INTERNATIONAL DOCTORAL SCHOOL OF THE USC

Carmen
González González

PhD Thesis

Peptide Engineering: From Fluorescent Biomarkers to Bioorthogonal Catalysts

Santiago de Compostela, 2025

Doctoral Programme in Chemical Science and Technology

ESCOLA DE DOUTORAMENTO
INTERNACIONAL DA USC

DOCTORAL THESIS

PEPTIDE ENGINEERING: FROM FLUORESCENT BIOMARKERS TO BIOORTHOGONAL CATALYSTS

Carmen González González

Director: M. Eugenio Vázquez

Tutor: M. Eugenio Vázquez

PHD PROGRAMME IN CHEMICAL SCIENCE AND TECHNOLOGY

SANTIAGO DE COMPOSTELA

AGRADECIMIENTOS

En primer lugar, me gustaría agradecer a mi director Eugenio, por haberme dado la oportunidad de realizar esta tesis y descubrir el mundo de la investigación. Gracias por confiar en mí desde el principio y acompañarme en cada etapa de este proceso. También quiero agradecer a Miguel, por la ayuda y motivación en los seminarios, fundamentales tanto para mi formación como para que esta tesis saliese adelante.

También quiero agradecer al proyecto europeo que hizo posible mi contratación: *European Union's Horizon 2020 FET Open*. Esta financiación fue fundamental para llevar a cabo esta investigación y seguir creciendo profesionalmente.

En tercer lugar, quiero dar las gracias a mis compañeros de laboratorio, con los que he compartido tantos buenos momentos durante estos años. A David, por su infinita paciencia (como él dice) y por enseñarme desde lo más básico hasta los trucos del día a día en el lab. También a Ana, Diego, Axel y Alberto, con quienes he vivido esta aventura del doctorado codo con codo. Gracias por estar siempre ahí, por los cafés, las risas, las charlas y por echarnos una mano siempre que hacía falta. Ha sido un lujo teneros como compañeros y hacer equipo con vosotros. Y, por supuesto, también a los últimos en llegar, David Alvar, Patri y Sara.

También quiero dedicar unas palabras de agradecimiento a otros compañeros que a pesar de no coincidir tanto tiempo en el laboratorio su ayuda fue indispensable para que esta tesis saliera adelante. En primer lugar, a Roi, por estar siempre dispuesto a echar una mano y enseñarnos tanto. A Soraya, por su ayuda tan valiosa en temas de catálisis y por confiarme algunos de sus proyectos cuando se marchó. Y a Laura, por su colaboración en las partes computacionales de los proyectos de catálisis.

También quiero agradecer a mis amigas de química: Paula, Carlota, Patri, Lucía y, en especial, Laura. Gracias por todas las risas, los desahogos científicos (y no tan científicos) y por estar siempre ahí para desconectar cuando más hacía falta.

Por último, quiero agradecer a mi familia, por poder contar siempre con ellos. A mis padres, por siempre estar orgullosos de mí, por animarme en los momentos difíciles y por hacer posible, desde el principio, que llegara hasta aquí. En especial, esta tesis se la quiero dedicar a mi madre, quien estuvo siempre pendiente de cada paso de todos mis proyectos, esta tesis también es suya. También agradecer a mi hermana Concha, por siempre escucharme y por hacerme sentir que, sin importar lo que pase, siempre puedo contar con ella. Y por supuesto, a Javi, por ser mi apoyo incondicional durante estos años. Gracias por tu infinita paciencia, escuchando mis problemas y proyectos una y otra vez, y por aprenderte todos los detalles como si también fueran tuyos. No sé qué hubiera hecho sin ti.

A mi madre

Table of Contents

Abbreviations

ε	*Molar extinction coefficient*
δ	*Chemical shift*
λ	*Wavelength*
ϕ	*Quantum yield*
τ	*Luminescence lifetime*
ADEPT	*Antibody-directed enzyme-driven prodrug therapy*
AF	*Alexa Fluor*
alloc	*allyloxycarbonyl*
ArM	*Artificial metalloenzymes*
a.u.	*Arbitral units*
Boc	*tert-butyloxycarbonyl*
BODIPY	boron-dipyrromethene
bpy	*2,2'-bipyridine*
CD	*Circular dichroism*
CL	*Clusteroluminescence*
COD	*1,5-cyclooctadiene*
COSY	*COrrelation SpectroscopY*
CRISP	*Clustered Regularly Interspaced Short Palindromic Repeats*
CTC	*2-Chlorotrityl Chloride*
CTE	*Clustering-triggered emission*
CuAAC	*Copper-Catalyzed Azide-Alkyne Cycloaddition*
DCM	*Dichloromethane*
DFT	*Density Functional Theory*
DIC	*N,N'-Diisopropylcarbodiimide*
DIEA	*N,N-Diisopropylethylamine*
DMEN	*Dulbecco's Modified Eagle medium*
DMF	*Dimethylformamide*
DMSO	*Dimethyl sulfoxide*
DNA	*Deoxyribonucleic acid*
DOTA	*2,2′,2′′,2′′′-(1,4,7,10-Tetraazacyclododecane-1,4,7,10-tetrayl)tetraacetic acid*
EEM	*Excitation–Emission Matrix*
en	*Ethylenediamine*
ESI-MS	*Electrospray Ionization Mass Spectrometry*
FBS	*Fetal Bovine Serum*
FITC	Fluorescein isothiocyanate
Fmoc	*Fluorenyl methoxycarbonyl*
FRET	*Förster Resonance Energy Transfer*
GFP	*Green Fluorescent Protein*

GndAcO	*Guanidinum acetate*
HATU	*Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium*
HBI	*4-(p-hydroxybenzylidene)imidazolidin-5-one*
HEPES	*4-(2-hidroxietil)-1-piperazinaetanosulfonato (buffer)*
HOBt	*Hydroxybenzotriazole*
HOMO	*Highest Occupied Molecular Orbital*
HPLC	*High Performance Liquid Chromatography*
IC	*Internal conversion*
ICT	*Intramolecular Charge Transfer*
ISC	*Intersystem conversion*
K_D	*Dissociation constant*
LOMO	*Lowest Unoccupied Molecular Orbital*
MA	*Maleic Acid*
MW	*Microwave*
NaAsc	*Sodium ascorbate*
NAF	*Non-Aromatic Fluorescence*
NHS	*N-Hydroxysuccinimide*
NMR	*Nuclear Magnetic Resonance*
OBOC	*One Bead-One Compound*
PBS	*Phosphate Buffered Saline*
PET	*Photoinduced electron transfer*
PDB	*Protein Data Base*
phen	*Phenanthroline*
PhSi	*Phenylsilane*
ProCharTS	*Protein Charge Transfer Spectra*
PyBOP	*Benzotriazole-1-yloxytripyrrolidinophosphonium hexafluorophosphate*
RP-HPLC	*Reversed-phase HPLC*
rt	*Room temperature*
RU	*Raman units*
SAH	*Single Alpha Helix*
SHB	*Short Hydrogen Bond*
SPPS	*Solid Phase Peptide Synthesis*
tBu	*Tert-Butyl*
TFA	*Trifluoroacetic acid*
TIS	*Triisopropylsilane*
TMR	*Tetramethyl rhodamine*
TLC	*Thin-layer chromatography*
TRIS	*tris(hydroxymethyl)aminomethane*
UV/Vis	*Ultraviolet/visible*
Wzip	*Tryptophan zipper motif*

1. RESUMO

O campo da bioloxía química evolucionou significativamente nas últimas décadas, integrando ferramentas da química orgánica, bioquímica e nanotecnoloxía para abordar cuestións biolóxicas complexas mediante o deseño de moléculas funcionais. Neste contexto, os péptidos emerxeron como plataformas versátiles tanto para a detección óptica como para a catálise en medios biolóxicos. Esta tese enmarca parte deste enfoque interdisciplinar, propoñendo novas estratexias para o deseño e modificación estrutural de péptidos co fin de mellorar as súas propiedades fluorescentes e catalíticas.

O traballo divídese en tres capítulos principais que abordan diferentes aplicacións das plataformas peptídicas: o desenvolvemento de sistemas autofluorescentes baseados en péptidos, a enxeñaría racional de catalizadores peptídicos artificiais con actividade intracelular e o uso de bibliotecas combinatorias para seleccionar secuencias funcionais de metalopeptídicos catalíticos.

Emisión anómala de péptidos α-helicoidais simples en solución

Ademais das proteínas fluorescentes verdes (GFPs) e das proteínas que incorporan cromóforos extrínsecos, é xeralmente aceptado que a absorción UV/vis e a fluorescencia das proteínas orixinan da presenza de residuos aromáticos na súa secuencia. Non obstante, ademais da fluorescencia convencional dos cromóforos intrínsecos, investigadores informaron de fluorescencia en péptidos sen residuos aromáticos nin grupos prostéticos, como solucións concentradas de polilisina ou ensamblaxes amiloides. A pesar do interese fundamental e das posibles aplicacións desta luminescencia anómala, coñecida como fluorescencia non aromática (NAF, polas súas siglas en inglés), a nosa comprensión deste fenómeno aínda é imperfecta e non existen sistemas modelo nin regras de deseño xerais para obter luminóxenos peptídicos non aromáticos.

Partindo deste concepto, propuxemos que os motivos α-helicoidais simples (SAH), formados por repeticións de catro residuos aniónicos (Glu) e catro catiónicos (Lys/Arg), $[E_4(K/R)_4]_n$, que soportan unha conformación α-helicoidal estable na ausencia de interaccións terciarias adicionais mediante unha rede de pontes salinas, poderían constituír unha plataforma ideal para xerar péptidos fluorescentes sen necesidade de residuos aromáticos. Con base nesta idea, propoñemos sintetizar $(E_4K_4)_4$ como modelo para os nosos estudos. Tras análise espectroscópica, confirmamos que este péptido é activo en UV e luminescente en longitudes de onda próximas ao UV en solución (λ_{exc} = 320 nm; $\lambda_{em} \approx$ 440 nm).

Despois de observar a presenza de emisión NAF no péptido modelo SAH, (E4K4)4, preguntámonos se tal emisión se vería afectada polo contido helicoidal, polo que sintetizamos un péptido SAH máis longo, $(E_4K_4)_9$, e un péptido zwitteriónico de control, $(E_2K_2)_8$, que ten a mesma composición zwitteriónica que as secuencias SAH modelo, pero disposta nun patrón diferente que se sabe que adopta unha configuración de liña aleatoria. Logo, medimos as súas Matrices de Excitación-Emisión (EEM) nas mesmas condicións e confirmamos que $(E_4K_4)_9$ exhibiu unha banda ampla e intensa centrada en aproximadamente 414 nm ao excitar a 320 nm, mentres que o péptido de liña aleatoria $(E_2K_2)_8$ apenas mostrou

emisión fotoluminescente. Isto demostrou que a estrutura helicoidal é crítica para xerar emisión NAF.

Ademais, avaliouse o efecto de diferentes modificacións, como a substitución de lisina por ornitina, $(E_4O_4)_4$; esta secuencia mantivo o plegado pero produciu un lixeiro aumento na emisión. Máis notable foi a variante de arxinina, $(E_4R_4)_4$, que mostrou un desprazamento ao vermello significativo e un aumento na intensidade de emisión. Ademais, grazas ás melloras na emisión, foi posible observar este péptido en células HeLa mediante microscopía confocal, destacando o seu potencial como sonda bioóptica. Ademais, a incorporación de residuos aromáticos, como Tyr, tamén mellorou as propiedades destas bandas anómalas.

Confirmamos que a fluorescencia era dependente da estrutura con experimentos de control adicionais, xa que desapareceu completamente tras a desnaturalización térmica da hélice, ao modificar o pH fóra do rango fisiolóxico ou cando se trataron os péptidos con axentes desestabilizadores da estrutura como o guanidinio. Tamén se observou unha diminución na fluorescencia en solucións de D_2O, suxerindo a implicación de pontes de hidróxeno no mecanismo de excitación. Do mesmo xeito, a intensidade de emisión mostrou unha resposta lineal coa concentración, suxerindo fortemente que a emisión non se debe a procesos de agregación intermolecular.

Ademais, cando incubamos unha solución do péptido $(E_4K_4)_9$ con proteinasa K, observamos unha diminución na intensidade de emisión co paso do tempo, xa que o péptido foi progresivamente dixerido pola encima. Este resultado apoia fortemente tanto a natureza anómala como o orixe peptídico da emisión, xa que calquera fluoróforo orgánico presente na mostra como impureza sería improbable que se vería afectado pola proteasa.

Para comprender mellor o orixe electrónico da NAF, realizáronse simulacións mecánico-cuánticas en modelos simplificados destes péptidos. Analizáronse as conformacións dobradas e desorganizadas, e calcularon as enerxías dos orbitais HOMO e LUMO (colaboración co Prof. Manuel Melle-Franco, na Universidade de Aveiro). Os resultados suxeriron que, nas estruturas helicoidais, as cadeas laterais cargadas (NH_3^+ e COO^-) poderían estabilizar os estados excitados mediante acoplamentos espaciais coa columna vertebral do péptido, permitindo transicións electrónicas na gama UV-visible. Aínda que o modelo non proporciona unha explicación definitiva, apoia a idea de que a organización tridimensional, a distribución de cargas e as pontes de hidróxeno xogan un papel esencial na xeración da NAF.

Outro achado notable foi a capacidade destes péptidos NAF para actuar como fluoróforos tradicionais en procesos de transferencia de enerxía, por exemplo como antenas que median a luminescencia de ións de lantánidos. Ao acoplar unha unidade DOTA no extremo N-terminal, posibilitouse a coordinación de ións de lantánidos como Tb^{3+}, Eu^{3+}, Sm^{3+} e Dy^{3+}. Ao excitar o péptido $(E_4K_4)_4$, observouse a emisión característica destes metais, mentres que para o péptido de liña aleatoria non se detectou tal emisión, demostrando unha transferencia de enerxía eficaz e dependente da helicidade. Estes resultados abren a porta a aplicacións en sensores ópticos, bioimaxe e microscopía avanzada.

Finalmente, implementáronse estratexias de macrocilización para aumentar a estabilidade estrutural do sistema. Xeráronse variantes cíclicas de $(E_4K_4)_9$ e dun análogo con residuos aromáticos cada catro posicións $(E_3YK_3Y)_9$ mediante síntese en fase sólida e macrocilización sobre resina. Estas versións cíclicas non só mantiveron a fluorescencia NAF senón que tamén expandiron a xanela espectral de emisión ata os 450 nm. Isto suxire que o ambiente tridimensional restrinxido favorece a estabilización dos estados excitados responsables da emisión.

En conxunto, este capítulo demostra que é posible deseñar péptidos fluorescentes sen residuos aromáticos, se adoptan estruturas ordenadas que permiten interaccións espaciais entre grupos cargados. A fluorescencia non só depende do pregamento, senón que tamén pode ser modulada por mutacións e ampliada mediante estratexias non convencionais como a ciclación ou a funcionalización con quelantes metálicos. Ademais, ábrense novas perspectivas para o uso destes sistemas en bioimaxe, transferencia de enerxía, sensores ópticos e deseño de sondas fluorescentes funcionais.

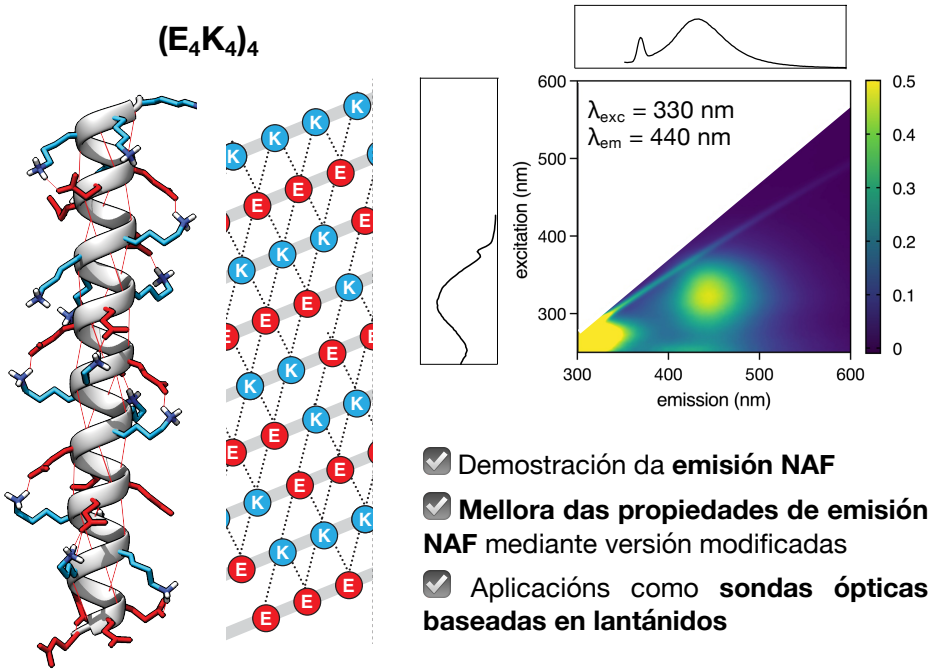


Figura 1. Representación gráfica do fenómeno de emisión anómala en péptidos de α-hélice simple en solución, representando a estrutura e o EEM de $(E_4K_4)_4$.

Enxeñaría *De Novo* de Pd-Metaloproteínas e o seu Uso como Catalizadores Intracelulares

O desenvolvemento de catalizadores artificiais que funcionen en ambientes biolóxicos complexos, como o interior das células vivas, representa un dos grandes desafíos da química biolóxica moderna. Aínda que os metais de transición demostraron unha eficiencia catalítica notable en medios orgánicos, a súa aplicación directa en sistemas biolóxicos adoita estar

limitada pola súa toxicidade, inestabilidade e rápida desactivación en presenza de especies redutoras ou nucleofílicas intracelulares.

Neste contexto, os complexos de Pd(II), amplamente utilizados na síntese orgánica, demostraron un gran potencial para realizar reaccións bioortogonais como a desprotección de grupos alquino, pero a súa efectividade en medios celulares é moi baixa debido á rápida desactivación do centro metálico. Este capítulo aborda estes desafíos mediante o deseño racional de novas arquitecturas peptídicas capaces de estabilizar centros metálicos e preservar a súa actividade catalítica no ambiente intracelular.

O traballo parte da hipótese de que estruturas peptídicas ben definidas, como as láminas β, poderían proporcionar un ambiente estruturado ideal para protexer o Pd(II) e, ao mesmo tempo, conferir á molécula características favorables para a súa internalización celular. Para este fin, utilizouse o dominio WW, unha miniproteína de aproximadamente 40 residuos composta por tres láminas β antiparalelas conectadas por dúas voltas, como andamiaxe. A partir da secuencia de referencia descrita por Macías *et al.* (PDB: 1E0M), deseñouse unha versión modificada (**WW0**), optimizada para a síntese química e compatible cos nosos estudos de coordinación metálica. Logo, deseñouse unha serie de variantes con residuos de His colocados estratexicamente na cara cóncava da lámina β, co obxectivo de crear sitios de coordinación bidentados para Pd(II), en rexións estruturalmente ríxidas e accesibles ao disolvente.

As secuencias peptídicas deseñadas foron sintetizadas mediante SPPS asistida por microondas e incubadas con $PdCl_2$(COD) para formar os complexos correspondentes. Observouse mediante HPLC-MS que só as secuencias con pares de His correctamente colocados deron lugar a especies coordinadas estables. De todas as variantes, o péptido **WW13/19**, que incorpora residuos de His nas posicións 13 e 19 da volta β1-β2, mostrou o mellor comportamento, mantendo a súa estrutura β tras a coordinación co metal, como se evidenciou nos espectros de dicroísmo circular. Este resultado non se replicou nos controis negativos, incluíndo **WW0** e **WW19** (péptido cun só His), que non modificaron a súa estrutura tras a adición do metal.

A análise de fluorescencia mostrou unha afinidade significativa entre **WW13/19** e Pd(II), cunha constante de disociación aparente de aproximadamente 12,7 μM, moito máis alta que a observada para a secuencia sen residuos de His (WW0). Para obter máis información estrutural sobre a estrutura secundaria das paladoproteínas sintéticas, foron estudadas por RMN ^{1}D a temperatura ambiente na ausencia e presenza de [Pd(COD)Cl_2]. O péptido **WW13/19** na ausencia de metal estaba maioritariamente desordenado, pero tras a adición de Pd(II) adquiriu unha conformación estruturada, con resonancias ben definidas tanto da estrutura principal do péptido como dos ligandos coordinados ao metal.

A continuación, avaliouse a capacidade catalítica do paladopeptido en medio acuoso mediante unha reacción modelo de desproparxilación dunha sonda fluoroxénica (2). As probas mostraron que WW13/19[Pd(II)] foi capaz de catalizar a reacción cunha eficiencia comparable á do [Pd(COD)Cl_2] libre (≈ 62 % de conversión), mesmo tras a eliminación do

exceso de Pd libre mediante ultradiafiltración. En contraste, os péptidos de control sen disposición bidentada (**WW0** e **WW19**) perderon completamente a súa actividade tras o mesmo tratamento, demostrando que a actividade catalítica está asociada á estrutura coordinada do péptido e non a trazas de paladio libre en disolución. Os estudos cinéticos confirmaron estes resultados, mostrando unha constante de reacción pseudo de primeira orde comparable entre o catalizador libre e **WW13/19**[Pd(II)].

Un dos aspectos máis destacables do sistema deseñado foi a súa sorprendente capacidade de internalización celular. Para avaliala, sintetizáronse versións marcadas cun fluoróforo TMR e analizouse a súa entrada en células HeLa mediante microscopía e citometría. Observouse que **WW13/19**[Pd(II)] entraba eficientemente nas células, cunha intensidade de sinal ata 6000 veces maior que a da súa versión non coordinada. Este fenómeno asociouse á rixidez conformacional inducida polo metal e á exposición de superficies cargadas dispostas favorablemente.

Unha vez confirmado que o complexo podía internalizarse en células vivas, probouse a súa actividade catalítica intracelular empregando de novo a sonda **1**. As células tratadas con **WW13/19**[Pd(II)] mostraron un sinal fluorescente intenso. En contraste, nin o Pd libre nin os péptidos de control foron capaces de xerar fluorescencia intracelular, reforzando a idea de que o complexo **WW13/19** actúa como unha protometaloproteína artificial activa e estable. Ademais, este comportamento foi replicado en células HepG2, suxerindo que o sistema é versátil e compatible con distintos tipos celulares.

Para explorar máis a aplicabilidade do sistema, probáronse outras sondas fluoroxénicas con diferentes estruturas, como unha baseada en resorufina (**4**) e outra en violeta cresílico (**6**). En ensaios *in vitro*, a actividade catalítica de WW13/19[Pd(II)] foi lixeiramente inferior á do Pd libre no caso da resorufina, pero significativamente superior ao fondo dos controis. No caso do violeta cresílico, a catálise *in vitro* resultou en rendementos moi baixos tanto para o complexo de paladio libre como para o noso sistema. Non obstante, a catálise en células vivas, tanto para a resorufina como mesmo para o violeta cresílico, xerou un sinal fluorescente apreciable cando se incubou con **WW13/19**[Pd(II)].

En conxunto, os resultados deste capítulo sitúan o complexo **WW13/19**[Pd(II)] como unha ferramenta versátil e eficaz para realizar transformacións químicas controladas no interior de células vivas. O seu deseño racional, baseado en dominios peptídicos pregados capaces de coordinar metais de forma específica e estable, permite combinar actividade catalítica, biocompatibilidade e capacidade de internalización nun só sistema. Esta estratexia inaugura unha nova clase de protometaloproteínas artificiais funcionais, expandindo as fronteiras da catálise peptídica en ambientes biolóxicos e abrindo novas oportunidades para o desenvolvemento de sistemas catalíticos programables con aplicacións en biotecnoloxía, terapias dirixidas e ferramentas moleculares avanzadas.

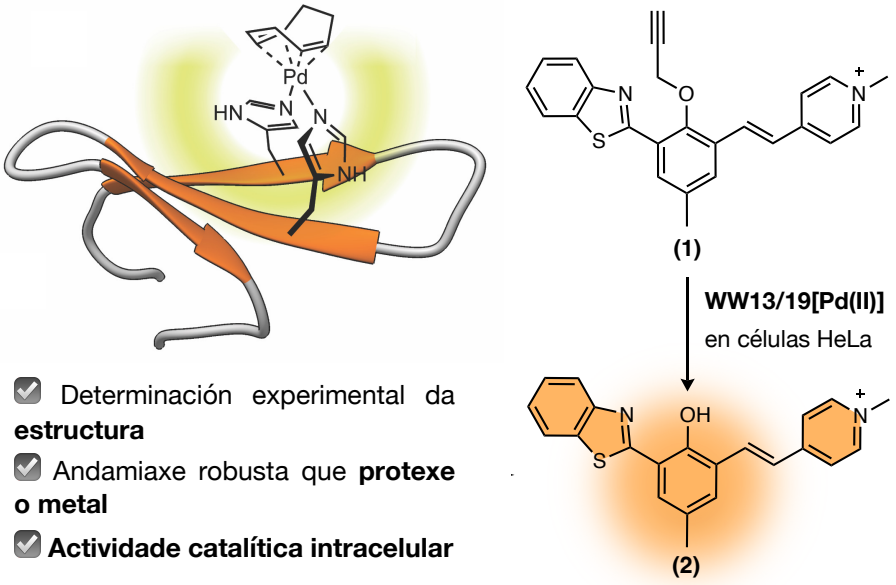


Figura 2. Representación gráfica da enxeñaría *de novo* de Pd-metaloproteínas e o seu uso como catalizadores intracelulares, amosando a estrutura e actividade catalítica intracelular do complexo **WW13/19**[Pd(II)].

Identificación de Metalopéptidos para a Catálise Intracelular Usando Bibliotecas Combinatorias

Neste capítulo retomamos a idea desenvolvida no capítulo dous, pero cun novo enfoque: seleccionar pequenos metalopéptidos que combinen o mellor dos complexos metálicos e das encimas artificiais, empregando métodos combinatorios no canto dun deseño racional. Para isto, recorremos á tecnoloxía SPOT, que permite sintetizar bibliotecas de péptidos en superficies planas e simplifica moito o cribado grazas ao uso de substratos fluoroxénicos.

Traballamos especificamente con estruturas do tipo β-hairpin debido á súa sinxeleza estrutural e á súa capacidade para organizar residuos quelantes, clave para axustar a reactividade do centro metálico.

Seleccionamos como base estrutural o zipper de triptófano (Wzip). Esta estrutura peptídica compacta, composta por 12 a 16 aminoácidos, está estabilizada polas interaccións entre cadeas laterais de triptófano. Demostrou ser ideal para a coordinación de metais, o que nos permitiu deseñar unha biblioteca combinatoria específica. Entre as posibles variantes, eliximos Wzip1 como punto de partida.

No deseño da biblioteca, mantivéronse os elementos esenciais de Wzip1 (o núcleo GN e os residuos de triptófano), e modificáronse as seis posicións restantes que apuntan á cara cóncava da hélice. Xeráronse combinacións con residuos de histidina en dúas posicións adxacentes e engadíronse aminoácidos como alanina, valina, treonina e arxinina nas outras catro. Isto resultou nunha biblioteca manexable de 264 péptidos.

Para avaliar a súa actividade catalítica, as placas foron incubadas primeiro co complexo de paladio [Pd(COD)Cl_2] e logo cun derivado proparxilado. A reacción foi monitorizada mediante microscopía confocal.

Os resultados do cribado amosaron patróns claros de preferencia por certos residuos en posicións específicas. As combinacións máis eficientes incluían histidinas nas posicións 3 e 10, valina na posición 5, e arxinina nas posicións 5 ou 8. Estes datos suxiren que estes aminoácidos desempeñan un papel importante na estabilidade estrutural ou na modulación do ambiente catalítico.

Entre todas as secuencias candidatas, seleccionouse o péptido **D14** ($A^1W\mathbf{H}^3WR^5GNVW\mathbf{H}^{10}WT$) para estudos posteriores. A variante **D4** ($R^1W\mathbf{H}^3WA^5GNVW\mathbf{H}^{10}WT$), estruturalmente similar pero con dous residuos intercambiados, foi escollida como control negativo. Segundo os resultados do cribado SPOT, esta variante reduciu significativamente a súa actividade catalítica.

Realizáronse ensaios de espectrometría de masas, fluorescencia e dicroscopía circular para avaliar a capacidade de coordinación. O péptido **D14** mostrou unha afinidade forte e rápida polo Pd(II), cunha constante de disociación no rango micromolar, formando un complexo ben definido capaz de manter unha estrutura secundaria compacta en solución. En contraste, **D4** non foi capaz de formar un complexo estable.

A actividade catalítica destes complexos foi posteriormente avaliada en condicións acuosas. **D14** mostrou unha actividade moderada en PBS, pero o máis notable foi que, ao transferirse a un medio biolóxico máis complexo como DMEM, onde o Pd libre perde efectividade, o complexo **D14**[Pd(II)] mantivo a súa actividade case intacta. Isto demostra o efecto protector do péptido, que estabiliza o centro catalítico fronte ao ambiente celular.

As probas en células vivas reforzaron estes achados. Cando se incubaron células HeLa co complexo **D14**[Pd(II)] e o derivado proparxilado, detectouse unha desprotección eficiente, evidenciada pola formación dun produto fluorescente. En contraste, nin o complexo libre de Pd(II) nin o péptido **D4** mostraron actividade detectable. Ademais, descubriuse que **D14** tamén pode funcionar con outros precursores metálicos como [Pd(bpy)Cl_2], demostrando a súa versatilidade como ligando adaptable.

Para explorar outras configuracións, estudáronse secuencias con His nas posicións 10/12 e 3/12. Péptidos **F14** ($A^1WR^3WV^5GNTW\mathbf{H}^{10}W\mathbf{H}$) e **E10** ($T^1W\mathbf{H}^3WA^5GNVWR^{10}W\mathbf{H}$) tamén demostraron boa actividade catalítica, tanto *in vitro* como en células vivas, validando a eficacia do enfoque SPOT como ferramenta de cribado.

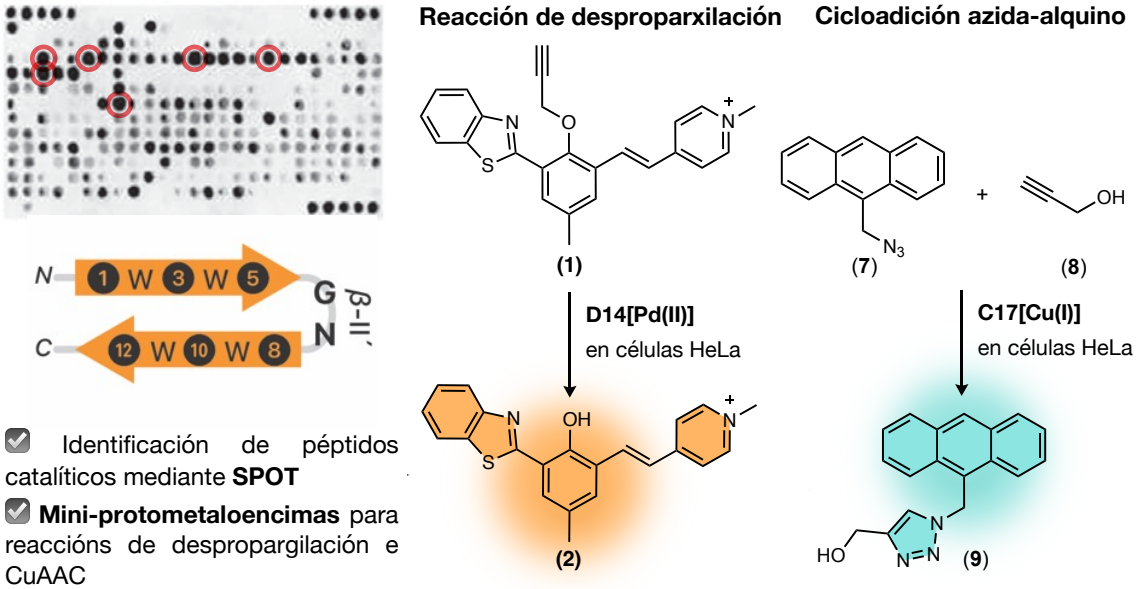
Na segunda parte do estudo, aplicouse o mesmo enfoque para identificar péptidos capaces de estabilizar ións Cu(I) e promover a reacción de clic azida-alquino (CuAAC), esencial na química bioortogonal. Deseñouse unha nova biblioteca, tamén baseada en Wzip1, con tres residuos de histidina ou dúas histidinas e un residuo ácido (Asp ou Glu), xunto con outros cinco aminoácidos para definir o ambiente catalítico. O cribado realizouse cunha sonda de

azida fluoroxénica que emite fluorescencia ao formar o triazol, permitindo detectar a actividade catalítica baixo condicións fisiolóxicas.

Os resultados destacaron a **C17** (Y[1]W**H**[3]WV[5]GNRW**H**[10]W**H**) como a secuencia máis activa, mentres que **C10** (A[1]W**H**[3]WY[5]GNVW**H**[10]W**H**), a pesar de ser estruturalmente similar, mostrou unha actividade moi limitada. Os ensaios de coordinación confirmaron que **C17** forma complexos de cobre de maneira eficiente, cunha afinidade moderada, axeitada para aplicacións biolóxicas. En medios acuosos, o complexo **C17**[Cu(I)] mostrou unha actividade catalítica significativamente maior que **C10**.

Finalmente, avaliouse a actividade intracelular destes novos sistemas. Cando as células HeLa foron tratadas con reactivos típicos para a reacción CuAAC, observouse unha intensa fluorescencia distribuída por todo o interior celular, seguindo a acción do complexo **C17**[Cu(I)]. En contraste, o complexo **C10** e o cobre libre mostraron pouca ou ningunha actividade, reforzando o papel do péptido como ligando estabilizador e facilitador do proceso catalítico.

En resumo, este capítulo demostra por primeira vez o potencial das bibliotecas SPOT para identificar sistematicamente metalopéptidos con actividade catalítica en condicións biolóxicas. A combinación do deseño estrutural racional con tecnoloxías de cribado eficaces e a validación en medios celulares permitiu illar complexos como **D14**[Pd(II)] e **C17**[Cu(I)], que, aínda que están lonxe de igualar a eficiencia das encimas naturais, representan modelos prometedores de protometaloenzimas funcionais en ambientes celulares. Estas plataformas poderían abrir novas vías para a manipulación química de procesos biolóxicos con aplicacións terapéuticas, como activación dirixida de fármacos, edición metabólica ou enxeñaría de sinais intracelulares.



- ☑ Identificación de péptidos catalíticos mediante **SPOT**
- ☑ **Mini-protometaloencimas** para reaccións de despropargilación e CuAAC

Figura 3. Representación gráfica da identificación de metalopeptidos para catálise intracelular mediante bibliotecas combinatorias, incluíndo a biblioteca SPOT de péptidos Wzip1 e as reaccións de desproparxilación e clic azida–alquino.

2. SUMMARY

The field of chemical biology has evolved significantly over the past decades, integrating tools from organic chemistry, biochemistry and nanotechnology to address complex biological questions by designing functional molecules. In this context, peptides have emerged as versatile platforms for both optical sensing and catalysis in biological media. This thesis is part of this interdisciplinary approach, proposing new strategies for the design and structural modification of peptides to enhance their fluorescent and catalytic properties.

This work was divided into three main chapters focusing on different application of peptide platforms: the development of peptide-based autofluorescent systems, the rational engineering of artificial peptide catalysts with intracellular activity, and the use of combinatorial libraries to select functional catalytic metallopeptide sequences.

Anomalous Emission from Single α-Helical Peptides in Solution

Besides GFPs and proteins incorporating extrinsic chromophores, it is generally accepted that the UV/vis absorption and fluorescence of proteins originates from the presence of aromatic residues in their sequence. However, in addition to the conventional fluorescence from intrinsic chromophores, researchers have reported fluorescence from peptides lacking aromatic residues or prosthetic groups, such as concentrated solutions of poly-lysine, or amyloid assemblies. Despite the fundamental interest and potential applications of this anomalous luminescence, commonly known as Non-Aromatic Fluorescence (NAF), our understanding of this phenomenon is still imperfect, and no model system or general design rules to obtain non-aromatic peptide luminogens are available.

Based on this concept, we proposed that single α-helix α-helix (SAH) motifs, formed by repeats of four anionic (Glu) and four cationic (Lys/Arg) residues, $[E_4(K/R)_4]_n$, that support a stable α-helical conformation in the absence of additional tertiary interactions through a network of salt bridges, could constitute an ideal platform for generating fluorescent peptides without the need for aromatic residues. Based on this idea, we propose to synthesize $(E_4K_4)_4$ as a model for our studies. After spectroscopic analysis, we confirmed that this peptide is UV active and luminescent at near-UV wavelengths in solution (λ_{exc} = 320 nm; $\lambda_{em} \approx$ 440 nm)

Having observed the presence of NAF emission from the model SAH peptide, $(E_4K_4)_4$, we wondered whether such emission would be affected by the helical content, and therefore we synthesized a longer SAH peptide, $(E_4K_4)_9$, and a control zwitterionic peptide, $(E_2K_2)_8$, which has the same zwitterionic composition than the model SAH sequences, but arranged in an different pattern that is known to adopt a random coil configuration. We then measured their Excitation-Emission Matrices (EEM) under the same conditions and confirmed that $(E_4K_4)_9$ displayed a broad and intense band centered at c.a. 414 nm upon excitation at 320 nm, while the random coil peptide $(E_2K_2)_8$ barely showed any photoluminescence emission. This demonstrated that the helical structure is critical for generating NAF emission.

Furthermore, the effect of different modifications was evaluated, such as the substitution of lysine for ornithine, $(E_4O_4)_4$, this sequence maintained the folding but produced a slight increase in emission. More notable was the arginine variant, $(E_4R_4)_4$, which showed a

significant red shift and increased emission intensity. Furthermore, thanks to the improvements in emission, it was possible to observe this peptide in HeLa cells by confocal microscopy, highlighting its potential as a biooptical probe. In addition, the incorporation of aromatic residues, such as Tyr, also improved the properties of these anomalous bands.

We confirmed that the fluorescence was structure-dependent with additional control experiments, as it disappeared completely after thermal denaturation of the helix, upon modifying the pH outside the physiological range, or when treating the peptides with structure-destabilizing agents such as guanidinium. A decrease in fluorescence was also observed in D_2O solutions, suggesting the involvement of hydrogen bonds in the excitation mechanism. Likewise, the emission intensity showed a linear response with concentration, strongly suggesting that the emission is not due to intermolecular aggregation processes.

Furthermore, when we incubated a solution of the peptide $(E_4K_4)_9$ with proteinase K, we observed a decrease in the emission intensity over time, as the peptide was progressively digested by the enzyme. This result strongly supports both the anomalous nature and the peptide origin of the emission, since any organic fluorophore present in the sample as an impurity would be unlikely to be affected by the protease.

To better understand the electronic origin of NAF, quantum mechanical simulations were performed on simplified models of these peptides. The folded and disorganized conformations were analyzed, and the energies of the HOMO and LUMO orbitals were calculated (collaboration with Prof. Manuel Melle-Franco, at the University of Aveiro). The results suggested that, in helical structures, charged side chains (NH_3^+ and COO^-) could stabilize excited states through spatial couplings with the peptide backbone, allowing electronic transitions in the UV-visible range. Although the model does not provide a definitive explanation, it does support the idea that three-dimensional organization, charge distribution and hydrogen bonds play an essential role in the generation of NAF.

Another notable finding was the ability of these NAF peptides to act as traditional fluorophores in energy transfer processes, such as antennas that mediate the luminescence of lanthanide ions. By coupling a DOTA unit at the N-terminal end, coordination of lanthanide ions such as Tb^{3+}, Eu^{3+}, Sm^{3+} and Dy^{3+} was enabled. Upon excitation of the $(E_4K_4)_4$ peptide, the characteristic emission of these metals was observed, while for the random coil peptide it was not, demonstrating an effective and helicity-dependent energy transfer. These results open the door to applications in optical sensors, bioimaging and advanced microscopy.

Finally, macrocyclization strategies were implemented to increase the structural stability of the system. Cyclic variants of $(E_4K_4)_9$ and an analogue with aromatic residues every four positions $(E_3YK_3Y)_9$ were generated through solid-phase synthesis and on-resin cyclization. These cyclic versions not only retained NAF fluorescence but also expanded the emission spectral window to 450 nm. This suggests that the restricted three-dimensional environment favors the stabilization of the excited states responsible for emission.

Taken together, this chapter demonstrates that it is possible to design fluorescent peptides without aromatic residues, if they adopt ordered structures that allow spatial interactions

between charged groups. Fluorescence not only depends on folding but can also be modulated by mutations and extended by non-conventional strategies such as cyclization or functionalization with metal chelators. In addition, new perspectives are opened for the use of these systems in bioimaging, energy transfer, optical sensors and design of functional fluorescent probes.

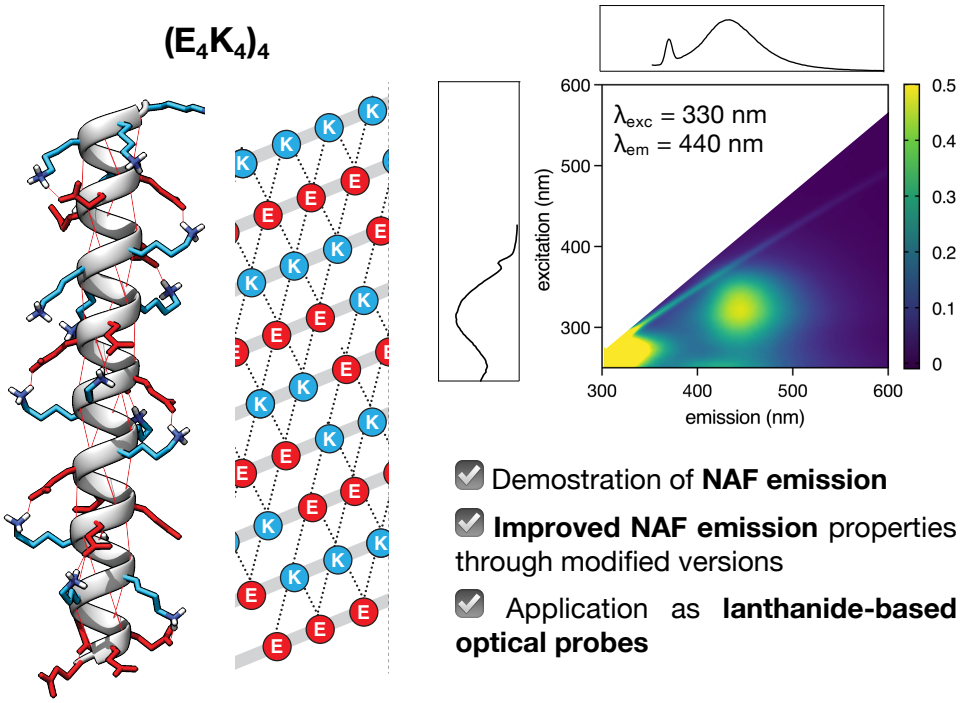


Figure 1. Summary figure of anomalous emission from single α-helical peptides in solution, the structure and EEM of $(E_4K_4)_4$ are represented.

De Novo Engineering of Pd-Metalloproteins and Their Use as Intracellular Catalysts

The development of artificial catalysts that function in complex biological environments, such as inside living cells, represents one of the great challenges of modern biological chemistry. While transition metals have demonstrated remarkable catalytic efficiency in organic media, their direct application in biological systems is often limited by their toxicity, instability and rapid deactivation in the presence of intracellular reductive or nucleophilic species.

In this context, Pd(II) complexes, widely used in organic synthesis, have shown great potential to carry out bioorthogonal reactions such as the deprotection of alkyne groups, but their effectiveness in cellular media is very low due to the rapid deactivation of the metal center. This chapter addresses these challenges through the rational design of new peptide architectures capable of stabilizing metal centers and preserving their catalytic activity in the intracellular environment.

The work starts from the hypothesis that well-defined peptide structures, such as β-sheets, could provide an ideal structured environment to protect Pd(II) and, at the same time, confer

to the molecule favorable characteristics for its cellular internalization. For this purpose, the WW domain, a miniprotein of approximately 40 residues composed of three antiparallel β-sheets connected by two turns, was used as a scaffold. From the reference sequence described by Macías *et al.* (PDB: 1E0M), we designed a modified version (**WW0**), optimized for chemical synthesis and compatible with our studies of metal coordination. Then, a series of variants were designed with strategically positioned His residues on the concave face of the β-sheet, with the aim of creating bidentate coordination sites for Pd(II), in structurally rigid and solvently accessible regions.

The designed peptide sequences were synthesized by microwave-assisted SPPS and incubated with $PdCl_2$(COD) to form the corresponding complexes. It was observed by HPLC-MS that only sequences with correctly positioned His pairs gave rise to stable coordinated species. Of all the variants, peptide **WW13/19**, which incorporates His residues at positions 13 and 19 of the β1-β2 turn, showed the best behavior, retaining its β-structure after coordination with metal, as evidenced by circular dichroism spectra. This result was not replicated in the negative controls, including **WW0** and **WW19** (peptide with a single His), which did not modify their structure after addition of the metal.

Fluorescence analysis showed significant affinity between **WW13/19** and Pd(II), with an apparent dissociation constant of approximately 12.7 μM, much higher than that observed for the sequence without His residues (**WW0**). To obtain further structural information on the secondary structure of the synthetic paladoproteins, they were studied by 1D NMR at room temperature in the absence and in the presence of [Pd(COD)Cl_2]. The **WW13/19** peptide in the absence of metal was mostly disordered, but after the addition of Pd(II) it acquired a structured conformation, with well-defined resonances of both the main structure of the peptide and the ligands coordinated to the metal.

Next, the catalytic ability of the paladopeptide was evaluated in aqueous medium by a model reaction of depropargylation of a fluorogenic probe (**2**). Tests showed that **WW13/19**[Pd(II)] was able to catalyze the reaction with an efficiency comparable to that of free [Pd(COD)Cl_2] (≈ 62 % conversion), even after removal of excess free Pd by ultradiafiltration. In contrast, the control peptides without bidentate arrangement (**WW0** and **WW19**) completely lost their activity after the same treatment, demonstrating that the catalytic activity is associated with the coordinated structure of the peptide and not with traces of free palladium in solution. Kinetic studies confirmed these results, showing a comparable pseudo-first order reaction constant between the free catalyst and **WW13/19**[Pd(II)].

One of the most remarkable aspects of the designed system was its surprising capacity for cell internalization. To evaluate it, TMR fluorophore-labeled versions were synthesized and their entry into HeLa cells was analyzed by microscopy and cytometry. **WW13/19**[Pd(II)] was found to efficiently enter the cells, with a signal intensity up to 6000-fold higher than that of its uncoordinated version. This phenomenon was associated with metal-induced conformational rigidity and exposure to favorably disposed charged surfaces.

Once it was confirmed that the complex could be internalized in living cells, its intracellular catalytic activity was tested using probe **1** again. Cells treated with **WW13/19**[Pd(II)] showed an intense fluorescent signal. In contrast, neither free Pd nor control peptides were able to generate intracellular fluorescence, reinforcing the idea that the **WW13/19** complex acts as an active and stable artificial proto-metalloprotein. Moreover, this behavior was replicated in HepG2 cells, suggesting that the system is versatile and compatible with different cell types.

To further explore the applicability of the system, other fluorogenic probes with different structures were tested, such as one based on resorufin (**4**) and another on cresyl violet (**6**). In *in vitro* assays, the catalytic activity of **WW13/19**[Pd(II)] was slightly lower than that of free Pd in the case of resorufin, but significantly higher than the background of the controls. In the case of cresyl violet, *in vitro* catalysis resulted in very low yields for both the free palladium complex and our system. However, catalysis in live cells, both for resorufin and even for cresyl violet, generated an appreciable fluorescent signal when incubated with **WW13/19**[Pd(II)].

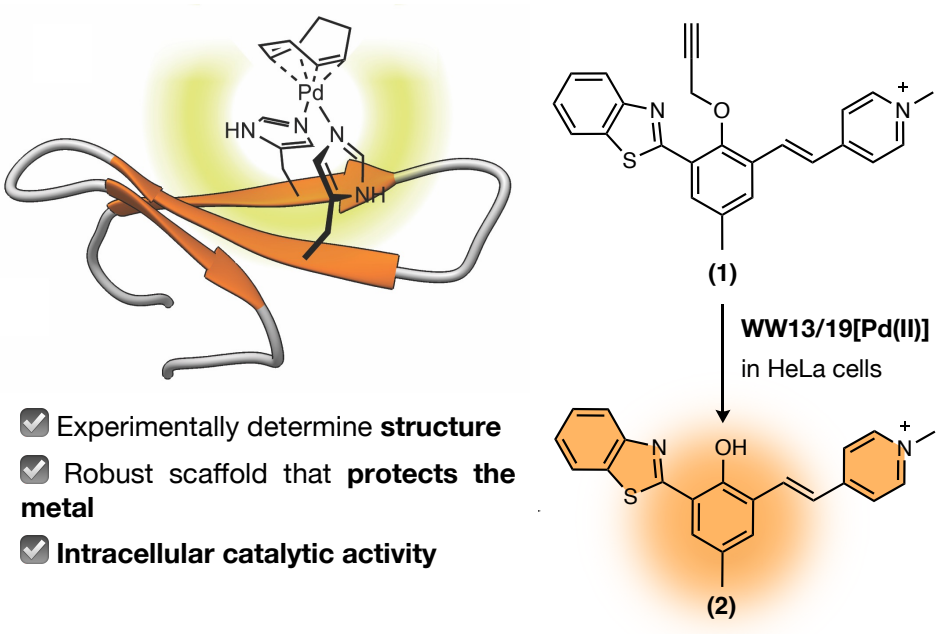


Figure 2. Summary figure of de novo engineering of Pd-metalloproteins and their use as intracellular catalysts, the structure and intracellular catalytic activity of the WW13/19[Pd(II)] complex are represented.

Taken together, the results of this chapter position the **WW13/19**[Pd(II)] complex as a versatile and effective tool for performing controlled chemical transformations inside living cells. Its rational design, based on folded peptide domains capable of coordinating metals in a specific and stable manner, allows combining catalytic activity, biocompatibility, and internalization capacity in a single system. This strategy inaugurates a new class of functional artificial proto-metalloproteins, expanding the frontiers of peptide catalysis in biological environments and opening new opportunities for the development of programmable catalytic systems with applications in biotechnology, targeted therapies, and advanced molecular tools.

Identification of Metallopeptides for Intracellular Catalysis Using Combinatorial Libraries

In this chapter, we revisit the idea developed in chapter two, but with a new approach: selecting small metallopeptides that combine the best of metal complexes and artificial enzymes using combinatorial methods instead of a rational approach. For this, we turned to SPOT technology, which allows peptide libraries to be synthesized on flat surfaces and greatly simplifies screening thanks to the use of fluorogenic substrates. Specifically, we work with β-hairpin scaffolds due to their structural simplicity and their ability to organize chelating residues, which is key to fine-tuning the reactivity of the metal center.

Specifically, we selected the tryptophan zipper (Wzip) as our structural base. This compact peptide structure consisting of 12 to 16 amino acids is stabilized by interactions between tryptophan side chains. This structure has proven ideal for coordinating metals, allowing us to design a specific combinatorial library. Among the possible variants, we chose Wzip1 as our starting point.

In designing the library, we maintained the essential elements of Wzip1, the GN core and the tryptophan residues, and modified the remaining six positions facing the concave face of the helix. Combinations were generated with histidine residues at two adjacent positions, and amino acids such as alanine, valine, threonine, and arginine were introduced at the other four sites. This resulted in a manageable library of 264 peptides. To evaluate their catalytic activity, the plates were first incubated with the palladium complex $Pd(COD)Cl_2$ and then with a propargylated derivative. The reaction was monitored and activity assessed using confocal microscopy.

The screening results showed clear patterns of preference for certain residues at specific positions. The most efficient combinations were those with histidines at positions 3 and 10, valine at position 5, and arginine at positions 5 or 8. These data suggest that these amino acids play an important role in structural stability or in modulating the catalytic environment. Of all the candidate sequences, the peptides **D14** ($A^1W\mathbf{H}^3WR^5GNVW\mathbf{H}^{10}WT$) was selected for further study. **D4** ($R^1W\mathbf{H}^3WA^5GNVW\mathbf{H}^{10}WT$), a similar variant but with two swapped residues, was chosen as a negative control. This variant, according to the SPOT screening results, significantly reduced its catalytic activity.

Mass spectrometry, fluorescence, and circular dichroism assays were performed to assess coordination capacity. Peptide **D14** displayed a strong and rapid affinity for Pd(II), with a dissociation constant in the micromolar range, forming a well-defined complex capable of maintaining a compact secondary structure in solution. In contrast, **D4** failed to form a stable complex.

The catalytic activity of these complexes was subsequently evaluated under aqueous conditions. **D14** showed moderate activity in PBS, but what was most notable was that when transferred to a more complex biological medium such as DMEM, where free Pd loses effectiveness, the **D14**[Pd(II)] complex maintained its activity almost intact. This demonstrates

the protective effect of the peptide, which stabilizes the catalytic center against the cellular environment.

Tests in live cells reinforced these findings. When HeLa cells were incubated with the **D14**[Pd(II)] complex and the propargylated derivative, efficient deprotection was detected, evidenced by the formation of the fluorescent product. In contrast, neither the free Pd(II) complex nor peptide **D4** showed detectable activity. Furthermore, it was found that **D14** could also work with other metal precursors such as $Pd(bpy)Cl_2$, demonstrating its versatility as an adaptable ligand.

To explore other configurations, sequences with His at positions 10/12 and 3/12 were studied. Peptides such as **F14** ($A^{1}WR^{3}WV^{5}GNTW\mathbf{H}^{10}W\mathbf{H}$) and **E10** ($T^{1}W\mathbf{H}^{3}WA^{5}GNVWR^{10}W\mathbf{H}$) also demonstrated good catalytic activity, both in vitro and in living cells, validating the effectiveness of the SPOT approach as a screening tool.

In a second part of the study, the same approach was applied to identify peptides capable of stabilizing Cu(I) ions and promoting the azide-alkyne click reaction (CuAAC), essential in bioorthogonal chemistry. A new library was designed, also based on Wzip1, with three histidine residues or two histidines and one acidic residue (Asp or Glu), along with five additional amino acids to define the catalytic environment. Screening was performed with a fluorogenic azide probe that becomes fluorescent upon forming the triazole, allowing the catalytic activity of the different peptides to be detected under physiological conditions.

The results highlighted **C17** ($Y^{1}W\mathbf{H}^{3}WV^{5}GNRW\mathbf{H}^{10}W\mathbf{H}$) as the most active sequence, while **C10** ($A^{1}W\mathbf{H}^{3}WY^{5}GNVW\mathbf{H}^{10}W\mathbf{H}$), despite being structurally similar, showed very limited activity. Coordination assays confirmed that **C17** forms copper complexes efficiently, with moderate affinity, suitable for biological applications. In tests in aqueous media, the **C17**[Cu(I)] complex showed significantly higher catalytic activity than **C10**.

Finally, the intracellular activity of these new systems was evaluated. When HeLa cells were treated with typical reagents for the CuAAC reaction, intense fluorescence was observed, distributed throughout the cell interior, following the action of the **C17**[Cu(I)] complex. In contrast, the **C10** complex and Cu(I) showed little catalysis, reinforcing the role of the peptide as a stabilizing ligand and facilitator of the catalytic process.

Overall, this chapter demonstrates for the first time the potential of SPOT libraries to systematically identify metallopeptides with catalytic activity under biological conditions. The combination of rational structural design with efficient screening technologies and validation in cellular media has allowed the isolation of complexes such as **D14**[Pd(II)] and **C17**[Cu(I)], which, while still far from matching the efficiency of natural enzymes, represent promising models of functional Cu(I) metalloenzymes in cellular environments. These platforms could open new avenues for the chemical manipulation of biological processes with therapeutic applications, such as targeted drug activation, metabolic editing, or intracellular signal engineering.

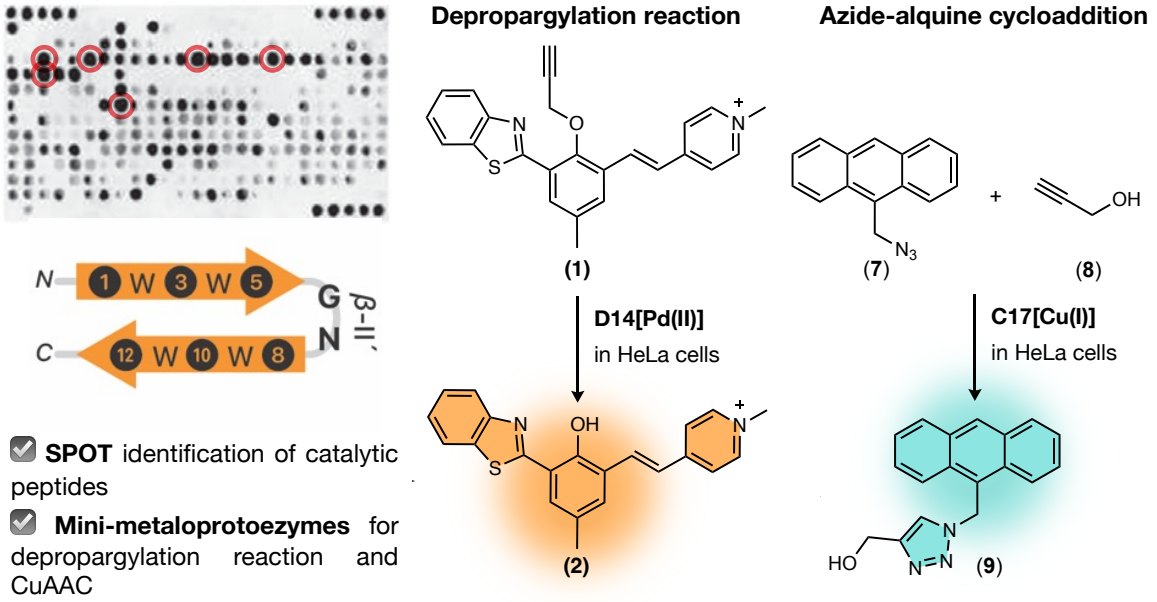


Figure 3. Summary figure of identifying metallopeptides for intracellular catalysis using combinatorial libraries, SPOT library of Wzip1 peptides and the despropargylation and azide-alquine cycloaddition reactions are represented.

3. INTRODUCTION

3.1. General Context of Chemical Biology

Chemical biology is a scientific discipline that studies the relationship between chemistry and biology, attempting to describe biological phenomena at the molecular level using chemical tools. Its main goal is to understand biological processes through chemical techniques, giving rise to innovative approaches to solve complex biological problems. [1] Given its multidisciplinary nature, delimiting the scope of chemical biology is not straightforward, as it encompasses a wide variety of areas, such as enzymatic catalysis, molecular recognition, supramolecular and synthetic chemistry, cell signaling, internalization of molecules and spectroscopy, among others. As a result, the definition of chemical biology has been the subject of debate since the term was coined in the mid-20th century.

The first relevant advances in this field date back to the 19th century. Pioneers such as the German chemist Friedrich Wöhler achieved a milestone in 1828 when he synthesized urea for the first time, demonstrating that biological compounds could originate from inorganic chemical materials.[2] In the late 19th century, the Swiss biochemist Friedrich Miescher used chemical compounds to isolate and analyze cell nuclei.[3] Although he was looking for proteins, he discovered a phosphorus-rich substance in white blood cells, which he called *nuclein*. Later, it began to be called nucleic acid, and today we know that these compounds make up DNA, the basis of the cell's genetic information.

During this period, cell staining agents were also developed.[4] These were fundamental to the studies of Nobel laureate Paul Ehrlich, who pioneered the field of medicinal chemistry with the development of *Salvarsan*, an arsenic-based drug to treat syphilis that greatly improved the previous mercury-based treatment, which caused serious adverse effects such as tooth darkening and weakening, kidney failure, and even death due to poisoning. [5]

In the second half of the 20th century, the relationship between chemistry and biology was reinforced, giving rise to the term "*chemical biology*" in 1945. This concept was coined by Linus Pauling and George W. Beadle in a grant proposal to the *Rockefeller Foundation*, which underlined the importance of chemistry in addressing "*the great problems of biology*".[6] However, it was not until 1954 that the first formal definition of the discipline was offered in the journal *Engineering and Science* (**Figure 4**).[7] Despite this first step, the term did not become popular until the 1990s.

---

[1] Ostler, E. L. *Chem. Cent. J.* **1**, 5 (2007).

[2] Wöhler, F. *Ann. Phys.* **88**, 253–256 (1828).

[3] Miescher, F. *Medicinisch-chemische Untersuchungen*, 4º edition (1871)

[4] Musumeci, G. *J. Histol. Histopathol.* **1**, 5 (2014).

[5] Drews, J. *Nat. Rev. Drug Disco.* **3**, 797–801 (2004).

[6] L. E. Kay, The Molecular Vision of Life. Caltech, *The Rockefeller Foundation, and the Rise of the New Biology*, Oxford Univ. Press ISBN: 9780195111439 (pp. 225-242) and *@CaltechArchives* (1997).

[7] *Eng. & Sci.* **17**, 9-13 (1954).

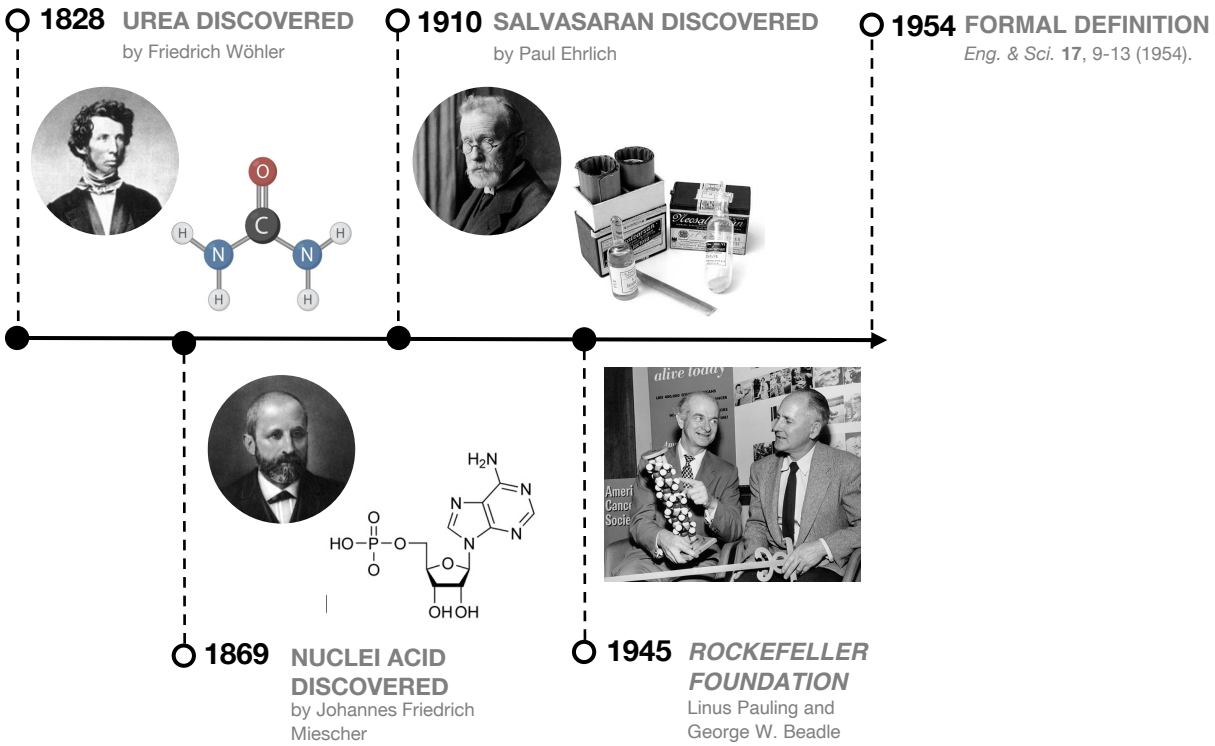


Figure 4. A timeline showing the main milestones in chemical biology, from 1828 (discovery of urea), 1869 (discovery of nucleic acids), 1910 (start of medicinal chemistry), 1945 (coining of the term) to 1954 (formal definition).

In the 21st century, chemical biology has established itself as a key research field at the intersection between chemistry and biology, driving significant advances. Today, many biological challenges can be addressed from a molecular perspective, thanks to the development of chemical sciences, which have provided tools to synthesize, modify and analyze increasingly complex systems.

The impact of this discipline has been recognized with multiple Nobel prizes over the years. In 2018, Frances H. Arnold, George P. Smith and Gregory P. Winter were awarded for their studies on the directed evolution of enzymes. Two years later, in 2020, Emmanuelle Charpentier and Jennifer A. Doudna received the award for their groundbreaking work on gene editing with CRISPR. More recently in 2022, Carolyn R. Bertozzi, Morten Meldal and K. Barry Sharpless were recognized for their contributions to the development of click and bioorthogonal chemistry.

Peptides and proteins play a crucial role in chemical biology, not only because of their structural and catalytic functions indispensable for life, but also because their design, modification and analysis have been central to important advances in biomedicine, biotechnology and materials chemistry. These advances have established chemical biology as a dynamic and evolving field.

3.2. Peptides and Proteins: Structure, Properties and Function

3.2.1. Historical Background

Proteins are large and complex molecules that play essential roles in the cell, acting as structural components, motors and catalysts for biochemical reactions. They consist of one or more chains of amino acids, called polypeptides, whose sequence is determined by the DNA of the gene that encodes them. Their functional diversity is due to the wide variety of three-dimensional structures they can adopt, as each protein has a specific conformation based on its amino acid sequence.

The history of proteins dates to the 18th century, when the Italian scientist Iacopo Beccari discovered a protein component in wheat flour, known today as gluten.[8] In the 19th century, the study of proteins advanced significantly with the observation that heating them in concentrated acid allowed the isolation and study of amino acids, the fundamental units of proteins. An important milestone occurred in 1820, when Henri Braconnot prepared glycine from gelatine to test whether proteins behaved, like starch, and were broken down by acids to produce sugars.[9] As a result, the importance of these molecules began to be recognized by chemists of the time, including Mulder and Berzelius, who in 1838 coined the term *protein*, derived from the Greek *prōteios*, meaning "*primary substance*".[10]

Although the basic chemical structure of proteins had been known for some time, it wasn't until 1951 that the main elements of protein secondary structure were proposed. Linus Pauling, Robert Corey, and Herman Branson described the α-helix and β-sheet to explain the X-ray diffraction patterns of fibrous and crystalline proteins such as hemoglobin and myoglobin. Seven years later, in 1958, Kendrew *et al.* confirmed the three-dimensional structure of proteins by solving the first atomic structure of myoglobin using X-ray crystallography.[11]

Today, protein research remains a major scientific focus, with ongoing discoveries in protein structure, function, and engineering driving innovations in fields such as biotechnology and medicine.

---

[8] Beach, E. F. *J. Hist. Med. Allied. Sci.* **16**, 354–373 (1961).

[9] Vickery, H. B., Schmidt, C. L. A. *Chem. Rev.* **9**, 169–318 (1931).

[10] Hartley, H. *Nature* **168**, 244 (1951).

[11] Kendrew, J. C. *et al. Nature* **181**, 662–666 (1958).

3.2.2. Molecular Structure and Organization of Proteins

3.2.2.1. Primary Structure: Amino Acids and Peptide Bond Formation

Proteins are composed of fundamental units called amino acids, small organic molecules consisting of an alpha carbon bonded to an amino group (–NH_2), a carboxyl group (–COOH), a hydrogen atom and a variable side chain (–R) that distinguishes each amino acid, attached to the same carbon atom, known as α carbon (Cα).

Proteins are composed of twenty naturally occurring amino acids, commonly represented by a letter, according to the notation proposed by Šorm in 1961,[12] but can also be found represented by a three-letter abbreviation. These are trypically classified according to the physicochemical properties of their side chains as apolar (G, A, V, L, I, M, F, W, P); uncharged polar (S, T, C, Y, N, Q) and charged, either acidic (D, E) or basic (H, K, R)[13] (**Figure 5**).

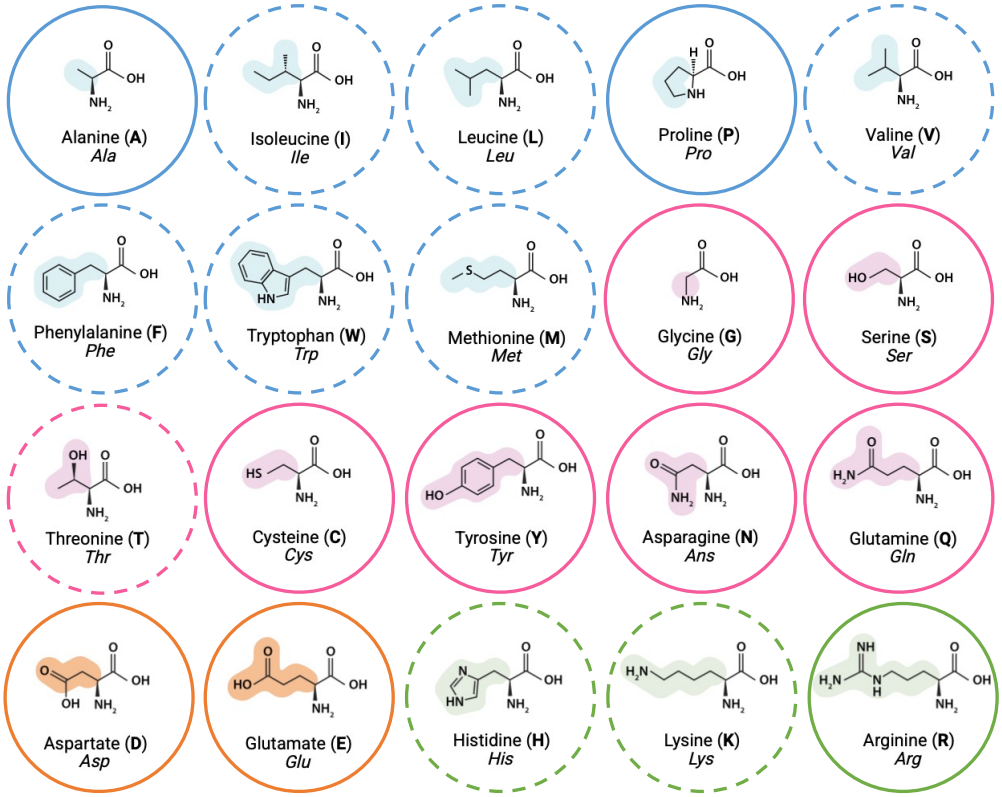


Figure 5. Classification of the 20 natural amino acids: apolar (blue), uncharged polar (pink), acidic (orange) and basic (green). The essential amino acids indicated with a dashed line. The chemical structure, the full name, the three-letter abbreviation and the one-letter abbreviation are shown. Image created with BioRender.com.

---

[12] Šorm, F *et al. Collect. Czech. Chem. Commun.* **26**, 531-578 (1961).

[13] (a) Voet, D., Voet, J. G., Pratt, C. W. *Fundamentals of Biochemistry,* John & Sons, New York (2007); (b) Alberts, B. *et al. Molecular Biology of the Cell*, Garland Publishing, New York, Chapter 3 (2008).

Structurally, proteins are polymers of amino acids, linked by amide bonds, known as peptide bonds. Each amino acid unit within the chain is also known as a *residue*, as it represents a repeating unit of the polymer. These bonds are formed through a condensation reaction, in which a water molecule is released when the carboxyl group of one amino acid reacts with the amino group of the next.

Peptide bond formation requires energy, since the reaction thermodynamically favors hydrolysis. However, once formed, the peptide bond is kinetically stable.[14] Notably, the C–N bond in the peptide is shorter than the typical C–N single bond found in amines, indicating a partial double-bond character between the carbonyl carbon and the amide nitrogen. As a result, the atoms involved in the peptide bond are coplanar, restricting their free rotation, conferring structural rigidity.[15]

The specific sequence of amino acids and the peptide bonds that link them together is what is known as the *primary structure* of a protein and is directly related to the protein's folding and biological function (**Figure 6**).

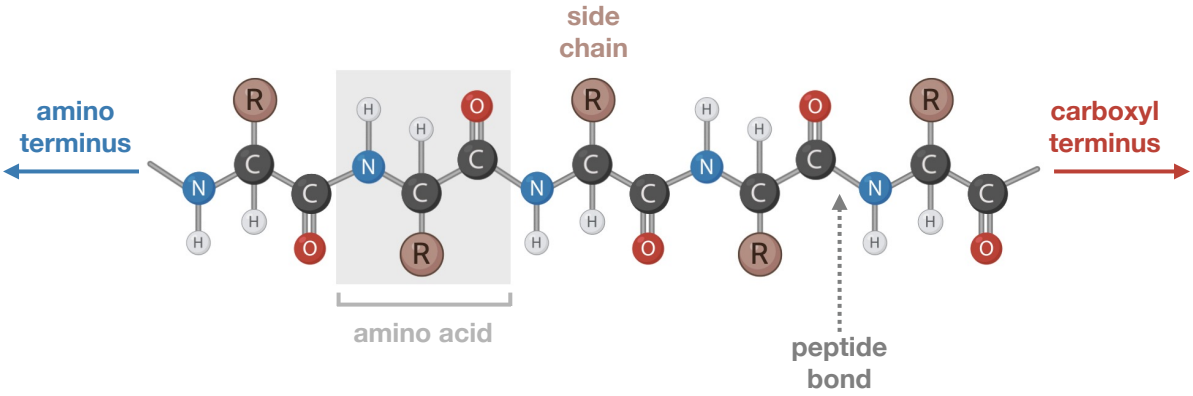


Figure 6. Primary structure of proteins. The image shows a chain of amino acids linked by peptide bonds, with amino (N-terminal) and carboxyl (C-terminal) ends indicated. Each amino acid has a side chain (R) that determines its chemical properties. Image created with BioRender.com.

3.2.2.2. Secondary Structure Motifs: α-helix and β-sheet

The secondary structure of a protein refers to the local spatial arrangement of the backbone atoms in a polypeptide chain, excluding the side chains and their relationship with other segments. This local folding adopts particularly stable patterns known as *secondary structure motifs*, with the most common being the α-helix and β-sheet, along with turns and, in the absence of a regular pattern, the so-called random coil.

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

[15] Nelson, D. L., Cox, M. M. *Lehninger Principles of Biochemistry*, W. H. Freeman Company, New York, Chapter 4 (2017).

The α-helix is one of the most common and significant secondary structure elements in proteins, with approximately 30% of residues in proteins adopting this conformation.[16] The α-helix is a right-handed helix with 3.6 amino acid residues per turn, where the polypeptide backbone coils around an imaginary axis, and the side chains extend outward. This structure is stabilized by hydrophobic effects from the backbone compaction and by hydrogen bonds formed between the carbonyl (CO) group of a given residue (*i*) and the amide (NH) group of the residue located four positions ahead (*i+4*)[17] (**Figure 7**a).

The transition from a random coil to an α-helix is entropically unfavorable, and peptides forming helices in solution often exist as a complex mixture of partly folded helices with frayed ends.[18] However, the stability of the α-helix depends on various factors, including the amino acid composition, sequence position, and residues near its ends.

Certain amino acids, such as proline and glycine, destabilize the helix due to structural constraints or excessive flexibility, while alanine and arginine have the highest tendencies to form α-helical conformations. Additionally, steric hindrance and hydrogen bond competition from residues like valine, threonine, serine, aspartate, and asparagine can also disrupt the α-helical conformation.[19]

β-sheets are regular secondary structures in proteins, composed of two or more short polypeptide strands (5-8 residues) known as β-strands. These strands adopt a fully extended zigzag conformation with a residue spacing of 3.5 Å. Adjacent β-strands are stabilized by hydrogen bonds between the N–H and C=O groups of amide bonds, forming either parallel or antiparallel arrangements, with the antiparallel configuration being more stable due to minimal hydrogen bond distortion (**Figure 7**b). The β-sheet structure can be formed by sequential or distant segments within the same polypeptide chain or even by strands from different polypeptides.

Proteins also contain loops and turns, structures that connect α-helices and β-sheets that are usually located on the surface of the protein, sometimes involved in intermolecular interactions and protein folding.[20] Although most of these elements adopt irregular secondary conformations, there are exceptions such as β-turns, which are closed turns that link adjacent and antiparallel β-strands.

---

[16] Barlow, D. J., Thornton, J. M. *J. Mol. Biol.* **201**, 601–619 (1988).

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

[18] Clarke, D. T. *et al. Proc. Natl. Acad. Sci. USA* **96**, 7232-7237 (1999).

[19] Pace, C. N., Scholtz, J. M. *Biophys. J.* **75**, 422–427 (1998).

[20] Rose, G. D. *Nature* **272**, 586–590 (1978).

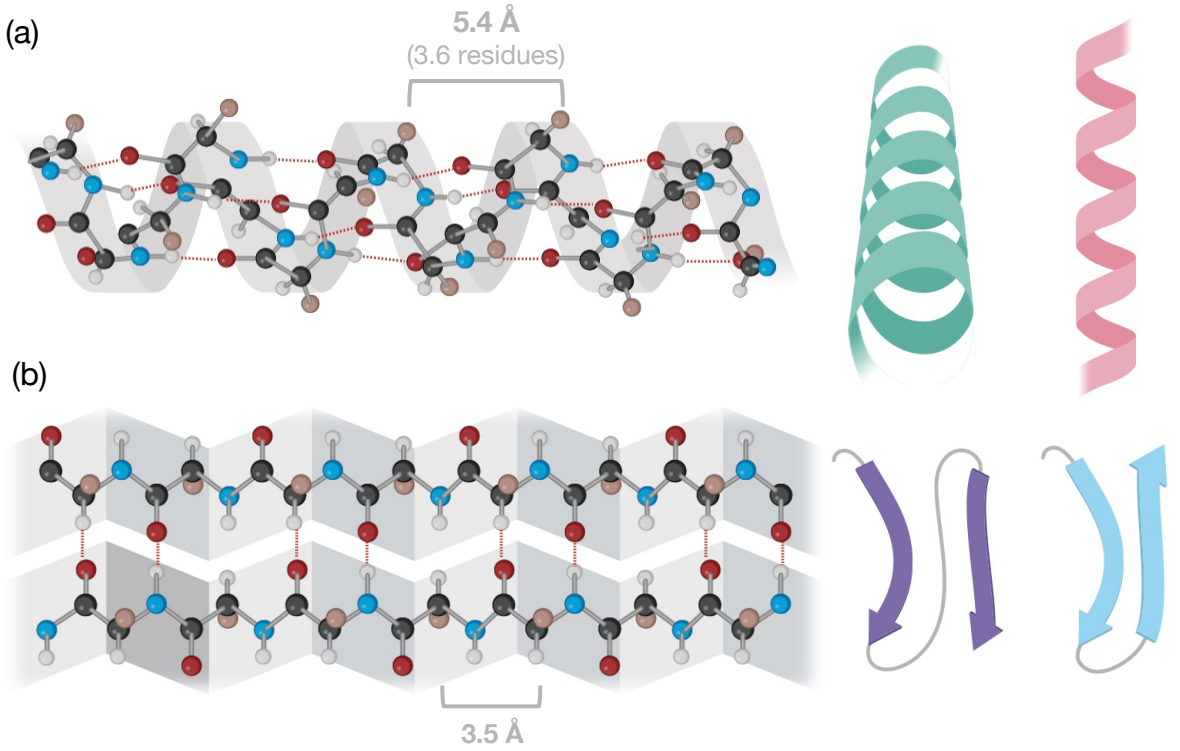


Figure 7. (a) Ball-and-stick model of the structure of an α-helical protein, showing intrachain hydrogen bonding (red dashed lines) and a cartoon of its view from one end, looking down the long axis (green), and from the side (pink); (b) ball-and-stick model of the β-sheet protein structure, showing hydrogen bond cross-links between adjacent chains (red dashed lines) and cartoons of a parallel (purple) and an antiparallel (blue) β-sheet. (Carbon atoms represented in black, oxygen in red, nitrogen in blue, hydrogen in white, and the side chain in brown). Image created with BioRender.com.

