

TESE DE DOUTORAMENTO

**TÁCTICAS QUÍMICAS PARA  
CONTROLAR A INTERACCIÓN  
DE SISTEMAS SINTÉTICOS CON  
ÁCIDOS NUCLEICOS**

Renata Calo Lapido

ESCOLA DE DOUTORAMENTO INTERNACIONAL

PROGRAMA DE DOUTORAMENTO EN CIENCIA E TECNOLOXÍA QUÍMICA

SANTIAGO DE COMPOSTELA

2019



**DECLARACIÓN DO AUTOR/A DA TESE**  
**TÁCTICAS QUÍMICAS PARA CONTROLAR A INTERACCIÓN DE**  
**SISTEMAS SINTÉTICOS CON ÁCIDOS NUCLEICOS**

Dna.: RENATA CALO LAPIDO.

Presento a miña tese, seguindo o procedemento axeitado ao Regulamento, e declaro que:

- 1) A tese abarca os resultados da elaboración do meu traballo.
- 2) De selo caso, na tese faise referencia ás colaboracións que tivo este traballo.
- 3) A tese é a versión definitiva presentada para a súa defensa e coincide coa versión enviada en formato electrónico.
- 4) Confirmo que a tese non incorre en ningún tipo de plaxio doutros autores nin de traballos presentados por min para a obtención doutros títulos.

*En Santiago de Compostela, a 21 de Outubro de 2019.*

Asdo. Renata Calo Lapidó





**AUTORIZACIÓN DO DIRECTOR / TITOR DA TESE**  
**TÁCTICAS QUÍMICAS PARA CONTROLAR A INTERACCIÓN DE**  
**SISTEMAS SINTÉTICOS CON ÁCIDOS NUCLEICOS**

D. José Luís Mascareñas Cid.

D. Marco Eugenio Vázquez Sentís.

INFORMA/N:

*Que a presente tese, correspóndese co traballo realizado por Dna. **Renata Calo Lapidó**, baixo a miña dirección, e autorizo a súa presentación, considerando que reúne os requisitos esixidos no Regulamento de Estudos de Doutoramento da USC, e que como director desta non incorre nas causas de abstención establecidas na Lei 40/2015.*

*En .Santiago de Compostela, 21 de Outubro de 2019.*

Asdo. José Luis Mascareñas Cid.

Asdo. Marco Eugenio Vázquez Sentís.





*mais soó os iñorantes  
e féridos e duros,  
imbéciles e escuros  
non nos entenden, non.*

[1]

*fillos dos nobres celtas,  
fortes e peregrinos,  
luitade polos destinos  
dos eidos de Breogán.*

(Eduardo Pondal, *Queixumes dos pinos*, 1886)



*Para miña nai, meu pai e Juansiño.*

*Por educarme, instruírme e apoiarme.  
Por compoñer cada pesa dos valores que me forman como persona hoxe en día.  
Por enseñarme a SER.  
FAMILIA.*





# ÍNDICE

<b>1. INTRODUCCIÓN .....</b>	<b>17</b>
1.1. A QUÍMICA BIOLÓXICA.....	19
1.2. O ADN, O FUNDAMENTO DA VIDA.....	21
1.2.1. <i>Estrutura</i> .....	21
1.2.2. <i>Función biolóxica</i> .....	23
1.3. AS PROTEÍNAS, O FUNCIONAMENTO DA VIDA. ....	24
1.3.1. <i>Estrutura</i> .....	24
1.3.2. <i>Síntese Química de Proteínas</i> .....	27
1.3.2.1. Síntese de proteínas en fase sólida.....	27
1.3.2.2. Síntese de Proteínas mediante Ligazón Química Nativa (LQN) .....	29
Síntese de tioésteres C-terminais en resinas: o <i>Nbz</i> .....	31
1.4. OS FACTORES DE TRANSCRICIÓN, A INTERACCIÓN CO ADN .....	33
1.4.1. <i>O Dogma Central da Bioloxía Molecular</i> .....	33
1.4.1.1. A replicación.....	33
1.4.2. <i>A Expresión Xénica</i> .....	34
1.4.2.1. A transcrición .....	34
1.4.2.2. A tradución.....	34
1.4.3. <i>Tipos de Factores de transcrición</i> .....	36
1.4.3.1. Dedos de Zinc.....	36
1.4.3.2. Hélice-xiro-hélice (HTH) e Homeodominio .....	37
1.4.3.3. Cremalleiras de Leucinas (bZip) .....	37
1.4.3.4. Dominios hélice-bucle-hélice (HLHZip) .....	38
1.4.4. <i>c-MYC</i> .....	38
1.4.5. <i>OMOMYC</i> .....	41
1.4.6. <i>Métodos de síntese de c-MYC e OMOMYC</i> .....	43
1.5. RECOÑECEMENTO DE ADN POR MOLÉCULAS PEQUENAS .....	45
1.5.1. <i>Unión covalente: axentes alquilantes</i> .....	45
1.5.2. <i>Unións non covalentes</i> .....	46
1.5.2.1. Intercalantes .....	46
1.5.2.2. Unión ao suco maior .....	47
1.5.2.3. Unión ao suco menor .....	48
Bisbenzamidinas .....	49
<b>2. OBXECTIVOS XERAIS .....</b>	<b>53</b>
<b>3. CAPÍTULO PRIMEIRO: SÍNTESE QUÍMICA DE ANÁLOGOS DE OMOMYC E MYC E RECOÑECEMENTO DO SUCO MAIOR DO ADN. ....</b>	<b>57</b>

3.1. PRECEDENTES DO GRUPO.....	59
3.2. OBXECTIVOS.....	61
3.3. PROPOSTA SINTÉTICA.....	63
3.4. SÍNTESE DO AMINOÁCIDO FMOC-DBZ (ÁCIDO 3-FMOC-4-DIAMINOBENZOICO).....	66
3.5. SÍNTESE DE DERIVADOS DE MYC MEDIANTE ESTRATEXIA 2-FR-LQN.....	67
3.5.1. Síntese dos fragmentos.....	67
3.5.2. Unión dos fragmentos por LQN e obtención de [AQ]MYC.....	74
3.6. SÍNTESE DE DERIVADOS DE OMOMYC.....	78
3.6.1. Síntese dos fragmentos.....	79
3.6.2. Unión dos fragmentos por LQN e desulfuración para obter [AQS]OMOMYC.....	87
3.7. SÍNTESE DE MINIPROTEINAS MYC MARCADAS CON FLUOROFOROS.....	89
3.7.1.1. Síntese de [AQ]MYC-FLU polo método 3 Fr-LQN.....	89
Síntese dos fragmentos peptídicos.....	89
Unión dos fragmentos por LQN e desulfuración para obter [AQK]MYC-Flu.....	92
3.7.1.2. Síntese de [AQ]OMOMYC-FLU polo método 3-Fr-LQN.....	95
Síntese dos fragmentos peptídicos.....	96
Unión dos fragmentos por LQN e desulfuración para obter [AQK]OMOMYC-FLU.....	97
3.7.2. Miniproteínas marcadas con fluoresceína: ESTRATEXIA B.....	98
3.7.2.1. Síntese polo método 3 Fr-LQN de FLU[AQ]OMOMYC.....	98
Síntese dos fragmentos.....	99
Unión dos fragmentos por LQN e desulfuración para obter FLU-[AQ]OMOMYC.....	101
3.7.2.3. Estudo da resistencia de diferentes fluoróforos á reacción de desulfuración.....	103
3.7.2.4. Miniproteínas marcadas con tetrametilrodamina.....	108
3.7.2.4.1. Unión dos fragmentos por LQN e desulfuración para obter TMR-[AQS]OMOMYC.....	110
3.8. FORMACIÓN DE PONTES DISULFURO EN MINIPROTEINAS MARCADOS CON RODAMINA.....	117
3.9. ESTUDOS ESPECTROSCÓPICOS DE INTERACCIÓN CO ADN DAS MINIPROTEINAS SINTÉTICAS.....	120
3.10. ESTUDO DE INTERNALIZACIÓN CELULAR DAS MINIPROTEINAS FLUORESCENTES.....	124
3.10.1. Estudos internalización a diferentes concentracións das miniproteínas marcadas.....	125
3.10.2. Estudos de internalización a diferentes tempos de incubación.....	127
3.10.3. Estudos de internalización en diferentes liñas celulares e comparativa.....	130
3.10.4. Localización.....	133
3.11. CONCLUSIÓNS.....	136

#### **4. CAPÍTULO SEGUNDO: CONTROL DO RECOÑECIMENTO DO SUCO MENOR DO ADN CON MOLÉCULAS PEQUENAS FOTOMODULABLES.....137**

4.1. O GRUPO AZOBENCENO: UN SISTEMA FOTOCONMUTABLE.....	139
4.1.1. As bandas de absorción.....	140
4.1.2. Isomerización de azocompostos para aplicacións en sistemas biolóxicos.....	140



4.1.3. <i>Algunhas aplicacións de azobencenos</i> .....	145
4.2. OBTIVOS.....	146
4.3. CHAVES FOTOCONMUTABLES PARA O RECOÑECEMENTO SELECTIVO DE ADN: O AZOBENCENO.....	148
4.3.1. <i>Síntese: primeira aproximación</i> .....	148
4.3.1.1. <i>Síntese: terceira e cuarta dianas sintéticas</i> .....	155
4.3.2. <i>Comportamento fotoquímico e Isomerizacións</i> .....	157
4.3.3. <i>Tempo de vida media do isómeros</i> .....	162
4.3.4. <i>Recoñecemento de ADN</i> .....	164
4.3.4.1. <i>Dicroísmo circular</i> .....	164
4.3.4.2. <i>Fluorescencia</i> .....	169
4.3.4.3. <i>Cálculo das constantes de disociación por UV</i> .....	171
4.4. CONCLUSIÓN.....	172
<b>5. CAPÍTULO TERCEIRO: RECOÑECEMENTO DO SUCO MENOR DO ADN CON OLIGÓMEROS</b>	
<b>FLUORESCENTES DE MOI ALTA AFINIDADE .....</b>	<b>173</b>
5.1. OBTIVOS.....	175
5.2. SÍNTESE: .....	176
5.2.1. <i>Estudos de oligomerización</i> .....	176
5.2.2. <i>Uso doutros conectores</i> .....	178
5.2.3. <i>Uso doutros monómeros de BBA</i> .....	180
5.2.4. <i>Outras probas de obtención dos oligómeros</i> .....	181
5.3. RECOÑECEMENTO DE ADN .....	185
5.3.1. <i>Dicroísmo circular</i> .....	185
5.3.2. <i>Fluorescencia</i> .....	186
5.3.3. <i>Experimentos de desnaturalización por temperatura: Meltings</i> .....	191
5.4. CONCLUSIÓN.....	194
<b>6. GENERAL CONCLUSIONS .....</b>	<b>197</b>
<b>7. SUMMARY .....</b>	<b>201</b>
<b>8. EXPERIMENTAL SECTION.....</b>	<b>205</b>
8.1. REAGENTS AND MATERIALS: .....	207
8.2. NMR, HPLC, MASS SPECTROMETRY, UV AND CD MEASUREMENTS .....	208
8.2.1. <i>Reverse phase high performance liquid chromatography (RP- HPLC)</i> .....	208
8.2.2. <i>Electrospray Ionization Mass Spectrometry (ESI/MS)</i> .....	209
8.2.3. <i>NMR</i> .....	209
8.2.4. <i>Matrix-assisted laser desorption/ionization mass spectrometry (MALDI/MS)</i> .....	209
8.2.5. <i>Circular Dichroisms (CD)</i> .....	209

8.2.6. Quantification of proteins, small molecules and oligonucleotides using UV.....	210
8.2.7. Photoisomerization studies of azoderivatives .....	213
8.2.8. Fluorescence measurements .....	213
8.2.9. $K_D$ calculations for azoderivatives.....	214
8.2.10. $K_D$ calculations for oligomers.....	215
8.3. CELL UPTAKE EXPERIMENTS.....	216
8.3.1. Incubation concentration studies .....	217
8.3.2. Incubation time studies .....	217
8.4. REACTIONS OF CHAPTER 1.....	218
8.4.1. Protocol for the synthesis of peptides.....	218
8.4.1.1. General protocol for solid phase peptide synthesis (SPPS).....	218
8.4.1.2. Cleavage/ deprotection step.....	218
8.4.1.3. Synthesis of peptides with Nbz at C-terminus .....	218
8.4.1.4. Coupling of TAMRA and Fluorescein at the N-terminus .....	219
8.4.1.5. Selective alloc deprotection. ....	220
8.4.1.6. Coupling of fluorescein at C-terminus.....	220
8.4.2. Synthesis of 3-Fmoc-4-diaminobenzoic acid (Fmoc-Dbz-OH) .....	220
8.4.3. CHARACTERIZATION OF THE FRAGMENTS .....	221
8.4.4. General Protocol for Native Chemical Ligations .....	227
8.4.4.1. Thiazolidine deprotection protocol.....	227
8.4.4.2. TWO-FRAGMENT APPROACH: Synthesis of [AQC <sup>537</sup> ] MYC.....	228
8.4.4.3. Three fragment approach: synthesis of [AQS]OMOMYC(SH).....	230
8.4.4.4. THREE-FRAGMENT APPROACH. Synthesis of TMR-[AQ]MYC(SH) .....	233
8.4.4.5. three fragment approach. synthesis of TMR-[AQS]OMOMYC(SH) .....	236
8.4.5. Desulfurization of the resulting miniproteins .....	238
8.4.5.1. Desulfurization of [AQC]MYC obtained by the 2-fragment approach.....	238
8.4.5.2. Desulfurization of [AQS]OMOMYC(SH) obtained by the 3-fragment approach.....	239
8.4.5.3. Desulfurization of TMR-[AQ]MYC(SH) obtained by the 3-fragment approach.....	239
8.4.5.4. Desulfurization of TMR-[AQS]OMOMYC(SH) obtained by the 3-fragment approach .....	240
8.5. REACTIONS OF CHAPTER 2 .....	241

## ABREVIATURAS E ACRÓNIMOS

$\delta$	Desplazamiento químico
$\mu\text{L}$	Microlitro
$\mu\text{M}$	Micromolar
$^{\circ}\text{C}$	Grao centígrado
A	Adenina
Ar	Argon
AcOH	Ácido acético
AcOEt	Acetato de etilo
ADN	Ácido desoxirribonucleico
ARN	Ácido ribonucleico
Boc	<i>tert</i> -Butiloxicarbonilo
C	Citosina
d	Dobrete
Dbz	Ácido 4-(methylamino)benzoico
DCM	Diclorometano
DMF	<i>N,N'</i> -Dimetilformamida
DMP	Periodinano de Dess Martin
DMSO	Dimetilsulfóxido
DMSO- $d_6$	Dimetilsulfóxido deuterado
Et	Etilo
EtOH	Etanol
Et <sub>2</sub> O	Eter etílico
ESI	Espectrometría de masas de electroesprai
G	Guanina
h	hora
HPLC-RP/EM	Espectroscopía de masas e cromatografía líquida de alta resolución en fase reversa

HATU	hexafluorofosfato de 2-(1H-7-azabenzotriazol-1-il)-1,1,3,3-tetrametiluronio
m	Multiplete
min	minutos
M	Molar
MeOH	Metanol
MeCN	Acetonitrilo
mL	Mililitro
MPAA	4-(carboximetil) tiofenol
EM	Espectrometría de masas
N	Nitróxeno
NBS	<i>N</i> -Bromosuccinimida
NMM	<i>N</i> -metilmorfolina
nm	Nanómetro
RMN	Resonancia magnética nuclear
s	Singlete
t	Triplete
T	Timina
t.a	Temperatura ambiente
TFA	ácido trifluoroacético
THF	Tetrahidrofurano
TMR	tetrametilrodamina
t <sub>R</sub>	Tempo de retención
UV	Ultravioleta
VIS	Visible



## **1. INTRODUCCIÓN**



## 1.1. A QUÍMICA BIOLÓXICA

Non é doado establecer de maneira precisa o significado do termo *Química Biolóxica*, xa que estas dúas palabras abarcan milleiros de conceptos. Podería dicirse que a Química Biolóxica é unha nova rama multidisciplinar da ciencia na que se aplican técnicas e coñecementos químicos para o estudo de sistemas biolóxicos, e que adapta o coñecemento das ciencias da vida para progresar na Química

A Química crea as ferramentas que permiten entender a bioloxía a nivel molecular. Exemplos destas ferramentas son a creación de pequenas moléculas capaces de interactuar cas biomoléculas que constitúen os seres vivos (ADN, proteínas, ARNj) ; os biomarcadores para detectar e explicar o funcionamento de estas biomoléculas, ou os medicamentos, para tratar determinadas enfermidades, combater a resistencia a antibióticos, deseñar novos antivirais etc. A finalidade última da Química Biolóxica consiste en entender os problemas biolóxicos e propoñer novas solucións no ámbito biomédico. Para isto, esta disciplina abrangue o deseño de modelos e sondas sintéticas, a introdución, liberación e estudo do seu comportamento a nivel celular, o control sobre a súa interacción con biomoléculas naturais ata o punto de poder manipular a expresión xénica, o ciclo celular, ou outros procesos.

Aínda que as súas orixes poden considerarse no século XIX, a Química Biolóxica asentou os seus piares e acadou un progreso exponencial no século XX da man dos avances tecnolóxicos en técnica sintéticas e analíticas e as sinerxías co progreso na bioloxía molecular e celular. Poderíamos nomear como dous dos pioneiros deste campo a Joseph Priestley e Humprey Davy que, a finais do 1700 e comezos do 1800, comprobaban os efectos terapéuticos de gases químicos que eles illaban en organismos vivos: aplicaban a química para mellorar no eido da bioloxía. O primeiro usaba animais vivos, neste caso ratos, e testaba a súa supervivencia despois de expoñelos a diversos gases que illaba, como o óxido de nitróxeno. O segundo pola contra, experimentaba en si mesmo os efectos secundarios de inhalar ditos gases.<sup>1</sup>

Sen dúbida, foi uns anos máis adiante, en 1828, cando Friedrich Whöler marcou un antes e un despois na fronteira entre a química e a bioloxía. Por casualidade, este químico alemán deu

---

<sup>1</sup> a) Priestly, J. Experiments and Observations on Different kinds if Air (Johnson Press, London, 1774). b) K. L. Morrison, G. A. Weiss, *Nat. Chem. Biol.* **2006**, 2, 3.

sintetizado urea a partir só de materiais inorgánicos,<sup>2</sup> demostrándose así que non se precisaba de organismos vivos para sintetizar compostos biolóxicos.

En 1856, o inglés Perkin sintetizou o primeiro colorante químico, a *mauveína*.<sup>3</sup> Un derivado da anilina cunha rechamante cor malva (de aquí o nome) que abriu paso ao uso dos colorantes sintéticos na industria téxtil amais de iniciar o uso de ditos derivados da anilina para a tinguidura de células específicas. Grazas a isto, e baseándose en que unhas células eran capaces de tingirse de maneira diferente que outras co mesmo colorante, podemos hoxe en día denominar ao que foi premio Nobel Paul Ehrlich como o pai da quimioterapia, pois deu sintetizado un composto, o *salvarsán*,<sup>4</sup> capaz de actuar selectivamente sobre os patóxenos causantes da sífilis e que non afectaba ao resto do organismo. Ponse así de manifesto o impacto e importancia que ten o controlar procesos celulares con pequenas moléculas sintetizadas no laboratorio para o tratamento de enfermidades.

Todo isto, sumado á elucidación da estrutura do ADN, o descifrado do código xenético e o deseño da metodoloxía para a síntese química de péptidos e proteínas, xa na segunda metade do século XX contribuíron a que **química biolóxica** se reafirmase como unha disciplina científica.

---

<sup>2</sup> F. Wöhler, *Ann. Chim. Phys.* **1828**, 37, 330.

<sup>3</sup> W. H. Perkin, *J. Chem. Soc., Trans.* **1896**, 69, 596.

<sup>4</sup> P. Ehrlich, A. Bertheim, *Berichte.* **1912**, 45, 756.



## 1.2. O ADN, O FUNDAMENTO DA VIDA.

### 1.2.1. Estrutura

O ADN, ou ácido desoxirribonucléico, é unha macromolécula formada por nucleótidos unidos a través do grupo hidroxilo 5' do fosfato dunha pentosa co grupo hidroxilo 3' da pentosa seguinte mediante enlaces fosfodiéster.

Un nucleótido está formado por un monosacárido (a  $\beta$ -D-2'-desoxirribosa), unha base nitrogenada e un grupo fosfato.<sup>5</sup> As bases están unidas ao carbono 1' do azucre mediante enlace *N*- $\beta$ -glicosídico. Hai catro bases diferentes, dúas de tipo púrica que son a adenina(A) e a guanina (G) e dúas de tipo pirimidínica: a timina (T) e a citosina (C). Tódolos nucleótidos que forman este polímero teñen a mesma desoxirribosa, polo que a diferenza entre diferentes ADN radica tan só na orde das bases nitrogenadas unidas a estas desoxirribosas. As propiedades das bases e a disposición destas inflúen no tipo de estrutura secundaria do ácido nucleico e son as que permiten xerar unha gran variabilidade. Os grupos fosfato están desprotonados a pH fisiolóxico polo que o ADN ten carácter ácido.<sup>6</sup>

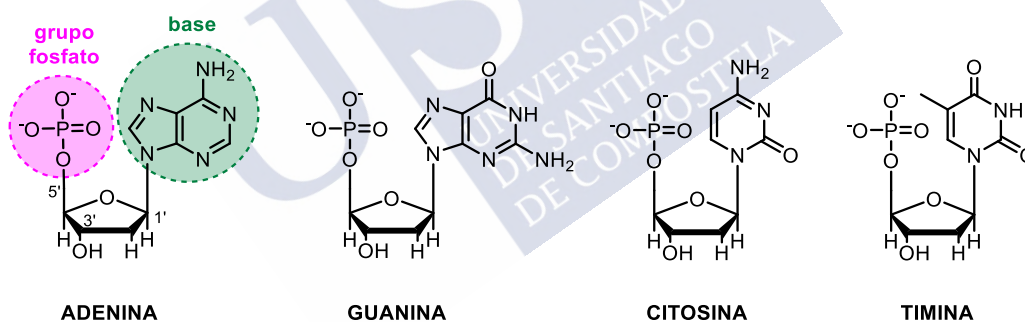


Figura 1. Representación dos catro nucleótidos que forman o ADN. Debaixo escrito o nome da base que o forma. Na primeira molécula indícase a numeración estándar dos átomos de carbono da aldopentosa.

Baseados nos experimentos de difracción de raios X de Rosalind Franklin e na *lei de equivalencia de bases* de Erwing Chargaff, Francis Crick e James Watson postulaban nun artigo en 1953 un modelo para a estrutura tridimensional do ADN.<sup>7</sup> Anos despois, estes dous científicos mais Maurice Wilkins gañaban o Premio Nobel de Medicina grazas a este artigo. Iso si, a pesares de que foi Rosalind Franklin quen realizou a difracción de Raios X da dobre hélice na que se fundamentaron o Watson e Crick, nunca obtivo recoñecemento por parte dos seus

<sup>5</sup>D. L. Nelson, M. M. Cox, *Lehninger. Principios de Bioquímica*. Quinta edición, 2009, Editorial Omega.

<sup>6</sup> R. Lavery, C. Zadecki, J. Westbrook, *Oxford Handbook of Nucleic Acid Structure*, Oxford 1999.

<sup>7</sup> J. D. Watson, F. H. C. Crick, *Nature* 1953, 171, 737.

colegas. É mais, crese que a investigadora nunca chegou a saber que os biólogos viran a súa imaxe de raios X antes de faceren o postulado da dobre hélice.<sup>8</sup>

Grazas a estes traballos e posteriores estudos, quedou establecido que o ADN está constituído por dúas cadeas antiparalelas de desoxirribonucleótidos que se asocian, unha coa outra, a través de pontes de hidróxeno entre pares de bases complementarias, formando unha dobre hélice dextróxima. Concretamente, a adenina forma dúas pontes de hidróxeno coa timina e a guanina forma tres coa citosina. As bases tentan minimizar a superficie en contacto coa auga e dispóñense no interior do dúplex mediante interaccións de amoreamento  $\pi$ . O esqueleto de pentosa-fosfatos quedan cara o exterior e asóciase con catións para neutralizar parcialmente a carga e estabilizar a estrutura.<sup>9,5</sup>

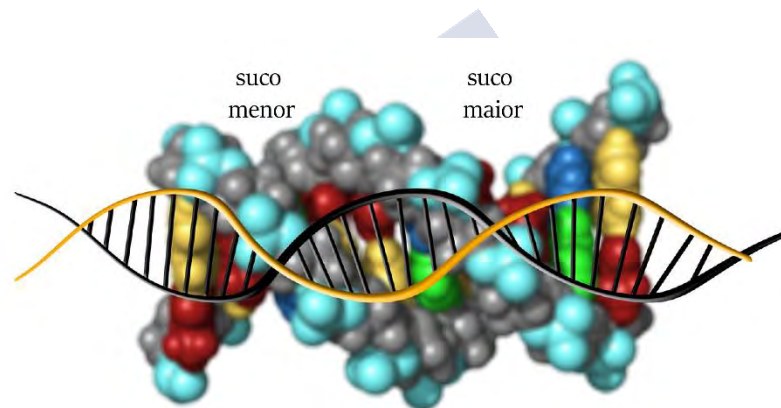


Figura 2. Representación da estrutura dunha cadea de ADN con conformación B. Amósanse tanto o suco maior (cunha anchura de 22 Å) e o suco menor (de 12 Å de ancho).

A pH fisiolóxico a forma máis estable na que se atopa unha molécula de ADN denomínase *Forma B*. Esta estrutura está caracterizada por ter unha anchura de hélice de 2 nm e conter 10,5 pares de bases e unha alzada de 3,6 nm por volta completa. A disposición entre as cadeas e a xeometría dos azucres forman dous sucros de tamaños desiguais e con atributos xeométricos diferentes, chamados suco maior e suco menor. O suco maior é ancho e pouco profundo (7,5 Å), mentres que o suco menor é máis estreito pero máis profundo (8,5 Å).

<sup>8</sup> J. D. Watson, *The double helix*, 1968, Editorial Weindenfeld & Nicolson.

<sup>9</sup> V. A. Bloomfield, D. M. Crothers, I. Tinoco, *Nucleic Acids; structures, properties and functions*, 2000, University Science Books, Sausalito.

### 1.2.2. Función biolóxica

O ADN é, sen dúbida, a biomolécula orgánica máis importante, xa que garda a información codificada a partir da cal se forma un organismo vivo. As funcións do ADN son o almacenamento e a transmisión da información xenética.

A secuencia liñar de nucleótidos conforma a estrutura primaria do ADN. A través dos procesos descritos no *Dogma Central da Bioloxía Molecular*,<sup>10</sup> que se explicará máis adiante, esta secuencia de nucleótidos transfórmase en ARN e, finalmente nunha proteína funcional.

A continuidade na existencia dunha especie esixe que a súa información xenética se manteña estable e íntegra. O ADN é o responsable de manter esta integridade intacta e de transmitir todo o material xenético dunha xeración de células á seguinte. Esta información xenética contén instrucións para formar tódolos compoñentes celulares, ademais de actuar como molde para producir moléculas idénticas de ADN que serán distribuídas á proxenie tras dividirse a célula. É por iso que o ADN é a única biomolécula que dispón de mecanismos de reparación propios

As cadeas de ADN son enormes, polo que este adoita atoparse superenrolado. De toda a cantidade de ADN que teñen os organismos, tan só unha pequena porcentaxe se chega a traducir en polipéptidos ou ARN funcionais. Os pequenos fragmentos que codifican estas secuencias primarias denomínanse xenes, e o conxunto de todos eles dentro dun individuo é o seu xenoma.<sup>11</sup>Cada xene ten encriptado a secuencia para formar unha única proteína.

---

<sup>10</sup> F.H.C Crick, *Nature* **1970**, 227, 561.

<sup>11</sup> H. Pearson, *Nature* **2006**, 441, 398.

### 1.3. AS PROTEÍNAS, O FUNCIONAMENTO DA VIDA.

#### 1.3.1. Estrutura

As proteínas son as biomoléculas máis abundantes nos seres vivos. Están formadas pola unión covalente de varios aminoácidos mediante un enlace que se chama comunmente *enlace peptídico*. Un aminoácido está constituído, alomenos, por un grupo amino e un grupo ácido (carboxilo) unidos a unha cadea carbonada. Segundo a posición do grupo amino, estas moléculas poden clasificarse en  $\alpha$ ,  $\beta$ ,  $\gamma$ ...-aminoácidos.

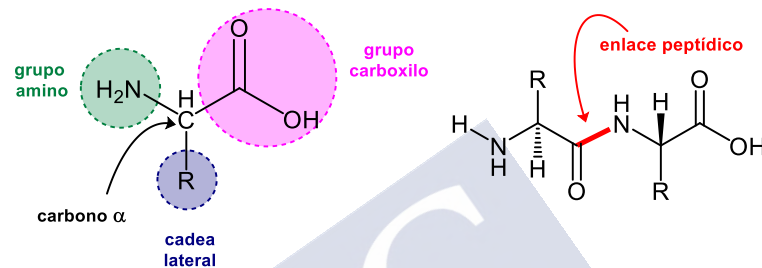


Figura 3. Esquerda: Estrutura xeral dun aminoácido. Dereita: Representación dun enlace peptídico entre dous L-aminoácidos.

Os que compoñen as proteínas son os  $\alpha$ -aminoácidos e hai vinte diferentes amplamente estendidos. Para facilitar a escritura das proteínas, a estes vinte aminoácidos estándar asígnóuselles unha abreviatura de tres letras ou símbolos dunha soa letra.<sup>12</sup>

O que os diferencia entre si é o outro grupo unido ao carbono  $\alpha$ , que se denomina *cadea lateral*, que lles confire as propiedades individuais a cada un dos aminoácidos. Segundo a natureza do grupo R nas cadeas laterais, os aminoácidos proteicos clasifícanse en tres grupos: básicos (posúen grupos amino na cadea lateral. Son os Arg, Lys, His), ácidos (teñen carboxilos na cadea lateral: Asp e Glu) e neutros (a carga neta é nula a pH neutro porque non teñen grupos ionizables nas cadeas laterais. Estes son: Gly, Ser, Thr, Tyr, Cys, Gln, Asn).

Debido a que o carbono  $\alpha$  é quiral, estes monómeros presentan dous enantiómeros. A pesares diso, as proteínas só se compoñen de residuos na súa forma estereoisomérica L.

As proteínas sintetízanse a través do acoplamento dun aminoácido cuxo extremo carboxílico reacciona co extremo amino doutro formando un enlace de tipo amida (tamén coñecido como enlace peptídico). Este enlace ten dúas formas resoantes que lle confiren

<sup>12</sup> M. Saffran, *Biochem. Educ.* **1998**, 26, 116.

estabilidade e carácter parcial de dobre enlace. Debido a isto, o enlace peptídico non pode rotar e os átomos de C, O, N colócanse no mesmo plano, quedando os grupos laterais ‘colgando’ do polipéptido.<sup>13,5</sup>

Segundo a cantidade de residuos que se unan para formar a cadea, os polímeros obtidos reciben o nome de oligopéptidos (poucas unidades de aminoácidos) ata péptidos ou proteínas (milleiros de residuos). Por mor da lonxitude destas macromoléculas e a diversidade de funcións que desenvolven dentro do organismo, as proteínas organízanse en catro niveis estruturais diferentes:

Estrutura primaria: é a secuencia de aminoácidos dun péptido (ou proteína). É dicir, o número, tipo e orde dos residuos na cadea. Sempre hai un extremo co grupo amino libre que se denomina extremo *N*-terminal e outro co grupo carboxilo libre que se coñece como *C*-terminal. Os péptidos numéranse a partir do extremo *N*-terminal, e sempre se escribe a súa secuencia a partir de ese extremo.

Estruturas secundaria: é a disposición espacial de determinados fragmentos do esqueleto polipeptídico que se repiten ao longo da proteína. En gran parte é consecuencia das pontes de hidróxeno entre os enlaces C=O e N-H e no que non interveñen as cadeas laterais. Dentro desta pregamento destacan dous tipos básicos: a hélice  $\alpha$ <sup>14</sup> e a lámina  $\beta$ .<sup>15</sup> A primeira consiste nun enrolamento en espiral da cadea sobre si mesma ao longo dun eixe lonxitudinal imaxinario no sentido de xiro das aguias dun reloxo. Cada volta contén 3.6 aminoácidos. As pontes de hidróxeno fórmanse entre un aminoácido e o cuarto que lle segue na cadea. As cadeas laterais dispóñense cara o exterior da hélice. Na lámina  $\beta$ , dous fragmentos peptídicos acóplanse mediante pontes de hidróxeno transversais. As cadeas adxacentes poden dispoñerse no mesmo sentido (paralelas) ou en sentidos opostos (láminas antiparalelas). A forma resultante semella unha folla pregada en *zigzag*. Os grupos R sobresaen da estrutura dispoñéndose alternativamente por riba ou por debaixo, en sentidos opostos, do plano da folla.

Estrutura terciaria: É a organización que resulta de tódolos pregamentos secundarios ao longo de toda a cadea, e a disposición espacial final que esta adopta para dar lugar á estrutura

---

<sup>13</sup> A) L. Pauling, R. B. Corey, H. R. Branson, *Proc. Nat. Acad. Sci.* **1951**, 37, 205. B) L. Stryer, J. M. Berg, J. L. Tymoczko, *Bioquímica con aplicacións clínicas*, **2012**, Editorial Reverté.

<sup>14</sup> D. Eisenberg, *Proc. Nat. Acad. Sci.* **2003**, 100, 11207.

<sup>15</sup> L. Pauling, R. B. Corey, H. R. Branson, *Proc. Nat. Acad. Sci.* **1951**, 37, 251.

pregada. Débese ás interaccións que se dan entre os grupos R dos residuos. É a esencial para levar a cabo a súa función biolóxica.

Estrutura cuaternaria: Defínese como a disposición relativa que toman diferentes subunidades proteicas no caso de que a proteína estea formada por varias cadeas polipeptídicas.

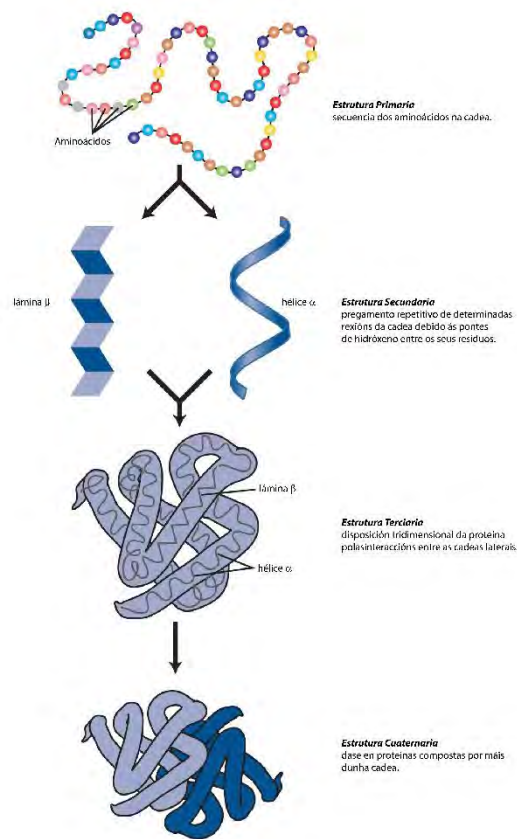


Figura 4. Niveis de organización das proteínas. Imaxe modificada de National Human Genome Research Institute [16]

As proteínas son as encargadas de levar a cabo practicamente tódalas funcións precisas para a vida. Hai milleiros de estruturas de proteínas, cada unha ca súa función asociada. O tipo de función que desempeñan dentro do organismo vén dada pola súa secuencia de aminoácidos. Son as responsable da organización da célula, da súa forma, do sostén da membrana lipídica, dos produtos que se sintetizan dentro da mesma e dos residuos que se desbotan. Ademais, actúan de conectores entre o medio extra e intra-celular captando sinais do medio e transferíndoas ao interior da célula. Outras interveñen nos movementos, nos procesos

<sup>16</sup> National Human Genome Research Institute, 2019, <https://www.genome.gov/>



homeostáticos para manter o pH, transportan outras moléculas, hormonas etc. Algunhas teñen función estrutural e forman parte da membrana lipídica, outras función encimática, e dirixen as reaccións químicas que se producen nos seres vivos e incluso se encargan de regular a expresión xénica.

### 1.3.2. Síntese Química de Proteínas

Debido á gran variedade de funcións que desempeñan, as proteínas son as biomoléculas máis interesantes á hora de analizar, comprender e poder modificar as súas funcións celulares. Cada ano incrementa en centos, ou miles, o número de proteínas identificadas na especie humana, estando preto dunhas 17.000.<sup>17</sup> Cada ano crece, tamén, o número de proteínas que son usadas como fármacos tanto no tratamento como na detección de moitas enfermidades grazas á súa especificidade e potencia.<sup>18</sup>

Debido a esta importancia é fundamental dispoñer de métodos para a preparación de proteínas. A síntese de proteínas non é nada doada debido ao tamaño e complexidade de grupos funcionais que presentan. A maioría das veces, o acceso a estas proteínas lévase a cabo mediante expresión biolóxica en bacterias, pero esta presenta limitacións: as purificacións adoitan ser tediosas e os rendementos baixos, a produción adoita ser lenta,<sup>19</sup> ademais, a introdución de modificacións na cadea peptídica (por exemplo con fluoróforos, aminoácidos non naturais, modificacións postraducionais etc)<sup>20</sup> resulta moi difícil.<sup>21</sup>

Grazas aos avances na síntese de péptidos durante os últimos dous séculos, a síntese química tense convertido nunha alternativa moi interesante para preparar proteínas naturais, non naturais e modificadas.<sup>22</sup>

#### 1.3.2.1. Síntese de proteínas en fase sólida

En 1963, Merrifield deu a coñecer un método que revolucionaría a síntese de péptidos e proteínas e que se utiliza aínda hoxe en día, a denominada *Síntese de Péptidos en Fase Sólida*

<sup>17</sup> Y.-K. Paik, C. M. Overall, E. W. Deutsch, W. S. Hancock, G. S. Omen, *J. Proteome Res.* **2016**, *15*, 3945.

<sup>18</sup> S. S. Usmani, G. Bedi, J. S. Samuel, S. Singh, S. Kalra, *PLoS ONE* **2017**, *12*, e0181748.

<sup>19</sup> A) M.-E. Beaulieu, F.-O. McDuff, V. Frappier, M. Montagne, J.-F. Naud and P. Lavigne, *J. Mol. Recognit.* **2012**, *25*, 414. B) M. E. Beaulieu, F. O. McDuff, M. Bédard, M. Montagne, P. Lavigne, *Methods Mol. Biol.* **2013**, *1012*, 7.

<sup>20</sup> C. D. Spicer, B. G. Davis, *Nat. Commun.* **2014**, *5*, 4740.

<sup>21</sup> S. B. H Kent, Y. Sohma, S. Liu, D. Bang, B. Pentelute, K. Mandal, *J. Pept. Sci.* **2012**, *18*, 428.

<sup>22</sup> A) B. L. Nilsson, M. B. Soellner, R. T. Raines, *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 91. B) V. Agouridas, O. El Mahdi, V. Diemer, M. Cargoët, J.-C. M. Monbaliu, O. Melnyk, *Chem. Rev.* **2019**, *119*, 7328. C) L. Liu, *Isr. J. Chem.* **2019**, *59*, 64. D) W. Hou, X. Zhang, C.-F. Liu, *Trans. Tianjin Univ.* **2017**, *23*, 401. D) R. Behrendt, P. White, J. Offer, *J. Pept. Sci.* **2016**, *22*, 4.

(SPFS).<sup>23</sup> A metodoloxía baséase en unir covalentemente un aminoácido  $N_\alpha$ -protexido a un soporte sólido polo extremo carboxílico. O soporte, denominado resina, debe ser insoluble nos disolventes que se usan nas posteriores reaccións. De seguido, desprotéxese selectivamente o grupo protector do grupo amino e ancorase o seguinte aminoácido, tamén  $N_\alpha$ -protexido da mesma maneira que o anterior. E así sucesivamente. Deste xeito o polipéptido crece dende o extremo  $C$ -terminal cara o  $N$ -terminal. Entre acoplamento e acoplamento lávase a resina, eliminándose facilmente subprodutos que non estean ancorados. As cadeas laterais están protexidas con grupos protectores chamados *semipermanentes* que resistan as condicións empregadas durante toda a síntese. O último paso é desancorar o péptido da resina e desprotexer estes grupos semipermanentes das cadeas laterais ao mesmo tempo.

Esta técnica permite a introdución de aminoácidos non naturais e modificacións postraducionais dunha maneira versátil e doada de regular. Actualmente, o método máis estendido é o denominado síntese con Fmoc, xa que os seus reactivos son doados de adquirir e a un prezo barato. Denomínase así porque os aminoácidos teñen o grupo  $\alpha$ -amino protexido con Fmoc. Outra estratexia moi usada é a Boc, onde os  $\alpha$ -amino están protexidos co grupo protector que lle da o nome. As dúas técnicas están meticulosamente desenvolvidas e asentadas. Sen dúbida, a máis usada é a Fmoc. Grazas a que non precisa do uso de TFA (corrosivo) nas etapas intermedias de desprotección, incluso adoita realizarse en aparellos especialmente deseñados para tal fin, denominados comunmente *sintetizadores de péptidos* e existe a posibilidade de utilizar microondas para mellorar e axilizar a síntese.

---

<sup>23</sup> R. B. Merrifield, *J. Am. Chem. Soc.*, **1963**, 85, 2149.



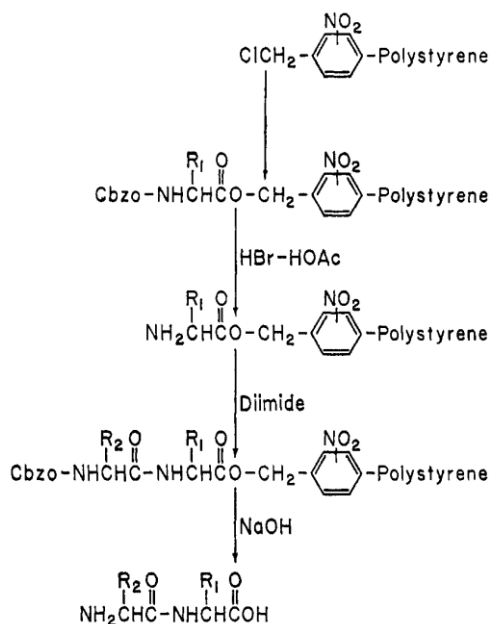


Figura 5. Esquema da SPFS que usou Merrifield, adaptada da ref [23].

A pesares de ser unha técnica excelente, rápida e moi eficiente, a eficiencia da síntese química en fase sólida está limitada a péptidos relativamente pequenos, duns 50-60 aminoácidos. As limitacións máis importantes derivan da agregación dos fragmentos peptídicos protexidos e o aumento dos subprodutos xerados a medida que aumenta a cadea sintetizada.

### 1.3.2.2. Síntese de Proteínas mediante Ligazón Química Nativa (LQN)

Para resolver os problemas da SPFS e acadar a síntese de péptidos máis longos desenvolvéronse métodos que conseguían unir dous segmentos peptídicos, sintetizados independentemente mediante SPFS, a través dun enlace tipo amida<sup>24,22b)</sup> Estes métodos combinan a SPFS coa modificación de péptidos en disolución.<sup>25</sup> O máis efectivo é o denominado *Ligazón Química Nativa* (LQN, en inglés *Native Chemical Ligation*).<sup>26</sup> O grupo de Kent publicaba en 1994 un artigo no que aplicaba este método dunha forma práctica á síntese de péptidos con enlaces de tipo amida.<sup>27</sup> A LQN une mediante enlace de tipo amida dous fragmentos peptídicos totalmente desprotexidos, en disolución acuosa a pH neutro. A reacción ten lugar entre un tioéster C-terminal dun fragmento e unha cistina en posición N-terminal

<sup>24</sup>A) B. L. Bray, *Nat. Rev. Drug Discov.* **2003**, 2, 587. B) T. W. Muir, D. Sondhi, P. Cole, *Proc. Natl Acad. Sci.* **1998**, 95, 6705. C) J. Bode, R. M. Fox, K. D. Baucom, *Angew. Chem. Int. Ed.* **2006**, 45, 1248. D) T. W. Muir, P. E. Dawson, S. B. H. Kent, *Methods Enzymol.* **1997**, 289, 266.

<sup>25</sup> P. A. Cistrone, M. J. Bird, D.T. Flood, A. P. Silvestri, J. C. Hintzen, D. A. Thompson, P.E. Dawson, *Curr. Protoc. Chem. Biol.* **2019**, 11, e61.

<sup>26</sup> S. B. H. Kent, *Bioorg. Med. Chem.*, **2017**, 25, 4926. B) S. B. H. Kent, *Chem. Soc. Rev.* **2009**, 38, 338.

<sup>27</sup> P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science*, **1994**, 266, 776.

doutro fragmento. En primeiro lugar, o sulfhidriilo libre da cistina *N*-terminal realiza un ataque nucleofílico sobre o tioéster, producíndose unha *trans*-tioesterificación. A continuación, o tioéster xerado sofre un reordenamento de acilo S→N onde se forma o enlace amida nativo. O primeiro paso pode ser reversible pero o segundo é irreversible, desprazando a reacción cara os produtos de ligazón.

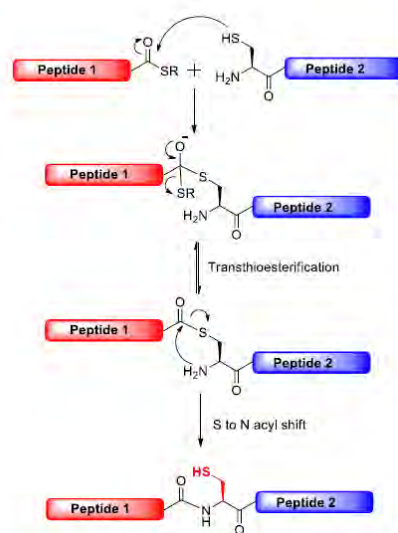


Figura 6. Esquema do mecanismo da LQN.

Unha das vantaxes desta reacción é a alta quimioselectividade que presenta, que se realiza en condicións moi suaves, a racemización normalmente é menor ao 1 %, <sup>28</sup> e os produtos pódense obter en bos rendementos. Por outro lado, a LQN tamén presenta algúns inconvenientes: Se se ten en conta que as proteínas naturais tan só presentan un 1,8 % de cistina, a necesidade dun residuo cistina no extremo *N*-terminal dun dos fragmentos dificulta a análise retrosintética e a elección do sitio de desconexión. Unha das solucións a este problema veu da man do equipo de Dawson, <sup>29</sup> o que propuxo realizar unha desulfuración catalítica final tras a LQN. Deste xeito a cistina pasa a ser alanina, moito máis abundante nas proteínas (8,9 %). Esta proposta e a súa mellora sen necesidade de uso de metais e a través dunha reacción radicalaria selectiva, <sup>30</sup> ampliaron o rango de aplicación da reacción de ligazón. <sup>31</sup>

<sup>28</sup> W. Lu, M. A. Qasim, S. B. H. Kent, *J. Am. Chem. Soc.* **1996**, *118*, 8518.

<sup>29</sup> L. Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.*, **2001**, *123*, 526.

<sup>30</sup> A) Q. Wan, S. J. Danishefsky, *Angew. Chem. Int. Ed.* **2007**, *46*, 9248. B) K. Jin, T. Li, H. Y. Chow, H. Liu, X. P. B. Li, *Angew. Chem. Int. Ed.* **2017**, *56*, 14607.

<sup>31</sup> L. R. Malins, R. J. Payne, *Curr. Opin. Chem. Biol.* **2014**, *22*, 70.

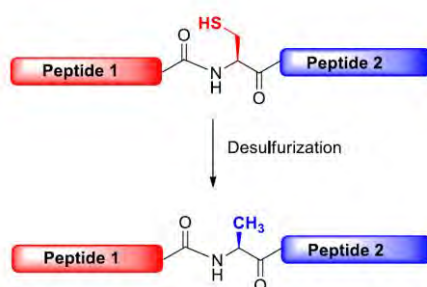


Figura 7. Estratexia de desulfuración.

### Síntese de tioésteres C-terminais en resinas: o *Nbz*

Para que a LQN poida ter lugar, precísase un tioéster no extremo C-terminal do fragmento N-terminal. Os fragmentos peptídicos sintetízanse mediante SPFS. Se a síntese se leva a cabo pola estratexia Fmoc (a máis estendida e automatizada a día de hoxe), introducir un tioéster C-terminal resulta moi complicado posto que son lábiles no medio básico que se usa neste método. Para acadar un tioéster terminal, desenvolvéronse múltiples estratexias como a uso de cócteles de desprotección diferente para eliminar o Fmoc,<sup>32</sup> a *trans*-tioesterificación de fragmentos completamente protexidos,<sup>33</sup> o uso dun linker especial para ancorar o péptido polo  $\alpha$ -amino,<sup>34</sup> o uso de N-metilsulfonamidas,<sup>35</sup> o enmascaramento dun precursor de tiol mediante hidracidas,<sup>36</sup> ou a utilización de *orto*-aminoanilidas<sup>37</sup> etc.

Unha das propostas máis estendidas é a creación dun precursor de tioésteres no extremo C-terminal dunha cadea peptídica en forma de *N*-acilbencimidazolinona ou máis coñecido como *Nbz*.<sup>38</sup> Neste método, colócase un aminoácido derivado do ácido benzoico como conector entre a resina e o péptido, para logo medrar o resto da cadea peptídica de xeito habitual. Por último, o diaminobenzoico (*Dbz*) transfórmase en *N*-acilbencimidazolinona (*Nbz*) mediante a acilación da amina libre da *orto* aminoanilida por tratamento con *para*-nitrocloroformato seguido de base, o que provoca unha ciclación intramolecular que forma o *Nbz*. Por último, desancórase todo o péptido da resina co *Nbz* en posición terminal usando TFA, e o derivado que resulta pode facilmente sufrir tiólise cun sulfhidriilo externo para xerar o tioéster para a LQN. Este método produce moi pouca epimerización, pero si se comprobou que a amina da *orto*-aminoanilida libre

<sup>32</sup> X. Q. Li, T. Kawakmi, S. Aimoto, *Tetrahedron Lett.* **1998**, 39, 8669.

<sup>33</sup> S. Futaki, K. Sogawa, J. Maruyama, T. Asahara, M. Niwa, H. Hojo, *Tetrahedron. Lett.* **1997**, 38, 6237.

<sup>34</sup> J. Alsina, T. S. Yokum, F. Albericio, G. J. Barany, *J. Org. Chem.* **1999**, 64, 8761

<sup>35</sup> Y. Shin, K. A. Winans, B. J. Backes, S. B. H. Kent, J. A. Ellman, C. R. Bertozzi *J. Am. Chem. Soc.* **1999**, 121, 11684.

<sup>36</sup> G. M. Fang, Y. M. Li, F. Shen, Y. C. Huang, J. B. Li, Y. Lin, H. K. Cui, L. Liu *Angew. Chem. Int. Ed.* **2011**, 50, 7645.

<sup>37</sup> a) J. Weidmann, E. Dimitrijevic, J. D. Hoheisel, P. E. Dawson *Org. Lett.* **2016**, 18, 164. b) J. X. Wang, G. M. Fang, Y. He, D. L. Qu, M. Yu, Z. Y. Hong, L. Liu *Angew. Chem. Int. Ed.* **2015**, 54, 2194.

<sup>38</sup> J. B. Blanco-Canosa, P. E. Dawson, *Angew. Chem. Int. Ed.* **2008**, 47,6851.

pode dar lugar a péptidos ramificados nesa posición. Para solucionar tal problema, no 2015, o mesmo grupo de Dawson publicou a mesma metodoloxía mellorada, na que a amina en *orto* estaba metilada e ó que denominaron *MeNbz*, un *Nbz* de segunda xeración.<sup>39</sup>

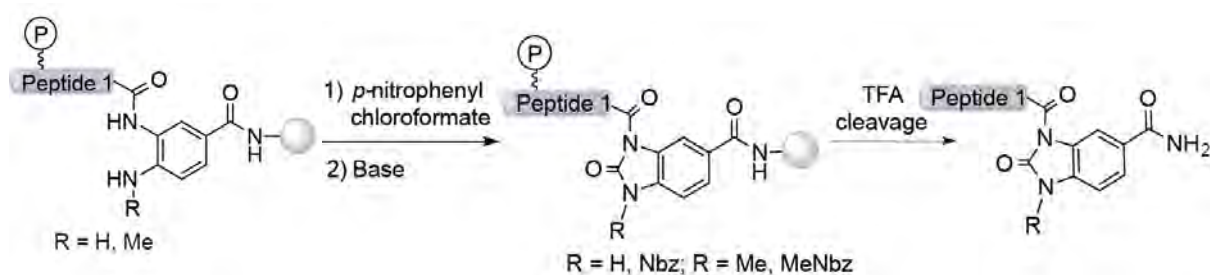


Figura 8. Proceso de formación do RNbz (primeira e segunda xeración) a partir do RDbz en fase sólida. Adaptado da ref. [40].



<sup>39</sup> J. B. Blanco-Canosa, B. Nardone, F. Albericio, P. E. Dawson, *J. Am. Chem. Soc.* **2015**, *137*, 7197.

<sup>40</sup> J. Palà-Pujadas, F. Albericio, J. B. Blanco-Canosa, *Angew. Chem. Int. Ed.* **2018**, *57*, 16120.

## 1.4. OS FACTORES DE TRANSCRICIÓN, A INTERACCIÓN CO ADN

### 1.4.1. O Dogma Central da Bioloxía Molecular

A información xenética que contén o ADN é a causa da síntese de proteínas específicas, responsables de case tódalas funcións e características dun organismo. Amais, a continuidade nunha especie esixe a transmisión fiel desta información dende as células nai ás células fillas. Unha célula usa tres procesos principais para manter intacta e converter a información xenética codificada no ADN en produtos funcionais, que poden ser ben proteínas, ben ARN. Estes procesos denomínanse: replicación, transcripción e tradución. O fluxo da información xenética dende o ADN a ARN ou proteínas descríbese mediante o coñecido como *Dogma central da Bioloxía Molecular*, que deu a coñecer Francis Crick nun simposio a finais do 57<sup>41</sup> e que sería publicado no 1970.<sup>42</sup> Polo tanto, estableceuse que a información xenética flúe da seguinte maneira: o ADN copia a súa mensaxe en forma de ARN mensaxeiro (transcripción) e, posteriormente, esta mensaxe é usada para sintetizar proteínas (tradución). Estas dúas etapas son coñecidas co nome de *Expresión Xénica*.

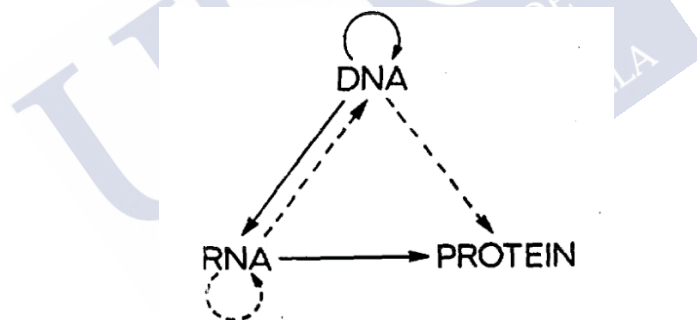


Figura 9. Esquema do dogma central da bioloxía molecular. As liñas continuas indican os pasos permitidos. As liñas descontinuas, só aqueles casos permitidos en ocasións excepcionais e a ausencia de liña indica que ese sentido de fluxo está prohibido. Adaptado da ref. [42]

A maiores, hai outras rutas de intercambio de información posibles, que se representan na **Figura 9**.

#### 1.4.1.1. A replicación

Cando unha célula se divide, o seu material xenético debe pasar ás células fillas de forma exacta. É por iso que o primeiro paso fundamental do *Dogma Central* é a duplicación ADN en dúas copias iguais. Para isto, a dobre hélice ábrese e o enzima *ADN Polimerasa III* realiza unha

<sup>41</sup> F. H. Crick, *Symp. Soc. Exp. Biol.* **1958**, 12, 138.

<sup>42</sup> F. H. Crick, *Nature* **1970**, 227, 561.

copia complementaria de cada cadea xerando dúas dobres hélices novas de maneira semiconservativa, é dicir, no que cada dobre hélice posúe unha cadea da febra nai.

## 1.4.2. A Expresión Xénica

### 1.4.2.1. A transcripción

Nesta etapa, a secuencia de ADN do xene que se vai expresar é rescrita en forma de ARN, que se denomina pre-ARN mensaxeiro (ARNm). En células eucariotas ten lugar no núcleo celular. O proceso lévase a cabo cando a *ARN polimerasa* se fixa a rexión específicas do ADN (xene promotor) e crea unha cadea de ARN complementaria e antiparalela ao ADN que quere copiar. Nesta cadea de pre-ARNm, a T é substituída por uracilo (U). O pre-ARNm sofre algunhas modificacións que o converten nun ARNm maduro e este pasa ao citoplasma

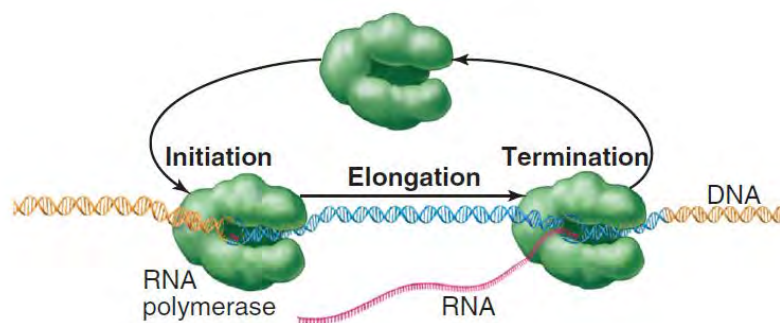


Figura 10. Representación do proceso de transcripción. A ARN polimerase (en verde) únese ao promotor e vai xerando o ARN mensaxeiro (vermello) durante a elongación. Adaptada da ref. [43]

### 1.4.2.2. A tradución

O último paso da expresión xénica consiste en descodificar o ARN mensaxeiro creado na transcripción, e convertelo na secuencia de aminoácidos determinada. Como neste paso se cambia da linguaxe dos nucleótidos aos aminoácidos, denomínase *tradución*. O ARNm viaxa dende o núcleo ata o citoplasma, onde se adhira aos ribosomas que se encargan de ler os nucleótidos do ARN e traducilos a aminoácidos para formar a proteína. Os nucleótidos son lidos en grupos de tres, chamados *codóns*. Hai 61 codóns que codifican os 20 aminoácidos e 3 codóns que se usan para marcar o fin da tradución. Os codóns son recoñecidos por moléculas de ARN especiais, denominadas ARN de transferencia (ARNt), os cales levan enganchados nun extremo

<sup>43</sup> *Cell Biology*, third edition, 2017, capítulo 10 pp 165-187.



o aminoácido correspondente co codón que len. Os ARNt axeitados únense aos codóns no ribosoma e provocan a que dous aminoácidos dos seus extremos próximos se unan, formando o polipéptido específico para ese ARNm. Cando os ARNt atopan o codón de terminación, as subunidades ribosómicas sepáranse e o péptido formado é liberado.

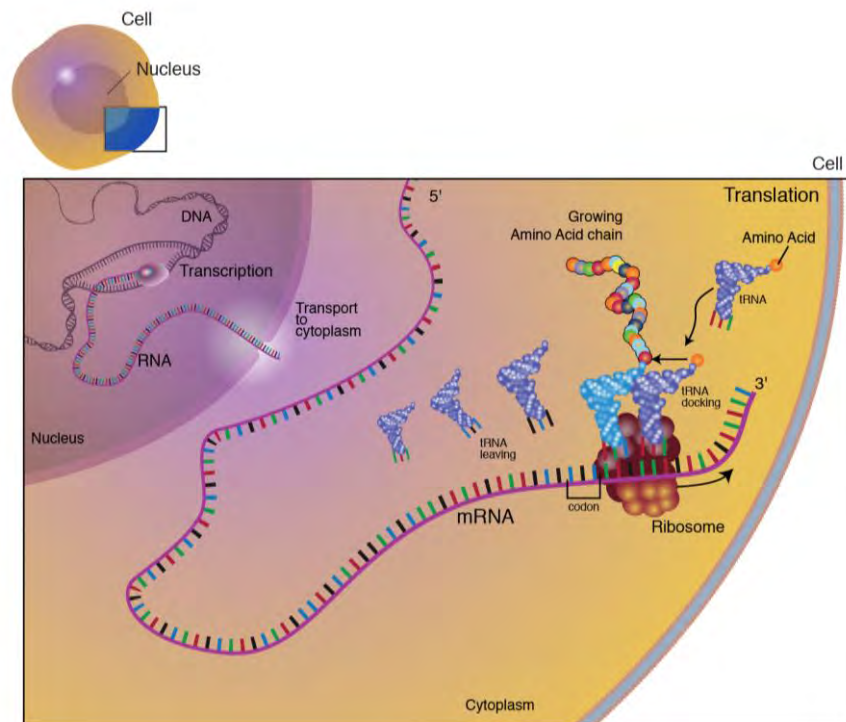


Figura 11. Representación da expresión xénica nunha célula eucariota. Adaptada da ref. [16]

A síntese de proteínas depende das necesidades celulares, só se realiza cando é preciso para non despilfarrar moléculas nin enerxía. Por outra banda, unha expresión descontrolada de xenes pode afectar a múltiples de procesos do ciclo celular e desencadear enfermidades. Dedúcese pois que a expresión dos xenes está altamente regulada, sobre todo en organismos pluricelulares para conseguir a especialización celular. A etapa máis controlada é o inicio da transcripción. De feito, é comunmente dito que se se controla a transcripción, contrólase a expresión xénica. Existen proteínas reguladoras que teñen o papel de inhibir ou activar esta etapa ao unirse a secuencias específicas de ADN dos promotores dos xenes que regulan. Estas proteínas denomínanse *Factores de Transcripción*.

### 1.4.3. Tipos de Factores de transcripción

Os factores de transcripción son as proteínas encargadas de activar ou inhibir a expresión dun xene mediante a súa unión a secuencias de ADN denominadas reguladoras. Esta unión permite controlar a transcripción dos xenes de forma espacial e temporal. Adoitan clasificarse segundo a súa estrutura de recoñecemento de ADN en:

#### 1.4.3.1. Dedos de Zinc

Este é o grupo de proteínas máis abundante. En células eucariotas, caracterízanse por ter unha rexión duns trinta aminoácidos que poden unirse tanto a ADN, como ARN ou outras proteínas ou lípidos, en forma de oligómeros. Adoitan ser estruturas de hélice  $\alpha$  combinada con dúas rexións de lámina  $\beta$  (chamadas  $\beta\beta\alpha$ ).<sup>44</sup> Estes dominios presentan cistinas e histidinas cuxas cadeas laterais se coordinan ao ión metálico de  $Zn^{2+}$ , coordinación que estabiliza a estrutura secundaria do dominio proteico, e facilita que a proteína adopte a conformación necesaria para a interacción co ADN.

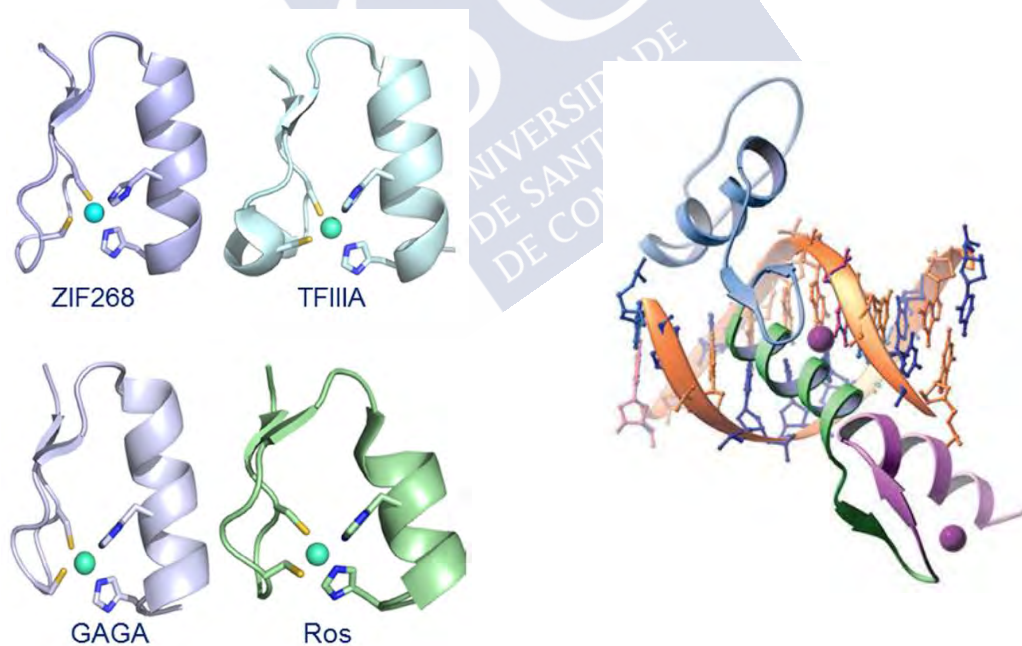


Figura 12. Esquerda: Estruturas tridimensionais dalgúns dominios de dedos de Zn de importantes factores de transcripción. As cistinas e as histidinas que coordinan o ión Zn están resaltadas na estrutura. Adaptada da ref. [45b]. Dereita: Proteína con tres dedos de cinc que recoñecen rexións con tres pares de bases consecutivas: GCG-TGG-GCG. Adaptada da ref. [43]

<sup>44</sup> J. M. Matthews, M. Sunde, *IUBMB Life* **2002**, *54*, 351.



### 1.4.3.2. Hélice-xiro-hélice (HTH) e Homeodominio

Este dominio é característico das células procariotas. Consta de dúas hélices  $\alpha$  de aproximadamente 60 aminoácidos separadas por un xiro  $\beta$ . Unha das hélices, *hélice de recoñecemento*, interacciona co suco maior do ADN recoñecendo unha secuencia de 6 pares de bases. Os monómeros HTH non son estruturas estables por si mesmos, polo que poden chegar a formar complexos con maior número de hélices. En bastantes casos, do extremo N-terminal colga un brazo flexible, con residuos básicos, que interacciona co suco menor aportando estabilidade e maior afinidade. Este tipo de estruturas en eucariotas denomínanse **Homeodomínios**.

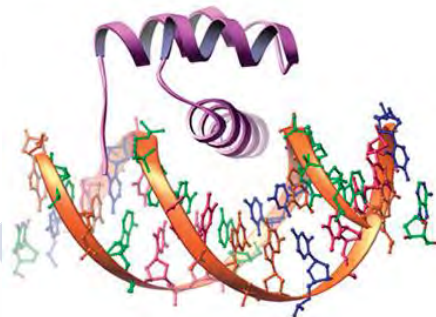


Figura 13. Exemplo de homeodominio recoñecendo a secuencia TAATGG. Adaptada da ref. [43]

### 1.4.3.3. Cremalleiras de Leucinas (bZip)

Estes factores de transcrición caracterízanse por ter dous motivos diferenciados: unha rexión rica en aminoácidos leucina, situada no extremo C-terminal, que permiten a dimerización, e outra rexión denominada *básica*, situada no extremo N-terminal, que recoñece ADN. Normalmente as Leu sitúanse cada sete aminoácidos (dúas voltas de hélice) da secuencia C-terminal. Unha proteína con este patrón de repetición das leucinas interacciona de forma hidrofóbica con outra hélice do mesmo tipo, creando dímeros (*coiled coil*). A rexión adxacente ás leucinas é moi básica, comprende varios residuos de Arg ou Lys que establecen interaccións iónicas favorables cos fosfatos e coas bases do ADN e permiten así a unión con este. É pois que se denomina habitualmente *cremalleira básica de leucinas (bZip)* en lugar de *LeuZip*.

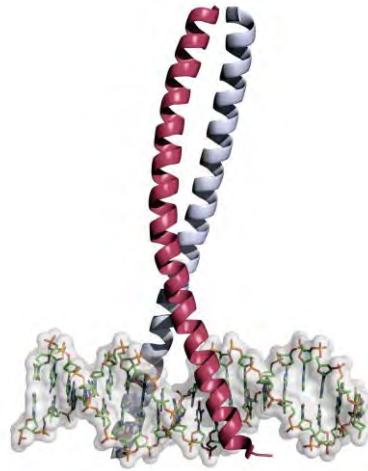


Figura 14. Raios X do dímero bZip de *c-Jun* (vermello) e *c-Fos* (azul) recoñecendo a secuencia consenso de ADN. O dímero Fos-Jun está envolto en procesos de regulación da proliferación celular, da apoptose ou diferenciación.

#### 1.4.3.4. Dominios hélice-bucle-hélice (HLHZip)

O dominio de HLH comprende aproximadamente uns 60 aminoácidos con dúas rexións diferentes: unha é a rexión básica e a outra o motivo HLH. A parte básica adoita estar no extremo *N*-terminal e consta duns 15 aminoácidos cuns seis residuos básicos. Esta parte é a que reconece ADN. A outra rexión, máis longa, contén dúas hélices  $\alpha$  de carácter anfipático, separadas por un pequeno bucle, xeralmente de 8 aminoácidos. Esta parte hidrofóbica permite a dimerización entre proteínas,<sup>45</sup> de forma similar a unha proteína bZIP.

#### 1.4.4. *c*-MYC

*c*-MYC é unha proteína de 439 aminoácidos da familia dos factores de transcrición MYC, que son cremalleiras de leucinas cunha rexión básica e con dominios hélice-bucle-hélice (HLH). Esta familia está formada por tres membros: *c*-MYC, N-MYC e L-MYC, pero moitas proliferacións celulares dependen unicamente de *c*-MYC. *c*-MYC (a partir d'agora chamaráselle MYC) está involucrada en múltiples funcións biolóxicas das células somáticas: na regulación do crecemento, proliferación, no metabolismo, na apoptose e na diferenciación celular.<sup>46</sup> Unha das funcións é promover a transcrición de determinados xenes e reprimir a expresión doutros: Por un lado, MYC encárgase de activar a transcrición de xenes que promoven o ciclo celular<sup>47</sup> e, por outro, inhibe a transcrición de xenes que deteñen o ciclo celular, é por iso que pode

<sup>45</sup> a) W. R. Atchley, W. Terhalle, A. Dress, *J. Mol. Evol.* **1999**, 48, 501. b) S. Jones, *Genome Biol.* **2004**, 5, 226.

<sup>46</sup> N. Meyer, L. Z. Penn, *Nat. Rev. Cancer* **2008**, 8, 976.

<sup>47</sup> a) S. Adhikary, M. Eilers, *Nat. Rev. Mol. Cell Biol.* **2005**, 6, 635. b) M. D. Cole, M. A. Nikiforov, *Curr. Top. Microbiol. Immunol.* **2006**, 302, 33.

promover a proliferación celular.<sup>48</sup> A célula controla estritamente o mecanismo de regulación de MYC mais pequenos fallos provocan a súa sobreexpresión que se traduce nunha actividade proliferativa aumentada.<sup>49</sup>

A rexión básica de MYC encárgase de facer posible a unión ao suco maior do ADN, mentres que o dominio de cremalleira de leucinas rexe a súa capacidade de dimerización. De feito, para que MYC leve a cabo as súas funcións, debe heterodimerizar con outra bHLHZip chamada MAX.<sup>50</sup> Este heterodímero MYC/MAX enlázase a unha rexión específica do xene promotor diana, o *E-box* (ou rexión E), que presenta a secuencia diana *CACGTG*.<sup>51</sup>

MAX é capaz de formar homodímeros MAX/MAX e tamén heterodímeros con outras proteínas: MYC/MAX, MXD/MAX, MNT/MAX etc (ver **Figura 16**). Pola contra, MYC só heterodimeriza con MAX e os homocomplexos MYC/MYC son moi inestables.

Grazas ás estruturas cristalinas e á elucidación dos raios X dos dímeros unidos á secuencia de ADN *CACGTG*, puido descifrarse a topoloxía e a forma coa que estas familias de factores de transcrición interaccionan co ADN.<sup>52</sup> Cada proteína individualmente presenta dous dominios de hélices  $\alpha$  separados polo bucle. A primeira hélice  $\alpha$  contén residuos da rexión básica e da denominada hélice 1 (H1), a segunda hélice  $\alpha$  abrangue a H2 e a cremalleira de leucinas. Os complexos de recoñecemento consisten en dous monómeros que se dispoñen paralelamente para dar un dímero con catro hélices  $\alpha$  separadas por dous bucles. As rexións básicas insírense no suco maior do ADN mediante contactos específicos coa *E-box*. Os dous segmentos C-terminais interaccionan mediante forzas iónicas e hidrofóbicas estabilizando a unión das dúas proteínas.<sup>53</sup>

<sup>48</sup> D. Kleine-Kohlbrecher, S. Adhikary, M. Eilers, *Curr. Top. Microbiol. Immunol.* **2006**, 302, 51.

<sup>49</sup> M. Conacci-Sorrell, L. McFerrin, R. N. Eisenman, *Cold Spring Harbor Perspect. Med.* **2014**, 4, a014357.

<sup>50</sup> a) B. Amati, M. W. Brooks, N. Levy, T. D. Littlewood, G. I. Evan, H. Land, *Cell* **1993**, 72, 233. b) E. M. Blackwood, B. Luscher, R. N. Eisenman, *Genes Dev.* **1992**, 6, 71. c) A. L. Gartel, K. Shchors, *Exp. Cell Res.* **2003**, 283, 17.

<sup>51</sup> a) P. C. Fernández, S. R. Frank, L. Wang, M. Schroeder, S. Liu, J. Greene, A. Cocito, B. Amati, *Genes Dev.* **2003**, 17, 1115. b) G. C. Prendergast, D. Lawe, E. B. Ziff, *Cell* **1991**, 65, 395.

<sup>52</sup> A) A. R. Ferre-D'Amare, G. C. Prendergast, E. B. Ziff, S. K. Burley, *Nature* **1993**, 363, 38. B) S. K. Nair, S. K. Burley, *Cell* **2003**, 112, 193

<sup>53</sup> L. A. Jung, A. Gebhardt, W. Koelmel, C. P. Ade, S. Walz, J. Kuper, B. von Eyss, S. Letschert, C. Redel, L. d'Artista, A. Biankin, L. Zender, M. Sauer, E. Wolf, G. Evan, C. Kisker, M. Eilers, *Oncogene* **2017**, 36, 1911.

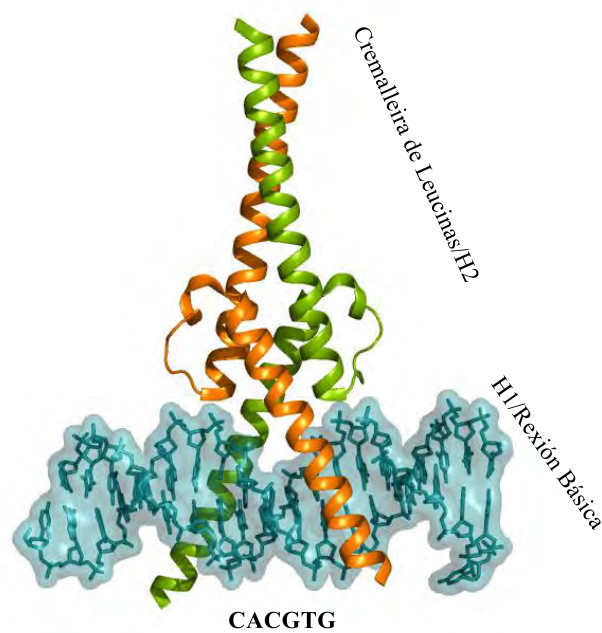


Figura 15. Representación da interacción de heterodímero MYC (verde)/MAX (laranja) co ADN. PDB 1NKP.

MYC e MAX pertencen a unha rede moi ampla de factores de transcrición do tipo bHLHZip que se encargan de regular diferentes aspectos do comportamento da celular. A maiores da familia de proteínas de MYC, nesta rede transcripcional tamén interveñen membros de MXD (coñecidos orixinalmente como MAD) ou MNT.

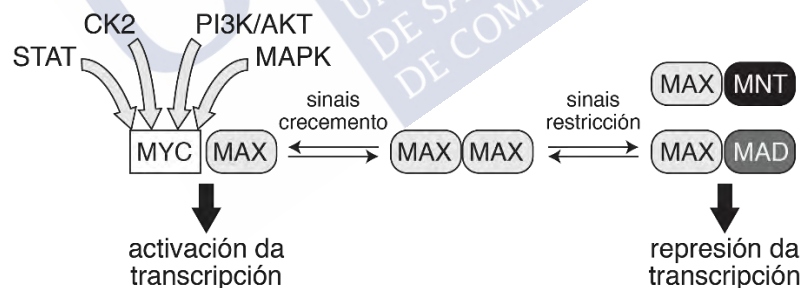


Figura 16. Esquema simplificado do mecanismo de activación de xenes promotores da proliferación celular debido á dimerización de MYC/MAX.

Para activar a proliferación celular, o dímero MYC/MAX únese á *E-box* dos xenes promotores para iniciar a súa transcrición. Os membros de Mxd actúan como antagonistas de MYC, así, as proteínas MXD e MNT forman heterodímeros con MAX e compiten co heterodímero MYC/MAX por unirse ás mesmas rexións de ADN (*E-boxes*) dos xenes promotores (ver **Figura 16**).<sup>54</sup> O mecanismo para a inhibición da transcrición dos xenes

<sup>54</sup> A) M. Conacci-Sorrell, L. McFerrin, R. N. Eisenman, *Cold Spring Harbor Perspect. Med.* **2014**, *4*, a014357. B) C. Grandori, S. M. Cowley, L. P. James, R. N. Eisenman, *Annu. Rev. Cell Dev. Biol.* **2000**, *16*, 653.

involucrados na supresión do ciclo celular non está tan claro, pero tamén depende da interacción do dímero MYC/MAX con outros co-reguladores e factores de transcrición.<sup>55</sup>

MYC está considerada como unha das principais dianas terapéuticas no cancro, de feito, estímase que MYC está desregulado no 75 % do tumores.<sup>56</sup> Ao mesmo tempo, é unha das dianas máis difíciles de abordar directamente, por varios motivos: o primeiro é que MYC se localiza no núcleo celular, localización complicada de acadar para anticorpos ou moléculas grandes deseñados para interaccionar con MYC; o segundo, que non posúe un sitio activo característico cunha conformación determinada, como o das enzimas, que poida servir como diana para pequenas moléculas.<sup>57</sup>

Como aínda no existe un método efectivo de interaccionar con MYC directamente, cada vez son máis os métodos indirectos para bloquear as funcións desta proteína, actuando sobre outros compostos esenciais que interveñen na desregulación da oncoproteína. Porén, hai moito interese en bloquear a dimerización entre MYC-MAX que é necesaria para unirse ao ADN e activar a transcrición. De feito, xa se desenvolvéron varios compostos que inhiben esta dimerización: a molécula pequena IIA6B17,<sup>58</sup> o composto 10058-F4<sup>59</sup> ou o péptido OMOMYC, creado e amplamente estudado pola Dr. Laura Soucek.

#### 1.4.5. OMOMYC

OMOMYC é unha miniproteína de 93 aminoácidos que xorde da mutación do dominio bHLHZip de MYC, e que actúa coma o seu dominante negativo.<sup>60</sup> Ca mesma secuencia que o extremo C-terminal da rexión Zip de MYC, diferéncianse só en catro mutacións na cremalleira de leucinas.

Este análogo de MYC é quen de homodimerizar e heterodimerizar formando tanto MYC/OMOMYC como MAX/OMOMYC. O homodímero OMOMYC/OMOMYC ( $\Delta G = -41$

<sup>55</sup> A) S. Adhikary, M. Eilers, *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 635. B) M. Eilers, R. N. Eisenman, *Genes Dev.* **2008**, *22*, 2755.

<sup>56</sup> M. Kalkat, J. De Melo, K. A. Hickman, C. Lourenco, C. Redel, D. Resetca, A. Tamachi, W. B. Tu, L. Z. Penn, *Genes* **2017**, *8*, 151.

<sup>57</sup> A) L. Soucek, J. Whitfield, C. P. Martins, A. J. Finch, D. J. Murphy, N. M. Sodik, A. N. Karnezis, L. B. Swigart, S. Nasi, G. I. Evan, *Nature* **2008**, *455*, 679. B) L. A. Cabaret, P. S. Rennie, A. Cherkasov, *Int. J. Mol. Sci.* **2019**, *20*, 120.

<sup>58</sup> T. Berg, S. B. Cohen, J. Desharnais, C. Sonderegger, D. J. Maslyar, J. Goldberg, D. L. Boger, P. K. Vogt, *Proc. Natl Acad. Sci.* **2002**, *99*, 3830.

<sup>59</sup> H. Wang, D. I. Hammoudeh, A. V. Follis, B. E. Reese, J. S. Lazo, S. J. Metallo, E. V. Prochownik, *Mol. Cancer Ther.* **2007**, *6*, 2399.

<sup>60</sup> A) L. Soucek, M. Helmer-Citterich, A. Sacco, R. Jucker, G. Cesareni, S. Nasi, *Oncogene*. **1998**, *17*, 2463. B) L. Soucek, R. Jucker, L. Panacchia, R. Ricordy, F. Tato, S. Nasi, *Cancer Research*, **2002**, *62*, 3507.



Kcal/mol) é máis estable que MYC/MAX ( $\Delta G = -33$  Kcal/mol). Amais, o complexo OMOMYC/OMOMYC atópase máis estabilizado que os seus respectivos heterodímeros grazas a varias interaccións adicionais debidas ás mutacións, e que non están presentes nos outros complexos.<sup>55</sup>

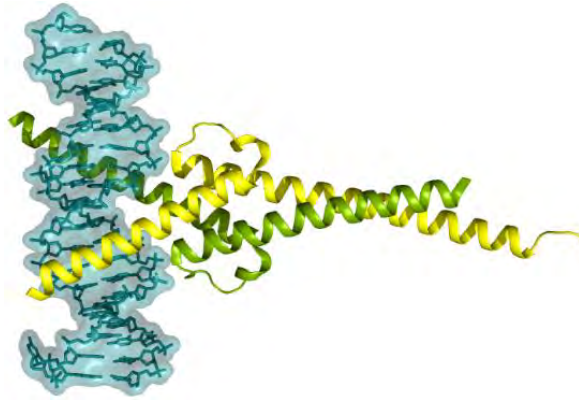


Figura 17. Representación do heterodímero MYC (verde)/OMOMYC (amarelo) co ADN.

