

# Conjugation and cyclization of polyamides derivatized with protected maleimides

Xavier Elduque Busquets



Aquesta tesi doctoral està subjecta a la llicència Reconeixement- NoComercial – SenseObraDerivada 3.0. Espanya de Creative Commons.

Esta tesis doctoral está sujeta a la licencia Reconocimiento - NoComercial – SinObraDerivada 3.0. España de Creative Commons.

This doctoral thesis is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0. Spain License.

# **Conjugation and cyclization of polyamides derivatized with protected maleimides**

Manuscript submitted by:

**Xavier Elduque Busquets**

Directed and revised by:

**Prof. Anna Grandas Sagarra**

Departament de Química Orgànica

Programa de Doctorat de Química Orgànica

Universitat de Barcelona, Juliol de 2014



*Satisfaction lays in the effort, not in the attainment,  
full effort is full victory*

*Mahatma Gandhi*

*Try not. Do... or do not. There is no try*

*Yoda*



## Agraïments

Antes de entrar en lo personal, quiero agradecer al Ministerio de Educación, Cultura y Deporte la concesión de la beca FPU para la realización de la tesis, y el suplemento para la estancia breve.

Aquesta tesi s'ha pogut dur a terme gràcies a moltes persones i entitats. Primer, gràcies a l'Anna, per dirigir-me la tesi, sin prisa pero sin pausa, sent sempre propera i accessible. També a l'Enrique, per acollir-me al grup i supervisar determinades parts de la feina. L'agraïment és extensiu també al Vicente, el Jordi i la Núria. Fora del grup, voldria agrair també a l'Ernesto, per ajudar-me en els primers passos i encoratjar-me a canviar de grup.

I'd also like to thank Prof. Engels for allowing me into his Arbeitskreis in Frankfurt, and Olga and Jasmin for helping me in the lab. Also Prof. Per Lincoln for all the help during my stay in Göteborg, and Graciela, Stefan, Anna and Bobo for easing the chemistry and the life in Sweden. Tack så mycket!

Voldria mencionar també els grups i entitats que ens han ajudat, arribant allí on nosaltres no arribem: David Sunyol, i al servei de masses de l'IRB, pels estudis d'estequiometria, al grup de la Dra. Martínez-Salas per les proves *in vitro* amb AD2, i al grup del Dr. Valcárcel per les proves amb els pèptids cíclics. També a la Irene i la Laura, de masses, per l'ajuda amb ESIs i MALDIs.

També voldria agrair a l'Albert, per obrir camí amb les maleïmides. Als altres membres d'oligos, Clément, Omar, Natàlia i Àlex, per fer un ambient de treball excel·lent i estar sempre disposats a ajudar. Ànims amb la tesi, i sort amb la feina. Vull fer l'agraïment extensiu als ex-oligos, doctorands, post-docs, alumnes i màsters que han anat passant pel grup, i a membres d'altres grups, com l'Aleix, o la gent de pèptids o del fut.

Molt especialment vull agrair a les nenes del 500, Ana i Eli, per tota l'ajuda, ànims i suport. Dues grandíssimes persones que han fet millors aquests anys dins i fora del lab.

Vull incloure a la gent de fora de l'ambient acadèmic, que han ajudat sense saber-ho a que tot anés endavant i a no perdre el món de vista, com la Lili, l'Arnau, la Roser, i la gent del JCT en general.

Per acabar, vull agrair a la Cris, als meus pares i al meu germà per tot l'anterior, i més. Pel seu suport, paciència i ànims incondicionals. I per fer-me els tupper.

Gràcies a tots



# Index

About the structure and contents of this work

Abbreviations

## PART A

1. Introduction and Goals	1
1.1. Why do chemists synthesize peptides and oligonucleotides?	1
1.2. Tuning compound stability: Introducing backbone modifications	2
1.3. Cyclization of biomolecules	4
1.4. Conjugation of biomolecules: Better together	6
1.5. Cyclization and conjugation strategies	8
1.6. Maleimides as useful moieties for both cyclization and conjugation reactions	12
1.7. Goals	13
2. Results and Discussion	15
2.1. Synthesis of maleimide-containing monomers	15
2.2. Conjugation	18
Single conjugations. Deprotection optimization and conjugate stability assessment	18
Double conjugations	22
2.3. Cyclization	24
Monocycle formation	24
Bicyclic peptides	29
2.4. Conjugated cyclic peptides: Proof of concept	32
2.5. Cyclic peptide conjugates as potential splicing modulators	42
3. Concluding Remarks	45
4. References and Notes	47

## PART B

5. Introduction and Goals	55
5.1. RNA: Structure, function and pharmacological potential	55
5.2. Viral IRES: A target for foot-and-mouth disease	57
5.3. AD2 as a GNRA binder	59
5.4. Goals	59
6. Results and Discussion	61
6.1. AD2 synthesis and characterization, and RNA synthesis	61



6.2. <i>In vitro</i> results	63
6.3. Absorption spectroscopy assays	65
6.4. Isothermal titration calorimetry assays	68
6.5. Fluorescence assays	69
6.6. Stoichiometry determination by mass spectrometry	71
7. Concluding Remarks	75
8. Experimental section	77
9. References and Notes	83
APPENDIXES	
Appendix 1: Equivalence table	87
Appendix 2: Experiments and compounds not included in the papers	89
Appendix 3: Papers	101
Contribution Report	101
Bioconjugate Chemistry	103
Organic Letters	141
Journal of Organic Chemistry	165
Appendix 4: Materials and Methods	193
A4.1. Reagents	193
A4.2. Instruments	193
A4.3. Methods	194
Appendix 5: AD2 and quinacrine binding to ct-DNA	199
Linear dichroism experiments	199
Fluorescence qualitative results. Quinacrine and AD2 displacement	201
Data processing with the McGhee-Von Hippel approach	202
Summary	204
Experimental section	204
References and Notes	205
Resum en Català	207

## **About the structure and contents of this work**

The work hereby presented is divided in two separate parts that correspond to two different projects, even though the thesis title only reflects the first. Both projects have been developed in parallel.

Part A is a methodological work in the field of conjugation and cyclization of polyamides. It is the main part of this thesis, and has yielded published results. It consists of an introduction, a discussion of the results, the final remarks and the reference collection.

Part B is shorter and corresponds to a completely different topic: the synthesis and evaluation of a small molecule as an RNA ligand. Even though the results have not been optimal it has been very formative, and we have decided to include it in this thesis. It also contains an introduction, a discussion of the results, the final remarks and the reference collection, and includes an experimental section at the end.

In this manuscript, five appendices are included. The first three appendices are related with Part A. The first is an equivalence table, which correlates the compound numbers of the thesis with the numbers of the publications. The second includes experiments performed within this work not reported in any publication. The third contains the published papers related with this work. The fourth appendix describes the materials, instruments and methodology employed for both Part A and Part B. Some of this information is directly referred to the publications in Appendix 3.

Finally, Appendix 5 is related to the work of the Part B and details the work carried out during a short stay at the Chalmers University of Technology in Gothenburg.



## Abbreviations

Ac	Acetyl
Acm	Acetamidomethyl
ACN	Acetonitrile
AMP	Adenosine monophosphate
Ap	2-Aminopurine
AZT	3'-Azido-3'-deoxythymidine
Boc	<i>N-tert</i> -Butoxycarbonyl
bs	Broad signal
CA	Ammonium citrate
CAT	Chloramphenicol acetyltransferase
CD	Circular dichroism
CPG	Controlled pore glass
CPP	Cell-penetrating peptide
ct-DNA	Calf thymus DNA
d	Doublet
DHB	2,5-Dihydroxybenzoic acid
DIAD	Diisopropyl azodicarboxylate
DMAP	4-Dimethylaminopyridine
DMF	<i>N,N</i> -dimethylformamide
DMSO	Dimethylsulfoxide
DMT	4,4'-Dimethoxytrityl
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EC <sub>50</sub>	Half maximal effective concentration
ESI	Electrospray ionization
Fm	Fluorenylmethyl

FMDV	Foot-and-mouth disease virus
Fmoc	Fluorenylmethoxycarbonyl
Glut	Glutathione
HATU	1-[Bis(Dimethylamino)methylene]-1 <i>H</i> -1,2,3-triazolo[4,5- <i>b</i> ]pyridinium 3-oxid hexafluorophosphate
HOAt	1-Hydroxy-7-azabenzotriazole
HOBt	1-Hydroxybenzotriazole
HPLC	High-performance liquid chromatography
HRMS	High-resolution mass spectrometry
IEDDA	Inverse electron-demand Diels-Alder
IM-MS	Ion-mobility mass spectrometry
IRES	Internal ribosomal entry site
ITC	Isothermal titration calorimetry
LD	Linear dichroism
LNA	Locked nucleic acid
lncRNA	Long non-coding RNA
LUC	Luciferase
m	Multiplet
Mal	Maleimide
MALDI	Matrix-assisted laser desorption/ionization
MRI	Magnetic resonance imaging
mRNA	Messenger RNA
MW	Microwave
NMP	1-Methyl-2-pyrrolidone
NMR	Nuclear magnetic resonance
Pac	Phenoxyacetyl
Pbf	2,2,3,6,7-Pentamethyldihydrobenzofuran-5-sulfonyl
PCR	Polymerase chain reaction

PDA	Photodiode array
PNA	Peptide nucleic acid
ProtMal	Protected maleimide
Rf	Retardation factor
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
RNase	Ribonuclease
s	Singlet
SHAPE	Selective 2'-hydroxyl acylation analyzed by primer extension
siRNA	Small interference RNA
snRNA	Small nuclear RNA
SPPS	Solid phase peptide synthesis
TBDMS	<i>tert</i> -Buthyldimethylsilyl
TBTA	Tris[(1-benzyl-1 <i>H</i> -1,2,3-triazol-4-yl)methyl]amine
<sup>t</sup> Bu	<i>tert</i> -Buthyl
TCEP	Tris(2-carboxyethyl)phosphine
TEA	Triethylamine
TEAA	Triethylammonium acetate
TFA	Trifluoroacetic acid
THAP	2',4',6'-Trihydroxyacetophenone
THF	Tetrahydrofuran
TIS	Triisopropylsilane
T <sub>m</sub>	Melting temperature
TOF	Time of flight
t <sub>R</sub>	Retention time
tRNA	Transfer RNA

Trt	Trityl
UNA	Unlocked nucleic acid
UV-Vis	Ultraviolet visible

## Nucleosides, PNA monomers and amino acids

Nucleobases (and nucleosides)	Code	PNA monomers	Code
Adenine	A	Aminoethylglycine-adenine	a
Cytosine	C	Aminoethylglycine-cytosine	c
Guanine	G	Aminoethylglycine-guanine	g
Thymine	T	Aminoethylglycine-thymine	t
Uracil	U		

Amino acids	Three letter code	One letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V
Ornithine	Orn	

# **PART A**





## 1. Introduction and Goals

### 1.1. Why do chemists synthesize peptides and oligonucleotides?

Peptides and oligonucleotides are natural oligomers composed of amino acids and nucleotides, respectively. Their relevance in metabolic processes is beyond question, for they have all possible functions from structural to regulatory, and actively play a role in almost all processes of life.

Peptides can be used to address different biological problems. As an illustration, peptides can be used as vaccines<sup>1</sup> or for structure-activity relationship studies.<sup>2</sup> They can also be used to mimic certain substructures of high molecular weight proteins or complex molecules. Furthermore, peptides can also be extremely helpful for the study of complex biological systems, like protein-protein interactions.<sup>3</sup>

Peptides are relevant in the chemical industry. A huge amount of generic peptides, like somatostatin or leuprelide, are synthesized at the kilogram scale, and there are several companies that produce customized peptides upon request.<sup>4</sup>

Just a few natural peptides have succeeded as drugs. It is the case of cyclic peptides like cyclosporines, with immunosuppressant activity, antibiotics like polymyxins and bacitracin or octeotide, an hormone inhibitor. Gramicidin D is an example of linear peptide drug.<sup>5</sup> Also remarkable is the biological activity of cyclotides, cyclic peptides with internal disulfide knots. One of the best known is Kalata, which can be extracted from a plant and facilitates childbirth.<sup>6</sup> Peptides mechanism of action varies from disrupting the bacterial membrane to direct interaction with a receptor.

Synthetic oligonucleotides are largely employed as primers for PCR.<sup>7</sup> The use of oligonucleotides in the development of supramolecular structures,<sup>8</sup> nanomaterials<sup>9</sup> and chips<sup>10</sup> has also been developed recently. Also, their extraordinary molecular recognition capacity provides oligonucleotides with a great potential as drugs. Some examples of the different

## 1. Introduction and Goals

approaches for the employment of oligonucleotides as drugs are antisense therapy, anti-gene therapy, and small interference RNA (siRNA) and microRNA mimics.

Antisense therapy consists in introducing into a cell an oligonucleotide whose sequence is complementary to an mRNA. Both chains hybridize, and the heteroduplex most often promotes destruction of the mRNA by RNase H, thus avoiding translation of the mRNA.<sup>11</sup> Contrarily, anti-gene-strategy designed oligonucleotides aim at DNA double chains and place themselves in the major groove, forming Hoogsteen type bonds and affecting transcription processes.<sup>12</sup> siRNAs and microRNAs are small RNA duplexes that are incorporated into a multiprotein complex (RISC) that activates RNases in presence of fully complementary mRNA sequences.<sup>13,14</sup> Researchers hold high expectations for siRNA because of their great potential in gene silencing therapy.<sup>15,16</sup>

Even though these strategies could seem potentially useful for all kinds of disease, the outcome has not matched the expectative, since very few have been approved as drugs.

The problem when using natural polymers such as peptides or oligonucleotides for drug development is their poor bioavailability. This means that even though they are active and selective, they are easily degraded *in vivo* and they are not properly transported through the membranes. This is not surprising since these big biomolecules do not match Lipinski's rule of 5, a list of ideal molecular features to render good pharmacokinetic properties.<sup>17</sup> According to this rule, no small peptide or oligonucleotide would be an appropriate drug candidate.

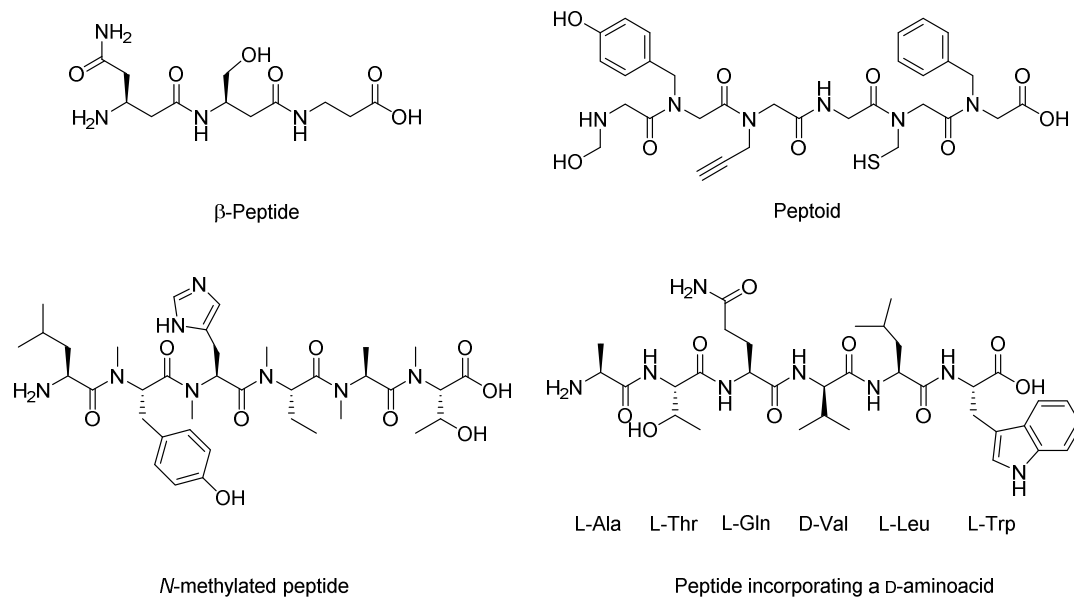
In order to address this problem, biomolecules can be modified in their backbone or conjugated.

### 1.2. Tuning compound stability: Introducing backbone modifications

Most modifications toward better stability are performed in the backbone. The main reason is to leave the sequence unaltered, in order not to disturb the recognition by the target. Also, enzymes like *endo* and *exo* proteases and nucleases that degrade oligomeric biomolecules, in order to reuse or recycle monomers, usually interact with the backbone and hydrolyze it. Another way of backbone modification is cyclization, but this will be discussed later.

Even though the changes in the structure may positively affect the activity and selectivity of the biomolecules, these changes may interfere with recognition by enzymes.<sup>18,19</sup>

A lot of modifications have been made in order to improve peptide stability, absorption properties and even activity. These new polymers are called peptidomimetics. As examples,  $\beta$ -peptides, peptoids, and the use of *N*-methylated or *D*-amino acids are briefly explained (**Fig 1.2.1**).



**Figure 1.2.1:** Structures of a  $\beta$ -peptide, a peptoid, an *N*-methylated peptide and a peptide incorporating a D-amino acid

$\beta$ -peptides are peptides formed by  $\beta$ -amino acids, in which the amine group is bound to the  $\beta$ - instead of the  $\alpha$ -carbon.  $\beta$ -amino acids can be found in nature<sup>20</sup> and they show high proteolytic resistance, so their use is a common strategy to avoid enzymatic degradation.  $\beta$ -peptide synthesis on a solid phase is straightforward, since  $\beta$ -amino acids are expensive but commercially available.<sup>21</sup>

In the case of peptoids the distance between side chains is not altered in comparison to natural peptides. The difference is that in peptoids the side chain is bound to the amine instead of the  $\alpha$ -carbon, so the building blocks are no longer chiral. This modification provides the biomolecule with a greater resistance to enzymes and a higher flexibility, due to the fact that the *cis-trans* exchange energy barrier is lower for tertiary than for secondary amides.<sup>22</sup> This flexibility can be an advantage if the peptoid is to be cyclized.<sup>23</sup> Their synthesis can be carried out in different ways on a solid support, and will be further discussed.

Peptides with *N*-methylated amino acids have the advantage of a greater flexibility and resistance to enzymes like peptoids, but the chirality of the building blocks and the location of the side chains remain unaltered. They can be obtained either by employing *N*-methylated amino acids, which are commercially available, or by a methylation reaction during peptide elongation in SPPS.<sup>24,25</sup>

D-amino acids are found in nature and their introduction in a sequence does not represent a change in the connectivity of the different atoms, although the chirality change of a single stereocenter may entail an important variation in the structure and thus a huge improvement of the activity.<sup>26</sup> D-Amino acids are common building blocks for the synthesis of modified peptides in SPPS and are also commercially available.

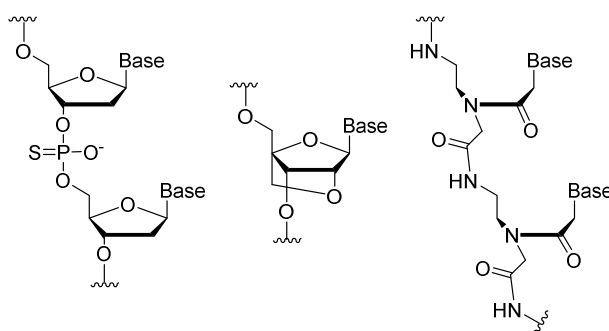
In the case of oligonucleotides, three examples of oligonucleotide analogs are briefly summarized (**Fig 1.2.2**). An interesting modification can be performed on the phosphate moiety, by converting it into a phosphorothioate.<sup>27</sup> For this purpose, the oxidation step of

## 1. Introduction and Goals

phosphite to phosphate in standard synthesis procedures has to be replaced by a sulfurization reaction. Despite the fact that sulfurization will create a new chiral center, and consequently the product obtained will be a mixture of diastereomers, this modification has proven to be useful for drug development.

Modifications at the sugar moiety have also been developed. By way of illustration, locked nucleic acids (LNAs) are oligonucleotides containing 2'-O,4'-C-methylene- $\beta$ -D-ribofuranose.<sup>28</sup> These oligonucleotides can be synthesized with a regular oligonucleotide synthesizer since monomers are commercially available. Toward the same goal but in a completely different strategy, unlocked nucleic acids (UNAs), which have no bond between C2' and C3', have also been developed. An interesting comparison between them has been published recently.<sup>29</sup>

Another remarkable modification of oligonucleotide backbones are peptide nucleic acids (PNAs). PNAs are made up of (2-aminoethyl)glycines with nucleobases attached via methylenecarbonyl linkers to the secondary amines. This achiral structure provides a distance of 6 bonds between the amines where nucleobases are attached, and 3 bonds between the amine and the nucleobase, mimicking natural oligonucleotides. PNAs have been widely studied for their tendency to hybridize with oligonucleotides with high affinity.<sup>30</sup> One of the main reasons for this high affinity is the absence of electrostatic repulsion between the backbones, the phosphate backbone being negatively charged and the PNA neutral.<sup>31</sup> They can be synthesized easily on a solid phase with the same methodology as for solid phase peptide synthesis (SPPS).



**Figure 1.2.2:** Structures of phosphorothioate oligonucleotides, LNAs and PNAs

Remarkably, the three oligonucleotides that have been approved by the FDA have some kind of backbone modification. Fomivirsen is a phosphorothioate oligonucleotide for retina inflammation, Mipomersen is also a phosphorothioate oligonucleotide that acts as a cholesterol reducing drug and Pegaptanib is an anti-angiogenic modified oligonucleotide.<sup>5</sup>

### 1.3. Cyclization of biomolecules

Cyclization is a different way of backbone modification also widely employed for activity, stability and absorption properties regulation.

By cyclic biomolecules (**Fig 1.3.1**) we refer to the compounds in which two monomers of a linear oligomer are bound. Cyclic compounds can be obtained by a head-to-tail cyclization, in

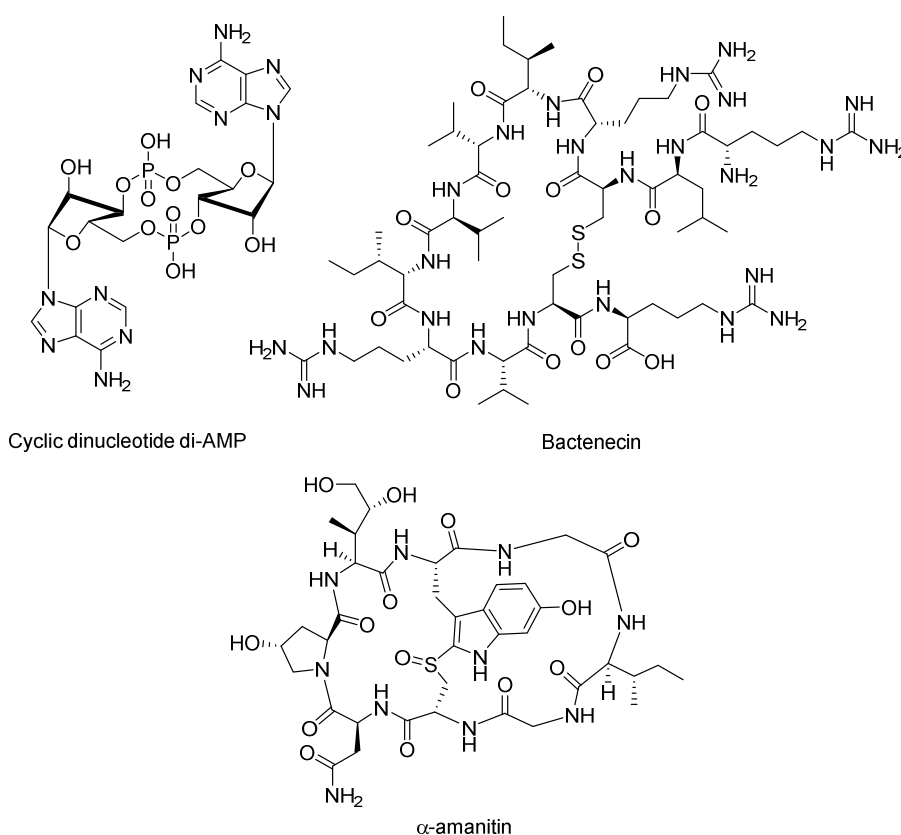
which one end of the oligomer reacts with the other end and thus incorporates all residues into the cycle, or through a head-to-side chain, tail-to-side chain or even side chain-to-side chain cyclization. In this case, one or two branches of linear oligomer append from the cycle.

Cyclic peptides are found in nature performing different and relevant activities, such as antifungal,<sup>32</sup> antimicrobial<sup>33</sup> or immunosuppressant.<sup>34</sup> They can be found in all kingdoms of life<sup>35</sup> and some are synthesized through non-ribosomal pathways.<sup>36</sup> For this reason, researchers and pharmaceutical companies have been attracted by these molecules.<sup>37</sup>

Peptide cyclization can also be employed to mimic certain protein substructures, simplifying the system and allowing for studies like protein folding or structure-activity relationships,<sup>38</sup> or for the construction of nanotubes by alternating D- and L-amino acids that self-assemble in  $\beta$ -sheet fashion.<sup>39</sup>

Cyclic nucleic acids can be found in nature. For example, bacteria usually have a single chromosome consisting in a large circular double stranded DNA.<sup>40</sup> Plasmids are commonly found as circular DNA double strands as well.<sup>41</sup> Furthermore, cyclic dinucleotides can also be found in some cellular pathways.<sup>42,43</sup> Synthetic cyclic oligonucleotides have also been extensively studied due to their potential as RNA and DNA ligands for gene silencing.<sup>44</sup>

By cyclizing a biomolecule two major improvements in their properties can be achieved: a better bioavailability and a better activity.<sup>45</sup> A recent review by Bock and coworkers<sup>46</sup> illustrates how structure constraints affect peptide bioavailability and activity.



**Figure 1.3.1:** Example of naturally occurring cyclic dinucleotide (di-AMP), a cyclic peptide (Bactenecin) and a bicyclic peptide ( $\alpha$ -amanitin)

## 1. Introduction and Goals

The action of exoproteases on peptides or exonucleases on oligonucleotides begins with the identification of an end of the biomolecule, but the ends are inexistent for head-to-tail cycles,<sup>47</sup> so the enzyme does not have a starting point and cannot act. It has also been reported that the activity of endoproteases and endonucleases in cycles is also affected, probably because the structure of the cycle is constrained and does not fit perfectly into the active site of the enzyme.<sup>48</sup>

When a biomolecule is cyclized the number of degrees of freedom is reduced in comparison to the linear oligomer. If the molecule is properly designed and by cyclizing it the functional groups responsible for the activity are suitably placed, the activity will be enhanced because the entropic penalty of the interaction with the target will decrease.<sup>49</sup>

Further cyclization of a macrocycle stabilizes it even more and can greatly improve its activity. Bicyclic peptides like  $\alpha$ -amanitin or sunflower trypsin inhibitor I are examples of constrained structures that can be found in nature. In nature there are also multicycles, with a definite pattern of disulfide bridges between their side chains.<sup>47</sup>

As an illustration of its applicability, the Heinis research group developed a bicyclic peptide that inhibited human urokinase-type plasminogen activator with an activity 200 fold higher than the best single macrocycle attempted.<sup>50</sup> To obtain even more constrained molecules, Dekan and coworkers managed to synthesize a multicyclic peptide with 4 different disulfide bridges, whose formation was directed by the use of cysteines with orthogonal protecting groups.<sup>51</sup>

### 1.4. Conjugation of biomolecules: Better together

Derived from the Latin word *conjugare*, which means “to link together”, in science we understand the word conjugation in many different ways. It can refer to the interaction of p-orbitals through  $\sigma$ -bonds, or to the transfer of genetic material between bacteria through direct contact. In this work, we use the word conjugation mainly to define the action of covalently binding two or more different molecules that usually have two different functions.

Closely related to conjugation is the term “prodrug”.<sup>52</sup> Prodrugs are inert molecules that are converted into active drugs through metabolic transformations *in vivo*. They are developed to enhance their bioavailability and selectivity, or to ease their administration avoiding problems such as taste, irritation or pain. Nevertheless, the frontier between prodrugs and conjugates is difficult to establish.

Common prodrugs show minor modifications related to the drug itself, like an alkylated amine or an ester instead of an acid or alcohol. As an illustration, Irinotecan is a prodrug consisting of an active molecule called SN-38 that inhibits topoisomerase I and is analogous to camptothecin, esterified with a biperidine carboxylic acid. This ester bond is hydrolyzed within the cell, thus releasing the active compound.<sup>53</sup>

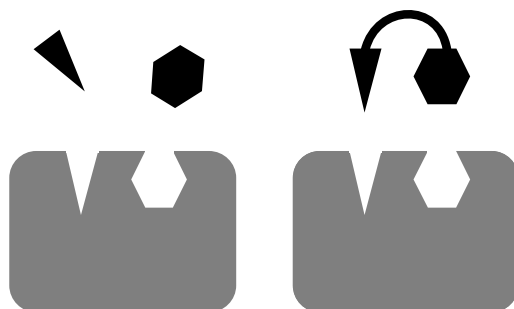
The different compounds that form the bioconjugate tend to have a unique specific function. Usually one of the compounds is a carrier, which allows for the introduction of the entire

bioconjugate into cells. The other compound is the one interacting with the target once released from the carrier. For example, Dasari and coworkers developed an anticancer prodrug by conjugating Gemcitabine with the Hoechst stain through a linker, reducing the toxicity and poor bioavailability of Gemcitabine. The drug was freed after reduction of a disulfide bond and hydrolysis of an amide bond.<sup>54</sup> Another anticancer prodrug was developed by conjugating taxol with hyaluronic acid. In this case taxol was released after an ester bond hydrolysis.<sup>55</sup> Finally, Wang and coworkers conjugated somatostatin to doxorubicin through two thioether bonds. Somatostatin acted as a carrier to tumor cells, due to the fact that somatostatin membrane receptors are overexpressed in such cells. Within the cell the thioether bond was cleaved by glutathione and the drug was thus released.<sup>56</sup>

Conjugation can be useful for improving drug bioavailability. For example, oligonucleotide stability can be improved through conjugation.<sup>57</sup> Conjugation can also facilitate cellular uptake, for example by linking together molecules with antitumoral activity and polysaccharides that act as carriers and target tumor cells.<sup>58,59</sup> Antibodies are also often conjugated for better specificity, either to anticancer drugs<sup>60</sup> or to protamines that complex siRNA.<sup>61</sup> Also very popular are cell penetrating peptides (CPPs), whose use as carriers of siRNA and antisense oligonucleotides is extensively described by Nakase and collaborators.<sup>62</sup>

Conjugation can also improve the affinity of the biomolecule in comparison to its single counterparts. As an illustration, consider a multivalent approach,<sup>63</sup> in which 2 different ligands can interact with different regions of a given receptor. If these molecules are bound in a way that the structure of the pharmacophores is unaltered, and can fit into the target without extra strain, the affinity of the conjugate will be much higher than the affinity of the individuals due to a decrease of the entropy penalty, aside from cooperation considerations (**Figure 1.4.1**).

Conjugation has even more applications, like imaging and diagnostic. That is the case of metal nanoparticle conjugation.<sup>64,65</sup> Metal nanoparticles can selectively signal specific targets for magnetic resonance imaging (MRI) or radioimaging when linked to biomolecules such as proteins, oligonucleotides, antibodies, etc.<sup>66,67</sup> Fluorescence dyes such as small molecules<sup>68</sup> or quantum dots<sup>69</sup> have also been employed for imaging purposes.



**Figure 1.4.1:** Schematic representation of a receptor with two binding sites and two individual ligands or a conjugate

Conjugation can also be employed for the synthesis of a variety of biocompatible materials, such as dendrimers or hydrogels among others,<sup>70,71</sup> and can also pave the way to the access to complex supramolecular structures. For example, the synthesis of C<sub>5</sub>-symmetric corannulenes



## 1. Introduction and Goals

functionalized with peptides, sugars and oligonucleotides for the assembly of icosahedral architectures was described by Mattarella and coworkers.<sup>72</sup>

### 1.5. Cyclization and conjugation strategies

Cyclization and conjugation strategies for biomolecules are similar, since both require two functional groups to selectively react with each other. The main difference is that while conjugations generally proceed better at high concentrations, cyclizations have to be carried out at high dilution or on a solid phase, under the so-called pseudodilution conditions, to avoid polymerization.<sup>73</sup> In the case of peptides, cyclizations are easier if the sequence incorporates turn-forming elements such as proline, glycine or D-amino acids.<sup>74</sup>

We divide the different synthetic strategies in two main groups, namely the reactions that yield natural linkages and the reactions that yield non-natural linkages.

#### Reactions yielding natural linkages

Some strategies for oligonucleotide and peptide conjugation or cyclization are designed to yield natural bonds, namely phosphate bonds, amides, esters or disulfide bridges.

Phosphate bonds (**Table 1.5.1, Entry 1**) link together the different nucleotides in oligonucleotides, and can be used for conjugation. For example, Pavlova and collaborators synthesized an oligonucleotide linked to an acridine compound at the 3' end via a phosphate bond, and a bleomycin, known to bind to the minor groove of DNA, at the 5' position via a phosphoramidate bond.<sup>75</sup> Peptides have also been conjugated through phosphate bonds, like in the case of nucleopeptides.<sup>76</sup> Phosphate bonds can also be employed for cyclization. For example, Kool reported a template-mediated enzymatic ligation for the circularization of a 56mer DNA oligonucleotide.<sup>77</sup>

Several methods generating amide bonds or ester bonds have been developed to obtain conjugates or cycles. In the case of oligonucleotide conjugation, amides can easily be generated in solution by reaction between an amine, which can be incorporated in the last step of oligonucleotide elongation by introducing a commercial phosphoramidite, and an active ester like a succinimido ester.<sup>66</sup> For peptide conjugation, amide bonds have been largely used, often taking advantage of the free amine of lysine side chains. Soudy and coworkers formed an amide when conjugating a peptide to doxorubicin through a linker.<sup>78</sup> Most of peptide cyclizations proceed through the attachment to the solid support of the first amino acid of the peptide through its side chain, followed by elongation of the peptide, deprotection of N- and C- ends and cyclization using standard ester or amide bond forming reagents. For example, for amide bond formation (**Table 1.5.1, Entry 2**) the most common strategy is the reaction between amines and carboxylic acids through activation of the acid group. This can be achieved with a carbodiimide and additives such as HOBt or HOAt, or with phosphonium salts like HATU.<sup>79</sup> Amide bond formation can be performed either on a solid phase<sup>80</sup> or in solution.<sup>81</sup> For ester formation (**Table 1.5.1, Entry 3**) the employment of DMAP/carbodiimide mixtures, N-hydroxysuccinimido ester amino acids, Mitsunobu or Yamaguchi reactions has been reported.<sup>82</sup>

Nevertheless, other methodologies of amide bond formation have been proposed that exploit thiol reactivity, for example the traceless Staudinger reaction or native chemical ligation.

In traceless Staudinger reaction (**Table 1.5.1, Entry 4**) an azide reacts with a phosphinothioether to generate a new amide bond. This reaction has been used in oligonucleotide conjugation,<sup>83</sup> peptide conjugation<sup>84</sup> and peptide cyclization. In cyclizations, the peptide must contain the azide moiety at the *N*-terminal position and a phosphinothiol has to be attached in solution at the *C*-terminus through a thioether bond after elongation and cleavage of the peptide. Then, in the presence of a base and a coupling agent, the new amide bond is produced.<sup>85</sup>

Native chemical ligation (**Table 1.5.1, Entry 5**) is a two-step reaction in which first a thioether undergoes a transthioesterification by a free thiol, and then the free amine of the thiol-containing residue displaces the thiol to irreversibly generate an amide bond, whose greater stability in comparison with the thioether is the driving force for the ultimate reaction.<sup>38</sup> As an illustration, Stetsenko conjugated a peptide with an oligonucleotide through native chemical ligation.<sup>86</sup> In the case of cyclization, the peptide usually incorporates the thioether at the *C*-terminus and a cysteine at the *N*-terminus.<sup>87</sup>

Disulfide bridges (**Table 1.5.1, Entry 6**) are the most common bonds for cyclic peptides in nature. Their ease of formation and their reversibility are often used to stabilize the 3D structure of the proteins.<sup>88</sup> With the employment of different orthogonal protecting groups for cysteines, disulfide bridges can be generated “à la carte”. In an oxidative environment, provided by molecular oxygen or iodine, disulfide bridges are fairly easy to create, but they are unstable in the presence of other free thiols.<sup>89</sup> These features make this bonding a good candidate for conjugation in prodrugs assembly.<sup>56</sup> Disulfide bridges have been used, for example, for the conjugation of peptide dendrimers with a PNA oligonucleotide analog.<sup>90</sup>

Disulfide bridges have also been employed for the synthesis of multicyclic peptides, by combining different orthogonal cysteine protecting groups. By stepwise deprotection of the cysteines, disulfide bond formation can be controlled.<sup>51</sup> An alternative when synthesizing natural multicycle compounds like cyclotides is to allow for peptide self-assembly. In this case, all the cysteines are deprotected at once, and the peptide folds directly into the most stable structure.<sup>91</sup> Nevertheless, this strategy often requires several purification steps.<sup>47</sup>

#### Reactions yielding non natural linkages

Chemists have also used other reactions for conjugation or cyclization to afford very complex molecules. The main difference between the reactions used to obtain natural and non-natural linkages is that, whereas the reactions yielding non-natural linkages are more chemoselective, they usually require the introduction of one or two “foreign” functional groups.

Some of these reactions are classified under the category of “click chemistry”. This concept was defined by Sharpless<sup>92</sup> in 2001 to refer to fast modular reactions being highly atom economic, high yielding, feasible in benign solvents like water, with a wide scope and with inoffensive by-products. Even though it is almost impossible for a reaction to match all these

## 1. Introduction and Goals

requirements, most reactions that will be explained below fit very well into the category of “click chemistry”.

“Click chemistry” has also become a useful tool for the generation of chemical libraries by combinatorial chemistry, which combined with high throughput screening allows for fast and systematic affinity assays of a massive number of different molecules.

Crossed metathesis (**Table 1.5.1, Entry 7**) has been used in cyclizations to replace disulfide bridges, producing a linkage with the same number of atoms. In the case of cyclic peptides two modified amino acids containing terminal double bonds have to be introduced in the sequence.<sup>48</sup>

Huisgen’s copper(I) catalyzed 1,3 dipolar cycloaddition (**Table 1.5.1, Entry 8**) is one of the most famous “click” reactions, and implies the formation of a 1,2,3-triazole ring from an alkyne and an azide. This reaction was developed in 1963 by Rolf Huisgen and was further improved with the introduction of the copper(I) catalysis, which confers the 1,4 regioselectivity, in 2002 independently by Sharpless<sup>93</sup> and Meldal,<sup>94</sup>. More recently, this reaction has been further developed by the introduction of ruthenium catalysis, which affords 1,5 disubstituted triazoles.<sup>95</sup>

The Huisgen cycloaddition is usually fast and clean, being this its main advantage, but has the drawback that two different functional groups have to be introduced in the respective molecules, and the azide may be incompatible with free thiols. Furthermore, both literature and the experience in our research group reveal that it requires fine tuning and is strongly substrate-dependent.<sup>96</sup> Nevertheless, this reaction has been widely used, for example for oligonucleotide conjugation<sup>97</sup> and cyclization,<sup>98</sup> or peptide conjugation<sup>99,100</sup> and cyclization.<sup>101</sup>

Another “click” reaction that has been recently used for peptide cyclization and conjugation is the radical thiol-ene reaction (**Table 1.5.1, Entry 9**). In this method a new thioether bond is produced through a radical mechanism. The requirements are a free thiol and a double bond.<sup>102</sup> Whereas a double bond may not be trivial to introduce in a peptide sequence, Aimetti and coworkers, as a matter of example, found a smart and practical solution by benefiting from an Alloc protecting group to obtain a side chain-to-side chain cyclic peptide.<sup>103</sup>

Oxime bond formation (**Table 1.5.1, Entry 10**) is also a “click chemistry” alternative for conjugation, the mild conditions of reaction being its main advantage. Nevertheless, the compounds afforded are easily hydrolyzed and thus not very stable. The Defranq research group, for example, managed to synthesize an oligonucleotide double conjugate by reacting an alkyne at the 5’ position with an azide, and an aldehyde at the 3’ position with an aminoxy-functionalized peptide or carbohydrate.<sup>97</sup> Oximes can also be employed for oligonucleotide cyclization,<sup>104</sup> peptide conjugation<sup>105</sup> and peptide cyclization.<sup>106</sup>

Michael addition (**Table 1.5.1, Entry 11**) is the generic name for the 1,4 nucleophilic additions to  $\alpha,\beta$ -unsaturated carbonyl compounds. This reaction was first developed in 1887 by Arthur Michael by reacting cinnamic acid ethyl ester with diethyl malonate in basic media.<sup>107</sup> This reaction is especially fast with thiols and electron deficient double bonds, allowing the double bond to react selectively with the thiol even in the presence of other nucleophiles in mild

conditions. Thanks to its chemoselectivity and versatility it has been employed in a wide range of applications.<sup>108</sup>

The Diels-Alder cycloaddition (**Table 1.5.1, Entry 12**) is probably one of the most famous cycloadditions in organic chemistry.<sup>109</sup> This reaction was developed in 1928 by Paul Hermann Diels and Kurt Alder, who would be awarded for it a Nobel Prize 22 years later. The traditional 2+4 cycloaddition proceeds through reaction between an electron-rich diene and an electron-deficient dienophile, even though the inverse electron-demand Diels-Alder reaction, with an electron-deficient diene and an electron-rich dienophile, has been extensively reported and studied lately.<sup>110</sup> The Diels-Alder reaction has also been widely used in conjugation due to its high chemoselectivity and the complete lack of byproducts.<sup>111</sup>

Entry	Component 1 (R <sub>1</sub> )	Component 2 (R <sub>2</sub> )	Conjugation reagent	Product
1		HO-R <sub>2</sub>	Enzyme	
2		H <sub>2</sub> N-R <sub>2</sub>	Carbodiimide, phosphonium or uronium salt	
3		HO-R <sub>2</sub>	Carbodiimide (or active ester intermediate)	
4		N <sub>3</sub> -R <sub>2</sub>	HS-CH <sub>2</sub> -PPh <sub>2</sub>	
5				
6	R <sub>1</sub> -SH	HS-R <sub>2</sub>	[Ox]	
7			Ruthenium catalyst	
8	R <sub>1</sub> -C≡C-	N <sub>3</sub> -R <sub>2</sub>	Cu(I)	
9	R <sub>1</sub> -SH		Catalyst, hv	
10		H <sub>2</sub> N-O-R <sub>2</sub>		
11	R <sub>1</sub> -SH			
12				

**Table 1.5.1:** Summary of conjugation and cyclization reactions. R<sub>1</sub> and R<sub>2</sub> are different molecules in case of conjugations and two parts of the same molecule in case of cyclizations

## 1. Introduction and Goals

### 1.6. Maleimides as useful moieties for both cyclization and conjugation reactions

In this work we have focused mainly in the two last “click” reactions: the Michael addition and the Diels-Alder cycloaddition.

Maleimides are commonly used moieties for conjugations, mainly via Michael addition but also via Diels-Alder cycloaddition with conjugated dienes. They have an electron deficient double bond in an  $\alpha,\beta$ -unsaturated carbonyl, which provides them with an excellent reactivity toward thiols via a Michael addition, and also with a high reactivity toward dienes via Diels-Alder reaction. As mentioned before, both Michael and Diels-Alder reactions fit perfectly into the concept of “click” reactions.

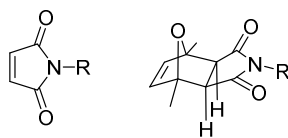
Maleimides are traditionally obtained by reaction of maleic anhydride with a primary amine, followed by dehydration of the resulting compound.<sup>112–114</sup> They have been extensively used in polymers,<sup>115</sup> as spacers for bioconjugation,<sup>116,117</sup> for labeling and imaging,<sup>118</sup> and also have been object of study for asymmetric reactions to obtain, for example, optically active succinimides<sup>119</sup> or quaternary amino acids.<sup>120</sup> Conjugations with several kinds of biomolecules have been reported, such as the formation of peptide-oligonucleotide conjugates,<sup>121</sup> protein-oligonucleotide conjugates,<sup>122</sup> protein-peptide conjugates<sup>123</sup> and many other combinations.

As stated before, an important feature of maleimides is their high reactivity toward nucleophiles. Maleimides can only be incorporated at the *N*-terminus of polyamides assembled by stepwise solid-phase procedures, because they are not stable to piperidine, the main reagent for amine deprotection employed after every coupling step. Also maleimides cannot be incorporated into oligonucleotide chains prior to cleavage because aqueous ammonia hydrolyzes them, and because ammonia can perform a Michael addition to the double bond.

A way to avoid the instability problem for oligonucleotides is to incorporate the maleimide moiety in solution, once the oligonucleotide is deprotected. This reaction requires the oligonucleotide to be derivatized with a linker that contains a free amine. This amine reacts with a molecule containing a maleimide and an active ester, like a succinimide ester, to finally yield the desired maleimide-functionalized oligonucleotide.<sup>124</sup> Nevertheless, oligonucleotide post-synthetic modifications in solution are often troublesome and low-yielded.<sup>125</sup>

In our research group different Diels-Alder reactions were performed between an *N*-terminal maleimido-peptide and a diene-oligonucleotide, obtained by the introduction of a diene phosphoramidite at the last synthetic step.<sup>121,126</sup>

To expand the range of synthetic possibilities and to overcome the lack of compatibility between maleimides and standard oligonucleotide synthesis conditions, in our research group a new methodology for protection-deprotection of maleimides was developed, completely orthogonal to either basic- or acid-based protection schemes. Maleimides were protected through a Diels-Alder reaction with 2,5-dimethylfuran (**Figure 1.6.1**). The *exo* cycloadduct proved to be stable to all the reagents of oligonucleotide and polyamide solid-phase synthesis. This protecting group is thermolabile. The cycloadduct undergoes a retro-Diels-Alder reaction at high temperature, which affords the fully reactive maleimide.<sup>127</sup>



**Figure 1.6.1:** Free maleimide and protected maleimide, *exo* adduct

Based on this maleimide protection methodology, a protected maleimide phosphoramidite is currently commercially available.<sup>128</sup> A collection of three different phosphoramidites allow for the incorporation of the maleimide moiety at any position of the chain,<sup>129</sup> and thus for the synthesis of cyclic oligonucleotides,<sup>130</sup> conjugates of oligonucleotides with natural linkages<sup>127</sup> or phosphorothioate oligonucleotide conjugates.<sup>131</sup>

### 1.7. Goals

One of the main goals of this thesis (Part A) was to further exploit the newly developed maleimide protecting group for the development of methodologies to access complex structures.

Given the current interest in conjugation, we wanted to develop a robust protocol for the assembly of polyamide conjugates through different “click chemistry” reactions, namely Michael additions and Diels-Alder cycloadditions. We also aimed for the synthesis of double conjugates, using peptide sequences containing two maleimides that could be conjugated selectively to two different biomolecules such as peptides or oligonucleotides.

The preparation of conjugates and double conjugates required the synthesis of new building blocks containing a protected maleimide. These building blocks were designed to be incorporated at any stage of peptide or peptoid synthesis, either within the sequence or at the *N*-terminal position.

We also wanted to exploit the maleimide-thiol reaction in the synthesis of cyclic and bicyclic peptides.

By combining these conjugation and cyclization strategies, we planned the synthesis of cyclic polyamides – peptides or peptoids – with a functional group available for conjugation, like a free maleimide, a free thiol or an alkyne. Further conjugation of these cyclic polyamides served as a proof of principle, thus validating this methodology of cyclic polyamide conjugates synthesis.

Finally, we aspired to apply the developed methodology to the synthesis of cyclic peptide – PNA conjugates with potential biological activity, to be further tested in collaboration with other research groups.



## 2. Results and discussion

### 2.1. Synthesis of maleimide-containing monomers

One of the requirements to fulfill the goals of this work is to have access to a variety of maleimide-containing building blocks.

For peptide synthesis we needed an amino acid derivatized with a protected maleimide, and we decided to prepare it by reacting Fmoc-lysine with [protected maleimido]propanoic acid.

Maleimidopropanoic acid (**1**, Fig. 2.1.1) is commercially available, but considering the price, around 230 € per gram, it seemed worth trying to prepare it from  $\beta$ -alanine and maleic anhydride that are cheap starting materials.

We first attempted a published protocol that claimed to be extremely easy and high yielded, which afforded the pure compound by precipitation in water.<sup>132</sup> After several attempts, in our hands the reaction did not work and we only obtained  $\beta$ -alanine maleate. Following a different protocol with much harsher reaction conditions and a silica gel column chromatography we obtained the desired compound in 49 % yield.<sup>133</sup>

The protection of maleimidopropanoic acid with 2,5-dimethylfuran had already been developed in the research group;<sup>127</sup> may therefore the following lines serve as a brief summary.

The reason behind the use of 2,5-dimethylfuran versus other furans is that the maleimide cycloadduct is completely stable at room temperature and the retro Diels-Alder reaction can be fulfilled at a reasonable temperature. Less substituted furans form too stable cycloadducts.

Reacting maleimidopropanoic acid (**1**) with an excess of 2,5-dimethylfuran in acetonitrile at 65°C for 6 hours yielded a mixture of *endo* and *exo* diastereomers. Dr. Albert Sánchez, who developed this methodology for oligonucleotides, employed the mixture of *endo* and *exo* protected maleimide diastereomers to attach the protected maleimide to oligonucleotides. He



## 2. Results and Discussion

observed that after the ammonia treatment that deprotects the oligonucleotide, a fraction underwent hydrolysis. He also noticed that the proportion of hydrolyzed and non-hydrolyzed functionalized oligonucleotide correlated with the proportion of *endo* and *exo* cycloadducts in the mixture of protected maleimide diastereomers. The conclusion was that the *endo* diastereomer was not stable under the deprotection conditions, and only the *exo* one was adequate for the preparation of maleimido-oligonucleotides.

Due to the need of using the pure *exo* cycloadduct, Dr. Sánchez took advantage of the different reactivity of the two diastereomers and reacted the reaction crude, containing both diastereomers, with concentrated aqueous ammonia overnight. After acidification and extraction with methylene chloride, the protected maleimide *exo* cycloadduct was recovered pure and in high yield. This compound (**2**) was used for the synthesis of maleimide-containing phosphoramidites<sup>129,127</sup> and for the incorporation of a protected maleimide at the *N*-terminal position of a polyamide.

Ammonia treatment to obtain the pure *exo* cycloadduct has also proved to be useful with other substrates in the synthesis of the maleimide building block for peptoids.

In order to assess the possibility of using both diastereomers in solid-phase peptide or peptoid synthesis, we performed a simple assay with a mixture of *endo* and *exo* diastereomers of [protected maleimido]propanoic acid. Upon the treatment with the Fmoc-removal solution (20 % piperidine in DMF), both diastereomers remained undegraded. However, we observed that the *endo* diastereomer was not stable to further derivatization of [protected maleimido]propanoic acid, and thus the employment of the mixture of diastereomers was ruled out.

[Protected maleimido]propanoic acid (**2**) was employed in this work for the synthesis of a maleimido-lysine building block (**4**) (**Fig. 2.1.1**), suitable to introduce protected maleimides into peptide chains.

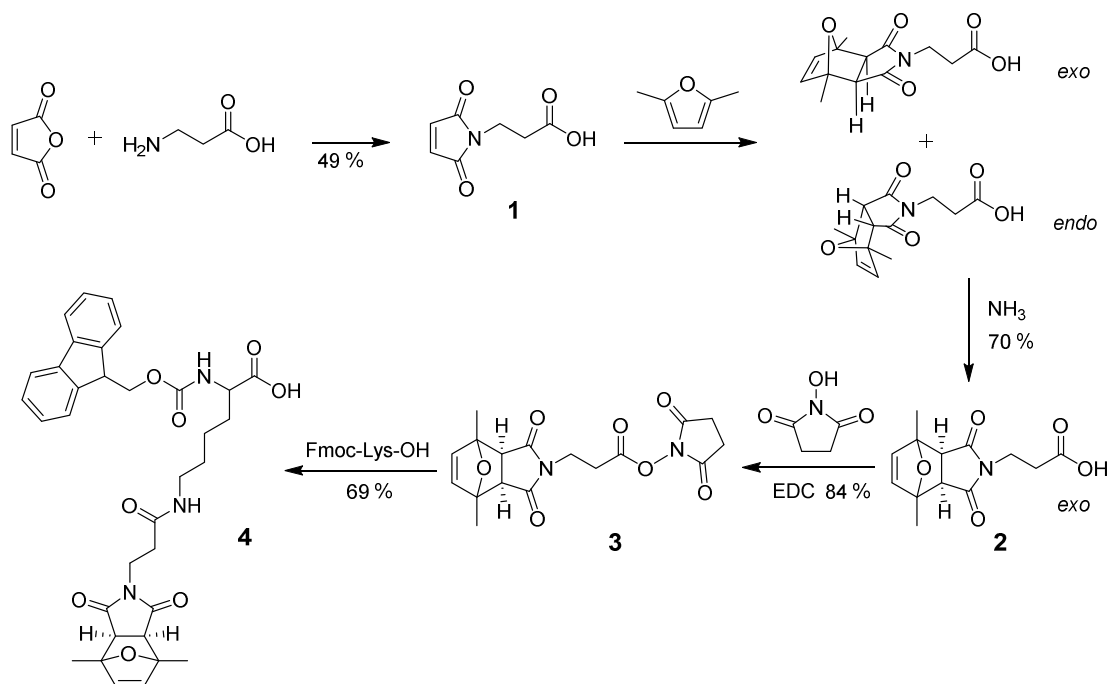
In order to link the lysine side chain to the carboxyl of [protected maleimido]propanoic acid (**2**), and to avoid the self condensation of lysine, [protected maleimido]propanoic acid was transformed into an active ester. It was reacted with *N*-hydroxysuccinimide, yielding the desired *N*-hydroxysuccinimide ester **3** in 84 % yield and pure enough to proceed with the next synthetic step without further purification.

This compound was further reacted with Fmoc-Lys-OH in the presence of *N,N*-ethyldiisopropylamine, affording the desired product **4** in 69 % yield after silica gel column chromatography. Fmoc-Lys[ProtMal]-OH can be employed for the incorporation of a protected maleimide at any position of the peptide chain.

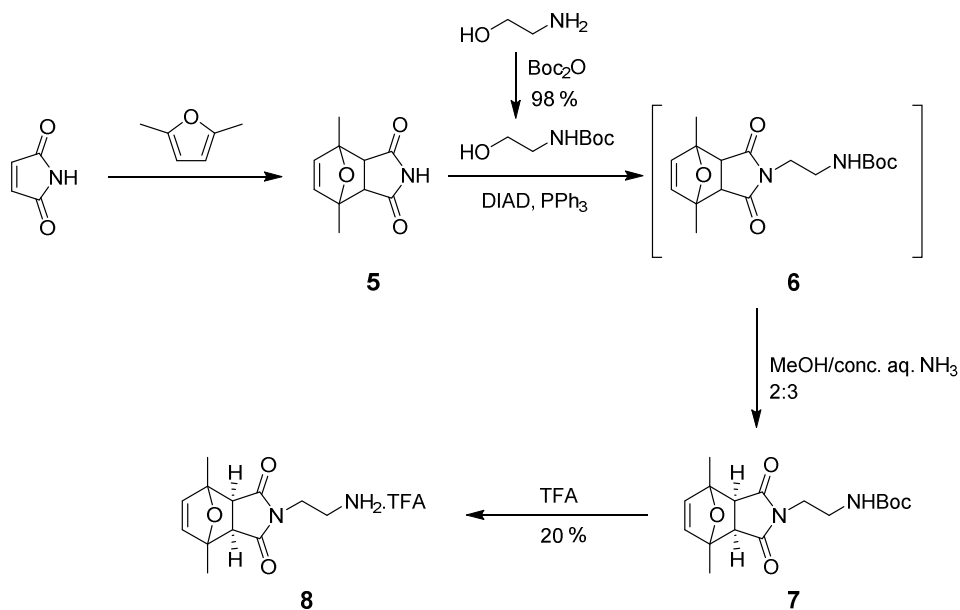
The monomer required for peptoid synthesis was an amino-functionalized protected maleimide (**8**). The synthesis depicted in **Fig. 2.1.2** was carefully optimized so it required a minimum amount of purification steps. It started with the protection of maleimide with 2,5-dimethylfuran. The mixture of diastereomers (**5**) underwent a Mitsunobu reaction with *N*-Boc-protected ethanolamine, DIAD and triphenylphosphine in THF. After the solvent was removed, the crude was directly treated with an ammonia/methanol mixture to hydrolyze the *endo*

product. After basification and extraction with methylene chloride, the amine **7** was deprotected with a TFA/methylene chloride mixture. As all the attempts to obtain the neutral desired product afforded the hydrolyzed maleimide, monomer **8** was obtained as a TFA salt after precipitation with diethyl ether.

Compound **8** can be employed for the introduction of a protected maleimide at any position of a peptoid chain.



**Figure 2.1.1:** Synthesis of maleimidopropanoic acid (**1**), [protected maleimido]propanoic acid (**2**) and Fmoc-Lys[ProtMal]-OH (**4**)



**Figure 2.1.2:** Synthesis of *N*-[protected maleimido]-ethanediamine (TFA salt) (**8**)

## 2. Results and Discussion

### 2.2. Conjugation

#### Single conjugations. Deprotection optimization and conjugate stability assessment

Once we were in possession of the different maleimide-containing building blocks, and before attempting the synthesis of more complex structures, we decided to test the behavior of the protected maleimides incorporated into peptides, in particular their deprotection and the stability of the conjugates generated.

For the first deprotection assays we synthesized a very simple peptide on a Wang resin. It was a sequence containing serine, tyrosine, glycine and our [protected maleimido]-derivatized lysine monomer (**9**, Fig. 2.2.1). Serine was introduced to increase hydrophilicity and thus improve the peptide solubility in water. Tyrosine was included to provide the peptide with a higher extinction coefficient (280 nm). The *N*-terminal position was acetylated in order to avoid obtaining a peptide with a global neutral charge, which might be sparingly soluble.

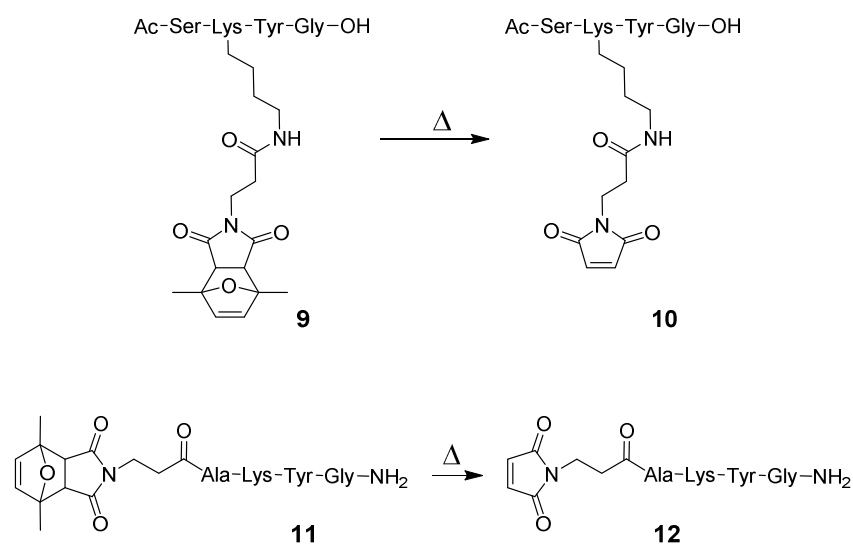


Figure 2.2.1: Peptides used in maleimide deprotection assays

Two different deprotection alternatives were employed. One consisted in heating compound **9** for shorter times in a methanol/water mixture in a microwave oven, and the other in heating it in a metal block for longer times suspended in toluene. For optimization we varied time, temperature, concentration and scale of the reaction. The best conditions we found were heating at 90°C for 90 minutes in the methanol/water mixture in the microwave oven, or to heat at 90°C for 3 hours in anhydrous toluene. A 25  $\mu$ M peptide concentration gave the best results.

These conditions were tested on another peptide (**11**, Fig. 2.2.1), in which the protected maleimide was located at the *N*-terminal position. This peptide was synthesized on a Rink amide resin and contained alanine, lysine, tyrosine and glycine. Lysine was introduced to provide a global positive charge to the peptide.

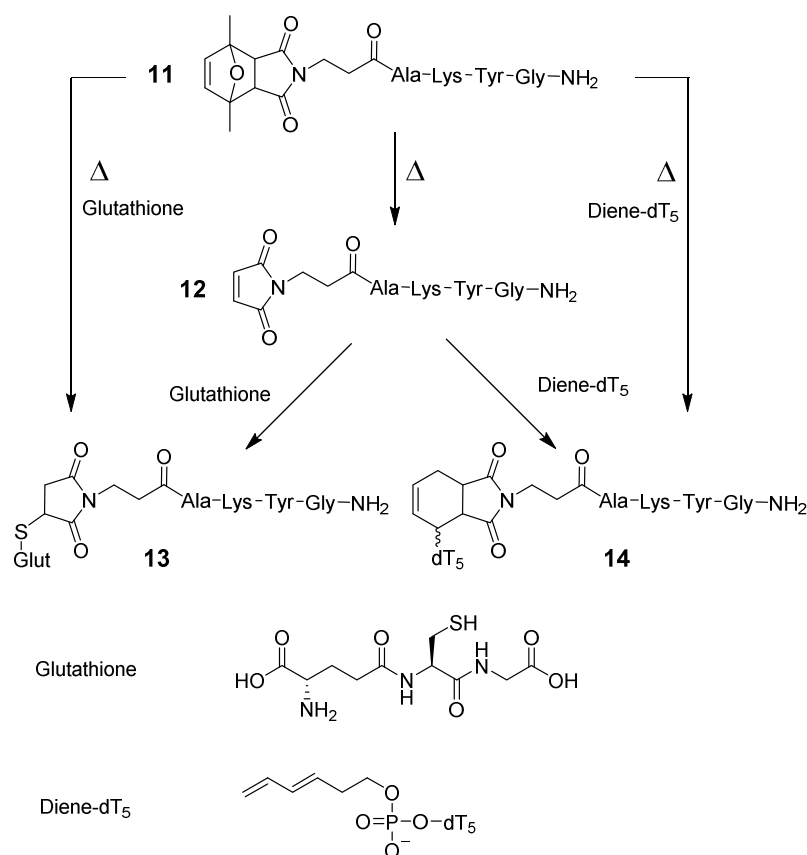
In this case maleimide deprotection by heating in a microwave did not afford the desired peptide, but a mixture of unknown compounds. Toluene deprotection conditions proved to be

better, the desired compound **12** was found in the mixture, but the reaction was not complete. Also, longer reaction times deteriorated the crude.

Nevertheless, the best deprotection crude, from **11**, was allowed to react with glutathione through a Michael addition. HPLC analysis showed that virtually all **12** available was conjugated, affording **13**. This confirmed that deprotection had afforded a fully reactive maleimide (Fig. 2.2.2).

Since we wanted to test the stability of both Michael adducts and Diels-Alder cycloadducts to maleimide deprotection conditions, a conjugate obtained from the Diels-Alder reaction was also prepared. **14** was obtained from the reaction of diene-dT<sub>5</sub> with **12**, in this case synthesized directly with a free maleimide instead of a protected maleimide at the *N*-terminal position.

Michael and Diels-Alder conjugates' stability was assessed by submitting them to both microwave and toluene heating deprotection conditions. Both conjugates remained undegraded under the two different conditions, which proved their stability.

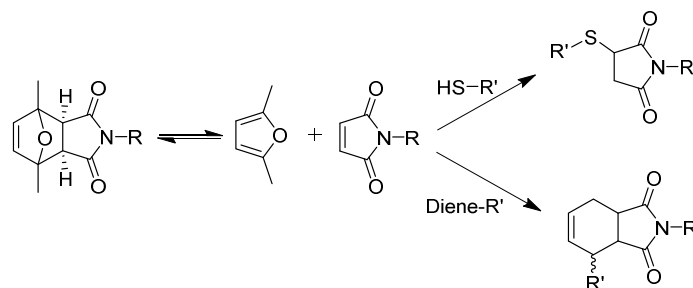


**Figure 2.2.2:** Synthesis of conjugates **13** and **14** from **11**, either in two steps (**11** → **12** → **13** or **14**) or one-pot (deprotection and conjugation at the same time)

Considering the great stability of the conjugates to maleimide deprotection conditions, we decided to attempt a one-pot deprotection and conjugation reaction. We mixed **11** with an excess of glutathione and heated it at 90°C in toluene for 6 hours, producing the desired conjugate in much higher yield than in the two-step sequence attempted before, with no detection of unreacted **11**. Likewise, we performed the same reaction with **11** and diene-dT<sub>5</sub>, with similar results (Fig. 2.2.2).

## 2. Results and Discussion

Without a compound to react with, namely a diene or a thiol, the deprotection reaction of a [protected maleimide] reaches an equilibrium between the free and protected maleimide, and consequently doesn't go to completion. Besides, maleimides apparently are not fully stable at high temperatures and they slowly degrade, directly affecting the overall deprotection yield. Yet in the presence of a conjugating agent (which traps the maleimide), at high temperature, maleimides react very fast yielding succinimides, which are much more stable and do not degrade under these conditions. Furthermore, the amount of maleimide decreases, so the equilibrium between protected and deprotected maleimide is displaced. In other words, conjugation is the driving force of the deprotection reaction (**Fig. 2.2.3**).



**Figure 2.2.3:** Conjugation drives maleimide deprotection

As stated in the introduction of this thesis, peptoids are peptidomimetics in which the side chain appends from the nitrogen atom of polyglycine chains. There are two main strategies for the synthesis of peptoids. The first consists in the synthesis in solution of *N*-protected, *N*-alkyl glycine monomers, which are assembled together in the same way as for solid-phase peptide synthesis. The second, known as the submonomer strategy, was developed by Zuckermann<sup>134</sup> and consists in a two-step protocol in which bromoacetic acid is first coupled to the solid support, and then a primary amine carrying the desired side chain is added in excess, performing a nucleophilic attack on the methylene and displacing bromide.

Using the submonomer procedure, we synthesized a short peptoid incorporating the [protected maleimido]-ethanediamine at the *N*-terminal position (**15**, **Fig. 2.2.4**). This peptoid was deprotected in toluene under very similar conditions as for peptides, but this time the reaction was completed without byproducts within 2.5 hours.

Deprotected peptoid **16** was allowed to react with glutathione, thiol-functionalized biotin and thiocholesterol to yield **18**, **19** and **20**, respectively (**Fig. 2.2.4**). So far, all the succinimides of the conjugates, generated after reaction of the maleimide via Michael addition or via Diels-Alder cycloaddition, were pretty stable in aqueous media. To our surprise, this was not the case of conjugate **19**, which underwent hydrolysis within minutes. The conjugate could be obtained by using a 10-fold excess of biotin and freezing the crude after 5 minute reaction.

Peptoid **16** was also allowed to react with diene-dT<sub>5</sub>. Due to the fact that Diels-Alder reactions proceeded more slowly than Michael additions, we decided to take advantage of the one-pot deprotection and conjugation strategy for the synthesis of **17**. **17** was obtained in 89 % yield through the one-pot strategy, by reaction of diene-dT<sub>5</sub> with **15**, whereas a conjugation attempt with **16** and diene-dT<sub>5</sub> at 37°C in water afforded only 23 % of the desired conjugate after 6 days of reaction.

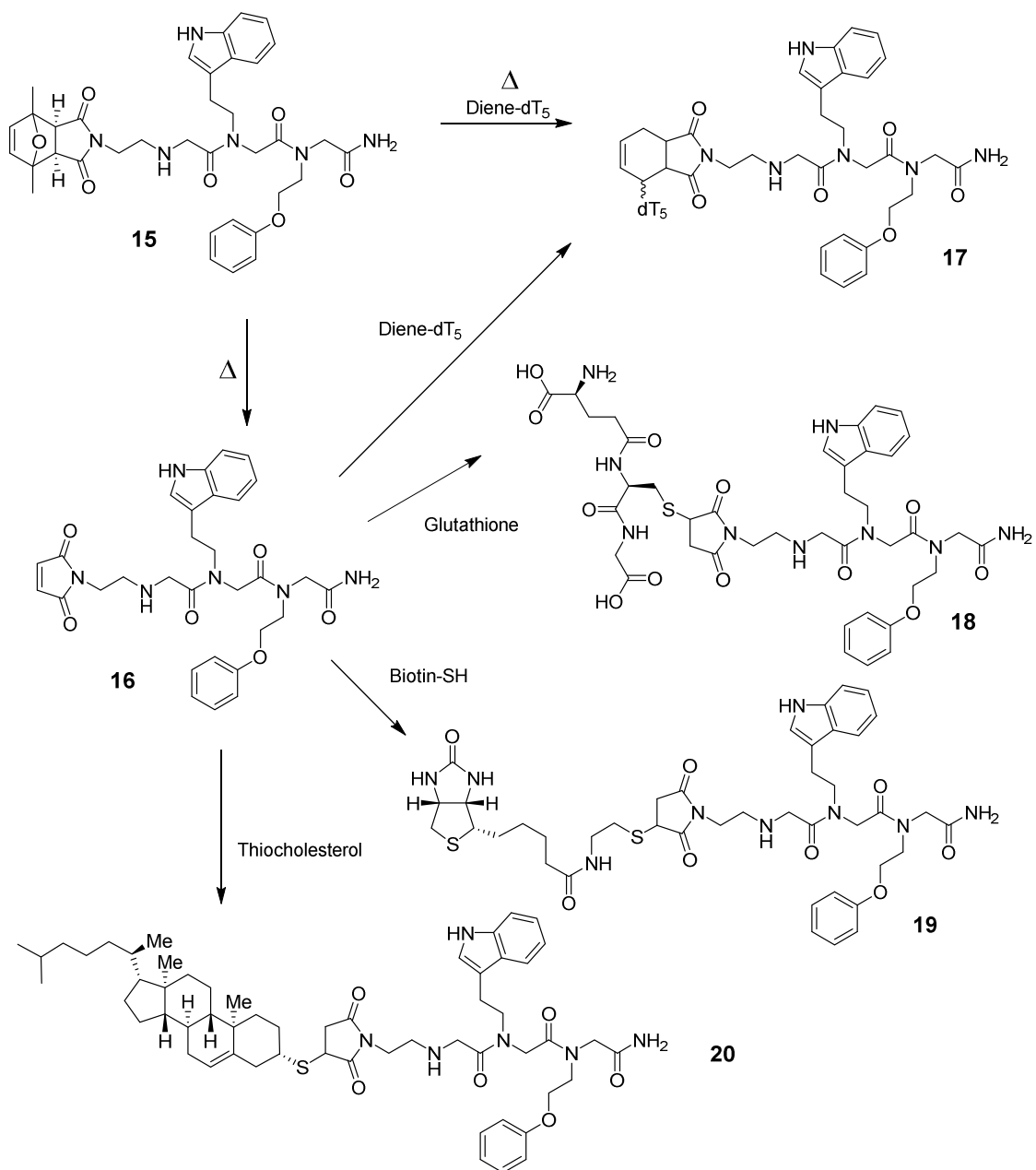


Figure 2.2.4: Maleimido-peptoid deprotection and conjugation

## 2. Results and Discussion

### Double conjugations

Once we had assessed the stability of the conjugates formed by Michael addition and Diels-Alder cycloaddition to maleimide deprotection conditions, we decided to tackle the synthesis of double conjugates.

Several strategies have been developed for the synthesis of double (or triple) conjugates via “click” reactions in biomolecules.

When using oligonucleotides as scaffolds, the most common reaction is the Huisgen cycloaddition. Some research groups have developed suitable deprotection conditions for differently-protected alkynes attached at different oligonucleotide nucleobases.<sup>135</sup> The double conjugates can be synthesized in three steps, first reacting a free alkyne, then deprotecting the protected alkyne and finally performing a second conjugation reaction. A different option is to incorporate a free alkyne and a functional group that can be transformed into an azide. A first Huisgen reaction is performed on the alkyne, then the azide is generated, and finally this new azide undergoes a second Huisgen cycloaddition.<sup>136,137</sup> Other alternatives are to introduce an azide and a masked azide, which can be transformed into a full reactive azide after a first conjugation,<sup>138</sup> or to tune azide reactivity by introducing a chelating azido group.<sup>139</sup> Also Sanders and col. have studied other dipoles with different reactivity for directed double “click” conjugations.<sup>140</sup>

The combination of Huisgen cycloadditions with other “click” reactions has also been exploited. As an illustration, the Jäschke group performed two different “click” reactions on a resin-linked oligonucleotide incorporating a terminal alkyne and a cyclooctene. They mixed the functionalized oligonucleotide with an azide and a tetrazine, thus obtaining the desired Huisgen and Diels-Alder cycloadducts.<sup>125</sup> Also, as briefly discussed in the introduction, the Defranq research group conjugated an oligonucleotide containing an alkyne and an aldehyde to different biomolecules through Huisgen cycloadditions and oxime bond formation.<sup>97</sup> Finally, Meyer and coworkers combined a Michael addition with a Huisgen cycloaddition to access to oligonucleotide-monosaccharide conjugates.<sup>141</sup>

For peptides and peptidomimetics as conjugation scaffolds, the Huisgen cycloaddition has also been largely employed. Different protecting groups for alkynes,<sup>142,143</sup> or the introduction of alkynes with different reactivity (like a regular alkyne and a cyclooctyne)<sup>144</sup> allow for a stepwise, double conjugate synthesis.

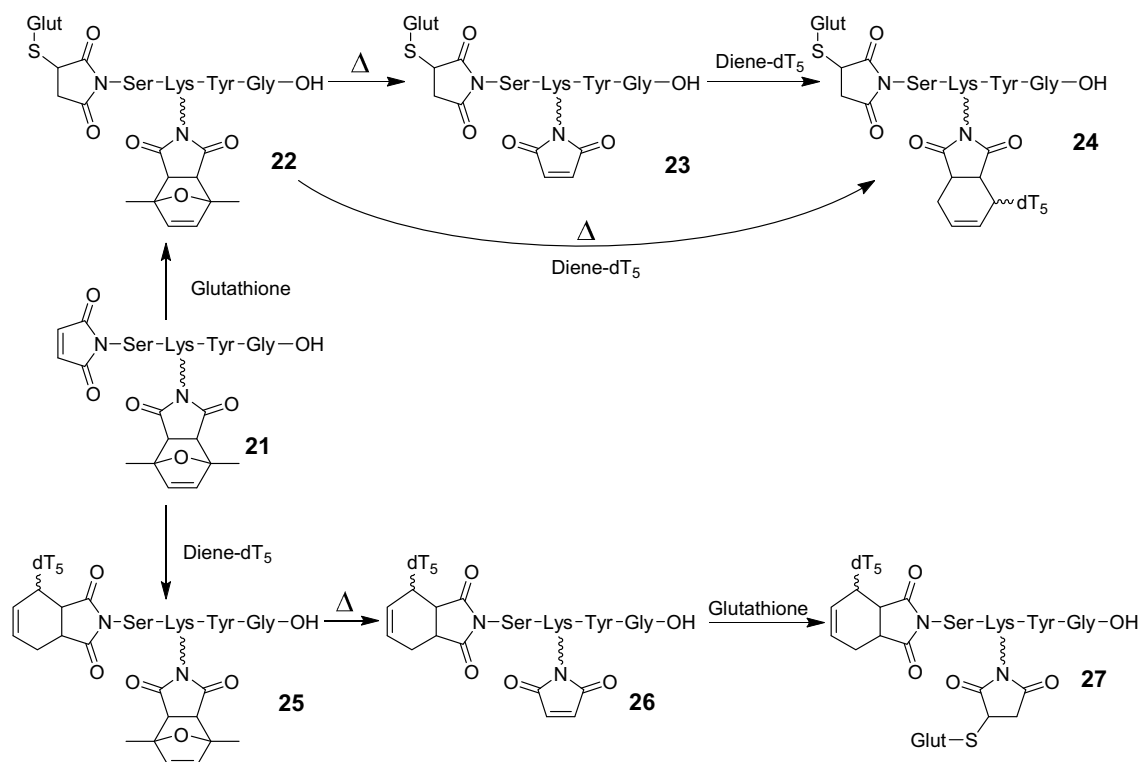
The employment of orthogonally protected thiols is a curious case. Differently protected cysteines have been used for the synthesis of multicyclic peptides, which can help to direct the formation of disulfide bonds. However, to the best of our knowledge, this has not been employed for the synthesis of double conjugates. A peptide containing two cysteines protected with trityl and S<sup>t</sup>Bu could be assembled. The first cysteine would be deprotected after cleavage, and it could undergo a reaction to yield a thioether. Deprotection of the second cysteine would afford a free thiol, which could undergo a reaction to yield either another thioether or a disulfide bond.

Other “click” reaction combinations have also been employed. As an illustration, a triple conjugate was synthesized on a protein, by electrophilic substitution on tyrosine by reaction with a triazoline-dione linker, generating an amide bond with a lysine side chain and performing a Michael addition between a maleimido-compound and a cysteine of the protein.<sup>145</sup> Another example of multiconjugation is the work of Galibert and coworkers.<sup>146</sup> They synthesized a cyclic peptide scaffold and, after slight postsynthetic modifications, they conjugated it to three different molecules one-pot through oxime bond formation, a Michael addition on a maleimide and a Huisgen cycloaddition.

In summary, several combinations of “click” reactions have been employed for the synthesis of double (or triple) conjugates. Nevertheless, to our knowledge, not very often the same functional group has been or can be used for two subsequent conjugation reactions on the same scaffold. Azides and alkynes have been exploited, and thiols could be used.

One of the goals of this thesis was to include maleimides in the same group as azides, alkynes and thiols, by obtaining different double conjugates through directed reactions on the different maleimide moieties of a peptide.

The strategy followed was to synthesize a peptide that contained a protected maleimide, by incorporating the lysine derivative monomer (**4**) synthesized previously, and a free maleimide at the *N*-terminal position by coupling maleimidopropanoic acid at the last step of chain elongation. This peptide could undergo a first conjugation with the free maleimide at the *N*-terminal position. Then, the protected maleimide on the lysine residue could be deprotected, and finally conjugated, affording a double conjugate (**Fig. 2.2.5**).



**Figure 2.2.5:** Double conjugations on a peptide



## 2. Results and Discussion

As a proof of principle, we assembled **21**, a simple peptide containing a free maleimide at the *N*-terminal position, the [protected maleimido] lysine derivative, serine, tyrosine and glycine on a Wang resin. After cleavage and purification, we conjugated it with either glutathione or diene-dT<sub>5</sub> and uneventfully obtained **22** and **25**.

**25** was deprotected under the same conditions as for peptoid **15**. Heating **25** in toluene at 90°C for 2.5 hours yielded the desired compound **26**, in high yield, ready for the second conjugation. Finally, **26** was mixed with an excess of glutathione to afford the desired double conjugate **27** in high yield.

The deprotection of **22** was much more problematic, as **22** did not undergo more than 50 % deprotection under the standard heating conditions in toluene. With the crude obtained we performed the second conjugation by mixing it with diene-dT<sub>5</sub>. We observed that the unprotected fraction reacted properly with the diene, affording **24**, whereas the protected obviously remained nonreactive.

We attempted the one-pot strategy, by mixing **22** with diene-dT<sub>5</sub> and heating it in toluene for 6 hours. The outcome of this reaction improved the 2 step strategy, because we observed full conversion of **22** to **24** and almost no byproducts.

Considering all our results in the attempts of maleimide deprotection, and also other results in the research group, it seems that deprotections proceed much better for substrates like oligonucleotides or PNAs than for peptides. This observation is also validated for conjugates. Peptides conjugated with oligonucleotides are deprotected more easily than those conjugated with other peptides. Peptoids do not behave as peptides since they perform much better in maleimide deprotection.

In summary, maleimide deprotections are strongly substrate dependant, and have to be optimized for every different kind of biopolymer and conjugate.

### 2.3. Cyclization

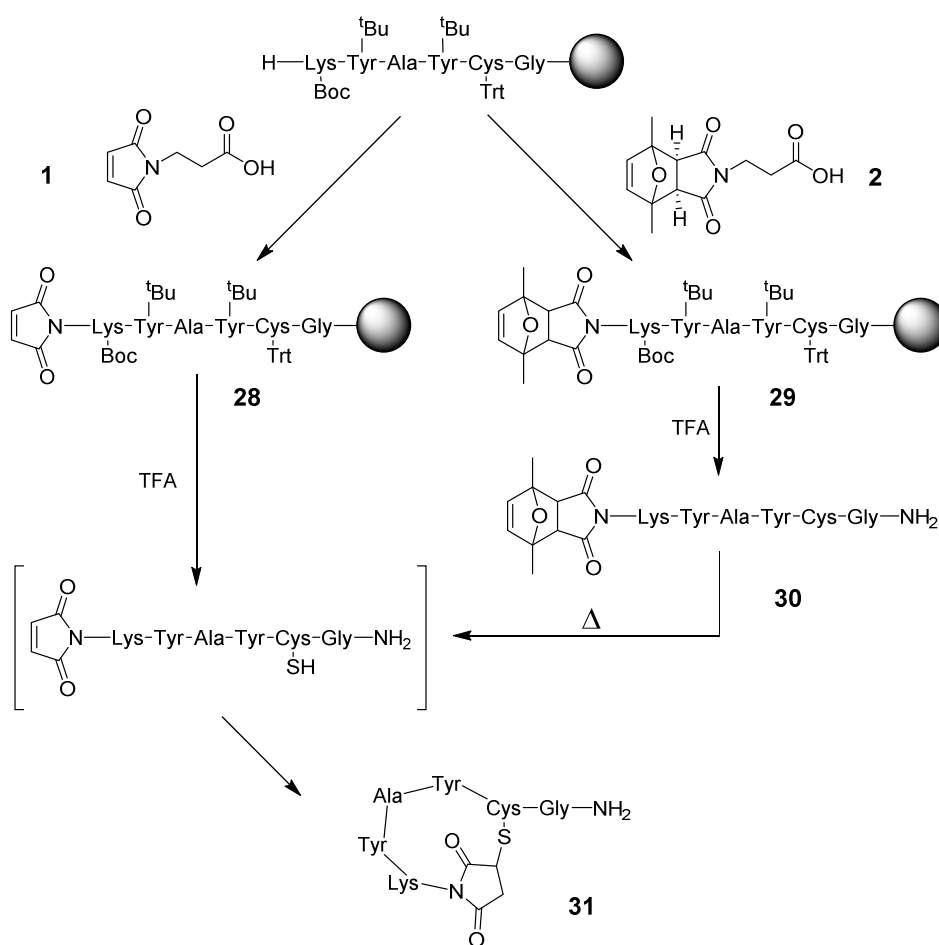
#### Monocycle formation

As stated in the introduction, cyclization is a common strategy to increase the activity and the stability of linear compounds. Peptide cyclization via amide bond formation has the main advantage that leaves no "scar" on the final compound, meaning that all the backbone bonds are identical. The main problem is that to perform this reaction all the other amino acid side chains have to be protected.

Michael addition of thiols on maleimides proceeds quickly, generates no byproducts, can be performed in water, and requires no protection of the side chains and no catalyst. Besides, thiols are excellent nucleophiles for Michael addition reactions and are readily available in cysteines. For these reasons we decided to exploit this reaction in the synthesis of cyclic and bicyclic peptides.

Surprisingly, to our knowledge only two papers have been published about the employment of maleimides for peptide cyclization. Sharma and coworkers employed a free maleimide at the *N*-terminal position to cyclize a short peptide on a resin.<sup>147</sup> The work from Koehler and coworkers, published during the development of this thesis and after our manuscript was accepted, takes advantage of the furan protection of maleimides to synthesize maleimido peptides and conjugate them to microarrays or to synthesize cyclic peptides, also on a solid support.<sup>148</sup>

All our cyclic peptides were obtained by reaction between maleimides and cysteines. The simplest way to introduce a maleimide in a peptide sequence is to incorporate it at the *N*-terminal position, because the monomers required are commercially available. But then, the question was whether the maleimide should be protected or deprotected (**Fig. 2.3.1**).



**Figure 2.3.1:** Alternatives for the synthesis of cyclic peptides with *N*-terminal maleimides

To answer this question, we decided to synthesize a short peptide terminated, in one case, with a free maleimide (**28**), and in the other, with a protected maleimide (**29**). The sequence also included a lysine to positively charge the peptide, two tyrosines to allow for the detection at low concentrations, and a cysteine protected with trityl to introduce the thiol moiety. The trityl protecting group is acid-labile. Therefore, during the acidolytic deprotection and cleavage of the peptide from the solid support, the thiol of the cysteine gets deprotected.

## 2. Results and Discussion

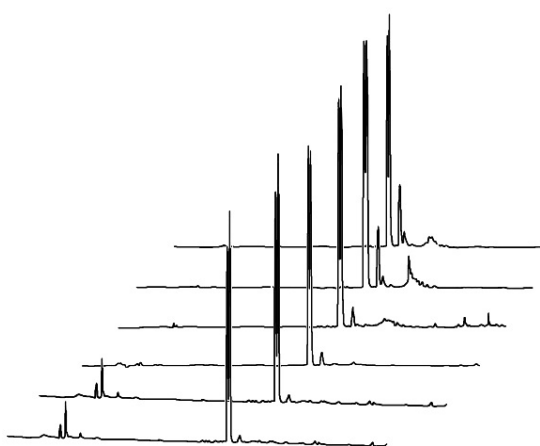
In the case of the peptide incorporating the free maleimide (**28**), the HPLC of the crude after cleavage showed a main peak corresponding to the desired cyclic peptide, meaning that the cyclization reaction is so fast (because the maleimide and free thiol are very reactive) that the Michael addition takes place even at low pH. The problem is that, even though the desired product corresponded to the largest peak in the HPLC chromatogram, there was a remarkable proportion (17 %) of cyclic dimer, resulting from a double interchain maleimide-thiol reaction. Also a 5 % of cyclic trimer was detected, as well as a small percentage of other unknown compounds.

In the case of the peptide incorporating the protected maleimide (**29**), the outcome was different. The TFA cleavage afforded a clean crude, with the main peak corresponding to the desired linear peptide **30**. Peptide deprotection and cyclization was performed in the microwave oven, dissolving **30** in a methanol/water mixture and heating at 90°C for 90 minutes.

This reaction was performed at a variety of concentrations in order to assess which range of concentrations was appropriate for peptide cyclization (**Fig. 2.3.2**). Starting at 25 µM, the desired cyclic peptide was obtained in 94 % yield (as assessed by HPLC). The cyclic dimer amounted 2 % and the rest corresponded to unknown byproducts. Increasing concentrations of **30** correlated with an increase in cyclic dimer and trimer. Above 10 mM concentration, the shape of the cyclization crude chromatogram strongly resembled that of cleavage of **28**.

These results are not surprising, taking into account that the concentration of **28** during the deprotection and cleavage event was fairly high.

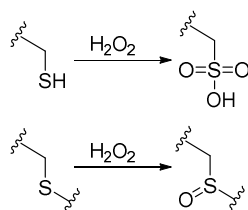
Concentration of <b>30</b>	Cycle ( <b>31</b> )	Dimer	Trimer	Other
25 µM	94 %	2 %	-	4 %
100 µM	97 %	3 %	-	0 %
500 µM	96 %	4 %	-	0 %
1 mM	95 %	5 %	-	0 %
5 mM	71 %	11 %	2 %	16 %
10 mM	69 %	13 %	3 %	15 %



**Figure 2.3.2:** HPLC chromatograms of the cyclization crudes at increasing concentrations of **30**

On the basis of these results, it is recommended to use protected instead of free maleimides at the *N*-terminal position of the peptides to obtain the corresponding cycles whenever the target, and its protection group scheme, allows for it.

A linear sequence with a free maleimide and a thiol has exactly the same molecular weight as the cyclization product. In order to assess that a cycle is obtained, and to make sure that we are not detecting a linear compound, we applied the same methodology previously developed in the group for cyclic oligonucleotides. A small aliquot of the peptide solution is reacted with diluted hydrogen peroxide for 1 hour. Hydrogen peroxide oxidizes the sulfur atom differently depending on the functional group in which it is involved. A free thiol, present in linear peptides, is oxidized to sulfonic acid, thus incorporating three oxygen atoms. Conversely, if the sulfur atom is forming a thioether like in cyclic peptides, it is only oxidized to sulfoxide, incorporating only one oxygen atom. Upon analyzing the crudes by MALDI-TOF, linear peptides show a mass increase of 48 UMA, whereas cyclic peptides show an increase of only 16 UMA. In this way, linear and cyclic peptides can easily be distinguished (**Fig 2.3.3**). Unfortunately, this methodology is not appropriate when the polyamide contains disulfide bridges, since they get oxidized to a variety of compounds, incorporating 1, 2 or 3 oxygen atoms (the disulfide bridge may also be cleaved). In these cases this methodology is not reliable and clear conclusions cannot be drawn.



**Figure 2.3.3:** Sulfur oxidation with H<sub>2</sub>O<sub>2</sub>

A different strategy employed for cyclization assessment was digestion with thrombin. Thrombin is an endoprotease that hydrolyzes selectively the arginine – glycine peptide bond in GRG sequences. Linear peptides containing this sequence are hydrolyzed within an hour, but if the GRG motif is included in a cycle it remains undegraded. This strategy is obviously limited to peptides containing the GRG motif, so, for example, it could not be used with cyclic peptide **31**.

Once we had succeeded with the cyclization of a relatively small and easy peptide, we decided to attempt the cyclization of longer peptides with more relevant and complex sequences (**Fig 2.3.4**).

We first synthesized the cyclic peptide Mal\*-PF<sub>D</sub>FNQYV-Orn-C\*-NH<sub>2</sub> (**33**) (the \* indicates the residues linked during the cyclization reaction). This 10 amino acid sequence is an analog of the natural antibiotic tyrocidine A (cyclic<sub>[D<sub>5</sub>PF<sub>D</sub>FNQYV-Orn-L]</sub>).<sup>149</sup> Tyrocidine is a mixture of very similar cyclic peptides that is produced by *Bacillus Brevis*. Its mechanism of action consists in disrupting the cell membrane of bacteria. This is especially interesting in the development of new antibiotics, because bacterial colonies are much less prone to develop resistance against this kind of mechanism of action.<sup>150</sup>

We assembled the peptide sequence on a Rink amide resin, coupling trityl-protected cysteine as the first amino acid, and introducing [protected maleimido]propanoic acid at the *N*-terminal

## 2. Results and Discussion

position. The acidic treatment afforded a linear peptide with all the side chains deprotected in high yield (**32**). The peptide was then dissolved in a methanol/water mixture to a 100  $\mu\text{M}$  concentration and was heated in the microwave oven to carry out the maleimide deprotection and cyclization. This reaction was attempted at two different scales, 1 and 10 mL, to test the scalability of the cyclization. Both attempts yielded more than 85 % of cyclic peptide **33** (as assessed by HPLC).

We then aimed for an even longer peptide. In this case, we synthesized a loop analog of an oxidoreductase, concretely dihydrolipoamide dehydrogenase.<sup>151</sup> This protein is known to participate in enzyme complexes as a homodimer in several metabolic pathways, and to act as a protease in its monomeric form.<sup>152</sup>

The linear precursor synthesized was ProtMal-EARERILPTYDC-NH<sub>2</sub> (**34**). This 12 amino acid sequence was assembled on a solid phase, and [protected maleimido]propanoic acid (**2**) was incorporated at the *N*-terminal position. After the acidic treatment, the HPLC chromatogram showed 2 main peaks, one corresponding to the desired linear peptide **34** and the other resulting from formation of the disulfide bond between cysteine and ethanethiol, included in the cleavage cocktail as a scavenger. Both compounds were collected, and the one containing the disulfide bond was reduced with TCEP and repurified (the combined fractions finally yielded 21 % of the desired linear peptide **34**).

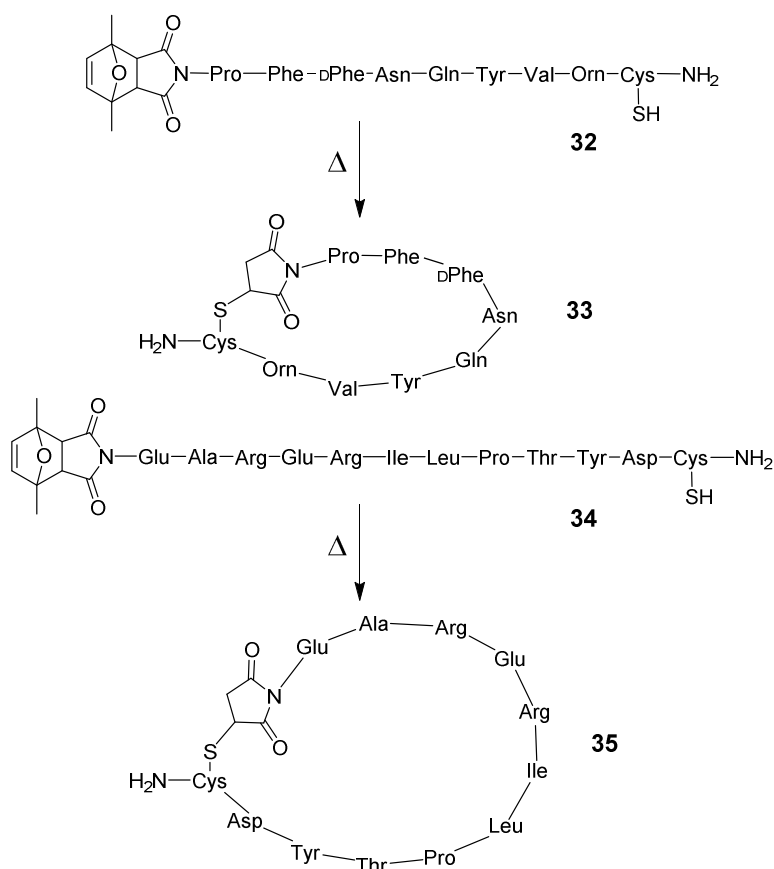


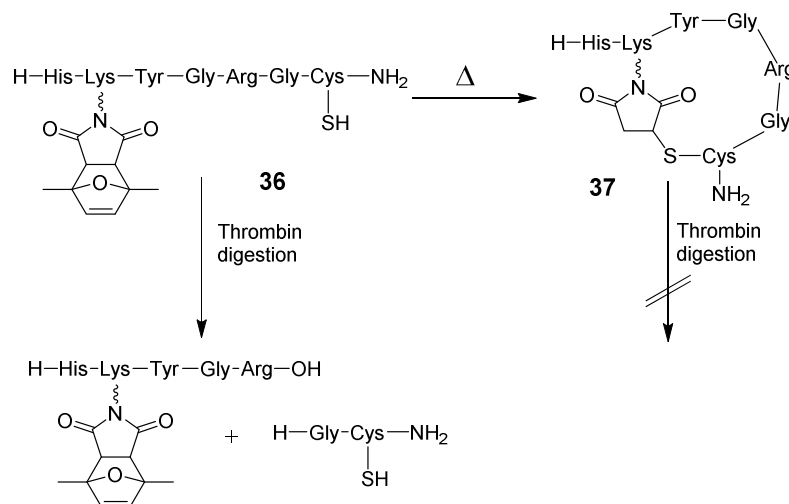
Figure 2.3.4: Thiol-maleimide cyclization with 9- and 12-mer peptides

The purified compound was redissolved in a methanol/water mixture and heated in the microwave oven at 90°C for 90 minutes. Two different concentrations were assayed, 25  $\mu\text{M}$  and 100  $\mu\text{M}$ . The reaction at 25  $\mu\text{M}$  concentration afforded the desired cyclic peptide **35** in a 45 % reaction and purification yield, whereas the HPLC trace of the reaction at 100  $\mu\text{M}$  concentration showed only unidentified compounds.

We also carried out, instead of head-to-side chain, side chain-to-side chain cyclizations. For this purpose we employed the lysine monomer (**4**). We also introduced a histidine residue next to the maleimide, to ultimately demonstrate the compatibility of this cyclization reaction with all the different amino acid side chains. Finally, we also decided to introduce a GRG motif in the sequence as a tool for peptide cyclization assessment with thrombin. To sum up, the sequence synthesized was H-Lys[ProtMal]-YGRGC-NH<sub>2</sub>. The synthesis proceeded smoothly, and after cleavage and purification the linear peptide **36** was obtained in 35 % yield.

Maleimide deprotection and cyclization was also carried out in the microwave oven at a 100  $\mu\text{M}$  peptide concentration. The quality of the cyclization crude allowed us to conclude that histidine does not interfere with the cyclization reaction, even if it is close in space to the maleimide. This is not surprising, since in a previous assay only cysteamine and maleimide reacted when a mixture of [protected maleimido]propanoic acid (**2**), cysteamine hydrochloride and histamine hydrochloride was heated in the microwave oven (see Appendix 2).

Both linear (**36**) and cyclic (**37**) compounds were digested with thrombin and analyzed by mass spectrometry. Whereas the linear peptide **36** underwent hydrolysis, the cyclic compound **37** remained unaltered, which confirmed that cyclization had taken place (Fig. 2.3.5).



**Figure 2.3.5:** Histidine-containing peptide side chain-to-side chain cyclization and digestion with thrombin

### Bicyclic peptides

In order to further exploit the maleimide possibilities for peptide cyclization, we devised the option of synthesizing bicyclic peptides.

Bicyclic peptides are interesting because they are even more constrained than monocyclic peptides, providing them with an even higher stability and in some cases improved activity.

## 2. Results and Discussion

The most common strategy for the synthesis of bicyclic peptides is to generate an amide bond between the ends of the backbone and then oxidize the free thiols to internal disulfide bonds (see Section 1.5).

Fasan and coworkers reported the synthesis of bicyclic peptides by oxime bond formation.<sup>153</sup> They introduced a ketone-modified tyrosine residue and a trifunctional building block with an amine, a thiol and an oxyamine. This molecule underwent a native chemical ligation-like reaction at the C-terminal position of the peptide, and the oxyamine reacted with the ketone to yield an oxime. Finally, the free thiol was oxidized to generate a disulfide bond with the thiol of another cysteine residue. In a different approach, May and Perrin employed the Savigne-Fontana reaction to afford a bicyclic peptide, with a linkage between the thiol of a cysteine and a tryptophan precursor.<sup>154</sup> The problem with this approach is that the tryptophan precursor had to be included as a tripeptide assembled in solution. Ring closing metathesis has also been employed for the synthesis of bicyclic peptides, by introducing a suitably derivatized tyrosine and allyl serines.<sup>155</sup>

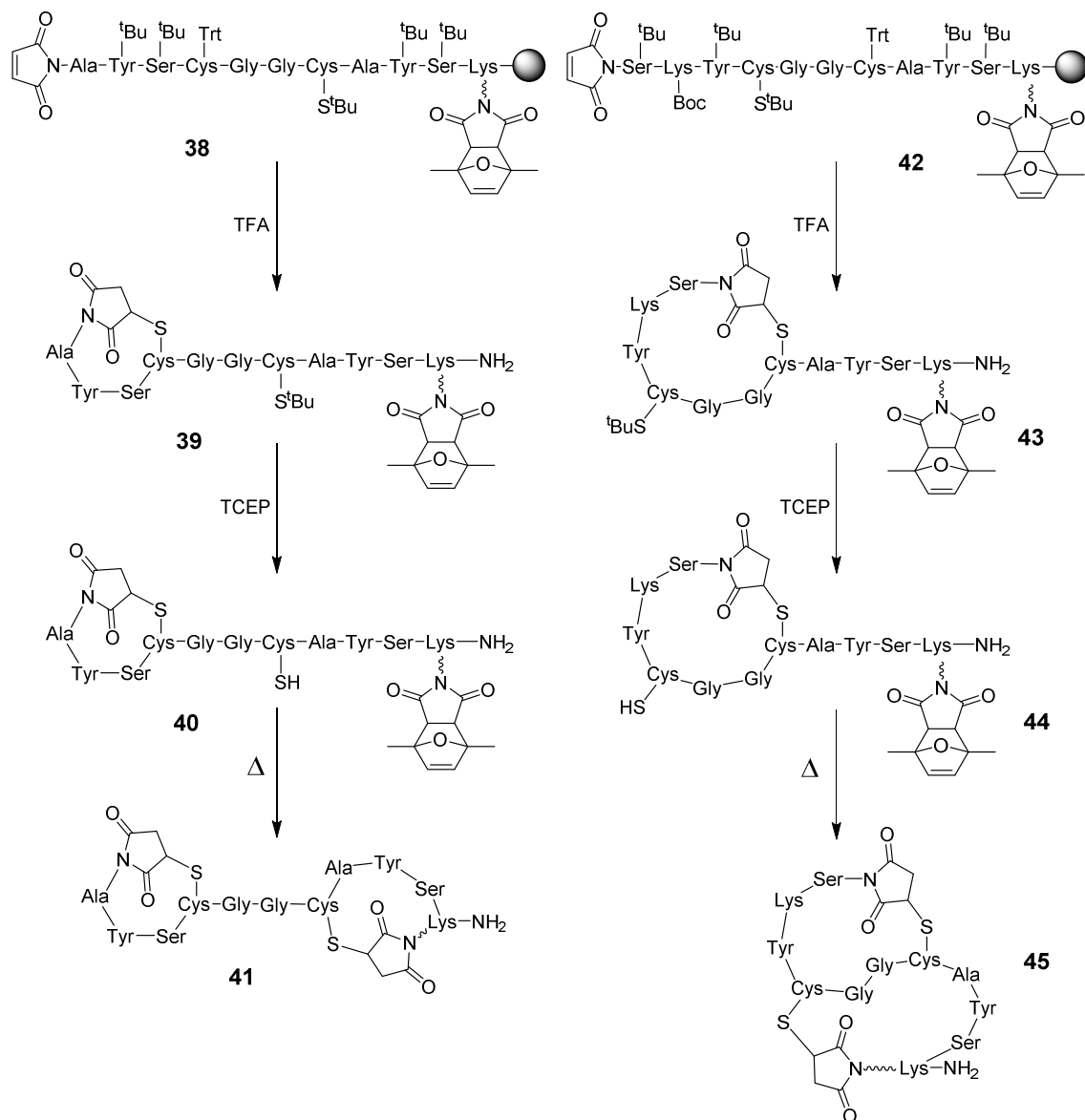
To the best of our knowledge the synthesis of bicyclic peptides using the same “click” reaction for the formation of each cycle has not been reported. The use of maleimides for this purpose, considering the compatibility with all the peptide side chains, seemed promising.

In the first monocycle prepared in this work we showed that we could prepare cyclic peptides by using either free or protected maleimides. By combining these two options in the same sequence and introducing differently protected cysteines we could easily synthesize bicyclic peptides (**Fig. 2.3.6**).

The strategy is to incorporate a free maleimide at the *N*-terminal position and a protected maleimide within the chain, by using the [protected maleimido] lysine derivative (**4**) described before. The sequence should also contain a cysteine protected with an acid-labile protecting group, and a differently protected cysteine, for example with a disulfide bridge. During the acidic treatment the first cycle would be generated by reacting the free maleimide at the *N*-terminal position with the free thiol of the cysteine. In a second step, the second cysteine would be deprotected with a reducing agent, cleaving the disulfide bridge. The last step would be to heat the peptide, thus deprotecting the second maleimide and generating the second cycle by reaction with the second cysteine. Reduction of the disulfide has to be carried out before maleimide deprotection, and the reagent used for this purpose thoroughly eliminated to prevent reaction with the maleimide generated at the subsequent step.

We first synthesized Mal-AYS-Cys(Trt)-GG-Cys(S<sup>t</sup>Bu)-AYS-Lys[ProtMal]-Rink (**38**). As expected, after acidic treatment the first cycle was generated. The overall synthesis, cleavage, cyclization and purification yield of **39** was only of 4 %. One of the reasons for such low yield is that the cyclization reaction happens at high concentration, so a remarkable proportion of dimer, trimer and other unidentified byproducts is generated. Another reason is that the solubility in water of this peptide was very poor, greatly complicating its purification. This poor solubility might be a consequence of a repeated pattern in the sequence, AYS. It is known that repetitive sequenced peptides may be prone to aggregation and thus have poor solubility.

The  $S^t\text{Bu}$  protecting group was removed by reaction with an excess of TCEP. The yield was again very low, only 5 %, for the same solubility issues stated before. Once the second cysteine of the cyclic peptide **39** was deprotected, **40** was redissolved in a methanol/water mixture to a concentration of 25  $\mu\text{M}$  and heated at 90°C for 90 minutes. This reaction was high yielded, 73 % from the HPLC trace, and afforded the desired bicyclic compound (**41**).



**Figure 2.3.6:** Bicyclic peptide syntheses

Due to the relative positions of the maleimide-containing lysine derivative and the two differently protected cysteines, the bicycle synthesized (**41**) displayed two isolated cycles linked by two glycine residues. We devised that if we changed the positions of the cysteines, by following the same deprotection and reaction scheme, we could obtain a different pattern of cyclization.

We synthesized a similar peptide sequence, but exchanging the positions of the two cysteines and introducing minor changes in the sequence, such as a serine instead of alanine, for solubility reasons (**42**). The first cycle formation, which occurred during the cleavage and side



## 2. Results and Discussion

chain deprotection, yielded an 11 % of the desired compound (**43**) after purification. Even though the yield was not high, it was much better than for the previous bicycle. Cys(S<sup>t</sup>Bu) deprotection yield (35 %) also surpassed the one of the previous bicycle, but the second cycle formation did not proceed so smoothly (25 %, HPLC trace). The final bicycle (**45**) was obtained, this time with both of the cycles fused instead of isolated.

These results illustrate the fact that proper peptide solubility is crucial for good yields, and that even though cyclization reaction yields may be sequence dependant, this methodology for bicyclic peptide synthesis is virtually applicable to any sequence.

### 2.4. Conjugated cyclic peptides. Proof of concept

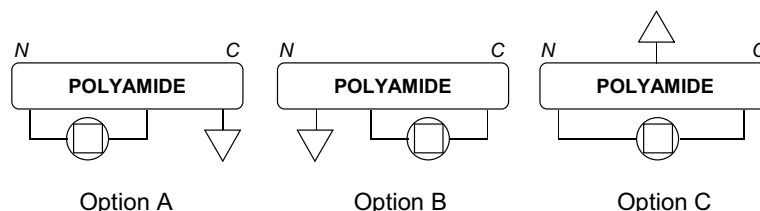
At this point we had obtained different conjugates and double conjugates of peptides taking advantage of both Michael additions and Diels-Alder cycloadditions, and we also had synthesized cyclic and bicyclic peptides. Therefore, we decided to go a step forward and try the synthesis of cyclic polyamides derivatized for conjugation. By conjugating cyclic peptides the benefits of conjugation and cyclization can be combined and thus active and resistant molecules can be obtained.

Plenty of authors have conjugated cyclic peptides for different purposes with different strategies. Dumy's research group synthesized a template cyclic peptide that contained hydroxylamines, which were conjugated through oxime bonds to cyclic RGD peptides, which target  $\alpha_v\beta_3$  integrin associated with tumor cells. A serine residue of the template peptide was then oxidized to aldehyde, to further conjugate it through an oxime bond to other molecules, like a 15mer oligonucleotide for duplex stability assessment.<sup>156</sup> Imma Güell and collaborators synthesized a cyclic peptide by attaching a glutamic acid side chain to a Rink amide resin, elongating the chain, deprotecting both the C- and N- terminus and generating an amide bond between them. The sequence, known to have antibacterial activity, incorporated either an azide or an alkyne. The subsequent conjugation took place on the solid phase when a Huisgen Cu(I)-catalyzed cycloaddition was carried out between the cyclic peptide and an alkyl- or azide-containing compound.<sup>157</sup> Finally, Conti's research group developed a cyclic peptide conjugate for imaging. The cyclic peptide also contained the RGD motif mentioned before, and a cysteine residue, which was conjugated to a bifunctional group with a maleimide and a chelator. This chelator could trap Cu<sup>64</sup>, commonly used in radiopharmaceuticals for imaging.<sup>158</sup>

In this work, the first syntheses of cyclic peptide conjugates taking advantage essentially of maleimide reactivity aim to serve as proof of concept and broaden the scope of the already existing possibilities for the synthesis of cyclic polyamide conjugates. Since we have worked both with peptides and peptoids, we will use the word polyamide here to refer to both.

Considering the structure, we envisioned three different options for the synthesis of cyclic polyamides for conjugation. The first option (A) we considered was to generate the cycle between the N-terminal position and a side chain in the middle of the sequence, leaving the conjugation spot, this is, the functional group to react with another molecule, outside the cycle and next to the C-terminal position. The second option (B) was the opposite. It consisted in

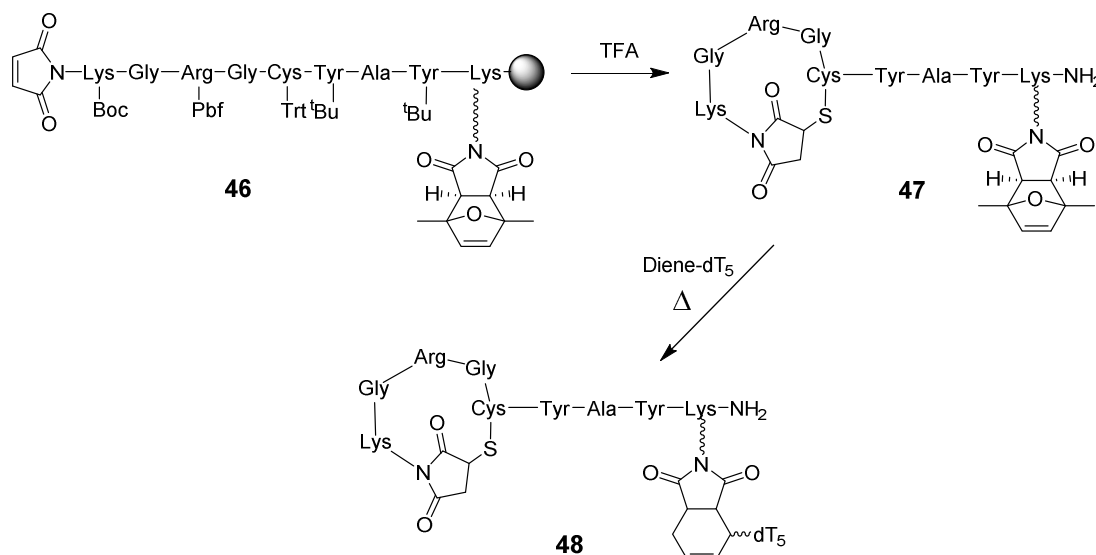
generating the cycle between the C-terminal position (or a side chain of an amino acid next to it) and a side chain of an amino acid in the middle of the sequence, leaving the conjugation spot outside the cycle, at the N-terminal position. The last option (C) was to create a cycle involving both the C- and N-terminal positions that also included the conjugation spot (**Fig 2.4.1**).



**Figure 2.4.1:** Three options for the synthesis of cyclic peptides derivatized for conjugation. The circle with the square represents a succinimido-thioether linkage, and the triangle is the functional group that will be used at the conjugation step

For option A, the strategy consisted in incorporating a thiol and two maleimides into the peptide chain, one free and one protected. The idea was to generate the cycle during the cleavage, so a free maleimide was attached to the N-terminal position of the peptide, and a cysteine with an acid-labile protecting group placed within the sequence. A protected maleimide was introduced into the sequence by coupling of the lysine derivative **4** next to the C-terminal position, prior to the cysteine, in order not to be included in the cycle (**Fig 2.4.2**).

We designed a 9-amino acid peptide, namely Mal-KGRG-Cys(Trt)-YAY-Lys[ProtMal]-Rink (**46**). Lysine was introduced to positively charge the sequence and make the peptide more hydrophilic. The GRG motif was introduced in order to assess cyclization by thrombin digestion assays.



**Figure 2.4.2:** Synthesis of a cyclic peptide – oligonucleotide conjugate (option A)

Synthesis, cleavage and purification afforded **47** in 17 % yield. When treating compound **47** with thrombin (**Fig. 2.4.2**), it remained undegraded, thus confirming that the GRG sequence was within the cycle. A conjugate of the cyclic peptide **47** was easily obtained in one more

## 2. Results and Discussion

step, by mixing it with an equimolar quantity of diene-dT<sub>5</sub> and heating in the microwave oven for 90 minutes at 90°C. Conjugate **48** was obtained in 90 % yield (from the HPLC chromatogram).

A was the most straightforward of the three options, because it only required one step of cleavage and one step of conjugation to obtain a cyclic peptide-oligonucleotide conjugate.

A different alternative, using option A, to obtain cyclic peptide conjugates, would be to introduce a cysteine protected with an orthogonal protecting group, like Cys(S<sup>t</sup>Bu), instead of the protected maleimide. This would allow for the conjugation not only with maleimido-containing molecules, but also with compounds prone to react with thiols, such as gold nanoparticles.<sup>159</sup> Nevertheless, this alternative would require an extra step, because disulfide bridge reduction and maleimide conjugation cannot be done one-pot.

Option B was the most laborious one. The strategy planned was to introduce the lysine derivative **4** at the C-terminal position and two orthogonally protected cysteines, one at the N-terminal position, which would be used for conjugation, and the other within the chain, protected with a trityl group, for cyclization. During the cleavage, this cysteine would be deprotected. Then, when heating the linear peptide at 90°C, the cycle would be formed by reaction between the free cysteine and the maleimide. Finally, the cysteine at the N-terminal position would be deprotected and conjugated, yielding the desired cyclic peptide conjugate.

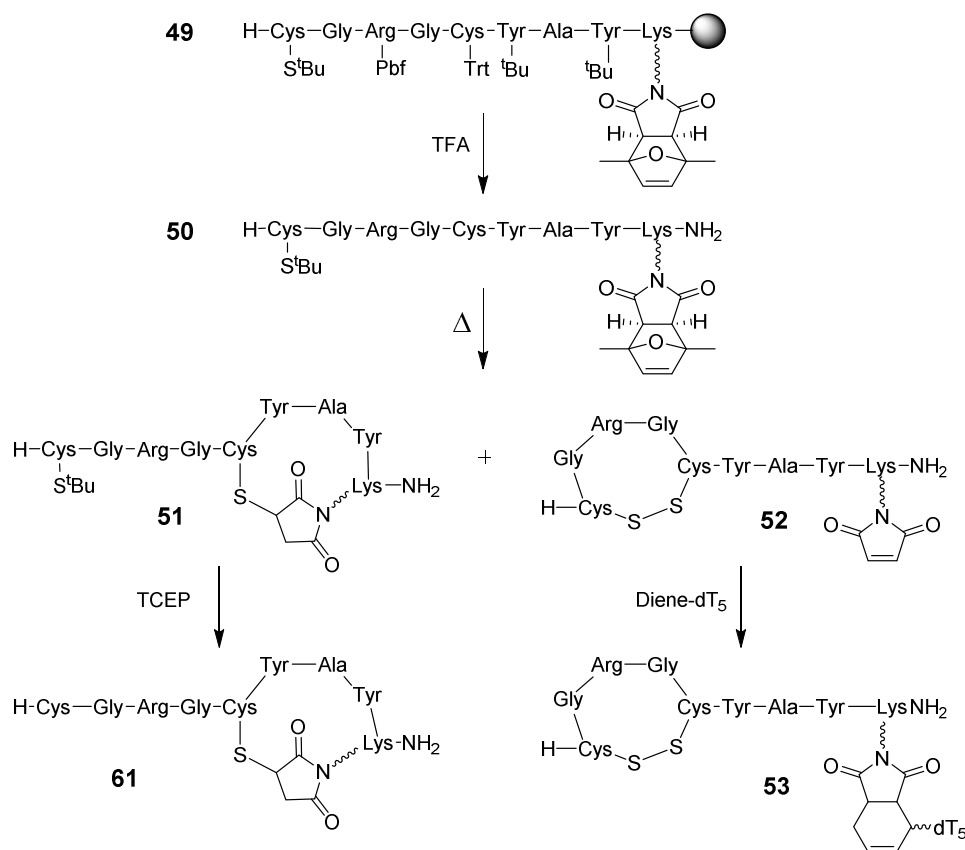
The challenging point of this option was that the protecting group of the cysteine located at the N-terminal position had to remain stable under the maleimide deprotection conditions. Given that limitation, prior to peptide synthesis we decided to test three common cysteine protecting groups. The compounds assayed were Fmoc-Cys(S<sup>t</sup>Bu)-OH, Fmoc-Cys(Acm)-OH and Boc-Cys(Fm)-OH. They were dissolved in a methanol/water mixture and heated in the microwave oven for 90 minutes at 90°C. All of them remained undegraded, except from the partial loss of the Boc group from Boc-Cys(Fm)-OH. This fact did not compromise the usage of Cys(Fm), since the Boc group was already removed from the peptide at the maleimide deprotection step. The conclusion of these assays was that, a priori, the three protecting groups could be employed with the N-terminal cysteine.

We synthesized a peptide with the same sequence as for option A, only replacing the terminal maleimido-lysine by a cysteine residue.

The synthesis with the S<sup>t</sup>Bu protected cysteine proceeded uneventfully, affording the desired linear peptide (**50**) in 23 % yield (**Fig. 2.4.3**). The problem appeared at the cyclization step. After heating the linear peptide in the microwave oven, two main products in similar ratio were obtained, **51** and **52**. The desired compound (**51**) had the expected mass and could be digested with thrombin. It reacted with TCEP, yielding 76 % (in the HPLC trace) of the desired deprotected peptide ready for conjugation (**61**).

As to the other product obtained in the deprotection (**52**), its mass matched with the loss of the S<sup>t</sup>Bu protecting group. Nevertheless, it remained undigested after treatment with thrombin. This result suggested that the peptide was cyclic, and that the cycle included the GRG motif. This reduced the alternatives to two: either both the maleimide and Cys(S<sup>t</sup>Bu) had

been deprotected and reacted, thus generating a big cycle, or only Cys(S<sup>t</sup>Bu) was deprotected and reacted with the other cysteine forming a disulfide bridge. In case it was the latter option, the maleimide would be free and ready for conjugation, so we allowed it to react with diene-dT<sub>5</sub>. The crude showed a main peak corresponding to conjugate **53**, confirming that we had obtained an intramolecular disulfide bridge and assessing that the S<sup>t</sup>Bu protecting group is not suitable for the synthesis strategy we were aiming for.



**Figure 2.4.3:** Assay to obtain a cyclic peptide derivatized with Cys(S<sup>t</sup>Bu) for conjugation (option B)

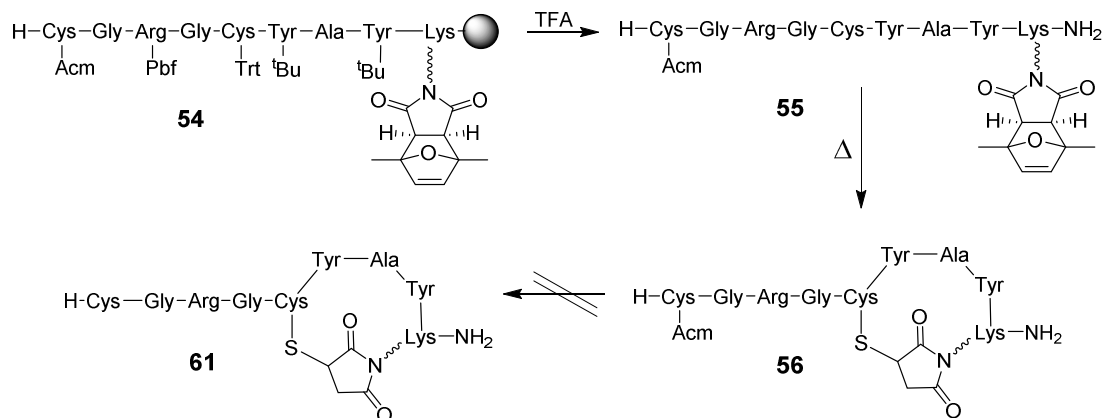
We next attempted the synthesis with the cysteine acetamidomethyl (Acm) protecting group (**Fig 2.4.4**). The linear peptide (**55**) was obtained in 34 % yield, and the cyclization step also proceeded smoothly to yield **56** (cyclization and purification yield: 30 %). In this case, problems appeared at the Cys(Acm) deprotection step.

We first tried the deprotection with silver triflate in trifluoroacetic acid. After 30 minutes, the reaction was quenched with DTT and the mixture analyzed by HPLC. Unfortunately the desired product (**61**) was detected in less than 5 % in the HPLC trace of the crude. We then tried the deprotection with iodine in an acidic medium. After 5 hours the reaction was quenched with sodium ascorbate and analyzed. The desired product (**61**) was not detected by either HPLC or mass spectrometry, but masses of the peptide plus one to four iodine atoms were observed. The last attempt was the deprotection with thallium(III) acetate. In this case none of the peaks in the HPLC and mass spectra could be identified.

These results suggested that the Acm protecting group was neither appropriate for our purposes because we could not remove it in an efficient way without damaging the cycle. The

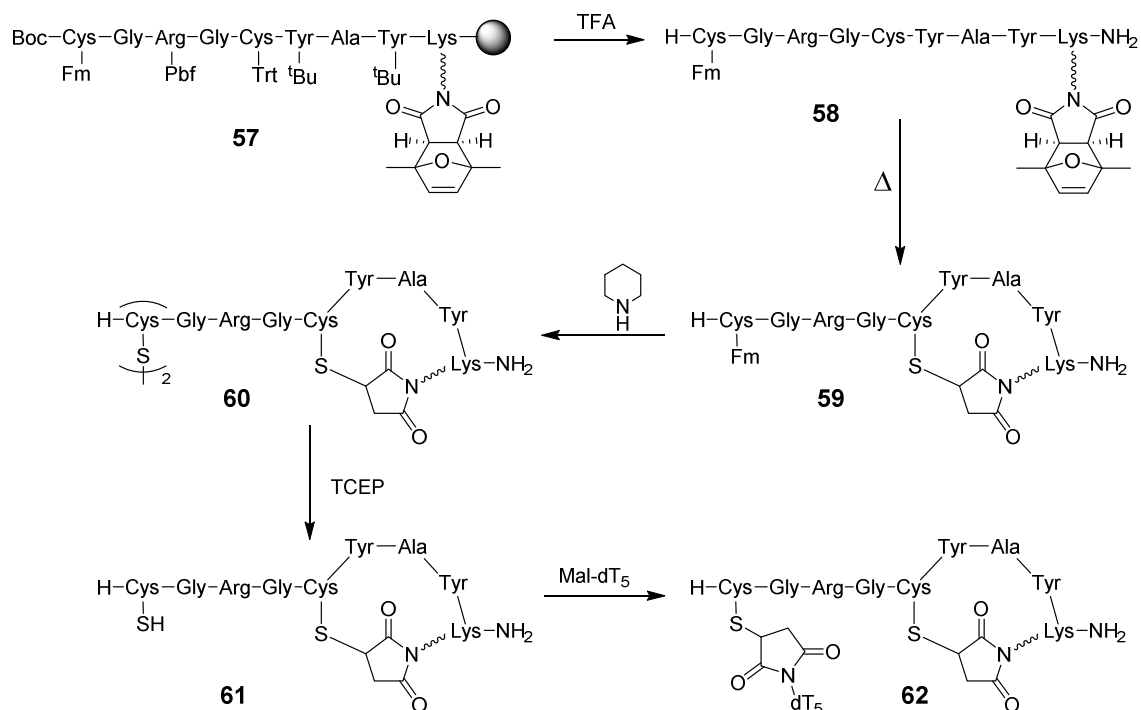
## 2. Results and Discussion

reason could be that the thioether bond resulting from cyclization is not stable to the oxidizing conditions of AcM removal. Protection of the thioether as a sulfoxide, as commonly performed in sequences with methionine residues, might be a solution, but would require an extra deprotection step.<sup>160</sup>



**Figure 2.4.4:** Attempt to prepare a cyclic peptide derivatized for conjugation (option B) employing Cys(Acm)

The last cysteine protecting group tested was fluorenylmethyl (Fm) (Fig 2.4.5). Boc-Cys(Fm)-OH was incorporated for this purpose. The Boc group would be removed during the deprotection and cleavage of the peptide from the solid support. The employment of this amino acid in an Fmoc-based solid-phase peptide synthesis has an obvious limitation. It has to be the last amino acid of the sequence, because it would be deprotected with the Fmoc removal solution, which contains piperidine, and the free thiol would perform side reactions in the next coupling step.



**Figure 2.4.5:** Synthesis of a cyclic peptide – oligonucleotide conjugate (option B) using Cys(Fm) for conjugation

The synthesis and cyclization also proceeded smoothly, affording the desired linear peptide **58** in 31 % yield and the cyclic compound **59** in 43 % yield. Removal of the Fm group was performed by dissolving the cyclic peptide in a solution of methylene chloride containing 33 % of piperidine and stirring for 45 minutes. This reaction afforded the peptide dimer (**60**) in 45 % yield, and this compound was further reduced with an excess of TCEP to finally obtain the desired functionalized cyclic peptide ready for conjugation (**61**) in 89 % yield (HPLC trace).

Compound **61** was reacted with a previously deprotected maleimido-dT<sub>5</sub> to afford the desired cyclic peptide-oligonucleotide conjugate **62**.

Overall, this strategy proved not to be the best, since it required many more steps than option A. An alternative could have been to functionalize the *N*-terminal position with an alkyne instead of a thiol, which needs not to be protected.

Option C first allowed for the synthesis of a peptide cycle with the conjugation spot within the cycle (Fig 2.4.6). In this case, the strategy was similar to option A. The cycle was generated during the cleavage event, by reaction between a free maleimide at the *N*-terminal position and a cysteine previously protected with an acid-labile group at the *C*-terminal position. The sequence incorporated an orthogonally protected cysteine, which was deprotected and used as conjugation spot.

The sequence synthesized was the same as in option A, but including Cys(S<sup>t</sup>Bu) instead of Cys(Trt) in the middle of the sequence, and Cys(Trt) instead of the maleimido derivatized lysine monomer **4** at the *C*-terminus. The final sequence was Mal-KGRG-Cys(S<sup>t</sup>Bu)-YAY-Cys(Trt)-Rink (**63**). One of the advantages of this strategy is that all the residues incorporated in the sequence are commercially available, so there is no need to synthesize special monomers for its assembly.

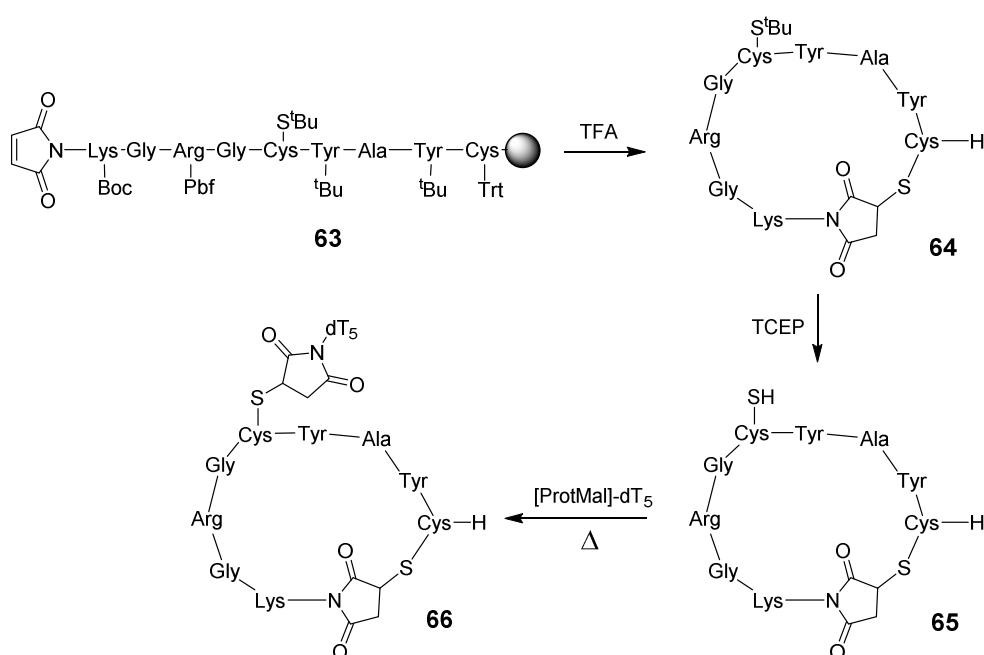


Figure 2.4.6: Synthesis of a cyclic peptide – oligonucleotide conjugate (option C)

## 2. Results and Discussion

The synthesis, cleavage, cyclization and purification afforded the desired cyclic compound (**64**) in 11 % yield. This 34-membered ring, 9 residues plus the maleimide unit, was stable to thrombin digestion conditions. This suggested that thrombin is sensitive to structure constraints even in relatively big cycles, and confirms the potential of cyclization for improving the resistance to degradation by enzymes.

The cyclic peptide was deprotected with TCEP, furnishing compound **65** in 80 % yield (HPLC trace), which was conjugated with [protected maleimido]-dT<sub>5</sub>. Equimolar quantities of both compounds were mixed and heated in a microwave oven, the protected maleimide undergoing deprotection and conjugation in a one-pot mode, yielding 77 % of the desired cyclic peptide conjugate (**66**) (HPLC trace).

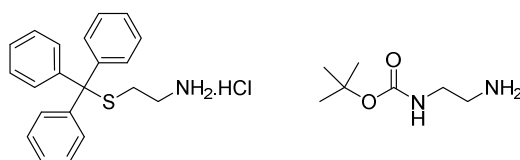
An alternative would have been to use the same protecting group scheme as in option A, just by replacing the S<sup>t</sup>Bu protected cysteine by the maleimide-containing lysine derivative (**4**). This compound could have been conjugated with a diene in one step, reducing the number of steps by 1 with respect to the synthesis performed.

As in the case of peptide conjugates described before, we decided to synthesize a peptoid analog to broaden the scope of applications of maleimide cyclization and conjugation strategy. Furthermore, we decided to use a different “click” reaction for the conjugation reaction, to prove the compatibility of the maleimide-thiol cyclization reaction with another alternative. We introduced an alkyne in the sequence, because it can undergo a Huisgen dipolar cycloaddition and potentially a Diels-Alder reaction with inverse electron demand.

As stated before, peptoids were assembled by the submonomer approach. Therefore, we needed different primary amines, incorporating the “side chains” of interest, suitably protected when required. We synthesized two different monomers for the introduction of thiols and amines into the peptoid chain.

For the introduction of a thiol we prepared *S*-trityl-cysteamine hydrochloride (**Fig. 2.4.7**), by reacting trityl chloride with cysteamine hydrochloride in DMF and precipitating the desired compound with water.

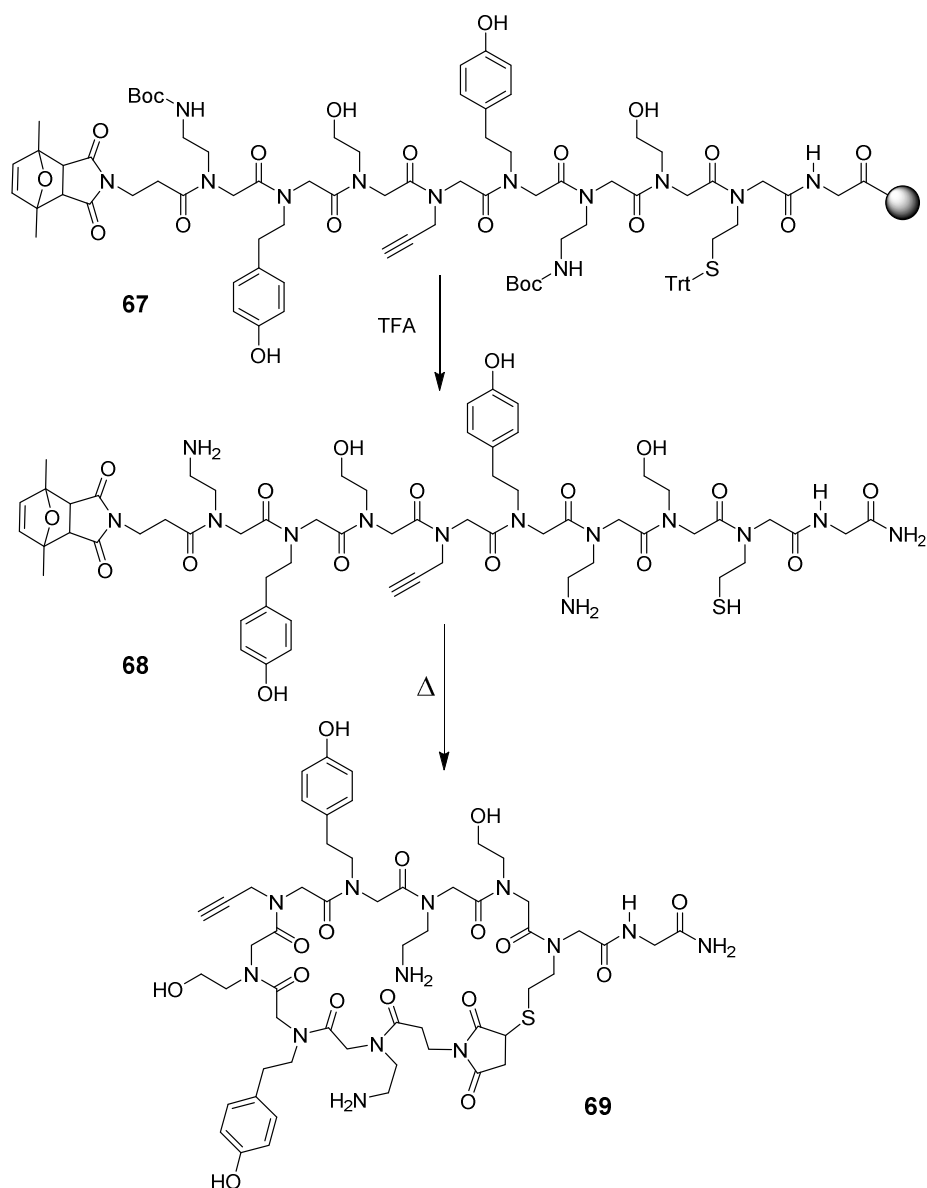
For the introduction of an amine we prepared *N*-Boc-ethylenediamine (**Fig. 2.4.7**), by adding dropwise a solution of Boc anhydride onto a solution containing an excess of ethylenediamine. The excess of ethylenediamine was removed under reduced pressure and after basification and extraction the desired compound was obtained.



**Figure 2.4.7:** Structures of *S*-trityl-cysteamine hydrochloride and *N*-Boc-ethylenediamine

The peptoid sequence synthesized is depicted in **Fig 2.4.8**. At the *C*-terminal position we incorporated a glycine residue as a spacer. The sequence included a variety of functional groups such as phenols, useful for detection and quantification, amines and alcohols. It also

contained an alkyne for conjugation and a thiol and a maleimide for cyclization. In this case, the maleimide placed at the *N*-terminal position was protected, because as stated before, cyclization crudes are much cleaner when the reaction does not occur during the cleavage and deprotection event, and its use is recommended when allowed by the synthetic scheme.



**Figure 2.4.8:** Synthesis of a cyclic peptoid derivatized for conjugation (option C)

Linear peptoid **68** was obtained in only 4 % yield. While we do not have an ultimate explanation for it, it is worth noticing that the peptoid synthesized is rather long considering the peptoids found in literature. Also the diversity of residues and the fact that alcohol residues were unprotected throughout the synthesis may have led to different side reactions. Besides, solid-phase peptoid synthesis is not as good-yielding as peptide synthesis, and maybe the synthetic protocols employed could be further optimized. Last but not least, this was one of the first peptoids synthesized, and experience and know-how is extremely valuable in these cases.



## 2. Results and Discussion

The cyclization reaction proceeded much better, the desired compound **69** was obtained in 79 % yield (HPLC trace) after heating in the microwave oven (100  $\mu$ M concentration). To our knowledge this is the largest cyclic peptoid derivative reported to date. This cycle includes 31 atoms, whilst the largest previously reported was a 27-membered macrocycle with a repetitive pattern of 1-phenylethyl side chain at all the positions.<sup>161</sup>

We decided to try two different “click” reactions with the alkyne of the cyclic peptoid, namely an inverse electron-demand Diels-Alder reaction (IEDDA) and a Cu(I)-catalyzed Huisgen dipolar cycloaddition (**Fig 2.4.9**).

In IEDDA reactions the electronic features of the diene and the dienophile are exchanged. Dienes are electron-deficient compounds, usually tetrazines, and dienophiles are electron-rich alkenes or alkynes. In the case of a tetrazine reacting with an alkene, the final compound loses aromaticity. Conversely, in the case of reaction with an alkyne, the reaction product is aromatic.<sup>162</sup> In both cases, the loss of N<sub>2</sub> renders the reaction irreversible, this fact being the driving force of the reaction.

Tetrazine substituents tune its reactivity. The more electron-withdrawing the substituent is, the more reactive the tetrazine will be. In IEDDA reaction optimization a balance has to be established, because even though tetrazines with more electron-withdrawing groups are more reactive, they are also more unstable and degrade more quickly.<sup>110</sup>

IEDDA reaction has already been used for a variety of applications, like the preparation of carbohydrate arrays<sup>163</sup> or for imaging, by incorporating tetrazine-derivatized fluorescent tags to modified antibodies.<sup>164</sup>

We attempted the IEDDA reaction with 3,6-di(2-pyridyl)-1,2,4,5-tetrazine. We mixed the cyclic peptoid with the tetrazine in equimolar quantities (125  $\mu$ M concentration) and allowed them to react for 90 minutes at room temperature. As the reaction did not advance, we heated at 50°C and left it progress overnight. HPLC analysis showed only degradation of the tetrazine, and the desired conjugate **70** was not observed.

In a different attempt, after heating the mixture at 90°C for an hour in the microwave oven, we observed a dirty HPLC crude chromatogram and the desired conjugate could neither be detected.

We think that this IEDDA reaction did not work because the alkyne was not reactive enough. First, it has been extensively reported that in IEDDA reactions strain in the dienophiles has a huge effect on the reactivity for both alkenes and alkynes,<sup>165,166</sup> and in our case the alkyne was not strained. Similar cases have been reported, in which it takes up to 20 days for reaction completion.<sup>165</sup> Second, alkynes are much less reactive than alkenes. Such is the difference in reactivity, that the Hilderbrand research group could perform a simultaneous double “click” experiment with a strained alkyne and an azide, and a strained alkene and a tetrazine, obtaining selectively two out of the four possible products.<sup>167</sup> Finally, a steric effect due to the environment of the alkyne and the geometry of the tetrazine is not to be discarded.

We also attempted a Cu(I) catalyzed Huisgen dipolar cycloaddition between the cyclic peptoid and AZT. The relevance and widespread use of this reaction has already been briefly discussed

before and will not be further examined. AZT is a 2'-deoxythymidine analog that years ago was used as a drug for the treatment of the HIV disease. It inhibits the HIV-specific reverse transcriptase, essential for the replication of retroviruses.<sup>5</sup>

Cu(I) was generated in situ by reduction of copper(II) sulfate with sodium ascorbate. Then AZT, cyclic peptoid **69** and a tristriazolylamine that stabilizes Cu(I) and prevents disproportion were added and heated in the microwave oven. Luckily, our research group had experience with this reaction for the conjugation of aminoglycosides,<sup>168</sup> and the previously optimized protocol worked properly also for cyclic peptoid **69**, yielding a 90 % of the desired conjugate **71** (HPLC trace).

The success of the Cu(I)-catalyzed Huisgen dipolar cycloaddition between a nucleoside and a cyclic peptoid confirms the compatibility between the Huisgen cycloaddition and the polyamides cyclized using the thiol-maleimide Michael-type addition. Compounds incorporating the resulting thiosuccinimide can be further conjugated under different conditions.

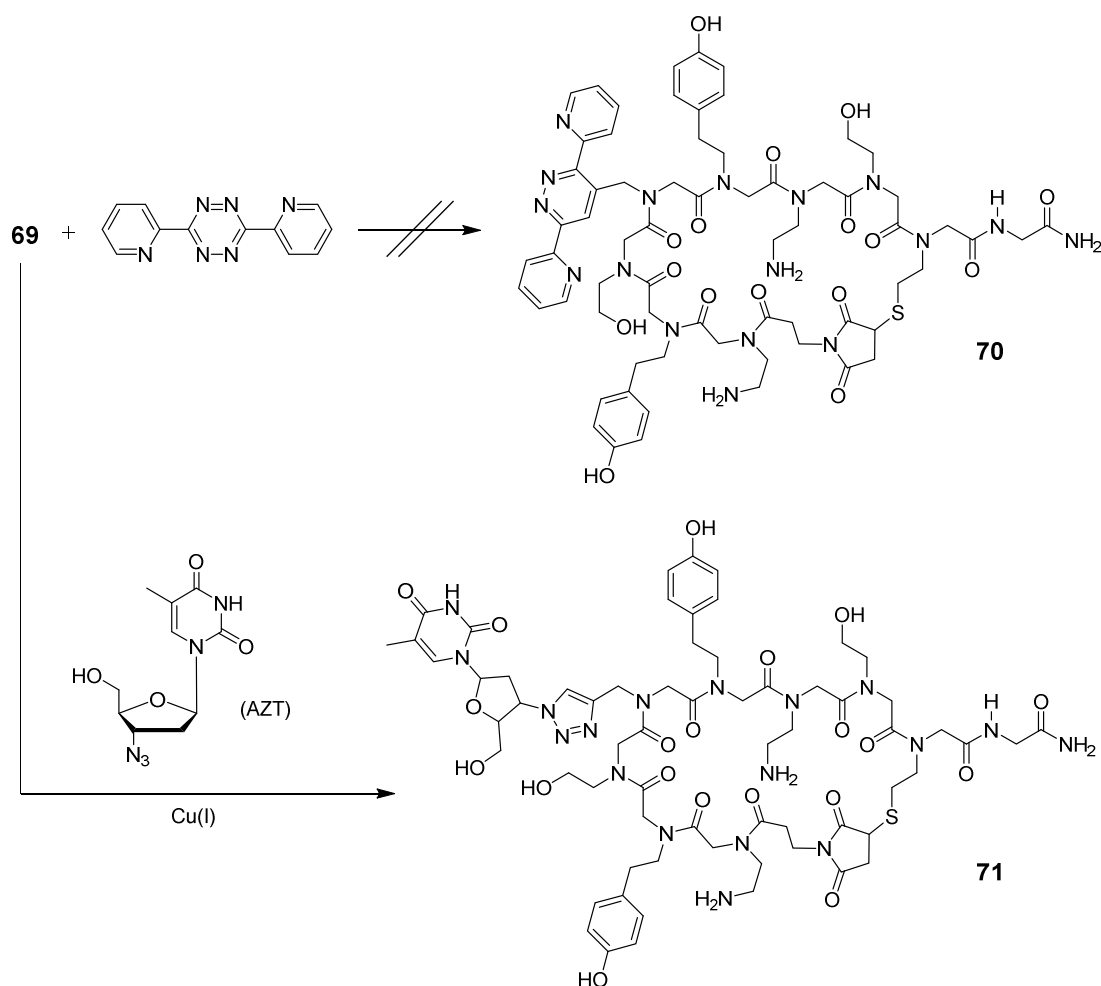


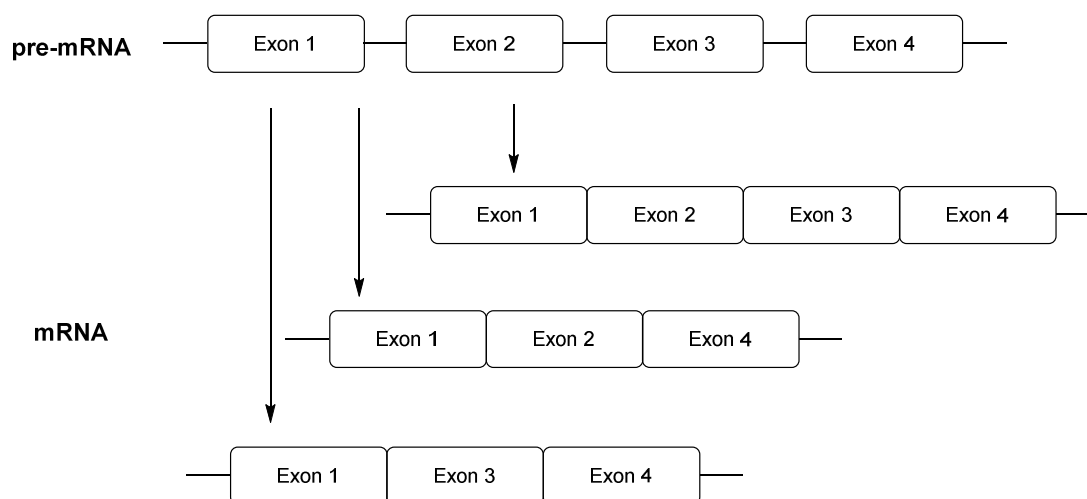
Figure 2.4.9: Alternatives tested for the conjugation of a cyclic peptoid (option C)

### 2.5. Cyclic peptide conjugates as potential splicing modulators

When DNA is transcribed, the RNA generated is still not ready for ribosomal translation. This RNA contains introns and exons, which are non-coding and coding sequences, respectively. In order to generate a mature mRNA from this pre-mRNA a process named RNA splicing has to take place,<sup>169</sup> in addition to other modifications.

The regular splicing mechanism consists in the removal of introns and fusion of exons, and is carried out by a complex machinery called spliceosome. Nevertheless, in order to generate various proteins from a single DNA precursor, one or several exons can be skipped. This process is named alternative splicing (**Fig 2.5.1**).<sup>170</sup>

Mis-splicing of pre-mRNAs can lead to a variety of diseases, either derived from genetic disorders or cancer. Hence targeting the spliceosome or any stage of the alternative splicing mechanism is of therapeutic interest.



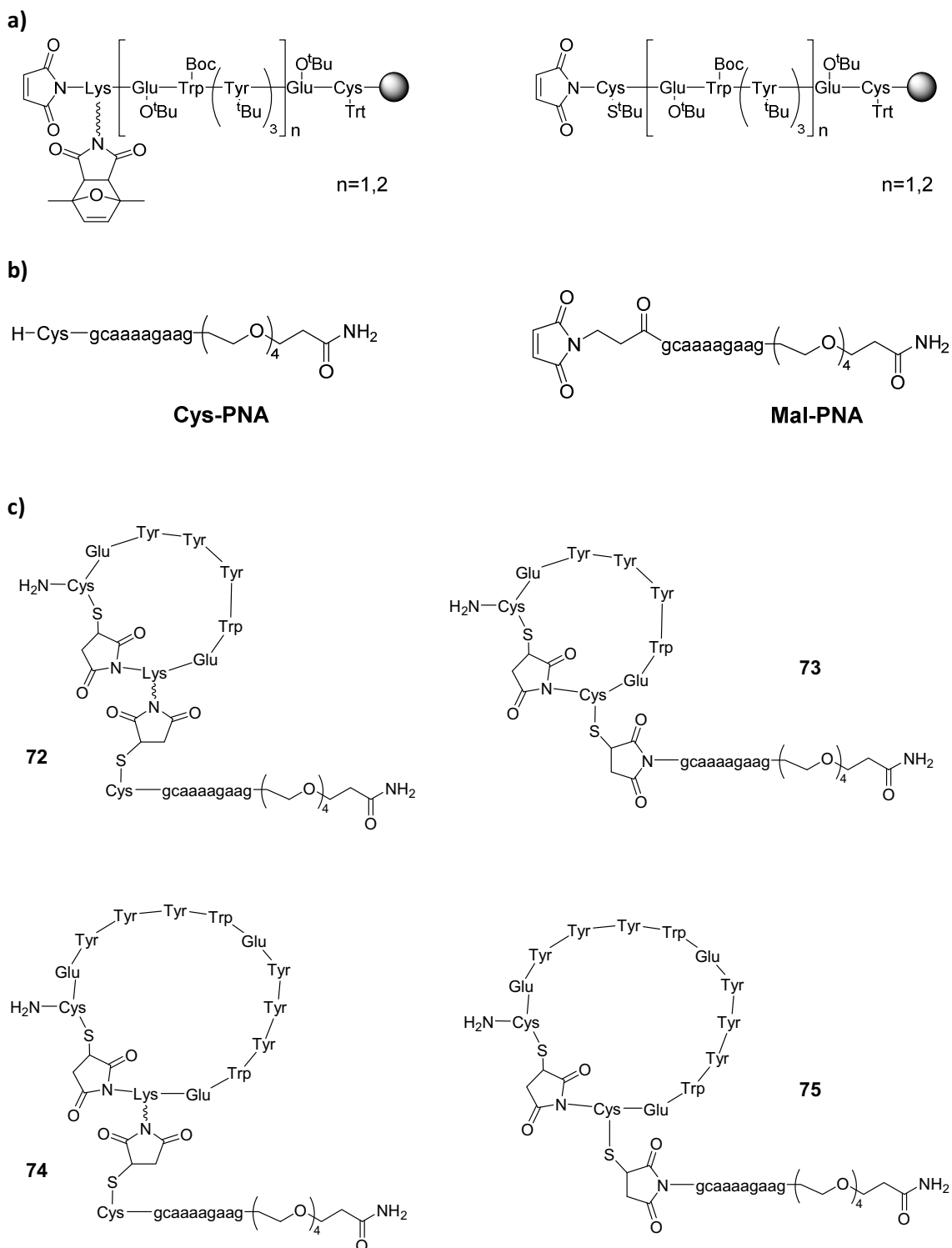
**Figure 2.5.1:** Different mRNAs may be obtained from a single pre-mRNA by alternative splicing

Previously, in our research group and in collaboration with the group of Dr. Juan Valcárcel (CRG), it was observed that PNA – peptide conjugates affected the alternative splicing of the *FAS* gene by inducing the skipping of exon 6, whereas the peptide or the PNA alone seemed to have no effect. The *FAS* gene codes for the Fas protein, which is a membrane receptor that leads to cell apoptosis, the programmed cell death, when binding a ligand. It is strongly associated with cancer, since its expression is related to tumor growth.<sup>171</sup>

Taking advantage of all the methodology developed in this part of the work, we decided to synthesize four different cyclic peptide – PNA conjugates (**72-75**), to assess whether linking a cyclic instead of a linear peptide to a PNA complementary to the Fas pre-mRNA had a differential effect on biological activity (**Fig 2.5.2**).

Given the fact that the peptide sequence known to be active contains a short sequence repeated twice, [(EYYW)<sub>2</sub>E], we produced four cyclic peptides, two of them with one repeat, and the other two with the entire sequence. Linear precursors incorporated a trityl-protected

cysteine at the C-terminal position and a free maleimide at the N-terminal position for cycle formation, and cyclization was accomplished during the cleavage and deprotection event.



**Figure 2.5.2:** **a)** Precursors of the cyclic peptides, **b)** PNAs used for conjugation, and **c)** conjugates synthesized

The precursors of one of the short-sequence cycles (**72**) and one of the entire-sequence cycles (**74**) incorporated the [protected maleimido] lysine monomer **4**. After cyclization, the maleimides were deprotected by heating in the microwave oven, and then allowed to react

## 2. Results and Discussion

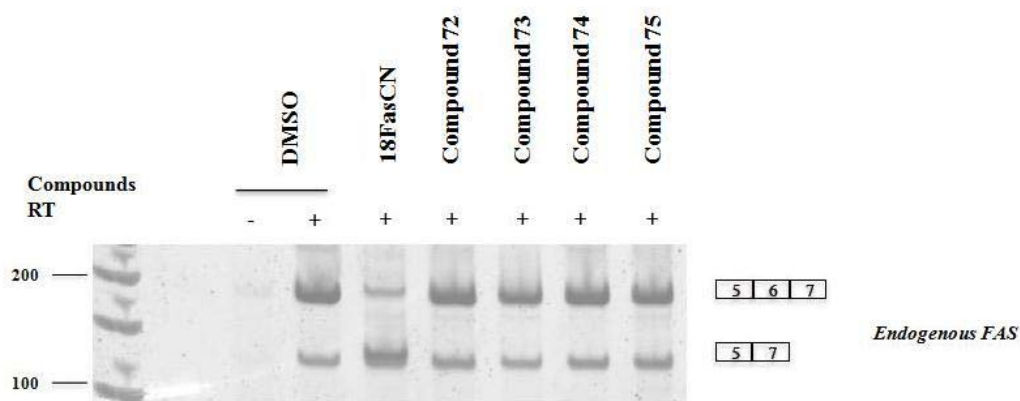
with the PNA, incorporating a free thiol at the *N*-terminal position. The protection scheme is the same as for option A in **Fig. 2.4.2**.

The other two cycles, **73** and **75**, incorporated a cysteine protected with  $S^t$ Bu instead of the [protected maleimido] lysine monomer. They were synthesized and cyclized in the same way as **72** and **74**, and cysteines were deprotected by reaction with TCEP to yield the cyclic compounds ready for conjugation. These peptides were allowed to react with the maleimido-PNA, affording the desired conjugates. The protection scheme is the same as for option C in **Fig. 2.4.6**.

For the four cyclic peptides the conjugation reaction was remarkably slow. In our previous experience with Michael reactions we had observed that the reactions took place within minutes or hours at room temperature, but in this case all of them required at least 7 days at 50°C to reach completion. We believe that the fact that there was a very small excess of maleimide or thiol, 1-2 equivalents rather than the 5-10 equivalents used in previous experiments, and the low solubility of the cyclic peptides in aqueous media compromised the kinetics of the reaction.

These four compounds were assayed in the laboratory of Dr. Valcárcel (Centre for Genomic Regulation, Barcelona). Unfortunately, the cyclic peptide – PNA conjugates appeared less active than the conjugates with the linear peptide previously tested in the group.

In the electrophoresis gel shown below (**Fig. 2.5.3**) it is clear that, even though the four compounds (**72** – **75**) display a band corresponding to the spliced RNA including exons [5] and [7], the band corresponding to the RNA with the three exons ([5][6][7]) is much more intense.



**Figure 2.5.3:** Alternative splicing assay with compounds **72** – **75**. 18FasCN corresponds to a previously assayed conjugate acting as a control

### 3. Concluding Remarks

1. 2,5-dimethylfuran is a suitable protecting group for maleimides. The *exo* cycloadduct has proven to be stable to both the Fmoc-based polyamide synthesis and the cleavage and deprotection conditions. Three building blocks, namely [protected maleimido]propanoic acid, the lysine derivative Fmoc-Lys[ProtMal]-OH and *N*-[protected maleimido]-ethanediamine have been synthesized. The first allows for the introduction of a protected maleimide at the *N*-terminal position of any polyamide. The second allows for the introduction of a protected maleimide at any position of a peptide chain, and the third is a primary amine that can be incorporated at any position of a peptoid chain.

2. Michael-type and Diels-Alder “click” conjugation reactions were carried out with peptides and peptoids containing either an *N*-terminal protected maleimide or a protected maleimide within the sequence. The extent of maleimide deprotection, either by heating the peptide or peptoid in a methanol/water mixture in the microwave oven or by heating it in toluene for longer times, was strongly substrate-dependant, but afforded fully reactive maleimides in all cases. The succinimides generated by reaction both between a maleimide and a thiol and between a conjugated diene and a maleimide were stable to maleimide deprotection conditions. One-pot deprotection and conjugation proved to be an interesting alternative to the conjugation – maleimide deprotection – conjugation methodology in case of difficult maleimide deprotections. It has the additional advantage that the number of purification steps is reduced by one.

3. By introducing both a maleimide and a thiol in a polyamide sequence (peptide or peptoid) cyclization took place via Michael addition. Peptides with protected maleimides at the *N*-terminal position afforded cleaner crudes than those with free maleimides. This cyclization strategy proved to be compatible with the unprotected side chains of all the amino acids (except cysteine, because of regioselectivity issues). Bicyclic peptides were obtained in three steps by incorporation into the chain of both a free and a protected maleimide, and two

### 3. Concluding Remarks

orthogonally protected thiols. Fused and non fused cycles could be obtained, depending on the relative positions of [protected maleimido] lysine and the differently protected cysteines.

**4.** Peptides and peptoids appended with maleimides, cysteines and a functional group for conjugation (thiol, maleimide or alkyne, suitably protected when required), were cyclized by the maleimide-thiol reaction. These cycles were further conjugated with oligonucleotides or nucleosides via “click chemistry” reactions, to afford conjugates of cyclic peptides.

**5.** Four cyclic peptide – PNA conjugates were synthesized, and their potential as splicing modulators was tested by Dr. Valcárcel’s research group. Unfortunately, none of them improved the activity of a linear peptide – PNA conjugate previously prepared in the group.

#### 4. References and Notes

- (1) Yamada, A.; Sasada, T.; Noguchi, M.; Itoh, K. *Cancer Sci.* **2013**, *104*, 15.
- (2) *Synthetic peptides: A user's guide*; Grant, G., Ed.; Oxford University Press, **2002**.
- (3) Azzarito, V.; Long, K.; Murphy, N. S.; Wilson, A. J. *Nat. Chem.* **2013**, *5*, 161.
- (4) BCN Peptides [www.bcnpeptides.com](http://www.bcnpeptides.com).
- (5) U.S. Food and Drug Administration [www.fda.gov](http://www.fda.gov).
- (6) Gran, L.; Sandberg, F.; Sletten, K. *J. Ethnopharmacol.* **2000**, *70*, 197.
- (7) Krainer, F. W.; Gmeiner, C.; Neutsch, L.; Windwarder, M.; Pletzenauer, R.; Herwig, C.; Altmann, F.; Glieder, A.; Spadiut, O. *Sci Rep* **2013**, *3*, 1.
- (8) Maune, H. T.; Han, S.; Barish, R. D.; Bockrath, M.; Iij, W. A. G.; Rothemund, P. W. K.; Winfree, E. *Nat. Nanotechnol.* **2010**, *5*, 61.
- (9) Seeman, N. C. *Annu. Rev. Biochem.* **2010**, *79*, 65.
- (10) Pirrung, M. C. *Angew. Chem.-Int. Edit.* **2002**, *41*, 1276.
- (11) Martínez, T.; Wright, N.; López-Fraga, M.; Jiménez, A. I.; Pañeda, C. *Hum. Genet.* **2013**, *132*, 481.
- (12) Praseuth, D.; Guieysse, A. L.; Helene, C. *Biochim. Biophys. Acta BBA-Gene Struct. Expr.* **1999**, *1489*, 181.
- (13) Ambesajir, A.; Kaushik, A.; Kaushik, J. J.; Petros, S. T. *Saudi J. Biol. Sci.* **2012**, *19*, 395.
- (14) Jackson, A. L.; Linsley, P. S. *Nat. Rev. Drug Discov.* **2010**, *9*, 57.
- (15) Kole, R.; Krainer, A. R.; Altman, S. *Nat. Rev. Drug Discov.* **2012**, *11*, 125.
- (16) Burnett, J. C.; Rossi, J. J. *Chem. Biol.* **2012**, *19*, 60.
- (17) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3.
- (18) Liskamp, R. M. J.; Rijkers, D. T. S.; Kruijtzter, J. A. W.; Kemmink, J. *ChemBioChem* **2011**, *12*, 1626.
- (19) Weiler, J.; Gausepohl, H.; Hauser, N.; Jensen, O. N.; Hoheisel, J. D. *Nucleic Acids Res.* **1997**, *25*, 2792.
- (20) Steer, D. L.; Lew, R. A.; Perlmutter, P.; Ian, S. A.; Aguilar, M.-I. *Curr. Med. Chem.* **2002**, *9*, 811.
- (21) Iris Biotech [www.iris-biotech.de](http://www.iris-biotech.de).
- (22) Olsen, C. A.; Montero, A.; Leman, L. J.; Ghadiri, M. R. *ACS Med. Chem. Lett.* **2012**, *3*, 749.
- (23) Roy, O.; Faure, S.; Thery, V.; Didierjean, C.; Taillefumier, C. *Org. Lett.* **2008**, *10*, 921.



#### 4. References and Notes

- (24) Chatterjee, J.; Gilon, C.; Hoffman, A.; Kessler, H. *Acc. Chem. Res.* **2008**, *41*, 1331.
- (25) Chatterjee, J.; Laufer, B.; Kessler, H. *Nat. Protoc.* **2012**, *7*, 432.
- (26) Kanthala, S.; Gauthier, T.; Satyanarayanajois, S. *Biopolymers* **2014**, *101*, 693.
- (27) Eckstein, F. **1983**, *22*, 423.
- (28) Obika, S.; Nanbu, D.; Hari, Y.; Morio, K.; In, Y.; Toshimasa, I.; Imanishi, T. *Tetrahedron Lett.* **1997**, *38*, 8735.
- (29) Campbell, M. A.; Wengel, J. *Chem. Soc. Rev.* **2011**, *40*, 5680.
- (30) Gambari, R. *Expert Opin. Ther. Patents* **2014**, *24*, 267.
- (31) Xu, F.; Pellino, A. M.; Knoll, W. *Thin Solid Films* **2008**, *516*, 8634.
- (32) Guo, J.; Hu, H.; Zhao, Q.; Wang, T.; Zou, Y.; Yu, S.; Wu, Q.; Guo, Z. *ChemMedChem* **2012**, *7*, 1496.
- (33) Viñas, M.; Rabanal, F.; Benz, R.; Vinuesa, T.; Fuste, E. In *Antimicrobial Compounds*; Villa, T. G.; Veiga-Crespo, P., Eds.; Springer Berlin Heidelberg, **2014**; pp. 269–284.
- (34) Joo, S.-H. *Biomol. Ther.* **2012**, *20*, 19.
- (35) Thorstholm, L.; Craik, D. J. *Drug Discov. Today Technol.* **2012**, *9*, e13.
- (36) Bockus, A. T.; McEwen, C. M.; Lokey, R. S. *Curr. Top. Med. Chem.* **2013**, *13*, 821.
- (37) A search of the term “cyclic peptide” results in more than 5000 hits from the last 2 years and some cyclic compounds are currently commonly used drugs.
- (38) Tulla-Puche, J.; Barany, G. *J. Org. Chem.* **2004**, *69*, 4101.
- (39) Chapman, R.; Bouten, P. J. M.; Hoogenboom, R.; Jolliffe, K. A.; Perrier, S. *Chem. Commun.* **2013**, *49*, 6522.
- (40) Nakabachi, A.; Yamashita, A.; Toh, H.; Ishikawa, H.; Dunbar, H. E.; Moran, N. A.; Hattori, M. *Science* **2006**, *314*, 267.
- (41) Plasmid <http://www.nature.com/scitable/definition/plasmid-plasmids-28>.
- (42) Danilchanka, O.; Mekalanos, J. J. *Cell* **2013**, *154*, 962.
- (43) Woodward, J. J.; Iavarone, A. T.; Portnoy, D. A. *Science* **2010**, *328*, 1703.
- (44) Kool, E. T. *Acc. Chem. Res.* **1998**, *31*, 502.
- (45) Huang, M. L.; Shin, S. B. Y.; Benson, M. A.; Torres, V. J.; Kirshenbaum, K. *ChemMedChem* **2012**, *7*, 114.
- (46) Bock, J. E.; Gavenonis, J.; Kritzer, J. A. *ACS Chem. Biol.* **2013**, *8*, 488.
- (47) Conibear, A. C.; Craik, D. J. *Isr. J. Chem.* **2011**, *51*, 908.
- (48) White, C. J.; Yudin, A. K. *Nat. Chem.* **2011**, *3*, 509.
- (49) Glas, A.; Bier, D.; Hahne, G.; Rademacher, C.; Ottmann, C.; Grossmann, T. N. *Angew. Chem.-Int. Edit.* **2014**, *53*, 2489.
- (50) Quartararo, J. S.; Wu, P.; Kritzer, J. A. *ChemBioChem* **2012**, *13*, 1490.
- (51) Dekan, Z.; Mobli, M.; Pennington, M. W.; Fung, E.; Nemeth, E.; Alewood, P. F. *Angew. Chem.-Int. Edit.* **2014**, *53*, 2931.
- (52) Singh, Y.; Palombo, M.; Sinko, P. J. *Curr. Med. Chem.* **2008**, *15*, 1802.
- (53) Camptosar prescribing information <http://labeling.pfizer.com/ShowLabeling.aspx?id=533>.
- (54) Dasari, M.; Acharya, A. P.; Kim, D.; Lee, S.; Lee, S.; Rhea, J.; Molinaro, R.; Murthy, N. *Bioconjugate Chem.* **2013**, *24*, 4.
- (55) Luo, Y.; Ziebell, M. R.; Prestwich, G. D. *Biomacromolecules* **2000**, *1*, 208.
- (56) Wang, T.; Ng, D. Y. W.; Wu, Y.; Thomas, J.; TamTran, T.; Weil, T. *Chem. Commun.* **2014**, *50*, 1116.
- (57) Manoharan, M. *Antisense Nucleic Acid Drug Dev.* **2002**, *12*, 103.
- (58) Migianu-Griffoni, E.; Chebbi, I.; Kachbi, S.; Monteil, M.; Sainte-Catherine, O.; Chaubet, F.; Oudar, O.; Lecouvey, M. *Bioconjugate Chem.* **2014**, *25*, 224.
- (59) Yan, H.; Tram, K. *Glycoconjugate. J.* **2007**, *24*, 107.
- (60) Chari, R. V. J.; Miller, M. L.; Widdison, W. C. *Angew. Chem.-Int. Edit.* **2014**, *53*, 3796.
- (61) Juliano, R.; Alam, M. R.; Dixit, V.; Kang, H. *Nucleic Acids Res.* **2008**, *36*, 4158.
- (62) Nakase, I.; Tanaka, G.; Futaki, S. *Mol. Biosyst.* **2013**, *9*, 855.

- (63) Krishnamurthy, V. M.; Estroff, L. A.; Whitesides, G. M. In *Fragment-based approaches in drug discovery*; Wiley-VCH, **2006**; pp. 11–53.
- (64) Rahim, M. K.; Kota, R.; Lee, S.; Haun, J. B. *Nanotechnol. Rev.* **2013**, *2*, 215.
- (65) Biju, V. *Chem. Soc. Rev.* **2014**, *43*, 744.
- (66) Singh, Y.; Murat, P.; Defrancq, E. *Chem. Soc. Rev.* **2010**, *39*, 2054.
- (67) Kalia, J.; Raines, R. T. *Curr. Org. Chem.* **2010**, *14*, 138.
- (68) Luo, S.; Zhang, E.; Su, Y.; Cheng, T.; Shi, C. *Biomaterials* **2011**, *32*, 7127.
- (69) Gunasekera, U. A.; Pankhurst, Q. A.; Douek, M. *Target. Oncol.* **2009**, *4*, 169.
- (70) Van Dijk, M.; Rijkers, D. T. S.; Liskamp, R. M. J.; van Nostrum, C. F.; Hennink, W. E. *Bioconjugate Chem.* **2009**, *20*, 2001.
- (71) DeForest, C. A.; Anseth, K. S. *Nat. Chem.* **2011**, *3*, 925.
- (72) Mattarella, M.; Berstis, L.; Baldrige, K. K.; Siegel, J. S. *Bioconjugate Chem.* **2014**, *25*, 115.
- (73) Roberts, K. D.; Lambert, J. N.; Ede, N. J.; Bray, A. M. *J. Pept. Sci.* **2006**, *12*, 525.
- (74) Klose, J.; Ehrlich, A.; Bienert, M. *Lett. Pept. Sci.* **1998**, *5*, 129.
- (75) Pavlova, A. S.; Vorobyev, P. E.; Zarytova, V. F. *Russ. J. Bioorganic Chem.* **2009**, *35*, 197.
- (76) Robles, J.; Pedrosa, E.; Grandas, A. *Nucleic Acids Res.* **1995**, *23*, 4151.
- (77) Kool, E. T. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 1.
- (78) Soudy, R.; Chen, C.; Kaur, K. *J. Med. Chem.* **2013**, *56*, 7564.
- (79) *Solid-phase synthesis. A practical guide*; Albericio, F.; Kates, S. A., Eds.; CRC Press, **2000**.
- (80) Di Maro, S.; Pong, R.-C.; Hsieh, J.-T.; Ahn, J.-M. *J. Med. Chem.* **2008**, *51*, 6639.
- (81) Long, Y.-Q.; Lee, S.-L.; Lin, C.-Y.; Enyedy, I. J.; Wang, S.; Li, P.; Dickson, R. B.; Roller, P. P. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2515.
- (82) Stawikowski, M.; Cudic, P. In *Peptide characterization and application protocols*; Methods in Molecular Biology; Springer, **2007**; Vol. 386, pp. 321–339.
- (83) Wang, C. C.-Y.; Seo, T. S.; Li, Z.; Ruparel, H.; Ju, J. *Bioconjugate Chem.* **2003**, *14*, 697.
- (84) Vallée, M. R. J.; Majkut, P.; Wilkening, I.; Weise, C.; Müller, G.; Hackenberger, C. P. R. *Org. Lett.* **2011**, *13*, 5440.
- (85) Kleineweischede, R.; Hackenberger, C. P. R. *Angew. Chem.-Int. Edit.* **2008**, *47*, 5984.
- (86) Stetsenko, D. A.; Gait, M. J. *J. Org. Chem.* **2000**, *65*, 4900.
- (87) Clark, R. J.; Craik, D. J. *Biopolymers* **2010**, *94*, 414.
- (88) Ishikawa, H.; Kim, S.; Kwak, K.; Wakasugi, K.; Fayer, M. D. *Proc. Natl. Acad. Sci.* **2007**, *104*, 19309.
- (89) Gray, W. R.; Luque, A.; Olivera, B. M.; Barrett, J.; Cruz, L. J. *J. Biol. Chem.* **1981**, *256*, 4734.
- (90) Saleh, A. F.; Arzumanov, A.; Abes, R.; Owen, D.; Lebleu, B.; Gait, M. J. *Bioconjugate Chem.* **2010**, *21*, 1902.
- (91) Colgrave, M. L.; Craik, D. J. *Biochemistry (Mosc.)* **2004**, *43*, 5965.
- (92) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem.-Int. Edit.* **2001**, *40*, 2004.
- (93) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem.-Int. Edit.* **2002**, *41*, 2596.
- (94) Tornøe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057.
- (95) Boren, B. C.; Narayan, S.; Rasmussen, L. K.; Zhang, L.; Zhao, H.; Lin, Z.; Jia, G.; Fokin, V. V. *J. Am. Chem. Soc.* **2008**, *130*, 8923.
- (96) Lallana, E.; Riguera, R.; Fernandez-Megia, E. *Angew. Chem.-Int. Edit.* **2011**, *50*, 8794.
- (97) Meyer, A.; Spinelli, N.; Dumy, P.; Vasseur, J.-J.; Morvan, F.; Defrancq, E. *J. Org. Chem.* **2010**, *75*, 3927.
- (98) El-Sagheer, A. H.; Brown, T. *Acc. Chem. Res.* **2012**, *45*, 1258.
- (99) Martin, M. E.; Parameswarappa, S. G.; O'Dorisio, M. S.; Pigge, F. C.; Schultz, M. K. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 4805.
- (100) Ahmad Fuaad, A.; Azmi, F.; Skwarczynski, M.; Toth, I. *Molecules* **2013**, *18*, 13148.

#### 4. References and Notes

- (101) Jagasia, R.; Holub, J. M.; Bollinger, M.; Kirshenbaum, K.; Finn, M. G. *J. Org. Chem.* **2009**, *74*, 2964.
- (102) Hoyle, C. E.; Bowman, C. N. *Angew. Chem.-Int. Edit.* **2010**, *49*, 1540.
- (103) Aimetti, A. A.; Shoemaker, R. K.; Lin, C.-C.; Anseth, K. S. *Chem. Commun.* **2010**, *46*, 4061.
- (104) Sturm, M. B.; Roday, S.; Schramm, V. L. *J. Am. Chem. Soc.* **2007**, *129*, 5544.
- (105) Forget, D.; Boturyn, D.; Defrancq, E.; Lhomme, J.; Dumy, P. *Chem.-Eur. J.* **2001**, *7*, 3976.
- (106) Lambert, J. N.; Mitchell, J. P.; Roberts, K. D. *J. Chem. Soc. Perkin Trans 1* **2001**, 471.
- (107) Tokoroyama, T. *Eur. J. Org. Chem.* **2010**, *2010*, 2009.
- (108) Mather, B. D.; Viswanathan, K.; Miller, K. M.; Long, T. E. *Prog. Polym. Sci.* **2006**, *31*, 487.
- (109) Kagan, H. B.; Riant, O. *Chem. Rev.* **1992**, *92*, 1007.
- (110) Devaraj, N. K. *Synlett* **2012**, *23*, 2147.
- (111) Steven, V.; Graham, D. *Org. Biomol. Chem.* **2008**, *6*, 3781.
- (112) Wang, Y.; Zhao, J.; Yuan, Y.; Liu, S.; Feng, Z.; Zhao, Y. *Polym. Degrad. Stab.* **2014**, *99*, 27.
- (113) Cava, M. P.; Deana, A. A.; Muth, K.; Mitchell, M. J. *Org. Synth.* **1961**, *41*, 93.
- (114) Nirogi, R.; Dwarampudi, A.; Kambhampati, R.; Bhatta, V.; Kota, L.; Shinde, A.; Badange, R.; Jayarajan, P.; Bhyrapuneni, G.; Dubey, P. K. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 4577.
- (115) Kuroda, S.; Hagiwara, T. *Polymer* **2011**, *52*, 1869.
- (116) Le Sann, C. *Nat. Prod. Rep.* **2006**, *23*, 357.
- (117) Scutaru, A. M.; Wenzel, M.; Scheffler, H.; Wolber, G.; Gust, R. *Bioconjugate Chem.* **2010**, *21*, 1728.
- (118) Kim, Y.; Ho, S. O.; Gassman, N. R.; Korlann, Y.; Landorf, E. V.; Collart, F. R.; Weiss, S. *Bioconjugate Chem.* **2008**, *19*, 786.
- (119) Kokotos, C. G. *Org. Lett.* **2013**, *15*, 2406.
- (120) Alba, A.-N. R.; Valero, G.; Calbet, T.; Font-Bardía, M.; Moyano, A.; Rios, R. *New J. Chem.* **2012**, *36*, 613.
- (121) Marchán, V.; Ortega, S.; Pulido, D.; Pedroso, E.; Grandas, A. *Nucleic Acids Res.* **2006**, *34*, e24.
- (122) Rabe, K. S.; Niemeyer, C. M. *Molecules* **2011**, *16*, 6916.
- (123) Podust, V. N.; Sim, B.-C.; Kothari, D.; Henthorn, L.; Gu, C.; Wang, C. -w.; McLaughlin, B.; Schellenberger, V. *Protein Eng. Des. Sel.* **2013**, *26*, 743.
- (124) El-Sagheer, A. H.; Cheong, V. V.; Brown, T. *Org. Biomol. Chem.* **2011**, *9*, 232.
- (125) Schoch, J.; Staudt, M.; Samanta, A.; Wiessler, M.; Jäschke, A. *Bioconjugate Chem.* **2012**, *23*, 1382.
- (126) Marchán, V.; Grandas, A. In *Current Protocols in Nucleic Acid Chemistry*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, **2001**; Vol. 31, pp. 4.32.1–4.32.31.
- (127) Sánchez, A.; Pedroso, E.; Grandas, A. *Org. Lett.* **2011**, *13*, 4364.
- (128) 5'-Maleimide-Modifier Phosphoramidite. Glen Research  
<http://www.glenresearch.com/ProductFiles/10-1938.html>.
- (129) Sánchez, A.; Pedroso, E.; Grandas, A. *Org. Biomol. Chem.* **2012**, *10*, 8478.
- (130) Sánchez, A.; Pedroso, E.; Grandas, A. *Chem. Commun.* **2013**, *49*, 309.
- (131) Sánchez, A.; Pedroso, E.; Grandas, A. *Bioconjugate Chem.* **2012**, *23*, 300.
- (132) Song, H. Y.; Ngai, M. H.; Song, Z. Y.; MacAry, P. A.; Hogley, J.; Lear, M. J. *Org. Biomol. Chem.* **2009**, *7*, 3400.
- (133) Pearson, R. J.; Kassianidis, E.; Slawin, A. M. Z.; Philp, D. *Org. Biomol. Chem.* **2004**, *2*, 3434.
- (134) Zuckermann, R. N.; Kerr, J. M.; Kent, S. B.; Moos, W. H. *J. Am. Chem. Soc.* **1992**, *114*, 10646.

- (135) Gramlich, P. M. E.; Warncke, S.; Gierlich, J.; Carell, T. *Angew. Chem.-Int. Edit.* **2008**, *47*, 3442.
- (136) Pourceau, G.; Meyer, A.; Vasseur, J.-J.; Morvan, F. *J. Org. Chem.* **2009**, *74*, 1218.
- (137) Meyer, A.; Pourceau, G.; Vasseur, J.-J.; Morvan, F. *J. Org. Chem.* **2010**, *75*, 6689.
- (138) Guiard, J.; Fiege, B.; Kitov, P. I.; Peters, T.; Bundle, D. R. *Chem. - Eur. J.* **2011**, *17*, 7438.
- (139) Ingale, S. A.; Seela, F. *J. Org. Chem.* **2013**, *78*, 3394.
- (140) Sanders, B. C.; Friscourt, F.; Ledin, P. A.; Mbua, N. E.; Arumugam, S.; Guo, J.; Boltje, T. J.; Popik, V. V.; Boons, G.-J. *J. Am. Chem. Soc.* **2011**, *133*, 949.
- (141) Meyer, A.; Vasseur, J.-J.; Morvan, F. *Eur. J. Org. Chem.* **2013**, *2013*, 465.
- (142) Aucagne, V.; Leigh, D. A. *Org. Lett.* **2006**, *8*, 4505.
- (143) Valverde, I. E.; Delmas, A. F.; Aucagne, V. *Tetrahedron* **2009**, *65*, 7597.
- (144) Kele, P.; Mezö, G.; Achatz, D.; Wolfbeis, O. S. *Angew. Chem.-Int. Edit.* **2009**, *48*, 344.
- (145) Ban, H.; Nagano, M.; Gavriilyuk, J.; Hakamata, W.; Inokuma, T.; Barbas, C. F. *Bioconjugate Chem.* **2013**, *24*, 520.
- (146) Galibert, M.; Renaudet, O.; Dumy, P.; Boturyn, D. *Angew. Chem.-Int. Edit.* **2011**, *50*, 1901.
- (147) Sharma, S. K.; Wu, A. D.; Chandramouli, N. *Tetrahedron Lett.* **1996**, *37*, 5665.
- (148) Koehler, K. C.; Alge, D. L.; Anseth, K. S.; Bowman, C. N. *Int. J. Pept. Res. Ther.* **2013**, *19*, 265.
- (149) Sieber, S. A.; Marahiel, M. A. *J. Bacteriol.* **2003**, *185*, 7036.
- (150) Qin, C.; Bu, X.; Wu, X.; Guo, Z. *J. Comb. Chem.* **2003**, *5*, 353.
- (151) Li, W.; Liu, Z.; Lai, L. *Biopolymers* **1999**, *49*, 481.
- (152) National Center for Biotechnology Information  
<http://www.ncbi.nlm.nih.gov/gene/1738>.
- (153) Smith, J. M.; Hill, N. C.; Krasniak, P. J.; Fasan, R. *Org. Biomol. Chem.* **2014**, *12*, 1135.
- (154) May, J. P.; Perrin, D. M. *Chem. - Eur. J.* **2008**, *14*, 3404.
- (155) Ten Brink, H. T.; Rijkers, D. T. S.; Kemmink, J.; Hilbers, H. W.; Liskamp, R. M. *J. Org. Biomol. Chem.* **2004**, *2*, 2658.
- (156) Garanger, E.; Boturyn, D.; Renaudet, O.; Defrancq, E.; Dumy, P. *J. Org. Chem.* **2006**, *71*, 2402.
- (157) Güell, I.; Vilà, S.; Micaló, L.; Badosa, E.; Montesinos, E.; Planas, M.; Feliu, L. *Eur. J. Org. Chem.* **2013**, *2013*, 4933.
- (158) Liu, S.; Li, D.; Huang, C.-W.; Yap, L.-P.; Park, R.; Shan, H.; Li, Z.; Conti, P. S. *Theranostics* **2012**, *2*, 589.
- (159) Häkkinen, H. *Nat. Chem.* **2012**, *4*, 443.
- (160) Andreu, D.; Albericio, F.; Solé, N. A.; Munson, M. C.; Ferrer, M.; Barany, G. In *Methods in Molecular Biology*; Humana Press Inc, **1995**; Vol. 35: Peptide Synthesis Protocols, pp. 91–169.
- (161) Butterfoss, G. L.; Yoo, B.; Jaworski, J. N.; Chorny, I.; Dill, K. A.; Zuckermann, R. N.; Bonneau, R.; Kirshenbaum, K.; Voelz, V. A. *Proc. Natl. Acad. Sci.* **2012**, *109*, 14320.
- (162) Domingo, L. R.; Picher, M. T.; Sáez, J. A. *J. Org. Chem.* **2009**, *74*, 2726.
- (163) Beckmann, H. S. G.; Niederwieser, A.; Wiessler, M.; Wittmann, V. *Chem. - Eur. J.* **2012**, *18*, 6548.
- (164) Devaraj, N. K.; Weissleder, R.; Hilderbrand, S. A. *Bioconjugate Chem.* **2008**, *19*, 2297.
- (165) Sauer, J.; Heldmann, D. K.; Hetzenegger, J.; Krauthan, J.; Sichert, H.; Schuster, J. *Eur. J. Org. Chem.* **1998**, 2885.
- (166) Chen, W.; Wang, D.; Dai, C.; Hamelberg, D.; Wang, B. *Chem. Commun.* **2012**, *48*, 1736.
- (167) Karver, M. R.; Weissleder, R.; Hilderbrand, S. A. *Angew. Chem.-Int. Edit.* **2012**, *51*, 920.
- (168) Alguacil, J.; Defaus, S.; Claudio, A.; Trapote, A.; Masides, M.; Robles, J. *Eur. J. Org. Chem.* **2010**, *2010*, 3102.
- (169) Kornblihtt, A. R.; Schor, I. E.; Alló, M.; Dujardin, G.; Petrillo, E.; Muñoz, M. J. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 153.

#### 4. References and Notes

- (170) Black, D. L. *Annu. Rev. Biochem.* **2003**, *72*, 291.
- (171) Chen, L.; Park, S.-M.; Tumanov, A. V.; Hau, A.; Sawada, K.; Feig, C.; Turner, J. R.; Fu, Y.-X.; Romero, I. L.; Lengyel, E.; Peter, M. E. *Nature* **2010**, *465*, 492.

# **PART B**



## 5. Introduction and Goals

### 5.1. RNA: Structure, function and pharmacological potential

Ribonucleic acids are complex molecules that possess a wide variety of functions and structures.

RNAs can be divided into coding and non-coding RNAs. Coding RNAs, which are essentially messenger RNAs (mRNAs), are products of DNA transcription destined to be translated by the ribosomes to assemble proteins. One of the best studied non-coding RNAs is transfer RNA (tRNA), which carries the appropriate amino acid to the ribosome during translation. Ribosomal RNAs (rRNAs) are also non-coding RNAs that act as enzymes, catalyzing amide bond formation in ribosomes. Further examples of non-coding RNAs are small nuclear RNAs (snRNAs), which are located in cell nucleus and play a vital role in the processing of mRNAs, and the more recently discovered interference RNAs (RNAis) and long non-coding RNAs (lncRNAs).<sup>1</sup>

Chemically speaking, the only differences between most RNAs and DNAs are that RNAs have hydroxyl groups at the 2' position of the ribose, whereas DNAs have hydrogen atoms, and that DNA thymines have one more methyl group than RNA uracils. Even though these differences do not seem remarkable, they confer on RNA a wide variety of activities and possible structures.

Double-stranded RNA usually adopts the A-form geometry, but it is common to find RNA as a single strand. RNA structures include loops, bulges, pockets, etc., providing a variety of complex protein-like three dimensional folding structures that are the key for their diverse biological function (**Fig. 5.1.1**).

Due to this wide range of functions and structures, RNAs are promising yet challenging targets, especially considering that there are very few structural data available.



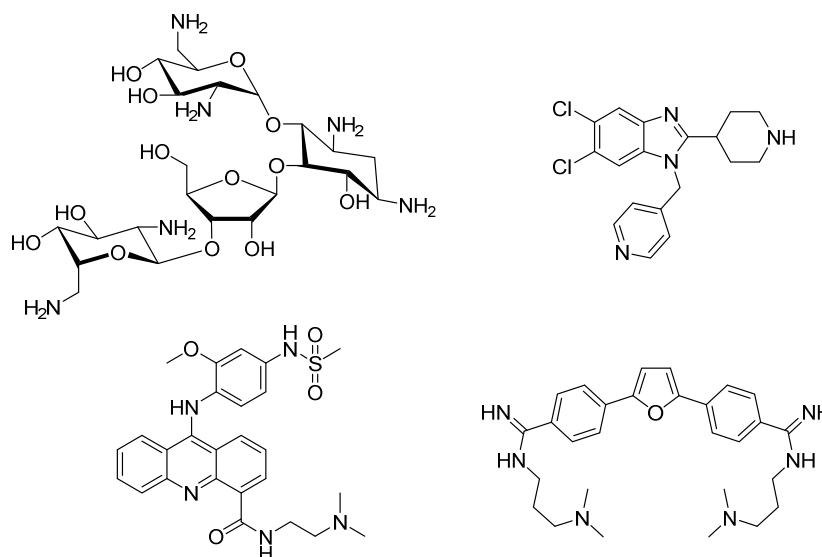
## 5. Introduction and Goals



**Figure 5.1.1:** Protein-like three dimensional structure of phenylalanine tRNA, and cartoon depicting its secondary structure<sup>2</sup>

There are several approaches for targeting RNAs and modulating their function. The employment of oligonucleotides as drugs, like in antisense or siRNA strategies, has already been briefly discussed in Section 1.1.

Another alternative is the use of small ligands (**Fig. 5.1.2**). A lot of different kinds of small molecules have been tested as potential RNA binders. Aminoglycosides, heterocyclic compounds like benzimidazoles, intercalators like acridines or furan derivatives are among them.<sup>3,4</sup> A common feature in their structure is the presence of protonable amines that can induce favorable electrostatic interactions with the RNA backbone.



**Figure 5.1.2:** Small molecules identified as RNA ligands: Neomycin B, a benzimidazole derivative, an acridine derivative and a furan-containing compound, respectively

## 5.2. Viral IRES: A target for foot-and-mouth disease

Foot-and-mouth disease (FMD) (*Aphthae epizooticae*) is a viral infectious disease that affects cloven hoofed ungulates and whose clinical symptoms are the appearance of blisters in mouth and feet. Even though the mortality of this disease is not high, it is very contagious and millions of animals have to be slaughtered when an epidemic outbreak occurs, with important economic consequences.<sup>5</sup>

Several approaches for the prevention and treatment of this disease have been considered.

Classical vaccination with an inactivated virus has been employed, but incomplete inactivation occasionally occurred, causing new focuses of infection.<sup>6</sup> Classical vaccination has evolved to the employment of epitopes, short peptide sequences that activate immunologic response without any risk of infection.<sup>7-9</sup>

A different alternative is to genetically modify animals to express an isoform of a key protein for the machinery linked to cellular translation initiation (eIF4G). This isoform has higher resistance to proteolytic degradation by an FMD virus (FMDV) protease than the natural isoform, and thus the normal cellular translation is less affected by the virus.<sup>10</sup>

The virus expresses its genetic material by producing a large polypeptide, which is cleaved afterwards by a protease to yield active peptides. Small peptide analogs have been tested, which possess an  $\alpha,\beta$ -unsaturated carbonyl that acts as Michael acceptor at the active site of this protease.<sup>11</sup>

Targeting the genetic material of the virus has also been attempted. Vagnozzi and collaborators synthesized arginine-rich peptides conjugated with DNA oligonucleotides for this purpose. The DNA sequence was complementary to mRNA sequences, either included or not in the IRES (see below), thus forming a double helix and disrupting its activity.<sup>12</sup>

Eukaryotic mRNA translation is usually cap-dependent,<sup>13</sup> which means that initiation of the translation depends on a protein machinery activated by the 5' cap. The 5' cap is a particular structure in which a triphosphate bond links the 5' end of the mRNA and the 5' position of *N*<sup>7</sup>-methylguanosine (**Fig. 5.2.1**).

The strategy of some viruses is to stop the cap-dependent translation of the hosts and initiate a different mechanism for the translation of their own genetic material, the non-cap-dependent translation. In this mechanism, the internal ribosomal entry site (IRES) within the mRNA sequence plays a key role.

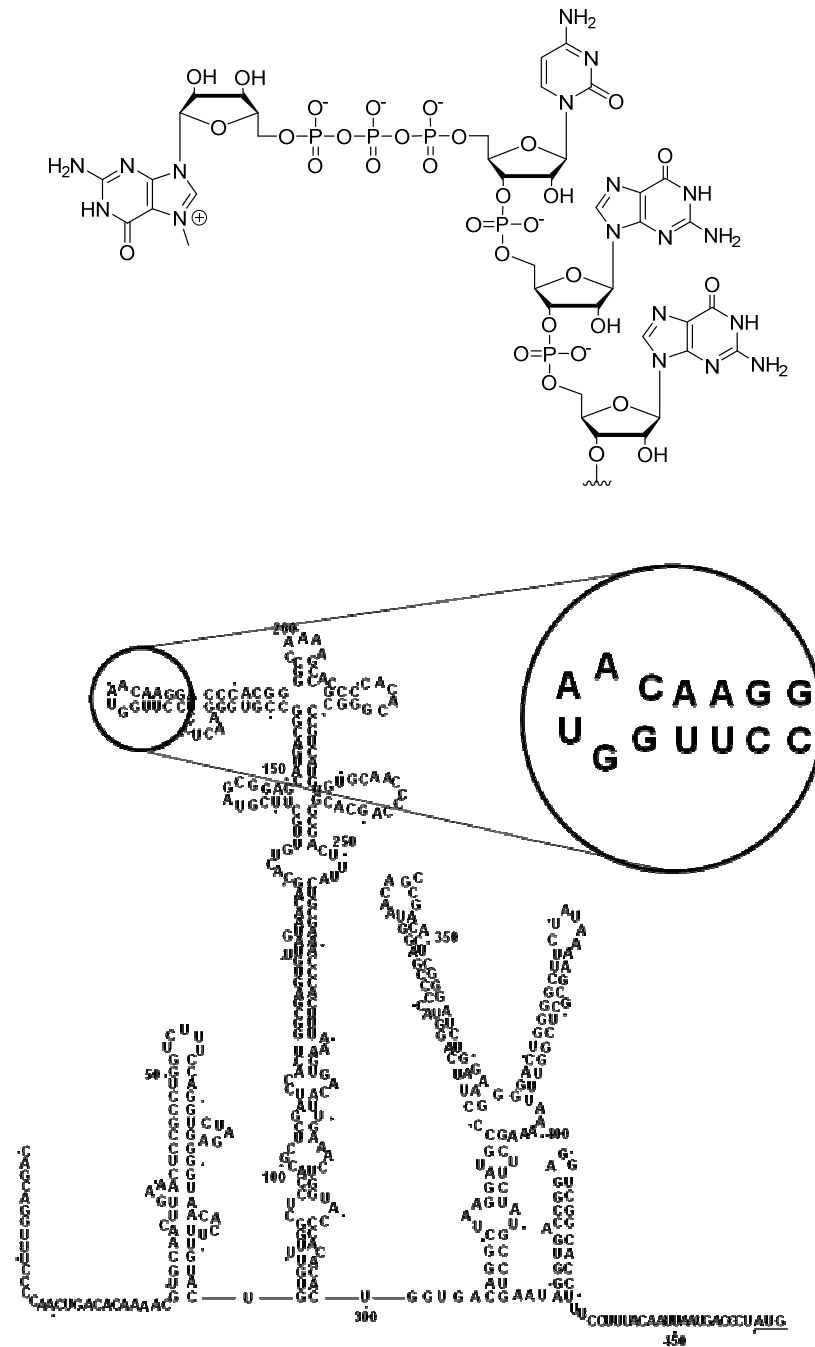
The viral IRES is a long RNA sequence with a complex three dimensional structure and several domains that binds to the 40S ribosomal subunit, and allows for the translation of mRNA without the need of a 5' cap (**Fig. 5.2.1**).

Viral IRESs are therefore interesting targets, because if a ligand could bind to a crucial part of the RNA structure, the genetic material of the virus would not be expressed and thus the infection would stop. Among all the FMDV IRES structural domains, the apical region of domain 3 is the most important. Even though little is known about the three dimensional

## 5. Introduction and Goals

structure of domain 3,<sup>14</sup> the information available indicates that it contains a GNRA loop (N = any nucleotide, R = purine), which stabilizes the RNA tertiary structure and is crucial for the IRES activity. GNRA loops have been largely studied, specially their structure and thermodynamics.<sup>15</sup>

Several GNRA loop interactions with other molecules have been reported. Among these, RNA sequences,<sup>16</sup> riboswitches,<sup>17</sup> proteins<sup>18</sup> or arginine-rich peptides<sup>19,20</sup> have been studied, but there is not much literature about small molecules that bind to GNRA tetraloops.<sup>21</sup>



**Figure 5.2.1:** 5'cap 0 structure, for cap-dependent translation (top), and FMDV IRES (bottom) with the GNRA loop highlighted

### 5.3 AD2 as a GNRA binder

One of the few reports on small molecules as GNRA binders was published in 2007 by Baranger and collaborators.<sup>22</sup> It described a series of GNRA loop binders obtained from the computational screening of a National Cancer Institute compound library. After several screening steps, a few ligands were assayed by proton NMR spectroscopy. **AD2** provided the best results (**Fig.5.3.1**).<sup>22</sup> Complexes between **AD2** and different RNA sequences containing GNRA loops were studied, from which a stoichiometry of two **AD2** molecules per loop and dissociation constants at the micromolar level were inferred.

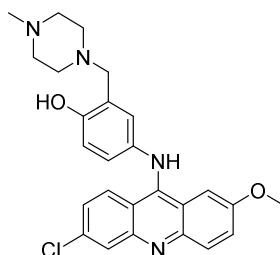


Figure 5.3.1: **AD2**

Besides its employment as an RNA loop binder, **AD2** has been also studied for other purposes. **AD2** has been tested for its potential as anticancer drug. It has been found to bind to HPSA5, a protein validated as an anticancer target,<sup>23</sup> and to interfere with Kaposi's sarcoma-associated herpes virus DNA synthesis,<sup>24</sup> or with protein-protein interactions also relevant for cancer.<sup>25</sup> **AD2** has also been tested as anti-cystic fibrosis disease drug,<sup>26</sup> as antimicrobial agent<sup>27</sup> or as antiviral, by inhibiting topoisomerase I,<sup>28</sup> but there is concern about its selectivity.

### 5.4 Goals

The main goal of this Part B was the evaluation of **AD2** as a ligand of the GUAA loop of the apical region of subdomain 3 of the IRES of the foot-and-mouth disease virus. For this purpose, the first requirement was the synthesis of **AD2**, as well as that of model RNAs.

The activity of **AD2** in cap- and IRES-dependent mRNA translation was studied, as well as the binding site of **AD2** in the FMDV IRES, which was assessed by SHAPE assays. These experiments were carried out at the laboratory of Prof. Encarnación Martínez-Salas (CBMSO, Madrid).

For a better understanding of the **AD2** interaction with RNA, several biophysical experiments were planned, including RNA thermal denaturation experiments, circular dichroism, isothermal titration calorimetry (ITC), electrospray ionization (ESI) or fluorescence titration. Only the latter required RNA to be labeled.

It is finally worth mentioning that during my short stay at Prof. Lincoln's lab (Chalmers University of Technology, Gothenburg), experiments aimed at examining the affinity of **AD2** for DNA were carried out. These experiments will be discussed in Appendix 5.



## 6. Results and discussion

### 6.1. AD2 synthesis and characterization, and RNA synthesis

**AD2** synthesis was the first requirement for this project. It was planned as a 4 step route, of which only parts had been previously described in the literature but not the whole scheme. It is depicted in **Fig 6.1.1**, and it starts with *p*-aminophenol. The first two steps were already described and could be easily reproduced.<sup>29</sup> The amino group of *p*-aminophenol was protected by acetylation. Acetic anhydride is a cheap reagent, the reaction is easy and the acetyl group is known to be much more stable than other protecting groups like *Boc* (known to be partially removed when heated). **Prec-1**, also known as paracetamol, was obtained in high yield uneventfully.

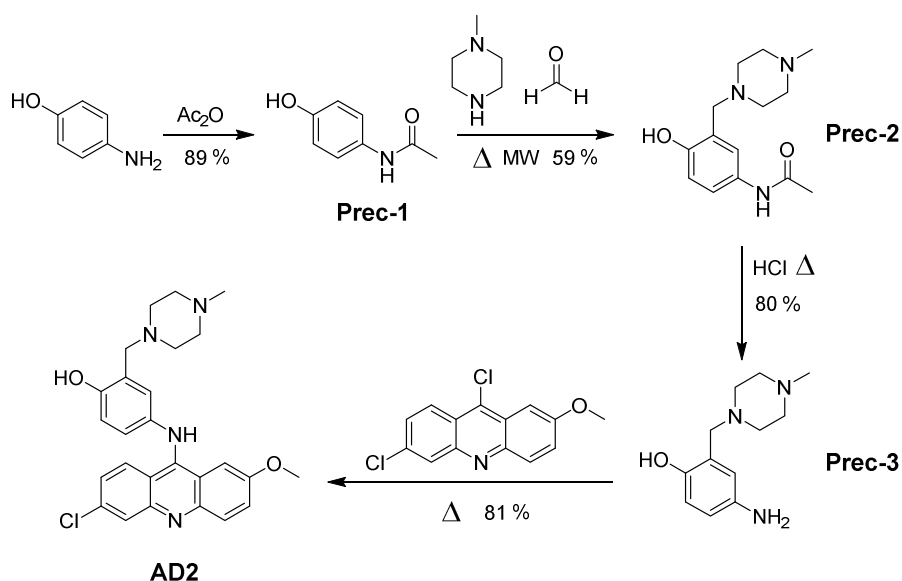


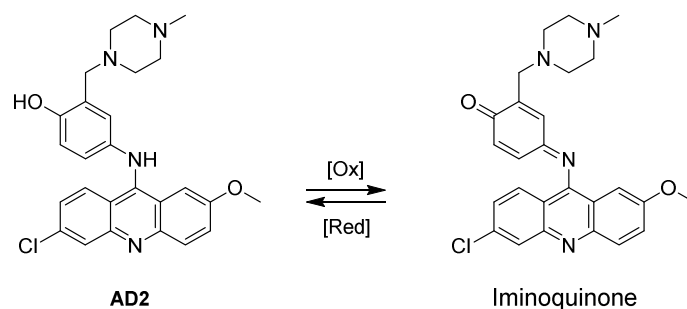
Figure 6.1.1: AD2 synthesis in four steps

## 6. Results and Discussion

The second step was a Mannich reaction between the acetylated aminophenol (**Prec-1**), *N*-methylpiperazine and formaldehyde. The three compounds were mixed in ethanol and heated in a microwave oven at 100 W for 3 minutes. After work-up and column chromatography purification, the desired compound (**Prec-2**) was obtained in moderate yield. Further hydrolysis of the acetamido group in acidic conditions yielded **Prec-3**, which reacted with 6,9-dichloro-2-methoxyacridine through an aromatic nucleophilic substitution at carbon 6.<sup>30</sup> This reaction afforded an orange compound that corresponded to **AD2**. **AD2** was obtained in 34 % overall yield.

A singularity found upon **AD2** structural characterization is its mass spectrum. The **AD2** structure was unmistakably confirmed by NMR spectroscopy, but in mass spectrometric analysis, by employing both electrospray and MALDI ionization techniques, we observed two peaks with similar intensities, one with the expected mass and the other with the expected mass minus two units. We suspected that the phenol ring of **AD2** was being oxidized to iminoquinone (**Fig. 6.1.2**) during ionization, because it has been reported that this kind of oxidation reactions happens rather easily.<sup>31</sup>

The oxidation reaction during ionization could be avoided in MALDI-TOF MS analysis by including ascorbic acid, a mild reducing agent, in the matrix mixture.



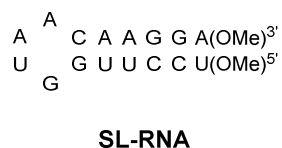
**Figure 6.1.2:** Oxidation of **AD2** to iminoquinone

Further characterization of **AD2** was carried out by the determination of the pKa value of the nitrogen of the acridine. The pKa value was determined by UV-Vis spectroscopy, taking advantage of the fact that protonated and deprotonated acridines absorb light at different wavelengths, thus yielding different UV-Vis spectra. The methodology employed<sup>32</sup> is explained at the experimental section. The pKa value obtained was 6.1.

In order to characterize the **AD2** binding with the GUAA tetraloop, we synthesized the stem-loop fragment of the apical region of the domain 3 of FMDV IRES, which contains this tetraloop. We employed this synthetic RNA in UV-monitored thermal denaturation curves, circular dichroism (CD), isothermal titration calorimetry (ITC) and electrospray ionization (ESI) experiments.

The sequence synthesized was a 16mer hairpin, in which, once annealed, four bases constituted the tetraloop and the other 12 formed 6 base pairs. Furthermore, nucleosides *O*-methylated at the 2' position were introduced at both the 3' and 5' ends of the sequence. This methylation confers higher stability towards enzymatic degradation. The sequence synthesized

(Stem-Loop RNA, **SL-RNA**) in its predicted secondary structure is shown in **Fig. 6.1.3**. The RNA synthesis protocol and compound data (HPLC and yield) are described in Section 8.

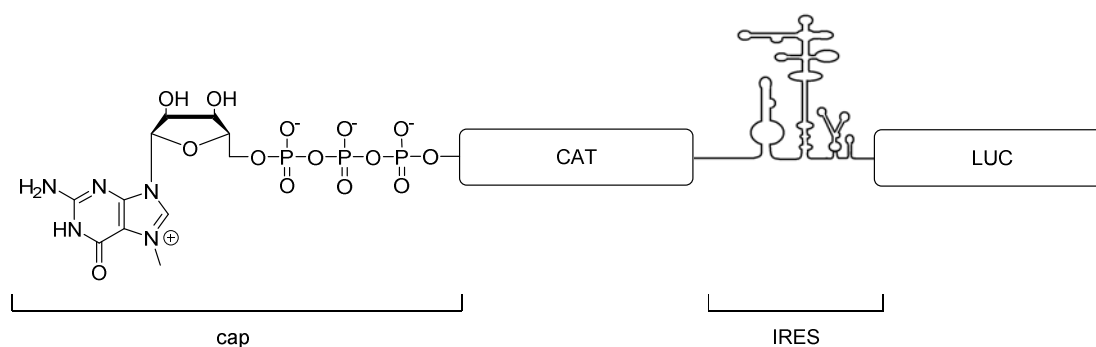


**Figure 6.1.3:** RNA sequence synthesized

## 6.2. *In vitro* results

Thanks to a collaboration established with the laboratory of Prof. Encarnación Martínez-Salas from *Centro de Biología Molecular Severo Ochoa (Universidad Autónoma de Madrid)*, FMDV IRES-dependent translation inhibition by **AD2** could be tested *in vitro*. They carried out several experiments with the entire IRES sequence, first to assess whether it inhibited IRES-mediated translation and then to determine the **AD2** binding site.

**AD2** translation inhibition was tested by employing a bicistronic RNA, depicted in **Fig. 6.2.1**. The bicistronic RNA codes for two different proteins, namely chloramphenicol acetyltransferase (CAT) and luciferase (LUC), which are two widely used reporting proteins.<sup>33,34</sup> In the absence of an IRES-interacting ligand, both CAT and LUC should be expressed, because both cap- and IRES- dependent translation would be taking place. Upon addition of **AD2**, a differential effect on cap- and IRES- dependent translation should be observed. This would mean that **AD2** blocks preferentially the IRES-dependent translation.



**Figure 6.2.1:** Schematic structure of the bicistronic RNA used in *in vitro* assays

What was observed was a decrease in both LUC and CAT expression, slightly more accentuated for LUC (**Fig. 6.2.2**). This may be explained as the result of unspecific interactions of **AD2** with different RNA regions, likely single-stranded as well as double-stranded, which would affect the translation by the ribosomes.



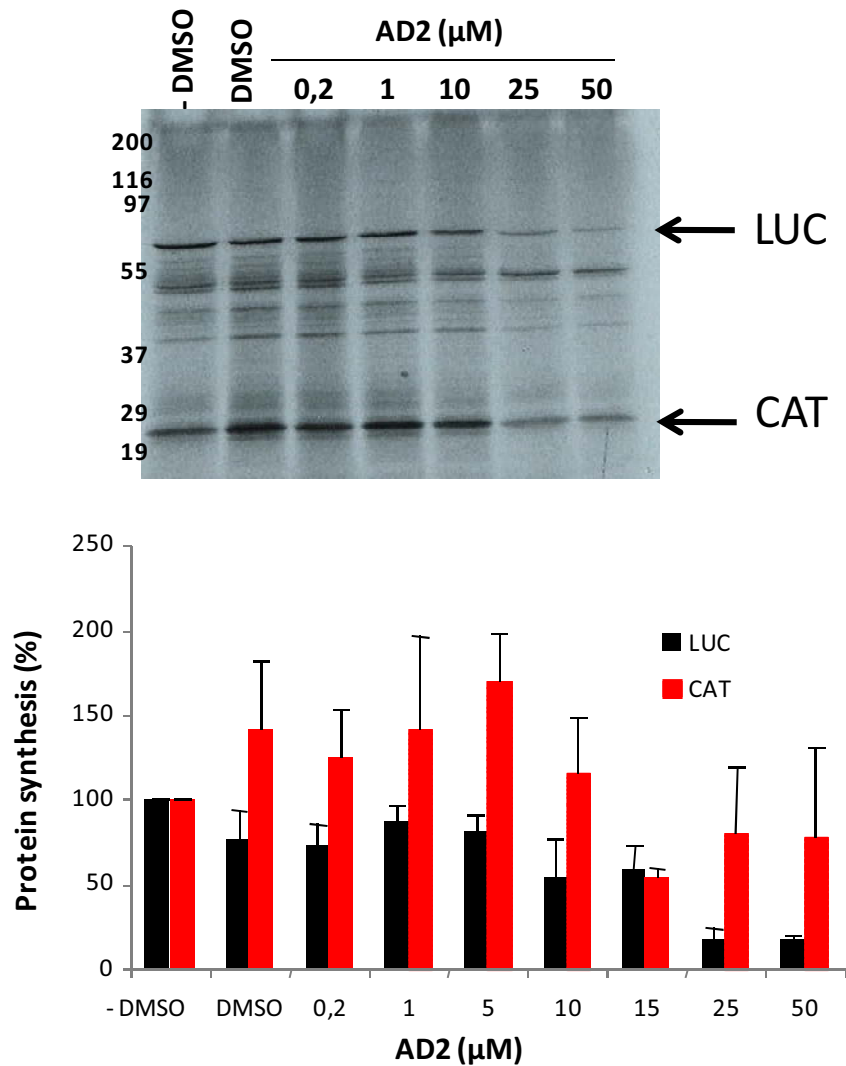


Figure 6.2.2: Translation inhibition assay with the bicistronic RNA and AD2

Given the fact that the translation inhibition assay displayed some differential activity for cap- and IRES-mediated translations, it was decided that AD2 binding site determination was worth the effort. It was performed with an assay named SHAPE. SHAPE stands for “selective 2'-hydroxyl acylation analyzed by primer extension”, and it is a technique employed for long RNA secondary and tertiary structure determination. The idea is to react the structured RNA in the presence and absence of the ligand with an acylating agent, namely *N*-methylisatoic anhydride (NMIA) (Fig. 6.2.3). This reagent reacts with non-constrained 2'-hydroxyls. The acylated IRES is then annealed with a DNA primer, and a reverse transcriptase is added to produce RNA chains until a modification point. The structural information is finally obtained by analysis of the length and abundance of transcribed DNA sequences.<sup>35</sup>

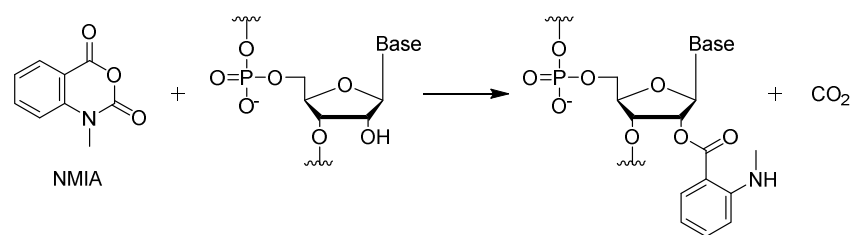


Figure 6.2.3: Reaction occurring in SHAPE assays

Unfortunately, SHAPE could not unequivocally establish the **AD2** binding site (Fig. 6.2.4). This can be related with the promiscuous nature of **AD2**. It binds to several places, thus not yielding a clear and unique signal.

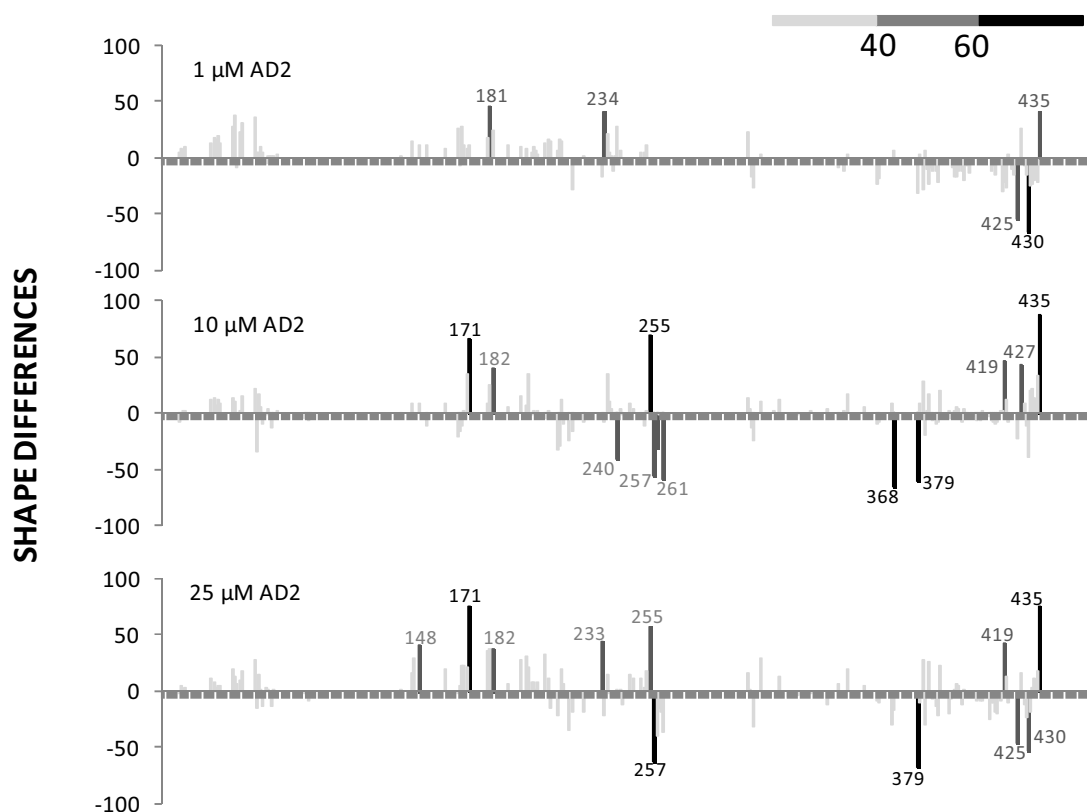


Figure 6.2.4: SHAPE differences upon addition of **AD2**

### 6.3. Absorption spectroscopy assays

In order to assess if a certain ligand binds to an oligonucleotide sequence, a very common experiment is the UV-monitored thermal denaturation. Upon heating, the two strands of a nucleic acid duplex are not separated and converted into single strands instantly. It is a cooperative process that takes place in a range of about  $30^\circ\text{C}$ . The temperature at which half of the oligonucleotide duplex is melted is considered the  $T_m$ . Ultimately, the  $T_m$  gives an idea of the double helix stability. Buffer, pH and saline content of the sample play a crucial role in the stability of the oligonucleotide and have to be chosen carefully.

## 6. Results and Discussion

As stated above, this process can be monitored by UV-Vis spectroscopy. Oligonucleotide chains extinction coefficient are lower when bases are paired in double helices. By constantly irradiating the sample at  $\lambda = 260$  nm, and heating or cooling it progressively (from 20 to 90°C and vice versa), a series of experimental points are obtained, which fit nicely into a sigmoid curve.

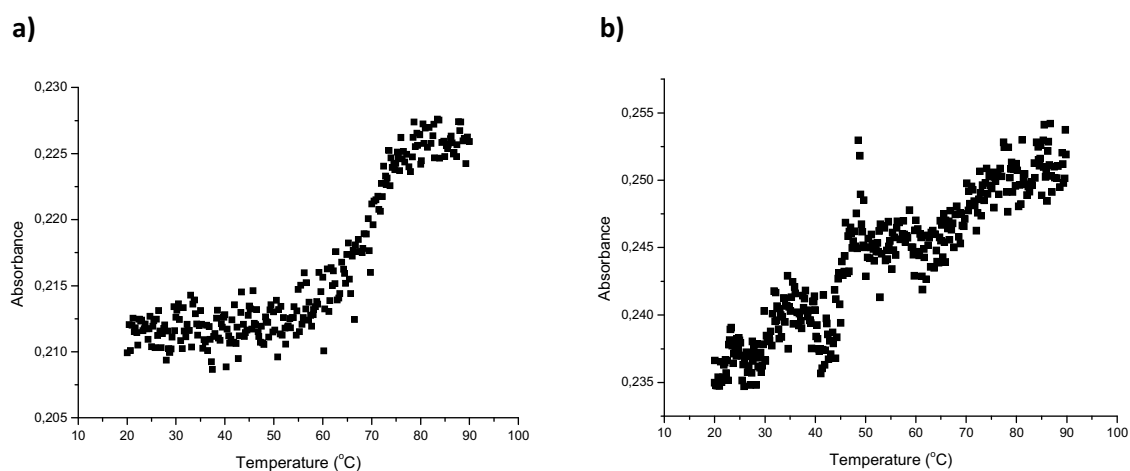
Depending on the binding mode and the structure, some ligands stabilize or destabilize the double helix structure, thus increasing or decreasing its  $T_m$ , respectively. Nevertheless, some ligands do not bind directly to the double helix or do not modify its stability and, for this reason, the experiment outcome has to be considered cautiously.

We performed several  $T_m$  measurements with increasing amounts of **AD2**, from 0 to 2 equivalents (see Appendix 4 for experimental details). The experimental results are summarized in **Table 6.3.1**.

Equivalents of <b>AD2</b>	$T_m$ values obtained (°C)			Average	Standard Deviation
0	71.7	71.1	70.3	71.0	0.7
0.2	66.8	69.7	75.2	70.6	4.3
0.5	59.5	60.9	70.1	63.5	5.8
1	69.0	71.7	64.7	68.5	3.5
2	69.7	/	/	69.7	/

**Table 6.3.1:** Melting point determination (results of three experiments)

The experimental points of the melting curves of **SL-RNA** upon addition of **AD2** presented a huge dispersion, and the last measurements did not adjust properly to a typical sigmoid equation (**Fig. 6.3.1, b**). We also did not observe a clear tendency in the changes of  $T_m$  values upon addition of **AD2**. There seems to be a slight tendency to decrease, but the standard deviation of the measurements is larger than the variations of the  $T_m$  values of the different titration experiments. Examples of the raw data from **SL-RNA** before addition of **AD2** and **SL-RNA** with two equivalents of **AD2** are shown in **Fig. 6.3.1**.



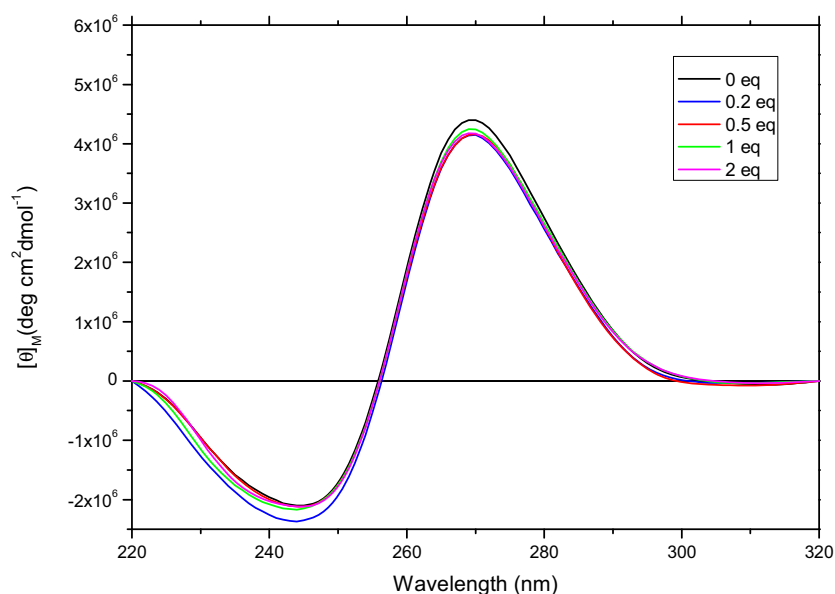
**Figure 6.3.1:** Thermal denaturation experiments with **SL-RNA**, a) without **AD2**, and b), with two equivalents of **AD2**.

Even though the electrospray ionization experiments, which will be explained in due time, assessed that more than two **AD2** molecules bind to **SL-RNA**, we decided not to continue increasing the amount of **AD2** in this experiment, because the data obtained with this technique were not providing any valuable information. We hypothesize that the dispersion observed upon addition of **AD2** is a consequence of the formation of different **SL-RNA – AD2** complexes.

We next carried out circular dichroism experiments. Circular dichroism (CD) is a technique in which structure variations of macromolecules like proteins or oligonucleotides can be observed. The technique consists in measuring the difference between the absorbance of a sample irradiated with left and right circularly polarized light. Given a chiral secondary or tertiary structure, left and right circularly polarized lights are not equally absorbed. A variation of the macromolecule structure can produce a variation of the absorption and thus different CD spectra.<sup>36</sup> CD spectra are recorded over a range of wavelengths, which in our case was between 220 and 320 nm.

We recorded the CD spectrum of **SL-RNA** with increasing amounts of **AD2**, from 0 to 2 equivalents (see Appendix 4 for experimental details), but we did not observe any significant variation in shape or intensity (**Fig 6.3.2**).

The fact that the CD spectra do not vary upon addition of ligand does not mean that the ligand does not bind to the target. On the basis of Prof. Baranger's results, **AD2** is expected to bind the loop, but it could also behave as intercalator because it contains an acridine moiety. It is described that the CD spectra of double-stranded RNAs changes upon addition of intercalators,<sup>37</sup> so in our case we hypothesize that, upon addition of **AD2**, the different complexes generated from interaction with both the loop and the stem may have different effects on the CD spectra, the combination of whom does not show major variations.



**Figure 6.3.2:** CD spectra of **SL-RNA** with increasing amounts of **AD2**

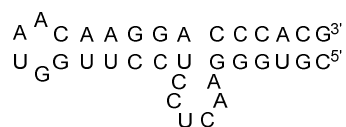
#### 6.4. Isothermal titration calorimetry experiments

In standard ITC experiments the target compound (oligonucleotide, protein...) is located in the sample cell. A burette automatically adds the ligand, a few microliters at each step. The interaction between ligand and target has an associated heat, which is related to a variation of the cell temperature.

ITC instruments measure the electrical power that has to be applied to keep the reference cell and the sample cell at the same temperature at each titration step. This electrical power is converted into enthalpy values.

ITC is a tag-free technique, that is, there is no need to label the ligand or the target. With an ITC experiment all the thermodynamic parameters of a ligand-target interaction can be obtained. The area of the peaks correlates with the binding enthalpy. By adjusting the curve obtained from the areas of the peaks to binding model equations, stoichiometry and binding constant values can be determined. Finally, if the binding constant is converted to a  $\Delta G$  value, the entropy variation of the interaction event can also be disclosed.

We first attempted a titration with **SL-RNA** (Fig. 6.1.3) at 25°C (see Appendix 4 for experimental conditions). We observed a hint of a tendency, the areas of the peaks progressively decreased, but these changes were tiny and we could not extract any information (Fig. 6.4.2, a). We next decided to employ a longer RNA sequence, available in the group, to see if we could obtain more information on the binding of **AD2** to a larger section of the viral IRES. The longer sequence, a stem-loop-bulge RNA (**SLB-RNA**) is depicted in Fig. 6.4.1. By employing **SLB-RNA** we observed a slightly better ITC thermogram than with **SL-RNA**, but the heat rate variation was extremely small (Fig. 6.4.2, b).

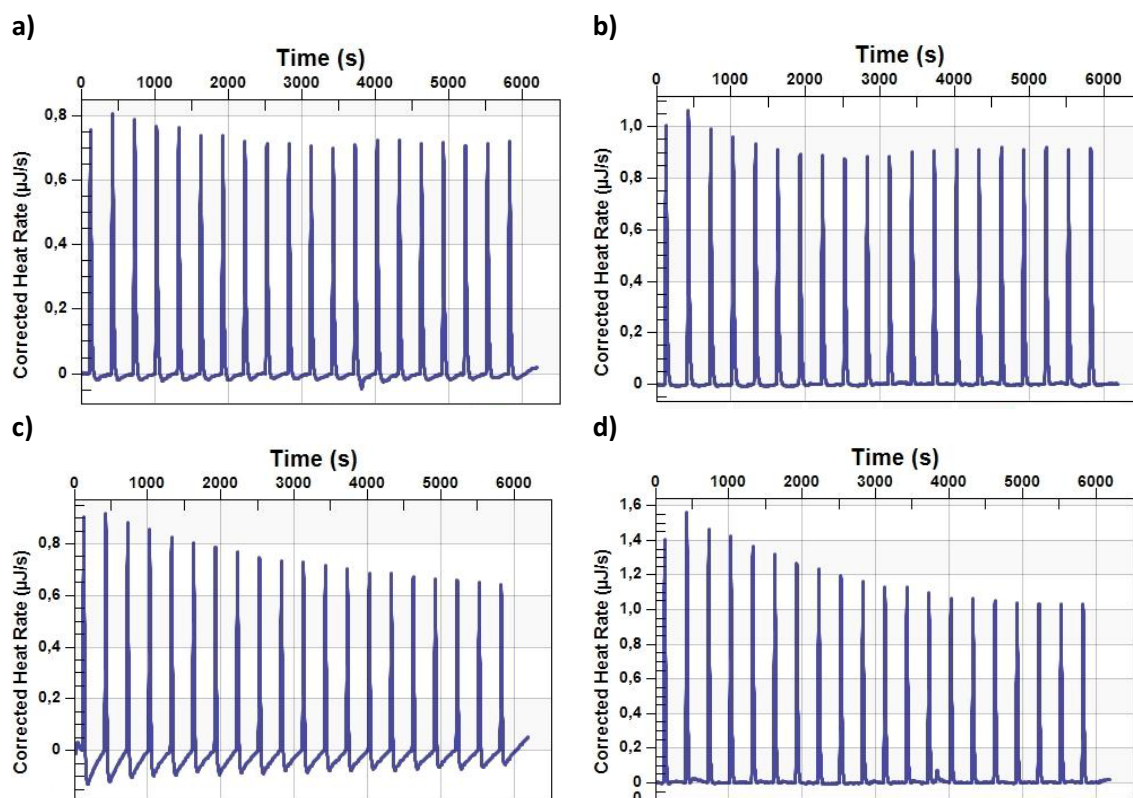


**SLB-RNA**

**Figure 6.4.1:** Longer RNA sequence, containing a stem-loop and a bulge

The same assays were then performed at a higher temperature, because the heat capacity, and thus the heat of binding, is temperature-dependent, and a slightly better signal-to-noise ratio can be obtained. When repeated at 40°C, both **SL-RNA** and **SLB-RNA** afforded better thermograms (Fig. 6.4.2, c and d), but heat changes were very small. This suggests that the interaction is entropically-driven, and thus not detected, because the ITC instrument solely detects heat, which is related to the  $\Delta H$  term.

Despite the fact that the thermograms obtained at 40°C were much better, they were still not good enough to be properly adjusted to a mathematical model and no quantitative data could be obtained.



**Figure 6.4.2:** ITC experiments thermograms: **a)** and **b)** correspond to **SL-RNA** and **SLB-RNA** respectively upon addition of 10 equivalents of **AD2** at **25°C** **c)** and **d)** are the same as **a)** and **b)** but at **40°C**

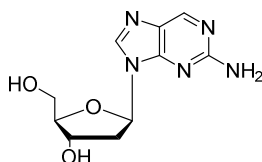
### 6.5. Fluorescence assays

Fluorescence is a physical phenomenon in which a given compound absorbs light and thus achieves an excited state, but instead of relaxing by losing the energy excess as heat, the compound expels it as lower frequency light.

Fluorescence spectroscopy is an extremely sensitive technique, which can be employed for the study of target – ligand interactions. This technique is based on the fact that the fluorescence intensity of a given sample, this is, the amount of light that the sample emits (perpendicularly to an incident light), changes depending on the environment of the fluorophore. Hence, if during the interaction between ligand and target the environment of the fluorescent moiety changes, so will do the fluorescence intensity.

The option chosen was RNA labeling. We decided to employ 2-aminopurine (Ap) (**Fig. 6.5.1**), which can replace an adenine residue without much distortion with respect to the original target structure. Ap fluorescence is quenched by stacking in the presence of aromatic rings, such as purines or pyrimidines. On the other hand, Ap fluorescence is known to be higher when surrounded by water.<sup>38</sup> If a ligand interacts with the RNA nearby the Ap and expels the Ap residue from within the structure, the fluorescence intensity is expected to increase. Conversely, if the ligand is incorporated into the structure so that stacking interactions with the Ap residue are established, the fluorescence is expected to be quenched and thus the signal will decrease.

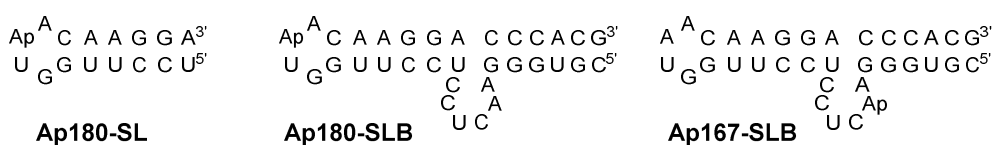
## 6. Results and Discussion



**Figure 6.5.1:** 2-Aminopurine 2'-deoxynucleoside

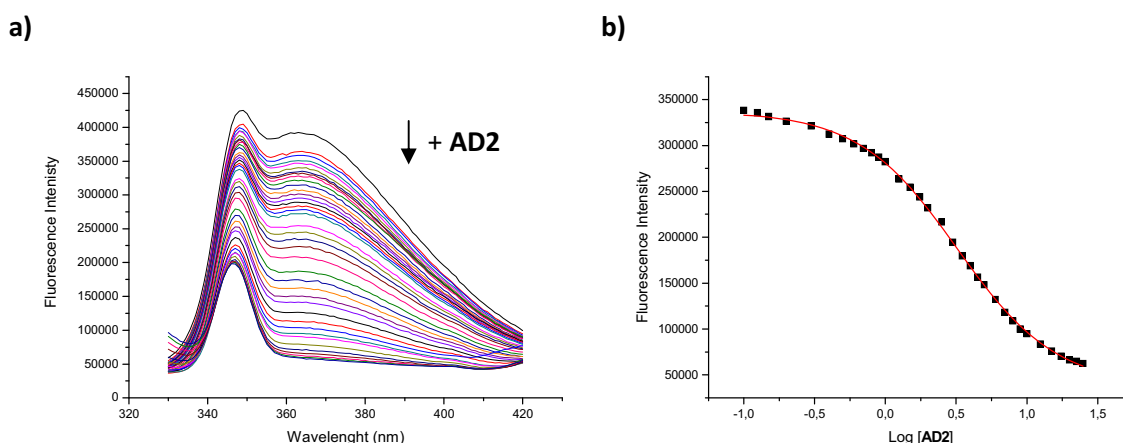
The second scenario was the one observed for **AD2** titrations (see Appendix 4 for experimental conditions), as the fluorescence intensity decreased upon addition of **AD2**. The experiment consisted in placing the labeled RNA in the cell and measuring the fluorescence intensity at a range of wavelengths after additions of aliquots of an **AD2** solution.

As we had observed interaction for both **SL-RNA** and **SLB-RNA** in ITC experiments, we decided to study these sequences using their labeled analogs, **Ap180-SL** and **Ap180-SLB**, respectively. We also decided to examine the same **SLB-RNA** sequence but labeled in the bulge (**Ap167-SLB**), to evaluate the affinity of **AD2** for the bulge. In summary, three different fluorescent targets were tested, namely **Ap180-SL**, **Ap180-SLB** and **Ap167-SLB** (Fig. 6.5.2). An example of the spectra obtained upon titration is shown in Fig. 6.5.3, a.



**Figure 6.5.2:** Labeled RNA sequences for fluorescence assays

The data obtained for the three sequences were processed in order to obtain  $EC_{50}$  values.  $EC_{50}$  stands for half maximal effective concentration, and its value corresponds to the ligand concentration required for a response halfway between the maximum and the baseline. The data fitted nicely to a sigmoid equation (Fig. 6.5.3, b), and the  $EC_{50}$  values obtained are collected in Table 6.5.1.



**Figure 6.5.3:** Typical fluorescence spectra (a) upon titration, and fitting the data at  $\lambda = 365$  nm into a sigmoid equation (b). These graphs correspond to an assay with **AD2** and **Ap180-SLB**

RNA	EC <sub>50</sub> values obtained (μM)			Average	Standard Deviation
<b>Ap180-SL</b>	3.63	3.54	4.90	4.03	0.76
<b>Ap180-SLB</b>	3.24	2.88	4.07	3.40	0.61
<b>Ap167-SLB</b>	1.86	1.41	1.35	1.54	0.28

**Table 6.5.1:** EC<sub>50</sub> values obtained from fluorescence assays

The fluorescence experiments results with **Ap180-SL** showed an average EC<sub>50</sub> value of 4.03 μM, which is reasonably good. The observed change in fluorescence at room temperature further confirms the fact that even though the oligonucleotide melting point and the circular dichroism spectra remain unchanged, an interaction does indeed take place.

The fluorescence experiments with the longer RNA labeled at the same position (**Ap180-SLB**) yielded an average EC<sub>50</sub> value of 3.40 μM. This EC<sub>50</sub> value is lower than the one from the previous experiment (with **Ap180-SL**), but considering the standard deviations, we inferred that the difference between the two EC<sub>50</sub> values is not significant.

The EC<sub>50</sub> obtained with **Ap167-SLB** was 1.54 μM. This result means that **AD2** binds preferably to the bulge than to the GUAA tetraloop, so during the titration **AD2** first binds to the bulge, and then to the loop.

Altogether, these results confirm that **AD2** is a rather promiscuous ligand. It binds to the GUAA loop, but it also binds, and even with better affinity, to the bulge. This fact was also discussed by Baranger, who described, after this work was undertaken, that under certain conditions **AD2** binds strongly also to single- and double-stranded RNAs.<sup>39</sup>

In relation with the ITC results, it could be hypothesized that the interaction between **AD2** and the bulge has a higher Δ*H* than with the GUAA loop, which might explain why the assay with the **SLB-RNA** at 40°C afforded the best thermogram.

### 6.6. Stoichiometry determination by mass spectrometry

Fluorescence spectroscopy assays suggested several binding sites for each oligonucleotide sequence. Hence, it seemed interesting to determine the stoichiometry of the complexes, in other words, how many **AD2** molecules can bind the target RNA.

A common strategy for stoichiometry determination is to carry out a fluorescence-monitored Job Plot,<sup>40</sup> but a proper assay demanded too much fluorescently-labeled RNA. Besides, if an **AD2** molecule binds far from the fluorescent nucleotide, it may not produce a variation in fluorescence intensity and thus yield misleading results.

ITC experiments can also give an idea about the binding stoichiometry, by adjusting the data into different binding equations, but in our case the data obtained were not good enough for stoichiometry determination.

Consequently, we decided to attempt the stoichiometry determination by electrospray ionization mass spectrometry. This experiment consists in analyzing the RNA – **AD2** complex at



## 6. Results and Discussion

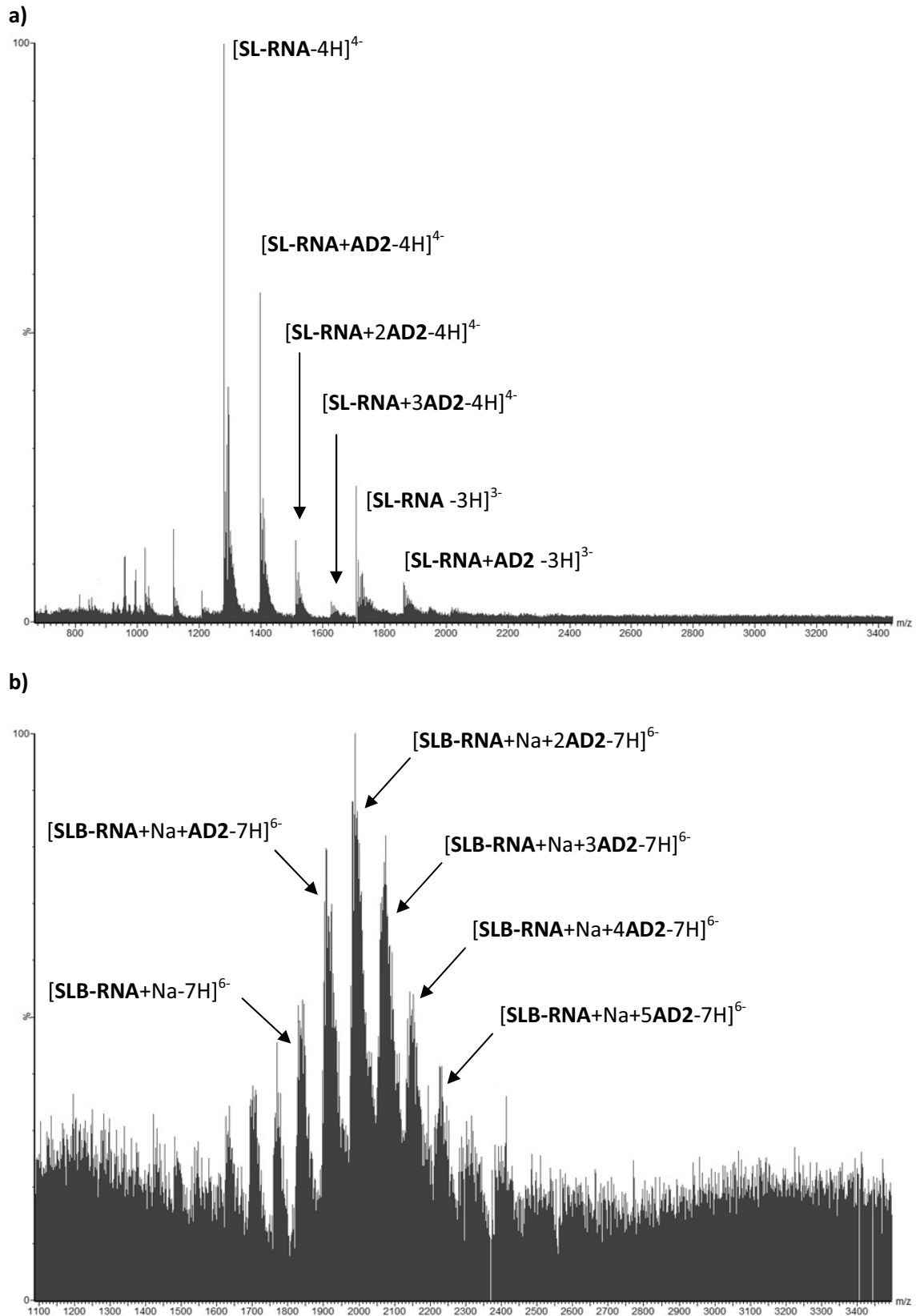
the minimum voltage, in order to ionize the sample without disrupting the complex, and detecting the mass of the target plus the ligands bound to it.<sup>41-44</sup>

We first assayed the **SL-RNA** with a huge excess of **AD2**, up to 20 equivalents, in order to generate the maximum amount of **AD2**-saturated target and detect it more easily. Unfortunately, the excess of **AD2** suppressed the ionization of the complexes, thus yielding only **AD2** peaks. We next repeated the analysis with 4 equivalents of **AD2**, and observed a major peak that corresponded to **SL-RNA** with 4 negative charges, as well as peaks corresponding to the complex with one, two and even three **AD2** molecules. The [**SL-RNA** + 3 **AD2**] peak was very small, near the detection limit of the technique. This means that **SL-RNA** binds at least three **AD2** molecules, but a higher stoichiometry can't be completely discarded. We also observed the same pattern of peaks with complexes with 3 and with 5 negative charges, but they were not as intense as those with 4 charges (**Fig. 6.6.1, a**).

We also performed this experiment with the longer RNA sequence, **SLB-RNA**. The peak of this RNA when analyzed alone did not match with the molecular weight calculated. It matched with an RNA incorporating a sodium atom. We think that this RNA sequence incorporated that metallic cation within its structure, likely into the bulge. Given the size of this RNA, in this case the sensitivity and the resolution of the method were not as good as with **SL-RNA**. Nevertheless, when we performed the experiment with 8 equivalents of **AD2**, we could observe the peak corresponding to **SLB-RNA** without ligands (with a global charge of -6), and the complexes with one to five **AD2** molecules (**Fig. 6.6.1, b**).

According to the data obtained, we can assume that **AD2** binds to **SLB-RNA** with a ratio of at least 5:1. This result is in agreement with the previous ESI experiment with **SL-RNA**, and with the results in fluorescence titrations, which suggested that **AD2** binds to both the loop and the bulge. Given the fact that Baranger research group described an stoichiometry of 2 **AD2** for each GNRA loop,<sup>22</sup> it can be hypothesized that two ligands bind to the loop and one or two more ligands likely bind to the bulge. The other ligands may interact with the double-stranded region of the RNA molecule, probably by intercalation, far from the fluorescent nucleotide and thus almost "invisible" in fluorescence assays.

Altogether, the results of the biophysical assays further confirmed what was already observed in the biological assays. **AD2** may be a good ligand for GNRA loops, but it also binds unspecifically to other loops and RNA regions. Further optimization of **AD2** would be required in order to improve its specificity.



**Figure 6.6.1:** Stoichiometry determination of the RNA-AD2 complexes by ESI. a) SL-RNA with 4 equivalents of AD2, and b) SLB-RNA with 8 equivalents of AD2



## 7. Concluding Remarks

1. Assessment of the **AD2** – target RNA interaction required synthesis of both **AD2** and the RNA sequence reproducing a stem-loop of the apical region of domain 3. **AD2** was synthesized in four steps, including an acetylation, a Mannich reaction, a deacetylation and an aromatic nucleophilic substitution, with an overall yield of 34 %. **AD2** has a pKa value of 6.1.

2. *In vitro* assays suggested that **AD2** exhibits some preference for the FMDV IRES RNA over other RNA regions, because a slight differential inhibition in cap- and IRES-mediated translation was observed. Nevertheless, the SHAPE assay assessed that **AD2** binds to multiple sites of the FMDV IRES RNA.

3. The assays with unlabeled RNA yielded scarcely valuable information. The thermal denaturation curves lost their sigmoid shape upon addition of up to 2 equivalents of **AD2**, and in the circular dichroism spectra no variation was observed. ITC assays with both the stem-loop RNA and the stem-loop-bulge RNA at 25°C did not suggest any remarkable interaction either. At 40°C some interaction could be observed, ultimately suggesting that the binding of **AD2** to the RNA is entropically-driven. The heats detected with the longer RNA sequence (**SLB-RNA**), were higher than with the shorter (**SL-RNA**), so we hypothesize that the binding of **AD2** with the bulge has a higher enthalpy than with the loop. Nevertheless, the thermograms were not good enough to obtain binding parameters.

4. Fluorescence afforded better results. The data obtained by employing three differently labeled RNA sequences suggested that **AD2** has a higher affinity for the bulge than for the GUAA tetraloop.

5. Electrospray ionization mass spectrometry assays suggested an **AD2**-RNA binding ratio of at least 3:1 for **SL-RNA** and 5:1 for **SLB-RNA**. Given the fact that the stoichiometry described for **AD2**-loop interactions is 2:1, these data suggest other interactions of **AD2** with the studied RNA sequences, probably by intercalation into the double-stranded region in addition to interaction with the bulge.

## 7. Concluding Remarks

6. In the light of the results obtained with all the *in vitro* and biophysical assays, it can be inferred that **AD2** is an active molecule that binds to GNRA tetraloops, but also to a lot of other RNA secondary structures and with even higher affinities. This promiscuity is the most important problem to be solved should this compound be optimized for drug development.

## 8. Experimental section

### ***N*-(4-Hydroxyphenyl)acetamide (Prec-1)**

*p*-Aminophenol (1 g, 9.2 mmol) was dissolved in EtOH (100 mL), and acetic anhydride (1 mL, 10.1 mmol) was added dropwise. The solution was heated at 50°C for 3 hours. The solvent was then removed under reduced pressure, and water (100 mL) was added. The acetic acid generated was neutralized with sodium bicarbonate, up to pH = 8. The desired compound was extracted with AcOEt (100 mL), the organic fractions combined, dried over MgSO<sub>4</sub>, filtered and the solvent removed under reduced pressure. **Prec-1** was obtained as a white solid (1.25 g, 89 %). R<sub>f</sub> = 0.4 (Hexanes/AcOEt 1:1); <sup>1</sup>H NMR (DMSO-d<sub>6</sub>, 400 MHz): δ 9.63 (s, 1H), 9.11 (s, 1H), 7.33 (d, *J* = 8.88 Hz, 2H), 6.66 (d, *J* = 8.88 Hz, 2H), 1.97 (s, 3H) ppm; <sup>13</sup>C NMR (DMSO-d<sub>6</sub>, 100 MHz): δ 167.5, 153.1, 131.0, 120.8, 115.0, 23.7 ppm; HRMS (ESI, positive mode): *m/z* 110.0601 [M-Ac+H]<sup>+</sup>, 152.0703[M+H]<sup>+</sup>, 174.0503 [M+Na]<sup>+</sup>; M calcd. for C<sub>8</sub>H<sub>9</sub>NO<sub>2</sub> 151.0633.

### ***N*-(4-Hydroxy-3-[(4-methylpiperazin-1-yl)methyl]phenyl)acetamide (Prec-2)**

*N*-(4-Hydroxyphenyl)acetamide (**Prec-1**) (2 g, 13.2 mmol), 1-methylpiperazine (1.5 mL, 13.2 mmol) and formaldehyde (1.6 mL, 21.2 mmol) were dissolved in absolute ethanol (2 mL). The mixture was heated in a MW oven for 3 minutes at 100W and 7 bar. The solvent was removed under reduced pressure, and the crude was purified by silica gel column chromatography eluting with AcOEt/NEt<sub>3</sub> 10:0.2 and increasing amounts of MeOH (until AcOEt/MeOH/NEt<sub>3</sub> 10:1:0.2), to yield the desired product, **Prec-2** (2.06 g, 59 %). R<sub>f</sub> = 0.2 (AcOEt/MeOH/NEt<sub>3</sub> 10:1:0.2); <sup>1</sup>H NMR (DMSO-d<sub>6</sub>, 400 MHz): δ 10.19 (s, 1H), 9.63 (s, 1H), 7.28 (m, 2H), 6.64 (d, *J* = 4.00 Hz, 1H), 3.55 (s, 2H), 2.3 (bs, 8H), 2.16 (s, 3H), 1.97 (s, 3H) ppm; <sup>13</sup>C NMR (DMSO-d<sub>6</sub>, 100

## 8. Experimental Section

MHz):  $\delta$  167.4, 152.3, 130.9, 122.0, 120.4, 119.3, 115.0, 58.6, 54.5, 52.0, 45.5, 23.7 ppm; HRMS (ESI, positive mode):  $m/z$  164.0689 [M-(1-methylpiperazine)+H]<sup>+</sup>, 264.1677 [M+H]<sup>+</sup>, M calcd. for C<sub>14</sub>H<sub>21</sub>N<sub>3</sub>O<sub>2</sub> 263.1634.

### 4-Amino-2-[(4-methylpiperazin-1-yl)methyl]phenol (Prec-3)

*N*-{4-Hydroxy-3-[(4-methylpiperazin-1-yl)methyl]phenyl}acetamide (**Prec-2**) (200 mg, 0.76 mmol) was dissolved in 5 M aqueous HCl (20 mL, 0.1 mol) and heated to reflux for 4 hours. The solution was then basified with NaOH to pH = 11, and extracted with AcOEt several times, until the organic fraction was colorless. The combined organic fractions were dried over MgSO<sub>4</sub>, filtered, and the solvent was removed under reduced pressure. **Prec-3** was obtained as a brown solid (137 mg, 80 %). R<sub>f</sub> = 0.2 (AcOEt/MeOH/NEt<sub>3</sub> 80:20:20); <sup>1</sup>H NMR (DMSO-d<sub>6</sub>, 400 MHz):  $\delta$  6.41 (m, 1H), 6.34 (m, 2H), 3.46 (s, 2H), 2.3 (bs, 8H), 2.16 (s, 3H) ppm; <sup>13</sup>C NMR (DMSO-d<sub>6</sub>, 100 MHz):  $\delta$  147.6, 140.5, 122.1, 115.5, 115.2, 114.0, 59.3, 54.6, 52.0, 45.5 ppm; ESI MS (positive mode):  $m/z$  101.4 [M-(1-methylpiperazine)+H]<sup>+</sup>, 222.6 [M+H]<sup>+</sup>, calcd. for C<sub>12</sub>H<sub>19</sub>N<sub>3</sub>O 221.15.

### 4-[(6-Chloro-2-methoxyacridin-9-yl)amino]-2-[(4-methylpiperazin-1-yl)methyl]phenol (AD2)

6,9-Dichloro-2-methoxyacridine (0.75 g, 2.71 mmol) and 4-amino-2-[(4-methylpiperazin-1-yl)methyl]phenol (0.60 g, 2.71 mmol) (**Prec-3**) were dissolved in EtOH (180 mL) and heated to reflux for 3.5 hours. The solvent was removed under reduced pressure, and the crude was purified by silica gel column chromatography eluting with AcOEt/NEt<sub>3</sub> 9.5:0.5 and increasing amounts of MeOH (until AcOEt/MeOH/NEt<sub>3</sub> 9:1:0.5), to yield an orange solid (1.02 g, 81 %). R<sub>f</sub> = 0.4 (AcOEt/MeOH/NEt<sub>3</sub> 9:1:0.5); <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz):  $\delta$  8.08 (s, 1H), 7.99 (d, *J* = 7.6 Hz, 1H), 7.87 (d, *J* = 9.2 Hz, 1H), 7.38 (dd, *J*<sub>1</sub> = 9.4 Hz, *J*<sub>2</sub> = 2.4 Hz, 1H), 7.23 (d, *J* = 9.1 Hz, 1H), 7.06 (s, 1H), 6.82 (dd, *J*<sub>1</sub> = 8.6 Hz, *J*<sub>2</sub> = 2.6 Hz, 1H), 6.76 (d, *J* = 8.6 Hz, 1H), 6.54 (d, *J* = 2.6 Hz, 1H), 3.72 (s, 3H), 3.58 (s, 2H), 2.3 (bs, 8H), 2.30 (s, 3H) ppm; <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz):  $\delta$  156.2, 153.4, 135.0, 125.2, 124.9, 122.1, 120.1, 119.9, 117.8, 116.9, 100.2, 61.3, 55.4, 55.0, 52.5, 46.0 ppm; HRMS (ESI, positive mode):  $m/z$  232.1009 [M+2H]<sup>2+</sup>, 363.0924 [M-(1-methylpiperazine)+H]<sup>+</sup>, 463.1929 [M+H]<sup>+</sup>; M calcd. for C<sub>26</sub>H<sub>27</sub>ClN<sub>4</sub>O<sub>2</sub>: 462.1823. Analytical HPLC (A: H<sub>2</sub>O (0.045 % TFA), B: ACN (0.036 % TFA), Gracesmart C18 column, 5  $\mu$ m, 250 x 4.6 mm, 1 mL/min, linear gradient from 5 to 30 % of B in 30 min): t<sub>R</sub> = 12.5 min (**Fig. 8.1**).

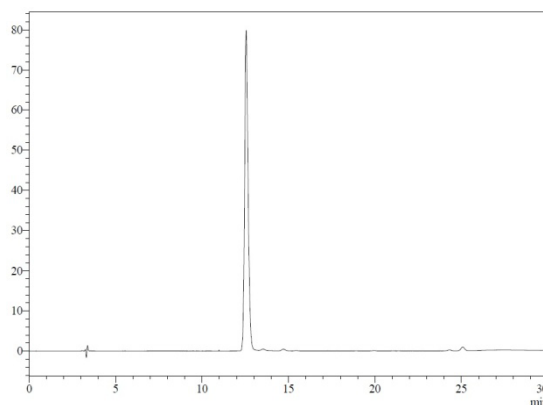


Figure 8.1: Pure AD2

### $5'$ U(OMe)CC-UUG-GUA-ACA-AGG-A(OMe) $3'$ SL-RNA

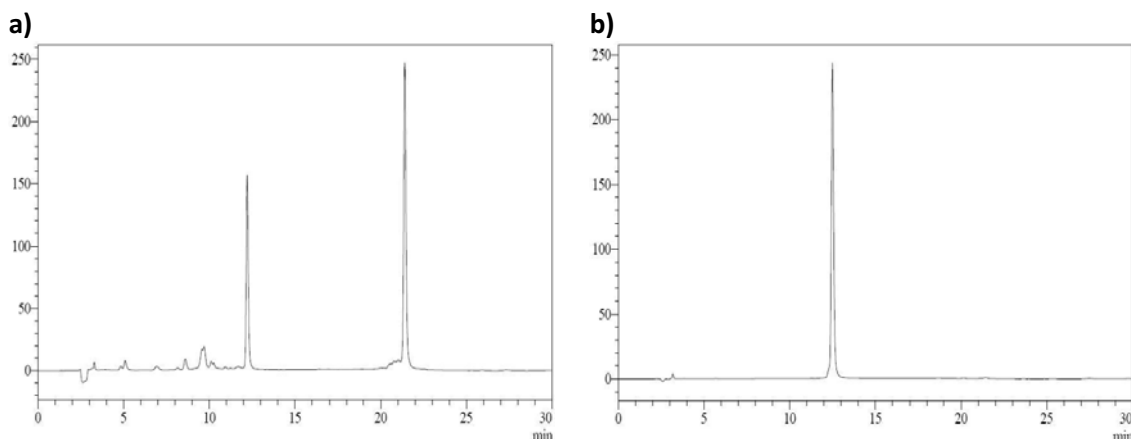
The synthesis was carried out at the 1  $\mu$ mol scale on a functionalized CPG solid support, using standard phosphite triester approach synthesis cycles. The extension of the coupling reaction was controlled by DMT quantification.

The phosphoramidites employed were derivatives of A<sup>Pac</sup>, C<sup>Ac</sup>, G<sup>iPrPac</sup> and U, protected with the DMT group at the 5' position and with the TBDMS group at the 2' position, and the 2'OMe analogs for the terminal nucleotides. After chain elongation, treatment with a 1:1 mixture of conc. aq. ammonia and a 8 M solution of methylamine in ethanol (2 mL/ $\mu$ mol) at 65°C for 30 minutes cleaved the oligonucleotide and removed all protecting groups except the 2'-TBDMS groups and 5'-DMT. After filtration and washing with ethanol and water, the solvent was removed under reduced pressure and the crude was treated with a 6:3:4 mixture of NMP, TEA and TEA-3HF at 65°C for 2.5 hours (600  $\mu$ L/ $\mu$ mol). The RNA precipitated upon addition of 2 volumes of trimethylsilyl isopropyl ether. Diethyl ether (5 mL/ $\mu$ mol) was added, the suspension was centrifuged at 2200 rpm at 5°C for 5 minutes and the supernatant was discarded. This process was repeated two more times.

The RNA crude (DMT-On) was purified by reversed-phase HPLC at the semipreparative scale (A: H<sub>2</sub>O/TEAA 0.1M; B: ACN, Phenomenex Jupiter C18 column, 10  $\mu$ m, 250 x 10 mm, 3 mL/min, linear gradient from 10 to 50 % of B in 30 min,  $t_R$  = 22.4 min) (**Fig. 8.2, a**). The pure oligonucleotide was redissolved in a 3 % acetic acid solution in water, and repurified after an hour of reaction (A: H<sub>2</sub>O/TEAA 0.1M; B: ACN, Phenomenex Jupiter C18 column, 10  $\mu$ m, 250 x 10 mm, 3 mL/min, linear gradient from 10 to 50 % of B in 30 min,  $t_R$  = 10.9 min). The final compound was quantified by UV spectroscopy and characterized by mass spectrometry. RNA extinction coefficients were calculated with an online calculator, which takes into account the different bases, their neighbors and the structure (single-stranded or double-stranded).<sup>45</sup> Overall yield (synthesis and purification): 21 %. Analytical HPLC (A: H<sub>2</sub>O/TEAA 0.05M; B: ACN, Kromasil C18 column, 10  $\mu$ m, 250 x 4 mm, 1 mL/min, linear gradient from 10 to 50 % of B in 30 min):  $t_R$  = 12.4 min (**Fig. 8.2, b**); MALDI-TOF MS (THAP/CA, negative mode):  $m/z$ : 5132.3 [M-H]<sup>-</sup>, 2566.5 [M-2H]<sup>2-</sup>, M calcd. for C<sub>155</sub>H<sub>193</sub>N<sub>62</sub>O<sub>109</sub>P<sub>15</sub> 5130.8.



## 8. Experimental Section



**Figure 8.2:** a) DMT-On crude of **SL-RNA**. b) DMT-Off pure **SL-RNA**. The peak at  $t_R = 12.3$  min in chromatogram a) corresponds to DMT-Off **SL-RNA**. It was collected and combined with the other fraction after DMT removal

### AD2 pKa determination<sup>32</sup>

First the absorbance of a solution of **AD2** (16.2  $\mu\text{M}$ ) was measured at pH 3.9 and 7.8, obtaining  $d_i$  and  $d_m$  respectively. Then a series of solutions of different pH were prepared, each one with **AD2** (16.2  $\mu\text{M}$ ), phosphate buffer (20 mM) and acetic acid/acetate buffer (20 mM). pHs were adjusted with NaOH or HCl. The different pH and absorbance data obtained at  $\lambda = 280$  nm are listed below.

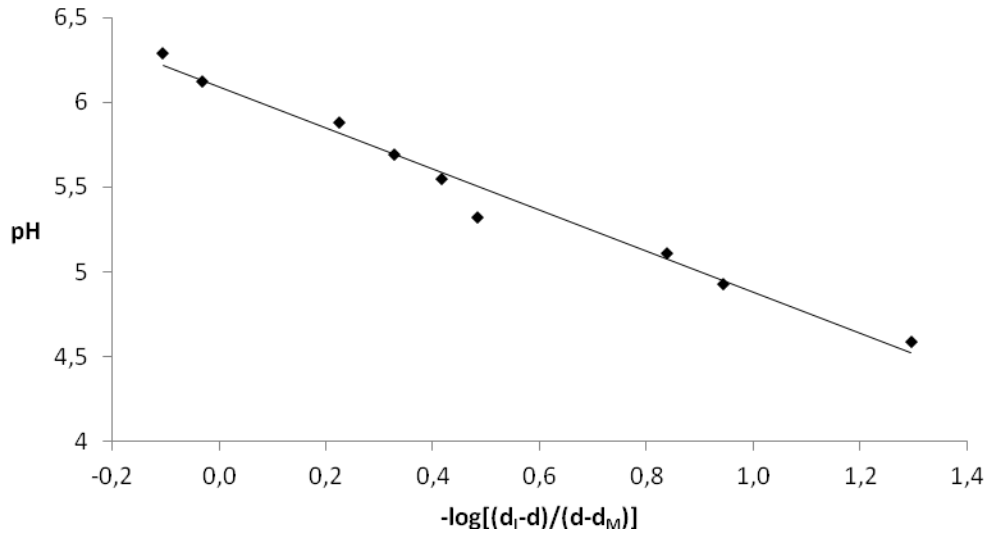
pH	Absorbance	pH	Absorbance
4.59	0.327	5.69	0.282
4.93	0.318	5.88	0.273
5.11	0.314	6.12	0.249
5.32	0.294	6.29	0.242
5.55	0.289		

The data were fitted to this equation:

$$pH = pKa - \log \frac{d_i - d}{d - d_m}$$

Where  $d_i$  is the absorption of the ionized compound,  $d_m$  is the absorption of the deprotonated compound and  $d$  is the absorption at different pH. Interception with the Y-axis corresponds with pKa.

For **AD2**,  $pKa = 6.1$ , and  $R^2$  (coefficient of determination) = 0.98



### Stoichiometry determination by mass spectrometry

**SL-RNA** and **SLB-RNA** were mixed with 4 and 8 equivalents of **AD2**, respectively, and were dissolved in a 50 mM ammonium acetate solution to a final RNA concentration of 25  $\mu\text{M}$ . The samples were annealed by heating at 90 °C and cooling down slowly. 75  $\mu\text{L}$  of the sample solution were directly injected.

The samples were negatively ionized by electrospray, mode IM-MS (V mode, Mobility-TOF mode). Other settings of the instrument are detailed below:

- Sampling cone: 45 v
- Source temperature: 100 °C
- Trap Collision Energy: 10 V
- Transfer Collision Energy: 10 V
- IM-Gas Flow: 24 ml/min
- Trap Gas Flow: 5 ml/min
- Bias: 15 V
- IM Wave height: 8.0 V
- IM Wave velocity: 300 m/s
- Pressures: Backing 5.75e0 mbar

Trap 4.00e-2 mbar

IMS 4.67e-1 mbar

TOF 2.17 e-6 mbar

## 8. Experimental Section

- $m/z$  range: 500 to 8000
- Scan time 1s
- RF (Offset/Limit): Source 350/450, IMS 350/380, Trap 350/380, Transfer 350/380
- External calibration with Csl,  $m/z$  range: 500 to 8000

Data were acquired and processed with the MassLynx software v 4.1 (SCN 704). MS spectra were deconvoluted to the average masses with integrated algorithms in MassLynx. IM-MS data were processed with Driftscope software vs. 2.4.

**SL-RNA** (ESI, negative mode)  $m/z$ : 1282.02  $[M-4H]^{4-}$ , 1397.85  $[M+AD2-4H]^{4-}$ , 1513.64  $[M+2AD2-4H]^{4-}$ , 1629.58  $[M+3AD2-4H]^{4-}$ , M calcd. for  $C_{155}H_{193}N_{62}O_{109}P_{15}$  5133.2,  $M+AD2$  calcd. for  $C_{181}H_{220}ClN_{66}O_{111}P_{15}$  5596.2,  $M+2AD2$  calcd. for  $C_{207}H_{247}Cl_2N_{70}O_{113}P_{15}$  6059.1,  $M+3AD2$  calcd. for  $C_{233}H_{274}Cl_3N_{74}O_{115}P_{15}$  6522.1.

**SLB-RNA** (ESI, negative mode)  $m/z$ : 1829.3  $[M+Na-7H]^{6-}$ , 1903.2  $[M+Na+AD2-7H]^{6-}$ , 1989.3  $[M+Na+2AD2-7H]^{6-}$ , 2063.4  $[M+Na+3AD2-7H]^{6-}$ , 2145.9  $[M+Na+4AD2-7H]^{6-}$ , 2214.4  $[M+Na+5AD2-7H]^{6-}$ , M calcd. for  $C_{323}H_{403}N_{130}O_{234}P_{33}$  10872.6,  $M+AD2$  calcd. for  $C_{349}H_{430}ClN_{134}O_{236}P_{33}$  11335.6,  $M+2AD2$  calcd. for  $C_{375}H_{457}Cl_2N_{138}O_{238}P_{33}$  11798.5,  $M+3AD2$  calcd. for  $C_{401}H_{484}Cl_3N_{142}O_{240}P_{33}$  12261.5,  $M+4AD2$  calcd. for  $C_{427}H_{511}Cl_4N_{146}O_{242}P_{33}$  12724.5,  $M+5AD2$  calcd. for  $C_{453}H_{538}Cl_5N_{150}O_{244}P_{33}$  13187.5.

## 9. References and Notes

- (1) Eddy, S. R. *Nat. Rev. Genet.* **2001**, *2*, 919.
- (2) Credit to Yikrazuul, [www.wikipedia.org](http://www.wikipedia.org), Data extracted from PDB 1ehz (<http://www.rcsb.org/pdb/explore/explore.do?structureId=1ehz>) rendered with PyMOL.
- (3) Thomas, J. R.; Hergenrother, P. J. *Chem. Rev.* **2008**, *108*, 1171.
- (4) Kumar, G. S. *J. Biosci.* **2012**, *37*, 539.
- (5) Grubman, M. J.; Baxt, B. *Clin. Microbiol. Rev.* **2004**, *17*, 465.
- (6) Domingo, E.; Baranowski, E.; Escarmís, C.; Sobrino, F. *Comp. Immunol. Microbiol. Infect. Dis.* **2002**, *25*, 297.
- (7) Cubillos, C.; de la Torre, B. G.; Jakab, A.; Clementi, G.; Borrás, E.; Barcena, J.; Andreu, D.; Sobrino, F.; Blanco, E. *J. Virol.* **2008**, *82*, 7223.
- (8) Andreu, D.; Barcena, J.; Blanco, E.; Cubillos, C.; García, B.; Monsó, M.; Sobrino, F. Peptide vaccines for the prevention of foot-and-mouth disease. EP2647390A1, **2013**.
- (9) Wei, H.; Fang, M.; Wan, M.; Wang, H.; Zhang, P.; Hu, X.; Wu, X.; Yang, M.; Zhang, Y.; Zhou, L.; Jiao, C.; Hua, L.; Diao, W.; Xiao, Y.; Yu, Y.; Wang, L. *Biotechnol. Lett.* **2014**, *36*, 723.
- (10) Carlson, D. F.; Fahrenkrug, S. C. Production of fmdv-resistant livestock by allele substitution. US 20140041066A1, **2013**.
- (11) Roqué Rosell, N. R.; Mokhlesi, L.; Milton, N. E.; Sweeney, T. R.; Zunszain, P. A.; Curry, S.; Leatherbarrow, R. J. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 490.
- (12) Vagnozzi, A.; Stein, D. A.; Iversen, P. L.; Rieder, E. *J. Virol.* **2007**, *81*, 11669.
- (13) Martínez-Salas, E.; Piñeiro, D.; Fernández, N. *Comp. Funct. Genomics* **2012**, *2012*, 1.
- (14) Jung, S.; Schlick, T. *Nucleic Acids Res.* **2013**, *41*, 1483.
- (15) Fiore, J. L.; Nesbitt, D. J. *Q. Rev. Biophys.* **2013**, *46*, 223.
- (16) Geary, C.; Baudrey, S.; Jaeger, L. *Nucleic Acids Res.* **2007**, *36*, 1138.
- (17) Fujita, Y.; Tanaka, T.; Furuta, H.; Ikawa, Y. *J. Biosci. Bioeng.* **2012**, *113*, 141.
- (18) Thapar, R.; Denmon, A. P.; Nikonowicz, E. P. *Wiley Interdiscip. Rev. RNA* **2014**, *5*, 49.
- (19) Horiya, S.; Koh, C.-S.; Matsufuji, S.; Harada, K. *Nucleic Acids Symp. Ser.* **2008**, *52*, 209.
- (20) Furusawa, H.; Fukusho, S.; Okahata, Y. *ChemBioChem* **2014**, *15*, 865.

## 9. References and Notes

- (21) Thomas, J. R.; Liu, X.; Hergenrother, P. J. *Biochemistry (Mosc.)* **2006**, *45*, 10928.
- (22) Yan, Z.; Sikri, S.; Beveridge, D. L.; Baranger, A. M. *J. Med. Chem.* **2007**, *50*, 4096.
- (23) Huang, M.; Li, Z.; Li, D.; Walker, S.; Greenan, C.; Kennedy, R. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 3044.
- (24) Dorjsuren, D.; Burnette, A.; Gray, G. N.; Chen, X.; Zhu, W.; Roberts, P. E.; Currens, M. J.; Shoemaker, R. H.; Ricciardi, R. P.; Sei, S. *Antiviral Res.* **2006**, *69*, 9.
- (25) Jordheim, L. P.; Barakat, K. H.; Heinrich-Balard, L.; Matera, E.-L.; Cros-Perrial, E.; Bouledrak, K.; El Sabeh, R.; Perez-Pineiro, R.; Wishart, D. S.; Cohen, R.; Tuszyński, J.; Dumontet, C. *Mol. Pharmacol.* **2013**, *84*, 12.
- (26) Odolczyk, N.; Fritsch, J.; Norez, C.; Servel, N.; da Cunha, M. F.; Bitam, S.; Kupniewska, A.; Wiszniewski, L.; Colas, J.; Tarnowski, K.; Tondelier, D.; Roldan, A.; Sausseureau, E. L.; Melin-Heschel, P.; Wiczorek, G.; Lukacs, G. L.; Dadlez, M.; Faure, G.; Herrmann, H.; Ollero, M.; Becq, F.; Zielenkiewicz, P.; Edelman, A. *EMBO Mol. Med.* **2013**, *5*, 1484.
- (27) Vehar, B.; Hrast, M.; Kovač, A.; Konc, J.; Mariner, K.; Chopra, I.; O'Neill, A.; Janežič, D.; Gobec, S. *Bioorg. Med. Chem.* **2011**, *19*, 5137.
- (28) Bond, A.; Reichert, Z.; Stivers, J. T. *Mol. Pharmacol.* **2005**, *69*, 547.
- (29) Mahesh, R.; Venkatesha Perumal, R. *Indian J. Chem.* **2004**, *43B*, 1012.
- (30) Ferlin, M. G.; Marzano, C.; Chiarelto, G.; Baccichetti, F.; Bordin, F. *Eur. J. Med. Chem.* **2000**, *35*, 827.
- (31) Lee, H. H.; Palmer, B. D.; Denny, W. A. *J. Org. Chem.* **1988**, *53*, 6042.
- (32) Albert, A.; Serjeant, E. P. In *Ionization constants of acids and bases: a laboratory manual*; John Wiley & Sons, Inc., **1962**; Vol. 4, pp. 69–92.
- (33) Promega. CAT enzyme assay system. Technical Bulletin.
- (34) Promega. Luciferase assay system. Technical Bulletin.
- (35) Steen, K.-A.; Malhotra, A.; Weeks, K. M. *J. Am. Chem. Soc.* **2010**, *132*, 9940.
- (36) Nordén, B.; Rodger, A.; Dafforn, T. *Linear dichroism and circular dichroism - a textbook on polarized light spectroscopy*; The Royal Society of Chemistry: Cambridge, **2010**.
- (37) Hoener, B.-A.; Sokoloski, T. D.; Mitscher, L. A. *Antimicrob. Agents Chemother.* **1973**, *4*, 455.
- (38) Jean, J. M.; Hall, K. B. *Proc. Natl. Acad. Sci.* **2001**, *98*, 37.
- (39) Warui, D. M.; Baranger, A. M. *J. Med. Chem.* **2009**, *52*, 5462.
- (40) Huang, C. Y. *Methods Enzymol.* **1982**, *87*, 509.
- (41) Sannes-Lowery, K. A.; Griffey, R. H.; Hofstadler, S. A. *Anal. Biochem.* **2000**, *280*, 264.
- (42) Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E. *Nucleic Acids Res.* **2002**, *30*, e82.
- (43) Brodbelt, J. S. *Annu. Rev. Anal. Chem.* **2010**, *3*, 67.
- (44) Hilton, G. R.; Benesch, J. L. P. *J. R. Soc. Interface* **2012**, *9*, 801.
- (45) IDT oligonucleotide extinction coefficient calculator  
<http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/default.aspx>.

# **APPENDICES**



## Appendix 1: Equivalence table

Compound number in this thesis	Compound number in the papers*
1	1 (OL)
2	1 (BC); 2 (OL)
3	2 (BC)
4	3 (BC); 11 (OL)
5	6 (BC)
6	-
7	8 (BC)
8	9 (BC)
9	-
10	-
11	10 (BC)
12	11 (BC)
13	12 (BC)
14	13 (BC)
15	18 (BC)
16	19 (BC)
17	23 (BC)
18	20 (BC)
19	21 (BC)
20	22 (BC)
21	24 (BC)
22	25 (BC)
23	26 (BC)
24	27 (BC)
25	28 (BC)
26	29 (BC)
27	30 (BC)
28	3 (OL)
29	4 (OL)
30	5 (OL)
31	7 (OL)
32	-
33	8 (OL)
34	-
35	9 (OL)
36	-
37	10 (OL)
38	12 (OL)
39	13 (OL)
40	-
41	14 (OL)
42	15 (OL)
43	16 (OL)
44	-
45	17 (OL)
46	1 (JOC)
47	2 (JOC)



Appendix 1: Equivalence table

48	3 (JOC)
49	4 (JOC)
50	5 (JOC)
51	6 (JOC)
52	7 (JOC)
53	8 (JOC)
54	9 (JOC)
55	10 (JOC)
56	11 (JOC)
57	12 (JOC)
58	13 (JOC)
59	14 (JOC)
60	15 (JOC)
61	16 (JOC)
62	17 (JOC)
63	18 (JOC)
64	19 (JOC)
65	20 (JOC)
66	21 (JOC)
67	22 (JOC)
68	23 (JOC)
69	24 (JOC)
70	-
71	25 (JOC)
72	-
73	-
74	-
75	-

---

\* BC: Bioconjugate Chemistry, OL: Organic Letters, JOC: Journal of Organic Chemistry

## Appendix 2: Experiments and compounds not included in the papers

### Assessment of the stability of protected maleimides

A 75:25 mixture of the *exo* and *endo* cycloadducts of **2** was dissolved in a 4:1 (v/v) THF/piperidine solution and stirred for 1 hour. The solvent was removed under reduced pressure and the crude analyzed by  $^1\text{H}$  NMR in  $\text{CDCl}_3$ . The spectra showed the same set of signals and the same relative areas before and after the treatment.

### Ac-Ser-Lys(protected maleimide)-Tyr-Gly-OH (**9**)

The synthesis was carried out on the Wang Resin (475 mg,  $f = 1.1$  mmol/g) according to the protocol described in Paper 1 (Bioconjugate Chemistry). Coupling of the first amino acid reduced the functionalization to  $f = 0.34$  mmol/g. Peptide **9** was purified (A:  $\text{H}_2\text{O} + 0.1\%$  TFA, B: ACN +  $0.1\%$  TFA, Jupiter Proteo C12 column  $10\ \mu\text{m}$ ,  $250 \times 10$  mm,  $3\ \text{mL/min}$ , linear gradient from 5 to 50 % of B in 30 min) and obtained in 7 % overall yield (synthesis and purification). Analytical HPLC (A:  $\text{H}_2\text{O} + 0.045\%$  TFA, B: ACN +  $0.036\%$  TFA, Gracesmart C18 column,  $5\ \mu\text{m}$ ,  $250 \times 4.6$  mm,  $1\ \text{mL/min}$ , linear gradient from 5 to 50 % of B in 30 min):  $t_{\text{R}} = 12.9$  min (**Fig. A2.1, a**); ESI MS (negative mode):  $m/z$ : 741.1  $[\text{M}-\text{H}]^-$ , 855.1  $[\text{M}+\text{TFA}]^-$ , M calcd. for  $\text{C}_{35}\text{H}_{46}\text{N}_6\text{O}_{12}$  742.3

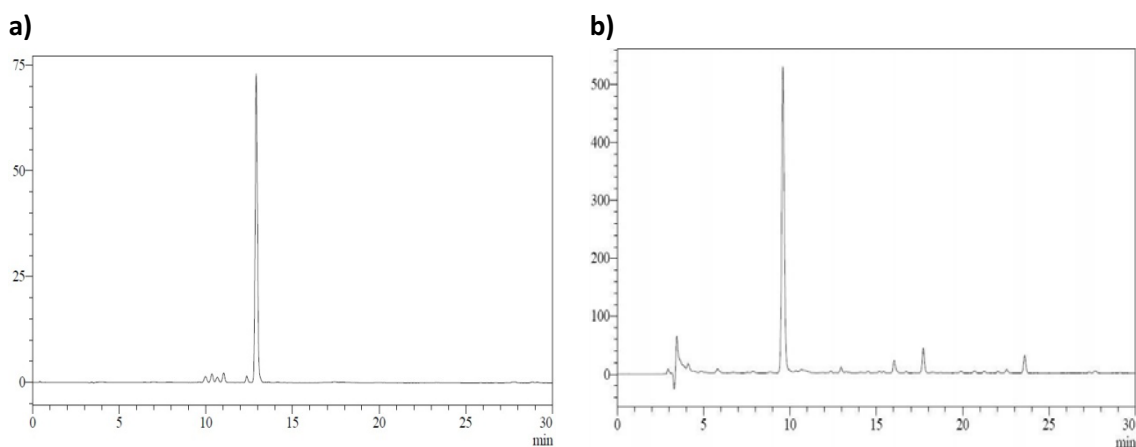
### Ac-Ser-Lys(maleimide)-Tyr-Gly-OH (**10**)

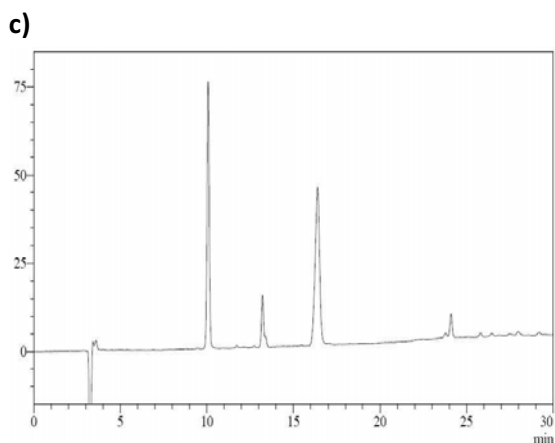
#### Deprotection by heating in the microwave oven

This protocol is described in Paper 1 (Bioconjugate Chemistry). Deprotection of **9**: 25 nmol,  $25\ \mu\text{M}$ ,  $1\ \text{mL}$  1:1 (v/v) methanol/water,  $90^\circ\text{C}$ , 90 min; deprotection yield (from the HPLC trace): 87 %. Analytical HPLC (A:  $\text{H}_2\text{O} + 0.045\%$  TFA, B: ACN +  $0.036\%$  TFA, Gracesmart C18 column,  $5\ \mu\text{m}$ ,  $250 \times 4.6$  mm,  $1\ \text{mL/min}$ , linear gradient from 5 to 50 % of B in 30 min):  $t_{\text{R}} = 9.6$  min (**Fig. A2.1, b**); ESI MS (negative mode):  $m/z$ : 645.2  $[\text{M}-\text{H}]^-$ , M calcd. for  $\text{C}_{29}\text{H}_{38}\text{N}_6\text{O}_{11}$  646.3

#### Deprotection by heating in toluene

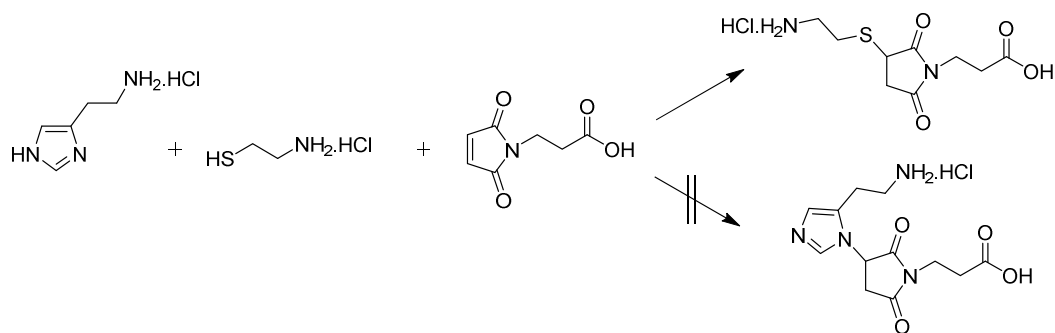
This protocol is described in Paper 1 (Bioconjugate Chemistry). Deprotection of **9**: 100 nmol,  $25\ \mu\text{M}$ ,  $4\ \text{mL}$  toluene, 4 h; deprotection yield (from the HPLC trace): 82 %. Analytical HPLC (A:  $\text{H}_2\text{O} + 0.045\%$  TFA, B: ACN +  $0.036\%$  TFA, Gracesmart C18 column,  $5\ \mu\text{m}$ ,  $250 \times 4.6$  mm,  $1\ \text{mL/min}$ , linear gradient from 5 to 50 % of B in 30 min):  $t_{\text{R}} = 10.1$  min (**Fig. A2.1, c**).





**Figure A2.1:** a) Pure peptide **9**, b) deprotection of **9** using the microwave-promoted deprotection protocol, and c) deprotection of **9** by heating in toluene. The peak eluting at 16.2 min corresponds to toluene

### Histamine and cysteamine competition assay



**Figure A2.2:** Histamine and cysteamine competition assay

Protected maleimidopropanoic acid (95 mg, 0.36 mmol), histamine hydrochloride (75 mg, 0.4 mmol) and cysteamine hydrochloride (45 mg, 0.4 mmol) were dissolved in a 1:1 (v/v) methanol/water mixture (4 mL) and heated in the microwave oven for 90 min at 90°C. The solvent was removed under reduced pressure and the crude was redissolved in a 10 % solution of Na<sub>2</sub>CO<sub>3</sub> in water and extracted with DCM. The aqueous fraction was acidified to pH = 3 with HCl, and reextracted with DCM. The solvent of the aqueous fraction and the two organic fractions was removed under reduced pressure and the crudes were analyzed by <sup>1</sup>H-NMR and ESI.

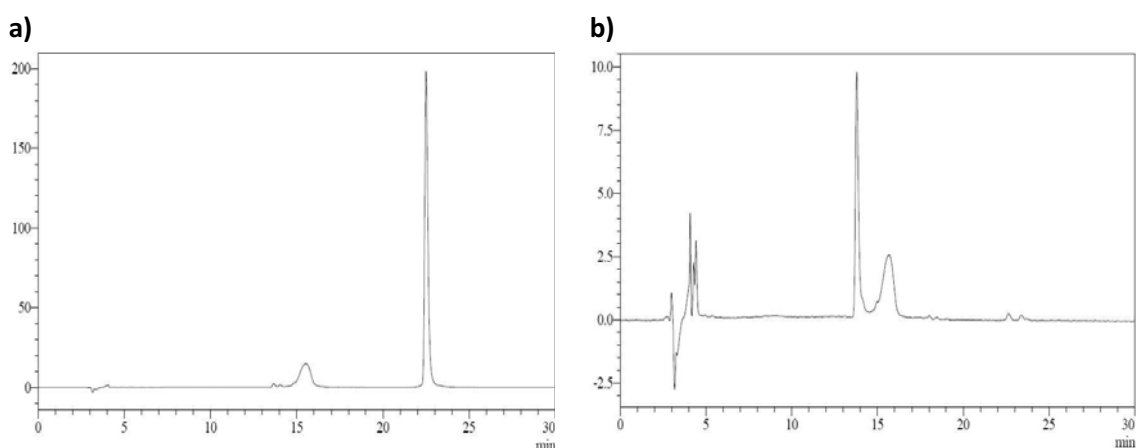
NMR of the organic fraction showed tiny signals of protected maleimidopropanoic acid. In the aqueous fraction a lot of signals were found between 2 and 4 ppm, and a single set of signals in the aromatic region, corresponding to free histamine.

ESI-MS of the aqueous fraction showed free histamine and succinimide-cysteamine, but no succinimide-histidine was detected (**Fig. A2.2**). ESI MS (positive mode): m/z: 111.9 [histamine + H]<sup>+</sup>, 246.5, [cysteamine-maleimidopropanoic acid + H]<sup>+</sup>, histamine calcd. for C<sub>5</sub>H<sub>9</sub>N<sub>3</sub> 111.1, cysteamine-maleimidopropanoic acid calcd. for C<sub>9</sub>H<sub>14</sub>N<sub>2</sub>O<sub>4</sub>S 246.1.

## Attempts to prepare conjugate **70**

### Attempt 1

An equimolar mixture of **69** (25 nmol) and 3,6-di(2-pyridyl)-1,2,4,5-tetrazine (25 nmol) was dissolved in a 5:4:1 (v/v) water/acetonitrile/methanol solution (1 mL) and allowed to react at room temperature for 90 min. The crude was analyzed by HPLC and the chromatogram showed only starting materials (**Fig. A2.3, a**). The solution was then heated at 50°C and stirred overnight. Analytical HPLC only showed tetrazine degradation and no trace of conjugate **70** (**Fig. A2.3, b**). Analytical HPLC chromatograms are shown below (A: H<sub>2</sub>O + 0.045 % TFA, B: ACN + 0.036 % TFA, Phenomenex Jupiter Proteo C12 column, 4 μm, 250 x 4.6 mm, 1 mL/min, linear gradient from 10 to 40 % of B in 30 min).



**Figure A2.3:** Crude of the conjugation of **69** with 3,6-di(2-pyridyl)-1,2,4,5-tetrazine **a)** after 90 min at room temperature, and **b)** overnight at 50°C. The broad peak eluting at  $t_R = 15.4$  min corresponds to cyclic peptoid **69**, the peak eluting at 22.5 min corresponds to the tetrazine employed, and the peak eluting at 13.8 min corresponds to tetrazine degradation

### Attempt 2

An equimolar mixture of **69** (25 nmol) and 3,6-di(2-pyridyl)-1,2,4,5-tetrazine (25 nmol) was dissolved in a 5:4:1 (v/v) water/acetonitrile/methanol solution (1 mL) and heated in the microwave oven at 90°C for 60 min. The crude was analyzed by HPLC and the chromatogram showed only starting cyclic peptoid and unidentified peaks. **70** Was also not detected by mass spectrometry.

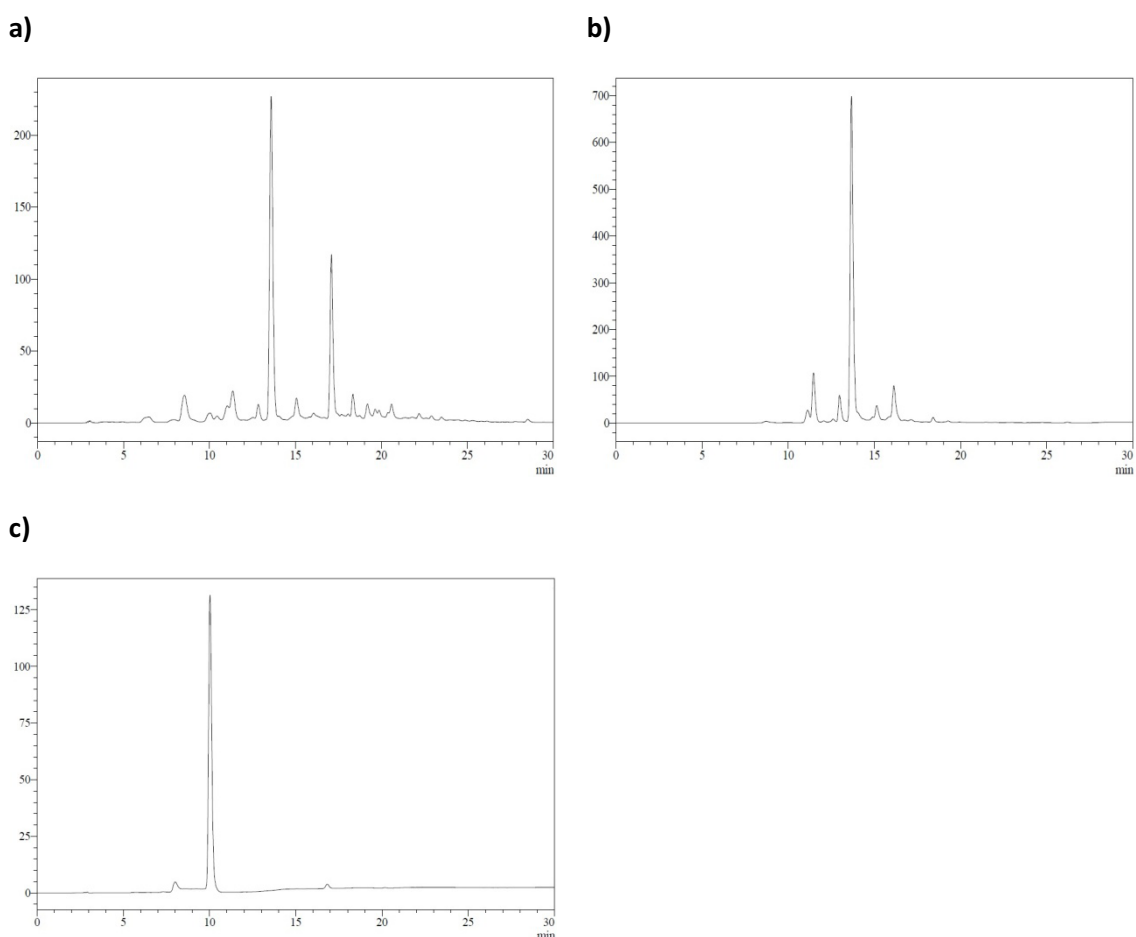
## Synthesis of cyclic peptide-PNA conjugates with potential biological activity

The synthesis of **Pre-72**, **Pre-73**, **Pre-74** and **Pre-75** was carried out on 1 g of Rink Amide MHBA Resin with  $f = 0.65$  mmol/g. At different stages of the synthesis, the resin was split into several batches.

### Cyclic Maleimido\*-Lys(protected maleimide)-Glu-Trp-Tyr-Tyr-Glu-Cys\*-NH<sub>2</sub> (**Pre-72**)

**4** and **1** were incorporated onto 115 μmol of Fmoc-Glu(O<sup>t</sup>Bu)-Trp(Boc)-Tyr(<sup>t</sup>Bu)-Tyr(<sup>t</sup>Bu)-Tyr(<sup>t</sup>Bu)-Glu(O<sup>t</sup>Bu)-Cys(Trt)-resin. The protected peptide-resin was treated with TFA/TIS/H<sub>2</sub>O 90:5:5, (2 hours) and purified (A: H<sub>2</sub>O + 0.1 % TFA, B: ACN + 0.1 % TFA, Phenomenex Jupiter Proteo C12 column, 10 μm, 250 x 10 mm, 3 mL/min, linear gradient from 40 to 60 % of B in 30 min,  $t_R = 9.6$  min). Overall yield (synthesis and purification): 24 %, 73 % purity (**Fig. A2.4, a** and

**b).** An analytical sample of the compound was repurified under the same conditions. Analytical HPLC (A: H<sub>2</sub>O + 0.045 % TFA, B: ACN + 0.036 % TFA, Phenomenex Jupiter Proteo C12 column, 4 μm, 250 x 4.6 mm, 1 mL/min, linear gradient from 30 to 60 % of B in 30 min):  $t_R$  = 13.7 min. The last analytical HPLC was performed under the same conditions at 80°C (**Fig. A2.4, b** and **c**). MALDI-TOF MS (THAP, negative mode):  $m/z$ : 1483.3 [M-furan-H]<sup>-</sup>; M calcd. for C<sub>77</sub>H<sub>89</sub>N<sub>13</sub>O<sub>22</sub>S 1579.6, M-furan calcd. for C<sub>71</sub>H<sub>81</sub>N<sub>13</sub>O<sub>21</sub>S 1483.5. MALDI-TOF MS after reaction with H<sub>2</sub>O<sub>2</sub> (THAP, positive mode):  $m/z$ : 1500.9 [M-furan+16+H]<sup>+</sup>, 1523.1 [M-furan+16+Na]<sup>+</sup>; M+16 calcd. for C<sub>77</sub>H<sub>89</sub>N<sub>13</sub>O<sub>23</sub>S 1595.6, M-furan+16 calcd. for C<sub>71</sub>H<sub>81</sub>N<sub>13</sub>O<sub>22</sub>S 1499.6.



**Figure A2.4:** a) Crude cyclic peptide **Pre-72**, b) purified **Pre-72**, and c) repurified **Pre-72** (80°C)

#### **Cyclic Maleimido\*-Lys(maleimide)-Glu-Trp-Tyr-Tyr-Glu-Cys\*-NH<sub>2</sub> (Dep-Pre-72)**

This protocol is described in Paper 1 (Bioconjugate Chemistry). Deprotection of **Pre-72** (73 % purity): 100 nmol, 100 μM, 1 mL 1:1 (v/v) methanol/water mixture, MW heating, 90°C, 90 min. The product was purified under the same analytical conditions stated below. Overall yield (deprotection and purification): 46 %. Analytical HPLC (A: H<sub>2</sub>O + 0.045 % TFA, B: ACN + 0.036 % TFA, Phenomenex Jupiter Proteo C12 column, 4 μm, 250 x 4.6 mm, 1 mL/min, linear gradient from 30 to 60 % of B in 30 min):  $t_R$  = 8.8 min (**Fig. A2.5**). MALDI-TOF MS (THAP, negative mode):  $m/z$ : 1482.9 [M-H]<sup>-</sup>; M calcd. for C<sub>71</sub>H<sub>81</sub>N<sub>13</sub>O<sub>21</sub>S 1483.5.

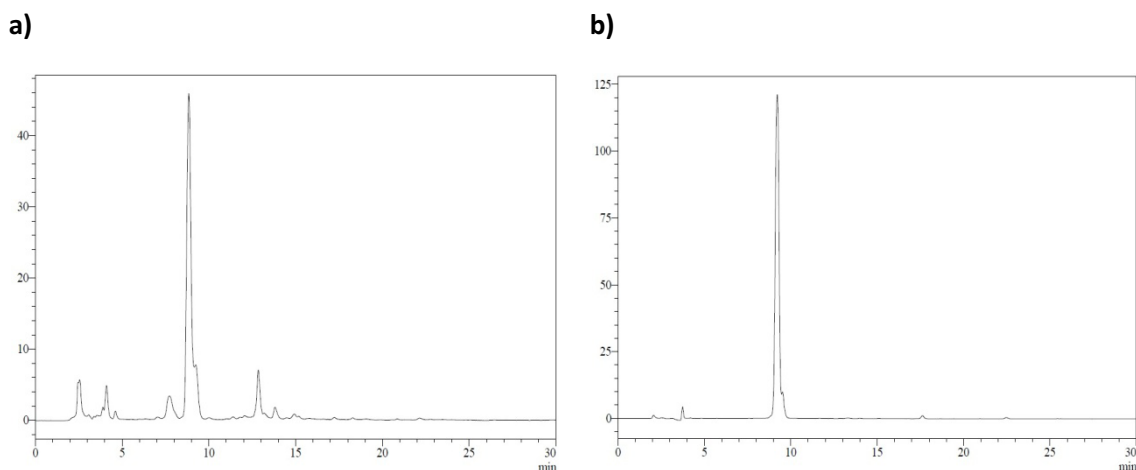


Figure A2.5: a) Crude deprotected cyclic peptide **Dep-Pre-72**, and b) pure **Dep-Pre-72**

### Conjugate **72**

**Dep-Pre-72** (90 nmol) and **Cys-PNA** (prepared by Natalia Masnou) (50 nmol) were dissolved in a 2:1 (v/v) water/ACN mixture to a 150  $\mu$ M final PNA concentration and heated at 50°C for 7 days. Conjugate **72** was purified under the same analytical conditions stated below. Overall yield (conjugation and purification): 12 %. Analytical HPLC (A: H<sub>2</sub>O + 0.045 % TFA, B: ACN + 0.036 % TFA, Phenomenex Jupiter Proteo C12 column, 4  $\mu$ m, 250 x 4.6 mm, 1 mL/min, linear gradient from 10 to 50 % of B in 30 min):  $t_R$  = 17.9 min (**Fig. A2.6**). MALDI-TOF MS (THAP, negative mode):  $m/z$ : 4629.9 [M-H]<sup>-</sup>; M calcd. for C<sub>194</sub>H<sub>240</sub>N<sub>84</sub>O<sub>51</sub>S<sub>2</sub> 4625.8.

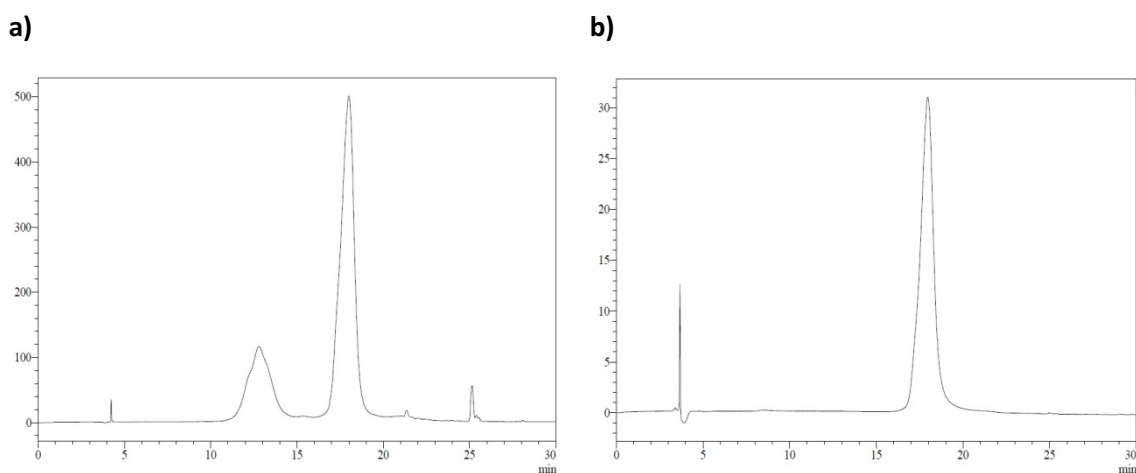


Figure A2.6: a) Crude **72**, and b) pure **72**. The broad peak with  $t_R$  = 13 min corresponds to **Cys-PNA**

### Cyclic Maleimido\*-Cys(S<sup>t</sup>Bu)-Glu-Trp-Tyr-Tyr-Tyr-Glu-Cys\*-NH<sub>2</sub> (**Pre-73**)

Fmoc-Cys(S<sup>t</sup>Bu)-OH and **1** were incorporated onto 115  $\mu$ mol of Fmoc-Glu(O<sup>t</sup>Bu)-Trp(Boc)-Tyr(<sup>t</sup>Bu)-Tyr(<sup>t</sup>Bu)-Tyr(<sup>t</sup>Bu)-Glu(O<sup>t</sup>Bu)-Cys(Trt)-resin. The protected peptide-resin was treated with TFA/TIS/H<sub>2</sub>O 90:5:5 (2 hours) and purified (A: H<sub>2</sub>O + 0.1 % TFA, B: ACN + 0.1 % TFA, Phenomenex Jupiter Proteo C12 column, 10  $\mu$ m, 250 x 10 mm, 3 mL/min, isocratic gradient 45 % of B in 30 min,  $t_R$  = 6.1 min). Overall yield (synthesis and purification): 37 %. Analytical HPLC (A: H<sub>2</sub>O + 0.045 % TFA, B: ACN + 0.036 % TFA, Phenomenex Jupiter Proteo C12 column, 4  $\mu$ m, 250 x 4.6 mm, 1 mL/min, 80°C, linear gradient from 30 to 60 % of B in 30 min):  $t_R$  = 15.6 min

(Fig. A2.7). MALDI-TOF MS (THAP, negative mode):  $m/z$ : 1395.1 [M-H]<sup>-</sup>; M calcd. for C<sub>65</sub>H<sub>77</sub>N<sub>11</sub>O<sub>18</sub>S<sub>3</sub> 1395.5. MALDI-TOF MS after reaction with H<sub>2</sub>O<sub>2</sub> (THAP, negative mode):  $m/z$ : 1394.8 [M-StBu+48+16-2H+Na]<sup>-</sup>; M-S<sup>t</sup>Bu+48+16 calcd. for C<sub>61</sub>H<sub>69</sub>N<sub>11</sub>O<sub>22</sub>S<sub>2</sub> 1371.4.

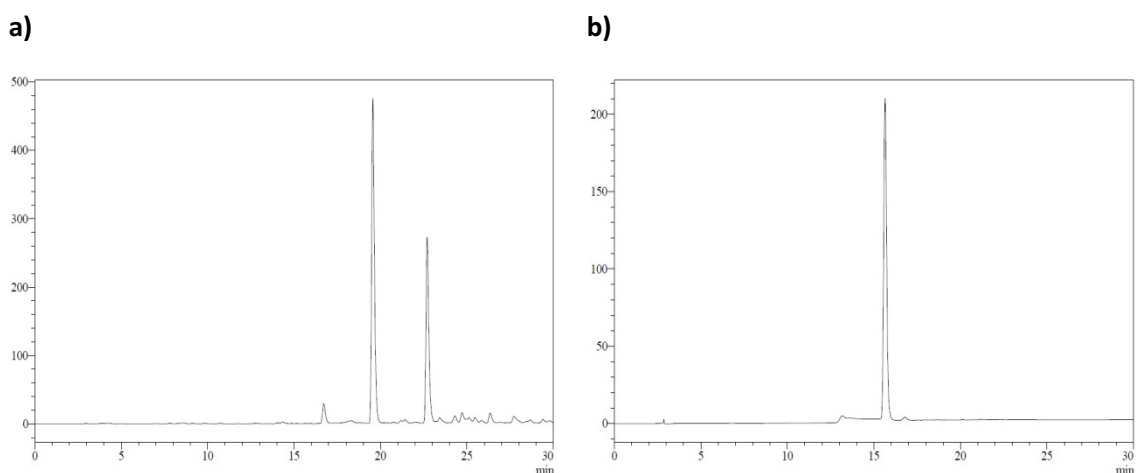


Figure A2.7: a) Crude cyclic peptide **Pre-73**, and b) pure **Pre-73** (80°C)

#### Cyclic Maleimido\*-Cys-Glu-Trp-Tyr-Tyr-Glu-Cys\*-NH<sub>2</sub> (Dep-Pre-73)

This protocol is described in both Paper 2 (Organic Letters) and Paper 3 (J. Org. Chem.). Deprotection of **Pre-73**: 200 nmol, 0.3 mM in a 2:1 (v/v) water/ACN mixture, TCEP (100 eq), pH = 5, 40°C, overnight. The product was purified under the same analytical conditions stated below. Overall yield (deprotection and purification): 10 %. Analytical HPLC (A: H<sub>2</sub>O + 0.045 % TFA, B: ACN + 0.036 % TFA, Phenomenex Jupiter Proteo C12 column, 4 μm, 250 x 4.6 mm, 1 mL/min, linear gradient from 30 to 60 % of B in 30 min):  $t_R$  = 9.3 min (Fig. A2.8). MALDI-TOF MS (THAP, negative mode):  $m/z$ : 1306.8 [M-H]<sup>-</sup>; M calcd. for C<sub>61</sub>H<sub>69</sub>N<sub>11</sub>O<sub>18</sub>S<sub>2</sub> 1307.4.

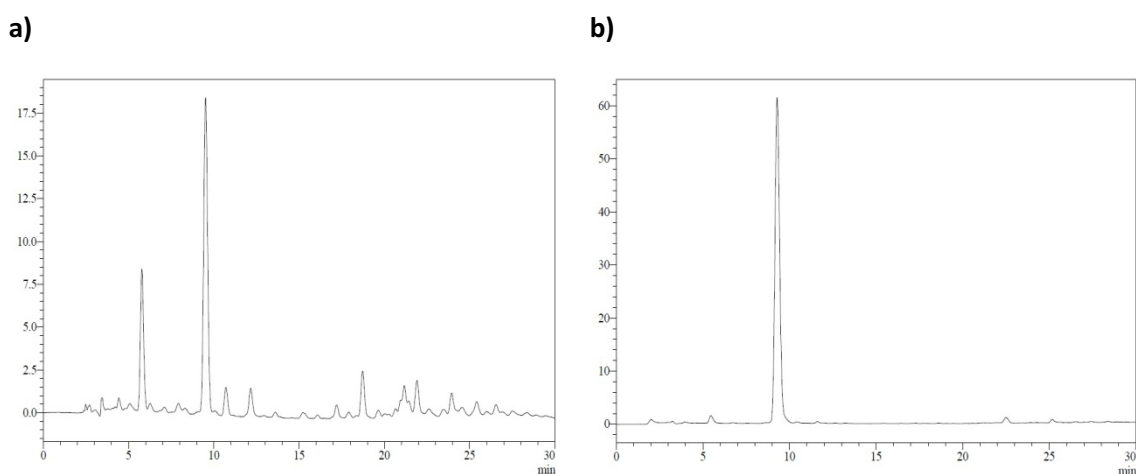
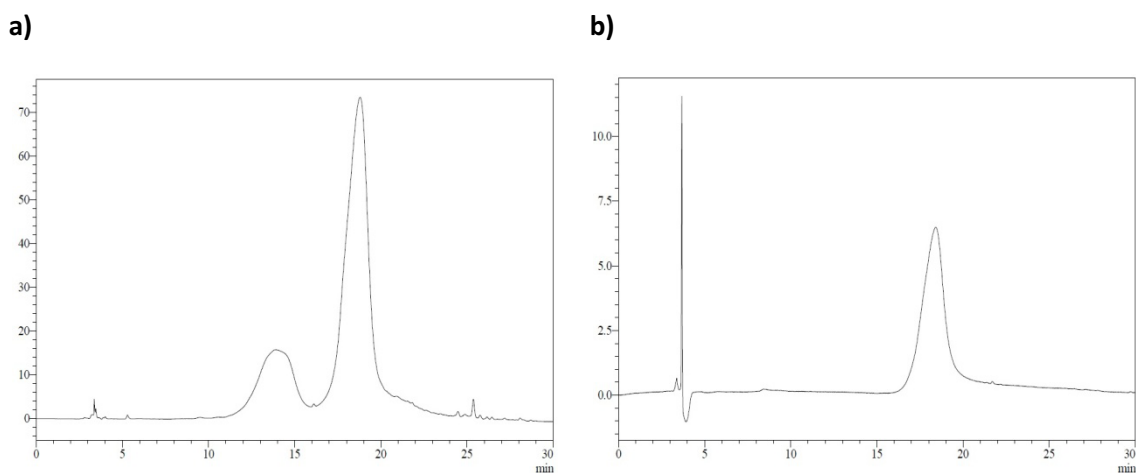


Figure A2.8: a) Deprotected cyclic peptide **Dep-Pre-73** crude, and b) pure **Dep-Pre-73**

**Conjugate 73**

**Dep-Pre-73** (50 nmol) and **Mal-PNA** (prepared by Natalia Masnou) (50 nmol) were dissolved in a 2:1 (v/v) water/ACN mixture to a 150  $\mu$ M final PNA concentration and heated at 50°C for 10 days. Conjugate **73** was purified under the same analytical conditions stated below. Overall yield (conjugation and purification): 24 %. Analytical HPLC (A: H<sub>2</sub>O + 0.045 % TFA, B: ACN + 0.036 % TFA, Phenomenex Jupiter Proteo C12 column, 4  $\mu$ m, 250 x 4.6 mm, 1 mL/min, 80°C, linear gradient from 10 to 50 % of B in 30 min):  $t_R$  = 18.8 min (**Fig. A2.9**). MALDI-TOF MS (THAP, positive mode):  $m/z$ : 4498.8 [M+H]<sup>+</sup>; M calcd. for C<sub>128</sub>H<sub>228</sub>N<sub>82</sub>O<sub>50</sub>S<sub>2</sub> 4497.7.

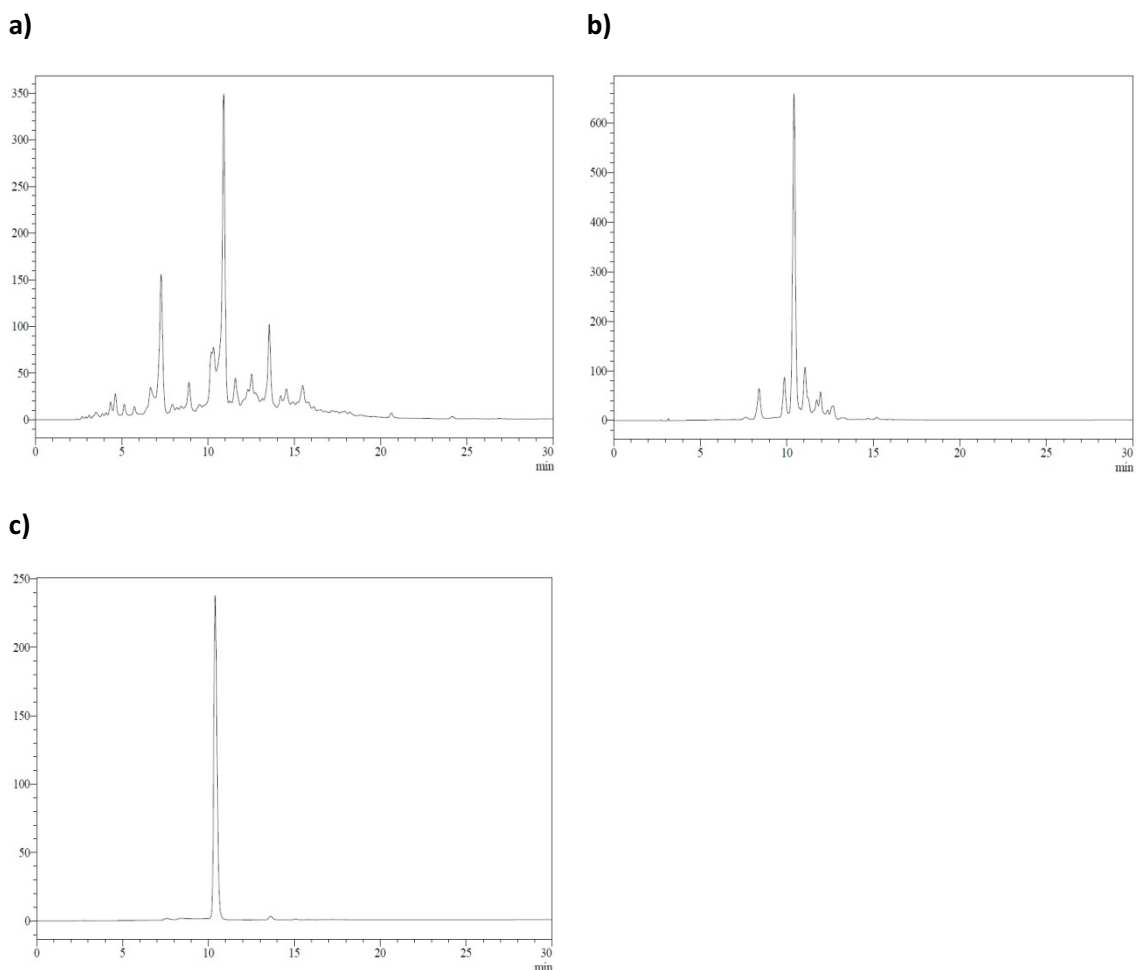


**Figure A2.9:** a) Crude **73**, and b) pure **73**. The peak eluting at 14 min corresponds to **Mal-PNA**

**Cyclic Maleimido\*-Lys(protected maleimide)-Glu-Trp-Tyr-Tyr-Tyr-Glu-Trp-Tyr-Tyr-Tyr-Glu-Cys\*-NH<sub>2</sub> (Pre-74)**

**4** and **1** were incorporated onto 37  $\mu$ mol of Fmoc-Glu(O<sup>t</sup>Bu)-Trp(Boc)-Tyr(<sup>t</sup>Bu)-Tyr(<sup>t</sup>Bu)-Tyr(<sup>t</sup>Bu)-Glu(O<sup>t</sup>Bu)-Trp(Boc)-Tyr(<sup>t</sup>Bu)-Tyr(<sup>t</sup>Bu)-Tyr(<sup>t</sup>Bu)-Glu(O<sup>t</sup>Bu)-Cys(Trt)-resin. The protected peptide-resin was treated with TFA/*m*-cresol/TIS/H<sub>2</sub>O 85:5:5:5 (2 x 1 hour) and purified (A: H<sub>2</sub>O + 0.1 % TFA, B: ACN + 0.1 % TFA, Phenomenex Jupiter Proteo C12 column, 10  $\mu$ m, 250 x 10 mm, 3 mL/min, 80°C, linear gradient from 35 to 65 % of B in 30 min,  $t_R$  = 10.5 min). Overall yield (synthesis and purification): 31 %, 72 % purity (**Fig. A2.10, a and b**). An analytical sample of the compound was repurified under the same conditions. Analytical HPLC (A: H<sub>2</sub>O + 0.045 % TFA, B: ACN + 0.036 % TFA, Phenomenex Jupiter Proteo C12 column, 4  $\mu$ m, 250 x 4.6 mm, 1 mL/min, 80°C, linear gradient from 35 to 65 % of B in 30 min):  $t_R$  = 10.4 min (**Fig. A2.10, c**). MALDI-TOF MS (THAP, negative mode):  $m/z$ : 2287.9 [M-furan-H]<sup>-</sup>; M calcd. for C<sub>120</sub>H<sub>133</sub>N<sub>19</sub>O<sub>32</sub>S 2383.9; M-furan calcd. for C<sub>114</sub>H<sub>125</sub>N<sub>19</sub>O<sub>31</sub>S 2287.8. MALDI-TOF MS after reaction with H<sub>2</sub>O<sub>2</sub> (THAP, positive mode):  $m/z$ : 2327.4 [M-furan+16+Na]<sup>+</sup>; M+16 calcd. for C<sub>120</sub>H<sub>133</sub>N<sub>19</sub>O<sub>33</sub>S 2399.9, M-furan+16 calcd. for C<sub>114</sub>H<sub>125</sub>N<sub>19</sub>O<sub>32</sub>S 2303.8.





**Figure A2.10:** a) Crude cyclic peptide **Pre-74**, b) purified **Pre-74**, and c) re-purified **Pre-74**. The peak eluting at 8.4 min (a) corresponds to the desired compound with the maleimide already deprotected

**Cyclic Maleimido\*-Lys(maleimide)-Glu-Trp-Tyr-Tyr-Tyr-Glu-Trp-Tyr-Tyr-Tyr-Glu-Cys\*-NH<sub>2</sub> (Dep-Pre-74)**

This protocol is described in Paper 1 (Bioconjugate Chemistry). Deprotection of **Pre-74** (72 % purity): 100 nmol, 100  $\mu$ M, 2 mL 1:1 (v/v) methanol/water mixture, MW heating, 90°C, 90 min. The product was purified under the same analytical conditions stated below. Overall yield (deprotection and purification): 13 %. Analytical HPLC (A: H<sub>2</sub>O + 0.045 % TFA, B: ACN + 0.036 % TFA, Phenomenex Jupiter Proteo C12 column, 4  $\mu$ m, 250 x 4.6 mm, 1 mL/min, 80°C, linear gradient from 35 to 65 % of B in 30 min):  $t_R$  = 6.2 min (**Fig. A2.11**). MALDI-TOF MS (THAP, negative mode):  $m/z$ : 2287.5 [M-H]<sup>-</sup>; M calcd. for C<sub>114</sub>H<sub>125</sub>N<sub>19</sub>O<sub>31</sub>S 2287.9.

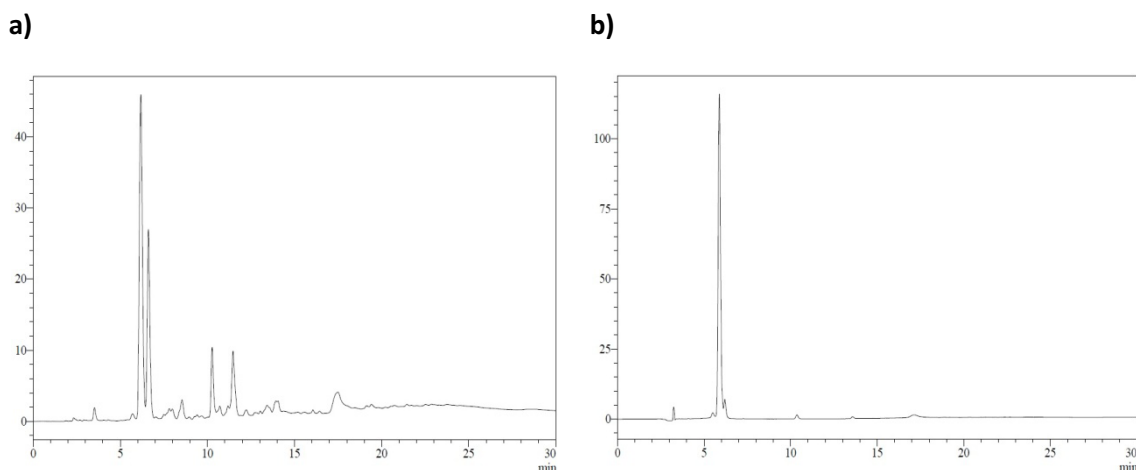


Figure A2.11: a) Crude deprotected cyclic peptide **Dep-Pre-74**, and b) pure **Dep-Pre-74**

### Conjugate 74

**Dep-Pre-74** (100 nmol) and **Cys-PNA** (prepared by Natalia Masnou) (50 nmol) were dissolved in a 2:1 (v/v) water/ACN mixture to a 150  $\mu$ M final PNA concentration and heated at 50°C for 8 days. Conjugate **74** was purified under the same analytical conditions stated below. Overall yield (conjugation and purification): 23 %. Analytical HPLC (A: H<sub>2</sub>O + 0.045 % TFA, B: ACN + 0.036 % TFA, Phenomenex Jupiter Proteo C12 column, 4  $\mu$ m, 250 x 4.6 mm, 1 mL/min, linear gradient from 15 to 55 % of B in 30 min):  $t_R$  = 17.8 min (**Fig. A2.12**). MALDI-TOF MS (THAP, negative mode):  $m/z$ : 5430.1 [M-H]<sup>-</sup>; M calcd. for C<sub>237</sub>H<sub>284</sub>N<sub>90</sub>O<sub>61</sub>S<sub>2</sub> 5435.1.

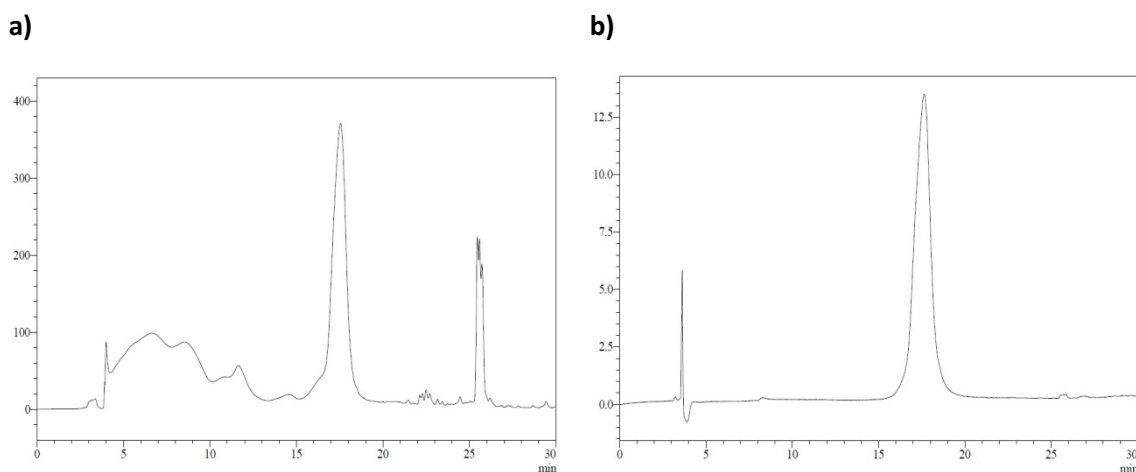


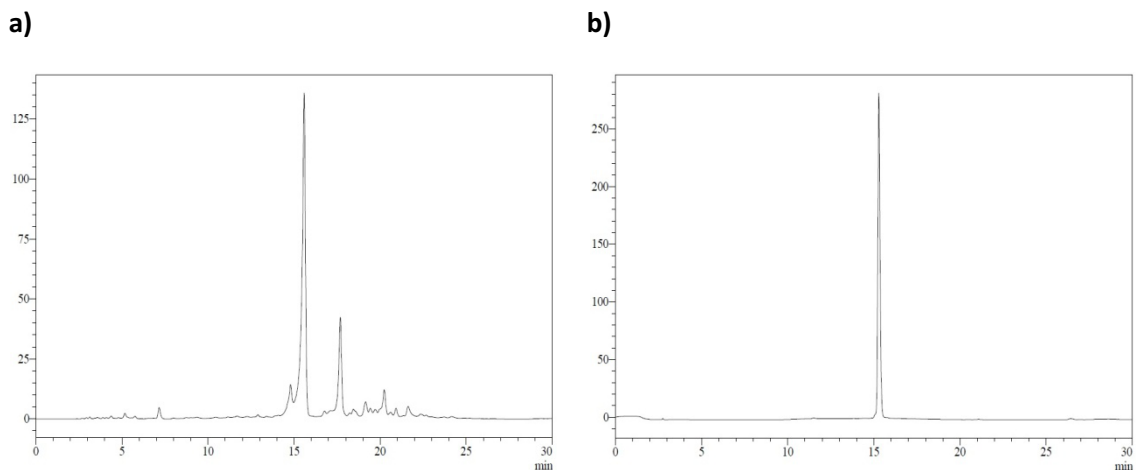
Figure A2.12: a) Crude **74**, and b) pure **74**. The peak with  $t_R$  = 25.5 min (a) corresponds to unreacted **Dep-Pre-74**

### Cyclic Maleimido\*-Cys(S<sup>t</sup>Bu)-Glu-Trp-Tyr-Tyr-Tyr-Glu-Trp-Tyr-Tyr-Tyr-Glu-Cys\*-NH<sub>2</sub> (Pre-75)

Fmoc-Cys(S<sup>t</sup>Bu)-OH and **1** were incorporated onto 37  $\mu$ mol of Fmoc-Glu(O<sup>t</sup>Bu)-Trp(Boc)-Tyr(<sup>t</sup>Bu)-Tyr(<sup>t</sup>Bu)-Tyr(<sup>t</sup>Bu)-Glu(O<sup>t</sup>Bu)-Trp(Boc)-Tyr(<sup>t</sup>Bu)-Tyr(<sup>t</sup>Bu)-Tyr(<sup>t</sup>Bu)-Glu(O<sup>t</sup>Bu)-Cys(Trt)-resin. The protected peptide-resin was treated with TFA/*m*-cresol/TIS/H<sub>2</sub>O 85:5:5:5 (2 x 1 hour) and purified (A: H<sub>2</sub>O + 0.1 % TFA, B: ACN + 0.1 % TFA, Phenomenex Jupiter Proteo C12 column, 10  $\mu$ m, 250 x 10 mm, 3 mL/min, linear gradient from 40 to 70 % of B in 30 min,  $t_R$  = 10.5 min). Overall yield (synthesis and purification): 39 %. Analytical HPLC: (A: H<sub>2</sub>O + 0.045 % TFA, B: ACN

## Appendix 2: Experiments and compounds not included in the papers

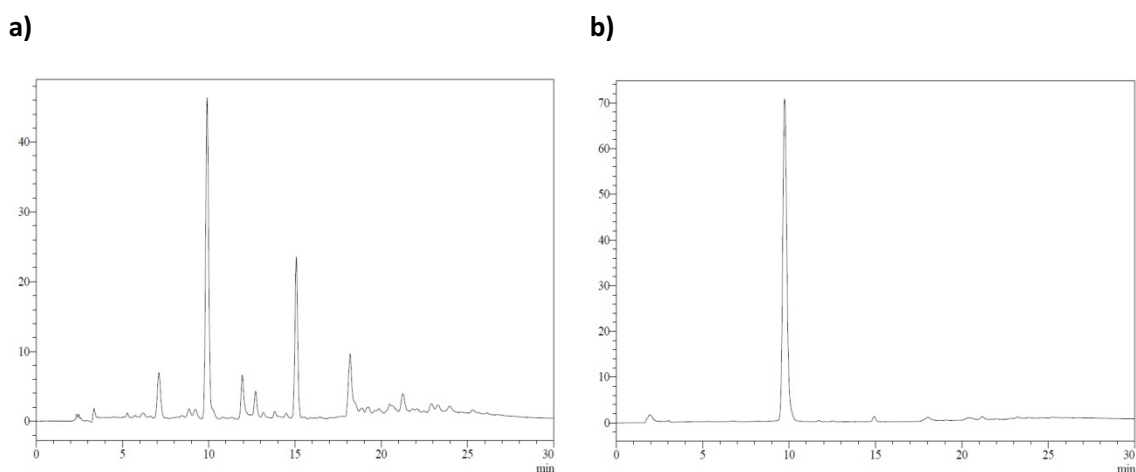
+ 0.036 % TFA, Phenomenex Jupiter Proteo C12 column, 4  $\mu$ m, 250 x 4.6 mm, 1 mL/min, 80°C, linear gradient from 35 to 65 % of B in 30 min):  $t_R$  = 15.3 min (**Fig. A2.13**). MALDI-TOF MS (THAP, negative mode):  $m/z$ : 2199.7 [M-H]<sup>-</sup>; M calcd. for C<sub>108</sub>H<sub>121</sub>N<sub>17</sub>O<sub>28</sub>S<sub>3</sub> 2199.8. MALDI-TOF MS after reaction with H<sub>2</sub>O<sub>2</sub> (THAP, negative mode):  $m/z$ : 2247.1 [M+16+32-H]<sup>-</sup>; M+16+32 calcd. for C<sub>108</sub>H<sub>121</sub>N<sub>17</sub>O<sub>31</sub>S<sub>3</sub> 2247.7.



**Figure A2.13:** a) Crude cyclic peptide **Pre-75**, and b) pure **Pre-75**

### Cyclic Maleimido\*-Cys-Glu-Trp-Tyr-Tyr-Glu-Trp-Tyr-Tyr-Tyr-Glu-Cys\*-NH<sub>2</sub> (Dep-Pre-75)

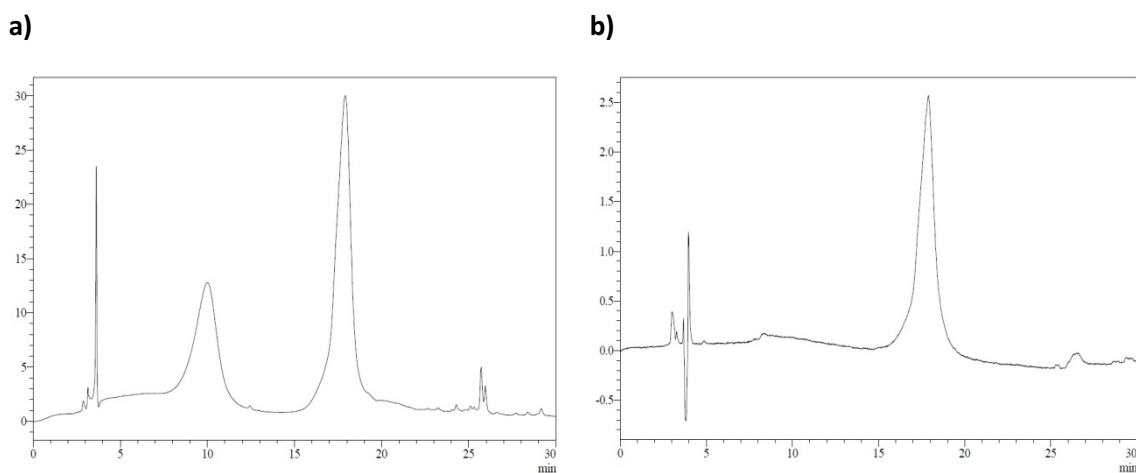
This protocol is described in Paper 2 (Organic Letters) and Paper 3 (J. Org. Chem.). Deprotection of **Pre-75**: 200 nmol, 0.3 mM in a 2:1 (v/v) water/ACN mixture, TCEP (100 eq), pH = 5, 40°C, 60 h. The product was purified under the same analytical conditions stated below. Overall yield (deprotection and purification): 6 %. Analytical HPLC (A: H<sub>2</sub>O + 0.045 % TFA, B: ACN + 0.036 % TFA, Phenomenex Jupiter Proteo C12 column, 4  $\mu$ m, 250 x 4.6 mm, 1 mL/min, linear gradient from 35 to 65 % of B in 30 min):  $t_R$  = 9.9 min (**Fig. A2.14**). MALDI-TOF MS (THAP, negative mode):  $m/z$ : 2110.9 [M-H]<sup>-</sup>; M calcd. for C<sub>104</sub>H<sub>113</sub>N<sub>17</sub>O<sub>28</sub>S<sub>2</sub> 2111.7.



**Figure A2.14:** a) Crude deprotected cyclic peptide **Dep-Pre-75**, and b) pure **Dep-Pre-75**

**Conjugate 75**

**Dep-Pre-75** (48 nmol) and **Mal-PNA** (prepared by Natalia Masnou) (50 nmol) were dissolved in a 2:1 (v/v) water/ACN mixture to a 150  $\mu$ M final PNA concentration and heated at 50°C for 8 days. Conjugate **75** was purified under the same analytical conditions stated below. Overall yield (conjugation and purification): 23 %. Analytical HPLC (A: H<sub>2</sub>O + 0.045 % TFA, B: ACN + 0.036 % TFA, Phenomenex Jupiter Proteo C12 column, 4  $\mu$ m, 250 x 4.6 mm, 1 mL/min, linear gradient from 15 to 55 % of B in 30 min):  $t_R$  = 17.9 min (**Fig. A2.15**). MALDI-TOF MS (THAP, negative mode):  $m/z$ : 5303.7 [M-H]<sup>-</sup>; M calcd. for C<sub>231</sub>H<sub>272</sub>N<sub>88</sub>O<sub>60</sub>S<sub>2</sub> 5302.0.



**Figure A2.15:** a) Crude **75**, and b) pure **75**. The peak eluting at  $t_R$  = 9 min (a) corresponds to **Mal-PNA** and the peak with  $t_R$  = 25.5 min (a) corresponds to **Dep-Pre-75**



## Appendix 3: Publications

### Experimental contribution report

Paper I: *Bioconjugate Chemistry*. In this paper I did the synthesis of the monomers for peptide and peptoid synthesis, and all the work related with peptides and peptoids. Albert Sánchez synthesized the PNA monomer and performed the experiments with PNAs. Kapil Sharma carried out the synthesis of the peptoid monomer using another procedure.

Paper II: *Organic Letters*. I carried out all the experimental work.

Paper III: *Journal of Organic Chemistry*: I carried out all the experimental work.



# PAPER I





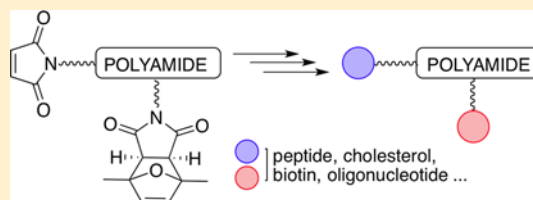
## Protected Maleimide Building Blocks for the Decoration of Peptides, Peptoids, and Peptide Nucleic Acids

Xavier Elduque, Albert Sánchez, Kapil Sharma, Enrique Pedroso, and Anna Grandas\*

Departament de Química Orgànica and IBUB, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1-11, 08028 Barcelona, Spain

### Supporting Information

**ABSTRACT:** Monomers allowing for the introduction of [2,5-dimethylfuran]-protected maleimides into polyamides such as peptides, peptide nucleic acids, and peptoids were prepared, as well as the corresponding oligomers. Suitable maleimide deprotection conditions were established in each case. The stability of the adducts generated by Michael-type maleimide–thiol reaction and Diels–Alder cycloaddition to maleimide deprotection conditions was exploited to prepare a variety of conjugates from peptide and PNA scaffolds incorporating one free and one protected maleimide. The target molecules were synthesized by using two subsequent maleimide-involving click reactions separated by a maleimide deprotection step. Carrying out maleimide deprotection and conjugation simultaneously gave better results than performing the two reactions subsequently.



### INTRODUCTION

Controlled ligation of different building blocks provides access to complex structures, libraries of products, and biomolecules modified à la carte. Biocompatible materials,<sup>1,2</sup> polymers with pharmacological application,<sup>3</sup> new drugs with higher affinity and specificity for the target,<sup>4,5</sup> biomolecules with optimized properties (increased half-life in biological fluids, enhanced uptake, incorporating reporter groups for imaging, etc.),<sup>6</sup> and complex nanostructures and supramolecular systems<sup>7,8</sup> have been prepared taking advantage of those chemical tools.

In the past few years different alternatives have been described for the modular construction of sophisticated molecules making use of so-called “click” reactions.<sup>9</sup> One of the options consisted of attaching different alkynes to a scaffold, combinations either of strained and terminal alkynes<sup>10</sup> or of unprotected and orthogonally protected alkynes.<sup>11–14</sup> Subsequent Huisgen reactions with azides, with or without Cu(I) or after the required deprotection treatment, allowed new building blocks to be placed where desired.

Other experiments also involving Cu(I)-catalyzed Huisgen cycloadditions have been made using scaffolds with two different appending groups, an azide and a group that can be transformed into either an azide<sup>15,16</sup> or an azide and a latent 1,3-dipole.<sup>17</sup> The alkyne–azide reaction was followed by functional group conversion (that is, transformation of the azide precursors into azides or of the latent 1,3-dipole into dipole), and this by the second cycloaddition.

Chemoselective reactions have been subsequently performed on scaffolds derivatized with two or three different functional groups. Huisgen cycloadditions, oxime and hydrazone formation, and thiol–maleimide reactions have been among the most commonly chosen orthogonal reactions.<sup>18–20</sup> The thiol–ene and inverse electron-demand Diels–Alder reactions are

becoming increasingly popular and have also been used in combination with other chemoselective processes.<sup>21–23</sup>

Finally, a combination of stepwise synthesis and attachment of different building blocks at different elongation stages also has allowed different moieties to be attached to a polyamide chain.<sup>24,25</sup>

Maleimides are versatile molecules susceptible to use in “click” reactions with thiols, as already mentioned, and with dienes. Unprotected maleimides are not compatible with basic, nucleophilic reagents, which may promote maleimide hydrolysis and/or add to the double bond. We have recently described that the *exo* cycloadduct obtained by reaction of an *N*-alkylated maleimide with 2,5-dimethylfuran is stable to concentrated aqueous ammonia at room temperature, which can be exploited for the on-resin synthesis of maleimido-oligonucleotides.<sup>26,27</sup> 5'-Maleimido-oligonucleotides were prepared by coupling derivatives incorporating protected maleimide cycloadducts followed by deprotection with ammonia and retro-Diels–Alder reaction.

The goal of this work was the preparation of maleimido-containing peptides, peptide nucleic acids (PNAs), and peptoids, and their use in conjugation reactions (Scheme 1).

Solid-phase synthesis of these polyamides requires reaction with bases, either to remove the Fmoc (9-fluorenylmethoxycarbonyl) temporary protecting group in the case of peptides and PNAs or to incorporate the construction units during peptoid assembly using the submonomer strategy.<sup>28</sup> Therefore, use of [2,5-dimethylfuran]-protected maleimide building blocks

Received: January 31, 2013

Revised: April 12, 2013

Published: April 14, 2013



542.3 [M+Na-Me<sub>2</sub>furan]<sup>+</sup>, 520.3 [M+H-Me<sub>2</sub>furan]<sup>+</sup>, M calcd. for C<sub>34</sub>H<sub>37</sub>N<sub>3</sub>O<sub>8</sub> 615.3. HRMS (ESI, negative mode): *m/z* 614.2481 [M-H]<sup>-</sup>, M calcd. for C<sub>34</sub>H<sub>37</sub>N<sub>3</sub>O<sub>8</sub> 615.2581.

***N*-[Protected maleimido]-*N*-(2-Fmoc-aminoethyl)glycine *tert*-Butyl Ester 4.** *N*-(2-Fmoc-aminoethyl)glycine *tert*-butyl ester·HCl (890 mg, 2.06 mmol) was dissolved in DCM (120 mL) and washed with a saturated NaHCO<sub>3</sub> aqueous solution (3 × 50 mL). The organic solution was dried over anhyd. Na<sub>2</sub>SO<sub>4</sub> and filtered, and the solvent was removed in vacuo. The resulting yellowish oil (810 mg of *N*-(2-Fmoc-aminoethyl)glycine *tert*-butyl ester, quantitative recovery) was dissolved in DMF (*N,N*-dimethylformamide, 3 mL) and bubbled with Ar. To this solution was added a mixture containing **1** (610 mg, 2.30 mmol), HOBt·H<sub>2</sub>O (1-hydroxybenzotriazole, 544 mg, 3.55 mmol), HATU (2-(7-aza-1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate, 873 mg, 2.30 mmol), and *N*-methylmorpholine (410 μL, 3.73 mmol) previously stirred at room temperature under an Ar atmosphere for 20 min. The resulting mixture was stirred for 60 h at room temperature under an Ar atmosphere. Solvent was removed under reduced pressure, and the resulting crude was purified by silica gel column chromatography eluting with a 1:1 (v/v) hexanes/AcOEt mixture. **4** was obtained as a pale yellowish oil (1.06 g, 80% overall yield).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ 7.75 (d, *J* = 7.6 Hz, 2H), 7.61 (t, *J* = 8.4 Hz, 2H), 7.39 (t, *J* = 7.4 Hz, 2H), 7.31 (t, *J* = 7.4 Hz, 2H), 6.28 (s, 2H), 5.97 (t, *J* = 6.2 Hz) and 5.64 (t, *J* = 5.4 Hz) (rotamers, 1H), 4.37 (m, 2H), 4.23 (m, 1H), 3.89 (s, 1H), 3.80 (m, 2H), 3.47 (m, 2H), 3.35 (m, 2H), 2.80 (s, 6H), 2.64 (t, *J* = 7.6 Hz), 2.52 (t, *J* = 7.1 Hz) (rotamers, 2H), 1.68 (s, 9H) ppm; <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>): δ 176.2, 166.1, 63.4, 61.7, 57.0, 43.9, 41.0, 36.7, 29.8, 29.5, 26.8, 24.5 ppm; HRMS (ESI, positive mode): *m/z* 644.2955 [M+H]<sup>+</sup>, 661.3227 [M+NH<sub>4</sub>]<sup>+</sup>, 666.2782 [M+Na]<sup>+</sup>, 1304.6093 [2M+NH<sub>4</sub>]<sup>+</sup>, 1309.5680 [2M+Na]<sup>+</sup>, M calcd. for C<sub>36</sub>H<sub>41</sub>N<sub>3</sub>O<sub>8</sub> 643.2887.

***N*-[Protected maleimido]-*N*-(2-Fmoc-aminoethyl)glycine 5.** A solution of **4** (500 mg, 0.777 mmol) in DCM (10 mL) was chilled in an ice bath. TFA (trifluoroacetic acid, 5 mL) and TIS (triisopropylsilane, 2 mL) were added, and the mixture was left to react at 0 °C for 30 min and then at room temperature until TLC analysis (hexanes/AcOEt 1:4, R<sub>f</sub> = 0.21) showed complete disappearance of **4** (3 h reaction time). Solvent and reagents were removed in vacuo, and the resulting oil was dissolved in the minimal amount of DCM and precipitated over cold hexanes. A white solid was obtained after decantation (455 mg, quantitative yield), which was thoroughly dried in a desiccator before its use in PNA synthesis.

IR (ATR),  $\nu$ : 3349, 2937, 1768, 1694, 1523, 1446, 1403, 1170 cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ 7.74 (d, *J* = 7.5 Hz, 2H), 7.59 (d, *J* = 5.5 Hz, 2H), 7.38 (t, *J* = 7.3 Hz, 2H), 7.30 (t, *J* = 6.7 Hz, 2H), 6.27 and 6.26 (2s, 2H), 5.78 (brs) and 5.69 (brs) (rotamers, 1H), 4.60 (brs) and 4.54 (brs) (rotamers, 1H), 4.37 (m, 2H), 4.21 (t, *J* = 6.4 Hz, 1H), 4.03 (m, 2H), 3.76 (brs, 4H), 3.49 (brs, 2H), 3.35 (brs, 2H), 2.80 (m, 2H), 2.68 (brs) and 2.55 (brs) (rotamers, 2H), 1.66 (s, 6H) ppm; <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>): δ 174.7, 172.6, 171.7, 156.7, 143.9, 141.3, 140.9, 127.7, 127.1, 125.1, 120.0, 87.6, 67.0, 52.5, 49.5, 49.2, 47.1, 39.4, 34.9, 30.1, 15.9 ppm; HRMS (ESI, positive mode): *m/z* 588.2328 [M+H]<sup>+</sup>, 605.2599 [M+NH<sub>4</sub>]<sup>+</sup>, 610.2159 [M+Na]<sup>+</sup>, 1175.4609 [2M+H]<sup>+</sup>, 1192.4849 [2M+NH<sub>4</sub>]<sup>+</sup>, 1197.4380 [2M+Na]<sup>+</sup>, M calcd. for C<sub>32</sub>H<sub>33</sub>N<sub>3</sub>O<sub>8</sub> 587.2261.

**2,5-Dimethylfuran-Protected Maleimide (exo + endo) 6.** Maleimide (3.0 g, 30.9 mmol) was dissolved in ACN

(acetonitrile, 55 mL). The solution was heated to 60 °C, and 2,5-dimethylfuran (8.3 mL, 77.3 mmol) was added. After overnight reaction at 60 °C, the mixture was evaporated to dryness to yield **6** as a 4:1 mixture of *exo* and *endo* isomers.

*Exo* isomer: <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ 6.31 (s, 2H), 2.88 (s, 2H), 1.73 (s, 6H) ppm; <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>): δ 175.0, 140.9, 87.7, 53.8, 15.8 ppm.

*Endo* isomer: <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ 6.31 (s, 2H), 3.27 (s, 2H), 1.78 (s, 6H) ppm; <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>): δ 174.4, 138.2, 135.1, 54.9, 18.5 ppm.

HRMS (ESI, positive mode): *m/z* 194.0915 [M+H]<sup>+</sup>, M calcd. for C<sub>10</sub>H<sub>11</sub>NO<sub>3</sub> 193.0739.

***N*-Boc-ethanolamine 7.**<sup>29</sup> Ethanolamine (6.1 g, 0.1 mol) was added to Boc<sub>2</sub>O (Boc = *tert*-butoxycarbonyl, 21.8 g, 0.1 mol) in three portions at 0 °C over 10 min, and the mixture was stirred for 2 h. A colorless precipitate formed with the evolution of heat. After filtration, the solid was purified by silica gel column chromatography eluting with a 6:1 hexanes/AcOEt mixture. The desired fractions (as assessed by TLC: hexanes/AcOEt 2:3, R<sub>f</sub> = 0.5, detection: reaction with anisaldehyde) were pooled and taken to dryness, which yielded pure **7** as a colorless oil (15.8 g, 98%).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ 5.05 (bs, 1H), 3.68 (m, 2H), 3.26 (t, 2H, *J* = 4.8 Hz), 2.83 (bs, 1H), 1.43 (s, 9H) ppm; <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>): δ 156.81, 79.63, 62.47, 43.10, 28.34 ppm.

***N*-[Protected maleimido]-ethanediamine-TFA 9 (exo).**

A mixture of protected maleimide **6** (500 mg, 2.59 mmol), *N*-Boc-ethanolamine **7** (278 mg, 1.73 mmol), and triphenylphosphine (679 mg, 2.59 mmol) was coevaporated with anhyd THF, dissolved in anhyd THF (20 mL), and chilled in an ice bath. DIAD (diisopropylazodicarboxylate, 511 μL, 2.59 mmol) was added dropwise. The mixture was reacted for 2 h at 0 °C and overnight at room temperature, and then the solvent was removed under reduced pressure. The resulting residue (which contained the mixture of *exo* and *endo* isomers) was treated with a 2:3 v/v MeOH/conc aq ammonia mixture overnight, and then concentrated under reduced pressure. CHCl<sub>3</sub> (20 mL) was added, and the pH of the aqueous phase was adjusted to 12–13. The two phases were separated, and the organic phase was reextracted with 0.01 M NaOH (20 mL, 3×), dried over MgSO<sub>4</sub>, and taken to dryness. <sup>1</sup>H NMR analysis showed that the resulting residue contained **8** as the *exo* isomer, plus byproducts resulting from the Mitsunobu reaction (triphenylphosphine oxide and DIADH<sub>2</sub>). Crude **8** was treated with a 3:7 TFA/DCM mixture (30 mL) for 1 h at room temperature, and then was taken to dryness. The desired product precipitated upon addition of anhydrous diethyl ether (2 mL). After centrifugation the supernatant was discarded and diethyl ether was added again to the precipitate. This step was repeated three times. The product was dried thoroughly under reduced pressure, to afford pure **9** (119 mg, 20%).

**8:** <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ 6.12 (s, 2H), 3.44 (t, *J* = 5.61 Hz, 2H), 3.10 (m, 2H), 2.65 (s, 2H), 1.51 (s, 6H), 1.22 (s, 9H) ppm.

**9:** IR (ATR),  $\nu$ : 3574, 2985, 1769, 1694, 1674, 1402, 1200, 1167, 1124 cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, acetone-*d*<sub>6</sub>): δ 6.38 (s, 2H), 3.96 (t, *J* = 5.33 Hz, 2H), 3.83 (t, *J* = 5.36 Hz, 2H), 2.95 (s, 2H), 1.59 (s, 6H) ppm; <sup>13</sup>C NMR (101 MHz, acetone-*d*<sub>6</sub>): δ 176.8, 142.7, 89.2, 54.8, 46.3, 38.3, 17.2 ppm; HRMS (ESI, positive mode): *m/z* 473.2388 [M+H]<sup>+</sup>, M calcd. for C<sub>24</sub>H<sub>32</sub>N<sub>4</sub>O<sub>6</sub> 472.2317.

**Syntheses of Conjugates.** Synthesis involved the following series of steps:

- Monoderivatized polyamides. (i) Elongation of polyamide chains by solid-phase synthesis, using commercially available building blocks and the protected maleimido monomers described here (3, 5, and 9 were used to introduce the protected maleimide moieties at any position within the chain, 3-maleimidopropanoic acid 1 was only used for *N*-terminal derivatization); (ii) acid treatment to remove all protecting groups except that on the maleimide moiety; (iii) maleimide deprotection following one of the procedures described below; (iv) conjugation (reaction with either a thiol or a diene, see below).
- Doubly derivatized polyamides. Steps (i) and (ii) were the same as above; (iii) first conjugation (reaction of the free maleimide with either a thiol or a diene); (iv) maleimide deprotection; (v) second conjugation (second reaction with either a thiol or a diene).
- Alternatively, in some cases maleimide deprotection and conjugation were carried out simultaneously (see below).

**Microwave-Promoted Maleimide Deprotection.** This procedure was used to deprotect maleimides appending from PNAs.

A solution (1000  $\mu\text{L}$ ) of [protected maleimido]-containing PNA in a 1:1 (v/v) MeOH/H<sub>2</sub>O mixture (25  $\mu\text{M}$  concentration) was introduced in a microwave vial and irradiated for 60 min at 100 °C. Solvent was removed under vacuum, and the resulting crude was dissolved in water for HPLC analysis and characterization by MALDI-TOF MS. The crude was used at the subsequent conjugation step without any purification.

**Maleimide Deprotection by Heating in Toluene.** This procedure was used to deprotect maleimides appending from peptides and peptoids.

The [protected maleimido]-containing peptide/peptoid was dried by coevaporation with anhydrous toluene (3–4 $\times$ ), and a new batch of anhydrous toluene was added (the amount that would be required to obtain a 25  $\mu\text{M}$  solution). The mixture was heated at 90 °C, toluene was removed under reduced pressure, and the crude was dissolved in water for HPLC analysis and mass spectrometric characterization. The crude was used at the subsequent conjugation step without any purification.

**Conjugation: General Procedure for Michael-Type Reactions.** Aliquots of aqueous (or aqueous–organic) solutions containing the required amounts of maleimido-containing oligomer and the corresponding thiol-containing compound (5- to 10-fold molar excess) were mixed, and the mixture was diluted with 0.5 M triethylammonium acetate, pH = 7.8–7.9 (final concentration of oligomer: 50–150  $\mu\text{M}$ ). The mixture was stirred at room temperature under an Ar atmosphere. Initially all experiments were carried out overnight; yet, we have verified that in most cases the reaction is complete in 1 h or even less. Shortening the reaction time prevents hydrolysis of the succinimide formed after the thiol-maleimide reaction, which has occasionally been observed (compound 21). Reaction crudes were analyzed by HPLC. Conjugates were purified by HPLC and characterized by MALDI-TOF MS.

When reagents were not soluble in water the reaction was carried out in organic solvent–aqueous buffer mixtures (see Supporting Information, synthesis of conjugate 22).

**Conjugation: General Procedure for Diels–Alder Cycloadditions.** The maleimido-containing oligomer and the diene-oligonucleotide were dissolved in water (final concentration of maleimido-containing oligomer: 50–250  $\mu\text{M}$ ). The molar ratio varied between 1:1 and 1:5, the component in excess generally being the least valuable one, in other words, the most easily accessible (which of the two reagents is used in excess was found to have no influence on the reaction yield; data not shown). The mixture was reacted at 37 °C, generally overnight. Reaction crudes were analyzed by HPLC. Conjugates were purified by HPLC and characterized by MALDI-TOF MS.

**One-Pot [Maleimide Deprotection + Conjugation] Reactions.** After coevaporation with anhydrous toluene (3 $\times$ ), the protected maleimido-containing polyamide, and either the diene (1–2 equiv) or the thiol-containing compound (10-fold molar excess) were dissolved or suspended in anhydrous toluene (concentration of the polyamide in the mixture: 25  $\mu\text{M}$ ). The mixture was reacted for 6 h at 90 °C, and then taken to dryness. HPLC analysis of the crudes (dissolved in water) showed complete deprotection of maleimides and high to quantitative conjugation yields.

This procedure was found to be a good alternative to accelerate cycloaddition reactions not completed even after days reaction time (preparation of conjugate 27; see Supporting Information).

## RESULTS AND DISCUSSION

**Synthesis of [Protected Maleimido]-Containing Monomers.** For the synthesis of the lysine derivative 3 (Scheme 2a), the carboxyl group of 1<sup>26</sup> was activated to give *N*-hydroxysuccinimido ester 2, and subsequent reaction of 2 with the  $\epsilon$ -amine of Fmoc-*L*-lysine linked the protected maleimide to the amino acid side chain and gave the target derivative 3.

The PNA monomer 5 (Scheme 2b) was prepared by first reacting 1 with the secondary amine of *N*-(*N*-Fmoc-2-aminoethyl)glycine *tert*-butyl ester, which was followed by an acid treatment to deprotect the carboxyl group of 4.

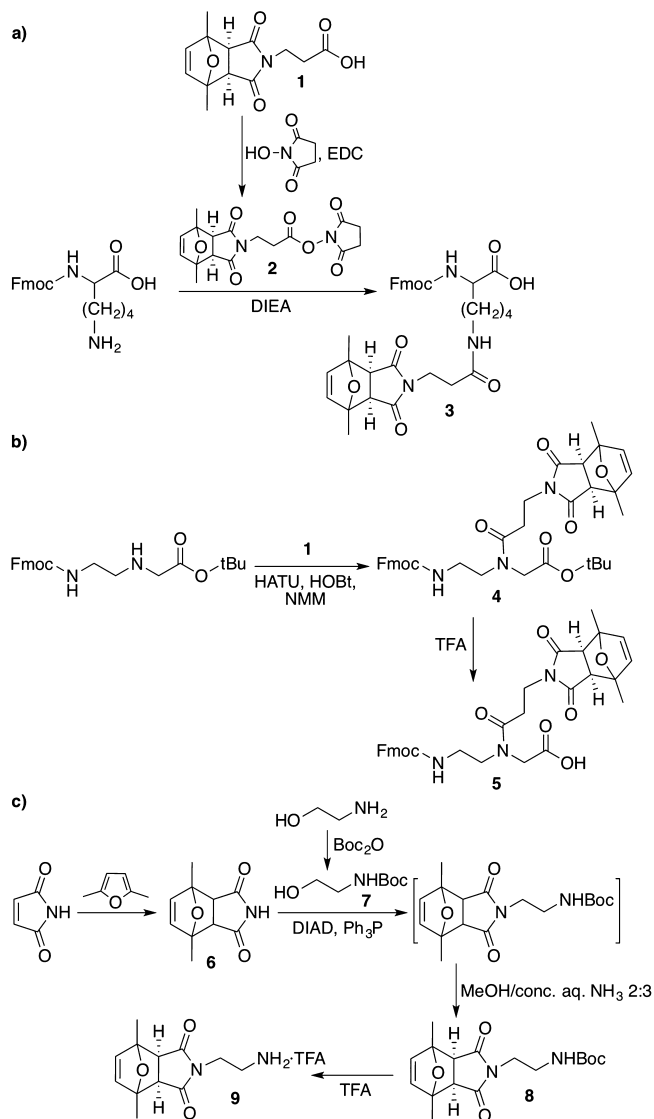
Scheme 2c summarizes the reactions carried out to obtain [protected maleimido]-amine 9. In the preliminary steps, 2,5-dimethylfuran-protected maleimide (6, mixture of *exo* and *endo* isomers) and *N*-Boc-aminoethanol (7) were prepared. Then a DIAD/Ph<sub>3</sub>P-mediated Mitsunobu reaction between 6 and 7, followed by reaction with a 2:3 MeOH/conc aq ammonia mixture to get rid of the *endo* isomer<sup>26</sup> afforded the fully protected maleimido-amine 8. Finally, removal of the Boc group by treatment with trifluoroacetic acid furnished amine 9 (trifluoroacetate salt).

These [protected maleimido]-containing monomers were used in the solid-phase synthesis of peptide, PNA, and peptoid polyamides following well-known procedures (see Supporting Information for experimental conditions, and Schemes S1–S3). Syntheses proceeded smoothly, showing that all derivatives behaved as standard building blocks.

**Maleimide Deprotection and Conjugation Reactions with Polyamides.** Conditions for maleimide deprotection (retro-Diels–Alder reaction) were examined for the different polyamides. Peptide 10, PNA 14, and peptoid 18 (Scheme 3) were heated either in a metal block (suspension in toluene) or in a microwave oven (MeOH/H<sub>2</sub>O solution), and the resulting crudes were analyzed by HPLC.

Peptide 10 afforded fairly heterogeneous crudes. Best results were achieved by heating in toluene (Scheme 3a, a.1; Figure

**Scheme 2. Synthesis of Compounds Suitable for the Introduction of Protected Maleimides into Peptides (a), PNAs (b), and Peptoids (c)**



S2), but the yield was not very good, and the longer the reaction time, the lower the quality of the crude.

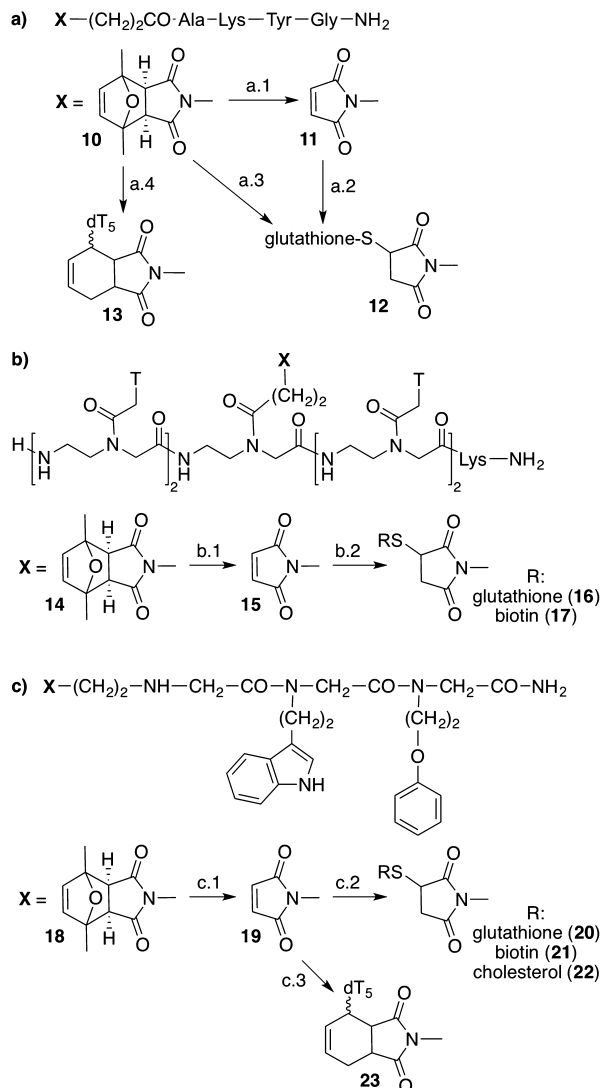
In the case of the PNA chain (**14**, Scheme 3b, b.1), highly homogeneous maleimido-PNA crudes (**15**) were obtained with any of the two deprotection procedures, the yield being higher when using the microwave oven (>95%, Figure S6) than when heating in toluene (ca. 70%).

As to peptoid **18** (Scheme 3c, c.1), microwave irradiation did not afford a sufficiently satisfactory crude. Yet, heating in toluene provided the target maleimido-peptoid (**19**) in good yield (>90%) and high purity (Figure S9).

Reaction (Scheme 3) of the maleimido-containing polyamides (peptide **11**, PNA **15**, and peptoid **19**) with thiol- and diene-containing compounds (Figure 1) afforded the target conjugation products (**12**, **16**, **17**, and **20–23**; Figures S3b, S7 and S10) and confirmed that maleimide deprotection had afforded fully reactive maleimides in all cases.

With the aim to improve the low yields associated with maleimide deprotection in peptides, we considered the possibility of carrying out one-pot deprotection and con-

**Scheme 3. Structures of Peptide (a), PNA (b) and Peptoid (c) Incorporating a Protected Maleimide Unit, And Reactions Subsequently Carried Out<sup>a</sup>**

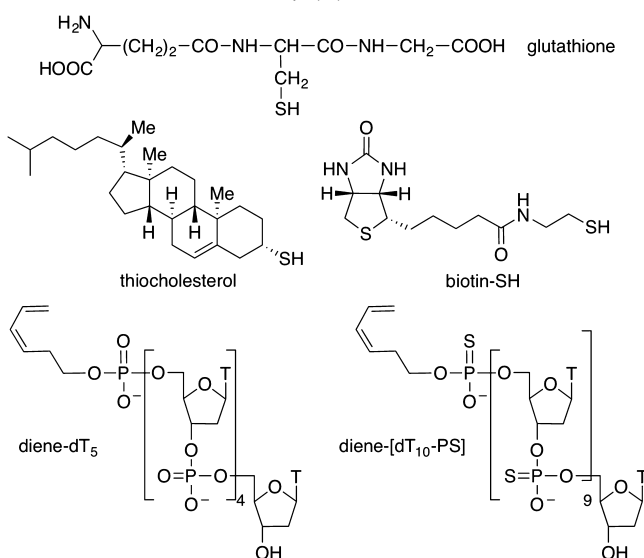


<sup>a</sup>Reaction conditions/reagents: a.1, heating in toluene; a.2, + glutathione; a.3, heating in toluene in the presence of glutathione; a.4, heating in toluene in the presence of diene-dTs; b.1, microwave irradiation; b.2, + glutathione/biotin-SH; c.1, heating in toluene; c.2, + glutathione/biotin-SH/thiocholesterol; c.3, + diene-dTs.

jugation reactions, in other words, of deprotecting the maleimide in the presence of either the thiol or the diene. Previously, the stability of the adducts resulting from conjugation reactions (Michael thiol-ene and Diels-Alder) to maleimide deprotection conditions was verified using standard samples of conjugates **12** and **13** (see Supporting Information, section 4). Both **12** and **13** remained undegraded, which indicated that maleimide deprotection conditions are compatible with maleimide-involving conjugation reactions (Figure S13).

When peptide **10** was suspended in toluene and heated in the presence of glutathione over a period of 6 h (Scheme 3a, a.3), the target conjugate **12** was obtained in high yield and purity (Figure S3b). Likewise, conjugate **13** was the main product in the crude (Figure S4) when **10** and diene-dTs were reacted under the same conditions (Scheme 3a, a.4). These

H-CysLysGluThrAlaAlaAlaLysPheGluArgGlnHisMetAspSerSerThrSerAlaAla-OH  
Cys-peptide



**Figure 1.** Structures of thiols and dienes used in conjugation reactions.

results suggest that the thiol or the diene present in the reaction medium drive the equilibrium of the retro-Diels–Alder reaction by trapping the free maleimide, and that the conjugation reaction is quick enough so as to prevent the maleimide from being degraded. Therefore, when the two reactions are simultaneously carried out the overall yield is higher, and the final conjugate is much more homogeneous. This result has also been exploited for the preparation of cyclic oligonucleotides.<sup>30</sup>

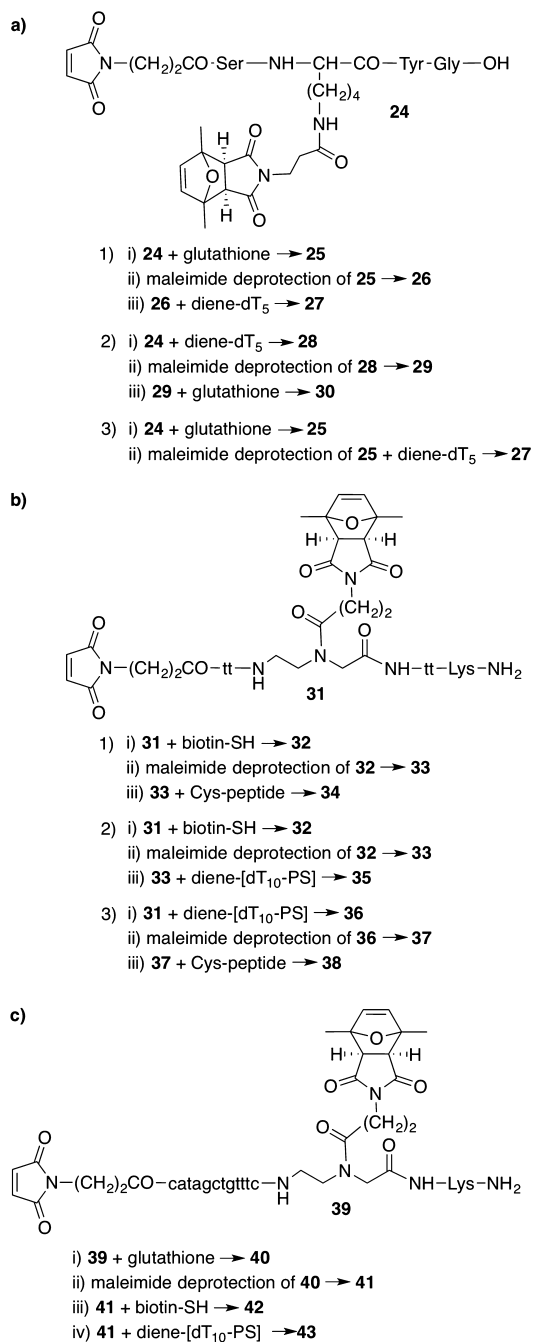
Both the Michael-type and Diels–Alder reactions provide conjugates containing succinimide rings (see structures in Schemes 1 and 3). These succinimides were fairly stable in aqueous media, with hydrolysis taking place only after several cycles of dissolving in water and lyophilization, except in the case of conjugate **21**. **21** could only be obtained if **19** was reacted with an excess (10 equiv) of biotin-SH for 5 min, after which time the mixture had to be immediately frozen and lyophilized. Longer reaction times gave crudes in which the main product was the conjugate with the hydrolyzed succinimide. Conjugate **28** (see Scheme 4 and below) also underwent some hydrolysis (37%) upon microwave-promoted maleimide deprotection.

Two maleimide-involving consecutive click reactions were carried out on peptide **24**, PNA **31**, and a PNA sequence described to exhibit antibactericidal antisense effects<sup>31</sup> (**39**) (Scheme 4).

Synthesis of conjugates differing in the position to which glutathione and diene-dT<sub>5</sub> were linked to peptide **24** (Scheme 4a, 1.i and 2.i, Figures S15 and S16) involved, first, reactions with the *N*-terminal free maleimide, which yielded conjugates **25** and **28**. **25** and **28** were then heated in toluene to deprotect the internal maleimides, providing compounds **26** and **29**. Finally, reaction of **26** and **29** with diene-dT<sub>5</sub> and glutathione, respectively, afforded the doubly derivatized regioisomeric conjugates **27** and **30**.

As in previous experiments, the maleimide deprotection yield was lower for the peptide–peptide conjugate (**25**) than for the peptide–oligonucleotide one (**28**). Yet, carrying out the deprotection and conjugation reactions simultaneously (Scheme 4a, 3.ii) significantly ameliorated (Figure S15d)

#### Scheme 4. Structures of Peptide (a) and PNA Scaffolds (b, c) Incorporating Two Maleimides, One Protected and One Unprotected, Used for the Attachment of Two Different Moieties, and Subsequent Deprotection and Conjugation Reactions



both the deprotection and the conjugation yield (the room temperature cycloaddition providing **27** from **26** was extremely slow). Hence, and in agreement with the outcome of the experiments described above, the yield of maleimide deprotection in the presence of a maleimide-trapping compound was higher than in its absence. It is also worth noting that maleimide deprotection (retro-Diels–Alder reaction) and Diels–Alder conjugation can be simultaneously carried out because of the higher stability of diene-maleimide cycloadducts with respect to 2,5-dimethylfuran-maleimide adducts.

As to PNA **31**, the final conjugates resulted from combining either two Michael-type reactions or, again, Michael and Diels–Alder reactions in either order (Scheme 4b). In the first case, conjugate **34** incorporated biotin and a 21-amino-acid peptide (Cys-peptide, Figure 1). In conjugate **35** the PNA scaffold was attached to biotin and a 10-mer phosphorothioate oligonucleotide, while in **38** the 10-mer phosphorothioate oligonucleotide and the Cys-peptide were linked to the PNA chain (Figures S18 and S19).

Finally, the two ends of PNA **39** were successfully modified (Scheme 4c) using the same three-step conjugation/maleimide deprotection/conjugation procedure (Figure S20). Hence, this methodology is compatible with PNAs containing all of the four nucleobases, and allows PNAs to be derivatized and keep their full recognition potential.

The fact that no side reactions were detected during the preparation of conjugates **35**, **38**, and **43** further supports our previous conclusions on reactions involving thiophosphate diesters and maleimides.<sup>27</sup> Both HPLC and mass spectrometric analyses (see Supporting Information) are consistent with (i) phosphorothioate diesters not interfering with maleimide–diene Diels–Alder cycloadditions (reactions affording **35** and **43**), (ii) microwave-promoted maleimide deprotection of the [phosphorothioate oligonucleotide]-PNA conjugate cleanly affording the target compound (**37**), and (iii) successful reaction of the resulting free maleimide with the Cys-peptide (to yield **38**).

## CONCLUSIONS

2,5-Dimethylfuran-protected maleimide moieties suitable for use in the synthesis of [protected maleimido]-containing peptides, PNAs, and peptoids were prepared and incorporated into the corresponding oligomers.

Best conditions for maleimide deprotection were found to vary depending on the oligomer, but in all cases fully reactive maleimides were obtained. Deprotection of peptide-linked maleimides was troublesome, since the reaction took place in low yield and was less clean than with PNAs and peptoids. Yet, these problems could be overcome by carrying out simultaneously maleimide deprotection and conjugation, since the adducts generated from Michael-type and Diels–Alder reactions remain stable under maleimide deprotection conditions. One-pot maleimide deprotection and conjugation furnished crudes substantially more homogeneous, increased the conjugation yield, and accelerated slow cycloadditions.

Finally, peptides and PNAs incorporating the new [protected maleimido]-containing derivatives and an unprotected maleimide at the *N*-terminal were employed as scaffolds to which two different units were appended. This dual conjugation was accomplished using either a three-step conjugation/maleimide deprotection/conjugation procedure or the two-step conjugation/[maleimide deprotection + conjugation] alternative process.

The methodology here described broadens the scope of possibilities available so far for the ligation of multiple components, the synthesis of complex molecular systems and libraries of differently decorated scaffolds, or to attach different labels or reporter groups to biomolecules (or their ligands). Decoration of a scaffold with different appendices is made possible simply by the introduction of two maleimides, one free and one protected. Moreover, incorporation of other functional groups can also be envisaged to allow for further derivatization,

since the thiol–maleimide and the Diels–Alder reactions can be combined with other conjugation chemistries.

## ASSOCIATED CONTENT

### Supporting Information

General materials and methods; details on the synthesis of the different compounds and mass spectrometric characterization; assessment of the stability of Michael-type adducts and Diels–Alder cycloadducts to maleimide deprotection conditions; HPLC traces of crude maleimido-polyamides and derived conjugates. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*Phone: + 34934021263. Fax: + 34933397878. E-mail: [anna.grandas@ub.edu](mailto:anna.grandas@ub.edu).

### Author Contributions

X.E. and A.S. contributed equally to this work.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work was supported by funds from the Ministerio de Economía y Competitividad (grant CTQ2010-21567-C02-01, and the project RNAREG, grant CSD2009-00080, funded under the programme CONSOLIDER INGENIO 2010), and the Generalitat de Catalunya (2009SGR-208). A. S. and X. E. were recipient fellows of the Generalitat de Catalunya and the MINECO, respectively.

## REFERENCES

- (1) DeForest, C. A., and Anseth, K. S. (2011) Cytocompatible click-based hydrogels with dynamically based tunable properties through orthogonal photoconjugation and photocleavage reactions. *Nat. Chem.* 3, 925–931.
- (2) Xu, J., Filion, T. M., Prifti, F., and Song, J. (2011) Cytocompatible poly(ethylene glycol)-*co*-polycarbonate hydrogels cross-linked by copper-free, strain-promoted click chemistry. *Chem.—Asian J.* 6, 2730–2737.
- (3) van Dijk, M., Rijkers, D. T. S., Liskamp, R. M. J., van Nostrum, C. F., and Hennink, W. E. (2009) Synthesis and applications of biomedical and pharmaceutical polymers via click chemistry methodologies. *Bioconjugate Chem.* 20, 2001–2016.
- (4) Shelke, S. V., Cutting, B., Jiang, X., Koliwer-Brandl, H., Strasser, D. S., Schwarzt, O., Kelm, S., and Ernst, B. (2010) A fragment-based in situ combinatorial approach to identify high-affinity ligands for unknown binding sites. *Angew. Chem., Int. Ed.* 49, 5721–5725.
- (5) Guiard, J., Fiege, B., Kitov, P. I., Peters, T., and Bundle, D. R. (2011) “Double-click” protocol for synthesis of heterobifunctional multivalent ligands: toward a focused library of specific norovirus inhibitors. *Chem.—Eur. J.* 17, 7438–7441.
- (6) Kalia, J., and Raines, R. T. (2010) Advances in bioconjugation. *Curr. Org. Chem.* 14, 138–147.
- (7) Schmidt, B. V. K. J., Fechner, N., Falkenhagen, J., and Lutz, J.-F. (2011) Controlled folding of synthetic polymer chains through the formation of positionable covalent bridges. *Nat. Chem.* 3, 234–238.
- (8) Fahrenbach, A. C., and Stoddart, J. F. (2011) Reactions under the click chemistry philosophy employed in supramolecular and mechanostereochemical systems. *Chem.—Asian J.* 6, 2660–2669.
- (9) Kolb, H. C., Finn, M. G., and Sharpless, K. B. (2001) Click chemistry: diverse chemical function from a few good reactions. *Angew. Chem., Int. Ed.* 40, 2004–2021.



- (10) Kele, P., Mezö, G., Achatz, D., and Wolfbeis, O. S. (2009) Dual labeling of biomolecules by using click chemistry: a sequential approach. *Angew. Chem., Int. Ed.* 48, 344–347.
- (11) Aucagne, V., and Leigh, D. A. (2006) Chemoselective formation of successive triazole linkages in one pot: “click-click” chemistry. *Org. Lett.* 8, 4504–4507.
- (12) Gramlich, P. M. E., Warncke, S., Gierlich, J., and Carell, T. (2008) Click-click-click: single to triple modification of DNA. *Angew. Chem., Int. Ed.* 47, 3442–3444.
- (13) Valverde, I. E., Delmas, A. F., and Aucagne, V. (2009) Click à la carte: robust semi-orthogonal alkyne protecting groups for multiple successive azide/alkyne cycloadditions. *Tetrahedron* 65, 7597–7602.
- (14) Ledin, P. A., Friscourt, F., Guo, J., and Boons, G.-J. (2011) Convergent assembly and surface modification of multifunctional dendrimers by three consecutive click reactions. *Chem.—Eur. J.* 17, 839–846.
- (15) Pourceau, G., Meyer, A., Vasseur, J.-J., and Morvan, F. (2009) Synthesis of mannose and galactose oligonucleotide conjugates by bi-click chemistry. *J. Org. Chem.* 74, 1218–1222.
- (16) Meyer, A., Pourceau, G., Vasseur, J.-J., and Morvan, F. (2009) 5'-Bis-conjugation of oligonucleotides by amidative oxidation and click chemistry. *J. Org. Chem.* 74, 6689–6692.
- (17) Sanders, B. C., Friscourt, F., Ledin, P. A., Mbua, N. E., Arumugam, S., Guo, J., Boltje, T. J., Popik, V. V., and Boons, G.-J. (2011) Metal-free sequential [3 + 2]-dipolar cycloadditions using cyclooctynes and 1,3-dipoles of different reactivity. *J. Am. Chem. Soc.* 133, 949–957.
- (18) Galibert, M., Dumy, P., and Boturyn, D. (2009) One-pot approach to well-defined biomolecular assemblies by orthogonal chemoselective ligations. *Angew. Chem., Int. Ed.* 48, 2576–2579.
- (19) Meyer, A., Spinelli, N., Dumy, P., Vasseur, J.-J., Morvan, F., and Defrancq, E. (2010) Oligonucleotide sequential bis-conjugation via click-oxime and click-Huisgen procedures. *J. Org. Chem.* 75, 3927–3930.
- (20) Galibert, M., Renaudet, O., Dumy, P., and Boturyn, D. (2011) Access to biomolecular assemblies through one-pot triple orthogonal chemoselective ligations. *Angew. Chem., Int. Ed.* 50, 1901–1904.
- (21) Gupta, N., Lin, B. F., Campos, L. M., Dimitriou, M. D., Hikita, S. T., Treat, N. D., Tirrell, M. V., Clegg, D. O., Kramer, E. J., and Hawker, C. J. (2010) A versatile approach to high-throughput microarrays using thiol-ene chemistry. *Nat. Chem.* 2, 138–145.
- (22) Hoyle, C. E., and Bowman, C. N. (2010) Thiol-ene click chemistry. *Angew. Chem., Int. Ed.* 49, 1540–1573.
- (23) Schoch, J., Staudt, M., Samanta, A., Wiessler, M., and Jäschke, A. (2012) Site-specific one-pot dual labeling of DNA by orthogonal cycloaddition chemistry. *Bioconjugate Chem.* 23, 1382–1386.
- (24) Holub, J. M., Jang, H., and Kirshenbaum, K. (2006) Clickity-click: highly functionalized peptoid oligomers generated by sequential conjugation reactions on solid-phase support. *Org. Biomol. Chem.* 4, 1497–1502.
- (25) Pourceau, G., Meyer, A., Chevlot, Y., Souteyrand, E., Vasseur, J.-J., and Morvan, F. (2010) Oligonucleotide carbohydrate-centered galactosyl cluster conjugates synthesized by click and phosphoramidite chemistries. *Bioconjugate Chem.* 21, 1520–1529.
- (26) Sánchez, A., Pedroso, E., and Grandas, A. (2011) Maleimide-dimethylfuran *exo* adducts: effective maleimide protection in the synthesis of oligonucleotide conjugates. *Org. Lett.* 13, 4364–4367.
- (27) Sánchez, A., Pedroso, E., and Grandas, A. (2012) Conjugation reactions involving maleimides and phosphorothioate oligonucleotides. *Bioconjugate Chem.* 23, 300–307.
- (28) Zuckermann, R. N., Kerr, J. M., Kent, S. B. H., and Moos, W. H. (1992) Efficient method for the preparation of peptoids [oligo(N-substituted glycines)] by submonomer solid-phase synthesis. *J. Am. Chem. Soc.* 114, 10646–10647.
- (29) Jia, X., Huang, Q., Li, J., Li, S., and Yang, Q. (2007) Environmentally benign N-Boc protection under solvent- and catalyst-free conditions. *Synlett* 5, 806–808.
- (30) Sánchez, A., Pedroso, E., and Grandas, A. (2013) Oligonucleotide cyclization: the thiol-maleimide reaction revisited. *Chem. Commun.* 49, 309–311.
- (31) Good, L., Awasthi, S. K., Dryselius, R., Larsson, O., and Nielsen, P. E. (2001) Bactericidal antisense effects of peptide-PNA conjugates. *Nat. Biotechnol.* 19, 360–364.

## SUPPORTING INFORMATION TO:

### Protected maleimide building blocks for the decoration of peptides, peptoids and peptide nucleic acids

Xavier Elduque, Albert Sánchez, Kapil Sharma, Enrique Pedroso and Anna Grandas\*

1. General materials and methods.
2. Synthesis of maleimido-propanoic acid (**1**) and diene-dT<sub>5</sub>.
3. Reaction conditions, yield, analysis, purification and characterization data for polyamides incorporating one protected maleimide moiety and derived conjugates.
4. Synthesis of standard samples of **11**, **12** and **13**. Assessment of the stability of Michael-type adducts and Diels-Alder cycloadducts to maleimide deprotection conditions.
5. Reaction conditions, yield, analysis, purification and characterization data for polyamides incorporating two maleimide moieties and derived double conjugates.
6. Abbreviations and references.

## 1. General materials and methods.

Nucleoside phosphoramidites (dT), CPG supports, and oligonucleotide synthesis reagents were from either Link Technologies, Glen Research Corporation or Applied Biosystems. PNA Monomers were from PolyOrg Inc. Fmoc-Ala-OH, Fmoc-Gly-OH, Fmoc-Lys-OH, Fmoc-Lys(Boc)-OH, Fmoc-Ser(<sup>t</sup>Bu)-OH, Fmoc-Tyr(<sup>t</sup>Bu)-OH, HATU, Rink Amide MBHA support and Wang resin were from Novabiochem. 3-Maleimidopropanoic acid was either from Bachem, or synthesized as described below.<sup>[1]</sup> [Protected maleimido]propanoic acid (**1**, *exo* adduct) was prepared as previously described.<sup>[2]</sup> *N*-(2-Fmoc-aminoethyl)glycine *tert*-butyl ester·HCl was from Fluka. Biotin-SH<sup>[3]</sup> and H-Cys-Lys-Glu-Thr-Ala-Ala-Ala-Lys-Phe-Glu-Arg-Gln-His-Met-Asp-Ser-Ser-Thr-Ser-Ala-Ala-OH (Cys-peptide) were prepared following described procedures.<sup>[2]</sup> Bromoacetic acid, DIPC, glutathione, thiocholesterol, tryptamine and the water-soluble carbodiimide (EDC) were from Sigma-Aldrich. HOBt·H<sub>2</sub>O and 2-phenoxyethylamine were purchased to Acros Organics and Alfa Aesar, respectively. Acid-free DCM was obtained by filtration through basic alumina.

TLC was carried out on silica gel plates 60 F<sub>254</sub> from Merck. Microwave irradiation was carried out in a Biotage Initiator<sup>TM</sup> oven, using reaction vessels of 0.2-0.5 mL or 0.5-2 mL as required. Samples were lyophilized in a FreezeMobile Virtis instrument.

<sup>1</sup>H and <sup>13</sup>C NMR spectra were recorded on a Varian Mercury 400 MHz spectrometer, and IR spectra (ATR) on a Nicolet 6700 FT-IR (Thermo Scientific).

The amount of free thiols in thiol-containing compounds was quantified by the Ellman test, as described in the Supporting Information of reference 2.

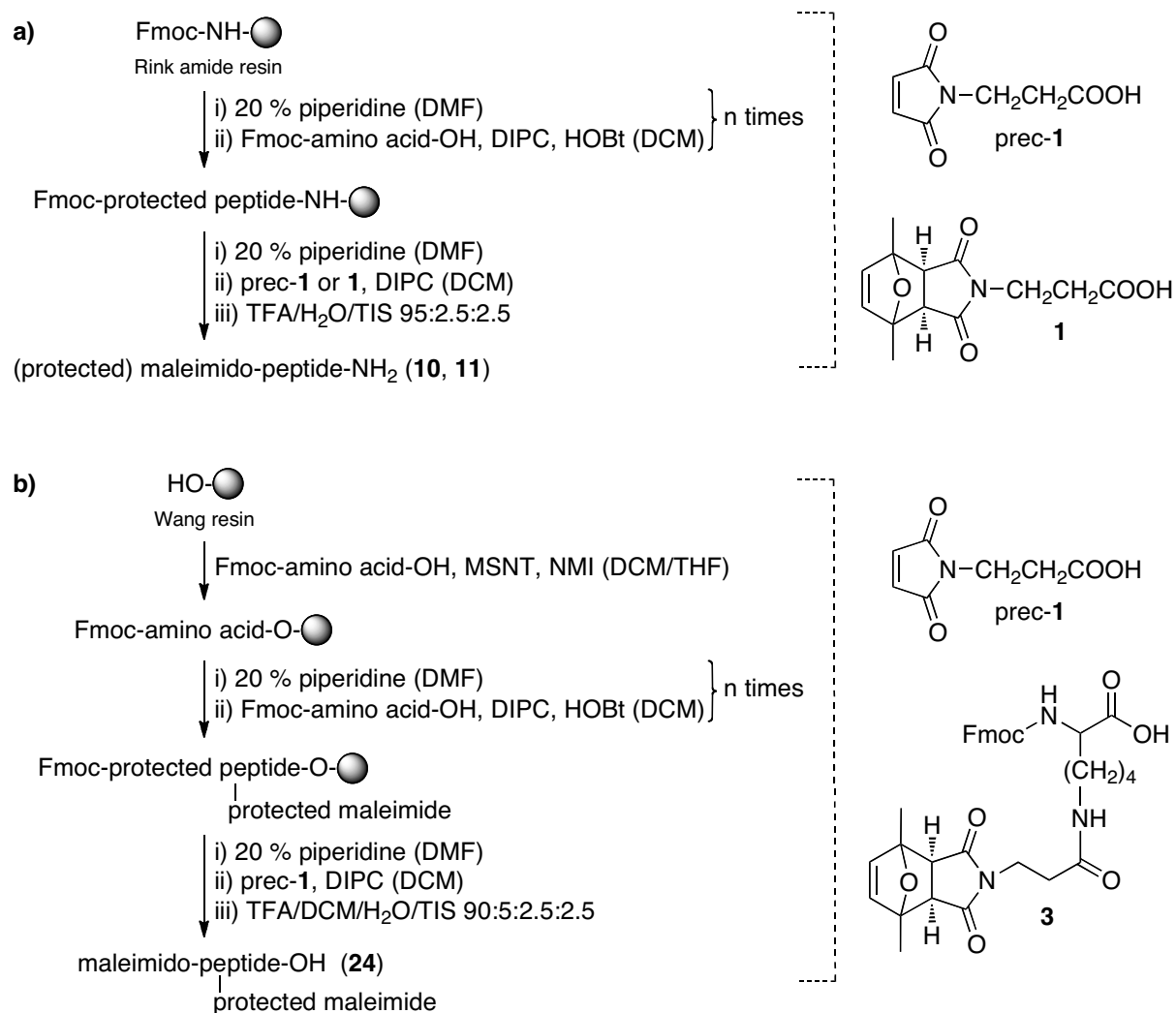
Syntheses of *N*-hydroxysuccinimido ester **8** and PNA monomer **11** were carried out following procedures previously used in the group and adapted from references 4 and 5, respectively.

### Oligonucleotide synthesis.

<sup>5</sup>Diene-oligonucleotides were assembled in a 3400 ABI automatic synthesizer at the 1 μmol scale, using standard phosphite triester approach synthesis cycles, as previously described.<sup>[3,6]</sup> When required, phosphite triesters were sulfurized by reaction with the Beaucage reagent. After chain elongation, treatment with conc. aq. ammonia at room temperature (1 h) removed protecting groups from the oligonucleotide. After filtration and washing, ammonia was evaporated under reduced pressure, and the sample lyophilized. All oligonucleotides were purified by reversed-phase HPLC, quantified by UV spectroscopy (λ = 260 nm), and characterized by mass spectrometry.

**All polyamides** were elongated manually in polyethylene syringes fitted with polypropylene discs using standard procedures (see below).

## Solid-phase peptide synthesis (Fmoc/tBu strategy, Scheme S1).



**Scheme S1.** Main steps in the solid-phase synthesis of peptides. Assembly on the Rink amide resin provides peptides with C-terminal carboxamides (a), while use of the Wang resin affords C-terminal peptide acids (b). The structures of non-standard monomers are shown on the right (maleimido-propanoic acid is the precursor of compound **1**, and is thus designed as prec-1). Full structures of peptides **10**, **11** and **24** are shown on Schemes 3 and 4 of the manuscript.

Peptides with carboxamide groups at the C-terminal were assembled on the Rink Amide resin, which was treated and washed with DCM (3 ×), DMF (3 ×), 20 % piperidine/DMF (2 × 15 min), DCM (3 ×), DMF (3 ×) and MeOH (3 ×) prior to chain elongation. Incorporation of the first amino acid on this resin and peptide elongation were accomplished by using 3 equiv. of both Fmoc-amino acid, HOBT·H<sub>2</sub>O and DIPC (dissolved in the minimum amount of DCM and a few drops of DMF, 90 min), which was followed by washing with DCM, DMF and MeOH (3 ×). In case the coupling was not complete, as assessed by the Kaiser test,<sup>[7]</sup> it was repeated

using 2 equiv. of the reagents. Removal of the Fmoc groups was effected by reaction with 20 % piperidine/DMF (1 × 3 min + 1 × 10 min), followed by washing (DMF, DCM). Activation of 3-maleimidopropanoic acid (prec-1) was carried out with DIPC (90 min reaction time). Even though activation with DIPC and HOBT·H<sub>2</sub>O seemed not to cause any harm on the maleimide, use of only the carbodiimide is safer to prevent addition of HOBT to the maleimide of prec-1).

Deprotection and cleavage from the Rink Amide resin was carried out by reaction with a 95:2.5:2.5 TFA/H<sub>2</sub>O/TIS mixture for 2.5 h at room temperature. Filtrate and washings (DCM) were collected and concentrated under a N<sub>2</sub> stream, and cold diethyl ether was added to the resulting oil. The mixture was centrifuged (10 min, 5 °C, 4800 rpm) and decanted. This procedure was repeated 3 times.

Peptides with C-terminal carboxyl groups were assembled on the Wang resin, which was first washed with DCM (3 ×), 50 % TFA/DCM (2 × 5 min), DCM (3 ×), DMF (3 ×), 20 % piperidine/DMF (2 × 5 min), DCM (3 ×), DMF (3 ×) and MeOH (3 ×). Esterification of the hydroxyl groups of the Wang resin was carried out by reaction with the Fmoc-amino acid (3 equiv), *N*-methylimidazole (3 equiv), and MSNT (4 equiv) in a 7:2 mixture of anh. DCM/THF for 2 h (room temperature), followed by filtration and washing with DCM, DMF and MeOH (3 ×). The degree of substitution was determined on an aliquot of Fmoc-amino acid-resin by quantification of the *N*-(9-fluorenylmethyl)piperidine formed ( $\lambda = 301 \text{ nm}$ ,  $\epsilon = 7800 \text{ M}^{-1}\text{cm}^{-1}$ ) after Fmoc deprotection with 20 % piperidine/DMF. If the substitution degree was suitable, unreacted hydroxyl groups were capped by reaction with a 1.1:1 mixture of Ac<sub>2</sub>O and pyridine (2 × 15 min). Peptide elongation was carried out as described above.

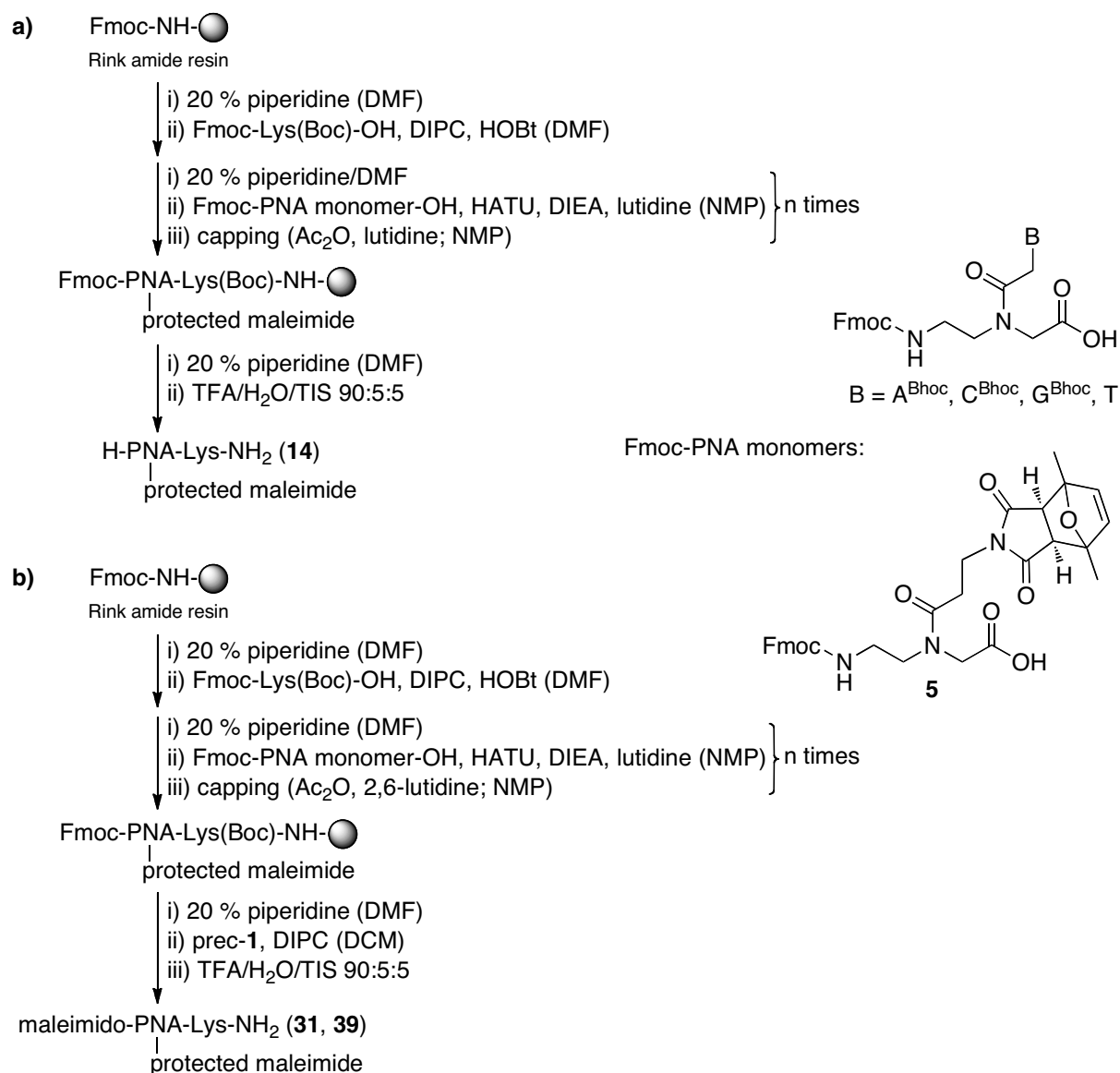
Deprotection and cleavage from the Wang resin was carried out by reaction with a 90:5:2.5:2.5 TFA/DCM/H<sub>2</sub>O/TIS mixture for 2 h at room temperature. Filtrate and washings (DCM) were collected and taken to dryness under reduced pressure.

Peptide crudes were purified by HPLC, quantified by UV spectroscopy ( $\lambda = 280 \text{ nm}$ ; Trp:  $\epsilon_{280} (\text{M}^{-1}\text{cm}^{-1}) = 5540$ , Tyr:  $\epsilon_{280} (\text{M}^{-1}\text{cm}^{-1}) = 1480$ ), and characterized by mass spectrometry.

### **Solid-phase PNA synthesis (Fmoc/Bhoc protocol, Scheme S2).**

PNA Synthesis was carried out on the Rink Amide resin using standard protocols similar to those described for peptide elongation. After the preliminary treatment and washing of the resin, and incorporation of Fmoc-Lys(Boc)-OH as described above, addition of PNA monomers differed from that of amino acids in the washings after removal of the Fmoc group (DMF, DCM and anh. NMP, 3 × 1 mL each), and the coupling conditions: 3 equiv. of PNA-monomer, 2.7 equiv. of HATU, 4 equiv. of DIEA and 4 equiv. 2,6-lutidine were dissolved in 200  $\mu\text{L}$  of anh. NMP under an Ar atmosphere. After 2 min preactivation the mixture was

added to the resin, and the coupling reaction left to proceed for 45 min. Coupling was followed by filtration and washings (DMF, DCM, MeOH, 5 × 1 mL each). To prevent formation of oligomers lacking one residue, capping was carried out after all coupling steps (3 mL of a 5:6:89 Ac<sub>2</sub>O/2,6-lutidine/DMF solution, 2 × 10 min). 3-Maleimidopropanoic acid was coupled as described above.



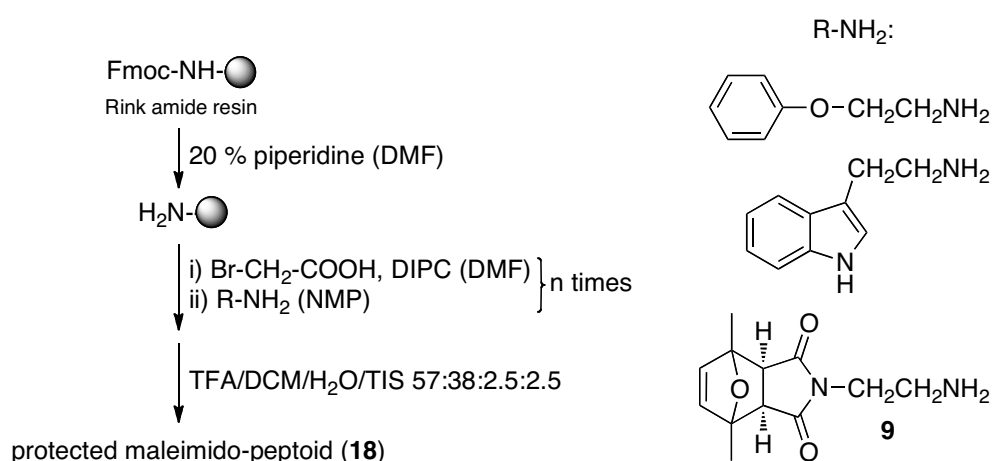
**Scheme S2.** Main steps in the solid-phase synthesis of peptide nucleic acids (PNAs) **14**, **31** and **39**. The structures of the monomers are shown on the right (and that of prec-1 in Scheme S1). Full structures of PNAs **14**, **31** and **39** are shown on Schemes 3 and 4 of the manuscript.

Deprotection and cleavage was carried out by treating the PNA-resin with a 90:5:5 mixture of TFA/H<sub>2</sub>O/TIS (800 μL) for 2 h at room temperature, and subsequent washes with TFA (3 ×

300  $\mu$ L). The combined solutions were concentrated to dryness under a  $N_2$  stream, which was followed by precipitation on cold diethyl ether. After centrifugation the supernatant was discarded and diethyl ether was added again to the precipitate. This step was repeated twice.

Crudes were analyzed and purified by HPLC, quantified by UV spectroscopy (using the same absorption coefficients as for natural nucleosides), and characterized by mass spectrometry.

### Solid-phase peptoid synthesis (submonomer procedure, Scheme S3).<sup>[8]</sup>



**Scheme S3.** Main steps in the solid-phase synthesis of peptoid **18**. The structures of the amines ( $R-NH_2$ ) are shown on the right. Full structure of peptoid **18** is shown on Scheme 3 of the manuscript.

Peptoid **18** was assembled on the Rink Amide resin, which was treated and washed as described above for peptide synthesis. For the acylation step, bromoacetic acid (25 equiv.) was dissolved in anh. DMF in a round-bottom flask, the solution was cooled to 0  $^{\circ}C$ , and DIPC (25 equiv.) was added dropwise. After 5 min, the solution was poured onto the resin in the syringe, and the mixture was stirred (shaker) under an Ar atmosphere for 45 min. The resin was filtered and washed (DCM, DMF, iPrOH and DCM, 3  $\times$  each), and the amination step was carried out by reaction with the corresponding amine (20 equiv.) dissolved in anh. NMP (500  $\mu$ L) for 3 h. After this the resin was washed again (DCM, DMF and MeOH, 3  $\times$  each). Cleavage from the resin was accomplished by reaction with a 57:38:2.5:2.5 mixture of TFA/DCM/ $H_2O$ /TIS for 1 h at room temperature, which was followed by washings with the cleavage mixture and DCM. Solvents and reagents were removed from the filtrate by means of a  $N_2$  stream until the solution became an oil. Cold diethyl ether (10 volumes) was added, and the mixture was centrifuged (10 min, 5  $^{\circ}C$ , 4800 rpm) and decanted. The ether addition plus centrifugation procedure was repeated 3 times. The crude was analyzed and purified by

HPLC, quantified by UV spectroscopy (phenoxy group:  $\epsilon_{270}$  ( $M^{-1}cm^{-1}$ ) = 1450, indole:  $\epsilon_{280}$  ( $M^{-1}cm^{-1}$ ) = 5540), and characterized by mass spectrometry.

**Conjugation procedures** are described in the Experimental Section of the manuscript.

#### **HPLC.**

Reversed-phase HPLC analysis and purification was performed using analytical and semipreparative Waters or Shimadzu systems. Unless otherwise indicated, analysis and purification conditions were:

**Oligonucleotide analysis conditions:** Kromasil C18 column (10  $\mu$ m, 100 Å, 250  $\times$  4.0 mm) from Akzo Nobel; solvent A: 0.05 M triethylammonium acetate, solvent B: H<sub>2</sub>O/ACN 1:1 (v/v), flow: 1 mL/min, detection wavelength: 254 nm.

**Oligonucleotide purification conditions (semipreparative scale):** Jupiter C18 column (10  $\mu$ m, 300 Å, 250  $\times$  10.0 mm) from Phenomenex; solvent A: 0.1 M triethylammonium acetate, solvent B: H<sub>2</sub>O/ACN 1:1 (v/v), unless otherwise indicated flow: 3 mL/min, detection wavelength: 260 nm.

**Polyamide analysis conditions:** Jupiter Proteo column (4  $\mu$ m, 90 Å, 250  $\times$  4.6 mm) from Phenomenex, or Gracesmart C18 column (5  $\mu$ m, 120 Å, 250  $\times$  4.6 mm); solvent A: 0.045 % TFA in water, solvent B: 0.036 % TFA in ACN, flow: 1 mL/min, detection wavelength: 260 (PNA) or 280 (peptide, peptoid with Ar-O groups) nm.

**Polyamide purification conditions (semipreparative scale):** Jupiter Proteo column (10  $\mu$ m, 90 Å, 250  $\times$  10.0 mm) from Phenomenex; solvent A: 0.1 % TFA in water, solvent B: 0.1 % TFA in ACN, flow: 3-4 mL/min, detection wavelength: 260 or 280 nm (depending on the groups present in the oligomer).

**Cholesterol-containing compounds analysis conditions:** Kromasil C4 column (10  $\mu$ m, 100 Å, 250  $\times$  4.6 mm); solvent A: 0.045 % TFA in water, solvent B: 0.036 % TFA in ACN, or solvent A: 0.05 M triethylammonium acetate, solvent B: H<sub>2</sub>O/ACN 1:1 (v/v), flow: 1 mL/min, detection wavelength according to the type of compound.

#### **Mass spectrometry.**

MALDI-TOF mass spectra were recorded on a 4800 *Plus* ABSciex instrument. Unless otherwise indicated, reflector was used. Typical oligonucleotide analysis conditions: 1:1 (v/v) 2,4,6-trihydroxyacetophenone/ammonium citrate (THAP/CA), negative mode; typical polyamide (peptide, peptoid, PNA) analysis conditions: 2,5-dihydroxybenzoic acid (DHB) + 0.1 % TFA, positive mode. ESI (low and high resolution) mass spectra were obtained using an LC/MSD-TOF spectrometer from Agilent Technologies.



## 2. Synthesis of maleimido-propanoic acid (1) and diene-dT<sub>5</sub>.

**3-Maleimidopropanoic acid (prec-1).**<sup>[1]</sup> Maleic anhydride (5.5 g, 56 mmol) and β-alanine (5.0 g, 56 mmol) was dissolved in AcOH (40 mL) and heated at 170 °C for 90 min. The solution was then cooled, and the solvent was removed under reduced pressure. The residue was coevaporated with dioxane to remove acetic acid, and the resulting oil was purified by silica gel column chromatography, eluting with DCM and increasing amounts of MeOH (0 to 15 %). A white solid (4.6 g, 49% yield) was obtained.

R<sub>f</sub> = 0.55 (DCM/MeOH 85:15); <sup>1</sup>H-NMR (400 MHz, CDCl<sub>3</sub>): δ 6.72 (s, 2H), 3.83 (t, *J*=7.19 Hz, 2H), 2.70 (t, *J*=7.20 Hz, 2H) ppm; <sup>13</sup>C-NMR (101 MHz, CDCl<sub>3</sub>): δ 176.4, 170.3, 134.2, 33.2, 32.5 ppm; ESI MS (negative mode): *m/z* 168.3 [M+H]<sup>-</sup>, M calcd. for C<sub>7</sub>H<sub>7</sub>NO<sub>4</sub> 169.0.

**<sup>5</sup>Diene-dT<sub>5</sub>.** After solid-phase assembly and deprotection with ammonia, diene-dT<sub>5</sub> was purified by HPLC at the semipreparative scale (Jupiter C18 column, 4 mL/min, gradient from 5 to 60 % of B in 30 min) and obtained in 60 % overall yield (synthesis and purification). Analytical HPLC (Kromasil C18 column, same gradient): t<sub>R</sub> = 18.4 min; MALDI-TOF MS (negative mode): *m/z* 1617.3 [M-H]<sup>-</sup>, 1639.2 [M-2H+Na]<sup>-</sup>, 1655.2 [M-2H+K]<sup>-</sup>, M calcd. for C<sub>56</sub>H<sub>75</sub>N<sub>10</sub>O<sub>36</sub>P<sub>5</sub> 1618.3.

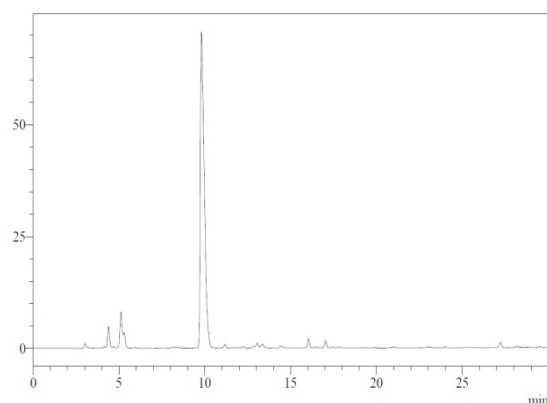
### 3. Reaction conditions, yield, analysis, purification and characterization data for polyamides incorporating one protected maleimide moiety and derived conjugates.

Conditions for maleimide deprotection and conjugation are described in the Experimental section of the manuscript.

#### Peptides and conjugates.

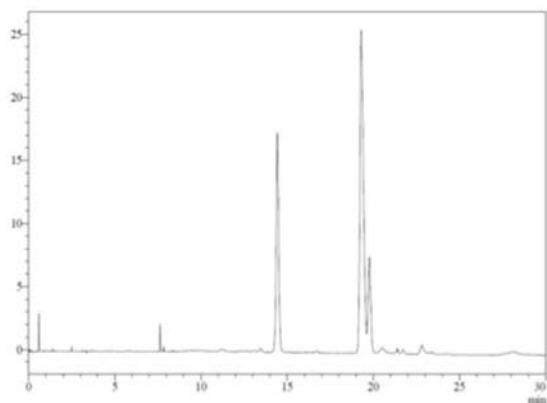
**[Protected maleimido]-Ala-Lys-Tyr-Gly-NH<sub>2</sub> 10.** Peptide **10** (assembled by coupling **1** onto H-Ala-Lys(Boc)-Tyr(*t*-Bu)-Gly-NH-resin) was purified at the semipreparative scale (Jupiter Proteo column, 4 mL/min, linear gradient from 5 to 50 % of B in 30 min) and obtained in 43 % overall yield (synthesis and purification). Analytical HPLC (Jupiter Proteo column, gradient from 5 to 50 % of B in 30 min):  $t_R = 9.8$  min (Figure S1); ESI MS (positive mode):  $m/z$  684.2 [M+H]<sup>+</sup>, 706.2 [M+Na]<sup>+</sup>, 1367.5 [2M+H]<sup>+</sup>, M calcd. for C<sub>33</sub>H<sub>45</sub>N<sub>7</sub>O<sub>9</sub> 683.3.

a)



**Figure S1.** Crude peptide **10** (Scheme 3a).

**Maleimido-peptide 11 from [protected maleimido]-peptide 10.** In addition to the standard sample prepared by incorporating 3-maleimidopropanoic acid at the *N*-terminal of the resin-linked peptide (see section 4), peptide **11** was also obtained by deprotection of the maleimide moiety of peptide **10** (50 nmol), after heating in toluene (20 mL) for 2 h at 90 °C (as explained in the Experimental Section of the manuscript). Deprotection yield (from the HPLC trace of the crude) 28 % (Figure S2). Deprotection yield after 4 h: 50 %. Analytical HPLC (Gracesmart column, gradient from 0 to 20 % of B in 30 min):  $t_R = 14.4$  min; MALDI-TOF MS (positive mode):  $m/z$  588.3 [M+H]<sup>+</sup>, M calcd. for C<sub>27</sub>H<sub>37</sub>N<sub>7</sub>O<sub>8</sub> 583.3.



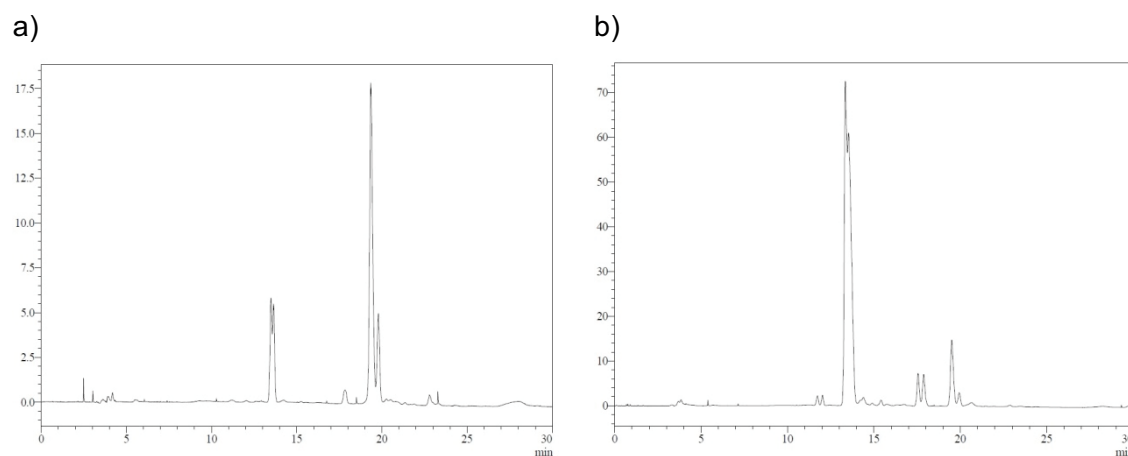
**Figure S2.** Crude obtained after deprotection of the maleimide of peptide **10** by heating in toluene (Scheme 3a, a.1). **11** is the peak with  $t_R \sim 14.5$  min and **10** that with  $t_R \sim 19.5$  min.

### Conjugate **12**.

- Synthesis of **12** by the two-step maleimide deprotection/conjugation process: when deprotection of the maleimide moiety of **10** was followed by reaction with glutathione (**11**/glutathione molar ratio: 1:10, TEAAc buffer, 1 h), virtually all of the peptide **11** available in the deprotection crude was consumed, but the area of the peaks corresponding to the target conjugate in the HPLC of the conjugation crude was 25 % (Figure S3a).

- One-pot deprotection (of **10**) and conjugation reaction (**10**/glutathione 1:10 molar ratio): 3 h heating in toluene at 90 °C: 50 %, an additional 3 h heating raised the yield to 70-75 %; 6 h heating without interruption: 82 % (from the HPLC of the crude, Figure S3b).

Analytical HPLC (Gracesmart column, gradient from 0 to 20 % of B in 30 min):  $t_R = 13.3$  min; MALDI-TOF MS (positive mode):  $m/z$  895.4  $[M+H]^+$ , 917.4  $[M+Na]^+$ , 933.4  $[M+K]^+$ , M calcd. for  $C_{37}H_{54}N_{10}O_{14}$  894.35.



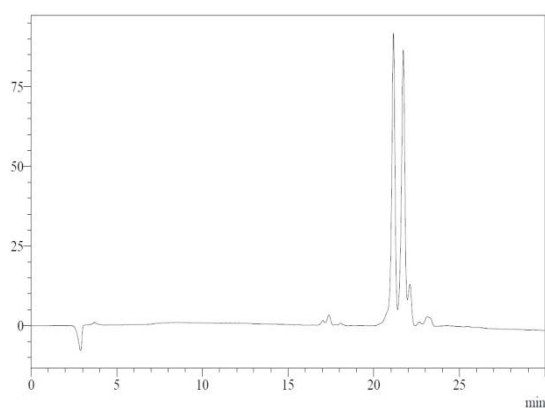
**Figure S3.** a) Crude obtained by reacting **11** (prepared by deprotection of the maleimide of peptide **10**, Scheme 3a, a.1) with glutathione (Scheme 3a, a.2); b) crude obtained after the

one-pot deprotection and conjugation reaction (**10** + glutathione,  $\Delta$  for 6 h; Scheme 3a, a.3). **12** is the peak with  $t_R \sim 13.5$  min.

**Conjugate 13.** One-pot deprotection and conjugation reaction: **10**/diene-dT<sub>5</sub> molar ratio 1:1.2, 6 h, toluene, 90 °C; yield (from the HPLC trace): 93 % (Figure S4).

Analytical HPLC (oligonucleotide analysis conditions, Kromasil C18 column, gradient from 5 to 60 % of B in 30 min):  $t_R = 19.9, 20.4$  min; MALDI-TOF MS (negative mode, both peaks):  $m/z$ : 2205.5 [M-H]<sup>-</sup>, 2227.5 [M-2H+Na]<sup>-</sup>, 2243.5 [M-2H+K]<sup>-</sup>, 2265.3 [M-3H+Na+K]<sup>-</sup>, 2284.1 [M-3H+2K]<sup>-</sup>, 2302.4 [M-4H+Na+2K]<sup>-</sup>, 2318.3 [M-4H+3K]<sup>-</sup>, M calcd. for C<sub>83</sub>H<sub>112</sub>N<sub>17</sub>O<sub>44</sub>P<sub>5</sub> 2205.57.

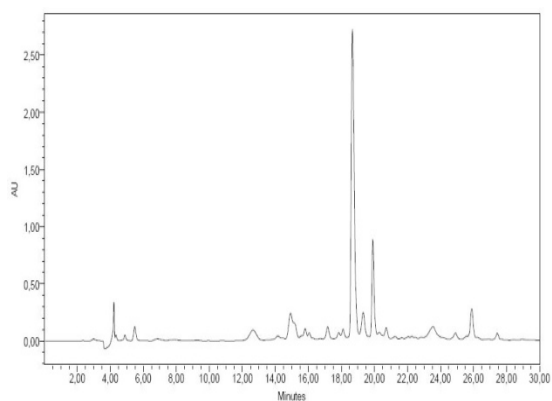
a)



**Figure S4.** Crude **13** obtained after the one-pot deprotection + conjugation procedure (**10** + diene-dT<sub>5</sub> +  $\Delta$ ) (Scheme 3a, a.4).

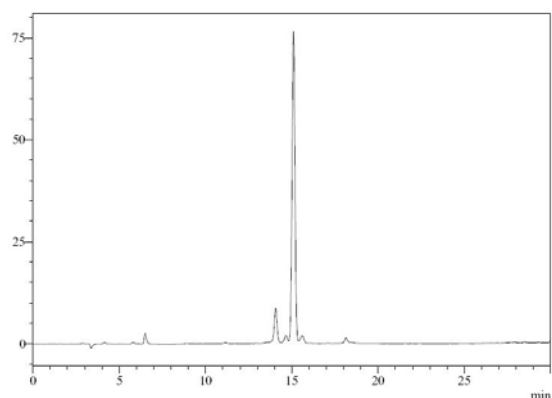
## PNA and conjugates.

**H-t-t-(protected maleimide PNA monomer)-t-t-Lys-NH<sub>2</sub> 14.** PNA **14** was purified at the semipreparative scale (Jupiter Proteo column, 3 mL/min, linear gradient from 10 to 40 % of B in 30 min) and obtained in 64 % overall yield (synthesis and purification). Analytical HPLC (Jupiter Proteo column, gradient from 10 to 40 % of B in 30 min):  $t_R = 18.7$  min (Figure S5); MALDI-TOF MS (positive mode):  $m/z$  1557.9 [M-H]<sup>+</sup>, M calcd. for C<sub>67</sub>H<sub>92</sub>N<sub>22</sub>O<sub>22</sub> 1556.7.



**Figure S5.** Crude PNA **14** (Scheme 3b).

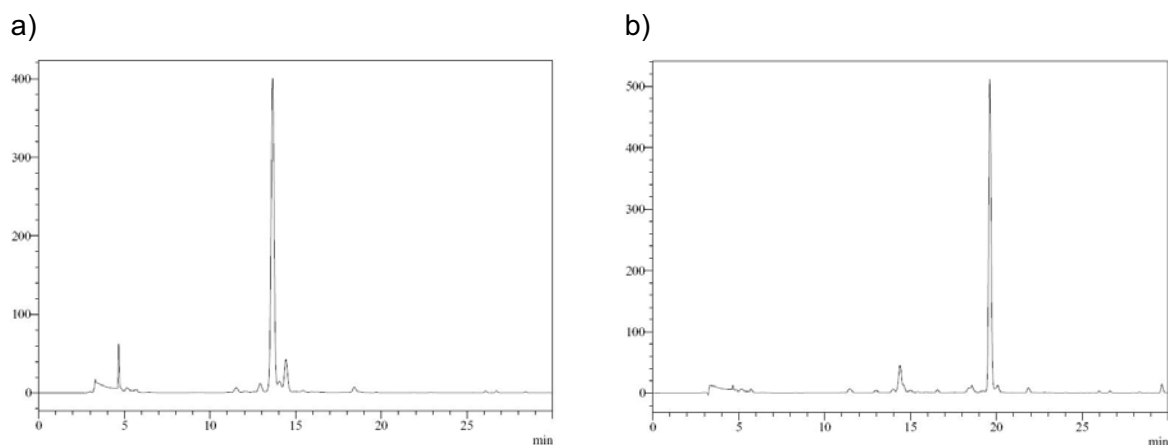
**PNA 15.** Deprotection of PNA **14** was carried out at 100 °C for 1 h (MW), deprotection yield (from the HPLC trace of the crude): 98 % (Figure S6). Analytical HPLC (Jupiter Proteo column, gradient from 5 to 35 % of B in 30 min):  $t_R = 15.2$  min; MALDI-TOF MS (positive mode):  $m/z$  1461.9  $[M+H]^+$ , M calcd. for  $C_{61}H_{84}N_{22}O_{21}$  1460.6.



**Figure S6.** Crude H-t-t-(maleimido PNA monomer)-t-t-Lys-CONH<sub>2</sub> **15** (Scheme 3b, b.1) after microwave-promoted deprotection of **14**.

**Conjugate 16.** Conjugation reaction: **15**/glutathione molar ratio: 1:10, 1 h; conjugation yield (from the HPLC trace of the crude): 90 % (Figure S7a). Analytical HPLC (Jupiter Proteo column, gradient from 5 to 35 % of B in 30 min):  $t_R = 13.7$  min; MALDI-TOF MS (negative mode):  $m/z$  1766.7  $[M-H]^-$ , M calcd. for  $C_{71}H_{101}N_{25}O_{27}S$  1767.7.

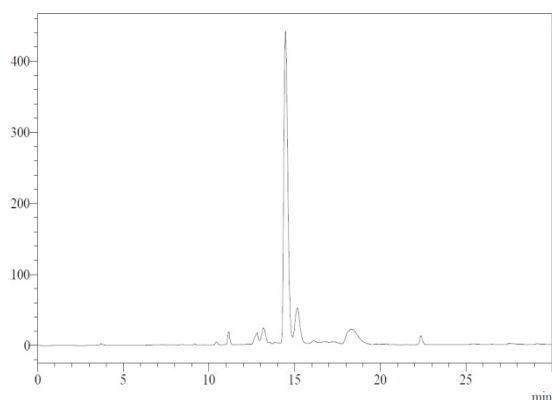
**Conjugate 17.** Conjugation reaction: **15**/biotin-SH molar ratio: 1:10, 1 h; conjugation yield (from the HPLC trace of the crude): 87 % (Figure S7b). Analytical HPLC (Jupiter Proteo column, gradient from 5 to 35 % of B in 30 min):  $t_R = 19.6$  min; MALDI-TOF MS (negative mode):  $m/z$  1762.7  $[M-H]^-$ , M calcd. for  $C_{73}H_{105}N_{25}O_{23}S_2$  1763.7.



**Figure S7.** Crude conjugates **16** (a) and **17** (b) (Scheme 3b, b.2).

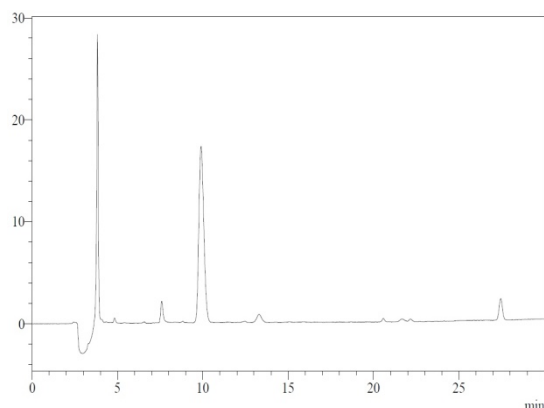
### Peptoids and conjugates.

**[Protected maleimido]-[N-(CH<sub>2</sub>CH<sub>2</sub>Indole)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>OPh)Gly]-NH<sub>2</sub> **18**.** Peptoid **18** was purified at the semipreparative scale (Jupiter Proteo column, 3 mL/min, linear gradient from 5 to 50% of B in 30 min) and obtained in 34 % overall yield (synthesis and purification). Analytical HPLC (Jupiter Proteo column, gradient from 20 to 100 % of B in 30 min):  $t_R$  = 14.4 min (Figure S8); MALDI-TOF MS (positive mode):  $m/z$ : 575.2 [M-furan+H]<sup>+</sup>, 597.2 [M-furan+Na]<sup>+</sup>, 613.2 [M-furan+K]<sup>+</sup>, 671.3 [M+H]<sup>+</sup>, 693.3 [M+Na]<sup>+</sup>, M calcd. for C<sub>36</sub>H<sub>42</sub>N<sub>6</sub>O<sub>7</sub> 670.3.



**Figure S8.** Crude peptoid **18** (main peak). The structure of peptoid **18** is shown in Scheme 3c.

**Maleimido-peptoid 19:** Deprotection: 100 nmol of peptoid **18**, 4 mL toluene, 2.5 h; deprotection yield (from the HPLC trace): 87 % (Figure S9); analytical HPLC (Jupiter Proteo column, gradient from 30 to 60 % of B in 30 min):  $t_R = 9.7$  min; MALDI-TOF MS (positive mode):  $m/z$  1063.6  $[M+H]^+$ , 1085.6  $[M+Na]^+$ , 1101.6  $[M+K]^+$ , M calcd. for  $C_{44}H_{58}N_{10}O_{19}S$  1062.36.



**Figure S9.** Crude peptoid **19** (Scheme 3c, c.1).

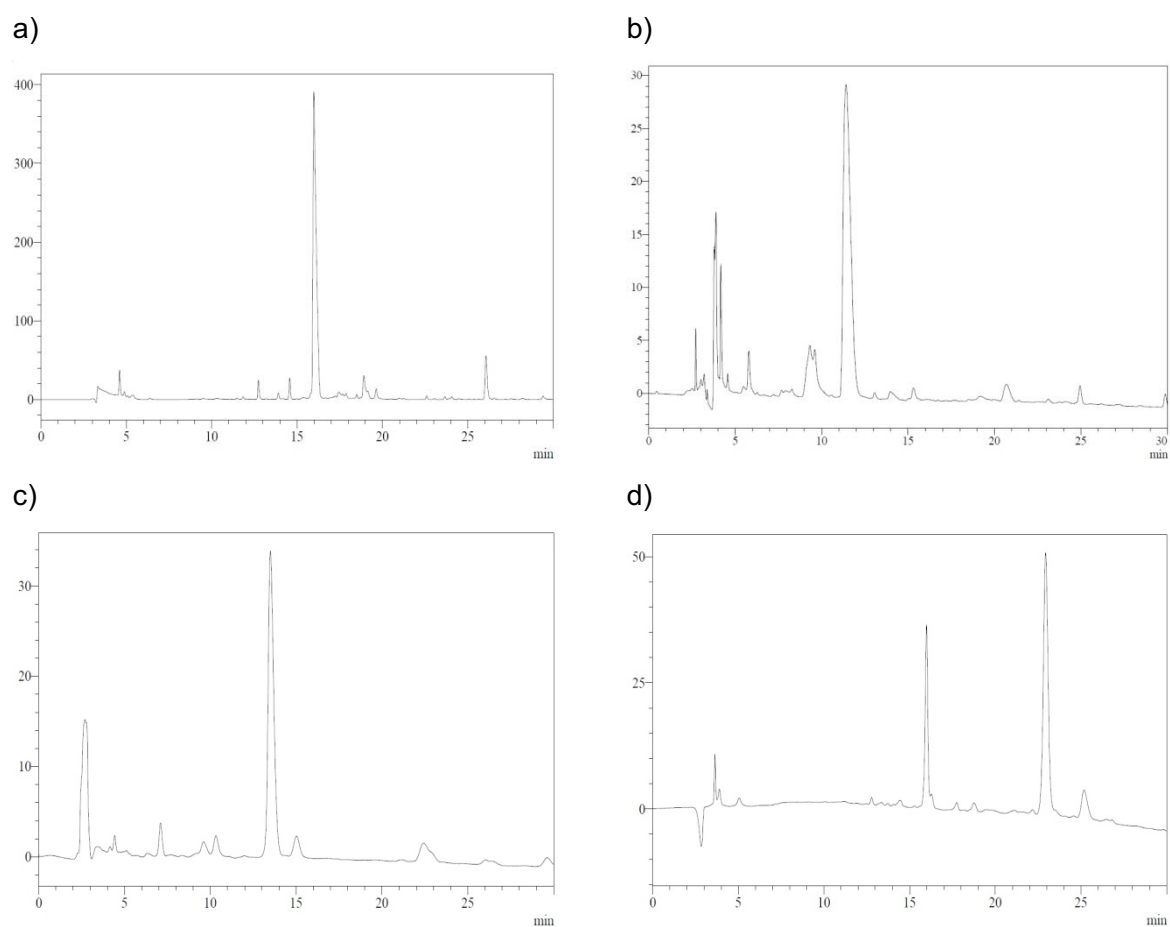
**Conjugate 20.** Conjugation reaction: **19**/glutathione molar ratio: 1:10, 5 min; conjugation yield (from the HPLC of the crude): 81 % (Figure S10a). Analytical HPLC (polyamide analysis conditions, Jupiter Proteo column, gradient from 5 to 100 % of B in 30 min):  $t_R = 16.0$  min; MALDI-TOF MS (positive mode):  $m/z$ : 882.4  $[M+H]^+$ , 904.4  $[M+Na]^+$ , 920.4  $[M+K]^+$ , M calcd. for  $C_{40}H_{51}N_9O_{12}S$  881.3.

**Conjugate 21.** Conjugation reaction: **19**/biotin-SH molar ratio: 1:10, 5 min; conjugation yield (from the HPLC of the crude): 76 % (Figure S10b). Analytical HPLC (polyamide analysis conditions, Jupiter Proteo column, gradient from 30 to 60 % of B in 30 min):  $t_R = 11.8$  min; MALDI-TOF MS (positive mode):  $m/z$  878.5  $[M+H]^+$ , M calcd. for  $C_{42}H_{55}N_9O_8S_2$  877.4.

**Conjugate 22.** Conjugation reaction: **19**/thiocholesterol molar ratio: 1:5, THF/ACN/TEAAc buffer 60:10:30, 1 h; conjugation yield (from the HPLC of the crude): 75 % (Figure S10c). Analytical HPLC (polyamide analysis conditions, Kromasil C4, gradient from 80 to 100 % of B in 30 min):  $t_R = 13.6$  min; MALDI-TOF MS (positive mode):  $m/z$  977.6  $[M+H]^+$ ; 999.6  $[M+Na]^+$ ; 1015.6  $[M+K]^+$ ; calcd. for  $C_{57}H_{80}N_6O_6S$  976.6.

**Conjugate 23.** Conjugation reaction: **19**/diene-dT<sub>5</sub> molar ratio: 1:1.4, 3 h, toluene, 90 °C; conjugation yield (from the HPLC of the crude): 89 % (Figure S10d). After 6 days in water at 37 °C the conjugation yield had been 23 % (HPLC profile not shown). Analytical HPLC (oligonucleotide analysis conditions, Kromasil C18, gradient from 5 to 100 % of B in 30 min):

$t_R = 22.9$  min; MALDI-TOF MS (negative mode):  $m/z$ : 2194.7  $[M+H]^+$ , M calcd. for  $C_{86}H_{109}N_{16}O_{42}P_5$  2192.6.

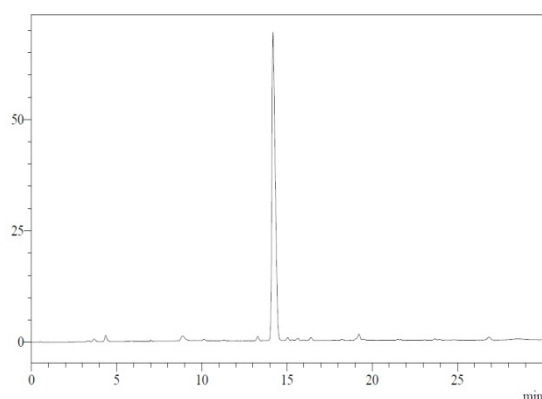


**Figure S10.** Crude conjugates (Scheme 3c, c.2) **20** (profile a), **21** (profile b), **22** (profile c) and **23** (profile d). The target conjugate is the main peak in all cases. Profile d: the peak with  $t_R \sim 16$  min corresponds to excess diene-dT<sub>5</sub>.



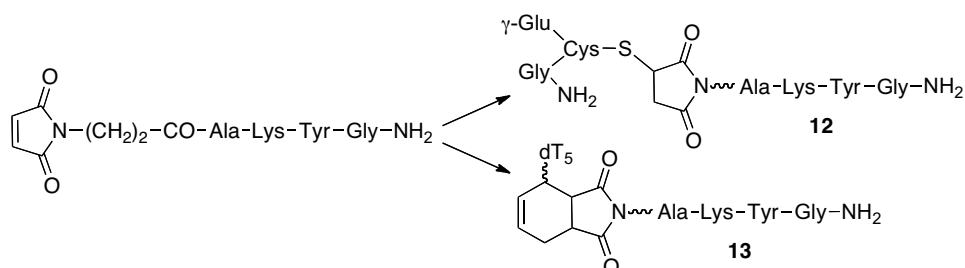
#### 4. Synthesis of standard samples of 11, 12 and 13. Assessment of the stability of Michael-type adducts and Diels-Alder cycloadducts to maleimide deprotection conditions.

**Maleimido-Ala-Lys-Tyr-Gly-NH<sub>2</sub> 11.** To obtain standard samples of conjugates **12** and **13** (see section 4A), peptide **11** was prepared by coupling 3-maleimidopropanoic acid (prec-1) to H-Ala-Lys(Boc)-Tyr(*t*-Bu)-Gly-NH-resin. **11** Was isolated in 65 % overall synthesis and purification yield after the TFA treatment and purification by semipreparative HPLC (Jupiter Proteo, 4 mL/min, linear gradient from 5 to 50 % of B in 30 min). Analytical HPLC (Gracesmart column, gradient from 0 to 20 % of B in 30 min):  $t_R = 14.3$  min (Figure S11); ESI MS (positive mode):  $m/z$  588.1 [M+H]<sup>+</sup>, 1175.3 [2M+H]<sup>+</sup>, M calcd. for C<sub>27</sub>H<sub>37</sub>N<sub>7</sub>O<sub>8</sub> 587.3.



**Figure S11.** Crude standard peptide **11**.

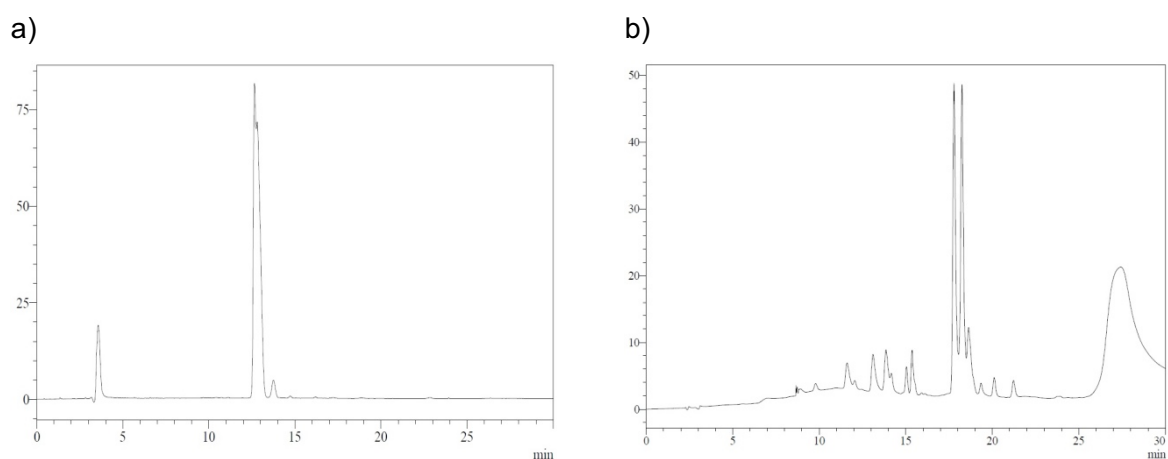
**Standard samples of conjugates 12 and 13.** These samples were prepared by reacting maleimido-Ala-Lys-Tyr-Gly-NH<sub>2</sub> (sample obtained as described above) with glutathione and 5' diene-dT<sub>5</sub>, respectively:



**Scheme S4.** Preparation of standard samples of conjugates **12** and **13** from the standard sample of peptide **11** (obtained by coupling maleimido-propanoic acid to the resin-linked protected peptide).

**Conjugate 12.** Conjugation reaction to synthesize the standard sample of **12: 11** (standard sample)/glutathione molar ratio: 1:10, TEAAc buffer, overnight; conjugation yield (from the HPLC of the crude): 85 % (Figure S12a).

**Conjugate 13.** Conjugation reaction to synthesize the standard sample of **13: 11** (standard sample)/diene-dT<sub>5</sub> molar ratio: 1:1, overnight; conjugation yield (from the HPLC of the crude): 64 % (Figure S12b).



**Figure S12.** a) Crude **12** obtained by reacting standard peptide **11** with glutathione; b) crude **13** obtained by reacting standard peptide **11** with diene-dT<sub>5</sub>.

#### **Stability to maleimide deprotection conditions.**

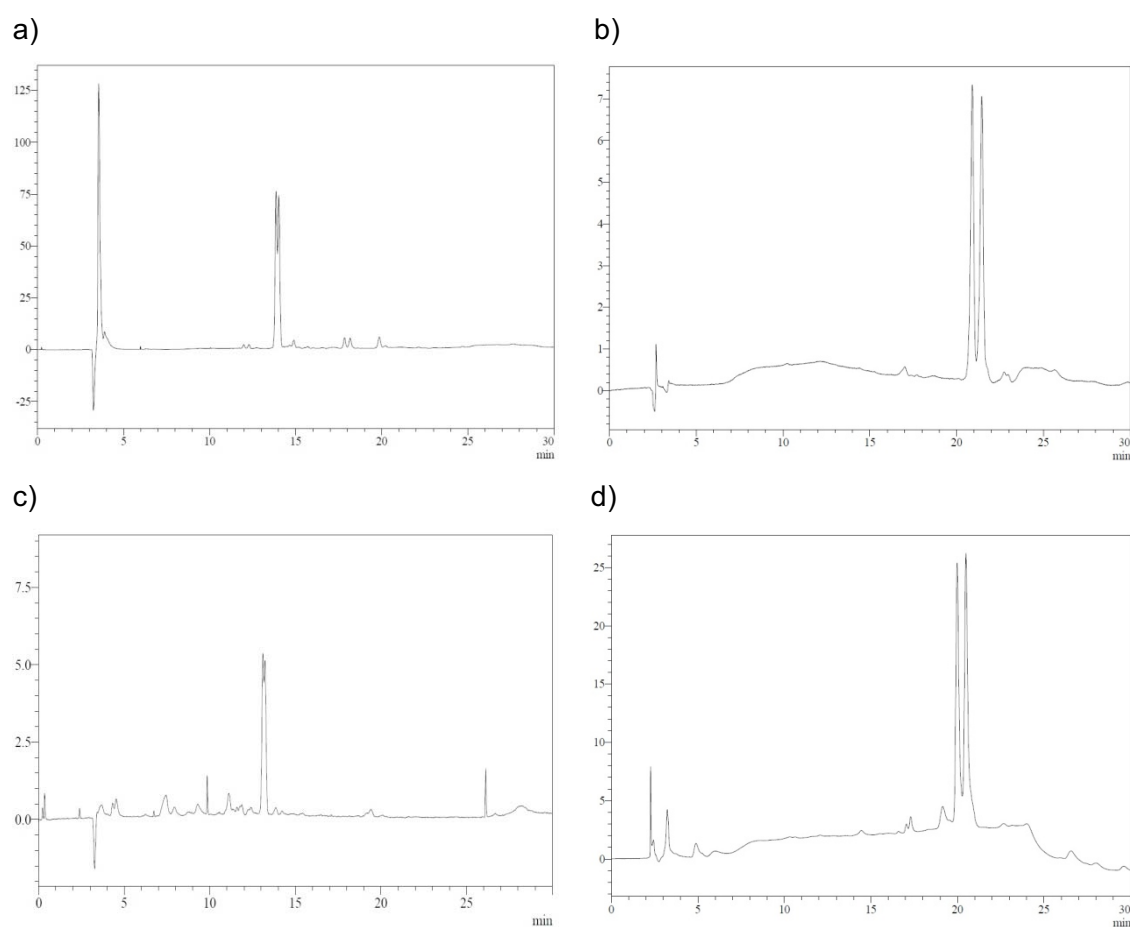
**Stability to heating in toluene.** In the first set of experiments, each of the conjugates (25 nmol) was dissolved in anh. toluene (1 mL), and the solution was heated at 90 °C for 6 h. Solvent was removed under reduced pressure, and the resulting residues were dissolved in 50 µL of water. HPLC analysis (Figure S4,a,b) showed that both **12** and **13** were completely stable.

**Stability to heating in a microwave oven.** Alternatively, each of the standard conjugates (**12** and **13**, 25 nmol) was dissolved in a 1:1 MeOH/H<sub>2</sub>O mixture (1 mL), and heated in a microwave oven at 90 °C for 90 min. Methanol (and some water) was removed under reduced pressure, and the resulting aqueous solution was analyzed by HPLC (Figure S4,c,d). Again, the two conjugates remained undegraded.

HPLC data for these experiments (differences in retention times are due to analyses being carried out in different days):

**Conjugate 12:** Analytical HPLC (Gracesmart column, gradient from 0 to 20 % of B in 30 min):  $t_R = 13.9$  min (after heating in toluene), 13.2 (after heating in the microwave oven).

**Conjugate 13:** Analytical HPLC (Kromasil C18 column, gradient from 5 to 60 % of B in 30 min):  $t_R = 20.9$  and 21.4 min after heating in toluene;  $t_R = 20.0$  and 20.5 min after heating in the microwave oven.

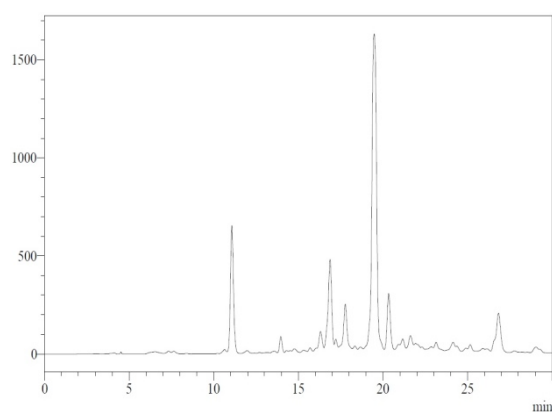


**Figure S13.** HPLC profiles of conjugates **12** and **13** after heating in toluene (a and b, respectively) and after heating in the microwave oven (c and d, respectively). HPLC profiles were obtained from analyses carried out in different days, and small differences in retention times were found. In all cases compounds were collected and identified as the corresponding conjugates by MALDI-TOF mass spectrometric analysis.

## 5. Reaction conditions, yield, analysis and characterization data for polyamides incorporating two maleimide moieties and derived double conjugates.

### Peptide and conjugates.

**Maleimido-Ser-Lys(protected maleimide)-Tyr-Gly-OH 24.** Peptide **24** was purified at the semipreparative scale (Jupiter Proteo column, 4 mL/min, linear gradient from 5 to 50 of B in 30 min) and obtained in 7 % overall yield (synthesis and purification). Analytical HPLC (Gracesmart column, gradient from 5 to 50 % of B in 30 min):  $t_R = 19.5$  min (Figure S14); ESI MS (positive mode):  $m/z$  852.3  $[M+H]^+$ , 874.3  $[M+Na]^+$ , 890.3  $[M+K]^+$ , M calcd. for  $C_{40}H_{49}N_7O_{14}$  851.3.



**Figure S14.** Crude peptide **24** (c) (the target peptide is the main peak). The structure of peptide **24** is shown in Scheme 4a.

**Conjugate 25.** Conjugation reaction: **24**/glutathione molar ratio: 1:10, overnight; conjugation yield (from the HPLC of the crude): 69 % (Figure S15a). Analytical HPLC (polyamide analysis conditions, Gracesmart column, gradient from 5 to 50 % of B in 30 min):  $t_R = 12.6$  min; MALDI-TOF MS (negative mode):  $m/z$  1061.4  $[M-furan-H]^-$ , 1157.5  $[M-H]^-$ , 1229.4  $[M-furan+THAP-H]^-$ ; M calcd. for  $C_{50}H_{66}N_{10}O_{20}S$  1158.42.

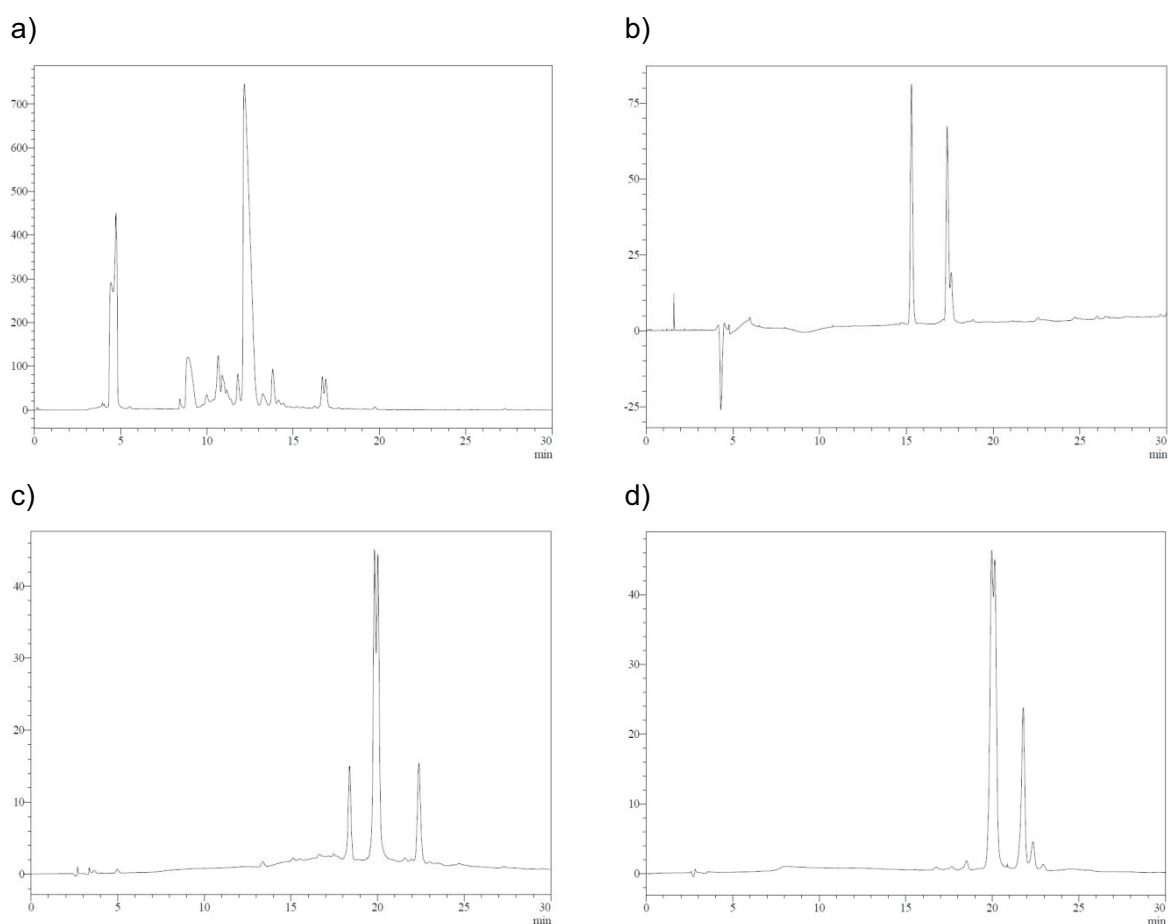
**Conjugate 26:** Deprotection: 100 nmol of conjugate **25**, 4 mL 1:2 (v/v) hexanes/toluene, 4 h; deprotection yield (from the HPLC trace): 50 % (Figure S15b). Analytical HPLC (polyamide analysis conditions, Gracesmart column, gradient from 0 to 50 % of B in 30 min):  $t_R = 15.3$  min; MALDI-TOF MS (positive mode):  $m/z$  1063.6  $[M+H]^+$ , 1085.6  $[M+Na]^+$ , 1101.6  $[M+K]^+$ , M calcd. for  $C_{44}H_{58}N_{10}O_{19}S$  1062.36.

## Conjugate 27.

· Conjugation reaction: **26**/diene-dT<sub>5</sub> molar ratio 5:1 (crude **26**, which still contained protected maleimide as described above, was not very soluble in water), 6 days; conjugation yield (from the HPLC trace of the crude): 74 % (Figure S15c).

· One-pot deprotection and conjugation: **25**/diene-dT<sub>5</sub> molar ratio 1:1.3, 6 h, 90 °C; yield (from the HPLC trace): 95 % (Figure S15d).

Analytical HPLC (oligonucleotide analysis conditions, Kromasil C18 column, gradient from 5 to 60 % of B in 30 min): t<sub>R</sub> = 19.9 min; MALDI-TOF MS (negative mode): m/z 2071.6 [M-2dT-H]<sup>-</sup>, 2376.6 [M-Glut-H]<sup>-</sup>, 2551.7 [M-glutamine-H]<sup>-</sup>, 2679.7 [M-H]<sup>-</sup>, M calcd. for C<sub>100</sub>H<sub>133</sub>N<sub>20</sub>O<sub>55</sub>P<sub>5</sub>S 2680.66.



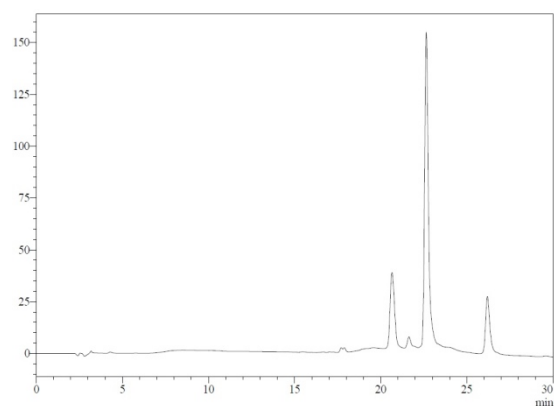
**Figure S15.** a) Crude conjugate **25** (**25** is the main peak; Scheme 4a, 1.i & Scheme 4a, 3.i); b) crude conjugate **26** (Scheme 4a, 1.ii), where the first main peak is **26** (t<sub>R</sub> ~ 15 min), and the second (t<sub>R</sub> ~ 18 min) is **25** (incomplete maleimide deprotection). c) And d): HPLC profiles of conjugation crude **27** obtained after the two-step process (deprotection of **25** followed by conjugation, profile c, Scheme 4a, 1.iii), and after the one-pot deprotection and conjugation procedure (profile d, Scheme 4a, 3.ii). Profiles c and d: The product eluting at ~ 20 min is conjugate **27**, and that with t<sub>R</sub> ~ 22 min is diene-dT<sub>5</sub>.

**Conjugate 28.** Conjugation reaction: **24**/diene-dT<sub>5</sub> molar ratio: 5:1, overnight; conjugation yield (from the HPLC of the crude): 68 % (Figure S16a). Analytical HPLC (oligonucleotide analysis conditions, Kromasil C18 column, gradient from 5 to 60 % of B in 30 min): t<sub>R</sub> = 22.6 min; MALDI-TOF MS (positive mode): m/z: 2372.4 [M-furan+H]<sup>+</sup>, 2469.4 [M+H]<sup>+</sup>, M calcd. for C<sub>96</sub>H<sub>124</sub>N<sub>17</sub>O<sub>50</sub>P<sub>5</sub> 2469.64.

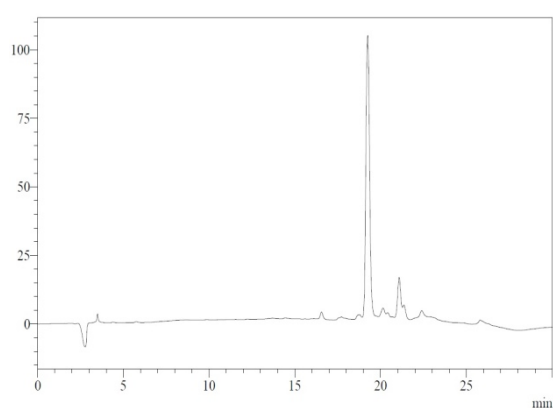
**Conjugate 29:** Deprotection: 24 nmol of conjugate **28**, 1 mL toluene, 2.5 h; deprotection yield (from the HPLC trace): 87 % (Figure S16b). Analytical HPLC (oligonucleotide analysis conditions, Kromasil C18 column, gradient from 5 to 60 % of B in 30 min): t<sub>R</sub> = 19.2 min; MALDI-TOF MS (negative mode): m/z 2372.4 [M-H]<sup>-</sup>, M calcd. for C<sub>90</sub>H<sub>116</sub>N<sub>17</sub>O<sub>49</sub>P<sub>5</sub> 2373.58.

**Conjugate 30.** Conjugation reaction: **29**/glutathione molar ratio: 1:10, overnight; conjugation yield (from the HPLC of the crude): 66 % (Figure S16c). Analytical HPLC (oligonucleotide analysis conditions, Kromasil C18 column, gradient from 5 to 60 % of B in 30 min): t<sub>R</sub> = 19.1 min; MALDI-TOF MS (positive mode): m/z: 2439.1 [M-T+H]<sup>+</sup>, 2682.1 [M+H]<sup>+</sup>, M calcd. for C<sub>100</sub>H<sub>133</sub>N<sub>20</sub>O<sub>55</sub>P<sub>5</sub>S 2680.66.

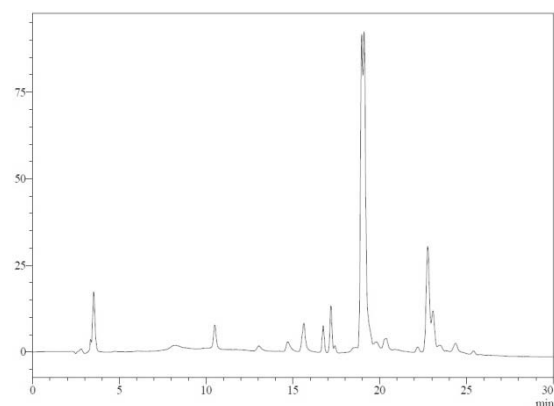
a)



b)



c)

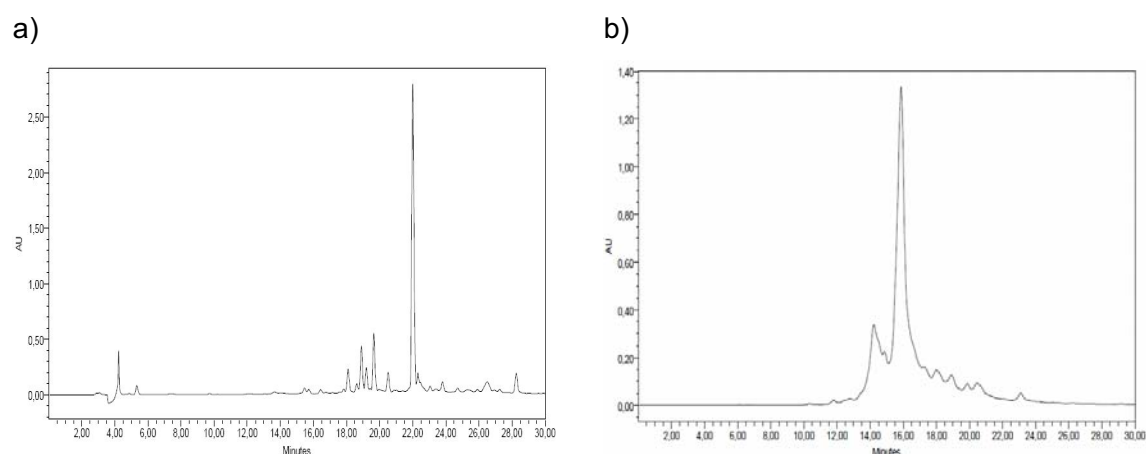


**Figure S16.** a) Crude conjugate **28** (Scheme 4a, 2.i). The main peak is **28** ( $t_R \sim 23$  min), and the first (here,  $t_R \sim 20$ -21 min) is diene-dT<sub>5</sub>. b) And c): Crude conjugates **29** (profile b, Scheme 4a, 2.ii) and **30** (profile c, Scheme 4a, 2.iii) (**29** and **30** are the main peak in both cases).

## PNAs and conjugates.

**Maleimido-t-t-(protected maleimido PNA monomer)-t-t-Lys-NH<sub>2</sub> 31.** PNA **31** was purified at the semipreparative scale (Jupiter Proteo column, 3 mL/min, linear gradient from 10 to 40 % of B in 30 min) and obtained in 57 % overall yield (synthesis and purification). Analytical HPLC (Jupiter Proteo column, gradient from 10 to 40 % of B in 30 min):  $t_R = 22.0$  min (Figure S17a); MALDI-TOF MS (positive mode):  $m/z$  1708.9  $[M+H]^+$ ,  $M$  calcd. for C<sub>74</sub>H<sub>97</sub>N<sub>23</sub>O<sub>25</sub> 1707.7.

**Maleimido-catagctgtttc-(protected maleimido PNA monomer)-Lys-NH<sub>2</sub> 39.** PNA **39** was analyzed and purified at the semipreparative scale (Jupiter Proteo column, 3 mL/min, linear gradient from 10 to 40 % of B in 30 min):  $t_R = 15.9$  min (Figure S17b) and obtained in 43 % overall yield (synthesis and purification); MALDI-TOF MS (DHB matrix, positive mode):  $m/z$  3788.0  $[(M\text{-dimethylfuran})+Na]^+$ ,  $m/z$  3804.0  $[(M\text{-dimethylfuran})+K]^+$ ,  $M$  calcd. for C<sub>159</sub>H<sub>201</sub>N<sub>69</sub>O<sub>49</sub> 3860.5; (M-dimethylfuran) calcd. for C<sub>153</sub>H<sub>193</sub>N<sub>69</sub>O<sub>48</sub> 3764.5.



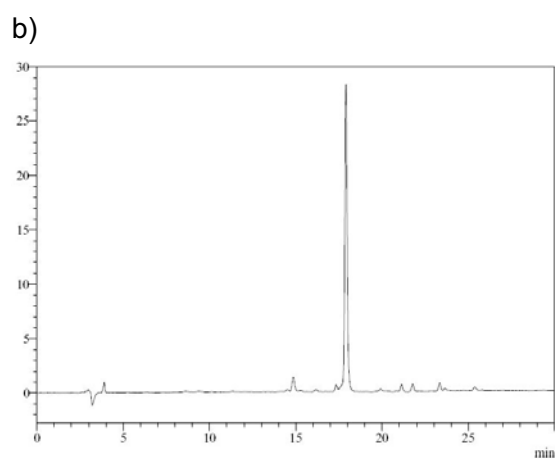
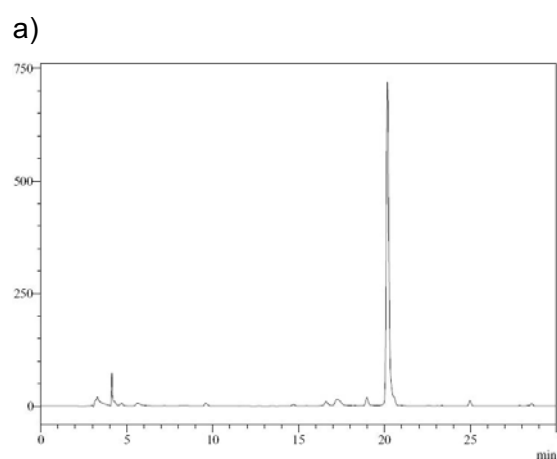
**Figure S17.** **31** (a) and **39** (b) (the target PNAs are the main peaks of each trace). The structures of PNAs **31** and **39** are shown in Schemes 4b and 4c, respectively.

**Conjugate 32.** Conjugation reaction: **31**/biotin-SH molar ratio: 1:10, 1 h; conjugation yield (from the HPLC trace of the crude): quantitative (Figure S18a). Analytical HPLC (polyamide analysis conditions, Jupiter Proteo column, gradient from 10 to 40 % of B in 30 min):  $t_R = 20.1$  min; MALDI-TOF MS (negative mode):  $m/z$  2009.7  $[M-H]^-$ , M calcd. for  $C_{86}H_{118}N_{26}O_{27}S_2$  2010.8.

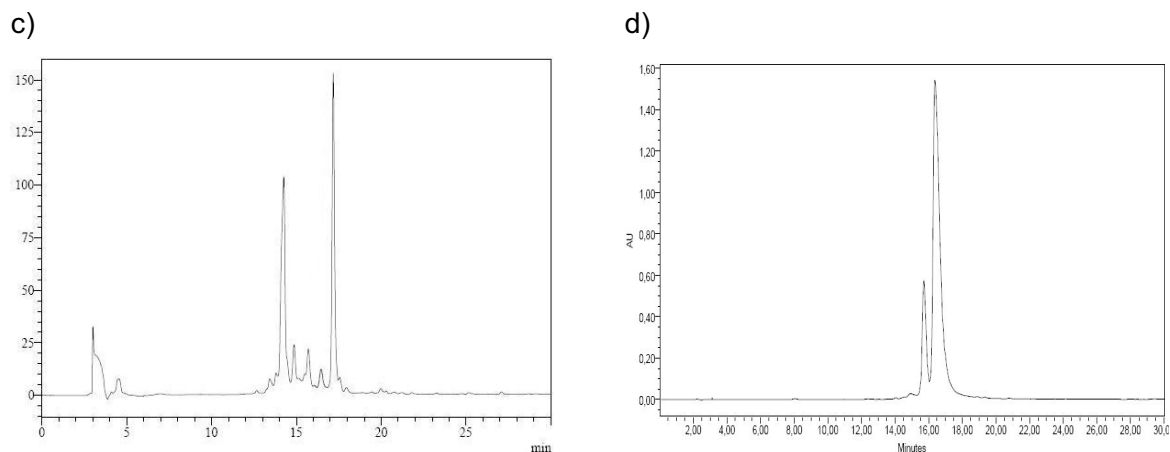
**Conjugate 33.** Deprotection of conjugate **32** was carried out by MW irradiation at 100 °C for 1 h, deprotection yield (from the HPLC trace): quantitative (Figure S18b); analytical HPLC (polyamide analysis conditions, Jupiter Proteo column, gradient from 10 to 40 % of B in 30 min):  $t_R = 17.9$  min; MALDI-TOF MS (negative mode):  $m/z$  1913.7  $[M-H]^-$ , M calcd. for  $C_{80}H_{110}N_{26}O_{26}S_2$  1914.8.

**Conjugate 34.** Conjugation reaction: **33**/Cys-peptide molar ratio: 1:5, overnight, conjugation yield (from the HPLC trace of the crude): quantitative (Figure S18c). Analytical HPLC (polyamide analysis conditions, Jupiter Proteo column, gradient from 10 to 40 % of B in 30 min):  $t_R = 17.3$  min; MALDI-TOF MS (negative mode):  $m/z$  4181.8  $[M-H]^-$ , M calcd. for  $C_{172}H_{259}N_{55}O_{60}S_4$  4182.8.

**Conjugate 35.** Conjugation reaction: **33**/diene-dT<sub>10</sub>-PS molar ratio: 1:1.2, overnight, conjugation yield (from the HPLC trace of the crude): quantitative (Figure S18d). Analytical HPLC (oligonucleotide analysis conditions, Kromasil C18 column, gradient from 10 to 80 % of B in 30 min):  $t_R = 16.4$  min; MALDI-TOF MS (negative mode):  $m/z$  5212.1  $[M-H]^-$ , M calcd. for  $C_{186}H_{250}N_{46}O_{87}P_{10}S_{12}$  5213.1.





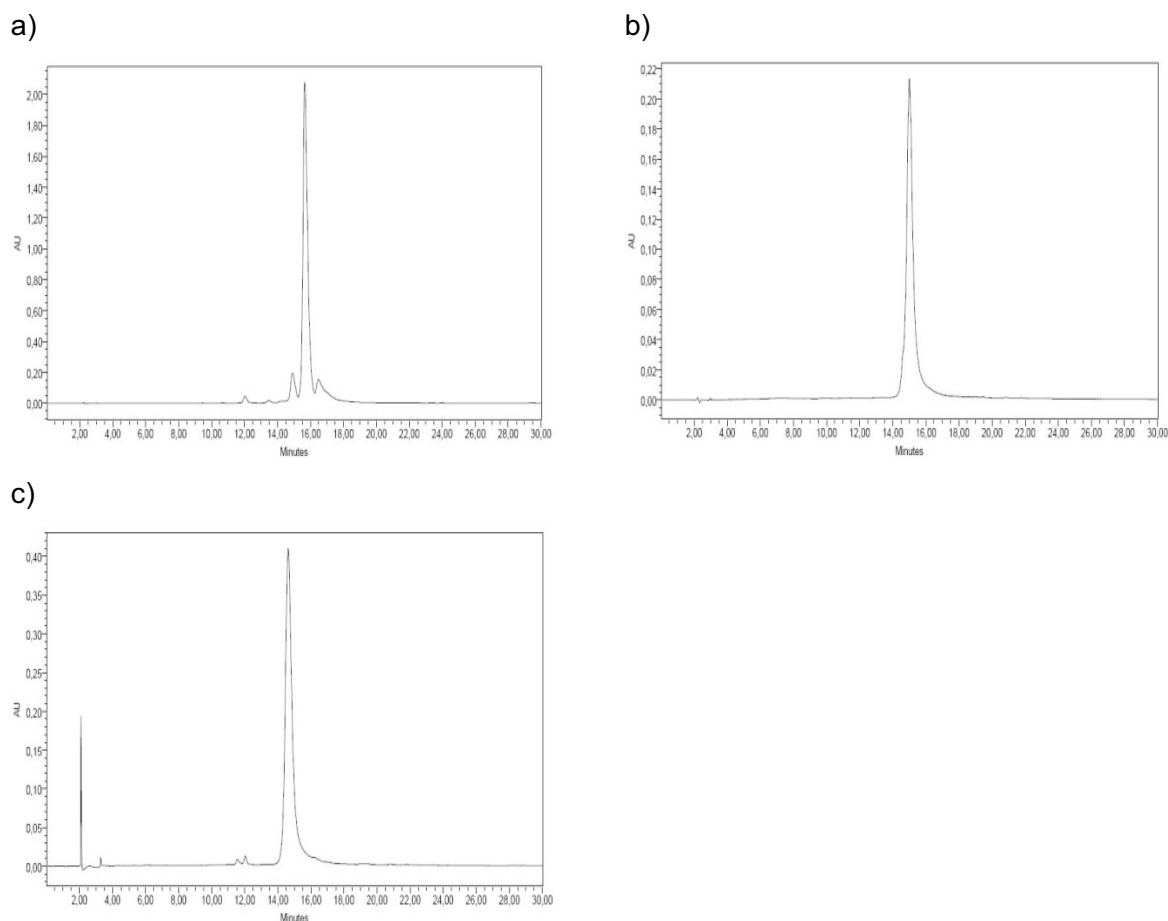


**Figure S18.** a) Crude conjugate **32** (Scheme 4b, 1.i & 2.i); b) crude conjugate **33** after microwave-promoted deprotection (Scheme 4b, 1.ii & 2.ii); c) crude conjugate **34** (**34** is the main peak, and the peak at 14.3 min corresponds to excess Cys-peptide; Scheme 4b, 1.iii); d) crude conjugate **35** (**35** is the main peak, and the peak with  $t_R \sim 16$  min corresponds to excess  $5'$  diene-dT<sub>10</sub>-PS; Scheme 4b, 2.iii).

**Conjugate 36.** Conjugation reaction: **31**/diene-dT<sub>10</sub>-PS molar ratio: 1:1, overnight, conjugation yield (from the HPLC trace of the crude): 95 % (Figure S19a). Analytical HPLC (oligonucleotide analysis conditions, Kromasil C18 column, gradient from 10 to 80 % of B in 30 min):  $t_R = 15.7$  min; MALDI-TOF MS (negative mode):  $m/z$  5004.7 [M-H]<sup>-</sup>, M calcd. for C<sub>180</sub>H<sub>237</sub>N<sub>43</sub>O<sub>86</sub>P<sub>10</sub>S<sub>10</sub> 5006.0.

**Conjugate 37.** Deprotection of conjugate **36** was carried out by MW irradiation at 100 °C for 1 h, deprotection yield (from the HPLC trace): quantitative (Figure S19b); analytical HPLC (oligonucleotide analysis conditions, Kromasil C18 column, gradient from 10 to 80 % of B in 30 min):  $t_R = 15.0$  min; MALDI-TOF MS (negative mode):  $m/z$  4908.6 [M-H]<sup>-</sup>, M calcd. for C<sub>174</sub>H<sub>229</sub>N<sub>43</sub>O<sub>85</sub>P<sub>10</sub>S<sub>10</sub> 4910.0.

**Conjugate 38.** Conjugation reaction: **37**/Cys-peptide molar ratio: 1:5, overnight, conjugation yield (from the HPLC trace of the crude): quantitative (Figure S19c). Analytical HPLC (oligonucleotide analysis conditions, Kromasil C18 column, gradient from 10 to 80 % of B in 30 min):  $t_R = 14.6$  min; MALDI-TOF MS (negative mode):  $m/z$  7177.0 [M-H]<sup>-</sup>, M calcd. for C<sub>266</sub>H<sub>378</sub>N<sub>72</sub>O<sub>119</sub>P<sub>10</sub>S<sub>12</sub> 7178.0.



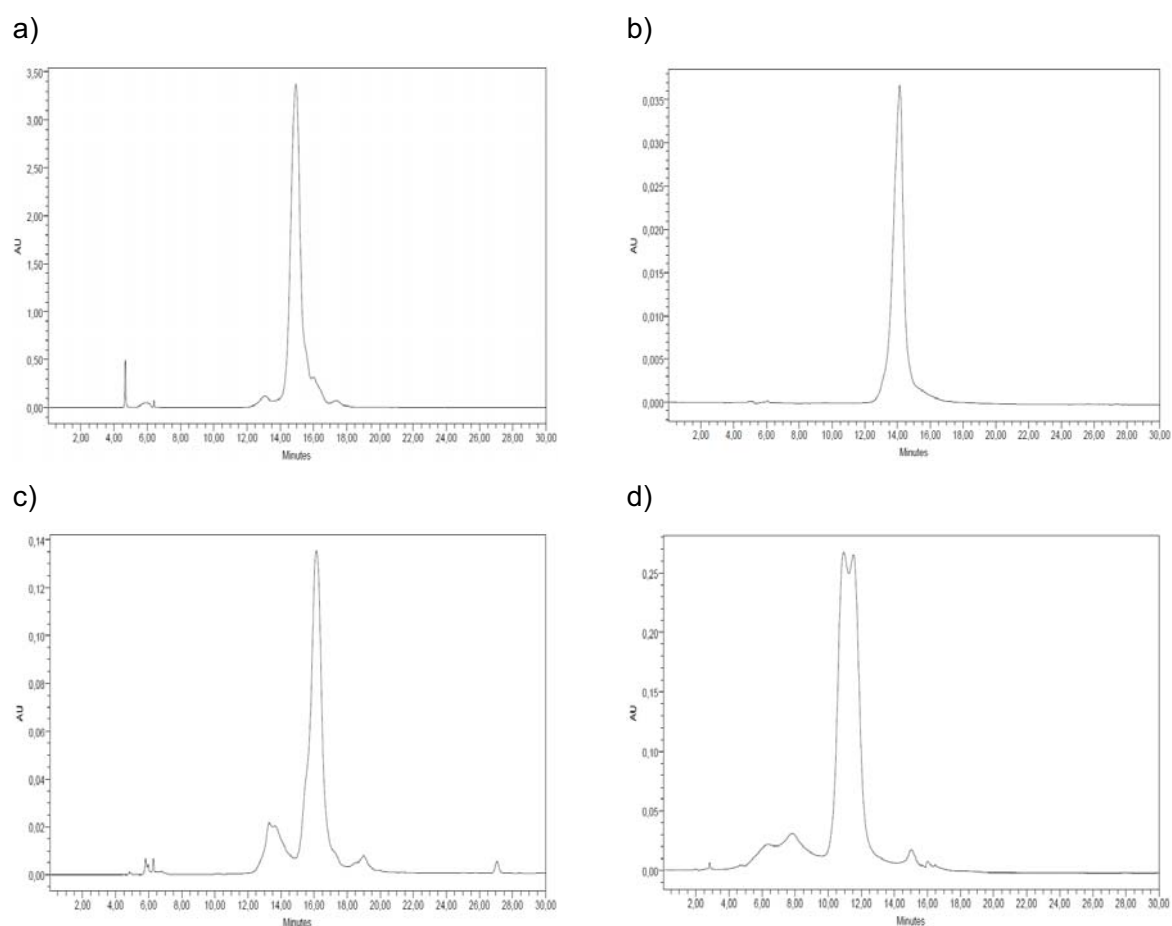
**Figure S19.** a) Crude conjugate **36** (the small peaks eluting at ~ 15 min and ~ 16.5 min correspond to unreacted **31** and diene-dT<sub>10</sub>-PS, respectively; Scheme 4b, 3.i); b) conjugate **37** after microwave-promoted deprotection (Scheme 4b, 3.ii); c) conjugate **38** (Scheme 4b, 3.iii).

**Conjugate 40.** Conjugation reaction: **39**/glutathione molar ratio: 1:10, overnight; conjugation yield (from the HPLC of the crude): 92 %. **39** Was analyzed and purified at the semipreparative scale (Jupiter Proteo column, 3 mL/min, linear gradient from 10 to 40 % of B in 30 min):  $t_R = 14.9$  min (Figure S20a); MALDI-TOF MS (negative mode):  $m/z$  4166.4 [M-H]<sup>-</sup>, M calcd. for C<sub>169</sub>H<sub>218</sub>N<sub>72</sub>O<sub>55</sub> 4167.6.

**Conjugate 41.** Deprotection of conjugate **40** was carried out by MW irradiation at 100 °C for 1 h, deprotection yield (from the HPLC trace): quantitative (Figure S20b); analytical HPLC (Jupiter Proteo column, 3 mL/min, linear gradient from 10 to 40 % of B in 30 min):  $t_R = 14.1$  min; MALDI-TOF MS (negative mode):  $m/z$  4070.8 [M-H]<sup>-</sup>, M calcd. for C<sub>163</sub>H<sub>210</sub>N<sub>72</sub>O<sub>54</sub>S 4071.6.

**Conjugate 42.** Conjugation reaction: **41**/biotin molar ratio: 1:10, overnight; conjugation yield (from the HPLC of the crude): 84 % (Figure S20c). **42** Was analyzed and purified at the semipreparative scale (Jupiter Proteo column, 3 mL/min, linear gradient from 10 to 40 % of B in 30 min):  $t_R = 16.1$  min (Figure S20c); MALDI-TOF MS (negative mode):  $m/z$  4373.3 [M-H]<sup>-</sup>, M calcd. for C<sub>175</sub>H<sub>231</sub>N<sub>75</sub>O<sub>56</sub>S<sub>3</sub> 4374.7.

**Conjugate 43.** Conjugation reaction: **41**/diene-dT<sub>10</sub>-PS molar ratio: 1:1, overnight, conjugation yield (from the HPLC trace of the crude): 92 % (Figure S20d). **43** Was analyzed and purified at the semipreparative scale (Jupiter Proteo column, 3 mL/min, linear gradient from 20 to 80 % of B in 30 min; A: 0.1 M triethylammonium acetate, solvent B: ACN):  $t_R = 10.9$  min and  $t_R = 11.5$  min (two diastereomers) (Figure S20d); MALDI-TOF MS (negative mode):  $m/z$  7369.4 [M-H]<sup>-</sup>, 3684.1 [(M-2H)/2]<sup>2-</sup>, M calcd. for C<sub>269</sub>H<sub>350</sub>N<sub>92</sub>O<sub>115</sub>P<sub>10</sub>S<sub>11</sub> 7369.9.



**Figure S20.** a) Crude conjugate **40**; b) conjugate **41** after microwave-promoted deprotection; c) crude conjugate **42**; d) crude conjugate **43**.

## 6. Abbreviations and references.

**Abbreviations:** Ac=acetyl, ACN=acetonitrile, Bhoc=benzhydryloxycarbonyl, Boc=*tert*-butoxycarbonyl, CNE=2-cyanoethyl, DCC= *N,N'*-dicyclohexylcarbodiimide, DIAD=diisopropylazodicarboxylate, DCM=dichloromethane, DIEA=*N,N*-diisopropylethylamine, DIPC=*N,N'*-diisopropylcarbodiimide, DHB=2,5-dihydroxybenzoic acid, DMAP=*N,N*-dimethylamino-pyridine, DMF=*N,N*-dimethylformamide, Dmf=dimethylaminomethylene (dimethylformamidine), DMT=4,4'-dimethoxytrityl, EDC=*N*-ethyl-*N'*-(3-dimethylaminopropyl)-carbodiimide·HCl, Fmoc=9-fluorenylmethoxycarbonyl, HATU=2-(7-aza-1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate, HOBT=1-hydroxybenzotriazole, iPrPac=isopropylphenoxyacetyl, LCAA-CPG=long chain aminoalkyl controlled pore glass beads, MSNT=1-(mesitylene-2-sulfonyl)-3-nitro-1,2,4-triazole, NMI=*N*-methylimidazole, NMM=*N*-methylmorpholine, NMP=*N*-methylpyrrolidone, Pac=phenoxyacetyl, PNA=peptide nucleic acid, TCA=trichloroacetic acid, TEA=triethylamine, THF=tetrahydrofuran, TIS=triisopropylsilane, TFA=trifluoroacetic acid.

### References:

1. Pearson, R. J., Kassianidis, E., Slawin, A. M. Z., and Philp, D. (2004) *Org. Biomol. Chem.* 2, 3434-3441.
2. Sánchez, A., Pedroso, E., and Grandas, A. (2011) *Org. Lett.* 13, 4363-4367.
3. Sánchez, A., Pedroso, E., and Grandas, A. (2012) *Bioconjugate Chem.* 23, 300-307.
4. Thibadeau, K., Léger, R., Huang, X., Robitaille, M., Quraishi, O., Soucy, C., Bousquet-Gagnon, N., van Wyk, P., Paradis, V., Castaigne J.-P., and Bridon, D. (2005) *Bioconjugate Chem.* 16, 1000-1008.
5. a) Gasser, G., Hüsken, N., Köster, S. D., and Metzler-Nolte, N. (2008) *Chem. Commun.* 3675-3677; b) Hüsken, N., Gasser, G., Köster, S. D., and Metzler-Nolte, N. (2009) *Bioconjugate Chem.* 20, 1578-1586.
6. Marchán, V., Ortega, S., Pulido, D., Pedroso, E., and Grandas, A. (2006) *Nucleic Acids Res.* 34, e24.
7. Kaiser, E., Colescott, R. L., Bossinger, C. D., and Cook, P. I. (1970) *Anal. Biochem.* 34, 595-598.
8. Zuckermann, R. N., Kerr, J. M., Kent, S. B. H., and Moos, W. H. (1992) *J. Am. Chem. Soc.* 114, 10646-10647.



# **PAPER II**



# Straightforward Synthesis of Cyclic and Bicyclic Peptides

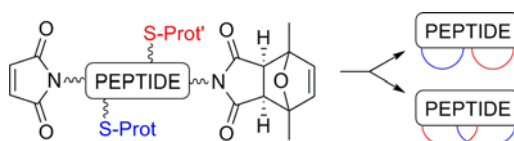
Xavier Elduque, Enrique Pedroso, and Anna Grandas\*

Departament de Química Orgànica and IBUB, Facultat de Química,  
Universitat de Barcelona, Martí i Franquès 1-11, 08028 Barcelona, Spain

anna.grandas@ub.edu

Received March 18, 2013

## ABSTRACT



Cyclic peptide architectures can be easily synthesized from cysteine-containing peptides with appending maleimides, free or protected, through an intramolecular Michael-type reaction. After peptide assembly, the peptide can cyclize either during the trifluoroacetic acid treatment, if the maleimide is not protected, or upon deprotection of the maleimide. The combination of free and protected maleimide moieties and two orthogonally protected cysteines gives access to structurally different bicyclic peptides with isolated or fused cycles.

Cyclization has been long recognized as providing peptides with increased stability to enzymatic degradation and better cell permeability (see ref 1 for recent reviews). Cyclization imposes structural constraints and allows for structural preorganization of functional groups, but the degree of flexibility still permitted is expected to enhance and facilitate interaction with the receptor target. In addition to their potential role as enzyme inhibitors, cyclic peptides are considered useful tools to interrogate complex structures and show promise in the interference of protein–protein interactions.<sup>1</sup> Cyclic peptides can also mimic protein loops. In this respect, appending conformationally constrained peptides from a scaffold can simulate the distribution of protein loops in space,<sup>2</sup> and this is of interest for immunological studies. In a different context, a Zinc-finger-type phosphorylated peptide with a cycle formed by metal chelation was able to distinguish between DNAs incorporating one of the two cytosines involved in epigenetic regulation, namely 5-hydroxymethylcytosine and 5-methylcytosine.<sup>3</sup>

Peptide macrocycles can be obtained by bridging the two ends of the peptide chain (head-to-tail), internal positions,

or both. Chemical synthesis has provided the most frequently occurring natural bridges (disulfides, macrolactams, macrolactones),<sup>1,4</sup> as well as rings with biaryls and diaryl ethers.<sup>5</sup> Moreover, ring-closing metathesis, the Cu(I)-catalyzed azide–alkyne cycloaddition,<sup>1,4</sup> the reaction between a nucleophile and an activated pyridine-*N*-oxide,<sup>6</sup> and sulfur-mediated reactions<sup>7</sup> have also afforded peptide cycles. When the latter involve a thiol and an alkene, thioether formation may take place through either radical<sup>8</sup> or Michael-type processes.<sup>9</sup> Even though the addition of a thiol to a maleimide is one of the oldest and most extensively used “click” reactions, it has hardly ever been exploited to obtain cyclic biomolecules.<sup>10</sup> There is one

(4) (a) Jiang, S.; Li, Z.; Ding, K.; Roller, P. P. *Curr. Org. Chem.* **2008**, *12*, 1502–1542. (b) Driggers, E. M.; Hale, S. P.; Lee, J.; Terret, N. K. *Nat. Rev. Drug Discovery* **2008**, *7*, 608–624.

(5) (a) Pitsinos, E. N.; Vidali, V. P.; Coladourous, E. A. *Eur. J. Org. Chem.* **2011**, 1207–1222. (b) Meyer, F.-M.; Collins, J. C.; Borin, B.; Bradow, J.; Liras, S.; Limberakis, C.; Mathiowitz, A. M.; Philippe, L.; Price, D.; Song, K.; James, K. *J. Org. Chem.* **2012**, *77*, 3099–3114.

(6) Londregan, A. T.; Farley, K. A.; Limberakis, C.; Mullins, P. B.; Piotrowski, D. W. *Org. Lett.* **2012**, *14*, 2890–2893.

(7) Hoyle, C. E.; Lowe, A. B.; Bowman, C. N. *Chem. Soc. Rev.* **2010**, *39*, 1355–1387.

(8) Aimetti, A. A.; Shoemaker, R. K.; Lin, C.-C.; Anseth, K. S. *Chem. Commun.* **2010**, *46*, 4061–4063.

(9) (a) Polinski, A.; Cooney, M. G.; Toy-Palmer, A.; Ösapay, G.; Goodman, M. *J. Med. Chem.* **1992**, *35*, 4185–4194. (b) Galande, A. K.; Trent, J. O.; Spatola, A. F. *Biopolymers* **2003**, *71*, 534–551. (c) Zhu, Y.; Gieselmann, M. D.; Zhou, H.; Averin, O.; van der Donk, W. *Org. Biomol. Chem.* **2003**, *1*, 3304–3315. (d) Matteucci, M.; Bhalay, G.; Bradley, M. *Tetrahedron Lett.* **2004**, *45*, 1399–1401.

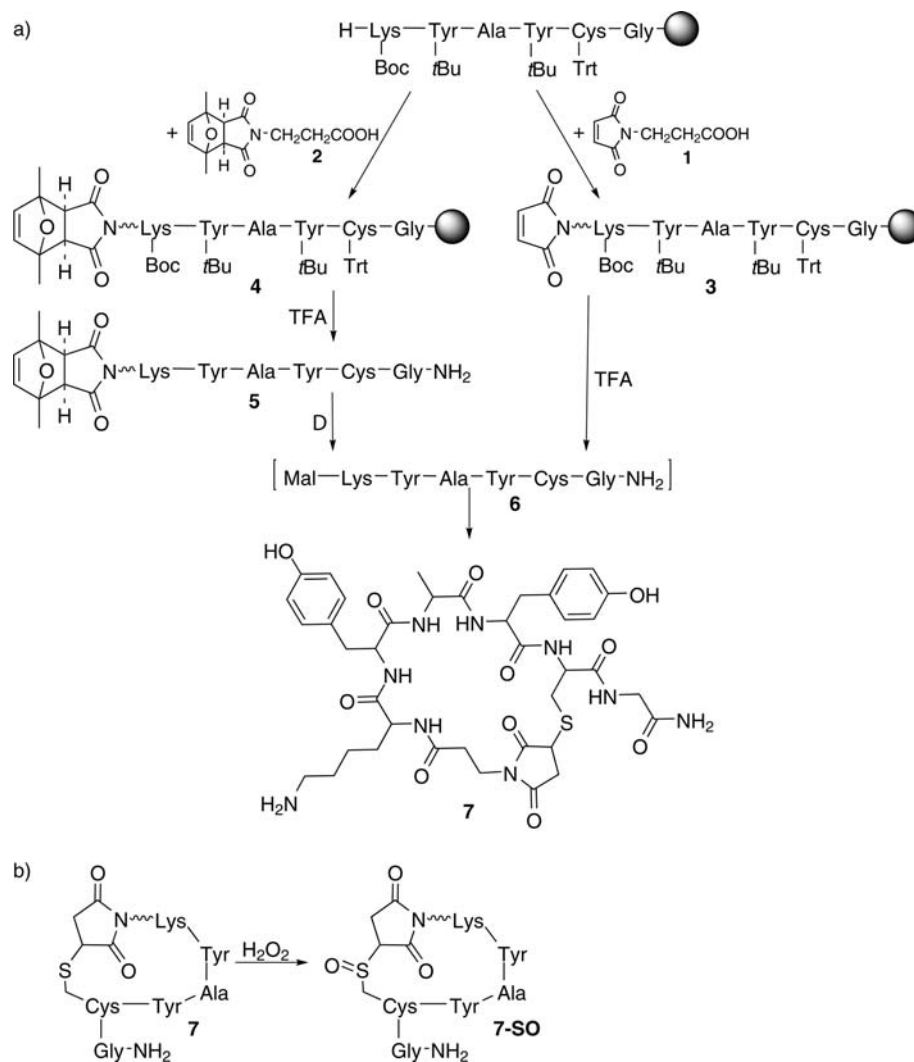
(1) (a) Marsault, E.; Peterson, M. L. *J. Med. Chem.* **2011**, *54*, 1961–2004. (b) Madsen, C. M.; Clausen, M. H. *Eur. J. Org. Chem.* **2011**, 3107–3115. (c) White, C. J.; Yudin, A. K. *Nat. Chem.* **2011**, *3*, 509–524.

(2) (a) Timmerman, P.; Beld, J.; Puijk, W. C.; Meloren, R. H. *ChemBioChem* **2005**, *6*, 821–824. (b) Heinis, C.; Rutherford, T.; Freund, S.; Winter, G. *Nat. Chem. Biol.* **2009**, *5*, 502–507. (c) Ghosh, P. S.; Hamilton, A. D. *J. Am. Chem. Soc.* **2012**, *134*, 13208–13211.

(3) Nomura, A.; Sugizaki, K.; Yanagisawa, H.; Okamoto, A. *Chem. Commun.* **2011**, *47*, 8277–8279.



**Scheme 1.** First Syntheses of a Cyclic Peptide (a) and Assessment of Cyclization (b)



report dealing with the on-resin cyclization of a peptide incorporating an *N*-terminal maleimide<sup>10a</sup> and a very recent one on the synthesis of cyclic oligonucleotides.<sup>10b</sup> Here, we describe its application to the preparation of peptide mono- and bicycles in solution.

To prove the suitability of the thiol–ene Michael reaction for this purpose, peptide sequence KYAYCG was assembled on the Rink amide resin (Scheme 1a; for details see Supporting Information). Maleimides are labile to nucleophilic bases and thus not compatible with the piperidine treatments that remove the Fmoc group in standard Fmoc/*t*Bu peptide chemistry. Hence, the maleimide moiety has to be incorporated either after the chain has been fully assembled or during chain elongation provided that it is protected.<sup>11</sup> For comparison purposes, part of the protected peptide resin was derivatized with 3-maleimidopropanoic

acid (1) and partly with 3-[2,5-dimethylfuran protected maleimido]propanoic acid (2).<sup>12</sup> Trifluoroacetic acid treatment of peptide-resin 3 provided a crude whose main product (Figure S1) was isolated and treated with H<sub>2</sub>O<sub>2</sub> to investigate whether it was linear (6) or circular (7). Under the reaction conditions,<sup>10b</sup> the free thiol of the linear precursor (6) is oxidized to sulfonic acid (mass 48 units higher), whereas the thioether of the cyclic compound becomes a sulfoxide (mass 16 units higher). After treatment with H<sub>2</sub>O<sub>2</sub> it was found that the main product in the crude was the sulfoxide of the cyclic molecule (7-SO, Scheme 1b), which showed that the Michael-type thiol-maleimide reaction can take place even in the strongly acidic deprotection medium, either during the deprotection procedure or upon elimination of the deprotecting mixture when the solution becomes more concentrated. 7 was found to be accompanied by some cyclic dimeric and trimeric peptides (~24% in the crude; see Figure S1).

(10) (a) Sharma, S. K.; Wu, A. D.; Chandramouli, N. *Tetrahedron Lett.* **1996**, *37*, 5665–5668. (b) Sánchez, A.; Pedroso, E.; Grandas, A. *Chem. Commun.* **2013**, *49*, 309–311.

(11) Elduque, X.; Sánchez, A.; Sharma, K.; Pedroso, E.; Grandas, A. *Bioconjugate Chem.*, accepted for publication.

(12) Sánchez, A.; Pedroso, E.; Grandas, A. *Org. Lett.* **2011**, *13*, 4363–4367.

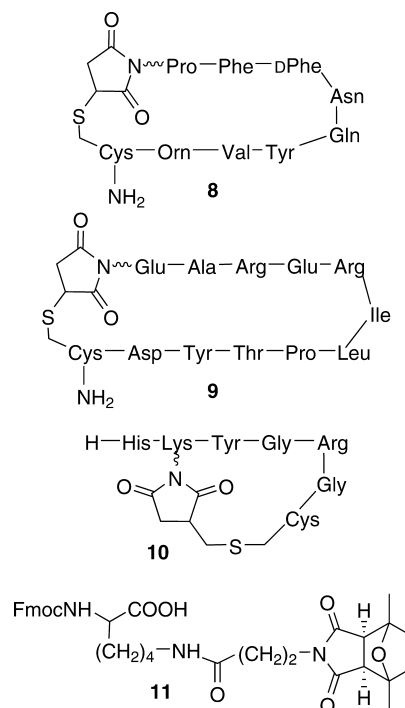
Acidolysis (trifluoroacetic acid) of peptide-resin **4** removed permanent protecting groups on the peptide, affording [protected maleimido]-peptide **5** (Scheme 1a). The maleimide was then deprotected by heating in a microwave oven at 90 °C. In agreement with previous results,<sup>10,11</sup> the retro-Diels–Alder reaction and cyclization took place simultaneously, furnishing a thiosuccinimide that remained stable to the heating conditions.

Different peptide concentrations were tested to examine under which maleimide deprotection conditions the target monocycle **7** was exclusively formed from **5**, or at the least formation of **7** was favored over that of oligomers. As summarized in Table S1 (see also Figure S3), in no case was it possible to fully prevent the formation of the cyclic dipeptide. Yet, in concentrations below 1 mM the extent of cyclic dimer formation did not exceed 5%, and the formation of cyclic tripeptide was virtually nil. Hence, protection of the maleimide allows for greater control of the dilution conditions in which cyclization takes place, so that fully protected [maleimido peptide]-resins such as **4** are better precursors of peptide rings than peptide resins with a free maleimide such as **3**. Moreover, the overall yield of the **4** → **7** process (21%, two steps with two purifications) was higher than that of the **3** → **7** transformation (12%, one purification).

Cyclic peptides of larger size, incorporating different amino acids, were then prepared (Figure 1). The amino acid sequence of **8** (9 amino acid cycle) reproduces most of the sequence of antibiotic tyrocidine,<sup>13</sup> and **9** (12 amino acid cycle) reproduces the loop sequence of an oxidoreductase (dihydrolipoamide dehydrogenase, UNiProtKb P09063).<sup>14</sup> **10** Was synthesized to assess whether histidine interferes with the thiol-maleimide reaction, since it has been described that *N*-terminal maleimido-derivatized, histidine-containing peptides can easily undergo an intramolecular imidazole-maleimide Michael-type reaction.<sup>15</sup> However, *N*-terminal maleimido-derivatized, histidine-containing peptides have been successfully used in conjugation with thiol-derivatized oligonucleotides.<sup>16</sup>

Cyclic peptides **8** and **9** were prepared by attaching **2** to the peptide resin. Compound **11** (Figure 1, ref 11), in which the protected maleimide moiety is linked to the side chain of lysine, was used for the synthesis of **10**. In all cases solid-phase assembly was first followed by an acidolytic treatment to remove protecting groups on the peptide chain, and after purification the retro-Diels–Alder reaction that deprotects the maleimide was carried out. As in the case of **7**, the latter was accompanied by cyclization, as assessed by the reaction with H<sub>2</sub>O<sub>2</sub>. These experiments showed that the thiol–maleimide reaction provides peptide cycles with at least 12 amino acids with very good [maleimide deprotection + cyclization] yields (~ 50%), both with head-to-side chain and side chain-to-side chain linkages, and including a

variety of trifunctional amino acids. They also showed that histidine-containing peptides can be cyclized making use of the Michael-type reaction even if the histidine residue is very close to the maleimide unit. In agreement with other authors' results,<sup>8</sup> cyclic peptides were found to elute before the linear precursors upon HPLC analysis.



**Figure 1.** Structures of cyclic peptides **8**–**10** and of the lysine derivative **11** used in the synthesis of **10**, **14**, and **17**.

To further explore the scope of the method, two bicyclic peptides, **14** and **17** (Scheme 2), were synthesized. In the former (“spectacles-like”) the two cycles are separated by two residues of the peptide chain, while in the latter the cycles have part of the chain in common. For their preparation we combined 3-maleimidopropanoic acid (**1**), the lysine derivative **11**, and two orthogonally protected cysteine residues, Cys(Trt) and Cys(*S*-*t*Bu). Since **1** can only be installed at the *N*-terminal (see above), by adequately placing the other three residues either type of bicycle can be obtained.

Trifluoroacetic acid treatment of peptide-resins **12** and **15** removed all protecting groups except those of Cys(*S*-*t*Bu) and the maleimido-derivatized lysine, so that the reaction between the *N*-terminal maleimide and the free thiol furnished the first cycle (compounds **13** and **16**, respectively). The *S*-*t*Bu group was subsequently removed from the still protected cysteine residue by reaction with tris(carboxyethyl)phosphine (TCEP). After purification to remove excess phosphine and prevent a reaction between the phosphine and the maleimide,<sup>10b</sup> heating in a microwave oven deprotected the maleimide attached to the lysine side chain and promoted the

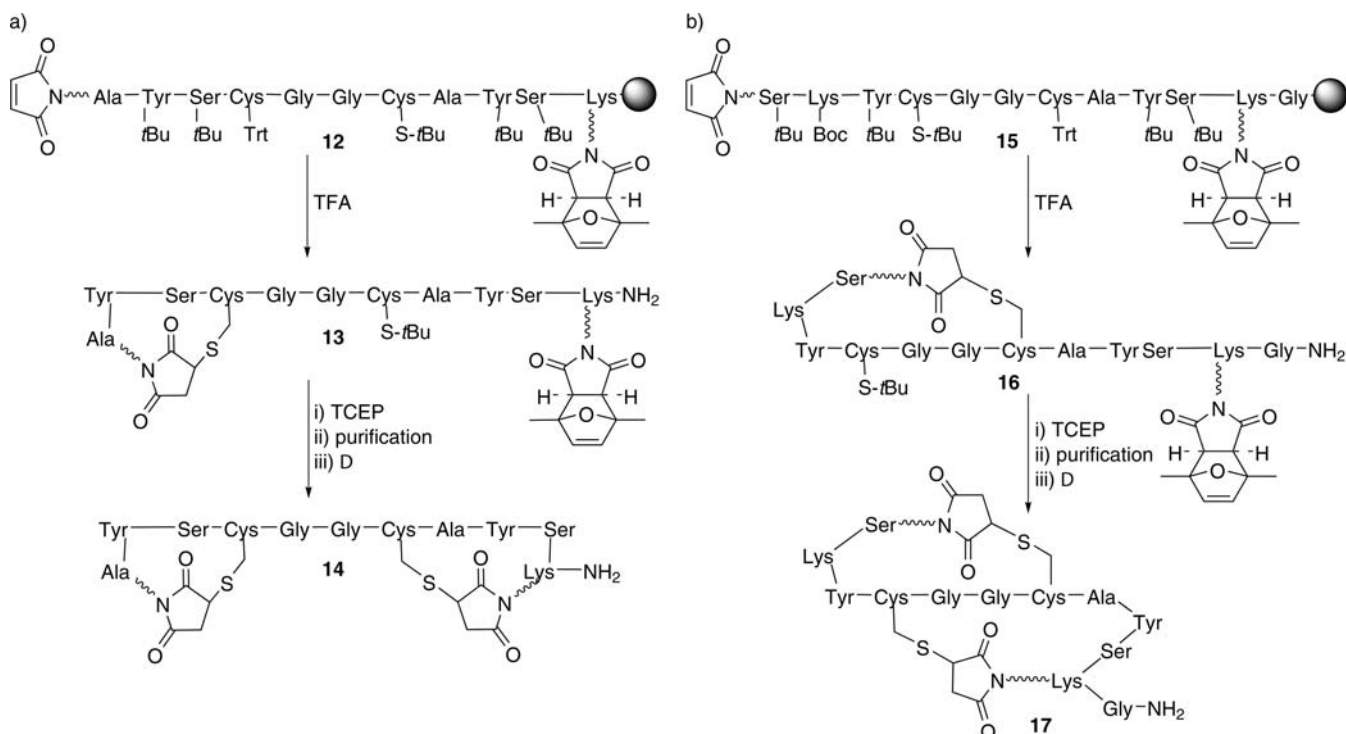
(13) Sieber, S. A.; Marahiel, M. A. *J. Bacteriol.* **2003**, *185*, 7036–7043.

(14) Li, W.; Liu, Z.; Lai, L. *Biopolymers* **1999**, *49*, 481–495.

(15) Papini, A.; Rudolph, S.; Sigmüller, G.; Musiol, H.-G.; Gohring, W.; Moroder, L. *Int. J. Peptide Protein Res.* **1992**, *39*, 348–355.

(16) Ming, X.; Alam, M. R.; Fisher, M.; Yan, Y.; Chen, X.; Juliano, R. L. *Nucleic Acids Res.* **2010**, *38*, 6567–6576.

**Scheme 2.** Synthesis of Bicyclic Peptides with Either Two Peptide-Linked Rings (a) or Two Cycles with Shared Amino Acids (b)



second cyclization. In both cases the target bicycle (**14** or **17**) was obtained. With this sequence of steps it is not possible to prevent the formation of oligomeric cyclic peptides during the first cyclization.

In summary, the Michael-type thiol–maleimide reaction is a straightforward procedure for the synthesis of cyclic peptides. If the peptide is appended with a free maleimide, cyclization takes place during the acid-promoted removal of peptide protecting groups, and formation of some cyclic dimers and trimers cannot be fully avoided. In case the peptide is derivatized with a 2,5-dimethylfuran-protected maleimide, which can be placed anywhere in the sequence, cyclization and maleimide deprotection take place simultaneously and in good yield, and formation of oligomers can be kept to a much lower extent. The method has proved suitable to synthesize up to 43-membered rings. It is compatible with most trifunctional amino acids, namely those bearing hydroxyl, acid, amide, and basic functional groups.

An adequate choice of the protection scheme allowed structurally different bicyclic peptides, with the two rings fused or isolated, to be obtained. For this purpose peptide chain precursors incorporated two differently protected cysteines, a free maleimide and a protected maleimide.

**Acknowledgment.** This work was supported by funds from the Ministerio de Economía y Competitividad (Grant CTQ2010-21567-C02-01, and the project RNAREG, Grant CSD2009-00080, funded under the programme CONSOLIDER INGENIO 2010) and the Generalitat de Catalunya (2009SGR-208). X.E. was a recipient fellow of the MINECO.

**Supporting Information Available.** Experimental procedures, compound characterization data and spectra, and HPLC profiles. This material is available free of charge via the Internet at <http://pubs.acs.org>.

The authors declare no competing financial interest.

**SUPPORTING INFORMATION TO:**

**Straightforward synthesis of cyclic and bicyclic peptides**

by Xavier Elduque, Enrique Pedroso and Anna Grandas\*

	Page
1. General materials and methods.	S2
2. First assays: Cyclic peptide <b>7</b> .	S5
3. Other monocyclic peptides.	S9
4. Bicyclic peptides.	S14
5. Abbreviations.	S18

## 1. General materials and methods.

Fmoc-Ala-OH, Fmoc-Asn(Trt)-OH, Fmoc-Asp-OH, Fmoc-Cys(S-*t*Bu)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Gly-OH, Fmoc-Ile-OH, Fmoc-Leu-OH, Fmoc-Lys-OH, Fmoc-Lys(Boc)-OH, Fmoc-Orn(Boc)-OH, Fmoc-Phe-OH, Fmoc-DPhe-OH, Fmoc-Pro-OH, Fmoc-Ser(*t*Bu)-OH, Fmoc-Thr(*t*Bu)-OH, Fmoc-Val-OH and the Rink Amide MBHA support were from Novabiochem. Fmoc-Arg(Pbf)-OH, Fmoc-Cys(Trt)-OH and Fmoc-Tyr(*t*Bu)-OH were from Iris Biotech. 3-Maleimidopropanoic acid (**1**) was from Bachem. [Protected maleimido]propanoic acid (**2**, *exo* adduct) and Fmoc-Lys[protected maleimide]-OH (**11**, *exo* adduct) were prepared as previously described (references 10b and 11 of the manuscript, respectively). DIPC and trityl chloride were from Sigma-Aldrich. HOBt·H<sub>2</sub>O was purchased to Acros Organics. Acid-free DCM was obtained by filtration through basic alumina.

TLC was carried out on silica gel plates 60 F<sub>254</sub> from Merck. Microwave irradiation was carried out in a Biotage Initiator™ oven, using reaction vessels of 0.2-0.5 mL or 0.5-2 mL as required. Samples were lyophilized in a FreezeMobile Virtis instrument.

<sup>1</sup>H and <sup>13</sup>C NMR spectra were recorded on a Varian Mercury 400 MHz spectrometer.

### Solid-phase peptide synthesis (Fmoc/*t*Bu strategy).

Peptides were elongated manually in polyethylene syringes fitted with polypropylene discs using standard procedures (see below).

The Rink Amide resin was treated and washed with DCM (3 ×), DMF (3 ×), 20 % piperidine/DMF (2 × 15 min), DCM (3 ×), DMF (3 ×) and MeOH (3 ×) prior to chain elongation. Incorporation of the first amino acid on the resin and peptide elongation were accomplished by using 3 equiv. of both Fmoc-amino acid, HOBt·H<sub>2</sub>O and DIPC (dissolved in the minimum amount of DCM and a few drops of DMF, 90 min), which was followed by washing with DCM, DMF and MeOH (3 ×). In case the coupling was not complete, as assessed by the Kaiser test (Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. *Anal. Biochem.* **1970**, *34* 595-598), it was repeated using 2 equiv. of the reagents. Amine groups on the resin still free after 2-3 couplings were capped. Washing with DCM (3 × 10 s) was followed by reaction with a 1.1:1 Ac<sub>2</sub>O/pyr mixture (2 × 15 min), and this by washing with DCM and MeOH (3 × 10 s each).

Removal of the Fmoc groups was effected by reaction with 20 % piperidine/DMF (1 × 3 min + 1 × 10 min), followed by washing (DMF, DCM). To determine the substitution degree, the filtrates from the deprotection treatments and washings were collected in a volumetric flask. This solution was properly diluted with DCM to achieve an

approximate concentration of 0.1  $\mu\text{M}$ . Its absorbance was measured at  $\lambda=301\text{ nm}$ , where  $\epsilon=7800\text{ M}^{-1}\text{cm}^{-1}$ <sup>19</sup>.

Activation of 3-maleimidopropanoic acid (**1**) was carried out with DIPC (90 min reaction time). Even though activation with DIPC and HOBt·H<sub>2</sub>O seemed not to cause any harm to the maleimide, use of only the carbodiimide is safer to prevent addition of HOBt.

Deprotection and cleavage from the resin was carried out by reaction with either 95:2.5:2.5 TFA/H<sub>2</sub>O/TIS or 87.5:5:5:2.5 TFA/H<sub>2</sub>O/EtSH/TIS mixtures (2 × 1 h, room temperature). Filtrate and washings (DCM) were collected and concentrated under a N<sub>2</sub> stream, and cold diethyl ether was added to the resulting oil. The mixture was centrifuged (10 min, 5 °C, 4800 rpm) and decanted. This procedure was repeated 3 times.

Peptide crudes were purified by HPLC, quantified by UV spectroscopy (Tyr:  $\epsilon_{280}$  ( $\text{M}^{-1}\text{cm}^{-1}$ ) = 1480), and characterized by mass spectrometry.

#### **HPLC.**

Reversed-phase HPLC analysis and purification was performed using analytical and semipreparative Shimadzu systems.

**Analysis conditions:** Jupiter Proteo column (4  $\mu\text{m}$ , 90 Å, 250 × 4.6 mm) from Phenomenex; solvent A: 0.045 % TFA in water, solvent B: 0.036 % TFA in ACN, flow: 1 mL/min, detection wavelength: 280 nm.

**Purification conditions (semipreparative scale):** Jupiter Proteo column (10  $\mu\text{m}$ , 90 Å, 250 × 10.0 mm) from Phenomenex; solvent A: 0.1 % TFA in water, solvent B: 0.1 % TFA in ACN, flow: 3 mL/min, detection wavelength: 280 nm.

#### **Mass spectrometry.**

MALDI-TOF mass spectra were recorded on a 4800 *Plus* ABSciex instrument. Unless otherwise indicated, reflector was used. Analysis conditions: 2',4',6'-Trihydroxyacetophenone monohydrate (THAP) with ammonium citrate (CA), or 2,5-dihydroxybenzoic acid (DHB) + 0.1 % TFA, positive or negative mode. Spectra of peptides incorporating 2,5-dimethylfuran-protected maleimides and S-*t*Bu-protected thiols typically showed loss of the corresponding protecting groups due to the ionization method.

ESI (low and high resolution) mass spectra were obtained using an LC/MSD-TOF spectrometer from Agilent Technologies.

**Assessment of cyclization.**

Cyclization was confirmed by reacting the peptide, either linear or cyclic, with a 3.5 % solution of H<sub>2</sub>O<sub>2</sub> for 1 h, and recording the mass spectrum of the resulting crude.

When the peptide is linear, H<sub>2</sub>O<sub>2</sub> oxidizes the free thiol to sulfonic acid, whereas when the peptide is cyclic the thioether is oxidized to sulfoxide. Therefore, if the peptide is linear the original mass increases 48 units (+ 3 oxygens), while if cyclic it only increases 16 units (+ 1 oxygen).

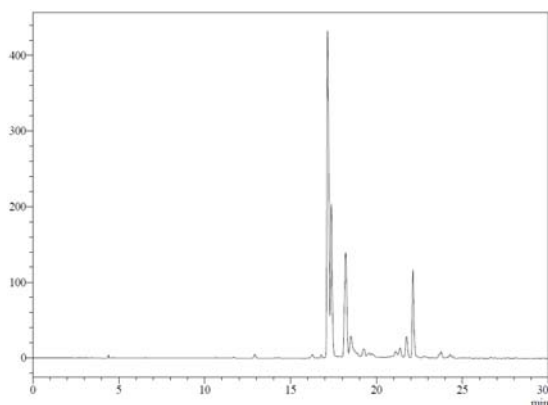
So far, in HPLC analyses run over the period several hours or in different days we have not found changes in retention times. This is in agreement with the cyclic peptides being stable in aqueous solutions, but the stability may vary depending on peptide composition and length.

## 2. First assays: Cyclic peptide 7.

**Cyclic maleimido\*-Lys-Tyr-Ala-Tyr-Cys\*-Gly-NH<sub>2</sub> 7 from 3** (the label \* refers to the units linked in the cycle).

3-Maleimidopropanoic acid (**1**) was coupled onto H-Lys(Boc)-Tyr(*t*Bu)-Ala-Cys(Trt)-Gly-NH-resin to give **3**, which was treated with the 95:2.5:2.5 TFA/H<sub>2</sub>O/TIS mixture. The main compound in the crude (later identified as **7**, Figure S1) was isolated in 12 % overall synthesis and purification yield after purification by semipreparative HPLC (linear gradient from 20 to 50 % of B in 30 min). Analytical HPLC (gradient from 5 to 50 % of B in 30 min): *t<sub>R</sub>* = 16.7, 16.9 min (2 diastereomers); MALDI-TOF MS (THAP/CA, positive mode): *m/z* 854.3 [M+H]<sup>+</sup>, M calc. for C<sub>39</sub>H<sub>51</sub>N<sub>9</sub>O<sub>11</sub>S 853.3.

MALDI-TOF MS analysis after treatment with H<sub>2</sub>O<sub>2</sub> (THAP/CA, positive mode): *m/z* 869.9 [M+H]<sup>+</sup>, M calc. for C<sub>39</sub>H<sub>51</sub>N<sub>9</sub>O<sub>12</sub>S 869.3.



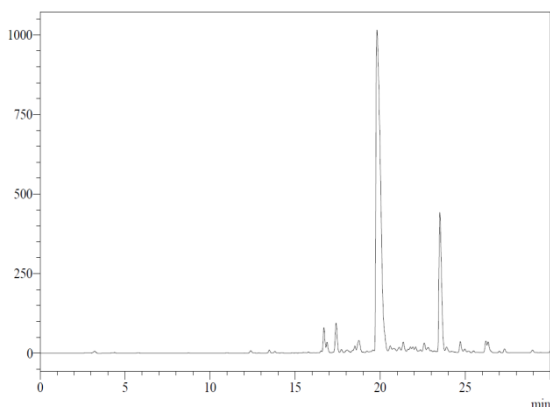
**Figure S1.** Crude peptide **7** from **3** (**7** is the main product in the crude).

### **Cyclic maleimido\*-Lys-Tyr-Ala-Tyr-Cys\*-Gly-NH<sub>2</sub> 7 from 4.**

**[Protected maleimido]-Lys-Tyr-Ala-Tyr-Cys-Gly-NH<sub>2</sub> 5.** Peptide-resin **4** was assembled by coupling **2** onto H-Lys(Boc)-Tyr(*t*Bu)-Ala-Tyr(*t*Bu)-Cys(Trt)-Gly-NH-resin. Removal of peptide protecting groups was carried out by reaction with the 95:2.5:2.5 TFA/H<sub>2</sub>O/TIS mixture, which furnished **5**. **5** (Figure S2) was purified at the semipreparative scale (linear gradient from 20 to 100 % of B in 30 min) and obtained in 42 % overall yield (synthesis and purification). Analytical HPLC (linear gradient from 5 to 50 % of B in 30 min): *t<sub>R</sub>* = 20.1 min; MALDI-TOF MS (THAP/CA, positive mode): *m/z*: 854.3 [M-furan+H]<sup>+</sup>, 950.4 [M+H]<sup>+</sup>; calc. for C<sub>45</sub>H<sub>59</sub>N<sub>9</sub>O<sub>13</sub>S 949.4.

MALDI-TOF MS after reaction with H<sub>2</sub>O<sub>2</sub> (THAP/CA, negative mode): *m/z*: 900.5 [M-furan+48-H]<sup>-</sup>; 996.5 [M+48-H]<sup>-</sup>; calc. for C<sub>45</sub>H<sub>60</sub>N<sub>9</sub>O<sub>16</sub>S 997.4.





**Figure S2.** Crude peptide **5** - linear precursor of **7** (**5** is the main product in the crude).

**Maleimide deprotection and cyclization: Cyclic maleimido\*-Lys-Tyr-Ala-Tyr-Cys\*-Gly-NH<sub>2</sub> **7**.** Peptide **5** was dissolved in a 1:1 H<sub>2</sub>O/MeOH solution (1 mL), and heated in a microwave at 90 °C for 90 min. This process was carried out with solutions of different concentration of **5** (see Figure S3 and Table S1). After the 90 min reaction, methanol was removed under reduced pressure and the resulting crude analyzed by HPLC (linear gradient from 5 to 50 % of B in 30 min):  $t_R = 17.0, 17.2$  (2 diastereomers). The structures of the products in the different fractions collected were assessed by MALDI-TOF MS.

Maleimide deprotection+cyclization and purification yield (100  $\mu$ M deprotection scale) was 51%.

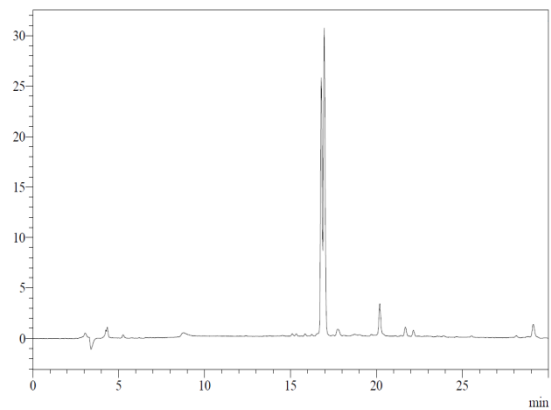
[Maleimide deprotection + cyclization] assays attempted by heating in toluene (reference 12 of the manuscript) gave similar or worse results than the microwave-promoted reaction in water/MeOH (data not shown).

**7:** MALDI-TOF MS (THAP/CA, positive mode):  $m/z$ : 854.4 [M+H]<sup>+</sup>; calc. for C<sub>39</sub>H<sub>51</sub>N<sub>9</sub>O<sub>11</sub>S 853.3.

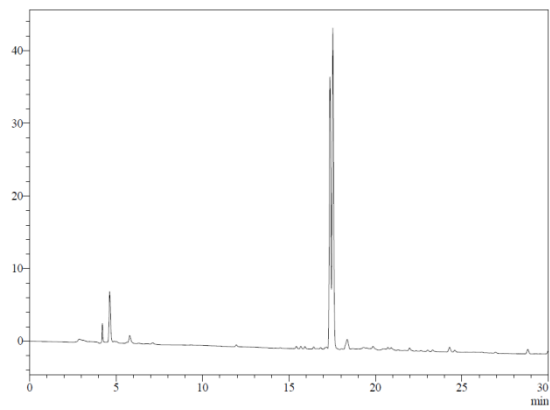
MALDI-TOF MS analysis after treatment with H<sub>2</sub>O<sub>2</sub> (DHB, positive mode): 870.4 [M+H]<sup>+</sup>; calc. for C<sub>39</sub>H<sub>51</sub>N<sub>9</sub>O<sub>12</sub>S 869.3.

The [maleimide deprotection+cyclization] process was carried out at different concentrations of [protected maleimido]-linear peptide precursor, to find reaction conditions minimizing the formation of cyclic dimer and cyclic trimer peptide. The results of these experiments are shown in Figure S3 (HPLC profiles a-f) and Table S1.

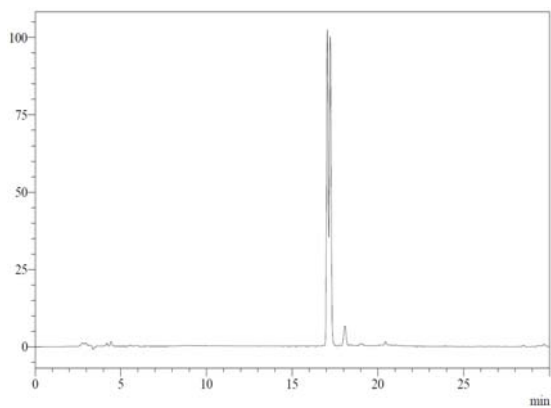
a)



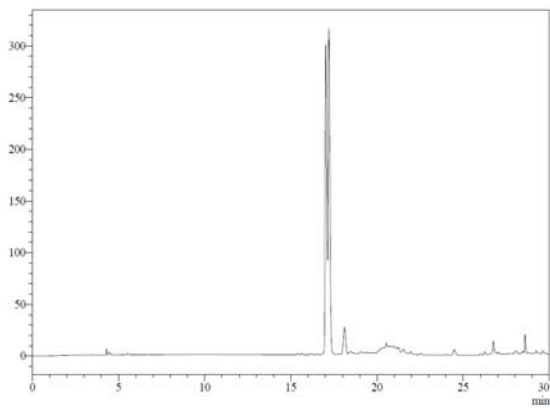
b)



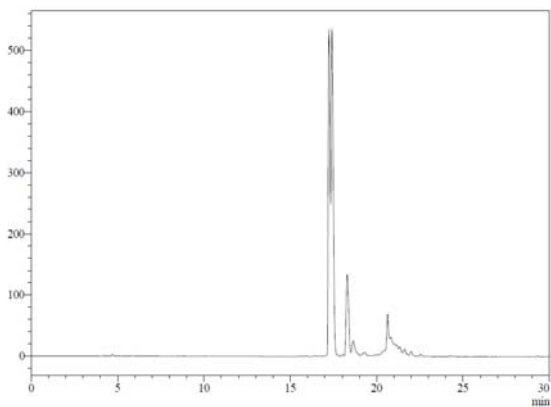
c)



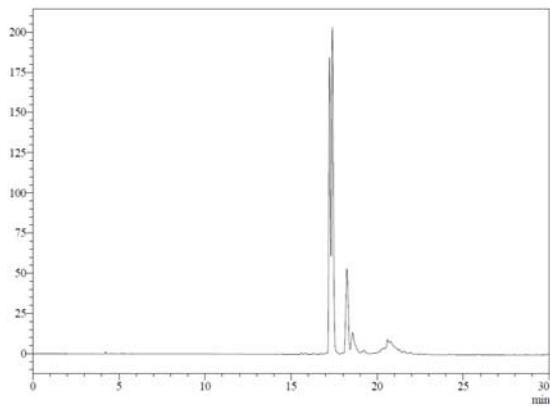
d)



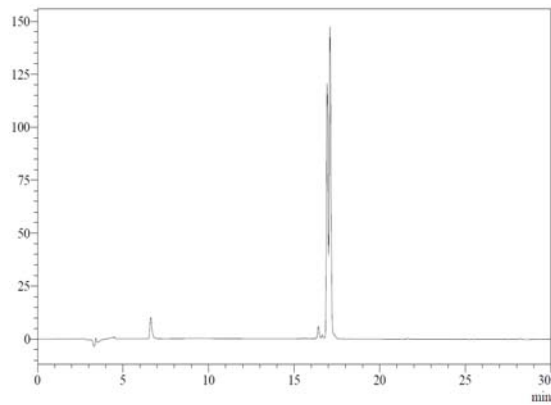
e)



f)



g)



**Figure S3.** HPLC traces of the crudes obtained after the deprotection and cyclizations assays carried out at different concentrations of **5**: 25  $\mu$ M (a), 100  $\mu$ M (b), 500  $\mu$ M (c), 1 mM (d), 5 mM (e), 10 mM (f) (the target peptide is the main peak of each profile). HPLC trace f looks very much like the trace of the crude obtained upon treatment of **3** with the acidolysis reagent (see Figure S1). The HPLC profile of purified **7** is shown in g.

**Cyclic dimer:**  $t_R$ : 18.1 min; MALDI-TOF MS (THAP/CA, positive mode): 1707.6 [M+H]<sup>+</sup>; calc. for C<sub>78</sub>H<sub>102</sub>N<sub>18</sub>O<sub>22</sub>S<sub>2</sub> 1706.7.

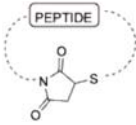
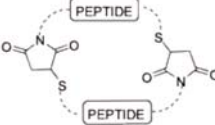
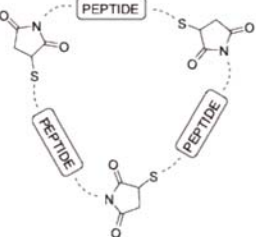
MALDI-TOF MS analysis after treatment with H<sub>2</sub>O<sub>2</sub> (DHB, positive mode): 1738.7 [M+H]<sup>+</sup>; calc. for C<sub>78</sub>H<sub>102</sub>N<sub>18</sub>O<sub>24</sub>S<sub>2</sub> 1738.7.

**Cyclic trimer:**  $t_R$ : 18.5 min; MALDI-TOF MS (THAP/CA, positive mode): 2560.6 [M+H]<sup>+</sup>; calc. for C<sub>117</sub>H<sub>153</sub>N<sub>27</sub>O<sub>33</sub>S<sub>3</sub> 2560.0.

MALDI-TOF MS analysis after treatment with H<sub>2</sub>O<sub>2</sub> (DHB, positive mode): 2608.5 [M+H]<sup>+</sup>; calc. for C<sub>117</sub>H<sub>153</sub>N<sub>27</sub>O<sub>36</sub>S<sub>3</sub> 2608.0.

**Table S1.**

Relative ratios of the products found in the crudes after maleimide deprotection, as assessed by the relative areas of the peaks at the HPLC traces.

Conc. of <b>5</b>	<b>7</b>	Dimer	Trimer	Other
				
25 $\mu$ M	94 %	2 %	-	4 %
100 $\mu$ M	97 %	3 %	-	0 %
500 $\mu$ M	96 %	4 %	-	0 %
1 mM	95 %	5 %	-	0 %
5 mM	71 %	11 %	2 %	16 %
10 mM	69 %	13 %	3 %	15 %

### 3. Other monocyclic peptides.

#### i. Linear precursors.

**[Protected maleimido]-Pro-Phe-DPhe-Asn-Gln-Tyr-Val-Orn-Cys-NH<sub>2</sub>** (linear precursor of **8**). The [protected peptide]-resin was derivatized with **2**, treated with the 87.5:5:5:2.5 TFA/H<sub>2</sub>O/EtSH/TIS mixture, and [protected maleimido]-Pro-Phe-DPhe-Asn-Gln-Tyr-Val-Orn-Cys-NH<sub>2</sub> (Figure S4a) purified at the semipreparative scale (linear gradient from 40 to 100 % of B in 30 min). Overall synthesis and purification yield: 44 %. Analytical HPLC (linear gradient from 30 to 100 % of B in 30 min): *t<sub>R</sub>* = 11.4 min; MALDI-TOF MS (THAP/CA, positive mode): *m/z*: 1281.6 [M-furan+H]<sup>+</sup>, 1377.6 [M+H]<sup>+</sup>; calc. for C<sub>67</sub>H<sub>88</sub>N<sub>14</sub>O<sub>16</sub>S 1376.6.

MALDI-TOF MS analysis after treatment with H<sub>2</sub>O<sub>2</sub> (THAP/CA, negative mode): 1327.7 [M (non oxidized peptide)-furan-H]<sup>-</sup>; 1423.8 [M-H]<sup>-</sup>; calc. for C<sub>67</sub>H<sub>89</sub>N<sub>14</sub>O<sub>19</sub>S 1424.6.

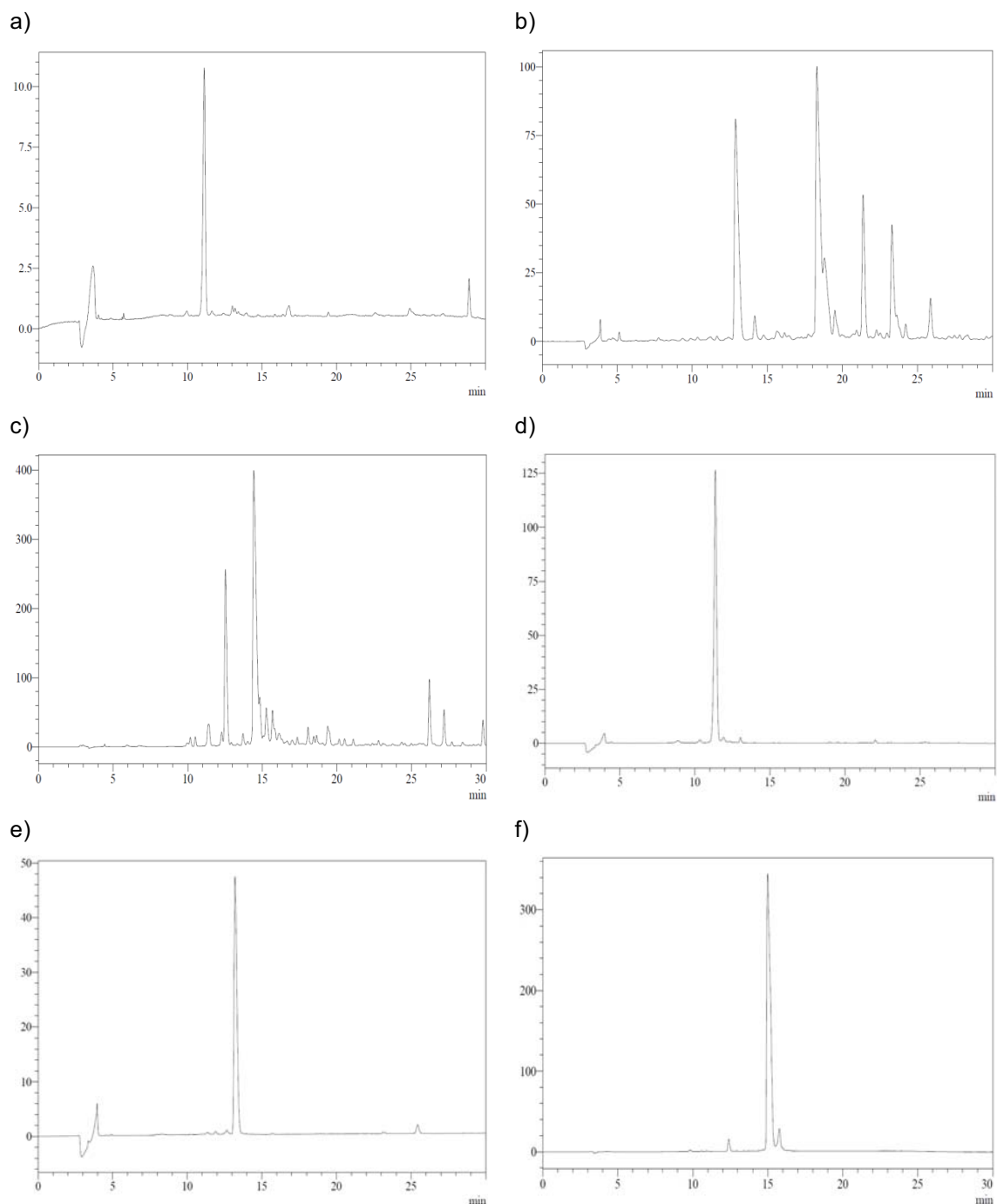
**[Protected maleimido]-Glu-Ala-Arg-Glu-Arg-Ile-Leu-Pro-Thr-Tyr-Asp-Cys-NH<sub>2</sub>** (linear precursor of **9**). The [protected peptide]-resin was derivatized with **2**, treated with the 87.5:5:5:2.5 TFA/H<sub>2</sub>O/EtSH/TIS mixture, and [protected maleimido]-Glu-Ala-Arg-Glu-Arg-Ile-Leu-Pro-Thr-Tyr-Asp-Cys-NH<sub>2</sub> (Figure S4b) purified at the semipreparative scale (linear gradient from 25 to 45 % of B in 30 min). In this case two fractions were obtained, one corresponding to the desired product and the other to the peptide that had formed a disulfide bond with ethanethiol. The latter was reacted with TCEP (715 mg, 2.5 mmol) at pH=5 for 1 h and purified again (in the same conditions). The combined fractions afforded the desired peptide in 21 % yield (synthesis and purification). Analytical HPLC (linear gradient from 20 to 40 % of B in 30 min): *t<sub>R</sub>* = 12.6 min; MALDI-TOF MS (THAP/CA, positive mode): *m/z*: 1615.8 [M-furan+H]<sup>+</sup>, 1711.9 [M+H]<sup>+</sup>; calc. for C<sub>75</sub>H<sub>114</sub>N<sub>20</sub>O<sub>24</sub>S 1710.8.

MALDI-TOF MS analysis after treatment with H<sub>2</sub>O<sub>2</sub> (THAP/CA, negative mode): 1757.6 [M-H]<sup>-</sup>; 1795.6 [M+K-2H]<sup>-</sup>; calc. for C<sub>75</sub>H<sub>114</sub>N<sub>20</sub>O<sub>27</sub>S 1758.8.

**H-His-Lys[protected maleimide]-Tyr-Gly-Arg-Gly-Cys-NH<sub>2</sub>** (linear precursor of **10**). The [protected peptide]-resin (assembled using derivative **11**) was treated with the 87.5:5:5:2.5 TFA/H<sub>2</sub>O/EtSH/TIS mixture, and H-His-Lys[protected maleimide]-Tyr-Gly-Arg-Gly-Cys-NH<sub>2</sub> (Figure S4c) purified at the semipreparative scale (linear gradient from 15 to 65 % of B in 30 min). Overall synthesis and purification yield: 35 %. Analytical HPLC (linear gradient from 5 to 50 % of B in 30 min): *t<sub>R</sub>* = 14.4 min; MALDI-

TOF MS (DHB, positive mode):  $m/z$ : 970.9 [M-furan+H]<sup>+</sup>, calc. for C<sub>47</sub>H<sub>67</sub>N<sub>15</sub>O<sub>12</sub>S 1065.5.

MALDI-TOF MS analysis after treatment with H<sub>2</sub>O<sub>2</sub> (DHB, negative mode):  $m/z$ : 1016.6 [M-furan-H]<sup>-</sup>; calc. for C<sub>47</sub>H<sub>67</sub>N<sub>15</sub>O<sub>15</sub>S 1113.5.



**Figure S4.** Linear precursors of **8** (profile a), **9** (profile b) and **10** (profile c): crudes obtained after acidolysis of peptide-resins. The target peptides are the main peaks of each profile, except in the case of (c), where the main peak is the hybrid disulfide

resulting from reaction between the peptide and ethanethiol. Profiles d, e and f: purified linear precursors of **8**, **9** and **10**, respectively.

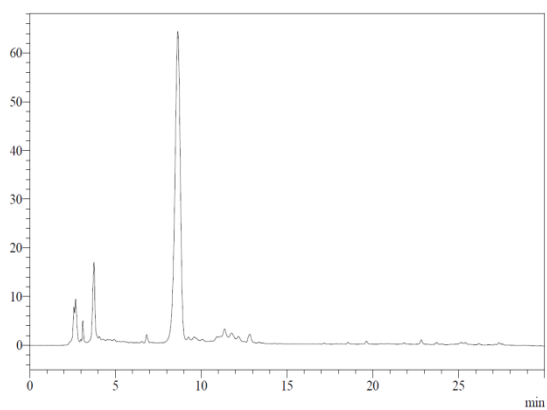
## ii. Maleimide deprotection and cyclization.

**Cyclic maleimido\*-Pro-Phe-DPhe-Asn-Gln-Tyr-Val-Orn-Cys\*-NH<sub>2</sub> 8.** The microwave-promoted [maleimide deprotection and cyclization] was carried out with a 100  $\mu$ M solution of the linear precursor at two scales, namely 1 mL and 10 mL solutions. HPLC analysis of the crudes (linear gradient from 30 to 100 % of B in 30 min) showed 85 and 96 % of **8** to be present in the two crudes, respectively (Figures S5a and S5b). **8** Was obtained after purification at the analytical scale (same gradient) in 58 % overall [maleimide deprotection+cyclization] and purification yield (10 mL-scale deprotection reaction). Analytical HPLC:  $t_R$  = 8.6 min; MALDI-TOF MS (THAP/CA, positive mode):  $m/z$ : 1281.4 [M+H]<sup>+</sup>; calc. for C<sub>61</sub>H<sub>80</sub>N<sub>14</sub>O<sub>15</sub>S 1280.6. MALDI-TOF MS analysis after treatment with H<sub>2</sub>O<sub>2</sub> (DHB, positive mode): 1297.4 [M+H]<sup>+</sup>; calc. for C<sub>61</sub>H<sub>80</sub>N<sub>14</sub>O<sub>16</sub>S 1296.6.

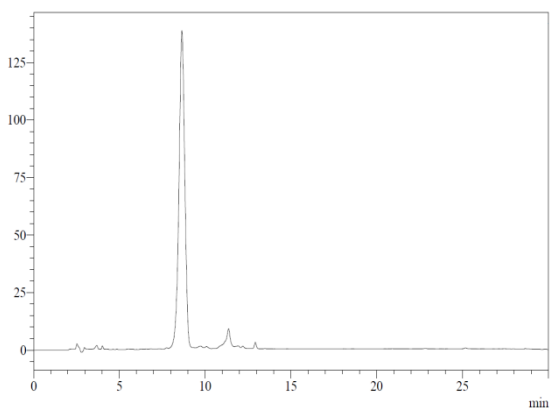
**Cyclic maleimido\*-Glu-Ala-Arg-Glu-Arg-Ile-Leu-Pro-Thr-Tyr-Asp-Cys\*-NH<sub>2</sub> 9.** The microwave-promoted [maleimide deprotection and cyclization] was carried out with a 25  $\mu$ M solution of the linear precursor (1 mL). **9** (Figure S5c, 63 % in the crude) was obtained after purification at the analytical scale in 45 % overall [maleimide deprotection+cyclization] and purification yield. Analytical HPLC (linear gradient from 20 to 40 % of B in 30 min):  $t_R$  = 11.3 min; MALDI-TOF MS (THAP/CA, positive mode):  $m/z$ : 1615.9 [M+H]<sup>+</sup>; calc. for C<sub>69</sub>H<sub>106</sub>N<sub>20</sub>O<sub>23</sub>S 1614.8. MALDI-TOF MS analysis after treatment with H<sub>2</sub>O<sub>2</sub> (DHB, positive mode): 1631.8 [M+H]<sup>+</sup>; calc. for C<sub>69</sub>H<sub>106</sub>N<sub>20</sub>O<sub>24</sub>S 1630.7.

**Cyclic H-His-Lys[maleimido\*]-Tyr-Gly-Arg-Gly-Cys\*-NH<sub>2</sub> 10.** The microwave-promoted [maleimide deprotection and cyclization] was carried out with a 100  $\mu$ M solution of the linear precursor (1 mL). **10** (77 % yield in the crude, Figure S5d) was obtained after purification at the analytical scale in 50 % overall [maleimide deprotection+cyclization] and purification yield. Analytical HPLC (linear gradient from 5 to 50 % of B in 30 min):  $t_R$  = 10.4, 10.6 min (2 diastereomers). MALDI-TOF MS (DHB, positive mode):  $m/z$ : 970.8 [M+H]<sup>+</sup>; calc. for C<sub>41</sub>H<sub>59</sub>N<sub>15</sub>O<sub>11</sub>S 969.4. MALDI-TOF MS analysis after treatment with H<sub>2</sub>O<sub>2</sub> (DHB, positive mode): 986.9 [M+H]<sup>+</sup>; calc. for C<sub>41</sub>H<sub>59</sub>N<sub>15</sub>O<sub>12</sub>S 985.4.

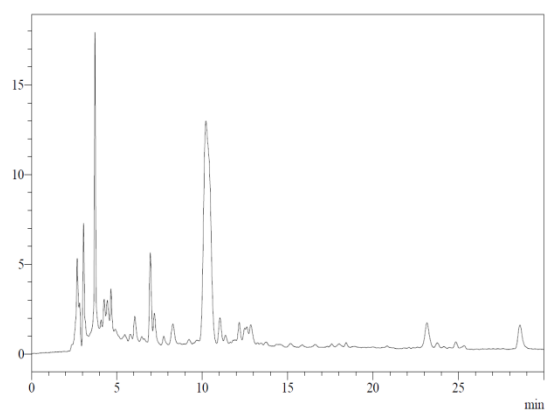
a)



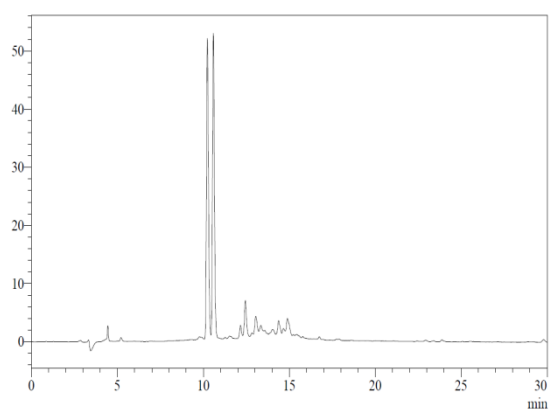
b)



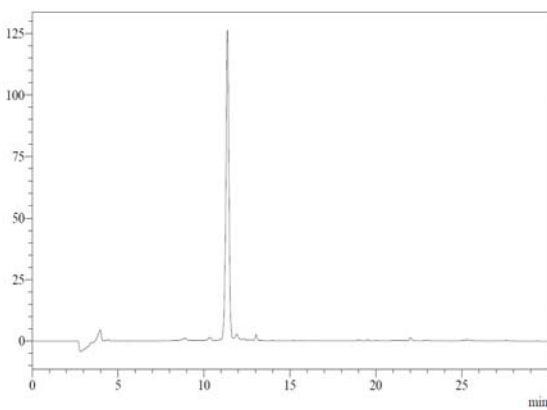
c)



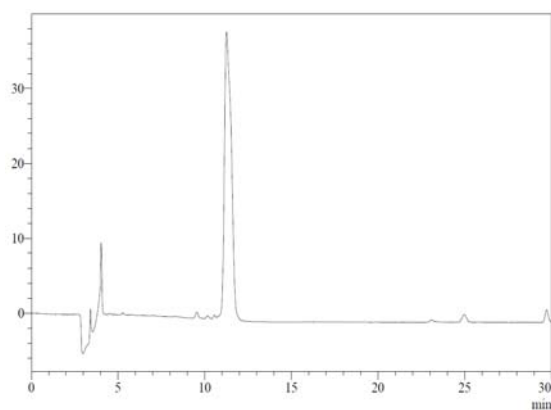
d)



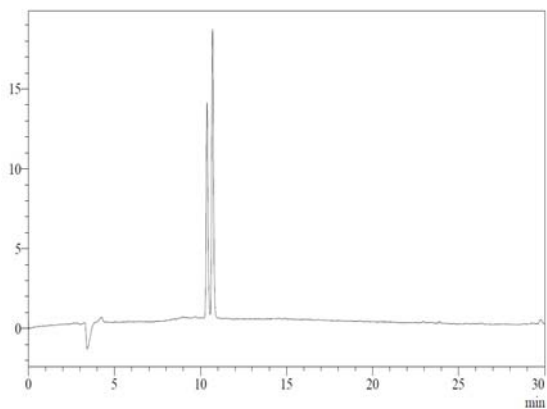
e)



f)



g)



**Figure S5.** Crude cyclic peptides after microwave-promoted maleimide deprotection and cyclization: **8** (1 mL scale, profile a, and 10 mL scale, profile b), **9** (profile c) and **10** (profile d). The target cyclic peptide is the main product in the crude in all cases. Profiles e, f and g: purified **8**, **9** and **10**, respectively.



## 4. Bicyclic peptides.

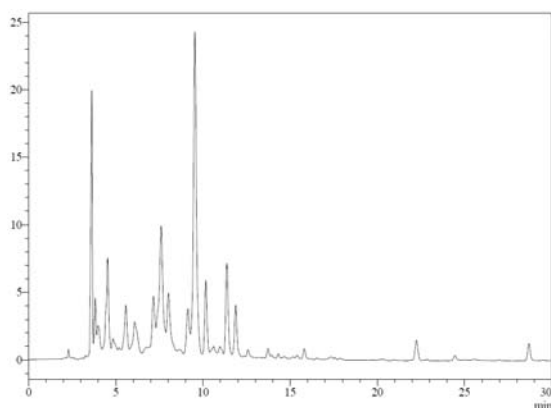
### i. Monocyclic precursors.

**Monocyclic Maleimido\*-Ala-Tyr-Ser-Cys\*-Gly-Gly-Cys(S-tBu)-Ala-Tyr-Ser-Lys[protected maleimide]-NH<sub>2</sub> 13.** Peptide-resin **12** was treated with the 95:2.5:2.5 TFA/H<sub>2</sub>O/TIS mixture, which afforded **13** (Figure S6a). **13** Was purified at the semipreparative scale (linear gradient from 30 to 60 % of B in 30 min) and obtained in 4 % overall yield (synthesis and purification). Analytical HPLC (linear gradient from 30 to 60 % of B in 30 min):  $t_R = 9.6$  min; ESI MS (positive mode):  $m/z$ : 1594.9 [M+H]<sup>+</sup>, 1616.9 [M+Na]<sup>+</sup>; calc. for C<sub>70</sub>H<sub>95</sub>N<sub>15</sub>O<sub>22</sub>S<sub>3</sub> 1593.6.

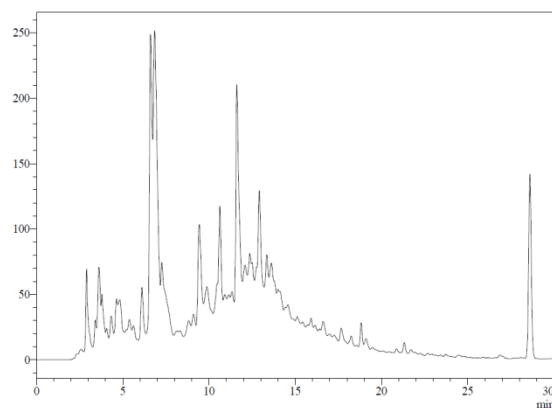
**Monocyclic Maleimido\*-Ser-Lys-Tyr-Cys(S-tBu)-Gly-Gly-Cys\*-Ala-Tyr-Ser-Lys[protected maleimide]-NH<sub>2</sub> 16.** Peptide-resin **15** was treated with the 95:2.5:2.5 TFA/H<sub>2</sub>O/TIS mixture, which afforded **16** (Figure S6b). **16** was purified at the semipreparative scale (linear gradient from 30 to 60 % of B in 30 min) and obtained in 11 % overall yield (synthesis and purification). Analytical HPLC (linear gradient from 30 to 60 % of B in 30 min):  $t_R = 6.6, 6.8$  min; ESI MS (positive mode):  $m/z$ : 1652.0 [M+H]<sup>+</sup>, calc. for C<sub>73</sub>H<sub>102</sub>N<sub>16</sub>O<sub>22</sub>S<sub>3</sub> 1650.7.

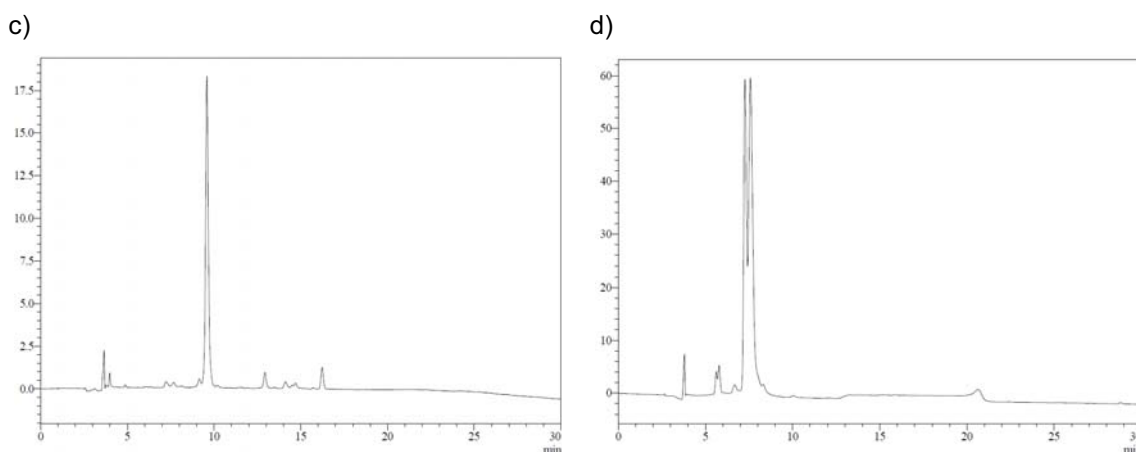
The structure of these two compounds (**13** and **16**) could not be confirmed by mass spectrometric analysis after reaction with H<sub>2</sub>O<sub>2</sub>. Reaction of H<sub>2</sub>O<sub>2</sub> with the disulfide of Cys(S-tBu) gave a complicated mixture of products (data not shown), so that the data were not conclusive.

a)



b)





**Figure S6.** Crude peptides **13** (a) and **16** (b). The target peptides are the main peaks of each profile ( $t_R \sim 9.5$  min and 2 peaks with  $t_R \sim 6.7$  min, respectively). Profiles c and d correspond to purified **13** and **16**, respectively.

## ii. Cys(S-*t*Bu) deprotection.

The peptide (**13**, **16**) was dissolved in H<sub>2</sub>O to yield a 0.3 mM solution, to which tris-(carboxyethyl)phosphine (500 eq) was added. The pH was adjusted to 5 with aqueous diluted NaOH and the mixture was stirred at room temperature for 4 h. Water was removed by lyophilization and the crude was analyzed by HPLC.

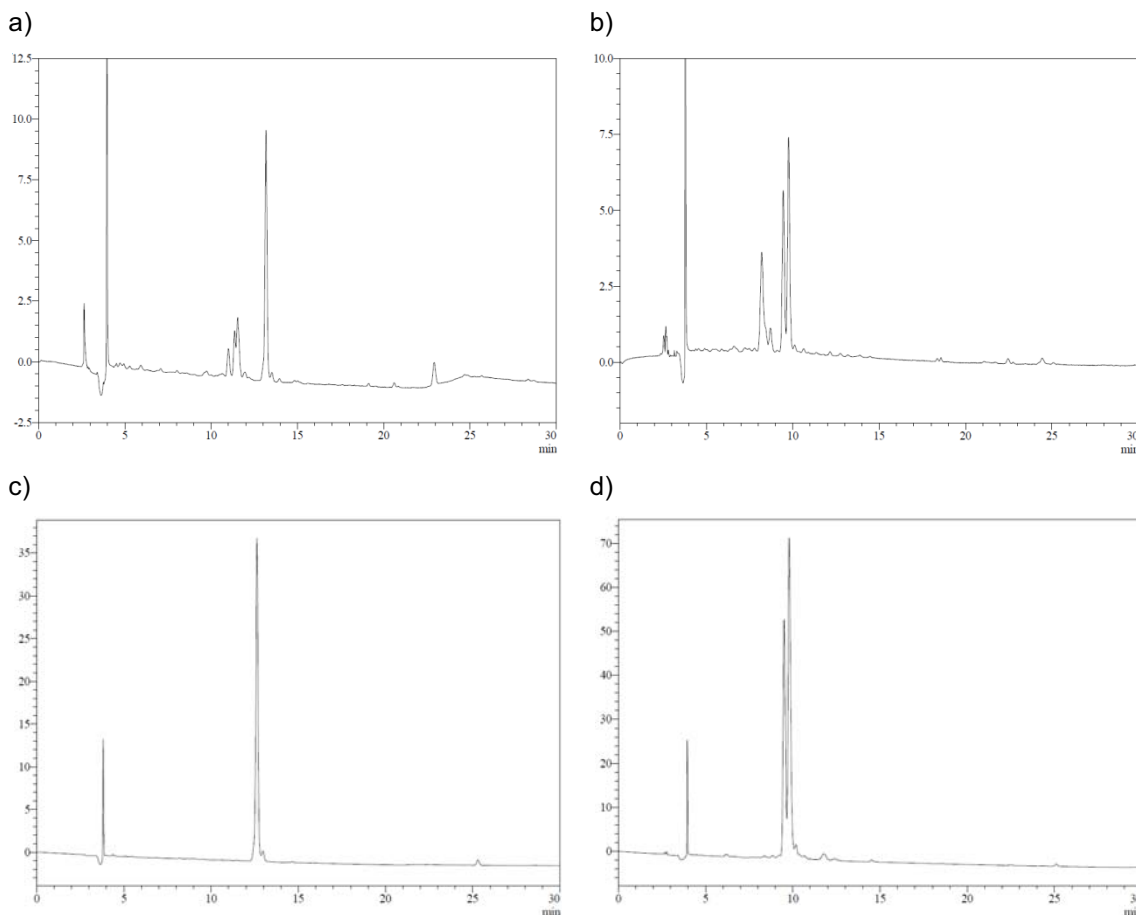
**Monocyclic maleimido\*-Ala-Tyr-Ser-Cys\*-Gly-Gly-Cys-Ala-Tyr-Ser-Lys[protected maleimide]-NH<sub>2</sub>.** 65 % yield in the crude (HPLC trace, Figure S7a). This peptide intermediate was purified at the semipreparative scale (linear gradient from 30 to 60 % of B in 30 min) and obtained in 5 % overall yield (thiol deprotection and purification). Analytical HPLC (linear gradient from 20 to 50 % of B in 30 min):  $t_R = 13.2$  min; ESI MS (positive mode):  $m/z$ : 1506.6 [M+H]<sup>+</sup>; calc. for C<sub>66</sub>H<sub>87</sub>N<sub>15</sub>O<sub>22</sub>S<sub>2</sub> 1505.6.

MALDI-TOF MS analysis after treatment with H<sub>2</sub>O<sub>2</sub> (DHB, negative mode): 1473.0 [M-furan-H]<sup>-</sup>; calc. for C<sub>66</sub>H<sub>89</sub>N<sub>15</sub>O<sub>26</sub>S<sub>2</sub> 1569.5.

**Monocyclic maleimido\*-Ser-Lys-Tyr-Cys-Gly-Gly-Cys\*-Ala-Tyr-Ser-Lys[protected maleimide]-NH<sub>2</sub>.** 73 % yield (HPLC trace, Figure S7b). This peptide intermediate was purified at the semipreparative scale (linear gradient from 20 to 50 % of B in 30 min) and obtained in 35 % overall yield (thiol deprotection and purification). Analytical HPLC

(linear gradient from 20 to 50 % of B in 30 min):  $t_R = 9.5, 9.8$  min (2 diastereomers); ESI MS (positive mode):  $m/z$ : 1563.8  $[M+H]^+$ ; calc. for  $C_{69}H_{94}N_{16}O_{22}S_2$  1562.6.

MALDI-TOF MS analysis after treatment with  $H_2O_2$  (DHB, negative mode): 1530.0  $[M-furan-H]^-$ ; calc. for  $C_{69}H_{96}N_{16}O_{24}S_2$  1626.6.



**Figure S7.** Crude monocyclic maleimido\*-Ala-Tyr-Ser-Cys\*-Gly-Gly-Cys-Ala-Tyr-Ser-Lys[protected maleimide]-NH<sub>2</sub> (a), and monocyclic maleimido\*-Ser-Lys-Tyr-Cys-Gly-Gly-Cys\*-Ala-Tyr-Ser-Lys[protected maleimide]-NH<sub>2</sub> (b) (the target peptides are the main peaks of each profile, respectively). The HPLC profiles of the purified products are shown in c and d, respectively.

### iii. Maleimide deprotection and formation of the second cycle.

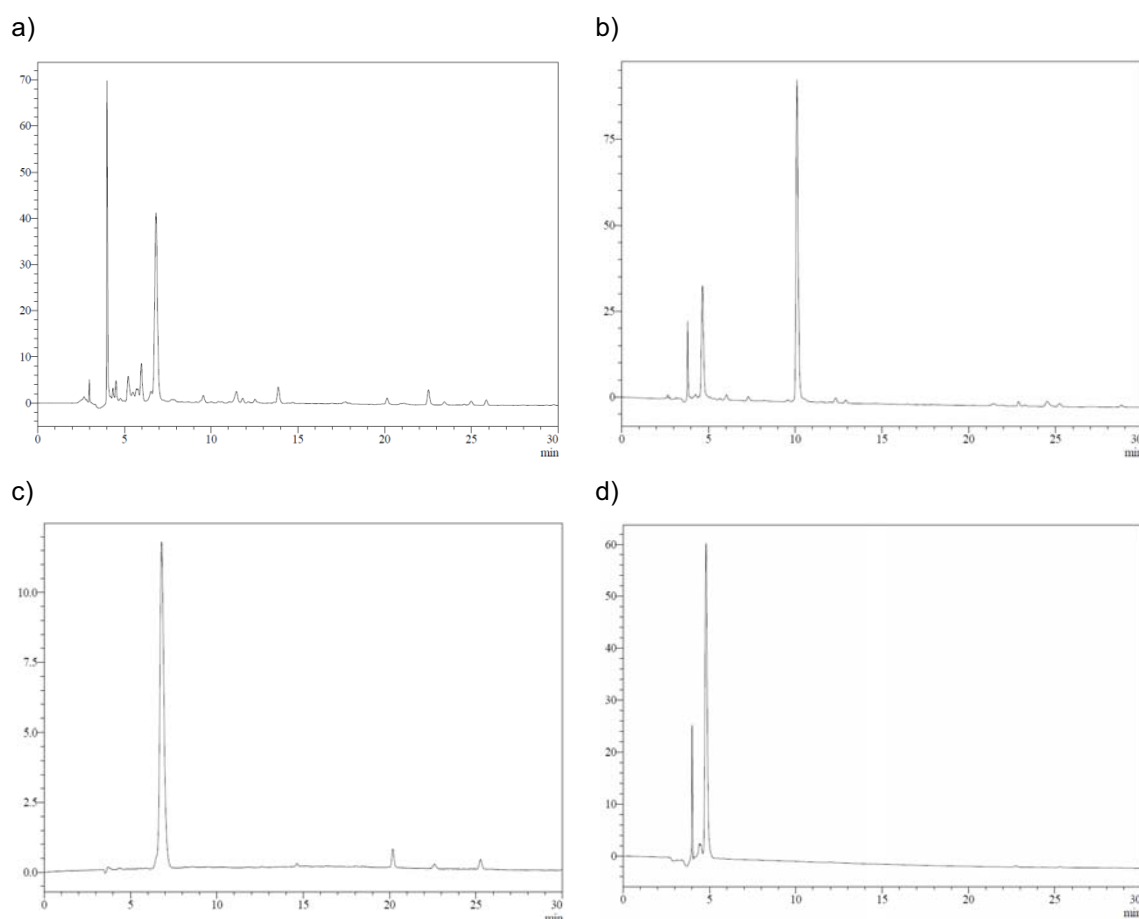
**Bicyclic maleimido\*-Ala-Tyr-Ser-Cys\*-Gly-Gly-Cys<sup>&</sup>-Ala-Tyr-Ser-Lys[maleimido<sup>&</sup>]-NH<sub>2</sub> 14** (groups labeled with \* are linked and form a cycle, and the same applies to groups labeled with <sup>&</sup>). Peptide **14** was obtained by heating cyclic maleimido\*-Ala-Tyr-Ser-Cys\*-Gly-Gly-Cys-Ala-Tyr-Ser-Lys[protected maleimide]-NH<sub>2</sub> (25  $\mu$ M in 1:1

MeOH/H<sub>2</sub>O, 1 mL) (73 % yield, HPLC trace, Figure S8a). Analytical HPLC (linear gradient from 20 to 50 % of B in 30 min):  $t_R = 6.8$  min; ESI MS (positive mode):  $m/z$ : 1410.3 [M+H]<sup>+</sup>; calc. for C<sub>60</sub>H<sub>79</sub>N<sub>15</sub>O<sub>21</sub>S<sub>2</sub> 1409.5.

MALDI-TOF MS analysis after treatment with H<sub>2</sub>O<sub>2</sub> (THAP/CA, positive mode): 1465.2 [M+Na]<sup>+</sup>; calc. for C<sub>60</sub>H<sub>79</sub>N<sub>15</sub>O<sub>23</sub>S<sub>2</sub> 1441.5.

**Bicyclic maleimido<sup>®</sup>-Ser-Lys-Tyr-Cys<sup>®</sup>-Gly-Gly-Cys<sup>®</sup>-Ala-Tyr-Ser-Lys[maleimido<sup>®</sup>]-NH<sub>2</sub> 17.** Peptide **17** was obtained by reacting cyclic maleimido<sup>®</sup>-Ser-Lys-Tyr-Cys-Gly-Gly-Cys<sup>®</sup>-Ala-Tyr-Ser-Lys[protected maleimide]-NH<sub>2</sub> (25 μM in 1:1 MeOH/H<sub>2</sub>O, 1 mL) (25 % yield, HPLC trace, Figure S8b). Analytical HPLC (linear gradient from 20 to 50 % of B in 30 min):  $t_R = 4.7$  min; MALDI-TOF MS (THAP/CA, positive mode):  $m/z$ : 1468.2 [M+H]<sup>+</sup>; calc. for C<sub>63</sub>H<sub>86</sub>N<sub>16</sub>O<sub>21</sub>S<sub>2</sub> 1466.6.

MALDI-TOF MS analysis after treatment with H<sub>2</sub>O<sub>2</sub> (THAP/CA, positive mode): 1500.2 [M+H]<sup>+</sup>; calc. for C<sub>63</sub>H<sub>86</sub>N<sub>16</sub>O<sub>23</sub>S<sub>2</sub> 1498.5.



**Figure S8.** Profiles a and b: crude bicyclic peptides **14** and **17** (the target peptides are the peaks at 6.8 and 4.7 min, respectively). Profiles c and d: purified **14** and **17**.

## 5. Abbreviations.

**Abbreviations:** ACN=acetonitrile, Boc=*tert*-butoxycarbonyl, CA=ammonium citrate, DCM=dichloromethane, DHB=2,5-dihydroxybenzoic acid, DIPC=*N,N'*-diisopropylcarbodiimide, DMF=*N,N*-dimethylformamide, ESI=electrospray ionization, Fmoc=9-fluorenylmethoxycarbonyl, HOBt=1-hydroxybenzotriazole, MALDI-TOF=matrix assisted laser desorption ionization – Time of flight, Pbf=2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl, TCEP=tris(carboxyethyl)phosphine, TFA=trifluoroacetic acid, THAP=2,4,6-trihydroxyacetophenone, TIS=triisopropylsilane, Trt=trityl.

# **PAPER III**

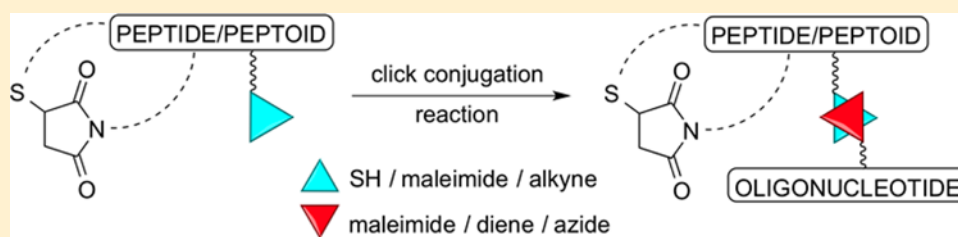


# Orthogonal Protection of Peptides and Peptoids for Cyclization by the Thiol–Ene Reaction and Conjugation

Xavier Elduque, Enrique Pedroso, and Anna Grandas\*

Departament de Química Orgànica i IBUB, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1-11, 08028 Barcelona, Spain

**S** Supporting Information



**ABSTRACT:** Cyclic peptides and peptoids were prepared using the thiol–ene Michael-type reaction. The linear precursors were provided with additional functional groups allowing for subsequent conjugation: an orthogonally protected thiol, a protected maleimide, or an alkyne. The functional group for conjugation was placed either within the cycle or in an external position. The click reactions employed for conjugation with suitably derivatized nucleoside or oligonucleotides were either cycloadditions (Diels–Alder, Cu(I)-catalyzed azide–alkyne) or the same Michael-type reaction as for cyclization.

## INTRODUCTION

With respect to the linear counterparts, conformationally constrained peptides offer increased bioavailability and affinity for the target.<sup>1</sup> Cell uptake has been shown to increase if cyclization is accompanied by the formation of internal H-bonds, which reduces the cost of amide bond desolvation.<sup>2,3</sup> Suitably derivatized cyclic peptides can be used as scaffolds (cyclic decapeptide RAFT<sup>4</sup> is an example) for the assembly of different moieties (epitopes, drugs, ligands, labels, etc.) at predetermined positions, affording biomolecules with a degree of structural organization much higher than the one accessible through derivatization of a linear peptide. Furthermore, cyclic peptides hold promise to interfere with protein–protein or protein–RNA interactions (refs 5–7 are some examples of recent publications on these subjects).

A survey of the latest literature shows that the synthesis of conjugates of cyclic peptides is an active subject of research. In a large number of papers, the affinity of  $\alpha_v\beta_3$  integrin<sup>8</sup> for RGD-containing cyclic peptides<sup>9</sup> (RGD = ArgGlyAsp) is exploited to develop new methodologies both for the selective targeting and the imaging of tumor cells/tissues and vessels expressing that cell adhesion protein. Additionally, conjugates of cyclic peptides with biological interest (immunosuppressive, epitope mimicking, antimicrobial, antifungal, or cell penetrating, among others) or nanotube precursors have been prepared to improve pharmacological properties<sup>10–15</sup> and to investigate new materials.<sup>16,17</sup>

Amide-linked units form most of these peptides, while some incorporate unnatural linkages. As to the conjugation chemistry, formation of peptide bonds and the Michael-type reaction between a maleimide and a thiol are the most

common alternatives, but other “click” reactions, and particularly the Cu(I)-catalyzed azide–alkyne cycloaddition, are becoming increasingly frequent.

There are very few reports on the synthesis and application of [cyclic peptide]–oligonucleotide conjugates. Hamilton and co-workers<sup>18</sup> linked cyclic tetrapeptides to guanine-rich oligonucleotides that assembled onto a parallel tetraplex. The distribution of the peptide ligands on the vertexes of the top tetrad provided an  $\alpha$ -chymotrypsin inhibitor that effectively interacted with the protein surface. In a related approach, the covalent attachment of peptides to the 5' ends of two guanine-rich oligonucleotides provided, in the presence of potassium cations, a parallel G-quadruplex with two loop mimics on its surface.<sup>19</sup> Cyclic peptides have been covalently linked to the sense strand of a siRNA duplex, which facilitated transport of the oligonucleotide through the cell membrane and inhibited gene expression.<sup>20</sup> More recently, disulfide-containing citrullinated peptides have been linked to various oligonucleotides and immobilized on plates for diagnostic ELISA assays.<sup>21</sup>

In this paper, we describe methodology allowing for the straightforward synthesis of [cyclic peptide]–oligonucleotide conjugates. Model peptides (and a peptoid) cyclized using the Michael-type maleimide–thiol reaction<sup>22</sup> and incorporating a functional group allowing for subsequent conjugation were prepared. Then, oligonucleotides were linked to different positions (internal or external) of the cyclic polyamide using various click reactions, namely the Diels–Alder and Cu(I)-

Received: November 5, 2013

Published: March 11, 2014



catalyzed Huisgen cycloadditions or, again, the Michael-type thiol–ene. We believe that easy access to [cyclic peptide]–oligonucleotide conjugates will pave the way for new studies aiming at assessing whether cyclic peptides facilitate the cell uptake of oligonucleotides and promote interactions with cell machineries different from those driven by the linear oligomers. To our knowledge, whether the gene silencing or splice-switching properties of oligonucleotides linked to cyclic peptides differ from those of [linear peptide]–oligonucleotide conjugates has never been assessed.

## RESULTS

As recently described by the group,<sup>22</sup> peptides can be cyclized by reaction between a maleimide and a thiol. Cyclic peptides can be obtained from linear precursors after the TFA treatment that deprotects the peptide chain following stepwise elongation on a solid matrix, using S-trityl-protected cysteine and a maleimide moiety (Scheme 1A). Since the maleimide moiety is labile to nucleophiles such as piperidine, which is used at every synthesis cycle to remove the Fmoc temporary protecting group, it can only be installed at the *N*-terminus. For introduction at any other position of the chain the maleimide has to be protected, for which purpose an Fmoc-lysine derivative with 2,5-dimethylfuran-protected 3-maleimidopropanoic acid linked to the  $\epsilon$ -amine can be used.<sup>23</sup> In this case, the cyclic peptide is not obtained after the reaction with TFA but after the thermal treatment that deprotects the maleimide.

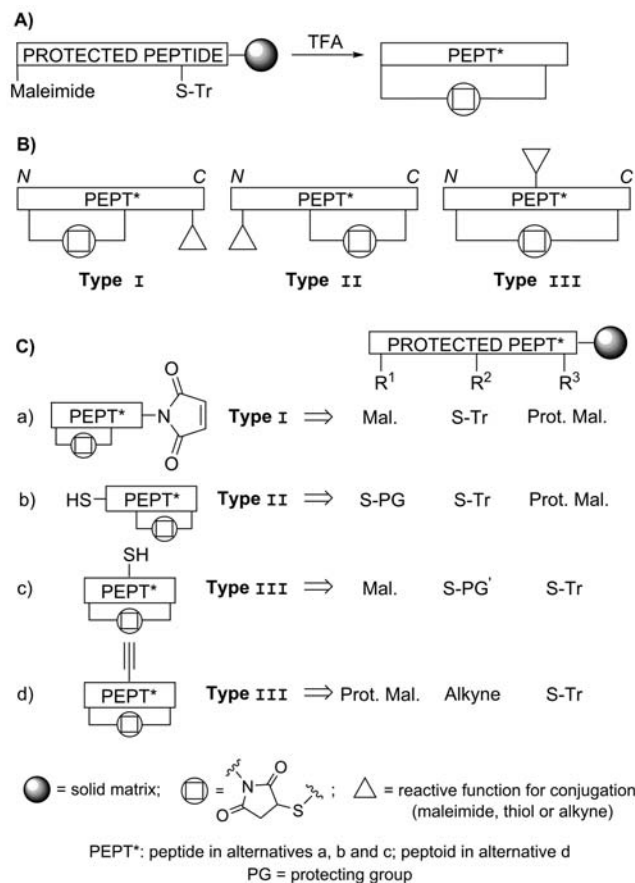
Cyclic peptides can be derivatized for conjugation in three different ways (see Scheme 1B). In cyclic peptides types I and II, the functional group for conjugation (depicted as a triangle) is outside the cycle, and the cycle involves, respectively, amino acids placed at either the *N*-terminus of the peptide chain (the group for conjugation is then placed at the *C*-terminus) or the *C*-terminus (and the group for conjugation is at the *N*-terminus). In type III cyclic peptides the group for conjugation is within the cycle.

Linear precursors incorporating a free and a protected maleimide are needed to synthesize maleimido-derivatized cyclic peptides. The free maleimide, which will form the cycle, can only be placed at the *N*-terminal end, while the protected maleimide can occupy any other position. As stated above, if the protected maleimide is closer to the *C*-terminus than the thiol involved in cyclization, the result is a type I cyclic peptide (Scheme 1C, alternative a). These peptides can be conjugated with either thiol- or diene-derivatized molecules.

The precursors of cyclic peptides derivatized with thiols must incorporate two orthogonally protected cysteine residues. The cysteine residue involved in cyclization can be protected with the trityl group (Scheme 1C, alternatives b and c), but the best option to protect the cysteine residue involved in conjugation depends on its position within the chain (see below). Both thiol-derivatized cyclic peptides types II and III have been prepared and conjugated with a maleimido-containing compound.

The use of different reactions for cyclization and conjugation is less complicated from a chemical point of view and has been evaluated with a peptoid chain (peptoid units are *N*-alkylglycines). Even though the thiol–ene reaction is probably much quicker than the reduction of azides by thiols,<sup>24,25</sup> we preferred not to expose the azide to the presence of a free thiol. Hence, an azide-containing molecule was used for conjugation with the alkyne-derivatized, type III

**Scheme 1.** (A) General Procedure for the Synthesis of Peptides Cyclized Using the Thiol–Maleimide Reaction. (B) General Structure of the Three Types of Cyclic Peptides (or Peptoid) Prepared in This Work, Derivatized for Conjugation.<sup>a</sup> (C) Schematic Structure of the Cyclic Molecules Showing the Functional Group for Conjugation and Combinations of Functional Groups ( $R^1$ ,  $R^2$ , and  $R^3$ , Suitably Protected when Required) on the Resin-Linked Peptide/Peptoid Precursor for Each of the Alternatives Tested (a–d)<sup>b</sup>



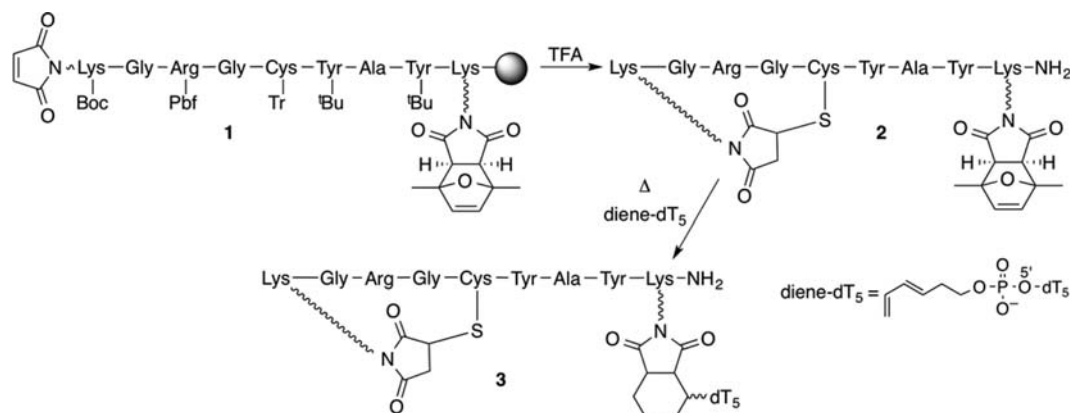
<sup>a</sup>Cyclization was accomplished by a Michael-type maleimide-thiol reaction in all cases. <sup>b</sup>Two of those groups had to react with each other and generate the cycle, and the third provided the functionality for conjugation with another molecule (in this work, nucleoside or oligonucleotide). Their relative position within the chain determined the type of cyclic molecule that was obtained (I, II or III). Abbreviations: Fm = 9-fluorenylmethyl; Mal = maleimide; Prot mal = (2,5-dimethylfuran)-protected maleimide; Tr = trityl.

cyclic peptide. Cyclization can be carried out with no risk in the presence of an alkyne (Scheme 1C, alternative d), which will not react with either the thiol or the maleimide.

Peptides were prepared by solid-phase synthesis using the standard Fmoc/<sup>t</sup>Bu methodology. All were assembled on the Rink amide MBHA (*p*-methylbenzhydrylamine) resin and were thus obtained as *C*-terminal carboxamides.

Acylation with 3-maleimidopropanoic acid after chain elongation allowed a free maleimide to be attached to the peptide *N*-terminus.

These compounds were conjugated with suitably derivatized oligonucleotides (or nucleoside), as shown in Schemes 2–5 in more detail. Diene- and maleimido-containing oligonucleo-

Scheme 2. Synthesis of Type I Cyclic Peptide 2 and Conjugation with Oligonucleotide dT<sub>5</sub><sup>a</sup>

<sup>a</sup>Abbreviations: Pbf = 2,2,4,6,7-pentamethylidihydrobenzofuran-5-sulfonyl.

tides (dT<sub>5</sub>, dT = 2'-deoxythymidine) were synthesized following the described procedures.<sup>26,27</sup>

Schemes 2–4 illustrate in more detail the different steps required to obtain cyclic peptides types I to III and the corresponding conjugation products.

**Synthesis and Conjugation of a Type I Cyclic Peptide.** In the first alternative (aimed to provide type I cyclic peptides, a in Scheme 1B), the fully protected peptide precursor (1) incorporated two maleimides, one free (at the *N*-terminus) and one protected (at the *C*-terminus), and a trityl-protected cysteine residue (Scheme 2). Therefore, the TFA treatment that was carried out after assembly of the linear precursor removed all protecting groups except that on the *C*-terminal maleimide and, as previously described,<sup>22</sup> furnished the partially protected cyclic peptide 2. Maleimide thermal deprotection (retro-Diels–Alder reaction) was performed subsequently in the presence of 5'-diene-dT<sub>5</sub>, which afforded the target conjugate 3.

**Synthesis and Conjugation of a Type II Cyclic Peptide.** The goal of the second alternative (b in Scheme 1B) was to obtain a peptide with a free thiol group for conjugation at the *N*-terminus of the chain (type II cyclic peptide). This required a precursor with two differently protected cysteines and a protected maleimide (Scheme 3). The cysteine that had to participate in cyclization was protected with the TFA-labile trityl group, and it was necessary to assess which thiol protecting group was suitable for the other cysteine. This group had to remain stable during the treatments required to obtain the cycle, that is, TFA treatment and maleimide deprotection.

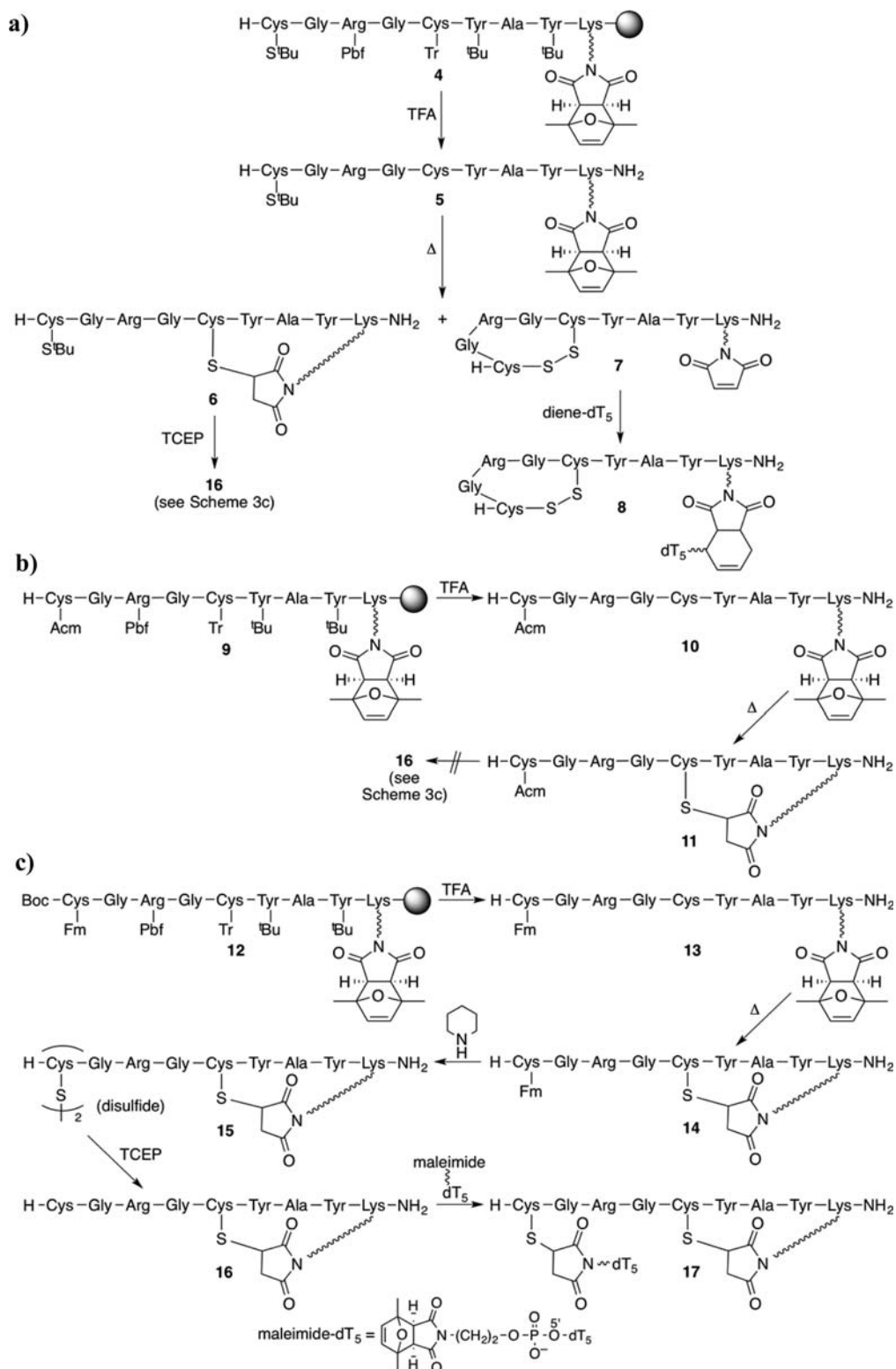
We first evaluated the possibility of protecting the side chain of the *N*-terminal cysteine with the *S*<sup>t</sup>Bu and acetamidomethyl groups, whose Fmoc derivatives are commercially available and which are TFA-stable (Scheme 3a,b). In preliminary experiments, both Fmoc-Cys(*S*<sup>t</sup>Bu)-OH and Fmoc-Cys(Acm)-OH proved to be stable to maleimide deprotection conditions (heating in a microwave oven a solution of the compound in 1:1 MeOH/H<sub>2</sub>O for 90 min at 90 °C). Yet, when peptide 5 was submitted to these reaction conditions, two products were obtained (Scheme 3a): the target compound (6) with the thiosuccinimide-containing cycle and a peptide (7) with a disulfide bridge and a free maleimide (the 6:7 ratio was 1:1 when the peptide was heated for 30 min, and 3:2 when the reaction was extended to the 90 min typically required for full maleimide deprotection<sup>27</sup>).

Mass spectrometric analysis confirmed presence of the *S*<sup>t</sup>Bu group in 6, its loss after reaction with tris(2-carboxyethyl)-phosphine (TCEP) to yield 16 (see structure in Scheme 3c), and loss of the three *N*-terminal amino acids when 6 was digested with thrombin. Instead, the mass of 7 was lower than that of 6, and the product remained undigested upon treatment with thrombin. Final confirmation of the presence of a free maleimide in 7 was obtained from reaction with diene-dT<sub>5</sub>, which afforded the [disulfide cyclic peptide]–oligonucleotide conjugate 8. As expected, the Diels–Alder reaction is fully compatible with the presence of the disulfide. Even though the thiol–maleimide reaction slightly predominated over disulfide formation, the relative amounts of the products obtained after maleimide deprotection are too similar to consider this alternative a good synthetic method.

In the second group of assays (Scheme 3b), the *N*-terminal cysteine was protected with the acetamidomethyl (Acm) group. Here the problem was that none of the conditions tested for its deprotection (reactions with either Ag<sup>+</sup>, iodine, or TI<sup>3+</sup>) provided clean crudes with the target cyclic peptide (16) as the main product (see the Experimental Section).

We next evaluated the combination of protecting groups shown for peptide-resin 12 (Scheme 3c), using commercially available *N*-Boc-Cys(Fm)-OH. In this case, treatment with TFA deprotected the *N*-terminus and the other functional groups, except those on the *N*-terminal cysteine and the maleimide, to yield 13. Maleimide deprotection and the thiol–maleimide reaction that furnished the cycle (14) took place simultaneously, and as described in the literature,<sup>28</sup> treatment with piperidine removed the Fm protecting group and provided the disulfide (15). The cyclic peptide with the free thiol (16) was obtained after reduction with TCEP, and subsequent reaction with maleimido-dT<sub>5</sub> afforded conjugate 17. Hence, the 9-fluorenylmethyl group is fully compatible with maleimide deprotection and peptide cyclization by the Michael-type thiol–ene reaction. Even though removal of the Fm group is accompanied by thiol to disulfide oxidation, which adds a reduction step, this group is suitable to mask the cysteine side chain provided that the so-protected cysteine is placed at the *N*-terminus. Peptide elongation using the Fmoc/*S*<sup>t</sup>Bu approach rules out any other possibility.

**Synthesis and Conjugation of Type III Cyclic Polyamides (Peptide and Peptoid).** Two different type III cyclic polyamides were synthesized (alternatives c and d in Scheme 1B). Here the functional group for conjugation

Scheme 3. Assays Carried out To Obtain Cyclic Peptide 16 (Type II, Alternative b in Scheme 1) Using Different Thiol Protecting Groups for the *N*-Terminal Cysteine (the One To Be Involved in Conjugation)<sup>a</sup>

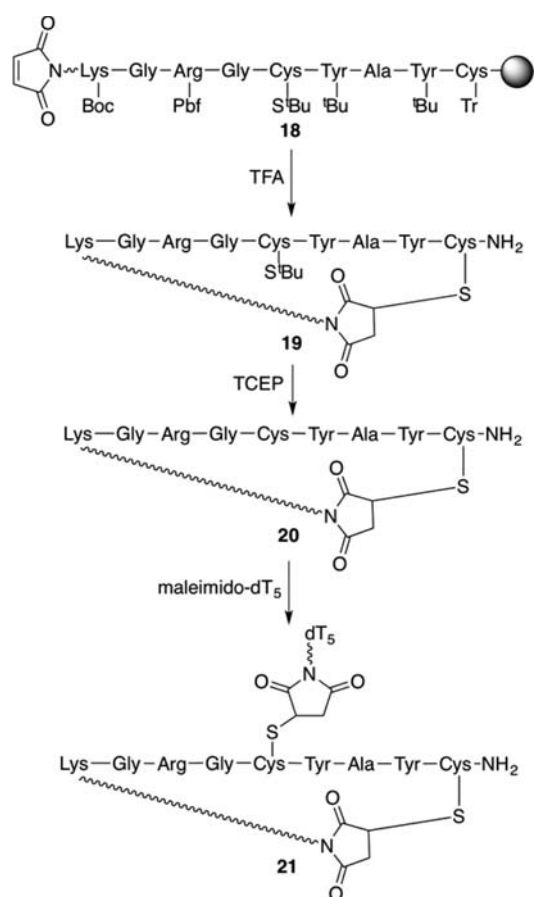
<sup>a</sup>(a) Deprotection of the maleimide of peptide 5 afforded two products, 6 and 7, in a similar ratio. The latter furnished conjugate 8 by reaction with diene-dT<sub>5</sub>. (b) Removal of the Acm cysteine protecting group from 11 yielded complex mixtures of products in all the conditions tested. (c) Cyclic peptide 16 was prepared by combining Tr and Fm as cysteine protecting groups and conjugated with maleimido-dT<sub>5</sub>. Abbreviations: Acm = acetamidomethyl; Fm = 9-fluorenylmethyl.

appended from a residue placed within the cycle. In one case, the polyamide was a peptide and the functional group for

conjugation a thiol (20), while in the other it was a peptoid derivatized with an alkyne (24).

Scheme 4 depicts the preparation and conjugation of cyclic peptide **20**. The C-terminal cysteine was protected with the

**Scheme 4. Steps Involved in the Preparation of Cyclic Peptide **20** and Conjugate **21****



trityl group, which reacted with the N-terminal, free maleimide upon treatment with TFA. Under these conditions, the free thiol gave the Michael-type reaction and left untouched the internal, S<sup>t</sup>Bu-protected cysteine residue. In other words, the free thiol reacted much faster with the maleimide than with the S<sup>t</sup>Bu-protected cysteine. The side chain of the internal cysteine of the cyclic peptide (**19**) was deprotected by reaction with TCEP and, after purification, reacted with 5'-maleimido-dT<sub>5</sub> to yield conjugate **21**.

Scheme 5 shows the preparation of a cyclic peptoid and the structure of the derived conjugate. On-resin assembly of the peptoid was carried out using the submonomer procedure,<sup>29</sup> also on the Rink amide MBHA resin. As shown in the upper part of Scheme 5, chain elongation involves two subsequent steps, namely acylation with bromoacetic acid and reaction with an amine. Three of the amines required to synthesize peptoid **23** were commercially available (see the list in the Supporting Information), and S-Tr-protected cysteamine and N-Boc-1,2-diaminoethane were prepared following described procedures.<sup>30,31</sup> Cyclization with an N-terminal maleimide required incorporation of a maleimide-containing monomer, and 3-maleimidopropanoic acid is the cheapest alternative. However, since we had observed that the quality of cyclic peptide crudes was a bit lower when cyclization took place upon treatment with TFA,<sup>22</sup> we incorporated the 2,5-dimethylfuran-protected maleimidopropanoic acid.<sup>27</sup> To in-

roduce maleimide moieties at other positions of the chain, 2-[(2,5-dimethylfuran-protected) maleimide]-1-aminoethane<sup>23</sup> should be employed. In the absence of two orthogonal maleimide protecting groups, this precludes use of maleimides for subsequent conjugation.

Chain assembly furnished the fully protected peptoid **22**, from which all protecting groups except that on the maleimide were removed by treatment with TFA. Maleimide deprotection and cyclization took place simultaneously to afford **24**. Finally, the Cu(I)-catalyzed Huisgen reaction between cyclic peptoid **24** and AZT (2',3'-dideoxy-3'-azidothymidine) provided the target conjugate (**25**). The azide-containing molecule was in this case a simple nucleoside analogue, but the high success of the Cu(I)-catalyzed alkyne-azide cycloaddition has prompted many groups to develop methodology for the preparation of azide-derivatized oligonucleotides (see, for instance, refs 32–34) circumventing the incompatibility between azides and phosphite-based reagents.<sup>35,36</sup>

## DISCUSSION

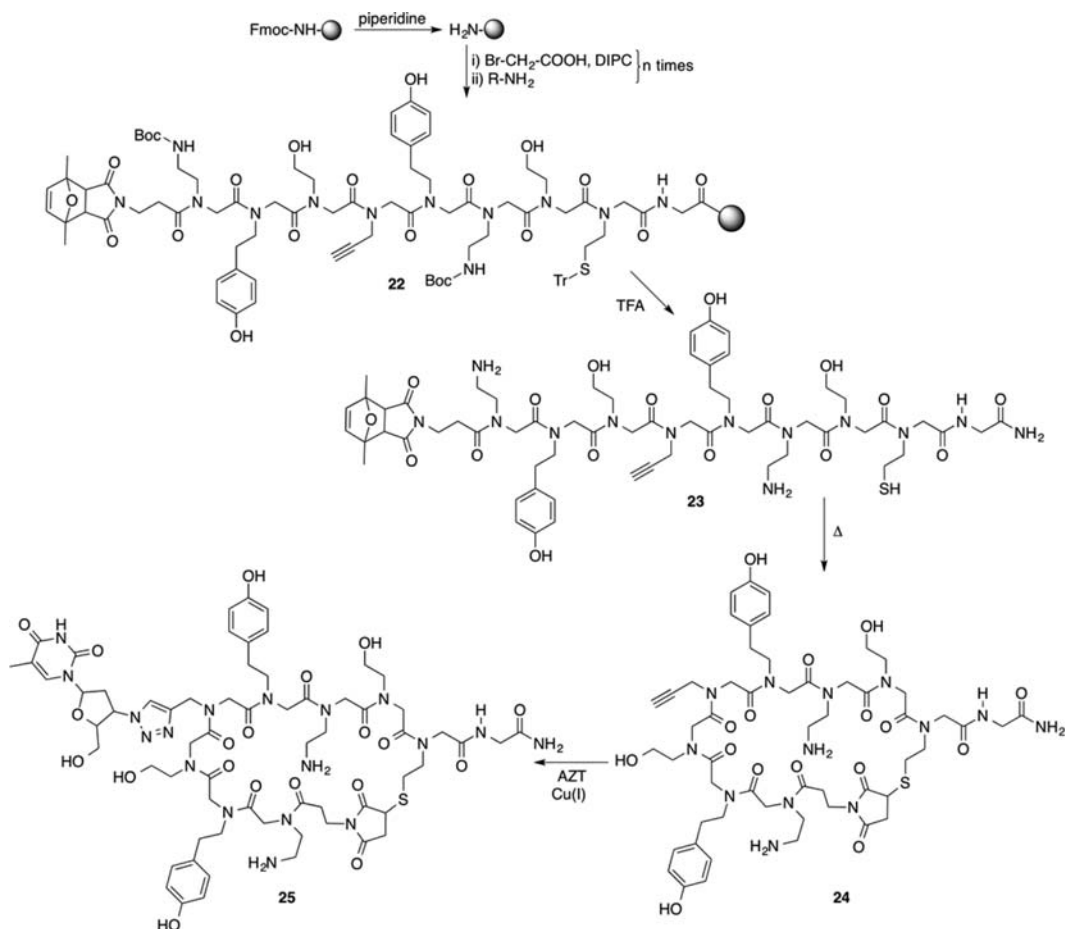
We have shown that, by adequately combining differently protected cysteines, free and protected maleimides, and an alkyne and choosing their relative position in the chain, it is possible to easily obtain cyclic peptides derivatized with a functional group allowing for subsequent conjugation, for instance, with oligonucleotides.

In all cases, cyclization was carried out using the thiol-ene Michael-type reaction. For this purpose, in all peptide precursors one of the cysteine residues was masked with the TFA-labile trityl group, so that the acidic treatment that follows chain elongation deprotected its thiol. When a free maleimide had been appended to the N-terminus of the chain, this unmasked thiol reacted with the maleimide to form the cycle (peptides **2** and **19**) during the TFA deprotection step.<sup>22</sup> Whenever a masked maleimide was incorporated into the chain, an additional deprotection step was required. Then, thermal removal of the maleimide protecting group (retro-Diels-Alder reaction) and cyclization took place simultaneously (peptide **14** and peptoid **24**).

As required for conjugation with a suitably derivatized molecule, cyclic peptides had to incorporate an additional functional group. In this work, this group was either a maleimide, a thiol, or an alkyne.

With the cyclization strategy that we have used, the only possibility to use a maleimide for conjugation is to assemble a peptide precursor with a free maleimide at the N-terminus (to form the cycle) and a protected maleimide at any other position. Placing the lysine carrying the protected maleimide at the C-terminus (as in **1**) furnishes type I cyclic peptides, with amino acids out of the cycle and the maleimide for conjugation at the C-terminal end. It is reasonable to surmise, on the basis of the results here described, that placing that lysine within the cycle will successfully provide cyclic peptides with an internal maleimide for conjugation. Maleimide-derivatized cyclic peptides (and any maleimide-containing compound) can be conjugated with diene-derivatized molecules, as in the synthesis of **3**, where maleimide deprotection and conjugation took place simultaneously. It is also worth mentioning that, even though the proof of principle experiments here described were carried out with simple oligonucleotide sequences, the compatibility of oligonucleotides of different length and composition<sup>27</sup> with

Scheme 5. Synthesis of Cyclic Peptoid 24 and Conjugate 25



the reaction conditions in which maleimides are deprotected suggests that any oligonucleotide can be linked to cyclic peptides. Alternatively, maleimide-derivatized cyclic peptides could be conjugated with thiols.

When the cyclic peptoid is to be derivatized with a thiol for conjugation, the choice of protecting group for the cysteine providing this thiol depends on its position in the chain. If the goal is a type II cyclic peptoid, the best alternative among those we have examined is to block the thiol with the Fm group, which is compatible with maleimide deprotection and cyclization and can be removed with no formation of byproducts. The only drawback is that removal of the Fm group by reaction with piperidine is accompanied by disulfide formation, an additional reduction step being thus necessary prior to conjugation. *N*-Terminal derivatization of a peptide with a thiol was here accomplished by introducing a suitably protected cysteine residue, but an *S*-blocked mercaptoacetic acid should also work. Use of cysteine derivatives has the advantage that the amino function can be further derivatized, allowing for the attachment a fluorophore, for instance, or any reporter group.

Conversely, the synthesis of cyclic peptides with the cysteine that provides the thiol for conjugation at any other position of the chain (not at the *N*-terminus) is much more straightforward (this is the case of type III peptide 20, but the same protection scheme could provide type I peptides). That cysteine can indeed be protected with the *S*<sup>t</sup>Bu group, which remains stable during the acidic treatment that removes most of the protecting groups from the peptide and does not react

with the free thiol resulting from the deprotection of Cys(Tr) (or if it does, the extent of the reaction is negligible). As stated before, at room temperature and under these strongly acidic conditions, formation of the cycle by reaction between the free thiol and the free maleimide is much quicker. Yet (see above and Scheme 3a), this is not the case when the peptide has to be heated.

Thiol-derivatized cyclic peptides can be conjugated with maleimido-oligonucleotides, as in the case of 17 and 21.

Type III cyclic peptoids derivatized for conjugation by the Cu(I)-catalyzed Huisgen cycloaddition were also synthesized. In this case, the key monomers were 3-(2,5-dimethylfuran-protected)maleimidopropanoic acid,<sup>27</sup> *S*-tritylcysteamine,<sup>30</sup> and propargylamine. Use of *S*-Fm-cysteamine instead of propargylamine would provide peptoids cyclized by the thiol-maleimide reaction and derivatized with a thiol group for conjugation.

As to the use of alkynes for conjugation, the triple bond can be installed at any position of the polyamide chain. It is neither affected by any of the peptide/peptoid synthesis reagents nor by the thiol-maleimide cycle-forming reaction. We have prepared a [cyclic peptoid]–nucleoside conjugate to verify that synthesis proceeded as expected, and the triple bond was introduced in an internal position by reacting propargylamine with the haloacylated resin-linked, growing peptoid chain. Commercially available alkyne-containing L-amino acids or alkynoic acids allow triple bonds to be attached to peptides in a straightforward manner, which is also the case for derivatization with azides.

## CONCLUSIONS

In summary, we have broadened the scope of applications of peptides cyclized making use of the maleimide–thiol reaction by introducing an additional group that allows for conjugation with oligonucleotides (or other molecules) utilizing click reactions and shown that this methodology can also be employed with peptoids. The method here described is based on the adequate choice of protection for maleimides and thiols, as well as suitably placing maleimides, thiols, or alkynes within the peptide/peptoid chain. These synthetic alternatives could also provide other cyclic polyamide scaffolds derivatized for conjugation, such as peptide nucleic acids.

It has been reported that subtle differences in the conjugation chemistry may have considerable impact on the biological outcome.<sup>37–39</sup> Therefore, it is of interest to develop chemical tools giving access to differently linked conjugates. This manuscript describes methodology allowing oligonucleotides to be linked to cyclic peptides, either to an internal position, or to any of the two ends of a chain in which only some of the residues of the peptide are part of a chemically stable cycle. Both cyclic peptides and oligonucleotides (natural or modified) are relevant chemical goals that hold promise from the therapeutic and diagnostics point of view. So far there is lack of information as to the differential effect of linking oligonucleotides to cyclic peptides rather than the linear counterparts. Work on this subject is now in progress in the group, and results will be published in due time.

## EXPERIMENTAL SECTION

**General Methods. Peptide Synthesis.** All peptides were assembled on the Rink Amide resin, which was treated and washed with DCM (3×), DMF (3×), 20% piperidine/DMF (2 × 15 min), DCM (3×), DMF (3×), and MeOH (3×) prior to chain elongation. Incorporation of the first amino acid on the solid matrix and peptide elongation were carried out using 3 equiv of Fmoc-amino acid, HOBt·H<sub>2</sub>O, and DIPC (DIPC = *N,N'*-diisopropylcarbodiimide) dissolved in the minimum amount of DCM and a few drops of DMF (90 min), which was followed by washing with DCM, DMF, and MeOH (3×). In case the coupling was not complete, as assessed by the Kaiser test, it was repeated using 2 equiv of the reagents. Fmoc groups were removed by reaction with 20% piperidine/DMF (1 × 3 min + 1 × 10 min) followed by washing (DMF, DCM). Activation of 3-maleimidopropanoic acid was carried out with DIPC (90 min reaction time). Even though activation with DIPC and HOBt·H<sub>2</sub>O seems not to cause any harm on the maleimide, use of only the carbodiimide is safer to prevent addition of HOBt to the maleimide of maleimidopropanoic acid.

Reaction with a TFA/H<sub>2</sub>O/TIS mixture (TIS = triisopropylsilane) at room temperature cleaved the peptide from the resin and removed most permanent protecting groups (see below and Schemes 2–4 for details). Filtrate and washings (DCM) were collected and concentrated under a N<sub>2</sub> stream, and cold diethyl ether was added to the resulting oil. The mixture was centrifuged (10 min, 5 °C, 4800 rpm) and decanted. This procedure was repeated three times. Peptides were quantified by UV spectroscopy from the absorbance of tyrosine residues at 280 nm ( $\epsilon_{280} \text{ M}^{-1} \text{ cm}^{-1} = 1480$ ).

**Peptoid synthesis.** Peptoid **22** was also assembled on the Rink Amide resin, which was treated and washed as described above for peptide synthesis. For the acylation step, bromoacetic acid (25 equiv) was dissolved in anhyd. DMF in a round-bottom flask, the solution was cooled to 0 °C, and DIPC (25 equiv) was added dropwise. After 5 min, the solution was poured onto the resin in the syringe, and the mixture was stirred (shaker) under an Ar atmosphere for 45 min. The resin was filtered and washed (DCM, DMF, iPrOH, and DCM, 3× each), and the amination step was carried out by reaction with the corresponding amine (20 equiv)

dissolved in anhyd NMP (500  $\mu\text{L}$  per 100 mg of resin) for 3 h. After this reaction, the resin was washed again (DCM, DMF and MeOH, 3× each). Cleavage from the resin and removal of most permanent protecting groups was accomplished by reaction with TFA at room temperature (see details below, and Scheme 5), which was followed by washings with the cleavage mixture and DCM. Solvents and reagents were removed from the filtrate by means of a N<sub>2</sub> stream until the solution became an oil. Cold diethyl ether (10 volumes) was added, and the mixture was centrifuged (10 min, 5 °C, 4800 rpm) and decanted. The ether addition plus centrifugation procedure was repeated three times. The crude was analyzed and purified by HPLC, quantified by UV spectroscopy (phenoxy group:  $\epsilon_{270} \text{ (M}^{-1} \text{ cm}^{-1}) = 1450$ ), and characterized by mass spectrometry.

**HPLC.** Unless otherwise stated, the following conditions were used. Semipreparative HPLC, acidic conditions: Jupiter Proteo column; 3 mL/min; solvent A: H<sub>2</sub>O + 0.1% TFA; solvent B: ACN + 0.1% TFA. Analytical HPLC, acidic conditions: Phenomenex C18, 4  $\mu\text{m}$ , Jupiter Proteo column; 1 mL/min; solvent A: H<sub>2</sub>O + 0.045% TFA; solvent B: ACN + 0.036% TFA. Analytical HPLC (also used for purification at a small scale), buffered medium: Kromasil C18, 10  $\mu\text{m}$  column; 1 mL/min; solvent A: 0.05 M TEAA (triethylammonium acetate) in water; solvent B: 1:1 H<sub>2</sub>O/ACN.

**MW-Promoted Polyamide Deprotection and Cyclization.** The [protected maleimido]-containing polyamide was dissolved in a 1:1 (v/v) MeOH/H<sub>2</sub>O mixture at a concentration between 25 and 100  $\mu\text{M}$  in a microwave flask. The flask was sealed and heated in a microwave oven (Biotage Initiator™ Classic) at 90 °C for 90 min (with stirring and internal probe for temperature monitoring). Methanol was then removed under reduced pressure, and the resulting aqueous solution was analyzed by HPLC. Alternatively, methanol removal was followed by lyophilization, and the crude was redissolved in water for HPLC analysis.

**Assessment of Cyclization: Reaction with H<sub>2</sub>O<sub>2</sub>.** Formation of a cyclic product was confirmed by reacting the peptide, either linear or cyclic, with a dilute solution of H<sub>2</sub>O<sub>2</sub> (3.5%) for 1 h, and subsequent MS analysis.<sup>23</sup> When the peptide is linear, H<sub>2</sub>O<sub>2</sub> oxidizes the free thiol to sulfonic acid (and the mass increases 48 units), whereas when the peptide is cyclic the thioether is oxidized to sulfoxide (and the mass increases 16 units). This method cannot be used when the peptide to be analyzed incorporates a Cys(S<sup>t</sup>Bu) residue because this disulfide is oxidized to different products. ESI mass spectrometric analysis showed formation of mixtures of products with and without loss of the S<sup>t</sup>Bu group. When analysis was performed in the negative mode, the main group of peaks corresponded to loss of the S<sup>t</sup>Bu group and oxidation of the thiol to sulfonic acid (+ 3 oxygen atoms,  $M + 48$ ).

**Assessment of Cyclization: Digestion with Thrombin.** Reaction with thrombin can also be used to assess circularity. Thrombin cleaves the Arg-Gly bond in linear GRG peptide sequences, but not when this sequence is within a cycle. Twenty micrograms of peptide and 5 u of thrombin, dissolved in a 1:2 PBS/H<sub>2</sub>O solution (PBS buffer: 140 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub> and 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, pH = 7.23), were incubated at room temperature overnight. The crude was readily analyzed by MALDI-TOF MS.

**Deprotection of Cys(S<sup>t</sup>Bu) and Reduction of Disulfide Bridges.** TCEP (500 equiv) was added to a 0.3 mM solution of the polyamide in H<sub>2</sub>O (1–3 mL). The pH was adjusted to 5 with an aqueous dilute NaOH solution, and the mixture was stirred at room temperature for 3 h. Water was removed by lyophilization and the crude analyzed by HPLC.

**Diels–Alder Cycloadditions.** The maleimido- or protected maleimido-containing peptide and diene-dT<sub>5</sub> were dissolved in water, and the amounts required to carry out the reaction (typical reagent ratio was 1:1) were taken from each of these solutions and mixed. This mixture was diluted with water (or water and methanol, to end up with a 1:1 proportion, if MW-promoted maleimide deprotection and conjugation were to take place simultaneously) to reach a reagent concentration between 30 and 300  $\mu\text{M}$ . The maleimide + diene mixture was reacted overnight at 37 °C. The [protected maleimide] + diene mixture was reacted in a MW oven

for 90 min at 90 °C. Solvents were removed by evaporation under reduced pressure and/or lyophilization, and reaction crudes were analyzed by HPLC. Conjugates were purified by HPLC and characterized by MALDI-TOF MS.

**Michael-Type Reactions.** Aliquots of aqueous solutions containing the required amounts of maleimido-containing compound and the corresponding thiol derivative were mixed (final concentration of polyamide: 50–150  $\mu\text{M}$ ). The mixture was stirred at room temperature under an Ar atmosphere overnight. After lyophilization, reaction crudes were analyzed by HPLC. Conjugates were purified by HPLC and characterized by MALDI-TOF MS.

**Assays To Assess the Stability of Cysteine Derivatives to Microwave-Promoted Maleimide-Deprotection Conditions.** The stabilities of Fmoc-Cys(S<sup>t</sup>Bu)-OH, Fmoc-Cys(Acm)-OH, and Boc-Cys(Fm)-OH were assessed by dissolving some 5–10 mg of the product in MeOH/H<sub>2</sub>O 1:1 (v/v) and heating the resulting solution at 90 °C for 90 min in the microwave oven. Solvent was removed under reduced pressure, and the crude was analyzed <sup>1</sup>H NMR. In all cases, a spectrum with the same set of signals as for the commercial compound was obtained, with one exception: partial loss of the tBu signal in the case of Boc-Cys(Fm)-OH, due to partial loss of the Boc group upon heating.

**Cyclic Maleimido\*-Lys-Gly-Arg-Gly-Cys\*-Tyr-Ala-Tyr-Lys-(Prot.Mal.)-NH<sub>2</sub>, 2 (the Label \* Refers to the Units Linked in the Cycle).** Peptide 2 was assembled on 110 mg of resin ( $f = 0.65$  mmol/g). Treatment of peptide-resin 1 with TFA/TIS/H<sub>2</sub>O 90:5:5, 2  $\times$  1 h, followed by purification at the semipreparative scale (linear gradient from 15 to 35% of B in 30 min,  $t_R = 15.1$  min), afforded 2 as a white solid (12.1  $\mu\text{mol}$ , 17%). Analytical HPLC: acidic conditions, linear gradient from 10 to 50% of B in 30 min,  $t_R = 16.2$  min. MALDI-TOF MS (DHB, positive mode, DHB = 2,5-dihydroxybenzoic acid):  $m/z$  1347.4 [M – furan + H]<sup>+</sup>, 1369.4 [M – furan + Na]<sup>+</sup>; M calcd for C<sub>66</sub>H<sub>91</sub>N<sub>17</sub>O<sub>18</sub>S 1441.6, M – furan calcd for C<sub>60</sub>H<sub>83</sub>N<sub>17</sub>O<sub>17</sub>S 1345.6. MALDI-TOF MS after reaction with H<sub>2</sub>O<sub>2</sub> (DHB, positive mode):  $m/z$  1362.8 [M – furan + 16 + H]<sup>+</sup>; M – furan + 16 calcd for C<sub>60</sub>H<sub>83</sub>N<sub>17</sub>O<sub>18</sub>S 1361.6. MALDI-TOF MS after digestion with thrombin (DHB, positive mode):  $m/z$  1347.3 [M + H]<sup>+</sup>, 1369.3 [M + Na]<sup>+</sup>, 1385.2 [M + K]<sup>+</sup>; M calcd for C<sub>66</sub>H<sub>91</sub>N<sub>17</sub>O<sub>18</sub>S 1441.6, M – furan calcd for C<sub>60</sub>H<sub>83</sub>N<sub>17</sub>O<sub>17</sub>S 1345.6. No digestion was observed.

**Conjugate 3 [Cyclic Maleimido\*-Lys-Gly-Arg-Gly-Cys\*-Tyr-Ala-Tyr-Lys(Mal.)-NH<sub>2</sub>] + [Diene-dT<sub>5</sub>].** Diene-dT<sub>5</sub> and 2 (15 nmol each) were reacted in a 1:1 MeOH/H<sub>2</sub>O solution (500  $\mu\text{L}$ , 30  $\mu\text{M}$ ) in a MW oven for 90 min at 90 °C. HPLC Analysis of the crude showed that the extent of the cycloaddition was 90%. Analytical HPLC: buffered medium, gradient from 5% to 60% of B in 30 min,  $t_R = 22.3$  min. MALDI-TOF (THAP, negative mode, THAP = 2,4,6-trihydroxyacetophenone):  $m/z$  2962.7 [M – H]<sup>–</sup>; M calcd for C<sub>116</sub>H<sub>158</sub>N<sub>27</sub>O<sub>53</sub>P<sub>5</sub>S 2963.9.

**H-Cys(S<sup>t</sup>Bu)-Gly-Arg-Gly-Cys-Tyr-Ala-Tyr-Lys(Prot.Mal.)-NH<sub>2</sub>, 5.** Fmoc-Cys(S<sup>t</sup>Bu)-OH was incorporated onto 121  $\mu\text{mol}$  of Fmoc-Gly-Arg(Pbf)-Gly-Cys(Tr)-Tyr(<sup>t</sup>Bu)-Ala-Tyr(<sup>t</sup>Bu)-Lys(Prot.Mal.)-resin, and the Fmoc group removed to yield 4. 4 Was deprotected and cleaved by reaction with TFA/TIS/H<sub>2</sub>O 95:2.5:2.5, 2  $\times$  1 h, and purified at the semipreparative scale (linear gradient from 25 to 50% of B in 30 min,  $t_R = 9.6$  min). A white solid corresponding to the desired peptide (5) was obtained (28  $\mu\text{mol}$ , 23%). Analytical HPLC: acidic conditions, linear gradient from 5 to 50% of B in 30 min,  $t_R = 22.7$  min. MALDI-TOF MS (DHB, positive mode):  $m/z$  1258.9 [M – furan + H]<sup>+</sup>; M calcd for C<sub>60</sub>H<sub>87</sub>N<sub>15</sub>O<sub>15</sub>S<sub>3</sub> 1353.6, M – furan calcd for C<sub>54</sub>H<sub>79</sub>N<sub>15</sub>O<sub>14</sub>S<sub>3</sub> 1257.5. MALDI-TOF MS after reaction with H<sub>2</sub>O<sub>2</sub> (DHB, negative mode):  $m/z$  1216.6 [M – furan – S<sup>t</sup>Bu + 48 – H]<sup>–</sup>, 1264.5 [M – furan – S<sup>t</sup>Bu + 48 + 48–H]<sup>–</sup>; M + 48 calcd for C<sub>60</sub>H<sub>87</sub>N<sub>15</sub>O<sub>18</sub>S<sub>3</sub> 1401.6, M – furan – S<sup>t</sup>Bu + 48 calcd for C<sub>50</sub>H<sub>71</sub>N<sub>15</sub>O<sub>17</sub>S<sub>2</sub> 1217.5. MALDI-TOF MS after digestion with thrombin (DHB, positive mode):  $m/z$  876.4 [M – furan – [Cys(S<sup>t</sup>Bu)-Gly-Arg] + Na]<sup>+</sup>, 892.4 [M – furan – [Cys(S<sup>t</sup>Bu)-Gly-Arg] + K]<sup>+</sup>; M calcd for C<sub>45</sub>H<sub>59</sub>N<sub>9</sub>O<sub>12</sub>S 949.4, M – furan – [Cys(S<sup>t</sup>Bu)-Gly-Arg] calcd for C<sub>39</sub>H<sub>51</sub>N<sub>9</sub>O<sub>11</sub>S 853.3.

**Deprotection of the Maleimide of 5.** Compound 5 (100 nmol) was dissolved in H<sub>2</sub>O/MeOH (4 mL) to yield a 25  $\mu\text{M}$  solution, which was heated in a MW oven as described in the General Methods. Two different products, namely 6 and 7, were obtained (no trace of 5 was detected in any case) in a ratio that varied depending on the reaction time: the 6:7 ratio was 49:51 after 30 min heating, 56:44 after 60 min, and 59:41 after 90 min. Compounds 6 and 7 were separated using analytical acidic HPLC conditions and characterized by MALDI-TOF MS before and after reaction with thrombin (see below). For an additional proof of structure, samples of 6 and 7 were subsequently reacted as described below to yield compounds 16 and 8, respectively, which were also characterized by MS.

**Cyclic H-Cys(S<sup>t</sup>Bu)-Gly-Arg-Gly-Cys\*-Tyr-Ala-Tyr-Lys\*-NH<sub>2</sub>, 6.** Analytical HPLC: acidic conditions, linear gradient from 5 to 50% of B in 30 min,  $t_R = 20.1$  min. MALDI-TOF MS (DHB, positive mode):  $m/z$  1258.9 [M + H]<sup>+</sup>, 1280.9 [M + Na]<sup>+</sup>, 1296.8 [M + K]<sup>+</sup>; M calcd for C<sub>54</sub>H<sub>79</sub>N<sub>15</sub>O<sub>14</sub>S<sub>3</sub> 1257.5. MALDI-TOF MS after reaction with H<sub>2</sub>O<sub>2</sub> (DHB, positive mode):  $m/z$  1186.7 [M – S<sup>t</sup>Bu + 16 + H]<sup>+</sup>, 1234.7 [M – S<sup>t</sup>Bu + 16 + 48 + H]<sup>+</sup>, 1290.7 [M + 32 + H]<sup>+</sup>; M + 32 calcd for C<sub>54</sub>H<sub>79</sub>N<sub>15</sub>O<sub>16</sub>S<sub>3</sub> 1289.5, M – S<sup>t</sup>Bu + 16 calcd for C<sub>50</sub>H<sub>71</sub>N<sub>15</sub>O<sub>15</sub>S<sub>2</sub> 1185.5. MALDI-TOF MS after reaction with thrombin (DHB, positive mode):  $m/z$  876.4 [M – [Cys(S<sup>t</sup>Bu)-Gly-Arg] + Na]<sup>+</sup>, 892.4 [M – [Cys(S<sup>t</sup>Bu)-Gly-Arg] + K]<sup>+</sup>; M – [Cys(S<sup>t</sup>Bu)-Gly-Arg] calcd for C<sub>39</sub>H<sub>51</sub>N<sub>9</sub>O<sub>11</sub>S 853.3.

**Cyclic H-Cys-Gly-Arg-Gly-Cys\*-Tyr-Ala-Tyr-Lys\*-NH<sub>2</sub>, 16.** Compound 6 (67 nmol) was reacted with TCEP according to the general procedures. Compound 16 was obtained in 76% yield (based on the HPLC trace). Analytical HPLC: acidic conditions, linear gradient from 5 to 50% of B in 30 min,  $t_R = 16.1$  min. MALDI-TOF MS (DHB, positive mode):  $m/z$  1170.7 [M + H]<sup>+</sup>, 1192.7 [M + Na]<sup>+</sup>, 1208.7 [M + K]<sup>+</sup>; M calcd for C<sub>50</sub>H<sub>71</sub>N<sub>15</sub>O<sub>14</sub>S<sub>2</sub> 1169.5. MALDI-TOF MS after reaction with H<sub>2</sub>O<sub>2</sub> (DHB, negative mode):  $m/z$  1232.6 [M + 48 + 16 – H]<sup>–</sup>; M + 48 + 16 calcd for C<sub>50</sub>H<sub>71</sub>N<sub>15</sub>O<sub>18</sub>S<sub>2</sub> 1233.5. MALDI-TOF MS after digestion with thrombin (DHB, positive mode):  $m/z$  876.5 [M – [Cys-Gly-Arg] + Na]<sup>+</sup>, 892.5 [M – [Cys-Gly-Arg] + K]<sup>+</sup>; M – [Cys-Gly-Arg] calcd for C<sub>39</sub>H<sub>51</sub>N<sub>9</sub>O<sub>11</sub>S 853.4.

**Cyclic H-Cys\*-Gly-Arg-Gly-Cys\*-Tyr-Ala-Tyr-Lys(Prot.Mal.)-NH<sub>2</sub>, 7.** Analytical HPLC: acidic conditions, linear gradient from 5 to 50% of B in 30 min,  $t_R = 16.7$  min. MALDI-TOF MS (DHB, positive mode):  $m/z$  1168.8 [M + H]<sup>+</sup>, 1190.7 [M + Na]<sup>+</sup>, 1206.7 [M + K]<sup>+</sup>; M calcd for C<sub>50</sub>H<sub>69</sub>N<sub>15</sub>O<sub>14</sub>S<sub>2</sub> 1167.5. MALDI-TOF MS after reaction with H<sub>2</sub>O<sub>2</sub> (DHB, negative mode):  $m/z$  1264.7 [M + 48 + 48 – H]<sup>–</sup>; M + 48 + 48 calcd for C<sub>50</sub>H<sub>71</sub>N<sub>15</sub>O<sub>20</sub>S<sub>2</sub> 1265.4. MALDI-TOF MS after digestion with thrombin (DHB, positive mode):  $m/z$  1168.8 [M + H]<sup>+</sup>, 1190.7 [M + Na]<sup>+</sup>, 1206.7 [M + K]<sup>+</sup>; M calcd for C<sub>50</sub>H<sub>69</sub>N<sub>15</sub>O<sub>14</sub>S<sub>2</sub> 1167.5. No digestion was detected.

**Conjugate 8, [Cyclic H-Cys\*-Gly-Arg-Gly-Cys\*-Tyr-Ala-Tyr-Lys(Mal.)-NH<sub>2</sub>] + [diene-dT<sub>5</sub>].** Compound 7 (15 nmol) and diene-dT<sub>5</sub> (25 nmol) were dissolved in water (65  $\mu\text{L}$ ) and allowed to react overnight at room temperature, yielding the desired product 8 (8.8 nmol, 59% conjugation and purification yield). 8 Was purified using acidic analytical HPLC conditions. Analytical HPLC: buffered medium, linear gradient from 5 to 60% of B in 30 min,  $t_R = 23.0$  min. MALDI-TOF MS (THAP, negative mode):  $m/z$  2785.0 [M – H]<sup>–</sup>, 2822.9 [M – 2H + K]<sup>–</sup>; M calcd for C<sub>106</sub>H<sub>144</sub>N<sub>25</sub>O<sub>50</sub>P<sub>5</sub>S<sub>2</sub> 2785.8.

**H-Cys(Acm)-Gly-Arg-Gly-Cys-Tyr-Ala-Tyr-Lys(Prot.Mal.)-NH<sub>2</sub>, 10.** Fmoc-Cys(Acm)-OH was incorporated onto 34  $\mu\text{mol}$  of peptide-resin Fmoc-Gly-Arg(Pbf)-Gly-Cys-Tyr(<sup>t</sup>Bu)-Ala-Tyr(<sup>t</sup>Bu)-Lys(Prot.Mal.)-resin and the Fmoc group removed to yield 9. Treatment with TFA/TIS/H<sub>2</sub>O 95:2.5:2.5, 2  $\times$  1 h and purification at the semipreparative scale (linear gradient from 25 to 70% of B in 30 min,  $t_R = 6.2$  min) afforded a white solid corresponding to the desired peptide 10 (11.4  $\mu\text{mol}$ , 34%). Analytical HPLC: acidic conditions, linear gradient from 5 to 50% of B in 30 min,  $t_R = 19.0$  min. MALDI-TOF MS (DHB, positive mode):  $m/z$  1170.7 [M – furan – Acm + H]<sup>+</sup> 1241.8 [M – furan + H]<sup>+</sup>; M calcd for C<sub>59</sub>H<sub>84</sub>N<sub>16</sub>O<sub>16</sub>S<sub>2</sub> 1336.6, M – furan calcd for C<sub>53</sub>H<sub>76</sub>N<sub>16</sub>O<sub>15</sub>S<sub>2</sub>

1240.5. MALDI-TOF MS after reaction with H<sub>2</sub>O<sub>2</sub> (DHB, negative mode):  $m/z$  1169.0 [M – furan – AcM–H]<sup>–</sup>, 1217.0 [M – furan – AcM + 48–H]<sup>–</sup>, 1233.0 [M – furan – AcM + 48 + 16–H]<sup>–</sup>; M + 48 calcd for C<sub>59</sub>H<sub>84</sub>N<sub>16</sub>O<sub>19</sub>S<sub>2</sub> 1384.5, M – furan – AcM + 48 calcd for C<sub>50</sub>H<sub>71</sub>N<sub>15</sub>O<sub>17</sub>S<sub>2</sub> 1217.5.

**Cyclic H-Cys(AcM)-Gly-Arg-Gly-Cys\*-Tyr-Ala-Tyr-Lys\*-NH<sub>2</sub>, 11.** Compound 10 (100 nmol) was dissolved in the H<sub>2</sub>O/MeOH mixture (1 mL, 100 μM) and heated in a MW oven according to the general procedures, yielding the desired product 11 (30 nmol, 30% cyclization and purification yield). Compound 11 was purified using acidic analytical HPLC conditions. Analytical HPLC: acidic conditions, linear gradient from 5 to 50% of B in 30 min,  $t_R$  = 16.0, 16.2 min (2 diastereomers). MALDI-TOF MS (DHB, positive mode):  $m/z$  1241.7 [M + H]<sup>+</sup>, 1263.7 [M + Na]<sup>+</sup>, 1279.7 [M + K]<sup>+</sup>; M calcd for C<sub>53</sub>H<sub>76</sub>N<sub>16</sub>O<sub>15</sub>S<sub>2</sub> 1240.5. MALDI-TOF MS after reaction with H<sub>2</sub>O<sub>2</sub> (DHB, positive mode):  $m/z$  1203.2 [M – AcM + 32 + H]<sup>+</sup>; M + 32 calcd for C<sub>53</sub>H<sub>76</sub>N<sub>16</sub>O<sub>17</sub>S<sub>2</sub> 1272.5, M – AcM + 32 calcd for C<sub>50</sub>H<sub>71</sub>N<sub>15</sub>O<sub>16</sub>S<sub>2</sub> 1201.6.

**Assays To Remove the AcM Group from Peptide 11.** *Reaction with AgOTf.* Compound 11 (30 nmol) was dissolved in TFA (1 mL), and AgOTf (1 mg, 3.9 μmol) and anisole (430 μL, 3.9 μmol) were added. The mixture was allowed to react for 1 h, after which time solvent was removed under reduced pressure and the crude redissolved with DTT (1 mg, 6.5 μmol) in a 1:1 mixture of H<sub>2</sub>O/AcOH (1.5 mL). After 30 min, the solution was frozen and lyophilized. The crude was then redissolved in water and analyzed, yielding a very complex HPLC profile. Integration of the HPLC trace and MALDI-TOF MS analysis showed that the desired product was obtained in less than 5% yield.

*Reaction with I<sub>2</sub>.* Compound 11 (500 nmol) was dissolved in a 9:1 H<sub>2</sub>O/AcOH mixture (100 μL), and I<sub>2</sub> (100 μL of a 50 mM solution in MeOH) was added. The mixture was allowed to react for 4 h. The reaction was quenched with 0.2 M sodium ascorbate (20 μL), and methanol was removed under reduced pressure. Neither HPLC analysis nor MALDI-TOF MS of the crude showed the presence of the desired product, but instead one resulting from the addition of one to four iodines to the peptide.

*Reaction with Ti(OAc)<sub>3</sub>.* Compound 11 (50 nmol) was dissolved in TFA (500 μL), and a solution of Ti(OAc)<sub>3</sub> (390 μg, 100 nmol) and anisole (10.9 μL, 100 nmol) in TFA (10 μL) was added. The mixture was reacted for 1 h at 0 °C, and the solvent was removed with a N<sub>2</sub> stream. 10% acetic acid in water (1 mL) and chloroform (1 mL) were added, and the two fractions were collected separately. HPLC and MALDI-TOF analysis of the two fractions showed the presence of multiple products, none of which could be identified.

**H-Cys(Fm)-Gly-Arg-Gly-Cys-Tyr-Ala-Tyr-Lys(Prot.Mal.)-NH<sub>2</sub>, 13.** Boc-Cys(Fm)-OH was incorporated onto 31 μmol of Fmoc-Gly-Arg(Gly)-Gly-Cys(Tr)-Tyr(Bu)-Ala-Tyr(Bu)-Lys(Prot.Mal.)-resin. Treatment with TFA/TIS/H<sub>2</sub>O 95:2.5:2.5, 2 × 1 h, and purification at the semipreparative scale (linear gradient from 30 to 60% of B in 30 min,  $t_R$  = 9.5 min) afforded a white solid corresponding to the desired peptide 13 (9.6 μmol, 31%). Analytical HPLC: acidic conditions, linear gradient from 30 to 60% of B in 30 min,  $t_R$  = 7.0 min. MALDI-TOF (DHB, positive mode):  $m/z$  1348.9 [M – furan + H]<sup>+</sup>, 1370.8 [M – furan + Na]<sup>+</sup>, 1386.8 [M – furan + K]<sup>+</sup>; M calcd for C<sub>70</sub>H<sub>89</sub>N<sub>15</sub>O<sub>15</sub>S<sub>2</sub> 1443.6, M – furan calcd for C<sub>64</sub>H<sub>81</sub>N<sub>15</sub>O<sub>14</sub>S<sub>2</sub> 1347.6. MALDI-TOF MS after reaction with H<sub>2</sub>O<sub>2</sub> (DHB, negative mode):  $m/z$  1214.4 [M – furan – Fm + 48 – H]<sup>–</sup>, 1232.6 [M – furan – Fm + 48 + 16 – H]<sup>–</sup>; M + 48 + 16 calcd for C<sub>70</sub>H<sub>89</sub>N<sub>15</sub>O<sub>19</sub>S<sub>2</sub> 1507.6, M – furan – Fm + 48 + 16 calcd for C<sub>50</sub>H<sub>71</sub>N<sub>15</sub>O<sub>18</sub>S<sub>2</sub> 1233.5.

**Cyclic H-Cys(Fm)-Gly-Arg-Gly-Cys\*-Tyr-Ala-Tyr-Lys\*-NH<sub>2</sub>, 14.** Compound 13 (400 nmol) was dissolved in the H<sub>2</sub>O/MeOH mixture (4 mL, 100 μM), heated in a MW oven according to the general procedures, and purified using analytical HPLC conditions to yield 14 (184 nmol, 46% cyclization and purification yield). Analytical HPLC: acidic conditions, linear gradient from 30 to 60% of B in 30 min,  $t_R$  = 4.8 min. MALDI-TOF MS (DHB, positive mode):  $m/z$  1348.8 [M + H]<sup>+</sup>, 1370.8 [M + Na]<sup>+</sup>, 1386.7 [M + K]<sup>+</sup>; M calcd for C<sub>64</sub>H<sub>81</sub>N<sub>15</sub>O<sub>14</sub>S<sub>2</sub> 1347.6. MALDI-TOF after reaction

with H<sub>2</sub>O<sub>2</sub> (DHB, positive mode):  $m/z$  1202.7 [M – Fm + 32 + H]<sup>+</sup>, 1380.9 [M + 32 + H]<sup>+</sup>; M + 32 calcd for C<sub>64</sub>H<sub>81</sub>N<sub>15</sub>O<sub>16</sub>S<sub>2</sub> 1379.5.

**Disulfide Dimer of Cyclic H-Cys-Gly-Arg-Gly-Cys\*-Tyr-Ala-Tyr-Lys\*-NH<sub>2</sub>, 15.** Compound 14 (92 nmol) was dissolved in 33% piperidine in DCM (500 μL) and the mixture reacted for 45 min. The solvent was then removed under reduced pressure and coevaporated with acetonitrile. The crude was purified using analytical HPLC acidic conditions (linear gradient from 5 to 50% of B in 30 min,  $t_R$  = 17.8 min), yielding 15 (41.6 nmol, 45% reaction and purification yield). MALDI-TOF MS (DHB, positive mode):  $m/z$  1170.6 [monomer + H]<sup>+</sup>, 2338.3 [M + H]<sup>+</sup>; M calcd for C<sub>100</sub>H<sub>114</sub>N<sub>30</sub>O<sub>28</sub>S<sub>4</sub> 2336.9.

**Cyclic H-Cys-Gly-Arg-Gly-Cys\*-Tyr-Ala-Tyr-Lys\*-NH<sub>2</sub>, 16.** Compound 15 (27 nmol) was reacted with TCEP as described above. Compound 16 was obtained in 89% yield (based on the HPLC trace) and purified using analytical HPLC conditions. Analytical HPLC: acidic conditions, linear gradient from 5 to 50% of B in 30 min,  $t_R$  = 15.8 min. MALDI-TOF MS (DHB, positive mode):  $m/z$  1171.1 [M + H]<sup>+</sup>, 1193.0 [M + Na]<sup>+</sup>; M calcd for C<sub>50</sub>H<sub>71</sub>N<sub>15</sub>O<sub>14</sub>S<sub>2</sub> 1169.5.

**Conjugate 17 [Cyclic H-Cys-Gly-Arg-Gly-Cys\*-Tyr-Ala-Tyr-Lys\*-NH<sub>2</sub>] + [Maleimido-dT<sub>5</sub>].** [Protected maleimido]-dT<sub>5</sub> (25 nmol) was dissolved in the 1:1 MeOH/H<sub>2</sub>O solution (1 mL, 25 μM) and heated in the microwave oven for 60 min at 90 °C. Methanol was removed under reduced pressure. HPLC analysis of the crude (buffered medium) showed that maleimide deprotection had taken place at an 85% extent. This crude was used immediately for the conjugation reaction. Analytical HPLC: buffered medium, linear gradient from 5 to 60% of B in 30 min,  $t_R$  = 18.3 min. MALDI-TOF MS (THAP, negative mode):  $m/z$  1661.7 [M – H]<sup>–</sup>; M calcd for C<sub>56</sub>H<sub>72</sub>N<sub>11</sub>O<sub>38</sub>P<sub>5</sub> 1661.3.

Mal-dT<sub>5</sub> (5 nmol) was mixed with 16 (4 nmol) dissolved in water (375 μL) and reacted overnight. The desired product (17) was obtained in 73% yield (based on the HPLC trace). Analytical HPLC: buffered medium, linear gradient from 5 to 60% of B in 30 min,  $t_R$  = 22.0 min. MALDI-TOF MS (THAP, negative mode):  $m/z$  2830.2 [M – H]<sup>–</sup>; M calcd for C<sub>106</sub>H<sub>143</sub>N<sub>26</sub>O<sub>52</sub>P<sub>5</sub>S<sub>2</sub> 2830.8.

**Cyclic Lys\*-Gly-Arg-Gly-Cys(S'Bu)-Tyr-Ala-Tyr-Cys\*-NH<sub>2</sub>, 19.** The synthesis of 18 was carried out on 110 mg of resin ( $f$  = 0.65 mmol/g). After reaction with TFA/TIS/H<sub>2</sub>O 90:5:5, 2 × 1 h, and purification at the semipreparative scale (linear gradient from 25 to 50% of B in 30 min,  $t_R$  = 9.8 min), 19 was obtained as a white solid (8.7 μmol, 11%). Analytical HPLC: acidic conditions, linear gradient from 20 to 60% of B in 30 min,  $t_R$  = 11.7 min. MALDI-TOF MS (DHB, positive mode):  $m/z$  1259.1 [M + H]<sup>+</sup>; M calcd for C<sub>54</sub>H<sub>79</sub>N<sub>15</sub>O<sub>14</sub>S<sub>3</sub> 1257.5. MALDI-TOF MS after reaction with H<sub>2</sub>O<sub>2</sub> (DHB, positive mode):  $m/z$  1186.6 [M – S'Bu + 16 + H]<sup>+</sup>, 1306.7 [M + 32 + 16 + H]<sup>+</sup>, M – S'Bu + 16 calcd for C<sub>50</sub>H<sub>71</sub>N<sub>15</sub>O<sub>15</sub>S<sub>2</sub> 1185.5; M + 32 + 16 calcd for C<sub>54</sub>H<sub>79</sub>N<sub>15</sub>O<sub>17</sub>S<sub>3</sub> 1305.5. MALDI-TOF MS after digestion with thrombin (DHB, positive mode):  $m/z$  1259.2 [M + H]<sup>+</sup>, 1281.2 [M + Na]<sup>+</sup>, 1297.2 [M + K]<sup>+</sup>; M calcd for C<sub>54</sub>H<sub>79</sub>N<sub>15</sub>O<sub>14</sub>S<sub>3</sub> 1257.5. No digestion was observed.

**Cyclic Lys\*-Gly-Arg-Gly-Cys-Tyr-Ala-Tyr-Cys\*-NH<sub>2</sub>, 20.** Compound 19 (100 nmol) was reacted with TCEP as described in the General Methods. Compound 20 was obtained in 80% yield (based on the HPLC trace) and purified using acidic analytical HPLC conditions. Analytical HPLC: acidic conditions, linear gradient from 10 to 50% of B in 30 min  $t_R$  = 13.5 min. MALDI-TOF MS (DHB, positive mode):  $m/z$  1170.9 [M + H]<sup>+</sup>; M calcd for C<sub>50</sub>H<sub>71</sub>N<sub>15</sub>O<sub>14</sub>S<sub>2</sub> 1169.5.

**Conjugate 21 [Cyclic Lys\*-Gly-Arg-Gly-Cys-Tyr-Ala-Tyr-Cys\*-NH<sub>2</sub>] + [Maleimido-dT<sub>5</sub>].** Compound 20 (10 nmol) and [protected maleimido]-dT<sub>5</sub> (10 nmol) were dissolved in a 3:1 MeOH/H<sub>2</sub>O solution (500 μL, 20 μM) and heated in a MW oven for 90 min at 90 °C. Methanol was removed under reduced pressure. The desired product (21) was obtained in 77% yield (based on the HPLC trace). Analytical HPLC: buffered medium, linear gradient from 5 to 50% of B in 30 min,  $t_R$  = 14.9 min. MALDI-TOF MS



(THAP, negative mode):  $m/z$  2829.5  $[M - H]^-$ ; M calcd for  $C_{106}H_{143}N_{26}O_{52}P_5S_2$  2830.7.

**Protected maleimido**-[N-(CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>Phenol)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>OH)Gly]-[N-(Propargyl)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>Phenol)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>OH)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>SH)Gly]-Gly-NH<sub>2</sub>, **23**. Peptoid **23** was assembled on 100 mg of resin ( $f = 0.72$  mmol/g). Treatment of the peptoid-resin **22** with TFA/*m*-cresol/TIS/H<sub>2</sub>O 90:5:2.5:2.5, 2 × 1 h, followed by purification at the semipreparative scale (linear gradient from 20 to 40% of B in 30 min,  $t_R = 12.0$  min), afforded **23** as a white solid (3.0 μmol, 4.1%). Analytical HPLC: acidic conditions, linear gradient from 20 to 40% of B in 30 min,  $t_R = 9.7$  min. MALDI-TOF MS (THAP/CA, positive mode, CA = citric acid):  $m/z$  1194.7  $[M - furan + H]^+$ , 1290.8  $[M + H]^+$ ; M calcd for  $C_{60}H_{83}N_{13}O_{17}S$  1289.6, M - furan calcd for  $C_{54}H_{75}N_{13}O_{16}S$  1193.5. MALDI-TOF MS after reaction with H<sub>2</sub>O<sub>2</sub> (THAP/CA, negative mode):  $m/z$  1240.7  $[M - furan + 48 - H]^-$ ; M - furan + 48 calcd for  $C_{54}H_{75}N_{13}O_{19}S$  1241.5.

**Cyclic Maleimido**\*-[N-(CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>Phenol)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>OH)Gly]-[N-(Propyne)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>Phenol)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>OH)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>SH)Gly]\*-Gly-NH<sub>2</sub>, **24**. Compound **23** (100 nmol) was dissolved in the H<sub>2</sub>O/MeOH solution (1 mL, 100 μM) and heated in a MW oven according to the general procedures. MeOH was removed under reduced pressure. The desired product (**24**) was obtained in 79% yield (based on the HPLC trace). Analytical HPLC: acidic conditions, linear gradient from 20 to 40% of B in 30 min,  $t_R = 15.8$  min. MALDI-TOF MS (THAP/CA, positive mode):  $m/z$  1194.7  $[M + H]^+$ ; M calcd for  $C_{54}H_{75}N_{13}O_{16}S$  1193.5. MALDI-TOF MS after reaction with H<sub>2</sub>O<sub>2</sub> (THAP/CA, positive mode):  $m/z$  1210.7  $[M + 16 + H]^+$ ; M + 16 calcd for  $C_{54}H_{75}N_{13}O_{17}S$  1209.5.

**Conjugate 25** [Cyclic Maleimido\*-[N-(CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>Phenol)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>OH)Gly]-[N-(Propyne)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>Phenol)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>OH)Gly]-[N-(CH<sub>2</sub>CH<sub>2</sub>SH)Gly]\*-Gly-NH<sub>2</sub>] + AZT. A 0.5 mM solution of CuSO<sub>4</sub> in water (50 μL, 25 nmol) and a 12 mM solution of sodium ascorbate in water (50 μL, 600 nmol) were mixed in a microwave flask under an Ar atmosphere and stirred during 5 min. Then TBTA (tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine) (50 μL of a 0.5 mM solution in MeOH:H<sub>2</sub>O 3:2, 24 nmol), AZT (3'-azido-3'-deoxythymidine) (110 μL of a 2 mM aqueous solution, 200 nmol), and compound **24** (100 μL of a 290 μM aqueous solution, 29 nmol) were added. The mixture was heated in a microwave oven for 1 h at 90 °C. MeOH was then removed under reduced pressure. The desired product (**25**) was obtained in 90% yield (based on the HPLC trace, calculated from compound **24**). Analytical HPLC: acidic conditions, linear gradient from 5 to 50% of B in 30 min,  $t_R = 17.1$  min. MALDI-TOF MS (DHB, positive mode):  $m/z$ : 1462.4  $[M + H]^+$ , 1484.3  $[M + Na]^+$ ; M calcd for  $C_{64}H_{88}N_{18}O_{20}S$  1460.6.

**S-Triyl Cysteamine Hydrochloride**. Triyl chloride (279 mg, 1.6 mmol) and cysteamine hydrochloride (200 mg, 1.8 mmol) were dissolved in DMF (1.5 mL) and stirred overnight at room temperature. The reaction was then quenched with cold water (10 mL), and the precipitated formed was filtered and washed with cold water (3 × 5 mL). A white solid (480 mg, 85% yield) was obtained:  $R_f$  (hexanes/AcOEt/TEA 75:20:5) = 0.68; <sup>1</sup>H NMR (CD<sub>3</sub>OD, 400 MHz)  $\delta$  7.5–7.1 (m, 15H, aromatics), 2.58 (t,  $J = 6.9$ , 2H, CH<sub>2</sub>-NH<sub>2</sub>), 2.47 (t,  $J = 6.8$ , 2H, S-CH<sub>2</sub>) ppm; <sup>13</sup>C NMR (CD<sub>3</sub>OD, 100 MHz)  $\delta$  145.6, 130.7, 129.2, 128.2, 68.4, 39.7, 30.1; ESI MS (positive mode)  $m/z$  319.9  $[M + H]^+$ , 639.0  $[2M + H]^+$ ; M calcd for C<sub>21</sub>H<sub>21</sub>NS 319.14.

**N-Boc-ethylenediamine**. A solution of Boc anhydride (1.2 g, 5.5 mmol) in DCM (80 mL) was added dropwise during 15 min to a solution of ethylenediamine (2.2 mL, 33.5 mmol) in DCM (10 mL). The mixture was stirred for 1 h at 0 °C and overnight at room temperature. Solvent was removed under reduced pressure, and the crude was dissolved in aqueous potassium carbonate (50 mL). This solution was extracted with DCM (3 × 50 mL). The organic fraction was dried over MgSO<sub>4</sub> and filtered and the solvent removed under reduced pressure. A colorless oil (857 mg, 60% yield) was obtained:  $R_f$  (hexanes/AcOEt/TEA 47.5:47.5:5) = 0.2; <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz)  $\delta$  3.16 (q,  $J = 5.9$ , 2H, CO-NH-CH<sub>2</sub>); 2.79 (t,  $J = 5.7$ , 2H,

CH<sub>2</sub>-NH<sub>2</sub>); 1.44 (s, 9H, CH<sub>3</sub>) ppm; <sup>13</sup>C NMR CDCl<sub>3</sub>, 100 MHz)  $\delta$  156.2, 79.2, 43.2, 41.8, 28.4 ppm; ESI MS (positive mode)  $m/z$  161.13  $[M + H]^+$ ; M calcd for C<sub>7</sub>H<sub>16</sub>N<sub>2</sub>O<sub>2</sub> 160.12.

## ■ ASSOCIATED CONTENT

### ● Supporting Information

HPLC traces of the compounds prepared, commercially available building blocks used for peptide and peptoid assembly, and <sup>1</sup>H and <sup>13</sup>C NMR spectra of compounds synthesized according to described procedures. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [anna.grandas@ub.edu](mailto:anna.grandas@ub.edu).

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was supported by funds from the Ministerio de Economía y Competitividad (Grant No. CTQ2010-21567-C02-01 and the project RNAREG, Grant No. CSD2009-00080, funded under the programme CONSOLIDER INGENIO 2010), and the Generalitat de Catalunya (2009SGR-208). X.E. was a recipient fellow of the MINECO.

## ■ REFERENCES

- (1) Bock, J. E.; Gavenonis, J.; Kritzer, J. A. *ACS Chem. Biol.* **2013**, *8*, 488–499.
- (2) Rezaei, T.; Bock, J. E.; Zhou, M. V.; Kalyanaraman, C.; Lokey, R. S.; Jacobson, M. P. *J. Am. Chem. Soc.* **2006**, *128*, 14073–14080.
- (3) Rezaei, T.; Yu, B.; Millhauser, G. L.; Jacobson, M. P.; Lokey, R. S. *J. Am. Chem. Soc.* **2006**, *128*, 2510–2511.
- (4) Boturin, D.; Coll, J.-L.; Garanger, E.; Favrot, M.-C.; Dumy, P. *J. Am. Chem. Soc.* **2004**, *126*, 5730–5739.
- (5) Birts, C. N.; Nijjar, S. K.; Mardle, C. A.; Hoakwie, F.; Duriez, P. J.; Blaydes, J. P.; Tavassoli, A. *Chem. Sci.* **2013**, *4*, 3046–3057.
- (6) Zhou, H.; Liu, L.; Huang, J.; Bernard, D.; Karatas, H.; Navarro, A.; Lei, M.; Wang, S. *J. Med. Chem.* **2013**, *56*, 1113–1123.
- (7) Spurr, I. B.; Birts, C. N.; Cuda, F.; Benkovic, S. J.; Blaydes, J. P.; Tavassoli, A. *ChemBioChem* **2012**, *13*, 1628–1634.
- (8) Smith, J. W.; Ruggeri, Z. M.; Kunicki, T. J.; Cheresch, D. A. *J. Biol. Chem.* **1992**, *265*, 12267–12271.
- (9) Gurrath, M.; Müller, G.; Kessler, H.; Aumailley, M.; Timpl, R. *Eur. J. Biochem.* **1992**, *210*, 911–921.
- (10) Kiptoo, P.; Büyüktimkin, B.; Badawi, A. H.; Stewart, J.; Ridwan, R.; Siahaan, T. J. *Clin. Exp. Immunol.* **2013**, *172*, 23–36.
- (11) Mulder, G. E.; Kruijtzter, J. A. W.; Liskamp, R. M. J. *Chem. Commun.* **2012**, *48*, 10007–10009.
- (12) Güell, I.; Vilà, S.; Micaló, L.; Badosa, E.; Montesinos, E.; Planas, M.; Feliu, L. *Eur. J. Org. Chem.* **2013**, 4933–4943.
- (13) Pal, S.; Mitra, K.; Azmi, S.; Ghosh, J. K.; Chakraborty, T. K. *Org. Biomol. Chem.* **2011**, *9*, 4806–4810.
- (14) Guo, J.; Hu, H.; Zhao, Q.; Wang, T.; Zou, Y.; Yu, S.; Wu, Q.; Guo, Z. *ChemMedChem* **2012**, *7*, 1496–1503.
- (15) Shirazi, A. N.; Tiwari, R.; Chhikara, B. S.; Mandal, D.; Parang, K. *Mol. Pharmacol.* **2013**, *10*, 488–499.
- (16) Chapman, R.; Bouten, P. J. M.; Hoogenboom, R.; Jolliffe, K. A.; Perrier, S. *Chem. Commun.* **2013**, *49*, 6522–6524.
- (17) Chapman, R.; Koh, M. L.; Warr, G. G.; Jolliffe, K. A.; Perrier, S. *Chem. Sci.* **2013**, *4*, 2581–2589.
- (18) Cai, J.; Rosenzweig, B. A.; Hamilton, A. D. *Chem.—Eur. J.* **2009**, *15*, 328–332.
- (19) Ghosh, P. S.; Hamilton, A. D. *J. Am. Chem. Soc.* **2012**, *134*, 13208–13211.

- (20) Alam, Md. R.; Ming, X.; Fisher, M.; Lackey, J.; Rajeev, K. G.; Manoharan, M.; Juliano, R. *Bioconjugate Chem.* **2011**, *22*, 1673–1681.
- (21) Aviñó, A.; Gómara, M. J.; Malakoutikhah, M.; Haro, I.; Eritja, R. *Molecules* **2012**, *17*, 13825–13843.
- (22) Elduque, X.; Pedroso, E.; Grandas, A. *Org. Lett.* **2013**, *15*, 2038–2041.
- (23) Elduque, X.; Sánchez, A.; Sharma, K.; Pedroso, E.; Grandas, A. *Bioconjugate Chem.* **2013**, *24*, 832–839.
- (24) Staros, J. V.; Bayley, H.; Standring, D. N.; Knowles, J. R. *Biochem. Biophys. Res. Commun.* **1978**, *80*, 568–572.
- (25) Handlon, A. L.; Oppenheimer, N. J. *Pharm. Res.* **1988**, *5*, 297–299.
- (26) Marchán, V.; Ortega, S.; Pulido, D.; Pedroso, E.; Grandas, A. *Nucleic Acids Res.* **2006**, *34*, e24.
- (27) Sánchez, A.; Pedroso, E.; Grandas, A. *Org. Lett.* **2011**, *13*, 4364–4367.
- (28) Ponsati, B.; Giralt, E.; Andreu, D. *Tetrahedron* **1990**, *46*, 8255–8266.
- (29) Zuckermann, R. N.; Kerr, J. M.; Kent, S. B. H.; Moos, W. H. J. *Am. Chem. Soc.* **1992**, *114*, 10646–10647.
- (30) Ma, B.; Li, G.; Kanter, P.; Lamonica, D.; Grossman, Z.; Pandey, R. K. *J. Porphyrins Phthalocyanines* **2003**, *7*, 500–507.
- (31) Riva, E.; Comi, D.; Borrelli, S.; Colombo, F.; Danieli, B.; Borlak, J.; Evensen, L.; Lorens, J. B.; Fontana, G.; Gia, O. M.; Dalla Via, L.; Passarella, D. *Bioorg. Med. Chem.* **2010**, *18*, 8660–8668.
- (32) Fauster, K.; Hartl, M.; Santner, T.; Aigner, M.; Kreutz, C.; Bister, K.; Ennifar, E.; Micura, R. *ACS Chem. Biol.* **2012**, *7*, 581–589.
- (33) Pourceau, G.; Meyer, A.; Vasseur, J.-J.; Morvan, F. *J. Org. Chem.* **2009**, *74*, 6837–6842.
- (34) Kiviniemi, A.; Virta, P.; Lönnberg, H. *Bioconjugate Chem.* **2008**, *19*, 1726–1734.
- (35) Wada, T.; Mochizuki, A.; Higashiya, S.; Tsuruoka, H.; Kawahara, S.; Ishikawa, M.; Sekine, M. *Tetrahedron Lett.* **2001**, *42*, 9215–9219.
- (36) Coppola, C.; Simeone, L.; De Napoli, L.; Montesarchio, D. *Eur. J. Org. Chem.* **2011**, 1155–1165.
- (37) Putta, M. R.; Zhu, F.-G.; Wang, D.; Bhagat, L.; Dai, M.; Kandimalla, E. R.; Agrawal, S. *Bioconjugate Chem.* **2010**, *21*, 39–45.
- (38) Shen, B.-Q.; et al. *Nat. Biotechnol.* **2012**, *30*, 184–189.
- (39) Monsó, M.; de la Torre, B. G.; Blanco, E.; Moreno, N.; Andreu, D. *Bioconjugate Chem.* **2013**, *24*, 578–585.



## SUPPORTING INFORMATION

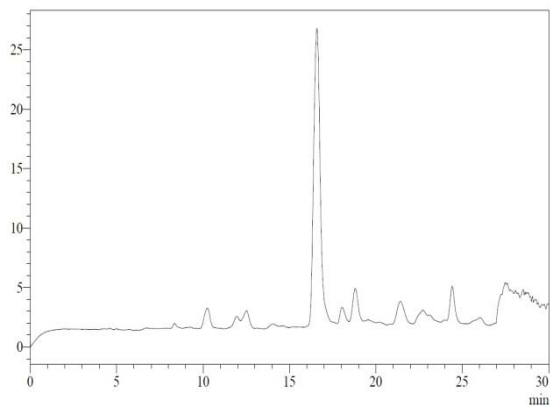
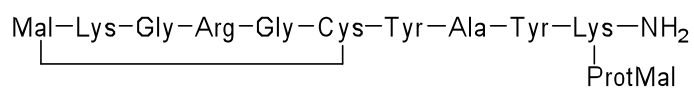
### **Orthogonal protection of peptides and peptoids for cyclization by the thiol-ene reaction and conjugation**

by Xavier Elduque, Enrique Pedroso and Anna Grandas\*

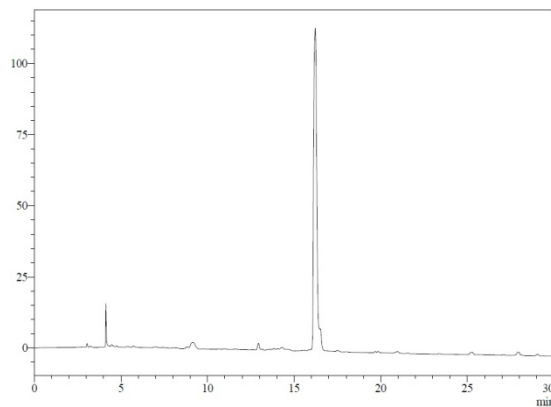
	Page
1. HPLC profiles of the products prepared.	2
3. Commercially available derivatives used for peptide and peptoid synthesis.	12
3. NMR Spectra of compounds prepared following described procedures.	13

## 1. HPLC profiles of the products prepared.

### Cyclic peptide 2:

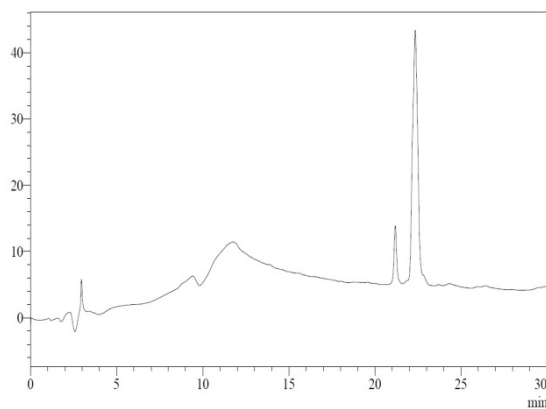
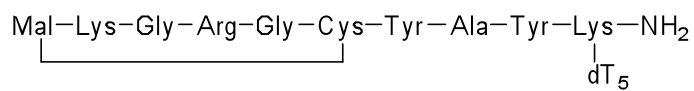


Crude



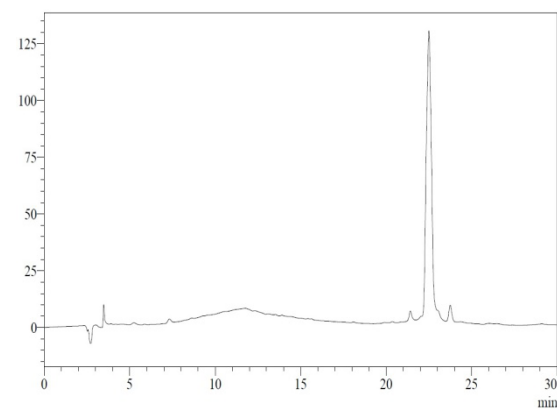
Purified

### Conjugate 3:



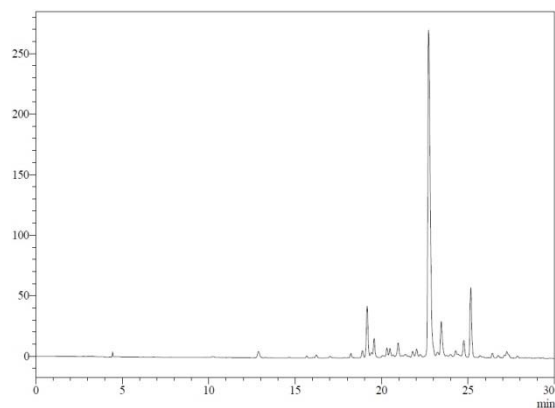
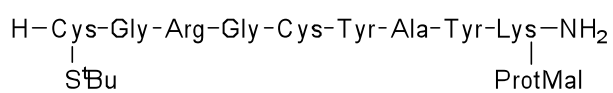
Crude

The peak eluting at 21.2 min corresponds to unreacted diene-dT<sub>5</sub>

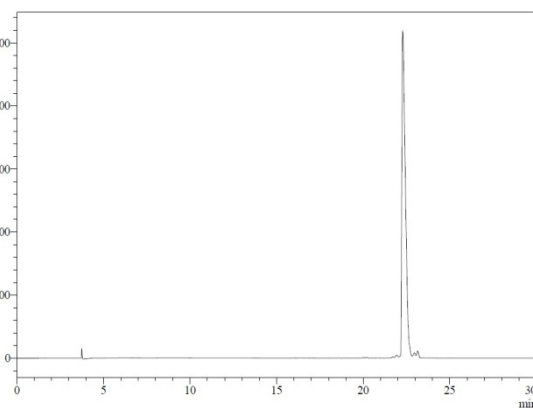


Purified

## Peptide 5:

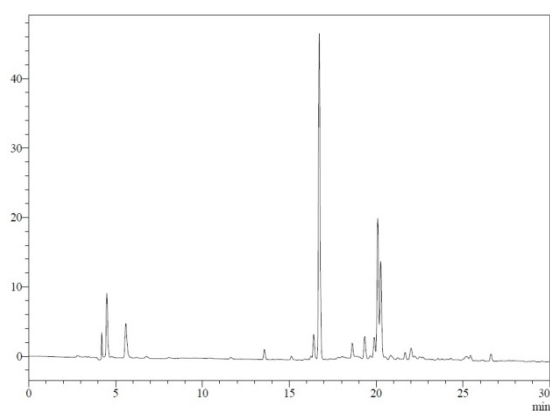


Crude



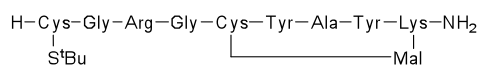
Purified

## Maleimide deprotection of peptide 5:

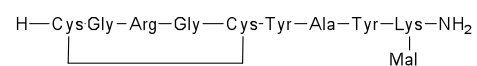


Crude after heating for 90 min

Peaks with  $t_R=16.7$  and  $20.1$  correspond to compounds **7** and **6**, respectively

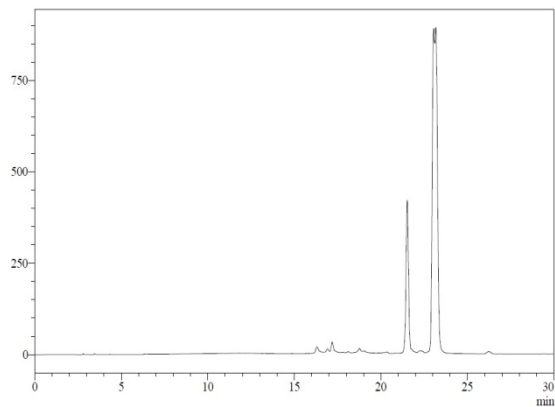
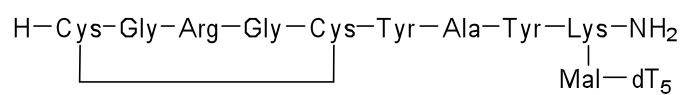


**6**



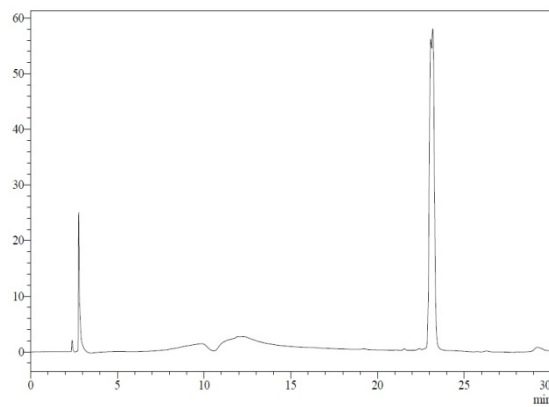
**7**

### Conjugate 8:



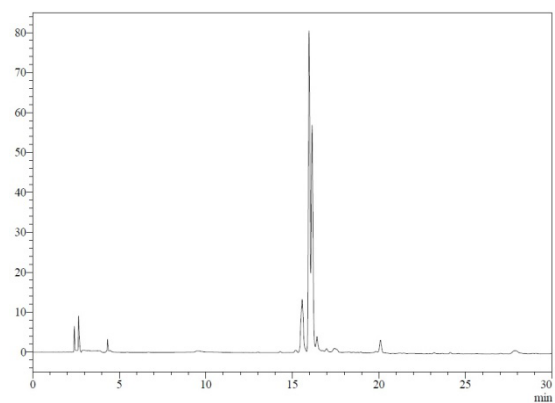
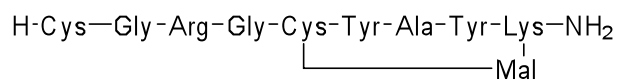
**Crude 8**

Conjugate 8 elutes at 23.0 min. The peak at 22.3 corresponds to unreacted diene-dT<sub>5</sub>.



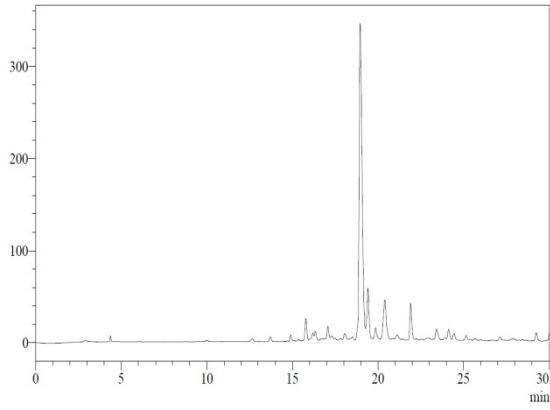
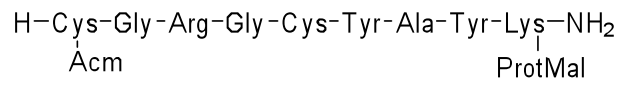
**Purified 8**

### Cyclic peptide 16 (from 6):

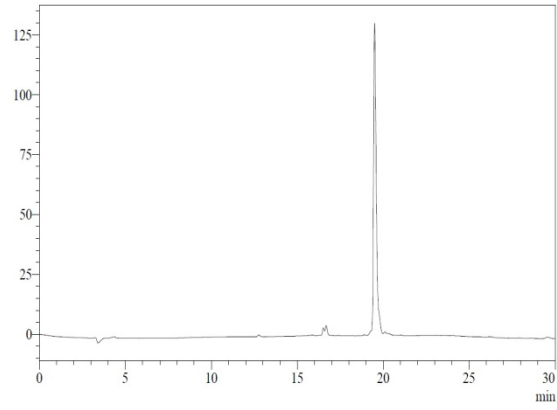


**Crude**

**Peptide 10:**

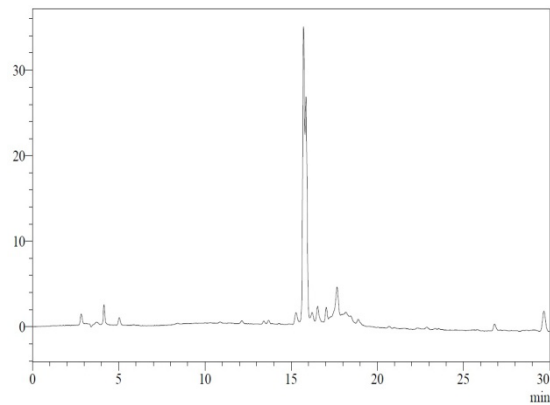
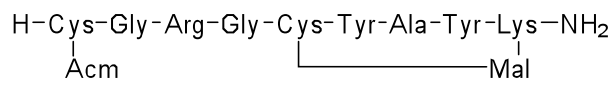


Crude

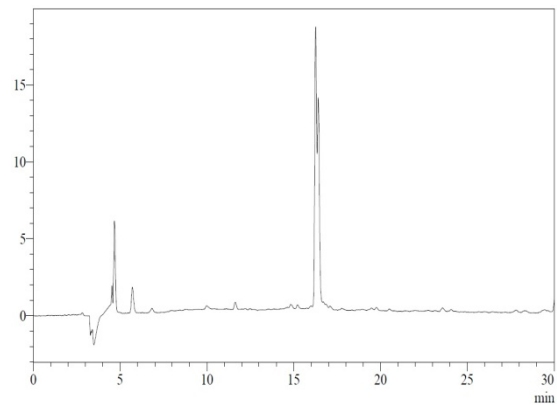


Purified

**Cyclic peptide 11:**



Crude

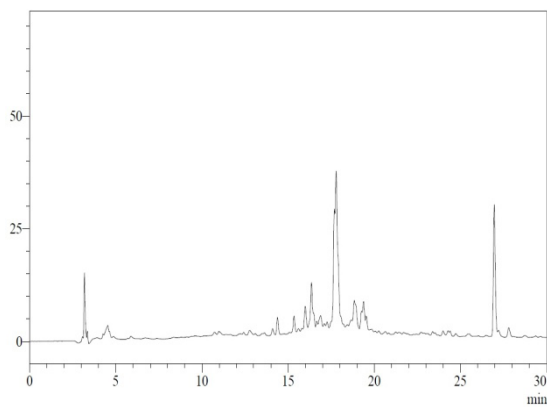
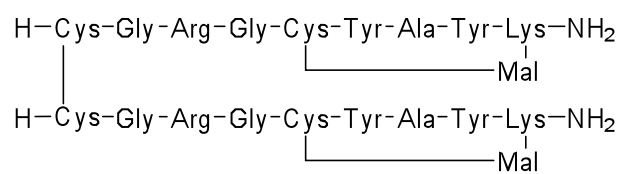


Purified



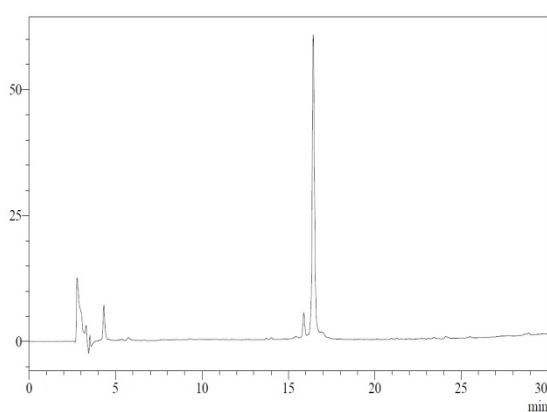
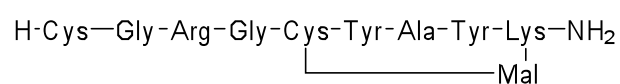


**Compound 15, disulfide dimer of cyclic peptide 16:**

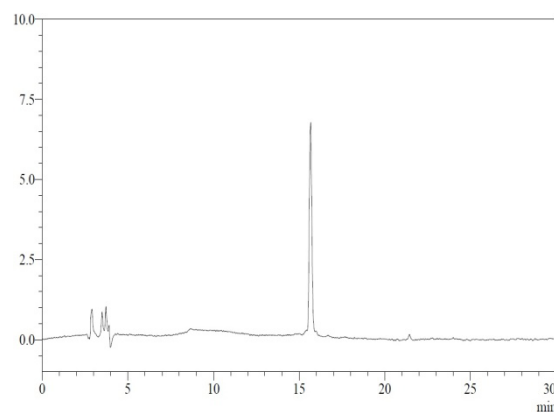


Crude

**Cyclic peptide 16 (from 15, after reduction of the disulfide):**

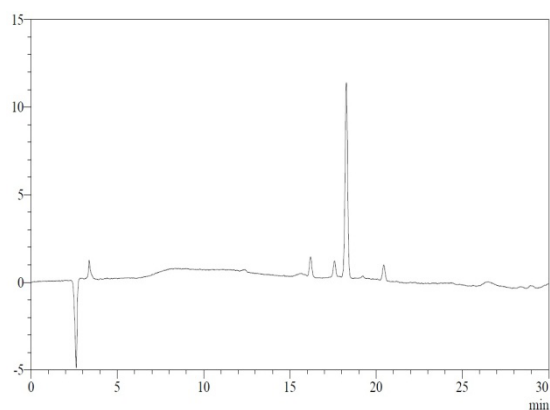


Crude



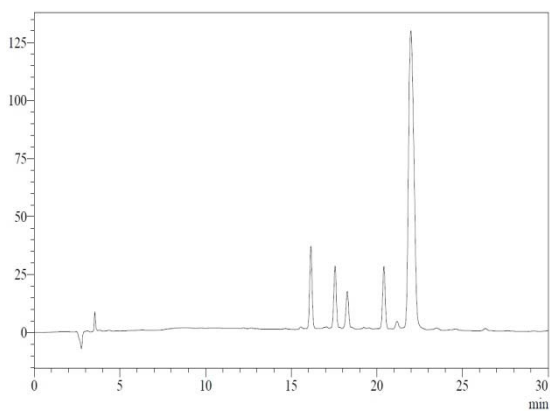
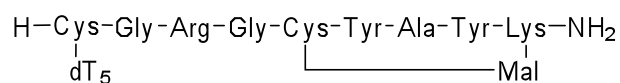
Purified

**Mal-dT<sub>5</sub> (from protected maleimido-dT<sub>5</sub>):**

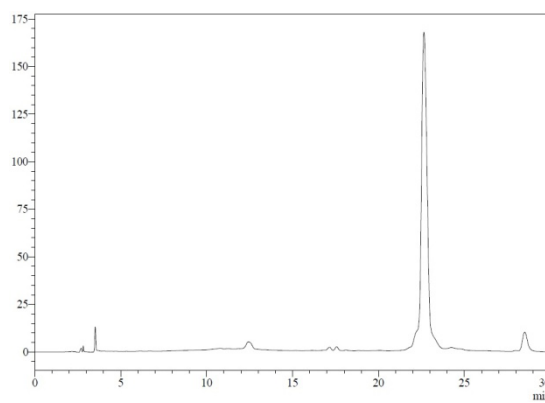


**Crude Mal-dT<sub>5</sub>**

**Conjugate 17:**



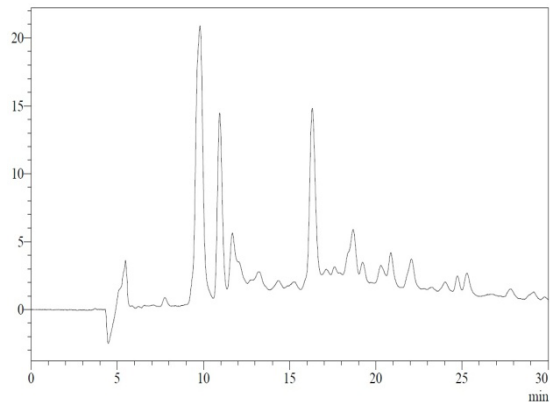
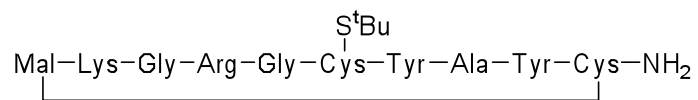
**Crude 17**



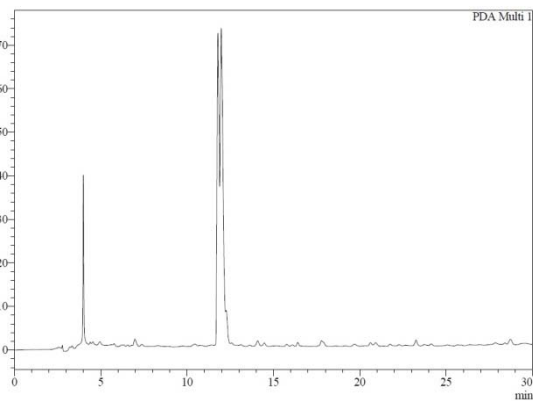
**Purified 17**

Conjugate **17** elutes at 22.0 min. The small peak at 18.3 corresponds to unreacted Mal-dT<sub>5</sub>, and the peaks at 16.1, 17.6 and 20.4 to byproducts already detected in the Mal-dT<sub>5</sub> crude (compound **16** is not detected).

### Cyclic peptide 19:

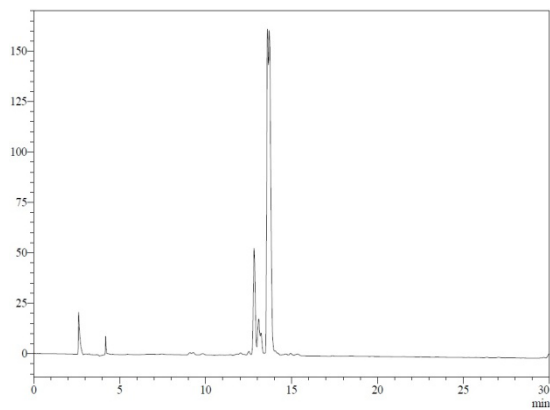
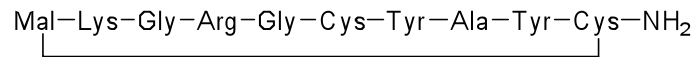


Crude  
(analyzed under purification conditions)

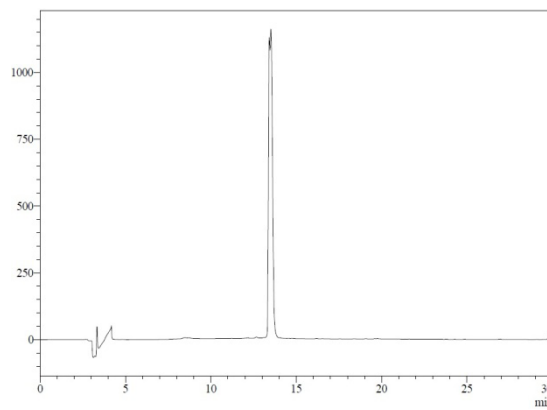


Purified

### Cyclic peptide 20:

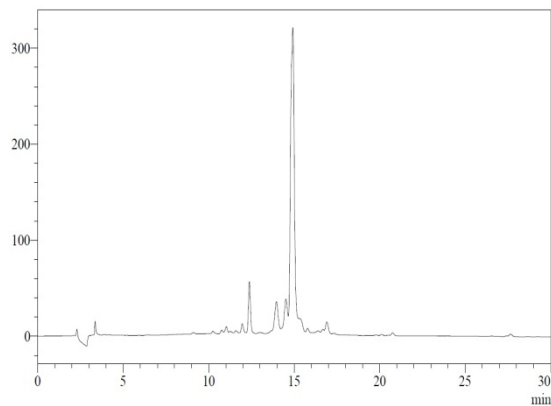
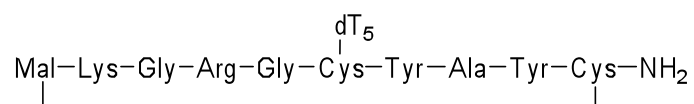


Crude



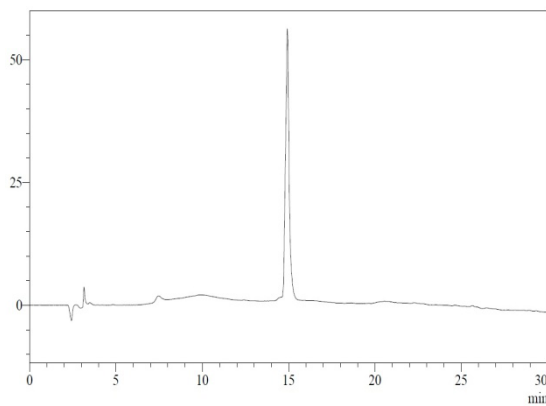
Purified

## Conjugate 21:



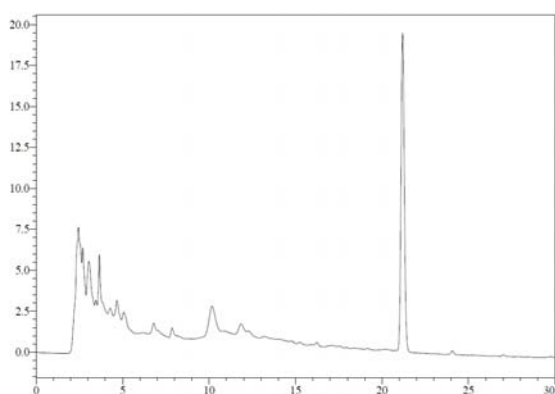
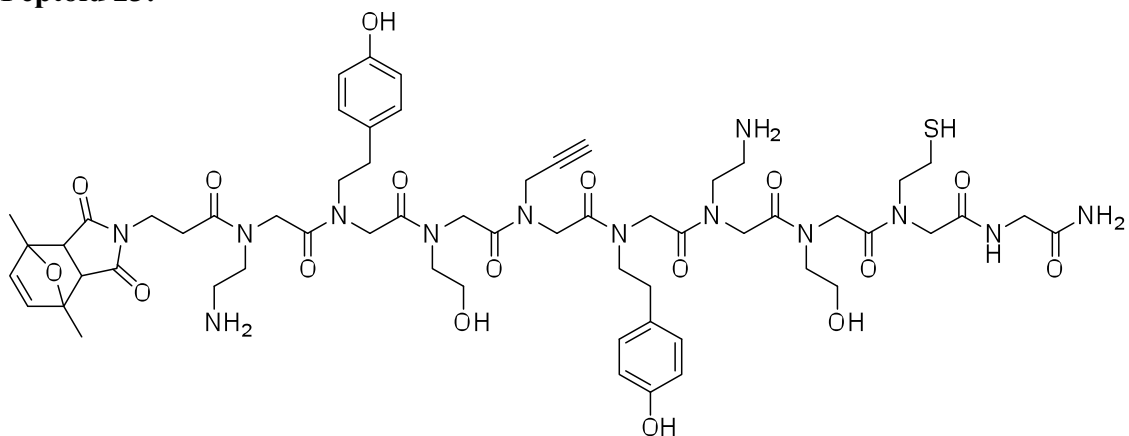
Crude

The peak at 12.4 min corresponds to deprotected maleimide-dT<sub>5</sub>



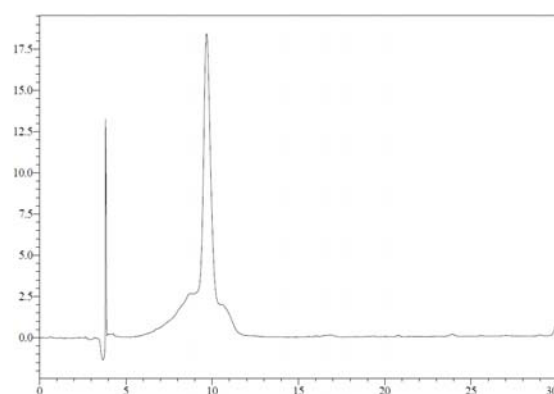
Purified

### Peptoid 23:



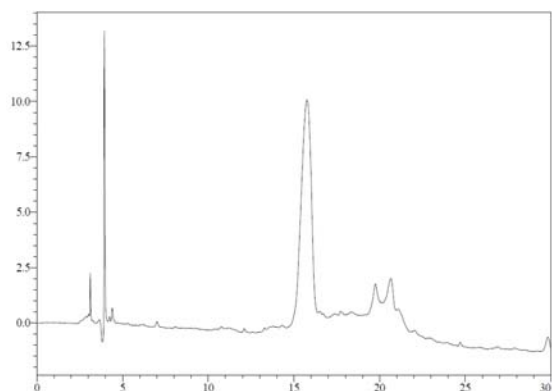
Crude

The peak eluting at 21 min corresponds to *m*-cresol

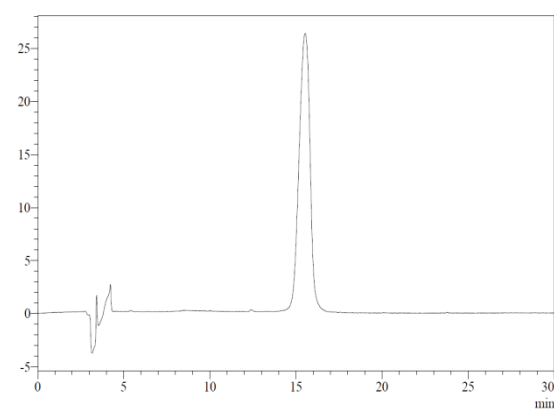


Purified

### Cyclic peptoid 24:

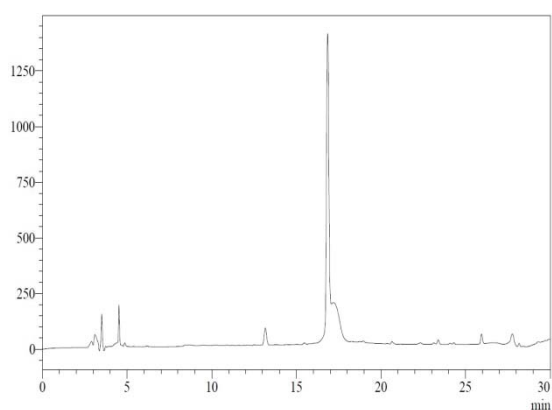
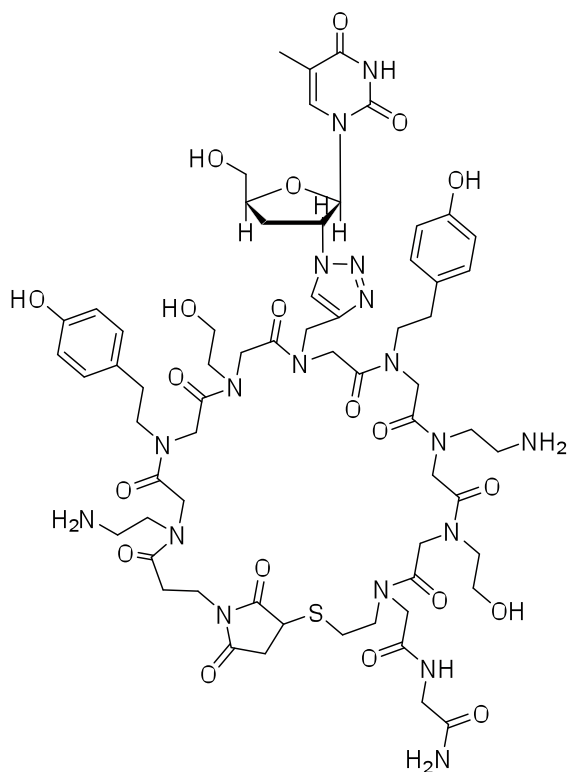


Crude



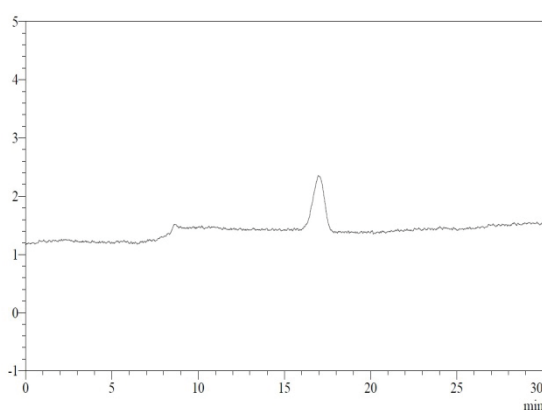
Purified

## Conjugate 25:



Crude

The sharp peak at 16.9 min corresponds to the excess of AZT, and was not taken in consideration for yield calculations

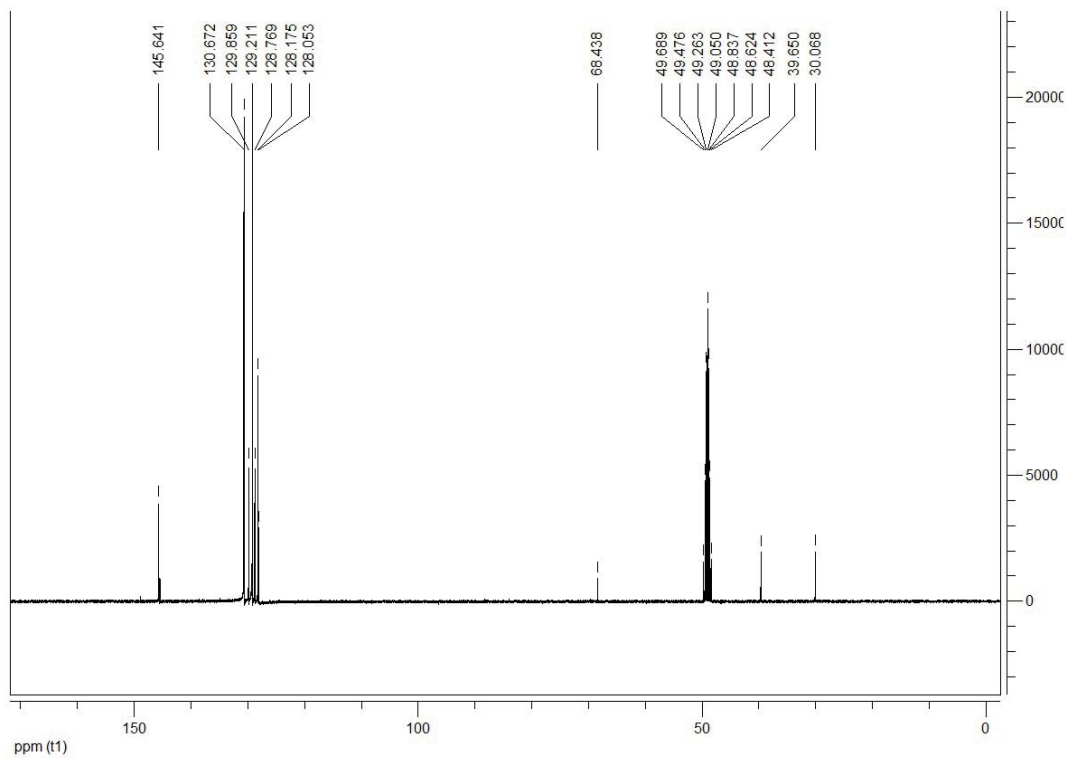
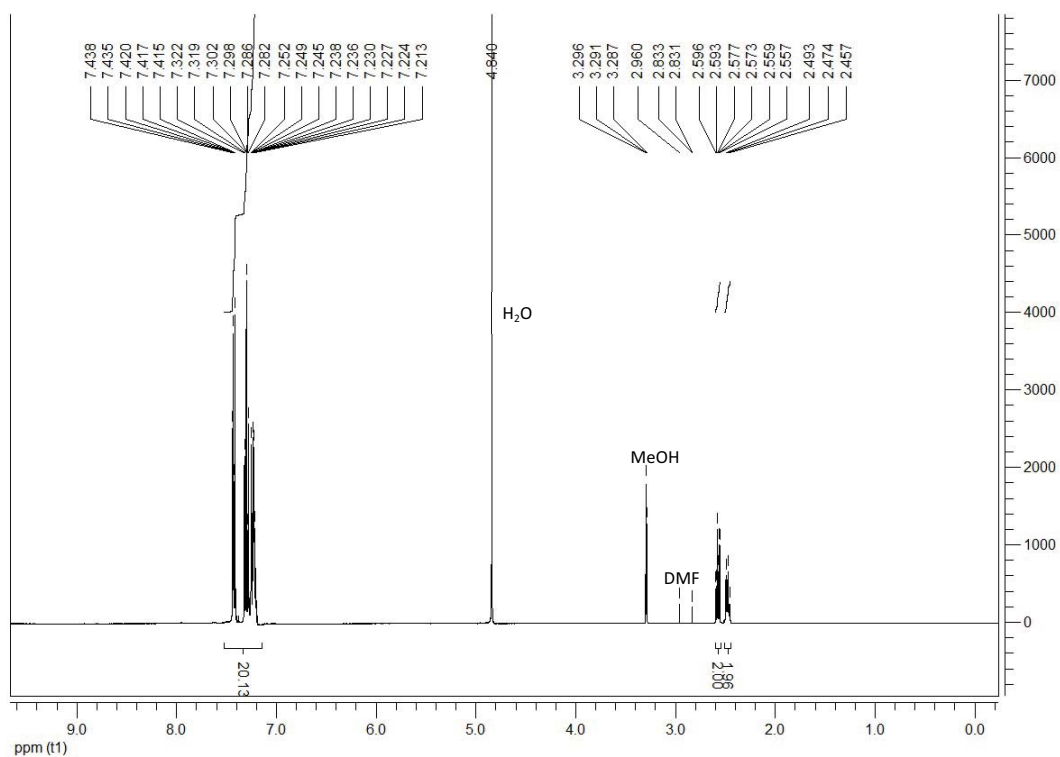


Purified

**2. Commercially available derivatives used for peptide and peptoid synthesis:** 3-maleimidopropanoic acid, Fmoc-Gly-OH, Fmoc-Ala-OH, Fmoc-Tyr(<sup>t</sup>Bu)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Cys(Acm)-OH, Fmoc-Cys(Trt)-OH, Fmoc-Cys(S<sup>t</sup>Bu)-OH, Boc-Cys(Fm)-OH, 4-(2-aminoethyl)phenol, 2-aminoethanol, 2-propyn-1-amine.

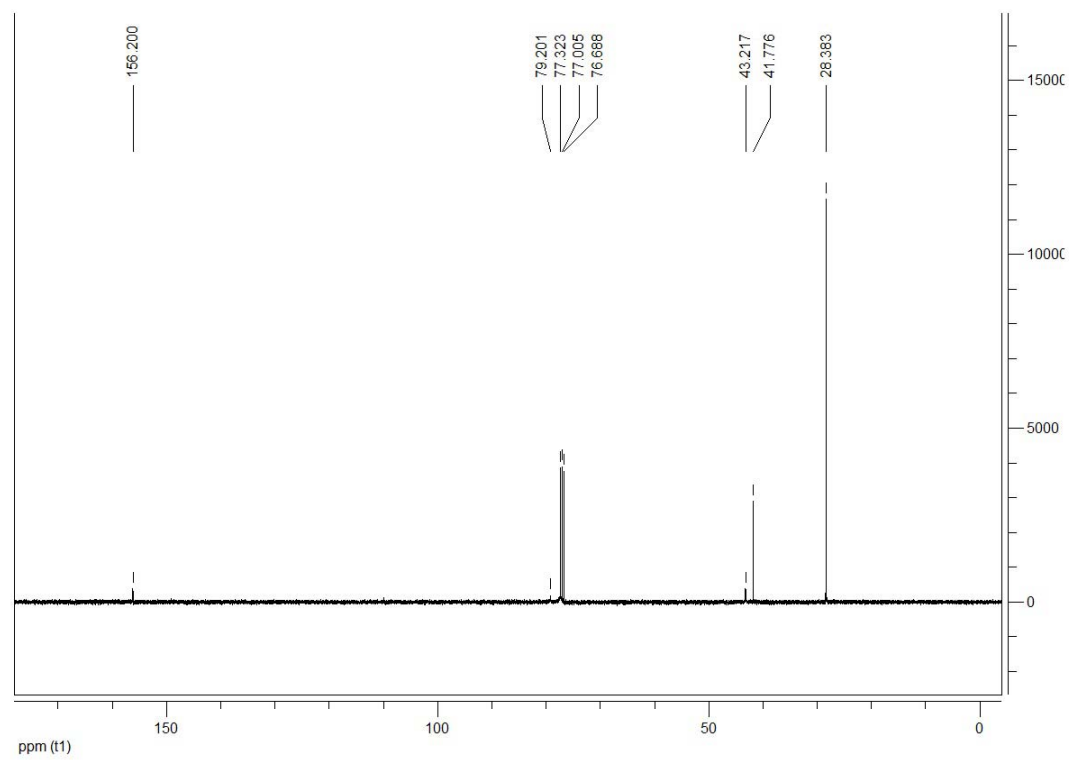
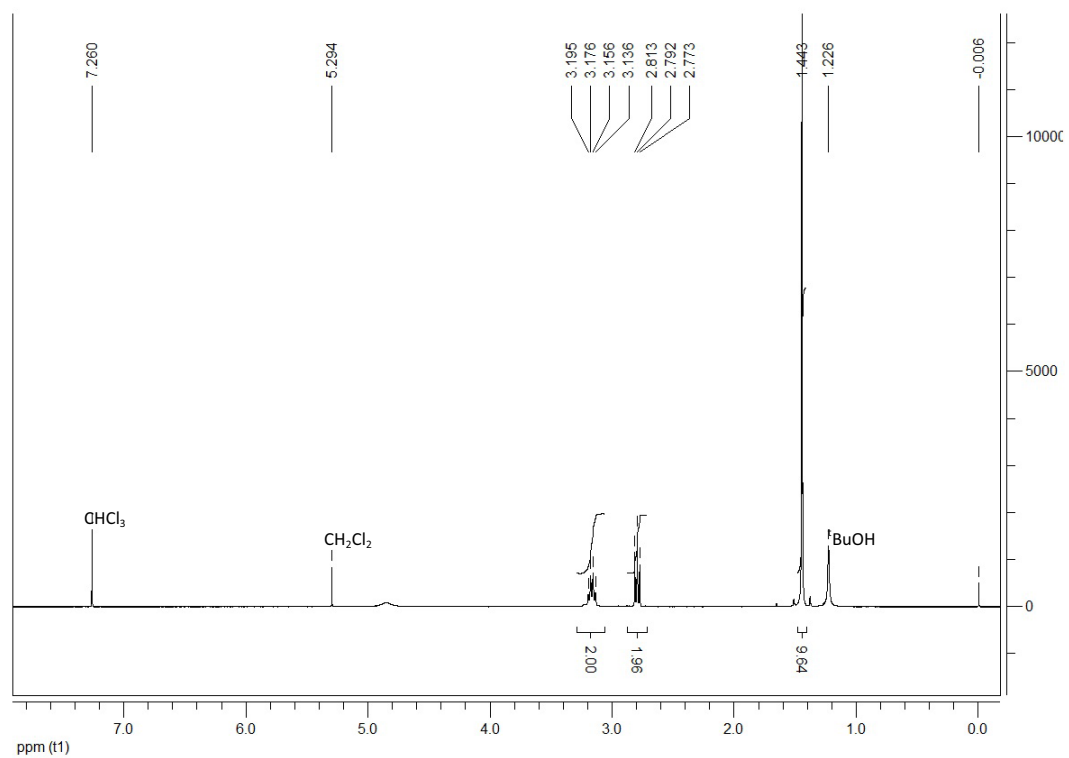
3.  $^1\text{H}$ - and  $^{13}\text{C}$ -NMR spectra (400 and 100.56 MHz, respectively) of compounds prepared following described procedures:

*S*-Tritylcysteamine hydrochloride (solvent:  $\text{CH}_3\text{OD}$ )





***N*-Boc-1,2-ethandiamine** (solvent: CDCl<sub>3</sub>)



## Appendix 4: Materials and Methods

### A4.1 Reagents

Reagents for maleimide monomers synthesis and primary amines synthesis for peptoid assembly are listed in Paper I (*Bioconjugate Chemistry*) and Paper III (*J. Org. Chem.*), respectively. 3,6-di(2-pyridyl)-1,2,4,5-tetrazine and all the reagents for the synthesis of AD2 were purchased from Sigma-Aldrich.

Oligonucleotide synthesis reagents are described in Paper I (*Bioconjugate Chemistry*). For RNA oligonucleotide synthesis, phosphoramidites and reagents were purchased from either Glen Research or Link Technologies. All the material was sterile, either purchased from VWR or sterilized in an autoclave. Fluorescently-labeled RNA sequences were purchased from Integrated DNA Technologies (IDT).

Peptide synthesis reagents are described in Paper I (*Bioconjugate Chemistry*), Paper II (*Organic Letters*) and Paper III (*J. Org. Chem.*).

#### Buffer preparation

Phosphate buffer: To prepare a pH = 6.8 phosphate buffer solution (100 mM, 100 mL), a 0.2 M solution of Na<sub>2</sub>HPO<sub>4</sub> (872 mg of salt in 24.5 mL of water) was mixed with a 0.2 M solution of NaH<sub>2</sub>PO<sub>4</sub> (796 mg of salt in 25.5 mL of water), and water was added up to 100 mL.

Acetic acid / Sodium acetate buffer: To prepare a pH = 4.75 acetic acid/acetate buffer solution (100 mM, 50 mL), acetic acid (286 µL) was dissolved in water, the pH was adjusted with NaOH and water was added up to 50 mL.

### A4.2 Instruments

#### NMR spectroscopy

NMR spectra were recorded in a Varian Mercury 400 MHz instrument. 1-10 mg of sample was dissolved in 0.7 mL of the appropriate deuterated solvent, CDCl<sub>3</sub>, DMSO-d<sub>6</sub>, acetone-d<sub>6</sub> or CD<sub>3</sub>OD depending on the sample.

#### Mass spectrometry

ESI and MALDI-TOF mass spectrometry instruments are detailed in Paper I (*Bioconjugate Chemistry*) and Paper II (*Organic Letters*). For **AD2** stoichiometry determination assays a Synapt G1 HDMS (Waters, Manchester, UK) mass spectrometer with direct infusion was employed.

## Appendix 4: Materials and Methods

### Spectrophotometry

UV absorption was measured in a Jasco V-550 spectrophotometer with quartz cells from Hellma. A Jasco ETC-505T Peltier was employed for temperature control UV-monitored thermal denaturation experiments.

Circular dichroism spectra were recorded in a Jasco-J810.

Spectra of fluorescent RNAs (titration with **AD2**) were recorded in a PTI instrument with a LPS-220B lamp.

### Other instruments

Suspensions were centrifuged in an Eppendorf centrifuge 5430R.

Labware was sterilized in an Autester ST P Selecta autoclave.

RNA was synthesized in a 3400 ABI automatic synthesizer.

Samples were lyophilized in a Virtis Freezemobile lyophilizer or in a Labconco lyophilizer.

ITC experiments were performed in a NanoITC instrument from TA instruments.

The microwave oven employed was a Biotage Initiator. For MW-promoted reactions sealed vials of volumes ranging from 0.2 mL to 20 mL were used.

pHs were measured with a Crison pHmeter Basic 20.

### HPLC

Analytical and preparative HPLCs were performed in a Shimadzu 20A system with either a SPD-M20A PDA detector or a SPD-20A double wavelength detector. A CTO-10AS oven was employed when heating was required.

Gradients, flows, solvents and columns depend on the case and are listed in the experimental section of Paper I (*Bioconjugate Chemistry*), Paper II (*Organic Letters*) and Paper III (*J. Org. Chem.*).

Unless otherwise stated, all HPLC chromatograms were performed with acidic elution conditions (Analytical: A: H<sub>2</sub>O + 0.045 % TFA, B: ACN + 0.036 % TFA; Preparative: A: H<sub>2</sub>O + 0.1 % TFA, B: ACN + 0.1 % TFA).

## **A4.3 Methods**

### Polyamide synthesis and quantification

General protocols for peptide synthesis and quantification are described in Paper I (*Bioconjugate Chemistry*), Paper II (*Organic Letters*) and Paper III (*J. Org. Chem.*). Cleavage cocktail composition and cleavage time vary slightly depending on the substrate.