3.2.2.3. Tertiary and Quaternary Structure: Protein Folding and Assembly

The *tertiary structure* of a protein refers to the three-dimensional arrangement of all its atoms, integrating the precise spatial coordination of secondary structure elements and the positioning of functional groups within a single polypeptide chain. Protein folding is primarily driven by non-specific hydrophobic interactions, where hydrophobic residues are buried away from water.[21] However, the stability of the structure is determined by specific tertiary interactions, such as salt bridges, hydrogen bonds, close packing of side chains or disulfide bonds.

In many cases, especially in proteins with a mass greater than 100 kDa, the functional structure is not limited to a single polypeptide chain but consists of multiple subunits. The spatial arrangement of these subunits within a protein is known as the *quaternary structure.*[22]

---

[21] Rose, G. D., Roy, S., *Proc. Natl. Acad. Sci. USA* **77**, 4643–4647 (1980).

[22] Wang, W., Nema, S., Teagaden, *D. Int. J. Pharm*. **390**, 89–99 (2010).

3.2.3. Synthetic Peptides

A peptide is a biopolymer composed of two or more amino acids joined by peptide bonds. Chains of less than 50-100 amino acid residues are generally considered peptides, while larger structures are classified as proteins. However, threre is no clear definition of the boundaries between proteins and peptides.

Over the past three decades, a wide range of peptides with essential functions in biological systems have been identified and characterized.[23] These molecules play crucial roles as hormones, neurotransmitters, antibiotics and regulators of various metabolic pathways.[24] Due to their broad biological activity and their potential as therapeutic agents, peptides have become valuable in the development of new drugs, vaccines and diagnostic tools. As a result, the study and synthesis of peptides that mimic their natural analogues has become consolidated as a fundamental field at the intersection of chemical biology, medicinal chemistry and biotechnology.[25]

Beyond their biomedical applications, synthetic peptides have gained increasing attention in materials science and nanotechnology. Functionalized peptides with customized properties, such as fluorescent peptides, have been widely explored in bioimaging, biodetection and diagnostic applications,[26] providing highly specific and tunable tools to track biological processes at the molecular level. In parallel, peptides with catalytic properties have emerged as promising alternatives to traditional enzymes, offering advantages such as increased stability, tunability and reusability in chemical and biotechnological processes.[27] These advances position synthetic peptides as versatile building blocks for research in the field of synthetic biology and chemistry.

3.2.3.1. Strategies for Peptide Synthesis and Modification

Peptide synthesis is a fundamental technique in chemical biology that allows the production of bioactive peptides for therapeutic and research applications.

Most current methods for the chemical synthesis of peptides use Merrifield's *Solid-Phase Peptide Synthesis* (SPPS) approach,[28] which allows the synthesis of peptides and small proteins up to about 100 amino acids step-by-step on an insoluble resin, which facilitates automation and purification (**Figure 8**). SPPS relies on protecting group strategies, such as

---

[23] Fosgerau, K., Hoffmann, T. *Drug Discov. Today* **20**, 122–128 (2015).

[24] A. Giannis, A., Kolter, T. *Angew. Chem. Int. Ed. Engl.* **32**, 1244–1267 (1993).

[25] (a) Shepherd, N. E. *et al. J Am Chem Soc* **127**, 2974–2983 (2005); (b) Groß, A., Hashimoto, C., Sticht, H., Eichler, J. *Front. Bioeng. Biotechnol.* **3** (2016).

[26] (a) Tung, C. H. *Biopolymers (Pept. Sci.)* **76**, 391–403 (2004); (b) Pazos, E. *et al. Chem. Soc. Rev.* **38**, 3348–3359 (2009).

[27] (a) Maeda, Y. *et al. Annu. Rev. Biomed. Eng.* **18**, 311–328 (2016); (b) Carvalho, S. *et al. Isr. J. Chem.* **62**, e202200029 (2022)

[28] Merrifield, R. B. *J. Am. Chem. Soc.* **85**, 2149–2154 (1963).

Fmoc (fluorenylmethyloxycarbonyl) and Boc (tert-butyloxycarbonyl), to control selective amino acid coupling and avoid unwanted side reactions, in particular the most used strategy is Fmoc/tBu. Peptides can be modified to improve their stability, bioavailability or biological activity. Common modifications include N-terminal acetylation and C-terminal amination, which improve resistance to enzymatic degradation.

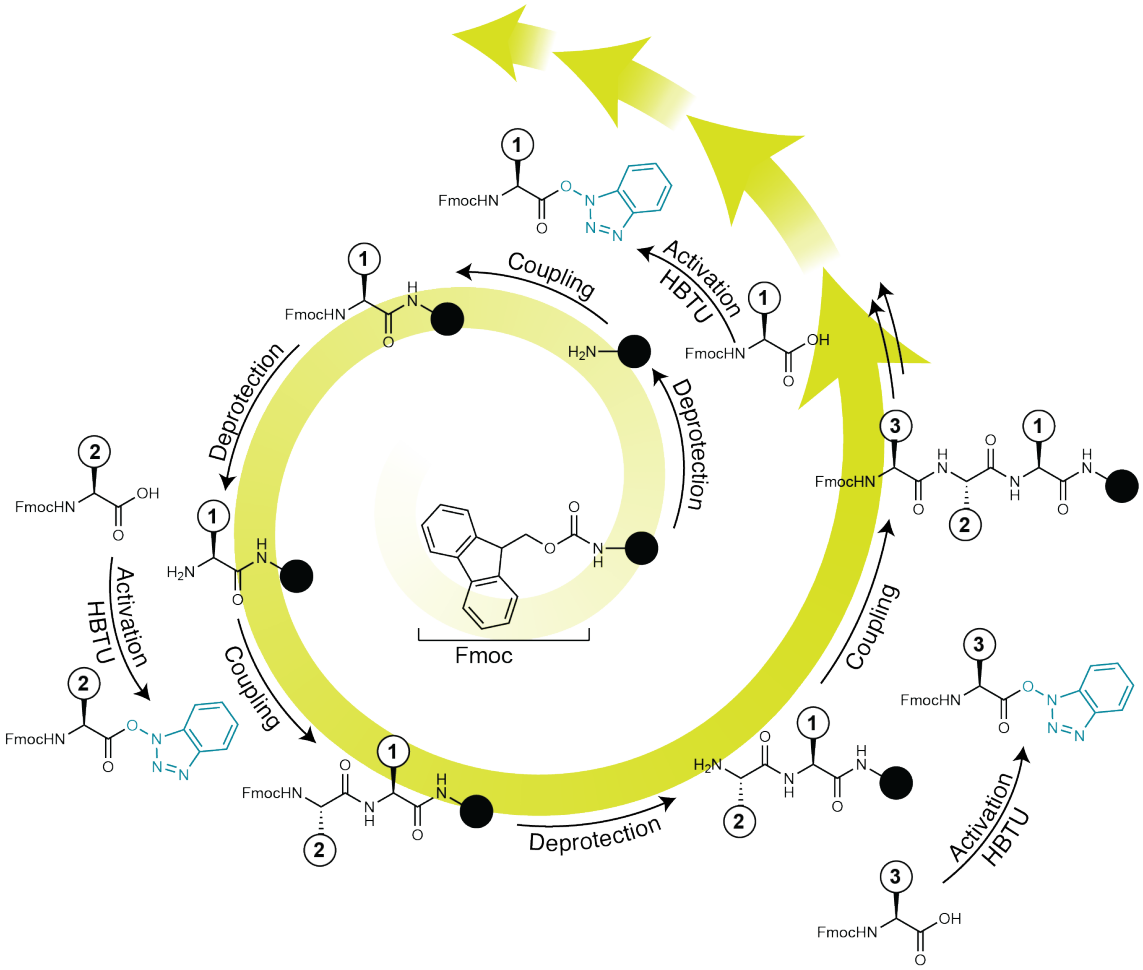


Figure 8. Schematic representation of Solid Phase Peptide Synthesis (SPPS).

A major advantage of SPPS, compared to recombinant protein expression, is the expanded set of non-natural amino acids and other building blocks that can be incorporated, including a wide range of non-proteinogenic amino acids or fluorescent tagging for imaging applications. These strategies expand the utility of synthetic peptides in drug development, diagnostics and biomaterials research.

3.2.3.2. Synthetic Peptide Libraries

A peptide library is a fundamental tool for studying proteins, consisting of a diverse range of peptides with systematic amino acid combinations. These libraries are widely used in drug design, protein-protein interaction analysis, and various biochemical and pharmaceutical applications. However, the creation of such libraries using conventional technologies is often

a relatively slow and costly process. To overcome these limitations, various modified peptide synthesis methods have been developed, such as OBOC.[29] These techniques allow for improved synthesis efficiency and scalability, facilitating the discovery and characterization of bioactive peptides.

SPOT synthesis was developed by Ronald Frank and co-workers in 1990, and further automation of this method was promoted by Jens Schneider-Mergener.[30] SPOT libraries are prepared by synthesizing peptides in a flat matrix, typically cellulose sheets, by spotting small aliquots of the different reagents in defined positions on the solid support;[31] the resulting matrix with the whole set of peptide sequences can be positionally addressed, thus greatly simplifying the selection process (**Figure 9**).[32]

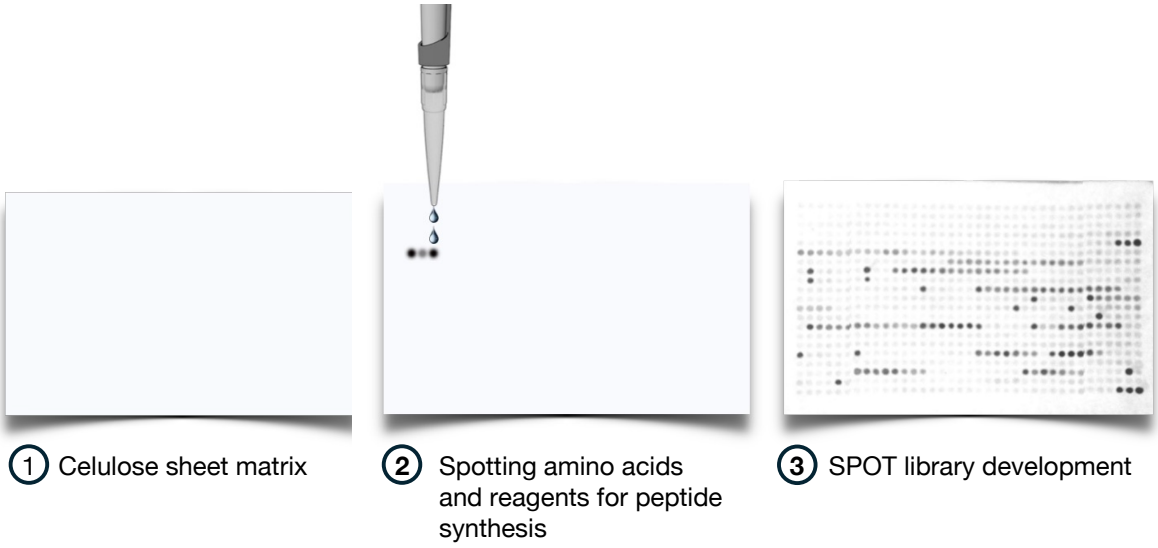


Figure 9. Preparation of the cellulose membrane. (1) Untreated cellulose membrane; (2) spotting of the different amino acids and reagents for the peptide synthesis; (3) development for the SPOT library.

The use of cellulose sheets has several advantages, as it is an inexpensive material that is resistant to the organic solvents and reagents used in SPPS. In addition, cellulose is stable in aqueous solutions and, because it is non-toxic, it is suitable for screening biological samples. Another advantage of this technique is that it allows the synthesis of chains of up to 50 amino

---

[29] (a) Houghten, R. A. *Proc. Natl. Acad. Sci. USA* **82**, 5131–5135 (1985); (b) Pellois, J. P. *et al. Nat. Biotechnol.* **20**, 922–926 (2002); (c) Geysen, H. M., Meloen, R. H., Barteling, S. J. *Proc. Natl. Acad. Sci. USA* **81**, 3998–4002 (1984); (d) Frank, R. *Tetrahedron* **48**, 9217–9232 (1992).

[30] Kramer, A., Schneider-Mergener, J. *Methods Mol. Biol.* **87**, 25–39 (1998).

[31] (a) Hilpert, K., Winkler, D. F. H., Hancock, R. E. W. *Nat. Protoc.* **2**, 6, 1333–1349 (2007); (b) Hilper, K., Winkler, D. F. H., Hancock, R. E. W. *Biotechnol. Genet. Eng. Rev.* **24**, 31–106 (2007); (c) Reineke, U., Volkmer-Engert, R., Schneider-Mergener, J. *Curr. Opin. Biotechnol.* **12**, 1, 59–64 (2001).

[32] Frank, R. *et al.* Combinatorial Synthesis on Membrane Supports by the SPOT Technique: Imaging Peptide Sequence and Shape Space. In *Combinatorial Peptide and Nonpeptide Libraries*; Wiley-VCH Verlag GmbH: Weinheim, Germany, pp 363–386 (2007).

acids,[33] and because it is based on Fmoc chemistry, it facilitates the incorporation of non-natural amino acids, increasing the chemical diversity of peptides.[34] Finally, this technique allows us to modify the peptides both before and after cleavage of the protective group of the final side chain. All these features make this method an excellent tool for screening large numbers of peptides for many different purposes. Potential applications range from simple binding assays[35] to more sophisticated enzymatic assays[36] and studies with living microbes,[37] or cells.[38]

[33] Toepert, F. *et al. Angew. Chem. Int. Ed. Engl.* **42**, 1136–1140 (2003).

[34] (a) Heine, N. *et al. Tetrahedron* **59**, 9919–9930 (2003); (b) Blackwell, H. E. *Curr. Opin. Chem. Biol.* **10**, 203–212 (2006).

[35] Kramer, A. *et al. Cell* **91**, 799–809 (1997).

[36] Hilpert, K. *et al. J. Biochem.* (Tokyo) **128**, 1051–1057 (2000).

[37] Hilpert, K. *et al. Nat. Biotechnol.* **23**, 1008–1012 (2005).

[38] Kamradt, T., Volkmer-Engert, R. *Mol. Divers* **8**, 271–280 (2004).

3.3. Luminescent Peptides for Bioimaging and Sensing

3.3.1. Fundamentals of Photoluminescence

The term *luminescence* comes from the Latin *lumen* (light) and was first introduced in 1888 by Eilhard Wiedemann as *luminescenz.*[39] It was used to describe all light-emitting phenomena that do not depend exclusively on a rise in temperature, thus differentiating it from incandescence. Today, luminescence is defined as "the spontaneous emission of radiation by an electronically or vibrationally excited species that is not in thermal equilibrium with its surroundings".[40] Luminescence can be classified according to the excitation mechanism of the emitting species, including phenomena such as photoluminescence (light), chemiluminescence (a chemical reaction), bioluminescence (enzymatic reaction), electroluminescence, etc. The most studied of these is photoluminescence, which is triggered when a molecule absorbs a photon in the visible region, causing one of its electrons to move from the ground state to a higher energy level. The light absorbed by the sample is converted into the excitation energy of the species, which in turn is transformed into chemical energy and/or dissipated in the form of radiant or heat energy (**Figure 10**).

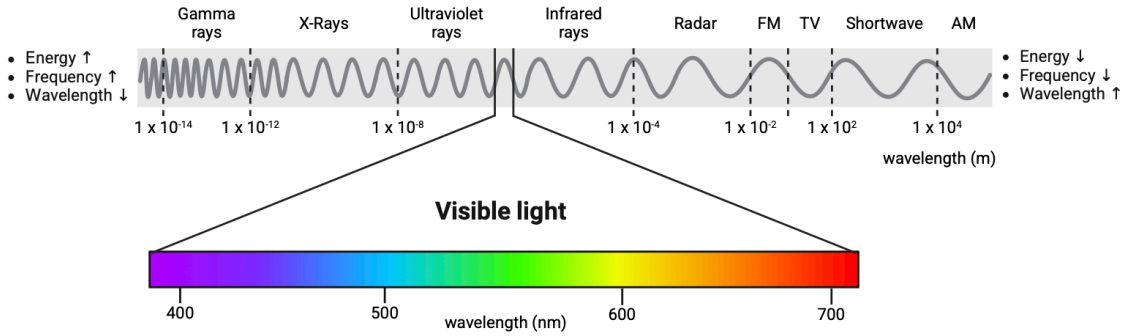


Figure 10. Electromagnetic spectrum showing the relationship between energy, frequency and wavelength. It highlights the visible light range, which spans approximately 400 to 700 nm, between ultraviolet and infrared radiation. Image created with BioRender.com.

The best tool to interpret the nature of photophysical processes is the so-called Perrin-Jablonski diagram that shows the relevant energy levels involved in the electronic transitions, together with the different radiative and non-radiative processes involved in the photoluminescence[41] (**Figure 11**). In general, if the emission of light occurs without a change in spin multiplicity, i.e., when a photon is emitted when the electrons are transferred from the

---

[39] Wiedemann, E. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **28**, 175, 493–495 (1889).

[40] Braslavsky, S. E. *et al. Pure App.l Chem.* **79**, 293–465 (2007).

[41] Valeur B., Brochon, J. New Trends in Fluorescence Spectroscopy: Applications to Chemical and Life Sciences, *Springer* (2001).

lowest-energy singlet excited state (S_1) to the fundamental state (S_0), the process is called *fluorescence*. In contrast, when the luminescence involves a triplet (T_1) to singlet transition or other change in spin multiplicity, it is known as *phosphorescence*. This process is formally forbidden by the quantum selection rules, which makes it rare at room temperature and in liquid solutions, where non-radiative deactivation mechanisms are usually faster. However, under low temperature conditions or in rigid media, phosphorescence becomes observable due to the long lifetime of the triplet state, allowing emissions that can extend for seconds, minutes or even longer. In addition to these radiative processes, non-radiative deactivation mechanisms can occur, such as internal conversion, where energy is dissipated without light emission, or interactions with nearby molecules, which can result in electron transfer or the formation of excimers and exciplexes.

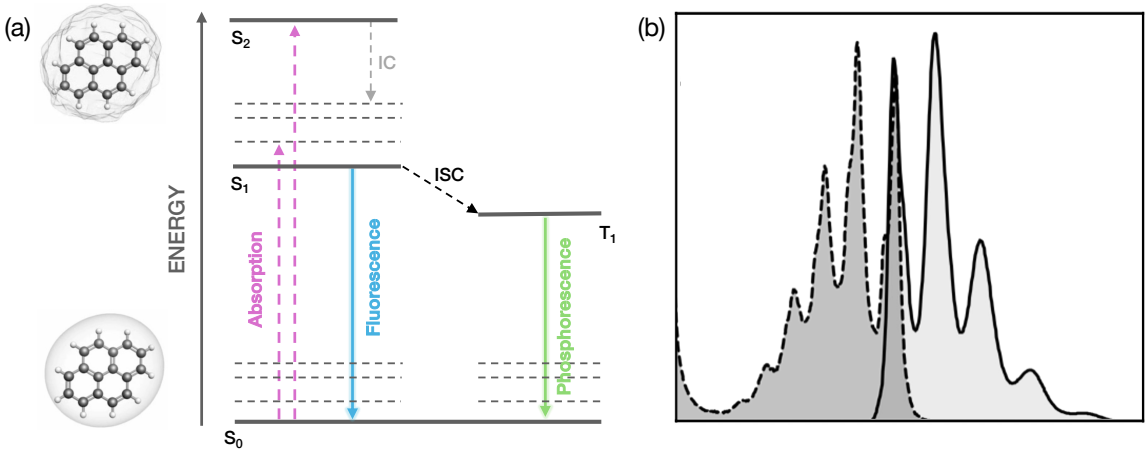


Figure 11. (a) Perrin-Jablonski diagram representing the main photophysical processes: absorption (purple), fluorescence (blue) and phosphorescence (pink), internal conversion (IC) and intersystem conversion (ISC) as non-radiative deactivation mechanisms are also shown; (b) representation of absorption and emission spectra of one molecule.

Photoluminescence techniques have become fundamental tools in biodetection and bioimaging,[42] mainly in the fields of molecular cell biology and protein biochemistry. Their ability to offer high sensitivity, reliability, and spatial and temporal resolution has enabled breakthroughs in the study of biological systems.[43] These techniques are used in light-emitting devices,[44] chemosensors[45] and biosensors,[46] facilitating rapid and accurate localization, trafficking and activity of biomolecules of interest.

---

[42] (a) Specht, E. A., Braselmann, E., Palmer, A. E. A. *Annu. Rev. Physiol.* **79**, 93–117 (2017); (b) Lavis, L. D., Raines, R. T. *ACS Chem. Biol.* **9**, 855–866 (2014).

[43] (a) Lakowicz, J. R. Principles of Fluorescence Spectroscopy, 3rd edition. *Springer US* (2006); (b) Park, S. J. *et al*. *Stem Cell Res.* **12**, 730–741 (2014); (c) Chan, J., Dodani, S. C., Chang, C. J. *Nature Chem* **4**, 973–984 (2012).

[44] Hwang, J., Nagaraju, P., Cho, M. J., Choi, D. H. *Aggregate* **4** (2023).

[45] Zhou, P., Han, K. *Aggregate* **3** (2022).

[46] Xu, S., Duan, Y., Liu, B. *Adv. Mater.* **32** (2020).

3.3.2. Sensors and Fluorescent Markers

3.3.2.1. Organic Fluorophores

Photoluminescence techniques rely on special molecules, known as fluorophores, which are based on conjugated structures that allow efficient absorption and emission of light. These structures include aromatic ring systems, conjugated double bonds and specific functional groups that influence their spectral properties. The first specifically recognized organic fluorescent molecule was quinine sulphate, identified by Sir John Herschel in 1845, which emits blue fluorescence when irradiated with ultraviolet (UV) light.[47] Since then, numerous other organic compounds have been discovered and synthesized, that emit fluorescence in a variety of colors. Among the most significant are xanthene derivatives, such as fluorescein[48] and rhodamine,[49] which were first reported in the late XIX century, and are still widely used today in fluorescent labelling and probes for bioimaging. In addition, dyes based on pyrones and oxazines, like BODIPY,[50] as well as cyanine dyes, such as Cy3 or Cy5, which stand out for their absorption in the red and near-infrared,[51] are also often used for this purpose.

Fluorophores with emissions spanning the entire visible spectrum are now available, enabling a broad range of applications in biological labeling and imaging (**Figure 12**).

Fluorophores have played a pivotal role for *in vitro* imaging since the 1940s, when Albert H. Coons and coworkers published a landmark paper describing the use of a synthetic fluorophore, fluorescein isothiocyanate (FITC), to label an antibody for diagnostic purposes.[52] This was the beginning of immunofluorescence, a technique that is still widely used today. This breakthrough demonstrated how fluorophores could be conjugated to biomolecules to create highly specific and sensitive imaging tools. These conjugates improve biocompatibility and targeting ability,[53] while allowing precise manipulation of fluorescence properties such as emission wavelength and lifetime. However, despite their advantages, many artificial fluorophores present problems such as, complicated synthesis with high cost and pollutant generation and inevitable biotoxicity and immunotoxicity; [54] meanwhile, most of the fluorophores with planar structures suffer from an aggregation-caused quenching effect, which impede their development for practical applications.[55]

---

[47] Herschel, Sir J. F. W. *Philos. Trans. R. Soc. Lond.* **135**, 143–145 (1845).

[48] Baeyer, A. *Ber. Dtsch. Chem. Ges.* **4**, 555–558 (1871).

[49] Ceresole, M. *US Patent* **377**, 349 (1888).

[50] Loudet, A., Burgess, K. *Chem. Rev.* **107**, 4891–4932 (2007).

[51] Luo, S. *et al. Biomaterials* **32**, 7127–7138 (2011).

[52] Coons, A. H., Creech, H. J., Jones, R. N. *Proc. Soc. Exp. Biol. Med.* **47**, 200–202 (1941).

[53] Naghibi, S., Chen, T., Jamshidi Ghahfarokhi, A., Tang, Y. *Aggregate* **2** (2021).

[54] Zhang, Z. *et al*. *Luminescence* **32**, 1488–1493 (2017).

[55] (a) Deng, X. *et al*. *Aggregate* **2**, e35 (2021); (b) Mei, J. *et al. Chem. Rev.* **115**, 11718–11940 (2015).

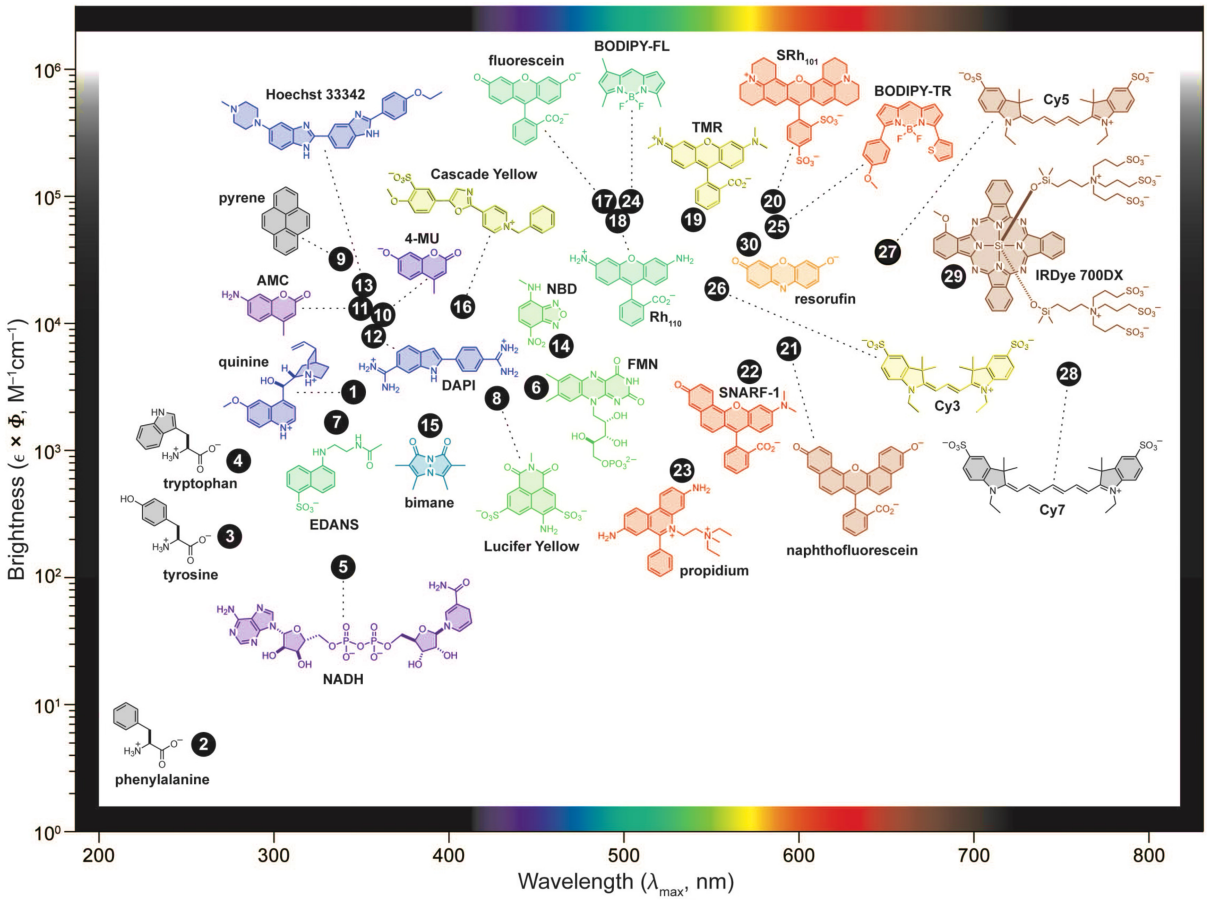


Figure 12. Emission spectra and brightness of various fluorescent molecules used in biology and chemistry.[56] The horizontal axis represents the maximum wavelength of emission (λ_{max}), ranging from the ultraviolet to the near infrared, while the vertical axis indicates the relative brightness of each fluorophore, expressed as the product of the quantum efficiency (Φ_f) and the molar extinction coefficient (ε). These include natural fluorophores such as aromatic amino acids (phenylalanine, tyrosine and tryptophan), as well as synthetic dyes widely used in fluorescence microscopy and molecular biology, such as fluorescein, BODIPY, cyanines and rhodamines (image adapted from reference [58a]).

3.3.2.2. Förster Resonance Energy Transfer (FRET) and The Antenna Effect

Despite significant advances in bioimaging and detection, technical challenges remain that limit our ability to visualize biological processes at the molecular level. There are several techniques to exploit additional properties of fluorophores, such as fluorescence resonance energy transfer (FRET), developed to study intermolecular interactions and metabolic processes. This photophysical phenomenon, described by the physicist Theodor Förster in 1948,[57] is based on the transfer of non-radiative energy between a donor dye and an acceptor

[56] (a) Lavis, L. D., Raines, R. T. *ACS Chem. Biol.* **3**, 142–155 (2008); (b) Grimm, J. B., Lavis, L. D. *Nat. Methods* **19**, 149–158 (2022); (c) Liu, X., Zhu, C., Tang, B. Z, *Acc. Chem. Res.* **55**, 197–208 (2022).

[57] Sener, M. *et al. ChemPhysChem* **12**, 518–531 (2011).

dye when they are close to each other. For this to occur, the emission spectrum of the donor must overlap with the absorption spectrum of the acceptor, and the energy difference between their electronic states must be compatible (**Figure 13**).[58] Since its discovery, the FRET process has been a breakthrough in fluorescence detection of biomolecules.[59]

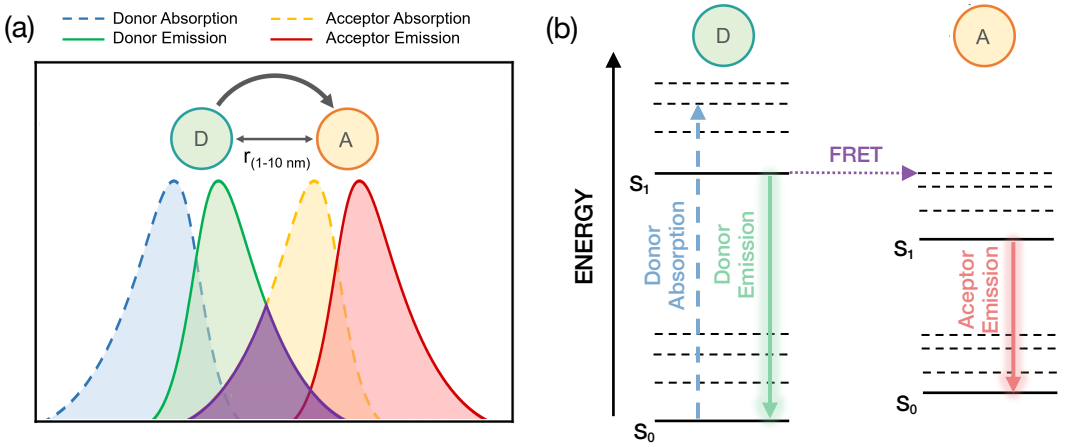


Figure 13. Schematic representation of the Förster Resonance Energy Transfer (FRET) mechanism. (a) Spectral overlay between donor (D) and acceptor (A); Perrin-Jablonski diagram representing the energy levels involved in the FRET process.

In addition to organic fluorophores, Lanthanide ion complexes of europium (Eu^{3+}) and terbium (Tb^{3+})[60] possess unique photophysical properties that make them very versatile molecular probes for the study of biological systems. Their excited states last longer than those of most conventional fluorophores, which is especially beneficial for time-gated detection techniques that reduce background signals from biological samples.

The emission from lanthanide ions is generated differently from fluorescence, as it originates from a forbidden parity transition within the 4*f* atomic orbitals. This gives them their unique properties, such as long lifetime, narrow band emission and large Stokes shifts, allowing experiments that would be impossible with organic fluorophores or fluorescent proteins. The forbidden nature of these transitions causes lanthanide ions to exhibit very weak light absorption, with molar absorption coefficients usually below 3.5 $M^{-1}cm^{-1}$ which makes their direct optical excitation inefficient. To overcome this, scientists often use an indirect excitation strategy called the "antenna effect" or sensitization. This involves exciting an organic chromophore, which then transfers energy to the lanthanide ion, ultimately producing luminescence. The energy transfer from the chromophore to the lanthanide can happen through two main mechanisms: Dexter transfer, which relies on orbital overlap and electron

[58] Algar, W. R. *et al. Coord Chem Rev* **263**, 65–85 (2014).

[59] (a) Algar, W. R. *et al*. *Nat. Methods* **16**, 815–829 (2019); (b) Bakar, B. T. *et al. Sensors (Basel)* **16**, 1488 (2016); (c) Hochreiter, B., Garcia, A. P., Schmid, J. A. *Sensors (Basel)* **15**, 26281–26314 (2015).

[60] (a) Selvin, P. R. *Annu. Rev. Biophys. Biomol. Struct.* **31**, 275-302 (2002); (b) Chen, J., Selvin, P. R. *Bioconjugate Chem.* **10**, 311-315 (1999).

exchange, or Förster transfer, which involves dipole–dipole interactions between the excited chromophore and the lanthanide ion. Due to this process, most lanthanide-based imaging tools and biosensors include three key components: the luminescent lanthanide ion, a ligand to hold it in place, and a sensitizing chromophore to absorb and pass along the energy.

While synthetic fluorophores and lanthanide-based probes have expanded the possibilities of fluorescence-based detection, their application in living systems can be limited by factors such as photostability, cytotoxicity, and complex conjugation procedures. These limitations have driven researchers to explore naturally occurring fluorophores, such as fluorescent proteins, which offer intrinsic biocompatibility and unique photophysical properties.

3.3.2.3. Fluorescent Proteins: Green Fluorescent Protein

Since the early studies in the 1950s-1960,[61] it has been accepted that the UV/vis absorption and fluorescence of proteins results exclusively from the presence of aromatic residues in their sequenc.[62] The natural aromatic amino acids are tryptophan (Trp), tyrosine (Tyr) and phenylalanine (Phe), which have absorption maxima around 280, 275 and 257 nm, respectively.[63] Among them, tryptophan stands out as the intrinsic fluorophore with the best properties including a molar absorption coefficient (ε_{278}) of 5.579 $M^{-1}cm^{-1}$ and a quantum yield (ϕ) of 0.12.[64]

However, in the 1960s, Osamu Shimomura made a major breakthrough by isolating green fluorescent proteins (GFP) from the jellyfish *Aequorea victoria*[65] (**Figure 14**a), which is capable of forming a 4-(p-hydroxybenzylidene)-5-imidazolinone (HBI) fluorophore autocatalytically from a combination of serine-tyrosine-glycine residues.[66] The main advantage of GFP and other fluorescent proteins over organic dyes is their ability to be expressed exogenously in cells as genetically encodable tags with any protein of interest, allowing the identification of a wide variety of biological events, such as subcellular localization and expression patterns in real time.[67] Recognizing the profound impact of GFP on biological research, the 2008 Nobel Prize in Chemistry was awarded to Osamu Shimomura, Martin Chalfie, and Roger Tsien for their pioneering work in its discovery and development (**Figure 14**b).

---

[61] (a) Debye, P.; Edwards, J. O. *Science* **116**, 143–144 (1952); (b) Wetlaufer, D. B. *Adv. Protein Chem.* **17**, 303–390 (1963); (c) Teale, F. W. *Biochem. J.* **76**, 381–388 (1960).

[62] (a) Bao, J., Tong, C., He, M., Zhang, H. *Luminescence* **39** (2024); (b) Morzan, U. N. *et al. J. Phys. Chem. B* **126** 7203–7211 (2022).

[63] Hunt, H. D., Simpson, W. T. *Product Information Bulletin.* J. R. Leader and J. F. Gormley, ibid **2** (1953).

[64] Chen, Y., Barkley, M. D. *Biochemistry* **37**, 9976–9982 (1998).

[65] Shimomura, O., Johnson, F. H., Saiga, Y. *J Cell. Physiol.* **59**, 223–239 (1962).

[66] Ormö, M. *et al*. *Science (1979)* **273**, 1392–1395 (1996).

[67] (a) Chalfie, M. *et al. Science (1979)* **263**, 802–805 (1994); (b) Zimmer, M. *Chem. Rev.* **102**, 759–781 (2002); (c) Remington, S. J. *Protein. Sci.* **20**, 1509–1519 (2011).

The intense green fluorescence of GFP is highly dependent on the proper folding of the GFP structure. However, upon isolation of HBI from GFP, it loses its ability to emit fluorescence (**Figure 14**c). This behavior resembled the photophysical mechanism of aggregation-induced emission (AIE).[68]

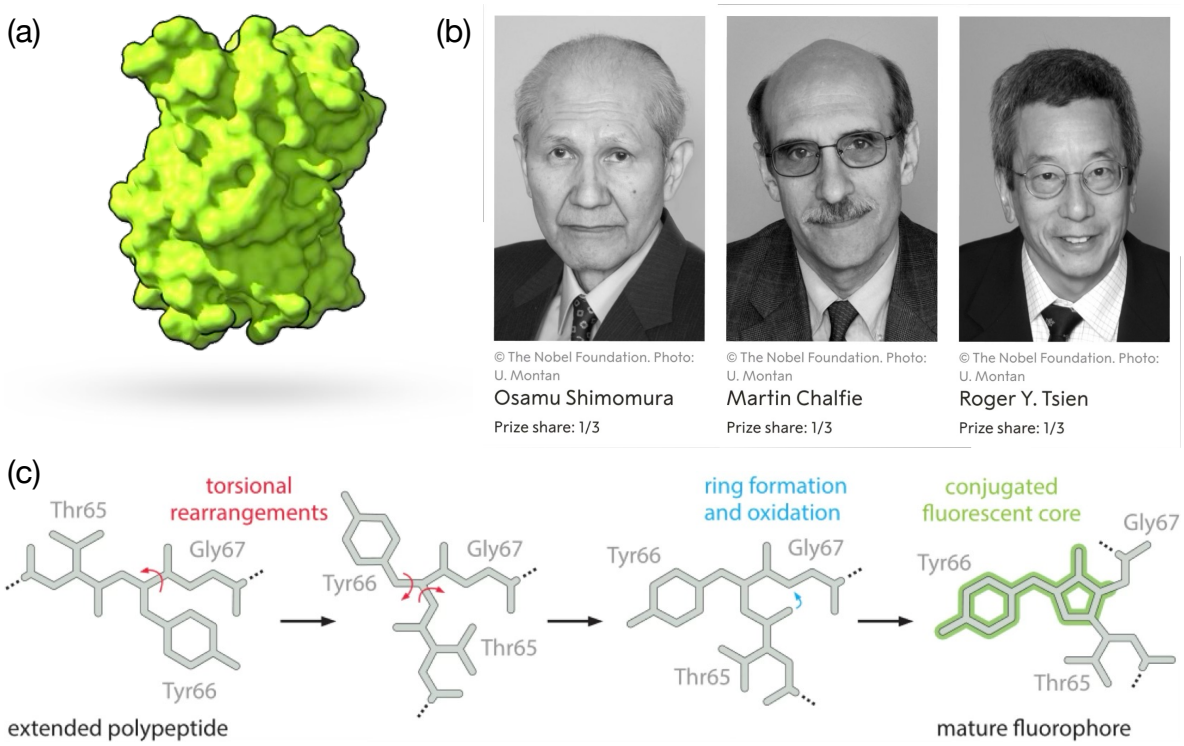


Figure 14. (a) Three-dimensional structure of the green fluorescent protein (GFP); (b) Osamu Shimomura, Martin Chalfie and Roger Y. Tsien, winners of the 2008 Nobel Prize in Chemistry; (c) maturation mechanism of the GFP fluorophore (HBI), involving torsional rearrangement, followed by cyclisation and oxidation to form the conjugated fluorescent core.

In recent years, one of the main goals of this field has been the identification and development of fluorescent proteins homologous to GFP. A key advance was introduced by Roger Tsien in 1995, who discovered that the S65T mutation improved both fluorescence and photostability, adjusting the excitation peak to 488 nm.[69] Subsequently, the F64L mutation gave rise to an enhanced GFP (EGFP), allowing its use in mammalian cells.[70] Other variants, such as superfolder GFP (sfGFP), facilitated folding under adverse conditions.[71]

---

[68] Zhao, Z., Zhang, H., Lam, J. W. Y., Tang, B. Z. *Angewandte Chemie - International Edition* **59**, 9888–9907 (2020).

[69] Heim, R., Cubitt, A. B., Tsien, R. Y. *Nature* **373**, 663–664 (1995).

[70] McRae, S. R., Brown, C. L., Bushell, G. R. *Protein Expr. Purif.* **41**, 121–127 (2005).

[71] Pédelacq, J. D. *et al. Nat. Biotechnol.* **24**, 79–88 (2006).

To date, significant progress has been made in the development of these variants, with a powerful set of fluorescent proteins that cover almost the entire visible spectrum, such as blue fluorescent protein (EBFP, EBFP2, Azurite, mKalama1), cyan fluorescent protein (ECFP, Cerulean, CyPet, mTurquoise2), and yellow fluorescent protein derivatives (YFP, Citrine, Venus, YPet), offering numerous alternatives for multicolour labelling and studies of protein-protein interactions.[72] Despite the enormous breakthrough in the industry, GFPs have a number of limitations for the study of anaerobic biosystems or hypoxia,[73] or rapid processes such as gene dynamics;[74] GFPs are large proteins (~28kDa) that can interfere with the function of a coexpressed biomolecule;[75] are fragile proteins and prone to (di/tetra)merization, which might lead to significant artifacts,[76] and are also immunogenic, which complicates its application *in vivo.*[77] To address these limitations, it is essential to develop novel genetically encodable tags for optical fluorescence microscopy.

[72] Pakhomov, A. A., Martynov, V. I. *Chem. Biol.* **15**, 755–764 (2008).

[73] Coralli, C., Cemazar, M., Kanthou, C., Tozer, G. M. Dachs, G. U. *Cancer Res.* **61**, 4784–4790 (2001).

[74] Balleza, E.,Kim, J. M., Cluzel, P. *Nat. Methods* **15**, 47–51 (2018).

[75] Hink, M. A. *et al. J. Biol. Chem.* **275**, 17556–17560 (2000).

[76] Swenson, E. S. *et al. Stem Cells* **25**, 2593–2600 (2007).

[77] Ansari, A. M. *et al. Stem Cell Rev* **12**, 553–559 (2016).

3.3.3. Emerging Phenomenon: Non-Aromatic Fluorescence (NAF)

In addition to the traditional fluorescence associated with intrinsic chromophores, anomalous photoluminescence, known as *Non-Aromatic Fluorescence* (NAF), has been observed in peptides with no aromatic residues. A notable example dates back to 1928, when Wels reported that serum albumin exhibited blue fluorescence at 450 nm upon excitation at 365 nm, which cannot be explained by the regular fluorescence of the aromatic residues.[78] Nearly a century later, in 2017, T. K. Mukherjee's group confirmed this observation and proposed that the anomalous emission originates from oligomeric aggregates.[79] Likewise, anomalous NAF has been widely described in many other systems, including peptides lacking aromatic residues, concentrated solutions of *poly*-Lys (**Figure 15**b),[80] or amyloid assemblies (**Figure 15**c), [81] as well as long-wavelength emission from aromatic amino acid aggregates,[82] and peptide condensates.[83]

Despite the current evidence of this interesting phenomenon, our understanding remains very limited. In fact, there is no working model or fundamental rules for the design and creation of non-aromatic peptide luminogens. Therefore, the detailed study of these systems represents a promising avenue for better understanding the principles of NAF and expanding their applications in fields such as biomedicine and materials science.

Various research groups have proposed different photophysical mechanisms to explain NAF, highlighting both its complexity and the current gaps in our understanding. For instance, studies of the autofluorescence of amyloid-like cross-β-sheet structures formed by the self-assembly of the octapeptide GVGVAGVVG, S. Sharpe *et al.* suggest that the intrinsic blue emission (λ_{exc} = 358 nm; λ_{em} = 460 nm) arises from the extensive delocalization of peptide bond electrons through the hydrogen bond networks within the amyloid structure.[84] Similarly, N. Amdursky *et al.* emphasized the role quantum confinement in the blue luminescence (λ_{exc} = 370 nm; λ_{em} = 450 nm) observed in H-Phe-Phe-OH dipeptide aggregates.[85]

---

[78] (a) Wels, P. *Pflügers Arch.* **219**, 738–752 (1928); (b) Konev, S. V. *Fluorescence and Phosphorescence of Proteins and Nucleic Acids* (Plenum Press, New York, NY, 1967).

[79] Bhattacharya, A. *et al. Langmuir* **33**, 10606–10615 (2017).

[80] (a) Stagi, L. *et al. Macromol. Chem. Phys.* **222**, (2021); (b) Homchaudhuri, L., Swaminathan, R. *Chem. Lett.* **30**, 844–845 (2001); (c) Cadeddu, M. *et al*. *Macromolecules* **57**, 514–527 (2024).

[81] Y. Gong, *et al. Sci. China Chem.* **56**, 1178–1182 (2013).

[82] (a) Babar, D. G., Sarkar, S. *Appl. Nanosci. (Switzerland)* **7**, 101–107 (2017); (b) Fan, Z., Sun, L., Huang, Y., Wang, Y., Zhang, M. *Nat. Nanotechnol*. **11**, 388–394 (2016).

[83] Sementa, D. *et al. Angew Chem. Int. Ed.* **62**, (2023).

[84] Sharpe, S., Simonetti, K., Yau, J., Walsh, P. *Biomacromolecules* **12**, 1546–1555 (2011)

[85] Amdursky, N. *et al. Nano Lett.* **9**, 3111–3115 (2009).

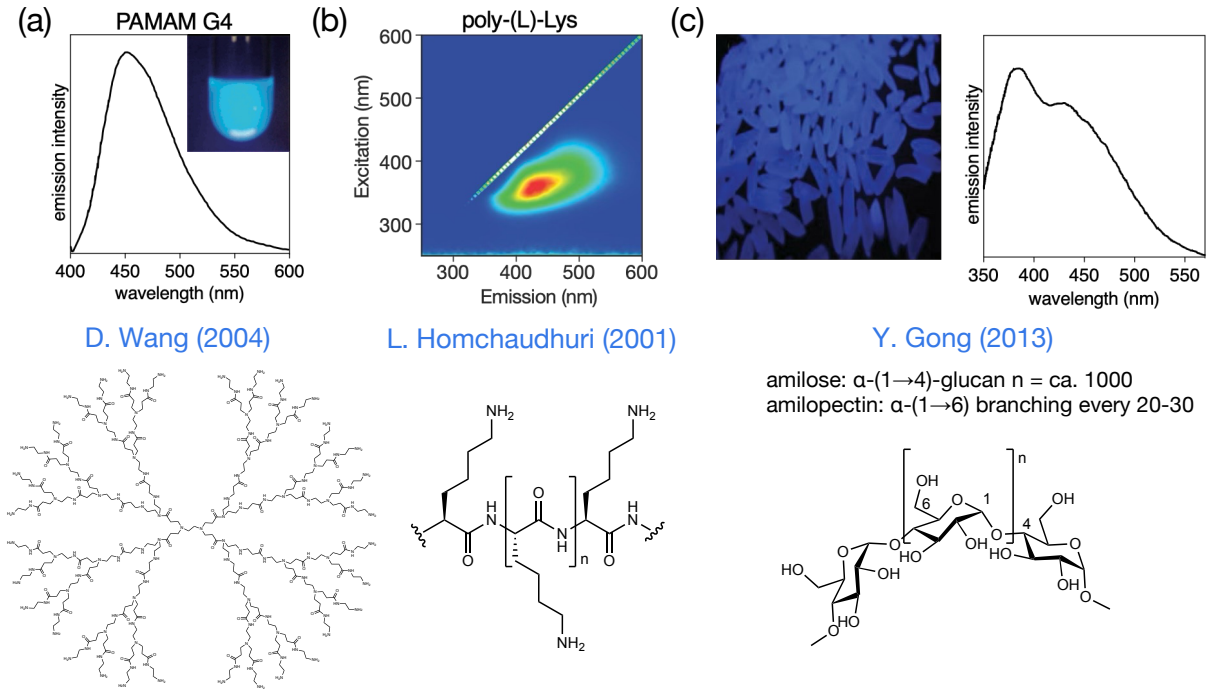


Figure 15. Examples of systems exhibiting non-aromatic fluorescence (NAF); (a) emission of a fourth generation (G4) PAMAM dendrimer, reported by Wang *et al.* (2004);[86] (image adapted form reference *[86]*) (b) excitation-emission matrix of poly-*L*-lysine demonstrating intrinsic fluorescence, according to Homchaudhuri *et al.* (2001) (image adapted form reference [82a]); (c) emission of natural biopolymers, such as starch, in a crystalline state, reported by Gong *et al.* (2013) (adapted form reference [83]).

An alternative photophysical explanation put forward by A. Hassanali's group is that NAF originates from the formation of Short Hydrogen Bonds (SHBs), which inhibit non-radiative decay pathways associated with carbonyl stretching vibrational modes.[87] Meanwhile, G. S. Kaminsky Schierle argued that proton transfer through hydrogen bonds reduces excitation energy levels, shifting them to the near-UV or visible range.[88] Furthermore, R. Swaminathan's group introduced an alternative explanation based on photoinduced charge transfer, where electron-rich donors (e.g., carboxylates) transfer charge to the LUMO of electron-poor acceptors (e.g., protonated amines or peptide chains), with subsequent charge recombination producing the emission known as Protein Charge Transfer Spectra (ProCharTS).[89] Finally, the clusteroluminescence (CL) mechanism, proposed by B. Z. Tang in 2019, suggests that clusteroluminogens (CLgens) follow band theory, and their energy gap is determined by the

[86] Wang, D.; Imae, T. *J. Am. Chem. Soc.* **126**, 13204–13205 (2004).

[87] (a) Stephens, A. D. *et al*. *Proc. Natl. Acad. Sci. USA* **118**, e2020389118 (2021); (b) Mirón, G. D. *et al. Nat. Commun.* **14**, (2023).

[88] Pinotsi, D. *et al*. *J. Am. Chem. Soc.* **138**, 3046–3057 (2016).

[89] Alom, S. E., Swaminathan R. *Phys. Chem. Chem. Phys.* **25**, 16626–16642 (2023).

energy difference between the conduction and valence bands.[90] The CL comes from electronic delocalization across space, induced by the physicochemical confinement of electron-rich functional groups (e.g., C=O, C-O, C-N, C(=O)-N) or small aromatic molecules, generating energy levels suitable for UV/Vis emission. Due to this mechanism, CLgens usually have an excitation spectrum different from the absorption spectrum, since the emission depends on the excitation due to the different sizes and conformations of the clusters and the emission is affected by the environment (**Figure 16**). To complicate things even more, in addition to these studies proposing NAF as a new photophysical mechanism, many researchers continue to support the traditional view that explains NAF as fluorescence emission from photooxidized amino acid side chains or impurities.[91]

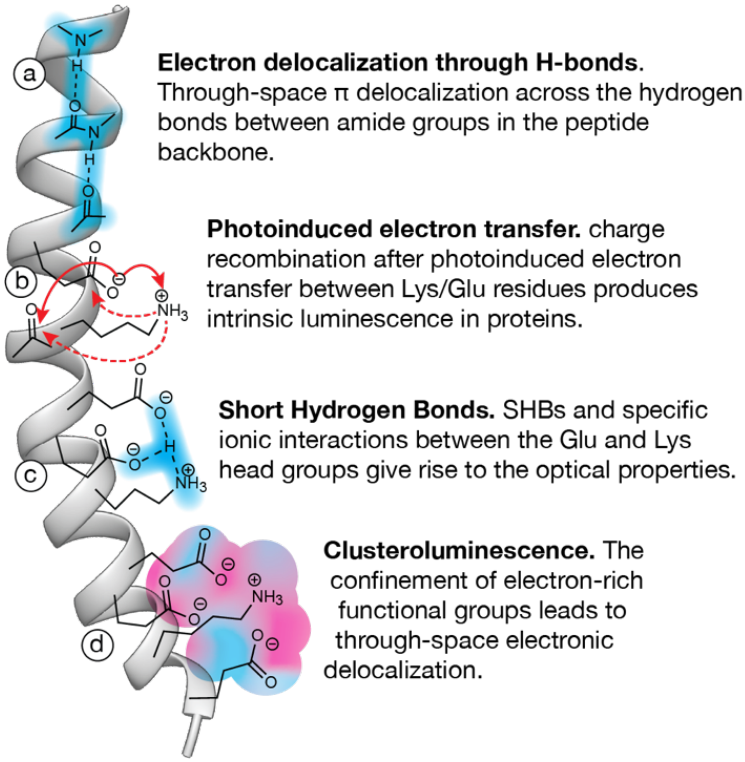


Figure 16. Main mechanisms proposed to explain NAF in peptides, represented as an α-helix cartoon. (a) Electronic delocalization through the hydrogen bonds; (b) photoinduced charge transfer involving the headgroups in Lys/Glu residues and the peptide backbone; (c) formation of short hydrogen bonds (SHBs); (d) clusteroluminescence (CL) mechanism leading to small energy caps by the through-space interaction of electron donor and acceptor groups.

---

[90] (a) Zhang, H. *et al. Materials Today* **32**, 275–292 (2020); (b) Zhang, H., Tang, B. Z. *JACS Au* **1**, 1805–1814 (2021).

[91] (a) Tikhonova, T. N. *et al. Arch. Biochem. Biophys.* **651**, 13–20 (2018); (b) Yakimov, B. P. *et al. Environ. Sci. Technol. Lett.* **9**, 452–458 (2022).

One of the main challenges in developing a unifying theoretical framework for NAF is the widely prevalent belief that it occurs exclusively in aggregated systems. This assumption has led research efforts to focus primarily on β-sheet fibers and oligomeric assemblies that assume important in characterizing them by conventional analytical and structural techniques. Another important limitation is the lack of systematic studies exploring the relationship between peptide sequence, structure and emission properties. Understanding these correlations would provide crucial experimental insights into the factors governing NAF, paving the way for rational design strategies. Advancing the understanding of this phenomenon could bring numerous advantages, such as the development of new bioimaging tools, label-free detection methods, and new biomaterials with tunable optical properties.

3.4. Catalytic Metallopeptides for Bioorthogal Transformations

3.4.1. Metalloproteins: Structure, Function, and Biological Significance

The term *metalloprotein* refers broadly to any protein that contains a metal cofactor in its structure.[92] These proteins represent a substantial part of the human proteome (between one-third and one-half) and play essential roles in cells and whole organisms.[93] In fact, it is estimated that between a quarter and a third of all proteins require the presence of metals to perform their function.

In metalloproteins, metal ions coordinate with amino acid residues to form a metal center, which can act as an enzymatic active site, facilitating essential chemical reactions, or as a structural element, stabilizing protein folding and assembly. Metal binding sites in proteins generally comprise a polar core, where the metal ion is coordinated by nitrogen, oxygen or sulphur centres in the side chains of hydrophilic residues or in the main structure of the peptide, surrounded by a layer of non-polar groups that prevent contact between water and the metal complex.[94] In particular, the imidazole ring of histidine, the thiolate groups of cysteine and the carboxylate groups of aspartates play a central role in the stabilization and reactivity of the metal center.[95]

While metals can play structural and regulatory roles in many proteins, their most prominent role is in enzymatic catalysis.[96] Such metalloproteins are known as metalloenzymes and catalyze a wide variety of biological reactions with remarkable efficiency and selectivity. These include redox reactions, essential in processes such as photosynthesis and nitrogen fixation;[97] detoxification reactions, carried out by hemoproteins such as *cytochromes P-450* (**Figure 17**a)[98] and by antioxidant enzymes such as *superoxide dismutases;*[99] as well as the conversion of CO_2 to bicarbonate, a reaction catalyzed by *carbonic anhydrase*, one of the fastest known enzymes, whose activity depends on a Zn(II) ion in its active site.[100]

---

[92] (a) Banci, L. *Metallomics and the Cell* **12**, 1–13 (Springer, 2013); (b) Shriver, D. F., Atkins, P. W. *Inorganic Chemistry* (3rd ed.), Oxford University Press (1999).

[93] Waldron, K. J., Robinson, N.J. *Nat. Rev. Microbiol.* **7**, 25–35 (2009).

[94] Yamashita, M. M. et al. *Proc. Natl. Acad. Sci. USA* **87**, 5648–5652 (1990).

[95] Messerschmidt, A. *et al. Handbook of Metalloproteins* (Wiley, 2001).

[96] Andreini, C. *et al. J. Biol. Inorg. Chem.* **13**, 1205-1218 (2008).

[97] Lu, Y. *et al. Nature* **460**, 855–862 (2009).

[98] Bernhardt, R. *J. Biotechnol.* **124**, 128–145 (2006).

[99] Perry, J. J. P. *et al. Biochim. Biophys. Acta* **1804**, 245–262 (2010).

[100] Supuran, C. T. *Biochem. J.* **473**, 2023–2032 (2016).

Metalloproteins play a key role in a wide variety of biological processes, including enzymatic catalysis, metal transport and storage, and signal transduction.[101] For example, iron and copper act as redox-active metals in proteins such as *cytochromes,*[102] *iron-sulphur proteins*[103] and *blue copper proteins,*[104] which are essential for cellular respiration and photosynthesis. In addition, proteins such as *hemoglobin* and *myoglobin* incorporate iron into their structure to facilitate oxygen transport and storage (**Figure 17**a),[105] while *hemocyanin*, use binuclear iron and copper sites to carry out this function.[106] Other metalloproteins are specialized in metal regulation and storage, such as *transferrins* that solubilizes and transports Fe(III) into cells,[107] or *ferritins* that store Fe(III) in the form of iron-oxo.[108] Zinc also plays a key role in the regulation of gene expression by stabilizing the conformation of certain proteins through a structure known as the *zinc finger*, a structural module in which a region of the protein folds around this ion to facilitate its function as a transcription factor (**Figure 17**b).[109] Interestingly, despite the wide variety of metals available in the periodic table, nature has evolved to use only a few of them as cofactors in the formation of metalloenzymes essential for life. In fact, most metallic proteins are based on alkali and alkaline earth elements, or on transition metals of the first row, such as Fe, Cu, Mn, Co or Zn,[110] each with specific coordination properties and reactivity. In fact, the nature of the metal ion is precisely what determines both the binding site on the protein, according to their coordination preferences based on the hard-soft theory of acids and bases, and the coordination number and geometry, which are dictated by its oxidation state.[111] Some of these reactions have been adapted to the synthetic chemistry laboratory, driving the development of biocatalysis and its application in chemistry and biotechnology. This has consolidated biocatalysis as an ever-growing field of research, with a significant impact on diverse scientific and industrial areas.[112]

---

[101] (a) Nies, D. H., Silver, S. *Molecular Microbiology of Heavy Metals, Springer-Verlag*, Berlin (2007); (b) Zheng, H. *et al. J. Inorg. Biochem.* **102**, 1765–1776 (2008); (c) Waldron, K. J., Robinson, N. *J Nat. Rev. Microbiol.* **7**, 25–35 (2009); (d) Waldron, K. J. *et al. Nature* **460**, 823–830 (2009).

[102] T. Yamanaka, K. Okunuki *Microbial Iron Metabolism* (Ed.: J.B. Neilands), Elsevier, 349–400, San Diego, CA, (1974).

[103] Cammack, R. *Adv. Inorg. Chem. (Ed.: Cammack, R.), Acad. Press*, 281–322 (1992).

[104] Penfield, K. W. *et al. J. Am. Chem. Soc.* **103**, 4382–4388 (1981).

[105] Momenteau, M, Reed, C. A. *Chem. Rev.* **94**, 659–698 (1994).

[106] (a) Stenkamp, R. E. *Chem. Rev.* **94**, 715–726 (1994); (b) Coates, C J., Decker, H., *Cell. Mo.l Life Sci.* **74**, 293–317 (2017).

[107] Baker, E. N. *Adv. Inorg. Chem. (Ed.: Sykes, A. G.), Acad. Press*, 389–463 (1994).

[108] Harrison, P. M. *et al. Adv. Inorg. Chem.* (*Ed.: Sykes, A. G.), Acad. Press*, 449–486 (1991).

[109] Berg, J. M. *Annu. Rev. Biophys. Biophys. Chem.* **19**, 405–421 (1990).

[110] (a) Andreini, C. *et al. J. Biol. Inorg. Chem.* **13**, 1205–1218 (2008); (b) Pordea, A. *Curr. Opin. Chem. Biol.* **25**, 124–132 (2015); (c) H. Eom, H., Song, W. J. *J. Biol. Inorg. Chem.* **24**, 517-531(2019); (d) Lewis, J. C. *ACS Catal* **3**, 1–22 (2013); (e) Valdez, C. E. *et al. Acc. Chem. Res.* **47**, 3110–3117 (2014).

[111] Thomson, A. J., Gray, H. B. *Curr. Opin. Chem. Biol.* **2**, 155–158 (1998).

[112] (a) Drauz, K., Waldmann, H. *Enzyme Catalysis in Organic Synthesis* (Wiley-VCH Verlag GmbH, 2002); (b) Fessner, W. D., Anthonsen, T. *Modern Biocatalysis* (Wiley-VCH Verlag GmbH & Co. KGaA, 2008).

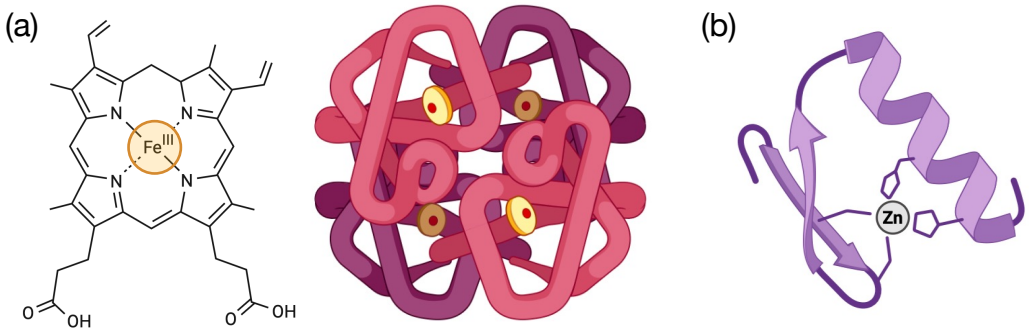


Figure 17. Representation of different metalloproteins and their associated metal ions. (a) Structure of the heme group *(cytochrome P-450*) with an Fe^{3+} ion and the three-dimensional representation of hemoglobin, an oxygen transport protein; (b) structure of a *zinc finger* domain, where the Zn^{2+} ion stabilizes the protein conformation. Image created with BioRender.com.

3.4.2. Development of Bioorthogonal Chemistry

Bioorthogonal chemistry involves reactions that can occur in living systems without interfering with natural biological processes, allowing biological systems to be modulated in a controlled manner.[113] Since its first conceptualization by C. Bertozzi, the field has revolutionized areas such as imaging, drug delivery and biotechnology by allowing the selective activation of molecules within living organisms.[114] In particular, it is very promising in the field of biomedicine, as it could allow us to treat diseases through reactions that natural enzymes cannot carry out.[115]

Precious metals such as ruthenium, gold, palladium and platinum can catalyze reactions under physiological conditions, which has allowed the development of bioorthogonal strategies for controlled drug activation.[116] In this approach, the prodrug is initially inactive due to a masking group, which is selectively removed by the metal catalyst, allowing activation only where it is needed (**Figure 18**). This concept is like antibody-directed enzyme-driven prodrug therapy (ADEPT), proposed in the 1970s, where enzymes were used to activate prodrugs at the site of a tumour.[117] However, unlike ADEPT, the use of metal catalysts avoids immunogenicity issues and broadens the possibilities for therapeutic applications. As a result, metal complexes have proven to be valuable tools for prodrug activation and functional modification of proteins in biological settings.

One of the first examples of metal-catalyzed reactions in biological systems was published in 1985 by Cséplö *et al.*[118] In this study, they used a homogeneous, water-soluble ruthenium complex that was able to hydrogenate the C-C double bonds of fatty acids in the presence of H_2 when incorporated into mesophyll protoplasts. However, the manuscript lacked experimental details, and extensive cellular damage was observed after only 1 h of exposure.

It was not until 2002 that this field began to develop with the discovery by Sharpless and Meldal *et al.* of the copper(I)-catalyzed [3+2]-azide alkyne cycloaddition (CuAAC) reaction,[119] which selectively transforms organic azides and terminal alkynes into the corresponding *trans*-1,4-disubstituted 1,2,3-triazoles under mild reaction conditions and in aqueous solution. This discovery paved the way for their use in different biomedical applications, such

---

[113] (a) Sletten, E., Bertozzi, C. R. *Angew Chemie. Int. Ed.* **48**, 38, 6974–6998 (2009); (b) Bertozzi, C. R. *Acc. Chem. Res.* **44**, 9, 651–653 (2011); (b) Prescher, J. A., Bertozzi, C. R. *Nat. Chem. Biol.* **1**, 1, 13–21 (2005).

[114] Zhang, X. *et al. Trends Chem.* **1**, 90–98 (2019).

[115] Li, J., Chen, P. R. *Nat. Chem. Biol.* **12**, 3, 129–137 (2016).

[116] Martinez-Calvo, M., Mascareñas, J. L. *Coord. Chem. Rev.* **359**, 57–59 (2018).

[117] Bagshawe, K. D., Sharma, S. K., Begent, R. H. J. *Expert. Opin. Biol. Ther.* **4**, 1777–1789 (2004).

[118] Vigh, L., Joó, F. Cséplö, Á. *Eur. J. Biochem.* **146**, 241–244 (1985).

[119] (a) Rostovtsev, V. V. *et al. Angew Chem. Int. Ed.* **41**,2596–2599 (2002); (b) Tornøe, C. W., Christensen, C., Meldal, M. *J. Org. Chem.* **67**, 3057–3064 (2002).

as target identification[120] or for drug discovery and delivery.[121] However, the use of this chemistry in cellular environments has been severely limited due to the toxicity of Cu ions.[122]

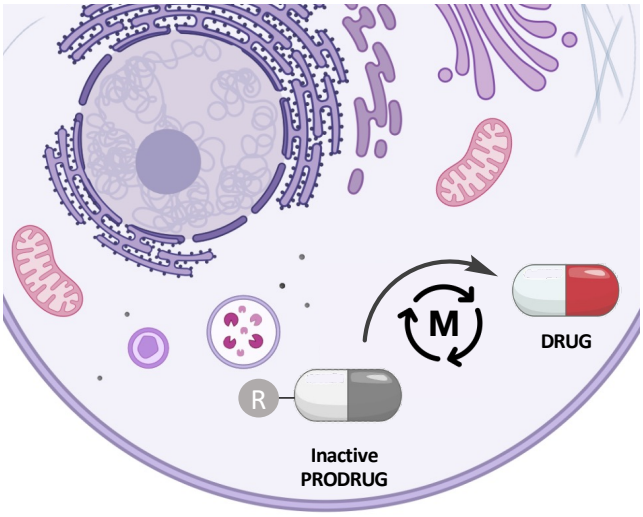


Figure 18. Bioorthogonal strategy for the controlled activation of drugs by metal catalysis. A protecting group on the prodrug (R) is selectively removed in the presence of a metal catalyst (M), allowing activation only at the desired site within the cell. Image created with BioRender.com.

In 2006, Streu and Meggers introduced the first Ru(II)-catalyzed desalination reaction in the presence of thiols, which could be carried out in water, in the air and, most importantly, even in the presence of living mammalian cells.[123] Their study involved the generation of a fluorescent product by cleavage of the allyloxycarbonyl (alloc) group of a rhodamine in the presence of the organometallic Ru(II) complex in HeLa cells (**Figure 19**a). Subsequently, Volker *et al.* followed the same strategy to release bioactive agents.[124]

In addition, it was subsequently demonstrated that specific cellular compartments could be localized through deprotection reactions without affecting the function of the rest of the cell. Mascareñas and co-workers developed a ruthenium complex with triphenylphosphonium-derived ligands, which selectively accumulates in mitochondria, where it can still promote not only the release of the allocyclically protected rhodamine, but also the release of a prodrug that functions as uncoupler of the mitochondrial membrane potential (**Figure 19**b).[125]

---

[120] (a) Su, Y. *et al. Sci. Rep.* **5**, 7724 (2015); (b) Wright, M. H., Sieber, S. A. *Nat. Prod. Rep.* **33**, 681–708 (2016); (c) Nakamoto, K., Akao, Y., Ueno, Y. *Bioorg. Med. Chem. Lett.* **28**, 2906–2909 (2018).

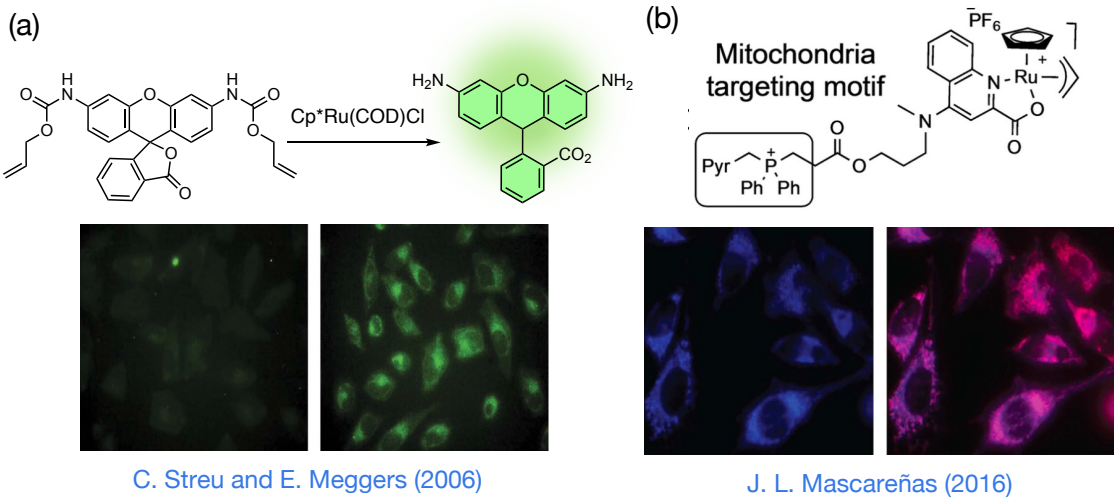
[121] (a) Jiang, X. *et al. Expert. Opin. Drug Discov.* **14**, 779–789 (2019); (b) K. Kacprzak, *et al. Chem. Rev.* **116**, 5689–5743 (2016); (c) Dong, S. *et al. Mol. Pharm.* **16**, 3770–3779 (2019); (d) Dong, S. *et al. ACS Appl. Mater. Interfaces* **11**, 8740–8748 (2019).

[122] (a) Kennedy, D. C. *et al J. Am. Chem. Soc.* **133**, 17993–18001 (2011); (b) D. Soriano del Amo, D. *et al. J. Am. Chem. Soc.* **132**, 16893–16899 (2010).

[123] Streu, C., Meggers, E. *Angew Chem. Int. Ed.* **45**, 5645–5648 (2006).

[124] Volker, T., *Angew Chem. Int. Ed.* **53**, 10536–10540 (2014).

[125] Tomás-Gamasa, M. *et al. Nat. Commun.* **7**, 12538-12548 (2016).



C. Streu and E. Meggers (2006)

J. L. Mascareñas (2016)

Figure 19. Examples of Ru-mediated bioorthogonal catalysis in chemical biology: (a) cleavage of allyl carbamate in biological media, described by Streu and Meggers (2006); (b) subcellular localization and catalytic activity of ruthenium complexes, described by Mascareñas *et al.* (2016) (images adapted form reference [127]).

Orthogonal catalysis with palladium is a powerful tool for modifying biomolecules in biocompatible environments.[126] Pd is one of the most widely used transition metals for both bond formation and bond breaking. One of the most prominent examples is the Suzuki-Miyaura reaction, used in 2009 by Davis and colleagues for protein modification.[127]

Subsequently, Lin *et al.* used the Pd-mediated Sonogashira reaction to functionalize proteins with alkyne functional groups in aqueous solution and in bacteria.[128] Chen *et al.* extended the use of this reaction to tag GFP and other proteins in live enteric pathogens, demonstrating its potential in cell biology studies and specific labeling in complex biological environments.[129]

Among the different Pd-mediated reactions, depropargylation is probably the most widely used in biorthogonal chemistry, with Unciti-Broceta and his team pioneering its application for drug delivery to cells. They developed pro-drugs, such as gemcitabine,[130] 5-fluorouracil,[131] floxuridine[132] and vorinostat,[133] each modified with propargylic groups to generate inactive precursors (**Figure 20**a).[134] In addition, advances in nanostructured materials have expanded the applications of Pd in biological systems. For example, in 2011, also the Unciti-Broceta's

---

[126] Hartings, M. *Nat. Chem.* **4**, 9, 764 (2012).

[127] Chalker, J. M., Wood, C. S. C., Davis, B. G. *J. Am. Chem. Soc.* **131**, 45, 16346–16347 (2009).

[128] Li, N. *et al. J. Am. Chem. Soc.* **133**, 39, 15316–15319 (2011).

[129] Li, J. *et al. J. Am. Chem. Soc.* **135**, 7330–7338 (2013).

[130] Weiss, J. T. *et al. J. Med. Chem.* **57**, 5395–5404 (2014).

[131] Weiss, J. T. *et al. Nat. Commun.* **5**, 3277 (2014).

[132] Weiss, J. T. *et al. Sci. Rep.* **5**, 9329 (2015).

[133] Rubio-Ruiz, B., Weiss, J. T., Unciti-Broceta, A. *J. Med. Chem.* **59**, 9974–9980 (2016).

[134] (a) Hazra, S. *et al. Biochemistry* **49**, 31, 6784–6790 (2010.

group used polymer-embedded palladium nanoparticles to perform cross-coupling reactions in cells and remove allylcarbamate groups. However, the reproducibility of these systems remains a challenge.[135] A major breakthrough came in 2014, when Chen *et al.* achieved the first metal-mediated protein activation in living cells.[136] Their study demonstrated that a genetically modified lysine containing a proc group could be efficiently deprotected by allylPd_2Cl_2, restoring the enzymatic activity of bacterial phosphothreonine lyase OspF.

Beyond fundamental studies, Pd-catalyzed depropargylation has been successfully integrated into theragnostic strategies. Wu *et al.* demonstrated its use in a prodrug that, in the presence of Pd, depropargylates and simultaneously releases a coumarin fluorophore and a potent anticancer drug.[137] By encapsulating the catalyst and precursor in separate liposomes, targeted delivery to tumors in mice was achieved, enabling fluorescence-based detection and effective suppression of tumor growth. Further expanding Pd applications, Weissleder and colleagues explored its role in depopargylation reactions in living environments. However, they found that Pd catalysts are prone to deactivation in complex cellular settings.[138] Meanwhile, the Mascareñas group developed discrete Pd(II) complexes with specifically designed phosphine ligands, achieving a balance between reactivity and stability. These complexes effectively promoted depropargylation reactions and alloc group cleavages in live HeLa and Vero cells (**Figure 20**b).[139]

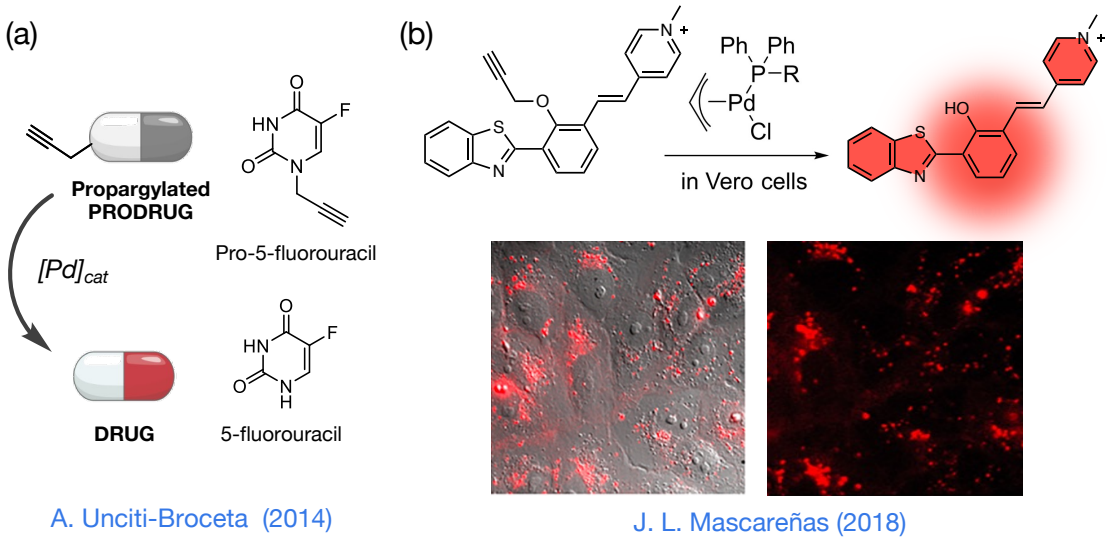


Figure 20. Examples of Pd-catalysed depropargylation in cells and in vivo: (a) General scheme of depropargylation mechanism for drug delivery, described by A. Unciti-Broceta *et al.* (2014); (b) intracellular depropargylation reaction mediated by palladium complexes equipped with phosphine ligands designed by Mascareñas *et al.* (2018) (image adapted form reference [141]).

---

[135] Yusop, R. M. *et al. Nat. Chem.* **3**, 3, 239–243 (2011).

[136] Li, J., *et al. Nat. Chem.* **6**, 352–361 (2014).

[137] Li, B. *et al. Biomaterials* **138**, 57–68 (2017).

[138] Miller, M. A. *et al. Nat. Commun.* **8**, 15906–15919 (2017).

[139] Martínez-Calvo, M. *et al. ACS Catal.* **8**, 6055–6061 (2018).

In addition to ruthenium and palladium, copper has shown great potential in bioorthogonal catalysis. However, as mentioned earlier, its application in living cells remains limited due to its toxicity. To overcome these challenges, researchers have designed azides with chelating groups that stabilize the catalyst and improve reactivity. For example, Ting's group proposed the use of neighboring pyridines, which improve the orthogonality of the reaction and allow the tagging of specific cell surface proteins with bright, photostable fluorophores (**Figure 21**a). However, this approach compromises catalytic turnover, as the product can sequester the copper complex.[140]

An alternative approach has focused on using bulky tristriazole-type ligands to optimize reactivity while reducing the toxicity of copper compounds. A notable example is the work of Tirrell *et al.*, who employed these ligands for the surface labeling of mammalian cells.[141] Similarly, Mascareñas' group developed water-soluble Cu(I) complexes to enhance intracellular CuAAC catalysis (**Figure 21**b).[142] Although these complexes did not show high efficiency, they did manage to solve problems like the excess of ligands or the cytotoxic effects of the ascorbate. While further optimization is needed, these advances represent significant progress for the use of copper catalysts in intracellular applications.

