4.3 Synthesis of the model peptide $(Ala)_{12}$ (46a and 46b)

The peptide Ala_{12} was synthesised using CBMIT (**46a**) or HATU (**46b**) as activators. The procedure used was the same for both syntheses, only the activation step changed.
50 mg of Wang resin loaded with N^{α}-FMOC-L-alanine (FMOC-Ala, loading 0.41 mmol/g, 0.0205 mmol) were placed in a Teflon syringe and rinsed with DMF (2ml X 8). After washing, the FMOC group was cleaved by shaking the resin for 10 min. and for 5 min. using a solution of 20% piperidine in DMF (2 ml). The deprotection solution was used to monitor the syntheses. 200 µl of this solution were transferred into a measuring flask (V=2 ml) and

diluted with DMF to the final volume of the flask. The absorbance of this solution was measured at λ=300 nm and DMF was used as blank. This procedure was repeated at each cycle after the FMOC deprotection.

The coupling reactions were performed using 10 eq. of FMOC-Ala in 1 ml DMF and the reaction time was of 30 minutes. A second coupling was always performed to ensure maximum coupling yields.

Activation procedure for **46a** (CBMIT): the activator CBMIT was prepared as described above using 806 mg (4.97 mmol) of carbonyldiimidazole and 1.09 ml (9.94 mmol) of methyl triflate in CH_3NO_2 (5 ml). The solution of CBMIT was added to a suspension of FMOC-Ala (1.55 g, 4.97 mmol in 5 ml of CH_3NO_2). After the precipitation of the activated FMOC-amino acid, DMF (14 ml) was added to redissolve the precipitate. The solution of activated FMOC-alanine was divided in small portions (1 ml each) and stored at –22 °C.

Activation procedure for **46b** (HATU): at every coupling a new portion of FMOC-amino acid was activated using 64 mg (0.205 mmol) of FMOC-Ala, 80 mg (0.205 mmol) of HATU and 54 μl (0.41 mmol) of collidine in 1 ml of DMF (~0.2 M), no preactivation was performed.

After the second coupling, a capping step was performed using 1 ml of a solution 0.5 M acetic anhydride, 0.5 M NMM in DMF (capping time: 10 minutes). FMOC-deprotection started a new cycle.

At the end of the synthesis, the resin was placed in a Falcon tube (40 ml) and the peptide was cleaved from the resin using the mixture TFA/H_2O 95/5 (5 ml). The cleavage reaction was performed at r.t. under magnetic stirring for 2.5 hours. After this time, the solution was filtered directly into a Falcon tube containing 30 ml of cold *tert*-butyldimethylether allowing the precipitation of the peptide. After centrifugation (10 min.) the supernatant solution was decanted and the crude peptide redissolved in neat TFA. Since addition of water caused the precipitation of part of the product, no lyophilisation was possible and the crude material was dried under vacuum.

MALDI-Tof analysis was performed using a mixture of 1 μl of crude peptide (dissolved in TFA) with 10 μl of a matrix solution (saturated solution of α-cyano-4-hydroxycinnamic acid in CH_3CN/H_2O 50/50). 1 μl of the mixture was placed on a gold MALDI plate and air dried before the measure.

3.4.4 Synthesis of the model peptide $(Ala)_{12}ArgAla(Arg)_2$ (47a , 47b and 47c)

$(Ala)_{12}ArgAla(Arg)_2$ (**47a**: CBMIT; **47b**: HATU; **47c**: DEPBT) was synthesised using the same procedure for the three syntheses exept the activation.
150 mg of Wang resin loaded with N^{α}-FMOC-N^{ω}-(2,2,5,7,8-pentamethylchroman-6-sulfonyl)-L-arginine (FMOC-Arg(Pmc), loading 0.5 mmol/g, 0.075 mmol) were placed in a Teflon syringe and rinsed with DMF (3ml X 8). After washing, the FMOC group was cleaved by shaking the resin for 10 min. and for 5 min. using a solution of 20% piperidine in DMF (3 ml). The FMOC-deprotecion level was used to monitor the synthesis of **47a** whereas **47b** and **47c** were monitored by the quantitative ninhydrin test. For **47a** 100 μl of the deprotection solution were used for the monitoring following the same procedure as described above.

The coupling reactions were performed using 5 eq. of FMOC-amino acid in 1 ml DMF and the reaction time was of 35 minutes. A second coupling was always performed to ensure maximum coupling yields. For every coupling 117 mg (0.375 mmol) of FMOC-Ala or 248 mg (0.375 mmol) of FMOC-Arg(Pmc) were activated using: HATU (142 mg, 0.375 mmol), NMM (83 μl, 0.75 mmol) or DEPBT (224 mg, 0.75 mmol), NMM (83 μl, 0.75 mmol). DEPBT was used with a preactivation of 30 minutes. As for the previous synthesis, with CBMIT, the total amount of FMOC-amino acid was activated at the beginning of the synthesis and then stored in portions of 1 ml (0.37 M) at -22 °C.

After the second coupling, a capping step was performed using 2 ml of a solution 0.5 M acetic anhydride, 0.5 M NMM in DMF (capping time: 5 minutes). FMOC-deprotection started a new cycle. The cleavage procedure was the same as described above. The three crude products were soluble in 50% aqueous CH_3COOH and could be lyophilised.

MALDI-ToF was performed as described above. HPLC analysis was carried out using: Buffer A = 0.1% TFA in H_2O/CH_3CN 95/5; Buffer B = 0.1% TFA in H_2O/CH_3CN 5/95. Gradient: 0-80% B in 80 minutes (linear).

3.4.5 Synthesis of the model peptide $(Ala)_{12}$ on AMP solid support (48a , 48b)

This model peptide was synthesised using CBMIT **48a** and DEPBT **48b** as activators. AMP solid support (409 mg, 11.4 μmol) was placed in a Teflon syringe. The following procedure was used for both synthesis: the FMOC group was cleaved by shaking the resin for 7 min. and for 3 min. using a solution of 20% piperidine in DMF (3 ml); coupling time (25 minutes); for the first four FMOC-amino acids a single coupling (15 eq.) was performed, for the 5^{th} and 6^{th}

amino acids a double coupling procedure was performed using 10 eq. of FMOC-Ala for each coupling and for the rest of the synthesis a double coupling with 15 eq. of FMOC-Ala each time was used. The volume of the coupling solution was of 0.6 ml. The activation was performed *in situ* with both activators using the adequate amounts of reagents. The syntheses were monitored by quantitative ninhydrin test at every step.

3.5 Conjugate synthesis and secondary structure determination

3.5.1 General procedure for the synthesis of the conjugates: peptide moiety synthesis

The peptidic moieties of the conjugates were synthesized prior the RNA synthesis following a stepwise, FMOC-based solid support oligopeptide synthesis. They were synthesized on batches of resin **43** or **44** (~120 mg, loading ~11 μmol/g) prepared as described above. The quality of the syntheses was monitored by measuring the FMOC-deprotection solution at 300 nm (10 % v/v in DMF against pure DMF) after every coupling and by the quantitative ninhydrin test every four couplings. The allyl group was used as side-chain protection for glutamic acid.