Estudos nos que se inducía a expresión de OMOMYC en ratos, demostraron unha regresión rápida de tumores pancreáticos, glioblastomas e adenocarcinomas de pulmón con efectos secundarios reversibles en tecidos normais.<sup>61</sup> A maiores, un estudo que diferenciaba os efectos e a actividade de OMOMYC en ambientes tumorais e non tumorais demostrou que mentras o recoñecemento de rexións específicas de ADN por parte da proteína MYC case non se alteraba, en medios tumorais a expresión de OMOMYC reduce a unión de MYC a secuencias polas que mostra baixa afinidade.<sup>55</sup>

Barállanse varios mecanismos de actuación de OMOMYC, un deles é que o homocomplexo OMOMYC/OMOMYC compite polo sitio de unión ao promotor con MYC/MAX, suprimindo a transcripción. Outro mecanismo proposto baséase na formación do dímero OMOMYC/MYC, que diminúe os niveis de MYC no medio, ‘secuestrándoo’ e reducindo a asociación MYC/MAX.

<sup>61</sup> L. Soucek, J. Whitfield, C.P. Martins, A. J. Finch, D. J. Murphy, N. M. Sodar, *Nature* **2008**, *455*, 679. B) D. Annibali, J. R. Whitfield, E. Favuzzi, T. Jauset, E. Serrano, I. Cuartas *Nat. Commun.* **2014**; *5*, 4632. C) N. M. Sodar, L. B. Swigart, A. N. Karnezis, D. Hanahan, G. I. Evan, L. Soucek *Genes Dev.* **2011**, *25*, 907.

### 1.4.6. Métodos de síntese de *c*-MYC e OMOMYC

Aínda que o interese e potencial biomédico tanto de MYC como de OMOMYC é inmenso, existen moi poucas aproximacións sintéticas para a súa obtención, á parte das técnicas recombinantes.

En 1995, o grupo de Kent deu a coñecer a síntese dun derivado do heterodímero MYC/MAX.<sup>62</sup> Este complexo consistía en unir as dúas hélices de cada monómero a través dun enlace de tipo tioéster e unir os monómeros entre si cun enlace oxima. Desafortunadamente, debido ao tioéster, esta estrutura resultou bastante inestable.

Máis recente, no 2017, o grupo de Madder acadou a síntese dun derivado da rexión básica do dímero MYC/MAX que era capaz de recoñecer a secuencia de ADN diana do dímero natural.<sup>63</sup> Nesta síntese introducen un espaciador de natureza esteroidea para unir as rexións básicas das dúas proteínas, polo que non se trata dun dímero proteico propiamente dito.

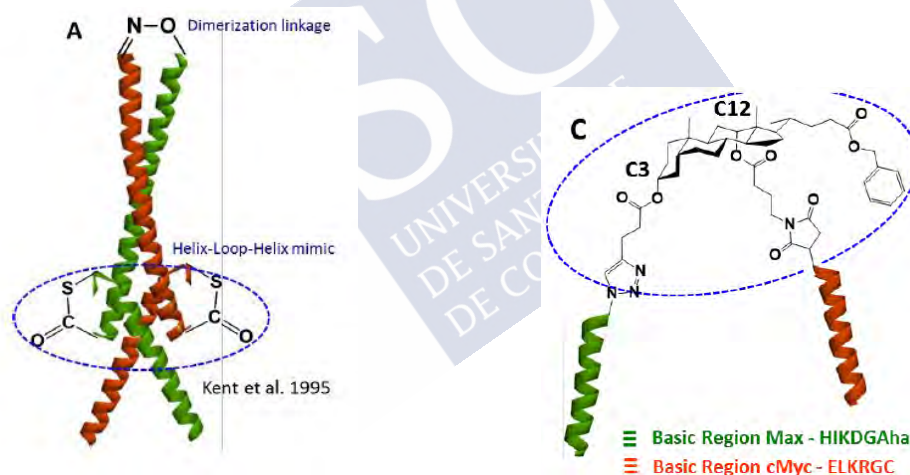


Figura 18. Representación dos miméticos do dímero MYC/MAX acadados por Kent (esquerda) e por Madder (dereita). Adaptada da ref. [63].

Para OMOMYC non se coñecen derivados ou análogos sintetizados quimicamente, ata o de agora.

Un dos maiores impedimentos á hora de explorar as funcións terapéuticas e as propiedades biolóxicas destas proteínas deriva da falta de un método eficiente de síntese. Os métodos recombinantes teñen un alcance moi limitado e non permiten a introdución de modificacións

<sup>62</sup> L. E. Canne, A. R. Ferre- D'Amare, S. K. Burley, S. B. H. Kent, *J. Am. Chem. Soc.* **1995**, *117*, 2998.

<sup>63</sup> Y. R. García, Y. V. Pabon -Martínez, C. I. E. Smith, A. Madder, *Chem. Commun.* **2017**, *53*, 6653.

non naturais na secuencia dos péptidos. Tampouco son suficientemente versátiles para facer derivados e conxugados, por exemplo con fluoróforos.<sup>18,19,20,21</sup>

É pois de gran importancia, traballar na creación de métodos sintéticos versátiles que permitan producir estas proteínas e distintos derivados.





## 1.5. RECOÑECIMENTO DE ADN POR MOLÉCULAS PEQUENAS

Como se viu antes, o ADN intervén en múltiples procesos fundamentais para a integridade da célula, e o almacenamento e traspaso da información xenética. En moitos casos, estes procesos conlevan a interacción con distintos compoñentes celulares, interaccións que están perfectamente coordinadas no tempo e no espazo, e teñen lugar, normalmente, en sitios específicos do ADN. Se algunha destas interaccións ten lugar de forma irregular, a célula descompénsase, e poden xurdir enfermidades coma o cancro.<sup>64</sup>

Dada a importancia do recoñecemento molecular do ADN, ao longo destas décadas os químicos e biólogos traballaron para deseñar ligandos non naturais capaces de interaccionar de forma selectiva e controlada con secuencias específicas do ADN, con obxectivo final de obter novos axentes terapéuticos.<sup>65</sup> Existe unha gran variedade de pequenas moléculas que interaccionan con ácidos nucleicos<sup>66,67</sup> e, segundo o xeito de interacción pódense clasificar en:

### 1.5.1. Unión covalente: axentes alquilantes

Estas moléculas adoitan ser electrófilos fortes que forman enlaces covalentes coas bases do ADN. O tipo de unión é normalmente irreversible, e pode ter consecuencias nas funcións de replicación e transcrición, levando incluso á morte celular. A maioría destas moléculas teñen selectividade moi baixa, que se observa na alta toxicidade en células normais. Polo xeral, a reacción de alquilación ten lugar por ataque nucleofílico do N<sup>7</sup> ou C-NH<sub>2</sub> da guanina.<sup>68</sup>

Dentro deste grupo atópanse tanto moléculas puramente orgánicas como a mitomicina C ou antramicina ou compostos metálicos famosos como o cisplatino ou carboplatino.

<sup>64</sup> R. L. Gurung, H. K. Lim, S. Venkatesan, P. S. Lee, M. P. Hande, *Mol. Cancer* **2014**, *13*, 232

<sup>65</sup> S. Bhaduri, N. Ranjan, D. P. Arya, *Beilstein J. Org. Chem.* **2018**, *14*, 1051.

<sup>66</sup> A) M. Wang, Y. Yu, C. Liang, A. Lu, G. Zhang, *Int. J. Mol. Sci.* **2016**, *17*, 779. B) Y. H. Du, J. Huang, X. C. Weng, X. Zhou, *Curr. Med. Chem.* **2010**, *17*, 173.

<sup>67</sup> S. Neidle, *Principles of Small Molecule-DNA Recognition* **2008**.

<sup>68</sup> L. F. Povirk, D. E. Shuker, *Mutat. Res.* **1994**, *318*, 205.

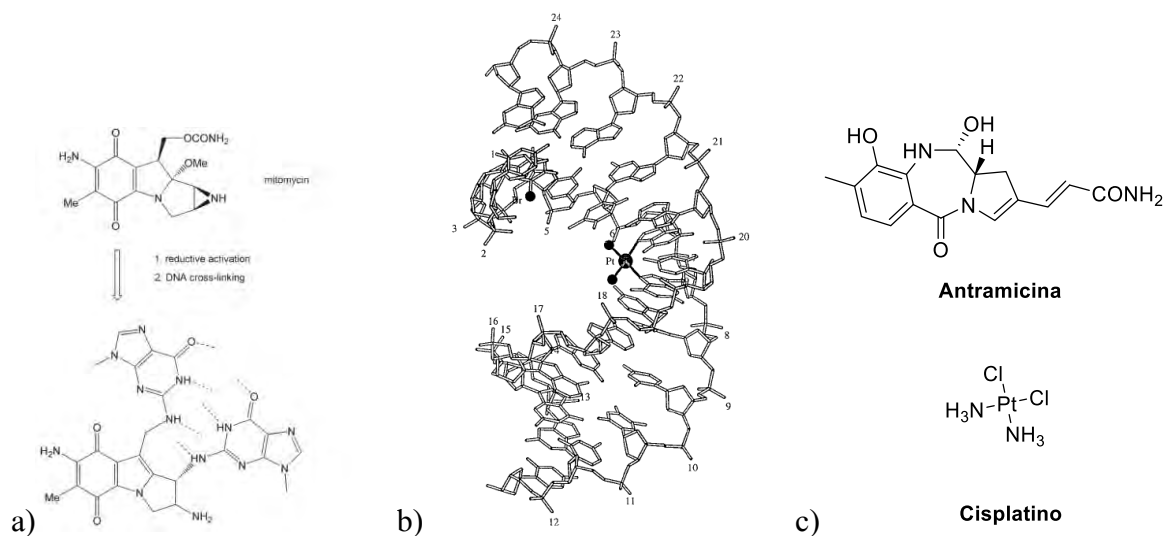
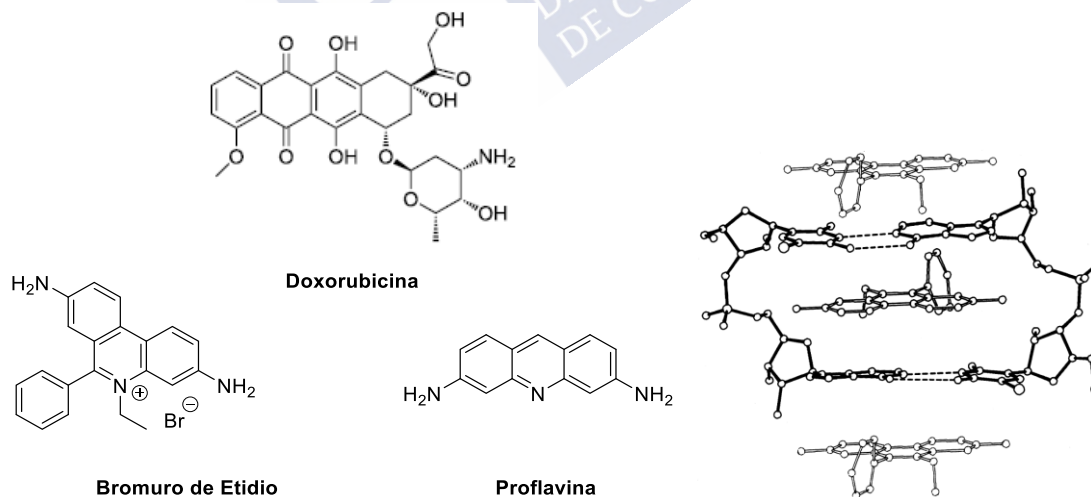


Figura 19. A) Esquema de como a mitomicina se une a dúas guaninas das dúas febras do ADN. Imaxe adaptada da ref. [67]. B) modelo do cisplatino unido a dúas guaninas do ADN de dobre hélice, ref. [69] C) estruturas do cisplatino e antramicina.

## 1.5.2. Unións non covalentes

### 1.5.2.1. Intercalantes

Polo xeral, os compostos intercalantes son moléculas planas aromáticas que se intercalan entre os pares de bases adxacentes do ADN mediante interaccións de apilamento  $\pi$ - $\pi$ , interaccións de Van der Waals, transferencias de carga ou interaccións hidrofóbicas.<sup>70</sup>



<sup>69</sup> P. M. Takahara, C. A. Frederick, S. J. Lippard, *J. Am. Chem. Soc.* **1996**, *118*, 12309

<sup>70</sup> A) J. B. Chaires, *Biopolymers* **1997**, *44*, 201. B) R. Martinez, L. Chacon-Garcia, *Curr. Med. Chem.* **2005**, *12*, 127. C) A. Rescifina, Z. Zagni, M. G. Varrica, V. Pistarà, A. Corsaro, *Eur. J. Med. Chem.* **2014**, *74*, 95.

Figura 20. Esquerda: estruturas dalgúns axentes intercalantes coñecidos. Dereita: estrutura cristalina do etidio intercalado entre os pares de bases. Adaptado de [71]

A intercalación provoca cambios conformacionais locais no ADN, tales como a contracción ou estiramento da hélice, que pode afectar aos procesos de reparación, replicación ou transcrición. Grazas a isto, poden usarse como antibióticos, antitumorais, antibacterianos. ou axentes antivirais. Moitos destes compostos son ben coñecidos: bromuro de etidio, acridinas, doxorubicina, antramincinas, etc.

### 1.5.2.2. Unión ao suco maior

Aínda que a maioría dos compostos que se unen ao suco maior son macromoléculas de tipo proteico mediante pontes de hidróxeno,<sup>72</sup> tamén existen ligandos de pequeno tamaño que se unen a esta rexión da dobre hélice. Habitualmente, inhiben as interaccións proteína-ADN dos factores de transcrición, bloqueando a súa función. Existen multitude de produtos naturais que realizan este tipo de interaccións: pluramicinas, aflatoxinas, leinamicinas etc.<sup>73</sup> Quizáis os máis coñecidos son os aminoglicósidos, coma a neomicina ou a paromomicina, que interaccionan cos grupo fosfato a través dos seus policationís. Tamén se reportaron varios híbridos de neomicina-HOECHST e neomicina-HOECHST-pireno comprobándose a súa unión no suco maior.<sup>74</sup>

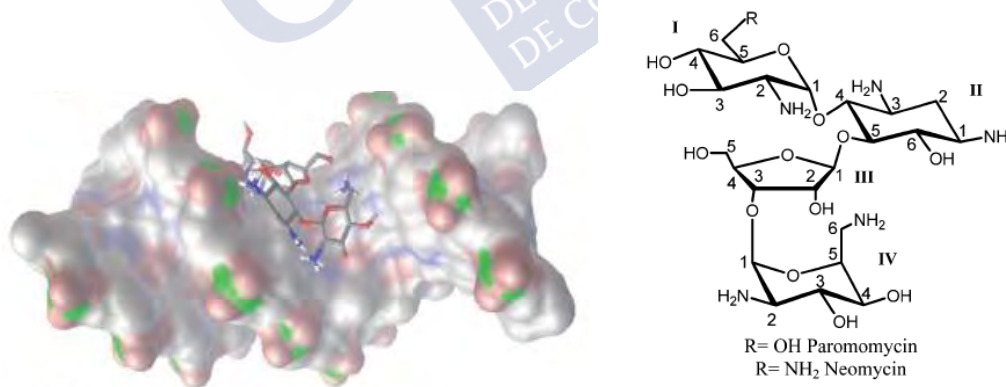


Figura 21. Esquerda: Modelo computacional da neomicina unida a ADN. Dereita: estruturas neomicina e paromomicina. Adaptada da ref. [73]

<sup>71</sup> C.-C. Tsai, S. C. Jain, H. M. Sobell, *Proc. Nat. Acad. Sci.* **1975**, 72, 628.

<sup>72</sup>A) N. N. Singh, A. M. Lambowitz, *J. Mol. Biol.* **2001**, 309, 361. B) N. M. Mamoon, Y. Song, S. E. Wellman, *Biochemistry* **2002**, 41, 9222.

<sup>73</sup> P. L. Hamilton, D. P. Arya, *Nat. Prod. Rep.* **2012**, 29, 134.

<sup>74</sup> A) B. Willis, D. P. Arya, *Biochemistry* **2010**, 49, 452. B) B. Willis, D. P. Arya, *Chem. Biochem.* **2006**, 60, 251.

### 1.5.2.3. Unión ao suco menor

Este é o grupo maioritario de axentes que se unen ao ADN. Os compostos máis propensos a realizar insercións no suco menor presentan heterociclos aromáticos (coma os polipirrois) con dicatións, poliaminas ou poliamidas. Adoitan ter estruturas ríxidas e planas compatibles co tamaño e forma do suco e desprazan moléculas de auga ao inserirse.<sup>75</sup> A conformación da dobre hélice non se ve demasiado afectada despois da inserción. A unión estabilízase por enlaces de hidróxeno, forzas de Van der Waals e interaccións hidrofóbicas. Estes compostos teñen preferencia por rexións ricas en A/T, sendo os máis coñecidos a neotripsina, o berenil, a penta e a propamidina, a distamicina, Hoechst ou o DAPI.

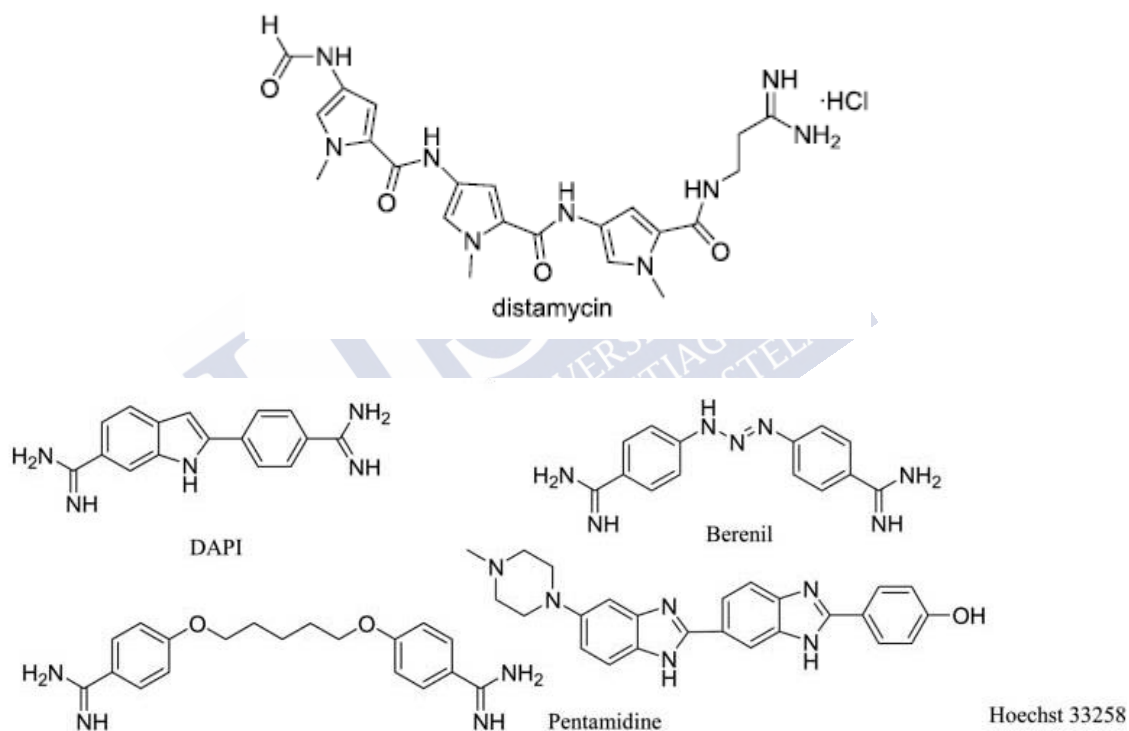


Figura 22. Estructuras dalgúns compostos que se insiren no suco menor.

<sup>75</sup> A) W. S. Somers, S. E. Phillips, *Nature* **1992**, 359, 387. B) T. Li, Y. Jin, A. K. Vershon, C. Wolberger, *Nucleic Acids Res.* **1998**, 26, 5707.

A química sintética permitiu a preparación de numerosos derivados da distamicina, neotripsina ou propamidina,<sup>76</sup> algún dos cales presentan propiedades de recoñecemento interesantes.

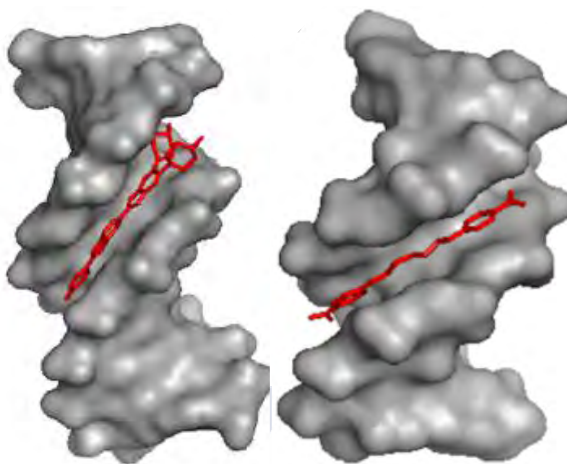


Figura 23 Estrutura cristalina formada por Hoescht 33258(esquerda) e pentamidina (dereita) co suco menor do ADN. Adaptada da ref. [65].

### Bisbenzamidinas

As bisamidinas son os axentes de recoñecemento do suco menor con aplicacións médicas máis antigos e dos máis estudados: o DAPI (usado como tinguidor de núcleos celulares), a furamidina ou a pentamidina son os máis relevantes, e téñense usado incluso na medicina como antiparasitarios.<sup>77</sup> A pentamidina, ao unirse ao ADN, dispón a súa parte central inserida no interior do suco, mentres que os extremos amidínicos quedan interaccionando co esqueleto fosfato aniónico.<sup>78</sup> A pesares de que estas moléculas son pequenas e relativamente simples, a súa obtención sintética non é trivial e normalmente conleva rendementos pobres.<sup>79</sup>

<sup>76</sup> A) C. Sessa, O. Pagani, M. G. Zurlo, J. de Jong, C. Hofmann, M. Lassus, P. Marrari, M. Strolin Benedetti, F. Cavalli, *Ann. Oncol.* **1994**, *5*, 901. B) M. Brogini, H. M. Coley, N. Mongelli, E. Pesenti, M. D. Wyatt, J. A. Hartley, M. D'Incalci, *Nucleic Acids Res.* **1995**, *23*, 81.

<sup>77</sup> A) W. D. Wilson, F. A. Tanius, A. Mathis, D. Tevis, J. E. Hall, W. D. Boykin, *Biochimie* **2008**, *90*, 999. B) J. K. Thuita, S. M. Karanja, T. Wenzler, R. E. Mdachi, J. M. Ngotho, J. M. Kagira, R. Tidwell, R. Brun, *Acta Trop.* **2008**, *108*, 6. C) S. Neidle, L. R. Kelland, J. O. Trent, I. J. Simpson, D. W. Boykin, A. Kumar, W. D. Wilson, *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1403. D) G. Susbielle, R. Blattes, V. Brevet, C. Monod, E. Kas, *Curr. Med. Chem.-Anti-Cancer Agents*, **2005**, *5*, 409.

<sup>78</sup> M. R. Conte, T. C. Jenkins, A. N. Lane, *Eur. J. Biochem.* **1995**, *229*, 433.

<sup>79</sup> F. Rodríguez, I. Rozas, M. Kaiser, R. Brun, B. Nguyen, W. D. Wilson, R. N. Garcia, C. Dardonville, *J. Med. Chem.* **2008**, *51*, 909.

No noso grupo de investigación descubriuse que a substitución dos osíxenos da propamidina por nitróxenos provoca o desprazamento cara maiores lonxitudes de onda de excitación ( $\lambda = 329$  nm) e de emisión ( $\lambda = 387$  nm). Amais, a intensidade de emisión fluorescente ten gran sensibilidade á polaridade do seu entorno. Deste xeito, a *bis(4-amidinofenil)propan-1,3-diamina* (BAPPA), mostrou un incremento de fluorescencia ao incubarse en unha disolución 0.5  $\mu$ M en tampón Tris·HCl pH 7.5 20 mM, 100 mM NaCl co o seu ADN diana (5'-GCGC AAATTT GCGTTTTTCGC AAATTT GCGC-3'). Este novo ligando do ADN, selectivo para secuencias ricas en A/T, tamén foi quen de internalizar células HeLa conservando a súas propiedades fluoroxénicas.<sup>80</sup> As propiedades solvatocrómicas permiten estudar a unión a ADN sen recorrer a técnicas máis agresivas como as que se viñan usando ata o momento, como a desnaturalización térmica de complexos ligando-ADN.<sup>81</sup>

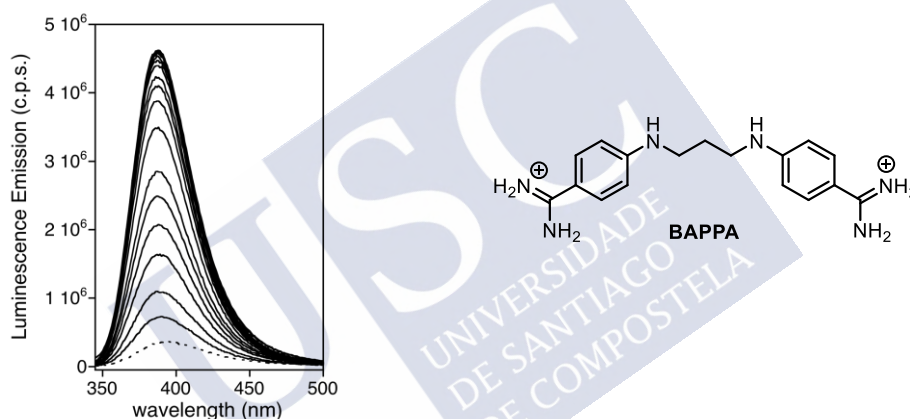


Figura 24 Esquerda: estrutura da azapropamidina (BAPPA). Dereita: espectro de emisión do BAPPA tras realizar unha valoración fluorescente engadindo a unha disolución 0,5  $\mu$ M de BAPPA, alicuotas do seu ADN diana.

O noso grupo de investigación desenvolveu un método moi sinxelo para obter este tipo de bisbenzamidinas fluoróforas nun único paso.<sup>82</sup> A reacción consiste nunha simple aminación redutora entre a 4-aminobenceno carboximidamida, reactivo comercial, e diferentes tipos de dialdehidos (ver **Figura 25**).

Deste xeito, pódese acceder, dunha forma rápida e versátil, a numerosos derivados bisbenzamidínicos, capaces de interaccionar con secuencias de ADN ricas en A/T. Todos estes derivados mostraron ser selectivos para secuencias ricas en A/T, cunhas constantes de afinidade

<sup>80</sup> O. Vazquez, M. I. Sanchez, J. Martinez-Costas, M. E. Vazquez, J. L. Mascareñas, *Org. Lett.* **2010**, *12*, 216.

<sup>81</sup> W. D. Wilson, F. A. Tanius, M. Fernandez-Saiz, C. T. Rigl, *Methods Mol. Biol.* **1997**, *90*, 219.

<sup>82</sup> M. I. Sánchez, O. Vázquez, J. Martínez-Costas, M. E. Vázquez, J. L. Mascareñas, *Chem. Sci.*, **2012**, *3*, 2383.



arredor de baixo micromolar. Esta aproximación permitiu deseñar facilmente varios sistemas fluorescentes baseados en marcadores con bisbenzamidinas.<sup>83</sup>

A pesar de tódolos avances, aínda hai moitos desafíos e retos neste ámbito da investigación, tales como a preparación de compostos que poidan responder a estímulos externos, o derivados con afinidades mais altas e boas selectividades.

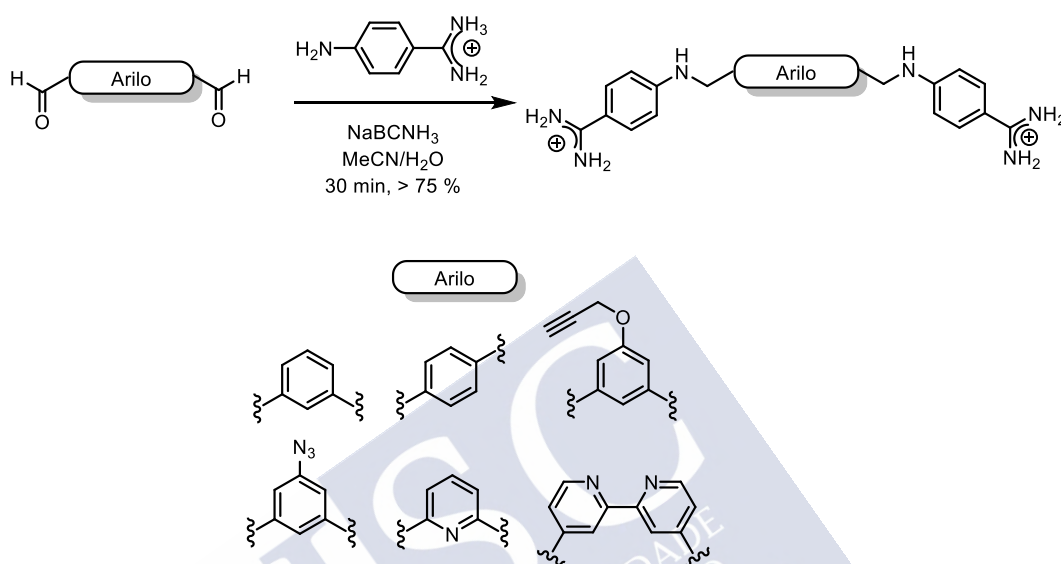


Figura 25. Reacción de aminación redutora entre a 4-aminobenceno carboximidamida e diferentes dialdehidos descuberta no grupo para obter bisbenzamidinas. Abaixo: Diferentes substituíntes arilo que poden conter os dialdehidos.

<sup>83</sup> A) M. I. Sánchez, O. Vázquez, M. E. Vázquez, J. L. Mascareñas, *Chem. Eur. J.* **2013**, *19*, 9923 B) M. I. Sánchez, C. Pernas, E. Vázquez Sentís, J. L. Mascareñas, *Chem. Sci.*, **2014**, *5*, 1901. C) J. Bordello, M. I. Sánchez, M. E. Vázquez, J. L. Mascareñas, W. Al-Soufi, M. Novo, *Chem. Eur. J.* **2015**, *21*, 1609. D) M. I. Sánchez, G. Rama, R. Calo-Lapido, K. Ucar, P. Lincoln, M. Vázquez, M. Melle-Franco, J. L. Mascareñas, M. E. Vázquez, *Chem. Sci.* **2019**, doi: 10.1039/C9SC03053K.







## **2. OBXECTIVOS XERAIS**



Nesta tese de doutoramento preténdese deseñar e sintetizar sistemas capaces de interaccionar de forma específica cos ácidos nucleicos, fundamentalmente co ADN.

Un dos obxectivos é descubrir métodos para manipular a interacción con ácidos nucleicos de forma eficiente tanto de moléculas grandes como de moléculas pequenas. Desenvolver sistemas que incorporen elementos moleculares específicos e permitan controlar a vontade a transición entre un estado inactivo e outro activo, en resposta a determinados sinais ou estímulos externos non daniños para sistemas biolóxicos.

Outro obxectivo fundamental é obter unha síntese eficiente de derivados da rexión de recoñecemento de ADN de oncoproteínas. Levarase a cabo a análise das propiedades de interacción co ADN destas miniproteínas así como o estudo das súas propiedades de internalización celular.

Por último, pénsase deseñar biomarcadores fluorescentes capaces de recoñecer secuencias longas de nucleótidos características de diferentes enfermidades.

A maiores, propónse o reto de recoñecer tanto suco maior como o suco menor do ADN con diferentes sistemas.

Para isto, será preciso a posta a punto da metodoloxía de obtención de ditos sistemas, a optimización das súas sínteses e o estudo da interacción destes sistemas co ADN, así como a internalización en células vivas de ser o caso.



**3. CAPÍTULO PRIMEIRO: SÍNTESE QUÍMICA DE ANÁLOGOS  
DE OMOMYC E MYC E RECOÑECIMENTO DO SUCO MAIOR  
DO ADN.**



### 3.1. PRECEDENTES DO GRUPO

O enorme interese da proteína MYC, e a falta de precedentes na súa síntese química, levou a que no grupo de investigación se iniciara un proxecto dirixido a desenvolver métodos sintéticos que permitan obter non só a proteína natural, senón tamén outros derivados funcionais.

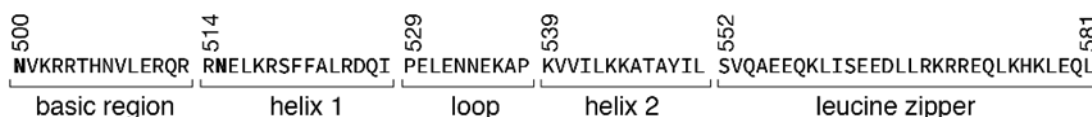


Figura 26. Secuencia da rexión de unión de MYC ao ADN

De feito, fai uns anos no grupo acadouse a primeira síntese química dun derivado da rexión básica de hélice-bucle-hélice da cremalleira de leucinas do dominio de unión ao ADN da proteína MYC, ao que se chamou [AQ]MYC.<sup>84</sup> A estratexia de síntese baseábase na unión por Ligazón Química Nativa (LQN) de tres fragmentos da proteína que se prepararon sinteticamente mediante Síntese de Péptidos en Fase Sólida (SPFS). En traballo preliminar foi importante para coñecer que os residuos de Asn<sup>500</sup> e Asn<sup>515</sup> tenden a formar aspartamidas nos procesos de síntesis en fase sólida, polo que é conveniente emprazalos por outros aminoácidos (Asn<sup>500</sup>→Ala e Asn<sup>515</sup>→Gln).

A pesares de que o rendemento da síntese foi moi modesto, comprobouse que o produto, que contén cistinas en lugar das alaninas Ala<sup>523</sup> e Ala<sup>549</sup> é quen de dimerizar con MAX, e interaccionar con ADN, información moi importante que nos animou a continuar este proxecto.

A maiores, tamén se comprobou que un dos intermedios da reacción, o cal presentaba dous residuos de cistina libres [AQ]MYC(SH), podía recoñecer a secuencia diana de ADN en cooperación con MAX. Cando estas Cys foron tratadas durante unha hora co reactivo de Ellman (DTNB), formouse un composto cun enlace disulfuro entre os residuos Cys<sup>523</sup> e Cys<sup>548</sup>, [AQ]MYC(S-S). O pregamento desta proteína ca ponte disulfuro presentaba unha conformación incapaz de dimerizar con MAX para recoñecer o ADN. Do mesmo xeito, ao expoñer o disulfuro [AQ]MYC(S-S) a un ambiente redutor con TCEP, dito enlace disulfuro era revertido ás cistinas libres e recuperábase a interacción co ADN.

<sup>84</sup> Traballo realizado pola Dr C. Penas Tallón.

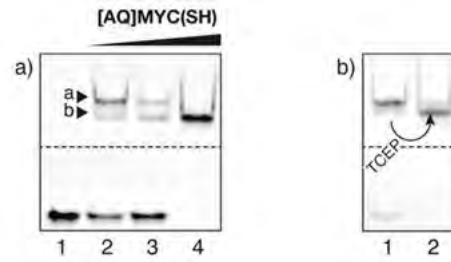


Figura 27. a) Análise electroforética da interacción de [AQ]MYC(SH) co ADN diana en Tris-HCl pH 7.5, 90 mM KCl, 1.8 mM MgCl<sub>2</sub>, 1.8 mM EDTA, 9 % glicerol, 0.11 mg/mL BSA, 2.25 % NP-40 (20 °C). Rúas 1-4: 50 nM E-ADN, 500 nM de MAX e 250, 500 e 1000 nM de [AQ]MYC(SH). A banda denominada a representa o complexo MAX<sub>2</sub>/E-DNA e a denominada b, o complexo [AQ]MYC(SH)/MAX/E-DNA; b) Todas as rúas conteñen 50 nM E-DNA; rúa 1: 500 nM MAX e 1000 nM [AQ]MYC(S-S); rúa 2: despois de engadir TECP, obtense [AQ]MYC(SH) que se une a ADN.





### 3.2. OBXECTIVOS

Como se comentou na *Introdución*, a pesares da enorme valía biomédica que posúe a oncoproteína MYC e o seu inhibidor máis potente OMOMYC,<sup>85</sup> até o de agora, non existen precedentes da súa síntese química, e por iso os métodos de produción baséanse na expresión de fragmentos recombinantes en bacterias.<sup>86</sup> Estas técnicas biolóxicas resultan bastante lentas, presentan problemas ca solubilidade das proteínas expresadas, e adoitan ser moi pouco eficientes. A maiores, non permiten a incorporación sinxela de modificacións non naturais coma fluoróforos. Ata o de agora, as formas máis comúns para a marcaxe de proteínas involucran o uso de residuos cistina e os marcadores adoitan ser outras proteínas fluorescentes (GFP, YFP, R-phycoerythrin)<sup>87</sup>

Todo isto supón un gran inconveniente á hora de estudar as propiedades e a diversidade biolóxica destas proteínas, e fai desexable o desenvolvemento de métodos sintéticos alternativos.

A síntese química de proteínas é versátil e permite introducir modificacións nas secuencias peptídicas: ben sexan fluoróforos, residuos non naturais, ou mutacións en posicións seleccionadas. Obter análogos de OMOMYC e MYC marcados fluorescentemente ou introducir a vontade modificacións de determinados residuos, permitiría estudar en profundidade o xeito de interacción con outras proteínas e co ADN, e avanzar no mecanismo de funcionamento a nivel molecular, incluíndo a súa localización celular. A síntese química tamén abriría a porta ao deseño de novos inhibidores de MYC. Por iso, os obxectivos fundamentais deste capítulo serán:

- A síntese química do dominio de unión a ADN de OMOMYC utilizando o método de 2 LQN .
- A síntese química de derivados do dominio bHLHZip de unión a ADN de OMOMYC marcados fluorescentemente.

<sup>85</sup> a) L. Soucek, M. Helmer-Citterich, A. Sacco, R. Jucker, G. Cesareni, S. Nasi, *Oncogene*, **1998**, *17*, 2463. b) L. Soucek, R. Jucker, L. Panacchia, R. Ricordy, F. Tatò, S. Nasi, *Cancer Res.*, **2002**, *62*, 3507

<sup>86</sup> M. E. Beaulieu, F. O. McDuff, V. Frappier, M. Montagne, J. F. Naud, P. Lavigne, *J. Mol. Recognit.* **2012**, *25*, 414.

<sup>87</sup> a) C. P. Toseland, *J. Chem. Biol.* **2013**, *6*, 85. b) X. Chen, Y.-W. Wu, *Org. Biomol. Chem.* **2016**, *14*, 5417.

- A síntese química de derivados do dominio bHLHZip de MYC polo método de 1 LQN e comparación co método 2 LQN.
- Comprobar as propiedades de dimerización e interacción co ADN das miniproteínas obtidas.
- Averiguar as propiedades de internalización celular dos derivados marcados.



### 3.3. PROPOSTA SINTÉTICA

A síntese química destas proteínas abordarase utilizando reaccións de tipo LQN, xa que son demasiado longas para o uso da SPFS. Para facer a LQN é fundamental fixar unha ruta retrosintética onde se elixan os puntos de ligazón. É dicir, segundo a secuencia de cada proteína, escoller os residuos máis axeitados onde poidan realizarse este tipo de reaccións e que dividan ás proteínas en fragmentos curtos, de entre 30-50 aminoácidos, que poidan ser sintetizados quimicamente de forma sinxela. A desconexión idónea sería en residuos Cys naturais. A maiores, debe terse en conta que aínda que a LQN é susceptible de realizarse con calquera combinación Xaa-Cys, onde Xaa representa calquera dos 20 aminoácidos naturais, a ligazón resulta moi lenta se o Xaa é un aminoácido impedido, especialmente no caso de Ile, Pro ou Val.<sup>88</sup>

Nas proteínas bHLHzip que nos ocupan (OMOMYC e MYC), a zona de unión a ADN consta dunha rexión básica no extremo *N* terminal, dúas hélices alfa unidas por un bucle, na parte central, e unha rexión básica de cremalleira de leucinas no extremo *C* terminal.

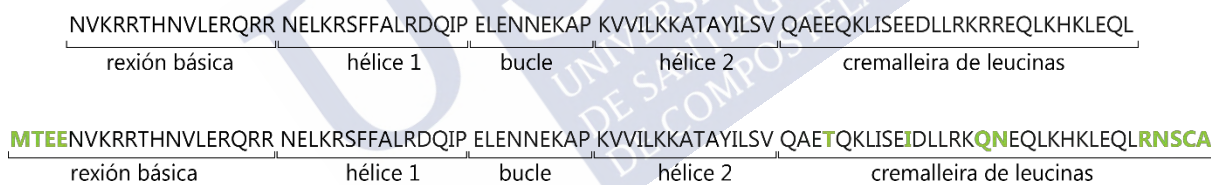


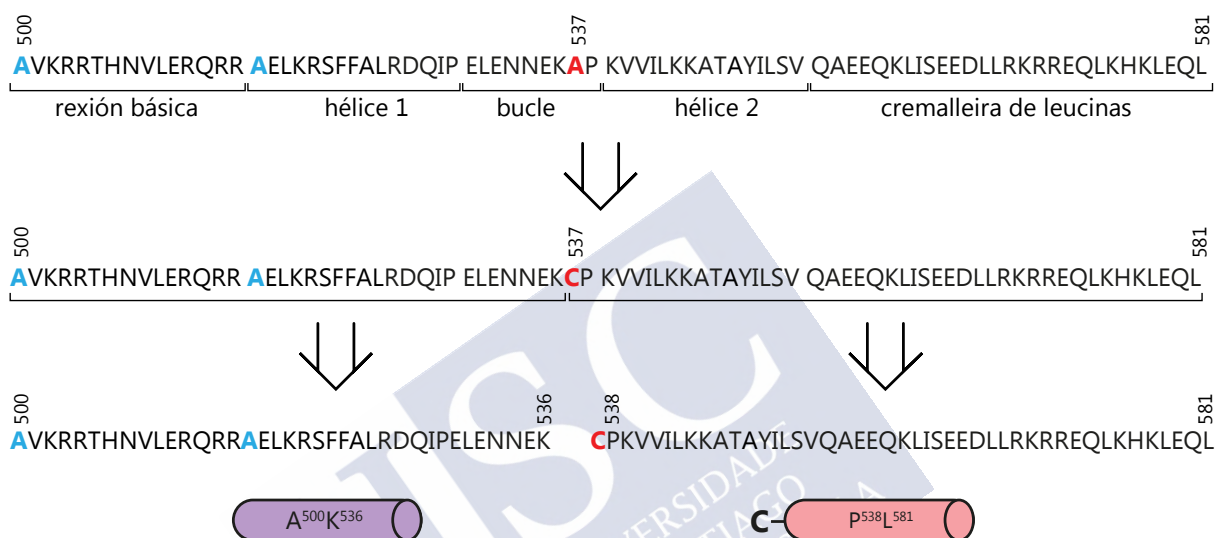
Figura 28. Secuencias da rexión de unión a ADN das proteínas MYC (arriba) e OMOMYC (abaixo). Aparecen resaltados os residuos diferentes entre as secuencias.

Botando unha ollada ás cadeas destes dominios de OMOMYC e MYC, obsérvase que a primeira contén un total de 91 aminoácidos (aa) e, a segunda, 82. A maiores, pode verse que ámbalas dúas proteínas carecen de residuos de cistina necesarios para poder facer as LQN. A única Cys de OMOMYC queda descartada de ser un punto de ancoraxe factible, pois está no extremo. Polo tanto, determinouse situar os puntos de corte en residuos naturais de Ala en lugar de Cys, que despois de realizar as LQN poderían converterse nos residuos de Ala naturais cunha redución por desulfuración.<sup>89</sup>

<sup>88</sup> T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 10068.

<sup>89</sup> L. Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.*, **2001**, *123*, 526.

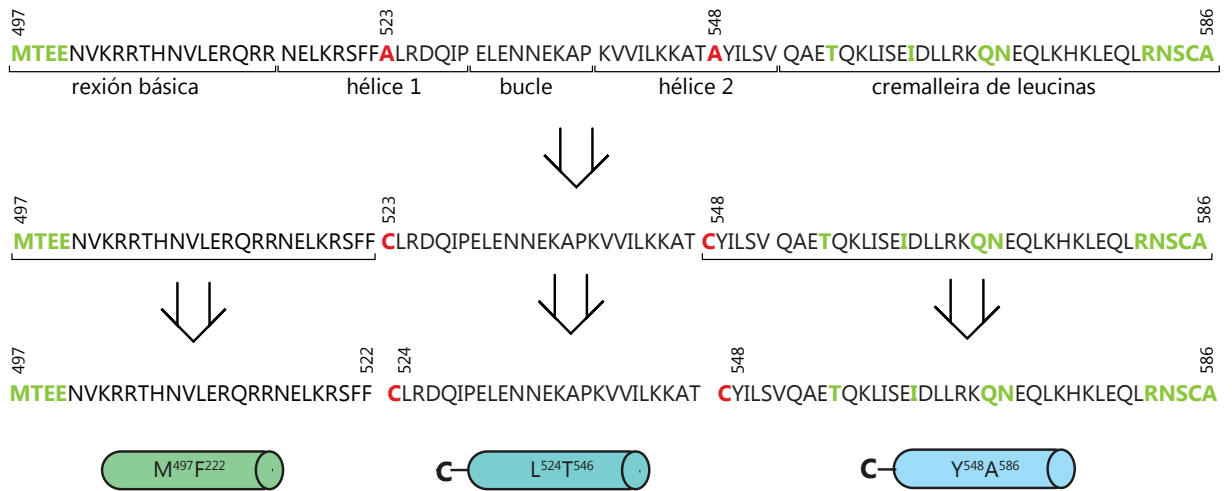
Dado que o grupo xa desenvolvera unha síntese de [AQ]MYC en tres fragmentos, decidiuse comprobar se se podía levar a cabo cunha simple desconexión (dous fragmentos, **2-Fr-LQN**). Para iso elixiuse como punto de corte a Ala<sup>537</sup> do bucle HLH pois está perfectamente situada para romper o péptido en dous fragmentos **A<sup>500</sup>K<sup>536</sup>** e **C<sup>537</sup>L<sup>581</sup>** de 37 e 45 aminoácidos. Trala LQN, unha desulfuración dará a proteína desexada. Nótese que se cambiarán os residuos Asn<sup>500</sup>→Ala e Asn<sup>515</sup>→Gln debido aos problemas de aspartimidación comentados anteriormente.



Esquema 1. Esquema retrosintético para a obtención de [AQ]MYC mediante a estratexia de dous fragmentos (2-Fr-LQN).

En canto a OMOMYC, preferiuse ensaiar primeiro a estratexia baseada en tres fragmentos realizando dúas LQN e, finalmente, unha desulfuración para obter o produto natural. As rexións de desconexión elixidas sitúanse nas Ala<sup>523</sup> e Ala<sup>548</sup> que xeran tres fragmentos: **M<sup>497</sup>F<sup>522</sup>**, **C<sup>523</sup>T<sup>546</sup>** e **C<sup>547</sup>A<sup>586</sup>** de 27, 25 e 39 residuos respectivamente.

Por último, os derivados peptídicos marcados realizaranse mediante as estratexias correspondentes de tres fragmentos con LQN (3 Fr-LQN), pois son mais convenientes para introducir un fluoróforo, xa que se sintetizarían fragmentos pequenos.



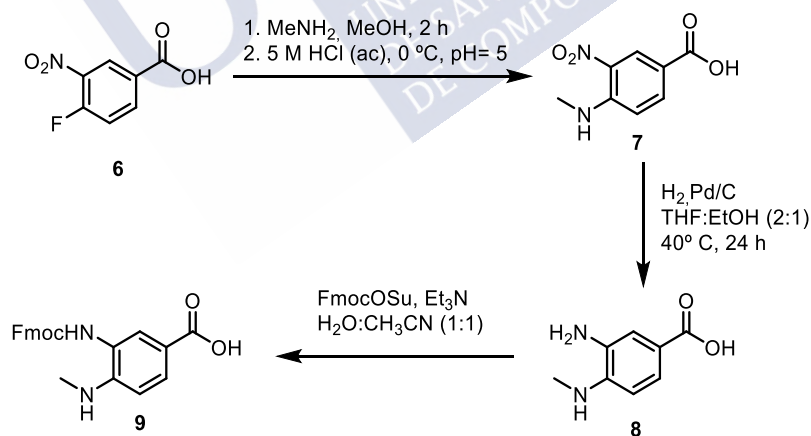
Esquema 2. Esquema retrosintético para a obtención deOMOMYC mediante a estratexia de tres fragmentos. A numeración da secuencia mantén a de MYC para maior coherencia



### 3.4. SÍNTESE DO AMINOÁCIDO FMOC-DBZ (ÁCIDO 3-FMOC-4-DIAMINOBENZOICO)

Para que unha reacción de LQN teña lugar, é preciso que se forme un tioéster no extremo C-terminal do fragmento *N*-terminal que participa na reacción, como se comentou na introdución (ver 1.3.2.2). Este tioéster fórmase *in situ* na reacción e para iso, débese introducir na secuencia peptídica en forma dun precursor deste tioéster. No caso deste traballo, empregouse o método das *N*-metilbencimidazolinonas (Nbz) para obter os tioésteres.<sup>39</sup> O Nbz fórmase en fase sólida a partir dun derivado do ácido benzoico, denominado Dbz. Polo tanto, o primeiro paso realizado foi a síntese do aminoácido Fmoc-Dbz-OH para ancoralo á resina.

O ácido **9** foi obtido seguindo un procedemento xa escrito na bibliografía.<sup>90</sup> A ruta de síntese do Dbz, tal como se amosa no **Esquema 3**, comeza ca substitución electrófila aromática do flúor no ácido 4-fluoro-3-nitrobenzoico **6** con metilamina e precipitar o composto **7** a pH 5. O grupo nitro do produto obtido foi reducido cunha hidroxenación catalizada por Pd/C. Por último, e xa que este aminoácido se vai usar para SPFS co protocolo Fmoc/*t*-Bu, protexeuse a amina primaria con Fmoc-hidroxisuccinimida (Fmoc-OSu) acadándose o aminoácido desexado **9** cun alto rendemento global do 69 %.



Esquema 3. Ruta sintética para obter o aminoácido Fmoc-Dbz-OH

<sup>90</sup> N. C. Kuzniweski, J. Gertsch, M. Wartmann, H. K. Altmann *Org. Lett.* **2008**, *10*, 1183.

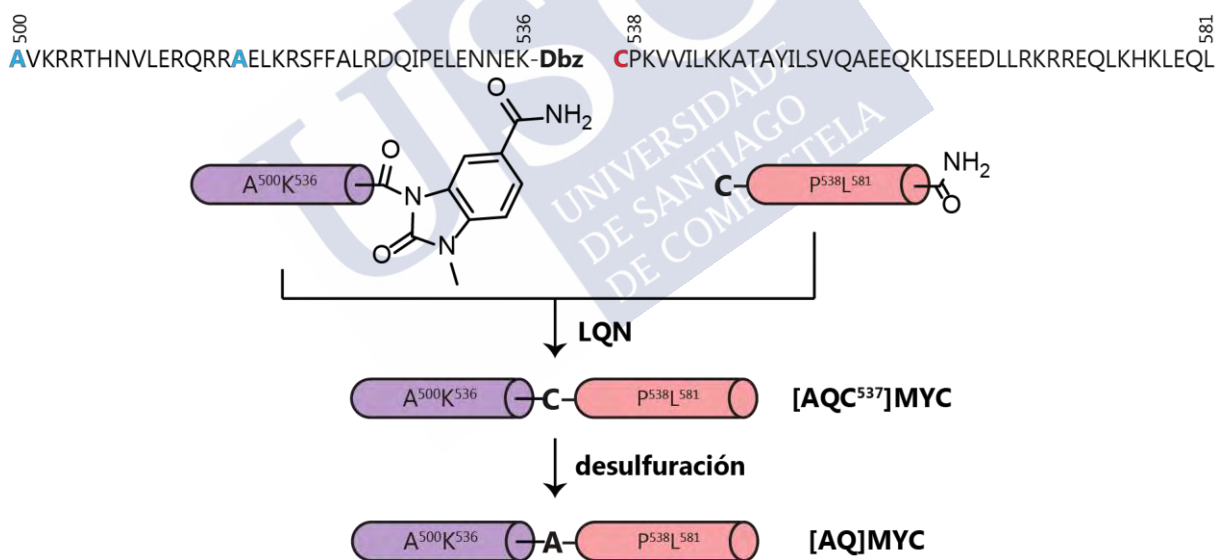
### 3.5. SÍNTESE DE DERIVADOS DE MYC MEDIANTE ESTRATEXIA 2-FR-LQN

Vendo a secuencia de [AQ]MYC sintetizada previamente no grupo, e tal como se indicou na proposta sintética, pode deducirse que un dos sitios de desconexión viables para a súa síntese é a Ala<sup>537</sup> situada na rexión do xiro. Este corte divide á proteína en dous fragmentos de lonxitude comparable, viables de someterse á ligazón química nativa: A<sup>500</sup>K<sup>536</sup>-Dbz (37 residuos + Dbz) e C<sup>537</sup>L<sup>581</sup> (45 residuos).



Figura 29. Secuencia de aminoácidos da rexión de unión de MYC ao ADN. En azul as mutacións introducidas con respecto á proteína natural. En vermello os sitios de desconexión que se usarán para a síntese 2 FrLQN.

É por isto que se decidiu realizar a síntese de [AQ]MYC a partir de dous fragmentos, cunha única reacción de LQN a través da A<sup>537</sup> e comparar as vantaxes/desvantaxes con respecto á síntese de tres fragmentos.



Esquema 4. Esquema da ruta sintética total para obter [AQ]MYC a través dunha única LQN.

#### 3.5.1. Síntese dos fragmentos

A preparación dos fragmentos realizouse en fase sólida<sup>91</sup> asistida por microondas, usando o método Fmoc/*t*-Bu. Como soporte sólido elixiuse a resina Fmoc-Rink Amida ChemMatrix, a cal libera como amida o extremo C-terminal dos péptidos unha vez desancorados da resina.

<sup>91</sup> I. Coin, M. Beyermann, M. Bienert, *Nat. Protoc.* **2007**, *2*, 3247.

Na síntese do fragmento N-terminal **A<sup>500</sup>K<sup>536</sup>-Dbz** o primeiro paso é o acoplamento manual do precursor do tioéster necesario para a LQN, o chamado Dbz. Para isto, desprotéxese a amina da resina (protexida con Fmoc) con 20 % de piperidina en DMF durante 20 min. De seguido, acóplase o aminoácido derivado do *N*-metildiaminobenzoico usando HATU como activador do ácido, DIEA como base, e DMF como disolvente. Unha vez acoplado o primeiro residuo, complétase a síntese do resto da secuencia de forma automática, co sintetizador de péptidos. Debido a que o grupo protector Fmoc é máis lábil, o último aminoácido acoplado estará protexido con Boc, para evitar que a amina *N*-terminal quede libre e susceptible de reaccionar no seguinte paso de formación de Nbz.

Despois de precipitar unha pequena cantidade de resina e comprobar por HPLC-EM que a secuencia peptídica está correcta, fórmase o derivado activado Nbz (*N*-acilmetilbencimidazolinona) para que poida ter lugar a reacción de LQN. A transformación do Dbz en Nbz transcorre por acilación da *N*-metilamina e formación dunha *o*-aminoanilida,<sup>92</sup> que de seguido por tratamento con base provoca o ataque intramolecular da anilida cara o carbamato e a subsecuente formación dunha *N*-acilurea aromática (Esquema 5)

---

<sup>92</sup> J.B. Blanco-Canosa, J.B. Dawson, *Angew.Chem.Int.Ed.* **2008**, *47*, 6851.





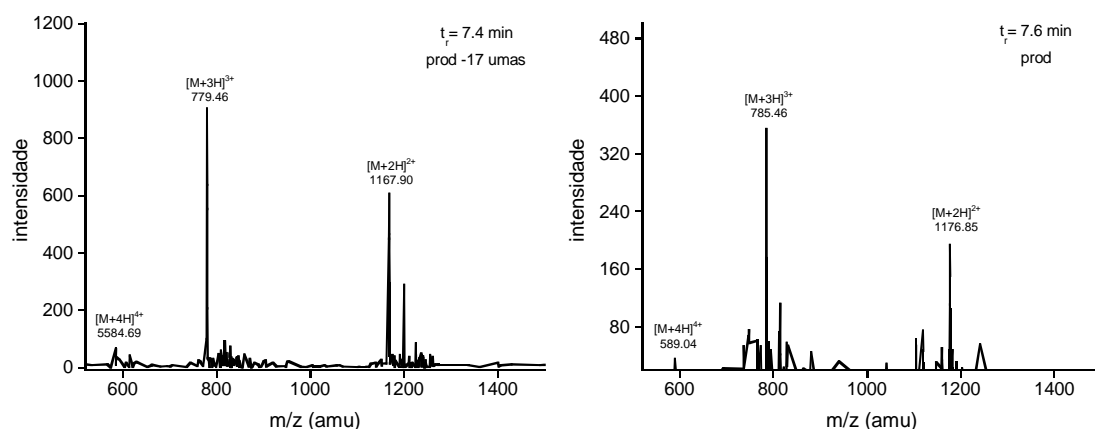
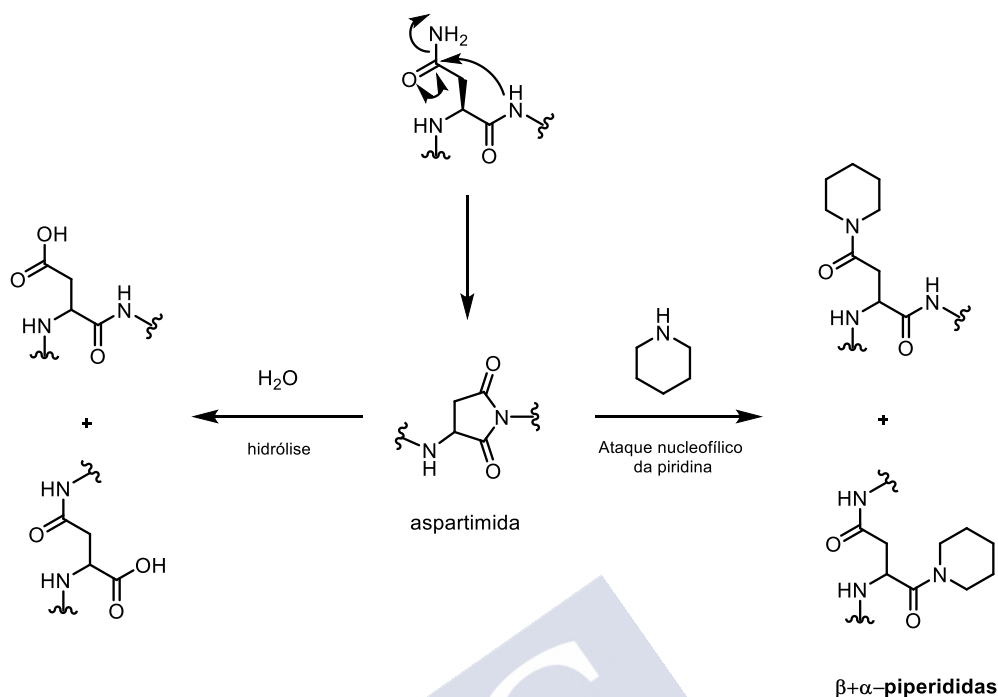


Figura 30. Arriba: cromatograma dos 19 primeiros aminoácidos do fragmento  $A^{500}K^{536}$ -Dbz, a 222 nm. Abaixo dereita: espectro de masas para o pico do produto a 7.6 min. Abaixo esquerda: espectro masas obtido para o pico a 7.4 min que se corresponde co subproduto de aspartimidación.

Esta diferenza de masa atribúese a unha reacción lateral experimentada polos residuos de Asn ou Asp, denominada desamidación ou aspartimidación.<sup>94</sup> Esta reacción secundaria ten lugar no paso de  $N_{\alpha}$ -desprotección con base do grupo Fmoc. Os subprodutos orixinados neste paso caracterízanse por ter un tempo de retención moi semellante ao produto principal da síntese, nos mellores casos o tempo de retención é lixeiramente diferente e pode separarse esta impureza na purificación, mais non é o habitual. Porén na maioría das ocasións, estes subprodutos detéctanse como unha anomalía no espectro de masas do produto principal xa que as masas do produto desexado se superpoñen cas masas das impurezas.

<sup>94</sup> R. Subiros-Funosas, A. El-Faham, F. Albericio, *Tetrahedron* **2011**, 67, 8595.



Esquema 6. Reacción lateral de aspartimidación e subproduto da desamidación

A asistencia á síntese de péptidos con microondas xera moitas vantaxes como o incremento das velocidades de reacción e os crús dos produtos son moito máis limpos.<sup>95</sup> A irradiación de microondas adoita aplicarse tamén no paso da  $N\alpha$ -desprotección do Fmoc. Nestes casos os tempos de reacción pasan de 20 min, para as desproteccións en condicións estándar a temperatura ambiente, a 0.5 ou 3 min usando as microondas a 50-90 °C.<sup>96</sup> Mais hai estudos que corroboran que o uso de radiación incrementa até nun 10 % as reaccións laterais de aspartimidación naqueles péptidos susceptibles de sufrir estas transformacións, é dicir, os que conteñen residuos de Asp ou Asn.<sup>97</sup>

Está comprobado que o uso de HOBT como aditivo na disolución de desprotección reduce a reacción lateral de desamidación.<sup>98</sup> Polo tanto, probouse de novo a síntese do fragmento **A<sup>500</sup>K<sup>536</sup>-Dbz** pero esta vez usando como disolución de eliminación do grupo Fmoc unha mestura de piperidina ao 20 % en DMF e 0.1 M de HOBT.

<sup>95</sup> M. Erdeyli, A. Gogoll, *Synthesis*, **2002**, *11*, 1592.

<sup>96</sup> S. Park, S. Gunasekera, T. Aboye, U. Goransson, *Int. J. Pept. Res. Ther.* **2010**, *16*, 167.

<sup>97</sup> a) Palasek, S.A.; Cox, Z.J.; Collins, J.M. *J. Pept. Sci.* **2007**, *13*, 143. b) S. L. Pedersen, *Chem. Soc. Rev.* **2012**, *14*, 1826.

<sup>98</sup> J. L. Lauer, C. G. Fields, G. B. Fields, *Lett. Pept. Sci.* **1994**, *1*, 197.

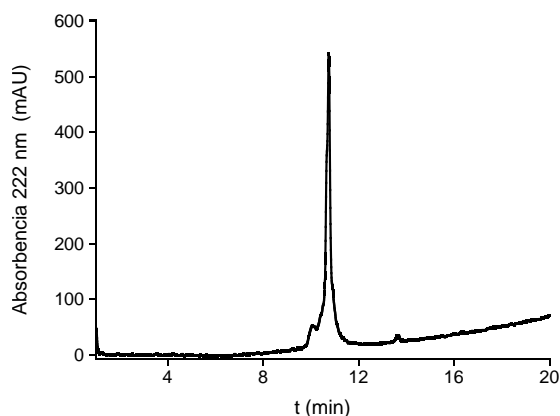


Figura 31. Cromatograma do fragmento sintetizado con 0.1 M de HOBT

Como se observa na Figura 31. Cromatograma do fragmento sintetizado con 0.1 M de HOBT, a impureza de desamidación diminuíu ao engadir HOBT como aditivo mais non se eliminou por completo xa que puideron observarse as masas correspondentes á formación de dúas piperididas (+134). A maiores, tamén se comprobou que a reacción de aspartimidación parece que estaba favorecida dalgún xeito polo aminoácido Dbz. Sintetizouse un péptido que contiña os cinco primeiros aminoácidos do fragmento *N*-terminal pero sen o Dbz, con secuencia Fmoc-ENNEK cunha masa esperada era 853.9, e atopouse unha masa despois do seu análise por HPLC-EM de 854.8 (ver Figura 32). Este fragmento non presentaba ningunha clase de reacción lateral.

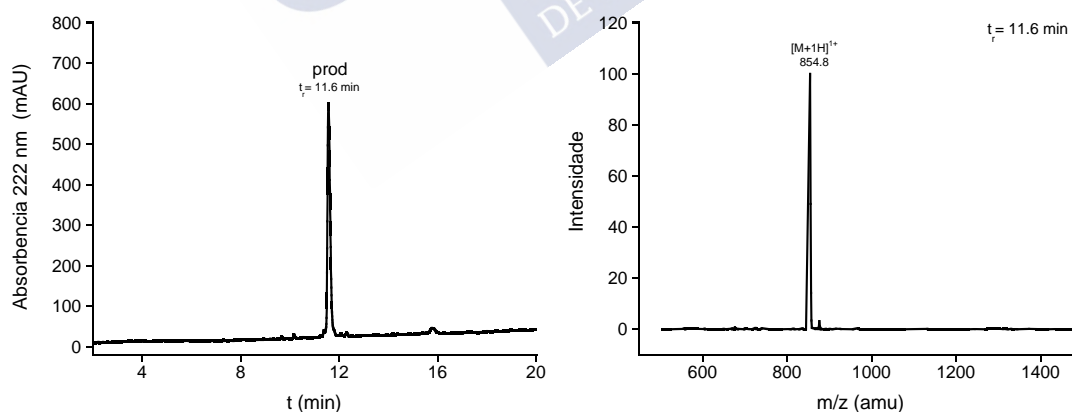


Figura 32. Espectro de masas dos cinco primeiros aminoácidos sen o Dbz.

Por outra banda, sintetizouse co Dbz o mesmo fragmento Fmoc-LENNEK-Dbz. Neste caso si se observou unha impureza de desamidación a maiores do pico do produto desexado (ver Figura 33). Isto indica que o Dbz ten algunha influencia na desamidación.

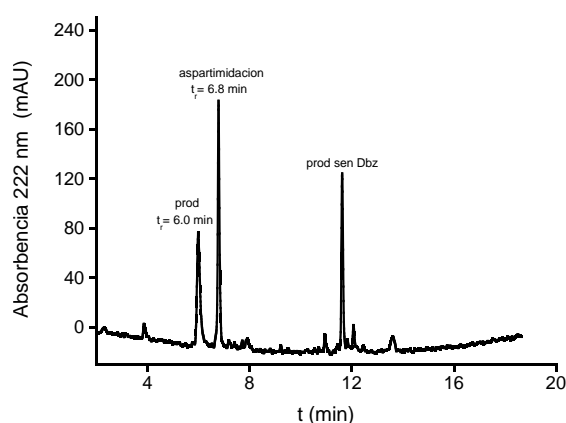


Figura 33. Cromatograma de Fmoc-LENNEK-Dbz. Vese pico de impureza de desamidación a 6.8 min.

Unha búsqueda na bibliografía amosou que cambiar a base piridina por piperazina reduce a formación das aspartimidadas.<sup>99</sup> A maiores de reducir esta reacción lateral, a piperazina é un produto comercial máis barato e fácil de conseguir que a piperidina xa que este último é un precursor da droga fenilciclohexilpiperidina (comunmente coñecida como *Po de Anxo*) e cuxo uso está moi restrinxido e regulado. Polo tanto, decidiuse cambiar a disolución de desprotección por piperazina ao 5 % (m/v) en DMF mais 0.1 M de HOBT. Observándose, esta vez, a síntese do produto desexado satisfactoriamente, sen subprodutos como se observa na **Figura 34**.

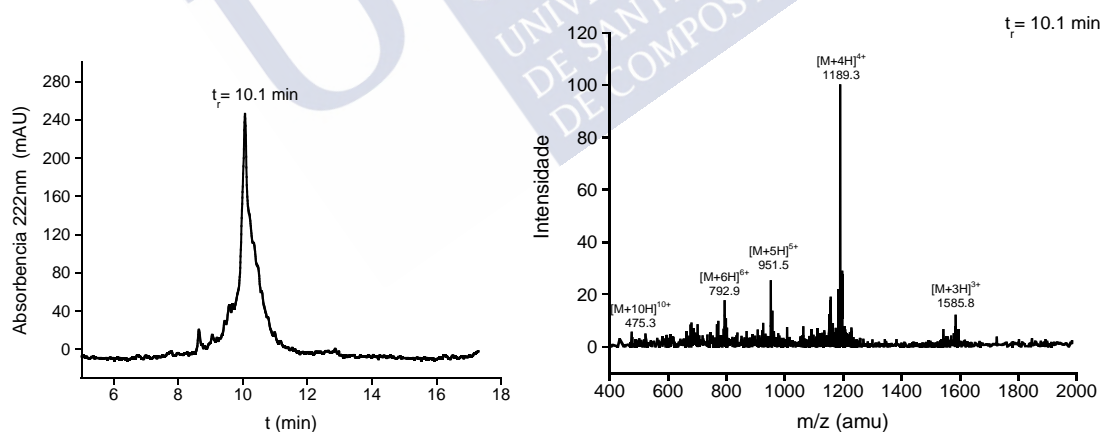


Figura 34. Esquerda cromatograma do cru da síntese do fragmento A<sup>500</sup>K<sup>536</sup>-Dbz a 222 nm. Dereita: espectro de masas do pico do produto a tempo de retención 10.1 min.

Por último, o residuo Dbz transformouse en Nbz seguindo os pasos explicados anteriormente, desancórase da resina e purifícase por cromatografía en fase reversa.

<sup>99</sup> J. D. Wade, M. N. Mathieu, M. Macris, G. W. Tregear, *Let. Pept. Sci.* **2000**, 7, 107.

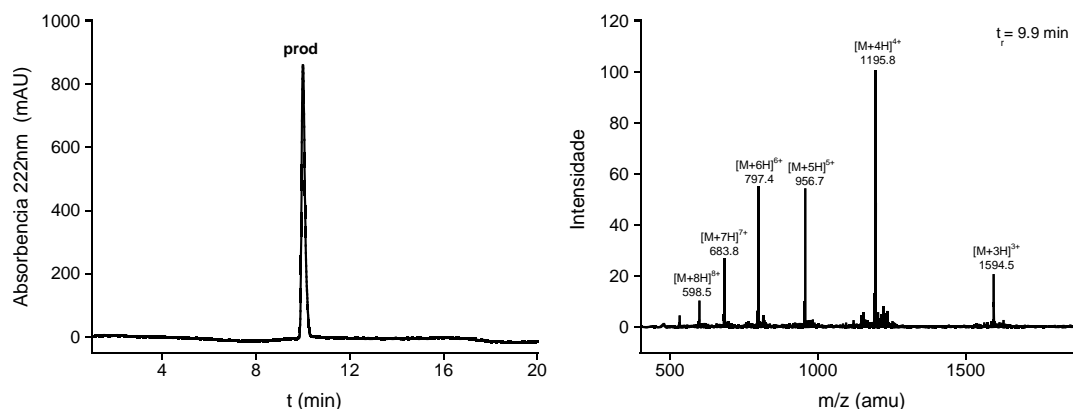


Figura 35. Esquerda: Cromatograma do produto puro  $A^{500}K^{536}$ -Nbz a 222 nm. Dereita: espectro de masas do produto puro onde se comproba que non ten desamidación

Por outro lado, para a síntese do fragmento C-terminal  $C^{537}L^{581}$  seguiu-se o procedemento estándar de síntese en fase sólida co método Fmoc.

Obténdose de maneira sinxela o fragmento desexado:

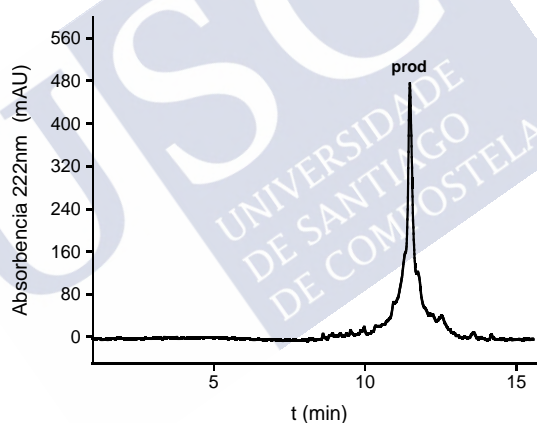
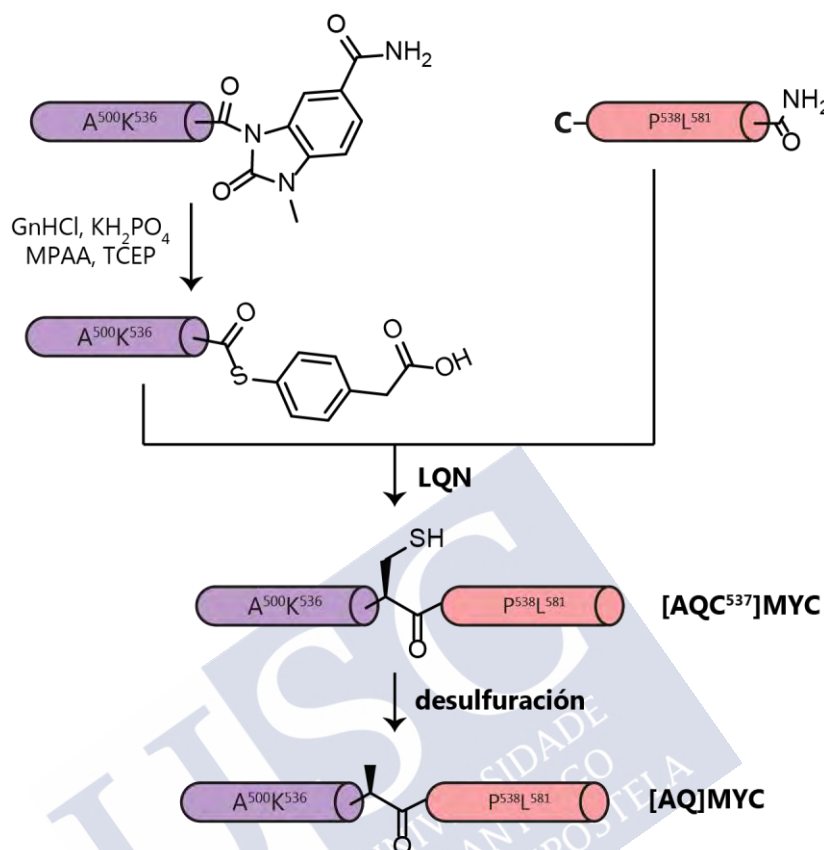


Figura 36. Cromatograma do cru do fragmento C-terminal  $C^{537}L^{581}$ .

### 3.5.2. Unión dos fragmentos por LQN e obtención de [AQ]MYC

A unión dos dous fragmentos sintetizados levouse a cabo a través dunha reacción de Ligazón Química Nativa (LQN). Para iso, atopouse que as condicións óptimas consistían en disolver o fragmento  $A^{500}K^{536}$ -Nbz en tampón desgasificado 6 M de  $Gn \cdot HCl$ , 200 mM de  $K_2HPO_4$ , 100 mM de ácido mercatofenilacético (MPAA), e 40 mM de TCEP co pH axustado a 7.5, que contiña 1.5 equiv. do outro fragmento,  $C^{537}L^{581}$ . O MPAA reacciona coa bencimidazolinona (Nbz) a través dunha reacción de adición-eliminación para formar un ariltioéster máis reactivo. Este tioéster C-terminal é atacado pola  $Cys^{537}$  libre do outro fragmento peptídico, producíndose unha transtioesterificación. Unha vez formado o enlace de

tipo tioéster que une os dous fragmentos, prodúcese a reordenación do acilo dende o S da cistina até o N para dar un enlace tipo amida.



Esquema 7. Ruta sintética para obter [AQ]MYC. Móstrase o intermedio de reacción da LQN e o proceso de desulfuración que se comentan neste apartado.

A mestura de reacción deixouse axitando durante 4.5 h a temperatura ambiente cando se recolleu unha alícuota do cru e se comprobou por HPLC-EM o avance da reacción. Como vemos na parte esquerda da **Figura 37**, a reacción completouse posto que o reactivo limitante A<sup>500</sup>K<sup>536</sup>-Nbz consumíuse por completo. Pasouse pois a parar a reacción engadindo o triplo de volume de tampón desnaturizante de Gn·HCl e conxelala en N<sub>2</sub> líquido.

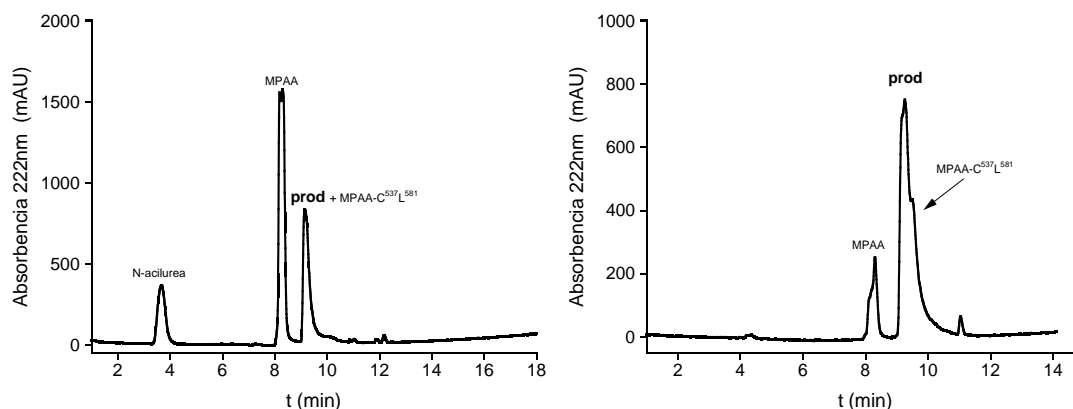


Figura 37. Esquerda: cromatograma do cru despois de 4.5 h de reacción antes de ser filtrado con amicons. Dereita: cromatograma despois de filtrarse o cru de reacción con amicons

No cromatograma do cru de reacción a 222 nm pódese apreciar un lombo no pico do produto con tempo de retención aproximado de 9.3 min. Este lombo identificouse como o aducto disulfuro formado pola reacción do tiol MPAA ca cistina terminal do fragmento en exceso  $C^{537}L^{581}$ . Denominouse  $MPAA-C^{537}L^{581}$ .

Finalmente, o produto desexado purificouse por cromatografía de líquidos de fase revertida. Previa purificación da mestura, engadiuse unha punta de espátula de TCEP para desfacer a ponte disulfuro do  $MPAA-C^{537}L^{581}$ . Isto, unido a un estudo do gradiente da fase móbil, fixo que os produtos se separasen perfectamente ao purificar a reacción nun HPLC semipreparativo como se mostra na **Figura 38**, obténdose cun excelente rendemento do 55 % o péptido desexado  $AQC^{537}[MYC]$ , con secuencia: **AVKRRTHNVL ERQRRQELKR SFFALRDQIP ELENNEKCPK VVILKKATAY ILSVQAEQK LISEEDLLRK RREQLKHKLEQL-NH<sub>2</sub>**

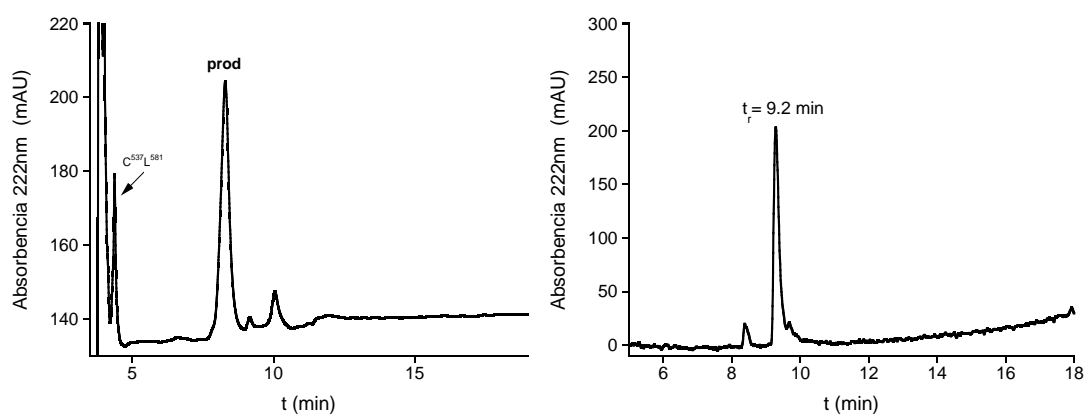


Figura 38. Esquerda: cru de reacción con gradiente do 25 % ao 45 % de MeCN + 0.1 % TFA en 30 min. Dereita: Cromatograma a 222 nm do produto puro  $[AQC^{537}] MYC$ , obtido polo método 2FrLQN



Por último, a miniproteína obxectivo [AQ] MYC obtívose de forma doada por tratamento de [AQC<sup>537</sup>] MYC cunha mestura de EtSH, TCEP e VA-044 como iniciador de radicais en tampón fosfato a pH 7. A mestura axitou a 37 °C durante 5 h e o produto purificouse por HPLC de fase revertida unha vez completada a reacción. Despois de liofilizar os produtos puros, obtívose un sólido branco que resultou ser a miniproteína desulfurada [AQ]MYC, con secuencia: AVKRRTHNVL ERQRRQELKR SFFALRDQIP ELENNEKAPK VVILKKATAY ILSVQAEQK LISEEDLLRK RREQLKHKLEQL-NH<sub>2</sub>. O rendemento para esta etapa foi dun 76 % (Rendemento global do acoplamento e desulfuración dun 41-42 %).

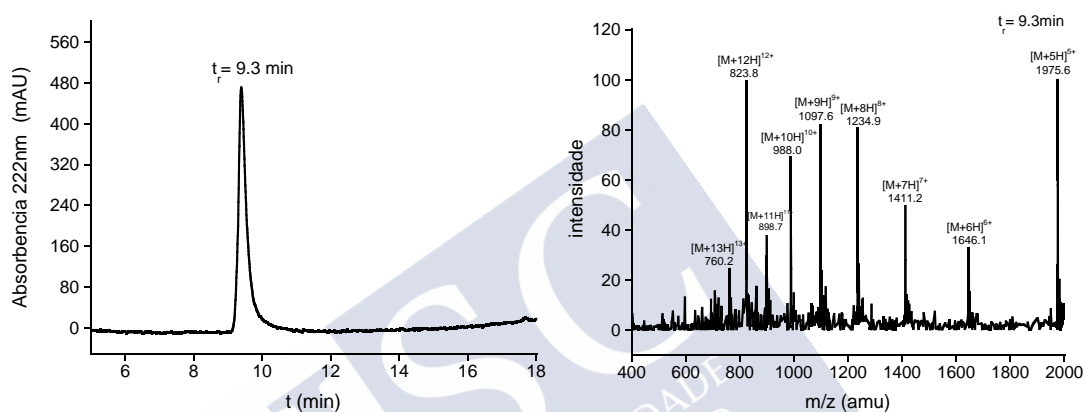


Figura 39. Esquerda: Cromatograma do produto puro [AQ] MYC, obtido trala desulfuración. Dereita: espectro de masas do produto

Namentres, a obtención de [AQ]MYC mediante o método de tres fragmentos levado a cabo anteriormente no laboratorio deu como rendemento global arredor dun 4 %. A etapa limitante deste proceso resultou ser a primeira das ligazóns na que só se illou un 20 % do produto desexado. Hai que ter en conta, ademais que este método de síntese comprende unha reacción a maiores, ca súa respectiva purificación por HPLC. Estes dous factores tiveron un forte impacto que se ve reflexado no rendemento global baixo.