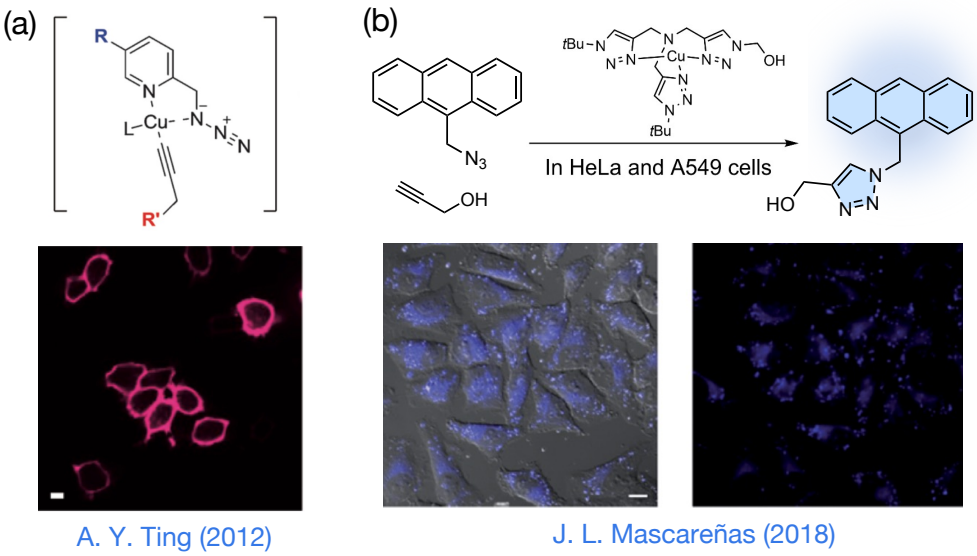


Figure 21. Examples of bioorthogonal CuAAC catalysis: (a) General scheme of the pyridine-driven CuAAC reaction for protein tagging, according to Ting (2012) (image adapted form reference [142]); (b) copper-tristriazole complexes for small-molecule cycloaddition reaction inside mammalian cells, described by Mascareñas (2018) (image adapted form reference [144]).

[140] Uttamapinant, C. *et al. Angew Chem. Int. Ed. Engl.* **51**, 24, 5852–5856 (2012).

[141] Link, A. J., Tirrell, D. A. *J. Am. Chem. Soc.* **125**, 11164–11165 (2003).

[142] Miguel-Ávila, J. *et al. Chem. Sci.* **9**, 1947–1952 (2018).

Other transition metals have also expanded the repertoire of bioorthogonal reactions. NHC-Au(I) complexes have facilitated the release of anticancer drugs via intramolecular reactions.[143] Moreover, in a complementary approach described by Mascareñas *et al.* demonstrated that Au-mediated catalysis enabled the formation of C–C bonds in biological systems, even in combination with ruthenium-catalyzed reactions inside living cells.[144] Similarly, Pt(II) and Pt(IV) complexes have enabled the cleavage of alkyne-bearing masking groups in living systems, facilitating the release of secondary amines or the activation of antibody-drug conjugates.[145] A prominent example is the use of cisplatin to activate 5-fluorouracil, which resulted in tumor reduction in a zebrafish xenograft model.

Transition metal catalysts are therefore excellent candidates for bioorthogonal transformations due to their high catalytic efficiency, a key requirement for the *in-situ* generation of therapeutics.[146] This localized activation allows precise control over when and where a drug is activated, thereby reducing unwanted side effects.[147] However, the use of unmodified transition metal catalysts for prodrug activation faces several challenges, including low water solubility, limited biocompatibility, and poor catalytic stability in biological environments.[148] Although progress has been made with transition metal-based catalysts such as palladium and ruthenium, their implementation in living systems remains limited. Therefore, metal complex engineering offers a possible solution to these challenges and improve their stability and controlled delivery.

[143] Vong, K., *et al. Chem. Sci.* **11**,10928–10933 (2020).

[144] Vidal, C. *et al. Nat. Commun.* **9**, 1913–1923 (2018).

[145] Oliveira, B. L. *et al. J. Am. Chem. Soc.* **142**, 10869–10880 (2020).

[146] Chankeshwara, S. V., Indrigo, E., Bradley, M. *Curr. Opin. Chem. Biol.* **21**, 128–135 (2014).

[147] Völker, T., Meggers, E. *Curr. Opin. Chem. Biol.* **25**, 48–54 (2015).

[148] Zhang, X. *et al. ChemBioChem* **21**, 19, 2759–2763 (2020).

3.4.3. Engineering of Artificial Metalloenzymes

As mentioned above, nature depends on metalloenzymes, which act as catalysts in numerous biological processes. However, their use is limited to substrates, metals and ligands present in cells, which considerably restricts their repertoire. In addition, transition metals are highly susceptible to the biological environment, which hinders their application in biomedicine. To overcome these limitations, efforts have been devoted to the creation of hybrid catalysts, combining synthetic metal catalysts with protein structures.[149] This has given rise to a new field of research known as "*artificial metalloenzymes"* (ArMs).[150]

The first description of an ArM dates back to 1956, when Fuji *et al.* reported that a Pd(II) salt adsorbed on silk fiber catalyzed the enantioselective hydrogenated reduction of amino acids.[151] Although the reproducibility of this experiment could not be confirmed later, this milestone is considered the starting point in the development of ArMs.[152]

A pioneering study was conducted by Kaiser *et al.* in 1976, who used carboxypeptidase A as an oxidase, replacing the Zn(II) center with Cu(II) to catalyze the oxidation of ascorbic acid.[153] Another important study was carried out by Wilson and Whitesides in 1978, who attempted to attach an abiotic metal center to a protein by assembling an artificial hydrogenase, through the incorporation of a biotinylated diphosphine-rhodium complex into the avidin protein.[154] However, the true potential of ArMs was not realized until the development of recombinant protein production techniques. In 1997, Distefano and Davies reported the modification of the scaffold of a recombinant lipid-binding protein in adipocytes by introducing iodoacetamido-1,10-phenanthroline as the Cu(II) chelator to carry out the stereoselective hydrolysis of racemic esters.[155]

Since their conception, the field of ArM has experienced remarkable growth so that ArMs capable of catalysing more than 40 reactions have been described, including desalination,[156] hydroamination, [157] olefin metathesis, [158] phenylacetylene polymerization, [159] carbene

---

[149] Schwizer, F. *et al. Chem. Rev.* **118**, 142–231 (2018).

[150] (a) Davis, H. J., Ward, R. T. *ACS Cent. Sci.* **7**, 1120–1136 (2019); (b) Tanaka, K., Vong, K. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* **96**, 79–94 (2020).

[151] Akabori, S., *et al. Nature* **178**, 323–324 (1956).

[152] Kagan, H. B. *Compr. Asym. Catal.* **1**, 9–30 (1999).

[153] Yamamura, K., Kaiser, E. T. *J. Chem. Soc., Chem. Commun.* **20**, 830–831(1976).

[154] Wilson, M. E., Whitesides, G. M. *J. Am. Chem. Soc.* **100**, 306–307 (1978).

[155] Davies, R. R., Distefano, M. D. *J. Am. Chem. Soc.* **119** 48, 11643–11652 (1997).

[156] Baiyoumy, A. *et al. ACS Catal.* **11**, 10705–10712 (2021).

[157] Christoffel, F. *et al. Nat. Catal.* **4**, 643–653 (2021).

[158] (a) Sauer, D. F., et al. *ACS Catal.* **5**, 7519–7522 (2015); (b) Jeschek, M. *et al. Nature* **537**, 661–665 (2016).

[159] Grimm, A. R. *et al. ACS Catal.* **8**, 2611–2614 (2018).

transfer[160] and transfer hydrogenation.[161] Many of these reactions have no analogues in nature, significantly extending the scope of classical biocatalysis.

Among the strategies used to construct ArMs are: (i) direct coordination of unsaturated metals with basic amino acid residues, such as histidine; (ii) substitution of the native metal in metalloproteins by others with novel catalytic activity; (iii) supramolecular anchoring of metal cofactors by inhibitors or high-affinity substrates; and (iv) covalent immobilization of cofactors to proteins by irreversible bonds.[162] These methodologies have been fundamental to the design of both repurposed and de novo ArMs.

Overall, ArMs possess properties that mimic those of natural enzymes and, in certain cases, could even overcome some of the limitations associated with metal-mediated bioorthogonal catalysis. However, one of the main challenges still to be solved is the poor intracellular activity of ArMs, mainly attributed to the limited cell penetration of proteins, and deactivation caused by intracellular species, such as thiols, which can coordinate and deactivate the metal center.[163]

Within this field, catalytic metallopeptides emerge as an intermediate option between conventional metal catalysts and protein-based ArMs (**Figure 22**). Metallopeptides become very attractive scaffolds for intracellular catalysis, as they are large enough to fold into structures that protect the metal center from reactive metabolites, but with a structure that is much simpler and easier to manipulate. Moreover, peptides are biocompatible, modular and easily synthesized by solid-phase methodologies. Therefore, in this context, peptide engineering will play a key role in optimizing these systems, providing them with biomimetic characteristics and improving their catalytic performance.

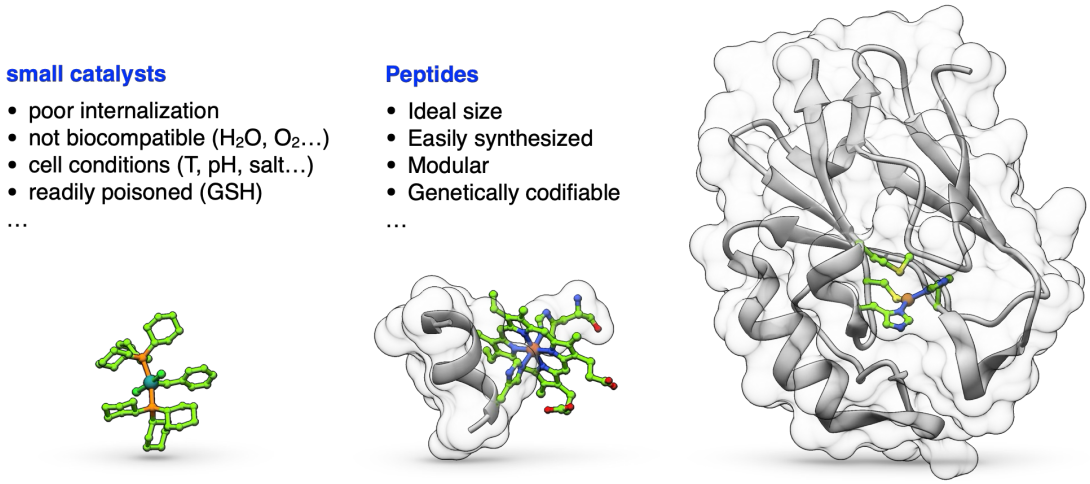


Figure 22. Comparison between traditional organic catalysts, ArM and catalytic metallopeptide.

---

[160] (a) Yu, K. *et al. J. Am. Chem. Soc.* **145**, 16621–16629 (2023); (b) Gu, Y. *et al. Angew Chem. Int. Ed.* **61**, e202110519 (2022); (c) Liu, Z. *et al. J. Am. Chem. Soc.* **144**, 883–890 (2022).

[161] Stein, A. *et al. ACS Cent. Sci.* **7**, 1874–1884 (2021).

[162] Schwizer, F. *et al. Chem. Rev.* **118**, 142–231 (2018).

[163] Sabatino, V., Unnikrishnan, V. B., Bernardes, G. J. L. *Chem. Catalysis* **2**, 39–51 (2022).

4. OBJECTIVES

Historically, peptides have been used in the pharmaceutical industry and in biomolecular research for various applications. However, recent advances have demonstrated their potential in areas such as enzyme catalysis and the development of fluorescent sensors.

The main objective of this thesis is to **explore and optimize the design of peptides with applications in fluorescence and catalysis**, to improve their performance and versatility in different applications. To this end, computational design methodologies, chemical synthesis and structural and functional characterization of the peptides obtained will be employed.

- In view of the precedents and current progress in the development of new non-aromatic fluorescent materials (NAFs), the main objective of the first chapter is the design and synthesis of a set of zwitterionic peptides based on the single α-helix (SAH) model as simple peptide platforms that could generate NAFs. Through an approach combining chemical synthesis and spectroscopic characterization, we aim to optimize the fluorescence intensity and stability of these systems. In addition, the influence of the peptide structure on the optical properties is investigated, providing key information for the design of new fluorophores with applications in biotechnology and analytical chemistry.
- In the second chapter, we seek to develop palladium-based artificial metalloenzymes by designing and engineering β-sheet proteins and explore their potential as intracellular catalysts in living cells. Through a combination of experimental studies and theoretical modelling, we seek to establish design principles that will enable the creation of more efficient and selective peptide catalysts for applications in green chemistry and sustainable synthesis.
- In the third chapter, the main objective is the development of a combinatorial methodology to obtain catalytic metallopeptides in an efficient and fast way. For this purpose, we will leverage the SPOT methodology, which allows us to generate a manageable peptide library organized by positions.

5. RESULTS AND DISCUSSION

5.1. Anomalous Emission from Single α-Helical Peptides in Solution

Summary

Traditionally, the fluorescence of proteins has generally been attributed to the presence of aromatic residues or prosthetic groups in their sequence. However, recent studies have revealed photoluminescence in protein and peptide aggregates without aromatic or conjugated structures. This phenomenon, known as *Non-Aromatic Fluorescence* (NAF), challenges the idea that defined fluorophores are necessary for luminescence in organic systems. Here, we describe the first systematic analysis of NAF-emitting peptides in solution and show that short peptides derived from zwitterionic single α-helices (SAHs), formed exclusively by non-aromatic lysine and glutamic acid residues, are UV-active and luminescent at near-UV wavelengths in solution (λ_{exc} = 320 nm; λ_{em} ≈ 440 nm). We also show that their emission depends on the α helical folding, which favors intramolecular through-space interactions between the Lys/Glu side chains and that conservative mutations, such as the replacement of Lys by Orn or Arg and the incorporation of aromatic residues strongly influence the NAF emission.

Results of this section have been accepted to be published in Cell Reports Physical Science:

Carmen González-González,[a] Roi Lopez-Blanco,[a] Juan A. González-Vera,[c] Sara D'Ingiullo,[a] David Bouzada,[a] Manuel Melle-Franco,[b] Ángel Orte,[c] M. Eugenio Vázquez[a]

Identification code: CR-PHYS-SCI-D-25-00062R4

Affiliations:

a. Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Departamento de Química Orgánica, Universidade de Santiago de Compostela. Santiago de Compostela 15782, Spain.

b. CICECO - Aveiro Institute of Materials, Department of Chemistry, University of Aveiro, Aveiro 3810-193, Portugal.

c. Departamento de Fisicoquímica, Unidad de Excelencia de Química Aplicada a Biomedicina y Medioambiente, Facultad de Farmacia, Universidad de Granada, Campus la Cartuja, Granada, 18071, Spain.

5.1.1. Background

Since the 1950s, it has been accepted that the photoluminescent properties in (poly)peptides are due to the presence of aromatic cofactors or residues in their sequence. However, in contrast to this dogma, a handful of studies have reported emission in the aggregate state of proteins and peptides lacking aromatic or conjugated functionalities, a phenomenon referred to as *Non-Aromatic Fluorescence* (NAF). The odd nature of NAF, a fascinating fundamental research challenge that defies the commonly accepted views of photoluminescence, together with the lack of systematic studies in solution exploring the relationship between the sequence/structure of polypeptides and their anomalous emission properties, has hindered our fundamental understanding of this phenomenon.

Previous reports on NAF peptides provide useful information for designing optimal emissive structures. In 2004, R. Swaminathan's team found that NAF emission in Lys-rich proteins decreased significantly upon unfolding, suggesting that a stable peptide structure capable of clustering amino acid side chains to promote through-space interactions could lead to the creation of emissive species.[164] Later, in 2020, the same group demonstrated that the anionic carboxylate and cationic ammonium side chains of Lys/Glu residues act as electron donors and acceptors, facilitating photoinduced electron transfer and resulting in weak NAF emission.[165] These recent advancements have shed light on the role of interactions in polymers and peptides in achieving unique emissive properties. Building on these findings, M. Guo showed that increasing ionic and hydrogen bond interactions in NAF polymers leads to red-shifted emission through the formation of emissive clusters.[166] This highlights the potential of non-covalent interactions as a key strategy for tuning fluorescence properties in polymeric and peptide-based systems.

Taking all these precedents into account, the α-single helix (SAH), also known as the α-helical E-R/K motif, could be a simple peptide platform that would allow us to combine all the described requirements to produce efficient NAF emission. SAHs are formed by repeats of four consecutive anionic (Glu) and four cationic (Lys/Arg) residues that support a stable α-helical conformation through a network of salt bridges in the absence of additional tertiary interactions.[167] Systematic structural studies of de novo E-R/K peptides show that the zwitterionic SAH motif tolerates the incorporation of other residues,[168] making it a versatile platform for the design of CL peptides combining charged Lys/Glu in structured α-helices.

---

[164] Homchaudhuri, L., Swaminathan, R. *Bull. Chem. Soc. Jpn* **77**, 765–769 (2004).

[165] Kumar, A. *et al. Biophys. J.* **118**, 468a (2020).

[166] Huang, Q. *et al*. *Macromol. Rapid Commun.* **42**, (2021).

[167] (a) Sivaramakrishnan, S. *et al. Proc. Natl. Acad. Sci. U. S. A.* **105**, 13356–13361 (2008); (b) Marqusee, S., Baldwin, R. L. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 8898–8902 (1987).

[168] Wolny, M. *et al. Sci. Rep.* **7**, (2017).

5.1.2. Objective

Non-aromatic fluorescence (NAF) in polypeptides has emerged as a promising tool for bioimaging, yet its underlying mechanisms remain poorly understood. Previous studies suggest that structured peptide environments, such as Lys-rich proteins, enhance fluorescence by stabilizing charge-transfer interactions.

To explore this, we focused on the single α-helix (SAH), or E-R/K α-helical motif, as an ideal peptide platform for efficient non-aromatic fluorescence (NAF) emission. SAHs are characterized by repeats of four anionic (Glu) and four cationic (Lys/Arg) residues, which form α-helical structures stabilized through salt bridges. Furthermore, this motif has been previously described for the incorporation of additional residues into the chain, making it a platform with high potential for designing NAF peptides. To explore its properties, we will synthesize a model SAH, $(E_4K_4)_4$, using microwave-assisted solid-phase peptide synthesis, to evaluate its structural and fluorescent properties to establish a foundation for engineered NAF emitters in bioimaging.

5.1.3.1. NAF Emission in SAH

α-Helices have been recognized as a common protein secondary structure since their identification by Pauling *et al.*[169] Single α-helix (SAH) unlike other helices, remain monomeric and structurally stable across various salt concentrations and pH levels without requiring tertiary interactions. [170] Although many potential SAHs have been identified in diverse proteins, only a few have been extensively studied experimentally.[171]

SAHs typically form in sequences rich in acidic (Glu) and basic (Arg, Lys) residues, lacking a hydrophobic seam, and featuring extensive intrahelical salt bridges (E–K or E–R). Studies have identified the E–R/K α-helices as the most stable SAH, which are formed by repeats of four consecutive anionic (Glu) and cationic (Lys/Arg) residues, which maintain their conformation through a robust salt bridge network without requiring tertiary interactions., as shown in the helical network diagram in **Figure 23**.[172]

First, we synthesized a 32-residue model SAH, $(E_4K_4)_4$, using standard

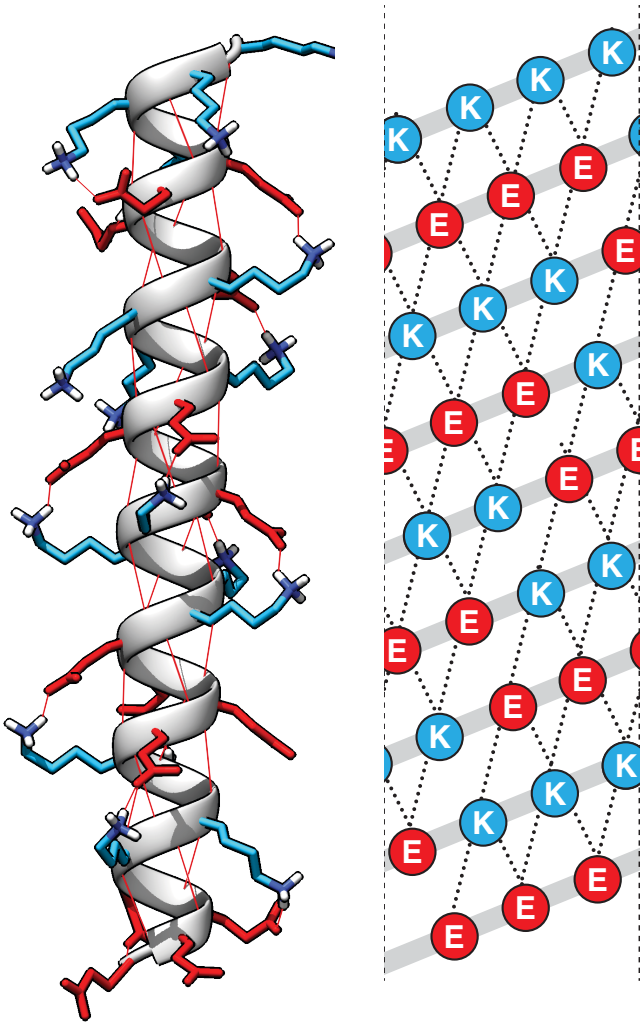


Figure 23. Structure of the peptide $(E_4K_4)_4$ as proposed by RoseTTAFold through the Robetta structure prediction service, and helical net diagram of the α-helix showing the stabilizing salt bridges as dashed lines.

[169] Pauling, L., Corey, R. B., Branson, H. R. *Proc. Natl. Acad. Sci. USA* **37**, 205–211 (1951).

[170] Wolny, M. *et al. J. Biol. Chem.* **289**, 27825–27835 (2014).

[171] (a) Suveges, D. *et al. Proteins* **74**, 905–916 (2009); (b) Sivaramakrishnan, S. *et al. Biophys. J.* **97**, 2993–2999 (2009); (c) Spink, B. J. *et al. Nat. Struct. Mol. Biol.* **15**, 591–597 (2008).

[172] (a) Dunnill, P. *Biophys. J.* **8**, 865–875 (1968); (b) Kim, D. E., Chivian, D., Baker, D. *Nucleic Acids Res.* **32**, (2004).

microwave-assisted SPPS methods,[173] as a simple peptide platform with NAF emission.

Since $(E_4K_4)_4$ lacks conventional chromophores that allow its quantification by UV spectroscopy, it is necessary to find a method to accurately determine its concentration. For this purpose, it was decided to calibrate the peptide concentration by ^{1}H NMR spectroscopy using maleic acid as an external standard in a sealed glass capillary, since it had been previously described by Orfi *et al.* [174] The external standard for quantification of the purified peptide offers the advantage of recovering analytes without contamination. Therefore, qNMR with external standards is more suitable for the concentration calculation of our samples.[175]

First, a relationship was established between the signal of the external standard, maleic acid, and the concentration of a known standard sample. In this case, valine was selected as it had well-defined and non-overlapping signals for its protons (**Figure 24**a). The peptide solution was then analyzed by quantifying the signal intensity of the Cα proton and comparing it with that of maleic acid, the concentration of which is known and allows us to establish a relationship with the concentration of the peptide stock solution (**Figure 24**b).

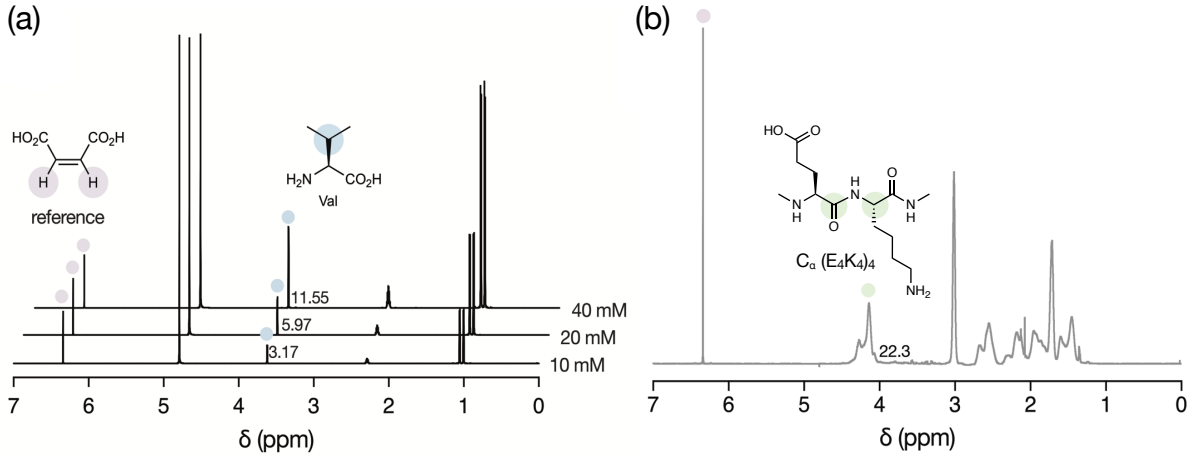


Figure 24. (a) ^{1}H NMR spectrum of 10, 20 and 40 mM of valine solution, using a 50 mM solution of MA as external patron in D_2O; (b) ^{1}H-NMR spectrum of $(E_4K_4)_4$ with maleic acid as internal standard in D_2O.

Once the concentration of our peptide stock solution was determined, we proceeded to measure the structural and spectral properties. As expected for a model SAH sequence, the circular dichroism spectrum of $(E_4K_4)_4$ showed a clear α-helical signature (**Figure 25**a), with an estimated helical content of ~42% based on the intensity of the 222 nm band.[176]

---

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

[174] Larive, C. K., Jayawickrama, D., Orfi, L. *Appl. Spectrosc* **51**, 1531–1536 (1997).

[175] Burton, I. W., Quilliam, M. A., Walter J. A. *Anal Chem* **77**, 3123–3131 (2005).

[176] Rohl, C. A., Baldwin, R. L. *Biochemistry* **36**, 8435–8442 (1972).

However, the most remarkable feature is observed when analyzing the spectroscopic properties of this peptide. Unexpectedly, $(E_4K_4)_4$ exhibited a near-UV absorption maximum at 270 nm and a broad shoulder around 320 nm (**Figure 25**b), despite lacking a traditional chromophore.

Next, the excitation-emission matrix (EEM) was analyzed, which consists of a 3D scan that generates a contour plot of excitation wavelength versus emission wavelength versus fluorescence intensity. EEMs are plotted in Raman units (RU), as they are quantitatively independent of instrument specificities and therefore comparable with measurements from other instruments or from the same instrument, but with different configurations. [177] Interestingly, in the EEM of $(E_4K_4)_4$ a distinct long-wavelength NAF emission band at 440 nm was revealed after excitation at 320 nm (0.47 RU) (**Figure 25**c). This emission had a quantum yield (ϕ) of 0.022 ± 0.002 and a luminescence lifetime (τ) of 4.42 ± 0.12 ns, values consistent with those reported for other NAF polypeptides.[178]

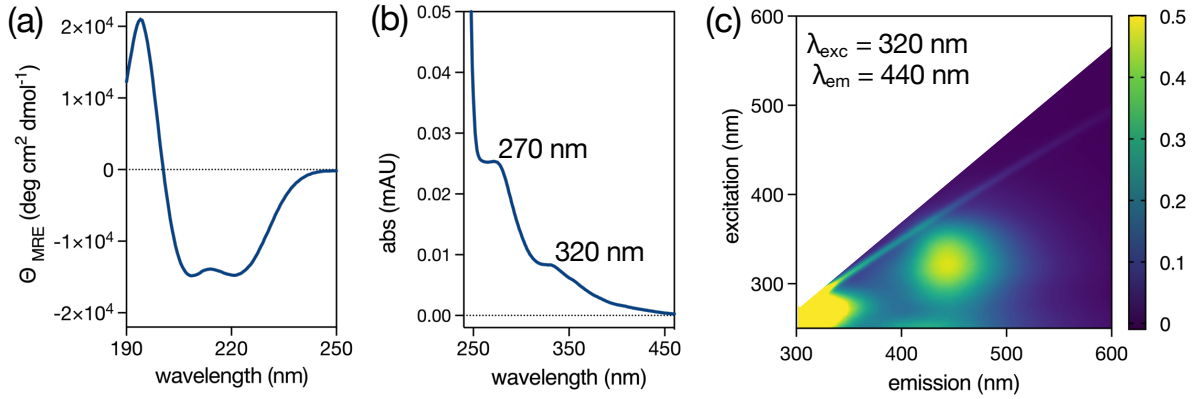


Figure 25. Structural and fluorescence characterization of $(E_4K_4)_4$. (a) Circular dichroism spectra of a 10 μM solution; (b) UV/Vis absorption of a 100 μM solution; (c) Excitation-Emission Matrix (EEM) in RU of a 50 μM solution. All experiments in 25 mM HEPES 150 mM NaCl pH 7.5 buffer at 20 °C.

Having observed NAF emission from the model SAH, $(E_4K_4)_4$, we wondered whether such emission would be affected by the helical content, and thus we synthesized a longer SAH peptide, $(E_4K_4)_9$, and a control zwitterionic peptide, $(E_2K_2)_8$, known to adopt random coil configuration.[179]

$(E_4K_4)_9$ exhibited a broad, intense emission band centered around 420 nm upon excitation at 330 nm (2.18 RU) (**Figure 26**a), with a luminescence lifetime (τ) of 5.08 ± 0.02 ns, comparable to that of $(E_4K_4)_4$. In contrast, the random coil peptide $(E_2K_2)_8$ showed significantly weaker long-wavelength emission (0.08 RU) (**Figure 26**b).

---

[177] Lawaetz, A. J., Stedmon, C. A. *App. Spectrosc.* **63**, 8 (2009).

[178] (a) Cadeddu, M. *et al*. *Macromolecules* **57**, 514–527 (2024); (b) Alom, S. E., Swaminathan R. *Phys. Chem. Chem. Phys.* **25**, 16626–16642 (2023).

[179] Baker, E., Bartlett, G., Crump, M. *et al. Nat. Chem. Biol.* **11**, 221–228 (2015).

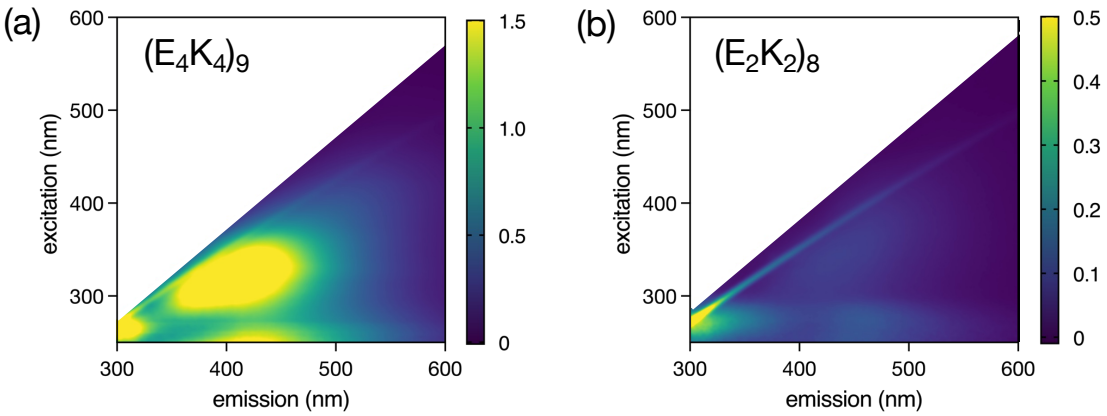


Figure 26. Excitation-Emission matrices (in Raman Units, RU) of 50 µM solutions in 25 mM HEPES buffer 150 mM NaCl, pH 7.5 at 20 °C. of peptides (a) $(E_4K_4)_9$ and (b) $(E_2K_2)_8$. Note the different scale for $(E_4K_4)_9$.

Moreover, this behavior clearly correlated with the absorption profile of each of these structures. In the case of $(E_4K_4)_9$, a much more intense absorption band was observed at both 270 and 320 nm, together with the appearance of a new signal around 350 nm. However, for the random coil type structure, no absorption is observed at any wavelength (**Figure 27**a).

Notably, the plot of the NAF emission intensity per residue versus the calculated average ellipticity of the residue showed a clear correlation (**Figure 27**b and c), highlighting the crucial role of the α-helical secondary structure and intramolecular clustering in the formation of emissive Lys/Glu clusters. This is consistent with the shorter distance between the side chains of Glu and Lys residues in the more compact and organized α-helical conformation, which favors the formation of hydrogen bonds and electronic interactions across space, essential to sustain the NAF phenomenon.

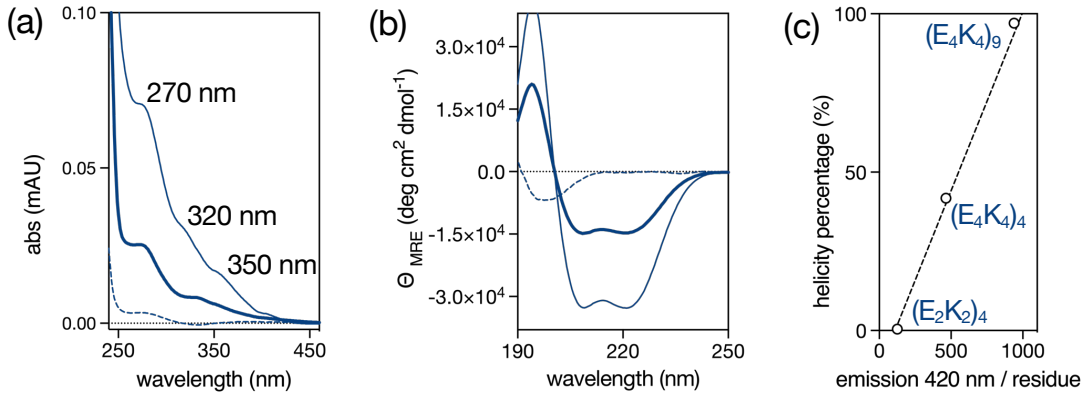


Figure 27. (a) UV/Vis of 100 µM solutions and (b) circular dichroism of 10 µM solutions of $(E_4K_4)_4$ (thick solid line), $(E_4K_4)_9$ (solid line), and $(E_2K_2)_8$ (dashed line); (c) correlation between the normalized emission intensity at 420 nm per residue with the calculated helicity. All experiments in 25 mM HEPES 150 mM NaCl pH 7.5 buffer at 20 °C.

Encouraged by these promising results, we decided to further investigate the properties of SAH by synthesizing modified versions of $(E_4K_4)_4$. Our first approach was to replace all cationic lysine residues in the model sequence of SAH $(E_4K_4)_4$ with ornithine, $(E_4O_4)_4$, a non-essential amino acid structurally similar to lysine, but with a shorter side chain as it contains one less methylene group.

This modification had virtually no impact on the secondary structure of the peptide, maintaining a helical content of around 33%, as can be seen in **Figure 28**a. But, interestingly, $(E_4O_4)_4$ shows both an absorption and NAF emission band with a slightly longer excitation wavelength (λ_{exc} = 330 nm) than that of the peptide $(E_4K_4)_4$ (λ_{exc} = 320 nm) and a significantly higher NAF emission intensity, reaching 0.53 RU (**Figure 28**b and c).

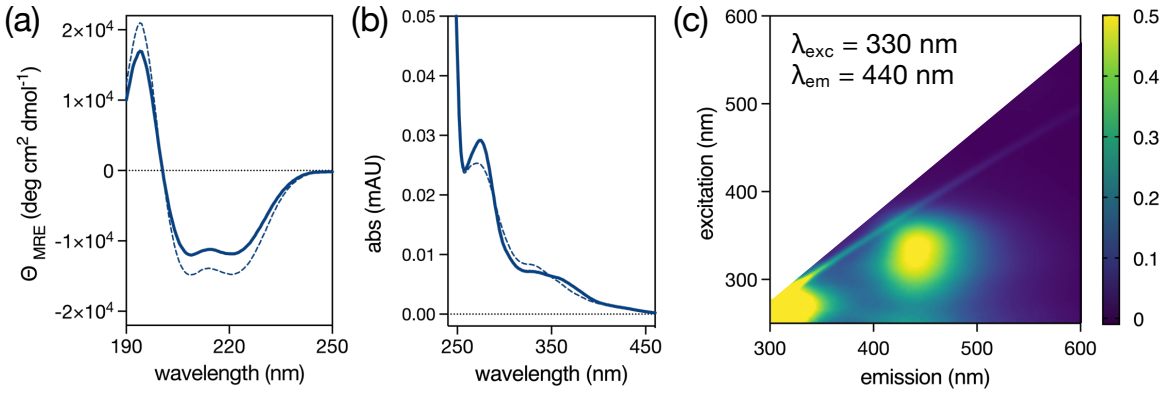


Figure 28. (a) Circular dichroism spectra of a 10 µM solution of $(E_4K_4)_4$ (dashed line) and $(E_4O_4)_4$ (thick solid line); (b) UV/Vis absorption of a 100 µM solution of $(E_4K_4)_4$ (dashed line) and $(E_4O_4)_4$ (thick solid line); (c) EEM in RU of a 50 µM solution of $(E_4O_4)_4$. All experiments were measured in 25 mM HEPES 150 mM NaCl pH 7.5 buffer at 20 °C.

However, our most remarkable result was obtained by replacing all lysine residues with arginine (R), giving rise to the peptide $(E_4R_4)_4$. As expected, this new peptide fully retained the helical structure when compared to our reference peptide, $(E_4K_4)_4$ (**Figure 29**a).

Most notably, analyses by UV/vis and EEM spectroscopy revealed that this replacement generates a NAF fluorescence-emitting peptide, with red-shifted excitation and emission wavelengths (λ_{exc} = 370 nm, λ_{em} = 450 nm) and significantly enhanced emission intensity, achieving 0.97 RU (**Figure 29**b and c).

Due to this substantial improvement in emission, it was decided to evaluate its behavior in living cells by confocal microscopy. For this purpose, HeLa cells were incubated for 1 hour with a 50 µM dilution of the peptide. As seen in the confocal images (**Figure 29**d), its efficient uptake in live cells was confirmed, highlighting its potential as a label-free fluorescent probe for biological imaging.

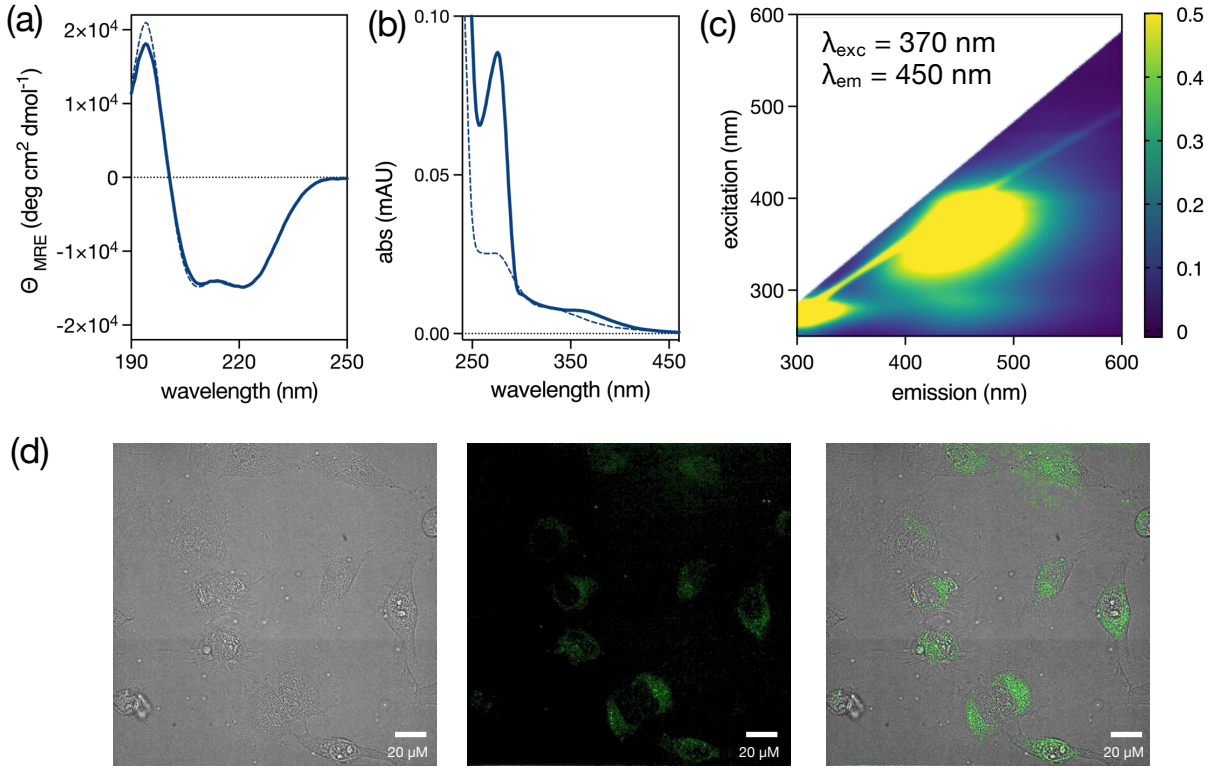


Figure 29. (a) Circular dichroism spectra of a 10 μM solution of $(E_4K_4)_4$ (dashed line) and $(E_4R_4)_4$ (thick solid line); (b) UV/Vis absorption of a 100 μM solution of $(E_4K_4)_4$ (dashed line) and $(E_4R_4)_4$ (thick solid line); (c) EEM in RU of a 50 μM solution of $(E_4R_4)_4$. All experiments in 25 mM HEPES 150 mM NaCl pH 7.5 buffer at 20 °C. (d) Fluorescence microscopy of HeLa cells incubated with a 50 μM solution of $(E_4R_4)_4$ for 1 h.

In order to explore the versatility of this new fluorescence, we synthesized a new peptide sequence, $(E_3YK_3Y)_4$, which incorporates Tyr residues appropriately spaced every (*i, i+4*) residues. This arrangement aims to generate a continuous strip of aromatic side chains aligned along the α-helix, which could favor orbital overlap and π-π interaction between the aromatic rings (**Figure 30**a).

As expected, $(E_3YK_3Y)_4$ exhibited the characteristic emission band of tyrosine around 300 nm. In addition, it also showed an emission band centered at 400 nm after excitation at 300 nm, which could be due to the excitonic effect of the chain-aligned tyrosine residues. But more importantly, we found an anomalous emission band centered at 430 nm after excitation at 340 nm (0.70 RU) (**Figure 30**b). Moreover, this behavior was also reproduced in the Arg analog, $(E_3YR_3Y)_4$, where we found another anomalous band, this time centered at 450 nm after excitation at 370 nm (0.81 RU), equivalent to that observed for the peptide without aromatic residues (**Figure 30**c). These results suggest that the 400 nm band probably arises from tyrosine-tyrosine interactions facilitated by helical folding, and that the unaltered NAF emission highlights its robustness to sequence modifications.

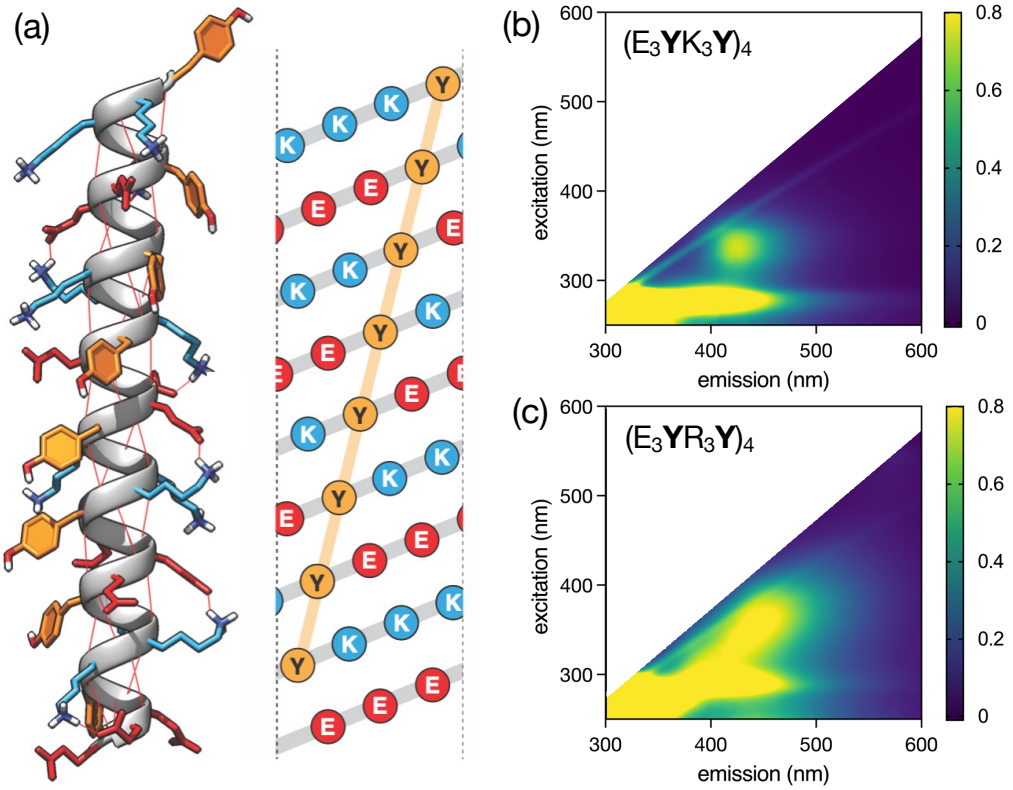


Figure 30. (a) Structure of the peptide $(E_3YK_3Y)_4$ and helical net diagram of the α-helix showing the continuous strip of aromatic side chains aligned along the α-helix; EEM in RU of a 50 μM solution of (b) $(E_3YK_3Y)_4$ and (c) $(E_3YR_3Y)_4$ both in 25 mM HEPES 150 mM NaCl pH 7.5 buffer at 20 °C. Note the different scale.

To test whether structural organization was key, we synthesized a sequence of 32 residues in which tyrosine residues were incorporated every three positions (EE**Y**EK**Y**KK**Y**…, Y3), thus altering the typical helical geometry (**Figure 31**a). In this unstructured variant, the anomalous band disappeared completely (**Figure 31**b).

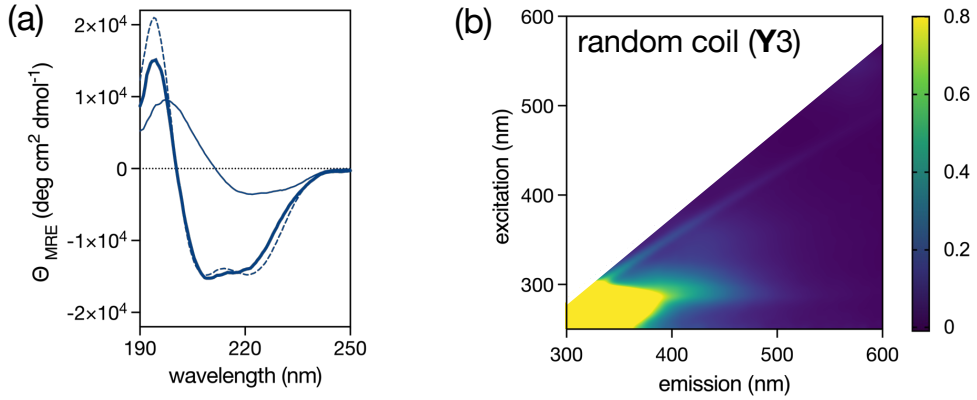


Figure 31. (a) Circular dichroism of 10 μM solutions of $(E_3YK_3Y)_4$ (thick solid line), random coil Y3 (solid line), and $(E_4K_4)_4$ (dashed line); (b) EEM in RU of a 50 μM solution of random coil Y3, both experiments in 25 mM HEPES, 150 mM NaCl, pH 7.5.

Our approach demonstrates how the rational design of peptide sequences allowed us to explore clusteroluminescence, opening new opportunities to develop simple but powerful optical probes based on fully synthetic peptides.

5.1.3.2. Control Experiments of Anomalous NAF Emission

To confirm that the source of this anomalous emission was indeed from the SAH peptides in their folded conformation, we performed a series of control experiments. First, we evaluated how the helical structure influences the fluorescence emission. A particularly revealing result was obtained by thermal unfolding of the peptide $(E_4K_4)_9$, during which the fluorescence maximum and circular dichroism were simultaneously monitored along a temperature gradient (10-90 °C). As shown in **Figure 32**, progressive increases in temperature lead to a loss of the circular dichroism signal ($T_m \approx 60$ °C according to the first derivative), indicating a clear denaturation and thus loss of the helical structure of the peptide (**Figure 32**a). In parallel, a continuous decrease in emission intensity is observed at 420 nm, showing that NAF fluorescence is directly related to the folded state of the peptide (**Figure 32**b). These results underline the importance of a well-defined helical structure for efficient emission.

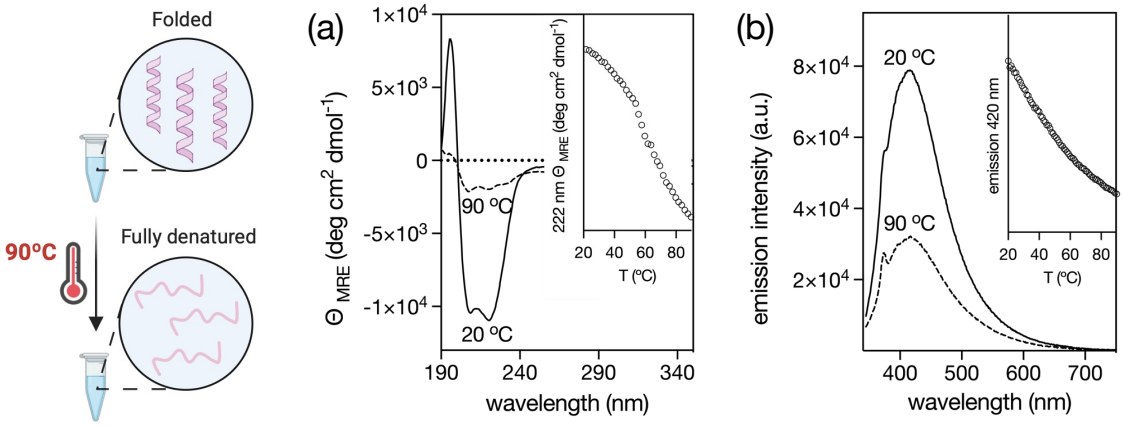


Figure 32. (a) Fluorescence spectra of 50 µM $(E_4K_4)_9$ at 90 °C (solid line) and 20 °C (dashed lines); inset shows emission intensity vs. temperature; (b) circular dichroism spectra under the same conditions, with the inset showing the intensity of the band at 222 nm vs. temperature; both in 25 mM HEPES 150 mM NaCl pH 7.5 buffer.

Further experiments were carried out in support of this hypothesis. Firstly, the effect of pH on the helical structure of the peptide $(E_4K_4)_4$ was evaluated, since its conformational stability is highly dependent on it. When analyzing helicity at both acidic and basic pH, a significant reduction of the secondary structure was observed (**Figure 33**a). Consistently, the EEMs showed a notable loss of the NAF emission band, with intensities reduced to 0.21 RU at pH 2 and 0.19 RU at pH 12 (**Figure 33**b).

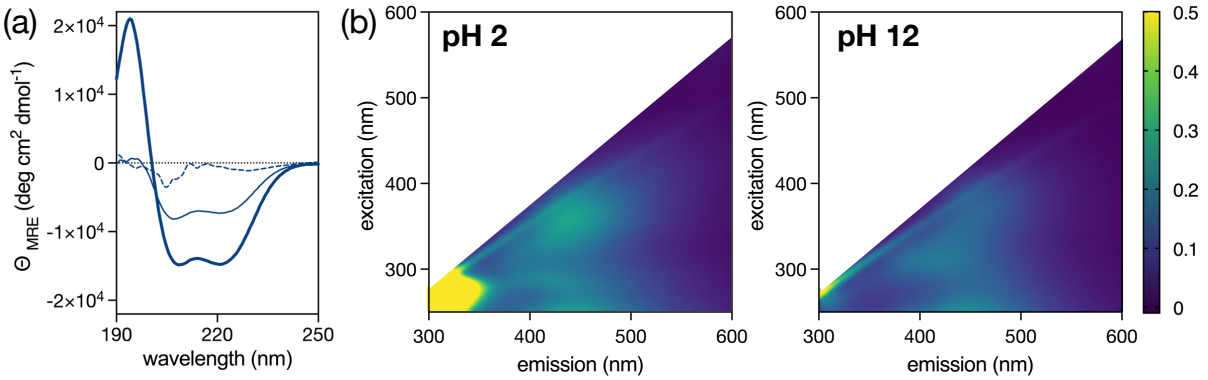


Figure 33. (a) Circular dichroism of 10 μM solutions of $(E_4K_4)_4$ at pH 7.5 (thick solid line), pH 2 (solid line), and pH 12 (dashed line); (b) EEM in RU of a 50 μM solution of $(E_4K_4)_4$ at pH 2 and pH 12, in 25 mM HEPES, 150 mM NaCl.

Additionally, an experiment was performed with guanidinium acetate (GndAcO), a chaotropic agent known to induce protein unfolding. At high concentrations (6.0 M), guanidinium effectively disrupts the secondary structure of proteins, promoting their transition to a random coiled-coil conformation. In this context, a solution of $(E_4K_4)_4$ (50 μM) in 6.0 M GndAcO was analyzed, again showing both loss of helicity (**Figure 34**a) and disappearance of NAF fluorescence (**Figure 34**b).

Taken together, these results reinforce the close relationship between helical folding of the peptide and the appearance of NAF fluorescence, suggesting that this emission is critically dependent on the structural integrity of the α-helix.

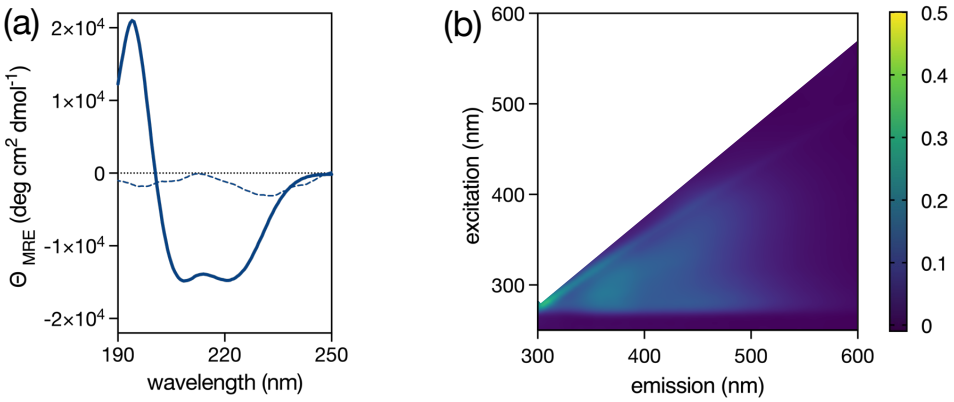


Figure 34. (a) Circular dichroism of 10 μM solutions of $(E_4K_4)_4$ at standard buffer (thick solid line) and in 6.0 M GndAcO buffer (dashed line); (b) EEM in RU of a 50 μM solution of $(E_4K_4)_4$ in 6.0 M GndAcO buffer.

To better understand this phenomenon, a new experiment was performed to investigate the role of hydrogen bonds in NAF emission. In this experiment, the emission of our model peptide $(E_4K_4)_4$ was measured in HEPES buffer, but this time prepared with D_2O. Interestingly, a clear reduction in fluorescence intensity was observed upon hydrogen/deuterium exchange,

generating a band with an intensity of 0.29 RU (**Figure 35**a). This suggests that hydrogen bonds could play a role in stabilizing the excited state responsible for NAF emission.

Also, the effect of ionic strength on the emission was investigated. For this, the emission of $(E_4K_4)_4$ was measured at different salt concentrations, and it was found that the emission increases at high salt concentrations (**Figure 35**b). This enhancement may be explained by salt-induced screening of electrostatic interactions, which could stabilize specific conformations or modulate hydrogen bond networks within the helix. Together, these findings highlight the sensitivity of NAF emission to the local physicochemical environment, particularly to factors that modulate non-covalent interactions such as hydrogen bonds and salt bridges.

Moreover, it is noteworthy that the emission intensity al 440 nm of $(E_4K_4)_4$ showed a linear response with concentration, demonstrating that the emission is not a result of intermolecular aggregation processes (**Figure 35**c). This confirms that the observed emission is an intrinsic property of the folded peptide structure, rather than a result of intermolecular aggregation processes. It should also be noted that these experiments were replicated using the modified sequences, including both the aromatic and non-aromatic variants, and in all cases consistent results were obtained, reinforcing the robustness of the phenomenon across different peptide designs.

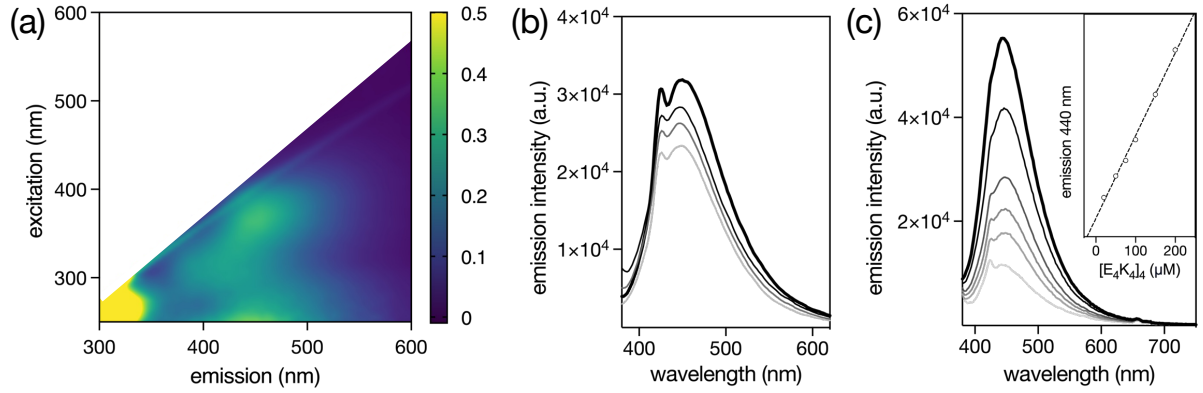


Figure 35. (a) EEM in RU of a 50 μM solution of $(E_4K_4)_4$ in 25 mM HEPES, 150 mM NaCl pH 7.5 prepared in D_2O; (b) emission spectra of 50 μM solutions of $(E_4K_4)_4$ at different salt concentrations in a 25 mM HEPES buffer pH 7.5, 150 mM NaCl (thick solid line) and thinner lines in progressively darker shades of gray at 0, 50 and 100 mM NaCl; (c) emission of $(E_4K_4)_4$ at 20, 50, 75, 100, 150 and 200 μM (from light grey to black) in 25 mM HEPES pH 7.5, 150 mM NaCl at 20 °C; inset shows the plot of emission intensity vs. concentration and the best linear fit (R^2 = 0.9950).

Finally, an important result was obtained by incubating a 50 μM solution of the peptide $(E_4K_4)_9$ with proteinase K for 20h at 37 °C, a broad-spectrum subtilisin-related serine protease. [180] The emission at 420 nm was monitored over time (0, 4, 8 and 20 h), and a progressive decrease in fluorescence intensity was observed as the peptide was digested by the enzyme (**Figure**

[180] Saenger, W. *Handbook of Proteolytic Enzymes*. (Elsevier Ltd., 2013).

36a). This degradation was further confirmed by HPLC analysis, which showed a reduction in the $(E_4K_4)_9$ peak area, consistent with proteolytic cleavage (**Figure 36**b). This result strongly supports both the anomalous nature and the peptide origin of the emission, as any organic fluorophore present in the sample as an impurity would probably not be affected by the protease.

In conclusion, these findings validate the intrinsic origin of the fluorescence, further highlighting the potential of these synthetic peptides as simple, label-free optical reporters.

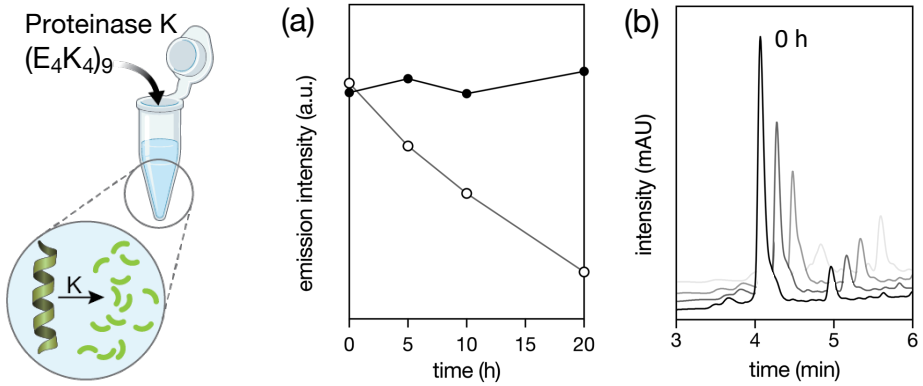


Figure 36. $(E_4K_4)_9$ peptide digestion with Proteinase K in 10 mM TRIS 50 mM NaCl pH 8 buffer (a) emission intensity at 420 nm of $(E_4K_4)_9$ (black circles) and another sample in the presence of Proteinase K (white circles) (b) HPLC of the mixture of $(E_4K_4)_9$ with Proteinase K after 0, 4, 8, and 20 h.

5.1.3.3. Quantum Calculations to Determine NAF Origin

To investigate the molecular basis of NAF emission in SAH peptides, quantum mechanical calculations of various conformations of the $(E_4K_4)_4$ and $(E_2K_2)_8$ sequences in water were performed in collaboration with Prof. Manuel Melle-Franco (University of Aveiro, Portugal). For this purpose, peptide structures were generated using various deep learning-based tools, such as RoseTTAFold[181] and AlphaFold3,[182] as well as a variant of AlphaFold3 with sidechain optimization using the SCWRL4 program.[183] For comparative purposes, α-helical and extended structures were also generated using the Tinker suite,[184] and their sidechain conformations were tested by Monte Carlo simulations using the Amber14SB forcefield to improve accuracy.[185]

[181] Baek, M. *et al. Science* **373**, 871–876 (2021).

[182] Abramson, J. *et al*. *Nature* **630**, 493–500 (2024).

[183] Krivov, G. G., Shapovalov, M. V., Dunbrack, R. L. *Proteins: Struct., Funct., Bioinf.* **77**, 778–795 (2009).

[184] Rackers, J. A. *et al. J. Chem. Theory Comput.* **14**, 5273–5289 (2018).

[185] Maier, J. A. *et al. J. Chem. Theory Comput.* **11**, 3696–3713 (2015).

In total, 31 unique conformations were obtained. They were initially optimized using the eXtended Tight-Binding (xTB) Hamiltonian,[186] and subsequently analyzed at the M06-2X 6-31G(d,p)-water level to energetically discriminate between competing conformations within each procedure. From this initial set, nine representative conformers (one per generation method) were selected for further analysis at the M06-2X/6-311++G(d,p)-water level (**Table 1**).

The calculated conformational energies show that RosettaFold structures are the most stable, while extended forms and $(E_2K_2)_8$ are less favorable energetically, suggesting they are unlikely to be significantly populated. Notably, analysis of the HOMO-LUMO gaps reveals that the linear $(E_4K_4)_4$ conformation has the widest gap (8.3 eV), whereas $(E_2K_2)_8$ shows the narrowest (7.6 eV). In general, HOMO and LUMO energies fluctuate within ~0.4 eV, and the gap varies by 0.7 eV. Interestingly, the lowest LUMO energy (–0.5 eV) also corresponds to $(E_2K_2)_8$, which could be related to electrostatic or structural effects at the E–K boundaries (e.g., EEKK or KKEE motifs), more prevalent in this peptide than in $(E_4K_4)_4$.

Table 1. Normalized relative energies of SAH peptides with different methods. HOMO and LUMO eigenvalues and corresponding gap computed with the M06-2X 6-311++G(d,p)-water level on conformations generated by the Tinker suite, RoseTTAFold, AlphaFold3 and AlphaFold3 with SCWRL4 side-chain optimization procedures. All values in eV and peptide conformations are α-helical except when noted. Abbreviations: Amber, Amber forcefield; MC, Monte Carlo side chain optimization; AF3, AlphaFold3; SC, SCWRL4 side chain optimization.

[§] Linear conformation
[#] units in kcal/mol/atom.

Peptide	Model	Rel. E[#]	LUMO	HOMO	gap
$(E_4K_4)_4$	Amber+MC[§]	0.24	–0.2	–8.4	8.3
	Amber[§]	0.38	–0.3	–8.4	8.1
	RoseTTAFold	0.00	–0.2	–8.3	8.0
	Amber+MC	0.14	–0.2	–8.2	8.0
	AF3 + SC	0.15	–0.2	–8.2	7.9
	AF3	0.12	–0.3	–8.1	7.8
	Amber	0.18	–0.2	–8.0	7.8
$(E_2K_2)_8$	AF3 + SC	0.26	–0.5	–8.1	7.7
	AF3	0.28	–0.5	–8.1	7.6

Moreover, the energies of the HOMO and LUMO orbitals vary notably across the different conformations, without a clear overarching trend (**Figure 37**a). This variability suggests a strong influence of side chain positioning on frontier orbital energetics. Indeed, the data indicate that side chains play a key role in determining both the location and energy of HOMO and LUMO levels. Typically, the HOMO electron density is localized along the peptide

[186] Bannwarth, C. *et al. Wiley Interdiscip. Rev. Comput. Mol. Sci.* **11** (2021).

backbone, whereas the LUMO is mainly associated with the lysine side chains and their terminal ammonium groups (**Figure 37**b and c).

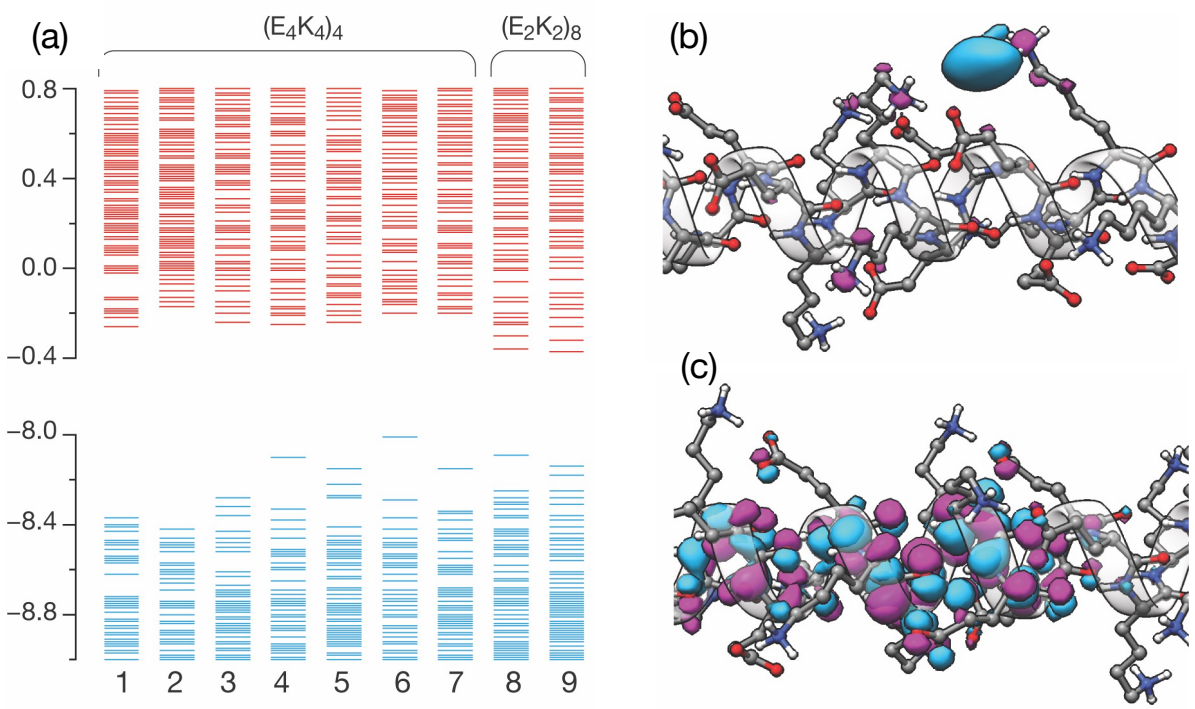


Figure 37. (a) Occupied (blue) and virtual (red) Kohn-Sham eigenvalues computed at the M06-2X 6-311++G(d,p)-water level on conformations generated with: 1: Tinker suite (linear conformation); 2: Tinker suite (linear conformation with Monte Carlo side chain optimization; 3: RoseTTAFold; 4: AlphaFold3; 5: AlphaFold3 with SCWRL4 side chain optimization; 6: Tinker suite; 7: Tinker suite and Monte Carlo side chain optimization; 8: $(E_2K_2)_8$ with AlphaFold3; 9: $(E_2K_2)_8$ with AlphaFold3 and SCWRL4 side chain optimization. (b) LUMO and (c) HOMO electron densities on the RoseTTAFold generated conformation. All energies in eV.

5.1.3.4. $(E_4K_4)_4$ as FRET Donor of Alexa Fluor

Having demonstrated NAF emission in zwitterionic α-helices, we considered whether these systems could be employed as traditional fluorophores in light-induced energy transfer processes, specifically as donors in FRET (Förster Resonance Energy Transfer) mechanisms. For this purpose, we synthesized a derivative of the model peptide $(E_4K_4)_4$ functionalized at the N-terminal end with Alexa Fluor™ 430 (AF), a commercial fluorophore whose absorption band overlaps with NAF emission of $(E_4K_4)_4$[187] and can be easily conjugated via its NHS ester.[188] As a negative control, we prepared the AF-$(E_2K_2)_2$ peptide, a non-emissive sequence that adopts a random coil conformation.

---

[187] Al Attar, H. A., Monkman, A. P. *Biomacromolecules* **10**, 1076–1084 (2009).

[188] (a) Lewis, B., Rathman. S., McMahon, R. J. *J. Nutr. Biochem.* **14**, 196–202 (2003); (b) Amancha, P. K., *et al. J. Immunol.* **191,** 6060–6070 (2013).

To evaluate the possible energy transfer process, we excited AF-$(E_4K_4)_4$ at 320 nm. In case of FRET, the emission should go from 420 nm (corresponding to the NAF signal of the peptide) to 545 nm, characteristic of the AF fluorophore (**Figure 38**a and b). When we performed this experiment in the presence of the AF-$(E_4K_4)_4$ peptide, a clear emission was observed at 545 nm, however when performed with the random coiled AF-$(E_2K_2)_4$ peptide, the intensity of this band was much lower (**Figure 38**c), which confirms the occurrence of the FRET process and supports its potential as a fluorescent probe without the need for internal aromatic markers.

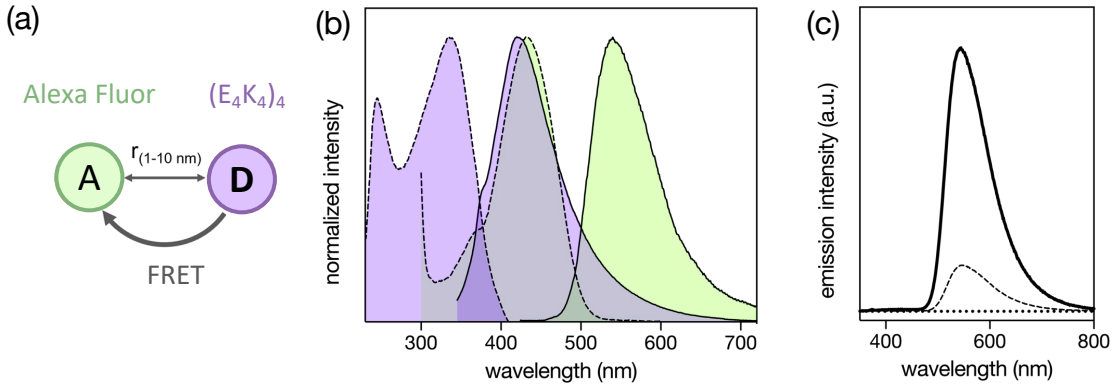


Figure 38. (a) Schematic representation of FRET process between $(E_4K_4)_4$ and Alexa Fluor; (b) normalized excitation and emission spectra of $(E_4K_4)_4$ and Alexa Fluor™ 430; (c) fluorescence spectra of 50 μM AF-$(E_4K_4)_4$ (solid line) and AF-$(E_2K_2)_4$ (dashed lines).

In conclusion, these findings open a promising avenue towards the exploration of photoinduced energy transfer processes, with the aim of broadening the functional applications of these peptides. Therefore, we propose to deepen the design of systems based on α-zwitterionic helices capable of acting as modulatory elements or optical sensors in complex biological environments.

5.1.3.5. NAF SAH Peptides as Lanthanide-Based Optical Probes

Lanthanides, such as europium, terbium and ytterbium, possess distinctive photophysical properties that make them versatile tools for studying biological systems. Their long-lasting photoluminescence, narrow-band emission and large Stokes shifts offer advantages over organic fluorophores and fluorescent proteins, allowing experiments that would otherwise be unfeasible. In addition, their ability to participate in processes such as luminescence resonance energy transfer and photon upconversion further extends their usefulness as fluorescent probes.

Based on this idea, we wondered whether our NAF peptide could serve as an antenna to mediate the luminescence of lanthanide ions. To this end, we synthesised a derivative of our model peptide $(E_4K_4)_4$ with a lanthanide-chelating DOTA macrocycle at the N-terminal end, DOTA–$(E_4K_4)_4$ (**Figure 39**).[189]

[189] Sánchez-Fernández, R. *et al. ACS Sens.* **9**, 5052–5057 (2024).

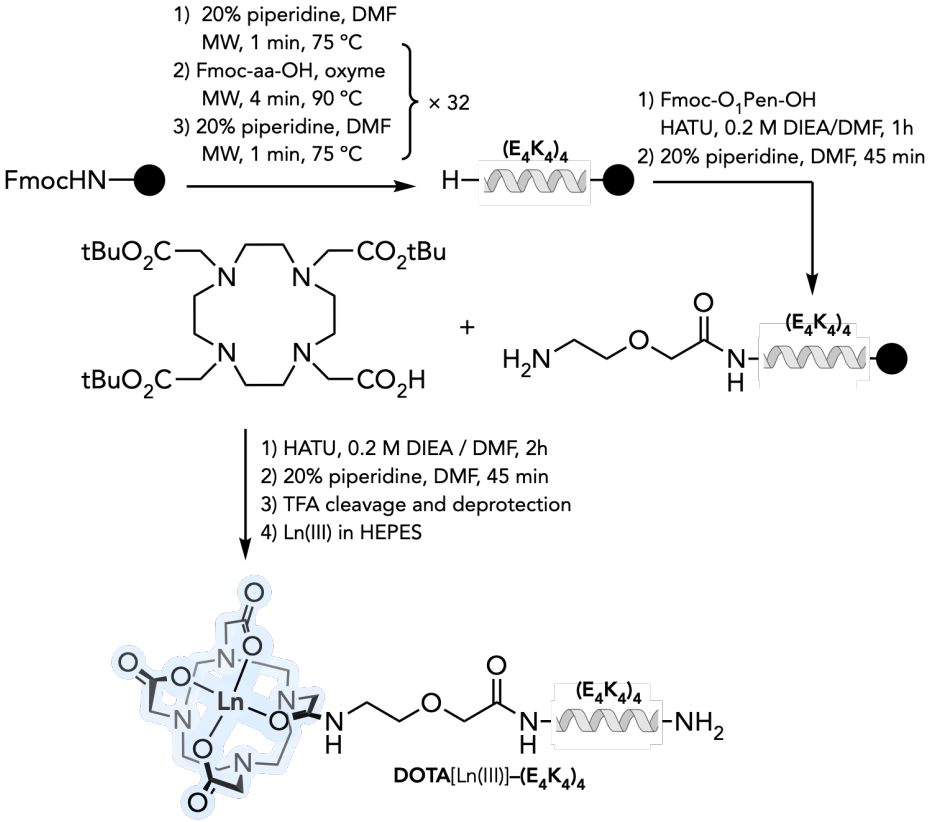


Figure 39. Synthetic route for obtaining DOTA–$(E_4K_4)_4$.

As expected, a 50 μM solution of DOTA–$(E_4K_4)_4$ showed no emission in time-gated experiments when irradiated at 330 nm. However, when we measured the emission from a mixture of DOTA–$(E_4K_4)_4$ with 1.0 equiv. of $Tb(CF_3SO_3)_3$, we observed the typical luminescence profile of the Tb(III) ion (**Figure 40**a). Given the stringent requirements for lanthanide sensitisation,[190] the most likely explanation is that this signal results from intramolecular sensitisation of the Tb(III) ion coordinated with the DOTA macrocycle and the SAH acting as an antenna. In addition, the same experiment was repeated using the non-emissive NAF control peptide DOTA–$(E_2K_2)_4$ and in this case, we did not observe the Ln(III) luminescence.

Importantly, we were also able to sensitize Eu(III) (**Figure 40**b), Sm(III) (**Figure 40**c) and Dy(III) (**Figure 40**d) ions in a similar manner using this NAF peptide, so that the SAH antenna, $(E_4K_4)_4$**,** has a significant advantage over the Trp antenna, which is typically used to sensitize Tb(III) ions, but cannot effectively sensitize other lanthanide ions luminescence.[191] These results highlight the potential of the NAF motif as a versatile antenna for the sensitization of multiple lanthanide ions, thus expanding the possibilities of designing peptide-based fluorescent probes for biochemical and biomedical applications.

---

[190] Hemmilä, I., Laitala, V. *J. Fluoresc.* **15**, 529–542 (2005).

[191] Horrocks, W. D., Collier, W. E. *J. Am. Chem. Soc.* **103**, 2856–2862 (1981).

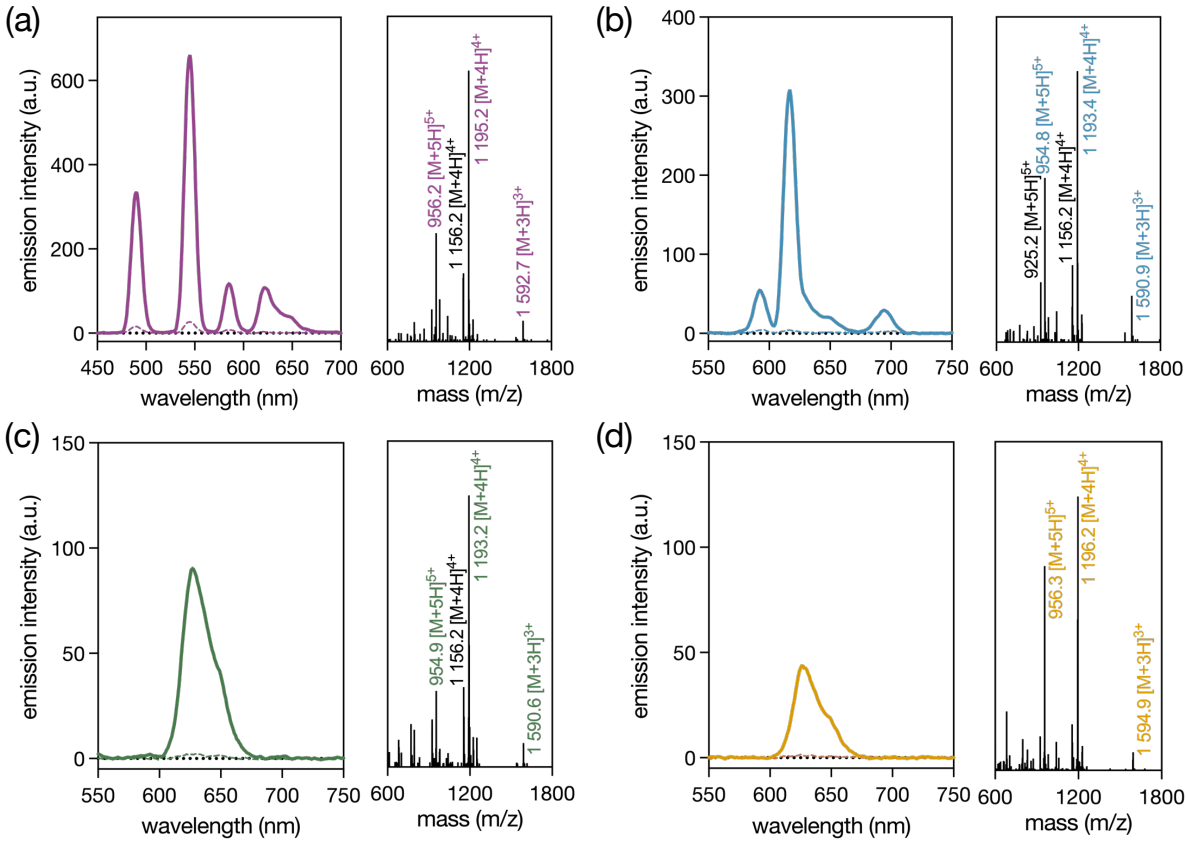


Figure 40. Time gated emission and ESI-MS spectra of 50 μM solutions of DOTA–$(E_4K_4)_4$ (solid line) and DOTA–$(E_2K_2)_4$ (dashed lines) in the presence of 1.0 equiv. of (a) $Tb(CF_3SO_3)_3$, (b) $EuCl_3$ (c) $SmCl_3$ and (d) $DyCl_3$. All spectra recorded in 25mM HEPES, 150 mM NaCl, pH 7.5 at 25°C, λ_{exc} = 320 nm, gate time 5 ms, total decay time 0.2 s, delay time 0.2 ms.