The following coupling cycles were used for the synthesis of all the conjugates: a) FMOC-deprotection: treatment with 2% piperidine, 2% DBU in DMF for 8 min followed by a second deprotection step of 5 min; b) Amino acid coupling: DEPBT used as activator with a preactivation time of 30 minutes; the coupling mixture (deep yellow), added to the solid support, was shaken for 1-2 h and the reaction repeated a second time. The quantities of FMOC-amino acid used varied from one synthesis to the other:

Synthesis of **50b** and **50f**: first five couplings—FMOC-Ala (200 eq.), DEPBT (400 eq.), NMM (400 eq.) dissolved in dry DMF to a final concentration of 0.4 M; from the 6^{th}-7^{th} amino acid—FMOC-Ala (400 eq.), DEPBT (800 eq.), NMM (800 eq.) dissolved in dry DMF to a final concentration of 0.8-1.2 M resin and the coupling time prolonged to 3 h.

Synthesis of **50c**: first nine couplings—FMOC-Ala (100 eq.), DEPBT (200 eq.), NMM (200 eq.) dissolved in dry DMF to a final concentration of 0.2 M; from the 10^{th} amino acid—FMOC-Ala (200 eq.), DEPBT (400 eq.), NMM (400 eq.) dissolved in dry DMF to a final concentration of ~0.5 M, the coupling time prolonged to 3 h.

Synthesis of **50d**: all couplings carried out with— FMOC-Ala (200 eq.), DEPBT (400 eq.), NMM (400 eq.) dissolved in dry DMF to a final concentration of 0.4-0.5 M.

Synthesis of **50e**, **50g** and **50h**: all couplings carried out with— FMOC-Ala (200 eq.), DEPBT (400 eq.), NMM (400 eq.) dissolved in dry DMF to a final concentration of ~0.3 M.

Capping procedure: 2 min treatment using an equivalent volume of Cap. A and Cap. B solutions.

The quantitative ninhydrin test gave an estimation of the peptide synthesis yields: **50b**, 71%; **50c** 30%, **50d** 36%, **50e** 24%, **50f** 57%, **50g** and **50h** ~40%.

3.5.2 General procedure for the synthesis of the conjugates: oligoribonucleotide moiety synthesis

Immediately after the coupling of the last amino acid, a part of the solid support **50** was transferred to a '1 μmole-reaction column' for ABI DNA/RNA Synthesizer (70 mg maximum capacity for aminomethyl polystyrene based resin). 35 mg (~0.4 μmol) for **52b**, **52f**, **52h** and 70 mg (~0.7 μmol) for **52c**, **52d**, **52e**, **52g** of solid support **50** were used for the RNA strand synthesis. The following 2'-*O*-*tert*-butyldimethylsilyl ribonucleoside β-cyanoethyl phosphoramidites were used for the synthesis of **52b**, **52c**, **52f**: A^{Pac}, G^{iPrPac}, C^{Ac}, U. **52d**, **52e**, **52g** and **52h** were synthesised using the 2'-*O*-TOM ribonucleoside β-cyanoethyl phosphoramidites: A^{Ac}, C^{Ac}, G^{Ac}, U. The reagents used were: 0.25 M 5-ethylthio-1*H*-tetrazole/CH_3CN (activator), 3% Cl_3CCOOH/CH_2Cl_2 (detritylation), Ac_2O/lutidine/THF (Cap A), NMM/THF (Cap B); 0.02 M I_2/H_2O/pyridine/THF (oxidation), 40% aq. CH_3NH_2/33% ethanolic CH_3NH_2 (1:1) (final cleavage and deprotection).

Coupling yields per step were monitored by automated conductivity integration. Coupling cycles: the standard 1 μmol-RNA coupling cycle (ABI, version 2.01) was modified into a 'couple-cap-ox-cap' cycle: coupling (WAIT 420 sec), 1st capping (WAIT 5 sec), oxidation (WAIT 45 sec), thorough CH_3CN/Ar wash (2 x), 2nd identical capping, another CH_3CN/Ar wash (2 x), detritylation. Average stepwise yields (~97.5%). The overall RNA yields were the following: **52b** 67%, **52c** no detection (technical problem); **52d** 67%, **52e** 55%, **52f** 60%, **52g** 55%, **52h** 43%.

For the conjugates **52c**, **52d**, **52e**, **52g**, **52h** containing glutamic acid residues a Pd(0) treatment followed the end of the synthesis to remove the allyl side chain protecting group. The deprotection was accomplished under Ar atmosphere with $Pd(PPh_3)_4$ (15 mg, ~20 eq) and $PhSiH_3$ (2 mg, 24 eq.) as an allyl group scavenger in CH_2Cl_2 (1 ml, 2 x 15 min). After pushing out of the column the Pd solution, the resin was extensively washed with CH_2Cl_2.

The polymer material was finally washed with ammonium N,N-diethyldithiocarbamate (0.5% w/w) in DMF to ensure the removal of any contaminating palladium residue.

3.5.3 Cleavage/Deprotection/Work up

After treatment of the solid support on the synthesiser with CH_3NH_2 at room temperature for 2 h, the collected liquid (~1.5 ml) was immediately evaporated in a SpeedVac.

Desilylation (2'-TBDMS-monomers): the crude product (persilylated RNA) from 0.7 μmol (0.4 μmol) syntheses was treated with 200 μl (150 μl) neat $Et_3N{\cdot}3HF$ and 60 μl (40 μl) DMF, heated and vortexed at 65 °C, and quenched at RT after 1.5 h with 30 μl (20 μl) sterile water. The crude deprotected RNA was precipitated by adding 2 ml (1 ml) *n*-butanol. After 2 h at –20 °C, the precipitate was spinned down and the pellets dissolved in 1 ml buffer A for strong anion exchange HPLC.

Desilylation (2'-TOM-monomers): the crude product (persilylated RNA) from 0.7 μmol (0.4 μmol) syntheses was treated with 350 μl (200 μl) 1M TBAF in THF at r.t. for 16 hours. At the end of the reaction, these solutions were filtered through NAP-10 columns to eliminate the fluoride ions.

Crude material: **52b:** 52 O.D.; **52c:** 55 O.D.; **52d:** 70 O.D.; **52e:** 72 O.D.; **52f** 28 O.D.; **52g:** 92 O.D.; **52h:** 35 O.D. (**o**ptical **d**ensity at 260 nm).

3.5.4 Purification

Strong Anion Exchange: After heating the RNA solution in a closed Eppendorf vial for 1 min at 100 °C, 500 ml crude material were injected (still warm) onto the SAX HPLC column. Buffer A: 20 mM Na/K phosphate, pH 7.0; buffer B: 0.6 N NaCl in buffer A. Gradient used for **52b**, **52f**, **52h**,: 0-65% B in 10 min / 8 min isocratic section at 65% B / 65-100% B in 3 min / stay at 100% for 20 min. Gradient used for **52c**, **52d**, **52e**, **52g**: 0-20% B in 6 min / 20-60% B in 7 min/ 15 min isocratic section at 65% B / 65-100% B in 2 min / stay at 100% for 20 min. For all the conjugates the product eluted as two main peaks, the first one at around 60-65% buffer B and the second one at 100% B.

