UNIVERSIDADE DA CORUÑA
Facultade de Ciencias

Grado en Química

Memoria del Trabajo de Fin de Grado

Viologen-peptide conjugates in supramolecular chemistry

Conjugados viológeno-péptido en química supramolecular

Conxugados violóxeno-péptido en química supramolecular

Vanesa López Sobrado

Curso: 2019 - 2020. Convocatoria: Septiembre

Director 1: Carlos Peinador Veira
Director 2: Elena Pazos Chantrero

Index

Abbreviations

Å	Ángstrom
Ac_2O	Acetic anhydride
Ac	Acetyl
MeCN	Acetonitrile
AcOEt	Ethyl acetate
Alloc	Allyloxycarbonyl
Boc	*Tert*-butoxycarbonyl
*t*Bu	*Tert*-butyl
Trt	Trityl
Bn	Benzyl
^{13}C-NMR	Carbon Nuclear Magnetic Resonance
δ	Chemical shift
CB[*n*]	Cucurbit-[*n*]-uril
DCM	Dichloromethane
DCC	*N,N'*-dicyclohexylcarbodiimide
DEDTC	*N,N*-diethyldithiocarbamate
DIEA	*N,N'*-diisopropylethylamine
DMF	Dimethylformamide
Et_2O	Diethyl ether
ESI	Electro Spray Ionization
EtOH	Ethanol
Fmoc	9-fluorenylmethoxycarbonyl
HPLC	High Performance Liquid Chromatography

HOBt	1-hydroxybenzotrizole
MeI	Iodomethane
MS	Mass Spectrometry
MeOH	Methanol
NMM	*N*-methylmorpholine
$Pd(OAc)_2$	Palladium acetate
Pbf	2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-sulfonyl
$PhSiH_3$	Phenylsilane
^{1}H-NMR	Proton Nuclear Magnetic Resonance
t_R	Retention time
rt	Room temperature
SPPS	Solid Phase Peptide Synthesis
HBTU	2-[(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate]
PPh_3	Triphenylphosphine
TLC	Thin Layer Chromatography
TFA	Trifluoroacetic acid
TIS	Triisopropylsilane
TNBS	2,4,6-trinitrobenzenesulfonic acid
UV	Ultra Violet

Abstract

The aim of this project is to couple a viologen derivative to a short peptide sequence, from the basic region of the GCN4 transcription factor, using a solid phase Zincke reaction. Due to the chemical properties of viologens, the desired conjugate could form supramolecular complexes with cucurbit[8]uril that allow to control their dimerization and biological function in response to external stimuli.

With this purpose in mind, the required activated viologen and peptide derivative were successfully synthesized. The coupling of the activated viologen to the orthogonally deprotected Lys side chain located at the *C*-terminus of the peptide sequence was carried out using a solid phase Zincke reaction.

In order to follow all the reactions and characterize intermediates and products, NMR and HPLC-MS were the techniques used for this purpose. The final viologen-peptide conjugate was purified by semipreparative reversed-phase HPLC.

Keywords: solid phase peptide synthesis, supramolecular chemistry, viologen, Zincke reaction.

Resumen

El objetivo de este proyecto es el acoplamiento de un derivado de un viológeno a una secuencia corta de un péptido, derivado de la región básica del factor de transcripción GCN4, usando una reacción de Zincke en fase sólida. Gracias a las propiedades químicas de los viológenos, el conjugado deseado puede formar complejos supramoleculares con el cucurbit[8]uril, lo que permitiría controlar su dimerización y función biológica en respuesta a estímulos externos.

Teniendo esto en mente, el viológeno activado y el derivado del péptido fueron sintetizados satisfactoriamente. El acoplamiento entre el viológeno activado y la cadena lateral desprotegida ortogonalmente de la lisina, situada en el extremo *C*-terminal de la secuencia del péptido, se llevó a cabo utilizando una reacción de Zincke en fase sólida.

Para seguir todas las reacciones y caracterizar los intermedios y los productos, RMN y HPLC-MS fueron las técnicas utilizadas para este fin. El conjugado final viológeno-péptido fue purificado mediante HPLC semipreparativo en fase reversa.

Palabras clave: síntesis de péptidos en fase sólida, química supramolecular, viológeno, reacción de Zincke.

Resumo

O obxectivo deste proxecto é o acoplamento dun derivado dun violóxeno a unha secuencia corta dun péptido, derivado da rexión básica do factor de transcripción GCN4, usando unha reacción de Zincke en fase sólida. Grazas ás propiedades químicas dos violóxenos, o conxugado desexado pode formar complexos supramoleculares co cucurbit[8]uril, o que permitiría controlar a súa dimerización e función biolóxica en resposta a estímulos externos.

Tendo isto en mente, o violóxeno activado e o derivado do péptido foron sintetizados satisfactoriamente. O acoplamento entre o violóxeno activado e a cadea lateral desprotexida ortogonalmente da lisina, situada no extremo *C*-terminal da secuencia do péptido, levouse a cabo utilizando unha reacción de Zincke en fase sólida.

Para seguir todas as reaccións e caracterizar os intermedios e os productos, RMN e HPLC-MS foron as técnicas utilizadas para este fin. O conxugado final violóxeno-péptido foi purificado mediante HPLC semipreparativo en fase reversa.

Palabras clave: síntese de péptidos en fase sólida, química supramolecular, violóxeno, reacción de Zincke.

Introduction

Proteins and peptides: Structure and bonding

Peptides and proteins are polymers formed by monomeric subunits that are called amino acids. There are 20 essential amino acids and all of them are characterized by a common structure: a carboxylic acid with an amino group attached to the C_α, which is the reason why they are called α-amino acids. For each one of the α-amino acids, the α-carbon is bonded to different side chains, which is key to differentiate one amino acid from other. The standard amino acid structure is shown in **Figure 1**.

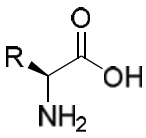


Figure 1. Structure of a standard amino acid

There are two types of amino acids, D or L, the difference between them is the position of the amine group respect to the C_α (see **Figure 2**); they are stereoisomers. In proteins, all amino acids are L.

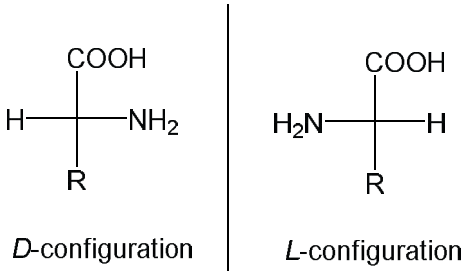


Figure 2. D and L configurations of amino acids

Amino acids can be classified depending on the polarity and acid/basic character of their side chain. In this way, we find acid amino acids (D, E), basic amino acids (K, R, H), polar amino acids (S, T, N, Q, Y, C) and non-polar amino acids (G, A, V, L, I, M, F, P, W), as shown in **Table 1**.

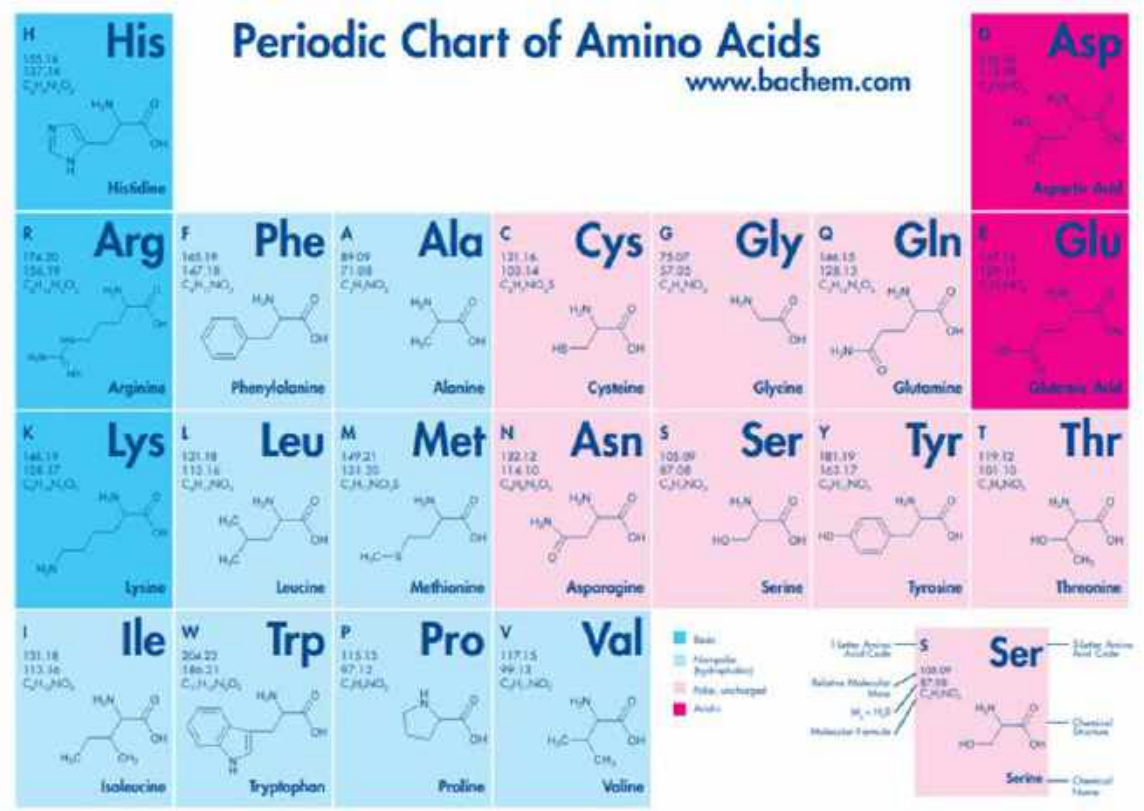
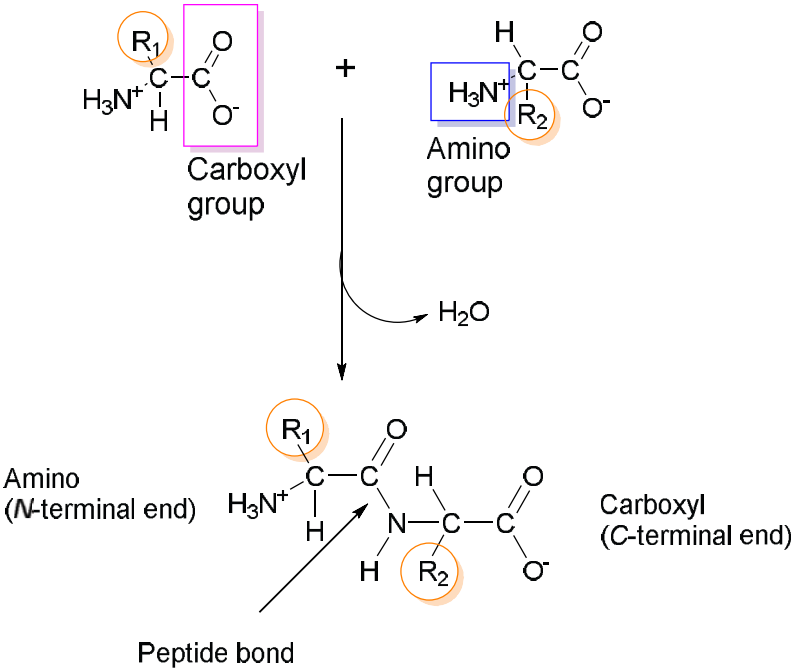


Table 1. Periodic chart of amino acids[1]

Amino acids have at least 2 pK_a values as all of them have two ionizable functional groups, the carboxylic acid and the amine. Some amino acids can have more than 2 ionizable groups because of their side chains. The isoelectric point (pI) is the pH value in which the overall net charge is equal to zero.

Amino acids are bonded between them by a peptide bond in order to form the primary structure of proteins and peptides. The bond is formed by the condensation of the carboxyl group of one amino acid and the amine of other, forming the corresponding amide bond and liberating one water molecule (**Scheme 1**). The amino acid coupled by its carboxyl group would stay at the *N*-terminus and the amino acid coupled by its amine would rest at the *C*-terminus in the formed peptide.

[1] Obtained from www.bachem.com (accessed Jul 22, 2020)



Scheme 1. Peptide bond formation

The peptide bond has a certain double bond character, due to the resonance of the electronic pair of the N atom with the C=O double bond (as can be seen in **Figure 3**).

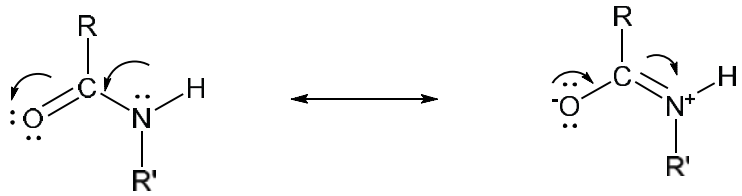


Figure 3. Resonance form of the peptide bond

Due to this resonance on the structure, the peptide bond is rigid and there is no free rotation (ω angle is fixed). On the other hand, the σ bonds formed between C_α-N and C_α-$C_{carbonyl}$, have free rotation and ϕ and ψ angles can adopt values from -180º to +180º (**Figure 4**).

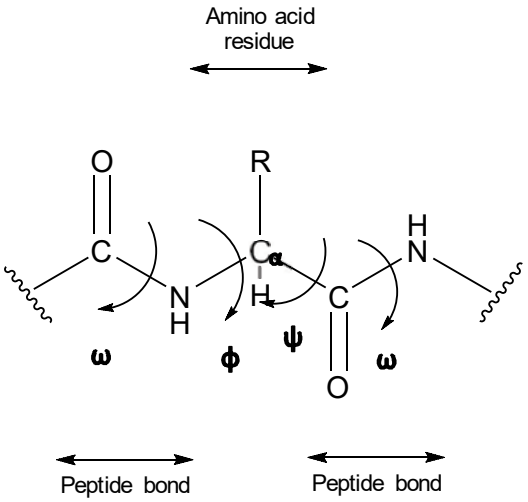


Figure 4. Different bonds rotation

By agreement, the structures are represented from the *N*-terminus to the *C*-terminus. The representation of the L-amino acids, linked by a trans peptide bond, is represented in **Figure 5**[2].