5.1.3.6. Macrocyclization of NAF SAH Peptides

Among the strategies used to modulate peptide properties, cyclization, understood as the covalent bonding between the ends of a polypeptide chain, has consolidated as an effective tool to improve both its stability and functionality.[192] Cyclic peptides obtained using this technique present significant advantages, such as increased resistance to proteolytic degradation,[193] enhanced thermal stability[194] and superior binding affinity,[195] making them promising candidates for biomedical and biotechnological applications.[196]

---

[192] (a) Zhang, R. *et al. BMCL* **26**, 1135–1150 (2018); (b) Thakkar, A., Trinh, T., Pei. D. *ACS Comb. Sci.* **15**, 2, 120–129 (2013).

[193] Li, P., Roller, P. P. *Curr. Top. Med. Chem.* **2**, 3, 325-341 (2002).

[194] Gentilucci, L., De Marco, R., Cerisoli, L. *Curr. Pharm. Des.* **16**, 28, 3185-3203 (2010).

[195] Millward, S. W. *et al. ACS Chem. Biol.* 2, 9, 625–634 (2007).

[196] (a) Driggers, E. M. *et al. Nat. Rev. Drug Discovery* **7**, 608–624 (2008); (b) Mallinson, J., Collins, I. *Future Med. Chem.* **4**, 1409–1438 (2012); (c) Marsault, E., Peterson, M. L. *J. Med. Chem.* **54**, 1961–2004 (2011).

These characteristics have driven the creation of natural product-inspired cyclic peptide libraries, both as tools for biomedical research and to address various therapeutic needs. However, macrocyclization still represents a considerable synthetic challenge,[197] as factors such as sequence and ring size critically influence reaction efficiency.[198]

Previous work has reported efficient on-resin cyclization of medium to large peptides in as little as 2 hours, employing the PyBOP/HOBt/DIEA system in DMF.[199] Inspired by this strategy, we set out to develop an efficient cyclization methodology for our NAF-emitting SAH derivatives.

To this end, we propose the synthesis of $(E_4K_4)_9$ as a linear precursor. This choice is based on the geometry of an ideal α-helix, which completes one turn every 3.6 residues, and the basic repeating unit of an SAH is 8 residues $(E_4K_4)_n$, so that a chain length of 72 amino acids corresponds to the least common multiple of 3.6 and 72, and is exactly 20 complete helical turns with a full set of (E_4K_4) repeats. This allows the peptide to close the ring-shaped peptide without generating unfavorable twists or tensions in the main chain. Furthermore, with these 9 repeats of the E_4K_4 motif, we ensure that this loading pattern repeats symmetrically and continuously around the ring (**Figure 41**). This length should allow the formation of a stable bent α-helix, according to studies that demonstrated that a 26-mer cyclic peptide is the smallest possible peptide that can sustain a bent α-helix.[200]

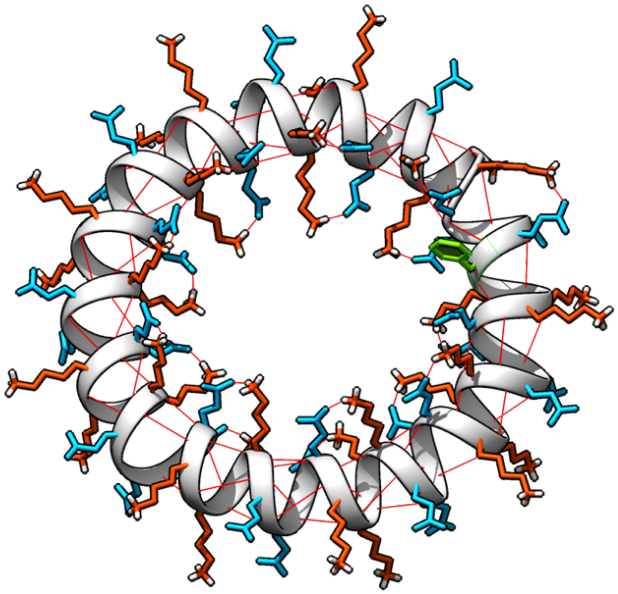


Figure 41. Proposed structure of the cyclopeptide $(E_4K_4)_9$.

[197] White, C. J., Yudin, A. K. *Nat. Chem.* **3**, 509–524 (2011).

[198] (a) Brady, S. F. *et al. J. Org. Chem.* **44**, 3101–3105 (1979); (b) Wadhwani, P. *et al. J. Org. Chem.* **71**, 55–61 (2006); (c) Qin, C. *et al. J. Med. Chem.* **46**, 4830–4833 (2003); (d) Perlman, Z. E. *et al. Bioorg. Med. Chem. Lett.* **15**, 5329–5334 (2005).

[199] Thakkar, A., Trinh, T. B., Pei, D. *ACS Comb. Sci.* **15**, 120–129 (2013).

[200] Kim, D. I., Han, S. J., Lim, Y. B. *Chem. Commun.* **58**, 27, 4368–4371 (2022).

The synthesis of the linear peptide $(E_4K_4)_9$ was completed by SPPS, using the Fmoc/tBu strategy on a previously activated 2-chlorotryptyl (CTC) resin. In the first coupling, a glutamic acid residue with Alloc protection was incorporated at the C-terminal was coupled to the resin through its side chain, with the aim of allowing site-specific cyclization upon completion of peptide chain elongation.

Once the synthesis of the linear peptide was completed, the Alloc group was selectively deprotected without affecting the rest of the side protections. For this purpose, the resin was treated with a solution of $Pd(PPh_3)_4$ (0.1 eq) and phenylsilane (PhSi, 10 eq) in dichloromethane (DCM) for 15 min at room temperature under inert nitrogen atmosphere. This procedure was repeated three consecutive times, thus ensuring complete removal of the palladium catalyst. With the C-terminal carboxylate group already free, intramolecular cyclization with the N-terminal amine was carried out on resin using a mixture of PyBOP (3.0 equiv.), HOBt (3.0 equiv.) and DIEA (6.0 equiv.) in DMF for 2 hours at room temperature. The efficiency of the cyclization was evaluated by HPLC-MS.

In addition to the synthesis of the cyclic peptide $(E_4K_4)_9$, we also wanted to explore the applicability of this cyclization strategy to peptide sequences with higher complexity, incorporating aromatic residues that could influence their structural and optical properties. For this purpose, we chose the sequence $(E_3YK_3Y)_9$, which introduces tyrosine residues every four amino acids. The synthesis of this aromatic variant was carried out following the same protocol described before, maintaining both the site-specific deprotection strategy with $Pd(PPh_3)_4$/PhSi, as well as the activation and resin-locking conditions with PyBOP/HOBt/DIEA. This approach allowed to directly compare both systems and to study the impact of tyrosine on the cycle formation and its possible functional properties.

Once our new peptide systems were successfully synthesized and characterized, we proceeded to record their EEMs to evaluate how cyclization influenced their NAF properties. The linear peptide $(E_4K_4)_9$ displayed excitation and emission bands centered at the same wavelengths as its shorter counterpart, $(E_4K_4)_4$, with comparable intensity (2.26 RU). Interestingly, the cyclized peptide displayed longer excitation/emission wavelengths, showing a signal at 380/450 nm with an intensity of 1.39 RU, almost twice that of the linear analog at the same wavelengths (0.70 RU) (**Figure 42**a). This suggests that cyclization enhances the stability or accessibility of the longer wavelength emissive states.

A similar trend was observed for the cyclic peptide $(E_3YK_3Y)_9$, which presented a band centered at 440 nm, after excitation at 360 nm (1.27 RU), red-shifted compared to its linear version (**Figure 42**b).

In conclusion, peptide cyclization not only preserves the anomalous fluorescence associated with NAF but can also extend the emission window at red-shifted wavelengths. This effect could be advantageous for future applications in optical sensing or imaging.

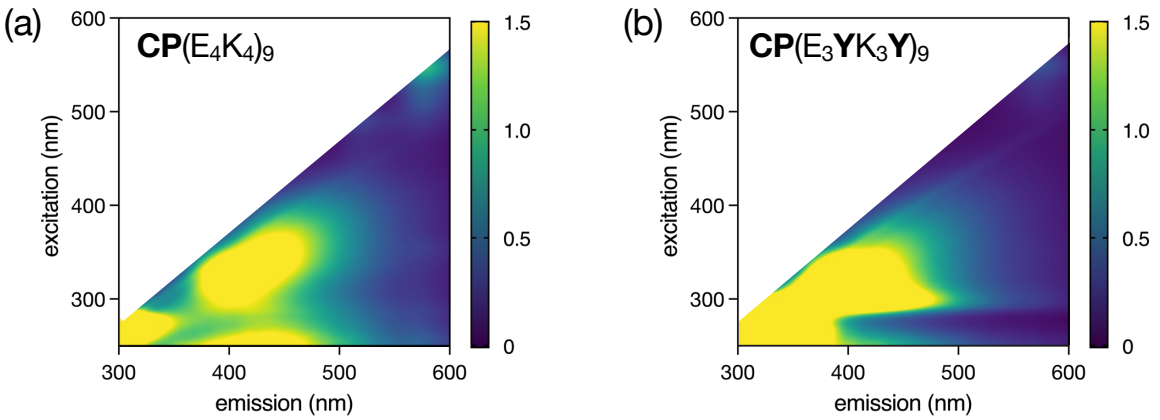


Figure 42. Excitation-Emission matrices (in Raman Units, RU) of 50 μM solutions in 25 mM HEPES buffer 150 mM NaCl, pH 7.5 at 20 ºC. of peptides (a) $CP(E_4K_4)_9$ and (b) $CP(E_3YK_3Y)_4$.

5.1.4. Conclusion

In this chapter we present, we study, for the first time, rationally designed non-aromatic fluorescence (NAF) emitting peptides designed on a zwitterionic SAH platform. We demonstrate that the α-helical conformation is a key structural element for the appearance of this emission, revealing a clear relationship between peptide folding and its photophysical properties.

We also show that the substitution of cationic residues in the SAH platform allows modulating the optical properties of the system, generating NAF luminogens with more extended emission and excitation wavelengths. Additionally, the implementation of macrocyclization strategies allowed not only to preserve the emission, but also to extend its spectral window, reinforcing the structural stability of the system.

Furthermore, we demonstrated that our systems could be applied in photoinduced energy transfer processes, such as acting as effective antennas for lanthanide ion sensitization, highlighting their potential in imaging technologies and optical sensors.

Currently, efforts are focused on better understanding the origin of NAF as well as developing genetically encodable versions that allow their integration as fluorescent markers in fluorescence microscopy and in vivo cellular studies.

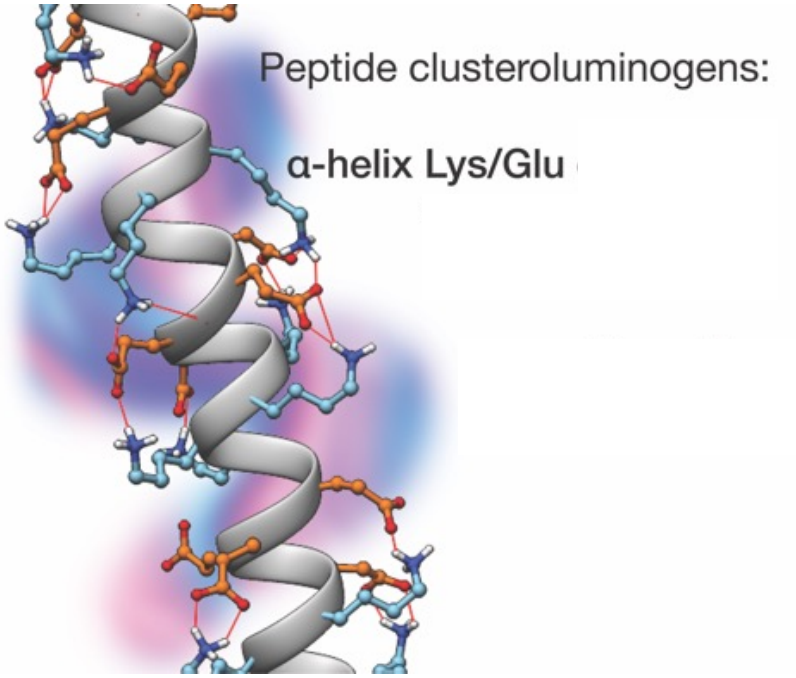


Figure 43. Illustration of the anomalous luminescence in our single α-helical peptides.

5.2. *De Novo* Engineering of Pd-Metalloproteins and Their Use as Intracellular Catalysts

Summary

The development of transition metal-based catalytic platforms that promote biorthogonal reactions inside living cells remains a major challenge in chemical biology. This is particularly true for palladium-based catalysts, which are very powerful in organic synthesis but perform poorly in cellular environments primarily due to their rapid deactivation. We now demonstrate that grafting Pd(II) complexes into engineered β-sheets of a model WW domain results in cell-compatible palladominiproteins that effectively catalyze depropargylation reactions inside HeLa cells. The concave shape of the WW domain β-sheet proved especially well-suited for accommodating the metal center and for preserving it from a rapid deactivation within the cellular environment.

Results of this section have been published in JACS Au as:

Soraya Learte-Aymamí,[a] Laura Martínez-Castro,[b] Carmen González-González,[a] Miriam Condeminas,[c,d] Pau Martin-Malpartida,[c] Sandra Baúlde,[a] José R. Couceiro,[a] Jean-Didier Maréchal,[b] Maria J. Macias,[c,e] José L. Mascareñas,[a] M. Eugenio Vázquez[a]

JACS Au **4**, 7, 2630–2639 (2024) https://doi.org/10.1021/jacsau.4c00379

Affiliations:

a. Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Departamento de Química Orgánica. Universidade de Santiago de Compostela, Santiago de Compostela 15705, Spain.

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

c. Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology (BIST), Baldiri Reixac, 10, Barcelona 08028, Spain.

d. Department of Medicine and Life Sciences, Universitat Pompeu Fabra (MELIS-UPF), Carrer del Doctor Aiguader 88, Barcelona 08003, Spain.

e. Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, Barcelona 08010, Spain.

[201] Learte-Aymamí, S., *et al. JACS Au* **4**, 7, 2630–2639 (2024).

5.2.1. Background

The development of artificial organometallic catalysts capable of promoting biocompatible reactions in living cells represents a key challenge in the field of chemical biology. However, despite several outstanding examples in the last two decades,[202] these catalysts often exhibit low efficiency in cellular environments, mainly due to the deactivation of their metal centers by reactive metabolites present in the cytoplasm.[203]

A promising alternative to discrete metal complexes is artificial metalloenzymes that incorporate late transition metals into protein scaffolds. These hybrid entities not only offer increased active site protection but also open new venues for developing more efficient and sustainable industrial processes. In parallel, natural metabolism, based on the action of highly specialized enzymes and metalloenzymes, carries out essential reactions with outstanding efficiency. Expanding this natural repertoire by introducing human-engineered abiotic transformations could have a significant impact on both basic science and biomedical applications. However, efficiently assembling these complex catalytic architectures within cells remains a major challenge.[204]

Beyond classical ligation reactions using high-energy reagents, there has been a growing interest in adapting reactions catalyzed by transition metals such as Ni, Ru, Rh, Pd or Ir to cellular contexts.[205] Recently, several protein-transition metal hybrids capable of promoting otherwise difficult reactions have been described.[206] In most cases, prosthetic groups have been used as metal coordination units, although examples have also been described in which artificial metalloproteins result from direct coordination of the metal with amino acid side chains.[207]

---

[202] (a) Völker, T., *et al. Angew Chem. Int. Ed.* **53**, 10536–10540 (2014); (b) Tomás-Gamasa, M., *et al. Nat. Commun.* **7**, 12538 (2016); (c) Sadler, J. C., *et al. RSC Chem. Biol.* **2**, 1073–1083 (2021); (d) Wang, W. *et al. Adv. Drug Deliv. Rev.* **176**, 113893 (2021); (e) Chang, T. C. Tanaka, K. *Bioorg. Med. Chem.* **46**, 116353 (2021).

[203] (a) Li, J. *et al. Nat. Chem.* **6**, 352–361 (2014); (b) Martínez-Calvo, M. *et al. ACS Catal.* **8**, 6055–6061 (2018); (c) Brewster, R. C., Klemencic, E., Jarvis, A. G. *J. Inorg. Biochem.* **215**, 111317 (2021); James, C. C., de Bruin, B., Reek, J. N. H. *Angew Chem. Int. Ed.* **62**, e202306645 (2023).

[204] (a) Ward, T. R. *Acc. Chem. Res.* **44**, 47–57 (2011); (b) Köhler, V. *et al. Nat. Chem.* **5**, 93–99 (2013) (c) Bos, J., Roelfes, G. *Curr. Opin. Chem. Biol.* **19**, 135–143 (2014); (d) Jeschek, M. *et al. Nature* **537**, 661–665 (2016); (e) Heinisch, T., Ward, T. R. *Acc. Chem. Res.* **49**, 1711–1721 (2016); (f) Schwizer, F. *et al. Chem. Rev.* **118**, 142–231 (2018).

[205] (a) Destito, P. *et al. Chem.* **27**, 4789–4816 (2021); (b) van de L'Isle, M. O. N., Ortega-Liebana, M. C., Unciti-Broceta, A. *Curr. Opin. Chem. Biol.* **61**, 32–42 (2021); (c) Miguel-Ávila, J., Tomás-Gamasa, M., Mascareñas, J. L. *Trends Chem.* **5**, 474–485 (2023); (d) Gupta, A. *et al. Acc. Chem. Res.* **56**, 2151–2169 (2023); (e) James, C. C., de Bruin, B., Reek, J. N. H. *Angew Chem. Int. Ed.* **135** (2023).

[206] Matsuo, T., Miyake, T., Hirota, S. *Tetrahedron* **60**, 151226–151233 (2019).

[207] (a) Nastri, F. *et al. Trends Biochem. Sci.* **44**, 1022–1040 (2019); (b) Studer, S. *et al. Science* **362**, 1285–1288 (2018).

In this field, palladium-based catalysts are of particular relevance due to their fundamental role in synthetic organic chemistry.[208] However, as with other metals, they have the limitation that these complexes tend to deactivate rapidly in the cellular environment.[209] A clear example of this problem is $PdCl_2$(COD), which, despite efficiently catalyzing depropargylation reactions in aqueous solutions *in vitro*, fails to achieve comparable performance inside cells.[210]

Taking all this into account, we hypothesized that integrating the metal center into a peptide scaffold could enhance catalyst stability and facilitate cellular internalization. Peptides are inherently biocompatible, modular, and easily synthesized using solid-phase methodologies, which enables precise tuning of their structure and properties to optimize metal coordination and their catalytic function.[211]

In this context, it was demonstrated that grafting a pair of His residues onto consecutive helical turns of a synthetic α-helical peptide creates an effective coordination site for Pd(II), enabling depropargylation reactions in living cells (**Figure 44**a).[212] However, when cells are preincubated with the metallopeptide before adding the substrate, the reaction fails. This suggests that the catalyst becomes deactivated in the intracellular environment, likely due to the lability of the peptide-metal coordination and/or the exposure of the metal center to the solvent.

Despite these challenges, the success of the metal-stapling strategy motivated us to explore alternative platforms to design more robust and durable palladium-peptide complexes capable of functioning as cell-compatible catalysts. We propose that multi-stranded β-sheet structures could provide a promising alternative. Due to their concave geometry, β-sheets may offer better protection to the metal ion compared to the α-helix, potentially preventing or delaying the deactivation of the catalytic center in the complex intracellular environment. Moreover, while the catalytic α-helix requires two His residues positioned on adjacent turns (i.e., relative positions *i, i+4*) to coordinate the metal, β-sheets allow for greater flexibility in arranging chelating residues. This structural versatility also enables the incorporation of additional functionalities in the second coordination sphere, providing opportunities to fine-tune the catalytic properties of the complex (**Figure 44**b).

---

[208] Powers, D. C., Ritter, T. Palladium(III) in Synthesis and Catalysis. in *Higher Oxidation State Organopalladium and Platinum Chemistry. Topics in Organometallic Chemistry*, vol. 35 129–156 (Springer, Berlin, Heidelberg, 2011).

[209] Martínez-Calvo, M. *et al. ACS Catal* **8**, 7, 6055−6061 (2018); (b) Brewster, R. C., Klemencic, E., Jarvis, A. G. *J. Inorg. Biochem.* **215**, 111317 (2021). (c) Li, J., *et al. Nat. Chem.* **6**, 4, 352−361 (2014).

[210] (a) Latocheski, E. *et al. Chem. Soc. Rev.* **49**, 7710−7729 (2020); (b) Coelho, S. E. *et al. ACS Catal.* **9**, 3792−3799 (2019); (c) Yusop, R. M. *et al. Nat. Chem.* **3**, 239−243 (2011).

[211] (a) Akagawa, K., Satou, J., Kudo, K. *J. Org. Chem.* **81**, 9396−9401 (2016); (b) Zheng, L. *et al. Angew Chem. Int. Ed.* **53**, 7599−7603 (2014); (c) Metrano, A. J. *et al. Chem. Rev.* **120**, 11479−11615 (2020); (d) Kinghorn, M. J. *et al. ACS Catal.* **7**, 7704−7708 (2017).

[212] Learte-Aymamí, S. *et al*. *Angew Chem. Int. Ed.* **59**, 9149−9154 (2020).

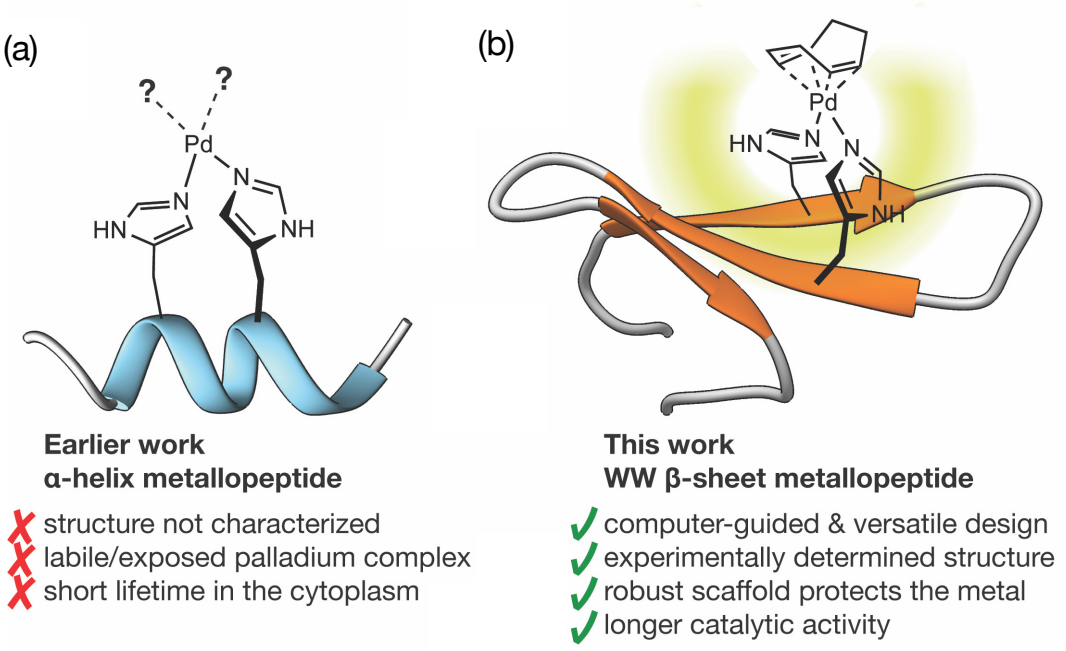


Figure 44. (a) Palladopeptide frameworks with grafted bis histidine residues. (b) The concave β-sheet structure provides greater protection to the catalytic metal center.

5.2.2. Objective

This chapter focuses on the engineering of catalytic metallopeptides using small β-folds derived from WW domains. WW domains are small, compactly folded protein modules of approximately 40 residues that mediate protein-protein interactions through their compact, antiparallel β-sheet structure. Their stability and the absence of disulfide bonds make them an ideal scaffold for the study of protein folding and catalytic engineering. In this project, we will synthesize and characterize the rationally designed WW domain that incorporate His pairs at positions that create an optimal coordination site for Pd(II) while preserving WW folding. The catalytic activity of these modified metallopeptides is assessed by a fluorogenic depropargylation model reaction. Peptides that exhibit high catalytic efficiency *in vitro* will be further tested for their internalization properties and intracellular stability to develop robust Pd-based metalloenzymes capable of catalyzing bioorthogonal transformations in living cells.

5.2.3. Results and Discussion

5.2.3.1. Design of the β-sheet Platform

The WW domain is a small protein characterized by an antiparallel β-sheet structure formed by three chains. This arrangement allows it to act as a signaling domain, since it has the ability to specifically bind to sequences that are rich in polyproline.[213] WW domains are the smallest protein modules found in nature and are composed of approximately 40 amino acids; they are also compactly folded and very stable.[214] All these properties make the WW domain a good model for protein engineering,[215] and an attractive platform to test our metal grafting strategy to generate catalytic metallopeptides.[216]

The β-sheet that is part of the WW domain structure is particularly relevant in this context, since it offers a structured surface that allows the design of pairs of histidine residues that act as metal chelation sites. These sites can precisely coordinate metal ions, which is essential for the catalytic activity of the resulting metallopeptides.

Another notable feature of the WW domain is its concave shape, as it could envelop the catalytic center limiting its exposure to the surrounding environment. This property may be particularly advantageous, as it protects the catalytic center from potential unwanted interactions with other molecules or with the solvent, which could affect the efficiency of catalysis.

For our designs, we took as a reference the WW domain prototype described by Macias *et al.* (PDB entry 1E0M).[217] In order to avoid any possible interference in the coordination of the metal ion, we made modifications to the 1E0M sequence, replacing the original residues His14 and His23 by Pro and Thr, respectively, which are residues commonly found at these positions in natural WW domains.

Additionally, during preliminary testing, we observed that the Asp9 residue present in the 1E0M sequence showed a high propensity for aspartimide formation during the solid-phase peptide synthesis process.[218] To avoid this problem, we replaced Asp9 with a Thr residue, which is also typical at this position in other native WW domains.

Finally, we also truncated the 1E0M sequence by removing two residues from the N-terminal region (Ser-Met) and three from the C-terminal region (Met-Ser-Ser). This decision was based

---

[213] Macias, M. J., Wiesner, S., Sudol, M. *FEBS Lett.* **513**, 30–37 (2002).

[214] Macias, M. *et al. Nature* **382**, 646–649 (1996).

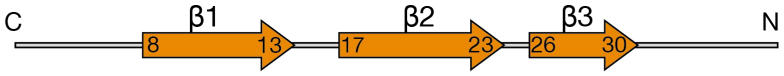
[215] (a) Dalby, P. A., Hoess, R. H., Degrado, W. F. *Protein Sci.* **9**, 2366–2376 (2000); (b) Deechongkit, S., Kelly, J. W. *J. Am. Chem. Soc.* **124**, 4980–4986 (2002); (c) Russ, W. P. *et al. Nature* **437**, 579–583 (2005); (d) Ferguson, N. *et al. PNAS* **98**, 13002–13007 (2001); (e) Zhou, R. *et al*. *Proc. Natl. Acad. Sci. USA* **111**, 18243–18248 (2014).

[216] Pham, T. L., Kovermann, M., Thomas, F. *ACS Synth. Biol.* **11**, 254–264 (2022).

[217] Macias, M. *et al. Nat. Struct. Mol. Biol.* **7**, 375–379 (2000).

[218] Subirós-Funosas, R., El-Faham, A., Albericio, F. *Tetrahedron* **67**, 8595–8606 (2011).

on previous studies showing that synthetic WW domains can maintain their natural folding structure even without these residues, and to avoid the presence of the Met residue that is prone to oxidation. As a result of these modifications, we obtained our reference sequence **WW0** (**Figure 45**).



	1	11	21	31
1E0M	SMGLPPGW*D*E	YKT*H*NGKTYY	YN*H*NTKTSTW	TDPRMSS
WW0	--GLPPGW**T**E	YKT**P**NGKTYY	YN**T**NTKTSTW	TDPR---
WW11/13	--GLPPGWTE	**H**K**H**PNGKTYY	YNTNTKTSTW	TDPR---
WW11/19	--GLPPGWTE	**H**KTPNGKT**H**Y	YNTNTKTSTW	TDPR---
WW13/19	--GLPPGWTE	YK**H**PNGKT**H**Y	YNTNTKTSTW	TDPR---
WW9/21	--GLPPGW**H**E	YKTPNGKTYY	**H**NTNTKTSTW	TDPR---
WW9/23	--GLPPGW**H**E	YKTPNGKTYY	YN**H**NTKTSTW	TDPR---
WW21/23	--GLPPGWTE	YKTPNGKTYY	**H**N**H**NTKTSTW	TDPR---

Figure 45. Sequences of the WW reference 1E0M and the designed WW peptides. The metal-coordinating His residues in bold and the rest of the sequence in grey for clarity. Peptides are named according to the positions of the grafted His residues.

Next, we combined visual inspection with computational prediction techniques and 3D modeling, in collaboration with Prof. Jean-Didier Maréchal and Laura Martinéz-Castro from the group *Insilichem* at the Universitat Autónoma of Barcelona, to identify the most suitable positions to graft histidine (His) residue pairs onto the β-sheet of the **WW0** sequence, with the goal of creating $Pd(His)_2$ binding sites.[219]

Specifically, they analyzed the **WW0** scaffold using our recently developed predictor for metal-binding sites, *BioMetAll.*[220] This analysis led them to conclude that the best sites for incorporating His residue pairs on the concave face of the β-sheet of **WW0** were adjacent *i, i+3* positions within the same chain, such as **WW11/13** and **WW21/23**; facing residue pairs, such as **WW9/21** and **WW11/19**; and residue pairs placed diagonally across two adjacent chains, such as **WW9/23** and **WW13/19**.

These engineered sequences can be grouped into two families, each of which defines metal-binding sites in different regions of the β-sheet. The first family is in the β1-β2 turn, highlighting positions 11, 13 and 19 (marked in green in **Figure 46**). The second family is located at the opposite end of the β1-β2 hairpin, close to the β2-β3 turn, highlighting positions 9, 21 and 23 (indicated in light blue in **Figure 46**).

[219] Alonso-Cotchico, L. *et al. Acc. Chem. Res.* **53**, 896–905 (2020).

[220] Sánchez-Aparicio, J. E. *et al. J. Chem. Inf. Model* **61**, 311–323 (2021).

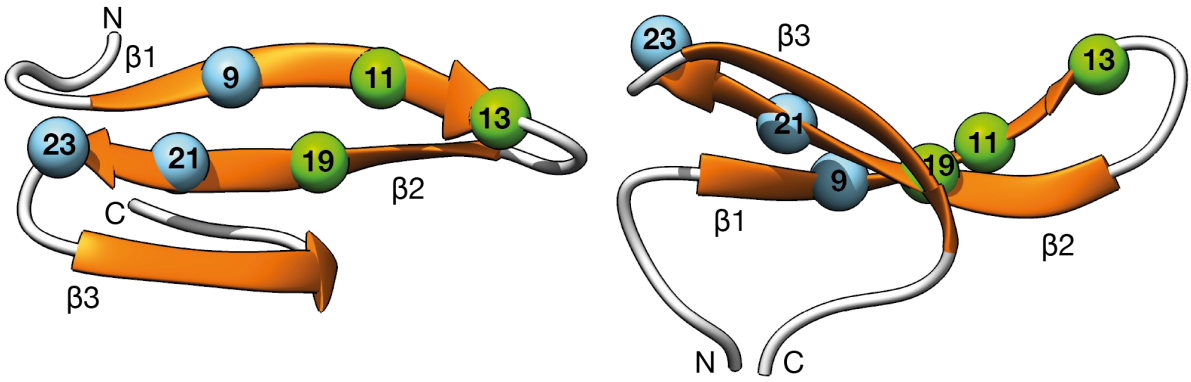


Figure 46. Top and side view of the reference 1E0M WW domain showing the concave structure of the β-sheet, and indicating the positions 9, 11, 13, 19, 21, and 23, where the His residues are grafted. Residue numbering according to the reference PDB structure 1E0M.

5.2.3.2. Pd(II) Binding to the His-grafted WW Peptides

The designed peptides were synthesized using standard microwave-assisted solid-phase peptide synthesis methods, following the Fmoc/tBu protocol.[221] With the WW peptides in hand, we evaluated their ability to coordinate metal ions. Each peptide was incubated in water with a solution of the palladium precursor dichloro(1,5-cyclooctadiene)palladium(II), $PdCl_2$(COD), at a 1:1 molar ratio for 1 hour, and the resulting mixtures were analyzed by HPLC-MS(ESI).

As expected, the control peptide **WW0**, which lacks histidine residues for metal coordination, showed no MS peaks indicative of complex formation. In contrast, all peptides engineered with histidine residue pairs displayed new MS peaks consistent with the formation of palladium coordination complexes. To further support our hypothesis, we included a control peptide, **WW19**, containing only a single histidine residue at position 19. This peptide did not produce any detectable coordination complexes by MS, reinforcing the idea that bidentate coordination, i.e., the involvement of two histidine residues, is essential to form stable complexes (**Figure 47**).

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

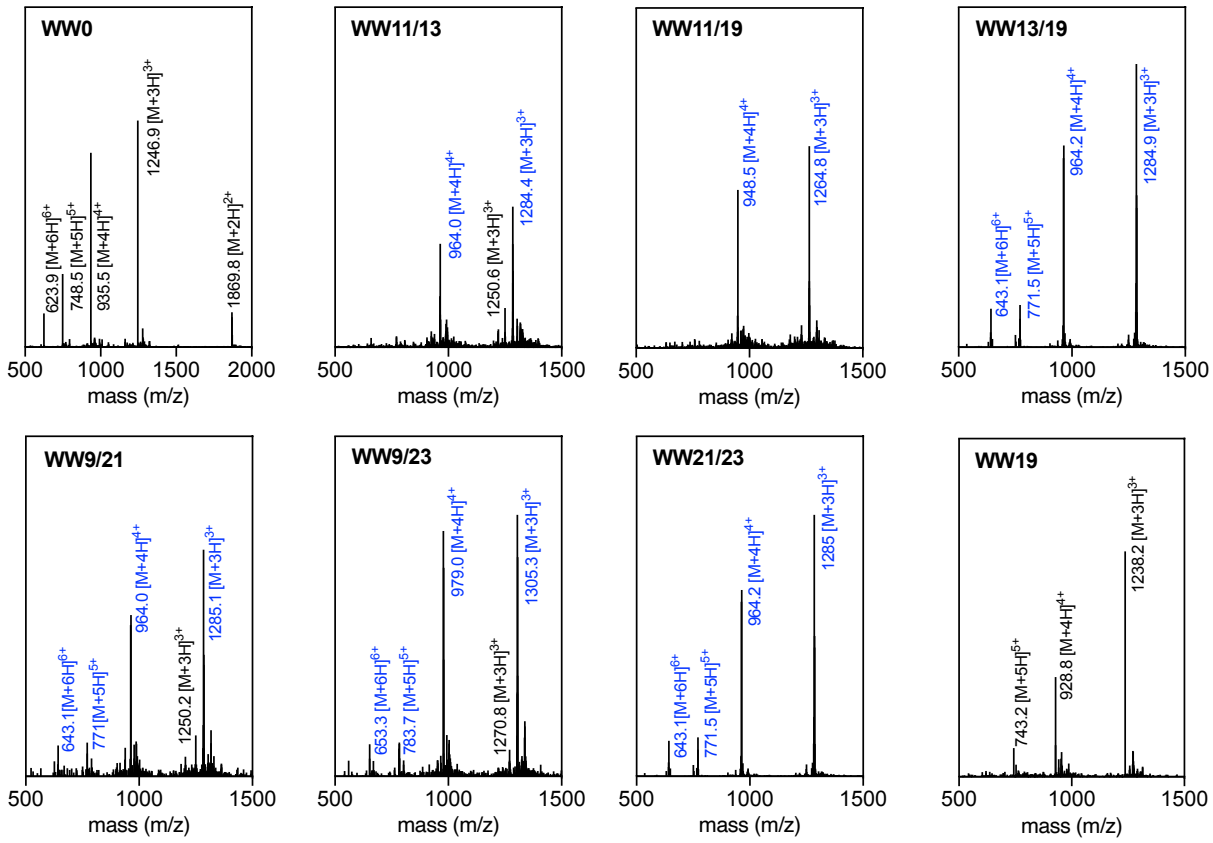


Figure 47. ESI-MS spectra of the mixture of WW peptides used in this study with 1.0 equiv. of $PdCl_2$(COD). The peaks corresponding to the apo-peptides are labeled in black, those of the corresponding palladopeptides, are in blue.

Despite the known limitations of CD spectroscopy for analyzing β-structures,[222] this technique provided additional evidence for the formation of Pd(II) complexes (**Figure 48**a) and allowed us to assess how mutations affected the structure of different histidine (His)-grafted peptides.

The CD spectrum of the reference sequence, **WW0**, showed the typical features of a well-structured WW domain, with a positive ellipticity band at 230 nm and a negative one near 210 nm.[223] As expected, the spectrum of **WW0** showed no significant changes after the addition of an equivalent of $PdCl_2$(COD) (**Figure 48**b). Similar behavior was reported for the mono-His control peptide, **WW19**. This is in agreement with previous HPLC-MS results where we saw that no palladium complex was formed without the presence of the two His residues.

[222] Micsonai, A. *et al*. *Proc. Natl. Acad. Sci. USA* **112**, E3095–E3103 (2015).

[223] (a) Jäger, M., Dendle, M., Kelly, J. W. *Protein Sci.* **18**, 1806–1813 (2009); (b) Koepf, E. K. *et al. Protein Sci* **8** (4), 841–853 (1999).

In contrast, the circular dichroism spectra of most of the His-modified peptides indicated that they did not adopt a structure like canonical WW domains upon addition of $PdCl_2(COD)$. The only exception was the **WW13/19** peptide, whose CD spectrum, after incorporation of palladium salt, showed a clear β-sheet signal (**Figure 48**c). For this reason and considering the importance of maintaining the WW folding structure for the stability of the complex and the protection of the metal center in intracellular quenchers, we decided to continue with the structural and catalytic studies of the **WW13/19** peptide.

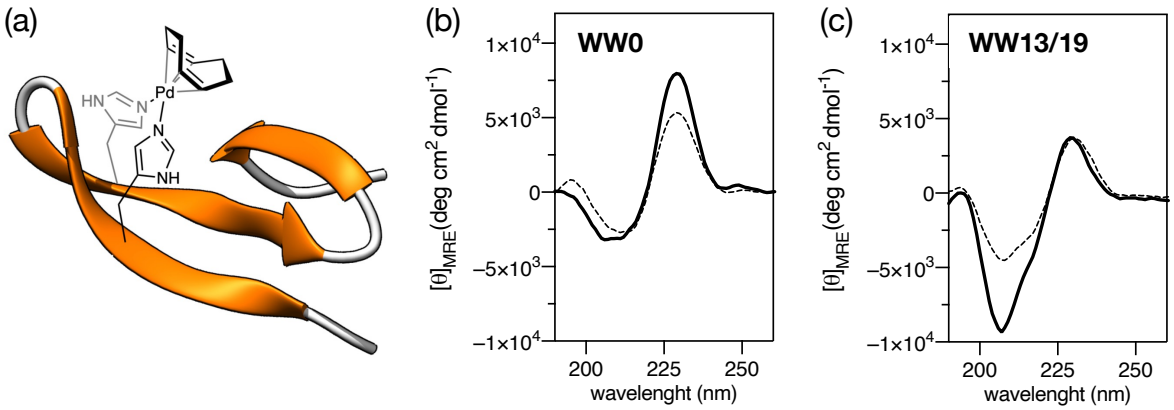


Figure 48. (a) Schematic representation of a WW domain mutant with His residues at appropriate positions to coordinate a $PdCl_2(COD)$ catalytic center (PDB: 1E0N); (b) circular dichroism spectra of 5 μM solutions of the non-coordinative **WW0** peptide and (c) of the His-modified **WW13/19** peptide, before (thick solid lines), and after incubation with 1 equiv. of Pd $Cl_2(COD)$ (dashed lines).

To monitor the coordination and formation of the palladoprotein complex in solution, the decrease in the emission intensity of tryptophan residues at 350 nm after progressive addition of Pd(II) ions was assessed (**Figure 49**a). For this purpose, a 5 μM solution of the peptide was titrated in 25 mM HEPES buffer (150 mM NaCl, pH 7.5) with increasing concentrations of Pd(II). The samples were incubated overnight to ensure equilibration of the mixture, and their respective emission spectra were subsequently recorded.

The results confirmed that **WW13/19** has a significantly higher affinity for Pd(II) compared to **WW0**. In the case of **WW13/19**, a marked decrease in fluorescence intensity was observed, which fitted a 1:1 binding pattern, with an approximate dissociation constant of K_D = 12.7 μM (**Figure 49**b). In contrast, for **WW0** virtually no quenching of the tryptophan signal (K_D > 1 mM) was observed (**Figure 49**c), indicating that the complex does not form.

To obtain further structural information on the secondary structure of the synthetic palladoproteins, the **WW0** and **WW13/19** peptides were analyzed by ^{1}D NMR at room temperature, both in the absence and presence of $PdCl_2(COD)$.

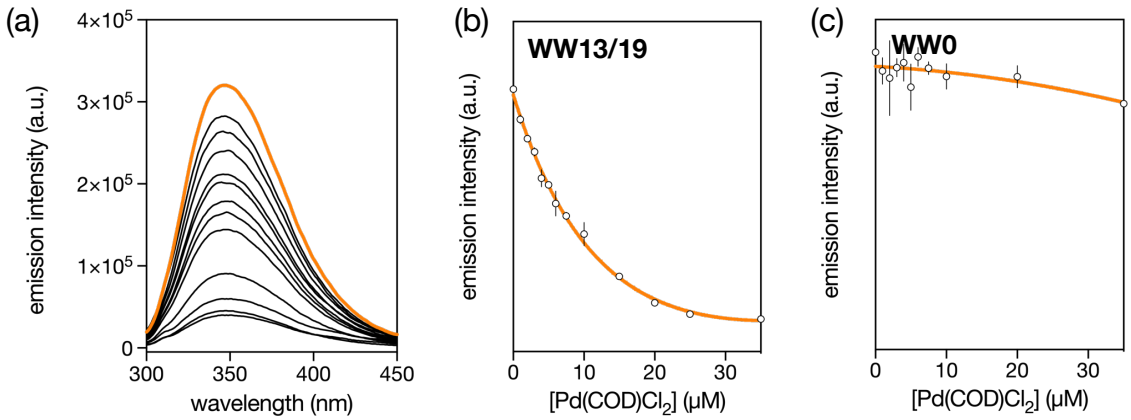


Figure 49. Pd(II) fluorescence titration of **WW13/19** and **WW0**. (a) fluorescence quenching of a 5 μM solution of **WW13/19** (orange thick line) and after incubation overnight with Pd(II); tryptophan emission profile at 350 nm for 5 μM solutions of (b) **WW13/19** and (c) WW0, incubated overnight with in increasing concentrations of $PdCl_2$(COD) and the best fit to 1:1 binding model (K_D = 12.66 μM for **WW13/19** and K_D = 2.7 mM for **WW0**, respectively). All the measurements were conducted in 25 mM HEPES buffer, 150 mM NaCl, pH = 7.5.

To obtain further structural information on the secondary structure of the synthetic palladoproteins, the **WW0** and **WW13/19** peptides were analyzed by ^{1}D NMR at room temperature, both in the absence and presence of $PdCl_2$(COD).

As expected, the **WW0** peptide showed good peak dispersion in the absence of the palladium salt, indicating a well-folded WW domain (**Figure 50**a, black trace). The Hε peaks of both Trp residues appeared well defined, and no changes were observed upon the addition of five equivalents of $PdCl_2$(COD), as expected for a peptide lacking coordinating His (**Figure 50**a, red trace). The only difference between the two spectra is the presence of a peak at 5.8 ppm, corresponding to the double bond protons of the COD ligand, indicating that $PdCl_2$(COD) is present but not interacting with the peptide.

of the metal, with only a single signal for the two Trp indoles and no methyl/methylene resonances below 1 ppm. These features suggest that **WW13/19** is largely unfolded without palladium (**Figure 50**b, black trace). Upon addition of five equivalents $PdCl_2$(COD), the **WW13/19** spectrum underwent a notable transformation. Increased peak dispersion, distinct Hε signals for each Trp, and long–range interactions between Trp8 and Pro33 (resonances around –0.5 ppm) indicated that metal coordination promotes peptide folding.[224] Additional resonances appeared in the 5–6 ppm range, corresponding to the COD ligand now in a structured environment. Notably, the presence of new signals at 12.5 and 12.8 ppm confirmed coordination of the His Hε2 protons, which were absent in **WW0** (**Figure 50**b, red trace). Taken together, these NMR results confirm that palladium binding to **WW13/19** induces a folding-

[224] (a) Pham, T. L., Kovermann, M., Thomas, F. *ACS Synth. Biol.* **11**, 1, 254–264 (2022); (b) Li, W. *et al. J. Am. Chem. Soc.* **130**, 3, 892–900 (2008).

upon-binding process, generating a well-folded palladoprotein where the metal center is coordinated both by the histidine residues and the COD ligand.

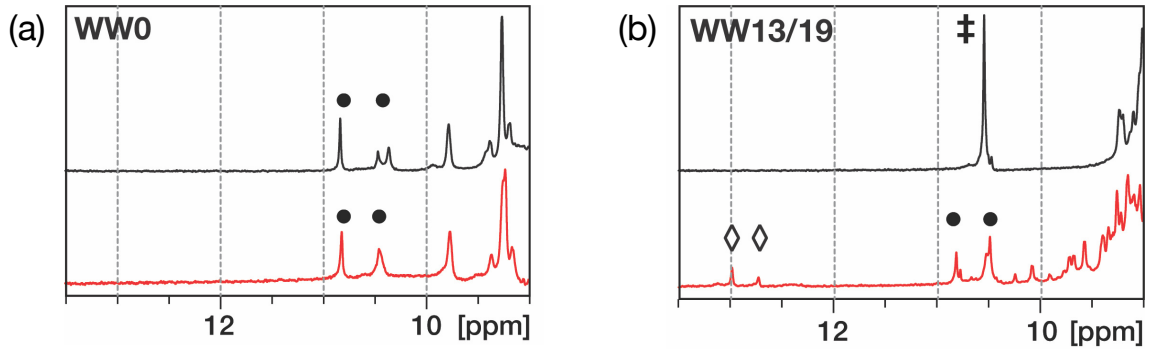


Figure 50. (a) ^{1}H NMR spectra of the peptide **WW0** (black) and the same peptide after the addition of $PdCl_2$(COD) (red); (b) ^{1}H NMR spectra of the peptide **WW13/19** His double mutant (black) and the same peptide after the addition of $PdCl_2$(COD) (red). Indicated resonance for the Trp indole group (double dagger), Trp resonances (black circles) and the new imidazole resonances at ~13 ppm corresponding to the H_ε of the His residues (labeled with diamonds).

5.2.3.3. In Vitro Depropargylation Reaction

Having confirmed that WW peptides grafted with histidine form coordination complexes in the presence of $PdCl_2$(COD), we set out to investigate whether these complexes can catalyze reactions.

As a proof of concept, we explored the depropargylation of the probe **1**,[225] which has been previously described for its high sensitivity, excellent selectivity and low detection limit for Pd^{2+} in aqueous solution.[226] Furthermore, the propargylated product can be synthesized from *2-hydroxy-5-methylbenzaldehyde* via a four-step synthetic route with an overall yield of 20%, according to a procedure previously described by Mascareñas *et al.* (**Figure 51**a).[227] More important, this propargylated product is non-fluorescent but, upon uncaging, generates a product (2) that emits light at 655 nm when excited in the far-ultraviolet region (λ_{ex} = 400 nm) (**Figure 51**b and c). This makes this probe a promising option for analyzing the catalytic activity of palladoproteins.

[225] M. Martínez-Calvo *et al. ACS Catal.* **8**, 6055 – 6061 (2018).

[226] Gao, T. *et al. Chem. Asian J.* **10**, 5, 1142–1145 (2015).

[227] Learte-Aymamí, S. *et al. Angew Chem. Int. Ed. Engl.* **59**, 9149–9154 (2020).

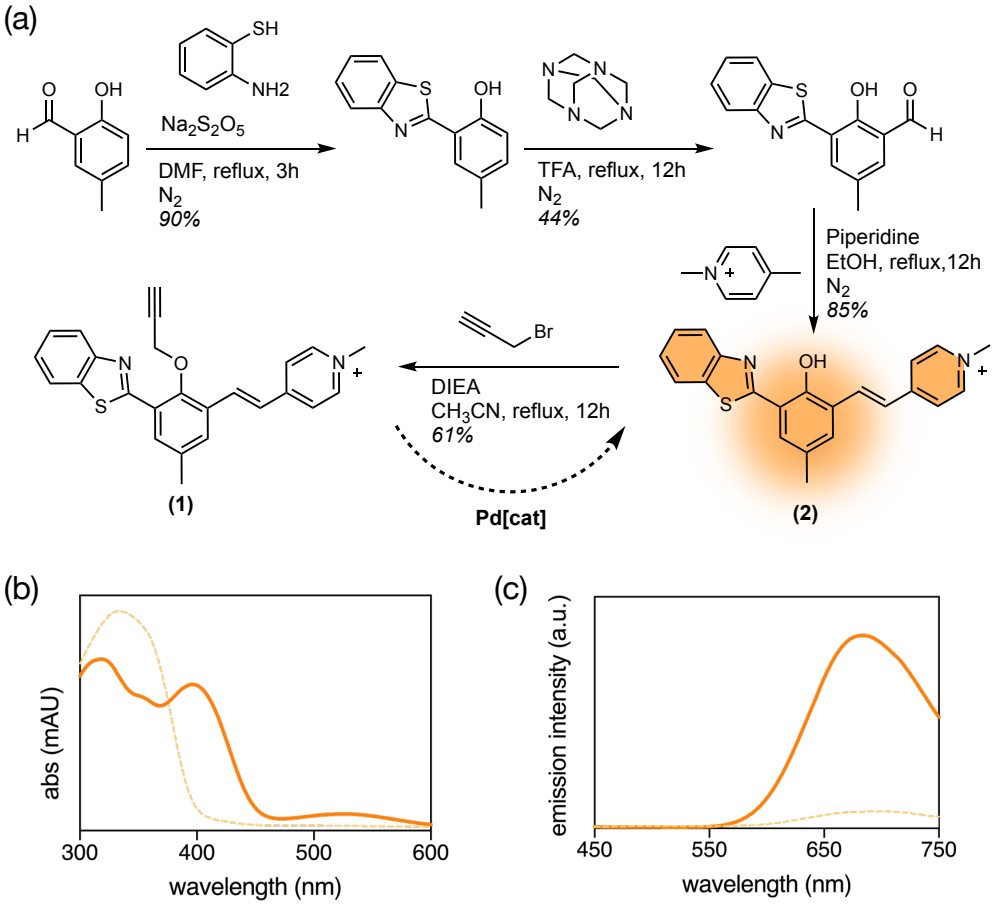


Figure 51. (a) Synthetic route to obtain the propargylated fluorogenic probe (**1**); (b) absorbance spectra of the propargylated fluorogenic probe **1** (dashed line) and the deprotected product **2** (thick solid line); (b) fluorescence emission spectra of **1** (dashed line) and **2** (thick solid line). Measurements taken at 20 μM, 1:1 v/v DMSO/H_2O, λ_{exc} = 400 nm.

After synthesis of our probe and confirmation that the peptides were able to form complexes with Pd(II) salts, the next step was to evaluate their catalytic ability. For this, 20 μM solutions of each WW peptide were incubated in Milli-Q water with $PdCl_2$(COD) at a 1:1 ratio for one hour. The obtained palladium solutions were then added to a mixture containing 200 μM of propargylated fluorogenic probe **1** in PBS (10 mM phosphate buffer, NaCl 100 mM, pH 7.5), and the reaction was kept stirred at 37 °C for 24 hours.

As shown in **Figure 52**a, HPLC-MS(ESI) analysis of the mixtures and quantification of the depropargylated product **2** showed that the **WW13/19**[Pd(II)] complex was the most active catalyst, as expected from the palladium coordination results. This complex yielded the depropargylation product with a moderate yield of 62 %, comparable to that achieved with $PdCl_2$(COD) in the absence of additional ligands.

Using the same protocol, but employing the **WW0** peptide, the formation of the depropargylation product was also observed, which was expected given the intrinsic catalytic activity of the free catalyst $PdCl_2$(COD). However, by incorporating an additional

ultradiafiltration step with 3 kDa Amicon centrifugal filters to remove the uncoordinated palladium salt, the catalytic activity of both pure $PdCl_2$(COD) and its mixture with **WW0** was drastically reduced. In contrast, the reaction catalyzed by the **WW13/19**[Pd(II)] complex maintained significant efficiency even after removal of free palladium. The control peptide **WW19,** containing a single histidine, showed very limited catalytic activity under these conditions, consistent with its inability to form a stable complex with Pd(II) (**Figure 52**b).

In parallel, kinetic studies were carried out to compare the catalytic activity of the **WW13/19**[Pd(II)], **WW0**[Pd(II)] and $PdCl_2$(COD) complexes (note that for **WW0**[Pd(II)] we used the 3 kDa Amicon centrifugal protocol to remove the free palladium), monitoring fluorescence at 620 nm upon excitation at 405 nm every 15 minutes over a 24-hour period at 37 °C. The reaction profiles for **WW13/19**[Pd(II)] and $PdCl_2$(COD) are similar, with conversion progressing for approximately the first 8 hours, after which it plateaued, remaining virtually constant until the end of the experiment. In contrast, the **WW0**[Pd(II)] complex showed significantly lower conversion (**Figure 52**c).

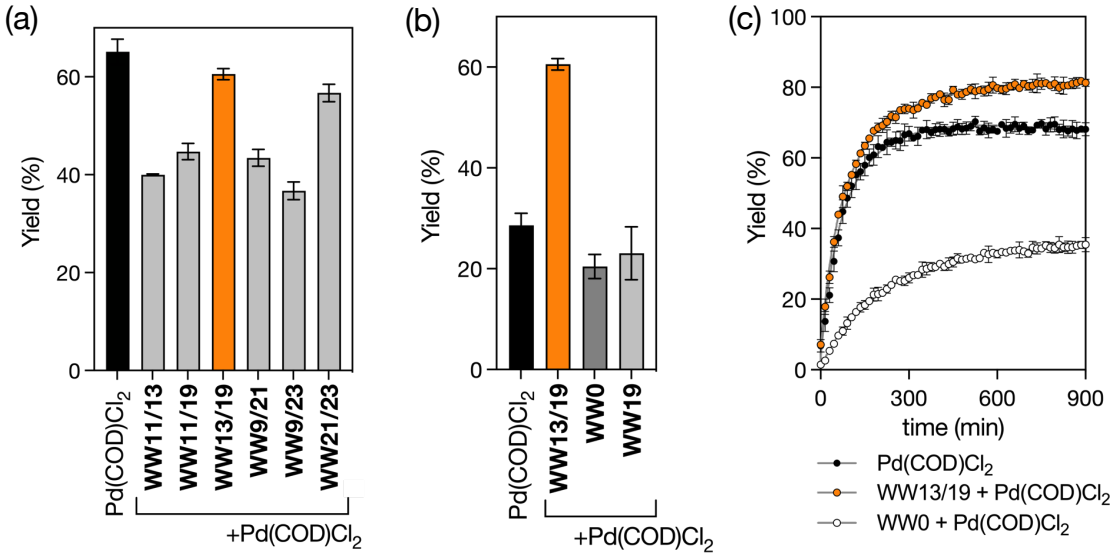


Figure 52. *In vitro* catalysis of WW peptides. (a) Bar diagram representation of the yields obtained for each catalyst. Reaction conditions: mixing for 24 h a 200 μM solution of probe **1** with 20 μM mixtures of each peptide mixed with $PdCl_2$(COD) in a 1:1 ratio for 1 h. Control experiments using $PdCl_2$(COD) in dark black, **WW13/19**[Pd(II)] the best catalyst in orange and other WW peptides in grey; (b) same experiment but with ultradiafiltration with 3 kDa Amicon centrifugal filters to remove the remaining free $PdCl_2$(COD), before the reaction; (c) kinetic traces of product **1** in the presence of **WW13/19** + $PdCl_2$(COD); **WW0** + $PdCl_2$(COD); $PdCl_2$(COD). The lines show the best fit to a pseudo-first kinetic model of the depropargylation profiles (% yield $Y = Y_0 + (Y_{max} - Y_0) \times (1 - e^{-k\,t})$, as described in Domingos *et al.*[228] Catalysis was performed in 10 mM phosphate buffer, 100 mM NaCl, pH = 7.5 and yields were calculated by HPLC-MS(ESI) using internal standards.

[228] Dal Forno, G. M. *et al. Chem. Sci.* **15**, 4458–4465 (2024).

Kinetic analysis under pseudo-first-order conditions confirmed that the reaction rates of **WW13/19**[Pd(II)] and $PdCl_2$(COD) are comparable, with rate constants of $k = 9.6 \times 10^{-3}$ and 12.2×10^{-3} min^{-1}, respectively. On the other hand, **WW0**[Pd(II)] exhibited a significantly lower constant of $k = 4.6 \times 10^{-4}$ min^{-1}, reflecting its limited catalytic ability.

5.2.3.4. Cellular Internalization of WW13/19 and its Palladium Complex

As a preliminary step to study the catalytic activity of the palladoprotein **WW13/19**[Pd(II)] inside mammalian cells, we synthesized the tetramethylrhodamine-labeled peptide TMR-**WW13/19**, which would allow us to evaluate its internalization by fluorescence microscopy. Thus, we incubated HeLa cells for 1 h with 5 μM solutions of the apopeptide TMR-**WW13/19** and its Pd(II) complex, TMR-**WW13/19**[Pd(II)], which was preassembled by mixing TMR-**WW13/19** with $PdCl_2$(COD) in water for 1 h. Interestingly, the palladoprotein TMR-**WW13/19**[Pd(II)] is better internalized than the palladium-free precursor TMR-**WW13/19**, as shown by the intensity of the intracellular fluorescence (**Figure 53**a).

Quantitative analysis of the internalization by cell cytometry confirmed this observation. The palladoprotein TMR-**WW13/19**[Pd(II)] internalized over 6000 times more efficiently than TMR-**WW13/19**, according to the recorded emission intensity of the TMR. The critical role of Pd(II) coordination for the internalization of these peptides is reinforced by the results with the mono-His control peptide, TMR-**WW19**, which is unable to translocate into the cellular interior, even in the presence of $PdCl_2$(COD) (**Figure 53**b).

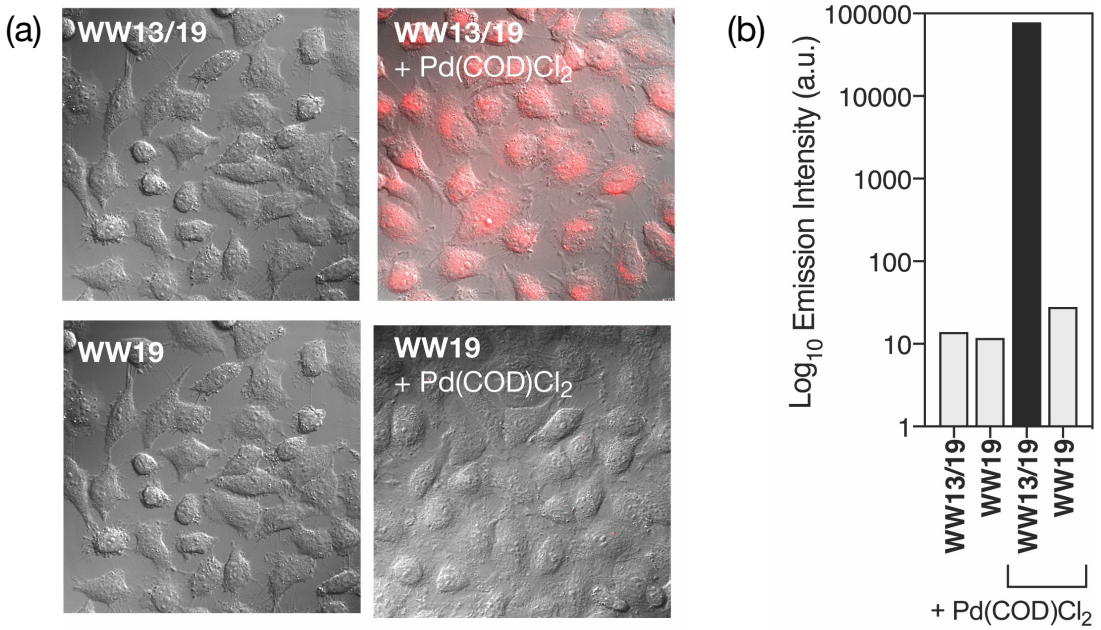


Figure 53. Internalization of TMR-labeled peptide **WW13/19** and its palladoprotein derivative **WW13/19**[Pd(II)]. (a) Fluorescence microscopy of HeLa cells with 5 μM solutions of the peptides **WW13/19** and **WW19** (left) and of their mixtures with $PdCl_2$(COD) in water (right) incubated for 1h; (b) semilogarithmic plot of the intracellular TMR emission measured by cell cytometry.

Although at this time, we cannot provide a definitive explanation for the increased cellular penetration of the palladium-stapled miniprotein, it is consistent with previous observations of increased penetration as a result of conformational constraints imposed by cyclization,[229] stapling,[230] or coordination.[231] Additional assays in the presence of various endocytosis inhibitors, suggest that the internalization of TMR-**WW13/19**[Pd(II)] is energy-dependent and most likely occurs by micropinocytosis, accumulating inside endosomal vesicles.

5.2.3.5. Intracellular Depropargylation Reaction Mediated by WW13/19[Pd(II)]

Having demonstrated the good internalization of **WW13/19**[Pd(II)], we were eager to test whether the palladoprotein could promote the depropargylation reaction inside cells. Therefore, we took advantage of the fluorogenic probe **1** tested in the *in vitro* experiments, which exhibits low emission in its protected form, but becomes highly fluorescent upon depropargylation, therefore allowing for easy monitoring of the catalytic deprotection reaction by fluorescence microscopy.[232]

The cellular reactions were carried out by incubation of HeLa cells with a 50 μM solution of the probe **1** for 1 h at 37 ºC. The medium was removed, and the cells washed twice with FBS-DMEM to remove any extracellular probe. Then a 50 μM solution of freshly preformed palladoprotein **WW13/19**[Pd(II)] were added, after 1 h incubation after 1 h and two washes with PBS the cells were imaged by wide-field fluorescence microscopy. The cells showed an intense fluorescence emission under the microscope, which must be associated with the formation of the desired depropargylation product **2**, since the protected probe is nonfluorescent under the conditions of our experiment. Given that the palladopeptide is preformed just before the cellular incubations, it could not be discarded that the intracellular reaction was promoted by discrete palladium entities derived from the palladium source. However, in the absence of the peptide, the complex $PdCl_2$(COD) is incapable of generating any intracellular fluorescence. Importantly, control experiments with non-coordinative peptides, **WW0** and **WW19**, mixed with $PdCl_2$(COD), neither induce any intracellular emission (**Figure 54**).

As mentioned before, earlier work has shown that related α-helical palladopeptides are rather unstable and only catalyze the depropargylation reaction when the cells are incubated with the probe before adding the catalyst. Our new metalloprotein **WW13/19**[Pd(II)] is more robust and thus it catalyzes the depropargylation reaction using a reverse order of addition. Indeed, when HeLa cells were incubated with the preformed complex **WW13/19**[Pd(II)] for 1 h, washed to remove extracellular compounds, and mixed with the probe **1** for 1 h, we observed similar intracellular emission (**Figure 54**a, bottom right).

---

[229] Lättig-Tünnemann, G. *et al. Nat. Commun.* **2**, 453 (2021).

[230] Li, S. *et al. Chem. Commun.* **56**, 100, 15655–15658 (2020).

[231] Learte-Aymamí, S. *et al. Angew Chem. Int. Ed. Engl.* **59**, 23, 9149–9154 (2020).

[232] Learte-Aymamí, S. *et al. J. Am. Chem. Soc.* **139**, 45, 16188–16193 (2017).

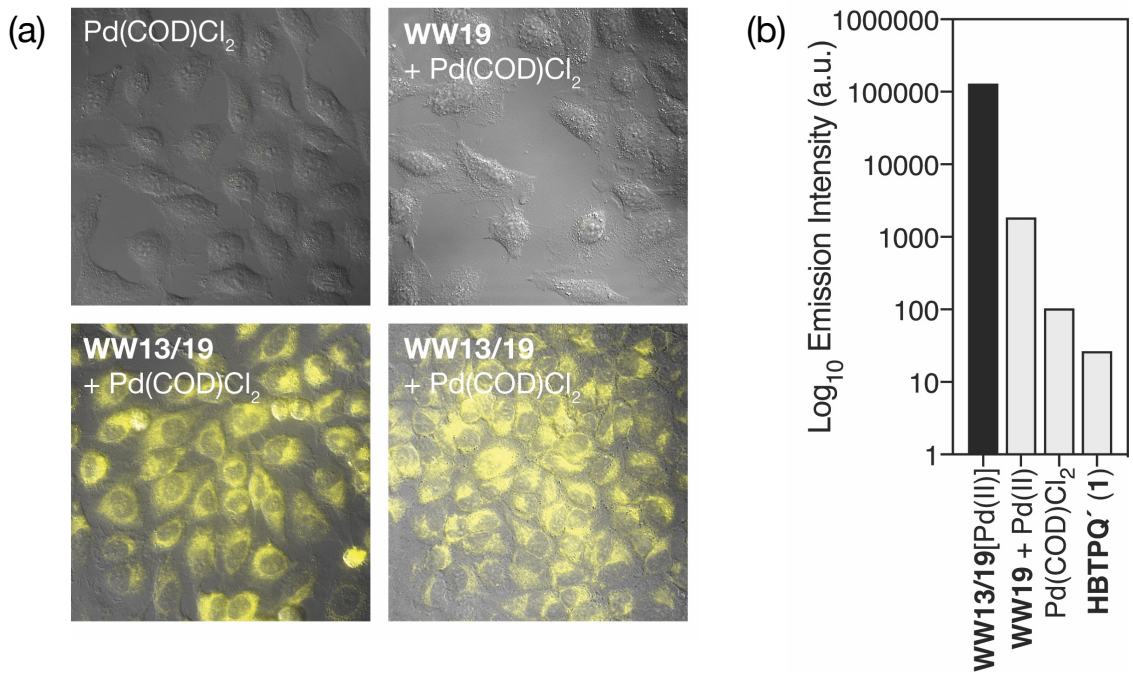


Figure 54. (a) Fluorescence microscopy of HeLa cells incubated with a 50 μM solution of **1** for 1 h washed twice with PBS and incubated for 1 h with a with $PdCl_2$(COD), a mixture of **WW19** with $PdCl_2$(COD), and a mixture of **WW13/19** with $PdCl_2$(COD). The bottom right image corresponds to the reverse experiment in which the cells are incubated first with the palladoprotein and then with the probe. Micrographs show the emission intensity at 620–680 nm upon irradiation at 405 nm; (b) quantification of the intracellular emission. Note the logarithmic scale in the intensity.

It is also important to note that this intracellular depropargylation reaction could also be carried out efficiently with the TMR-**WW13/19**[Pd(II)] analog (**Figure 55**).

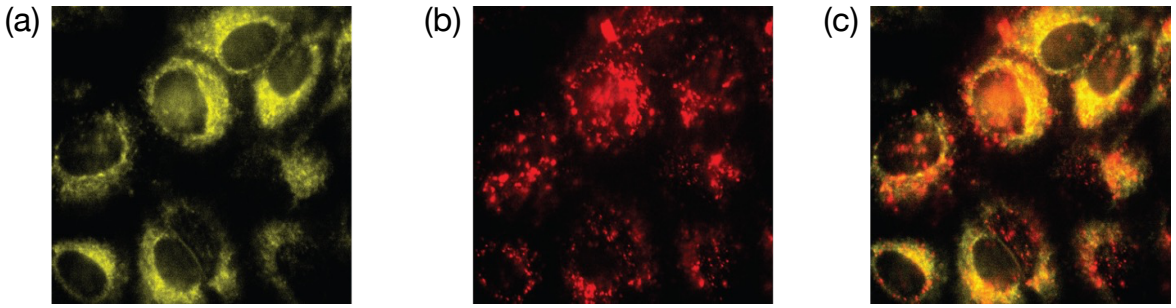


Figure 55. Fluorescence microscopy of HeLa cells incubated with a 50 μM solution of the fluorogenic probe **1** for 30 min, washed twice with PBS, and incubated for 1 h with TMR-**WW13/19**[Pd(II)], (a) emission of the deprotected probe **2**; (b) emission of the TMR fluorophore attached to **WW13/19**; (c) colocalization of both images.

Encouraged by the apparent robustness of the new β-sheet-based catalyst, **WW13/19**[Pd(II)], we investigated whether it could catalyse the reaction in other cell lines, such as HepG2 cells, a human liver cancer cell line.

The first step was to study the internalisation of the **WW13/19** miniprotein and its palladium complex, whose behaviour is analogous to that previously observed in HeLa cells. This means that the TMR-**WW13/19**[Pd(II)] palladoprotein internalises better than the TMR-**WW13/19** precursor without palladium, as shown by the intracellular fluorescence intensity (**Figure 56**).

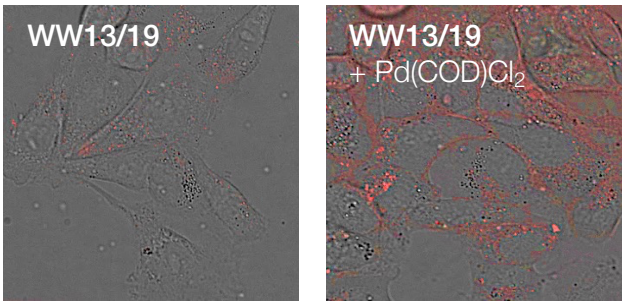


Figure 56. Fluorescence microscopy of internalization of 5 μM solutions of TMR-labeled peptide **WW13/19** and its palladoprotein derivative **WW13/19**[Pd(II)].

Depropargylation reaction was then carried out in cell medium following the same protocol described for HeLa cells. Although the result was not as favourable as in the previous case, a similar behaviour was obtained, which supports the generality of the approach. After one hour of incubation, we observed an increase in fluorescence intensity in cells treated with our **WW13/19**[Pd(II)] catalyst, compared to those treated with the non-coordinating mono-His peptide, WW19 (**Figure 57**).

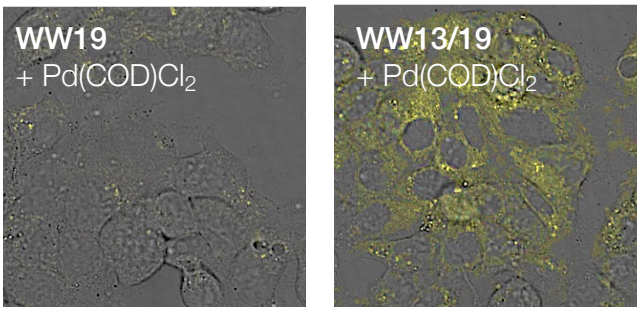


Figure 57. Fluorescence microscopy of HepG2 cells incubated with a 50 μM solution of **1** for 1 h, washed twice with PBS and incubated for 1 h with a mixture of **WW19** with $PdCl_2$(COD), and a mixture of **WW13/19** with $PdCl_2$(COD).

Importantly, previous studies have shown that standard palladium sources, such as $PdCl(allyl)_2$ or $Pd(OAc)_2$, fail to promote this type of reaction inside HeLa cells,[233] probably due to the low cellular penetration of palladium species and their rapid deactivation in the presence of cellular components. Therefore, in this work we have developed an efficient strategy for the encapsulation of Pd(II), which has allowed us to carry out catalytic reactions inside the cells.

[233] Miller, M. A. *et al. Nat. Commun.* **8**, 15906–15919 (2027).

5.2.3.6. Intracellular Depropargylation Reaction with other Probes

Encouraged by these results, we set out to evaluate the versatility of our catalytic system against other fluorogenic probes, thus opening the way to potential applications in drug development.

As a first alternative model, we selected resorufin, a widely used fluorophore due to its exceptional photophysical properties, such as long excitation and emission wavelength (572/585 nm), high fluorescence quantum yield (ϕ = 0.74 in water), [234] and biocompatibility.[235] Although the synthesis procedure of resorufin is not very clear in recent literature, this fluorophore is commercially available and affordable. Moreover, its chemical structure allows easy functionalization by modification of the 7-hydroxy group, which weakens its intramolecular charge transfer (ICT) process, leading to fluorescence quenching. However, the deprotection of the substituent groups of the probes allows the push-pull process to be recovered (**Figure 58**). [236]

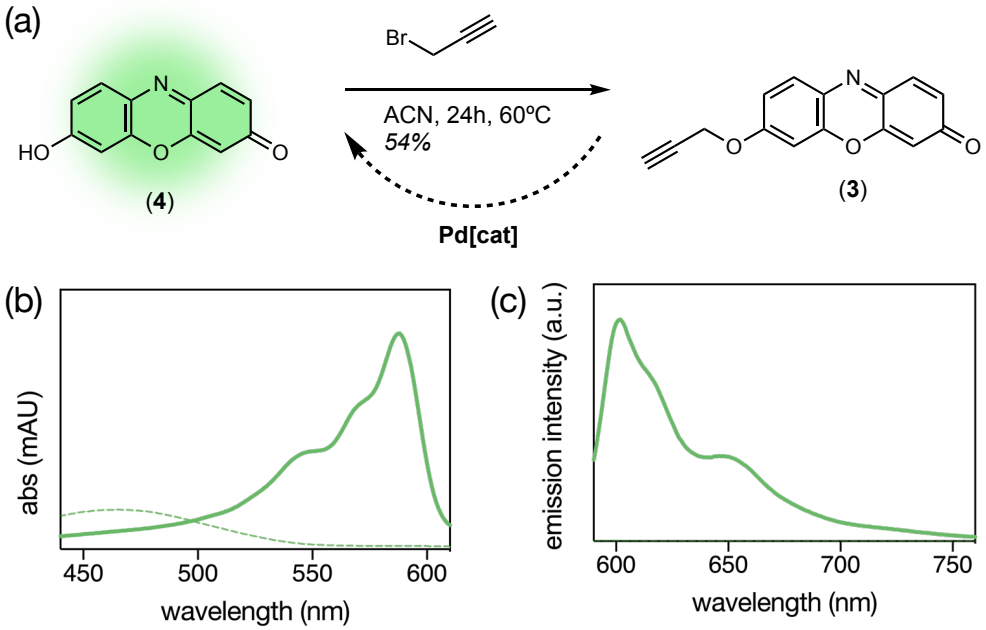


Figure 58. (a) Synthetic route to obtain the propargylated fluorogenic probe **3**; (b) absorbance spectra of the propargylated fluorogenic probe **3** (dashed line) and the deprotected product **4** (thick solid line); (b) fluorescence emission spectra of **3** (dashed line) and **4** (thick solid line). Measurements taken at 20 μM, 1:1 v/v DMSO/H_2O, λ_{exc} = 572 nm.

---

[234] Im, H. G. *et al. Org. Biomol. Chem.* **11**, 18, 2966–2971 (2013).

[235] Hitomi, Y., Takeyasu, T. Kodera, M. *Chem. Commun.* **49**, 85, 9929–9931 (2013).

[236] (a) Simpson, D. J. *et al. J. Org. Chem.* **56**, 18, 5391–5396 (1991); (b) Li, Z. *et al. Anal. Chem.* **85**, 8, 3926–3932 (2013); (c) Y. Shiraishi, *et al. Chem. Commun.* **49**, 99, 11680–11682 (2013).

The *in vitro* assays were first performed following the same protocol as previously described for probe **1**. As shown in **Figure 59**a, the most efficient catalyst was the $PdCl_2$(COD) complex, with a yield of 43%. However, the **WW13/19**[Pd(II)] complex was also able to catalyze the depropargylation of the resorufin probe, achieving a comparable 36 % yield. Importantly, when the non-coordinating peptides **WW0** and **WW19** were used and subjected to an ultra-diafiltration step using 3 kDa Amicon filters to remove free palladium, no fluorescence was detected, confirming the need for a stable palladium complex for efficient catalysis.

Given the positive result obtained *in vitro*, we decided to assess the feasibility of performing depropargylation of the resorufin probe in a cellular environment. For this purpose, HeLa cells were incubated with a 50 μM solution of probe **3** for one hour at 37 °C. Subsequently, a freshly prepared 50 μM solution of the **WW13/19**[Pd(II)] complex was added. After an additional hour of incubation, fluorescence signal was observed by microscopy, indicating intracellular formation of resorufin. In contrast, control experiments performed with the $PdCl_2$(COD) complex, as well as with the **WW0** and **WW19** peptides mixed with $PdCl_2$(COD), showed no intracellular emission, in agreement with our earlier observevations for probe **1** (**Figure 59**b).