Reversed Phase Desalting: The central part of each of the two peaks was collected (~10-15 ml) and injected into a C_{18} reverse phase HPLC column (8 x 250 mm, ODS, buffer A: nanopure water, flow rate 2 ml/min) to be desalted. Multiple injections of 1.5 ml volume were necessary to load all the product onto the column. After all the conjugate had been adsorbed,

the excess salts were washed off the column with water (up to 6 ml/min no conjugate desorbtion was observed). The product was eluted by a step gradient to 90% CH_3CN/H_2O (= 100% B, 3 ml/min) into a ~4 ml fraction. The fraction was subsequentely concentrated on a Rotavap system under reduced pressure to a volume of ~1 ml (heating bath 40 °C, interior of the flask silanized with ~2% $(CH_3)_2SiCl_2/CCl_4$). Concentration using a SpeedVac caused the precipitation of a part of the product that was impossible to redissolve aftewards.

Reversed Phase Chromatography: The two concentrated and desalted fractions from SAX-HPLC were then analysed and purified (when possible) by RP-HPLC. Buffer A: 0.1 M aq. NaOAc; buffer B: CH_3CN/buffer A (9:1). Gradient used: 0-60% B in 60 min (linear). For all the conjugates, the target product proved to be present only in the first peak of the SAX injection. RP-HPLC profiles of all of the conjugates revealed the precence of a very broad peak eluting in front of all the other peaks (t_R ~ 10-15 min). This peak is predominant for the conjugates containing very hydrophobic peptides (**52f**, **52g**, **52h**) leading to the impossibility to isolate a single species.

The fractions collected from RP-HPLC purification were desalted (as described above) and, before desorption of the product from the column, the K^+/Na^+ counterions were exchanged with NH_4^+ ions by injecting 1.5 ml 0.25 M NH_4OAc (sterile-filtered). The excess ammonium salt was washed off with water (6 ml/min) and the desalted conjugate (NH_4^+ form) was eluted within 4-5 ml as described above. This fraction was concentrated and analysed by MALDI-ToF. 1-3 O.D. were obtained as purified material. Aqueous stock solutions of desalted material could be stored for over 12 months in RNAse-free Eppendorf tubes (Ambion) at 4 °C without degradation (do not freeze!) and were directly used for CD spectroscopy and thermal denaturing measurements.

3.5.5 Mass spectrometry

MALDI-ToF mass spectra were acquired on a Bruker Riflex III spectrometer in negative ion, linear mode. A mixture of 2,4,6-trihydroxy acetophenone (0.3 M in ethanol), diammonium citrate (0.1 M in water), and CH_3CN (5:2:3) was used as matrix. 1 μl of the matrix were mixed with 1 μl of the product (~3 O.D.) and 0.3 μl of this solution were deposited on the plate and air dried. Results: **52a** [pRNA-$(Ala)_8$: $C_{232}H_{303}N_{91}O_{163}P_{22}$] $[M–H]^–$: calcd 7654.8, found 7649.8; **52b** [RNA-$(Ala)_{15}$: $C_{253}H_{337}N_{98}O_{167}P_{21}$] $[M–H]^–$: calcd 8072.4, found 8072.2; **52b** [RNA-$(Ala)_{16}$: $C_{256}H_{342}N_{99}O_{168}P_{21}$] $[M–H]^–$: calcd 8143.5, found 8139.9; **52c** [RNA-$(Ala)_{10}$Glu: $C_{243}H_{319}N_{94}O_{165}P_{21}$] $[M–H]^–$: calcd 7846.1, found 7839.8; **52d** [RNA-

$(Ala)_{18}(Glu_2pGlu)$: $C_{277}H_{371}N_{104}O_{178}P_{21}$] $[M–H]^-$: calcd 8654.9, found 8654.0; **52e** [RNA-$(Ala)_7Glu(Ala)_7Glu(Ala)_4(Glu)_2$: $C_{282}H_{380}N_{105}O_{182}P_{21}$] $[M–H]^-$: calcd 8802.1, found 8819.5; **7f** [RNA-$(Ala)(Leu)_{18}Glu$: $C_{324}H_{472}N_{103}O_{174}P_{21}$] $[M–H]^-$: calcd 9259.2, found 7138, 7220 and other peaks around these values; **7g** [RNA-$(Ala)(Leu)_{18}Glu_3$: $C_{334}H_{486}N_{105}O_{181}P_{21}$] $[M–H]^-$: calcd 9517.5, found 7248 (very broad peak) ; **7h** [RNA-$(Ala)_{20}$: $C_{268}H_{361}N_{103}O_{172}P_{21}$]/[RNA-$(Ala)_{21}$: $C_{271}H_{367}N_{104}O_{173}P_{21}$]/ [RNA-$(Ala)_{22}$: $C_{274}H_{371}N_{105}O_{174}P_{21}$] $[M–H]^-$: calcd 8427.8/8498.9/8569.9, found 8429.9/8501.6/8571.8 (approximately 43:37:20).

3.5.6 CD spectroscopy

CD measurements were performed in 100 mM NaCl, 10 mM NaH_2PO_4 buffer, pH 7.5 (NaOH) using an Aviv Model 62 A DS instrument equipped with a thermoelectric temperature controller. Spectra were registered at 0°, 25° and 60 °C using a 1 cm-3 ml Teflon-stoppered quarz cuvette with 2 ml solution volume (1.75 ml of buffer and 0.25 ml of conjugate stock solution: ~2.4 μM conjugate solution, $A_{260nm,25°C} \approx 0.35$). The region scanned was 200-350 nm with 0.5 nm wavelength increments. Spectra are the average of 3 scans, baseline corrected, normalised with respect to the number of nucleotide residues and to concentration ($\lambda_{max,\,RNA}$ at 260 nm and $\varepsilon_{260,calcd}$=145'100 $M^{-1}cm^{-1}$) and smoothed.

3.5.7 Denaturation profiles and thermodynamic analysis

Melting experiments were carried out on a Perkin-Elmer Lambda Bio 40 spectrophotometer equipped with a heating-cooling block thermocontroller PTP-6. The data were collected using the software UV WinLab® (for Lambda) and WinTemp® (for PTP-6). The melting profiles were registered using 1 cm/1 ml Teflon-stoppered quarz cuvette using as a buffer a solution of 100 mM NaCl/10 mM NaH_2PO_4, pH 7.5 (NaOH). 1 ml of conjugate solution (1.0 ml buffer + few μl conjugate stock solution) were instantaneously pre-heated to 100 °C for 2 minutes and cooled down slowly before measuring. The absorbance at 260 nm was recorded every 0.1 ° between 20 and 102 °C with a heating rate of 1.0 °/min. For stability reasons all profiles were measured in the 'up' mode and every sequence was analyzed at least three times.

The 'optimized temperature-window fitting' method developed by us, which takes into account only the main transition of the RNA hairpin above ~85° C, was used to abstract the van't Hoff thermodynamics from the denaturation profiles. The fitting procedure is described in detail in the Supporting Information of ref. [122b].