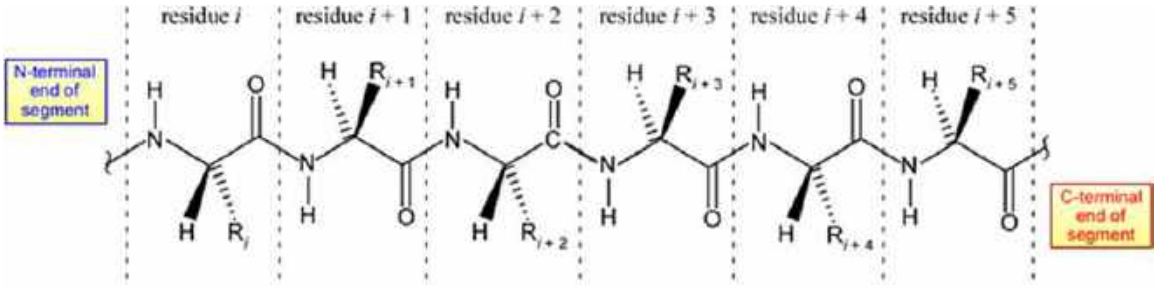


Figure 5. Six-residue polypeptide segment representation[2]

Polypeptide fragments of a protein can fold and adopt different secondary structures. The first one was proposed by Pauling and is the most observed in proteins; it is called the α-helix. This structure has helicoidal shape and C=O and NH groups are linked by H bonds parallel to the main axis; this bond is formed between the amine group of one residue *i* and the carbonyl group from another amino acid *i+4* of the sequence (**Figure 6 (a)**). Another common secondary structure is the β-sheet conformation, in which the different chains interact between them by H bonds perpendicular to the main axis of the polypeptide chains (**Figure 6 (b)**). The tertiary structure is the geometrical and spatial disposition (3D) of the different secondary structures within the protein sequence and the quaternary structure is formed by the interactions of two or more protein sequences

[2] Obtained from http://guweb2.gonzaga.edu/faculty/cronk/CHEM440pub/L09.html (accessed on Jul 23, 2020)

via non-covalent bonds[2]. These non-covalent bonds can be H bonds between amino acid side chains, ionic interactions between one carboxylic group and one amine of different amino acid side chains, hydrophobic interactions which implies the interaction between apolar groups of the amino acid side chains that would be in some way "hided" from the medium or the water in which usually are the proteins, and Van der Waals forces. Disulfide bridges (S-S) between Cys residues, that is, one sulfur atom of a Cys is bonded with another sulfur atom of another Cys of other secondary structure, is a type of covalent bonds that can be formed.

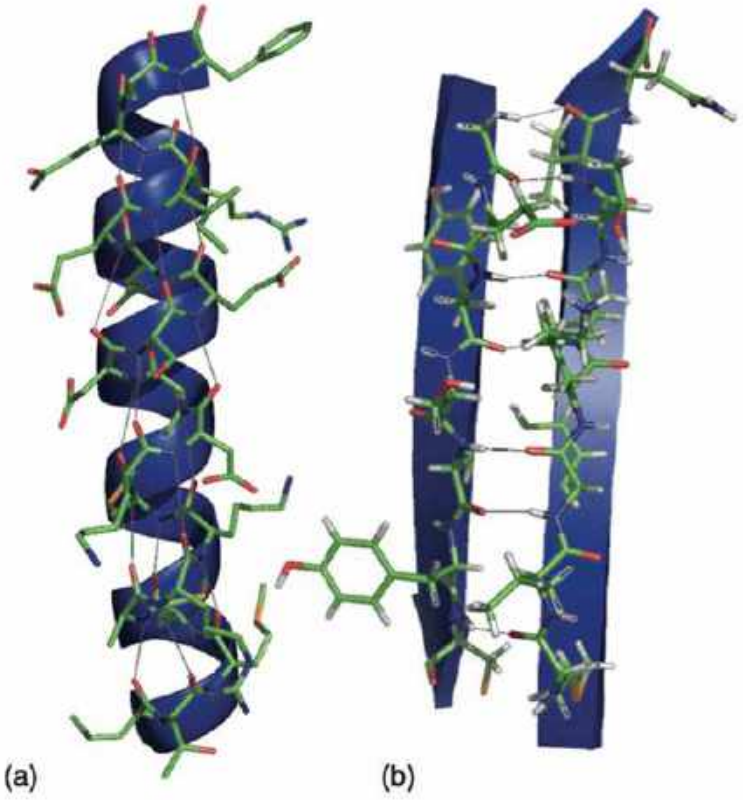


Figure 6. (a) α-helix and (b) β-sheet conformations[3]

Solid Phase Peptide Synthesis (SPPS)

Peptides can be synthesized in solution, in solid phase or by a combination of both. The most common one is the Solid Phase Peptide Synthesis (SPPS) developed by Merrifield in 1965[4]. This method is based on the anchoring of the first amino acid, located at the *C*-terminus of a peptide sequence, to an insoluble polymer resin and, after that, the following amino acids are coupled to the resin with their amine groups and side chains protected (if some reactive group is present), continuing the synthesis

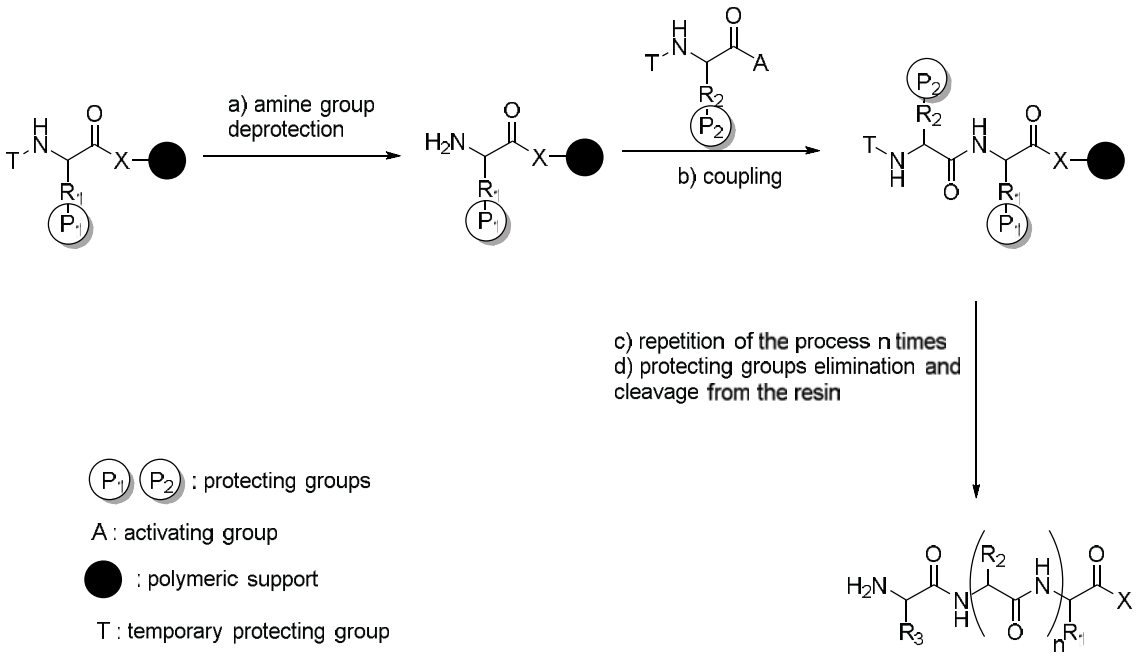
[3] Obtained from https://www.chemtube3d.com/images/aleximages/26-3.png (accessed on Jul 23, 2020)
[4] Merrifield, R. B. Automated Synthesis of Peptides. *Science* **1965**, 150 (3693), 178-185

from the *C*-terminus to the *N*-terminus. For the coupling of each amino acid a 4 step cycle consisting on the cleavage of the amine-protecting group, washing, coupling of a protected amino acid and washing is repeated. The excess of reactants and soluble by-products can be removed by filtering and washing to ensure their correct elimination. In order to achieve a successful synthesis is fundamental to do a good choice of the solid support and protecting groups.

The main characteristic of SPPS is that the product is bonded all the time to a polymeric support. Therefore, the election of the polymeric resin is important because it has to be well solvated in the reaction solvent to allow the efficient entering of the reagents to the immobilized polypeptide chain within the resin. The linker that connects the peptide with the resin is also important because it provides a reversible connection between the peptide and the resin and, at the same time, protects the *C*-terminus while the peptide is synthesized. Furthermore, the linker also determines the functionality of the peptide *C*-terminus.

To carry out effective peptide couplings avoiding side reactions with the amino acid side chains, those must be permanently protected during the peptide synthesis. For this purpose the amino acids have protecting groups that are orthogonal to the amine protecting group, so they are not affected during the deprotection step[5]. As summary of the peptide synthesis methodology, the **Scheme 2** shows the main steps of the SPPS.

[5] Albericio, F.; Isidro-Llobet, A.; Mercedes, A. Amino Acid-Protecting Groups. *Chem. Rev.* **2009**, 2455-2504



Scheme 2. Methodology for the SPPS

For SPPS there are two methodologies that differentiate in the combination of protecting groups used for the peptide synthesis: Boc/Bn or Fmoc/*t*Bu[6].

In Boc/Bn methodology, a Boc (*tert*-butoxycarbonyl) protecting group is used for the temporary protection of the α-amine nitrogen and Bn (benzyl) group is used to protect the amino acids side chains. Boc is deprotected in acid media with TFA (**Figure 7 a)**).

In Fmoc/*t*Bu methodology, an Fmoc (9-fluorenylmethoxycarbonyl) protecting group is used for the temporary protection of the amine and *t*Bu (*tert*-butyl) is used to protect the amino acid side chains[7]. In this case, the Fmoc is easily removed in basic media with piperidine (**Figure 7 b)**).

Amino acid side chain protecting groups are generally selective so that they are cleaved under the same conditions used for the release of the peptide from the resin. In the case of the Boc/Bn strategy, the cleavage requires strong conditions (treatment with HF or trifluoromethanesulfonic acid), while in the case of Fmoc/*t*Bu strategy it is sufficient with TFA treatment.

[6] *Solid-Phase Peptide Synthesis*, 2002363; Global Marketing, Bachem Group, **2016**

[7] Fields, G. B.; Noble, R. L. Solid Phase Peptide Synthesis Utilizing 9-Fluorenylmethoxycarbonyl Amino Acids. *Int. J. Pept. Protein Res.* **2009**, 35, 161-214

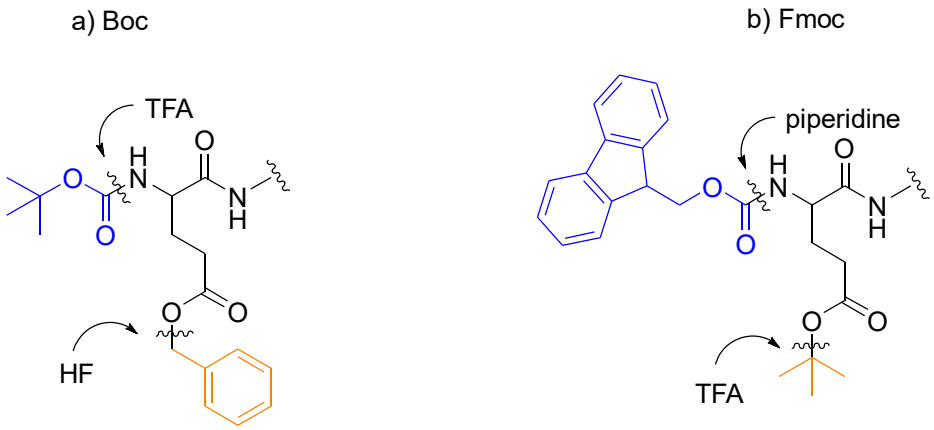
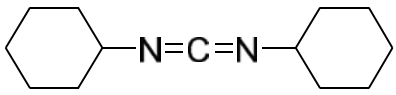


Figure 7. Strategies for the cleavage of Boc/Bn and Fmoc/tBu protecting groups[7]

The Fmoc/*t*Bu methodology was developed because of the problems that the continuous use of TFA could generate on the amino acid side chains. In this strategy, the TFA is only used in the last step of the synthesis to liberate the peptide of the resin and the deprotection of the Fmoc group is carried out with piperidine that is a mild base. In addition, the Boc/Bn strategy presents another disadvantage; the HF that must be used at the final deprotection step is highly dangerous[7].

As it has been mentioned before, the main step of the peptide synthesis is the formation of the peptide bond between amino acids and, in order to make efficient couplings, the carboxyl group must be activated. There are 4 different ways to carry out this activation:

1- Carbodiimides: Are one of the more used activating agents. Their main disadvantage is that they can cause dehydration of Asn and Gln residues, that has to be avoided by adding HOBt to the reaction mixture. One of these activating agents is the *N,N'*-dicyclohexylcarbodiimide (DCC):



DCC

2- Acid anhydrides: Mainly used for Boc synthesis, they are usually generated *in situ* using 1 eq of DCC and 2 eq of the corresponding protected amino acid.
3- Active esters: The peptide coupling requires the use of the derivatized amino acids, (highly reactive esters), as for example with HOBt.
4- *In situ* formation of active esters: Is the most important method because it is easy to use and allows to carry out clean and fast reactions without side reactions. The carboxylic acid is activated *in situ* in order to avoid its

degradation in presence of DIEA, which is the most used base in SPPS[8]. Some of them can be seen in **Figure 8**.