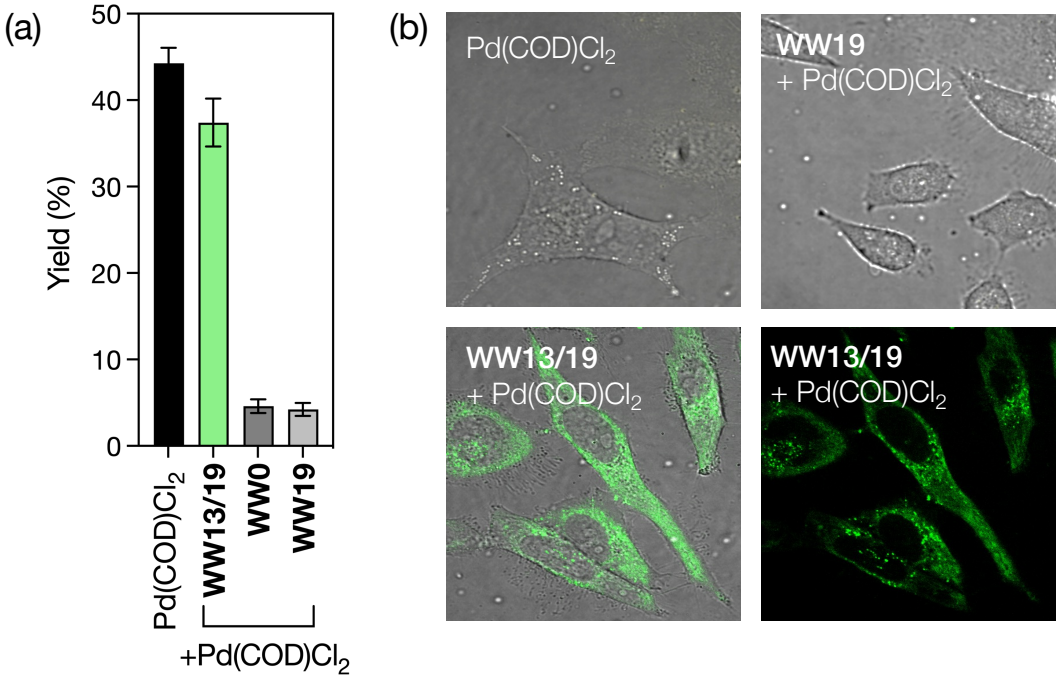


Figure 59. (a) Bar diagram representation of the yields obtained for each catalyst. Reaction conditions: mixing for 24 h a 200 μM solution of probe **3** with 20 μM mixtures of each peptide mixed with $PdCl_2$(COD) in a 1:1 ratio for 1 h in 10 mM phosphate, 100 mM NaCl, pH 7.5 buffer. Yields were calculated by HPLC-MS(ESI) using internal standards; (c) Fluorescence microscopy of HeLa cells incubated with a 50 μM solution of **3** and incubated for 1 h with $PdCl_2$(COD), a mixture of **WW19** with $PdCl_2$(COD), and a mixture of **WW13/19** with $PdCl_2$(COD). Micrographs show the emission intensity at 525–575 nm upon irradiation at 488 nm.

In addition, we aimed to investigate whether this chemistry could be applied to more complex substrates, such as the protected bis-propargyl form of cresyl violet (**5**) [237] (**Figure 60**a). Deprotection of this compound releases cresyl violet (**6**), a xanthene-type fluorophore widely used for its excellent photostability and high fluorescence quantum yield.[238] This fluorophore emits in the red region of the spectrum (λ_{em} = ~ 635 nm), which allows for quantitative self-calibrating detection, especially useful in biological environments[239] (**Figure 60**b and c).

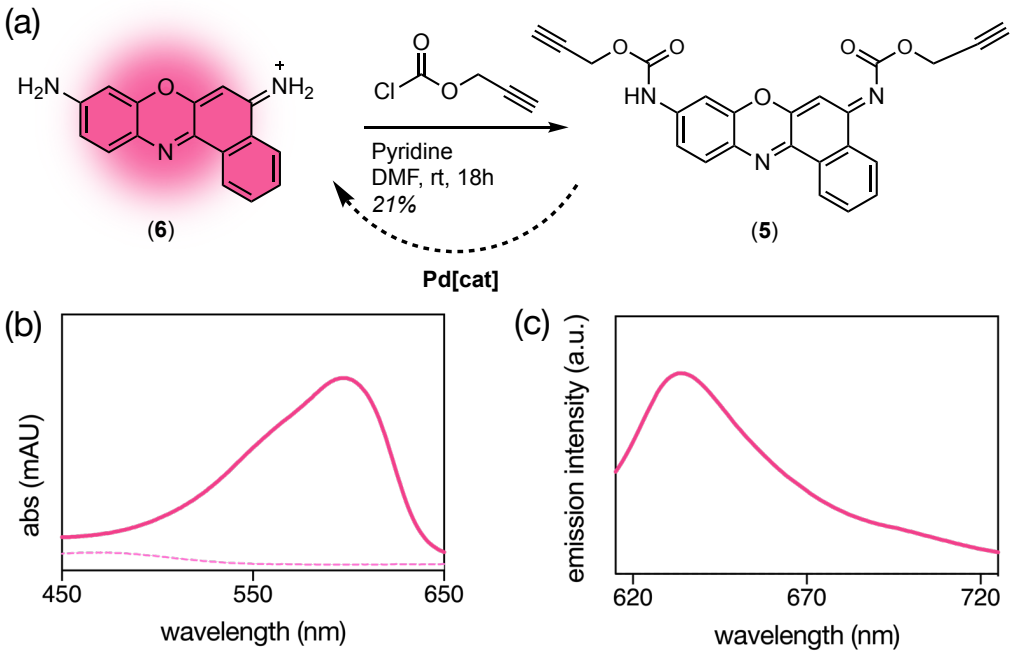


Figure 60. (a) Synthetic route to obtain the propargylated fluorogenic probe **5**; (b) absorbance spectra of the propargylated fluorogenic probe **5** (dashed line) and the deprotected product **6** (thick solid line); (b) fluorescence emission spectra of **5** (dashed line) and **6** (thick solid line). Measurements taken at 20 μM, 1:1 v/v DMSO/H_2O, λ_{exc} = 600 nm.

As shown in **Figure 61**a, this reaction shows very low yields under *in vitro* conditions. In fact, the $PdCl_2$(COD) complex shows no appreciable catalytic activity, and our **WW13/19**[Pd(II)] system barely achieves a yield of more than 5 %. However, when incubating HeLa cells with probe **5** (50 μM, 1 h), followed by the same protocol described for the other probes, a marked increase in cell fluorescence was observed in samples treated with **WW13/19**[Pd(II)],

[237] Li, X. *et al. Biosens. Bioelectron.* **200**, 113929 (2022).

[238] (a) Urrutia, M., Ortiz, C. *Biotech. Histochem.* **90**, 159–166 (2015); (b) Magde, D. *et al. J. Phys. Chem.* **83**, 696–699 (1979); (c) Gong, Q. *et al. Chem. Sci.* **7**, 788–792 (2016); (d) Zhang, J. *et al. J. Am. Chem. Soc.* **133**, 14109–14119 (2011).

[239] (a) Bae, S. K. *et al. J. Am. Chem. Soc.* **135**, 9915–9923 (2013); (b) Haidekker, M. A. *et al. J. Am. Chem. Soc.* **128**, 398–399 (2006); (c) Han, X. *et al. Chem. Sci.* **7**, 5098–5107 2016). (d) Li, X. Y. et al. *Anal. Chem.* **91**, 11409–11416 (2019).

compared to those treated with the non-coordinating mono-His peptide **WW19** or with the $PdCl_2$(COD) complex (**Figure 61**b).

This contrast between the results *in vitro* and in cells could be explained by several factors. In the intracellular environment, it could stabilize the probe in its active form, leaving the propargyl groups more exposed. In addition, internalization of the **WW13/19**[Pd(II)] catalytic system could improve its proximity and availability for the reaction. Taken together, these results once again underline the efficacy and potential applicability of **WW13/19**[Pd(II)] as a bioorthogonal catalyst in cellular contexts.

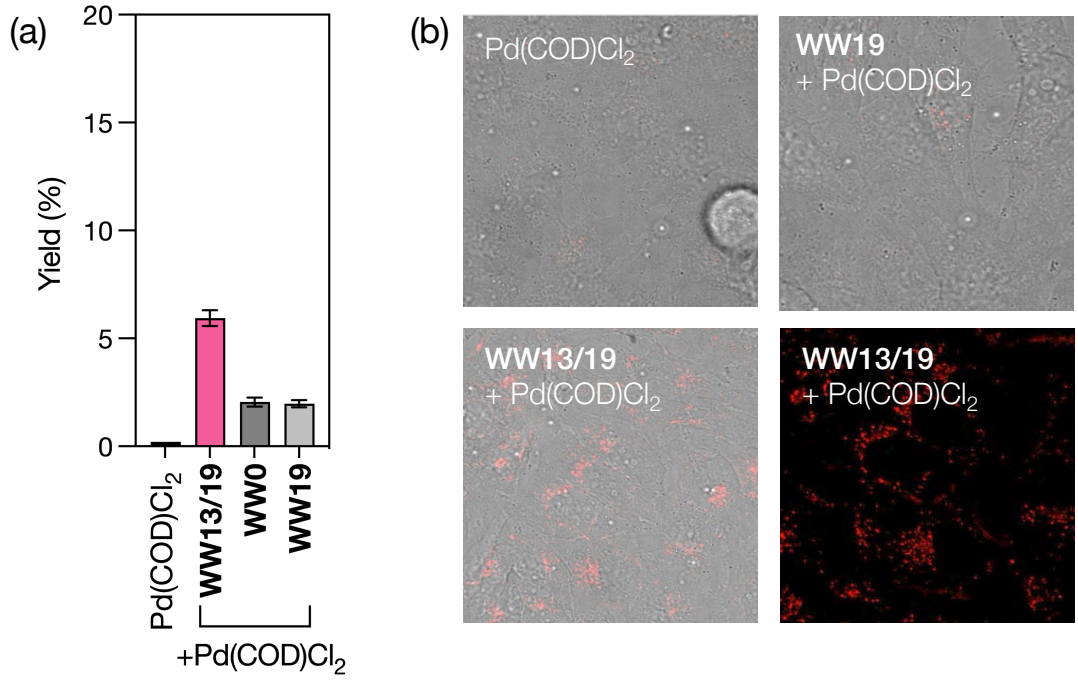


Figure 61. (a) Bar diagram representation of the yields obtained for each catalyst. Reaction conditions: mixing for 24 h a 200 μM solution of derivative **5** with 20 μM mixtures of each peptide mixed with $PdCl_2$(COD) in a 1:1 ratio for 1 h in 10 mM phosphate, 100 mM NaCl, pH 7.5 buffer. Yields were calculated by HPLC-MS(ESI) using internal standards; (b) fluorescence microscopy of HeLa cells incubated with a 50 μM solution of **5** and incubated for 1 h with $PdCl_2$(COD), a mixture of **WW19** with $PdCl_2$(COD), and a mixture of **WW13/19** with $PdCl_2$(COD). Micrographs show the emission intensity at 620–700 nm upon irradiation at 561 nm.

5.2.4. Conclusion

In this work we have described an unprecedented artificial metalloprotein that could be considered as a non-natural proto-metalloenzyme, since it promotes a catalytic depropargylation reaction in living mammalian cells, although still with modest activity. Through protein engineering, we succeeded in integrating Pd(II) complexes into WW domain-based miniproteins, which allowed us to stabilize the metal and protect it from rapid deactivation in cellular media.

Our structural studies, combining HPLC-MS(ESI), CD, fluorescence and NMR spectroscopy analysis, confirmed the correct incorporation of the metal center and the preservation of protein folding. In addition, we demonstrated that these palladoproteins can catalyze depropargylation reactions inside HeLa cells, which resulted in increased fluorescence in assays with fluorogenic probes.

Compared to conventional palladium complexes, our catalysts showed significantly improved activity in cellular environments, highlighting their potential as bioorthogonal tools for applications in biocatalysis and biomedicine. These results open new perspectives for the design of more robust and versatile artificial metalloproteins capable of expanding bioorthogonal chemistry in complex biological environments.

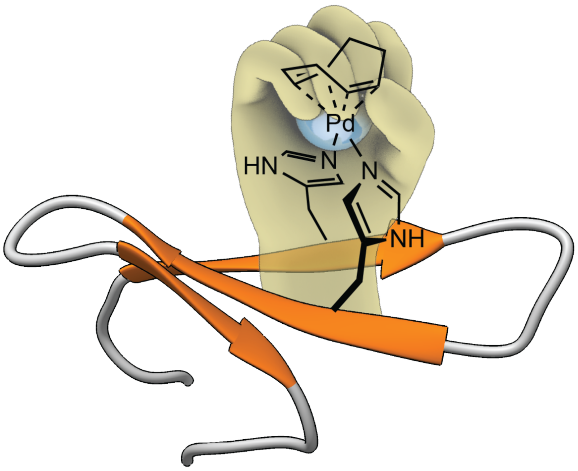


Figure 62. Illustration of our unnatural proto-metalloenzyme, highlighting its ability to stabilize Pd(II) complexes and protect them from rapid deactivation in cellular media.

5.3. Identifying Metallopeptides for Intracellular Catalysis Using Combinatorial Libraries

Summary

The discovery and development of artificial catalysts to carry out biorthogonal reactions in living cells is a primary goal at the interface of Chemistry and Biology. Current approaches rely on time-consuming trial-and-error methods that often fail to yield an optimal catalyst. In contrast, we show that positionally addressable combinatorial libraries (SPOT libraries) provide a significant advantage for the efficient identification of novel catalytic metallopeptides. Using these libraries, we were able to rapidly identify catalytic β-hairpin metallopeptides capable of promoting efficient depropargylation and click reactions, even in challenging intracellular environments.

Results of this section have been published in ACS Catalysis:

Carmen González-González,[a,] Laura Martínez-Castro,[b,] Soraya Learte-Aymamí,[a,] Clara Pose-Insua,[a] José R. Couceiro,[c] Pau Martin-Malpartida,[d] Maria J. Macias,[d,e] Jean-Didier Maréchal,[b] José L. Mascareñas,[a] M. Eugenio Vázquez[a] *ACS Catal.* **15**, 8624–8632 (2025) https://doi.org/10.1021/acscatal.5c00525

Affiliations:

a Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Departamento de Química Orgánica, Universidade de Santiago de Compostela. Santiago de Compostela 15705, Spain.

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

c Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Universidade de Santiago de Compostela. Santiago de Compostela 15705, Spain.

d. Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Baldiri Reixac, 10, Barcelona 08028, Spain.

e. Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, Barcelona 08010, Spain.

Information in this chapter is reproduced from Ref. [240] with permission from *Americal Chemical Society.*

[240] González-González, C., *et al. ACS Catal* **15**, 8624–8632 (2025).

5.3.1. Background

Bioorthogonal transformations are invaluable tools for studying and manipulating biological systems.[241] Therefore, the development of *artificial metalloenzymes* capable of catalyzing novel transformations and extending the range of biocompatible reactions in living cells represents a major challenge at the crossroads of catalysis, protein engineering and chemical biology.[242]

As mentioned in chapter two, catalytic peptides represent a promising alternative to ArM. These metallopeptides offer significant advantages,[243] such as chiral scaffolding,[244] proximity-enhanced selectivity[245] and tunable physicochemical properties. Although artificial amino acids with metal-binding ligands can improve coordination, their complexity in protein expression limits their practical applicability.[246]

Despite these advantages, these catalytic metallopeptides still lag behind natural enzymes.[247] One of the main challenges in the development of catalytic metallopeptides is the creation of a functional active site. A common strategy is to incorporate catalytically active groups, typically (organo)metal complexes, into proteins with well-defined and stable structures. These systems are optimised by mutagenesis and combinatorial methods, an efficient, but often time-consuming and laborious process.[248] Although combinatorial methods represent a potentially more practical strategy, their application has been limited by the difficulties involved in selecting, within large libraries, those sequences that exhibit catalytic activity.[249]

SPOT libraries, introduced in 1992,[250] involve the synthesis of peptides on flat matrices, usually cellulose sheets, by localized application of reagents at defined positions on the solid

---

[241] (a) Devaraj, N. K. *ACS Cent. Sci.* **4**, 952−959 (2018); (b) Li, Y., Fu, H. *ChemistryOpen* **9**, 835−853 (2020); (c) Seoane, A., Mascareñas, J. L. *Eur. J. Org. Chem.* **32** (2022); (d) Scinto, S. L. *et al. Nat. Rev. Methods. Primers.* **1**, 30−23 (2021); (e) Madec, H. *et al. Chem. Sci.* **14**, 409−442 (2023).

[242] (a) Davis, H. J., Ward, T. R. *ACS Cent. Sci.* **5**, 1120–1136 (2019); (b) Wittwer, M. *et al. Nature Catal.* **4**, 814–827 (2021); (c) Bos, J., Roelfes, G. *Curr. Opin. Chem. Bio.l* **19**, 135–143 (2014).

[243] Jarvo, E. R., Miller, S. J. *Tetrahedron* **58**, 2481–2495 (2002).

[244] (a) Akagawa, K. *et al. J. Org. Chem.* **81**, 9396–9401 (2016); (b) Zheng, L. *et al. Angew Chem. Int. Ed. Engl.* **53**, 7599–7603 (2014).

[245] Kinghorn, M. J. *et al. ACS Catal.* **7**, 7704–7708 (2017).

[246] Metrano, A. J. *et al. Chem. Rev.* **120**, 11479–11615 (2020).

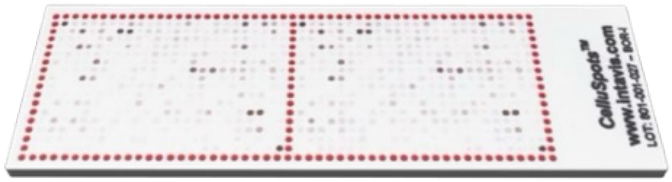
[247] Learte-Aymamí, S. *et al. Commun. Chem.* **5**, 1, 75 (2022).

[248] Akcapinar, G. B., Sezerman, O. U. *Biosci. Rep.* **37***,* 2, BSR20160179 (2017).

[249] (a) Huang, X., Pieczko, M. E., Long, E. C. *Biochemistry* **38**, 2160–2166 (1999); (b) Berkessel, A., Hérault, D. A. *Angew. Chem. Int. Ed.* **38**, 102–105 (1999); (c) Francis, M. B., Jacobsen, E. N. *Angew. Chem. Int. Ed.* **38**, 937–941 (1999); (d) Copeland, G. T., Miller, S. J. *J. Am. Chem. Soc.* **123**, 6496–6502 (2001); (e) Hagemeyer, A. *et al. Appl. Catal. A* **221**, 23–43 (2001); (f) Berkessel, A. *Curr. Opin. Chem. Biol.* **7**, 409–419 (2003); (g) Kerton, F. M., Whitwood, A. C., Willans, C. E. *Dalton Trans.* **15**, 2237–2244 (2004); (h) Sambasivan, R., Ball, Z. T. *Angew. Chem. Int. Ed.* **51**, 8568–8572 (2012); (i) Maeda, Y. *et al. Annu. Rev. Biomed. Eng.* **18**, 311–328 (2016).

[250] Frank, R. *Tetrahedron* **48**, 9217–9232 (1992).

support[251] (**Figure 63**). This design allows precise placement of each library component within the matrix, which greatly simplifies the screening process.[252] Thanks to their simplicity and versatility, SPOT peptide libraries have become a widely used tool in biology, especially for the study of protein-protein interactions, such as epitope mapping, enzyme substrate analysis, antibody specificity or interactome profiling.[253] Interestingly, despite their potential, these libraries have been barely explored in the field of catalysis, with only one isolated case reported in which they were used to identify an organocatalytic peptide esterase.[254]



2 × 384 spots
76 × 26 mm microscope slide
~20 pmol peptide/spot (~0.8 mm)

Figure 63. Example of a SPOT library plate.

For all these reasons, SPOT technology could offer an exceptionally versatile platform for the identification of catalytic peptides, allowing the generation of positioned libraries of manageable size that can be efficiently screened using fluorogenic substrates.

As a starting point, we selected β-hairpins as the structural scaffold for our libraries due to their simple yet efficient architecture. These structures feature a concave and extended geometry that offers multiple advantages: not only do they facilitate the precise arrangement of the chelating residues around the metal ion, but they also allow the introduction of additional side chains close to the active site, which is key to optimize the reactivity and stability of the catalytic complex. Within this structural family, we focus on *Tryptophan zipper motif* (Wzip)[255] (**Figure 64**), short monomeric β-hairpins (12 or 16 residues) stabilized by Trp-Trp interactions, which have demonstrated reversible and cooperative folding and exhibit well-characterized NMR structures (PDB: 1LE0, 1LE1, 1LE3). These features make Wzip an ideal platform for the rational design of combinatorial libraries using SPOT technology.

---

[251] (a) Hilpert, K., Winkler, D. F. H., Hancock, R. E. W. *Nat. Protoc.* **2**, 1333–1349 (2007); (b) Hilpert, K., Winkler, D. F. H., Hancock, R. E. W. *Biotechnol. Genet. Eng. Rev.* **24**, 31–106 (2007); (c) Reineke, U., Volkmer-Engert, R., Schneider-Mergener, J. *Curr. Opin. Biotechnol.* **12**, 59–64 (2001).

[252] Frank, R. *et al. In Combinatorial Peptide and Nonpeptide Libraries* 363–386 (Wiley-VCH Verlag GmbH, 2007).

[253] (a) Thiele, A., Stangl, G. I., Schutkowski, M. *Mol. Biotechnol.* **49**, 283–305 (2011); (b) Volkmer, R., Tapia, V., Landgraf, C. *FEBS Lett.* **586**, 2780–2786 (2012); (c) Engelmann, B. W. *Methods Mol. Biol.* **1555**, 375–394 (2017); (d) Sandmann, C. L. et al. *Mol. Cell* **83**, 994–1011.e18 (2023).

[254] Biswas, R., Maillard, N., Kofoed, J., Reymond, J. L. *Chem. Commun.* **46**, 8746–8748 (2010).

[255] Cochran, A. G., Skelton, N. J., Starovasnik, M. A. *Proc. Natl. Acad. Sci. USA* **98**, 5578–5583 (2001).

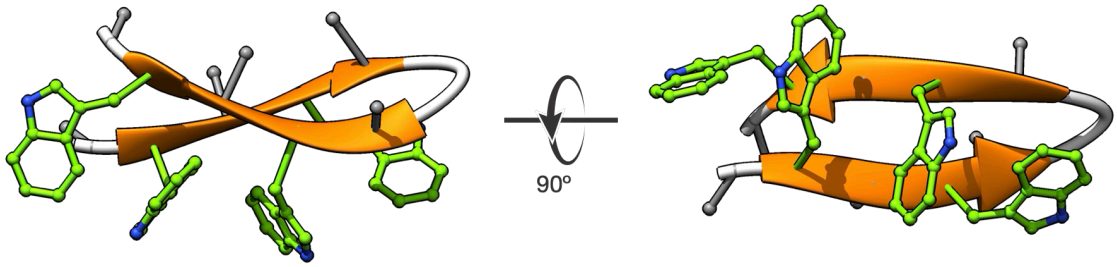


Figure 64. Two-sided representation of the structure of the tryptophan zip motif (Wzip), a short monomeric β-hairpin (12 residues) stabilised by Trp-Trp interactions (PDB: 1LE0).

As a proof of concept, we build on our previous chapter 2 and using these libraries succeed in identifying catalytic β-hairpin paladopeptides capable of promoting efficient depropargylation reactions in complex intracellular environments.

This same strategy holds particular promise for the design and selection of copper metallopeptides capable of catalyzing azide-alkyne cycloaddition (CuAAC), a fundamental transformation in bioorthogonal chemistry due to its high efficiency and selectivity, which has been widely applied in drug development, biomaterials and chemical biology studies.[256] However, the implementation of CuAAC in living cells remains a challenge, mainly due to the toxicity of free Cu(I) and its inactivation by cellular biomolecules.[257] To overcome these limitations, Wzip hairpin-based SPOT libraries can be effective tools for the identification of peptide ligands that efficiently stabilize Cu(I), reduce oxidative stress and promote catalysis under cellular conditions. The use of peptide ligands ensures biocompatibility and provides extraordinary versatility to tailor the biological properties of such ligands by increasing cell permeability, [258] crossing the blood-brain barrier, [259] targeting tumors [260] or encoding intracellular localization by introducing specific markers.[261]

---

[256] Barrow, A. S. *et al.*, *Chem. Soc. Rev.* **48**, 4731–4758 (2019).

[257] (a) Pickens, C. J. *et al. Bioconjug. Chem.* **29**, 686–701 (2018); (b) Uttamapinant, C. *et al. Angew. Chem. Int. Ed.* **51**, 5852–5856 (2012); (c) Bevilacqua, V. *et al. Angew. Chem. Int. Ed.* **53**, 5872–5876 (2014); (d) Besanceney-Webler, C. *et al. Angew. Chem. Int. Ed*. **50**, 8051–8056 (2011).

[258] Pescina, S. *et al. J. Control. Release* **284**, 84–102 (2018).

[259] Raucher, D. *Curr. Opin. Pharmacol.* **47**, 14–19 (2019).

[260] Zhao, N. *et al. Anticancer Agents Med. Chem.* **18**, 74–86 (2018).

[261] (a) Pap, E. H. *et al. Exp. Cell Res.* **265**, 288–293 (2001); (b) Rajendran, L., Knölker, H. J., Simons, K. *Nat. Rev. Drug Discov.* **9**, 29–42 (2010).

5.3.2. Objective

Low molecular weight peptides are promising catalysts for bioorthogonal reactions due to their modularity, which enables fine-tuning of reactivity and selectivity. This project leverages our expertise in peptide engineering and coordination chemistry to develop catalytic β-hairpin metallopeptides for intracellular reactions. Our approach utilizes SPOT positionally addressable combinatorial libraries combined with fluorogenic screening to efficiently identify catalytic β-hairpins based on the Tryptophan zipper motif. These novel peptide catalysts could expand cellular metabolism by enabling exogenous chemical transformations, potentially leading to new selective therapeutic strategies with activatable drugs.

In this section, we show the application of this methodology for the discovery of functional catalytic metallopeptides, focusing our efforts on two main objectives: the development of palladium metallopeptides for bioorthogonal depropargylation-type reactions, and copper metallopeptides capable of efficiently catalyzing azide-alkyne cycloaddition (CuAAC) in complex biological media.

5.3.3. Results and Discussion

5.3.3.1. SPOT Library Design and Screening

To design our SPOT library, we selected the tryptophan zip peptide (Wzip) family as β-hairpin scaffold. These are short peptides (12-16 residues) with well-defined, monomeric β-hairpin structure stabilized by the stacking of Trp side chains.[262] The structures of several of these peptides have been determined by NMR, and the Trpzip has been found to be an efficient scaffold for metal binding, [263] facilitating the structurally guided design of specific combinatorial libraries.

Within this family of peptides, we chose to design the library taking the Wzip1 dodecamer, $S^1WT^3WE^5GNK^8WT^{10}WK^{12}$, as a starting point. This model peptide has a central β-II' turn (E^5GNK^8), with two WTW motifs on either side, whose Trp residues intertwine their side chains, forming the hydrophobic core of the hairpin. In designing our SPOT library, we sought to retain both the central GN sequence, which defines the β-turn, and the tryptophan core, considering them essential structural elements. Therefore, the positions explored to identify optimal catalytic sequences are 1, 3, 5, 8, 10 and 12 ($x^1Wx^3Wx^5GNx^8Wx^{10}Wx^{12}$), all oriented towards the concave side of the hairpin, which allows efficient bidentate coordination with a Pd(II) complex, while providing steric shielding to prevent its rapid deactivation in complex aqueous media (**Figure 65**).

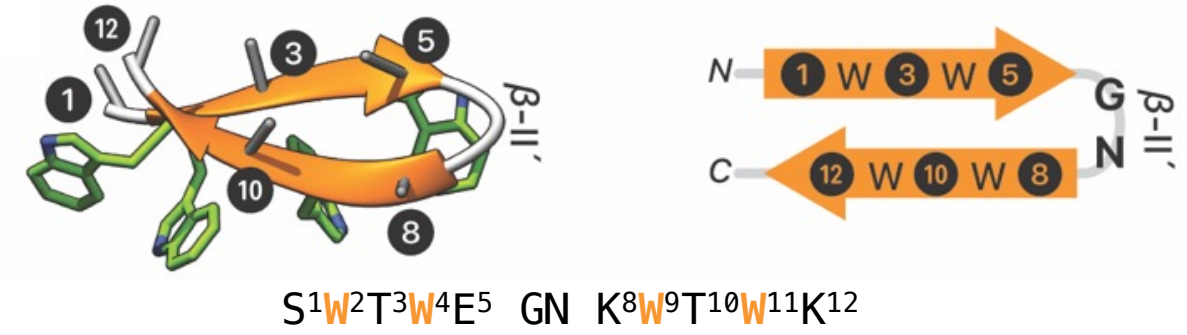


Figure 65. Left: The structure of a Wzip hairpin (PDB: 1LE0) indicating as stubs the positions mutated in the screening (1 to 12 from *N*- to *C*- terminus), right: schematic representation of the Wzip hairpin indicating the positions of the screening.

We evaluated a total of 11 possible ways of introducing two adjacent His residues within the β-hairpin of the *Wzip1* peptide (e.g. His^1/His^3, His^1/His^{12}, His^1/His^{10}, His^3/His^{12}, His^3/His^{10}, His^3/His^8, among others). In each of these designs, the other four remaining positions of the peptide were randomized. However, performing a full saturation mutagenesis with 14 common amino acids—excluding His, Trp and others with coordination capacity such as Cys,

---

[262] Cochran, A. G., Skelton, N. J., Starovasnik, M. A. *Proc. Natl. Acad. Sci. USA* **98**, 10, 5578–5583 (2001).

[263] Pham, T. L. *et al. ChemBioChem* **24**, 3, e202200588 (2023).

Met, Asp and Glu—would have generated a large library, exceeding one 100,000 peptides. Therefore, we opted for a model library using only four amino acids (Ala, Val, Thr, and Arg), which were selected for their tendency to form β-sheet structures,[264] their physicochemical properties, and their good behavior during solid-phase peptide synthesis.[265] In this way, we generated a more manageable library of 264 peptides using a bespoke Python script, written by Laura Martínez-Castro from the *Insilichem* group at the UAB. This program allows combining different histidine pairs at coordinating positions and generating all possible combinations of the four selected amino acids at the remaining positions, thus obtaining all the library sequences. Once synthesized, the SPOT library was printed on a cellulose support and transferred to a slide using the CelluSPOT technique, allowing the screening assay to be miniaturized.[266]

For screening, the same probe described in Chapter 2 (**1**) was used, as it exhibits low emission when protected, but becomes highly fluorescent after depropargylation (**Figure 66**). To carry out the experiment, CelluSPOT slides with the complete library were incubated with a solution of the palladium precursor dichloro(1,5-cyclooctadiene)palladium(II), $PdCl_2(COD)$, in PBS buffer for 1 h, with the aim of forming the metallopeptides directly on the matrix. Then, a thorough wash with PBS was performed to remove any unbound catalytically active salts.[267]

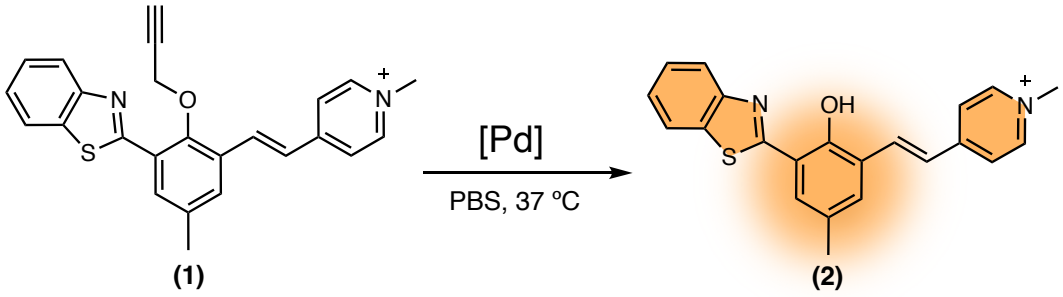


Figure 66. Propargylated probe **1** and its deprotection reaction to obtain the fluorescent product **2**.

The matrix was then incubated with fluorogenic probe **1** (200 μM in PBS) and the progress of the depropargylation reaction was monitored by measuring the emission at 635 nm after excitation at 330 nm (**Figure 67**a). The screening was repeated three times and showed excellent reproducibility, especially in the identification of the most active and inactive sequences (with a standard deviation of 6.5 % among the top 25 results). As a preliminary test of the stability of catalytic metallopeptides in biologically relevant media, the screening was also carried out in DMEM, obtaining qualitative results like those obtained with PBS.

---

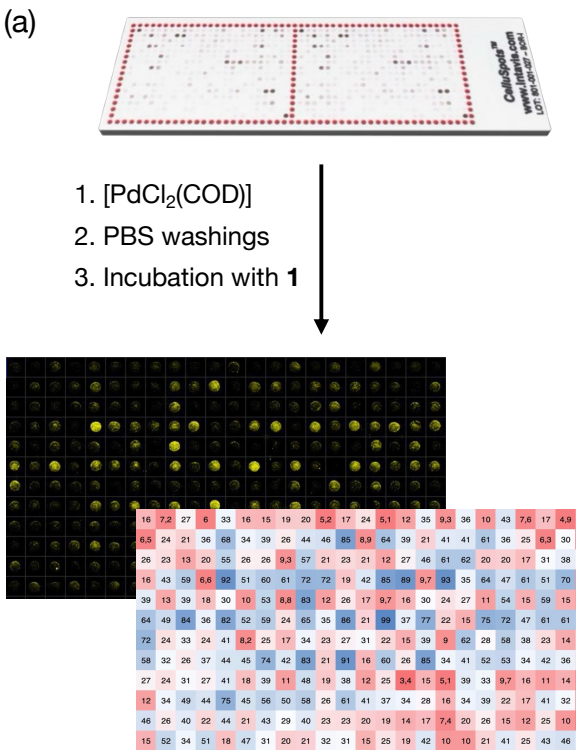
[264] Minor, D. L., Jr, Kim, P. S. *Nature* **367**, 6464, 660–663 (1994).

[265] Guixer, B. *et al. J. Pept. Sci.* **22**, 9, 577–591 (2016).

[266] Dikmans, A. *et al. QSAR Comb. Sci.* **25**, 11, 1069–1080 (2006).

[267] Learte-Aymamí, S. *et al. Angew. Chem. Int. Ed.* **59**, 23, 9149–9154 (2020).

An analysis of the frequency of different residues at each position within the top fifteen sequences[268] showed a clear preference for the location of the pair of His residues at positions His^3/His^{10}, facing each other across the β-hairpin, as evidenced by four of the most active catalysts (**D17**, **D6**, **D15** and **D14**) sharing this arrangement. A pattern for simpler peptides with two adjacent His residues in the C-terminal strand, located at positions His^{10}/His^{12} (**F14**, **F12**, **F6**, **F16** and **F4**), were also identified. The remaining six peptides do not belong to either of these two families, nor do they share among them an obvious common organization; however, even within this structural diversity, there is a marked tendency to use the His^3 or His^{10} positions as preferred sites for metal center coordination (**Figure 67**b).



(b)

		3	10	12	
1	99	A W R W V G N T W **H** W **H**			**F14**
2	93	A W **H** W V G N R W **H** W T			**D17**
3	92	R W **H** W V G N A W **H** W T			**D6**
4	91	T W **H** W V G N **H** W A W R			**H12**
5	89	A W **H** W T G N R W **H** W V			**D15**
6	85	T W V W A G N R W **H** W **H**			**F12**
7	85	**H** W T W V G N A W **H** W R			**B12**
8	84	A W **H** W R G N V W **H** W T			**D14**
9	84	A W **H** W T G N **H** W V W R			**H16**
10	84	R W A W V G N T W **H** W **H**			**F4**
11	83	T W **H** W A G N **H** W V W R			**H10**
12	83	T W **H** W A G N V W R W **H**			**E10**
13	82	R W V W A G N T W **H** W **H**			**F6**
14	77	A W T W V G N R W **H** W **H**			**F16**
15	75	R W V W **H** G N A W **H** W T			**J6**

Figure 67. (a) Screening of the library. The CelluSPOT slide was incubated for 1 h with a 1 mM solution of $PdCl_2$(COD), followed by a 10-min wash with PBS to remove any uncoordinated palladium. Next, the slide was incubated with a 200 μM solution of probe **1**, and the slide emission was recorded at 0, 4, and 24 h. The catalytic metallopeptides are identified based on their position in the array, and their relative catalytic activity quantified by the emission intensity of each spot; (b) the fifteen best sequences ranked by the amount of depropargylated product **2** after 24 h (green column, normalized average of the relative fluorescence emission intensity in three experiments), and their positions in the array.

In addition, a preference was observed for the Val residue at position 5, present in seven of the fifteen most active peptides identified in the screening (**F14**, **D17**, **D6**, **H12**, **B12**, **F4** and **F16**). On the other hand, the Thr residue was generally the least favored at most positions, except for the N- and C-terminal ends, as seen in peptides such as **D17**, **D6**, **H12**, **F12**, **F12**,

[268] Davis, H. J., Ward, T. R. *ACS Cent. Sci.* **5**, 7, 1120–1136 (2019).

D14, **H10**, **E10** and **J6**, where likewise the hydrophobic and bulky Val residue showed a tendency to be negatively selected.

Taking these results into account, a more detailed analysis of the sequences presenting the His^3/His^{10} combination was performed, since it was the most favored arrangement according to the screening. Within this family, a recurrent selection of the Arg residue was observed, predominantly at positions 8 and 5, while Val tended to occupy the opposite position. In addition, these combinations were frequently associated with the presence of Thr at the N- and C-terminal ends of the hairpin, positions 1 or 12. This trend is clearly reflected in sequences **D17** and **D14**, which ranked among the most active catalysts in the screening, second and eighth, respectively.

5.3.3.2. Investigating Palladium Coordination with Structured Peptides

Following the screening experiments, two representative β-hairpin peptides were synthesized to characterize their physicochemical properties and to evaluate in detail their coordinative and catalytic behavior in solution. **D14** (*A*^{1}W**H**^{3}W*R*^{5}GNVW**H**^{10}WT) was selected as one of the most active peptides within the family with the His^3/His^{10} combination. As a negative control, peptide **D4** (*R*^{1}W**H**^{3}W*A*^{5}GNVW**H**^{10}WT) was chosen. This peptide also belongs to the His^3/His^{10} family, but exhibits significantly reduced catalytic activity in CelluSPOT assays, quantified as the relative emission intensity at the corresponding spot. Both peptides have an almost identical sequence, differing only in an exchange of position between Ala^1 and Arg^5 on **D14**, which appear as Arg^1 and Ala^5 on **D4** (**Figure 68**).

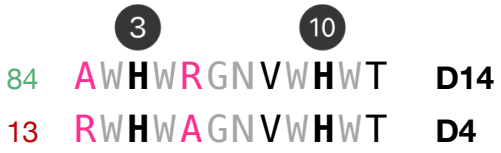


Figure 68. D14 and **D4** sequences of the His^3/His^{10} family, with the normalized average of three experiments on the left, sequence modifications are marked in pink

To study the metal-coordinating capacity of peptides **D14** and **D4**, equimolar amounts of each (500 μM) were incubated with $PdCl_2$(COD) in water, and the mixtures were analyzed by mass spectrometry (MS) at different incubation times. In the case of **D14**, a main peak corresponding to the metallopeptide complex **D14**[Pd(II)] was detected only 15 min after incubation, indicating efficient and rapid coordination with the metal. In contrast, the mixture of **D4** with $PdCl_2$(COD) showed only the peak associated with the free peptide, even after 24 h of incubation, suggesting practically no coordination capacity (**Figure 69**).

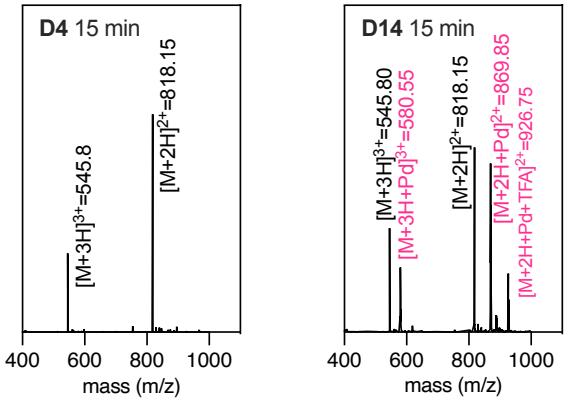


Figure 69. ESI-MS spectra of the mixture of **D4** and **D14** with 1.0 equiv. of $PdCl_2$(COD) for 15 min. The peaks corresponding to the apo-peptides are labeled in black, those of the corresponding palladopeptides, are in pink.

To monitor the coordination and formation of the palladoprotein complex in solution, we measured the decrease in emission intensity of tryptophan residues at 350 nm following the progressive addition of Pd(II) ions (**Figure 70**a). To this end, solutions of peptides **D14** and **D4** were prepared at 5 μM in HEPES buffer and titrated with increasing concentrations of Pd(II). The mixtures were incubated overnight to ensure equilibration, and their emission spectra were subsequently recorded.

The results confirmed that **D14** exhibits a significantly higher affinity for Pd(II) compared to **D4**. In the case of **D14**, the fluorescence intensity decreased sharply, following a 1:1 binding model, with an estimated dissociation constant of $K_D \approx 1.0$ μM (**Figure 70**b). In contrast, for **D4**, hardly any reduction in the fluorescence signal was detected, suggesting a very weak interaction with the metal ($K_D > 1$ mM) and, therefore, poor complex formation (**Figure 70**c).

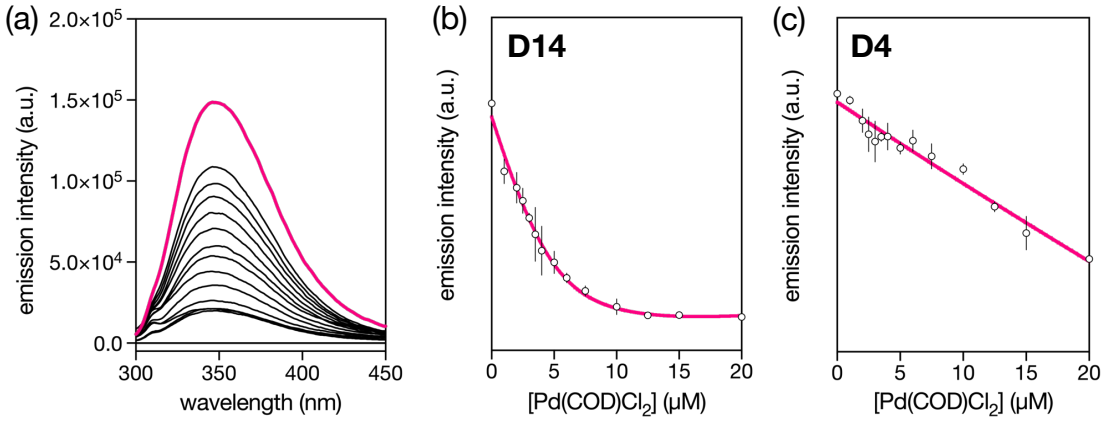


Figure 70. Pd(II) fluorescence titration of **D14** and **D4**. (a) fluorescence quenching of a 5 μM solution of **D14** (pink thick line) and after incubation overnight with Pd(II); tryptophan emission profile at 350 nm for 5 μM solutions of (b) **D14** and (c) **D4**, incubated overnight with in increasing concentrations of $PdCl_2$(COD) and the best fit to 1:1 binding model (K_D = 1.019 μM for **D14** and K_D = 3.865 mM for **D4**). All the measurements were conducted in 25 mM HEPES buffer, 150 mM NaCl, pH = 7.5.

To better understand the metal coordination of **D14**, its circular dichroism spectrum was measured, which, in the far UV region, exhibited a pair of positive excitons due to electronic coupling between the indole side chains.[269] This pattern changed significantly after the addition of $PdCl_2(COD)$ (**Figure 71**a).

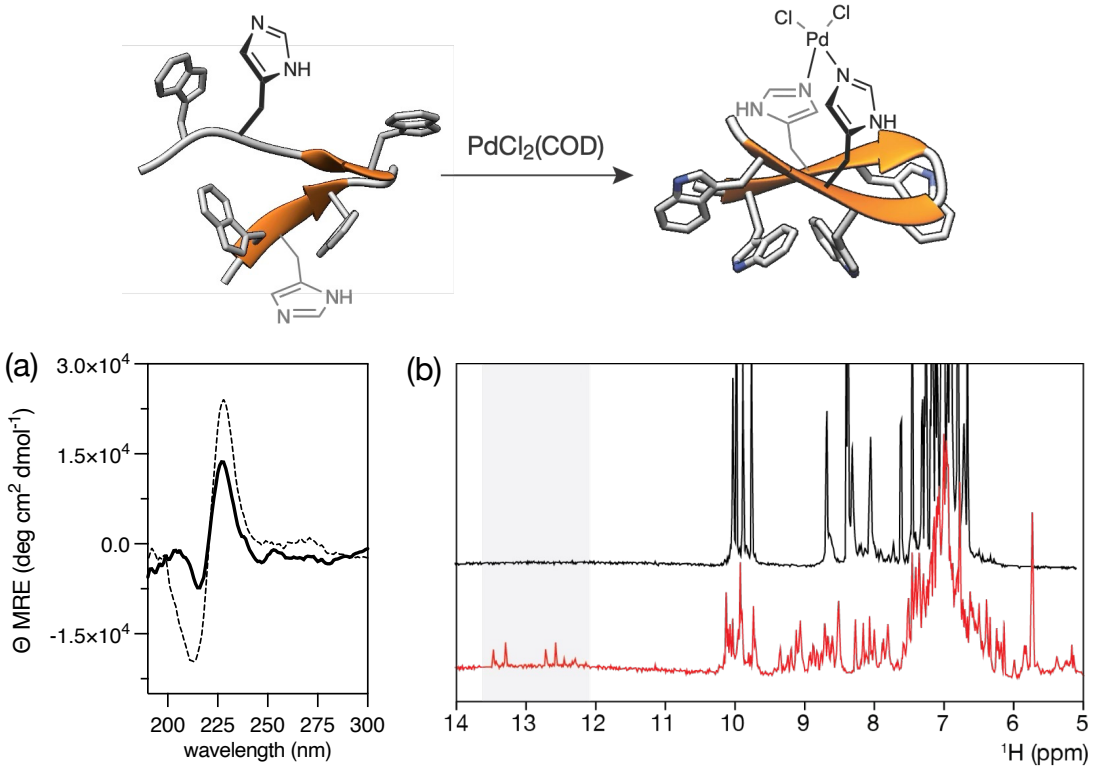


Figure 71. (a) CD spectra of 5 μM **D14** (dashed line) and 5 μM **D14** in the presence of and 1 equiv. of $PdCl_2(COD)$ (thick solid line), measurements in 10 mM phosphate buffer pH 7.5 and 100 mM of NaCl. (b) ^{1}D NMR data of the free peptide **D14** (black), and in the presence of 4.0 equiv. of Pd(II) (red). The His resonance region, which shows the metal coordination in the red trace, is highlighted in gray, and the intensity in that region magnified for clarity.

Additionally, NMR data were acquired for the peptide alone and in the presence of $PdCl_2(COD)$ (**Figure 71**b) by Dr. Maria Macías at the Institute for Research in Biomedicine in Barcelona. The ^{1}D NMR data for the free peptide at a concentration of 1.0 mM showed minimal chemical shift dispersion in the amide region (from 10 ppm to 8 ppm), suggesting that the peptide adopts an extended conformation. However, upon adding increasing concentrations of metal, well-dispersed amide signals appeared, along with chemical shifts characteristic of the histidine side chain, resonating between 12 and 13.5 ppm after coordination with Pd(II). These signals indicate that the peptide is no longer flexible or

[269] Roy, A., Bour, P., Keiderling, T. A. *Chirality* **21 Suppl 1**, E163-71 (2009).

extended, but adopts a defined secondary structure, consistent with the β-hairpin Wzip folding, as observed in other designed β-sheet metal-peptides.[270]

5.3.3.3. Cellular Uptake of β-hairpin Paladopeptide D14

With these results, we were ready to investigate the viability of these palladiopeptides in cells. To this end, we studied the internalization of the **D14** and **D4** peptides by fluorescence microscopy, which required the synthesis of TMR-**D14** and TMR-**D4** peptides labeled with TAMRA at the N-terminal end.

Initial experiments involved incubating the apo forms of both peptides (i.e., without metal) with cultured HeLa cells. In both cases, we observed low or negligible levels of intracellular fluorescence (**Figure 72**a), suggesting that the uncoordinated peptides have limited ability to cross the cell membrane under these conditions. Notably, preincubation of TMR-**D14** with 1 equivalent of $PdCl_2$(COD) resulted in the appearance of bright intracellular emission (**Figure 72**b). In contrast, the same treatment applied to TMR-**D4** did not internalize, as evidenced by the continued absence of fluorescence within the cells (**Figure 72**c).

Taken together, these results suggest that metal-induced stapling can significantly enhance the ability of short peptides to penetrate cell membranes. This effect appears to be closely related to the structural properties of the peptide and the formation of a stable and compact metallopeptide complex.[271]

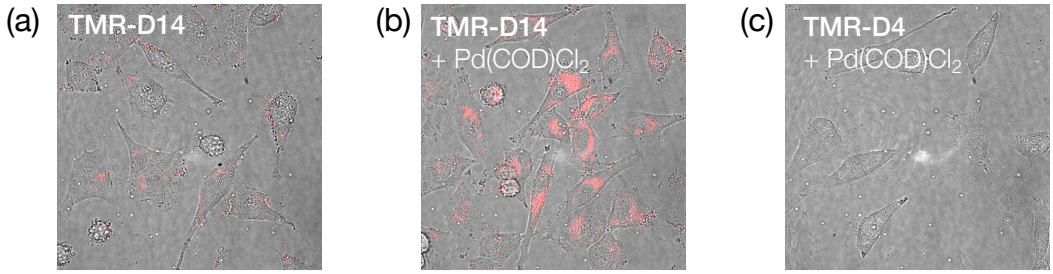


Figure 72. Cellular uptake of TMR-**D14** and TMR-**D4** in HeLa cells. Fluorescence micrographics recording the emission at 579 nm upon irradiation at 546 nm of (a) 10 μM of TMR-**D14**; (b) 10 μM of a mixture of TMR-**D14** with $PdCl_2$(COD) for 15 min; and (c) 10 μM of a mixture of TMR-**D4** with $PdCl_2$(COD) for 15 min.

[270] Pham, T. L., Kovermann, M., Thomas, F. *ACS Synth. Biol.* **11**, 1, 254–264 (2022).

[271] (a) Learte-Aymamí, S. *et al. J. Am. Chem. Soc.* **139**, 16188–16193 (2017); (b) Lättig-Tünnemann, G. *et al. Nat. Commun.* **2**, 453 (2011); (c) Li, S. *et al. Chem. Commun.* **56**, 15655–15658 (2020).

5.3.3.4. Study of the Catalytic Properties of β-hairpin Paladopeptides

The good internalization of TMR-**D14**[Pd(II)] encouraged us to study its catalytic properties in living cells. To do this, 20 μM solutions of peptides **D14** and **D4** were incubated with $PdCl_2$(COD) in a 1:1 ratio in Milli-Q water for one hour, thus forming the respective palladium complexes. These solutions were added to a mixture containing 200 μM of fluorogenic probe 1 in PBS, and the reaction was allowed to stir at 37 °C for 24 hours.

As shown in **Figure 73**a, HPLC-MS(ESI) analysis and quantification of the depropargylated product revealed that the non-coordinating peptide **D4** exhibited high catalytic activity, comparable to that of free $PdCl_2$(COD), whereas the metallopeptide **D14** reached a yield close to 40%. This result was attributed to the presence of free, non-coordinated palladium in solution.

To exclude this effect, the experiment was repeated in DMEM, where the Pd(II) salt is catalytically inactive in the absence of a ligand that stabilizes the catalytic complex. Under these conditions, both the palladium complex without peptide and the complex with **D4** showed significantly reduced activity, with yields less than 15%. In contrast, **D14** retained its catalytic activity practically intact (**Figure 73**a, marked with *), which demonstrates its ability to form a stable complex with Pd(II) and effectively protect the metal center from the environment, allowing it to maintain its catalytic function even in more demanding media.

The positive *in vitro* results of **D14**[Pd(II)] encouraged us to study its catalytic properties in living cells. To this end, HeLa cells were incubated for 1 hour at 37°C with a 50 μM solution of fluorogenic probe 1 in FBS-DMEM medium. After this incubation, the medium was removed, and the cells were washed twice with FBS-DMEM to remove any remaining uninternalized probe. The cells were then treated with a 50 μM solution of the preformed catalyst **D14**[Pd(II)] in fresh medium, incubating for 1.5 hours. After two washes with PBS, images were obtained by wide-field fluorescence microscopy. Encouragingly, an intense intracellular fluorescent signal corresponding to product **2** was observed, indicating that the depropargylation reaction was carried out efficiently within the cells (**Figure 73**b).

In contrast, when the experiment was repeated with a mixture of **D4** and $PdCl_2$(COD), or with $PdCl_2$(COD) alone, no fluorescence was detected, confirming that the **D14** peptide plays a crucial role in stabilizing the catalytic center and protecting the complex from the cellular environment.

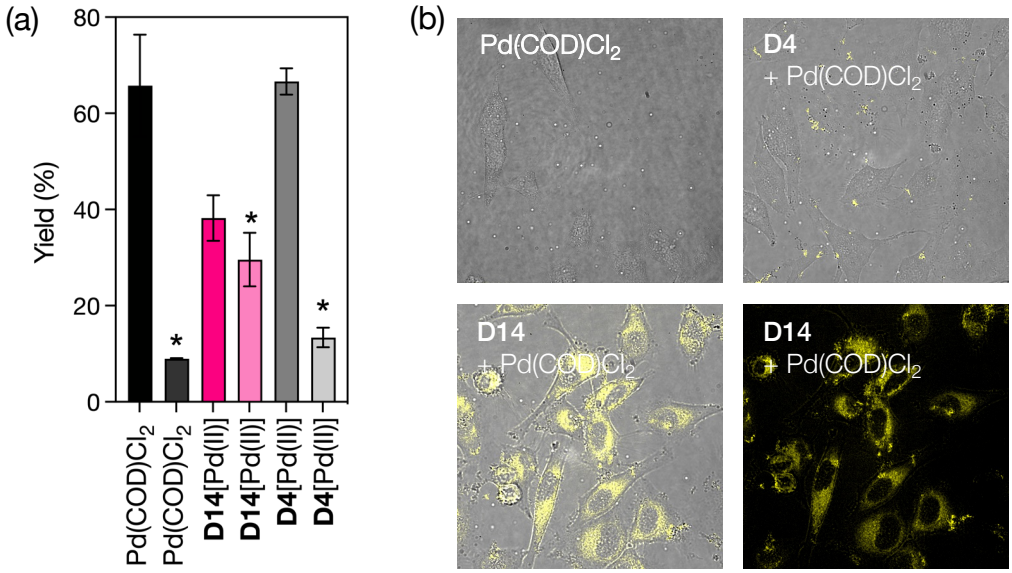


Figure 73. (a) Bar diagram representation of the yields obtained for each catalyst. Reaction conditions: mixing for 24 h a 200 μM solution of probe **1** with 20 μM mixtures of $PdCl_2$(COD) and **D14** and **D4** mixed with $PdCl_2$(COD) in a 1:1 ratio for 1 h in PBS buffer or in DMEN (*). Yields were calculated by HPLC-MS(ESI) using internal standards; (b) fluorescence microscopy of HeLa cells incubated with a 50 μM solution of **1** for 1 h washed twice with PBS and incubated for 1 h with 50 μM solution of $PdCl_2$(COD), a mixture of **D4** with $PdCl_2$(COD), and a mixture of **D14** with $PdCl_2$(COD). Micrographs show the emission intensity at 620–680 nm upon irradiation at 405 nm.

With the aim of broadening the scope of this chemistry, we set out to investigate whether it could be extended to other Pd(II) salts. To this end, we repeated the experiments using $PdCl_2$(en)], $PdCl_2$(phen)], and $PdCl_2$(bpy)], all of which are active in PBS buffer and capable of catalyzing the depropargylation of probe **1** with efficiencies comparable to $PdCl_2$(COD). However, when forming complexes with our peptides, the *in vitro* catalytic yields were generally lower than those obtained with $PdCl_2$(COD). The only exception was the complex formed between **D14** and $PdCl_2$(bpy)], which displayed a yield of approximately 30% (**Figure 74**a).

Encouraged by these results, we decided to assess whether this new complex could also retain its activity in a cellular environment. Following the same protocol described previously, we observed the formation of the fluorescent product **2** within treated cells, demonstrating that the **D14**[Pd(bpy)] complex is also catalytically active under intracellular conditions (**Figure 74**b).

These findings highlight not only the versatility of the **D14** peptide in coordinating different Pd(II) precursors but also its ability to maintain catalytic activity in cellular environments, reinforcing its potential as an adaptable platform for bioorthogonal catalysis applications.

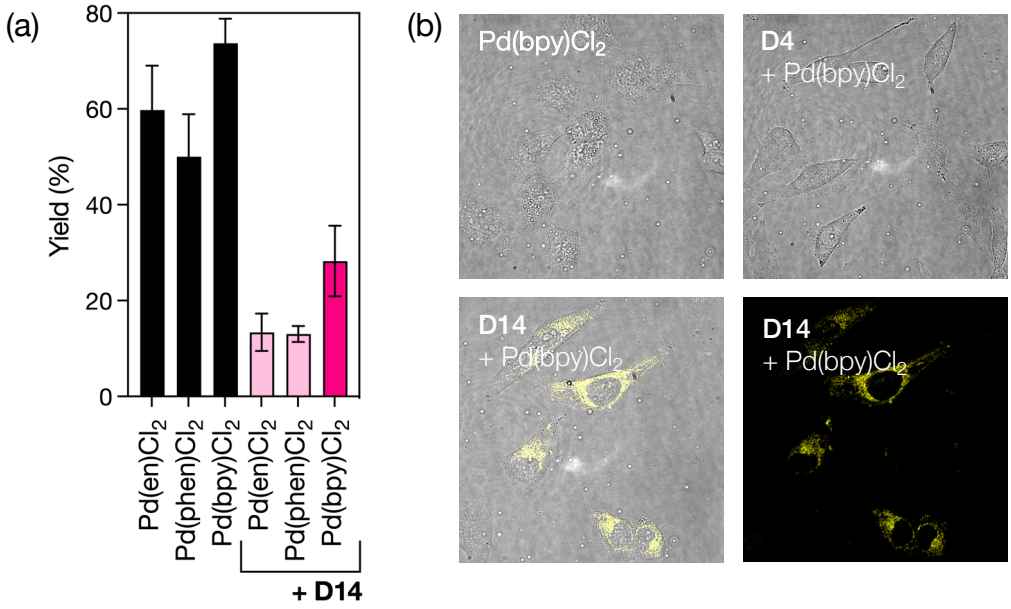


Figure 74. (a) Bar diagram representation of the yields obtained for each catalyst in PBS buffer (same reaction conditions as describe above); (b) fluorescence microscopy of HeLa cells incubated with a 50 μM solution of **1** for 1 h washed twice with PBS and incubated for 1 h with 50 μM solution of $PdCl_2$(bpy)], a mixture of **D4** with $PdCl_2$(bpy)] and **D14** with $PdCl_2$(bpy)].

After demonstrating that the peptide family His^3/His^{10} was capable of catalyzing the depropargylation reaction in living cells, we wondered whether other histidine arrangements favored in the SPOT screening, such as the His^{10}/His^{12} family, could also catalyze this reaction. To explore this possibility, we selected peptide **F14** ($A^1WR^3W\mathit{V}^5GN\mathit{T}W\mathbf{H}^{10}W\mathbf{H}$), which showed the highest catalytic activity within the His^{10}/His^{12} family according to the screening results (99%). As a negative control, we synthesized peptide **F13** ($A^1WR^3W\mathit{T}^5GN\mathit{V}W\mathbf{H}^{10}W\mathbf{H}$), a very similar sequence to **F14** with only an exchange of residues at positions Val^5 and Thr^8, which led to significantly lower catalytic activity in the SPOT assay (21%) (**Figure 75**a).

Furthermore, since positions 3, 10, and 12 were identified as the most favorable for histidine coordination, we also included a second set of peptides bearing His residues at positions 3/12. Among them, peptide **E10** ($T^1W\mathbf{H}^3WA^5GN\mathit{V}W\mathit{R}^{10}W\mathbf{H}$) was selected for its high yield (83%), while **E9** ($T^1W\mathbf{H}^3WA^5GN\mathit{R}W\mathit{V}^{10}W\mathbf{H}$), with positions 8 and 10 swapped, was used as a negative control (9%) (**Figure 75**b).

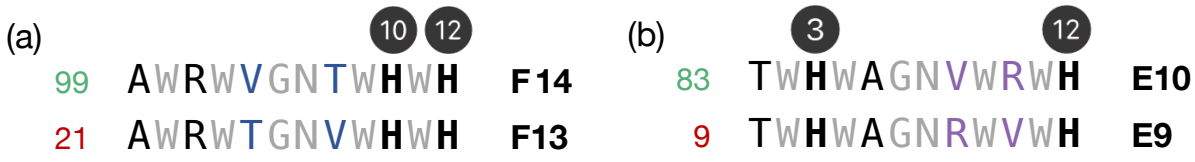


Figure 75. (a) F14 and F13 sequences; (b) E10 and E9 sequences; sequence modifications marked.

To assess *in vitro* catalytic activity, 20 μM solutions of each peptide were incubated with $PdCl_2$(COD) at a 1:1 ratio for 1 hour. Subsequently, these mixtures were added to a dilution containing 200 μM of fluorogenic probe 1 in PBS buffer. After 24 h incubation at 37 °C, HPLC-MS(ESI) analysis revealed catalytic yields close to 40 % for the **F14**[Pd(II)] and **E10**[Pd(II)] complexes, values that were maintained when the reaction was repeated in DMEM medium. In contrast, their respective negative controls, **F13** and **E9**, did not (**Figure 76**a).

Furthermore, when evaluating their catalytic activity in a cellular environment, fluorescence microscopy images revealed intense intracellular emission in cells treated with **F14**[Pd(II)] and **E10**[Pd(II)], indicating efficient formation of product **2**. In contrast, controls F13 and E9 did not generate detectable fluorescent signal (**Figure 76**b). These results demonstrate that other positions of the His residue also allow the formation of catalytically active Pd(II) complexes in living cells, validating the efficacy of SPOT screening as a tool to identify *Wzip1* sequences with catalytic activity.

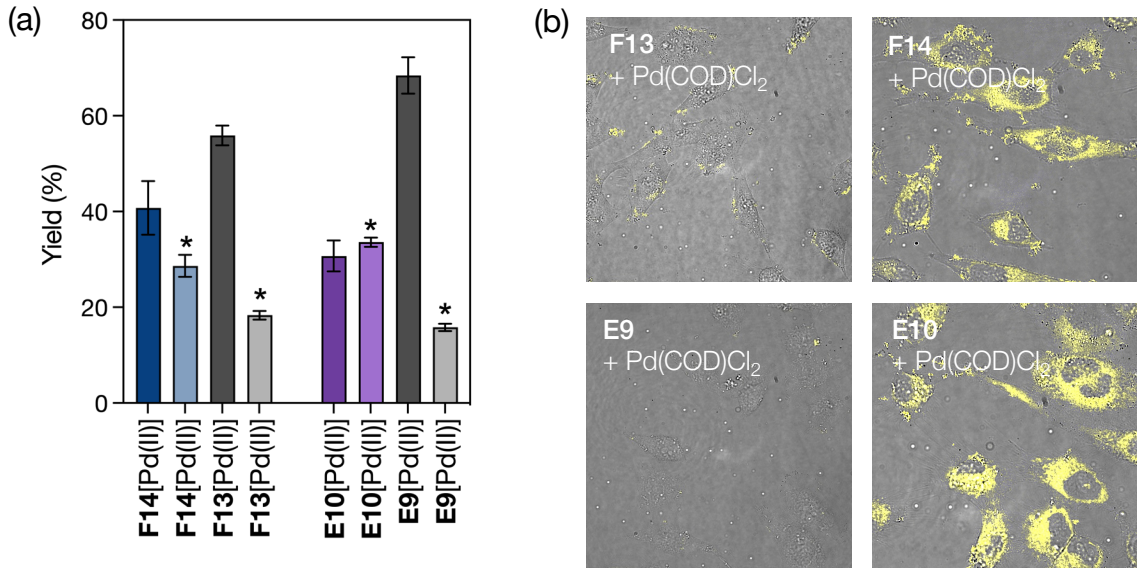


Figure 76. (a) Bar diagram representation of the yields obtained for each catalyst in PBS buffer or in DMEN (*) (same reaction conditions as describe above); (b) fluorescence microscopy of HeLa cells incubated with a 50 μM solution of 1 for 1 h, washed twice with PBS, and incubated for 1 h with a 50 μM solution of a mixture of $PdCl_2$(COD) with **F13**, **F14**, **E10**, and **E9**.

5.3.3.5. SPOT Selection of Peptide Ligands for CuAAC

Our next objective was to apply the previously described SPOT methodology for the identification of Cu(I) stabilizing ligands of Wzip1 that facilitate the CuAAC reaction in complex biological media and in living cells.

As substrates, we selected fluorogenic azides that increase their emission upon reaction with the corresponding alkynes.[272] In particular, we opted for 9-(azidomethyl)anthracene (**7**, **Figure**

[272] Golas, L. *et al. Chem. Soc. Rev.* **39**, 1233 (2010).

77a), a virtually non-fluorescent molecule that undergoes a fluorescence increase, up to 150-fold, upon conversion to triazole *via* the CuAAC reaction (**Figure 77**b and c), and which has been previously employed to monitor this transformation in intracellular contexts.[273] This increase in emission is attributed to the suppression of photoinduced electron transfer (PET) upon replacement of the azide group by a triazole ring.[274]

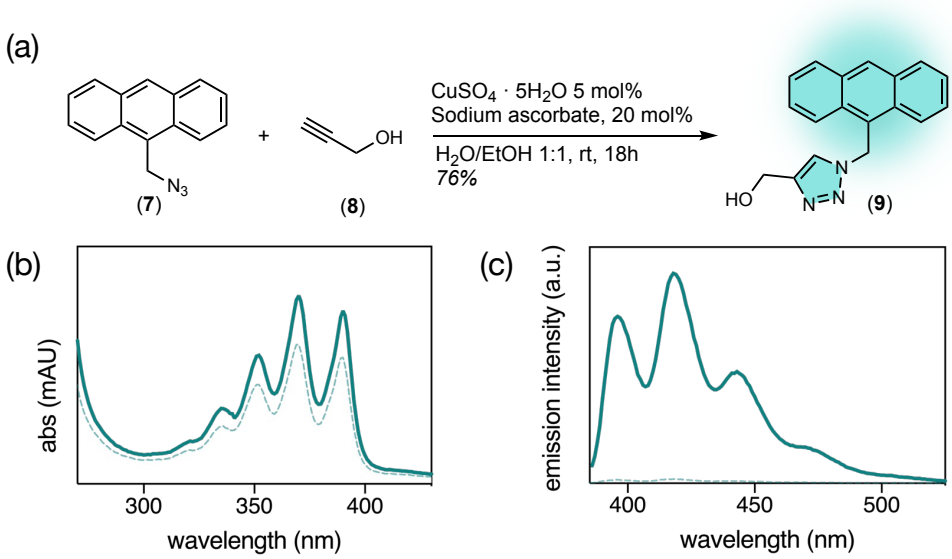


Figure 77. (a) Synthetic route to obtain the fluorogenic probe **9**; (b) absorbance spectra of the azide product **7** (dashed line) and the probe **9** (thick solid line); (c) fluorescence emission spectra of **7** (dashed line) and **9** (thick solid line). Measurements taken at 20 μM, 1:1 v/v DMSO/H_2O, λ_{exc} = 370 nm.

For the design of this new library, Wzip1-type ligands with tridentate architecture in which three His residues are located at positions 3, 10 and 12 were used to efficiently coordinate Cu(I) ions. This choice is based on results previously obtained with palladium catalytic complexes. In addition to the classical arrangement with three His residues, it was also proposed to combine two His residues with a third acidic residue, which can be either Glu or Asp, which is also able to coordinate to copper.[275] The remaining four positions of each peptide were defined by random selection among five representative amino acids (Ala, Val, Thr, Arg and Tyr) with the aim of exploring a diverse chemical environment around the catalytic center. This strategy resulted in a library of 374 distinct sequences. In addition, negative controls were included in sequences **A1**-**A10**. For example, **A8** corresponds to the sequence **D14** and **A10** to **F14**, described above. These sequences contain only two coordination residues and are therefore expected to exhibit low or no ability to stabilize the Cu(I) ion.

Once the library was designed, the previously described protocol for the identification of ligands active in CuAAC catalysis was followed. Specifically, peptides were incubated with a

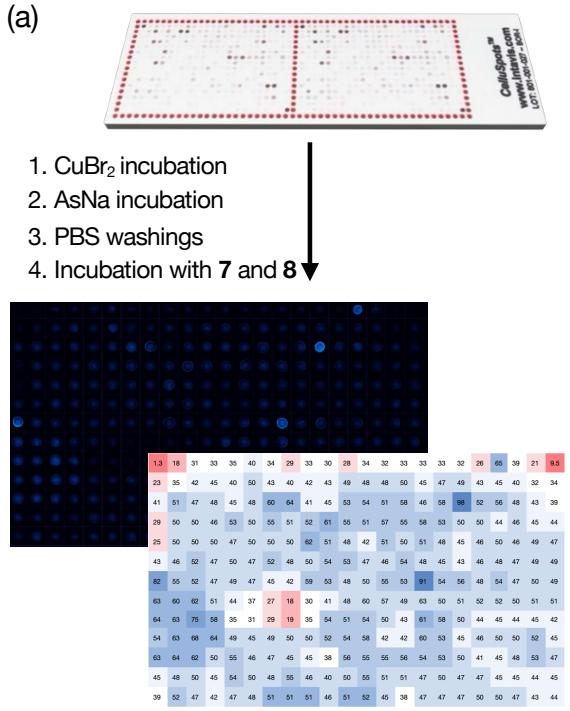
[273] Martínez-Calvo, M. *et al. ACS Catal.* **8**, 6055–6060 (2018).

[274] Xie, F. *et al. Tetrahedron* **64**, 2906 (2008).

[275] Drienovská, I. *et al. Chem. Sci.* **8**, 7228–7235 (2017).

fluorogenic azide (**7**) under standard CuAAC conditions ($CuBr_2$, sodium ascorbate, PBS), the formation of the triazole product **9** was monitored by fluorescence microscopy (**Figure 78**a).

Comparison of the peptide sequences reveals that there are clear differences in catalytic activity depending on the environment of these residues. Three of the most active peptides (**C17**, **A19** and **C8**) present the three His residues for copper coordination. However, others include an Asp residue at position 13, which could be acting as a general base favoring proton transfer during catalysis (**G15**, **G1**, **I3** and **J3**). On the contrary, the substitution of this residue by Glu, does not give rise to peptides with good catalytic activity. Furthermore, a tendency is observed in the most active peptides to contain hydrophobic or aromatic residues, such as Tyr or Val, in position 1. On the other hand, the incorporation of charged residues, such as Arg in positions 5 or 8, also seems to play a relevant role. It is also important to note that the negative controls (**A8** and **A10**), designed with only two coordinating residues, presented significantly lower activities of around 30%, confirming the importance of the tridentate arrangement for the stabilization of the catalytic center (**Figure 78**b).



(b)

		3	10	12	
1	98	Y W **H** W V G N R W **H** W **H**			**C17**
2	91	V W **H** W A G N Y W **H** W **D**			**G15**
3	81	V W **H** W R G N A W **H** W **D**			**G1**
4	75	A W **H** W V G N Y W **H** W **D**			**I3**
5	68	Y W **H** W R G N A W **H** W **D**			**J3**
6	65	R W **H** W A G N Y W **H** W **H**			**A19**
7	64	A W **H** W Y G N R W **H** W **H**			**C8**
8	64	Y W **D** W T G N R W **H** W **H**			**K2**
9	64	V W **D** W R G N Y W **H** W **H**			**H1**
10	62	R W **D** W Y G N V W **H** W **H**			**E9**
	41	A W **H** W Y G N V W **H** W **H**			**C10**
	30	A W **H** W R G N V W **H** W T			**A10**
	29	A W **H** W R G N V W **H** W T			**A8**

Figure 78. (a) Library analysis. The CelluSPOT slide was first incubated for 1 h with a 1 mM $CuBr_2$ solution, followed by incubation for 1 h with a 10 mM NaAsc solution, to reduce copper, and finally washed with PBS for 10 min to remove any copper non-coordination. It was then incubated with a 200 µM solution of **7** and 400 µM solution of **8** and dot blot emission was recorded at 0 and 4 h. Catalytic metallopeptides were identified according to their position on the array, and their relative catalytic activity was quantified by the emission intensity of each spot over time; (b) the top ten sequences ranked by catalytic efficiency (green column, normalized average of three experiments) and their corresponding positions on the CelluSPOT array, the control sequences A8 and A10, and an example of a peptide with poor catalytic efficiency are also represented in red.

From these results, peptide **C17** (Y[1]W**H**[3]WV[5]GNRW**H**[10]W**H**) was selected as a top candidate for synthesis and validation under biological conditions, due to its high activity (98%) and representative sequence. In addition, a peptide of the same family but with much lower catalytic activity, **C10** (A[1]W**H**[3]WY[5]GNVW**H**[10]W**H**) (41%), was synthesized as a negative control.

To evaluate the coordination ability of Cu(II) by the new β-hairpin-like peptide systems, equimolar amounts (500 μM) of the **C17** peptide and $CuBr_2$ were incubated in PBS buffer, and the mixtures were analyzed by mass spectrometry (MS) after 1 h of incubation. As shown in **Figure 79**a, peaks corresponding to the **C17**[Cu(II)] peptide-metal complex were detected, confirming the formation of the metallopeptide.

To monitor this interaction with higher resolution, the suppression of the typical emission band of tryptophan residues after the progressive addition of Cu(II) ions was studied (**Figure 79**b). For this purpose, solutions of the **C17** peptide were prepared at 1 μM HEPES buffer, which were titrated with increasing concentrations of Cu(II). The mixtures were incubated overnight to ensure equilibrium and subsequently their fluorescence spectra were recorded. The titration profile, measured at 420 nm, was fitted to a 1:1 binding model, resulting in a K_D of approximately 3.36 μM, indicating a moderate affinity interaction between the peptide and the metal ion.

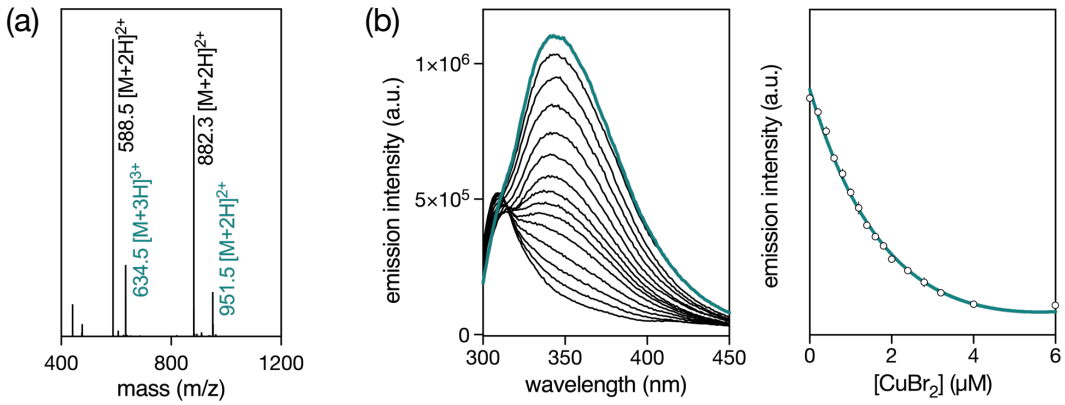


Figure 79. (a) ESI-MS spectra of the mixture of **C17** with 1.0 equiv. of $CuBr_2$ for 15 min. The peaks corresponding to the apo-peptides are labeled in black, those of the corresponding copperpeptides, are in blue; (b) Cu(II) fluorescence titration of **C17**, left fluorescence quenching of a 1 μM solution (blue thick line) and after incubation overnight with $CuBr_2$ and right tryptophan emission profile at 350 nm for a 1 μM solutions incubated overnight with in increasing concentrations of $CuBr_2$ and the best fit to 1:1 binding model (K_D = 3.358 μM). All the measurements were conducted in 25 mM HEPES buffer, 150 mM NaCl, pH = 7.5.

Following the coordination results, we set out to evaluate whether the copper complex of **C17** could be effectively internalized by cells. As in previous experiments, a fluorescently labeled derivative, TMR-**C17**, was synthesized by attaching a TAMRA moiety to the N-terminal end of the peptide. When HeLa cells were incubated with the apo form of TMR-**C17**, we observed very limited intracellular fluorescence (**Figure 80**a), indicating that the uncoordinated peptide

alone exhibits minimal cellular permeability under these conditions. However, after preincubation of TMR-**C17** with $CuBr_2$ and sodium ascorbate (NaAsc), which together generate the catalytically active Cu(I) species, a clear increase in intracellular fluorescence was observed (**Figure 80**b). As with the palladium analogues, this can be explained by the fact that metal coordination leads to a more compact structure that enhances cellular uptake.

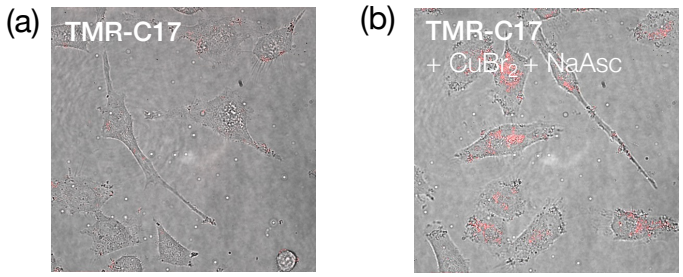


Figure 80. Cellular uptake of TMR-**C17** in HeLa cells. Fluorescence micrographics recording the emission at 579 nm upon irradiation at 546 nm of (a) 10 μM TMR-**C17**; and (b) 10 μM TMR-**C17** and $CuBr_2$ mixed for 30 mins with NaAsc.

Before evaluating the catalytic activity in cellular environments, the performance of the **C17** and **C10** complexes was evaluated in aqueous media. The reactions were carried out using 100 mol % Cu(II) by premixing $CuBr_2$ with one equivalent of the corresponding peptide in PBS at room temperature for 1 hr. Next, sodium ascorbate (NaAsc, >30 equivalents) was added to reduce copper to its +1 state (Cu(I)), and the solution was incubated for 30 min at room temperature. Finally, anthracenylated azide **7** (100 μM) and propargylic alcohol **8** (200 μM) solutions were added, and the reaction mixture was allowed to incubate for 24 h at 37 °C.

Under free copper conditions (i.e., in the absence of peptide, only $CuBr_2$ and NaAsc), the reaction proceeded with moderate yields, around 40 %. Most relevant, however, was that the **C17**[Cu(I)] complex showed catalytic activity, reaching a yield close to 25 %. In contrast, the **C10**[Cu(I)] complex did not promote the appreciable formation of product **9** (**Figure 81**a).

Next, the catalytic behavior of the peptide complexes was explored in a cellular environment. For this purpose, HeLa cells were incubated for 1 hour at 37 °C with a 100 μM solution of azide **7** and a 200 μM solution of propargyl **8** in FBS-DMEM medium. After incubation, the medium was removed, and the cells were washed twice with FBS-DMEM to remove any remaining uninternalized products. The cells were then treated with a 50 μM solution of the peptide complexes **C17**[Cu(I)] and **C10**[Cu(I)] together with the respective free copper salt in medium. For the formation of the copper complexes, they were incubated at equimolar concentrations for 1 h and then treated with excess NaAsc for 30 min. After two washes with PBS, images were obtained by wide-field fluorescence microscopy.

As shown in **Figure 81**b, the fluorescence signal in cells treated with the **C17**[Cu(I)] complex was intense and widely distributed, indicating efficient intracellular catalytic activity. In contrast, cells treated with the **C10**[Cu(I)] complex showed no detectable signal, consistent with the low yields observed in solution. The control condition with $CuBr_2$ without peptide

showed weak and localized fluorescence, highlighting the role of the **C17** peptide in stabilizing the copper catalytic center in a biological environment.

