INTERNATIONAL DOCTORAL
SCHOOL OF THE USC

Ana
Alcalde Ordóñez

PhD Thesis

Selective recognition of three-way DNA junctions with designed peptidomimetics

Santiago de Compostela, 2023

Doctoral Programme in Chemical Science and Technology

TESIS DE DOCTORADO

SELECTIVE RECOGNITION OF THREE-WAY DNA JUNCTIONS WITH DESIGNED PEPTIDOMIMETICS

Ana Alcalde Ordóñez

ESCUELA DE DOCTORADO INTERNACIONAL DE LA UNIVERSIDAD DE SANTIAGO DE COMPOSTELA

PROGRAMA DE DOCTORADO EN CIENCIA Y TECNOLOGÍA QUÍMICA

SANTIAGO DE COMPOSTELA

2023

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

DECLARACIÓN DEL AUTOR/A DE LA TESIS

D./Dña. **Ana Alcalde Ordóñez**

Título de la tesis: **Selective recognition of three-way DNA junctions with designed peptidomimetics**

Presento mi tesis, siguiendo el procedimiento adecuado al Reglamento y declaro que:

1) La tesis abarca los resultados de la elaboración de mi trabajo.
2) De ser el caso, en la tesis se hace referencia a las colaboraciones que tuvo este trabajo.
3) Confirmo que la tesis no incurre en ningún tipo de plagio de otros autores ni de trabajos presentados por mí para la obtención de otros títulos.
4) La tesis es la versión definitiva presentada para su defensa y coincide la versión impresa con la presentada en formato electrónico.

Y me comprometo a presentar el Compromiso Documental de Supervisión en el caso que el original no esté depositado en la Escuela.

En **Santiago de Compostela**, **06 de Junio de 2023**.

Firma electrónica

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

AUTORIZACIÓN DEL DIRECTOR / TUTOR DE LA TESIS

Selective recognition of three-way DNA junctions with designed peptidomimetics

D./Dª. M. Eugenio Vázquez Sentís
D./Dª. Miguel Vázquez López

INFORMA/N:

Que la presente tesis, se corresponde con el trabajo realizado por D/Dª. Ana Alcalde Ordóñez, bajo mi dirección/tutorización, y autorizo su presentación, considerando que reúne los requisitos exigidos en el Reglamento de Estudios de Doctorado de la USC, y que como director de esta no incurre en las causas de abstención establecidas en la Ley 40/2015.

De acuerdo con lo indicado en el Reglamento de Estudios de Doctorado, declara también que la presente tesis doctoral es idónea para ser defendida en base a la modalidad de Monográfica con reproducción de publicaciones, en los que la participación del doctorando/a fue decisiva para su elaboración y las publicaciones se ajustan al Plan de Investigación.

En Santiago de Compostela, 06 de Junio de 2023

Fdo.: M. Eugenio Vázquez Sentís

Fdo.: Miguel Vázquez López

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

A meus pais e meus tíos.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Table of contents

Abbreviations ..... **iv**
Resumo **1**
Summary **7**
Introduction **14**
1. Chemical biology 14
2. DNA: basic concepts 16
2.1. Biological functions 16
2.1.1. Central Dogma of Biology 16
2.1.2. Specific functions of DNA 16
2.2. Structures of DNA 17
2.2.1. DNA three-way junction (3WJs) 18
3. Amino acids, peptides and proteins 21
3.1. Short introduction 21
3.2. Structure of proteins 21
3.2.1. Secondary structure motifs: α-helix, β-sheet and β-turn 23
3.3. Metalloproteins 27
3.3.1. Synthetic metallopeptides 30
4. Peptide helicates 47
4.1. Helicates 47
4.1.1. Structural characteristics of helicates 47
4.1.2. Fields of application 50
4.2. Potential of helicates in DNA recognition 51
4.2.1. DNA binders and their potential in therapy and diagnosis 51
4.2.2. Helicates as DNA binders 53
General Objectives **59**
Chapter I **61**
Design and Synthesis of Cu(II) Peptide Helicates that Selective Cleavage DNA Replication Foci in Mammalian Cells 61
Summary 61
Introduction 62
Objetive 65
Results and discussion 65
Conclusions 75
Chapter II **77**
Chiral parallel peptide helicates with trigonal symmetry 77
Summary 77

UCC UNIVERSIDADE

Introduction 78
Results and discussion 79
Conclusions 89
Chapter III 91
Efficient and Selective Cleave of 3WJ by ATCUN-Functionalized Peptide Helicates . 91
Summary 91
Introduction 92
Objetive 94
Results and discussion 94
Conclusions 103
Chapter IV 105
Selective recognition of A/T-rich DNA 3-way junctions with a three-fold symmetric tripeptide 105
Summary 105
Introduction 106
Objetive 108
Results and discussion 108
Conclusions 115
General Conclusions 117
Experimental Section 119
1. Reactives 119
2. Procedures and experimental techniques 119
Chapter I 131
Chapter II 140
Chapter III 144
Chapter IV 149
List of publications 157

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Abbreviations

5Bpy	5'-amino-2,2'-bipyridyl-5-carboxylic acid
3WJ	three-way DNA junction
βAla	β-Alanine
δ	chemical shift (ppm)
Θ	molar ellipticity
°C	celsius degree
aa	amino acid
Ac	acetyl
Ac_2O	acetic anhydride
Alloc	allyloxycarbonyl
DNA	deoxyribonucleic acid
RNA	ribonucleic acid
Boc	tert-butyloxycarbonyl
Bpy	2,2'-bipyridine
CD	circular dichroism
c.p.s	counts per second
d	doublet (in NMR spectra)
dd	double doublet (in NMR spectra)
DCM	diclorometane
DIC	*N,N'*-Diisopropylcarbodiimide
DIEA	*N,N'*-diisopropylethylamine
DMF	*N,N'*-dimethylformamide
DMSO	dimethyl sulfoxide
dsDNA	double-stranded DNA
EPR	electron paramagnetic resonance
eq	equivalent
ESI	Electro Spray Ionization
Fmoc	9-fluorenylmethoxycarbonile
HATU	1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo [4,5-b]pyridinium 3-oxide hexafluorophosphate
HBTU	(2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

	hexafluorophosphate
HEPES	4-(2-hydroxyethyll)-1-piperazineethanesulfonic acid
HPLC	High Performance Liquid Chromatography
J	coupling constant (Hz)
K_D	dissociation constant
m	multiplet (in NMR spectra)
MALDI	Matrix Assisted Laser Desorption Ionization
MeOH	methanol
mdeg	millidegrees
min	minute
MQ	milli-Q water
MS	mass spectrometry
MW	molecular weight
NMR	Nuclear Magnetic Resonance
PBS	Phosphate Buffer Saline
ppm	parts per million
q	quartet (in NMR spectra)
RP	Reverse Phase (in HPLC chromatography)
s	singlet (in NMR spectra)
SPPS	Solid-Phase Peptide Synthesis
t	triplet (in NMR spectra)
TOF	Time Of Flight
t_R	retention time
rt	room temperature
TFA	trifluoroacetic acid
TIS	triisopropylsilane
UV-Vis	Ultraviolet-Visible

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Resumo

Capítulo I. Deseño e síntese de helicatos peptídicos de Cu(II) que escinden selectivamente os focos de replicación de ADN en células mamíferas

Neste capítulo discútese a obtención estereoselectiva de helicatos peptídicos dinucleares de Cu(II), mediante o pregamento dirixido polos ións metálicos de péptidos lineais, equipados con residuos coordinantes derivados de 2,2'-bipiridina, que se conectan a través de secuencias heteroquirais promotoras de xiros β de tipo II e II'. Estes novos metalopéptidos de Cu(II) presentan actividade nucleasa selectiva polas estruturas 3WJ.

Os ligandos precursores de helicatos **LLD** e o seu enantiómero **DDL**, previamente publicados, sintetizáronse seguindo a metodoloxía clásica de síntese de péptidos en fase sólida, incorporando seis unidades coordinantes de βAla5Bpy. Os péptidos finais purificáronse mediante HPLC en fase reversa, e a súas propiedades coordinativas foron estudadas monitorizando a desactivación na intensidade de emisión de fluorescencia do residuo de Bpy, a consecuencia da súa coordinación aos catións Cu(II). Os resultados obtidos axustáronse a un modelo de unión 1:2 (**LLD**:Cu^{II}) con unhas constantes de disociación aproximadas a 0.3 e 0.15 μM, correspondentes coa entrada do primeiro e o segundo ión Cu(II), respectivamente. Adicionalmente leváronse a cabo estudos de dicroísmo circular para determinar a quiralidade final dos helicatos. Cabe destacar que a adición de Cu(II) sobre unha disolución de **LLD** ou **DDL** xenera a unha inversión no efecto *Cotton* da banda de 2,2'-bipiridina centrada a 320 nm, consistente coa formación dos helicatos **ΔΔ-Cu^{II}_2LLD** e **ΛΛ-Cu^{II}_2DDL**. Esta inversión na configuración parece estar asociada a unha distorsión *Jahn-Teller*, característica dos complexos de Cu(II) d^9 octaédricos.

Unha vez se caracterizaron os helicatos de Cu(II), procedeuse a estudar as súas propiedades de recoñecemento de 3WJs, rexistrando a emisión dun 3WJ marcado con fluoresceína (Flu-3WJ) tras adicións consecutivas dos helicatos **ΔΔ-Cu^{II}_2LLD** e **ΛΛ-Cu^{II}_2DDL**. Os perfís obtidos axustáronse a un modelo de unión 1:1 (Flu-3WJ/helicato), que inclúe interaccións non específicas, cunhas constantes de disociación aparentes K_D = 260 nM para **ΔΔ-Cu^{II}_2LLD** e K_D = 654 nM para **ΛΛ-Cu^{II}_2DDL**. Adicionalmente, comprobouse que a afinidade de **ΔΔ-Cu^{II}_2LLD** polas estruturas 3WJ non está limitada a unha soa secuencia, posto que seguindo a mesma metodoloxía, se calculou a K_D deste helicato peptídico de Cu(II) con un 3WJ de diferente secuencia, Flu-GC-3WJ, resultando nun valor aproximado a 300 nM, moi similar ao previamente obtido para Flu-3WJ.

Para analizar en maior profundidade as propiedades de unión ao ADN do helicato **ΔΔ-Cu^{II}_2LLD**, leváronse a cabos ensaios de electroforese en xel de poliacrilamida (PAGE). Nestes experimentos observouse a aparición dunha nova banda retardada asociada á formación do complexo 3WJ/**ΔΔ-Cu^{II}_2LLD**, mentres que non se mostra ningunha modificación na banda correspondente ao ADN canónico, o que confirma a baixa afinidade deste helicato peptídico de Cu(II) polo ADN-B.

Adicionalmente, realizáronse diferentes experimentos de electroforese co obxectivo de analizar o perfil de ruptura de ADN de **ΔΔ-Cu^{II}_2LLD**. A análise de PAGE

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

en presenza de Na-*L*-ascorbato revelou que o metalopéptido de Cu(II) rompe selectivamente estruturas de ADN 3WJ, como se puido deducir a partir da diminución na intensidade da banda asociada á formación do complexo 3WJ/**ΔΔ-Cu$^{II}_2$LLD** en presenza de cantidades crecentes de helicato de Cu(II). Porén, non se observaron cambios significativos na banda asociada ao ADN-B regular, confirmándose de novo a alta selectividade deste helicato polas 3WJs. A modo de control, este mesmo ensaio repetiuse en presenza do helicato **ΛΛ-Fe$^{II}_2$LLD**, que segundo o esperado, non presentou actividade nucleasa, malia unirse selectivamente ao ADN 3WJ.

De seguido, confirmouse que o mecanismo responsable da actividade nucleasa de **ΔΔ-Cu$^{II}_2$LLD** está mediado por ROS, mediante experimentos de electroforese nos que se observou que a presenza de *scavengers* selectivos de ROS, como *tiron*, poden inhibir a ruptura de ADN.

Finalmente, caracterizouse a capacidade do helicato **ΔΔ-Cu$^{II}_2$LLD** para recoñecer e romper selectivamente estruturas 3WJ en células Vero funcionais, previamente permeabilizadas con dixitonina. Observouse que o helicato peptídico de Cu(II) marcado con TAMRA co-localiza con focos de replicación de ADN ricos en estruturas 3WJ, que foron previamente marcados. Isto confirmou o potencial deste novo metalocilindro de Cu(II) para asociarse a 3WJs do ADN celular. Para rematar, a súa capacidade para levar a cabo acción nucleasa selectiva de 3WJs en células funcionais, incubadas en presenza de H_2O_2, confirmouse mediante un ensaio *TUNEL*, no cal as células tratadas con **ΔΔ-Cu$^{II}_2$LLD**, presentaron maior emisión de fluorescencia, o que se traduce nun maior número de grupos fosfato de ADN libres, e polo tanto, unha maior actividade nucleasa.

Capítulo II. Helicatos peptídicos paralelos e quirais con simetría trigonal

Neste capítulo demostrouse que un ligando peptídico con simetría C_3, equipado con unidades coordinantes derivadas de 2,2'-bipiridina, pode comportarse como un sistema dinámico que en presencia de Co(II) se auto-ensambla en helicatos peptídicos enantiomericamente puros con propiedades de unión ao ADN. Pola contra, o ligando libre agrega en fibras supramoleculares en ausencia de metal. En suma, os helicatos peptídicos de Co(II), de natureza dinámica, puideron ser oxidados para dar o correspondente metalopéptido de Co(III) cineticamente inerte. Os helicatos de Co(II) e Co(III) demostraron unión selectiva a 3WJs fronte a un ADN canónico de dobre cadea.

O novo ligando **L** está formado por tres cadeas peptídicas ancladas a un bloque central de benceno tri-substituído, cada unha delas incorporando dous residuos βAla5Bpy e un residuo de L-Arg, que controlan termodinamicamente a formación enantioselectiva dos correspondentes helicatos. As cadeas peptídicas sintetizáronse seguindo a metodoloxía clásica de síntese de péptidos en fase sólida, e fixéronse reaccionar en resina con 1,3,5-tris(bromometil)benceno para dar o ligando **L** de simetría trigonal. O péptido final **L** purificouse mediante HPLC en fase reversa, e caracterizáronse as súas propiedades coordinativas. Para isto, explotouse de novo a diminución na emisión de fluorescencia da 2,2'-bipiridina exercida polos ións Co(II) ao coordinarse. O perfil resultante axustouse a un modelo de unión 1:2 (L:Co^{II}), cunhas constantes de disociación sorprendentemente débiles en comparación coas reportadas para sistemas similares, apuntando a que **L** podería estar agregado. Esta hipótese foi apoiada pola diferenza de

USC DE SANTIAGO

intensidade observada nos espectros de CD de dúas stocks de **L** a diferente concentración, de novo suxerindo un diferente estado de agregación do péptido.

Para conseguir medidas reproducibles, preparouse unha nova *stock* de **L** en DMSO que se empregou para estudar de novo a formación do helicato de Co(II) correspondente, seguindo a mesma metodoloxía baseada en monitorizar a desactivación da fluorescencia do residuo de Bpy por coordinación aos ións metálicos. O novo perfil puido axustarse a un modelo 1:1 (L:Co^{II}_2) cunha constante de disociación aparente $K_D \approx$ 1.32 μM, consistente cos sistemas descritos.

Decidimos afondar na capacidade de agregación do péptido **L** realizando diferentes estudos usando a nova *stock* de ligando en DMSO. Os espectros de CD de **L** a 10 μM medidos incrementando a temperatura de 0 a 95 ºC mostraron unha diminución na intensidade da banda a 320 nm, consistente cun proceso de desagregación do péptido, ao que tamén apunta a transición observada ao redor de 60 ºC na sinal do *melting* medido a 320 nm. Nos estudos de UV-Vis do ligando **L** libre, aparece un ombro na banda de absorbancia da Bpy centrada a 308 nm, que se asocia ao agregado supramolecular, e como esperábamos, desaparece ao incrementar a temperatura, así como en presenza de Co(II), confirmando que os ións metálicos favorecen a formación de helicatos discretos. A quiralidade destes helicatos peptídicos caracterizouse mediante estudos de dicroísmo circular, cuns resultados consistentes coa formación do metalopéptido enantiomericamente puro **ΛΛ-Co^{II}_2L**, como indicou a banda de efecto Cotton positivo da 2,2'-bipiridina centrada a 320 nm.

Para caracterizar a natureza dos agregados supramoleculares realizáronse estudos de microscopía e DLS, a partir de disolucións do ligando **L** libre, e en presenza de 15 eq. de ións Co(II), empregando auga e DMSO como disolventes. Os resultados de DLS parecen indicar que o ligando **L** agrega en disolución acuosa en ausencia de metal, e forma complexos discretos con Co(II), polo contra, en DMSO non se observa a formación de agregados. Porén, a polidispersión nas medidas de DLS é demasiado alta e non permite obter resultados precisos. Nas imaxes de microscopía electrónica STEM e TEM, confirmouse a formación de fibras helicoidais nunha disolución acuosa do ligando peptídico **L** en condicións libres de metal. Adicionalmente, as imaxes de TEM demostraron que en presenza de Co(II) se inhibe o proceso de formación destas fibras, fomentando a síntese de helicatos peptídicos discretos. De igual modo, nas imaxes obtidas para as disolucións de control en DMSO, non se observou a formación de fibras por agregación de **L,** tanto en ausencia como en presenza de ións Co(II). Estes resultados, apoiaron a nosa hipótese inicial, derivada dos resultados dos ensaios previos de fluorescencia, CD e UV-Vis.

Unha vez se confirmou a formación do helicato peptídico **ΛΛ-Co^{II}_2L**, ensamblado baixo control termodinámico, este oxidouse empregando $(NH_4)_2Ce(NO_3)_6$, dando lugar á obtención do metalopéptido dinuclear cineticamente inerte de Co(III), que puido ser caracterizado mediante espectrometría de masas e dicroísmo circular. O espectro de CD de **Co^{III}_2L**, confirmou a retención na configuración do centro metálico no proceso de oxidación, e polo tanto, a formación do helicato **ΛΛ-Co^{III}_2L**.

USC UNIVERSIDADE DE SANTIAGO

Posteriormente, procedeuse a estudar as propiedades de unión a 3WJs de **ΛΛ-Co^{II}_2L** e **ΛΛ-Co^{III}_2L**, monitorizando a diminución na intensidade da banda de emisión dun 3WJ marcado con fluoresceína tras adición de cantidades crecentes dos helicatos

peptídicos de Co(II) e Co(III). O perfil resultante para o helicato de Co(II) axustouse a un modelo de unión 1:1 (Flu-3WJ/helicato), cunha constante de disociación aparente K_D de 359 nM, mentres que con Co(III), o perfil da valoración axustouse a un modelo máis complexo, 1:1+1:3, con K_{D1} = 0.128 μM e K_{D2} = 0.825 μM, respectivamente. Desafortunadamente, a constante de disociación dos helicatos cun ADN de dobre cadea non puido ser calculada debido a problemas de precipitación, que suxiren interaccións non específicas entre os helicatos e o dsDNA.

Para concluír, o modo de interacción entre o ADN e os metalopéptidos de Co(II) e Co(III) caracterizouse mediante ensaios de electroforese en xel de poliacrilamida. A análise do xel confirmou que **ΛΛ-Co^{II}_2L** e **ΛΛ-Co^{III}_2L** presentan unión selectiva e de alta afinidade ao 3WJ, como indicou a aparición dunha nova banda retardada en presenza desta estrutura non canónica con cantidades crecentes de ambos helicatos, que se asocia á formación dos complexos 3WJ/**ΛΛ-Co^{II}_2L** e 3WJ/**ΛΛ-Co^{III}_2L**, que pola contra, non se observa na incubación cun ADN-B regular de dobre cadea. Este resultado pon de manifesto a baixa afinidade de ambos metalopéptidos polo dsDNA.

Capítulo III. Ruptura eficiente e selectiva de 3WJ por helicatos peptídicos funcionalizados con ATCUN

Neste capítulo descríbese a síntese de helicatos peptícos quirais de Cu(II) e Ni(II), equipados cun novo ligando derivado do péptido **LLD**, descrito no Capítulo I. **LLD** funcionalizouse no seu extremo amino co motivo tripeptídico ATCUN, que se caracteriza por formar complexos plano-cadrados de alta afinidade cos ións Cu(II) e Ni(II), que presentan propiedades nucleasa de ADN. Os metalopéptidos de Cu(II) e Ni(II) derivados deste novo ligando demostraron unión selectiva a estruturas 3WJ. En particular, o helicato peptídico de Cu(II) presentou actividade nucleasa altamente selectiva polo ADN 3WJ en comparación cun ADN-B regular.

O novo ligando precursor de helicatos **RGH-LLD**, combina a secuencia de motivo ATCUN **L-Arg–Gly–L-His** (RGH), coa secuencia do ligando **LLD**, conectadas a través de tres residuos de lisina acoplados mediante as súas cadeas laterais. **RGH-LLD** sintetizouse seguindo a metodoloxía estándar de SPPS e purificouse por HPLC en fase reversa.

Unha vez se illou o péptido, as súas propiedades coordinativas caracterizáronse en presenza de ións metálicos Cu(II) e Ni(II) de forma análoga aos capítulos anteriores, medindo a diminución na intensidade de emisión dos residuos de Bpy tras a adición dos correspondentes catións metálicos Posto que esta estratexia só contempla a coordinación ao fragmento LLD, asumimos que o primeiro equivalente de M(II) con respecto ao ligando peptídico se une ao motivo ATCUN, que presenta constantes de disociación reportadas de orde picomolar. O perfil resultante para Cu(II) axustouse a un modelo de unión 1:1 con constante de disociación aparente en torno a 0.1 μM para os dous ións metálicos, mentres que para Ni(II), o perfil se axustou a un modelo 1:2 con K_{D1} = 0.00208 μM e K_{D2} = 0.333 μM. Estes estudos, xunto aos resultados de espectrometría de masas MALDI suxiren a formación de metalopéptidos cunha relación global 1:3 (**RGH-LLD**:Cu^{II}/Ni^{II}).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Os estudos de dicroísmo circular confirmaron a obtención de helicatos quirais, que no caso particular do metalopéptido de Cu(II), ao igual que se observou no Capítulo

I, presentou unha configuración invertida consistente coa obtención do helicato **ΔΔ-Cu$^{II}_3$RGH-LLD**, posiblemente consecuencia dunha distorsión *Jahn-Teller*. O Ni(II), pola súa banda, conduce á formación do helicato diasteroméricо **ΛΛ-Ni$^{II}_3$RGH-LLD**, como indicou a banda de efecto *Cotton* positivo da Bpy a 320 nm

Unha vez se caracterizaron os helicatos de Cu(II) e Ni(II), estudáronse as súas propiedades de unión a 3WJs, rexistrando a emisión dun 3WJ marcado con fluoresceína (Flu-3WJ) tras adicións consecutivas dos helicatos **ΔΔ-Cu$^{II}_3$RGH-LLD** e **ΛΛ-Ni$^{II}_3$RGH-LLD**. Os perfís obtidos axustáronse a un modelo de unión 1:1 (Flu-3WJ/helicato), que inclúe interaccións non-específicas, cunhas constantes de disociación aparentes de 170 nM para **ΔΔ-Cu$^{II}_3$RGH-LLD** e 140 nM para **ΛΛ-Ni$^{II}_3$RGH-LLD**. Ao igual que no Capítulo I, calculouse a constante de disociación destes metalopéptidos a 3WJs de diferente secuencia, obténdose uns valores similares que confirman que os helicatos de Cu(II) e Ni(II) recoñecen as estruturas 3WJs nun modo independente da súa secuencia. A K_D para a unión a unha estrutura de ADN-B regular non puido ser calculada debido á precipitación da mostra durante a realización do experimento, suxerindo a posibilidade de interaccións non específicas entre os metalocilindros e o dsDNA.

De seguido, a capacidade de **ΔΔ-Cu$^{II}_3$RGH-LLD** e **ΛΛ-Ni$^{II}_3$RGH-LLD** para romper selectivamente 3WJs *in vitro* estudouse mediante ensaios de electroforese en xel de poliacrilamida. A análise de PAGE revelou que o metalopéptido de Cu(II) une e rompe selectivamente estruturas de ADN 3WJ en presenza de H_2O_2, como se pode deducir a partir da diminución na intensidade da banda asociada á formación do complexo 3WJ/**ΔΔ-Cu$^{II}_3$RGH-LLD**. Esta actividade nucleasa non se observou na banda asociada ao ADN-B regular, así como tampouco en presenza do helicato de Ni(II). **ΛΛ-Ni$^{II}_3$RGH-LLD** presentou un modo de actuación diferente, uníndose con alta selectividade ao 3WJ, pero sen levar a cabo actividade nucleasa. En ambos casos, a banda asociada ao ADN de dobre cadea permaneceu invariable, confirmando a baixa afinidade destes helicatos peptídicos quimera pola estrutura canónica do ADN.

Finalmente, caracterizouse a capacidade dos helicatos **ΛΛ-Ni$^{II}_3$RGH-LLD** e **ΔΔ-Cu$^{II}_3$RGH-LLD** para recoñecer selectivamente estruturas 3WJ en células HeLa funcionais, previamente permeabilizadas con dixitonina. Observouse que estes metalopéptidos de Ni(II) e Cu(II) marcados con TAMRA colocalizan con focos de replicación de ADN marcados en verde, ricos en estruturas 3WJ, confirmando o potencial destes novos sistemas para asociarse selectivamente a 3WJs en células.

Capítulo IV. Recoñecemento selectivo de 3WJs ricos en A/T cun péptido trimérico simétrico

Neste capítulo, descríbese una nova estratexia de recoñecemento de 3WJs, que está baseada na unión cooperativa e con selectividade de secuencia aos tres brazos de ADN dúplex que constitúen as estruturas 3WJ. A eficiencia desta metodoloxía demostrouse mediante o deseño dun novo sistema peptídico, con simetría trigonal, formado por tres cadeas de péptidos AT-Hook acopladas a un corazón de 1,3,5-tristirilbenceno. Este novo ligando presentou propiedades de unión selectiva a 3WJs ricos en secuencias de A-T, que se caracterizaron explotando o aumento na intensidade de emisión de fluorescencia do núcleo central de benceno, derivado da restrición conformacional causada pola súa unión ao ADN.

USC DE SANTIAGO

Resumo

Para o deseño do novo ligado trimérico AT-Hook, construíuse un modelo computacional hipotético formado por tres péptidos AT-Hook, insertados simultaneamente no suco menor dos tres brazos de ADN dúplex no 3WJ, e un 1,3,5-tristirilbenceno situado sobre a cavidade central desta estrutura do ADN. A distancia de 4 Å entre os extremos amino dos péptidos AT-Hook e os grupos carboxilato no núcleo de benceno, suxeriu que un residuo de lisina acoplado a través da súa cadea lateral encaixaría perfectamente como residuo de unión. Este modelo hipotético foi investigado en profundidade mediante *molecular modelling* empregando a plataforma *GaudiMM*, que confirmou un excelente axuste entre o modelo deseñado e unha estrutura 3WJ.

O péptido AT-Hook sintetizouse seguindo a metodoloxía estándar de SPPS asistida por microondas, e as cadeas foron simultaneamente acopladas ao corazón de 1,3,5-tristirilbenceno en resina, obténdose o péptido final que se purificou mediante HPLC en fase reversa.

Unha vez se confirmou a obtención do péptido AT-Hook trimérico **1**, estudamos a súa interacción co ADN, medindo o aumento na emisión de fluorescencia inherente ao núcleo 1,3,5-tristirilbenceno, a causa da restrición conformacional que implica a súa unión ao ADN. En función disto, rexistrouse a emisión de fluorescencia de **1**, tras adicións sucesivas de diferentes estruturas de ADN, TA-3WJ (rico en A-T), GC-3WJ (rico en G-C) como control negativo, e un dsDNA (rico en A-T). Para o caso do ADN diana TA-3WJ, o perfil resultante axustouse a un modelo mixto que combina a formación dun complexo 1:1 con K_D ao redor de 20 nM, coa contribución menor dun termo de unión non específico. O perfil da valoración en presenza de ADN de dobre cadea rico en A-T, axústase a este mesmo modelo, pero con peor afinidade de unión, o que se traduce nunha constante de disociación ata 10 veces máis débil. Como era de esperar, no perfil obtido para o ADN control GC-3WJ, non se observou aumento na emisión de fluorescencia. Cabe destacar que a adición de 1 eq. de TA-3WJ sobre esta disolución resultou nun aumento da intensidade, o que confirmou a alta selectividade de unión do AT-Hook trimérico **1** por 3WJs ricos en A/T.

A interacción de **1** co suco menor dos brazos de ADN dúplex en TA-3WJ, foi apoiada cos resultados de dicroísmo circular, onde se observou unha nova banda en presenza do péptido **1**. Porén o espectro de CD de GC-3WJ permaneceu invariable trala adición de **1**, confirmando que non existe unión de alta afinidade.

Experimentos adicionais de electroforese confirmaron a formación dun complexo de alta afinidade entre **1** e TA-3WJ, como indicou a aparición dunha nova banda retardada, así como a formación de asociacións de baixa afinidade con GC-3WJ e dsDNA.

Finalmente, avaliouse a capacidade do péptido trimérico **1** para identificar selectivamente estruturas 3WJs ricas en secuencias A/T en células HeLa funcionais, previamente permeabilizadas con dixitonina. Observouse que o trímero de AT-Hook **1**, que aparece na canle de emisión azul, mostraba unha forte correlación coa localización dos focos de replicación ricos en 3WJs, que foron previamente marcados en verde, confirmando o gran potencial deste novo sistema **1** para recoñecer e marcar estruturas 3WJ no ADN celular.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Summary

Chapter I. Design and synthesis of Cu(II) peptide helicates that selective cleavage DNA replication foci in mammalian cells

This chapter analyses the stereoselective synthesis of dinuclear Cu(II) peptide helicates by metal ion-directed folding of linear peptides sequences equipped with coordinating residues derived from 2,2'-bipyridine, which are linked through β-turns type II and II' promoter heterochiral sequences. These novel Cu(II) metallopeptides exhibited selective nuclease activity towards 3WJ structures.

The previously reported **LLD** and **DDL** peptide ligands were synthesized following conventional solid-phase peptide synthesis methodology, featuring six βAla5Bpy coordinating units. The final peptides were purified by RP-HPLC and their coordination properties were studied by monitoring the deactivation in the fluorescence emission intensity of the Bpy residue, as a consequence of the coordination to Cu(II) ions. The obtained profiles were fitted to a 1:2 (**LLD**:Cu^{II}) binding model, with dissociation constants around 0.3 and 0.15 μM, for the first and second Cu(II) ions, respectively. In addition, circular dichroism studies were conducted to determine the final chirality in the helicates. It is noteworthy that addition of Cu(II) to a **LLD** or **DDL** solution leads to an inversion in the Cotton effect band of the 2,2'-bipyridine residue centered at 320 nm, which is consistent with the formation of **ΔΔ-Cu^{II}_2LLD** and **ΛΛ-Cu^{II}_2DDL**. This inversion in the configuration seems to be associated with a Jahn-Teller distortion, which is characteristic of the d^9 octahedral Cu(II) complexes.

Once the Cu(II) helicates were characterized, we studied their 3WJ recognition properties by measuring the emission of a fluorescein-labelled 3WJ (Flu-3WJ) after addition of increasing amounts of **ΔΔ-Cu^{II}_2LLD** and **ΛΛ-Cu^{II}_2DDL**. The resulting profiles were fitted to a 1:1 (Flu-3WJ/helicate) binding model, plus non-specific interactions, with apparent dissociation constants of K_D = 260 nM for **ΔΔ-Cu^{II}_2LLD** and K_D = 654 nM for **ΛΛ-Cu^{II}_2DDL**. Additionally, it was found that the affinity of **ΔΔ-Cu^{II}_2LLD** for 3WJ structures is not limited to a single sequence, since following the same methodology, it was calculated the K_D of this Cu(II) peptide helicate with a 3WJ of different sequence, Flu-GC-3WJ, resulting in a value of approximately 300 nM, which is very similar to that previously calculated for Flu-3WJ.

To further analyze the DNA binding properties of the **ΔΔ-Cu^{II}_2LLD** helicate, electrophoresis assays were undertaken on polyacrylamide gel (PAGE). In these experiments, we observed the appearance of a new slow-migrating band consistent with the formation of the 3WJ/**ΔΔ-Cu^{II}_2LLD** complex, while no change was observed in the band associated to the dsDNA, thus confirming the low affinity of this Cu(II) peptide helicate for canonical B-DNA.

Moreover, different electrophoresis experiments were carried out to analyse the DNA cleavage profile of **ΔΔ-Cu^{II}_2LLD**. PAGE analysis performed in the presence of Na-(L)-ascorbate revealed that the Cu(II) metallopeptide selectively cleaves 3WJ DNA structures, as can be concluded from the decrease in the intensity of the band associated with the formation of the 3WJ/**ΔΔ-Cu^{II}_2LLD** complex in the presence of increasing amounts of Cu(II) helicate. However, no significant changes were observed in the band

associated with regular B-DNA, confirming once again the high selectivity of this helicate for 3WJs. As a control, the same assay was repeated in the presence of **ΛΛ-Fe^{II}_2LLD** helicate, which, as expected, did not exhibit nuclease activity, despite binding selectively to the 3WJ DNA.

Further, it was confirmed by electrophoresis experiments that the mechanism responsible for the nuclease activity of **ΔΔ-Cu^{II}_2LLD** is mediated by ROS, since the presence of selective ROS scavengers, such as tiron, inhibited this DNA cleavage process.

Finally, we tested the ability of **ΔΔ-Cu^{II}_2LLD** helicate to selectively recognize and cleave 3WJ structures in functional Vero cells, previously permeabilized with digitonin. The TAMRA-labelled Cu(II) peptide helicate was observed to colocalize with the green-labelled DNA replication foci, enriched in 3WJ structures, which were previously labelled. This result supported the potential of the novel Cu(II) metallocylinder for associating with 3WJ in cellular DNA. We also demonstrated the ability of **ΔΔ-Cu^{II}_2LLD** to carry out selective 3WJ nuclease activity in functional cells in the presence of H_2O_2, as confirmed by a *TUNEL* assay, in which those cells treated with **ΔΔ-Cu^{II}_2LLD**, showed higher fluorescence emission, which translates into a higher number of free DNA phosphate groups, and therefore, a greater nuclease activity.

Chapter II. Chiral parallel peptide helicates with trigonal symmetry

In this chapter it was shown that a peptide ligand with *C_3* symmetry, equipped with coordinating units derived from 2,2'-bipyridine, behaves as a dynamic system, which self-assembles into enantiomerically pure peptide helicates with DNA-binding properties in the presence of Co(II) ions, but in the absence of metal, this ligand aggregates into supramolecular fibers. Moreover, the Co(II) peptide helicate was successfully oxidized to give the corresponding kinetically inert Co(III) metallopeptide. Co(II) and Co(III) helicates exhibited selective binding towards 3WJs versus the canonical double-stranded DNA.

The new **L** ligand consists of three peptide strands attached to a central block of benzene, each of them bearing two βAla5Bpy units and one L-Arg residue, which thermodynamically control the enantioselective assembly of the corresponding helicates. The peptide chains were synthesized using standard Fmoc-SPPS protocols, and subsequently, they were reacted on resin with 1,3,5-tris(bromomethyl)benzene resulting in the **L** ligand with trigonal symmetry. The final peptide was purified by RP-HPLC and its coordination properties were characterized by measuring the quenching in the fluorescence emission of the 2,2'-bipyridine residue upon coordination to Co(II) ions. The resulting profile was fitted to a 1:2 (**L**:Co^{II}) binding model, with dissociation constants weaker than those reported in similar systems, suggesting that **L** might be aggregated. Our hypothesis was supported by the different intensity observed between the CD spectra of two stocks with **L** at different concentration, which appears to confirm a distinct degree of aggregation.

To obtain reproducible measurements, a new stock solution of **L** was prepared in DMSO and then used to investigate the formation of the corresponding Co(II) helicate, following the same methodology as before. The new profile could be fitted to a 1:1 (**L**:Co^{II}_2) model with an apparent dissociation constant of $K_D \approx 1.32$ μM, consistent with similar systems.

We decided to further investigate the aggregation behaviour of **L** by conducting different studies. The CD spectra of **L** at 10 μM was recorded by increasing the temperature from 0 to 95 ºC, showing a decrease in the band intensity at 320 nm, consistent with a disaggregation process, which is also pointed out by the transition observed around 60 ºC in the melting assay. UV-Vis studies of the free **L** ligand show a shoulder in the Bpy band at 308 nm, associated with a supramolecular aggregate, that disappears by heating up the sample, as well as in the presence of Co(II) ions. These results indicate that metal ions promote the formation of discrete helicates. The obtained results in the CD studies, were consistent with the formation of the enantiomerically pure **ΛΛ-Co^{II}_2L**, as indicated by the positive Cotton effect band at 320 nm.

The nature of the supramolecular aggregates was characterized by microscopy and DLS studies performed on solutions of the free **L** ligand and in the presence of 15 eq. of Co(II) ions, using water and DMSO as solvents. The DLS results appear to suggest that ligand aggregates in aqueous solution in the absence of metal, and it forms discrete complexes with Co(II), while no aggregate formation is observed in DMSO. However, the polydispersion in DLS measurements is too high to obtain accurate results. STEM and TEM electron microscopy images confirmed the formation of helical fibers in the aqueous solution of **L** peptide ligand under metal-free conditions. Additionally, TEM images showed that the presence of Co(II) inhibits the process of fiber formation, promoting the synthesis of discrete peptide helicates. Also, in the images obtained for the control solutions in DMSO, no fiber formation was observed, both in the absence and presence of Co(II) ions. These results supported our initial hypothesis, derived from the previous fluorescence, CD and UV-Vis studies.

Having confirmed the assembly of the **ΛΛ-Co^{II}_2L** peptide helicate under thermodynamic control, it was oxidized using $(NH_4)_2Ce(NO_3)_6$ into the kinetically inert Co(III) metallopeptide, which was then characterized by mass spectrometry and circular dichroism. The CD spectrum of **Co^{III}_2L** confirmed the retention of the configuration in the metal centre during the oxidation process, and therefore, the formation of the **ΛΛ-Co^{III}_2L** helicate.

Afterwards, we investigated the 3WJ binding properties of **ΛΛ-Co^{II}_2L** and **ΛΛ-Co^{III}_2L**, by monitoring the decrease in the intensity of the emission band of a fluorescein-labelled 3WJ upon addition of increasing amounts of the Co(II) and Co(III) peptide helicates. The resulting profile for the Co(II) helicate was fitted to a 1:1 binding model (Flu-3WJ/helicate), with an apparent dissociation constant of 359 nM, whereas for Co(III), the titration profile was fitted to a more complex model, 1:1+1:3, with K_{D1} = 0.128 μM and K_{D2} = 0.825 μM, respectively. Unfortunately, the dissociation constant for these helicates with a double-stranded DNA could not be measured due to precipitation problems suggesting non-specific interactions between the helicates and the dsDNA.

Finally, the interaction between DNA and Co(II) and Co(III) peptide helicates was analyzed by polyacrylamide gel electrophoresis. The PAGE analysis confirmed that **ΛΛ-Co^{II}_2L** and **ΛΛ-Co^{III}_2L** displayed similar binding affinity and selectivity towards 3WJ, as shown by the appearance of a new slow-migrating band in the presence of this non-canonical DNA structure, which is associated with the formation of the complexes 3WJ/**ΛΛ-Co^{II}_2L** and 3WJ/**ΛΛ-Co^{III}_2L** and was not observed in the incubation with a

regular double-stranded B-DNA, thus confirming the poor affinity of both metallopeptides for the dsDNA.

Chapter III. Efficient and selective cleavege of 3WJ by ATCUN-functionalized peptide helicates

This chapter discusses the synthesis of Cu(II) and Ni(II) chiral peptide helicates, equipped with a new designed ligand derived from the **LLD** peptide described in Chapter I. **LLD** is functionalized with the tripeptide ATCUN motif at the N-terminus, which is known to form square-planar complexes of high affinity with Cu(II) and Ni(II) ions, that have shown nuclease properties. The Cu(II) and Ni(II) metallopeptides derived from this new ligand exhibited selective binding to 3WJ structures. In particular, the Cu(II) peptide helicate displayed highly selective nuclease activity for 3WJ DNA over regular B-DNA.

The new **RGH-LLD** ligand combines the ATCUN motif **L-Arg–Gly–L-His** (RGH) with the sequence of the **LLD** ligand, joined together by three lysine residues coupled via their side chains. **RGH-LLD** was synthesized following standard SPPS procedures and purified by RP-HPLC.

Once the peptide was isolated, its coordinative properties were characterized in the presence of Cu(II) and Ni(II) metal ions in an analogue manner to the previous chapters, by measuring the decrease in the emission intensity of the Bpy residues after the addition of the corresponding metal cations. Given that this strategy only considers the metal coordination to the **LLD** fragment, we assumed that the first equivalent of M(II) with respect to the peptide ligand is taken up by the ATCUN motif, with dissociation constants in the order of picomolar. The resulting profile for Cu(II) was fitted to a 1:1 binding model with an apparent dissociation constant around 0.1 μM for both metal ions, whereas for Ni(II), the profile was fitted to a 1:2 model with K_{D1} = 0.00208 μM and K_{D2} = 0.333 μM. These results, together with those of MALDI, suggest the formation of metallopeptides with an overall ratio of 1:3 (**RGH-LLD**:Cu^{II}/Ni^{II}).

Circular dichroism studies confirmed the formation of chiral helicates, which in the particular case of the Cu(II) metallopeptide, as observed in Chapter I, displayed an inverted configuration consistent with the formation of the **ΔΔ-Cu^{II}_3RGH-LLD** helicate, probably as a consequence of Jahn-Teller distortion. On the other hand, Ni(II) leads to the formation of the diastereoisomer **ΛΛ-Ni^{II}_3RGH-LLD** helicate, as suggested by the positive Cotton effect band at 320 nm.

Once the Cu(II) and Ni(II) helicates were characterized, their 3WJ-binding properties were investigated by measuring the emission of a fluorescein-labelled 3WJ (Flu-3WJ) after successive additions of the **ΔΔ-Cu^{II}_3RGH-LLD** and **ΛΛ-Ni^{II}_3RGH-LLD** helicates. The obtained profiles were fitted to a 1:1 (Flu-3WJ/helicate) binding model, plus non-specific interactions, resulting in apparent dissociation constants of 170 nM for **ΔΔ-Cu^{II}_3RGH-LLD** and 140 nM for **ΛΛ-Ni^{II}_3RGH-LLD**. As in Chapter I, the dissociation constants for these Cu(II) and Ni(II) metallopeptides with 3WJs of different sequences were also calculated, resulting in similar K_D values, thus confirming that **ΔΔ-Cu^{II}_3RGH-LLD** and **ΛΛ-Ni^{II}_3RGH-LLD** recognize 3WJ structures in a sequence-independent manner. The K_D for the interaction with a canonical dsDNA could not be obtained due to the precipitation of the sample during the experiment, thus suggesting non-specific interactions between the metallocylinders and the dsDNA.

Following, the ability of **ΔΔ-Cu$^{II}_3$RGH-LLD** and **ΛΛ-Ni$^{II}_3$RGH-LLD** to selectively cleave 3WJs *in vitro* was determined using electrophoresis assays. PAGE analysis revealed that the Cu(II)-metallopeptide selectively binds and cleaves 3WJ structures in the presence of H_2O_2, as can be concluded from the decrease in the intensity of the band associated with the 3WJ/**ΔΔ-Cu$^{II}_3$RGH-LLD** adduct. This nuclease activity was not observed in the regular B-DNA, nor in the presence of the Ni(II) helicate. The Ni(II) metallopeptide showed a different behavior, **ΛΛ-Ni$^{II}_3$RGH-LLD** selectively recognizes 3WJs with high affinity, as indicate for the appearance of a new slow-migrating band consistent with the formation of the 3WJ/**ΛΛ-Ni$^{II}_3$RGH-LLD** complex, but it is not capable of damaging the DNA. In both cases, the band associated with the double-stranded DNA remains unaffected, thus confirming the poor affinity of the chimeric peptide helicates for this canonical DNA structure.

Finally, we tested the ability of **ΛΛ-Ni$^{II}_3$RGH-LLD** and **ΔΔ-Cu$^{II}_3$RGH-LLD** helicates to selectively recognize 3WJ structures in functional HeLa cells, previously permeabilized with digitonin. The TAMRA-labelled Ni(II) and Cu(II) peptide helicate colocalize with the green-labelled DNA replication foci, enriched in 3WJ structures, thus demonstrating the potential of these new systems for selective associating with 3WJs in cells.

Chapter IV. Selective recognition of A/T-rich DNA 3-way junctions with a three-fold symmetric tripeptide

In this chapter, we describe a new strategy for the recognition of 3WJs, which relies on the cooperative and sequence-selective binding to the three duplex DNA arms in the 3WJ structure. The efficacy of this approach was demonstrated through the design of a new peptide system, with trigonal symmetry, which consists of three AT-Hook peptide strands attached to a 1,3,5-tristyrylbenzene core. This new trimeric peptide showed 3WJ-binding properties with selectivity for A/T-rich sequences, which were characterized by exploiting the increased intensity in the fluorescence emission of the central benzene core, derived from the conformational restriction arising from its DNA binding.

A hypothetical computational model was built to design the new trimeric AT-Hook ligand, consisting of three AT-Hook peptides, simultaneously inserted into the three duplex DNA minor grooves in the 3WJ, and a 1,3,5-tristyrylbenzene placed over the central hydrophobic cavity. The distance of 4 Å between the N-terminus of the AT-Hook peptides and the carboxylate groups in the benzene core suggested that a lysine residue would perfectly fit as a linker residue. Our hypothetical model was further investigated by molecular modelling using the platform *GaudiMM*, that verified an excellent fit between the designed model and a 3WJ structure.

The AT-Hook peptide was assembled following standard Fmoc-solid-phase MW-assisted peptide synthesis protocols. Next, the chains were simultaneously attached to the 1,3,5-tristyrylbenzene core on resin, leading to the final peptide that was purified by RP-HPLC.

Once the trimeric AT-Hook peptide **1** was isolated, we investigated its interaction with different DNA by measuring the increase in the intrinsic fluorescence emission of the 1,3,5-tristyrylbenzene, promoted by a conformational restriction associated to the

DNA binding. Based on this, the fluorescence emission of **1** was recorded after successive additions of different DNA structures, TA-3WJ (A-T-rich), GC-3WJ (G-C-rich) as control and dsDNA (A-T-rich). For the target TA-3WJ, the resulting profile was fitted to a composite model that combines the formation of a specific 1:1 complex, with K_D in the order of 20 nM, with a minor contribution of a nonspecific binding term. The titration profile in the presence of A/T-rich double-stranded DNA was fitted to the same model, but with a lower binding affinity, reflected in an approximately 10 times weaker dissociation constant. As expected, in the profile obtained for the GC-3WJ control, there was no increase in the fluorescence emission. It is noteworthy that the addition of 1 eq. of TA-3WJ to this solution resulted in an increase in the fluorescence intensity, thus confirming the high binding selectivity of trimeric AT-Hook **1** for A/T-rich 3WJs.

The interaction of **1** with the minor groove of the duplex branches in TA-3WJ was supported by circular dichroism, with the appearance of a new band in the presence of peptide **1**. Nevertheless, the CD spectrum of GC-3WJ remained unchanged after the addition of **1**, proving the poor binding affinity.

Additional experiments by electrophoretic gel shift assays confirmed the formation of a single high-affinity complex between **1** and TA-3WJ, as indicated by the appearance of a new slow-migrating band, as well as the formation of lower affinity associations with GC-3WJ and dsDNA.

Finally, we tested the ability of trimeric peptide **1** to selectively label A/T-rich 3WJs in functional HeLa cells, previously permeabilized with digitonin. It was observed that **1**, appearing in the blue emission channel, has a strong correlation with the localization of the green-labelled DNA replicating sites, where the 3WJs are accumulated. These results confirm the great potential of **1** to label 3WJs in the cellular DNA.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Introduction

1. Chemical biology

The definition of *chemical biology* has been controversial since the term emerged in the mid-20th Century. However, broadly speaking, it can be defined as a field of research that lies at the intersection of chemistry and biology, with a focus on understanding biological phenomena at a molecular level through chemical tools and applying this knowledge to advance the chemical sciences.[1] Due to the multidisciplinary nature of chemical biology, it is difficult to define its scope, which comprises many fields, including enzyme catalysis, biomolecular recognition, supramolecular and synthetic chemistry, cell signalling and internalization, or spectroscopy, among others.

Although chemical biology was not defined as an independent discipline until the 20th century,[2] the first scientific advances in this field date back to the 19th century, with early pioneers like the German chemist Freidrich Wöhler, who first synthesized urea in 1828, thus showing that biological compounds could be obtained from inorganic chemical materials.[3] A few years later, in 1868, Friedrich Miescher made the discovery of nuclein while he was conducting experiments on the chemical composition of leukocytes.[4] In addition, this period saw the development of specific cell staining agents,[5] which played a major role in the studies carried out by the Nobel Price Paul Ehrlich, focused on the methodical search for a specific drug to treat syphilis. His research led to the obtention of the arsenic-based drug *Salvarsan* in 1910, a landmark discovery that represents the beginning of *medicinal chemistry*.[6] In the latter half of the 20th century, there were many examples of the intimate connection between the field of Chemistry and Biology, which prompted the birth of the term *chemical biology* in 1945, first introduced by Linus Pauling and George W. Beadle in a grant proposal to the *Rockefeller Foundation*, that outlined *"the great problems of Biology"* and the key role of Chemistry in addressing them.[2] However, it was not until 1954 that we find the first definition of the scope and aims of this new independent discipline in the journal *Engineering and Science*.[7] Nevertheless, despite its early beginnings, in the 1980s there was still some debate about the concept of *chemical biology*, as reflected in the Arthur Kornberg´s 1987 publication advocating the value of Chemistry in biological research: "*Much of life can be understood in rational terms if expressed in the language of chemistry*".[8]

Thankfully, despite some resistance from the biological community, in the 21st century chemical biology has evolved into a well-established field of research, even having been recognized with the Nobel Prize in Chemistry 2022 awarded jointly to

---

[1] E. L. Ostler, *Chem. Cent. J.* **2007**, *1*, 5.
[2] L. E. Kay, The Molecular Vision of Life. Caltech, *The Rockefeller Foundation, and the Rise of the New Biology*, Oxford Univ. Press **1997**, ISBN: 9780195111439 (pp. 225-242) and *@CaltechArchives.*
[3] F. Wöhler, *Ann. Phys.* **1828**, *88*, 253–256.
[4] R. Dahm, *Hum. Genet.* **2008**, *122*, 565–581.
[5] G. Musumeci, *J. Histol. Histopathol.* **2014**, *1*, 5.
[6] J. Drews, *Nat. Rev. Drug Discov.* **2004**, *3*, 797–801.
[7] *Eng. & Sci.* **1954**, *17*, 9-13.
[8] A. Kornberg, *Biochemistry* **1987**, *26*, 6888-6891.

Carolyn R. Bertozzi, Morten Meldal and K. Barry Sharpless. This field still defies a closed definition because its constant growth and expansion fuelled by the development of new tools and approaches, such as new bioconjugation methods, or the synergies with the fields of nanotechnology and materials science. This diverse and constantly evolving nature means that *chemical biology* will be at the forefront of major advances in science.[9]

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

[9] G. von Kiedrowski, *ChemBioChem* **2001**, *2*, 597-598.

2. DNA: basic concepts

2.1. Biological functions

2.1.1. Central Dogma of Biology

The genetic information of living organisms, namely, the specific sequence of the nucleic acids that codes for the synthesis of their proteins, defines their nature.[10] The mechanism of genetic information transmission was postulated by Francis Crick in 1957. According to Crick´s original scheme, known today as the *Central Dogma of Biology*, the genetic information flows in a single direction from the DNA to proteins via RNA, thus excluding the possibility of genetic modification carried out by proteins.[11] This conceptual framework marked the birth of the field of *molecular biology* in the 1950s. The unidirectionality of the process has been extensively questioned, thus resulting in several updates to this dogma that introduced new pathways for the flux of information that account for the epigenetic modification of gene expression by proteins and the regulatory role of some RNAs.[12]

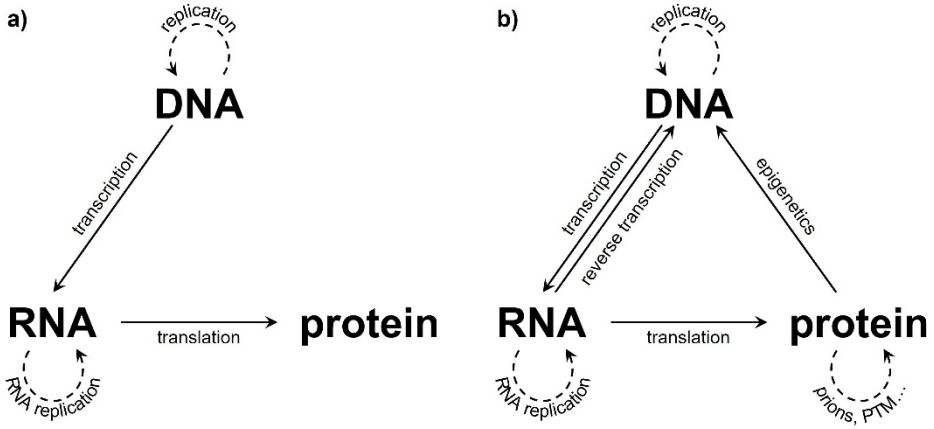


Figure 1. a) Original *Central Dogma of Biology* proposed by F. Crick in 1957; b) updated version of this dogma.[11,12]

2.1.2. Specific functions of DNA

Deoxyribonucleic acid (DNA) is the cell's hereditary material and carries instructions for the development, function, and reproduction of a living organism. As stated in the *Central Dogma of Biology*, its main biological functions include the storage of genetic information, protein coding (transcription and translation) and self-replication to ensure the transference of information to daughter cells. The genetic information is encoded by the linear sequence of four different nucleotides, that make up the DNA strands: adenine (A), cytosine (C), guanine (G) and thymine (T). The order in which these nucleotides are arranged into the DNA chain defines the amino acid sequence of the coded proteins, since the unit constituted by three adjacent nucleotides (codon) codes for each of twenty amino acids. At the functional level, DNA is organized into genes, segments of the molecule that code for a specific protein, through a process that mainly involves two

[10] H. Lodish, A. Berk, L. S. Zipursky, P. Matsudaira, D. Baltimore, J. Darnell, *Molecular Cell Biology*, W. H. Freeman, **2000**, chapter 9.
[11] F. H. C. Crick, *Symp. Soc. Exp. Biol*. **1958**, *12*, 138-163.
[12] a) F. Crick, *Nature* **1970**, *227*, 561–563; b) M. Cobb, *PLoS Biol.* **2017**, *15*, e2003243.

steps: *transcription* of the gene into a messenger ribonucleic acid (mRNA) molecule, and its *translation* into a protein via an intermediate mRNA.[12,13] Despite being a major role of DNA, just over 1.5% of the human genome is protein-coding, and the specific biological functions of the additional 98.5%, known as non-coding DNA, remain to be elucidated. However, it is known that it is involved in the regulation of gene expression.[14]

2.2. Structures of DNA

In eukaryotic cells, DNA is stored in their nuclei packed into tight bundles called *chromosomes*, which are part of a dense DNA-protein complex, known as *chromatin*. The *nucleosome* is the minimal structural unit of chromatin, comprising a DNA sequence wound around an octamer of histone proteins.[15]

In terms of structure, DNA can be described as a biopolymer resulting from the condensation of monomeric units known as deoxyribonucleotides. Nucleotides consist of three primary components: a *pentose sugar* (β-D-2'-deoxyribose) and a highly negatively charged *phosphodiester backbone,* both with structural function, and a *nitrogenous base* responsible for the genetic codification, which can be either pure (adenine-A and guanine-G) or pyrimidine (thymine-T and cytosine-C).

Although the DNA double helix can adopt different conformations depending on the conditions of the environment such as pH, ionic strength or solvent, at physiological conditions, the canonical DNA or B-DNA is the most common conformation and it is also the one most physiologically relevant and often recognized by transcription factors. B-DNA is a double-stranded DNA (dsDNA) with a regular structure of a right-handed double helix with a helix diameter of approximately 20 Å; the strands form antiparallel dimers that stick together through hydrogen bonds between complementary bases (A-T and G-C). The helix has 10 pairs of bases per each turn with 3.4 Å between successive pairs, and the rotation per residue is 36º. This structure is the result of the geometry and conformational preferences of the sugars in combination with the hydrophobic character of the nitrogenous bases, oriented towards the interior of the helix to reduce their contact surface within the aqueous media.[16] The B-DNA structure is very regular and conserved, but exhibits some sequence-dependent variations, which are of great relevance for protein-DNA interactions.[17] Also noteworthy is the presence of two grooves of different sizes in the helix (Figure 2): the *minor groove*, narrow and deep, whose width depends on the nucleotide sequence (regions rich in A-T are narrower than those with G-C), and the *major groove*, wide and shallow, where the proteins interact preferentially, because of its greater accessibility and variability, thus encoding more information.[18]

---

[13] H. Pearson, *Nature* **2006**, *441*, 398–401.
[14] E. Pennisi, *Science* **2007**, *316*, 1556–1557.
[15] a) Y.-J. Li, X.-H. Fu, D.-P. Liu, C.-C. Liang, *Int. J. Biochem. Cell Biol.* **2004**, *36*, 1411–1423; b) S. Khorasanizadeh, *Cell* **2004**, *116*, 259–272.
[16] S. Neidle, *Nucleic Acid Structure and Recognition*, Oxford University Press, USA, **2002**.
[17] a) G. M. Blackburn, M. J. Gait, D. Loakes, D. M. Williams, Eds., *Nucleic Acids in Chemistry and Biology*, RCS Pub., Cambridge, **2006**; b) R. E. Dickerson, H. R. Drew, *J. Mol. Biol.* **1981**, *149*, 761–786.
[18] a) P. L. Privalov, A. I. Dragan, C. Crane-Robinson, K. J. Breslauer, D. P. Remeta, C. A. S. A. Minetti, *J. Mol. Biol.* **2007**, *365*, 1–9; b) S. Boga, D. Bouzada, D. García Peña, M. Vázquez López, M. E. Vázquez, *European J. Org. Chem.* **2018**, *2018*, 249–261.

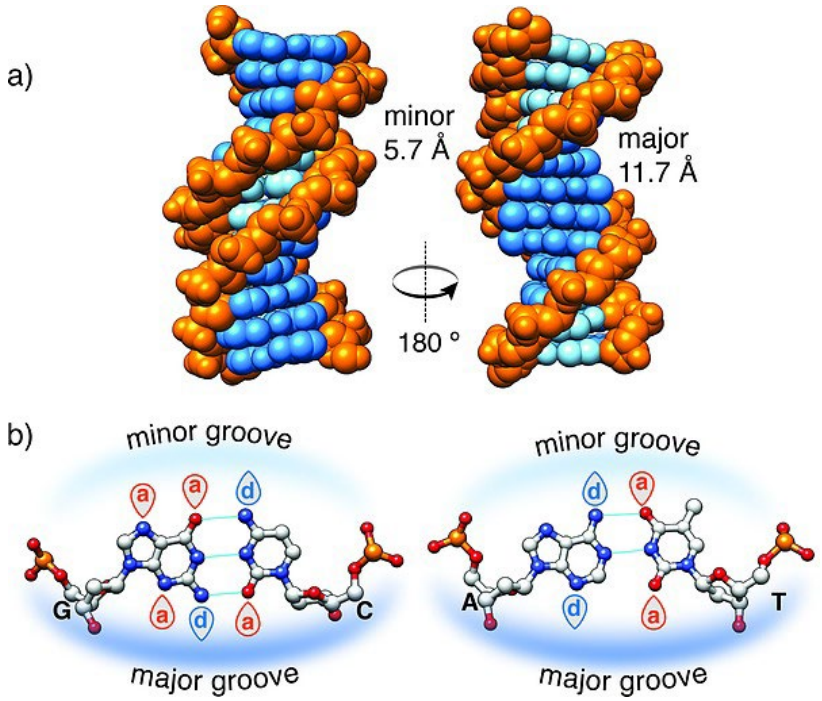


Figure 2. a) Structure of canonical DNA, or B-DNA, showing the difference in size between the major groove, in dark blue, and the minor groove, in light blue; b) Structure of the nitrogenous bases highlighting the hydrogen bonds between complementary bases.[18b]

Beyond the canonical antiparallel double-stranded B-DNA conformation, the DNA can adopt different non-canonical secondary structures, such as G-quadruplexes, which are overrepresented in gene promoter regions,[19] the related telomeric and promoter i-motifs,[20] Holliday junctions,[21] and three-way-junctions (3WJs).[22] The DNA strands that make up these non-canonical structures can be joined together in an antiparallel or parallel manner, giving them a large structural diversity, which added to the relevance of the biological functions that they play (genetic regulation, transcription and recombination or the chromosome assembly),[23] makes them promising therapeutic DNA targets. In fact, the main topic of this work is the development of new 3WJ binding agents.

2.2.1. DNA three-way junction (3WJs)

3WJs, the simplest branched DNA structure, are symmetric assemblies composed by three convergent B-DNA arms. The central area of 3WJs, where their three B-DNA arms join, is known as the *branch point*, and features a hydrophobic cavity of ~ 12 Å in

[19] a) S. Balasubramanian, L. H. Hurley, S. Neidle, *Nat. Rev. Drug Discov.* **2011**, *10*, 261–275; b) S. Neidle, *Nat. Rev. Chem.* **2017**, *1*, 0041.
[20] a) H. A. Day, P. Pavlou, Z. A. E. Waller, *Bioorg. Med. Chem.* **2014**, *22*, 4407–4418; b) J. Amato, N. Iaccarino, A. Randazzo, E. Novellino, B. Pagano, *ChemMedChem* **2014**, *9*, 2026–2030.
[21] A. L. Brogden, N. H. Hopcroft, M. Searcey, C. J. Cardin, *Angew. Chem. Int. Ed Engl.* **2007**, *46*, 3850–3854.
[22] K. J. Neelsen, M. Lopes, *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 207–220.
[23] a) N. B. Ramsing, K. Rippe, T. M. Jovin, *Biochemistry* **1989**, *28*, 9528–9535; b) V. R. Parvathy, S. R. Bhaumik, K. V. R. Chary, G. Govil, K. Liu, F. B. Howard, H. T. Miles, *Nucleic Acids Res.* **2002**, *30*, 1500–1511.

diameter with a prismatic shape, which is delimited by the terminal base pairs at the end of each of their three B-DNA arms (Figure 3).[24,25]

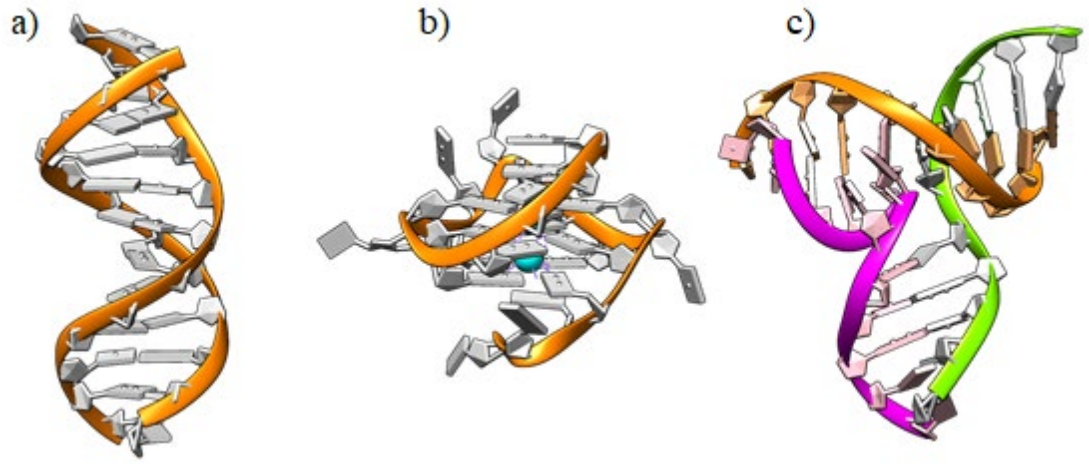


Figure 3. Some secondary structures of DNA: a) B-DNA, b)-G-quadruplex; c) Three-way junction.