3.5.8 AFM analysis

AFM imaging was carried out using a Thermomicroscope Explorer AFM (Santa Clara, USA) equipped with a 100 μm tripod scanner, in non-contact mode, using high frequency (F_0=320 KHz) pyramidal cantilevers with silicon probes at a scan frequency of 1 Hz. Images are processed with the SPMLab 5.01 software package. Samples were prepared by depositing a volume of 5-15 μl of each compound solution on glass plates, imaging was processed after drying the samples overnight at 37 °C.

4. REFERENCES

References

[1] M. Nomura, P. Traub, H. Bechmann, *Nature* **1968**, *219*, 793.

[2] a) H. Hampl, H. Schulze, K. H. Nierhaus, *J. Biol. Chem.* **1981**, *256*, 2284.
b) F. J. Franceschi, K. H. Nierhaus, *J. Biol. Chem.* **1990**, *265*, 16676.

[3] a) H. F. Noller, V. Hoffarth, L. Zimniak, *Science* **1992**, *256*, 1416.
b) P. Khaitovich, A. S. Mankin, R. Green, L. Lancaster, H. F. Noller, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 85.

[4] R. K. Agrawal, J. Frank, *Curr. Opin. Struct. Biol.* **1999**, *9*, 215.

[5] N. Ban, P. Nissen, J. Hansen, P. B. Moore, T. A. Steitz, *Science* **2000**, *289*, 905.

[6] P. Nissen, J. Hansen, N. Ban, P. B. Moore, T. A. Steitz, *Science* **2000**, *289*, 920.

[7] M. Welch, J. Chastang, M. Yarus, *Biochemistry* **1995**, *34*, 385.

[8] a) N. Polacek, M. Gaynor, A. Yassin, A. S. Mankin, *Nature* **2001**, *411*, 498.
b) J. Thompson, D. F. Kim, M. O' Connor, K. R. Liebermann, M. A. Bayfield, S. T. Gregory, R. Green, H. F. Noller, A. E. Dahlberg, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9002.

[9] a) W. Gilbert, *Nature* **1986**, *395*, 260.
b) M. Yarus, *Curr. Opin. Chem. Biol.* **1999**, *3*, 260.
c) G. F. Joyce, *Nature* **2002**, *418*, 214.
d) K. Kruger, P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, T. R. Cech, *Cell* **1982**, *31*, 147.

[10] a) P. A. Lohse, J. W. Szostak, *Nature* **1996**, *381*, 442.
b) T. W. Wiegand, R. C. Janssen, B. E. Eaton, *Chem. Biol.* **1997**, *4*, 675.
c) P. J. Unrau, D. P. Bartel, *Nature* **1998**, *395*, 260.

[11] S. W. Fox, K. Harada. *Science* **1958**, *128*, 1214.

[12] a) G. Ourisson, Y. Nakatani, *Chem. Biol.* **1994**, *1* 11.
b) A. D. Miller, *ChemBioChem.* **2002**, *3*, 45.

[13] a) D. W. Deamer, *Nature* **1985**, *317*, 792.
b) D. W. Deamer, *Microbiol. Mol. Biol. Rev.* **1997**, 239.

[14] a) G. Blobel, *Proc. Natl. Acad. Sci. USA* **1980**, *77*, 1496.
b) T. Cavalier-Smith, *J. Mol. Evol.* **1990**, *30*, 7.

[15] A. C. Chakrabarti, D. W. Deamer, *J. Mol. Evol.* **1994**, *39*, 1.

[16] W. R. Hargreaves, D. W. Deamer, *Biochemistry* **1978**, *17*, 3759.

[17] a) J. D. Lear, Z. R. Wassermann, W. F. DeGrado, *Science* **1988**, *240*, 1177.

b) T. J. Gibson, A. I. Lamond, *J. Mol. Evol.* **1990**, *30*, 7.

[18] a) P. A. Bachmann, P. L. Luisi, J. Lang, *Nature* **1991**, *357*, 57.

b) E. Blöchliger, M. Blocher, P. Walde, P. L. Luisi, *J. Phys. Chem. B* **1998**, *102*, 10383.

c) J. W. Szostak, D. P. Bartel, P. L. Luisi, *Nature* **2001**, *409*, 387.

[19] R. W. Holley, J. Apgar, G. A. Everett, J. T. Madison, M. Marquisee, S. H. Merrill, J. R. Penswick, A. Zamir, *Science* **1965**, *147*, 1462.

[20] a) S. H. Kim, G. F. Quigley, F. L. Suddah, A. McPherson, D. Sneden, J. J. Kim, J. Weinziel, A. Rich, *Science* **1973**, *179*, 285.

b) J. D. Robertus, J. E. Ladner, J. T. Finch, D. Rhodes, R. S. Brown, B. F. C. Clark, A. Klug, *Nature* **1974**, *250*, 546.

[21] T. H. Fraser, A. Rich, *Proc. Natl. Acad. Sci. USA* **1975**, *45*, 31.

[22] J. Preiss, P. Berg, E. J. Ofengand, F. H. Bergman, M. Dieckmann, *Proc. Natl. Acad. Sci. USA* **1959**, *45*, 319.

[23] S. Chladek, M. Sprinzl, *Angew. Chem.* **1985**, *97*, 377.

[24] a) S. Pitsch, *Chimia* **2001**, *55*, 60.

b) A. Stutz, C. Höbartner, S. Pitsch, *Helv. Chim. Acta* **2000**, *83*, 2477.

c) A. Stutz, S. Pitsch, *Synlett* **1999**, *S1*, 930.

[25] a) R. K. Agrawal, A. B. Heagle, P. Penczek, R. A. Grassucci, J. Frank, *Nat. Struct. Biol.* **1999**, *6*, 643.

b) H. Stark, M. V. Rodnina, H.-J. Wieden, M. van Heel, W. Wintermeyer, *Cell* **2000**, *100*, 301.

[26] R. E. Monro, K. A. Marker, *J. Mol. Biol.* **1967**, *25*, 347.

[27] a) T. M. Schmeing, A. C. Seyla, J. L. Hansen, B. Freedborn, J. K. Soukup, S. A. Scaringe, S. A. Strobel, P. B. Moore, T. A. Steitz, *Nat. Struct. Biol.* **2002**, *9*, 225.

b) J. L. Hansen, T. M. Schmeing, P. B. Moore, T. A. Steitz, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11670.

[28] K. M. Choi, R. Brimacombe, *Nucleic Acids Res.* **1998**, *26*, 887.

[29] R. Beckmann, C. M. T. Spahn, N. Eswar, J. Helmers, P. A. Penczek, A. Sali, J. Frank, G. Blobel, *Cell* **2001**, *107*, 361.

[30] G. W. Ramachandran, C. Ramakrisnan, V. Sasisekharan, *J. Mol. Biol.* **1963**, *7*, 95.

[31] L Pauling, R. B. Corey, H. R. Branson, *Proc. Natl. Acad. Sci. USA* **1951**, *37*, 205.

[32] a) D. Kihara, Y. Zhang, H. Lu, A. Kolinski, J. Skolnick, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5993.

b) E. S. Huang, R. Samudrala, J. W. Ponder, *J. Mol. Biol.* **1999**, *290*, 267.

[33] O. Rathore, D. Y. Sogah, *J. Am. Chem. Soc***. 2001**, *123*, 5231.