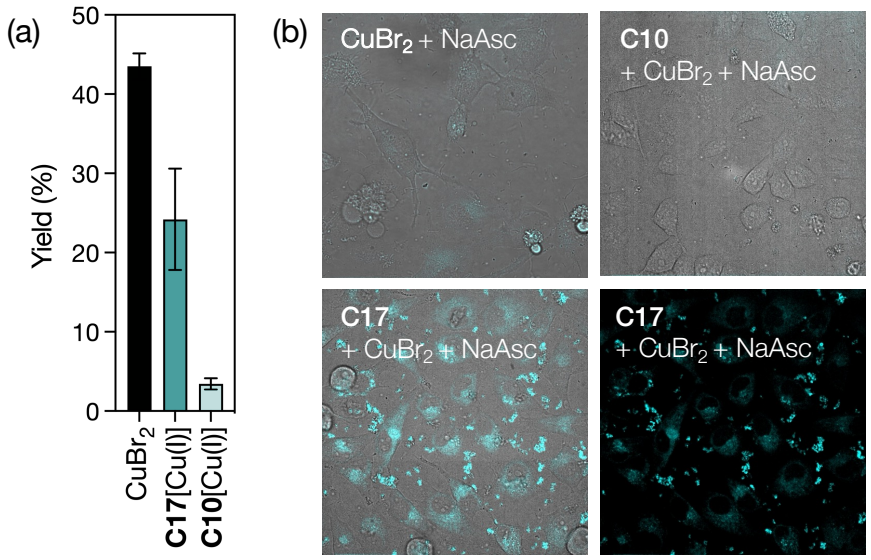


Figure 81. (a) Bar diagram representation of the yields obtained for each catalyst. Reaction conditions: mixing for 24 h a 100 μM solution of **7** and 200 μM of **8** with 50 μM mixtures of $CuBr_2$ and **C17** and **C10** mixed with $CuBr_2$ in a 1:1 ratio for 1 h in PBS buffer followed by 30 mins incubation with a NaAsc. Yields were calculated by HPLC-MS(ESI) using internal standards; (b) fluorescence microscopy of HeLa cells incubated with a 50 μM solution of **7** and 100 μM of **8** for 1 h washed twice with PBS and incubated for 1 h with 50 μM solution of $CuBr_2$, a mixture of **C10** with $CuBr_2$ and **C17** with $CuBr_2$. Micrographs show the emission intensity at 405–450 nm upon irradiation at 385 nm.

These results confirm the ability of the **C17** peptide to facilitate CuAAC reactions in complex cellular media, acting as an efficient ligand for the generation of functional bioorthogonal catalysts under physiological conditions. Over the past decade, abiotic transition metal-catalyzed reactions have successfully been introduced into living cells and organisms, allowing unprecedented biorthogonal transformations with exquisite selectivity and efficiency.[276] These new systems could serve as the next generation of powerful chemical tools for modulating biology, with applications in biomedicine, such as for disrupting cell-cell communication, targeted release of drugs, or interfacing synthetic catalysts with cellular metabolism.[277]

[276] Rebelein, J. G., Ward, T. R. *Curr. Opin. Biotechnol.* **53**, 106–114 (2018).

[277] Patra, M., Gasser, G. *Chembiochem* **13**,1232–1252 (2012).

5.3.4. Conclusion

In conclusion, we have demonstrated for the first time the potential of SPOT libraries for the systematic identification of catalytically active metallopeptides. We identified a β-hairpin palladopeptide, **D14**[Pd(II)], capable of catalyzing a depropargylation reaction, as well as a second complex, **C17**[Cu(I)], that catalyzes an azide-alkyne cycloaddition (CuAAC), both in aqueous media and in live mammalian cells.

Although the catalytic efficiency of these systems still falls short of that observed in natural metalloenzymes, the resulting complexes can be considered miniature protometalloenzymes, capable of functioning in cellular environments due to their high stability and favorable internalization properties.

These findings lay the groundwork for the development of spatially addressable peptide platforms aimed at discovering efficient bioorthogonal catalysts for the selective release of active molecules, using Trp-zipper β-hairpins or even alternative small peptide scaffolds.

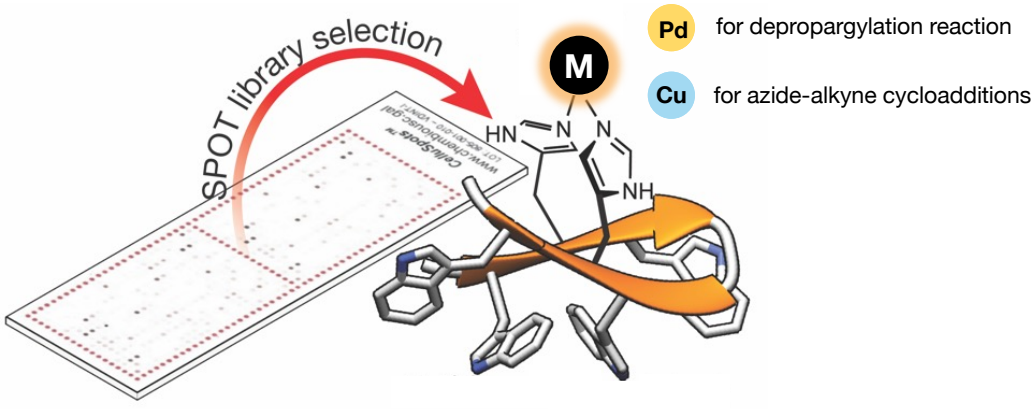


Figure 82. Illustration of the application of the SPOT methodology for screening peptide sequences with catalytic activity.

6. CONCLUSIONS

This thesis robustly demonstrates that peptides can be rationally designed to acquire unique photophysical and catalytic properties even in the absence of traditional aromatic elements or cofactors. Our main results are:

- The discovery and systematic characterization of NAF emission from monomeric peptides in solution, establishing its dependence on helical folding and non-covalent interactions.
- The successful application of a NAF fluorophore into an energy transfer process for lanthanide sensitization, extending its applicability in optoelectronics.
- The improvement of the stability and functionality of these systems by cyclization, which allowed to widen the emission window and its potential application in bioimaging.
- The creation of artificial metalloproteins based on β-sheets, capable of stabilizing Pd(II) complexes and protecting them from their rapid deactivation in the cellular environment, allowing bioorthogonal catalysis to be carried out efficiently.
- The validation of combinatorial peptide libraries as efficient tools for the identification of bioactive catalysts capable of operating in complex cellular environments.

Overall, this work opens new possibilities for the development of functional peptide materials, with applications in the design of fluorescent biosensors, biocatalysis and tools for the biomedical field, constituting a relevant contribution to the advancement of modern biological chemistry.

7. METHODOLOGY

7.1. General Information

7.1.1. General protocols

Chemical were purchased from *Sigma Aldrich*, *Alfa Aesar*, and *Iris Biotech* and used without further purification. Reaction mixtures were stirred using Teflon-coated magnetic stir bars. The abbreviation “rt” refers to reactions carried out at room temperature (approximately at 23 °C). The temperature was maintained using Thermowatch-controlled heating blocks. Thin-layer chromatography (TLC) was performed on silica gel plates (*Merck 60 silica gel F254*) and the reaction components were visualized by observation under UV light and/or by treating the plates with $KMnO_4$ or ninhydrine followed by heating. Flash chromatography was carried out on silica gel (*Merck Geduran Su 60*, 40 - 63 μM silica gel, normal phase). Concentration refers to the removal of volatile solvents via distillation using a rotary evaporator *Büchi R-210* equipped with a *thermostated bath B-491*, a vacuum regulator *V-850,* followed by residual solvent removal under high vacuum.

7.1.2. General Peptide Synthesis Procedures

All peptide synthesis reagents and amino acid derivatives were purchased from *Sigma Aldrich* or *Iris Biotech GmbH*; amino acids were purchased as Fmoc-protected derivatives with the standard side-chain protection scheme: Fmoc-Ala-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Asp(OtBu)-OH, Fmoc-Glu(OtBu)-OH, Fmoc-Gly-OH, Fmoc-His(Trt)-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Orn(Boc)-OH, Fmoc-Pro-OH, Fmoc-Ser(tBu)-OH, Fmoc-Thr(tBu)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Tyr(tBu)-OH and Fmoc-Val-OH.

Solid Phase Peptide Synthesis Protocols. Peptides were synthesized on a 0.05 mmol scale according to standard microwave-assisted Fmoc/tBu solid-phase peptide synthesis protocols on a *Liberty Lite* peptide synthesizer (*CEM Corporation*) using a *Rink Amide ChemMatrix* resin (100-200 mesh) with a loading of 0.62 mmol/g for all peptides. *Rink Amide ProTide* resin 0.18 mmol/g was used for peptides with more than 40 amino acids in their chain. Each residue was coupled in a 5-fold excess using DIC as activator, oxime as base, and DMF as solvent according to CEM recommended protocols. Couplings were performed at 90 °C for 4 minutes. Deprotection of the temporal Fmoc protecting group was performed by treating the resin with 20% piperidine in DMF for 1 min at 75 °C. For the cleavage/deprotection step the resin-bound peptides were treated for 2 h with 900 μL TFA, 50 μL DCM, 25 μL H_2O, and 25 μL TIS (1 mL of cocktail/40 mg resin). The cleavage mixture was filtered from the resin and the crude peptide was precipitated with ice-cold diethyl ether for 10-30 min and then centrifuged.

Semi-preparative HPLC Purification. The residue was redissolved in water for purification using reversed-phase HPLC (RP-HPLC) on a *Waters 1500 series Liquid Chromatograph* preparative RP-HPLC with a single wavelength UV/Vis detector at 222 nm, using a reverse-

phase column *Sunfire Prep C_{18} OBD* (5 µm, 19 × 150 mm) with mixtures of 0.1% aqueous TFA and 0.1% TFA in acetonitrile as mobile phase. The combined fractions with the peptides were freeze-dried.

UHPLC-MS Characterization. The isolated peptides were analyzed by analytical UHPLC-MS with an *Agilent 1200* series LC/MS using a *SB C_{18}* (1.8 µm, 2.1 × 50 mm) analytical column from *Phenomenex*. Standard conditions for analytical UHPLC consisted of a linear gradient from 5 to 95% of solvent B for 15 min at a flow rate of 0.35 mL/min (A: water with 0.1% TFA, B: acetonitrile with 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. The mass spectrometry data of the different synthesized compounds have been written using the following format: HPLC-MS (ESI) (m/z, t_R): calculated for $C_xH_yN_zO_k$ $[M+H]^+$ =; found $[M+H]^+$ =, $[M+2H]^{2+}$ =, …

7.1.3. Experimental Techniques

UV/Vis Measurements. Measurements of UV/Vis 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: 250-750 nm range, scan speed of 200 nm/min, resolution of 1.0 nm.

Fluorescence Measurements. All fluorescence measurements were performed with an *Edinburgh Instruments FS5* spectrofluorometer coupled with module Standard Cuvette Holder SC-05, using a high precision cell Hellma (114F-QS) semi-micro cuvette.

Circular Dichroism Measurements. CD experiments were made with a *Jasco-715* coupled with a thermostat *Nestlab RTE-111*. Acquisition range: 350-195 nm; band width: 2.0 nm; resolution: 0.2 nm; accumulation: 3 scans; sensitivity 100 mdeg; response time: 0.25 s, speed: 100 nm/min. Measurements were made in a 2 mm cell at 20 °C. The ellipticity (θ /mdeg) was converted to mean molar ellipticity (Φ /mdeg $M^{-1}cm^{-1}$) using the equation reported by Kelly *et al.*[278]

$$\Phi = \frac{\theta \cdot M}{10 \cdot n \cdot l \cdot C}$$

θ is the ellipticity in mdeg, M is the molecular mass of the sample, n is the number of amide bonds in the structural formula, l is the optical path length in cm, and C is the concentration in mg/mL.

[278] Kelly, S., Price, N. *CPPS* **1**, 349–384 (2000).

^{1}H and ^{13}C NMR Experiments. NMR experiments were carried out using a *Varian Mercury* 300 MHz NMR spectrometer. Chemical shift values are reported in ppm with the solvent resonance as the internal standard ($CDCl_3$: δ 7.26 for ^{1}H, δ 77.2 for ^{13}C; DMSO: δ 2.50 for ^{1}H, δ 39.5 for ^{13}C). Coupling constants *J* are given in Hertz (Hz). Multiplicities are reported as follows: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet or as a combination of them. The chemical shifts for protons (δ) are reported in parts per million downfield from tetramethyl silane and are referenced to residual protium in the NMR solvent ($CHCl_3$ δ = 7.26). Chemical shifts for carbon are reported in parts per million downfield from tetramethyl silane and are referenced to the carbon resonances of the solvent ($CDCl_3$ δ = 77.0). NMR spectra were analyzed using *MestreNova* NMR data processing software (*www.mestrelab.com*).

7.1.4. Cell assays

All steps were performed on a sterile clean bench *Tesltar* AV-100 at room temperature. Solutions stored in a fridge were warmed beforehand in a water bath (37 °C). Unless otherwise specified, all incubations were performed in DMEM containing 5% of fetal bovine serum (FBS-DMEM).

Cell Culture. All cell lines were cultured in DMEM (Dulbecco´s modified Eagle´s medium), 5 mM glutamine, penicillin (100 units/mL) and streptomycin (100 units/mL) (all from Invitrogen). Proliferating cultures were maintained in a 5% CO_2 humidified incubator at 37 °C. For all the experiments, cells were seeded in the corresponding well at the indicated concentration two days before treatment.

Fluorescence microscopy. All images were obtained with an *Andor Zyla* mounted on a *Nikon TiE*. Confocal images were acquired in an *Andor Dragonfly High Speed Confocal Platform*. Images were further processed with *Image J* or *NIS software* (*Nikon*). Microscopy settings: The filter sets for the observation of the fluorescence of the products were as follows: HBTP: Widefield: LED λ excitation: 385 nm. Filter cube: BP 375/28x nm, LP 515lp nm and DM 415 nm; TMR: Widefield: LED λ excitation: 550 nm. Filter cube TRITC-B-000 (Semrock): BP 543/22 nm, LP 593/40 nm and DM 562 nm. Confocal: Laser excitation: 561 nm. LP 620/60 and DM 567 nm.

7.2. Anomalous Emission from Single α-Helical Peptides in Solution

7.2.1. Synthesis and Characterization of Peptides

Table 2. Name and sequence of the peptides used in this study. §O = ornithine. All peptides are N-terminal free amine, C-terminal amide.

Peptide	Sequence
$(E_4K_4)_4$	EEEEKKKKEE EEKKKKEEEE KKKKEEEEKK KK
$(E_4K_4)_9$	EEEEKKKKEE EEKKKKEEEE KKKKEEEEKK KKEEEEKKKK EEEEKKKKEE EEKKKKEEEE KKKKEEEEKK KK
$(E_2K_2)_8$	EEKKEEKKEE KKEEKKEEKK EEKKEEKKEE KK
$(E_4O_4)_4$§	EEEE0000EE EE0000EEEE 0000EEEE00 00
$(E_4R_4)_4$	EEEERRRREE EERRRREEEE RRRREEEERR RR
$(E_3YK_3Y)_4$	EEEYKKKYEE EYKKKYEEEY KKKYEEEYKK KY
$(E_3YR_3Y)_4$	EEEYRRRYEE EYRRRYEEEY RRRYEEEYRR RY
Y3	EEYEKYKKYE EYKKYKEYEE YKKYEEYEKY KK

Synthesis of labeled AF-$(E_4K_4)_4$ and AF-$(E_2K_2)_4$. Once the peptides were fully assembled and still attached to the solid support, we coupled Fmoc-5-amino-3-oxapentanoic acid (Fmoc-O_1Pen-OH) as spacer using standard solid-phase methods (activation with 1.0 equiv. of HATU in 2 mL DMF with 3 mL 0.2 M DIEA/DMF). The terminal Fmoc protecting group was removed (20% piperidine/DMF) and a mixture of *Alexa Fluor™ 430 NHS Ester* (AF, 3.0 equiv., 0.15 mmol), 3 equiv. of HATU and 5 equiv. of DIEA 0.2 M in DMF was added onto the resin and shaken for 2h.

Synthesis of labeled DOTA-$(E_4K_4)_4$ and DOTA-$(E_2K_2)_4$. Once the peptides were fully assembled and still attached to the solid support, we coupled Fmoc-5-amino-3-oxapentanoic acid (Fmoc-O_1Pen-OH) as spacer using standard solid-phase methods (activation with 1.0 equiv. of HATU in 2 mL DMF with 3 mL 0.2 M DIEA/DMF). The terminal Fmoc protecting group was removed (20% piperidine/DMF) and a mixture of DOTA-*tris*-(tert-butyl ester) (5.0 equiv., 0.25 mmol), 5 equiv. of HATU and 5 equiv. of DIEA 0.2 M in DMF was added onto the resin and shaken for 2 h.

On-resin cyclization. To enable site-specific cyclization, we designed a peptide sequence incorporating an amino acid with a side chain that can be selectively deprotected (Alloc). For the deprotection step, we use $Pd(PPh_3)_4$/PhSi in DCM for 15 min. This method allows precise control over the cyclization site, utilizing PyBOP/HOBt/DIEA in DMF for 2 h.

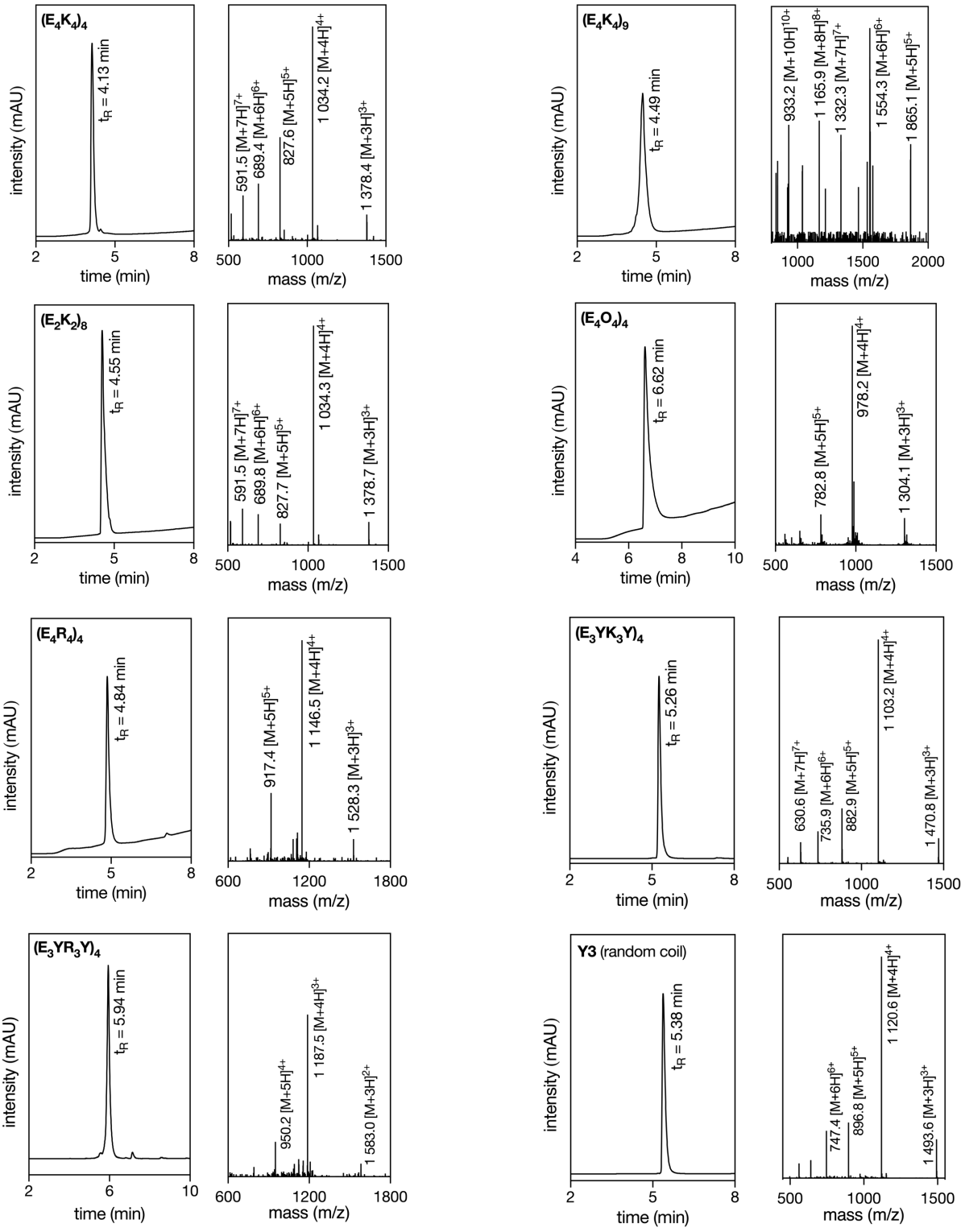


Figure 83. HPLC traces of the purified single alpha-helix peptides and their corresponding ESI-MS spectra showing in each case peaks matching the expected masses.

$(E_4K_4)_4$ HPLC-MS (ESI) (m/z, t_R = 4.13 min): calculated for $C_{176}H_{307}N_{49}O_{64}$ $[M+H]^+$ = 4 133.7; found $[M+3H]^{3+}$ = 1 378.4, $[M+4H]^{4+}$ = 1 034.2, $[M+5H]^{5+}$ = 827.6, $[M+6H]^{6+}$ = 689.4, $[M+7H]^{7+}$ = 591.5. **$(E_4K_4)_9$** HPLC-MS (ESI) (m/z, t_R = 4.49 min): calculated for $C_{396}H_{687}N_{109}O_{144}$ $[M+H]^+$ = 9 279.5; found $[M+5H]^{5+}$ = 1 869.1, $[M+6H]^{6+}$ = 1 554.3, $[M+7H]^{7+}$ = 1 332.3, $[M+8H]^{8+}$ = 1 165.9 $[M+10H]^{10+}$ = 933.2. **$(E_2K_2)_8$** HPLC-MS (ESI) (m/z, t_R = 4.13 min): calculated for $C_{176}H_{307}N_{49}O_{64}$ $[M+H]^+$ = 4 133.7; found $[M+3H]^{3+}$ = 1 378.7, $[M+4H]^{4+}$ = 1 034.3, $[M+5H]^{5+}$ = 827.7, $[M+6H]^{6+}$ = 689.8, $[M+7H]^{7+}$ = 591.5. **$(E_4O_4)_4$** HPLC-MS (ESI) (m/z, t_R = 6.62 min): calculated for $C_{160}H_{275}N_{49}O_{64}$ $[M+H]^+$ = 3 909.2; found $[M+3H]^{3+}$ = 1 304.1, $[M+4H]^{4+}$ = 978.2, $[M+5H]^{5+}$ = 782.8. **$(E_4R_4)_4$** HPLC-MS (ESI) (m/z, t_R = 4.84 min): calculated for $C_{176}H_{307}N_{81}O_{64}$ $[M+H]^+$ = 4 581.9; found $[M+3H]^{3+}$ = 1 528.3, $[M+4H]^{4+}$ = 1 146.5, $[M+5H]^{5+}$ = 917.4. **$(E_3YK_3Y)_4$** HPLC-MS (ESI) (m/z, t_R = 5.26 min): calculated for $C_{204}H_{303}N_{45}O_{64}$ $[M+H]^+$ = 4 409.9; found $[M+3H]^{3+}$ = 1 470.8, $[M+4H]^{4+}$ = 1 103.2, $[M+5H]^{5+}$ = 882.9, $[M+6H]^{6+}$ = 735.9. **$(E_3YR_3Y)_4$** HPLC-MS (ESI) (m/z, t_R = 5.94 min): calculated for $C_{204}H_{303}N_{69}O_{64}$ $[M+H]^+$ = 4 746.1; found $[M+3H]^{3+}$ = 1 583.0, $[M+4H]^{4+}$ = 1 187.5, $[M+5H]^{5+}$ = 950.2. **Y3** HPLC-MS (ESI) (m/z, t_R = 5.38 min): calculated for $C_{211}H_{302}N_{44}O_{64}$ $[M+H]^+$ = 4 479.0; found $[M+3H]^{3+}$ = 1 493.6, $[M+4H]^{4+}$ = 1 120.6, $[M+5H]^{5+}$ = 896.8, $[M+6H]^{6+}$ = 747.4.

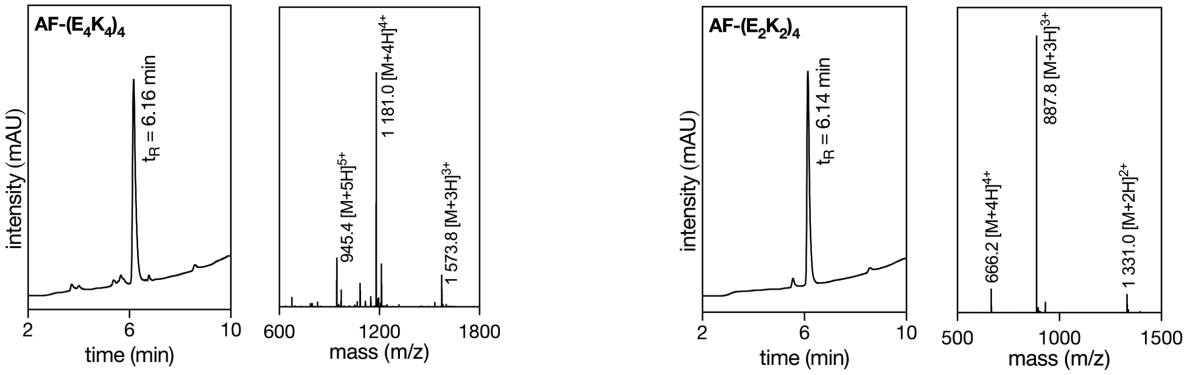


Figure 84. HPLC traces of the purified AF labeled peptides and their corresponding ESI-MS spectra showing in each case peaks matching the expected masses.

AF-$(E_4K_4)_4$ HPLC-MS (ESI) (m/z, t_R = 6.16 min): calculated for $C_{202}H_{336}F_3N_{51}O_{72}S$ $[M+H]^+$ = 4 720.2; found $[M+3H]^{3+}$ = 1 573.8, $[M+4H]^{4+}$ = 1 181.0, $[M+5H]^{5+}$ = 945.4. **AF-$(E_2K_2)_4$** HPLC-MS (ESI) (m/z, t_R = 6.14 min): calculated for $C_{114}H_{184}F_3N_{27}O_{40}S$ $[M+H]^+$ = 2 660.3; found $[M+2H]^{2+}$ = 1 331.0, $[M+3H]^{3+}$ = 887.8, $[M+4H]^{4+}$ = 666.2.

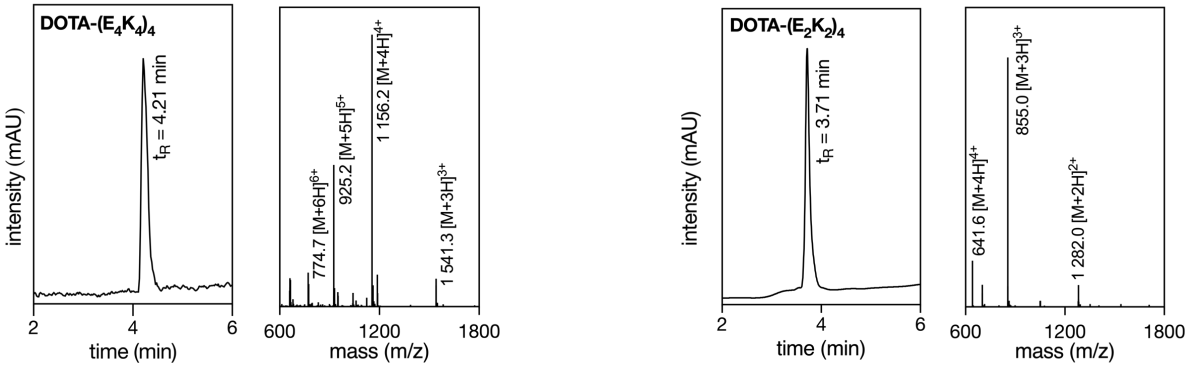
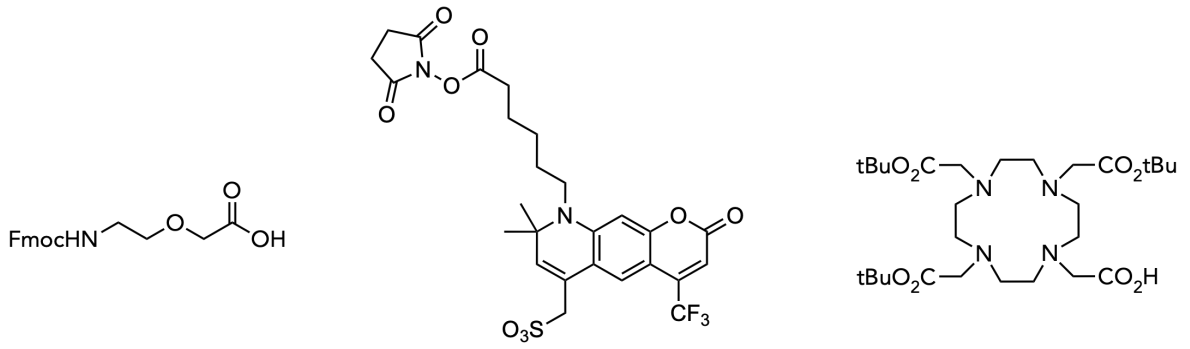


Figure 85. HPLC traces of the purified DOTA labeled peptides and their corresponding ESI-MS spectra showing in each case peaks matching the expected masses.

DOTA-(E_4K_4)$_4$ HPLC-MS (ESI) (m/z, t_R = 4.21 min): calculated for $C_{196}H_{340}N_{54}O_{73}$ $[M+H]^+$ = 4 621.2; found $[M+3H]^{3+}$ = 1 541.3, $[M+4H]^{4+}$ = 1 156.2, $[M+5H]^{5+}$ = 925.2, $[M+6H]^{6+}$ = 774.7. **DOTA-(E_2K_2)$_4$** HPLC-MS (ESI) (m/z, t_R = 3.71 min): calculated for $C_{108}H_{188}N_{30}O_{41}$ $[M+H]^+$ = 2 562.8; found $[M+2H]^{2+}$ = 1 282.0, $[M+3H]^{3+}$ = 855.0, $[M+4H]^{4+}$ = 641.6.



Fmoc-O_1Pen-OH **Alexa FluorTM 430 (AF) NHS Ester** **DOTA-tris(tert-butyl ester)**

Figure 86. Structures of Fmoc-5-amino-3-oxapentanoic acid (Fmoc-O_1Pen-OH), *Alexa Fluor™ 430 NHS Ester* (AF) and DOTA-tris(tert-butyl ester).

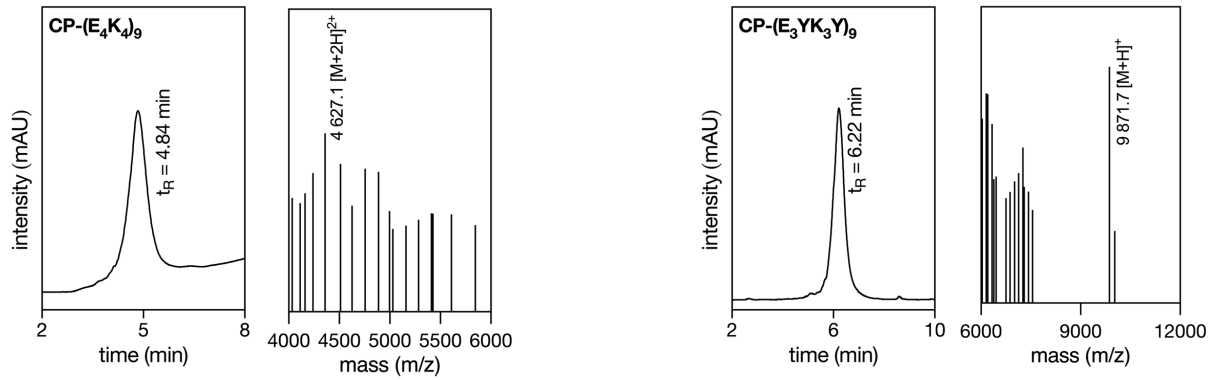


Figure 87. HPLC traces of the purified cyclopeptides and MALDI spectra showing in each case the peak matching the expected mass.

CP-(E_4K_4)$_9$ MALDI: calculated for $C_{396}H_{684}N_{108}O_{144}$ $[M+H]^+$ = 9 256.9; found $[M+2H]^{2+}$ = 4 627.1. **CP-(E_3YK_3Y)$_4$** MALDI: calculated for $C_{459}H_{675}N_{99}O_{144}$ $[M+H]^+$ = 9 877.8; found $[M+H]^+$ = 9 871.7.

Lanthanide coordinated peptides. DOTA labeled peptides (50 μM) were mixed with equimolar amounts of lanthanide ions $Tb(CF_3SO_3)_3$, $EuCl_3$, $SmCl_3$ and $DyCl_3$ in 25 mM HEPES, NaCl 150 mM pH 7.5 and the resulting mixtures were analyzed after 15 min. by Electrospray Ionization Mass Spectrometry (ESI/MS) with an *Agilent 6120 Quadrupole* LC/MS model in positive scan mode using direct injection of the purified peptide solution into the MS detector. The peaks corresponding to the apo-peptides are shown in black, and those corresponding to the complexes, that show the mass of the peptide and Ln(III), are in different colors.

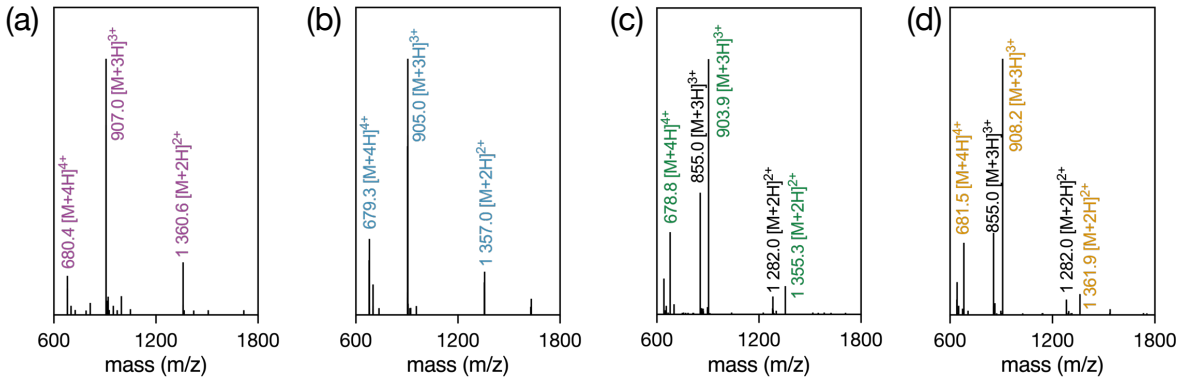


Figure 88. ESI-MS spectra of the mixture of DOTA-$(E_2K_2)_4$ (50 μM) with 1.0 equiv. of (a) $Tb(CF_3SO_3)_3$, (b) $EuCl_3$ (c) $SmCl_3$ and (d) $DyCl_3$ in 25 mM HEPES, 150 mM NaCl pH 7.5 after 15 min incubation. The peaks corresponding to the apo-peptides are labeled in black, those of the corresponding coordinated peptides are colored.

7.2.2. Determination of the Concentration by NMR

External Standard for Quantitation by Proton NMR. Valine/D_2O solutions of 10, 20 and 40 mM, respectively, were measured and quantified by the ^{1}H-NMR using 50 mM MA/D_2O solution as an external standard. The relationship between valine concentrations and MA signal was calculated by linear regression with a coefficient of 0.9992 and the slope was 0.2869 (y = 0.2869 x + 0.152).

Peptides Concentration Calculation. We must identify an isolated and assignable signal in the ^{1}H-NMR spectrum of the peptide. So, a ^{2}D-NMR COSY (COrrelation SpectroscopY) experiment was measured. The Cα signal of the peptide chain for both glutamic acid (E) and lysine (K) residues will be the same, as the difference in chemical shift is practically zero. Furthermore, the experiment's selectivity for detecting couplings between hydrogens separated by two and three bonds establishes a correlation between the Cα and other signals manifest in the NMR spectrum. Taking all this information, it can be reasonably concluded that the Cα of our peptide corresponds to the signal observed at 4.15 ppm.

We know that Cα must be equal to the number of residues in the peptide sequence. Consequently, we can derive the intensity per proton for each chain. Upon acquiring this information, we substitute it into the previously derived equation correlating proton intensity to concentration. Through this process, we can quantitatively ascertain the concentration of the stock solutions for our peptides, despite lacking any residues exhibiting UV absorption.

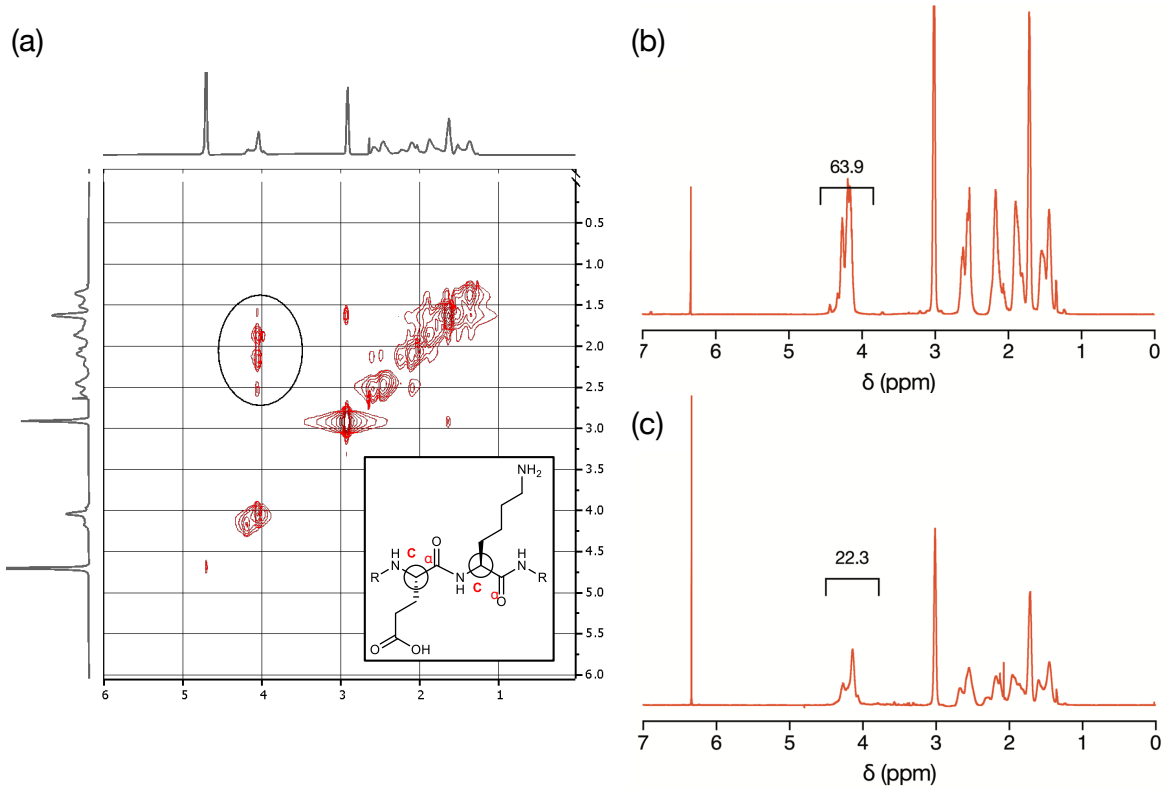


Figure 89. (a) 2D-NMR COSY spectrum of $(E_4K_4)_9$ in D_2O. ^{1}H-NMR spectrum of (b) $(E_4K_4)_9$ and (c) $(E_4K_4)_4$ with maleic acid as internal standard in D_2O.

7.2.3. Experimental Protocols

Excitation-Emission Matrices (EEMs). EEMs were made with a peptide concentration of 50 μM at 20 °C. The settings used were the following: excitation at different wavelengths (λ_{exc}), from 250-600 nm, recording every 5 nm; acquisition range: 270-750 nm; step: 5.0 nm, dwell: 0.1 s; repeats 1; the excitation/emission bandwidth was 4.0/4.5 nm.

To represent the EEM in Raman units (R.U.), first the spectrum of the HEPES buffer is measured at 350 nm, since this wavelength is already used for signal-to-noise ratio determinations and is within the range of many fluorescence spectrophotometers. The Raman peak integral at a fixed wavelength (350 nm) is used to normalize the entire spectrum or EEM.

Quantum Yield and Lifetime Determination. Quantum yield determination was performed using a *Jasco FP8300* fluorometer and a *Cary 4000* UV/Vis spectrophotometer, using quinine sulphate (in aqueous H_2SO_4 0.1 M) as reference, as described elsewhere (*3*). Time-resolved fluorimetry for lifetime determination was performed on a *PicoQuant FluoTime 200*, using a 320-nm pulsed LED (PLS-320) as the excitation source for $(E_4K_4)_9$ and $(E_4K_4)_4$. The excitation sources were controlled by a PDL-800 driver (*PicoQuant*) working at 20 MHz repetition rate. Three different decay traces were collected for each sample, at 390, 400, and 410 nm. The decay traces were globally fitted to a tri-exponential decay function, keeping the decay times

at global parameters. **Table S2** shows the results of the fittings, averaged at the three emission wavelengths. **Figure S8** shows representative decay traces from each peptide. The intensity-weighted average lifetime values were reported for each peptide in the main text.

Table 3. Results from global fittings of luminescence decay traces of $(E_4K_4)_4$ and $(E_4K_4)_9$.

	τ /ns[a]	Fractional amplitude (%)[b]	Fractional intensity (%)[c]	$\langle\tau\rangle_{int}$ /ns[d]
$(E_4K_4)_4$				
Comp. 1	0.68 ± 0.02	58.5 ± 0.8	19.8 ± 0.1	
Comp. 2	2.35 ± 0.04	29.3 ± 1.5	34.3 ± 2.0	4.42 ± 0.12
Comp. 3	7.57 ± 0.08	12.2 ± 0.6	45.9 ± 1.9	
$(E_4K_4)_9$				
Comp. 1	0.62 ± 0.02	50.6 ± 0.2	12.7 ± 0.1	
Comp. 2	2.55 ± 0.04	31.9 ± 0.2	32.9 ± 0.3	5.08 ± 0.02
Comp. 3	7.66 ± 0.05	17.5 ± 0.2	54.4 ± 0.4	

[a] The associated error values are the standard deviations from the fittings. [b] The values are averaged from results of the three different wavelengths, and the associated error values are the corresponding standard deviations. [c] The intensity values estimated as $A_i \cdot \tau_i$ are averaged from results of the three different wavelengths, and the associated error values are the corresponding standard deviations. [d] Intensity-weighted average lifetime.

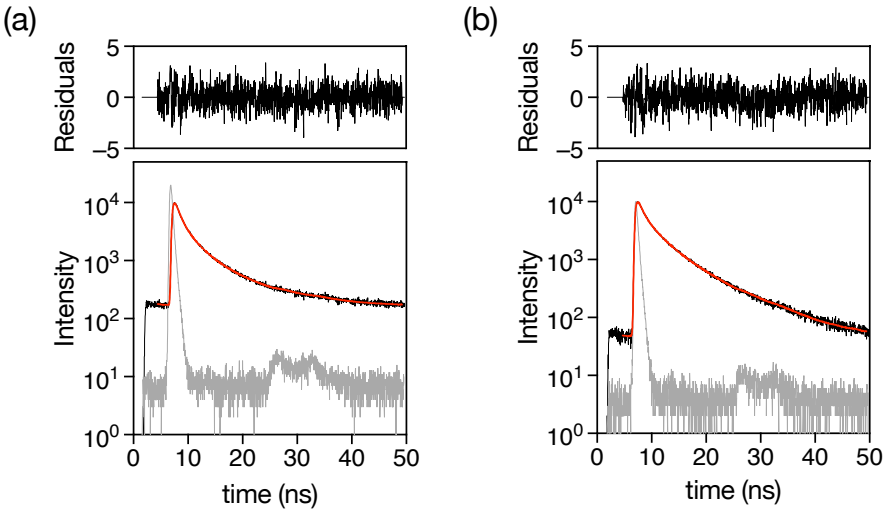


Figure 90. Fluorescence decay traces (light blue) of (a) **(**$(E_4K_4)_4$ and (b) $(E_4K_4)_9$ peptides. The convoluted fitted function (red line) and the impulse response function (gray) are also shown in the plot. The top panels show the weighted residuals from the tri-exponential fittings.

Thermal Melting Assay. The emission of the peptide $(E_4K_4)_9$ was measured at different temperatures between 20-90 °C in 10 mM HEPES pH 7.5, 100 mM NaCl with the following settings: excitation wavelength: 335/250 nm; acquisition range: 350-750/270-750 nm; step: 1.0 nm, dwell: 0.1 s; repeats 1; excitation/emission band width: 4.5/4.0.

Effect of the concentration in emission intensity. We measured the fluorescence spectra of the peptides at different concentrations with the following parameters: excitation wavelength: 335 nm; acquisition range: 350-800 nm; step: 1.0 nm, dwell: 0.25 s; repeats 1; excitation/emission band width: 2.5/2.0.

Peptide Digestion with Proteinase K. A 50 μM solution of $(E_4K_4)_9$ was incubated at 37 ºC with Proteinase K enzyme (100 μg/mL), that cleaves peptide bonds adjacent to the carboxylic group of amino acids and is useful for the general digestion of proteins in biological samples.[279] The mixture was excited at 325 nm and the fluorescence maximum was measured at 0, 5, 10 and 20 h in 20 mM TRIS pH 8.0, 50 mM NaCl.[280] The settings used were the following: excitation wavelength: 335 nm; emission wavelength: 420 nm; excitation/emission band width: 4.5/4.0.

To confirm the evolution of the reaction, it was followed by HPLC-MS. The chromatogram of the peptide, enzyme and reaction were also measured, using linear gradient from 5 to 95% of solvent B (A: water with 0.1% TFA, B: acetonitrile with 0.1% TFA) over 15 mins at a flow rate of 0.35 mL/min.

Time gated experiments. Time gated experiments were run in 25mM HEPES, 150 mM NaCl, pH 7.5 at 25ºC. Setting: excitation wavelength 320 nm; excitation slit width 20 nm, emission slit width 10 nm; increment 1 nm; dwell time 0.5 s, gate time 5 ms, total decay time 0.2 s, delay time 0.2 ms; PMT detector voltage 1000 V.

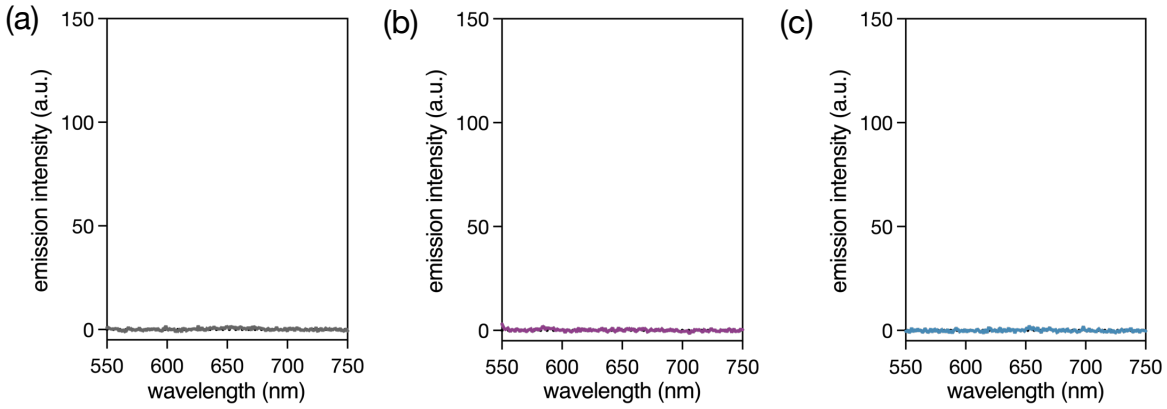


Figure 91. Time gated emission spectra of (a) 50 μM solutions of DOTA-$(E_4K_4)_4$ with no lanthanide added; (b) 50 μM solution of DOTA with 1.0 eq of $Tb(CF_3SO_3)_3$, and (c) 50 μM DOTA with 1.0 eq of $EuCl_3$.

[279] Ebeling, W. *et al. Eur. J. Biochem.* **47**, 91–7 (1974).

[280] (a) Schwartz, D. C., Cantor, C. R. *Cell* **37**, 67–75 (1984); (b) Cleveland, D. W. *et al. J. Biol. Chem.* **252**, 1102–6 (1977); (c) Herrmann, B. G., Frischauf, A. M. *Meth. Enzymol.* **152**, 180–3 (1987); (d) Lee, J. J., Costlow, N. A. *Meth. Enzymol.* **152**, 633–48 (1987).

7.2.4. DFT Models

The peptide initials structures were obtained by different software protein structure prediction methodologies. Namely, the five top predictions of RoseTTAFold, with the deep-learning-based modelling method and the ab initio Rosetta fragment assembly method, both available in the Robetta server, and AlphaFold3 as available in the AlfaFold3 server. In addition, AlphaFold3 results were also augmented by the SCWLR4 methodology to produce optimized side-chain interactions. As a reference, linear and Alpha-helical polymers were also built with the utilities available in Tinker with side chains optimized by the Monte Carlo (MC) procedure in z-matrix space with the Amber14SB ForceField (Amber FF) in a water continuum. The so-produced geometries were optimized with a procedure aimed to conserve the input molecular conformation in a water continuum with the eXtended Tight Binding XTB-GFN2 semiempirical methodology. The electronic properties of the resulting over 30 peptides were computed and energetically screened with the M06-2X-6-31G(d,p)-water methodology (**Table 4**). The lowest energy structure for each methodology was selected (**Figure 93**) to be computed with a larger basis set, as presented in the main text.

Table 4. Normalized relative energies of SAH peptides with different methods, HOMO and LUMO eigenvalues and corresponding gap computed with the M06-2X 6-31G(d,p)-water level on conformations generated by the tinker suite, RoseTTAFold, AlphaFold3 and AlphaFold3 with SCWRL4 side-chain optimization procedures. All values in eV but for the Relative Energies (RE) which are in kcal/mol/atom.

Model	Conf.	Conf #	Seq.	RE / kcal/mol /atom	LUMO	HOMO	GAP
Amber FF	Linear	NA	$(E_4K_4)_4$	0.55	1.4	-7.8	9.2
Amber FF, MC 10K steps	Linear	NA	$(E_4K_4)_4$	0.37	1.3	-7.8	9.1
Amber FF, MC 1K steps	Linear	NA	$(E_4K_4)_4$	0.39	1.1	-7.8	9.0
RosettaFold	α-helix	4	$(E_4K_4)_4$	0.05	0.9	-7.8	8.8
RosettaFold	α-helix	1	$(E_4K_4)_4$	0.00	0.9	-7.8	8.8
AlphaFold3 + SCWRL4	α-helix	3	$(E_4K_4)_4$	0.22	1.1	-7.6	8.7
RosettaFold (Ab Initio)	α-helix	3	$(E_4K_4)_4$	0.04	1.1	-7.6	8.7
AlphaFold3 + SCWRL4	α-helix	2	$(E_4K_4)_4$	0.22	1.1	-7.6	8.7
RosettaFold (Ab Initio)	α-helix	1	$(E_4K_4)_4$	0.01	1.1	-7.6	8.7
AlphaFold3 + SCWRL4	α-helix	4	$(E_4K_4)_4$	0.18	1.1	-7.6	8.6
Amber FF, MC 10K steps	α-helix	NA	$(E_4K_4)_4$	0.14	1.0	-7.6	8.6
AlphaFold3	α-helix	2	$(E_4K_4)_4$	0.30	1.1	-7.6	8.6
RosettaFold (Ab Initio)	α-helix	4	$(E_4K_4)_4$	0.04	1.1	-7.5	8.6
AlphaFold3 + SCWRL4	α-helix	1	$(E_4K_4)_4$	0.20	1.0	-7.6	8.6
AlphaFold3	α-helix	3	$(E_4K_4)_4$	0.20	1.0	-7.6	8.6
RosettaFold (Ab Initio)	α-helix	1	$(E_4K_4)_4$	0.09	1.0	-7.6	8.5
RosettaFold (Ab Initio)	α-helix	4	$(E_4K_4)_4$	0.10	0.9	-7.6	8.5
RosettaFold (Ab Initio)	α-helix	2	$(E_4K_4)_4$	0.07	1.1	-7.5	8.5
AlphaFold3 + SCWRL4	α-helix	3	$(E_2K_2)_8$	0.38	1.0	-7.5	8.5
RosettaFold (Ab Initio)	α-helix	2	$(E_4K_4)_4$	0.07	1.1	-7.5	8.5
AlphaFold3	α-helix	3	$(E_2K_2)_8$	0.40	1.0	-7.6	8.5
Amber FF, MC 1K steps	α-helix	NA	$(E_4K_4)_4$	0.21	0.9	-7.6	8.5
AlphaFold3 + SCWRL4	α-helix	0	$(E_2K_2)_8$	0.34	1.0	-7.5	8.5
AlphaFold3	α-helix	0	$(E_2K_2)_8$	0.38	1.0	-7.5	8.5
AlphaFold3 + SCWRL4	α-helix	4	$(E_2K_2)_8$	0.42	0.9	-7.5	8.5
AlphaFold3	α-helix	1	$(E_2K_2)_8$	0.17	0.9	-7.6	8.4
AlphaFold3 + SCWRL4	α-helix	0	$(E_2K_2)_8$	0.22	0.8	-7.6	8.4
AlphaFold3	α-helix	1	$(E_2K_2)_8$	0.38	0.8	-7.6	8.4
Amber FF	α-helix	NA	$(E_4K_4)_4$	0.22	0.8	-7.5	8.4
AlphaFold3 + SCWRL4	α-helix	2	$(E_2K_2)_8$	0.40	0.8	-7.5	8.3
AlphaFold3 + SCWRL4	α-helix	1	$(E_2K_2)_8$	0.34	0.6	-7.6	8.2

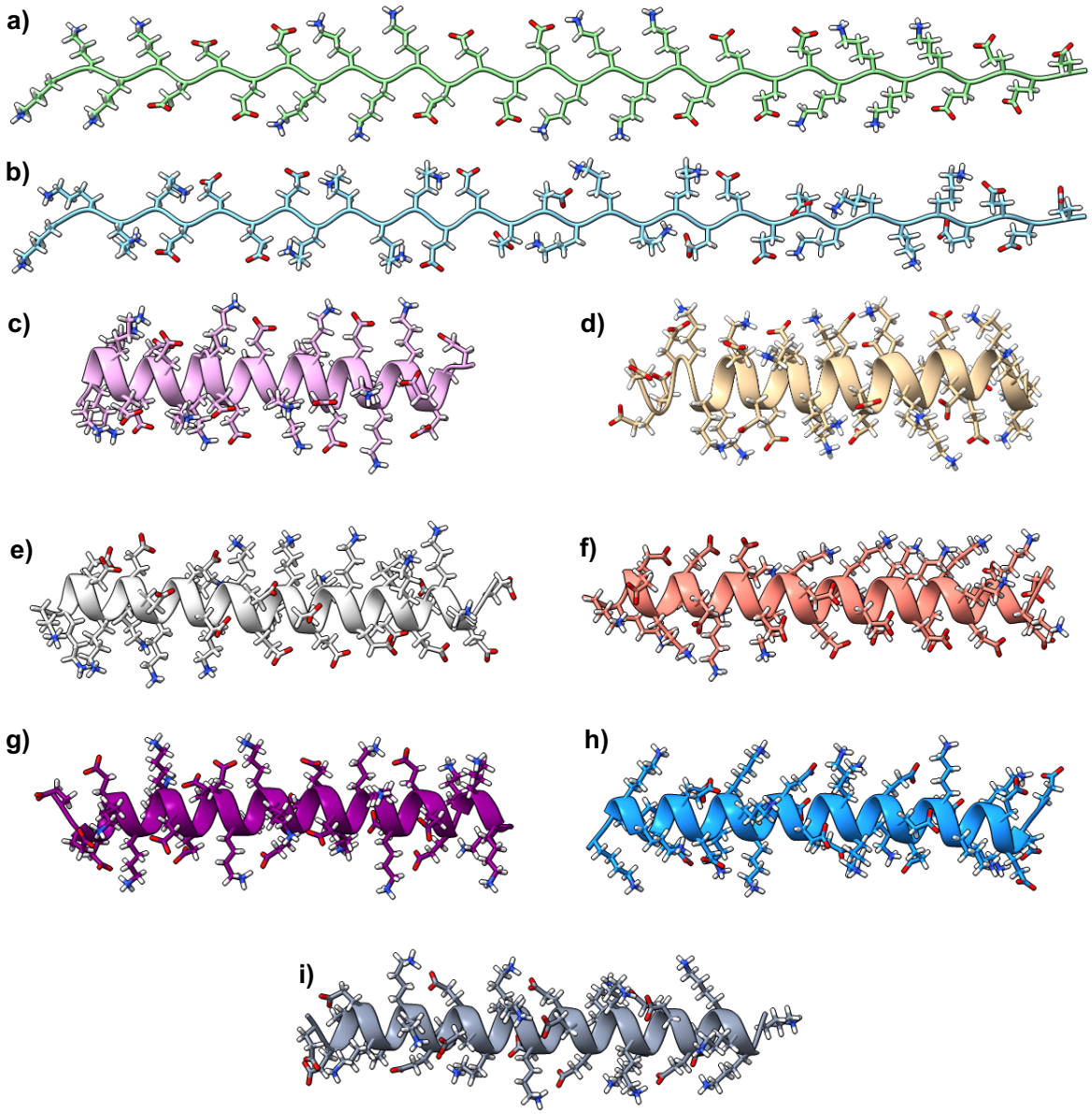


Figure 92. Minimized structures of extended conformations of $(E_4K_4)_4$ built with (a) Tinker and (b) Tinker with Monte Carlo optimization, α-helix conformations of $(E_4K_4)_4$ built with (c) Tinker and (d) Tinker with Monte Carlo optimization, α-helix conformations of $(E_2K_2)_8$ built with (e) AlphaFold3 and (f) AlphaFold3 with SCWRL4 and α-helix conformations of $(E_4K_4)_4$ built with (g) AlphaFold3, (h) AlphaFold3 with SCWRL4, and (i) RoseTTAFold.

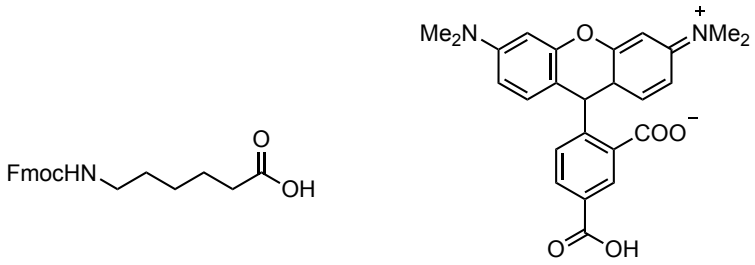
7.3. *De Novo* Engineering of Pd-Metalloproteins and Their Use as Intracellular Catalysts

7.3.1. Synthesis and Characterization of Peptides

Table 5. Name and sequence of the peptides used in this study. All peptides are N-terminal free amine, C-terminal amide. In grey indicated the 1EOM origin sequence.

Peptide	Sequence
1E0M	SMGLPPGWDE YKTHNGKTYY YNHNTKTSTW TDPRMSS
WW0	––GLPPGWDE YKTPNGKTYY YNTNTKTSTW TDPR–––
WW11/13	––GLPPGWDE HKHPNGKTYY YNTNTKTSTW TDPR–––
WW11/19	––GLPPGWDE HKTPNGKTHY YNTNTKTSTW TDPR–––
WW13/19	––GLPPGWDE YKHPNGKTHY YNTNTKTSTW TDPR–––
WW9/21	––GLPPGWHE YKTPNGKTYY HNTNTKTSTW TDPR–––
WW9/23	––GLPPGWHE YKTPNGKTYY YNHNTKTSTW TDPR–––
WW21/23	––GLPPGWDE YKTPNGKTYY HNHNTKTSTW TDPR–––
WW19	––GLPPGWDE YKTPNGKTHY YNTNTKTSTW TDPR–––

Synthesis of labeled TMR-WW13/19 and TMR-WW19. Once the peptides were fully assembled and still attached to the solid support, we coupled Fmoc-6-aminohexanoic acid (Fmoc-Ahx-OH) as spacer using standard solid-phase methods (activation with 1.0 equiv. of HATU in 2 mL DMF with 3 mL 0.195 M DIEA/DMF). The terminal Fmoc protecting group was removed (20% piperidine/DMF) and a mixture of 5-carboxytetramethylrhodamine (TMR-OH, 3.0 equiv., 0.15 mmol, 64.5 mg), 3 equiv. of HATU and 5 equiv. of DIEA 0.195 M in DMF was added onto the resin and shaken for 60 min.



Fmoc-Ahx-OH **5-carboxytetramethylrhodamine (TMR-OH)**

Figure 93. Structures of Fmoc-Ahx-OH, and 5-carboxytetramethylrhodamine (TMR-OH).

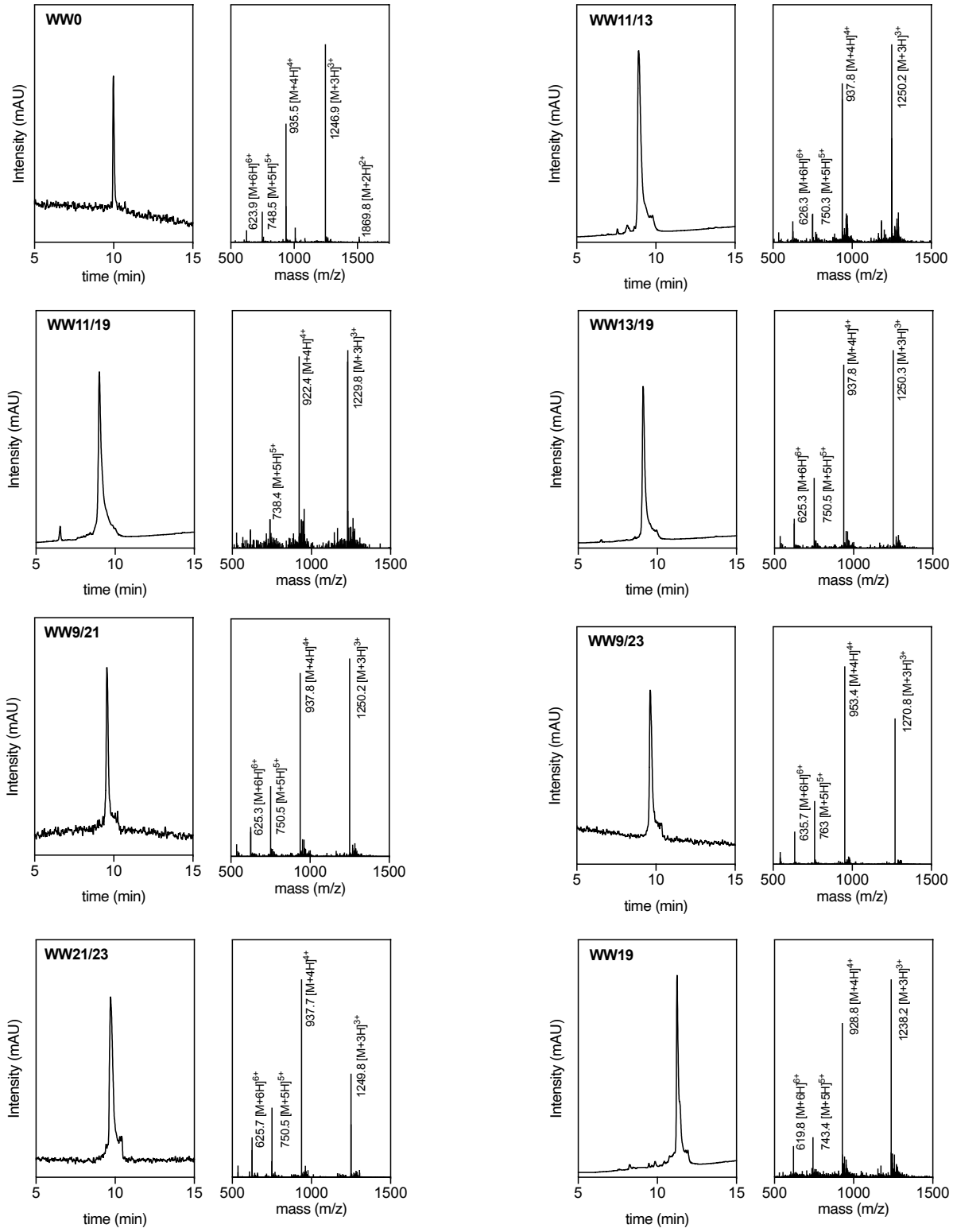


Figure 94. HPLC traces of the purified WW peptides used in this study and their corresponding ESI-MS spectra showing in each case peaks matching the expected masses.

WW0 HPLC-MS (ESI): calculated for $C_{170}H_{243}N_{43}O_{54}$ $[M+H]^+$ = 3 753.1; found $[M+2H]^{2+}$ = 1 869.8, $[M+3H]^{3+}$ = 1 246.9, $[M+4H]^{4+}$ = 935.5, $[M+5H]^{5+}$ = 748.5, $[M+6H]^{6+}$ = 623.9. **WW11/13** HPLC-MS (ESI): calculated for $C_{169}H_{241}N_{47}O_{52}$ $[M+H]^+$ = 3 763.1; found $[M+3H]^{3+}$ = 1 250.2, $[M+4H]^{4+}$ = 937.8, $[M+5H]^{5+}$ = 750.3, $[M+6H]^{6+}$ = 626.3. **WW11/19** HPLC-MS (ESI): calculated for $C_{164}H_{239}N_{47}O_{52}$ $[M+H]^+$ = 3 701.0; found $[M+3H]^{3+}$ = 1 229.8, $[M+4H]^{4+}$ = 922.4, $[M+5H]^{5+}$ = 738.4. **WW13/19** HPLC-MS (ESI): calculated for $C_{169}H_{241}N_{47}O_{52}$ $[M+H]^+$ = 3 763.1; found $[M+3H]^{3+}$ = 1 250.3, $[M+4H]^{4+}$ = 937.8, $[M+5H]^{5+}$ = 750.5, $[M+6H]^{6+}$ = 625.3. **WW9/21** HPLC-MS (ESI): calculated for $C_{169}H_{243}N_{47}O_{51}$ $[M+H]^+$ = 3 749.1; found $[M+3H]^{3+}$ = 1 250.2, $[M+4H]^{4+}$ = 937.8, $[M+5H]^{5+}$ = 750.5, $[M+6H]^{6+}$ = 625.3. **WW9/23** HPLC-MS (ESI): calculated for $C_{174}H_{245}N_{47}O_{51}$ $[M+H]^+$ = 3 811.1; found $[M+3H]^{3+}$ = 1 270.8, $[M+4H]^{4+}$ = 953.4, $[M+5H]^{5+}$ = 763.0, $[M+6H]^{6+}$ = 635.7. **WW21/23** HPLC-MS (ESI): calculated for $C_{169}H_{241}N_{47}O_{52}$ $[M+H]^+$ = 3 763.1; found $[M+3H]^{3+}$ = 1 249.8, $[M+4H]^{4+}$ = 937.7, $[M+5H]^{5+}$ = 750.5, $[M+6H]^{6+}$ = 625.7. **WW19** HPLC-MS (ESI): calculated for $C_{167}H_{241}N_{45}O_{53}$ $[M+H]^+$ = 3 727.0; found $[M+3H]^{3+}$ = 1 238.2, $[M+4H]^{4+}$ = 928.8, $[M+5H]^{5+}$ = 743.4, $[M+6H]^{6+}$ = 619.8.

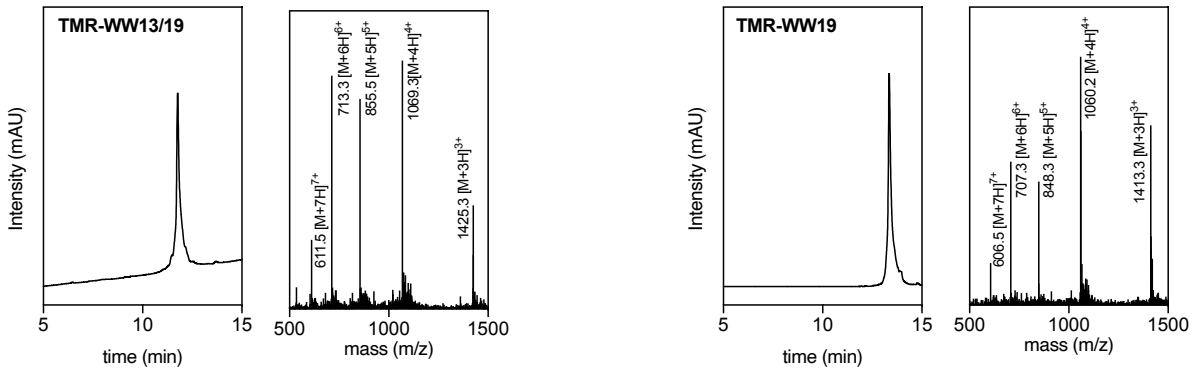


Figure 95. HPLC traces of the purified TMR labeled peptides used in this study and their corresponding ESI-MS spectra showing in each case peaks matching the expected masses.

TMR-WW13/19 HPLC-MS (ESI): calculated for $C_{200}H_{278}N_{51}O_{55}$ $[M+H]^+$ = 4 276.7; found $[M+3H]^{3+}$ = 1 425.3, $[M+4H]^{4+}$ = 1 069.3, $[M+5H]^{5+}$ = 855.5, $[M+6H]^{6+}$ = 713.3, $[M+7H]^{7+}$ = 611.5. **TMR-WW19** HPLC-MS (ESI): calculated for $C_{198}H_{280}N_{49}O_{56}$ $[M+H]^+$ = 4 242.7; found $[M+3H]^{3+}$ = 1 413.3, $[M+4H]^{4+}$ = 1 060.2, $[M+5H]^{5+}$ = 848.3, $[M+6H]^{6+}$ = 707.3, $[M+7H]^{7+}$ = 606.5.

Palladium coordinated peptides. Peptides (1 mM) were mixed with equimolar amounts of $[Pd(COD)Cl_2]$ in H_2O and the resulting mixtures were analyzed after 15 min. by Electrospray Ionization Mass Spectrometry (ESI/MS) with an *Agilent 6120 Quadrupole* LC/MS model in positive scan mode using direct injection of the purified peptide solution into the MS detector. The peaks corresponding to the apo-peptides are shown in black, and those corresponding to the complexes, that show the mass of the peptide and Pd(II), are in blue.

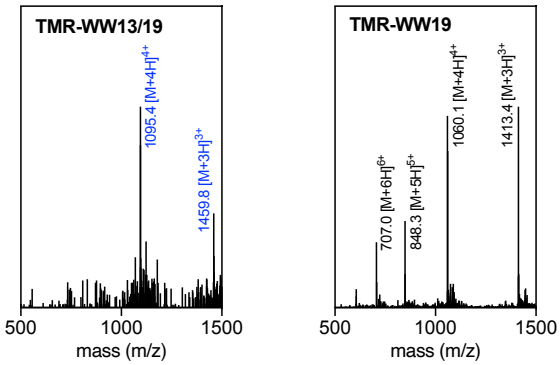


Figure 96. ESI-MS spectra of the mixture of TMR labeled peptides (1 mM) with 1.0 equiv. of $Pd(COD)Cl_2$ in water after 15 min incubation. The peaks corresponding to the apo-peptides are labeled in black, those of the corresponding palladopeptides, are in blue.

7.3.2. Synthesis and Characterization of Probes

Synthesis of (1). The synthesis was made according to a previously published procedure, described by Mascareñas *et al.*[281].

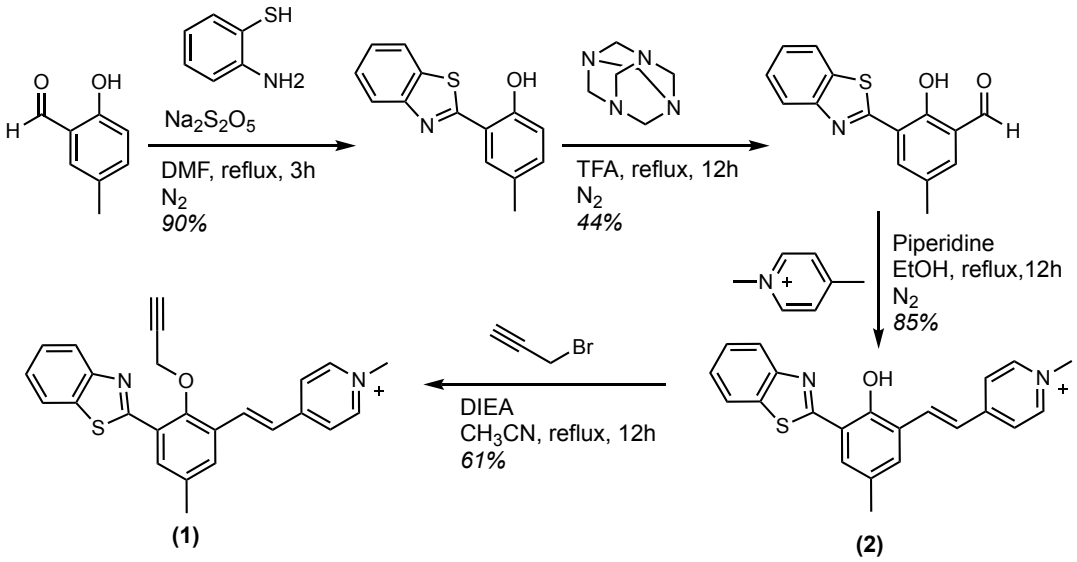


Figure 97. Synthetic route to obtain propargylated derived product (**1**).

2-(benzo[d]thiazol-2-yl)-4-methylphenol. 2-hydroxy-5-methylbenzaldehyde (1.50 g, 11.0 mmol, 1.1 equiv.) and $Na_2S_2O_5$ (1.80 g, 9.5 mmol, 1.0 equiv.) were dissolved in 15 mL anhydrous DMF in a Schlenk flask under an inert atmosphere. 2-aminobenzenethiol (1.20 ml, 11.0 mmol, 1.1 equiv.) was added, and the reaction was then stirred at 110 °C for 3 h. 10 mL mili-Q water were added to the reaction mixture leading to the formation of a white precipitate. The precipitate was collected by filtration to afford the desired compound as a white powder

[281] Learte-Aymamí, S. *et al. Angew. Chem. Int. Ed Engl.* **59**, 9149–9154 (2020).

(2.57 g, 90%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 12.28 (s, 1H), 7.96 (d, *J* = 8.1 Hz, 1H), 7.87 (d, *J* = 7.9 Hz, 1H), 7.54 – 7.32 (m, 3H), 7.18 (d, *J* = 8.6 Hz, 1H), 7.01 (d, *J* = 8.5 Hz, 1H), 2.35 (s, 4H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 169.5, 155.9, 152.0, 133.8, 132.7, 128.7, 128.4, 126.7, 125.5, 122.2, 121.5, 117.7, 116.4, 20.5.

3-(benzo[d]thiazol-2-yl)-2-hydroxy-5-methylbenzaldehyde. 2-(benzo[d]thiazol-2-*yl*)-4-methyl phenol (2.57 g, 10.7 mmol, 1 equiv.) and hexamethylenetetramine (4.51 g, 32.0 mmol, 3.0 equiv.) were dissolved in TFA (18 mL) under inert atmosphere, the reaction was heated at reflux overnight. 4 M HCl (180 mL) was added to the reaction mixture and the resulting mixture was extracted with Et_2O (2 × 150 mL). The organic layer was dried wit $MgSO_4$, filtered, and the solvent evaporated. The crude was redissolved in cold Et_2O and to afford the desired compound as a white solid (1.28 g, 44%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 10.47 (s, 1H), 8.03 (d, *J* = 8.1 Hz, 1H), 7.93 (m, 2H), 7.70 (s, 1H), 7.54 (t, *J* = 7.2 Hz, 1H), 7.44 (t, *J* = 7.6 Hz, 1H), 2.40 (s, 3H).^{13}C NMR (75 MHz, $CDCl_3$) δ 191.2, 167.0, 158.7, 151.5, 135.4, 133.3, 132.9, 129.1, 127.0, 126.0, 123.8, 122.5, 121.7, 118.85, 20.4.

(E)-4-(3-(benzo[d]thiazol-2-yl)-2-hydroxy-5-methylstyryl)-1-methylpyridin-1-ium ***(2)****.* 3-(benzo[*d*]thiazol-2-*yl*)-2-hydroxy-5-methylbenzaldehyde (1.00 g, 3.70 mmol, 1.0 equiv.), 1,4-dimethylpyridinium and piperidine (385.0 µL, 3.70 mmol, 1.0 equiv.) were dissolved in EtOH (75 mL) and refluxed overnight under N_2 atmosphere. The solvents were removed under vacuum and the resulting solid is redissolved in the minimal amount of MeOH and precipitated with cold Et_2O to afford the compound (**2**) after filtration as an orange solid (1.52 g, 85%). ^{1}H NMR (300 MHz, DMSO-d^6) δ 13.05 (s, 1H), 8.84 (d, *J* = 6.9 Hz, 2H), 8.21 (d, *J* = 6.9 Hz, 3H), 8.16 – 8.04 (m, 2H), 7.76 (d, *J* = 9.6 Hz, 2H), 7.69 – 7.46 (m, 3H), 4.27 (s, 3H), 2.38 (s, 3H). ^{13}C NMR (75 MHz, DMSO-d^6) δ 168.2, 154.0, 152.6, 150.7, 145.0, 134.7, 132.4, 130.6, 128.9, 127.1, 126.0, 124.3, 123.5, 123.4, 122.3, 121.9, 117.0, 46.9, 19.9.

(E)-4-(3-(benzo[d]thiazol-2-yl)-5-methyl-2-(prop-2-yn-1-yloxy)styryl)-1-methylpyridin-1-ium ***(1)*****.** To a solution of compound **2** (972.0 mg, 2.0 mmol, 1.0 equiv.) in CH_3CN (7 mL) propargyl bromide (350. 0 µL, 4.0 mmol, 2.0 equiv.) and DIEA (700.0 µL, 4.0 mmol, 2.0 equiv.) were added, and the reaction mixture was refluxed overnight. Cold MeOH was added, and the precipitate was collected by filtration. Compound (**1**) was obtained as bright yellow solid after washing with MeOH (583.0 mg, 61%)[282]. ^{1}H NMR (300 MHz, DMSO-d^6) δ 8.70 (d, *J* = 6.7 Hz, 2H), 8.06 (d, *J* = 7.1 Hz, 2H), 8.01 (s, 1H), 7.95 (d, *J* = 7.9 Hz, 1H), 7.91 – 7.80 (m, 2H), 7.68 (s, 1H), 7.46 – 7.31 (m, 2H), 7.26 (t, *J* = 7.7 Hz, 1H), 4.59 (s, 2H), 4.08 (s, 3H), 3.08 (s, 1H), 2.24 (s, 3H). ^{13}C NMR (75 MHz, DMSO-d^6) δ 161.9, 152.2, 152.0, 151.7, 145.2, 135.5, 135.0, 133.9, 131.0, 130.6, 129.8, 127.4, 126.4, 125.5, 125.4, 123.9, 122.7, 122.0, 80.2, 78.1, 62.8, 47.0, 20.4.

[282] Gao, T. *et al. Chem. Asian J.* **10**, 1142–1145 (2015).