Por outra banda, para o método de 2Fr-LQN optimizáronse arduamente as condicións de purificación. Introduciuse a filtración con amicóns previa á separación cromatográfica e axustáronse os gradientes e columnas usadas para a purificación. Isto provocou a diminución dos tempos de retención mantendo unha boa separación. Ao diminuír o tempo que o péptido pasa dentro da columna recupérase máis cantidade deste. Polo tanto, aumentouse notablemente o rendemento global (41-42%).

### 3.6. SÍNTESE DE DERIVADOS DE OMOMYC

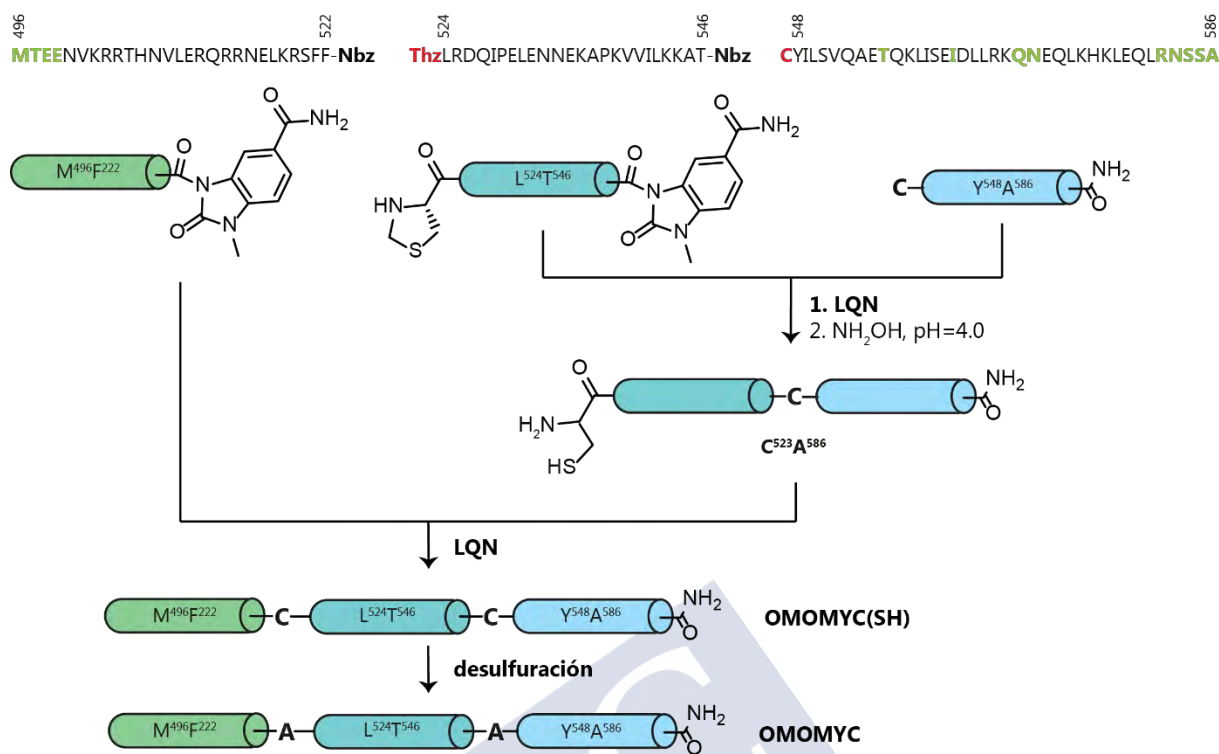
OMOMYC é unha proteína obtida por métodos recombinantes e que consta da secuencia indicada na **Figura 40**. En esta Figura márcanse ademais en azul os dous residuos de asparaxina que se mutaban na síntese de tres fragmentos de MYC levada a cabo polo grupo.

Como se razoou no apartado *de proposta sintética* deste capítulo, a falta de residuos Cys na secuencia desta proteína fai que os aminoácidos máis axeitados para actuar de sitios de ligazón sexan as Ala 523 e 547.



Figura 40. Secuencia da rexión de unión a ADN (bHLHZip) de OMOMYC. En vermello os puntos de desconexión para as LQN.

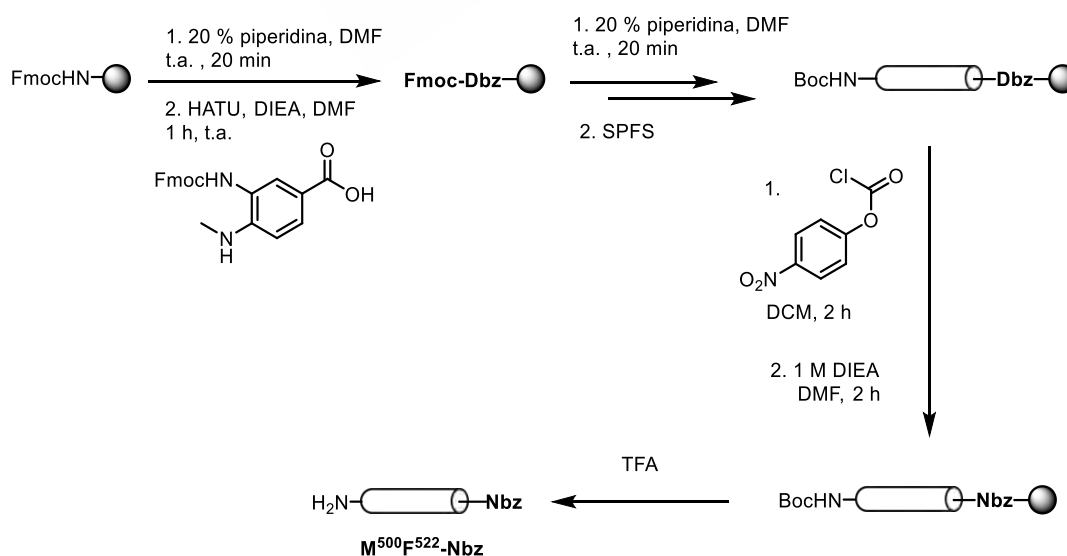
Dado que neste proxecto de tese usarase un sintetizador de péptidos asistido por microondas (non usado para sínteses anteriores do grupo) e, a maiores, a secuencia contén catro residuos iniciais MTEE diferentes da de MYC, decidiuse tentar a síntese deste fragmento para OMOMYC sen mutar ningunha das devanditas asparaxinas. Como consecuencia, planeouse unha análise retrosintética baseada en dúas LQN que implica a conexión secuencial de tres péptidos: o fragmento *N*-terminal  $M^{500}F^{522}$ -Nbz (27 residuos + Nbz), o fragmento intermedio, que se designará por  $Thz^{523}T^{546}$ -Nbz (25 residuos + Nbz) e o *C*-terminal  $C^{547}A^{586}$  (39 residuos)



Esquema 8. Esquema sintético total para a obtenção de OMOMYC.

### 3.6.1. Síntese dos fragmentos.

O fragmento *N*-terminal, ao que se denominou **M<sup>500</sup>F<sup>522</sup>-Nbz** consta de 27 aminoácidos mais o precursor do tioéster Nbz.



Esquema 9. Ruta sintética para a obtenção do fragmento *N*-terminal de OMOMYC, M<sup>500</sup>F<sup>522</sup>-Nbz.

O primeiro paso desta ruta é introducir a anilida non natural xa coñecida como Dbz. Este paso realízase manualmente sobre a resina Rink-Amide ChemMatrix usando a sal de uronio HATU como axente activante. O resto dos aminoácidos acopláronse co sintetizador automático *Liberty Lite*. O último aminoácido introdúcese protexido como Boc para evitar a desprotección do Fmoc e calquera reacción lateral no paso seguinte, de acilación do Dbz. Por último, unha vez toda a secuencia está ancorada ao soporte sólido, transfórmase o Dbz en Nbz mediante a acilación da anilida con *para*-nitroclorofornato, como se explicou anteriormente. Despois de comprobar cunha pequena alícuota da resina que se obtivo o produto desexado, pásase a desancorar a secuencia peptídica da resina por tratamento con TFA.

O cru de reacción amósase a continuación onde se pode ver o pico de 7 min maioritario atribúese ao fragmento desexado, mentres que o pico con tempo de retención 6 min posúe unha masa 147 umas menor que o produto requerido. Esta perda de masa atribúese a que non se acoplou o aminoácido Fmoc-Dbz. Mais non se observa en ningún caso indicio de desamidación.

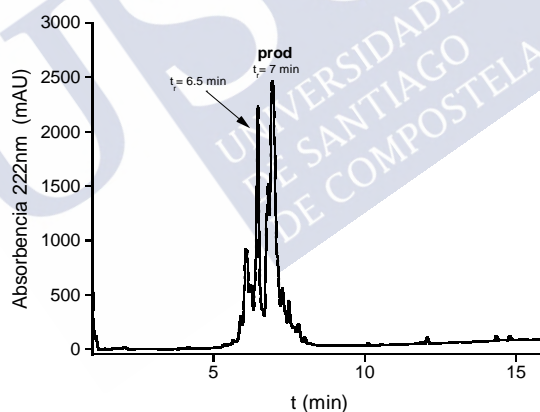


Figura 41. Cromatograma do cru da síntese do fragmento de OMOMYC M<sup>500</sup>F<sup>525</sup>-Dbz.

Cando se procedeu á transformación do Dbz na *N*-acil-*N*-methylacilurea (Nbz) e posterior análise por HPLC-EM dunha alícuota de resina, non apareceu o pico do produto desexado e, en vez deste, xurdiu un pico maioritario con tempo de retención 8 min dunhas 29 unidades de masa superior ao produto buscado. A maiores, observouse o pico correspondente ao péptido acetilado pero sen o Dbz a 7.4 min.

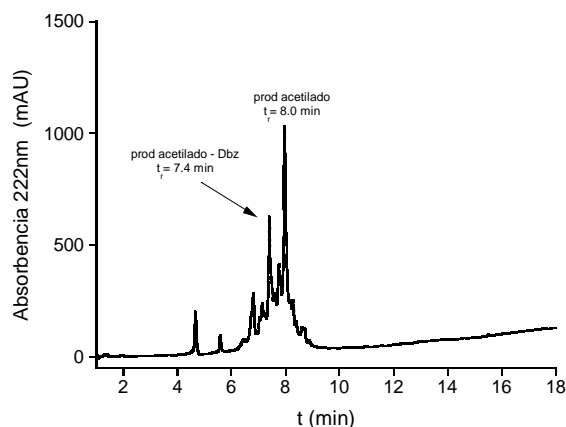


Figura 42. Cromatograma despois da acilación con cloroformato

Unha hipótese que explicaría este fenómeno pode ser que o *para*-nitrocloroformato está acilando non só o Dbz, senón que tamén pode estar atacando ao xofre da metionina *N*-terminal para ciclar intramolecularmente ca amina da mesma.

Con este resultado, e tendo en conta que a secuencia MTEE da oncoproteína deriva da súa transfección en bacterias, non é unha secuencia propia da proteína nin indispensable para a súa síntese química,<sup>100</sup> mais que non se apreciou desamidación neste fragmento, determinouse eliminala e sintetizar o fragmento sen estes catro aminoácidos iniciais. Polo tanto, a nova secuencia a sintetizar sería NVKRRTHNVLERQRRNELKRSFF-Nbz e que se denominará de aquí en diante N<sup>500</sup>F<sup>522</sup>-Nbz .

O procedemento é exactamente o mesmo que o exposto no **Esquema 8**. Polo tanto, volveuse crecer en resina a secuencia peptídica correspondente mais cando se analizou por HPLC-EM o cru de reacción distinguíronse dous picos practicamente inseparables. O primeiro maioritario, con tempo de retención 10.0 min, ten unha masa igual ao produto de reacción menos 17 unidades. O pico a 10.3 min pertence ao produto buscado N<sup>500</sup>F<sup>522</sup>-Dbz, aínda que o espectro de masas presenta unha impureza de masa +116. Estas diferenzas nas masas moleculares correspóndense cos procesos de desamidación. No caso do pico a 10.0 min, produciuse unha aspartimidación con perda de NH<sub>3</sub> e no segundo caso a masa da impureza cadra ca formación de dúas piperididas e unha aspartimida.

<sup>100</sup>L. Soucek, M. Helmer-Citterich, A. Sacco, R. Jucker, G. Cesareni, S. Nasi, *Oncogene* **1998**, *17*, 2463

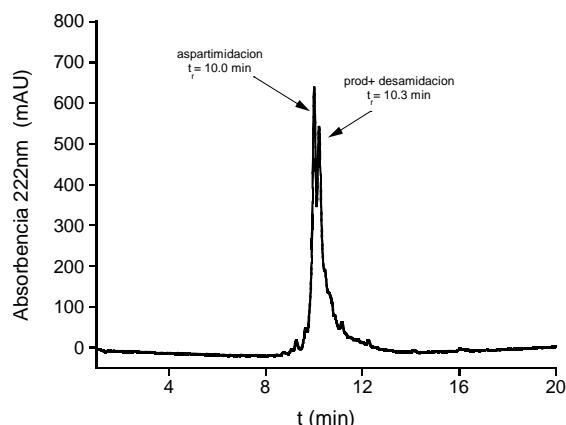


Figura 43. Cromatograma do cru de reacción obtido para a síntese de  $N^{500}F^{522}$ -Dbz

Coa inesperada desamidación obtida, e tendo en conta que no anterior apartado se solventou cambiando a disolución de desprotección, probouse aquí tamén a sintetizar de novo o fragmento pero usando a metodoloxía optimizada anteriormente para estes casos (piperazina ao 5 % en g/L en DMF mais 0.1 M de HOBT). Desafortunadamente, nesta ocasión a aplicación desta metodoloxía para eliminar o Fmoc non tivo efecto esperado. Despois desta terceira tentativa de síntese en SPFS do fragmento  $N^{500}F^{522}$ -Dbz volveu a analizarse por HPLC-EM e volveu observarse un cromatograma do cru parecido ao anterior cun pico de desamidación inseparable do produto desexado, cuxo pico do produto tratou de purificarse por fase reversa sen éxito.

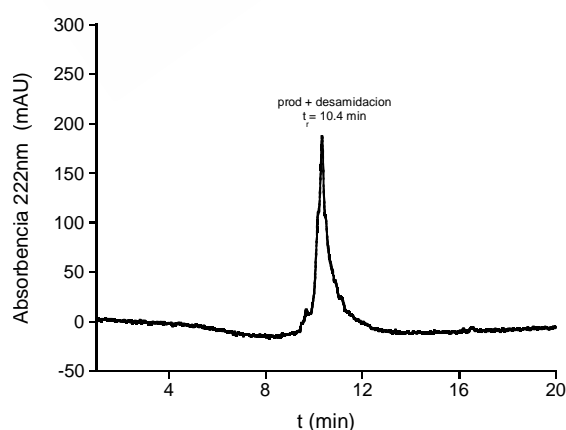


Figura 44. Cromatograma do produto obtido despois da purificación, para sintetizar  $N^{500}F^{522}$ -Dbz cambiando o cóctel no paso de desprotección.

Xa que estas estratexias non solucionaron os problemas atopados, determinouse realizar o fragmento  $N$ -terminal mutando as asparaxinas das posicións 500 e 516 ao igual que se fixera

anteriormente no grupo, quedando a secuencia obxectivo final como: AVKRRTHNVLERQRRQELKRSFF (**A<sup>500</sup>F<sup>522</sup>-Dbz**). Agora si, seguindo a ruta marcada para o resto das tentativas, esta vez obtívose o fragmento sen detectar desamidacións. Tanto **A<sup>500</sup>F<sup>522</sup>-Dbz** como **A<sup>500</sup>F<sup>522</sup>-Nbz** se sintetizaron perfectamente. Como último paso, desancorouse da resina **A<sup>500</sup>F<sup>522</sup>-Nbz**, e purificouse, obtendo o fragmento cun rendemento final do 17 % aproximadamente.

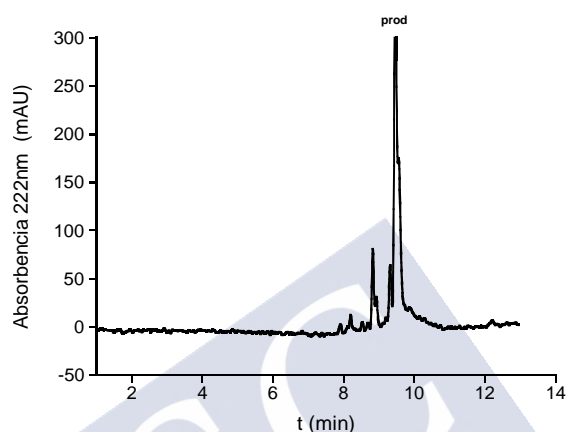


Figura 45. Cromatograma do fragmento N-terminal mutado **A<sup>500</sup>F<sup>522</sup>-Dbz** e sen desamidación.

A síntese do fragmento peptídico intermedio tamén se realizou en fase sólida, co soporte Rink Amide Chem Matrix e seguindo a metodoloxía Fmoc habitual.

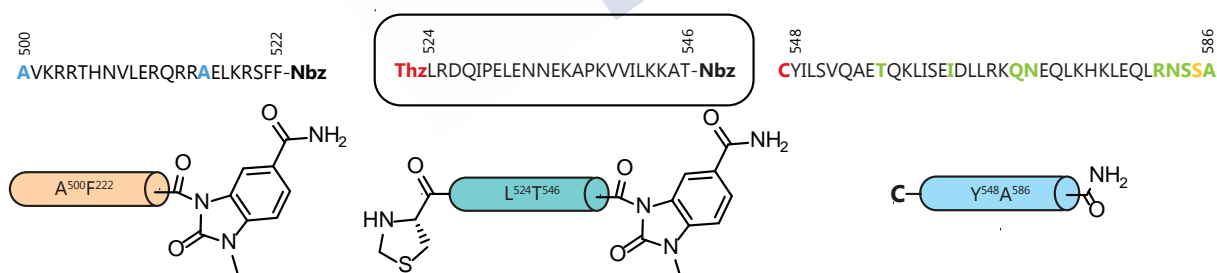
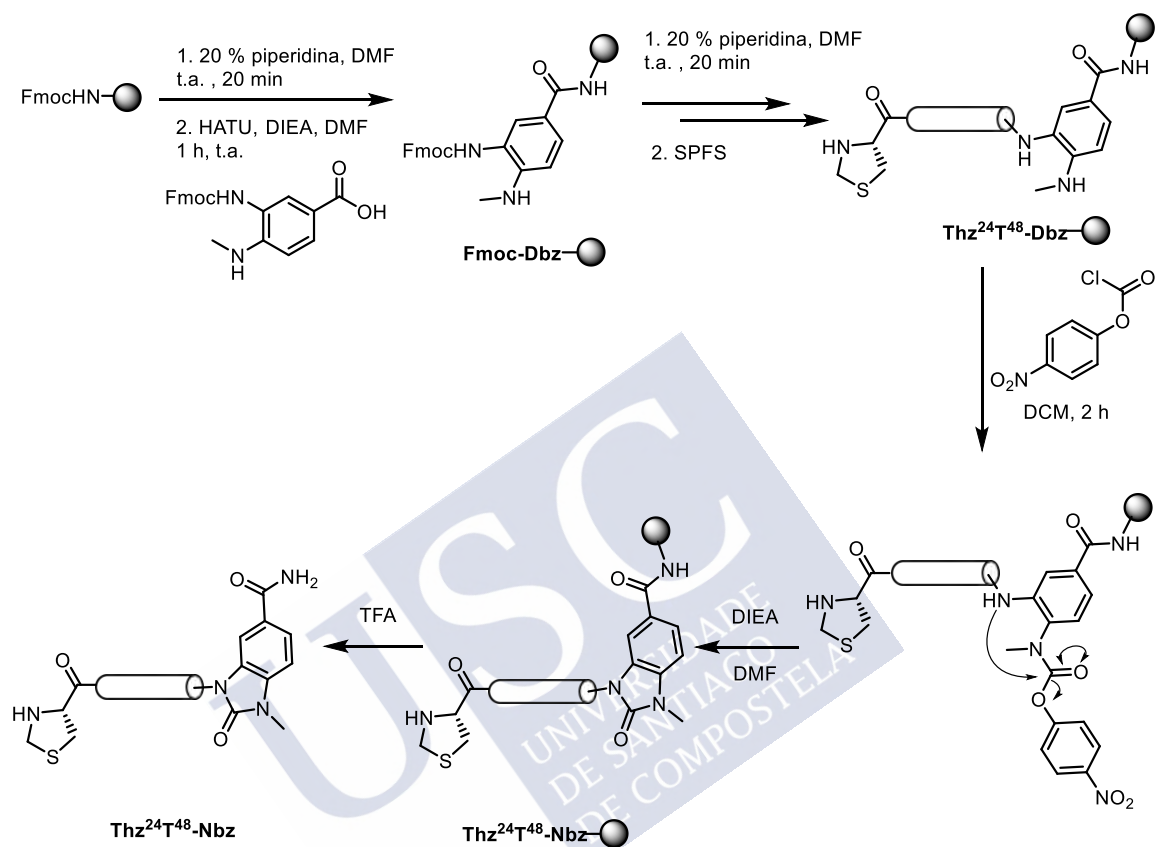


Figura 46. Representación dos fragmentos de OMOMYC. Rodeado aparece o fragmento cuxa síntese se trata neste apartado.

Como se observa da secuencia, debe introducirse unha Cys na posición N-terminal deste fragmento. Para evitar oligomerizacións e/ou ciclacións indesexadas durante as reaccións de LQN, dito residuo protexerase como Thz (4-carboxo-1,3-tiazolidina) cuxa desprotección ortogonal se pode realizar en disolución unha vez rematada a LQN.<sup>101</sup> O procedemento para

<sup>101</sup> D. Bang, S. B. H. Kent, *Angew. Chem. Int. Ed.*, **2004**, *43*, 2534.

este anaco require a ruta sintética explicada no **Esquema 10**. Neste caso, o primeiro aminoácido Dbz, sintetizado previamente, acóplase á man, seguido da síntese en fase sólida do resto da secuencia. Como último aminoácido introdúcese a cistina 27 enmascarada como Thz en condicións estándar e, finalmente, transfórmase o Dbz en Nbz con cloroformato.



Esquema 10. Ruta sintética para obter o fragmento intermedio  $\text{Thz}^{523}\text{T}^{546}\text{-Nbz}$ .

Por último, compróbase que se obtivo o produto precipitando unha pequena alícuota de resina e analizándoa por HPLC-EM. Identificouse o produto desexado no cru de reacción, polo que se desancorou do soporte sólido o resto do péptido con tratamento acedo de TFA, precipitouse en éter frío e liofilizouse. O cru ten poucas impurezas, polo que o produto obtido se usou nas reaccións posteriores sen mais purificación. Obtívose a secuencia desexada cun rendemento aproximado dun 29 %, tendo en conta a cantidade que se precipitou de resina con respecto da carga desta,



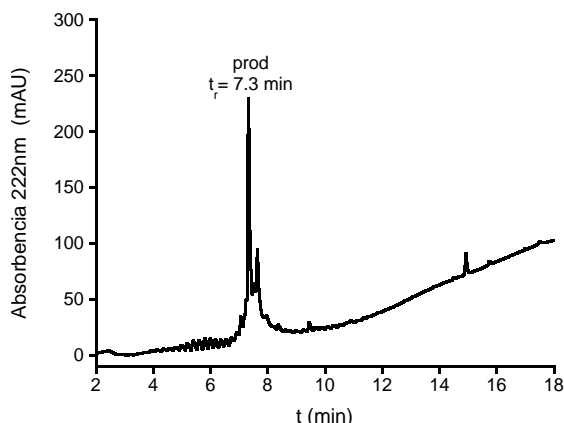


Figura 47. Cromatograma do cru do fragmento intermedio Thz<sup>523</sup>T<sup>546</sup>-Nbz.

Con respecto á secuencia que corresponde ao anaco C-terminal, atópase unha cistina na posición 585.

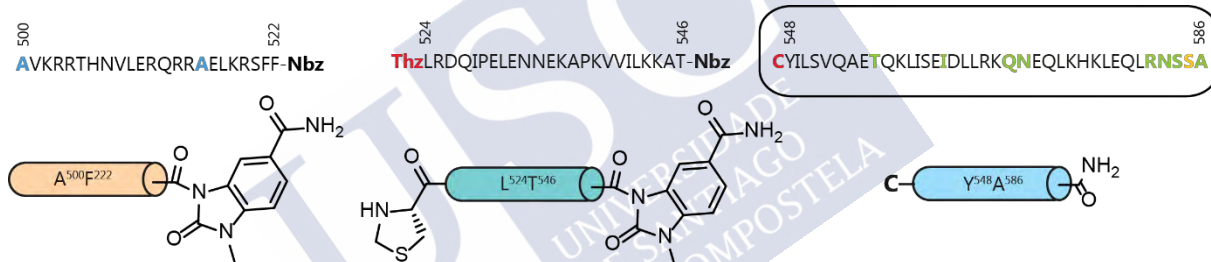


Figura 48. Representación dos fragmentos de OMOMYC. Rodeado aparece o fragmento cuxa síntese se explica a continuación.

Como a reacción de LQN é moi selectiva, posiblemente non houbera problemas en manter esa Cys, pois a fase da tioesterificación é reversible e o reagrupamento N-S non sería posible para este residuo posto que non ten grupos amino en posición beta con respecto do sulfhidrilo. A pesares desta circunstancia, considerouse substituír este residuo para evitar problemas de reaccións laterais ou impurezas que complicasen a metodoloxía. Así, a Cys<sup>86</sup> foi trocada por unha Ser na mesma posición. Este cambio é debido a que a serina é o aminoácido con propiedades estruturais máis parecidas á cistina. Ámbolos dous son residuos neutros, aproximadamente do mesmo tamaño. Só difiren no intercambio dun átomo de xofre (Cys) por un de osíxeno (Ser) e a aminoácido conserva as súas propiedades hidrofílicas.<sup>102</sup>

<sup>102</sup> M. J. Betts, R. B. Russell, Bioinforma Genet Wiley Chichester. 2003, 317, 289. Chapter 14--

Polo tanto, a síntese do fragmento de 39 aminoácidos, agora renumerado  $C^{547}A^{586}$ , realizouse mediante SPFS asistida por microondas, seguindo o método Fmoc estándar. Como soporte sólido escolleuse a resina Rink Amida Chem Matrix con carga aproximada 0.5 mmol/g para obter unha boa cantidade de péptido. Para acoplar os aminoácidos utilizouse unha disolución 0.5 M de *N,N'*-diisopropilcarbodiimida (DIC) como axente activante e 1 M de acetato de etil-2-ciano-2-(hidroxiiminio), (oxima), como aditivo, en DMF. A oxima actúa como base convertido rapidamente en éster reactivo o intermedio O-acilisourea. Ditos ésteres son bastante estables e diminúen os subprodutos de racemización ou outras reaccións laterais. A combinación DIC/oxima é unha alternativa ao uso de HOBt como aditivo. Comparado co HOBt, a oxima é un reactivo máis barato, ten baixo risco de explosión e maior capacidade para inhibir a racemización.<sup>103</sup>

Despois de ensamblar o fragmento no sintetizador de péptidos, unha pequena alícuota do cru analizouse por HPLC-EM corroborando que a síntese transcorrera perfectamente. Pasouse entón, a desancorar o péptido do soporte sólido por tratamento ácido e precipitouse en éter frío. Como o cru de reacción resultou moi limpo por HPLC, este fragmento usouse sen maior purificación. Despois de redissolver o péptido nunha mestura 10 % de MeCN en auga e liofilizalo, determinouse un rendemento global a partir da rotura da resina dun 12 % aproximadamente. Algo baixo, seguramente por mor de que se perde produto na etapa de precipitación.

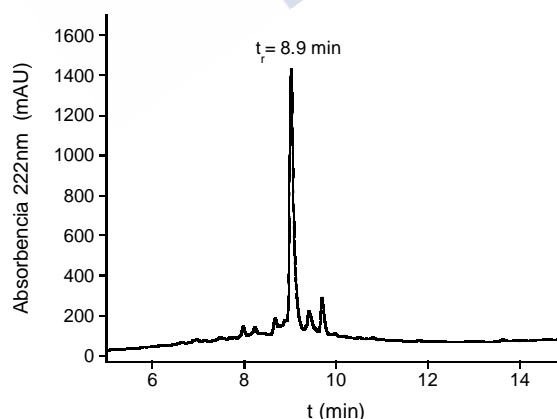
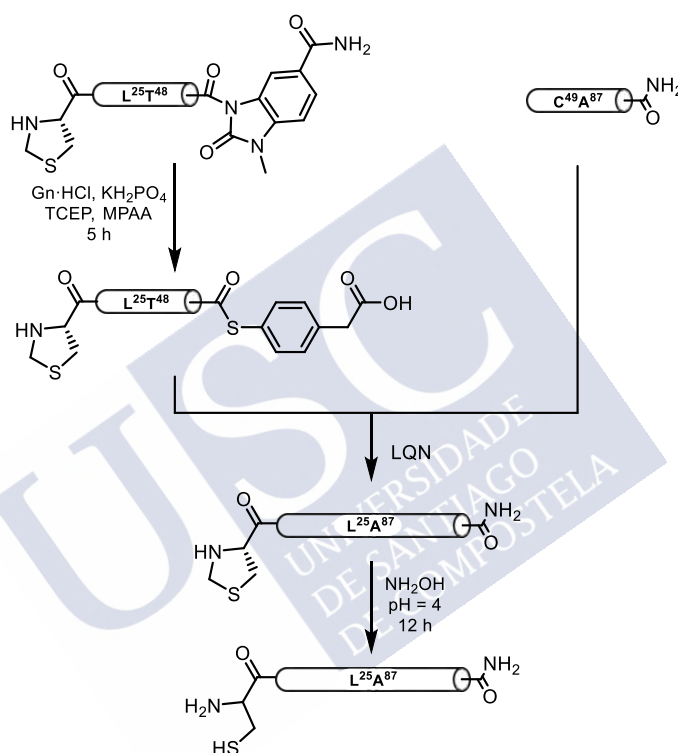


Figura 49. Cromatograma do cru da síntese por SPFS do fragmento  $C^{49}A^{87}$

<sup>103</sup> R. Subirós-Furanos, R. Prohens, R. Barbas, A. El-Faham, F. Albericio, *Chem. Eur. J.* **2009**, *15*, 9394.

### 3.6.2. Unión dos fragmentos por LQN e desulfuración para obter [AQS]OMOMYC

Unha vez obtidos os fragmentos, o seguinte paso é unilos mediante LQN. Usarase a ligazón química nativa para obter péptidos máis longos crecendo a proteína de C a N-terminal en disolución. Para iso, comézase xuntando a través dunha primeira reacción de LQN o fragmento C-terminal co fragmento intermedio. Neste caso atopouse que as condicións óptimas para este paso son engadir 1 equiv do fragmento intermedio **Thz<sup>523</sup>T<sup>546</sup>-Nbz** xunto con 1.3 equiv do fragmento C-terminal **C<sup>547</sup>A<sup>586</sup>**.



Esquema 11 Esquema dos pasos para a primeira LQN entre o fragmento C-terminal e o fragmento intermedio.

As dúas substancias de partida uníronse seguindo o procedemento estándar de LQN explicado anteriormente, para dar o fragmento **C<sup>523</sup>A<sup>587</sup>**. O procedemento consiste en axitar a mestura dos dous fragmentos unhas 5 h hasta que se observa a consumición do reactivo limitante (HPLC). Logo desprotexeuse selectivamente a cistina C<sup>523</sup> engadindo hidroxilamina sobre o cru de reacción anterior disolto no tampón da LQN e axustando o pH a catro. Deixando reaccionar toda a noite a temperatura ambiente observouse por HPLC a aparición de un novo pico que corresponde co produto coa Cys natural desprotexida.

O cru de reacción é moi limpo, o que permitiu recuperar o produto de ligazón cun excelente rendemento do 43 % .

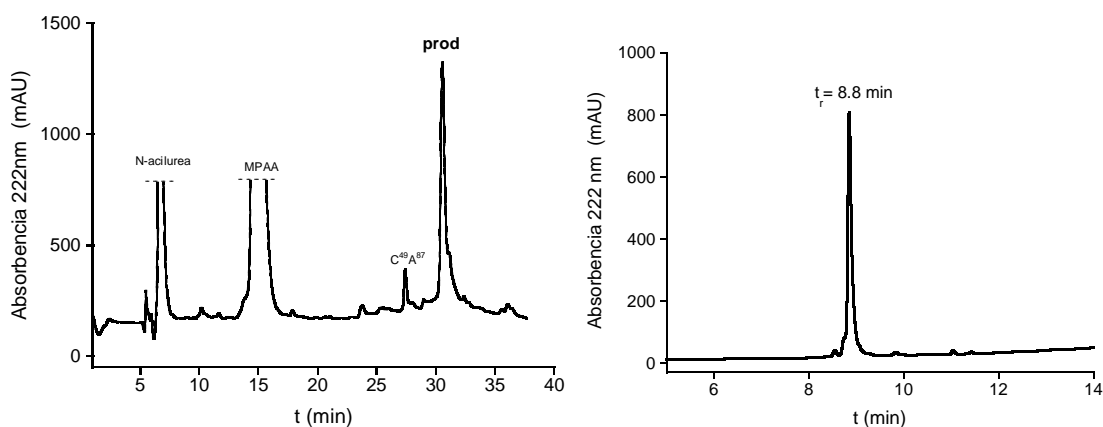


Figura 50. Esquerda: Cromatograma do cru de reacción despois de engadir hidroxilamina e deixalo axitando toda a noite. Dereita: Cromatograma do produto puro obtido.

Unha vez se liofilizou e illou o fragmento C<sup>523</sup>A<sup>586</sup> puro, pasouse a efectuar a segunda ligazón. Os reactivos neste paso foron o fragmento *N*-terminal A<sup>500</sup>F<sup>522</sup>-Dbz que actuará como limitante nesta etapa, e C<sup>523</sup>A<sup>586</sup> obtido da primeira LQN. Co mesmo procedemento da primeira LQN, realizouse a 2<sup>a</sup> LQN axitando a mestura durante 4.5 h a temperatura ambiente. Trala purificación, obtívose un rendemento desta segunda etapa dun 49 %.

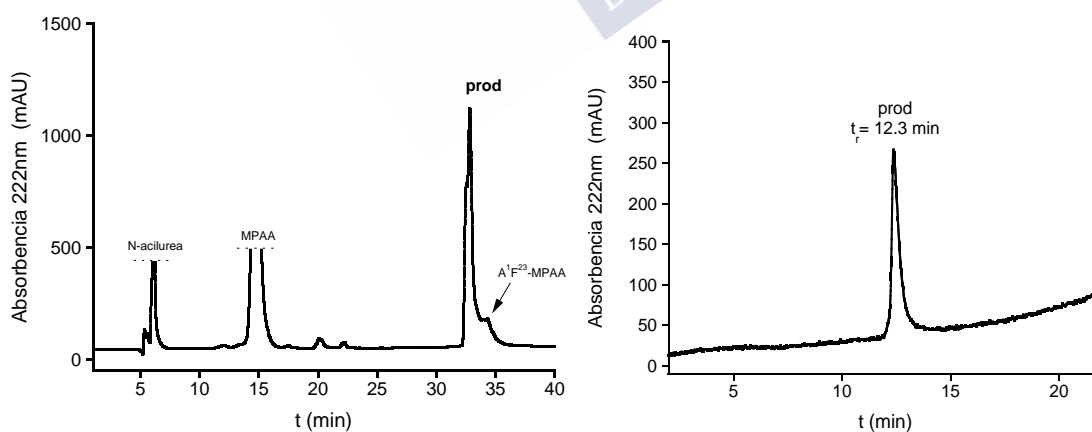


Figura 51. Esquerda: Cromatograma a 222 nm do cru de reacción da segunda LQN , con gradiente do 20 % ao 60 % de MeCN + 0.1 % TFA en 30 min. Dereita: Cromatograma do produto puro obtido a 222 nm con gradiente do 5 % ao 95 % de MeCN + 0.1 % TFA en 19 min.

### 3.7. SÍNTESE DE MINIPROTEINAS MYC MARCADAS CON FLUOROFOROS

#### 3.7.1.1. Síntese de [AQ]MYC-FLU polo método 3 Fr-LQN

Para levar a cabo a síntese das proteínas marcadas cun fluoróforo elixiuse a vía de síntese en tres fragmentos que serían unidos a través de dúas LQN, xa que é máis doado modificar un fragmento pequeno (coma os usados na estratexia 3 FrLQN) para introducir o fluoróforo, que modificar un fragmento cunha secuencia máis longa (como son os da estratexia 2FrLQN). Deste xeito a estratexia 3-FrLQN é máis versátil para introducir os marcadores fluorescentes. Primeiro considerouse a ancorar a fluoresceína no extremo C-terminal do péptido, pois era onde o equipo de Laura Soucek. introducía os fluoroforos para OMOMYC, aínda que no seu caso estaban obrigados a acoplalo a través de unha cistina introducida no método biosintético.

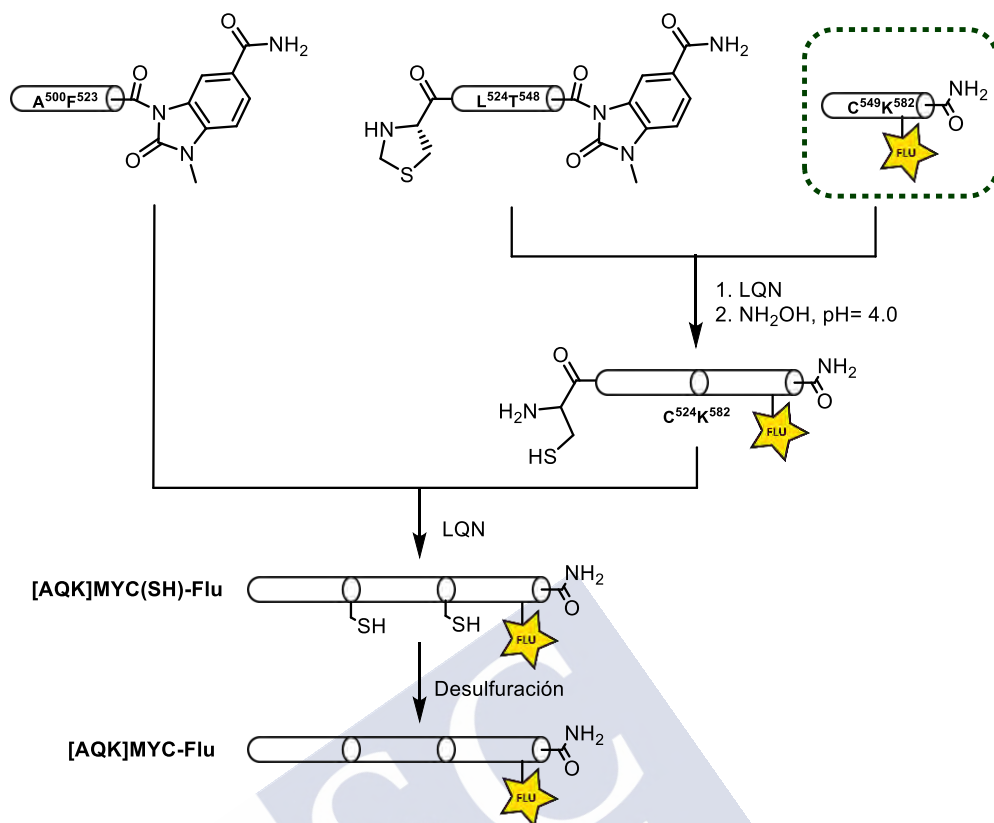
Pensouse que a mellor estratexia para introducir o marcador nesa posición consistiría en usar unha lisina C-terminal coa amina da cadea lateral protexida polo grupo *alloc* (aliloxicarbonilo). Este grupo protector podería ser eliminado ortogonalmente e de maneira sinxela en fase sólida, con catálise de paladio. Por iso plantexouse unha desconexión a través da alanina 524 e 548, coloreadas en vermello na figura 11. A maiores, a miniproteína contén a lisina *alloc* en posición C-terminal, residuo 582, para acoplarlle o fluoróforo a través da desprotección selectiva deste residuo. As mutacións nos residuos de Ala<sup>500</sup> e Gln<sup>516</sup> con respecto da proteína natural están resaltadas en azul claro na ilustración.



Figura 52. Secuencia do dominio de unión de MYC a ADN e representación da unión dos fragmentos peptídicos

#### Síntese dos fragmentos peptídicos

O camiño sintético para acadar os péptidos diana comeza por sintetizar en fase sólida, por separado, os tres fragmentos peptídicos.



esquema 12. Ruta sintética total para obter [AQK]MYC-Flu. O fragmento rodeado con liña descontinua é do que se vai discutir a síntese a continuación.

En primeiro lugar, levouse a cabo a síntese do fragmento C-terminal marcado coa fluoresceína seguindo os pasos do esquema seguinte:

Así pois, a síntese realizouse en fase sólida pola estratexia Fmoc e usando Rink Amide ChemMatrix. No extremo C-terminal deste fragmento acoplouse, a maiores, unha lisina protexida co grupo alloc. Unha vez que o péptido estivo completamente ancorado ao soporte sólido, esta lisina foi desprotexida ortogonalmente.<sup>104</sup>

Para verificar que se produciu a desprotección usouse o test do ácido 2,4,6-trinitrobenzenosulfónico (TNBS).<sup>105</sup> Este test consiste en mesturar unha pequena alícuota da resina cunha gota de DIEA ao 10 % en DMF mais unha gota de TNBS ao 1 % en DMF. O sulfonato do TNBS condensa cas aminas libres que poida haber na resina, xerando unha sulfonamida altamente cromoxénica a simple vista. Polo tanto, se despois duns minutos de reacción o soporte sólido se torna dun vermello rechamante, dise que o test é positivo e que a

<sup>104</sup> N. Thieriet, J. Alsina, E. Giralt, F. Guibe, F. Albericio, *Tetrahedron Lett.* **1997**, 38, 7275.

<sup>105</sup> W. S. Hancock, E. Battersby, *Anal. Biochem.* **1976**, 71, 260.

resina presenta aminas libres. Se a resina non cambia de cor, o test dise que é negativo e non hai aminas primarias libres.

Unha vez comprobado que a lisina estaba desprotexida, procedeuse a acoplar o fluoróforo. Para isto suspendeuse a resina nunha disolución de DIEA, DMF e o éster de succinimida da fluoresceína e a mestura axitouuse durante unhas 4 h a temperatura ambiente nun axitador magnético.

Unha vez precipitada unha pequena alicuota de produto para analizar o acoplamento do fluorescente por HPLC-EM sorprendentemente, no cromatograma a 222 nm observáronse varios picos con tempos de retención moi similares, mais ningún deles presentaba a masa esperada do produto.

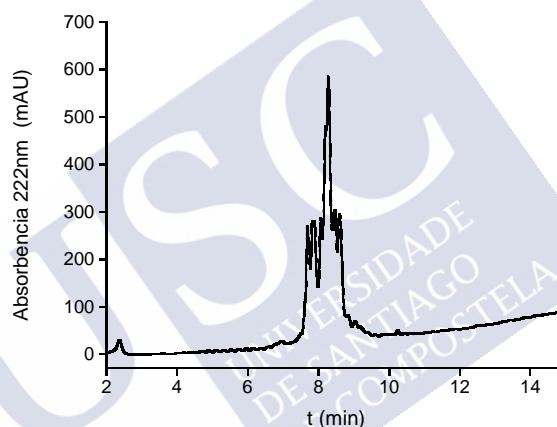
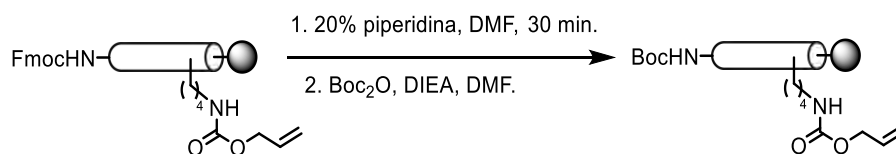


Figura 53. Cru de reacción despois acoplamento do fluoróforo.

Despois dunha análise detallada, asociáronse as masas dalgúns picos ca masa correspondente ao péptido de partida mais o acoplamento de dúas moléculas de fluoróforo. Logo de repetir a marcade unha vez máis e obter o mesmo resultado, pensouse que a base DIEA usada para ancorar o fluoróforo á resina, debido os longos tempos usados para a reacción de marcade, podía estar promovendo a liberación lenta do grupo protector Fmoc da amina en posición *N*-terminal.<sup>106</sup> Sendo así, o fluoróforo reaccionaría indiscriminadamente tanto en *C*-terminal (Lys<sup>582</sup>) como en *N*-terminal, dando unha mestura de péptidos marcados en unha ou nas dúas posicións. Polo tanto, decidiuse cambiar a estratexia e protexer o extremo *N*-terminal do fragmento con Boc, un grupo protector máis robusto que o Fmoc. Para isto, e aproveitando o fragmento xa sintetizado anteriormente, desprotexeuse o Fmoc *N*-terminal totalmente por

<sup>106</sup> M. Bodanszky, S. S. Deshmane, J. Martínez, *J. Org. Chem.* **1979**, *44*, 1622.

tratamento con piperidina e posteriormente protexeuse cun ataque carbonílico dende a amina cara o anhídrido de Boc, en DMF. A reacción deixouse axitando a temperatura ambiente durante unhas 5-6 h e comprobouse co test TNSB que tódalas aminas estaban protexidas.



Esquema 13. Desprotección do Fmoc C-terminal e protección con Boc.

A continuación, desprotexeuse o grupo alloc con catálise de Pd como se indicou anteriormente e, por último, acoplouse o éster de *N*-hidroxisuccinimida (NHS) da fluoresceína coa amina primaria libre da lisina C-terminal. Tras reaccionar, o marcador queda suxeito mediante un enlace amida ao péptido e libérase o hidróxido de succinimida como grupo saínte.

Despois de comprobar por HPLC-EM que se obtivo o péptido desexado, pasouse a desancorar a secuencia de aminoácidos da resina (**Figura 54**). Para isto, a resina tratouse co cóctel de desprotección para despois precipitar o produto en éter frío. Finalmente, o produto purificouse por HPLC e obtívose o fragmento desexado como un po amarelo nun rendemento aproximado dun 12 %. Os fragmentos peptídicos *N*-terminal, **A<sup>500</sup>F<sup>522</sup>-Nbz** e central, **Thz<sup>523</sup>T<sup>547</sup>-Nbz** sintetizáronse do xeito xa comentado no apartado de [AQS]OMOMYC .

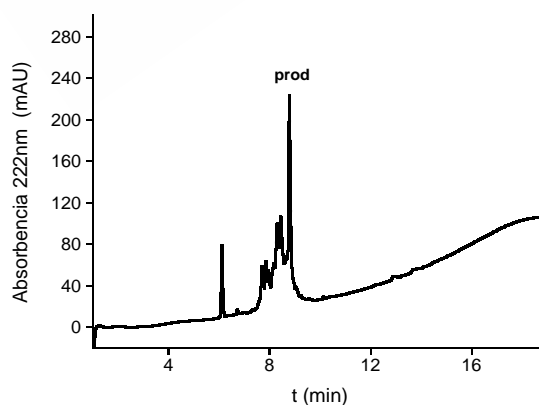
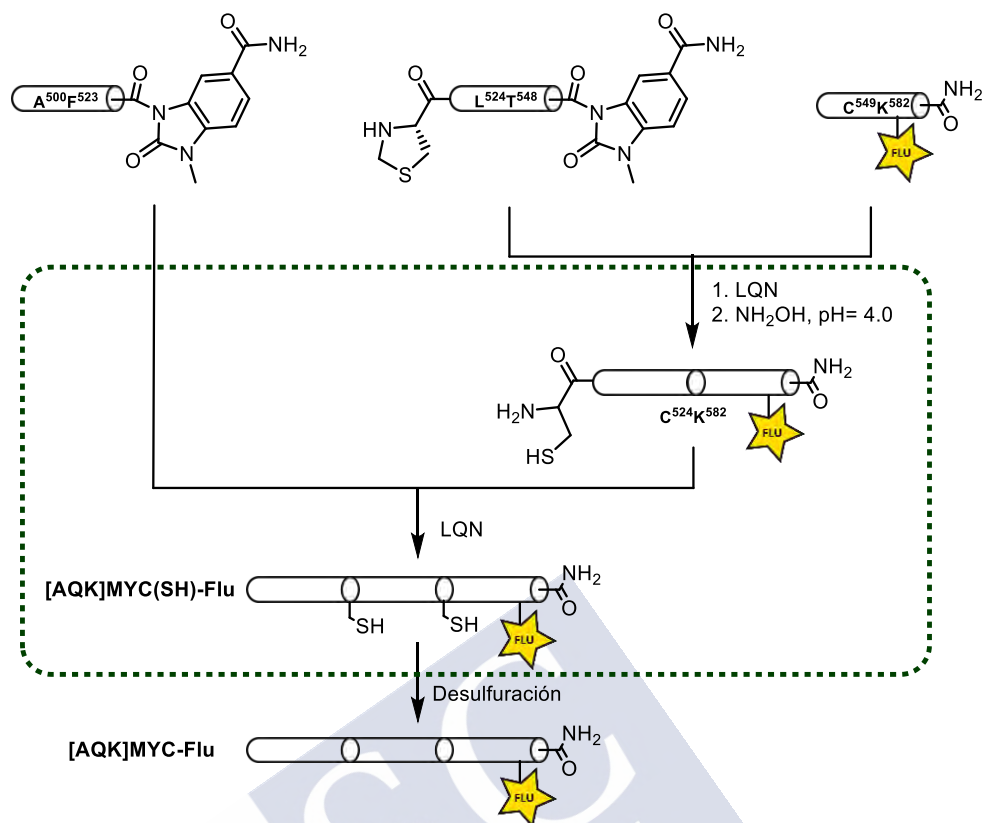


Figura 54. Cromatograma do cru de C<sup>548</sup>K<sup>582</sup>-Flu.

### Unión dos fragmentos por LQN e desulfuración para obter [AQK]MYC-Flu

A unión dos fragmentos peptídicos realizouse mediante LQN que fai crecer a proteína en sentido C-terminal a N-terminal.





Esquema 14. Esquema de síntese total. Resaltada a parte correspondente ás LQN.

Primeiramente uníronse os fragmentos **Thz<sup>523</sup>T<sup>546</sup>-Nbz** e **C<sup>547</sup>K<sup>582</sup>-Flu**. Despois de facer a ligazón en condicións *standard* e proceder á subseguinte desprotección da tiazolidina con hidroxilamina, analizouse unha alicuota do cru de reacción observándose que o pico do produto de acoplamento superpoñíase con varios picos doutros compostos e subprodutos. Malia probar diversos gradientes de disolventes para a fase móbil o proceso de purificación resultou complicado. A maiores, despois de analizar os picos recollidos no proceso de purificación, comprobouse que quedaba aínda moita cantidade de substancia de partida sen reaccionar. Os restos do fragmento inicial **Thz<sup>523</sup>T<sup>547</sup>-Nbz** que quedou sen reaccionar recolleuse xunto co pico do MPAA. De todas formas, se conseguimos obter o produto desexado nun 18 % de rendemento aproximado.

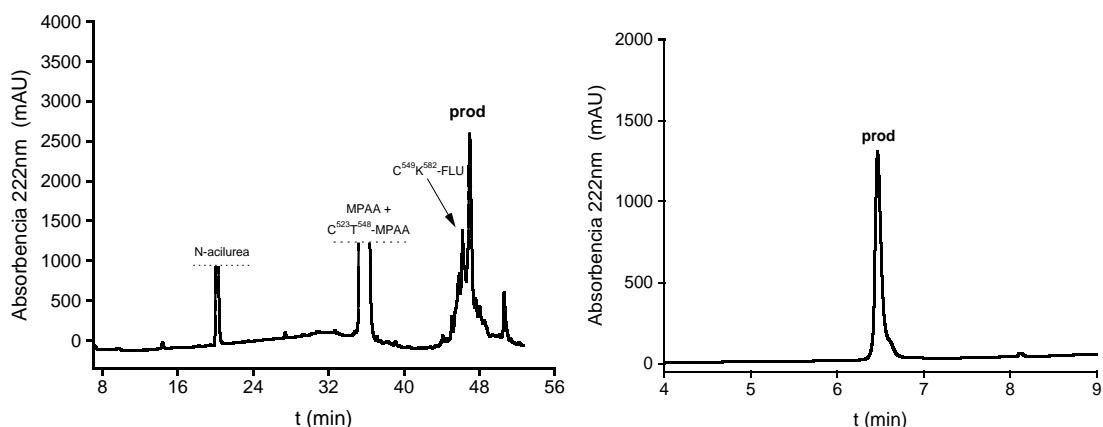


Figura 55. Esquerda: cru de reacción da primeira LQN despois da desprotección da tiazolidina. Dereita: cromatograma do produto puro C<sup>523</sup>K<sup>582</sup>-Flu a 222 nm.

De seguido, o produto obtido da primeira LQN xa ca cistina libre, fíxose reaccionar co fragmento A<sup>500</sup>F<sup>522</sup>-Nbz. O avance da reacción seguiu-se por HPLC-EM tanto ás 4.5 h como as 6 h sen observarse cambios significativos no cromatograma. Deste xeito, decidiuse purificar o cru, o que permitiu recoller o produto desexado [AQK]MYC(SH)-Flu como un pó amarelo despois de ser liofilizado (sobre 28 % de rendemento).

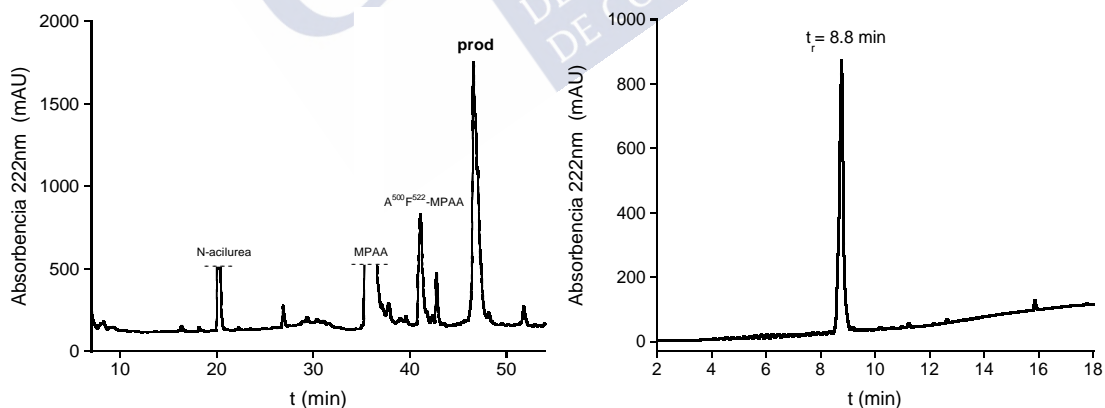


Figura 56. Esquerda: cru da segunda ligazón. Dereita: cromatograma de [AQK]MYC(SH)-Flu.

Para rematar, levouse a cabo a reacción de desulfuración. Probáronse diferentes tempos de reacción, diferentes concentracións de péptido e diferentes temperaturas, mais non puido observarse en ningún momento a masa do produto desexado. A maiores, os produtos recuperados da purificación nin tan sequera presentaban cor amarela ou laranxa característica

da fluoresceína. Nalgún caso atopouse un subproduto cunha masa de 44 umas menor que a do péptido desexado. Dita masa coincide ca descarboxilación da fluoresceína.

### 3.7.1.2. Síntese de [AQ]OMOMYC-FLU polo método 3·Fr-LQN

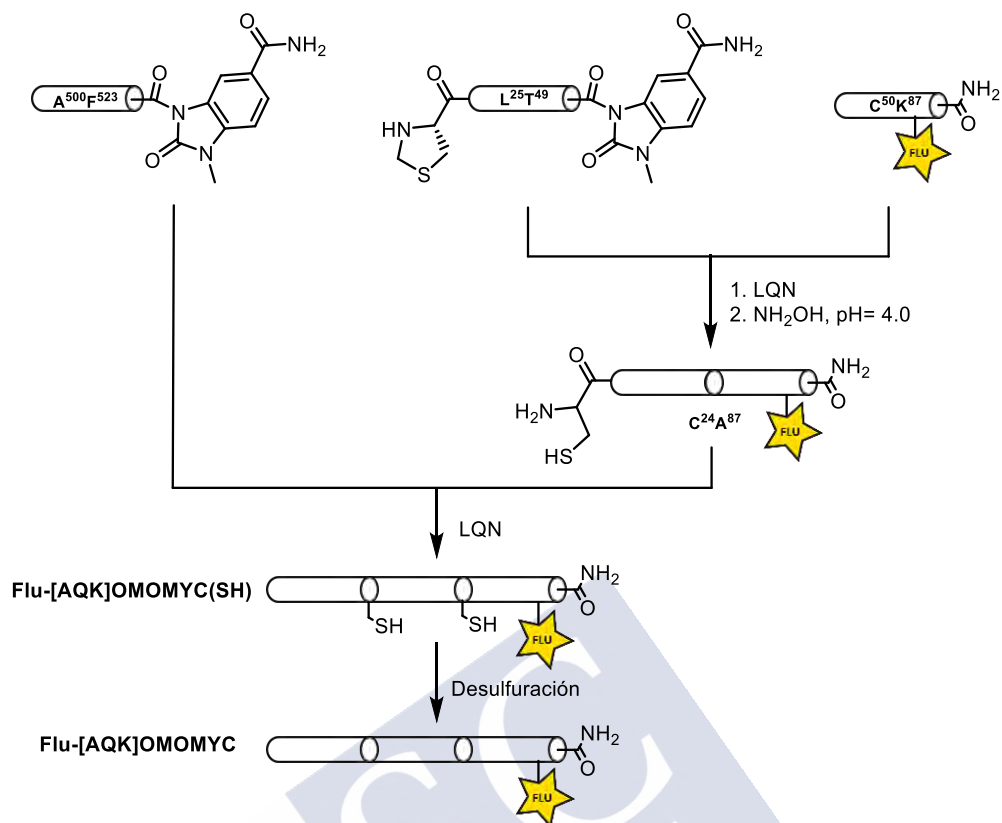
Paralelamente á síntese de [AQK]MYC-Flu, tamén se realizou a síntese de [AQK]OMOMYC-Flu. Esta miniproteína é un derivado de [AQS]OMOMYC marcado co fluoróforo fluoresceína no extremo C-terminal. Para esta preparación seguiuuse tamén a estratexia de unión de tres fragmentos mediante dúas LQN.

Segundo a proposta de retrosíntese para esta secuencia, os catro residuos nos que difire de MYC (marcados en verde na **Figura 57**) e os cinco aminoácidos a maiores, sitúanse todos no fragmento C-terminal, C<sup>49</sup>A<sup>87</sup>-FLU. En posición 86 a proteína orixinal contén un aminoácido cistina. Para evitar calquera problema co paso final de desulfuración e reaccións de polimerización durante as LQN, aproveitouse para mutar esta cistina 86 por lisina (alloc) para engancharlle o marcador fluorescente nesta posición.



Figura 57. Secuencia do dominio de unión de OMOMYC a ADN e representación da unión dos fragmentos peptídicos

A síntese realizouse de maneira análoga a outros apartados, de acordo co esquema sintético seguinte:



Esquema 15. Síntese total do análogo de OMOMYC marcado con fluoresceína en N-terminal.

### Síntese dos fragmentos peptídicos

Tanto o fragmento *N*-terminal como o fragmento intermedio xa os tiñamos preparados das sínteses anteriores ( $A^{500}F^{522}-Nbz$  e  $Thz^{523}T^{547}-Nbz$  que pasan a ser  $A^1F^{23}-Nbz$  e  $Thz^{24}T^{48}-Nbz$  respectivamente). Despois de crecer a secuencia usando síntese de péptidos en fase sólida ca estratexia Fmoc, o extremo *N*-terminal deste fragmento protexeu-se con Boc. A continuación elimínase o grupo *alloc*, protector da  $Lys^{586}$  durante toda a noite. A resina resultante tratouse con HATU e con pon o reactivo, e entón procedeu-se a desancorar o produto peptídico da resina con TFA. O cru de reacción precipitouse en éter frío, liofilizouse e o produto, denominado C49A87-Flu, purificouse no HPLC semipreparativo (aproximadamente 15 % de rendemento).

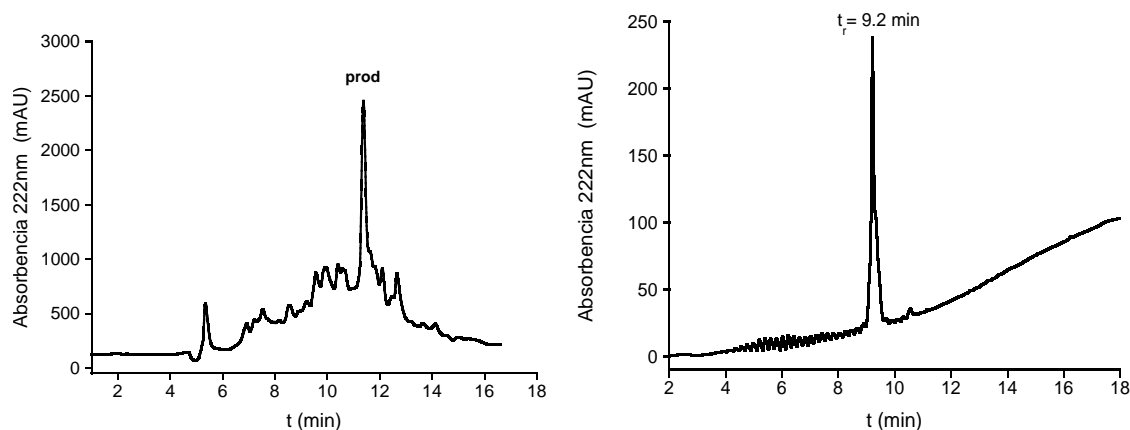


Figura 58. Esquerda: Cromatograma a 222 nm do cru de reacción para obter C<sup>50</sup>-A<sup>87</sup>-Flu. Dereita: cromatograma do produto puro a 222 nm despois da purificación

### Unión dos fragmentos por LQN e desulfuración para obter [AQK]OMOMYC-FLU

Cos fragmentos xa disponibles, procedeuse primeiro a realizar unha tioesterificación do fragmento intermedio **Thz<sup>24</sup>T<sup>48</sup>-Nbz** por tratamento con 40 mM de MPAA en tampón GnHCl 6 M seguido de adición de **C<sup>49</sup>A<sup>87</sup>-Flu** para converter a tiazolidina na cistina.

O produto resultante do acoplamento illouse tras purificarse por fase revertida. Como se pode ver no cromatograma a 222 nm, o cru de reacción sae bastante sucio e o pico do produto aparece no medio dunha montaña de outros picos cun tempo de retención de 40 min aproximadamente. Aínda así, puido illarse o produto desexado, cun rendemento final de un 16 % aproximadamente, considerando as dúas etapas de ligazón.

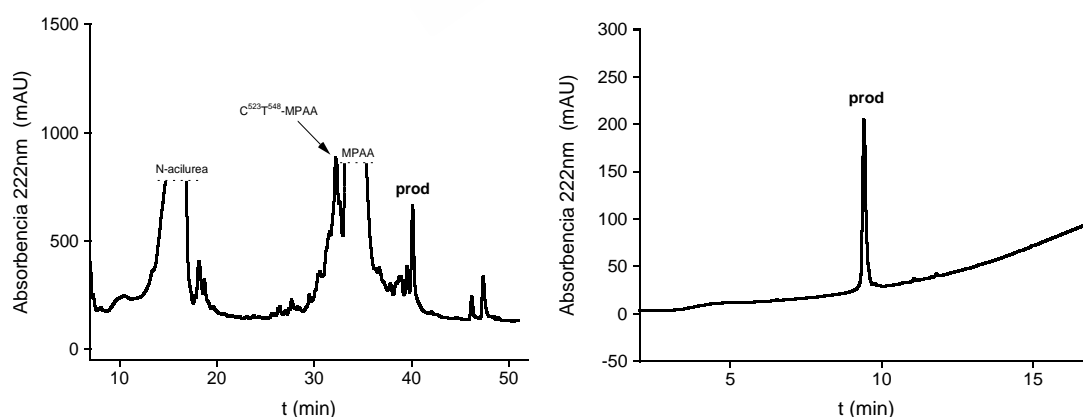


Figura 59. Esquerda: cromatograma do cru de reacción. Dereita: cromatograma do produto puro C<sup>24</sup>K<sup>87</sup>-Flu.

De seguido, o produto puro da primeira ligazón usouse como reactivo para efectuar a unión do último fragmento peptídico. Deste xeito, e ao igual que en ocasións anteriores, os fragmentos  $A^{1F^{23}}\text{-Nbz}$  e  $C^{24}K^{87}\text{-FLU}$  disolvéronse en tampón desnaturalizante e deixáronse reaccionar durante 4 h a temperatura ambiente. De novo, o cromatograma do cru da reacción resultou ser moi complicado de purificar, pero conseguiuase illar o produto desexado (rendemento aproximado do 21%)

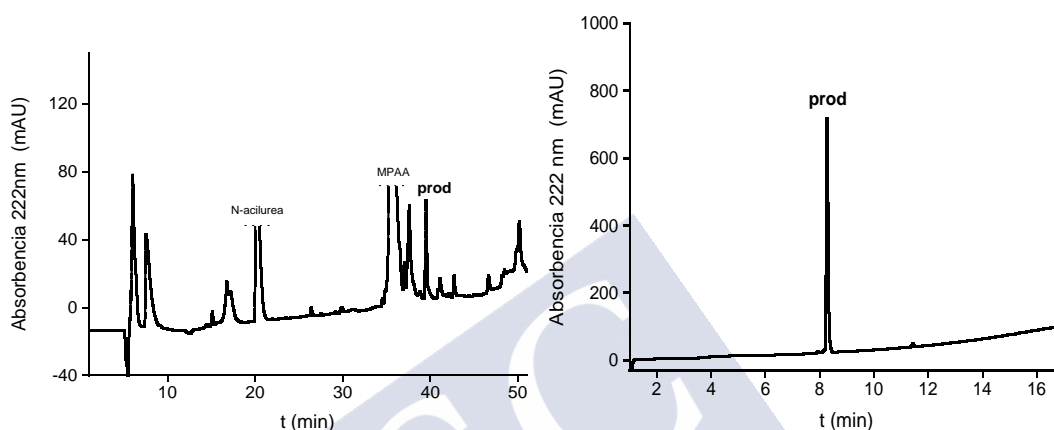


Figura 60. Esquerda: cru de reacción a 222 nm trala segunda LQN. Dereita: pico produto puro [AQK]OMOMyc(SH)-Flu a 222 nm.

Xa por último, ensaiouse a desulfuración do péptido [AQK]OMOMYC(SH)-Flu usando as condicións *standard*. Desafortunadamente, despois de 5 h de reacción, a análise HPLC-EM de unha alicuota do cru non mostrou a masa do produto indicado.

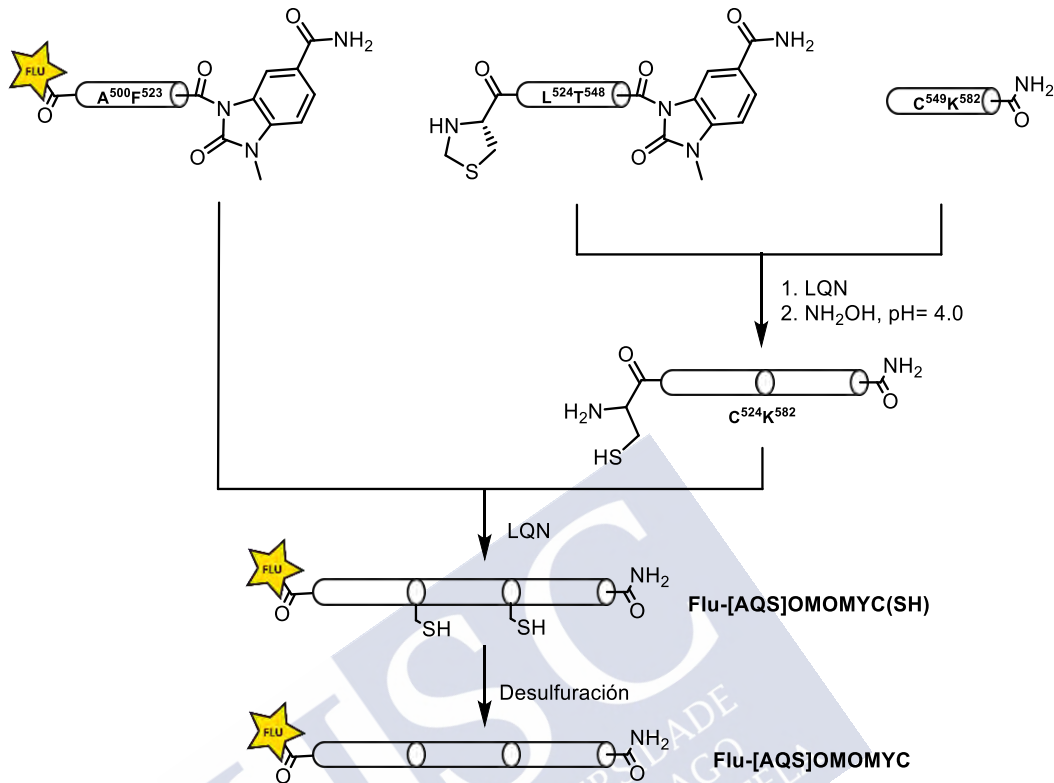
En vista de que non fomos quen de realizar a desulfuración desexada, que os crús de reacción seguindo esta estratexia de marcaxe estaban resultando máis complexos do esperado, e que os rendementos eran un pouco baixos, cambiouse a estratexia de síntese destes derivados marcados. A nova estratexia *B*, explícase no seguinte apartado.

### 3.7.2. Miniproteínas marcadas con fluoresceína: ESTRATEXIA B

#### 3.7.2.1. Síntese polo método 3 Fr-LQN de FLU[AQ]OMOMYC

Nesta nova estratexia, o fluoróforo elixido segue sendo a fluoresceína e a ruta tamén envolve o acoplamento de tres fragmentos peptídicos. A novidade consiste en incorporar o fluoróforo no fragmento N-terminal das miniproteínas en lugar do C-terminal. Deste xeito, a síntese é máis doada porque se introduce o fragmento fluorescente na segunda reacción de LQN, co que se evita levar o reactivo limitante (o fragmento marcado) dende o primeiro paso.

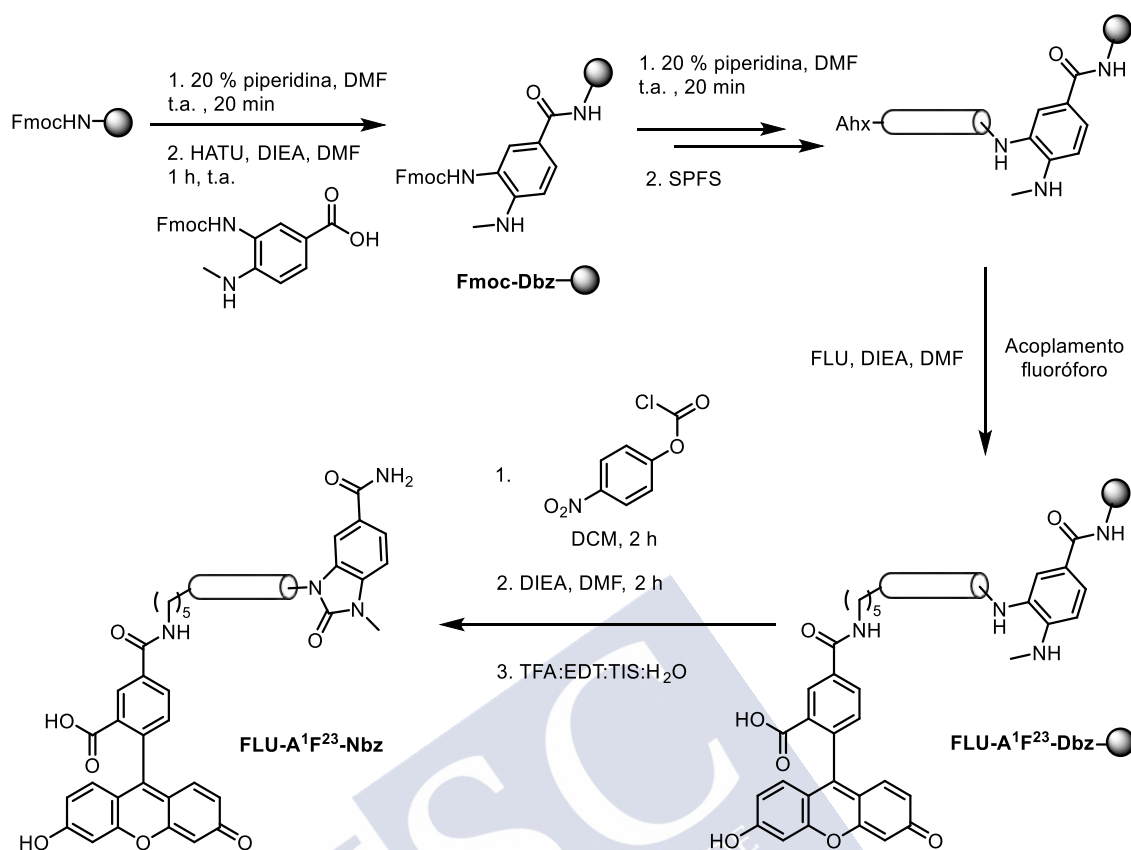
Amais, esperan obterse uns crús de reacción moito máis limpos. O esquema xeral da síntese total dos péptidos con esta estratexia amósase no **Esquema 16**.



Esquema 16. Ruta sintética para obter as miniproteínas marcadas no extremo N-terminal mediante a estratexia B.

### Síntese dos fragmentos

Para acoplar o grupo fluorescente no fragmento que se corresponde ca parte N-terminal foi preciso sintetizar un fragmento novo. O esquema que segue, reflexa os pasos sintéticos que foron precisos para obter dito fragmento, nomeado como **Flu-A<sup>1</sup>F<sup>23</sup>**:



Esquema 17. Ruta sintética para a obtención do fragmento *N*-terminal marcado **Flu-A<sup>1</sup>F<sup>23</sup>-Nbz**. Ahx = ácido aminoheptanoico que actúa como separador entre a cadea e o fluoróforo.

A secuencia peptídica creceuse no sintetizador de péptidos seguindo os protocolos estándar. Introduciuse un aminoácido (Ahx) en posición *N*-terminal que serve como espaciador entre a secuencia peptídica e o fluoróforo,<sup>107</sup> co fin de facilitar o acoplamento e evitar posibles interferencias indesexadas tales como que o impedimento estérico do marcador que poida chegar a condicionar a función bioactiva da secuencia

Unha vez crecida a secuencia peptídica no sintetizador e ancorado o espaciador Ahx, introduciuse o fluoróforo seguindo o método estándar con DIEA en DMF. De seguido, desancorouse unha alícuota pequena de péptido da resina por tratamento con TFA, precipitouse en éter dietílico frío e analizouse por HPLC-EM, observándose o produto **Flu-A<sup>1</sup>F<sup>23</sup>-Dbz** desexado sen problema, tanto co fluoróforo como co Dbz intactos.

Seguidamente, fíxose reaccionar unha pequena alícuota da resina que contiña **Flu-A<sup>1</sup>F<sup>23</sup>-Dbz** con *p*-nitrofenilclorofomato. Sorprendentemente, a reacción deu unha conversión completa no produto **Flu-A<sup>1</sup>F<sup>23</sup>-Nbz**, sen ningún indicio de degradación do fluoróforo ou do

<sup>107</sup> M. Jullian, A. Hernandez, A. Maurras, K. Puget, M. Amblard, J. Martinez, G. Subra, *Tetrahedron Lett.* **2009**, 50, 260.



Nbz, así como un cromatograma do cru bastante limpo. Despois de purificarse, illouse o fragmento cun 15 % de rendemento global.

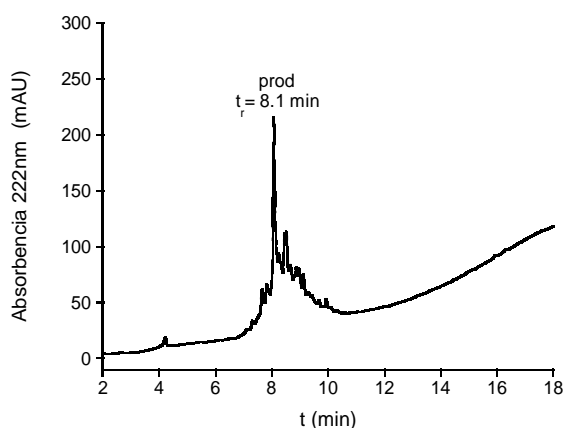


Figura 61. Cromatograma do cru de reacción para obter Flu-A<sup>1</sup>F<sup>23</sup>-Nbz.

Os fragmento intermedios **Thz<sup>24</sup>T<sup>48</sup>-Nbz** e **C-terminal C<sup>49</sup>A<sup>87</sup>**, son os mesmos que se utilizaron anteriormente na síntese de OMOMYC (**apartado 3.6**).

### Unión dos fragmentos por LQN e desulfuración para obter FLU-[AQ]OMOMYC

Cando os tres fragmentos estiveron sintetizados, efectuouse a primeira LQN entre os péptidos **Thz<sup>24</sup>T<sup>48</sup>-Nbz** e **C<sup>49</sup>A<sup>87</sup>** para formar a miniproteína **C<sup>24</sup>A<sup>87</sup>**, seguindo os procedementos xa usados na síntese de [AQS]OMOMYC (**3Fr-LQN**) (ver **apartado 3.6**).

Deseguido, fíxose a LQN de **C<sup>24</sup>A<sup>87</sup>** co fragmento **Flu-A<sup>1</sup>F<sup>23</sup>-Nbz**, observándose a formación do produto desexado **Flu-[AQS]OMOMYC(SH)**. O cru de reacción non era limpo, e quedaba moita sustancia de partida (**Figura 62**). De tódolos xeitos, o produto puido separarse cun rendemento do 14 %.

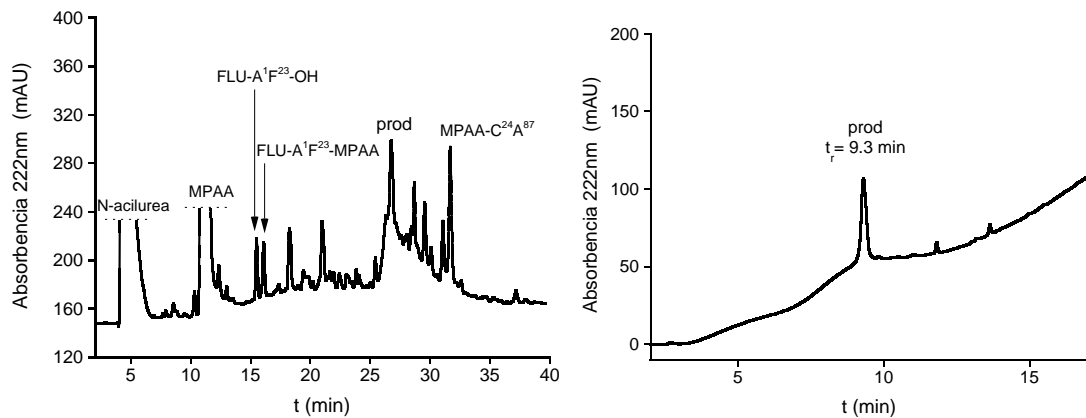


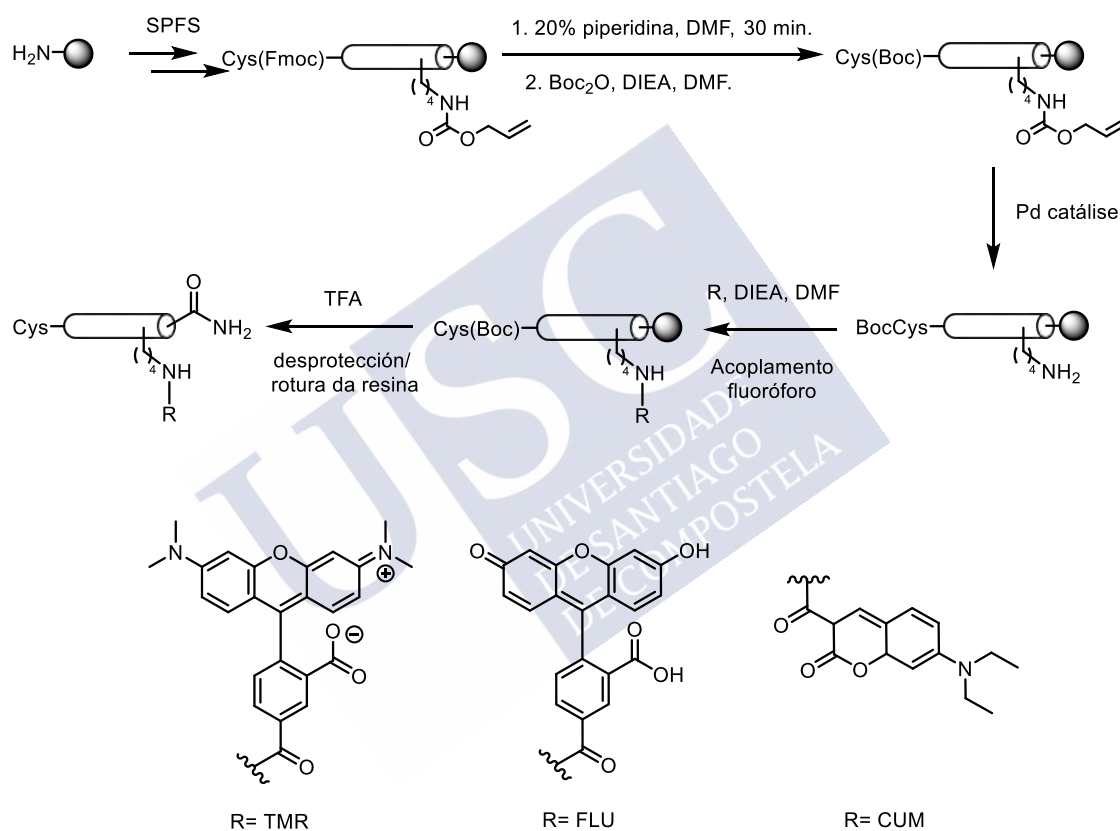
Figura 62. Esquerda: Cromatograma do cru de reacción ( 222 nm). Dereita: Cromatograma despois do produto Flu-[AQ]OMOMYC(SH) obtido trala purificación.

