TESE DE DOUTORAMENTO

DNA RECOGNITION WITH SELF-ASSEMBLED PEPTIDE HELICATES

Jacobo Gómez González

ESCOLA DE DOUTORAMENTO INTERNACIONAL

PROGRAMA DE DOUTORAMENTO EN CIENCIA E TECNOLOXÍA QUÍMICA

SANTIAGO DE COMPOSTELA / LUGO

AÑO 2020

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

DECLARACIÓN DEL AUTOR DE LA TESIS

DNA recognition with self-assembled peptide helicates

D. Jacobo Gómez González

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) *En su caso, en la tesis se hace referencia a las colaboraciones que tuvo este trabajo.*
3) *La tesis es la versión definitiva presentada para su defensa y coincide con la versión enviada en formato electrónico.*
4) *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.*

En Santiago de Compostela, 31 de Enero de 2020

Fdo. Jacobo Gómez Gonzalez

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

AUTORIZACIÓN DEL DIRECTOR / TUTOR DE LA TESIS

DNA recognition with self-assembled peptide helicates

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

INFORMA/N:

Que la presente tesis, corresponde con el trabajo realizado por D. Jacobo Gómez González bajo mi dirección, 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 ésta no incurre en las causas de abstención establecidas en Ley 40/2015.

En Santiago de Compostela, 31 de Enero de 2020.

Fdo.: M. Eugenio Vázquez Sentís

Fdo.:Miguel Vázquez López

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Acknowledgements

En primer lugar me gustaría agradecer a mis directores, Miguel y Eugenio, por haberme dado la oportunidad de realizar esta tesis y descubrir el mundo de la investigación durante estos últimos 4 años de mi vida. Ha sido un largo camino que seguro habría sido mucho más complicado de no haber gozado de vuestro apoyo, paciencia y dedicación.

En segundo lugar me gustaría agradecer a mi familia, especialmente a mi madre y abuelo, por haber sido siempre de gran ayuda en todos los aspectos de mi carrera académica y personal, empujándome a continuar y ofreciéndome los medios necesarios para poder hacerlo sin pedir a cambio nada más que mi éxito personal; me gustaría también resaltar en este apartado a Clara, apoyo incondicional y soporte moral durante estos tres últimos años de mi vida. Sin tu compañía todo habría sido mucho más difícil.

Gracias también a las personas que encendieron mi llama investigadora y me dieron la oportunidad de pisar por primera un laboratorio de química: Eugenio, Lourdes, Saleta, María, André, Xosé... sin vosotros esta tesis no habría sido posible.

Me gustaría dedicar un agradecimiento especial a las personas más cercanas a mi en toda esta etapa. A David , porque a pesar de tanto *bullying*, puedo sentirme orgulloso de considerarte un hermano más; a Sonia por tantos buenos momentos en el laboratorio y fuera de él; a Soraya por tantos vaciles y cotilleos; a Diego por acogerme en mi llegada y guiarme hasta hacer de mi un poquito mejor químico; a Ana por su buen carácter siempre; a Andrés por los seis peores meses de mi vida, en los que ni un segundo he gozado de silencio; y por último a Gus por el apoyo y ayuda en este tramo final. Sin todos vosotros estos cuatro años hubieran sido eternos; gracias por hacerlos inolvidables.

Gracias también al prof. José Martínez Costas por la ayuda prestada en todos los experimentos realizados en células, así como por los buenos momentos compartidos. También como no, agradecer a la gente del P3L6-Rebeca, Natalia y Tomás- los buenos ratos, risas y, por supuesto, la comida compartida conmigo.

También me gustaría dedicar algunas palabras de agradecimiento a quienes a pesar de no formar parte de mi grupo, han amenizado muchas mañanas, tardes, comidas, cafés y noches de juerga: Jesús, Roi, Arcadio, Lamas, Eva, Álvaro, Marta, Sabela, Sara, Fátima, Pozo, Félix, Saleta, Felipe, Jaime, Sandra, Marc... y tantos otros que probablemente me haya dejado por el camino y que espero puedan perdonarme por ello.

Mención especial también a mis amigos coruñeses de la infancia y adolescencia: Dani, Yerai, Anxo, Antón, Álex, Abraio, Javi, Kas, Kike, Marcos... por tan buenos momentos habidos y por haber. Gracias por ayudarme a forjar quien soy a día de hoy.

Por supuesto no puedo olvidarme de todos aquellos que me acompañaron a lo largo del comienzo de mi vida académica y con los que he compartido los que probablemente hayan sido los mejores años de mi vida hasta el momento: Movilla, Pintado, Noguer, Fer, Brais, Nico... ¡E os que nos quedan aínda!

Por supuesto no debo olvidarme de la gente de Louro, con la que llevo compartiendo veranos interminables desde que mi memoria alcanza. Álvaro, Juan, Santi, Laura, Cristian, Iñigo, Jorge, Gonzalo, Antía, Gonzaga, Marta... gracias por amenizar los meses de descanso de esta tesis y ayudarme a coger fuerzas cada año para continuar.

Agradecer también a mis compañeros de equipo, cuerpo técnico y directiva del Racing San Lorenzo, que me han ayudado a evadirme de los días largo y difíciles, así como a encontrar de nuevo las ganas de jugar al fútbol.

Finally, I would like to thank Professor Gunnlaugsson for giving me the opportunity to spend 4 wonderful months doing chemistry in his laboratory. And, of course, thanks to all the people there (Adam, Bruno, Ema, Sachi, Oxana, Sam, Ana, Jason, Eoin, Charlyne…) for making me feel like home.

A mis abuelos.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Table of contents

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Abbreviations

5Bpy	5'-amino-2,2'-bipyridyl-5-carboxylic acid
βAla	β-Alanine
δ	chemical shift (ppm)
Θ	molar elipcity
°	Celsius degree
A	Adenine
aa	amino acid
Ac	acetyl
Ac_2O	acetic anhydride
alloc	allyloxycarbonyl
DNA	deoxyribonucleic acid
Boc	*t*ert-butyloxycarbonyl
Bpy	2,2´-bipyridine
br	broad signal (in NMR spectra)
CD	circular dichroism
c.p.s	counts per second
d	doublet (in NMR spectra)
dd	double doublet (in NMR spectra)
DCM	dicloromethane
DIEA	*N*,*N*´-diisopropylethylamine
DMF	*N*,*N*´-dimethylformamide
DMSO	dimethylsulfoxide
dsDNA	double-stranded DNA
MS	mass spectrometry
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 hexafluorophosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	High Performance Liquid Chromatography
J	coupling constant (Hz)
K_D	dissociation constant
µM	micromolar
m	multiplet (in NMR spectra)
MALDI	Matrix Assisted Laser Desorption Ionization
MeOH	methanol
mdeg	millidegrees
min	minute
mM	milimolar
MQ	milli-Q water
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
twDNA	three-way DNA junction
rt	room temperature
TFA	trifluoroacetic acid
TIS	triisopropylsilane
UV-Vis	Ultraviolet-visible

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Introduction

Biological Chemistry

Chemistry is the branch of science concerned with the substances of which matter is composed, their properties and reactions, and the use of such reactions to form new substances. On the other hand, Biology deals with living organisms and their vital processes.[1] In 1828, Friedrich Wöhler published a paper entitled "*On the artificial formation of urea*", which demonstrated that inorganic starting materials could be used to synthesize substances that were associated only with living organisms.[2] Despite this early demonstration of the intimate connections between Chemistry and Biology, these two disciplines evolved independently until the 20th century. In 1910, Paul Ehrlich introduced the arsenic-based drug *Salvarsan* as a remedy for syphilis. His methodical search for a specific drug to treat a specific disease marked the beginning of Medicinal Chemistry.[3] A couple of decades later, in 1948, Sydney Farber published a revolutionary clinical study in *The New England Journal of Medicine*, showing that *Aminopterin* (4-aminopteroylglutamic acid) induced remissions in children with acute leukemia.[4] It was clear by then that Chemistry and Biology had a common playground studying the molecular basis of life, and Chemistry would be instrumental in the treatment of diseases, as S. Farber said: "*I think that cancer will be controlled eventually by chemical means*". The conceptual integration of both disciplines was finally proposed in 1945 by Linus Pauling and George W. Beadle, who included the term *Chemical Biology* in a grant proposal to the *Rockefeller Foundation* that outlined *"the great problems of Biology"*, which, according to Pauling, were only addressable thanks to the extraordinary recent advances in Physics and Chemistry.[5]

---

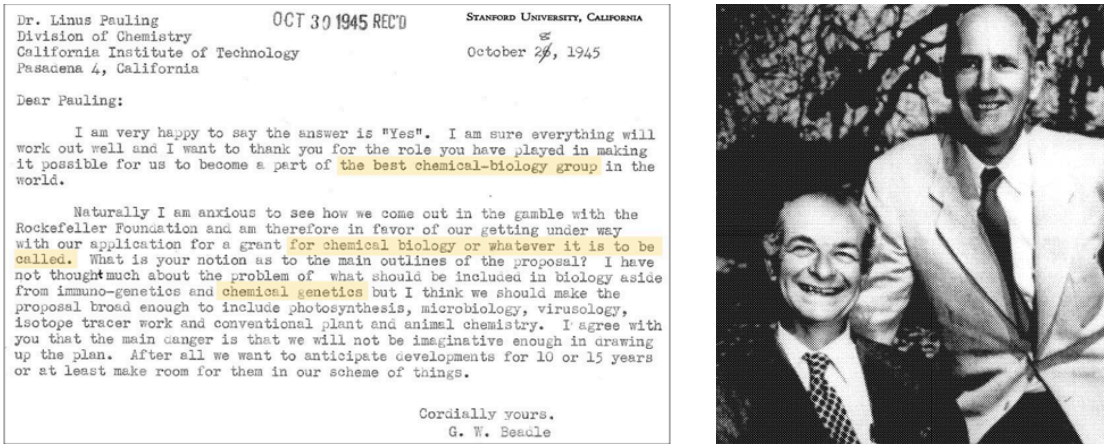
1 Oxford English Dictionary.

2 F. Wöhler, *Ann. Phys.* **1828**, 88, 253-256.

3 J. Drews, *Nat. Rev. Drug Discov.* **2004**, 3, 797-801.

4 S. Farber, L. K. Diamond, *N. Engl. J. Med.* **1948**, *238*, 787-793.

5 Lily 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*



Dr. Linus Pauling
Division of Chemistry
California Institute of Technology
Pasadena 4, California

OCT 30 1945 REC'D

STANFORD UNIVERSITY, CALIFORNIA

October 26, 1945

Dear Pauling:

I am very happy to say the answer is "Yes". I am sure everything will work out well and I want to thank you for the role you have played in making it possible for us to become a part of the best chemical-biology group in the world.

Naturally I am anxious to see how we come out in the gamble with the Rockefeller Foundation and am therefore in favor of our getting under way with our application for a grant for chemical biology or whatever it is to be called. What is your notion as to the main outlines of the proposal? I have not thought much about the problem of what should be included in biology aside from immuno-genetics and chemical genetics but I think we should make the proposal broad enough to include photosynthesis, microbiology, virusology, isotope tracer work and conventional plant and animal chemistry. I agree with you that the main danger is that we will not be imaginative enough in drawing up the plan. After all we want to anticipate developments for 10 or 15 years or at least make room for them in our scheme of things.

Cordially yours,
G. W. Beadle

Figure 1. Left: the grant proposal submitted by L. Pauling where the term "*Chemical Biology*" was first used (from @*CaltechArchives*). Right: Linus Pauling and George Beadle (taken from ref. 5).

The concept of *Chemical Biology* as an independent discipline reappears in a 1954 Editorial in the Caltech magazine *Engineering and Science* that, for the first time, defines the scope and aims of this new branch of Chemistry. "*[Chemical Biology] is merely the name given to that work in chemistry which is of biological interest—and to that work in biology in which the chemical approach is used to solve biological problems*".[6] Still, more than thirty years after this editorial, Arthur Kornberg, one of the biggest advocates of the integration of Chemistry and Biology, was still trying to convince his fellow biochemists of the value of Chemistry in biological research: "*Much of life can be understood in rational terms if expressed in the language of chemistry. It is an international language, a language for all of time, and a language that explains where we came from, what we are, and where the physical world will allow us to go*".[7]

Fortunately, at the beginning of the 21st century, Chemical Biology is not just a buzzword, but it has become an established field of research at the crossroads between Chemistry and Biology that seeks a deep molecular understanding of biological phenomena by chemical means, while at the same time aiming to the application of that biological knowledge for the advance of the chemical sciences.

Just as when in 1964 a US National Academy of Sciences committee defined Chemistry as "*the research activities of chemistry departments in this country*", Chemical Biology is a rich and diverse field that defies a closed definition and that constantly grows as new tools and approaches open new research venues. From the development of new bioconjugation methods for the selective modification of biomolecules, bringing the power of modern catalytic methods into the creation of bioactive species with extended properties, to the design of new sensing strategies that allow the visualization of hidden phenomena, including the application of the rich properties of biological molecules, protein and nucleic acids, outside their realm and into materials science or nanotechnology, Chemical Biology will be without any doubt at the core of the greatest advances in chemical (and biological) sciences.[8]

---

6 *Eng. & Sci.* **1954**, 17, 9-13

7 a) A. Kornberg, *Chem. Biol.* **1996**, 3, 3-5; b) A. Kornberg, *Biochemistry* **1987**, *26*, 6888-6891.

8 G. von Kiedrowski, *ChemBioChem* **2001**, 2, 597-598.

Amino acids, peptides and proteins

Introduction

It was Beccari of Bologna, a largely forgotten Italian philosopher, physician, and scientist, who in the 18th century described gluten as a protein component of wheat flour,[9] but it was not until the 19th century when chemists started the systematic and scientific studies of proteins as a unique family of biomolecules. Proteins were too complex to be studied by the analytic techniques available at that time, but scientists found that heating the proteins in concentrated acid afforded simple crystalline substances, called *amino acids*, which could be isolated and studied with their rudimentary methods. Modern protein chemistry can be dated back to 1820, when Henri Braconnot prepared *glycine* from gelatin in the course of his attempt to see if proteins behave like starch and are decomposed by acids with the production of sugar. By the beginning of the 20th century, 18 amino acids were identified as components of proteins.[10]

Peptides and proteins are polymers of amino acids linked through amide—also called *peptide*—bonds. The protein structure, reactivity, and function are defined by the sequence of its amino acids, also called *primary structure* of the protein. Amino acids contain a carboxyl group (–COOH), an amino group ($-NH_2$), a hydrogen atom, and a variable side chain (–R). Each amino acid has a unique side chain with a defined chemical structure—and therefore, reactivity. Amino acids are typically classified according to the chemical nature of their side chain.[11]

Soon, the three-letter abbreviations became tiresome to write, and a one-letter abbreviation system was finally established by Šorm in 1961.[12,13] The one-letter code followed mostly an alphabetical order, but with some colorful changes. Thus, for example, Arginine was abbreviated to *R*, because arginine sounds like *R*-ginine. Aspartic acid was abbreviated to *D*, because the amino acid's name sounds like aspar-*D*-ic acid. Similarly, glutamine became *Q*, sounds like *Q*lutamine, and phenylalanine was abbreviated to *F*, for *F*-enylalanine. Asparagine was abbreviated to *N* for asparagi-*N*-e, and glutamic acid represented by *E* because aspartic acid is *D*, and glutamic acid is larger than aspartic acid by one carbon atom, hence glutamic acid is the next letter of the alphabet, *E*. While Leucine is *L*, Lysine became *K* because *L* was already taken for leucine.

---

9 E. F. Beach, *J. Hist. Med. Allied Sci.* **1961**, *16*, 354-373.

10 H. B. Vickery, C. L. A. Schmidt, *Chem. Rev.* **1931**, *9*, 169-318.

11 M. Saffran, *Biochem. Ed.* **1998**, 26, 116-118.

12 F. Šorm, B. Keil, J. Vaněček, V. Tomášek, O. Mikeš, B. Meloun, V. Kostka, V. Holeyšovský, *Collect. Czech. Chem. Commun.* **1961**, *26*, 531-578.

13 a) D. Voet, J. G. Voet, C. W. Pratt, *Fundamentals of Biochemistry*, John & Sons, New York, **2007**, 76-81; b) B. ALberts, A. Johnson, J. Lewis, M. Raff, K Roberts, P. Walter, *Molecular Biology of the Cell* **2008**, Garland Publishing, New York, Chapter 3.

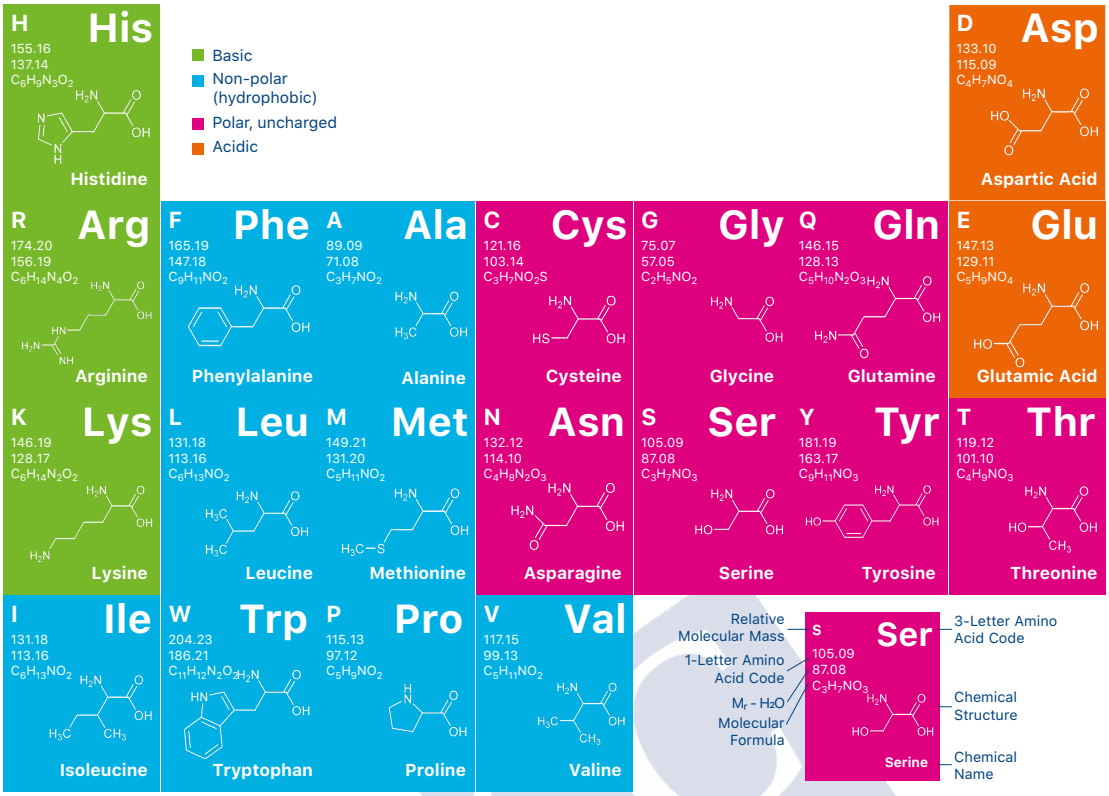


Figure 2. The "*periodic*" table of the 20 natural amino acids. Adapted from *www.bachem.com*.

Structure of proteins

The peptide bond

X-ray diffraction studies of amino acids and small peptides carried out by Linus Pauling and Robert Corey showed that peptide C–N bond is somewhat shorter than a typical C–N bond in amines, which indicates that the bond between the carbonyl carbon and the amide nitrogen has partial double bond character; as a consequence, the atoms associated with the peptide bond are coplanar and the peptide bond cannot rotate freely. Thus, the backbone of a polypeptide chain can be described as a series of rigid planes with a common point of rotation at Cα. The spatial orientation between the two planes—the conformation between consecutive residues—is characterized by two dihedral angles, ϕ (N–Cα) and ψ (Cα–C).[14] Certain combinations of ϕ and ψ angles lead to steric clashes between the atoms, and these define forbidden conformations for the peptide chains (in fact, most of the ϕ/ψ combinations that define conformational space are sterically excluded). The allowed conformations for a peptide chain can be plotted in a graph known as the *Ramachandran plot*, which shows defined regions in the ϕ/ψ space for each of the secondary structure elements that describe the local arrangement of the main-chains atoms.

14 D. L. Nelson, M.M. Cox, *Lehninger Principles of Biochemistry* **2017**, W.H. Freeman Company, NY; Chapter 4.

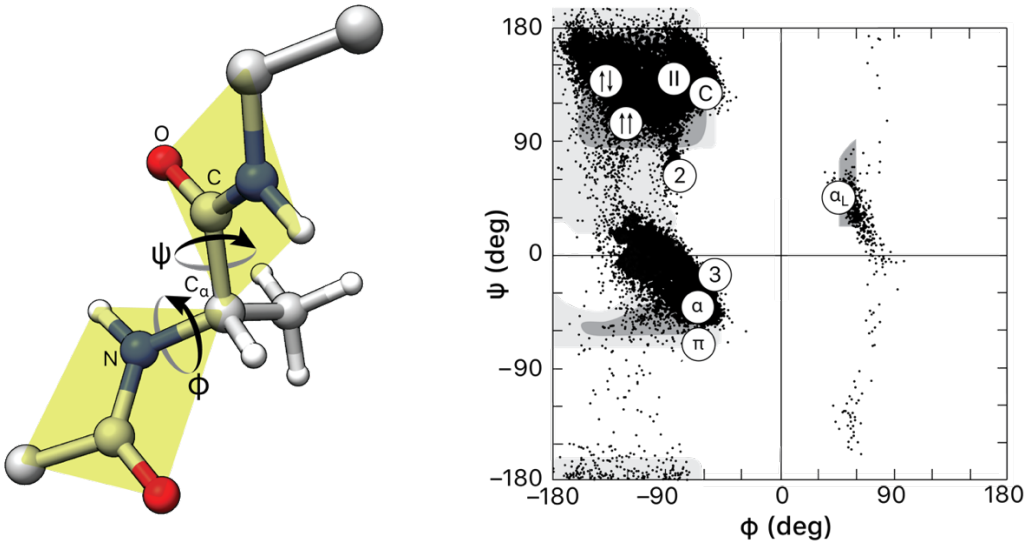


Figure 3. Left: Representation of the torsional freedom in a peptide chain. The partial double bond in the peptidic bond restricts the degrees of freedom of a peptide. Right: Ramachandran diagram showing the distribution of allowed ϕ (N–Cα) and ψ (Cα–C) conformation angles in proteins. The regions of allowed angles are shaded dark grey, whereas the light-gray regions correspond to conformations having borderline crashing van der Waals distances. The diagram also shows the distribution of angles for 25,327 residues (excluding Gly) in 207 high-resolution X-ray structures (<1.2 Å).[15]

Secondary structure elements: the α-helix and the β-sheet

The *α-helix* and the *β-sheet* were proposed by Linus Pauling and Robert Corey in 1951, before detailed structural data of proteins were even available, to explain X-ray diffraction patterns observed in some fibrous proteins and in some crystalline proteins such as hemoglobin and myoglobin; the experimental determination of the atomic structure of myoglobin a few years later in 1958 by Kendrew *et al.*[16] confirmed Pauling's chemical intuition.[17] The α-helix is the most common secondary structure element, so that about 30% of the residues in proteins are in α-helical conformation.[18] The α-helix is characterized by a particular set of dihedral angles ($\phi = -47°$ and $\psi = -57°$) that define a conformation in which the backbone of the polypeptide is wound around an imaginary axis and the side chains of the amino acids are oriented to the exterior of the helix. α-helices are right-handed, with 3.6 amino acid residues per helical turn.

The polarization of the peptide bond also has important electrostatic consequences, because the alignment of all the peptide bonds with the α-helix axis generates a net macro-dipole moment for the whole helix, with its positive pole at the N-terminus, and the negative pole at the C-terminus. The effect of the helix dipole for a helix of 20 residues

15 Adapted from D. Voet, J. G. Voet, *Biochemistry* **2011**, 4th Ed. John Wiley & Sons, New York.

16 J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, D. C. Phillips, *Nature* **1958**, *181*, 662-666.

17 D. Eisenberg, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 11207-11210.

18 D. J. Barlow, J. M. Thornton, *J. Mol. Biol.* **1988**, *201*, 601-619.

is about 63 D, which can be approximated by a 0.5-0.7 positive charge near the N-terminus and 0.5-0.7 negative charge units near the C-terminus of the α-helix.[19]

Thermodynamically, the α-helix is stabilized by hydrophobic effects resulting from the compaction of the of the peptide backbone, as well as hydrogen bonds between the N-H and C=O groups between residues i and i+4 in the main chain.[14,20] The N- and C-terminal residues in α-helices cannot form the standard (i,i+4) hydrogen bond pattern, so the amide NH hydrogen bond donors at the helix N-terminus are satisfied predominantly by local acceptors in nearby side-chains whereas carbonyl C=O groups at the C-terminus are satisfied by backbone NH groups from the sequence following the helix. This defines preferred sequences at both ends called N- and C-caps. For example, the N-terminal *Capping box*, and *Probox* motifs are characterized by the pXXp and hPXXVh residue patterns (p, polar; h, hydrophobic residue; P, proline; X, any amino acid).[21] Side chains at positions (i,i+3), (i,i+4) and (i,i+7) are close in space, and interactions between them can affect helix stability. 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, thereby, avoiding any interaction between them.

Brown and Klee reported for the first time in 1971 the formation of an α-helix by a short peptide (the C-peptide of ribonuclease A).[22] The study of this peptide is of great historic relevance, because it would reveal the importance of side chain interactions in the stabilization of α-helical conformations, including salt bridges between side chains, as exemplified by the (i,i+8) bridge between Glu^{2} and Arg^{10} across two turns of the α-helix, a salt bridge between His^{12} and Glu^{9}, and the stabilizing interaction between the helix dipole with the protonated residue His^{12} at the C-terminus of the α-helix,[23] and a cation-π interaction between this residue with Phe^{8}. The work on this peptide also revealed the relevance of Ala residues in positions 4-6 in stabilizing the α-helical conformation.

The folding of an α-helix from a random coil configuration is an entropically unfavorable process, indeed, peptides that form helices in solution are typically found as a complex mixture of all helix, all coil or, most frequently, central helices with frayed-coil ends. The conformational freedom of peptide chains can be reduced by introducing geometric constraints; the more extended strategy involves introducing bonds between distant side chains in the sequence (i.e., stapled peptides) using a variety of chemical modifications

19 a) W. G. J. Hol, Adv. Biophys. **1985**, *19*, 133-165; b) D. Sali, M. Bycroft, A. R. Fersht, *Nature* **1988**, 335, 740-743.

20 a) C. Branden, J. Tooze, *Introduction to protein structure* **1999**, Garland Publishing Inc., New York; b) N. Sewald, H-D. Jakubke, *Peptides: Chemistry and Biology* **2002**, Wiley-VCH GmbH & Co, Weinheim.

21 A. J. Doig, in *Progress in Molecular Biology and Translational Science*, ed. P. M. Conn, Academic Press, **2008**, vol. 83, pp. 1-52.

22 J. E. Brown, W. A. Klee, *Biochemistry* **1971**, *10*, 470-476.

23 K. R. Shoemaker, R. Fairman, D. A. Schultz, A. D. Robertson, E. J. York, J. M. Stewart, R. L. Baldwin, *Biopolymers* **1990**, *29*, 1-11.

including triazole stapled, hydrocarbon stapled, disulfide cross-linked, lactam cross-linked helices, as well as metal coordination.[24]

β-sheets are the secondary structure element predicted by Pauling and Corey. β-sheets are constituted by two or more short polypeptide fragments, called β-strands. It is a more extended conformation of the polypeptide chains than the α-helix and, as in the previous case, its structure is defined by a characteristic set of dihedral angles.

In β-sheets the strands are fully extended in a zig-zag form with a distance between residues of 3.5 Å. The residues of one strand are linked to the residues of a second strand *via* hydrogen bonding between N–H and C=O of the amide bonds. We can find different substructures depending on the parallel or antiparallel orientation between the adjacent strands, being the antiparallel the most stable due to the minimal distortion of the inter-chain hydrogen bonding. In this β-sheet structure the side chains are oriented perpendicularly to the plane of the sheet in an alternating pattern.[14,20] β-sheet are much less defined than α-helices, and their classification as secondary structure elements (and not as tertiary) is somewhat arbitrary, since β-sheets typically involve distant elements in the sequence.

24 a) S. Roy, P. Ghosh, I. Ahmed, M. Chakraborty, G. Naiya, B. Ghosh, *Biomedicines* **2018**, *6*, 118-134; b) Y. H. Lau, P. de Andrade, Y. Wu, D. R. Spring, *Chem. Soc. Rev.* **2015**, 44, 91-102.

Metalloproteins

General concepts

Some natural amino acids can coordinate metal ions due to the presence of coordinating groups in their side chains. Indeed, a variety of peptide models that bind metal ions have been studied.[25] In proteins metals usually appear associated to a core of polar and hydrophilic residues surrounded by a shell of non-polar groups,[26] which isolate the metal complex from water molecules. Each metal ion favors different sets of protein ligands. These preferences are frequently dictated by the hard-soft theory of acids and bases as well as the coordination number and geometry of each metal site is going to be determined by the oxidation state of the metal. Metal ions are required for the maintenance of all life forms. In fact, about one third of the structurally characterized proteins contain a metal ion in their structure.[27] These metals can play different roles inside the proteins including the transport and the storage of the metal ions, or catalytic function as enzymes. These are key in some of the most important biochemical processes, such as respiration, nitrogen fixation and oxygenic photosynthesis.[25]

Importantly, due to the unique chemical properties of each metal, a particular metal ion can be apt for different types of biological functions. For example, iron and copper are redox-active metals that can be switched between their oxidized and reduced forms, and normally they participate in electron transfer processes.[28] Despite that, they are also involved in dioxygen storage and carriage via metalloproteins such as hemoglobin, myoglobin and hemocyanin. Also, storage of these metals is facilitated in cells by specific proteins (ceruloplasmine, ferritin, transferrin). Another example apart from iron and copper is zinc that acts normally as a superacid center in different metalloproteins, promoting different chemical reactions (carbonic anhydrase, alcohol dehydrogenase).[29]

In addition to catalytic, transport, and storage functions, metal ions are involved in structural stabilization of proteins, particularly in small peptide and protein domains that cannot bury enough residues to form a large hydrophobic core that can sustain a folded structure. Binding to a metal ion cross-links specific residues within a polypeptide chain, effectively stabilizing a particular local structure.[30] A prototypical case involves binding of Zn(II) to **Cys_2His_2 zinc finger (ZF) domains**. Cys_2His_2 ZFs are short domains of about 30 residues that contain two invariant pairs of His and Cys residues.[31] The peptide chain is folded around a Zn(II) ion, which is tetrahedrally coordinated forming a small independent structural domain or module. Thus, Zn(II) is essential for the structure formation due to the Cys-Cys and His-His interaction with the metal ion. In addition to these highly conserved pairs each repeated sequence also contains other conserved amino

25 R. Jernigan, G. Raghunathan, I. Bahar, *Curr. Opin. Struct. Biol.* **1994**, 4, 256-263.
26 Yamashita, M. M.; Wesson, L.; Eisenman, G.; Eisenberg, D. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 5648-5642.
27 A. J. Thomson, H. B. Gray, *Curr. Opin. Struct. Biol.* **1992**, *2*, 155-158.
28 D. E. Fenton, *Biocoordination Chemistry*, Oxford University Press, UK, **1995**.
29 H. B. Gray, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3563-3568.
30 S. S. Krishna, I. Majumdar and N. V. Grishin, *Nucleic Acids Res.* **2003**, *31*, 532-550.
31 A. M Ginsberg, B. O. King, R. G. Roeder, *Cell* **1984**, *39*, 479-489.

acids such as Tyr[6] (or Phe[6]), Phe[17], and Leu[23], all of them hydrophobic residues that might interact to form a hydrophobic cluster that stabilizes the finger structure.[32]

Another classic example of metal-mediated folding is Calmodulin (CaM), which is a small acidic protein of approximately 17 kDa that is present in all eukaryotic cells, acting as a primary receptor for intracellular Ca(II).[33] Binding of CaM to four Ca(II) ions through its **EF-hand motifs** induces large structural changes that expose hydrophobic surfaces that enable CaM to interact with and regulate a wide variety of enzymes.[34] Babu and co–workers published the first X-ray crystallographic structure of Ca(II)-saturated mammalian calmodulin in 1986.[35] Calmodulin's Ca(II)-binding sites have a typical helix-loop-helix conformation with a number of acidic residues such as Asp and a bidentate Glu that provide hard ligands for Ca(II).[33,36]

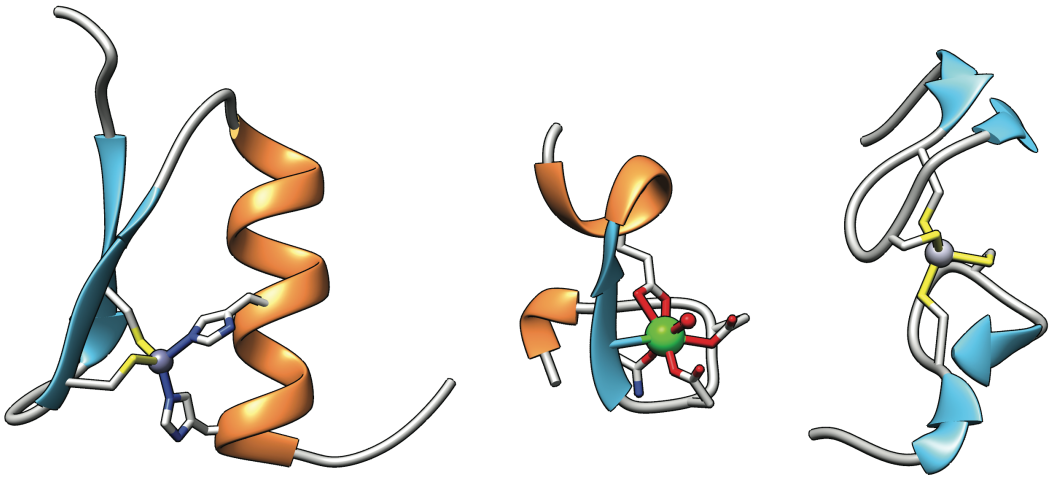


Figure 4. Left: X-ray structure of the first zinc finger of the Cys_2His_2 zinc finger Zif268 bound to the eukaryotic TATA box sequence (PDB: 1G2F);[37] center: X-ray structure of one EF-hand loop in Calmodulin (PDB: 3CLN), note the water bound to the free coordination site of Ca(II);[38] right: X-ray structure of the Rad50 coiled-coil Zn hook which pairs of conserved Cys-X-X-Cys motifs form interlocking hooks that bind one Zn(II) ion (PDB: 1L8D).[39]

In contrast with the Cys_2His_2 ZFs and CaM EF-Hands, in the **Zinc Hook (ZH)**, which mediates the homodimerization of Rad50 proteins, the coordination of the metal ion is achieved by dimerization of two short peptide loops. These two hooks feature a conserved Cys^{444}-X-X-Cys^{447} (CXXC) motif and together coordinate tetrahedrally the Zn(II) ion;

---

32 N. P. Pavletich, C. O. Pabo, *Science* **1991**, *252*, 809-817.
33 M.A. Wilson, A. T. Brunger, *J. Mol. Biol.* **2000**, *301*, 1237-1258.
34 J. J. Chou, S. Li, C. B. Klee, A. Bax, *Nat. Struc. Biol.* **2001**, *8*, 990-997.
35 Y. S. Babu, C. E. Bugg, W. J. Cook, *J. Mol. Biol.* **1988**, *204*, 191-204.
36 R. Chattopadhaya, W. E Meador, A. R Means, F. A. Quiocho, *J. Mol. Biol.* **1999**, *228*, 1177-1192.
37 S. A. Wolfe, R. A. Grant, M. Elrod-Erickson, C. O. Pabo, *Structure* **2001**, *9*, 717-723.
38 Y. Sudhakar Babu, C. E. Buggy, W. J. Cook, *J. Mol. Biol.* **1988**, *294*, 191-204.
39 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.

mutation of either of the Cys residues by Gly reduced the viability of the cells by a factor of 3–8 relative to the wild-type transformants. Interestingly, the zinc hook has been studied as an independent tag that allows the controlled dimerization with high affinity (femtomolar range).[40]

Viruses are the most numerous organisms in the biosphere. Nowadays they are exploited in an emergent array of applications such as typing bacteria,[41] peptide display,[42] and experimental phage therapy.[43] Apart from this, another important and interesting field is their application in the study of protein folding and assembly. With this aim, to solve X-ray structures of different parts of these viruses results really interesting.

Van Raaij and co-workers have solved different structures that are stabilized by the presence of coordinated metal ions inside these protein buildings. One of these structures corresponds to the long tail fiber receptor-binding tip of a Bacteriophage T4.[44] This structure showed an elaborate interwoven trimer formed by a globular domain of around 4.5 Å wide, an elongated needle domain of around 15 Å wide and a small head domain of around 25 Å. Total length of the structure is about 200 Å. The needle domain is a 150-Å-long six stranded antiparallel right-handed twist β-barrel, and it is the most interesting one in terms of coordinating properties. Inside the 15-Å-wide cylinder there are seven Fe(II) ions, each of them octahedrally coordinated by three pairs of His residues that stabilize the whole structure. Finally, each monomer forms an interwoven head, where each strain is wounded around a neighbor peptide chain.

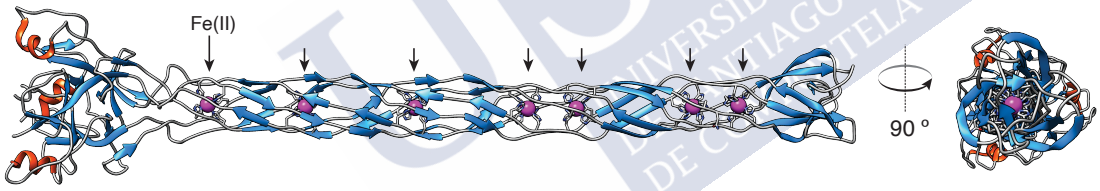


Figure 5. Structure of the receptor-binding tip of the T4 long tail fiber containing seven iron ions coordinated by His residues arranged colinearly along the core of the biological unit (ions indicated with arrows).[44]

40 T. Kochańczyk, P. Jakimowicz, A. Krężel, *Chem. Commun.* **2012**, 49, 1312-1314.

41 S. Hagens, M.J. Loessner, *Appl. Microbiol. Biotechnol.* **2007**, *76*, 513-519.

42 V. A. Petrenko, V. J. Vodyanov, *J. Microbiol Methods* **2003**, 53, 253-262.

43 a) A. Parisien, B. Allain, J. Zhang, R. Mandeville, C. Q. Lan, *J. Appl. Microbiol.* **2008**, 104, 1-13; b) A. Wright, C. H. Hawkins, E. E. Anggard, D. R. Harper, *Clin. Otolaryngol.* **2009**, 34, 349-357.

44 S. G. Bartual, J. M. Otero, C. Garcia-Doval, A. L. Llamas-Saiz, R. Kahn, G. C. Fox, M. J. van Raaij, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 20287-20292.

Artificial metallopeptides

Peptides (polymers of amino acids of less than 100 residues) represent an important group of natural molecules. Their function depends on their ability to adopt well-defined secondary structures, which are stabilized by weak inter- and intra- molecular forces such as hydrogen bonding, electrostatic interactions, disulphide bridges or metal coordination.[45] Misfolding or conformational transitions could have drastic consequences on the properties of these systems, such as in the case of Alzheimer disease where an α-helix conversion into a β-sheet leads to the aggregation of β-amyloid plaques.[46] In addition to the intramolecular covalent bonds of disulfide bridges that stabilize many small α-helix natural peptides, such as conotoxins,[47] the strength and directionality of coordination bonds can provide effective structural support to small peptidic structures. There are a number of functional groups that can act as potential ligands in peptide side chains, those that are more frequently observed playing this role are the imidazole in His, the carboxylate groups in Asp and Glu or the thiolate in Cys. Additionally, the oxygen atoms in the main chain peptidic bond can often be found within the first coordination sphere of hard metal ions. In this section we will give a brief outline of metal ion-directed stabilization of the most important protein secondary structures.

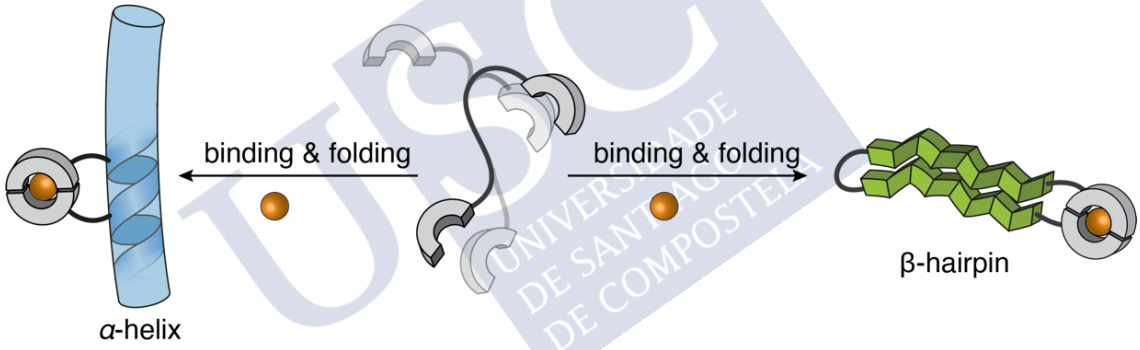


Figure 6. Metal-ion-directed secondary structure stabilization by appropriate positioning of metal chelating ligands (e.g., at positions (*i*,*i*+4) to stabilize α-helical conformation upon chelation of the metal ion).[45]

α-helix

These secondary structures are the most typical in peptides and proteins, and are involved in many protein-protein interactions, therefore mimicking them is of great practical relevance in the search of new bioactive species. There are many examples in the literature using different coordinating functional groups that can stabilize this secondary conformation through metal coordination.

45 M. Albrecht, P. Stortz, *Chem. Soc. Rev.* **2005**, 34, 496-506.

46 A. Kapurniotu, A. Buck, M. Weber, A. Schmauder, T. Hirsch, J. Bernhagen, M. Tatarek-Nossol, *Chem. Biol.* **2003**, *10*, 149-159.

47 A.-H. Jin, M. Muttenthaler, S. Dutertre, S. W. A. Himaya, Q. Kaas, D. J. Craik, R. J. Lewis, P. F. Alewood, *Chem. Rev.* **2019**, *119*, 11510-11549.

Isolated helices

Ghadiri *et al.*, showed for the first time in 1990 how to stabilize short α-helical structures through metal-ion coordination by introducing in the sequence a pair of His residues at positions (i,i+4). This arrangement places the imidazole groups at the appropriate distance to chelate transition metal ions, thereby locking the peptide backbone in a helical conformation. Furthermore, by replacing one of the His residues by a Cys, the authors could modify the coordination preferences of the peptide from Cu(II) to Cd(II).[48] The same peptide was also stabilized using $[Ru(NH_3)_4]^{2+}$ showing good helical conformation.[49]

Following this (i,i+4) principle to stabilize the helical conformation, David P. Fairlie and co-workers demonstrated that the pentapeptide Ac-HAAAH-NH_2 adopted an α-helical conformation upon coordination of $[Pd(en)]^{2+}$ in water, and studied its folding by 2D NMR.[50] This system was also studied with $[Ru(NH_3)_4]^{2+}$ and showed a higher α-helical conformation than his analogue of $[Pd(en)]^{2+}$.[51] Importantly, these coordination staple can also stabilize longer peptides, such as fragments of the protein thermolysin,[52] and can be used to stabilize helical conformations of 3-4 α-turns in oligopeptides of 10-15 amino acids length.[53]

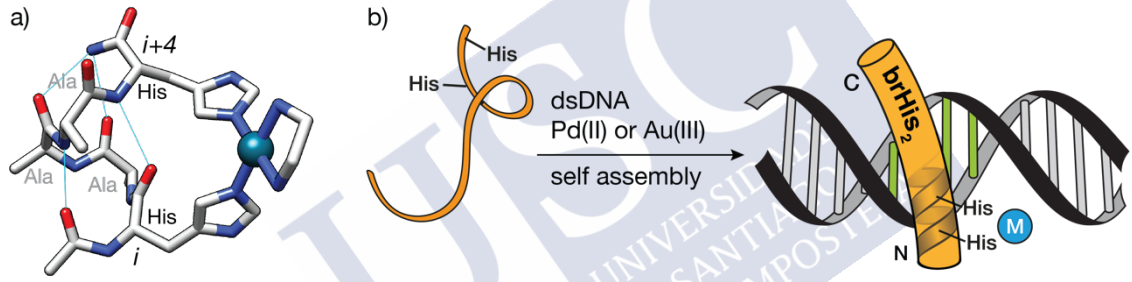


Figure 7. a) Structure of the short α-helical peptide Ac-HAAAH-NH_2 coordinated with $[Pd(en)]^{2+}$, showing intrahelical hydrogen bonds (light blue) and the coordination of the His side chains that locks the helical conformation. b) Schematic representation of a DNA-binding peptide $brHis_2$, derived from the basic region of the GCN4 transcription factor, containing two His residues to enhance its α-helical conformation upon metal binding. The stabilized peptide binds to the DNA half-site GATC.[54]

A recent example was reported by Learte-Aymamí *et al.*, where a fragment of the basic region of a bZIP transcription factor was modified to include two histidine residues at designed (i,i+4) position of its N-terminus (**$brHis_2$**). The resulting monomeric peptide is

48 M. R. Ghadiri, C. Choi, *J. Am. Chem. Soc.* **1990**, *112*, 1630-1632.

49 R. Ghadiri, A. K. Fernholz, *J. Am. Chem. Soc.* **1990**, *112*, 9633-9635.

50 M. J. Kelso, H. N. Hoang, T. G. Appleton, D. P. Fairlie, *J. Am. Chem. Soc.* **2000**, *122*, 10488-10489.

51 M. T. Ma, H. N. Hoang, C. C. G. Scully, T. G. Appleton, D. P. Fairlie, *J. Am. Chem. Soc.* **2009**, *131*, 4505-4512.

52 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.