Synthesis of (3)

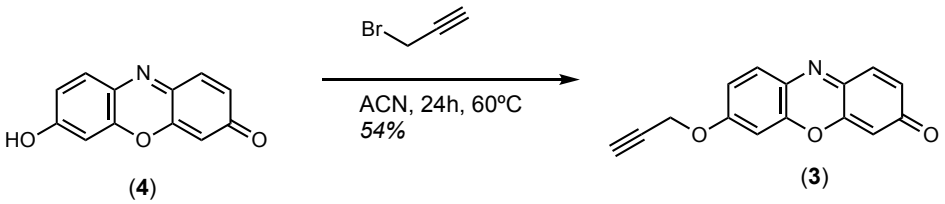


Figure 98. Synthetic route to obtain propargylated derived product (**3**).

2-(benzo[d]thiazol-2-yl)-4-methylphenol. Resorufin (**4**) (353.0 mg, 1.5 mmol, 1.0 equiv.) and DBU (456.6 mg, 3.0 mmol, 2.0 equiv.) were dissolved in 5 mL anhydrous DMF in a Schlenk flask under an inert atmosphere. Propargyl bromide (0.26 ml, 3.0 mmol, 2.0 equiv.) was added dropwise, and the reaction was then stirred at room temperature for 18 h. The solvent was removed in vacuo and DCM was added, and the precipitate was collected by filtration. The resulting crude was purified by silica gel column chromatography using a DCM/MeOH (9:1) mixture. The compound **3** was obtained as an orange solid (204.1 mg, 54%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.73 (d, J = 8.9 Hz, 1H), 7.42 (d, J = 9.8 Hz, 1H), 7.01 (dd, J = 8.9, 2.7 Hz, 1H), 6.93 (d, J = 2.7 Hz, 1H), 6.84 (dd, J = 9.8, 2.0 Hz, 1H), 6.33 (d, J = 2.0 Hz, 1H), 4.81 (d, J = 2.4 Hz, 2H), 2.61 (t, J = 2.4 Hz, 1H). ^{13}C NMR (75 MHz, DMSO-d^6) δ 185.3, 161.0, 149.7, 149.6, 135.1, 134.9, 133.8, 131.2, 128.2, 114.1, 105.7, 101.4, 79.2, 78.2, 56.5.

Synthesis of (5). The synthesis was made according to a previously published procedure, described by Bradely *et al.*[283].

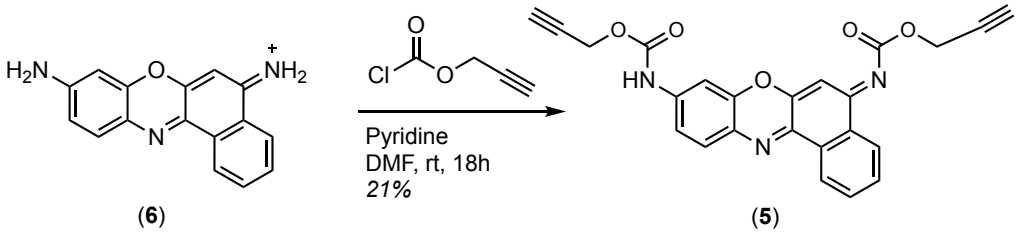


Figure 99. Synthetic route to obtain propargylated derived product (**5**).

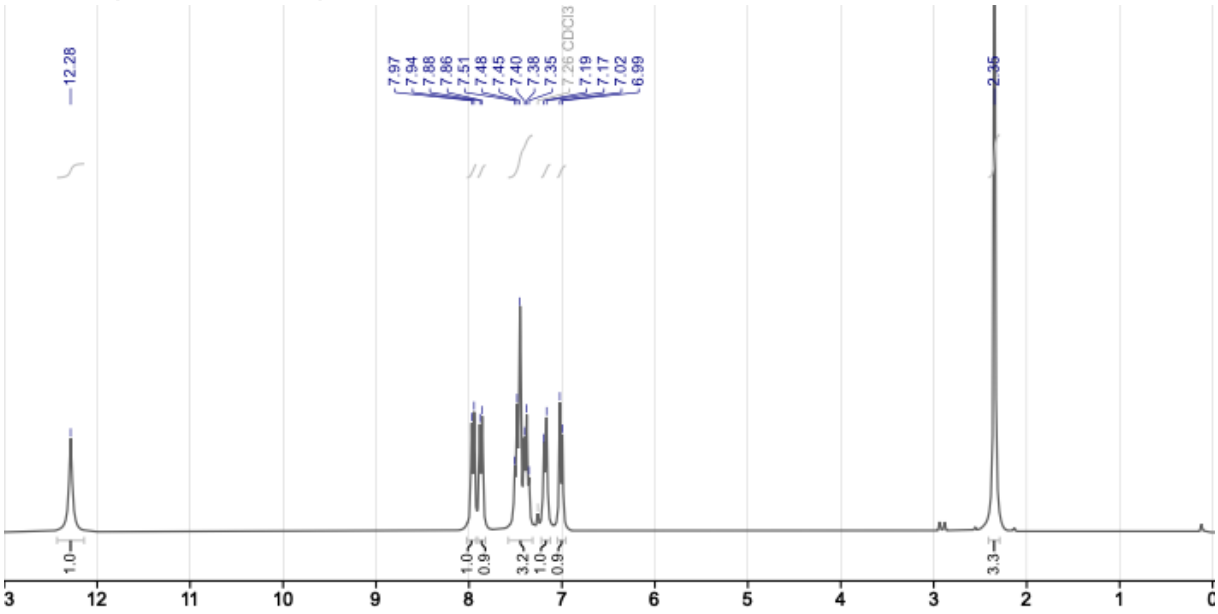
*Prop-2-yn-1-yl(E)-(5-(((prop-2-yn-1-yloxy)carbonyl)imino)-5H-benzo[a]phenoxazin-9-yl)carbamate (**5**).* Cresyl violet perchlorate (**6**) (100 mg, 0.27 mmol) was dissolved in anhydrous DMF (1.0 mL) and cooled to 0 ˚C. Separately, pyridine (140 µL, 2.8 mmol, 10.0 equiv.) and propargyl chloroformate (116 µL, 1.4 mmol, 5 equiv.) were each dissolved in anhydrous DMF (0.5 mL) and the solutions added dropwise to the cresyl violet. The reaction mixture was stirred at room temperature for 18 h, after which it was diluted with iPrOH/$CHCl_3$ (25:75, 10.0 mL), washed with 5% HCl (2 × 10.0 mL) and $NaHCO_3$ (2 × 5.0 mL) and dried over MgSO4. The solvent was removed in vacuo and the resulting crude was purified by silica gel column chromatography using a hexane/ethyl acetate (2:1) mixture. The compound **5** was obtained

[283] Clavadetscher, J. *et al. Angew. Chem. Int. Ed.* **56**, 6864–6868 (2017).

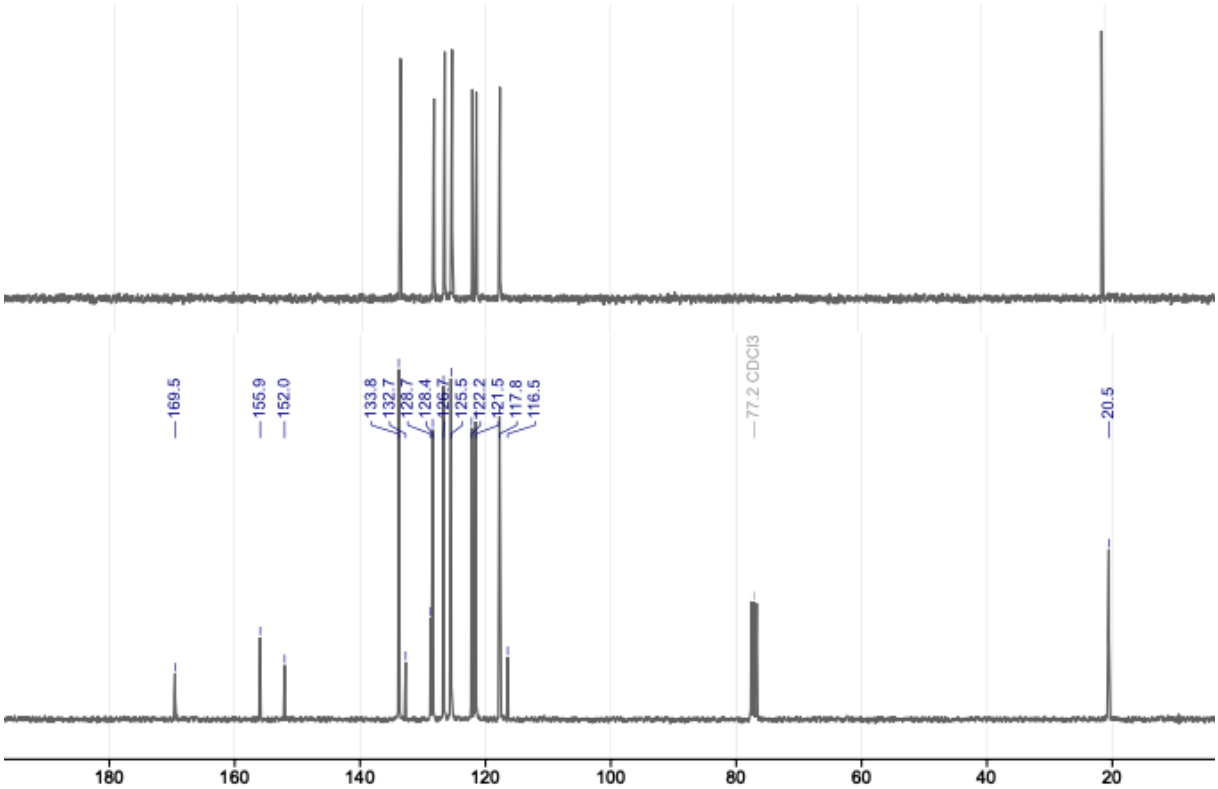
as an orange solid (21%). ^{1}H NMR (300 MHz, DMSO-d^{6}) δ 10.46 (s, 1H), 8.58 (d, *J* = 9.3 Hz, 1H), 8.30 (d, *J* = 9.4 Hz, 1H), 7.93 – 7.67 (m, 3H), 7.58 (s, 1H), 7.42 (d, *J* = 8.7 Hz, 1H), 4.87 (dd, *J* = 26.4, 2.4 Hz, 4H), 3.63 (dt, *J* = 15.6, 2.4 Hz, 2H). ^{13}C NMR (75 MHz, DMSO-d^{6}) δ 162.1, 159.3, 152.3, 148.6, 144.3, 142.2, 131.7, 131.3, 130.9, 129.9, 128.4, 124.8, 124.1, 115.3, 103.8, 100.1, 78.6, 78.5, 77.9, 53.6, 52.5.

2-(benzo[d]thiazol-2-yl)-4-methylphenol NMR characterization:

^{1}H NMR (300 MHz, $CDCl_3$)



^{13}C NMR (75 MHz, $CDCl_3$)

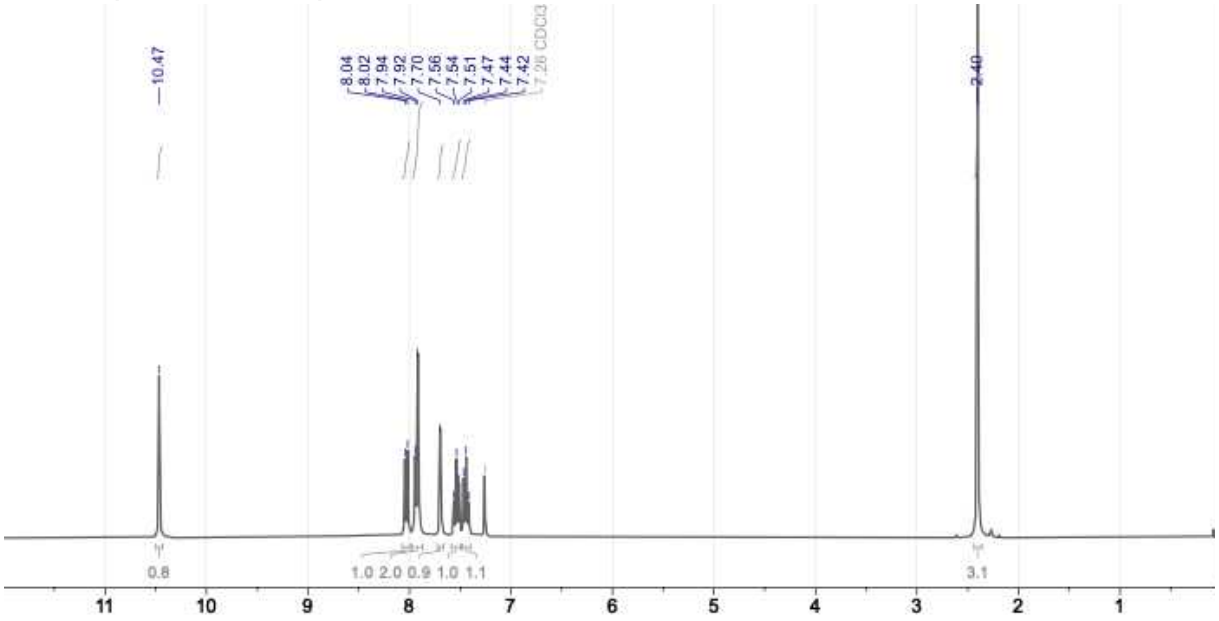


^{1}H NMR (300 MHz, $CDCl_3$) δ 12.28 (s, 1H), 7.96 (d, *J* = 8.1 Hz, 1H), 7.87 (d, *J* = 7.9 Hz, 1H), 7.54 – 7.32 (m, 3H), 7.18 (d, *J* = 8.6 Hz, 1H), 7.01 (d, *J* = 8.5 Hz, 1H), 2.35 (s, 4H). **^{13}C NMR** (75 MHz, $CDCl_3$) δ 169.5, 155.9, 152.0, 133.8, 132.7, 128.7, 128.4, 126.7, 125.5, 122.2, 121.5, 117.7, 116.4, 20.5.

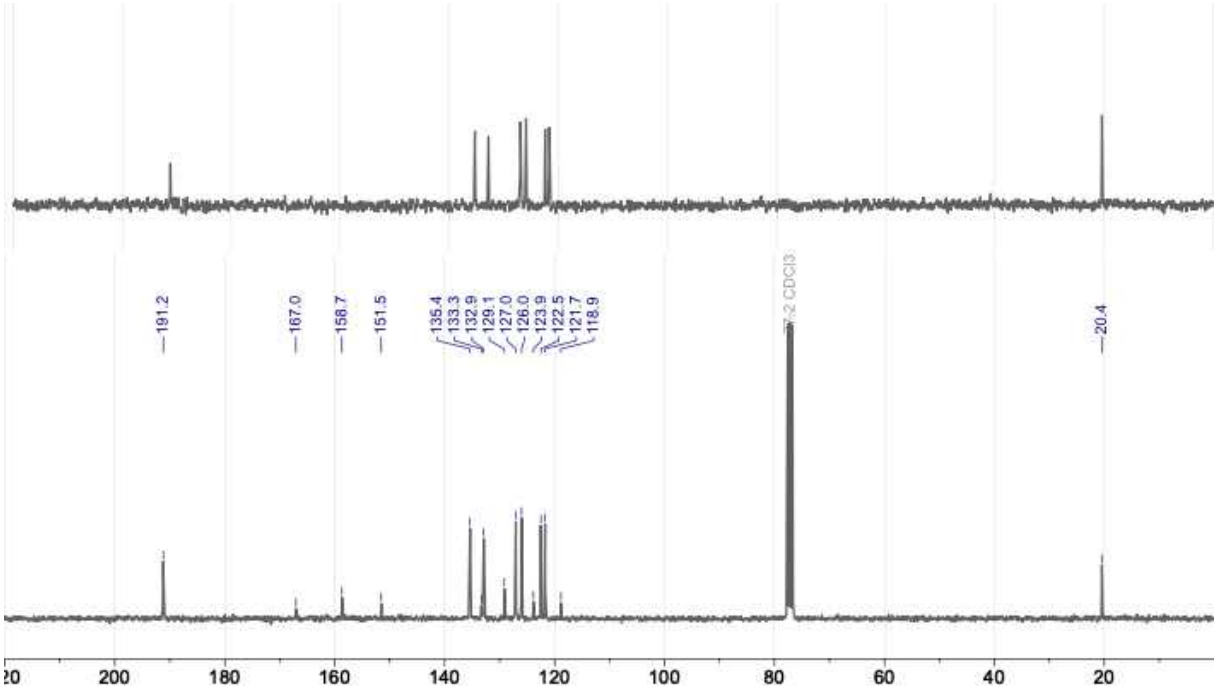
Figure 100. ^{1}H and ^{13}C NMR characterization of 2-(benzo[d]thiazol-2-yl)-4-methylphenol in $CDCl_3$.

3-(benzo[d]thiazol-2-yl)-2-hydroxy-5-methylbenzaldehyde NMR characterization:

^{1}H NMR (300 MHz, $CDCl_3$)



^{13}C NMR (75 MHz, $CDCl_3$)

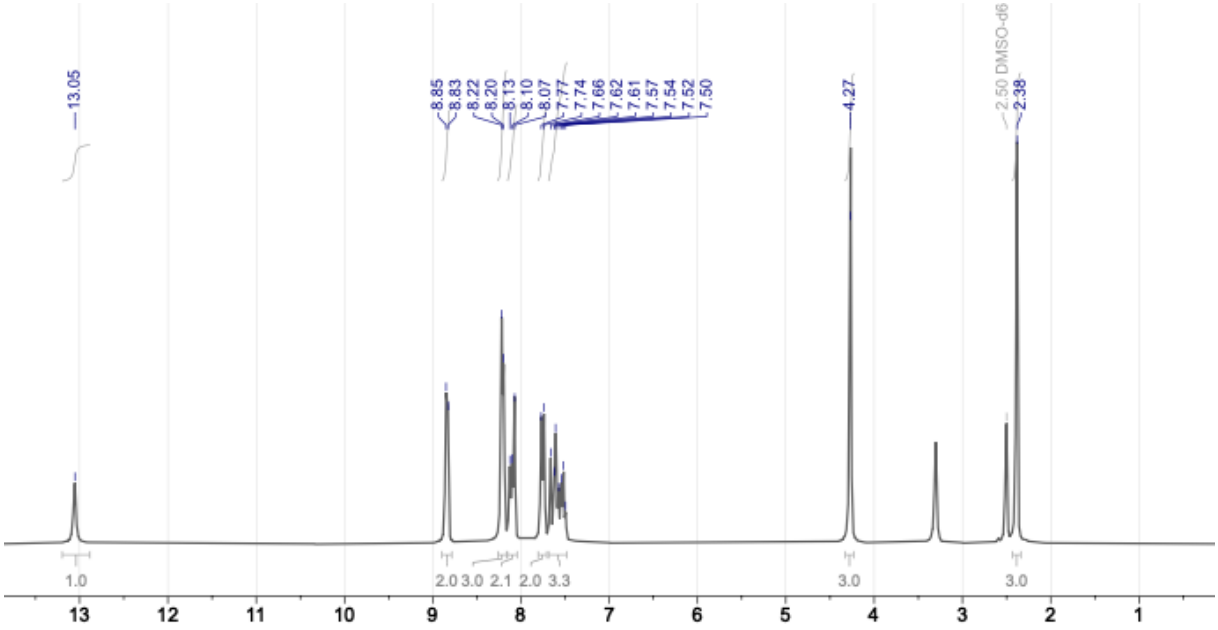


^{1}H NMR (300 MHz, $CDCl_3$) δ 10.47 (s, 1H), 8.03 (d, J = 8.1 Hz, 1H), 7.93 (m, 2H), 7.70 (s, 1H), 7.54 (t, J = 7.2 Hz, 1H), 7.44 (t, J = 7.6 Hz, 1H), 2.40 (s, 3H).**^{13}C NMR** (75 MHz, $CDCl_3$) δ 191.2, 167.0, 158.7, 151.5, 135.4, 133.3, 132.9, 129.1, 127.0, 126.0, 123.8, 122.5, 121.7, 118.85, 20.4.

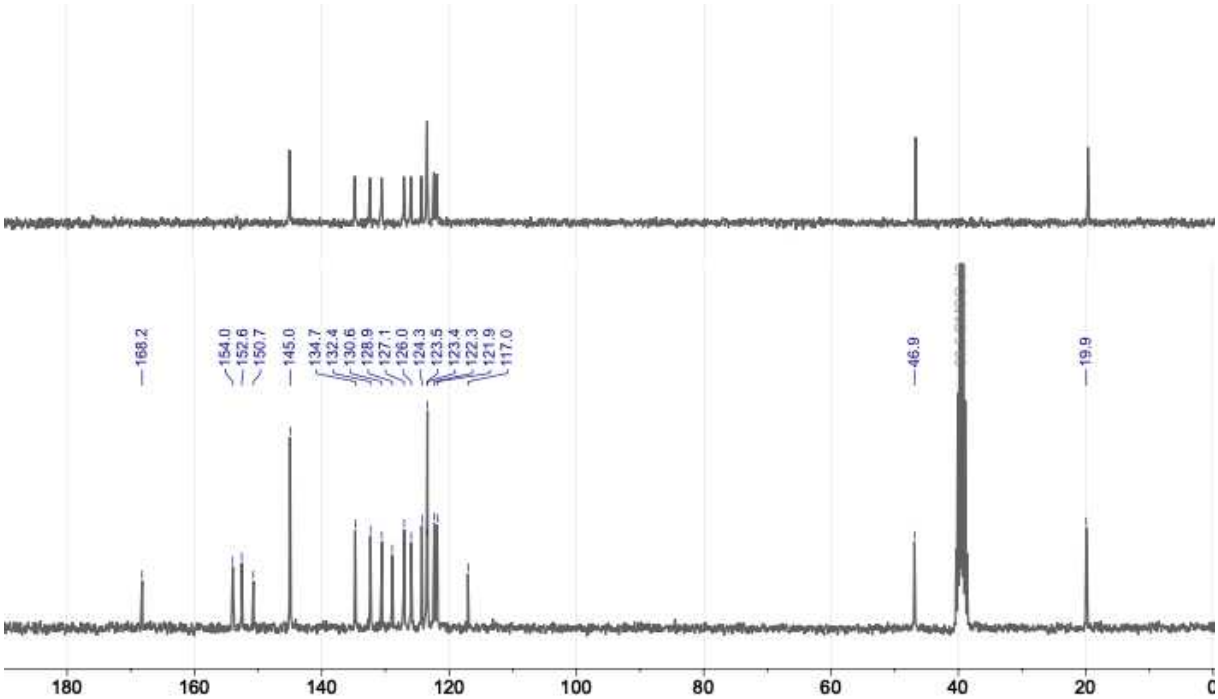
Figure 101. ^{1}H and ^{13}C NMR characterization of 3-(benzo[d]thiazol-2-yl)-2-hydroxy-5-methylbenzaldehyde in $CDCl_3$.

*Probe (**2**) NMR characterization*:

^{1}H NMR (300 MHz, DMSO)



^{13}C NMR (75 MHz, DMSO)

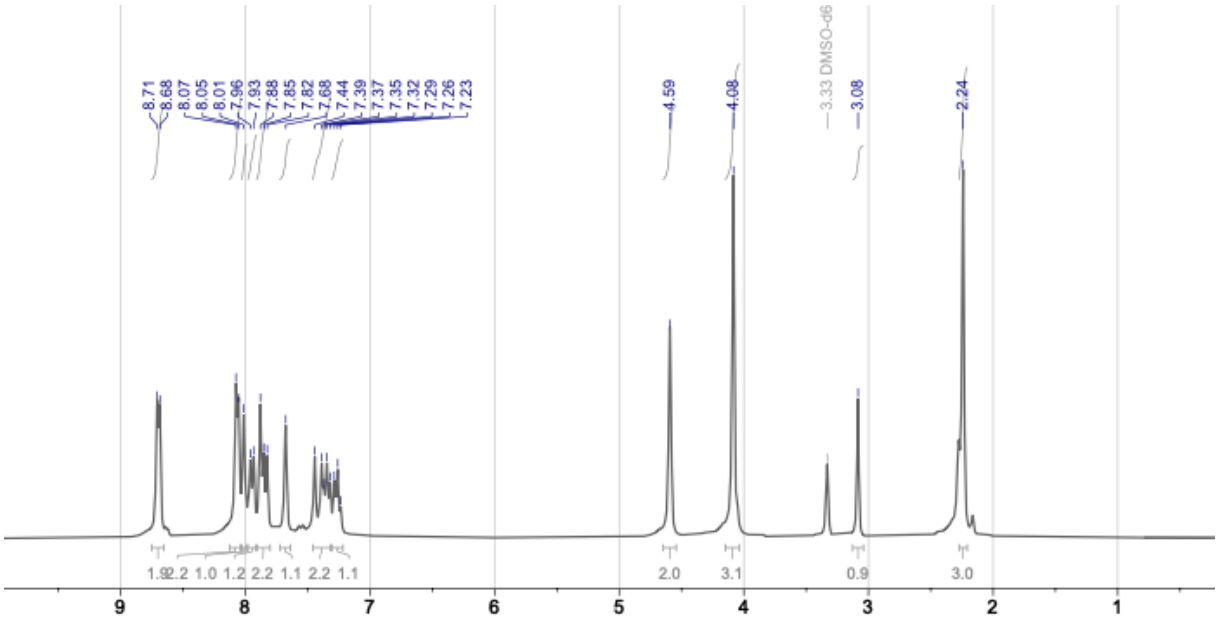


^{1}H NMR (300 MHz, DMSO-d^{6}) δ 13.05 (s, 1H), 8.84 (d, *J* = 6.9 Hz, 2H), 8.21 (d, *J* = 6.9 Hz, 3H), 8.16 – 8.04 (m, 2H), 7.76 (d, *J* = 9.6 Hz, 2H), 7.69 – 7.46 (m, 3H), 4.27 (s, 3H), 2.38 (s, 3H). **^{13}C NMR** (75 MHz, DMSO-d^{6}) δ 168.2, 154.0, 152.6, 150.7, 145.0, 134.7, 132.4, 130.6, 128.9, 127.1, 126.0, 124.3, 123.5, 123.4, 122.3, 121.9, 117.0, 46.9, 19.9.

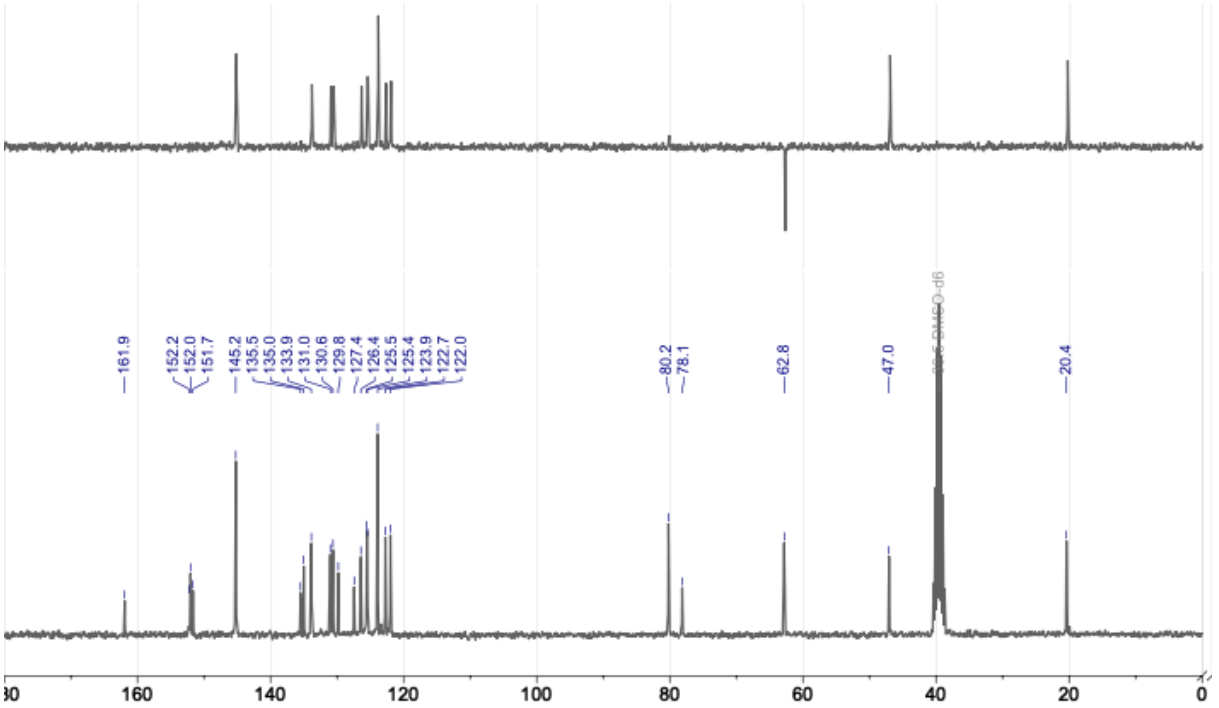
Figure 102. ^{1}H and ^{13}C NMR characterization of (E)-4-(3-(benzo[d]thiazol-2-yl)-2-hydroxy-5-methylstyryl)-1-methylpyridin-1-ium (2) in DMSO.

*Propargylated probe (**1**) NMR characterization*:

^{1}H NMR (300 MHz, DMSO)



^{13}C NMR (75 MHz, DMSO)

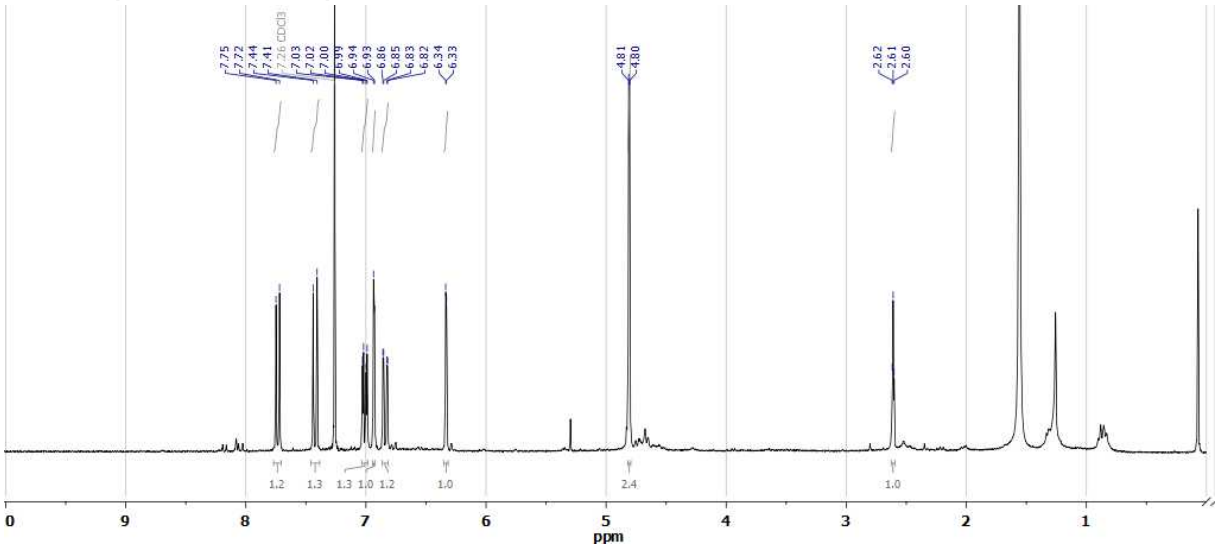


^{1}H NMR (300 MHz, DMSO-d^{6}) δ 8.70 (d, *J* = 6.7 Hz, 2H), 8.06 (d, *J* = 7.1 Hz, 2H), 8.01 (s, 1H), 7.95 (d, *J* = 7.9 Hz, 1H), 7.91 – 7.80 (m, 2H), 7.68 (s, 1H), 7.46 – 7.31 (m, 2H), 7.26 (t, *J* = 7.7 Hz, 1H), 4.59 (s, 2H), 4.08 (s, 3H), 3.08 (s, 1H), 2.24 (s, 3H). **^{13}C NMR** (75 MHz, DMSO-d^{6}) δ 161.9, 152.2, 152.0, 151.7, 145.2, 135.5, 135.0, 133.9, 131.0, 130.6, 129.8, 127.4, 126.4, 125.5, 125.4, 123.9, 122.7, 122.0, 80.2, 78.1, 62.8, 47.0, 20.4.

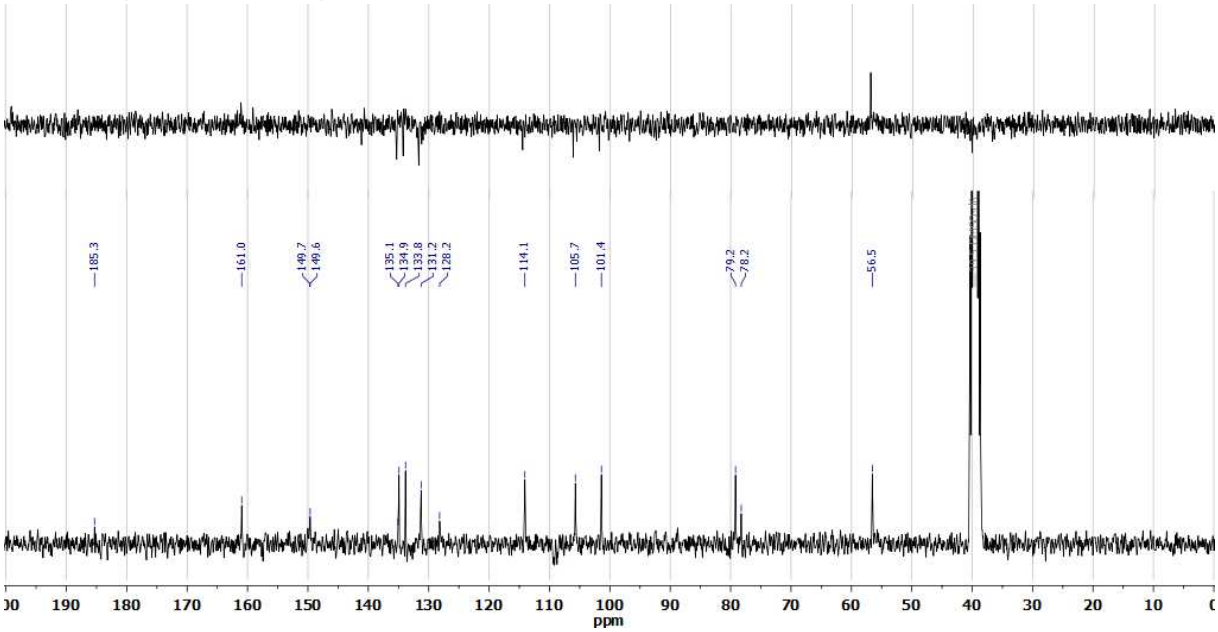
Figure 103. ^{1}H and ^{13}C NMR characterization of (**1**) in DMSO.

*Propargylated probe (**3**) NMR characterization:*

^{1}H NMR (300 MHz, DMSO)



^{13}C NMR (75 MHz, DMSO)

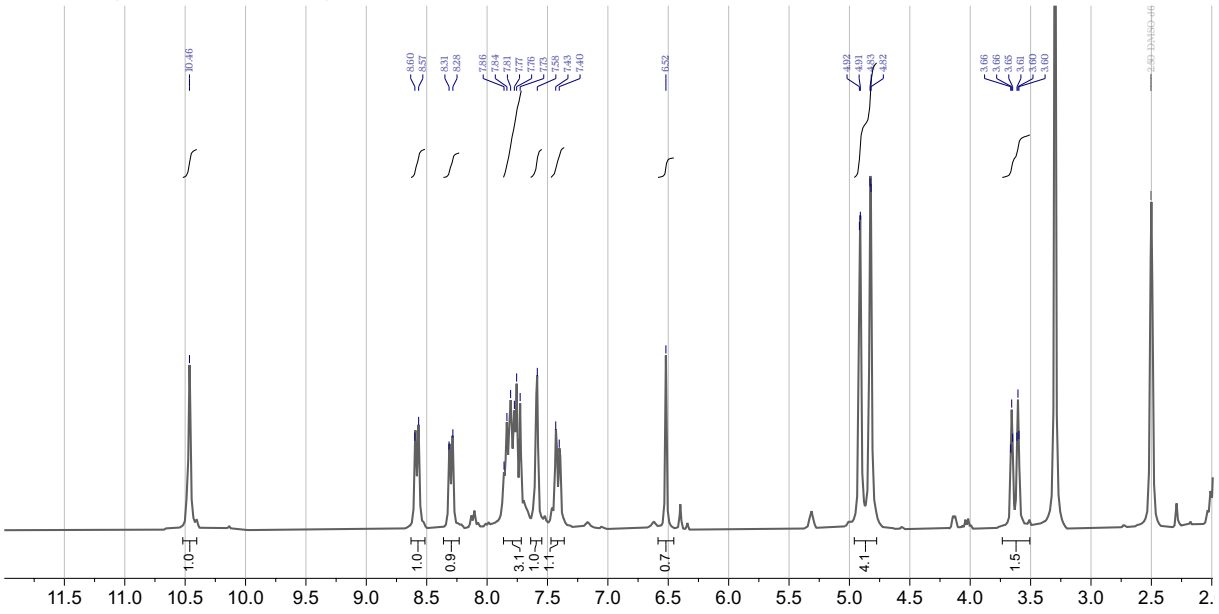


^{1}H NMR (300 MHz, $CDCl_3$) δ 7.73 (d, J = 8.9 Hz, 1H), 7.42 (d, J = 9.8 Hz, 1H), 7.01 (dd, J = 8.9, 2.7 Hz, 1H), 6.93 (d, J = 2.7 Hz, 1H), 6.84 (dd, J = 9.8, 2.0 Hz, 1H), 6.33 (d, J = 2.0 Hz, 1H), 4.81 (d, J = 2.4 Hz, 2H), 2.61 (t, J = 2.4 Hz, 1H). **^{13}C NMR** (75 MHz, DMSO-d^6) δ 185.3, 161.0, 149.7, 149.6, 135.1, 134.9, 133.8, 131.2, 128.2, 114.1, 105.7, 101.4, 79.2, 78.2, 56.5.

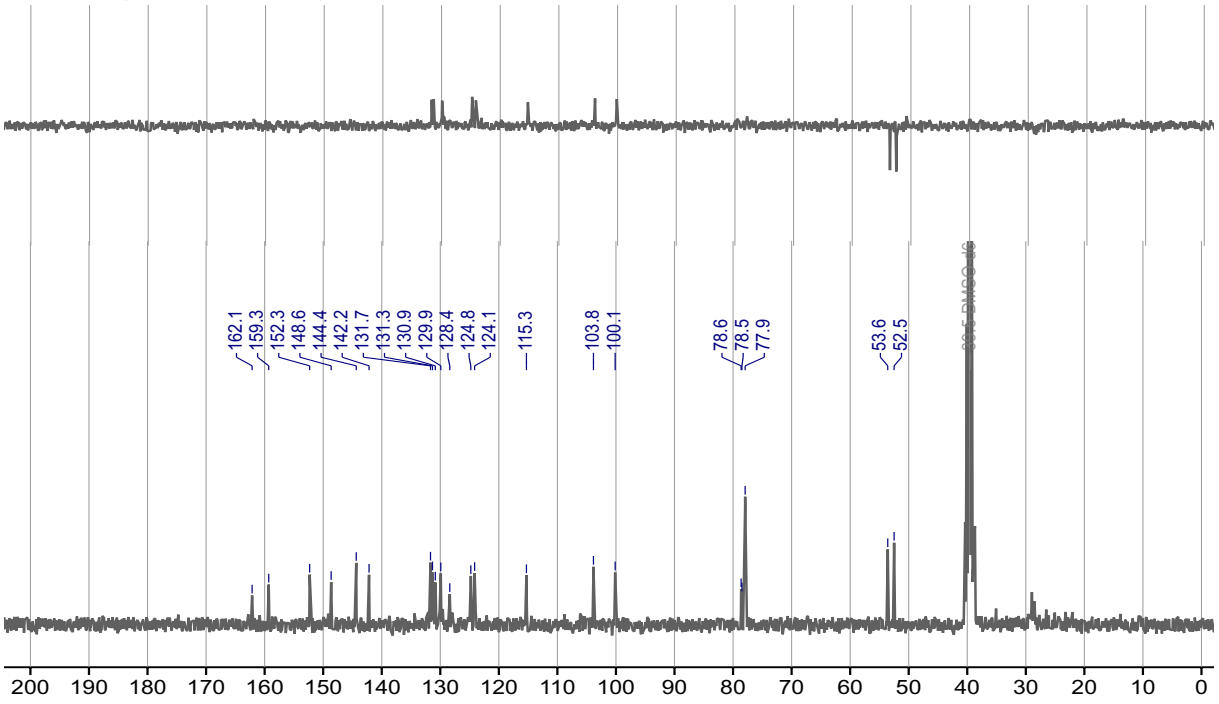
Figure 104. ^{1}H and ^{13}C NMR of (**5**) in $CDCl_3$/DMSO.

*Propargylated probe (**5**) NMR characterization:*

^{1}H NMR (300 MHz, DMSO)



^{13}C NMR (75 MHz, DMSO)



^{1}H NMR (300 MHz, DMSO-d^{6}) δ 10.46 (s, 1H), 8.58 (d, *J* = 9.3 Hz, 1H), 8.30 (d, *J* = 9.4 Hz, 1H), 7.93 – 7.67 (m, 3H), 7.58 (s, 1H), 7.42 (d, *J* = 8.7 Hz, 1H), 4.87 (dd, *J* = 26.4, 2.4 Hz, 4H), 3.63 (dt, *J* = 15.6, 2.4 Hz, 2H). **^{13}C NMR** (75 MHz, DMSO-d^{6}) δ 162.1, 159.3, 152.3, 148.6, 144.3, 142.2, 131.7, 131.3, 130.9, 129.9, 128.4, 124.8, 124.1, 115.3, 103.8, 100.1, 78.6, 78.5, 77.9, 53.6, 52.5.

Figure 105. ^{1}H and ^{13}C NMR of (**5**) in DMSO.

7.3.3. Experimental Protocols

Circular Dichroisms. The mixtures were allowed to stand for 15 min before registering the spectra. The final spectra are the average of 5 scans and were processed using the "*smooth*" macro implemented in the program *Kaleidagraph* (v 3.5 by Synergy Software). Samples contained 10 mM phosphate buffer pH 7.5 and 100 mM of NaCl, 5 μM peptide and 1 equiv. of $PdCl_2$(COD) (when present).

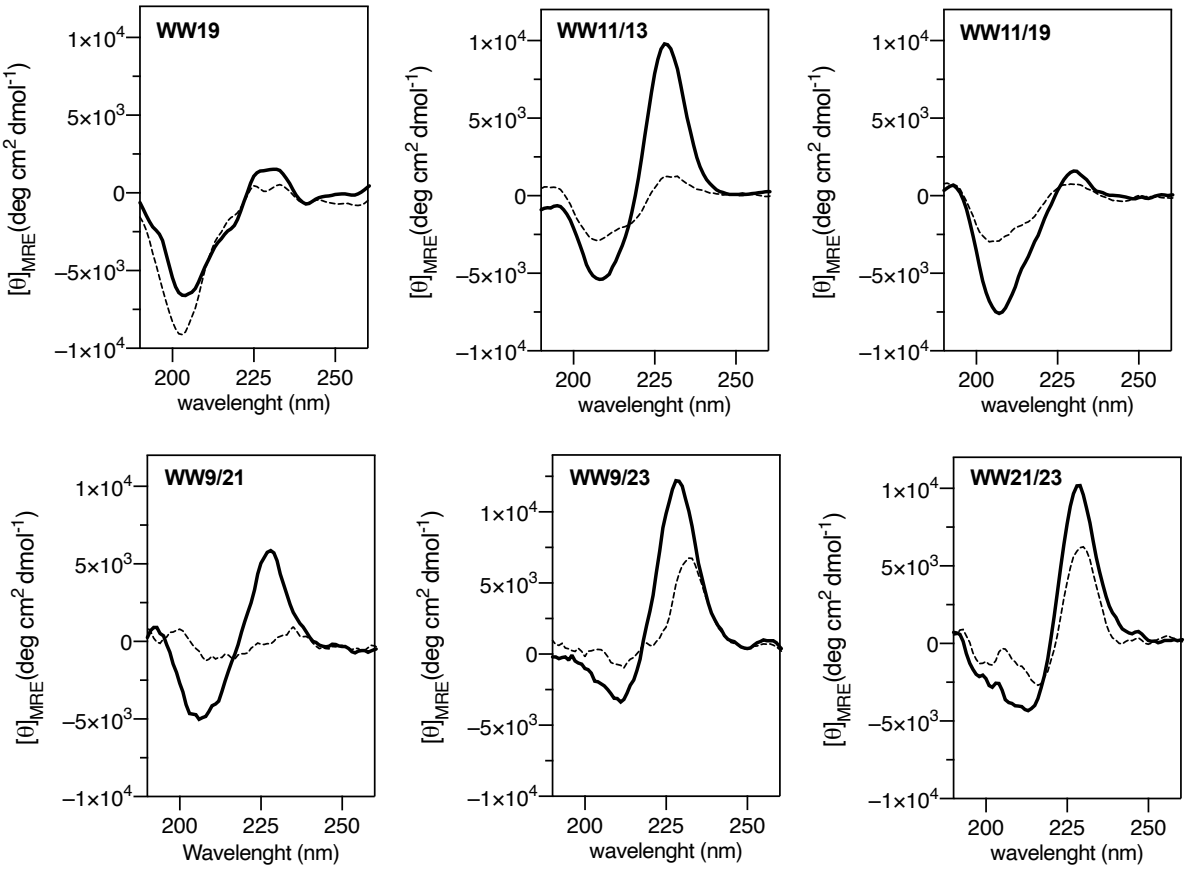


Figure 106. Circular dichroism spectra of 5 μM solutions of the WW peptides, before (thick solid lines), and after incubation with 1 equiv. of $PdCl_2$(COD) (dashed lines) and formation of their corresponding coordination complexes. Measurements were made in a 2 mm cell at 25 ºC.

NMR experiments. The experiments were recorded on a *Bruker Avance III* 600-MHz spectrometer (IRB Barcelona) equipped with a quadruple (^{1}H, ^{13}C, ^{15}N, ^{31}P) resonance cryogenic probe head and a z-pulse field gradient unit at 298 K using a 1mM solution of each peptide—in the presence or absence of $Pd(COD)Cl_2$—^{1}D proton spectra were recorded with a sweep width of 12000 Hz and 32 k data points. A total of 16 scans were accumulated with an acquisition time of 2.05s. A Watergate w5 composite pulse was used to suppress the water signal.

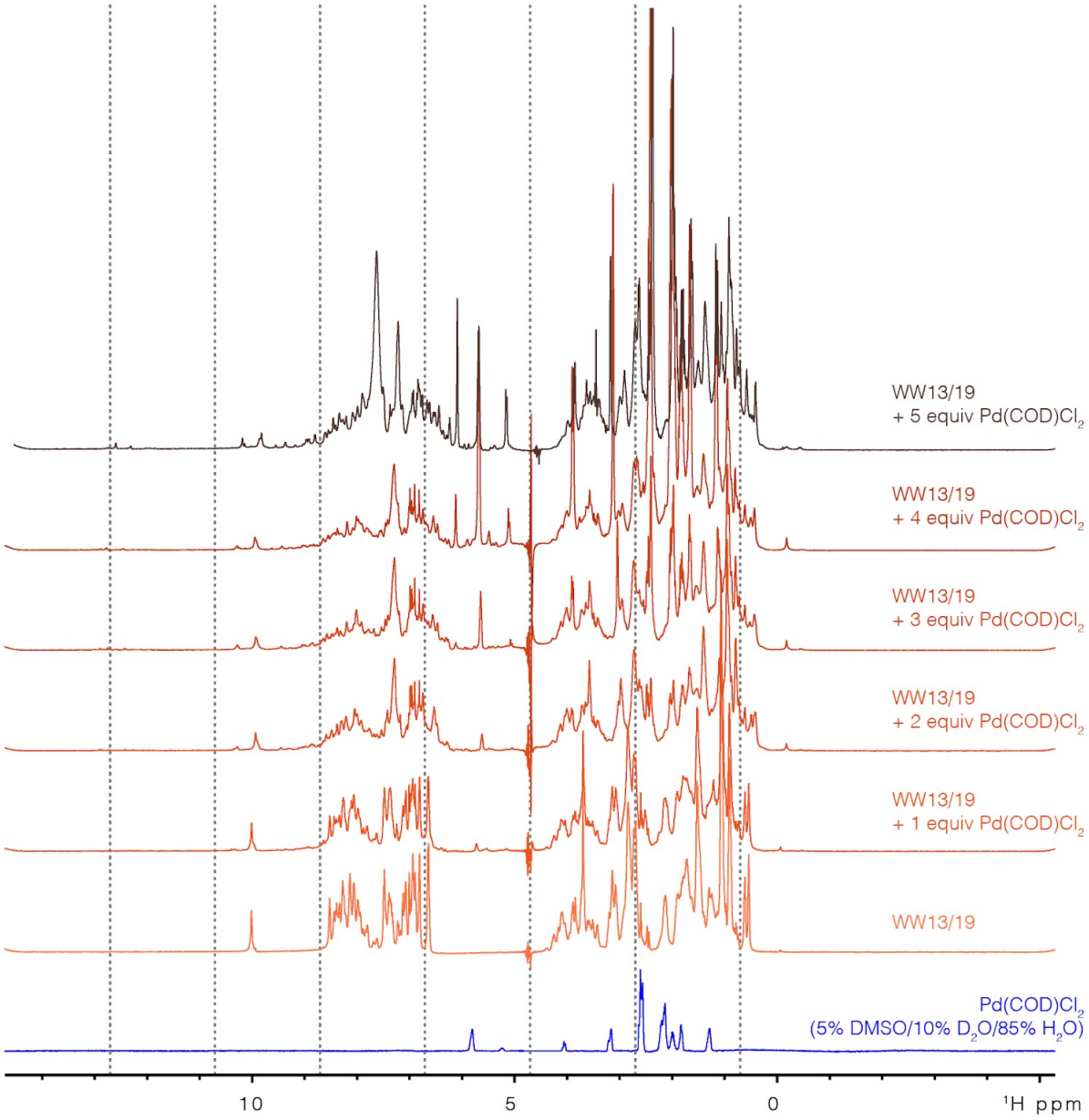


Figure 107. Titration of WW13/19 with up to five equivalents of [$Pd(COD)Cl_2$]. Progressive folding of the peptide can be followed by Trp Hε peaks at ~10 ppm and general dispersion of the amide region (~ 6-9ppm).

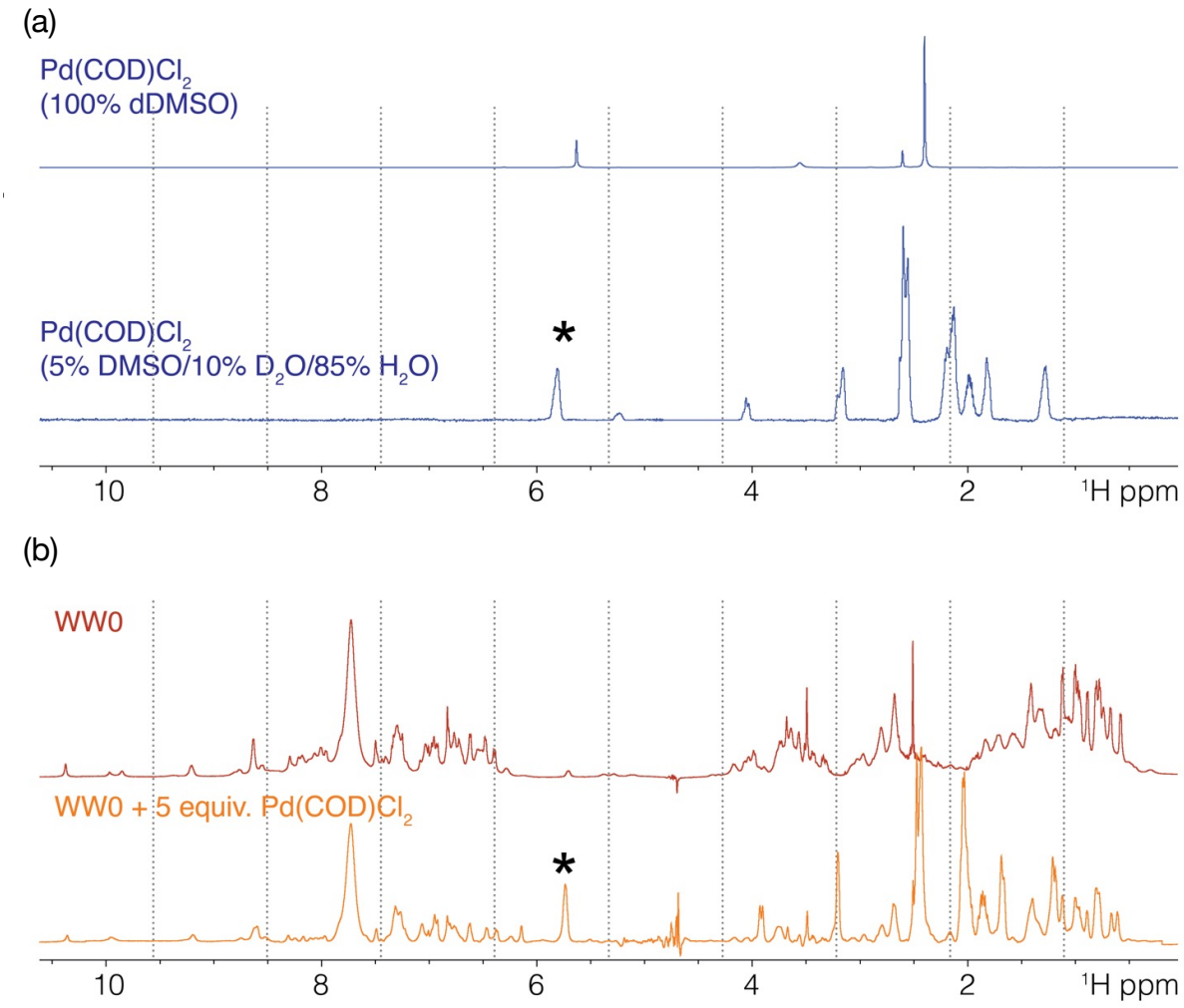


Figure 108. (a) [Pd(COD)Cl_2] ^{1}D NMR spectra in 100% dDMSO and 5% dDMSO/10% D_2O/85% H_2O. (b) WW0 ^{1}D NMR spectrum in the absence (TOP) and presence (BOTTOM) of 5 equivalents of [Pd(COD)Cl_2]. Asterisk shows the unbound [Pd(COD)Cl_2].

***In vitro* catalytic studies.** Catalytic deprotection of (**1**) to release the unencapsulated product **2** was performed in a 1.5 mL Eppendorf. For this purpose, a fresh solution of **1** (20 mM in DMSO, 1.0 equiv.) was added to PBS and to the resulting mixture a solution of the palladium complex (1 mM in PBS, 0.1 equiv.) was added. The reaction mixture was kept for 24 h at 37 °C with stirring at 800 rpm in a thermoshaker. After this time, 100 µL of the reaction was taken, diluted to 400 µL with H_2O and 50 µM coumarin and analyzed by reverse phase HPLC-MS. The results were processed according to the calibration curve, in which coumarin was used as internal standard. Each value is the average of three independent measurements. In the case of using ultradiafiltration with 0.5 mL Amicon® Ultra centrifugal filters (3k device) to remove the remaining free [Pd(COD)Cl_2], the experiment is the same as described above, but before the reactions and after coordination, we filtered the crude peptides and metal ion with the filters (x3 times) in the same PBS buffer. The same procedure was followed with the other two probes.

7.3.4. Biological Experiments

Cell internalization studies. Cells were seeded on glass-bottom plates 48 h before treatment. Culture medium was removed and DMEM containing 5% fetal bovine serum (FBS-DMEM) and peptides (5 μM) or metallopeptides (5 μM) were added. Before the addition to cells, peptides were pre-incubated with metal complexes (1:1) in PBS for 15 min. After 30 min, cells were washed twice with PBS and replace with fresh FBS-DMEM to observe under the microscope with appropriate filters. Digital pictures of the different samples were taken under identical conditions of gain and exposure.

Intracellular reactions. HeLa cells were seeded on glass-bottom plates 48 h before treatment. Culture medium was removed and FBS-DMEM containing the indicated probe (50 μM) was added. After 1 hour incubation, cells were washed twice with FBS-DMEM and a solution of metal complexes or metallopeptides (peptides were preincubated with metal complexes (1:1) in PBS for 15 min before the addition to cells) in FBS-DMEM were added. After 1 h incubation, cells were washed twice in PBS and replace with fresh FBS-DMEM to observe under the microscope with appropriate filters. Digital pictures of the different samples were taken under identical conditions of gain and exposure.

7.4. Identifying Metallopeptides for Intracellular Catalysis Using Combinatorial Libraries

7.4.1. SPOT Libraries

SPOT peptide libraries and CelluSPOT slides were synthesized *by Intavis Peptide Services GmbH*, Waldhäuser Straße 64, 72076 Tübingen, Germany, following the procedures described in the literature.

The following protocol was followed to perform the SPOT library screening: The CelluSPOT slide was incubated for 1 h with shaking with 10 mL of a 1 mM dichloro(1,5-cyclooctadiene) palladium(II), [Pd(COD)Cl_2] or copper bromide, $CuBr_2$, solution for click. The slide is washed 3 times with PBS for 10 min per wash to remove any uncoordinated palladium. For copper catalysis, a previous step is done by incubating with 10 mL of a 10 mM solution of sodium ascrobate, NaAsc. Next, 50 µL of a 200 µM solution of probe **1** and for the click 200 µM of azide **7** and 400 µM of propargyl **8** are added, the slide is covered with a coverslip and the edges are sealed with cytosyl. The slide is brought to the microscope and the emission of the spots at 0, 4 and 24 h is recorded. The plate was kept in a humid atmosphere while the studies were performed under the microscope to prevent drying.

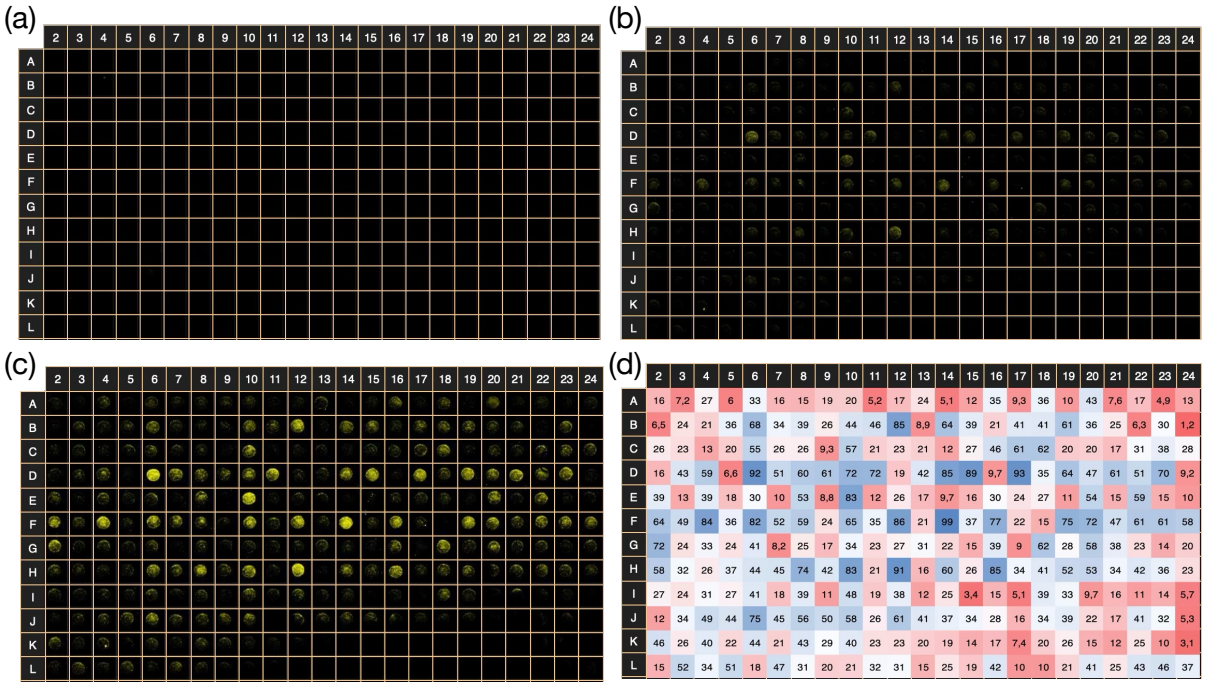


Figure 109. Confocal microscopy images of CelluSPOT slide after (a) 0 h; (b) 4 h and (c) 24 h of incubation with probe 1; (d) representation of its relative catalytic activity, quantified by the emission intensity of each point over time after 24 h of incubation.

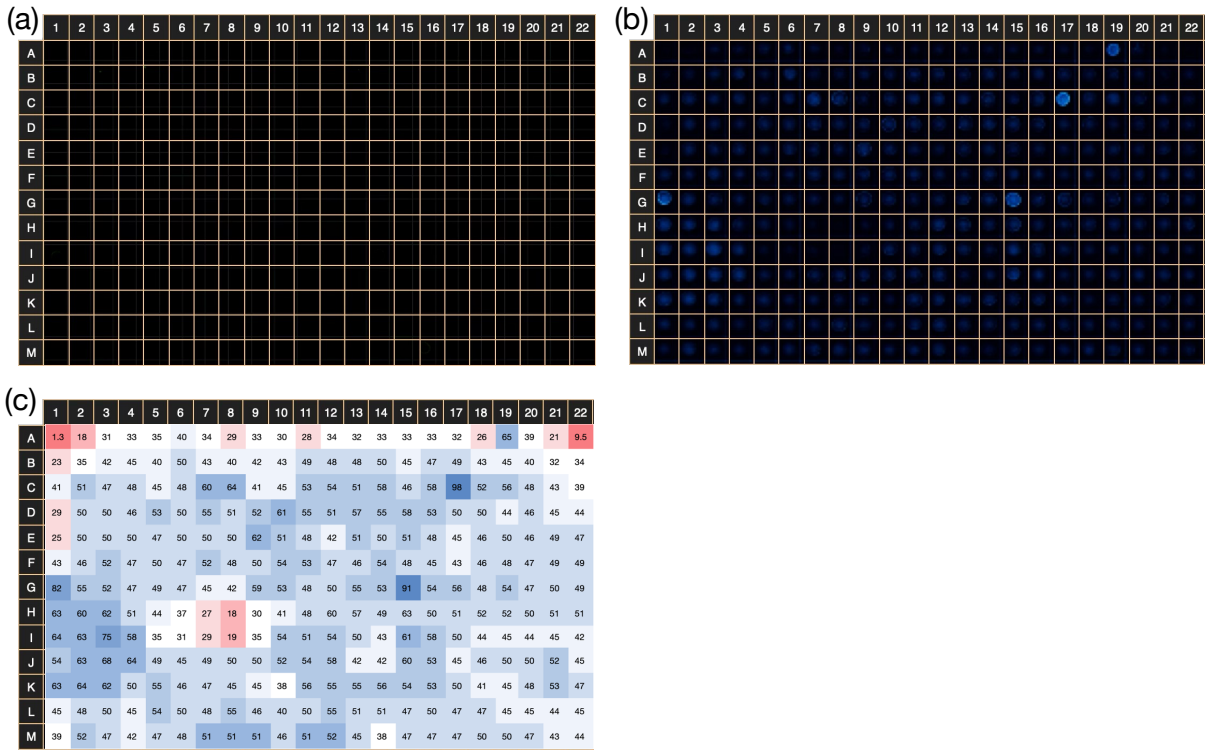


Figure 110. Confocal microscopy images of CelluSPOT slide after (a) 0 h and (b) 4 h of incubation with azide 7 and propgargyl 8; (d) representation of its relative catalytic activity, quantified by the emission intensity of each point over time after 4 h of incubation.

7.4.2. Synthesis and Characterization of Peptides

Table 6. Name and sequence of the peptides used in this study. All peptides are N-terminal free amine, C-terminal amide. In grey indicated the WZip1 origin sequence.

Peptide	Sequence
WZip1	SWTWE GNKWTWK
D14	AWHWR GNVWHWT
D4	RWHWA GNVWHWT
F14	AWRWV GNTWHWH
F13	AWRWT GNVWHWH
E10	TWHWA GNVWRWH
E9	TWHWA GNRWVWH
C17	YWHWV GNRWHWH
C10	AWHWY GNVWHWH

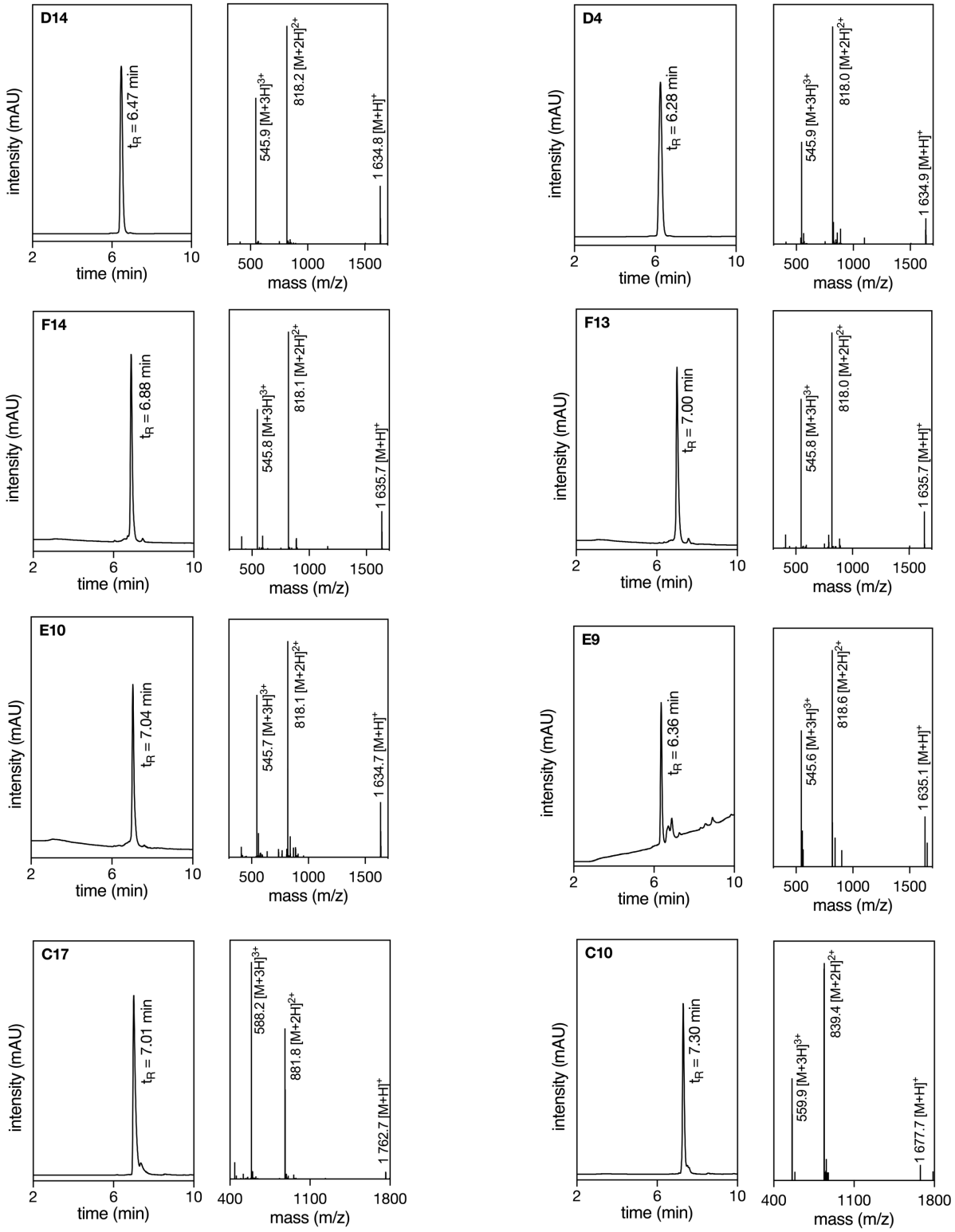


Figure 111. HPLC traces of the purified SPOT peptides used in this study and their corresponding ESI-MS spectra showing in each case peaks matching the expected masses.

D14 HPLC-MS (ESI) (m/z, t_R = 6.47 min): calculated for $C_{80}H_{99}N_{25}O_{14}$ $[M+H]^+$ = 1 634.8; found $[M+H]^+$ = 1 634.8, $[M+2H]^{2+}$ = 818.2, $[M+3H]^{3+}$ = 545.9. **D4** HPLC-MS (ESI) (m/z, t_R = 6.28 min): calculated for $C_{80}H_{99}N_{25}O_{14}$ $[M+H]^+$ = 1 634.8; found $[M+H]^+$ = 1 634.9, $[M+2H]^{2+}$ = 818.0, $[M+3H]^{3+}$ = 545.9. **F14** HPLC-MS (ESI) (m/z, t_R = 6.88 min): calculated for $C_{80}H_{99}N_{25}O_{14}$ $[M+H]^+$ = 1 634.8; found $[M+H]^+$ = 1 635.7, $[M+2H]^{2+}$ = 818.1, $[M+3H]^{3+}$ = 545.8. **F13** HPLC-MS (ESI) (m/z, t_R = 7.00 min): calculated for $C_{80}H_{99}N_{25}O_{14}$ $[M+H]^+$ = 1 634.8; found $[M+H]^+$ = 1 635.7, $[M+2H]^{2+}$ = 818.0, $[M+3H]^{3+}$ = 545.8. **E10** HPLC-MS (ESI) (m/z, t_R = 7.04 min): calculated for $C_{80}H_{99}N_{25}O_{14}$ $[M+H]^+$ = 1 634.8; found $[M+H]^+$ = 1 634.7, $[M+2H]^{2+}$ = 818.1, $[M+3H]^{3+}$ = 545.7. **E9** HPLC-MS (ESI) (m/z, t_R = 6.36 min): calculated for $C_{80}H_{99}N_{25}O_{14}$ $[M+H]^+$ = 1 634.8; found $[M+H]^+$ = 1 635.1, $[M+2H]^{2+}$ = 818.6, $[M+3H]^{3+}$ = 545.6. **C17** HPLC-MS (ESI) (m/z, t_R = 7.01 min): calculated for $C_{88}H_{103}N_{27}O_{14}$ $[M+H]^+$ = 1 763.0; found $[M+H]^+$ = 1 762.7, $[M+2H]^{2+}$ = 881.8, $[M+3H]^{3+}$ = 588.2. **C10** HPLC-MS (ESI) (m/z, t_R = 7.30 min): calculated for $C_{85}H_{96}N_{24}O_{14}$ $[M+H]^+$ = 1 677.9; found $[M+H]^+$ = 1 677.7, $[M+2H]^{2+}$ = 839.4, $[M+3H]^{3+}$ = 559.9.

Synthesis of labeled TMR-D14, TMR-D4 and TMR-C17. Once the peptides were fully assembled and still attached to the solid support, we coupled Fmoc-6-aminohexanoic acid (Fmoc-Ahx-OH) as spacer using standard solid-phase methods (activation with 1.0 equiv. of HATU in 2 mL DMF with 3 mL 0.195 M DIEA/DMF). The terminal Fmoc protecting group was removed (20% piperidine/DMF) and a mixture of 5-carboxytetramethylrhodamine (TMR-OH, 3.0 equiv., 0.15 mmol, 64.5 mg), 3 equiv. of HATU and 5 equiv. of DIEA 0.195 M in DMF was added onto the resin and shaken for 60 min.

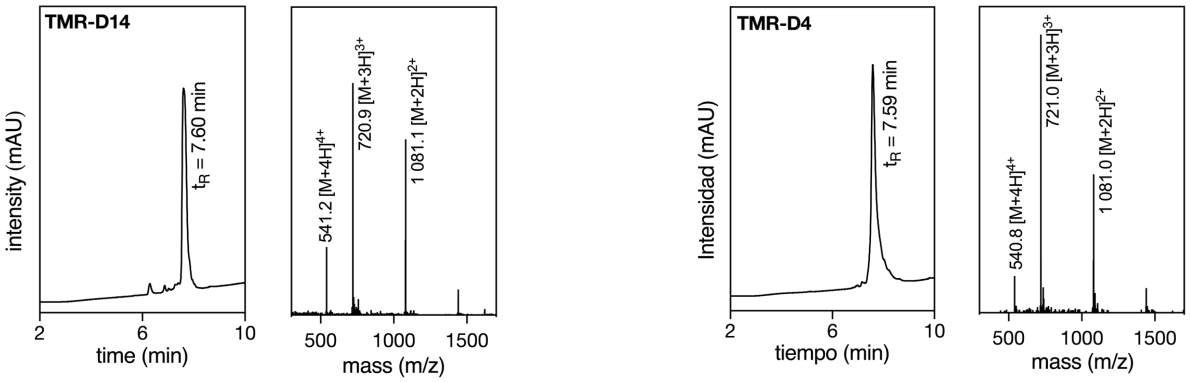


Figure 112. HPLC traces of the purified TMR labeled peptides D14 and D4 and their corresponding ESI-MS spectra showing in each case peaks matching the expected masses.

TMR-D14 HPLC-MS (ESI) (m/z, t_R = 7.60 min): calculated for $C_{111}H_{132}N_{28}O_{19}$: 2 161.0, found: $[M+2H]^{2+}$ = 1 081.1, $[M+3H]^{3+}$ = 720.9, $[M+4H]^{4+}$ = 541.2. **TMR-D4** HPLC-MS (ESI) (m/z, t_R = 7.59 min): calculated for $C_{111}H_{132}N_{28}O_{19}$: 2 161.0, found: $[M+2H]^{2+}$ = 1 081.0, $[M+3H]^{3+}$ = 721.0, $[M+4H]^{4+}$ = 540.8.

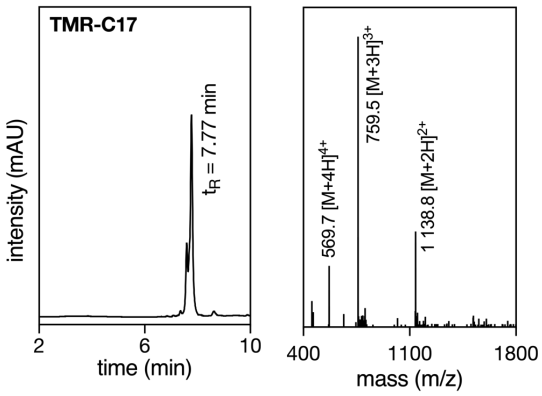


Figure 113. HPLC traces of the purified TMR labeled peptide C17 and its corresponding ESI-MS spectra showing in each case peaks matching the expected masses.

TMR-C17 HPLC-MS (ESI) (m/z, t_R = 7.77 min): calculated for $C_{133}H_{128}N_{30}O_{20}$: 2 275.6, found: $[M+2H]^{2+}$ = 1 138.8, $[M+3H]^{3+}$ = 759.5, $[M+4H]^{4+}$ = 569.7.

Methal coordinated peptides. Peptides (1 mM) were mixed with equimolar amounts of $[Pd(COD)Cl_2]$ or $CuBr_2$ in PBS and the resulting mixtures were analyzed (after 15 min. palladium complexes and 1 h for copper complexes) by Electrospray Ionization Mass Spectrometry (ESI/MS) with an *Agilent 6120 Quadrupole* LC/MS model in positive scan mode using direct injection of the purified peptide solution into the MS detector. The peaks corresponding to the apo-peptides are shown in black, and those corresponding to the complexes are in colors.

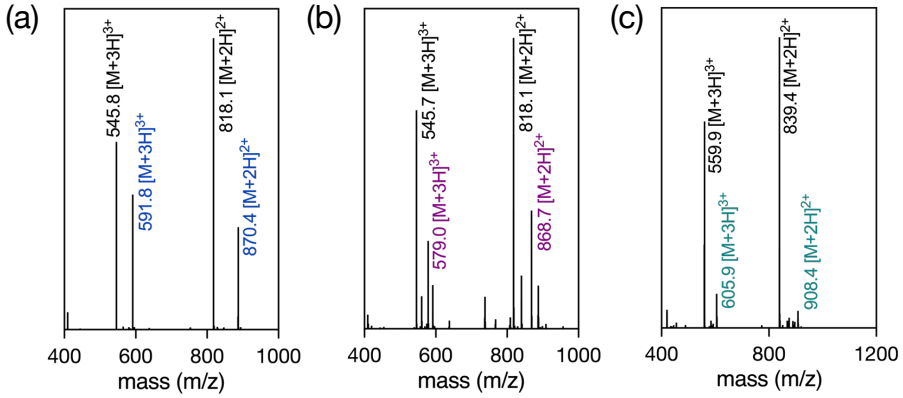


Figure 114. ESI-MS spectra of (a) the mixture of F14 (1 mM) with 1.0 equiv. of $Pd(COD)Cl_2$ in PBS after 15 min incubation; (b) the mixture of E10 (1 mM) with 1.0 equiv. of $Pd(COD)Cl_2$ in PBS after 15 min incubation and (c) the mixture of C10 (1 mM) with 1.0 equiv. of $CuBr_2$ in PBS after 1 h incubation. The peaks corresponding to the apo-peptides are labeled in black, those of the corresponding complex are in colors.

7.4.3. Synthesis and Characterization of Probes

Synthesis of (9)

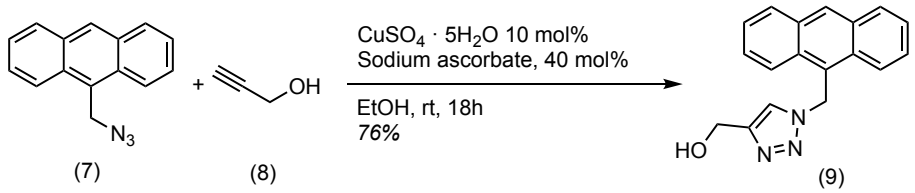
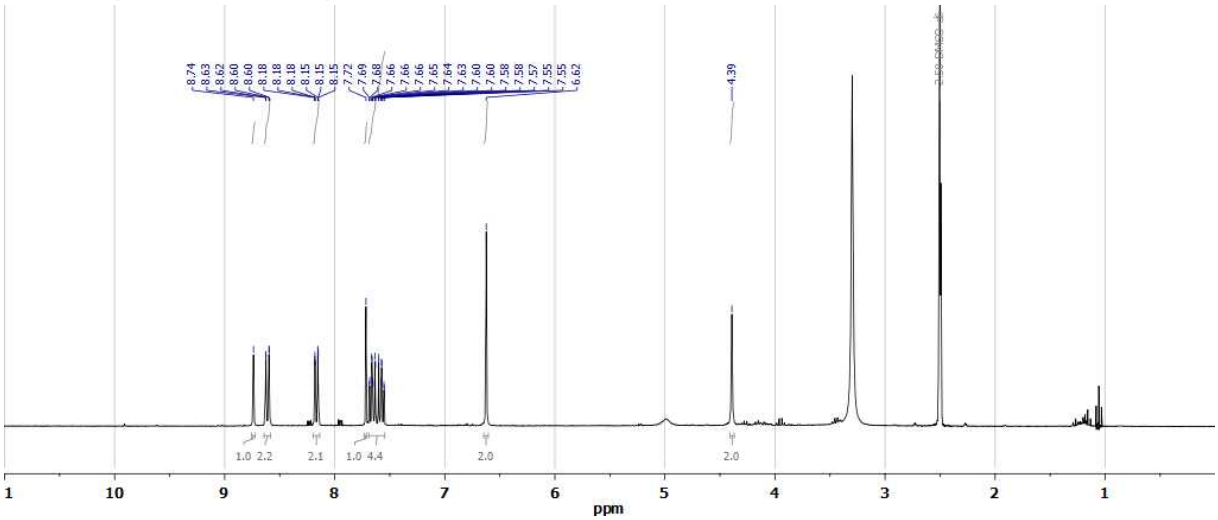


Figure 115. Synthetic route to obtain probe (**9**).