A miniproteína cas cistinas libres, someteuse á reacción de desulfuración para dar a secuencia peptídica con cadansúa alanina no lugar das cistinas, ao igual que o OMOMYC natural. Mais, de novo, e a pesares de probar a reacción varias veces, en ningún dos casos se chegou a detectar por HPLC-EM o produto desexado.

Neste momento non quedou outra que realizar un estudo máis profundo sobre a compatibilidade das condicións de desulfuración cos diferentes fluoróforos.

### 3.7.3. Estudo da resistencia de diferentes fluoróforos á reacción de desulfuración

Sintetizáronse en fase sólida catro secuencias peptídicas curtas, de 14 aminoácidos cada unha, que contiñan unha cistina *N*-terminal, e que se correspondían co extremo *C*-terminal de MYC. A tres delas acoplóuselles cadanseu fluoróforo na lisina inicial. Os fluoróforos escollidos foron fluoresceína (FLU), cumarina (CUM) e tetrametilrodamina (TMR) posto que eran dos que se dispoñían no laboratorio naquel momento.



Esquema 18. Ruta sintética seguida para a obtención dos péptidos modelo marcados con diferentes fluoróforos. A secuencia do péptido ca cistina terminal é CGGEQLKHKLEQLK(alloc).

A maiores, preparáronse un péptido similar pero que conten unha proparxilglicina en lugar da lisina na posición *C*-terminal. Este péptido podería funcionalizarse *a posteriori* a través dunha reacción de tipo *click* entre o triplo enlace da poteína e a azida dun fluoróforo.

Os péptidos sintetizáronse e purificáronse en condicións habituais. Os crús foron sometidos ás condicións de desulfuración engadindo 40 mM do iniciador de radicais VA-044, 4,6 % en volume de EtSH e 500 mM TCEP·HCl en 6M de Gn·HCl e 200 mM de tampón fosfato a pH= 7. As reaccións levaronse a cabo durante 4,5 h a 37 °C. Como a reacción xera gran

cantidade de residuos de moléculas pequenas, as masas dos péptidos non puideron detectarse a partir de análíticas dos crús, polo que tivo que procederse á súa purificación mediante cromatografía líquida en fase reversa e posterior liofilización dos produtos recollidos. Tra-la purificación, para o péptido modelo proparxilado, recuperouse un composto cuxa masa coincidía coa masa do péptido mais a fosfina xuntas ( péptido proba 2 + TCEP).

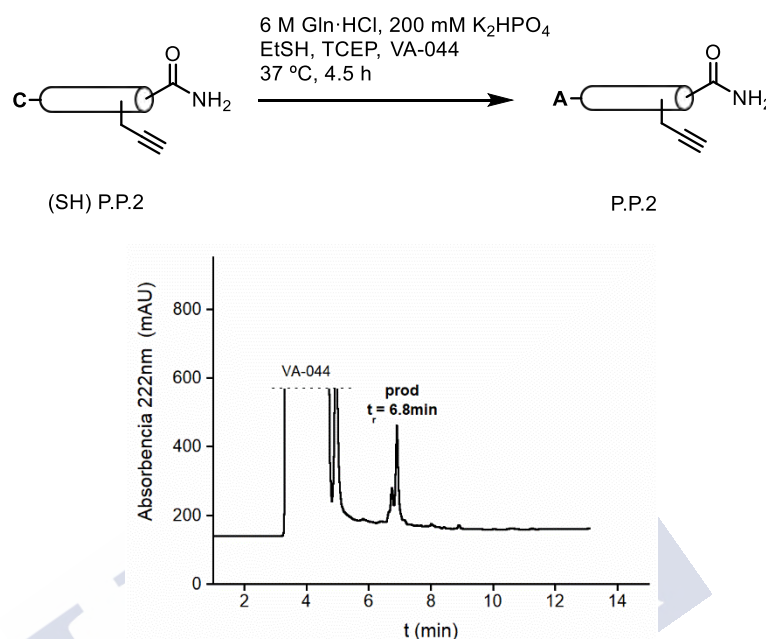


Figura 63. Arriba: Esquema da reacción de desulfuración do péptido proparxilado, para obter o correspondente péptido cunha alanina. Abaixo: cromatograma do cru onde se detecta o péptido esperado ( 222 nm)

Para o péptido con fluoresceína non se puido detectar o produto de desulfuración en ningún momento, recuperándose da purificación un único subproduto con tempo de retención de 11.4 min que non posuía cor e cuxa masa non puido asociarse con ningún composto coñecido.

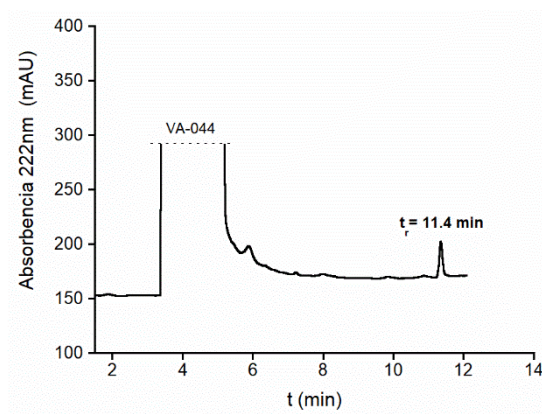
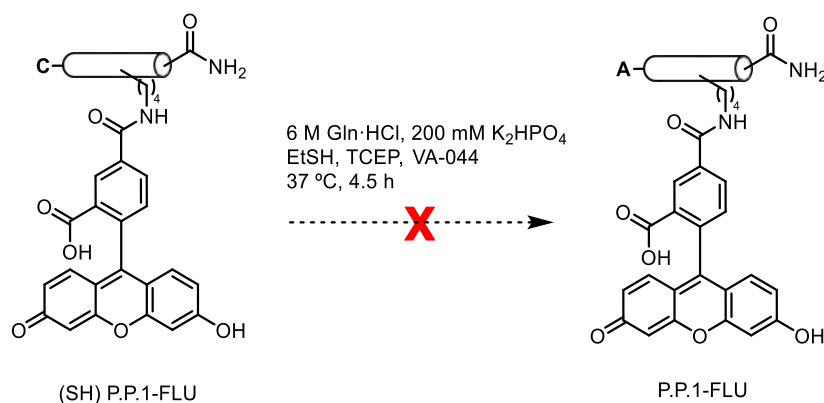


Figura 64. Arriba: esquema reacción da desulfuración co péptidos proba 1 con fluoresceína. Abaixo: cromatograma do cru de reacción ( $\lambda$ : 222 nm.).

Namentres, nos casos aos que se lle acoplara a TMR e a CUM si conseguiron illarse os produtos de desulfuración desexados e as reaccións transcorrían case de forma cuantitativa. Para o caso do péptido marcado con rodamina, recolleuse un produto cunha forte cor rosada a tempo de retención 6.5 min. Este resultou ser o produto de desulfuración esperado. A maiores, no caso do péptido con cumarina, tamén se conseguiu illar facilmente un sólido alaranxado cuxa masa coincidía perfectamente co produto de reacción desexado

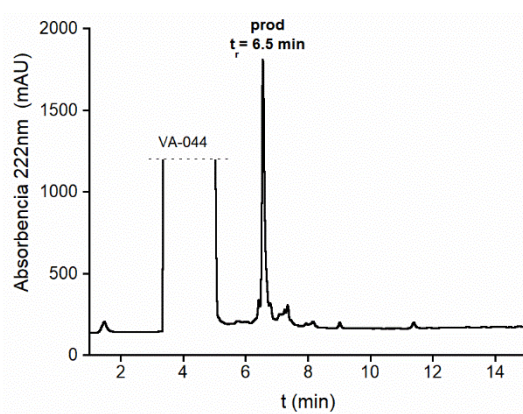
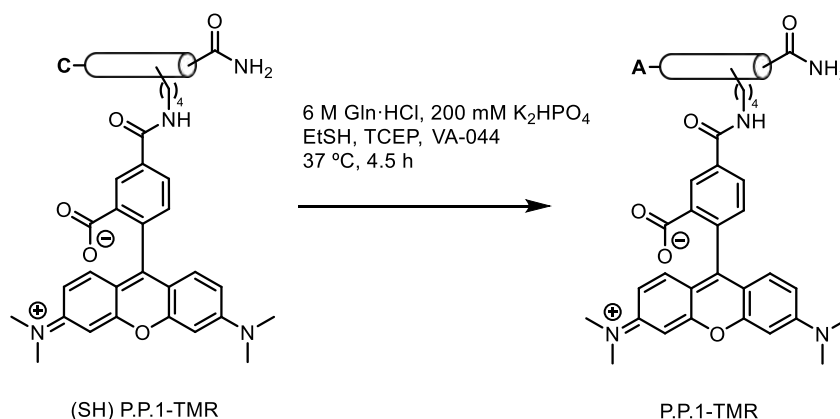


Figura 65. Arriba: esquema reacción da desulfuración co péptido proba 1 marcado con rodamina. Abaixo: cromatograma do cru de reacción.

Aínda que a cumarina soporta ben tódalas condicións de síntese, esta adoita ser menos luminosa e ter menor rendemento cuántico ca rodamina, polo que se debe usar o péptido marcado en maior concentración para ter a mesma intensidade de luz. As súas lonxitudes de onda de emisión e excitación son relativamente baixas, tirando cara o ultravioleta, onde moitas moléculas orgánicas absorben, e ademais, adoita ter problemas de escurecemento por autofluorescencia, polo que poden xurdir interferencias nos ensaios microscópicos.<sup>108</sup> A rodamina, pola contra, ten as bandas máis desprazadas cara infravermello o que evita interferencias nas medidas.<sup>109</sup> Por último, e se cadra máis importante, a fotoestabilidade dos derivados da rodamina é de 3 a 4 ordes de magnitude maior que os derivados da cumarina.<sup>110</sup> Deste xeito, o uso da cumarina é limitado cando se precisan tempos de irradiación prolongados ou de alta intensidade que a fai non axeitada para estudos en microscopía de células vivas.

<sup>108</sup> a) S. K. Grant, J. G. Sklar, R. T. Cummings, *Journal of Biomolecular Screening*, 2002, 7, 531.

<sup>109</sup> T. C. Turek-Etienne, E. C. Small, S. C. Soh, T. A. Xin, P. V. Gaitonde, E. B. Barrabee, R. W. Bryant, R. W., *Journal of Biomolecular Screening*, 2003, 8, 176.

<sup>110</sup> C. Eggeling, J. Widengren, R. Rigler, C. A. M. Seidel, *Anal. Chem.*, **1998**, 70, 2651.

En base a estes resultados decidiuse usar como fluoróforo a TMR pois tamén soportaba ben tódalas condicións de síntese e a súa fotoestabilidade é boa.

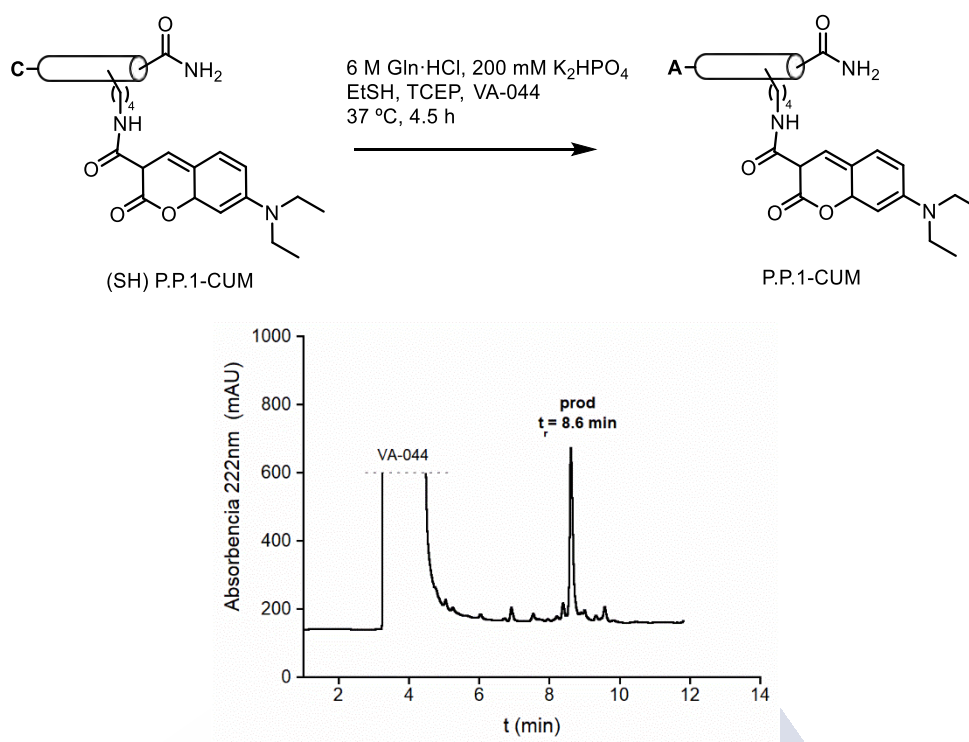


Figura 66. Arriba: esquema reacción da desulfuración co péptido proba 1 marcado con cumarina. abaixo: cromatograma do cru de reacción.

A maiores, cambiouse a estratexia para acoplar o marcador fluorescente no fragmento do extremo *N*- terminal dos péptidos a sintetizar, deste xeito, o proceso global de síntese sería máis doado xa que o fragmento marcado sempre actúa como reactivo limitante e non se introduciría na síntese ata o derradeiro paso. Amais, para simplificar a purificación dos produtos nos crús de reacción, só se usaría un dos isómeros do marcador. Neste caso elixiuse a 5-CarboxiTMR.

### 3.7.4. Miniproteínas marcadas con tetrametilrodamina

A síntese das oncoproteínas marcadas con TMR levouse a cabo seguindo a mesma estratexia 3Fr-LQN que a usada para trata de obter **Flu-[AQS]OMOMYC**. O fluoróforo introducirase no extremo *N*-terminal, o que simplificará tódolos pasos.

Observando as secuencias e os estudos realizados anteriormente, determinouse cortar a secuencia total das proteínas exactamente nos mesmos residuos de alanina que se usaran para os péptidos sen marcar. Na **Figura 67** as alaninas de corte están marcadas en vermello para as dúas proteínas MYC e OMOMYC. Desta forma obtéñense tres fragmentos en cada proteína de tamaño similar e axeitados para a súa síntese en fase sólida.



Figura 67. Arriba: secuencia do dominio de unión ao ADN de MYC. Abaixo: secuencia do dominio de unión a ADN OMOMYC.

Como os puntos de corte son os mesmos que os usados nas sínteses dos dous péptidos sen marcar, os fragmentos intermedios e *C*-terminais tamén son os mesmos usando anteriormente. O único fragmento diferente sería o que leva o fluoróforo, o *N*-terminal. Mais botando un ollo á secuencia vese a simple vista que este fragmento coincide para ámbolos dous péptidos polo que só se sintetizará unha única vez e empregárase en ámbalas dúas sínteses.

O fragmento *C*-terminal marcado, que consta dun total de 24 residuos, preparouse segundo o seguinte esquema de síntese:



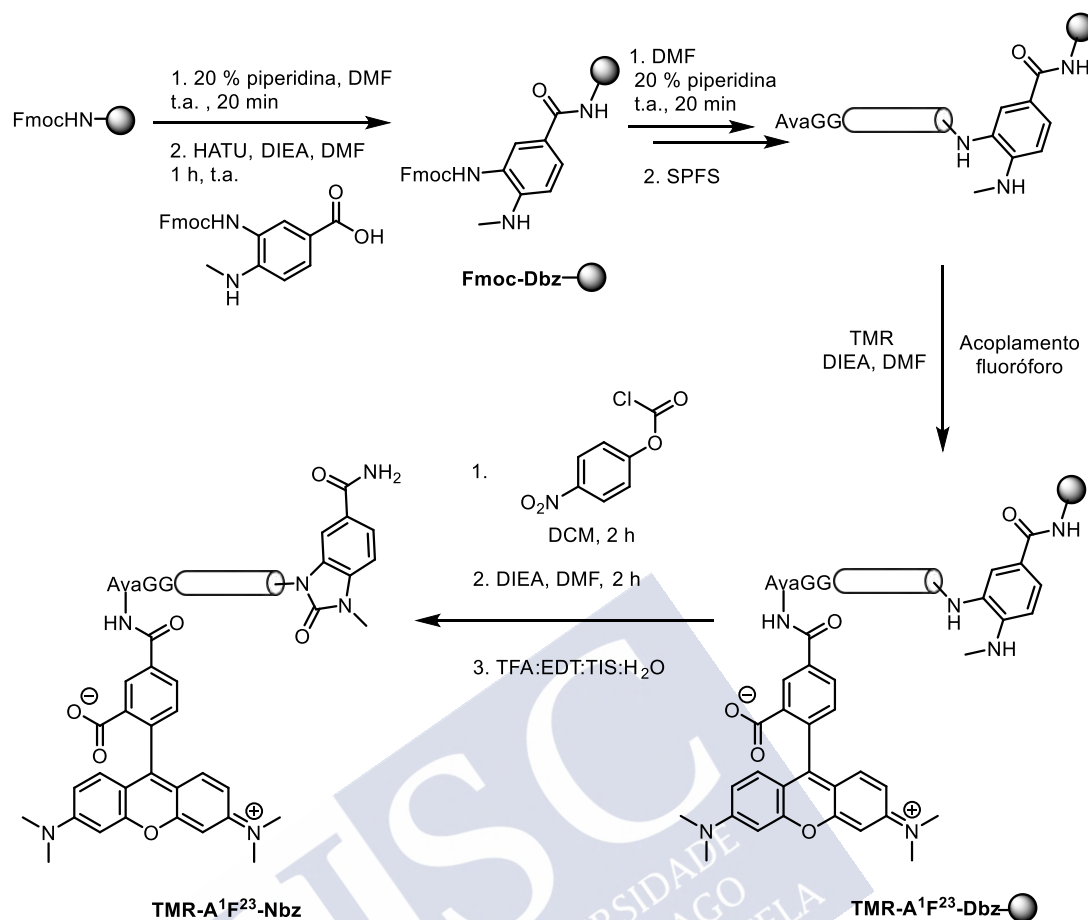


Figura 68. Ruta sintética para obter o fragmento N-terminal marcado con TMR. Ava= ácido aminoaléxico

Primeiramente, ancorouse á resina Rink-Amide ChemMatrix o aminoácido Dbz non natural previamente sintetizado. De seguido, desprotexeuse o Fmoc e realizouse o crecemento da cadea peptídica introducindo no extremo N-terminal do fragmento tres aminoácidos a maiores, o ácido 5-aminopentanoico (Ava) mais dúas glicinas (G), para asegurar que o fluoróforo quede ben separado da cadea peptídica. En segundo lugar, e ca secuencia peptídica aínda enganchada á resina, introduciuse o fluoróforo TMR. Para isto usouse DMF como disolvente, DIEA como base, e a tetrametilrodamina en forma de éster de succinimida. A suspensión deixouse axitando a temperatura ambiente unhas 3 h e analizouse unha pequena alícuota por HPLC, comprobando que o fluoróforo se ancorara perfectamente ao extremo N-terminal do fragmento a través dun enlace peptídico. Finalmente, transformouse o aminoácido Dbz no precursor do tioéster, Nbz, engadindo cloroformato ao péptido en resina, seguido dunha mestura de DIEA en DMF para acilar a aminoanilida. Este paso analizouse precipitando unha pequena alícuota da resina tratada, e examinándoa mediante HPLC-EM. O proceso resultou ser cuantitativo, polo que se procedeu a desancorar todo o péptido da resina por tratamento con

TFA durante unhas 4 h para logo precipitar o péptido en éter frío. Liofilizouse o sólido obtido e purificouse por HPLC de fase revertida, acadando o produto requirido cun 15 % de rendemento global.

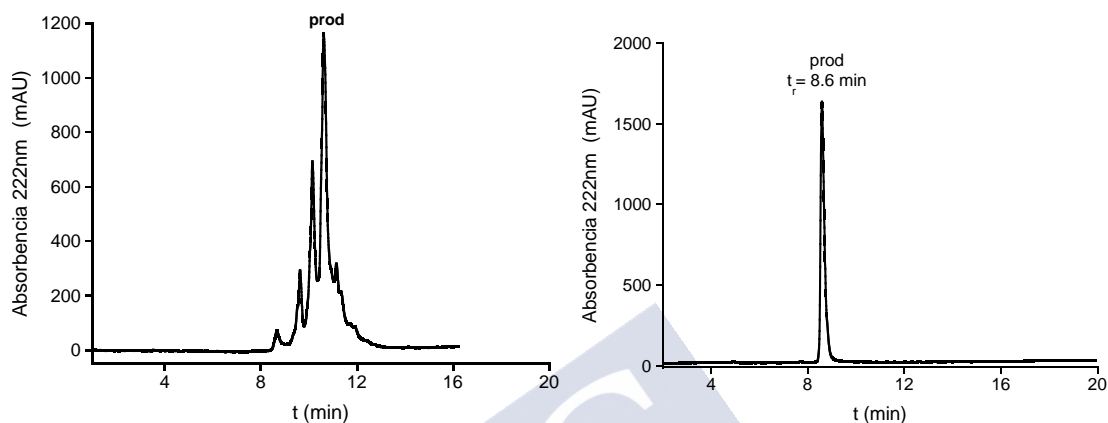
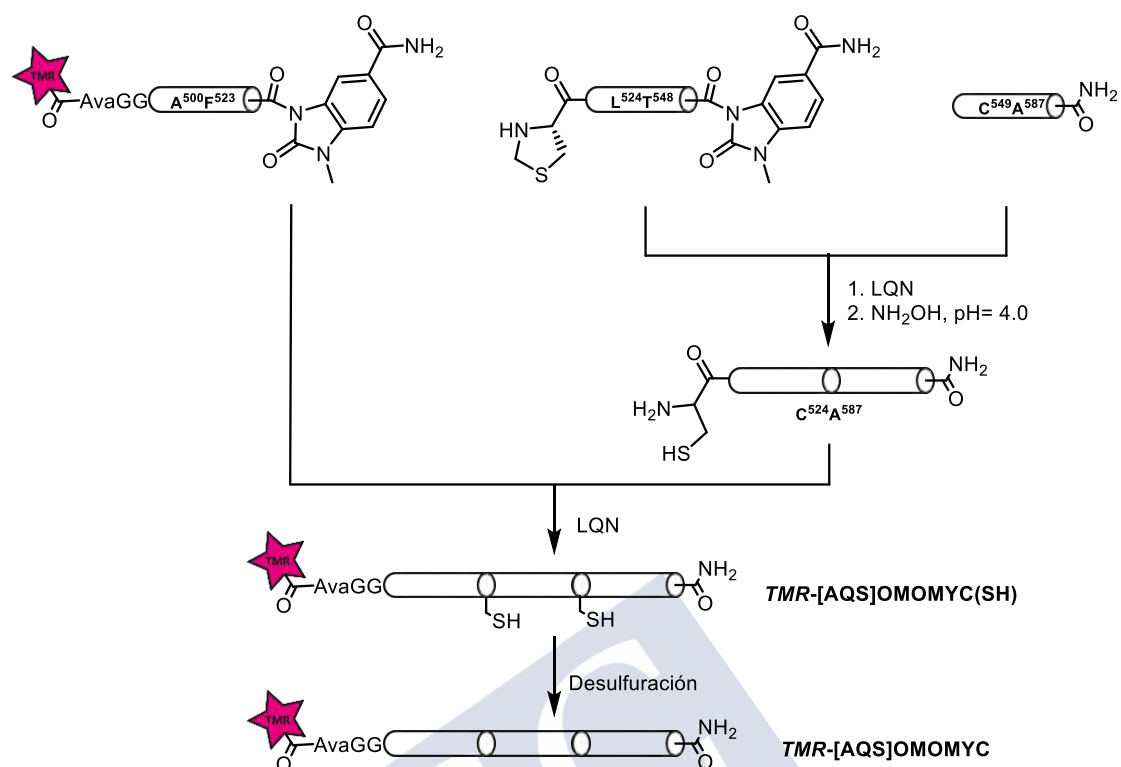


Figura 69. Esquerda: Cromatograma do cru despois da conversión a Nbz. Dereita: cromatograma do fragmento TMR-A<sup>1</sup>F<sup>23</sup>-Nbz puro.

#### 3.7.4.1. Unión dos fragmentos por LQN e desulfuración para obter TMR-[AQS]OMOMYC

Como se mostra no **Esquema 19**, a preparación do análogo de OMOMYC marcado con TMR levarase a cabo cunha estratexia que implica dúas ligazón químicas nativas que xerarán dúas cistinas, mais unha desulfuración final para transformar estas cistinas nos residuos de alanina presentes na secuencias dos péptidos naturais.



Esquema 19. Ruta sintética para obter polo método 3Fr-LQN TMR-[AQS]OMOMYC. Estrela rosa representa o fluoróforo rodamina.

A primeira ligazón química nativa para esta diana sintética realizouse usando os fragmentos **Thz<sup>24</sup>T<sup>48</sup>-Nbz** e **C<sup>49</sup>A<sup>87</sup>**, seguido da eliminación ortogonal do Thz para convertilo en Cys. Esta ligazón conduciu ao produto **C<sup>24</sup>A<sup>87</sup>** que coincide co obtido para a síntese de **[AQS]OMOMYC** sen marcar. Neste paso conseguiu-se un rendemento do 43 % do produto illado.

Posteriormente, realizouse a segunda LQN entre o fragmento marcado **TMR-A<sup>1</sup>F<sup>23</sup>-Nbz**, do cal se engadiron 1.5 equivalentes e o produto da ligazón anterior **C<sup>24</sup>A<sup>87</sup>** que se usou como reactivo limitante. Neste caso acadouse unha boa separación dos produtos cun gradiente da fase móbil lento (indicado na figura 31xxx) e puido detectarse **TMR-[AQS]OMOMYC(SH)** como produto maioritario da reacción. Ademais, tamén se identificaron os picos correspondentes ao fragmento de partida hidrolizado **TMR-A<sup>1</sup>F<sup>23</sup>-OH**, ademais do intermedio precursor da ligazón nativa, o tioéster **TMR-A<sup>1</sup>F<sup>23</sup>-MPAA**. Neste caso, antes da purificación, o cru filtrouse con amíóns. Deste xeito puido retirarse a maior parte das moléculas pequenas como o MPAA e a N-acilurea facendo a súa separación por HPLC de fase reversa moito máis doada.

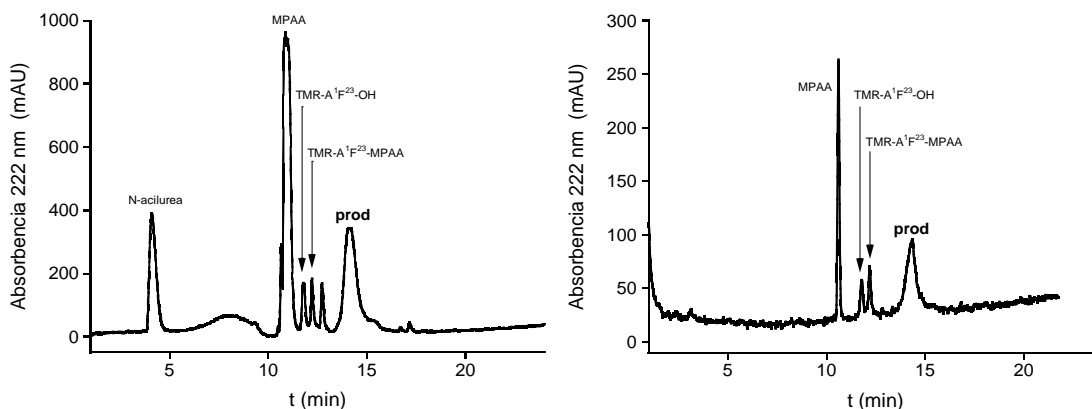


Figura 70. Esquerda: cru de reacción ás 4.5 h antes de ser filtrado con amicóns. Dereita: cru de reacción despois de pasar por amicóns. Método usado 5 % ó 75 % de MeCN e 0.1 % TFA en 19 min.

Despois de illar por HPLC e liofilizar o produto desexado, determinouse un rendemento para esta etapa dun 39 %.

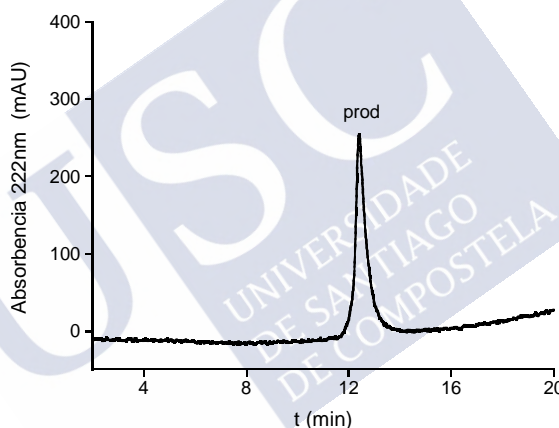


Figura 71. Cromatograma do produto puro. Gradiente do 5 % ó 95 % de MeCN e 0.1 % TFA en 19 min.

Pasamos agora a levar a cabo a reacción de desulfuración. A diferenza do que pasaba cando o péptido tiña como fluoroforo a fluoresceína, observamos que reacción funciona de forma moi eficiente. Despois dunhas 5 h de reacción a 37 °C, o cru purificouse e liofilizouse, para dar a miniproteína buscada con secuencia: **TMR-Ava-GGAVKRRTH NVLERQRRQE LKRSFFALRD QIPELENNEK APKVVILKKA TAYILSVQAE TQKLISEIDL LRKQNEQLKH KLEQLRNSSA-NH<sub>2</sub>** cun rendemento do 80 %.

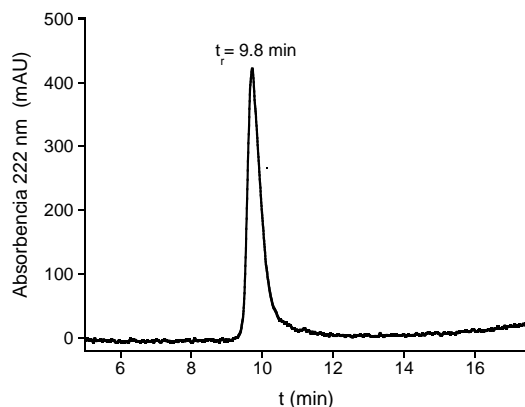
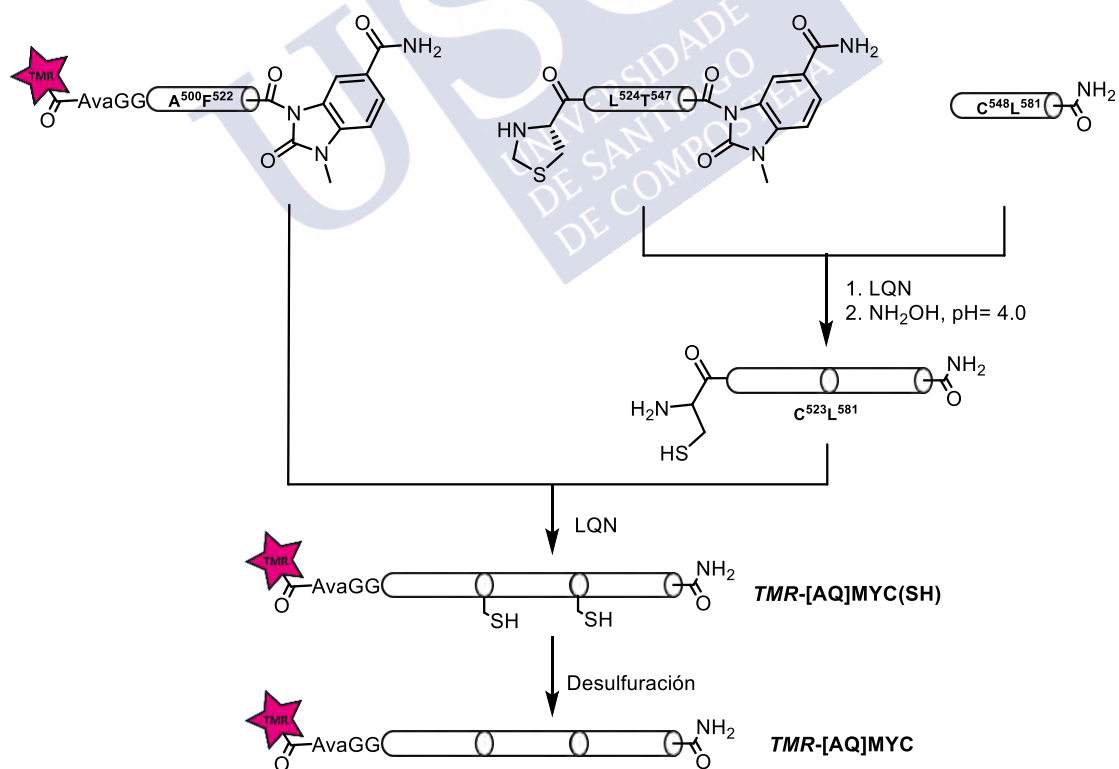


Figura 72. Cromatograma do produto puro TMR-[AQS]OMOMYC com método 5 % → 95 % de MeCN em 15 min.

### Unión dos fragmentos por LQN e desulfuración para obter TMR-[AQ]MYC

Para preparar o derivado de MYC marcado con rodamina seguiuise un esquema similar o caso anterior:



esquema 20. Esquema representativo da ruta sintética que produce TMR-[AQ]MYC.

O primeiro paso consistiu en unir o fragmento intermedio  $\text{Thz}^{523}\text{T}^{547}\text{-Nbz}$  co *N*-terminal  $\text{C}^{548}\text{L}^{581}$  por LQN (4.5 h , 37°C) para, de seguido, desprotexer ortogonalmente a tiazolidina por

tratamento con hidroxilamina a pH= 4 durante toda a noite. Ao día seguinte o cru de reacción analizouse, observándose o pico do produto e a desaparición case completa dos sustratos de partida. Pasouse pois, a elaborar a mestura. Cas condicións xa optimizadas, engadiuse un volume de tampón guanidinio 6 M, 200 mM de  $K_2HPO_4$  tal que triplicaba a dilución da disolución inicial. Esta mestura final, filtrouse a través de amicons e recuperouse a proteína desexada en auga e no mesmo tampón.

Despois disto, pasouse a purificar o recuperado da filtración nun HPLC semipreparativo onde puido separarse mellor o produto desexado das impurezas que formaban unha lomba co mesmo tempo de retención. A pesares disto, tan só se acadou un rendemento dun 20 % para este paso.

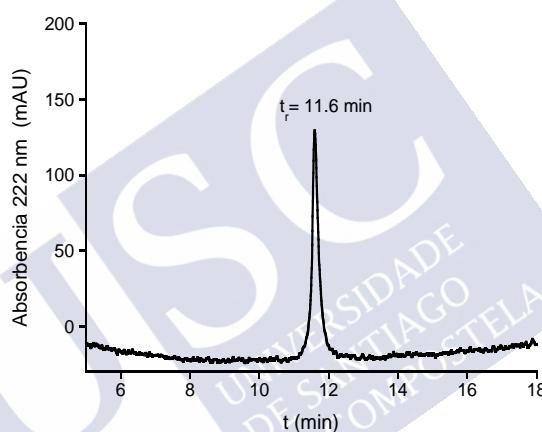


Figura 73. Cromatograma do produto da primeira LQN,  $C^{523}L^{581}$ , puro.

O segundo paso consistiu na creación da esqueleto peptídico completo de **TMR-[AQ]MYC(SH)**. Para o que se realizou unha segunda ligazón entre o precursor do tioéster **TMR-A<sup>1</sup>F<sup>23</sup>-Nbz** activado co MPAA e o produto puro da primeira LQN,  $C^{523}L^{581}$ . Unha vez rematada a reacción pasouse a filtrar o cru a través de amicons, recuperándose unha mestura moi limpa e fácil de purificar. Despois da purificación e correspondente liofilización, obtívose un sólido rosa que resultou ser o produto, cun rendemento excelente do 45 %.

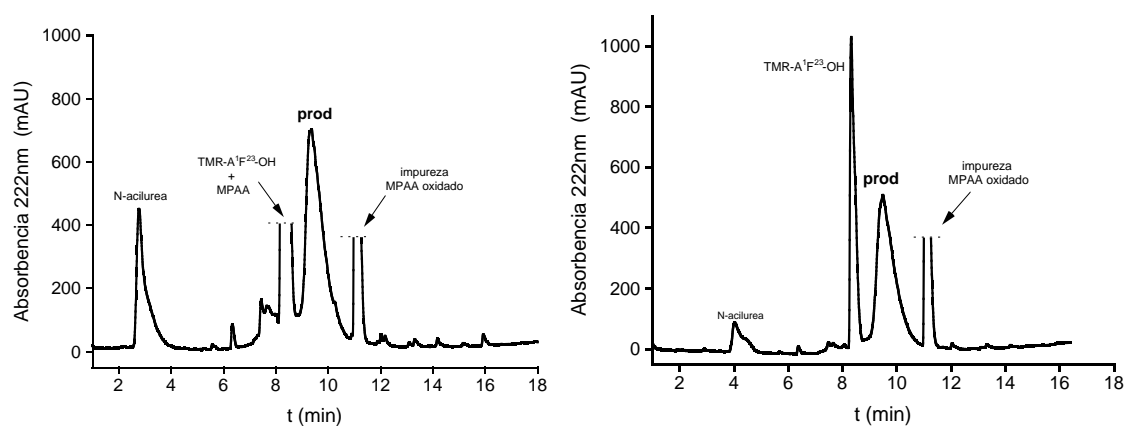


Figura 74. Esquerda: cru de reacción ás 4.5 h antes de ser filtrado con amicóns. Dereita: cru de reacción despois de pasar por amicóns. Método usado 5 % → 95 % de MeCN + 0.1 % TFA en 19 min.

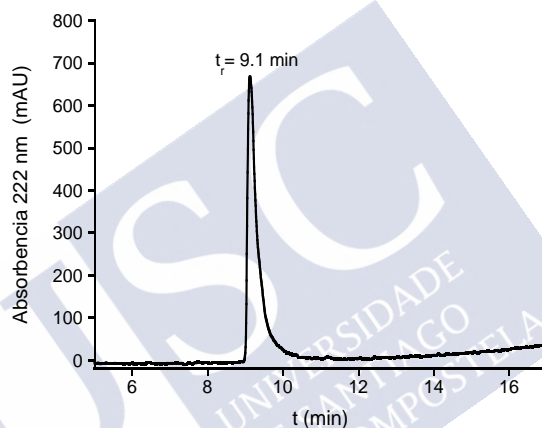


Figura 75. Cromatograma trala purificación de TMR-[AQ]MYC(SH).

Por último, levouse a cabo a desulfuración final na cal as dúas cistinas da miniproteína se convertiron en dúas alaninas, como na secuencia natural de MYC. Esta etapa final, acadouse cun rendemento do 70 %, obténdose o produto coa secuencia: **TMR-Ava-GGAVKRRTHN VLERQRRQELK RSFFALRDQIP ELENNEKAPKV VILKKATAY ILSVQAEEQKL ISEEDLLRKRR EQLKHKLEQL-NH<sub>2</sub>**

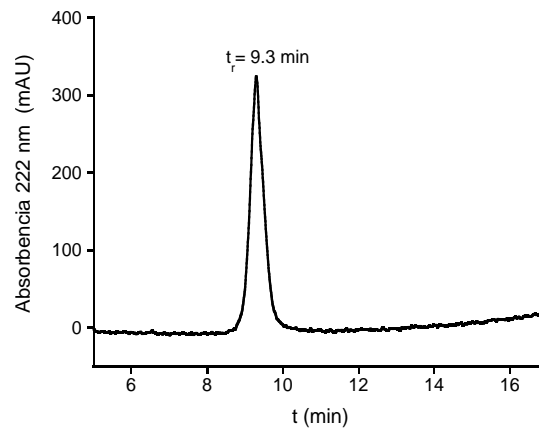


Figura 76. Cromatograma do produto TMR-[AQ]MYC puro.





### 3.8. FORMACIÓN DE PONTES DISULFURO EN MINIPROTEINAS MARCADOS CON RODAMINA

Como se comentou na introdución deste capítulo, no grupo de investigación demostrárase en experimentos preliminares que a formación dunha ponte disulfuro intramolecular na proteína con dúas cistinas inhibía a súa capacidade para unirse a ADN en cooperación con MAX. Sobre esta base, tratamos de realizar una oxidación similar nas miniproteínas que conteñen a rodamina, co fin de forma o derivado indicado na seguinte figura.

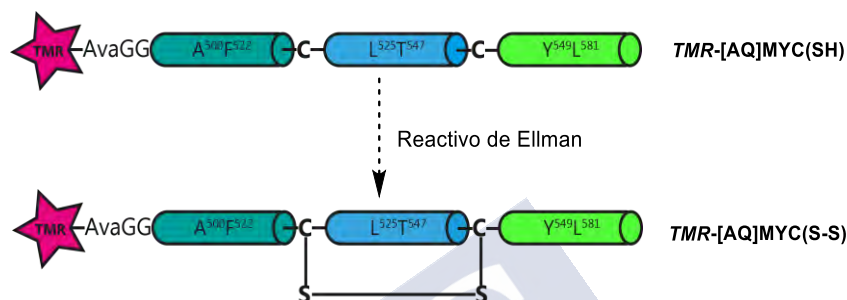


Figura 77. Representación da reacción de formación do disulfuro de TMR-[AQ]MYC(SH)

En primeiro lugar, probouse a oxidación de **TMR-[AQ]MYC(SH)** usando unha disolución 1 mM do péptido en 100 mM de tampón Tris·HCl a pH 7,5, e engadindo 10 mM de DTNB, ou reactivo de Ellman (ácido 5,5-ditiobis-(2-nitrobenzoico)). A mestura axitouuse a temperatura ambiente, burbullando ar e con axitación magnética. Despois de dúas horas, a reacción analizouse por HPLC-MS pero non se detectou o produto desexado nin tampouco o pico correspondente ao substrato de partida. A mesma reacción probouse con **TMR-[AQS]OMOMYC(SH)** e usando como control a mesma proteína sen o fluoróforo **[AQS]OMOMYC(SH)**. Mentres que para esta miniproteína control, ás 2 h de reacción xa se detectara a formación do disulfuro, para o caso do péptido con TAMRA non se observou a formación do produto nin ás 24 h. A pesares de que o cru de reacción aparece moi limpo para o caso dos péptidos sen marcar xa que tan só se diferencian dous picos, un correspondente ao péptido de partida e o outro ao produto de oxidación ( ver Figura 78), para o caso dos péptidos con fluoróforo non se observa o mesmo patrón. Para **TMR-[AQ]MYC(SH)** aprécianse dous picos no cromatograma do cru, os cales se purificaron e se analizaron por separado. Sorprendentemente, o pico con tempo de retención en torno a 7 min ten unha masa que coincide coa masa do fluoróforo TMR desancorado do resto do péptido. O espectro de masas que se

obtivo para o outro pico, con tempo de retención 9 min, non se puido asociar con ningún péptido coñecido nin con intermedios de reacción.

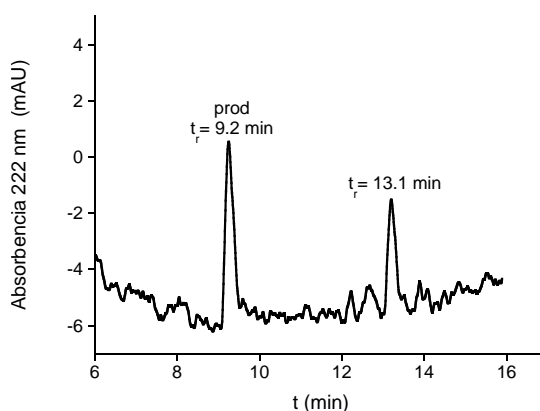


Figura 78. Cromatograma a 222 nm do cru de reacción de oxidación para obter [AQS]OMOMYC(SS). O pico con tempo de retención 13.1 min correspóndese coa substancia de partida. O de 9.2 min é o produto oxidado [AQS]OMOMYC(SS).

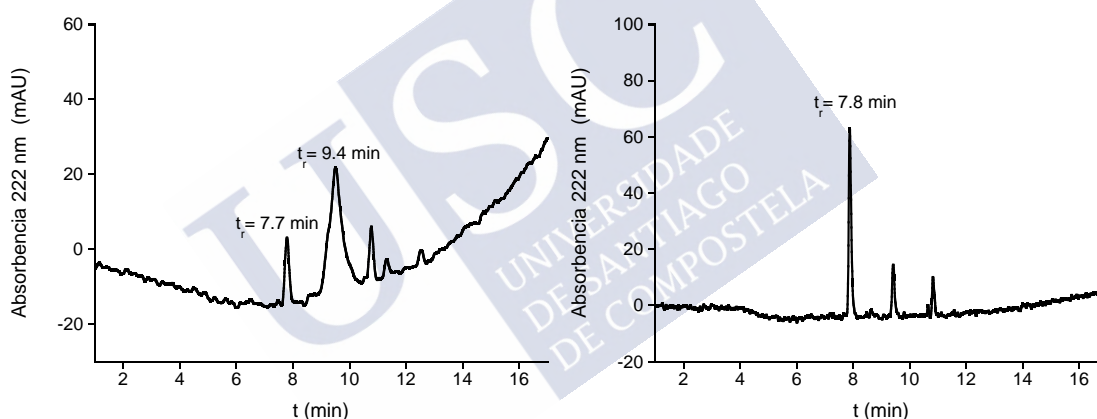


Figura 79. Esquerda: cromatograma a 222 nm do cru de reacción de oxidación a TMR-[AQS]OMOMYC(SS). Dereita: pico de 7,8 min despois da purificación. Ningún pico se corresponde ca substancia de partida nin co produto de reacción oxidado

Por mor destes resultados negativos, decidiuse facer un estudo de diferentes condicións para acadar a oxidación variando os tampóns e os pH. Así mesmo, probouse un método de obtención do S-S atopado en bibliografía engadindo como codisolvente DMSO. Este protocolo demostraba que engadindo un 10 % de DMSO ao tampón fosfato a pH 7.5 promovíase a oxidación dos tiois nun tempo de reacción reducido.<sup>111</sup> Desafortunadamente, ningunha das condicións probadas xerou o produto disulfuro para ningún dos dous análogos de MYC ou

<sup>111</sup> J. Atcher, I. Alfonso, *RSC Adv.*, **2013**, 3, 25605.

OMOMYC equipados con TMR, A seguinte táboa presenta un resumo das condicións que se probaron e os tampóns e variacións introducidas en cada proba.

Táboa 1. Optimización da reacción de oxidación para diferentes péptidos e diferentes condicións.

péptido	Tampón -pH	observacións
TMR-[AQ]MYC(SH)	Tris-HCl 7.5	Non se observa produto
TMR-[AQ]MYC(SH)	Tris-HCl 8.3	Non se observa produto
TMR-[AQ]MYC(SH)	Fosfato 7.5+10 % DMSO	Non se observa produto
TMR-[AOS]OMOMYC(SH)	Tris-HCl 7.5	Non se observa produto
TMR-[AOS]OMOMYC(SH)	Tris-HCl 8.3	Non se observa produto
TMR-[AOS]OMOMYC(SH)	Fosfato 7.5+10 % DMSO	Non se observa produto
[AOS]OMOMYC(SH)	Fosfato 7.5+10 % DMSO	Só péptido de partida
[AOS]OMOMYC(SH)	Tris-HCl 8.3	Produto
[AQ]MYC(SH)	Tris-HCl 7.5	Produto



### 3.9. ESTUDOS ESPECTROSCÓPICOS DE INTERACCIÓN CO ADN DAS MINIPROTEINAS SINTÉTICAS

Os péptidos [AQ]MYC e [AQS]OMOMYC obtidos por síntese química presentan alteracións na súa estrutura primaria, concretamente na rexión básica que reconece o ADN. Polo tanto, é preciso comprobar se estas proteínas sintéticas conservan as súas propiedades de pregamento e dimerización, así como de interacción co ADN. Para iso, realizáronse experimentos espectroscópicos de dicroísmo circular a temperatura 25 °C en tampón fosfato 10 mM a pH 7.4 con, 150 mM de KCl e 0.25 mM EDTA. 1 equivalente refírese a unha concentración na cubeta de 5  $\mu$ M de ADN ou de concentración de péptido monómero. O ADN de dobre cadea usado contén a secuencia diana denominada *E-box* (resaltada en cursiva na secuencia do oligo 5'- CTGTAGGC *CACGTG* ACCGGATG- 3', esta apareceu coa complementaria para facer os estudos).

No primeiro experimento, mostrado na **Figura 80**, comprobouse o pregamento e a interacción co ADN de [AQ]MYC

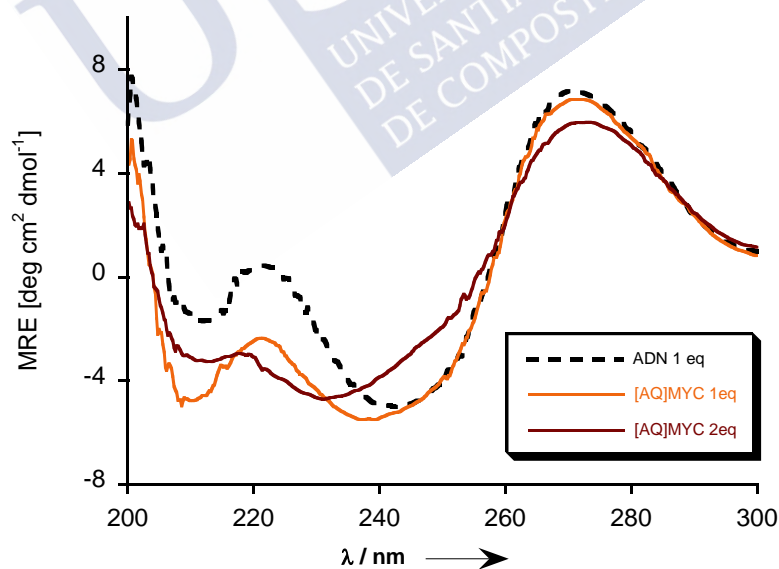
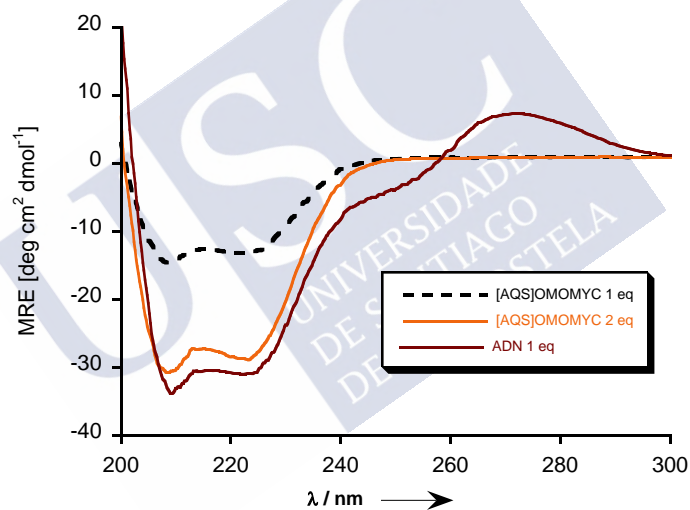


Figura 80 Medidas de dicroísmo circular para comprobar o pregamento de [AQ]MYC e a súa interacción co ADN consenso.

A liña punteada representa 1 equivalente de ADN. Obsérvase que a cantidade de hélice  $\alpha$  máxima acádase cando a relación entre [AQ]MYC : ADN é de 2:1, banda marrón. Mentres que

a adición dun equivalente do péptido [AQ]MYC sobre un equivalente de ADN (liña laranxa) forma unha banda con dous mínimos a 209 nm e 240 nm, o que indica que existen segmentos helicoidales illados pero a maior parte da proteína non está pregada. En ningún caso se observan bandas claras características das hélices alfa a 208 e 222 nm.<sup>112</sup> As bandas formadas son pouco intensas en tódolos casos e os mínimos en torno a 230-240 nm son anchos, polo tanto, conclúese que [AQ]MYC forma homodímeros inestables e amosa interacción débil co ADN, como era de esperar segundo a bibliografía.<sup>113</sup> A formación dos complexos está máis estabilizada en presenza do ADN diana, pois a banda a 240 sofre un desprazamento cara 230 nm.

No segundo experimento de DC (dicroísmo circular) estudouse a homodimerización de [AQS]OMOMYC.(Figura 81).



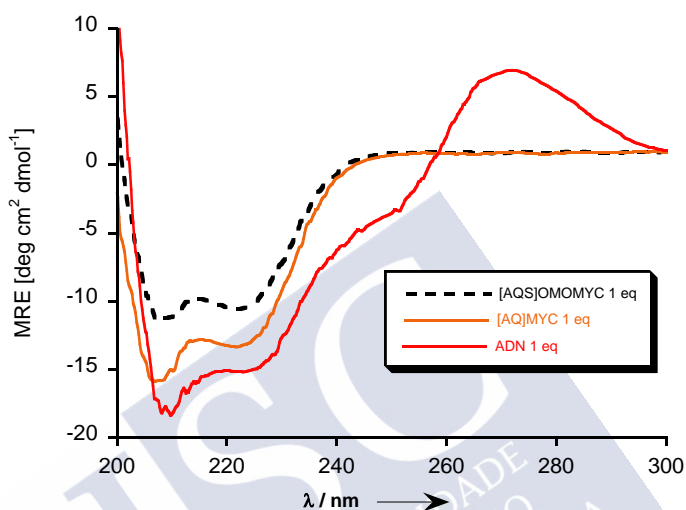
**Figura 81.** Experimento de dicroísmo circular para comprobar o pregamento e a interacción coa secuencia diana de ADN de [AQS]OMOMYC.

[AQS]OMOMYC por si mesmo xa contén dúas bandas a 208 nm e a 222 nm características dun pregamento en hélice alfa. Cando se engade outro equivalente deste péptido, o mínimo das bandas aumenta confirmándose que forman un homodímero estable. Ao adicionar un equivalente de ADN diana sobre esta disolución, as bandas aparecen máis pronunciadas

<sup>112</sup> S. M. Kelly, T. J. Jess, N. C. Price, *Biochim Biophys Acta*, **2005**, 1751, 119.

<sup>113</sup> A)M.E. Beaulieu, F.O. McDuff, V. Frappier, M. Montagne, J.F. Naud, P.Lavigne *J Mol Recognit.* **2012**; 25, 414. B) F.O. McDuff, J-F. Naud, M. Montagne, S. Sauvé, P. Lavigne, *J. Mol. Recognit.*, **2009**, 22, 261.

confirmándose que o homodímero se estabiliza ao interactuar co ADN diana. Pode concluírse que o [AQS]OMOMYC presenta unha estrutura secundaria máis definida, cun pregamento maior e máis estable que [AQ]MYC, formando heterodímeros en ausencia de ADN que recoñecen a rexión específica da *E-box* do ADN diana. Por último, analizouse a formación do heterodímero [AQS]OMOMYC/[AQ]MYC e a interacción deste co '*E-box*'.



**Figura 82.** Experimento de dicroísmo circular para comprobar a formación do heterodímero de [AQS]OMOMYC/[AQ]MYC e a súa interacción coa secuencia diana de ADN.

Sobre unha disolución que contiña 1 equivalente de [AQS]OMOMYC agregouse un equivalente da proteína [AQ]MYC, formándose unha nova banda cos dous mínimos a 208 nm e 222 nm que caracterizan ás hélices alfa e cunha forma diferente ás correspondentes a ningún homodímero. Deste xeito, esta nova banda (laranja) atribuíuse ao heterodímero [AQS]OMOMYC/[AQ]MYC. Se se amece un equivalente de ADN sobre o heterodímero, os mínimos aumentan a súa intensidade indicando unha interacción do complexo proteico co ADN.

Coa intención de ir máis aló, tentouse elucidar que complexo é máis estable, se ben o homodímero [AQS]OMOMYC/[AQS]OMOMYC ou o heterodímero [AQS]OMOMYC/[AQ]MYC. Polo tanto, sobre unha mestura que contiña 1 equivalente do heterodímero [AQS]OMOMYC/[AQ]MYC con 1 equiv de ADN diana, agregouse exceso de

[AQS]OMOMYC monomérico formándose unha banda moi similar en aspecto e intensidade ao homodímero de [AQS]OMOMYC. A pesares de engadirle máis equivalentes do péptido monomérico [AQ]MYC, a banda que se observaba en intensidade e aspecto seguía semellando a do homodímero. Polo tanto, deste experimento certificase que o heterodímero [AQS]OMOMYC/[AQ]MYC se forma de maneira estable e recoñece o ADN. A maiores, cando 1 equivalente de ADN se trata con exceso de [AQS]OMOMYC ou [AQ]MYC, a banda resultante semella máis próxima á do homodímero [AQS]OMOMYC que á do heterodímero [AQS]OMOMYC/[AQ]MYC.

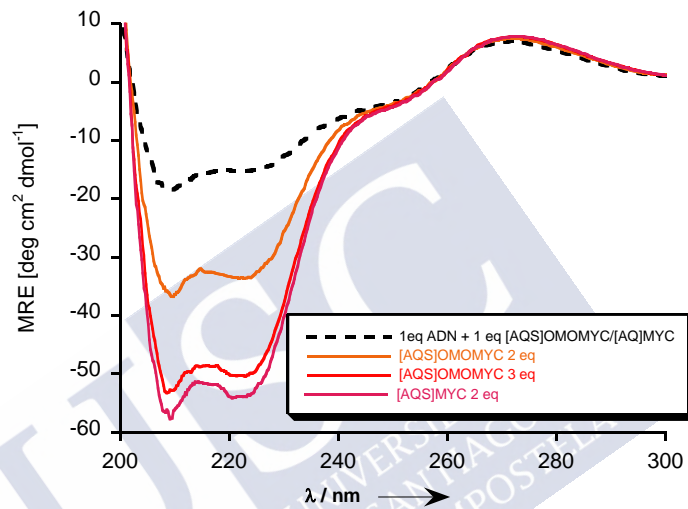


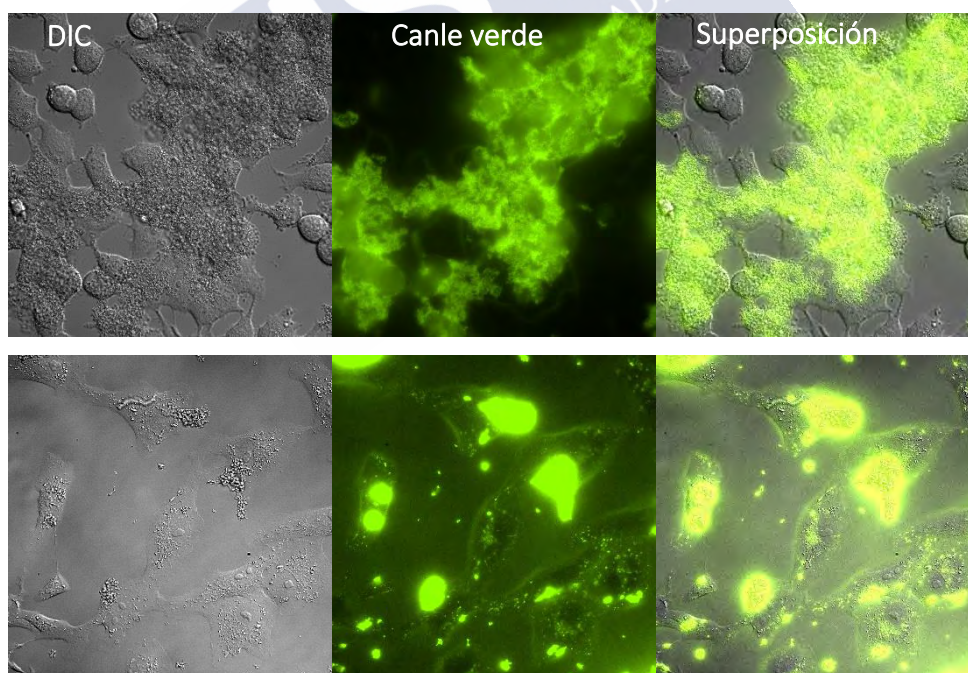
Figura 83. DC que estuda a variación na banda do complexo [AQS]OMOMYC/[AQ]MYC/ADN fronte a exceso de [AQS]OMOMYC ou [AQ]MYC.



### 3.10. ESTUDIO DE INTERNALIZACIÓN CELULAR DAS MINIPROTEINAS FLUORESCENTES

Unha vez sintetizados e caracterizadas os derivados das oncoproteína MYC e OMOMYC, leváronse a cabo experimentos de internalización celular. O grupo da doutora Laura Soucek descubriu que OMOMYC marcado con fluoresceína no extremo C-terminal é quen de penetrar en células vivas de tipo A549 en 2 h, cunha concentración 32  $\mu$ M e incluso chegar, en certa medida, ao núcleo. Para a análise, fixan os cultivos celulares cun 3 % de PFA (paraformaldehído) e analizan con microscopía confocal a fluorescencia das células.<sup>114</sup>

Deste xeito, probamos a incubar cultivos de diferentes liñas celulares cos nosos péptidos sintéticos. Para as probas iniciais usáronse os derivados con fluoresceína, e que teñen as dúas cistinas, tales como [AQK]MYC(SH)-FLU. Para estes primeiros ensaios utilizamos células de tipo HeLa (células de cancro cérvico-uterino humanas) e VeRo (células epiteliais de ril de mono verde africano)



*Ilustración 1. Incubación de diferentes liñas celulares durante 30 min con 5  $\mu$ M de [AQK]MYC(SH)-FITC Arriba: células HeLa. Abaixo: células VeRo.*

<sup>114</sup> Soucek, L.; Beaulieu, M.-E. (2014), España, *Methods and Compositions for the Treatment of Cancer*, WO 2014/180889 A1.

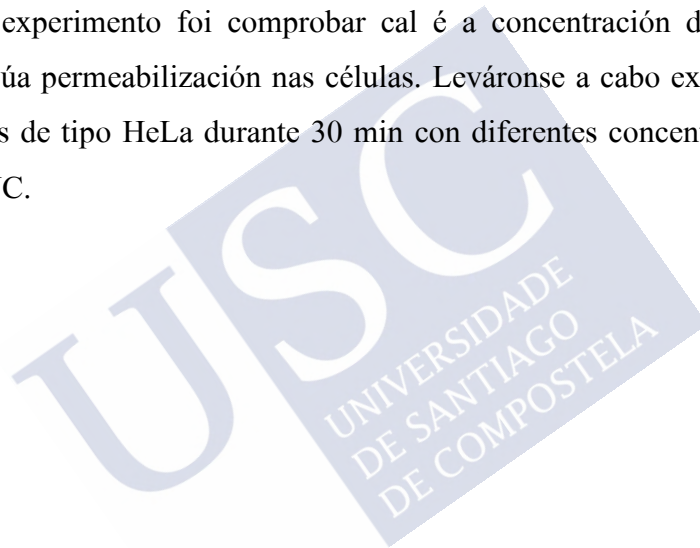


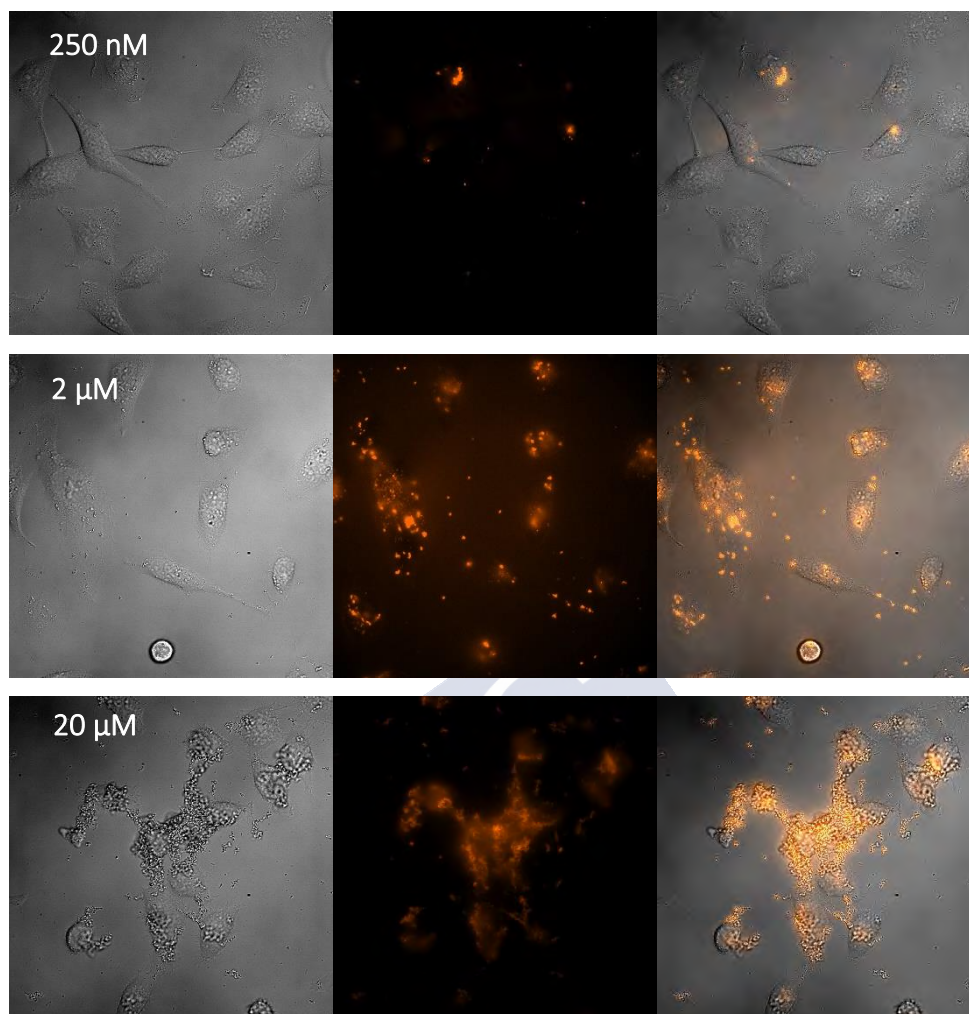
As células incubáronse con 5  $\mu$ M de [AQK]MYC(SH)-FITC en DMEM sen FBS durante 30 min, lavadas tres veces con medio fresco e analizadas baixo o microscopio de fluorescencia. Como pode observarse na ilustración 18, a concentración usada para estas proteínas foi demasiado elevada observándose unicamente agregados do péptido que chegaban a cubrir as células. Aínda así, o péptido era capaz de penetrar no interior destas.

### **3.10.1. Estudos internalización a diferentes concentracións das miniproteínas marcadas**

Os seguintes estudos de internalización realizáronse cas proteínas marcadas con rodamina, pois téñense as dúas versións de cada péptido (con cistinas libres e reducidas).

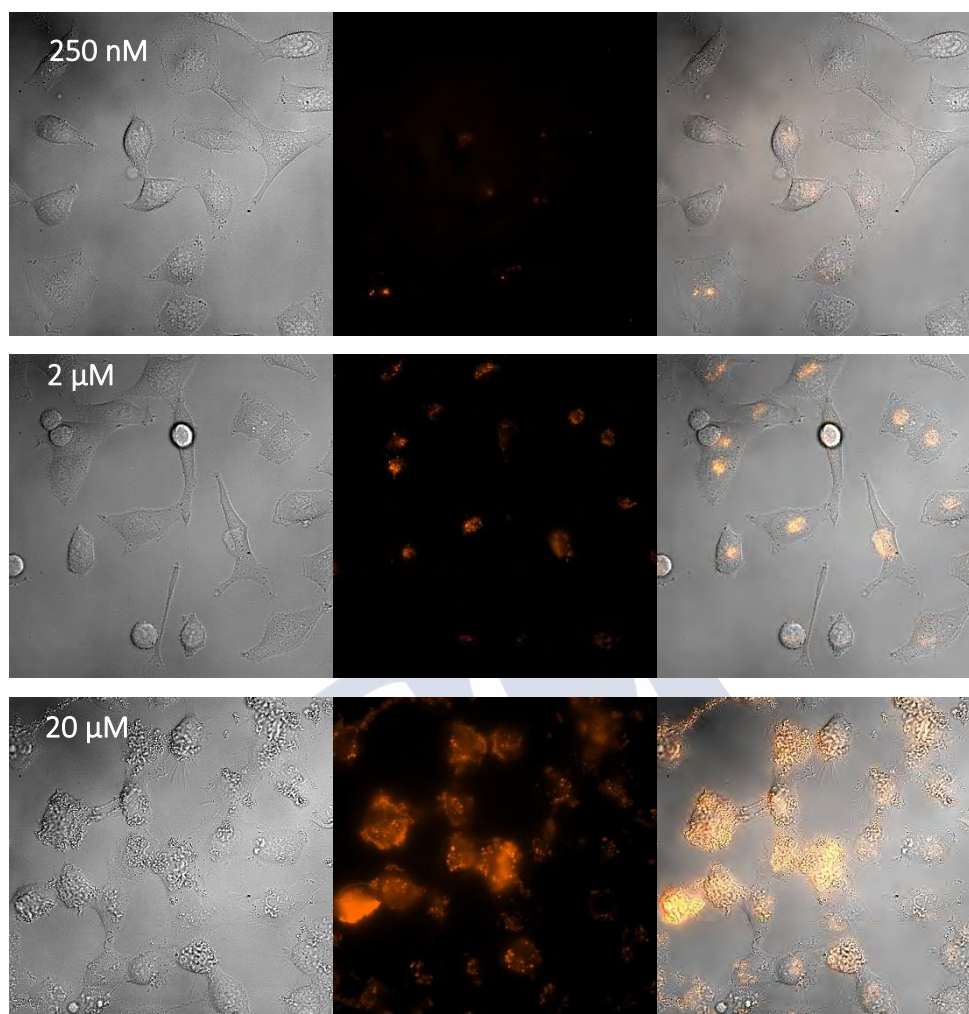
O primeiro experimento foi comprobar cal é a concentración dos péptidos marcados axeitada para a súa permeabilización nas células. Leváronse a cabo experimentos nos que se incubaron células de tipo HeLa durante 30 min con diferentes concentracións dos derivados marcados de MYC.





**Ilustración 2.** Incubación de **TMR-[AQ]MYC(SH)** en células *HeLa* durante 30 min a diferentes concentraciones, en DEMEM sen FBS. Esquerda: brightfield. Medio: canle vermella. Dereita: superposición das dúas canles. Fotografías tomadas no microscopio de fluorescencia.

A pesares de que os péptidos entran na célula a tódalas concentracións testadas, obsérvase a formación de agregados de gran tamaño dependendo da concentración usada. A 20  $\mu$ M hai demasiada agregación dos péptidos, o que se reflexa nun efecto drástico de morte celular. Para 2  $\mu$ M tamén se aprecian agregados aínda que en menor medida, nesta concentración as proteínas penetran no interior celular, mais a integridade destas vese algo comprometida, probablemente pola alta concentración do péptido. Polo tanto, a concentracións de baixo micromolar e nanomolar é cando se produce a entrada máis limpa, con menor agregación e na que a membrana celular non se ve tan afectada. Destas imaxes dedúcese tamén que o tempo de incubación non foi suficiente para que o péptido penetrara ben na célula e que, a simple vista, **TMR-[AQ]MYC(SH)** tende a agregar máis que o seu análogo sen SH.



**Ilustración 3.** Incubación de TMR-[AQ]MYC en células HeLa durante 30 min a diferentes concentracións, en DMEM sen FBS. Esquerda: brightfield. Medio: canle vermella. Dereita: superposición das dúas canles. Fotografías tomadas no microscopio de fluorescencia.

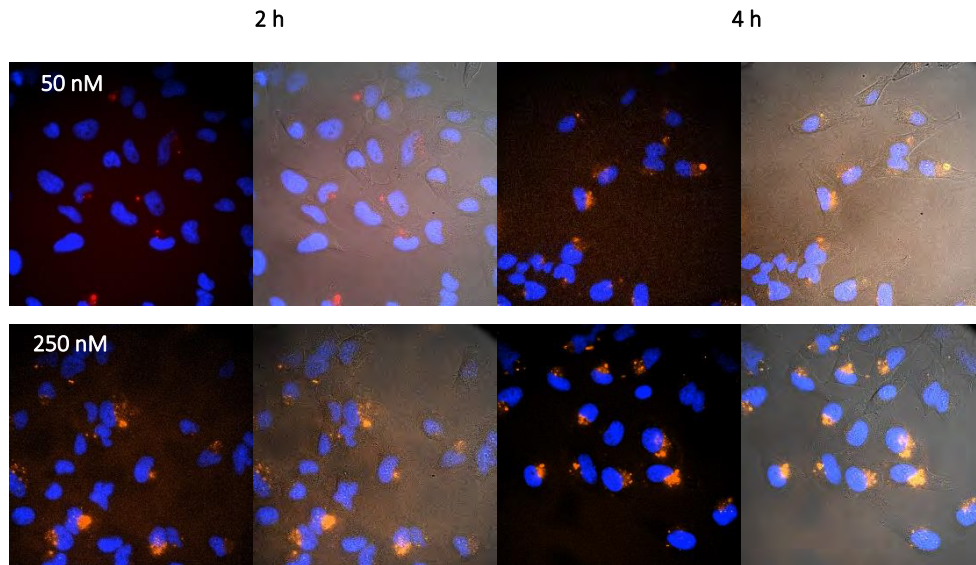
### 3.10.2. Estudos de internalización a diferentes tempos de incubación

Para coñecer o tempo de incubación axeitado dos péptidos marcados, realizáronse diferentes experimentos ca mesma concentración de péptido e a igual condicións, pero variando o tempo de incubación.

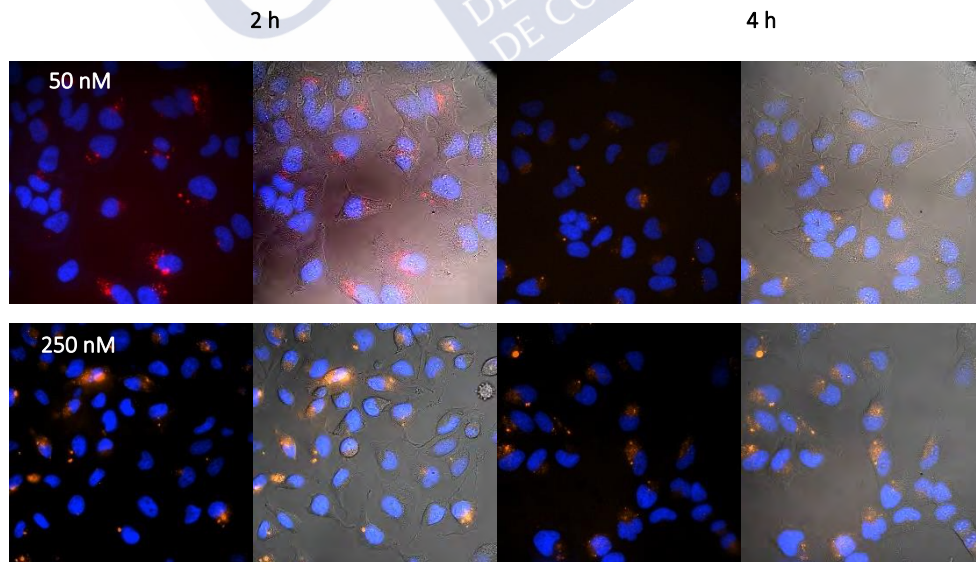
Do apartado anterior concluíuse que a internalización debía realizarse a concentracións de nanomolar. Por isto, se fixeron probas con concentracións de 50 e 250 nM dos derivados de MYC. Como se mostra nas **Ilustración 4 e Ilustración 5** a 250 nM obsérvase mellor penetración das proteínas marcadas con 4 h que con 2 h de incubación. Cando os péptidos se



incubando a 50 nM tamén se aprecia internalización. Mais neste caso, a fluorescencia é tenue, precisando do aumento da exposición e intensidade do láser na toma das fotos. Por mor disto, decidiuse que a concentración óptima é de 250 nM.

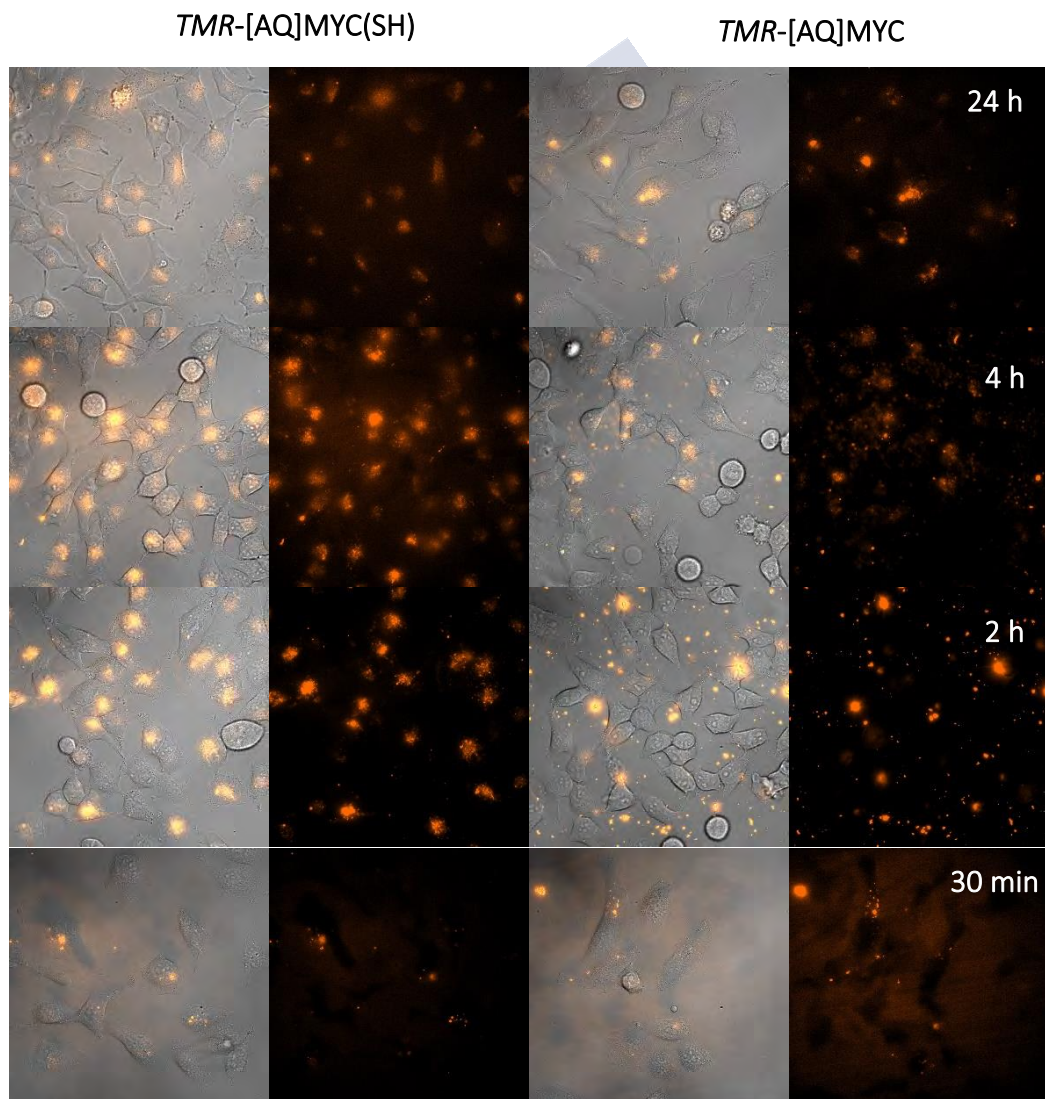


**Ilustración 4.** Células HeLa incubadas con **TMR-[AQ]MYC(SH)** en DMEM sen FBS. Os núcleos foron tinguidos con 10  $\mu$ M de HOESCHT durante 15 minutos antes de lavar con DMEM fresco.



**Ilustración 5.** Células HeLa incubadas con **TMR-[AQ]MYC** en DMEM sen FBS. Os núcleos foron tinguidos con 10  $\mu$ M de HOESCHT durante 15 minutos antes de lavar con DMEM fresco

Unha vez atopada a concentración axeitada, leváronse a cabo medidas de incubación de péptidos a diferentes tempos, fixando a concentración en 250 nM. Nestes experimentos observouse que a tempos de incubación curtos, os péptidos non son capaces de penetrar correctamente, atopándose moi pouco péptido no interior. Para tempos de incubación de 2 h, boa cantidade de proteína marcada está no interior da célula pero hai bastantes agregados. Namentres, se se deixan tempos de incubación maiores, de 4 h, non se aprecian agregados, os péptidos entran mellor, a luminiscencia observada é máis homoxénea e a distribución máis regular do péptido pola célula, asociándose este fenómeno á menor agregación. Amais, 4 h é o tempo de incubación co que se observa máis cantidade de péptido internalizado.