In principle, 3WJs should fold by pairwise coaxial stacking of helices, on the contrary, it has been widely reported that three-way DNA junctions with full base-pairing assume a symmetric "Y" structure, with trigonal pyramidal geometry rather than planar shape, characterized by inter-helical angles of 60 to 90°.[26] Evidence for this has been provided by comparative EMSA assays, which demonstrate that the three angles between the arms of perfect 3WJs are similar, even in the presence of ions that are found to induce coaxial stacking.[27] These results were backed up by FRET experiments, showing that end-to-end distances between the three branches were also similar,[28] as well as by NMR studies,[29] and through AFM images, where symmetrical molecules are seen.[30] According to D. M. J. Lilley, who reviewed branched DNA structures, *"The perfect 3H three-way junction appears to break the general folding principle, i.e. that helical junctions tend to undergo coaxial helical stacking and ion-dependent folding. The rigid framework of the fully paired three-way junction evidently makes such folding very difficult".*[31] Nevertheless, by relaxing the structural constraints on a perfect 3WJ with the addition of conformational flexibility, such as unpaired bases, or in RNA junctions (RNA is known to be much more conformationally flexible than DNA), 3WJs undergo a conformational

[24] K. C. Woods, S. S. Martin, V. C. Chu, E. P. Baldwin, *J. Mol. Biol.* **2001**, *313*, 49–69.
[25] D. M. J. Lilley, *Biopolymers* **1998**, *48*, 101–112.
[26] a) L. S. Shlyakhtenko, E. Appella, R. E. Harrington, I. Kutyavin, Y. L. Lyubchenko, *J. Biomol. Struct. Dyn.* **1994**, *12*, 131–143; b) L. S. Shlyakhtenko, D. Rekesh, S. M. Lindsay, I. Kutyavin, E. Appella, R. E. Harrington, Y. L. Lyubchenko, *J. Biomol. Struct. Dyn.* **1994**, *11*, 1175–1189.
[27] D. R. Duckett, D. M. Lilley, *EMBO J.* **1990**, *9*, 1659–1664.
[28] F. Stühmeier, J. B. Welch, A. I. Murchie, D. M. Lilley, R. M. Clegg, *Biochemistry* **1997**, *36*, 13530–13538.
[29] a) N. B. Leontis, M. T. Hills, M. Piotto, A. Malhotra, J. Nussbaum, D. G. Gorenstein, *J. Biomol. Struct. Dyn.* **1993**, *11*, 215–223; b) M. A. Rosen, D. J. Patel, *Biochemistry* **1993**, *32*, 6563–6575; c) M. Zhong, M. S. Rashes, N. B. Leontis, N. R. Kallenbach, *Biochemistry* **1994**, *33*, 3660–3667.
[30] L. S. Shlyakhtenko, V. N. Potaman, R. R. Sinden, A. A. Gall, Y. L. Lyubchenko, *Nucleic Acids Res.* **2000**, *28*, 3472–3477.
[31] D. M. Lilley, *Q. Rev. Biophys.* **2000**, *33*, 109–159.

change induced by coaxial helix-helix stacking in the presence of ions, leading to a stacked conformation. In these specific cases, 3WJs follow the same general principle of folding as the four-way junctions.[27]

3WJs play important roles in different cellular processes;[32] they are a common motif in RNA-protein complexes as well as in structural and catalytic RNAs,[31] since they occur during splicing[33] and translation processes.[34] In addition, 3WJs are transiently formed during the processes of DNA recombination, retroviral integration and transposition, and also during DNA replication at the meeting of leading and lagging strands.[35,36,37] Their roles in these major biological processes make 3WJs a promising target for the development of new DNA binding agents with therapeutic properties.[38]

[32] a) J. Atkinson, P. McGlynn, *Nucleic Acids Res.* **2009**, *37*, 3475–3492; b) R. R. Sinden, M. J. Pytlos-Sinden, V. N. Potaman, *Front. Biosci.* **2007**, *12*, 4788–4799.
[33] C. Guthrie, B. Patterson, *Annu. Rev. Genet.* **1988**, *22*, 387–419.
[34] a) B. T. Wimberly, D. E. Brodersen, W. M. Clemons Jr, R. J. Morgan-Warren, A. P. Carter, C. Vonrhein, T. Hartsch, V. Ramakrishnan, *Nature* **2000**, *407*, 327–339; b) N. Ban, P. Nissen, J. Hansen, P. B. Moore, T. A. Steitz, *Science* **2000**, *289*, 905–920.
[35] M. R. Singleton, S. Scaife, D. B. Wigley, *Cell* **2001**, *107*, 79–89.
[36] a) N. C. Seeman, N. R. Kallenbach, *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 53–86; b) C. Altona, J. A. Pikkemaat, F. J. Overmars, *Curr. Opin. Struct. Biol.* **1996**, *6*, 305–316.
[37] C. H. Hu, F. Y. Xu, K. Wang, A. N. Pearson, G. D. Pearson, *Gene* **1992**, *110*, 145–150.
[38] a) R. D. Wells, *Trends Biochem. Sci.* **2007**, *32*, 271–278; b) A. Bacolla, R. D. Wells, *Mol. Carcinog.* **2009**, *48*, 273–285; c) H. Tateishi-Karimata, N. Sugimoto, *Chem. Commun.* **2020**, *56*, 2379–2390.

3. Amino acids, peptides and proteins

3.1. Short introduction

Proteins play fundamental roles in most biological processes that take place in all living cells. The history of proteins can be traced back to the 18th century, when the italian scientist Iacopo Beccari of Bologna found that wheat flour contains a protein component, known today as gluten.[39] This discovery was followed, in the 19th century, by the systematic and scientific study of proteins. This was made possible by the discovery that heating proteins in concentrated acid led to the isolation of simple crystalline substances, the *amino acids*, which could be studied using the rudimentary analytic techniques available at that time. In this way, Henri Braconnot obtained in 1820 *glycine* from gelatin, thus marking the beginning of modern protein chemistry.[40] However, it was not until 1838 that Mulder, as suggested by Berzelius, coined the term *protein* with the meaning of primary substance.[41] By the early 20th century, eighteen amino acids had been identified in proteins,[40] known with three-letter abbreviations, that became tedious to use, thus driving Šorm to the establishment of a one-letter abbreviation system in 1961.[42]

Although the different amino acids that make up a protein, its chemical structure, had been known for a long time, the major protein secondary structure elements were not proposed until 1951, when Linus Pauling, Robert Corey and German Branson described the α-helix and the β-sheet to explain X-ray diffraction patterns of some fibrous and crystalline proteins, such as hemoglobin and myoglobin; seven years later, in 1958, Kendrew *et al.* confirmed the three-dimensional structure of proteins, by solving the first atomic structure of myoglobin using X-ray crystallography.[43]

Nowadays, the study of proteins is still a main area of research, with new advances in protein structure, function and engineering that lead to the innovation in other fields such as biotechnology or medicine.

3.2. Structure of proteins

Peptides and proteins are biopolymers made up of α-amino acids. An α-amino acid consists of a carboxyl group (–COOH), an amino group (–NH_2), a hydrogen atom, and a variable side chain (–R) attached to the same carbon atom, known as α carbon (Cα). Most proteins are composed of a combination of twenty of these α-amino acids, known as *standard amino acids*, which differ in the chemical nature, and hence reactivity, of the functional groups of their side chains. Based on their chemical nature, determined by the side chain, amino acids are classified as hydrophobic (G, A, V, L, I, M, F, P, W), polar (S, T, N, Q, Y, C), acid (D, E) and basic (K, R, H).[44] As mentioned above, these amino

[39] E. F. Beach, *J. Hist. Med. Allied Sci.* **1961**, *16*, 354–373.
[40] H. B. Vickery, C. L. A. Schmidt, *Chem. Rev.* **1931**, *9*, 169–318.
[41] H. Hartley, *Nature* **1951**, *168*, 244.
[42] F. Šorm, B. Keil, J. Vaněček, V. Tomášek, O. Mikeš, B. Meloun, V. Kostka, V. Holeyšovský, *Collect. Czech. Chem. Commun.* **1961**, *26*, 531–578.
[43] J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, D. C. Phillips, *Nature* **1958**, *181*, 662–666.
[44] a) D. Voet, J. G. Voet, C. W. Pratt, *Fundamentals of Biochemistry: Life at the Molecular Level*, John Wiley & Sons, New York, **2016**, chapter 4; b) B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter, *Molecular Biology of the Cell*, Garland Science, New York, **2008**, chapter 3.

acids are named with a three-letter or one-letter abbreviations, used to indicate the sequence of proteins.



Figure 4. The "Periodic Chart of Amino Acids", where all the twenty proteinogenic amino acids are classified by different colors. Adapted from www.bachem.com.

Proteins are formed by condensation of the carboxy group of one amino acid with the amino group of the following amino acid, thus leading to the formation of an amide bond, also known as *peptide bond* (Figure 5). To synthesize this peptide bond, which is kinetically stable once formed, an energy input is required because the equilibrium of this reaction is shifted towards hydrolysis.[45]

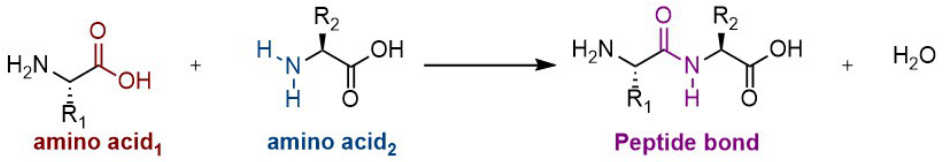


Figure 5. Condensation reaction between two consecutive amino acids forming a peptide bond.

X-ray diffraction studies of amino acids and small peptides, conducted by Linus Pauling and Robert Corey, demonstrated that the C–N peptide bond is shorter than the single C–N bond of amines, thus concluding that the peptide bond has partial double bond character. This has an important structural consequence, which is that the atoms associated with the amide bond are coplanar, and therefore the C–N peptide bond is unable to rotate freely. This allows us to picture the polypeptide backbone as a sequence of rigid planes sharing a rotation point at Cα. Additionally, the rigid peptide bonds limit the range of conformations that can be assumed by a polypeptide chain, which are defined trough the two dihedral angles, φ (N-Cα) and ψ (Cα-C). that characterize the spatial

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

[45] R. B. Martin, *Biopolymers* **1998**, *45*, 351–353.

orientation between two contiguous planes (Figure 6). According to this, the allowed values for φ and ψ, so the unrestricted conformations, are graphically revealed when ψ is plotted versus φ in the *Ramachandran plot*, introduced by G. N. Ramachandran.[46]

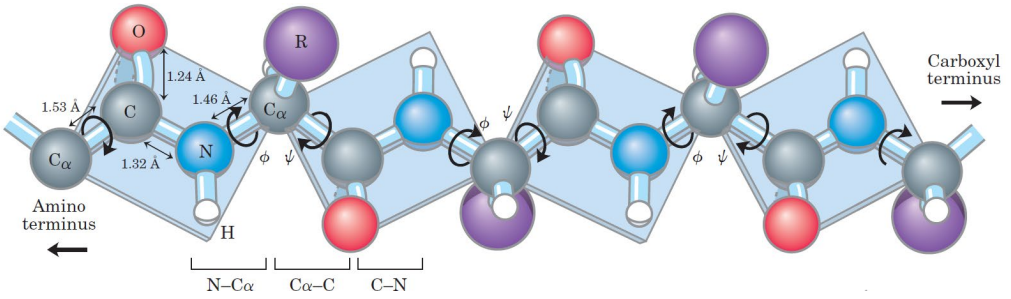


Figure 6. Conformation of a polypeptide chain defined by the dihedral angles φ and ψ. Image taken from Ref. 46 (p. 119).[46]

Four different levels of complexity are commonly used to describe the structure of proteins:

- *Primary structure*: it describes all the covalent bonds, mostly peptide and disulfide bonds, that connect the amino acid residues in a polypeptide chain, with the most important element being the amino acids sequence, because it defines the structure, reactivity, and function of the protein.
- *Secondary structure*: corresponding to the local spatial arrangement of the main chain atoms in a particular segment of the polypeptide, independently of the side chains and the relationships with other segments of the chain. Local folding can adopt certain particularly stable patterns known as *secondary structure motifs*, including the α-helix, β-sheet, β-turn and random coil, with the latter not exhibiting a regular pattern.
- *Tertiary structure*: it refers to the three-dimensional arrangement of all the atoms in a protein, because of the spatial distribution of secondary structure motifs, which is usually stabilized via hydrophobic interactions between the nonpolar side chains or disulfide bonds.
- *Quaternary structure*: the arrangement in space of the different polypeptide subunits that make up a protein.[46]

3.2.1. Secondary structure motifs: α-helix, β-sheet and β-turn

3.2.1.1. α-helix

The α-helix is one of the most prevalent and important protein structures, since about 30% of the residues in a protein are in α-helical conformation.[47] This structure is characterized by two unique dihedral angles (φ = −57º and ψ = −47º), that define a conformation in which the polypeptide backbone is wound around an imaginary axis, with the side chains groups orienting outward from the helix. The ideal α-helix is right-handed, with each helical turn containing approximately 3.6 amino acid residues (Figure 7). This conformation is thermodynamically stable due to the hydrophobic effects resulting from the compaction of the peptide backbone such as the hydrogen bonds

[46] D. L. Nelson, M. M. Cox, *Lehninger Principles of Biochemistry*, W. H. Freeman Company, New York, **2008**, chapter 4.
[47] D. J. Barlow, J. M. Thornton, *J. Mol. Biol.* **1988**, *201*, 601–619.

between the acyl oxygen (C=O) of residue *i* (acceptor) and the amide group (N–H) of residue *i*+4 (donor).[46,48] However, the stability of the α-helix is strongly dependent on different factors such as the amino acids that make up this structure,[49] their position in the peptide sequence,[50] as well as the residues placed close to the ends of the helix.[51]

- The intrinsic tendency of each amino acid to stabilize an α-helical structure is dependent on its side chain. Thus, proline and glycine have lower tendencies due to their structural limitations. Proline lacks a N–H group because of its cyclic structure, while glycine has high conformational flexibility that makes folding into an α-helix entropically unfavorable. Branching at the β-carbon atom of amino acids like valine, threonine and isoleucine, as well as the presence of hydrogen bond donors or acceptors close to the main chains in serine, aspartate and asparagine, destabilize α-helices due to steric hindrance or competition for the C=O and N–H groups of the main chain. Therefore, alanine and arginine present the highest and second-highest tendencies, respectively, to form α-helical conformations.[49]
- The position of a residue with respect to their neighbors plays important roles in helix stabilization since the side chains interactions can stabilize or destabilize this structure. Side chains at positions (*i, i*+3) and (*i, i*+4) are close in space, so interactions between them have major impact on helix stability, being the most significant, the electrostatic interactions between amino acids with different charges and the hydrophobic interactions between aromatic residues. On the other hand, residues at positions (*i, i*+2), (*i, i*+5) and (*i, i*+6) place the side-chain pairs on opposite sides of the helix, so they do not interact.[50]
- The stability of the α-helix is also influenced by the residues located near its ends. This is because the alignment of all the peptide bonds with helical axis generates a net macrodipole for the whole helix (due to the inherent polarization of the peptide bond), placing partial positive and negative charges at the N-terminus and the C-terminus, respectively. As a result, the positively charged residues are usually found near the C-terminal end while the negatively charged ones at the N-terminal end, in order to exert a stabilizing interaction with the dipole charge.[46,51]

3.2.1.2. β-sheet

β-sheets are regular secondary structures in proteins that consist of two or more β-strands, which are short polypeptide chains of 5-8 residues that lie in an extended zig-zag form, resulting in a distance of 3.5 Å between consecutive residues. Adjacent β-strands are linked together through hydrogen bonds, between the N–H and C=O groups of the amide bonds, in either a parallel or antiparallel manner, leading to different substructures (Figure 7). Among these, the antiparallel orientation is the most stable as it causes minimal distortion of the inter-chain hydrogen bonding. It is also noteworthy that in this β-sheet structure the side chains of subsequent amino acids are oriented

[48] a) C. I. Branden, J. Tooze, *Introduction to Protein Structure*, Garland Science, New York, **2012**, chapter 2; b) N. Sewald, H.-D. Jakubke, *Peptides: Chemistry and Biology*, John Wiley & Sons, New York, **2015**, chapter 2.
[49] C. N. Pace, J. M. Scholtz, *Biophys. J.* **1998**, *75*, 422–427.
[50] J. E. Brown, W. A. Klee, *Biochemistry* **1971**, *10*, 470–476.
[51] a) W. G. J. Hol, *Adv. Biophys.* **1985**, *19*, 133–165; b) D. S˘ali, M. Bycroft, A. R. Fersht, *Nature* **1988**, *335*, 740–743.

perpendicularly to the plane of the sheet in an alternating pattern. Actually, the classification of β-sheets as secondary structure motifs is not accurate, since this conformation typically involves individual segments that are quite distant from each other in the linear sequence of the polypeptide,[46,48] (although they are close in the space), which refers to the *tertiary structure*.

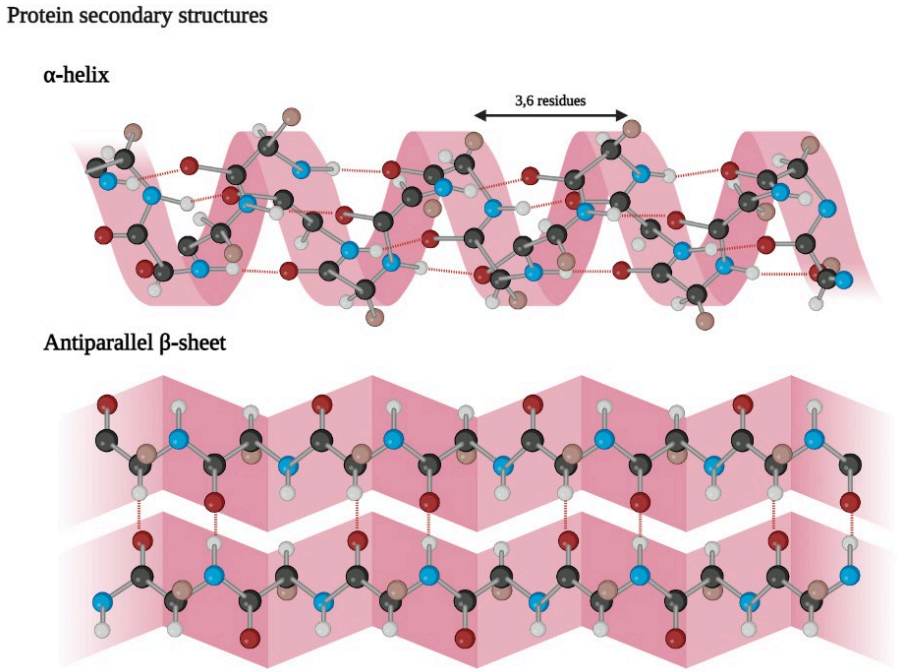


Figure 7. Top: ball-and-stick model of an α-helix protein structure, showing intrachain hydrogen bonds (red dashed lines); bottom: ball-and-stick model of an antiparallel β-sheet protein structure, showing hydrogen-bond crosslinks between adjacent chains (red dashed lines) (C = black, H = white, O = red, N = blue, R group = brown).

3.2.1.3. β-turn

Reverse turns, which are regions of the polypeptide chain where an overall change of direction occurs, are a major structural feature in globular proteins, comprising around one third of the total amino acids.[52] Turns connect different secondary structure elements and are typically located on the protein surface, sometimes involved in intermolecular interactions, although probably for us their most attractive function is their main role in the protein folding.[53] Reverse turn peptide sequences are dominated by Gly and Pro residues, due to their flexibility and constrained cyclic structure respectively, and to a lesser extent, by Asn and Asp. There are several different types of reverse turns, such as *γ-turns* and *α-turns*, being the most common type the *β-turns*,[52] which are found connecting the ends of two adjacent strands in an antiparallel β-sheet.[46]

β-turns are made up of four residues, stabilized through an intramolecular hydrogen bond between the *i* and *i*+3 residues, resulting in a 10-membered ring[54] that places Cα1 and Cα4 at less than 7 Å of distance. In 1968, the β-turns were classified by

USC

[52] A. M. C. Marcelino, L. M. Gierasch, *Biopolymers* **2008**, *89*, 380–391.
[53] G. D. Rose, *Nature* **1978**, *272*, 586–590.
[54] K. Suat, S. D. S. Jois, *Curr. Pharm. Des.* **2003**, *9*, 1209–1224.

Venkatachalam according to the φ and ψ torsion angles of the *i*+1 and *i*+2 residues, known as "*corner residues*", into type I and I' (mirror images) and type II and II':[55]

- **Type I β-turn:** the major turn type in native proteins, composed of L-residues at both positions *i*+1 and *i*+2. The Pro residue frequently occurs at position *i*+2, although it can also appear at *i*+1.
- **Type II β-turn:** in this case, the L-Pro residue is mainly located at *i*+1. The *i*+2 position is typically occupied by Gly, D-residues or small L-residues.
- **G1 β-bulge turn:** this is a type I turn featuring an extra residue at the *i*+3 position, which places the hydrogen bond between *i* and *i*+4.
- **Type VI turn:** the peptide bond between *i*+2 and *i*+3 displays a *cis* conformation, while retaining the hydrogen bond between *i* and *i*+3.[56,57]
- **Open turn:** a reverse turn not stabilized by a hydrogen bond.[58]

In *chemical biology*, the stabilization of these β-turns has gained great interest because they are responsible for nucleating the folding of β-hairpins, of biological importance due to their presence in β-sheets and in many bioactive cyclic peptides. This has led to the development of different strategies to mimic the β-turn motif in short synthetic peptides, including: amide-bond substitution in the peptide strand;[59] use of N-methylated residues and α-methyl amino acids, such as aminoisobutyric acid (Aib);[60] cyclization;[61] replacement of the *i*+1 and *i*+2 residues by organic linkers;[62] and substitution of L-amino acids by D-amino acids.[61] Regarding this latter approach, it has been widely reported the use of heterochiral sequences, combining L- and D-residues, to nucleate type I and II β-turns in artificial peptidomimetics.[63,64] In fact, several examples have already been described of type II β-turn formation driven by sequences bearing the conformationally restricted L-Pro and a D-residue, at *i*+1 and *i*+2 positions respectively,[65] while the enantiomeric sequence, D-Pro–L-residue, has been shown to induce the type II' β-turn mirror image.[66]

[55] C. M. Venkatachalam, *Biopolymers* **1968**, *6*, 1425–1436.
[56] K. S. Rotondi, L. M. Gierasch, *Biopolymers* **2006**, *84*, 13–22.
[57] J. S. Richardson, E. D. Getzoff, D. C. Richardson, *Proc. Natl. Acad. Sci. U. S. A.* **1978**, *75*, 2574–2578.
[58] G. D. Rose, L. M. Gierasch, J. A. Smith, *Adv. Protein Chem.* **1985**, *37*, 1–109.
[59] A. Geyer, G. Mueller, H. Kessler, *J. Am. Chem. Soc.* **1994**, *116*, 7735–7743.
[60] a) W. G. Rajeswaran, S. J. Hocart, W. A. Murphy, J. E. Taylor, D. H. Coy, *J. Med. Chem.* **2001**, *44*, 1416–1421; b) Y. Takeuchi, G. R. Marshall, *J. Am. Chem. Soc.* **1998**, *120*, 5363–5372; c) S. Ro, H.-J. Lee, I.-A. Ahn, D.-K. Shin, K.-B. Lee, C.-J. Yoon, Y.-S. Choi, *Bioorg. Med. Chem.* **2001**, *9*, 1837–1841.
[61] V. J. Hruby, F. al-Obeidi, W. Kazmierski, *Biochem. J* **1990**, *268*, 249–262.
[62] M. MacDonald, J. Aube, *Curr. Org. Chem.* **2001**, *5*, 417–438.
[63] M. G. Bomar, B. Song, P. Kibler, K. Kodukula, A. K. Galande, *Org. Lett.* **2011**, *13*, 5878–5881.
[64] V. Brenner, F. Piuzzi, I. Dimicoli, B. Tardivel, M. Mons, *Angew. Chem. Int. Ed Engl.* **2007**, *46*, 2463–2466.
[65] B. Imperiali, R. A. Moats, S. L. Fisher, T. J. Prins, *J. Am. Chem. Soc.* **1992**, *114*, 3182–3188.
[66] T. S. Haque, S. H. Gellman, *J. Am. Chem. Soc.* **1997**, *119*, 2303–2304.

3.3. Metalloproteins

A significant number of proteins and enzymes, approximately one-third of all known proteins, contain metal ions in their structure, mainly transition metals such as iron, cobalt, nickel, copper or zinc.[67] These are known as *metalloproteins*, whose functions are determined by their coordinated metal, or *metal-binding proteins*, in which the metal ions offer thermodynamic stability, but do not play a role in their function.[68] Metal-binding sites in proteins usually comprise a polar core, where the metal ion is coordinated by nitrogen, oxygen or sulfur centers in the side chains of hydrophilic residues or in the peptide backbone, surrounded by a shell of non-polar groups that avoid contact between water and the metal complex.[69] It is the nature of the metal ion itself that governs both the binding site on the protein, based on its coordinative preferences according to the hard-soft theory of acids and bases, as well as the coordination number and geometry, dictated by its oxidation state.[70] These metal sites in protein structures serve structural, regulatory or catalytic roles, being responsible for catalyzing major biological processes such as photosynthesis, respiration, nitrogen fixation and transport.[68,71] In fact, more comprehensively, protein-bound metal sites can be classified according to their function into five categories:

- **Structural**: the correct conformation of a protein is stabilized by coordination to the metal ion, thus involving in some cases a regulatory role. This functional category will be further discussed below.
- **Transport and storage**: refer to the uptake, binding and release of metal ions in soluble form. This group comprises three types of proteins:[72] *transferrins*, solubilize Fe(III) by binding, thus inhibiting the formation of $Fe(OH)_3$, transport and deliver the iron cation to cells;[73] *ferritins* store Fe(III) in aerobic cells, in large iron-oxo aggregates enclosed by a protein shell;[74] and *metallothioneins* which, due to their high cysteine content, sequester Zn(II) for homeostatic regulation, and also Cd(II), Hg(II) and Cu(I) for detoxification.[75]
- **Electron transfer**: metal sites act as a redox center, enabling the uptake, storage, and transfer of electrons. These are usually iron and cooper ions, whose redox potential depends both on their own properties and on the conformation of the protein around them. Electron-transfer proteins are divided into three groups:[72] *cytochromes*, which contain a heme group accommodating Fe(II) or Fe(III) and work in respiration and

[67] Y. Gao, C. Chen, Z. Chai, *J. Anal. At. Spectrom.* **2007**, *22*, 856–866.
[68] J. S. Garcia, C. S. de Magalhães, M. A. Z. Arruda, *Talanta* **2006**, *69*, 1–15.
[69] M. M. Yamashita, L. Wesson, G. Eisenman, D. Eisenberg, *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 5648–5652.
[70] A. J. Thomson, H. B. Gray, *Curr. Opin. Chem. Biol.* **1998**, *2*, 155–158.
[71] Y. Lu, N. Yeung, N. Sieracki, N. M. Marshall, *Nature* **2009**, *460*, 855–862.
[72] R. H. Holm, P. Kennepohl, E. I. Solomon, *Chem. Rev.* **1996**, *96*, 2239–2314.
[73] E. N. Baker, in *Advances in Inorganic Chemistry* (Ed.: A.G. Sykes), Academic Press, **1994**, pp. 389–463.
[74] P. M. Harrison, S. C. Andrews, P. J. Artymiuk, G. C. Ford, J. R. Guest, J. Hirzmann, D. M. Lawson, J. C. Livingstone, J. M. A. Smith, A. Treffry, S. J. Yewdall, in *Advances in Inorganic Chemistry* (Ed.: A.G. Sykes), Academic Press, **1991**, pp. 449–486.
[75] M. J. Stillman, *Coord. Chem. Rev.* **1995**, *144*, 461–511.

photosynthesis;[76] *iron-sulfur proteins*, these are nonheme iron-containing proteins characterized by the presence of sulfur, that can be separated into *rubredoxins* and *ferredoxins* based on the type of metal site;[77] and *blue copper proteins*, as the name suggests, they are redox proteins containing Cu(I)/(II) ions whose classical site is found in the *plastocyanins*, which are involved in oxygenic photosynthesis.[78]

- **Dioxygen binding**: the metal ion transiently coordinate to the dioxygen (O_2).[72] The main representatives of this group of dioxygen-carrying proteins are *hemoglobin* and *myoglobin*, both of them contain a heme iron group (tetrapyrrole cyclic compound that tightly coordinates iron) where O_2 is bound for transport in the blood and storage in muscles, respectively.[79] Additionally, there are two others dioxygen-carrying proteins that lack the heme group, these are the *hemerythrins* and *hemocyanins*, which have binuclear active sites with iron and copper, in that order.[80,81]
- **Catalytic**: the metal site enables substrate binding, activation and reaction.[72] There is a wide variety of metalloenzymes, which contain different metal ions with a key role in their active sites, and catalyze a wide range of reactions including: redox reactions, such as photosynthesis and nitrogen fixation;[71] detoxification reactions, catalyzed by *cytochromes P-450*, so called because they are hemoproteins,[82] and *superoxide dismutases*[83] among others; as well as the hydration of CO_2 to bicarbonate, HCO_3^-, catalyzed by *carbonic anhydrase* (among the fastest known enzymes), for whose function a Zn(II) ion in its active site is indispensable.[84]

Beyond these categories, other metal-binding sites have been identified in natural protein crystal structures with apparently none of these five types of function.[72] Among all these functions that can be played by protein-bound metal ions, the structural one is the most relevant for this work, hence it will be discussed in detail below.

Many biological processes are mediated by DNA-protein or protein-protein interactions, which involve specific structural domains in these macromolecules.[85] In the specific case of proteins, these peptide domains are sometimes too small to fold into their bioactive conformation by themselves, making it stabilized through binding to metal ions, which leads to an effective cross-link between specific residues with a stabilizing effect.[86]

---

[76] T. Yamanaka, K. Okunuki, in *Microbial Iron Metabolism* (Ed.: J.B. Neilands), Elsevier, San Diego, CA, **1974**, pp. 349–400.
[77] R. Cammack, in *Advances in Inorganic Chemistry* (Ed.: R. Cammack), Academic Press, **1992**, pp. 281–322.
[78] K. W. Penfield, R. R. Gay, R. S. Himmelwright, N. C. Eickman, V. A. Norris, H. C. Freeman, E. I. Solomon, *J. Am. Chem. Soc.* **1981**, *103*, 4382–4388.
[79] M. Momenteau, C. A. Reed, *Chem. Rev.* **1994**, *94*, 659–698.
[80] R. E. Stenkamp, *Chem. Rev.* **1994**, *94*, 715–726.
[81] C. J. Coates, H. Decker, *Cell. Mol. Life Sci.* **2017**, *74*, 293–317.
[82] R. Bernhardt, *J. Biotechnol.* **2006**, *124*, 128–145.
[83] J. J. P. Perry, D. S. Shin, E. D. Getzoff, J. A. Tainer, *Biochim. Biophys. Acta* **2010**, *1804*, 245–262.
[84] C. T. Supuran, *Biochem. J* **2016**, *473*, 2023–2032.
[85] A. D. de Araujo, H. N. Hoang, W. M. Kok, F. Diness, P. Gupta, T. A. Hill, R. W. Driver, D. A. Price, S. Liras, D. P. Fairlie, *Angew. Chem. Int. Ed Engl.* **2014**, *53*, 6965–6969.
[86] C. M. Gomes, P. Wittung-Stafshede, *Protein Folding and Metal Ions: Mechanisms, Biology and Disease*, CRC Press, **2016**, chapter 1.

Several examples of metal-mediated folding are found in nature, including the *zinc finger* motifs (ZFs), *calmodulin* (CaM) protein or *zinc hook* motifs (ZH), among others.

Most enzymes use the Zn(II) metal ions with catalytic roles, acting as a superacid centers.[84] However, due to its ability to form structure-stabilizing cross-links with no added chemical reactivity, zinc ions also plays strictly structural roles in a wide variety of metallobiomolecules. The most well-known example is the classical *Cys_2His_2 zinc finger motif*, found in most eukaryotic transcription factor proteins involved in nucleic acid binding and regulation of gene expression.[87] These DNA binding proteins contain several zinc finger motifs in tandem, with each of them bearing the following general sequence (Y/F)-X-Cys-X_{2-4}-Cys-X_3-F-X_5-Leu-X_2-His-X_{3-4}-His-X_{2-6} (X being any amino acid).[88] Zinc finger motifs constitute independent structural domains, consisting of two antiparallel β-strands followed by an α-helix, held together by tetrahedral coordination to the Zn(II) ion via cysteine (β-strands) and histidine (α-helix) side chains. The other conserved hydrophobic residues pack together to provide a hydrophobic core.[89]

Besides intraprotein Zn(II)-binding sites, zinc ions can also play a purely structural role in Zn(II) interprotein sites. A classic example is the *zinc hook domain*, which is a site unique for Rad50 protein, an essential component of the MRN(X) complex responsible for repairing double-stranded DNA damage.[90] Rad50 proteins comprise a complex made up of two monomeric units, linked together via their minimal zinc hook motifs, which consist of a 14 amino acids peptide sequence containing two invariant Cys residues, within a CXXC motif, able to bind Zn(II) in a tetrahiolate coordination sphere (ZnS_4), resulting in the complex's organization in a highly stable β-hairpin like structure.[91] It is noteworthy that mutation of any of the invariant Cys residues leads to loss of coordination ability to Zn(II), thus reducing cell viability.[92]

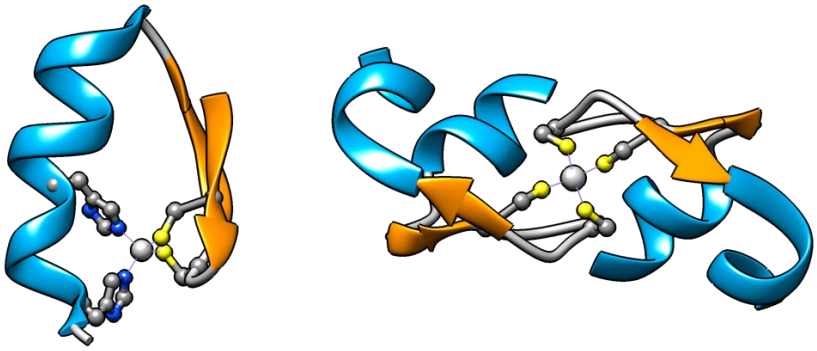


Figure 8. Left: structure of a monomeric Cys_2His_2 zinc finger (PDB 1ZNF). Right: structure of the dimeric zinc hook from protein Rad50 (PDB 1L8D).

[87] J. M. Berg, Y. Shi, *Science* **1996**, *271*, 1081–1085.
[88] E. H. Cox, G. L. McLendon, *Curr. Opin. Chem. Biol.* **2000**, *4*, 162–165.
[89] M. Padjasek, A. Kocyła, K. Kluska, O. Kerber, J. B. Tran, A. Krężel, *J. Inorg. Biochem.* **2020**, *204*, 110955.
[90] K.-P. Hopfner, L. Craig, G. Moncalian, R. A. Zinkel, T. Usui, B. A. L. Owen, A. Karcher, B. Henderson, J.-L. Bodmer, C. T. McMurray, J. P. Carney, J. H. J. Petrini, J. A. Tainer, *Nature* **2002**, *418*, 562–566.
[91] T. Kochańczyk, P. Jakimowicz, A. Krężel, *Chem. Commun.* **2013**, *49*, 1312–1314.
[92] J. J. W. Wiltzius, M. Hohl, J. C. Fleming, J. H. J. Petrini, *Nat. Struct. Mol. Biol.* **2005**, *12*, 403–407.

However, Zn(II) ion is not the only one that plays a structural function, since one of the major examples of conformational changes in metal-bound proteins is found in Ca(II) trigger proteins, such as calmodulins, which respond to Ca(II) concentration changes in the cytoplasm.[93] These proteins carry the *Ef-hand binding motif*, which consists of a helix-loop-helix topology, composed of 12 residues, including 6 highly conserved at positions 1,3,5,7,9 and 12, which, together with a single water molecule, provide seven ligands coordinating a single Ca(II) ion in a pentagonal bipyramid coordination.[94,95] The motif typically occurs in pairs, as in the case of *calmodulin* (CaM), that is formed by two lobes linked by a central helix, each of them presenting two Ef-hand motifs, which induce a conformational change upon binding to four Ca(II) ions, that enables CaM to interact with a wide variety of enzymes.[96]



Figure 9. Structure of human calmodulin protein (PDB 1CLL) containing four nearly identical high-affinity Ca(II) binding sites featuring a characteristic loop flanked by two α-helices, known as Ef-hands.

3.3.1. Synthetic metallopeptides

Peptides, short biopolymers with fewer than 50-100 amino acid residues, which are involved in many key physiological mechanisms in biological systems.[97] Over the last thirty years, a large number of biologically active peptides have been discovered and characterized, ranging from hormones to neuropeptides.[98] Their function, as well as that of short peptide sequences in proteins, is governed by the unique secondary structure acquired (α-helix and β-sheet), known as their *bioactive conformation*, which is stabilized by weak forces, including hydrogen bonding, metal coordination, electrostatic interactions or disulfide bridges.[48b,99]

Due to their relevant biological activity, and therefore their great potential as therapeutic agents (novel pharmaceuticals, vaccines, diagnostic agents...), the development of synthetic peptides that mimic the function of natural peptides is a relevant

[93] M. R. Nelson, W. J. Chazin, *Biometals* **1998**, *11*, 297–318.
[94] S. Drmota Prebil, U. Slapšak, M. Pavšič, G. Ilc, V. Puž, E. de Almeida Ribeiro, D. Anrather, M. Hartl, L. Backman, J. Plavec, B. Lenarčič, K. Djinović-Carugo, *Sci. Rep.* **2016**, *6*, 27383.
[95] Y. Kim, J. T. Welch, K. M. Lindstrom, S. J. Franklin, *J. Biol. Inorg. Chem.* **2001**, *6*, 173–181.
[96] C. Yang, G. S. Jas, K. Kuczera, *Biochim. Biophys. Acta* **2004**, *1697*, 289–300.
[97] K. Fosgerau, T. Hoffmann, *Drug Discov. Today* **2015**, *20*, 122–128.
[98] A. Giannis, T. Kolter, *Angew. Chem. Int. Ed Engl.* **1993**, *32*, 1244–1267.
[99] N. Sewald, H.-D. Jakubke, *Peptides: Chemistry and Biology*, John Wiley & Sons, New York, **2015**, chapter 3.

field at the interface of chemical biology and medicinal chemistry.[100] Indeed, since the use of insulin in therapy began in the 1920s, more than 60 peptide drugs have been commercialized.[101] However, the use of peptides as therapeutics is often limited by their low membrane permeability, rapid proteolysis, but mainly because of the inherent flexibility of short and medium-sized peptides (less than 30 residues), which complicates the stabilization of the right conformation.[102] In an attempt to improve the therapeutic potential of peptides, the natural sequences of biologically active peptides are taken as a reference followed by structural and functional modifications in order to optimize their chemical properties or stabilize their bioactive conformation.[103]

Regarding the stabilization of the secondary structure in short synthetic peptides, several strategies have been tried to constrain the flexibility of the chain: incorporation of non-natural amino acids, nucleation templates, covalent and non-covalent side chain interactions as well as metal ion coordination. Unfortunately, there is currently no consensus on the most effective strategy.[100] Nevertheless, given the above mentioned ability of polypeptides to coordinate metal ions with the consequent structure stabilization effect,[86] in this section it will be listed some of the most representative examples of synthetic peptides with metal-stabilized conformations that have been developed since the beginning of the 1990s.

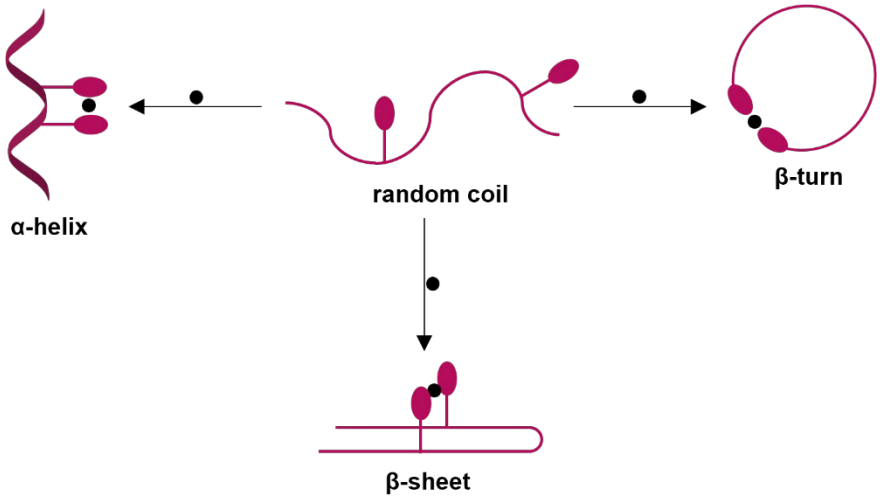


Figure 10. Scheme of the metal-binding-induced stabilization of secondary structures in artificial peptides.

3.3.1.1. α-Helical metallopeptides

Numerous biological processes are regulated by DNA-protein or protein-protein interactions that mainly involve a protein α-helix structure of 1 to 4 helical turns, equivalent to a peptide sequence of 4 to 15 residues, so the development of short synthetic

[100] N. E. Shepherd, H. N. Hoang, G. Abbenante, D. P. Fairlie, *J. Am. Chem. Soc.* **2005**, *127*, 2974–2983.
[101] J. L. Lau, M. K. Dunn, *Bioorg. Med. Chem.* **2018**, *26*, 2700–2707.
[102] a) Y. A. Haggag, A. A. Donia, M. A. Osman, S. A. El-Gizawy, *Biomed. J.* **2018**, *1*; b) S. Marqus, E. Pirogova, T. J. Piva, *J. Biomed. Sci.* **2017**, *24*, 21; c) C. Adessi, C. Soto, *Curr. Med. Chem.* **2002**, *9*, 963–978.
[103] a) A. Begum, D. Sujatha, K. V. S. R. G. Prasad, K. Bharathi, *Asian J. Chem.* **2017**, *29*, 1879–1887; b) H. M. Werner, C. C. Cabalteja, W. S. Horne, *Chembiochem* **2016**, *17*, 712–718.

peptides that mimic the structure of these α-helices, with the aim of regulating these molecular processes, is a major challenge in chemical biology.[85,104] Accordingly, different short peptides mimicking the structure of those α-helical recognition motifs in proteins have been synthesized, by stabilizing their structure through coordination to different metal ions via the side chain of natural amino acids (histidine, cysteine, aspartic or glutamic acid) or non-natural coordinating residues.

a) Monomeric α-helices

In 1990, Ghadiri *et al.* reported a simple method to stabilize short α-helical peptides in water, by introducing histidine and cysteine residues at positions *i*, *i*+4 into the peptide sequence to coordinate to transition metal ions (Cu(II), Zn(II), Cd(II) and Ni(II)). The resulting bidentate complexes constrain the peptide backbone in an α-helical structure.[105] With the aim of expanding the scope and usefulness of their methodology, they prepared highly stable α-helical peptides using an exchange-inert metal complex. For this purpose, they synthesized the acetyl-AEAAAKEAAAK**H**AAA**H**A-CONH$_2$ **1** and acetyl-AEAAAKHAAA**H**EAAAKA-CON**H**$_2$ **2** peptides, that showed a random coil structure in water but, upon coordination to the $[Ru(NH_3)_4]^{2+}$ complex, folded into a stable α-helical conformation.[106]

Likewise, using histidine residues at positions *i* and *i*+4, in 2000 Fairlie and co-workers reported the first known stable single α-helical turn in solution. They synthesized the pentapeptide Ac-His-Ala-Ala-Ala-His-NH$_2$ and folded it into an α-helical turn upon coordination to a $Pd(en)^{2+}$ metal clip.[107] To extend the application of the $[Pd(en)(H^*xxxH^*)]^{2+}$ macrocycle, Fairlie *et al.* proved the effective stabilization of 3- and 4-α turns in larger unstructured peptides, 10 to 15 residues, involving residues from the Zn(II)-binding α-helical domain in thermolysin, through the in situ synthesis of multiples of these Pd(II) cyclic metallopeptides.[108]

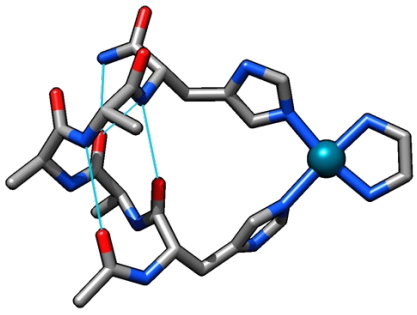


Figure 11. NMR structure of the pentapeptide Ac-His-Ala-Ala-Ala-His-NH$_2$ showing an α-helical turn conformation stabilized via histidine coordination to $Pd(en)^{2+}$ (intrahelical hydrogen bonds in light blue).[107]

[104] G. Dimartino, D. Wang, R. N. Chapman, P. S. Arora, *Org. Lett.* **2005**, *7*, 2389–2392.
[105] M. R. Ghadiri, C. Choi, *J. Am. Chem. Soc.* **1990**, *112*, 1630–1632.
[106] M. R. Ghadiri, A. K. Fernholz, *J. Am. Chem. Soc.* **1990**, *112*, 9633–9635.
[107] M. J. Kelso, H. N. Hoang, T. G. Appleton, D. P. Fairlie, *J. Am. Chem. Soc.* **2000**, *122*, 10488–10489.
[108] a) M. J. Kelso, H. N. Hoang, W. Oliver, N. Sokolenko, D. R. March, T. G. Appleton, D. P. Fairlie, *Angew. Chem. Int. Ed.* **2003**, *42* 421–424; b) M. J. Kelso, R. L. Beyer, H. N. Hoang, A. S. Lakdawala, J. P. Snyder, W. V. Oliver, T. A. Robertson, T. G. Appleton, D. P. Fairlie, *J. Am. Chem. Soc.* **2004**, *126*, 4828–4842.

The group of J. L. Mascareñas has recently applied this methodology for the synthesis of a peptide derived from the basic region (br) of the **GCN4 bZIP** transcription factor equipped with two histidine residues (**brHis$_2$**) at strategic *i* and *i*+4 positions; **brHis$_2$** is unable to bind to its DNA target. Addition of $PdCl_2$(en) leads to high-affinity and selective binding to DNA, which can be reverted with an excess of palladium chelator and induces efficient internalization in living mammalian cells. On this basis, they proved that once internalized, this stapled metallopeptide is catalytically active, conducting palladium-promoted depropargylation reactions in cells.[109] Using the same approach, in 2022, our group reported the first switchable KRAS oncoprotein binding agent, which is a peptide derived from the αH helix of its **SOS1** cofactor, modified with two histidines that coordinate Pd(II), to stabilize the helical conformation to increase the binding affinity for KRAS, also reversible in the presence of metal chelator.[110]

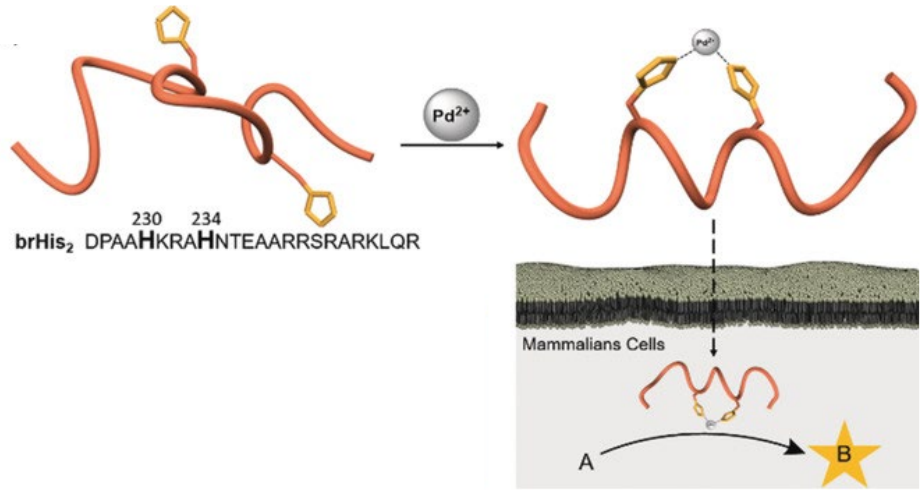


Figure 12. Stabilization of the α-helical structure in **brHis$_2$** upon Pd(II) coordination. The stapled metallopeptide is able to internalize in mammalian cells and catalyze depropargylation reactions.[109b]

The same concept has been used in the α-helix stabilization using non-natural coordinating residues. In this field, Sasaki and Hopkins took as a model the EDTA ligand, a chelating agent that binds to a large number of metal ions with high stability and introduced two non-natural residues with aminodiacetic acid (EDTA analogue) bearing side chains, separated by one turn in the α-helix. Simultaneous coordination of the two unnatural aminodiacetic acid residues to a single transition metal ion (Co(II), Ni(II), Cu(II), Zn(II) and Cd(II)) forms a labile cross-link that stabilizes or directs the folding of the peptide into an α-helical conformation, as shown by the CD spectra.[111]

In all these examples, metal binding is used to stabilize helical structures, however, the opposite effect, metal chelation-induced helix destabilization of stable

[109] a) S. Learte-Aymamí, N. Curado, J. Rodríguez, M. E. Vázquez, J. L. Mascareñas, *J. Am. Chem. Soc.* **2017**, *139*, 16188–16193; b) S. Learte-Aymamí, C. Vidal, A. Gutiérrez-González, J. L. Mascareñas, *Angew. Chem. Int. Ed.* **2020**, *59*, 9149–9154.

[110] S. Learte-Aymamí, P. Martin-Malpartida, L. Roldán-Martín, G. Sciortino, J. R. Couceiro, J.-D. Maréchal, M. J. Macías, J. L. Mascareñas, M. E. Vázquez, *Commun Chem* **2022**, *5*, 75.

[111] a) F. Ruan, Y. Chen, P. B. Hopkins, *J. Am. Chem. Soc.* **1990**, *112*, 9403–9404; b) F. Ruan, Y. Chen, K. Itoh, T. Sasaki, P. B. Hopkins, *J. Org. Chem.* **1991**, *56*, 4347–4354.

peptide helices, can be useful for switching protein function. Sugiura *et al.* synthesized a 17-residue helical peptide, provided with two iminodiacetic acid derivatives of lysine placed at *i*, *i*+2 positions, whose helical structure can be controlled by oxidation of Fe(II) to Fe(III), resulting in a decrease of the helical content and, subsequently, the modulation of its binding affinity to the target molecule.[112]

To achieve a methodology that provides selective catalysts, Gilbertson and co-workers developed a system for the solid-phase peptide synthesis of sequences equipped with phosphine ligands, that can coordinate metal ions and adopt an α-helical conformation. This approach takes advantage of the predictable structure of the peptide ligands to control the selectivity of the catalytic reactions carried out by the transition metal bound to phosphines. In their first example, they synthesized a 12-residue peptide containing two diphenylphosphinoserines at *i* and *i*+4, which upon biding to Rh(I) ion assumes an α-helical structure, being able to catalyze the olefin hydrogenation reaction in solid phase.[113]

Beyond the stabilization of the α-helical conformation in short peptides, Tezcan *et al*. have recently developed a new strategy that, apart from inducing α-helicity in unstructured peptides using metal-mediated crosslinkinks across two helix turns, also allows the incorporation of metal-based functionalities, such as luminescence, catalysis or oligomerization, modulated by the type of chelating ligand, specific metal, or pH. To this end, they designed hybrid metal coordination motifs (**HCMs**), which are tridentate metal coordination modules consisting of a natural coordinating residue, placed at *i* position, and a synthetic bidentate chelating group attached to the side chain at *i*+7. In 2013, they reported the synthesis of a 10-residue peptide with an (*i*, *i*+7) His/8-hydroquinoline (Quin) HCM motif that binds divalent cations, such as Co(II), Ni(II), Cu(II) or Zn(II). The α-helix content of the resulting metal complexes is dependent on the length and flexibility of the linker, as well as the chosen metal ion. Additionally, this approach allows simultaneous generation of luminescence upon coordination to Zn(II) or $Re^{I}(Quin)(CO)_3$ moiety.[114] To increase the family of HCM motifs, the 8-quinoline group was changed to a phenanthroline group, leading to a new peptide ligand that, upon binding to divalent metal ions, is able to stabilize an α-helical structure. It is important to highlight the greater resistance to proteolysis of the resulting metallopeptides. They proved the usefulness of their approach by introducing an HCM into the peptide sequence of the DNA recognition domain of a bZIP protein, that was able to dimerize in a α-helical conformation to bind its target DNA.[115]

b) α-Helical coiled coils

The *coiled coil* is a prevalent structural motif found in a wide range of proteins, typically consisting of two to five amphipathic α-helices that wind around each other into

[112] S. Futaki, T. Kiwada, Y. Sugiura, *J. Am. Chem. Soc.* **2004**, *126*, 15762–15769.
[113] a) S. R. Gilbertson, G. Chen, M. McLoughlin, *J. Am. Chem. Soc.* **1994**, *116*, 4481–4482; b) S. R. Gilbertson, X. Wang, G. S. Hoge, C. A. Klug, J. Schaefer, *Organometallics* **1996**, *15*, 4678–4680.
[114] S. J. Smith, K. Du, R. J. Radford, F. A. Tezcan, *Chem. Sci.* **2013**, *4*, 3740–3747.
[115] S. J. Smith, R. J. Radford, R. H. Subramanian, B. R. Barnett, J. S. Figueroa, F. A. Tezcan, *Chem. Sci.* **2016**, *7*, 5453–5461.

a left-handed supercoil.[116] This structure is mainly stabilized by interchain packing of hydrophobic side chains along the interface, in addition to interhelical electrostatic interactions between complementary charged residues. The opposite coiling of the right-handed α-helices and the left-handed supercoil lowers the number of residues per helical from 3.6 to 3.5 to optimize the packing of the side chains in a "knobs-into-holes manner", as described by F. Crick in 1953.[117,118] Each helix in the coiled coil exhibits a seven-residue pattern (*abcdefg*)*n*, known as *heptad repeat*, where *a* and *d* positions usually correspond to the nonpolar residues at the interface (L, V, I), while *e* and *g* are occupied by charged residues exposed to the solvent (e.g. E, K), maintaining interhelical electrostatic interactions. The other hydrophilic *b*, *c* and *f* positions, are placed on the helical surfaces exposed to the solvent.[119]

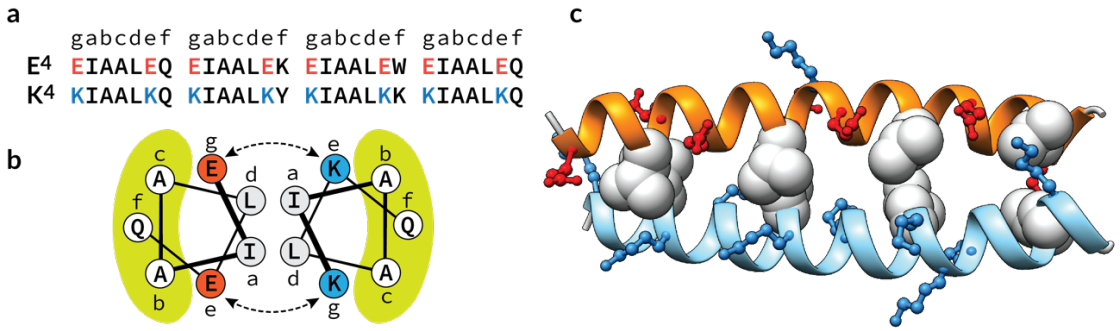


Figure 13. a) Peptide sequences of a designed heterodimer coiled-coil; b) helical wheel diagram showing the disposition of the residues along the axis of the α-helices; c) model structure of the dimeric coiled-coil, where the hydrophobic Leu/Ile at the interface are highlighted in grey, with the charged residues, Glu and Lys, are in red and blue respectively.

Coiled-coil domains in proteins play major biological functions,[120] and regulate many specific protein-protein interactions, connecting, spacing and orienting different functional domains in cells dynamically.[120,121] The importance and diversity of their functions, together with their simple and well-characterized structure, make of coiled coils a versatile model for *de novo* protein design.[122] In this context, with a view to developing a system with metal-regulated assembly, Tanaka *et al*. were inspired by the sequence of the de novo designed triple-stranded parallel α-helical coiled coil **IZ**, [YGG($I_dE_eK_fK_gI_aE_bA_c$)$_4$], where two Ile at *d* and *a* positions were replaced by His residues, yielding an octahedral metal binding site within the hydrophobic core of the coiled coil. The new peptide displays a random coil structure in aqueous solution, but coordination with transition metal ions such as Co(II), Ni(II), Cu(II) or Zn(II) leads to the

[116] J. M. Mason, K. M. Arndt, *Chembiochem* **2004**, *5*, 170–176.
[117] C. Cohen, D. A. Parry, *Proteins* **1990**, *7*, 1–15.
[118] F. H. C. Crick, IUCr, *Acta Crystallogr.* **1953**, *6*, 689–697.
[119] a) E. K. O'Shea, K. J. Lumb, P. S. Kim, *Curr. Biol.* **1993**, *3*, 658–667; b) T. J. Graddis, D. G. Myszka, I. M. Chaiken, *Biochemistry* **1993**, *32*, 12664–12671.
[120] H. Robson Marsden, A. Kros, *Angew. Chem. Int. Ed Engl.* **2010**, *49*, 2988–3005.
[121] A. Rose, I. Meier, *Cell. Mol. Life Sci.* **2004**, *61*, 1996–2009.
[122] X. I. Ambroggio, B. Kuhlman, *J. Am. Chem. Soc.* **2006**, *128*, 1154–1161.

formation of a triple-stranded parallel α-helical coiled coil.[123,124] In the study of metal-binding selectivity, two Ile rediues in **IZ** were replaced by Cys and Ala residues, leading to the stabilization of a triple-stranded coiled coil conformation in the presence of soft metals ions such as Cu(II) ,Cd(II) or Hg(II).[125]

3.3.1.2. β-hairpin meatallopeptides

β-sheet structures are also involved in protein-protein interactions, although to a much lesser degree than α-helices. Actually, β-sheets have been found mediating interactions associated with autoimmune diseases, cancer or neurodegenerative diseases, such as Alzheimer's disease. Given this context, designed β-sheets emerge as promising inhibitors of these undesired PPIs.[126] Unfortunately, the development of artificial β-sheets is complicated because of their propensity for self-aggregation and high context-dependent stability.[127] To overcome these drawbacks, different strategies have been developed, such as the use of β-strand or β-turn templates for nucleation, [128] β-strand mimics bearing unnatural amino acids,[129] and β-hairpin peptides cycled through disulfide-bridge[130] or triazolyl-bridge.[131] Metal ions have also been employed to induce and stabilize β-hairpin structures, as a turn-nuclei or as an intramolecular cross-linker.

An early example of metal-mediated β-sheet stabilization was reported by Schneider and Kelly in 1995, who introduced a coordinating residue derived from bipyridine, acting as a β-sheet nucleator, at *i*+1 and *i*+2 positions of a β-turn; upon Cu(II) coordination in a square planar symmetry, these residues successfully enforce a reverse-turn conformation, leading to an anti-parallel β-hairpin folding of the peptide strands.[132] Likewise, Searle and co-workers synthesized two different metalloproteins based on an artificial 16-residue peptide that folds into a β-hairpin (K**K**YTVSINGKKKITV**S**I).[133] In one of the derived peptides, Lys and Ser residues cross-linking two strands at N- and C-termini, respectively (in bold in the sequence), are replaced by two His residues to provide a Zn(II)-binding site. For the second derived peptide, a third His residue is appended to the N-terminus to improve Zn(II) coordination in a tetrahedral geometry, saturated with the carboxylate-terminal group or a solvent molecule, thus mimicking the catalytic site in carbonic anhydrase. Both peptides display a partially pre-organized β-hairpin

[123] K. Suzuki, H. Hiroaki, D. Kohda, H. Nakamura, T. Tanaka, *J. Am. Chem. Soc.* **1998**, *120*, 13008–13015.
[124] T. Tanaka, T. Mizuno, S. Fukui, H. Hiroaki, J.-I. Oku, K. Kanaori, K. Tajima, M. Shirakawa, *J. Am. Chem. Soc.* **2004**, *126*, 14023–14028.
[125] T. Kiyokawa, K. Kanaori, K. Tajima, T. Tanaka, *Biopolymers* **2000**, *55*, 407–414.
[126] J. Laxio Arenas, J. Kaffy, S. Ongeri, *Curr. Opin. Chem. Biol.* **2019**, *52*, 157–167.
[127] D. L. Minor Jr, P. S. Kim, *Nature* **1994**, *367*, 660–663.
[128] J. P. Schneider, J. W. Kelly, *Chem. Rev.* **1995**, *95*, 2169–2187.
[129] M. A. Etienne, J. P. Aucoin, Y. Fu, R. L. McCarley, R. P. Hammer, *J. Am. Chem. Soc.* **2006**, *128*, 3522–3523.
[130] A. Spanopoulou, L. Heidrich, H.-R. Chen, C. Frost, D. Hrle, E. Malideli, K. Hille, A. Grammatikopoulos, J. Bernhagen, M. Zacharias, G. Rammes, A. Kapurniotu, *Angew. Chem. Int. Ed Engl.* **2018**, *57*, 14503–14508.
[131] L. E. Hanold, K. Oruganty, N. T. Ton, A. M. Beedle, N. Kannan, E. J. Kennedy, *PLoS One* **2015**, *10*, e0118796.
[132] J. P. Schneider, J. W. Kelly, *J. Am. Chem. Soc.* **1995**, *117*, 2533–2546.
[133] a) A. J. Maynard, G. J. Sharman, M. S. Searle, *J. Am. Chem. Soc.* **1998**, *120*, 1996–2007; b) S. R. Griffiths-Jones, A. J. Maynard, M. S. Searle, *J. Mol. Biol.* **1999**, *292*, 1051–1069.

conformation at pH 7, being destabilized at acidic pH due to electrostatic repulsions between the protonated His residues, together with protonation of the carboxyl terminus. Zn(II) chelation via residues placed near the amino- and carboxyl- terminal, enhanced the stability of the β-sheet structure, with the most pronounced effect seen in the metallopeptide bearing three His residues.[134]

Attempting to reproduce the biological activity of the naturally occurring cyclopeptides *segetalin A* and *B*, Albrecht et *al.* synthesized a new metallacyclopeptide composed of the WAG sequence, corresponding to the active part in the natural product, equipped with 2,3-dihydroxybenzoic acid attached to the amino terminus, and 2,3-dihydroxybenzylamine coupled to the C-terminal. In the presence of *cis*-dioxomolybdenum bis(acetylacetonate), the ligand yields a dicatechol Mo(VI) complex, that effectively fixes the short peptide in the same loop-type conformation as segetalins.[135]

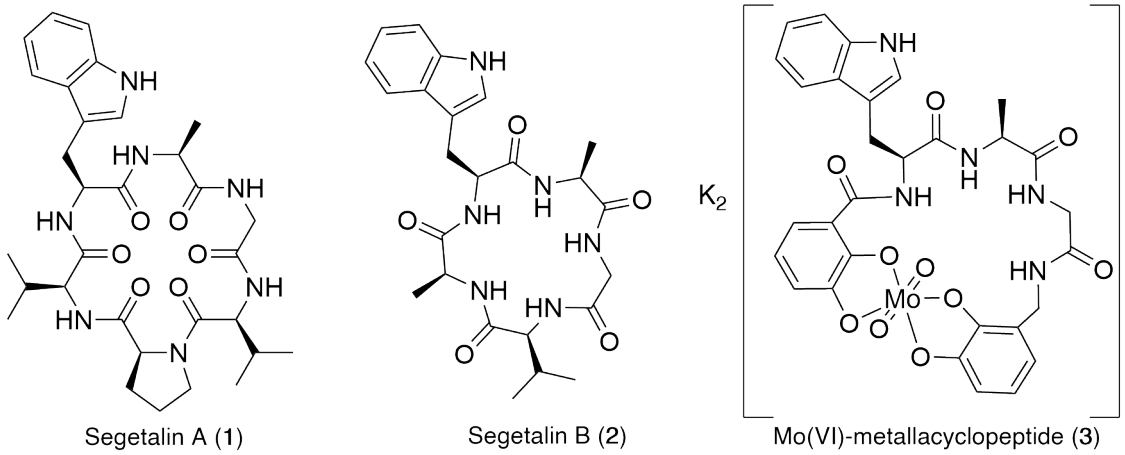


Figure 14. Structure of the native cyclopeptides *segetalin A* (**1**) and *B* (**2**) and the synthetic Mo(VI)-metallacyclopeptide (**3**), corresponding to their bioactive site, featuring the same loop-type conformation.[135]

The stabilization of loops has also been explored by the group of B. Imperiali, who in 1993 published one of the first examples of a β-hairpin structure induced by metal coordination. They placed a sequence, that has been proven to form β-turns (Val-Pro-(D)-Ser-His), in the core of a peptide containing His or Glu coordinating residues to promote Zn(II) binding.[136] Later on, the same approach was followed in the synthesis of Zn(II) fluorescent chemosensors based on short peptides that incorporate the Zn(II) chelation-sensitive fluorophore, 8-hydroxy-5-(*N*,*N*-dimethylsulfonamido)-2-methylquinoline, plus an additional coordinating residue, linked by a β-turn sequence that preorganizes the metal-binding site, favoring the Zn(II) coordination and subsequent detection.[137]

[134] G. Platt, C.-W. Chung, M. S. Searle, *Chem. Commun.* **2001**, 1162–1163.
[135] M. Albrecht, P. Stortz, P. Weis, *Supramol. Chem.* **2003**, *15*, 477–483.
[136] B. Imperiali, T. M. Kapoor, *Tetrahedron* **1993**, *49*, 3501–3510.
[137] M. D. Shults, D. A. Pearce, B. Imperiali, *J. Am. Chem. Soc.* **2003**, *125*, 10591–10597.

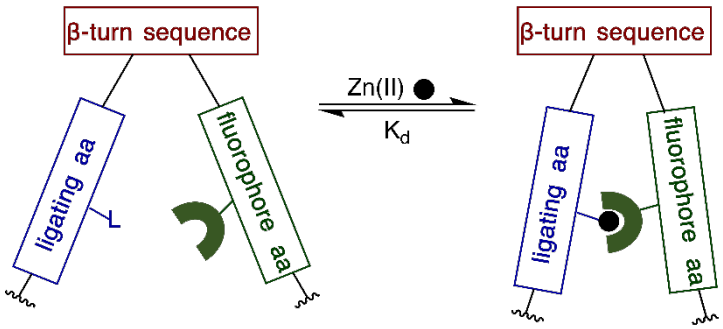


Figure 15. Scheme of the Zn(II) fluorescent chemosensor mechanism. The peptide is preorganized into a type II β-turn, promoting the Zn(II) binding to the quinoline ligand, thus switching its fluorescence emission.[137]

Recently, our group developed a luminescent sensor for BIV *TAR* RNA, inspired by the sequence of a natural peptide (**Tat**) which folds into a β-hairpin structure upon binding to the target RNA. In the design, an EDTA chelator was coupled to the N-terminal end of Tat, and a phenanthroline chromophore was attached to the side chain of a C-terminal Lys. In the presence of *TAR* RNA, the peptide folds into a β-hairpin, which promotes the coordination to Tb(III) or Eu(III) through the EDTA and the phenanthroline antenna, leading to effective sensitization of the complexed lanthanide ion and an increase in its luminescence.[138]

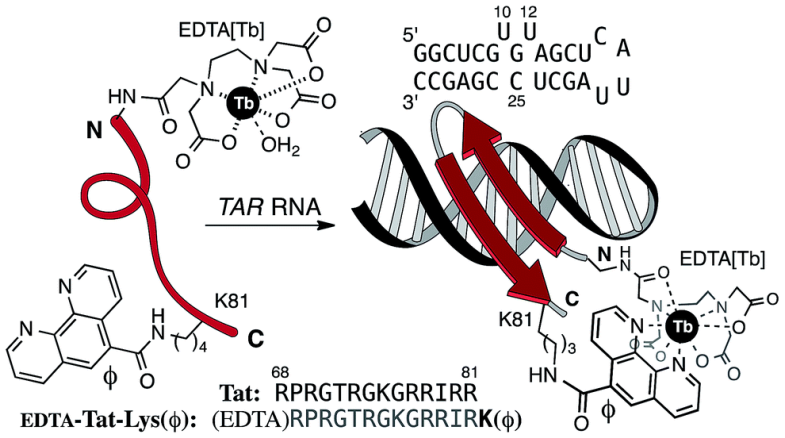


Figure 16. Outline of the *TAR* RNA sensing mechanism: Upon binding to the target RNA, the peptide folds into a β-hairpin structure that promotes the coordination of the phenanthroline antenna to the EDTA(Tb) complex.[138]

In 2018, the group of M. Fujita reported the first example of a *de novo* synthesized β-barrel tertiary structure, featuring a hydrophobic pore, with potential for channeling or storing small molecules. For this purpose, they designed an eight-residue peptide sequence combining the β-sheet-prone sequence FVFV with the PGP loop sequence, connected via a 3-amino benzoyl as a spacer, and bearing two pyridine metal-binding sites at the carboxyl and amino terminus. The Zn(II) addition results in a macrocycle with an antiparallel β-sheet structure, which self-assembles into a cylindrical β-barrel made up of three metallomacrocycles.[139]

[138] C. Penas, J. L. Mascareñas, M. E. Vázquez, *Chem. Sci.* **2016**, *7*, 2674–2678.
[139] M. Yamagami, T. Sawada, M. Fujita, *J. Am. Chem. Soc.* **2018**, *140*, 8644–8647.

3.3.1.3. 2,2'-Bipyridine as non-natural coordinating residue in artificial metallopeptides

The 2,2'-bipyridine unit (Bpy) is a privileged ligand in coordination and supramolecular chemistry;[140] its excellent coordinative properties arise from the presence of the two nitrogen donors, which provide a chelating effect, coupled with its ability to behave as a π-acceptor.[141] Moreover, owing to its frontier Lewis base character, Bpy binds with high affinity and kinetically inert metal complexes that exhibit attractive photophysical properties, such as luminescence. These unique features, confer great potential for Bpy to be used as a non-natural coordinating residue in the synthesis of new metallopeptides, being functionalized with carboxylic acid or amino groups, according to its location on the peptide strand.

a) Bpy ligand at the N-terminus

To attach a residue derived from Bpy to the amino end of the peptide, the aromatic rings are functionalized with one or two carboxylic acid groups, leading to the corresponding mono- or di-substituted derivatives.

Monosubstituted Bipyridines

Carboxylic acid group can be incorporated at four different positions in the aromatic rings of Bpy (Bpac), thus yielding a coordinating residue that can be coupled to the N-terminal end of any peptide chain using the standard protocols of solid-phase peptide synthesis. The 4Bpac and 5Bpac derivatives (Figure 17) are the most widely employed in peptide ligands derived from Bpy.[142,143]

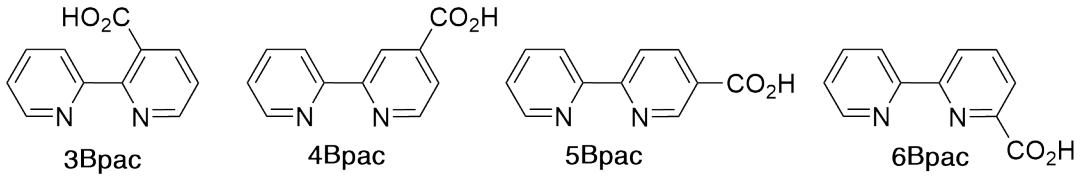


Figure 17. Non-natural coordinating residues derived from monosubstituted Bpy.

The group of R. Ghadiri has been a pioneer in the use of these Bpac ligands, since in 1992 they reported the synthesis of a 15-residue metallopeptide equipped with the

[140] a) W. W. Brandt, F. P. Dwyer, E. D. Gyarfas, *Chem. Rev.* **1954**, *54*, 959–1017; b) C. Kaes, A. Katz, M. W. Hosseini, *Chem. Rev.* **2000**, *100*, 3553–3590.

[141] F. A. Cotton, G. Wilkinson, *Advanced inorganic chemistry.* John Wiley & Sons**,** New York, NY, **1988**, pp. 682-728.

[142] a) W. S. Aldridge 3rd, B. J. Hornstein, S. Serron, D. M. Dattelbaum, J. R. Schoonover, T. J. Meyer, *J. Org. Chem.* **2006**, *71*, 5186–5190; b) T. Koide, M. Yuguchi, M. Kawakita, H. Konno, *J. Am. Chem. Soc.* **2002**, *124*, 9388–9389; c) D. G. McCafferty, B. M. Bishop, C. G. Wall, S. G. Hughes, S. L. Mecklenberg, T. J. Meyer, B. W. Erickson, *Tetrahedron* **1995**, *51*, 1093–1106; d) D. J. Wilger, S. E. Bettis, C. K. Materese, M. Minakova, G. A. Papoian, J. M. Papanikolas, M. L. Waters, *Inorg. Chem.* **2012**, *51*, 11324–11338; e) I. Hamachi, S. Tanaka, S. Tsukiji, S. Shinkai, S. Oishi, *Inorg. Chem.* **1998**, *37*, 4380–4388; f) J. G. Vos, J. M. Kelly, *Dalton Trans.* **2006**, 4869–4883.

[143] a) L. Roy, M. A. Case, *J. Phys. Chem. B* **2011**, *115*, 2454–2464; b) M. H. Filby, J. Muldoon, S. Dabb, N. C. Fletcher, A. E. Ashcroft, A. J. Wilson, *Chem. Commun.* **2011**, *47*, 559–561;c) H. K. Munch, S. T. Heide, N. J. Christensen, T. Hoeg-Jensen, P. W. Thulstrup, K. J. Jensen, *Chem. Eur. J.* **2011**, *17*, 7198–7204.

5Bpac ligand at the N-terminal end, which displays a random coil structure in the absence of metal and, upon binding to divalent transition metal ions—Fe(II), Ni(II) or Ru(II)—folds into a three-stranded α-helical coiled coil.[144]

Disubstituted Bipyridines

Ligands derived from Bpy substituted with two carboxylic acid groups (Bpda) in all symmetrical positions (except for 3,3'-Bpy) (Figure 18), have been extensively utilized in coordination chemistry, and in the synthesis of new metallopeptides, acting as linkers between the amino ends of two peptide strands.[145,146]

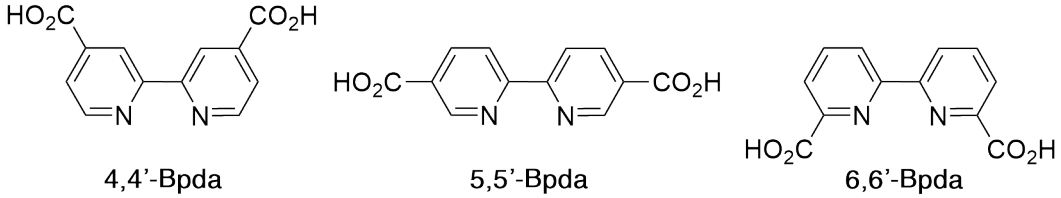


Figure 18. Non-natural coordinating residues derived from Bpy disubstituted with carboxilyc acids.

Our group published in 2010 a work in collaboration with J.-C. G. Bünzli describing the synthesis of a new heterometallic peptide complex in which a 5,5'-Bpda residue coordinates a Ru(II) ion and connects two peptide strands that in turn coordinate to Ln(III) ions through DOTA residues attached to their C-terminus (Figure 19).[147]

Yashima and co-workers employed the same 5,5'-Bpda ligand in the synthesis of a dendritic metallopeptide assembled by Fe(II) coordination into a $[Fe(5,5'\text{-Bpda})_3]^{2+}$ complex, whose chirality is controlled by the stereochemistry of the amino acids.[148]

[144] M. R. Ghadiri, C. Soares, C. Choi, *J. Am. Chem. Soc.* **1992**, *114*, 825–831.
[145] a) K.-H. Chang, J.-H. Liao, C.-T. Chen, B. K. Mehta, P.-T. Chou, J.-M. Fang, *J. Org. Chem.* **2005**, *70*, 2026–2032; b) M. Lieberman, T. Sasaki, *J. Am. Chem. Soc.* **1991**, *113*, 1470–1471; c) J. Ohkanda, R. Satoh, N. Kato, *Chem. Commun.* **2009**, *45*, 6949–6951; d) D. Hesek, Y. Inoue, S. R. Everitt, H. Ishida, M. Kunieda, M. G. Drew, *Inorg. Chem.* **2000**, *39*, 308–316; e) H. Ishida, Y. Inoue, *Biopolymers* **2000**, *55*, 469–478; f) H. Mürner, P. Belser, A. von Zelewsky, *J. Am. Chem. Soc.* **1996**, *118*, 7989–7994; g) S. Watanabe, O. Onogawa, Y. Komatsu, K. Yoshida, *J. Am. Chem. Soc.* **1998**, *120*, 229–230.
[146] a) D. R. Ahn, T. W. Kim, J. I. Hong, *J. Org. Chem.* **2001**, *66*, 5008–5011; b) E. C. Constable, E. L. Dunphy, C. E. Housecroft, M. Neuburger, S. Schaffner, F. Schaper, S. R. Batten, *Dalton Trans. J. Inorg. Chem.* **2007**, 4323–4332.
[147] M. Vázquez López, S. V. Eliseeva, J. M. Blanco, G. Rama, M. R. Bermejo, M. E. Vázquez, J.-C. G. Bünzli, *Eur. J. Inorg. Chem*. **2010**, 4532-4545.
[148] N. Ousaka, Y. Takeyama, H. Iida, E. Yashima, *Nat. Chem.* **2011**, *3*, 856–861.

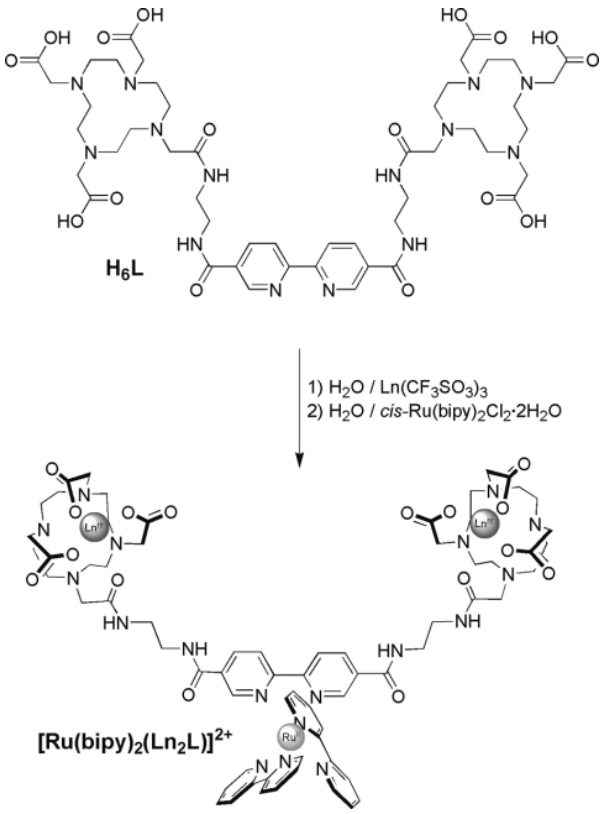


Figure 19. Synthesis of the heterometallic lanthanide-ruthenium Ln_2LRu derived from the 5,5'-Bpda coordinating residue.[147]

b) Bpy ligand as non-natural amino acids

All the previously mentioned Bpy-derived ligands can only be attached at the amino end of the peptide sequence, which largely limits their use in the design of new metallopeptides. However, different strategies have been described to obtain Bpy amino acid derivatives, which allows the Bpy to be coupled at any position within the peptide chain using the conventional SPPS methodology.

Bipyridine in α-amino acid side chains

There are many reported examples of non-natural amino acid synthesis featuring a Bpy unit in their side chain (Bpa), attached through positions 4, 5, or 6 of their aromatic rings (Figure 20).[149] The main drawback of these amino acids, is the difficulty associated with their stereoselective synthesis.[150] However, these amino acids are powerful building blocks for the synthesis of metallopeptides. Thus, Barbara Imperiali and co-workers

[149] a) B. Imperiali, S. L. Fisher, *J. Org. Chem.* **1992**, *57*, 757–759; b) B. Imperiali, T. J. Prins, S. L. Fisher, *J. Org. Chem.* **1993**, *58*, 1613–1616.

[150] a) N. Park, J. Ryu, S. Jang, H. S. Lee, *Tetrahedron* **2012**, *68*, 4649–4654; b) T. H. D. Nguyen, K. Ozawa, M. Stanton-Cook, R. Barrow, T. Huber, G. Otting, *Angew. Chem. Int. Ed Engl.* **2011**, *50*, 692–694; c) H. S. Lee, P. G. Schultz, *J. Am. Chem. Soc.* **2008**, *130*, 13194–13195; d) J. Xie, W. Liu, P. G. Schultz, *Angew. Chem. Int. Ed.* **2007**, *46*, 9239–9242; e) K. L. Haas, K. J. Franz, *Chem. Rev.* **2009**, *109*, 4921–4960; g) S. R. Wilson, A. Yasmin, Y. Wu, *J. Org. Chem.* **1992**, *57*, 6941–6945.

published the successful stabilization of a partially preorganized β-hairpin structure by Zn(II) or Co(II) coordination via two unnatural α-amino acids of 6Bpa.[151]

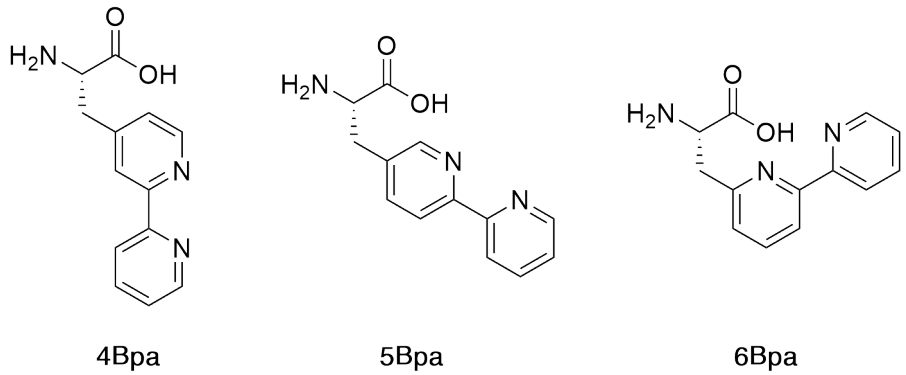


Figure 20. Non-natural amino acids bearing a Bpy in the side chain.

Asymmetric bipyridines (5Bpy and 4Bpy)

The asymmetric derivatization of Bpy unit in the form of amino acids such as 5Bpy and 4Bpy (Figure 21), allows the incorporation of the ligand into the peptide backbone, enhancing the conformational coupling between the metal ion complex and the backbone of the peptide chain.[152,153]

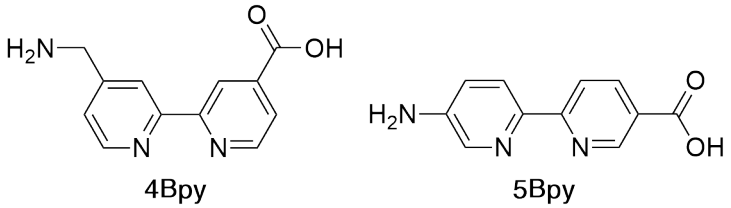


Figure 21. Non-natural amino acids derived from asymmetrically substituted Bpy with amino and carboxylic acid groups at positions 5 or 4.

The synthetic route for the non-natural coordinating amino acid 5Bpy, was developed in the group of B. Imperiali in 1996. To obtain this Bpy unit, they performed a Stille cross-coupling reaction between a nitro-functionalized pyridine and a second pyridine derivative with a methyl ester group. Reduction of the nitro group, using a catalytic hydrogenation, yields the corresponding amine, which is then reacted to afford the final Fmoc-amino acid, qualified for the SPPS methodology (Figure 22). A set of peptide ligands derived from the 5Bpy residue were synthesized, which are also provided with a fluorophore and a quencher at their N- and C-terminal ends, respectively. Formation of the corresponding metallopeptides could be characterized by monitoring the quenching effect on the fluorophore emission resulting from metal ion-induced conformational change.[154]

[151] R. P. Cheng, S. L. Fisher, B. Imperiali, *J. Am. Chem. Soc.* **1996**, *118*, 11349-11356.
[152] a) B. M. Bishop, D. G. McCafferty, B. W. Erickson, *Tetrahedron,* **2000**, *56*, 4629-4638; b) K. Heinze, K. Hempel, *Chem. Eur. J.* **2009**, *15*, 1346–1358.
[153] G. R. Newkome, J. Gross, A. K. Patri, *J. Org. Chem.* **1997**, *62*, 3013–3014.
[154] A. Torrado, B. Imperiali, *J. Org. Chem.* **1996**, *61*, 8940–8948.

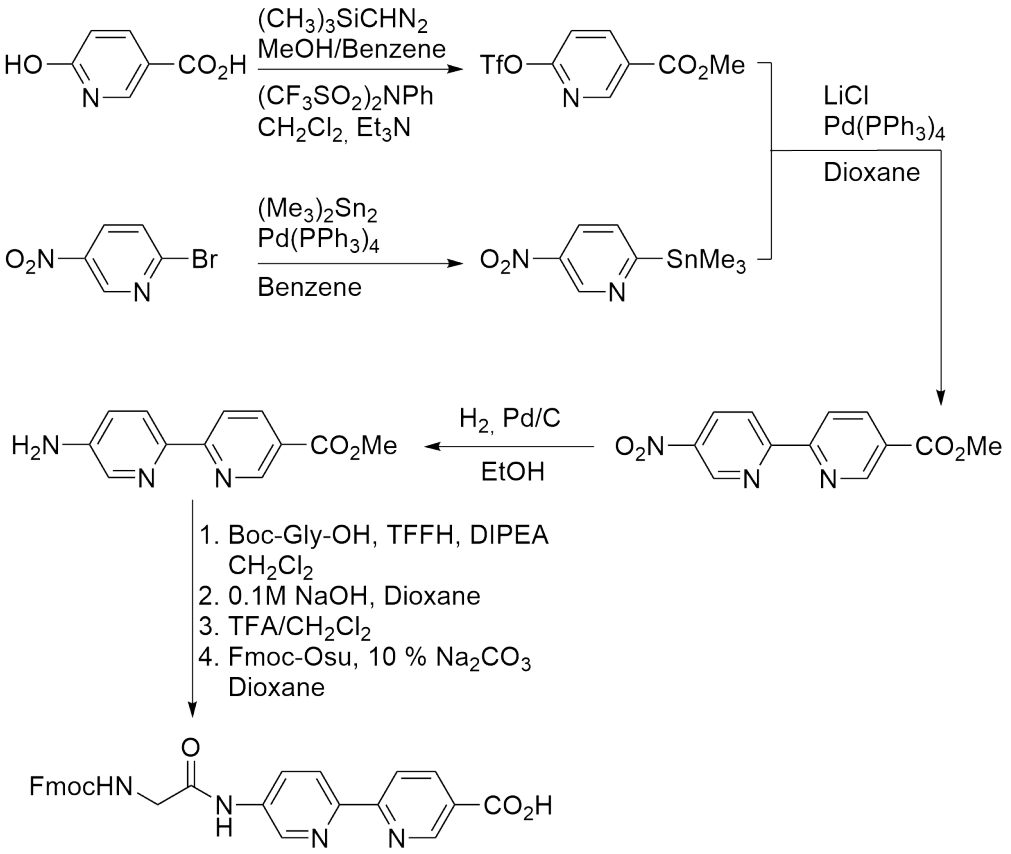


Figure 22. Synthetic route for obtaining the coordinating residue derived from 5Bpy, developed by Imperiali, and suitable for its use in a SPPS methodology.[154]

Later in 1997, Newkome *et al.* reported a different pathway for the synthesis of the 5Bpy ligand. In this case, the the key step of the synthesis is the desymmetrization of diethyl [1,1'-bipirdyl]-4,4'-dicarboxylate by treatment with hydrazine monohydrate, which leads to the monocarbohydrazide product through selective precipitation. Afterwards, the hydrazide is oxidized to the azide group, which is then converted into a carbamate by a Curtius rearrangement. The carbamate is finally hydrolyzed together with the ester group to obtain the final Bpy-derived amino acid, in an overall yield of 21% (Figure 23).[153]

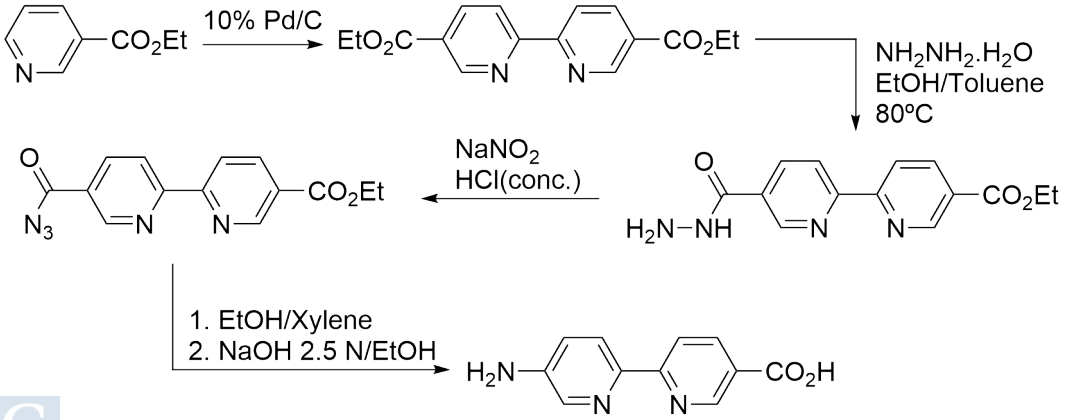


Figure 23. Synthetic route developed by Newkome for obtaining the coordinating amino acid 5Bpy.[153]

USC DE SANTIAGO DE COMPOSTELA

The route developed by Newkome was later redesigned by Darbre and co-workers, using the 5,5'-dimethyl-2,2'-bipyridine as starting material, instead of pyridine derivatives. This precursor was oxidized to give the corresponding diacid, which is subsequently esterified in acidic media. Through selective deprotection of one ethyl ester group, the mono carboxylate product is obtained, that, by treatment with diphenylphosphoryl azide, yields the corresponding azide, which following the route described by Newkome leads to the Fmoc-5Bpy amino acid (Figure 24).[155]

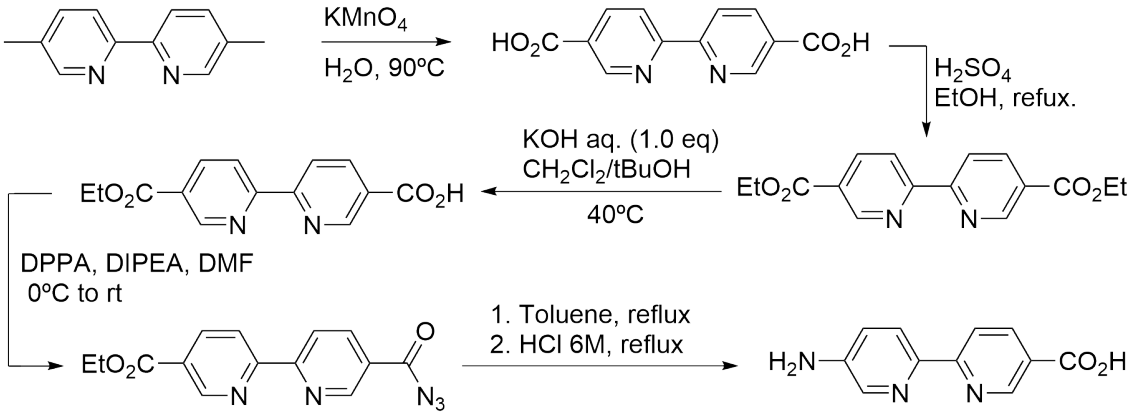


Figure 24. Synthetic route developed by Darbre for obtaining the coordinating amino acid 5Bpy.[155]

More recently, our research group has developed a new route for the synthesis of a dipeptide building block derived from the 5Bpy unit, compatible with the solid-phase peptide synthesis (SPPS).[156] Our new pathway is based on the above described synthetic routes.[153,154,155,157]

Briefly, our approach starts with the oxidation and subsequent esterification of the commercially available 5,5'-dimethyl-2,2'-bipyridine (**1**) to give the corresponding diethylester (**3**), which is then reacted with hydrazine monohydrate in a solution of ethanol:toluene (3:1) at 80º for 3 h, giving the monocarbohydrazide product (**4**) by selective precipitation. This desymmetrization reaction is the key step of the process. Next, following Newkome's route, the hydrazide group is oxidized to carbazide (**5**), which is then transformed into a carbamate (**6)** through a Curtius rearrangement, carried out in ethanol:xylene 1:1. The final hydrolysis of the carbamate and ester groups in **6**, gives rise to the 5Bpy amino acid (**7**). Regretfully, **7** is not suitable for the SPPS methodology, due to its low solubility, poor nucleophilicity and high tendency to undergo oxidation of the 5' amine.[154,132,158] To overcome these drawbacks, the core amino-acid **7** is derivatized as a di-peptide building block by coupling it with Fmoc-βAla-OH (**8a**) or Fmoc-O1Pen-OH (**8b**), thus yielding, Fmoc-βAla5Bpy-OH (**9a**) and Fmoc-O1Pen5Bpy-OH (**9b**), respectively, on a multigram scale and in 41% overall yield (Figure 25).[156]

[155] N. A. Uhlich, P. Sommer, C. Bühr, S. Schürch, J.-L. Reymond, T. Darbre, *Chem. Commun.* **2009**, 6237–6239.
[156] G. Rama, A. Ardá, J.-D. Maréchal, I. Gamba, H. Ishida, J. Jiménez-Barbero, M. E. Vázquez, M. Vázquez López, *Chem. Eur. J.* **2012**, *18*, 7030–7035.
[157] C. P. Whittle, *J. Heterocycl. Chem.* **1977**, *14*, 191–194.
[158] C. Yu, J. W. Taylor, *Bioorg. Med. Chem.* **1999**, *7*, 161–175.

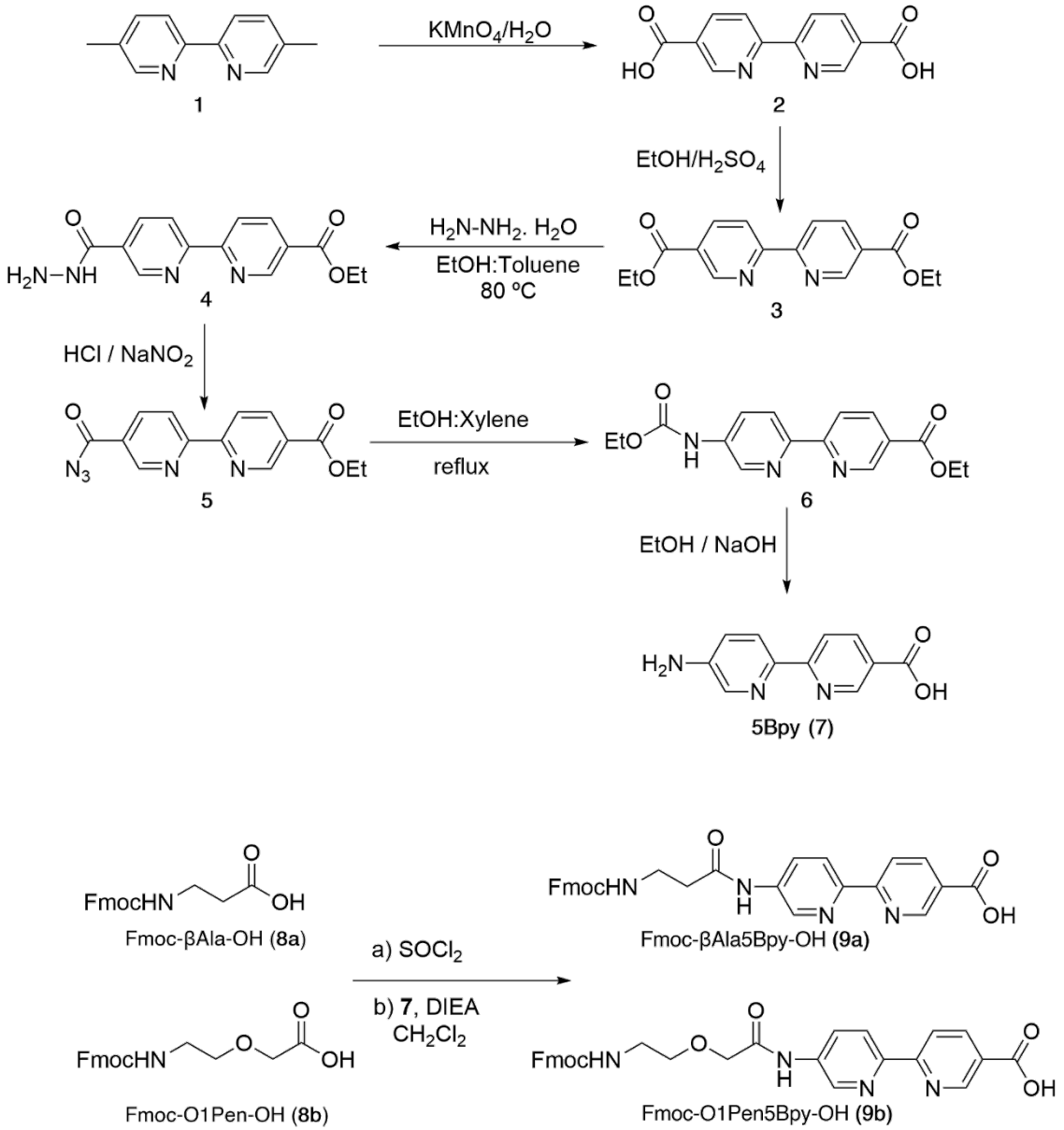


Figure 25. Synthetic route developed in our research group for obtaining the non-natural coordinating residues 9a and 9b, derived from 5Bpy and suitable for their use in a SPPS methodology.[156]

There are a number of examples in literature showing the effective incorporation of this 5Bpy units into peptide backbones, following the standard protocols of the SPPS methodology, as well as their potential in the synthesis of new metallopeptides with attractive features.[159] In 2006, Ishida and co-workers reported the synthesis of a new Ru(II) tris(bipyridyl) metallopeptide with luminescent properties, which, upon metal ion binding, folds into a predictable conformation induced by two β-turn sequences and the chirality of the amino acids coupled to the 5Bpy unit.[160]

[159] a) M. Kyakuno, S. Oishi, H. Ishida, *Chem. Lett.* **2005**, *34*, 1554–1555; b) M. Albrecht, P. Stortz, *Chem. Soc. Rev.* **2005**, *34*, 496–506; c) M. Albrecht, P. Stortz, R. Nolting, *Synthesis* **2003**, *2003*, 1307–1320.

[160] H. Ishida, Y. Maruyama, M. Kyakuno, Y. Kodera, T. Maeda, S. Oishi, *Chembiochem* **2006**, *7*, 1567–1570.

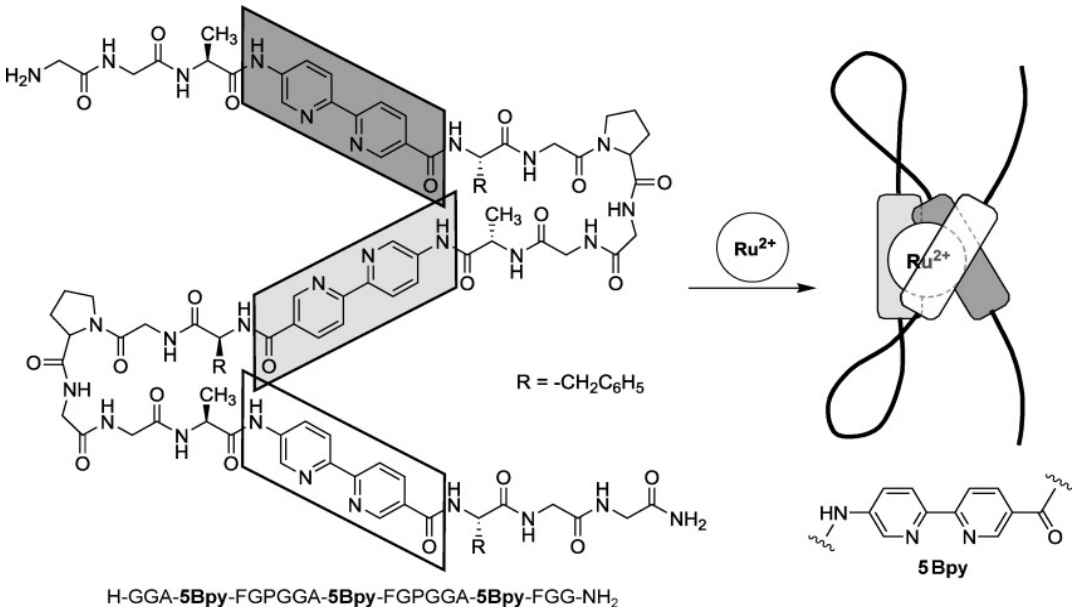


Figure 26. Synthesis of the luminescent Ru(II) complex derived from a peptide ligand equipped with three 5Bpy coordinating residues, designed by Ishida.[160]

Likewise, our research group has widely used Fmoc-βAla5Bpy-OH (**9a**) and Fmoc-O1Pen5Bpy-OH (**9b**) building blocks[156] in the synthesis of new DNA-binding metallopeptides. As a first approach, polypyridyl Ru(II) and Ir(III) complexes with a βAla5Bpy coordinating residue, coupled to a peptide strand, were prepared.[161] Different mononuclear chiral metallopeptides have also been synthesized, with the coordination sphere of Fe(II), Co(II), Ni(II) and Zn(II) ions being saturated trough two or three 5Bpy residues placed at specific positions in the peptide sequence.[156,162] The DNA-binding affinity of these metallopeptides, as well as their cell internalization have been improved through easy modifications in the structure of the peptide chains, simplified by the synthetic versatility of the SPPS.

[161] a) I. Gamba, I. Salvadó, G. Rama, M. Bertazzon, M. I. Sánchez, V. M. Sánchez-Pedregal, J. Martínez-Costas, R. F. Brissos, P. Gamez, J. L. Mascareñas, M. Vázquez López, M. E. Vázquez, *Chem. Eur. J.* **2013**, *19*, 13369–13375; b) I. Gamba, I. Salvadó, R. F. Brissos, P. Gamez, J. Brea, M. I. Loza, M. E. Vázquez, M. Vázquez López, *Chem. Commun.* **2016**, *52*, 1234-1237; c) I. Salvadó, I. Gamba, J. Montenegro, J. Martínez-Costas, J. M. Brea, M. I. Loza, M. Vázquez López, M. E. Vázquez, *Chem. Commun.* **2016**, *52*, 11008–11011; d) D. Bouzada, I. Salvadó, G. Barka, G. Rama, J. Martínez-Costas, R. Lorca, Á. Somoza, M. Melle-Franco, M. E. Vázquez, M. V. López, *Chem. Commun.* **2018**, *54*, 658–661.

[162] I. Gamba, G. Rama, E. Ortega-Carrasco, R. Berardozzi, V. M. Sánchez-Pedregal, L. Di Bari, J.-D. Maréchal, M. E. Vázquez, M. V. López, *Dalton Trans. J. Inorg. Chem.* **2016**, *45*, 881–885.

4. Peptide helicates

4.1. Helicates

The term *helicate* was first introduced in 1987 by Jean Marie Lehn to describe a double-helical strand polymetallic complex displaying the features of an inorganic double helix.[163] In other words, a *metallosupramolecular helicate* refers to a chiral helical compound made up of one or more organic ligands wrapped around and coordinated to two or more metal cations, which define the rotational axis in the helix (Figure 27).[164] The definition of helicates comprises from classical coordination complexes to more complex supermolecules,[165] that meet two conditions: a) at least one acyclic organic chain is required to be wound around and coordinated to metal cations through coordinated bonds (dative bonds); b) they form discrete polynuclear oligomers, since more than two metal cations are necessary to define the helical axis.

4.1.1. Structural characteristics of helicates

As discussed above, metallo-supramolecular helicates are composed of metal centres and organic ligands known as strands. Such ligands appear as a succession of binding domains joined by spacers.[164] Within these binding domains are located the donor atoms (N, S, O, etc.) able to form dative bonds with metal cations; although they may also be found in the spacers.

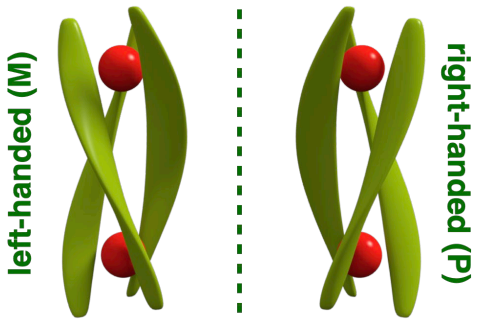


Figure 27. Illustration of the two enantiomers (M and P) of a chiral supramolecular dinuclear trihelicate. The red balls represent the metal centres and the green ribbons the organic ligand strands.

Helicates can be classified following different criteria:

- According to the number of constituent strands, they are divided into *monohelicates*, *dihelicates*, *trihelicates*, etc.[166]
- Depending on the nature and number of the metal ions, they are respectively classified as *homonuclear* or *heteronuclea*r, and *dinuclear*, *trinuclear* or *polynuclear*.
- Regarding the binding domains, they are divided into homoleptic or heteroleptic, bearing the same or different coordinating units, respectively.[164a,167]

[163] J. M. Lehn, A. Rigault, J. Siegel, J. Harrowfield, B. Chevrier, D. Moras, *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 2565–2569.
[164] a) C. Piguet, G. Bernardinelli, G. Hopfgartner, *Chem. Rev.* **1997**, *97*, 2005–2062; b) E. C. Constable, in *Progress in Inorganic Chemistry*, John Wiley & Sons, Inc., Hoboken, NJ, USA, **2007**, pp. 67–138.
[165] D. S. Lawrence, T. Jiang, M. Levett, *Chem. Rev.* **1995**, *95*, 2229–2260.
[166] H. Miyake, H. Tsukube, *Chem. Soc. Rev.* **2012**, *41*, 6977–6991.
[167] M. Albrecht, *Chem. Rev.* **2001**, *101*, 3457–3498.

- Depending on whether the coordination sphere of the metal ion is completed with the binding domains, *saturated* and *unsaturated* helicates can be distinguished.

Very few examples of *monohelicates* have been reported in literature, being polynuclear complexes consisting of a single helical strand coordinated to the metallic centers,[168] making these complexes much less common than dihelicates and trihelicates.

Dinuclear dihelicates are one of the most common in this group, mostly derived from polipyridiyl ligands or Schiff bases.[169] The predominant metal ions involved are Fe(II),[170] Co(II),[170b,171] Ni(II),[172] Cu(II),[173] Zn(II),[170b,173c,174] Ag(I),[169b,175] Cu(I),[163,173a,175d,176] and, to a lesser degree, Cd(II),[169c,177] which exhibit coordination indices of four, five or six. Although more exotic, trinuclear and polynuclear dihelicates have also been described.[178]

---

[168] a) J. Sanmartín, A. M. García-Deibe, M. R. Bermejo, F. Novio, D. Navarro, M. Fondo, *Eur. J. Inorg. Chem.* **2003**, *2003*, 3905–3913; b) M. Vázquez López, G. Zaragoza, R. Pedrido, G. Rama, M. R. Bermejo, *Inorg. Chem. Commun.* **2008**, *11*, 995–998; c) M. Vázquez López, G. Zaragoza, M. Otero, R. Pedrido, G. Rama, M. R. Bermejo, *Cryst. Growth Des.* **2008**, *8*, 2083–2086; d) U. Phukon, B. Shankar, M. Sathiyendiran, *Dalton Trans. J. Inorg. Chem.* **2022**, *51*, 16307–16315.

[169] a) B. Hasenknopf, J. M. Lehn, G. Baum, D. Fenske, *Proceedings of the National Academy of Sciences* **1996**, *93*, 1397–1400; b) M.-H. Shu, W.-Y. Sun, C.-Y. Duan, W.-X. Tang, W.-J. Zhang, *Transition Met. Chem.* **1999**, *24*, 628–632; c) M. Martínez-Calvo, A. M. González-Noya, R. Pedrido, M. J. Romero, M. I. Fernández, G. Zaragoza, M. R. Bermejo, *Dalton Trans.* **2010**, *39*, 1191–1194 d) L. Carbonaro, M. Isola, V. Liuzzo, F. Marchetti, F. Balzano, C. Silvio Pomelli, A. Raffaelli, *Eur. J. Inorg. Chem.* **2001**, *2001*, 353–357.

[170] a) G. Rapenne, J.-P. Sauvage, B. T. Patterson, F. Richard Keene, *Chem. Commun.* **1999**, *0*, 1853–1854; b) S. D. Reid, A. J. Blake, C. Wilson, J. B. Love, *Inorg. Chem.* **2006**, *45*, 636–643.

[171] S. G. Telfer, T. Sato, R. Kuroda, *Angew. Chem. Int. Ed Engl.* **2004**, *43*, 581–584.

[172] a) M. Vázquez, M. R. Bermejo, M. Fondo, A. M. González, J. Mahía, L. Sorace, D. Gatteschi, *Eur. J. Inorg. Chem.* **2001**, *2001*, 1863–1868; b) M. Isola, V. Liuzzo, F. Marchetti, A. Raffaelli, *Eur. J. Inorg. Chem.* **2002**, *2002*, 1588–1590.

[173] a) V. Amendola, M. Boiocchi, V. Brega, L. Fabbrizzi, L. Mosca, *Inorg. Chem.* **2010**, *49*, 997–1007; b) Y. Song, I. A. Koval, P. Gamez, G. A. van Albada, I. Mutikainen, U. Turpeinen, J. Reedijk, *Polyhedron* **2004**, *23*, 1769–1775; c) M. R. Bermejo, A. M. González-Noya, R. M. Pedrido, M. J. Romero, M. Vázquez, *Angew. Chem. Int. Ed Engl.* **2005**, *44*, 4182–4187; d) P. Halder, P. R. Banerjee, E. Zangrando, T. K. Paine, *Eur. J. Inorg. Chem.* **2008**, *2008*, 5659–5665.

[174] a) T. E. Wood, A. C. Ross, N. D. Dalgleish, E. D. Power, A. Thompson, X. Chen, Y. Okamoto, *J. Org. Chem.* **2005**, *70*, 9967–9974; b) A. Jarzebski, C. Tenten, C. Bannwarth, G. Schnakenburg, S. Grimme, A. Lützen, *Chem. Eur. J.* **2017**, *23*, 12380–12386.

[175] a) E. C. Constable, S. M. Elder, M. J. Hannon, A. Martin, P. R. Raithby, D. A. Tocher, *J. Chem. Soc. Dalton Trans.* **1996**, 2423–2433; b) H. Oshio, M. Yamamoto, T. Ito, *J. Chem. Soc. Dalton Trans.* **1999**, *0*, 2641–2644; d) U. Kiehne, T. Weilandt, A. Lützen, *Org. Lett.* **2007**, *9*, 1283–1286.

[176] a) G. Baum, E. C. Constable, D. Fenske, C. E. Housecroft, T. Kulke, M. Neuburger, M. Zehnder, *J. Chem. Soc. Dalton Trans.* **2000**, 945–959; b) P. K. Pal, S. Chowdhury, P. Purkayastha, D. A. Tocher, D. Datta, *Inorg. Chem. Commun.* **2000**, *3*, 585–589; c) N. K. Solanki, A. E. H. Wheatley, S. Radojevic, M. McPartlin, M. A. Halcrow, *J. Chem. Soc. Dalton Trans.* **1999**, *0*, 521–524.

[177] D. J. Cooke, J. M. Cross, R. V. Fennessy, L. P. Harding, C. R. Rice, C. Slater, *Chem. Commun.* **2013**, *49*, 7785–7787.

[178] a) N. Fatin-Rouge, S. Blanc, A. Pfeil, A. Rigault, A.-M. Albrecht-Gary, J.-M. Lehn, *Helv. Chim. Acta* **2001**, *84*, 1694–1711; b) D. Dragancea, P. Bourosh, S. Shova, V. Kravtsov, in *Materials Science and Condensed Matter Physics*, Ibn.idsi.md, **2014**, pp. 183–183; c) B. Schoentjes, J.-M. Lehn, *Helv. Chim. Acta* **1995**, *78*, 1–12; d) J. Tang, J. S. Costa, G. Aromí, I. Mutikainen, U. Turpeinen, **2007**.

Dinuclear trihelicates also constitute the majority among their group and derive from polypyridine ligands and Schiff bases.[179] They are mainly built up from Fe(II),[179f] Co(II),[179c,180] Ni(II),[180,181] and Ru(II) ions,[182] displaying an octahedral geometry. Other examples of trinuclear and polynuclear trihelicates have also been reported.[183]

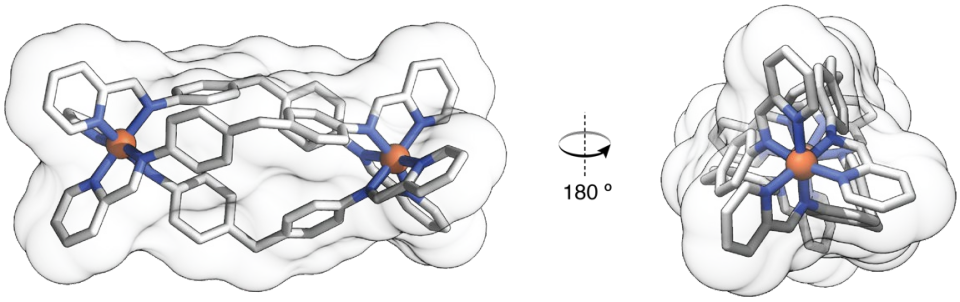


Figure 28. Structure of a Fe(II) dinuclear trihelicate derived from a bis-pyridinylimine ligand, reported by Hannon and co-workers (PDB 2ET0).[184]

4.1.1.1. Helicity and chirality

All of the metal cations within a helicate exhibit the same absolute configuration, (*P* = plus or *M* = minus), which defines the helicity of the complex, i.e., whether the helix turns in a clockwise (*P*) or anti-clockwise (*M*) direction (Figure 27).[185]

The synthesis of enantiomerically pure helicates is essential for the development of new drugs and molecular materials, due to the distinct properties exhibited by the different chiral isomers of the same molecule,[186] as well as the key role of chirality in major biological processes such as the regulation of gene expression, where DNA recognition affinity is strongly dependent on chirality.[187] The resolution of racemic mixtures of helicates, using chromatography or spontaneous separation through

[179] a) J. Libman, Y. Tor, A. Shanzer, *J. Am. Chem. Soc.* **1987**, *109*, 5880–5881; b) B. Kersting, M. Meyer, R. E. Powers, K. N. Raymond, *J. Am. Chem. Soc.* **1996**, *118*, 7221–7222; c) C. Piguet, G. Bernardinelli, B. Bocquet, O. Schaad, A. F. Williams, *Inorg. Chem.* **1994**, *33*, 4112–4121; d) S. G. Sreerama, S. Pal, *Inorg. Chem.* **2005**, *44*, 6299–6307; e) M. Albrecht, *Chem. Eur. J.* **2000**, *6*, 3485–3489; f) K. J. Howard-Smith, A. R. Craze, H. Zenno, J. Yagyu, S. Hayami, F. Li, *Chem. Commun.* **2020**, *56*, 8838–8841.

[180] H. Cheng, D. Chun-ying, F. Chen-jie, M. Qing-jin, *J. Chem. Soc. Dalton Trans.* **2000**, 2419–2424.

[181] a) M. J. Hannon, C. L. Painting, A. Jackson, J. Hamblin, W. Errington, *Chem. Commun.* **1997**, *0*, 1807–1808; b) J. Malina, H. Kostrhunova, V. Brabec, *Inorg. Chem. Front.* **2022**, *9*, 5597–5606.

[182] a) J. Malina, M. J. Hannon, V. Brabec, *Chem. Eur. J.* **2008**, *14*, 10408–10414; b) G. I. Pascu, A. C. G. Hotze, C. Sanchez-Cano, B. M. Kariuki, M. J. Hannon, *Angew. Chem. Int. Ed.* **2007**, *46*, 4374–4378.

[183] a) R. Krämer, J.-M. Lehn, A. De Cian, J. Fischer, *Angew. Chem. Int. Ed Engl.* **1993**, *32*, 703–706; b) T. Riis-Johannessen, G. Bernardinelli, Y. Filinchuk, S. Clifford, N. Dalla Favera, C. Piguet, *Inorg. Chem.* **2009**, *48*, 5512–5525; c) T. Riis-Johannessen, N. Dalla Favera, T. K. Todorova, S. M. Huber, L. Gagliardi, C. Piguet, *Chem. Eur. J.* **2009**, *15*, 12702–12718.

[184] A. Oleksi, A. G. Blanco, R. Boer, I. Usón, J. Aymamí, A. Rodger, M. J. Hannon, M. Coll, *Angew. Chem. Int. Ed Engl.* **2006**, *45*, 1227–1231.

[185] C. Piguet, G. Hopfgartner, B. Bocquet, O. Schaad, A. F. Williams, *J. Am. Chem. Soc.* **1994**, *116*, 9092–9102.

[186] J.-P. Zhang, X.-C. Huang, X.-M. Chen, *Chem. Soc. Rev.* **2009**, *38*, 2385–2396.

[187] G. Wang, L. A. Christensen, K. M. Vasquez, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 2677–2682.

crystallization, is a tedious and inefficient process,[188] therefore a the development of stereoselective synthesis routes of helicates is key for the development of the applied chemistry of these interesting species. In this context, our research group has proposed a new methodology, based on the *Pfeiffer effect*,[189] for the stereoselective synthesis of helicates derived from chiral ligands, in particular peptides composed of chiral residues,[190] which transfer their own chiral information to the metal ion, thus effectively controlling the final chirality of the complex.[190,191]

4.1.2. Fields of application

The unusual structural properties of supramolecular helicates, determined by the nature of their ligands and metal ions, give them great potential to be used in the fields of chemical biology, supramolecular chemistry, and molecular materials. In fact, the value of helicates as chemical sensors,[192] magnetic materials,[193] anti-tumour drugs,[194] catalysts,[195] or as molecular machines,[196] among others, has already been proven. However, for our research, their most attractive activity is related to their DNA binding properties,[190,197] being the field of study of this thesis.

---

[188] M. J. Romero, V. Suárez, S. Fernández-Fariña, M. Maneiro, E. Martínez-Núñez, G. Zaragoza, A. M. González-Noya, R. Pedrido, *Chem. Eur. J.* **2017**, *23*, 4884–4892.

[189] S. Kirschner, N. Ahmad, C. Munir, R. J. Pollock, *J. Macromol. Sci. Part A Pure Appl. Chem.* **1979**, *51*, 913–923.

[190] a) I. Gamba, G. Rama, E. Ortega-Carrasco, J.-D. Maréchal, J. Martínez-Costas, M. E. Vázquez, M. V. López, *Chem. Commun.* **2014**, *50*, 11097–11100; b) J. Gómez-González, D. G. Peña, G. Barka, G. Sciortino, J.-D. Maréchal, M. Vázquez López, M. E. Vázquez, *Front. Chem.* **2018**, *6*, 520; c) J. Gómez-González, Y. Pérez, G. Sciortino, L. Roldan-Martín, J. Martínez-Costas, J.-D. Maréchal, I. Alfonso, M. Vázquez López, M. E. Vázquez, *Angew. Chem. Int. Ed.* **2021**, *60*, 8859–8866; d) J. Gómez-González, D. Bouzada, L. A. Pérez-Márquez, G. Sciortino, J.-D. Maréchal, M. Vázquez López, M. E. Vázquez, *Bioconjug. Chem.* **2021**, *32*, 1564–1569

[191] F. Stomeo, C. Lincheneau, J. P. Leonard, J. E. O'Brien, R. D. Peacock, C. P. McCoy, T. Gunnlaugsson, *J. Am. Chem. Soc.* **2009**, *131*, 9636–9637.

[192] a) M. Wenzel, S. R. Bruere, Q. W. Knapp, P. A. Tasker, P. G. Plieger, *Dalton Trans*., **2010**, *39*, 2936; b) C. D. B. Vandevyver, A.-S. Chauvin, S. Comby, J.-C. G. Bünzli, *Chem. Commun.* **2007**, 1716–1718; c) J.-C. G. Bünzli, *Chem. Rev.* **2010**, *110*, 2729–2755.

[193] a) M. Vázquez, A. Taglietti, D. Gatteschi, L. Sorace, C. Sangregorio, A. M. González, M. Maneiro, R. M. Pedrido, M. R. Bermejo, *Chem. Commun.* **2003**, 1840–1841; b) R. Pedrido, M. V. López, L. Sorace, A. M. González-Noya, M. Cwiklinska, V. Suárez-Gómez, G. Zaragoza, M. R. Bermejo, *Chem. Commun.* **2010**, *46*, 4797–4799; c) C. J. Matthews, S. T. Onions, G. Morata, L. J. Davis, S. L. Heath, D. J. Price, *Angew. Chem. Int. Ed Engl.* **2003**, *42*, 3166–3169.

[194] R. A. Kaner, S. J. Allison, A. D. Faulkner, R. M. Phillips, D. I. Roper, S. L. Shepherd, D. H. Simpson, N. R. Waterfield, P. Scott, *Chem. Sci.* **2016**, *7*, 951–958.

[195] a) C. Dietrich-Buchecker, G. Rapenne, *Chem. Commun.* **1997**, *0*, 2053–2054; b) M. Martínez-Calvo, M. Vázquez López, R. Pedrido, A. M. González-Noya, M. R. Bermejo, E. Monzani, L. Casella, L. Sorace, *Chem. Eur. J.* **2010**, *16*, 14175–14180.

[196] a) K. Miwa, Y. Furusho, E. Yashima, *Nat. Chem.* **2010**, *2*, 444–449; b) S. Yamamoto, H. Iida, E. Yashima, *Angew. Chem. Int. Ed.* **2013**, *52*, 6949–6953; c) Y. Suzuki, T. Nakamura, H. Iida, N. Ousaka, E. Yashima, *J. Am. Chem. Soc.* **2016**, *138*, 4852–4859.

[197] a) I. Meistermann, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletten, S. Khalid, P. M. Rodger, J. C. Peberdy, C. J. Isaac, A. Rodger, M. J. Hannon, *Proceedings of the National Academy of Sciences* **2002**, *99*, 5069–5074; b) U. McDonnell, M. R. Hicks, M. J. Hannon, A. Rodger, *J. Inorg. Biochem.* **2008**, *102*, 2052–2059.

4.2. Potential of helicates in DNA recognition

4.2.1. DNA binders and their potential in therapy and diagnosis

In recent years, one of the main objectives in the field of chemical biology has been the recognition of specific DNA sequences, whose diagnostic and clinical potential relies on the fact that many diseases are originated by the genotype. The main strategy followed so far has been the hybridization of single-stranded DNA fragments by using complementary sequences functionalized with fluorophores.[198,199] In spite of this, selective DNA detection remains a challenge.[200]

Nowadays, the main DNA binders are small organic molecules[201] and transition metal complexes,[150e,202] although research concerning the latter is still underdeveloped in comparison, given the typically poor biocompatibility of the ligands, their lability under physiological conditions and the synthetic complexity for their structural derivatization, in addition to other limitations, also present in the small molecules, such as low chemical stability and cell internalization, together with poor DNA selectivity, which leads to higher toxicity. In addition, the minor groove, with less therapeutic interest, is their main target.

The mechanism of interaction with the DNA is defined by the specific features of each compound, including:[203]

- ***Electrostatic interactions:*** group I and II metal cations, as well as polycationic organic molecules, such as *spermine* or *spermidine*, stabilize the folded conformation of DNA, by interacting, in a non-specific manner, with its negatively charged surface, thus neutralizing the electrostatic repulsions between phosphate groups,[204] as occurs with a family of trinuclear oligomers of Pt(II).[205]

---

[198] a) W. T. Mason, Fluorescent and Luminescent Probes for Biological Activity: A Practical Guide to Technology for Quantitative Real-Time Analysis, Elsevier, **1999**, pp. 51-58; b) E. Pazos, O. Vázquez, J. L. Mascareñas, M. E. Vázquez, *Chem. Soc. Rev.* **2009**, *38*, 3348–3359; c) S. B. VanEngelenburg, A. E. Palmer, *Curr. Opin. Chem. Biol.* **2008**, *12*, 60–65.

[199] a) I. Ghosh, C.I. Stains, D.J. Segal, *Mol. Biosyst*., **2006**, *2*, 551; b) P. E. Nielsen, *Curr. Med. Chem.* **2001**, *8*, 545–550; c) J. M. Levsky, R. H. Singer, *J. Cell Sci.* **2003**, *116*, 2833–2838; d) L. Bissonnette, M. G. Bergeron, *Expert Rev. Mol. Diagn.* **2006**, *6*, 433–450; e) K. K. Jain, *Expert Rev. Mol. Diagn.* **2003**, *3*, 153–161; f) R. Zhao, *Clin. Lab. Sci.* **2005**, *18*, 254–262.

[200] a) S. Janssen, T. Durussel, U. K. Laemmli, *Mol. Cell* **2000**, *6*, 999–1011; b) K. Maeshima, S. Janssen, U. K. Laemmli, *EMBO J.* **2001**, *20*, 3218–3228; c) V. C. Rucker, S. Foister, C. Melander, P. B. Dervan, *J. Am. Chem. Soc.* **2003**, *125*, 1195–1202; d) V. C. Rucker, A. R. Dunn, S. Sharma, P. B. Dervan, H. B. Gray, *J. Phys. Chem. B* **2004**, *108*, 7490–7494; e) C. L. Warren, N. C. S. Kratochvil, K. E. Hauschild, S. Foister, M. L. Brezinski, P. B. Dervan, G. N. Phillips, A. Z. Ansari, *Proceedings of the National Academy of Sciences* **2006**, *103*, 867–872.

[201] W. C. Russell, C. Newman, D. H. Williamson, *Nature* **1975**, *253*, 461–462.

[202] a) V. Fernández-Moreira, F. L. Thorp-Greenwood, M. P. Coogan, *Chem. Commun.* **2010**, *46*, 186–202; b) A. C. Komor, J. K. Barton, *Chem. Commun.* **2013**, *49*, 3617–3630.

[203] a) L. M. Wilhelmsson, P. Lincoln, B. Nordén, *Sequence-Specific DNA Binding Agents*, Royal Society Of Chemistry, **2006**; b) A. L. Molina, M. J. Fernández, in *Anales de La Real Sociedad Española de Química*, Real Sociedad Española De Química, **2008**, pp. 280–289; c) M. J. Hannon, *Chem. Soc. Rev.* **2007**, *36*, 280–295.

[204] H. Deng, V. A. Bloomfield, J. M. Benevides, G. J. Thomas. Jr, *Nucleic Acids Res.* **2000**, *28*, 3379–3385.

[205] D. R. Boer, A. Canals, M. Coll, *Dalton Trans.* **2009**, 399–414.

- ***Intercalation:*** typical of aromatic cations with at least two six-membered rings, which are inserted between consecutive DNA base pairs establishing π-stacking interactions together with electrostatic or charge transfer, and hydrogen bonds.[206] Unfortunately, although they tend to interact with regions rich in guanine-cytosine pairs, their sequence selectivity is low.[207] Among the intercalating agents described so far are ethidium bromide, widely used in fluorescent DNA labeling,[208] and several metal complexes, such as Rhodium(III) Δ-α-[Rh[(R,R)-Me_2trien]phi]$^{3+}$ (Figure 29).[209]

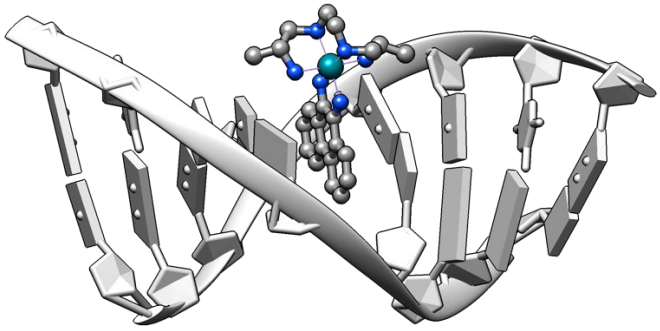


Figure 29. X-ray crystal structure of the DNA intercalating agent Δ-α-[Rh[(R,R)-Me_2trien]phi]$^{3+}$ bound to duplex DNA, showing sequence selectivity for 5'-TGCA-3'(NDB: UD0005).[209b]

- ***Insertion between base pairs:*** is the DNA binding mode displayed by those metal complexes derived from aromatic ligands, with a surface area too large for intercalation. Instead, they are placed between the *mismatches* (non-complementary base pairs) in DNA.[202b] An example of this kind of interaction is that showed by Rh(III) Δ-[Rh$(bpy)_2$(chrysi)]$^{3+}$.[210]
- ***Insertion in the DNA grooves:*** small molecules typically insert into the minor groove of DNA, with selectivity towards A/T-rich sequences because they form a narrower and deeper groove than G-C base pairs, leading to better van der Waals and electrostatic, hydrophobic interactions with the bottom of the groove, and hydrogen bonding with the bases. A representative example withing this group is the cell nuclei marker DAPI.[211] On the other hand, several binding agents targeting the major groove

[206] a) K. Nakamoto, M. Tsuboi, G. D. Strahan, *Drug-DNA Interactions: Structures and Spectra*, John Wiley & Sons, Inc., Hoboken, USA, **2008**, chapter 2; b) A. Rescifina, C. Zagni, M. G. Varrica, V. Pistarà, A. Corsaro, *Eur. J. Med. Chem.* **2014**, *74*, 95–115.
[207] a) B. A. D. Neto, A. A. M. Lapis, *Molecules* **2009**, *14*, 1725–1746; b) H. M. Berman, P. R. Young, *Annu. Rev. Biophys. Bioeng.* **1981**, *10*, 87–114; c) H.-K. Liu, P. J. Sadler, *Acc. Chem. Res.* **2011**, *44*, 349–359.
[208] W. A. Franklin, J. D. Locker, *J. Histochem. Cytochem.* **1981**, *29*, 572–576.
[209] a) B. P. Hudson, C. M. Dupureur, J. K. Barton, *J. Am. Chem. Soc.* **1995**, *117*, 9379–9380; b) C. L. Kielkopf, K. E. Erkkila, B. P. Hudson, J. K. Barton, D. C. Rees, *Nat. Struct. Biol.* **2000**, *7*, 117–121.
[210] a) B. M. Zeglis, V. C. Pierre, J. T. Kaiser, J. K. Barton, *Biochemistry* **2009**, *48*, 4247–4253; b) A. C. Komor, C. J. Schneider, A. G. Weidmann, J. K. Barton, *J. Am. Chem. Soc.* **2012**, *134*, 19223–19233.
[211] a) P. G. Baraldi, A. Bovero, F. Fruttarolo, D. Preti, M. A. Tabrizi, M. G. Pavani, R. Romagnoli, *Med. Res. Rev.* **2004**, *24*, 475–528; b) B. Nguyen, M. P. H. Lee, D. Hamelberg, A. Joubert, C. Bailly, R. Brun, S. Neidle, W. D. Wilson, *J. Am. Chem. Soc.* **2002**, *124*, 13680–13681; c) D. Hamelberg, L. D. Williams, W. D. Wilson, *J. Am. Chem. Soc.* **2001**, *123*, 7745–7755.

of DNA have been reported, known as triple helix-forming oligonucleotides (TFO), which show great potential in the regulation of genetic transcription.[212]

- ***Covalent interaction with the nucleotides:*** most of the drugs currently used in chemotherapy, known as *alkylating agents*, rely on this mechanism that forms covalent (including coordinative covalent) bonds with the bases of the DNA acting as a nucleophile.[213] *Cisplatin*, $[Pt(NH_3)_2Cl_2]$, is one of the most potent antitumor drugs within this group.[214]

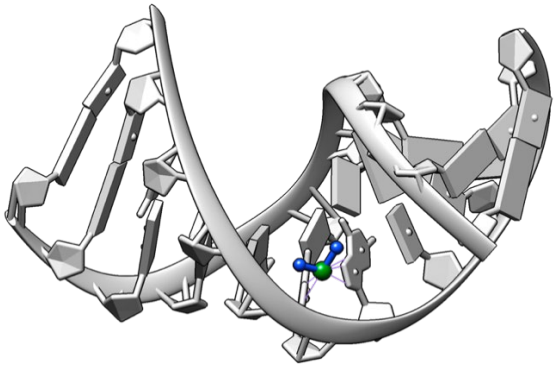


Figure 30. X-ray crystal structure of the antitumor agent cisplatin coordinated to canonical DNA (PDB: 1AIO).[214e]

4.2.2. Helicates as DNA binders

As previously mentioned, those molecules with the ability to bind and/or modulate DNA processing have shown great potential in the development of new diagnostic agents and drugs.[203c,215] In biological systems, the *transcription factors*, specialized proteins equipped with α-helices, are responsible for carrying out the DNA transcription process through non-covalent molecular recognition of the B-DNA major groove. Sadly, reproducing the chemical composition of these proteins with a *de novo* design still entails many difficulties.

Taking as reference the pioneering work carried out in the research group of Prof. J. M. Lehn, who in 1995 reported the effective insertion into the DNA major groove of a set of polynuclear Cu(I) helicates derived from bipyridine ligands,[178c] In an elegant

[212] a) D. Praseuth, A. L. Guieysse, C. Hélène, *Biochim. Biophys. Acta* **1999**, *1489*, 181–206; b) F. Lecubin, M. Devys, J. L. Fourrey, J.-S. Sun, R. Benhida, *Nucleosides Nucleotides Nucleic Acids* **2003**, *22*, 1281–1284; c) M. Jazouli, D. Guianvarc'h, K. Bougrin, M. Soufiaoui, P. Vierling, R. Benhida, *Nucleosides Nucleotides Nucleic Acids* **2003**, *22*, 1277–1280; d) T. D. Ros, G. Spalluto, M. Prato, T. Saison-Behmoaras, A. Boutorine, B. Cacciari, *Curr. Med. Chem.* **2005**, *12*, 71–88.
[213] a) D. Fu, J. A. Calvo, L. D. Samson, *Nat. Rev. Cancer* **2012**, *12*, 104–120; b) D. Karati, K. R. Mahadik, P. Trivedi, D. Kumar, *Anticancer Agents Med. Chem.* **2022**, *22*, 1478–1495.
[214] a) T. Boulikas, M. Vougiouka, *Oncol. Rep.* **2004**, *11*, 559–595; b) R. A. Alderden, M. D. Hall, T. W. Hambley, *J. Chem. Educ.* **2006**, *83*, 728-734; c) S. Ghosh, *Bioorg. Chem.* **2019**, *88*, 102925; d) L. Qi, Q. Luo, Y. Zhang, F. Jia, Y. Zhao, F. Wang, *Chem. Res. Toxicol.* **2019**, *32*, 1469–1486; e) P. M. Takahara, A. C. Rosenzweig, C. A. Frederick, S. J. Lippard, *Nature* **1995**, *377*, 649–652.
[215] a) P. C. A. Bruijnincx, P. J. Sadler, *Curr. Opin. Chem. Biol.* **2008**, *12*, 197–206; b) L. Strekowski, B. Wilson, *Mutat. Res.* **2007**, *623*, 3–13; c) N. Hadjiliadis, E. Sletten, Eds., *Metal Complex - DNA Interactions*, John Wiley & Sons, Chichester, UK, **2009**; d) J. W. Højfeldt, A. R. Van Dyke, A. K. Mapp, *Chem. Soc. Rev.* **2011**, *40*, 4286–4294; e) M. R. Gill, J. A. Thomas, *Chem. Soc. Rev.* **2012**, *41*, 3179–3192.

approach, Prof. M. J. Hannon proposed in 2001 to reproduce the DNA recognition properties of peptide α-helices by mimicking their shape and size, rather than their chemical composition.[216,217] Based on this idea, he proved that helicates were promising candidates as DNA binding agents, because of their unique chirality, helicity, and cylindrical shape.[185] He published the synthesis of a dinuclear Fe(II) trihelicate bearing bis-pyridylimine ligands, which exhibits chiral discrimination in its binding affinity towards both grooves of B-DNA.[197a,218] The results observed suggest that metal ions not only induce the structural self-organization of the helicate, but they are active actors in their biological activity, as they are the responsible of stabilizing the adduct formed between the metal cylinder and the DNA *via* electrostatic interactions between the net positive charge that they conferred to the helicate and the negatively charged phosphate chain of DNA. Importantly, subsequent studies have shown that this dinuclear Fe(II) helicate is able to inhibit cell growth without causing DNA strand breaks, which means that it is neither mutagenic nor genotoxic,[219] as happens with cisplatin. This offers a great advantage for the use of helicates in the design of new antitumor drugs.

The importance of chirality on the DNA binding of helicates was demonstrated by Cherie Motti and co-workers, who separated a racemic mixture of a Ru(II) dinuclear trihelicate by elution through a stationary phase with an attached dodecanucleotide, which binds with higher affinity to the *P* isomer than to the *M* isomer.[220]

In recent years, Prof. P. Scott has made significant advances in the field of helicates, leading to the synthesis of enantiomerically pure, water-soluble and stable Fe(II) trihelicates with high binding affinity (nM dissociation constants) and selectivity towards AT-rich sequences in the B-DNA major groove. In particular, these *flexicates* show good antibacterial activity against gram positive and gram negative bacteria.[221]

As previously discussed in this introduction, DNA can adopt other secondary structures besides the canonical B-DNA conformation, including G-quadruplexes[19,222] or DNA three-way junctions.[22,24,25] which occur at specific sites in the genome and at certain moments in the cell cycle. Recently, these non-canonical DNA structures have been associated to the development of several diseases,[223] including myotonic dystrophy type

---

[216] M. J. Hannon, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletten, I. Meistermann, C. J. Isaac, K. J. Sanders, A. Rodger, *Angew. Chem. Int. Ed.* **2001**, *40*, 879–884.
[217] H. Song, M. Postings, P. Scott, N. J. Rogers, *Chem. Sci.* **2021**, *12*, 1620–1631.
[218] C. Uerpmann, J. Malina, M. Pascu, G. J. Clarkson, V. Moreno, A. Rodger, A. Grandas, M. J. Hannon, *Chem. Eur. J.* **2005**, *11*, 1750–1756.
[219] A. C. G. Hotze, N. J. Hodges, R. E. Hayden, C. Sanchez-Cano, C. Paines, N. Male, M.-K. Tse, C. M. Bunce, J. K. Chipman, M. J. Hannon, *Chem. Biol.* **2008**, *15*, 1258–1267.
[220] C. R. K. Glasson, G. V. Meehan, J. K. Clegg, L. F. Lindoy, J. A. Smith, F. R. Keene, C. Motti, *Chem. Eur. J.* **2008**, *14*, 10535–10538.
[221] S. E. Howson, A. Bolhuis, V. Brabec, G. J. Clarkson, J. Malina, A. Rodger, P. Scott, *Nat. Chem.* **2012**, *4*, 31–36.
[222] a) J. L. Huppert, *Chem. Soc. Rev.* **2008**, *37*, 1375–1384: b) M. L. Bochman, K. Paeschke, V. A. Zakian, *Nat. Rev. Genet.* **2012**, *13*, 770–780; c) E. H. Blackburn, *Nature* **1991**, *350*, 569–573; d) A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, L. H. Hurley, *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11593–11598.
[223] a) C.-Y. Chen, Q. Wang, J.-Q. Liu, Y.-H. Hao, Z. Tan, *J. Am. Chem. Soc.* **2011**, *133*, 15036–15044; b) S. Cosconati, A. Rizzo, R. Trotta, B. Pagano, S. Iachettini, S. De Tito, I. Lauri, I.

1 or Huntington's disease, being human neurodegenerative diseases related to trinucleotide expansions, which are 3WJ-forming mutations.[224] This is why one of the main challenges in chemical biology is the development of selective binders to these secondary structures. In this field, helicates have already demonstrated their potential in the recognition of G-Quadruplexes[225] and 3WJs.[226] It was Prof. Hannon who first demonstrated in 2006 that helicates were able to attach to 3WJ structures by inserting themselves at their branching point (Figure 31).[184]

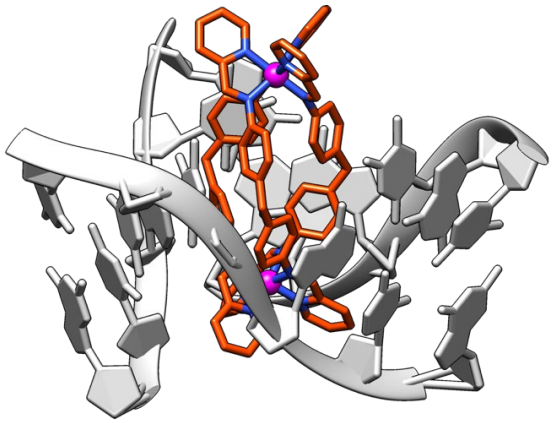


Figure 31. X-ray crystal structure of the Hannon´s Fe(II) trihelicate placed into the DNA three-way junction branch point (PDB: 2ET0).[184]

4.2.2.1. Background of peptide helicates as DNA-binding agents

Despite the great potential of helicates in the field of chemical biology, their development has been slow, due to the lack of versatile synthetic methods for their structural and functional optimization, together with difficulties associated with their enantioselective synthesis and poor stability in physiological media.[226] Helicates derived from peptide ligands emerge as a potential solution to these limitations, but so far, just a few examples have been reported in the bibliography. Prof. M. Albretch pioneered this field publishing in 2002 the synthesis of a dinuclear triple-stranded Ti(IV) helicate, derived from ligand strands made up of valine residues (1-3) connecting two catechol non-natural coordinating residues.[227] Some years later, following the same strategy, Yashima and co-workers synthesized dinuclear Fe(II) helicates consisting of three strands, each of them equipped with two coordinating residues derived from bipyridine (5Bpac), bridged through a peptide sequence.[228]

Fotticchia, M. Giustiniano, L. Marinelli, C. Giancola, E. Novellino, A. Biroccio, A. Randazzo, *J. Med. Chem.* **2012**, *55*, 9785–9792; c) M. Bejugam, S. Sewitz, P. S. Shirude, R. Rodriguez, R. Shahid, S. Balasubramanian, *J. Am. Chem. Soc.* **2007**, *129*, 12926–12927; d) S. N. Georgiades, N. H. Abd Karim, K. Suntharalingam, R. Vilar, *Angew. Chem. Int. Ed Engl.* **2010**, *49*, 4020–4034.

[224] a) C. E. Pearson, M. Tam, Y. H. Wang, S. E. Montgomery, A. C. Dar, J. D. Cleary, K. Nichol, *Nucleic Acids Res.* **2002**, *30*, 4534–4547; b) R. R. Sinden, *Nature*, **2001**, *411*, 757-758.

[225] a) H. Yu, X. Wang, M. Fu, J. Ren, X. Qu, *Nucleic Acids Res.* **2008**, *36*, 5695–5703; b) J. Wang, Y. Chen, J. Ren, C. Zhao, X. Qu, *Nucleic Acids Res.,* **2014**, *42*, 3792.

[226] J. Malina, M. J. Hannon, V. Brabec, *Chem. Eur. J.* **2007**, *13*, 3871–3877.

[227] a) M. Albrecht, M. Napp, M. Schneider, P. Weis, R. Fröhlich, *Chem. Eur. J.* **2001**, *7*, 3966–3975; b) M. Albrecht, O. Spieß, M. Schneider, P. Weis, *Chem. Commun.* **2002**, 786–787.

[228] N. Ousaka, Y. Takeyama, E. Yashima, *Chem. Sci.* **2012**, *3*, 466–469.

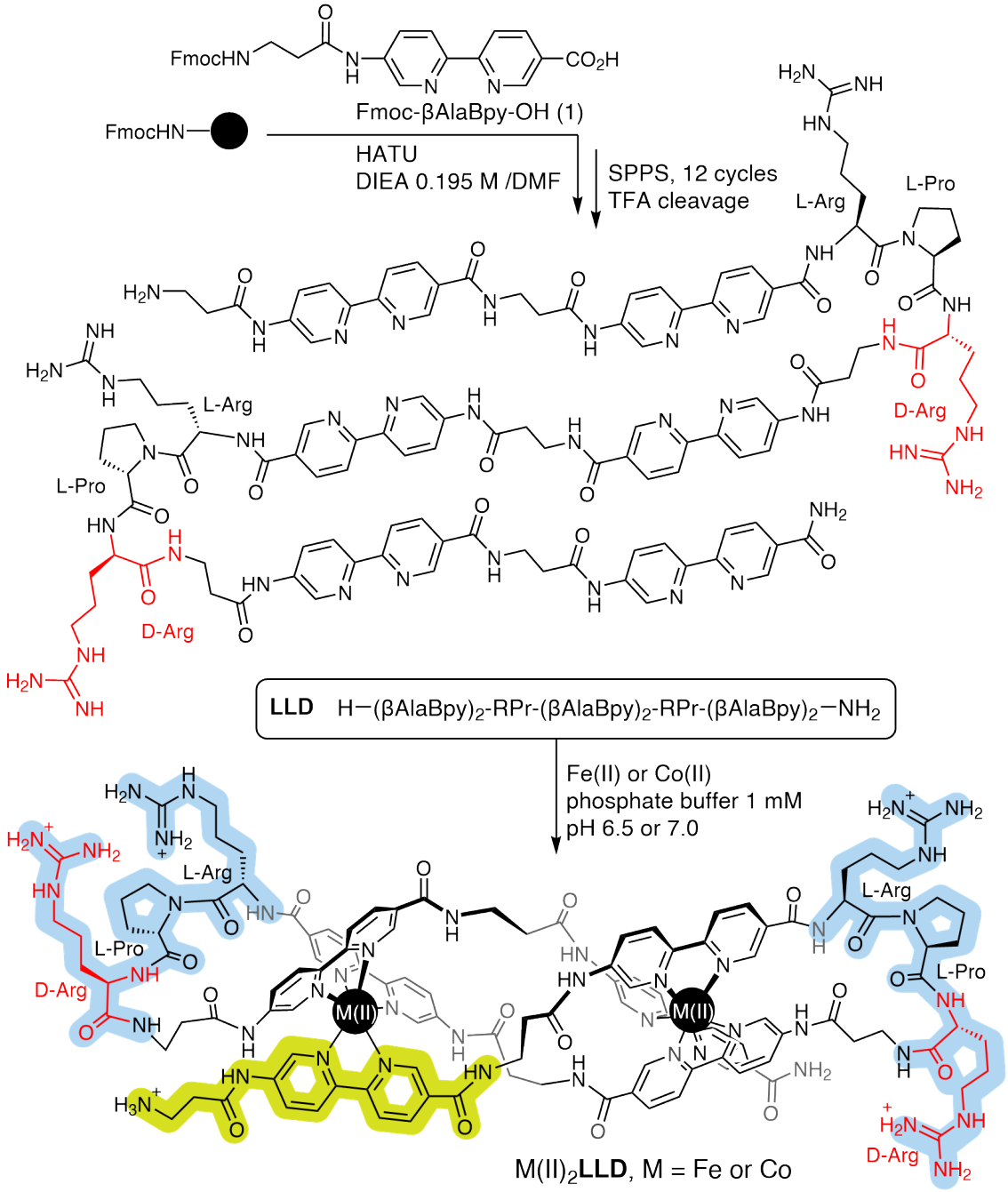
Prof. Hannon, with the aim of increasing the binding selectivity towards specific sequences in the B-DNA major groove displayed by his iron(II) helicates, modified in 2009 the bis-pyridylimine ligand precursor by anchoring a GGS-$CONH_2$ tripeptide at its terminal ends. The peptide included a Ser residue, chosen for its abundance in protein-DNA interactions, and a two-Gly spacer.[229] Subsequently, based on the ability of chiral ligands to determine the final chirality of helicates, Hannon decided to exchange this tripeptide for a single L- and D-Arginine amino acid, giving rise to the enantioselective synthesis of chirally pure M and P isomers of the Fe(II) helicate, respectively. Moreover, the arginine residues also improved the water solubility of the helicates.[230]

All the peptide ligands described above were synthesized on solution without following a SPPS methodology. However, to obtain DNA-binding helicates with improved chemical and biological properties, our research group developed ten years ago a completely different approach based on the solid-phase peptide synthesis (SPPS) methodology, specifically by using peptide ligands as precursors of helicates. This novel approach combines the biocompatibility, modularity and structural control of peptides with the versatility and synthetic flexibility of the SPPS methodology, thus enabling the rapid construction and modification of helicates, which facilitates the optimization of their biophysical properties.

The structural design of these peptide helicates is based on the preorganization of the peptide ligands through the incorporation of chiral amino acids strategically located between non-natural coordination residues derived from 2,2'-bipyridine. These chiral amino acids will be responsible for encoding the final chirality in the helicates. The studies conducted with the resulting complexes, such as fluorescence titrations or EMSA, confirmed the efficient synthesis of chirally-pure helicates, showing high binding affinity and selectivity for 3WJs against canonical DNA, as well as very promising labeling properties *in cellulo*.[190]

USC

[229] L. Cardo, M. J. Hannon, *Inorganica Chim. Acta* **2009**, *362*, 784–792.
[230] L. Cardo, V. Sadovnikova, S. Phongtongpasuk, N. J. Hodges, M. J. Hannon, *Chem. Commun.* **2011**, *47*, 6575–6577.



Scheme 1. Scheme of the methodology developed in our group for the synthesis of chirally pure helicates derived from peptide ligands based in the SPPS methodology.[190c]

Nowadays, our research group is still working on optimization of these peptide helicates, by introducing modifications in the peptide sequence, aiming to improve their *in vitro* and *in cellulo* 3WJ binding properties as well as to endow them with interesting properties from a biological and therapeutic point of view. In fact, this is one of the main objectives of this this doctoral thesis.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

General Objectives

In view of the precedents and the current progress in obtaining new agents for the recognition of specific DNA sequences with antitumor potential, **the main objective of this thesis is the development of new peptide-based binders, capable of selectively recognize 3WJs with high affinity**, towards this objective, we have followed two different strategies.

The same approach was used in the first three chapters, consisting on the synthesis of chirally pure helicates derived from peptide ligands bearing six 5Bpy coordinating units. These ligands stereoselectively self-assemble in solution in the presence of octahedral metal ions to give the corresponding dinuclear peptide helicates:

- In the first chapter, we exploit the known DNA cleavage properties displayed by polypyridine copper complexes to synthesize Cu(II) peptide helicates with selective and efficient 3WJ nuclease activity. Moreover, aiming to improve the nuclease activity shown by these systems, in the third chapter we describe the design of a peptide ligand precursor of peptide helicates functionalized by an ATCUN motif at its N-terminus. This ligand self-assemble in the presence of Cu(II) and Ni(II) ions into the corresponding Cu(II) and Ni(II) peptide helicates chimeras, exhibiting high ability to selective recognize and cleave 3WJs.
- In the second chapter, we describe to design a new dynamic peptide system with C_3 symmetry that will be able to self-assemble into fibers or discrete helicates depending on the presence of Co(II) ions. Moreover, the Co(II) metallopeptide can be oxidized to yield the corresponding kinetically-inert Co(III) helicate.

In the fourth chapter, the selective recognition of 3WJs is approached in a different way, through the cooperative binding to these non-canonical structures by a new peptide designed with three AT-Hook peptides strands attached to a central core, which will selectivity recognize the A/T rich duplex-DNA arms in 3WJs.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter I

Design and Synthesis of Cu(II) Peptide Helicates that Selective Cleavage DNA Replication Foci in Mammalian Cells

Summary

One of the challenges posed using copper complexes capable of cleaving DNA as potential anticancer drugs is that they are not very selective in terms of the oxidation process that breaks DNA strands. A possible solution to solve this problem is to design systems that can selectively target and damage non-canonical DNA structures. We have designed a Cu(II) peptide helicate that can selectively bind 3WJs and cause oxidative DNA damage through a ROS-mediated pathway *in vitro* and *in cellulo*, specifically at DNA replication sites in cell nuclei, where this DNA structure is transiently formed. In this chapter we describe the first example of a targeted chemical nuclease that can distinguish 3WJs from other forms of DNA in cells. Since DNA replication is deregulated in cancer cells, this strategy may pave the way for the development of a new class of anticancer drugs based on copper-based artificial nucleases.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Introduction

In 1965 Barnett Rosenberg accidentally discovered the cytotoxic properties of cisplatin,[214b] an event that placed the focus on DNA-targeted metallodrugs for cancer therapy.[231] Over the past few years, the attention in this research field has shifted towards non-canonical DNA structures owing to greater knowledge regarding their structural polymorphism and biology.[161d,184,190,232]. Specially, in the development of new anticancer drugs, DNA three-way junctions (3WJs), the simplest branched DNA structures composed by three duplex DNA arms converging in a central cavity,[22,24,25] emerge as promising targets with unique therapeutic properties.[38,233] This is because the transient formation of 3WJs in the replication forks during DNA replication[35,36,37] suggests that 3WJ ligands could be able to stall the replication fork and elicit genotoxic replication stress in rapidly dividing cells, such as tumoral ones.[234] In fact, it has already been demonstrated that the stabilization of such non-canonical structures through small molecules binding leads to DNA damage that results in cell death.[233c]

In the search for new chemotherapeutic drugs, synthetic agents capable of hydrolysing DNA by cleavage of phosphate ester bonds, mimicking natural nucleases, could block gene expression at the DNA level, and have therefore been the subject of much attention in the last twenty years.[235] In biological systems, nuclease enzymes carry out different key control functions, such as degradation of viral RNA and RNA splicing, together with synthesis, repair and packaging of DNA.[236] In many of them, the active site is occupied by divalent cations, mainly Mg(II), Ca(II), Mn(II) or Zn(II), which are essential for their catalytic activity,[237] thus prompting the investigation of metal complexes as potential artificial metallonucleases.[238] Indeed, a few lanthanide and transition metal ion complexes have been reported to accelerate the rate of DNA hydrolysis.[239]

---

[231] E. J. Anthony, E. M. Bolitho, H. E. Bridgewater, O. W. L. Carter, J. M. Donnelly, C. Imberti, E. C. Lant, F. Lermyte, R. J. Needham, M. Palau, Others, *Chem. Sci.* **2020**, *11*, 12888–12917.

[232] a) A. Shivalingam, M. A. Izquierdo, A. L. Marois, A. Vyšniauskas, K. Suhling, M. K. Kuimova, R. Vilar, *Nat. Commun.* **2015**, *6*, 8178; b) R. Vilar, in *Medicinal Chemistry* (Eds.: P.J. Sadler, R. van Eldik), Elsevier, San Diego, CA, **2020**, pp. 425–445; c) C. A. J. Hooper, L. Cardo, J. S. Craig, L. Melidis, A. Garai, R. T. Egan, V. Sadovnikova, F. Burkert, L. Male, N. J. Hodges, D. F. Browning, R. Rosas, F. Liu, F. V. Rocha, M. A. Lima, S. Liu, D. Bardelang, M. J. Hannon, *J. Am. Chem. Soc.* **2020**, *142*, 20651–20660.

[233] a) L. A. Howell, M. Searcey, *Chembiochem* **2009**, *10*, 2139–2143; b) H. Tateishi-Karimata, N. Sugimoto, *Nucleic Acids Res.* **2021**, *49*, 7839–7855; c) J. Zell, F. Rota Sperti, S. Britton, D. Monchaud, *RSC Chem Biol* **2021**, *2*, 47–76.

[234] a) H. Liao, F. Ji, T. Helleday, S. Ying, *EMBO Rep.* **2018**, *19*, 1-18; b) K. Duskova, P. Lejault, É. Benchimol, R. Guillot, S. Britton, A. Granzhan, D. Monchaud, *J. Am. Chem. Soc.* **2020**, *142*, 424–435.

[235] J. Borowska, M. Sierant, E. Sochacka, D. Sanna, E. Lodyga-Chruscinska, *J. Biol. Inorg. Chem.* **2015**, *20*, 989–1004.

[236] W. Yang, *Q. Rev. Biophys.* **2011**, *44*, 1–93.

[237] J. A. Cowan, *Chem. Rev.* **1998**, *98*, 1067–1088.

[238] F. Mancin, P. Scrimin, P. Tecilla, U. Tonellato, *Chem. Commun.* **2005**, 2540–2548.

[239] a) J. Rammo, R. Hettich, A. Roigk, H.-J. Schneider, *Chem. Commun.* **1996**, 105–107; b) S. Takenaka, T. Ihara, M. Takagi, *J. Mol. Recognit.* **1990**, *3*, 156–162.

Copper complexes have been the most widely explored as potential chemical metallonuclease agents.[240] These systems are capable of inducing duplex DNA cleavage *via* nucleobase oxidation, hydrolysis of the phosphodiester backbone or oxidative H-atom abstraction from the deoxyribose sugar. This cleavage is mediated by the formation of *ROS* (*reactive oxygen species*) generated either in the presence of an oxidizing (hydrogen peroxide or molecular oxygen) or a reducing agents (sodium ascorbate), by photo-induction or even in a self-activated mode.[241] To date, many examples of copper complexes as synthetic nucleases with significant anticancer activity and lower off-target toxicity than platinum drugs have already been reported in the bibliography.[242] Unfortunately, metallonucleases in general, and copper-based derivatives in particular, suffer from lack of DNA sequence selectivity. [243,244] To overcome this important drawback, they have been covalently attached to sequence-selective DNA binding agents, such as distamycin[245] and triplex-forming oligonucleotides,[246] or their targets have been shifted to non-canonical DNA structures.[247] In relation to the latter, in 2015 Cowan *et al.* reported the synthesis of a new catalytic Cu(II)-acridine-ATCUN metallopeptide, which showed selective binding and cleavage for G-quadruplex telomeric DNA, with promising anticancer properties.[248] However, to date, no nucleases that specifically target 3WJs have been described.

As explained in the introduction, helicates are chiral metal supramolecular complexes where one or more organic ligands are wound around and coordinate to two or more metal cations.[163,164] Some of this systems have shown high selectivity and binding affinity for 3WJs,[184,190,226] but, although these systems have shown nuclease properties against canonical DNA,[249] their potential as selective 3WJ nucleases have never been studied.

---

[240] T. J. P. McGivern, S. Afsharpour, C. J. Marmion, *Inorganica Chim. Acta* **2018**, *472*, 12–39.
[241] a) F. Mancin, P. Scrimin, P. Tecilla, *Chem. Commun.* **2012**, *48*, 5545–5559; b) M. Pitié, G. Pratviel, *Chem. Rev.* **2010**, *110*, 1018–1059; c) C. Liu, M. Wang, T. Zhang, H. Sun, *Coord. Chem. Rev.* **2004**, *248*, 147–168.
[242] C. Santini, M. Pellei, V. Gandin, M. Porchia, F. Tisato, C. Marzano, *Chem. Rev.* **2014**, *114*, 815–862.
[243] a) Z.-R. Li, J. Li, W. Cai, J. Y. H. Lai, S. M. K. McKinnie, W.-P. Zhang, B. S. Moore, W. Zhang, P.-Y. Qian, *Nat. Chem.* **2019**, *11*, 880–889; b) C. Wende, C. Lüdtke, N. Kulak, *Eur. J. Inorg. Chem.* **2014**, *2014*, 2597–2612; c) N. Z. Fantoni, Z. Molphy, S. O'Carroll, G. Menounou, G. Mitrikas, M. G. Krokidis, C. Chatgilialoglu, J. Colleran, A. Banasiak, M. Clynes, S. Roche, S. Kelly, V. McKee, A. Kellett, *Chem. Eur. J.* **2021**, *27*, 971–983.
[244] K. J. Humphreys, K. D. Karlin, S. E. Rokita, *J. Am. Chem. Soc.* **2001**, *123*, 5588–5589.
[245] M. Pitié, C. J. Burrows, B. Meunier, *Nucleic Acids Res.* **2000**, *28*, 4856–4864.
[246] a) T. Lauria, C. Slator, V. McKee, M. Müller, S. Stazzoni, A. L. Crisp, T. Carell, A. Kellett, *Chem. Eur. J.* **2020**, *26*, 16782–16792; b) B. McGorman, N. Z. Fantoni, S. O'Carroll, A. Ziemele, A. H. El-Sagheer, T. Brown, A. Kellett, *Nucleic Acids Res.* **2022**, *50*, 5467–5481; c) N. Z. Fantoni, B. McGorman, Z. Molphy, D. Singleton, S. Walsh, A. H. El-Sagheer, V. McKee, T. Brown, A. Kellett, *Chembiochem* **2020**, *21*, 3563–3574.
[247] M. Nadai, F. Doria, M. Scalabrin, V. Pirota, V. Grande, G. Bergamaschi, V. Amendola, F. R. Winnerdy, A. T. Phan, S. N. Richter, M. Freccero, *J. Am. Chem. Soc.* **2018**, *140*, 14528–14532.
[248] Z. Yu, M. Han, J. A. Cowan, *Angew. Chem. Int. Ed.,* **2015**, *54*, 1901–1905.
[249] a) R. K. Mishra, K. K. Upadhyay, S. Shukla, R. Mishra, *Chem. Commun.* **2012**, *48*, 4238-4240; b) L. J. Childs, J. Malina, B. E. Rolfsnes, M. Pascu, M. J. Prieto, M. J. Broome, P. M. Rodger, E. Sletten, V. Moreno, A. Rodger, M. J. Hannon, *Chem. Eur. J.* **2006**, *12*, 4919–4927.

Despite their promising DNA-binding properties,[190,197] the use of helicates in chemical biology continues to develop slowly, due to the lack of a general strategy for their efficient and versatile stereoselective synthesis, which hinders the optimization of their biological properties.[226] To address these challenges, our research group has developed a methodology for the stereoselective synthesis of helicates using peptide ligands bearing non-natural coordinating residues derived from 2,2'-bipyridine, specifically Fmoc-βAla5Bpy-OH[190b,c,d] and Fmoc-O1Pen5Bpy-OH.[156,190a]

In 2021, we report the design and preparation of new oligocationic helicates derived from peptide ligands equipped with six artificial coordinating building blocks Fmoc-βAla5Bpy-OH and two heterochiral β-turn sequences, –L-Arg–L-Pro–D-Arg– (or –D-Arg–D-Pro–L-Arg–),[190c] that solve the solubility and aggregation problems in the original design (with –Gly–D/L-Pro–Gly– as β-turn loop sequences).[156,190a,250] These new peptide ligands folded predictably in the presence of Fe(II) ions yielding stable, and water-soluble, chiral dinuclear helicates. In particular, the Fe(II) peptide helicate **ΛΛ-Fe^{II}_2LLD** (Scheme 1) exhibited high affinity and selectivity towards 3WJs compared with dsDNA *in vitro*, highlighting its ability to selectively label DNA replication foci in the cell nuclei.[190c] This was a major breakthrough in the recognition of 3WJs structures at the cellular level, and prompts further studies regarding the potential of these new helicates as artificial metallonucleases targeting this non-canonical DNA structure.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

[250] A. M. C. Marcelino, L. M. Gierasch, *Biopolymers* **2008**, *89*, 380–391.

Objetive

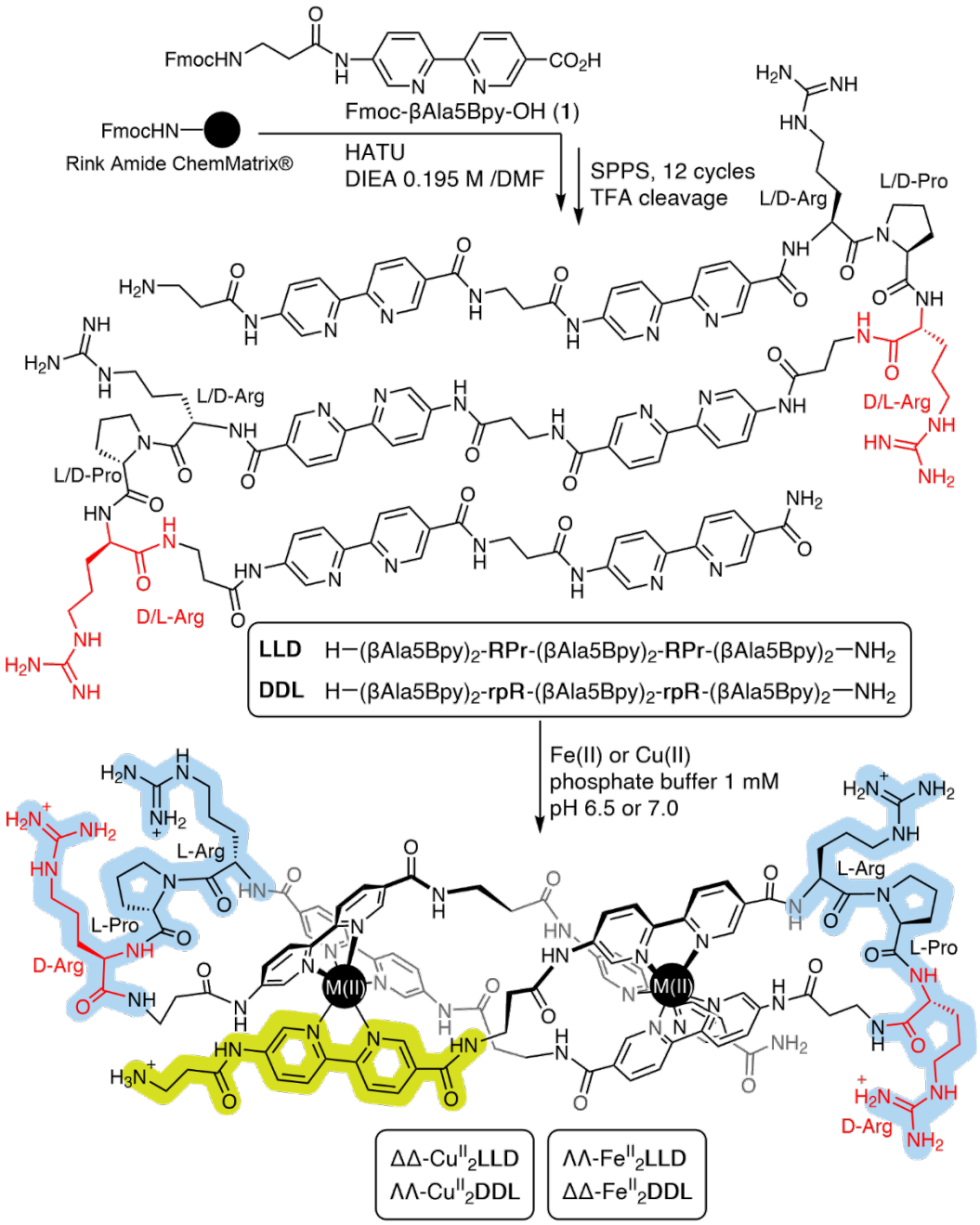
Based on the well-known nuclease properties of copper complexes,[240-248,249b] and the *in vitro* and *in cellulo* 3WJ-binding selectivity shown by **ΛΛ-Fe$^{II}_2$LLD**,[190c] the objective of this chapter is the development and study of the first 3WJ-specific artificial metallonuclease based on a Cu(II) peptide helicate derived from the previously described peptide ligands **LLD/DDL**. Since DNA replication is deregulated in cancer cells, and 3WJ is a non-canonical DNA structure transiently generated during this key cell process, we believe this strategy may pave the way for the development of a new class of anticancer drugs based on copper-based artificial nucleases.

Results and discussion

Synthesis and characterization of Cu(II) peptide helicates

The peptide ligands **LLD** and **DDL** were assembled using standard Fmoc solid-phase peptide synthesis procedures.[190c] The Bpy ligand was introduced as an Fmoc-protected dipeptide building block, Fmoc-βAla5Bpy-OH (**1**, Scheme 2), which was synthesized following the reported route that we have described in the introduction.[156] The final peptide ligands were purified by RP-HPLC, and once isolated, they were characterized by mass spectrometry (Figure 61 and Figure 62).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



Scheme 2. Sequences of the peptide ligands **LLD** and **DDL** equipped with six units of the Fmoc-βAla5Bpy-OH (**1**) building block. Synthesis of the peptides following the protocols of the SPPS methodology. Heterochiral β-turn sequences (RPr or rpR) are highlighted in light blue. After cleavage and purification, the Fe(II) and Cu(II) helicates are stereoselectively folded in the presence of the corresponding divalent metal ions. L-amino acids are indicated in upper-case while D-amino acids are in lower case (i.e., R, P refer to L-Arg and L-Pro, r and p to D-Arg and D-Pro).

The formation in solution of the peptide helicate $\mathbf{Cu^{II}_2LLD}$ (Scheme 2) was characterized by monitoring the quenching of the fluorescence emission band of the Bpy residues in **LLD** upon metal ion coordination (Figure 32). This was possible because the 5'-amido-[2,2'-bipyridine]-5-carboxamide unit in **1** is highly emissive ($\Phi \approx 0.37$),[251] with a maximum emission peak centered at 420 nm upon excitation at 308 nm. Thus, the

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

[251] A. T. Rhys Williams, S. A. Winfield, J. N. Miller, *Analyst* **1983**, *108*, 1067–1071.

titration of a 2 µM solution of the **LLD** peptide in 1mM phosphate buffer (10 mM NaCl, pH 7.0) with increasing amounts of $CuCl_2 \cdot 2\ H_2O$ lead to a decrease in the emission intensity. The resulting titration profile at 420 nm was fitted to a 1:2 (L:M) binding mode with dissociation constants of approximately 0.3 and 0.15 µM for the first and the second Cu(II) ions.[252] MALDI spectra of the solution at saturating concentrations of Cu(II) ions showed a peak at m/z = 2573.731, consistent with the formation of $\mathbf{[Cu^{II}_2LLD]^{4+}}$ (Figure 65). $\mathbf{Fe^{II}_2LLD}$ was synthesized following reported procedures.[190c]

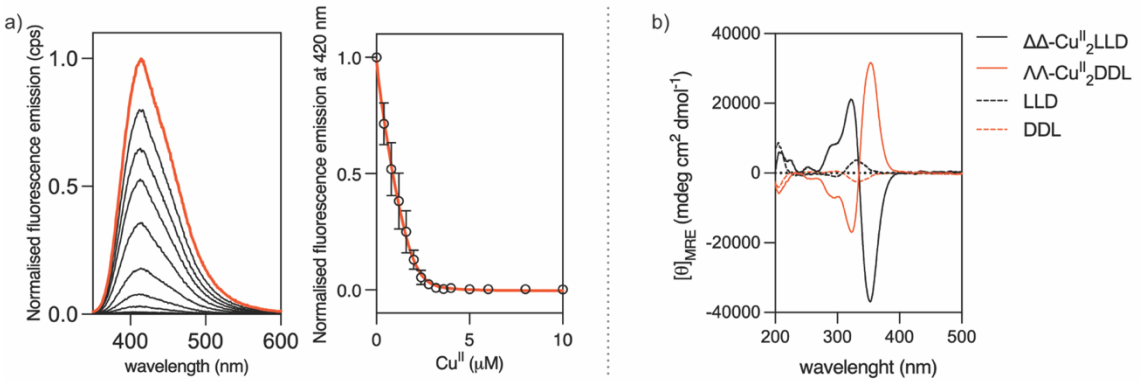


Figure 32. a) Left: Normalized emission spectrum of a 2 µM solution of **LLD** in 1 mM phosphate buffer NaCl 10 mM, pH 7.0 (orange), and spectra of the same solution in the presence of increasing concentrations of Cu(II) ions (black lines). Right: Titration profile of three independent experiments at 420 nm and best fit according to the 1:2 model in *DynaFit* (same range as main plot). b) CD spectra of 5 µM solutions of **LLD** (black dashed line) and **DDL** (orange dashed line) ligands in 1 mM phosphate buffer, 10 mM NaCl, pH 7.0, and in the presence of 25 µM Cu(II) ions (same colours, continuous lines).

The CD spectrum of the peptide ligand **LLD** is dominated by a positive Cotton effect band at *ca*. 320 nm. This indicated an important preorganization of the peptide strand and the effective chiral induction of the heterochiral β-turn loops, in combination with π-stacking interactions between the bipyridines. Addition of Cu(II) ions to a solution of **LLD** resulted in an inverted Cotton effect at this wavelength, coupled with a large increase in its intensity. This is consistent with the formation of a ΔΔ- (or *P*) helicate (Figure 32b). Surprisingly, this is the opposite chirality to that induced by Fe(II) ions with the same peptide ligand, which forms the diasteromeric (*M*) helicate.[190c] We also synthesized the isomeric $\mathbf{Cu^{II}_2DDL}$ helicate (Scheme 2) which, as expected, showed a mirror image CD spectra to that of $\mathbf{Cu^{II}_2LLD}$, consistent with a ΛΛ configuration of the Cu(II) complex (*i.e.*, with the formation of a *M* helicate) (Figure 32b).

It is known that the d^9 configuration of Cu(II) complexes results in pronounced Jahn-Teller distortion for octahedral six-coordinate systems and subsequent kinetic lability of axial ligands, even in chelated systems.[253] It appears that the differences in the coordination behavior of Fe(II) and Cu(II) lead to different arrangements of the peptide ligand around the metal centers, resulting in the opposite helical chirality of the corresponding peptide helicates. To better define the coordination of Cu(II) ions in the helicates, we performed EPR spectra at low (120 K) and room temperature (300 K) using

[252] a) P. Kuzmič, *Anal. Biochem.* **1996**, *237*, 260–273; b) P. Kuzmič, in *Methods in Enzymology* (Eds.: M.L. Johnson, L. Brand), Academic Press, **2009**, pp. 247–280.
[253] a) D. H. Powell, L. Helm, A. E. Merbach, *J. Chem. Phys.* **1991**, *95*, 9258–9265; b) M. A. Halcrow, *Chem. Soc. Rev.* **2013**, *42*, 1784–1795.

a 200 µM solution of **ΔΔ-Cu$^{II}_2$LLD** in phosphate buffer. At low temperature, such as a frozen solution, the spectra of this metallopeptide displayed a relatively symmetric signal, with an isotropic g = 2.197, suggesting regular coordination, which corresponds to octahedral or trigonal prismatic environments. In contrast, at room temperature, the narrowing of the lines suggests a possible rhombic distortion (Figure 66), as observed in other Cu(II) trisbidentate octahedral complexes due to Jahn-Teller distortion.[254]

DNA-binding properties and nuclease activity studies *in vitro*

Some helicates selectively bind to 3WJs by inserting into the central hydrophobic cavity.[184,190,226] To study the DNA binding properties of the new Cu(II) peptide helicates, we measured the fluorescence emission of a 2 µM solution of a fluorescein-labelled 3WJ (Flu-3WJ) in 1mM phosphate buffer, NaCl 10 mM, pH 7, after addition of increasing amounts of the pre-formed **ΔΔ-Cu$^{II}_2$LLD** or **ΛΛ-Cu$^{II}_2$DDL** peptide helicates (λ_{exc} = 490 nm). The progressive quenching of the fluorescein emission at 515 nm could be fitted to a 1:1 binding mode (Flu-3WJ:helicate) plus nonspecific interactions, with apparent dissociation constant K_D = 260 nM for **ΔΔ-Cu$^{II}_2$LLD** and K_D = 654 nM for **ΛΛ-Cu$^{II}_2$DDL** (Figure 33).[252] These results point to a higher affinity of the *P* helicate, **ΔΔ-Cu$^{II}_2$LLD**, for the 3WJ structures as compared to the isomeric *M*-helicate.

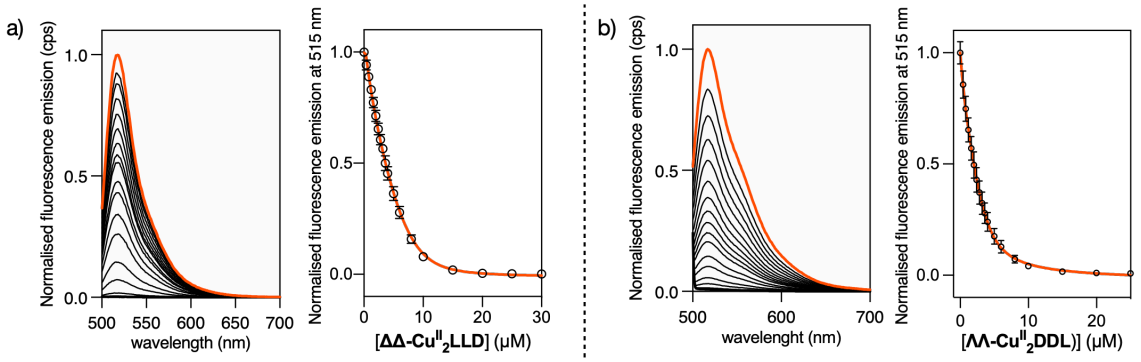


Figure33. a) Left: full emission spectra of a 2 µM solution of the fluorescein-labelled 3WJ (Flu-3WJ) in absence (orange line) and in the presence of increasing concentrations of **ΔΔ-Cu$^{II}_2$LLD** (black lines of decreasing intensity). Right: profile of the full titration with the best fit (orange line) according to the 1:1 model plus unspecific interaction in *DynaFit*, with an apparent dissociation constant K_D = 260 nM. b) Left: full emission spectra of a 2 µM solution of fluorescein-labelled 3WJ (Flu-3WJ) in absence and in presence of increasing concentrations of **ΛΛ-Cu$^{II}_2$DDL** (decreasing intensity). Right: profile of the full titration with the best fit (orange line) according to the 1:1 model plus unspecific interaction in *DynaFit*, with an apparent dissociation constant K_D = 654 nM. Flu-3WJ oligonucleotide sequences: *5'–FAM–TTTT CAC CGC TCT GGT CCT C–3'; 5'–CAG GCT GTG AGC GGT G–3'; 5'–GAG GAC CAA CAG CCT G–3'*.

Following the same methodology, we also tried, unsuccessfully, to calculate the K_D numbers of the **ΔΔ-Cu$^{II}_2$LLD** binding to dsDNA, since the mixture precipitated during the titration, even prior to reaching the helicate saturation concentration, thus suggesting a nonspecific interaction between the peptide helicate and the canonical dsDNA structure (See 4.1 of Chapter I in Experimental Section). However, we measured the K_D of **ΔΔ-Cu$^{II}_2$LLD** with a different 3WJ sequence (Flu-GC-3WJ), resulting in a very similar value to that obtained for Flu-3WJ (300 nM), which indicates that the affinity of this Cu(II) peptide helicate for 3WJs is not limited to a single DNA sequence (Figure 67).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

[254] M. V. Veidis, G. H. Schreiber, T. E. Gough, G. J. Palenik. *J. Am. Chem. Soc.* **1969**, *91*, 1859–1860.

Furthermore, control experiments performed with the free peptide ligand **LLD** demonstrate the need for the Cu(II) helicate to be fully assembled, that is, coordinated to the metal ions, for it to acquire 3WJ recognition properties (Figure 70).[190c]

To further analyze the binding of **ΔΔ-Cu$^{II}_2$LLD** with 3WJ and dsDNA, we performed electrophoretic mobility shift assays (EMSA) in polyacrylamide gel (PAGE) under non-denaturing conditions[255] using *SYBRGold* for DNA staining.[256] In agreement with the fluorescence studies, we observed that incubation of 3WJ with increasing concentrations of the pre-formed **ΔΔ-Cu$^{II}_2$LLD** helicate gave rise to the appearance of a new slower-migrating band with no smearing, corresponding to the formation of the 3WJ/**ΔΔ-Cu$^{II}_2$LLD** adduct (Figure 34a_**B**). As can be seen, almost all the 3WJ has been complexed with **ΔΔ-Cu$^{II}_2$LLD**, even at the lowest concentration of helicate tested, as expected for a high-affinity interaction. The intensity of this band (Figure 34a_**B**) increases with the concentration of the Cu(II) helicate, while the intensity of the lower band (Figure 34a_**A**), related to unassembled 3WJ strands, decreases in consonance. In contrast, no changes in the position or intensity of the band associated to dsDNA (Figure 34b_**C**) were observed upon incubation of the pre-formed **ΔΔ-Cu$^{II}_2$LLD** helicate with *off*-target dsDNA in the presence of target 3WJ. These results confirm the poor affinity of **ΔΔ-Cu$^{II}_2$LLD** for canonical B-DNA (Figure 34b).

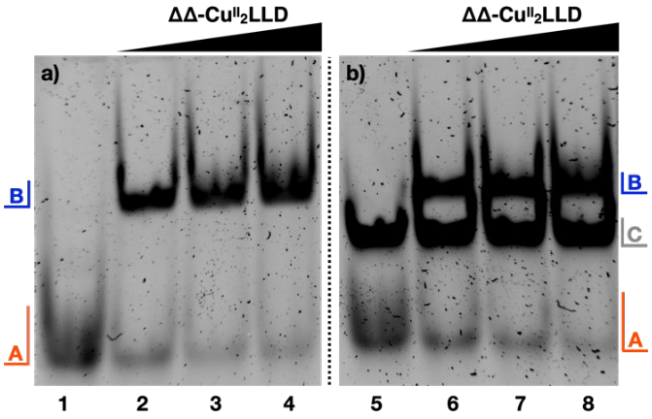


Figure 34. a) PAGE analysis of 200 nM of 3WJ in the absence (lane 1) and in the presence of 200 (lane 2), 400 (lane 3) and 600 nM (lane 4) of **ΔΔ-Cu$^{II}_2$LLD**; b) PAGE analysis of 200 nM of 3WJ + 200 nM of dsDNA in the absence (lane 5) and in the presence of 200 (lane 6), 400 (lane 7) and 600 (lane 8) nM of **ΔΔ-Cu$^{II}_2$LLD**. In these experimental PAGE conditions, the lower intensity bands of lane 1 and 5 (**A**) are unassembled ssDNA strands of 3WJ. The new, slow-migrating band (**B**) in lanes 2-4 and 6-8 corresponds to the 3WJ structure assembled in the presence of **ΔΔ-Cu$^{II}_2$LLD**. The intensity of band **B** in a) and b) increases with the concentration of the Cu(II) helicate, while the intensity of the lower band (**A**) decreases. In contrast, the position and intensity of the band corresponding to dsDNA in lanes 5-8 (**C**) remains unaffected in the presence of increasing concentrations of the Cu(II) helicate, demonstrating the low affinity of this metallopeptide for canonical DNA. Note that, using these experimental PAGE conditions (See 2.9.1 of Procedures and experimental techniques in Experimental Section), 3WJ is only assembled in the presence of the Cu(II) peptide helicate. 3WJ oligonucleotide sequences: *5'–CAC CGC TCT GGT CCT C–3'; 5'–CAG GCT GTG AGC GGT G–3'; 5'–GAG GAC CAA CAG CCT G–3'*. dsDNA oligonucleotide sequence: *5'–AAC ACA TGC AGG ACG GCG CTT–3'* (only one strand shown).

[255] a) L. M. Hellman, M. G. Fried, *Nat. Protoc.* **2007**, *2*, 1849–1861; b) D. Lane, P. Prentki, M. Chandler, *Microbiol. Rev.* **1992**, *56*, 509–528.

[256] R. S. Tuma, M. P. Beaudet, X. Jin, L. J. Jones, C.-Y. Cheung, S. Yue, V. L. Singer, *Anal. Biochem.* **1999**, *268*, 278–288.

We performed PAGE experiments with labelled oligos to simplify the visualization of target 3WJ and *off*-target dsDNA (Figure 35a). 3WJ strands were labelled at their 5' end with FAM, ROX and Cy5, while the dsDNA strands were labelled with FAM and Cy5. Fluorophore-labelled 3WJ and dsDNA strands were then incubated with **ΔΔ-Cu$^{II}_2$LLD** and visualized using a multiplex fluorescent assay (Figure 35a). The blue (FAM), green (ROX) and red (Cy5) signals were combined to produce a yellow 3WJ band, while the blue (FAM) and red (Cy5) dsDNA strands are seen as pink.

The same multiplex binding assay was subsequently employed to analyse the DNA-cleavage profile of **ΔΔ-Cu$^{II}_2$LLD**. With this aim, the ability of this metallopeptide to selectively recognize and cleave 3WJs *in vitro* was also tested through PAGE (Figure 35b). The fluorescently labelled target 3WJ and *off*-target dsDNA oligonucleotides were incubated with increasing equivalents of **ΔΔ-Cu$^{II}_2$LLD** and Na-(L)-ascorbate at 37 ºC for 24 hours. PAGE analysis revealed that this Cu(II) peptide helicate can selectively cleave 3WJ, while the canonical dsDNA in the sample remains undamaged when exposed up to 7.5 equivalents of **ΔΔ-Cu$^{II}_2$LLD**.

In addition, a control cleavage experiment was also performed using **ΛΛ-Fe$^{II}_2$LLD**. In this case, 3WJ and dsDNA were incubated with the Fe(II) peptide helicate and Na-(L)-ascorbate under the same conditions to that of the Cu(II) derivative assay, followed by visualization of the cleavage profile using PAGE (Figure 35c). Band densitometry revealed that **ΔΔ-Cu$^{II}_2$LLD** successfully displays high cleavage selectivity for the 3WJ target against dsDNA, as for example, just 2% of the 3WJ remained versus 80% of *off*-target dsDNA following exposure to 5 eq. of the Cu(II) peptide helicate (Figure 35d and Figure 76). The next step in our work was the determination of the incubation period required to achieve 3WJ selective damage using a time point study (Figure 77). The observed results suggest that cleavage initiates within 1 hour of incubation and increases up to 12 hours (from 12 to 24 hours there was no significant increase in the cleavage). Accordingly, it can be concluded that **ΔΔ-Cu$^{II}_2$LLD** can rapidly and selectively damage 3WJs, and that *off*-target dsDNA damage would not be expected after longer incubation times.

Combining the results showed in Figures 34 and 35 it seems clear that **ΔΔ-Cu$^{II}_2$LLD** can discriminate between 3WJs and dsDNA and induce selective cleavage of the 3WJ *in vitro*. To our knowledge, this is the first example of a chemical nuclease that performs such activity.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

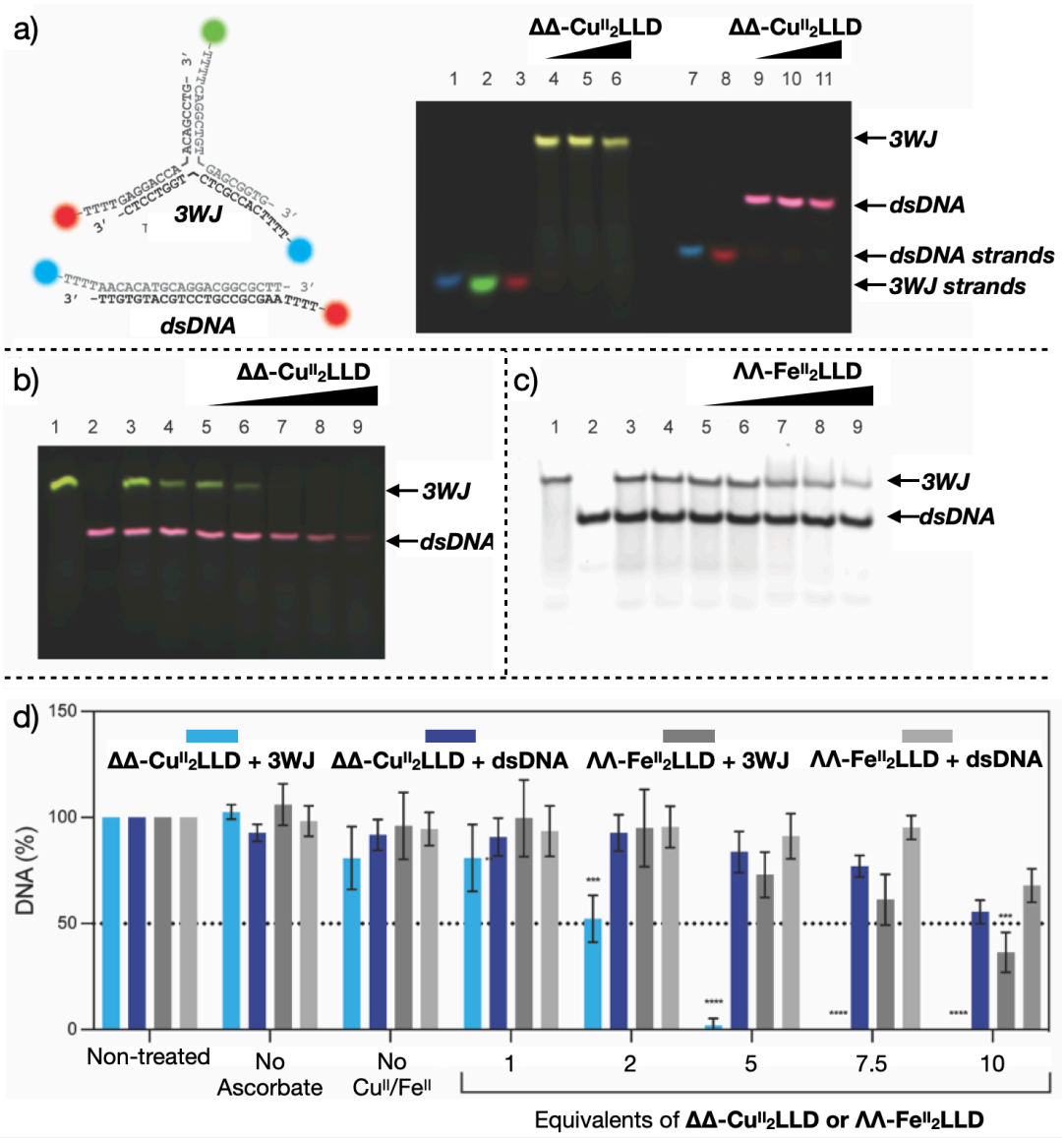


Figure 35. a) Fluorescently labelled target 3WJ and *off*-target dsDNA. Blue = FAM, Green = ROX and Red = Cy5 (See 2.9.2 of Procedures and experimental techniques in Experimental Section). Multiplex image from PAGE analysis of the formation of 3WJ DNA in the presence of increasing concentrations of **ΔΔ-Cu$^{II}_2$LLD**. Lane 1; FAM-labelled single strand of 3WJ. Lane 2; ROX-labelled single strand. Lane 3; 3WJ strand labelled with Cy5. Lane 4-6; 3WJ strands incubated with 1, 2 or 3 equivalents of **ΔΔ-Cu$^{II}_2$LLD**, respectively. Lane 7; single strand of dsDNA labelled with FAM. Lane 8: Cy5-labelled strand of dsDNA. Lanes 9-11; dsDNA incubated with 1, 2 or 3 equivalents of **ΔΔ-Cu$^{II}_2$LLD**, respectively (Figure 72). b) PAGE analysis of the DNA damage inflicted by **ΔΔ-Cu$^{II}_2$LLD**. Fluorophore labelled DNA was utilized to ensure clarity of target 3WJ DNA (yellow) and *off*-target dsDNA (pink). Lane 1; 3WJ. Lane 2; dsDNA. Lane 3; 3WJ and dsDNA incubated with **ΔΔ-Cu$^{II}_2$LLD**. Lane 4: target and *off*-target with **LLD** peptide ligand (no Cu(II) present). Lanes 5-9; 3WJ and dsDNA incubated with 1, 2, 5, 7.5 or 10 eq. of **ΔΔ-Cu$^{II}_2$LLD** and Na-(L)-ascorbate (10 eq. to **ΔΔ-Cu$^{II}_2$LLD)** (Figure 73). c) Control cleavage experiment performed with **ΛΛ-Fe$^{II}_2$LLD**. Lane 1; 3WJ DNA. Lane 2; dsDNA *off*-target. Lane 3; control of 3WJ and dsDNA incubated with **ΛΛ-Fe$^{II}_2$LLD**. Lane 4: 3WJ and dsDNA with non-metalated **LLD** peptide ligand. Lanes 5-9; 3WJ and dsDNA incubated with 1, 2, 5, 7.5 or 10 eq. of **ΛΛ-Fe$^{II}_2$LLD** and Na-(L)-ascorbate (10 eq. to **ΛΛ-Fe$^{II}_2$LLD)**. d) Band densitometry analysis of cleavage induced upon 3WJ and dsDNA by **ΔΔ-Cu$^{II}_2$LLD** (Figure 74) and **ΛΛ-Fe$^{II}_2$LLD** (Figure 35c and Figure 75). * = $P \leq 0.05$ ** = $P \leq 0.01$ *** = $P \leq 0.001$ and **** = $P \leq 0.0001$. Fluorescently labelled 3WJ oligonucleotide sequences: *5'–FAM–TTTT CAC CGC TCT GGT CCT C–3'; 5'–ROX–TTTT CAG GCT GTG AGC GGT G–3'; 5'– Cy5–TTTT GAG GAC CAA CAG CCT G–3'*. Fluorescently labelled dsDNA oligonucleotide sequence: *5'–FAM–TTTT AAC ACA TGC AGG ACG GCG CTT–3'*; *5'–Cy5–TTTT AAG CGC CGT CCT GCA TGT GTT–3'*.

In order to shed light on the cleavage mechanism induced by **ΔΔ-Cu$^{II}_2$LLD**, 3WJ was incubated with the Cu(II) peptide helicate, Na-(L)-ascorbate, and different selective scavengers for specific ROS (L-histidine for $^1O_2$; D-mannitol for $^{\bullet}OH$; L-methionine for $^{\bullet}OH$, HOCl and H_2O_2; or sodium catechol sulfate, known as tiron, for $O_2^{\bullet-}$),[246b,257] and the resulting cleavage profiles were visualized by PAGE (Figure 36). Figure 36b shows that the presence of tiron significantly reduced the activity of the Cu(II) metallopeptide, thus indicating that superoxide ($O_2^{\bullet-}$) plays a key role in the 3WJ cleavage inflicted by **ΔΔ-Cu$^{II}_2$LLD**. L-histidine also impacted the 3WJ damage, which suggested that singlet oxygen ($^1O_2$) was also involved in the cleavage mechanism. However, since $^1O_2$ is not usually generated under photoactivation conditions, we decided to further investigate the role of this ROS in the Cu(II) peptide helicate activity. Therefore, the scavenger assay was repeated including deuterium oxide (D_2O) and sodium azide (NaN_3), known for being a $^1O_2$ stabilizer and a $^1O_2$ scavenger, respectively (Figure 36a). Surprisingly, the presence of D_2O and NaN_3 did not disrupt 3WJ cleavage, hence it is not clear at this point how $^1O_2$ mediates the **ΔΔ-Cu$^{II}_2$LLD** nuclease activity. Nevertheless, it should be noted that several recent studies reporting the dsDNA damaging effects of metallonucleases based on polynuclear Cu(II) complexes have predicted $^1O_2$-mediated DNA damage and identified L-histidine as viable scavenger, both *in vitro* and *in cellulo*.[258]

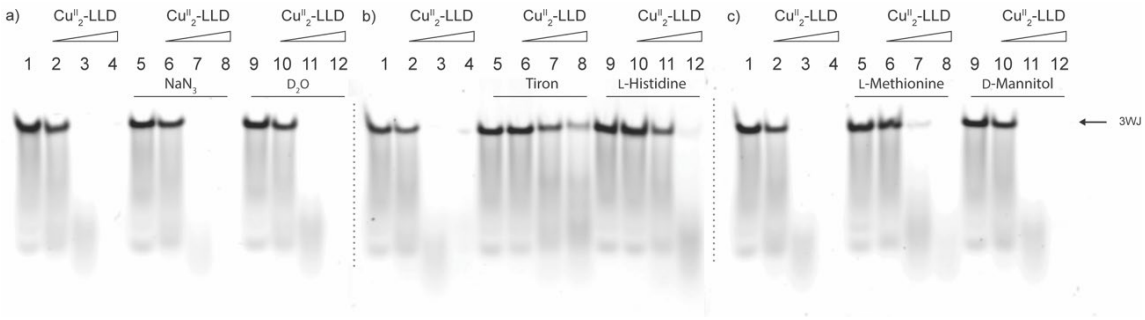


Figure 36. 20% native PAGE showing the cleavage profile of the **ΔΔ-Cu$^{II}_2$LLD** in the presence of ROS scavengers; a) sodium azide (NaN_3) and deuterium oxide (D_2O); b) tiron and L-histidine; c) L-methionine and D-mannitol. Lane 1-4: control (no scavenger). Lanes 5-8: a) NaN_3, b) tiron, c) L-methionine. Lanes 9-12: a) D_2O, b) L-histidine, c) L-methionine. Lanes 1,5,9: 3WJ control. Lanes 2,6,10: 1 eq. **ΔΔ-Cu$^{II}_2$LLD**. Lanes 3,7,11: 5 eq. **ΔΔ-Cu$^{II}_2$LLD**. Lanes 4,8,12: 10 eq. **ΔΔ-Cu$^{II}_2$LLD**. All lanes (excluding controls) contain Na-(L)-ascorbate (10 eq. to **ΔΔ-Cu$^{II}_2$LLD**).

DNA-binding properties and nuclease activity studies in mammalian cells

In preliminary experiments, a TAMRA-labelled Cu(II) peptide helicate, **ΔΔ-Cu$^{II}_2$TAMRA-LLD**, failed to internalize in Vero and HeLa cells. However, it readily translocated in digitonin-treated cells, where this helicate showed a diffused distribution in the cytosol, and a strong punctuated distribution in the nucleus, as well as an evident concentration in the nucleolus (Figure 37 and Figure 79). Most of the nuclear spots co-localized with GFP-tagged Proliferating Cellular Nuclear Antigen (PCNA), which

[257] A. Kellett, Z. Molphy, C. Slator, V. McKee, N. P. Farrell, *Chem. Soc. Rev.* **2019**, *48*, 971–988.
[258] a) N. McStay, C. Slator, V. Singh, A. Gibney, F. Westerlund, A. Kellett, *Nucleic Ac. Res.* **2021**, *49*, 10289‑10308; b) C. Slator, Z. Molphy, V. McKee, C. Long, T. Brown, A. Kellett, *Nucleic Acids Res.* **2018**, *46*, 2733–2750.

specifically labels DNA replication foci in actively dividing cells.[259] While its cytosolic distribution and association with the nucleoli might reflect some *off*-target RNA-binding of the helicate, replisomes labelled by PCNA are enriched in 3WJs due to the presence of replication forks. Thus, our results strongly suggest the direct association of the Cu(II) peptide helicate with 3WJs structures in functional cell nuclei. Furthermore, control experiments showed that the **LLD** peptide ligand in absence of Cu(II) metal ions does not localize in the DNA replication centers (Figure 78). This control experiment also demonstrates that **ΔΔ-Cu^{II}_2TAMRA-LLD** is stable in cell medium and that it is only capable of binding to DNA replication centres in the nucleus as a discrete metallopeptide.

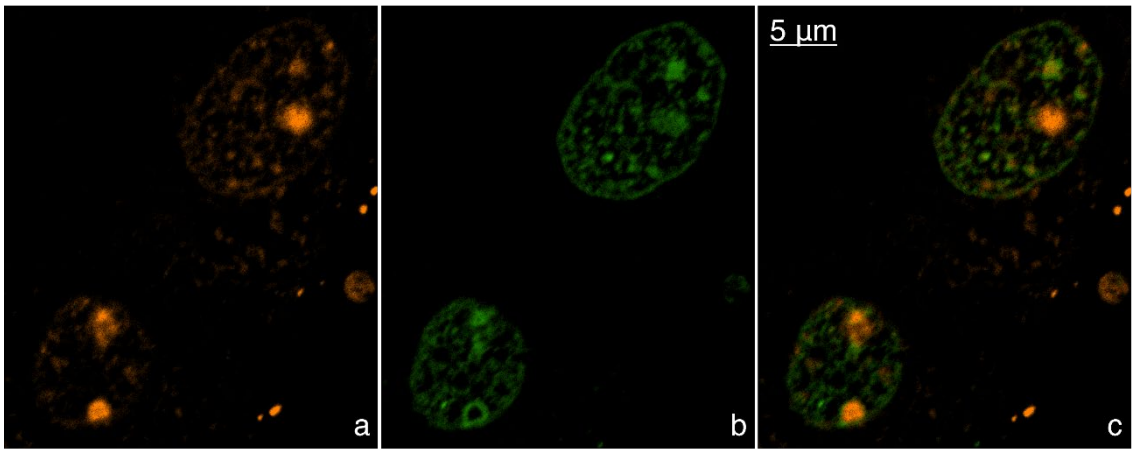


Figure 37. ΔΔ-Cu^{II}_2TAMRA-LLD selectively stains DNA replication sites. Vero cells expressing protein GFP-PCNAL2 were incubated in the presence of 25 µg/mL of digitonin for 3min, and then in the presence of 5 µM of **ΔΔ-Cu^{II}_2TAMRA-LLD** for 15 min. The Cu(II) peptide helicate was pre-formed by adding 5 equivalents of $CuCl_2$ • 2 H_2O to an aqueous solution of **TAMRA-LLD** peptide ligand. a) Red channel emission showing the distribution of the TAMRA-labelled helicate; b) green channel, corresponding to the emission of the GFP-PCNAL2 probe that labels DNA replication foci; c) overlay of the green and red channels. Nucleoli are clearly labelled as big red spots inside the cell nuclei. The image shows a representative confocal section of the Vero cells.

With the aim to investigate the stability of **ΔΔ-Cu^{II}_2LLD** in water media in order to assess its ability to carry out its nuclease activity *in cellulo*, we designed a competitive assay with EDTA, an organic ligand that readily forms complexes with copper ions of extraordinary thermodynamic stability.[260] This assay showed that the ligand substitution reaction for **ΔΔ-Cu^{II}_2LLD** takes longer than five minutes to occur in absence of 3WJ, but it slows down to more than 20 minutes in the presence of the target 3WJ, thus indicating that **ΔΔ-Cu^{II}_2LLD** possesses remarkably kinetic inertness in water media and suggests that it also exhibits very large stability *in cellulo*, especially, in the cell nucleus when bound to its target 3WJs (Figure 69).

Having demonstrated that **ΔΔ-Cu^{II}_2LLD** associates with 3WJ-enriched foci in cellular DNA, we studied if it can exerts its nuclease activity upon chromatin-associated 3WJs in cells. To maximize the number of cells undergoing DNA replication, and thus increase the chances of detecting free DNA ends obtained by nuclease cleavage, we used

[259] a) H. Leonhardt, H.-P. Rahn, P. Weinzierl, A. Sporbert, T. Cremer, D. Zink, M. C. Cardoso, *J. Cell Biol.* **2000**, *149*, 271–280; b) R. Yokoyama, T. Hirakawa, S. Hayashi, T. Sakamoto, S. Matsunaga, *Sci. Rep.* **2016**, *6*, 29657.
[260] S. Chaberek Jr, A. E. Martell, *J. Am. Chem. Soc.* **1955**, *77*, 1477–1480.

of Vero cells that were previously synchronized by 42 h of serum deprivation.[261] 16 h after resupplementation of the cultures with 10% serum, when most of the cells were in *S phase*, the cells were pre-treated with digitonin and then incubated with H_2O_2 in the presence or absence of **ΔΔ-Cu$^{II}_2$LLD**. A control experiment was also performed with **ΛΛ-Fe$^{II}_2$LLD** in the presence of H_2O_2, as this peptide helicate was earlier identified to have very poor nuclease activity *in vitro* (Figures 35c and 35d). Cells were then incubated for 2 h at 37 ºC and subjected to a TUNEL assay to detect free DNA ends in the cell nuclei.[262] Results show that incubation of **ΔΔ-Cu$^{II}_2$LLD** in the presence of H_2O_2 produced strong and clear increase on the TUNEL signal (Figure 38), indicating the presence of free DNA ends and thus, nuclease activity *in cellulo*. Importantly, this increase in signal was not observed in the H_2O_2 and H_2O_2+**ΛΛ-Fe$^{II}_2$LLD** control experiments.

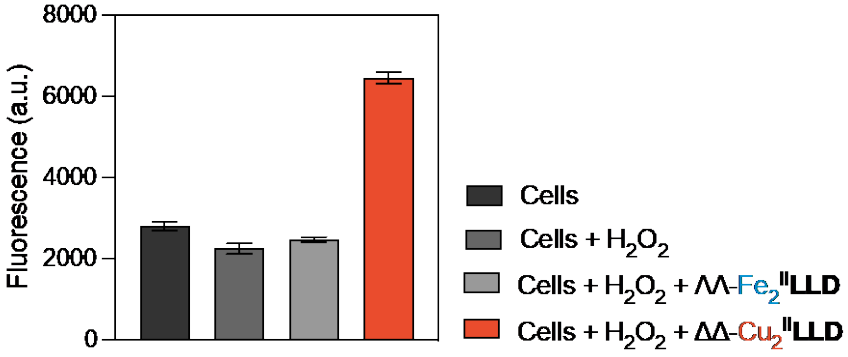


Figure 38. Detection of free DNA ends by TUNEL assay. S-phase synchronized Vero cells were incubated with H_2O_2 in the absence or presence of the Fe(II) or Cu(II) peptide helicates indicated in the figure at 37 ºC for 2 hours, and then subjected to TUNEL analysis for the detection of free DNA ends. Results are represented as the mean with standard deviation of a triplicate experiment.

[261] R. M. Bringhurst, P. A. Schaffer, *J. Virol.* **2006**, *80*, 4528–4537.

[262] TUNEL Assay (commercial) Kit detects the DNA fragmentation of apoptotic cells by exploiting the fact that the DNA breaks expose many 3'-hydroxyl ends. These hydroxyl groups can then serve as starting points for terminal deoxynucleotidyl transferase (TdT), which adds deoxyribonucleotides in a template-independent fashion. Addition of the deoxythymidine analog 5-bromo-2'-deoxyuridine 5'-triphosphate (BrdUTP) to the TdT reaction serves to label the break sites. Once incorporated into the DNA, BrdU can be detected by an anti-BrdU antibody using standard immunohistochemical techniques.

Conclusions

Using standard protocols of the SPPS methodology, we synthesized new oligocationic Cu(II) peptide helicates derived from a peptide ligand bearing 2,2'-bipyridine chelating residues. **ΔΔ-Cu^{II}_2LLD** displays high affinity and binding selectivity for 3WJ both *in vitro* and in functional cells. In addition, this new Cu(II) peptide helicate can selectively cleave 3WJs through the generation of superoxide radicals. Noteworthy, selective cleavage of DNA replication centers was also observed wit high efficiency in functional mammalian cells. To our knowledge, this is the first example of a synthetic chemical nuclease able to selectively cleave 3WJ both *in vitro* and *in cellulo*. It is also important to highlight that just by changing the metal cation in the helicate, Fe(II) or Cu(II), while maintaining the same peptide ligand, **LLD**, metallodrugs with unique and distinctive biological properties are obtained, which can be applied in different scenarios. For instance, the Fe(II) peptide helicate selectively recognizes and stabilizes replication centers in the cell nucleus, while the Cu(II) helicate, on the other hand, can selectively cleave these centres. The work described in this chapter can be considered as a proof of concept and lay the basis for the development of a new class of anticancer metallodrugs, targeting 3WJ structures at the DNA replication foci of cell nuclei with great precision.

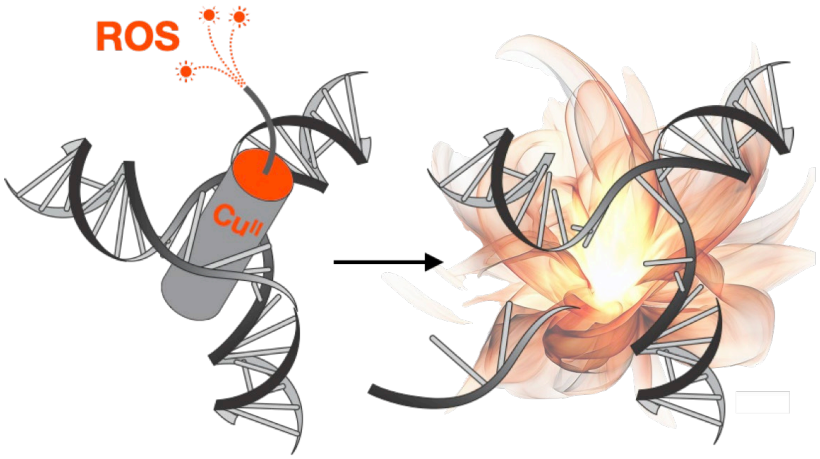


Figure 39. Illustration of the 3WJ-specific metallonuclease approach described in this chapter.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter II

Chiral parallel peptide helicates with trigonal symmetry

Summary

In this chapter we show that a C_3 symmetric peptidomimetic containing 2,2'-bipyridine chelating residues folds in the presence of Co(II) ions into a parallel triple-stranded peptide helicate with a preferred chirality defined by the residues in its sequence. Additionally, *in situ* oxidation of the thermodynamically-selected Co(II) chiral helicate affords the kinetically inert Co(III) helicate. Both Co(II) and Co(III) complexes display good binding selectivity towards 3WJs against canonical DNA.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Introduction

Metallosupramolecular helicates stand out as promising agents for recognition of non-canonical three-way junctions DNA (3WJs) due to their high affinity and selectivity for these structures.[190,197] One of the key factors for their selectivity is the complementarity between the structure and trigonal symmetry of the helicates with the cavity in the branch point of the 3WJs.[184,263] This has been exploited for the development of other 3WJ binders, such as triptycene derivatives,[264] whose *threefold symmetry* and non-planar π-surface,[265] similar in size to the hydrophobic core of 3WJs allows effective π-stacking interactions with the nucleobases at the junction interface.[24,184,263,266] Likewise, C_3-symmetric cationic azacryptands and self-assembled supramolecular $Fe^{II}_4L_4$ tetrahedral cages also exhibit high affinity and selectivity for 3WJs over dsDNA.[267,268]

With these precedents in mind, we envisioned that a C_3 symmetric core could be conjugated to three peptide sequences equipped with coordinating residues derived from 2,2'-bipyridine. The resulting trigonal ligand should be able to self-assemble in the presence of Co(II) ions into enantiomerically pure triple-stranded dinuclear helicates with binding selectivity towards 3WJs.

Furthermore, it has been shown that the C_3-symmetric tripeptides attached to aromatic cores, such as benzenetricarboxamide, can self-assemble into one-dimensional columnar aggregates with amplified chirality (such as fibers and rods), stabilized by intermolecular hydrogen bonding of amide groups and π-π stacking interactions between the aromatic cores,[269] and the bipyridinyl groups in the peptide chains can further stabilize these aggregates through additional π-π stacking interactions.[270] The resulting supramolecular polymers have a great potential in drug delivery, immunoengineering and regenerative therapies,[269b] as well as for chiral recognition, separation of enantiomers,

---

[263] S. Phongtongpasuk, S. Paulus, J. Schnabl, R. K. O. Sigel, B. Spingler, M. J. Hannon, E. Freisinger, *Angew. Chem. Int. Ed Engl.* **2013**, *52*, 11513–11516.

[264] a) S. A. Barros, D. M. Chenoweth, *Angew. Chem. Int. Ed Engl.* **2014**, *53*, 13746–13750; b) S. A. Barros, I. Yoon, D. M. Chenoweth, *Angew. Chem. Int. Ed Engl.* **2016**, *55*, 8258–8261.

[265] a) P. D. Bartlett, M. J. Ryan, S. G. Cohen, *J. Am. Chem. Soc.* **1942**, *64*, 2649–2653; b) H. Hart, S. Shamouilian, Y. Takehira, *J. Org. Chem.* **1981**, *46*, 4427–4432.

[266] T. M. Long, T. M. Swager, *J. Am. Chem. Soc.* **2002**, *124*, 3826–3827.

[267] a) J. Novotna, A. Laguerre, A. Granzhan, M. Pirrotta, M.-P. Teulade-Fichou, D. Monchaud, *Org. Biomol. Chem.* **2015**, *13*, 215–222; b) K. Duskova, J. Lamarche, S. Amor, C. Caron, N. Queyriaux, M. Gaschard, M.-J. Penouilh, G. de Robillard, D. Delmas, C. H. Devillers, A. Granzhan, M.-P. Teulade-Fichou, M. Chavarot-Kerlidou, B. Therrien, S. Britton and D. Monchaud, *J. Med. Chem.* **2019**, *62*, 4456-4466; c) L. Guyon, M. Pirrotta, K. Duskova, A. Granzhan, M.-P. Teulade-Fichou, D. Monchaud, *Nucleic Acids Res.* **2018**, *46*, e16.

[268] J. Zhu, C. J. E. Haynes, M. Kieffer, J. L. Greenfield, R. D. Greenhalgh, J. R. Nitschke, U. F. Keyser, *J. Am. Chem. Soc.* **2019**, *141*, 11358–11362.

[269] a) O. Zagorodko, V. J. Nebot, M. J. Vicent, *Polym. Chem.* **2020**, *11*, 1220–1229; b) M. A. VandenBerg, J. K. Sahoo, L. Zou, W. McCarthy, M. J. Webber, *ACS Nano* **2020**, *14*, 5491–5505; c) J. Wang, L. Chen, J. Wu, W. Li, K. Liu, T. Masuda, A. Zhang, *Chem. Asian J.* **2018**, *13*, 3647–3652.

[270] J. J. Van Gorp, J. A. J. M. Vekemans, E. W. Meijer, *J. Am. Chem. Soc.* **2002**, *124*, 14759–14769.

asymmetric catalysis or conductive systems, among others.[271] These precedents suggest that this new ligand would be prone to dynamically aggregate into one-dimensional columnar structures, due to its trigonal symmetry and aromaticity.[269,270] and this aggregation could be exploited for controlling the 3WJ binding.

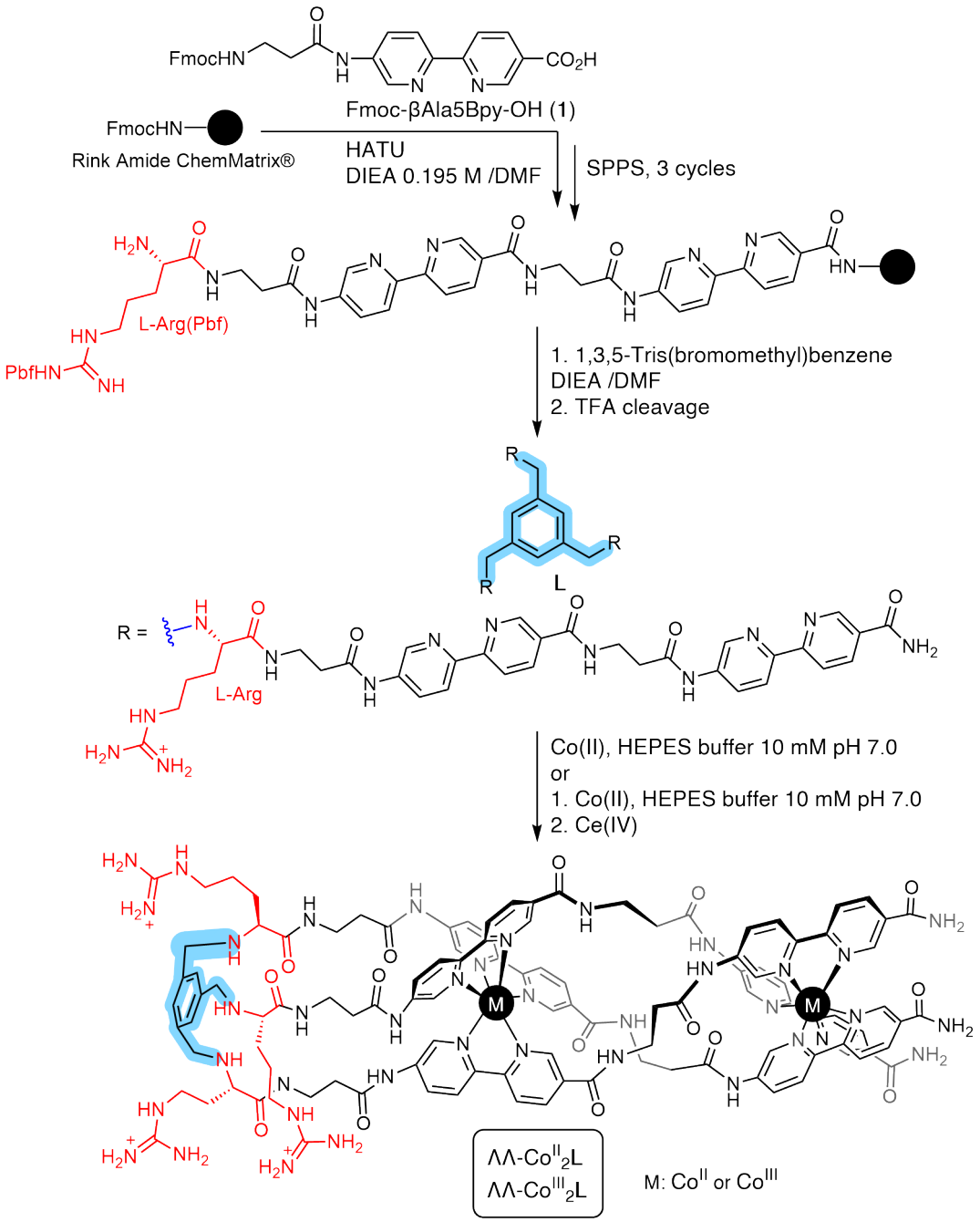
Results and discussion

Design, synthesis, and characterization of the peptide helicates and helical fibers

The newly designed peptide ligand **L** consists of three peptide strands attached to a central trisubstituted benzene core, that provides the system with trigonal symmetry. Each of these three peptide strands is composed of two Fmoc-βAla5Bpy-OH (**1**, Scheme 3) building blocks in tandem,[156] that together will allow the coordinating needs of two octahedral metal ions, and one L-Arg residue acting as a linker between the coordinating residues and the aromatic core. These chiral L-Arg residues should thermodynamically control the formation of chirally pure helices and, thanks to their cationic nature, improve the solubility of the ligand in aqueous media in absence of metal ions, and increase the overall affinity of the helicate for the DNA by providing additional electrostatic contacts with the DNA backbone.[272,273,274]

The peptide chains were synthesized using standard Fmoc-SPPS protocols. Once the peptide sequences were fully assembled, their N-terminal amines were reacted with 1,3,5-tris(bromomethyl)benzene via a S_N2 reaction while the peptides were still attached to the solid support. The combined deprotection and cleavage step of the final peptide ligand was performed with an acidic cocktail of TFA, and the desired ligand was purified by RP-HPLC, and its identity confirmed by mass spectrometry (Figure 80).

[271] E. Yashima, N. Ousaka, D. Taura, K. Shimomura, T. Ikai, K. Maeda, *Chem. Rev.* **2016**, *116*, 13752–13990.
[272] J. DeRouchey, B. Hoover, D. C. Rau, *Biochemistry* **2013**, *52*, 3000–3009.
[273] S. M. West, R. Rohs, R. S. Mann, B. Honig, *J. Biomol. Struct. Dyn.* **2010**, *27*, 861–866.
[274] D. P. Mascotti, T. M. Lohman, *Biochemistry* **1997**, *36*, 7272–7279.



Scheme 3. Solid-phase peptide synthesis of the peptide ligand **L** equipped with six Fmoc-βAla5Bpy-OH (**1**) units, three (L)-arginine residues and the benzene aromatic core, which confers a C_3 symmetry to the system. After cleavage and purification, the Co(II) peptide helicate is stereoselectively folded under thermodynamic control in the presence of Co(II) metal ions. This Co(II) peptide helicate can be oxidized *in situ* with Ce(IV) to give rise the kinetically inert Co(III) metallopeptide.

In order to monitor the coordination and formation in solution of the $\mathbf{Co^{II}_2L}$ helicate (Scheme 3) we measured the quenching effect in the emission intensity of the Bpy units, at c.a. 420 nm,[251] upon addition of increasing amounts of Co(II) ions (Figure 40a). Thus, a 2 μM solution of the **L** peptide ligand in 10 mM HEPES buffer (100 mM NaCl, pH 7.0) was titrated with increasing amounts of $Co(ClO_4)_2 \cdot 6\ H_2O$ and the emission spectra were measured after each addition. This resulted in a decrease in the fluorescence intensity at 420 nm, which was fitted to a 1:2 (L:M) binding mode with

dissociation constants of approximately K_{D1} = 11.2 μM and K_{D2} = 315 μM for the first and the second Co(II) ions.[252] These binding constants were much weaker than those measured with other self-assembled peptide helicates,[190c] which suggested that the peptide ligand **L** may be aggregated.

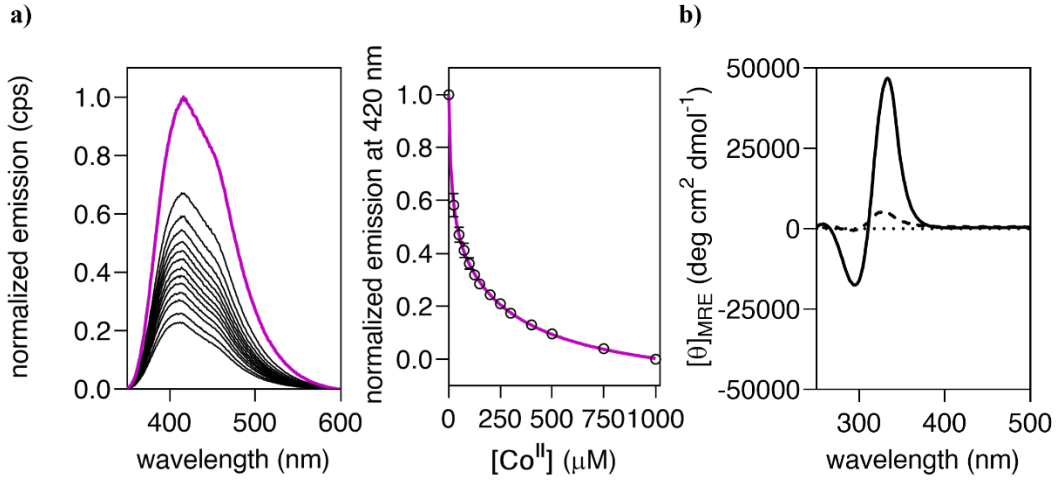


Figure 40. a) Titration of the **L** peptide ligand from a MQ water stock solution. Left: normalized emission spectrum of a 2 μM solution of **L** (purple), and spectra of the same solution in the presence of increasing amounts of Co(II) ions (black lines). Right: titration profile at 420 nm of three independent experiments with the best fit according to a 1:2 binding model in *DynaFit*.[252] b) Circular Dichroism spectra of 50 μM solutions of **L** derived from a 1.1 mM (continuous line) and 240 μM (dashed line) stock solutions. Conditions: 10 mM HEPES buffer, NaCl 100 mM, pH 7.0.

To test our hypothesis, we measured the CD spectra of two 50 μM solutions of **L** (Figure 40b), prepared from stocks with different peptide concentrations (240 μM and 1.1 mM), in which **L** should exist in different aggregation states. Both CD spectra display a positive Cotton effect band at *ca*. 320 nm, which points to a significant preorganization of the peptide ligand **L**, together with the effective chiral induction by the L-Arg residue. However, the most remarkable aspect is the difference in the intensity between both CD spectra, which appears to confirm a higher degree of aggregation of **L** in the more concentrated stock solution, resulting in the amplified chirality observed in the CD spectrum.

To obtain reproducible measurements, and avoid differences in the peptide aggregation state, a new peptide stock solution was prepared using DMSO as solvent to disrupt the intermolecular hydrogen bonds that promote the formation of supramolecular assemblies.[275] Next, we studied the formation of the Co(II) helicate using this new peptide ligand stock. To this end, a fluorescence titration was carried out using independent solutions of **L** at 2 μM in 10 mM HEPES buffer (100 mM NaCl, pH 7.0) with increasing different concentrations of Co(II) ions. These samples were incubated overnight to ensure that the mixture reached the equilibrium, and their respective emission spectra were then recorded. The new titration profile at 420 nm saturated at much lower concentrations of Co(II) (10 eq) and could be fitted to a 1:1 (L:M_2) binding model, resulting in an apparent dissociation constant of $K_D \approx 1.32$ μM,[252] consistent with the previously reported systems.

[275] a) K.-I. Oh, K. Rajesh, J. F. Stanton, C. R. Baiz, *Angew. Chem. Int. Ed.* **2017**, *56*, 11375–11379; b) M. Jackson, H. H. Mantsch, *Biochim. Biophys. Acta* **1991**, *1078*, 231–235.

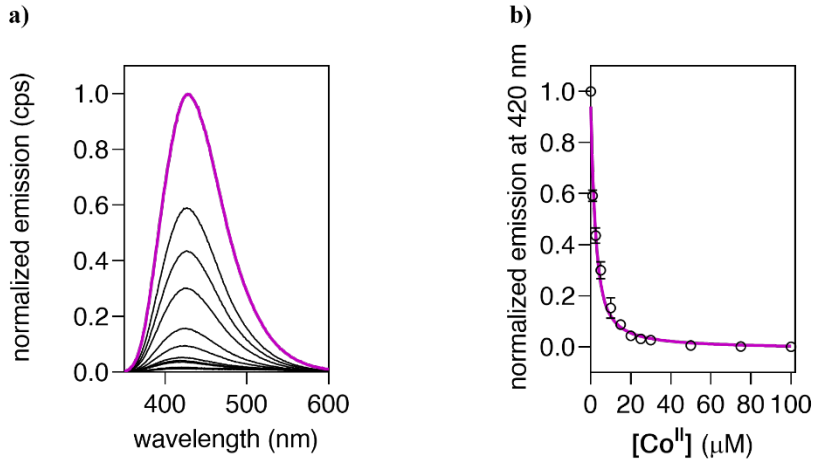


Figure 41. Titration of the **L** peptide ligand from a DMSO stock solution. a) Normalized emission spectrum of a 2 μM solution of **L** (purple), and spectra of different solutions of **L** at 2 μM incubated overnight in the presence of different amounts of Co(II) ions (black lines), in 10 mM HEPES buffer, NaCl 100 mM, pH 7.0. b) Titration profile at 420 nm of three independent experiments, and the best fit according to a 1:1 ($L:M_2$) binding model in *DynaFit*.[252]

Once we confirmed that **L** self-assembles into supramolecular aggregates with amplified chirality, we decided to take a closer look into this behavior. Thus, using the DMSO stock solution of **L**, we measured the CD spectra of a 10 μM solution of the peptide ligand in 10 mM HEPES buffer (100 mM NaCl, pH 7.0) at increasing temperatures from 0 to 95 ºC (Figure 42a), which resulted in a diminution of the band intensity at 320 nm, thus indicating the loss of aggregation upon heating up the sample. The melting profile measured at 320 nm (Figure 42b), showed wo transitions around 25 and 60 ºC, which are consistent with alignment and disaggregation processes, respectively. To gain further information and to analyse the influence of metal ions on the system organization, UV-Vis studies of a 10 μM solution of **L,** in the absence and presence of Co(II) ions, were also performed at different temperatures (Figure 42c). In absence of metal ions, the intensity of the 2,2'-bipyridine absorbance band at 308 nm decreases and the shoulder around 340 nm, associated to the supramolecular assembly, disappears as temperature goes up. This confirms the loss of aggregation at high temperatures observed by CD. The shoulder is again missing in the spectra measured with 15 eq. of Co(II), thus indicating that presence of metal ions promotes the formation of discrete metal complexes rather than supramolecular aggregates. It should be noted that the increased intensity at 95 ºC may be associated with a more concentrated sample due to an evaporation process.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

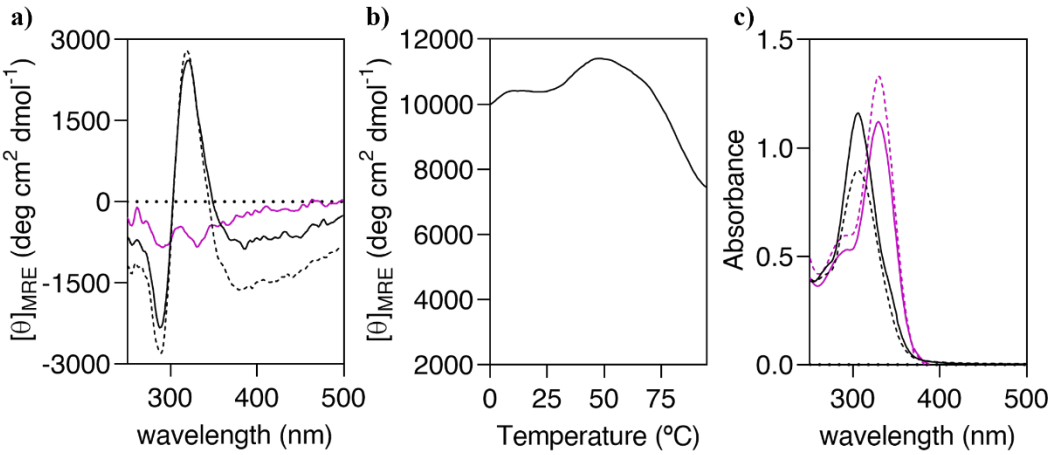


Figure 42. a) CD spectra of a 10 µM solution of **L** at 0 ºC (black dashed line), 20º C (black continuous line) and 95 ºC (purple line). b) Melting profile at 320 nm measured by increasing the temperature of a 10 µM solution of **L** from 0 to 95 ºC. c) UV-Vis spectra of a 10 µM solution of **L** at 20 ºC (black continuous line) and 95 ºC (black dashed line), and in the presence of 150 µM of Co(II) ions at 20º C (purple continuous line) and 95 ºC (purple dashed line). Conditions: 10 mM HEPES buffer, NaCl 100 mM, pH 7.0.

Finally, to study the stereoselectivity in the synthesis of the Co(II) dinuclear peptide helicate, we measured the CD spectra of a 10 µM solution of free **L** ligand, and the same solution incubated overnight with 150 µM of Co(II) ions. The self-assembly of the Co(II) metallopeptide resulted in a slight shift in the positive Cotton effect band from 320 to 330 nm, thus confirming the effective Co(II) coordination, yielding to the formation of the corresponding discrete **ΛΛ-Co^{II}_2L** peptide helicate (Figure 43).

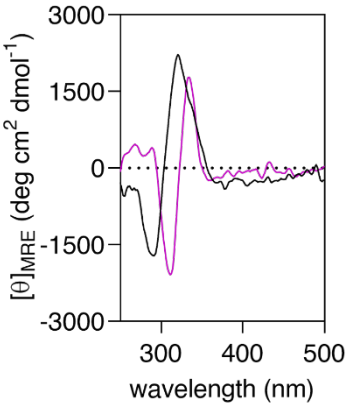


Figure 43. CD spectra of **L** peptide ligand at 10 µM (black line) from the DMSO stock solution, and in the presence of 150 µM of Co(II) ions (purple line). The measurements were performed in 10 mM HEPES buffer, 100 mM NaCl, pH 7.0 at 20 ºC.

To determine the nature of the aggregates formed by the **L** peptide in aqueous solution, DLS assays and microscopy studies were carried out using solutions of the free **L** peptide ligand at 10 µM, and in the presence of 15 eq. of Co(II) ions, both of them in water and DMSO (Figure 44a and Figure 44b). DLS studies provide information about the behavior of aggregates in solution. The data obtained through this technique are the hydrodynamic diameter (diameter of the aggregate + solvent) and the PDI, which gives information about the similarity of the aggregates: the lower it is, the lower the polydispersities of the aggregates are. In the first case (Figure 44a, left), the aqueous solution of **L** (10 µM) shows the presence of a DLS curve with a majority signal whose

size is 1647 nm and with a PDI = 0.389. This indicates the presence of aggregates of considerable size and with a rather high polydispersity, perfectly compatible with the formation of fibers. The presence of Co(II) ions (15 eq.) in the above solution (Figure 44a, right) causes the formation of larger aggregates and an increase in polydispersity at the same time. This behavior is perfectly compatible with a coordination of Co(II) ions to the peptide ligand and the formation of the corresponding helicates. This is an assumption that will be confirmed by the results of the microscopy images, as the polydispersity of these measurements is so high that it is not possible to know with precision the type of aggregate present in the solution. Finally, control experiments were performed using DMSO as solvent in the absence and presence of Co(II) ions (Figure 44b). Under these conditions no aggregates should form. The DLS curves indicate the presence of small aggregates with an average size of 482 nm and very high polydispersity. These results confirm that no distinct aggregates form in DMSO that can be detected by this technique.

Moreover, STEM and TEM microscopy images demonstrate the formation of helical fibers in aqueous solution at low peptide concentration (10 μM) under metal-free conditions (Figure 44c and 44d). The formation of the fibers is likely driven by intermolecular π-π stacking interactions between the benzene central blocks and the Bpy residues in **L**, as well as by hydrogen bonding between amide groups.[269,270] As we expected, the presence of Co(II) ions was shown to inhibit the peptide aggregation process as seen in TEM images (44d). When using DMSO as solvent, in control experiments, the formation of fibers is also inhibited. These results support our hypothesis based on the fluorescence, CD and UV-Vis assays.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

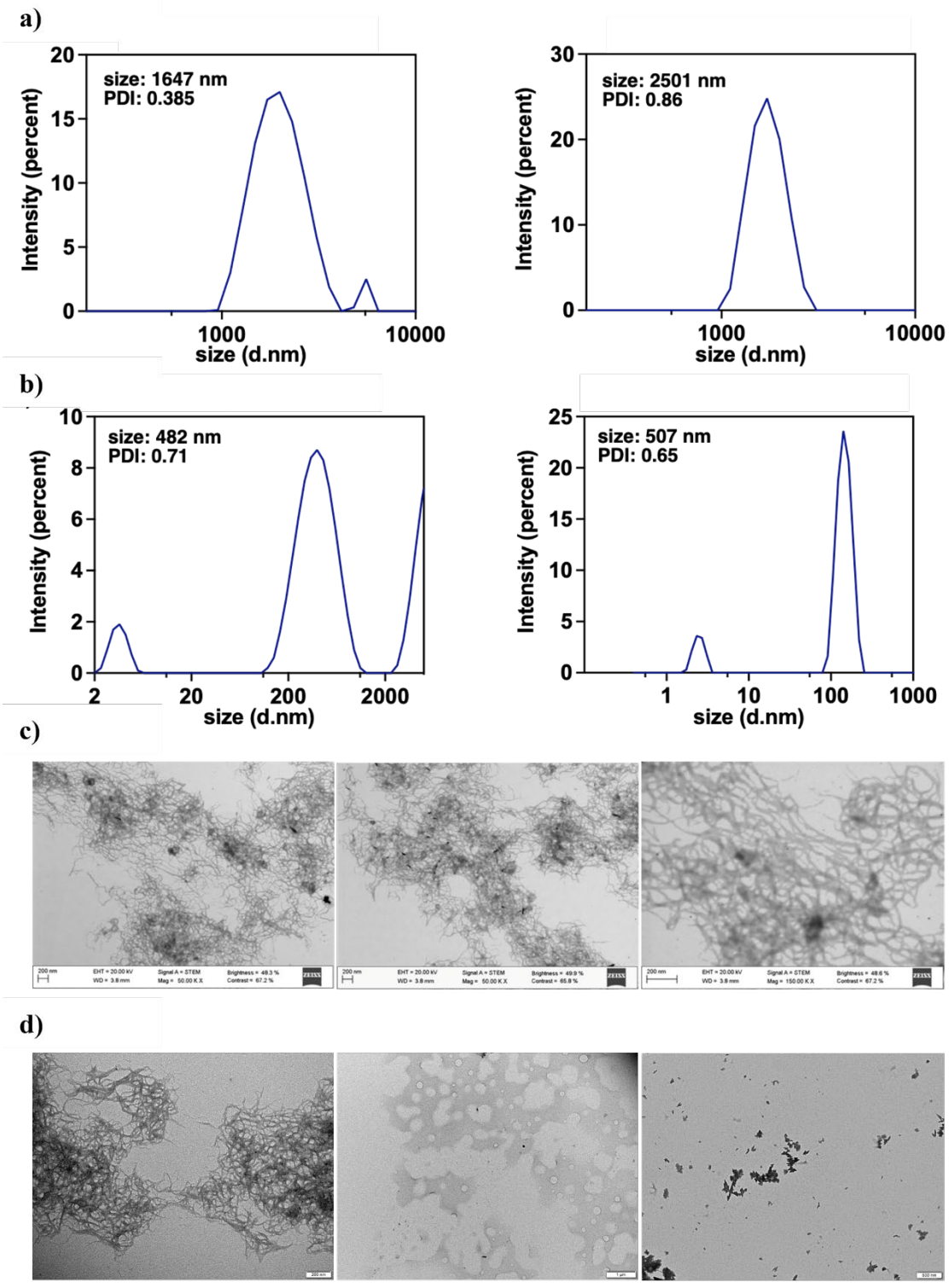


Figure 44. a) DLS curve of a 10 μM solution of **L** in water in absence (left) and presence (right) of 15 eq. of Co(II) ions. b) Same experiment but in DMSO. c) Electron microscopy (STEM) studies were performed to determine the shape and size of the aggregates formed by **L** (10 μM) in water in the absence of Co(II) ions. This provides a 2D image where the darkest areas will be those of organic matter. Carbon coated grids and drop casting deposition were used to perform these studies. The STEM images show the formation of very elongated fibers with an average width of 20 nm and SD = 5 nm (50 measurements). d) Left: TEM electron microscopy image of a 10 μM aqueous solution of **L**, showing fibers with an average width of 18 nm, consistent with the STEM results; Middle: TEM electron microscopy image of a 10 μM aqueous solution of **L** and 15 eq. of Co(II) ions, where no fibers are observed, but rather polydisperse aggregates; Right: TEM electron microscopy image of a 10 μM solution of **L** in DMSO.

Regarding the **ΛΛ-Co$^{II}_2$L** helicate, the results presented above prove that this self-assembled Co(II) metallopeptide is intrinsically dynamic, thus restricting its potential applications.[276] In order to achieve a more stable system, once **ΛΛ-Co$^{II}_2$L** has been dynamically assembled, Co(II) was oxidized *in situ*, without a change in the coordination geometry, to give the kinetically inert Co(III) metallopeptide, using the assembly-followed-by-oxidation protocol previously reported by Lusby and co-workers.[277] Thus, a 200 μM aqueous solution of **L** was incubated with 3 mM of $Co(ClO_4)_2 \cdot 6H_2O$. Thereafter, the sample was treated with 1.1 eq. of $(NH_4)_2Ce(NO_3)_6$ (3.3 mM) with respect to Co(II), acting as oxidant, which lead to the corresponding dinuclear Co(III) complex (**Co$^{III}_2$L**), whose identity was confirmed, upon HPLC purification, by UHPLC-MS (Figure 81) and MALDI-TOF mass spectrometry (Figure 83), where the peak with m/z = 2359.1 is consistent with the successful synthesis of the Co(III) peptide helicate by *in situ* oxidation and after the self-assembly of the corresponding Co(II) metallopeptide precursor. The chirality of the kinetically inert **Co$^{III}_2$L** was tested by circular dichroism, where the positive Cotton effect band around 336 nm, is consistent with the formation of the corresponding **ΛΛ-Co$^{III}_2$L** peptide helicate (Figure 82).

DNA-binding properties of the Co(II) and Co(III) peptide helicates

Once we demonstrated the synthesis of Co(II) and Co(III) peptide helicates derived from **L**, we characterized their 3WJ binding properties.[184,190,226] In order to do this, we monitored the quenching of the fluorescein emission in a 2 μM solution of fluorescein-labelled 3WJ (Flu-3WJ) in 10 mM HEPES buffer (100 mM NaCl, pH 7.0) upon addition of increasing amounts of the preformed **ΛΛ-Co$^{II}_2$L** or **ΛΛ-Co$^{III}_2$L** helicates (λ_{exc} = 490 nm). In the case of the Co(II) metallopeptide, the resulting titration profile at 515 nm could be fitted to a 1:1 binding mode (Flu-3WJ:**ΛΛ-Co$^{II}_2$L**) plus nonspecific interactions with an apparent K_D of 359 nM (Figure 45a).[252] However, the titration profile at 515 nm differs for the interaction between Flu-3WJ with the kinetically inert Co(III) helicate, which was fitted to a 1:1+1:3 binding model (Flu-3WJ:**ΛΛ-Co$^{III}_2$L**) in *Dynafit*. This model is characterized by two different dissociation constants, K_{D1} = 0.128 μM and K_{D2} = 0.825 μM, likely corresponding to the insertion of **ΛΛ-Co$^{III}_2$L** into the central cavity of 3WJ and to its interactions with the hydrophobic surface of the three equivalent blunt ends of 3WJ, respectively (Figure 45b and Figure 85).[252] Following the same strategy, we could not measure the K_D of the **ΛΛ-Co$^{II}_2$L** binding to double-stranded DNA, given that the mixture precipitated before reaching the Co(II) helicate saturation concentration, which points to a nonspecific/electrostatic interaction between the metallopeptide and the canonical DNA (Figure 84).

[276] W. M. Bloch, G. H. Clever, *Chem. Commun.* **2017**, *53*, 8506–8516.

[277] a) D. A. Leigh, P. J. Lusby, R. T. McBurney, A. Morelli, A. M. Z. Slawin, A. R. Thomson, D. B. Walker, *J. Am. Chem. Soc.* **2009**, *131*, 3762–3771; b) D. A. Leigh, P. J. Lusby, A. M. Z. Slawin, D. B. Walker, *Chem. Commun.* **2012**, *48*, 5826–5828; c) P. R. Symmers, M. J. Burke, D. P. August, P. I. T. Thomson, G. S. Nichol, M. R. Warren, C. J. Campbell, P. J. Lusby, *Chem. Sci.* **2015**, *6*, 756–760.

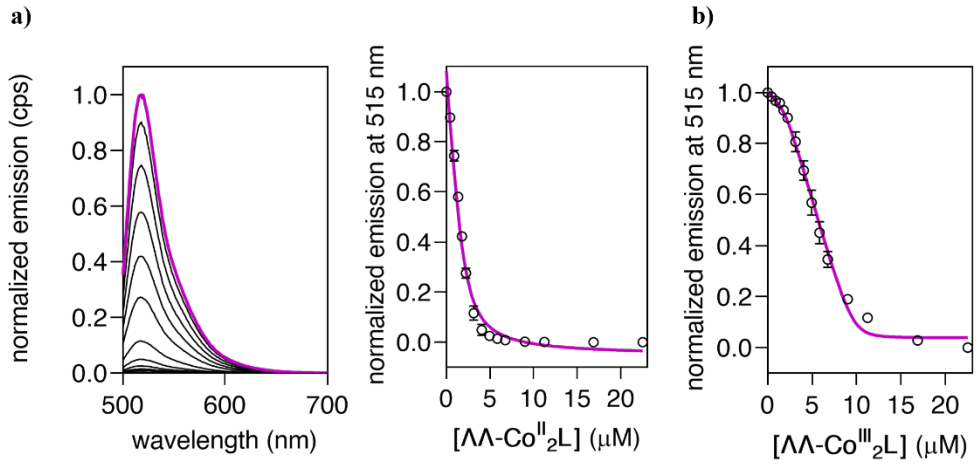


Figure 45. a) Fluorescence titration of a fluorescein-labelled 3WJ DNA (Flu-3WJ) with the Co(II) peptide helicate. Left: normalized emission spectra of the 2 μM solution of Flu-3WJ in absence (purple line) and in the presence of increasing concentrations of **ΛΛ-Co$^{II}_2$L** (black lines of decreasing intensity). Right: profile at 515 nm of the same fluorescence titration experiment with the best fit to a 1:1 binding model (purple line);[252] b) Fluorescence titration profile at 515 nm of Flu-3WJ at 2 uM upon addition of increasing amounts of **ΛΛ-Co$^{III}_2$L**, with the best fit to a 1:1+1:3 binding mode in *DynaFit* (purple line).[252] Titration data are the average of three independent experiments. Conditions: 10 mM HEPES buffer, 100 mM NaCl, pH 7.0; λ_{exc} = 490 nm; Flu-3WJ oligonucleotide sequences: *5'–FAM–TTTT CAC CGC TCT GGT CCT C–3'; 5'–CAG GCT GTG AGC GGT G–3'; 5'–GAG GAC CAA CAG CCT G–3'*.

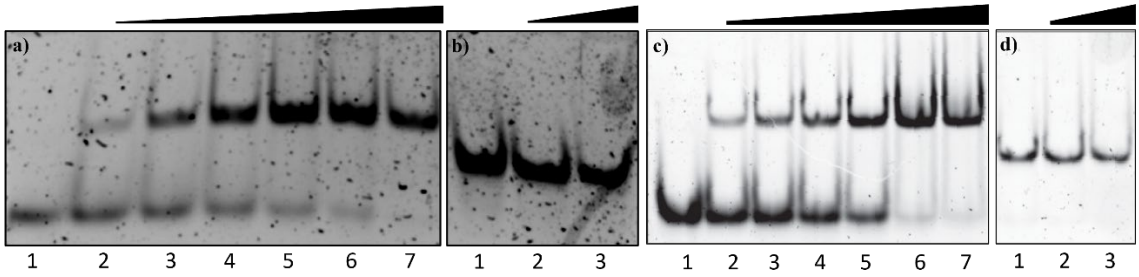


Figure 46. EMSA DNA binding studies results for **ΛΛ-Co$^{II}_2$L** and **ΛΛ-Co$^{III}_2$L** helicates. a) Lanes 1-7: 200 nM of 3WJ with 0, 25, 50, 100, 150, 500 and 750 nM of **ΛΛ-Co$^{II}_2$L** in each lane; b) Lanes 1-3: 200 nM of dsDNA with 0, 500 and 1000 nM of **ΛΛ-Co$^{II}_2$L** in each lane. c) Lanes 1-7: 200 nM of 3WJ with 0, 25, 50, 100, 150, 500 and 750 nM of **ΛΛ-Co$^{III}_2$L** in each lane; d) Lanes 1-3: 50 nM of dsDNA with 0, 500 and 1000 nM of **ΛΛ-Co$^{III}_2$L** in each lane. 3WJ oligonucleotide sequences: *5'–CAC CGC TCT GGT CCT C–3'; 5'–CAG GCT GTG AGC GGT G–3'; 5'–GAG GAC CAA CAG CCT G–3'*. dsDNA oligonucleotide sequence: *5'–AAC ACA TGC AGG ACG GCG CTT–3'* (only one strand shown).

To further characterize the DNA-binding properties of the Co(II) and Co(III) peptide helicates, we run electrophoretic mobility shift assays (EMSA) in polyacrylamide gel (PAGE) under non-denaturing conditions,[255] by using *SYBRGold*, a fluorescent dye that intercalates into DNA.[256] In agreement with the results of the fluorescence assays, incubation of the target DNA three-way junction with increasing concentrations of the pre-formed **ΛΛ-Co$^{II}_2$L** helicate resulted in the appearance of a new slow-migrating band, associated with the formation of the 3WJ/**ΛΛ-Co$^{II}_2$L** adduct (Figure 46a, lanes 1-7). As expected, the intensity of this new band increases with the concentration of the Co(II) helicate, while the intensity of the lower free 3WJ band decreases. Note that a single band with no smearing is observed, even at high **ΛΛ-Co$^{II}_2$L** concentrations, which indicates the formation of an unique complex. On the contrary, incubation of **ΛΛ-Co$^{II}_2$L** with a model dsDNA does not induce shifts in the position of the band (Figure 46b, lanes 1-3),

thus confirming the poor affinity of this Co(II) helicate for regular B-DNA, in comparison to the 3WJ. The EMSA assay performed in the presence of the kinetically inert Co(III) metallopeptide showed affinity for the target 3WJ (Figure 46c, lanes 1-7) similar to that of the Co(II) helicate, which is consistent with the results obtained by fluorescence. Once again, incubation of **ΛΛ-Co^{III}_2L** with double-stranded DNA did not lead to the appearance of a new slow-migrating band (Figure 46d, lanes 1-3), thus supporting the high selectivity of these peptide helicates towards 3WJs versus canonical DNA *in vitro*.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Conclusions

In this chapter we have described the synthesis of a new peptide ligand made up of three peptide strands, derived from the *βAla5Bpy* chelating building blocks, attached to a central core of 1,3,5-tris(bromomethyl)benzene, which confers a trigonal symmetry to the ligand. This gives rise to a dynamic system that self-assembles into supramolecular fibers or discrete peptide helicates in absence or presence of Co(II) metal ions, respectively. The corresponding triple-stranded Co(II) peptide helicate was dynamically synthesized, and subsequently oxidized to the kinetically inert Co(III) metallopeptide derivative, both of them displaying high binding affinity and selectivity towards 3WJs *in vitro*.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter III

Efficient and Selective Cleave of 3WJ by ATCUN-Functionalized Peptide Helicates

Summary

In the first chapter, we described how an oligocationic Cu(II) peptide helicate was able to promote the selective cleavage of 3WJs through ROS production both *in vitro* and *in cellulo*, breaking DNA replication foci in cell nuclei. On the other hand, the Cu(II)/Ni(II) ATCUN binding motif is a small peptide metal-binding site found at the N-terminus of many natural proteins that exhibits excellent dsDNA cleavage properties as well as antitumor activity due to ROS formation. We reasoned that the efficiency of 3WJ cleavage by the Cu(II) peptide helicates described in Chapter I could be substantially improved by adding an ATCUN motif at the N-terminal end of the peptide ligand chain. Therefore, in this chapter we address the synthesis through SPPS methodology of Ni(II) and Cu(II) peptide helicates functionalized with the ATCUN motif as well as the *in vitro* study of their 3WJ binding properties, their *in vitro* capabilities as nucleases and their cell-internalization properties.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Introduction

In the search for novel antitumor therapies that overcome the toxicity and chemoresistance issues associated with the currently available ones, the ATCUN motif emerges as a promising approach for the design of new metallodrugs with catalytic nuclease and protease activity.[278] The ATCUN motif is a small tripeptide capable of coordinate with high affinity a single Ni(II) or Cu(II) ion and which is found at the N-terminus of many natural proteins as albumin, among others. Their derived metallopeptides have been studied for more than 35 years due to its remarkable nucleic acid cleavage properties.[279] ATCUN tripeptides form very stable complexes of square-planar coordination geometry with Cu(II) and Ni(II),[280] offering for metal ion coordination an imidazole nitrogen donor atom of a His residue located at the third position of the tripeptide, a deprotonated amide nitrogen atom of the same His residue, a deprotonated amide nitrogen atom of a residue located at the second position of the tripeptide and the NH_2-terminal group of the peptide sequence.[279,281] The extremely high thermodynamic stability of ATCUN-metallopeptides —at or above physiological pH, Xaa-Xaa-His tripeptides bind avidly to Cu(II) or Ni(II) with dissociation constants of 10^{-16}–10^{-17} M— gives them significant therapeutic potential. This is because they can be administered in their apo form and then recruit intracellular Cu(II) or Ni(II), avoiding toxicity and regulatory issues.[278] In addition, these metallopeptides do not form isomers at the metal center, eliminating the need to resolve individual species prior to DNA recognition studies.

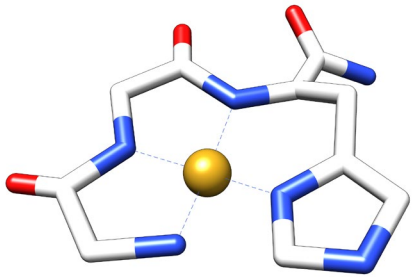


Figure 47. Structure of the square-planar Cu(II)-ATCUN (GGH) metallopeptide. The image was made using *Chimera* software.

There are so many examples in literature of ATCUN-derived metallopeptides with the generic formula M(II)-Xaa-Xaa-His (where M is Ni(II) or Cu(II) and Xaa is any amino acid) that induce DNA and RNA cleavage when activated with ascorbate or peroxide-based reagents,[282] as the nuclease activity of these metallopeptides is mediated by reactive oxygen species (ROS).[283]

---

[278] C. M. Agbale, M. H. Cardoso, I. K. Galyuon, O. L. Franco, *Metallomics* **2016**, *8*, 1159–1169.
[279] C. Harford, B. Sarkar, *Acc. Chem. Res.* **1997**, *30*, 123–130.
[280] K. P. Neupane, A. R. Aldous, J. A. Kritzer, *J. Inorg. Biochem.* **2014**, *139*, 65–76.
[281] N. Camerman, A. Camerman, B. Sarkar, *Can. J. Chem.* **1976**, *54*, 1309–1316.
[282] E. C. Long, C. A. Claussen in *Small Molecule and RNA Binders: From Synthesis to Nucleic Acid Complexes. Volume 1*. (Eds.: M. Demeunynck, C. Bailly, W. D. Wilson), Wiley-VCH, Weinheim, Germany. **2003**, *Chapter 5*, pp. 88–125
[283] Y. Jin, M. A. Lewis, N. H. Gokhale, E. C. Long, J. A. Cowan, *J. Am. Chem. Soc.* **2007**, *129*, 8353–8361.

ATCUN metallopeptides with M(II)-Xaa-Xaa-His sequence associate with the minor groove of the DNA helix and damage the DNA by interacting with the deoxyribose molecules in this groove. Their binding properties depend on the features of the peptide ligand, such as its side chains, chirality and flanking sequences. For instance, M(II)-Gly-Gly-His show moderate cleavage preferences for A/T-rich regions of dsDNA.[284] Despite their low selectivity Pauling et *al.* demonstrated that the Cu(II)-Gly-Gly-His ATCUN metallopeptide exhibits antitumor activity, being able to kill Ehrlich ascites tumor cells *in vitro*, in the presence of ascorbate, as well as to increase the lifetime of mice inoculated with these cells.[285]

A significant advantage of ATCUN derivatives is their peptidic nature, which allows their simple conjugation to additional tags to increase their target selectivity, thus reducing their toxicity and side effects.[278] So far, a few examples of chimeras combining the ATCUN motif with DNA recognition domains in natural proteins,[286] such as Zinc Fingers or the Fos protein, have already been shown to selectively bind and cleave their DNA target regions.[287] Likewise, it is possible to derivatize the the ATCUN motif to a ligand with selectivity for non-canonical DNA structures as, for example, the copper-acridine-ATCUN complex (Cu(II)-Gly-Gly-Gly-His-Lys-Acridine) reported by Cowan *et al*., selectively binds and cleaves G-quadruplex telomeric DNA, inhibiting the proliferation of cancer cells.[248]

[284] Q. Liang, P. D. Eason, E. C. Long. *J. Am. Chem. Soc*. **1995**, *117*, 9625–9631.
[285] E. Kimoto, H. Tanaka, J. Gyotoku, F. Morishige, L. Pauling, *Cancer Res.* **1983**, *43*, 824–828.
[286] a) D. P. Mack, B. L. Iverson, P. B. Dervan, *J. Am. Chem. Soc.* **1988**, *110*, 7572–7574; b) D. P. Mack, P. B. Dervan, *J. Am. Chem. Soc.* **1990**, *112*, 4604–4606; c) D. P. Mack, P. B. Dervan, *Biochemistry* **1992**, *31*, 9399–9405; d) S. L. Grokhovskiĭ, V. A. Nikolaev, V. E. Zubarev, A. N. Surovaia, A. L. Zhuze, B. K. Chernov, N. I. Sidorova, A. S. Zasedatelev, G. V. Gurskiĭ, *Mol. Biol.* **1992**, *26*, 1274–1297.
[287] a) M. Nagaoka, M. Hagihara, J. Kuwahara, Y. Sugiura, *J. Am. Chem. Soc.* **1994**, *116*, 4085–4086; b) C. Harford, S. Narindrasorasak, B. Sarkar, *Biochemistry* **1996**, *35*, 4271–4278.

Objetive

Based on the highly efficient and selective nuclease activity exhibited by chimera constructs composed of an ATCUN motif linked to a specific DNA-binding domain,[248,286,287]in this chapter we will describe the development of new chiral Cu(II) and Ni(II) peptide helicates featuring the peptide ligand **LLD** described in Chapter I,[190c] with the tripeptide ATCUN motif appended at their N-terminus. These peptide helicate chimeras should exhibit improved nuclease activity towards 3WJs compared to the Cu(II) peptide helicate previously described in Chapter I.

Results and discussion

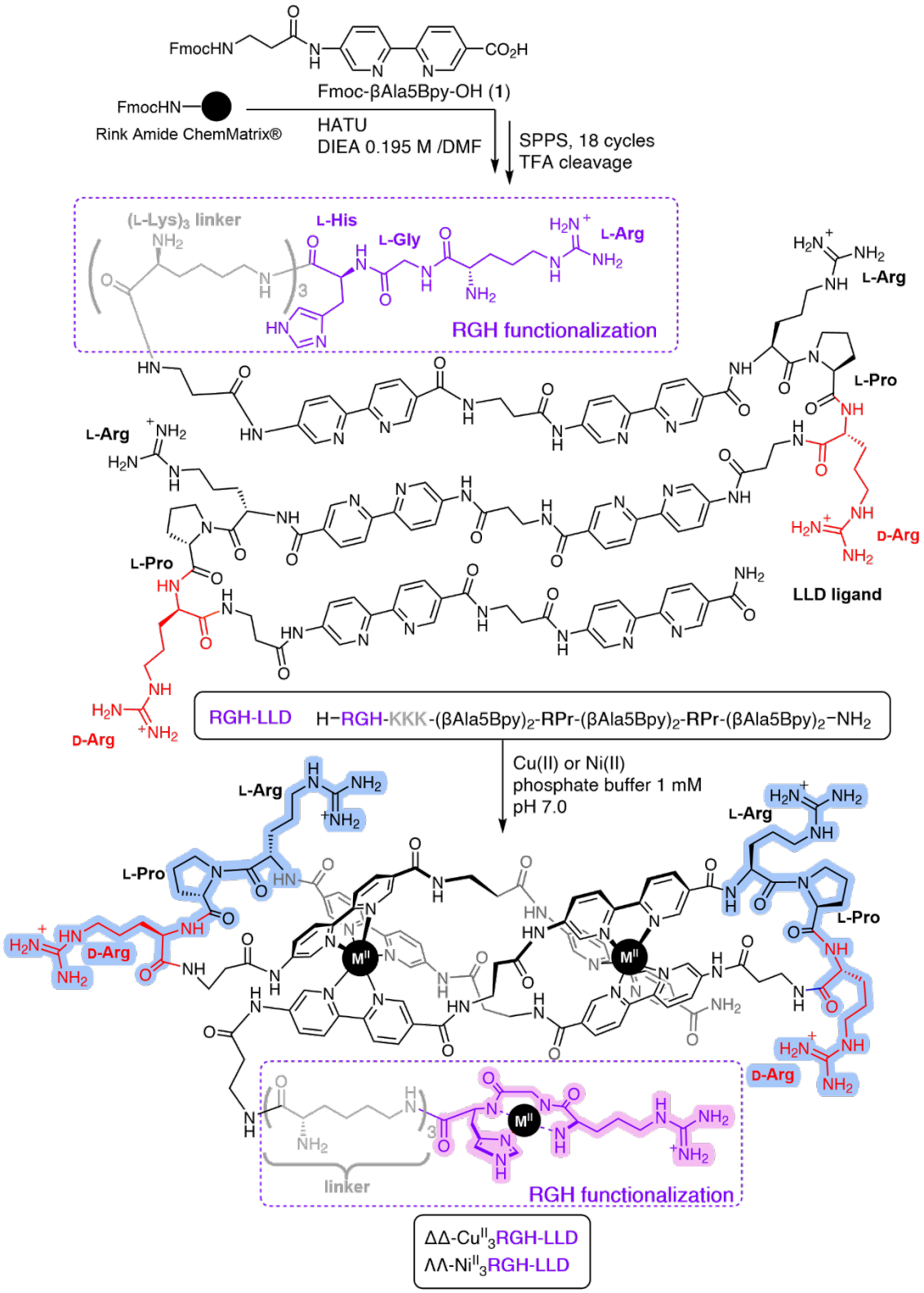
Design, synthesis and characterization of the Cu(II) and Ni(II) ATCUN-derived peptide helicate chimeras

The newly designed peptide ligand **RGH-LLD** features a tripeptide ATCUN motif at the N-terminus of the peptide ligand **LLD**, described in Chapter I. Aiming to improve the binding affinity of the corresponding helicates to their target DNA, we decided to use as ATCUN motif the tripeptide **L-Arg–Gly–L-His** (RGH) since it has been described that a basic residue (such as L-Arg) at the N-terminal position enhances the DNA binding properties of the ATCUN-derived metallopeptides due to the electrostatic interactions between the cationic Arg sidechain and the negatively charged phosphodiester linkages of the dsDNA.[288] The ATCUN complex is separated from the **LLD** peptide ligand[190c] by a peptide spacer composed of three laterally connected lysines, which should allow the ATCUN metallopeptide to freely bind to the minor groove of the 3WJ while the **M(II)LLD** peptide helicate is simultaneously inserted into the central hydrophobic cavity of the junction.

The peptide ligand **RGH-LLD** was synthesized using standard Fmoc solid-phase peptide synthesis protocols (Scheme 4), as described for the **LLD** derivative in Chapter I.[190c] The peptide ligand was simultaneously deprotected and cleaved from the solid support by treatment with an acidic TFA cocktail, and after RP-HPLC purification, its identity was confirmed by mass spectrometry (Figure 86).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

[288] R. Nagane, T. Koshigoe, M. Chikira, E. C. Long, *J. Inorg. Biochem.* **2001**, *83*, 17–23.



Scheme 4. Solid-phase peptide synthesis of the peptide ligand **RGH-LLD**, equipped with six βAla5Bpy coordinating residues, a trilysine spacer and a N-terminal ATCUN motif. The two heterochiral β-turn sequences (RPr) are highlighted in blue, the Lys_3 spacer is colored in grey and the ATCUN motif is highlighted in pink. After cleavage and purification, the peptide ligand stereoselectively self-assembly into the corresponding peptide helicates in the presence of Cu(II) and Ni(II) ions, which also coordinate with high affinity the N-terminal ATCUN tripeptide, generating in this case a metallopeptide unit with square planar coordination geometry. L-amino acids are indicated in upper-case while D-amino acids are in lower case (i.e., R, P refer to L-Arg and L-Pro, and r to D-Arg).

We monitored the formation of the Cu(II) and Ni(II) ATCUN-derived peptide helicate chimeras (Scheme 4) by measuring the quenching in the emission band of the Bpy units in the peptide ligand **RGH-LLD**.[251] We noticed that the coordination kinetics was very slow, and we had to wait a few minutes after the addition of the corresponding aliquot of metal salt solution for equilibration. Thus, to calculate the dissociation constants of the metallopeptide chimeras, we decided to perform the corresponding spectrofluorimetric titrations by incubating overnight independent solutions with a fixed concentration of the **RGH-LLD** ligand (2 μM in 1mM phosphate buffer, 10 mM NaCl, pH 7.0) with different amounts of the metal salts ($CuCl_2 \cdot 2\ H_2O$ or $NiCl_2$) (Figure 48).

The design of this titration only accounts for the metal coordination to the **LLD** fragment of the **RGH-LLD** peptide ligand, and does not monitor the coordination to the ATCUN motif, which is known to have an extremely high metal-binding affinity with dissociation constants in the order of picomolar or even lower.[278,289] Assuming that the first equivalent of M(II) with respect to the peptide ligand (i.e., 2 μM of M(II)) binds to the ATCUN motif prior to the **LLD** fragment, the resulting profile at 420 nm, starting from 2 μM of M(II), was fitted to a 1:1 binding model for Cu(II) ions and to a 1:2 for Ni(II). The apparent dissociation constant for **Cu^{II}_3RGH-LLD** is calculated to be $K_D \approx$ 0.1 μM for two Cu(II) ions coordinated to the **LLD** fragment, whereas for **Ni^{II}_3RGH-LLD**, the calculated dissociation constants are K_{D1} = 0.00208 μM and K_{D2} = 0.333 μM, for the first and second Ni(II) ions.[252] These results are consistent with the values reported for similar systems.[190c] Therefore, these studies suggest that both systems show global 1:3 (**RGH-LLD**:Cu^{II}/Ni^{II}) binding modes. MALDI spectra of peptide ligand solutions at saturation concentrations of Cu(II) or Ni(II) ions showed a peak at m/z = 3363.8521 and m/z = 3356.1284 consistent with the formation of **$[Cu^{II}_3$RGH-LLD$]^{6+}$** and **$[Ni^{II}_3$RGH-LLD$]^{6+}$** respectively (Figure 89 and 90), thus confirming our hypothesis.

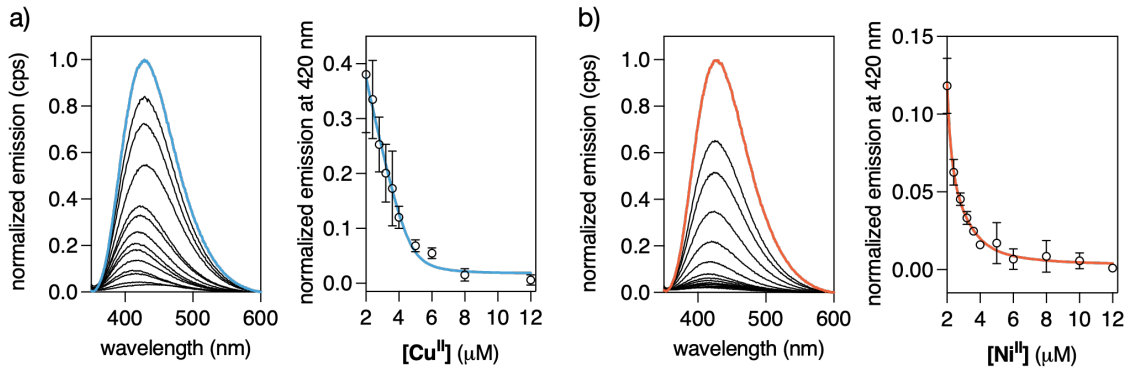


Figure 48. a) Spectrofluorimetric titration of the peptide ligand **RGH-LLD** with Cu(II) ions; left: fluorescence emission spectrum of a 2 μM solution of **RGH-LLD** (blue), and emission spectra of a series of 2 μM solutions of **RGH-LLD** independently incubated overnight in the presence of increasing amounts of Cu(II) ions (black lines); right: titration profile at 420 nm of three independent experiments, starting from [Cu(II)] at 2 μM, with the best fit (blue line) according to a 1:1 (L:M_2) binding model calculated by *DynaFit*.[252] b) Spectrofluorimetric titration of the peptide ligand **RGH-LLD** with Ni(II) ions; left: fluorescence emission spectrum of a 2 μM solution of **RGH-LLD** (orange), and emission spectra of a series of 2 μM solutions of **RGH-LLD** independently incubated overnight in the presence of increasing amounts of Ni(II) ions (black lines); right: titration profile at 420 nm of three independent experiments, starting from a [Ni(II)] at 2 μM, with the best fit (orange line) according to a 1:2 (L:M) binding model calculated by *DynaFit*. Conditions: 1 mM phosphate buffer, NaCl 10 mM, pH 7.0.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

[289] K. Bossak-Ahmad, T. Frączyk, W. Bal, S. C. Drew, *Chembiochem* **2020**, *21*, 331–334.

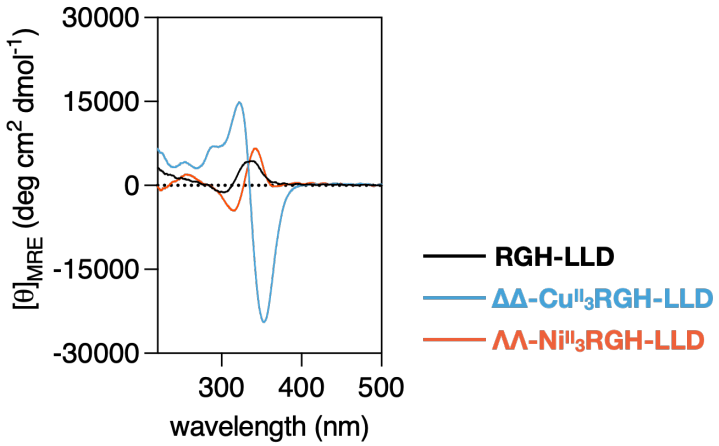


Figure 49. CD spectra of **RGH-LLD** peptide ligand at 5 μM (black line) in 1 mM phosphate buffer, 10 mM NaCl, pH 7.0, and in the presence of 7.5 eq. of Cu(II) (blue line) or Ni(II) (orange line) ions.

Once demonstrated that the peptide ligand **RGH-LLD** assembled in the corresponding peptide helicates in the presence of Cu(II) and Ni(II) ions, we studied their chirality by measuring the CD spectra of a 5 μM solution of the free peptide ligand **RGH-LLD** and in the presence of 7.5 eq. of Cu(II) or Ni(II) ions. The CD spectrum of **RGH-LLD** is dominated by CD band with a positive Cotton effect centered at 345 nm, confirming a significant preorganization of the peptide strand and the effective chiral induction of the heterochiral β-turn loops, as in the case of the **LLD** ligand.[190c] Consistent with the results discussed in Chapter I for **LLD** (Figure 32), the assembly of the Cu(II) metallopeptide resulted in an inversion of the Cotton effect of this CD band, together with a large increase in its intensity, which points to the formation of a ΔΔ- (or *P*) helicate. On the contrary, the addition of Ni(II) ions induces CD band with a positive Cotton effect at this wavelength that suggest the preferential formation of the diasteromeric (*M*) helicate (Figure 49). Such inverted chirality observed for the Cu(II) peptide helicate chimera with respect to the Ni(II) derivative seems to be promoted by a Jahn-Teller distortion, typical of octahedral coordination geometries in d^9 systems, such as Cu(II).[253]

In vitro DNA-binding studies

The 3WJ binding properties of the Cu(II) and Ni(II) peptide helicates chimeras derived from the peptide ligand **RGH-LLD** were studied by measuring the quenching in the emission band of a 2 μM solution of fluorescein-labelled 3WJ DNA (Flu-3WJ), induced by the addition of increasing amounts of the pre-formed **ΔΔ-Cu$^{II}_3$RGH-LLD** or **ΛΛ-Ni$^{II}_3$RGH-LLD** (λ_{exc} = 490 nm). The resulting titration profiles at 515 nm could be fitted to a mixed model including a 1:1 binding mode (Flu-3WJ:helicate) together with non-specific interactions using the *Dynafit* software, resulting in an apparent dissociation constant K_D of 170 nM for **ΔΔ-Cu$^{II}_3$RGH-LLD** and of 140 nM for **ΛΛ-Ni$^{II}_3$RGH-LLD** (Figure 50).[252] These values are lower to those calculated for the Fe(II) peptide helicate derived from **LLD**,[190c] as well as to those discussed in Chapter I for **ΔΔ-Cu$^{II}_2$LLD** (Figure 33), indicating that the ATCUN metallopeptide appended to the N-terminal end of the **LLD** peptide helicate increases the 3WJ binding affinity (Figure 50).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

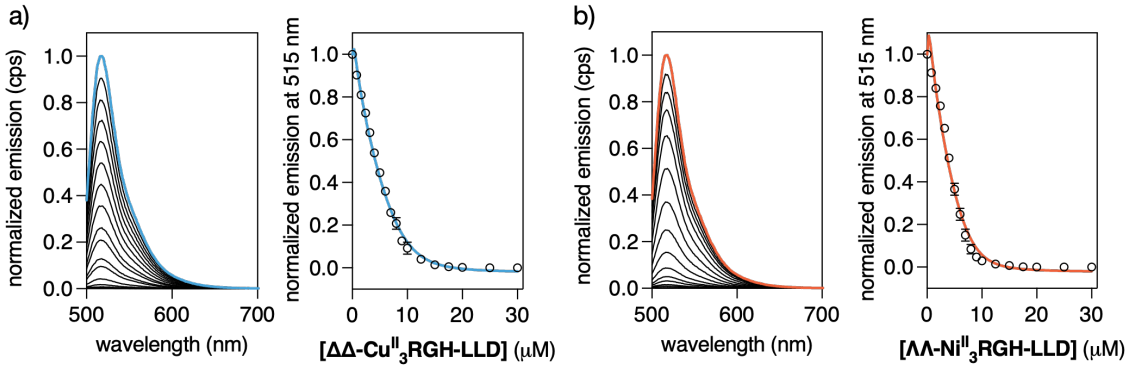


Figure 50. a) Left: normalized emission spectra of a 2 μM solution of the fluorescein-labelled 3WJ (Flu-3WJ) in absence (blue line) and in the presence of increasing concentrations of **ΔΔ-Cu$^{II}_3$RGH-LLD** (black lines of decreasing intensity); right: titration profile at 515 nm with the best fit (blue line) to a 1:1 binding model plus non-specific interactions calculated with *Dynafit*;[252] b) left: normalized emission spectra of a 2 μM solution of the fluorescein-labelled 3WJ (Flu-3WJ) in absence (orange line) and in the presence of increasing concentrations of **ΛΛ-Ni$^{II}_3$RGH-LLD** (black lines of decreasing intensity); right: titration profile at 515 nm with the best fit (orange line) to a 1:1 binding mode plus non-specific interactions calculated with *Dynafit.* Titration data are the average of three independent experiments. Conditions: phosphate buffer 1mM, NaCl 10 mM, pH 7.0; λ_{exc} = 490 nm; Flu-3WJ oligonucleotide sequences: *5'–FAM–TTTT CAC CGC TCT GGT CCT C–3'; 5'–CAG GCT GTG AGC GGT G–3'; 5'–GAG GAC CAA CAG CCT G–3'*.

Following the same approach described in Chapter I, we also calculated the K_D for the binding of **ΔΔ-Cu$^{II}_3$RGH-LLD** and **ΛΛ-Ni$^{II}_3$RGH-LLD** to other 3WJ oligonucleotides using the same binding model (Flu-GC-3WJ) (Figure 92). The resulting dissociation constants are in the same range (K_D = 260 nM and K_D = 380 nM for Cu(II) and Ni(II) peptide helicate chimeras, respectively). The binding of **ΔΔ-Cu$^{II}_3$RGH-LLD** to an AT-rich 3WJ (Flu-TA-3WJ) fits to a simple 1:1 binding model with $K_D \approx 230$ nM, (Figure 93). These data suggest that the Cu(II) and Ni(II) ATCUN-functionalized metallopeptides recognize 3WJ structures in a sequence-independent manner. Finally, spectrofluorimetric titration of **ΔΔ-Cu$^{II}_3$RGH-LLD** and **ΛΛ-Ni$^{II}_3$RGH-LLD** with canonical DNA resulted in the formation of a precipitate during the experiments, which is indicative of a non-specific interaction between the peptide helicate chimeras and dsDNA (Figure 91) and of their high selectivity for 3WJs *in vitro*.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

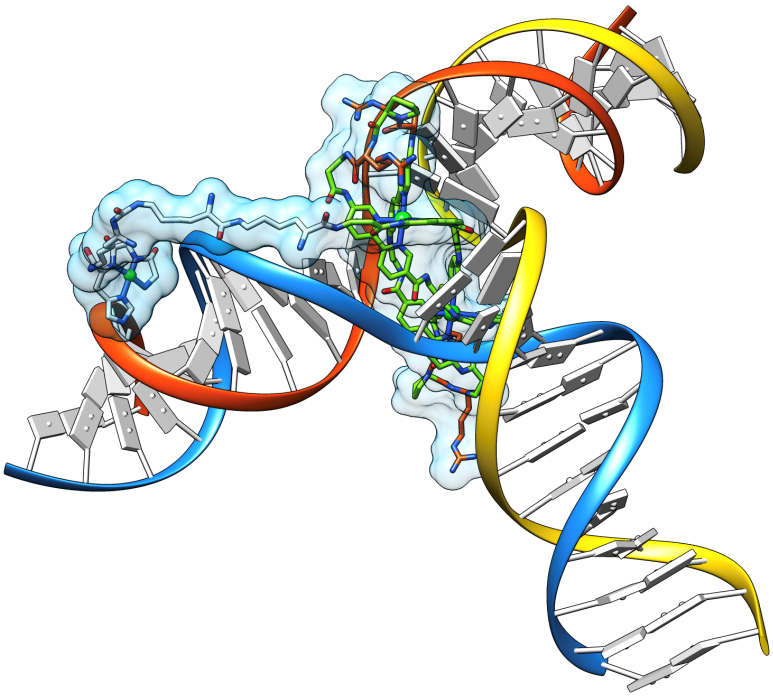


Figure 51. Illustration of **ΔΔ-Cu$^{II}_3$RGH-LLD** recognizing a 3WJ structure. The **LLD** peptide helicate unit is inserted in the branch point of the 3WJ, whereas the **RGH** metallopeptide unit is inserted in the minor groove of one of their double-stranded DNA arms. The image was made using *Chimera* software.

In vitro 3WJ nuclease studies

We analyzed the capability of **ΔΔ-Cu$^{II}_3$RGH-LLD** and **ΛΛ-Ni$^{II}_3$RGH-LLD** to selectively cleave 3WJs *in vitro* using electrophoretic mobility shift assays (EMSA) in polyacrylamide gel (PAGE) under non-denaturing conditions[255] and *SYBRGold* for DNA staining.[256] The target 3WJ and *off*-target dsDNA oligonucleotides were incubated with increasing concentrations of the pre-formed **ΔΔ-Cu$^{II}_3$RGH-LLD** or **ΛΛ-Ni$^{II}_3$RGH-LLD** in the presence of dihydrogen peroxide (H_2O_2) at room temperature for 1 hour. Control lines without H_2O_2 and the peptide helicate chimeras were also added. It has to be highlighted that the 3WJ oligonucleotide need the presence of the Cu(II) or Ni(II) peptide helicate chimera in order to be conveniently assembled (lines 1 and 4 in Figures 52a and 52b), but this is not the case for canonical dsDNA (line 2 in Figures 52a and 52b).

We found that **Cu$^{II}_3$RGH-LLD** selectively binds 3WJ in the presence of canonical dsDNA as suggested by the disappearance of the lower band associated with the unassembled 3WJ oligonucleotide (Figure 52a, line 1) together with the formation of the new slower-migrating band, with no smearing, ascribed to the formation of the 3WJ/**ΔΔ-Cu$^{II}_3$RGH-LLD** adduct (Figure 52a, line 4). Moreover, the data obtained suggest that **ΔΔ-Cu$^{II}_3$RGH-LLD** selective cleaves 3WJ in the presence of canonical DNA and H_2O_2, as suggested by the gradually decrease of the intensity of the band ascribed to the formation of the 3WJ/**ΔΔ-Cu$^{II}_3$RGH-LLD** adduct (Figure 52a, lines 5-8) proportionally to the concentration of metallopeptide added. Very interestingly, the canonical DNA in the sample remains undamaged even at high **ΔΔ-Cu$^{II}_3$RGH-LLD** concentration, thus confirming the poor affinity of the Cu(II) peptide helicate chimera for B-DNA (Figure 52a, line 8).

USC DE COMPOSTELA

PAGE studies with **ΛΛ-Ni$^{II}_3$RGH-LLD** (Figure 52b) indicate a different DNA cleavage behaviour. The incubation of 3WJ and dsDNA with increasing equivalents of the Ni(II) peptide helicate chimera leads to the appearance of a new slower-migrating band, with no smearing, associated to the formation of the 3WJ/**ΛΛ-Ni$^{II}_3$RGH-LLD** adduct. Interestingly, the intensity of this band increases with the concentration of the Ni(II) metallopeptide, contrary to what has been observed in the case of **ΔΔ-Cu$^{II}_3$RGH-LLD**. This indicates that **ΛΛ-Ni$^{II}_3$RGH-LLD** shows a great affinity for 3WJ, in accordance with the spectrofluorimetric data, but does not exhibit nuclease activity for 3WJ. Moreover, the band associated with canonical DNA remains unaffected, thus confirming the poor affinity of the Ni(II) peptide helicate chimera for B-DNA.

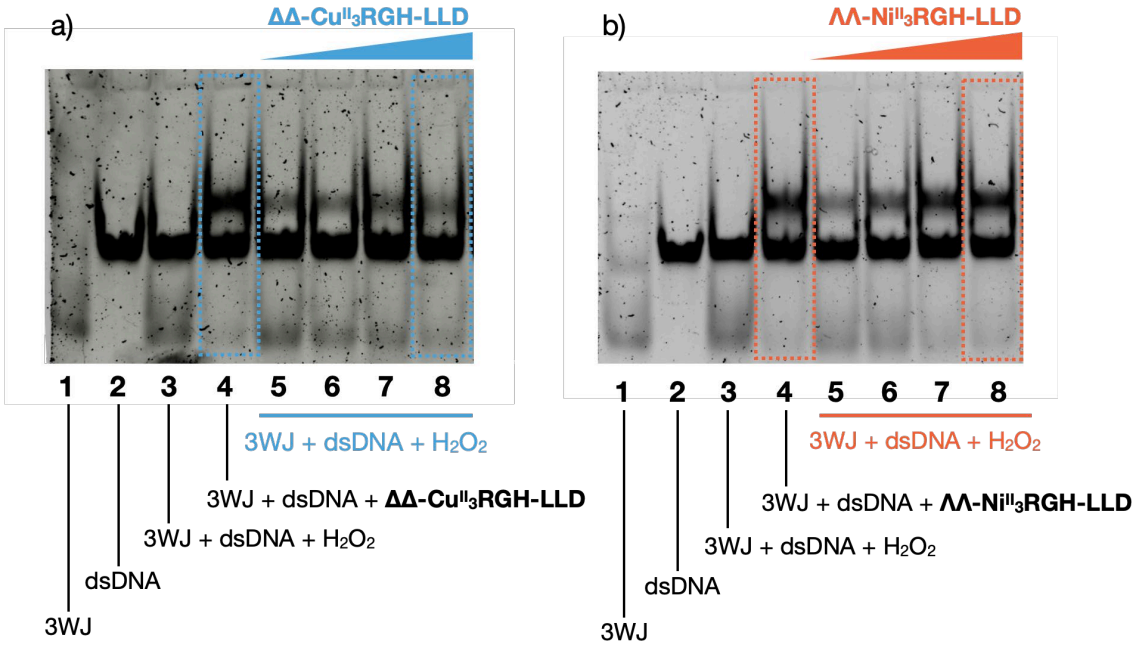


Figure 52. a) PAGE analysis of the DNA-binding and cleavage properties of **ΔΔ-Cu$^{II}_3$RGH-LLD**. Lane 1: 200 nM of 3WJ. Lane 2: 200 nM of dsDNA. Lane 3: 200 nM of 3WJ and dsDNA in the presence of 250 μM of H_2O_2. Lane 4: 200 nM of 3WJ and dsDNA in the presence of 600 nM of **ΔΔ-Cu$^{II}_3$RGH-LLD**. Lanes 5-10: 200 nM of 3WJ and dsDNA in the presence of 100, 200, 400, 600, 800 and 1000 nM of **ΔΔ-Cu$^{II}_3$RGH-LLD** and 250 μM of H_2O_2, respectively. To get an idea of the nuclease potential of this metallopeptide, please compare lanes 4 and 8. b) PAGE analysis of the DNA-binding and cleavage properties of **ΛΛ-Ni$^{II}_3$RGH-LLD**. Lane 1: 200 nM of 3WJ. Lane 2: 200 nM of dsDNA. Lane 3: 200 nM of 3WJ and dsDNA in the presence of 250 μM of H_2O_2. Lane 4: 200 nM of 3WJ and dsDNA in the presence of 600 nM of **ΛΛ-Ni$^{II}_3$RGH-LLD**. Lanes 5-10: 200 nM of 3WJ and dsDNA in the presence of 100, 200, 400, 600, 800 and 1000 nM of **ΛΛ-Ni$^{II}_3$RGH-LLD** and 250 μM of H_2O_2, respectively. To get an idea of the nuclease potential of this metallopeptide, please compare lanes 4 and 8. 3WJ oligonucleotide sequences: *5'–CAC CGC TCT GGT CCT C–3'; 5'–CAG GCT GTG AGC GGT G–3'; 5'–GAG GAC CAA CAG CCT G–3'*. dsDNA oligonucleotide sequence: *5'–AAC ACA TGC AGG ACG GCG CTT–3'* (only one strand shown).

Assuming that the nuclease capacity of these chimeras should be the sum of that shown by the **LLD** peptide helicate fragment and that of the **RGH** metallopeptide, we can draw some conclusions from the behavior offered by the Cu(II) and Ni(II) peptide helicates chimeras. The studies described in Chapter I demonstrate that **ΔΔ-Cu$^{II}_2$LLD** exhibits efficient and selective nuclease activity for 3WJ. On the other hand we have indicated in the introduction to this chapter that the Cu(II)- and Ni(II) metallopeptides derived from the ATCUN motif show notable nuclease activity for canonical DNA.

However, while **ΔΔ-Cu$^{II}_3$RGH-LLD** shows remarkable nuclease activity by 3WJ, **ΛΛ-Ni$^{II}_3$RGH-LLD** is almost inactive. This behaviour suggests that the Ni(II)-**LLD** fragment of the chimera does not exhibit nuclease activity on its own. The data obtained seem to indicate that the ATCUN fragments are not important in the nuclease activity of the chimeras, although they do play a significant role in their binding properties to 3WJ.

It is known that the nuclease activity of Cu(II) and Ni(II) metallopeptides derived from the ATCUN motif is mediated by ROS formation via $M^{3+/2+}$ redox couples after dsDNA minor groove insertion. In this context, the apparent lack of nuclease activity displayed by the ATCUN fragments of **ΔΔ-Cu$^{II}_3$RGH-LLD** and **ΛΛ-Ni$^{II}_3$RGH-LLD** may to be associated with issues related with a DNA-binding orientational effect of these fragments rather with problems in the production of ROS species. Moreover, whereas the structure of Cu in d^8 and d^9 configurations is assumed to be square planar in both redox states, Ni features different coordination geometries in its d^7 and d^8 configurations, which, in this particular case, suggests a metal-associated ROS position relative to the DNA bonds less optimal in mediating DNA cleavage chemistry, compared to the ROS position with the Cu(II) helicate.[283] However, more studies are needed to draw accurate conclusions on this issue.

DNA-binding studies in mammalian cells

Following the same approach described in Chapter I for the study of the intracellular localization of the **ΔΔ-Cu$^{II}_2$TAMRA-LLD** helicate (Figure 37), **ΛΛ-Ni$^{II}_3$TAMRA-RGH-LLD** readily translocated in digitonin-treated HeLa cells, showing a diffuse distribution in the cytosol, and a strong punctuated pattern in the nucleus, as well as an evident concentration in the nucleolus (Figure 53). Most of the nuclear spots co-localized with GFP-tagged Proliferating Cellular Nuclear Antigen (PCNA), which specifically labels DNA replication foci in actively dividing cells.[259]

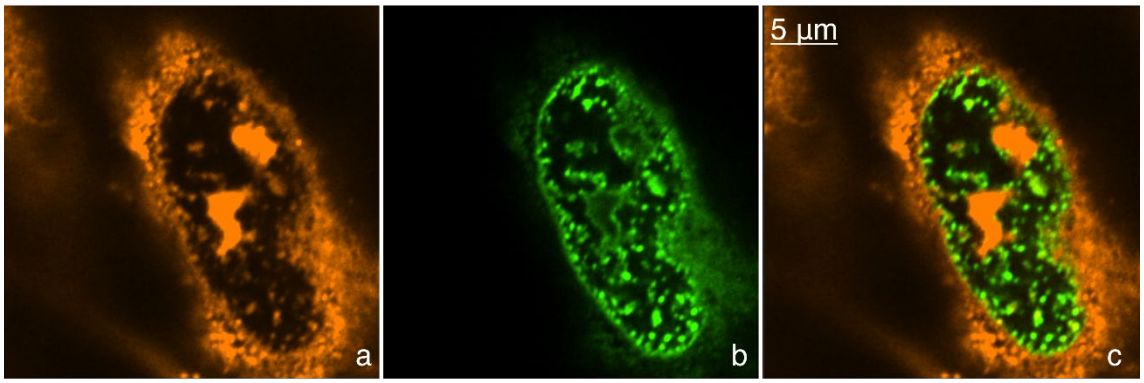


Figure 53. ΛΛ-Ni$^{II}_3$TAMRA-RGH-LLD selectively stains DNA replication sites. HeLa cells expressing protein GFP-PCNAL2 were incubated in the presence of 25 µg/mL of digitonin for 3min, and then in the presence of 5 µM of **ΛΛ-Ni$^{II}_3$TAMRA-RGH-LLD** for 15 min. The Ni(II) peptide helicate was pre-formed by adding 7.5 equivalents of $NiCl_2$ to an aqueous solution of **TAMRA-RGH-LLD** peptide ligand that was overnight incubated. A) Red channel emission showing the distribution of the TAMRA-labelled helicate; b) green channel, corresponding to the emission of the GFP-PCNAL2 probe that labels DNA replication foci; c) overlay of the green and red channels. Nucleoli are clearly labelled as big red spots inside the cell nuclei. The image shows a representative confocal section of the HeLa cells.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

While the cytosolic distribution and association with the nucleoli might reflect some *off*-target RNA-binding of the helicate, replisomes labelled by PCNA are enriched

in 3WJs due to the presence of replication forks. Therefore, our results strongly suggest the direct association of this Ni(II) peptide helicate chimera with 3WJs structures in functional cell nuclei. Furthermore, control experiments showed that the **TAMRA-RGH-LLD** peptide ligand in absence of metal ions does not localize in the DNA replication centers (Figure 94). This control experiment suggests that **ΛΛ-Ni^{II}_3TAMRA-RGH-LLD** is stable in cell medium during the assay, and that this new system is only able to recognize selectively DNA replication centres in the nucleus of mammalian cells as a discrete metallopeptide entity.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Conclusions

In this chapter we described the synthesis of two oligocationic Cu(II) and Ni(II) peptide helicates chimeras by SPPS methods. These metallopeptides are derived from the new designed peptide ligand **RGH-LLD**, which is composed by the **LLD** peptide ligand fragment,[190c] described in Chapter I, engineered at its amino terminus with an ATCUN motif. CD spectra of the metallopeptides indicate that the peptide ligand is highly organised prior to the metal ion coordination, and also that both systems display inverted chirality, probably due to the Jahn-Teller distortion experimented by the Cu(II) derivative. The dissociation constant calculated for the metal binding using spectrofluorimetric data are very similar in both systems and slightly better to those observed for the peptide helicates derived from **LLD**, indicating the positive influence of the attached ATCUN fragment.

Spectrofluorimetric and PAGE studies indicate that both **ΔΔ-Cu^{II}_3RGH-LLD** and **ΛΛ-Ni^{II}_3RGH-LLD** display high affinity and selectivity for 3WJ *in vitro*. Additionally, the Cu(II) derivative can efficiently and selectively cleave 3WJs in the presence of canonical DNA by ROS production, showing great potential as antitumor agent targeting non-canonical DNA structures. On the contrary, the Ni(II) peptide helicate chimera does not show such behavior. The difference in the nuclease activity between **ΔΔ-Cu^{II}_3RGH-LLD** and **ΛΛ-Ni^{II}_3RGH-LLD,** confirms the great therapeutic potential of this class of peptide metallodrugs, which is related to the functional versatility provided by the different metal ions that could be used in their synthesis.

Confocal microscopy studies indicate that the Ni(II) chimera is able to internalize into digitonin-treated HeLa cells and was localized in the nucleus. Its nuclear distribution matched the localization of DNA replication foci, demonstrating that these systems also bind with high affinity and selectivity to 3WJ structures *in cellulo*.

We are currently working on the quantification of their nuclease activity *in vitro*, their capability to carry it out *in cellulo* as well as the mechanism of action of their nuclease activity.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter IV

Selective recognition of A/T-rich DNA 3-way junctions with a three-fold symmetric tripeptide

Summary

3WJs have recently emerged as potential chemotherapeutic targets. Here, we present a new strategy to target 3WJs that relies on the cooperative and sequence-selective recognition of the three A/T-rich duplex DNA arms in the 3WJ structure by three AT-Hook peptide strands attached to a C_3-symmetrical fluorogenic core of 1,3,5-tristyrylbenzene.

Reproduced from Ref. J. Gómez-González, L. Martínez-Castro, J. Tolosa-Barrilero, A. Alcalde-Ordóñez, S. Learte-Aymamí, J. L. Mascareñas, J. C. García-Martínez, J. Martínez-Costas, J.-D. Maréchal, M. Vázquez López, M. E. Vázquez, *Chem. Commun.* **2022**, *58*, 7769-7772 with permission from the Royal Society of Chemistry.

Publisher: Royal Society of Chemistry

ISSN: 1364-548X

J. Gómez-González,[a] L. Martínez-Castro,[b] J. Tolosa-Barrilero,[c,d] A. Alcalde-Ordóñez,[a] S. Learte-Aymamí,[a] J. L. Mascareñas,[a] J. C. García-Martínez,[c,d] J. Martínez-Costas,[e] J.-D. Maréchal,[b] M. Vázquez López,[f] M. E. Vázquez[a]

[a] Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Departamento de Química Orgánica, Universidade de Santiago de Compostela, Spain.

[b] Insilichem, Departament de Química, Universitat Autònoma de Barcelona, 08193 Cerdanyola, Spain.

[c] Department of Inorganic, Organic Chemistry and Biochemistry, Faculty of Pharmacy, University of Castilla-La Mancha, 02071 Albacete, Spain.

[d] Regional Center for Biomedical Research (CRIB), 02071 Albacete, Spain.

[e] Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Departamento de Bioquímica y Biología Molecular, Universidade de Santiago de Compostela, Spain.

[f] Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Departamento de Química Inorgánica, Universidade de Santiago de Compostela, Spain.

Introduction

On August 25, 1942, Dr Gustaf Lindskog administered to a patient with non-Hodgkin´s lymphoma the first intravenous chemotherapy for cancer based on a DNA alkylating agent.[290] More than 80 years later, these cytotoxic agents that aim at the DNA, including cross-linking agents, remain essential weapons in the chemotherapeutic arsenal.[291] However, beyond the canonical antiparallel double-stranded B-DNA conformation, non-canonical DNA structures have emerged as promising alternative DNA targets for the development of new drugs with enhanced therapeutic properties.[38c,233,292] Among them, 3WJs, the simplest branched DNA structures transiently formed during DNA replication[22,24,25,31,35,36,37,234b] have already displayed great antitumour potential, since it has been reported that stabilization of such secondary structures through their ligands binding results in DNA damage, and consequent cell death[233c] probably driven by replication fork stalling and genotoxic replication stress.[234]

Since the group of Prof. M. Hannon discovered that metal helicates fit into the branch point of 3WJs,[184] researchers have described a number of molecules that exploit symmetry, hydrophobicity, shape and charge complementarity to target such central cavity in 3WJs.[293] In general, a *threefold symmetry* together with an aromatic π-surface improve the binding affinity towards this hydrophobic core by promoting π-stacking interactions with the nucleobases at the junction interface. Examples include peptide helicates,[190] azacryptands,[267] triptycenes,[264] cationic calix[3]carbazoles[294] and derivatives with appended peptides for additional nonspecific interactions with the duplex branches.[229,230] However, the branch point targeted by these recognition agents is basically a featureless hydrophobic hole that does not present any obvious structural elements to allow the sequence selective recognition of 3WJs. To gain sequence selectivity, an attractive strategy would be the design of new binders consisting of binding-agents to the 3WJ central cavity combined with sequence-selective DNA recognition motifs in natural proteins, such as the *AT-Hook* motif to target adjacent sequences in the dsDNA branches.

The AT-Hook is a nine-residue oligocationic peptide motif (RKP**RGR**PKK) found in eukaryotic HMG-I(Y) nuclear proteins.[295] Monomeric AT-Hook peptides display low affinity, in the order of millimolar, for A/T-rich DNA sequences,[296] but appropriately spaced AT-Hook units cooperatively bind to composite A/T rich sites with high affinity in the low micromolar to nanomolar range.[296b,297] Structural NMR and crystallography studies show that the central **RGR** segment of the AT-Hook, sits at the

[290] K. W. Kohn, *Cancer Res.* **1996**, *56*, 5533–5546.
[291] J. Sun, Q. Wei, Y. Zhou, J. Wang, Q. Liu, H. Xu, *BMC Syst. Biol.* **2017**, *11*, 87.
[292] S. M. Nelson, L. R. Ferguson, W. A. Denny, *Cell Chromosome* **2004**, *3*, 2.
[293] S. Vuong, L. Stefan, P. Lejault, Y. Rousselin, F. Denat, D. Monchaud, *Biochimie* **2012**, *94*, 442–450.
[294] Z. Yang, Y. Chen, G. Li, Z. Tian, L. Zhao, X. Wu, Q. Ma, M. Liu, P. Yang, *Chem. Eur. J.* **2018**, *24*, 6087–6093.
[295] L. Aravind, D. Landsman, *Nucleic Acids Res.* **1998**, *26*, 4413–4421.
[296] a) R. Reeves, M. S. Nissen, *J. Biol. Chem.* **1990**, *265*, 8573–8582; b) B. H. Geierstanger, B. F. Volkman, W. Kremer, D. E. Wemmer, *Biochemistry* **1994**, *33*, 5347–5355.
[297] a) J. F. Maher, D. Nathans, *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 6716–6720; b) J. P. Wagner, D. M. Quill, D. E. Pettijohn, *J. Biol. Chem.* **1995**, *270*, 7394–7398.

bottom of the DNA minor groove, being deeply inserted in an extended concave conformation, making extensive polar and hydrophobic contacts, while the basic residues flanking this core sequence participate in additional electrostatic contacts with the sugar-phosphate backbone.[298] Researchers have exploited the versatility conferred by the peptide nature and the simplicity of the AT-Hook module for the design of artificial duplex DNA binding agents.[299]

[298] a) J. R. Huth, C. A. Bewley, M. S. Nissen, J. N. S. Evans, R. Reeves, A. M. Gronenborn, G. M. Clore, *Nat. Struct. Biol.* **1997**, *4*, 657–665; b) E. Fonfría-Subirós, F. Acosta-Reyes, N. Saperas, J. Pous, J. A. Subirana, J. L. Campos, *PLoS One* **2012**, *7*, e37120.

[299] a) S. Learte-Aymamí, J. Rodríguez, M. E. Vázquez, J. L. Mascareñas, *Chem. Eur. J.* **2020**, *26*, 8875–8878; b) J. Rodríguez, J. Mosquera, R. García-Fandiño, M. E. Vázquez, J. L. Mascareñas, *Chem. Sci.* **2016**, *7*, 3298–3303.

Objetive

In this chapter we will describe how we have planned to reproduce the mechanism used by proteins that recognize 3WJs in biological systems, which interact with duplex arms by exploiting the triple symmetry of 3WJs.[24] In particular, we will describe the modification of a symmetrical 1,3,5-trisitylbenzene triple core, which is able to bind to the hydrophobic cavity of 3WJs, with three AT-Hook peptide units, which are able to selectively insert into A/T-rich sequences of the minor groove of the B-DNA arms of 3WJs. Taking advantage of the inherent fluorogenic features of the 1,3,5-tristyrylbenzene unit, we will study the DNA binding properties of the newly designed trimeric ligand, which selectively recognize A/T-rich 3WJs over canonical DNA and other 3WJ structures.

Results and discussion

Design and synthesis of the trimeric AT-Hook 3WJ binder

Taking as starting point the crystal structures of the AT-Hook peptide bound to a short A/T-rich oligonucleotide (PDB: 3UXW) and the trimeric Cre recombinase bound to a three-way Lox DNA junction (PDB: 1F44), we built a hypothetical model of three AT-Hook peptides simultaneously bound to the three DNA minor grooves adjacent to the branch point in the 3WJ. This rough model suggested that the N-terminus of the bound AT-Hook peptides would be about 4 Å from the carboxylates in a 1,3,5-tristyrylbenzene core at the center of the 3WJ, so that they could be connected through the side chain of the N-terminal Lysine residue (Figure 54).

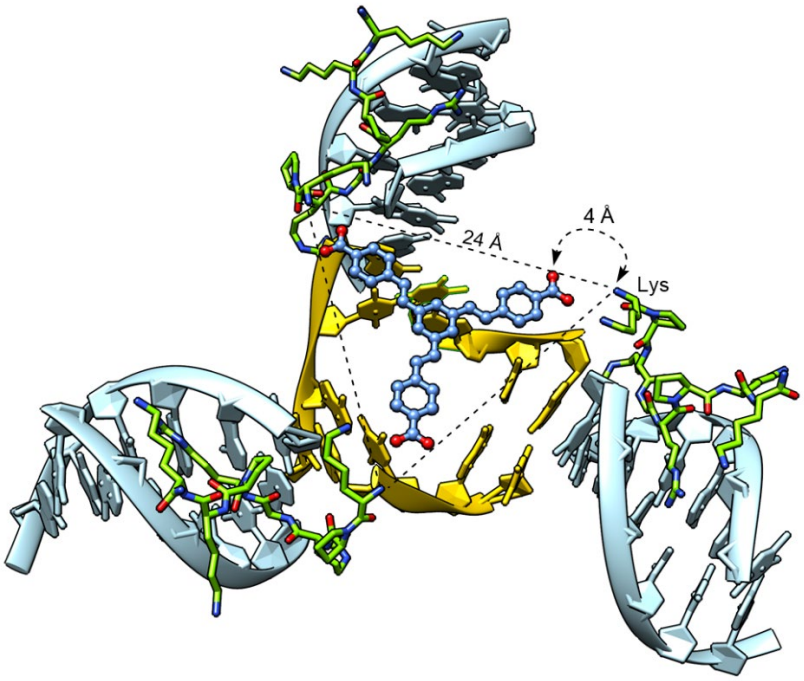


Figure 54. Hypothetical model of three AT-Hook peptides simultaneously inserted into the three duplex DNA minor grooves adjacent to the branching point in the 3WJ, with their N-terminus at 4 Å from the carboxylate groups in a 1,3,5-tristyrylbenzene core hovering over this central cavity.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

This hypothesis was further investigated by molecular modelling with the multi-objective genetic algorithm platform *GaudiMM*,[300] which allowed us to make an extensive conformational exploration while simultaneously evaluating several fitness functions (See 5 of Chapter IV in Experimental Section for computational details). The absolute best solution shows an excellent distance match between the 1,3,5-tristyrylbenzene and the AT-Hook, as well as great fitting with the 3WJ, with the central phenyl ring hovering over the large cavity of the 3-way junction branch point (Figure 55).

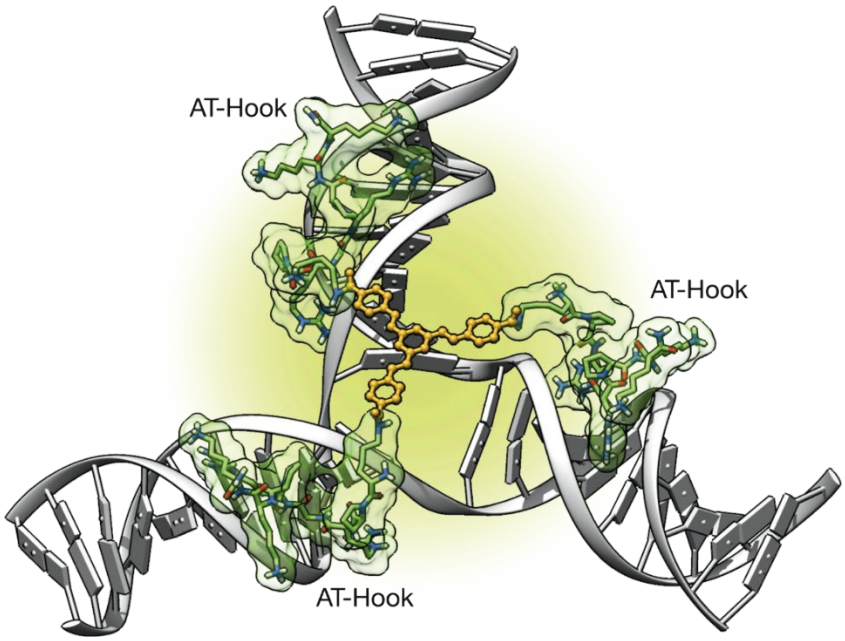
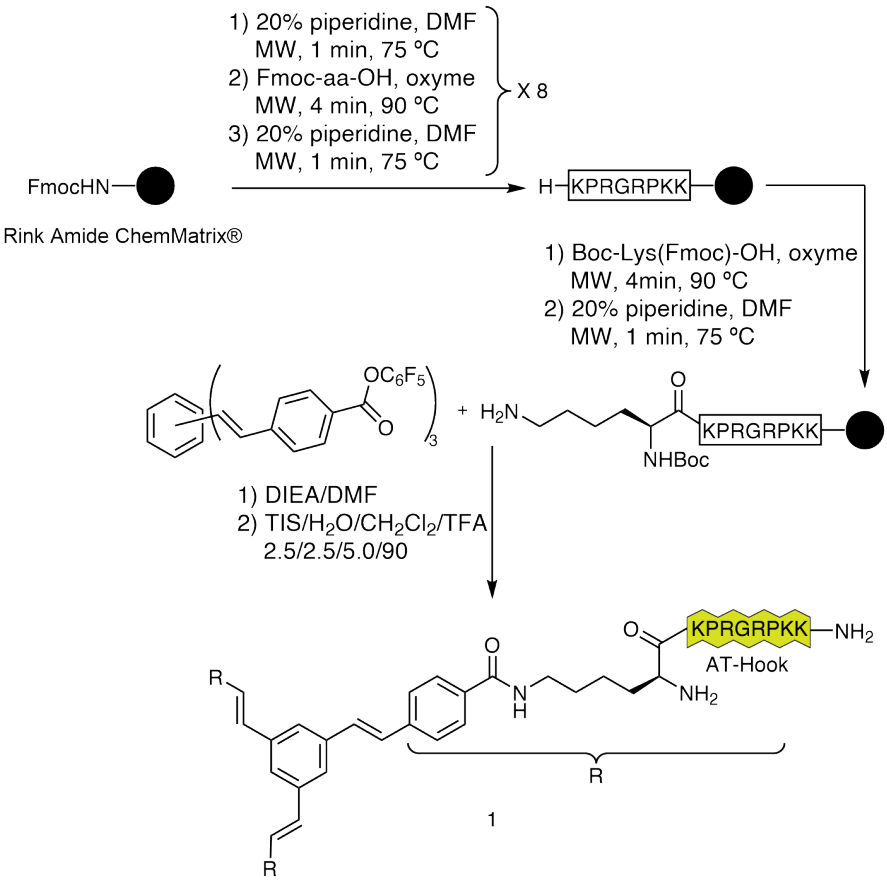


Figure 55. Best solutions of the *GaudiMM* calculation. Atoms represented in yellow ball and stick correspond to the docked 1,3,5-tristyrylbenzene moiety. AT-Hook peptides are represented in sticks and surface. The DNA scaffold in ribbons and the side chain of the nucleic acid in planks.

The AT-Hook peptide was assembled following standard Fmoc-solid-phase MW-assisted peptide synthesis protocols.[301] Next, the key step in the synthesis of the C_3-symmetric AT-Hook 3WJ binder (**1**, Scheme 5) was the on-resin simultaneous attachment of the three AT-Hook strands to the 1,3,5-tristyrylbenzene core. This was successfully achieved by incubation of the peptide, bearing the N-terminal lysine sidechain selectively deprotected, with 1,3,5-tristyrylbenzene pentafluorophenyl ester (Scheme 5). The final product was purified by reverse-phase HPLC, and once isolated, its identity was confirmed by ESI-MS (Figure 95).

[300] J. Rodríguez-Guerra Pedregal, G. Sciortino, J. Guasp, M. Municoy, J.-D. Maréchal, *J. Comput. Chem.* **2017**, *38*, 2118–2126.

[301] a) I. Coin, M. Beyermann, M. Bienert, *Nat. Protoc.* **2007**, *2*, 3247–3256; b) G. S. Vanier, *Methods Mol. Biol.* **2013**, *1047*, 235–249.



Scheme 5. Solid-phase peptide synthesis of the trimeric AT-Hook 3WJ binder, **1**, trough on-resin attachment of the three AT-Hook arms.

DNA-binding studies *in vitro*

Having at hand the desired 3WJ binder **1**, we studied its interaction with the DNA by taking advantage of the intrinsic fluorogenic properties of the 1,3,5-tristyrylbenzene core. It is well known that the excited state of these compounds is deactivated in solution by phoisomerization processes and intramolecular motions,[302] but their conformational restriction deactivates these nonradiative decay mechanisms, leading to an increased emission quantum yield.[303] Thus, a 1 μM solution of **1** was titrated with increasing concentrations of different DNAs, and the emission at 400 nm upon excitation at 320 nm was recorder after each addition. We selected two different 3WJs: TA-3WJ, containing a three high-affinity sites for the AT-Hook peptides (TTTAA), and a GC-rich 3WJ as control, GC-3WJ. In the case of TA-3WJ there is a quick drop in the emission intensity, probably due to the non-specific binding of the highly cationic **1** to the negatively charged DNA (Figure 56a, black circles at low DNA concentration). Then, in the second part of the titration, as the TA-3WJ concentration is increased and the AT-Hook peptides bind to

[302] R. Domínguez, M. Moral, M. P. Fernández-Liencres, T. Peña-Ruiz, J. Tolosa, J. Canales-Vázquez, J. C. García-Martínez, A. Navarro, A. Garzón-Ruiz, *Chem. Eur. J.* **2020**, *26*, 3373–3384.

[303] a) A. Garzón, M. P. Fernández-Liencres, M. Moral, T. Peña-Ruiz, A. Navarro, J. Tolosa, J. Canales-Vázquez, D. Hermida-Merino, I. Bravo, J. Albaladejo, J. C. García-Martínez, *J. Phys. Chem. C Nanomater. Interfaces* **2017**, *121*, 4720–4733; b) F. de Lera-Garrido, A. Sánchez-Ruiz, J. Rodríguez-López, J. Tolosa, J. C. García-Martínez, *Dyes Pigm.* **2020**, *179*, 108410.

the minor grooves in the 3WJ branches, the 1,3,5-tristyrylbenzene core is rigidified, and its emission increases. The titration profile could be fitted to a composite model that combines the formation of a specific 1:1 complex, with K_D in the order of 20 nM, with a minor contribution of a nonspecific binding term, linearly proportional to the DNA concentration (Figure 56a). On the contrary, the titration with the junction GC-3WJ also shows an abrupt drop in the emission intensity, probably resulting from the non-specific association of the probe with the polyanionic GC-3WJ and the efficient quenching of the 1,3,5-tristyrylbenzene fluorophore by the guanine bases (Figure 56b).[304] It is important to note that this quenching could be reverted by adding 1 equivalent of TA-3WJ, even in the presence of 4 equivalents of GC-3WJ, which demonstrates the high selectivity of **1** for A/T-rich 3WJs (Figure 100). The titration of **1** with a canonical DNA featuring a single AT-Hook binding site resembles that of TA-3WJ, although higher equivalents of dsDNA are required to reach the final titration plateau. This is reflected in an approximately 10 times weaker dissociation constant (Figure 56c). Furthermore, the final emission intensity is much weaker, likely due to poor restriction of intramolecular motions in the complex.

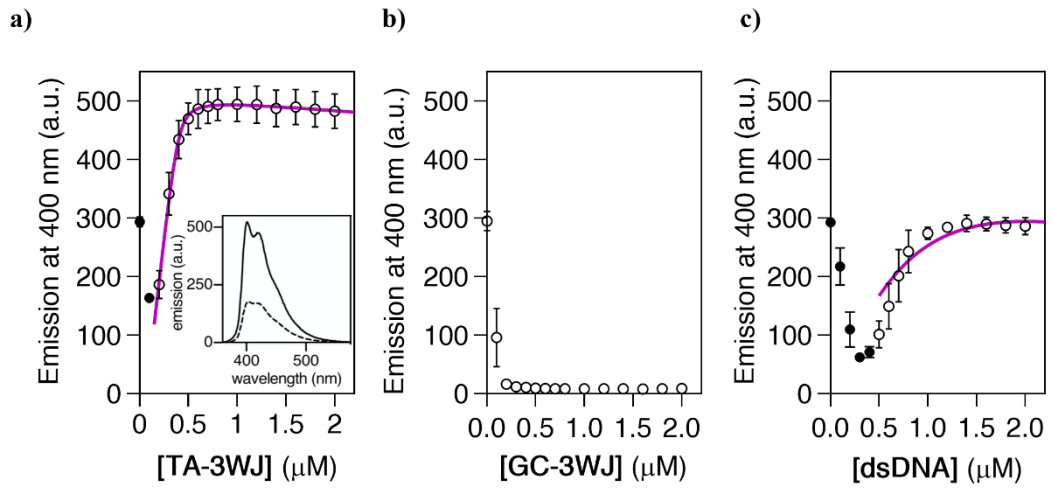


Figure 56. Fluorescence titration profiles at 400 nm of a 1 μM solution of **1** in HEPES buffer 10 mM, 100 mM NaCl, pH 7.0 with: (a) TA-3WJ; (b) GC-3WJ; (c) dsDNA. Graphs in (a) and (c) show the best fit to a mixed binding mode including 1:1 and a non-specific term. The initial points of the titrations shown as black circles were not included in the K_D calculation. Inset in (a) shows the third (dashed line) and last (solid line) emission spectra in the titration. Titration data are the average of three independent experiments DNA oligonucleotide sequences: TA-3WJ: *5'–CAC CTT TAA [AATT] TTA AAC CTC–3'; 5'–CAG GTT TAA [AATT] TTA AAG GTG–3'; 5'–GAG GTT TAA [AATT] TTA AAC CTG–3';* GC-3WJ: *5'–CAC CGG GCC [CCGG] GGC CCC CTC–3'; 5'–CAG GGG GCC [CCGG] GGC CCG GTG–3'; 5'–GAG GGG GCC [CCGG] GGC CCC CTG–3';* dsDNA: *5'–GAC GGA ATT TGA GAG CGT CG–3* (only one strand shown).

The interaction of **1** with the minor groove of the duplex branches in TA-3WJ was confirmed by Circular Dichroism (Figure 57) of a 1 μM solution of the different DNAs before and after the addition of 1 eq. of the peptide ligand **1** and the appearance of a new induced band *ca*. 325 nm (Figure 57a-b). Additionally, the same experiment in the presence of GC-3WJ did not show any new induced band (Figure 57c), thus confirming the selectivity of **1** for A/T rich sequences.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

[304] T. Heinlein, J.-P. Knemeyer, O. Piestert, M. Sauer, *J. Phys. Chem. B* **2003**, *107*, 7957–7964.

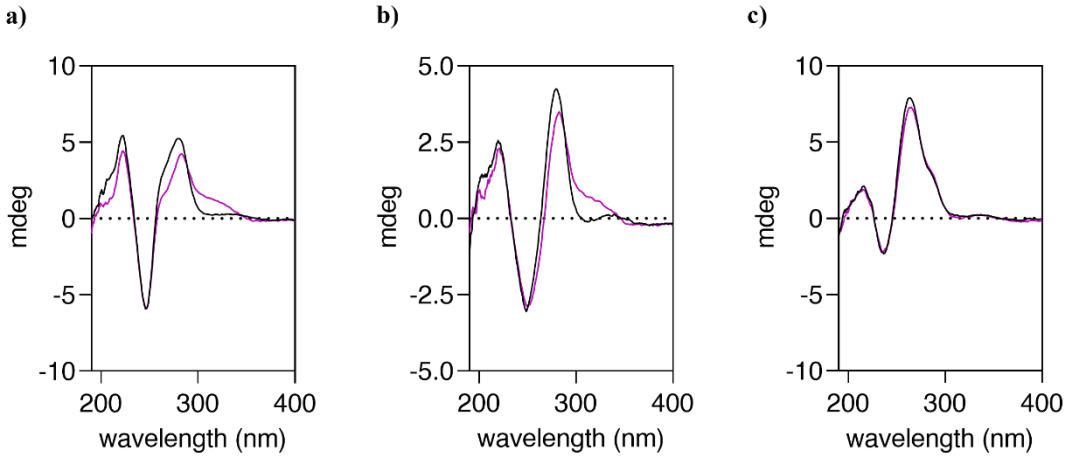


Figure 57. CD spectra of a 1 μM solution of DNA (black line) in HEPES buffer 10 mM, 100 mM NaCl, pH 7.0, and in the presence of 1 eq. of **1** (purple line). a) TA-3WJ; b) dsDNA and c) GC-3WJ.

Following the spectroscopic studies, we decided to characterize in more detail the DNA binding properties of the trimeric AT-Hook peptide **1**, by electrophoretic gel shift assays under non-denaturing conditions.[255,256] Thus, incubation of TA-3WJ with increasing concentrations of **1** gave rise to a new slow-migrating band, consistent with the formation of the expected **1**/TA-3WJ complex. Importantly, incubation gave rise only to one new band, which indicates the formation of a single new complex, presumably that in which the three AT-Hook arms are simultaneously interacting with the duplex branches of the junction (Figure 58a). A similar experiment with the non-target GC-3WJ junction also indicated the formation of a new complex, together with a reduction in the intensity of the bands and some smearing at higher concentrations of the probe, pointing to the formation of higher-order aggregates and/or multiple non-specific complexes (Figure 58b).[305] This is also consistent with the high electrostatic contribution to DNA binding with oligocationic peptides such as the AT-Hook.[306] Finally, incubation with dsDNA produced a new band only at very high concentrations of **1**, which is consistent with the lower binding affinity detected by fluorescence (Figure 58c). No binding was detected when incubating a single AT-Hook peptide with TA-3WJ or with dsDNA. In the case of dsDNA, we only observed precipitation at high peptide loadings.

[305] a) E. K. Liebler, U. Diederichsen, *Org. Lett.* **2004**, *6*, 2893–2896; b) G.-X. He, K. A. Browne, A. Blasko, T. C. Bruice, *J. Am. Chem. Soc.* **1994**, *116*, 3716–3725; c) C. Portela, F. Albericio, R. Eritja, L. Castedo, J. L. Mascareñas, *Chembiochem* **2007**, *8*, 1110–1114; d) O. Vázquez, M. E. Vázquez, J. B. Blanco, L. Castedo, J. L. Mascareñas, *Angew. Chem. Int. Ed Engl.* **2007**, *46*, 6886-6890.

[306] C. Crane-Robinson, A. I. Dragan, P. L. Privalov, *Trends Biochem. Sci.* **2006**, *31*, 547–552.

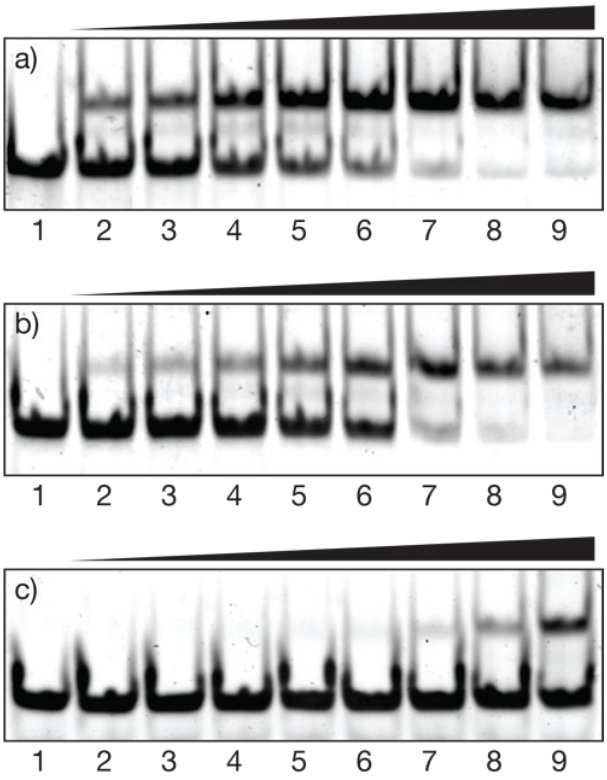


Figure 58 PAGE analysis of the DNA binding properties of **1**. a) Lane 1-9: 200 nM of TA-3WJ in the presence of 0, 100, 200, 300, 400, 500, 600, 700 and 800 nM of **1**, respectively; b) sames as in a) but with GC-3WJ; c) same as in a) but with dsDNA. DNA oligonucleotide sequences: TA-3WJ: *5'–CAC CTT TAA [AATT] TTA AAC CTC–3'; 5'–CAG GTT TAA [AATT] TTA AAG GTG–3'; 5'–GAG GTT TAA [AATT] TTA AAC CTG–3';* GC-3WJ: *5'–CAC CGG GCC [CCGG] GGC CCC CTC–3'; 5'–CAG GGG GCC [CCGG] GGC CCG GTG–3'; 5'–GAG GGG GCC [CCGG] GGC CCC CTG–3';* dsDNA: *5'–GAC GGA ATT TGA GAG CGT CG–3* (only one strand shown).

DNA-binding studies in functional cells

Having demonstrated that the trimeric peptide **1** selectively binds to A/T-rich 3WJs *in vitro*, we studied its application to label these structures in HeLa cells, by exploiting the inherent fluorogenic properties of its 1,3,5-tristyrylbenzene core. Preliminary experiments showed that **1** was poorly internalized (incubation for 30 min of 5 μM concentration of **1**). However, if the cells were pretreated with digitonin, then **1** was efficiently internalized,[307] and appeared in the blue emission channel as a punctuated pattern in the cytoplasm and the nucleus (Figure 59a). Remarkably, the punctuated distribution in the cell nucleus has a strong correlation (Pearson´s correlation coefficient 0.60)[308] with the localization of the DNA replicating sites, that were labelled with the proliferating cell nuclear antigen (PCNA) fused to GFP (Figure 59b-c).[259] Notably, while the *in vitro* dynamic range would hardly be enough to differentiate the free/bound states, by fluorescence microscopy the probe is greatly concentrated in the replication foci,

[307] a) M. Nishikawa, S. Nojima, T. Akiyama, U. Sankawa, K. Inoue, *J. Biochem.* **1984**, *96*, 1231–1239; b) M. P. Stewart, R. Langer, K. F. Jensen, *Chem. Rev.* **2018**, *118*, 7409–7531.

[308] a) K. W. Dunn, M. M. Kamocka, J. H. McDonald, *Am. J. Physiol. Cell Physiol.* **2011**, *300*, C723–42; b) J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, *Nat. Methods* **2012**, *9*, 676–682.

where the 3WJs from multiple replication forks are accumulated, thus allowing the clear labelling of these sites. The staining in the cytoplasm and of what appears to be the cell nucleolus is consistent with the binding of **1** to RNA structures, which are known potential binding targets of the AT-Hook peptides.[309]

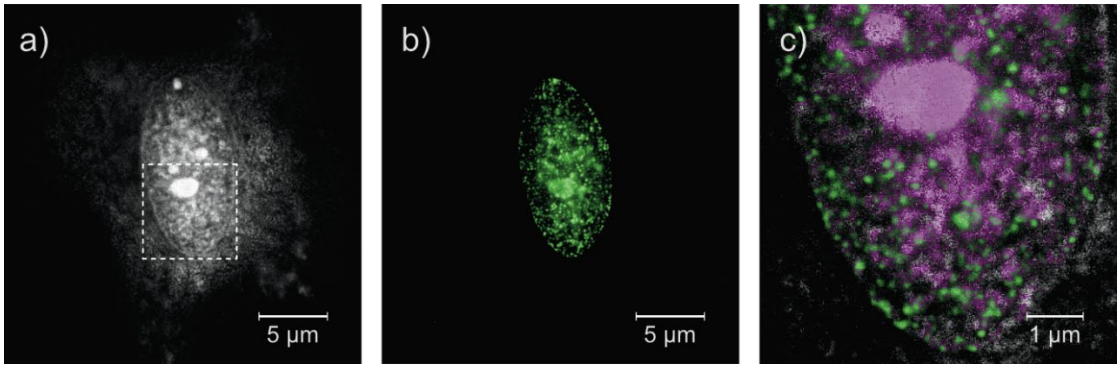


Figure 59. DNA replication sites in HeLa cell nucleus. HeLa cells expressing protein GFP-PCNAL2 were incubated in the presence of 25 µg/mL of digitonin for 3min, and then in the presence of 5 µM of **1** for 30 min. a) Blue channel emission showing the distribution of **1** (channel fake colored in white for easier view); b) green channel, corresponding to the emission of the GFP-PCNAL2 probe labeling the DNA replication foci in the cell nucleus; c) overlay of the green and blue channels with the difference filter applied, (so white areas in a) invert the green color and appear as purple) to highlight the overlay between the staining of GFP-PCNAL2 and that of **1**. The image is a 1000 x magnification confocal image of a representative cell in the slide, all of them displaying a similar emission pattern.

USC UNIVERSIDADE DE SANTIAGO

[309] M. Filarsky, K. Zillner, I. Araya, A. Villar-Garea, R. Merkl, G. Längst, A. Németh, *RNA Biol.* **2015**, *12*, 864–876.

Conclusions

In this chapter, we have described a new approach for selectively recognizing 3WJs through sequence-selective and cooperative binding of three AT-Hook peptides in their three duplex DNA arms. Moreover, we have exploited the fluorogenic properties of the 1,3,5-tristyrylbenzene core and the increased emission resulting from the conformational restrictions associated with 3WJ binding to selectively label DNA replication sites in functional cells.

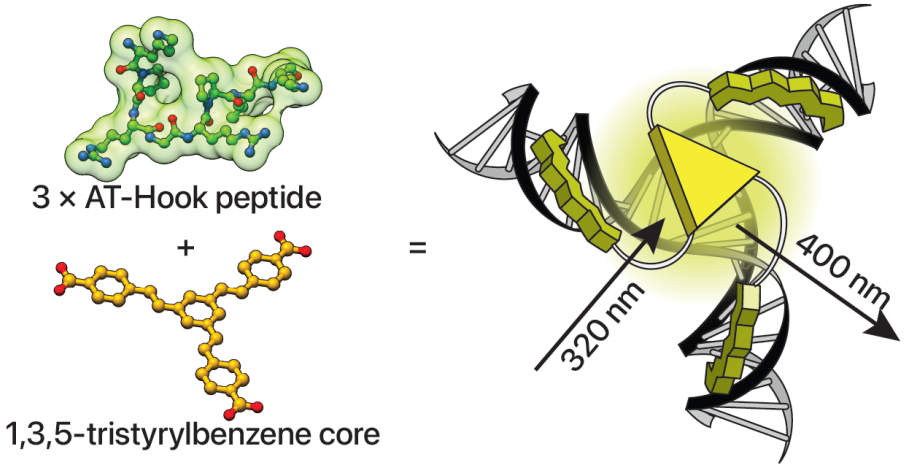


Figure 60. Schematic representation of our strategy to target 3WJs, based on the cooperative binding of three AT-Hook peptides to the minor grooves of the duplex DNA arms, resulting in an increased emission of the 1,3,5-tristyrylbenzene core.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

General Conclusions

1) We have demonstrated that the known nuclease properties of polypyridine copper(II) complexes can be exploited in peptide helicates to afford **selective nucleases targeting 3WJs** structures *in vitro* and in mammalian cells.

2) We have shown that **ΔΔ-Cu^{II}_2LLD** helicate **can selectively bind and cleave 3WJs** via a ROS-mediated mechanism, both *in vitro* and in functional Vero cells.

3) We have synthesized Cu(II) and Ni(II)-metallopeptides derived from a new chimeric ligand composed of a peptide sequence that forms helicates in the presence of octahedral metal ions, together with the ATCUN motif, which is known to display nuclease activity. **ΔΔ-Cu^{II}_3RGH-LLD** and **ΛΛ-Ni^{II}_3RGH-LLD** have shown high affinity and binding selectivity towards 3WJs both *in vitro* and in functional HeLa cells. However, only the Cu(II) metallopeptide exhibited cleavage properties, thus suggesting that the nuclease activity of these peptide helicate chimeras is independent of the ATCUN motif, being inherent to the peptide helicate unit.

4) We have engineered a new peptide ligand featuring trigonal symmetry, which gave rise to a **dynamic system with metal-controlled 3WJ recognition**. In the presence of Co(II), this peptide self-assembles into discrete helicates with DNA binding properties, whereas in the absence of metal ions, it aggregates into nanofibers.

5) A **kinetically inert Co(III)-metallocylinder** with selective 3WJ binding properties has been obtained through oxidation of these latter dynamic Co(II)-peptide helicate.

6) By reproducing the mechanism used in natural proteins, we designed a new system comprising three AT-Hook peptide strands, that **selectively and cooperatively recognizes A/T-rich sequences** in 3WJs.

7) All these systems have displayed high binding affinity to 3WJs with K_D values in the nanomolar range, as well as high selectivity towards this non-canonical structure versus regular B-DNA, as verified by PAGE analysis. Moreover, **ΔΔ-Cu^{II}_2LLD**, **ΛΛ-Ni^{II}_3RGH-LLD** and the trimeric **AT-Hook** peptide have been found to accumulate in the cell nuclei, mostly in **active replicating sites**.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Experimental Section

1. Reactives

All the solvents employed for the synthesis of the Fmoc-βAla5Bpy-OH were obtained from different commercial sources: *Fisher Chemical* (acetonitrile for HPLC, chloroform, toluene, DMF, methanol, NaOH), *Scharlau* (absolute ethanol), *Sigma Aldrich* (5,5'-dimethyl-2,2'-bipyridine, $KMnO_4$, $SOCl_2$, hydrazine monohydrate, xylene, DIPEA), *Panreac* (celite and $NaNO_2$).

For the synthesis of the peptide ligands all the amino acids were provided by *Sigma Aldrich*, except for the orthogonally protected Fmoc-Lys(Alloc)-OH, which was purchased from *Bachem*. All the amino acids were used as their Fmoc protected derivatives. The resin employed for the SPPS was H-Rink-Amide *ChemMatrix* 35-100 mesh particle size from *Sigma Aldrich*. For the Alloc deprotection, $Pd(PPh_3)_4$, $PhSiH_3$ were obtained from *Sigma Aldrich*. 6-Carboxytetramethylrhodamine succinimidyl ester (TAMRA) was provided by *Carbosynth.* The 1,3,5-tris(bromomethyl)benzene was purchased from *Alfa Aesar* and pentafluorophenol from *Sigma Aldrich.* DCC was provide by *Alfa Aesar*. Solvents for SPPS and organic synthesis were provided by Fischer Scientific.

The metal salts employed in the obtention of the metallopeptides: $CuCl_2 \cdot 2\ H_2O$ was acquired to *Acros Organics;* Mohr's salt, $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$, was provided by *Sigma Aldrich*; $Co(ClO_4)_2 \cdot 6\ H_2O$, $(NH_4)_2Ce(NO_3)_6$ and $NiCl_2$ were provided by *Alfa Aesar.*

2. Procedures and experimental techniques

2.1. Mass spectrometry

Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) was performed with a *Bruker Autoflex MALDI-TOF* model in positive scan mode by direct irradiation of the matrix-adsorbed peptide. 4-HCCA (α-cyano-4-hydroxycinnamic acid) was the selected matrix for all these experiments.

2.2. UV-Vis Spectroscopy

UV measurements were made in a Jasco V-630 spectrophotometer coupled to a Jasco ETC-717 temperature controller, using a standard Hellma semi-micro cuvette (108.002-QS) with a light path of 10 mm. Measurements were made at 20 °C.

Acquisition parameters were: 220-700 nm range, scan speed of 200 nm/min, resolution of 0.2 nm.

2.3. Circular Dichroism

Circular dichroism spectra measurements were made with a Jasco J-715 coupled to a Neslab RTE-111 thermostated water bath, using a Hellma 100-QS cuvette (2 mm light pass). Scan speed was 200 nm/min and the obtained spectra are the mean of three accumulations.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

2.4. EPR studies

EPR measurements were made in a *Bruker EMX* spectrometer.

2.5. DNA hybridization

Oligonucleotides were obtained from Biomers. Concentration of the oligonucleotides was measured by UV-VIS by using the extinction coefficient given by the Commercial House *Biomers*. For the hybridization process, a stochiometric mixture of the DNA strands was prepared in H2O MQ and heated at 90°C for 10 minutes. After this time the mixture was slowly cooled down until room temperature, obtaining the resulting 3WJ/ds hybridized DNA.