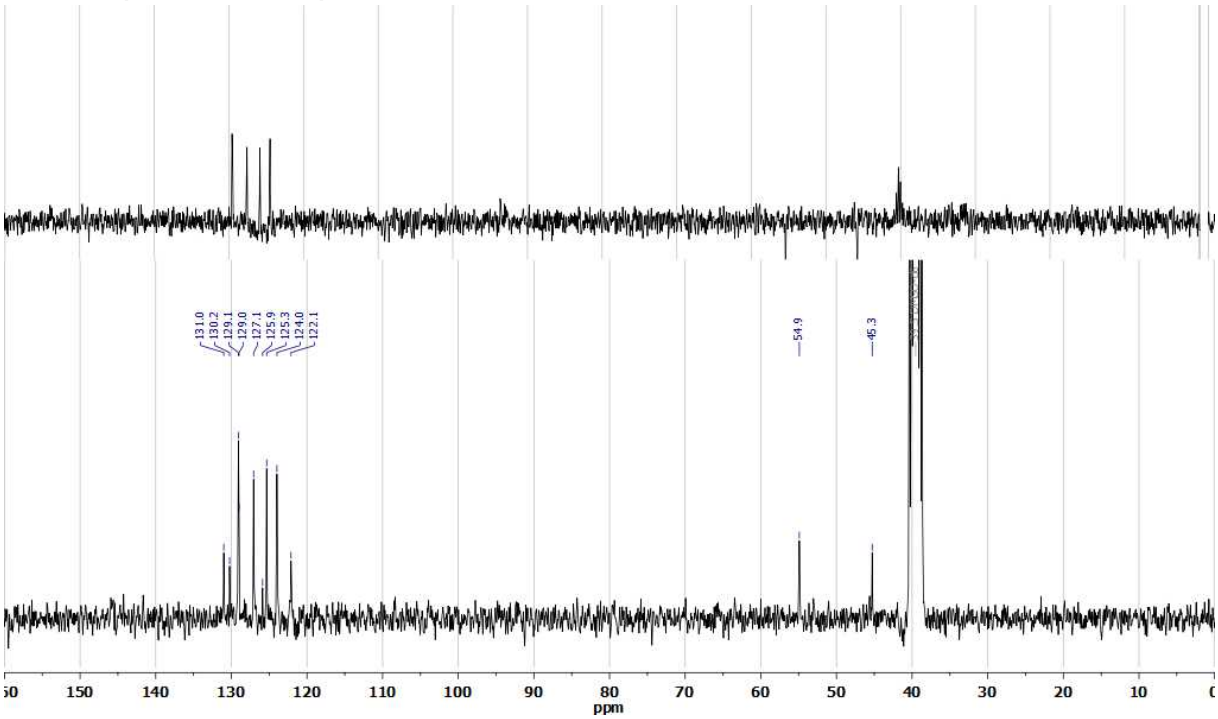
*(1-(anthracen-9-ylmethyl)-1H-1,2,3-triazol-4-yl)methanol (**9**).* $CuSO_4$ (12.2 mg, 0.05 mmol, 0.1 equiv.) in 10 mL of EtOH was added into a round-bottom flask. AscNa (29.6 mg, 0.2 mmol, 0.4 equiv.) was added and the reaction was stirred at room temperature for 10 min. After that time, 9-(azidomethyl)anthracene (7) (116.5 mg, 0.5 mmol, 1.0 equiv.) and propargyl alcohol (8) (43 µL, 0.75 mmol, 1.5 equiv.) were added and the reaction was stirred at room temperature overnight. The solvent was removed in vacuo and compound 9 was obtained as a yellow solid (110.2 mg, 76%). ^{1}H NMR (300 MHz, DMSO-d^{6}) δ 8.74 (s, 1H), 8.64 – 8.58 (m, 2H), 8.20 – 8.14 (m, 2H), 7.72 (s, 1H), 7.69 – 7.54 (m, 4H), 6.62 (s, 2H), 4.39 (s, 2H). ^{13}C NMR (75 MHz, DMSO-d^{6}) δ 131.0, 130.2, 129.1, 129.0, 127.1, 125.9, 125.3, 124.0, 122.1, 54.9, 45.3.

*Propargylated probe (**9**) NMR characterization:*

^{1}H NMR (300 MHz, DMSO)



^{13}C NMR (75 MHz, DMSO)



^{1}H NMR (300 MHz, DMSO-d^{6}) δ 8.74 (s, 1H), 8.64 – 8.58 (m, 2H), 8.20 – 8.14 (m, 2H), 7.72 (s, 1H), 7.69 – 7.54 (m, 4H), 6.62 (s, 2H), 4.39 (s, 2H). **^{13}C NMR** (75 MHz, DMSO-d^{6}) δ 131.0, 130.2, 129.1, 129.0, 127.1, 125.9, 125.3, 124.0, 122.1, 54.9, 45.3.

Figure 116. ^{1}H and ^{13}C NMR of (**9**) in DMSO.

7.4.4. Experimental Protocols

Circular Dichroisms. The mixtures were allowed to stand for 15 min before registering the spectra. The final spectra are the average of 5 scans and were processed using the "*smooth*" macro implemented in the program *Kaleidagraph* (v 3.5 by Synergy Software). Samples contained 10 mM phosphate buffer pH 7.5 and 100 mM of NaCl, 5 μM peptide and 1 equiv. of $PdCl_2$(COD) (when present).

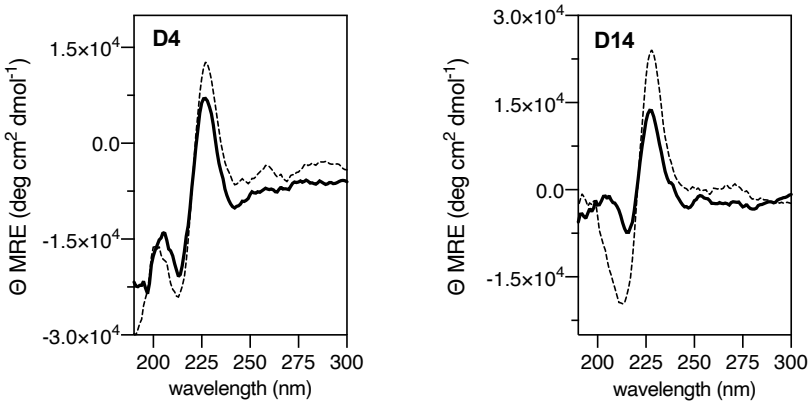


Figure 117. Left: CD spectra of 5 μM **D4** (dashed line) and 5 μM **D4** in the presence of and 1 equiv. of $PdCl_2$(COD) (thick solid line). Right: CD spectra of 5 μM **D14** (dashed line) and 5 μM **D14** in the presence of and 1 equiv. of $PdCl_2$(COD) (thick solid line). Measurements in 10 mM phosphate buffer pH 7.5 and 100 mM of NaCl using a 2 mm cell at 4 ºC.

***In vitro* catalytic studies.** Catalytic deprotection of (**1**) to release the unencapsulated product **2** was performed in a 1.5 mL Eppendorf. For this purpose, a fresh solution of **1** (20 mM in DMSO, 1.0 equiv.) was added to PBS and to the resulting mixture a solution of the palladium complex (1 mM in PBS, 0.1 equiv.) was added. The reaction mixture was kept for 24 h at 37 °C with stirring at 800 rpm in a thermoshaker. After this time, 100 μL of the reaction was taken, diluted to 400 μL with H_2O and 50 μM coumarin and analyzed by reverse phase HPLC-MS. The results were processed according to the calibration curve, in which coumarin was used as internal standard. Each value is the average of three independent measurements.

For the click reactions, an equivalent protocol was followed. To do this, a fresh solution of 7 (20 mM in DMSO, 1.0 equiv.) and 8 (20 mM in DMSO, 2.0 equiv.) was added to PBS and, to the resulting mixture, a solution of the copper complex (1 mM in PBS, 0.1 equiv.), previously incubated for 30 min with a 10 mM solution of NaAsc. The reaction mixture was kept for 24 h at 37 °C with shaking at 800 rpm in a thermoshaker. After this time, 100 μL of the reaction was taken, diluted to 400 μL with H_2O and 50 μM coumarin and analyzed by reversed-phase HPLC-MS. The results were processed according to the calibration curve, in which coumarin was used as an internal standard. Each value is the average of three independent measurements.

Intracellular Reactions. HeLa cells were seeded in glass-bottom dishes 48 h before treatment. The culture medium was removed, and FBS-DMEM containing probe 1 (50 μM) or compounds 7 (100 μM) and 8 (200 μM) was added. After 1 h of incubation, cells were washed twice with FBS-DMEM, and a solution of metal complexes or metallopeptides (peptides were preincubated with metal complexes (1:1) in water for 10 min before adding them to the cells; copper complexes were incubated for 30 min with AscNa) in FBS-DMEM was added. After 1 h of incubation, cells were washed twice with PBS and replaced with fresh FBS-DMEM for microscopic observation using appropriate filters. Digital images of the different samples were taken under identical gain and exposure conditions.

8. BIBLIOGRAPHY

1. Ostler, E. L. *Chem. Cent. J.* **1**, 5 (2007).
2. Wöhler, F. *Ann. Phys.* **88**, 253–256 (1828).
3. Miescher, F. *Medicinisch-chemische Untersuchungen*, 4º edition (1871)
4. Musumeci, G. *J. Histol. Histopathol.* **1**, 5 (2014).
5. Drews, J. *Nat. Rev. Drug Disco.* **3**, 797–801 (2004).
6. L. E. Kay, The Molecular Vision of Life. Caltech, *The Rockefeller Foundation, and the Rise of the New Biology*, Oxford Univ. Press ISBN: 9780195111439 (pp. 225-242) and *@CaltechArchives* (1997).
7. *Eng. & Sci.* **17**, 9-13 (1954).
8. Beach, E. F. *J. Hist. Med. Allied. Sci.* **16**, 354–373 (1961).
9. Vickery, H. B., Schmidt, C. L. A. *Chem. Rev.* **9**, 169–318 (1931).
10. Hartley, H. *Nature* **168**, 244 (1951).
11. Kendrew, J. C. *et al. Nature* **181**, 662–666 (1958).
12. Šorm, F *et al. Collect. Czech. Chem. Commun.* **26**, 531-578 (1961).
13. (a) Voet, D., Voet, J. G., Pratt, C. W. *Fundamentals of Biochemistry,* John & Sons, New York (2007); (b) Alberts, B. *et al. Molecular Biology of the Cell*, Garland Publishing, New York, Chapter 3 (2008).
14. Martin, R. B. *Biopolymers* **45**, 351–353 (1998).
15. Nelson, D. L., Cox, M. M. *Lehninger Principles of Biochemistry*, W. H. Freeman Company, New York, Chapter 4 (2017).
16. Barlow, D. J., Thornton, J. M. *J. Mol. Biol.* **201**, 601–619 (1988).
17. (a) Branden, C., Tooze, J. *Introduction to protein structure* Garland Publishing Inc., New York (1999); (b) Sewald, N., Jakubke, H-D. *Peptides: Chemistry and Biology* Wiley-VCH GmbH & Co, Weinheim (2002).
18. Clarke, D. T. *et al. Proc. Natl. Acad. Sci. USA* **96**, 7232-7237 (1999).
19. Pace, C. N., Scholtz, J. M. *Biophys. J.* **75**, 422–427 (1998).
20. Rose, G. D. *Nature* **272**, 586–590 (1978).
21. Rose, G. D., Roy, S., *Proc. Natl. Acad. Sci. USA* **77**, 4643–4647 (1980).
22. Wang, W., Nema, S., Teagaden, *D. Int. J. Pharm*. **390**, 89–99 (2010).
23. Fosgerau, K., Hoffmann, T. *Drug Discov. Today* **20**, 122–128 (2015).
24. A. Giannis, A., Kolter, T. *Angew. Chem. Int. Ed. Engl.* **32**, 1244–1267 (1993).
25. (a) Shepherd, N. E. *et al. J Am Chem Soc* **127**, 2974–2983 (2005); (b) Groß, A., Hashimoto, C., Sticht, H., Eichler, J. *Front. Bioeng. Biotechnol.* **3** (2016).
26. (a) Tung, C. H. *Biopolymers (Pept. Sci.)* **76**, 391–403 (2004); (b) Pazos, E. *et al. Chem. Soc. Rev.* **38**, 3348–3359 (2009).
27. (a) Maeda, Y. *et al. Annu. Rev. Biomed. Eng.* **18**, 311–328 (2016); (b) Carvalho, S. *et al. Isr. J. Chem.* **62**, e202200029 (2022)
28. Merrifield, R. B. *J. Am. Chem. Soc.* **85**, 2149–2154 (1963).
29. (a) Houghten, R. A. *Proc. Natl. Acad. Sci. USA* **82**, 5131–5135 (1985); (b) Pellois, J. P. *et al. Nat. Biotechnol.* **20**, 922–926 (2002); (c) Geysen, H. M., Meloen, R. H., Barteling, S. J. *Proc. Natl. Acad. Sci. USA* **81**, 3998–4002 (1984); (d) Frank, R. *Tetrahedron* **48**, 9217–9232 (1992).
30. Kramer, A., Schneider-Mergener, J. *Methods Mol. Biol.* **87**, 25–39 (1998).
31. (a) Hilpert, K., Winkler, D. F. H., Hancock, R. E. W. *Nat. Protoc.* **2**, 6, 1333–1349 (2007); (b) Hilper, K., Winkler, D. F. H., Hancock, R. E. W. *Biotechnol. Genet. Eng. Rev.* **24**, 31–106 (2007); (c) Reineke, U., Volkmer-Engert, R., Schneider-Mergener, J. *Curr. Opin. Biotechnol.* **12**, 1, 59–64 (2001).

32. Frank, R. *et al.* Combinatorial Synthesis on Membrane Supports by the SPOT Technique: Imaging Peptide Sequence and Shape Space. In *Combinatorial Peptide and Nonpeptide Libraries*; Wiley-VCH Verlag GmbH: Weinheim, Germany, pp 363–386 (2007).
33. Toepert, F. *et al. Angew. Chem. Int. Ed. Engl.* **42**, 1136–1140 (2003).
34. (a) Heine, N. *et al. Tetrahedron* **59**, 9919–9930 (2003); (b) Blackwell, H. E. *Curr. Opin. Chem. Biol.* **10**, 203–212 (2006).
35. Kramer, A. *et al. Cell* **91**, 799–809 (1997).
36. Hilpert, K. *et al. J. Biochem.* (Tokyo) **128**, 1051–1057 (2000).
37. Hilpert, K. *et al. Nat. Biotechnol.* **23**, 1008–1012 (2005).
38. Kamradt, T., Volkmer-Engert, R. *Mol. Divers* **8**, 271–280 (2004).
39. Wiedemann, E. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **28**, 175, 493–495 (1889).
40. Braslavsky, S. E. *et al. Pure App.l Chem.* **79**, 293–465 (2007).
41. Valeur B., Brochon, J. New Trends in Fluorescence Spectroscopy: Applications to Chemical and Life Sciences, *Springer* (2001).
42. (a) Specht, E. A., Braselmann, E., Palmer, A. E. A. *Annu. Rev. Physiol.* **79**, 93–117 (2017); (b) Lavis, L. D., Raines, R. T. *ACS Chem. Biol.* **9**, 855–866 (2014).
43. (a) Lakowicz, J. R. Principles of Fluorescence Spectroscopy, 3rd edition. *Springer US* (2006); (b) Park, S. J. *et al*. *Stem Cell Res.* **12**, 730–741 (2014); (c) Chan, J., Dodani, S. C., Chang, C. J. *Nature Chem* **4**, 973–984 (2012).
44. Hwang, J., Nagaraju, P., Cho, M. J., Choi, D. H. *Aggregate* **4** (2023).
45. Zhou, P., Han, K. *Aggregate* **3** (2022).
46. Xu, S., Duan, Y., Liu, B. *Adv. Mater.* **32** (2020).
47. Herschel, Sir J. F. W. *Philos. Trans. R. Soc. Lond.* **135**, 143–145 (1845).
48. Baeyer, A. *Ber. Dtsch. Chem. Ges.* **4**, 555–558 (1871).
49. Ceresole, M. *US Patent* **377**, 349 (1888).
50. Loudet, A., Burgess, K. *Chem. Rev.* **107**, 4891–4932 (2007).
51. Luo, S. *et al. Biomaterials* **32**, 7127–7138 (2011).
52. Coons, A. H., Creech, H. J., Jones, R. N. *Proc. Soc. Exp. Biol. Med.* **47**, 200–202 (1941).
53. Naghibi, S., Chen, T., Jamshidi Ghahfarokhi, A., Tang, Y. *Aggregate* **2** (2021).
54. Zhang, Z. *et al*. *Luminescence* **32**, 1488–1493 (2017).
55. (a) Deng, X. *et al*. *Aggregate* **2**, e35 (2021); (b) Mei, J. *et al. Chem. Rev.* **115**, 11718–11940 (2015).
56. (a) Lavis, L. D., Raines, R. T. *ACS Chem. Biol.* **3**, 142–155 (2008); (b) Grimm, J. B., Lavis, L. D. *Nat. Methods* **19**, 149–158 (2022); (c) Liu, X., Zhu, C., Tang, B. Z, *Acc. Chem. Res.* **55**, 197–208 (2022).
57. Sener, M. *et al. ChemPhysChem* **12**, 518–531 (2011).
58. Algar, W. R. *et al. Coord Chem Rev* **263**, 65–85 (2014).
59. (a) Algar, W. R. *et al*. *Nat. Methods* **16**, 815–829 (2019); (b) Bakar, B. T. *et al. Sensors (Basel)* **16**, 1488 (2016); (c) Hochreiter, B., Garcia, A. P., Schmid, J. A. *Sensors (Basel)* **15**, 26281–26314 (2015).
60. (a) Selvin, P. R. *Annu. Rev. Biophys. Biomol. Struct.* **31**, 275-302 (2002); (b) Chen, J., Selvin, P. R. *Bioconjugate Chem.* **10**, 311-315 (1999).
61. (a) Debye, P.; Edwards, J. O. *Science* **116**, 143–144 (1952); (b) Wetlaufer, D. B. *Adv. Protein Chem.* **17**, 303–390 (1963); (c) Teale, F. W. *Biochem. J.* **76**, 381–388 (1960).
62. (a) Bao, J., Tong, C., He, M., Zhang, H. *Luminescence* **39** (2024); (b) Morzan, U. N. *et al. J. Phys. Chem. B* **126** 7203–7211 (2022).

63. Hunt, H. D., Simpson, W. T. *Product Information Bulletin.* J. R. Leader and J. F. Gormley, ibid **2** (1953).
64. Chen, Y., Barkley, M. D. *Biochemistry* **37**, 9976–9982 (1998).
65. Shimomura, O., Johnson, F. H., Saiga, Y. *J Cell. Physiol.* **59**, 223–239 (1962).
66. Ormö, M. *et al*. *Science (1979)* **273**, 1392–1395 (1996).
67. (a) Chalfie, M. *et al. Science (1979)* **263**, 802–805 (1994); (b) Zimmer, M. *Chem. Rev.* **102**, 759–781 (2002); (c) Remington, S. J. *Protein. Sci.* **20**, 1509–1519 (2011).
68. Zhao, Z., Zhang, H., Lam, J. W. Y., Tang, B. Z. *Angewandte Chemie - International Edition* **59**, 9888–9907 (2020).
69. Heim, R., Cubitt, A. B., Tsien, R. Y. *Nature* **373**, 663–664 (1995).
70. McRae, S. R., Brown, C. L., Bushell, G. R. *Protein Expr. Purif.* **41**, 121–127 (2005).
71. Pédelacq, J. D. *et al. Nat. Biotechnol.* **24**, 79–88 (2006).
72. Pakhomov, A. A., Martynov, V. I. *Chem. Biol.* **15**, 755–764 (2008).
73. Coralli, C., Cemazar, M., Kanthou, C., Tozer, G. M. Dachs, G. U. *Cancer Res.* **61**, 4784–4790 (2001).
74. Balleza, E.,Kim, J. M., Cluzel, P. *Nat. Methods* **15**, 47–51 (2018).
75. Hink, M. A. *et al. J. Biol. Chem.* **275**, 17556–17560 (2000).
76. Swenson, E. S. *et al. Stem Cells* **25**, 2593–2600 (2007).
77. Ansari, A. M. *et al. Stem Cell Rev* **12**, 553–559 (2016).
78. (a) Wels, P. *Pflügers Arch.* **219**, 738–752 (1928); (b) Konev, S. V. *Fluorescence and Phosphorescence of Proteins and Nucleic Acids* (Plenum Press, New York, NY, 1967).
79. Bhattacharya, A. *et al. Langmuir* **33**, 10606–10615 (2017).
80. (a) Stagi, L. *et al. Macromol. Chem. Phys.* **222**, (2021); (b) Homchaudhuri, L., Swaminathan, R. *Chem. Lett.* **30**, 844–845 (2001); (c) Cadeddu, M. *et al*. *Macromolecules* **57**, 514–527 (2024).
81. Y. Gong, *et al. Sci. China Chem.* **56**, 1178–1182 (2013).
82. (a) Babar, D. G., Sarkar, S. *Appl. Nanosci. (Switzerland)* **7**, 101–107 (2017); (b) Fan, Z., Sun, L., Huang, Y., Wang, Y., Zhang, M. *Nat. Nanotechnol*. **11**, 388–394 (2016).
83. Sementa, D. *et al. Angew Chem. Int. Ed.* **62**, (2023).
84. Sharpe, S., Simonetti, K., Yau, J., Walsh, P. *Biomacromolecules* **12**, 1546–1555 (2011)
85. Amdursky, N. *et al. Nano Lett.* **9**, 3111–3115 (2009).
86. Wang, D.; Imae, T. *J. Am. Chem. Soc.* **126**, 13204–13205 (2004).
87. (a) Stephens, A. D. *et al*. *Proc. Natl. Acad. Sci. USA* **118**, e2020389118 (2021); (b) Mirón, G. D. *et al. Nat. Commun.* **14**, (2023).
88. Pinotsi, D. *et al*. *J. Am. Chem. Soc.* **138**, 3046–3057 (2016).
89. Alom, S. E., Swaminathan R. *Phys. Chem. Chem. Phys.* **25**, 16626–16642 (2023).
90. (a) Zhang, H. *et al. Materials Today* **32**, 275–292 (2020); (b) Zhang, H., Tang, B. Z. *JACS Au* **1**, 1805–1814 (2021).
91. (a) Tikhonova, T. N. *et al. Arch. Biochem. Biophys.* **651**, 13–20 (2018); (b) Yakimov, B. P. *et al. Environ. Sci. Technol. Lett.* **9**, 452–458 (2022).
92. (a) Banci, L. *Metallomics and the Cell* **12**, 1–13 (Springer, 2013); (b) Shriver, D. F., Atkins, P. W. *Inorganic Chemistry* (3rd ed.), Oxford University Press (1999).
93. Waldron, K. J., Robinson, N.J. *Nat. Rev. Microbiol.* **7**, 25–35 (2009).
94. Yamashita, M. M. et al. *Proc. Natl. Acad. Sci. USA* **87**, 5648–5652 (1990).
95. Messerschmidt, A. *et al. Handbook of Metalloproteins* (Wiley, 2001).
96. Andreini, C. *et al. J. Biol. Inorg. Chem.* **13**, 1205-1218 (2008).
97. Lu, Y. *et al. Nature* **460**, 855–862 (2009).
98. Bernhardt, R. *J. Biotechnol.* **124**, 128–145 (2006).

99. Perry, J. J. P. *et al. Biochim. Biophys. Acta* **1804**, 245–262 (2010).
100. Supuran, C. T. *Biochem. J.* **473**, 2023–2032 (2016).
101. (a) Nies, D. H., Silver, S. *Molecular Microbiology of Heavy Metals, Springer-Verlag*, Berlin (2007); (b) Zheng, H. *et al. J. Inorg. Biochem.* **102**, 1765–1776 (2008); (c) Waldron, K. J., Robinson, N. *J Nat. Rev. Microbiol.* **7**, 25–35 (2009); (d) Waldron, K. J. *et al. Nature* **460**, 823–830 (2009).
102. T. Yamanaka, K. Okunuki *Microbial Iron Metabolism* (Ed.: J.B. Neilands), Elsevier, 349–400, San Diego, CA, (1974).
103. Cammack, R. *Adv. Inorg. Chem. (Ed.: Cammack, R.), Acad. Press*, 281–322 (1992).
104. Penfield, K. W. *et al. J. Am. Chem. Soc.* **103**, 4382–4388 (1981).
105. Momenteau, M, Reed, C. A. *Chem. Rev.* **94**, 659–698 (1994).
106. (a) Stenkamp, R. E. *Chem. Rev.* **94**, 715–726 (1994); (b) Coates, C J., Decker, H., *Cell. Mo.l Life Sci.* **74**, 293–317 (2017).
107. Baker, E. N. *Adv. Inorg. Chem. (Ed.: Sykes, A. G.), Acad. Press*, 389–463 (1994).
108. Harrison, P. M. *et al. Adv. Inorg. Chem.* (*Ed.: Sykes, A. G.), Acad. Press*, 449–486 (1991).
109. Berg, J. M. *Annu. Rev. Biophys. Biophys. Chem.* **19**, 405–421 (1990).
110. (a) Andreini, C. *et al. J. Biol. Inorg. Chem.* **13**, 1205–1218 (2008); (b) Pordea, A. *Curr. Opin. Chem. Biol.* **25**, 124–132 (2015); (c) H. Eom, H., Song, W. J. *J. Biol. Inorg. Chem.* **24**, 517-531(2019); (d) Lewis, J. C. *ACS Catal* **3**, 1–22 (2013); (e) Valdez, C. E. *et al. Acc. Chem. Res.* **47**, 3110–3117 (2014).
111. Thomson, A. J., Gray, H. B. *Curr. Opin. Chem. Biol.* **2**, 155–158 (1998).
112. (a) Drauz, K., Waldmann, H. *Enzyme Catalysis in Organic Synthesis* (Wiley-VCH Verlag GmbH, 2002); (b) Fessner, W. D., Anthonsen, T. *Modern Biocatalysis* (Wiley-VCH Verlag GmbH & Co. KGaA, 2008).
113. (a) Sletten, E., Bertozzi, C. R. *Angew Chemie. Int. Ed.* **48**, 38, 6974–6998 (2009); (b) Bertozzi, C. R. *Acc. Chem. Res.* **44**, 9, 651–653 (2011); (b) Prescher, J. A., Bertozzi, C. R. *Nat. Chem. Biol.* **1**, 1, 13–21 (2005).
114. Zhang, X. *et al. Trends Chem.* **1**, 90–98 (2019).
115. Li, J., Chen, P. R. *Nat. Chem. Biol.* **12**, 3, 129–137 (2016).
116. Martinez-Calvo, M., Mascareñas, J. L. *Coord. Chem. Rev.* **359**, 57–59 (2018).
117. Bagshawe, K. D., Sharma, S. K., Begent, R. H. J. *Expert. Opin. Biol. Ther.* **4**, 1777–1789 (2004).
118. Vigh, L., Joó, F. Cséplö, Á. *Eur. J. Biochem.* **146**, 241–244 (1985).
119. (a) Rostovtsev, V. V. *et al. Angew Chem. Int. Ed.* **41**,2596–2599 (2002); (b) Tornøe, C. W., Christensen, C., Meldal, M. *J. Org. Chem.* **67**, 3057–3064 (2002).
120. (a) Su, Y. *et al. Sci. Rep.* **5**, 7724 (2015); (b) Wright, M. H., Sieber, S. A. *Nat. Prod. Rep.* **33**, 681–708 (2016); (c) Nakamoto, K., Akao, Y., Ueno, Y. *Bioorg. Med. Chem. Lett.* **28**, 2906–2909 (2018).
121. (a) Jiang, X. *et al. Expert. Opin. Drug Discov.* **14**, 779–789 (2019); (b) K. Kacprzak, *et al. Chem. Rev.* **116**, 5689–5743 (2016); (c) Dong, S. *et al. Mol. Pharm.* **16**, 3770–3779 (2019); (d) Dong, S. *et al. ACS Appl. Mater. Interfaces* **11**, 8740–8748 (2019).
122. (a) Kennedy, D. C. *et al J. Am. Chem. Soc.* **133**, 17993–18001 (2011); (b) D. Soriano del Amo, D. *et al. J. Am. Chem. Soc.* **132**, 16893–16899 (2010).
123. Streu, C., Meggers, E. *Angew Chem. Int. Ed.* **45**, 5645–5648 (2006).
124. Volker, T., *Angew Chem. Int. Ed.* **53**, 10536–10540 (2014).

125. Tomás-Gamasa, M. *et al. Nat. Commun.* **7**, 12538-12548 (2016).
126. Hartings, M. *Nat. Chem.* **4**, 9, 764 (2012).
127. Chalker, J. M., Wood, C. S. C., Davis, B. G. *J. Am. Chem. Soc.* **131**, 45, 16346–16347 (2009).
128. Li, N. *et al. J. Am. Chem. Soc.* **133**, 39, 15316–15319 (2011).
129. Li, J. *et al. J. Am. Chem. Soc.* **135**, 7330–7338 (2013).
130. Weiss, J. T. *et al. J. Med. Chem.* **57**, 5395–5404 (2014).
131. Weiss, J. T. *et al. Nat. Commun.* **5**, 3277 (2014).
132. Weiss, J. T. *et al. Sci. Rep.* **5**, 9329 (2015).
133. Rubio-Ruiz, B., Weiss, J. T., Unciti-Broceta, A. *J. Med. Chem.* **59**, 9974–9980 (2016).
134. (a) Hazra, S. *et al. Biochemistry* **49**, 31, 6784–6790 (2010.
135. Yusop, R. M. *et al. Nat. Chem.* **3**, 3, 239–243 (2011).
136. Li, J., *et al. Nat. Chem.* **6**, 352–361 (2014).
137. Li, B. *et al. Biomaterials* **138**, 57–68 (2017).
138. Miller, M. A. *et al. Nat. Commun.* **8**, 15906–15919 (2017).
139. Martínez-Calvo, M. *et al. ACS Catal.* **8**, 6055–6061 (2018).
140. Uttamapinant, C. *et al. Angew Chem. Int. Ed. Engl.* **51**, 24, 5852–5856 (2012).
141. Link, A. J., Tirrell, D. A. *J. Am. Chem. Soc.* **125**, 11164–11165 (2003).
142. Miguel-Ávila, J*. et al. Chem. Sci.* **9**, 1947–1952 (2018).
143. Vong, K., *et al. Chem. Sci.* **11**,10928–10933 (2020).
144. Vidal, C. *et al. Nat. Commun.* **9**, 1913–1923 (2018).
145. Oliveira, B. L. *et al. J. Am. Chem. Soc.* **142**, 10869–10880 (2020).
146. Chankeshwara, S. V., Indrigo, E., Bradley, M. *Curr. Opin. Chem. Biol.* **21**, 128–135 (2014).
147. Völker, T., Meggers, E. *Curr. Opin. Chem. Biol.* **25**, 48–54 (2015).
148. Zhang, X. *et al. ChemBioChem* **21**, 19, 2759–2763 (2020).
149. Schwizer, F. *et al. Chem. Rev.* **118**, 142–231 (2018).
150. (a) Davis, H. J., Ward, R. T. *ACS Cent. Sci.* **7**, 1120–1136 (2019); (b) Tanaka, K., Vong, K. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* **96**, 79–94 (2020).
151. Akabori, S., *et al. Nature* **178**, 323–324 (1956).
152. Kagan, H. B. *Compr. Asym. Catal.* **1**, 9–30 (1999).
153. Yamamura, K., Kaiser, E. T. *J. Chem. Soc., Chem. Commun.* **20**, 830–831(1976).
154. Wilson, M. E., Whitesides, G. M. *J. Am. Chem. Soc.* **100**, 306–307 (1978).
155. Davies, R. R., Distefano, M. D. *J. Am. Chem. Soc.* **119** 48, 11643–11652 (1997).
156. Baiyoumy, A. *et al. ACS Catal.* **11**, 10705–10712 (2021).
157. Christoffel, F. *et al. Nat. Catal.* **4**, 643–653 (2021).
158. (a) Sauer, D. F., et al. *ACS Catal.* **5**, 7519–7522 (2015); (b) Jeschek, M. *et al. Nature* **537**, 661–665 (2016).
159. Grimm, A. R. *et al. ACS Catal.* **8**, 2611–2614 (2018).

160. (a) Yu, K. *et al. J. Am. Chem. Soc.* **145**, 16621–16629 (2023); (b) Gu, Y. *et al. Angew Chem. Int. Ed.* **61**, e202110519 (2022); (c) Liu, Z. *et al. J. Am. Chem. Soc.* **144**, 883–890 (2022).
161. Stein, A. *et al. ACS Cent. Sci.* **7**, 1874–1884 (2021).
162. Schwizer, F. *et al. Chem. Rev.* **118**, 142−231 (2018).
163. Sabatino, V., Unnikrishnan, V. B., Bernardes, G. J. L. *Chem. Catalysis* **2**, 39–51 (2022).
164. Homchaudhuri, L., Swaminathan, R. *Bull. Chem. Soc. Jpn* **77**, 765–769 (2004).
165. Kumar, A. *et al. Biophys. J.* **118**, 468a (2020).
166. Huang, Q. *et al*. *Macromol. Rapid Commun.* **42**, (2021).
167. (a) Sivaramakrishnan, S. *et al. Proc. Natl. Acad. Sci. U. S. A.* **105**, 13356–13361 (2008); (b) Marqusee, S., Baldwin, R. L. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 8898–8902 (1987).
168. Wolny, M. *et al. Sci. Rep.* **7**, (2017).
169. Pauling, L., Corey, R. B., Branson, H. R. *Proc. Natl. Acad. Sci. USA* **37**, 205–211 (1951).
170. Wolny, M. *et al. J. Biol. Chem.* **289**, 27825–27835 (2014).
171. (a) Suveges, D. *et al. Proteins* **74**, 905–916 (2009); (b) Sivaramakrishnan, S. *et al. Biophys. J.* **97**, 2993–2999 (2009); (c) Spink, B. J. *et al. Nat. Struct. Mol. Biol.* **15**, 591–597 (2008).
172. (a) Dunnill, P. *Biophys. J.* **8**, 865–875 (1968); (b) Kim, D. E., Chivian, D., Baker, D. *Nucleic Acids Res.* **32**, (2004).
173. (a) Coin, I., Beyermann, M., Bienert, M. *Nat. Protoc.* **2**, 3247–3256 (2007); (b) G. S. Vanier, *Methods Mol. Biol.* **1047**, 235–249 (2013).
174. Larive, C. K., Jayawickrama, D., Orfi, L. *Appl. Spectrosc* **51**, 1531–1536 (1997).
175. Burton, I. W., Quilliam, M. A., Walter J. A. *Anal Chem* **77**, 3123–3131 (2005).
176. Rohl, C. A., Baldwin, R. L. *Biochemistry* **36**, 8435–8442 (1972).
177. Lawaetz, A. J., Stedmon, C. A. *App. Spectrosc.* **63**, 8 (2009).
178. (a) Cadeddu, M. *et al*. *Macromolecules* **57**, 514–527 (2024); (b) Alom, S. E., Swaminathan R. *Phys. Chem. Chem. Phys.* **25**, 16626–16642 (2023).
179. Baker, E., Bartlett, G., Crump, M. *et al. Nat. Chem. Biol.* **11**, 221–228 (2015).
180. Saenger, W. *Handbook of Proteolytic Enzymes*. (Elsevier Ltd., 2013).
181. Baek, M. *et al. Science* **373**, 871–876 (2021).
182. Abramson, J. *et al*. *Nature* **630**, 493–500 (2024).
183. Krivov, G. G., Shapovalov, M. V., Dunbrack, R. L. *Proteins: Struct., Funct., Bioinf.* **77**, 778–795 (2009).
184. Rackers, J. A. *et al. J. Chem. Theory Comput.* **14**, 5273–5289 (2018).
185. Maier, J. A. *et al. J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
186. Bannwarth, C. *et al. Wiley Interdiscip. Rev. Comput. Mol. Sci.* **11** (2021).
187. Al Attar, H. A., Monkman, A. P. *Biomacromolecules* **10**, 1076–1084 (2009).
188. (a) Lewis, B., Rathman. S., McMahon, R. J. *J. Nutr. Biochem.* **14**, 196–202 (2003); (b) Amancha, P. K., *et al. J. Immunol.* **191,** 6060–6070 (2013).
189. Sánchez-Fernández, R. *et al. ACS Sens.* **9**, 5052−5057 (2024).
190. Hemmilä, I., Laitala, V. *J. Fluoresc.* **15**, 529–542 (2005).

191. Horrocks, W. D., Collier, W. E. *J. Am. Chem. Soc.* **103**, 2856–2862 (1981).
192. (a) Zhang, R. *et al. BMCL* **26**, 1135–1150 (2018); (b) Thakkar, A., Trinh, T., Pei. D. *ACS Comb. Sci.* **15**, 2, 120–129 (2013).
193. Li, P., Roller, P. P. *Curr. Top. Med. Chem.* **2**, 3, 325-341 (2002).
194. Gentilucci, L., De Marco, R., Cerisoli, L. *Curr. Pharm. Des.* **16**, 28, 3185-3203 (2010).
195. Millward, S. W. *et al. ACS Chem. Biol.* 2, 9, 625–634 (2007).
196. (a) Driggers, E. M. *et al. Nat. Rev. Drug Discovery* **7**, 608–624 (2008); (b) Mallinson, J., Collins, I. *Future Med. Chem.* **4**, 1409–1438 (2012); (c) Marsault, E., Peterson, M. L. *J. Med. Chem.* **54**, 1961–2004 (2011).
197. White, C. J., Yudin, A. K. *Nat. Chem.* **3**, 509–524 (2011).
198. (a) Brady, S. F. *et al. J. Org. Chem.* **44**, 3101–3105 (1979); (b) Wadhwani, P. *et al. J. Org. Chem.* **71**, 55–61 (2006); (c) Qin, C. *et al. J. Med. Chem.* **46**, 4830–4833 (2003); (d) Perlman, Z. E. *et al. Bioorg. Med. Chem. Lett.* **15**, 5329–5334 (2005).
199. Thakkar, A., Trinh, T. B., Pei, D. *ACS Comb. Sci.* **15**, 120–129 (2013).
200. Kim, D. I., Han, S. J., Lim, Y. B. *Chem. Commun.* **58***,* 27, 4368–4371 (2022).
201. Learte-Aymamí, S., *et al. JACS Au* **4**, 7, 2630–2639 (2024).
202. (a) Völker, T., *et al. Angew Chem. Int. Ed.* **53**, 10536–10540 (2014); (b) Tomás-Gamasa, M., *et al. Nat. Commun.* **7**, 12538 (2016); (c) Sadler, J. C., *et al. RSC Chem. Biol.* **2**, 1073–1083 (2021); (d) Wang, W. *et al. Adv. Drug Deliv. Rev.* **176**, 113893 (2021); (e) Chang, T. C. Tanaka, K. *Bioorg. Med. Chem.* **46**, 116353 (2021).
203. (a) Li, J. *et al. Nat. Chem.* **6**, 352–361 (2014); (b) Martínez-Calvo, M. *et al. ACS Catal.* **8**, 6055–6061 (2018); (c) Brewster, R. C., Klemencic, E., Jarvis, A. G. *J. Inorg. Biochem.* **215**, 111317 (2021); James, C. C., de Bruin, B., Reek, J. N. H. *Angew Chem. Int. Ed.* **62**, e202306645 (2023).
204. (a) Ward, T. R. *Acc. Chem. Res.* **44**, 47–57 (2011); (b) Köhler, V. *et al. Nat. Chem.* **5**, 93–99 (2013) (c) Bos, J., Roelfes, G. *Curr. Opin. Chem. Biol.* **19**, 135–143 (2014); (d) Jeschek, M. *et al. Nature* **537**, 661–665 (2016); (e) Heinisch, T., Ward, T. R. *Acc. Chem. Res.* **49**, 1711–1721 (2016); (f) Schwizer, F. *et al. Chem. Rev.* **118**, 142–231 (2018).
205. (a) Destito, P. *et al. Chem.* **27**, 4789–4816 (2021); (b) van de L'Isle, M. O. N., Ortega-Liebana, M. C., Unciti-Broceta, A. *Curr. Opin. Chem. Biol.* **61**, 32–42 (2021); (c) Miguel-Ávila, J., Tomás-Gamasa, M., Mascareñas, J. L. *Trends Chem.* **5**, 474–485 (2023); (d) Gupta, A. *et al. Acc. Chem. Res.* **56**, 2151–2169 (2023); (e) James, C. C., de Bruin, B., Reek, J. N. H. *Angew Chem. Int. Ed.* **135** (2023).
206. Matsuo, T., Miyake, T., Hirota, S. *Tetrahedron* **60**, 151226–151233 (2019).
207. (a) Nastri, F. *et al. Trends Biochem. Sci.* **44**, 1022–1040 (2019); (b) Studer, S. *et al. Science* **362**, 1285–1288 (2018).
208. Powers, D. C., Ritter, T. Palladium(III) in Synthesis and Catalysis. in *Higher Oxidation State Organopalladium and Platinum Chemistry. Topics in Organometallic Chemistry*, vol. 35 129–156 (Springer, Berlin, Heidelberg, 2011).
209. Martínez-Calvo, M. *et al. ACS Catal* **8**, 7, 6055–6061 (2018); (b) Brewster, R. C., Klemencic, E., Jarvis, A. G. *J. Inorg. Biochem.* **215**, 111317 (2021). (c) Li, J., *et al. Nat. Chem.* **6**, 4, 352–361 (2014).

210. (a) Latocheski, E. *et al. Chem. Soc. Rev.* **49**, 7710–7729 (2020); (b) Coelho, S. E. *et al. ACS Catal.* **9**, 3792–3799 (2019); (c) Yusop, R. M. *et al. Nat. Chem.* **3**, 239–243 (2011).
211. (a) Akagawa, K., Satou, J., Kudo, K. *J. Org. Chem.* **81**, 9396–9401 (2016); (b) Zheng, L. *et al. Angew Chem. Int. Ed.* **53**, 7599–7603 (2014); (c) Metrano, A. J. *et al. Chem. Rev.* **120**, 11479–11615 (2020); (d) Kinghorn, M. J. *et al. ACS Catal.* **7**, 7704–7708 (2017).
212. Learte-Aymamí, S. *et al*. *Angew Chem. Int. Ed.* **59**, 9149–9154 (2020).
213. Macias, M. J., Wiesner, S., Sudol, M. *FEBS Lett.* **513**, 30–37 (2002).
214. Macias, M. *et al. Nature* **382**, 646–649 (1996).
215. (a) Dalby, P. A., Hoess, R. H., Degrado, W. F. *Protein Sci.* **9**, 2366–2376 (2000); (b) Deechongkit, S., Kelly, J. W. *J. Am. Chem. Soc.* **124**, 4980–4986 (2002); (c) Russ, W. P. *et al. Nature* **437**, 579–583 (2005); (d) Ferguson, N. *et al. PNAS* **98**, 13002–13007 (2001); (e) Zhou, R. *et al*. *Proc. Natl. Acad. Sci. USA* **111**, 18243–18248 (2014).
216. Pham, T. L., Kovermann, M., Thomas, F. *ACS Synth. Biol.* **11**, 254–264 (2022).
217. Macias, M. *et al. Nat. Struct. Mol. Biol.* **7**, 375–379 (2000).
218. Subirós-Funosas, R., El-Faham, A., Albericio, F. *Tetrahedron* **67**, 8595–8606 (2011).
219. Alonso-Cotchico, L. *et al. Acc. Chem. Res.* **53**, 896–905 (2020).
220. Sánchez-Aparicio, J. E. *et al. J. Chem. Inf. Model* **61**, 311–323 (2021).
221. (a) Coin, I., Beyermann, M., Bienert, M. *Nat. Protoc.* **2**, 3247–3256 (2007); (b) G. S. Vanier, *Methods Mol. Biol.* **1047**, 235–249 (2013).
222. Micsonai, A. *et al*. *Proc. Natl. Acad. Sci. USA* **112**, E3095–E3103 (2015).
223. (a) Jäger, M., Dendle, M., Kelly, J. W. *Protein Sci.* **18**, 1806–1813 (2009); (b) Koepf, E. K. *et al. Protein Sci* **8** (4), 841–853 (1999).
224. (a) Pham, T. L., Kovermann, M., Thomas, F. *ACS Synth. Biol.* **11**, 1, 254–264 (2022); (b) Li, W. *et al. J. Am. Chem. Soc.* **130**, 3, 892–900 (2008).
225. M. Martínez-Calvo *et al. ACS Catal.* **8**, 6055 – 6061 (2018).
226. Gao, T. *et al. Chem. Asian J.* **10**, 5, 1142–1145 (2015).
227. Learte-Aymamí, S. *et al. Angew Chem. Int. Ed. Engl.* **59**, 9149–9154 (2020).
228. Dal Forno, G. M. *et al. Chem. Sci.* **15**, 4458–4465 (2024).
229. Lättig-Tünnemann, G. *et al. Nat. Commun.* **2**, 453 (2021).
230. Li, S. *et al. Chem. Commun.* **56**, 100, 15655–15658 (2020).
231. Learte-Aymamí, S. *et al. Angew Chem. Int. Ed. Engl.* **59**, 23, 9149–9154 (2020).
232. Learte-Aymamí, S. *et al. J. Am. Chem. Soc.* **139**, 45, 16188–16193 (2017).
233. Miller, M. A. *et al. Nat. Commun.* **8**, 15906–15919 (2027).
234. Im, H. G. *et al. Org. Biomol. Chem.* **11**, 18, 2966–2971 (2013).
235. Hitomi, Y., Takeyasu, T. Kodera, M. *Chem. Commun.* **49**, 85, 9929–9931 (2013).
236. (a) Simpson, D. J. *et al. J. Org. Chem.* **56**, 18, 5391–5396 (1991); (b) Li, Z. *et al. Anal. Chem.* **85**, 8, 3926–3932 (2013); (c) Y. Shiraishi, *et al. Chem. Commun.* **49**, 99, 11680–11682 (2013).
237. Li, X. *et al. Biosens. Bioelectron.* **200**, 113929 (2022).

238. (a) Urrutia, M., Ortiz, C. *Biotech. Histochem.* **90**, 159–166 (2015); (b) Magde, D. *et al. J. Phys. Chem.* **83**, 696–699 (1979); (c) Gong, Q. *et al. Chem. Sci.* **7**, 788–792 (2016); (d) Zhang, J. *et al. J. Am. Chem. Soc.* **133**, 14109–14119 (2011).

239. (a) Bae, S. K. *et al. J. Am. Chem. Soc.* **135**, 9915–9923 (2013); (b) Haidekker, M. A. *et al. J. Am. Chem. Soc.* **128**, 398–399 (2006); (c) Han, X. *et al. Chem. Sci.* **7**, 5098–5107 2016). (d) Li, X. Y. et al. *Anal. Chem.* **91**, 11409–11416 (2019).

240. González-González, C., *et al. ACS Catal* **15**, 8624–8632 (2025).

241. (a) Devaraj, N. K. *ACS Cent. Sci.* **4**, 952−959 (2018); (b) Li, Y., Fu, H. *ChemistryOpen* **9**, 835−853 (2020); (c) Seoane, A., Mascareñas, J. L. *Eur. J. Org. Chem.* **32** (2022); (d) Scinto, S. L. *et al. Nat. Rev. Methods. Primers.* **1**, 30−23 (2021); (e) Madec, H. *et al. Chem. Sci.* **14**, 409−442 (2023).

242. (a) Davis, H. J., Ward, T. R. *ACS Cent. Sci.* **5**, 1120–1136 (2019); (b) Wittwer, M. *et al. Nature Catal.* **4**, 814–827 (2021); (c) Bos, J., Roelfes, G. *Curr. Opin. Chem. Bio.l* **19**, 135–143 (2014).

243. Jarvo, E. R., Miller, S. J. *Tetrahedron* **58**, 2481–2495 (2002).

244. (a) Akagawa, K. *et al. J. Org. Chem.* **81**, 9396–9401 (2016); (b) Zheng, L. *et al. Angew Chem. Int. Ed. Engl.* **53**, 7599–7603 (2014).

245. Kinghorn, M. J. *et al. ACS Catal.* **7**, 7704–7708 (2017).

246. Metrano, A. J. *et al. Chem. Rev.* **120**, 11479–11615 (2020).

247. Learte-Aymamí, S. *et al. Commun. Chem.* **5**, 1, 75 (2022).

248. Akcapinar, G. B., Sezerman, O. U. *Biosci. Rep.* **37***,* 2, BSR20160179 (2017).

249. (a) Huang, X., Pieczko, M. E., Long, E. C. *Biochemistry* **38**, 2160–2166 (1999); (b) Berkessel, A., Hérault, D. A. *Angew. Chem. Int. Ed.* **38**, 102–105 (1999); (c) Francis, M. B., Jacobsen, E. N. *Angew. Chem. Int. Ed.* **38**, 937–941 (1999); (d) Copeland, G. T., Miller, S. J. *J. Am. Chem. Soc.* **123**, 6496–6502 (2001); (e) Hagemeyer, A. *et al. Appl. Catal. A* **221**, 23–43 (2001); (f) Berkessel, A. *Curr. Opin. Chem. Biol.* **7**, 409–419 (2003); (g) Kerton, F. M., Whitwood, A. C., Willans, C. E. *Dalton Trans.* **15**, 2237–2244 (2004); (h) Sambasivan, R., Ball, Z. T. *Angew. Chem. Int. Ed.* **51**, 8568–8572 (2012); (i) Maeda, Y. *et al. Annu. Rev. Biomed. Eng.* **18**, 311–328 (2016).

250. Frank, R. *Tetrahedron* **48**, 9217–9232 (1992).

251. (a) Hilpert, K., Winkler, D. F. H., Hancock, R. E. W. *Nat. Protoc.* **2**, 1333–1349 (2007); (b) Hilpert, K., Winkler, D. F. H., Hancock, R. E. W. *Biotechnol. Genet. Eng. Rev.* **24**, 31–106 (2007); (c) Reineke, U., Volkmer-Engert, R., Schneider-Mergener, J. *Curr. Opin. Biotechnol.* **12**, 59–64 (2001).

252. Frank, R. *et al. In Combinatorial Peptide and Nonpeptide Libraries* 363–386 (Wiley-VCH Verlag GmbH, 2007).

253. (a) Thiele, A., Stangl, G. I., Schutkowski, M. *Mol. Biotechnol.* **49**, 283–305 (2011); (b) Volkmer, R., Tapia, V., Landgraf, C. *FEBS Lett.* **586**, 2780–2786 (2012); (c) Engelmann, B. W. *Methods Mol. Biol.* **1555**, 375–394 (2017); (d) Sandmann, C. L. et al. *Mol. Cell* **83**, 994–1011.e18 (2023).

254. Biswas, R., Maillard, N., Kofoed, J., Reymond, J. L. *Chem. Commun.* **46**, 8746–8748 (2010).

255. Cochran, A. G., Skelton, N. J., Starovasnik, M. A. *Proc. Natl. Acad. Sci. USA* **98**, 5578–5583 (2001).

256. Barrow, A. S. *et al.*, *Chem. Soc. Rev.* **48**, 4731–4758 (2019).
257. (a) Pickens, C. J. *et al. Bioconjug. Chem.* **29**, 686–701 (2018); (b) Uttamapinant, C. *et al. Angew. Chem. Int. Ed.* **51**, 5852–5856 (2012); (c) Bevilacqua, V. *et al. Angew. Chem. Int. Ed.* **53**, 5872–5876 (2014); (d) Besanceney-Webler, C. *et al. Angew. Chem. Int. Ed*. **50**, 8051–8056 (2011).
258. Pescina, S. *et al. J. Control. Release* **284**, 84–102 (2018).
259. Raucher, D. *Curr. Opin. Pharmacol.* **47**, 14–19 (2019).
260. Zhao, N. *et al. Anticancer Agents Med. Chem.* **18**, 74–86 (2018).
261. (a) Pap, E. H. *et al. Exp. Cell Res.* **265**, 288–293 (2001); (b) Rajendran, L., Knölker, H. J., Simons, K. *Nat. Rev. Drug Discov.* **9**, 29–42 (2010).
262. Cochran, A. G., Skelton, N. J., Starovasnik, M. A. *Proc. Natl. Acad. Sci. USA* **98**, 10, 5578–5583 (2001).
263. Pham, T. L. *et al. ChemBioChem* **24**, 3, e202200588 (2023).
264. Minor, D. L., Jr, Kim, P. S. *Nature* **367**, 6464, 660–663 (1994).
265. Guixer, B. *et al. J. Pept. Sci.* **22**, 9, 577–591 (2016).
266. Dikmans, A. *et al. QSAR Comb. Sci.* **25**, 11, 1069–1080 (2006).
267. Learte-Aymamí, S. *et al. Angew. Chem. Int. Ed.* **59**, 23, 9149–9154 (2020).
268. Davis, H. J., Ward, T. R. *ACS Cent. Sci.* **5**, 7, 1120–1136 (2019).
269. Roy, A., Bour, P., Keiderling, T. A. *Chirality* **21 Suppl 1**, E163-71 (2009).
270. Pham, T. L., Kovermann, M., Thomas, F. *ACS Synth. Biol.* **11**, 1, 254–264 (2022).
271. (a) Learte-Aymamí, S. *et al. J. Am. Chem. Soc.* **139**, 16188–16193 (2017); (b) Lättig-Tünnemann, G. *et al. Nat. Commun.* **2**, 453 (2011); (c) Li, S. *et al. Chem. Commun.* **56**, 15655–15658 (2020).
272. Golas, L. *et al. Chem. Soc. Rev.* **39**, 1233 (2010).
273. Martínez-Calvo, M. *et al. ACS Catal.* **8**, 6055–6060 (2018).
274. Xie, F. *et al. Tetrahedron* **64**, 2906 (2008).
275. Drienovská, I. *et al. Chem. Sci.* **8**, 7228–7235 (2017).
276. Rebelein, J. G., Ward, T. R. *Curr. Opin. Biotechnol.* **53**, 106–114 (2018).
277. Patra, M., Gasser, G. *Chembiochem* **13**,1232–1252 (2012).
278. Kelly, S., Price, N. *CPPS* **1**, 349–384 (2000).
279. Ebeling, W. *et al. Eur. J. Biochem.* **47**, 91–7 (1974).
280. (a) Schwartz, D. C., Cantor, C. R. *Cell* **37**, 67–75 (1984); (b) Cleveland, D. W. *et al. J. Biol. Chem.* **252**, 1102–6 (1977); (c) Herrmann, B. G., Frischauf, A. M. *Meth. Enzymol.* **152**, 180–3 (1987); (d) Lee, J. J., Costlow, N. A. *Meth. Enzymol.* **152**, 633–48 (1987).
281. Learte-Aymamí, S. *et al. Angew. Chem. Int. Ed Engl.* **59**, 9149–9154 (2020).
282. Gao, T. *et al. Chem. Asian J.* **10**, 1142–1145 (2015).
283. Clavadetscher, J. *et al. Angew. Chem. Int. Ed.* **56**, 6864–6868 (2017).

9. APPENDIX

9.1. Appendix 1. List of Publications

Non-Aromatic Fluorescence from Single α-Helical Peptides

Accepted to be published in Cell Reports Physical Science (*CR-PHYS-SCI-D-25-00062R4*)

Carmen González-González,[a] Roi Lopez-Blanco,[a] Juan A. González-Vera,[c] Sara D'Ingiullo,[a] David Bouzada,[a] Manuel Melle-Franco,[b] Ángel Orte,[c] M. Eugenio Vázquez[a]

Affiliations:

a. Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Departamento de Química Orgánica, Universidade de Santiago de Compostela. Santiago de Compostela 15782, Spain.
b. CICECO - Aveiro Institute of Materials, Department of Chemistry, University of Aveiro, Aveiro 3810-193, Portugal.
c. Departamento de Fisicoquímica, Unidad de Excelencia de Química Aplicada a Biomedicina y Medioambiente, Facultad de Farmacia, Universidad de Granada, Campus la Cartuja, Granada, 18071, Spain.

Specific contribution in the publication:

Synthesized the peptides and conducted preliminary steady state emission experiments; conducted the steady state photoluminescence experiments, including measuring the EEMs for all the peptides, the proteinase K digestion and other control experiments, and the lanthanide sensitization studies; validation, formal analysis, review and editing the manuscript.

Quality indexes:

Results of Chapter 1 were published in Cell Reports Physical Science, an open access journal from Cell Press publishing cutting-edge research across the physical sciences, including chemistry, physics, materials science, energy science, engineering, and related interdisciplinary work, currently has an impact factor of 7.9 (Clarivative Analytics, 2023), a CiteScore of 11.4 and the following positions in the following categories: quartile 1 (Q1) in Chemistry, Energy, Engineering, Material Science and Physics and Astronimy (SJR 2024 2.098) calculated by Scimago:
https://www.scimagojr.com/journalsearch.php?q=21101037113&tip=sid&clean=0

***De Novo* Engineering of Pd-Metalloproteins and Their Use as Intracellular Catalysts**

JACS Au **4**, 7, 2630–2639 (2024) https://doi.org/10.1021/jacsau.4c00379

Soraya Learte-Aymamí,[a] Laura Martínez-Castro,[b] Carmen González-González,[a] Miriam Condeminas,[c,d] Pau Martin-Malpartida,[c] Sandra Baúlde,[a] José R. Couceiro,[a] Jean-Didier Maréchal,[b] Maria J. Macias,[c,e] José L. Mascareñas,[a] M. Eugenio Vázquez[a] *JACS Au* **4**, 7, 2630–2639 (2024).

Affiliations:

a. Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Departamento de Química Orgánica. Universidade de Santiago de Compostela, Santiago de Compostela 15705, Spain.
b. Insilichem, Departament de Química, Universitat Autònoma de Barcelona, Cerdanyola 08193, Spain.
c. Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology (BIST), Baldiri Reixac, 10, Barcelona 08028, Spain.
d. Department of Medicine and Life Sciences, Universitat Pompeu Fabra (MELIS-UPF), Carrer del Doctor Aiguader 88, Barcelona 08003, Spain.
e. Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, Barcelona 08010, Spain.

Specific contribution in the publication:

Synthesized the peptides and conducted the catalytic studies in vitro and in cell cultures; synthesized the fluorogenic probes; data curation, formal analysis, investigation, methodology, visualization, writing-review & editing.

Quality indexes:

Results of Chapter 2 were published in JACS Au, an open access journal of the American Chemical Society, currently has a 2-year impact factor of 8.6 (2023), a CiteScore of 9.1 (2023) and the following positions in the following categories: quartile 1 (Q1) in Chemistry (SJR 2024 2.944) calculated by Scimago:
https://www.scimagojr.com/journalsearch.php?q=21101132410&tip=sid&clean=0

Journal authorization:

Streamlined Identification of Metallopeptides for Intracellular Catalysis Using Positionally-Addressable Combinatorial Libraries

ACS Catal. **15**, 8624−8632 (2025) https://doi.org/10.1021/acscatal.5c00525

Carmen González-González,[a,] Laura Martínez-Castro,[b,] Soraya Learte-Aymamí,[a,] Clara Pose-Insua,[a] José R. Couceiro,[c] Pau Martin-Malpartida,[d] Maria J. Macias,[d,e] Jean-Didier Maréchal,[b] José L. Mascareñas,[a] M. Eugenio Vázquez[a] *ACS Catal.* **15**, 8624−8632 (2025) https://doi.org/10.1021/acscatal.5c00525

Affiliations:

a. Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Departamento de Química Orgánica, Universidade de Santiago de Compostela. Santiago de Compostela 15705, Spain.
b. Insilichem, Departament de Química, Universitat Autònoma de Barcelona, Cerdanyola 08193, Spain.

c. Centro Singular de Investigación en Química Biolóxica e Materiais Moleculares (CiQUS), Universidade de Santiago de Compostela. Santiago de Compostela 15705, Spain.
d. Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Baldiri Reixac, 10, Barcelona 08028, Spain.
e. Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, Barcelona 08010, Spain.

Specific contribution in the publication:

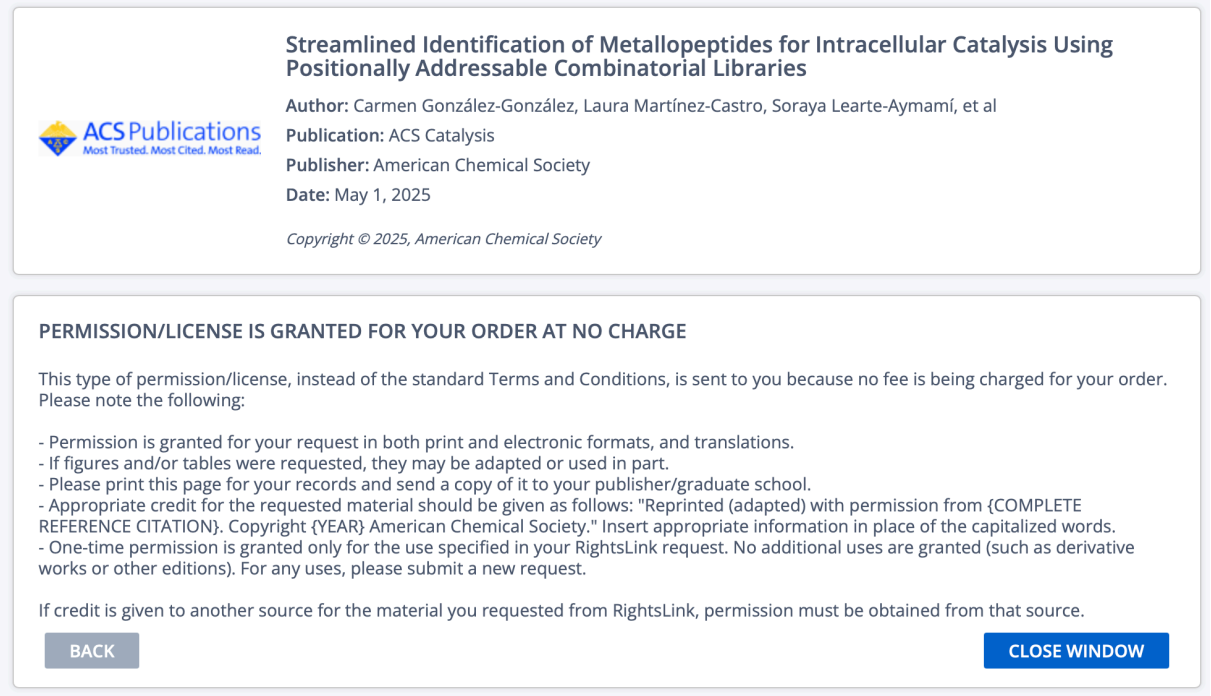
Conducted the catalysis experiments; participated in the data interpretation and manuscript preparation.

Quality indexes:

Results of Chapter 3 were published in ACS Catalsis, a journal of the American Chemical Society, currently has a 2-year impact factor of 11.7 (2023), a CiteScore of 20.8 (2023) and the following positions in the following categories: quartile 1 (Q1) in Chemical Engineering and Chemistry (SJR 2024 3.782) calculated by Scimago:
https://www.scimagojr.com/journalsearch.php?q=19700188320&tip=sid

Journal authorization:

Permission granted from *American Chemical Society.*



ACS Publications
Most Trusted. Most Cited. Most Read.

Streamlined Identification of Metallopeptides for Intracellular Catalysis Using Positionally Addressable Combinatorial Libraries

Author: Carmen González-González, Laura Martínez-Castro, Soraya Learte-Aymamí, et al
Publication: ACS Catalysis
Publisher: American Chemical Society
Date: May 1, 2025

Copyright © 2025, American Chemical Society

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms and Conditions, is sent to you because no fee is being charged for your order. Please note the following:

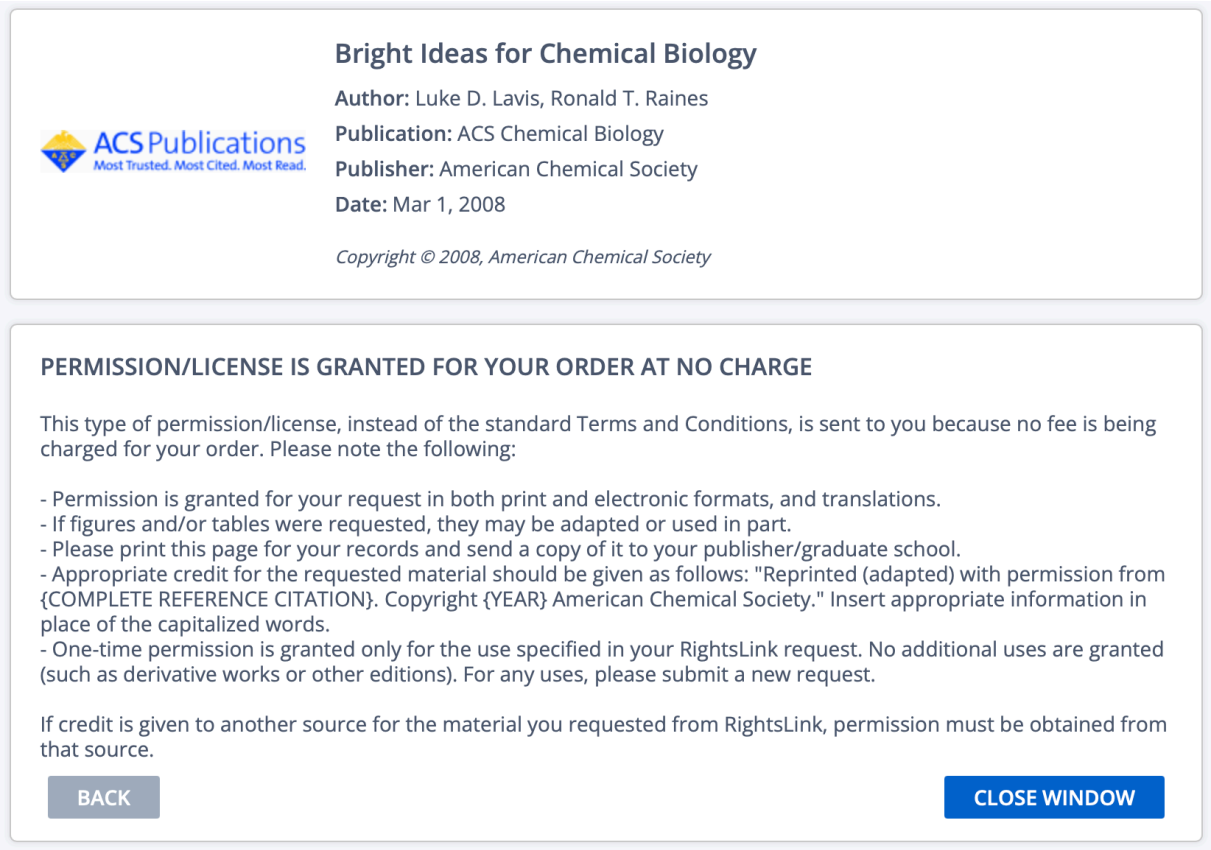
- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from {COMPLETE REFERENCE CITATION}. Copyright {YEAR} American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your RightsLink request. No additional uses are granted (such as derivative works or other editions). For any uses, please submit a new request.

If credit is given to another source for the material you requested from RightsLink, permission must be obtained from that source.

BACK CLOSE WINDOW

9.2. Appendix 2. Rights and Permissions of the Images

- **Figure 12.** Lavis, L. D., Raines, R. T. *ACS Chem. Biol.* **3**, 142–155 (2008) Permission granted from *American Chemical Society*



ACS Publications
Most Trusted. Most Cited. Most Read.

Bright Ideas for Chemical Biology

Author: Luke D. Lavis, Ronald T. Raines
Publication: ACS Chemical Biology
Publisher: American Chemical Society
Date: Mar 1, 2008

Copyright © 2008, American Chemical Society

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

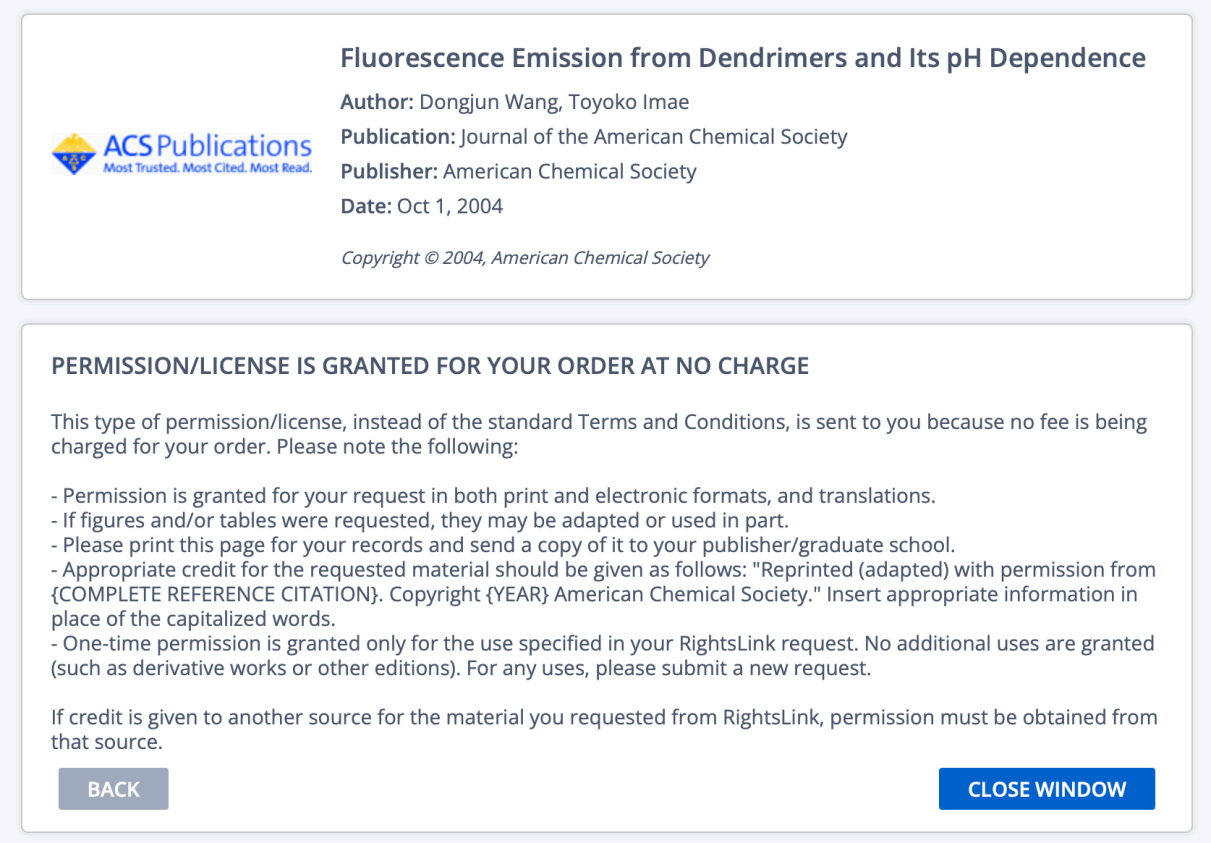
This type of permission/license, instead of the standard Terms and Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from {COMPLETE REFERENCE CITATION}. Copyright {YEAR} American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your RightsLink request. No additional uses are granted (such as derivative works or other editions). For any uses, please submit a new request.

If credit is given to another source for the material you requested from RightsLink, permission must be obtained from that source.

BACK CLOSE WINDOW

- **Figure 15a.** Wang, D.; Imae, T. *J. Am. Chem. Soc.* **126**, 13204–13205 (2004). Permission granted from *American Chemical Society*



ACS Publications
Most Trusted. Most Cited. Most Read.

Fluorescence Emission from Dendrimers and Its pH Dependence

Author: Dongjun Wang, Toyoko Imae
Publication: Journal of the American Chemical Society
Publisher: American Chemical Society
Date: Oct 1, 2004

Copyright © 2004, American Chemical Society

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms and Conditions, is sent to you because no fee is being charged for your order. Please note the following:

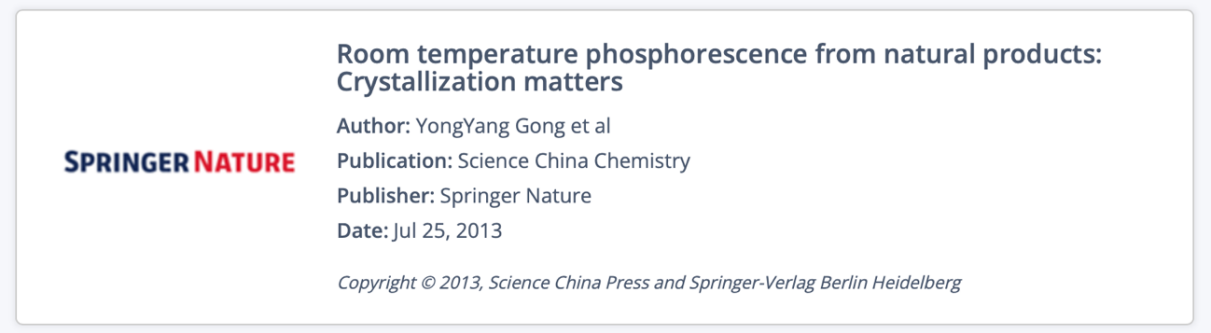
- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from {COMPLETE REFERENCE CITATION}. Copyright {YEAR} American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your RightsLink request. No additional uses are granted (such as derivative works or other editions). For any uses, please submit a new request.

If credit is given to another source for the material you requested from RightsLink, permission must be obtained from that source.

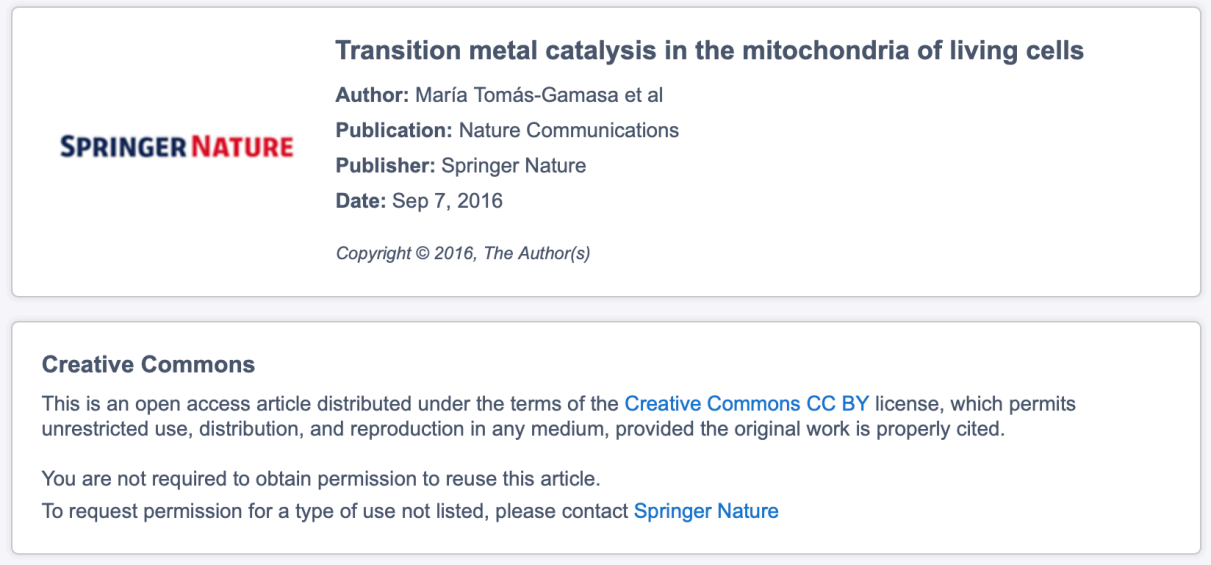
BACK CLOSE WINDOW

- **Figure 15b.** Stagi, L. *et al. Macromol. Chem. Phys.* **222**, (2021). © 2021 The Authors. Macromolecular Chemistry and Physics published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

- **Figure 15c.** Y. Gong, *et al. Sci. China Chem.* **56**, 1178–1182 (2013). Permission granted from *Springer Nature* https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=de504b22-6de7-4112-b715-2843129abcb3

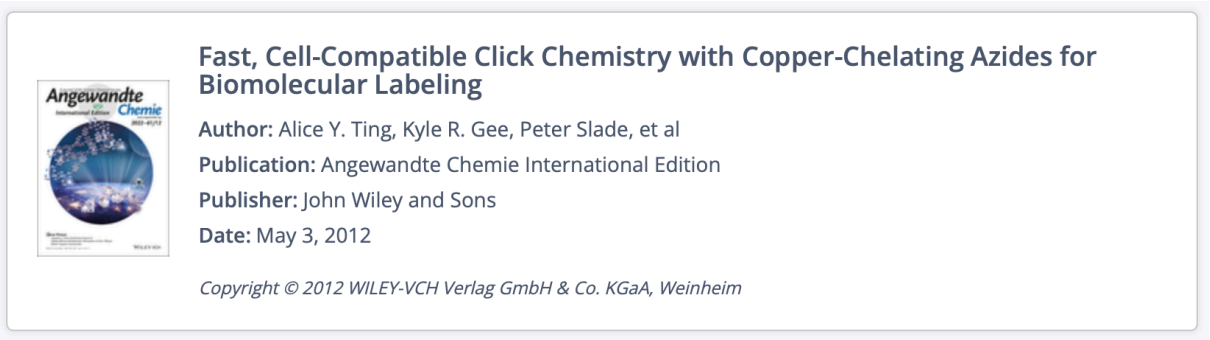


- **Figure 19.** Tomás-Gamasa, M. *et al. Nat. Commun.* **7**, 12538–12548 (2016). This is an open access article under the terms of the Creative Commons CC License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.



- **Figure 20b.** Martínez-Calvo, M. *et al. ACS Catal.* **8**, 6055–6061 (2018). © 2018 *American Chemical Society.* This is an open access article under the terms of the Creative Commons CC License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

- **Figure 21a.** Uttamapinant, C. *et al. Angew Chem. Int. Ed. Engl.* **51**, 24, 5852–5856 (2012).
  Permission granted from *John Wiley and Sons*
  https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=e78256ea-f5e0-4fb5-aa80-a14b86189a2f



- **Figure 21b.** Miguel-Ávila, J. *et al. Chem. Sci.* **9**, 1947–1952 (2018).
  © 2018 *Chemical Science*. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence. You can use material from this article in other publications without requesting further permissions from the RSC, provided that the correct acknowledgement is given.

- **Figures 5, 6, 7, 10, 17** and **18** were personally created with BioRender:
  https://www.biorender.com

- **All other images in this Doctoral Thesis** without the corresponding acknowledgment and proper reproduction have been personally elaborated.

Historically, peptides have been used in the pharmaceutical industry and in biomolecular research for various applications. However, recent advances have demonstrated their potential in areas such as enzymatic catalysis and the development of fluorescent sensors.

This thesis focuses on the design and optimization of peptides with applications in fluorescence and bioorthogonal catalysis. For this purpose, non-aromatic fluorescent peptides based on single α-helices, with potential in bioimaging; catalytic metallopeptides of β-sheet structure, active in cellular environments; and a combinatorial methodology using SPOT peptide libraries to identify optimal catalytic sequences is introduced.