53 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.

not able to recognize its consensus target DNA site, however, addition of $[Pd(en)]^{2+}$ promotes a high-affinity interaction with a great selectivity for this sequence. Furthermore, this interaction can be repeatedly switched on and off by addition of a chelating agent such as DEDTC or $[Pd(en)]^{2+}$ (Figure 7b).[54]

A similar approach to helix stabilization by coordination was used by Hopkins, who explored the potential of aminodiacetic (Ada) chelating side chains to induce the formation of α-helices upon metal chelation and demonstrated that the coordinative preferences of each metal are a fundamental for helix induction.[55] Some years later, this Ada chelating group was used again by Sugiura *et al.* In this case, using helices containing a pair of substituted Lys with these Ada chelators, they showed that metal-induced helix destabilization was a promising approach to functional switching, especially for helices that are intrinsically stable. By attaching this coordinative group in an (i,i+2) position in a 17-residue model peptide (Figure 8. a significant decrease in the helical content was observed after addition of Fe(III), whereas addition of Fe(II) had no influence in the stability of the helix. This modulates the helix conformation by changing the oxidation state of the Fe ions. This same approach was also tested with the transcription factor Jun and Fos leucine zippers segments.[56]

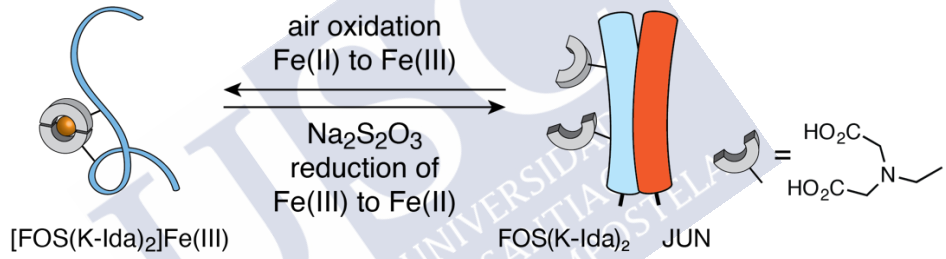


Figure 8. Schematic representation of Sugiura's destabilization model (*i*,*i*+2) position.[56]

Another coordinating unit that permits the stabilization of helices was proposed by Gilbertson *et al.* In this case, the introduction of a phosphorus as ligand allows the coordination of different metals such as Rh(I).[57] More specifically, they introduced a diphenylphosphine side chain in an artificial amino acid building block. The introduction of this ligand in a short peptide in (i,i+4) positions allowed the induction of an α-helical conformation by coordination with Rh(I).[58]

Recently, Tezcan and co-workers introduced new coordinating groups at position (i,i+7) with respect to a His residue; α-helical conformation is induced by metal coordination thanks to the length and flexibility of the linkers connecting these new coordinating groups. In their first report they described different 8-hydroxiquinoline (Quin)

54 S. Learte-Aymamí, N. Curado, J. Rodríguez, M. E. Vázquez, J. L. Mascareñas, *J. Am. Chem. Soc.* **2017**, *139*, 16188-16193.

55 F. Ruan, Y. Chen, P. B. Hopkins, *J. Am. Chem. Soc.* **1990**, *112*, 9403-9404.

56 S. Futaki, T. Kiwada, Y. Sugiura, *J. Am. Chem. Soc.* **2004**, *126*, 15762-15769.

57 S. R. Gilbertson, G. Chen, M. McLoughlin, *J. Am. Chem. Soc.* **1994**, *116*, 4481-4482.

58 S. R. Gilbertson, X. Wang, G. S. Hoge, C. A. Klug, J. Schaefer, *Organometallics* **1996**, *15*, 4678-4680.

derivatives attached to a Glu side chain; helical conformation is induced in the presence of Zn(II), Cu(II), Co(II) and Ni(II) ions as cross-linking agent. Moreover, introduction of $Re(CO)_3$-Quin derivatives lead to luminescent peptides in the visible region, showing that this coordinating motifs can introduce structure and functionality in the engineered peptides in order to study protein-protein interactions.[59] In a second work, a Cys side chain was derivatized with a chelating phenanthroline group at (*i*,*i*+7), relative to the His residue. Coordination of Zn(II), Cu(II), Co(II) and Ni(II) promotes helical conformation. This approach was then tested with a basic region of a leucine zipper, showing that the coordination stabilizes the complex with the DNA.[60]

Coiled coils

Coiled coils are a widely distributed supersecondary structure elements that perform a variety of functions in natural proteins,[61] from mediating protein-protein interactions,[62,63] to structural roles in protein filaments and chromosome organization,[64,65] and dynamic functions in motor proteins and protein machines.[66] The structural simplicity of coiled coils has allowed researchers to understand its folding, and apply this knowledge for the *de novo* design of artificial coiled coils, thus making of this protein motif an excellent platform for the development of discrete self-assembled units with tailored properties.

Coiled coils are constituted by the association of two to six α-helices that associate forming parallel or antiparallel multimeric bundles. The sequence of coiled coil peptides is characterized by a seven-residue periodicity—the *heptad* repeat $(abcdefg)_n$—with defined patterns of hydrophobic residues (e.g., Leu, Ile) in the first (*a*) and fourth (*d*) positions of the heptad; this pattern creates a hydrophobic strip along the helices that drives their oligomerization by packing their hydrophobic chains in a "*knobs-into-holes*" arrangement proposed by Crick in 1953.[67] Additionally, *e* and *g* positions are solvent-exposed, and often occupied by charged residues that define the specificity (and stoichiometry) through electrostatic interactions (Figure 9).[68]

59 S. J. Smith, K. Du, R. J. Radford, F. A. Tezcan, *Chem. Sci.* **2013**, 4, 3740-3747.

60 S. J. Smith, R. J. Radford, R. H. Subramanian, B. R. Barnett, J. S. Figueroa, F. A. Tezcan, *Chem. Sci.* **2016**, *7*, 5453-5461.

61 It has been calculated that up to 5% of the resiues in proteins are in coiled coil regions: E. Wolf, P. S. Kim, B. Berger, *Protein Sci.* **1997**, *6*, 1179-1189.

62 E. K. O'Shea, J. D. Klemm, P. S. Kim, T. Alber, *Science* **1991**, 254, 539-544.

63 J. N. Glover, S. C. Harrison, *Nature* **1995**, 373, 257-261.

64 H. Herrmann, U. Aebi, *Curr. Opin. Struc. Biol.* **1998**, 8, 177-185.

65 a) L. Truebestein and T. A. Leonard, *Bioessays* **2016**, 38, 903-916; b) 1 P. Burkhard, J. Stetefeld, S. V. Strelkov, *Trends Cell Biol.* **2001**, *11*, 82-88.

66 a) T. Koshiba, S. A. Detmer, J. T. Kaiser, H. Chen, J. M. McCaffery, D. C. Chan, *Science* **2004**, *305*, 858-862; b) M. Thormählen, A. Marx, S. Sack, E. Mandelkow, *J. Struct. Biol.* **1998**, *122*, 30-41.

67 F. H. S. Crick, *Acta Crystallogr.* **1953**, 6, 689-697.

68 P. Burkhard, J. Stetefeld, S. V. Strelkov, *Trends Cell Biol.* **2001**, *11*, 82-88.

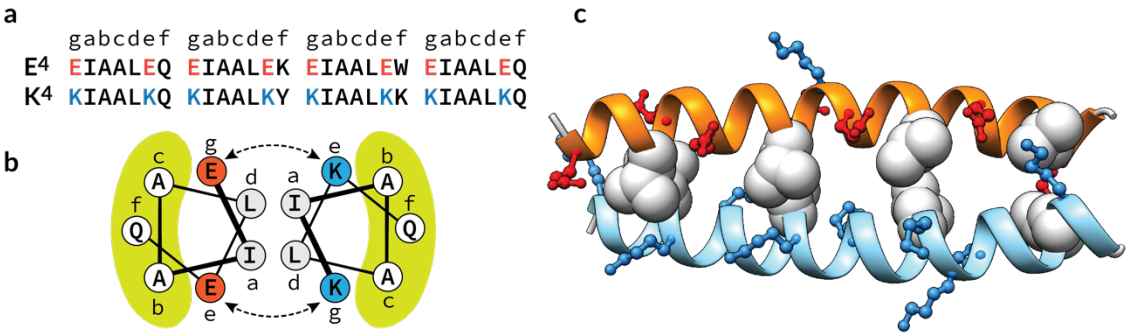


Figure 9. a) Peptide sequences of a designed coiled coil heterodimer where the heterotrimeric dimer is defined by the electrostatic interactions between Lys and Glu residues, b) its helical wheel diagram showing the arrangement of the amino acids along the axis of the α helices, and c) model structure of the heterodimer highlighting the hydrophobic Leu/Ile interface (grey), and the complementary charged residues, Glu (red) and Lys (blue).[68]

One of the first examples of metal-stabilization of coiled coil structures was proposed by Tanaka and co-workers, who designed a triple stranded parallel α-helical coiled coil, with the sequence YGG(IEKKIEA)$_4$. Replacement in this sequence of two isoleucine residues by two His residues in the third heptad lead to the formation of a coordination core in the middle of the coiled coil structure that can stabilize the three-stranded parallel supramolecular structure upon metal coordination.[69] Using the same model peptide, substitution of these Ile by an Ala and a Cys residues, results in the disruption of the triple-standed α-helix bundle, however, addition of Cd(II) or Hg(II), results again in the coiled coil formation through Cys coordination of this metals with a trigonal geometry.[70]

A more complex approach using the same previous sequence was achieved some years later. In this a new *de novo* design based on previous studies[69,71] was synthesized and tested. In this new sequence substitution of three Ile by three His residues results in the formation of the three stranded coiled coil, but in this case with two different metals ions–Cu(II) and Ni(II)–coordinated inside the hydrophobic core of the structure. First metal ion coordination promotes the formation of the coiled coil, thereby promoting the second metal ion coordination. This was the first example in literature of a peptide able to bind two different metal ions in its structure.[72]

An inverse approach was proposed by Futaki and co-workers. In this case, instead of stabilizing the helix conformation, they proposed to destabilize it in order to modulate the DNA binding properties of a leucine zipper. To achieve this, methodology proposed by Sugiura *et al.* was followed,[56] by introducing a chelating group in two amino acid side

69 K. Suzuki, H. Hiroaki, D. Kohda, H. Nakamura, T. Tanaka, *J. Am. Chem. Soc.* **1998**, *120*, 13008-13015.

70 X. Li, K. Suzuki, A. Kashiwada, H. Hiroaki, D. Kohda, T. Tanaka, *Protein Sci.* **2000**, *9*, 1327-1333.

71 T. Kiyokawa, K. Kanaori, K. Tajima, M. Koike, T. Mizuno, J.-I. Oku, T. Tanaka, *J. Pept. Res.* **2004**, *63*, 347-353.

72 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.

chains placed in positions (i,i+2) of the peptide sequence. These amino diacetic acid chelating units were introduced via disulphide bridge in Cys side chains to construct a building block that was integrated in the leucine zipper segment of the protein. Addition of Co(II) to the solution containing this peptide results in the partial losing of the helical conformation, showing also worse DNA binding properties as a result of the distortion of the secondary structure of the protein (Figure 10). [73]

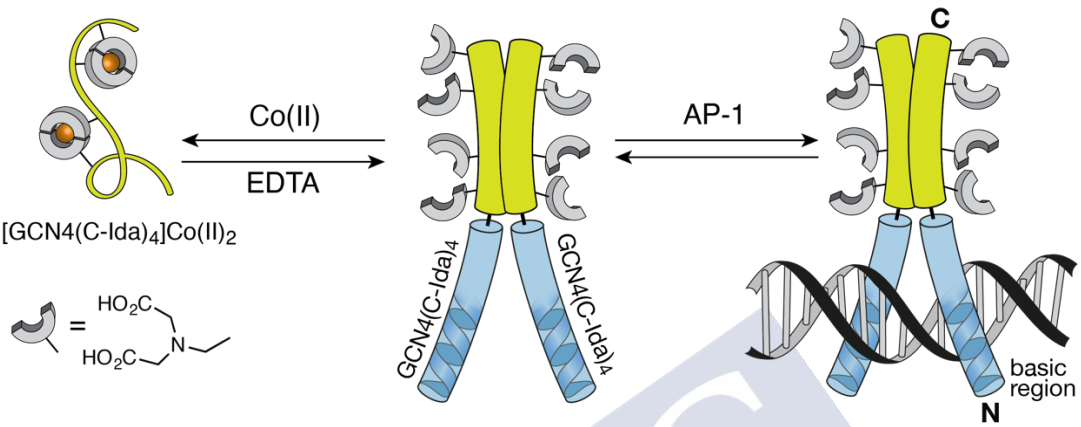


Figure 10. Model proposed by S. Futaki to control DNA binding of an engineered bZIP using Ida chelating units to promote α-helical conformations.[73]

The formation of metal complexes has also been exploited in combination with orthogonal stimuli, such as redox inputs, for the development of more sophisticated switches. Thus, in 2013 the group of J. L. Mascareñas described a DNA binding peptide based on the GCN4 transcription factor that contained both a metal-chelating 5,5″-dimethyl-2,2′:6′,2′-terpyridine ligand attached to the side chain of a C-terminal Lys[231] residue, and a redox-sensitive Cys residue on its *N*-terminus (**Cys)GCN4(tpy)** that modified its DNA binding preferences in response to specific chemical inputs (Figure 11).[74] Incubation of this bifunctional peptide with Ni(II) ions in the presence of the ATF/CREB site (5'–ATGA cg TCAT–3') gave rise to a complex between the Ni(II)-mediated dimer of the GCN4 fragment [**(Cys)GCN4(tpy)**]$_2$Ni(II) and the DNA, but under the same conditions, this peptide did not display significant affinity for the inverted sequence (5'–TCAT cg ATGA–3'). Remarkably, the sequence selectivity was reversed when the Cys residues were oxidized with DTNB, 5,5'-dithiobis-(2-nitrobenzoic acid), so that the disulfide dimer **(SS)[GCN4(tpy)]$_2$** preferentially bound to the inverted sequence (5'–TCAT cg ATGA–3'). Treatment of the **(SS)[GCN4(tpy)]$_2$**/DNA complex with *tris*-(2-carboxyethyl) phosphine (TCEP) to reduce the disulfide bond, resulted in the disassembly of the complex with the DNA.

73 Y. Azuma, M. Imanishi, T. Yoshimura, T. Kawabata, S. Futaki, *Angew. Chem. Int. Ed.* **2009**, 48, 6853-6856.

74 J. Mosquera, A. Jiménez-Balsa, V. I. Dodero, M. E. Vázquez, J. L. Mascareñas, *Nat. Commun.* **2013**, 4, 1874-1878.

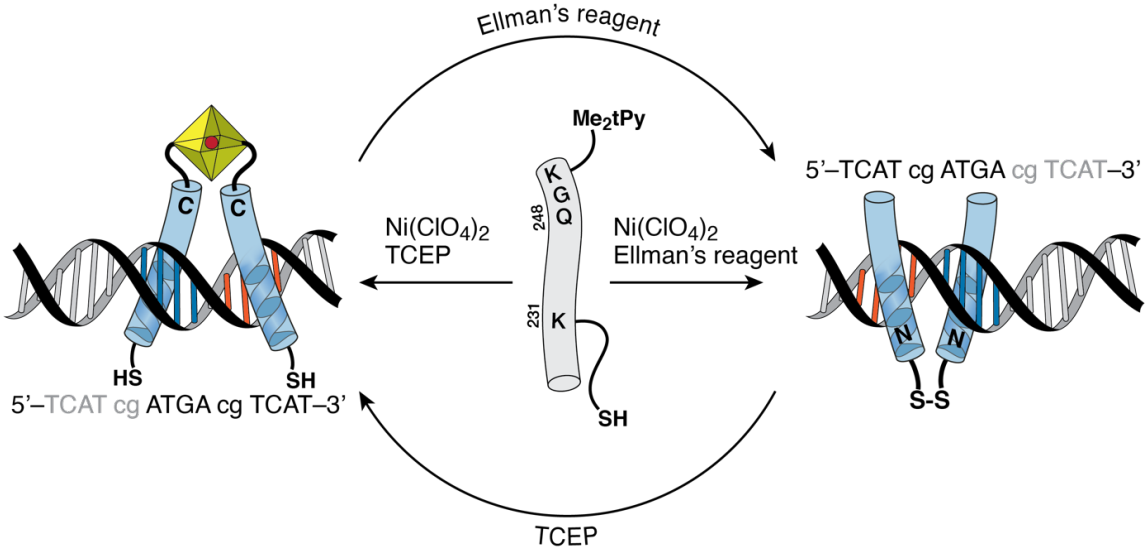


Figure 11. Stimuli-responsive DNA binding with a GCN4 peptide fragment. The (Cys)GCN4(tpy) peptide (center) can be oxidized with Ellman's reagent to afford a disulfide-mediated dimer (SS)[GCN4(tpy)]$_2$ (on the right hand side of the image), which binds to the inverted sequence 5'–TCAT cg ATGA–3' in the presence of Ni(II) ($K_D \approx 758$ nM); reduction of this peptide with TCEP yields an alternative dimer mediated by the terpyridine-Ni(II) complex (left) that binds with high affinity ($K_D \approx 670$ nM) to the consensus DNA (5'-ATGA cg TCAT-3'). The disulfide dimer can be recovered by oxidation with Ellman's reagent.[74]

A two-stranded coiled coil stabilized by metal chelation was described by Hodges *et al.* Specifically, they modified a model coiled coil sequence (VGALQ^{e}KQ^g)$_5$ with two Glu residues strategically placed in the peptide sequence in positions *g* and *e* of the center heptad. The carboxylate side chains in these two residues would introduce ionic repulsions at the coiled coil interface, destabilizing the association, but create a high-affinity metal binding site that would stabilize the formation of the coiled coil structure in the presence of Ln(III), Yb(III) or Zn(II).[75]

Conticello and co-workers reported the metal-ion-directed self-assembly of a three stranded coiled coil. Their system was based on a previous design in which the hydrophobic core of the three-stranded coiled coil featured His residues that in their protonated form would destabilize the association, but at high pH would allow the formation of fibers.[76] They envisioned that the His residues also create trigonal planar coordinating sites appropriate for the coordination of Ag(I) ions. Indeed, addition of this metal ion to a solution of the peptide results in a transition from random coil to an α-helical conformation.[77]

The group of Vincent L. Pecoraro has studied coiled coils as platforms for metal ion coordination for many years. An example of their work is the modification of the model leucine zipper sequence TRI, Ac-G(LKALEEK)$_4$G-NH$_2$, replacing the Leu in position 9

75 W. D. Kohn, C. M. Kay, B. D. Sykes, R. S. Hodges, *J. Am. Chem. Soc.* **1998**, *120*, 1124-1132.

76 Y. Zimenkov, S. N. Dublin, R. Ni, R. S. Tu, V. Breedveld, R. P. Apkarian, V. P. Conticello, *J. Am. Chem. Soc.* **2006**, *128*, 6770-6771.

77 S. N. Dublin, V. P. Conticello, *J. Am. Chem. Soc.* **2008**, *130*, 49-51.

with a L-Pen residue (L-β,β-dimethylcysteine) and Ile 23 with a His. This results in the obtention of a triple-stranded coiled coil with two different binding sites for different metals. His_3 coordination site is able to bind Zn(II) while L-Pen_3 is able to accommodate Hg(II) ions. This structure mimics natural carbonic anhydrase (AC) where we also can find a Zn(II) active site. HgS_3 site has a key role in the structural maintenance of the structure. This construction showed really good activity, occurring CO_2 hydration with an efficiency of around 500-fold better than in ACII. Also in the absence of Zn(II), it showed a good activity in the hydrolysis of the *p*-nitrophenylacetate.[78] Similarly, they redesigned a triple stranded coiled coil reported by DeGrado and co-workers,[79] introducing four residues of Cys, so that, they could reproduce a rubredoxin metal binding site. This architecture can truly mimic the natural protein conformation around the metal ion, and it could even function similarly.[80]

Another *de novo* coiled coil was proposed by Peacock *et al*. later. In this case a new peptide sequence (MB1) was designed to self-assemble in a three-stranded metallo-coiled coil in the presence of trivalent lanthanide cations. They engineered the metal binding site in the hydrophobic core of a model coiled coil Ac-G(I^aAAI^dEQK)$_5$G-NH_2, in which Ile residues at positions *a* and *d* of the repeating heptad are responsible for the generation of the hydrophobic core and favor the coiled coil formation, Ala residues in *b* and *c* positions are helix inducing, and interhelical salt bridges established between Glu and Lys residues placed at positions *e* and *g* stabilize the trimeric structure through ionic interactions. In order to make the final peptide able to sequester Ln(III) ions, two adjacent Ile residues were replaced with an Asn and an Asp. The Asn was introduced at the *d* site in the second heptad (position 12) because of the importance of these positions in the formation of three-stranded coiled coils. Introduction of the Asp in the *a* position directly below (residue 16) completes the Ln(III) binding site. Finally a tryptophan was introduced in the f position adjacent to the Ln(III) binding site to sensitize lanthanide luminescence. In the absence of Ln(III) the negative charges on the Asp residues repel each other, destabilizing the coiled coil. Addition of $GdCl_3$ resulted in an enhancement of the helicity. To gain a deeper knowledge of the coordination sphere, lifetime decay of the Tb(III) luminescence was measured revealing that there were no water molecules coordinated to the Tb(III). On the other hand, the Gd(III) coiled coil showed promising magnetic resonance capabilities (Figure 12).[81] A follow-up paper revealed that displacing the metal binding site across the coiled coil scaffold results in peptide architectures with different stabilities and coordination properties, thus allowing the modification of the inner coordination sphere of the lanthanide ions, which influences emissive and magnetic properties.[82]

---

78 M. L. Zastrow, A. F. A. Peacock, J. A. Stuckey, V. L. Pecoraro, *Nat. Chem.* **2012**, 4, 118-123.

79 J. W. Bryson, J. R. Desjarlais, T. M. Handel, W. F. Degrado, *Protein Sci.* **1998**, 7, 1404-1414.

80 A. G. Tebo, T. B. J. Pinter, R. García-Serres, A. L. Speelman, C. Tard, O. Sénéque, G. Blondin, J.-M. Latour, J. Penner-Hahn, N. Lehnert, V. L. Pecoraro, *Biochemistry* **2018**, 57, 2308-2316.

81 M. R. Berwick, D. J. Lewis, A. W. Jones, R. A. Parslow, T. R. Dafforn, H. J. Cooper, J. Wilkie, Z. Pikramenou, M. M. Britton, A. F. A. Peacock, *J. Am. Chem. Soc.* **2014**, 136, 1166-1169.

82 M. R. Berwick, L. N. Slope, C. F. Smith, S. M. King, S. L. Newton, R. B. Gillis, G. G. Adams, A. J. Rowe, S. E. Harding, M. M. Britton, A. F. A. Peacock, *Chem. Sci.* **2015**, 7, 2207-2216.

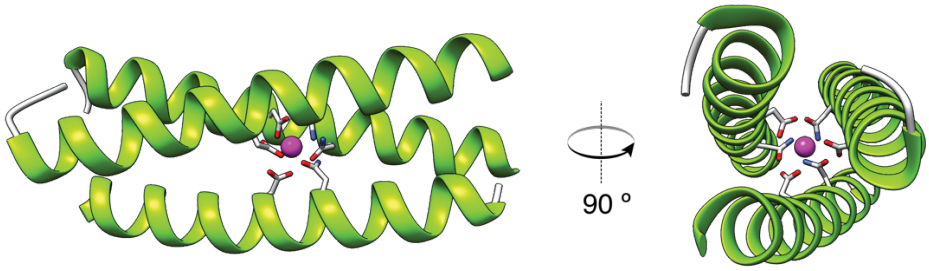


Figure 12. Model of the three-stranded coiled coil reported by Peacock *et al.* designed by mutation of the central Ile residues by coordinating Asn and Asp residues that create a coordination environment of a Tb(III) ion.[81]

Beyond the use of coiled coil scaffolds, He and co-workers reported a uranium-selective peptide with femtomolar binding constants based on a three α-helix bundle from a thermostable protein of *Methanobacterium thermoautotropicum*. A computational screening revealed that this protein could accommodate hexagonal bipyramid or pentagonal bipyramid uranyl binding site by introducing three mutations (Figure 13).[83]

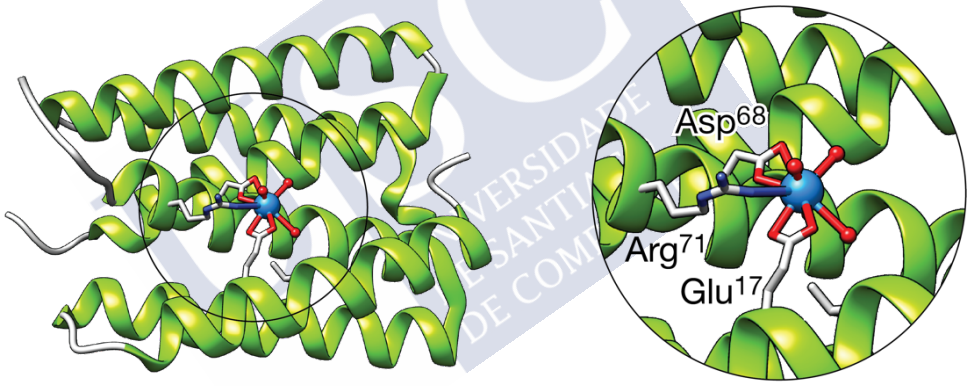


Figure 13. Structure of uranyl-protein complex (PDB ID 4FZP) showing the metal binding site in the surface of the four-helix bundle. Detail of the binding site on the right, highlighting the coordinating side chains.[83]

β-sheet and β-turns

β-sheets and hairpins structures are not so abundant as helices, but they are also fundamental secondary structure elements involved in many protein-protein interactions (e.g., WW domains that mediate recognition of prolyproline sequences).[84] There are fewer examples in literature reporting β-sheet structures with metal ions, which is not surprising, because of the inherent difficulties of the *de novo* design of β- structures, as well as the experimental difficulties associated with their study (i.e., their known tendency to aggregate in solution).

---

83 L. Zhou, M. Bosscher, C. Zhang, S. Ozçubukçu, L. Zhang, W. Zhang, C. J. Li, J. Liu, M. P. Jensen, L. Lai, C. He, *Nat. Chem.* **2014**, *6*, 236-241.

84 M. J. Macias, V. Gervais, C. Civera, H. Oschkinat, *Nat. Struct. Biol.* **2000**, *7*, 375-379.

One of the first examples of a metal-binding β-hairpin was proposed by Searle *et al.*, who describe two peptides His_2-b (K**H**YTVSINGKKITV**H**I) and His_3-b (**H**K**H**YTV-SINGKKITV**H**I), that bind Zn(II) through the His residues in the sequence. In absence of Zn(II), these peptides display a weak signal at 216 nm in CD experiments, which indicates an equilibrium between random coil and folded β-hairpin conformations. Addition of Zn(II) induces a more intense minimum at 216 nm (particularly for His_3-b), which suggests the formation of a more stable complex, probably due to the presence of the third His residue that saturates better the coordination sphere of the Zn(II) ion.[85]

He and co-workers described the assembly of peptidic nanoparticles by tagging bovine carbonic anhydrase (BCA) with a small peptide sequence that is able to form β-sheets and promote the self-assembly of these tagged proteins into large nanoparticles. The interaction between these peptide fragments can be controlled by changing the pH or promoting the chelation of Mg(II) cations; combining these two parameters (pH and ionic strength) the researchers were able to control the size of the assembled nanoparticles (Figure 14) .[86]

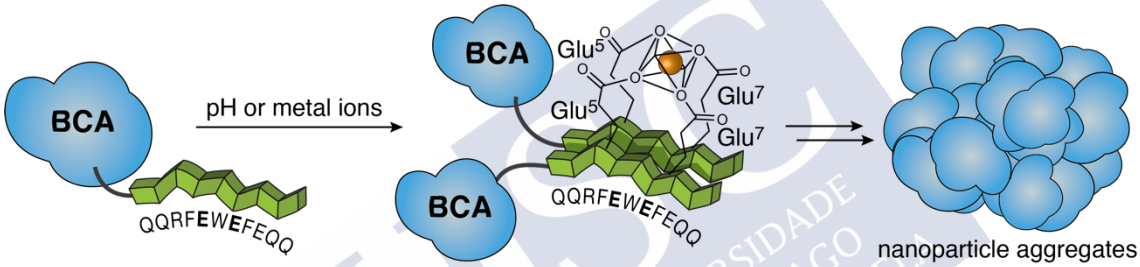


Figure 14. Schematic representation of the self-assembly of the BCA-tagged protein mediated by Mg(II) chelation of designed β-sheets.[86]

The group of José L. Mascareñas reported the rational modification of a natural RNA-binding peptide by coupling its metal-chelating properties with the formation of specific complexes with RNA. Thus, the peptide contains an EDTA chelator at the N-terminus, and a phenanthroline, which acts as an antenna, both coordinating and sensitizing the Ln(III) ions, at its C-terminus; once the peptide is folded into a β-hairpin upon RNA binding, the phenanthroline and EDTA are brought together, thus creating the coordination environment for Eu(III) or Tb(III) ions (Figure 15).[87]

Fujita and co-workers recently described the self-assembly of a β-barrel with a *de novo* designed peptide with two metal binding pyridines at both the N- and C-terminus, connected through a canonical PGP turn sequence to a short internal section (FVFV) that mediates the association into β-sheets.[88]

85 G. Platt, M. S. Searle, C.-W. Chung, *Chem. Commun.* **2001**, 1162-1163.
86 B. K. Shanbhag, C. Liu, V. S. Haritos, L. He, *ACS Nano* **2018**, *12*, 6956-6967
87 C. Penas, J. L. Mascareñas, M. E. Vázquez, *Chem. Sci.* **2016**, *7*, 2674-2678.
88 M. Yamagami, T. Sawada, M. Fujita, *J. Am. Chem. Soc.* **2018**, *140*, 8644-8647.

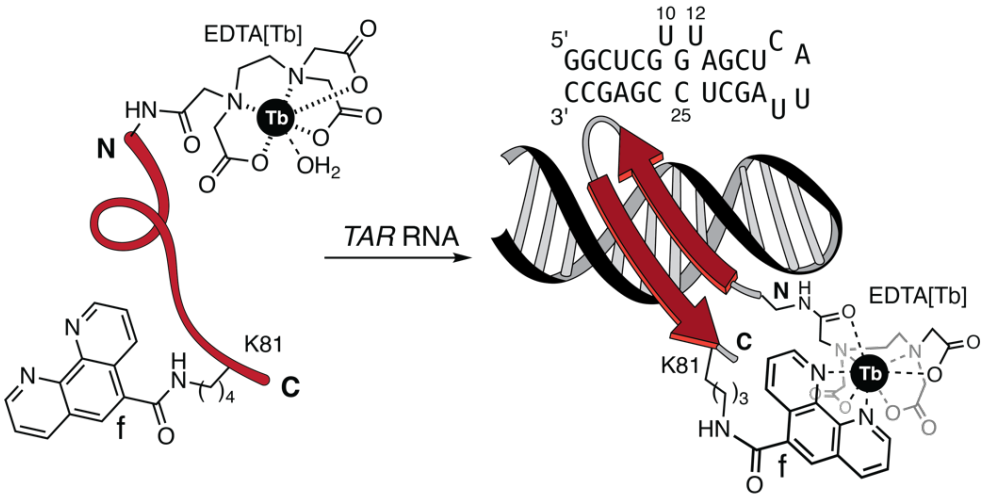


Figure 15. Schematic representation of RNA binding, β-sheet formation and Tb(III) coordination.[86]

Albrecht and co-workers proposed a new method for the conformational fixation of bioactive loop-type peptide structures. In order to simulate the loop-like conformation of the natural peptides Segetalin A and B, Albrecht *et al.* attached at the N and C termini of the sequence a coordinating unit derived from catechol that coordinated molybdenum(VI) dioxo. The resulting metallocyclic peptide retained the loop conformation in a very similar way than the one that the natural product displays.[89]

Another approach to obtain these loops or turns was proposed by Imperiali and co-workers, but in this case only by using natural amino acids. For this study three different peptides were prepared. In these sequences, there is just a change between peptide **1** and **2** in the coordinating units (His^1 instead of Glu^1), while the peptide **3** has a completely different composition in the loop sequence. CD studies showed that only peptides **1** and **2** display a β-turn-like spectra, almost superimposable in both cases, while CD spectra of peptide **3** was different. Furthermore, addition of Zn(II) ions resulted in an enhancement of the maximum at 206 nm in both peptides 1 and **2**, suggesting a strong stabilization of the secondary structure. In the case of the peptide **3**, no changes are observed in the secondary structure, demonstrating that the absence of the correct turn sequence led to an unstructured disposition.[90]

89 M. Albrecht, P. Stortz, P. Weis, *Supramol. Chem.* **2003**, 15, 477-483.

90 B. Imperiali, T. M. Kapoor, *Tetrahedron*, **1993**, 49, 3501-3510.

Metallopeptides derived from 2,2'-Bipyridine

2,2´-Bipyridine (Bpy) is one of the most widely used ligands in coordination chemistry.[91] Coordination compounds derived from Bpy, usually, present high association constants due to the chelating effect of the two nitrogen donors and the capacity of the bipyridine ligand to act as a π-acceptor.[92] These ligands show a frontier Lewis base character, so that they can form stable complexes with almost all transition metals, some of them kinetically inert and with interesting photophysical properties. Due to the coordinative preferences of the Bpy compared with natural amino acid side chains, Bpy has been extensively used as ligand in artificial metallopeptides. Introduction of this coordinative group can be achieved with different strategies depending on its position within the peptide chain.

N-capping bipyridines

N-capping with Bpy monosubstituted with a carboxylic acid (Bpac)

The Bpy unit can be monosubstituted in four different positions of the aromatic ring with a carboxylic acid group to afford derivatives that can be integrated into standard solid-phase peptide synthesis protocols and easily introduced at the *N*-terminus of peptide sequences. We can find many examples in the literature of peptide ligands derived from these Bpy derivatives, being 4Bpac and 5Bpac the most common (Figure 16).[93,94]

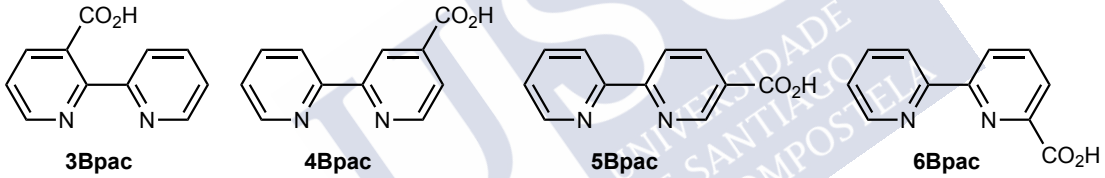


Figure 16. Bpy derivatives monosubstituted with a carboxylic acid.

In 1992 Ghadiri *et al.* published one of the first examples of an artificial metallopeptide containing one of these monosubstituted bipyridine ligands; more specifically a 15-mer peptide featuring a 5Bpac unit in its *N*-terminus. This peptide was initially in a random coil conformation in solution, but addition of divalent metal ions, such as Fe(II), Ni(II)

91 W. W. Brandt, *Chem. Rev.* **1954**, 54, 959-1011.

92 F. A. Cotton, G. Wilkinson, *Advanced Inorganic Chemistry*, **1988**, Jonh Wiley & Sons.

93 a) W. S. W. Aldridge, B. J. B. Hornstein, S. S. Serron, D. M. D. Dattelbaum, J. R. J. Schoonover, T. J. T. Meyer, *J. Org. Chem.* **2006**, *71*, 5186-5190; b) T. Koide, M. Yuguchi, M. Kawakita, *J. Am. Chem. Soc.* **2002**, *124*, 9388-9389; c) D. G. McCafferty, B. M. Bishop, C. G. Wall, S. G. Hughes, *Tetrahedron* **1995**, *51*, 1093-1106; d) D. J. Wilger, S. E. Bettis, C. K. Materese, M. Minakova, G. A. G. 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**, 35, 4869-4883.

94 a) L. Roy, M. A. Case, *J. Phys. Chem. B*, **2011**, 115, 2454-2464; b) A. J. Doerr, G. L. McLendon, *Inorg. Chem.* **2004**, 43, 7916-7925; c) M. H. Filby, J. Muldoon, S. Dabb, N. C. Fletcher, A. E. Ashcroft, A. J. Wilson, *Chem. Commun.* **2011**, *47*, 559-561; d) 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.

or Ru(II), induced its folding into an α-helical conformation and assembly of a triple stranded coiled coil.[95]

N-capping with Bpy disubstituted with two carboxylic acids (Bpda)

Bipyridine dicarboxylic acids (Bpda) have been widely used in coordination chemistry and, by extension, in the design of artificial metallopeptides. A number of examples with all the possible symmetric isomers (except the 3,3'-disubstituted) have been described in the literature over the years in which these Bpda ligands are used as a bridge or connector between two peptide chains (Figure 17).[96,97] A collaboration of our research group with that of J.-C. G. Bünzli reported in 2010 the application of 5,5′-Bpda for the construction of heterometallic Ln(III)-Ru(II) complexes in which the central 5,5′-Bpda ligand coordinates the Ru(II) ion, while both peptide arms arising from the Bpy unit are derivatized with a terminal DOTA macrocycle to coordinate the lanthanide ions.[98]

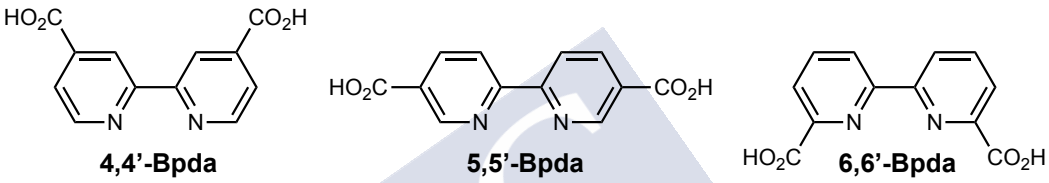


Figure 17. 2,2'-bipyridines disubstituted with carboxilyc acid groups.

Another example of this approach was reported by Yashima *et al.*, who synthesized a dendritic metallopeptide with a $[Fe(5,5'\text{-Bpda})_3]^{2+}$. The chirality of this system was determined by the chirality of the amino acids employed in the peptide sequence.[99]

95 M. R. Ghadiri, C. Soares, C. Choi, *J. Am. Chem. Soc.* **1992**, *114*, 825-831

96 a) K. H. Chang, J. H. Liao, C. T. Chen, B. K. B. 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. L. Everitt, H. Ishida, M. Kunieda, M. G. B. Drew, *Inorg. Chem.* **2000**, 39, 308-316; e) H. Ishida, Y. Inoue, *Biopolymers*, **2000**, 55, 469-478; f) J. G. Vos, J. M. Kelly, *Dalton Trans.* **2006**, 35, 4869-4883; g) H. Mürner, P. Belser, A. von Zelewsky, *J. Am. Chem. Soc.* **1996**, *118*, 7989-7994; h) S. Watanabe, O. Onogawa, Y. Komatsu, K. Yoshida, *J. Am. Chem. Soc.* **1998**, *120*, 229-230.

97 a) R. Schobert, B. Biersack, S. Knauer, M. Ocker, *Bioorg. Med. Chem.* **2008**, *16*, 6-9; b) D. R. Ahn, T. W. Kim, J. I. Hong, *J. Org. Chem.* **2001**, *66*, 5008-5011; c) E. C. Constable, E. L. Dunphy, C. E. Housecroft, M. Neuburger, S. Schaffner, F. Schaper, S. R. Batten, *Dalton Trans.* **2007**, *36*, 4323-4332.

98 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**, *2010*, 4532-4545.

99 N. Ousaka, Y. Takeyama, H. Iida, E. Yashima, *Nat. Chem.* **2011**, 3, 856-861.

Internal bipyridines

Unfortunately, the Bpy ligands described so far can only terminate the peptide chain; to introduce a Bpy ligand in the middle of a peptide, it has to be derivatized as an amino acid. This offers important advantages, such as better integration of this amino acid in the sequence and greater flexibility in the design of the metallopeptide architectures.

Bpy side chains (Bpa)

2,2´-Bpy can be functionalized in positions 4, 5 or 6 of the aromatic ring, so that it is possible to obtain amino acids equipped with the Bpy group in their side chain (Figure 18).[100] This methodology can also be applied to other polypyridine ligands.[101] However, the stereoselective synthesis of bipyridine amino acids is far from straightforward, which has hampered its application in the design of artificial metallopeptides.[102] An early example of the use of these amino acids was reported by the group of B. Imperiali, who described the synthesis and application of 6Bpa (Figure 18) to stabilize β-hairpin structures through metal coordination .[103]

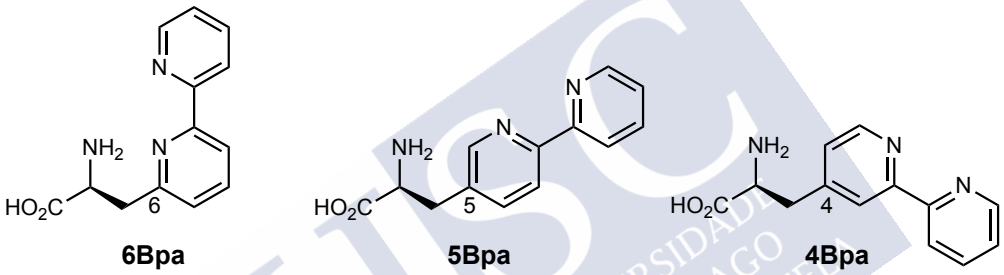


Figure 18. Amino acids equipped with Bpy in the side chain.

Asymmetrically substituted Bpy (acid 5´-amino-5-carboxybipyridine)

Introduction of the Bpy ligand in the peptide main chain ensures a tight coupling between the conformational preferences of the peptide and the coordination of the metal ion; this has been achieved with the amino acids 5Bpy and 4Bpy (Figure 19). [104,105]

---

100 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.

101 N. D. Jabre, T. Respondek, S. A. Ulku, N. Korostelova, J. J. Kodanko, *J. Org. Chem.* **2010**, 75, 650-659.

102 a) 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; b) N. Park, J. Ryu, S. Jang, H. S. Lee, *Tetrahedron* **2012**, 68, 4649-4654; c) T. H. D. Nguyen, K. Ozawa, M. Stanton-Cook, R. Barrow, T. Huber, G. Otting, *Angew. Chem. Int. Ed.* **2011**, 50, 692-694; d) H. S. H. Lee, P. G. P. Schultz, *J. Am. Chem. Soc.* **2008**, 130, 13194-13195; e) J. J. Xie, W. W. Liu, P. G. P. Schultz, *Angew. Chem. Int. Ed.* **2006**, 46, 9239-9242; f) 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.

103 R. P. Cheng, S. L. Fisher, B. Imperiali, *J. Am. Chem. Soc.* **1996**, 118, 11349-11356.

104 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.

105 G. R. Newkome, J. Gross, A. K. Patri, *J. Org. Chem.* **1997**, 62, 3013-3014.

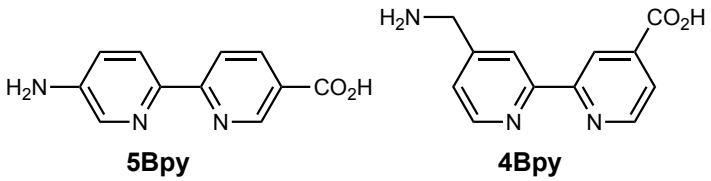


Figure 19. Structures corresponding to the two different bipyridines substituted in different positions with both an amino and an acid group.

5Bpy was firstly reported in 1996 by the Imperiali group (Figure 20). The key step in their synthetic route is a Stille coupling of two pyridine derivatives, one with a nitro group and another with a methyl ester. Hydrogenation in the presence of Pd/C afforded the corresponding amine, which after a series of transformations gave the desired Fmoc-amino acid suitable for solid-phase peptide synthesis.[106]

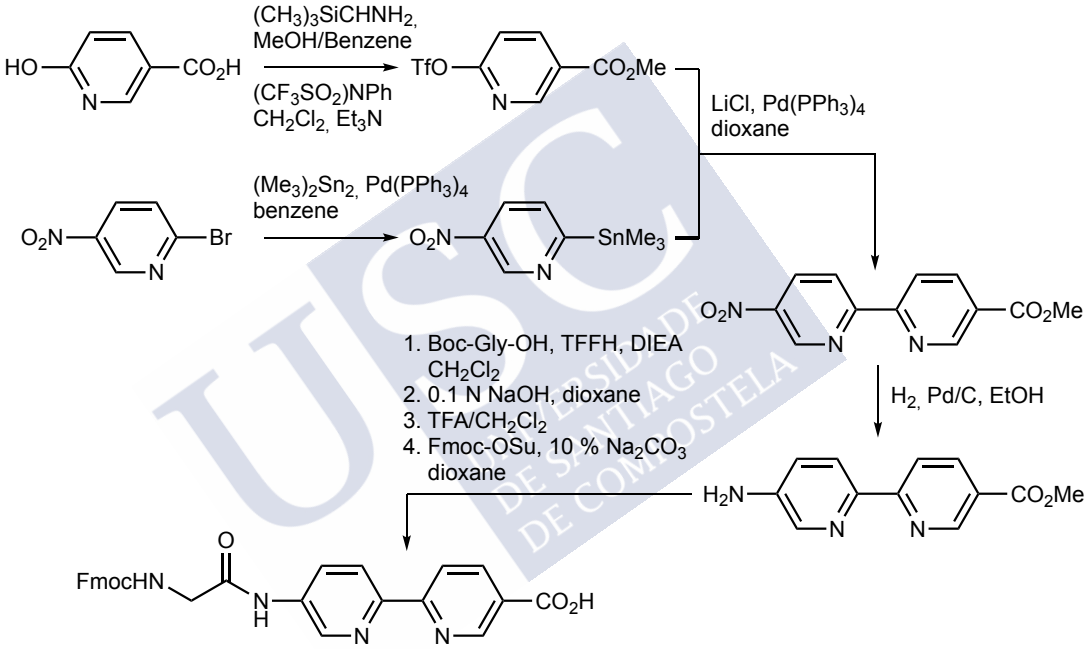


Figure 20. Synthetic scheme followed by B. Imperiali *et al*. to access the Fmoc-protected bipyridine building block for solid-phase peptide synthesis of metal-chelating peptides.[106]

These coordinating peptides are equipped with a fluorophore in their N-terminus and a quencher in the C-terminus. As a result of metal ion coordination, a conformational change in the peptide leads to the quenching of the fluorophore emission.[106] A few years later, George R. Newkome developed a new synthetic route for the obtention of the 5Bpy (Figure 21). The key step in their synthesis was the desymmetrization of a diethyl [1,1'-biphenyl]-4,4'-dicarboxylate intermediate with hydrazine monohydrate under conditions that allow the selective precipitation of the monocarbohydrazide, which is oxidized into the corresponding acyl azide, and then transformed into a carbamate through a Curtius rearrangement. Simultaneous hydrolysis of the carbamate and the ester group gives the desired bipyridine amino acid in a 21% overall yield.[107]

106 A. Torrado, B. Imperiali, *J. Org. Chem.* **1996**, *61*, 8940-8948.
107 G. R. Newkome, J. Gross, A. K. Patri, *J. Org. Chem.*, **1997**, *62*, 3013-3014.