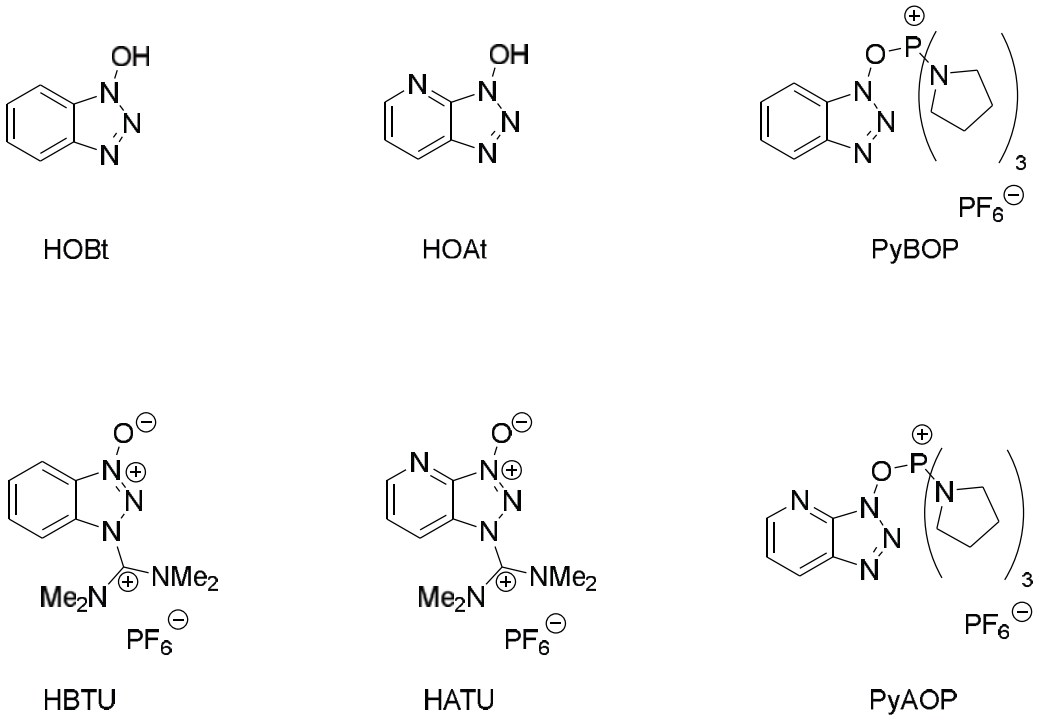
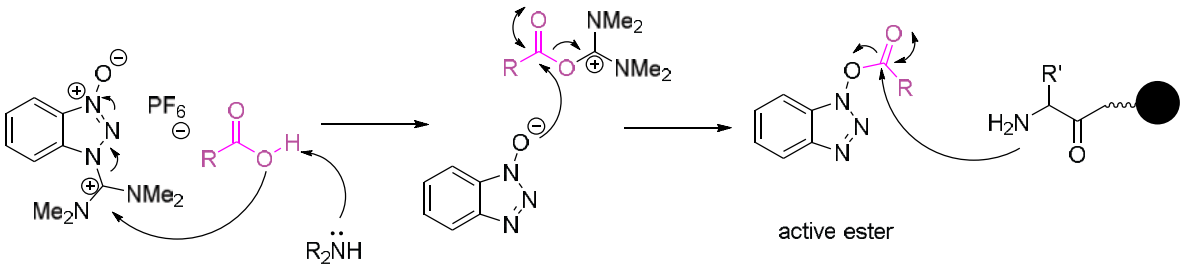


Figure 8. Common coupling agents used in SPPS

The mechanism of the activation reaction, *in situ* formation of the active ester and coupling with the amino acid of the peptide for the case of HBTU is presented in **Scheme 3**:



Scheme 3. Acid activation, in situ formation of the active ester and coupling with the amino acid of the peptide

The most important advantage in solid phase synthesis over the classical methodology in solution is that reagents can be used in excess to guarantee high yields in each step of the synthesis. The main limitation of this methodology is that the side products that could be generated on each step of the synthesis would be accumulated, and at the

[8] Albericio, F.; Bofill, J. M.; El-Faham, A.; Kates, S. A. Use of Onium Salt-Based Coupling Reagents in Peptide Synthesis. *J. Org. Chem.* **1998**, *63*, 9678-9683

end, the separation of the desired peptide from them will be difficult because their chromatographic characteristics are quite similar.

Viologens in supramolecular chemistry

Viologens are *N,N'*-disubstituted 4,4'-bipyridine derivatives commonly used as building blocks in supramolecular chemistry due to their interesting chemical properties (**Figure 9**).

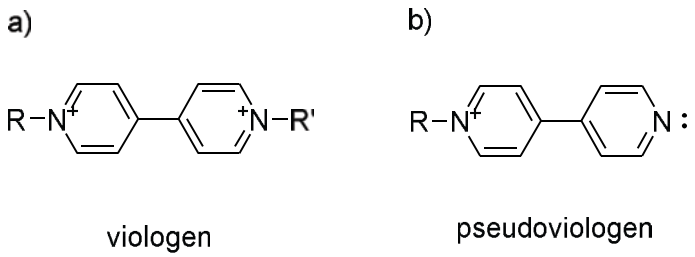
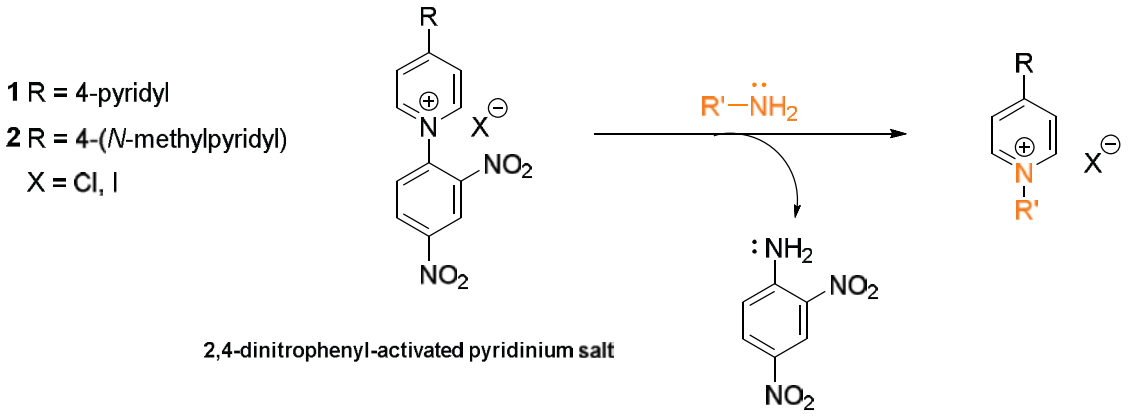


Figure 9. General structure of viologens and pseudoviologens

When both pyridines are substituted the structure is called viologen (**Figure 9 a)**) and when only one pyridine is substituted the structure is called pseudoviologen (**Figure 9 b)**).

Viologens can be easily prepared using nucleophilic aromatic substitution reactions with alkyl halides or by the Zincke reaction.

The well-known Zincke reaction involves the use of a primary amine and a Zincke salt (2,4-dinitrophenyl-activated pyridinium salts) yielding a new pyridinium salt in which the amine group in incorporated in the heterocycle (**Scheme 4**).

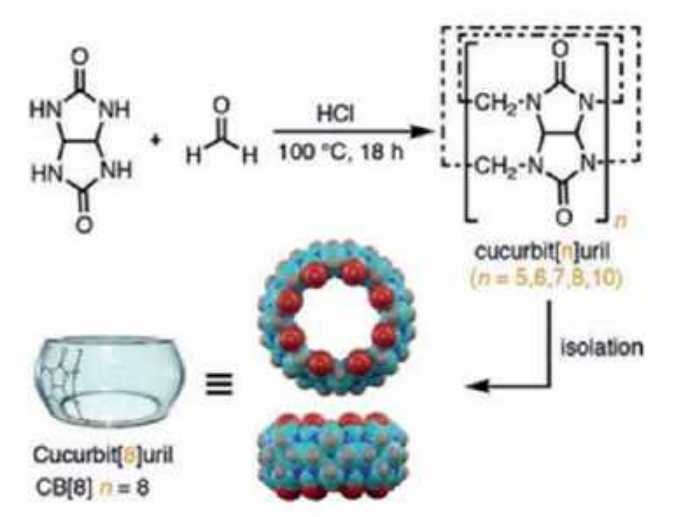


Scheme 4. General scheme of the Zincke reaction

Viologens are π-deficient aromatic molecules and good electron-acceptors widely used for the development of molecular switches due to their reversible and easily accessible

one- or two-electron reductions ($V^{2+} + e^- \leftrightarrows V^{+\cdot} + e^- \leftrightarrows V$). Furthermore, because of the positive charges located over both ends of the viologens, they can be encapsulated as guests and form supramolecular inclusion complexes with cucurbit[*n*]uril macromolecules[9].

Cucurbit[*n*]urils (CB[*n*]) are a family of pumpking-shaped macromolecules obtained by the reaction between formaldehyde and *n* glycoluril units joined together by 2*n* methylene bridges (**Scheme 5**)[10]. Their inner hydrophobic cavity is accessible through two identical carbonyl portals, and are capable of forming inclusion complexes with neutral or cationic organic guests by means of hydrophobic forces, cation-dipole interactions, and optimization of host-guest packing coefficients.



Scheme 5. Scheme of the reaction between formaldehyde and n glycoluril units[10]

CB[8] with a cavity of 493 Å^3 was first synthesized two decades ago[11]. It was the first reported molecule of the cucurbituril family capable of forming 1:2 complexes with appropriate guests.

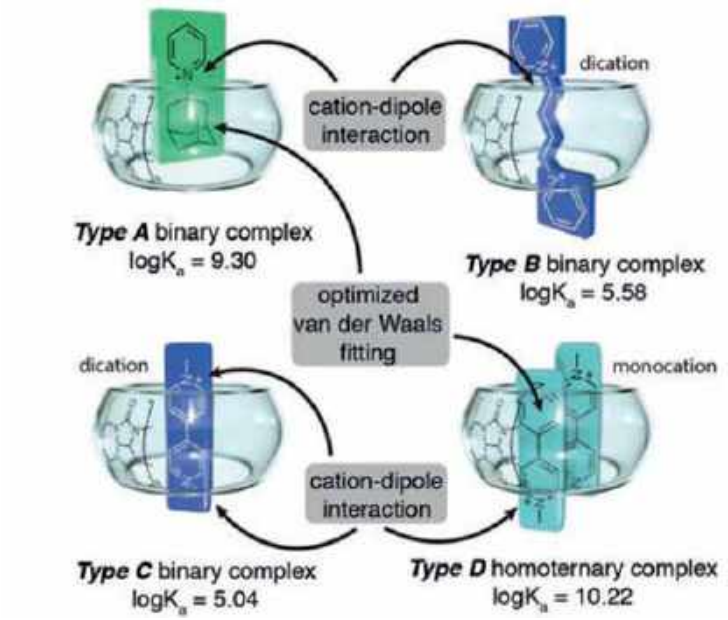
It forms Type A binary complexes with monocationic guests with bulky hydrophobic moieties. Type B and Type C binary complexes are formed with dicationic guests and differ on the nature of the aliphatic or electron-acceptor aromatic regions introduced in the cavity of the host. As mentioned before, due to the larger cavity of CB[8], Type D

[9] Cortón, P.; Novo, P.; López-Sobrado, V.; García, M. D.; Peinador, C.; Pazos, E. Solid-Phase Zincke Reaction for the Synthesis of Peptide-4,4'-bipyridinium Conjugates. *Synthesis* **2020**, 52

[10] Pazos, E.; Novo, P.; Peinador, C.; Kaifer, A. E.; García, M. D. *Angew. Chem. Int. Ed.* **2019**, 58, 403

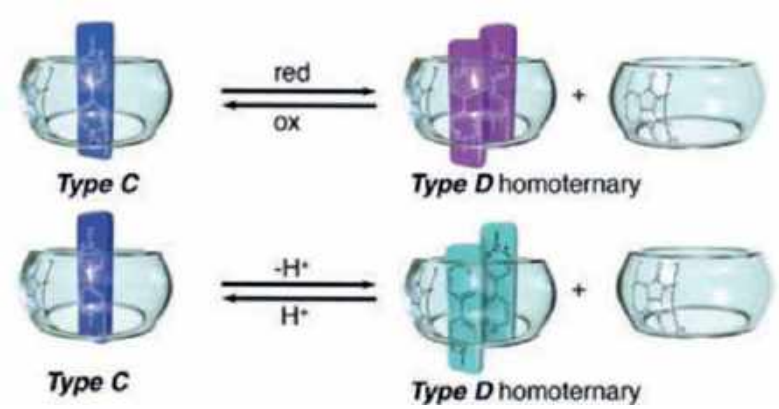
[11] Kim, J.; Jung, I.-S.; Kim, S.-Y.; Lee, E.; Kang, J.-K.; Sakamoto S.; Yamaguchi, K.; Kim, K. *J. Am. Chem. Soc.* **2000**, *122*, 540-541

homoternary 1:2 host-guest complexes can be formed as well, with two complementary, antiparallel and identical monocationic guests (**Scheme 6**)[10].



Scheme 6. Expected Type A-D CB[8]-guest complexes

Taking into account the properties of both molecules, CB[8] as host and bipyridine-based cations as guests, Kim *et al*. reported a Type C 1:1 complex between CB[8] and dimethyl-viologen (MV^{2+}) in aqueous media. These complexes explained by Kim are supramolecular switches, they are capable of interchange from a Type C binary complex to a Type D homoternary complex after the application of external stimuli. When CB[8] is combined with viologens, it is sensible to redox changes and when CB[8] is combined with pseudoviologens, it is sensible to pH (**Scheme 7**)[12].



Scheme 7. Intramolecular redox and acid-base Type C-D switches

[12] Jeon, W. S.; Kim, H.-J.; Lee, C.; Kim, K. *Chem. Commun.* **2002**, 1828-1829

Therefore, considering the huge potential of viologen:CB[8] inclusion complexes for the development of stimuli-responsive systems, and the interest in the introduction of dynamic moieties into peptide sequences that allow to reversibly control their function by means of host-guest chemistry, the development of straightforward synthetic methods that allow to introduce viologen derivatives in peptide sequences is of great interest.

Background

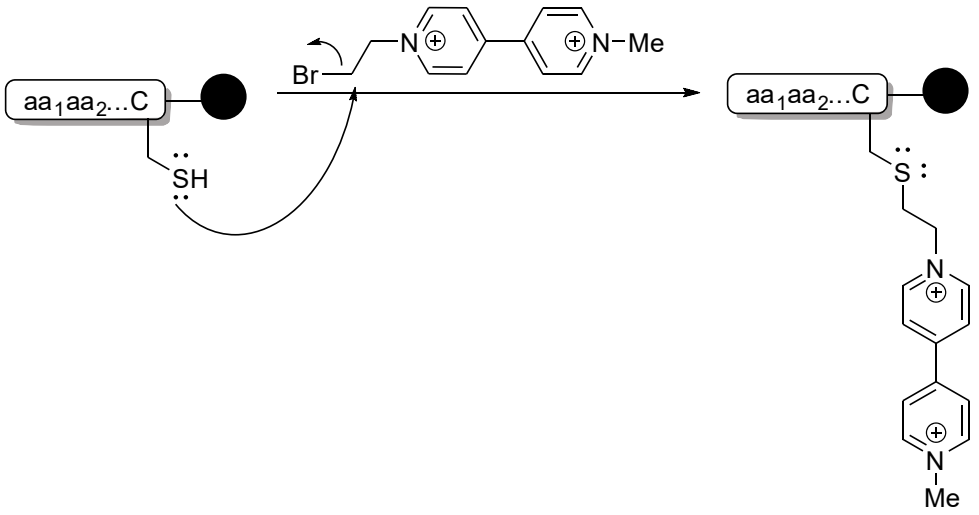
Peptides are ideal platforms for the development of biosensors, drugs, etc. due to their synthetic simplicity and functional and structural versatility. In this regard, peptide-based conjugates have demonstrated their huge potential in many research areas, such as in the biomedical field. For these reasons, over the last years it has increased the interest in the development of synthetic methodologies that allow an efficient and straightforward functionalization of peptides.