- Fluorescein-labelled 3WJ (Flu-3WJ):

  Y1: 5'–FAM-TTTT CAC CGC TCT GGT CCT C–3'

  Y2: 5'–CAG GCT GTG AGC GGT G–3'

  Y3: 5'–GAG GAC CAA CAG CCT G–3'

- 3WJ:

  Y1: 5'–CAC CGC TCT GGT CCT C–3'

  Y2: 5'–CAG GCT GTG AGC GGT G–3'

  Y3: 5'–GAG GAC CAA CAG CCT G–3'

- Fluorescein-labelled TA-3WJ:

  Y1: 5'–FAM-TTTT CAC CTT TAA AAT TTT AAA CCT C–3'

  Y2: 5'–GAG GTT TAA AAT TTT AAA CCT G–3'

  Y3: 5'–CAG GTT TAA AAT TTT AAA GGT G–3'

- TA-3WJ:

  Y1: 5'–CAC CTT TAA AAT TTT AAA CCT C–3'

  Y2: 5'–GAG GTT TAA AAT TTT AAA CCT G–3'

  Y3: 5'–CAG GTT TAA AAT TTT AAA GGT G–3'

- Fluorescein-labelled GC-3WJ:

  Y1: 5'–FAM-TTTT CAC CGG GCC CCG GGG CCC CCT C–3'

  Y2: 5'–CAG GGG GCC CCG GGG CCC GGT G–3'

  Y3: 5'–GAG GGG GCC CCG GGG CCC CCT G–3'

- GC-3WJ:

  Y1: 5'–CAC CGG GCC CCG GGG CCC CCT C–3'

  Y2: 5'–CAG GGG GCC CCG GGG CCC GGT G–3'

  Y3: 5'–GAG GGG GCC CCG GGG CCC CCT G–3'

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

- Fluorescein-labelled dsDNA:

5'–FAM-TTTT AAC ACA TGC AGG ACG GCG CTT–3' (only one strand shown)

- dsDNA:

5'–AAC ACA TGC AGG ACG GCG CTT–3' (only one strand shown)

- A/T-rich dsDNA:

5'–GAC GGA ATT TGA GAG CGT CG–3' (only one strand shown)

2.6. Solid phase peptide synthesis

2.6.1. LLD, DDL and RGH-LLD peptide synthesis

C-terminal amide **LLD**, **DDL** and **RGH-LLD** peptides were synthesized following standard Fmoc-peptide synthesis protocols on a 0.1 mmol scale using a 0.5 mmol/g loading H-Rink amide ChemMatrix resin (35–100 mesh size particle). The obtention of this peptide was carried out by hand, using HBTU as activating agent for the natural amino acids of the sequence and HATU for the Fmoc-βAla5Bpy-OH residues. Each amino acid was activated for 1 min in DIEA/DMF 0.2 M (4 mL) before being added onto the resin. These manual couplings were conducted for 60 min. Deprotection of the temporal Fmoc protecting group was performed by treating the resin with 20% piperidine in DMF for 20 min.

All the couplings were carried out in a 5-fold excess of the corresponding amino acid for the natural ones and in a 3.5-fold excess for the Fmoc-βAla5Bpy-OH.

For the synthesis of **TAMRA-LLD**, the N-terminal capping with TAMRA was carried out by reacting the N-terminal deprotected amine with the activated fluorophore—6-Carboxytetramethylrhodamine succinimidyl ester (TAMRA-OSu)—in 1 mL DIEA/DMF 0.2 M during 2 h (approximately 7.5 mg of the reactive fluorophore/40 mg of resin).

2.6.2. TAMRA-RGH-LLD peptide synthesis

In the case of the TAMRA-labelled RGH-LLD peptide, the fluorophore was coupled to a free amine in a lysine residue of the linker sequence. The first step was the orthogonal deprotection of the Alloc group in the lysine, which was carried out following the methodology proposed by Sainlos *et al.*[310] Thus, a suspension of the resin in DCM was degassed by bubbling of N_2 for 5 minutes. After that, $Pd(PPh_3)$ (0.8 eq.) and $PhSiH_3$ (25 eq.) were added and the flux of N_2 was maintained for 15 minutes. After this time, the resin was washed with DCM three times, and the previous process repeated twice. Finally, the resin was filtered and washed with DMF, DEDTC and DMF again. Afterwards, the deprotected amine in the lysine residues was reacted with 6-Carboxytetramethylrhodamine succinimidyl ester (TAMRA-OSu)—in 1 mL DIEA/DMF 0.2 M during 2 h (approximately 7.5 mg of the reactive fluorophore/40 mg of resin).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

[310] M. Sainlos, B. Imperiali, *Nat. Protoc.* **2007**, *2*, 3201-3209.

2.6.3. L trigonal peptide synthesis

The peptide strands were synthesized following standard Fmoc-peptide synthesis protocols on a 0.1 mmol scale using a 0.5 mmol/g loading H-Rink amide *ChemMatrix* resin (35–100 mesh size particle). All the amino acids were coupling by hand using HATU as the activating agent. Each amino acid was activated for 1 min in DIEA/DMF 0.2 M (4 mL) before being added onto the resin. These manual couplings were conducted for 60 min. The coupling of the (L)-Arg residue was carried out in a 5-fold excess whereas a 3.5-fold excess was used for the Fmoc-βAla5Bpy-OH. The deprotection of the final Fmoc protecting group was performed by treating the resin with 20% piperidine in DMF for 20 min. The key step in the synthesis of **L** was the attachment of the benzene core using the S_N2 reaction, that was carried out on the resin itself. Thus, a suspension of the resin in a solution of 1,3,5-tris(bromomethyl)benzene (0.33 eq.) and DIEA (3 eq.) in 5 mL of DMF, was stirred at room temperature for 24 h. Thereafter, the resin was filtered and washed with DMF three times.

2.6.4. Trimeric AT-Hook peptide 1 synthesis

C-terminal amide natural AT-Hook derivatives were synthesized following standard Fmoc-peptide synthesis protocols on a 0.1 mmol scale using a 0.5 mmol/g loading H-Rink amide *ChemMatrix* resin (35-100 mesh size particle) with a Liberty Lite automatic microwave assisted peptide synthesizer from CEM Corporation. The amino acids were coupled in 5-fold excess using oxyme as an activating agent. Couplings were conducted for 4 min at 90 °C. Deprotection of the temporal Fmoc protecting group was performed by treating the resin with 20% piperidine in DMF for 1 min at 75 °C. After the obtention of the monomer AT-Hook peptide, trimerization was carried out by hand employing 0.3 eq of the trispentafluorophenol ester in DIEA/DMF 0.2M during 2h.

2.6.5. Cleavage from the resin and final deprotection

Cleavage and deprotection of the peptides were simultaneously performed by treatment of the resin-bound peptide for 2.5 h with an acidic cocktail containing 50 µL of CH_2Cl_2, 25 µL of H_2O, 25 µL of TIS (triisopropylsilane), and 900 µL of TFA (1 mL of cocktail / 40 mg of resin). The resin was filtered, and the TFA filtrate was concentrated under a nitrogen stream to an approximate volume of 1 mL, and then added onto ice-cold diethyl ether (20 mL). After 10-30 min, the precipitate was centrifuged and washed again with 5 mL of ice-cold ether. The solid residues were dried under N_2, re-dissolved in a mixture of H_2O and CH_3CN and purified by preparative HPLC for the obtention of the pure compounds.

2.7. HPLC and UHPLC

2.7.1. UHPLC-MS

Peptide analysis was performed by analytical UHPLC-MS with an *Agilent 1200 series LC/MS* using a *SB C18* (1.8 µm, 2.1 × 50mm) analytical column from *Phenomenex*. Standard conditions for analytical UHPLC consisted of a linear gradient from 5% to 95% of solvent B in 12 min at a flow rate of 0.350 mL/min (A: water 0.1% TFA, B: acetonitrile 0.1% TFA). Compounds were detected by UV absorption at 222, 254, 280, 310, 330 and 550 nm. Electrospray Ionization Mass Spectrometry (ESI/MS) was performed with an

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Agilent 6120 Quadrupole LC/MS model in positive scan mode using direct injection of the purified peptide solution into the MS detector.

2.7.2. Preparative HPLC

Peptide purification was performed by preparative RP-HPLC with a *Waters 1500 series Liquid Chromatograph* using a *Sunfire Prep C18 OBD* (5 μm, 19 × 150 mm) reverse-phase column from *Waters*. Standard conditions for preparative RP-HPLC consisted of an isocratic regime during the first 5 min, followed by different linear gradients of solvent B during 30 min (A: water 0.1% TFA, B: acetonitrile 0.1% TFA). The gradients were adjusted for each peptide. Collected fractions with pure products were lyophilized with a *ThermoSavant Modulyo D lyophilizator* equipped with an *Edwars RV* high vacuum pum.

2.8. Fluorescence spectroscopy

Luminescence experiments were made with a *Varian Cary Eclipse* Fluorescence Spectrophotometer coupled to a Cary *Single Cell Peltier* accessory (Agilent Technologies) temperature controller. All measurements were made with a Hellma semi-micro cuvette (108F-QS) at 20 °C. The settings for these measurements were optimized depending on the studied system.

2.8.1. Data analysis in fluorescence titrations

Fluorescence titration data were analyzed using the *DynaFit 4.0* software, which characterizes the reacting system in terms of stoichiometric equations instead of mathematical notation; the chemical equations are then translated into the underlying mathematical equations using matrix theory.[252] *DynaFit* is available free of charge at http://www.biokin.com/dynafit/. The program requires plain text files, known as scripts, that contain the chemical model underlying the experimental data, the values of model parameters, such as starting concentrations of reactants, as well as information about location of the files. A typical script used in the analysis of these type of titrations is included below. The file has been commented to indicate the purpose of the keywords and sections, but the reader is recommended to review the *DynaFit* scripting manual distributed along the program or available at the *DynaFit* website.

```
[task]                    ;semicolons indicate comments
 task = fit               ;nature of the calculation to be performed
 data = equilibria
[mechanism]               ;Free-form 1:1 binding model with Kd
 R + L <==> RL : Kd dissoc;to be calculated as dissociation constant
[constants]               ;Initial Kd value for iteration
 Kd = 1.0 ?               ;the "?" indicates that this will be optimized
[concentrations]
 R = 2.0                  ;Fixed conc. of the DNA during the peptide titration
[responses]               ;contribution to the spectroscopic signal of each
 R = 0.1 ?                ;of the different components of the equilibrium
 RL = 1.5 ?               ;these will be optimized ("?" after the values)
[data]                    ;location of files and information about the data
 variable  L              ;the species that changes conc. during the titration
 offset auto ?
 directory ./exp/SH ;file path (relative to DynaFit program location)
 extension txt
 file f1                  ;name of the experimental data file
[output]
 directory ./exp/SH/out ;path indicating location of DynaFit output files
```

```
[settings]                ;cosmetic settings that control DynaFit graphics
 {Output}                 ;fits were exported & finally plotted with
 XAxisUnit = uM           ;GraphPad Prism 7.0c. GraphPad Software,
 BlackBackground = n      ;La Jolla California, www.graphpad.com
 XAxisLabel = [peptide]
 YAxisLabel = anisotropy
 WriteTXT = y
```

2.9. Polyacrylamide gel electrophoretic analysis

Polyacrylamide gel electrophoresis (PAGE) experiments with fluorophore-labelled DNA were performed using an *Invitrogen mini gel tank*. Gels were cast using a *SureCast*^{TM} gel casting system (*Invitrogen*) and powered by a *BioRad basic power pac*$^{TM.}$ PAGE gels were visualised using a *G:Box 9* mini gel documentation system and analysed with *GeneSys* software (*Syngene*). The multiplex assay was performed on *GeneSys* software (S*yngene*); an image with each the appropriate excitation and emission for wach fluorophore was captured and the images were overlayed, and a false colour applied for each fluorescent signal. Band densitometry was performed in triplicate on gels post-stained with *SYBR gold* (1X, *Invitrogen*) using *GeneTools* software (*Syngene*). Statistical analysis was performed using *GraphPad Prism*.

PAGE experiments with not-labellled DNA were performed with a *BioRad Mini Protean* gel system, powered by an electrophoresis power supplies *PowerPac Basic model*, maximum power 150 V, frequency 50–60 Hz at 130 V (constant V).

2.9.1. Not labelled 3WJ and dsDNA binding studies

The DNA sequences used in these studies can be found in section 2.5 of this file. Binding reactions were performed for 30 or 60 min in 18 mM Tris-HCl (pH 7.5), 90 mM KCl, 1.8 mM $MgCl_2$, 9% glycerol, 0.11 mg/mL BSA, and 2.2% NP-40. For the experiments we used 200 nM of the DNAs (3WJ and dsDNA), different quantities of the specific helicate or peptide **1** in a total incubation volume of 20 µL. After incubation for at room temperature, products were resolved by PAGE using a 10% non-denaturing polyacrylamide gel and 0.5 × TBE buffer (0.445 M Tris, 0.445 M Boric acid) for 35 min at 130 V at 25 °C, and analyzed by staining with SyBrGold (Molecular Probes: 5 µL in 50 mL of 1 × TBE) for 10 min, followed by fluorescence visualization (*BioRad* GelDoc XR+ molecular imager).

2.9.2. Fluorophore labelled 3WJ and dsDNA binding studies with ΔΔ-Cu^{II}_2LLD

These experiments were carried out in collaboration with the **Prof. Andrew Kellett** from the School of Chemical Sciences at the Dublin City University (DCU, Ireland)

Fluorescently labelled 3WJ (5 pmol) and dsDNA (5 pmol) sequences (see below) were incubated with 1, 2 or 3 eq. of **ΔΔ-Cu^{II}_2LLD** (5, 10 or 30 pmol) in HEPES buffer (10 mM HEPES, 750 mM NaCl, 50 mM $MgCl_2$, pH 7) at 37 °C for 2 h prior to 20% native PAGE analysis. The gel was run in 50 mM Tris-acetate buffer at 70 V for 240 min and visualized with a multiplex assay. FAM fluorescence was observed with blue light and 525 nm filter, ROX was excited by UV fluorescence and Cy5 was visualized with red light and a 705 nm filter. The DNA sequences used in these studies are the following:

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

- 3WJ:

  Y1: 5'–FAM-TTT TCA CCG CTC TGG TCC TC–3'

  Y2: 5'–ROX-TTT TCA GGC TGT GAG CGG TG–3'

  Y3: 5'–Cy5-TTT TGA GGA CCA ACA GCC TG–3'

- dsDNA:

  Z1: 5'–FAM-TTT TAA CAC ATG CAG GAC GGC GCT T–3'

  Z2: 5'-Cy5-TTT TAA GCG CCG TCC TGC ATG TGT T-3'

2.9.3. In vitro nuclease activity studies of ΔΔ-Cu^{II}_2LLD

These experiments were carried out in collaboration with the **Prof. Andrew Kellett** from the School of Chemical Sciences at the Dublin City University (DCU, Ireland)

Nuclease activity was analysed by PAGE. Equal concentrations of each strand of the 3WJ were mixed in HEPES buffer (10 mM HEPES, 750 mM NaCl, 50 mM $MgCl_2$, pH 7) and heated to 95 °C before being cooled in 5 °C increments to 4 °C. The annealed 3WJ (5 pmol), and an *off*-target DNA duplex (5 pmol), were incubated at 37 °C for 24 hours with **ΔΔ-Cu^{II}_2LLD** (1, 2, 5, 7.5 or 10 equivalents, that is, 5, 10, 25, 37.5 or 50 pmol) and Na-(L)-ascorbate (10 eq. to **ΔΔ-Cu^{II}_2LLD**, that is, 50, 125, 250, 375 or 500 pmol). The reaction was quenched with loading dye (6X, 10 mM Tris-HCL, pH 7.6, 0.03% bromophenol blue, 0.03% xylene cyanol, 60 mM EDTA) and loaded onto a 20% native PAGE gel (50 mM Tris-acetate, 75 mM NaCl, 5 mM $MgCl_2$, pH 6.1). The gel was run in 50 mM Tris-acetate buffer at 70 V for 240 min, and post-stained with SYBR gold (1X, Invitrogen). Gel was visualized under UV fluorescence. Band densitometry was performed in triplicate, and the significance of the data was analyzed by Two-way ANOVA with Tukey's multiple comparisons.

Cleavage activity on fluorescently labelled DNA. Samples were prepared as described above, but fluorophore labelled 3WJ and dsDNA strands were utilized. Cleavage assay was quenched and loaded onto a 20% native PAGE gel (50 mM Tris-acetate, 75 mM NaCl, 5 mM $MgCl_2$, pH 6.1). The gel was run in 50 mM Tris-acetate buffer at 70 V for 240 min and visualized using the multiplex assay described above.

ΛΛ-Fe^{II}_2LLD cleavage. Samples containing 3WJ and dsDNA were prepared as previously described but **ΔΔ-Cu^{II}_2LLD** was replaced with **ΛΛ-Fe^{II}_2LLD**. All other reaction components were kept consistent with the Cu^{II} helicate cleavage assay.

Time-point study. Samples containing 3WJ, dsDNA, **ΔΔ-Cu^{II}_2LLD** (1, 5 and 10 eq. to DNA) and Na-(L)-ascorbate were prepared as described above, and incubated a various time points (1 hour, 3 hours and 12 hours) at 37 °C before being quenched with loading dye (6X, 10 mM Tris-HCL, pH 7.6, 0.03% bromophenol blue, 0.03% xylene cyanol, 60 mM EDTA) and loaded onto a 20% native PAGE gel (50 mM Tris-acetate, 75 mM NaCl, 5 mM $MgCl_2$, pH 6.1). Gel was run at 70 V for 240 min, post-stained with SYBR gold (1X, Invitrogen) and visualized by UV fluorescence.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Reactive oxygen species analysis. The 3WJ and **ΔΔ-Cu^{II}_2LLD** samples were prepared as previously described, but the dsDNA *off*-target was replaced with a reactive

oxygen species scavenger (50 nmol); (L)-His, D-mannitol, (L)-Met, tiron (4,5-dihydroxy-1,3-benzendisulfonic acid), sodium azide (NaN_3) or deuterium oxide (D_2O) and incubated at 37 °C for 24 h. Products were analyzed by 20% Tris-acetate PAGE (50 mM Tris-acetate, 75 mM NaCl, 5 mM $MgCl_2$, pH 6.1).

2.9.4. In vitro nuclease activity studies of ΔΔ-Cu^{II}_3RGH-LLD and ΛΛ-Ni^{II}_3RGH-LLD

The same conditions as described in section 2.9.1 were used to analyse the nuclease activity of **ΔΔ-Cu^{II}_3RGH-LLD** and **ΛΛ-Ni^{II}_3RGH-LLD,** but in this case, the samples were incubated in the presence of 250 μM of H_2O_2.

2.10. Cell internalization studies

These experiments were carried out in collaboration with the **Dr. Natalia Barreiro Piñeiro** and **Prof. José Martínez Costas** from CiQUS.

Vero or HeLa cells were grown in DMEM (Dulbecco modified Eagle Medium) with 10% of fetal bovine serum in glass-bottom 35 mm tissue culture plates (IBIDI). One day before the internalization experiment, the cells were transfected with plasmid pc653-pENeGFPCNAL2mut[259] using PEI Polyethylenimine HCl MAX (PEImax®, obtained from Polysciences) following the instructions from the manufacturer. That plasmid drives the expression of GFP-tagged Proliferating Cellular Nuclear Antigen (PCNA) to label DNA replication factories. The next day the cells were washed twice with transfer buffer (25 mM HEPES, 125 mM KOAc, 2 mM $Mg(AcO)_2$, 1 mM EDTA). Then, cells were treated with a digitonin solution (25 μg/mL) in the same previous buffer at 0 °C for 3 minutes. After this time cells were washed again twice with transfer buffer containing 10 mg/mL of BSA. Cells were then incubated in transfer buffer with 5 μM of the specific TAMRA-labelled metallopeptide or peptide **1** for 15 minutes at 37 ºC. After this time cells were washed three times with transfer buffer containing 10 mg/mL of BSA and finally covered with DMEM supplemented with 10% of fetal bovine serum.

Images were obtained with an *Andor Dragonfly* spinning disk confocal system mounted on a *Nikon* TiE microscope equipped with a Zyla 4.2 PLUS camera (*Andor*).

2.11. Nuclease activity assays in cells

These experiments were carried out in collaboration with the **Dr. Natalia Barreiro Piñeiro** and **Prof. José Martínez Costas** from CiQUS.

Cell synchronization. $4x10^6$ Vero cells were grown in T175 flasks. After 24 hours of incubation at 37 ºC with 5% CO_2, the medium was removed, and cells were washed three times with PBS 1X and incubated in DMEM-HEPES with 1% antibiotics and without Glutamine and FBS. After 42 h, the medium was replaced by DMEM + 10% FBS and 1% glutamine + antibiotics. After 16 hours the cells were trypsinized and counted.

In cellulo nuclease assay. Vero cells were pelleted (500 x g, 5 min 4 ºC) and washed three times with Transfer Buffer (25 mM HEPES pH 7.9, 125 mM KOAc, 2 mM $Mg(AcO)_2$, 1 mM EDTA). Subsequently, cells were treated with 25 μg/ml digitonin in Transfer Buffer for 3 min on ice. After this time cells were centrifuged, and the supernatant were removed. Cells were resuspended according to the treatment in: Transfer buffer, Transfer buffer + 2.5 mM H_2O_2, Transfer buffer + 2.5 mM H_2O_2 + 5 μM **ΛΛ-Fe^{II}_2LLD** and Transfer buffer + 2.5 mM H_2O_2 + 5 μM **ΔΔ-Cu^{II}_2LLD** and incubated 2 hours at 37 ºC shaking every 15 minutes. After this time, cells were centrifuged and

washed twice with Transfer buffer + 10 mg/ml BSA and twice with PBS 1X. Finally, cells were fixed for the TUNEL Assay.

TUNEL Assay. Cells were fixed with 5 mL of 1% PFA (paraformaldehyde) in PBS for 15 min at 4 ºC. Subsequently, cells were centrifuged (500 x g, 5 min, 4 °C) and washed twice with PBS 1X. Then, the cells were resuspended in 5 mL of ice-cold 70% ethanol and incubate overnight 5 days at −20 ºC. Then, the cells were collected by centrifugation and subjected to TUNEL assay following the manufacturer´s instructions (APO-BrdU^{TM} TUNEL Assay Kit, Thermo Fisher Scientific). Briefly, they were washed with 1 mL of Wash buffer and then incubated with DNA-labeling solution overnight at RT on a nutator mixer. Next, the cells were washed twice with Rinse Buffer and incubated with Alexa Fluor 488^{TM} dye-labeled anti-BrdU antibody for 30 min at RT on a nutator and analyzed by flow cytometry using a Guava easyCyte HT Millipore flow cytometer equipped with a 488 nm blue laser coupled with a 525/30 nm filter.

2.12. DLS and microscopy studies

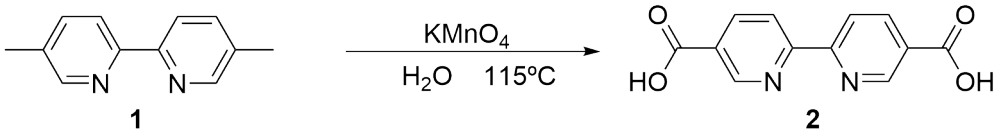
These experiments were carried out in collaboration with **Prof. Félix Freire** from CiQUS.

DLS studies were performed on a *Nano-ZS 90 (Malvern)* equipped with a *He-Ne laser* (l = 633 nm) under scattering angle of 173°. The samples were maintained at the designed temperature for 5 min before testing. DLS measurements were carried out in all cases at 0.1 mg / mL. STEM measurements were performed on a *LEO-435VP electron microscope* equipped with an energy dispersive X-ray (EDX) spectrometer. To study the morphology of the aggregate, a dispersion of the **L** at a concentration of 0.1 mg/mL was drop-casted onto carbon chip and allowed to dry at rt for 12 h.

2.13. Synthesis of Fmoc-βAla5Bpy-OH (9a)

The synthesis of this chelating residue was carried out following the methodology recently published by our group.[156] All the obtained products were analyzed by HPLC-MS to test their purity.

2.13.1. Synthesis of 2,2'-bipyridine-5,5'-dicarboxilic acid (2)



1 **2**

A mixture of 39 g of potassium permanganate and 7 g of 5,5'-dimethyl-2,2' -bipyridine in 250 mL of H_2O was heated for 2 h (115 ºC), cooled at room temperature and filtrated trough celite. The filtrate was cooled to 4 ºC and acidified with HCl until precipitation of a white solid, which was filtrated, washed with water and lyophilized to afford the desired product in a 93% yield (8.7 g)

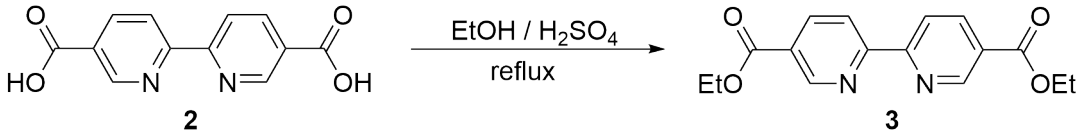
^{1}H-NMR (500 MHz, DMSO-d_6, δ): 13.50 (br); 9.18 (dd, ^{4}J = 2.15, 5J = 0.8 Hz 2H); 8.55 (dd, ^{3}J= 8.3; 5J = 0.8 Hz, 2H); 8.44 (dd, ^{3}J = 8.3; 4J= 2.15 Hz, 2H).

^{13}C-NMR (125 MHz, DMSO-d_6, δ): 165.49, 156.84, 149.82, 137.97, 126.65, 120.62.

MALDI-TOF (m/z) $[M+H]^+$ calculated for $[C_{12}H_8N_2O_4]$ 245.0; found 245.0.

USC DE COMPOSTELA

2.13.2. Synthesis of diethyl [2,2'-bipyridine]-5,5'-dicarboxylate (3)



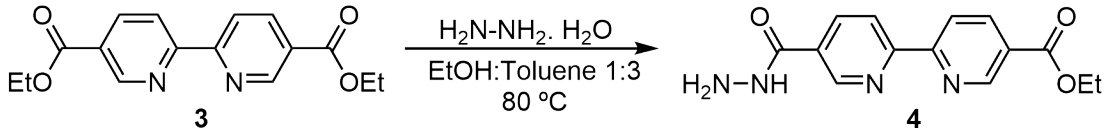
2,2'-bipyridine-5,5'-dicarboxylic acid (10.0 g, 41 mmol) was suspended in 150 mL of absolute ethanol. Concentrated sulfuric acid (20.0 mL) was slowly added, and the resulting mixture was refluxed for 18 h. The solution was cooled at room temperature and added over 400 mL of water at 4 ºC causing the precipitation of a white solid, which was filtered, washed with water and lyophilized. 11.5 g (93.0 %)

^{1}H-NMR (500 MHz, DMSO-d_6, δ): 9.20 (dd, ^{4}J = 2.15, 5J = 0.8 Hz 2H); 8.57 (dd, ^{3}J= 8.3; ^{5}J = 0.8 Hz, 2H); 8.46 (dd, ^{3}J = 8.3; ^{4}J = 2.15 Hz, 2H); 4.0 (q, ^{3}J = 7.1 Hz, 4H); 1.37 (t, ^{3}J = 7.1 Hz, 6H).

^{13}C-NMR (125 MHz, DMSO-d_6, δ): 164.2, 157.2, 149.8, 138.0, 126.2, 121.0, 61.1, 13.9.

MALDI-TOF (m/z) $[M+H]^+$ calculated for $[C_{16}H_{16}N_2O_4]$ 301.1; found 301.1.

2.13.3. Synthesis of Ethyl 5'-Carbohydrazido-2,2'-bipyridine-5-carboxylate (4)



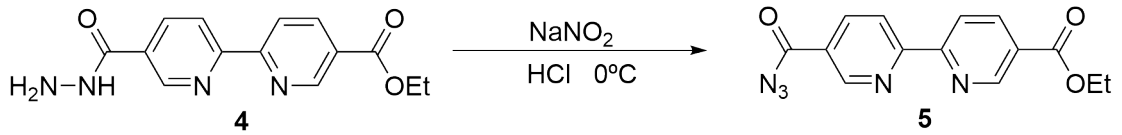
A mixture of diethyl 2,2'-bipyridine-5,5'-dicarboxylate (15.00 g, 50 mmol) and hydrazine hydrate (3.75 mL, 55 mmol) in a solution of EtOH (42 mL) and toluene (128 mL) was heated at 80 °C for 48 h. The precipitate was filtered, washed with $CHCl_3$ and dried under vacuum 11.4 g (80%). The unreacted diethyl ester was concentrated and mixed again to obtain a global yield of (90%).

^{1}H-NMR (500 MHz, DMSO-d_6, δ): 1.35 (t, 3H), 4.37 (q, 2H), 8.35 (d, ^{3}J = 8.4 Hz, 1H), 8.45 (dd, ^{3}J = 8.4 Hz, ^{4}J) 2 Hz, 1H), 8.52 (d, ^{3}J = 8.4 Hz, 1H), 8.57 (d, ^{3}J = 8.4 Hz, 1H), 9.1 (d, ^{4}J = 2 Hz, 1H), 9.2 (d, ^{4}J = 2 Hz, 1H), 10.1 (br, 1H).

^{13}C-NMR (125 MHz, DMSO-d_6,δ): 14.1, 61.3, 120.9, 120.9, 126.1, 129.6, 136.2, 138.2, 148.1, 150.0, 155.8, 157.8, 163.8, 164.5.

MALDI-TOF (m/z) $[M+H]^+$ calculated for $[C_{14}H_{14}N_4O_3]$ 287.1; found 287.1.

2.13.4. Synthesis of Ethyl 5'-Carbohydrazido-2,2'-bipyridine-5-carboxylate (5)



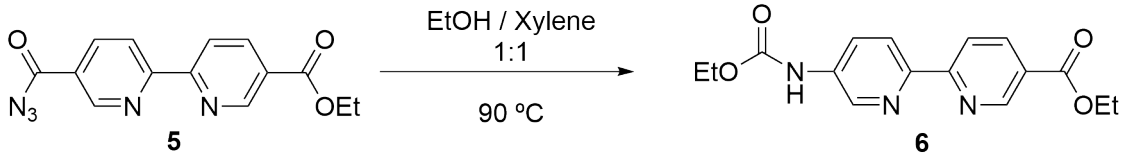
A stirred solution of 5-(ethoxycarbonyl)-5'-carbohydrazido-2,2'-bipyridine (5.7 g, 20 mmol) in concentrated HCl (100 mL) was cooled to 0 °C, and then an aqueous solution of $NaNO_2$ (1.73 g, 25 mmol; 15 mL) was added dropwise, maintaining the temperature at 0 °C. After 60 min, the yellow solution was diluted with water (300 mL) to precipitate the monoester 5 as a white powder, which was filtered, washed with water, and lyophilized 5.5 g (92 %).

^{1}H-NMR (500 MHz, Acetone-d^{6},δ): 9.13 (d, ^{4}J = 1.91 Hz, 1H), 9.10 (d, ^{4}J = 1.91 Hz, 1H), 8.71 (d, ^{3}J = 12.5 Hz, 1H), 8.69 (d, ^{3}J = 12.5 Hz, 1H), 8.517 (m, 2H; 4.31 (q, ^{3}J = 7.15 Hz, 2H); 1.29 (t, ^{3}J= 7.15, 3H).

^{13}C-NMR (125 MHz, Acetone-d_6,δ): 170.8, 164.5, 159.2, 157.7, 150.3, 150.0, 138.0, 137.9, 127.0, 121.3, 61.2, 13.6.

MALDI-TOF (m/z) $[M+H]^+$ calculated for $[C_{14}H_{11}N_5O_3]$ 298.1; found 298.1

2.13.5. Synthesis of Ethyl 5'-(ethoxycarbonyl)amino-2,2'-bipyridine-5-carboxylate (6)



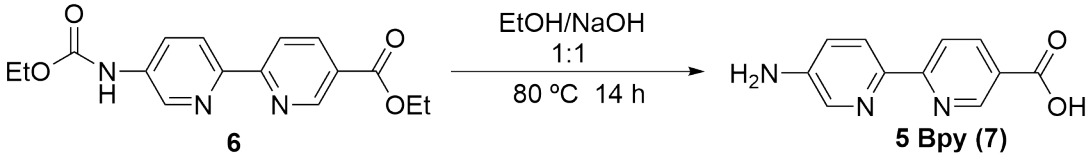
A solution of ethyl 5'-carbazido-2,2'-bipyridine-5-carboxylate (8.6 g, 29 mmol) in a mixture of EtOH (100 mL) and xylene (100 mL) was heated at 90 ºC for 4 h. The solvent was evaporated under reduced pressure, and the yellow residue was washed with EtOH and dried in vacuo. 8.21 g (90%).

^{1}H-NMR (500 MHz, DMSO-d_6, δ): 1.27 (t, J = 7.2 Hz, 3H), 1.35 (t, J = 7.2 Hz, 3H), 4.17 (q, J = 7.2 Hz, 2H), 4.36 (q, J = 7.2Hz, 2H), 8.08 (dd, ^{3}J = 8 Hz, ^{4}J = 2 Hz, 1H), 8.35 (dd, ^{3}J = 8.3 Hz, ^{4}J = 2 Hz, 1H), 8.40 (dd, ^{3}J = 8.3Hz, ^{5}J = 1 Hz, 1H), 8.45 (d, ^{3}J = 8 Hz, 1H), 8.70 (d, ^{4}J = 2 Hz, 1H), 9.12 (d, ^{4}J = 2 Hz, ^{5}J = 1 Hz, 1H), 10.13 (s, 1H).

^{13}C-NMR (125 MHz, DMSO-d_6, δ): 13.6, 13.9, 60.3, 60.6, 119.1, 121.2, 124.46, 124.9, 136.7, 137.4, 139.0, 147.4, 149.4, 153.1, 158.1, 164.2.

MALDI-TOF (m/z) $[M+H]^+$ calculated for $[C_{16}H_{17}N_3O_4]$ 316.1; found 316.1.

Synthesis of 5'-Amino-2,2'-bipyridine-5-carboxylic acid hydrochloride (7)



A stirred solution of ethyl 5'-[(ethoxycarbonyl)amino]-2,2'-bipyridine-5-carboxylate (14.55 g, 45.9 mmol) in a mixture of EtOH (50 mL) and 2.5 M aqueous NaOH (50 mL) was heated at 75 ºC for 14 h. The EtOH was concentrated in vacuum, and the aqueous solution was acidified with HCl to afford a bright yellow precipitate, which was filtered, washed with cold water, and lyophilized 9.80 g (85%).

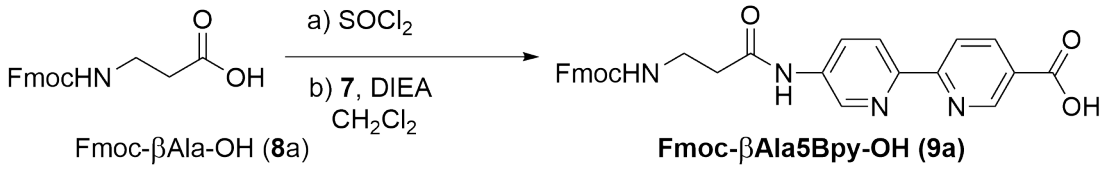
^{1}H-NMR (500 MHz, D_2O, δ): 7.00 (dd, ^{3}J = 8.5 Hz, ^{4}J = 2.8 Hz, 1H), 7.55, 7.61 (d, ^{3}J = 8.5 Hz, 2H), 7.81 (d, ^{4}J = 2.9 Hz, 1H), 7.98 (dd, ^{3}J = 8.5 Hz, ^{4}J = 2.9 Hz, 1H), 8.65 (d, ^{4}J = 2.8 Hz, 1H)

^{13}C-NMR (125 MHz, D_2O, δ): 123.3d, 126.2, 126.7, 133.8, 139.8, 141.34, 147.5, 147.4, 152.4, 159.6, 176.04

MALDI-TOF (m/z) $[M+H]^+$ calculated for $[C_{11}H_9N_3O_2]$ 216.1; found 216.1.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

2.13.6. Synthesis of Fmoc-βAla5Bpy-OH (9a)



Over 4.58 g (14.72 mmol) of Fmoc-βAla 10 mL of $SOCl_2$ was added carefully at room temperature. The solution was stirred magnetically for 30 min, the thionyl chloride was evaporated in vacuum and the yellow solid was washed with CH_2Cl_2 (3 × 10 mL), and dried under reduced pressure. The compound (**7**) 3.7 g (14.8 mmol), 20 mL of CH_2Cl_2 and 5 mL of DIEA were added. The suspension was stirred at room temperature overnight. The solvent was evaporated, the solid was suspended in acetonitrile 0.1 %TFA/H_2O (2:1), centrifuged and washed with acetonitrile and H_2O and lyophilized 6.70 g (89 %).

^{1}H–RMN (300 MHz, DMSO-d_6, δ): 10.44 (s, 1H); 9.13 (d, 4J = 1.36 Hz, 1H); 8.89 (d, ^4J = 2.12 Hz, 1H); 8.42 (m, 3H); 8.25(dd, ^3J = 8.48 Hz, ^4J = 2.12 Hz, 1H); 7.88 (d, ^3J = 7.46 Hz, 2H); 7.67 (d, ^3J = 7.38 Hz, 2H); 7.49 (t, ^3J = 5.34 Hz, NH); 7.40 (t, ^3J = 7.21 Hz, 2H); 7.30 (t, ^3J = 7.38 Hz, 2H); 4.30 (d, ^3J = 6.95 Hz, 2H); 4.21 (t, ^3J = 6.95 Hz, 1H) 3.32 (t, ^3J = 6.70 Hz, 2H); 2.60 (t, ^3J = 6.70 Hz, 2H).

^{13}C-NMR (75 MHz, DMSO-d_6, δ): 169.76 (C), 168.62 (C), 157.23 (C), 150.94 (C), 149.21 (C), 144.74 (C), 143.47 (C), 142.27 (C), 141.59 (C), 140.28 (CH), 138.29 (CH), 138.08 (CH), 129.73 (CH), 128.09 (CH), 125.96 (CH), 122.14 (CH), 120.87 (CH), 120.79 (CH), 118.73 (CH), 110.34 (CH), 67.38 (CH_2), 63.77 (CH_2), 47.62 (CH), 41.71 (CH_2).

MALDI-TOF (m/z) $[M+H]^+$ calculated for $[C_{29}H_{25}N_4O_5]$ 509.1; found 509.1.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Chapter I

1. UHPLC-MS of the purified peptide ligands

LLD: H-(βAla5Bpy)$_2$RPr(βAla5Bpy)$_2$RPr(βAla5Bpy)$_2$-NH$_2$

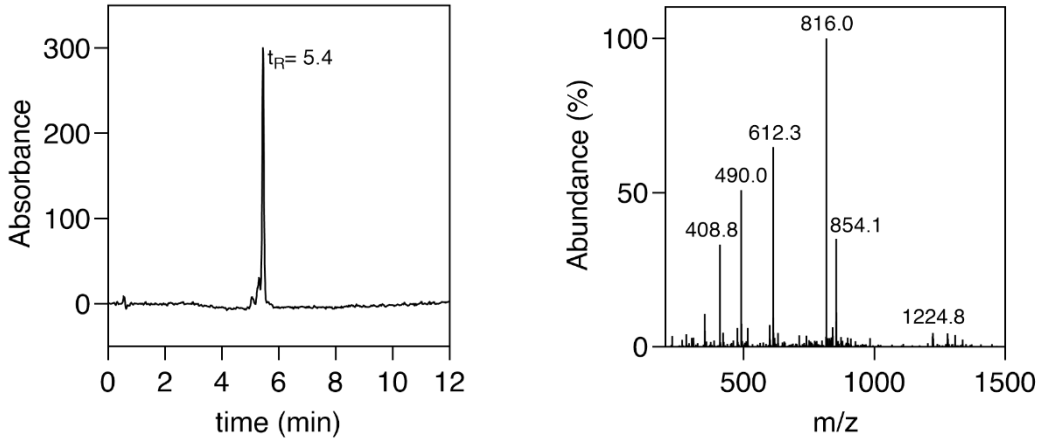


Figure 61. Left: chromatogram of the purified **LLD** ligand. Right: mass spectra corresponding to the peak with t_R of 5.4 min.

HPLC-MS (ESI): (5-95% B, t_R= 5.4 min) Calculated for $C_{118}H_{137}N_{43}O_{18}$ = 2444.11; found $[M+2H+TFAH]^{2+}$ = 1289.7; $[M+3H]^{3+}$ = 816.0; $[M+3H+TFAH]^{3+}$ = 854.1; $[M+4H]^{4+}$ = 612.3; $[M+5H]^{5+}$ = 490.0; $[M+6H]^{6+}$ = 408.8.

MS (MALDI-TOF): Calculated for $C_{118}H_{137}N_{43}O_{18}$ = 2444.11; found: $[M+H]^+$ = 2446.21.

DDL: H-(βAla5Bpy)$_2$rpR(βAla5Bpy)$_2$rpR(βAla5Bpy)$_2$-NH$_2$

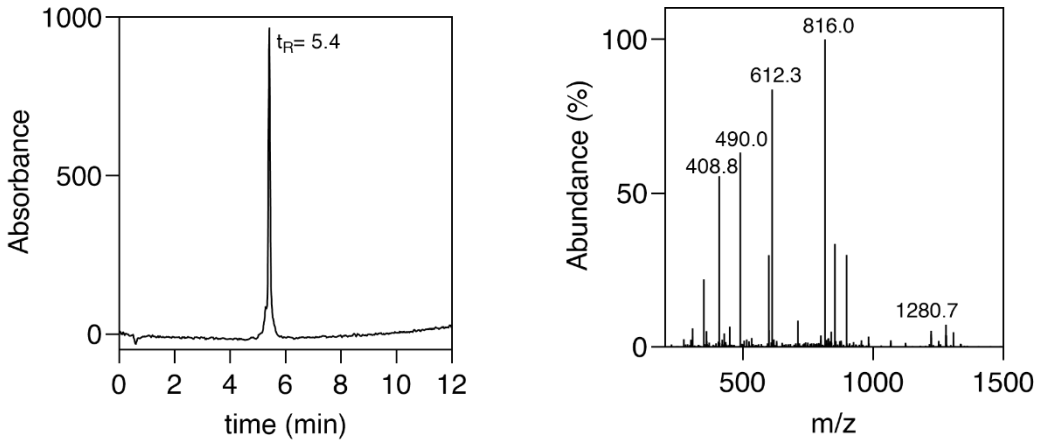


Figure 62. Left: chromatogram of the purified **DDL** ligand. Right: mass spectra corresponding to the peak with t_R of 5.4 min.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

HPLC-MS (ESI): (5-95% B, t_R= 5.4 min) Calculated for $C_{118}H_{137}N_{43}O_{18}$ = 2444.11; found $[M+2H+TFAH]^{2+}$ = 1289.7; $[M+3H]^{3+}$ = 816.0; $[M+3H+TFAH]^{3+}$ = 854.1; $[M+4H]^{4+}$ = 612.3; $[M+5H]^{5+}$ = 490.0; $[M+6H]^{6+}$ = 408.8.

MS (MALDI-TOF): Calculated for $C_{118}H_{137}N_{43}O_{18}$ = 2444.11; found: $[M+H]^+$ = 2446.21.

TAMRA-LLD: TAMRA-O1Pen(βAla5Bpy)$_2$RPr(βAla5Bpy)$_2$RPr(βAla5Bpy)$_2$-NH_2

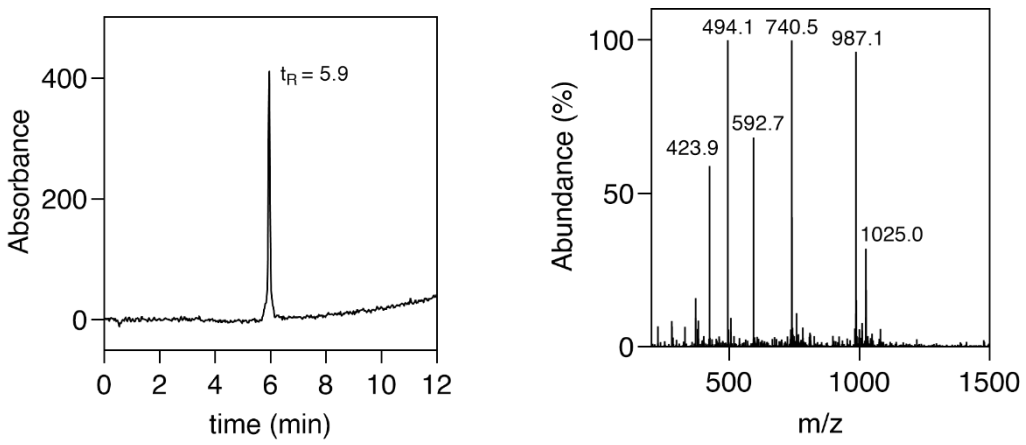


Figure 63. Left: chromatogram of the purified **TAMRA-LLD** ligand. Right: mass spectra corresponding to the peak with tR of 5.9 min.

HPLC-MS (ESI): (5-95% B, t_R= 5.9 min) Calculated for $C_{147}H_{166}N_{46}O_{24}$ = 2959.31; found $[M+3H+TFAH]^{3+}$ = 1025.0; $[M+3H]^{3+}$ = 987.1; $[M+4H]^{4+}$ = 740.5; $[M+5H]^{5+}$ = 592.7; $[M+6H]^{6+}$ = 494.1; $[M+7H]^{7+}$ = 423.9.

MS (MALDI-TOF): Calculated for $C_{147}H_{166}N_{46}O_{24}$ = 2959.31; found: $[M+H]^+$ = 2960.43.

2. Synthesis of the peptide helicates

2.1.Synthesis of the Fe^{II} peptide helicates

The synthesis of **ΛΛ-Fe^{II}_2LLD** and **ΛΛ-Fe^{II}_2TAMRA-LLD** metallopeptides were made following the procedure reported in our previous paper.[190c]

2.2.Synthesis of the Cu^{II} peptide helicates

Onto a solution of **LLD**, **DDL** or **TAMRA-LLD** in phosphate buffer (1 mM, 10 mM NaCl, pH 7.0), $CuCl_2$. • 2 H_2O was added as source of Cu^{II} ions until saturating conditions (5 eq).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

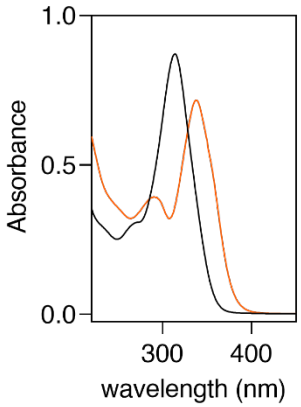


Figure 64. UV-VIS of a 5 µM solution of **LLD** ligand (1 mM phosphate buffer, 10 mM NaCl, pH 7.0) before (black line) and after the addition of 5 equivalents of $CuCl_2 \cdot 2\ H_2O$ (orange line).

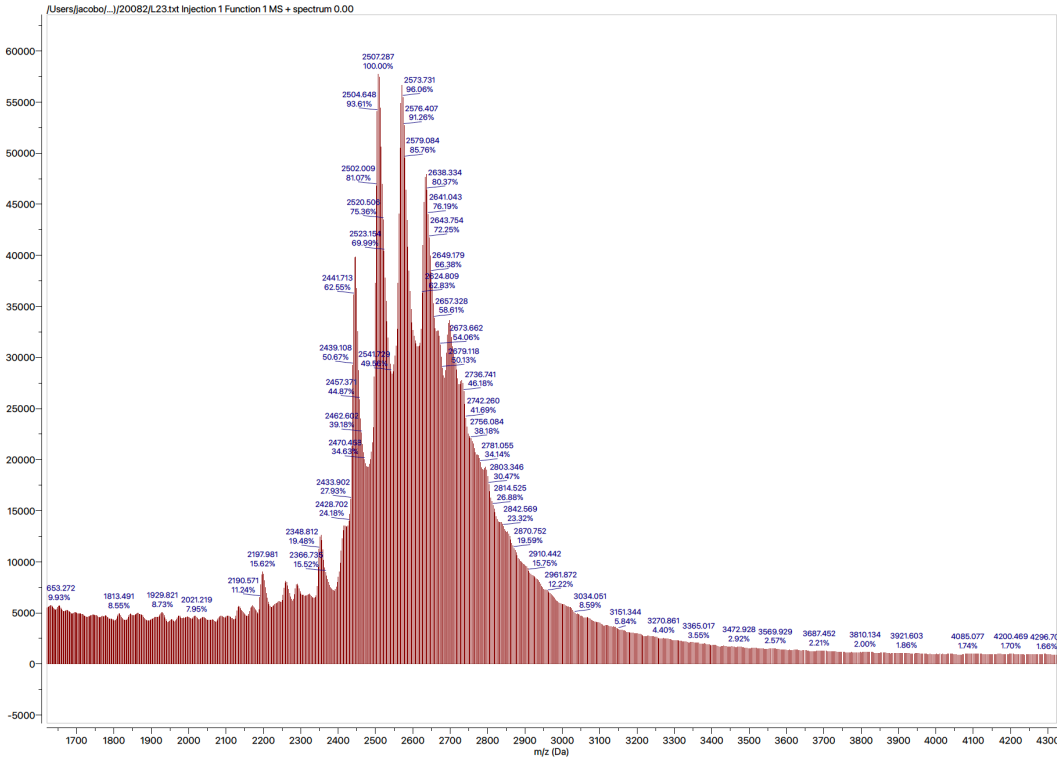


Figure 65. MALDI spectra of **ΔΔ-Cu^{II}_2LLD**.

MS (MALDI-TOF): Calculated for **ΔΔ-Cu^{II}_2LLD** $C_{118}Cu_2H_{137}N_{43}O_{18}$ = 2572.77; found: $[M+2Cu+H]^+$ = 2573.731.

MS (MALDI-TOF): Calculated for **ΛΛ-Cu^{II}_2DDL** $C_{118}Cu_2H_{137}N_{43}O_{18}$ = 2572.77; found: $[M+2Cu+H]^+$ = 2573.73.

MS (MALDI-TOF): Calculated for **ΔΔ-Cu^{II}_2TAMRA-LLD** $C_{147}Cu_2H_{166}N_{46}O_{24}$ = 3086.32; found: [M+2Cu] = 3086.75; $[M+Cu+H]^+$ = 3022.63; $[M+H]^+$ = 2959.13.

3. EPR studies

These experiments were carried out in collaboration with the **Prof. Francisco Rivadulla** from CiQUS (Universidade de Santiago de Compostela, Spain). X-band EPR

spectra of the **ΔΔ-Cu$_2^{II}$LLD** metallopeptide were obtained in solution [200 μM of metallopeptide in phosphate buffer (1 mM, 10 mM NaCl, pH 7.0)] at 300 K and 120 K.

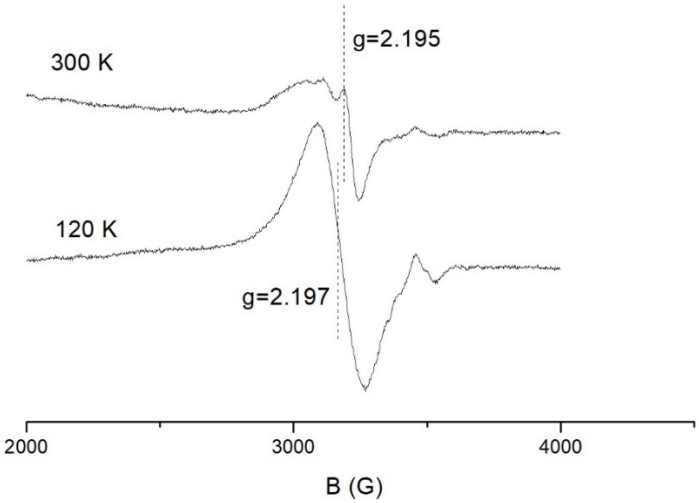


Figure 66. X-band Cu^{II} EPR spectra of the **ΔΔ-Cu$^{II}_2$LLD** peptide helicate measured in solution (200 μM of metallopeptide in phosphate buffer 1mM, 10 mM NaCl, pH 7.0) at 120 K (bottom) and 300 K (top).

4. Fluorescence titrations experiments with DNA

4.1. Fluorescein-labelled dsDNA with ΔΔ-Cu$^{II}_2$LLD

A fluorescence titration was performed (in triplicate) in which, on a 2 μM solution of fluorescein-labelled dsDNA (Flu-dsDNA) aliquots of a mother solution of **ΔΔ-Cu$^{II}_2$LLD** (300 μM) were added. The K_D of the interaction could not be calculated as a precipitation process occurs in the titration cuvette during the experiment. This result demonstrates the absence of specific interaction between the peptide helicate and dsDNA. Conditions: phosphate buffer 1 mM, NaCl 10 mM, pH 7.0.

4.2. Fluorescein-labelled GC-3WJ with ΔΔ-Cu$^{II}_2$LLD

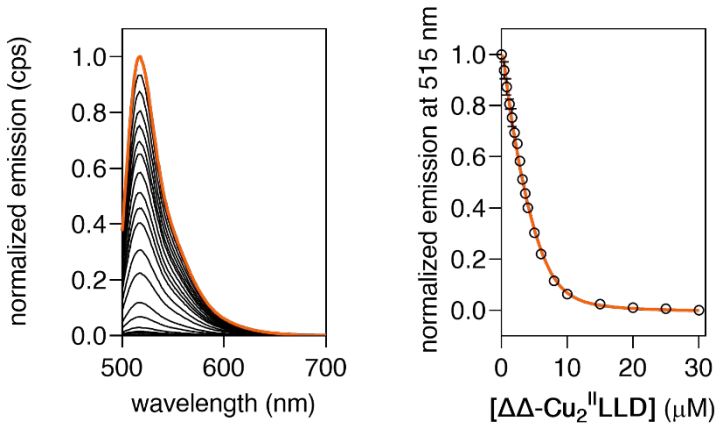


Figure 67. Left: Full emission spectra of a 2 μM solution of the fluorescein-labelled GC-3WJ (Flu-GC-3WJ) in absence (orange line) and in the presence of increasing concentrations of **ΔΔ-Cu$^{II}_2$LLD** (black lines of decreasing intensity). Right: Profile of the full titration with the best fit (orange line) according to the 1:1 model plus unspecific interaction in *DynaFit*, with an apparent dissociation constant K_D = 300 nM. Conditions: 1mM phosphate buffer, NaCl 10 mM, pH 7.0.

USC DE COMPOSTELA

4.3. Fluorescein-labelled 3WJ with ΛΛ-Cu^{II}_2DDL

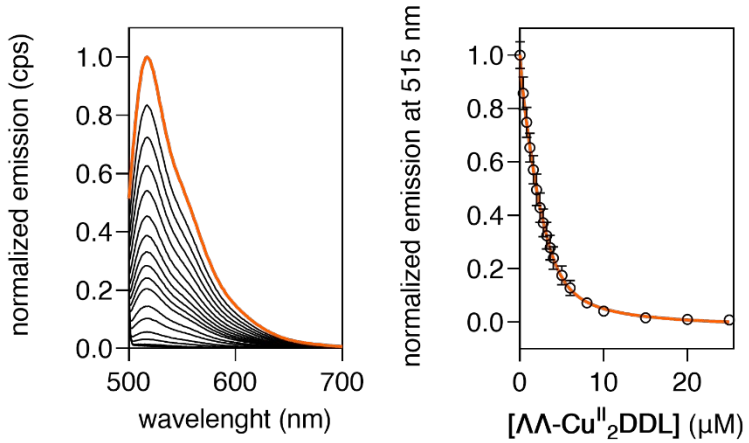


Figure 68. Left: Full emission spectra of a 2 μM solution of the fluorescein-labelled 3WJ (Flu-3WJ) in absence (orange line) and in the presence of increasing concentrations of **ΛΛ-Cu^{II}_2DDL** (black lines of decreasing intensity). Right: Profile of the full titration with the best fit (orange line) according to the 1:1 model plus unspecific interaction in *DynaFit*, with an apparent dissociation constant K_D = 654 nM. Conditions: 1mM phosphate buffer, NaCl 10 mM, pH 7.0.

5. Studies of the kinetic inertness of ΔΔ-Cu^{II}_2LLD in physiological media

The fluorescence emission at 420 nm of a 2 μM solution of **ΔΔ-Cu^{II}_2LLD** and 4 μM of $EDTA^{260}$ in phosphate buffer (1 mM, 10 mM NaCl, pH 7.0) was monitored during 60 min. The resulting complex half time was $t_{1/2}$ = 5.815 minutes and was calculated using Prism 9 *software* (by Dotmatics). The same experiment was repeated in the presence of 0.267 μM of 3WJ. The resulting complex half time was $t_{1/2}$ = 23.12 minutes and was also calculated using *Prism 9* software (by Dotmatics).

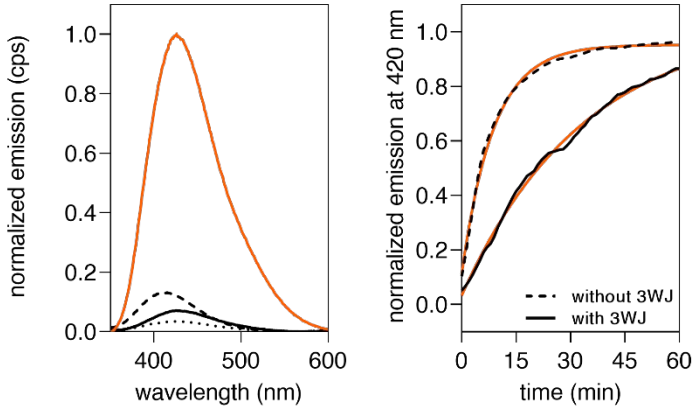


Figure 69. Left: full emission spectra of a 2 μM solution of **LLD** (orange line), and spectra of the same solution with 10 μM of Cu^{II} ions (black continuous line), and with of 0.267 μM of 3WJ (black dotted line), and with 4 μM of EDTA after 60 min of stabilization (black dashed line). Right: fluorescence emission at 420 nm of a 2 μM solution of **ΔΔ-Cu^{II}_2LLD** and 4 μM of EDTA without (black dashed line) and with (black continuous line) the presence of 0.267 μM of 3WJ for 60 minutes, together with their respective best fits (orange lines) resulting in a complex half time $t_{1/2}$ = 5.815 and 23.12 minutes, respectively. Conditions: phosphate buffer 1 mM, NaCl 10 mM, pH 7.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

6. 3WJ-binding properties of the free peptide ligand LLD

The key role played by metal ions in the recognition of 3WJ by these family of peptide helicates was demonstrated in our previous wotk[190c] through a fluorescence anisotropy titration of a 2 μM solution of a fluorescein-labelled **LLD** peptide ligand (**Flu-LLD**) with increasing concentrations of 3WJ. The descending and oscillating pattern observed in the titration showed the lack of interaction between the ligand and the 3WJ.

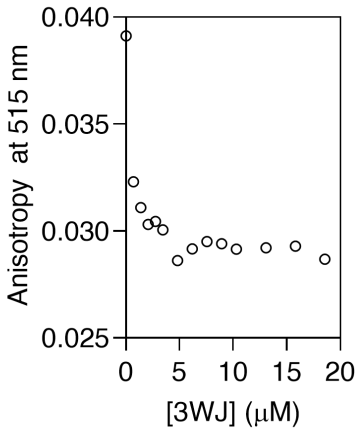


Figure 70. Anisotropy signal of a 2 μM solution of **Flu-LLD** in the presence of increasing concentrations of 3WJ. The absence of interaction between the two systems is shown to be evident due to the decreasing and oscillating pattern of the points on the graph.

7. Gel electrophoresis: DNA binding studies

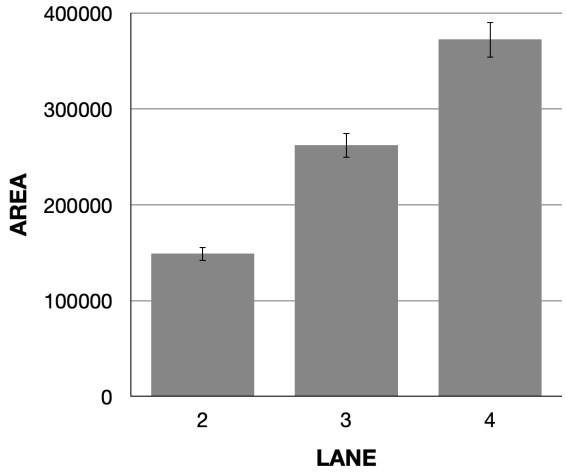


Figure 71. The area of each band in the gel of Figure 36 of the manuscript was measured using the *Fiji ImageJ implementation*[311] using the Gel Analysis method outlined in the ImageJ documentation (https://imagej.nih.gov/ij/docs/menus/analyze.html#gels). The area of each band clearly shows a saturation profile and concentration dependent signal, consistent with the progressive saturation of 3WJ binding with the Cu^{II} helicate

[311] a) C. A. Schneider, W. S. Rasband, K. W. Eliceiri, *Nat. Methods* **2012**, *9*, 671–675; b) J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, *Nat. Methods* **2012**, *9*, 676–682.

8. *Gel electrophoresis: In vitro nuclease activity studies*

These experiments were carried out in collaboration with the **Prof. Andrew Kellett** from the School of Chemical Sciences at Dublin City University (DCU).

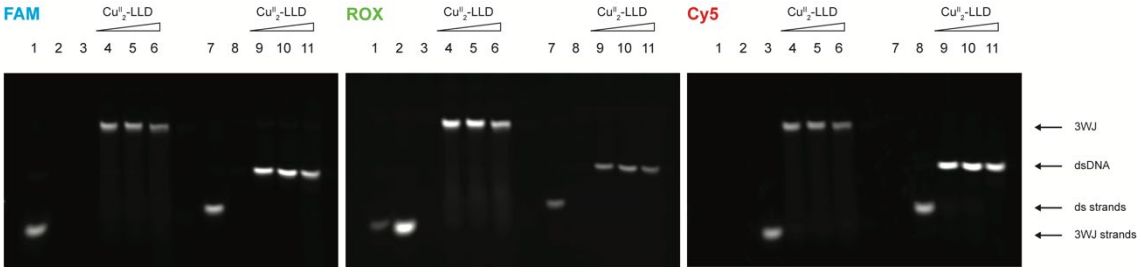


Figure 72. Images obtained during multiplex assay with FAM (left), ROX (middle) and Cy5 (right) excitation and emission filters.

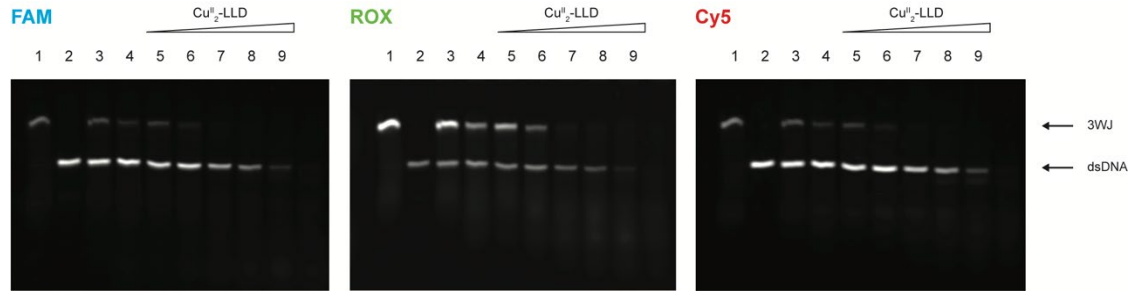


Figure 73. FAM (left), ROX (middle) and Cy5 (right) images for fluorophore cleavage assay.

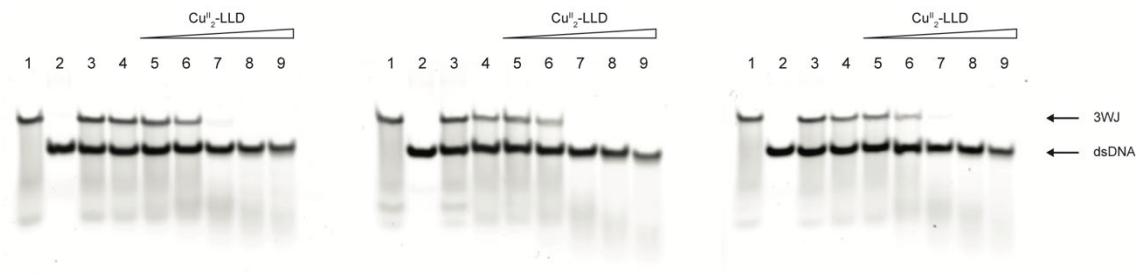


Figure 74. 20% native page of cleavage inflicted by **ΔΔ-Cu$^{II}_2$LLD** in triplicate for densitometry. Lane 1; 3WJ. Lane 2; dsDNA. Lane 3; 3WJ and dsDNA incubated with **ΔΔ-Cu$^{II}_2$LLD**. Lane 4: Target and off-target with **LLD** (no Cu^{II} present). Lanes 5-9; 3WJ and dsDNA incubated with 1, 2, 5, 7.5 or 10 eq. of **ΔΔ-Cu$^{II}_2$LLD** and Na-(L)-ascorbate (10 eq. to **ΔΔ-Cu$^{II}_2$LLD)**.

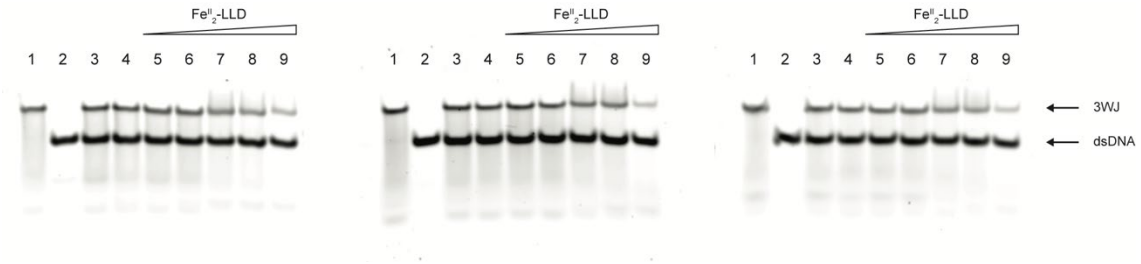


Figure 75. ΛΛ-Fe$^{II}_2$LLD cleavage for densitometry. **ΛΛ-Fe$^{II}_2$LLD** cleavage assay was performed in triplicate for band densitometry analysis. Lane 1; 3WJ. Lane 2; dsDNA. Lane 3; Target 3WJ and off-target dsDNA incubated with **ΛΛ-Fe$^{II}_2$LLD**. Lane 4; 3WJ and dsDNA with **LLD** (no Fe^{II} present). Lanes 5-9; Target 3WJ and off-target dsDNA with 1, 2, 5, 7.5 or 10 eq. of **ΛΛ-Fe$^{II}_2$LLD** and Na-(L)-ascorbate (10 eq. to **ΛΛ-Fe$^{II}_2$LLD**). Gel on the left is also shown in Figure 35c.