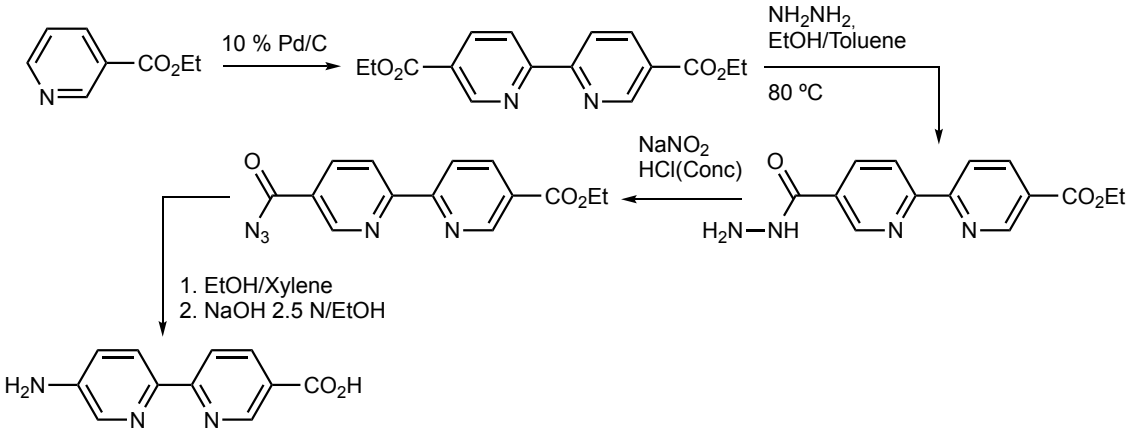


Figure 21. Synthetic route to 5Bpy proposed by Newkome *et al.* [107]

More recently, T. Darbre's group designed a new synthetic route for the obtention of a Fmoc-protected 5Bpy amino acid for its use in solid-phase peptide synthesis (Figure 22). Darbre started the synthetic pathway from the 5,5´-dimethyl-2,2´-bipyridine, which was oxidized and esterified before desymmetrization by selective deprotection of one of the ethyl esters. The resulting mono carboxylate was then reacted with azide. From this point onwards the synthetic route followed the procedure developed by Newkome.[108]

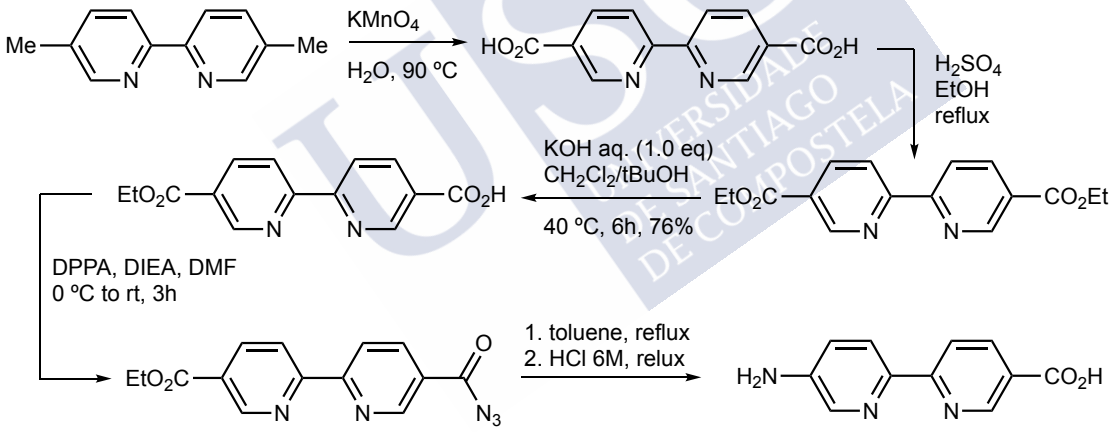


Figure 22. Synthetic route proposed by T. Darbre and co-workers.[108]

5Bpy was used in the synthesis of metallopeptides with interesting properties.[109] Ishida *et al.* used this amino acid for the construction of luminescent Ru(II) tris-bipyridyl complexes, including folded peptide structures induced by short β-turn sequences. [110]

Our group has used the 5Bpy building block extensively during the past few years for the construction of DNA-binding polypyridyl Ru(II) and Ir(III) metallopeptides. Thanks to

---

108 N. A. Uhlich, P. Sommer, C. Bühr, S. Schürch, J.L. Reymond, T. Darbre, *Chem. Commun.* **2009**, 45, 6237-6241.

109 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**, 9, 1307-1320.

110 a) H. Ishida, Y. Inoue, *Biopolymers* **2000**, 55, 469-478; b) H. Ishida, Y. Maruyama, M. Kyakuno, Y. Kodera, T. Maeda, S. Oishi, *ChemBioChem* **2006**, 7, 1567-1570.

the synthetic versatility of the SPPS, structural modifications were easily introduced in the structure of the peptide ligands in order to improve their DNA binding properties and cell internalization.[111] We have also used 5Bpy for the construction of mononuclear bisbipyridine and trisbipyridine Fe(II), Co(II), Ni(II) and Zn(II) chiral metallopeptides.[112]

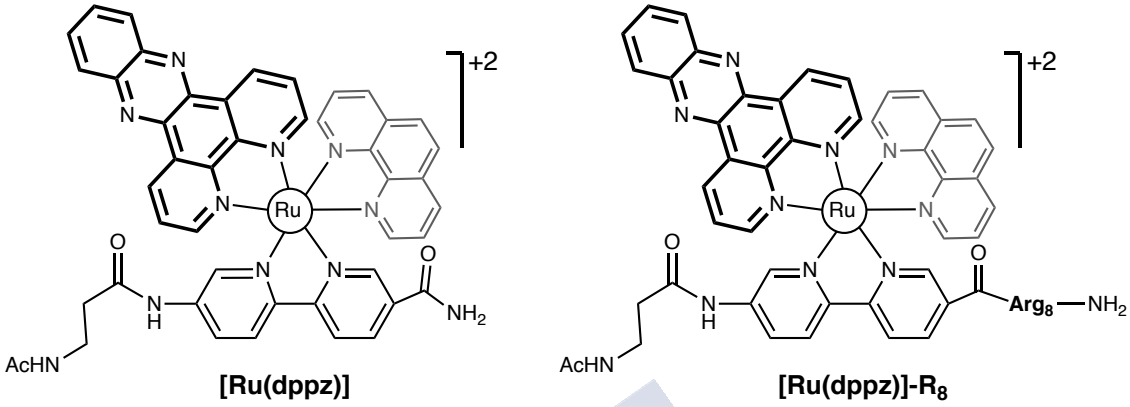


Figure 23. Polyarginine Ru(II) metallopeptides derived from 5Bpy and equipped with two dppz ligand units ($[Ru(5Bpy\text{-}R_8)(dppz)_2]^{2+}$) developed by our group. Simple modification of the peptide sequence endows the ruthenium dppz complex with an octaarginine domain for increased DNA binding and cell internalization. [111d]

111 a) I. Gamba, I. Salvadó, G. Rama, M. Bertazzon, M. Sánchez, V. M. Sánchez-Pedregal, 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. Martınez-Costas, R. Lorca, A. Somoza, M. Melle- Franco, M.E. Vázquez, M. Vázquez López, *Chem. Commun.* **2018**, 54, 658-661.

112 a) 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; b) I. Gamba, G. Rama, E. Ortega-Carrasco, R. Berardozzi, V. M. Sánchez-Pedregal, L. Di Bari, J. -D. Maréchal, M. E. Vázquez, M. Vázquez López, *Dalton Trans.* **2016**, 45, 881-885.

Supramolecular Helicates and DNA binding

Definition, structure and classification of helicates

Jean Marie Lehn introduced the term *helicate* in 1987 to describe a double-stranded polinuclear metal complex with helical structure.[113] Thus, a supramolecular helicate can be defined as a helical complex constituted by one or more strands wound around the helical axis defined by two (or more) metal ions.[114] This term includes classic coordination compounds as well as more complex helical assemblies.[115] However, there are two rules that all of them must fulfil to be considered as a helicate:

1. One or more non-cyclic strands must be wrapped and coordinated to one or more metal ions. This rule excludes helical structures that are stabilized by hydrogen bonding, such as DNA, and cyclic helical structures (knots, helical macrocycles, etc.).
2. The helix axis must be constituted by at least two metal ions, forming a discrete polynuclear molecule.

Helicates can be classified depending on the number of strands wound around the metal centers. Thus, *monohelicates*, *dihelicates* or *trihelicates* are composed by one, two or three strands, respectively.[116] Also, depending on the number of metallic centers they can be classified in *mononuclear*, *dinuclear*, *trinuclear* or *polinuclear*. Helicates can also be classified as *homoleptic* or *heteroleptic* if they are composed by association of the same coordinating units or not. Also, if the metal binding site is able to occupy all the coordinating positions, we could find *saturated* or *unsaturated* helicates.

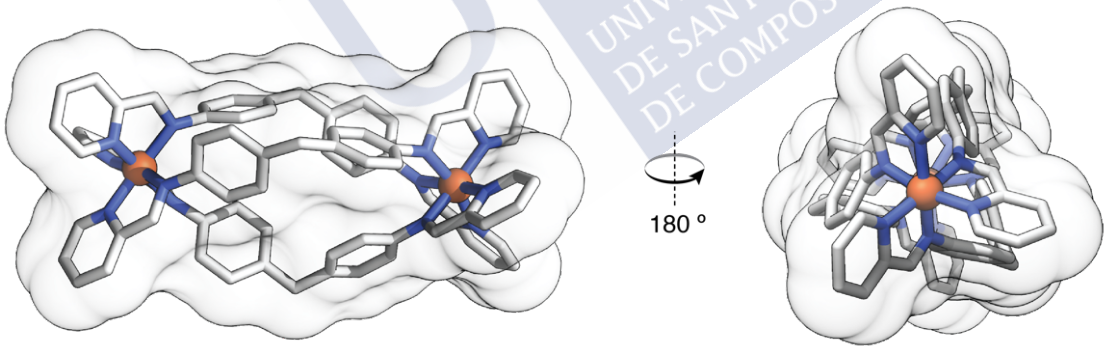


Figure 24. Structure of a left-handed (M) dinuclear trihelicate of Fe(II) described by Hannon *et al.* (PDB 2ET0) in which the three ligand strands wrap around the axis defined by two metal ions.[117]

113 J. M. Lehn, A. Rigault, J. Siegel, J. Harrowfield, B. Chevrier, D. Moras, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 2565-2569.

114 C. Piguet, G. Bernardinelli, G. Hopfgartner, *Chem. Rev.* **1997**, *97*, 2005-2062.

115 D. S. Lawrence, T. Jiang, M. Levett, *Chem. Rev.* **1995**, *95*, 2229-2260.

116 H. Miyake, H. Tsukube, *Chem. Soc. Rev.* **2012**, *41*, 6977-6991.

117 A. Oleksi, A. G. Blanco, R. Boer, I. Usón, J. Aymamí, A. Rodger, M. J. Hannon, M. Coll, *Angew. Chem. Int. Ed.* **2006**, *45*, 1227-1231.

The majority of the published literature concerning dinuclear dihelicates are complexes of Fe(II),[118] Co(II),[119] Ni(II),[120] Zn(II),[121] Ag(I)[122] and Cu(II).[123] Most of them are constituted by polipyridyl ligands or Schiff bases.[124, 125] In these examples, the metal ion can be found as tetra, penta or hexacoordinated. In the case of dinuclear trihelicates, Fe(II),[126] Co(II)[127] and Ni(II)[128] are the most abundant metal ions, although Ru(II) complexes have also been reported.[129] In these constructions, the metal ion is normally hexacoordinate with an

---

118 a) K. T. Potts, K. M. Keshavarz, F. S. Tham, H. D. Abruña, C. Arana, *Inorg. Chem.* **1993**, 32, 4436-4441; b) G. Rapenne, B. T. Patterson, J-P. Sauvage, F. R. Keene, *Chem. Commun.* **1999**, 1853-1855. c) E. C. Constable, M. Neuburguer, D. Smith, M. Zehnder, *Chem. Commun.* **1996**, 1917-1923; d) E. C. Constable, A. J. Edwards, R. Martínez-Manez, P. R. Raithby, *Dalton Trans.* **1995**, 3253-3263.

119 a) E. C. Constable, M. Neuburguer, D. Smith, M. Zehnder, *Chem. Commun.* **1996**, 1917-1923; b) E. C. Constable, A. J. Edwards, R. Martínez-Manez, P. R. Raithby, *Dalton Trans.* **1995**, 3253-3263.

120 a) M. Vazquez, M. R. Bermejo, M. Fondo, A. M. Gonzalez, J. Mahía, L. Sorace, D. Gatteschi, *Eur. J. Inorg. Chem.* **2001**, 1863-1868; b) B. Hasenknoptz, J-M. Lehn, *Helv. Chim. Acta*, **1996**, 79, 1643-1650; c) A. Bylik, M. M. Harding, P. Turner, T. W. Hambley, *Dalton Trans.* **1995**, 2549-2555; d) E. C. Constable, M. D. Ward, D. A. Tocher, *J. Am. Chem. Soc.* **1990**, *112*, 1256-1258.

121 M. Vázquez, M. R. Bermejo, M. Licchelli, A. M. González-Noya, R. M. Pedrido, C. Sangregorio, L. Sorace, A. M. García-Deibe, J. Sanmartín, *Eur. J. Inor. Chem.* **2005**, 3479-3490.

122 a) E. C. Constable, S. M. Elser, M. J. Hannon, A. Martin, P. R. Raithby, D. A. Tocher, *Dalton Trans.* **1996**, 25, 2423-2429; b) N. K. Solanki, A. E. H. Wheatley, S. Radojevic, M. McPartlin, M. A. Halcrow, *Dalton Trans.* **1999**, *28*, 521-562; c) H. Oshio, M. Yamamoto y T. Ito, *Dalton Trans.* **1999**, *28*, 2641-2648.

123 J. Sanmartín, M. R. Bermejo, A. García-Deibe, O. Piro, E. E. Castellano, *Chem. Commun.* **1999**, 1953-1955.

124 a) B. Hasenknopf, J-M. Lehn, G. Baum, D. Fenske, *Proc. Natl. Acad. Sci. USA* **1996**, 93, 1397-1403; b) M-H. Shu, W-Y. Sun, C-Y Duan, Y-J. Fun, W-J. Zhang, W-X. Tang, *Dalton Trans.* **1999**, *28*, 729-735; c) E. C. Constable, M. Neuburguer, D. R. Smith y M. Zehnder, *J. Chem. Soc. Dalton Trans.* **1996**, 25, 4207-4215.

125 L. Carbonaro, M. Isola, V. Liuzzo, F. Marchetti, F. Balzano, C. S. Pomelli, A. Raffaelli, *Eur. J. Inorg. Chem.* **2001**, 353-357; b) G. C. van Stein, G. van Koten, H. Passenier, O. Steinebach, *Inorg. Chim. Acta*, **1984**, *89*, 79-85; c) G. Paolucci, S. Stelluto, S. Sitran, D. Ajo, F. Benetollo, A. Polo, G. Bombieri, *Inorg. Chim. Acta*, **1992**, *193*, 57-66.

126 a) S. Ferrere, M. C. Elliot, *Inorg. Chem.* **1995**, 34, 5818-5824; b) S. L. Larson, S. M. Hendrikson, S. Ferrere, D. L. Derr, C. M. Elliot, *J. Am. Chem. Soc.* **1995**, *117*, 5881-5883.

127 a) H. Cheng, D. Chun-ying, F. Cheng-jie, M. Qing-jin, *Dalton. Trans.* **2000**, 2419-2430; b) C. Piguet, G. Benardinelli, B. Bocquet, O. Schaad, A. F. Williams, *Inorg. Chem.* **1994**, 33, 4112-4119.

128 a) R. Kramer, J-M. Lehn, A. DeCian, J. Fischer, *Angew. Chem. Int. Ed.* **1993**, 32, 704-709; b) M. J. Hannon, C. L. Painting, A. Jackson, J. Hamblin, W. Errington, *Chem. Commun.* **1997**, 1807-1812;

129 J. Malina, M. J. Hannon, V. Brabec, *Chem. Eur. J.* **2008**, 14, 10408-10414.

octahedral coordination geometry. Again, the typical ligands that can be found in literature are normally polipyridyl complexes and Schiff bases.[130,131]

Chirality of the helicates

One of the most important properties of supramolecular systems is chirality, which plays an important role in the transfer or information in living systems,[132] regulation of gene expression[133] or, as in the case of F_1 motor in ATP synthase, maintaining the homeostasis.[134] Due to their helical structure, helicates are inherently chiral species, which can be right-handed (P or plus), when the sense of the turn is clockwise, or left-handed (M or minus), when it is anti-clockwise. Control of the supramolecular chirality of the helicates is one of the most challenging aspects in their synthesis, so that the function and the properties of these metal complexes will be determined by this property, especially in DNA recognition.[135] Thus, as it was proposed previously in this group, a really effective strategy for the obtention of pure helicates derived from labile metal centers is the employment of chiral ligands that are able to transfer their chiral information to the metal center, thus generating preferentially a P or M enantiopure helix.[135a,136] This strategy based on the Pfeiffer effect[137] is a really exploited and efficient option to control the final chirality of the supramolecular systems. Despite that, more sophisticated methods are required, and a newer approach will be exposed in the discussion of this thesis.

Applications of the helicates

Properties of the helicates are defined by the metal ions and the ligands that constitute them. Selecting carefully the building blocks, chemical and structural characteristics can

130 J.-M. Lehn, A. A. Rigault, J. Siegel, J. Harrowfield, B. Chevrier, D. Moras, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 2565-2569.

131 a) H. Cheng, D. Chun-ying, F. Cheng-jie, M. Qing-jin, *Dalton Trans.* **2000**, 2419-2423; b) C. Piguet, G. Benardinelli, B. Bocquet, O. Schaad, A. F. Williams, *Inorg. Chem.* **1994**, *33*, 4112-4125; c) R. Kramer, J-M. Lehn, A. DeCian, J. Fischer, *Angew. Chem. Int. Ed.* **1993**, *32*, 704-709; d) M. J. Hannon, S. Bunce, A. J. Clarke, N. W. Alcock, *Angew. Chem. Int. Ed.* **1999**, *38*, 1277-1278; e) L. K. Thompson, C. J. Matthews, L. Zhao, C. Wilson, M. A. Leech, J. A. K. Howard, *Dalton Trans.* **2001**, 2258-2262.

132 a) H. Buschmann, M. Hauptmann, D. Niessing, C. W. Lloyd, A. R. Schäffnera, *Plant Cell* **2009**, *21*, 2090-2106; b) J. Capdevila, K. J. Vogan, C. J. Tabin, J. C. I. Belmonte, *Cell* **2000**, *101*, 9-21; c) T. Hashimoto, *Philos. Trans. R. Soc.Lond. B.* **2002**, 357, 799-808.

133 G. Wang, L. A. Christensen, K. M. Vasquez, *Proc. Natl. Acad. Sci. USA* **2006**, 103, 2677-2682.

134 a) F. A. Samatey, K. Imada, S. Nagashima, F. Vonderviszt, T. Kumasaka, T.; M. Yamamoto, K. Namba, *Nature* **2001**, *410*, 331-337; b) G. Song, J. Ren, *Chem. Commun.* **2010**, *46*, 7283-7294.

135 a) I. Gamba, G. Rama, E. Ortega-Carrasco, J.-D. Maréchal, J. Martínez-Costas, M. E. Vázquez, M. Vázquez López, *Chem. Commun.* **2014**, *50*, 11097-11100; b) M. I. Sánchez, G. Rama, R. Calo-Lapido, K. Ucar, P. Lincoln, M. Vázquez López, M. Melle-Franco, J. L. Mascareñas, M. E. Vázquez, *Chem. Sci.* **2019**, *10*, 8668-8674.

136 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.

137 S. Kirshner, N. Ahmad, R. C. Brasted, V. J. Landis, E. J. Kuhajek, P. E. R. Nordquist and L. Mayer, in *Coordination Chemistry*, ed. S. Kirshner, Plenum Press, New York, **1969**, pp. 42-64.

be modulated in order to display different functions. Helicates are really versatile systems and they have a lot of potential applications as sensors,[138] liquid crystals,[139] molecular machines,[140] logic gates,[141] magnetic materials,[142] or catalysts.[143] Also, and more interesting for our aim, they have some applications in the field of Chemical Biology as DNA binding agents, as we can see later.

Precedents of peptide helicates

There are not many examples in literature of peptide helicates. One of them was proposed by Albrecth and co-workers described for the first time the employment of natural and non-natural amino acids for the synthesis of dinuclear trihelicates of Ti(IV).[144] More recently, Yashima *et al.* synthesized dinuclear trihelicates of Fe(II) following a similar methodology than Albretch in the previous case, but in this case by changing the cathecol units by 5Bpac in the ends. Chirality of these helicates can be modulated by changing the solvent.[145]

Hannon and co-workers exploited the possibility of coupling peptides to their major groove-DNA-binding supramolecular metallocilinders in order to improve their selectivity. With this aim they introduced a tripeptide (GGS-$CONH_2$) in both ends of a bis-pyridylimine precursor. Glycine residues act as spacer and the serine residue was chosen because is a common amino acid in protein-DNA recognition motifs.[146] In a similar approach, Hannon *et al.* introduced chiral amino acids in order to control the final

---

138 a) S. Goetz, P. E. Kruger, *Dalton Trans.* **2006**, 35, 1277-1284; b) G. Bokolinis, T. Riis-Johannessen, L. P. Harding, J. C. Jeffery, N. McLay, C. R. Rice, *Chem. Commun.* **2006**, 42, 1980-1982; c) X. Zhu, C. He, D. Dong, Y. Liu, C. Duan, *Dalton Trans.* **2010**, 39, 10051-10055.

139 a) K. Binnemans, K. Lodewyckx, T. Cardinaels, T. N. Parac-Vogt, C. Bourgogne, D. Guillon, B. Donnio, *Eur. J. Inorg. Chem.* **2006**, *2006*, 150-157; b) K. Binnemans, Y. G. Galyametdinov, R. Van Deun, D. W. Bruce, S. R. Collinson, A. P. Polishchuk, I. Bikchantaev, W. Haase, A. V. Prosvirin, L. Tinchurina, I. Litvinov, A. Gubajdullin, A. Rakhmatullin, K. Uytterhoeven, L. Van Meervelt, *J. Am. Chem. Soc.* **2000**, *122*, 4335-4344; c) A. El-ghayoury, L. Douce, A. Skoulios, R. Ziessel, *Angew. Chem. Int. Ed.* **1998**, *37*, 2205-2208.

140 a) F. Cardinali, H. Mamlouk, Y. Rio, N. Armaroli, J. F. Nierengarten, *Chem. Commun.* **2004**, 1582-1583; b) D. Schultz, F. Biaso, A. R. M. Shahi, M. Geoffroy, K. Rissanen, L. Gagliardi, C. J. Cramer, J. R. Nitschke, *Chem. Eur. J.* **2008**, 14, 7180-7185.

141 M. Vázquez López, M. E. Vázquez, C. Gómez-Reino, R. Pedrido, M. R. Bermejo, *New J. Chem.* **2008**, 32, 1473-1477.

142 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. Vázquez 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.* **2003**, 42, 3166-3169.

143 a) C. Dietrich-Buchecker, G. Rapenne, *Chem. Commun.* **1997**, 33, 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.

144 M. Albrecht, O. Spieß, M. Schneider, P. Weis, *Chem Comm.* **2002**, 786-787.

145 N. Ousaka, Y. Takeyama, E. Yashima, *Chem. Sci.* **2012**, 3, 466-469.

146 L. Cardo, M. J. Hannon, *Inorg. Chim. Acta* **2009**, *362*, 784-792

chirality of the resulting helicate. Thus, they coupled a L/D-Arg residue at the ends of the previous bis-pyridylimine precursor, thus controlling the final chirality of the obtained helicates with the chirality of the Arg, and making them water soluble.[147]

In all of these cases, none of the peptide precursors were synthesized through a SPPS or in a modular phase from their N-terminus to the C-terminus. Furthermore, in some of these examples only one amino acid is used; that is why they cannot be considered as peptide helicates strictly. A closer approach will be discussed later in this passage.

DNA function and structure

Function of the DNA

Nature of the living organisms is determined by its genes, the concrete sequence of nucleic acids that codify the proteins.[148] Transference of this information is a crucial aspect for life. The flux of this information was postulated by Francis Crick in 1957. This idea is the so-called *Central Dogma of the Biology*.[149] This dogma has been actualized in the past years introducing new pathways for the flux of the information (Figure 25).

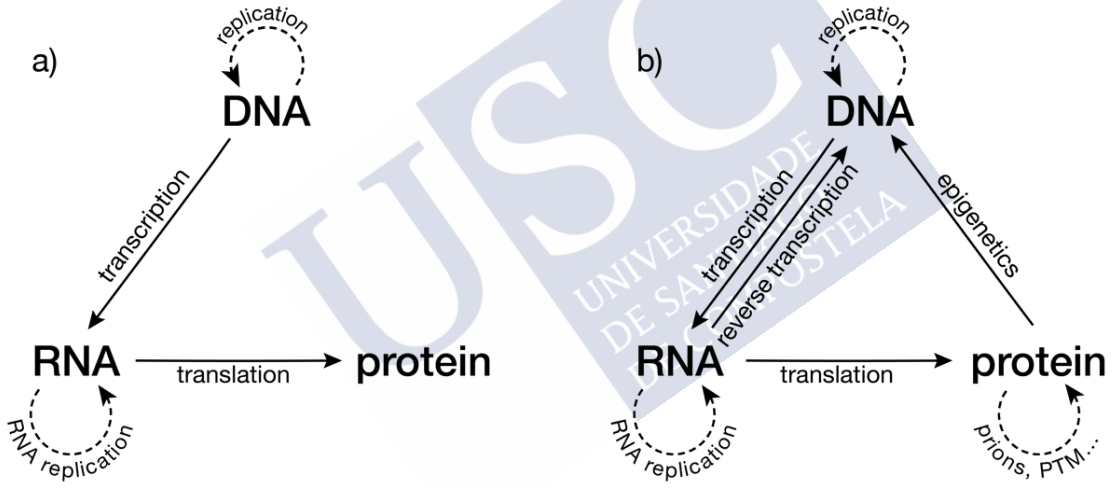


Figure 25. a) The original central dogma of molecular biology as originally proposed by F. Crick in 1957;[150] b) updated version of the dogma including other processes that influence the way in which information is transmitted/stored in cells.

Based on the previous scheme, genetic information is transcribed unidirectionally from the DNA to the RNA. This dogma gave to the DNA a key role in biology and promoted the appearance of a new molecular biology in the 1950's decade. However, it was also criticized because in this first proposal, information is translated into proteins and these

147 L. Cardo, V. Sadovnikova, S. Phongtongpasuk, N. J. Hodges, M. J. Hannon, *Chem. Commun.* **2011**, *47*, 6575-6577.

148 H. Lodish, A. Berk, L. S. Zipursky, P. Matsudaira, D. Baltimore, J. Darnell, *Molecular Cell Biology*, W. H. Freeman, **2000**, chapter 9.

149 F. H. C. Crick, *Symp. Soc. Exp. Biol.* **1958**, *12*, 138-163.

150 a) F Crick, *Symp. Soc. Exp. Biol.* **1958**, *12*, 138-163; b) F. Crick, *Nature* **1970**, *227*, 561-563; c) M. Cobb, *PLoS Biol.* **2017**, *15*, e2003243.

proteins cannot be modified or affect to the genes. This unidirectionality of the process is the point that caused more controversy and it is now still questioned.

Structure of B-DNA

DNA is a biopolymer composed by nucleotides and with a characteristic highly negatively charged phosphate backbone. It has a regular structure of a right-handed double helix where the strands remain united due to the hydrogen bonding interactions that exist between them. This helix can adopt different structures depending on the conditions of the environment such as pH, ionic strength, solvent, etc. But the most common and relevant structure in physiological conditions is the B-DNA (Figure 26). B-DNA has a diameter of 20 Å approximately with 10 pairs of bases per each turn. The distance between the consecutive pairs is around 3.4 Å and the rotation per residue is 36°. Despite being a very regular and conserved structure, there are some deviations in the torsion angle, that are characteristic of the sequence and really important in terms of protein-DNA interaction.[151] In the structure of the B-DNA we can distinguish two grooves of different sizes. Major groove is wide and not really deep, while the minor groove is narrow and deep. The width of this last one is going to be determined by the sequence, being the regions rich in A-T narrower while G-C regions are wider.

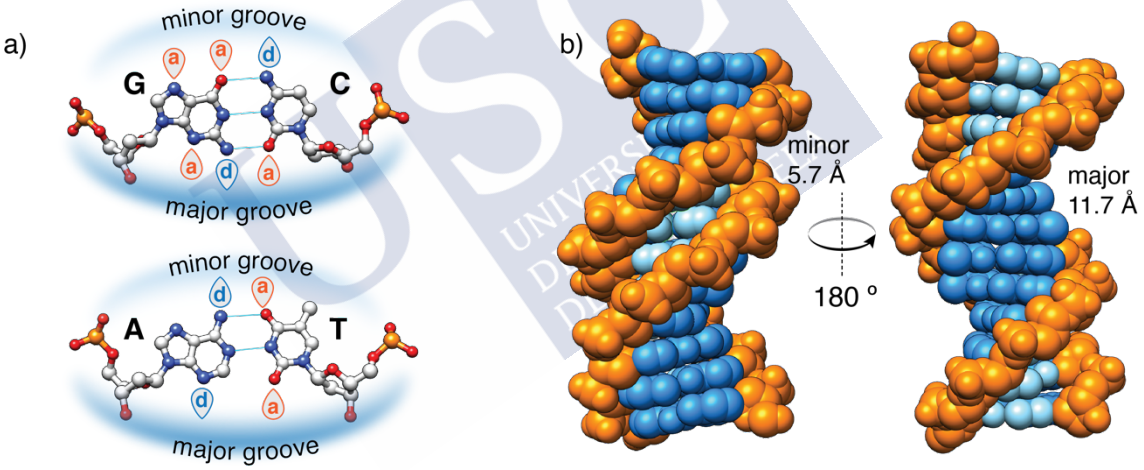


Figure 26. Left: structure of the four nitrogenous bases forming the corresponding base pairs, showing the different combination of hydrogen bond donors and acceptors in the major and minor grooves of the DNA. Right: structure of the Dickerson-Drew B-DNA dodecamer, where the minor groove (light blue) and major groove (dark blue) can be observed; the width of each groove is indicated for comparison.

It is important to mention that the variability of the major groove is much bigger than in the case of the minor groove. Also, studies about the hydration of these grooves showed important differences from the thermodynamic point of view. The addition of these

151 a) G. M. Blackburn, M. J. Gait, D. Loakes, D. M. Williams, *Nucleic Acids in Chemistry and Biology*, RSC Publishing, Cambridge, **2006**; b) Z. Otwinowski, R. W. Schevitz, R–G. Zhang, C. L. Lawson, A. Joachimiak, R. Q. Marmorstein, B. F. Luisi, P. B. Sigler, *Nature* **1988**, 335, 321-329; c) R. E. Dickerson, H. R. Drew, J. Mol. Biol. **1981**, 149, 761-786.

factors—size, hydration and variability—justify the fact that the proteins recognize the DNA preferentially through the interaction with the major groove.[152]

Structure of three-way DNA junction

DNA three-way-junction (twDNA) is composed of three DNA chains that converge at a junction point, generating a prism-shaped central hydrophobic hole just at the center of the superstructure.[117,153,154] twDNAs play important roles in various cellular processes.[155] They occur in both RNA during splicing[156] and translation processes,[157] as well as in DNA, where they form transiently during replication,[158] and recombination processes.[159] In addition to their important roles in various biological processes, branched junctions are very attractive superstructures in DNA-based nanotechnology.[160]

Interaction of the DNA and small molecules

In the last years, discovery of the genetic origin of several diseases as well as the importance of the genotype in the answer to different drugs, resulted in a crescent interest in the detection and sensing of specific sequences with a great clinic and diagnostic potential. Despite the big efforts in this field, selective recognition of DNA sequences is still a non-solved problem in Chemical Biology.

Between the classical DNA-binding agents we can highlight coordination compounds with heavy metals and small organic molecules that recognize the minor groove. These classical DNA-binding agents possess several limitations. One of these is the low selectivity against the sequence of the DNA and therefore the subsequent toxicity. Other problems that these molecules present are the poor stability in physiological media, problems to penetrate the cell and, in the case of the metal complexes, poor biocompatibility of the ligands, lability in physiological media, and the tedious derivatization that makes difficult to carry out systematic studies.

---

152 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) N. M. Luscombe, J. Laskowski, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, *Nucleic Acids Res.* **2000**, *28*, 235-242.

153 D. M. J. Lilley, *Quart. Rev. Biophys.* **2000**, 33, 109-159.

154 D. M. J. Lilley, *Biopolymers* **1998**, 48, 101-112.

155 a) P. A. Kitts, H. A. Nash, *Nature*, **1987**, 329, 346-348; b) T. L. Orr-Weaver, J. W. Szostak, R. J. Rothstein, *Proc. Natl. Acad. Sci. U.S.A.*, **1981**, *78*, 6354-6358.

156 C. Guthrie, B. Patterson, *Annu. Rev. Genet.* **1988**, *22*, 387-419.

157 a) B. T. Wimberly, D. E. Brodersen, W. M. Clemons, 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.

158 M. R. Singleton, S. Scaife, D. E. Wigley, *Cell* **2001**, 107, 79-89.

159 C. J. Hutchins, P. D. Rathjen, A. C. Forster, R. H. Symons, *Nucleic Acids Res.* **1986**, 14, 3627-3640.

160 a) N. C. Seeman, *Curr. Opin. Struct. Biol* **1996**, *6*, 519-526; b) H. T. LaBean, H. Yan, J. Kopatsch, F. R. Liu, E. Winfree, J. H. Rief, N. C. J. Seeman, *J. Am. Chem. Soc.* **2000**, *122*, 1848-1860.

These molecules can interact with the DNA in some different ways:[161]

- *Electrostatic interactions*: negative phosphate groups in the DNA interact with positive metal ions such as Na^+, K^+ or Mg^{2+}. These interactions stabilize the DNA structure, avoiding the electrostatic repulsion forces. This stabilization can also be achieved with polycationic molecules like spermine or spermidine.[162] Another example of this interaction is a group of Pt(II) compounds known as phosphate clamps. These molecules interact with the DNA in a similar way to natural polyamines.[163]
- *Intercalation/insertion*: this kind of interaction is typical of cationic molecules with aromatic groups in their structure. These molecules can intercalate the DNA between base pairs. The interaction is stabilized by an addition of different forces: π-stacking, charge transfer, hydrogen bonding, and electrostatic interactions.[164] There is a preference of binding in regions that are rich in G-C pairs, although they usually display poor sequence selectivity.[165] Some aromatic DNA-binders, due to their especially expanded surfaces, cannot intercalate but insert between pairs of mismatched DNA bases, displacing the natural base pair and replacing it inside the DNA helix.[166]
- *Insertion in the minor groove:* this kind of interaction is typical of small molecules that can introduce themselves in the minor groove of the DNA. These interactions are favored by the action of different forces such as electrostatic interactions, van der Waals forces or hydrogen bonding.[167] Also there is an entropic factor related with the capability of these molecules to displace water molecules and other ions when the insertion occurs. Small molecules that interact with the DNA in this way usually prefer A/T-rich regions, because the groove in this case is deeper and narrower than the G-C regions. A classic example of this molecules is the DAPI.[168]
- *Covalent/Coordinated covalent interaction*: some molecules are able to interact with the DNA in a covalent mode, establishing bonds with the nucleophilic regions of

---

161 a) L. M. Wilhelmsson, P. Lincoln, B. Nordén, *Sequence–specific DNA Binding Agents*, RSC Publishing, **2006**. b) A. Lorente, M. J. Fernández, *An. Quím.* **2008**, *104*, 280-289. c) M. J. Hannon, *Chem. Soc. Rev.* **2007**, *36*, 280-295.

162 H. Deng, V. A. Bloomfield, J. M. Benevides, G. J. Thomas Jr, *Nucl. Acids Res.* **2000**, *28*, 3379-3385.

163 R. Boer, A. Canals, M. Coll, *Dalton Trans.* **2009**, 399-414.

164 a) K. Nakamoto, M. Tsuboi, G. D. Strahan, *Intercalating Drugs, in Drug-DNA interactions: Structures and Spectra*, **2008**, John Wiley & Sons, Inc., Hoboken, USA; b) A. Rescifina, Z. Zagni, M.G. Varrica, V. Pistarà, A. Corsaro, *Eur. J. Med. Chem.* **2014**, *74*, 95-115.

165 a) B. A. D. Neto, A. A. M. *Lapis, Molecules*, **2009**, *14*, 1725-1746; b) H. M. Bernan, P. R. Young, *Annu. Rev. Biophys. Bioeng.* **1981**, *10*, 87-114.

166 A. C. Komor, J. K. Barton, *Chem. Commun.*, **2013**, 3617-3630.

167 P. G. Baraldi, A. Bovero, F. Fruttarolo, D. Petri, M. A. Tabrizi, M. G. Pavani, R. Romagnoli, *Med. Res. Rev.* **2004**, *24*, 475-528.

168 a) B. Nguyen, M. P. H. Lee, D. Hamelberg, A. Joubert, C. Bailly, R. Brun, S. Neidle, D. W. J. Wilson, *J. Am. Chem. Soc.* **2002**, *124*, 13680-13681; b) D. Hamelberg, L. D. Williams, W.D. Wilson, *J. Am. Chem. Soc.* **2001**, *123*, 7745-7755.

nitrogenous bases. Some of these molecules are also able to establish covalent bonds with more than one base, crosslinking the DNA and avoiding the replication process. For this reason, this kind of molecules are widely used in chemotherapy.[169] Some examples are the nitrogen mustards and complexes of Pt(II) such as cisplatin or carboplatin.[170]

Helicates as DNA-binding agents

In 1995, Prof. Jean Marie Lehn and co-workers, described for the first time the DNA binding properties of a family of Cu(I) polynuclear helicates derived from Bpy, which appeared to insert in the DNA major groove, a non-covalent interaction that might be specific and clearly dependent on the size and shape of the helicates.[171] Later, the group of M. J. Hannon described the properties of a new Fe(II) dinuclear trihelicate derived from a bis-pyridinylimine that was able to bind the DNA in the major groove, causing a significant bend of the double helix. Furthermore, in this work Hannon demonstrated that the chirality plays a key role in the binding mode of the helicates, so that the M isomer of the helicate binds preferably to the major groove, while the P isomer does it in the minor groove.[172]

In 2006, Hannon *et al.* published the X-ray structure of the interaction between this helicate and a three-way DNA junction. This non-canonical DNA structure consists in three strands with complementary ends that generate a prism-shaped central cavity where the Hannon's helicate fits perfectly.[117] Subsequent studies revealed its high activity against different cancer cell lines. The mechanism seems to be non genotoxic as in the case of cisplatin or other alkylating agents, but it acts reducing the mitochondrial activity, thus inhibiting the cellular growth.[173]

The importance of the chirality in the DNA recognition was highlighted in the previous discussion about the helicate designed by Hannon. Following this, Motti *et al.* were able to separate the isomers P and M of a Ru(II) dinuclear trihelicate. The method consists in the employment of a nucleotide of 12 bases pairs that in anchored to a stationary phase, through which, the racemic mixture is eluted. Thus, the P isomer, that interacts stronger with the DNA, is retained in the stationary phase, while the M isomer is almost eluted with solvent.[174]

---

169 a) D. Fu, J. A. Calvo, L. D. Samson, *Nat. Rev. Cancer* **2012**, *12*, 104-120;

170 a) T. Boulikas, M. Vougiouka, *Oncol. Rep.* **2004**, *11*, 559-595; b) R. A. Alderden, M. D. Hall, T. W. Hambley, *J. Chem. Ed.* **2006**, *83*, 728-734.

171 B. Schoentjes, J.-M. Lehn, *Helv. Chim. Acta* **1995**, *78*, 1-12.

172 a) J. M. C. A. Kerckhoffs, J. C. Peberdy, I. Meistermann, L. J. Childs, C. J. Isaac, C. R. Pearmund, V. Reudegger, S. Khalid, N. W. Alcock, M. J. Hannon, *Dalton Trans.* **2007**, 734-742; b) 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, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5069-5074.

173 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.

174 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.

The group of Prof. P. Scott is also a remarkable reference in the field of the helicates as DNA-binding agents. They described the synthesis of optically pure, water soluble and stable Fe(II) dinuclear dihelicates. These helicates recognize the B-DNA by binding the major groove, preferably A-T rich regions, with nM dissociation constants. These compounds have also shown good activity against gram positive and gram-negative bacteria.[175]

A closer approach to the final aim of this doctoral thesis was carried in our group in the past years. In a recent communication, the synthesis of enantiomerically pure peptide helicates was achieved. The followed methodology was to combine natural and a non-natural residue-derived from the asymmetrically substituted 5Bpy amino acid - to obtain the ligand precursors of the final helicates (Figure 27). The peptide ligand strands led to the obtention of tris-bipyridyl Fe(II) dinuclear helicates that showed good DNA-binding properties versus three-way DNA and also high selectivity versus B-DNA.

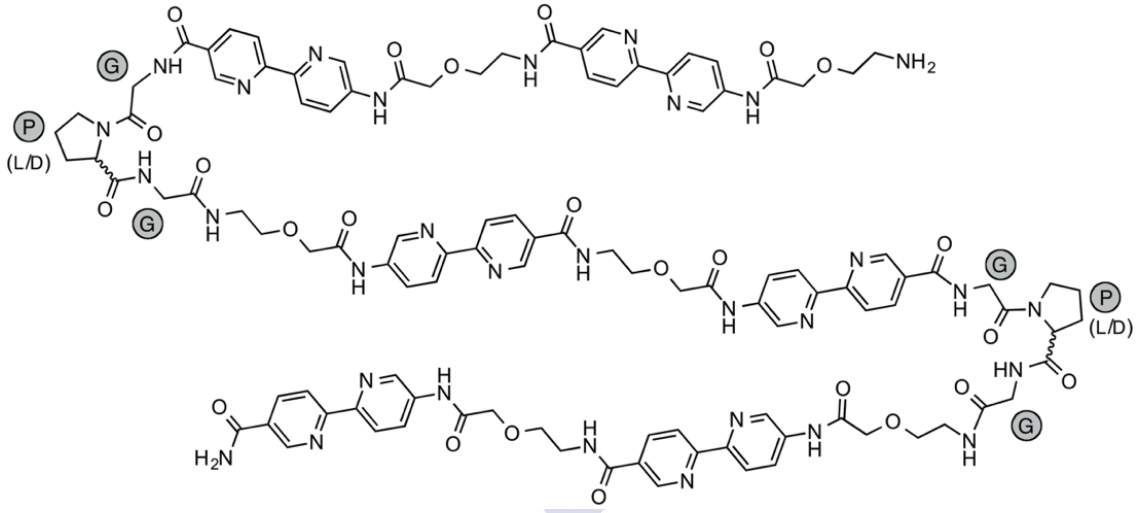


Figure 27. Structure of the peptide precursor of Fe(II) dinuclear helicates developed in our group. Chirality of the proline was selected depending on the desired supramolecular isomer.[135a]

This is the first precedent of real peptide helicates where the coordinating unit is perfectly integrated in the peptide sequence. For the synthesis of the ligands, solid phase peptide synthesis was carried out, providing all the advantages that this technique offers for the obtention of the ligands such as versatility, efficiency, reliability… In the design of the strands, chiral promoters of β-turn (GPG) were introduced in the sequence to connect the different coordinating subunits. Thus, just changing the isomer of the proline that is used, the final supramolecular chirality is also controlled.[135a]

Some other structures have also shown good DNA binding properties versus three-way junction-DNA. The most representative molecules that are able to bind selectively this kind of DNA structures are known as tryptycenes. These kind of molecules are characterized by possessing a threefold symmetric architecture with dimensions similar to those of the central helical interface of a perfectly base-paired nucleic acid three-way

175 S. E. Howson, A. Bolhuis, V. Brabec, G. J. Clarkson, J. Malina, A. Rodger, P. Scott, *Nat. Chem.* **2012**, 4, 31-36.

junction.[176] Group of Prof. Chenoweth have described in the recent years many approaches for the obtention of these kind of molecules that selectively bind the hydrophobic cavity of the twDNA.[177]

Other different compounds apart from these triptycene structures, have also shown good DNA binding properties against three-way junction. Some of these examples are the azacryptands proposed by Prof. Monchaud[178], or the recently proposed $Fe(II)_4L_4$ cage proposed by Prof. Jonathan Nitschke.[179]

176 S. A. Barros, D. M. Chenoweth, *Angew. Chem. Int. Ed.* **2014**, 53, 13746-13750.

177 a) S. A. Barros, D. M. Chenoweth, *Chem. Sci.* **2015**, *6*, 4752-4755; b)I. Yoon, S.-E. Suh, S. A. Barros, D. M. Chenoweth, *Org. Lett.* **2016**, *18*, 1096-1099; c) S. A. Barros, I. Yoon, S.-E. Suh, D. M. Chenoweth, *Org. Lett.* **2016**, *18*, 2423-2426; d) S. A. Barros, I. Yoon, D. M. Chenoweth, *Angew. Chem. Int. Ed.* **2016**, 55, 8258-8261.

178 J. Novotna, A. Laguerre, A. Granzhan, M. Pirrotta, M.-P. Teulade-Fichou, D. Monchaud, *Org. Biomol. Chem.* **2015**, 13, 215-222.

179 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.

General Objectives

The main objective of this thesis is the obtention of enantiomerically pure peptide helicates derived from 5Bpy by using two different approaches:

- In the first one we exploit the self-assembly properties of two different trimeric natural platforms (the Fibritin foldon of the bacteriophage T4, T4Ff and the β-annulus peptide of the Sesbania Mosaic virus capsid) in order to obtain Fe(II) bis(tris-bipyridyl) helicates. In both cases, the chirality of the metal centers will be defined by the supramolecular assembly of the peptide trimers.
- In the second approach, we design a lineal peptide with three bisbipyridine coordinating subunits. These subunits are connected by two mixed chiral sequences of arginine and proline residues able to promote the formation of β-turns, which support the assembly of the corresponding Fe(II) dinuclear helicates. The supramolecular chirality around the metal centers is going to be determined by the conformation induced by the chiral β-turns. Moreover, the synthesis of the kinetically-inert Co(III) helicates will be carried out starting from the corresponding Co(II) derivatives.