[34] a) D. A. Tirrell, *Science* **1996**, *271*, 39.

b) A. Lazaris, S. Arcidiacono, Y. Huang, J.-F. Zhou, F. Dungay, N. Chretien, E. A. Welsh, J. M. Soares, C. N. Karatzas, *Science* **2002**, *295*, 472.

c) S. Kubik, *Angew. Chem. Int. Ed.* **2002**, *41*, 2721.

[35] a) C. M. Dobson, *Trends Biochem. Sci.* **1999**, *24*, 329.

b) J. D. Sipe, A. S. Cohen, *J. Struct. Biol.* **2000**, *130*, 88.

[36] B. H. Zimm, J. K. Bragg, *J. Chem. Phys.* **1959**, *31*, 526.

[37] a) R. J. Kennedy, K.-Y. Tsang, D. S. Kemp, , *J. Am. Chem. Soc.* **2002**, *124*, 934.

b) B. S. Kinnear, M. F. Jarrold, *J. Am. Chem. Soc.* **2001**, *123*, 7907.

c) J. S. Miller, R. J. Kennedy, D. S. Kemp, *Biochemistry* **2001**, *40*, 305.

d) R. Improta, V. Barone, K. N. Kudin, G. E. Scuseria, *J. Am. Chem. Soc.* **2001**, *123*, 3311.

e) E. J. Spek, C. A. Olson, Z. Shi, N. R. Kallenbach, , *J. Am. Chem. Soc.* **1999**, *121*, 5571.

f) C. A. Rohl, W. Fiori, R. L. Baldwin, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 3682.

g) R. R. Hudgins, M. F. Jarrold, *J. Am. Chem. Soc.* **1999**, *121*, 3494.

h) D. S. Kemp, L. Oslick, T. J. Allen, *, J. Am. Chem. Soc.* **1996**, *118*, 4249.

i) K. Groebke, P. Renold, K. Y. Tsang, T. J. Allen, K. F. McClure, D. S. Kemp, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 4025.

j) P. C. Lyu, J. C. Sherman, A. Chen, N. R. Kallenbach, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 5317.

k) D. S. Kemp, J. G. Boyd, C. C. Muendel, *Nature* **1991**, *352*, 451.

l) A. Chakrabarrty, J. A. Schellman, R. L. Baldwin, *Nature* **1991**, *351*, 586.

m) P. C. Lyu, M. I. Liff, L. A. Marky, N. R. Kallenbach, *Science* **1990**, *250*, 669.

n) S. Marqusee, V. H. Robbins, R. L. Baldwin, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 5286.

o) J. R. Parrish, E. R. Blout, *Biopolymers* **1972**, *11*, 1001.

[38] a) M. Takano, T. Yamato, J. Higo, A. Suyama, K. Nagayama, *J. Am. Chem. Soc.* **1999**, *121*, 605.

b) X. Wu, S. Wang, *J. Phys. Chem. B* **2001**, *105*, 2227.

c) H. S. Son, B. H. Hong, C.-H. Lee, S. Yun, K. S. Kim, *J. Am. Chem. Soc.* **2001**, *123*, 514.

[39] a) V. Muñoz, L. Serrano, *Structural Biology* **1994**, *1*, 399.
b) V. Muñoz, L. Serrano*, J. Mol. Biol.* **1995**, *245*, 275.
c) V. Muñoz, L. Serrano, *Biopolymers* **1997**, *41*, 494.

[40] a) R. W. Woody, *Methods in Enzymology* **1995**, *246*, 34.
b) R. W. Woody, in *"Circular Dichroism*", (Eds. N. Koji, N. Berova), VCH, New York, **1994**, Chap.17.
c) W. C. Johnson, *Proteins* **1990**, *7*, 205.

[41] a) C. Toniolo, A. Polese, F. Formaggio, M. Crisma, J. Kamphuis, *J. Am. Chem. Soc.* **1996**, *118*, 2744.
b) T. S. Sudha, E. K. S. Vijayakumar, P. Balaram*, Int. J. Peptide Protein Res***. 1983**, *22*, 464.

[42] R. C. Manning, M. Illangasekare, R. W. Woody, *Biophysical Chemistry* **1988**, *31*, 77.

[43] a) R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149.
b) R. B. Merrifield, *Biochemistry* **1964**, *3*, 1385.

[44] L. A. Carpino, G. Y. Han, *J. Org. Chem.* **1972**, *37*, 3404.

[45] a) E. Atherton, H. Fox, D. Harkiss, C. J. Logan, R. C. Sheppard, B. J. Williams, *J. Chem. Soc. Chem. Comm.* **1978**, 537.
b) E. Atherton, H. Fox, D. Harkiss, R. C. Sheppard, *J. Chem. Soc. Chem. Comm.* **1978**, 539.
c) E. Atherton, C. J. Logan, R. C. Sheppard, *J. Chem. Soc. Perkin Trans. 1* **1981**, 538.

[46] S.-S. Wang, *J. Am. Chem. Soc.* **1973**, *95*, 1328.

[47] D. S. King, C. G. Fields, G. B. Fields, *Int. J. Peptide Protein Res.* **1990**, *36*, 255.

[48] E. Kaiser, R. L. Colescott, C. D. Bossinger, P. I. Cook, *Anal. Biochem.* **1970**, *34*, 595.

[49] E. Atherton, R. C. Sheppard, *Solid phase peptide synthesis: a practical approach,* IRL Press, **1989**, pp. 107-130.

[50] a) P. T. Gilham, H. G. Khorana, *J. Am. Chem. Soc.* **1958**, *80*, 6212.
b) G. Weimann, H. G. Khorana, *J. Am. Chem. Soc.* **1962**, *84*, 419.
c) H. Schaller, G. Weimann, B. Lerch, H. G. Khorana, *J. Am. Chem. Soc.* **1963**, *85*, 3821.
d) G. S. Ti, B. L. Gaffney, R. A. Jones, *J. Am. Chem. Soc.* **1982**, *104*, 1316.

[51] a) L. J. McBride, M. H. Caruthers, *Tetrahedron Lett.* **1983**, *24*, 2953.

b) A. Kume, R. Iwase, M. Sekine, T. Hata, *Nucleic Acids Res.* **1984**, *12*, 8525.

c) L. J. McBride, R. Kierzek, S. L. Beaucage, M. H. Caruthers*, J. Am. Chem. Soc.* **1986**, *108*, 2040.

[52] a) B. E. Griffin, C. B. Reese, *Tetrahedron Lett.* **1964**, *5*, 2925.

b) S. A. Scaringe, F. E. Wincott, M. H. Caruthers, *J. Am. Chem. Soc.* **1986**, *120*, 11820.

[53] a) K. K. Ogilvie, K. L. Sadana, E. A. Thompson, M. A. Quilliam, J. B. Westmore, *Tetrahedron Lett.* **1974**, *15*, 2861.

b) N. Usman, K. K. Ogilvie, M.-Y. Jiang, R. J. Cedergren, *J. Am. Chem. Soc.* **1987**, *109*, 7845.