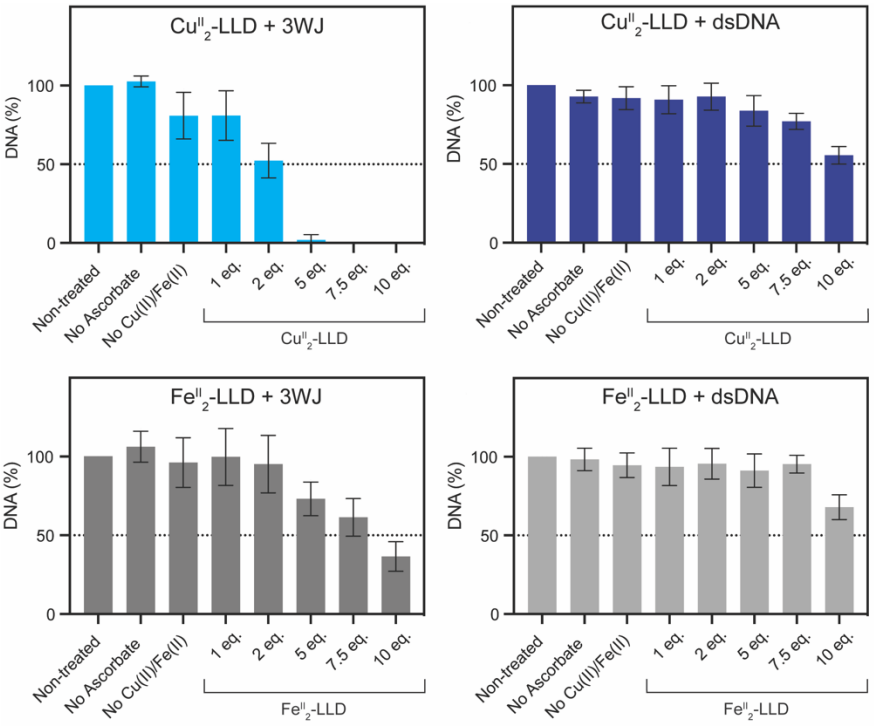


Figure 76. Band densitometry analysis of cleavage inflicted by **ΔΔ-Cu$^{II}_2$LLD** and **ΛΛ-Fe$^{II}_2$LLD** on target 3WJ and off-target dsDNA. Densitometry was performed in triplicate, average DNA remaining, and standard deviation were calculated and plotted. This data is also shown in Figure 35d.

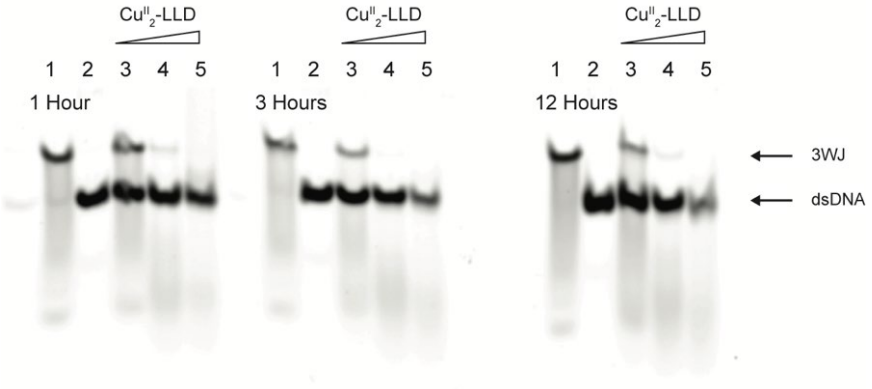


Figure 77. PAGE analysis of cleavage profile at 1, 3 and 12 hours. Lane 1; Target 3WJ. Lane 2; off-target dsDNA. Lane 3-5; 3WJ and dsDNA with 1, 5 or 10 eq. **ΔΔ-Cu$^{II}_2$LLD** and Na-(L)-ascorbate (10 eq. to **ΔΔ-Cu$^{II}_2$LLD**).

9. Cell-internalization studies

These experiments were carried out in collaboration with the **Dr. Natalia Barreiro Piñeiro** and **Prof. José Martínez Costas** from CiQUS.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

9.1. Cell-internalization of the peptide ligand TAMRA-LLD in Vero cells

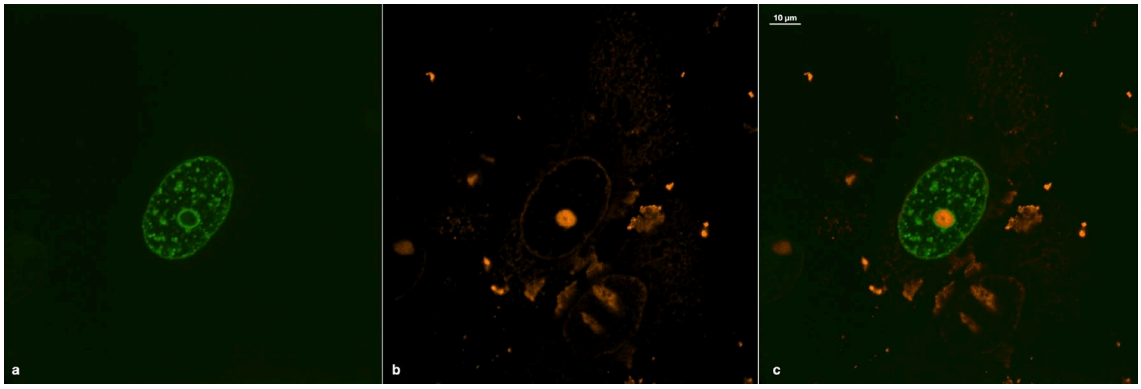


Figure 78. Free peptide ligand **TAMRA-LLD** alone enters the cell nucleus but cannot recognize DNA replication sites. Vero cells expressing protein GFP-PCNAL2 were incubated in the presence of 25 µg/mL of Digitonin for 3 min and then in the presence of 5 µM of **TAMRA-LLD** for 15 min (these reported concentrations of digitonin, and peptide ligand refer to the final amount of these compounds in the cell culture transfer buffer after their addition from more concentrated stock solutions). a) Green channel, corresponding to the emission of the GFP-PCNAL2 probe labelling the DNA replication foci; b) red channel emission showing the distribution of the rhodamine-labelled peptide ligand; c) overlay of the green and red channels. Nucleoli are clearly labelled as big red spots inside the cell nuclei. The images show a representative confocal section of the Vero cells.

9.2. Cell-internalization of the Cu^{II} peptide helicate in HeLa cells

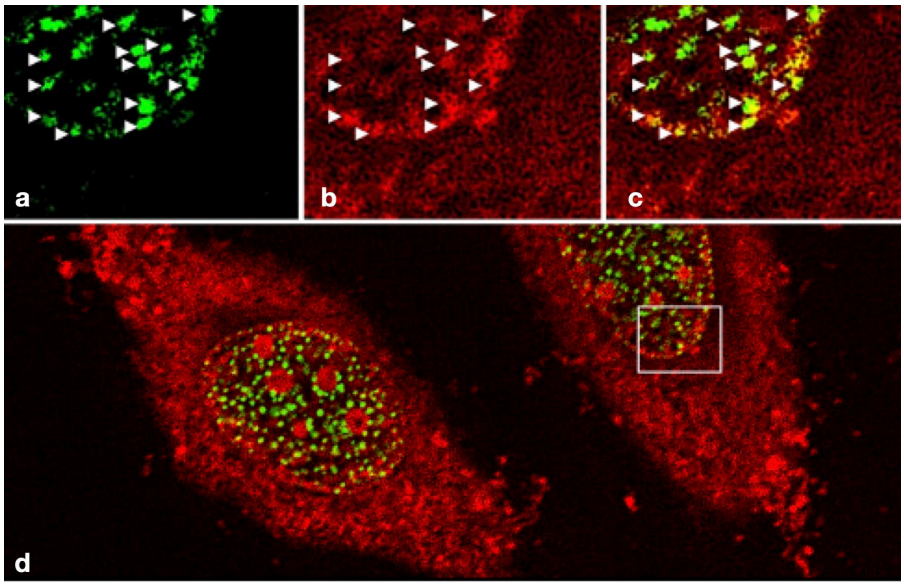


Figure 79. ΔΔ-Cu^{II}_2TAMRA-LLD selectively stains DNA replication sites. HeLa cells expressing protein GFP-PCNAL2 were incubated in the presence of 25 µg/mL of digitonin for 3 min and then in the presence of 5 µM of **ΔΔ-Cu^{II}_2TAMRA-LLD** for 15 min (these reported concentrations of digitonin and helicate refer to the final amount of these compounds in the cell culture transfer buffer after their addition from more concentrated stock solutions). The Cu^{II} peptide helicate used in these studies was pre-formed in an independent Eppendorf vial just before the internalization experiment by adding 5 equivalents of $CuCl_2.2H_2O$ to an aqueous solution of **TAMRA-LLD** peptide ligand. a) Green channel, corresponding to the emission of the GFP-PCNAL2 probe labeling the DNA replication foci; b) red channel emission showing the distribution of the rhodamine-labelled helicate; c) overlay of the green and red channels of the square region shown in d). Arrows highlight some foci where the staining of GFP-PCNAL2 and **ΔΔ-Cu^{II}_2TAMRA-LLD** overlap. Nucleoli are clearly labelled as big red spots inside the cell nuclei. The images show a representative confocal section of the HeLa cells.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Chapter II

1. UHPLC-MS of the purified peptide ligand

L:

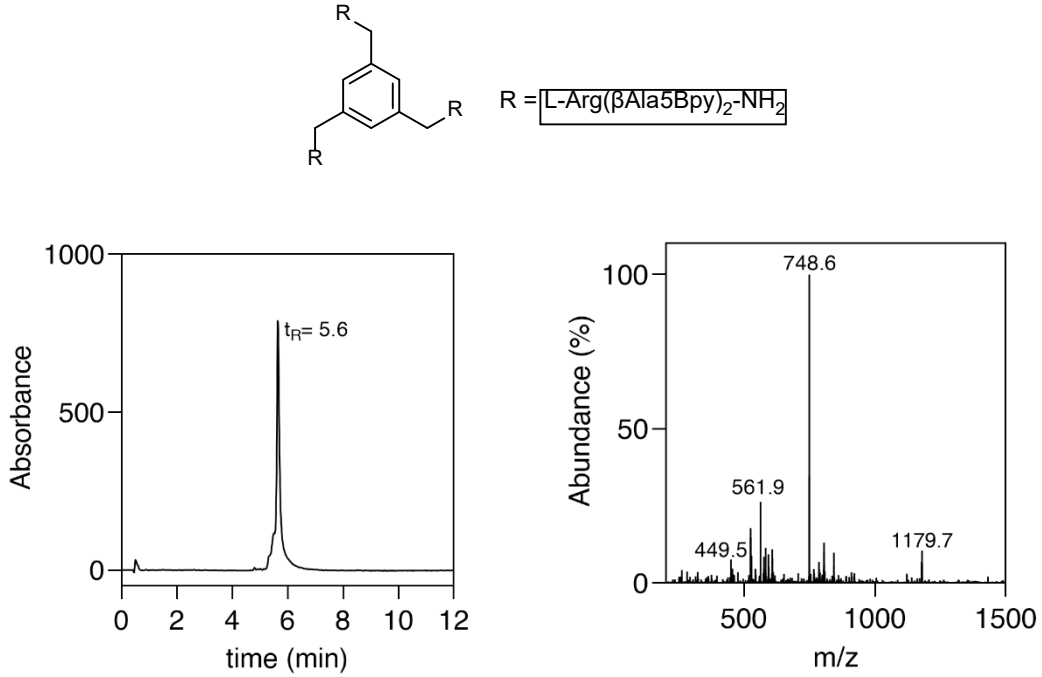


Figure 80. Left: chromatogram of the purified **L** ligand. Right: mass spectra corresponding to the peak with t_R of 5.6 min.

HPLC-MS (ESI): (5-95% B, t_R= 5.6 min) Calculated for $C_{111}H_{123}N_{39}O_{15}$ = 2242.01; found $[M+2H+TFAH]^{2+}$ = 1179.7; $[M+3H]^{3+}$ = 748.6; $[M+4H]^{4+}$ = 561.9; $[M+5H]^{5+}$ = 449.5.

MS (MALDI-TOF): Calculated for $C_{111}H_{123}N_{39}O_{15}$ = 2242.01; found: $[M+H]^{+}$ = 2244.1.

2. Synthesis of the Co(III) peptide helicate

A 200 μM aqueous solution of **L** was incubated with 15 equivalents of $Co(ClO_4)_2 \cdot 6H_2O$ (3 mM). Thereafter, the sample was treated with 1.1 eq. of $(NH_4)_2Ce(NO_3)_6$ (3.3 mM) with respect to Co(II). After 30 min of reaction, the obtained kinetically inert Co(III)-metallopeptide was purified by semipreparative HPLC in reverse phase (0-40% of AcN + 0.1% TFA in 30 min) and freeze dried. The resulting $\mathbf{Co^{III}_2L}$ helicate was then characterized by UHPLC-MS, MALDI-TOF mass spectrometry and circular dichroism.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

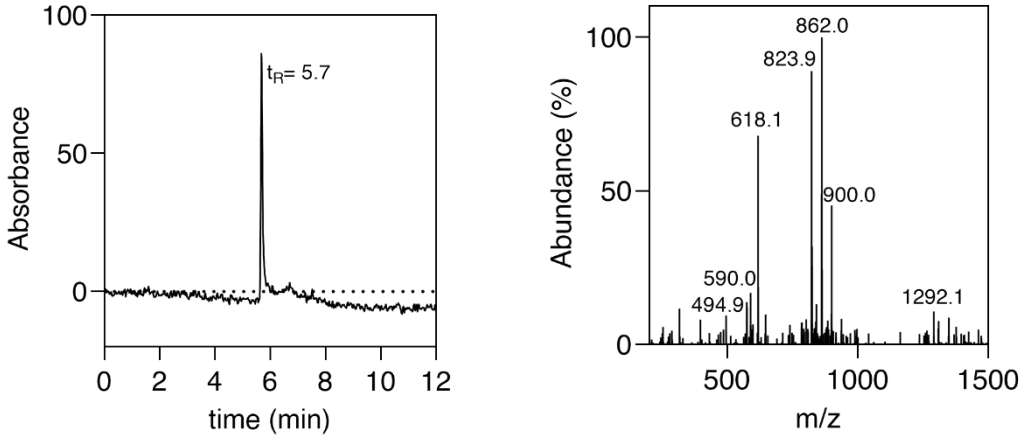


Figure 81. Left: chromatogram of the purified **Co^{III}_2L** peptide helicate. Right: mass spectra corresponding to the peak with t_R of 5.7 min.

HPLC-MS (ESI): (5-95% B, t_R= 5.7 min) Calculated for **ΛΛ-Co^{III}_2L** $C_{111}H_{123}N_{39}O_{15}Co_2$ = 2359.87; found $[M+2Co+2H+2TFAH]^{2+}$ = 1292.1; $[M+2Co+3H+3TFAH]^{3+}$ = 900.0; $[M+2Co+3H+2TFAH]^{2+}$ = 862.0; $[M+2Co+3H+TFAH]^{3+}$ = 823.9; $[M+2Co+4H+TFAH]^{4+}$ = 618.1; $[M+2Co+4H]^{4+}$ = 590.0; $[M+2Co+5H+TFAH]^{5+}$ = 494.9.

We measured the CD spectra of a 2 μM solution of **Co^{III}_2L** peptide helicate. The CD spectrum is dominated by a positive Cotton effect band centered at 336 nm, thus confirming the formation of the corresponding discrete **ΛΛ-Co^{III}_2L** helicate.

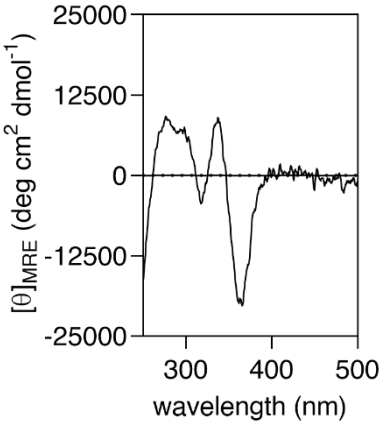


Figure 82. CD spectra of **Co^{III}_2L** at 2 μM. The measurement was performed in 10 mM HEPES buffer, 100 mM NaCl, pH 7.0 at 20 ºC.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

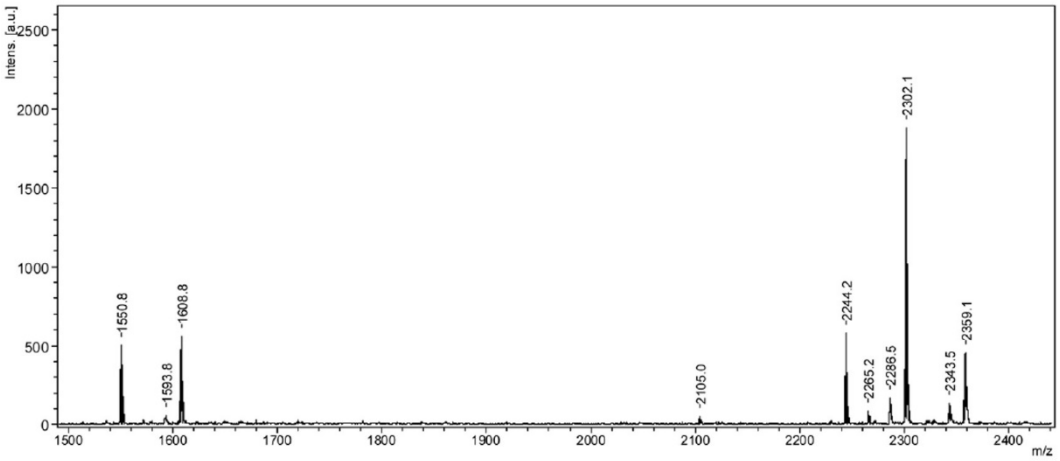


Figure 83. MALDI spectra of **ΛΛ-Co$^{III}_2$L**.

MS (MALDI-TOF): Calculated for **ΛΛ-Co$^{III}_2$L** $C_{111}H_{123}N_{39}O_{15}Co_2$ = 2359.87; found: $[M+2Co+H]^+$ = 2359.1; $[M+Co+H]^+$ = 2302.1; $[M+H]^+$ = 2244.2

3. Fluorescence titrations experiments with DNA

3.1. Fluorescein-labelled dsDNA with ΛΛ-Co$^{II}_2$L

A fluorescence titration was performed (in triplicate) in which, over a 2 μM solution of fluorescein-labelled dsDNA (Flu-dsDNA) aliquots of a mother solution of **ΛΛ-Co$^{II}_2$L** were added. The K_D of the interaction could not be calculated as a precipitation process occurs in the titration cuvette during the experiment. This result demonstrates the absence of specific interaction between the peptide helicate and dsDNA.

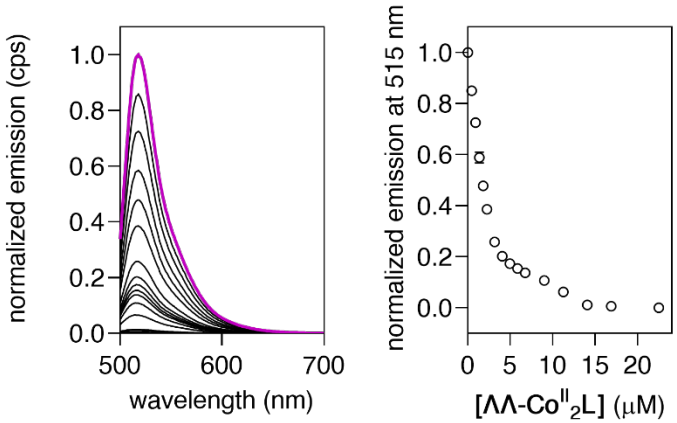


Figure 84. Left: Normalized emission spectra of the 2 μM solution of Flu-dsDNA in absence (purple line) and in the presence of increasing concentrations of **ΛΛ-Co$^{II}_2$L** (black lines of decreasing intensity); Right: profile at 515 nm of the same fluorescence titration experiment. Conditions: 10 mM HEPES buffer, 100 mM NaCl, pH 7.0; λ_{exc} = 490 nm

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

3.2. Fluorescein-labelled 3WJ with ΛΛ-Co^{III}_2L

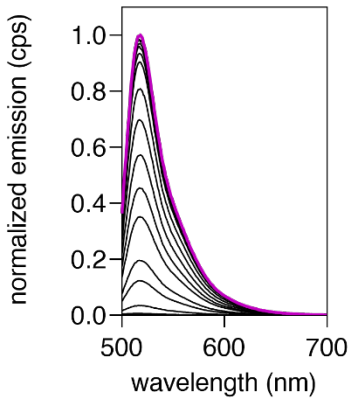


Figure 85. Full emission spectra of a 2 μM solution of the fluorescein-labelled 3WJ (Flu-3WJ) in absence (purple line) and in the presence of increasing concentrations of **ΛΛ-Co^{III}_2L** (black lines of decreasing intensity). Conditions: 1mM phosphate buffer, NaCl 10 mM, pH 7.0. Conditions: 10 mM HEPES buffer, 100 mM NaCl, pH 7.0; λ_{exc} = 490 nm.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Chapter III

1. UHPLC-MS of the purified peptide ligands

RGH-LLD:H-RGHKKK(βAla5Bpy)$_2$RPr(βAla5Bpy)$_2$RPr(βAla5Bpy)$_2$-NH$_2$

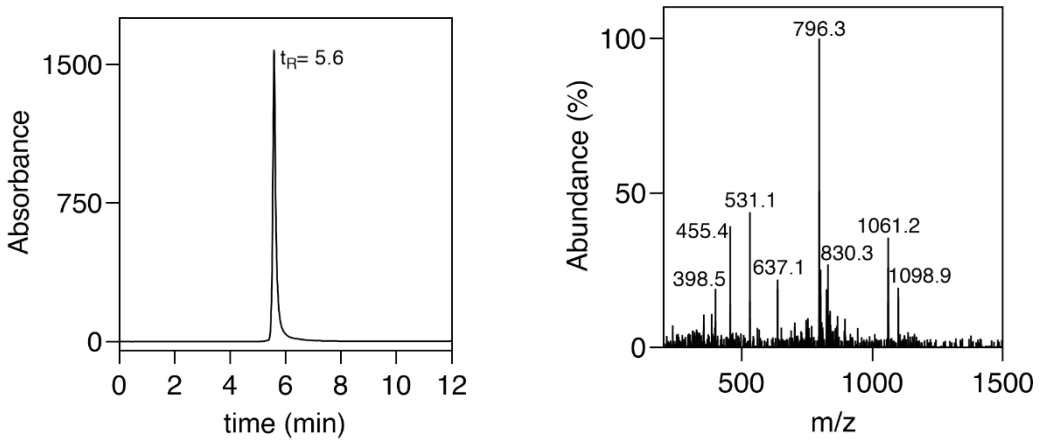


Figure 86. Left: chromatogram of the purified **RGH-LLD** ligand. Right: mass spectra corresponding to the peak with t_R of 5.6 min.

HPLC-MS (ESI): (5-95% B, t_R= 5.6 min) Calculated for $C_{150}H_{195}N_{57}O_{24}$ = 3178.58; found $[M+3H+TFAH]^{3+}$ = 1098.9; $[M+3H]^{3+}$ =1061.2; $[M+4H+TFAH+Na]^{4+}$ = 830.3; $[M+4H]^{4+}$ = 796.3; $[M+5H]^{5+}$ =637.1; $[M+6H]^{6+}$ =531.1; $[M+7H]^{7+}$ = 455.4; $[M+8H]^{8+}$ =398.5.

TAMRA-RGH-LLD:

H-RGHKKK(TAMRA)(βAla5Bpy)$_2$RPr(βAla5Bpy)$_2$RPr(βAla5Bpy)$_2$-NH$_2$

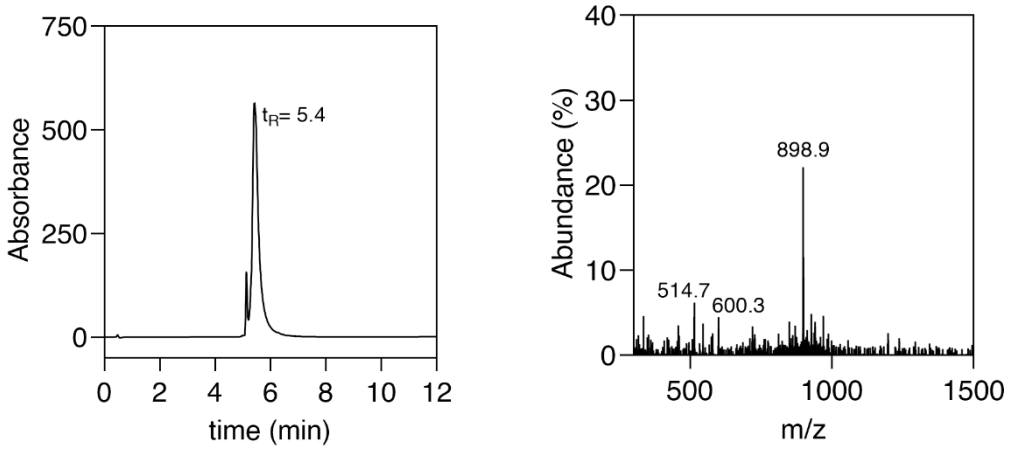


Figure 87. Left: chromatogram of the purified **TAMRA-RGH-LLD** ligand. Right: mass spectra corresponding to the peak with t_R of 5.4 min.

HPLC-MS (ESI): (5-95% B, t_R= 5.4 min) Calculated for $C_{175}H_{215}N_{59}O_{28}$ = 3590.72; found $[M+4H]^{4+}$ = 898.9; $[M+6H]^{6+}$ = 600.3; $[M+7H]^{7+}$ = 514.7.

2. *Synthesis of the Cu^{II} and Ni^{II} peptide helicates*

Onto a solution of **RGH-LLD** or **TAMRA-RGH-LLD** in phosphate buffer (1 mM, 10 mM NaCl, pH 7.0), $CuCl_2$. • 2 H_2O or $NiCl_2$ were added as source of Cu^{II} and Ni^{II}, ions, respectively, until saturating conditions (7.5 eq.), and the samples were incubated for 24 h.

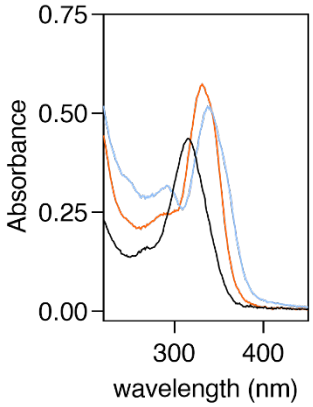


Figure 88. UV-Vis of a 5 µM solution of **RGH-LLD** ligand (1 mM phosphate buffer, 10 mM NaCl, pH 7.0) before (black line) and after the addition of 7.5 equivalents of of Cu(II) (blue line) or Ni(II) (orange line) ions.

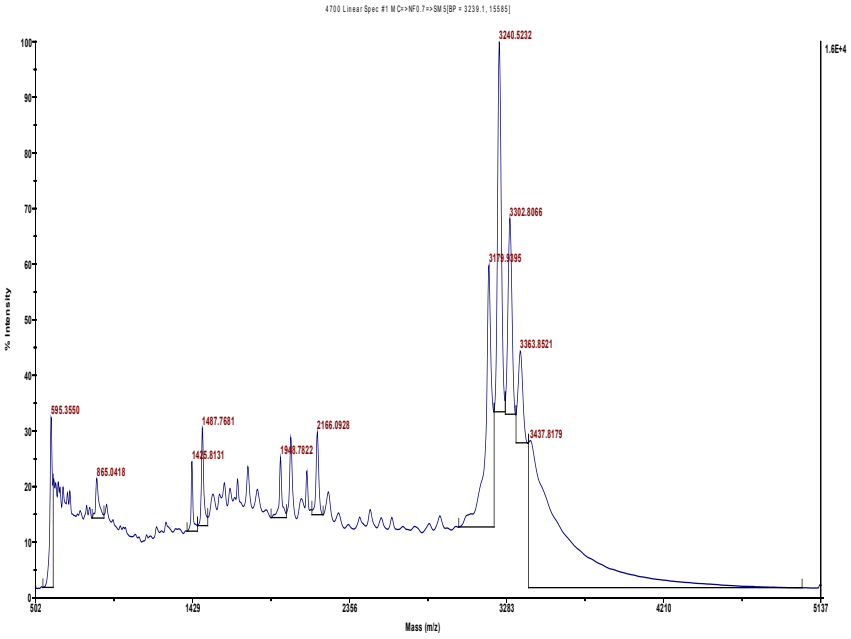


Figure 89. MALDI spectra of **ΔΔ-Cu^{II}_3RGH-LLD.**

MS (MALDI-TOF): Calculated for **ΔΔ-Cu^{II}_3RGH-LLD** $C_{150}H_{191}N_{57}O_{24}Cu_3$ = 3363.34; found: $[M+3Cu+H]^+$ = 3363.85; $[M+2Cu+H]^+$ = 3302.81; $[M+Cu+H]^+$ = 3240.52 $[M+H]^+$ = 3179.94.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

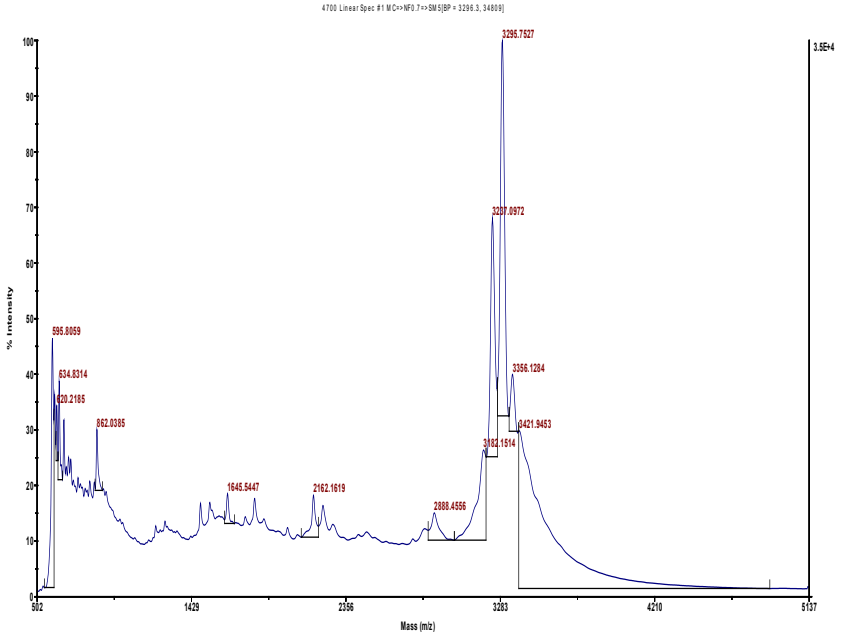


Figure 90. MALDI spectra of **ΛΛ-Ni$^{II}_3$RGH-LLD.**

MS (MALDI-TOF): Calculated for **ΛΛ-Ni$^{II}_3$RGH-LLD** $C_{150}H_{191}N_{57}O_{24}Ni_3$ = 3352.39; found: $[M+3Ni+H]^+$ = 3356.13; $[M+2Ni+H]^+$ = 3295.75; $[M+Ni+H]^+$ = 3237.10 $[M+H]^+$ = 3182.15.

3. Fluorescence titrations experiments with DNA

3.1. Fluorescein-labelled dsDNA with ΔΔ-Cu$^{II}_3$RGH-LLD and ΛΛ-Ni$^{II}_3$RGH-LLD

Fluorescence titrations were performed in which, over a 2 μM solution of fluorescein-labelled dsDNA (Flu-dsDNA) aliquots of a mother solution of **ΔΔ-Cu$^{II}_3$RGH-LLD** or **ΛΛ-Ni$^{II}_3$RGH-LLD** were added. In both cases, The K_D of the interaction could not be calculated as a precipitation process occurs in the titration cuvette during the experiment. This result demonstrates the absence of specific interaction between these peptide helicates and dsDNA.

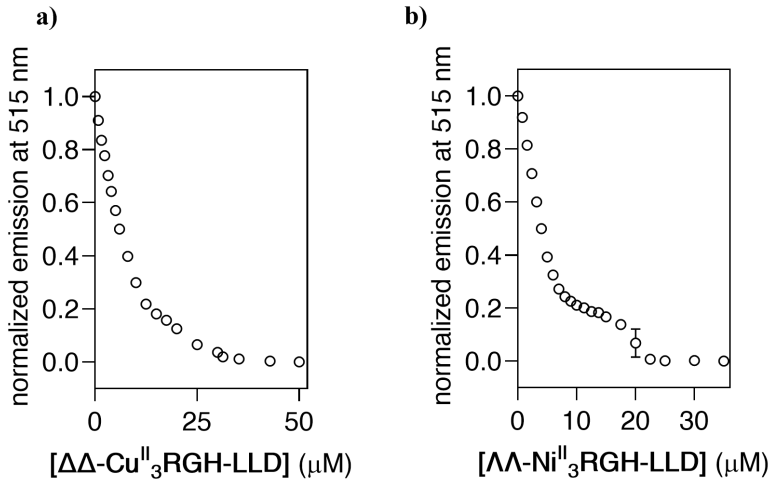


Figure 91. a) Fluorescence titration profile at 515 nm of a the 2 μM solution of Flu-dsDNA upon addition of increasing amounts of **ΔΔ-Cu$^{II}_3$RGH-LLD**; b) same as a) but in the presence of **ΛΛ-Ni$^{II}_3$RGH-LLD**. Conditions: phosphate buffer 1mM, NaCl 10 mM, pH 7.0; λ_{exc} = 490 nm.

3.2. Fluorescein-labelled GC-3WJ with ΔΔ-Cu^{II}_3RGH-LLD and ΛΛ-Ni^{II}_3RGH-LLD

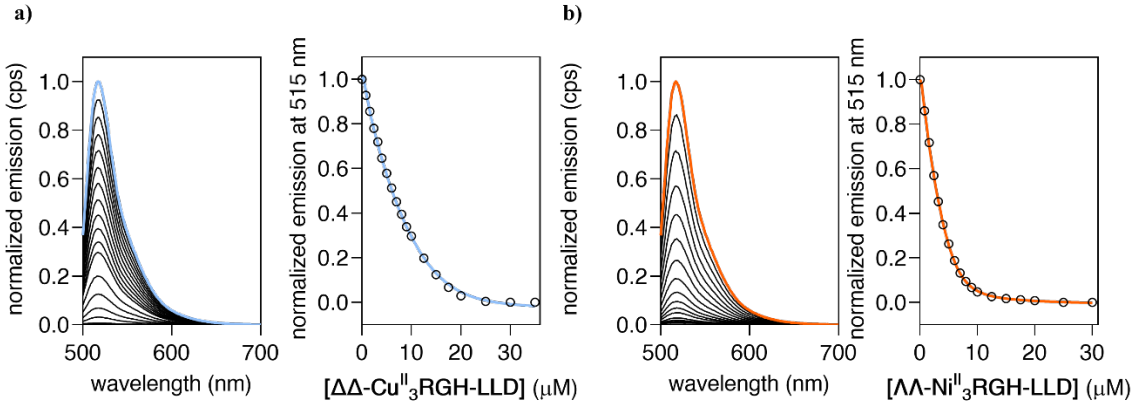


Figure 92. a) Left: normalized emission spectra of a 2 μM solution of the fluorescein-labelled GC-3WJ (Flu-GC-3WJ) in absence (blue line) and with increasing concentrations of **ΔΔ-Cu^{II}_3RGH-LLD** (black lines of decreasing intensity); right: titration profile at 515 nm with the best fit (blue line) to a 1:1 binding model plus non-specific interactions calculated with *Dynafit*; b) left: normalized emission spectra of a 2 μM solution of the fluorescein-labelled GC-3WJ (Flu-GC-3WJ) in absence (orange line) and in the presence of increasing concentrations of **ΛΛ-Ni^{II}_3RGH-LLD** (black lines of decreasing intensity); right: titration profile at 515 nm with the best fit (orange line) to a 1:1 binding mode plus non-specific interactions calculated with *Dynafit.* Titration data are the average of three independent experiments. Conditions: phosphate buffer 1mM, NaCl 10 mM, pH 7.0; λ_{exc} = 490 nm.

3.3. Fluorescein-labelled TA-3WJ with ΔΔ-Cu^{II}_3RGH-LLD

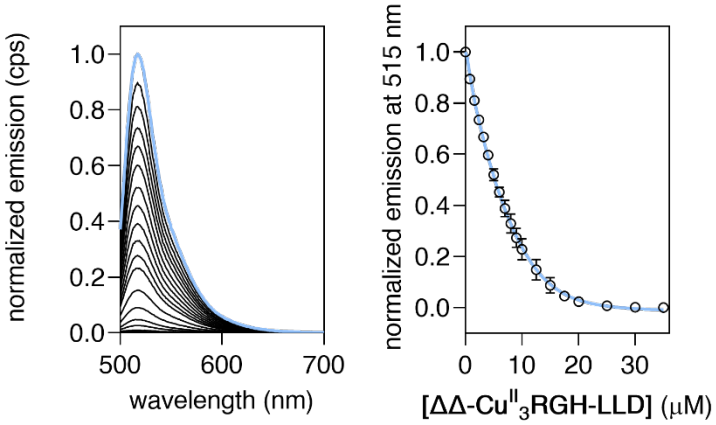


Figure 93. Left: normalized emission spectra of a 2 μM solution of the fluorescein-labelled TA-3WJ (Flu-TA-3WJ) in absence (blue line) and in the presence of increasing concentrations of **ΔΔ-Cu^{II}_3RGH-LLD** (black lines of decreasing intensity); right: titration profile at 515 nm with the best fit (blue line) to a 1:1 binding model plus non-specific interactions calculated with *Dynafit.* Titration data are the average of three independent experiments. Conditions: phosphate buffer 1mM, NaCl 10 mM, pH 7.0; λ_{exc} = 490 nm.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4. Cell-internalization studies

These experiments were carried out in collaboration with the **Dr. Natalia Barreiro Piñeiro** and **Prof. José Martínez Costas** from CiQUS.

4.1. Cell-internalization of the peptide ligand TAMRA-RGH-LLD in HeLa cells

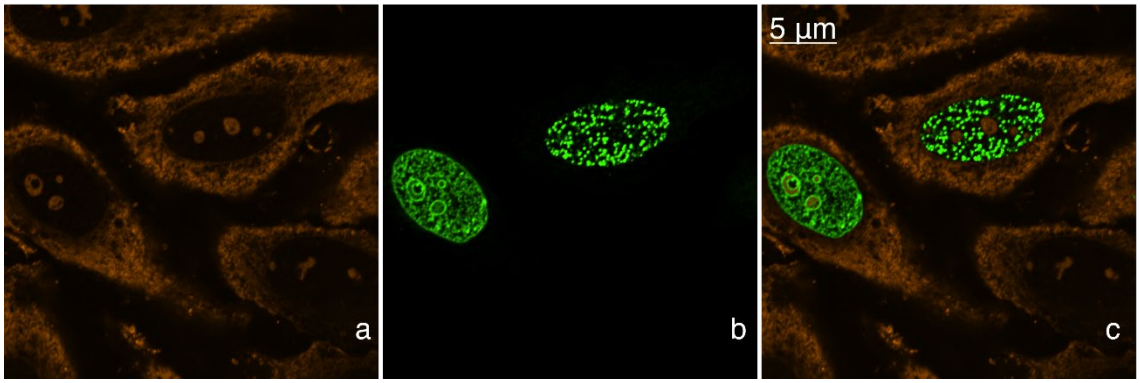


Figure 94. Free peptide ligand **TAMRA-RGH-LLD** alone enters the cell nucleus but cannot recognize DNA replication sites. HeLa cells expressing protein GFP-PCNAL2 were incubated in the presence of 25 µg/mL of Digitonin for 3 min and then in the presence of 5 µM of **TAMRA-RGH-LLD** for 15 min (these reported concentrations of digitonin, and peptide ligand refer to the final amount of these compounds in the cell culture transfer buffer after their addition from more concentrated stock solutions). a) Red channel emission showing the distribution of the rhodamine-labelled peptide ligand; b) green channel, corresponding to the emission of the GFP-PCNAL2 probe labelling the DNA replication foci; c) overlay of the green and red channels. Nucleoli are clearly labelled as big red spots inside the cell nuclei. The images show a representative confocal section of the Vero cells.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Chapter IV

1. UHPLC-MS of the purified peptide ligand 1

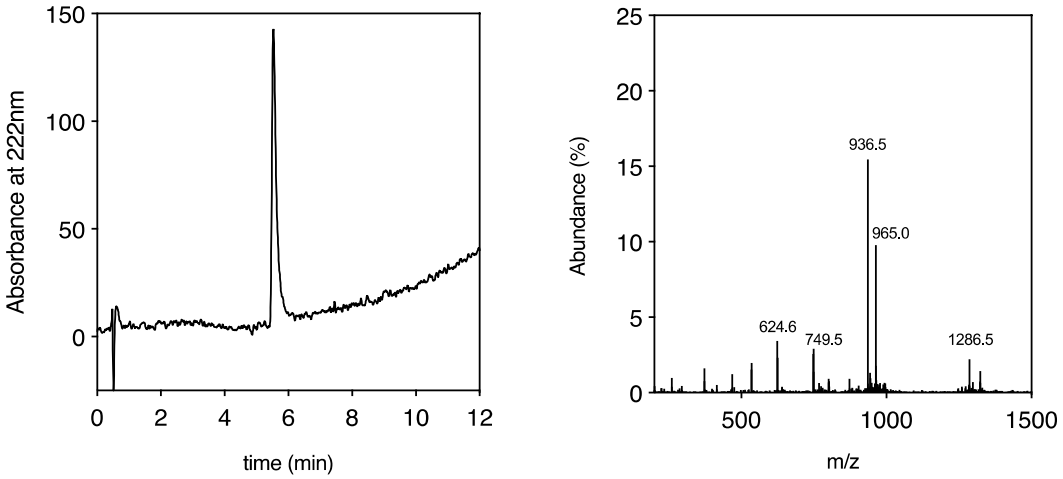


Figure 95. Left: chromatogram of the purified **1** ligand. Right: mass spectra corresponding to the peak with t_R of 5.54 min.

HPLC-MS (ESI): (5-95% B, t_R= 5.54 min) Calculated for $C_{177}H_{294}N_{60}O_{30}$ = 3740.33; found $[M+3H+TFAH]^{3+}$ = 1286.50; $[M+4H+TFAH]^{4+}$ = 965.0; $[M+4H]^{4+}$ = 936.5; $[M+5H]^{5+}$ = 749.5; $[M+6H]^{6+}$ = 624.6.

MS (MALDI-TOF): Calculated for $C_{177}H_{294}N_{60}O_{30}$ = 3740.33; found: $[M+H]^{+}$ = 3746.12.

2. UV-Vis Spectroscopy

UV-Vis absorbance of a 5 µM solution of the peptide ligand **1** in 10 mM HEPES buffer, 100 mM NaCl, pH 7.01 was measured, showing a maximum c.a. 320 nm

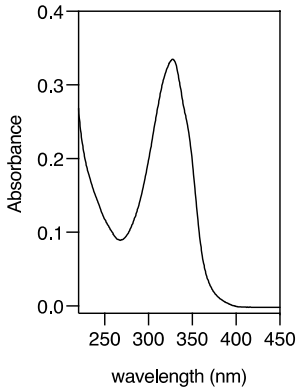


Figure 96. UV-Vis spectra of the AT-Hook trimer **1**.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

3. Fluorescence titrations experiments with DNA

Control experiments carried out in the presence of the TA-3WJ and the aromatic core of the peptide ligand were carried out. Thus, over a 1 μM solution of the aromatic core in HEPES buffer 10 mM, NaCl 100 mM, pH 7.01, TA-3WJ was added. In this study no DNA-binding can be appreciated in the oscillating pattern that the profile of the titration shows.

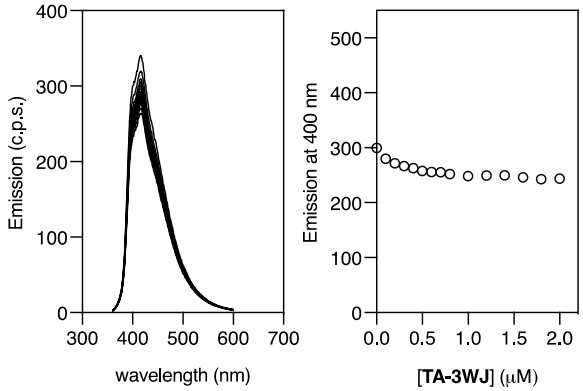


Figure 97. Left: full emission spectra of the aromatic core (more intense) and in the presence of increasing concentrations of the TA-3WJ (decreasing intensity). Right: profile of the full titration.

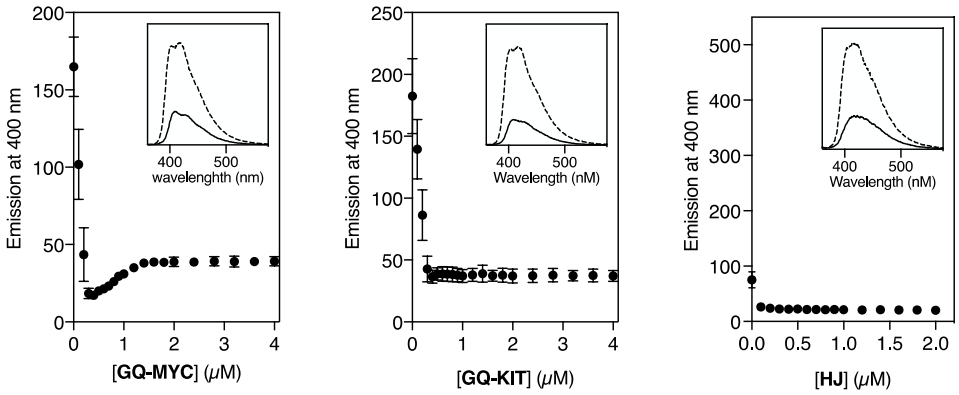


Figure 98. Selectivity of **1** towards different non-canonical DNA structures. Fluorescence titrations of a 1 μM solution of **1** in HEPES buffer 10 mM, 100 mM NaCl, pH 7.0 with: a) **GQ-MYC**; b) **GQ-KIT**; c) **HJ**. Insets show the first (dashed line) and last (solid line) emission spectra in the titrations.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

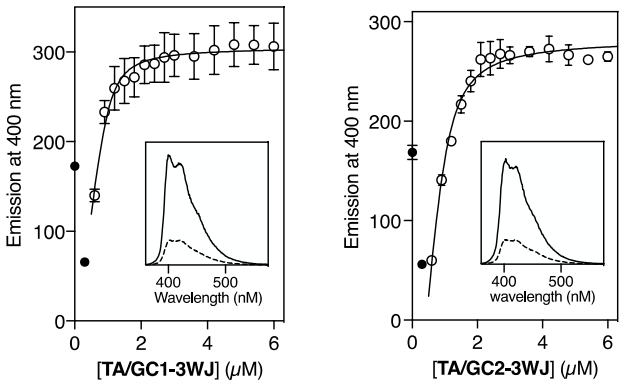


Figure 99. Selectivity of **1** towards mixed-sequence three-way junctions. Probe **1** also binds to asymmetric 3WJs containing two AT-Hook binding sites (TA/GC1-3WJ) and a single AT-Hook binding site in one of the branches (TA/GC2-3WJ). The dissociations constants for each of these partly complementary sites is about 60 and 100 nM, for TA/GC1-3WJ and TA/GC2-3WJ, respectively. supporting the relevance of the simultaneous recognition of the three dsDNA branches in the 3WJ to achieve the best binding. The fact that non-target 3WJs still present remarkably tight binding is probably due to the high contribution of ionic (Arg/phosphate) interactions, which can also take place even with G/C-rich sequences and has been estimated to represent up to 70% of the total energy of the AT-Hook/DNA. [312]

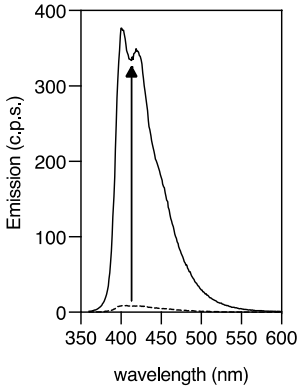
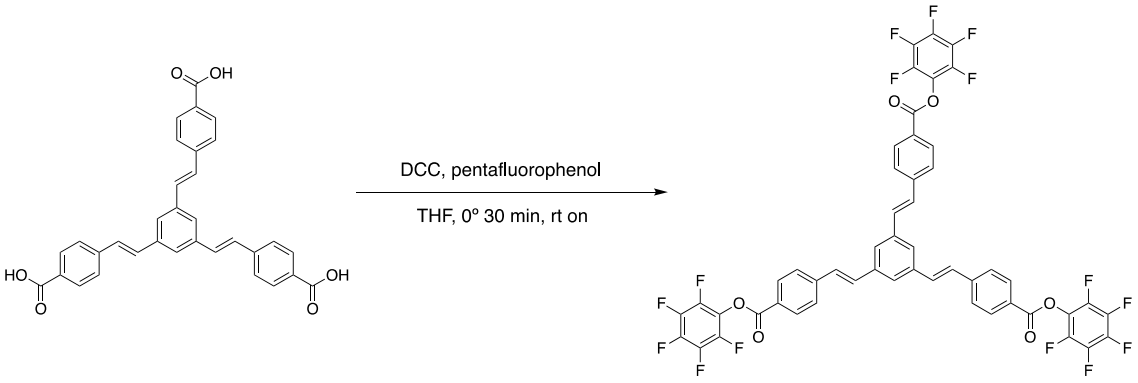


Figure 100. Emission spectra of of a 1 μM solution of **1** in HEPES buffer 10 mM, 100 mM NaCl, pH 7.0 in the presence of 2 μM **GC-3WJ** (dashed line) and after the addition of 1 eq. of **TA-3WJ** (solid line).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

[312] A. I. Dragan, J. R. Liggins, C. Crane-Robinson, P. L. Privalov, *J. Mol. Biol.* **2003**, *327*, 393–411.

4. *Synthesis of tris(perfluorophenyl) 4,4',4''-((1E,1'E,1''E)-benzene-1,3,5-triyltris(ethene-2,1-diyl)) tribenzoate (2)*



1 **2**

To a stirred solution of the tricarboxylic acid **1** (50 mg, 0.096 mmol) in THF (7 mL) under argon at 0°, DCC (72 mg, 0.36 mmol) was added. The reaction was allowed to reach room temperature, after which, pentaflurophenol (53 mg, 0.288 mmol) was added. The mixture was stirred overnight. After this time, the reaction was filtered, and the solvent was evaporated under reduced pressure. The resulting solid was purified by flash chromatography using DCM as eluent to yield the compound **2** as a white powder (45 mg, 46% yield). **^{1}H-NMR** (500 MHz, CDCl3, δ): 8.23 (d, *J* = 8.3 Hz, 6H), 7.72 (m, 9H), 7.33 (dd, *J* = 16, 6.5 Hz, 6H).**MS** (**APCI-TOF**): calculated for $C_{51}H_{21}F_{15}O_6$ = 1014.11; found: $[M+H]^+$= 1015.11.

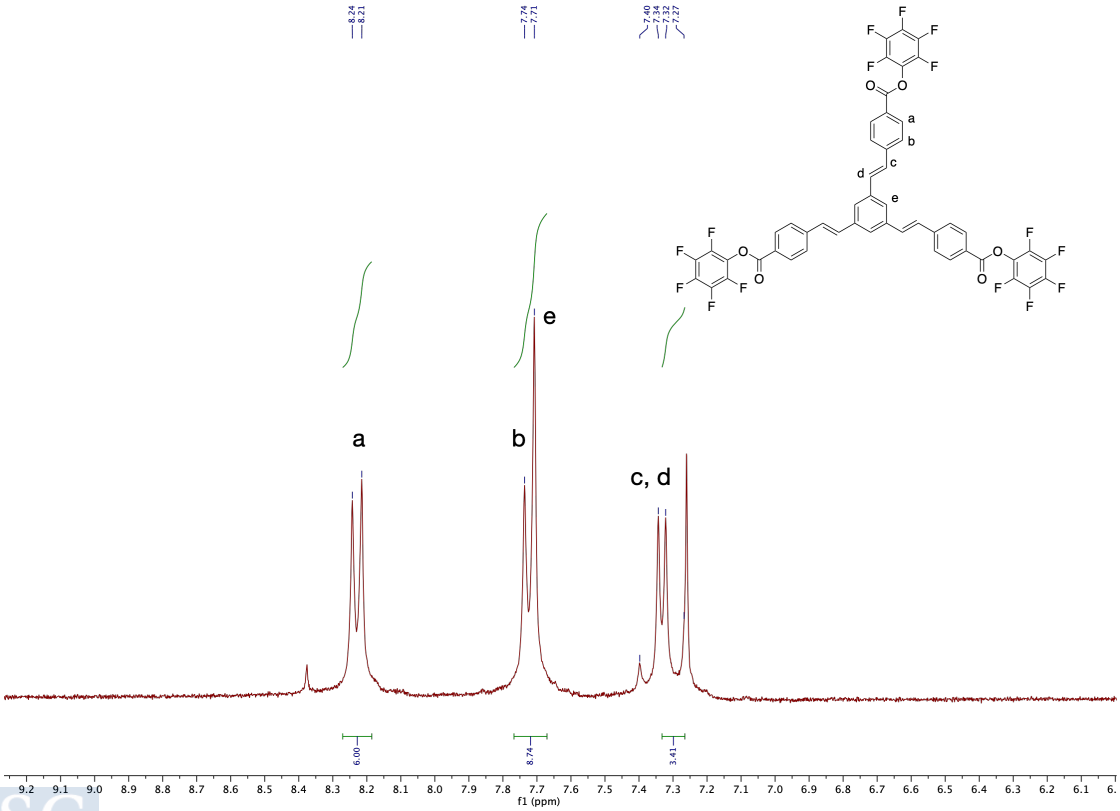


Figure 101. ^{1}H-NMR spectra of the compound **2** in $CDCl_3$.

5. *Molecular modelling*

These studies were carried out by **Prof. Dr. Jean-Didier Maréchal** at Universitat Autònoma de Barcelona.

They performed protein-ligand docking of 1,3,5-tristyrylbenzene to the 3WJ/AT-Hook model using four objectives during the simulation: the three distances between the styryl benzene carboxylates and the N-terminal Lys side chain of the AT-hook peptides, which should be set to an average of 1.35 Å, and minimization of the atomic clashes between the ligands and the receptor. From those calculations, they observed 13 poses that met the distance objective with a reasonable deviation of ± 1.5 Å (an acceptable deviation considering the likely geometric rearrangements of the 3WJ/AT-Hook complexes), two of the best poses also have very good contact values (lower than 100 Å).

System set up: The structure of a crystallographic three-way junction (PDB code: 1f44) was superimposed on the AT Hook region (PDB code: 3uxw) using UCSF Chimera software.[313] Then, the crystallographic waters and other small molecules were removed, as well as the three-way junction template after manual curation. The tristyryl benzene was set up with all carboxylic oxygens adjusted to the proper charge. Both systems were protonated using the implemented algorithm in UCSF Chimera.

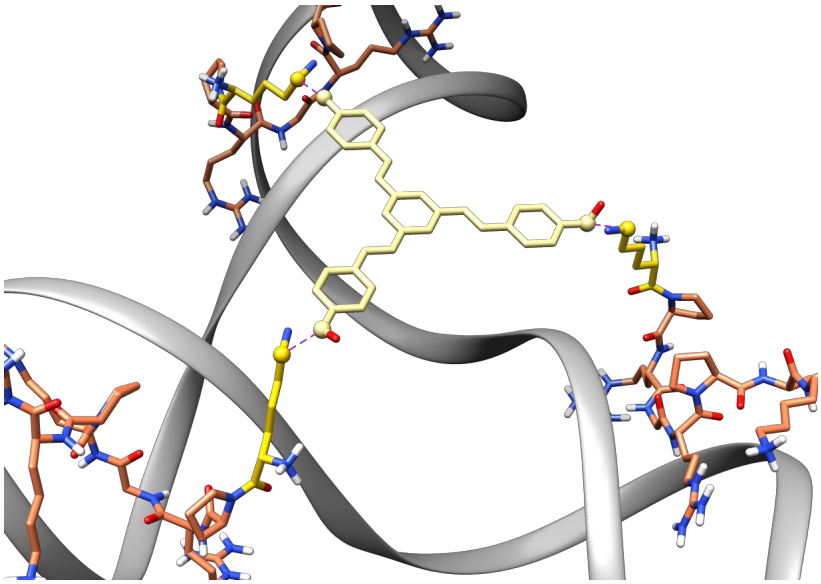


Figure 102. Set up system. The backbone scaffold of the DNA three-way junction (PDB code: 1F44) and AT rich region (PDB code: 3UXW) are depicted in gray; AT-Hook peptides in orange with the terminal lysines in dark yellow. The 1,3,5-tritrystyrilbenzene is depicted in light yellow. Atoms with ball shape indicate the points between which the distances were restricted in the *GaudiMM*.

Docking. The multi-objective genetic algorithm *GaudiMM* was used for the simulations, with an input (input.yaml file provided at the end of the section) containing the information on the objectives and how to optimize them. Calculations were performed considering four objectives: the three distances between each carbonyl C atom on the 1,3,5-tristyrylbenzene and the N on the lysine side chain at the N-termini of the AT-hook,

313 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–1612.

aiming to be matched to an average of 1.35Å and minimization of the atomic clashes between the ligands and the receptor. Moreover, the flexibility of the terminal lysines was allowed to facilitate accommodation of the linker.

GaudiMM calculations led to 13 solutions satisfying the distance objectives with ±1.5 Å (accounting for some fluctuation of the DNA structure and AT-hook peptides with respect to the original X-ray structure). Moreover, 2 of these solutions exhibit clashes lower than 100 Å; a good value as demonstrated in previous studies.[314]

Input.yaml file for GaudiMM

```
path: ./input.yaml
ga:
                cx_eta: 5
                cx_pb: 0.5
                generations: 100
                lambda_: 3.0
                mu: 1.0
                mut_eta: 5
                mut_indpb: 1.0
                mut_pb: 0.5
                population: 200
genes:
-module: gaudi.genes.molecule
                name: Protein
                path: ./protein.mol2

-module: gaudi.genes.molecule
                name: Cofactor
                path: ./linker.mol2

-module: gaudi.genes.torsion
                name: TorsionC
                target: Cofactor
                rotatable_bonds: [[1, 56, 1], [50, 51, 50], [53, 54, 53] ]

-module: gaudi.genes.rotamers
                name: Rotamers
                residues: [Protein/10, Protein/1, Protein/19 ]

-interpolation: 0.5
                module: gaudi.genes.search
                name: Search
                precision: 5
                radius: 3.0
                rotate: true
                target: Cofactor

objectives:
-module: gaudi.objectives.contacts
                name: Clashes
                probes: [Cofactor]
                radius: 5.0
                weight: -1.0
                which: clashes

-module: gaudi.objectives.distance
                name: Distance56-160
```

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

314 J.-E. Sánchez-Aparicio, G. Sciortino, D. V. Herrmannsdoerfer, P. O. Chueca, J. R.-G. Pedregal, J.-D. Maréchal, *Int. J. Mol. Sci.* **2019**, *20*, 3155.

```
                target: Protein/160
                probes: [Cofactor/56]
                threshold: 1.350
                weight: -1.0

-module: gaudi.objectives.distance
                name: Distance54-9
                target: Protein/9
                probes: [Cofactor/54]
                threshold: 1.350
                weight: -1.0

-module: gaudi.objectives.distance
                name: Distance51-313
                target: Protein/313
                probes: [Cofactor/51]
                threshold: 1.350
                weight: -1.0
output:
                check_every: 0
                compress: true
                history: false
                name: 3uxw_dist_clash_rotK
                pareto: true
                path: ./3uxw_dist_clash_rotK
                precision: 3
                prompt_on_exception: false
                verbose: true
similarity:
                args:
                -   - Cofactor
                - 0.5
                kwargs: {}
                module: gaudi.similarity.rmsd
```

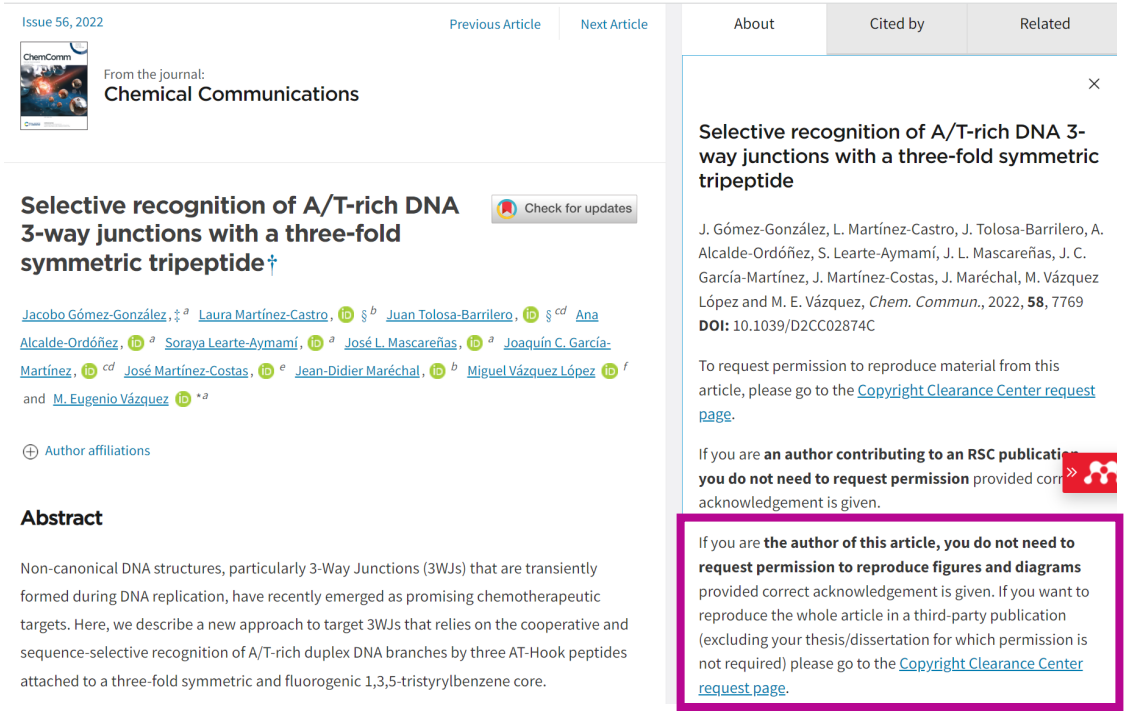
USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

List of publications

Chapter IV

Title	Selective recognition of A/T-rich DNA 3-way junctions with a three-fold symmetric tripeptide
Authors	Jacobo Gómez González, Laura Martínez Castro, Juan Tolosa Barrilero, Ana Alcalde Ordóñez, Soraya Learte Aymamí, José Luis Mascareñas Cid, Joaquín C. García Martínez, José Manuel Martínez Costas, Jean-Didier Maréchal, Miguel Vázquez López, Marco Eugenio Vázquez Sentís
Journal	Chemical Communications
Year	2022
Volume, pages	58, 7769-7772
Impact factor, Quartile	6.065, Q1 (2022-2023)
PhD student contribution	Experimental work: DNA binding studies. Data analysis and Manuscript preparation.



Issue 56, 2022 Previous Article Next Article

From the journal: Chemical Communications

Selective recognition of A/T-rich DNA 3-way junctions with a three-fold symmetric tripeptide†

Check for updates

Jacobo Gómez-González,‡[a] Laura Martínez-Castro,[b] Juan Tolosa-Barrilero,[cd] Ana Alcalde-Ordóñez,[a] Soraya Learte-Aymamí,[a] José L. Mascareñas,[a] Joaquín C. García-Martínez,[cd] José Martínez-Costas,[e] Jean-Didier Maréchal,[b] Miguel Vázquez López[f] and M. Eugenio Vázquez*[a]

Author affiliations

Abstract

Non-canonical DNA structures, particularly 3-Way Junctions (3WJs) that are transiently formed during DNA replication, have recently emerged as promising chemotherapeutic targets. Here, we describe a new approach to target 3WJs that relies on the cooperative and sequence-selective recognition of A/T-rich duplex DNA branches by three AT-Hook peptides attached to a three-fold symmetric and fluorogenic 1,3,5-tristyrylbenzene core.

About Cited by Related

Selective recognition of A/T-rich DNA 3-way junctions with a three-fold symmetric tripeptide

J. Gómez-González, L. Martínez-Castro, J. Tolosa-Barrilero, A. Alcalde-Ordóñez, S. Learte-Aymamí, J. L. Mascareñas, J. C. García-Martínez, J. Martínez-Costas, J. Maréchal, M. Vázquez López and M. E. Vázquez, *Chem. Commun.*, 2022, **58**, 7769
DOI: 10.1039/D2CC02874C

To request permission to reproduce material from this article, please go to the Copyright Clearance Center request page.

If you are **an author contributing to an RSC publicati you do not need to request permission** provided cor acknowledgement is given.

If you are **the author of this article, you do not need to request permission to reproduce figures and diagrams** provided correct acknowledgement is given. If you want to reproduce the whole article in a third-party publication (excluding your thesis/dissertation for which permission is not required) please go to the Copyright Clearance Center request page.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA





Non-canonical DNA structures, such as three-way DNA junctions (3WJs), formed during DNA replication, have emerged as promising alternative DNA targets for the development of new drugs with attractive therapeutic properties. In this context, supramolecular helicates have shown high selectivity and binding affinity for 3WJs, but their use in this field is hampered by the lack of synthetic versatility and a general strategy to their enantioselective synthesis. In this thesis, we describe the stereoselective self-assembly of helicates derived from peptide ligands equipped with 2,2'-bipyridine coordinating residues, as well as the implementation of a new strategy that relies on the cooperative 3WJ-binding mechanism of natural proteins, to yield new systems with selective 3WJs recognition properties.