Finally, ability to selectively recognize three-way DNA junction will be tested for the obtained systems by employing different techniques.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter 1

Stereoselective self-assembly of a trimeric helicate

Summary

Peptides are highly versatile platforms for the precise design of supramolecular metal architectures and, particularly, for the controlled assembly of helicates. In this context, we show that the trimeric bacteriophage T4 Fibritin foldon domain (T4Ff) can been engineered on its N-terminus with metal-chelating 2,2′-bipyridine units to yield a peptide ligand that was stereoselectively assembled in the presence of Fe(II) ions into a parallel, three-stranded peptide helicate with M chirality. Modeling studies support the proposed self-assembly and the stability of the final helicate. Furthermore, we also show that the designed trimeric peptide helicate selectively recognize three-way DNA junctions over double-stranded DNA with μM affinity.

This chapter was originally published in a special issue of *Frontiers* journal on *Supramolecular Metal-Based Entities for Biomedical and Biological Applications* as:

J. Gómez-González, D. G. Peña, G. Barka, G. Sciortino, J.-D. Maréchal, M. Vázquez López, M. E. Vázquez. Directed Self-Assembly of Trimeric DNA-Binding chiral Miniprotein Helicates. *Front. Chem.*, **2018**, 6, 520. doi: *10.3389/fchem.2018.00520*.

https://www.frontiersin.org/articles/10.3389/fchem.2018.00520/full

The T4 Fibritin Foldon

The T4 Fibritin foldon (T4Ff) is a short peptide of 27 residues that is found in the C-terminal region of *Fibritin*, a structural protein of bacteriophage T4. This small domain is fundamental for the trimerization of the Fibritin protein, and it is postulated that it provides correct alignment of the Fibritin monomers facilitating the formation of the trimeric coiled coils in the structure (Figure 28).

The structure of the T4 Fibritin protein has been studied by X-ray crystallography[180] and more recently by cryoelectron microscopy;[181] the structure of the T4Ff has also been determined by X-ray[182] and by NMR spectroscopy.[183]

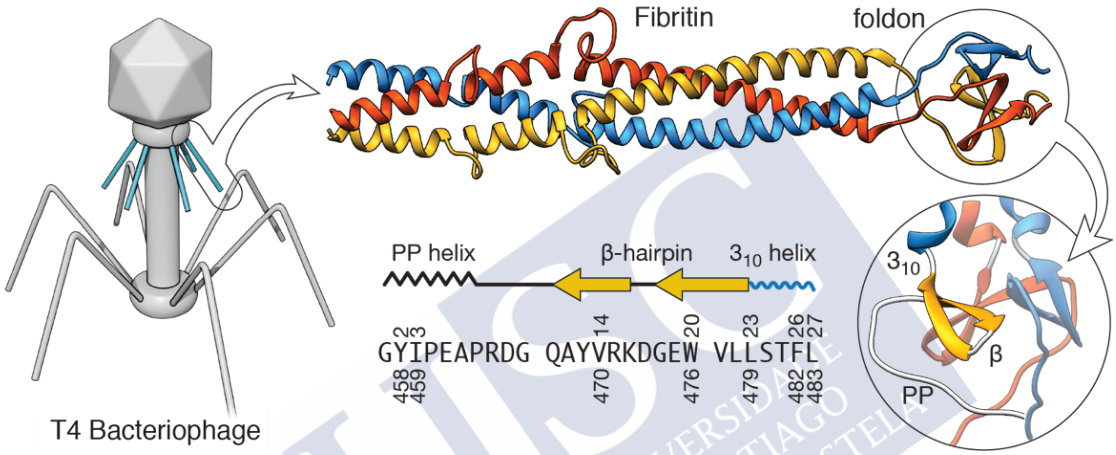


Figure 28. Fibritin is a homotrimer protein found in the Bacteriophage T4 virus collar that stimulate assembly of the long tail fibers (in blue) and their subsequent attachment to the tail baseplate during virus assembly. The structure of the C-terminal region of Fibritin, showing part of the trimeric α-helical coiled coil and the autonomous foldon unit (UniProtKB, P10104; PDB, 1AA0). Detail of the foldon domain showing the main structural elements: C-terminal 3_{10} helix, a β-hairpin, and the N-terminal polyproline (PP) helix (PDB, 1RFO). Schematic representation of the secondary structure elements and amino acid sequence of the foldon domain. Key resides in the foldon domain (1-27) are highlighted. Sequence numbers below correspond to the numbering in the X-ray structure of the T4 Fibritin in PDB 1AA0.

180 a) Y. Tao, S.V. Strelkov, V. V. Mesyanzhinov, M. G. Rossmann, *Structure* **1997**, *5*, 789-798; b) A. Mitraki, S. Miller, M. J. van Raaij, J. *Struct. Biol.* **2002**, *137*, 236-247.

181 a) V. A. Kostyuchenko, P. R. Chipman, P. G. Leiman, F. Arisaka, V. V. Mesyanzhinov, M. G. Rossmann, *Nat. Struct. Mol. Biol.* **2005**, *12*, 810-813; b) A. Fokine, Z. Zhang, S. Kanamaru, V. D. Bowman, A. A. Aksyuk, F. Arisaka, V. B. Rao, M. G. Rossmann, *J. Mol. Biol.* **2013**, *425*, 1731-1744.

182 A. Berthelmann, J. Lach, M. A. Gräwert, M. Groll, J. Eichler, *Org. Biomol. Chem.* **2014**, *12*, 2606-2614.

183 S. Güthe, L. Kapinos, A. Möglich, S. Meier, S. Grzesiek, T. Kiefhaber, *J. Mol. Biol.* **2004**, *337*, 905-915.

The T4Ff is a compact homotrimeric β-propeller-like structure. Each subunit is divided in three different regions, a 3_{10} helix at the C-terminus, followed by a β-hairpin and a N-terminal polyproline helix.[184] The β-hairpin—and the whole T4Ff structure—is stabilized by a set of hydrogen bonds, hydrophobic, and electrostatic interactions. Residues Y458, I459, V470, W476, L479, F482, and L483 form the hydrophobic core within the T4Ff,[185] and the three β-hairpins interact through two main chain hydrogen bonds between each pair of neighbor hairpins. W476 is fundamental for the structural integrity of the whole domain, and its mutation to non-aromatic or small residues prevents folding of the T4Ff. In addition to those interactions, there is a key intermolecular salt bridge between E461 and R471 from adjacent monomers. Thanks to this latter interaction the trimerization process of the natural T4Ff can be modulated by varying the pH.

T4Ff is a privileged scaffold for protein engineering that can overcome many of the drawbacks associated with other motifs currently used for rational protein design. Thus for instance, *α*-helical coiled coils, which probably are the most studied motif and have found multiple applications from drug delivery to nanotechnology, display relatively low thermodynamic stability and high structural polymorphism, regarding not only helix arrangement and orientation, but also in partner selection and stoichiometry.[186] In contrast, T4Ff displays a robust and well-defined three dimensional structure, extraordinary thermodynamic stability and fast two-state folding kinetics, promising a simplified approach to complex designs. Some examples in literature of T4Ff fusion proteins include a recombinant trivalent single-chain variable fragment antibody directed against rabies virus glycoprotein G,[187] a recombinant mimetics of the HIV gp41 prehairpin fusion intermediate,[188] a recombinant influenza H5N1 vaccine[189] or the construction of different nanostructures.[190]

184 J. Habazettl, A. Reiner, T. Kiefhaber, *J. Mol. Biol.* 2009, 389, 103-114.

185 Residue numbers according to the T4 Fibritin X-ray structure (PDB code 1AA0), as reported in 1a.

186 G. G. Rhys, C. W. Wood, J. L. Beesley, N. R. Zaccai, A. J. Burton, R. L. Brady, A. R. Thomson and D. N. Woolfson, J. Am. Chem. Soc. 2019, 141, 8787-8797; b) N. C. Burgess, T. H. Sharp, F. Thomas, C. W. Wood, A. R. Thomson, N. R. Zaccai, R. L. Brady, L. C. Serpell and D. N. Woolfson, J. Am. Chem. Soc. 2015, 137, 10554-10562.

187 I. Turki, A. Hammami, H. Kharmachi, M. Mousli, Mol. Immunol. 2014, 57, 66-73.

188 Z. Qi, C. Pan, H. Lu, Y. Shui, L. Li, X. Li, X. Xu, S. Liu, S. Jiang, Biochem. Biophys. Res. Commun. 2010, 398, 506-512.

189 L. Du, V.H. Leung, X. Zhang, J. Zhou, M. Chen, W. He, H.Y. Zhang, C.C. Chan, V.K. Poon, G. Zhao, S. Sun, L. Cai, Y. Zhou, B.J. Zheng, S. Jiang, PLoS One 2011, 6, e16555.

190 N. Kobayashi, K. Yanase, T. Sato, S. Unzai, M. H. Hecht, R. Arai, J. Am. Chem. Soc. 2015, 137, 11285-11293.

Objective

The objective of this chapter is the stereoselective assembly of a DNA-binding peptide helicate directed by the T4Ff domain. The foldon domain will define the number, orientation and preorganization of the ligand strands of the helicate, inducing a particular supramolecular chirality in the metal centers.

Results and discussion

As shown in the introduction, Bpy is a privileged chelator that forms stable complexes with a variety of metal ions and that has been extensively used in coordination and supramolecular chemistry.[191] Furthermore, we have previously described an Fmoc-protected Bpy dipeptide building block that can be readily implemented into standard Fmoc solid-phase peptide synthesis (SPPS) protocols. Moreover, we have shown that the structure of this chelating unit, in which the Bpy ligand is integrated in the peptide backbone, effectively couples the conformational preferences of the peptide chain with the geometry of the resulting metal complexes.[111,135a]

The 5Bpy unit is derivatized as a Fmoc-protected achiral amino acid featuring a β-Alanine residue for improved solubility. Starting from the commercially available 5,5'-dimethy-2,2'-bipyridine, and following previously reported procedures,[107,192] we were able to synthesize the Fmoc-protected coordinating amino acid (Fmoc-βAla5Bpy-OH, **1**) on a multigram scale with an overall yield of 46% (Scheme 1). In short, after initial 5,5'-dimethy-2,2'-bipyridine oxidation with potassium permanganate and esterification, the intermediate Bpy was desymmetrized by reaction with hydrazine monohydrate in a mixture of ethanol and toluene (3:1) at 80 °C for 30 h. This desymmetrization reaction is the critical step of the route, the proportion of the solvents and the reaction temperature being key parameters to achieve the selective precipitation of the intermediate monohydrazine (**4**). Curtius rearrangement of the corresponding carbazide by heating in a 1:1 mixture of ethanol and xylene, and subsequent deprotection of intermediate **6**, afforded the desired Bpy amino acid core unit (5Bpy, **7**). Unfortunately, this amino acid was found rather insoluble, and its 5'-amine displayed low nucleophilicity and high tendency to oxidize.[106,193] Therefore, in order to obtain a derivative that could be used in peptide synthesis in solid phase, **7** was derivatized in the form of a dipeptide that shows increased solubility, contains a more reactive amine, and avoids the problems derived from oxidation in 5'. In this way, coupling of **7** with Fmoc-βAla-OH, allows us to obtain

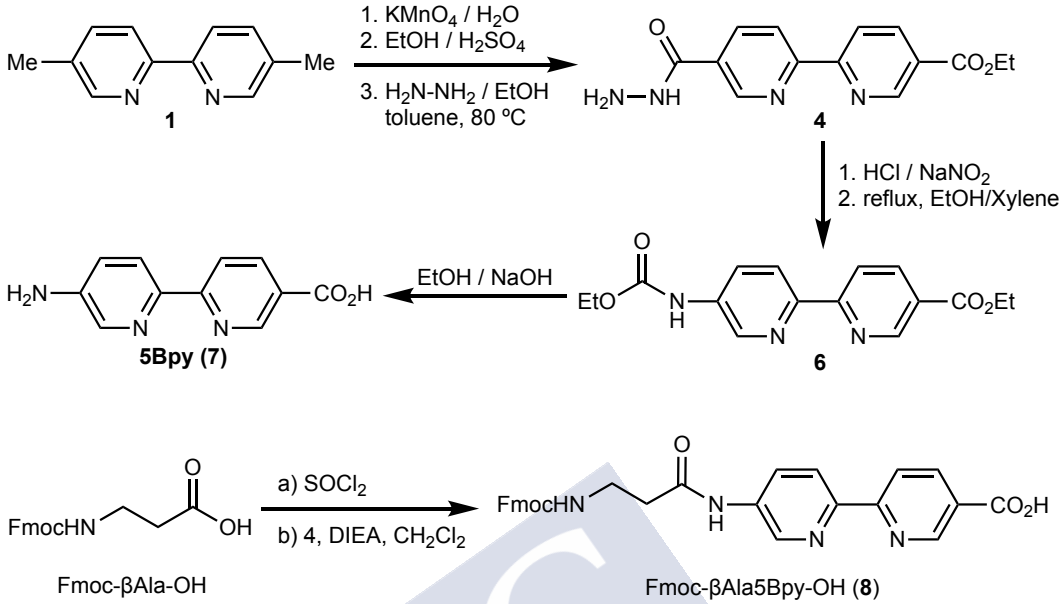
---

191 a) C. Kaes, A. Katz, M. W. Hossein, *Chem. Rev.* **2000**, *100*, 3553-3590; b) W. W. Brandt, *Chem. Rev.* **1954**, 54, 959-1007; c) M. Y. Ogawa, A. Gretchikhine, S. Soni, S. Davis, *Inorg. Chem.* **1995**, 34, 6423-6424; d) S. J. Metallo, A. Schepartz, *Chem. Biol.* **1994**, *1*, 143-151; e) N. Sardesai, S. C. Lin, K. Zimmerman, J. K. Barton, *Bioconjugate Chem.* **1995**, *6*, 302-312; f) B. M. Bishop, D. G. McCafferty, B. W. Erickson, *Tetrahedron* **2000**, 56, 4629-4638; g) K. J. Kise, B. E. Bowler, *Inorg. Chem.* **2002**, 41, 379-386.

192 C. P. Whittle, *J. Heterocyclic Chem.* **1977**, *14*, 191-194.

193 Aromatic amines typically display poor reactivity in peptide couplings: a) J. P. Schneider, J. W. Kelly, *J. Am. Chem. Soc.* **1995**, *117*, 2533-2546; b) C. Yu, J. W. Taylor, *Bioorg. Med. Chem.* **1999**, *7*, 161-175.

the derivative Fmoc-βAla5Bpy-OH (**8**), which is our basic building block for the solid phase synthesis of metallopeptides.



Scheme 1. Synthesis of the coordinating building block Fmoc-βAla5Bpy-OH (**8**).

Inspection of the structure of T4Ff (PDB IDs 4NCU or 1RFO), showed that the N-terminal Gly residues are relatively close to each other and could accommodate the chelating 2,2'-bipyridine units without noticeable distortion of the T4Ff scaffold upon metal ion coordination (Figure 29). Moreover, we envisioned that the natural twist of the N-terminal polyproline helices in the folded T4Ff trimer should induce a ΛΛ-configuration (M helicity) on its derived helicate, which would be the preferred chirality for the efficient recognition of three-way junction DNA.[117]

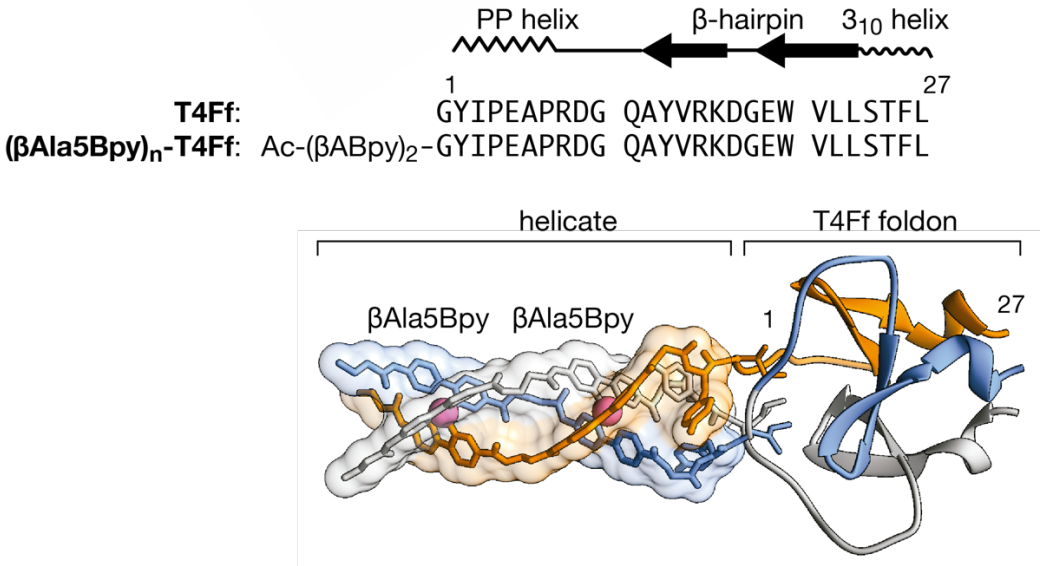


Figure 29. Top: sequence of the natural foldon and the ligand strand precursor of the peptide helicate containing the two 5Bpy units attached to the N-terminus of the natural sequence. Bottom: proposed structure for the $[Fe(II)_2[(\beta Ala5Bpy)_2\text{-}T4Ff]_3]^{4+}$ helicate.

The **(βAla5Bpy)$_2$-T4Ff** peptide strand was synthesized following standard Fmoc SPPS methods.[194] To avoid potential issues related with aspartimide formation in the two Asp-Gly stretches of the sequence, these residues were coupled using the commercially available Fmoc-Asp(OtBu)-(Dmb)Gly-OH dipeptide, where the amide nitrogen in the Gly residue is protected with a bulky group in order to avoid aspartimide formation by attack of the N of the Gly to the carboxylate group in the Asp side chain (Figure 30). Upon assembly of the full sequence and cleavage/deprotection, the final peptide ligand was purified by reverse-phase HPLC, and its identity confirmed by ESI-MS.

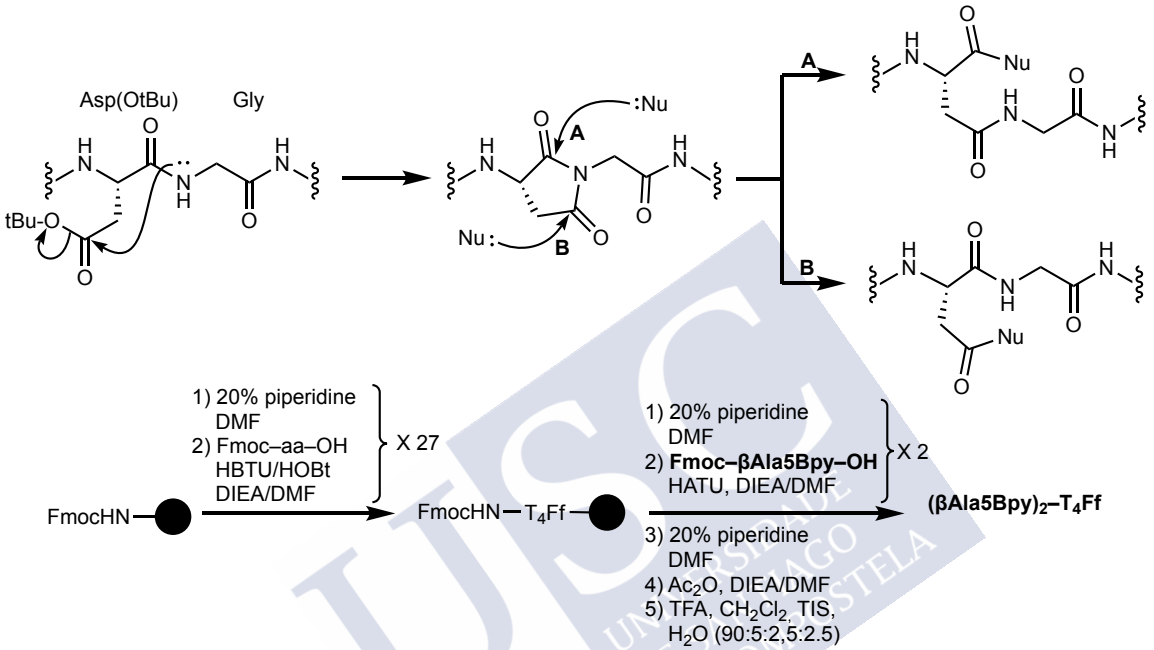


Figure 30. Top: scheme for the different pathways that a peptide can follow after the aspartimide formation and the latter nucleophile attack. Bottom: Solid-phase peptide synthesis of the T4Ff helicate precursor ligand.

Having at hand the desired peptides we proceeded with the study of their metal ion binding properties. Surprisingly, while bipyridine is reported to be weakly emissive, and is even considered non-fluorescent,[195,196] we found that the asymmetric 5'-amido-[2,2'-bipyridine]-5-carboxamide unit within the βAla residue was highly emissive, displaying intense band at c.a. 420 nm with a quantum yield of 0.37.[197] Importantly, its emission was quenched by coordination to Fe(II) ions, which could be exploited to monitor the formation of the T4Ff helicate. Thus, we recorded the emission spectra (λ_{exc} = 310 nm) of a 3 μM solution (9 μM monomer) of [**(βAla5Bpy)$_2$-T4Ff**]$_3$ in phosphate buffer (1 mM, 10 mM NaCl, pH 6.5) in the presence of increasing concentrations of $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ (Mohr's salt) as source of Fe(II) ions and observed a concentration-dependent quenching of the emission intensity of the

194 I. Coin, M. Beyermann, M. Bienert, *Nat. Protoc.* **2007**, *2*, 3247-3256.

195 S. Dhanya, P. K. Bhattacharyya, *J. Photochem. Photobiol. A Chem.* **199**2, *63*, 179-185.

196 M. Yagi, T. Kaneshima, Y. Wada, K. Takemura, Y. Yokoyama, *J. Photochem. Photobiol. A Chem.* **1994**, *84*, 27-32.

197 Y. Dong, T. Liu, X. Wan, H. Pei, L. Wu, Y. Yao, *Sens. Actuators B Chem.* **2007,** 241, 1139-1144.

bipyridine ligands. The emission intensity profile of the titration could be fitted to a 1:2 binding mode with dissociation constants for the first and second iron coordination of 5.5 and 6.6 µM, respectively (Figure 31, left).[198] UV/Vis titrations were also qualitatively consistent with the fluorescence data, showing a weak metal-to-ligand charge transfer band (MLCT) at about 535 nm in the presence of Fe(II) ions. The formation of the expected $[Fe(II)_2[(\mathbf{\beta Ala5Bpy)_2\text{-}T4Ff}]_3]^{4+}$ metallopeptide was also confirmed by mass spectrometry of the final solution of the titrations, which showed a peak at the expected mass of the molecular ion (m/z = 11084.6).

In order to study the chirality induction around the metal centers we measured the circular dichroism (CD) spectra of the assembled trimeric $[\mathbf{(\beta Ala5Bpy)_2\text{-}T4Ff}]_3$ scaffold and its dinuclear Fe(II) complex $[Fe(II)_2[(\mathbf{\beta Ala5Bpy)_2\text{-}T4Ff}]_3]^{4+}$. As expected from the original analysis, the observed positive Cotton effect at c.a. 330 nm is consistent with the formation of a ΛΛ- (or M) helicate ΛΛ-$[Fe(II)_2[(\mathbf{\beta Ala5Bpy)_2\text{-}T4Ff}]_3]^{4+}$. Furthermore, the intensity of the CD bands in absence of Fe(II) and the relatively small change in the CD spectra upon addition of iron(II) ions suggests that the Bpy units are strongly preorganized, even in absence of the metal ions, and that only a small rearrangement of the chromophores takes place upon coordination (Figure 31, right). This observation is also consistent with earlier computational studies with related bis-bipyridyl peptide ligands, which showed that the bipyridine residues have a large tendency to stack on top of each other; this stacking will presumably provide rigidity to the bis-bipyridyl trimer and facilitate the helical induction by the foldon domain.

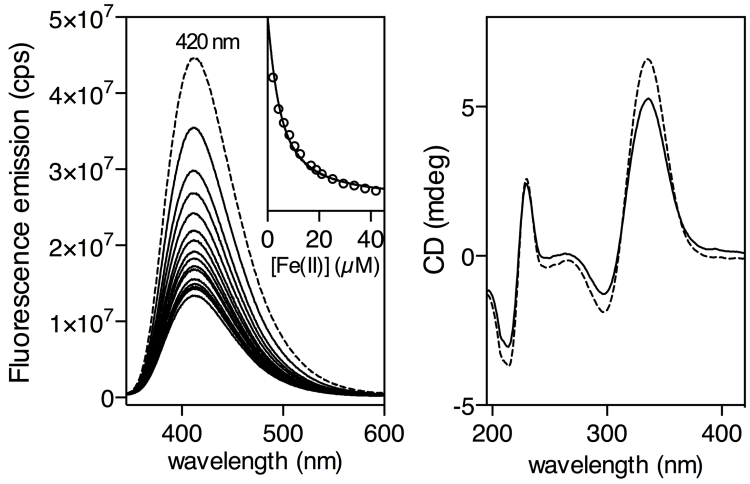


Figure 31. Left: fluorescence titration of a 3 µM (9 µM monomer) solution of the trimer $\mathbf{[(\beta Ala5Bpy)_2\text{-}T4Ff]_3}$ with increasing concentrations of Fe(II) ions. Inset shows emission at 420 nm upon excitation at 310 nm with increasing concentrations of Fe(II) ions, and the best fit to a 1:2 binding mode. Experiments were made in triplicate. Right: CD of a 6 µM solution (18 µM monomer) of $\mathbf{[(\beta Ala5Bpy)_2\text{-}T4Ff]_3}$ trimer (dashed line) and in the presence of Fe(II) ions (90 µM, solid line). All experiments in 1mM phosphate buffer, pH 6.5, 10 mM NaCl at 20 °C.

198 a) P. Kuzmič, *Anal. Biochem.* **1996**, 237, 260-273; b) P. Kuzmič, *Methods Enzymol.* **2009**, 467, 247-280

In order to gain some insight into the structure and stability of the peptide helicate the group of Professor Jean-Didier Maréchal at the *Universitat Autònoma de Barcelona* performed Molecular Dynamics (MD) simulations in explicit solvent and periodic boundary conditions. Analysis of the MD trajectories revealed that the structure of ΛΛ-$[Fe(II)_2[(\mathbf{\beta Ala5Bpy)_2\text{-}T4Ff}]_3]^{4+}$ appears to be highly stable, retaining its helicity, as well as the octahedral coordination geometry around the metal centers (Figure 32). Moreover, the T4Ff scaffold appears to be stable during the simulation showing no appreciable deformations as a result of the introduction of the artificial $(\beta Ala5Bpy)_2$ domain. The root-mean square deviation (RMSD) of the whole system was computed along the MD using the minimized initial structures as a reference, the trajectories attain relative stable RMSD after the first ~20 ns, that reach up to 1.99 ± 0.62 Å in average. A cluster analysis was performed on the full-length MD experiments showing a predominant conformations occupying about ~40% of the total conformation repartition. Overall, the results highlight that the computed model is very stable along the 100 ns of the MD and results consistent with the experimental data. Interestingly, the MD analysis revealed a hinge region with increased flexibility connecting the more rigid helicate and foldon domains, which suggests the replacement of the N-terminal Gly reside for a more conformationally restricted residue in future designs.

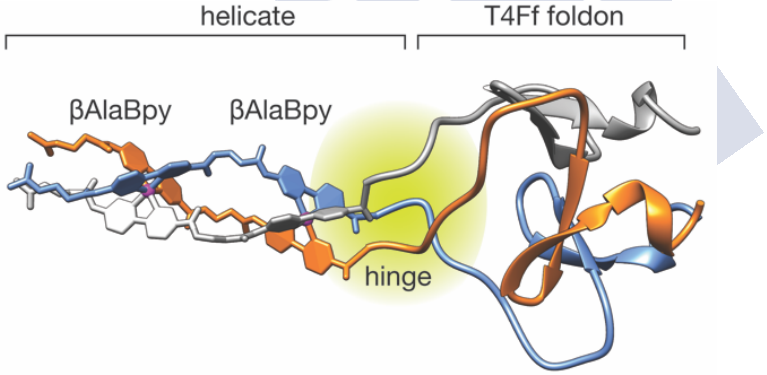


Figure 32. First cluster representative frame of the MD trajectory for ΛΛ-$[Fe(II)_2[\mathbf{(\beta Ala5Bpy)_2\text{-}T4Ff}]_3]^{4+}$ showing the stable T4Ff domain. Note the flexible hinge region between the rigid helicate and the T4Ff domain.

Having made a preliminary characterization of the T4Ff helicate, we studied its DNA binding properties by titrating a 2 μM solution of $[\mathbf{(\beta Ala5Bpy)_2\text{-}T4Ff}]_3$ peptide trimer (6 μM mononer) in the presence of saturating concentrations of Fe(II) according to the previous titrations (20 μM) with increasing concentrations of a three-way DNA junction (twDNA), and measuring the fluorescence anisotropy of the 5Bpy fluorophores at 420 nm after each addition of DNA. The titration profile could be fitted to a 1:1 binding mode, with a dissociation constant of ~2.1 μM. Titrations under the same conditions with a model double stranded DNA (dsDNA) led to a small, monotonic increase in the anisotropy, which is in tune with the formation of weak complexes or non-specific binding, which is consistent with previous studies with other helicates (Figure 33).[199]

199 J. C. Peberdy, J. Malina, S. Khalid, M. J. Hannon,J. Rodger, *Inorg. Biochem.* **2007**, *101*, 1937-1945.

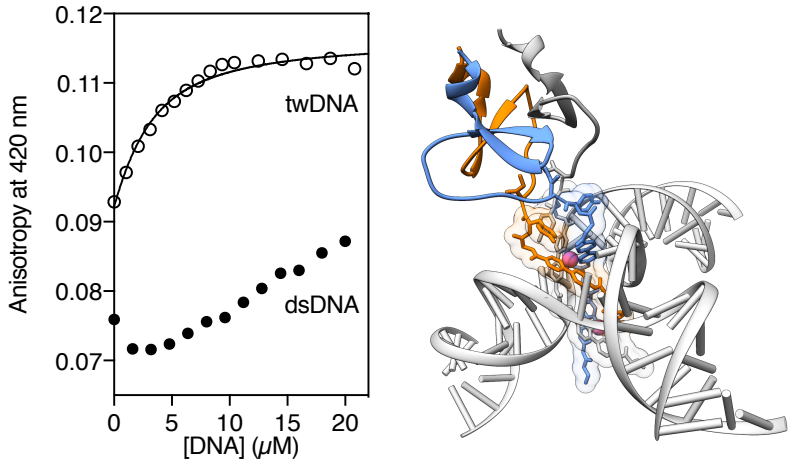


Figure 33. Left: anisotropy titration of ΛΛ-[Fe(II)$_2$[**(βAla5Bpy)$_2$-T4Ff**]$_3$]$^{4+}$ in 1 mM phosphate buffer, 10 mM NaCl with increasing concentrations of twDNA and best fit to a 1:1 binding mode (curve fitting done with *DynaFit*).[198] Right: model of the complex between ΛΛ-[Fe(II)$_2$[**(βAla5Bpy)$_2$-T4Ff**]$_3$]$^{4+}$ and the twDNA based on the reported PDB structure of a metal cylinder inserted in a twDNA.[117] twDNAs : 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: 5'–AAC ACA TGC AGG ACG GCG CTT–3' (only one strand shown).

Additional controls were carried out to verify the binding selectivity. Thus, a 2 μM solution of the trimer [**(βAla5Bpy)$_2$-T4Ff**]$_3$ (6 μM monomer) in the presence of saturating concentrations of Fe(II), was titrated with increasing concentrations of three DNAs featuring one, two or three mismatches.

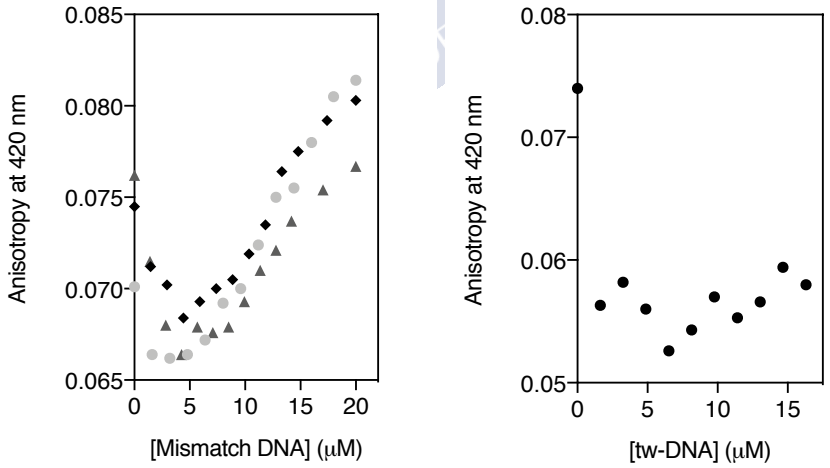


Figure 34. Left: fluorescence anisotropy titration of ΛΛ-[Fe(II)$_2$[**(βAla5Bpy)$_2$-T4Ff**]$_3$]$^{4+}$ in 1 mM phosphate buffer, 10 mM NaCl with increasing concentrations of mismatch G (one mistake, light grey circles), GA (two mistakes, black diamonds) and GGA (three mistakes, dark grey triangles) DNA's. Right: fluorescence anisotropy titration profile of [**(βAla5Bpy)$_2$-T4Ff**]$_3$ trimer peptide with twDNA. Lack of interaction is denoted in the oscillating pattern. twDNAs: 5'–CAC CGC TCT GGT CCT C–3'; 5'–CAG GCT GTG AGC GGT G–3'; 5'–GAG GAC CAA CAG CCT G–3'. Mismatch DNA: 5'–AAC ACA TGC **AGG** ACG GCG CTT–3' (common strand in all mismatch sequences). Mismatch G DNA: 5'– AAG CGC CGT **CGT** GCA TGT GTT–3'. Mismatch GA: 5'–AAG CGC CGT **CGA** GCA TGT GTT–3'. Mismatch GGA: 5'– AAG CGC CGT **GGA** GCA TGT GTT–3. Mismatched codons are highlighted in bold letters.

The response was similar to that observed in the presence of regular dsDNA, with a small and monotonic increment, suggesting the formation of nonspecific complexes. As additional control, incubation of [**(βAla5Bpy)$_2$-T4Ff**]$_3$ in absence of Fe(II) ions with increasing concentrations of the same twDNA showed no response, which indicates that the peptide by itself is incapable of high-affinity twDNA binding (Figure 34).

In addition to the spectroscopic studies, we also studied the DNA binding properties of the ΛΛ-[Fe(II)$_2$[**(βAla5Bpy)$_2$-T4Ff**]$_3$]$^{4+}$ helicate by electrophoretic mobility assays (EMSA) in polyacrylamide gel under non-denaturing conditions,[200] visualizing the DNA in the gel using *SybrGold* staining.[201] In agreement with the previously shown fluorescence titrations, incubation of the target twDNA with the helicate ΛΛ-[Fe(II)$_2$[**(βAla5Bpy)$_2$-T4Ff**]$_3$]$^{4+}$ resulted in the concentration-dependent appearance of a new retarded band, consistent with the formation of the expected complex between the DNA and the helicate (Figure 35, lanes 1-6). Additionally, the overall intensity of the lanes of the gel is progressively reduced in the presence of increasing concentrations of the metallopeptide, which suggests the formation of higher-order aggregates in the gel conditions.[202] On the other hand, incubation of a model double-stranded DNA with the peptide helicate did not show any new slow-migrating bands, which is in agreement with the expected low affinity for this form of DNA (Figure 35, lanes 7-10).

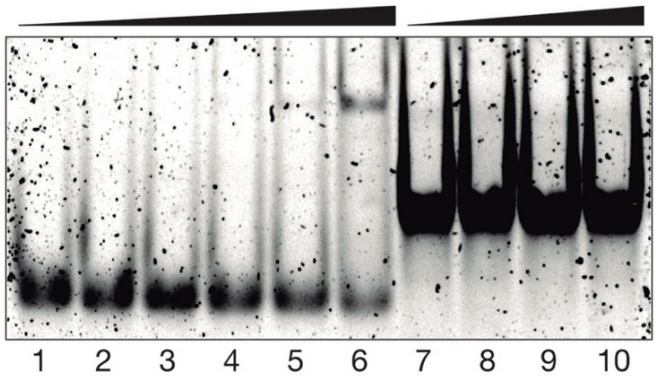


Figure 35. EMSA DNA binding of ΛΛ-[Fe(II)$_2$[**(βAla5Bpy)$_2$-T4Ff**]$_3$]$^{4+}$ Lanes 1–6, 200 nM twDNA with 0, 150, 250, 500, 1,000, and 2,000nM of [(βAla5Bpy)$_2$-T4Ff]$_3$ and 14 eq. of $(NH_4)_2Fe(SO_4)_2$ • 6 H_2O in each lane; lanes 7–10, 200 nM dsDNA with 0, 500, 1,000, and 2,000 nM of [**(βAla5Bpy)$_2$-T4Ff**]$_3$ and 14 eq. of $(NH_4)_2Fe(SO_4)_2$ • 6 H_2O in each lane. twDNAs: 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: 5'–AAC ACA TGC AGG ACG GCG CTT–3' (only one strand shown).

200 L. M. Hellman, M. G. Fried, *Nat. Protoc.* **2007**, *2*, 1849-1861.

201 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.

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

Conclusions

In summary, we have shown the potential of small protein domains for the precise structural organization of metal complexes. Modification of the T4 Fibritin foldon with metal-chelating bipyridine units allows the assembly of unique three-strand helicates in which the parallel orientation of the three helicate ligands is directed by the self-assembled T4Ff domain, and the chirality of the dinuclear helicate (*M* helicity or ΛΛ-configuration in the metal complexes) is selected by the relative orientation of the natural polyproline helices at the N-terminus of the T4Ff trimer. The final supramolecular peptide helicate ΛΛ-$[Fe(II)_2[(\mathbf{\beta Ala5Bpy})_2\mathbf{\text{-}T4Ff}]_3]^{4+}$ displays good *in vitro* affinity and selectivity towards three-way DNA junction.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Chapter 2

Stereoselective self-assembly of a DNA-binding helicate directed by the β-annulus motif

Summary

Peptides are highly versatile platforms for the precise design of supramolecular metal architectures and, particularly, for the controlled assembly of helicates. Herein, we show that the β-annulus peptide motif is capable of directing the stereoselective assembly of designed peptides containing 5Bpy chelating residues into three-stranded chiral peptide helicates. We show by a variety of methods that the resulting metallopeptides selectively bind with high affinity and selectivity to three-way junction DNA, thus demonstrating the potential of self-assembled short peptide motifs, such as the β-annulus, for the construction of DNA-binding supramolecular metallo-assemblies.

Introduction and precedents of β-annulus

Peptides are excellent supramolecular building blocks that can encode structural and functional information in their amino acid sequence. For this reason, some different peptide motifs—as coiled coils and β-hairpins—have been studied as different platforms for the construction of different systems with different applications.[54-56,86-87] Despite the enormous potential of these peptide motifs for the stereoselection of metal complexes, this application has not been exploited properly, and only some examples regarding coiled coils have been published.[69-83]

As an alternative and orthogonal platform to the leucine zipper we focus our attention in this chapter in the β-annulus peptide motif derived from Sesbania Mosaic Virus (SeMV), which consists in a 12 amino acid sequence. This short peptide is located in the center of the C subunit of the viral capsid and comprises the 12 residues located between the Gly^{48} and the Met^{59}(GISMAPSAQGAM).[203] This β-annulus motif is a trimeric structure that is stabilized by a network of hydrogen bonds stablished between the three monomers that form it.[204] This scaffold is considered the responsible of the maintenance and control of the curvature and the size of the assembled viral capsids (Figure 36). In SeMV is formed by three sets of β-interactions, where the residues 48-52 from a C-subunit stablished hydrogen bonds with the residues 55-58 of another symmetric subunit. In the polipeptide backbone is also present a turn of 120° that allows this interaction thanks to the presence of a proline residue in the sequence (Pro^{53}). This residue is also highly conserved in the sequence of other viruses such as TBSV, TCV, TNV, CarMV, CPMoV, demonstrating in this way the importance of this proline residue in the formation of the β-annulus in SeMV.[203]

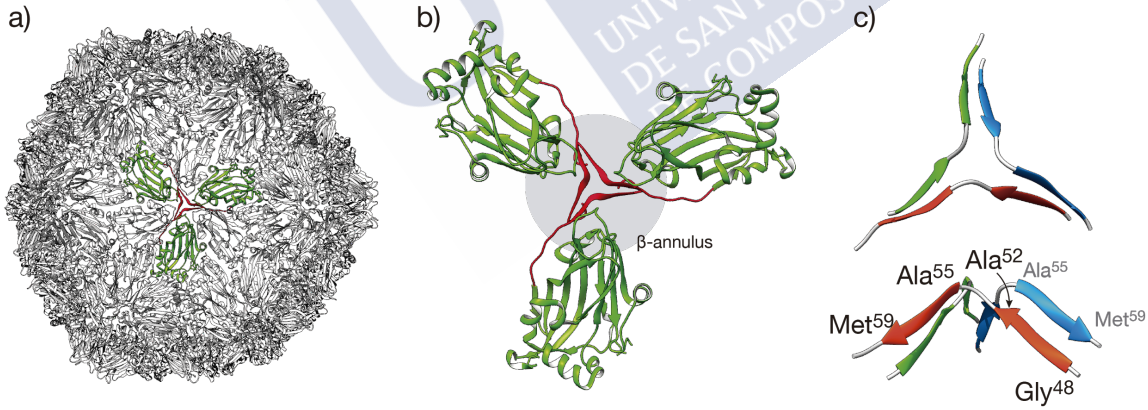


Figure 36. a) Structure of the SeMV capsid (PDB 1X35) where the β-annulus is highlighted in red and the C-subunit in green. b) Structure previously highlighted in the capsid of SeMV where the β-annulus is showed in red and the resulted trimer with the C-subunit highlighted in green. c) Structure of the β-annulus isolated from the rest of the structure showing the key residues involved in the trimerization of the structure

These β-annulus motifs are not exclusive of SeMV, and similar structures can be found in other viruses. Thus, for example, Matsuura *et al.* have described the application of a

203 P. S. Satheshkumar, G. L. Lokesh, M. R. N. Murthy, H. S. Savithri, *J. Mol. Biol.* **2005**, 353, 447-458.

204 A. M. Silva, M. G. Rossmann, *J. Mol. Biol.* **1987**, 197, 69-87.

β-annulus peptide fragment of the internal skeleton of tomato bushy stunt virus capsid for the assembly of virus-like nanocapsules with sizes of 30–50 nm in water,[205] as well as the assembly of modified versions to encapsulate anionic dyes, DNA, His-tagged proteins, or anionic CdTe quantum dots.[206] The SeMV β-annulus itself has been previously used for the formation of nanospheres by coupling dimerization peptide motifs to the ends of β-annulus. Among the different motifs used for the β-annulus oligomerization are the tryptophan zippers, or sequences that form β-sheets.[206] However, the potential of the β-annulus to encode the assembly of discrete structures has not been explored.

Objective

We propose that β-annulus could allow the programmed assembly of chiral dinuclear helicates, thus offering an attractive platform for the synergistic combination of coordination and peptide chemistry. In this chapter we will describe the synthesis of peptide derivatives of the β-annulus that integrate in their structure chelating ligands derived from 2,2'-bipyridine and that, in the presence of metal ions, are capable of selectively forming chiral, DNA-binding, dinuclear peptide helicates, directed by the self-assembly of the β-annulus chains.

Results and discussion

As an alternative platform orthogonal to the omnipresent coiled-coils, we focused our attention on the C_3-symmetric β-annulus motif of the Sesbania Mosaic Virus (SeMV), which consists of a short dodecapeptide comprising the residues between Gly^{48} and Met^{59} (GISMAPSAQGAM) of the coat protein of the SeMV virus capsid. The β-annulus is a symmetric three-way junction of two-stranded β-sheets formed between residues 48-52 of each strand with residues 55-58 of the symmetric peptide chain.[204] The backbone of the polypeptide displays a 120° turn that allows this arrangement thanks to the presence of a central proline residue (Pro^{53}).

As a coordinating unit we selected 2,2'-bipyridine, a ligand that has been frequently used in coordination chemistry and gives stable complexes with a variety of metal ions.[191] In particular, in our group we have described the synthesis of an amino acid derivative of 2,2'-bipyridine protected with the Fmoc group that can be integrated into standard solid-phase peptide synthesis protocols (Fmoc-βAla5Bpy-OH). Furthermore, the structure of this chelating amino acid, in which the Bpy unit is directly integrated into the peptide backbone, allows the effective coupling of the conformational preferences of the polypeptide chain with the coordination of the metal complexes, thus allowing the induction of chirality.

Inspection of the β-annulus structure (PDB code 1X33) clearly showed three symmetrically equivalent Ser^{54} residues located at the center of the β-annulus which are

---

205 K. Matsuura, K. Watanabe, T. Matsuzaki, K. Sakurai, N. Kimizuka, *Angew. Chem. Int. Ed Engl.* **2010**, 49, 9662-9665.

206 a) K. Matsuura, K. Watanabe, Y. Matsushita, N. Kimizuka, *Polym. J.* **2013**, 45, 529-534; b) K. Matsuura, T. Nakamura, K. Watanabe, T. Noguchi, K. Minamihata, N. Kamiya, N. Kimizuka, *Org. Biomol. Chem.* **2016**, 14, 7869-7874; c) S. Fujita, K. Matsuura, *Chem. Lett.* **2016**, 45, 922-924.

ideally positioned to serve as anchor points for the introduction of the coordinating βAla5Bpy residues. Exploratory molecular modeling studies confirmed that a mutated peptide featuring a set of two chelating βAla5Bpy residues attached to the side chain of Dab residue in that position could coordinate a pair of metal ions forming a dinuclear helicate without significantly distorting the β-annulus structure (Figure 37).