[54] a) S. Pitsch, P. A. Weiss, L. Jenny, A. Stutz, X. Wu, *Helv. Chim. Acta* **2001**, *84*, 3773.

b) X. Wu, S. Pitsch, *Nucleic Acids Res.* **1998**, *26*, 4315.

c) S. Pitsch, P. A. Weiss, X. Wu, D. Ackermann, T. Honegger, *Helv. Chim. Acta* **1999**, *82*, 1753.

[55] a) R. L. Letsinger, J. L. Finnan, G: A. Heavner, W. B. Lundsford*, J. Am. Chem. Soc.* **1975**, *97*, 3278.

b) M. H. Caruthers, A. D. Barone, S. L. Beaucage, D. R. Dodds, E. F. Fisher, L. J. McBride, M. Matteucci, Z. Stabinsky, J. Y. Tang, *Methods Enzymol.* **1987**, *154*, 287.

[56] M. P. Reddy, N. B. Hanna, F. Farooqui, *Tetrahedron Lett.* **1994**, *35*, 4311.

[57] a) P. C. Zamecnik, M. L. Stephenson, *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 280.

b) E. Uhlmann, A. Peyman, *Chem. Rev.* **1990**, *90*, 543.

c) D. D. Ma, T. Rede, N. A. Naqui, P. D. Cook, *Biotechnol. Annu. Rev.* **2000**, *5*, 155.

[58] B. A. Sullunger, E. Gilboa, *Nature* **2002**, *418*, 252.

[59] S.-H. Kang, E. L. Zirbes, R. Kole, *Antisense Nucl.Acid Drug Dev.* **1999**, *9*, 497.

[60] R. L. Letsinger, G. Zhang, D. K. Sun, T. Ikeuchi, P. S. Sarin, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 6553.

[61] a) M. Lemaitre, B. Bayard, B. Lebleu, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 648.

b) J. P. Bongartz, A. M. Aubertin, P. G. Milhaud, B. Lebleu, *Nucleic. Acids. Res.* **1994**, *22*, 4681.

c) S. B. Rajur, C. M. Roth, J. R. Morgan, M. L. Yarmush, *Bioconjugate Chem.* **1997**, *8*, 935.

[62] J.-P. Leonetti, G. Degols, B. Lebleu, *Bioconjugate Chem.* **1990**, *1*, 149.

[63] Z. Wie, C.-H. Tung, T. Zhu, W. A. Dickenhof, K. J. Breslauer, D. E. Georgopoulos, M. J. Leibowitz, S. Stein, *Nucleic Acids Res.* **1996**, *24*, 655.

[64] a) J. Goodchild, *Bioconjugate Chem.* **1990**, *1*, 165.

b) J. Nielsen, S. Brenner, K. D. Janda, *J. Am. Chem. Soc.* **1993**, *115*, 9812.

c) J. Matulic-Adamic, L. Beigelman, L. W. Dudycz, C. Gonzalez, N. Usman, *Bioorg. Med. Chem. Lett.* **1995**, *5*, 2721.

d) J. Olejnik, H.-C. Lüdemann, E. Krzymanska-Olejnik, S. Berkenkamp, F. Hillenkamp, K. J. Rothschild, *Nucleic Acids Res*. **1999**, *27*, 4626.

e) L. De Napoli, A. Messere, D. Montesarchio, G. Piccialli, E. Benedetti, E. Bucci, F. Rossi, *Bioorg. Med. Chem.* **1999**, *7*, 395.

f) V. Marchán, C. Rodríguez-Tanty, M. Estrada, E Pedroso, A. Grandas, *Eur. J. Org. Chem.* **2000**, 2495

g) A. Sakakura, Y. Hayakawa, *Tetrahedron* **2000**, *56*, 4427.

h) D. A. Stetsenko, A. A. Arzumanov, V. A. Korshun, M. J. Gait, *Molecular Biology* **2000**, *34*, 852.

i) M. de Champdoré, L. De Napoli, G. Di Fabio, A. Messere, D. Montesarchio, G. Piccialli, *Chem. Commun.* **2001**, 2598.

j) D. Forget, D. Boturyn, E. Defrancq, J. Lhomme, P. Dumy, *Chem. Eur. J.* **2001**, *7*, 3976.

k) T. Kubo, K. Dubey, M. Fujii, *Nucleosides, Nucleotides & Nucleic Acids* **2001**, *20*, 1321.

l) D. Forget, O. Renaudet, D. Boturyn, E. Defrancq, P. Dumy, *Tetrahedron Lett.* **2001**, *42*, 9171.

m) E. R. Civitello, R. G. Leniek, K. A. Hossler, D. M. Stearns, *Bioconjugate Chem.* **2001**, *12*, 459.

n) M. Ballico, S. Drioli, F. Morvan, L. Xodo, G. M. Bonora, *Bioconjugate Chem.* **2001**, *12*, 719.

o) T. S. Zatsepin, D. A. Stetsenko, A. A. Arzumanov, E. A. Romanova, M. J. Gait, T. S. Oretskaya, M. *Bioconjugate Chem.* **2002**, *13*, 822.

p) Z. J. Gartner, M. W. Kanan, D. R. Liu*, J. Am. Chem. Soc.* **2002**, *124*, 10304.

q) M. Adinolfi, L. De Napoli, G. Di Fabio, A. Iadonisi, D. Montesarchio, G. Piccialli, *Tetrahedron* **2002**, *58*, 6697.

[65] R. Eritja, A. Pons, M. Escarceller, E. Giralt, F. Albericio, *Tetrahedron* **1991**, *47*, 4113.

[66] C.-H. Tung, M. J. Rudolph, S. Stein, *Bioconjugate Chem.* **1991**, *2*, 464.

[67] a) K. Arar, M. Monsigny, R. Mayer, *Tetrahedron Lett.* **1993**, *34*, 8087.
b) B. G. de la Torre, F. Albericio, E. Saison-Behmoaras, A.Bachi, R. Eritja, *Bioconjugate Chem.* **1999**, *10*, 1005.

[68] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776.

[69] M. McPherson, M. C. Wright, P. A. Lohse, *Synlett* **1999**, *S1*, 978.

[70] a) D. A. Stetsenko, M. J. Gait, *J. Org. Chem.* **2000**, *65*, 4900.
b) D. A. Stetsenko, M. J. Gait, *Nucleosides & Nucleotides* **2001**, *20*, 801.

[71] a) S. Peyrottes, B. Mestre, F. Burlina, M. J. Gait, *Tetrahedron* **1998**, *54*, 12513.
b) S. Peyrottes, B. Mestre, F. Burlina, M. J. Gait, *Nucleosides & Nucleotides* **1999**, *18*, 1443.

[72] a) D. L. McMinn, T. J. Matray, M. M. Greenberg, *J. Org. Chem.* **1997**, *62*, 7074.
b) D. L. McMinn, M. M. Greenberg, *J. Am. Chem. Soc.* **1998**, *120*, 3289.
c) D. L. McMinn, M. M. Greenberg, *Bioorg. Med. Chem. Lett.* **1999**, *9*, 547.
d) J. D. Kahl, M. M. Greenberg, *J. Org. Chem.* **1999**, *64*, 507.