In this context, peptides have been extensively modified through the introduction on their structure of appropriately functionalized non-natural amino acids, and the subsequent use of a variety of highly efficient reactions, like Staudinger, Diels-Alder or Pd-catalyzed[9] cross coupling reactions, to incorporate the desired functionalities. Moreover, recent advances in organic reactions, as for example C-H activation, have also paved the way to the selective modification of natural amino acid residues.

However, despite these advances in peptide modification, there is a lack of methodologies for the non-covalent modification of peptides. In this regard, as it has been previously mentioned in the introduction, the incorporation of viologens in peptide sequences is highly interesting because those could act as stimuli-responsive guests in supramolecular complexes, allowing to reversibly control the functionality of peptides by host-guest chemistry.

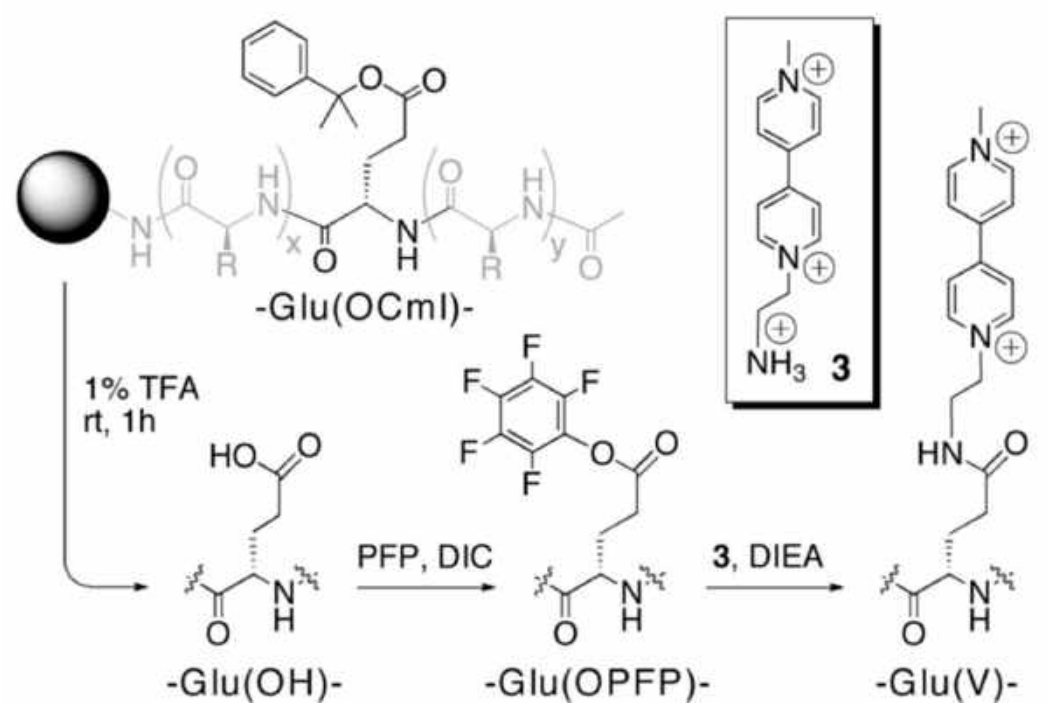
In the literature there are a few of examples that describe the modification of peptides with viologens. The first example was reported by the group of K. Plaxco in 2007 who described the conjugation of a viologen-functionalized bromide to a cysteine residue of a peptide sequence[13] (**Scheme 8**). The main limitation of this methodology comes from the need of having a thiol group (cysteine residue) in the peptide sequence, which are sometimes easily oxidized to a disulfide bond.

[13] Oh, K. J.; Cash, K. J.; Hugenberg, V.; Plaxco, K. W. *Bioconjugate Chem.* **2007**, 18, 607



Scheme 8. Reaction carried out by the group of K. Plaxco

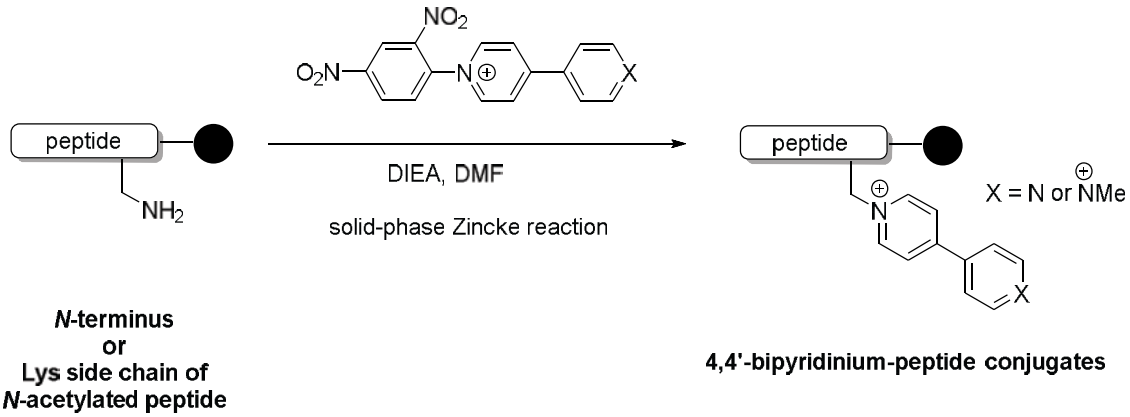
Three years later the group of A. Urbach developed a methodology for the conjugation of amine-functionalized viologens to the side chain of Glu residues by a peptide bond. This strategy results in viologen-peptide conjugates in which the viologen unit is connected to the peptide backbone by a relatively large linker coming from the alkyl amine connector and the Glu side chain (**Scheme 9**)[14].



Scheme 9. Viologen-peptide conjugate synthesis using Glu side chains

[14] Reczek, J. J.; Rebolini, E.; Urbach, A. R. Solid-Phase Synthesis of Peptide-Viologen Conjugates. *J. Org. Chem.* **2010**, 75 (6), 2111-2114

More recently, the research groups of Prof. Carlos Peinador and Dr. Elena Pazos, from Universidade da Coruña, have developed a new methodology for the preparation of viologen-peptide conjugates consisting on the use of a solid-phase Zincke reaction. In this case, they have shown that this reaction can be used to modified short peptides with viologens and pseudoviologens, both at their *N*-terminus and in a Lys side chain located at the *C*-terminus of the peptides. Furthermore, they have also demonstrated that pseudoviologen derivatives can be coupled in a Lys side chain located at the *C*-terminus of a long (28 amino acid) and highly charged peptide. Therefore, this method is a versatile and clean strategy that, *a priori*, allows the introduction of any activated viologen derivative in any amino group of a peptide (**Scheme 10**).



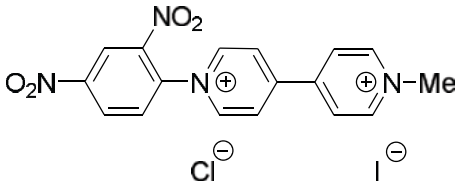
Scheme 10. General scheme of the coupling of the (pseudo)viologen to the peptide

Objectives

As part of a larger project, which aims to develop 4,4'bipyridinium-peptide conjugates as artificial transcription factors that allow to control their dimerization and function in response to external stimuli, the main objective of this work is to synthesize a 4,4'bipyridinium-peptide derivative that might respond to redox changes.

Therefore, the specific objectives of this project are:

- The synthesis of the *N*-activated viologen derivative (compound **2**) starting from 4,4'-bypiridine:

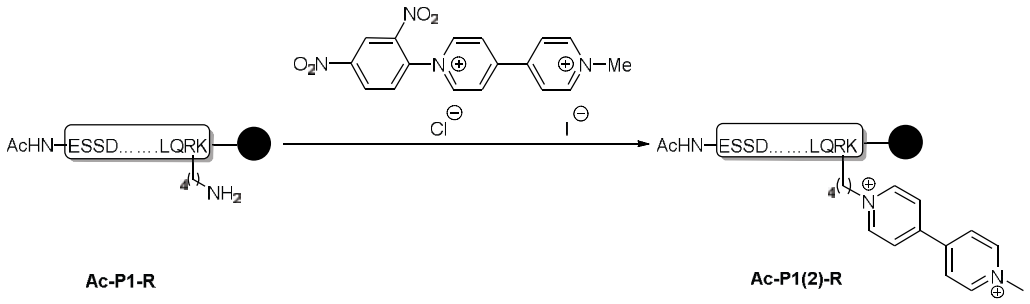


2

- The synthesis of an alloc-protected peptide derived from the basic region of the GCN4 transcription factor by SPPS (**Fmoc-P1(alloc)-R**) with the following sequence:

Fmoc-ESSDPAALKRARNTEAARRSRARKLQRK(alloc)-NH_2

- The coupling of the viologen derivative **2** to the orthogonally deprotected lysine side chain, located at the *C*-terminus, using a solid phase Zincke reaction to obtain the desired viologen-peptide conjugate **Ac-P1(2)**:

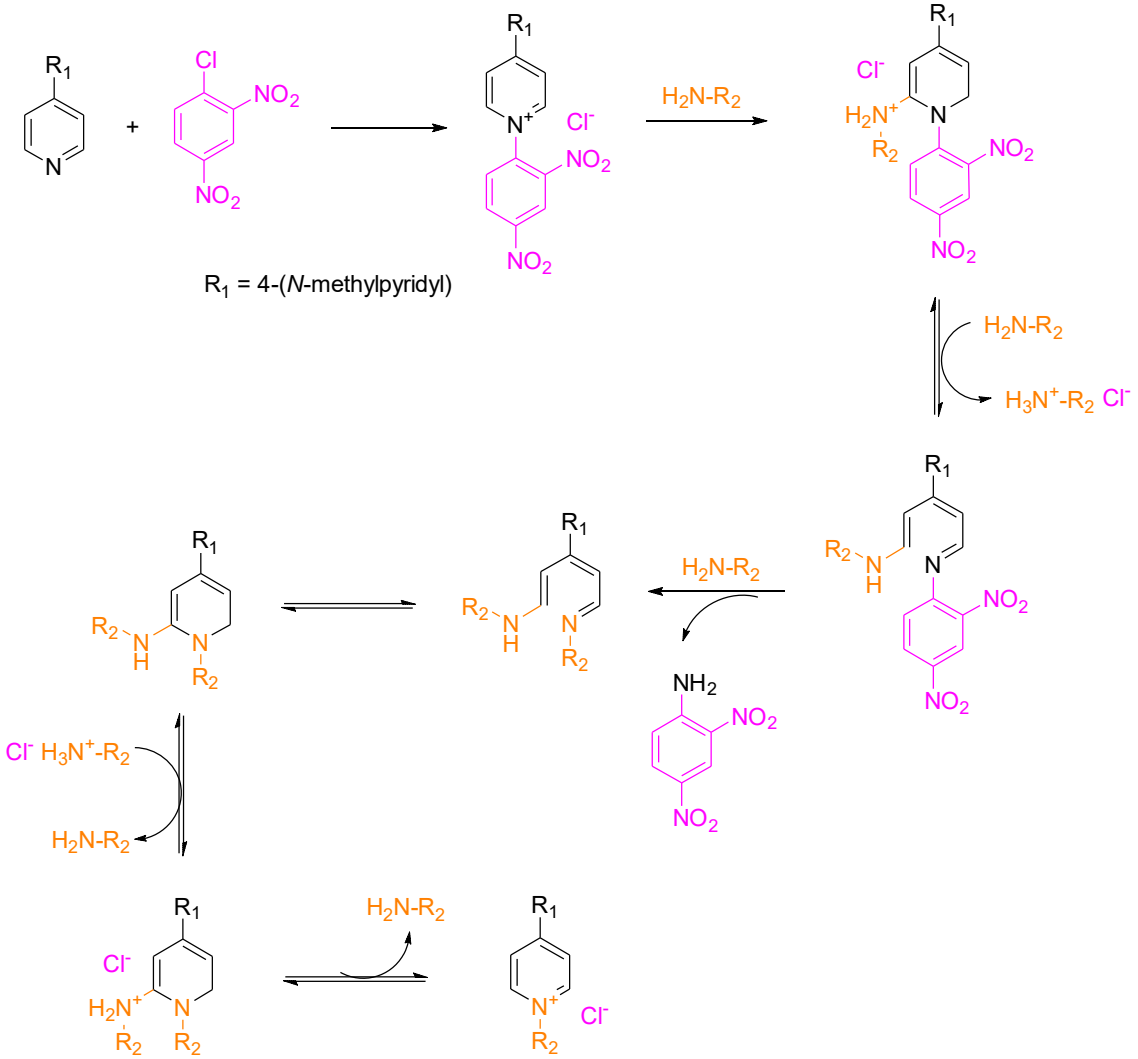


Ac-P1-R **Ac-P1(2)-R**

Results and discussion

Synthesis of the *N*-activated 4,4'-bipyridinium salt 2 as precursor for the Zincke reaction

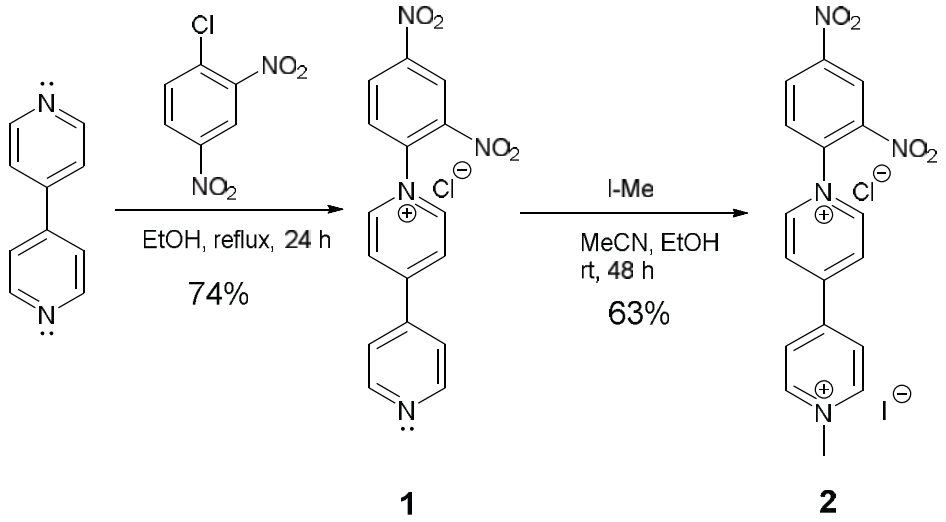
The *N*-activated 4,4'-bipyridinium salt **2**, required for the coupling to the peptide through a Zincke reaction that follows an ANROC mechanism (Nucleophilic Addition, Ring Opening and Ring Closure)[15], was obtained from the reaction of 1-chloro-2,4-dinitrobenzene with 4,4'-bipyridine. The mechanism presented in **Scheme 11**, is intended to explain more thoroughly the single steps of the Zincke reaction.