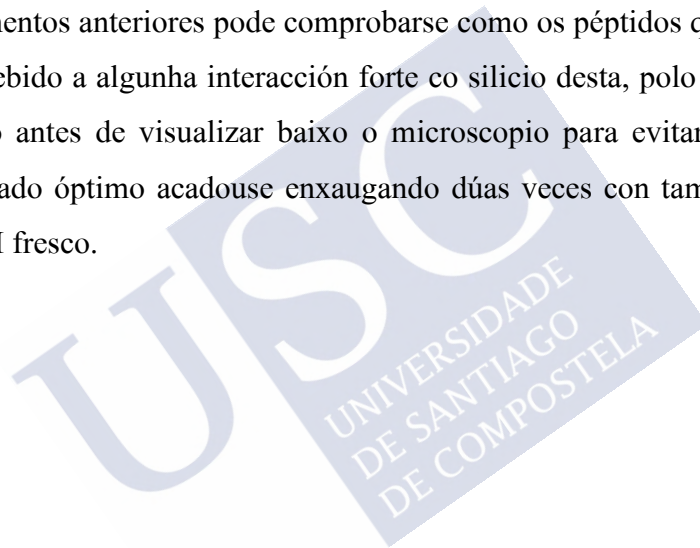


**Ilustración 6.** Incubación de células HeLa a diferentes tempos con 250 nM de péptido marcado, en DMEM sen FBS.

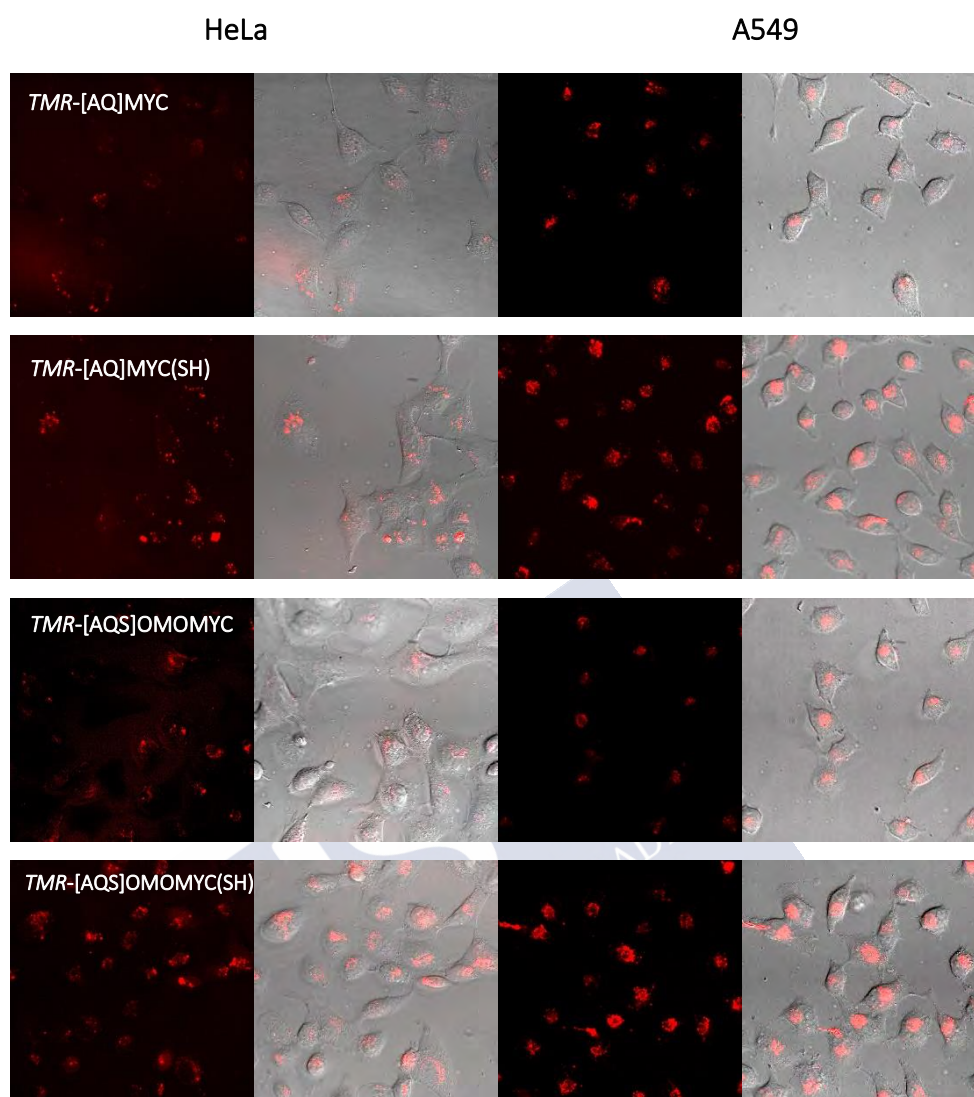
Os ensaios nos que se trataron os cultivos con disolucións de proteínas marcadas durante 24 h, presentaron menos agregados, boa penetración dentro do ambiente celular por parte de ámbolos dous péptidos e unha distribución homoxénea destes péptidos no interior da célula. Así mesmo, hai menor ruído de fluorescencia grazas a que as proteínas non se pegan tanto ao fondo da placa de vidro. Mais este tempo de incubación relativamente longo, pode conlevar á degradación das proteínas marcadas. Polo tanto, a duración óptima de incubación fixouse como 4 h a unha concentración 250 nM dos péptidos marcados.

### **3.10.3. Estudos de internalización en diferentes liñas celulares e comparativa**

Dos experimentos anteriores pode comprobarse como os péptidos quedan moi pegados ao vidro da placa debido a algunha interacción forte co silicio desta, polo que se debe lavar moi ben cada pocillo antes de visualizar baixo o microscopio para evitar o ruído de fondo. O protocolo de lavado óptimo acadouse enxaugando dúas veces con tampón PBS e tres veces máis con DMEM fresco.





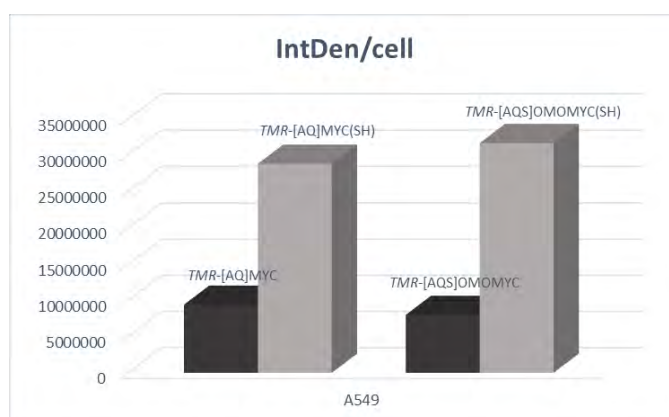


**Ilustración 7.** Incubación de 250 nM dos péptidos marcados en DMEM sen FBS durante 4 h. Comparación entre diferentes liñas celulares. Amósase a canle vermella e a superposición entre o DIC e a canle vermella.

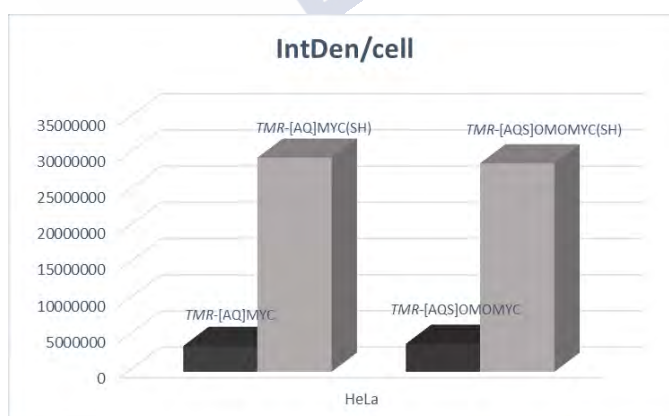
Para realizar este estudo comparativo e estimar a relación de penetración dunha proteína con respecto doutra, elimináronse os valores de ruído de fluorescencia de fondo, polo que só se tiveron en conta os datos dos píxeles das sinais do primeiro plano. Aínda que se pode observar perfectamente a simple vista, cando se analizou cuantitativamente a internalización relativa das proteínas, determinouse que aquelas que conteñen os grupos SH das Cys libres, son internalizados máis eficientemente que os seus análogos desulfurados.<sup>115</sup> Este feito é

<sup>115</sup> N. Chuard, G. Gasparini, D. Moreau, S. Lörcher, C. Palivan, W. Meier, N. Sakai, S. Matile, *Angew. Chem. Int. Ed* 2017, **56**, 2947.

consistente con antecedentes bibliográficos nos que se afirman que o intercambio tiol-disulfuro na membrana celular favorece a introdución de cargos moleculares.<sup>116</sup> A maiores, os análogos de OMOMYC adoitan penetrar mellor que os de MYC, formando vesículas máis homoxéneas e redondeadas arredor da rexión perinuclear.



**Figura 84.** Comparación da densidade de fluorescencia integrada de cada péptido marcado internalizado en células A459 por cada célula. **TMR-[AQ]MYC(SH)** penetra arredor dunhas 3 veces máis que o seu análogo sen cistinas. Amais, **TMR-[AQS]OMOMYC(SH)** penetra unhas 4 veces mellor que **TMR-[AQS]OMOMYC**.



**Figura 85.** Comparación da densidade de fluorescencia integrada de cada péptido marcado internalizado en células HeLa por cada célula. **TMR-[AQ]MYC(SH)** internaliza de

<sup>116</sup> a) X. Meng, T. Li, Y. Zhao, C. Wu, *ACS Chem. Biol.* 2018, **13**, 3078; b) T. Li, W. Gao, J. Liang, M. Zha, Y. Chen, Y. Zhao, C. Wu, *Anal. Chem.* 2017, **89**, 8501; c) N. Chuard, G. Gasparini, D. Moreau, S. Lörcher, C. Palivan, W. Meier, N. Sakai, S. Matile, *Angew. Chem. Int. Ed Engl.* 2017, **56**, 2947.



7 a 8 veces máis que o seu análogo sen cistinas. Amais, **TMR-[AQS]OMOMYC(SH)** penetra unhas 7 veces mellor que **TMR-[AQ]OMOMYC**.

Destes experimentos conclúese que pequenas variacións na secuencia peptídica de MYC pode ter consecuencias drásticas nas súas propiedades de internalización e agregación, así como na distribución dentro do ambiente celular.

#### 3.10.4. Localización

Estudos preliminares coas células HeLa indicaban que os péptidos sintetizados se quedaban en vesículas moi preto do núcleo. Para corroborar esta hipótese realizáronse incubacións das células que contiñan os polipéptidos marcados cun marcador de núcleo. O escollido foi o Hoescht 33342 que se trata dunha bis-benzamidina que se une ao suco menor de ADN de dobre cadea e que posúe fluorescencia na rexión azul. A marcación do núcleo confirmou que os péptidos se quedaban en vesículas arredor do núcleo. A maiores, coincubáronse as células con cloroquina, un derivado da quinoleína que serve para rebentar as vesículas e comprobar daquela, se as miniproteínas migran até acadar o núcleo cando non se atopan encapsuladas. Estes experimentos realizáronse tanto en células HeLa como en A549 mais en ningún dos casos se acadou a lise completa das vesículas. A cloroquina favorece que as vesículas teñan maior tamaño e sexan máis brillantes tanto para o caso de **TMR-[AQ]MYC** como para o derivado con cistinas.

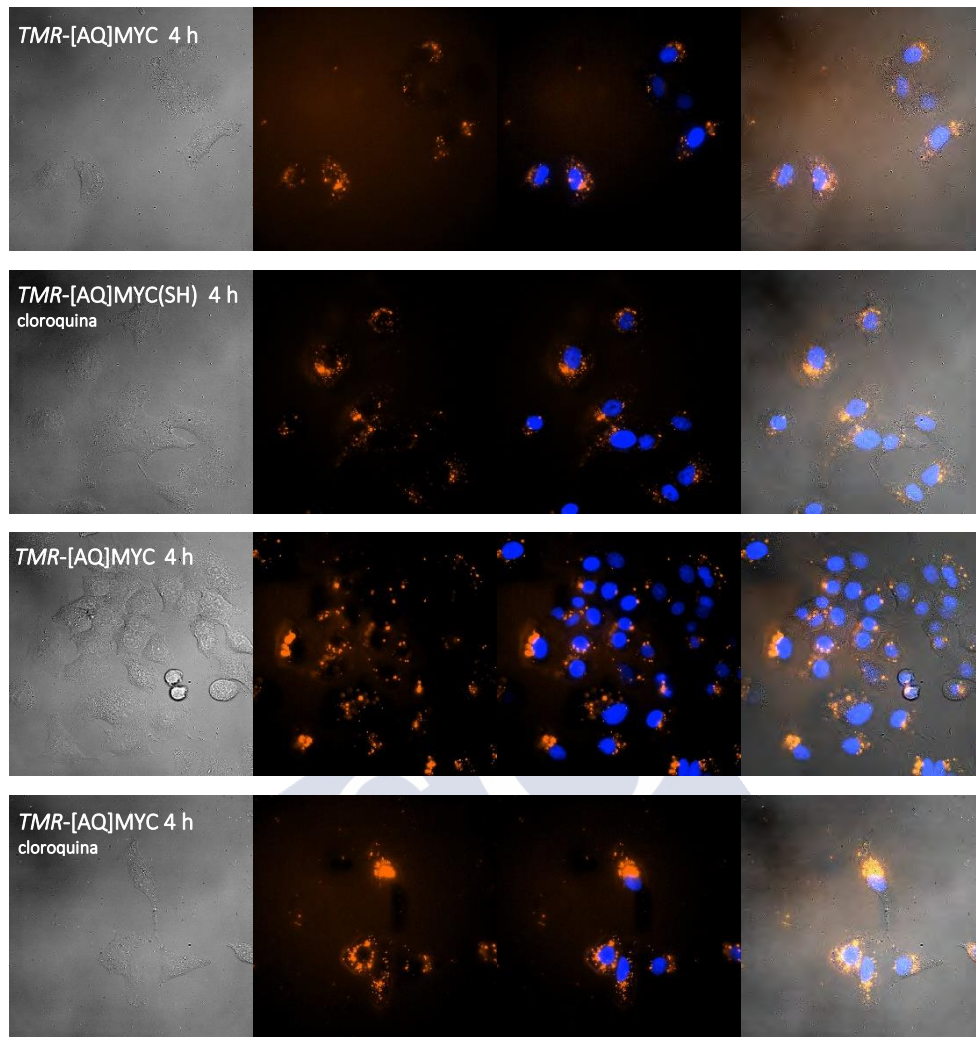


Ilustración 8. Células A549 incubadas con 250 nM en DMEM sin FBS dos péptidos marcados. O núcleo tinguiuse de azul incubando as células cunha disolución 10  $\mu$ M de HOESCHT 15 min antes dos lavados. A 2<sup>o</sup> e 4<sup>o</sup> liña representan a incubación das células con 100  $\mu$ M de cloroquina durante 4 h.

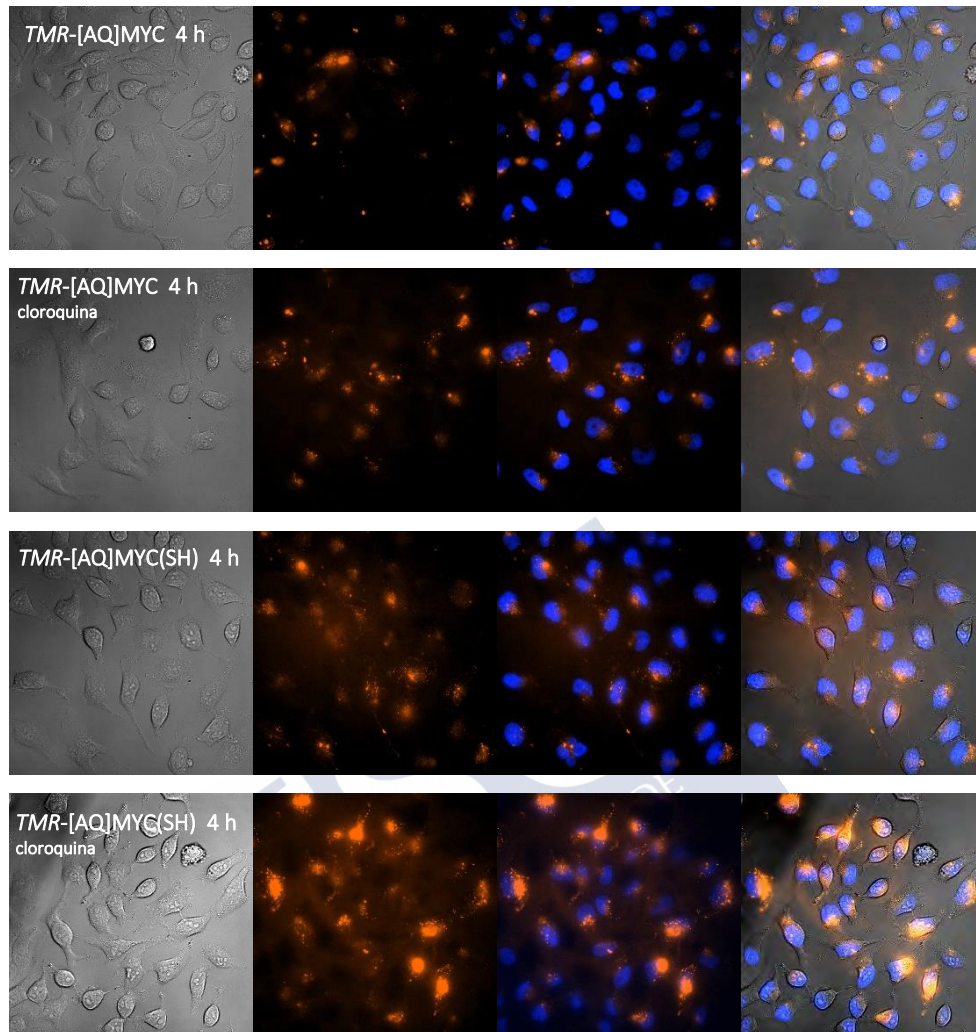


Ilustración 9. Células HeLa incubadas con 250 nM en DMEM sen FBS dos péptidos marcados. O núcleo tinguiuse de azul incubando as células cunha disolución 10  $\mu$ M de HOESCHT 15 min antes dos lavados. A 2<sup>o</sup> e 4<sup>o</sup> liña representan a incubación das células con 100  $\mu$ M de cloroquina durante 4 h

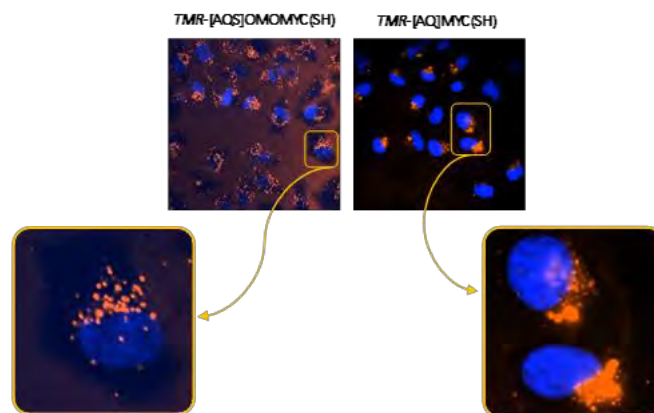


Ilustración 10. Superposición da canle vermella e azul da internalización dos derivados con cistinas libres. Móstrase unha ampliación dun cacho da imaxe na que se aprecia a localización perinuclear das vesículas dos péptidos.

### 3.11. CONCLUSIÓNS

- Sintetizáronse once dominios de unión a ADN das oncoproteínas MYC e OMOMYC seguindo dúas estratexias de retrosíntese diferentes. Nalgúns casos obténdose rendementos excelentes.
- [AQ]MYC interacciona co ADN diana pero de xeito moi débil, formando homodímeros moi inestables.
- [AQS]OMOMYC forma homodímeros con estrutura de hélice alfa estables tanto en presenza como en ausencia do seu ADN que recoñecen a rexión específica *E-box*.
- Tanto os derivados marcados de MYC como de OMOMYC presentan unha excelente permeabilidade celular a concentracións de nanomolar, nunca antes reportado tan baixa concentración.
- Os péptidos amosan unha localización perinuclear chegando a penetrar no núcleo en casos concretos.
- A presenza de cistinas libres nas estruturas favorece a entrada na célula.
- Os derivados de OMOMYC forman vesículas endosómicas máis homoxéneas e mellor formadas que os derivados de MYC, xerando unha internalización máis limpa e eficiente para estes derivados.

Os resultados máis relevantes deste capítulo foron publicados en *Org. Biomol. Chem.*, **2019**, *17*, 6748.

**4. CAPÍTULO SEGUNDO: CONTROL DO RECOÑECIMENTO  
DO SUCO MENOR DO ADN CON MOLÉCULAS PEQUENAS  
FOTOMODULABLES**





#### 4.1. O GRUPO AZOBENCENO: UN SISTEMA FOTOCONMUTABLE

Os azobencenos son moléculas fotocromáticas que se caracterizan por ter un dobre enlace N=N no que cada nitróxeno está unido a un anel aromático.<sup>117</sup> Este grupo funcional outórgalle a propiedade de existir en dúas formas isoméricas diferentes, facilmente interconvertibles entre elas: as formas *trans* (ou *E*) e *cis* (ou *Z*).<sup>118</sup> A forma *trans* é termodinamicamente máis estable. Irradiando coa lonxitude de onda axeitada, normalmente no ultravioleta, pode pasarse do isómero *trans* ao *cis*. O cambio do *cis* ao *trans* faise quentando ou con luz visible.<sup>119</sup>

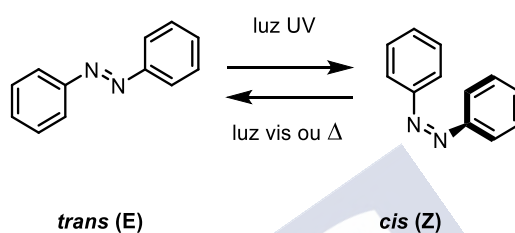


Figura 86. Representación da fotoisomerización do azobenceno.

A isomerización provoca cambios nas propiedades físicas destas moléculas. A variación na xeometría e disposición espacial e nos espectros de absorción son, sen dúbida, os máis notorios. O isómero *trans* asume unha disposición plana, mentres que un dos aneis bencénicos do *cis* rota un ángulo de 56° con respecto do outro para evitar repulsións estéricas. Esta disposición provoca que aumente o momento dipolar do isómero *Z*.

Estes foto interruptores teñen estruturas simples, un rendemento cuántico relativamente alto, e bastante resistencia á fatiga, sendo capaces de aguantar intactos numerosos ciclos de isomerización.<sup>119</sup> Estas vantaxes lévanos a ser dos interruptores moleculares máis usados hoxe en día. A pesares disto, o mecanismo polo cal un isómero se converte noutro non está aínda claro e barállanse varias hipóteses.<sup>120</sup>

<sup>117</sup> G. S. Hartley, *Nature* **1937**, *140*, 281.

<sup>118</sup> J. Griffiths, *Chem. Soc. Rev.* **1972**, *1*, 481.

<sup>119</sup> a) H. Rau, E. Lueddecke, *J. Am. Chem. Soc.* **1982**, *104*, 1616. b) H. M. D. Bandarab, S. C. Burdette, *Chem. Soc. Rev.* **2012**, *41*, 1809.

<sup>120</sup> a) E. Kazuma, M. Han, J. Jung, J. Oh, T. Seki, Y. Kim, *J. Phys. Chem. Lett.* **2015**, *6*, 4239. b) T. Fujino, S. Y. Arzhantsev, T. Tahara, *J. Phys. Chem. A* **2001**, *105*, 8123.

#### 4.1.1. As bandas de absorción.

O espectro de absorción dos azobencenos caracterízase por ter dúas bandas ben definidas para cada isómero. Por un lado, presentan unha banda moi intensa correspondente á transición  $\pi-\pi^*$  na rexión UV, en torno aos 320 nm para o *trans* e a 280 nm no caso do *cis*. A outra banda, de menor intensidade, asóciase á transición  $n-\pi^*$  e aparece na rexión visible arredor dos 440 nm. Esta banda coincide para ambas formas, sendo máis intensa para a forma *cis*.

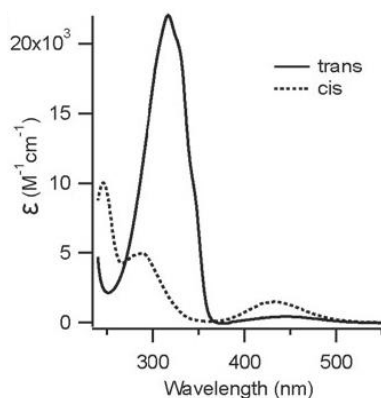


Figura 87. Espectro de absorción dos isómeros *cis* e *trans* do azobenceno. Imaxe adaptada da referencia[121]

#### 4.1.2. Isomerización de azocompostos para aplicacións en sistemas biolóxicos

Os espectros de absorbancia dos dous isómeros atópanse superpostos como se viu na figura anterior, isto contribúe a que a fotorreversión con luz visible non sexa completa. Un problema engadido que limita a súa aplicación en biomedicina deriva da necesidade de usar luz UV para inducir o desprazamento *trans*→*cis*. Este tipo de luz é daniño para medios biolóxicos ademais de non penetrar ben en tecidos.<sup>122</sup>

Modificar os substituíntes dos aneis aromáticos provoca cambios nos espectros electrónicos desprazando as bandas de absorción dos isómeros. Nos últimos anos, esta estratexia utilizouse para crear novos azocromóforos que poden ser isomerizados en condicións máis suaves con luz visible ou, incluso, con luz no infravermello.<sup>123,139</sup>

<sup>121</sup> A. A. Beharry, G. A. Woolley, *Chem. Soc. Rev.* **2011**, *40*, 4422.

<sup>122</sup>A) A. M. Smith, M. C. Mancini, S. Nie, *Nat. Nanotechnol.* **2009**, *4*, 710. B) C. Ash, M. Dubec, K. Donne, T. Bashfor *Lasers Med. Sci.* **2017**; *32*, 1909.

<sup>123</sup> H. A. Wegner, *Angew. Chem. Int. Ed.* **2012**, *51*, 4787.



A forma clásica de acadar espectros de absorción de azos con bandas desprazadas cara o infravermello é creando sistemas do tipo estilbena, tamén coñecidos como azos *push-pull*. Estes azoderivados fórmanse introducindo un grupo dador a un lado do núcleo azoico e un grupo aceptor no outro extremo. A estrutura resultante está moi polarizada, dando lugar ao desprazamento cara o infravermello da banda  $\pi-\pi^*$  que pode chegar, incluso, a solaparse coa  $n-\pi^*$ . O problema deste tipo de sistema reside no tempo de vida media corto dos isómeros *Z*. Nalgúns casos non chega nin a detectarse a isomerización porque a relaxación térmica do isómero *cis* tarda a penas nanosegundos.<sup>124</sup>

Estratexias máis recentes para isomerizar azocompostos con luz visible ou infravermella céntranse en separar as bandas  $n-\pi^*$  dos dous isómeros funcionalizando os aneis aromáticos do azobenceno. Se se acada a separación suficiente, a isomerización cara ámbalas dúas direccións poderá facerse irradiando sobre estas bandas, ou o que é o mesmo, empregando luz visible.

Un dos precedentes máis importantes neste eido foi a creación de azos tetrametoxisubstituídos levado a cabo no grupo de Woolley.<sup>125</sup> Introducendo catro grupos metoxilo nas posicións *orto* ao enlace N=N conseguiron desprazar a banda  $n-\pi^*$  do isómero *E* cara o vermello. O equilibrio destes derivados pode desprazarse para acadar un 80 % da forma *Z* irradiando con luz verde e un 85 % da forma *E* irradiando con luz azul. Segundo os autores, a forma non planar *trans* destes metoxiderivados aumenta a enerxía do orbital  $n$  por mor da repulsión electrónica que exercen os pares de electróns solitarios dos átomos de O e N que se atopan próximos, desprazando a súa banda  $n-\pi^*$  cara o vermello. Pola contra, a xeometría do isómero *cis* separa os grupos metoxilo dos pares electrónicos non enlazantes do nitróxeno, inhibido esta repulsión e mantendo a banda deste isómero sen variacións en comparación ao azobenceno simple. A maiores, devanditos azoderivados presentaron un tempo de vida media da forma *Z* de 14 días a temperatura ambiente.

<sup>124</sup> J. Garcia-Amoros, D. Velasco, *J. Org. Chem.* **2012**, *8*, 1003.

<sup>125</sup> a) S. Samanta, A. A. Beharry, O. Sadvovski, T. M. McCormick, A. Babalhavaeji, V. Tropepe, G. A. Woolley, *J. Am. Chem. Soc.* **2013**, *135*, 9777; b) A. A. Beharry, O. Sadvovski, G. A. Woolley, *J. Am. Chem. Soc.* **2011**, *14*, 19684.

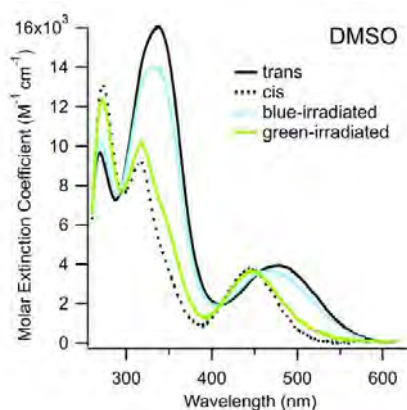
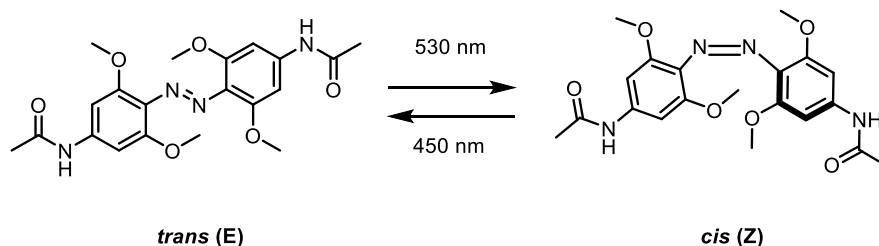


Figura 88. Arriba: reacción de isomerización dos tetrametoxibencenos. Abaixo: Espectro absorción dos azobencenos tetrametoxisustituídos. Imaxe adaptada da ref. [125]

Outro antecedente de grande interese foi publicado polo grupo de Hecht no 2012.<sup>139</sup> Neste traballo a funcionalización dos aneis aromáticos do azobenceno con átomos de flúor na posición *orto* ao dobre enlace N=N permitiulle desprazar o equilibrio dos isómeros irradiando con luz visible e acadando fotoisomerizacións case cuantitativas en ámbalas dúas direccións.

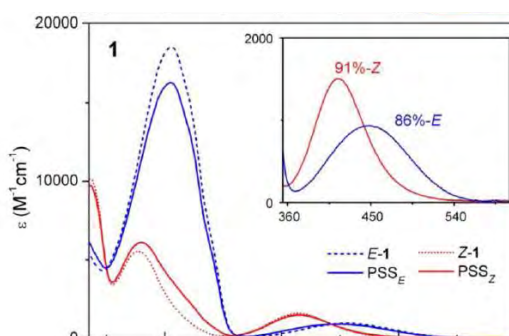
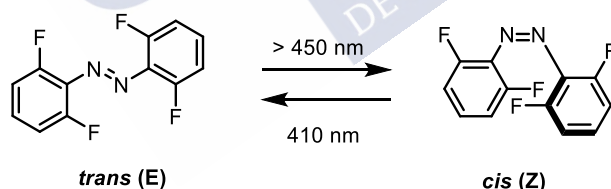


Figura 89. Arriba: reacción de isomerización dos azo orto-tetrafluorados. Abaixo: separación das bandas de absorción despois da isomerización destes derivados. Imaxe adaptada da ref. [139]

En oposición aos azos de Woolley metoxisubstituídos, os azos con substituíntes flúor conservan a planaridade do isómero *trans* porque os átomos de flúor teñen un tamaño moi pequeno. Esta calidade é moi importante para certos sistemas que precisan un amontoamento con interaccións  $\pi,\pi$  etc. Hai que resaltar qa excelente estabilidade termodinámica da forma *Z*, cun tempo de vida media duns 700 días a temperatura ambiente.

Ao introducir os flúor na posición *orto*, estes reducen a densidade electrónica arredor do enlace N=N por efecto indutivo.

Cálculos teóricos corroboraron que, en comparación co azobenceno simple, tódolos orbitais moleculares do azobenceno tetrafluor substituído son máis baixos en enerxía. O isómero *cis* ten o orbital n particularmente moi estabilizado. É dicir, o orbital n do isómero *Z* é baixo en enerxía, polo que a enerxía da transición  $n-\pi^*$  aumenta en maior proporción que para o *E*, resultando nun desprazamento a menor lonxitude de onda desta banda para o isómero *Z* (arredor de 414 nm en acetonitrilo, comparado con 456 nm no azobenceno simple). En cómputo global, a transición  $n-\pi^*$  para o isómero *E* apenas cambia. Polo tanto o resultado plásmase nunha separación de 42 nm das bandas  $n-\pi^*$  dos dous isómeros nos seus respectivos espectros de absorción. Deste xeito son capaces de desprazar o equilibrio para os dous lados con luz visible.

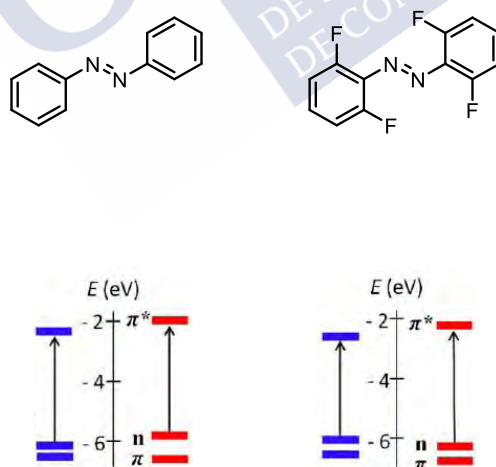


Figura 90. Representación do diagrama de enerxías para o azobenceno simple(esquerda) e o derivado orto-tetrafluorado(dereita). En azul os orbitais das formas *trans* e en vermello os das formas *cis*. Imaxe adaptada da referencia [139].

O mesmo grupo estudou o efecto que producía nos espectros de absorción funcionalizar o azo *orto*-tetrafluorado na posición *para*.<sup>126</sup> Estudaron tanto substituíntes electroattractores de electróns como electrodadores e descubriron que os grupos aceptores actúan en conxunto cos flúor, separando aínda máis as bandas  $n-\pi^*$ . Os efectos máis acusados observáronse para os derivados substituídos con grupos éster e amida. Nos dous casos obtiveron separacións entre as bandas  $n-\pi^*$  de 50 nm.



---

<sup>126</sup> C. Knie, M. Utecht, F. Zhao, H. Kulla, S. Kovalenko, A. M. Brouwer, P. Saalfrank, S. Hecht, D. Bléger, *Chem. Eur. J.* **2014**, *20*, 16492.

#### 4.1.3. Algunhas aplicacións de azobencenos

Grazas á propiedade da fotomodulación, os azobencenos téñense aplicado á síntese de máquinas moleculares<sup>127</sup>, ciencia de materiais e polímeros<sup>128</sup>, interruptores moleculares,<sup>129</sup> farmacoloxía,<sup>130</sup> etc. Especialmente interesante é o seu uso nas aplicacións biolóxicas.<sup>131</sup> O que abrangue o noso coñecemento, ata o de agora tódolos azoderivados que se usan para modificar as propiedades dalgunha biomacromolécula están covalentemente unidas a estas.<sup>132</sup> É dicir, precísase alterar quimicamente a estrutura primaria, e consecuentemente secundaria, da molécula a tratar polo que xa non se estaría ante un composto biolóxico 'puro' senón que ante un derivado sintético muado que poida que non exerza as mesmas funcións no organismo. O uso adoita estar máis estendido no ámbito das proteínas aínda que tamén se ten utilizado con oligonucleótidos. A maiores, a isomerización destes biocompostos modificados implica o uso de luz ultravioleta a excepción duns poucos exemplos derivados dos precedentes comentados,<sup>133</sup> un problema engadido se o fin último é traballar en organismos vivos.

Foi só durante o transcurso desta tese que o grupo de Matczyszyn publicou un caso no que sintetizaban azoderivados con extremos poliamina que eran capaces de modificar as propiedades do ADN a través de interaccións non covalentes.<sup>134</sup>

<sup>127</sup> a) X. M. Sun, W. Wang, L. B. Qiu, W. H. Guo, Y. L. Yu, H. S. Peng, *Angew. Chem. Int. Ed.* **2012**, *51*, 8520; b) Y.-J. Choi, J.-T. Kim, W.-J. Yoon, D.-G. Kang, M. Park, D.-Y. Kim, M.-H. Lee, S.-k. Ahn, K.-U. Jeong, *ACS Macro Lett.* **2018**, *7*, 576.

<sup>128</sup> a) M.-M. Russew, S. Hecht, *Adv. Mater.* **2010**, *22*, 3348; b) E. Blasco, B. V. K. J. Schmidt, C. Barner-Kowollik, M. Piñol, L., Oriol, *Macromol.* **2014**, *47*, 3693.

<sup>129</sup> E. Merino, M. Ribagorda, *Beilstein J. Org. Chem.* **2012**, *8*, 1071.

<sup>130</sup> M. J. Hansen, J. I. C. Hille, W. Szymanski, A. J. M. Driessen, B. L. Feringa, *Chem.* **2019**, *5*, 1293; B) W. A. Velema, J. P. van der Berg, M. J. Hansen, W. Szymanski, A. J. M. Driessen, B. L. Feringa, *Nat. Chem.* **2013**, *5*, 924.

<sup>131</sup> a) A. M. Caamano, M. E. Vazquez, J. Martinez-Costas, L. Castedo, J. L. Mascareñas, *Angew. Chem., Int. Ed.* **2000**, *39*, 3104; b) L. Guerrero, O. S. Smart, C. J. Weston, D. C. Burns, G. A. Woolley, R. K. Allemann, *Angew. Chem. Int. Ed.* **2005**, *44*, 7778; c) L. Albert, O. Vázquez, *Chem. Commun.* **2019**, doi:10.1039/C9CC03346G; d) M. Dong, A. Babalhavaeji, M. J. Hansen, L. Kálmán, G. A. Woolley, *Chem. Commun.* **2015**, *51*, 12981.

<sup>132</sup> A. S. Lubbe, W. Szymanski, B. L. Feringa, *Chem. Soc. Rev.* **2017**, *46*, 1052.

<sup>133</sup> a) A. Rullo, A. Reiner, A. Reiter, D. Trauner, E. Y. Isacoff, G. A. Woolley, *Chem. Commun.* **2014**, *50*, 14613; b) S. Samanta, T. M. McCormick, S. K. Schmidt, D. S. Seferos, G. A. Woolley, *Chem. Commun. (Camb.)* **2013**, *49*, 10314; c) C. Hoppmann, I. Maslennikov, S. Choe, L. Wang, *J. Am. Chem. Soc.* **2015**, *137*, 11218; d) J. B. Trads, J. Burgstaller, L. Laprell, D. B. Konrad, L. de la Osa de la Rosa, C. D. Weaver, H. Baier, D. Trauner, D. M. Barber, *Org. Biomol. Chem.* **2017**, *15*, 76; e) M. Dong, A. Babalhavaeji, M. J. Hansen, L. Kálmán, G. A. Woolley, *Chem. Commun.* **2015**, *51*, 12981.

<sup>134</sup> a) M. Deiana, Z. Pokladek, K. Matczyszyn, P. Mlynarz, M. Buckle, M. Samoc, *J. Mater. Chem. B* **2017**, *5*, 1028; b) M. Deiana, Z. Pokladek, J. Olesiak-Banska, P. Mlynarz, M. Samoc, K. Matczyszyn, *Sci. Rep.* **2016**, *6*, 28605; c) M. Dudek, M. Deiana, Z. Pokladek, K. Pawlik, K. Matczyszyn, *Chem Eur. J.* **2018**, *24*, 18963.

## 4.2. OBXECTIVOS<sup>135</sup>

A natureza posúe un grao de precisión no recoñecemento selectivo de secuencias de ADN difícil de mimetizar. Amais, moitos biosistemas modifican as súas propiedades de interacción segundo os sinais externos que perciban do seu entorno.<sup>136</sup> Sería importante xerar sistemas sintéticos que, ó igual que os naturais, poidan recoñecer secuencias específicas de forma controlada no tempo e no espazo, só cando reciben un estímulo externo. De feito, un gran reto no eido da química biolóxica é poder modular a actividade biolóxica de compostos sintéticos mediante un estímulo externo inocuo en medios celulares.

Sen dubida un dos estímulos externos mais atractivos é a luz, posto que se trata dunha fonte limpa, moi selectiva e compatible con ambientes biolóxicos. A maiores, a luz permite a manipulación precisa da intensidade e da lonxitude de onde de irradiación moi doadamente.<sup>137</sup>

A partir dos resultados obtidos no noso grupo de investigación sobre axentes fluoroxénicos que se unen ó ADN de tipo bisbenzamidina,<sup>138</sup> e baseándose nos precedentes sobre a química de azobencenos, un dos obxectivos que nos propuxemos é a “hibridación” entre un núcleo azobencénico e moléculas do tipo bisbenzamidinio co fin de obter azoderivados que deben unirse a secuencias específicas do ADN grazas á súa parte bisbenzamidínica. Espérase que o cambio de xeometría entre *cis* e *trans* teña unha influencia importante na interacción co ADN diana, ao facilitar ou impedir que os grupos amidinio se insiran no suco menor. O funcionamento do sistema bisbenzamidínico represéntase no seguinte esquema:

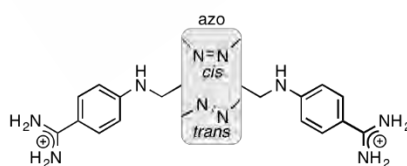


Figura 91. Representación do sistema azobencénico.

<sup>135</sup> Este proxecto realizouse en colaboración co doutor Mateo I. Sánchez.

<sup>136</sup> J. M. Requena, Stress Response in Microbiology **2012**, Caister Acad. Press.

<sup>137</sup> a) C. Brieke, F. Rohrbach, A. Gottschalk, G. Mayer, A. Heckel, *Angew. Chem., Int. Ed.* **2012**, *51*, 8446; b) G. M. and A. Heckel, *Angew. Chem., Int. Ed.*, **2006**, *45*, 4900.

<sup>138</sup> a) O. Vázquez, M. I. Sánchez, J. Martínez-Costas, M. E. Vázquez, J. L. Mascareñas, *Org. Lett.* **2010**, *12*, 216; b) O. Vázquez, M. I. Sánchez, M. E. Vázquez, J. L. Mascareñas, *Chem Commun.* **2010**, *46*, 5518. c) M. I. Sánchez, O. Vázquez, J. Martínez-Costas, M. E. Vázquez, J. L. Mascareñas, *Chem. Sci.* **2012**, *3*, 2383.

Polo tanto, propónse a creación dun sistema facilmente conmutable que se una especificamente a secuencias diana de ADN e que permita intercambiar entre dous estados *activo/inactivo*, de maneira reversible e a vontade, mediante a aplicación dun estímulo externo non nocivo.





### 4.3. CHAVES FOTOCONMUTABLES PARA O RECOÑECEMENTO SELECTIVO DE ADN: O AZOBENCENO.

#### 4.3.1. Síntese: primeira aproximación

Baseándose nos precedentes anteriormente descritos, decidiuse que o derivado amidínico que se mostra na seguinte figura serían os primeiros candidatos para comprobar se é posible acadar unha unión fotomodulable a secuencias específicas de ADN.

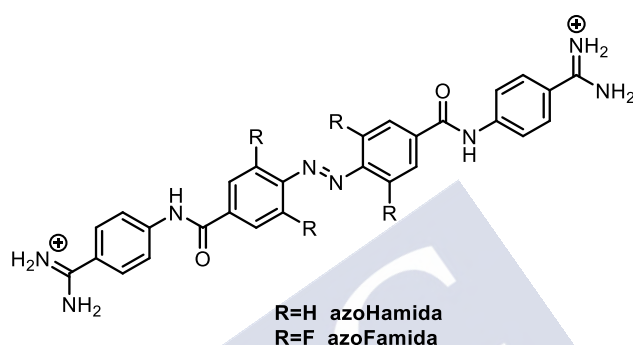


Figura 92. Estrutura das dianas sintéticas da primeira aproximación

Os grupos carbonilo en posición *para* ao grupo azo retiran carga do dobre enlace N=N, o que permite a separación entre as bandas do *cis* e do *trans* para a súa fácil isomerización. O grupo funcional amida posúe carácter parcial de dobre enlace é dicir, ten unha disposición atómica estabilizada por resonancia. Isto provoca unha restrición na liberdade de xiro do enlace CO-N, que fai que tódolos átomos implicados no enlace se atopen no mesmo plano. Esta característica aumenta a planaridade da forma *trans* desta diana sintética que, se cabe, favorecería a súa inserción no suco menor.

Decidiuse realizar a síntese tanto do derivado tetrafluorado (**azoFamida**) esperando que se poida isomerizar con luz visible en ámbalas direccións coma do non fluorado (**azoHamida**) por se os átomos de flúor do primeiro comprometen a interacción e/ou o recoñecemento do ADN.

Tal como demostrou o grupo de Stefan Hecht,<sup>139</sup> a alta electronegatividade dos flúor retira carga por efecto inductivo, polo que a densidade electrónica arredor do enlace N=N diminúe. Como consecuencia, a forma *cis* destes azoderivados tetrafluorados presenta unha estabilidade

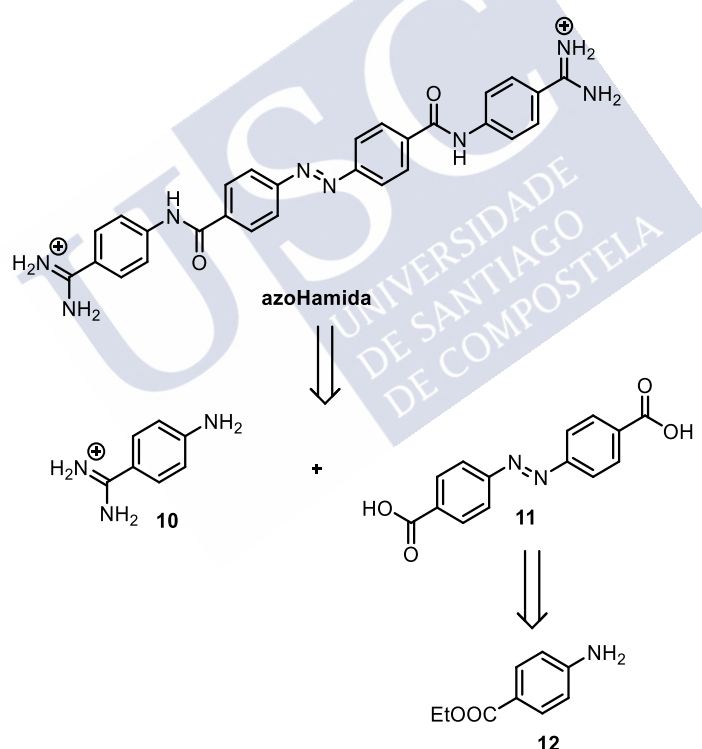
<sup>139</sup> D. Bléger, J. Schwarz, A. Brouwer, S. Hecht, *J. Am. Chem. Soc.* **2012**, *134*, 20597.



térmica excepcional, cun tempo de vida media duns 700 días. Esta *orto* funcionalización resulta tamén na separación das bandas  $n-\pi^*$  dos dous isómeros, que fai posible a fotoisomerización con luz visible bidireccionalmente e case cuantitativa.

A maiores, os átomos de flúor teñen un tamaño moi pequeno que non afecta á planaridade da molécula como si o fan outros grupos como os metoxilo. Estes últimos, deseñados no grupo de Woolley<sup>140</sup> tamén estabilizan a forma *cis* pero son moito máis voluminosos e distorsionan a forma plana do isómero *trans*. Amais, son vulnerables de sufrir redución en medio biolóxico. Estas características poderían dificultar a inserción no suco menor do ADN, polo que se decidiu usar os derivados fluorados.

Para a síntese do **azoHamida** penseuse nun acoplamento dobre entre o diácido **11** e a amina comercial *para*-aminobenzamidina como se mostra no Esquema 21 de retrosíntese.



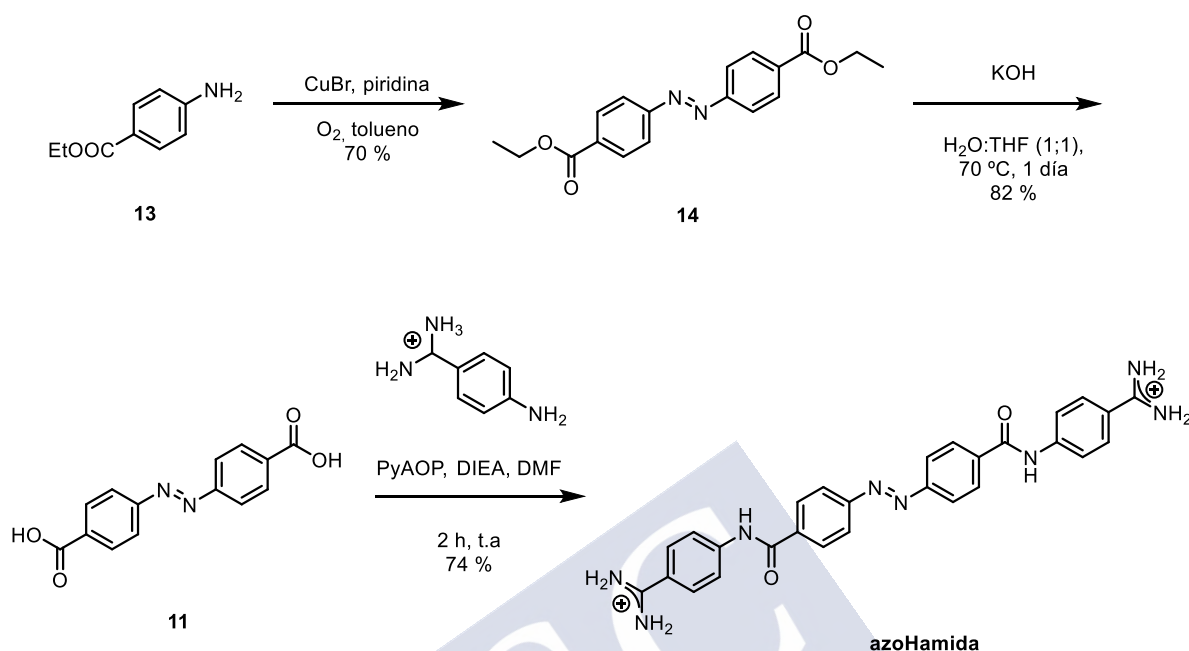
Esquema 21. Aproximación retrosintética para obter a diana sintética azoHamida.

En total, a ruta sintética consta só de tres pasos (ver **Esquema 22**). No primeiro, fórmase o enlace azo seguindo un procedemento descrito.<sup>141</sup> Mediante o acoplamento oxidativo de dúas

<sup>140</sup> A. Beharry, O. Sadovskii, G. A. Woolley, *J. Am. Chem. Soc.* **2011**, *133*, 19684.

<sup>141</sup> C. Zhang, N. Jiao, *Angew. Chem. Int. Ed.*, **2010**, *49*, 6174.

benzocaínas en condicións suaves usando reactivos comerciais baratos, obtívose **14** cun alto rendemento. De seguido, hidrolizáronse os ésteres quentando en medio básico.

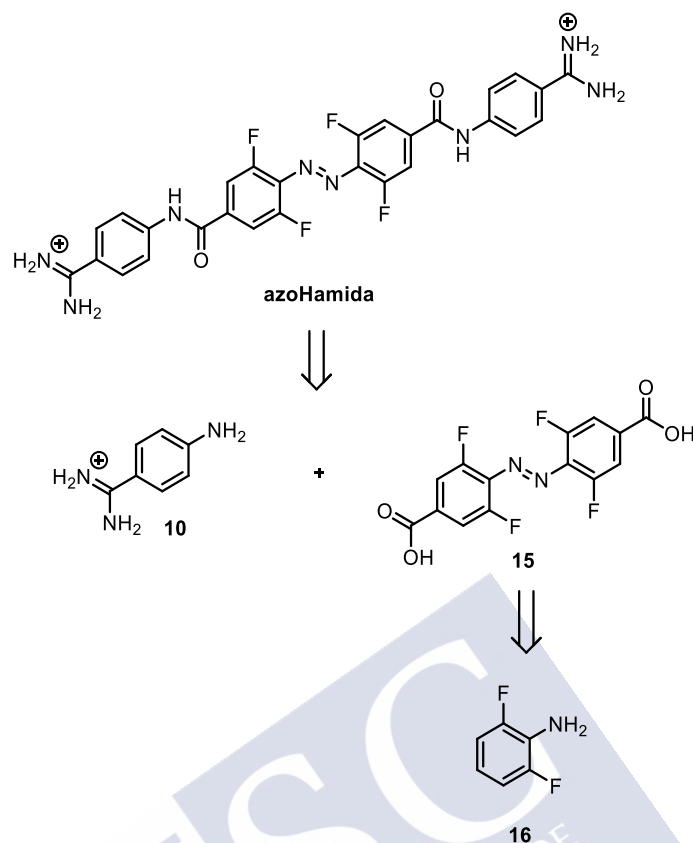


Esquema 22. Ruta sintética para acadar azoHamida.

Por último, realizouse o acoplamento entre a aminobenzamidina e o diácido **15**. Proboouse a condensación con HATU e DIEA en DCM e DMF a temperatura ambiente e, a pesares de observarse o produto en ámbalas dúas tentativas, este atopábase en forma minoritaria e o cru de reacción analizado por HPLC-MS tiña un aspecto moi sucio. Trocouse pois, o axente acoplante por PyAOP. Desta vez, a reacción transcorreu moito máis limpa en tan só unhas 3 h podendo illarse o produto **azoHamida** despois de purificar en fase revertida, cun 74 % de rendemento.

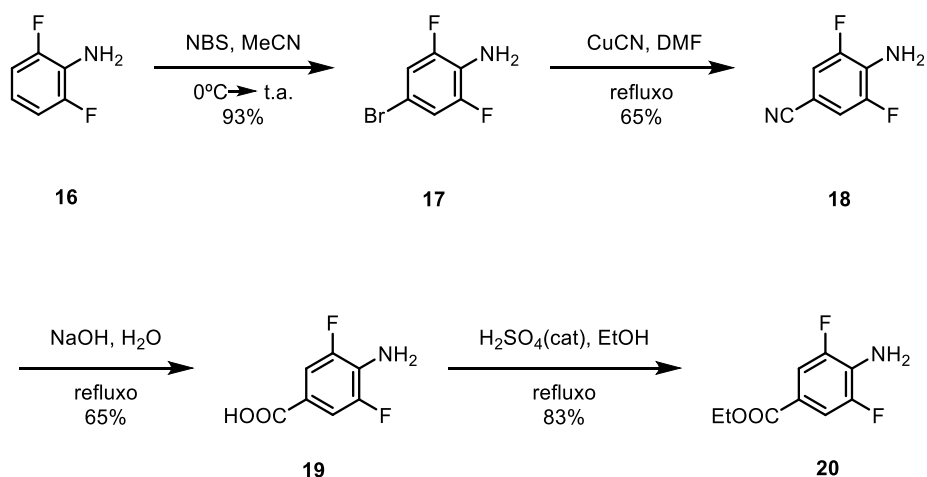
Ao irradiar unha disolución de **azoHamida** en condicións típicas de isomerización ( $\lambda = 365 \text{ nm}$ ), non se apreciou ningún cambio nos espectros de UV nin RMN. Tamén se probou a quentar a mostra e, de novo, sen indicios de isomerización.

Para a síntese do derivado fluorado, propúxose unha retrosíntese moi similar:



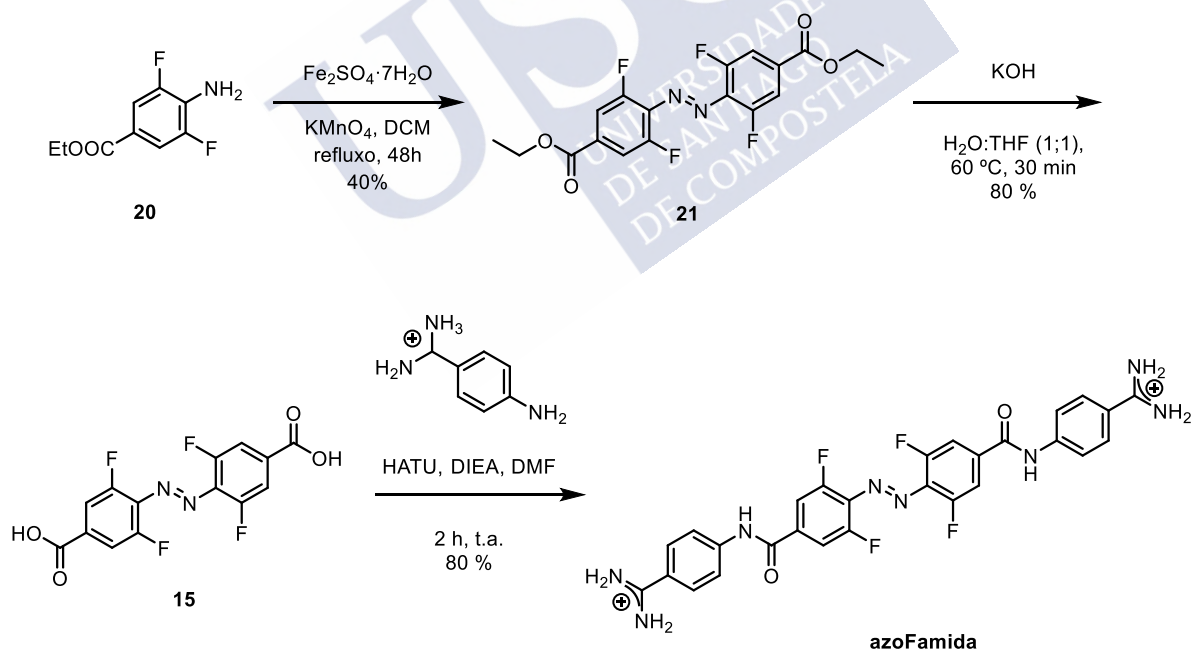
Esquema 23. Aproximación retrosintética para obter a diana sintética azoFamida.

Para formar o núcleo azoico **15** seguíuse a estratexia sintética descrita por Hetch ata o diéster **21**<sup>139</sup>. Na cal se broma rexiosselectivamente en *para* con *N*-bromosuccinimida (NBS) a 2,6-difluoroanilina **16** comercial para logo substituír o bromuro por un nitrilo empregando CuCN. Cunha hidrólise en medio básico transfórmase o grupo nitrilo en ácido e, a continuación, coa esterificación de **19**, acadouse a difluoroanilina **20**, tal como se mostra no Esquema 24.



Esquema 24. Primeira parte da ruta sintética seguida para acadar o núcleo azoico tetrafluorado.

O éster **20** dimerizouse por tratamento con sulfato de ferro heptahidratado mais permanganato de potasio obténdose o núcleo do azocomposto **21** cun rendimento moderado despois de deixalo reaccionando durante dous días.



Esquema 25. Segunda parte da ruta sintética para obter o azoFamida.

Esta etapa, con tan baixo rendemento, limita a síntese completa, de sete pasos en total. Xa que se tiñan os intermedios sintetizados, tratouse de optimizar devandito acoplamento. Probáronse paralelamente varias vías para obter o núcleo azoico.

Na primeira vía, nomeada vía C no **Esquema 26**, reproducíronse as mesmas condicións que no caso do **azoHamida**. Tratouse de dimerizar o ester fluorado (**9**) con CuBr en condicións oxidantes até obter o núcleo de azo. con tan pouca fortuna, que só se obtivo unha mestura complexa de compostos.

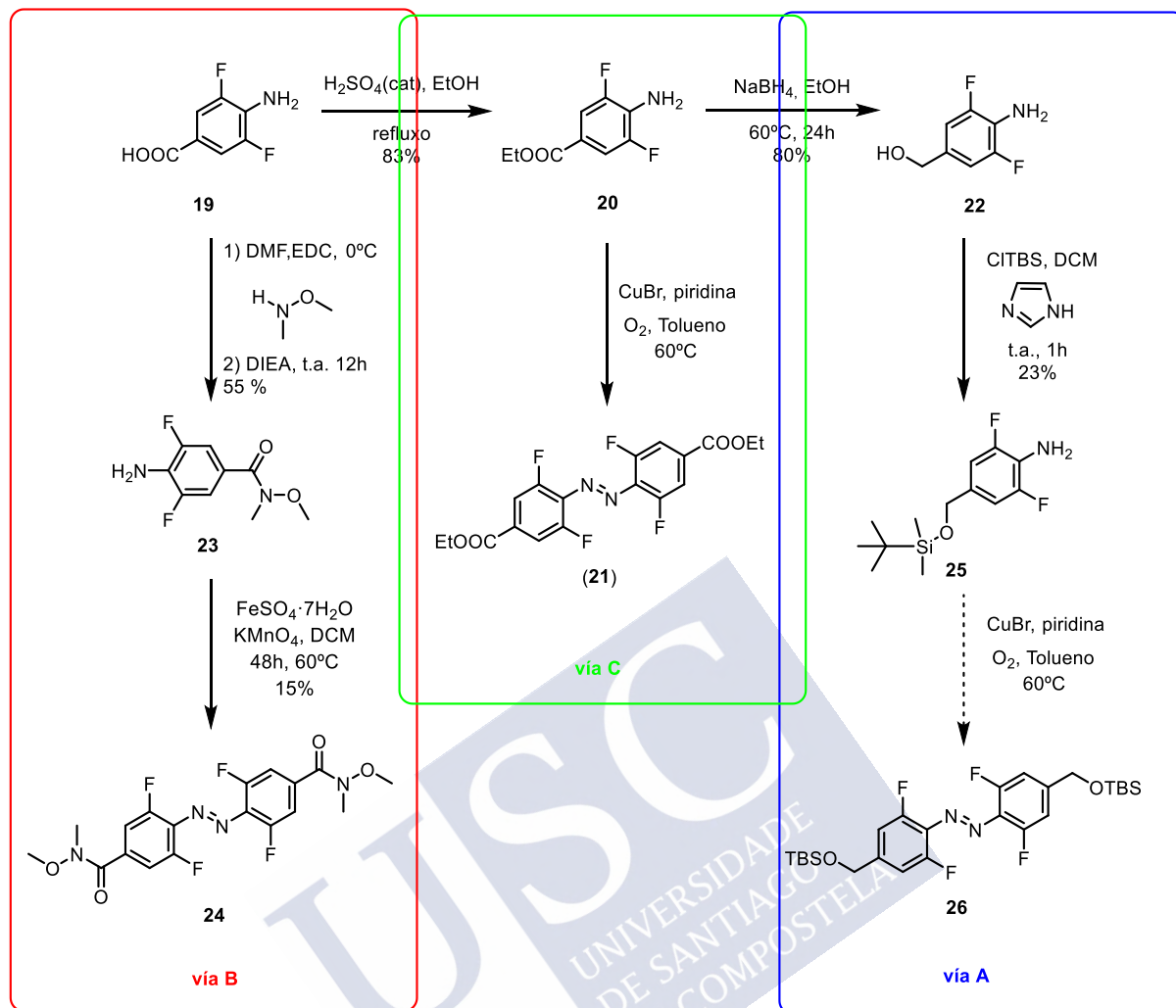
Por non desestimar este tipo de acoplamento oxidante que funcionara tan ben para o **azoHamida**, seguiuse a vía A. Así pois, en primeiro lugar, reduciuse o 4-amino-3,5-difluorobenzoato de etilo ao alcol **22** e protexeuse este con TBSCl para evitar calquera tipo de reacción de condensación secundaria en vindeiras etapas. Malia tódolos esforzos, o rendemento da protección de alcol foi pésimo, observándose que o produto se degradaba na etapa de purificación. Polo que se rexeitou esta vía sintética.

A terceira e última estratexia probada, vía B, para obter o enlace azo baseouse na creación dunha amida de Weinreb.<sup>142</sup> Esta comeza coa formación da amida de Weinreb **23** a través do acoplamento entre a *N,O*-dimetilhidroxilamina co ácido da 3,5-difluoroanilina **19** empregando o axente acoplante etildimetilcarbodiimida (EDC) e a diisopropiletilamina (DIEA) como base. Nun segundo paso, formouse o composto azoico **24** con sulfato de ferro heptahidratado en condicións oxidantes. só se logrou un 15% do dímero **24**. Pode que a conxugación electrónica da amida de Weinreb ó ser especialmente rica en carga, dificulte o acoplamento para dar o azo dímero **24** explicando un rendemento tan miúdo.

Consecuentemente, rexeitouse tamén esta terceira vía e decidiuse retomar a ruta do **Esquema 25** no afán de acadar o **azoFamida**. Deste xeito, hidrolizouse en medio básico o diester tetrafluorado **21** e, posteriormente realizouse o acoplamento entre o diácido obtido mais a 4-aminobenzamidina, con HATU como axente activante de ácidos. Este procedemento resultou na obtención da segunda diana sintética nun 80 % de rendemento trala súa purificación por fase revertida. O composto analizouse por HPLC-EM, observándose un só pico no cromatograma, o que semellaba a presenza dun único isómero.

---

<sup>142</sup> P.-Y Yang, M. Wang, C. Y. He, S. Q. Yao, *Chem. Commun.* **2012**, 48, 835.



Esquema 26. Rutas sintéticas alternativas probadas para optimizar o acoplamento do núcleo do azoFamida.

Despois de probar as condicións de isomerización coma ca diana sintética obtida anteriormente, non se detectaron cambios nos seus espectros. Polo tanto, esta segunda diana sintética tampouco é quen de levar a cabo a isomerización.

Con estes resultados experimentais propúxose que as moléculas deseñadas ata o de agora, cas amidas en posición *para* ao grupo azo, e unindo outro anel de benceno ao núcleo azoico, posúen elevada resonancia electrónica dende os extremos ata o enlace N=N intensificada polo enlace peptídico que serve como conector entre os aneis. Esta característica electrónica impide que o dobre enlace N=N se rompa para pasar da xeometría *cis* á *trans*.

Para solucionar este problema, propúxose outras dianas sintéticas, cuxa síntese se basea nunha reacción descuberta no noso grupo de investigación.

#### 4.3.1.1. Síntese: terceira e cuarta dianas sintéticas

A reacción trátase dunha simple aminación redutora entre un dialdehido aromático e a *para*-aminobenzamidina que permite obter de xeito rápido e versátil, numerosas aza-bisbenzamidinas.<sup>146c</sup>

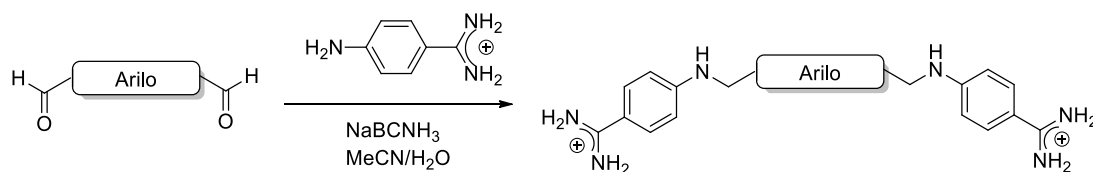


Figura 93. Reacción de aminación redutora para obter bisbenzamidinas fluoroxénicas desenvolta no noso grupo de investigación.

As novas dianas sintéticas propostas móstranse na **Figura 94**. Nestas moléculas, elimínase o enlace peptídico e introdúcese un grupo metilénico. Agárdase que este grupo rompa a conxugación do azo cos substituíntes e se poidan levar a cabo as isomerizacións De novo, realizouse a síntese tamén, da versión *orto*-tetrafluorada.

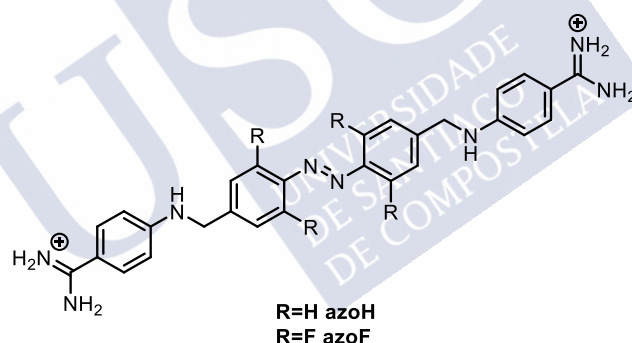
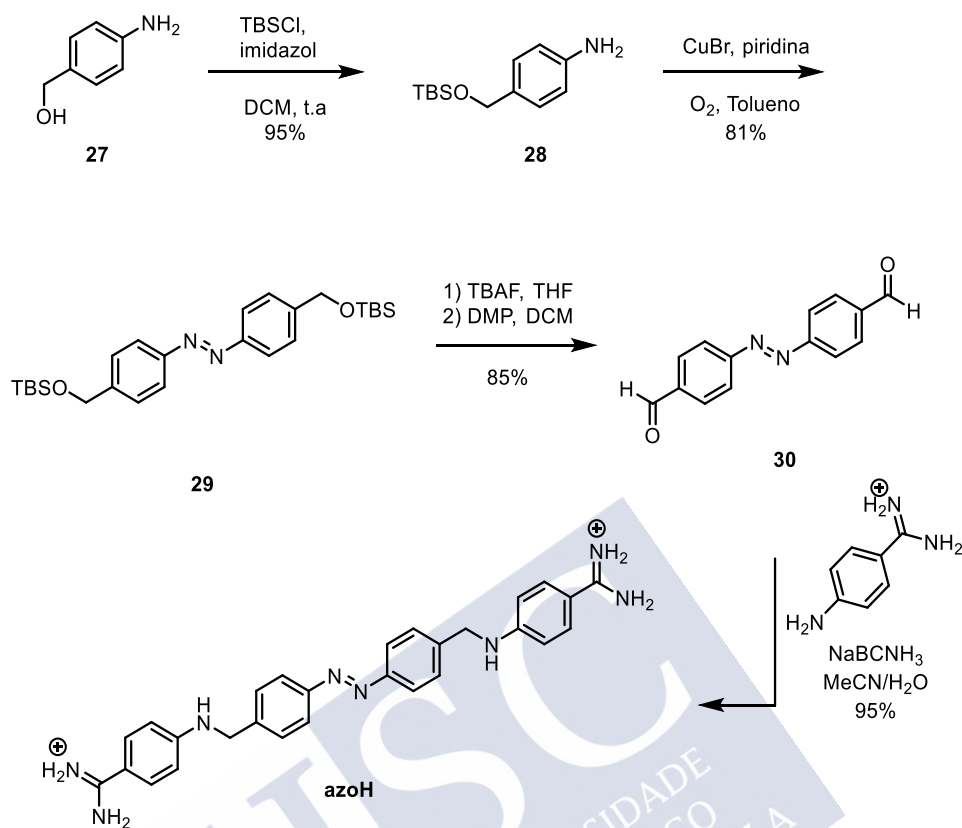


Figura 94. Estrutura dianas sintéticas 3 (azoH) e 4 (azoF).

Para a síntese do **azoH** partiuse do alcohol 4-aminobencílico **27**, protexeuse con TBSCl e, posteriormente, tivo lugar a formación do grupo azo a través dun acoplamento oxidante catalizado por CuBr (ver **Esquema 27**).

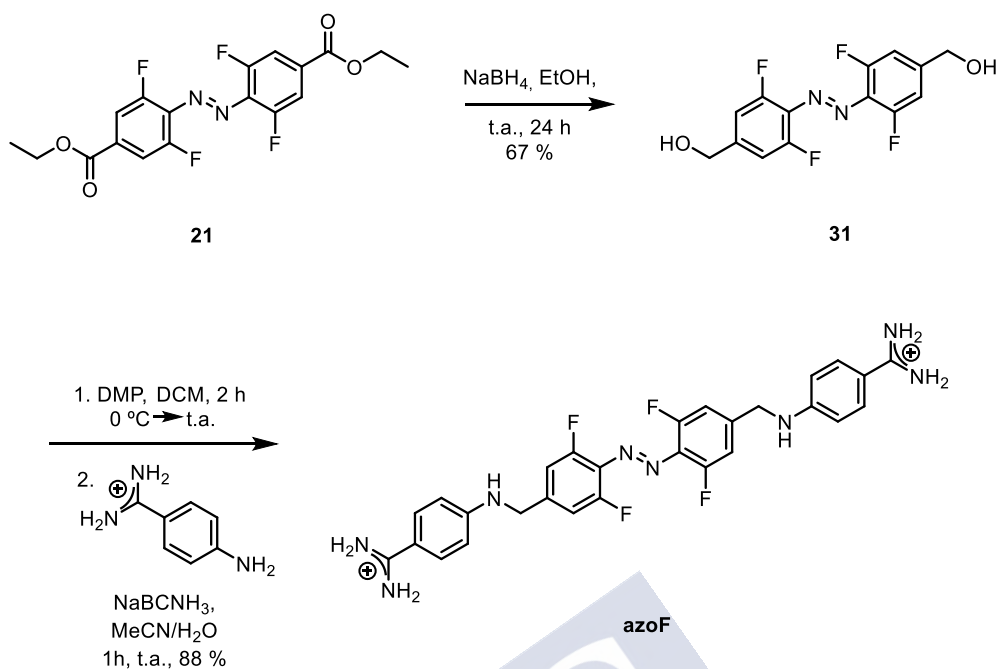




Esquema 27. Ruta sintética para obtención da terceira diana sintética azoH.

Unha vez illado o composto **29**, desprotexéronse os alcois para logo oxidalos con DMP, até obter o dialdehído **30**. Cabe destacar que a presenza do grupo protector é necesaria porque, nestas condicións de oxidación, un posible produto sería a oxidación do alcol a un aldehído, podendo este condensar coa anilina.

Co dialdehído desexado e a través da aminación redutora antes descrita, obtívose a diana sintética **azoH**. En paralelo, fíxose o azoF reutilizando o diéster **21** que xa se tiña sintetizado.



Esquema 28. Ruta de obtención da cuarta diana sintética, azoF.

Despois de optimizar as condicións, a redución do diéster **21** levouse a cabo con NaBH<sub>4</sub> dando o dialcol **31**. O cal se oxidou ao dialdehido con periodinano de Dess Martin (DMP), pero a purificación deste composto non foi satisfactoria. Bardante de tratar o cru de reacción con tiosulfato sódico para eliminar o iodo hipervalente, en tódolos intentos de purificación se arrastraban trazas deste. Debido a isto, acordouse facer a reacción *in situ*, é dicir, engadindo a 4-aminobenzamidina e o redutor NaBCNH<sub>3</sub> sobre o cru do aldehido. Este procedemento resultou na obtención do **azoF** cun 88 % de rendemento global trala súa purificación por fase revertida.

#### 4.3.2. Comportamento fotoquímico e Isomerizacións

Os espectros de absorción dos derivados azobencénicos obtidos constan de dúas bandas: unha moi intensa na rexión ultravioleta arredor dos 313 nm, que se corresponde coa transición  $\pi-\pi^*$ . Outra, asociada á transición  $n-\pi^*$  e bastante máis feble, situada na rexión visible sobre os 450 nm.

Cando se rexistrou o espectro para o **azoHtrans** puro obtido directamente da síntese, apreciouse unha banda de absorbencia moi forte con  $\lambda_{\text{máx}} \approx 323$  nm. Así como outra menos intensa na rexión visible con  $\lambda_{\text{máx}} \approx 439$  nm ( **Figura 95**). Esta mesma mostra irradiouse a 365 nm con luz ultravioleta provocando a súa isomerización a *cis*. No novo espectro, comprobouse que a banda  $\pi - \pi^*$  de absorción sufriu un pequeno efecto hipsocrómico ata  $\lambda_{\text{máx}} \approx 316$  nm concomitante cunha diminución na súa intensidade. A banda da rexión visible desprazouse lixeiramente cara o azul  $\lambda_{\text{máx}} \approx 426$  nm, resultando nunha separación de 13 nm entres as bandas  $n - \pi^*$  dos dous isómeros.

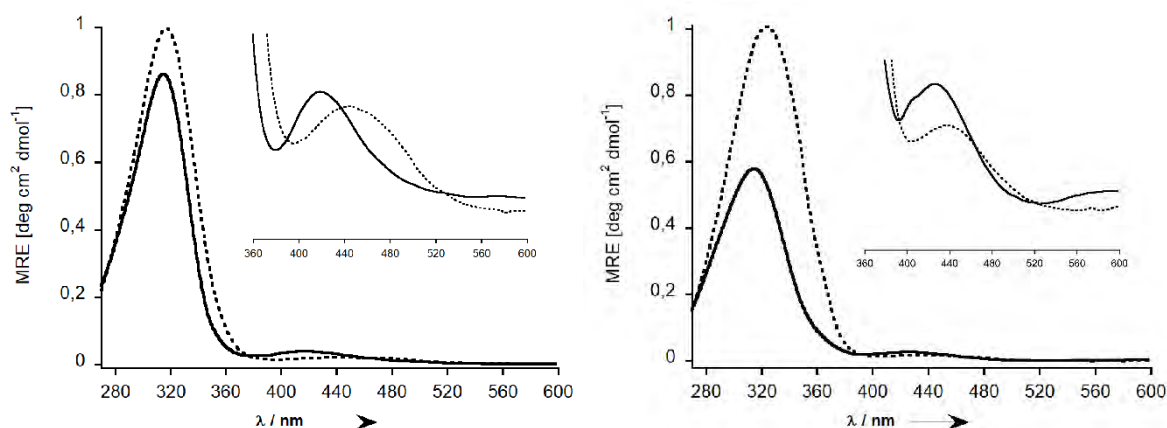


Figura 95. Esquerda: Espectros de absorción para o **azoFtrans** (liña descontinua) e o **azoFcis** (liña continua). Dereita: Espectros de absorción para o **azoHtrans** (liña descontinua) e o **azoHcis** (liña continua). Móstrase unha ampliación para as bandas da rexión do visible.

O *o*-tetrafluoroazoderivado na súa forma *trans* presenta unha banda de absorción intensa centrada aproximadamente en 317 nm correspondente á transición  $\pi - \pi^*$ ; o máximo de absorción para a banda  $n - \pi^*$  desta forma atópase sobre 450 nm. Namentres, o isómero *cis* de **azoF** posúe unha banda  $\pi - \pi^*$  a 314 nm tamén bastante intensa e outra banda no visible con máximo en torno a 417 nm. A variación entre os máximos das bandas dos dous isómeros na rexión do visible é duns 33 nm. Deste xeito, a introdución dos átomos de flúor en *orto* produciu un desprazamento cara o vermello da banda menos intensa  $n - \pi^*$  con respecto do seu homólogo sen flúor para o **azoFtrans**, así como un lixeiro cambio cara o azul de dita banda para o **azoFcis** como era de esperar. Estas variación permitirá isomerizar o **azoF** con luz visible en ámbolos dous sentidos.

Cabe destacar que a banda de absorción do grupo amidinio adoita localizarse en torno aos 312 - 319 nm, ao igual que a transición  $\pi-\pi^*$  do grupo azo. É por este motivo polo que a intensidade desta banda despois da isomerización cara *cis* segue sendo tan notable.

Os experimentos de UV-vis foron rexistrados a 20 °C en metanol, DMSO e auga pero non se observaron variacións significativas nas bandas ao cambiar de disolvente.

Para que os sistemas creados poidan usarse en aplicacións biolóxicas, é preciso que non sufran degradacións por mor da luz e do tempo usado para a súa isomerización, así como que resistan múltiples ciclos de radiación conservando a súa absorbencia e sen producir fotobranqueo aparente. Realizáronse medidas de absorbencia despois de alternar varios ciclos de isomerización *Z/E* con luz verde e azul respectivamente e non se observou degradación das propiedades fotoquímicas das moléculas. A maiores, deixouse irradiando durante 1 h con luz azul (transición do ciclo 8 ao 9) e outra hora con luz verde ou UV no caso do **azoH** (transición do ciclo 9 ao 10) para finalmente, confirmarse que ámbolos dous azo-interruptores creados son robustos e a súa fotoconversión é reversible. (ver **Figura 96**)

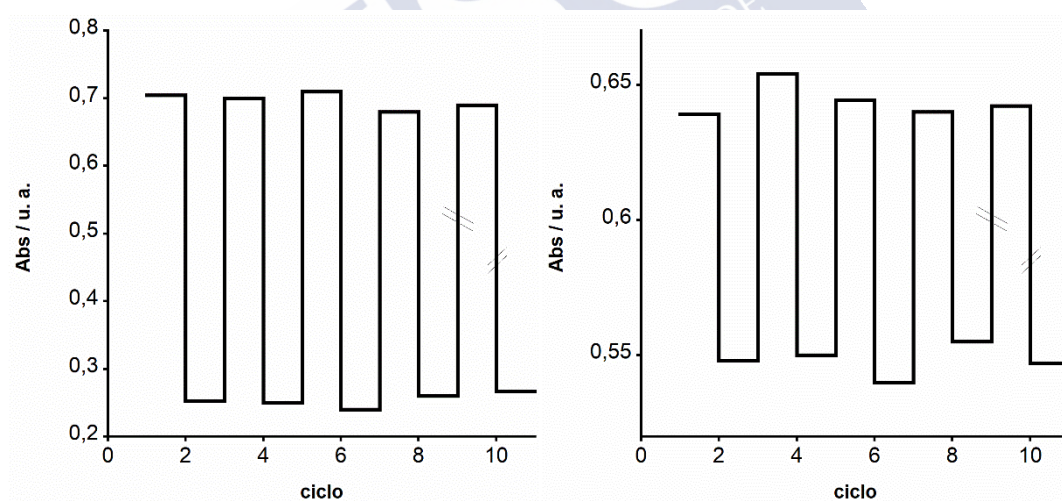


Figura 96. Múltiples roldas de interconversión alternando luz UV e luz azul ( $\lambda= 410$  nm) para azoH (esquerda) e luz verde( $\lambda= 520$  nm) e luz azul ( $\lambda= 410$  nm) para o azoF (dereita). Os cortes entre os ciclos 9 e 10 correspóndense cunha hora de irradiación continua. As absorbencias foron medidas nos máximos das transicións das bandas  $\pi-\pi^*$  dos isómeros E en disolución acuosa Tris·HCl pH 7.5 20 mM, 100 mM de NaCl.

O estudo da fotoisomerización seguiu rexistrando espectros de  $^1\text{H}$ -RMN, HPLC e UV antes e despois de cada irradiación. Desta forma puido calcularse a relación *E:Z* de cada isómero no seu estado fotoestacionario (EFE).

Para o **azoH** obtívose unha porcentaxe de > 88 % do isómero Z despois de irradiar a 365 nm durante 20 min. A isomerización Z→E con luz azul (410 nm) resultou nun desprazamento do 80 % cara o isómero E. Esta mesma isomerización quentando a 80 °C durante 40 min produciu tamén un 80 % de **azoHtrans**.