A general protocol for peptoid synthesis and quantification is described in Paper I (*Bioconjugate Chemistry*) and Paper III (*J. Org. Chem.*).

Cyclization assessment protocols are described in Paper III (*J. Org. Chem.*).

All polyamides were elongated manually on polyethylene syringes provided with a polypropylene filter.

#### Oligonucleotide synthesis and quantification

DNA synthesis and quantification is described in Paper I (*Bioconjugate Chemistry*).

#### UV-monitored thermal denaturation experiments

The sample solution contained RNA (5  $\mu\text{M}$ ), phosphate buffer (10 mM), pH = 6.8, and NaCl (100 mM).

RNA was dissolved in water (800  $\mu\text{L}$ ), and 100 mM phosphate buffer (100  $\mu\text{L}$ ) and 1M NaCl (100  $\mu\text{L}$ ) were added. 250  $\mu\text{L}$  of this solution were transferred to a 0.1 cm path length cell.

Provided with a continuous flow of  $\text{N}_2$  to avoid condensation upon cooling, the cell was kept at 90°C for 10 minutes and then cooled from 90°C to 20°C and heated from 20°C to 90°C at a 0.5°C/min rate (140 min/experiment). The absorption was measured continuously at 260 nm.

The sample was titrated with AD2, adding aliquots from a 0.5 mM AD2 solution in water. Several experiments were performed with 0, 0.2, 0.5, 1 and 2 equivalents of AD2.

The data were adjusted with the Origin® software to a sigmoid curve, and the inflexion point was considered the melting temperature.

#### Circular dichroism experiments

The sample solution contained RNA (2  $\mu\text{M}$ ), phosphate buffer (10 mM) at pH = 6.8 and NaCl (100 mM).

RNA was dissolved in water (800  $\mu\text{L}$ ), and 100 mM phosphate buffer (100  $\mu\text{L}$ ) and 1M NaCl (100  $\mu\text{L}$ ) were added. RNA was annealed by heating this solution in a sealed eppendorf at 90°C in a sand bath and allowing it to temper. The solution was transferred into a 1 cm path length cell.

The scans ranged from 220 to 320 nm at a rate of 100 nm/min, and each scan was repeated 8 times.

The sample was titrated with AD2, adding aliquots from a 0.5 mM AD2 solution in water. Several experiments were performed with 0, 0.2, 0.5, 1 and 2 equivalents of AD2.

The data were processed by converting it to molar ellipticity and plotting it against wavelength.

## Appendix 4: Materials and Methods

### Isothermal calorimetry titrations

A 10  $\mu\text{M}$  solution of RNA and a 500  $\mu\text{M}$  solution of **AD2** in 1 mM phosphate buffer were prepared (this solution was previously degassed at 25°C for 5 minutes under reduced pressure). The RNA was annealed as explained for circular dichroism experiments.

The cell was cleaned and dried, and the RNA solution (300  $\mu\text{L}$ ) was introduced carefully. The reference cell was filled with the same phosphate buffer solution. The syringe was filled with the **AD2** solution (50  $\mu\text{L}$ ), avoiding the formation of bubbles, and was coupled to the burette and introduced into the microcalorimeter. The experiment was performed at 25°C and 40°C with continuous stirring at 250 rpm. The ligand was added over 20 injections of 2.5  $\mu\text{L}$  each, with an idle time of 5 minutes between injections.

The data were adjusted with the instrument software to different binding modes.

### Fluorescence experiments with fluorescently-labeled RNA

Aminopurine-labeled RNA (5 nmol) was dissolved in 100 mM phosphate buffer (1 mL), and an aliquot was diluted with water to yield a 0.5  $\mu\text{M}$  RNA solution in 10 mM phosphate buffer, pH = 6.8. RNA was annealed by heating this solution in a sealed eppendorf at 90°C in a sand bath and allowing it to temper. The solution was transferred into a 1 cm path length fluorescence cell, and introduced in the fluorimeter. After irradiating at 310 nm, emission spectra were recorded from 330 nm to 420 nm. All the slits were set at 1.5 nm width.

Titration were performed by addition of aliquots from different **AD2** solutions in intervals of approx. 5 minutes as shown in the table below.

[AD2] stock ( $\mu\text{M}$ )	$\mu\text{L}$ added	Total equivalents	Total nmol
6.25	2	0.025	0.0125
6.25	2	0.05	0.025
6.25	2	0.075	0.0375
6.25	2	0.1	0.05
12.5	2	0.15	0.075
12.5	2	0.2	0.1
12.5	2	0.25	0.125
12.5	2	0.3	0.15
25	2	0.4	0.2
50	2	0.6	0.3
50	2	0.8	0.4
50	2	1	0.5
50	2	1.2	0.6
50	2	1.4	0.7
50	2	1.6	0.8
50	2	1.8	0.9
50	2	2	1
125	2	2.5	1.25
125	2	3	1.5
125	2	3.5	1.75
125	2	4	2
500	1	5	2.5

500	1	6	3
500	1	7	3.5
500	1	8	4
500	1	9	4.5
500	1	10	5
500	2	12	6
500	2	14	7
500	2	16	8
500	2	18	9
500	2	20	10
500	2	25	12.5
500	2	30	15
500	2	35	17.5
500	2	40	20
500	2	45	22.5
500	2	50	25

The fluorescence intensity was corrected by multiplying by a dilution factor, a fraction of the real volume divided by the initial volume.

The data were adjusted with the Origin® software to a sigmoid curve, by plotting the fluorescence intensity versus the logarithm of **AD2** concentration. The value at x-axis at the sigmoid inflection point was the logarithm of EC<sub>50</sub>.



## Appendix 5: AD2 and quinacrine binding to ct-DNA

During the three-month stay at Prof. Lincoln's lab (Chalmers University of Technology, Gothenburg) part of my time was dedicated to study the interaction of **AD2** with calf thymus DNA (ct-DNA). We first examined its binding mode with linear dichroism experiments and then we determined the binding stoichiometry and the affinity constant with fluorescence competition experiments between quinacrine and **AD2** (Fig. A5.1).

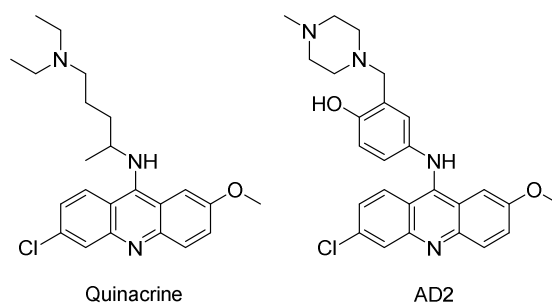


Figure A5.1: Quinacrine and **AD2** structures

### Linear dichroism experiments

Linear dichroism is a spectroscopy-based technique in which an oriented sample is irradiated with polarized light beams parallel and perpendicular to a reference axis (in our case, that of the DNA chain). The absorption values of the perpendicular light are subtracted from the values of the parallel one, to obtain the final LD spectrum.<sup>1</sup>

$$LD = A_{\parallel} - A_{\perp}$$

One of the requirements for linear dichroism measurements is sample orientation. There are several ways to orient the molecules, such as electric fields or stretched polymers. The method of choice was the flow orientation with Couette cells. These cells consist of two concentric quartz cylinders with a narrow space between them where the sample is placed. Rotation of the inner cylinder generates a continuous flow that orientates long molecules like DNA chains perpendicularly to the rotation axis (Fig A5.2).

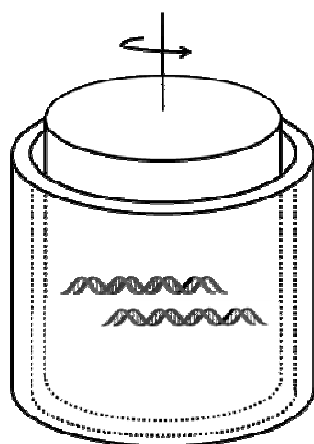


Figure A5.2: Schematic representation of a Couette cell



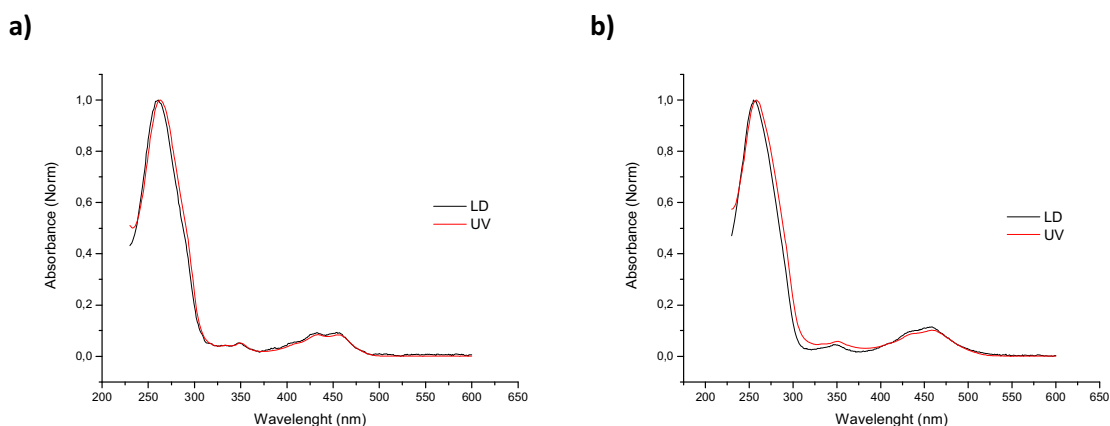
## Appendix 5: AD2 and Quinacrine experiments with ct-DNA

Light absorption is much higher when the incident light is oscillating parallel to the transition dipole moment of the sample than when it is perpendicular. For acridines, the transition dipole moment is parallel to the axis of the three ring structure. If the LD values obtained are negative, this means that the absorption is higher when the light is polarized perpendicularly to the reference axis ( $A_{\perp}$ ), hence the transition dipole moment is parallel to  $A_{\perp}$  and perpendicular to the reference axis. In our case, given that the transition dipole moment is placed alongside the acridine, it is safe to assume that if the signal is negative, the acridine is perpendicular to the DNA helix, most likely intercalated.

We performed LD experiments with ct-DNA alone, ct-DNA with quinacrine (5 base-pairs : 1 quinacrine molecule) and ct-DNA with **AD2** (5 base-pairs : 1 **AD2** molecule) (for DNA quantification, see experimental section in this appendix). At this ratio we assumed that all DNA binding sites were occupied by either quinacrine or **AD2**. Even though there was an excess of ligand, the ligand would be too small to be oriented and thus invisible in LD experiments. We performed the experiments with quinacrine mainly to assess that the binding mode is the same both for quinacrine and **AD2**, and thus to validate the employment of quinacrine for later fluorescence competition assays.

The LD spectra of ct-DNA showed a nice negative band. This makes sense, since the bases, the part that absorbs light of the DNA, are perpendicular to the reference axis, this is, the oligonucleotide chain.

LD spectra of ct-DNA in the presence of both quinacrine and **AD2** showed the negative band of the DNA ( $\lambda \approx 260$  nm), and also other negative bands at longer wavelengths, corresponding to the acridine moieties ( $\lambda \approx 450$  nm).



**Figure A5.3:** Superimposition of UV and LD spectra for **a)** quinacrine and **b)** **AD2**. Both LD and UV spectra are normalized, and LD spectrum is multiplied by -1

These data indicated that the acridines were sitting essentially perpendicular to the reference axis, namely the DNA sequence. In order to further assess this conclusion, we compared the linear dichroism spectra with the regular UV absorption spectra. We found that by multiplying the LD spectra by -1 and normalizing both UV and LD spectra all the spectra matched perfectly, both ct-DNA and quinacrine or **AD2** bands, meaning that the orientation of the DNA bases and the acridines is the same, and assessing an intercalative binding mode (**Fig A5.3**).

### Fluorescence qualitative results. Quinacrine and AD2 displacement

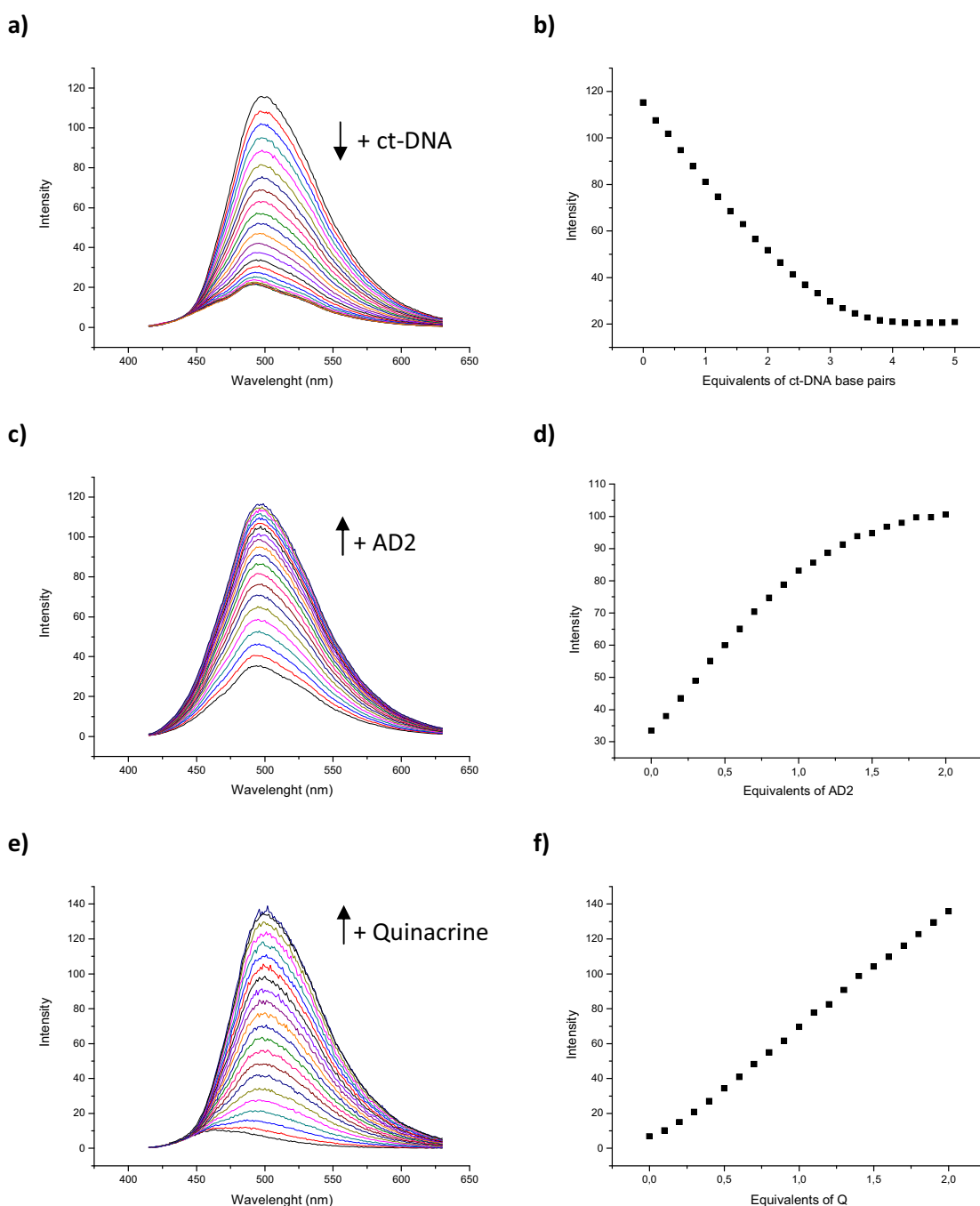
Quinacrine is a molecule that was employed as an antimalarial drug during the Second World War.<sup>2</sup> Its structure is very similar to that of **AD2**, but whereas **AD2** is not fluorescent, quinacrine emits nicely at 500 nm when irradiated at 400 nm. Quinacrine fluorescence is quenched when intercalated into DNA, so the quinacrine binding constant can be easily determined by a simple titration. We took advantage of the fact that quinacrine fluorescence changes upon binding, and that **AD2** and quinacrine share the same binding mode, to perform competition or displacement assays for **AD2** binding determination (**Fig. A5.4**).

We first evaluated the binding of quinacrine to ct-DNA. For this purpose we titrated a solution of quinacrine with increasing quantities of ct-DNA, up to 5 equivalents of base-pairs. We observed a nice decrease of the fluorescence, due to the fact that the free quinacrine intercalates progressively between the DNA base-pairs (**Fig. A5.4, a and b**).

In order to evaluate the binding of **AD2**, we performed competitive experiments with quinacrine. We prepared a sample with quinacrine and 3 equivalents of DNA base-pairs, and titrated with **AD2**. In this case we observed a fluorescence increase, meaning that **AD2** had a better affinity for DNA than quinacrine. **AD2** expelled quinacrine out of the DNA, thus unquenching quinacrine fluorescence (**Fig. A5.4, c and d**).

After the addition of 1 equivalent of **AD2** with respect to quinacrine almost all quinacrine was already expelled, and thus the fluorescence intensity did not further increase upon addition of more **AD2**.

We also performed this experiment the other way around, starting with **AD2** and 3 equivalents of ct-DNA base-pairs and adding quinacrine. In this case we observed a linear increase of the fluorescence, meaning that quinacrine does not expel **AD2** out of the DNA and that all the quinacrine added remains free (**Fig. A5.4, e and f**).

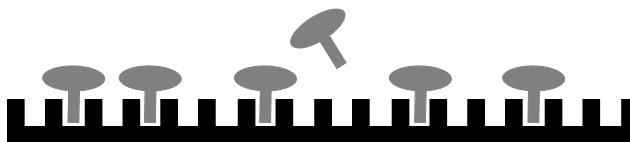


**Figure A5.4:** Quinacrine fluorescence upon addition of ct-DNA, **a)** the collection of spectra and **b)** the intensity values at  $\lambda = 500$  nm. Quinacrine – ct-DNA complex fluorescence upon addition of **AD2**, **c)** the collection of spectra and **d)** the values at  $\lambda = 500$  nm. **AD2** – ct-DNA complex fluorescence upon addition of quinacrine, **e)** the collection of spectra and **f)** the values at  $\lambda = 500$  nm

### Data processing with the McGhee-Von Hippel approach

The McGhee-Von Hippel approach for data processing is an extremely complicated matrix-based iterative probabilistic calculation. For its deep understanding, specialized literature is strongly recommended.<sup>3,4</sup> A simplified explanation of the basis of this approach is the following:

In the McGhee-Von Hippel approach the DNA chain is treated as a one-dimensional lattice of binding sites. These binding sites may not be accessible if the ligand occupies more than one base-pair when it is bound (**Figure A5.5**). Also, bound ligands may interact with each other, causing cooperative or anti-cooperative behaviors. Taking all this into account, considering the equilibria between the bound ligands and the pool of free ligands and mass balance equations, this approach calculates the fraction of bound ligands at each titration step.



**Figure A5.5:** Schematic representation of a DNA chain with ligands occupying more than one binding site

For the first data processing attempted we started by employing only the data from the first titration, ct-DNA over quinacrine, at 500 nm. We obtained a good fit with a stoichiometry near 2.5 and a submicromolar dissociation constant (**Table A5.1, Entry 1**). The experience of Prof. Lincoln's research group is that a 2.5 stoichiometry is common for intercalators, meaning that a ligand molecule binds every two or three base-pairs. The error value corresponds to the quadratic root of the summation of each point's quadratic error.

We next processed the data of the competition experiment at 500 nm, allowing for the optimization of the new constant and the stoichiometry of both quinacrine and **AD2**. The value obtained for quinacrine stoichiometry was the same as for the previous processing. We observed that the dissociation constant for **AD2** was extremely low, in the nanomolar range, in agreement with the qualitative results discussed before. The **AD2** stoichiometry value was 2.8, a reasonable value relatively close to the ideal 2.5 for intercalators (**Table A5.1, Entry 2**).

The third and fourth sets of data processing were performed by employing the entire fluorescence spectra instead of only one wavelength. In this way the values obtained should be affected by smaller errors, because as all the entire spectra are taken into account, the noise can be reduced. We performed a singular value decomposition and the first two singular values were considered. We obtained a  $K_Q$  of 414 nM, relatively similar to the value obtained at the first data processing (**Table A5.1, Entry 3**). The stoichiometry value obtained, 2.4, also matches with the previous one.

We finally processed the competition assay data also by performing a singular value decomposition and considering only the first two singular values. By fixing both the quinacrine dissociation constant at 414 nM and the stoichiometry at 2.3, the optimization of the other parameters produced the values stated in **Entry 4 (Table A5.1)**. We obtained a very similar stoichiometry value as for the previous competition data processing (**Table A5.1, Entry 2**), and even though the dissociation constant differs around 3 times from the previous one, it is still within the same order of magnitude.

	$K_Q$ (nM)	$n_Q$	$K_{AD2}$ (nM)	$n_{AD2}$	Error
Entry 1	303	2.3	-	-	16.4
Entry 2	303	2.3	2.9	2.8	7.5
Entry 3	414	2.4	-	-	59.3
Entry 4	414	2.3	9.4	3	39.8

**Table A5.1:** Data processing results from fluorescence experiments.  $K_Q$  and  $K_{AD2}$  are dissociation constants, and  $n$  is the number of base-pairs per ligand

## Summary

We have studied the interaction of **AD2** with ct-DNA by linear dichroism and fluorescence. Linear dichroism experiments indicated that **AD2** binds to DNA by intercalation, because the **AD2** signals in the spectrum suggest that it is placed perpendicularly to the DNA axis.

Stoichiometry and dissociation constants have been obtained from competition experiments between **AD2** and quinacrine, and the McGhee-Von Hippel based data processing. The binding stoichiometry of **AD2** is around 3, meaning that an **AD2** molecule binds every 3 base-pairs. The dissociation constants of quinacrine and **AD2** are about 300-400 nM and 3-10 nM, respectively, meaning that the affinity of **AD2** for ct-DNA is between 50 and 100 times higher than that of quinacrine. This could be explained by the reduction in conformational freedom and maybe additional stacking interactions provided by the phenol ring of **AD2**, not present in quinacrine. The additional positive charge provided by the *N*-methyl group of the piperazine ring likely reinforces ct-DNA – **AD2** complex formation by electrostatic interaction with the phosphates of the backbone.

## Experimental section

### UV-Vis absorption measurements

The measurements were carried out in a Varian Cary 50 Bio. ct-DNA was quantified assuming a nucleotide average extinction coefficient  $\epsilon_{\lambda=258} = 6600 \text{ M}^{-1}\text{cm}^{-1}$ . Sample concentration is given as base-pairs concentration.

### Linear dichroism

Linear dichroism spectra were recorded in an Applied Photophysics Chirascan instrument.

The instrument was turned on and purged with  $\text{N}_2$  for 10 minutes. The sample solution contained ct-DNA (100  $\mu\text{M}$ ) and acetate buffer (10 mM) at pH = 4.63. It could contain also 5 equivalents of either quinacrine or **AD2**. The sample was transferred into a 200  $\mu\text{L}$  Couette flow cell and rotation was set to 1000 r.p.m. The scans ranged from 230 to 600 nm at a rate of 30 nm/min.

Data were processed by extracting the baseline, recorded with the same sample without rotation.

### Fluorescence

Fluorescence spectra were recorded in a Varian Cary Eclipse spectrophotometer with a Cary temperature controller.

Experiments were performed in a 1 cm path length fluorescence cell. After irradiating at 400 nm, emission spectra were recorded from 415 nm to 630 nm, at 25°C. All the slits were set at 5 nm width, and the detector at medium or high voltage.

For quinacrine binding assays, a solution containing quinacrine (10 µM, 1 eq) and acetate buffer (10 mM) pH = 4.63 was prepared and transferred into a 2.5 mL cell. Additions of ct-DNA up to 10 equivalents were performed (125 µL from a 2 mM stock solution).

For **AD2** affinity assays, a solution containing ct-DNA (30 µM, 3 eq), quinacrine (10 µM, 1 eq) and acetate buffer (10 mM) pH = 4.63 was prepared and transferred into a 2.5 mL cell. Additions of **AD2** up to 2 equivalents were performed (250 µL from a 200 µM stock solution).

### Fluorescence data processing

Before processing, the data were smoothed with a Savitzky-Golay filter. Dilution was then corrected by multiplying it by a dilution factor, a fraction of the real volume divided by the initial volume. Finally, data were further corrected to avoid the inner filter effect by applying the following equation.

$$I_{corr} = \frac{I}{10^{-0.5 A}}$$

$I_{corr}$  is the corrected fluorescence intensity,  $I$  is the fluorescence intensity and  $A$  is the absorption of the sample.

Data were processed via matrix based McGhee – Von Hippel equations, either by employing fluorescence intensity values at  $\lambda = 500$  nm or the entire spectra matrix by performing a singular value decomposition.

### **References and Notes**

- (1) Nordén, B.; Rodger, A.; Dafforn, T. *Linear Dichroism and Circular Dichroism - A Textbook on Polarized-Light Spectroscopy*; The Royal Society of Chemistry: Cambridge, **2010**.
- (2) Baird, J. K. *Antimicrob. Agents Chemother.* **2011**, *55*, 1827.
- (3) McGhee, J. D.; von Hippel, P. H. *J. Mol. Biol.* **1974**, *86*, 469.
- (4) Lincoln, P. *Chem. Phys. Lett.* **1998**, *288*, 647.



## Resum en Català

### PART A

#### Introducció i objectius

Els polímers naturals com els oligonucleòtids o els pèptids tenen un gran potencial com a fàrmacs, atès que ambdós juguen papers essencials en la majoria dels processos cel·lulars, tant en tasques estructurals com de reconeixement, de regulació o metabòliques.<sup>1-5</sup>

El problema amb la utilització d'aquests compostos és la seva baixa biodisponibilitat. Com que són compostos naturals, els enzims com les nucleases i proteases els reconeixen fàcilment i els degraden ràpidament. Tanmateix, a causa de la seva mida i les seves característiques estructurals, aquests compostos no poden penetrar les membranes cel·lulars i per tant sovint no poden accedir a la diana amb què han d'interaccionar.

A fi i efecte de modificar les propietats i evitar aquests inconvenients, es poden emprar dues estratègies: la conjugació de la biomolècula o la modificació del seu esquelet, fent mimètics de la molècula o ciclant-la.<sup>6-10</sup>

Les maleïmides són grups funcionals molt reactius, que poden reaccionar amb tiols donant addicions de Michael, o amb diens, a través de cicloaddicions de Diels-Alder. El problema és que aquesta alta reactivitat limita la seva utilització, ja que emprant els mètodes de síntesi estàndard només es poden introduir als extrems d'oligonucleòtids i pèptids, i no sempre de manera directa.<sup>11,12</sup> Per resoldre aquest problema al grup es va desenvolupar un esquema de protecció per a les maleïmides, consistent en la utilització de 2,5-dimetilfurà com a protector. Aquest furà reacciona amb la maleïmida generant un cicloadducte estable. Aquest cicloadducte pot experimentar una reacció de retro-Diels-Alder, és a dir, que la maleïmida es desprotegeix a alta temperatura. Al grup de recerca es va aplicar aquesta metodologia per conjuguar i ciclar oligonucleòtids.<sup>13,14</sup>

Amb aquests antecedents l'objectiu principal de la tesi era la utilització de maleïmides (protegides i/o desprotegides) per a la síntesi de conjugats i cicles de poliamides, on s'engloben tant pèptids com peptoides. Combinant les estratègies de conjugació i ciclació, també ens vàrem plantejar la síntesi de conjugats oligonucleòtid – pèptid cíclic. L'objectiu final era la síntesi i avaluació de la possible activitat biològica de conjugats PNA – pèptid cíclic.

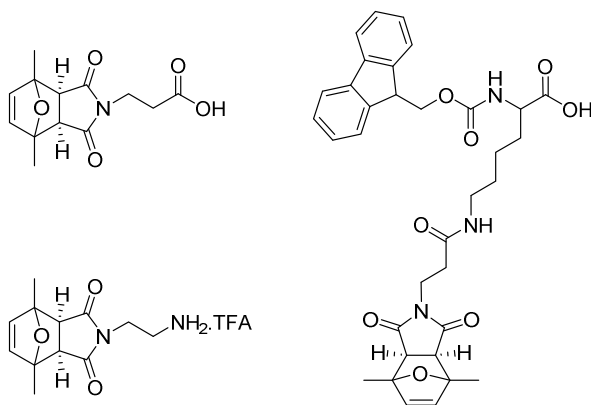
El pas previ, però, consistia en la síntesi de monòmers que continguessin maleïmides protegides, per tal de poder incorporar maleïmides a qualsevol punt de la cadena.

#### Síntesi de monòmers incorporant maleïmides protegides

Els monòmers que es van sintetitzar són l'àcid maleïmidopropanoic protegit, que serveix per incorporar maleïmides protegides a l'extrem *N*-terminal d'una poliamida; la Fmoc-Lys[ProtMal]-OH, que és un derivat de la lisina que permet la introducció de maleïmides protegides en qualsevol punt d'una cadena peptídica, i la *N*-[ProtMal]-etandiamina (que s'obté



com a sal), una maleimida protegida derivatitzada amb una amina que permet la introducció de maleïmides protegides a qualsevol punt de la cadena d'un peptòide (**Fig. RC1**).



**Figura RC1:** Estructures dels monòmers amb maleïmides protegides

Per a la síntesi de l'àcid amb la maleïmida protegida es podia partir de l'àcid maleïmidopropanoic comercial, que és car, o es podia sintetitzar en un pas a partir d'anhídrid maleïc i  $\beta$ -alanina. L'àcid es va fer reaccionar amb 2,5-dimetilfurà i es van obtenir els dos isòmers, *endo* i *exo*. Vàrem comprovar que l'isòmer *endo* era massa poc estable per ser utilitzat en posteriors etapes sintètiques, així que la mescla d'isòmers es va fer reaccionar amb amoníac aquós, el qual únicament hidrolitza l'isòmer *endo*. Mitjançant una elaboració senzilla, consistent en un canvi de pH i extraccions, es va obtenir l'isòmer desitjat (*exo*) de l'àcid amb la maleïmida protegida.

Per a la síntesi del derivat de lisina es va emprar l'àcid amb la maleïmida protegida. Aquest es va convertir en un èster actiu mitjançant la reacció amb *N*-hidroxi-succinimida. Aquest compost es va fer reaccionar amb Fmoc-Lys-OH per obtenir el monòmer desitjat.

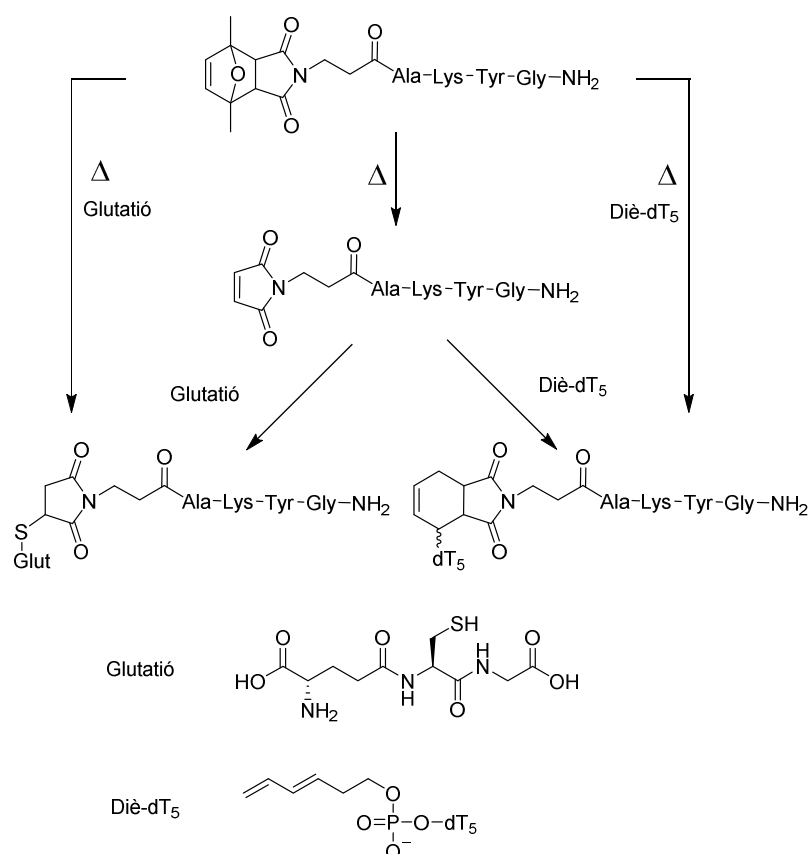
Finalment, per a la síntesi del monòmer de peptòide es va partir de maleïmida i *N*-Boc-etanolamina. La maleïmida es va protegir amb 2,5-dimetilfurà, i es va dur a terme una reacció de Mitsunobu amb *N*-Boc-etanolamina. Posteriors tractaments amb amoníac i TFA van proporcionar el monòmer desitjat (isòmer *exo*) en forma de sal de TFA.

## Conjugacions

Es varen sintetitzar dos pèptids curts que contenien una maleïmida protegida, un d'ells amb la maleïmida protegida a *N*-terminal, i l'altre incorporant el derivat de lisina enmig de la seqüència. Amb els dos pèptids es van assajar diferents condicions de desprotecció de la maleïmida, concretament l'escalfament durant tres hores en toluè anhidre a 90°C, o 90 minuts en aigua/metanol a 90°C en un microones. Les primeres semblaven ser les millors condicions per desprotegir una maleïmida unida a una cadena peptídica, tot i que la desprotecció no era quantitativa.

Després de la desprotecció, el pèptid amb la maleïmida a *N*-terminal es va conjuguar amb glutatí i diè-dT<sub>5</sub> i confirmà que la reacció retro-Diels-Alder donava una maleïmida

perfectament reactiva. També es va comprovar que, en sotmetre els dos conjugats obtinguts a les condicions de desprotecció de la maleimida, les molècules quedaven inalterades. Això demostrava l'estabilitat de les succinimides, els grups generats en les reaccions de conjugació, a les condicions de desprotecció de la maleimida. Per tant, les conjugacions també es podien fer de forma simultània a les desproteccions (*one-pot*), escalfant el pèptid a les condicions de desprotecció en presència del tiol o diè corresponent (**Fig. RC2**). Aquesta alternativa millora considerablement el rendiment global dels processos de desprotecció i conjugació quan la desprotecció de la maleimida presenta problemes.



**Figura RC2:** Síntesi de conjugats en dos passos o *one-pot* (desprotecció i conjugació simultànies)

De la mateixa manera, es va sintetitzar un peptíde amb una maleimida protegida, incorporant el monòmer que havíem preparat prèviament. El peptíde es va poder desprotegir i conjuguar amb biotina, tiocolsterol i glutatió via reacció de Michael, i amb diè-dT<sub>5</sub> via reacció de Diels-Alder (**Fig. RC3**).

Seguidament es va decidir dur a terme la síntesi de dobles conjugats. L'estratègia a seguir era la incorporació en un pèptid de dues maleimides, una protegida i l'altra no, per realitzar una primera conjugació, seguida d'una desprotecció, i finalment una segona conjugació.

En el primer cas, en què el pèptid primer es va conjuguar amb un oligonucleòtid, no hi va haver cap mena de problema durant les reaccions de desprotecció i segona conjugació. En el segon cas, però, la desprotecció del conjugat pèptid – glutatió no va funcionar tan bé, motiu pel qual es va optar per la reacció *one-pot* amb bons resultats (**Fig. RC4**).

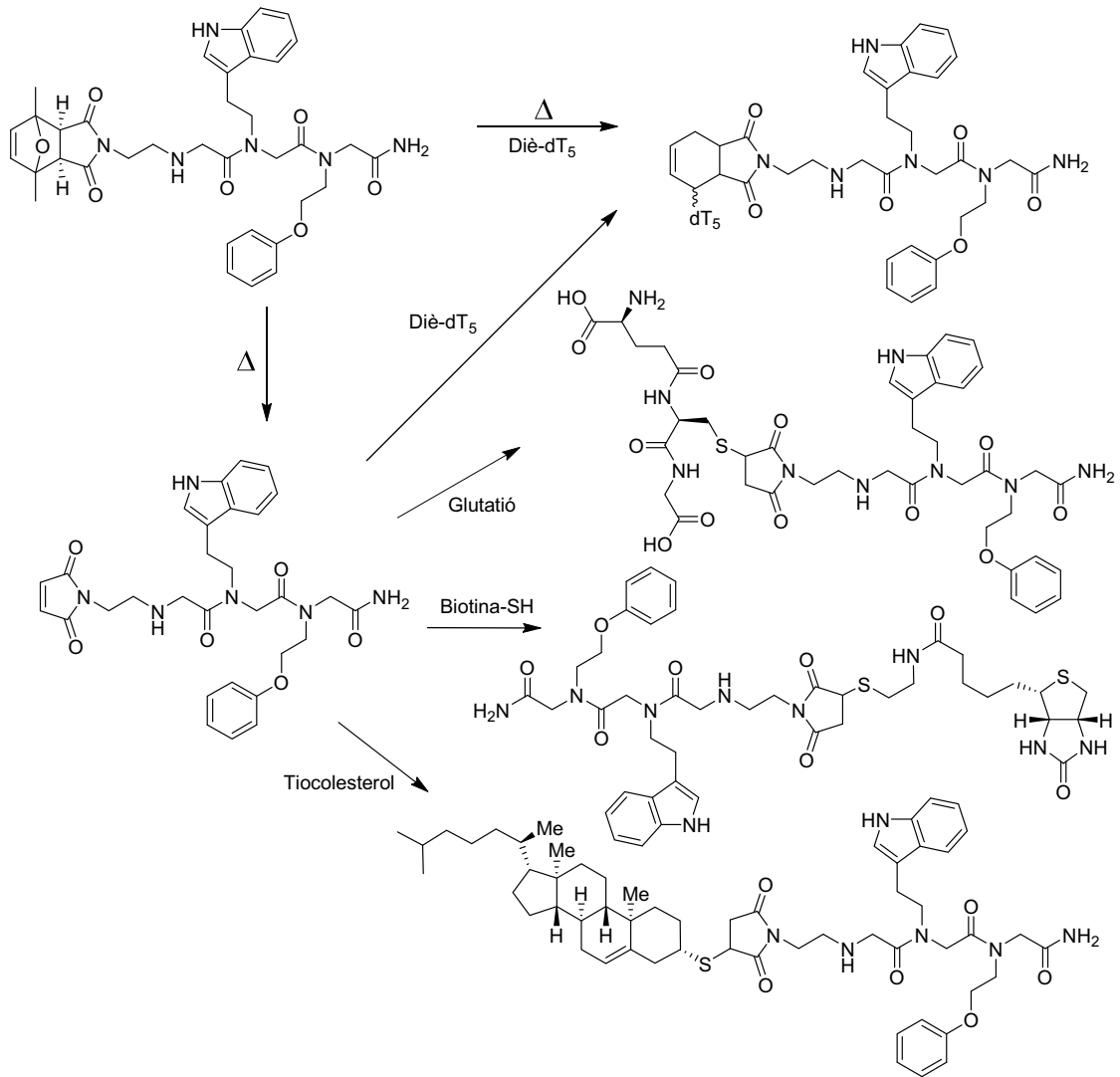


Figura RC3: Síntesi de conjugats d'un peptòide

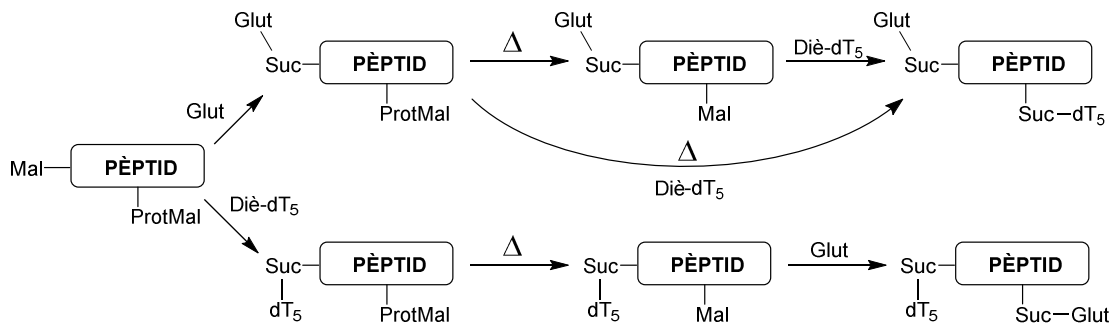


Figura RC4: Síntesi de conjugats dobles d'un peptid model

## Ciclacions

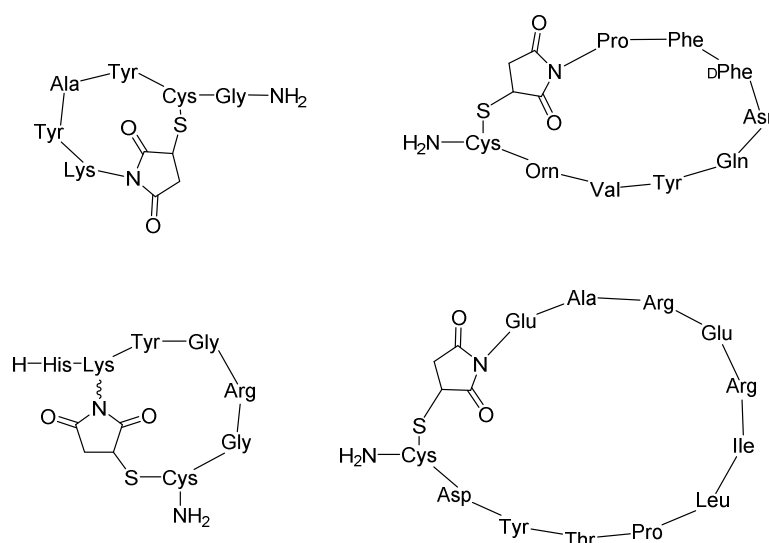
Atès que les conjugacions via addició de Michael funcionen correctament i de forma molt ràpida, vam decidir intentar la ciclació de pèptids fent ús d'aquesta reacció. La idea consistia en la incorporació de tiols (de residus de cisteïna) i de maleïmides a la cadena peptídica, per així dur a terme addicions de Michael intramoleculares.

Vàrem sintetitzar una seqüència peptídica curta, acabada a *N*-terminal amb àcid maleimidopropanoic bé protegit o bé desprotegit. Es va observar que, en fer el tractament de desancoratge i desprotecció del pèptid amb la maleimida desprotegida, el producte que s'obtenia corresponia directament al cycle desitjat. Tot i això, hi havia una proporció no menyspreable de dímer i trímer cíclics.

Per altra banda, quan es va treballar amb la maleimida protegida, després del tractament de desancoratge i desprotecció es va obtenir el corresponent pèptid lineal. La desprotecció de la maleimida i subseqüent ciclació d'aquest pèptid es va assajar en diverses concentracions, obtenint-se un cru molt net en concentracions baixes, i un cru cada cop més brut i amb més proporció de dímer i trímer cíclics en augmentar la concentració de pèptid.

Aquests resultats senyalaven que les reaccions de ciclació es poden dur a terme amb maleïmides protegides o desprotegides a *N*-terminal, però que sempre que l'esquema de protecció ho permeti és millor incorporar la maleimida protegida, ja que els crus són molt més nets.

Amb aquesta metodologia es van sintetitzar una sèrie de pèptids cíclics, alguns d'ells anàlegs de productes naturals com la tirocidina o un mímic de *loop* de l'oxidoreductasa, que incorporaven pràcticament tots els aminoàcids proteïnògens (**Fig. RC5**). Les ciclacions varen funcionar correctament en tots els casos i indicaren que el mètode desenvolupat és compatible amb virtualment tots els aminoàcids desprotegits, amb l'òbvia excepció de la cisteïna.



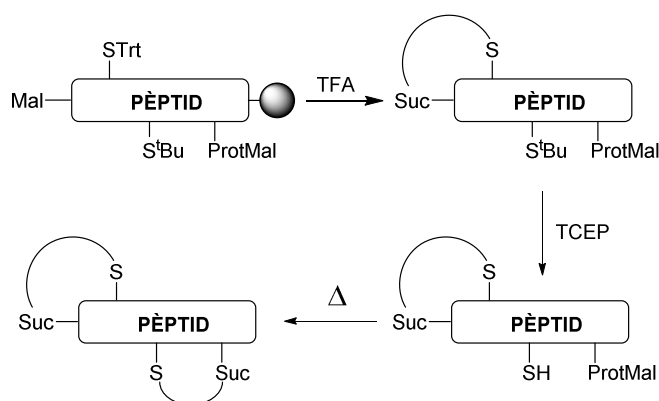
**Figura RC5:** Pèptids cíclics sintetitzats

Com que havíem vist que les ciclacions funcionaven, millor o pitjor, tant quan s'empraven maleïmides protegides com desprotegides, vàrem considerar que, si en un pèptid es podien introduir cisteïnes protegides ortogonalment entre elles, a més d'una maleïmida lliure i una protegida, es podrien obtenir bicicles.

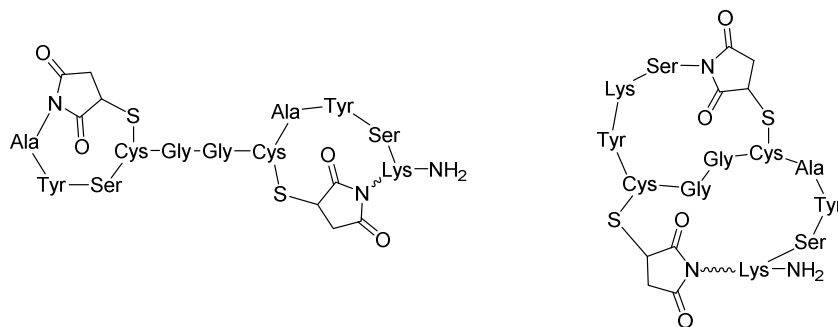
L'estratègia consistia a incorporar una maleïmida lliure a *N*-terminal i una maleïmida protegida en un altre punt de la seqüència, emprant el derivat de lisina, així com dues cisteïnes, una protegida amb tritil, que és làbil a àcids, i una protegida amb el grup *tert*-butiltio, làbil a reductors. Durant el tractament àcid de desprotecció i desancoratge del pèptid, es perdria el grup tritil d'una de les cisteïnes, i es formaria el primer cycle per reacció amb la maleïmida lliure de la posició *N*-terminal. Posteriorment es faria reaccionar el pèptid cíclic amb *tris*-(carboxietil)fosfina, per tal de desprotegir l'altra cisteïna. Finalment, el pèptid cíclic amb el tiol lliure s'escalfaria per tal de desprotegir la maleïmida del derivat de lisina, i així formar el segon cycle.

Seguint aquesta estratègia es van poder sintetitzar dos pèptids bicíclics de seqüències molt similars però amb patrons de ciclació diferents, un d'ells amb els cycles separats, i l'altre amb els cycles fusionats. Aquests diferents patrons de ciclació es van aconseguir canviant les posicions relatives de les dues cisteïnes diferentment protegides (**Fig. RC6**).

a)



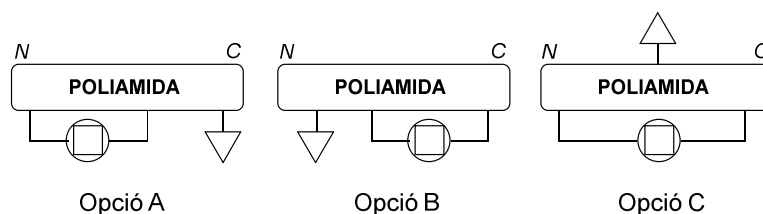
b)



**Figura RC6:** a) Esquema general de síntesi de pèptids bicíclics amb els cycles separats. b) Pèptids bicíclics sintetitzats

## Conjugats de pèptids cíclics. Prova de concepte

En aquest punt ja havíem estat capaços de fer conjugats senzills i dobles, cicles i bicicles, així que vàrem decidir intentar la síntesi de conjugats de pèptids cíclics. Vàrem imaginar tres possibles opcions, en les quals el grup funcional per conjugar (triangle, a la figura) estava fora del cicle, a C-terminal o a N-terminal, o dins del propi cicle (**Fig. RC7**).



**Figura RC7:** Opcions per la síntesi de conjugats de poliamides cíclics

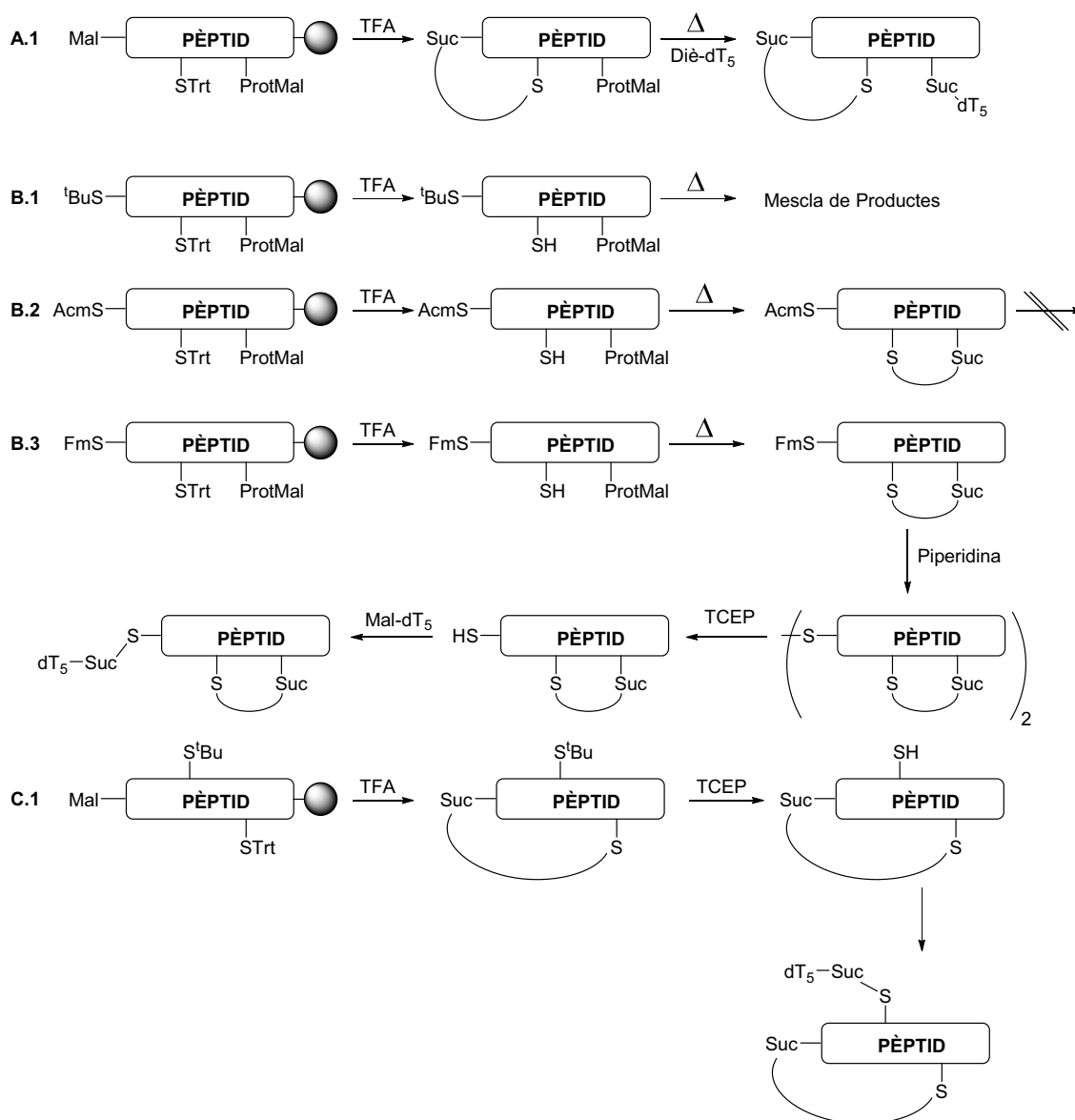
Per tal d'obtenir la primera estructura (**Fig. RC7, Opció A** (triangle = maleimida)), es va sintetitzar un pèptid amb una maleimida lliure a N-terminal, el derivat de lisina amb la maleimida protegida a C-terminal i una cisteïna protegida amb tritol al mig de la seqüència. Igual que amb els pèptids bicíclics, el cicle es va formar durant el tractament àcid de desprotecció i desancoratge del pèptid. En desprotegir la maleimida del pèptid cíclic resultant en presència d'un diè-oligonucleòtid es va obtenir en només dues etapes el conjugat de pèptid cíclic – oligonucleòtid desitjat (**Fig. RC8, A.1**).

L'obtenció de la segona estructura (**Fig. RC7, Opció B** (triangle = tiol)) va resultar ser la més complexa de totes. L'estratègia dissenyada consistia a incorporar a la seqüència el derivat de lisina amb la maleimida protegida i una cisteïna protegida amb tritol, per generar el cicle, i una cisteïna per conjugar a N-terminal, amb un protector estable a l'àcid i a la temperatura.

Els monòmers de cisteïna protegida amb  $S^tBu$ , Acm i Fm van resultar ser perfectament estables a les condicions de desprotecció de la maleimida. Malgrat això, després del tractament àcid, quan el pèptid que contenia Cys( $S^tBu$ ) es va escalfar per desprotegir la maleimida i ciclar, es va obtenir una mescla de dos productes majoritaris, un corresponent al cicle desitjat, i l'altre amb un pont disulfur entre els dos residus de cisteïna del pèptid (**Fig. RC8, B.1**). Això indicava que no es pot desprotegir la maleimida en presència d'un tiol lliure i una Cys( $S^tBu$ ), perquè el tiol lliure reacciona amb tots dos. La utilització de Cys(Acm), en canvi, va permetre l'obtenció del cicle desitjat amb bon rendiment. Tot i això, el pèptid amb el tiol lliure, necessari per a la conjugació, no es va poder detectar després de diferents assajos de desprotecció de la cisteïna, presumiblement perquè tots els reactius emprats per a la desprotecció implicaven reaccions redox que comprometien l'estabilitat del tioèter resultant de la formació del cicle (**Fig. RC8, B.2**).

Finalment, l'únic grup protector amb què es va poder obtenir el pèptid cíclic desitjat amb el tiol lliure va ser el Fm. La utilització d'aquest protector, però, té un desavantatge clar respecte als altres, i és que com que és làbil a la piperidina ha de ser, per força, l'últim aminoàcid de la seqüència (s'introdueix com a Boc-Cys(Fm)-OH). Un cop es va obtenir el pèptid cíclic amb el tiol lliure a N-terminal, aquest es va conjuguar amb un maleimido-oligonucleòtid via reacció de Michael, i s'obtingué, ara sí, el pèptid cíclic conjugat desitjat (**Fig. RC8, B.3**).

La síntesi d'un pèptid amb la tercera estructura (**Fig. RC7, Opció C** (triangle = tiol)) es va dur a terme incorporant una maleimida lliure a *N*-terminal, una cisteïna protegida amb el grup tritil a *C*-terminal i una cisteïna protegida amb S<sup>t</sup>Bu al mig de la seqüència. El tiol de la cisteïna protegida amb tritil va quedar lliure durant el tractament àcid de desprotecció i desancoratge del pèptid i va reaccionar molt ràpidament amb la maleimida, formant d'aquesta manera el cicle desitjat. Posteriorment, aquest pèptid es va fer reaccionar amb *tris*-(carboxietil)fosfina per desprotegir la segona cisteïna, i finalment es va conjuguar amb un maleimido-oligonucleòtid per formar el conjugat desitjat (**Fig. RC8, C.1**).



**Figura RC8:** Síntesi de conjugats de pèptids cíclics amb el punt de conjugació fora del cicle (**Opcions A i B**) o dins del cicle (**Opció C**)

De la mateixa manera que havíem sintetitzat peptoides per conjuguar, vàrem decidir sintetitzar un conjugat d'un peptoid cíclic. En aquest cas, però, la reacció de conjugació es va dur a terme amb una reacció "click" diferent, per demostrar la compatibilitat dels compostos ciclats via reacció de Michael entre tiol i maleimida amb un altre mètode de conjugació.

En aquest cas el peptoide es va dissenyar d'acord amb l'Opció C (**Fig. RC7** (triangle = alquí)). Incorporava una maleimida protegida a *N*-terminal, un tiol protegit amb el grup tritil prop de C-terminal (provinent de *S*-tritil-cisteamina, prèviament sintetitzada) i un alquí. Després de la síntesi en fase sòlida, la desprotecció i el desancoratge es va obtenir el peptoide lineal amb el tiol desprotegit. En escalfar aquest compost es va desprotegir la maleimida i es va formar el peptoide cíclic desitjat.

Sobre aquest peptoide es varen assajar dues reaccions de conjugació. Primer es va intentar conjuguar amb una tetrazina via una reacció de Diels-Alder amb demanda electrònica inversa. La tetrazina és deficient d'electrons, i l'alquí, actuant com a dienòfil, és ric en electrons. En aquest cas el conjugat desitjat no es va poder obtenir, ja que només s'observaven senyals de degradació de la tetrazina als crus de reacció. Creiem que el problema principal és la baixa reactivitat de l'alquí, ja que habitualment aquestes reaccions es duen a terme amb alquens i alquins tensionats.<sup>15,16</sup>

L'altra reacció de conjugació assajada va ser la cicloaddició de Huisgen catalitzada per Cu(I). En aquest cas el peptoide cíclic es va poder conjuguar amb èxit amb AZT, un nucleòsid amb una azida que s'havia utilitzat com a fàrmac contra la SIDA (**Fig. RC9**).<sup>17</sup>



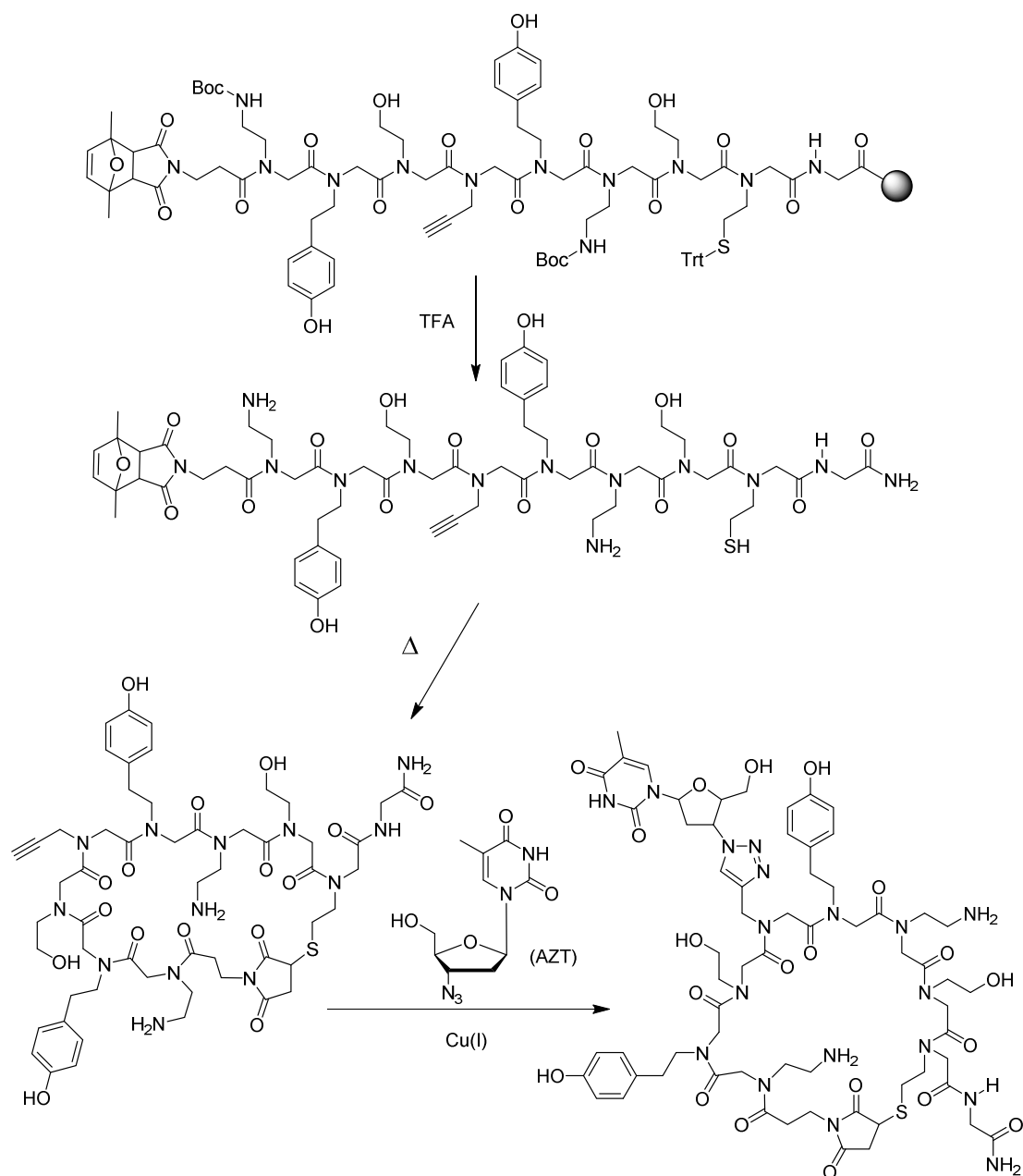


Figura RC9: Síntesi d'un conjugat d'un peptòide cíclic

### Pèptids cíclics conjugats com a potencials moduladors de l'*splicing*

L'*splicing* és un procés de maduració de l'RNA missatger en el qual els introns, parts no codificants de l'RNA, s'eliminen de la seqüència de pre-mRNA, tot ajuntant els diferents exons, les parts codificants. L'*splicing* alternatiu és un tipus d'*splicing* en el qual alguns dels exons són també eliminats de la seqüència. Aquest procés permet que a partir d'un sol pre-mRNA es puguin obtenir diferents mRNAs madurs, i per tant diferents proteïnes.<sup>18,19</sup>

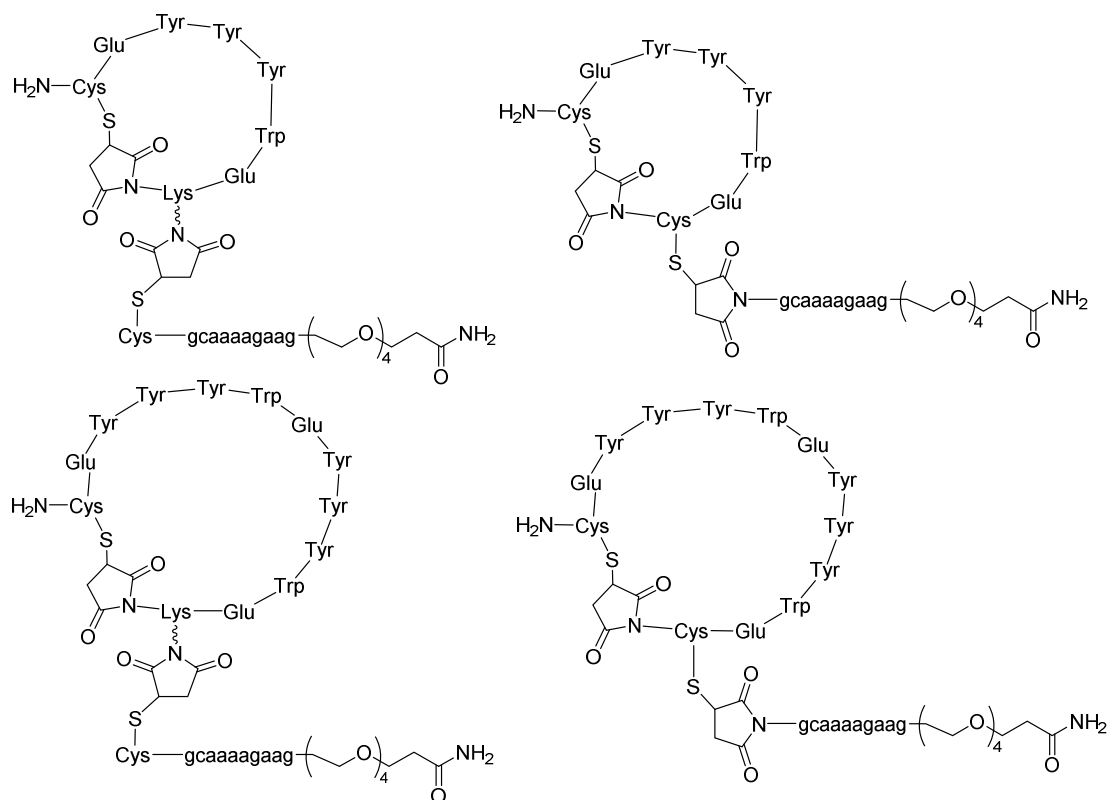
Moltes malalties genètiques o càncers provenen d'errors en aquest procés, així que la seva regulació és d'alt interès farmacològic.

Al grup de recerca s'havien sintetitzat conjugats del tipus pèptid – PNA amb activitat inductora d'*splicing* de l'RNA que codifica per la proteïna Fas, relacionada amb la senyalització per a la mort cel·lular programada.

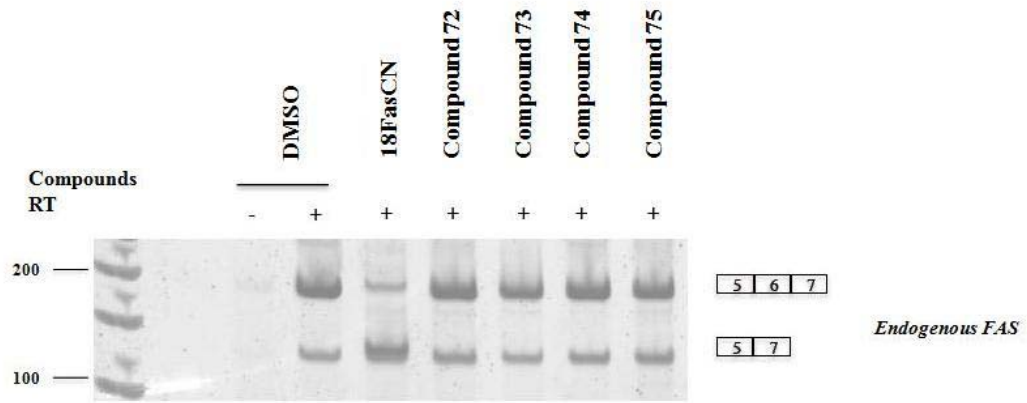
Aprofitant la metodologia desenvolupada per a la ciclació i conjugació de pèptids, es va decidir sintetitzar conjugats amb els pèptids cíclics en lloc de lineals. Com que el pèptid actiu contenia una seqüència més curta repetida dues vegades [(EYYW)<sub>2</sub>E], vàrem decidir sintetitzar quatre pèptids cíclics per conjuguar, dos d'ells amb una sola repetició i els altres dos amb la seqüència sencera. A més, un dels grans i un dels petits eren pèptids cíclics amb una maleimida per conjuguar amb un PNA-tiol, i els altres dos eren pèptids cíclics amb un tiol lliure, per conjuguar amb un PNA-maleimida.

La ciclació del pèptid en tots els casos es va dur a terme via una addició de Michael, incorporant una maleimida lliure a *N*-terminal i una cisteïna protegida amb tritol a *C*-terminal. En tots els casos el grup funcional per conjuguar estava dins del cicle (**Fig. RC7, Opció C**). Als pèptids cíclics que havien de contenir una maleimida per conjuguar s'hi va incorporar el derivat de lisina amb la maleimida protegida, seguint l'esquema de protecció de l'**Opció A.1** de l'apartat anterior (**Fig. RC8**). Per contra, quan el pèptid havia de contenir un tiol lliure per conjuguar, s'hi incorporava una cisteïna protegida amb S<sup>t</sup>Bu, com a l'**Opció C.1** de l'apartat anterior (**Fig. RC8**).

Els conjugats obtinguts (**Fig. RC9**) es van assajar al grup del Dr. Valcárcel (Centre de Regulació Genòmica, Barcelona), però malauradament es va comprovar que eren molt menys actius que el conjugat de referència pèptid lineal – PNA (**Fig. RC10**).



**Figura RC9:** Estructures dels conjugats pèptid cíclic – PNA



**Figura RC10:** Resultats d'assajos d'*splicing* alternatiu amb els conjugats de la **Fig. RC9 (Compounds 72-75)** (s'indica la numeració emprada a la PART A de la memòria). 18FasCN correspon a un conjugat pèptid lineal – PNA prèviament assajat, que actua com a referència

## PART B

### Introducció i objectius

L'RNA és una diana terapèutica interessant, ja que, a part de ser l'intermediari entre el genoma i la maquinària de síntesi de proteïnes, té un seguit d'altres funcions essencials en els cicles cel·lulars. A més, la seva capacitat de formar estructures secundàries i terciàries diferents a la doble hèlix fa que pugui adoptar estructures més semblants a les d'una proteïna que a les del DNA.<sup>20</sup>

L'IRES és una estructura particular d'RNA que serveix per iniciar la traducció de l'mRNA sense necessitat d'una estructura "cap" a l'extrem 5'.<sup>21</sup> Aquesta estructura es troba sovint en RNAs virals, com per exemple el del virus de la febre aftosa. La febre aftosa és una malaltia altament contagiosa que afecta ungulats com cavalls o vaques, i per culpa de la qual cada cop que sorgeix un brot centenars d'animals han de ser sacrificats, amb les seves corresponents pèrdues econòmiques.<sup>22</sup>

La part apical del domini 3 de l'IRES de la febre aftosa, que és essencial per a la seva funció, conté una estructura en forma de *tetra-loop* amb la seqüència GUAA.<sup>23</sup> El grup de recerca de la Prof. Baranger va descriure a l'any 2007 que un lligand anomenat **AD2 (Fig. RC11)** presentava bones afinitats per *loops* de tipus GNRA, dins dels quals s'inclou el *loop* de seqüència GUAA de l'IRES de la febre aftosa.<sup>24</sup>

L'objectiu principal de la PART B era, per tant, l'estudi de la interacció de l'**AD2** amb el *loop* GUAA de l'IRES del virus de la febre aftosa. Amb aquest fi els primers passos a seguir eren la síntesi de l'**AD2** i d'una seqüència d'RNA (**SL-RNA**) corresponent a la de l'IRES que conté el *loop* GUAA, per així fer estudis amb una col·lecció de tècniques biofísiques com la determinació de la temperatura de fusió per canvis en l'absorció de llum ultraviolada, el dicromisme circular, la ITC, l'ESI i la fluorescència. També es volia avaluar si l'**AD2** interferia en la traducció via IRES, i determinar el lloc d'unió de l'**AD2** a l'IRES a partir d'assajos *in vitro*.

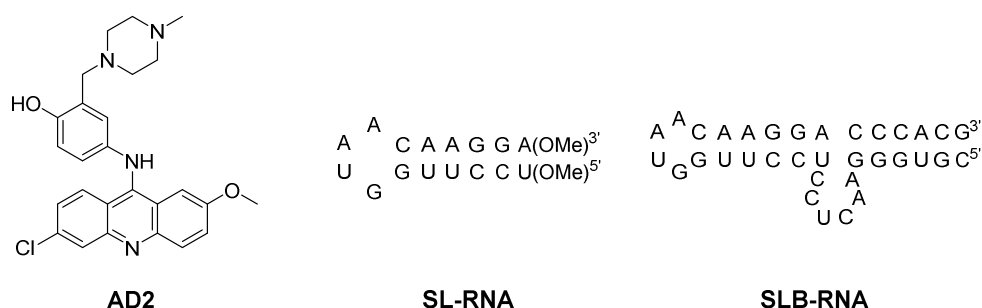
En relació amb aquest tema, durant l'estada realitzada al grup de recerca del Prof. Lincoln a Göteborg, es va decidir estudiar la interacció de l'**AD2** amb un DNA de seqüència no predefinida, també amb tècniques biofísiques com el dicromisme lineal o la fluorescència.

### Síntesi i caracterització de l'AD2, i síntesi de l'RNA

L'**AD2** es va sintetitzar seguint una ruta sintètica de quatre etapes, partint del *p*-aminofenol. La primera consistia en una protecció del grup amino per acetilació, la segona en una reacció de Mannich amb formaldehid i *N*-metil-piperazina,<sup>25</sup> la tercera en la desprotecció del grup amino i la quarta en una substitució nucleòfila aromàtica sobre un anell d'un derivat d'acridina.<sup>26</sup> El producte desitjat es va obtenir en un 34 % de rendiment global.

També se'n va determinar el pKa mitjançant el mètode descrit per Albert i Serjeant,<sup>27</sup> i es va obtenir un valor de 6.1.

La síntesi de l'**SL-RNA (Fig. RC11)** es va dur a terme pels procediments estàndard, incorporant grups metil als hidroxils 2' dels nucleòtids dels extrems, per dotar l'estructura d'una millor resistència a RNases.

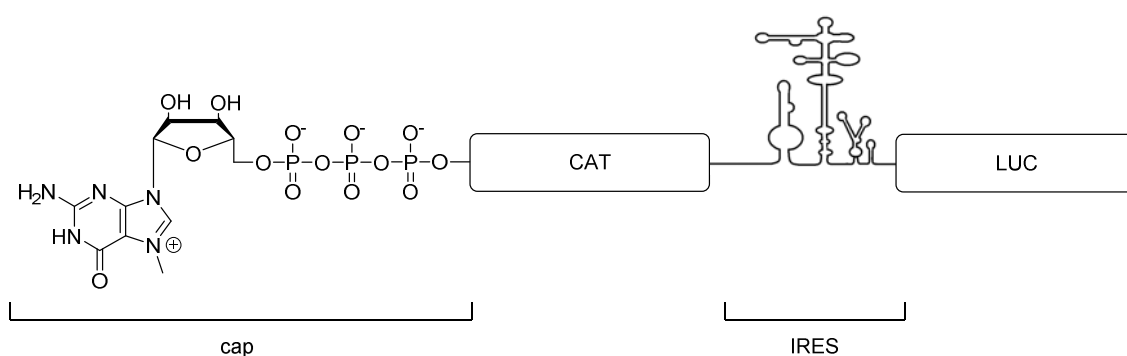


**Figura RC11:** Estructures de l'**AD2**, i dels RNAs emprats en aquest treball (**SL-RNA** i **SLB-RNA**)

### Resultats dels assajos biològics *in vitro*

Es van realitzar un seguit d'assajos *in vitro* al grup de la Prof. Martínez-Salas (Centro de Biología Molecular Severo Ochoa, Madrid).

El primer assaig consistia a determinar si l'**AD2** inhibia la traducció mediada per IRES, emprant un RNA bicistrònic (**Fig. RC12**). Aquest, a grans trets, contenia: una estructura tipus "cap" a l'extrem 5', un mRNA que codificava per la proteïna cloramfenicol acetiltransferasa (CAT), un IRES i finalment un altre mRNA que codificava per la proteïna luciferasa (LUC). En condicions normals, la maquinària cel·lular expressaria tant la proteïna CAT com la LUC. Idealment, en presència d'un lligand de l'IRES, s'observaria una inhibició diferencial de la traducció per a CAT i LUC, essent la segona més inhibida que la primera.



**Figura RC12:** RNA bicistrònic

El que es va observar va ser un descens de l'expressió de les dues proteïnes, però lleugerament més acusat per a la LUC. Aquest resultat indicava que l'**AD2** és poc selectiu, però com que afectava més a la traducció mediada per IRES es van realitzar assajos per mirar d'esbrinar amb quina zona de l'IRES interaccionava.

Per a tal fi es va realitzar un estudi per SHAPE, que és una tècnica emprada per a la determinació d'estructures secundàries i terciàries d'RNA.<sup>28</sup> En aquesta tècnica es fa reaccionar l'RNA, en presència i absència de lligand, amb un agent acilant que reacciona únicament amb els 2'-hidroxils més accessibles. Posteriorment, l'RNA acilat s'hibrida amb un

*primer* de DNA i s'hi afegeix una transcriptasa inversa per generar cadenes d'RNA fins als punts de modificació. La informació estructural s'obté finalment mitjançant l'anàlisi de la llargària i l'abundància de les seqüències de DNA transcrites.

El resultat d'aquest experiment emprant l'IRES de l'FMDV i **AD2** assenyalava que, si bé l'**AD2** semblava unir-se al *loop* GUAA esperat, semblava que s'unia també a d'altres punts de forma inespecífica.

### Assajos d'espectroscòpia d'absorció

Paral·lelament als assajos *in vitro*, es van començar a realitzar un seguit de proves biofísiques per tal d'estudiar més profundament la interacció de l'**AD2** amb l'RNA.

El primer estudi que es va dur a terme va ser la determinació de la temperatura de fusió de l'**SL-RNA** en presència de quantitats creixents d'**AD2**. Una variació a la temperatura de fusió de l'oligonucleòtid denota una estabilització o desestabilització de la seqüència, indicant que el lligand s'uneix a l'estructura i en modifica les propietats. En el nostre cas no es va observar una variació significativa de la temperatura de fusió de l'**SL-RNA**, ja que la dispersió dels valors obtinguts era molt més gran que les diferències entre els seus valors promig corresponents. A part, s'observava una pèrdua gradual de la forma sigmoïdal de les corbes, fent impossible l'ajust matemàtic a partir de 2 equivalents d'**AD2**. Creiem que aquesta dispersió pot ser fruit de la formació de diferents complexos **AD2 – SL-RNA**.

També es van dur a terme estudis per dicroisme circular. Aquesta tècnica, basada en l'absorció de llum polaritzada, permet detectar canvis en l'estructura tridimensional de biomolècules com oligonucleòtids o proteïnes. En el nostre cas tampoc vàrem observar una variació significativa dels espectres de l'**SL-RNA** en afegir concentracions creixents d'**AD2**. Tot i que per la naturalesa intercalant de l'**AD2** esperaríem observar una forta variació als espectres de dicroisme circular,<sup>29</sup> pensem que les interaccions que pot tenir l'**AD2** amb el *loop* i l'*stem* de la seqüència poden tenir efectes oposats a l'espectre de dicroisme, i això fa que no s'observin canvis significatius després de l'addició de fins a 2 equivalents d'**AD2**.

### Experiments de calorimetria (ITC)

Veient que les tècniques d'absorció estàndard no donaven resultats útils per a la caracterització de la unió **AD2 – SL-RNA** vàrem decidir assajar la calorimetria de valoració isotèrmica. Aquesta tècnica detecta la diferència de potència elèctrica que s'ha d'aplicar a una cel·la amb la diana després de cada addició de lligand per tal de mantenir-la a la mateixa temperatura que la cel·la de referència. Els valors obtinguts es poden transformar fàcilment a diferències d'entalpia. Amb la col·lecció de diferències d'entalpia també es pot obtenir una constant de dissociació, així com la resta de paràmetres termodinàmics.

En el nostre cas es varen realitzar experiments amb dos RNAs diana, l'**SL-RNA** i l'**SLB-RNA** (Fig. **RC11**). El darrer conté la mateixa seqüència que l'**SL-RNA** però és més llarg, i incorpora també

un *bulge*. A 25°C, però, les variacions de calor eren tan petites que a la pràctica només es podia intuir una tendència decreixent. En repetir l'experiment a 40°C els dos termogrames obtinguts milloraven força els anteriors i assenyalaven que existia una interacció, sobretot amb l'**SLB-RNA**. Ara, les dades indicaven que la interacció estava dirigida entròpicament, i que no era suficientment forta com per extreure'n valors termodinàmics fiables.

### Assajos de fluorescència

Per tal de determinar valors d'afinitat es van realitzar assajos de fluorescència amb RNAs marcats amb 2-aminopurina. La 2-aminopurina és una nucleobase fluorescent que pot substituir un residu d'adenina sense pertorbar en excés l'estructura de l'RNA. La fluorescència de l'aminopurina decreix quan estableix interaccions d'apilament amb altres nucleobases, o, en el nostre cas, amb la part aromàtica de l'**AD2**.

Es varen realitzar valoracions d'**AD2** amb tres seqüències d'RNA marcat: **Ap180-SL**, **Ap180-SLB** i **Ap167-SLB** (Fig. RC13). Els valors d'EC<sub>50</sub> obtinguts es troben a la taula inferior (Taula RC1).

Els valors obtinguts per l'**Ap180-SL** són raonablement bons, i indiquen que la interacció té lloc i es produeix amb una EC<sub>50</sub> de l'ordre micromolar baix. El fet que la EC<sub>50</sub> de la interacció amb l'**Ap167-SLB** sigui més baixa que les obtingudes pels dos RNAs marcats a l'**Ap180**, però, indica que el lligand és més afí pel *bulge* que pel *loop* de tipus GNRA de la seqüència. Això posa de manifest la promiscuïtat de l'**AD2** com a lligand, i concorda amb el que la mateixa Prof. Baranger descrivia (sense dades addicionals) un cop aquest treball ja estava en marxa.<sup>30</sup>

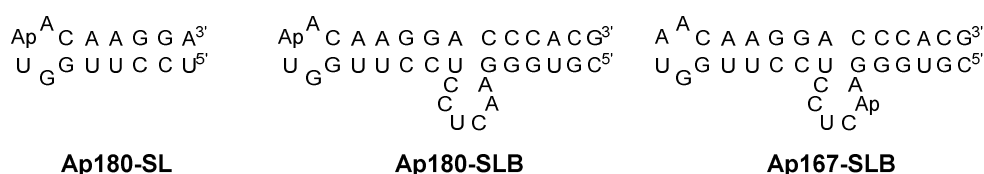


Figura RC13: Estructures dels RNAs emprats als assajos de fluorescència

RNA	Valors obtinguts d'EC <sub>50</sub> (μM)			Promig	Desviació estàndard
<b>Ap180-SL</b>	3.63	3.54	4.90	4.03	0.76
<b>Ap180-SLB</b>	3.24	2.88	4.07	3.40	0.61
<b>Ap167-SLB</b>	1.86	1.41	1.35	1.54	0.28

Taula RC1: Valors d'EC<sub>50</sub> obtinguts dels assajos de fluorescència

## Determinació de l'estequiometria per espectrometria de masses

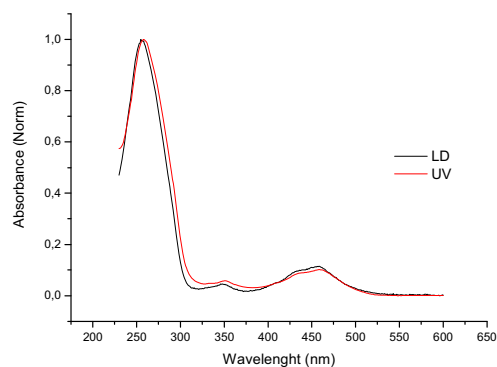
Un cop determinades les  $EC_{50}$  dels diferents RNAs marcats vàrem decidir determinar l'estequiometria dels complexos **AD2** – RNA. Existeixen diversos mètodes per determinar l'estequiometria, com per exemple el Job Plot,<sup>31</sup> però aquest requereix molta quantitat d'RNA fluorescent, i, com que als experiments de fluorescència només es detecten les interaccions dels voltants de la marca fluorescent, no és evident que s'obtingui l'estequiometria global. L'alternativa considerada va ser l'espectrometria de masses. Treballant amb el mínim voltatge es poden detectar les masses de la diana sola i les dels complexos diana – lligand.<sup>32–35</sup>

En el nostre cas vàrem observar una estequiometria mínima de 3:1 per al complex **AD2** – **SL-RNA**. La Prof. Baranger i col·laboradors havien determinat una estequiometria de 2:1 amb estructures molt similars per fluorescència,<sup>24</sup> motiu pel qual pensem que dos molècules d'**AD2** s'incorporen al *loop*, i almenys una altra molècula d'**AD2** s'intercala de forma inespecífica al llarg de la doble cadena.

En el complex **AD2** – **SLB-RNA** vàrem observar que almenys 5 molècules d'**AD2** s'unien a l'RNA diana, possiblement dues al *loop*, i la resta repartides entre el *bulge* i intercalades de forma inespecífica a l'*stem*.

## Estudi de la unió de l'AD2 al DNA

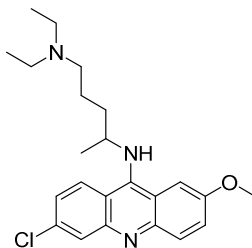
Durant l'estada al laboratori del Prof. Lincoln (Chalmers University of Technology, Göteborg) ens vàrem plantejar l'estudi de la unió de l'**AD2** al DNA. Primer es varen realitzar assajos de dicroisme lineal. Aquesta tècnica consisteix a mesurar l'absorbància d'una mostra irradiada amb llum polaritzada orientada paral·lela i perpendicularment respecte l'eix de la mostra. Els dos senyals es resten per obtenir l'espectre de dicroisme lineal definitiu. Per tal d'obtenir senyal cal orientar la mostra. En el nostre cas, la mostra s'orientava emprant cel·les tipus *Couette*, proveïdes d'un rotor que genera un flux que orienta les cadenes de DNA. Les mesures realitzades amb ct-DNA i **AD2** mostraven un espectre idèntic al d'absorció de llum ultraviolada però amb signe negatiu, la qual cosa significa que la interacció de l'**AD2** amb el DNA és per intercalació, ja que l'orientació de les nucleobases i de l'acridina eren idèntiques (**Fig. RC14**).



**Figura RC14:** Superposició dels espectres d'UV i dicroisme lineal d'**AD2**. Els dos espectres estan normalitzats, i el de dicroisme lineal està multiplicat per -1

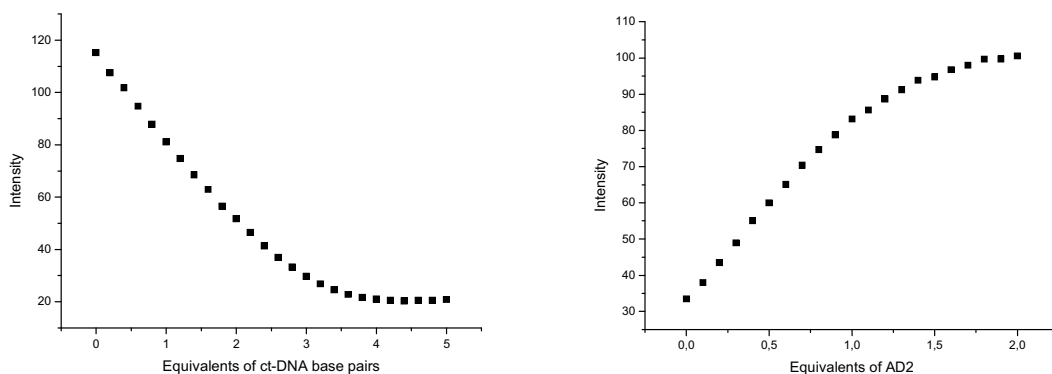


Un cop determinat el mode d'unió es va procedir a determinar la constant de dissociació de l'**AD2**. La tècnica escollida va ser la fluorescència, però com que l'**AD2** per si sol no és fluorescent es va decidir realitzar assajos de desplaçament de la quinacrina, un compost àmpliament estudiat, que es feia servir com a antimalàric durant la Segona Guerra Mundial,<sup>36</sup> i que és fluorescent quan no es troba intercalat al DNA (**Fig. RC15**).



**Figura RC15:** Estructura de la quinacrina

Primer es va determinar l'afinitat de la quinacrina pel ct-DNA, mesurant com decreixia la intensitat de la fluorescència de la quinacrina amb quantitats creixents de ct-DNA. Després es va preparar una mostra de ct-DNA saturat amb quinacrina, i es va mesurar l'increment de fluorescència a mida que s'afegia **AD2** i aquest desplaçava la quinacrina (**Fig. RC16**).



**Figura RC16:** Intensitat de la fluorescència a  $\lambda = 500$  nm de quinacrina amb quantitats creixents de ct-DNA (esquerra), i intensitat de la fluorescència a  $\lambda = 500$  nm d'una mescla de quinacrina – ct-DNA amb quantitats creixents d'**AD2**

Les dades generades als experiments de fluorescència es van processar com a intensitats discretes a una longitud d'ona determinada, i també com a espectres sencers, via *singular value decomposition*, emprant el software Matlab® i utilitzant l'aproximació matricial de McGhee – Von Hippel.<sup>37</sup> Ambdós processats apuntaven a una constant de dissociació de la quinacrina amb valors entre 300 i 400 nM i estequiometria al voltant de 2.5, i per a l'**AD2** unes constants entre 3 i 10 nM, amb estequiometries entre 2.5 i 3.

En conjunt, aquest treball va confirmar la capacitat de l'**AD2** per interaccionar amb dobles hèlixes d'àcid nucleic via intercalació, i, per primer cop, va posar de manifest que l'**AD2** pot interaccionar tant amb RNA com amb DNA.

## Conclusions Generals

1. El cicloadducte *exo* maleimida – 2,5-dimetilfurà és estable a les condicions de síntesi de poliamides en fase sòlida, però no *l'endo*. És per això que l'isòmer *exo* es va emprar per sintetitzar tres monòmers per a la incorporació de maleimides a diferents punts de cadenes peptídiques i peptoídiques.
2. Es van sintetitzar pèptids i peptoides amb maleimides protegides, els quals es van poder desprotegir i conjuguar via addició de Michael o cicloaddició de Diels-Alder amb glutatió o diè-dT<sub>5</sub>. En cas de desproteccions difícils es va comprovar que es podia desprotegir i conjuguar simultàniament, i així es millorava el rendiment de desprotecció i, per tant, també el de conjugació. Tanmateix, introduint una maleimida lliure i una protegida al mateix pèptid es van poder obtenir conjugats dobles, és a dir, pèptids conjugats amb glutatió i dT<sub>5</sub>.
3. Incorporant una maleimida (protegida o desprotegida) i un tiol en una mateixa seqüència es van poder obtenir una col·lecció de pèptids cíclics de diferents mides. Es va comprovar també la compatibilitat d'aquest mètode amb pràcticament tots els aminoàcids proteïnògens (excepte la Cys). Incorporant dues maleimides, una protegida i l'altra lliure, i dues cisteïnes protegides ortogonalment es van poder obtenir pèptids bicíclics, tant amb els anells separats com fusionats.
4. Pèptids i peptoides amb maleimides, cisteïnes i un grup funcional per conjuguar (tiol, maleimida o alquí, convenientment protegits quan era necessari), es varen ciclar via la reacció entre maleimida i tiol. Posteriorment aquests cicles es van conjuguar amb oligonucleòtids o nucleòsids via reaccions "click", per obtenir conjugats de poliamides cícliques.
5. Es van sintetitzar quatre conjugats del tipus pèptid cíclic – PNA, i es va assajar la seva activitat com a moduladors de l'*splicing*. Desafortunadament, cap d'ells va millorar l'activitat dels conjugats amb els pèptids lineals preparats prèviament al grup.
6. Es va sintetitzar el compost **AD2** amb un rendiment global del 34 %, i se'n va determinar el pKa, amb un valor de 6.1. També es va sintetitzar l'**SL-RNA**, amb el qual es van dur a terme els primers assajos biofísics d'absorció.
7. Els estudis *in vitro* realitzats al grup de la Prof. Martínez-Salas assenyalaven que l'**AD2** presentava certa activitat com a inhibidor de la traducció mediada per IRES, però era massa promiscu i s'unia de forma inespecífica a d'altres estructures d'RNA, diferents als *loops* GNRA.
8. Ni les determinacions de temperatura de fusió ni els espectres de diroisme circular van aportar cap informació útil per a la caracterització de la unió de l'**AD2** amb l'RNA. La informació obtinguda dels experiments de calorimetria tampoc va resultar útil, ja que, si bé es va detectar interacció, aquesta semblava dirigida entròpicament i els termogrames obtinguts no permetien l'obtenció de cap valor termodinàmic fiable.
9. Amb els experiments de fluorescència amb RNA marcat amb 2-aminopurina es van poder obtenir valors d'EC<sub>50</sub>, amb els quals es va confirmar el que s'havia observat als assajos *in vitro*, que l'**AD2** és un bon lligand de RNA però molt poc selectiu. Aquesta conclusió es va veure recolzada pels resultats dels experiments d'espectrometria de masses, que indicaven una

estequiometria mínima de 3 unitats d'AD2 als complexos amb l'SL-RNA i de 5 en el cas de l'SLB-RNA. Els resultats d'ambdós experiments indicaven, per tant, que l'AD2 no tan sols s'uneix a *loops*, sinó que també s'uneix al *bulge* i s'intercala a la zona de l'*stem*.

**10.** D'acord amb els assajos de dicroisme lineal i fluorescència realitzats al grup del Prof. Lincoln, l'AD2 s'uneix al DNA per intercalació amb una constant de dissociació de l'ordre nanomolar. Aquests experiments confirmen la seva capacitat per a actuar com a intercalant, la qual cosa recolza que pugui intercalar a zones de doble hèlix d'RNA malgrat que la seva estructura difereix de la de la doble hèlix de DNA.

## Referències i Notes

- (1) Gran, L.; Sandberg, F.; Sletten, K. J. *Ethnopharmacol.* **2000**, *70*, 197.
- (2) Yamada, A.; Sasada, T.; Noguchi, M.; Itoh, K. *Cancer Sci.* **2013**, *104*, 15.
- (3) Martínez, T.; Wright, N.; López-Fraga, M.; Jiménez, A. I.; Pañeda, C. *Hum. Genet.* **2013**, *132*, 481.
- (4) Jackson, A. L.; Linsley, P. S. *Nat. Rev. Drug Discov.* **2010**, *9*, 57.
- (5) Kole, R.; Krainer, A. R.; Altman, S. *Nat. Rev. Drug Discov.* **2012**, *11*, 125.
- (6) Manoharan, M. *Antisense Nucleic Acid Drug Dev.* **2002**, *12*, 103.
- (7) Krishnamurthy, V. M.; Estroff, L. A.; Whitesides, G. M. In *Fragment-based approaches in drug discovery*; Wiley-VCH, **2006**; pp. 11–53.
- (8) Huang, M. L.; Shin, S. B. Y.; Benson, M. A.; Torres, V. J.; Kirshenbaum, K. *ChemMedChem* **2012**, *7*, 114.
- (9) Bock, J. E.; Gavenonis, J.; Kritzer, J. A. *ACS Chem. Biol.* **2013**, *8*, 488.
- (10) Migianu-Griffoni, E.; Chebbi, I.; Kachbi, S.; Monteil, M.; Sainte-Catherine, O.; Chaubet, F.; Oudar, O.; Lecouvey, M. *Bioconjugate Chem.* **2014**, *25*, 224.
- (11) Marchán, V.; Grandas, A. In *Current Protocols in Nucleic Acid Chemistry*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, **2001**; Vol. 31, pp. 4.32.1–4.32.31.
- (12) Marchán, V.; Ortega, S.; Pulido, D.; Pedroso, E.; Grandas, A. *Nucleic Acids Res.* **2006**, *34*, e24.
- (13) Sánchez, A.; Pedroso, E.; Grandas, A. *Org. Lett.* **2011**, *13*, 4364.
- (14) Sánchez, A.; Pedroso, E.; Grandas, A. *Chem. Commun.* **2013**, *49*, 309.
- (15) Sauer, J.; Heldmann, D. K.; Hetzenegger, J.; Krauthan, J.; Sichert, H.; Schuster, J. *Eur. J. Org. Chem.* **1998**, 2885.
- (16) Chen, W.; Wang, D.; Dai, C.; Hamelberg, D.; Wang, B. *Chem. Commun.* **2012**, *48*, 1736.
- (17) U.S. Food and Drug Administration [www.fda.gov](http://www.fda.gov).
- (18) Black, D. L. *Annu. Rev. Biochem.* **2003**, *72*, 291.
- (19) Kornblihtt, A. R.; Schor, I. E.; Alló, M.; Dujardin, G.; Petrillo, E.; Muñoz, M. J. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 153.
- (20) Eddy, S. R. *Nat. Rev. Genet.* **2001**, *2*, 919.
- (21) Martínez-Salas, E.; Piñeiro, D.; Fernández, N. *Comp. Funct. Genomics* **2012**, *2012*, 1.
- (22) Grubman, M. J.; Baxt, B. *Clin. Microbiol. Rev.* **2004**, *17*, 465.
- (23) Fiore, J. L.; Nesbitt, D. J. *Q. Rev. Biophys.* **2013**, *46*, 223.
- (24) Yan, Z.; Sikri, S.; Beveridge, D. L.; Baranger, A. M. *J. Med. Chem.* **2007**, *50*, 4096.

- (25) Mahesh, R.; Venkatesha Perumal, R. *Indian J. Chem.* **2004**, *43B*, 1012.
- (26) Ferlin, M. G.; Marzano, C.; Chiarello, G.; Baccichetti, F.; Bordin, F. *Eur. J. Med. Chem.* **2000**, *35*, 827.
- (27) Albert, A.; Serjeant, E. P. In *Ionization constants of acids and bases: a laboratory manual*; John Wiley & Sons, Inc., **1962**; Vol. 4, pp. 69–92.
- (28) Steen, K.-A.; Malhotra, A.; Weeks, K. M. *J. Am. Chem. Soc.* **2010**, *132*, 9940.
- (29) Hoener, B.-A.; Sokoloski, T. D.; Mitscher, L. A. *Antimicrob. Agents Chemother.* **1973**, *4*, 455.
- (30) Warui, D. M.; Baranger, A. M. *J. Med. Chem.* **2009**, *52*, 5462.
- (31) Huang, C. Y. *Methods Enzymol.* **1982**, *87*, 509.
- (32) Sannes-Lowery, K. A.; Griffey, R. H.; Hofstadler, S. A. *Anal. Biochem.* **2000**, *280*, 264.
- (33) Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E. *Nucleic Acids Res.* **2002**, *30*, e82.
- (34) Brodbelt, J. S. *Annu. Rev. Anal. Chem.* **2010**, *3*, 67.
- (35) Hilton, G. R.; Benesch, J. L. P. *J. R. Soc. Interface* **2012**, *9*, 801.
- (36) Baird, J. K. *Antimicrob. Agents Chemother.* **2011**, *55*, 1827.
- (37) McGhee, J. D.; von Hippel, P. H. *J. Mol. Biol.* **1974**, *86*, 469.