[73] a) E. M. Zubin, E. A. Romanova, E. M. Volkov, V. N. Tashlitsky, G. A. Korshunova, Z. A. Shabarova, T. S. Oretskaya, *FEBS Lett***. 1999**, *456*, 59.
b) J.-T. Hwang, M. M. Greenberg, *Org. Lett.* **1999**, *1*, 2021.
c) J.-T. Hwang, M. M. Greenberg, *J. Org. Chem.* **2001**, *66*, 363.
d) E. S. Krider, J. E. Miller, T. J. Meade, *Bioconjugate Chem.* **2002**, *13*, 155.

[74] a) B.G. de la Torre, A. Aviño, G. Tarrason, J. Piulats, F. Albericio, R. Eritja, *Tetrahedron Lett.* **1994**, *35*, 2733.
b) J. Robles, M. Beltrán, V. Marchán, Y. Perez, I. Travesset, E. Pedroso, A. Grandas, *Tetrahedron* **1999**, *55*, 13251.

[75] a) J. Haralambidis, L. Duncan, G. W. Tregear, *Tetrahedron Lett.* **1987**, *28*, 5199.
b) J. Haralambidis, L. Duncan, K. Angus, G. W. Tregear, *Nucleic Acids Res.* **1990**, *18*, 493.
c) J.-C. Truffert, O. Lorthioir, U. Asseline, N. T. Thuong, A. Brack, *Tetrahedron Lett.* **1994**, *35*, 2353.
d) S. Soukchareun, G. W. Tregear, J. Haralambidis, *Bioconjugate Chem.* **1995**, *6*, 43.
e) S. Soukchareun, J. Haralambidis, G. W. Tregear, *Bioconjugate Chem.* **1998**, *9*, 466.
f) C.-P. Chen, X.-X. Li, L.-R. Zhang, J.-M. Min, J. Y.-W. Chan, K.-P. Fung, S.-Q. Wang, L.-H. Zhang, *Bioconjugate Chem.* **2002**, *13*, 525.

[76] a) F. Bergmann, W. Bannwarth, *Tetrahedron Lett.* **1995**, *36*, 1839.
b) C. N. Tetzlaff, I. Schwope, C. F. Bleczinski, J. A. Steinberg, C. Richert, *Tetrahedron Lett.* **1998**, *39*, 4215.
c) I. Schwope, C. F. Bleczinski, C. Richert, *J. Org. Chem.* **1999**, *64*, 4749.
d) C. F. Bleczinski, C. Richert, *Org. Lett.* **2000**, *2*, 1697.

[77] a) J. Robles, E. Pedroso, A. Grandas, *Nucleic Acids Res.* **1995**, *23*, 4151.
b) M. Beltrán, M. Maseda, Y. Perez, J. Robles, E. Pedroso, A. Grandas, *Nucleosides & Nucleotides* **1997**, *16*, 1487.
c) J.Robles, M. Maseda, M. Beltrán, M. Concernau, E. Pedroso, A. Grandas, *Bioconjugate Chem.* **1997**, *8*, 785.
d) M. Beltrán, E. Pedroso, A. Grandas, *Tetrahedron Lett.* **1998**, *39*, 4115.

[78] a) C. D. Duby, C. D. Richardson, R. Brousseau, , *Tetrahedron Lett.* **1991**, *32*, 879.
b) M. Antopolsky, A. Azhayev, *Helv. Chim. Acta* **1999**, *82*, 2130.
c) D. A. Stetsenko, M. J. Gait, *Bioconjugate Chem.* **2001**, *12*, 576.
d) M. Antopolsky, E. Azhayeva, U. Tengvall, A. Azhayev, *Tetrahedron Lett.* **2002**, *43*, 527.
e) D. A. Stetsenko, A. D. Malakhov, M. J. Gait, *Org. Lett.* **2002**, *4*, 3259.

[79] T. H. Fraser, A. Rich, *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 2671.

[80] a) O. Botta, P. Strazewski, *Nucleosides & Nucleotides* **1999**, *18*, 721.
b) O. Botta, *PhD thesis*, University of Basel, **1999**.

[81] a) N. Q. Nguyen-Trung, O. Botta, S. Terenzi, P. Strazewski, *J. Org. Chem.* **2003**, *68*, 2038.
b) N. Q. Nguyen-Trung, *PhD thesis*, University of Basel, **2003**.

[82] B. R. Baker, J. P. Joseph, J. H. Williams, *J. Am. Chem. Soc.* **1955**, *77*, 1.

[83] a) R. Mengel, H. Wiedner, *Chem. Ber.* **1976**, *109*, 433.
b) M. C. Samano, M. J. Robins, *Tetrahedron Lett.* **1989**, *30*, 2329.
c) M. J. Robins, R. W. Miles, M. C. Samano, R. L. Kaspar, *J. Org. Chem.* **2001**, *66*, 8204.

[84] K. K. Ogilvie, S. L. Beaucage, A. L. Schifman, N. Y. Theriault, K. L. Sadana, *Can. J. Chem*. **1978**, *56*, 2768.

[85] W. L. Sung, S. A. Narang, *Can. J. Chem*. **1982**, *60*, 111.

[86] a) J. Garegg, B. Samuelsson, *Carbohydrate Res.* **1978**, *67*, 267.
b) F. Hansske, D. Madej, M. J. Robins, *Tetrahedron* **1984**, *40*, 125.
c) M. J. Robins, S. Sarker, V. Samano, S. F. Wnuk, *Tetrahedron* **1997**, *53*, 447.

[87] a) L. J. McBride, M. H. Caruthers, *Tetrahedron Lett.* **1983**, *24*, 2953.
b) L. J. McBride, R. Kierzek, S. L. Beaucage, M. H. Caruthers, *J. Am. Che. Soc.* **1986**, *108*, 2040.

[88] B. C. Froehler, M. D. Matteucci, *Nucleic Acids Res.* **1983**, *11*, 8031.

[89] a) A. M. Ozols, A. V. Azhayev, N. B. Dyatkina, A. A. Krayevsky, *Synthesis* **1980**, 557.
b) M. J. Robins, S. D. Hawrelak, A. E. Hernández, S. F. Wnuk, *Nucleosides & Nucleotides* **1992**, *11*, 821.

[90] M. Vaultier, N. Knouzi, R. Carrié, *Tetrahedron Lett.* **1983**, *24*, 763.

[91] A. K. Saha, P. Schultz, H. Rapoport, *J. Am. Chem. Soc.* **1989**, *111*, 4856.

[92] M. Bartra, P. Romea, F. Urpí, J. Vilarrasa, *Tetrahedron* **1990**, *46*, 587.

[93] M. P. Reddy, J. B. Rampal, S. L. Beaucage, *Tetrahedron Lett*. **1987**, *28*, 23.

[94] P. Kumar, N. N. Gosh, K. L. Sadana, B. S. Garg, K. C. Gupta, *Nucleosides & Nucleotides* **1993**, *12*, 565.

[95] M. D. Matteucci, M. H. Caruthers, *J. Am. Chem. Soc.* **1981**, *103*, 3185.

[96] C. McCollum, A. Andrus, *Tetrahedron Lett.* **1991**, *32*, 4069.

[97] J. Katzhendler, S. Cohen, E. Rahamin, M. Weisz, I Ringel, J. Deutsch, *Tetrahedron* **1989**, *45*, 2777.