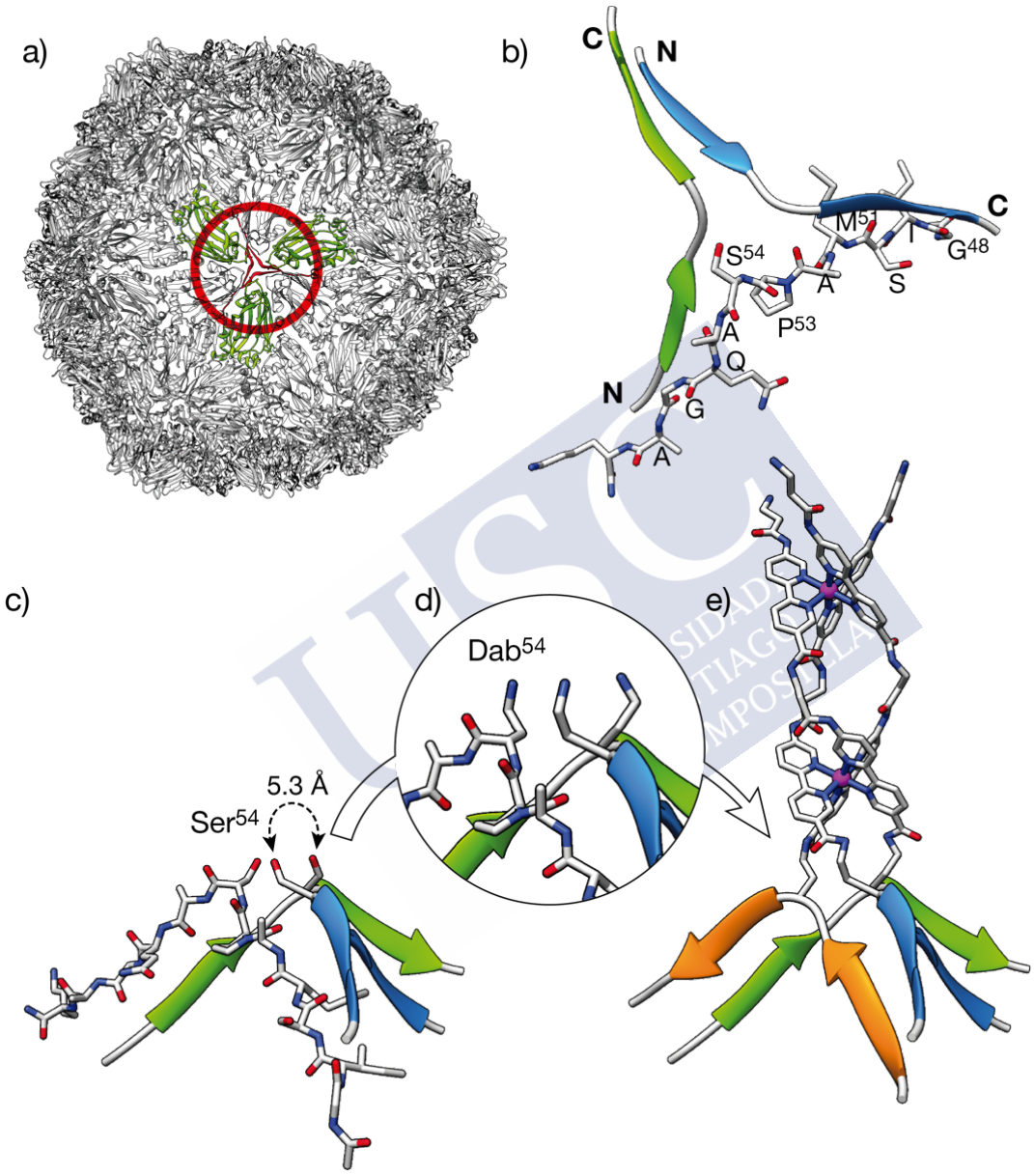
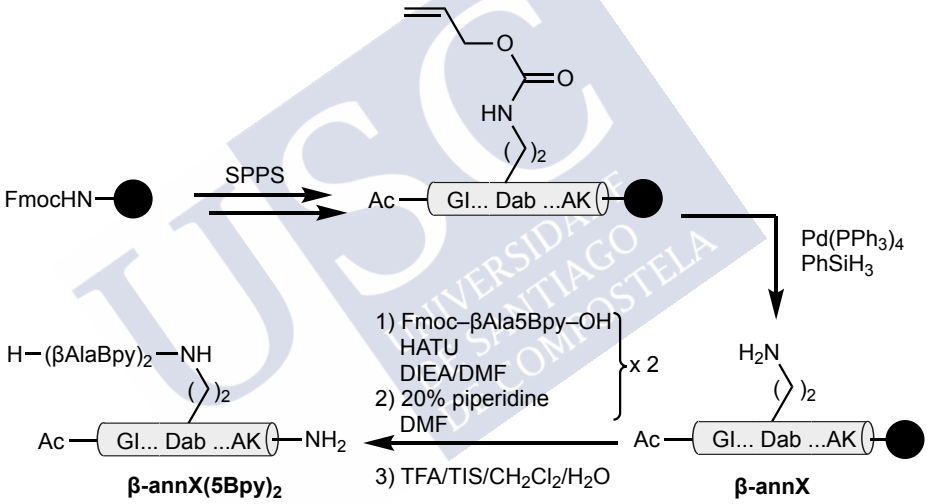


Figure 37. Design of the β-annulus helicate. Top a) Structure of the Sesbania Mosaic Virus (SeMV) capsid, highlighting the β-annulus motif (in red), at the center of the C subunit trimer shown in green (PDB ID 1X35); b) Isolated β-annulus showing the relative orientation of the three peptide chains, as well as the natural residues in one of the symmetric peptide strands. Note the position of Ser^{54} near the center of the three-way junction. Bottom: modification of the β-annulus to introduce the coordinating βAla5Bpy residues c) detail of the β-annulus highlighting the distance between Ser^{54} residues; d) Ser^{54} are mutated into Dab residues, that serve as anchor points for the introduction of the chelating βAla5Bpy residues to yield the **β-annX(5Bpy)$_2$** ligand precursor of the helicate shown in e).

Therefore, the final peptide ligand target sequence, Ac-G[48]IS-nL[51]-AP-Dab(Alloc)[54]-AQGAK[59]-NH_2, includes the orthogonally protected L-2,4-diaminobutyric acid residue in place of Ser[54] in the model β-annulus sequence (X(Alloc)[54]). Additionally, we replaced the residue Met[51] with an isosteric nor-Leucine residue (nL[51]) to avoid potential problems related to the oxidation of the thioether side chain of methionine.[207] Finally, also for synthetic reasons, and in order to promote the solubility of the helicate precursor peptide, we also replaced the C-terminal Met[59] by a ionizable Lys residue (K[59]).

The synthesis of the target peptide was performed following standard solid-phase peptide synthesis protocols.[194] Once the β-annulus strand was fully assembled, the Alloc group was selectively deprotected under catalytic conditions ($Pd(PPh_3)_4$ and $PhSiH_3$, Scheme 2) and the bipyridine chelating residues (Fmoc-βAla5Bpy-OH) were sequentially assembled onto the orthogonally-deprotected β-annulus sequence using HATU as coupling agent. Finally, the deprotection and release of the peptide from the resin was carried out using standard conditions by treatment with an acidic TFA cocktail. The peptide was purified by reverse-phase HPLC, and the identity confirmed by MS (MALDI-TOF).



Scheme 2. Scheme of the solid-phase peptide synthesis (SPPS) of the β-annulus helicate precursor peptide ligand **β-annX(5Bpy)$_2$**.

Taking into account that the parent β-annulus motif trimerizes in solution, we studied the Fe(II) binding properties of the **β-annX(5Bpy)$_2$** peptide, and its capacity for the formation of helicates. We found that the asymmetric 5'-amido-[2,2'-bipyridine]-5-carboxamide unit within the βAla5Bpy residue was highly emissive, and that its emission was quenched by its coordination to Fe(II) ions, which could be exploited to monitor the formation of helicates. Therefore, we added increasing amounts of $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ (Mohr's salt) to a buffered solution of the **β-annX(5Bpy)$_2$** peptide, and recorded the emission from the Bpy ligand after each addition. We observed a progressive decrease in the emission intensity of the 2,2'-bipyridine as the concentration of Fe(II) in the cuvette increased (Figure 38). That is, the coordination of the Fe(II) ions exerts a quenching effect on the Bpy ligands. This quenching could be used to monitor the formation of the

207 C. Schöneich, J. Yang, *J. Chem. Soc. Perkin Trans.* **1996**, 915-924.

helicate. Thus, the titration profile of the emission intensity at 420 nm could be adjusted to a 1:2 interaction model with the *DynaFit* program,[198] with tight dissociation constants: K_{D1} = 14.2 ± 5.5 μM for the first coordination and K_{D2} = 15.9 ± 1.3 μM for the second.

Having demonstrated that the **β-annX(5Bpy)$_2$** peptide self-assembles and coordinates Fe(II) ions, we studied whether the formation of the [Fe(II)$_2$(**β-annX(5Bpy)$_2$**)$_3$]$^{4+}$ metallopeptide was stereoselective. Thus, we recorded the circular dichroism (CD) spectra of a 21 μM solution of the **β-annX(5Bpy)$_2$** peptide and of the same solution in the presence of saturating concentration of Fe(II) ions. The spectrum of the **β-annX(5Bpy)$_2$** peptide by itself displayed the typical signature of a β-sheet structure. Furthermore, the spectrum in the presence of the Fe(II) ions displayed a clear positive CE band centered at c.a. 320 nm, indicating the formation of a helicate complex with M helicity. Also, strong preorganization that the system confers to the helicate region can be observed in the small variation of the CD spectra before and after the coordination process (Figure 38, right).

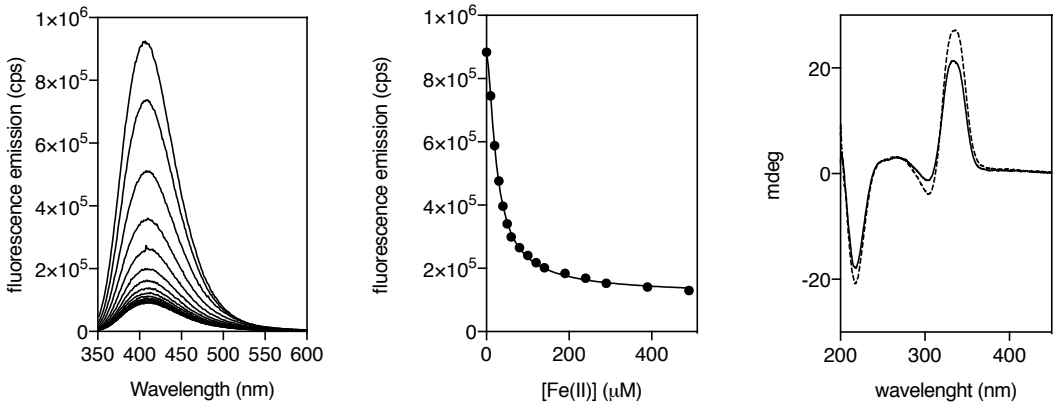


Figure 38. Left: emission spectra of a 20 μM solution of **β-annX(5Bpy)$_2$** with increasing concentrations of $(NH_4)_2Fe(SO_4)_2$ • 6 H_2O. Center: Titration profile of the maximum emission wavelength at 410 nm with increasing concentrations of Fe(II) ions and best fit according to the 1:2 model in *DynaFit*. Experimental data points corresponds to the average of three independent titrations; right: CD of a 7 μM solution (21 μM monomer) of [Fe(II)$_2$[**β-annX(5Bpy)$_2$**]$_3$]$^{4+}$ (dashed line) and in the presence of Fe(II) (200 μM, solid line). All experiments were made in 10 mM Hepes buffer, pH 6.5, 100 mM NaCl at 20 °C.

Helicates are known to selectively interact with three-way DNA junctions by inserting the metallocylinder into the hydrophobic center of the three-way junction.[117,135a] In order to study the binding of the β-annulus helicate to the DNA, we relied on the strong quenching effect of the helicate in nearby fluorophores.[179] Therefore, we prepared a 2 μM solution of a fluorescein-labeled three-way DNA junction (FAM-twDNA), in 10 mM HEPES buffer, 100 mM NaCl, pH 6.5, and recorded its emission spectrum upon excitation at 490 nm after addition of successive aliquots of a solution containing the **β-annX(5Bpy)$_2$** peptide and $(NH_4)_2Fe(SO_4)_2$ • 6 H_2O (400 μM and 6 mM, respectively). The titration profile of the emission at 515 nm could be fitted to an 1:1 binding mode (twDNA/ΛΛ-[Fe(II)$_2$(**β-annX(5Bpy)$_2$**)$_3$]$^{4+}$) with an apparent K_D of ≈ 2.82± 0.57 μM.

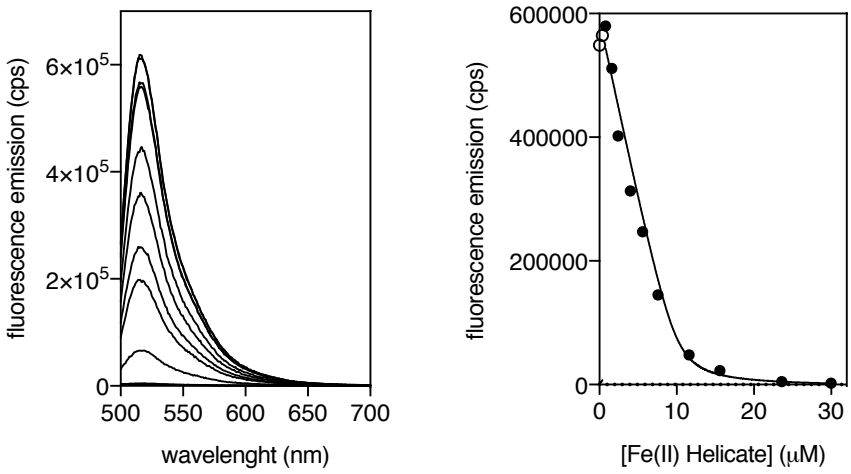


Figure 39. Left: Fluorescence spectra of a 2 μM solution of FAM-twDNA in the presence of increasing concentrations of ΛΛ-[Fe(II)2(**β-annX(5Bpy)$_2$)$_3$**]$^{4+}$. Right: titration profile and best fit to a 1:1 binding mode for twDNA (solid line, black circles); open circles were excluded for curve fitting. All experiments in 10 mM HEPES buffer, 100 mM NaCl, pH 6.5. twDNAs: 5'–CAC CGC TCT GGT CCT C–3'; 5'–CAG GCT GTG AGC GGT G–3'; 5'–GAG GAC CAA CAG CCT G–3'.

In addition to the spectroscopic studies, we also studied the DNA binding properties of the ΛΛ-[Fe(II)$_2$(**β-annX(5Bpy)$_2$)$_3$**]$^{4+}$ helicate by electrophoretic mobility assays (EMSA) in polyacrylamide gel under non-denaturing conditions,[200] and using *SybrGold* as the DNA stain.[201]

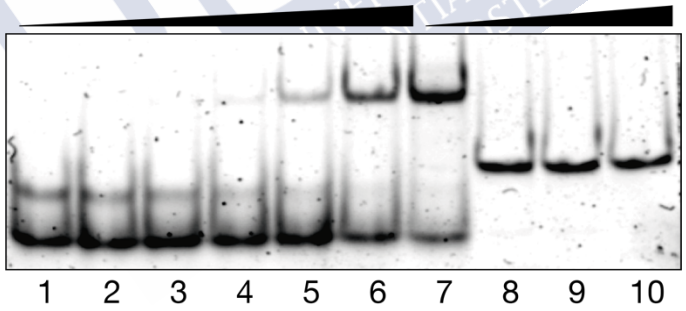


Figure 40. EMSA DNA binding studies results for the ΛΛ-[Fe$_2$(**β-annX(5Bpy)$_2$)$_3$**]$^{4+}$ helicate. Lanes 1-7, 200 nM twDNA with 0, 25, 50, 100, 250, 500 and 1000 nM of [**β-annX(5Bpy)$_2$**]$_3$ and 20 eq. of $(NH_4)_2Fe(SO_4)_2$ • 6 H_2O in each lane; lanes 8-10, 200 nM dsDNA with 0, 500 and 1000 nM of [**β-annX(5Bpy)$_2$**]$_3$ and 20 eq. of $(NH_4)_2Fe(SO_4)_2$ • 6 H_2O in each lane. Samples were resolved on a 10% nondenaturing polyacrylamide gel and 0.5 × TBE buffer over 40 min at 25 °C and stained with SyBrGold (5 μL in 50 mL of 1× TBE) for 10 min, followed by fluorescence visualization. twDNAs: 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: 5'–AAC ACA TGC AGG ACG GCG CTT–3' (only one strand shown).

In agreement with the fluorescence titration studies, incubation of the target three-way DNA, twDNA, with increasing concentrations of the peptide helicate resulted in the appearance of new slow-migrating bands, consistent with the formation of the expected twDNA/ΛΛ-[Fe(II)$_2$(**β-annX(5Bpy)$_2$)$_3$**]$^{4+}$ adduct (Figure 40, lanes 1-7). Remarkably, no smearing is observed, even at high concentrations of the ΛΛ-[Fe$_2$(**β-annX(5Bpy)$_2$)$_3$**]$^{4+}$ helicate, and only one band is formed, thus demonstrating the formation of a unique

complex.[202] On the other hand, incubation of a model dsDNA with ΛΛ-[Fe(II)$_2$(**β-annX(5Bpy)$_2$**)$_3$]$^{4+}$ helicate did not induce the formation of new retarded bands, even at high concentrations of the ΛΛ-[Fe(II)$_2$(**β-annX(5Bpy)$_2$**)$_3$]$^{4+}$ helicate, which clearly confirms the low affinity of this complex for regular B-DNA (Figure 40, lanes 8-10).

Conclusions

We have demonstrated that the short β-annulus motif from the Sesbania Mosaic Virus can be modified to direct the stereoselective self-assembly of peptide helicates with a M helicity. Furthermore, the resulting helicate displays high affinity and selectivity towards three-way DNA junction over regular dsDNA, as shown both by spectroscopic and electrophoretic studies.

Chapter 3

Dynamic stereoselection of inert DNA-binding peptide helicates

Summary

Designed artificial metal-chelating peptides encode in their sequence the stereoselective folding into three-stranded helicates. By using Co(II), the dynamically selected metallopeptides can be locked as kinetically-inert complexes by *in situ* oxidation to Co(III), thus providing a straightforward two-step approach for the stereoselective formation of kinetically inert Co(III) peptide helicates. Both the kinetically labile Fe(II) and Co(II) helicates, as well as the kinetically-inert Co(III) complexes, selectively recognize three-way DNA junction , and a rhodamine-labeled analog can be internalized in living cells to selectively stain DNA replication factories *in vivo*.

Introduction and precedents

Helicates are discrete metal complexes in which one or more organic ligands coil around—and coordinate—two or more metal ions.[114,208] As a result of ligand coiling, helicates are inherently chiral species that can appear as two enantiomers, according to the orientation in which the ligands twist around the helical axis defined by the metal centers. Besides their interest in supramolecular chemistry, helicates have shown promising DNA-binding properties that have been associated with antimicrobial and antitumoral effects.[175,209]However, despite the efforts of many groups, helicates are still considered as rather exotic DNA binding agents. One of the major obstacles for the development of bioactive helicates is the complexity of their synthesis, particularly regarding to the stereochemical control of the metallocylinder, which difficult the efficient access to structural variants required for the screening and optimization of their biological properties. Notwithstanding some noteworthy examples that describe the stereoselective synthesis of chiral helicates, [112,135,210,211] no general strategy for the synthesis of these system has been reported yet. Therefore, we decided to approach the stereoselective synthesis of helicates as a test case to demonstrate the advantages of using peptides for the design of metallo-supramolecular entities. Indeed, we have previously described the synthesis of a 2,2'-bipyridine amino acid derivative,[112a] and its incorporation into peptide ligands that predictably fold into complexes with defined chirality in the presence of metal ions, including self-assembled helicates.[112,135a] Unfortunately, in this original design of the helicate peptide ligands, two fairly hydrophobic β-turn sequences (Gly–(D/L)-Pro–Gly) were introduced in the sequence , making these peptide ligands poorly water soluble and prone to aggregation.[135a,212] Furthermore, the resulting peptide helicates, based on relatively labile Fe(II)-bipyridine coordination, were inherently dynamic, which limited their applications.

---

208 a) M. Albrecht, *Chem. Rev.* **2001**, *101*, 3457-3498; b) M. Albrecht, *Angew. Chem. Int. Ed.* **2005**, 44, 6448-6451.

209 a) D. H. Simpson, A. Hapeshi, N. J. Rogers, V. Brabec, D. G. J. Clarkson, J. Malina, H. Song, D. J. Fox, O. Hrabina, G. L. Kay, A. K. King, A. D. Millard, J. Moat, D. I. Roper, N. R. Waterfield, P. Scott, *Chem. Sci.*, **2019**, *10*, 9708-9720; b) 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; c) A. D. Faulkner, R. A. Kaner, Q. M. A. Abdallah, G. Clarkson, D. J. Fox, P. Gurnani, S. E. Howson, R. M. Phillips, D. I. Roper, D. H. Simpson, P- Scott, *Nat. Chem.*, **2014**, *6*, 797-803; d) H. Song, N. J. Rogers, S. J. Allison, V. Brabec, H. Bridgewater, H. Kostrhunova, L. Markova, R. M. Phillips, E. C. Pinder, S. L. Shepherd, L. S. Young, J. Zajac, P. Scott, *Chem. Sci.*, **2019**, *10*, 8547-8557.

210 a) S. E. Howson, L. E. N. Allan, N. P. Chmel, G. J. Clarkson, R. van Gorkum, P. Scott, *Chem. Commun.*, **2009**, 1727-1729; b) S. E. Howson, L. E. N. Allan, N. P. Chmel, G. J. Clarkson, R. J. Deeth, A. D. Faulkner, D. H. Simpson, P. Scott, *Dalton Trans.*, **2011**, *40*, 10416-10433; c) S. E. Howson, G. J. Clarkson, A. D. Faulkner, R. A. Kaner, M. J. Whitmore, P. Scott, *Dalton Trans.*, **2013**, *42*, 14967-14981.

211 a) T. Haino, H. Shio, R. Takano, Y. Fukazawa, *Chem. Commun.* **2009**, 2481-2483; b) W. Chen, X. Tang, W. Dou, B. Wang, L. Guo, Z. Ju, W. Liu, *Chem. Eur. J.* **2017**, *23*, 9804-9811.

212 A. M. C. Marcelino, L. M. Gierasch, *Biopolymers* **2008**, *89*, 380-391.

Objective

The –Pro–Gly– loops originally reported for chiral induction and proper folding were very hydrophobic and resulted in low solubility, and a marked tendency to form gels of the peptide ligands.[112,135a] As first step towards the development of well-behaved metallopeptide helicates (MPH), we propose the replacement of the original loops with alternative sequences that also induce β-turn formation within short peptides, while ensuring good solubility and coordination properties.

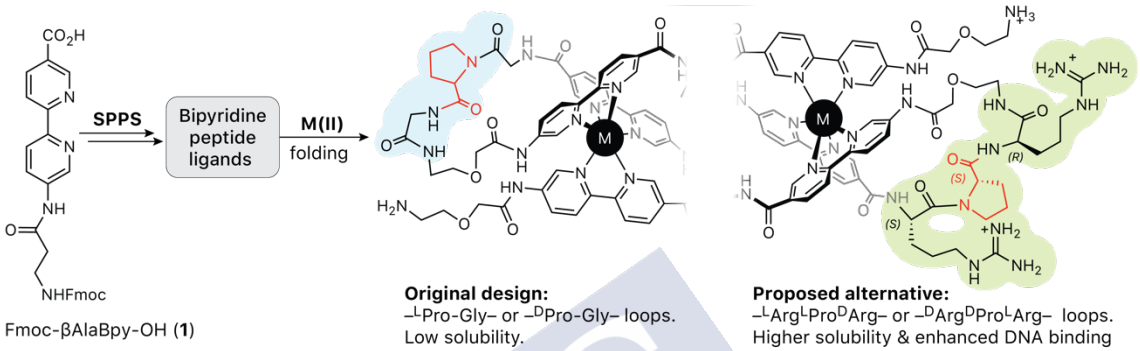


Figure 41. Proposed modifications in the β-turn inducing sequences to define soluble and well-behaved MPHs.

On the other hand, until now we have exploited dynamic processes in which mixtures of the peptide ligands and Fe(II) ions evolve towards the more stable Fe(II)-metallocylinders under thermodynamic control. This strategy introduces severe limitations for the application of such helicates, because these self-assembled helicates are intrinsically dynamic and could exchange ligands, change their structure or association state, and even disassemble in response to changes in their environment.[213] To overcome these limitations, a more robust and kinetically inert interaction would be desirable.[214] Therefore, with this aim in mind, we propose to dynamically assemble the metallocylinders in the presence of labile Co(II) ions, and then lock the equilibrium by oxidation of the complexes in the form of kinetically inert Co(III) species.

Results and discussion

Herein we describe an alternative peptide helicate design with improved solubility, and the synthesis of kinetically inert helicates through the oxidative modification of dynamically assembled Co(II) helicates to yield inert Co(III) metallocylinders with thermodynamically controlled chiral selection. To achieve this, methodology proposed by Lusby *et al.* was followed in order to oxidize the Co(II).[215]

---

213 W. M. Bloch, G. H. Clever, *Chem. Commun.* **2017**, 53, 8506-8516.

214 a) S. V. Wegner, J. P. Spatz, *Angew. Chem. Int. Ed.* **2013**, 52, 7593-7596; b) S. V. Wegner, F. C. Schenk, J. P. Spatz, *Chemistry* **2016**, 22, 3156-3162.

215 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.

Our first idea based on the original –Gly-(D/L)–Pro–Gly– loop (ligand **GPG**, Scheme 3) employed in the first helicate design–was to shorten this loop sequence by removing the N-terminal glycine residues, resulting in a new –(D/L)–Pro–Gly– connector. This should promote a tighter and more compact β-turn that might favor the metal coordination properties of the new ligands by reducing the flexibility of the hinge that connect the coordinating βAla5Bpy subunits. Unfortunately, the solubility of the resulting peptides was even worse than in the case of the –Gly-(D/L)–Pro–Gly– loop sequence, forming gels that hampers the purification of the obtained crude.

For this reason and in order to improve the solubility of the peptide ligands (and that of the final helicate) we decided to replace the β-turn modules –Gly-(D/L)–Pro–Gly–with alternative sequences that contained charged and/or polar amino acids and, just like the original modules, were capable of inducing a reverse β-turn to direct the folding of the peptide chain into the desired three-stranded helicates. Based on the known tendency of heterochiral sequences to promote the formation of type *II* (L-D) or type *II'* (D-L) β-turns,[216] we focused our attention on the combination residues with mixed chirality (D-D-L, or its mirror image L-L-D), which was found to increase turn formation and stability in short peptides.[217] This pattern was very attractive for our purposes because, in contrast with the previously used –Gly-(D/L)–Pro–Gly– sequence, or the widely used heterochiral diproline module, D-Pro–L-Pro,[218] it allowed us great freedom in the choice of the residues in the loop, as long as they maintained their relative chirality in the sequence. Thus, in order to maximize the solubility, and considering the prevalence of Arg residues in protein-DNA complexes,[219] as well as reports of the beneficial effects of arginine residues in DNA-binding metallopeptides[111] and metallocylinders,[147] we selected the tripeptide sequences –L-Arg–L-Pro–D-Arg– and –D-Arg–D-Pro–L-Arg– to define the β-turns in the two enantiomeric peptide ligands **LLD** and **DDL**, respectively. These turns direct the folding of the strand of the metallocylinder equipped with three sets of two βAla5Bpy units in tandem. The Fmoc-βAla5Bpy-OH unit is designed so that the 2,2'-bipyridine (Bpy) ligand is integrated into the peptide backbone, allowing a strong coupling between the conformational preferences of the peptide ligand and the geometry of the resulting complexes.

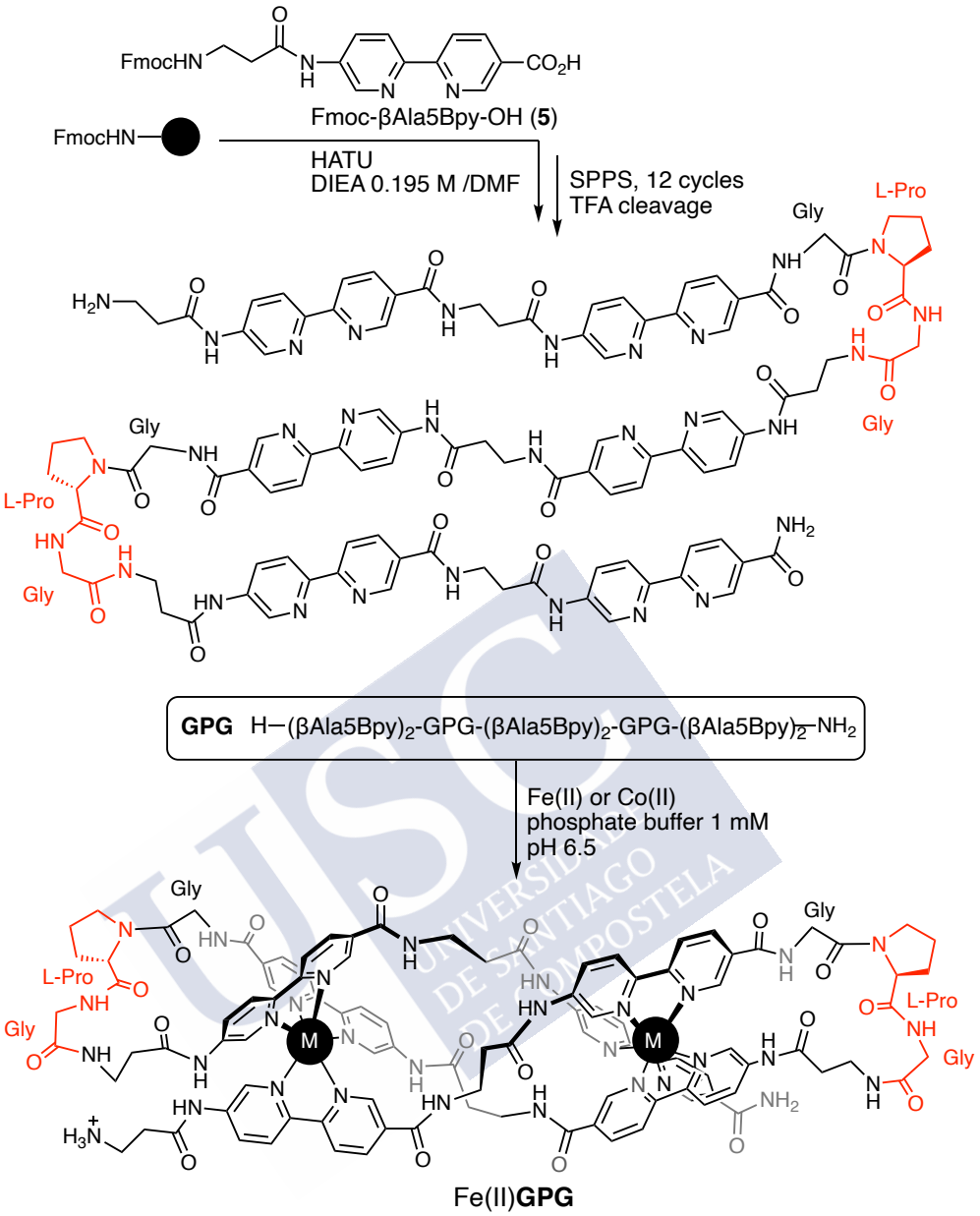
---

216 a) G. D. Rose, L. M. Glerasch, J. A. Smith, in *Adv. Prot. Chem.* (Eds.: C.B. Anfinsen, J.T. Edsall, F.M. Richards), Academic Press, **1985**, pp. 1-109; b) V. Brenner, F. Piuzzi, I. Dimicoli, B. Tardivel, M. Mons, *Angew. Chem. Int. Ed.* **2007**, *46*, 2463-2466.

217 M. G. Bomar, B. Song, P. Kibler, K. Kodukula, A. K. Galande, *Org. Lett.* **2011**, *13*, 5878-5881.

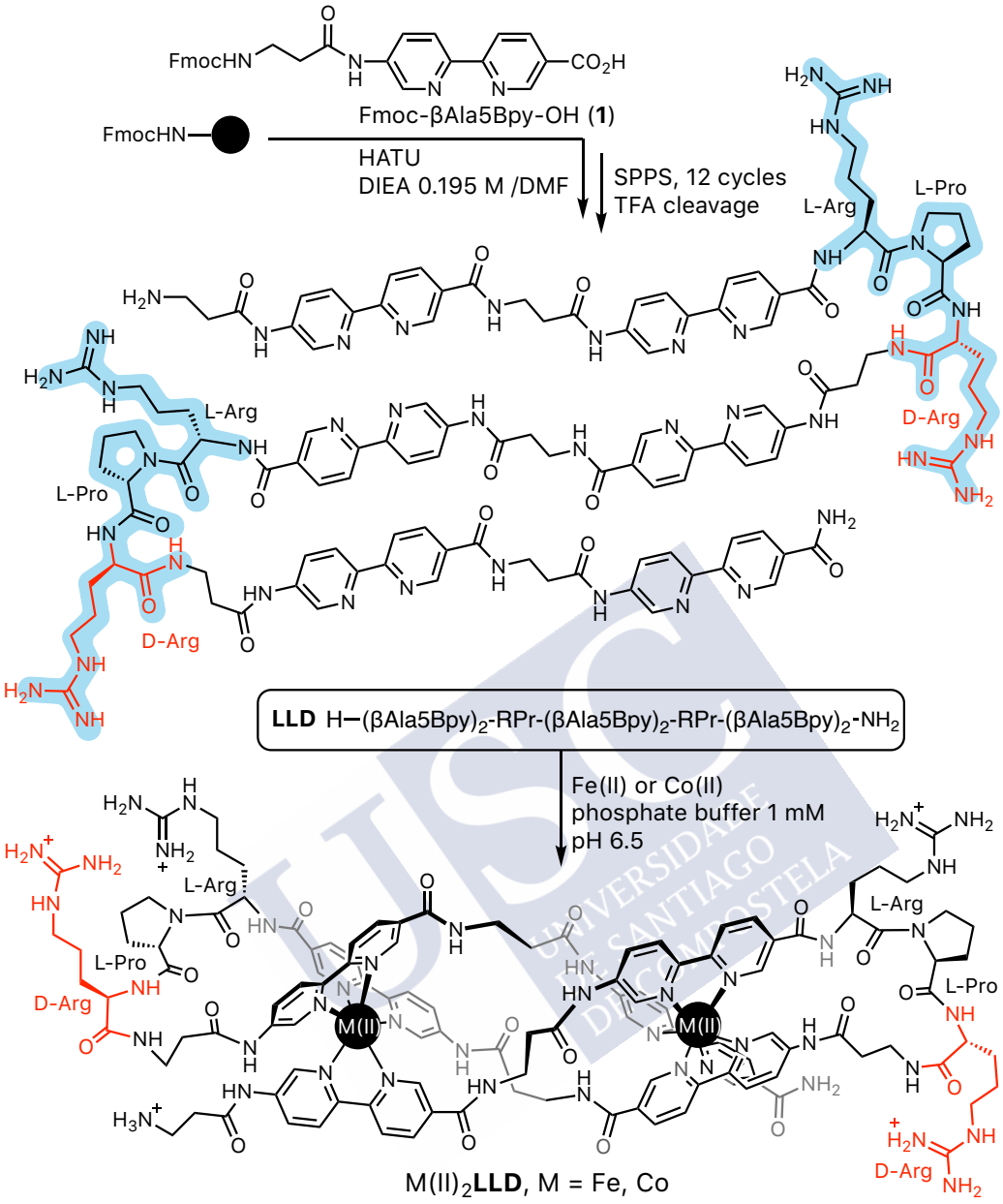
218 B. Chatterjee, I. Saha, S. Raghothama, S. Aravinda, R. Rai, N. Shamala, P. Balaram, *Chem. Eur. J.* **2008**, *14*, 6192-6204.

219 a) J. DeRouchey, B. Hoover, D. C. Rau, *Biochemistry*, **2013**, *52*, 3000-3009; b) S. M. West, R. Rohs, R. S. Mann, B. Honig, *J. Biomol. Struct. Dyn.* **2010**, *27*, 861-866; c) C. Crane-Robinson, A. I. Dragan, P. L. Privalov, *Trends Biochem. Sci.* **2006**, *31*, 547-552; d) D. P. Mascotti, T. M. Lohman, *Biochemistry* **1997**, *36*, 7272-7279.



Scheme 3. Solid-phase peptide synthesis of the helicate ligand **GPG** using the Fmoc-βAla5Bpy-OH building block and the β-turn sequence PG. The obtained compound showed really poor water solubility and other problems as gel formation.

The precursor peptide ligands were assembled by standard Fmoc solid-phase peptide synthesis protocols.[194] The Bpy ligands were introduced in the form of an Fmoc-protected unit Fmoc-βAla5Bpy-OH, which was synthesized with minor modifications of the previous reported procedures.[111] The peptides were then purified by HPLC, and their identity confirmed by mass spectrometry. As expected, the presence of four charged arginine residues in the sequence made these new peptide ligands with Arg-Pro-Arg turns readily soluble in water, radically improving the poor solubility of our previous design featuring hydrophobic –Gly-(D/L)–Pro–Gly– turns.



Scheme 4. Solid-phase peptide synthesis of the helicate ligand **LLD** using the Fmoc-βAla5Bpy-OH building block. β-turn sequences (RPr) are highlighted in light blue. After cleavage and purification, the helicate is stereoselectively folded under thermodynamic control in the presence of coordinating metal ions, Fe(II) or Co(II). The **LLD** ligand contains two loops featuring the heterochiral L-Arg–L-Pro–D-Arg sequence, while the **DDL** contains enantiomeric D-Arg–D-Pro–L-Arg loops (H-(βAla5Bpy)$_2$-rpR-(βAla5Bpy)$_2$-rpR-(βAla5Bpy)$_2$-NH$_2$). L-amino acids are indicated in upper-case, and D-amino acids in lower case (i.e., R, P for L-Arg and L-Pro, and r, p for D-Arg and D-Pro, respectively).

Having at hand the desired peptide ligands (**LLD** and **DDL**), our next step was to study their programmed folding in the presence of metal ions to stereoselectively form the desired helicates. To our surprise, while the Bpy ligand is weakly emissive, and for all practical purposes considered non-fluorescent,[195,196] our preliminary experiments with the **DDL** ligand showed that the 5'-amido-[2,2'-bipyridine]-5-carboxamide unit within the

βAla residue was highly emissive ($\Phi \approx 0.37$),[220] featuring an intense emission band at c.a. 420 nm upon excitation at 310 nm. Furthermore, we found that this band was quenched in the presence of metal ions, and that we could exploit this quenching effect to monitor the coordination and formation of the derived Fe(II) or Co(II) metallocylinders. Thus, a fluorescence titration of a 2 μM solution of the **DDL** peptide in phosphate buffer (1 mM, 10 mM NaCl, pH 6.5) with increasing concentrations of $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ resulted in a series of emission spectra with decreasing emission intensity due to the quenching of the 5'-amido-[2,2'-bipyridine]-5-carboxamide units by coordination of Fe(II) ions. The resulting titration profile at 420 nm could be fitted to a 1:2 binding mode with dissociation constants for the first, and second iron coordination of 0.55 and 0.41 μM respectively.[198] Likewise, titration of the same **DDL** peptide ligands with Co(II) also resulted in a significant quenching of the bipyridine emission band, which allowed us to study the formation of the equivalent metallocylinder with dissociation constants of 0.54 and 0.44 μM. Additionally, MALDI spectra of the solutions at saturating concentrations of metal ions showed peaks at 2554.04 and 2561.10 mass, corresponding to the formation of the $[Fe(II)_2\mathbf{DDL}]^{4+}$ and $[Co(II)_2\mathbf{DDL}]^{4+}$ helicates, respectively.

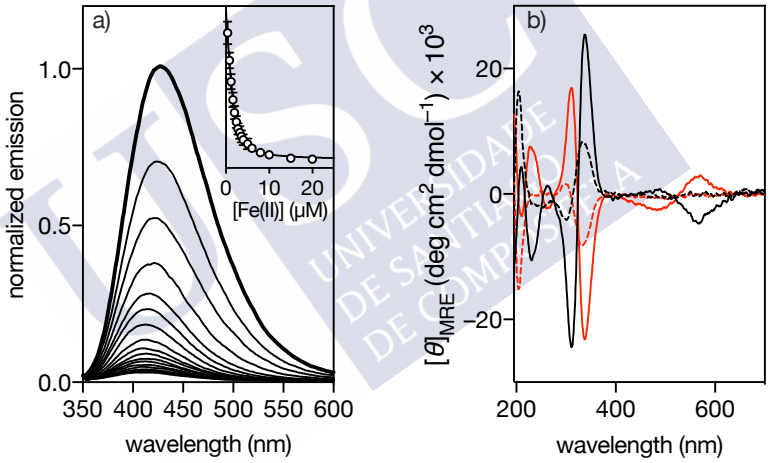


Figure 42. Titration of the **DDL** ligand strand. a) Normalized emission spectrum of a 2 μM solution of **DDL** in 1 mM phosphate buffer NaCl 10 mM, pH 6.5 (thick line), and spectra of the same solution in the presence of increasing concentrations of Fe(II) ions (thin lines); inset showing the titration profile of three independent experiments and best fit to a one to two binding model (same range as main plot); b) Circular dichroism spectra of 5 μM solutions of **LLD** (black dashed line) and **DDL** (red dashed line) in 1 mM phosphate buffer NaCl 10 mM, pH 6.5, and in the presence of 25 μM of Fe(II) ions (same colors, continuous lines).

Importantly, the **DDL** and **LLD** ligands displayed mirror image CD spectra, dominated by a strong positive Cotton effect band at c.a. 320 nm for **LLD** and an equivalent negative Cotton effect band for the **DDL** peptide, which pointed to a significant preorganization of the peptide chain and the effective chiral induction by the –D-Arg–D-Pro–L-Arg– and –L-Arg–L-Pro–D-Arg– β-turn modules, respectively. Addition of Fe(II) to **LLD** resulted in a large increase of the positive Cotton effect band at c.a. 320 nm, as well as the appearance

220 A. T. Rhys Williams, S. A. Winfield, J. N. Miller, *Analyst* **1983**, *108*, 1067-1071.

of a weaker negative Cotton effect band about 525 nm, both bands consistent with the formation of a ΛΛ-metallocylinder. Likewise, the enantiomeric peptide ligand **DDL** gave rise to opposite bands at 320 and 525 nm, as expected for the induction of the mirror image ΔΔ-helicate. Taken together, these data are consistent with a dynamic process in which the mixture of the peptide ligand and the Fe(II) ion evolves towards the more stable folded Fe(II)-metallocylinder(s) under thermodynamic control. This equilibrium is fundamental for the stereoselection and correct folding of the desired helicate structure, but at the same time introduces severe limitations for the application of such helicates, because these self-assembled helicates are intrinsically dynamic and could disassemble, exchange ligands, or modify their structure or association state in response to changes in their environment.[213]

To overcome these limitations, a more robust and kinetically inert interaction was desirable.[214a] With this aim in mind, we considered the possibility of dynamically assembling the metallocylinders in the presence of a labile Co(II) ions, which would then be locked by oxidation in the form of a kinetically inert Co(III) complexes. This strategy has been described for the immobilization of His-tagged proteins on solid supports and surfaces,[221] as well as for the synthesis of stable complexes,[127b,215c] but to our knowledge, never in the case of designed metallopeptides. Thus, a 600 μM solution of **LLD** (or **DDL**) was incubated with 3 mM concentration of $Co(ClO_4)_2$ in milli-Q water. After 15 min of equilibration, the mixture was treated with cerium ammonium nitrate as oxidant $(NH_4)_2Ce(NO_3)_6$ and analyzed by HPLC after 15 min.

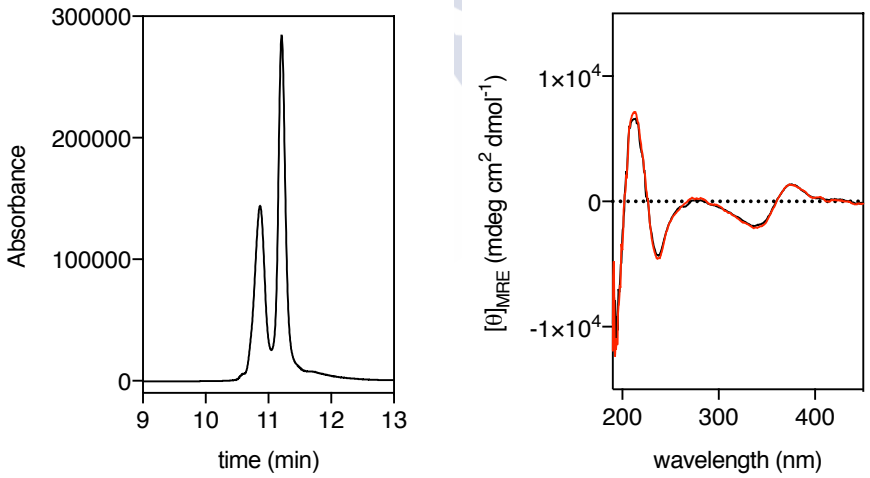


Figure 43. Left: chromatogram (0-40% of AcN + 0.1% TFA in 30 min) of the resulting mixture after the oxidation of the Co(II) helicate to Co(III). Each peak corresponds to a new formed diastereomer that in the case of the labile and dynamic Fe(II) and Co(II) was impossible to distinguish. Right: normalized CD spectra of the isolated isomer A (red line) and the isomer B (black line). Both of them shows a positive Cotton effect in the 5Bpy region (c.a. 351 nm), demonstrating the retention of the ΛΛ- (M) configuration around the metal center.

221 a) S. Auer, L. Azizi, F. Faschinger, V. Blazevic, T. Vesikari, H. J. Gruber, V. P. Hytönen, *Sens. Actuators B Chem.* **2017**, 243, 104-113; b) S. V. Wegner, F. C. Schenk, J. P. Spatz, *Chem. Eur. J.* **2016**, 22, 3156-3162; c) J. E. Hale, *Anal. Biochem.* **1995**, 231, 46-49.

Gratifyingly, while HPLC analysis of the labile **LLD** complexes with Fe(II) and Co(II) ions always showed a single peak, corresponding to the uncomplexed peptide ligand resulting from the disassembly of the metallocylinders in the acidic conditions of the reverse-phase HPLC analysis (0.1% TFA H_2O/CH_3CN), the HPLC trace of the oxidized Co(III) mixture showed two peaks with masses that were consistent with the formation of two isomeric $[Co(III)_2\mathbf{LLD}]^{6+}$ metallocylinders. The circular dichroism spectra of both isomeric complexes were superimposable, and in both cases the signs of the observed Cotton effects indicated that they correspond to the expected ΛΛ-metallocylinders (Figure 43).