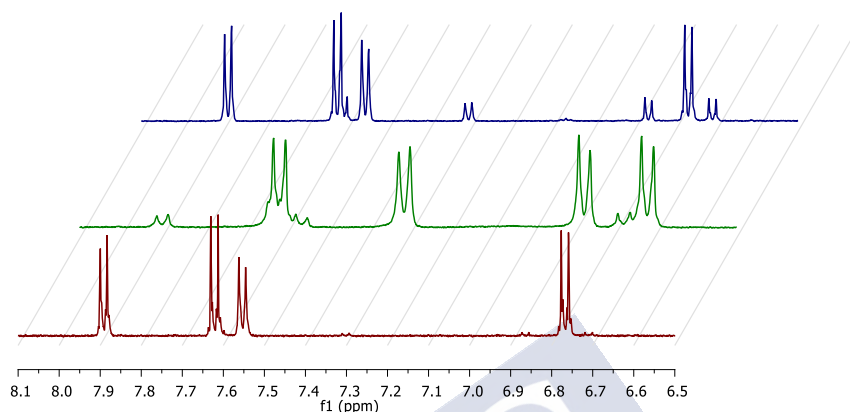


Figura 97. Estudo da isomerización de azoH mediante  $^1\text{H}$ -RMN. Represenaase a ampliación do espectro da mostra sen irradiar (vermello), irradiada durante 20 min a 365 nm (verde) e irradiada con luz azul durante 15 min (azul)

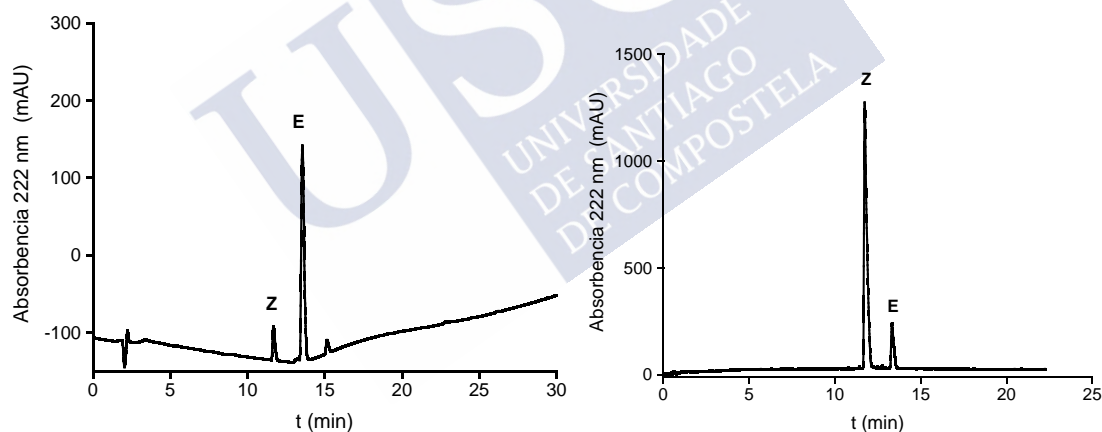


Figura 98. Esquerda: cromatograma tras irradiar o azoH a 410 nm durante 15 min. Dereita: cromatograma do equilibrio despois de irradiar 20 min a 365 nm.

En canto ao **azoF**, a isomerización con luz ultravioleta a 365 nm durante 20 min tan só dou un 60 % de **azoFcis** (ver **Figura 100**) Aumentando o tempo de irradiación a 365 nm ata 45 min non se mellorou o desprazamento do equilibrio. Mentres que facendo a mesma isomerización pero irradiando na banda  $n-\pi^*$  durante 15 min con luz verde (520 nm) obtívose máis dun 88 % do isómero Z.

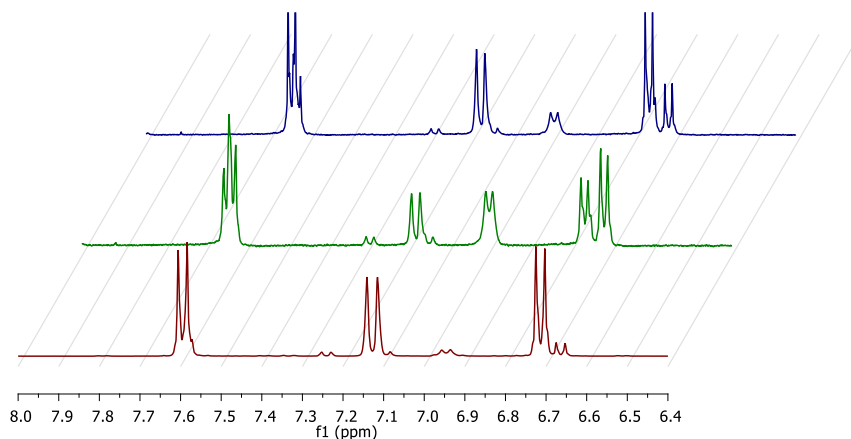


Figura 99. Ampliación do espectro de  $^1\text{H}$ -RMN do azo F. Represéntanse o espectro da mostra sen irradiar (vermello), irradiada durante 30 min a 365 nm (UV) e quente a 80 °C durante 40 min.

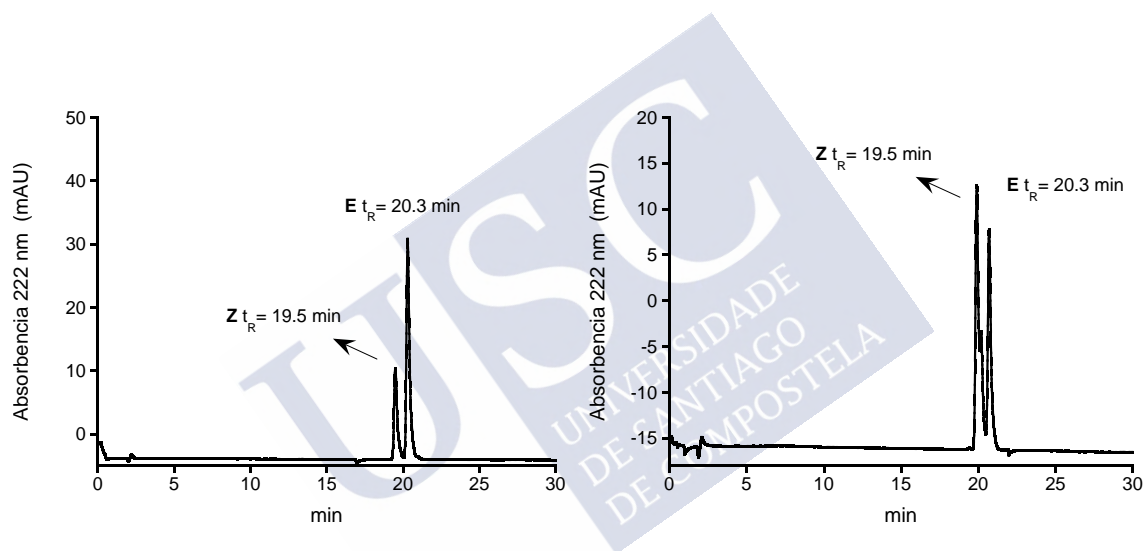


Figura 100. Esquerda: cromatograma tras quente o azoF a 80 °C durante 40 min. Dereita: cromatograma do equilibrio despois de irradiar 45 min a 365 nm.

A conversión en isómero *trans* cando se fai quente a mostra a 80 °C durante 40 min obtense un 80 % desta forma. Namentres, esta conversión se realiza irradiando con luz azul unha mostra durante 15 min, obtense un > 87 % do azoF*trans* (ver **Figura 101** e **Figura 100**).<sup>143</sup>

<sup>143</sup> Durante o transcurso deste proxecto, cambiouse o aparello de HPLC-EM e os isómeros do azoF non puideron separarse cas condicións usadas na nova maquinaria.

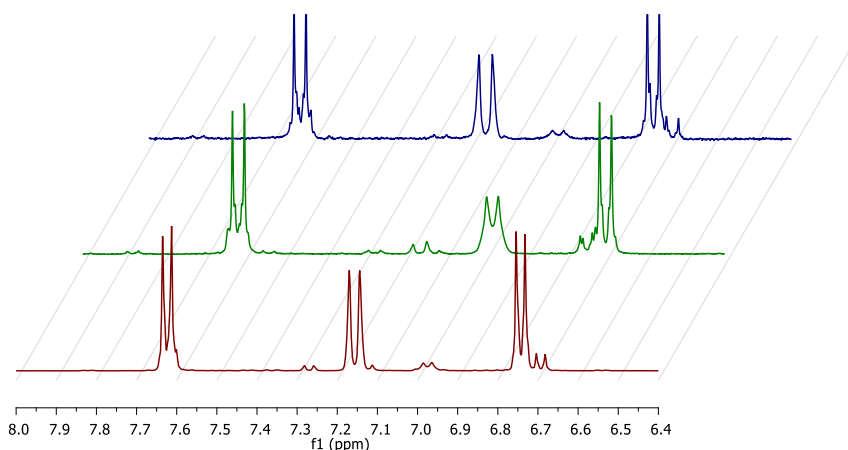


Figura 101. Ampliación do espectro de  $^1\text{H}$ -RMN do azoF. Represéntanse o espectro da mostra sen irradiar (vermello), irradiada durante 15 min con luz verde (520 nm) e irradiada con luz azul (410 nm) durante 15 min (azul).

#### 4.3.3. Tempo de vida media do isómeros

O tempo de vida media de cada isómero dende o seu EFE é outro punto clave para o bo funcionamento do interruptor que se desexa crear. Se un dos isómeros posúe un tempo de vida media demasiado curto implica que non teña tempo a levar a cabo a súa función sobre o ADN. Para estudar a lentitude da reversión dos equilibrios de isomerización realizáronselle  $^1\text{H}$ -RMN periodicamente. Irradiáronse cadansúa mostra de **azoH** e **azoF** a 410 nm e realizouse a análise dos isómeros *E* destas moléculas.

Para o **azoHcis**, dende un 88 % pasou a un 77 % nun día, o segundo día a porcentaxe do *cis* atopouse nun 66 % como se amosa na (**Figura 102**). Aos sete días a mestura tiña 61% de *cis* e aos 10 días o 57 % (espectro azul escuro **Figura 102**). A mestura 1:1 (*E:Z*) deuse aos 15 días despois da irradiación. E o equilibrio acadouse ao chegar ó 30 % de isómero *cis* en 30 días (espectro lila **Figura 102**). Onde se mantivo durante meses.



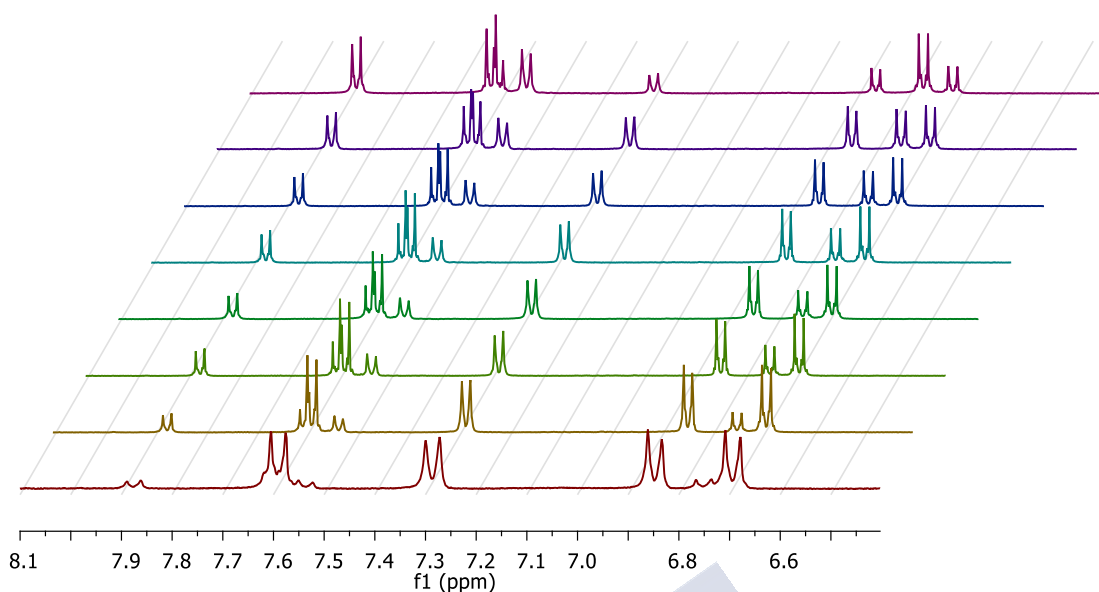


Figura 102. Ampliación do espectro de <sup>1</sup>H-RMN para o estudo do tempo de vida media de azoHcis.

Os espectros de RMN do isómero *cis* do **azoF** revelaron que este composto era termodinamicamente máis estable que o seu análogo **azoHcis**. Deste xeito, a composición da mestura contiña un 85 % de *Z* tras un día da irradiación con luz verde. Aos sete días presentaba un 80 % de *cis* e pasados 10 días a composición só se depraizara cara un 77% de *Z*. Ao cabo de 15 días disolución aínda contiña un 75 % de **azoFcis**. (Ver Figura 103)

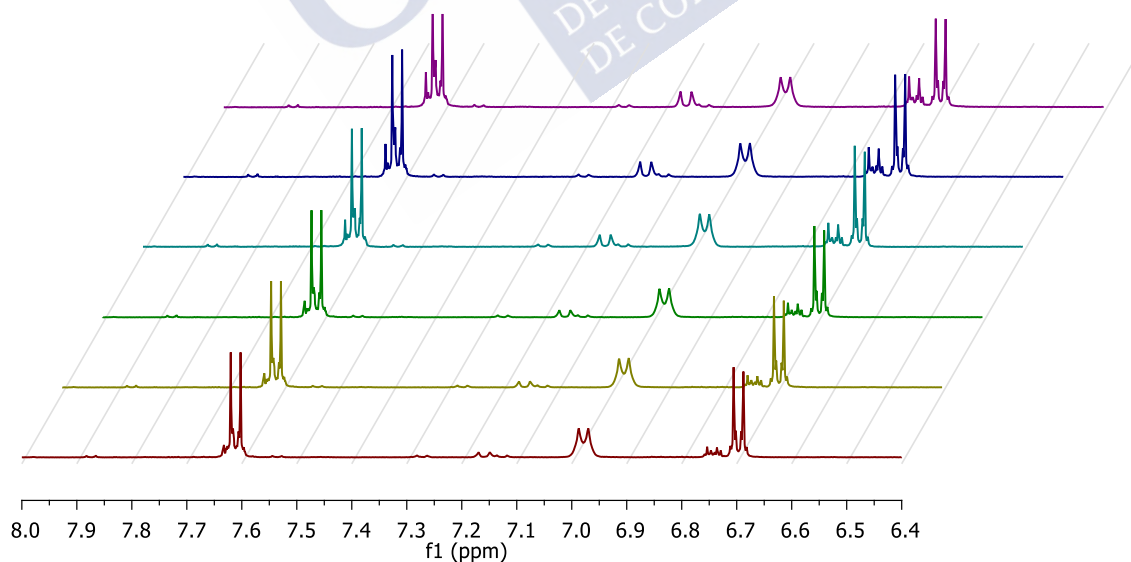


Figura 103. Ampliación do espectro de <sup>1</sup>H-RMN para o estudo do tempo de vida media de azoFcis. Espectros a 1 día(vermello), 2 días, 3 días(verde), 7 días, 10 días(azul escuro) e 15 días (lila) despois da irradiación con luz verde.



A relación do isómero do **azoFtrans** apenas cambiaba (**Figura 104**) Pasando dun 80:20 (*E:Z*) a un 78:22 (*E:Z*) despois de 120 h. Despois de 10 días diminuía a cantidade da forma *trans* ata 75:25 (*E:Z*). Non foi ata preto dos 90 días cando a mestura acadou o equilibrio de 60:40 (*E:Z*) manténdose nesta proporción durante meses.

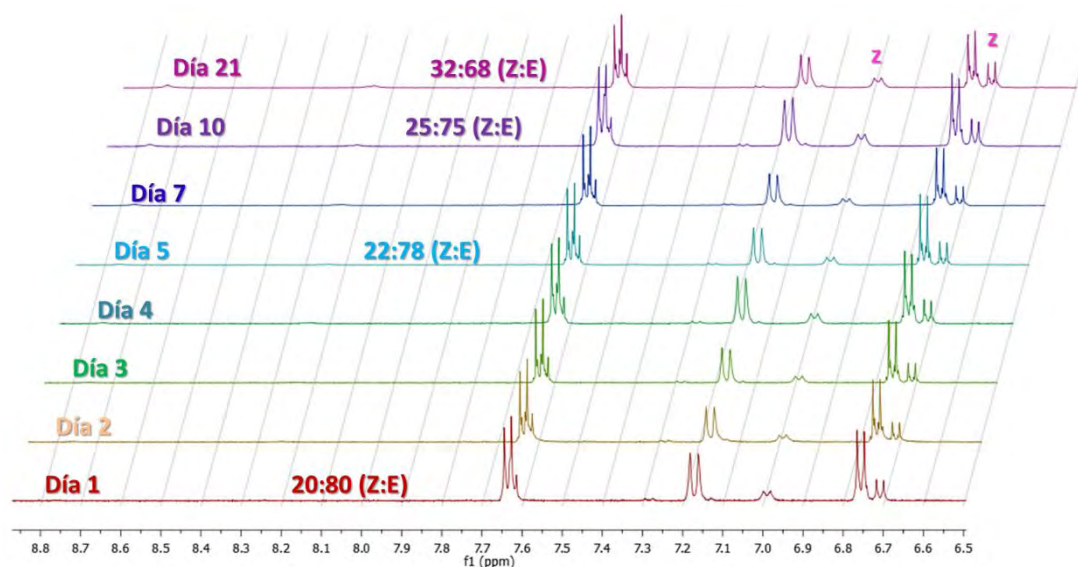


Figura 104. Ampliación do espectro de  $^1\text{H}$ -RMN para o estudo do tempo de vida media de azoFtrans.

Estes experimentos confirman que o equilibrio de reversión entre os isómeros tanto *cis*, como *trans* de ámbolos dous azos sintetizados, é moi lento. A maiores os isómeros *cis* presentan moita estabilidade termodinámica. Estas características fanos perfectos para aplicacións biolóxicas.

#### 4.3.4. Recoñecemento de ADN

##### 4.3.4.1. Dicroísmo circular

Aquelas moléculas que presentan quiralidade describen un espectro de dicroísmo circular característico ao interaccionar coa luz polarizada. Namentres, se un composto é aquiral, o seu espectro de DC (dicroísmo circular) non presenta variacións e débuxase como unha liña con valor constante, dise pois, que non dan sinal en dicroísmo circular. É ben sabido que a maioría das estruturas secundarias das moléculas de interese biolóxico son asimétricas, como é o caso do ADN. Pola contra, as dianas sintéticas obtidas neste capítulo son totalmente aquirais.

A técnica de dicroísmo circular é moi útil para caracterizar a interacción de pequenos ligandos inactivos en DC con macromoléculas quirais. Cando os ligandos se complexan coas biomoléculas provocan un cambio conformacional desta que xera un sinal inducido no espectro de dicroísmo (DCI). Para cada conformación do ADN existe un espectro característico de DC. A forma B caracterízase por presentar unha banda negativa a 245 nm e tres bandas positivas en torno aos 195, 220 e 280 nm.<sup>144</sup>

Co fin de elucidar o xeito de unión ao ADN dos azoderivados obtidos e se estes inducen algún cambio conformacional na estrutura secundaria do ADN, realizáronse varios experimentos de DC engadindo cantidades crecentes dos azos tanto na súa forma *trans* coma na súa forma *cis* sobre unha disolución 5  $\mu$ M do ADN diana e rexistráronse os respectivos espectros.

En tódolos casos se observa un efecto Cotton positivo en torno aos 330 nm, característico de moléculas que se insiren no suco menor do ADN.<sup>145</sup> A maiores, esta banda inducida só sofre un incremento na súa intensidade é dicir, non cambia de forma, polo que podemos concluír un modo de unión simple entre o ligando e o ADN.<sup>146</sup>

Sorprendentemente, para o caso do isómero *cis* do **azoF** a banda a 330 nm do espectro co ADN **A3T3** aparece moito máis intensa que para o seu análogo *trans*, o que reflexa unha unión máis forte do *Z*. A maiores, a banda negativa a 245 nm e a positiva a 280 nm desprázanse cara cero (diminúen a intensidade) a medida que aumenta a cantidade de isómero *cis* unido. Namentres, a unión do isómero *trans* provoca un cambio moito máis feble e discreto para a banda a 245 nm e a penas produce un cambio apreciable na banda a 280 nm. As diferencias entre un isómero e o outro son máis pronunciadas canto menor é a concentración de azo engadido.

Estes cambios no espectro de DC da dobre hélice de ADN producidos polo isómero *Z* reflicten un cambio conformacional na estrutura secundaria B do oligonucleótido que non se observan, case, para o isómero *trans*.<sup>147</sup> É dicir, irradiando con luz verde pode cambiarse

<sup>144</sup> a) N. C. Garbett, P. A. Ragazzon and J. B. Chaires, *Nat. Protoc.* **2007**, 2, 3166.

<sup>145</sup> N. Holmgaard List, J. Knoops, J. Rubio-Magnieto, J. Idé, D. Beljonne, P. Norman, M. Surin, M. Linares, *J. Am. Chem. Soc.* **2017**, 139, 14947.

<sup>146</sup>a) M. Rahimian, A.; Kumar, M.; Say, S. A.; Bakunov, D. W.; Boykin, R. R.; Tidwell, W.; Wilson, D. *Biochemistry* **2009**, 48, 1573. b) M. Munde, M. A.; Ismail, R.; Arafa, P.; Peixoto, C. J.; Collar, Y.; Liu, L.; Hu, M.-H.; David-Cordonnier, A.; Lansiaux, C.; Bailly, D. W.; Boykin, W.; Wilson, D. *J. Am. Chem. Soc.* **2007**, 129, 13732.

<sup>147</sup> J. Kypr, I. Kejnovska, D. Renciuik and M. Vorlickova, *Nucleic acids* **2009**, 37, 1713.

directamente a conformación de secuencias de ADN ricas en A/T debido á interacción non covalente de pequenas moléculas inseridas no suco menor destas secuencias específicas.



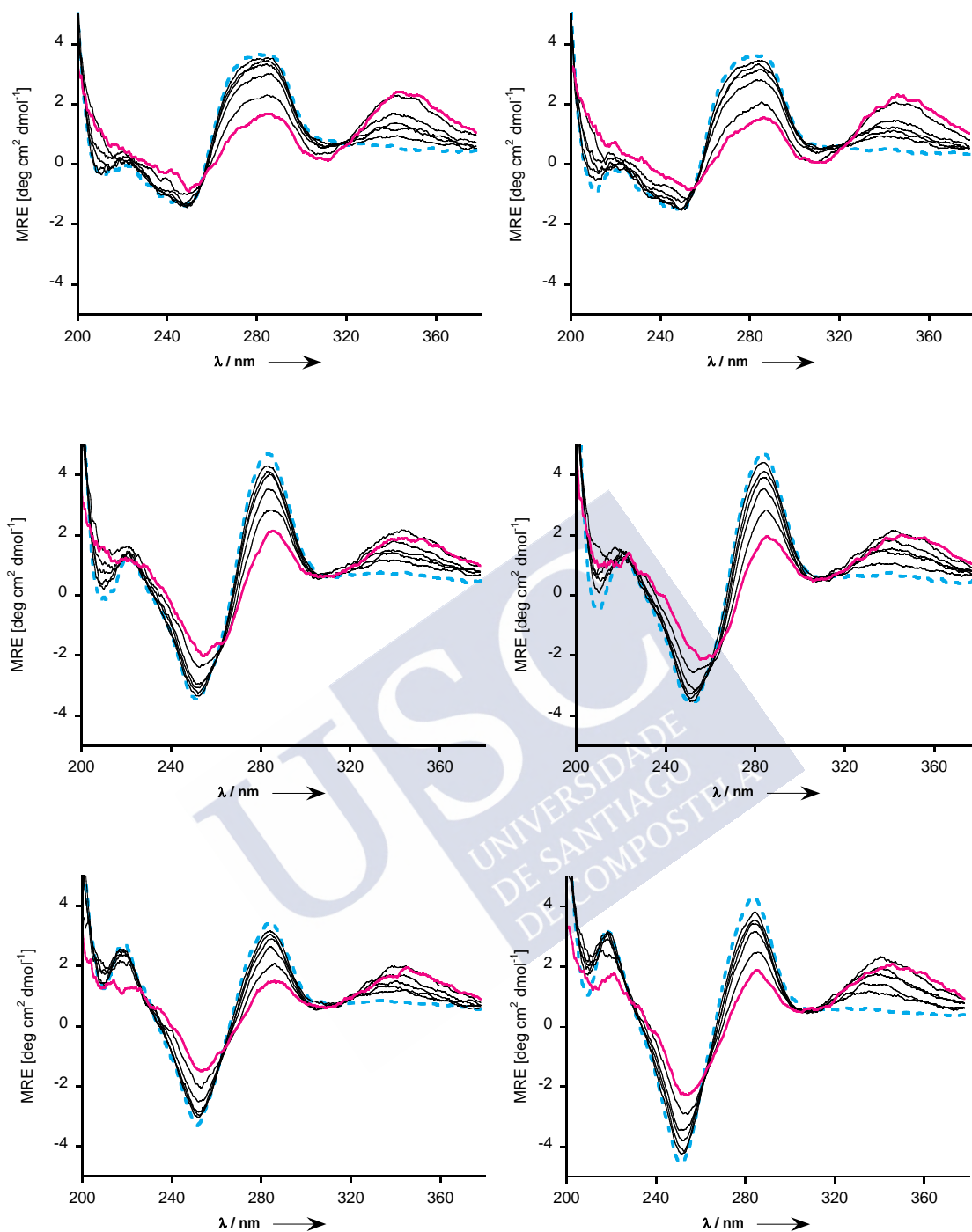


Figura 105. Comparación dos espectros de DC do azoHcis (esquerda) con azoHtrans (dereita) con diferentes secuencias de ADN. Primeira fila, con ADN GC, segunda fila con ADN A2T2 e terceira fila con ADN A3T3. Sobre unha disolución de ADN en tampón (liña azul descontinua) engádenselle sucesivamente 1 eq, 2 eq, 3 eq, 5 eq, 10 eq e 15 eq (liña rosa) do azoderivado correspondente.

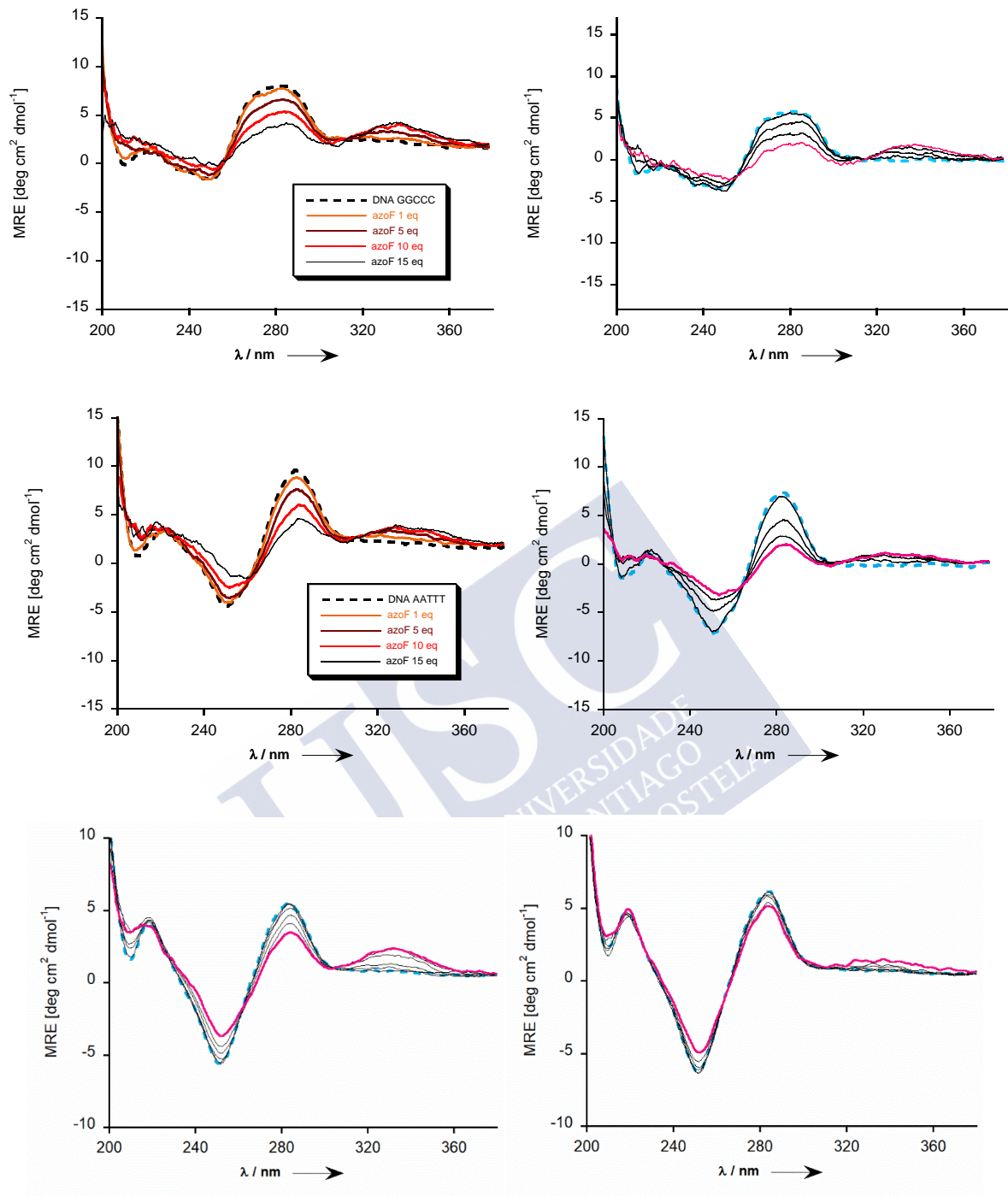


Figura 106. Comparación dos espectros de DC do azoFcis (columna esquerda) con azoHtrans (columna dereita) con diferentes secuencias de ADN. Primeira fila, con ADN GC, segunda fila con ADN A2T2 e terceira fila con ADN A3T3. Sobre unha disolución de ADN en tampón (liña azul descontinua) engádenselle sucesivamente 1 eq, 2 eq, 3 eq, 5 eq, 10 eq e 15 eq (liña rosa) do azoderivado correspondente

#### 4.3.4.2. Fluorescencia

Os compostos obtidos non presentaron fluorescencia nas concentracións usadas, nin en disolución acuosa nin en presenza de ADN. Polo que, para apoiar a hipótese de que a inserción é o tipo de mecanismo que utilizan os azoderivados sintetizados para interaccionar co ADN, realizáronse unhas competicións fluorescentes.

Este experimento baséanse na competición de dúas moléculas diferentes para acadar a súa unión a unha mesma rexión única de ADN (a unha secuencia específica). Ambos compostos deben ter a mesma secuencia diana. O modelo considera unha unión competitiva completa, o que significa que a interacción do axente non fluorescente e da sonda son exclusivas mutuamente.

Como sonda, molécula que emite o sinal fluorescente, utilizouse a BBA, un marcador solvatocrómico coñecido que se insire no suco menor.<sup>148</sup>

A unha disolución ADN saturado con BBA engadíronselle volumes crecentes doutra disolución de concentración coñecida do azo a estudar ata conseguir o desprazamento completo da sonda fluorescente.

Obsérvanse en tódolos casos un decaemento no máximo de emisión da fluorescencia da BBA, en torno aos 390 nm, a medida que aumenta a cantidade de azocomposto na disolución. Isto reforza a proposta de que o efectivamente, tanto o **azoH** como o **azoF** se insiren no suco menor do ADN.

A maiores, pode apreciarse unha maior afinidade cando se trata de secuencias ricas en A/T fronte a aquelas que presentan GC.

---

<sup>148</sup> Vázquez, O.; Sánchez, M. I.; Martínez-Costas, J.; Vázquez, M. E.; Mascareñas, J. L. *Org. Lett.* **2010**, *12*, 216. Sanchez, M. I.; Vazquez, O.; Martinez-Costas, J.; Vazquez, M. E.; Mascarenas, J. L. *Chem. Sci.* **2012**, *3*, 2383.



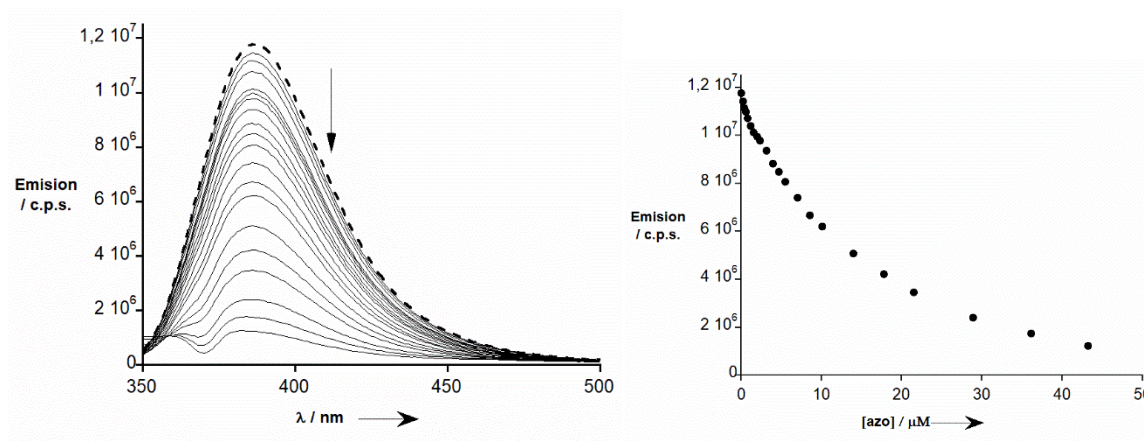


Figura 107. Esquerda: espectro da valoración de fluorescencia do complexo **BBA-A3T3 0.5 μM** despois da adición de sucesivas alícuotas do azoF en proporción maioritaria E. Dereita: Representación da intensidade de fluorescencia a 387 nm fronte á concentración de azo engadido.

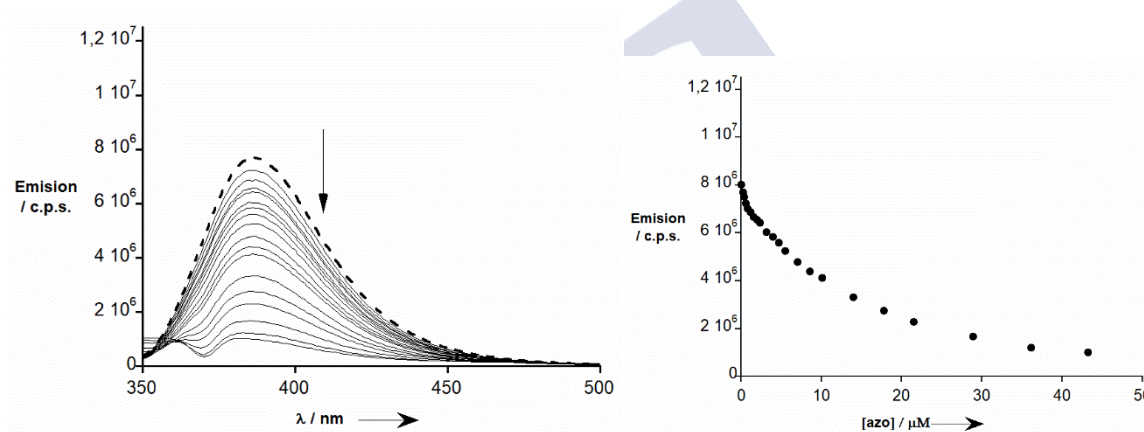


Figura 108. Esquerda: espectro da valoración de fluorescencia do complexo **BBA-A2T2 0.5 μM** despois da adición de sucesivas alícuotas do azoF en proporción maioritaria E. Dereita: Representación da intensidade de fluorescencia a 387 nm fronte á concentración de azo engadido.

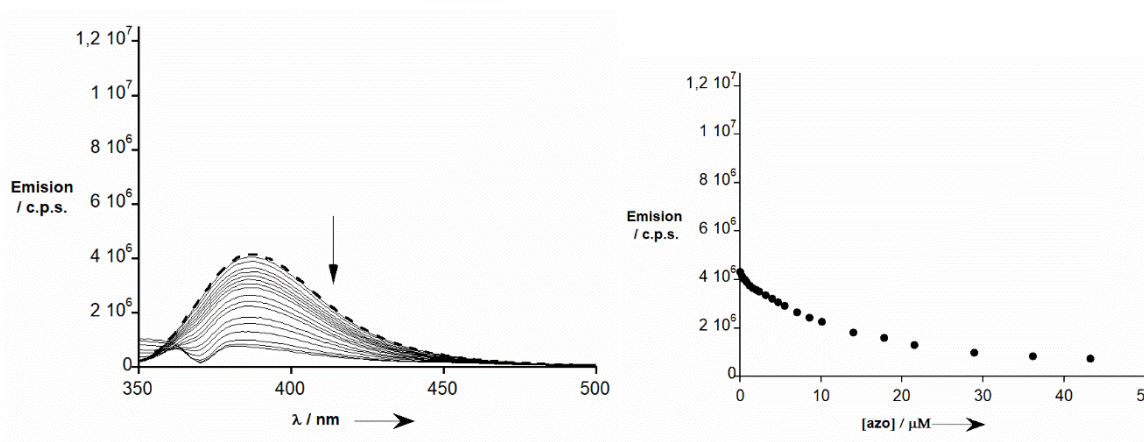


Figura 109. Esquerda: espectro da valoración de fluorescencia do complexo **BBA-GC 0.5 μM** despois da adición de sucesivas alícuotas do azoF en proporción maioritaria E. Dereita: Representación da intensidade de fluorescencia a 387 nm fronte á concentración de azo engadido.

#### 4.3.4.3. Cálculo das constantes de disociación por UV

O espectro de absorción dos compostos varía cando interaccionan co ADN. Grazas a esta propiedade, foi posible comparar cuantitativamente a afinidade de cada isómero con diferentes secuencias de ADN. A medida que aumenta a cantidade de ADN presente na disolución, a intensidade das bandas  $\pi$ - $\pi^*$  e  $n$ - $\pi^*$  diminúe como se mostra na Figura 110. Deste xeito, realizáronse valoracións de absorbencia engadindo cantidades crecentes de diferentes oligonucleótidos de dobre cadea sobre unha disolución de cada azocomposto.

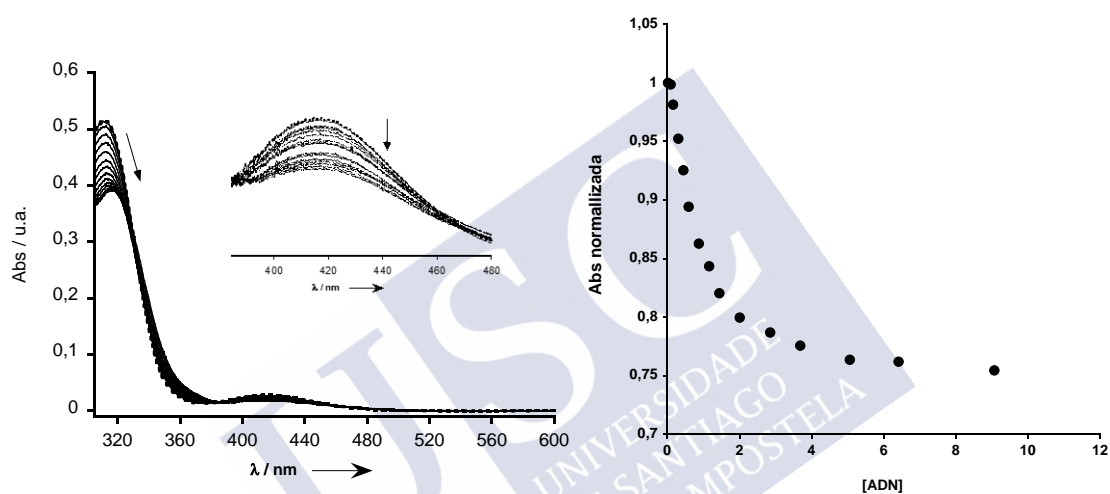


Figura 110. Esquerda: espectro da valoración de uv de azoF con A3T3. Inerise unha imaxe ampliada da rexión visible para observar mellor a variación desta banda. Dereita: regresión do decaemento da absorbencia a 313 nm normalizada, fronte á cantidade de ADN ( $\mu$ M) engadida.

Os valores das  $K_D$  calculados roldan valores de baixo micromolar ou incluso nanomolar. O que quere dicir que os azoderivados sintetizados presentan moi boa afinidade polas cadeas de ADN. Estas cifras están no rango reportado de bisbenzamidínios que presentan unión polo suco menor.<sup>148</sup> A maiores, os datos son acordes aos espectros de DC. **azoH** presenta unha maior afinidade polos ADN mentres que o **azoF** con algo menor afinidade, presenta maior selectividade por secuencias ricas en A/T sobre GC. Para oligonucleótidos do tipo **A3T3**, **azoFcis** presenta unha maior unión que o seu análogo **azoFtrans**.



Táboa 2.  $K_D$  (nM) calculadas a 20 °C para o azoH nas súas dúas formas para diferentes secuencias de ADN.

ADN	azoHtrans	azoHcis
A3T3	61.7 ± 7	128 ± 12
A2T2	367.9 ± 27	627 ± 49
GC	269 ± 36	381 ± 115

Táboa 3. Comparación das  $K_D$  (nM) calculadas a 20 °C para o azoF nas súas dúas formas, para diferentes secuencias de ADN

ADN	azoFtrans	azoFcis
A3T3	342 ± 85	236 ± 20
A2T2	1603 ± 143	1484 ± 91
GC	1541 ± 242	1385 ± 177

#### 4.4. CONCLUSIÓNS

- Sintetizáronse catro azoderivados diferentes.
- O derivado **azoH** é capaz de isomerizar baixo calor ou irradiación UV. Os tempos de vida media son bastante lentos en ámbolos dous sentidos de reconversión.
- O **azoF** pode isomerizar con calor e UV ou con luz visible. A isomerización con luz visible produce maior grao de desprazamento dos equilibrios cara o seus estados fotoestacionarios que a isomerización con UV ou calor.
- O equilibrio de reconversión do **azoFcis** cara o *trans* é moi lento, tras sete días da irradiación, aínda se atopa nun 80 % desprazado cara a forma *cis*.
- Tódolos azocompostos recoñecen ADN de dobre cadea con constantes de disociación de baixo micromolar. O modo de unión é inserción no suco menor e únense con maior afinidade a secuencias ricas en **A3T3** que CG.
- Sorprendentemente, o **azoFcis** ten máis afinidade por secuencias AAATTT que o seu análogo **azoFtrans**.
- Ao introducirse no suco menor, o **azoFcis** distorsiona en maior medida a estrutura tridimensional do ADN que o **azoFtrans**.

**5. CAPÍTULO TERCEIRO: RECOÑECIMIENTO DO SUCO  
MENOR DO ADN CON OLIGÓMEROS FLUORESCENTES DE  
MOI ALTA AFINIDADE**





## 5.1. OBXECTIVOS

As bisbenzamidinas catiónicas, como se comentou na introdución xeral, son pequenas moléculas que recoñecen secuencias de ADN ricas en pares de bases A/T sobre outras que conteñan G/C. Adoitan unirse inseríndose no suco menor da dobre hélice e posúen constantes de disociación moderadas no rango de baixo  $\mu$ M. Malia que non teñen moita selectividade dentro das rexións con diferentes combinacións de A/T, presentaron aplicacións terapéuticas como axentes antiparasitarios. Esta aplicación xorde, precisamente, de que as bisbenzamidinas interaccionan con ADN mitocondrial dos parasitos que posúen cadeas con pares de A/T repetidas.<sup>149</sup> Varias enfermidades hereditarias dexenerativas teñen en común que se repiten determinadas secuencias do xenoma. Ditas expansións das secuencias do xenoma desestabilizan e deforman a dobre hélice do ADN (forman nodos ou desemparellan pares de bases) provocando erros na etapa de replicación e que provocan graves afeccións, é o caso da distrofia muscular, a enfermidade de Huntington ou a ataxia de Friedreich.<sup>150</sup>

Previamente, no noso grupo de investigación desenvolveuse unha síntese fácil e eficiente para sintetizar *aza*-bisbenzamidinas nun único paso e con rendementos moi bos. Estas moléculas presentaban fluorescencia unha vez unidas a rexións de ADN ricas en A/T con afinidades en torno a 1  $\mu$ M.

En base a isto, o obxectivo principal deste capítulo é a síntese de oligómeros de bisbenzamidinas co gallo de recoñecer secuencias máis longas de ADN e con maior afinidade. Deste xeito, espérase que as subunidades de BBA actúen cun efecto cooperativo incrementando a intensidade de emisión e a afinidade por secuencias de ADN diana. Os oligómeros poderían usarse para marcar e localizar, ou incluso estabilizar, aquelas rexións de ADN con secuencias do tipo  $(A_xT_y)_n$  repetidas  $n$  veces, causantes de moitas enfermidades. Así como terían unha maior selectividade e levarían a cabo accións antiparasitarias con menos efectos secundarios.

---

<sup>149</sup> a) D.W. Boykin, *J. Braz. Chem. Soc.* **2002**, *13*, 763; b) A.H. Fairlamb, *Trends Parasitol.* **2003**, *19*, 488.

<sup>150</sup> A) P.I. Patel, G. Isaya, *Am. J. Hum. Genet.* **2001**, *69*, 15. B) S.M. Mirkin, *Nature* **2007**, *447*, 932. C) S.M. Mirkin, *Nature* **2007**, *447*, 932.

## 5.2. SÍNTESE:<sup>151</sup>

### 5.2.1. Estudos de oligomerización

Para os primeiros ensaios escolleuse como monómero a *aza*- bisbenzamidina cun anel fenílico entre os residuos de benzamidinos, **1a**, xa que se trata dun derivado sinxelo e con boa afinidade por secuencias de ADN ricas en A/T, como se reportara con anterioridade no grupo de investigación.<sup>82</sup>

De aquí en diante referiranse os derivados de bisbenzamidinas como BBA. Para unir os monómeros usouse como conector un dibromuro comercial barato, **2a**.

Comezouse optimizando as condicións da reacción de dimerización (**Táboa 4**).

A **mBBA** é soluble en disolventes acuosos, metanol ou acetonitrilo. Estes medios non son compatibles con moitas das bases utilizadas por iso se empregaron o DMSO e a DMF na maioría das entradas da táboa anterior.

Usando exceso de **mBBA** e sen empregar base, non se detectou a formación de ningún dos oligómeros (entrada 1). Usando K<sub>2</sub>CO<sub>3</sub>, piridina ou <sup>t</sup>BuOK a temperatura ambiente (entradas 2, 3 e 4), non se observou a formación do produto desexado ou só se observaron trazas. Se aumentamos a temperatura empregando de base <sup>t</sup>BuOK, si se chega a observar por HPLC-MS un pico minoritario que se corresponde co dímero desexado pero resulta imposible separalo dunha mestura complexa de outros subprodutos que non se chegan a identificar.

Ao utilizar bases máis fortes coma o *n*-BuLi (entrada 6), non se observan produtos por HPLC-MS. Neste caso, o cru de reacción presenta un cromatograma moi sucio con unha montaña de picos inseparables provenientes probablemente da polimerización da **mBBA**.

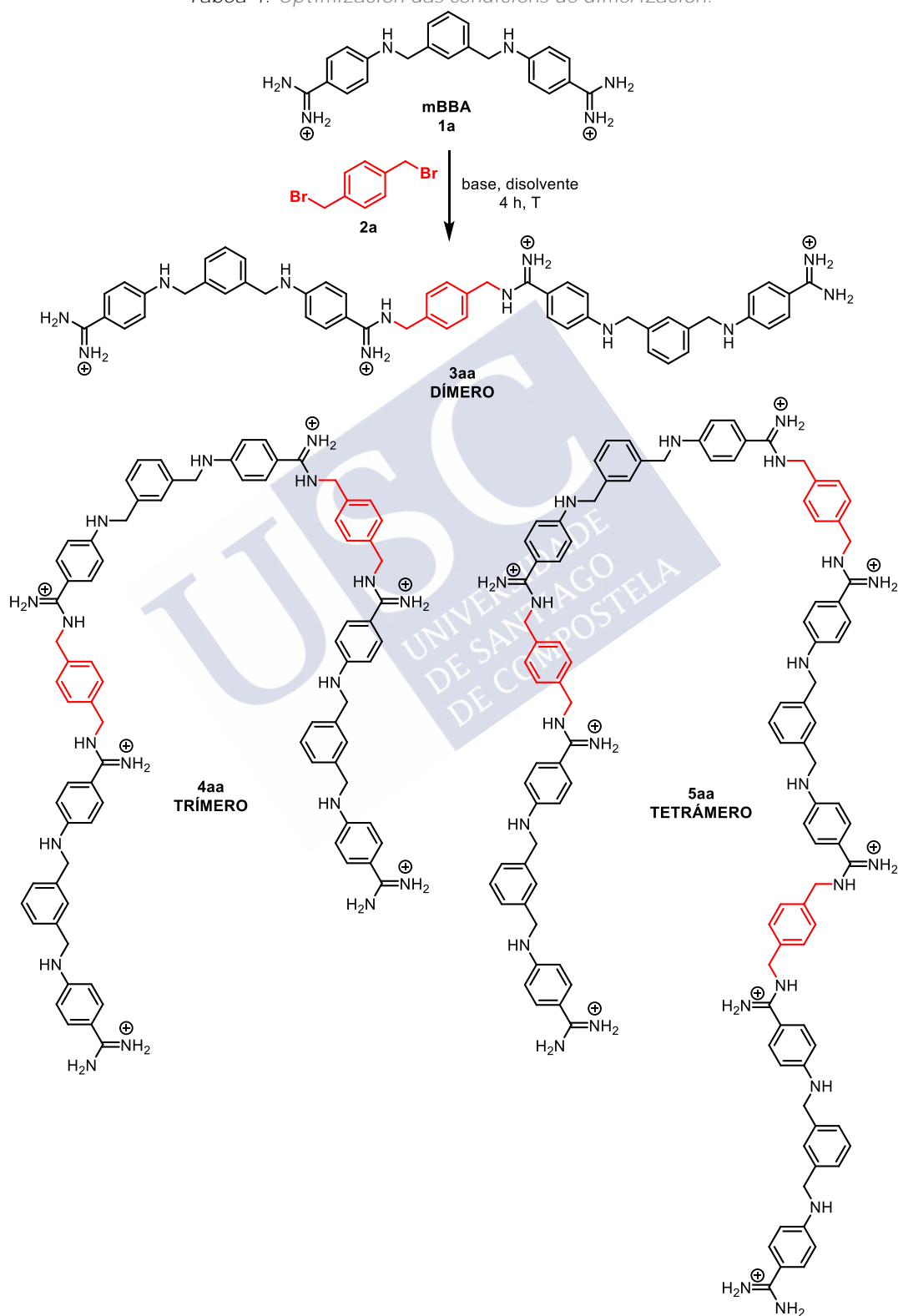
Cando se usa NaH sí se observa a formación de dímero como produto maioritario tanto ás 4 h como ás 24 h (entradas 7 e 8, **Táboa 4**). Ás 24 h de reacción cunha concentración 100 mM (entrada 8) consegue illarse un 20 % do dímero **3aa**, 2 % de trímero **4aa** e trazas de tetrámero **5aa**. O mellor resultado obtívose engadindo un lixeiro exceso de NaH en relación aos equivalentes de **mBBA** empregados (entrada 9, **Táboa 4**). Tras 4 h de reacción, mediante HPLC-MS detectouse a formación de tódolos oligómeros. Despois de purificar por HPLC en fase revertida, obtívose un 40 % de dímero **3aa** como produto maioritario e tamén se illou o

---

<sup>151</sup> En colaboración co doutor Mateo I. Sánchez.

trímero **4aa** nun 18 % de rendemento. Lamentablemente, só se lograron separa trazas do tetrámero **4aa**. Aumentando a temperatura nestas mesmas condicións non se obtivo ningún cambio significativo na obtención do dímero nin na formación do trímero.

Táboa 4. Optimización das condicións de dimerización.<sup>a</sup>



Entrada	Base	eq <sub>base</sub>	eq <sub>BBA</sub>	disolvente	Temperatura(°C)	3aa(%)
1	-	-	3	DMSO	t. a.	0
2	K <sub>2</sub> CO <sub>3</sub>	20	4	DMF	70	trazas
3	Piridina	9.3	3	DMF	100	0
4	<sup>t</sup> BuOK	9.3	3	DMF	t. a.	0
5	<sup>t</sup> BuOK	9.3	3	DMF	100	trazas
6 <sup>d</sup>	<i>n</i> -BuLi	6.1	2	THF	-78→t. a.	0
7	NaH	8	4	DMSO	t. a.	15
8 <sup>b</sup>	NaH	8	4	DMSO	t. a.	20
9 <sup>c</sup>	NaH	13	4	DMSO	t. a.	40
10 <sup>c</sup>	NaH	13	4	DMSO	70	38

<sup>a</sup>Reaccións levadas a cabo nunha escala de 0.02 mmol (1 eq) de dibromuro. e concentración 50 mM. Seguíronse por HPLC-MS. <sup>b</sup>24 h de reacción, 100 mM. Íllanse trazas de trímero e tetramero. <sup>c</sup>100 mM, illase trímero **4aa** e trazas de tetramero **5aa**. <sup>d</sup>Engádesse a base lentamente a -78 °C e déixase axitando durante 45 min. A continuación engádesse unha disolución de dibromuro gota a gota e déixase que acade t. a.

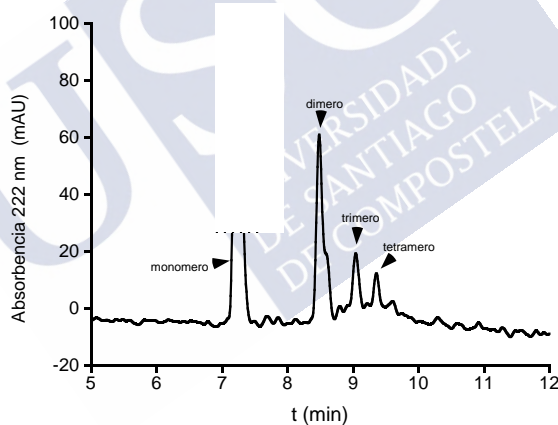


Figura 111. Cromatograma a 222 nm do cru de reacción para a entrada 9 da Táboa 4.

### 5.2.2. Uso doutros conectores

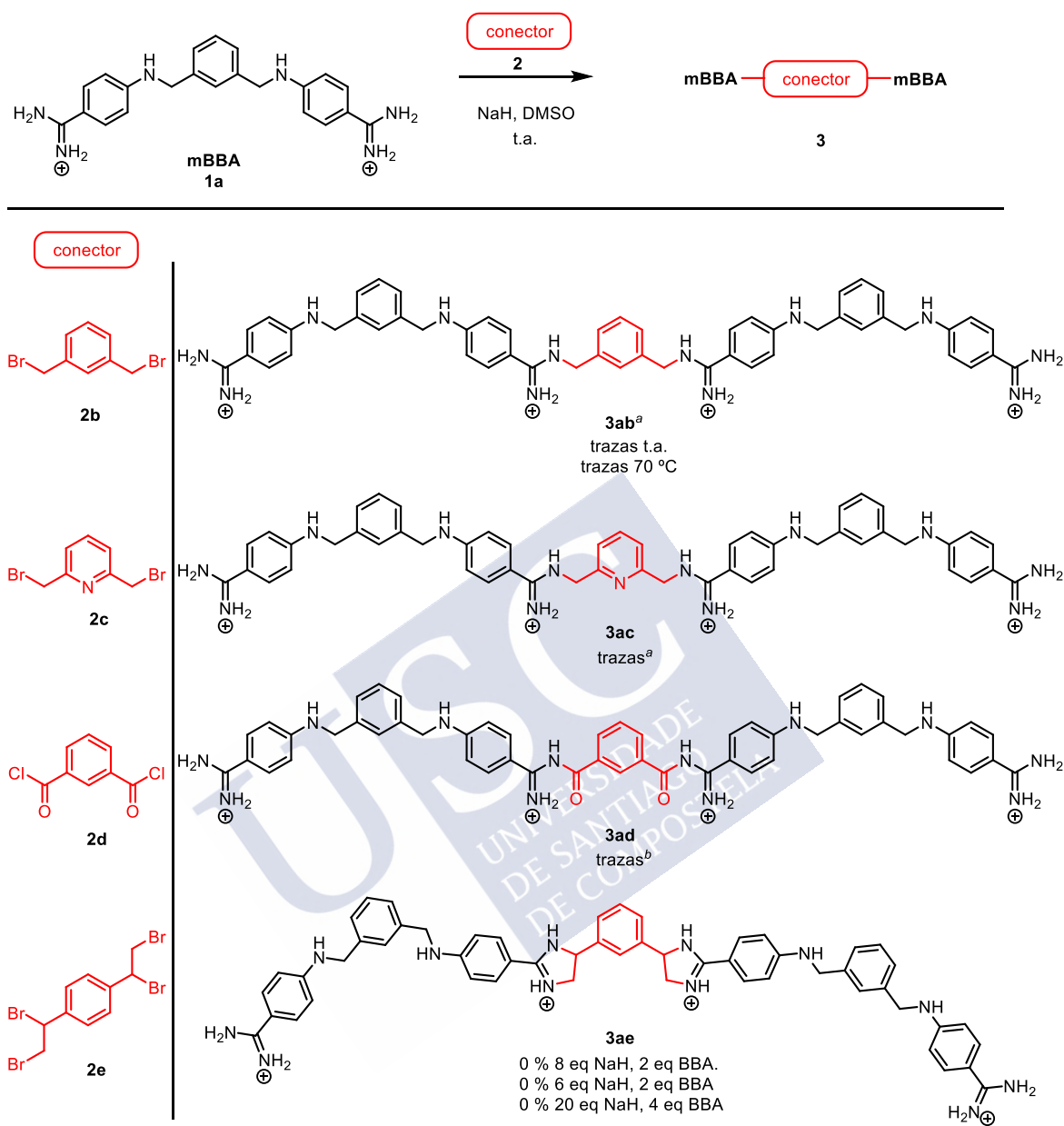
Unha vez atopadas as mellores condicións de reacción, procedeuse a estudar o seu alcance e limitacións empregando diferentes conectores entre as bisbenzamidinas (Táboa 5) para intentar aumentar o rendemento da reacción e/ou conseguir tamén, diferentes estruturas de oligómeros. Para isto, mantívose fixa a unidade de BBA e fíxose reaccionar con varios tipos de dibromuros comerciais aplicando as condicións óptimas, resaltadas en laranxa, da Táboa 4.



Como se comentou no anteriormente, usando o *1,4-Bis(bromometil)benceno* **2a** o dímero **3aa** estaba completamente formado ás 4 h, seguindo a reacción por HPLC-MS. Ás 24 h, o dímero seguía sendo o produto maioritario pero tamén se observaron picos ben diferenciados correspondentes á formación do trímero **4aa** e do tetrámero **5aa**. Namentres, ao utilizar o derivado do xileno substituído en meta (ver **Táboa 5**), dibromuro **2b**, a reacción non foi tan ben, observándose múltiples picos no cromatograma de HPLC. Despois da purificación en fase revertida, o composto **3ab** non puido separarse puro. Porén, repetiuse a reacción aumentando a temperatura ata 70 °C mais tan só puido illarse trazas do dímero **3ab**. Non se observou a formación doutros oligómeros. Ao empregar como conector o análogo de piridina disubstituído **2c**, obtívose o correspondente produto **3ac** pero nun rendemento ínfimo. De novo, obtívose unha mestura complexa de picos por HPLC-MS que pode deberse á polimerización excesiva das BBA. Do dímero **3ad** do dicloruro de ácido tan só se detectaron trazas por HPLC-MS e non se illou.

Co gallo de evitar posibles polimerizacións indesexadas, pensouse en utilizar o tetrabromuro **2e** como conector. Con esta estrutura formaríase un imidazol o cal bloquearía os dous nitróxenos amidínicos de posibles ramificacións. A pesares de que se probaron diversas condicións, o produto desexado **3ae** non chegou a observarse en ningún momento.

A excepción de para **4aa e 5aa**, en ningún dos outros casos se detectou formación de oligómeros máis longos que o dímero.

Táboa 5. Alcance do conector na reacción de oligomerización.<sup>c</sup>

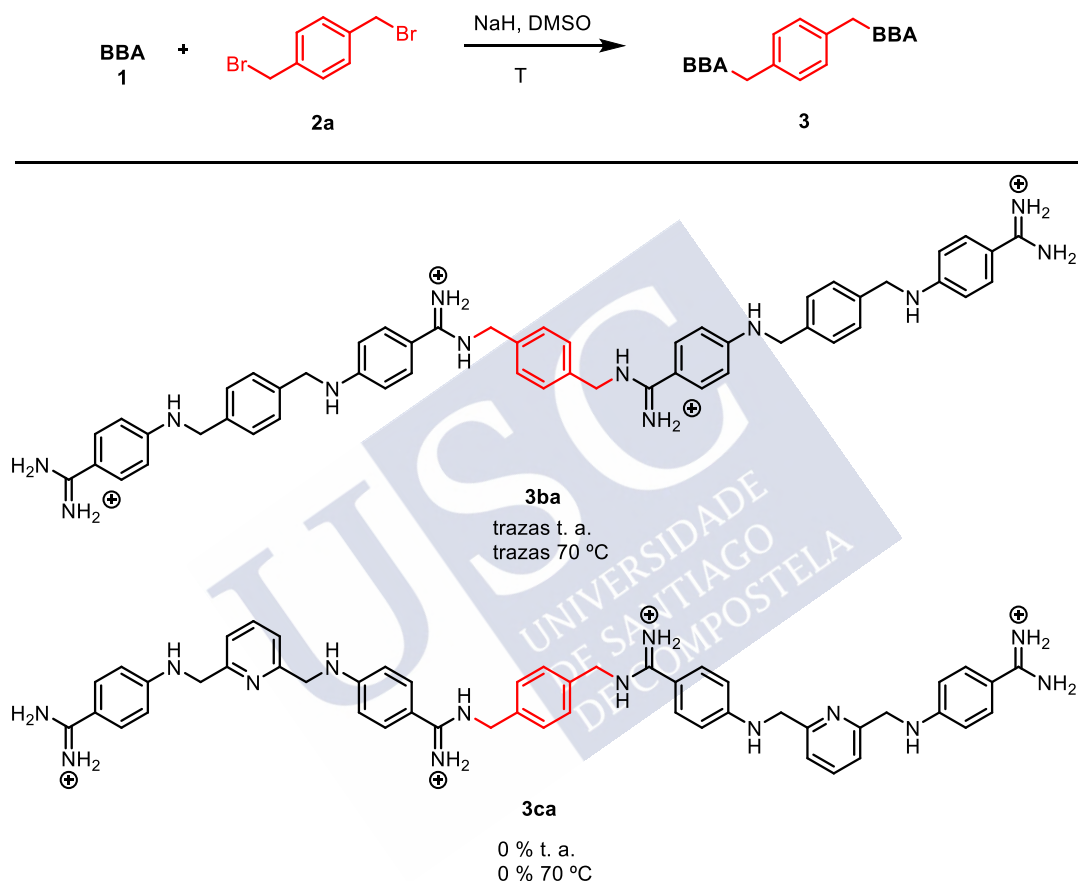
<sup>a</sup>Illado. <sup>b</sup> Detectado por HPLC-MS. <sup>c</sup> Reaccións levadas a cabo nunha escala de 0.02 mmol de conector. Seguironse por HPLC-MS ás 4 h e ás 24 h.

### 5.2.3. Uso doutros monómeros de BBA

Por último, comprobouse a capacidade de oligomerización doutros monómeros usando as mellores condicións de reacción atopadas ata o momento. Para isto, probouse o monómero da *p*BBA e o derivado de piridina da BBA. Como se observara no grupo de investigación previamente, este último composto é quen de unirse a secuencias do tipo *AAGTT* con maior afinidade que os análogos fenólicos. De acadarse o dímero **3ca**, poderían recoñecerse diferentes secuencias de ADN con moi alta afinidade.

Cando se empregou a **pBBA** como monómero **1b** tanto ás 4 h como ás 24 h a temperatura ambiente, tan só se apreciaron trazas do dímero correspondente **3ba** dentro dunha mestura complexa de subprodutos. Ao aumentar a temperatura non se observaron melloras no avance da reacción. Os produtos correspondentes ao análogo da BBA con piridina, non se observaron en ningún dos casos.

Táboa 6. Alcance do monómero na reacción de oligomerización.<sup>a</sup>



<sup>a</sup>Reaccións levadas a cabo nunha escala de 0.02 mmol de diBromuro e foron seguidas por HPLC-MS ás 4h e ás 24h.

#### 5.2.4. Outras probas de obtención dos oligómeros

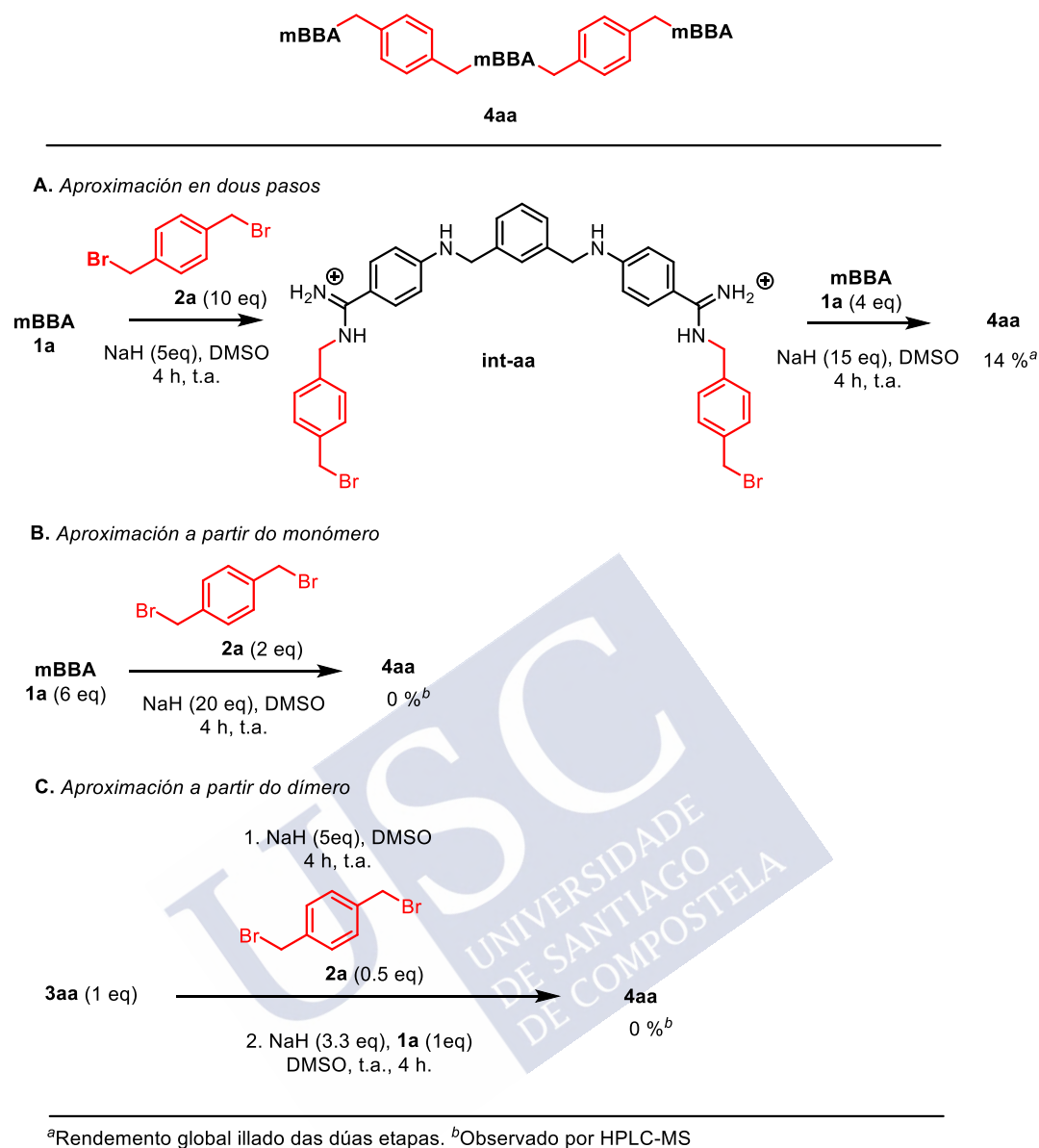
Como se comentou na **Táboa 5**, trala reacción de **1a** con **2a** puido identificarse a formación do derivado trimérico **4aa**. Nun intento de mellorar o rendemento e optimizar a obtención do trímero **4aa**, leváronse a cabo diferentes estratexias que se recollen no **Esquema 29**.

Na *aproximación en dous pasos* (**Esquema 29 A**) sintetizouse primeiramente un derivado dibromado da BBA, ao que se denominou **Int-aa**. Despois da purificación por cromatografía

en fase revertida, o intermedio **Int-aa** eluíu xunto cunha impureza co mesmo tempo de retención e que non absorbía no ultravioleta. A pesar disto, decidiuse empregar na segunda etapa. O **int-aa** actuaría como conector nesta segunda parte da aproximación. Para realizar a reacción seguironse as condicións descritas na entrada 8 da **Táboa 4**. Este paso tamén se purificou por fase revertida e obtívose o trímero desexado **4aa** nun rendemento global do 14 %. O baixo rendemento pode deberse á complexidade da purificación da primeira etapa.

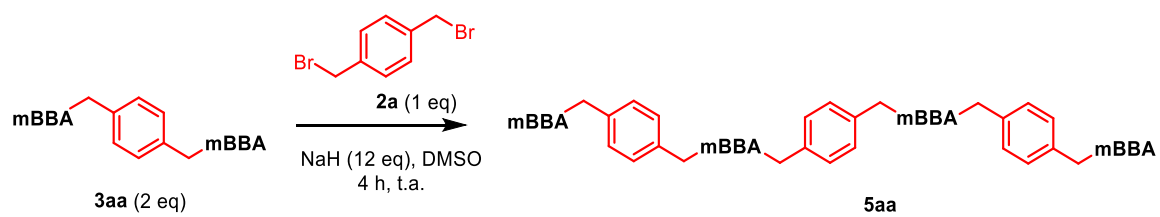
Por outra banda, tanto para a *aproximación a partir do dímero* como para a *aproximación a partir do monómero* (**Esquema 29B e C**), non se detectou por HPLC-MS a formación do trímero nin tetrámero en ningún dos casos. Estas dúas estratexias produciron crús de reacción bastante complexos, poida que fora debido a polimerizacións indesexadas.





Esquema 29. Probas para a obtención do trímero.

Para obtención do tetrámero probouse, sen éxito, a reacción de unión de dous dímeros usando como conector o dibromuro **2a** (ver **Esquema 30**). As condicións son análogas ás da entrada 9 da **Táboa 4** pero usando só 2 equivalentes de dímero, debido ao custo de obtención do mesmo.



Esquema 30. Proba de obtención do tetrámero.



### 5.3. RECOÑECIMENTO DE ADN

#### 5.3.1. Dicroísmo circular

Os experimentos de dicroísmo circular realizáronse para elucidar a forma de unión ao ADN, como se explicou no capítulo anterior. Os oligómeros non presentan quiralidade polo que non teñen espectros característicos de dicroísmo circular. Cando interaccionan co ADN aparece unha banda en torno a 330 nm característica da inserción do oligómero no suco menor da dobre hélice. Como se mostra nas **Figura 112** e **Figura 113**, ámbolos dous oligómeros presentan bandas inducidas máis pronunciadas con ADN ricos en A/T, do que se deduce que hai maior interacción con ditas secuencias que con secuencias ricas en GC.

Amais, obsérvanse pequenas variacións inducidas na rexión do espectro do ADN. Concretamente, a medida que aumenta a cantidade de oligómero presente na disolución, hai unha diminución da intensidade (desprazamento cara cero) das bandas en torno a 245 nm e a 280 nm, indicando perturbacións na estrutura do ADN, probablemente debido á inserción dos biomarcadores

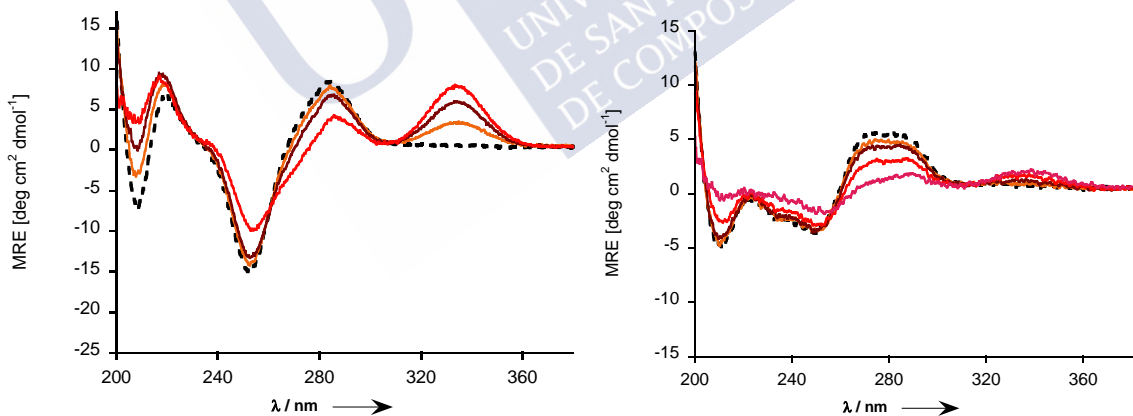


Figura 112. Espectros de dicroísmo circular dunha disolución 5  $\mu\text{M}$  de ADN en tampón Tris-HCl 20 mM, 100 mM NaCl, pH 7.5 (liña punteada) na presenza de 1, 2, 5 e 10 equivalentes de dímero 3aa (liñas sólidas) Esquerda: ADN ( $A_3T_3GC$ )<sub>2</sub>. Dereita ADN ( $G_3C_2$ )<sub>2</sub>.



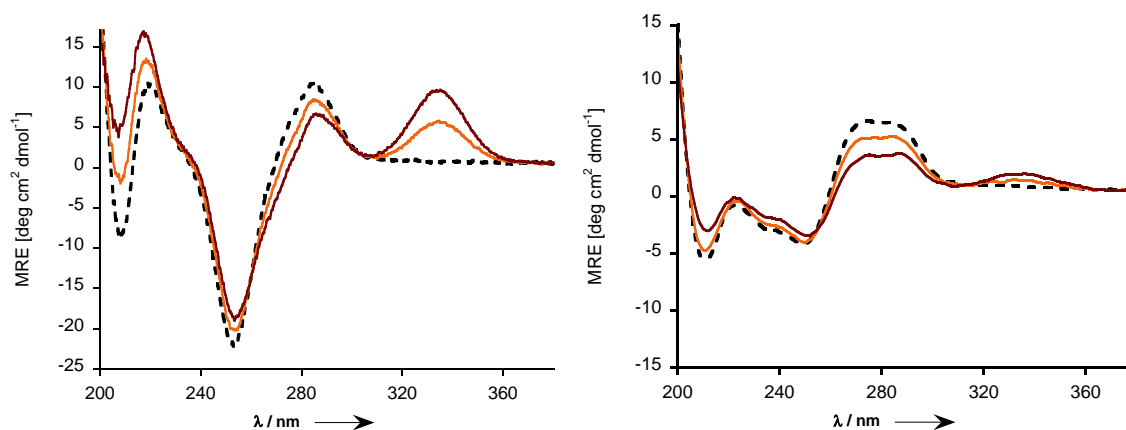


Figura 113. Espectros de dicroísmo circular dunha disolución 5  $\mu\text{M}$  de ADN en tampón Tris-HCl 20 mM pH 7.5, 100 mM NaCl (liña punteada) na presenza de 1 e 2 equivalentes de trímero 4aa (liñas sólidas) Esquerda: ADN ( $\text{A}_3\text{T}_3\text{GC}$ )<sub>3</sub>. Dereita ADN ( $\text{G}_3\text{C}_2$ )<sub>3</sub>.

### 5.3.2. Fluorescencia

Tanto o dímero **3aa** coma o trímero **4aa** foron sometidos a experimentos de espectroscopía de fluorescencia. Os oligómeros a penas presentaron emisión de fluorescencia residual detectable por si mesmos, mais resultaron ser fortemente emisivos cando se incubaron con alícuotas de ADN ricas en A/T. É dicir, actúan como biomarcadores solvatocrómicos que emiten luz unha vez que recoñecen o ADN diana.

Despois de realizar múltiples valoracións fluorescentes entre os oligómeros obtidos e diferentes secuencias de ADN, pode concluírse que:

1. *Os oligómeros teñen afinidades moi altas polos ADN diana. Para calcular o seu valor é preciso que a concentración á que se fai a valoración sexa moi pequena.*

Isto conleva unha diminución drástica na intensidade do sinal fluorescente chegando, ás veces, a verse afectada polo ruído de fondo dada a limitada sensibilidade dos aparellos.

2. *O equilibrio de unión entre os oligómeros e o ADN ten unha cinética moi lenta.*

Para que o composto **4aa**:ADN acade unha intensidade de emisión estable precísase unha media de 3 minutos. A **Figura 114** mostra un exemplo de cinética realizado co ADN ( $\text{A}_3\text{T}_3\text{GC}$ )<sub>2</sub>.

Para a complexación entre trímero e ADN, o equilibrio aínda é mais lento, acadándose entre os 4 - 5 min.

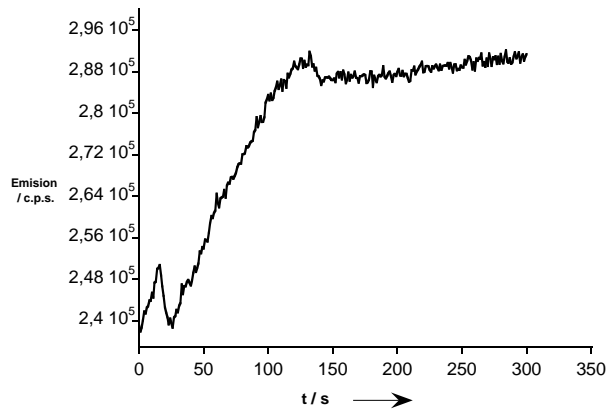


Figura 114. Cinética da emisión do equilibrio de complexación dímero:ADN realizada cunha disolución 15 nM de ADN e 30 nM de dímero en disolución tampón Tris-HCl 20 mM pH 7.5, 100 mM NaCl.

Porén as medidas de fluorescencia foron realizadas esperando 4 minutos de incubación para cada medida nas valoracións do dímero **3aa** e 6 minutos para valoracións do trímero **4aa**.

3. *As valoracións de fluorescencia chegan á saturación a diferentes ratios oligómero: ADN dependendo da concentración que se estea usando.*

Como se mostra na **Figura 115 esquerda**, ao valorar unha disolución 100 nM de dímero **3aa**, o sistema acada a saturación entre 40-50 nM de ADN. Este resultado coincide cun valor estequiométrico 2:1 de **3aa**:ADN. Namentres, se se diminúe a concentración de dímero ata 15 nM, o sistema acada a saturación en torno aos 12 nM de concentración de ADN. Desta forma, a estequiometría aproxímase máis a 1:1 (ver **Figura 115 dereita**)

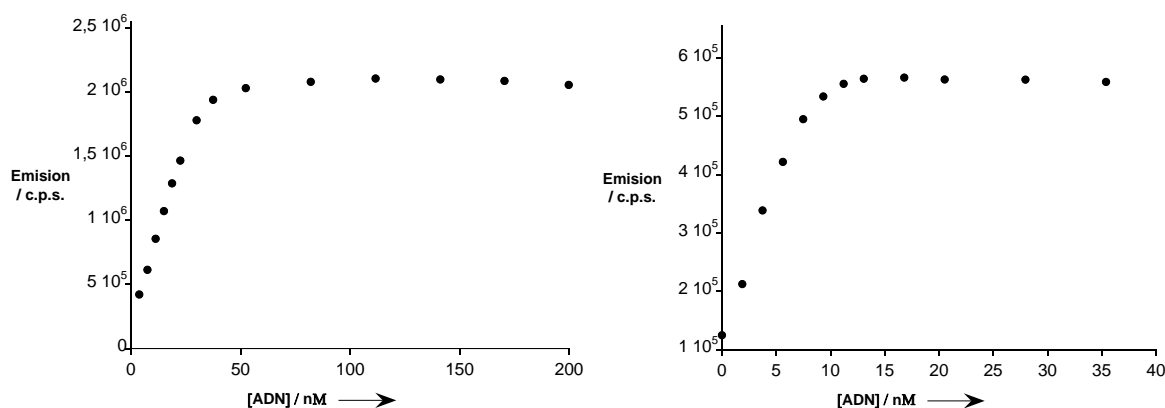


Figura 115. Representación da intensidade de fluorescencia de dímero **3aa** a 398 nm fronte á concentración de ADN  $(A_3T_3GC)_2$  engadida. Esquerda: disolución 100 nM de **3aa**. Dereita: disolución 15 nM de **3aa** en disolución tampón Tris-HCl 20 mM pH 7.5, 100 mM NaCl.

Este fenómeno pode que se deba ao recoñecemento de ADN mediante diferentes complexos de dímero, todos en equilibrio. A baixas concentración estaría favorecida a unión dunha única molécula de dímero **3aa** á secuencia de ADN diana, estequiometría 1:1. Namentres, para concentracións máis elevadas, un posible apilamento entre os compostos **3aa** faría que se dúas moléculas de dímero apiladas se insiran, á vez, dentro do mesmo suco menor, estequiometría 2:1 **3aa**:ADN.

4. *As curvas de valoración poden ser diferentes baixo as mesmas condicións experimentais.*

Se se observa a gráfica da **Figura 116**, contéplase que para reproducións baixo as mesmas condicións experimentais dunha valoración de dímero **3aa** con ADN ATTCT, obtéñense tres curvas bastante diferentes. Primeiro en intensidade de emisión, e segundo en valor de saturación. As tres curvas chegan á saturación do sinal a valores no intervalo dende 50 ata 100 nM de ADN, o que se correspondería con diferentes estequiometrías. Unha posible explicación para tal resultado é que os oligómeros interactúen co ADN formando varios complexos dentro dunha mesma estequiometría. Cada modelo de unión ten unhas propiedades fluorescentes diferentes, polo que dependendo do complexo que estea favorecido en cada momento a intensidade de emisión pode variar e dar lugar a interpretacións erróneas dos datos.

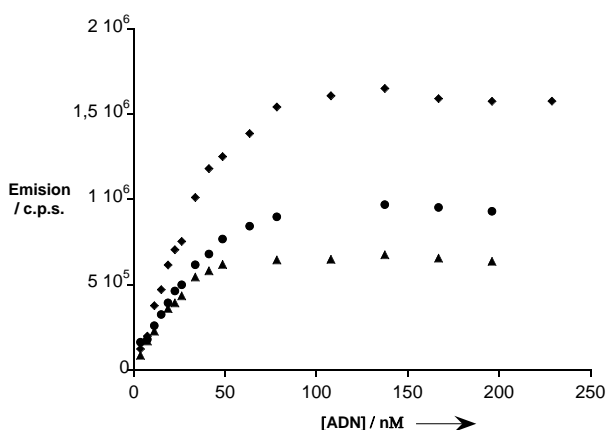


Figura 116. Representación da intensidade de fluorescencia dunha disolución 100nM de dímero 3aa a 398 nm, tampón Tris-HCl 20 mM pH 7.5, 100 mM NaCl fronte á concentración de ADN (ATTCT) engadida. Tres experimentos diferentes nas mesmas condicións.