Scheme 11. ANROC mechanism for the Zincke reaction[15]

[15] Cheng, W. C.;Kurth, M. J. The Zincke reaction. A review. *Org. Prep. Proced. Int.* **2002**, 34 (6), 585-608

Intermediate **1** was synthesized in a 74% yield by a nucleophilic aromatic substitution between 1-chloro-2,4-dinitrobenzene and 4,4'-bipyridine in EtOH under reflux (**Scheme 12**). The desired product was precipitated with Et_2O, washed with AcOEt and its identity was confirmed by ^{1}H-NMR.



Scheme 12. Synthesis of N-activated pyridinium salts 1 and 2

As shown in **Scheme 12**, compound **2** was synthesized by methylation of compound **1** with MeI in a mixture of MeCN and EtOH at room temperature. Compound **2** was synthesized in a 63% yield.

Peptide synthesis

The peptide sequence ESSDPAALKRARNTEAARRSRARKLQRK derived from the basic region of the GCN4 transcription factor, was chosen for the introduction of the viologen derivative **2** by a solid phase Zincke reaction. In order to incorporate the viologen at the *C*-terminus of the peptide sequence, an orthogonally-protected Lys (K) was introduced at this end with the aim of having the desired selectivity for the Zincke reaction. All amino groups of the amino acids are protected with Fmoc, which are removed during the peptide synthesis with piperidine (mechanism shown in **Scheme 13**). The amino acids that have reactive side chains were also introduced with their side chains protected: Arg side chains were protected with Pbf (2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-sulfonyl); Gln and Asn with Trt (trityl); Lys (except the one located at the *C*-terminus) with Boc; Ser, Thr, Asp and Glu with *t*Bu. Their structures are shown in **Figure 10**[16].

[16] Gregersen, S. Flourescent peptide-stabilized silver-nanoclusters, a solid-phase approach for high-throughput ligand discovery. Ph.D. Thesis, Aalborg University, Aalborg (Denmark), **2014**

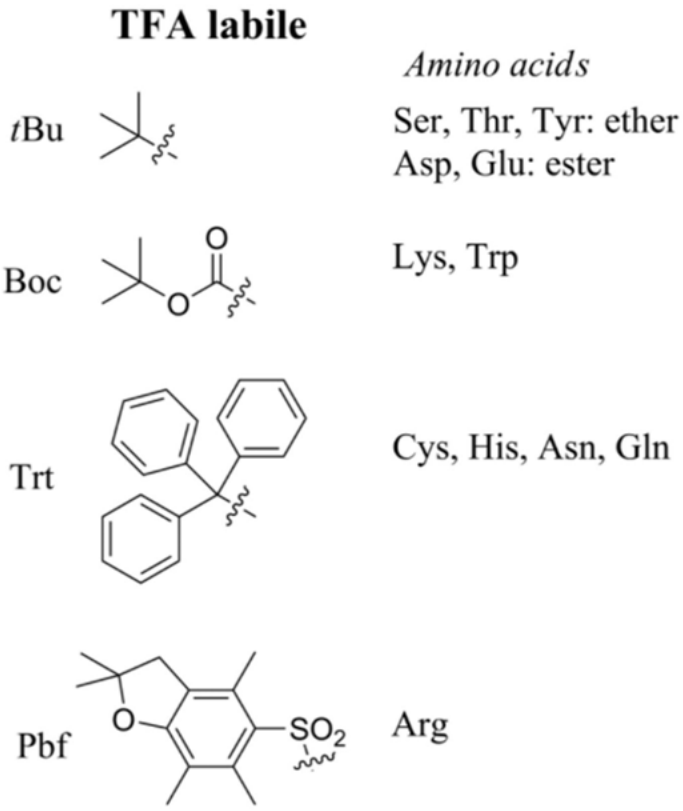
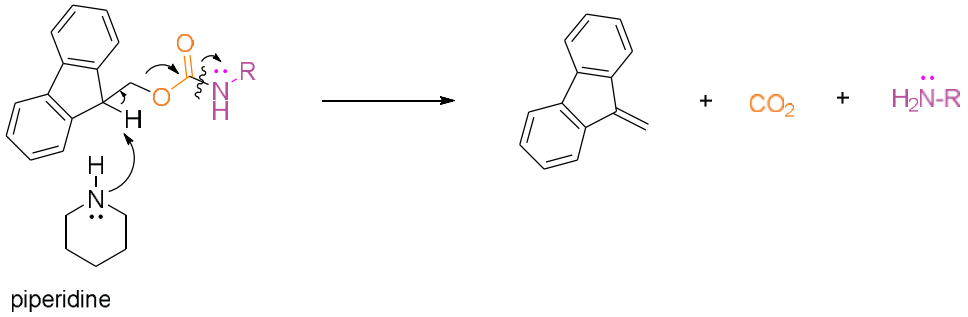


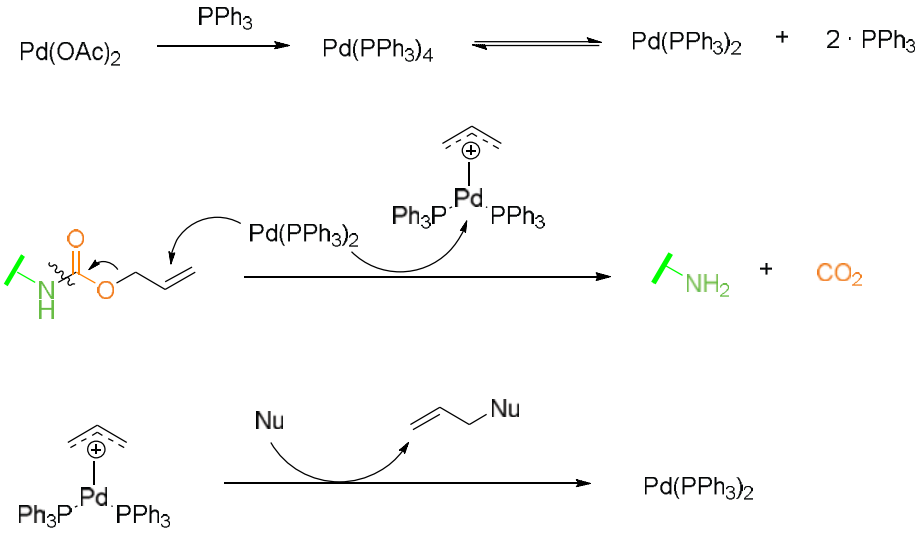
Figure 10. TFA labile protecting groups used

As all of them are TFA labile, they are going to be removed in the same step. For the Lys residue located at the *C*-terminus, which has its side chain protected with an alloc group, is orthogonally deprotected using palladium catalysis ($Pd(PPh_3)_4$ generated *in situ* in the reaction media). The Lys deprotection mechanism can be seen in **Scheme 13**.

Fmoc deprotection:



Alloc deprotection:



Scheme 13. Fmoc and Alloc deprotection mechanism schemes

The peptide sequence was successfully synthesized by SPPS using standard Fmoc solid-phase peptide synthesis protocols consisting on: 1) addition of the activated Fmoc-protected amino acids 2) removal of the Fmoc protecting group and 3) repetition of steps 1 and 2 with all the amino acids of the peptide sequence. Every coupling was carried out in DMF for 30 min.

After the synthesis of the whole peptide sequence Fmoc-ESSDPAALKRARNTEAARRSRARKLQRK-NH_2 was completed, a small aliquot of the peptide was cleaved from the resin and the crude was analyzed by HPLC-MS. As can be observed in the HPLC chromatogram shown in **Figure 11**, there is a main peak at t_R=12,2 min that was identified by MS (**Figure 12**) as the desired peptide.

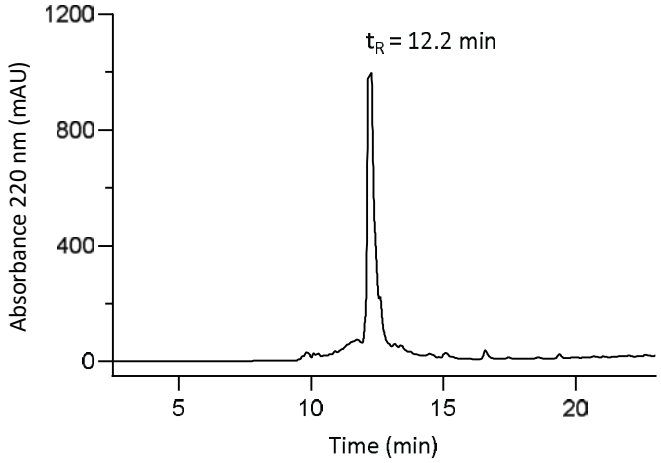


Figure 11. HPLC chromatogram of the reaction crude in which the peak at 12,2 min corresponds to Fmoc-P1(alloc)-R

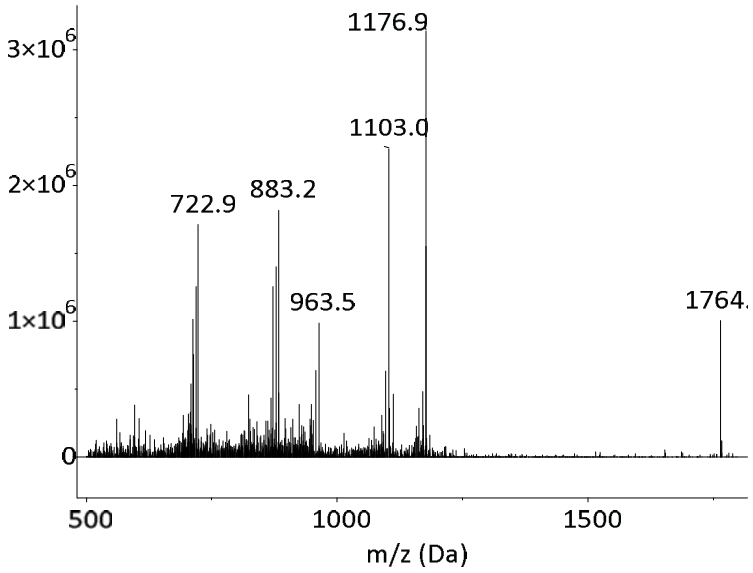
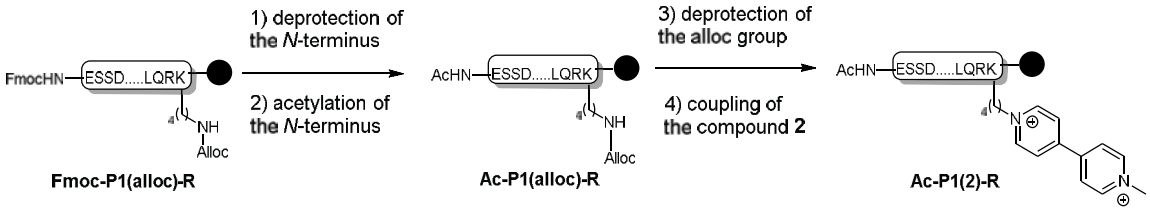


Figure 12. ESI-MS spectrum of the HPLC chromatogram peak at 12,2 min corresponding to Fmoc-P1(alloc)-R

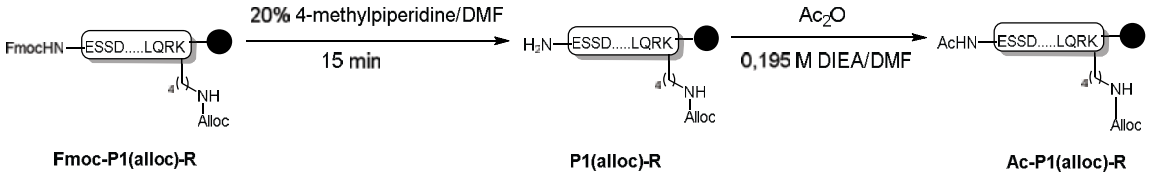
Viologen-peptide conjugation

In order to perform the Zincke reaction over the deprotected Lys side chain, the peptide has to be first acetylated at the *N*-terminus before proceeding with the orthogonal deprotection of the Lys side chain.



Scheme 14. General scheme of the viologen-peptide conjugation

- Acetylation



Scheme 15. Scheme showing the deprotection and acetylation of the N-terminus

In order to acetylate the *N*-terminus of the peptide, first a deprotection of the Fmoc protecting group has to be done. For this purpose, **Fmoc-P1(alloc)-R** was treated with 20% 4-methylpiperidine in DMF for 15 min and the product of this reaction was treated with Ac_2O in 0,195 M DIEA/DMF to obtain the acetylated peptide. After the acetylation step, a small aliquot of the peptide was cleaved from the resin and the crude was analyzed by HPLC-MS. As can be observed in **Figure 13**, the HPLC chromatogram of the reaction crude shows a main peak at t_R=10,9 min that was identified by MS (**Figure 14**) as **Ac-P1(alloc)-R**. The smallest peak at t_R=10,5 min could not be identified by MS.