Helicates are known to selectively interact with three-way DNA by inserting the metallocylinder into its hydrophobic center.[117,135a] In order to study the DNA-binding properties of these new helicates, we realized that we could monitor the quenching that both Fe(II) and Co(II) helicates display over fluorophores placed nearby.[179] Thus, for the Fe(II) peptide helicate, we prepared a 2 μM solution of a fluorescein-labelled three-way DNA junction (FAM-twDNA), in 1 mM PBS buffer, 10 mM NaCl, pH 6.5, and recorded its emission spectrum upon excitation at 490 nm after addition of successive aliquots of a solution containing the peptide ligand and $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ (200 μM and 1 mM, respectively). The progressive quenching of the emission intensity at 515 nm could be fitted to a 1:1 binding mode (twDNA/$[Fe(II)_2\mathbf{LLD}]^{4+}$) with an apparent K_D of ≈ 0.45 μM.[198] In the case of the Co(III) helicate, once again, we took advantage of the quenching properties of these metallopeptides. So, onto a 2 μM solution of a fluorescein-labelled three-way junction we add increasing amounts of our preformed helicate recording the emission of the fluorescein (λ_{exc} = 490 nm, λ_{em} = 515 nm) after each addition. In this case we decided to increase the ionic strength of the buffer (10 mM PBS, 100 mM NaCl, 10 mM $MgCl_2$ pH 6.5) in order to try to reduce the unspecific interactions between the helicate and the polyphosphate backbone of the DNA. The titration profile of the emission intensity at 515 nm could be fitted to a mixed 1:1/1:3 model characterized by two different binding constants. A lower-affinity binding site (presumably corresponding to the insertion of the $[Co(III)_2\mathbf{LLD}]^{6+}$ helicate in the hydrophobic cavity) with a K_D of ≈ 7.9 μM, and three equivalent higher-affinity sites (K_D of ≈ 4.9 μM), which, based on reported X-ray structures, could correspond to interactions of the complex with the hydrophobic surface of the blunt ends of the twDNA.

An additional fluorescence anisotropy titration of a 2 μM solution of the fluorescein-labelled **LLD** derivative (**FITC-LLD**) with increasing concentrations of twDNA, demonstrated the key role of metal coordination in twDNA recognition. In this case the descending and oscillating pattern observed in the titration showed the lack of interaction between the ligand and the twDNA.

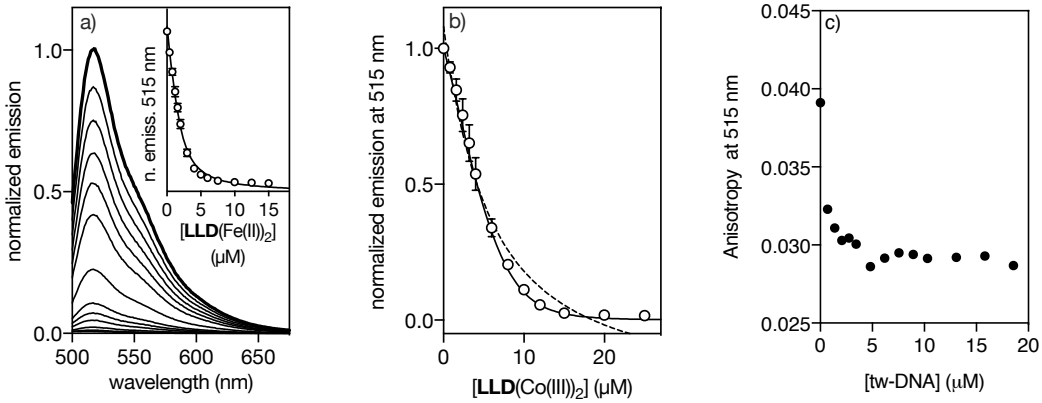


Figure 44. a) Normalized emission spectra of a titration of a 2 μM solution of the fluorescein-labeled twDNA, in 1 mM phosphate buffer, 10 mM NaCl, pH 6.5 (thick line) with increasing concentrations of the $[Fe(II)_2\mathbf{LLD}]^{4+}$ (thin lines); inset: corresponding normalized titration profile of the fluorescein quenching at 515 nm and best fit to a one to one binding model; b) Normalized emission of a 2 μM solution of the fluorescein-labelled twDNA, in 10 mM phosphate buffer, 100 mM NaCl, 10 mM $MgCl_2$, pH 6.5 with increasing concentrations of $[Co(III)_2\mathbf{LLD}]^{6+}$; the best fit to a simple one to one binding model is shown as a dashed line, and the fit to a model including non-specific binding is shown as a continuous line. Titration data are the average of three independent experiments; c) anisotropy control of a 2 μM solution of **FITC-LLD** with increasing concentrations of twDNA, where no interaction is appreciated in the decreasing and oscillating pattern of the data. twDNAs: 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: 5'–AAC ACA TGC AGG ACG GCG CTT–3' (only one strand shown).

To gain a deeper understanding of the DNA binding properties of both Fe(II) and Co(III) helicates, we carried out electrophoretic mobility shift assays (EMSA) in polyacrylamide gel under non-denaturing conditions[200] and using *SYBRGold* as the DNA stain.[201] Thus, in agreement with the fluorescence studies we found that incubation of twDNA with increasing concentrations of the **LLD** ligand and Fe(II) resulted in the appearance of new slow-migrating band, consistent with the formation of the twDNA/$[Fe(II)_2\mathbf{LLD}]^{4+}$ complex. As expected for a high-affinity interaction, a complete band displacement is observed already at the lowest concentrations tested, giving rise to a single new band with no smearing, even at high concentration of $[Fe(II)_2\mathbf{LLD}]^{4+}$. On the other hand, no new slow-migrating bands were observed upon incubation of $[Fe(II)_2\mathbf{LLD}]^{4+}$ with a double-stranded DNA (dsDNA), which confirms the low affinity of this complex for regular B-DNA. In agreement with the fluorescence titrations, the EMSA assays showed reduced affinity of $[Co(III)_2\mathbf{LLD}]^{6+}$ for the three-way DNA, so that we observe a complete disappearance of the twDNA band only at the highest concentrations of $[Co(III)_2\mathbf{LLD}]^{6+}$ tested (750 nM). Remarkably, these assays also confirmed the tendency of the highly charged Co(III) derivative to form non-specific complexes, as shown by the disappearance of the dsDNA band upon incubation with this helicate. This effect is consistent with the formation of multiple species in the gel (Figure 45).[202]

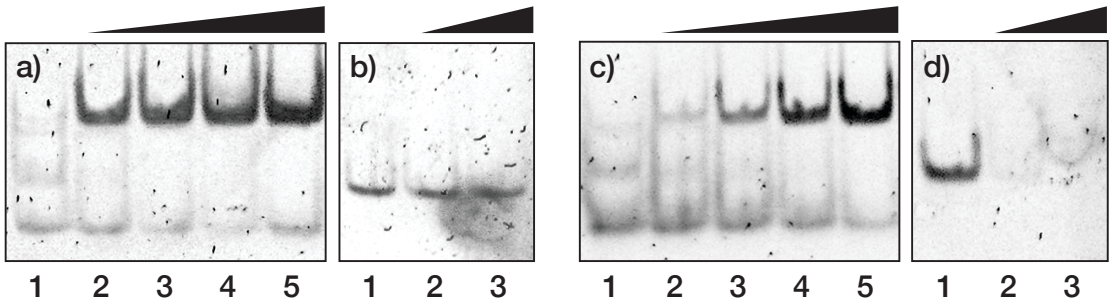


Figure 45. EMSA analysis of the **LLD** ligand binding to DNA. a) 200 nM twDNA in each lane. 0, 50, 100, 250, 500 nM of **LLD** and 5 eq of $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$. b) 75 nM dsDNA, with 0, 1000, and 1500 nM of **LLD** and 5 eq of $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$. c) and d) panels correspond to the same conditions as a) and b), but with the Co(III) helicate. twDNAs: 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: 5'–AAC ACA TGC AGG ACG GCG CTT–3' (only one strand shown).

Having demonstrated the binding to twDNA in vitro, we studied the application of these newly developed helicates to study these DNA structures in living cells. Preliminary experiments in which rhodamine- and fluorescein-labelled $[Fe(II)_2\mathbf{LLD}]^{4+}$ helicates (5 µM **Rho-LLD** or **FITC-LLD** and 25 µM $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$) were incubated overnight with HeLa cells showed that these complexes remain trapped in endosomes (Figure 46).

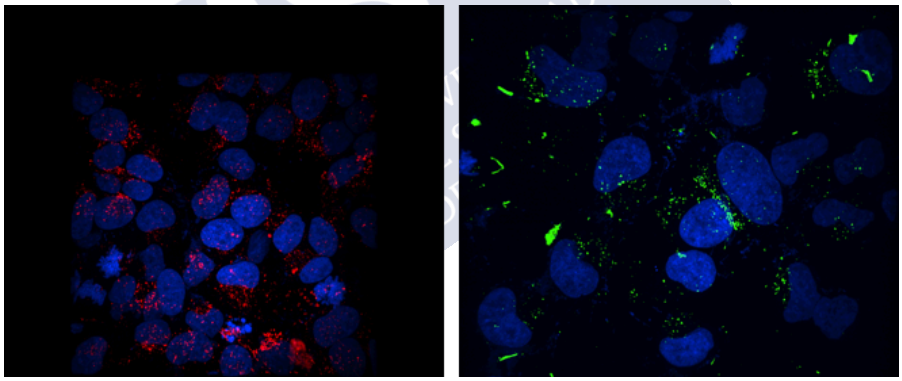


Figure 46. $[Fe(II)_2\mathbf{Rho\text{-}LLD}]^{4+}$ (left) and $[Fe(II)_2\mathbf{FITC\text{-}LLD}]^{4+}$ (right) helicates after overnight incubation with HeLa cells. The punctuated pattern is consistent with the endosome location of both peptides. In both pictures DAPI was used for the nucleus stain.

In order to try to know the final destination of our helicates inside the cell, we synthesized two new derivatives with two different N-terminal modifications that could allow these metallocylinders to escape from the endosome. First modification was to attach a Rho-O1Pen-CXC sequence at the N-terminal position of the peptide strand. It is known that thiols of the tripeptide can help to promote the internalization thanks to the interaction that establish with other SH groups of the cell membrane.[222] In this case, the X residue corresponds to a His, in order to contribute to the final peptide positive charge.

222 a) G. Gasparini, E.-K. Bang, J. Montenegro and S. Matile, *Chem. Commun.* **2015**, 51, 10389-10402; b) N. Chuard, G. Gasparini, D. Moreau, S. Lörcher, C. Palivan, W. Meier, N. Sakai and S. Matile, *Angew. Chem. Int. Ed.* **2017**, 56, 2947-2950.

O1Pen was used as a spacer between the Rho label and the peptide sequence. This new peptide helicate, [Fe(II)$_2$**Rho-LLD_CHC**]$^{4+}$, showed improved cell penetrating properties, being able to penetrate the cell in 30 minutes of incubation. Unfortunately, the final location of this peptide was again trapped in endosomes.

A second modification was carried out. In this case, we add to the original sequence an endosomal escape tag, the so called Aurein 1.2 (GLFDIIKKIAESF). This peptide is known for being able to disrupt the membrane of the endosomes—enhancing the endosomal escape—without disrupting mammalian cell membrane.[223] As in the previous case, after the Aurein 1.2 sequence, a O1Pen spacer was placed before the Rho fluorescent label. As in the previous case, Aurein promotes an upturn in the cell penetrating properties of the helicate, [Fe(II)$_2$**Rho-LLD_Au1.2**]$^{6+}$ but no signals of endosomal escape are appreciated after 4h of endosomal maturation process.

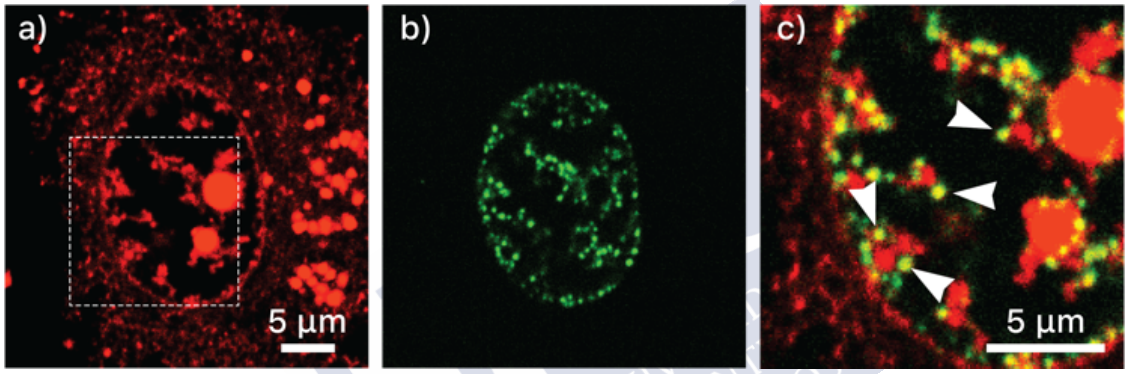


Figure 47. [Fe(II)$_2$**Rho-LLD**]$^{4+}$ selectively stains DNA replication sites. HeLa cells expressing protein GFP-PCNAL2 were incubated with 25 µg/ml Digitonin for 3 min, then 5 µM **Rho-LLD** and 25 µM $(NH_4)_2Fe(SO_4)_2 \cdot 6\,H_2O$ for 30 min. a) Red channel emission showing the distribution of the rhodamine-labeled helicate; b) green channel, corresponding to the emission of the GFP-PCNAL2 probe labeling the DNA replication foci; c) overlay of the green and red channels of the square region shown in a). Arrows highlight some of the foci where the staining of GFP-PCNAL2 and [Fe(II)$_2$**Rho-LLD**]$^{4+}$ overlap.

However, if the cells are pretreated with Digitonin,[224] then [**Fe(II)$_2$Rho-LLD**]$^{4+}$ was effectively internalized, and showed in the red emission channel as a punctuated pattern in the cytoplasm and the nucleus (Figure 47a). Remarkably, the nuclear distribution of the helicate matched the localization of the DNA replication sites labeled with the proliferating cell nuclear antigen (PCNA) fused to GFP,[225] which demonstrates that the helicate is capable of targeting the three-way junctions transiently formed during DNA

223 a) M. Li, Y. Tao, Y. Shu, J. R. LaRochelle, A. Steinauer, D. Thompson, A. Schepartz, Z.-Y. Chen, D. R. Liu, *J. Am. Chem. Soc.* **2015**, *137*, 14084-14093; b) D. I. Fernandez, A. P. Le Brun, T. C. Whitwell, M.-A. Sani, M. James, F. Separovic, *Phys. Chem. Chem. Phys.* **2012**, *14*, 15739-15751.

224 M. Nishikyaawa, S. Nojima, T. Akiyama, U. Sankawa, K. Inoue, *J. Biochem.* **1984**, *96*, 1231-1239.

225 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*, 29657-29665.

replication *in vivo* (Figure 47c).[226] The staining of the cell nucleolus is consistent with $[Fe(II)_2$**Rho-LLD**$]^{4+}$ binding to three-way RNA structures, which are also potential biological targets of helicates.[227]

Conclusions

We have demonstrated that peptide design rules can be effectively applied to encode both structural and chiral information in de novo designed peptide ligands. We report new oligocationic peptide sequences containing bipyridine ligands that fold into three-stranded metallocylinders and bind with high affinity and specificity to the central cavity in three-way DNA junction. Furthermore, we have applied an assembly-followed-by-fixing method to obtain inert Co(III) helicates, which display increased DNA binding. Finally, these Fe(II) systems have shown that are able to stain replication sites if they are able to reach the cell nucleus.

226 L. S. Shlyakhtenko, V. N. Potaman, R. R. Sinden, A. A. Gall, Y. L. Lyubchenko, *Nucleic Acids Res.* **2000**, *28*, 3472-3477.

227 Phongtongpasuk, S. Paulus, J. Schnabl, R. K. O. Sigel, B. Spingler, M. J. Hannon, E. Freisinger, *Angew. Chem. Int. Ed.* **2013**, *52*, 11513-11516.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

General Conclusions

1) We have demonstrated that the Fibritin foldon of the bacteriophage T4 (T4Ff) and the β-annulus peptide of the Sesbania Mosaic virus capsid can be used to enantioselectively obtain self-assembled helicates. In both cases the chirality of the final complex was the preferred for the DNA-binding (M, or left-handed).

2) We have obtained a second generation of Fe(II) metallopeptide cylinders by employing a new heterochiral β-turn promoting sequence and the βAla5Bpy chelating building block. The selected L-Arg–L-Pro–D-Arg sequence also improved the solubility of the systems in water.

3) We have obtained enantiomerically pure and kinetically inert metallopeptide cylinders by mild oxidation of the thermodynamically stable Co(II) helicates to the corresponding Co(III) complexes. This allowed us to discover the coexistence of a number of topological and conformational isomers, that in the case of the previous Fe(II) helicates were impossible to isolate and identify.

4) All these systems have shown promising DNA-binding properties versus three-way. DNA junction, with K_D values in the micromolar range (T4Ff and β-annulus) and even in the nanomolar for the **LLD** derivatives. In addition, electrophoretic mobility shift assays (EMSA) showed excellent selectivity of this helicate towards twDNA.

5) Internalization assays conducted with digitonin-permebealized HeLa cells showed that the $Fe(II)_2$**LLD** helicate accumulated in the cell nuclei. Furthermore, it accumulated preferentially in active replicating sites labelled with GFP-PCNAL2.

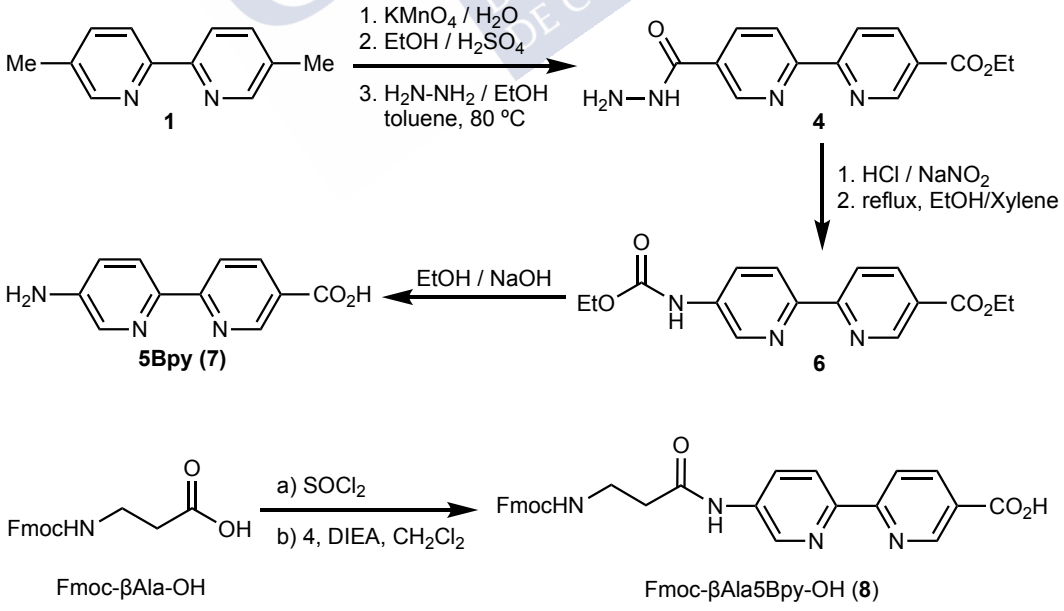
USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Resumen

Capítulo 1. Autoensamblaje estereoselectivo de un helicato trimérico

En este capítulo se demuestra que el péptido T4Ff modificado en su extremo *N*-terminal con dos unidades coordinantes de 2,2´-bipiridina dirige la formación de helicatos tricatenarios de Fe(II) con quiralidad *M* (levógira).

El primer paso en la obtención del ligando peptídico fue la síntesis de un residuo derivado del ligando 2,2´-bipiridina que se pueda integrar fácilmente en la secuencia peptídica mediante la estrategia estándar de síntesis de péptidos en fase sólida. Para ello, se ha puesto a punto una ruta sintética que permite la obtención de derivados asimétricos de la 2,2´-bipiridina con un grupo carboxilato y un grupo amino en las posiciones 5 y 5´. El núcleo asimétrico se ha modificado con un residuo de Fmoc-β-alanina para su utilización en síntesis de péptidos en fase sólida.



Esquema 1. Síntesis de la unidad de construcción coordinante Fmoc-βAla5Bpy-OH (**8**).

Una vez obtenido el residuo coordinante, se realizó la síntesis del péptido T4Ff, constituido por los últimos 27 amino ácidos del extremo C-terminal de la Fibritina, una

proteína trimérica del bacteriófago T4. La estructura secundaria de este péptido se divide en tres regiones claramente diferenciadas, una hélice 3_{10} en el extremo C-terminal, seguida de un giro β y, finalmente, una hélice de poliprolinas en el extremo N-terminal. Este dominio se eligió por su demostrada capacidad para trimerizar y servir de soporte de estructuras peptídicas triméricas.

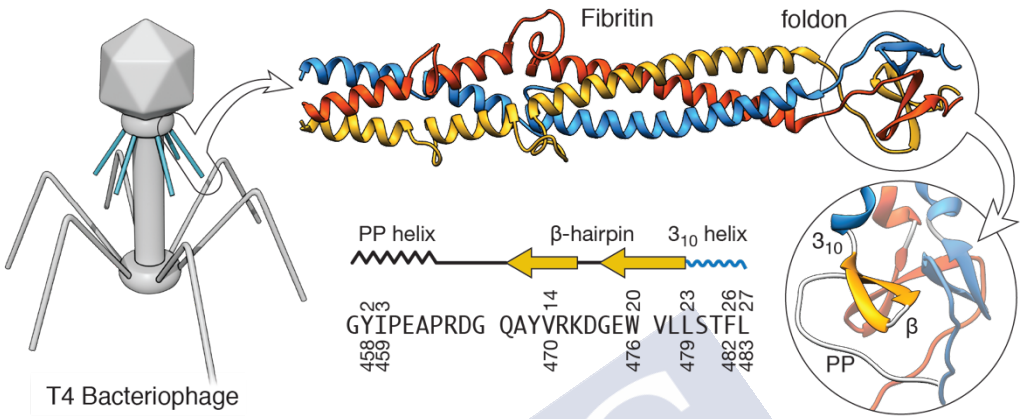


Figura 1. Estructura trimérica de la fibritina donde se aprecia en su extremo C-terminal el péptido T4FF o Foldon y las distintas partes de su estructura secundaria.

El estudio de la estructura secundaria de péptido T4Ff (PDB IDs 4NCU or 1RFO) reveló que la geometría del extremo N-terminal permitía acomodar las unidades coordinantes de 2,2′-bipiridina causando una distorsión mínima de la estructura peptídica tras la coordinación metálica. Además, la presencia de la hélice de poliprolinas en ese extremo podría inducir una quiralidad levógira sobre el helicato final, de manera que se decidió modificar el péptido natural con dos unidades coordinantes de βAla5Bpy en su extremo N-terminal.

Una vez sintetizado el péptido final **(βAla5Bpy)$_2$T4Ff** se purificó mediante HPLC en fase reversa y se llevaron a cabo estudios de coordinación. Para ello se aprovechó la desactivación de la emisión fluorescente que ejercen los iones de Fe(II) sobre el fluoróforo de 2,2′-bipirdinina como resultado de su coordinación. Los resultados obtenidos se ajustaron a un modelo de unión 1:2 con unas constantes de disociación de 5.5 µM para la entrada del primer ión metálico y 6.6 µM para el segundo. Además, se llevaron a cabo estudios de dicroísmo circular para confirmar la quiralidad del helicato final. Los resultados obtenidos son consistentes con la formación de un complejo M o levógiro, como se puede apreciar en el efecto Cotton positivo que presenta la banda de la 2,2′-bipiridina centrada a 320 nm.

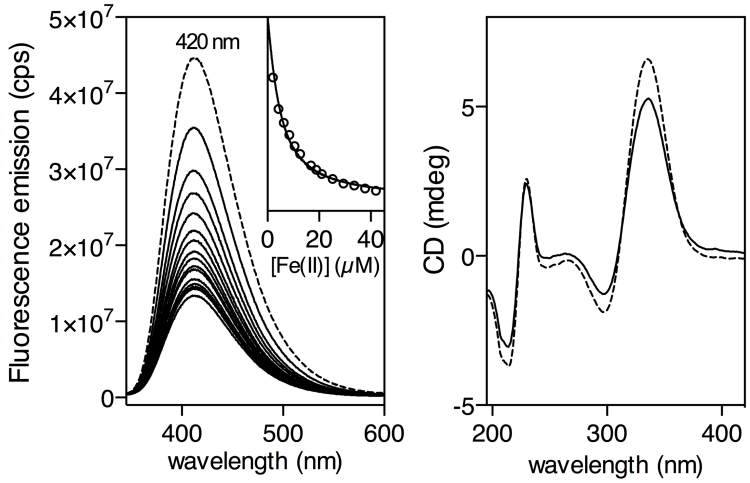


Figura 2. Izquierda: valoración de fluorescencia de una disolución 3 μM (9 μM de monómero) de **[(βAla5Bpy)$_2$-T4Ff]**$_3$ con concentraciones crecientes de Fe(II). El recuadro muestra la emisión a 420 nm tras la excitación a 310 nm con concentraciones crecientes de Fe(II), y el mejor ajuste para un modo de unión 1: 2. Los experimentos se realizaron por triplicado. Derecha: CD de una disolución 6 μM (18 μM de monómero) de **[(βAla5Bpy)$_2$-T4Ff]**$_3$ (línea discontinua) y en presencia de Fe(II) 90 μM (línea continua). Todos los experimentos se realizaron en tampón fosfato 1 mM, pH 6,5, NaCl 10 mM a 20 °C.

Una vez caracterizado el helicato final de Fe(II), el siguiente paso fue el estudio de sus propiedades de unión al ADN. Los estudios de anisotropía de fluorescencia muestran claramente una preferencia por el ADN de tres vías (twDNA) frente al ADN-B de doble cadena (dsDNA), mostrando una constante de disociación de ~2.1 μM.

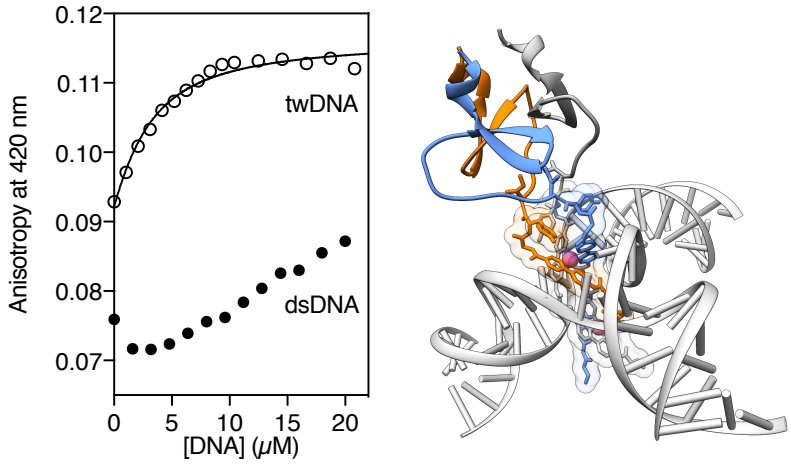


Figura 3. Izquierda: Valoración de anisotropía de ΛΛ-[Fe(II)$_2$[**(βAla5Bpy)$_2$-T4Ff**]$_3$]$^{4+}$en tampón de fosfato 1 mM, NaCl 10 mM, pH 6.5 con concentraciones crecientes de twDNA. Se muestra el ajuste a un modelo de unión 1:1. Derecha: Modelo esperado de la interacción entre el complejo de ΛΛ-[Fe(II)$_2$[**(βAla5Bpy)$_2$-T4Ff**]$_3$]$^{4+}$ y el twDNA.

Finalmente, también se analizaron las propiedades de unión del ADN mediante ensayos de electroforesis en gel de poliacrilamida. Estos ensayos revelaron la aparición de una nueva banda retardada que indica la interacción entre el helicato obtenido ΛΛ-$[Fe(II)_2[\mathbf{(\beta Ala5Bpy)_2\text{-}T4Ff}]_3]^{4+}$y el twDNA. Además, en las calles control con DNA de doble hebra no se aprecia la aparición de ninguna banda retardada, confirmando de esta manera la selectividad del helicato frente al twDNA.

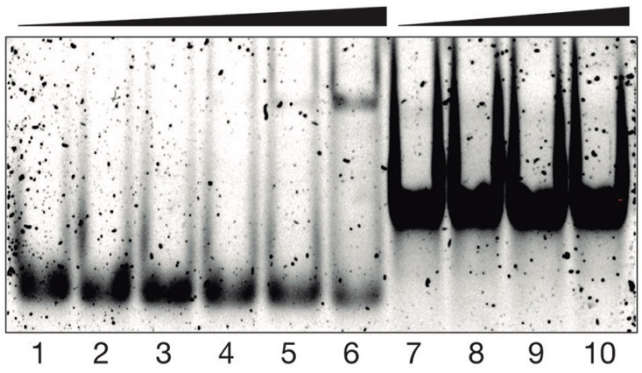


Figura 4. Resultados de los estudios de unión al ADN para ΛΛ-$[Fe(II)_2[\mathbf{(\beta Ala5Bpy)_2\text{-}T4Ff}]_3]^{4+}$. Calles 1–6, twDNA 200 nM con 0, 150, 250, 500, 1,000 y 2,000nM de $[\mathbf{(\beta Ala5Bpy)_2\text{-}T4Ff}]_3$ y 14 eq. de $(NH4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ en cada calle; calles 7-10, dsDNA 200 nM con 0, 500, 1,000 y 2,000 nM de $[\mathbf{(\beta Ala5Bpy)_2\text{-}T4Ff}]_3$ y 14 eq. de $(NH4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ en cada calle.

Capítulo 2. Autoensamblaje estereoselectivo de un helicato con propiedades de unión al ADN dirigido por el péptido β-Annulus.

En este capítulo se demuestra que el péptido β-Annulus derivado de la cápside del virus del mosaico de Sesbania selectivamente modificado y equipado en la cadena lateral de uno de sus amino ácidos con dos unidades coordinantes derivadas de la 2,2´-bipiridina, es capaz de promover el autoensamblaje estereoselectivo de helicatos tricatenarios de Fe(II) que presentan propiedades de unión al ADN. El β-Annulus es un péptido corto, compuesto por 12 amino ácidos (GISMAPS^{54}AQGAM), que se ensambla formando una estructura trimérica, soportada a través de una red de interacciones de puentes de hidrógeno entre cada uno de los monómeros que lo conforman. En la secuencia peptídica se puede apreciar un giro de 120°, fundamental para el mantenimiento de la estructura, que está promovido por la presencia de la Pro^{53} en la secuencia peptídica.

El estudio de la estructura del β-annulus mostró que el residuo de Ser^{54} podría servir como punto de anclaje para la introducción ortogonal de dos residuos coordinantes derivados de la 2,2´-bipiridina. Por tanto, se decidió sustituir este aminoácido por el ácido L-2,4-diaminobutírico ortogonalmente protegido en su cadena lateral con un grupo Alloc. Además de este cambio se realizaron también dos sustituciones más por motivos sintéticos. La secuencia final Ac-G^{48}IS-nL^{51}-AP-Dab(Alloc)^{54}-AQGAK^{59}-NH_2. fue sintetizada siguiendo la metodología clásica de síntesis de péptidos en fase sólida, a continuación se procedió a desproteger ortogonalmente la cadena lateral del residueo de Dab(Alloc) para permitir el acoplamiento de las dos subunidades de βAla5Bpy. En fase

sólida. Posteriormente el péptido se desprotegió/liberó de la resina y se purificó mediante HPLC en fase reversa.

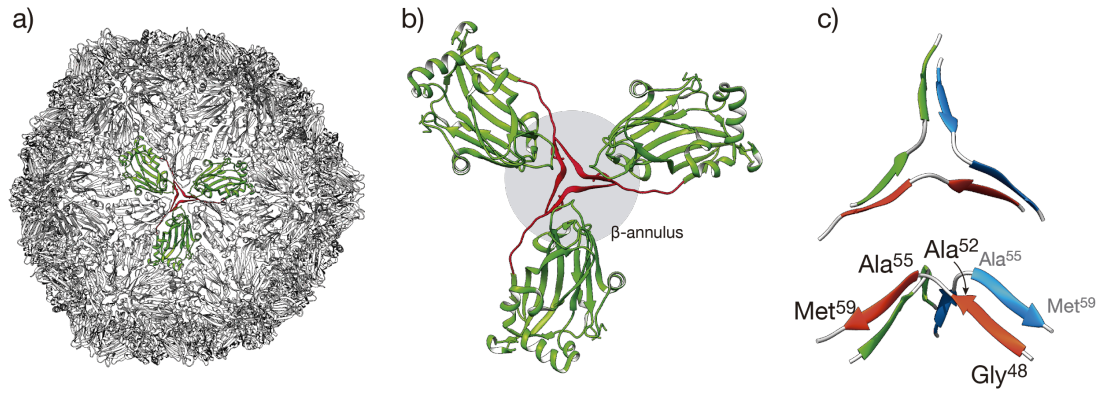


Figura 5. a) Estructura de la cápside de SeMV (PDB 1X35) donde el anillo β se resalta en rojo y la subunidad C en verde. b) Estructura previamente resaltada en la cápsida de SeMV donde el anillo β se muestra en rojo y el trímero resultante con la subunidad C resaltada en verde. c) Estructura del anillo β aislado del resto de la estructura que muestra los residuos clave involucrados en la trimerización de la estructura.

Una vez obtenido el péptido final, el siguiente paso fue la caracterización de sus propiedades coordinativas y de la quiralidad del helicato resultante. Para ello se empleó de nuevo la desactivación de la emisión fluorescente de la 2,2´-bipiridina que ejercen los iones de Fe(II) tras su coordinación.

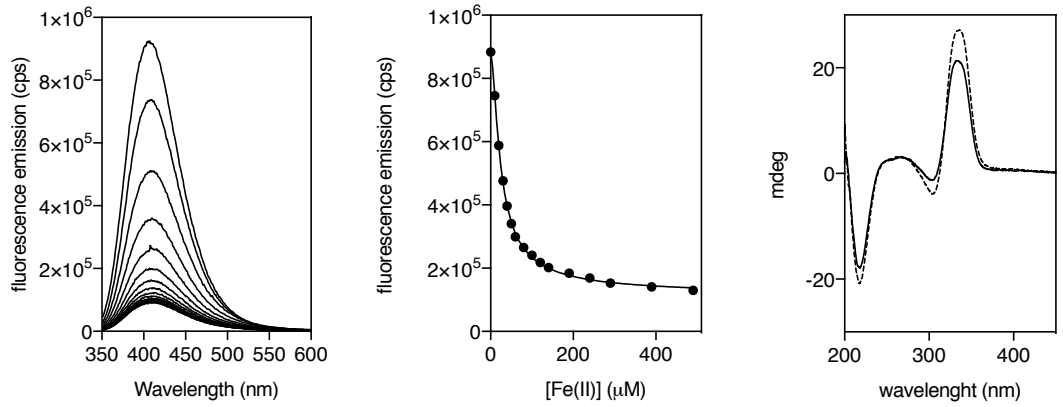


Figura 6. Izquierda: espectros de emisión de una disolución de péptido **β-annX(5Bpy)$_2$** 20 μM (tampón HEPES 10 mM, NaCl 100 mM, pH 6.5) con concentraciones crecientes de $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$. Centro: Perfil de titulación de la longitud de onda de emisión máxima a 410 nm con concentraciones crecientes de Fe(II); centro: se muestra el mejor ajuste según el modelo 1: 2 en DynaFit. Los puntos de datos experimentales corresponden al promedio de tres valoraciones independientes; derecha: CD de una disolución de 7 μM (monómero 21 μM) de ΛΛ-$[Fe(II)_2$(**β-annX(5Bpy)$_2$**$)_3]^{4+}$ (línea discontinua) y en presencia de Fe(II) 200 μM (línea continua).Todos los experimentos se realizaron en tampón Hepes 10 mM, pH 6,5, NaCl 100 mM a 20 °C. Los datos representados corresponden al promedio de tres acumulaciones.

Los datos resultantes se ajustaron a un modelo de unión 1:2 con una constantes de disociación de 14 y 15 µM para el primer y el segundo ion de Fe(II), respectivamente. Además, los estudios de dicroísmo circular muestran claramente la presencia de un helicato con quiralidad M, consistente con la presencia de una banda con efecto Cotton positivo centrada a 320 nm.

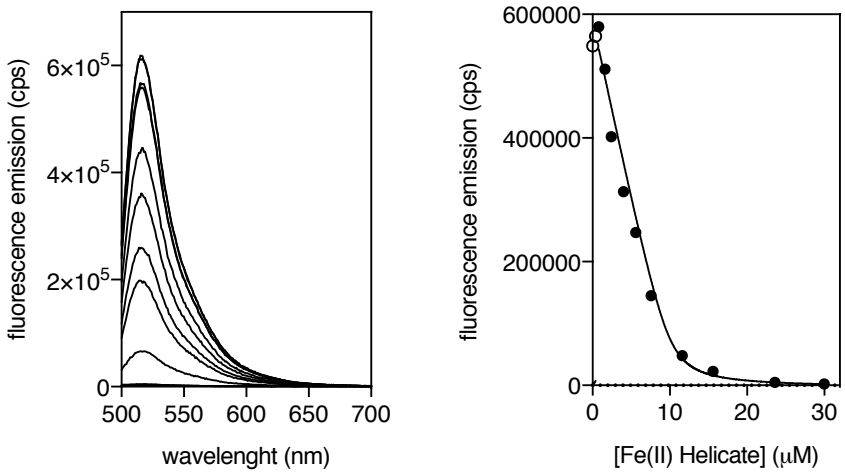


Figura 7. Izquierda: espectros de fluorescencia de una disolución 2 µM de FAM-twDNA (tampón HEPES 10 mM, NaCl 100 mM, pH 6.5) en presencia de concentraciones crecientes de ΛΛ-$[Fe(II)_2(\textbf{β-annX(5Bpy)}_2)_3]^{4+}$ helicado. Derecha: perfil de la valoración y mejor ajuste a un modo de unión 1:1 para twDNA (línea continua, círculos negros); los círculos abiertos representan valores que se excluyen el ajuste del modo de enlace 1:1.

Finalmente, se caracterizó el reconocimiento del twDNA midiendo la desactivación de la emisión que el helicato ejerce sobre los fluoróforos próximos. De esta, manera registró la emisión de un twDNA marcado con fluoresceína (FAM-twDNA) en concentraciones crecientes del helicato. Los datos obtenidos pudieron ajustarse a un modelo de unión 1:1 con una constante de disociación de 2.8 µM.

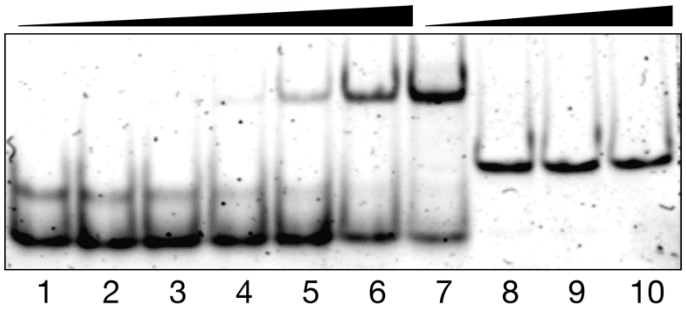


Figura 8. Resultados de los estudios de unión al ADN para ΛΛ-$[Fe(II)_2(\textbf{β-annX(5Bpy)}_2)_3]^{4+}$. Calles 1–6, twDNA 200 nM con 0, 150, 250, 500, 1,000 y 2,000nM de $[[\textbf{β-annX(Bpy)}_2]_3$ y 20 eq. de $(NH4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ en cada calle; calles 7-10, dsADN 200 nM con 0, 500, 1,000 y 2,000 nM de $[[\textbf{β-annX(Bpy)}_2]_3$ y 20 eq. de $(NH4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ en cada calle.

Para confirmar los resultados obtenidos por fluorescencia directa se realizaron también ensayos de unión a ADN por electroforesis en gel de poliacrilamida. Estos ensayos

revelaron la aparición de una nueva banda retardada como resultado de la interacción entre el helicato y el twDNA. Una vez más, en las calles control con DNA de doble hebra no se aprecia la aparición de ninguna banda retardada, confirmando de esta manera la selectividad del helicato frente al twDNA.

Capítulo 3. Estereoselección dinámica de helicatos peptídicos con propiedades de unión al ADN.

En este capítulo se demuestra la obtención de helicatos peptídicos enantiopuros y cinéticamente inertes a través del plegamiento dirigido por iones metálicos de péptidos lineales que integran en su secuencia unidades coordinantes de 2,2´-bipiridina conectadas entre sí por secuencias quirales promotoras de giros β que además dotan a las hebras peptídicas de una gran solubilidad en medio acuoso.

Se ha sintetizado la Fmoc-βAla5Bpy como residuo coordinante, integrándolo en una secuencia peptídica que reemplaza los giros GPG utilizdos previamente en nuestro grupo por una mezcla de residuos de Arg con quiralidad alterna que promueven la formación de giros β tipo II o II´. Así pues, las secuencias finalmente integradas en los nuevos sistemas están compuestos por L-Arg–L-Pro–D-Arg y su imagen especular D-Arg–D-Pro–L-Arg, que aportan solubilidad y ventajas a la hora del reconocimiento de ADN. Los ligandos precursores de helicatos, **DDL**: (H-(βAla5Bpy)$_2$-rpR-(βAla5Bpy)$_2$-rpR-(βAla5Bpy)$_2$-NH_2 y **LLD**: (H-(βAla5Bpy)$_2$-RPr-(βAla5Bpy)$_2$-RPr-(βAla5Bpy)$_2$-NH_2 se sintetizaron de acuerdo a la metodología clásica de síntesis de péptidos en fase sólida.

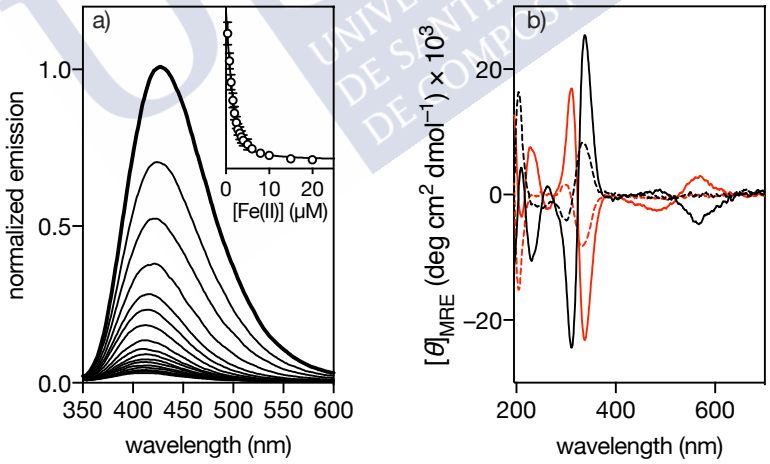


Figura 9. Valoración del ligando **DDL**. a) Espectro de emisión normalizado de una disolución 2 μM de **DDL** en tampón fosfato 1 mM NaCl 10 mM, pH 6.5 (línea gruesa) y espectros de la misma disolución en presencia de concentraciones crecientes de iones Fe(II) (líneas finas) ; recuadro que muestra el perfil de tres experimentos independientes y se ajusta mejor a un modelo de unión 1:2; b) Espectros de dicroísmo circular de disoluciones 5 μM de **LLD** (línea discontinua negra) y **DDL** (línea discontinua roja) en tampón fosfato NaCl 10 mM, pH 6,5 y en presencia de iones Fe(II) 25 μM (mismos colores, líneas continuas).

Una vez aislados los péptidos, se procedió a la caracterización de sus propiedades coordinativas en presencia de iones metálicos (Fe(II) o Co(II)) de forma análoga a los capítulos anteriores, así como también a la caracterización de la quiralidad de los helicatos

resultantes. El ajuste a un modelo 1:2 resultó en constantes de disociación muy similares tanto para Fe(II) como Co(II), en ambos casos en torno a 0.5 y 0.4 µM. Los estudios de dicroísmo circular confirmaron, además, la obtención de helicatos quirales.

Una vez caracterizadas las propiedades de unión a metales así como la quiralidad de sus helicatos se obtuvieron los helicatos cinéticamente inertes derivados de Co(III) con quiralidad M, siendo esta la preferida para la interacción con el ADN. Para ello, el helicato termodinámicamente más estable de Co(II) obtenido por la mezcla del ligando ***LLD*** con Co(II) se oxidó utilizando $Ce(NO_3)_6(NH_4)_2$, observándose por HPLC la aparición de dos picos correspondientes a dos nuevos isómeros previamente imposibles de aislar debido a la labilidad de los complejos de Fe(II) y Co(II). Ambos isómeros se aislaron y se caracterizaron por dicroísmo circular de nuevo, confirmando la retención de la configuración M presente en sus derivados de Co(II).

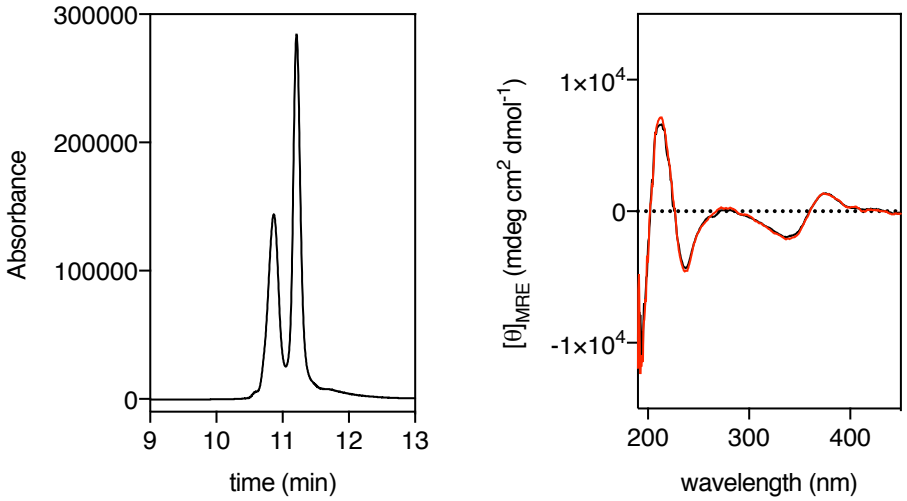


Figura 11. Izquierda: cromatograma (0-40% de AcN + 0.1% TFA en 30 min) de la mezcla resultante después de la oxidación del helicato de Co(II) a Co(III). Cada pico corresponde a un nuevo diastereómero formado. Derecha: espectros de CD normalizados de los dos isómeros aislados, que en ambos casos muestran un efecto Cotton positivo en la región de la bipiridina (aproximadamente 351 nm), que demuestra la retención de la configuración ΛΛ- (o *M*) alrededor del centro metálico.