[98] S. Gunzenhauser, E. Biala, P. Strazewski, *Tetrahedron Lett*. **1998**, *39*, 6277.

[99] T. Brown, C. E. Pritchard, G. Turner, S. A. Salisbury, *J. Chem. Soc. Chem. Commun.* **1989**, 891.

[100] a) R. B. Merrifield, J. Singer, B. T. Chait, *Anal. Biochem.* **1988**, *174*, 399.
b) R. C. de L. Milton, S. C. F. Milton, P. A. Adams, *J. Am. Chem. Soc.* **1990**, *112*, 6039.
c) S. E. Blondelle, B. Forood, R. A. Houghten, E. Pérez-Payá, *Biochemistry* **1997**, *36*, 8393.
d) R. Warrass, J.-M. Wieruszeski, C. Boutillon, G. Lippens, *J. Am. Chem. Soc.* **2000**, *122*, 1789.
e) J. C. Hendrix, K. J. Halverson, J. T. Jarret, P. T. Lansbury, *J. Org. Chem.* **1990**, *55*, 4517.
f) B. D. Larsen, A. Holm, *Int. J. Pept. Protein Res.* **1994**, *43*, 1.

[101] L. A. Carpino, *J. Am. Chem. Soc.* **1993**, *115*, 4397.

[102] a) C.-X. Fan, X.-L. Hao, Y.-H. Ye, *Synthetic Comm.* **1996**, *26*, 1455.

b) H. Li, X. Jiang, Y.-H. Ye, C. Fan, T. Romoff, M. Goodman, *Org. Lett.* **1999**, *1*, 91.

[103] S. C. Story, J. V. Aldrich, *Int. J. Pept. Protein Res.* **1994**, *43*, 292.

[104] a) R. B.Merrifield, A. R. Mitchell, J. E. Clarke, *J. Org. Chem.* **1974**, *39*, 660.
b) M. Bodanszky, J. Martinez, *Synthesis* **1981**, *5*, 333.

[105] R. Ramage, J. Green, *Tetrahedron Lett.* **1987**, *28*, 2287.

[106] a) A. Percot, X. X. Zhu, M. Lafleur, *Biopolymers* **1999**, *50*, 647.
b) R. N. A. H. Lewis, Y.-P. Zhang, R. S. Hodges, W. K. Subczynski, A. Kusumi, C. R. Flach, R. Mendelsohn, R. N. McElhaney, *Biochemistry* **2001**, *40*, 12103.

[107] S. Vauthey, S. Santoso, H. Gong, N. Watson, S. Zhang, *Proc. Natl. Acad. Sci. USA* **2001**, *53*, 555.

[108] M. Dettin, S. Pegoraro, P. Rovero, S. Bicciato, A. Bagno, C. Di Bello, *J. Peptide Res.* **1997**, *49*, 103.

[109] J. Bedford, C. Hyde, T. Johnson, W. Jun, D. O. M. Quibell, R. C. Sheppard, *Int. J. Peptide Protein Res.* **1992**, *40*, 300.

[110] a) M. Karas, F. Hillenkamp, *Anal. Chem.* **1988**, *60*, 1319.
b) K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida, *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151.

[111] a) R. M. Caprioli (ed.), Kluwer Academic Publishers, *Mass Spectrometry in Biomolecular Sciences* **1996**, 33-49.
b) K. Dreisewerd, *Chem. Rev.* **2003**, *103*, asap.

[112] O. N. Jensen, S. Kulkarni, J. V. Aldrich, D. F. Barofsky, *Nucleic Acids Res.* **1996**, *24*, 3866.

[113] J. G. Harrison, S. Balasubramanian, *Nucleic Acids Res.* **1998**, *26*, 3136.

[114] a) S. Carpasso, L. Mazzarella, F. Sica, A. Zagari, S. Salvadori, *J. Chem. Soc. Chem. Comm.* **1992**, 919.
b) R. Dölling, M. Beyermann, J. Haenel, F. Kernchen, E. Krause, P. Franke, M. Brudel, M. Bienert, , *J. Chem. Soc. Chem. Comm.* **1994**, 853.
c) J. L. Lauer, C. G. Fields, G. B. Fields, *Letters in Peptide Science* **1994**, *1*, 197.
d) Y. Yang, W. V. Sweeney, K. Schneider, S. Thörnqvist, B. T. Chait, J. P. Tam, *Tetrahedron Lett.* **1994**, *35*, 9689.

[115] W. C. Chang, B. W. Bycroft, D. J. Evans, P. D. White, *J. Chem. Soc. Chem. Comm.* **1995**, 2209.

[116] a) H. Waldmann, H. Kunz, *Liebigs Ann. Chem.* **1983**, 1712.

b) A. Trzeciak, W. Bannwarth, *Tetrahedron Lett.* **1992**, *33*, 4557.

c) n. Thieret, J. Alsina, E. Giralt, F. Guibé, F. Albericio, *Tetrahedron Lett.* **1997**, *38*, 7275.

d) F. Guibé, *Tetrahedron* **1998**, *54*, 2967.

[117] T. Jonhson, M. Liley, T. J. Cheeseright, F. Begum, *J. Chem. Soc. Perkin Trans. 1* **2000**, 2811.

[118] Y. Hayakawa, S. Wakabayashi, H. Kato, R. Noyori, *J. Am. Chem. Soc.* **1990**, *112*, 1691.

[119] a) D. Stanojevic, G. L. Verdine, *Nature Struct. Biol.* **1995**, *2*, 450.

[120] D. M. Gray, *Circular Dichroism: Principles and Applications, Second Edition*, Ed. by Nina Berova, Koji Nakanishi, Robert W. Woody, J. Wiley & Sons **2000**, Chapt. 27.

[121] W. C. Johnson, *Circular Dichroism: Principles and Applications, Second Edition*, Ed. by Nina Berova, Koji Nakanishi, Robert W. Woody, J. Wiley & Sons **2000**, Chapt. 24.

[122] a) P. Strazewski, E. Biala, K. Gabriel, W. H. McClain, *RNA* **1999**, *5*, 1490.

b) E. Biala, P. Strazewski, *J. Am. Chem. Soc.* **2002**, *124*, 3540.

[123] P. W. Wallimann, R. J. Kennedy, J. S. Miller, W. Shalongo, D. S. Kemp, *J. Am. Chem. Soc.* **2003**, *125*, asap.

[124] M. Sprinzl, *Biochem. Byophys. Res. Comm.* **1976**, *71*, 1025.

[125] A. Dorenbeck, M. Scheffler, M. Wüstefeld, G. von Kiedrowski, *Angew. Chem.* **2003**, in press.

[126] G. Binnig, C. F. Quate, C. Gerber, *Phys. Rev. Lett.* **1986**, *56*, 930.

[127] T. R. Albrecht, P. Grütter, D. Horne, D. Rugar, *J. Appl. Phys.* **1991**, *69*, 668.

[128] S. Fernanda-Lopez, H.-S. Kim, E. C. Choi, M. Delgado, J. R. Granja, A. Khasanov, M. R. Ghadiri, *Nature* **2002**, *412*, 452.