Os problemas son máis pronunciados para o trímero **5aa** posto que a unión é moito máis forte e o sistema de interacción co ADN máis complexo.

Tendo en conta tódolos puntos anteriores, puideron realizarse valoracións fluorescentes reproducibles con diferentes cadeas de ADN. Os datos experimentais para o dímero **3aa**, axustáronse coa ecuación de Hill e obtivéronse os valores que se mostran na **Táboa 7**. En tódolos casos o valor para o coeficiente de Hill ( $n$ ) é superior a 1, o que indica que hai cooperatividade entre as moléculas de dímero unidas ao ADN así como, que existe unión múltiple deste oligómero.<sup>152</sup>

As constantes de disociación obtidas atópanse no rango de baixo nanomolar en tódolos casos. Queda patente pois, que o dímero **3aa** presenta forte afinidade por secuencias con pares de bases A/T. A maiores, **3aa** reconece de maneira extraordinaria secuencias con repeticións do tipo (ATTCT)•(AGAAT) asociada á ataxia espinocerebelosa tipo 10 (SCA10).<sup>153</sup>

<sup>152</sup> H. Prinz, *J. Chem. Biol.* **2010**, 3, 37.

<sup>153</sup> V.N. Potaman, J.J. Bissler, V.I. Hashem, E.A. Oussatcheva, L. Lu, L.S. Shlyakhtenko, Y.L. Lyubchenko, T. Matsuura, T. Ashizawa, M. Leffak, C.J. Benham, R.R. Sinden, *J. Mol. Biol.* **2003**, 326, 1095.

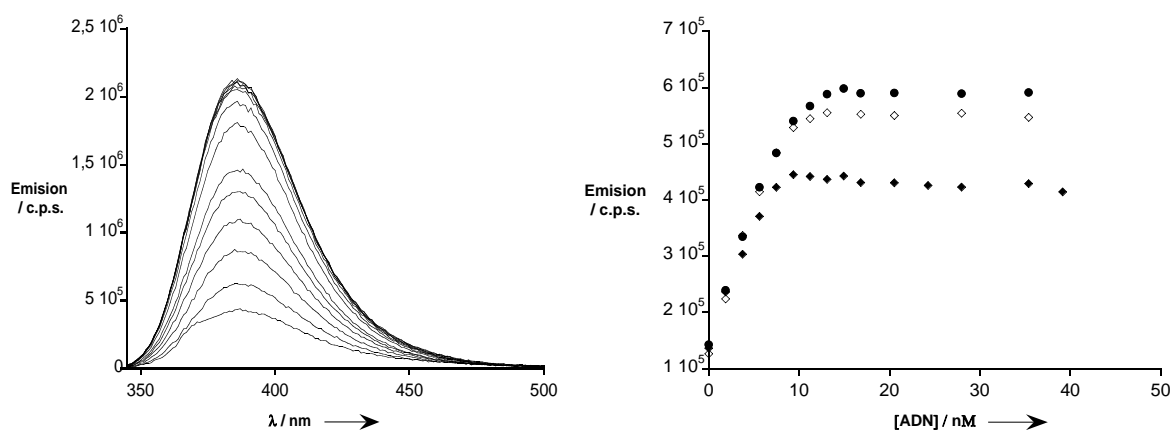


Figura 117. Esquerda: espectro emisión 3aa con  $(A_3T_3GC)_2$ . Dereita: Representación da intensidade de fluorescencia dunha disolución 15 nM de dímero 3aa en tampón Tris-HCl 20 mM pH 7.5, 100 mM NaCl fronte á concentración de ADN  $(A_3T_3GC)_2$  rombos brancos,  $(A_3T_3C)_2$  círculos e  $(A_2T_3GC)_2$  rombos negros.

Táboa 7. Valores constantes de disociación calculadas a partir das valoracións fluorescentes. Exprésanse en nM.  $n$  representa o coeficiente de Hill.

ADN	$(ATTCT)_2$	$(A3T3GC)_2$	$(A3T3C)_2$	$(A2T3C)_2$
$K_D$	$27.6 \pm 7.4$	$21.5 \pm 1.5$	$32.4 \pm 1.3$	$16.5 \pm 1$
$n$	1.8	2.43	1.6	2.8

Actualmente, estase traballando en determinar os valores de  $K_D$  para a unión entre o trímero **4aa** e os seus ADN diana correspondentes. Como se pode ver na **Figura 118** e como cabía esperar, **4aa** ten altísima afinidade por secuencias de ADN de dobre febra con tres rexións de unión ricas en A/T. Esta regresión non se axusta a un modelo 1:1 pero si se puido estimar da curva unha  $K_D$  que rolda o valor de 30 picomolar. Resultados experimentais preliminares, respaldan a hipótese de que se poden formar diferentes aductos trímero **4aa**:ADN. Ao igual que para **3aa**, segundo a concentración, estarían favorecidos un ou outro tipo de unión. O que se traduce nun sistema moi complexo de analizar, con diferentes modelos de unión con estequiometría maior que 1 e altísima afinidade.

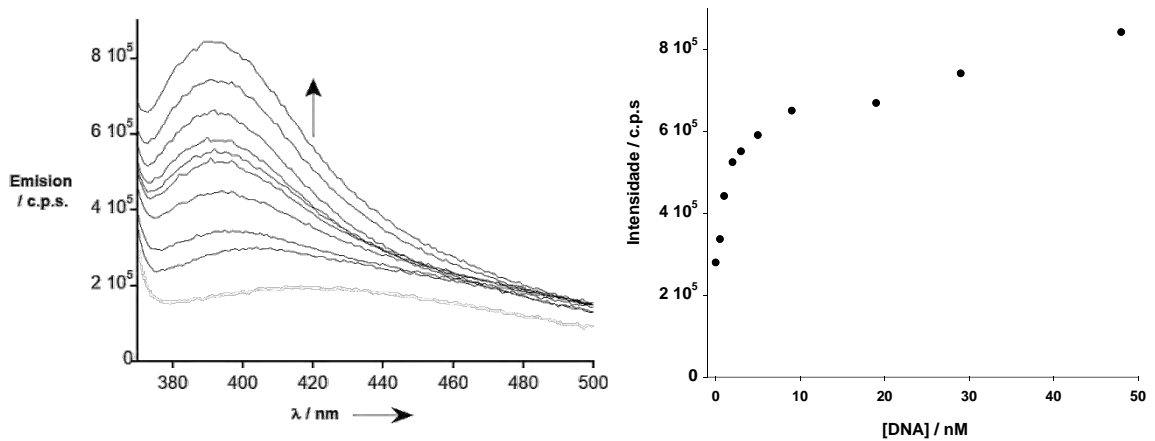


Figura 118. Esquerda: espectro da valoración de fluorescencia do complexo 10 nM de trímero 4aa despois da adición de sucesivas alícuotas do ADN  $(A_3T_3GC)_2$ . Dereita: Representación da intensidade de fluorescencia a 398 nm fronte á concentración de ADN engadida.

### 5.3.3. Experimentos de desnaturalización por temperatura: *Meltings*

Os resultados anteriores foron corroborados con experimentos de desnaturalización do ADN. Ao quentar unha mostra con ADN de dobre cadea hibridado, as pontes de hidróxeno que manteñen unidas as dúas cadeas debilítanse, provocando que a dobre hélice se desenrole e chegue a romperse e dividirse en dúas febras separadas. Este fenómeno coñécese como desnaturalización, fusión ou *melting* (en inglés) do ADN. A temperatura na que a metade das moléculas de ADN están desnaturalizadas denomínase **Temperatura de Fusión** ( $T_m$ ) do ADN.

Os aneis das bases das cadeas de ADN absorben a 260 nm. Cando as cadeas se atopan separadas, a absorbencia é maior que con respecto á dobre hélice. O proceso de desnaturalización, pois, pode seguirse doadamente realizando espectros de absorbencia do ADN a 260 nm. A medida que aumenta a temperatura e se separan as febras de ADN prodúcese un incremento na intensidade do sinal de UV. O espectro obtido ten forma de curva sigmoide e a temperatura que se corresponda co punto medio da curva é a  $T_m$ .

Para iso, rexistráronse os cambios nos espectros de absorbencia a 260 nm a medida que incrementaba a temperatura tanto dos ADN diana sos como dos complexos ADN/oligómero.

Os valores da  $T_m$  foron determinados a partir das derivadas das curvas de desnaturalización. Para aquelas curvas anómalas, a  $T_m$  estimouse a partir do gráfico.

As mostras sen ligando, só das cadeas de ADN, presentaron curvas de desnaturalización monofásicas.

Ao engadir un equivalente de dímero ás diferentes secuencias de ADN, obsérvanse curvas bifásicas (ver **Figura 119**). Despois de analizar as dúas transicións, atopouse que a primeira das curvas presentaba unha  $T_m$  igual á  $T_m$  experimental do ADN mentres que a segunda curva posuía unha  $T_m$  totalmente diferente e nova. Este fenómeno, en consonancia cas valoracións fluorescentes, xustificaría unha estequiometría de unión de tipo 2:1 debido á forte unión que presenta o dímero **3aa** e á cooperación positiva das unidades de BBA. Así pois, para o ratio 1:1 analizado, quedaría ADN libre en disolución, que se atribúe á primeira das curvas.

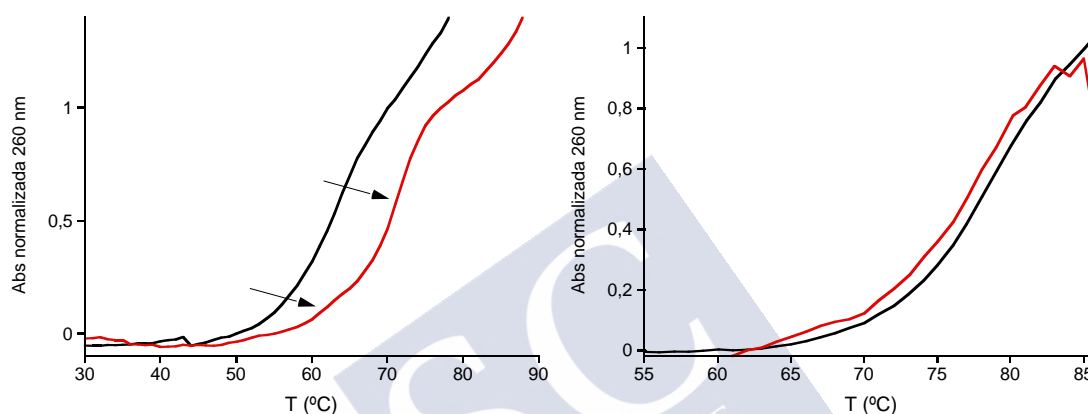


Figura 119. Espectros de uv de desnaturalización dunha disolución  $0.5 \mu\text{M}$  de ADN en tampón Tris-HCl 20 mM pH 7.5, 100 mM NaCl (liña punteada) e na presenza de 1 equivalente de dímero (liñas sólidas) Esquerda: ADN  $(A_3T_3GC)_2$ . Dereita ADN  $(GC)_2$ .

Como tendencia xeral, a temperatura de fusión do ADN increméntase nuns 10 graos cando o dímero se atopan unido ás secuencias ricas en A/T (ver **Táboa 8**). Pola contra, para cadeas ricas en GC, a  $T_m$  non varía. Entón, estes biomarcadores interactúan fortemente con secuencias específicas de ADN ricas en A/T mentres que non se unen a secuencias ricas en GC.

Táboa 8. Comparación das temperaturas de desnaturalización experimentais para o ADN e o complexo ADN-trímero obtidas das curvas de UV. Expresados en  $^{\circ}\text{C}$ . A  $\Delta T$  foi calculada como  $(T_m \text{ con dímero}) - (T_m \text{ sen dímero})$ .

ADN	$T_{\text{ADN}}$	$T_{\text{COMPLEJO}}$	$\Delta T$
$(A_3T_3GC)_2$	63	73	10
$(A_3T_3C)_2$	58	74	16
$(A_2T_3C)_2$	62	71	9
GC2	77	77	0

Por outra banda, para os complexos ADN/trímero con ratio 1:1 os espectros de *melting* deron curvas multifásicas (**Figura 120**). Do mesmo xeito que para o dímero, a primeira curva

correspóndese ca  $T_m$  do ADN libre en disolución e o resto das curvas con diferentes modos de unión do trímero ao ADN.

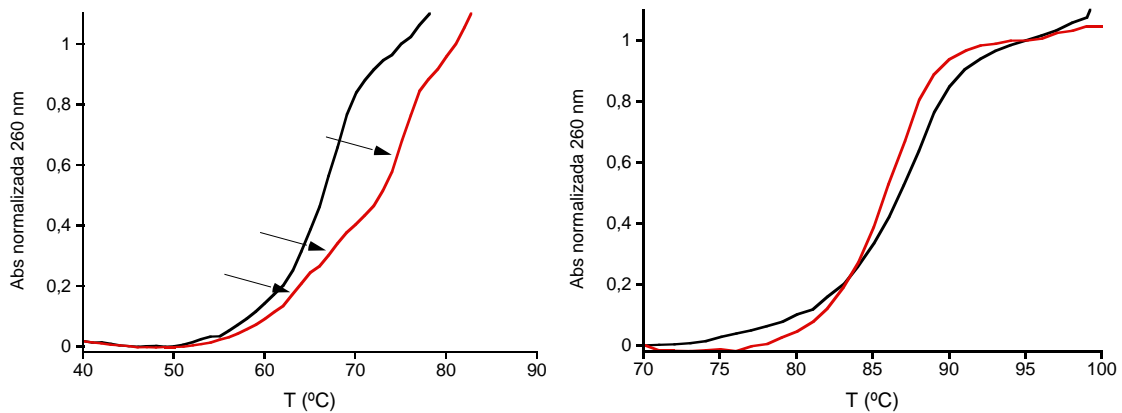


Figura 120. Superposición de espectros de uv de desnaturalización dunha disolución  $0.2 \mu\text{M}$  de ADN de dobre cadea en tampón Tris-HCl 20 mM pH 7.5, 100 mM NaCl (liña punteada) e na presenza de 1 equivalente de trímero (liñas sólidas) Esquerda: ADN  $(A_2T_3C)_3$ . Dereita ADN  $(GC)_3$ .

Os resultados obtidos dos experimentos de  $T_m$  ca estequiometría 1:1, reforzan a hipótese da formación pouco habitual de complexos 2:1 para dímero:ADN e modos de unión con estequiometrías máis complexas para o aducto trímero:ADN.

Táboa 9. Comparación das temperaturas de desnaturalización experimentais para o ADN e o complexo ADN-trímero obtidas das curvas de UV. Expresados en  $^{\circ}\text{C}$ . A  $\Delta T$  foi calculada como  $(T_m \text{ con trímero}) - (T_m \text{ sen trímero})$ .

ADN	$T_{\text{ADN}}$	$T_{\text{COMPLEJO}}$	$\Delta T$
$(A_3T_3GC)_3$	78	84	6
$(A_3T_3C)_3$	68	73	5
$(A_2T_3C)_3$	66	74	8
$GC_3$	88	87	-1

Con motivo de entender un pouco mellor a forma de interacción entre os oligómeros e os ADN, efectuáronse experimentos de desnaturalización dos complexos con estequiometría 1:1 a diferentes concentracións. Para concentracións baixas, obtivéronse curvas multifásicas. A medida que se incrementou a concentración de complexo dentro da cubeta, as fases tenderon a desaparecer e formarse unha única curva (ver **Figura 121** para exemplo). Tal comportamento podería ser debido ao recoñecemento de ADN por diferentes complexos. A medida que se



aumenta a concentración, o recoñecemento por parte dun dos complexos estaría favorecido por riba dos demais.

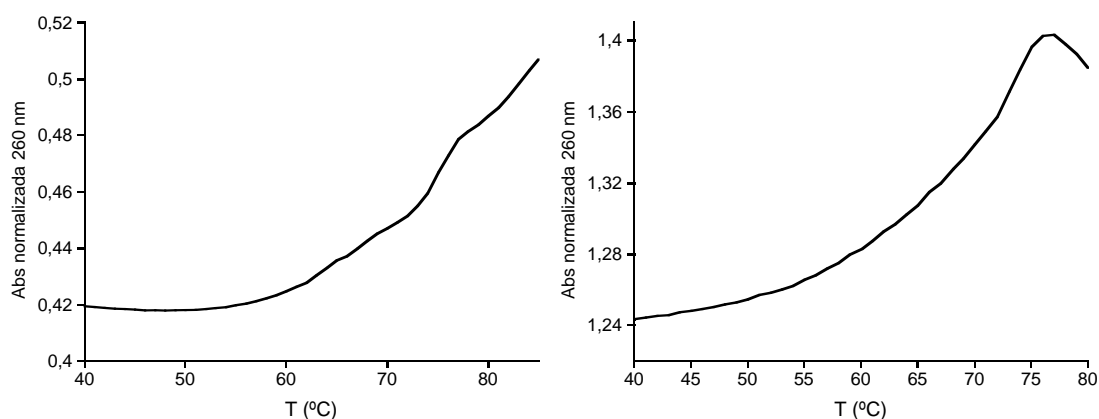


Figura 121. espectros de uv de desnaturalización dunha disolución 0.2  $\mu\text{M}$  (esquerda) e 1  $\mu\text{M}$  (dereita) de complexo trímero:ADN 1:1 en tampón Tris-HCl 20 mM pH 7.5, 100 mM NaCl. ADN  $(\text{A}_2\text{T}_3\text{C})_3$ .

Só se mostran as curvas de desnaturalización máis representativas, para maior información véxase a *Parte Experimental (EXPERIMENTAL SECTION)*.

Polo tanto, acorde a tódolos experimentos, propónse que os oligómeros obtidos recoñecen con moi alta afinidade ( $K_D$  no rango de nanomolar para dímero e picomolar para trímero) secuencias de ADN longas, ricas en diferentes combinacións de pares A/T sobre aquelas ricas en GC. O modo de unión é inserción no suco menor. En disolución, os oligómeros poden formar diferentes tipos de complexos co ADN. Segundo a concentración, o equilibrio desprázase cara un tipo determinado de complexo oligómero:ADN.

#### 5.4. CONCLUSIÓNS

- Diseñáronse e sintetizáronse oligómeros fluorescentes de bisbenzamidinas dunha forma moi sinxela e nun único paso.
- Comprobouse que recoñecen selectivamente secuencias de ADN ricas en pares de bases A/T sobre outras ricas en G/C.
- Os oligómeros insírense no suco menor do ADN.
- A modelo de unión axústase a unha estequiometría maior que 1:1 tanto para trímero como parao dímero.
- Fórmanse diferentes complexos oligómero:ADN segundo a concentración.
- Actúan como biomarcadores detectando secuencias longas de pares de bases A/T repetidas, coma os da ataxia de tipo 10.

- Ámbolos dous oligómeros illados incrementan a temperatura de desnaturalización do seus ADN diana, indicando unha forte interacción co ADN.
- Os oligómeros presentan moi altas afinidades polas secuencias diana: con constantes de disociación no rango de baixo nM para o dímero **3aa** e pM para o trímero **4aa**.







## **6. GENERAL CONCLUSIONS**



The main conclusions for each chapter of this thesis are:

**First Chapter: *Chemical synthesis of OMOMYC and MYC derivatives and its interaction with the major groove of the DNA.***

- Several analogues of the bHLHZip domain of MYC and OMOMYC were synthesized using two different chemical approaches based on *Native Chemical Ligation* (NCL).
- Those versatile chemical techniques allowed the introduction of designed mutations in the sequence of the peptides such as non natural residues, and fluorophores.
- Intracellular distribution of the fluorescently labelled derivatives was monitored demonstrating that they can act as cell penetrating peptides at nanomolar concentrations in various cell lines. Miniproteins containing free cysteines showed better internalization.
- Remarkably, those miniproteins maintained the DNA binding properties of the natural counterparts.

**Second Chapter: *Control of the recognition of the minor groove of the DNA with photo-switchable small molecules***

- Four different azoderivatives were rationally obtained. Two of them capable of photoisomerize under uv-vis irradiation.
- The *orto*-tetrafluorinated derivative is able to isomerize under visible light in both E/Z directions with nearly quantitative conversion.
- The Z form of the tetrafluoroazobenzenes showed a remarkably high half- life time.
- The azocompounds recognize specific AAATTT rich sequences of the DNA over those containing GC, by inserting in the minor groove. Particularly, the Z isomers displayed stronger affinities for those oligonucleotides than E isomers.

**Third Chapter: *Recognition of the minor groove of the DNA by fluorescent oligomers with very high affinities.***

- An easy one pot synthetic procedure to prepare bisbenzamidine oligomers was reported.
- The oligomers recognize long sequences of A/T base pair repeats by inserting in the minor groove of the DNA.
- The interaction process with the DNA showed synergistic effects between bisbenzamidine units and allowed an extremely efficient and selective DNA binding (low nM and pM dissociation constants).
- The resulting oligomers could be used as fluorescent biomarkers of spinocerebellar ataxia type 10 (SCA10) DNA repeats.





## **7. SUMMARY**

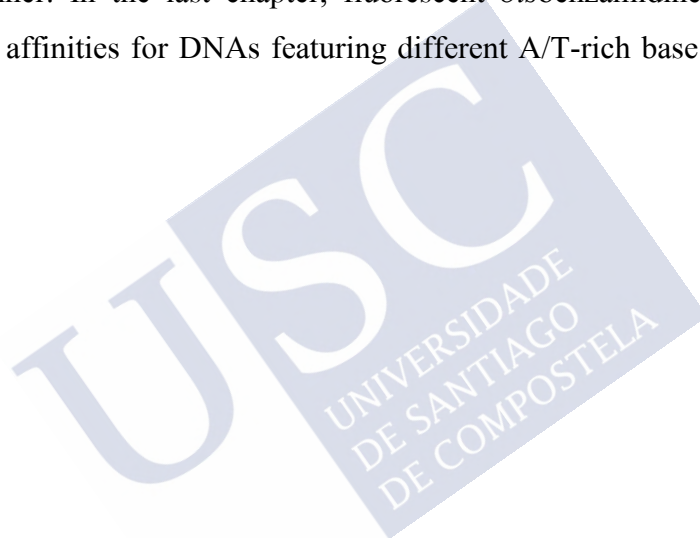




This thesis reports the design and synthesis of three different systems able to interact and recognize specific sequences of DNA.

In the first part, functional analogues of the oncoprotein MYC, of its negative dominant OMOMYC, as well as fluorescently labelled derivatives of these proteins were synthesized in excellent yields by way of two alternative strategies which involve either one or two *Native Chemical Ligation* reactions. These therapeutically relevant miniproteins are able to dimerize and recognize the major groove of the *E*-DNA and successfully internalize living cells.

The second part describes the synthesis of light-switchable systems capable of interacting specifically with the minor groove of A/T-rich sequences of DNA in a reversible and externally controllable manner. In the last chapter, fluorescent *bisbenzamidine* oligomers that show remarkably high affinities for DNAs featuring different A/T-rich base pair combinations are described.







## **8. EXPERIMENTAL SECTION**



### 8.1. REAGENTS AND MATERIALS:

Unless stated otherwise, reagents were purchased from commercial sources and used without further purification.

Standard Fmoc amino acids, Boc-Ala, *N*-[(1*H*-benzotriazol-1-yl)(dimethylamino)methylene]-*N*-methylmethanaminiumhexafluorophosphate (HBTU) and *N*-[(dimethylamino)-1*H*-1, 2, 3- tiazolo [4, 5- b]piridin-1-il-metilen]-*N*-methylmethanaminium hexafluorophosphate (HATU) from *GL Biochem (Shanghai)*, and Boc-L-thiazolidine-4-carboxylic acid from *Bachem*. TIS (triisopropilsilane), EDT (ethanedithiol), Tris(2-carboxyethyl) phosphine hydrochloride (TCEP•HCl), 4-mercaptophenilacetic acid (MPAA), *N,N*-dimethylformamide (DMF) and acetonitrile (CH<sub>3</sub>CN) were purchased from *Sigma-Aldrich*. *N,N*-Diisopropylethylamine (DIEA), dichloromethane and trifluoroacetic acid (TFA) were purchased from *Fischer Scientific UK*. 2-2'-Azobis[2-(2-imidazolin-2-yl) propane] dihydrochloride (VA-044) was purchased from *Wako Pure Chemical Industries*. All other reagents were purchased from *sigma Aldrich*.

For ligation experiments, Guanidine Hydrochloride Ultrapure (Gn•HCl, *MP Biomedicals*) was used.

Peptides were synthesized on Rink-amide-ChemMatrix resin (0.57 mmol/g loading) from *Biotage AB*, or on PAL-PEG-PS resin (0.19 mmol/g loading) from *Applied Biosystems*, depending on the case.

For Fmoc-Dbz-OH synthesis we used: 4-Fluoro-3-methylbenzoic acid, methylamine solution, palladium on carbon and 9-fluorenylmethyl *N*-succinimidyl carbonate (Fmoc-OSu). All of them were purchased from *Sigma Aldrich*.

Hairpin oligonucleotides were supplied by *Thermo Fischer* and their sequences are:

**AAATTT:** 5'-GGC AAATTT CAG TTTT CTG AAATTT GCC-3'

**AATTT:** 5'-GGCG AATTT CGC TTTT GCG AATT CGCC-3'

**AATT:** 5'-GGCG AATT CAGC TTTT GCTG AATT CGCC-3'

**AAGTT:** 5'-GGCG AAGTT CGC TTTT GCG AACTT CGCC-3'

**AGATT:** 5'-GGCG AGATT CGC TTTT GCG AATCT CGCC-3'

**AAGCTT:** 5'–GGC AAGCTT CGC TTTT GCG AAGCTT GCC–3'

**GGCCC:** 5'–GGCA GGCCCAGC TTTT GCTGGGCC TGCC–3'

**(A3T3)<sub>2</sub>:** 5'–GCGC AAATTT GC AAATTT GCGC–3'

**(A3T3)<sub>3</sub>:** 5'– GCGC AAATTT GC AAATTT GC AAATTT GCGC–3'

**(GC)<sub>2</sub>:** 5' –GCAC GGCCC AC GGCCC TCGC–3'

**(GC)<sub>3</sub>:** 5'–GCAC GGCCC AC GGCCC AC GGCCC TCGC–3'

Oligonucleotides were hybridized by mixing complementary sequences at equal molar concentration, heating at 90 °C for 10 min and then slowly cooling the mixture to rt.

## 8.2. NMR, HPLC, MASS SPECTROMETRY, UV AND CD MEASUREMENTS

### 8.2.1. Reverse phase high performance liquid chromatography (RP- HPLC)

**For the 1<sup>st</sup> chapter:** Reverse Phase High-Performance Liquid Chromatography (RP-HPLC) was performed using an *Agilent 1100 Series* Liquid Chromatograph system. Analytical HPLC was run using a *Jupiter C4 column* (5  $\mu$ m, 300 Å, 150 × 4.60 mm) at a flow rate of 1 mL/min with a gradient from 0 % to 70 % of B in 30 min (eluent A: 0.1 % TFA in H<sub>2</sub>O, eluent B: 0.1 % TFA in CH<sub>3</sub>CN) for the crudes, and a gradient from 5 % to 95 % of B in 15 min for the analysis of the pure peptides. Purification of the final peptides was performed using a semipreparative *Agilent Zorbax 300 SB-C3 column* (5  $\mu$ m, 300 Å, 250 × 9.4 mm) at a flow rate of 4 mL/min employing the following gradient: 0→5 min, 0 % B; 5→85 min, 0→70 % B. Eluent A: H<sub>2</sub>O, 0.1% TFA. Eluent B: CH<sub>3</sub>CN, 0.1 % TFA.

These standard conditions were used otherwise indicated. Compounds were monitored at 222, 270, 304 and 330 nm (and 555 nm for peptides containing TAMRA (5-Carboxytetramethylrhodamine) tags).

**For the 2<sup>nd</sup> chapter:** Reactions were monitored by TLC analysis, using 60 F<sub>254</sub> (*Merck*) silica gel and visualized under UV light (254 nm, 365 nm). Some reactions were also followed by analytical RP-HPLC with an *Agilent 1100 series* LC/MS using an *Eclipse XDBC18* (4.6 × 150 mm, 5  $\mu$ m) analytical column. Compounds were detected by UV absorption (220, 270, 304 and 330 nm). Standard conditions for analytical RP-HPLC consisted on a linear gradient from

5 % to 95% of solvent B for 30 min at a flow rate of 1 mL/min (A: H<sub>2</sub>O, 0.1 % TFA, B: CH<sub>3</sub>CN, 0.1% TFA). Purification of final products were performed by semipreparative RP-HPLC with an *Agilent 1100 series LC* equipped with a UV-visible detector using a *Luna 5u C18(2) 100Å* (5  $\mu$ m, 10  $\times$  250 mm) reverse-phase column from *Phenomenex*. Collected fractions with pure products were lyophilized with a *ThermoSavant Modulyo D* lyophilizator equipped with an *Edwards RV* high vacuum pump

### 8.2.2. Electrospray Ionization Mass Spectrometry (ESI/MS)

ESI-MS was performed with an *Agilent 1100 Series LC/MSD VL G1956A* model in positive scan mode, using direct injection of the purified peptide solution into the mass spectrometer. *PE Biosystems Voyager System 1089* or *Bruker Biotof II* were also used to obtain the MS characterization for small molecules from chapter two.

### 8.2.3. NMR

NMR spectra were recorded using a *Varian Mercury 300* (300 MHz for <sup>1</sup>H and 75 MHz for <sup>13</sup>C); a *Bruker DPX-250* (250 MHz for <sup>1</sup>H and 63 MHz for <sup>13</sup>C), a *Varian Inova 400* (400 MHz for <sup>1</sup>H and 101 MHz for <sup>13</sup>C) or a *Varian Inova 500* (500 MHz for <sup>1</sup>H and 127 MHz for <sup>13</sup>C) spectrophotometers and processed using the *MestReNova v6.1.1-6384* suite (Mestrelab Research). Solvents used were: CHCl<sub>3</sub>-*d*<sub>3</sub>, MeOH-*d*<sub>4</sub> or DMSO-*d*<sub>6</sub>, depending the compound. Chemical shifts units are  $\delta$  (ppm) with TMS as reference ( $\delta = 0.0$  ppm) and coupling constants (J) are shown in Hz

### 8.2.4. Matrix-assisted laser desorption/ionization mass spectrometry (MALDI/MS)

MALDI spectrometry was performed with a *Bruker Autoflex MALDI/TOF* model in positive scan mode by direct irradiation of the matrix-absorbed peptide

### 8.2.5. Circular Dichroisms (CD)

Circular dichroism measurements were performed on a *Jasco J-715* equipment coupled to a *Neslab RTE-111* thermostatic water bath, using a *Hellma* micro cuvette (100-QS, 2 mm light pass. The spectra are the average of 3 scans and they also were slightly smoothed for a better visualization using the macro "*smooth*" implemented in the *KaleidaGraph* program (Synergy Software version 4.1).



Measurements of 300  $\mu$ L of a 5  $\mu$ M solution of peptides from first chapter were run in 10 mM phosphate, 150 mM KCl, 1 mM DTT, and 0.25 mM EDTA buffer pH 7.4, at 20 ° C.

Measurements of 300  $\mu$ L of a 5  $\mu$ M solution of compounds from the second chapter were registered in Tris•HCl 20 mM, NaCl 100 mM buffer, pH 7.5, at 20 ° C.

Settings:

- Acquisition range: 380-195 nm.
- Band width: 2.0 nm.
- Resolution: 0.2 nm.
- Accumulation: 5-10 scans
- Sensitivity: 10 milidegrees.
- Response time: 0.25 seconds.
- Speed: 100 nm/min.

### 8.2.6. Quantification of proteins, small molecules and oligonucleotides using UV

UV measurements were performed on a *Varian Cary 100 Bio spectrophotometer* or a *JASCO UV-630* spectrophotometer coupled with a *PolyScience* thermostat, using a standard *Hellma* semi-micro cuvette (140.002-QS). The concentrations of the stock solutions of the compounds were obtained using the Lambert Beer's law.

DNA solutions were measured at 260 nm and the molar extinction coefficients (in  $M^{-1} cm^{-1}$ ) of single strand oligonucleotides were calculated with the following formula:<sup>154</sup>

$$\epsilon_{260nm} = \{(8.8 \times n^o T) + (7.3 \times n^o C) + (11.7 \times n^o G) + (15.4 \times n^o A)\} \times 0.8 \cdot 10^3$$

Where  $n^o$  represents the number of the nitrogenated bases in the sequence of the DNA,  $T$  the thymine type bases,  $C$  the cytosines,  $G$  the guanines and  $A$  adenines.

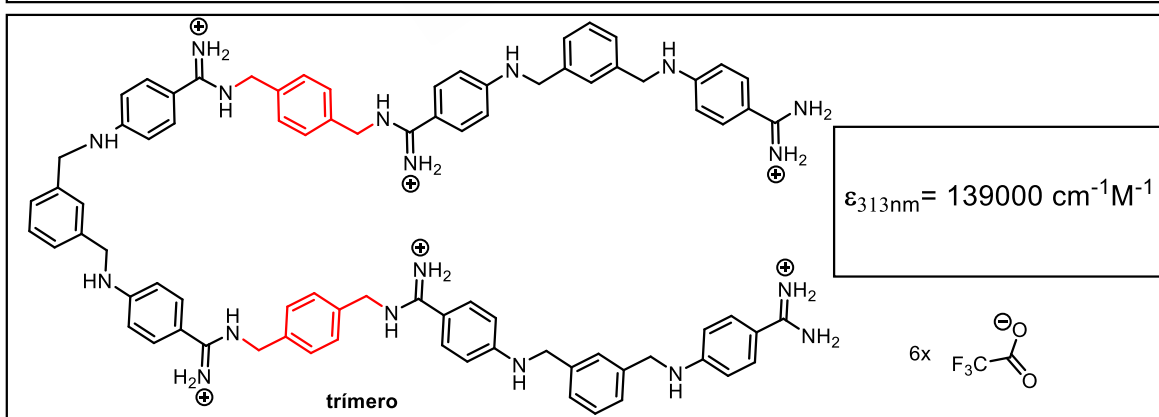
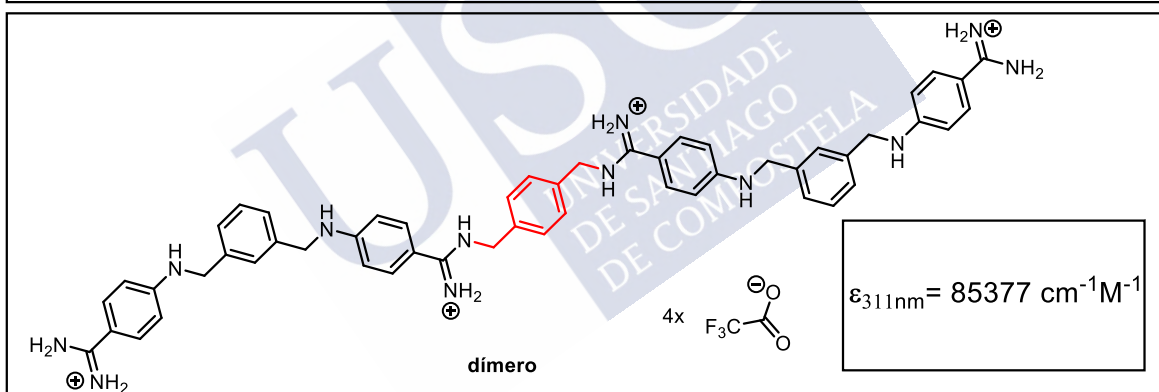
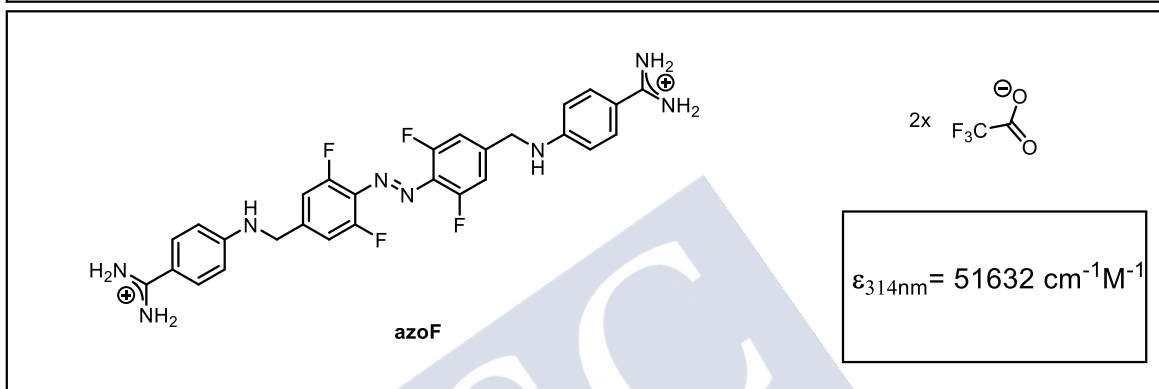
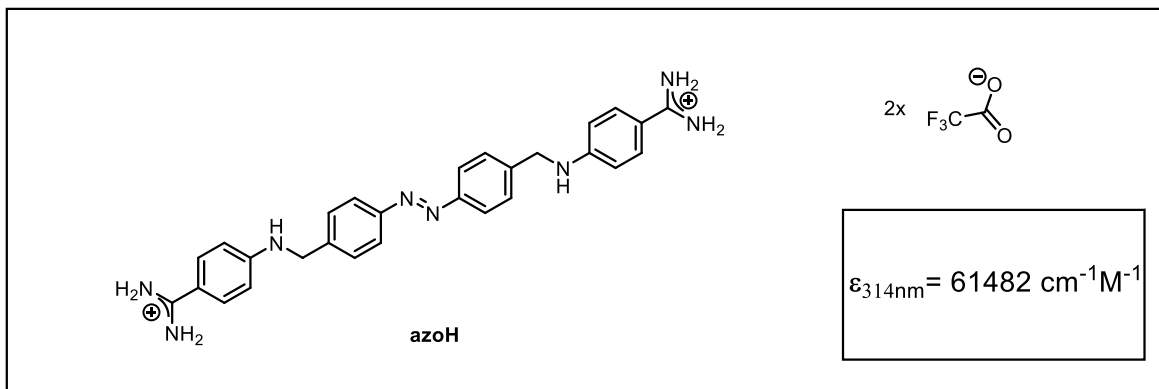
For peptides [AQ]MYC(SH), [AQ]MYC, Tyr was used as the chromophore with a molar extinction coefficient of  $1405 M^{-1}cm^{-1}$  at 274 nm

<sup>154</sup> a) K. C. Engman, P. Sandin, S. Osborne, T. Brown, M. Billeter, P. Lincoln, B. Norden, B. Albinsson and L. M. Wilhelmsson, *Nucl. Acids Res.* 2004, **17**, 5087. b) G. Kallansrud and B. Ward, *Anal Biochem.* 1996, **236**, 134.

For labelled peptides *TMR-[AQ]MYC*, *TMR-[AQ]MYC(SH)*, *TMR-[AQ]OMOMYC* and *TMR-[AQ]OMOMYC(SH)*, TAMRA was used as the chromophore with a molar extinction coefficient of  $80000 \text{ M}^{-1}\text{cm}^{-1}$  at 552 nm.

Molar extinction coefficients of other compounds are shown in the following picture. They were obtained applying the Lambert-Beer's equation.





### 8.2.7. Photoisomerization studies of azoderivatives

Solutions of the azo derivatives were irradiated at 365 nm for 20 min using a VL-4LC 4 W, 50 Hz lamp equipped with band-pass filters.

For visible irradiation, an Asahi Spectra Max 303, with a 300W short arc xenon lamp was used to irradiate samples of azo derivatives. band pass filters of:  $520 \pm 10$  nm and  $410 \pm 10$  nm, were used during 15 min for a monochromatic irradiation.

### 8.2.8. Fluorescence measurements

Measurements were made with a Jobin-Yvon *Fluoromax-3*, (DataMax 2.20) coupled to a *Wavelength Electronics* LFI-3751 temperature controller, using a *Hellma* micro cuvette of 1 mL.

All titrations were made following the same procedure: increasing amounts of the target dsDNA were successively added to a 1 mL solution of the selected compound in Tris-HCl buffer 20 mM, 100 mM NaCl, pH 7.5 and the fluorescence spectra was recorded after each addition.

For competition titrations: aliquots of the azo compound were successively added to a solution (1 mL) of BBA 1  $\mu$ M in Tris•HCl buffer 20 mM; 100 mM NaCl; pH 7.5, in the presence of 0.5  $\mu$ M target oligonucleotide. Emission of the DNA/BBA complex was recorded using an excitation wavelength of 329 nm.

Settings:

- Increment: 1 nm
- Integration time: 0.2 s
- Excitation slit width: 3.0 nm
- Emission slit width: 6.0 nm
- excitation wavelength: 329 nm
- Acquisition range from 345 to 500 nm
- T: 20 °C.

### 8.2.9. $K_D$ calculations for azoderivatives

UV-vis titrations were recorded on a *JASCO UV-630* spectrophotometer coupled with a *PolyScience* thermostat, using a standard *Hellma* semi-micro cuvette (140.002-QS).

All titrations were made following the same procedure: keeping the concentration of **azoH** or **azoF** constant, increasing amounts of the target dsDNA were successively added to a 700  $\mu$ L solution of the corresponding azo derivative in Tris-HCl buffer 20 mM, 100 mM NaCl, pH 7.5. After each addition, the decrease of the absorbance was monitored upon increments in DNA concentration.

Titrations were performed at 0.5  $\mu$ M solutions of azo derivatives in *trans* form and 1  $\mu$ M solutions of azos in *cis* form as their absorption bands and affinities are weaker.

The  $K_D$  values for azo derivatives in both its conformations were obtained from the slope of the Benesi-Hildebran equation:<sup>155</sup>

$$\frac{1}{\Delta A} = \frac{1}{\Delta A_{\max}} + \frac{1}{K_a [\Delta A_{\max}]} \times \frac{1}{[DNA]}$$

where  $\Delta A$  is the difference between the absorbance of azo in the absence and presence of DNA,  $\Delta A_{\max}$  is the final absorbance of the azo-DNA adduct. The *trans:cis* ratio was taken into account.

<sup>155</sup> A) Hu, Y.-J., Liu, Y. Xiao, X.-H. *Biomacromolecules* **2009**, *10*, 517. B) H. A Benesi, J. H Hildebrand, *J. Am. Chem. Soc.* **1949**, *71*, 2703.

### 8.2.10. $K_D$ calculations for oligomers

UV meltings of the DNA-oligomer adducts were performed on a *JASCO UV-630* spectrophotometer coupled with a *PolyScience* thermostat, using a standard *Hellma* semi-micro cuvette (140.002-QS).

Measurements of 600  $\mu$ L of 0.6  $\mu$ M, 1  $\mu$ M and 0.2  $\mu$ M solutions of DNA:oligomer (1:1) complexes were registered in Tris·HCl 20 mM, NaCl 100 mM, pH= 7,5 buffer.

Settings:

- Acquisition range: 20-100 °C.
- $\lambda$  registered: 260 nm.
- equilibria time: 30 s.
- Speed: 0.5 °C/min.

$$I = \frac{[DNA]^n}{K_D + [DNA]^n}$$

Where  $I$  is the intensity of the signal and  $n$  the Hill coefficient

### 8.3. CELL UPTAKE EXPERIMENTS

The experiments were carried out at room temperature on a sterile *Teslar AV-100 bench*. **HeLa** and **A459** cells were cultured in DMEM (Dulbecco Modified Eagle's Medium) containing 10 % of FBS (Fetal Bovine Serum) with 5 mM of L-glutamine and 1 % of Penicillin-Streptomycin-Glutamine Mix. The day before the cellular uptake experiments, cells were seeded in 24 well plates containing glass bottom dishes (15 mm).

Stock solutions of the peptides were dissolved in DMEM without FBS.

For the cell uptake experiments, cells were first washed 3 times with PBS and overlaid with 500  $\mu$ L of fresh DMEM without serum. Samples with the indicated concentration of the peptide were added, and the mixtures were incubated in an incubator at 37 °C; with 5 % CO<sub>2</sub> atmosphere. Then the medium was removed and the cells were washed with PBS (2  $\times$  1 mL) and fresh DMEM (3  $\times$  1 mL).

Digital pictures were captured with a *Nikon Ti-E microscope* equipped with an *Andor Zyla 4.2 digital camera*. Confocal pictures were taken with an *Andor Dragonfly High Speed Confocal Platform* under identical conditions of gain and exposure for all the samples. Images were further processed (cropping, resizing and global contrast and brightness adjustment) using *Image J* software.

The parameters for the fluorescent channels at the Nikon (Semrock) equipment were:

filter cube DAPI-1160B-000: BP 387/11 nm, LP 447/60 nm and DM 409 nm.

filter cube FITC-3540C-000: BP 482/35 nm, LP 536/40 nm and DM 506 nm.

filter cube TRITCB-000: BP 543/22 nm, LP 593/40 nm and DM 562 nm.

The parameters for the confocal equipment were:

DAPI-LP 450/50 nm and DM 418 nm.  $\lambda_{exc}$ = 405 nm

Cy3-LP 620/60 nm and DM 567 nm.  $\lambda_{exc}$ = 561 nm.

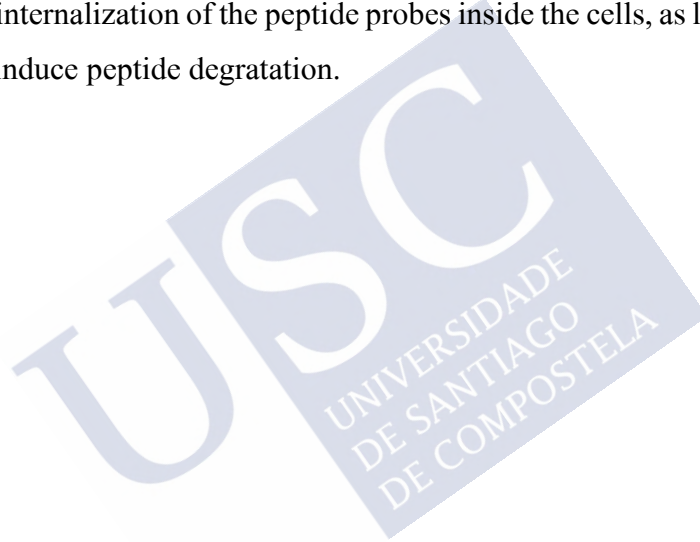
When necessary, co-staining experiments were performed to clarify the intracellular distribution of the peptides. **HOECHST 33342** was used as a *bona fide* DNA fluorescent probe.

### 8.3.1. Incubation concentration studies

The appropriate concentration for incubation of the *TMR*-[AQ]MYC and *TMR*-[AQ]MYC(SH) peptides was studied in HeLa cells for 30 min. Both tagged proteins internalize at nanomolar concentrations. For 50 nM the exposure time and laser intensity had to be prolonged to clearly detect the proteins. At micromolar range the peptides aggregate compromising the cell integrity. So, the best concentration was found to be 250 nM.

### 8.3.2. Incubation time studies

Incubation time was optimized in a series experiments using the optimized 250 nM concentration. While 2 h incubation was not enough to observe a clear internalization, 4 h was enough for clear internalization of the peptide probes inside the cells, as longer incubation times (24 h) appear to induce peptide degradation.





## 8.4. REACTIONS OF CHAPTER 1

### 8.4.1. Protocol for the synthesis of peptides

#### 8.4.1.1. General protocol for solid phase peptide synthesis (SPPS)

Peptide fragments were synthesized by SPPS using an **automated microwave Liberty Blue Peptide Synthesizer** from *CEM Corporation* following the manufacturer recommended procedures. DIC (*N, N'*-Diisopropylcarbodiimide), was used as activator and Oxyma (Ethyl(hydroxyimino)cyanoacetate) as base. Peptides were synthesized at 0.1 mmol scale using the Fmoc strategy and *Rink Amide-ChemMatrix* (loading 0.5 mmol/g) was used as solid support. The deprotection of the temporal Fmoc protecting group was performed by treating the resin with 20 % piperidine in DMF. For sequences where aspartimide formation was a concern, 0.1 M HOBt was added to the deprotection cocktail to minimize the side reaction.

#### 8.4.1.2. Cleavage/ deprotection step

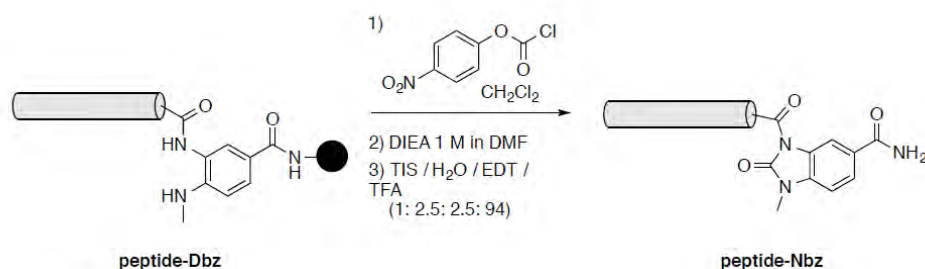
The final peptides were cleaved from the resin, and side-chain protecting groups were simultaneously removed using the following acidic cleavage cocktail: 94 % TFA, 2.5 % EDT, 2.5 % TIS and 1.0 % H<sub>2</sub>O (1 mL/40 mg peptidyl-resin) for 3 h. The cleavage solution containing the peptide was concentrated under vacuum to a minimal volume, added over cold ether for precipitation and separated by centrifugation. The supernatant was removed and the residue was dissolved in 30% CH<sub>3</sub>CN: H<sub>2</sub>O (0.1% TFA) and freeze-dried. The resulting solid was purified by HPLC using a semipreparative column to obtain the desired peptide as a solid. The yield for the purified peptides was calculated based on the original resin loading and the amount of peptidyl resin cleaved.

#### 8.4.1.3. Synthesis of peptides with Nbz at C-terminus

First, the Dbz moiety was coupled manually to the resin as the first *C*-terminal amino acid. For this, HATU (10 mmol) and Fmoc-Dbz-OH (1 mmol) were dissolved in DMF (0.2 M final concentration) and DIEA was added to the mixture (1.1 mmol). The amino acid was activated for 2 min before being added onto the resin. The resin was stirred using a nitrogen stream during the couplings (60 min).

Then, the next amino acids of the peptide were sequentially coupled using the general protocol for solid phase peptide synthesis described above. The last residue of this peptide

fragment was introduced as *N* $\alpha$ -Boc protected residue. The transformation of Dbz into Nbz was carried out following previously reported conditions, as shown in the next scheme.<sup>156</sup>



The C-terminus Dbz residue was activated through acylation with 4-nitrophenylchloroformate, followed by the addition of base to promote intramolecular attack of the anilide and cyclization to yield the resin-bound benzimidazolone. For this purpose, ~0.1 mmol of each peptide resin was swelled in CH<sub>2</sub>Cl<sub>2</sub> for 2 h. Then, the resin was filtered and a solution of 4-nitrophenylchloroformate (201 mg, 1 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (2.5 mL) was added onto it. The mixture was shaken for 2 h, and the resin was filtered and washed with CH<sub>2</sub>Cl<sub>2</sub> (2 × 5 mL × 3 min) and DMF (2 × 5 mL × 3 min). The filtered resin was swelled in DMF for 25 min, followed by filtration and addition of a solution of DIEA 1 M in DMF. N<sub>2</sub> (g) was passed through the resin suspension for 2 h. The resin was finally filtered and washed with DMF (2 × 5 mL × 3 min) and CH<sub>2</sub>Cl<sub>2</sub> (2 × 5 mL × 3 min).

The final Nbz-peptides were cleaved from the resin and the side-chain protecting groups were simultaneously removed using the standard TFA cleavage cocktail for 3 h. The TFA solution containing the peptide was concentrated under vacuum to a minimal volume, added onto cold ether and the resulting precipitate was centrifuged. The supernatant was removed and the residue was dissolved in CH<sub>3</sub>CN: H<sub>2</sub>O (3:7) and freeze-dried. The desired products were obtained after purification.

#### 8.4.1.4. Coupling of TAMRA and Fluorescein at the N-terminus

After SPPS synthesis, an extra Fmoc-Ava-Gly-Gly (Ava = aminovaleric acid) chain was added as a linker to the last *N*-terminal deprotected amino group of the corresponding fragment to avoid any side problem in the basic region of the peptides.

<sup>156</sup> Blanco-Canosa, B.J.; Dawson, E.P.; *Angew. Chem. Int. Ed.* **2008**, *47*, 6851.

Then, fluorophores (1.5 eq) was dissolved in DMF (20 mM). HATU (1.5 eq) and *N,N*-diisopropylethylamine (4 eq DIEA) were added to this solution. After two min, the resulting mixture was added over the Fmoc-deprotected peptide attached to the resin (approximately 0.025 mmol, 1 eq). N<sub>2</sub> was passed through the resin suspension for 4 h at rt. After that, the resin was filtered, washed with DMF (3 × 3 mL × 3 min) and with CH<sub>2</sub>Cl<sub>2</sub> (2 × 5 mL × 3 min). The reaction product was analyzed by cleavage of a small resin aliquot (4 mg) with 100  $\mu$ L of the cleavage reagent: TIS (25  $\mu$ L), H<sub>2</sub>O (25  $\mu$ L), CH<sub>2</sub>Cl<sub>2</sub> (50  $\mu$ L) and TFA (900  $\mu$ L) for 2 h, and its subsequent HPLC-MS.

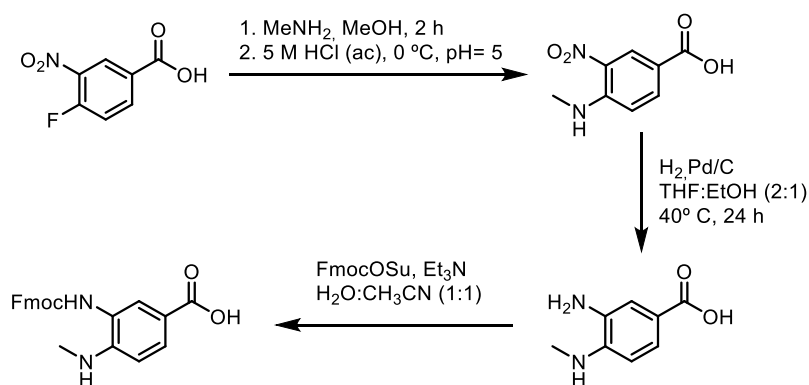
#### 8.4.1.5. Selective alloc deprotection.

Once SPPS synthesis of the peptides was completed, the selective deprotection of Lys(alloc) residues was carried out. PPh<sub>3</sub> (1.5 eq), *N*-methylmorpholine (10 eq) and phenylsilane (10 eq) in CH<sub>2</sub>Cl<sub>2</sub> (20mM) were added to 1 eq of the peptide attached to the resin and stirred at rt for 20 min. Then, Pd(OAc)<sub>2</sub> (0.3 eq) was added and the resulting mixture was stirred overnight. The resin was filtered and washed with DMF (2 × 1.5 mL, 2 min), diethyldithiocarbamate (DEDTC, 25 mg in 5 mL of DMF, 2 × 1.5 mL, 5 min), DMF (2 × 1.5 mL, 2 min) and CH<sub>2</sub>Cl<sub>2</sub> (2.5 mL, 2 min). Finally, the resin was dried and resulting product was checked by cleavage of a small aliquot of the resin and subsequent HPLC analysis.

#### 8.4.1.6. Coupling of fluorescein at C-terminus

After the orthogonal deprotection of amino group from Lys(alloc), the fluorophore was coupled following standard procedure described in *Coupling of TAMRA and Fluorescein at the N-terminus*.

#### 8.4.2. Synthesis of 3-Fmoc-4-diaminobenzoic acid (Fmoc-Dbz-OH)



Scheme S1. Synthetic scheme for Fmoc-Dbz-OH preparation.

Synthesis of intermediate **Fmoc-Dbz-OH** was performed as previously described.<sup>157</sup>

Final protection of the amino group was carried out using Fmoc-OSu: 1.7 g (10.5 mmol) of intermediate **3** and 3.9 g (11.57 mmol) of 9-Fluorenylmethyl *N*-succinimidyl carbonate (Fmoc-OSu) were dissolved in 100 mL (10.52 mmol) of CH<sub>3</sub>CN: H<sub>2</sub>O (1:1). After Et<sub>3</sub>N addition (1.5 mL) the solution changed to brown color. The reaction was stirred for 24 h, and then acidified with 1 M HCl (aq) to get a precipitate that was filtered and washed with CH<sub>3</sub>CN. The resulting solid was dried under vacuum and freeze-dried, affording the desired product as a white solid (3.2 g, 79 %).

<sup>1</sup>H NMR (300 MHz, *d*<sub>6</sub>-DMSO, 25°C, TMS): δ=4.25-4.45 (m, 3H; CH, CH<sub>2</sub>), 5.85 (s, 1H; NH), 6.60 (d, 1H; CH), 7.25-7.50 (m, 4H; CH), 7.65-7.80 (m, 4H; CH), 7.92 (d, 2H; CH), 8.70 (s, 1H; NH), 12.2 (s, 1H; OH).

ESI mass calculated for C<sub>23</sub>H<sub>20</sub>N<sub>2</sub>O<sub>4</sub> [M+1H]<sup>+</sup>: *m/z*=388.42, found 389.10

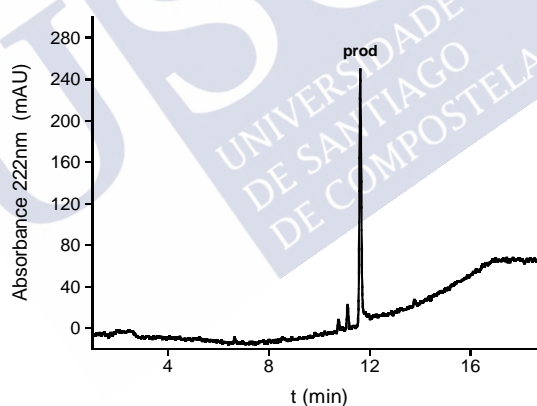


Figure S1. HPLC chromatogram of the pure Fmoc-Dbz-OH product at 222 nm

### 8.4.3. CHARACTERIZATION OF THE FRAGMENTS

**C**<sup>537</sup>**L**<sup>581</sup>-NH<sub>2</sub>

**C**PKVVILKKATAYILSVQAEEQKLISEEDLLRKRREQLKHKLEQL-NH<sub>2</sub>

<sup>157</sup> Kuzniweski, N. C; Gertsch, J.; Wartmann, M.; Altmann, H.K.; *Organic Letters* **2008**, *10*, 1183.

This product was obtained using the *general protocol for solid phase peptide synthesis* previously described. After cleavage from resin and purification by RP-HPLC, the peptide was isolated as a white powder (27 %).

**ESI-MS** (m/z): calculated for  $C_{237}H_{411}N_{68}O_{67}S= 5316.36$ ; found 1773.0  $[M+3H]^{3+}$ , 1329.9  $[M+4H]^{4+}$ , 1064.2  $[M+5H]^{5+}$ , 887.1  $[M+6H]^{6+}$ , 760.4  $[M+7H]^{7+}$ , 665.5  $[M+8H]^{8+}$ .

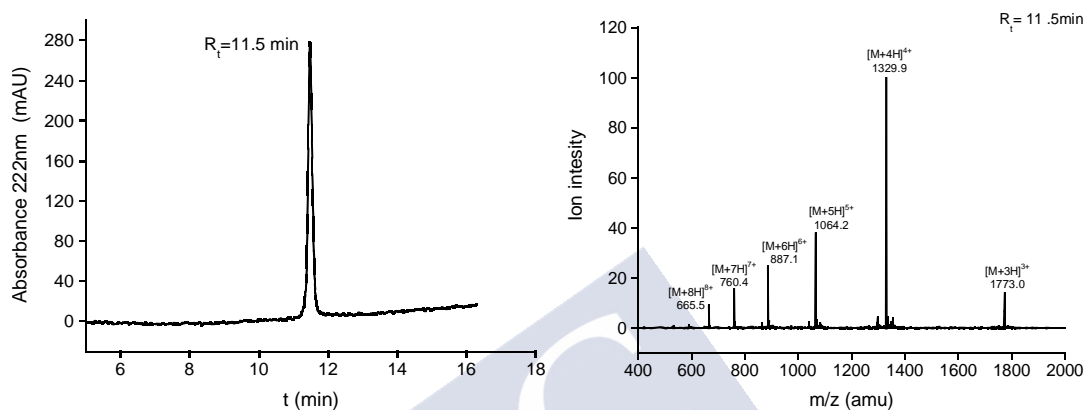


Figure S2. Left: HPLC chromatogram at 222 nm of the pure peptide  $C^{537}$ - $L^{581}$ - $NH_2$ . Right: ESI mass spectrum of the product.

**A<sup>500</sup>K<sup>536</sup>-Nbz**

**AVKRRTHNVLERQRRQELKRSFFALRDQIPELENNEK-Nbz**

**ESI-MS** (m/z): calculated for  $C_{207}H_{339}N_{72}O_{59}= 4779.44$ ; found 1593.8  $[M+3H]^{3+}$ , 1195.8  $[M+4H]^{4+}$ , 956.7  $[M+5H]^{5+}$ , 797.4  $[M+6H]^{6+}$ , 683.8  $[M+7H]^{7+}$ , 598.5  $[M+8H]^{8+}$

This fragment was obtained following the general protocol for *synthesis of peptides with Nbz at C-terminus* previously described.

Before the transformation into Nbz-peptide was performed, a sample with the Dbz moiety was cleaved from the resin and analyzed. The crude of the cleavage and MS containing the fragment **A<sup>500</sup>K<sup>536</sup>-Dbz** is shown in the next figure showing the desired product as a major peak. Finally, after the transformation into Nbz, cleavage from resin and purification by RP-HPLC, **A<sup>500</sup>K<sup>536</sup>-Nbz** was obtained as a white solid (28 %)

**A<sup>500</sup>K<sup>536</sup>-Dbz:**

**AVKRRTHNVLERQRRQELKRSFFALRDQIPELENNEK-Dbz**

ESI-MS ( $m/z$ ): calculated for  $C_{206}H_{339}N_{72}O_{58}$  = 4752.42; found 1585.8  $[M+3H]^{3+}$ , 1189.3  $[M+4H]^{4+}$ , 951.5  $[M+5H]^{5+}$ , 792.9  $[M+6H]^{6+}$ , 475.3  $[M+10H]^{10+}$ .

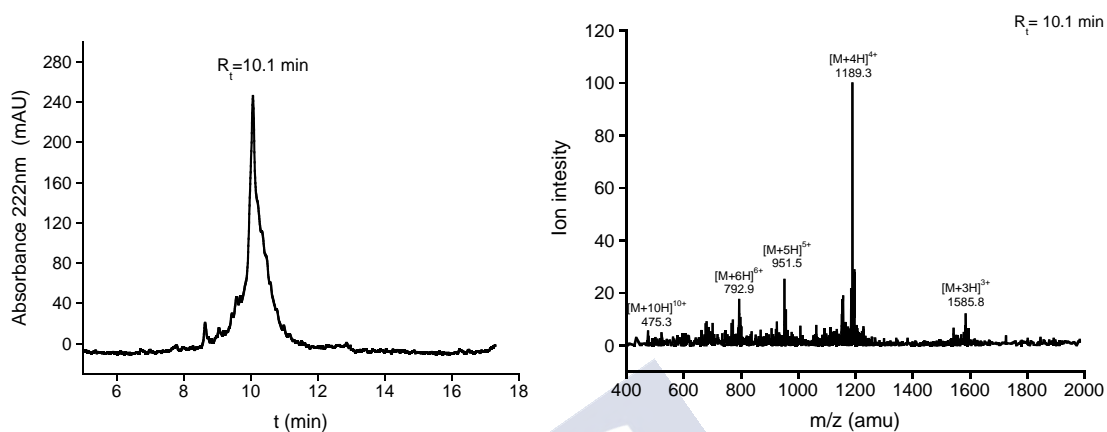


Figure S3. Left: HPLC chromatogram at 222 nm of the crude peptide with Dbz at C-terminus after cleavage from the resin. Right: ESI mass spectrum of the product.

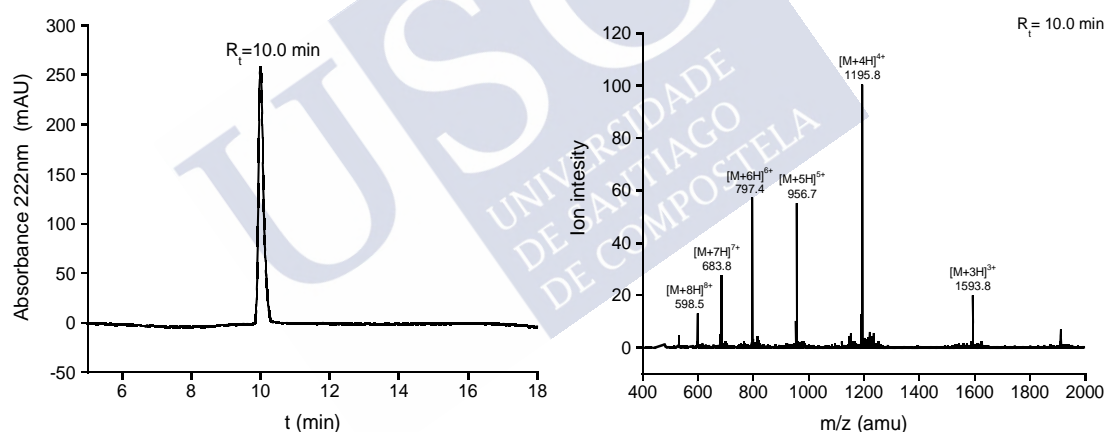


Figure S4. Left: HPLC chromatogram at 222 nm of the pure peptide with Nbz after purification. Right: ESI mass spectrum of the product.

**C<sup>548</sup>Leu<sup>581</sup>**

**CYILSVQAEEQKLISEEDLLRKRREQLKHKLEQL-NH<sub>2</sub>**

This fragment was obtained following the *general protocol for solid phase peptide synthesis* previously described. The cleavage of 0.02 mmol of resin afforded the product **C<sup>548</sup>L<sup>581</sup>** (37 mg, 45 %) as a white solid that was used without further purification.

ESI-MS (m/z): [M+H]<sup>+</sup> calculated for C<sub>182</sub>H<sub>310</sub>N<sub>54</sub>O<sub>55</sub>S= 4166.86, found 1464.2 [M+3H]<sup>3+</sup>, 1098.2 [M+4H]<sup>4+</sup>, 878.3 [M+5H]<sup>5+</sup>

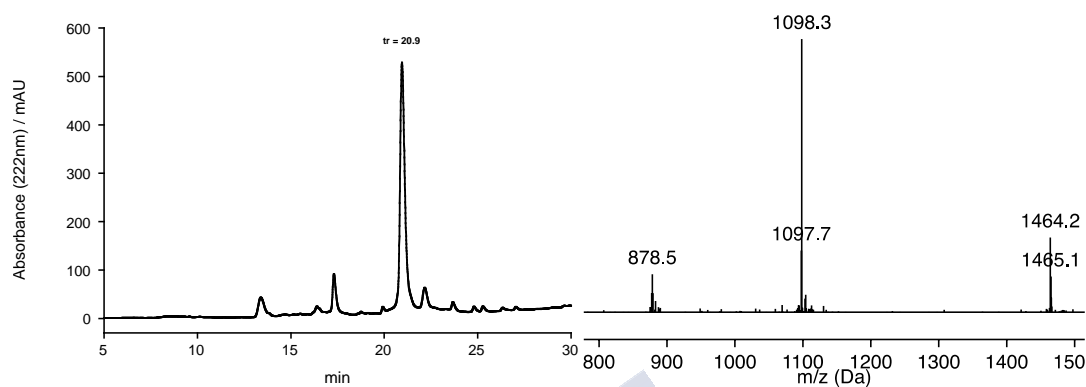


Figure S5. Left: HPLC chromatogram of the crude peptide C<sup>548</sup>-L<sup>581</sup> after cleavage from the resin at 222 nm. Right: ESI mass spectrum of the product

### Thz<sup>523</sup>T<sup>547</sup>-Nbz

### ThzPLRDQIPELENNEKAPKVVILKKAT-Nbz

Thz<sup>523</sup>T<sup>547</sup>-Nbz was synthesized following the previously described *general protocol for synthesis of peptides with Nbz at C-terminus*. After cleavage of 0.05 mmol from resin the resulting white solid was used without further purification (45 mg, 29 %).

ESI-MS (m/z): [MH]<sup>+</sup> calculated for C<sub>135</sub>H<sub>224</sub>N<sub>38</sub>O<sub>39</sub>S= 3035.56, found 1012. [M+3H]<sup>3+</sup>, 783.0 [M+4H]<sup>4+</sup>, 626.5 [M+5H]<sup>5+</sup>

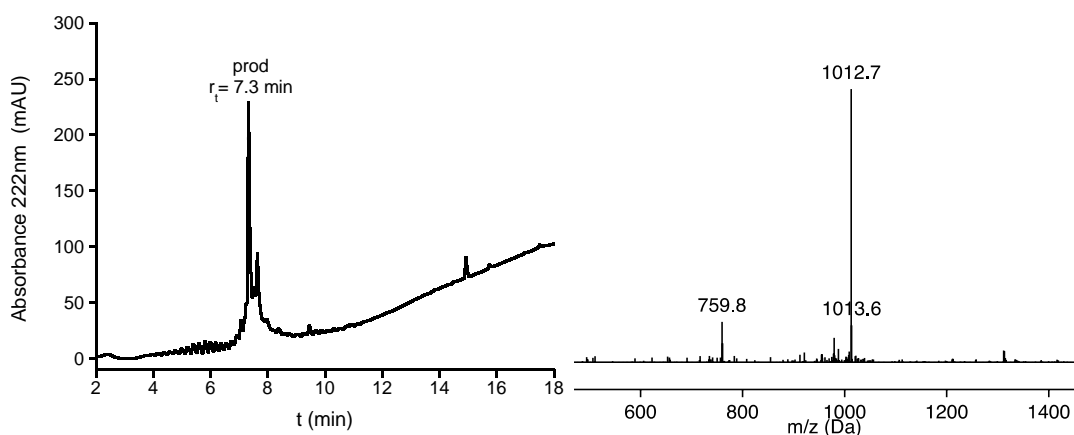




Figure S6. Left: HPLC chromatogram at 222 nm of the crude peptide Thz<sup>523</sup>T<sup>547</sup>-Nbz after cleavage from the resin. Right: ESI mass spectrum of the product.

**C<sup>548</sup>A<sup>586</sup>**

**CYILSVQAETQKLISEIDLLRKQNEQLKHKLEQLRNSA-NH<sub>2</sub>**

This fragment was obtained following the *general protocol for solid phase peptide synthesis* previously described. The cleavage of 0.06 mmol from resin afforded the product **C<sup>548</sup>-A<sup>586</sup>** (33 mg, 12 %) as a white solid that was used without further purification.

ESI-MS (m/z): calculated for C<sub>198</sub>H<sub>337</sub>N<sub>59</sub>O<sub>62</sub>S= 4568.28; found 1523.8 [M+3H]<sup>3+</sup>, 1142.8 [M+4H]<sup>4+</sup>, 914.5 [M+5H]<sup>5+</sup>.

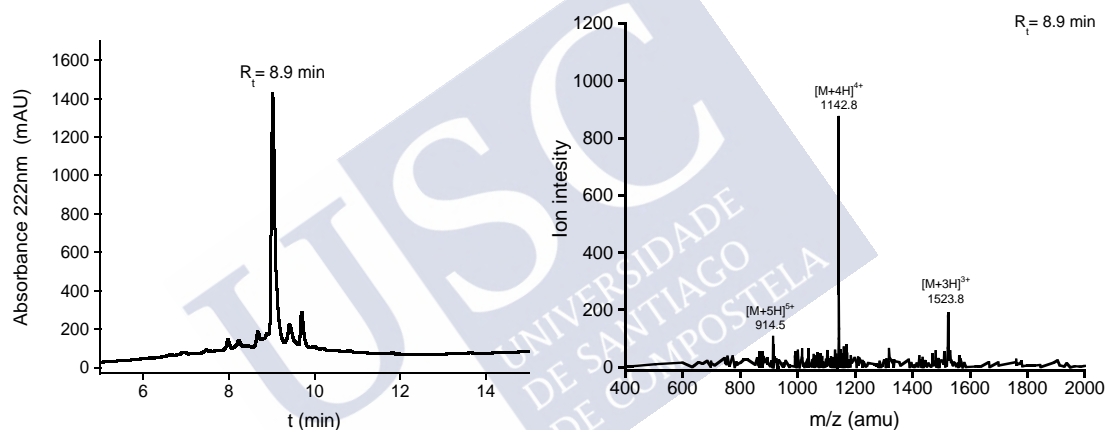


Figure S7. Left: HPLC chromatogram of the crude peptide C<sup>548</sup>A<sup>586</sup>-NH<sub>2</sub> after cleavage from the resin at 222 nm. Right: ESI mass spectrum of the product.

**TMR-AvaGG-A<sup>500</sup>F<sup>522</sup>-Nbz**

**TMR-AvaGGAKRRTHNVLERQRRQELKRSFF-Nbz**

ESI-MS (m/z): [MH]<sup>+</sup> calculated for C<sub>171</sub>H<sub>258</sub>N<sub>56</sub>O<sub>401</sub>= 3754.30, found 1252.3 [M+3H]<sup>3+</sup>, 939.5 [M+4H]<sup>4+</sup>, 751.8 [M+5H]<sup>5+</sup>, 626.8 [M+6H]<sup>6+</sup>, 537.2 [M+7H]<sup>7+</sup>

This fragment was obtained according to the previously described *general protocol for synthesis of peptides with Nbz at C-terminus* following the procedure of *Introducing of TAMRA at the N-terminus*



Before the transformation into *Nbz*-peptide was performed, a sample with the *Dbz* moiety was cleaved from the resin and analyzed. The crude of the cleavage and MS containing the peptide **AvaGG-A<sup>500</sup>F<sup>522</sup>-Dbz** is shown in the next figure. After TMR coupling and *Nbz* formation, the resulting peptide was cleavage from the resin. Purification by RP-HPLC afforded **TMR-AvaGG-A<sup>500</sup>F<sup>522</sup>-Nbz** as a pink solid (16 %)

### AvaGG-A<sup>500</sup>F<sup>522</sup>-Dbz:

### AvaGGAKRRTHNVLERQRRQELKRSFF-*Dbz*

ESI-MS (m/z): [MH]<sup>+</sup> calculated for C<sub>145</sub>H<sub>240</sub>N<sub>54</sub>O<sub>36</sub>= 3315.86, found 1106.3 [M+3H]<sup>3+</sup>, 830.0 [M+4H]<sup>4+</sup>, 664.2 [M+5H]<sup>5+</sup>, 553.6 [M+6H]<sup>6+</sup>

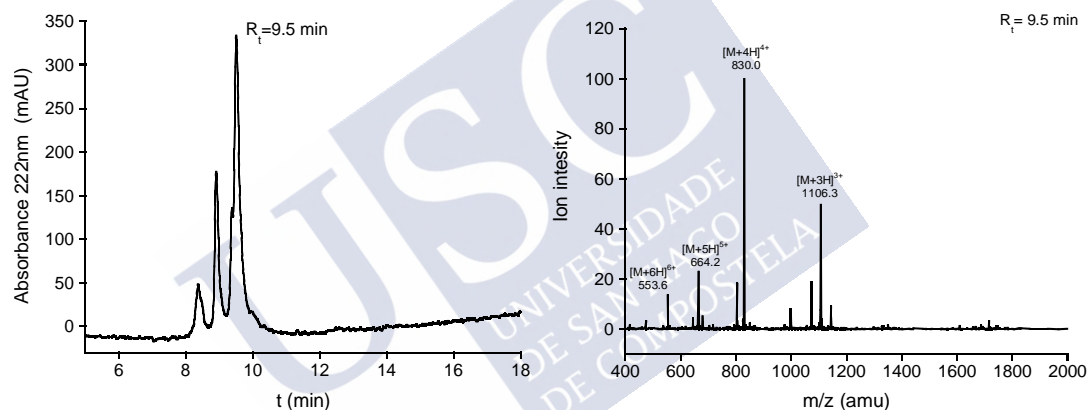


Figure S8. Left: HPLC chromatogram of the crude peptide AvaGG-A<sup>500</sup>F<sup>522</sup>-Dbz after cleavage from the resin at 222 nm. Right: ESI mass spectrum of the product.

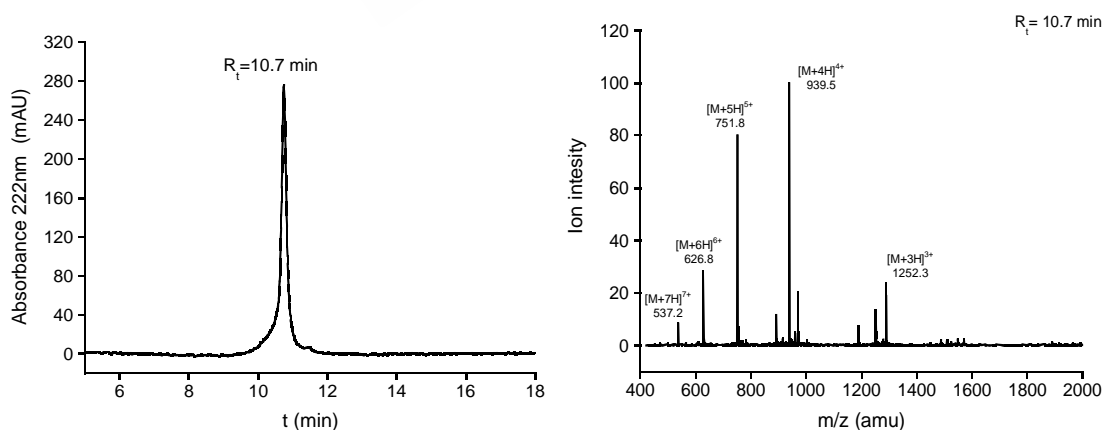


Figure S9. Left: HPLC chromatogram of the pure peptide TAMRA-AvaGG-A<sup>500</sup>F<sup>522</sup>-Nbz after cleavage from the resin and purification by RP-HPLC, at 222 nm. Right: ESI mass spectrum of the product.

#### 8.4.4. General Protocol for Native Chemical Ligations

The first *Native Chemical Ligations* were performed at room temperature dissolving the Nbz-peptide to a concentration of  $\sim 2.5$  mM and adding 1.5 eq of the cysteinyl peptide in 6 M Gn•HCl, 200 mM phosphate buffer (pH 7.0). Finally, 4-mercaptophenylacetic acid (MPAA) and tris(2-carboxyethyl) phosphine hydrochloride (TCEP•HCl) were added to the resulting mixture to final concentrations of 100 mM and 40 mM, respectively, and the pH was adjusted to 7.5 with NaOH (aq) and the buffer solution was degassed under N<sub>2</sub> for 15 min. For the second *Native Chemical Ligations*, the cysteinyl-peptide obtained from the first NCL was dissolved to a 2.5 mM solution, and followed by reaction using the same procedure as described above.

The progress of the reaction was followed by HPLC-MS. Once completed, it was quenched by diluting the crude up to 3 times using the buffer solution, and then frozen using liquid nitrogen. Finally, the product was purified by HPLC. When needed, the crude was filtrated using *Amicon Ultra-0.5 mL centrifugal filters* for protein purification.

Spin conditions for filtration: 40 ° fixed angle rotor, 14,000 x g, at rt, 500  $\mu$ L starting volume, 20-15  $\mu$ L final volume, 10-minute spin

Spin conditions for sample recovery: reverse spinning, 30  $\mu$ L starting volume, 40° fixed angle rotor, 1,000 x g, room temperature, 4 min spin.

##### 8.4.4.1. Thiazolidine deprotection protocol

When the orthogonal deprotection of the *N*-terminal Cys of the product resulting from the first NCL was required, it was performed as one-pot procedure right after the ligation took place. After the NCL was completed, hydroxylamine was added to a 200 mM final concentration; The pH of the solution was then adjusted to 4.0, and the solution was shaken for 14 h, and the desired product was purified by RP-HPLC. The appropriate fractions were collected and freeze-dried to afford the desired product

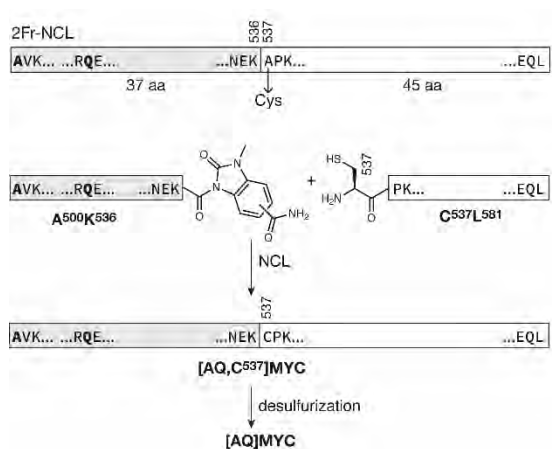
8.4.4.2. TWO-FRAGMENT APPROACH: Synthesis of [AQC<sup>537</sup>] MYC

figure S 11. Two-fragment disconnection for synthesizing the Asn<sup>500</sup>→Ala and Asn<sup>515</sup>→Gln MYC mutant

[AQC<sup>537</sup>] MYC

AVKRRTHNVLERQRRQELKRSFFALRDQIPELENNEKCPKVVLKKATAYILSVQ  
AEEQKLISEEDLLRKRREQLKHKLEQL-NH<sub>2</sub>

Synthesized from A<sup>500</sup>K<sup>536</sup>-Nbz (15 mg, 3.1  $\mu$ mol) and C<sup>537</sup>L<sup>581</sup> (25 mg, 4.7  $\mu$ mol) following the general procedure for NCL. A white powder was obtained after RP-HPLC purification (16 mg, 54 %).

ESI-MS (m/z): [M+H]<sup>+</sup> calculated for C<sub>435</sub>H<sub>739</sub>N<sub>137</sub>O<sub>124</sub>S= 9904.59; found 1981.9 [M+5H]<sup>5+</sup>, 1651.8 [M+6H]<sup>6+</sup>, 1238.9 [M+8H]<sup>8+</sup>, 1101.6 [M+9H]<sup>9+</sup>, 991.3 [M+10H]<sup>10+</sup>, 901.3 [M+11H]<sup>11+</sup>, 826.3 [M+12H]<sup>12+</sup>, 762.8 [M+13H]<sup>13+</sup>.