Posteriormente se procedió a la caracterización de las propiedades de unión al twDNA de los helicatos de Fe(II) y de Co(III) siguiendo los métodos descritos previamente. En este caso el perfil obtenido para el caso del Fe(II) se pudo ajustar a un modelo de unión 1:1 con una constante de disociación de 450 nM, mientras que en el caso del Co(III) el perfil obtenido se ajusta a un modelo más complejo que supone la formación de complejos 1:1 junto con interacciones no específicas más lábiles con constantes de unión de 4.5 µM y 7.9 µM.

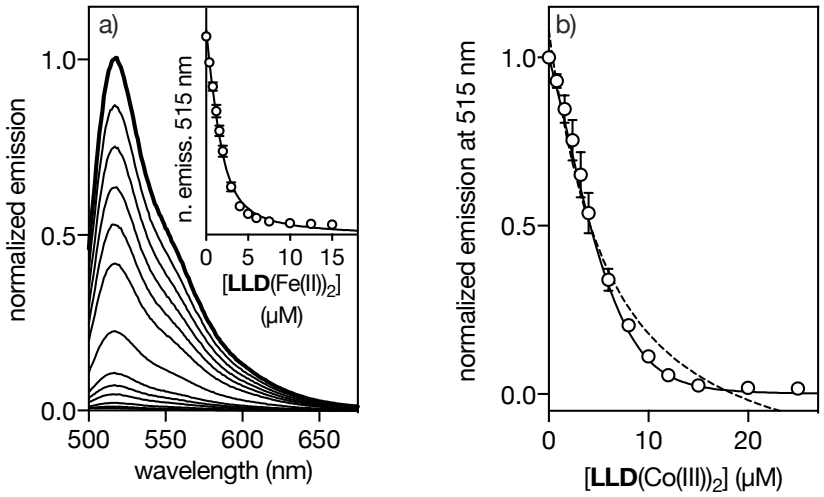


Figura11. a) Espectros de emisión normalizados de una valoración de una disolución 2 μM del ADN de unión de tres vías marcado con fluoresceína, FAM-twDNA, en tampón fosfato 1 mM, NaCl 10 mM, pH 6.5 (línea gruesa) con concentraciones crecientes de $[Fe(II)_2\mathbf{DDL}]^{4+}$ (líneas finas); recuadro: perfil de la valoración normalizado correspondiente a la desactivación de la emisión de la fluoresceína a 515 nm ajustado a un modelo de unión 1:1; b) Emisión normalizada de una disolución 2 μM del ADN de tres vías marcado con fluoresceína, FAM-twDNA, en tampón fosfato 10 mM, NaCl 100 mM, $MgCl_2$ 10 mM, pH 6.5 con concentraciones crecientes de $[Co(III)_2\mathbf{LLD}]^{6+}$; el mejor ajuste para un modelo de unión simple 1:1 se muestra como una línea discontinua, y el ajuste para un modelo que incluye interacción no específica se muestra como una línea continua. Los datos de valoración son el promedio de tres experimentos independientes.

Como en los casos anteriores también se estudiaron las propiedades de unión al ADN de los helicatos por electroforesis. En este caso ambos mostraron selectividad por twDNA, que se aprecia en la aparición de una nueva banda retardada en presencia del twDNA, pero no al incubarse con el ADN de doble cadena. Además, los resultados confirman la mayor afinidad de los complejos de Fe(II) en comparación con los de Co(III), apreciables en la rápida saturación apreciada en el caso del Fe(II).

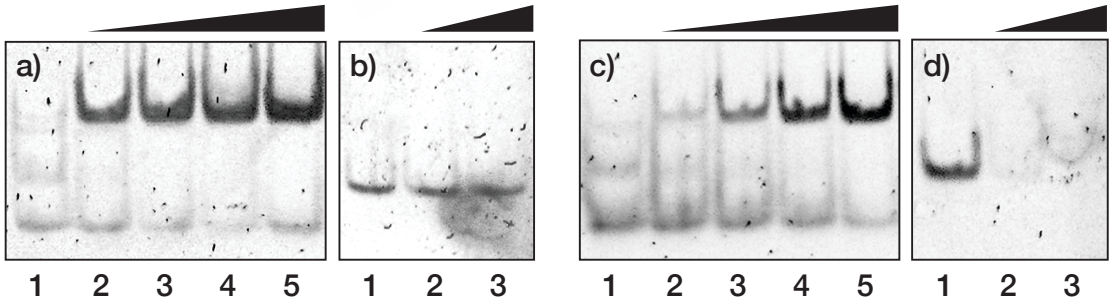


Figura 13. Análisis electroforético de la unión del ligando **LLD** al ADN. a) 200 nM de twDNA en cada calle. 0, 50, 100, 250, 500 nM de **LLD** y 5 eq de $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$. b) dsDNA 75 nM, con 0, 1000 y 1500 nM de **LLD** y 5 eq de $((NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$. c) y d) corresponden a las mismas condiciones que a y b, pero con el helicato de Co(III).

Finalmente se estudiaron las propiedades de internalización celular del compuesto **LLD**. Para ello se sintetizaron los análogos fluorescentes del isómero **LLD** añadiendo en el extremo N-terminal fluoresceína o rodamina para observar la influencia del fluoróforo en

su internalización. Desafortunadamente en todos los estudios realizados los helicatos resultantes permanecieron atrapados en endosomas. También se realizó un ensayo en el que se permeabilizaba previamente la membrana citoplasmática con digitonina, pudiéndose observar en ese caso la internalización y localización nuclear del helicato de rodamina. Además, utilizando un marcador de zonas de replicación, se pudo comprobar que estos helicatos marcan zonas de formación transitoria de twDNA en las factorías de replicación celular.

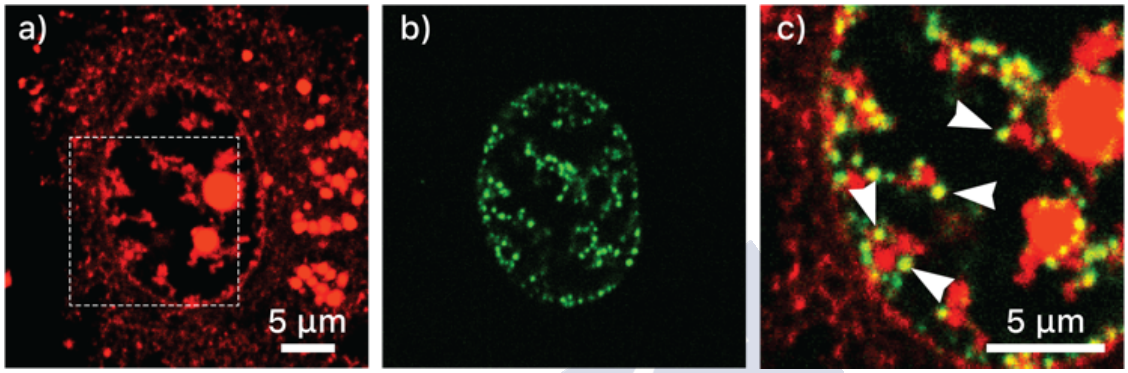


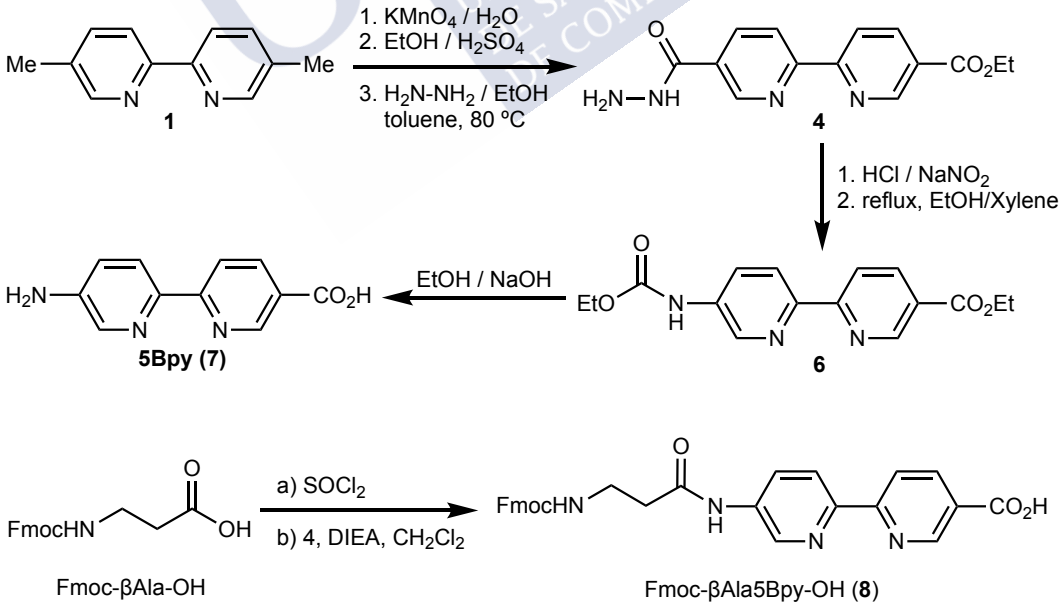
Figura 13. $[Fe(II)_2$**Rho-LLD**$]^{4+}$marca selectivamente los sitios de replicación de ADN. Las células HeLa que expresan la proteína GFP-PCNAL2 se incubaron con digitonina 25 µg / ml durante 3 minutos, luego con **Rho-LLD** 5 µM y 25 µM $(NH4)_2Fe(SO_4)_2$ • 6 H_2O durante 30 minutos. a) emisión de canal rojo que muestra la distribución del helicato marcado con rodamina; b) canal verde, correspondiente a la emisión de la sonda GFP-PCNAL2 que marca los focos de replicación de ADN; c) Superposición de los canales verde y rojo de la región cuadrada que se muestra en a). Las flechas resaltan algunos de los focos donde se superponen las manchas de GFP-PCNAL2 y $[Fe(II)_2$**Rho-LLD**$]^{4+}$

Summary

Chapter 1. Stereoselective self-assembly of a trimeric helicate

In this chapter we demonstrated the ability of the T4Ff, modified in its N-terminus with two chelating units of a 2,2´-bipyridine, to self-assemble into three-stranded helicates with M (left-handed) chirality .

The first step was the synthesis of a residue derived from 2,2´-bipyridine that can be easily integrated into a standard Solid-Phase Peptide Synthesis protocols. For this, we developed a synthetic route for obtaining asymmetric derivatives of 2,2´-bipyridine containing a carboxylic acid and an amino group in positions 5 and 5´. This asymmetric unit was derivatized with an Fmoc-protected β-alanine for its use in Solid-Phase Peptide Synthesis.



Scheme 1. Synthesis of the coordinating building block Fmoc-βAla5Bpy-OH (**8**).

Having at hand the coordinating residue, we synthesized the T4Ff peptide, containing the last 27 residues at the C-terminal end of Fibritin, a trimeric protein of the T4 phage. The secondary structure of this peptide is divided into three clearly differentiated regions, a helix 3_{10} at the C-terminal end, followed by a β-turn and finally a polyproline helix at the

N-terminal end. This domain was selected because of its ability to trimerize and support trimeric peptide structures.

Inspection of the structure of T4Ff (PDB IDs 4NCU or 1RFO), showed that the N-terminal Gly residues are relatively close to each other and could accommodate the chelating 2,2'-bipyridine units without noticeable distortion of the T4Ff scaffold upon metal coordination. In addition, the presence of the polyproline helix at this end could induce a left-handed chirality on the final helicate. Therefore, we decided to modify the core T4Ff structure at its N-terminus with two βAla5Bpy coordinating units.

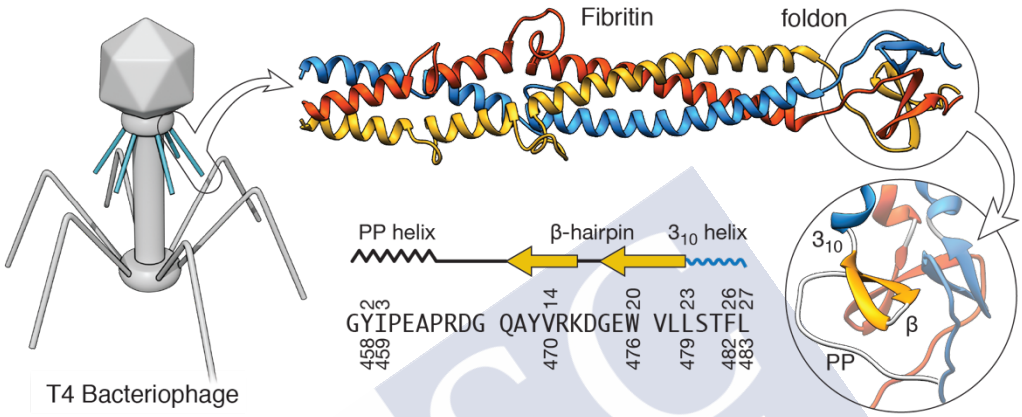


Figure 1. Fibritin is a homotrimer protein found in the Bacteriophage T4 virus collar that stimulate assembly of the long tail fibers (in blue), and their subsequent attachment to the tail baseplate during virus assembly. The structure of the C-terminal region of Fibritin showing part of the trimeric α-helical coiled coil, and the autonomous foldon unit (UniProtKB, P10104; PDB, 1AA0). Detail of the foldon domain showing the main structural elements: C-terminal 3_{10} helix, a β-hairpin, and the N-terminal polyproline (PP) helix (PDB, 1RFO). Schematic representation of the secondary structure elements and amino acid sequence of the foldon domain. Key resides in the foldon domain (1-27) are highlighted. Sequence numbers below correspond to the numbering in the X-ray structure of the T4 Fibritin in PDB 1AA0.

Once the final peptide **(βAla5Bpy)$_2$T4Ff** was synthesized, and purified by reverse phase HPLC, we carried out the metal coordination studies by monitoring the quenching of the fluorescence emission of the bipyridines exerted by Fe(II) ions upon coordination. The resulting profile could be fitted to a 1:2 binding mode with dissociation constants of 5.5 and 6.6 μM for the entry of the first and the second metal ions, respectively. In addition, circular dichroism studies were consistent with the formation of a M (or left-handed) complex, as can be seen in the positive Cotton effect of the bipyridine band.

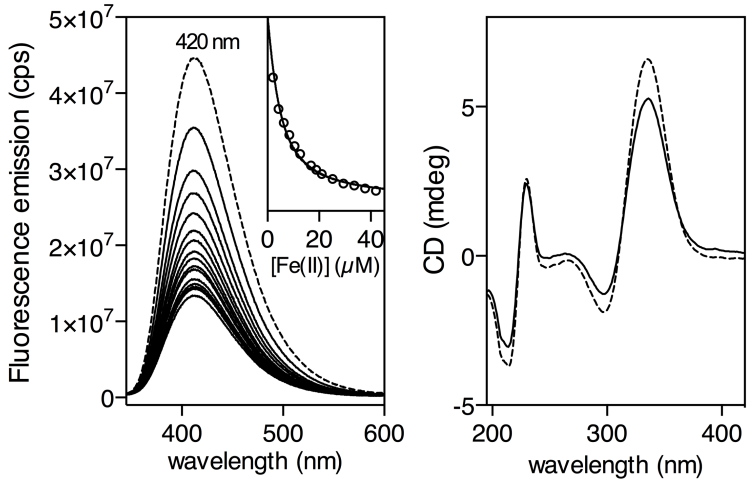


Figure 2. Left: fluorescence titration of a 3 μM (9 μM monomer) solution of **[(βAla5Bpy)$_2$-T4Ff]**$_3$ with increasing concentrations of Fe(II). Inset shows emission at 420 nm upon excitation at 310 nm with increasing concentrations of Fe(II), and the best fit to a 1:2 binding mode. Experiments were made in triplicate. Right: CD of a 6 μM solution (18 μM monomer) of **[(βAla5Bpy)$_2$-T4Ff]**$_3$ (dashed line) and in the presence of 90 μM Fe(II) (solid line). All experiments were made in 1 mM phosphate buffer, pH 6.5, 10 mM NaCl at 20 °C

Once the final Fe(II) helicate was characterized, the next step was the study of its DNA binding properties by fluorescence anisotropy studies, which clearly show a preference for twDNA over the classic double-stranded DNA (dsDNA), showing a K_D of ~ 2.1 μM.

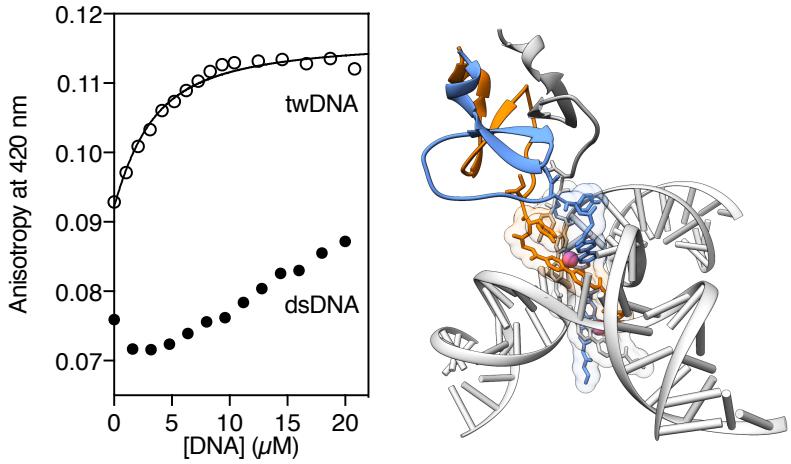


Figure 3. Left: Anisotropy titration of ΛΛ-[Fe(II)$_2$[(**βAla5Bpy)$_2$-T4Ff**]$_3$]$^{4+}$ in 1 mM phosphate buffer, 10 mM NaCl with increasing concentrations of twDNA. The best fit to a 1:1 binding mode is shown (curve fitting was performed using DynaFit). Right: Model of the interaction between the ΛΛ-[Fe(II)$_2$[(**βAla5Bpy)$_2$-T4Ff**]$_3$]$^{4+}$ and the three-way junction based on the reported pdb structure of a metal cylinder inserted in a three-way DNA.

Finally, we also analyzed the DNA binding properties by Electrophoretic Mobility Shift Assays (EMSA). These assays revealed the appearance of a new slow-migrating band consistent with the interaction between the helicate and the twDNA. In addition, no slow

migrating bands could be observed in the control lanes with dsDNA, confirming the selectivity of the helicate for twDNA.

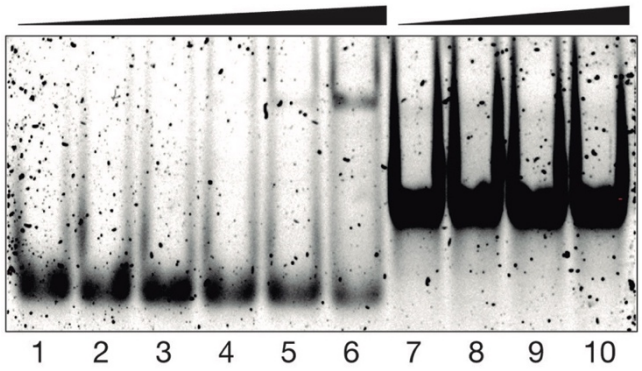


Figure 4. EMSA DNA binding studies results for ΛΛ-$[Fe(II)_2$**$[(\beta Ala5Bpy)_2$-T4Ff**$]_3]^{4+}$ helicate. Lanes 1–6, 200 nM twDNA with 0, 150, 250, 500, 1,000, and 2,000nM of **$[(\beta Ala5Bpy)_2$-T4Ff**$]_3$ and 14 eq. of $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ in each lane; lanes 7–10, 200 nM dsDNA with 0, 500, 1,000, and 2,000 nM of **$[(\beta Ala5Bpy)_2$-T4Ff**$]_3$ and 14 eq. of $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ in each lane.

Chapter 2. Stereoselective self-assembly of a DNA-binding helicate directed by the β-annulus motif

In this chapter we show that the β-Annulus peptide, derived from the capsid of the Sesbania Mosaic Virus, selectively modified and equipped in the side chain of one of its amino acids with two coordinating units derived from 2,2'-bipyridine is capable of promoting the stereoselective self-assembly of a Fe(II) three-stranded helicate that exhibits good DNA binding properties. The β-Annulus is a short peptide composed of 12 amino acids (GISMAPS^{54}AQGAM) that self-assembles into a trimeric structure that is held together by a network of hydrogen bonds between each of the monomer that comprise it. The presence of a 120° turn promoted by the Pro^{53} residue in the peptide sequence is fundamental for the stability of the trimeric structure.

Inspection of the structure of β-annulus showed that the three residues of Ser^{54} could serve as an anchor point for the introduction of two coordinating residues derived from 2,2´-bipyridine. Therefore, this Ser residue was replaced by a side chain Alloc-protected L-2,4-diaminobutyric acid. In addition to this change, two more substitutions were also made for synthetic reasons. The final sequence Ac-G^{48}IS-nL^{51}-AP-Dab(Alloc)^{54}-AQGAK^{59}-NH_2 was synthesized following standard Fmoc-solid-phase peptide synthesis methods, followed by the removal of the Alloc group to liberate the side chain of the Dab residue, where two subunits of βAla5Bpy were coupled in the solid phase. Finally, the peptide was purified by reverse phase HPLC.

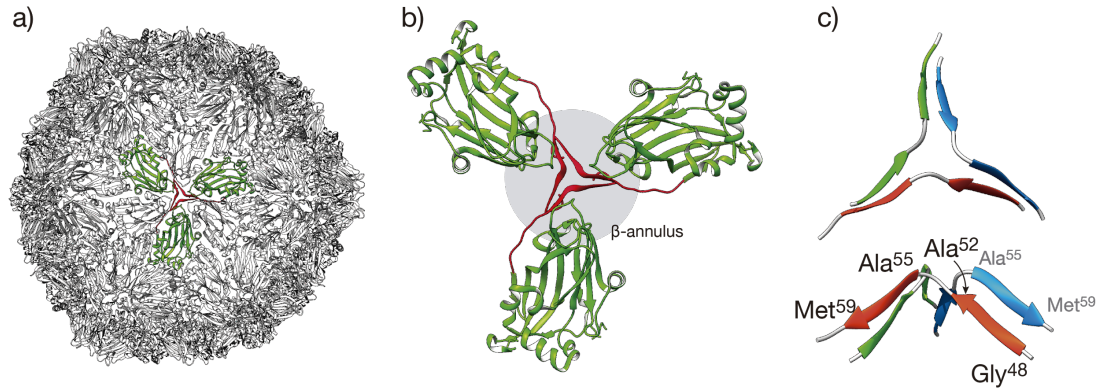


Figure 5. a) Structure of the SeMV capsid (PDB 1X35) where the β-annulus is highlighted in red and the C-subunit in green. b) Structure previously highlighted in the capsid of SeMV where the β-annulus is showed in red and the resulted trimer with the C-subunit highlighted in green. c) Structure of the β-annulus isolated from the rest of the structure showing the key residues involved in the trimerization of the structure.

Once the final peptide was obtained, the next step was the characterization of its coordinative properties as well as the chirality of the resulting helicate exploiting the quenching on the fluorescent emission of bipyridine exerted by Fe(II) ions upon coordination. The resulting data were adjusted to a 1:2 binding mode with dissociation constants of 14 and 15 μM for the entry of the first and the second Fe(II), respectively. In addition, circular dichroism studies showed a positive Cotton effect in the bipyridine at 320 nm, consistent with presence of a helicate with M chirality.

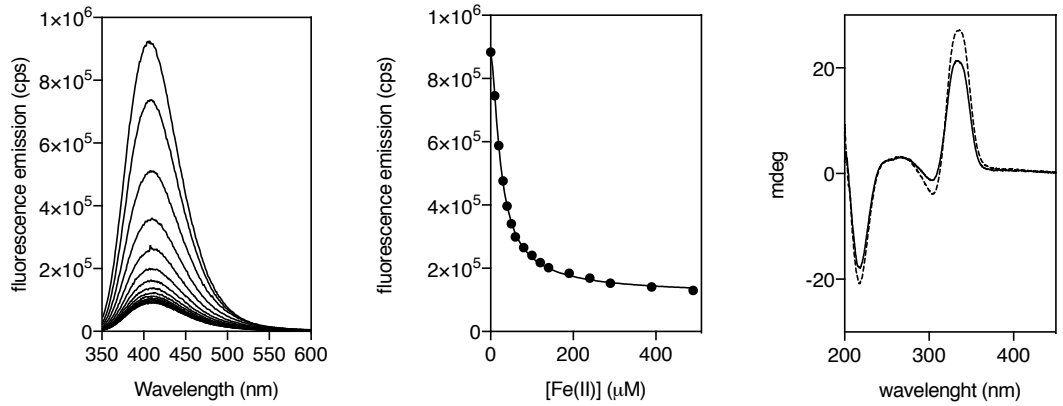


Figure 6. Left: emission spectra of a 20 μM **β-annX(5Bpy)$_2$** peptide solution (10 mM HEPES buffer, 100 mM NaCl, pH 6.5) with increasing concentrations of $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$. Center: Titration profile of the maximum emission wavelength at 410 nm with increasing concentrations of Fe(II); center: the best fit according to the 1:2 model in *DynaFit* is shown. Experimental data points corresponds to the average of three independent titrations; right: CD of a 7 μM solution (21 μM monomer) of ΛΛ-$[Fe(II)_2$(**β-annX(5Bpy)$_2$**$)_3]^{4+}$ (dashed line) and in the presence of 200 μM Fe(II) (solid line). All experiments were made in 10 mM Hepes buffer, pH 6.5, 100 mM NaCl at 20 °C. The represented data corresponds to the average of three accumulations.

Finally, the DNA binding properties of the helicate was characterized by measuring the quenching induced by the helicate on nearby fluorophores. Thus, we measured the

emission of a fluorescein-labelled twDNA (FAM-twDNA) in the preseence of increasing concentrations of the helicate. The data obtained could be adjusted to a 1:1 binding model, with a dissociation constant of 2.82 μM.

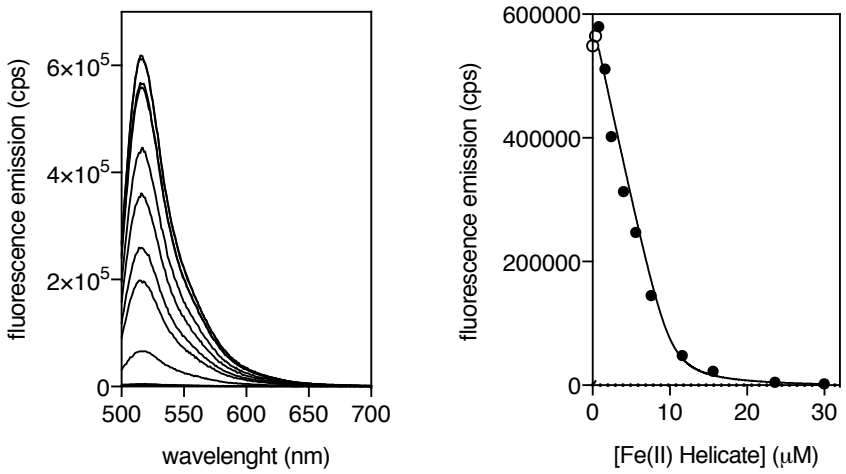


Figure 7. Left: Fluorescence spectra of a 2 μM solution of FAM-twDNA (10 mM HEPES buffer, 100 mM NaCl, pH 6.5) in the presence of increasing concentrations of the ΛΛ-$[Fe(II)_2(\textbf{β-annX(5Bpy)}_2)_3]^{4+}$ helicate. Right: titration profile and best fit to a 1:1 binding mode for twDNA (solid line, black circles); open circles represent values that are exclude of the1:1 binding mode fit.

In order to confirm the results obtained by direct fluorescence, DNA binding assays were also carried out by EMSA. These tests revealed the appearance of a new slow-migrating band as a result of the interaction between the obtained helicate and the twDNA. In addition, no slow-migrating band can be observed in the control lanes with dsDNA, thus confirming the selectivity of the helicate against the twDNA.

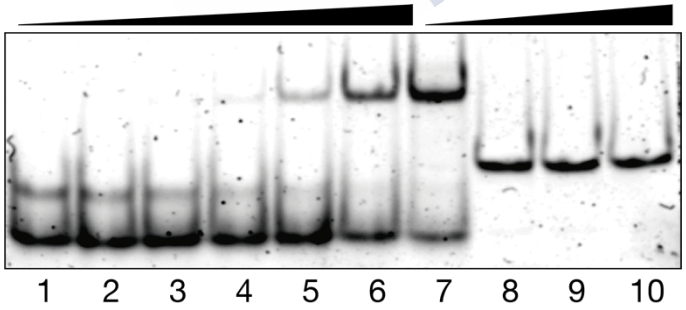


Figure 8. EMSA DNA binding studies results for ΛΛ-$[Fe(II)_2(\textbf{β-annX(5Bpy)}_2)_3]^{4+}$ helicate. Lanes 1-6, 200 nM twDNA with 0, 150, 250, 500, 1000, and 2000 nM of $[\textbf{β-annX(5Bpy)}_2]_3$ and 20 eq. of $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ in each lane; lanes 7-10, 200 nM dsDNA with 0, 500, 1000, and 2000 nM of $[\textbf{β-annX(5Bpy)}_2]_3$ and 20 eq. of $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ in each lane.

Chapter 3. Dynamic stereoselection of inert DNA-binding peptide helicates

This chapter demonstrates the obtention of enantiopure and kinetically inert peptide helicates through metal ion-directed folding of linear peptides that integrate 2,2'-bipyridine coordinating units in their sequence. These units are connected to each other by chiral sequences that promote β-turns and also improve the solubility of the final peptide ligands.

We synthesized the coordinating residue Fmoc-βAla5Bpy, which was integrated into a peptide sequence that replace the original GPG β-turns with a mixture of amino acids with alternating chiralities that promotes the formation of type II or II´ β-turns. Therefore, the sequences finally integrated into the new systems are composed of L-Arg–L-Pro–D-Arg and its mirror image D-Arg–D-Pro–L-Arg, which provide solubility in addition to improved DNA binding. The helicate precursor ligands **DDL**: (H-(βAla5Bpy)$_2$-rpR-(βAla5Bpy)$_2$-rpR-(βAla5Bpy)$_2$-NH_2 and **LLD**: (H-(βAla5Bpy)$_2$-RPr-(βAla5Bpy)$_2$-RPr-(βAla5Bpy)$_2$-NH_2 were synthesized following standard solid-phase peptide synthesis procedures

Once both sequences were synthesized and purified by reverse phase HPLC, their coordination properties and chiral structure were characterized in the presence of metal ions (Fe(II) or Co(II)) as explained before. The data were fitted to a 1:2 binding mode with dissociation constants in the order of 0.5 and 0.4 μM for both Fe(II) and Co(II). Circular dichroism studies also confirmed the selection of chiral helicates.

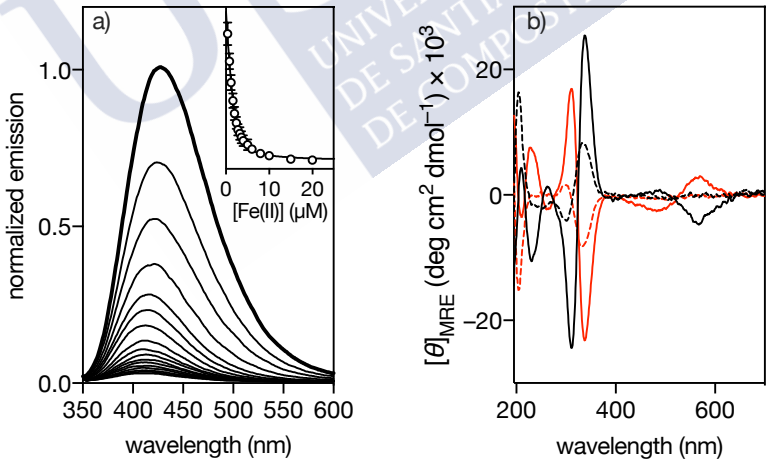


Figure 9. Titration of the **DDL** ligand. a) Normalized emission spectrum of a 2 μM solution of **DDL** in 1 mM phosphate buffer NaCl 10 mM, pH 6.5 (thick line), and spectra of the same solution in the presence of increasing concentrations of Fe(II) ions (thin lines); inset showing the titration profile of three independent experiments and best fit to a one to two binding model (same range as main plot); b) Circular dichroism spectra of 5 μM solutions of **LLD** (black dashed line) and **DDL** (red dashed line) in 1 mM phosphate buffer NaCl 10 mM, pH 6.5, and in the presence of 25 μM Fe(II) ions (same colors, continuous lines).

Once we characterized the metal binding properties and the chirality of the helicates, we synthesized their Co(III) kinetically inert derivatives with M chirality. To do this, the thermodynamically selected helicate resulting from the mixture of the **LLD** ligand and

Co(II) was oxidized using $Ce(NO_3)_6(NH_4)_2$, which resulted in the appearance of two new HPLC peaks corresponding to two new isomers, which were impossible to isolate previously due to the kinetic lability of the Fe(II) and Co(II) complexes.

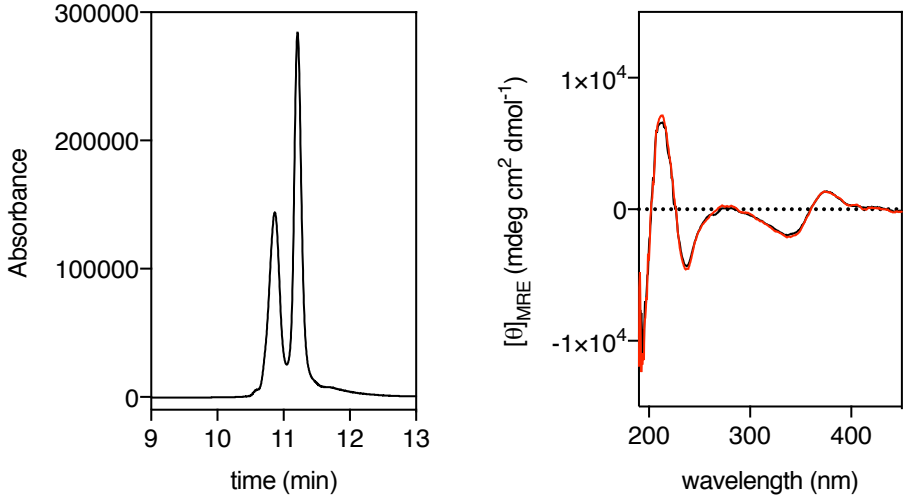


Figure 10. Left: chromatogram (0-40% of AcN + 0.1% TFA in 30 min) of the resulting mixture after the oxidation of the Co(II) helicate to Co(III). Each peak corresponds to a new formed diastereomer that in the case of the labile and dynamic Fe(II) and Co(II) was impossible to distinguish. Right: normalised CD spectra of the isolated isomer A (red line) and the isomer B (black line). Both of them shows a positive Cotton effect in the bipyridine band (c.a. 351 nm), demonstrating the retention of the ΛΛ- (*M*) configuration around the metal center.

Both isomers were isolated and characterized by circular dichroism, confirming the retention of the M configuration present in the Co(II) derivatives. We characterized the twDNA binding properties of the Fe(II) and Co(III) helicates as described before. In this case, the profile obtained in the case of Fe(II) could be adjusted to a 1:1 binding mode with a dissociation constant of 450 nM. Interestingly, in the case of Co(III) the profile obtained is not adjustable to a simple 1:1 model, but required the introduction of nonspecific interactions in the model, with dissociation constants of 7.9 and 4.5 μM.

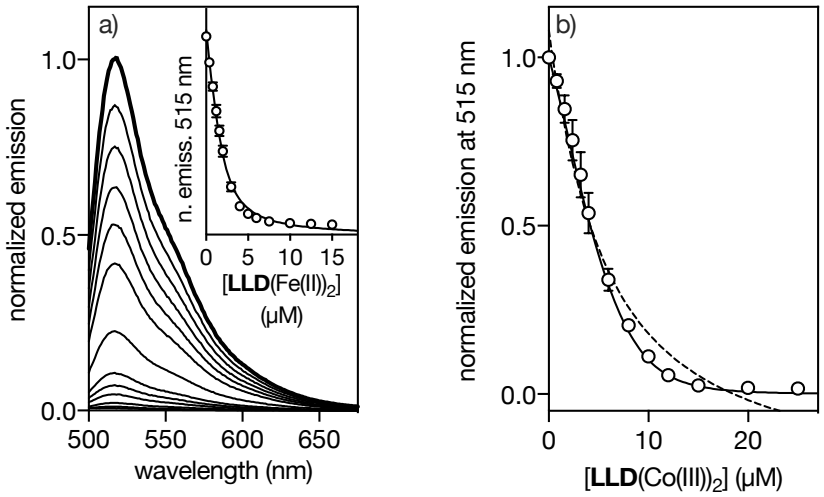


Figure 11. a) Normalized emission spectra of a titration of a 2 μM solution of the fluorescein-labeled three-way junction DNA, FAM-twDNA, in 1 mM phosphate buffer, 10 mM NaCl, pH 6.5 (thick line) with increasing concentrations of the $[Fe(II)_2$**Rho-LLD**$]^{4+}$ (thin lines); inset: corresponding normalized titration profile of the fluorescein quenching at 515 nm and best fit to a one to one binding model; b) Normalized emission of a 2 μM solution of the fluorescein-labelled three-way junction DNA, FAM-twDNA, in 10 mM phosphate buffer, 100 mM NaCl, 10 mM $MgCl_2$, pH 6.5 with increasing concentrations of the $[Co(III)_2$**Rho-LLD**$]^{6+}$; the best fit to a simple one to one binding model is shown as a dashed line, and the fit to a model including non-specific binding is shown as a continuous line. Titration data are the average of three independent experiments.

As in the previous cases, the DNA binding properties of the helicates were also studied by EMSA. In this case both helicates displayed good twDNA selectivity, which can be appreciated in the appearance of a new slow-migrating band in the presence of twDNA, and that no new bands were observed in the presence of ds-DNA. In addition, the results confirm the higher affinity of the Fe(II) complexes compared with those of Co(III), which can be appreciated in the rapid saturation observed in the case of Fe(II).

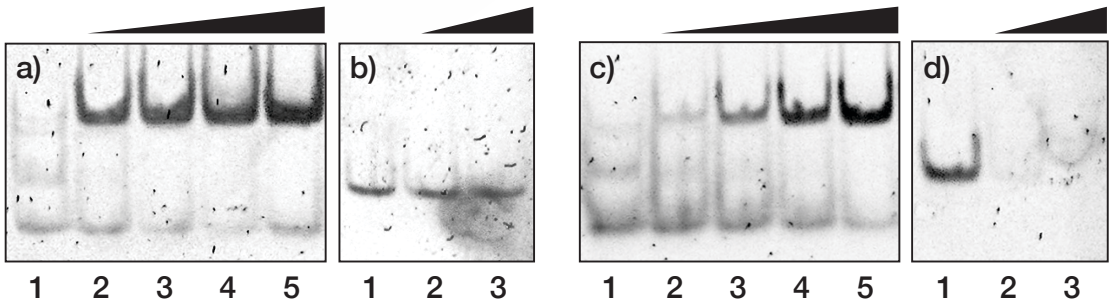


Figure 12. EMSA analysis of the **LLD** ligand binding to DNA. a) 200 nM tw DNA in each lane. 0, 50, 100, 250, 500 nM of **LLD** and 5 eq of $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$. b) 75 nM dsDNA, with 0, 1000, and 1500 nM of **LLD** and 5 eq of $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$. c) and d) panels correspond to the same conditions as a) and b), but with the Co(III) helicate.

Finally, the cell internalization properties of the **LLD** compound were studied. For this, two fluorescent analogs of the **LLD** ligand were synthesized with fluorescein or rhodamine at its N-terminus to observe the influence of the fluorophore over the internalization properties. Unfortunately, in all attempted assays, the resulting helicates remained trapped in endosomes. Assays performed by previously permeabilizing the

cytoplasmic membrane with digitonin demonstrated the nuclear location of the helicates. In addition, using a marker of replication zones, it was possible to confirm the ability of these helicates to mark zones of transient formation of twDNA in the cell replication

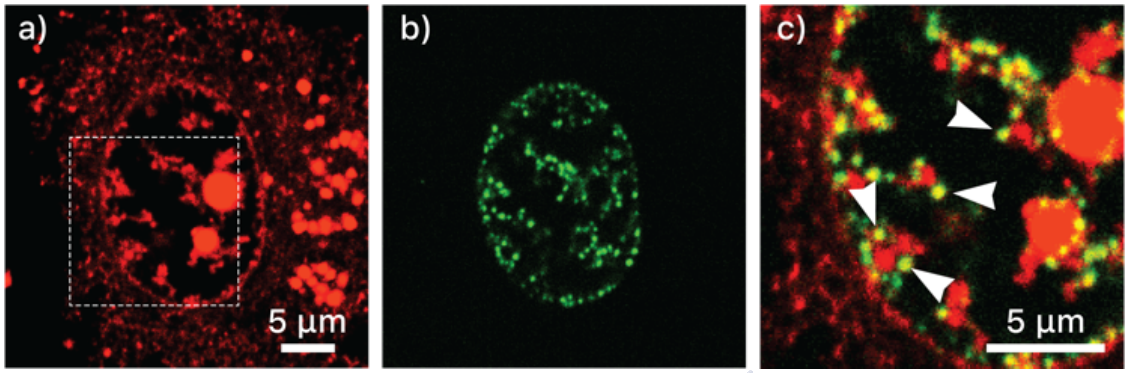


Figure 13. $[Fe(II)_2$**Rho-LLD**$]^{4+}$ selectively stains DNA replication sites. HeLa cells expressing protein GFP-PCNAL2 were incubated with 25 µg/ml Digitonin for 3 min, then 5 µM **Rho-LLD** and 25 µM $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ for 30 min. a) Red channel emission showing the distribution of the rhodamine-labeled helicate; b) green channel, corresponding to the emission of the GFP-PCNAL2 probe labeling the DNA replication foci; c) Overlay of the green and red channels of the square region shown in a). Arrows highlight some of the foci where the staining of GFP-PCNAL2 and $[Fe(II)_2$**Rho-LLD**$]^{4+}$overlap.

Experimental section

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 Fmoc-Nle-OH (*Nova Biochem*) and the Alloc derivatives employed in the synthesis of the four different annulus peptides (*Alfa Aesar*). 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 were obtained from *Sigma Aldrich*.

Reactives employed in the obtention of the metallopeptides were acquired *Sigma Aldrich* (Mohr's salt) and *Alfa Aesar* ($CoClO_4 \cdot 6H_2O$ and $(NH_4)_2Ce(NO_3)_6$).

Procedures and experimental techniques

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. Matrix was selected in function of the molecular weight of the peptides, being 4-HCCA (α-cyano-4-hydroxycinnamic acid) the chosen one for smaller peptides and sinapinic acid for those larger than 3 KDa.

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.

Fluorescence spectroscopy

Coordination experiments and direct quenching of fluorescein-labelled DNA's

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 adapted depending on the studied system.

Single point anisotropy titrations

Single point anisotropy was carried out in a *Jobin-Yvon Fluoromax-3* (*DataMax 2.20*) coupled to a temperature controller *Wavelength Electronics LFI–3751*. All measurements were made with a *Hellma* semi-micro cuvette (108F-QS) at 20 °C. The values are an accumulation of 3 measurements until the error was lower that 1%. The settings for these measurements were adapted depending on the system.

Circular Dichroism

Circular dichroism 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.

HPLC and UHPLC

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 on 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, 270 and 330 nm. Electrospray Ionization Mass Spectrometry (ESI/MS) was performed with an Agilent 6120 Quadrupole LC/MS model in positive scan mode using direct injection of the purified peptide solution into the MS detector.

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 on 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.

DNA hybridization

Oligonucleotides were obtained from *Biomers*. Concentration of the oligonucleotides was measured by UV-VIS by using the extinction coefficient given by the Comercial House Biomers.

For the hybridization process, a stochiometric mixture of the DNA strands was prepared in PBS buffer (10 mM, 100 mM NaCl, pH 7.0) and heated at 90 °C during 10 minutes. After

this time the mixture was slowly cooled down until room temperature, obtaining the resulting 3wj/ds hybridized DNA.

Cell Internalization studies

These experiments were carried out in collaboration with the **Prof. José Martínez Costas** from CIQUS. HeLa cells were grown in DMEM (Dulbecco modified Eagle Medium) with 10% of fetal bovine serum in a 12-plate box with a glass coverslip of 15 mm. Before carrying out the experiments, cells were washed three times with PBS and finally 0.5 mL of DMEM was added.

a) Experiments with $[Fe(II)_2$**FITC-LLD**$]^{4+}$ and $[Fe(II)_2$**Rho-LLD**$]^{4+}$: a determined volume was added to each plate in order to reach a final concentration of 5 μM of the compound. The resulting mixture was incubated overnight under an atmosphere of 5% of CO_2 at 37 °C. After this time the cells were washed three times with PBS and visualized.
b) Experiments with $[Fe(II)_2$**Rho-LLD_CHC**$]^{4+}$ and $[Fe(II)_2$**Rho-LLD_Au1.2**$]^{4+}$: a determined volume was added to each plate in order to reach a final concentration of 5 μM of the compound. The resulting mixture was incubated 30 min under an atmosphere of 5% of CO_2 at 37 °C. After this time the cells were washed three times with PBS and visualized.
c) Experiments with **Digitonin** and $[Fe(II)_2$**Rho-LLD**$]^{4+}$cells were washed twice with transfer buffer (25 mM HEPES, 125 mM KOAc, 2 mM $Mg(AcO)_2$, 1 mM EDTA). Afterthat cells were treated with a Digitonin solution (25 mg/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. Then, cells were incubated in transfer buffer with 5 μM of $[Fe(II)_2$**Rho-LLD**$]^{4+}$ for 15 minutes. After this time cells were washed three times with transfer buffer containing 10 mg/mL of BSA and resuspended in DMEM with 10% of fetal bovine serum. In this experiment, cells were previously transfected using lipofectamine to express proliferating cell nuclear antigen (PCNA) fused to GFP, acting as a label for the replication sites of the cell.

Images were alternatively obtained with an *Olympus* DP-71 digital camera mounted on an *Olympus* BX51 fluorescence microscope, or with an *Andor Dragonfly* spinning disk confocal system mounted on a *Nikon* TiE microscope equipped with a Zyla 4.2 PLUS camera (*Andor*).