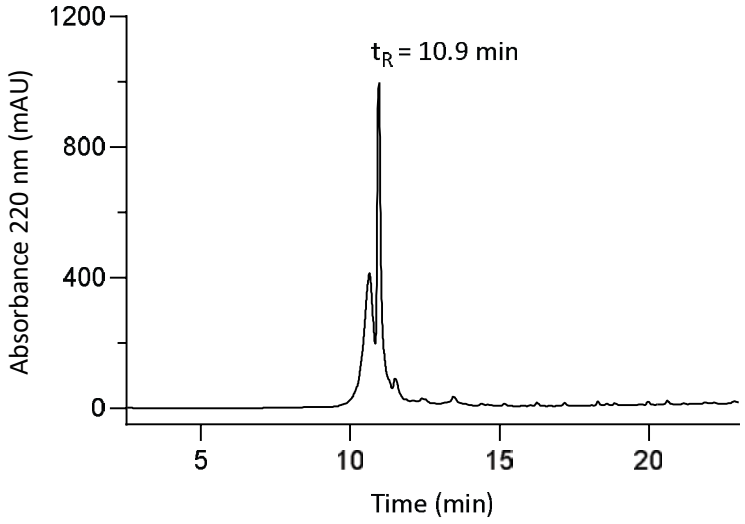


Figure 13. HPLC chromatogram of the reaction crude in which the peak at 10,9 min corresponds to Ac-P1(alloc)-R

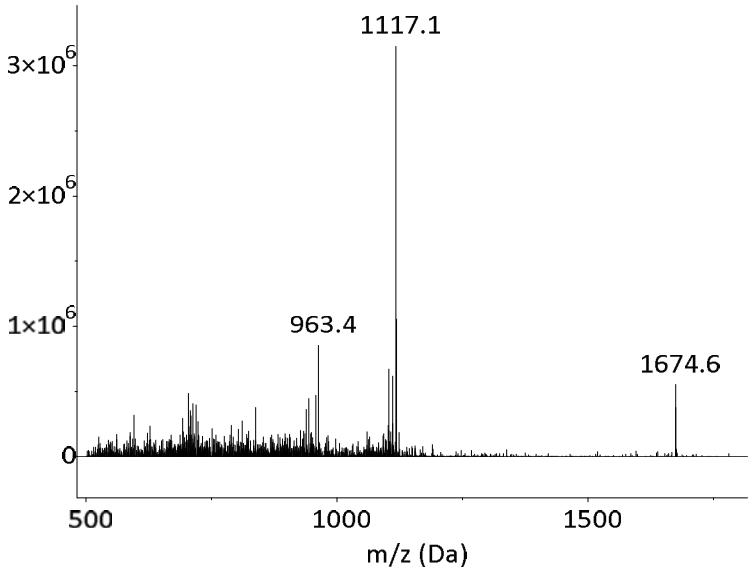


Figure 14. ESI-MS spectrum of the HPLC chromatogram peak at 10,9 min corresponding to Ac-P1(alloc)-R

- Deprotection of the Lys(alloc) side chain

Ac-P1(alloc)-R **Ac-P1-R**

Scheme 16. Deprotection of the Lys side chain

The orthogonal removal of the alloc group from the Lys side chain at the *C*-terminus of **Ac-P1(alloc)-R** was carried out by treating the resin with a mixture of 0,3 eq of $Pd(OAc)_2$, 1,5 eq of PPh_3, 10 eq of NMM (*N*-methylmorpholine) and 10 eq of $PhSiH_3$ (phenylsilane) in DCM (dichloromethane) overnight. After the deprotection step, a small aliquot of **Ac-P1-R** was cleaved from the resin and the crude was analyzed by HPLC-MS. As can be observed in **Figure 15**, the HPLC chromatogram of the reaction crude shows a main peak at t_R=9,9 min that was identified by MS as **Ac-P1-R**. The smallest peak at t_R=9,5 min could not be identified by MS.

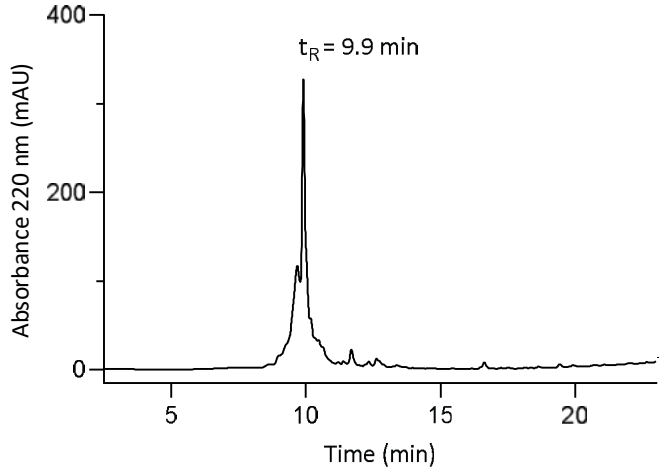


Figure 15. HPLC chromatogram of the reaction crude in which the peak at 9,9 min corresponds to Ac-P1-R

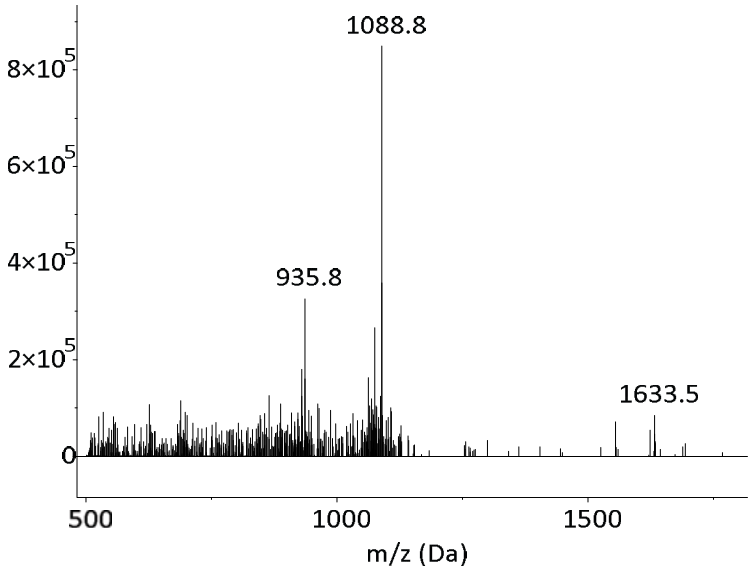
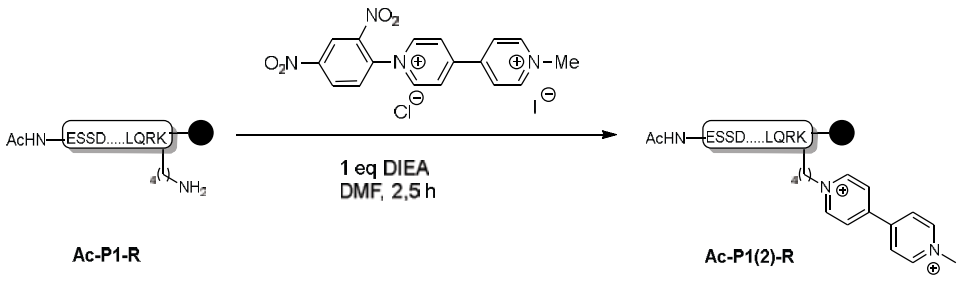


Figure 16. ESI-MS spectrum of the HPLC chromatogram peak at 9,9 min corresponding to Ac-P1-R

- Compound **2** coupling to the Lys side chain



Scheme 17. Viologen-peptide coupling at the Lys side chain

The coupling of compound **2** to the Lys side chain was performed by treating the resin **Ac-P1-R** with a mixture of **2** (3 eq) and DIEA (1 eq) in DMF for 2,5 h. After the coupling step, **Ac-P1(2)-R** was cleaved from the resin and purified.

As can be observed in **Figure 17**, the HPLC chromatogram of the purified peptide conjugate **Ac-P1(2)** only shows one peak at t_R=9,6 min that was identified by MS as the desired viologen-peptide conjugate.

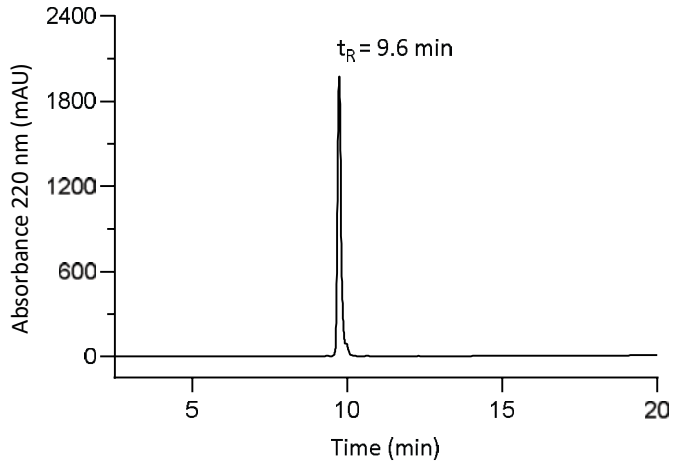


Figure 17. HPLC purified chromatogram of Ac-P1(2)

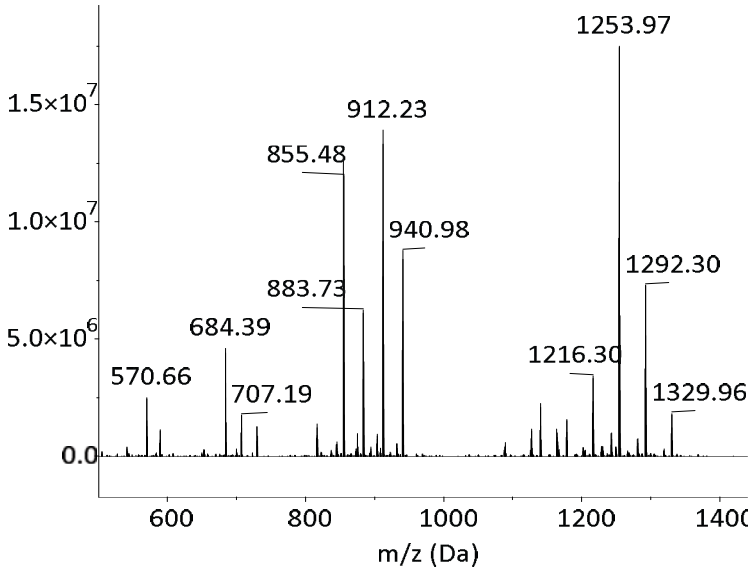


Figure 18. ESI-MS spectrum of the HPLC chromatogram peak at 9,6 min corresponding to Ac-P1(2)

Experimental procedure

General information

All solvents used were synthesis grade, except for dimethylformamide (DMF), trifluoroacetic acid (TFA), and diisopropylethylamine (DIEA) that were peptide synthesis grade. Water was purified using a Milli-Q system (*Millipore*).

TLC was the technique used in order to follow the reactions for pseudoviologen and viologen. Due to the polarity and the positive charge of these molecules, a mixture of H_2O (150 mL), MeOH (150 mL), MeCN (600 mL) and NaCl (5 g) was used as eluent for the chromatography. The same mixture was used for the purification method. Products were identified with UV light at 254 nm and 360 nm.

Small molecules were characterized by ^{1}H-NMR and ^{13}C-NMR in a *Bruker Avance 300* (300 MHz). Deuterated water (D_2O) and deuterated MeCN (CD_3CN) were used for the preparation of the samples.

Reversed-phase HPLC-MS analyses were performed using a *Thermo Scientific UltiMate 3000* connected to a single quadrupole mass spectrometer *Thermo Scientific MSQ Plus* and a PDA (Photo-Diode Array) detector.HPLC-MS quality solvents were used to prepare the eluents, A: 0,1% TFA, H_2O and B: 0,1% TFA, MeCN. The column used for these analyses was *Phenomenex Aeris* 3,6 µm peptide XB-C18 100 Å; 150 x 2,1 mm. The gradient used for reversed-phase HPLC-MS experiments was the following:

Time (min)	Flow (mL/min)	% B
0	0,300	5
2	0,300	5
25	0,300	95
26	0,300	100
32	0,300	100
33	0,300	5
40	0,300	5

Table 2. Gradient for HPLC and HPLC-MS analysis

Reversed-phase HPLC purification was carried out in an *Agilent 1200 series* and the column used was a *Phenomenex Luna* 5 µm C18 100 Å; 250 x 10 mm.

The injections volumes were increased from 10 µL to 100 µL. The gradient was also modified between injections in order to obtain better separation of the peaks, starting

from the gradient shown above for the analytical HPLC with a flow of 2,500 mL/min. For 100 µL injections, the gradient used was:

Time (min)	Flow (mL/min)	% B
0	2,500	15
5	2,500	15
45	2,500	50
55	2,500	95

Table 3. Final gradient for semipreparative HPLC purification

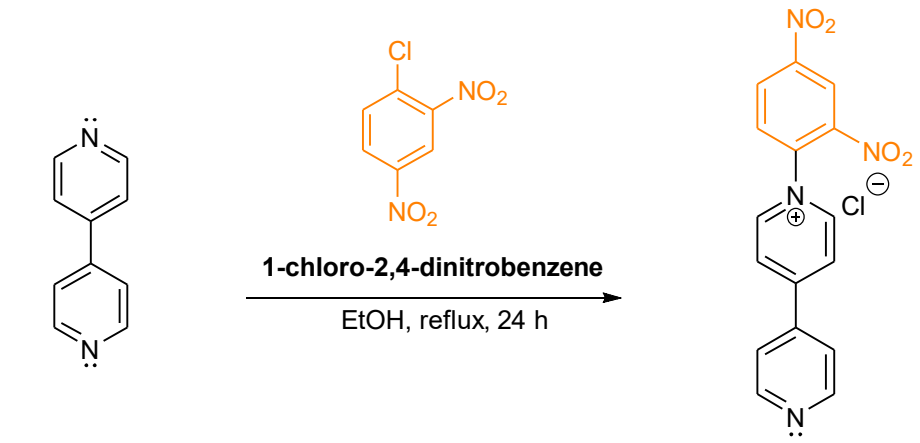
The purified fractions were collected and lyophilized using a *Teslar Cryodos -80* instrument.

The final viologen-peptide conjugate was characterized by HPLC-MS using a UHPLC *Thermo Scientific Accela* connected to a linear trap quadrupole mass spectrometer *Thermo Scientific LTQ-Orbitrap Discovery* and a PDA detector from the Research Support Facilities (SAI, Servizos de Apoio á Investigación) of the UDC.

Synthesis

1-(2,4-dinitrophenyl)-1'-methylbipyridinium chloride iodide synthesis

Synthesis of 1-(2,4-dinitrophenyl)-4,4'-bipyridinium chloride (1)

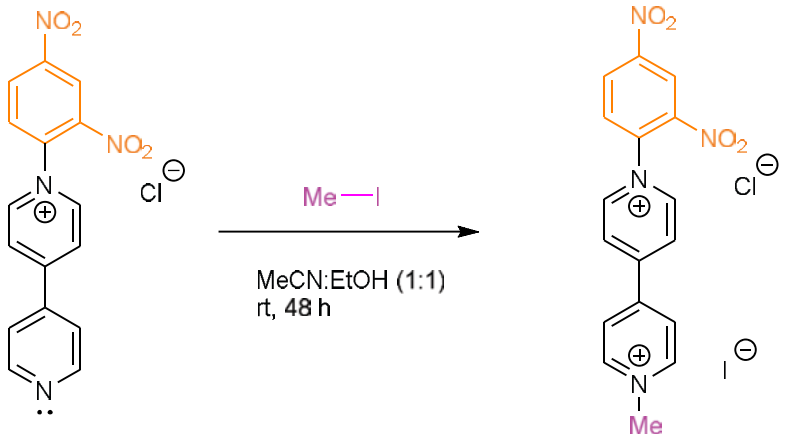


4,4'-bipyridine (1,9921 g, 12,8 mmol) and 1-chloro-2,4-dinitrobenzene (12,8 mmol, 2,5964 g) were dissolved in 25 mL of EtOH, and then the mixture was refluxed for 24 h. The reaction was followed by TLC and quenched by cooling down to room temperature. The product was precipitated with 200 mL of Et_2O yielding a white solid. The precipitate was filtrated and washed with AcOEt, then dissolved in MeOH and

concentrated to dryness to give compound **1** as a brown-yellowish solid in a 74% yield (3,396 g).

^{1}H-NMR (300 MHz, D_2O) δ (ppm): 9,41 (d, *J*=2,5 Hz, 1 H), 9,26 (d, *J*=7,0 Hz, 2 H), 8,95 (dd, J_1=8,7 Hz, J_2=2,5 Hz, 1 H), 8,86 (d, *J*=6,4 Hz, 1 H), 8,70 (d, *J*=7,0 Hz, 2 H), 8,28 (d, *J*=8,7 Hz, 1 H), 8,05 (d, *J*=6,3 Hz, 2 H).

Synthesis of 1-(2,4-dinitrophenyl)-1'-methylbipyridinium chloride iodide (2)



1 **2**

Compound **1** (0,5007 g, 1,396 mmol) was dissolved in a 1:1 mixture of MeCN and EtOH (30 mL). Then, MeI (1,736 mL, 27,89 mmol) was added and the resulting mixture was stirred at room temperature, observing a light orange color immediately after the addition of MeI. After 24 h, the color of the solution was dark red, but after checking by TLC the reaction was not completed. MeI (0,868 mL, 13,9 mmol) was added again to the mixture and the reaction was stirred again for 24 h. After this time, the solution was still red and a precipitate was formed. The mixture was vacuum filtered and the precipitate was washed with 15 mL of MeCN. Compound **2** was isolated as a solid in a 63% yield (0,4401 g).

^{1}H-NMR (300 MHz, D_2O) δ (ppm): 9,46 (m, 3 H), 9,17 (d, 2 H, *J*=6,5 Hz), 9,00 (dd, 1 H, J_1=2,5 Hz, J_2=8,7 Hz), 8,87 (d, 2 H, *J*=7,0 Hz), 8,69 (d, 2 H, *J*=6,9 Hz), 8,35 (d, 1 H, *J*=8,7 Hz), 4,58 (s, 3 H).

^{13}C-NMR (300 MHz, D_2O) δ (ppm): 153,3 (1 C), 149,9 (1 C), 149,3 (1 C), 146,6 (4 C), 142,9 (1 C), 138,3 (1 C), 131,2 (1 C), 130,8 (1 C), 127,2 (2 C), 127,1 (2 C), 122,8 (1 C), 48,6 (1 C).

HR-MS (ESI): m/z $[M]^+$ calculated for $C_{17}H_{14}N_4O_4$: 338,1009

found: 338,1023

Synthesis of the GCN4 peptide derivative

- Resin: H-Rink amide ChemMatrix resin (0,47 mmol/g)
- Coupling agent: HBTU/HOBt 0,2 M
- Base: DIEA/DMF 0,195 M
- TNBS test for the protection/deprotection of the amine group: 1% TNBS/DMF + 10% DIEA/DMF
- Fmoc deprotecting agent: 4-methylpiperidine
- Sequence: Glu-Ser-Ser-Asp-Pro-Ala-Ala-Leu-Lys-Arg-Ala-Arg-Asn-Thr-Glu-Ala-Ala-Arg-Arg-Ser-Arg-Ala-Arg-Lys-Leu-Gln-Arg-Lys

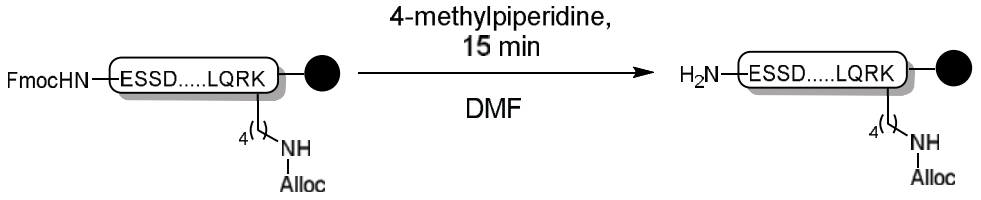
The peptide sequence **Fmoc-P1(alloc)-R** was assembled in a 0,05 mmol scale using a PS3 Automatic Peptide Synthesizer from *Protein Technologies Inc*. Amino acids were coupled to the resin in 4-fold excess using 2-[(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate] (HBTU) as activating agent. Each amino acid dissolved in DMF and activated with the coupling agent for 2 min before being added to the resin. Peptide couplings were carried out for 30 min. The temporal Fmoc protecting group was removed by treating the resin with a solution of 20% 4-methylpiperidine in DMF for 10 min.

Once the whole peptide sequence was synthesized, 3,5 mg of **Fmoc-P1(alloc)-R** were placed in an Eppendorf tube and resuspended in 200 µL of a cleavage cocktail (95% TFA, 2,5% H_2O and 2,5% triisopropylsilane, TIS) and the mixture was shaken for 1,5 h. Then, the resin was filtered, the TFA filtrate was added over 1 mL of cold ether, and centrifuged. The precipitate was dried with N_2. The solid was redissolved in MeCN:H_2O (1:1) and the crude analyzed by reversed-phase HPLC.

t_R=12,2 min (column *Phenomenex Aeris* 3,6 µm peptide XB-C18 100 Å, lineal gradient 5→95% MeCN, 0,1% TFA/H_2O, 0,1% TFA).

MS (ESI): m/z $[MH]^+$ calculated for $C_{133}H_{239}N_{55}O_{41}$: 3527,9; found: 1764,7 $[M + 2H]^{2+}$, 1176,9 $[M + 3H + 4TFA]^{3+}$, 883,2 $[M + 4H]^{4+}$.

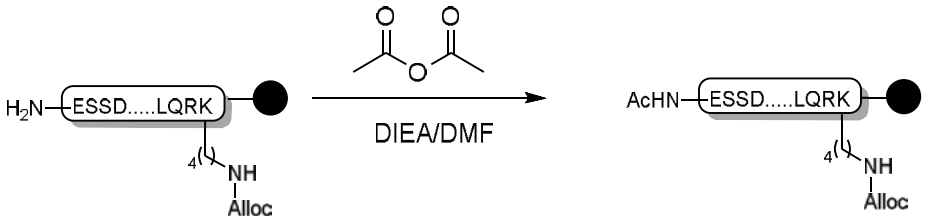
Deprotection of the *N*-terminus



Fmoc-P1(alloc)-R **P1(alloc)-R**

After checking that the peptide synthesis was satisfactory, the *N*-terminus of the peptide was deprotected with 4-methylpiperidine. To do so, 0,05 mmol of **Fmoc-P1(alloc)-R** were treated with a 20% 4-methlypiperidine solution in DMF for 15 minutes and washed with DMF (3 x 5 mL x 2 min).

Acetylation of the *N*-terminus



P1(alloc)-R **Ac-P1(alloc)-R**

0,05 mmol of **P1(alloc)-R** were treated with a mixture of 0,195 M of DIEA/DMF (1,54 mL) and Ac_2O (23,6 µL) for 30 minutes.

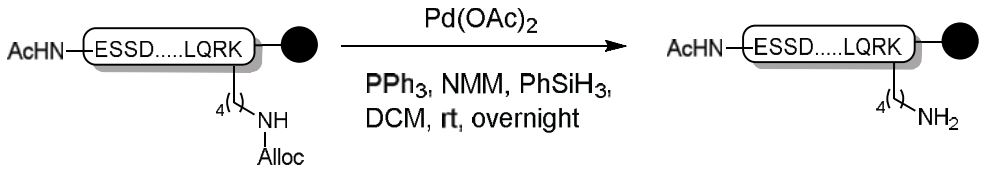
Then, the resin was filtered and washed with DMF (3 x 5 mL x 3 min), DCM (3 x 5 mL x 3 min) and dried.

3,5 mg of **Ac-P1(alloc)-R** were placed in an Eppendorf tube and resuspended in 200 µL of a cleavage cocktail (95% TFA, 2,5% H_2O and 2,5% TIS) and the mixture was shaken for 1,5 h. Then, the resin was filtered and the TFA filtrate was added over 1 mL of cold ether, and centrifuged. The precipitate was dried with N_2. The solid was redissolved in MeCN:H_2O (1:1) and the crude analyzed by reversed-phase HPLC.

t_R=10,9 min (column *Phenomenex Aeris* 3,6 µm peptide XB-C18 100 Å, lineal gradient 5→95% MeCN, 0,1% TFA/H_2O, 0,1% TFA).

MS (ESI): m/z $[MH]^+$ calculated for $C_{137}H_{242}N_{54}O_{44}$: 3347,84; found: 1674,6 $[M + 2H]^{2+}$, 1117,1 $[M + 3H]^{3+}$, 963,4 $[M + 3H + 4TFA]^{3+}$.

Deprotection of the Lys(alloc) side chain



Ac-P1(alloc)-R **Ac-P1-R**

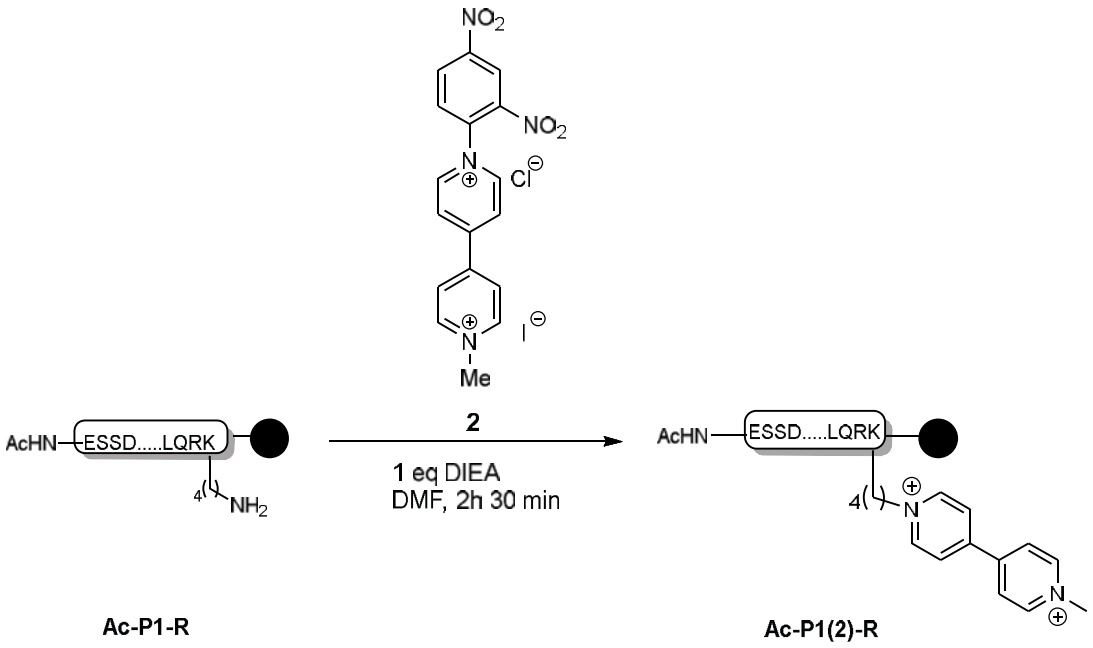
Ac-P1(alloc)-R (0,05 mmol) was treated with a mixture of $Pd(OAc)_2$ (1,7 mg, 0,00758 mmol), PPh_3 (9,8 mg, 0,0374 mmol), *N*-methylmorpholine (NMM, 27,5 µL, 0,25 mmol), $PhSiH_3$ (30,8 µL, 0,25 mmol) in DCM (8 mL) overnight. The day after, the resin was filtered and washed with DCM (3 x 5 mL x 3 min), DMF (3 x 5 mL x 3 min), diethyldithiocarbamate (DEDTC, 25,9 mg in 5 mL of DMF) (3 x 5 mL x 3 min), DMF (3 x 5 mL x 3 min) and DCM (3 x 5 mL x 3 min) and dried.

3,5 mg of **Ac-P1-R** were placed in an Eppendorf tube and resuspended in 200 µL of a cleavage cocktail (95% TFA, 2,5% H_2O and 2,5% TIS) and the mixture was shaken for 1,5 h. Then, the resin was filtered, the TFA filtrate was added over 1 mL of cold ether, and centrifuged. The precipitate was dried with N_2. The solid was redissolved in MeCN:H_2O (1:1) and the crude analyzed by reversed-phase HPLC.

t_R=9,9 min (column *Phenomenex Aeris* 3,6 µm peptide XB-C18 100 Å, lineal gradient 5→95% MeCN, 0,1% TFA/H_2O, 0,1% TFA).

MS (ESI): m/z $[MH]^+$ calculated for $C_{133}H_{238}N_{54}O_{42}$: 3263,81; found: 1633,5 $[M + 2H]^{2+}$, 1088,8 $[M + 3H]^{3+}$, 935,8 $[M + 4H+ 4TFA]^{4+}$.

Compound 2 coupling



Ac-P1-R **Ac-P1(2)-R**

Ac-P1-R (0,05 mmol) was resuspended in 5 mL of DMF and 182 µL of 0,195 M DIEA/DMF and 37,3 mg of compound **2** were added to the mixture. After 2,5 h the resin was washed with DMF (3 x 5 mL x 3 min) and DCM (3 x 5 mL x 3 min) and dried.

3,5 mg of **Ac-P1(2)-R** were placed in an Eppendorf tube and resuspended in 200 µL of a cleavage cocktail (95% TFA, 2,5% H_2O and 2,5% TIS) and the mixture was shaken for 1,5 h. Then, the resin was filtered, the TFA filtrate was added over 1 mL of cold ether, and centrifuged. The precipitate was dried with N_2.

Cleavage from the resin and deprotection of the semipermanent protecting groups

Ac-P1(2)-R (0,05 mmol) were treated with 6 mL of the cleavage cocktail for 3 hours. Then, the resin was filtered and the TFA filtrate was added over 45 mL of ice-cold ether. After 10 min the mixture was centrifuged for 15 minutes, the supernatant was discarded and the precipitate was dried under N_2. The precipitate was redissolved in 900 µL of a 1:1 H_2O:MeCN mixture and the crude was purified by reversed phase HPLC. The collected fractions were lyophilized yielding peptide **Ac-P1(2)** in a 21% yield (18 mg).

t_R=9,6 min (column *Phenomenex Aeris* 3,6 µm peptide XB-C18 100 Å, lineal gradient 5→95% MeCN, 0,1% TFA/H_2O, 0,1% TFA).

MS (ESI): m/z $[M]^{2+}$ calculated for $C_{144}H_{248}N_{56}O_{41}$: 1709,45; found: 1329,96 $[M + H + 5TFA]^{3+}$, 1292,30 $[M + H + 4TFA]^{3+}$, 1253,97 $[M + H + 3TFA]^{3+}$, 1216,30 $[M + H + 2TFA]^{3+}$, 1140,31 $[MH]^{3+}$, 940,98 $[M + 2H + 3TFA]^{4+}$, 912,23 $[M + 2H + 2TFA]^{4+}$, 883,73 $[M + 2H + TFA]^{4+}$, 855,48 $[MH_2]^{4+}$, 684,39 $[MH_3]^{5+}$, 570,66 $[MH_4]^{6+}$.

Conclusions

The following conclusions can be claimed regarding the work carried out in this project:

- The two viologen derivatives **1-(2,4-dinitrophenyl)-4-4'-bipyridinium chloride** (compound **1**) and **1-(2,4-dinitrophenyl)-1'-methylbipyridinium chloride iodide** (compound **2**) were successfully synthesized.
- The desired peptide **Fmoc-P1(alloc)-R** was synthesized following standard solid phase peptide synthesis protocols.
- After acetylation of the *N*-terminus of the peptide and the orthogonal deprotection of the Lys side chain, the coupling of compound **2** to the Lys side chain of **Ac-P1-R** by means of a solid phase Zincke reaction was carried out, and the desired peptide **Ac-P1(2)** was successfully isolated and characterized.

Conclusiones

Después de haber realizado este trabajo, se puede concluir lo siguiente:

- Los dos derivados de los viológenos **cloruro de 1-(2,4-dinitrofenil)-4-4'-bipiridinio** (compuesto **1**) y **cloroyoduro de 1-(2,4-dinitrofenil)-1'-metilbipiridinio** (compuesto **2**) se sintetizaron satisfactoriamente.
- El péptido deseado **Fmoc-P1(alloc)-R** se sintetizó siguiendo los protocolos de síntesis de péptidos en fase sólida.
- Después de la acetilación del extremo *N*-terminal del péptido y de la desprotección ortogonal de la cadena lateral de la lisina, se llevó a cabo, mediante una reacción de Zincke en fase sólida, el acoplamiento del compuesto **2** a la cadena lateral de la lisina de **Ac-P1-R** y el péptido deseado **Ac-P1(2)** fue aislado y caracterizado exitosamente.

Conclusións

Despois de ter realizado este traballo, pódese concluir o seguinte:

- Os dous derivados dos violóxenos **cloruro de 1-(2,4-dinitrofenil)-4-4'-bipiridinio** (composto **1**) e **cloroioduro de 1-(2,4-dinitrofenil)-1'-metilbipiridinio** (composto **2**) sintetizáronse satisfactoriamente.
- O péptido desexado **Fmoc-P1(alloc)-R** sintetizouse seguindo os protocolos de síntese de péptidos en fase sólida.
- Despois da acetilación do extremo *N*-terminal do péptido e da desprotección ortogonal da cadea lateral da lisina, levouse a cabo, seguindo unha reacción de Zincke en fase sólida, o acoplamento do composto **2** á cadea lateral da lisina de **Ac-P1-R**, e o péptido desexado **Ac-P1(2)** foi illado e caracterizado exitosamente.

Annex

Compound 1

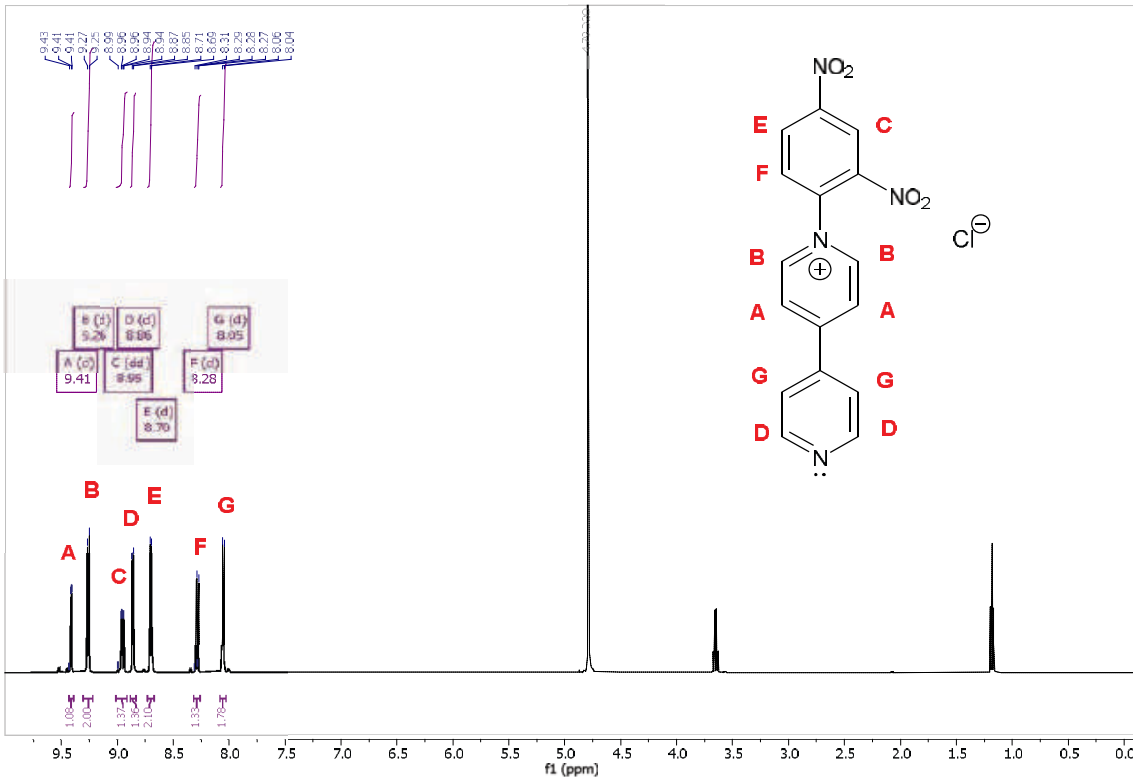


Figure 19. ^{1}H-NMR spectrum (300 MHz, D_2O) for compound 1

Compound 2

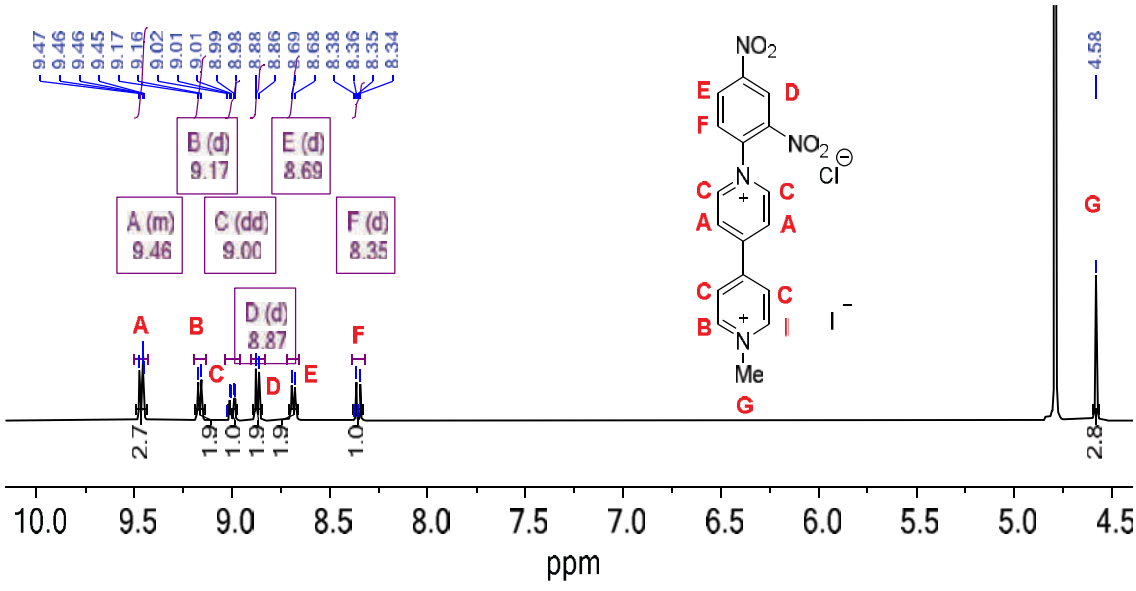


Figure 20. ^{1}H-NMR spectrum (300 MHz, D_2O) for compound 2

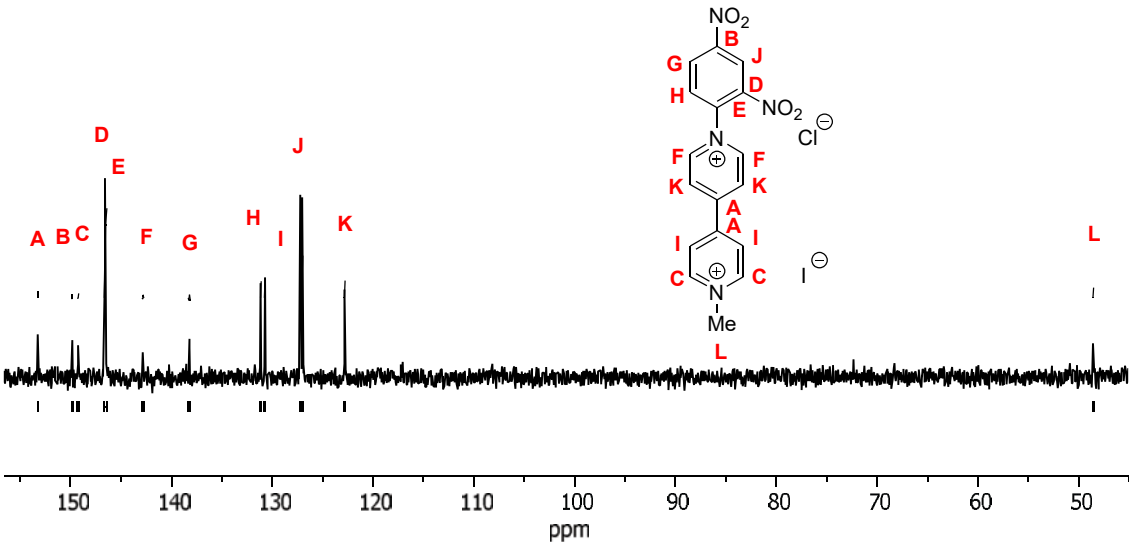


Figure 21. ^{13}C-NMR spectrum (300 MHz, D_2O) for compound 2

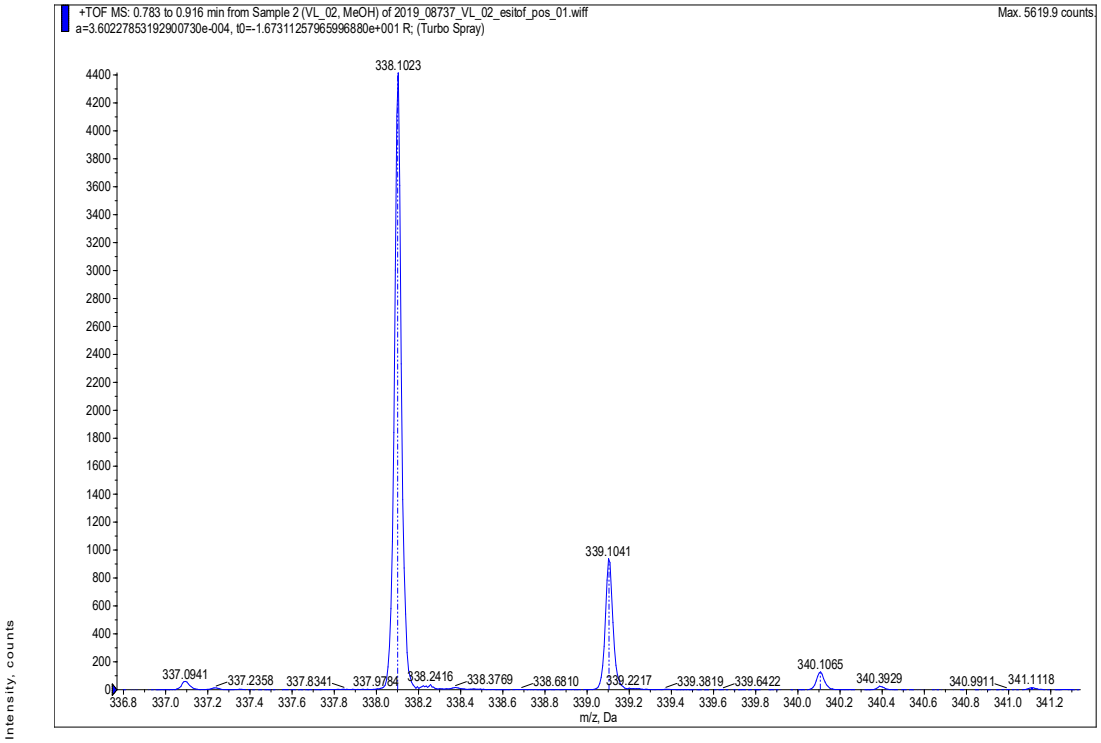


Figure 22. ESI-MS spectrum where the peak with m/z=338,1023 corresponds to compound 2