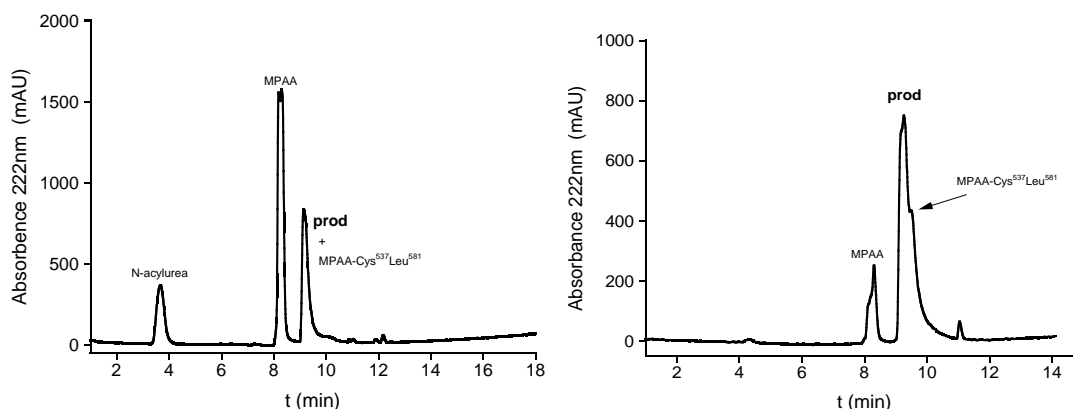


Figure S10. Left: HPLC at 222 nm of the crude reaction after 4.5 h. Right: HPLC of the crude reaction after filtered through amicon filters. prod refers to [AQC<sup>537</sup>] MYC peptide

The *N*-terminal cysteine of the fragment C<sup>537</sup>L<sup>581</sup> in excess formed a disulphide impurity with the thiol MPAA. Since the desired product and this adduct, formed by the fragment and the thiol MPAA, had the same retention time, a TCEP was added to the mixture prior to the purification. The solvent gradient was also changed in the semipreparative HPLC with respect to the gradient used in analytical runs. The best separation was achieved using a gradient from 25 % to 45 % of eluent B during 30 min.

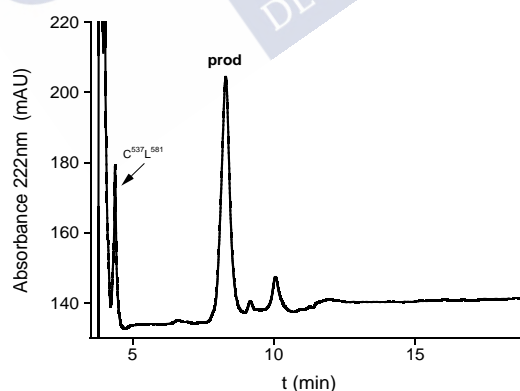


Figure S11. HPLC of the purification in the semipreparative HPLC after TCEP added. Gradient used: 25 % → 45 % of eluent B in 30 min. Prod refers to [AQC<sup>537</sup>]MYC peptide

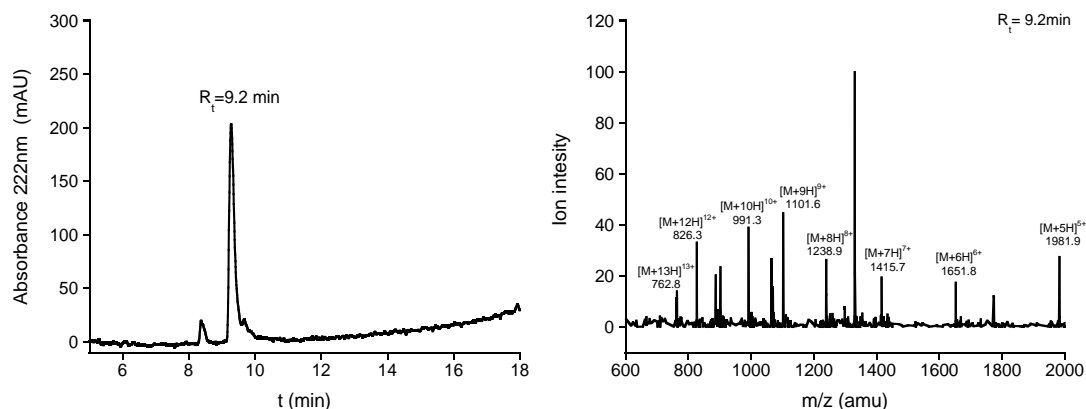
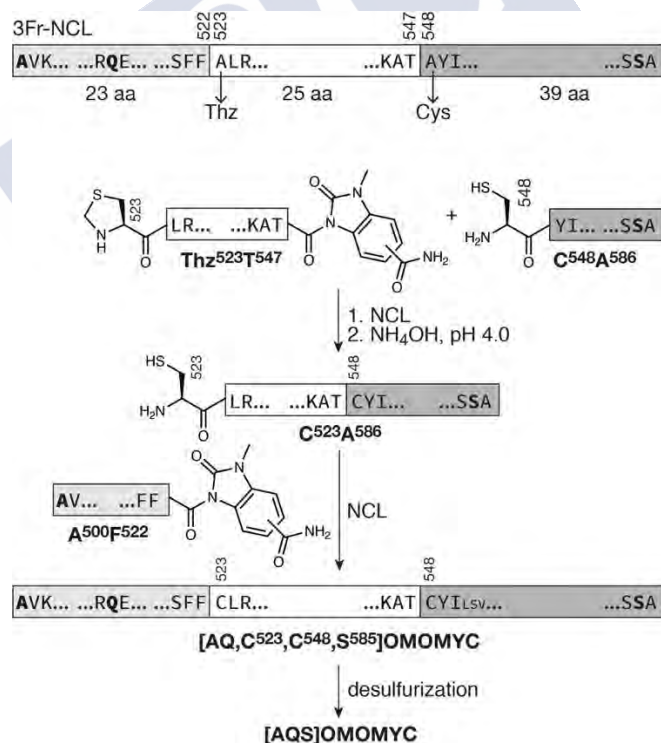


Figure S124. Left: HPLC of the pure  $[AQC^{537}]MYC$  peptide at 222 nm. Right: ESI mass spectrum of the product.

### 8.4.4.3. Three fragment approach: synthesis of $[AQS]OMOMYC(SH)$ .



Scheme 2. a) Sequence and structural elements of the peptide omomyc; in bold indicated the amino acids that are different from MYC, as well as the two N-terminal mutations we had already introduced in our synthetic MYC. b) Three-fragment disconnection for synthesizing the Asn<sup>500</sup>→Ala and Asn<sup>515</sup>→Gln omomyc mutant ( $[AQS]OMOMYC$ ). Numbering in the sequence maintains that of MYC for consistency.

## First Native Chemical Ligation

 $C^{523}A^{586}$ 

CLRDQIPELE NNEKAPKVVI LKKATCYILS VQAETQKLI SEIDLLRKQN  
 EQLKHKLEQL RNSSA-NH<sub>2</sub>

This ligation took place between fragments **Thz**<sup>523</sup>**T**<sup>546</sup>-**Nbz** (5 mg, 1.8  $\mu$ mol) and **C**<sup>548</sup>**A**<sup>586</sup> (10 mg, 2.2  $\mu$ mol) In this case, the best yield was achieved adding only 1.2 equivalents of the reactant in excess. After following the general procedure for NCL and thiazolidine deprotection, a white powder (6 mg, 43 %) was obtained.

ESI-MS (m/z): calculated for C<sub>323</sub>H<sub>552</sub>N<sub>94</sub>O<sub>99</sub>S<sub>2</sub>= 7400.65; found 1234.6 [M+6H]<sup>6+</sup>, 1058.1 [M+7H]<sup>7+</sup>, 925.9 [M+8H]<sup>8+</sup>, 823.2 [M+9H]<sup>9+</sup>, 741.1 [M+10H]<sup>10+</sup>.

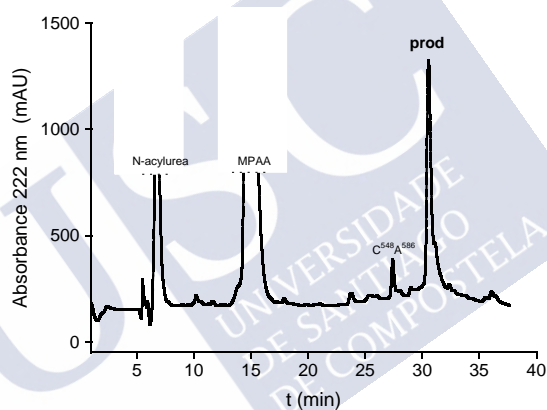


Figure 13. HPLC of the crude reaction after the 1<sup>st</sup> NCL and thiazolidine deprotection, at 222 nm. Gradient used: 20 %  $\rightarrow$  65 % in 70 min. prod refers to C<sup>523</sup>A<sup>586</sup> peptide.

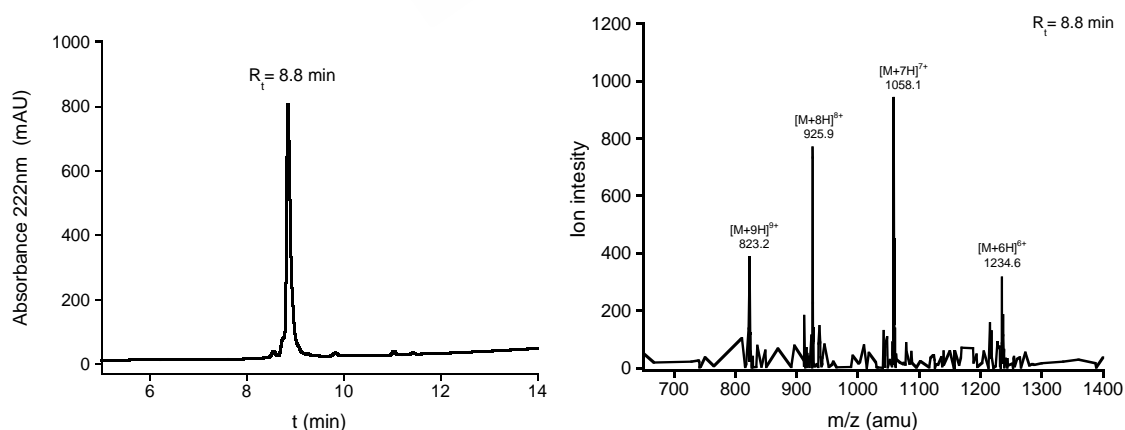


Figure S20. Left: Chromatogram at 222 nm of the pure peptide after the 1<sup>st</sup> NCL and thiazolidine deprotection. Right: ESI mass spectrum of pure product.

*Second Native Chemical Ligation*

**[AQC<sup>523</sup>C<sup>548</sup>S<sup>586</sup>] OMOMYC or -[AQS]OMOMYC(SH)**

**AVKRRTHNVL ERQRRQELKR SFFCLRDQIP ELENNEKAPK VVILKKATCY**  
**ILSVQAETQK LISEIDLLRK QNEQLKHKLE QLRNSSA-NH<sub>2</sub>**

This ligation took place between fragments [C<sup>523</sup>A<sup>587</sup>] (14 mg, 2  $\mu$ mol) and A<sup>500</sup>F<sup>522</sup>-Nbz (10 mg, 3  $\mu$ mol) according to the general procedure for NCL. The product was obtained as a white powder (10.2 mg, 49 %) after RP-HPLC purification and lyophilization

**ESI-MS** (m/z): calculated for C<sub>485</sub>H<sub>801</sub>N<sub>147</sub>O<sub>138</sub>= 10338.07; found 1723.4 [M+6H]<sup>6+</sup>, 1477.9 [M+7H]<sup>7+</sup>, 1293.4 [M+8H]<sup>8+</sup>, 1149.7 [M+9H]<sup>9+</sup>, 1034.6 [M+10H]<sup>10+</sup>, 940.5 [M+11H]<sup>11+</sup>, 862.5 [M+12H]<sup>12+</sup>, 796.2 [M+13H]<sup>13+</sup>

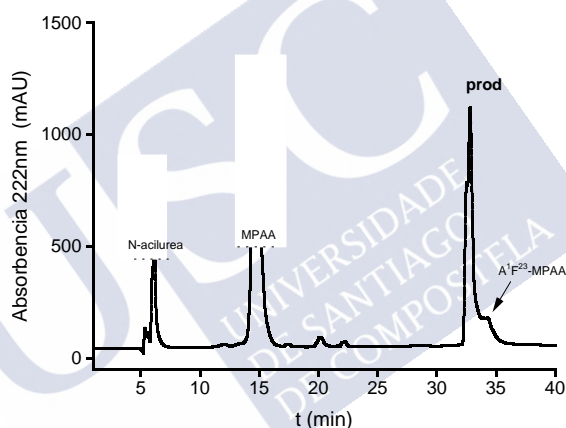


Figure 122. Chromatogram at 222 nm the crude NCL. Gradient: 20 % to 60 % of MeCN + 0.1 % TFA in 30 min.

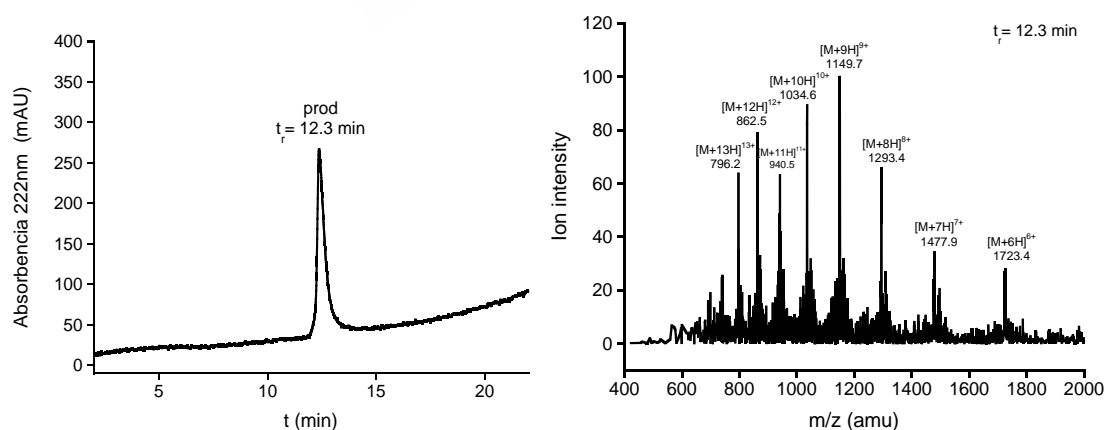
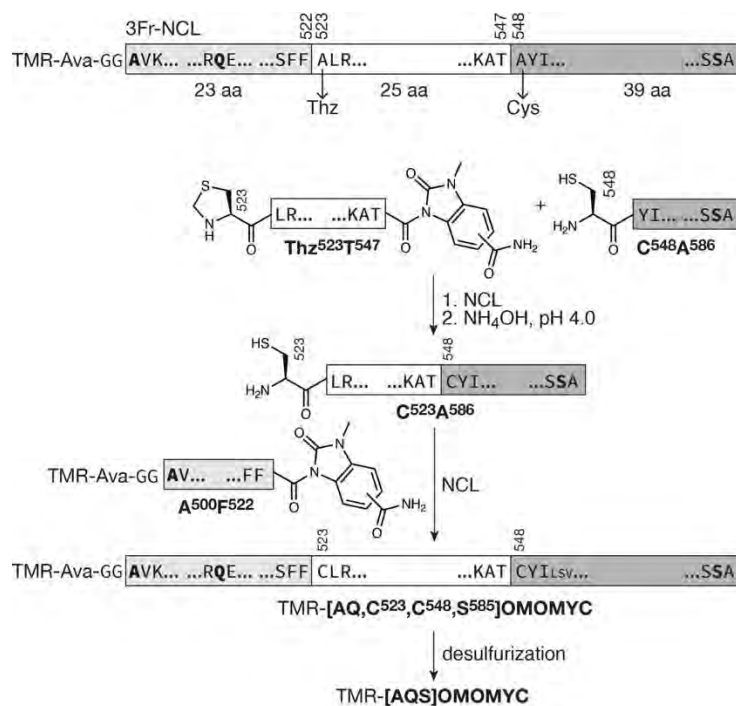


Figure S22. Left: HPLC at 222 nm of the pure product. Gradient used: 5 %  $\rightarrow$  95 % of B in 19 min. Right: MS spectrum of the product.

## 8.4.4.4. THREE-FRAGMENT APPROACH. Synthesis of TMR-[AQ]MYC(SH)



Scheme 3. Three-fragment disconnections used for synthesizing the Asn<sup>500</sup>→Ala and Asn<sup>515</sup>→Gln MYC mutant

## First Native Chemical Ligation

C<sup>523</sup>L<sup>581</sup>

CLRDQIPELE NNEKAPKVVI LKKATCYILS VQAEEQKLIS EEDLLRKRRE  
QLKHKLEQL-NH<sub>2</sub>

This ligation took place between fragments Thz<sup>523</sup>T<sup>546</sup>-Nbz (20 mg, 6.6  $\mu$ mol) and C<sup>549</sup>L<sup>581</sup> (41 mg, 9.9  $\mu$ mol) according to the general procedure for NCL. A white powder (9 mg, 20 %) was obtained after thiazolidine deprotection and RP-HPLC purification

ESI-MS (m/z): calculated for C<sub>307</sub>H<sub>525</sub>N<sub>89</sub>O<sub>92</sub>S<sub>2</sub>= 6999.23; found 1750.9 [M+4H]<sup>4+</sup>, 1400.8 [M+5H]<sup>5+</sup>, 1000.8 [M+7H]<sup>7+</sup>, 1167.4 [M+6H]<sup>6+</sup>, 875.8 [M+8H]<sup>8+</sup>, 778.3 [M+9H]<sup>9+</sup>.



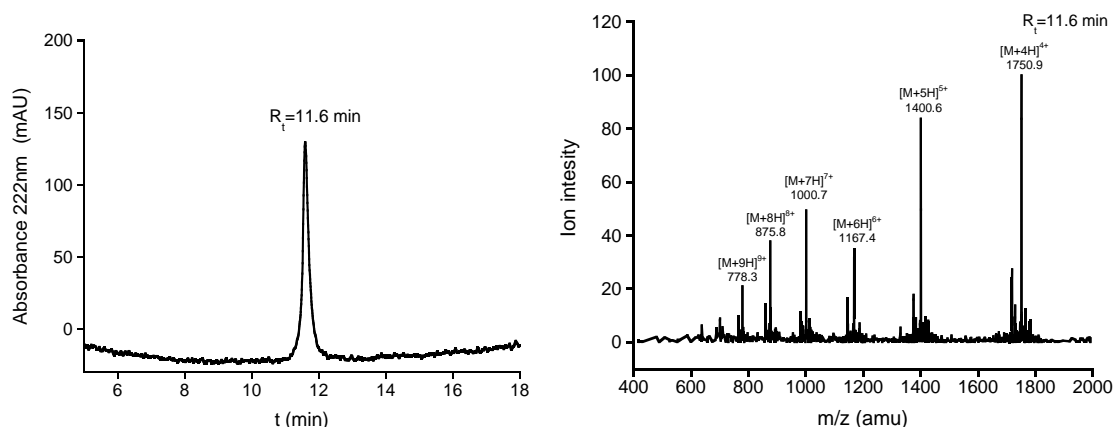


Figure S. Left: HPLC of the pure peptide at 222 nm. Right: ESI mass spectrum of product.

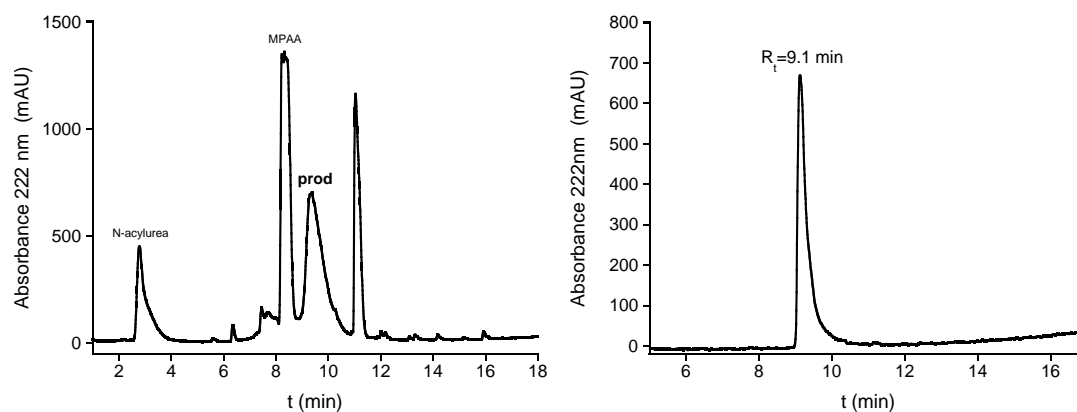
## Second Native Chemical Ligation

### TMR-[AQ]MYC(SH)

**TMR**-Ava-GGA VKRRTHN VLERQRRQEL KRSFFCLRDQ IPELENNEK APKVVILKKA  
**TC**YILSVQAE EQKLISEED LLRKRREQLK HKLEQL-NH<sub>2</sub>.

This ligation took place between fragments [C<sup>523</sup>L<sup>581</sup>]-NH<sub>2</sub> (14 mg, 2  $\mu$ mol) and **TMR**-AvaGG-A<sup>500</sup>F<sup>522</sup>-Nbz (11 mg, 3  $\mu$ mol) according to the general procedure for NCL. A pink powder (9.5 mg, 45 %) was obtained after RP-HPLC purification and lyophilization.

ESI-MS (m/z): calculated for C<sub>469</sub>H<sub>774</sub>N<sub>142</sub>O<sub>131</sub>S<sub>2</sub>= 10562.33; found 1761.39 [M+6H]<sup>6+</sup>, 1509.8 [M+7H]<sup>7+</sup>, 1321.0 [M+8H]<sup>8+</sup>, 1174.4 [M+9H]<sup>9+</sup>, 1056.9 [M+10H]<sup>10+</sup>, 960.9 [M+11H]<sup>11+</sup>, 881.0 [M+12H]<sup>12+</sup>, 813.5 [M+13H]<sup>13+</sup>



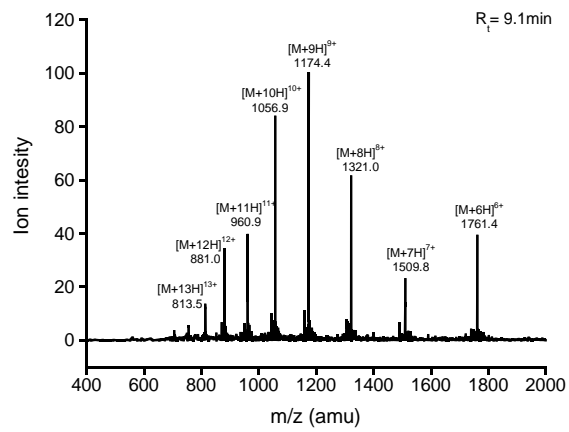
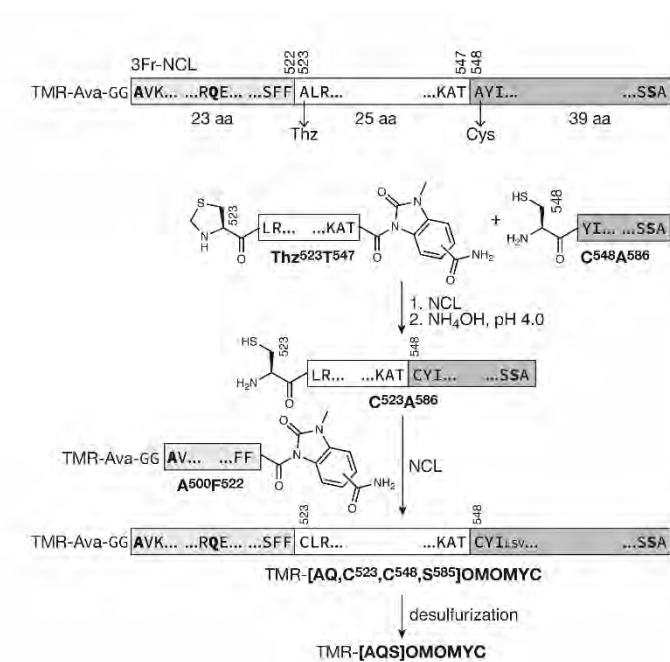


Figure S148. Top left: HPLC of the crude reaction after the 2<sup>nd</sup> NCL, at 222 nm; Top right: HPLC of the pure peptide at 222 nm. Bottom: ESI mass spectrum of product. Prod refers to TMR-[AQ][MYC(SH)]



## 8.4.4.5. three fragment approach. synthesis of TMR-[AQS]OMOMYC(SH)



Scheme 4. Three-fragment disconnection for synthesizing the Asn<sup>500</sup>→Ala and Asn<sup>515</sup>→Gln omomyc labelled mutant (TMR-[AQS]OMOMYC). Numbering in the sequence maintains that of MYC for consistency

### First Native Chemical Ligation

Same as C<sup>523</sup>A<sup>586</sup>

*Second Native Chemical Ligation*

**TMR-[AQC<sup>523</sup>C<sup>548</sup>S<sup>585</sup>] OMOMYC or TMR-[AQS]OMOMYC(SH)**

**TMR-AvaGG AVKRRTHNVL ERQRRQELKR SFFALRDQIP ELENNEKAPK  
VVILKKATAY ILSVQAE<sup>T</sup>QK LISEIDLLRK QNEQLKHKLE QLRNSSA-NH<sub>2</sub>**

This ligation took place between fragments [C<sup>523</sup>A<sup>586</sup>] (7 mg, 1  $\mu$ mol) and **TMR-AvaGG-A<sup>500</sup>F<sup>522</sup>-Nbz** (6 mg, 1.5  $\mu$ mol) according to the general procedure for NCL. The product was obtained as a pink powder (4.2 mg, 39 %) after RP-HPLC purification and lyophilization

**ESI-MS** (m/z): calculated for C<sub>485</sub>H<sub>801</sub>N<sub>147</sub>O<sub>138</sub>= 10899.63; found 1371.8 [M+8H]<sup>8+</sup>, 1219.2 [M+9H]<sup>9+</sup>, 1097.4 [M+10H]<sup>10+</sup>, 997.8 [M+11H]<sup>11+</sup>, 914.5 [M+12H]<sup>12+</sup>, 844.0 [M+13H]<sup>13+</sup>, 784.3 [M+14H]<sup>14+</sup>.

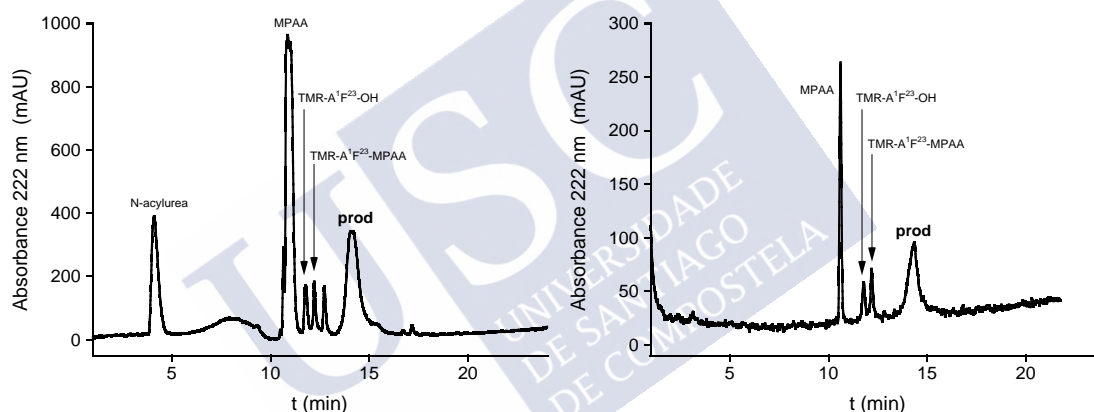


Figure S2115. Left: HPLC at 222 nm of the crude reaction after the 2<sup>nd</sup> NCL. Gradient used: 5 %  $\rightarrow$  75 % of B in 19 min. Right: Crude of the reaction after amicon filtered. Prod refers to TMR-[AQS]OMOMYC(SH).

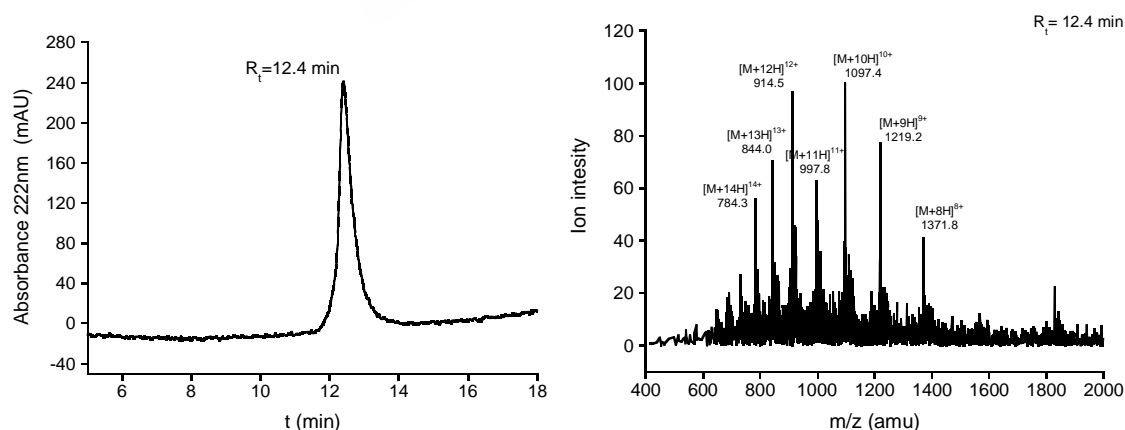


Figure S22. Left: HPLC at 222 nm of the pure product. Gradient used: 5 %  $\rightarrow$  95 % of B in 19 min. Right: MS spectrum of the product.

### 8.4.5. Desulfurization of the resulting miniproteins

0.3  $\mu$ mol of the corresponding peptide was dissolved in degassed 6 M  $\text{Gn}\cdot\text{HCl}$ , 200 mM  $\text{K}_2\text{HPO}_4$  and 500 mM TCEP (pH= 7) to a 100  $\mu$ M concentration. After degassing back the resulting mixture for 10 minutes, 4,6 % (v/v) of EtSH and 40 mM of VA-044 (aq) were added. Reaction mixture was shaken at 37 °C for 5 ½ hours. Then, it was quenched with TFA and directly purified by RP-HPLC to afford the desired product.

#### 8.4.5.1. Desulfurization of [AQC]MYC obtained by the 2-fragment approach

##### [AQ]MYC

AVKRRTHNVL ERQRRQELKR SFFALRDQIP ELENNEKAPK VVILKKATAY  
ILSVQAEQK LISEEDLLR KRREQLKHKL EQL-NH<sub>2</sub>

The product was obtained as a white solid (2.2 mg, 76 %) from pure starting peptide after purification.

ESI-MS (m/z): calculated for  $\text{C}_{435}\text{H}_{739}\text{N}_{137}\text{O}_{124}$  = 9872.53; found 1975.6  $[\text{M}+5\text{H}]^{5+}$ , 1646.1  $[\text{M}+6\text{H}]^{6+}$ , 1411.2  $[\text{M}+7\text{H}]^{7+}$ , 1234.9  $[\text{M}+8\text{H}]^{8+}$ , 1097.6  $[\text{M}+9\text{H}]^{9+}$ , 988.0  $[\text{M}+10\text{H}]^{10+}$ , 898.7  $[\text{M}+11\text{H}]^{11+}$ , 823.8  $[\text{M}+12\text{H}]^{12+}$ , 760.4  $[\text{M}+13\text{H}]^{13+}$

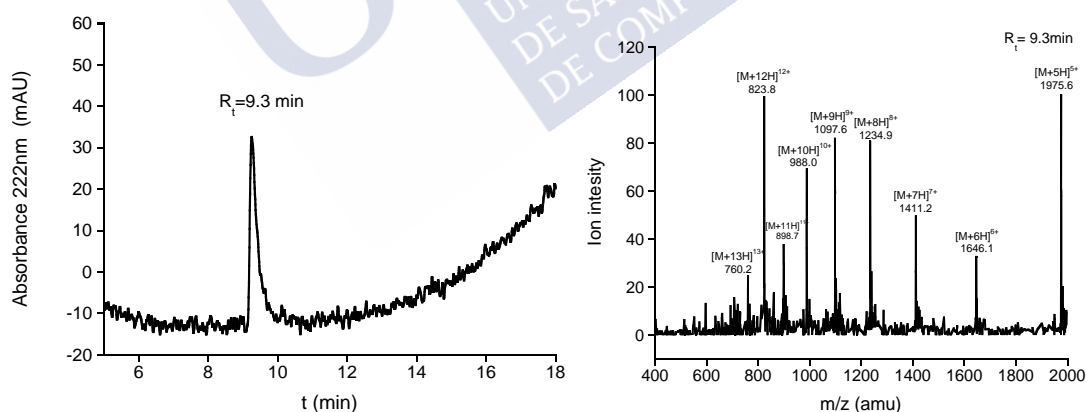


Figure S163. Left: HPLC of the pure peptide at 222 nm. Right: ESI mass spectrum of product.

### 8.4.5.2. Desulfurization of [AQS]OMOMYC(SH) obtained by the 3-fragment approach

#### [AQS]OMOMYC

AVKRRTHNVLERQRRQELKRSFFALRDQIPELENNEKAPKVVILKKATA YILSVQAET  
 QKLISEIDLLRKQNEQLKHKLEQLRNSSA-NH<sub>2</sub>

After RP-HPLC purification, pure peptide was obtained as a white solid (2 mg, 75 %).

ESI-MS (m/z): calculated for C<sub>451</sub>H<sub>766</sub>N<sub>142</sub>O<sub>131</sub>= 10267.76; found 1712.3 [M+6H]<sup>6+</sup>, 1467.9 [M+7H]<sup>7+</sup>, 1141.9 [M+9H]<sup>9+</sup>, 1027.8 [M+10H]<sup>10+</sup>, 934.5 [M+11H]<sup>11+</sup>, 856.7 [M+12H]<sup>12+</sup>, 790.8 [M+13H]<sup>13+</sup>, 734.4 [M+14H]<sup>14+</sup>

### 8.4.5.3. Desulfurization of TMR-[AQ]MYC(SH) obtained by the 3-fragment approach

#### TMR-[AQ]MYC

TMR<sub>A</sub>VaGGAVKRRTHNVLERQRRQELKRSFFALRDQIPELENNEKAPKVVILKK  
 ATA YILSVQAEEQKLISEEDLLRKRREQLKHKLEQL-NH<sub>2</sub>

After purification, the pure peptide (2.2 mg, 70 %) was isolated as a pink solid.

ESI-MS (m/z): calculated for C<sub>469</sub>H<sub>774</sub>N<sub>142</sub>O<sub>131</sub>= 10498.23; found 1750.8 [M+6H]<sup>6+</sup>, 1500.0 [M+7H]<sup>7+</sup>, 1313.3 [M+8H]<sup>8+</sup>, 1167.0 [M+9H]<sup>9+</sup>, 1051.4 [M+10H]<sup>10+</sup>, 954.2 [M+11H]<sup>11+</sup>, 876.1 [M+12H]<sup>12+</sup>, 751.9 [M+14H]<sup>14+</sup>.

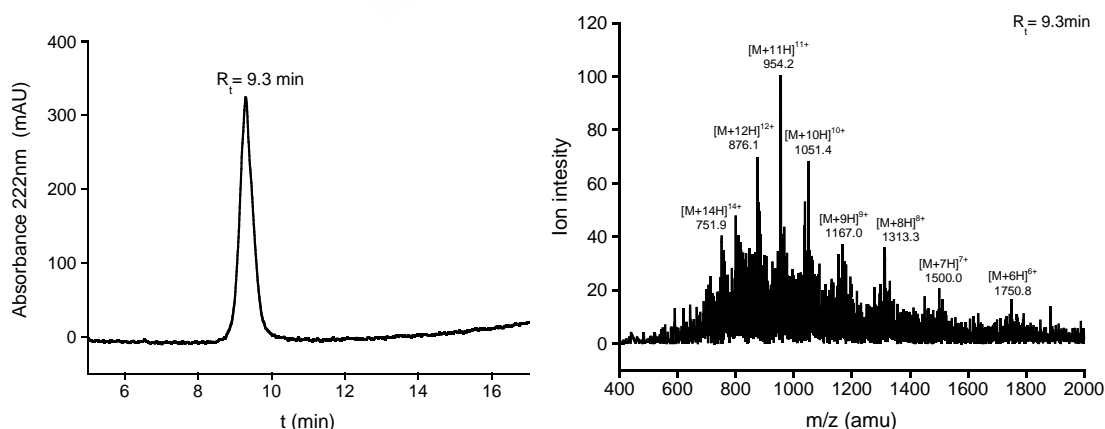


Figure S17. Left: HPLC of the pure peptide at 222 nm. Right: ESI mass spectrum of product.

#### 8.4.5.4. Desulfurization of TMR-[AQS]OMOMYC(SH) obtained by the 3-fragment approach

##### TMR-[AQS]OMOMYC

TMR-Ava-GGA-VKRRTHNVLE RQRRQELKRS FFALRDQIP ELENNEKAP  
KVVILKKATA YILSVQAET QKLISEIDLL RKQNEQLKHK LEQLRNSSA-NH<sub>2</sub>

After RP-HPLC purification, pure peptide was obtained as a pink solid (2.6 mg, 80 %).

ESI-MS (m/z): calculated for C<sub>485</sub>H<sub>801</sub>N<sub>147</sub>O<sub>138</sub>= 10899.63; found 1817.5 [M+6H]<sup>6+</sup>, 1558.0 [M+7H]<sup>7+</sup>, 1212.0 [M+9H]<sup>9+</sup>, 1091.0 [M+10H]<sup>10+</sup>, 991.7 [M+11H]<sup>11+</sup>, 909.4 [M+12H]<sup>12+</sup>, 839.3 [M+13H]<sup>13+</sup>, 779.5 [M+14H]<sup>14+</sup>

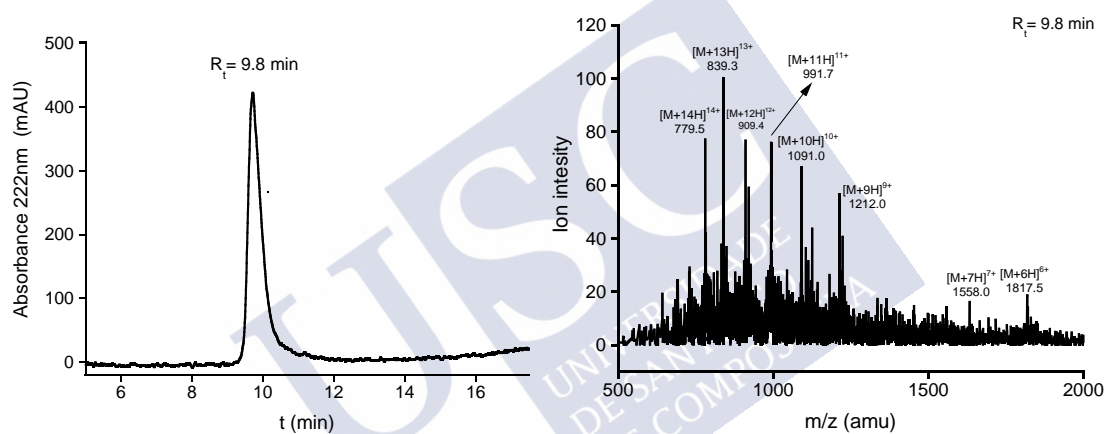
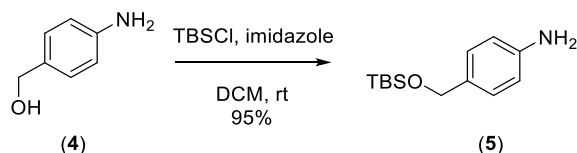


Figure S26. Left: HPLC chromatogram of the pure desulfurized peptide at 222 nm. Right: ESI mass spectrum of the peptide.

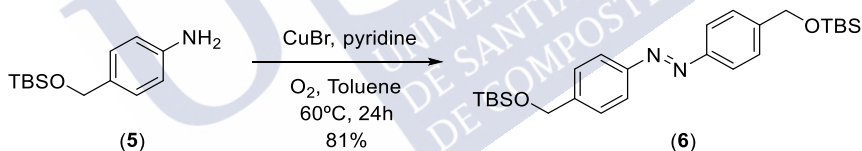
## 8.5. REACTIONS OF CHAPTER 2

## Synthesis of (5)



To a solution of **4** (1.01 g, 8.16 mmol) in  $\text{CH}_2\text{Cl}_2$  (40 mL) TBSCl (1.35 g, 8.98 mmol) and imidazole (611.35 mg, 8.98 mmol) were added, and the resulting solution was stirred at rt for 30 min.<sup>158</sup> The resulting suspension was concentrated under reduced pressure, and the residue was purified by flash chromatography (40 % EtOAc/ $\text{CH}_2\text{Cl}_2$ ) to afford **5** (1.84 g, 7.75 mmol, 95 %). As a white solid.

$^1\text{H NMR}$  (250 MHz,  $\text{CD}_2\text{Cl}_2$ ,  $\delta$ ): 0.08 (s, 6 H), 0.93 (s, 9 H), 3.76 (s, 2 H), 4.63 (s, 2 H), 6.68 (d,  $J = 7.3$  Hz, 2 H), 7.08-7.15 (m, 2 H).

Synthesis of (6)<sup>159</sup>

Pyridine (163  $\mu\text{L}$ , 2.0 mmol) was added to a mixture of **5** (1.60 g, 6.7 mmol), CuBr (0.096 g, 0.67 mmol) in toluene (20 mL), and the resulting mixture was stirred for 24 h at 60 °C. After cooling to rt, HCl 10 % (50 mL) was added and the residue was extracted with hexane (3  $\times$  50 mL). the combined organic extracts were washed with brine (10 mL), dried ( $\text{Na}_2\text{SO}_4$ ) and concentrated under reduced pressure. The resulting mixture was purified by  $\text{SiO}_2$  chromatography (50%  $\text{CH}_2\text{Cl}_2$ /hexane) to yield the desired product **6** as an orange solid (1.27 g, 2.70 mmol, 81 %).

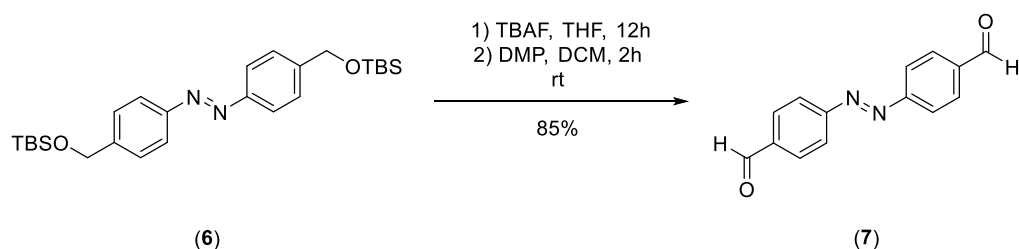
$^1\text{H NMR}$  (300 MHz,  $\text{CDCl}_3$   $\delta$ ): 0.23 (s, 12H), 1.09 (s, 18H), 5.05 (s, 4H), 7.50 (d,  $J = 8.61$  Hz, 4H), 7.78 (d,  $J = 8.61$  Hz, 4H)

<sup>158</sup> Hashmi, S.; Bührle, M.; Wölfle, M.; Rudolph, M.; Wietek, M.; Rominger F., Frey, W. *Chem. Eur. J.* **2010**, *16*, 9846.

<sup>159</sup> Chun, Z.; Ning, J. *Angew. Chem. Int. Ed.* **2010**, *49*, 6174.



### Synthesis of (7)



The azo-compound **6** (1.194 g, 2.54 mmol) was dissolved in a round bottom flask in THF (25 mL), TBAF (3.8 mL, 3.83 mmol). was added dropwise. The mixture was stirred for 12 h at rt. Then, 50 mL of 10% aqueous HCl was added and the crude was extracted with EtOAc (5 × 30) the combined organic layers were washed with brine (10 mL), dried (Na<sub>2</sub>SO<sub>4</sub>) and concentrated under reduced pressure to give the unprotected dialcohol as an orange powder. This solid (109 mg, 0.45 mmol) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (4.5 mL) and treated with DMP (420 mg 0.99 mmol) for 2 hours at rt. The reaction was quenched with 1 M Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> (3 mL), followed by addition of 30 mL of brine and extraction with EtOAc (3 × 30 mL). The combined organic extracts were dried with anhydrous Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. Finally, the crude was purified by RP-HPLC and freeze-dried to provide **7** (514 mg, 2.16 mmol, 85 % overall yield of the two steps) as a red solid.

**<sup>1</sup>H NMR:** (300 MHz, DMSO-*d*<sub>6</sub> δ): 8.07-8.16 (m, 8H), 10.21 (s, 2H).

**<sup>13</sup>C NMR, DEPT:** (75 MHz, DMSO-*d*<sub>6</sub> δ): 122.8 (CH), 130.2 (CH), 137.1 (C), 154.7 (C), 191.9 (CHO).

**HRMS (ESI<sup>+</sup>):** [M+H] calc. for C<sub>14</sub>H<sub>10</sub>N<sub>2</sub>O<sub>2</sub> = 238,0742. found 238,0746; C<sub>14</sub>H<sub>10</sub>N<sub>2</sub>O<sub>2</sub> (M.W. 238,0742).

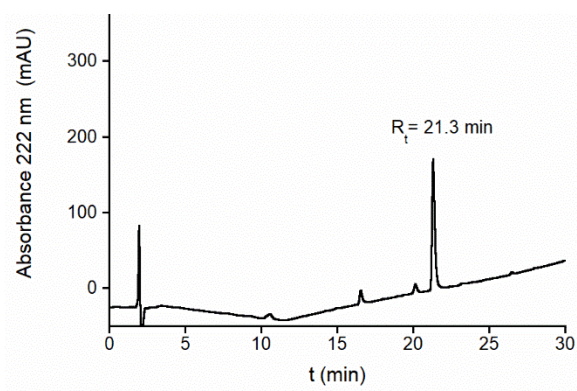
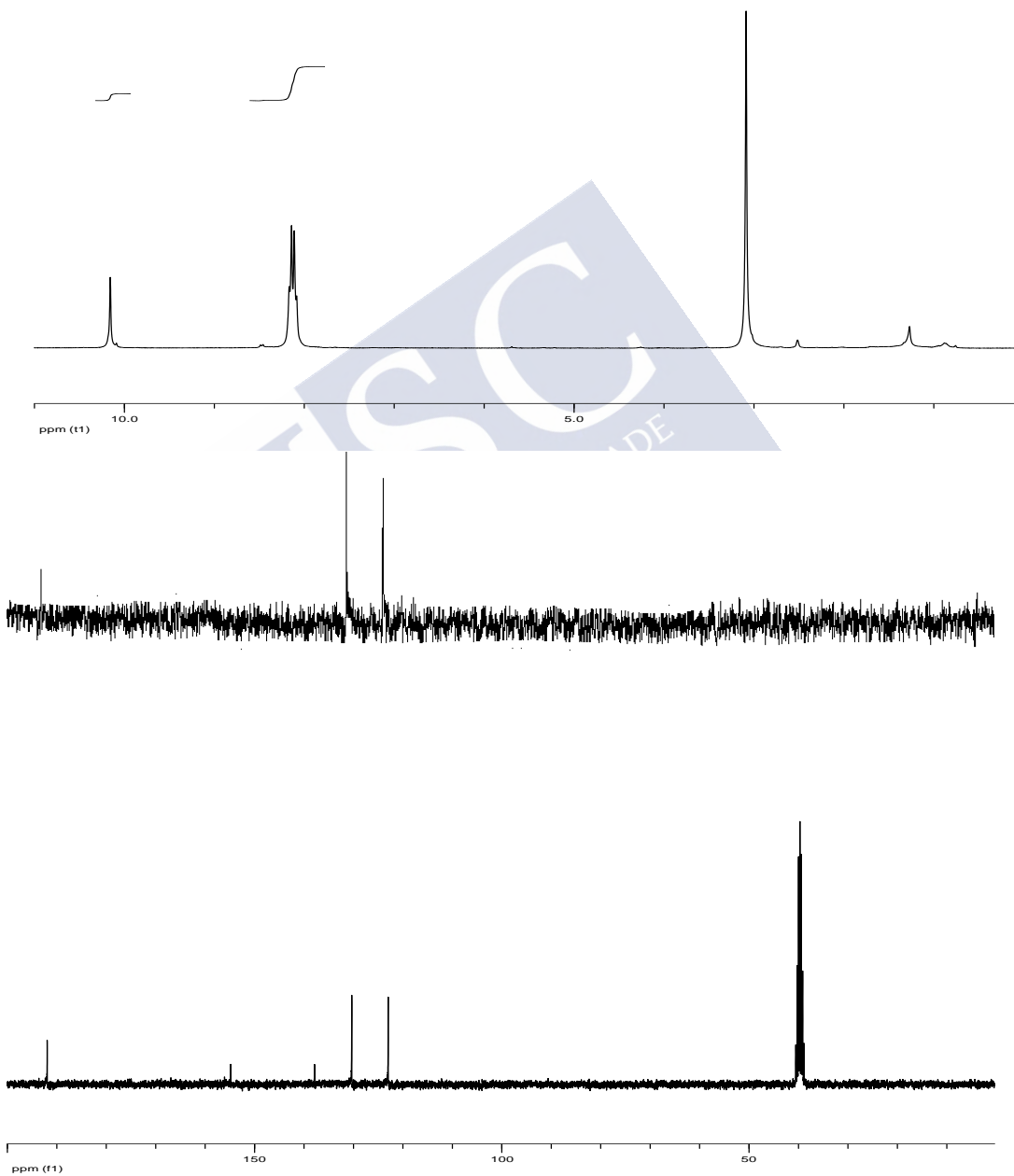
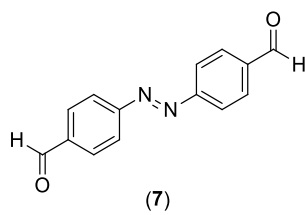
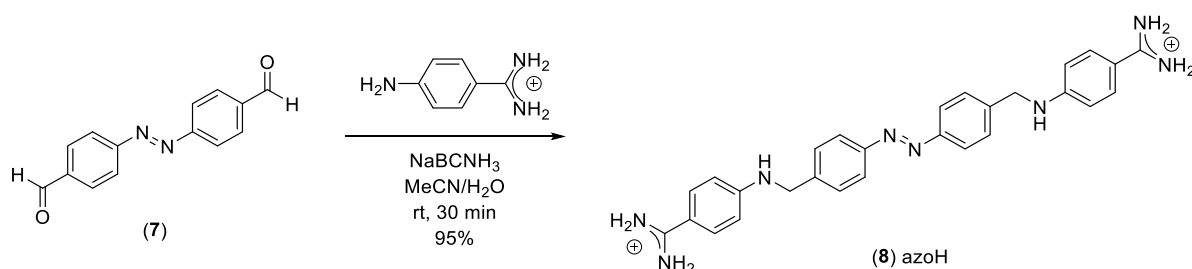


Figure 18. HPLC chromatogram at 222 nm of the pure product

(7)



## Synthesis of azoH



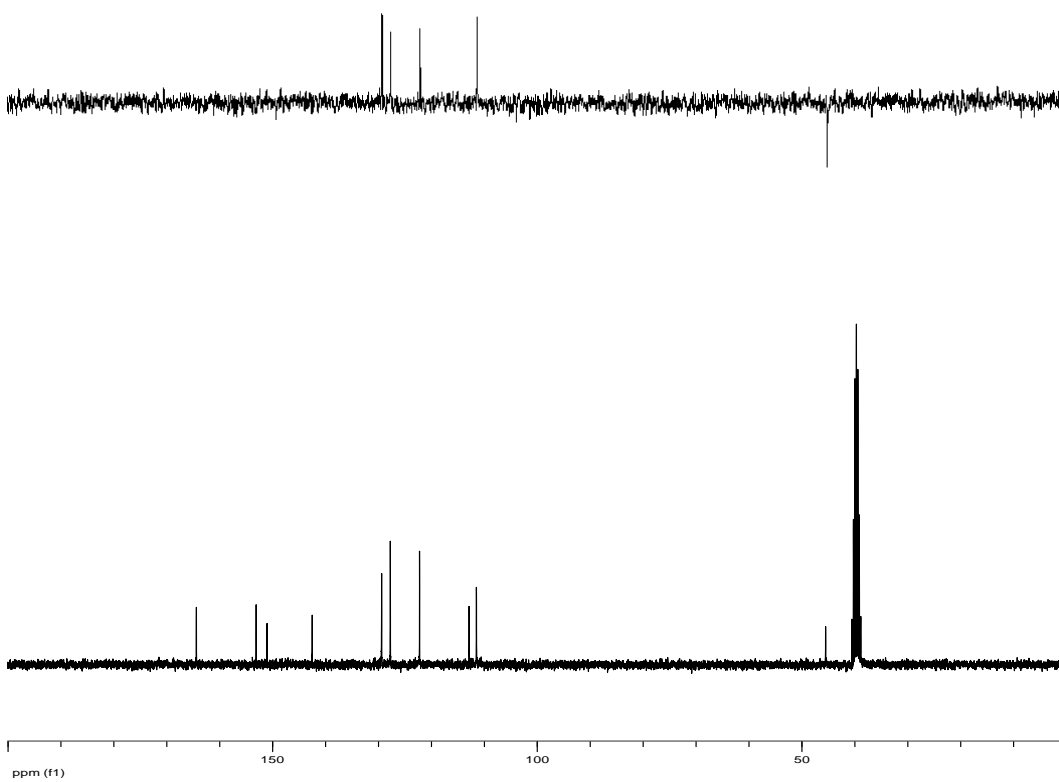
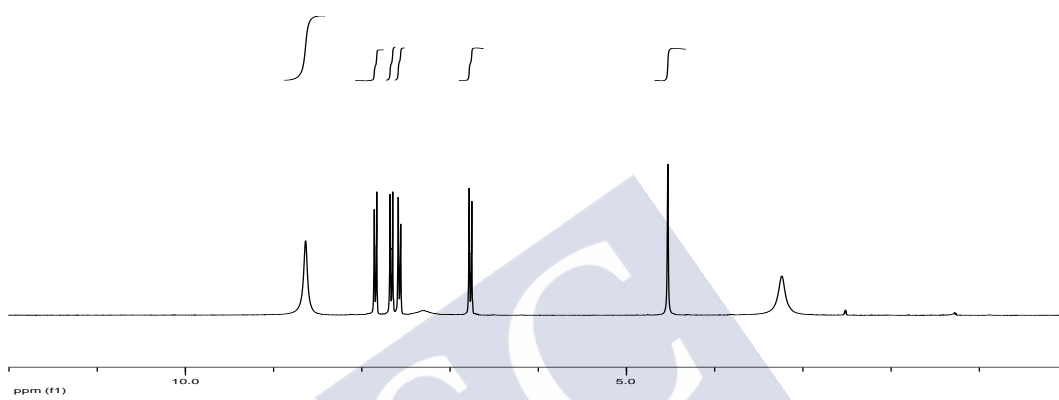
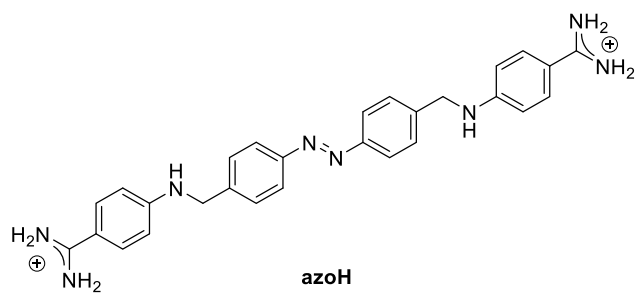
NaBCNH<sub>3</sub> (28.16 mg, 0.447 mmol) was added to a solution of the aminobenzimidine (130 mg, 0.630 mmol) and **7** (71 mg, 0.298 mmol) in MeCN/H<sub>2</sub>O (1:3) (3 mL). The resulting solution was stirred for 30 min at rt. The reaction mixture was directly purified by RP-HPLC. The appropriate fractions were lyophilized and an orange powder (200 mg, 0.283 mmol, 90 %) was identified as the trifluoroacetic salt of the **azoH**

**<sup>1</sup>H NMR:** (300 MHz, DMSO-*d*<sub>6</sub> δ): 4.51 (s, 4H), 6.76 (d, *J* = 8.87 Hz, 4H), 7.56 (d, *J* = 8.36 Hz, 4H), 7.65 (d, *J* = 8.87 Hz, 4H), 7.84 (d, *J* = 8.36 Hz, 4H), 8.63 (s, 8H).

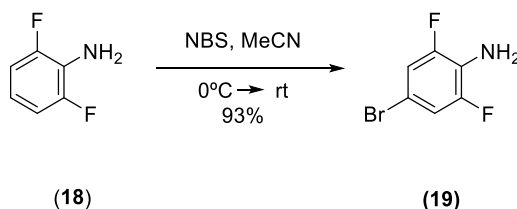
**<sup>13</sup>C NMR, DEPT:** (75 MHz, DMSO-*d*<sub>6</sub> δ): 45.2 (CH<sub>2</sub>), 111.3 (CH), 112.6 (C), 122.0 (CH), 127.5 (CH), 129.2 (CH), 142.3 (C), 150.9 (C), 152.9 (C), 164.2 (C).

**HRMS (ESI<sup>+</sup>)** [M+H] calc. for C<sub>28</sub>H<sub>31</sub>N<sub>8</sub> = 479.2566. Found 479.2569; C<sub>30</sub>H<sub>29</sub>F<sub>6</sub>N<sub>8</sub>O<sub>4</sub> (M.W.: 706.6365).

## azoH



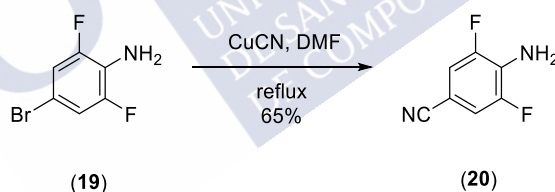
### Synthesis of (19)



Compound **(19)** was synthesized following a previous methodology.<sup>160</sup> A stirring solution of 2,6-difluoroaniline **(18)** (4 g, 31 mmol), in acetonitrile (62 mL) was cooled in an iced/water bath. NBS (5.5 g, 31 mmol) was added and the mixture was allowed to reach rt and stirred for 12 h. The solvent was removed under reduced pressure and the residue purified by flash chromatography on silica gel (50 % CH<sub>2</sub>Cl<sub>2</sub>/Hexanes) to afford **(19)** (6,00 g, 29 mmol, 93 %) as a purple solid.

<sup>1</sup>H NMR: (300 MHz, CDCl<sub>3</sub> δ): 7.00 (dd, *J* = 6.1, 1.3 Hz, 2 H), 3.75 (s, 2 H).

### Synthesis of (20)

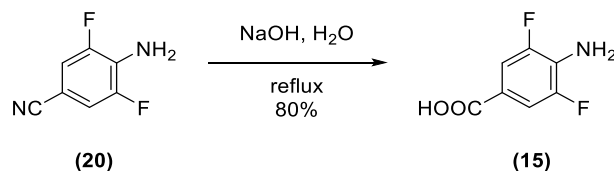


To a solution of 2,6-difluoro-4-bromoaniline **(19)** (5 g, 24.00 mmol) in DMF (50 mL) was added CuCN (6,40 g, 72.00 mmol), heated to reflux and stirred for 14 h under Ar.<sup>3</sup> TLC analysis showed that the starting material was consumed. The mixture was then poured into a NH<sub>4</sub>Cl aqueous solution and extracted with ethyl acetate. The combined organic layers were dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered, and the solvent was removed. The residue purified by flash column chromatography on silica gel (30 % CH<sub>2</sub>Cl<sub>2</sub>/hexanes) to obtain the desired product as a white powder **(20)** (2.40 g, 15.59 mmol, 65 %).

<sup>1</sup>H NMR: (300 MHz, CDCl<sub>3</sub> δ): 7.23 (dd, *J* = 6.0, 1.7 Hz, 2 H), 3.74 (s, 2 H).

<sup>160</sup> Bléger D., Schwarz J., Brouwer A., Hecht S. *J. Am. Chem. Soc.* **2012**, *134*, 20597.

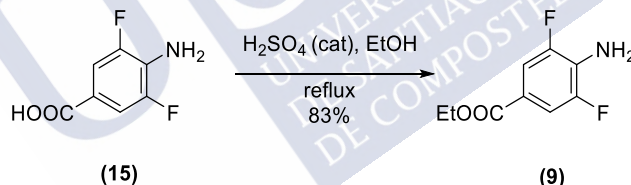
### Synthesis of (15)



4-amino-3,5-difluorobenzonitrile (**20**) (2 g, 13 mmol) was refluxed in 1 M NaOH (80 mL) for 12 h.<sup>161</sup> The mixture was then cooled to rt and the desired product (**15**) was precipitated with 1 M HCl (50 mL). After filtration, the precipitate was lyophilized and the hydrochloric salt of the acid was obtained as a brown solid (1.8 g, 10.40 mmol, 80 %).

<sup>1</sup>H NMR: (300 MHz, *DMSO-d*<sub>6</sub> δ) 3.81 (s, 2H), 7.40 (dd, *J* = 4.2, 1.4 Hz, 2 H).

### Synthesis of (9)

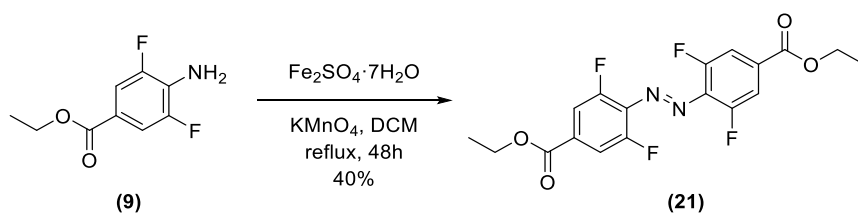


In a round bottom flask containing 4-amino-3,5-difluorobenzoic acid (**15**) (1.6 g, 9.15 mmol), EtOH (35 mL) and a few drops of H<sub>2</sub>SO<sub>4</sub> as catalyst were added. The mixture was refluxed for 5 h, neutralized with a saturated NaHCO<sub>3</sub> aqueous solution. The mixture was extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 × 60 mL). The organic fractions were dried over anhydrous Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure to afford the desired product (**9**) (1.55 g, 7.71 mmol, 83 %) as a brown solid without further purification.

<sup>1</sup>H RMN: (300 MHz, *CDCl*<sub>3</sub> δ) 7.50 (dd, *J* = 7.2, 2.2 Hz, 2 H), 4.30 (q, *J* = 7.1 Hz, 2 H), 4.07 (s, 2 H), 1.34 (t, *J* = 7.1 Hz, 3H).

<sup>161</sup> Bléger D., Schwarz J., Brouwer A., Hecht S. *J. Am. Chem. Soc.* **2012**, *134*,20597.

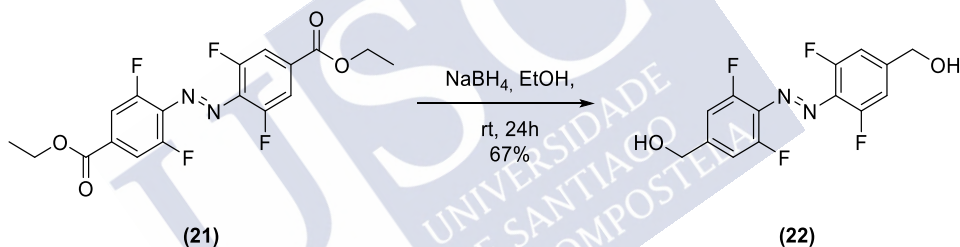
### Synthesis of (21)



A mixture of **9** (0.5 g, 2.18 mmol),  $\text{Fe}_2\text{SO}_4 \cdot 7\text{H}_2\text{O}$  (4.25 g, 13 mmol),  $\text{KMnO}_4$  (4.25 g, 27 mmol) in  $\text{CH}_2\text{Cl}_2$  (75 mL) was refluxed for 48 h. After cooling to rt, the crude filtered through celite, concentrated under reduced pressure and purified by  $\text{SiO}_2$  chromatography (50%  $\text{CH}_2\text{Cl}_2$ /hexane) to yield **(21)** (174 mg, 0,44 mmol, 40 %) as a red solid.

**$^1\text{H}$  NMR:** (300 MHz,  $\text{CDCl}_3$   $\delta$ ) 7.74-7.81 (m, 4H), 4.45 (q,  $J = 7.0$  Hz, 4H), 1.45 (t,  $J = 7.0$  Hz, 6H).

### Synthesis of (22)



$\text{NaBH}_4$  (152 mg, 4.02 mmol). was added dropwise to a solution of **(21)** (200 mg, 0.5 mmol) in EtOH (10 mL) The resulting solution was stirred for 24 at rt under Ar atmosphere.  $\text{H}_2\text{O}$  milli Q (50 mL) was added to quench the reaction and the crude was extracted with EtOAc (3 x 80 mL). The organic phase was dried with anhydrous  $\text{Na}_2\text{SO}_4$  and concentrated over reduced pressures. The product was purified by  $\text{SiO}_2$  chromatography (50 % EtOAc/Hexane) and an orange powder was identified as the dialcohol **(22)** (106 mg, 0.34 mmol, 67 %).

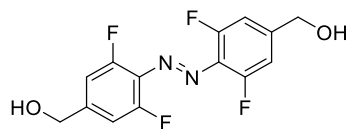
**E-isomer-  $^1\text{H}$  NMR** (250 MHz,  $\text{CD}_3\text{OD}$   $\delta$ ): 7.13 (d,  $J = 10.49$  Hz, 4H), 4.64 (s, 4H).

**Z-isomer-  $^1\text{H}$  NMR** (250 MHz,  $\text{CD}_3\text{OD}$   $\delta$ ): 6.93 (d,  $J = 8.41$  Hz, 1H), 4.52 (s, 1H).

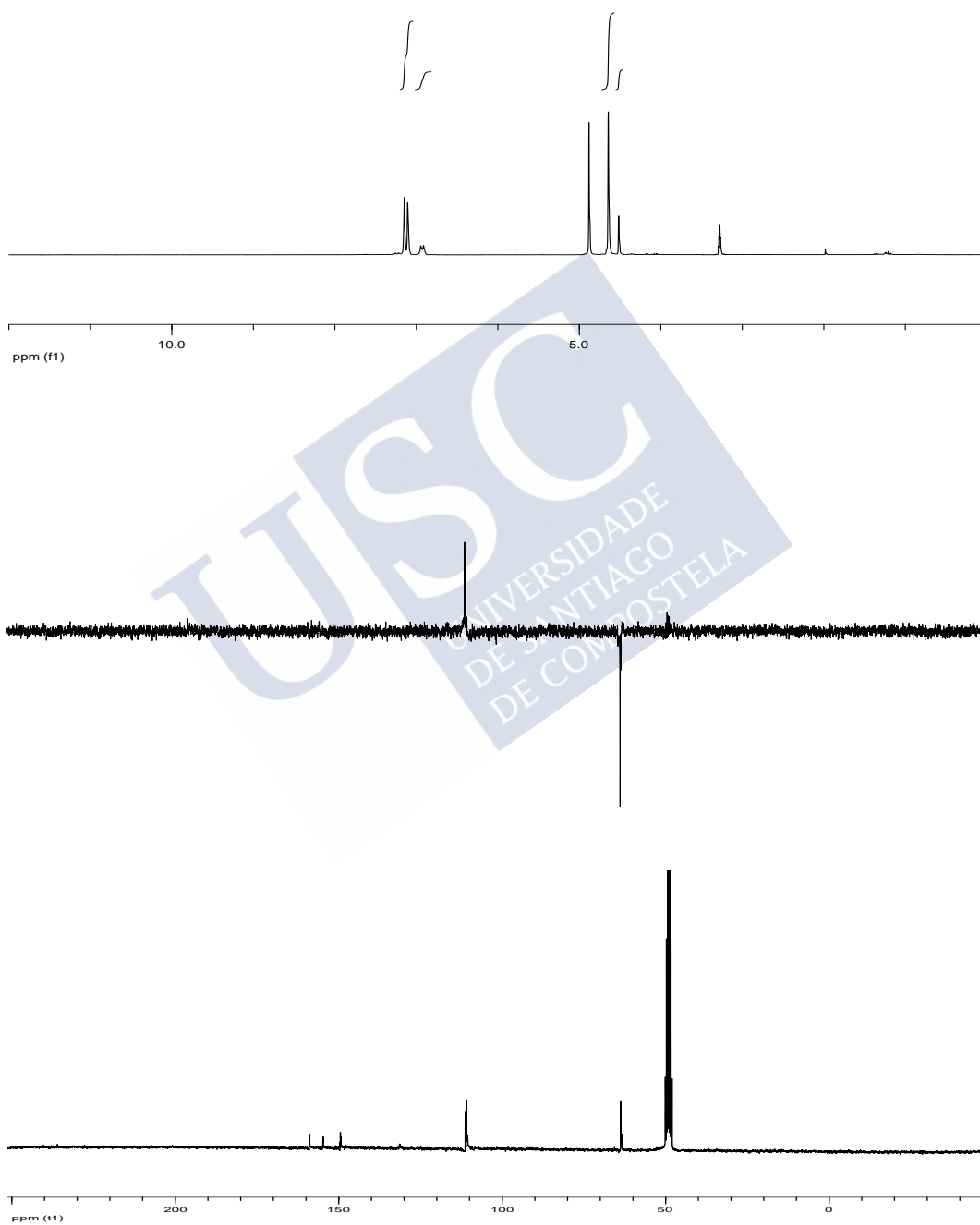
**$^{13}\text{C}$  NMR, DEPT** (63 MHz,  $\text{CD}_3\text{OD}$   $\delta$ ): 158.9 (C), 154.8 (C), 149.5 (C), 110.9 (CH), 63.70 (CH<sub>2</sub>)

**HRMS (ESI<sup>+</sup>-MS)**  $[\text{M}+\text{H}]$  calc. for  $\text{C}_{14}\text{H}_{11}\text{F}_4\text{N}_2\text{O}_2 = 315.0757$ . Found 315.0755  $\text{C}_{14}\text{H}_{10}\text{F}_4\text{N}_2\text{O}_2$  (M.W.: 314.2396).

(22)

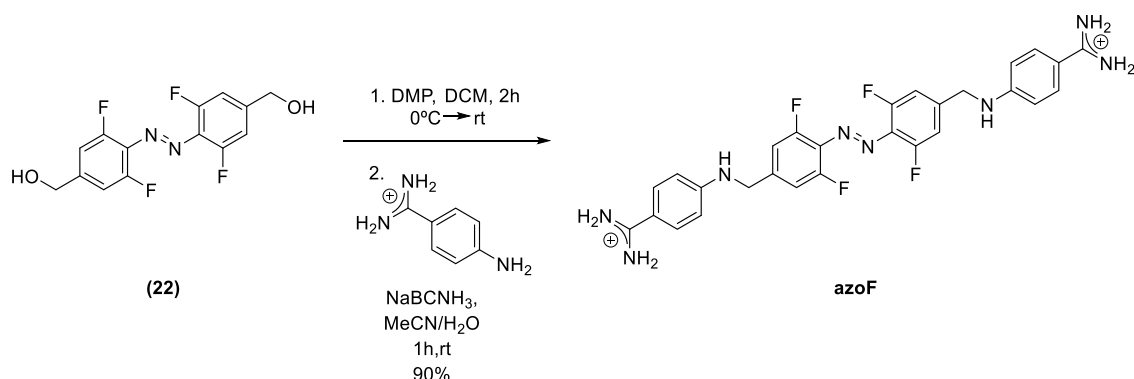


(22)





### Synthesis of azoF



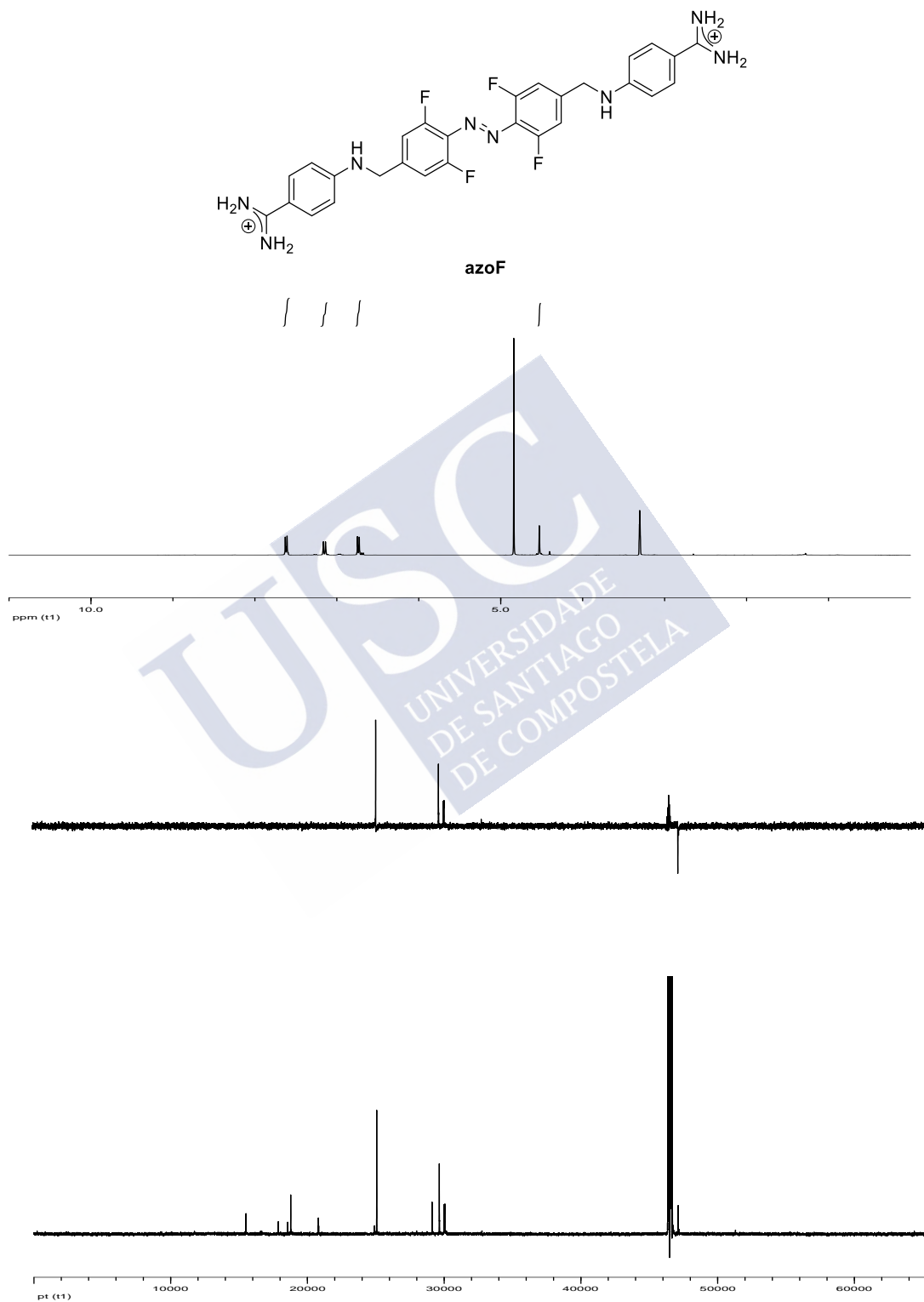
The solid (**22**) (105.4 mg, 0.34 mmol) was dissolved in  $\text{CH}_2\text{Cl}_2$  (10 mL) and poured onto a cooling bath of ice-water. DMP (313 mg, 0.74 mmol) was added dropwise. The mixture was then allowed to reach rt and was stirred for 2 h. The reaction was quenched with  $\text{Na}_2\text{S}_2\text{O}_3$  1 M (5 mL) and stirred for 30 more minutes. After that, brine solution (20 mL) was added and then extracted with EtOAc ( $3 \times 60$  mL). The combined organic extracts were dried with anhydrous  $\text{Na}_2\text{SO}_4$  and concentrated under reduced pressure. The residue was dissolved in MeCN/ $\text{H}_2\text{O}$  (1:3) (8 mL) and NaBCNH<sub>3</sub> (20.32 mg, 0.32 mmol) and aminobenzamidine (141 mg, 0.68 mmol) were added. The resulting solution was stirred for 1 h at rt. The solution was directly purified by RP-HPLC. The appropriate fractions were lyophilized, and an orange powder was identified as the trifluoroacetic salt of the **azoF** derivative (**14**) (238 mg, 0.31 mmol, 90 %).

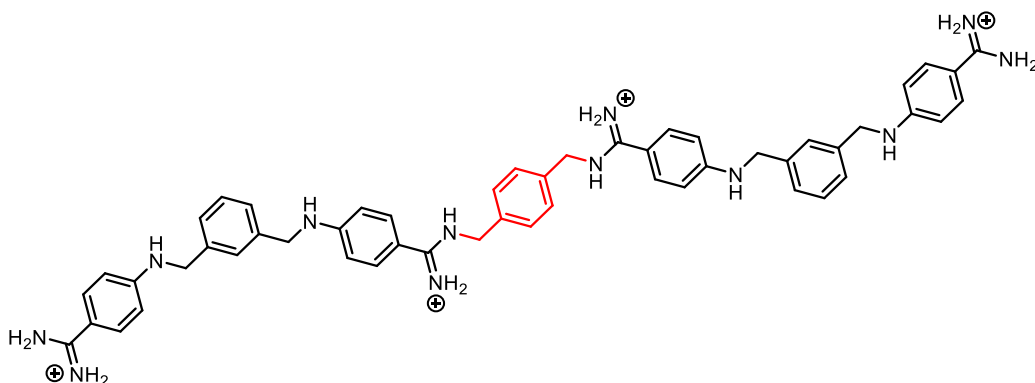
**E-isomer** <sup>1</sup>H NMR (400 MHz,  $\text{CD}_3\text{OD}$   $\delta$ ) 7.62 (d,  $J = 8.85$  Hz, 1H), 7.15 (d,  $J = 10.27$  Hz, 1H), 6.74 (d,  $J = 8.87$  Hz, 1H), 4.53 (s, 1H).

<sup>13</sup>C NMR, DEPT (101 MHz,  $\text{CD}_3\text{OD}$   $\delta$ ): 155.71 (C), 154.86 (C), 147.14 (C), 130.87 (C), 131.50 (C), 115.43 (C), 113.41 (C), 112.10 (C), 111.85 (C), 46.78 ( $\text{CH}_2$ ).

**HRMS (ESI<sup>+</sup>):**  $[\text{M}+\text{Na}^+]$  calc. for  $\text{C}_{28}\text{H}_{26}\text{F}_4\text{N}_8\text{Na} = 573,2260$ . Found 573,2262;  $\text{C}_{32}\text{H}_{26}\text{F}_{10}\text{N}_8\text{O}_4$  (M.W. 778,6120).

## AzoF



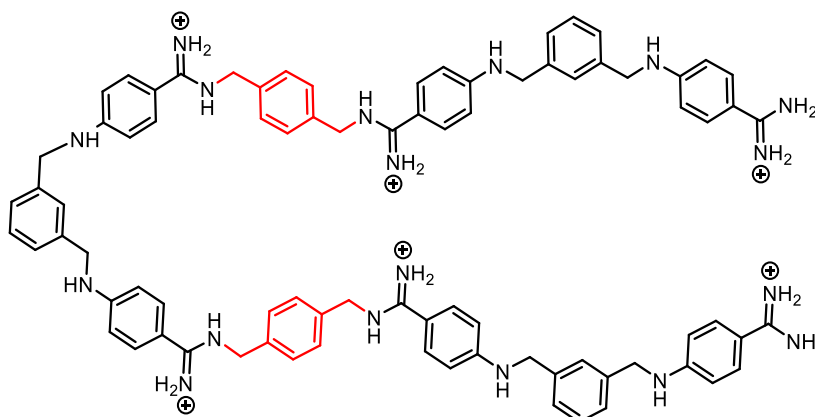
**Dimer**

**$^1\text{H}$  NMR** (400 MHz,  $\text{CD}_3\text{OD}$   $\delta$ ) 7.62 (d,  $J = 8.85$  Hz, 1H), 7.15 (d,  $J = 10.27$  Hz, 1H), 6.74 (d,  $J = 8.87$  Hz, 1H), 4.53 (s, 1H) (250 MHz,  $\text{CD}_3\text{OD}$ ): 7.13 (d,  $J = 10.49$  Hz, 4H), 4.64 (s, 4H).

**$^{13}\text{C}$  NMR, DEPT** (101 MHz,  $\text{CD}_3\text{OD}$   $\delta$ ): 155.71 (C), 154.86 (C), 147.14 (C), 130.87 (C), 131.50 (C), 115.43 (CH), 113.41 (C), 112.10 (CH), 111.85 (C), 46.78 ( $\text{CH}_2$ ).

**HRMS (ESI $^+$ )**: M.W. calc. for  $\text{C}_{52}\text{H}_{54}\text{N}_{12}$  = 846.46. Found 424.23 [ $\text{M}+2\text{H}$ ] $^{2+}$ , 283.15 [ $\text{M}+3\text{H}$ ] $^{3+}$ .

## Trimer



**$^1\text{H NMR}$**  (500 MHz,  $\text{CD}_3\text{OD}$   $\delta$ ) 7.58 (d,  $J = 9.1$  Hz, 2H), 7.56 – 7.52 (m, 4H), 7.44 (s, 4H), 7.37 (s, 2H), 7.31 – 7.23 (m,  $J = 12.9, 4.0$  Hz, 6H), 6.74 – 6.69 (m, 6H), 4.65 (s, 4H), 4.41 (s, 2H), 4.40 (s, 4H).

**$^{13}\text{C NMR, DEPT}$**  (126 MHz,  $\text{CD}_3\text{OD}$   $\delta$ ) 168.19 (C), 158.38 (C), 158.10 (C), 143.65 (C), 139.52 (CH), 133.59 (CH), 133.48 (CH), 132.86 (CH), 132.04 (CH), 130.09 (CH), 118.49 (C), 116.16 (CH), 50.57 ( $\text{CH}_2$ ), 49.64 ( $\text{CH}_2$ ).

**HRMS (ESI<sup>+</sup>):** calculated for  $\text{C}_{82}\text{H}_{84}\text{N}_{18}$  1320.7 found 441.5 [ $\text{M}+3\text{H}$ ]<sup>3+</sup>, 331.4 [ $\text{M}+4\text{H}$ ]<sup>4+</sup>, 265.3 [ $\text{M}+5\text{H}$ ]<sup>5+</sup>, 221.3 [ $\text{M}+6\text{H}$ ]<sup>6+</sup>.

**Trimer**