EMSA *experiments*

EMSA 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 140 V (constant V). Binding reactions were performed over 30 min in 1.8 mM Tris-HCl (pH 7.5), 90 mM KCl, 1.8 mM $MgCl_2$, 0.2 mM TCEP, 9% glycerol, 0.11 mg/mL BSA, and 2.2% NP-40. For the experiments we used 200 nM of the DNAs (twDNA and dsDNA), and a total incubation volume of 20 μL. After incubation for 30 min 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

25 °C, and analyzed by staining with SyBrGold (Molecular Probes: 5 µL in 50 mL of 1 × TBE) for 10 min and visualized by fluorescence (BioRad GelDoc XR+ molecular imager).

Solid phase peptide synthesis

Foldon and β-annulus synthesis

C-terminal amide natural T4Ff and β-annulus 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. Once the synthesis is finished, the peptide was acetylated with a solution of 0.8 ml AcOH, 2 ml of DIEA/DMF (0.2 M) and 3.2 ml of DMF.

Deprotection of the alloc group

In the case of β-annulus derivatives, deprotection of the Alloc group is necessary before the coupling of the last two non-natural residues of Fmoc-βAla5Bpy-OH. Deprotection of the alloc group was carried out following the methodology proposed by Sainlos *et al.*[228] Thus, a suspension of the resin in DCM was degassed by gently bubbling of N_2 for 5 minutes. After this time $Pd(PPh_3)$ (0.8 eq.) and PhSiH (25 eq) were added and the flux of N_2 was maintained for 15 min. After this time the resin was washed with DCM three times and the previous process repeated twice.

Coupling of the Fmoc-βAla5Bpy-OH

The last non-natural Fmoc-βAla5Bpy-OH residue units were coupled by hand in 4-fold excess using HATU as 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. Deprotection of the temporal Fmoc protecting group was performed by treating the resin with 20% piperidine in DMF for 20 min.

LLD and DDL ligands synthesis

C-terminal amide **LLD** and **DDL** 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). The obtention of these peptides were 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.

228 M. Sainlos, B. Imperiali, *Nat. Protoc.* **2007**, *2*, 3201–3209.

In the case of the **Rho-LLD_Au1.2**, after the synthesis of the **LLD** ligand, the coupling of the Aurein 1.2 sequence was carried out using the Liberty Lite automatic microwave assisted peptide synthesizer from CEM Corporation.

TMRA and FITC couplings were made by hand and conducted in DIEA/DMF 0.2M during 2h. In

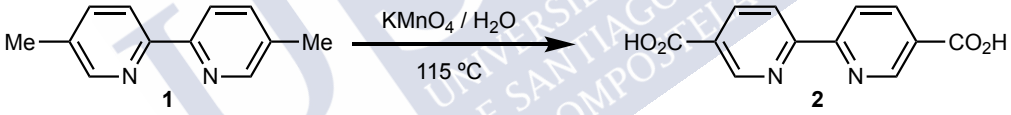
Cleavage from the resin and final deprotection

Cleavage and deprotection of the peptide were simultaneously performed using standard conditions by incubating the resin for 2.5 h with an acidic mixture containing 50 µL CH_2Cl_2, 25 µL of H_2O, 25 µL of TIS (triisopropylsilane), and 900 µL of TFA. 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 obtained residue was then dissolved in a mixture of H_2O and MeCN and purified by preparative HPLC for the obtention of the pure compounds.

Synthesis of Fmoc-βAla5Bpy-OH (8)

The synthesis of this chelating residue was carried out following the methodology recently published by our group with minor modifications.[111,112,135a] All the obtained products were analyzed by HPLC-MS to test their purity.

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



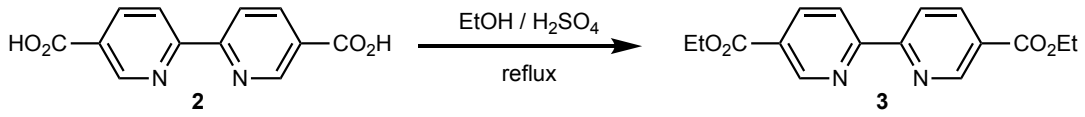
A mixture of 39 g of potassium permanganate and 7 g of 5,5'-dimethyl-2,2' -bipyridine in 250 mL of H_2O was heated for 2 h (115 ºC), cooled at room temperature and filtrated trough celite. The filtrate was cooled to 4 ºC and acidified with HCl until precipitation of a white solid, which was filtrated,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

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



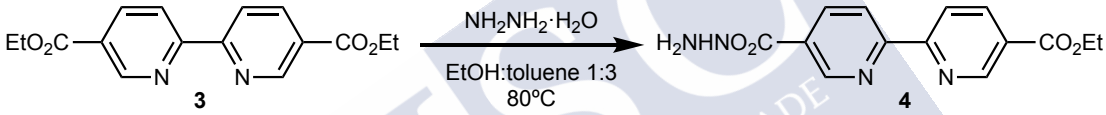
2,2'-bipyridine-5,5'-dicarboxylic acid (10.0 g, 41 mmol) was suspended in 150 mL of absolute ethanol. Concentrated sulfuric acid (20.0 mL) was slowly added, and the resulting mixture was refluxed for 18 h. The solution was cooled at room temperature and added over 400 mL of water at 4 ºC causing the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

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



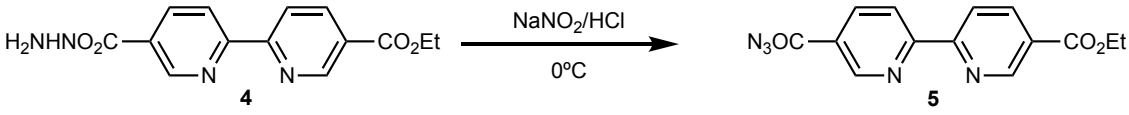
A mixture of diethyl 2,2'-bipyridine-5,5'-dicarboxylate (15.00 g, 50 mmol) and hydrazine hydrate (3.75 mL, 55 mmol) in a solution of EtOH (42 mL) and toluene (128 mL) was heated at 80 °C for 48 h. The precipitate was filtered, washed with $CHCl_3$ and dried under vacuum 11.4 g (80%). The unreacted diethyl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

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



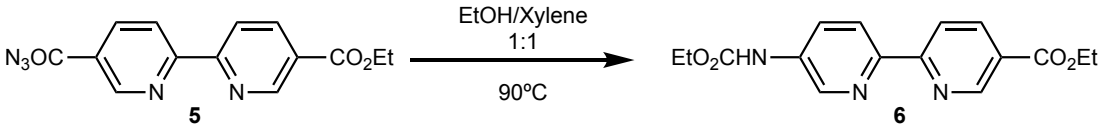
A stirred solution of 5-(ethoxycarbonyl)-5'-carbohydrazido-2,2'-bipyridine (5.7 g, 20 mmol) in concentrated HCl (100 mL) was cooled to 0 °C, and then an aqueous solution of $NaNO_2$ (1.73 g, 25 mmol; 15 mL) was added dropwise, maintaining the temperature at 0 °C. After 60 min, the yellow solution was 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

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



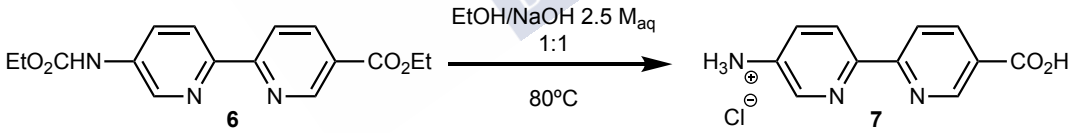
A solution of ethyl 5'-carbazido-2,2'-bipyridine-5-carboxylate (8.6 g, 29 mmol) in a mixture of EtOH (100 mL) and xylene (100 mL) was heated at 90 ºC for 4 h. The solvent was evaporated under reduced presure, and the yellow residue was washed with EtOH and dried in vacuo. 8.21 g (90%).

^{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)



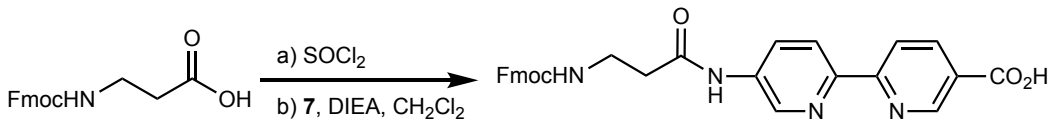
A stirred solution of ethyl 5'-[(ethoxycarbonyl)amino]-2,2'-bipyridine-5-carboxylate (14.55 g, 45.9 mmol) in a mixture of EtOH (50 mL) and 2.5 M aqueous NaOH (50 mL) was heated at 75 ºC for 14 h. The EtOH was concentrated in vacuum, and the aqueous solution was acidified with HCl to afford a bright yellow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

Synthesis of Fmoc-βAla5Bpy-OH (8)



Fmoc-βAla5Bpy-OH, **8**

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 (3x10 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/H2O (2:1), centrifuged and washed with acetonitrile and H_2O and lyophilized 6.70 g (89 %).

1H–RMN (300 MHz, DMSO-d_6, δ): 10.44 (s, 1H); 9.13 (d, 4*J* = 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).

13C-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.

Chapter 1

T4Ff: H-GYIPEAPRDGQAYVRKDGEWVLLSTFL-NH_2

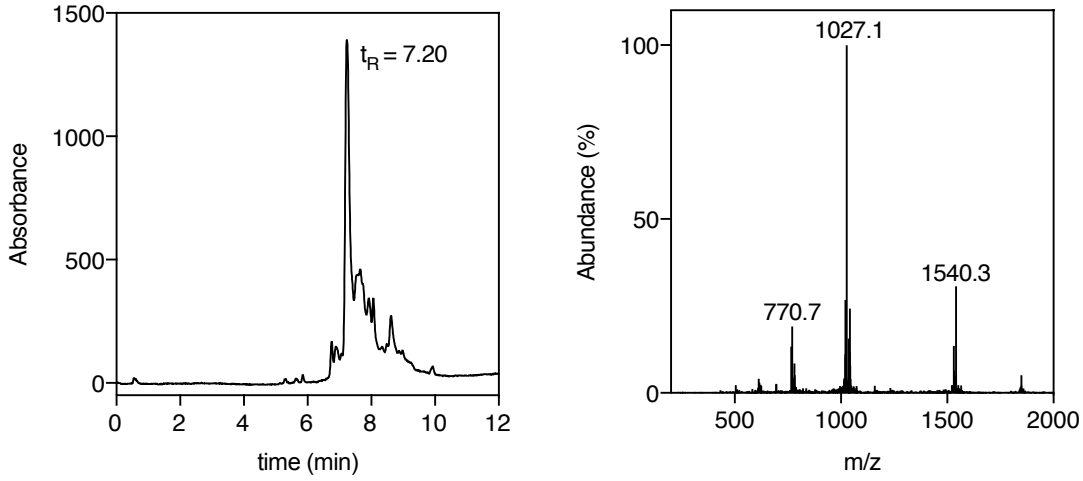


Figure 48. Right: crude chromatogram of T4Ff. Left: mass spectra corresponding to the peak wit t_R of 7.2 min.

HPLC-MS(ESI): (5-95% B, t_R= 7.2 min) Calculated for $C_{142}H_{215}N_{37}O_{40}$ = 3080.50; found $[M+2H]^{2+}$ = 1540.3; $[M+3H]^{3+}$ = 1027.1; $[M+4H]^{4+}$ = 770.7.

(βAla5Bpy)$_2$T4Ff: *Ac-(βAla5Bpy)$_2$GYIPEAPRDGQAYVRKDGEWVLLSTFL-NH_2*

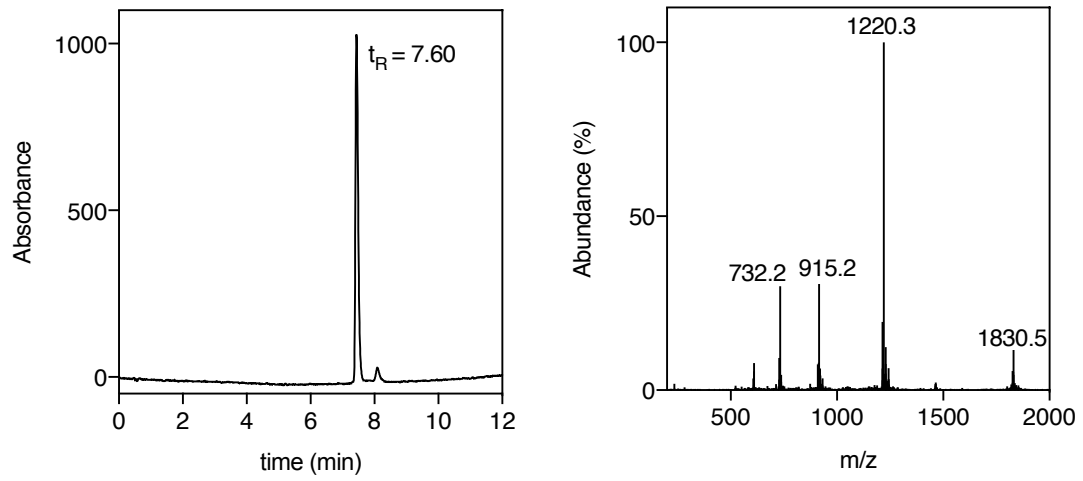


Figure 49. Left: chromatogram of the purified **(βAla5Bpy)$_2$T4Ff**. Right: mass spectra corresponding to the peak wit t_R of 7.6 min.

HPLC-MS (ESI): (5-95% B, t_R= 7.60 min) Calculated for $C_{172}H_{241}N_{45}O_{45}$ = 3659.09; found $[M+2H]^{2+}$ = 1830.5; $[M+3H]^{3+}$ = 1220.3; $[M+4H]^{4+}$ = 915.2; $[M+5H]^{5+}$= 732.2.

MS (MALDI-TOF): Calculated for $C_{172}H_{241}N_{45}O_{45}$ = 3659.09; found: $[M+H]^{+}$ = 3658.90.

Synthesis of the Fe(II) peptide helicate

Over a 3 μM solution of [(**βAla5Bpy)$_2$-T4Ff**]$_3$ in phosphate buffer (1 mM, 10 mM NaCl, pH 6.5), $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ (Mohr's salt) was added as source of Fe(II) ions until saturating conditions (15 eq).

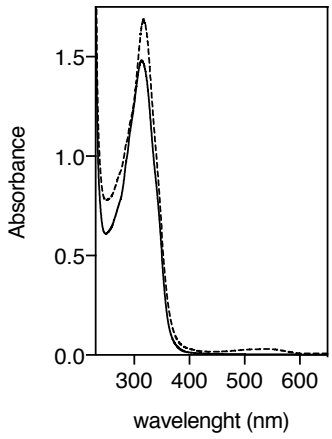


Figure 50. UV-VIS spectra of a 9 μM solution of [(**βAla5Bpy)$_2$-T4Ff**]$_3$ ligand before (black solid line) and after the addition of 15 equivalents of $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ (black dashed line). Formation of the corresponding Fe(II) complex can be confirmed with the appearance of a MLCT band centered at 535 nm.

MS (MALDI-TOF): Calculated for $[Fe(II)_2[(\textbf{βAla5Bpy)}_2\textbf{-T4Ff}]_3]^{4+}$ $C_{516}Fe_2H_{723}N_{135}O_{135}$ = 11082.25; found $[M\text{-}3H]^+$ = 11084.60.

Chapter 2

βannX(Alloc): *H-GISnLAPK(Alloc)AQGAM-NH₂*

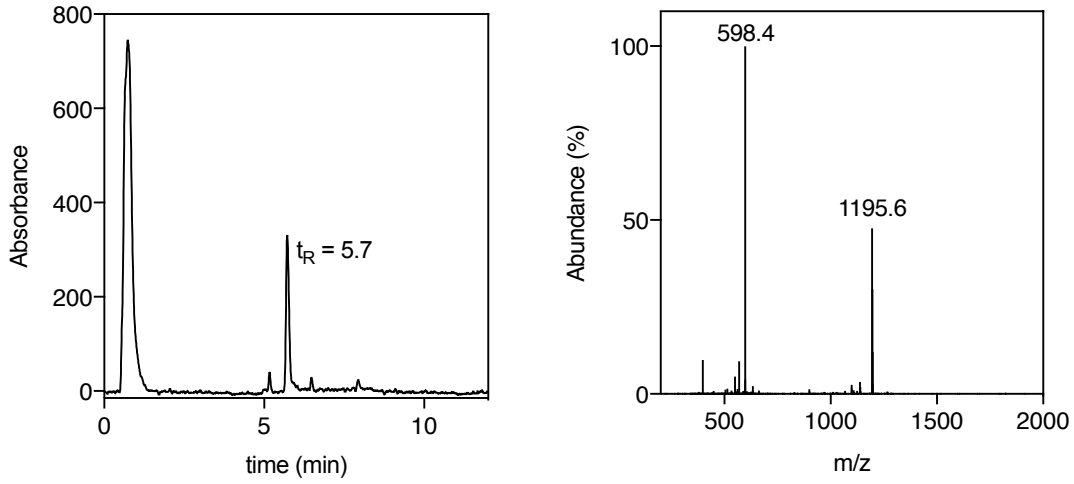


Figure 51. Left: chromatogram of the crude of βannX(Alloc). Right: mass spectra corresponding to the peak wit t_R of 5.7 min.

HPLC-MS (ESI): (5-95% B, t_R= 5.7 min) Calculated for $C_{52}H_{90}N_{16}O_{16}$ = 1194.67; found $[M+H]^+$ = 1195.6; $[M+2H]^{2+}$ = 598.4.

βannX(βAla5Bpy)₂: *AcNH-GISnLAPK(βAla5Bpy)₂AQGAM-NH₂*

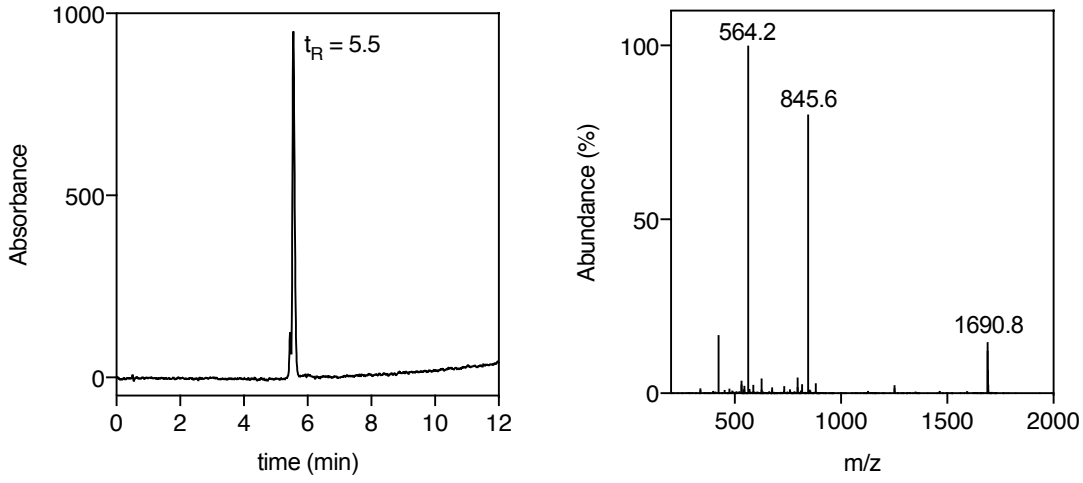


Figure 52. Left: chromatogram of the purified **βannX(βAla5Bpy)₂**. Right: mass spectra corresponding to the peak wit t_R of 5.5 min.

HPLC-MS (ESI): (5-95% B, t_R= 5.5 min) Calculated for $C_{78}H_{112}N_{24}O_{19}$ = 1688.85; found $[M+H]^+$ = 1690.8; $[M+2H]^{2+}$ = 845.6; $[M+3H]^{3+}$ = 564.2.

MS (MALDI-TOF): Calculated for $C_{78}H_{112}N_{24}O_{19}$ = 1688.85; found: $[M+H]^+$ =1689.90.

Synthesis of the Fe(II) peptide helicate

Over a 5 μM solution of [**βannX (βAla5Bpy)$_2$**]$_3$ in HEPES buffer (10 mM, 100 mM NaCl, pH 6.5), $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ (Mohr's salt) was added as source of Fe(II) ions until saturating conditions (15 eq).

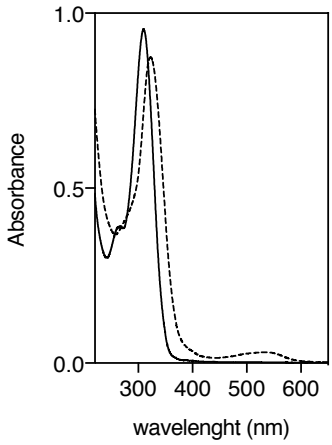


Figure 53. UV-VIS spectra of a 5 μM solution of [(**βAla5Bpy)$_2$-T4Ff**]$_3$ helicate precursor before (black solid line) and after the addition of 15 eq equivalents of $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ (black dashed line). Formation of the corresponding Fe(II) complex can be confirmed with the appearance of a MLCT band centered at 535 nm.

MS (MALDI-TOF): Calculated for $C_{78}H_{112}N_{24}O_{19}$ = 1688.85; found: $[M-2Fe+H]^+$ = 5065.74.

Chapter 3

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

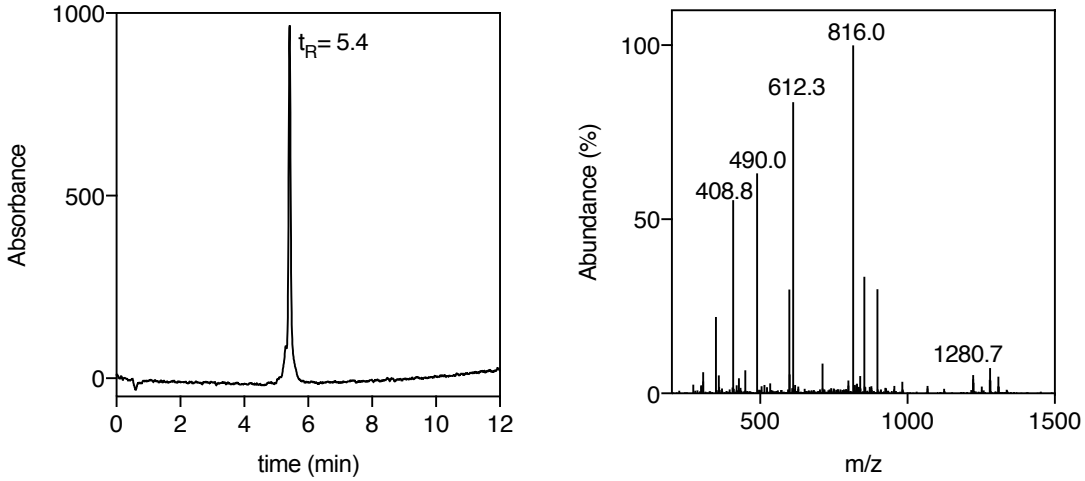


Figure 54. Left: chromatogram of the purified **DDL**. Right: mass spectra corresponding to the peak with t_R of 5.4 min.

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; $[M+4H]^{4+}$ = 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: $[M+H]^+$ = 2446.11.

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

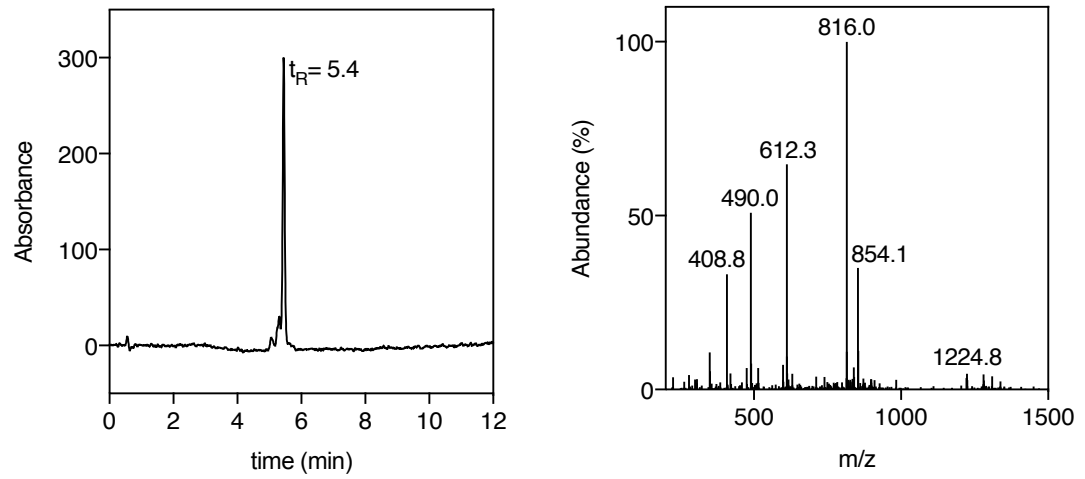


Figure 55. Left: chromatogram of the purified **LLD**. Right: mass spectra corresponding to the peak wit t_R of 5.4 min.

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.

FITC-LLD: FITC-Amhx(βAla5Bpy)$_2$RPR(βAla5Bpy)$_2$RPR(βAla5Bpy)$_2$-NH_2

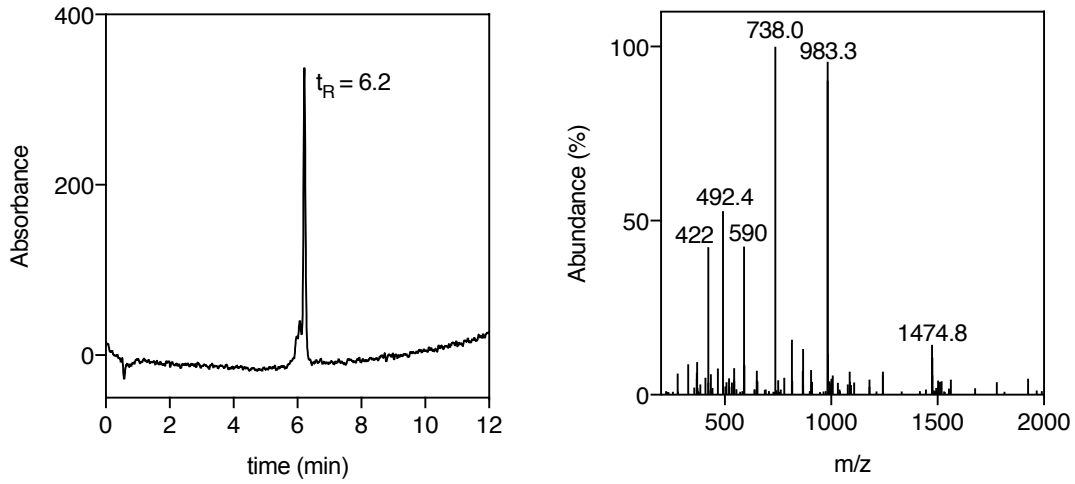


Figure 56. Left: chromatogram of the purified **FITC-LLD**. Right: mass spectra corresponding to the peak wit t_R of 6.2 min.

HPLC-MS (ESI): (5-95% B, t_R= 6.2 min) Calculated for $C_{145}H_{159}N_{45}O_{24}S$ = 2946.23; found $[M+2H]^{2+}$ = 1474.8; $[M+3H]^{3+}$ = 983.3; $[M+4H]^{4+}$ = 738.0; $[M+5H]^{5+}$ = 590.4; $[M+6H]^{6+}$ = 492.4; $[M+7H]^{7+}$ = 422.0.

MS (MALDI-TOF): Calculated for $C_{145}H_{159}N_{45}O_{24}S$ = 2946.23; found: $[M+H]^+$ = 2948.51.

Rho-LLD: Rho-O1Pen(βAla5Bpy)$_2$RPR(βAla5Bpy)$_2$RPR(βAla5Bpy)$_2$-NH_2

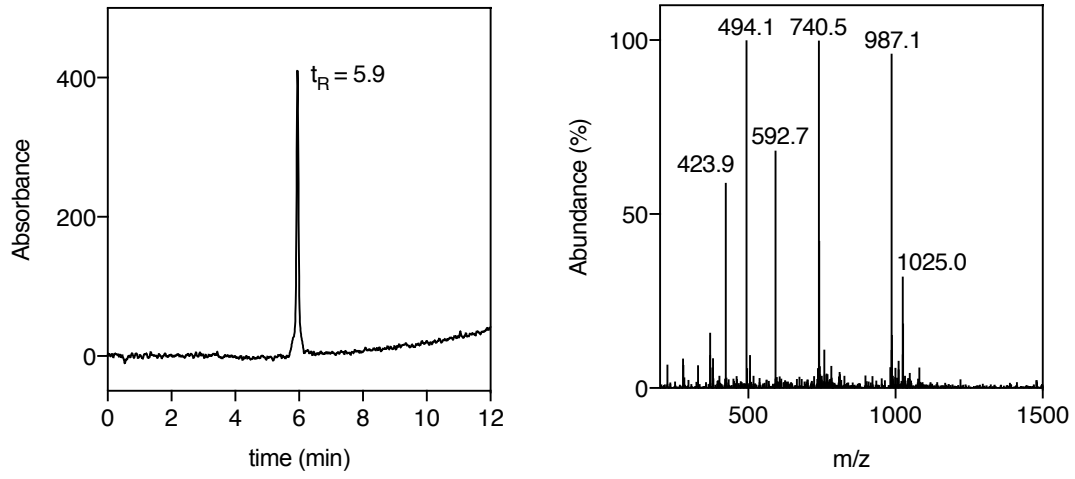


Figure 57. Left: chromatogram of the purified **Rho-LLD**. Right: mass spectra corresponding to the peak wit t_R of 5.9 min.

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.

Rho-LLD_Au1.2:

Rho-O1PenGLFDIIKKIAESF(βAla5Bpy)$_2$RPR(βAla5Bpy)$_2$RPR(βAla5Bpy)$_2$-NH_2

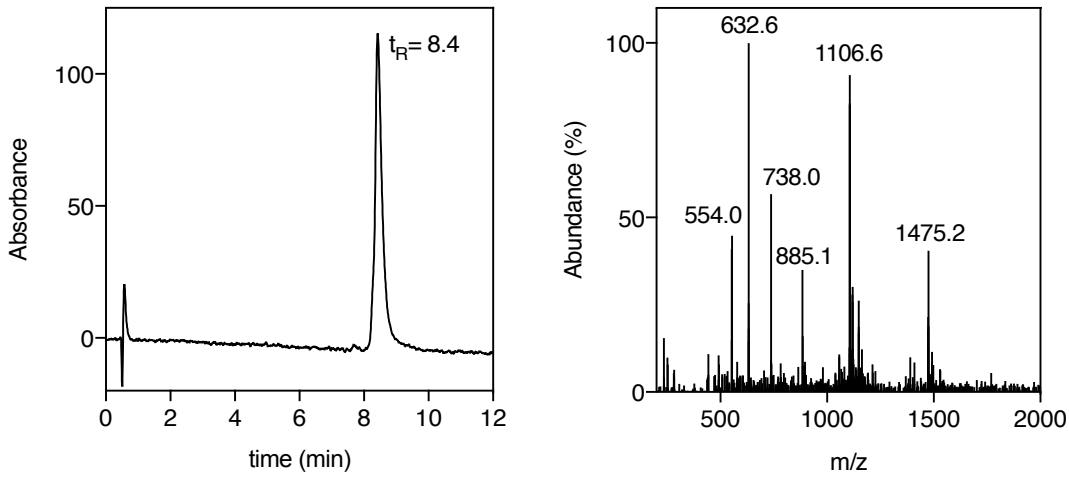


Figure 58. Left: chromatogram of the purified **Rho-LLD_Au1.2**. Right: mass spectra corresponding to the peak wit t_R of 8.4 min.

HPLC-MS (ESI): (5-95% B, t_R= 8.4 min) Calculated for $C_{218}H_{277}N_{69}O_{42}$ = 4421.14; found $[M+3H]^{3+}$ = 1475.2; $[M+4H]^{4+}$ = 1106.6; $[M+5H]^{5+}$ = 885.1; $[M+6H]^{6+}$ = 738.0; $[M+7H]^{7+}$ = 632.6; $[M+8H]^{8+}$ = 554.0;

MS (MALDI-TOF): Calculated for $C_{218}H_{277}N_{69}O_{42}$ = 4421.14; found: $[M+H]^+$ = 4422.42.

Rho-LLD_CHC: Rho-O1PenCHC(βAla5Bpy)$_2$RPR(βAla5Bpy)$_2$RPR(βAla5Bpy)$_2$-NH_2

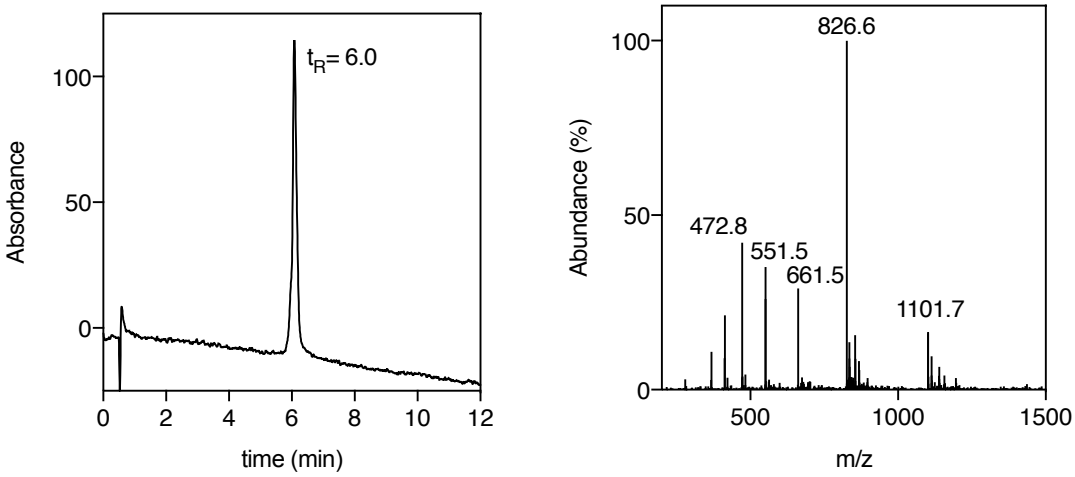


Figure 59. Left: chromatogram of the purified **Rho-LLD_CHC**. Right: mass spectra corresponding to the peak wit t_R of 6.0 min.

HPLC-MS (ESI): (5-95% B, t_R= 6.0 min) Calculated for $C_{159}H_{184}N_{51}O_{27}S_2$ = 3303.40; found $[M+3H]^{3+}$ = 1101.7; $[M+4H]^{4+}$ = 826.6; $[M+5H]^{5+}$ = 661.5; $[M+6H]^{6+}$ = 551.5; $[M+7H]^{7+}$ = 472.8.

MS (MALDI-TOF): Calculated for $C_{159}H_{184}N_{51}O_{27}S_2$ = 3304.40; found: $[M+H]^+$ = 3303.6.

Synthesis of the Fe(II) peptide helicates

Onto a 2 μM solution of any of the **LLD/DDL** derivatives in phosphate buffer (1 mM, 10 mM NaCl, pH 6.5), $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ (Mohr's salt) was added as source of Fe(II) ions until saturating conditions (5 eq).

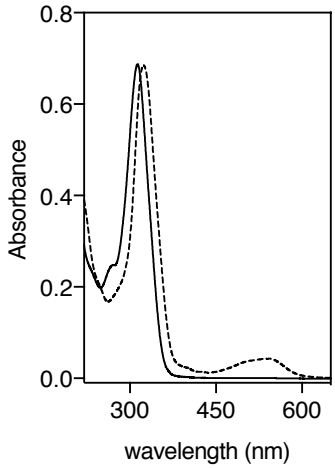


Figure 60. UV-VIS spectra of a 4 μM solution of **DDL** ligand before (black solid line) and after the addition of 5 eq equivalents of $(NH_4)_2Fe(SO_4)_2 \cdot 6\ H_2O$ (black dashed line). Formation of the corresponding Fe(II) complex can be confirmed with the appearance of a MLCT band centered at 535 nm.

MS (MALDI-TOF): Calculated for $[Fe(II)_2\mathbf{DDL}]^{4+}$ $C_{118}Fe_2H_{137}N_{43}O_{18}$ = 2555.97; found: $[M+2Fe-3H]^+$ = 2554.04.

MS (MALDI-TOF): Calculated for $[Fe(II)_2\mathbf{LLD}]^{4+}$ $C_{118}Fe_2H_{137}N_{43}O_{18}$ = 2555.97; found: $[M+2Fe-3H]^+$ = 2555.33.

MS (MALDI-TOF): Calculated for $[Fe(II)_2\mathbf{FITC\text{-}LLD}]^{4+}$ $C_{145}Fe_2H_{159}N_{45}O_{24}S$ = 3058.09; found: $[M+2Fe-3H]^+$ = 3056.24.

MS (MALDI-TOF): Calculated for $[Fe(II)_2\mathbf{Rho\text{-}LLD}]^{4+}C_{147}Fe_2H_{166}N_{46}O_{24}$ = 3071.17; found: $[M+2Fe-3H]^+$ = 3068.26.

MS (MALDI-TOF): Calculated for $[Fe(II)_2\mathbf{Rho\text{-}LLD\_Au1.2}]^{4+}$ $C_{218}Fe_2H_{277}N_{69}O_{42}$ = 4533.99; found: $[M+Fe-H]^+$ = 4476.34;

MS (MALDI-TOF): Calculated for $[Fe(II)_2\mathbf{Rho\text{-}LLD\_CHC}]^{4+}$ $C_{159}Fe_2H_{184}N_{51}O_{27}S_2$ = 3416.26; found: $[M+2Fe-3H]^+$ = 3411.45.

Synthesis of the Co(II) peptide helicates

A fluorescence titration of a 2 μM solution of the **DDL** peptide in phosphate buffer (1 mM, 10 mM NaCl, pH 6.5) with increasing concentrations of $Co(ClO_4)_2 \cdot 6\ H_2O$ resulted in a series of emission spectra with decreasing emission intensity due to the quenching of the 5'-amido-[2,2'-bipyridine]-5-carboxamide units by coordination of Co(II) ions. The resulting titration profile at 420 nm could be fitted to a 1:2 binding mode using *Dynafit* software with dissociation constants for the first, and second iron coordination of 0.54 and 0.44 μM.

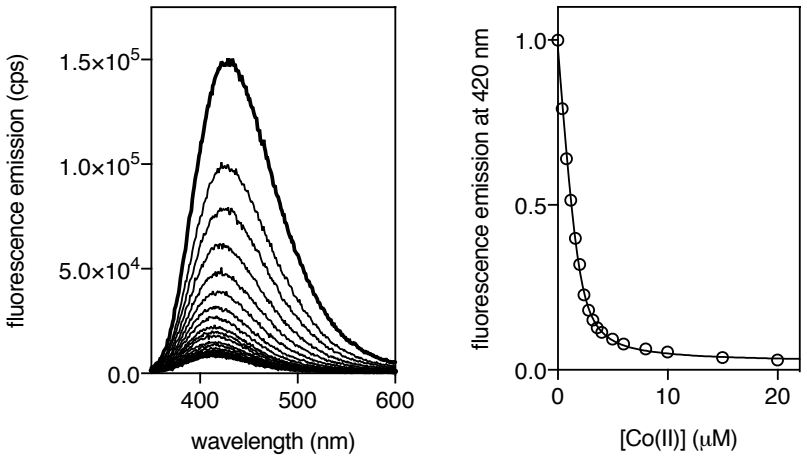


Figure 61. Left: emission spectra of a 2 μM **LLD** peptide solution (1 mM PBS buffer, 10 mM NaCl, pH 6.5) with increasing concentrations of $Co(ClO_4)_2 \cdot 6\ H_2O$. Right: Titration profile of the maximum emission wavelength at 410 nm with increasing concentrations of Co(II) ions. The best fit according to the 1:2 model in DynaFit is also shown. Experimental data points correspond to the average of three independent titrations.

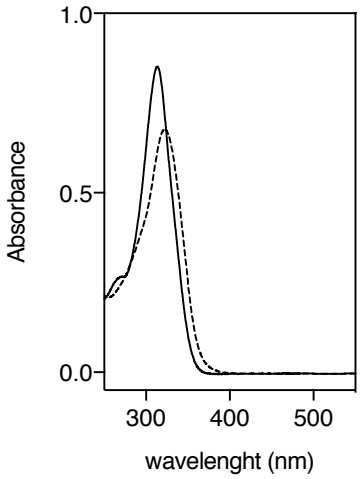


Figure 62. UV-VIS spectra of a 4 μM solution of **DDL** ligand before (black solid line) and after the addition of 5 eq equivalents of $Co(ClO_4)_2 \cdot 6\ H_2O$ (black dashed line).

MS (MALDI-TOF): Calculated for $[Co(II)_2\mathbf{LLD}]^{4+}$ $C_{172}Co_2H_{241}N_{45}O_{45}$ = 2561.97; found: $[M-3H]^+$ = 2561.10.

Synthesis of the Co(III) peptide helicate

Over a 600 µM solution of **LLD**, 5 equivalents of $Co(ClO_4)_2 \cdot 6\ H_2O$ (3 mM) were added from a 100 x stock solution. After 30 min of stabilization, 5.5 equivalents of $(NH_4)_2Ce(NO_3)_6$ (3.3 mM) were added over the preformed Co(II) helicate. After another 30 min, the two obtained isomers were purified by semipreparative HPLC in reverse phase (0-40% of AcN + 0.1% TFA in 30 min) and freeze dried, yielding a yellow solid (≈ 80% yield).

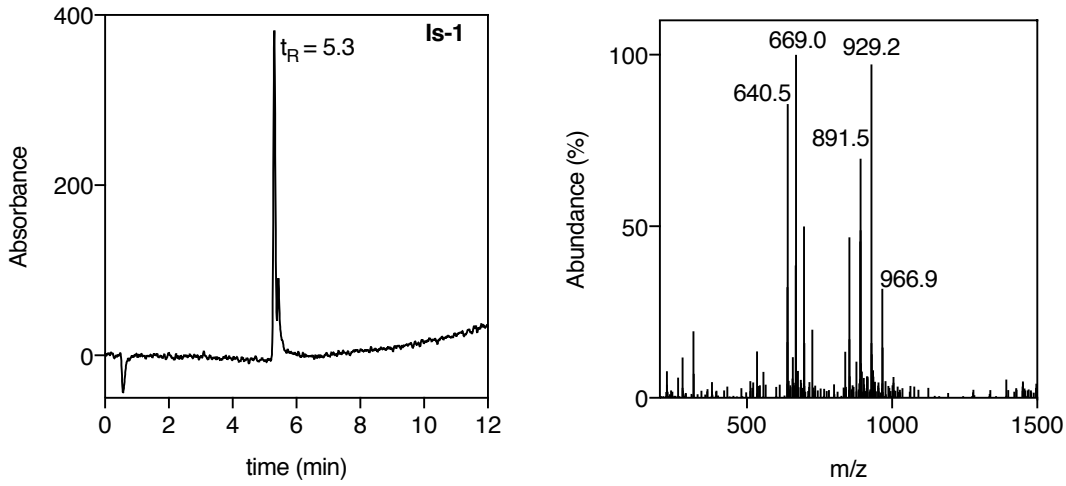


Figure 63. Left: chromatogram of the purified isomer I of $[Co(III)_2\mathbf{LLD}]^{6+}$. Right: mass spectra corresponding to the peak wit t_R of 5.3 min.

HPLC-MS (ESI): (5-95% B, t_R= 5.3 min) Calculated for $C_{172}Co_2H_{241}N_{45}O_{45}$ = 2561.97; found $[M+3H+TFAH]^{3+}$ = 891.5; $[M+3H+2TFAH]^{3+}$ = 929.2; $[M+3H+3TFAH]^{3+}$ = 966.9; [M+4H]4+ = 640.5; [M+4H+TFAH]4+ = 669.0.

MS (MALDI-TOF): Calculated for $C_{172}Co_2H_{241}N_{45}O_{45}$ = 2561.97; found: $[M+H]^+$ = 2561.11.

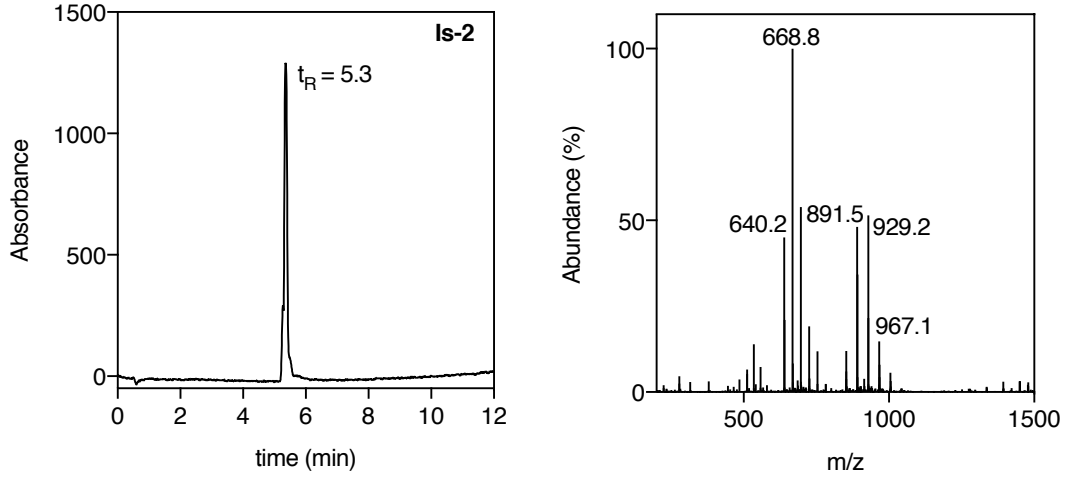


Figure 64. Left: chromatogram of the purified isomer 2 of $[Co(III)_2\mathbf{LLD}]^{6+}$. Right: mass spectra corresponding to the peak wit t_R of 5.3 min.

HPLC-MS (ESI): (5-95% B, t_R= 5.3 min) Calculated for $C_{172}Co_2H_{241}N_{45}O_{45}$ = 2561.97; found $[M+3H+TFAH]^{3+}$ = 891.5; $[M+3H+2TFAH]^{3+}$ = 929.2; $[M+3H+3TFAH]^{3+}$ = 967.1; $[M+4H]4+$ = 640.2; $[M+4H+TFAH]4+$ = 668.8.

MS (MALDI-TOF): Calculated for $C_{172}Co_2H_{241}N_{45}O_{45}$ = 2561.97; found: $[M+H]^+$ = 2563.20.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA