Universidade de Lisboa
Faculdade de Farmácia

Chemical site-selective functionalization of proteins with boronic acids

Roberto Russo

Orientadores: Professor Doutor Pedro Miguel Pimenta Gois
Doutor Gonçalo Jose Lopes Bernardes

Tese orientada pelo e coorientada pelo, especialmente elaborada para obtenção do grau de Doutor em Doutoramento em Farmacia, especialidade de Química Farmacêutica e Terapêutica

2019

Universidade de Lisboa
Faculdade de Farmácia

Chemical site-selective functionalization of proteins with boronic acids

Roberto Russo

Orientadores: Professor Doutor Pedro Miguel Pimenta Gois
Doutor Gonçalo Jose Lopes Bernardes

Tese orientada pelo e coorientada pelo, especialmente elaborada para obtenção do grau de Doutor em Doutoramento em Farmacia, especialidade de Química Farmacêutica e Terapêutica

Júri:
Presidente: Doutor António José Leitão das Neves Almeida, Professor Catedrático e Presidente do Concelho Científico da Faculdade de Farmácia da Universidade de Lisboa
Vogais:
- Doutor Francisco Corzana López, Professor Titular de Universidad, Universidad de la Rioja, Espanha;

- Doutora Maria Manuel Martinho Sequeira Barata Marques, Professora Auxiliaria com Agragação, Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa;

- Doutor Tiago Correia de Oliveira Rodrigues, Investigador Auxiliar, Faculdade de Medicina da Universidade de Lisboa;

- Doutor Carlos Alberto Mateus Afonso, Professor Catedrático, Faculdade de Farmácia da Universidade de Lisboa

- Doutor Pedro Miguel Pimenta Gois, Professor Auxiliari com Agragação, Faculdade de Farmácia da Universidade de Lisboa, Orientador

Este trabalho foi financiado através da bolsa de doutoramento MSCA-ITN-2015-ETN-675007.

2019

O presente trabalho foi desenvolvido sob orientação do Professor Doutor Pedro M. P. Góis do iMed.ULisboa (Instituto de Investigação do Medicamento) da Faculdade de Farmácia da Universidade de Lisboa e co-orientação do Doutor Gonçalo J. L. Bernardes, do Instituto de Medicina Molecular da Faculdade de Medicina da Universidade de Lisboa. Este trabalho foi financiado pela Marie Skłodowska-Curie Actions ITN - ProteinConjugates através da bolsa de doutoramento MSCA-ITN-2015-ETN-675007.

This work was developed under scientific supervision of Professor Dr. Pedro M. P. Góis from iMed.ULisboa (Instituto de Investigação do Medicamento), Faculty of Pharmacy, University of Lisbon and co-supervision of Dr. Gonçalo J. L. Bernardes, from the Instituto de Medicina Molecular, Faculty of Medicine, University of Lisbon. This work was financially supported by the Marie Skłodowska-Curie Actions ITN - ProteinConjugates, through the grant MSCA-ITN-2015-ETN-675007.

List of Publications and Awards

Papers in International Scientific Periodicals with Referees:

- J. P. M. António, **R. Russo**, C. P. Carvalho, P. M. S. D. Cal, P. M. P. Gois, "Boronic acids as building blocks for the construction of therapeutically useful bioconjugates". *Chem. Soc. Rev.* **2019**, DOI: 10.1039/c9cs00184k

Book Chapter:

- S. Aguiar, J. Dias, A. M. Manuel, **R. Russo**, P. M. P. Gois, F. A. da Silva and J. Goncalves. Chapter Five - Chimeric Small Antibody Fragments as Strategy to Deliver Therapeutic Payloads; in "Therapeutic Proteins and Peptides", ed. R. Donev, *Academic Press*, **2018**, vol. 112, pp. 143–182.

Patent Submission:

- **Roberto Russo**, Rita I. C. Padanha, Pedro M. P. Gois, "On the use of 3-Hydroxy Quinolinones as Boron Hot-Spot for the preparation of protein conjugates, their uses and methods for their preparation", **PT20181000078171.**

Oral Communications in Scientific Conferences:

- **Roberto Russo**, Rita I. C. Padanha, Fabio Fernandes, Luis F. Veiros, Pedro M. P. Gois "Engineering Boron Hot-Spots for the Site-Selective Formation of Iminoboronates on Peptide Chains". Chemistry and Biology CrossTalks, held in Lisbon, January 2019.
- **Roberto Russo**, Rita I. C. Padanha, Luis F. Veiros, Pedro M. P. Gois "Functionalized 3-Hydroxyquinolinones as Boronic Acid Hot-Spots". MedChemSicily, held in Palermo, July 2018.
- **Roberto Russo**, Rita I. C. Padanha, Luis F. Veiros, Pedro M. P. Gois "Functionalized 3-Hydroxyquinolinones as Boronic Acid Hot-Spots". ITN ProteinConjugateS Summer School Meeting, held in Vienna, July 2018.
- **Roberto Russo**, Luis F. Veiros, Pedro M. P. Gois "Bioconjugation of boronic acids via a 3-Hydroxy-Quinolinone platform". 2nd Encontro do Colégio de Química da Ulisboa, held in Lisbon, October 2017.
- **Roberto Russo**, Luis F. Veiros, Pedro M. P. Gois "Development of small molecular hotspots for boronic acid bioconjugation". ITN ProteinConjugateS mid-term

meeting, held in Bruxelles, October 2017.

- **Roberto Russo**, Luis F. Veiros, Pedro M. P. Gois “Development of small molecular hotspots for boronic acid bioconjugation”. ITN ProteinConjugateS meeting, held in London, July 2017.

Poster Communications in Scientific Conferences:

- **Roberto Russo**, Rita I. C. Padanha, Fabio Fernandes, Luis F. Veiros, Pedro M. P. Gois “Engineering Boron Hot-Spots for the Site-Selective Formation of Iminoboronates on Peptide Chains”. 2nd PSL Chemical Biology Symposium, held in Paris, January 2019.
- **Roberto Russo**, Roberta Paterna, Pedro M. P. Gois “3HQs as responsive handles for boronic acid bioconjugation”. 10th iMed.ULisboa Postgraduate Students Meeting & 3rd i3DU Meeting, held in Lisbon, July 2018.
- **Roberto Russo**, Roberta Paterna, Pedro M. P. Gois “3HQs as responsive handles for boronic acid bioconjugation”.12th ENQO and 5ENQT, held in Coimbra, January 2018.
- **Roberto Russo**, Roberta Paterna, Pedro M. P. Gois “Development of small molecular cores for boronic acid ligation in peptides”. XXVENSPQ, held in Lisbon, July 2017.

Agradecimentos

Table of contents

Figure Index

Table Index

Scheme Index

Abbreviations and Symbols

- **3HQ(s)** – 3-hydroxy quinolinone
- **ADC(s)** – Antibody drug conjugate(s)
- **Ab** – Antibody
- **AL** – Alizarin
- **BA(s)** – Boronic acid(s)
- **BE(s)** – Boronate Ester(s)
- **Bn** – Benzyl
- **^{11}B-NMR** – ^{11}Boron Nuclear magnetic resonance
- **BSA** – Bovine serum albumine
- **Btz** – Bortezomib
- **^{13}C-NMR** – ^{13}Carbon Nuclear magnetic resonance
- **2-ABBA** – 2-acetyl benzene boronic acid
- **2-CBBA(s)** – 2-carbonyl benzene boronic acid(s)
- **2-FBBA** – 2-formyl benzene boronic acid
- **DFT** – Density functional theory
- **DMF** – Dimethylformamide
- **DMSO** – Dimethyl sulfoxide
- **equiv.** – Equivalents
- **ESI-MS** – Electrospray ionization mass spectrometry
- **FDA** – Food and Drug Administration
- **GSH** – Reduced glutathione
- **^{1}H-NMR** – ^{1}Hydrogen Nuclear magnetic resonance
- **HPA** - Hyperphanylalaninemia
- **hPAH** – Human Phenylalanyne hydroxylase
- **HPLC** – High-performance liquid chromatography
- **HRMS** – High resolution mass spectrometry
- **IC_{50}** – half maximal inhibitory concentration
- **LC-MS** – Liquid-chromatography mass spectrometry
- **LRMS** – Low resolution mass spectrometry
- **mp** – Melting point
- **NBD** – 7-nitrobenzooxadiazole
- **NHS** – *N*-hydroxy succinimide
- **PBA** – Phenyl boronic acid

- **PBS** – Phosphate buffer saline
- **PEG** – Polyethylene glycol
- **Ph** – Phenyl
- **PKU** – Phenylketonuria
- **RNA(s)** – Ribonucleic acid(s)
- **ROS** – Reactive oxygen species
- **RT** – Room temperature
- **SI** – Supporting information
- **siRNA(s)** – Short interfering ribonucleic acid(s)
- **TBTU** – O-(Benzotriazol-1-yl)-N,N,N′,N′-tetramethyluronium tetrafluoroborate
- ***t*-Bu** – *tert*-butyl
- **TDC** – Targeting drug conjugates
- λ_{abs} – Longest-wavelength absorption maximum

Abstract

Chemical conjugation of a biomolecule, such as a peptide or a nucleic acid, with a functional molecule, for example a drug or a fluorophore, yields a bioconjugate, a polyfunctional entity that combines the properties of its components to elicit a biological effect. This kind of assemblies have found many successful applications, allowing the precise imaging of biological processes and delivering innovative treatments to the clinic, especially in oncology. The success of this class of constructs is determined by their ability to deliver the payload of choice to specific targets inside a biológical environment thanks to the targeting ability of the biomolecule. In order to have a functional conjugate, the payload must be linked to the targeting unit without altering the structure and function of the latter, preserving its ability to interact with the biological target. As a consequence, bioconjugation reactions must be performed in aqueous environment at physiological pH, avoiding strong reagents that might compromise the integrity of the biomolecule. Moreover, this ligation needs to be stable in physiological conditions while allowing the release of the payload in response to a stimulus that characterizes the biological target (pH, oxidative stress, enzymes).

Boronic acids are a class of biocompatible reagents with the ability to form reversible covalent ligations in water, leading to their successful application in the development of stimulus-responsive bioconjugates.

In this thesis, the 3-hydroxy quinolinone (3HQ) scaffold was studied as potential ligand for boronic acids in bioconjugation conditions, with the objective of using it as a functional unit for the reversible ligation of boronic acids to proteins. This scaffold was optimized for the desired purpose, yielding a structure that was successfully applied for the construction of a non-internalizing fluorescent bioconjugate. The resulting conjugate was tested *in vitro*, confirming its ability to deliver a borylated fluorescent probe to HT-29 cancer cells by targeting the 67 Laminin receptor.

Additionally, the 3HQ motif was explored as multivalent modulator of human phenylalanine hydroxylase (hPAH), a protein whose mutation is the cause of phenylketonuria (PKU). In this study, various assays were performed to assess the ability of 3 series of derivatives to stabilyze the structure of hPAH and to increase its catalytic activity. Finally, the best performing compounds from this study were docked onto the structure of hPAH to gain additional insight on their activity.

Keywords: **Boronic acids, Bioconjugation, Reversible ligations, Heterocyclic scaffolds, Biological chemistry**

Resumo

A conjugação de uma molécula funcional, como um fármaco ou um fluoróforo, com uma biomolécula com propriedades de direcionamento (p.ex. proteína ou ácido nucleico) origina uma estrutura denominada de bioconjugado. Este tipo de conjugados combinam eficazmente as propriedades dos dois componentes, com o objetivo de obter uma função biológica específica e direcionada, podendo esta ser esta terapêutica, imagiológica, teranóstica, entre outras. O sucesso deste tipo de assemblagens depende, como tal, da capacidade da própria biomolécula em exercer eficazmente as suas propriedades de direcionamento para um alvo definido, assim como da capacidade da molécula funcional em manter as suas propriedades desde a administração do conjugado, até ao alvo, onde vai cumprir as suas funções seletivamente. Este processo constitui portanto uma importante vantagem sobre técnicas e terapêuticas sem direcionamento e tem levado a um aumento na investigação das suas aplicações em investigação biológica e na terapia de precisão.

Consequentemente, é imperativo que a ligação da carga (fármaco ou fluoróforo) com a biomolécula não altere nem a sua estrutura nem as propriedades biológicas da biomolécula, a fim de preservar a sua habilidade de interagir com o próprio alvo. Por esta razão, a química utilizada na formação de bioconjugados deve ser compatível com meios biológicos, ou seja funcionar em ambientes aquosos, pH fisiológico e sem a utilização de reagentes que podem comprometer a estrutura da biomolécula. Além disso, a ligação entre os dois componentes do bioconjugado deve manter-se estável nas condições fisiológicas, mas ser reversível em resposta a determinados estímulos que caracterizam o ambiente alvo, de forma a libertarem a molécula funcional inalterada.

Os ácidos borónicos são uma classe de compostos orgânicos que contêm um átomo de boro trivalente, com dois substituintes hidroxilicos e um substituinte alquílico ou arílico. Este tipo de compostos têm várias propriedades adequadas para bioconjugação. Em primeiro lugar, a sua biocompatibilidade e estabilidade em meios biológicos. A reatividade dos ácidos borónicos está relacionada com a acidez de Lewis do átomo de boro que, com a própria orbital *p* vazia, pode aceitar pares de eletrões desemparelhados. Este comportamento permite a interconversão dos ácidos borónicos entre a forma trigonal planar e a forma tetraédrica depois de aceitar um ligando adicional, formando ligações dinâmicas e reversíveis.

O foco desta tese de doutoramento foi o desenvolvimento de novas estratégias para bioconjugação baseadas em ligações reversíveis promovidas por ácidos borónicos. Neste

contexto, a primeira modalidade de ligação explorada foram os ésteres borónicos, que resultam da reação entre ácidos borónicos e ligandos bidentados que podem coordenar o átomo de boro. Este tipo de ligandos bidentados geralmente contém grupos hidroxilo ou aminas para ligar o centro de boro, gerando estruturas que estão em equilíbrio entre a forma trigonal planar e a forma tetraédrica, onde o boro se encontra carregado negativamente.

O primeiro passo deste projeto foi a pesquisa de novas estruturas para a quelação de ácidos borónicos, com base em vários *scaffolds* de produtos naturais. Uma das estruturas moleculares que suscitou a nossa atenção foi a 3-hydroxiquinolinona (3HQ), um biciclo heteroaromático que constitui o esqueleto de vários alcaloides da Vinca. Esta molécula, apesar de simples, possui diversas características interessantes para a nossa pesquisa. Na verdade, esta molécula possui uma estrutura pseudopeptídica, sendo um isóstero do aminoácido glicina, e sobretudo, é conhecida por ser um bom ligando para vários catiões metálicos divalentes. Estudos preliminares sobre a reatividade da 3HQ simples com ácidos borónicos em condições de bioconjugação (meio aquoso, pH fisiológico, temperatura ambiente e alta diluição) mostraram a formação do éster borónico desejado após apenas 5 minutos de reação, confirmando as nossas expectativas.

Seguidamente, realizámos um estudo focando na modificação estrutural da 3HQ com o objetivo não só de otimizar a reatividade com ácidos borónicos, mas também, inserir uma unidade funcional na posição 4 do heterociclo para uma futura conjugação com proteínas. A execução deste estudo implicou o desenvolvimento de uma nova estratégia sintética para a obtenção da 3HQ com um substituinte carboxamida na posição 4, uma vez que as metodologias conhecidas na literatura apresentam limitações em termos de rendimento e degradação dos produtos desejados.

A nova via sintética, descrita nesta tese, tem como base a utilização de um intermediário versátil, o NHS diazo acetato que, através de uma sequência tandem de inserção/expansão de anel com vários derivados da isatina, permite a síntese rápida e eficiente de várias 3HQ com uma unidade de NHS em posição 4. A reação destes intermediários 3HQ-4NHS com aminas primárias e secundárias é capaz de originar uma série de derivados 4-carboxamida 3HQ de uma forma simples e rápida. A síntese de vários 3HQ, com diferentes substituintes tanto no biciclo, como na posição carboxamida, e a sua capacidade em complexar ácidos borónicos em condições de bioconjugação possibilitou uma analise estrutura-atividade de forma a obtermos a estrutura mais eficaz. Esta análise salientou a importância de colocar um substituinte electro-atractor na posição 7 da 3HQ, para aumentar a acidez do grupo fenólico e melhorar a afinidade para ácidos borónicos. Adicionalmente, observou-se um efeito detrimental na presença de

impedimento estérico na posição 4-carboxamida em relação à sua reatividade, tendo esta tendência sido confirmada com a medição das constantes de afinidade de dois derivados da 3HQ por meio de um ensaio de titulação competitiva da alizarina. De facto, neste ensaio, o derivado com uma glicina na posição 4 mostrou uma afinidade de quase 7 vezes superior do que registado para a 4-carboxileucina 3HQ, na complexação do ácido fenil borónico.

Tendo o derivado fluorado e com um substituinte 4-carboxiglicina emergido como a melhor estrutura para a complexação de ácidos borónicos, esta foi posteriormente modificada com a introdução de uma unidade de maleimida. Este composto, graças à sua unidade maleimida, foi utilizado para modificar um péptido modelo no seu resíduo de cisteína, gerando um derivado péptido-3HQ. No entanto a reatividade deste péptido modificado com a 3HQ resultou ser sub-ótima na complexação de ácidos borónicos em condições de bioconjugação, provavelmente devido ao forte impedimento estérico que a cadeia do péptido impõe na estrutura final.

Visando ultrapassar esta limitação encontrada na reatividade de péptidos modificados com a plataforma 3HQ em complexar ácidos borónicos simples, uma abordagem alternativa foi empregue de forma a aumentar o ganho energético da ligação entre 3HQ e ácidos borónicos. Com este objetivo, a plataforma 3HQ foi aplicada para direcionar e estabilizar a formação de iminoboronatos em resíduos de cisteína *N* - terminais.

Esta nova abordagem, baseada na utilização do mesmo derivado 3HQ-glicina-maleimida desenvolvido previamente, tem como objetivo a solução de duas importantes limitações de iminoboronatos clássicos: a estabilidade e a seletividade. Os iminoboronatos são derivados da reação de uma amina primária e um ácido borónico com um substituinte carbonilico na posição 2, ou seja, iminas estabilizadas por uma ligação dativa entre o átomo de azoto e o boro. Este tipo de assemblagens obtiveram sucesso no contexto da bioconjugação devido à sua cinética de formação rápida e biocompatibilidade, juntamente com o seu carácter reversível em presença de vários estímulos (açucares, dopamina, glutationa, pH acídico). Infelizmente, algumas aplicações não são apropriadas ao uso de iminoboronatos devido à sua estabilidade limitada em meios biológicos complexos, assim como a impossibilidade de controlar o local de modificação, uma vez que todas as aminas expostas podem ser modificadas com esta abordagem.

A viabilidade da utilização da plataforma 3HQ para promover a formação de iminoboronatos estabilizados em cisteínas *N*-terminais foi testada em varias reações modelo, utilizando em primeiro lugar a cisteína como parceiro reacional e, subsequentemente, vários dipéptidos com cisteínas *N*-terminais. As experiências modelo

confirmaram as nossas hipóteses, tendo os 3HQ-iminoboronatos formados por meio desta nova abordagem resultado em conjugados mais estáveis que os iminoboronates clássicos, inclusivamente na presença de pH acídico, glucose e albumina. É importante realçar que estes ainda possuem um carácter de reversibilidade na presença de glutationa. Adicionalmente, a reação com péptidos modelo confirmaram a seletividade deste método, uma vez que na presença de vários grupos amina disponíveis para formar iminoboronatos, a seletividade para a cisteína *N* - terminal modificada com a plataforma 3HQ foi observada. A formação do 3HQ-iminoboronato foi investigada e caracterizada por meio de espectroscopia RMN de azoto, boro e hidrogénio, confirmando a presença de uma estrutura de iminoboronato estabilizada por interação com a unidade de 3HQ.

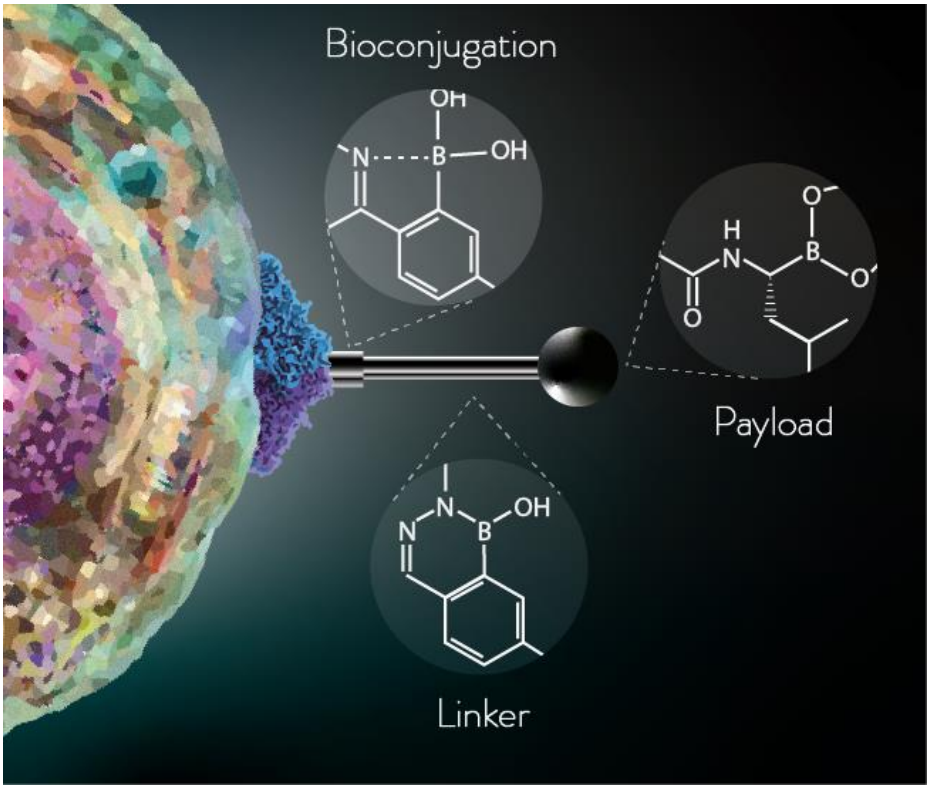
Esta plataforma inovadora foi aplicada na construção de um conjugado não-internalizante para a entrega de sondas fluorescentes boriladas em células cancerígenas HT29. Este conjugado foi sintetizado utilizando um fragmento de laminina como unidade de direcionamento, o que permitiu validar o receptor deste péptido, 67LR, como alvo viável para a entrega de moléculas ativas boriladas em células de carcinoma colo-rectal.

No último capitulo desta tese, a síntese de uma pequena bilioteca de 3-hidroxi quinolinonas foi aplicada no desenvolvimento de moduladores da atividade da fenilalanina hidroxilase, a enzima responsável pelo metabolismo da fenilalanina. Esta enzima é suscetivel a várias mutações genéticas, o que causa diversas disfunções resultantes da acumulação tóxica de fenilalanina no organismo. Como estas mutações causam uma falha na estrutura e na atividade da proteína, a abordagem farmacológica neste sentido consiste no desenvolvimento de pequenas moléculas, chamadas chaperonas, com a habilidade de restaurar a funcionalidade nativa desta enzima. Neste contesto, três famílias de 3HQ foram sintetizadas e testadas pela sua habilidade em estabilizar a estrutura correta da fenflalanina hidroxilase e, ao mesmo tempo, aumentar a sua atividade enzimática. Entre os vários derivados, os 4 compostos que deram os melhores resultados foram selecionados para estudos de docking molecular no centro ativo da enzima, o que permitiu a racionalização dos resultados obtidos.

Palavras-chave: : Ácidos boronicos, Bioconjugação, Ligações reversiveis, Estructuras heterocíclicas, Química biológica

Chapter I

State of Art - Boronic Acids in Bioconjugation



Abstract

The construction of therapeuthic conjugates has been one of the most powerful tools that entered the landscape of chemical biology in the last decades, allowing the precise imaging of biological processes and delivering innovative treatments to the clinic. In this framework, boronic acids have found success as functional components of bioconjugates due to their reversible nature, which makes them ideal for the development of stimulus-responsive constructs. Even though many biocompatible reactions based on boronic acids are known, the investigation in this area is still thriving thanks to the vast possibilities for tuning this functionality for different applications, yielding constructs with specific chemical behaviour.

I.1. Overview on Boronic Acids

Boron is the fifth element of the periodic table, a semi-metal with one less proton than carbon, making it an element with various interesting properties and a unique chemistry. The defining characteristic of boron's chemistry is the empty *p* orbital that this element possesses, which allows it to establish reversible covalent bonds with oxygen and nitrogen nucleophiles. This Lewis acid-base reaction results in the central boron atom interconversion between an uncharged trigonal planar structure and an anionic tetravalent species, with a geometry and an isoelectronic structure that are equivalent to a neutral sp^3-hybridized carbon (**Figure 1**). Also, the similarity of boron to carbon is particularly evident in aromatic structures, as B-O and B-N bonds are isoelectronic, isovalent and isosteric to C=N and C=C bonds, respectively.[1–4]

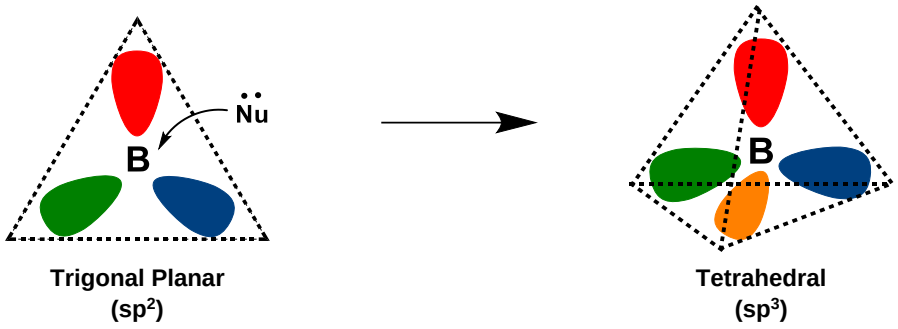


Figure 1 – Conformational shift of the Boron atom

In nature, boron is mostly found as boric acid, either in its mineral form sassolite, which is present in volcanic areas, or as a salt in seawater. Boric acid is a harmless (LD_{50} 5.14 g/Kg for rats)[5] compound that is naturally uptaken by plants as nutrient, hence several natural products deriving from its metabolism have been isolated. Even though many of these products are known (*e.g.* boromycin, tartrolon B, borophycin and AI-2) and have shown interesting properties (*e.g.* antibiotic, anti-HIV, anticancer, antimicrobial, antibacterial, antihepatitis), the labile nature of boron complexation makes the isolation of such entities difficult, and the disassembly of these structures during extraction and purification processes leaves us with just a partial understanding of boron's important role in nature.[6] Throughout the years, several families of boron-based compounds have found success in different areas, such as therapeutic applications, imaging, catalysis, hydrogen storage and many others but, for the scope of this thesis, the focus will stay on boronic acids (BAs), their properties and their application in bioconjugation.

Boronic acids are trivalent boron compounds bearing two hydroxyl substituents and an alkyl or aryl moiety which were reported for the first time in 1860[7] but only gained popularity once hydroboration methodologies made structurally diverse derivatives synthetically available.[8,9] Given their Lewis acid character, boronic acids have the ability to form dynamic covalent bonds with Lewis base donors, both in organic and aqueous media, which is one of the most important aspect of this element's unique reactivity. In time, BAs became one of the most prevalent reagents in modern synthesis, enabling the effective construction of different types of C-C and C-heteroatom bonds.[10]

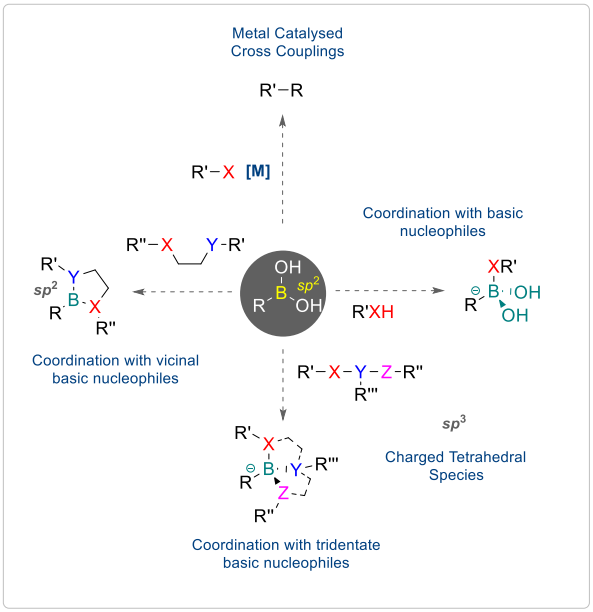


Figure 2 – General reactivity of Boronic acids

Apart from their use as reagents, BAs also exhibit unique properties that are now well appreciated in the areas of medicinal chemistry, chemical biology and material sciences, as BAs can bind nucleophiles in aqueous conditions to generate various boronate species. In recent years, these reactions became a powerful synthetic strategy to construct self-organizing systems, sensors and many different types of functional materials.[11,12]

Due to their mild Lewis acidity, BAs are typically uncharged at physiologic conditions and retain the ability to reversibly coordinate oxygen-based nucleophiles, namely at the active site of proteases.[13] The discovery that BAs may be used as transition-state analogue protein inhibitors triggered a burgeoning interest among the medicinal chemistry community that was using boronated compounds mostly for their ability to emit alpha particles under neutron bombardment (boron neutron capture therapy).[14–16]

Notably, boronic acids nowadays occupy an important niche in medicinal chemistry, with FDA-approved representatives in various therapeutic areas as well as several candidates in the pipeline.

I.2. Therapeutic Bioconjugates

Bioconjugates are multifunctional constructs in which biomolecules like peptides and proteins, small vitamins or nucleic acids are endowed with the properties of specific payloads (e.g., biological activity and fluorescence). This approach harnesses the targeting potential of biomolecules, exploiting the ones that are selectively recognized by receptors associated with a disease to deliver the desired cargo to the affected area. These entities have been successfully applied in the clinic, mainly in the field of cancer therapy, as they offer the possibility to selectively deliver cytotoxic drugs to tumoural tissues, a valuable feature in a field in which target engagement is crucial to avoid severe side-effects. In this framework, antibodies have been the first biomolecules to see success as targeting units, as demonstrated by the clinical approval of several antibody-drug conjugates (ADCs).[17–19] With the success gained from these initial applications, several other entities have been explored as targeting units, expanding the toolbox of therapeutic drug conjugates. Independently from the nature of the construct, the general structure of therapeutic bioconjugates can be reduced to 4 fundamental components: the targeting unit, the warhead, the linker and the payload (**Figure 3**).

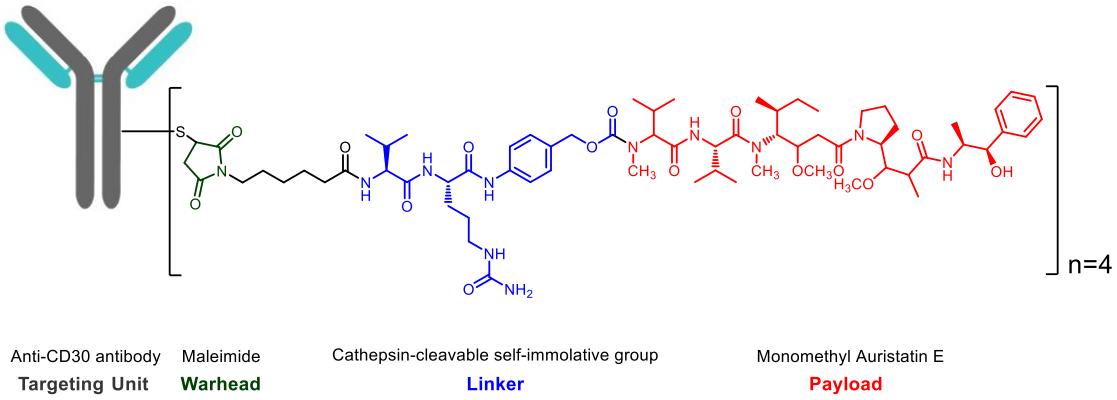


Figure 3 – Structure and functional components of the FDA-approved ADC Brentuximab Vedotin

As already mentioned, the targeting unit is constituted by a biomolecule (protein, vitamin, small peptide) that is selectively recognized by receptors that are overexpressed in the diseased area, and thus is able to direct the conjugate towards the desired location.[20] The payload is the cargo that will be released by the conjugate, which can have a therapeutic

(drugs, radionucleotides) and/or diagnostic (contrast agents, fluorescent probes) function.[21,22] The synthesis of bioconjugates by direct chemical installation of the molecular payload onto the biomolecule is often impossible due to the lack of suitable chemical functions that may be used in the bioconjugation process. In order to conjugate the payload to the targeting unit, it is necessary to have a moiety that is able to selectively modify proteins in biocompatible conditions, indicated as the warhead. This unit occupies a unique chemical space in the framework of organic reagents, as several properties are required for a successful warhead. First of all, the reactivity of chemical warheads needs to be compatible with biological conditions (aqueous environment, neutral pH, high dilution) and cannot perturb the structure and function of the protein upon modification. It is also imperative that the modification happens in a selective manner, in order to control the attachment site and the number of units appended to the protein, with the objective of obtaining homogenous conjugates.[23–27] Finally, the linker is the unit that ensures the proper delivery of the payload. This part of the conjugate, in fact, needs to be a stimulus-responsive moiety that is be stable in physiological conditions, and then disassemble in response to a certain trigger associated with the targeted environment (pH, oxidative stress, enzymes), releasing the payload in its active form.[28,29]

I.3. Boronic Acids in Bioconjugation

As the core functionality of a bioconjugate lies in its ability to selectively release the payload in the desired environment, the chemistry used to connect the cargo to the targeting unit is intimately related to clinical effectiveness. Namely, the function of the biomolecule should not be significantly altered by the attachment of the payload and the conjugate stability should be maintained in circulation until it reaches the diseased site, where the different chemical environment will trigger the payload activation. In the framework of improving the different aspects of bioconjugates' construction, many efforts have been pointed at discovering new conjugation strategies for better warheads and linker technologies, with the objective of developing more homogeneous and stimuli-responsive conjugates. In recent years, boronic acids' unique reactivity, dynamic coordination properties under physiological conditions and biocompatibility brought them into the limelight of chemical biology. In fact, this reversible covalent functionality enables the synthesis of bioconjugates with suitable properties to be used in the fields of bioconjugation, live-cell-imaging, theranostics and selective drug delivery. Also, having already two representatives in the clinic as cancer treatment and being a useful prodrug moiety, BAs constitute valuable payloads as well. In the framework of this thesis, this

chapter will give an introduction on the use of boronic acids as payloads and of two of the most used bioconjugation technologies involving boronic acids, namely boronate esters and iminoboronates.

I.3.1 Boronic Acids as Payloads

Thanks to protocols such as the Miyaura borylation, the boronic acid function can be synthetically introduced in the structure of organic molecules, vastly expanding the chemical space that this class of compounds can occupy.[30] This has not only expanded the research on BA-based drugs, which have already shown strong potential, but also on their incorporation in pre-existing structures to endow them with novel properties.[31]

Over the past decades, many boron-containing bioactive molecules have been developed, prevalently as protease inhibitors.[32] Presently, more than 400 protein structures have been crystallized with BAs and are available in the Protein Data Bank.[33] The analysis of these structures indicates that the activity of BAs is mainly related to the inhibition of serine proteases, as they are able to form a tetracoordinated boronate complex by coordination with the hydroxyl side chain of the active serine residue, mimicking the structure of the natural transition state. Despite the focus on serine proteases, BA's ability to coordinate nucleophilic residues makes them able to also bind threonine[34] and histidine.[35–38] Other nucleophilic residues, such as lysine, arginine, tyrosine and cysteine, could represent potential targets for the development of BA-based inhibitors, but so far have not been reported. Despite showing efficacy *in vitro*[4], only few drugs containing the BA moiety have made it to the clinic, namely the proteasome inhibitors Bortezomib (Velcade®)[39] and Ixazomib (Ninlaro®)[40] for the treatment of multiple myeloma, the antifungal agent Tavaborole (KERYDIN®)[41], Crisaborole (Eucrisa®)[42] for the topical treatment of eczema and the inhibitor of β-Lactamase, Vaborbactam[43] (**Figure 4**).

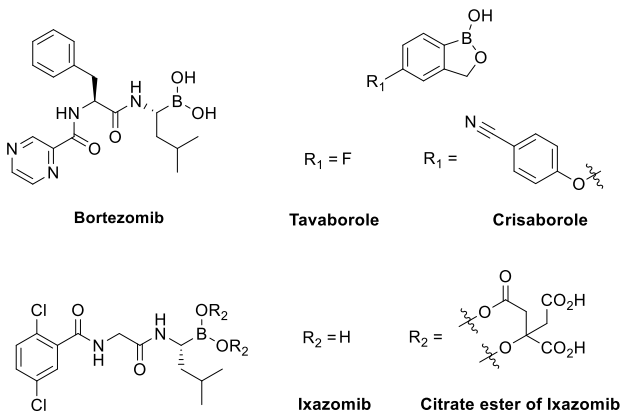


Figure 4 – Structure of FDA-approved drugs containing the BA moiety

The main problem in translating the positive *in vitro* results of BA-based drugs to the clinic is the poor pharmacokinetic profile of boronic acids in their open shell form, as they are able to interact non-specifically with different naturally-occurring nucleophiles.[44] One of the main strategies to avoid this pitfall is the derivatization of the boronic acid moiety as a boronate ester, shielding its reactivity. The comparison between the two FDA-approved proteasome inhibitors Bortezomib and Ixazomib is an excellent example to illustrate how this modification affects the pharmacokinetics of BA-based drug. Both drugs bind the same biological target, the β5-subunit of the 20S proteasome with comparable IC_{50} values in the low-nanomolar range. However, Bortezomib, which was approved in 2003, is used in the clinic in its simple BA form and requires intravenous or subcutaneous administration.[45] On the other hand, Ixazomib, approved in 2015, is administered with its BA moiety derivatized as citrate ester, thus restraining its activity and improving the pharmacokinetic properties of the drug. Thanks to this functionalization, Ixazomib became the first orally-available proteasome inhibitor, bringing a substantial improvement to patients' quality of life.[46]

Incorporating the BA moiety in heterocyclic structures, such as benzoxaboroles, offers another possibility to address BA's pharmacokinetic issues, as shown for Tavaborole and Crisaborole. In fact, recent reports demonstrate that many B-N heterocycles are able to improve the stability of boronic acids without interfering with their selectivity and inhibitory activity.[47].

Another interesting property of boronic acids is their susceptibility towards reactive oxygen species (ROS). In fact, aromatic boronic acids can be converted to phenols in the presence

of hydrogen peroxide or other ROS, following an analogous mechanism to the one of the oxidation step of the Brown hydroboration-oxidation reaction. Since the concentration of ROS is abnormally high in both inflammated and tumoural environment, BAs are useful for the development of prodrugs and imaging systems that can be applied in these contexts.[48] This unique reactivity has led many research groups to replace phenol moieties with boronic acids in drugs and fluorescent probes alike, obtaining ROS-responsive structures. One of the first instances of this approach comes from Lu and co-workers,[49] which replaced the phenol of SN-38, the active metabolite of the FDA-approved drug Irinotecan, with a BA, obtaining a ROS-activatable prodrug. (**Figure 5A**). The oxidation of this molecule was shown to happen with H_2O_2 concentrations as low as 0.5 mM, which led to better *in vivo* results than Irinotecan itself in glioblastoma xenograft mice models. As inflammated tissues also present high concentrations of ROS, Kang and co-workers reported the synthesis of a borylated precursor of *p*-hydroxybenzyl alcohol, an anti-inflammatory agent. Derivatization of *p*-hydroxybenzyl alcohol as a hydroxymethyl boronic ester improved its efficacy in the treatment of ischemia-reperfusion injury by not only being selectively activated at the desired location, but also by increasing its circulation half-life.[50] Exploiting the reactivity between boronic acids and reactive oxygen species, many research groups have developed ROS-activated turn-on fluorescent probes. Wang and co-workers, for example, developed a BA derivative of coumarin, a highly fluorescent molecule that bears a phenol group that is crucial for its fluorescence properties. The borylation of coumarin, in fact, turned off its fluorescent properties, which were restored in a concentration-dependent fashion by incubation with H_2O_2 (**Figure 5B**).[51] This approach[52–57] has been adopted by many other research groups that developed a plethora of borylated probes for the imaging of various ROS-related phenomena.[58–67] Boronic acid oxidation reactions can also be used as a trigger for initiating a self-immolative cascade for the libertation of active molecules. For example, Cohen and co-workers developed various systems in which the conversion of a BA into phenol triggers a self-immolative cascade that disassembles BA-based linkers (**Figure 5C**).[68]

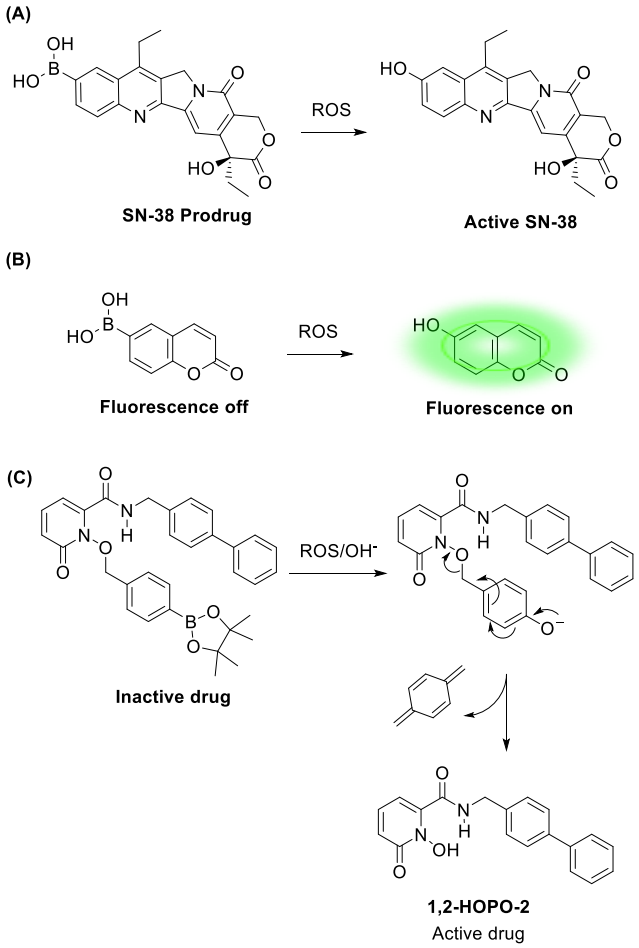


Figure 5 - ROS-mediated activation of prodrugs/quenched probes: SN-38 prodrug is transformed into SN-38 in the presence of ROS (A); A quenched coumarin derivative is activated in the presence of ROS (B); ROS trigger a self-immolative cascade to release the active drug (C).

These platforms have been employed both for the construction of prodrugs, as well as delivery systems in the construction of targeted drug conjugates. A more complex system was reported by Kim and co-workers that, combining some of the previously illustrated concepts, reported the synthesis of a hybrid moiety constituted by a coumarin unit and SN-38 for theranostic applications.[69] Incubation of this compound inside cancer cells led to the H_2O_2-mediated oxidation of the borylated coumarin which, in turn, triggered an electron cascade that ultimately released SN-38 and coumarin, providing localized fluorescence and cytotoxic effect at the same time. Despite the many technologies developed in this field, there is, however, still one fundamental challenge to be addressed, which is the development of BA probes that are ROS-specific. In fact, since the various reactive oxygen species are endowed with different roles in living organisms, it is valuable to devise

molecular platforms that can selectively respond to one specific reactive species. Until now, few BA-based methods have shown selectivity for a single species, though in the future this approach could constitute an important tool to visualize ROS signalling, understand its biomolecular basis and exploit it for more accurate drug delivery.

Moreover, boronic acids have been reported to passively permeate cells with remarkable efficiency, thanks to their ability to form transient interactions with the several glycans commonly present on cell surfaces. This behaviour opened the door to several applications, promoting the internalization of several bioactive molecules and functional systems.[70,71]

Given their properties, boronated payloads are valuable units in bioconjugation, combining stimulus-responsiveness, biocompatibility and versatility. Furthermore, as boronic acids are prone to react promiscuously with endogenous biomolecules, their insertion in bioconjugates is particularly useful to limit this behaviour and improve their pharmacokinetic properties.

I.3.2 Boronate Esters

As previously mentioned, in order to have a functioning bioconjugate, the properties of the biomolecule should not be altered by the attachment of the payload. This is a non-trivial challenge, as the majority of reaction known in organic synthesis require conditions that are not compatible with the preservation of the structure of biomolecules (i. e. high temperatures, organic solvents, oxidative/reductive conditions, extreme pH values, etc.). To address this issue, various biocompatible reactions were explored, leading to the development of several means of building covalent bonds onto biomolecules in aqueous conditions. Biocompatibility is a necessary feature for the building of a successful chemical warhead, however, it is not sufficient, as it is also important to have control over the site and the number of modifications performed, in order to produce homogenous conjugates that have rationalizable properties and batch-to-batch reproducibility.

Over the last years, several developments in this field have been reported, however, most of the chemical warhead used in bioconjugation were designed to yield covalent ligations, relying on the linker unit for the stimulus-responsive cargo release. Despite their potential to prepare conjugates that may respond to a predetermined stimulus, the use of reversible ligation methods has been rather overlooked in this area. The use of BAs as warheads for biomolecule functionalization is particularly attractive as it can be used to design

responsive bioconjugates, as BAs can selectively establish reversible covalent bonds with a diverse set of biological entities under biocompatible conditions.[72] Using BAs as warheads adds value to the construction of bioconjugates, as their reversible nature renders the use of a linker unit for the selective release of the payload superfluous, allowing for the construction of simpler, more compact conjugates.

One of the first functionalities to be explored in this context are boronate esters (BEs), which are the product of the nucleophilic attack of an alcohol onto a boronic acid. Notably, the ability of boronic acid to react with various naturally-occurring nucleophiles in aqueous conditions is particularly pronounced towards the complexation of vicinal hydroxyl groups, such as carbohydrates and catechols, yielding cyclic BEs with reaction rates of 10^2-10^3 $M^{-1}s^{-1}$.[73–76] Given the relatively fast formation kinetics and the good stability in aqueous conditions, boronate esters have found widespread application as warhead, linker and recognition units.

The equilibrium of boronate esters formation in water has been object of many studies, as it is a rather complex multi-step process, some of which are still matter of investigation.[77–79] In the context of bioconjugation, boronate esters are appreciated for their dynamic nature, as their formation-disassembly equilibrium depends on both pH and competition effects. BE formation is favoured at basic pH, as it allows for the formation of the tetrahedral boronate anion, which contributes to the stability of the product (**Figure 6**).[80] On the other hand, lower pH values trigger the hydrolysis of these constructs, allowing to build reversible conjugates that can respond to acidic conditions by disassembling.[81] As low pH values are found in tumoural environment, as well as in cytosolic compartments, the acid-mediated hydrolysis of BE constructs constitutes a useful property for selective payload release.

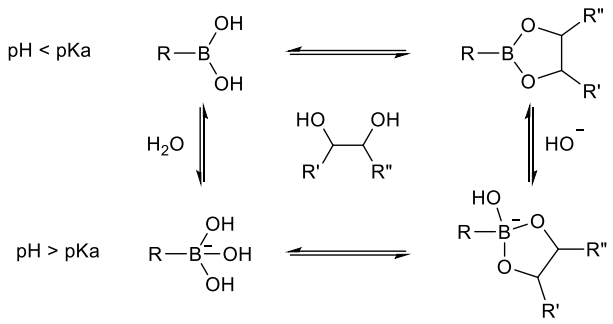


Figure 6 - pH-Dependent formation equilibrium of a boronic ester in aqueous solution.

Moreover, since boronate ester formation is a dynamic equilibrium, BEs can be disrupted by the presence of other BA ligands, as the competition between the available ligands will favour the formation of the most stable BE. Hence, even relatively stable BEs can be displaced by the interaction of the BA with a stronger ligand, which represents yet another possible release mechanism for payloads installed through this functionality. In general, the diols that show the best affinity for BAs are cyclic or hindered cis-1,2 diols, which yield rather stable complexes.[81] Given their reactivity with diols, boronic acids have also been widely employed for the detection and quantification of carbohydrates, however, such applications will not be discussed in this chapter.[82–84] In the field of protein modification, the interaction between boronic acids and carbohydrates has been exploited for the traceless labelling of glycoproteins, based on transient BE formation. As described by Lin and co-workers, the attachment of a clickable alkyne handle on a phenyl boronic acid moiety via a reactive tosyl spacer allowed the installation of the handle onto the protein by proximity-driven S_N2 reaction. This approach relies on the formation of a boronate ester between phenyl boronic acid (PBA) and the carbohydrates of the glycoproteins, which primes the tosyl-appended clickable handle to be attacked by a nucleophilic residue of the protein. Displacement of the BE *via* incubation with glycerol or sorbitol, which have higher affinity for PBA than the sugars of the glycoprotein, ensures the removal of the boronated warhead, affording the traceless modification of the target protein (**Figure 7**). The protein obtained after this process can easily be used for pull-down essays thanks to the alkyne handle, allowing for the discovery of new glycoproteins and studying their interaction with other proteins.[85]

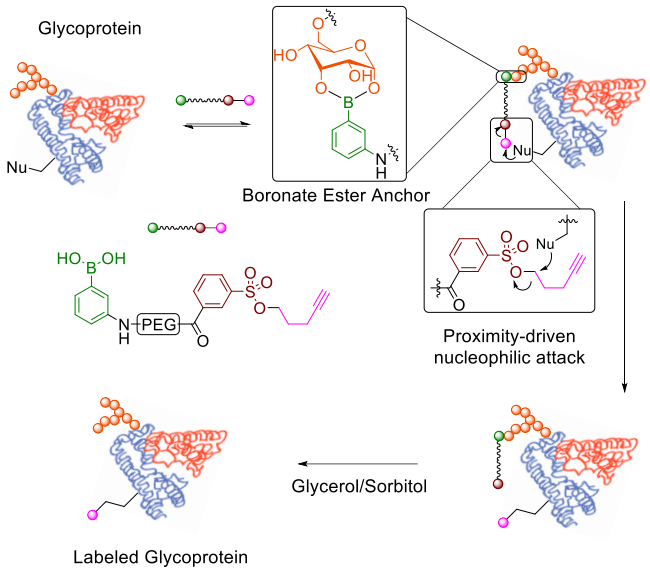


Figure 7 – BE-promoted attachment of a clickable handle on a glycoprotein through a proximity-driven S_N2 reaction.

The functionalization of proteins by attachment of BAs onto amino acids that possess hydroxylated side chains would be a convenient bioconjugation strategy, as it would allow to target serine, threonine and tyrosine residues. However, the formation of a single B-O bond is not sufficient to yield stable linkages in complex biological media, hence direct conjugation to these residues is not possible. Turning this limitation into an advantage, Schepartz and co-workers aimed at targeting tetraserine motifs, which are a recurring theme in protein structures, as more than 200 proteins within the human proteome contain the SSPGSS sequence. With this objective in mind, they leveraged a bis-BA pro-fluorescent rhodamine tag, named RhoBo, whose fluorescence is turned on upon complexation with proteins bearing the tetraserine motif (λ_{max}= 520 nm). This non-toxic small molecule not only exhibited cell-permeability, but also showed remarkable selectivity, as no side-reactivity was observed towards monosaccharides and other endogenous molecules (**Figure 8**).[86]

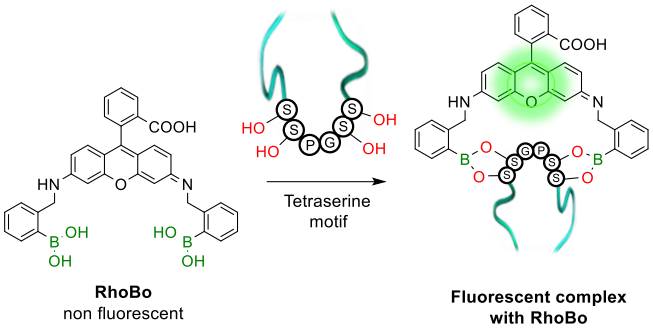


Figure 8 - Representation of RhoBo fluorescent complex upon tetraserine sequence recognition.

Similarly, tyrosine is equally unable to form stable 1:1 complexes with BAs, despite having a more acidic side chain. Tirelli and co-workers, however, described a methodology to circumvent this limitation, by leveraging the action of the tyrosinase enzyme. In fact, this enzyme is able to oxidize the side chain of tyrosine to a catechol, which displays higher affinity for boronic acids, leading to site-selective boronate ester formation.[87] Proteins that do not present solvent-accessible tyrosine residues in their native structure can still be modified with this approach using the Hemagglutinin-derived HA-tag to introduce exposed Tyr groups onto the target protein.[88] Biomolecule modification through boronate esters is not limited to proteins, as the ribose unit of RNA offers another potential handle for the construction of such structures. In fact, boronic acid's interaction with RNA has been applied, over the years, not only in separation and purification techniques, but also to design different bioconjugates with potential therapeutic use.[89–91] Specifically, one of the main appeals of BA-modified RNA units is their improved cross-membrane transport capacity compared to unmodified RNA chains.[92] Also, as ribose is not a strong BA chelator (K_{eq}= 24 M^{-1} for PBA), BEs built on RNA fragments can easily be displaced by competing ligands, allowing for the construction of responsive structures that disassemble in the presence of high concentrations of certain biomolecules. Taking advance of displacement-triggered BE hydrolysis, Kataoka and co-workers developed a PBA-modified poly-lysine co-polymer with a C-terminus PEG chain that is able to reversibly bind siRNAs and form micellar assemblies through BE formation and electrostatic interactions. The structures obtained with this approach were reported to cross cell membrane efficiently and then disassemble due to the high intracellular concentration of ribonucleosides that compete for BA chelation, leading to BE displacement and siRNA release.[93] Aside from applying boronate ester formation to native biomolecules, modern biotechnology approaches allowed to incorporate non-natural amino acids into the

structure of proteins, opening new opportunities for this area of investigation. In 2008, Schultz and co-workers leveraged the amber stop codon strategy to incorporate an artificial boronated amino acid (*p*-boron phenylalanine) inside the sequence of various proteins. The resulting borylated proteins not only displayed reactivity in classic boron organic chemistry (oxidation, reduction, Suzuki coupling, etc.) but also gained the ability to form BEs, allowing for traceless purification and detection (**Figure 9**).[94]

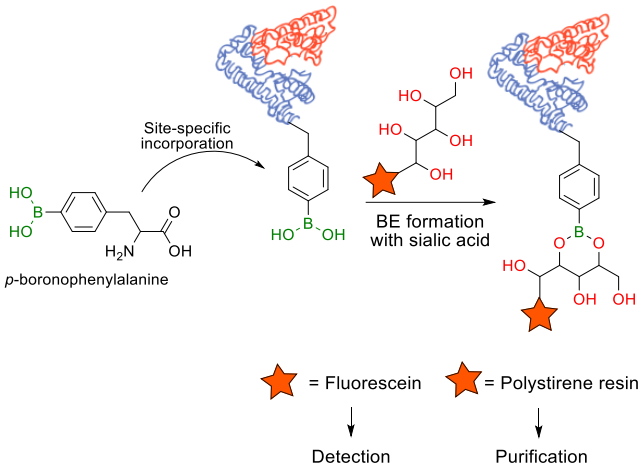


Figure 9 - Incorporation of a *p*-boronophenylalanine non-canonical amino acid into proteins allows their detection and purification through BE formation.

Following on their intial report, the authors further explored this approach and generated antibodies in *E. coli* that incorporated *p*-boronophenylalanine, which gained the ability to form reversible BEs with acyclic glycans.[95] BA incorporation into proteins also opened the door to the development of fluorescent protein sensors for H_2O_2[96] and peroxynitrite[97] that leverage the unique reactivity of BAs towards oxidizing species. Despite the many successful applications reported, boronate esters based on carbohydrates and other naturally-derived ligands often lack the stability to be employed in complex biological settings. Aiming at addressing this limitations, Hall and co-workers developed a click semi-bioorthogonal system for BE formation, employing a highly strained diol to form stabilized adducts. In fact, by tuning the structure of the BA-diol pair, they obtained a system, comprised of nopoldiol and 2-methyl-5-carboxyphenylboronic acid, that yields complete BE formation within minutes at low micromolar concentration in aqueous medium (ca. 7.7 M^{-1} s^{-1} and Keq≈ 1.2 × 105 M^{-1}).[98] With this approach, it was possible to modify thioredoxin and albumin with the BA ligand, which then reacted with a nopoldiol-bearing fluorescent probe under physiological conditions, yielding the corresponding fluorescently labeled proteins (**Figure 10A**). Bringing this concept further, the authors later described a

synergistic “click” boronate/ thiosemicarbazone platform that couples the fast nopoldiol BE formation, with hydrazine formation, yielding irreversible bioorthogonal conjugation that could be performed in live cells (**Figure 10B**).

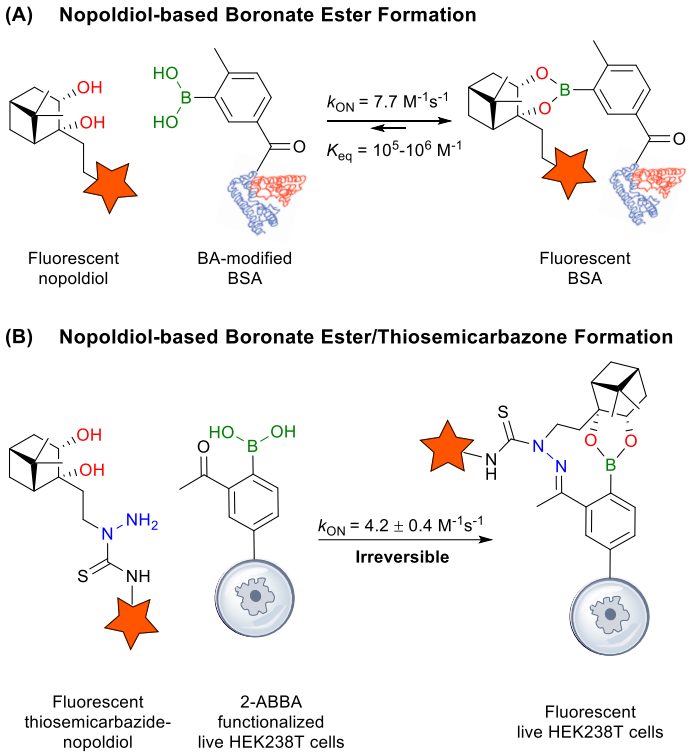


Figure 10 - Formation of stable BEs; (A) The highly strained nopoldiol binds tightly but reversibly to PBA; (B) The introduction of a synergistic thiosemicarbazone renders the BE with nopoldiol irreversible.

I.3.3 Iminoboronates

The reactivity of boronic acids with carbohydrates has been long known and a mainstay in both analytical and biological chemistry methodologies, as the applications reported so far demonstrate, but their application towards native protein modification was limited to glycoproteins. However, in 2012, Gois and co-workers reported the formation of BA-stabilized imines, called iminoboronates, arising from the reaction between primary amines and 2-carbonyl benzene boronic acids (2-CBBAs) in bioconjugation conditions.[99] This reaction was applied to the modification of lysine side chains and the *N*-terminus of many proteins, yielding the desired iminoboronates with fast kinetics. Bioconjugation with primary amines, either N-terminal residues or ε-amine side chains of lysines, is a common strategy for protein modification, given the nucleophilicity and solvent accessibility of these

moieties. Primary amines can be modified with electrophiles such as activated esters, sulfonyl chlorides, isocyanates or isothiocyanates.[100] Alternatively, primary amines can be modified thorugh the generation of imines with carbonyl compounds, such linkages are, however, very susceptible to hydrolysis, hence a subsequent reduction step is required (**Figure 11A**).[101,102]

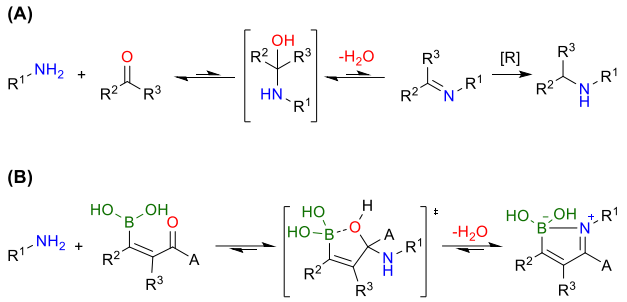


Figure 11 - Strategies for amine condensation with carbonyl compounds. (A) Imine formation, reversibility in aqueous media and reductive amination. (B) Iminoboronate formation and reversibility in aqueous media.

All of these approaches have the caveat of being irreversible methodologies for amine modification, yielding structures that are not able to respond to stimuli. In their pioneering study, Gois and co-workers reported that iminoboronates can be formed over a wide range of pH (6-9) and resulted stable over a week thanks to the stabilizing effect fo the B-N dative bond (**Figure 11B**). As already mentioned, the uniqueness of boron chemistry lies in its empy *p* orbital, which plays a pivotal role in iminoboronate formation, accepting a pair of electrons from nitrogen, thus stabilizing the final product. The impact of boron in iminoboronate stabilization has been investigated by DFT calculations, which support a mechanism in which the BA acts as an intramolecular Lewis acid, catalyzing the attack of the amine onto the carbonyl group, favouring the elimination of water to generate the imine and, more importantly, enabling the B-N bond formation (dB–N = 1.71 Å). This bond formation enhances the stability of the product (ΔE = -7.4 kcal mol^{-1}) and contributes to an overall favourable energy balance of -10.0 kcal mol^{-1}. Notably, the installation of iminoboronates is a reversible reaction, as this functionality can be removed in the presence of various biomolecules, such as dopamine, glutathione or fructose, offering a valuable stimulus-resposive release mechanism. Further studies by Anslyn and co-workers reported the formation of a zwitterionic solvent-insertion structure for iminoboronates, that can contribute to their stability.[103] Among 2-CBBAs, ketone derivatives afford the more stable iminoboronate than their aldehyde counterparts, as emerged from the direct comparison between 2-formyl benzeneboronic acid (2-FBBA) and

2-acetyl benzeneboronic acid (2-ABBA). For this reason, 2-ABBA was used in the successful functionalization of the primary amino groups exposed in the hormonal neuropeptide somatostatin and proteins like lysozyme, cytochrome c, ribonuclease A, and myoglobin (**Figure 12**).

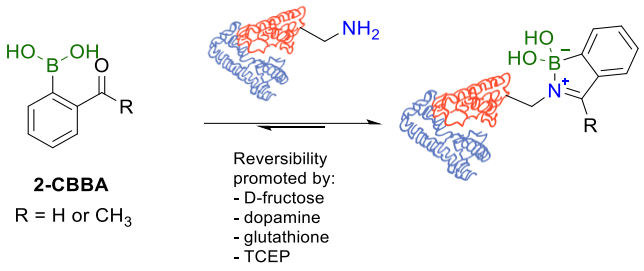


Figure 12 - Iminoboronate protein modification and reversibility in aqueous media

Following on their initial report regarding protein modification with iminoboronates, Gois and co-workers applied this technology to the construction of a variety of bioconjugates with therapeutic or imaging applications (**Figure 13**). In fact, the iminoboronate technology was first applied to prepare fluorescent folate derivatives which were used to selectively label NCI-H460 human non-small lung cancer cells, which are known to overexpress folate receptors.[104] After validating the ability of iminoboronate-based folate conjugates to delivery their payload to non-small lung cancer cells, the same authors prepared a folate-paclitaxel small molecule drug conjugate which elicited an IC_{50} of 20.7 nM against NCI-H460 cancer cells.

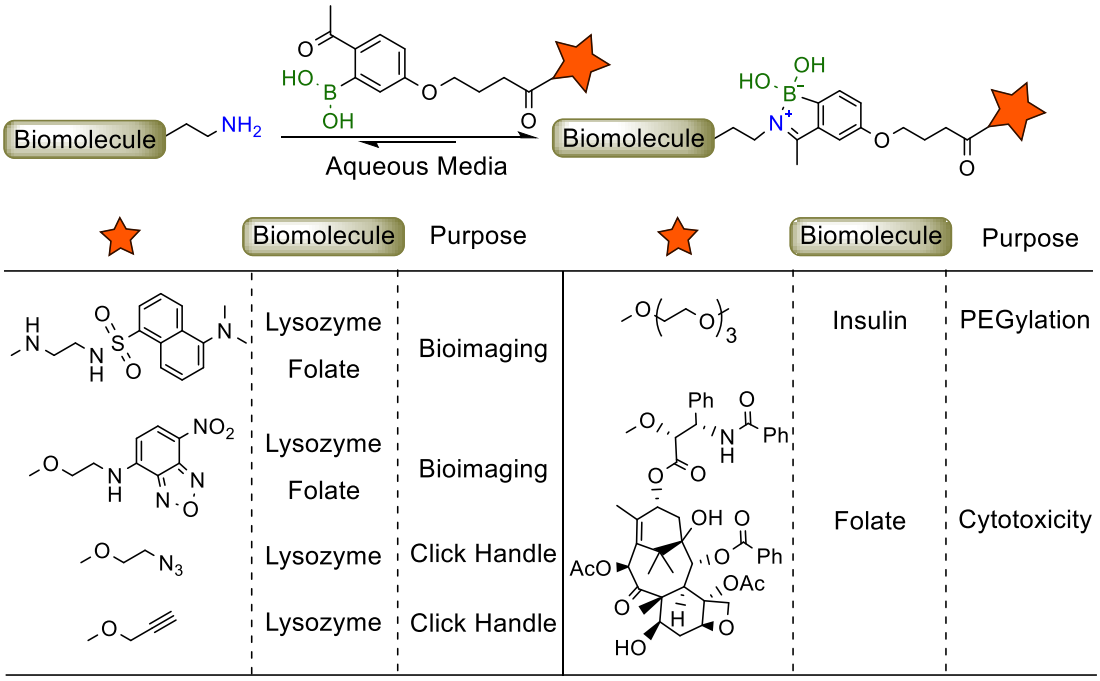


Figure 13 - Iminoboronate-based bioconjugation to anchor fluorophores, clickable handles, PEG derivatives and paclitaxel.

Additional developments on this system lead to the iminoboronate-mediated modification of lysozyme with clickable handles for Cu-assisted or strain-promoted azide-alkyne cycloaddition. Leveraging the fructose-promoted hydrolysis of iminoboronates allowed for the development of a stimulus-responsive platform for selective Insulin release. In this context, polyethylene (PEG) units functionalized with 2-ABBA were used to form iminoboronates onto the surface of Insulin, rendering it inactive while simultaneously increasing its circulation half-life. The modification of Insulin was reverted in the presence of fructose, which triggered the release of the active biomolecule, constituting a valuable approach for the mobilitization of insulin in response to the increase of sugar levels in the blood.[105] Another ingenous application of iminoboronate technology comes from the R&D division of AstraZeneca, which leveraged the favourable thermodynamics of iminoboronate formation to increase the affinity and on-target residence of a protein-protein interaction inhibitor. As reported, the insertion of a 2-FBBA warhead onto the scaffold of an indole-based myeloid cell leukaemia 1 (Mcl-1) inhibitor allowed for the formation of an iminoboronate with the non-catalytic Lys234 residue, which synergistically improved the binding by providing an additional point of interaction, while aligning the ligand to the binding pocket.[106] Although most applications of iminoboronates involve the

modification of amino groups present on protein surfaces, there are a plethora of biomolecules sporting primary amines that can be modified with this platform.

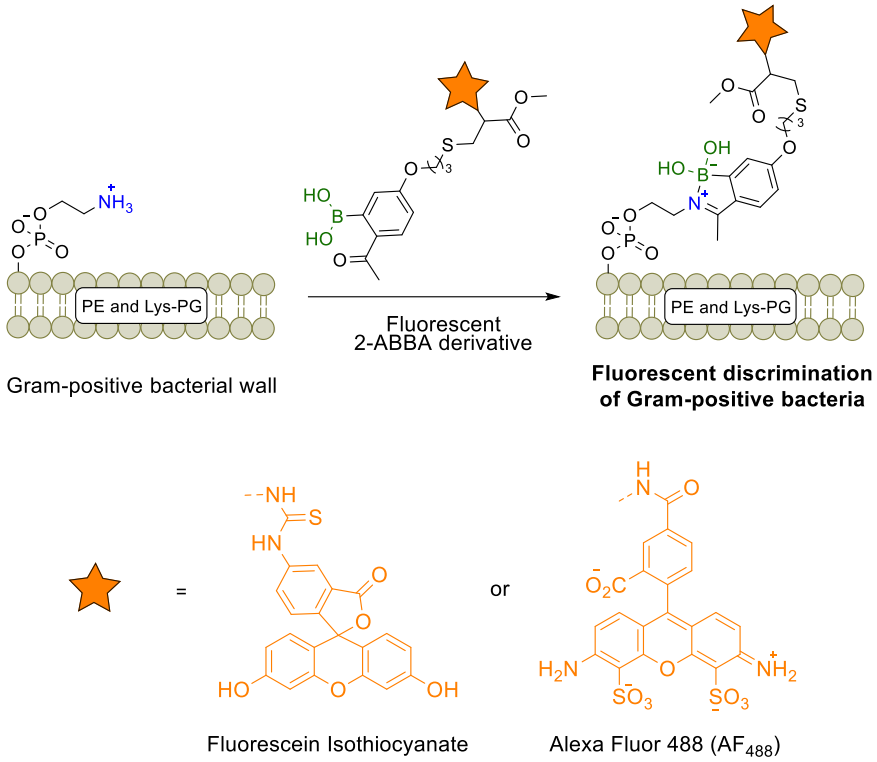


Figure 14 - Fluorescent Gram-positive bacteria discrimination based on a selective lipid-iminoboronates formation.

An exemple of these applications comes from Gao and co-workers, which exploited the presence of amino lipids on Gram-positive bacteria to develop an imaging method to detect specific kinds of bacterial infections. In fact, only the membrane of this class of bacteria contains phosphatidylethanolamine and lysine-modified phosphatidyl glycerol, both of which can be used for iminoboronate formation (**Figure 14**).

Continuing the exploration of the possibilities tied to the insertion of borylated non-natural amino acids inside the structure of a protein, Gao and co-workers reported the use of a 2-ABBA derivative of tyrosine to produce a peptidic scaffold that is able to form iminoboronates. In this context, it was reported that peptides bearing both a Lys and a 2-CBBA unnatural amino acid spaced by 2-5 other amino acids could efficiently yield macrocyclization products. These cyclic peptides retained the reversible nature of iminoboronates as they could be hydrolyzed back to the linear conformation after exposure to acidic or oxidizing environment, as well as in the presence of modulators such as hydrazine (**Figure 15**). The ability to produce these kind of structure is a valuable tool for

the study of biological processes, as conformational equilibria between cyclic and linear peptidic structures play an important role in substrate recognition and protein-protein interactions.

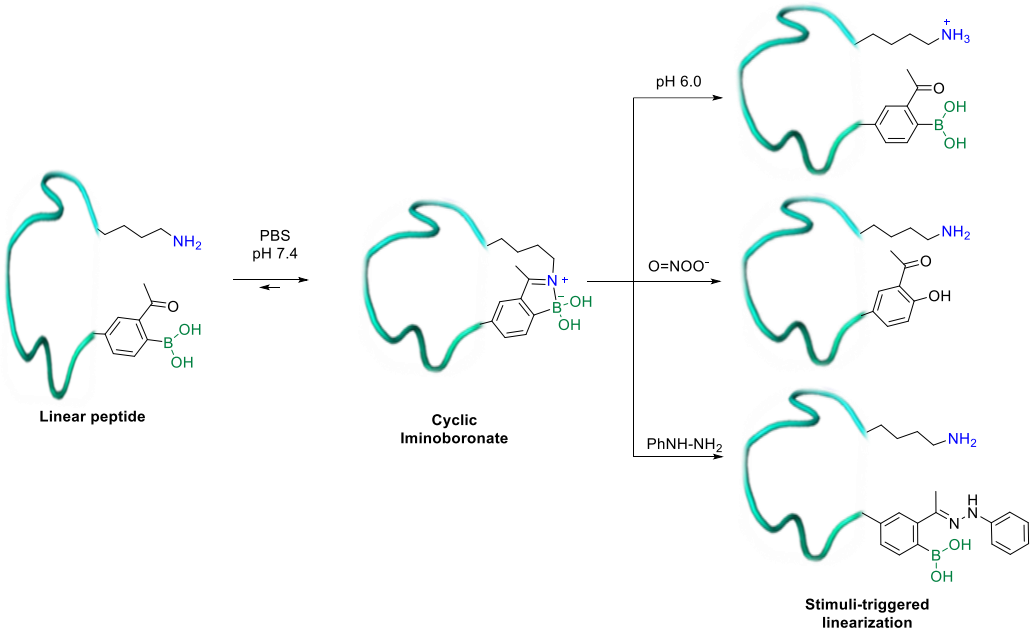


Figure 15 - Iminoboronate formation enables reversible peptide macrocyclization

One of these cases is the receptor-mediated internalization of RGD peptides by the $\alpha_v\beta_3$ integrin surface receptor, which is overexpressed in various cancer cell lines and is a widely targeted receptor for therapeuthic conjugates. As this receptor specifically binds and internalizes the RGD sequence when it is in cyclic conformation, Gao and co-workers prepared a cyclic RGD derivative based on the reported iminoboronate-based cyclization of Lys and 2-CBBA residues and its internalization was tested in ovarian cancer cell lines, using a commercially-available (cyclo-(RGDfC)) peptide as positive control. As expected, while both peptides were internalized under physiological conditions, the iminoboronate-based cyclic peptide failed to internalize at acidic pH, confirming the stimulus-responsive disassembly of the cyclized structure.[107]

Aiming at imporving the hydrolytic stability of imonoboronates, Gois and co-workers reported the synthesis of N,O-iminoboronates, a new class of iminoboronates that showed remarkable stability (<12% hydrolysis in 7 days).[108] The enchanced stability of these constructs, compared to classic iminoboronates, comes from the fact that they are

prepared from the reaction of 2-CBBAs with aminomethyl phenols, whose hydroxyl group offers an additional coordination point to the boron atom, affording a stabilized bicyclic structure. The improved stability profile of N, O iminoboronates makes them compatible with human plasma (<20% degradation in 48h), while still being reversible in acidic conditions, as required for the design of tumour-targeting bioconjugates. In fact, building on this unit as the linker, a folic acid conjugate was prepared with an aminophenol derivative of coumarin, whose fluorescence was shut off by its incorporation in the N, O-iminoboronate core. Incubation of this conjugate in breast cancer cells proved the validity of this approach, as fluorescence was detected only inside the cells, arising from the folic acid-mediated internalization and subsequent pH-triggered hydrolysis of the N,O-iminoboronate which, in turn, released the coumarin restoring its fluorescence (**Figure 16**).

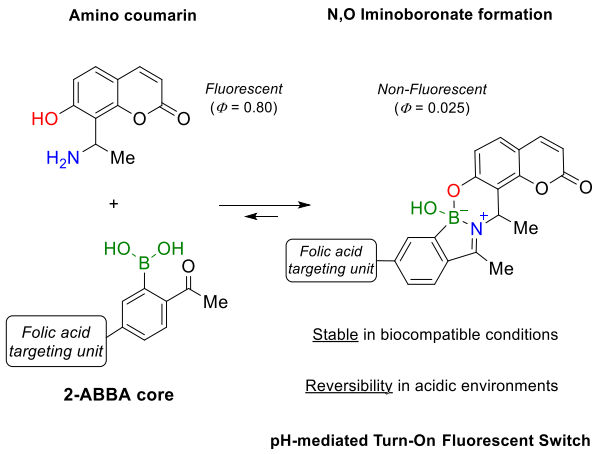
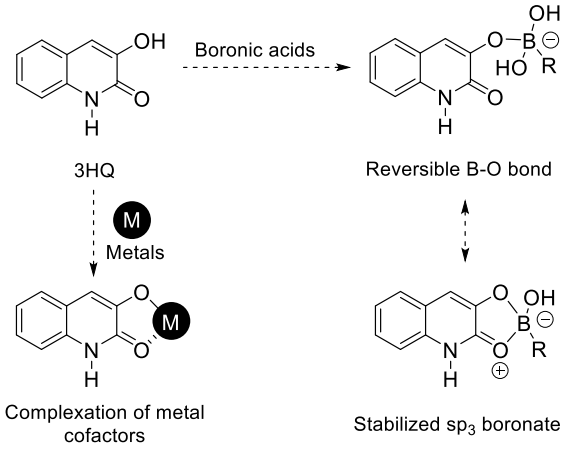


Figure 16 - Aminomethyl phenols and 2-ABBA derivatives react to form *N,O*-Iminoboronates

Despite the remarkable stability improvements that the N, O iminoboronate technology brought in the context of iminoboronate-based platforms, this technology is not useful for the direct modification of proteins. Given this, protein modification with iminoboronates still suffers from two main problems, namely the limited stability of this ligation in complex media and the lack of site-selectivity of the modification.

Chapter II

Modification of the 3-Hydroxy Quinolinone (3HQ) Scaffold for Boronic acid ligation



Abstract

Boronate esters have found widespread application in the landscape of bioconjugation thanks to their tuneable properties and reversibility, features that are directly related to the structure of the BA ligand. In this context, the development of new structures for boronic acid ligation is the key step towards new applications in chemical biology, augmenting the chemical space available to the application of this ligation.

The 3-Hydroxy Quinolinone (3HQ) scaffold is a recurring motif in natural products, found in alkaloid compounds of various plant species. This small molecular core is a bioisoster of the amino acid glycine and in nature was found to be able to bind metal cofactors.

In this chapter, the evolution of the 3HQ scaffold in order to tune its properties towards BA ligation is described, along with the synthetic route used to develop a library of 3HQ derivatives and their evaluation as BA ligands.

A grateful acknowledgement to all collaborators involved in this chapter:

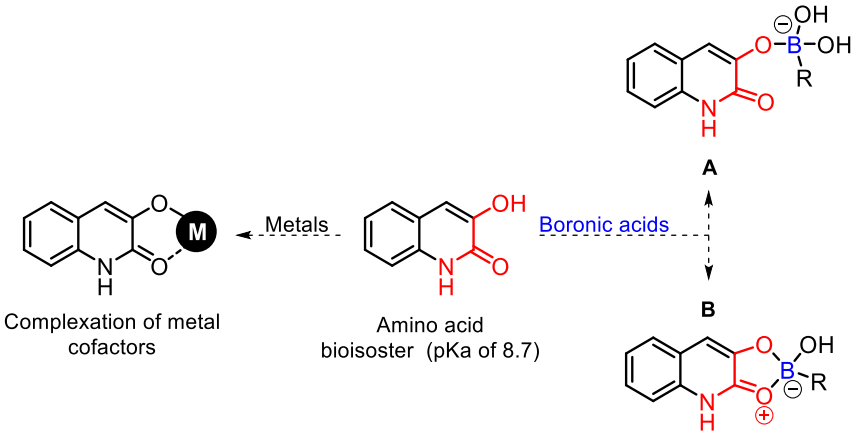
Roberta Paterna (Development of 3HQ synthesis)

Luis F. Veiros (DFT calculations)

II.1 Rationale and Goals

The properties for which BEs are appreciated (reaction kinetics, stability, stimulus-responsiveness) are intimately related to the nature of the ligand that binds the BA. Since BE formation is a dynamic equilibrium, the affinity of a ligand for the BA is a key factor in defining the properties of the system in which it is involved, as it is directly linked to its reversibility. As described in the first chapter of this thesis, BEs can be displaced by the presence of molecules with higher affinity for BAs than the original one, allowing for the design of systems that disassemble in the presence of certain biomolecules, such as sialic acid or fructose. An interesting feature in BE formation is the wide range of affinities covered by the ligands that are known up to date, spacing from low affinity diols such as D-glucose [73] (K_{eq} = 4.6 M^{-1}) to remarkably strong ones, such as nopoldiol [98] (K_{eq} = 3.3 x 10^5 M^{-1}). In this framework, it is important to discover new scaffolds for BA ligation, expanding the chemical space of BE and exploring properties that can be leveraged for novel applications of this functionality.

3-hydroxyquinolin-2(1H)-one (3HQ) is an important motif that appears in the structure of the viridicatin family of natural products.[109–112] This core has been used in the discovery of biological active compounds mainly because it acts as a bioisoster of the amino acid glycine and binds effectively to enzyme metal cofactors.[113–115]



Scheme 1 - 3HQ scaffold and its metal-chelating reactivity, along with expected reactivity with BAs

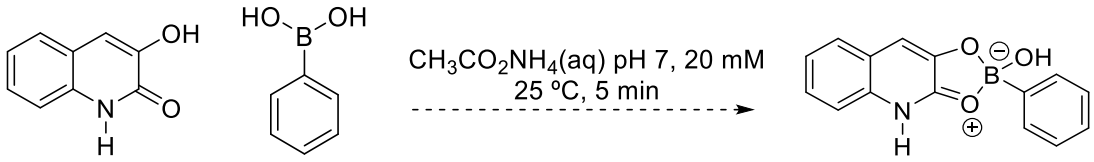
Considering this set of properties, the 3HQ scaffold constitutes a valid candidate for BA ligation, since the acidic hydroxyl function can form a B-O bond leading to the formation of

a sp^3 hybridised boronate species that may be stabilized by an intramolecular coordination with the oxygen of the amide carbonyl group (**Scheme 1**).

II.2 Results and discussion

II.2.1 Preliminary studies on boronic acid ligation with 3HQs

In order to test the applicability of 3HQs as BA ligands, the simple core **1** was synthesized by reaction of Isatin with trimethylsilyldiazomethane, using the protocol reported by Zhang and co-workers.[116] Compound **1** was then tested in a model reaction with 10 equivalents of phenyl boronic acid (PBA) in ammonium acetate pH 7 solution at 37º and 2 mM concentration (**Scheme 2**).



1 **PBA** **2**

Scheme 2 – Model reaction of 3HQ 1 with PBA

The reaction mixture was analysed after 5 minutes by taking up a 10 µL sample of the reaction mixture, diluting it in 190 µL of ammonium acetate pH 7 solution and then injecting it directly in the mass spectrometer. The resulting mass spectrum clearly shows the formation of complexation product **2**, corresponding to the peak with *m/z*= 248.1, while the signal of 3HQ **1**, with *m/z*= 162.0, has low intensity (**Figure 17**).

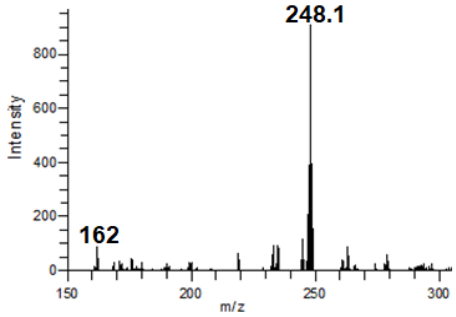
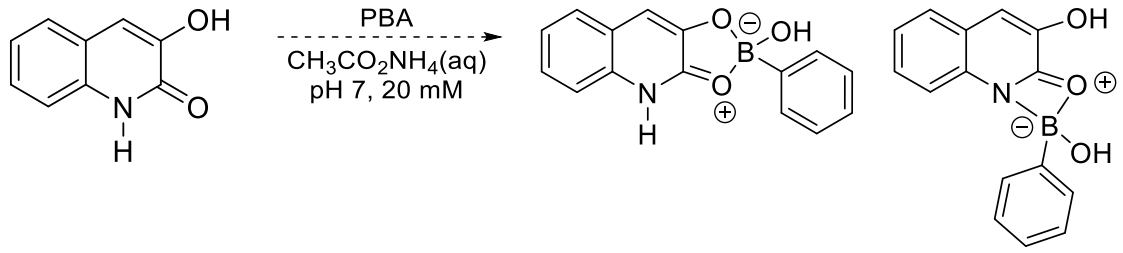


Figure 17 – ESI-MS spectrum of model reaction between 3HQ 1 and PBA

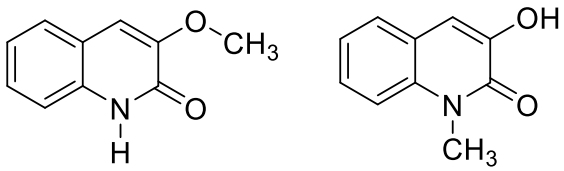
Having established the reactivity of the simple 3HQ core towards the ligation of PBA, it was necessary to elucidate the structure of the complex generated, as two distinct modes of ligation are available (**Scheme 3**) and MS experiments don't allow to gain any structural information on this matter.



1 **2** **3**

Scheme 3 – Possible coordination modes of PBA to 3HQ 1

In order to gain insight into the coordination mode of the complex observed, compounds **4** and **5** were prepared (**Scheme 4**) and tested in the model reaction with PBA. Compound **4** is methylated on the hydroxyl oxygen, hence it is expected to be unable to bind PBA as in structure **2**. On the other hand, compound **5** can only form a product with a structure analogous to **2** since the methylation on the N atom precludes it from binding PBA as shown in structure **3**.



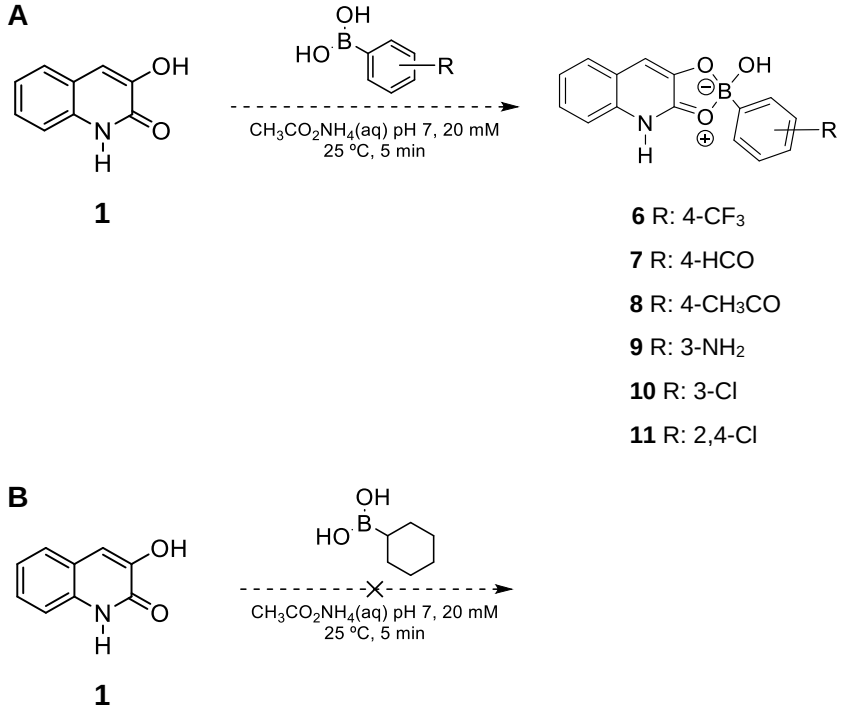
4 **5**

Scheme 4 – Methylated derivatives of the simple 3HQ scaffold

As expected, compound **4** did not react at all, confirming the key role of the hydroxyl group in binding BAs. On the other hand, the ESI-MS spectrum of the reaction of **5** with PBA shows the same pattern observed for the simple 3HQ **1**, suggesting that structure **2** is the likely coordination profile for the 3HQ-PBA complexation.

As this preliminary result confirmed the expected behaviour of 3HQs in binding BAs, a panel of BA derivatives was tested in the same model reaction in order to assess the

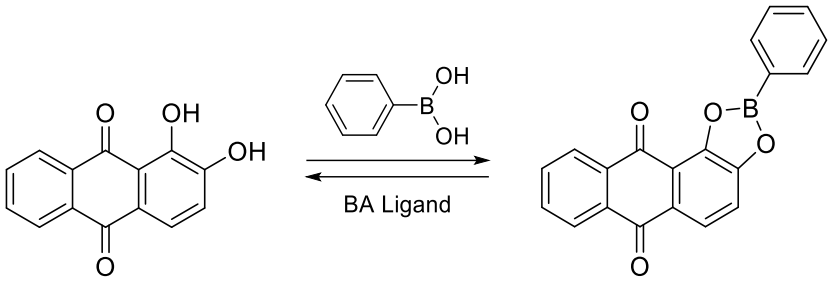
compatibility of 3HQs with BAs showing different electronic characteristics. Among the various BA derivatives, all of the aromatic BAs yielded the expected products (**Scheme 5A**), while the reaction with cyclohexyl boronic acid produced a complex mixture in which the desired product was undetectable (**Scheme 5B**). This behaviour is probably due to the less electron-deficient nature of the boron atom in alkyl BAs, which renders it less prone to this kind of ligation.



Scheme 5 - Complexes generated from the reaction of 1 with a panel of BAs

The following step in characterizing the complexation of BAs and 3HQ was the measurement of their binding affinity, a fundamental property to assess the stability and reversibility of the ligation. The most commonly used method for the determination of binding affinity towards BAs is the Alizarin (AL) competitive titration, which has been widely applied for the determination of the K_a of boronate ester formation with various compounds.[73,117] This methodology exploits the properties of Alizarin red, a polycyclic compound that, thanks to its catechol moiety, is able to form boronate esters. This ligation produces a fluorescent complex between AL and the boronic acid employed, that can be displaced by other ligands. As the presence of another binding partner for the boronic acid can displace the BA-AL complex, the fluorescence intensity decreases in a concentration-dependent fashion. Hence, by first determining the K_a of ALfor the chosen BA by direct

fluorimetric titration, it is then possible to calculate the K_a of the boronic acid ligand of choice by competitive titration of the AL-BA complex (**Scheme 6**) .



Alizarin — Alizarin-PBA complex

Fluorescence off — **Fluorescence on**

Scheme 6 – Fluorescent complex formation between AL and PBA and ligand-triggered complex displacement

Following this methodology, the K_a of complexation between AL and PBA was evaluated in a 1:4 mixture of DMF and ammonium acetate pH 7 solution by adding subsequent aliquots of PBA to a starting 50 μM solution of Alizarin. As the UV-VIS spectrum of the titration shows two isosbestic points, the chosen wavelength for excitation of the system was λ=420 nm (**Figure 18A**). By exciting the system at this wavelength, the fluorescence spectrum of the titration was recorded and the emission values at λ_{exc}=620 nm were used for the calculation of the K_a of AL-PBA complexation in the chosen solvent mixture (**Figure 18B**). The fluorescence intensity data were used to determine the K_a following the Benesi-Hildebrand method, which starts from **Equation 1** that states the relationship between the fluorescence intensity (I_f) and the K_a.

$$\Delta I_f = \frac{(\Delta k p_0 K_a)[L][I_0]}{1 + K_a[L]}$$

Equation 1 - Relationship between fluorescence intensity and association constant

The double reciprocal of **Equation 1**, **Equation 2** is the Benesi-Hildebrand equation, from which the K_a can be extrapolated as the quotient of the intercept and the slope in a plot of 1/[PBA] vs. $1/\Delta I_f$.

$$\frac{1}{\Delta I_f} = (\Delta k p_0 I_0 K_a)^{-1} \frac{1}{[L]} + (\Delta k p_0 I_0)^{-1}$$

Equation 2 – Benesi-Hildebrand equation

Following this protocol, the fluorescence intensity at λ=620 nm was plotted in a Benesi-Hildebrand plot to determine the K_a for Alizarin-PBA complexation in this solvent mixture, indicated as K_{AL}.

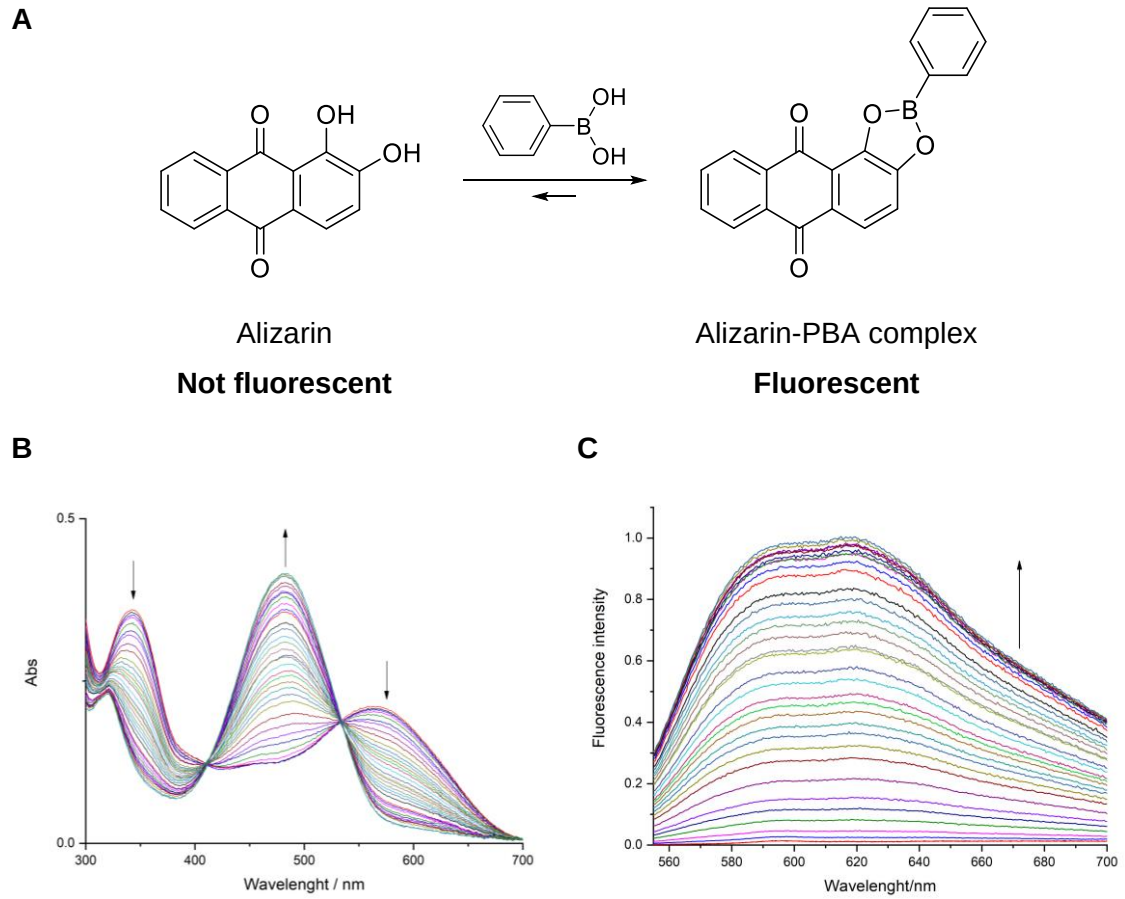
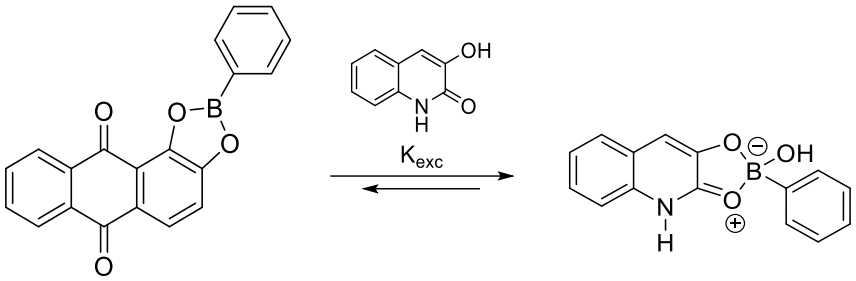


Figure 18 – Complexation equilibrium of Alizarin with PBA (A), UV-VIS spectrum of the titration of AL with PBA (B), Fluorescence emission spectrum of the titration, recorded with λ_{exc}=620 nm (C)

From this titration experiment, the K_{AL} for PBA was calculated to be 1224 ± 31 M^{-1}, which is coherent with what already reported for AL in aqueous conditions.[73] Having calculated the K_{AL} relative to our system, the subsequent step was the evaluation of the K_a of compound **1** with PBA. In order to do so, a 50 μM 1:1 solution of PBA and Alizarin in the same solvent mixture was prepared, to which subsequent aliquots of 3HQ **1** were added,

registering the UV-VIS and fluorescence spectra at each point of the titration. Even though a clear concentration–dependant decrease in fluorescence was observed during the titration (**Figure 19**), when the concentration of compound **1** reached 0.5 mM it started precipitating from the reaction medium, limiting the quantification of the K_a.

A



Alizarin-PBA complex **1** 3HQ-PBA complex **(2)**

Fluorescence on **Fluorescence off**

B

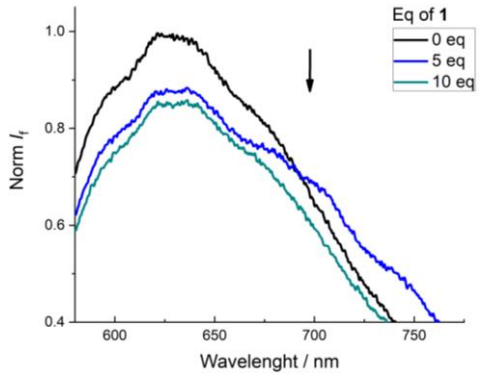
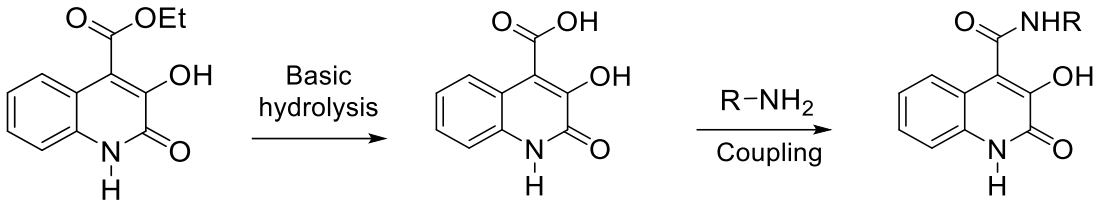


Figure 19 – Competitive equilibrium for the complexation of PBA by Alizarin and 3HQ 1 (A) Fluorescence emission of the titration of AL-PBA complex with 3HQ 1 (B)

II.2.2 Evolution of the 3HQ scaffold

As the final purpose of the 3HQ scaffold is its insertion in a peptidic chain as a bioconjugation handle, it was necessary to evolve this molecular core in order to accommodate an attachment point for protein functionalization. From both a structural and synthetic point of view, the most suitable modification was the insertion of an amide as bioconjugation handle in position 4 of the heterocycle. The synthesis of 4-carboxamide-3HQs is described in the literature via the Eister ring-expansion reaction of isatins with ethyl diazo acetate,[118] followed by the ester hydrolysis under basic conditions, and the

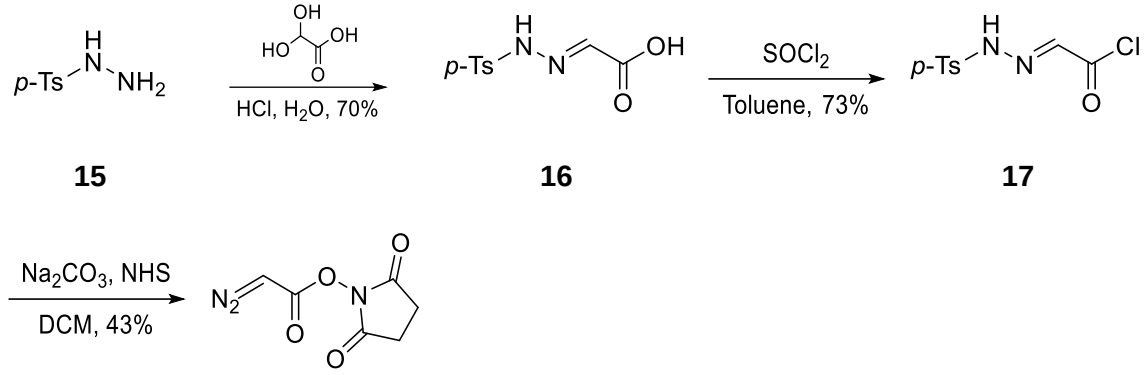
coupling with amines using the Steglich reaction (**Scheme 7**). Unfortunately, this protocol typically affords poor yields of the amide and, depending on the heterocycle substituents, leads to decarboxylation of the 3HQ due to the harsh conditions used during ester hydrolysis.[119] Aware of this limitation, a direct route to prepare 4-carboxamide-3HQs in milder conditions was envisioned, utilyzing a NHS-diazo acetate intermediate (compound **21**) to perform an Eister ring expansion reaction,[120] followed by a simple amidation step.



12 **13** **14**

Scheme 7 – Known procedure towards 4-amide 3HQs

As shown in **Scheme 8**, intermediate **18** was obtained starting from commercially-available glyoxylic acid and tosyl hydrazine (**15**) in three steps with 22% overall yield and without resorting to any chromatographic purification step.



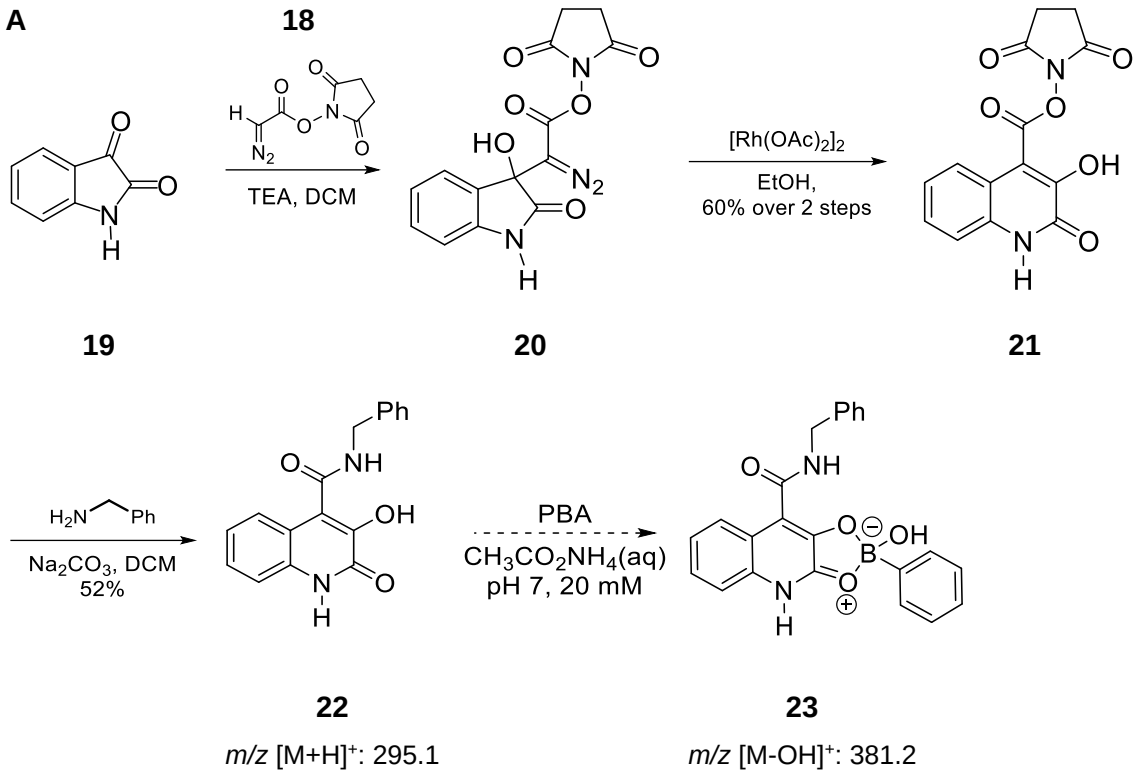
15 **16** **17**

18

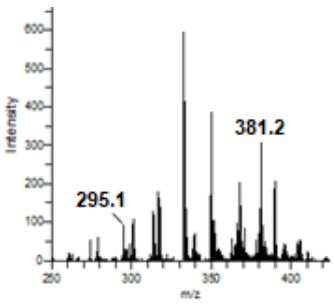
Scheme 8 – Synthetic route towards NHS-diazoacetate 18

Subsequently, compound **18** was used to react with isatin, affording the insertion intermediate **20**, which could readily be converted into 4-NHS-3HQ **21** by rhodium(II)-catalysed ring expansion. This intermediate can easily be converted into different amide derivatives by reaction with primary and secondary amines in the presence of a base (**Scheme 9**). By adopting this new synthetic route, benzyl amide derivative **22** was easily

synthesized and adopted as a model compound to test the reactivity of 4-carboxamide-3HQ scaffolds. As isatin **19** and many of its derivatives are commercially available, this synthetic protocol allows for easy access to various derivatives, with different amides and substituents on the 3HQ core itself.



B

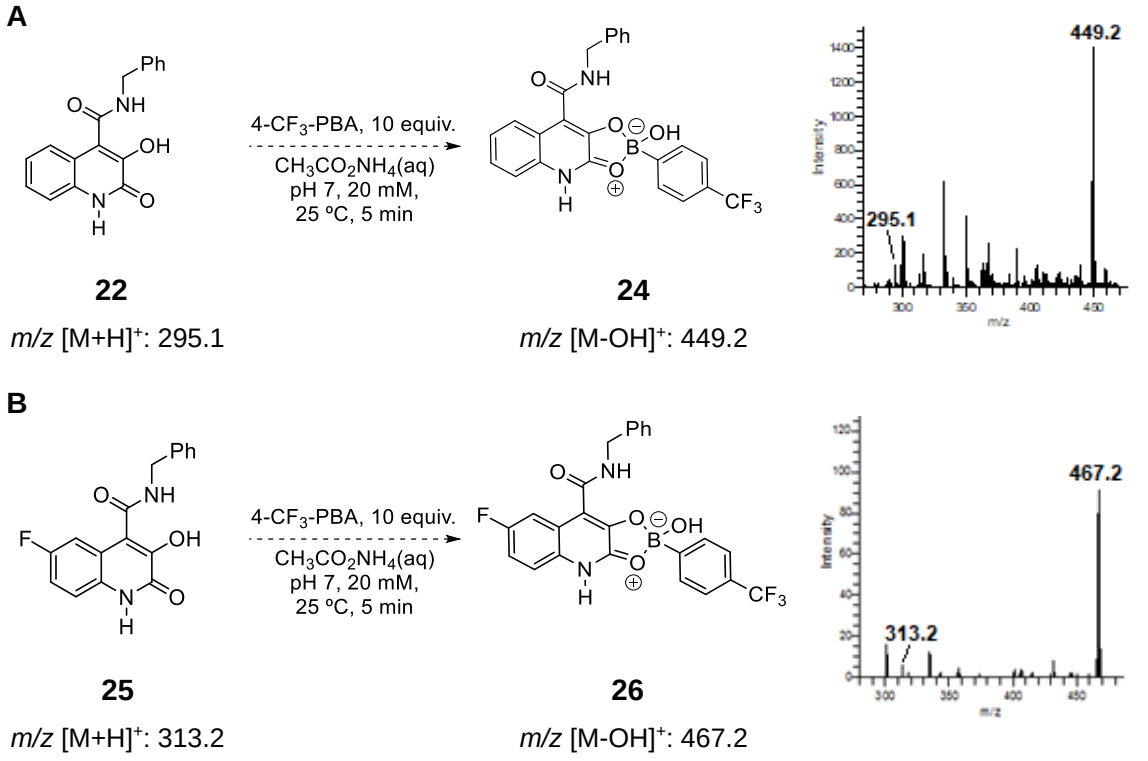


Scheme 9 – Synthesis of 4-carboxamide-3HQ 22 and subsequent reaction with PBA in bioconjugation conditions(A) along with the ESI-MS of the reaction (B)

The insertion of the carboxamide substituent on position 4 of the scaffold was expected to increase the acidity of the OH of this new family of compounds, leading to an improved reactivity profile compared to the simple 3HQ derivatives employed in the previous

paragraph. However, as shown in **Scheme 9B**, the reaction of compound **22** with PBA in bioconjugation conditions produced a complex mixture.

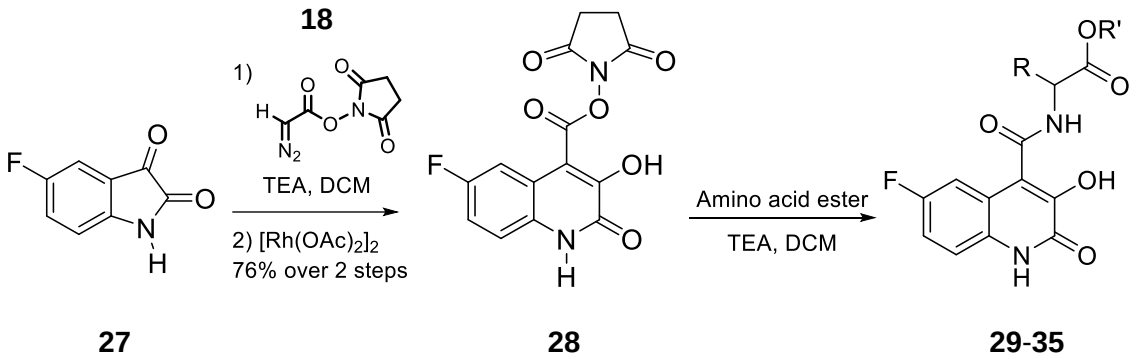
One of the reasons for the observed behaviour could be the sterical hindrance that the benzyl amide moiety brought onto the system, hampering the coordination of PBA. In order to overcome this limitation, the electronic properties of the 3HQ-BA pair were explored, aiming at maximizing the affinity of the binding partners. The first step in this direction was the use of a BA with a more electron deficient boron atom, 4-CF_3 PBA, which indeed showed improved reactivity compared to PBA (**Scheme 10A**). In order to further improve the reaction profile of the 3HQ, a fluorine atom was inserted at position 7 of the scaffold, by adopting 5-fluoro isatin **27** as the starting material in the synthetic route already reported in **Scheme 9**, yielding compound **25**. The increased acidity of the OH of this derivative, along with the adoption of 4-CF_3 PBA as binding partner, led to a good reactivity profile, thus these modifications were incorporated in the design of the 3HQ-BA system going forward (**Scheme 10**).



Scheme 10 – Reaction of 4-CF_3 PBA with 4-carboxamide-3HQ derivatives 22 (A) and 25 (B), along with relative ESI-MS spectra.

The initial studies on the 4-carboxamide-3HQ derivatives **22** and **25** indicated that steric effects of the 3HQ scaffold influence dramatically the coordination of BAs. Therefore, we initiated a study to identify the most suitable unit to insert in position 4 of the heterocycle without negatively impacting the coordination of BAs.

With this objective, a library of derivatives of the 7-fluoro-3HQ scaffold was synthesized, featuring amino acid esters as substituents on the amide position (**Scheme 11**). All of these derivatives were obtained with moderate to good yield after reacting intermediate **28** with different amino acid esters (**Table 1**). In order to identify the most suitable residue for building the spacer unit, the reactivity of these compounds was tested in the coordination of 4-CF_3 PBA in ammonium acetate pH 7 solution. After 5 minutes of stirring at 25 ºC, each reaction was analysed by MS spectrometry, in accordance with the protocol adopted in the initial model reaction.



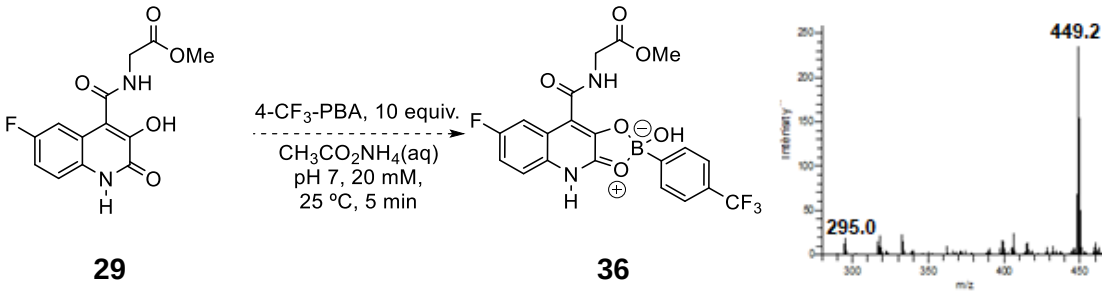
Scheme 11 – Synthesis of fluorinated NHS 3HQ derivative 28 and its use in the synthesis of a library of 3HQ amino acid derivatives

Compound	Amino acid	Structure	Yield
29	Glycine methyl ester	**R**=H **R'**=CH_3	72%
30	Alanine methyl ester	**R**=CH_3 **R'**=CH_3	71%
31	Leucine methyl ester	**R**=$CH_2CH(CH_3)_2$ **R'**=CH_3	71%
32	Phenylglycine methyl ester	**R**=Ph **R'**=CH_3	76%
33	Phenylalanine ethyl ester	**R**= CH_2Ph **R'**= CH_2CH_3	84%
34	Serine methyl ester	**R**=CH_2OH **R'**=CH_3	32%
35	Cysteine methyl ester	**R**=CH_2SH **R'**=CH_3	26%

Table 1 – Structure and yields relative to 3HQ amino acid derivatives 29-35

Similarly to what was previously observed, the ESI-MS spectra of the reactions showed that the reactivity profile of each amino acid derivative towards BA complexation was influenced by their sterical hindrance. In fact, while glycine derivative **29** showed excellent reactivity, the complexation with more hindered amino acids was clearly more difficult.

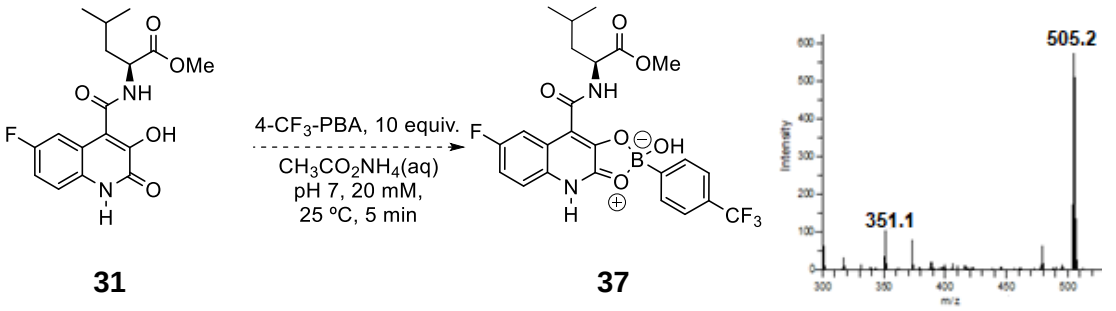
A



29 **36**

m/z [M+H]$^+$: 295.0 *m/z* [M-OH]$^+$: 449.2

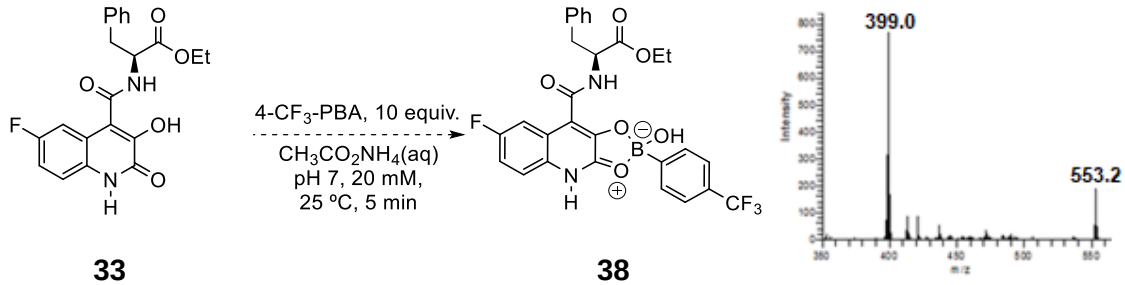
B



31 **37**

m/z [M+H]$^+$: 351.1 *m/z* [M-OH]$^+$: 505.2

C



33 **38**

m/z [M+H]$^+$: 399.0 *m/z* [M-OH]$^+$: 553.2

Scheme 12 – Reaction of glycine derivative 29 (A), leucine derivative 31 (B) and phenylalanine derivative 33 (C) with 4-CF3 PBA and relative ESI-MS spectra

As shown in **Scheme 12**, the intensity of the peak of the 3HQ-BA adduct decreases, relatively to the one of the starting 3HQ derivative, as the bulkiness of the amino acid substituent increases. Interestingly, even though they showed good reactivity towards BAs, both the serine and cysteine derivatives **34** and **35** shown to be prone to degradation, probably due to an intramolecular nucleophilic attack that causes the hydrolysis of the amide bond. After testing the reactivity of all of these derivatives *via* ESI-MS, compounds **29** and **31** were also submitted to the already mentioned competitive titration with the PBA-AL system, with the objective of measuring the affinity constants of these compounds. As shown in the previous paragraph, the determination of the K_a for PBA of compound **1** was impossible with this method, due to its aggregation and consequent precipitation during the titration (**Figure 19**). In this case, the two selected amino acid derivatives proved to be less prone to aggregation and enabled a complete titration, thus obtaining their respective association constants towards PBA. Following the same protocol described for compound **1**, a 50 μM 1:1 mixture of AL and PBA in a solvent mixture of DMF and ammonium acetate pH 7 solution (1:4) was titrated by subsequent addition of aliquots of either **29** or **31**. As expected, the subsequent additions of 3HQ derivatives caused a concentration-dependent decrease in fluorescence, corresponding to the displacement of the Alizarin-PBA complex (**Figure 20**). The values of fluorescence intensity at λ=626 nm arising from the titration performed with glycine derivative **29**, along with the previously calculated K_{AL} value, were used to calculate the K_a of this compound towards PBA, using the methodology reported by Tirelli and co-workers.[121] After the mathematical treatment of the data collected, the affinity constant of **29** towards PBA ligation resulted to be equal to 699 ± 17 M^{-1}. Interestingly, the same technique applied to leucine derivative **31** yielded a K_a of 117.7 ±3 M^{-1}, in line with the observation that more sterically hindered amide derivatives have lower affinity towards BAs. Referencing the work of Wang and co-workers on the affinity of natural molecules towards boronic acids,[73] the K_a of glycine derivative **29** resulted to be comparable to the one of dopamine, while being greater than the one of common carbohydrates. Because of this feature, the use of glycine as a spacer in the 3HQ platform for bioconjugation is expected to yield a stable boronate, that can still be displaced by high concentration of stronger BA ligands such as sialic acid and dopamine.

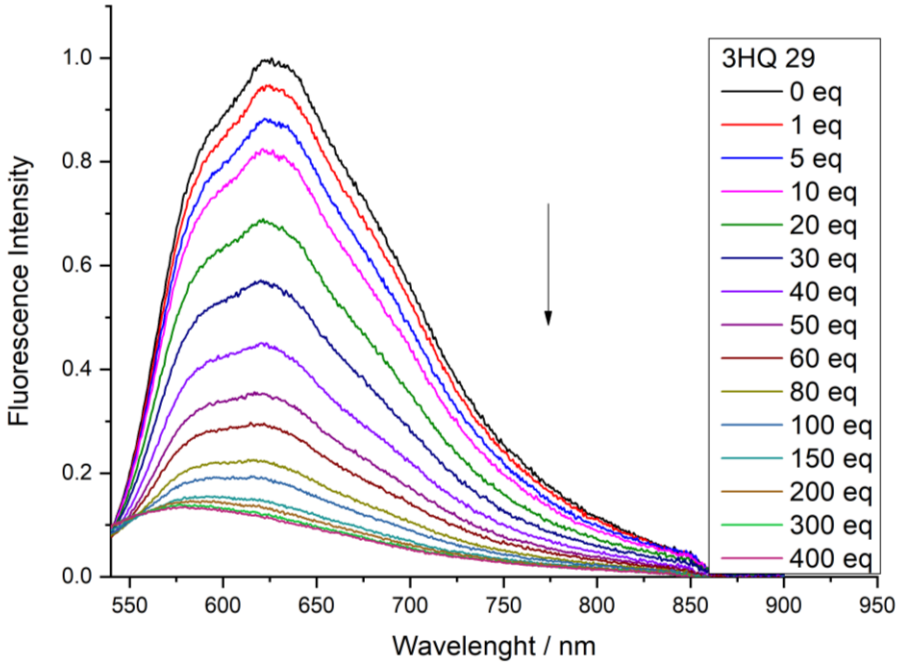
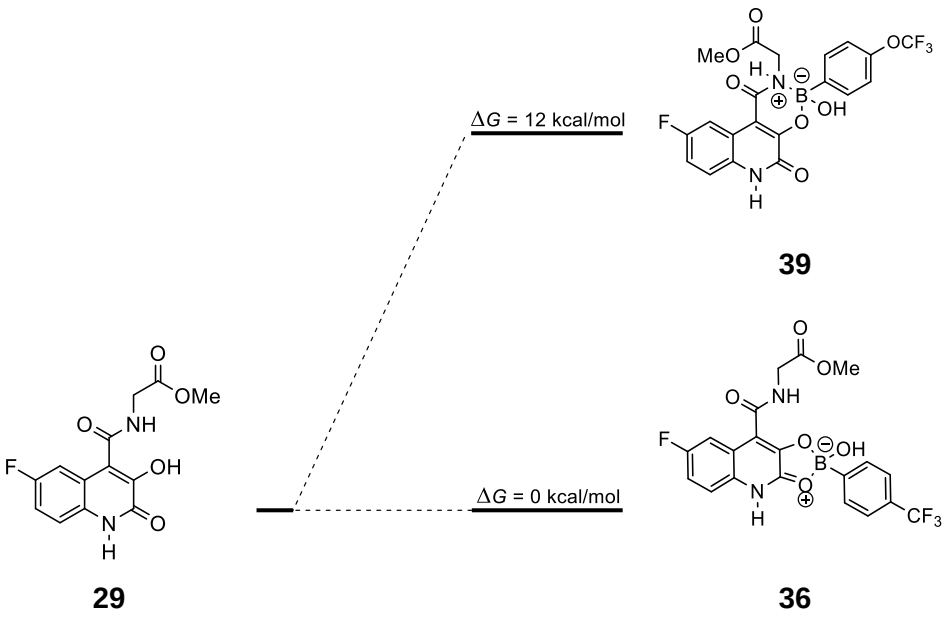


Figure 20 – Fluorescence spectrum of the titration of the AL-PBA complex with 3HQ glycine derivative 29

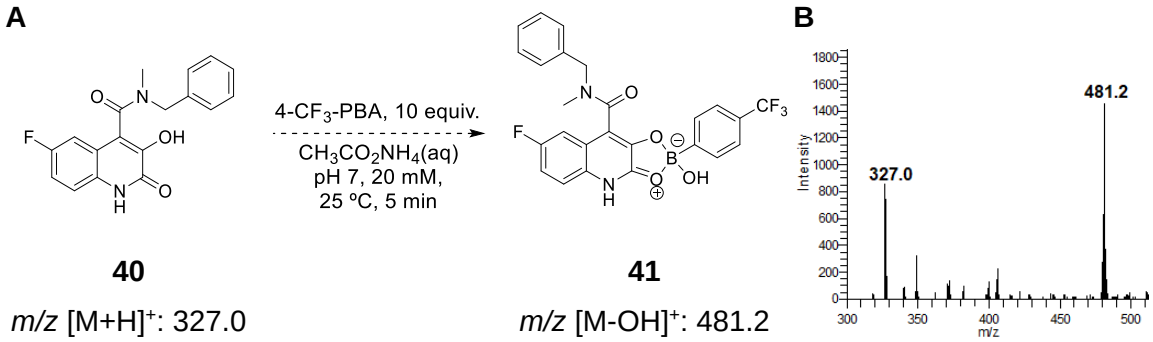
The results of compounds **29**-**35** in the model reaction with 4-CF_3-PBA in bioconjugation conditions confirmed the negative influence that sterical hindrance has on the ligation. In fact, the best results were obtained with glycine derivative **29** while, as shown in **Scheme 12**, 3HQ-carboxamide derivatives with larger side chains showed poor reactivity. Measurement of the affinity constants with the AL assay confirmed this behaviour, as the K_a of compound **29** resulted to be significantly higher than the one of the more branched leucine derivative **31**. Taking into account the results obtained in this paragraph, glycine derivative **29** was selected as starting structure for the construction of the 3HQ-based platform for bioconjugation.

However, the addition of the amide functionality to the original 3HQ scaffold opened up the possibility of having another mode of ligation for boronic acids. In fact, in addition to the 3HQ boronate ester formation observed in the simple core, the nitrogen atom of the amide could potentially partake in binding the boron centre. Using glycine derivative **29** as a model compound, DFT calculations were performed, aiming at evaluating the energy associated with the two possible ligation modes.



Scheme 13 – DFT calculations showing the energies related to the two possible coordination modes of 4-amide 3HQ derivatives

As shown in **Scheme 13**, the salicyl hydroxamate-like ligation product **39** is less energetically favoured compared to the one formed by 3HQ ligation, suggesting that this is the prevalent species in solution. Notably, these calculations showed that the formation of adduct **36** is a thermoneutral process, indicating that an excess of boronic acid is necessary to push the system towards complete complexation. Additionally, in order to confirm that the amide handle in position 4 does not directly coordinate the boron centre, a methylated analogue of compound **25**, 3HQ derivative **40** was synthesized.

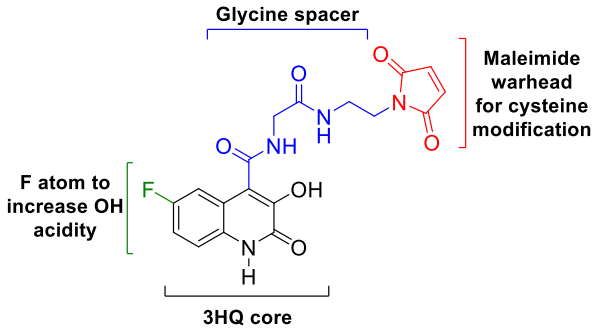


Scheme 14 – Reaction of N-methylated 3HQ 40 with 4-CF_3-PBA (A) and relative ESI-MS spectrum (B).

This compound was obtained by reacting N-methyl benzyl amine with intermediate **28** and was then employed in the model reaction with 4-CF_3-PBA (**Scheme 14**). As expected, complexation occurred in 5 minutes, albeit with lower efficiency probably due to the sterical hindrance that the methyl substituent brought to the structure.

II.2.3 Peptide modification with 3HQs

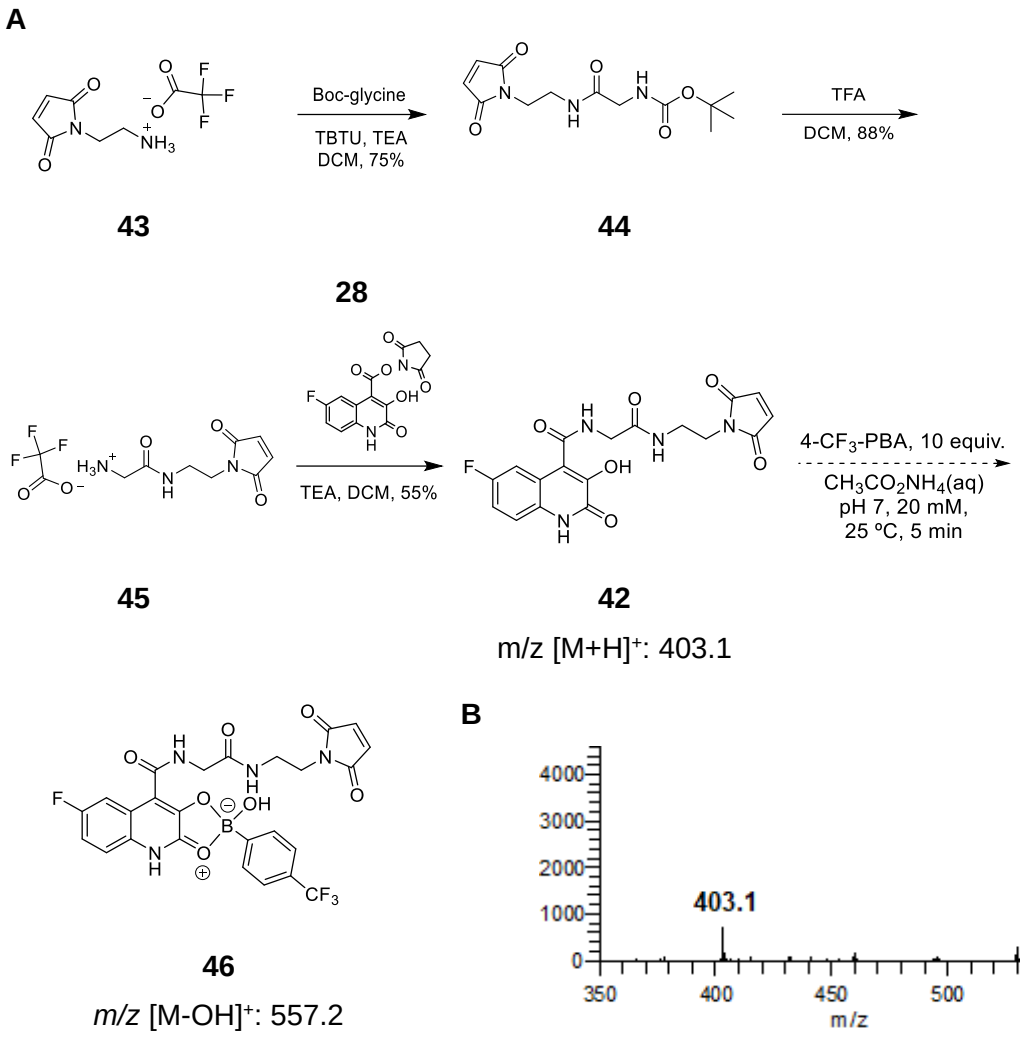
Finally, the last step towards obtaining a platform for BA bioconjugation based on the 3HQ scaffold was the insertion of a handle to install this structure onto peptides. Having this objective, we envisioned the modification of the structure of compound **29** to accommodate a maleimide unit as bioconjugation warhead. Maleimides have been widely used for peptide modification, as they act as efficient Michael acceptors for the functionalization of cysteine side chains in aqueous conditions.[25] Starting from the structure of 3HQ-glycine derivative **29**, we designed compound **42** in which the glycine chain is elongated with an ethylenodiamine maleimide unit.



42

Scheme 15 – Functional 3HQ unit for the construction of bioconjugates

The synthesis of compound **42**, as shown in **Scheme 16A**, was achieved through the reaction of NHS-3HQ intermediate **28** with a glycine-maleimide spacer that was obtained by coupling of glycine with maleimide **43**. This compound was first tested for its reactivity in binding 4-CF_3 PBA in bioconjugation conditions, as shown in **Scheme 16B**. Gratifyinlgy, despite increasing the size of the amide linker in position 4 of the 3HQ scaffold, this compound maintained the favourable reactivity profile of simpler glycine derivative **29**.



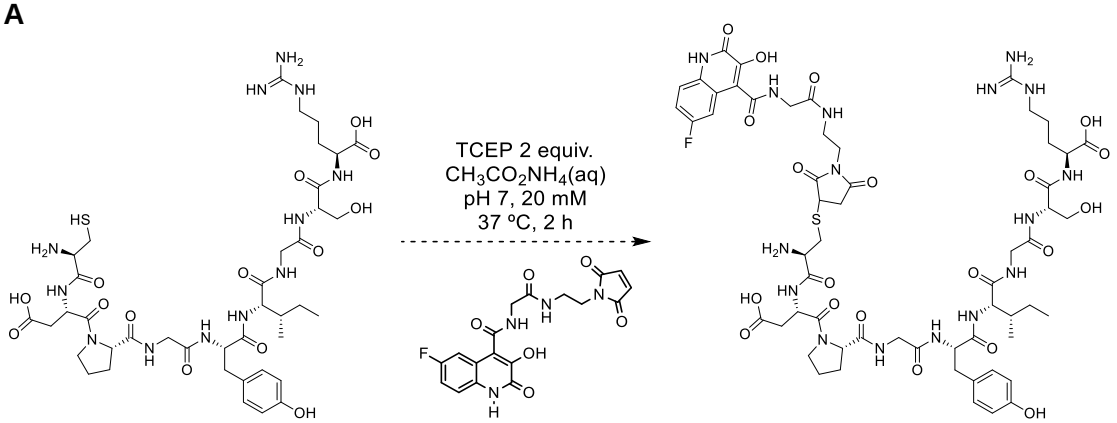
Scheme 16 – Synthesis of 3HQ-maleimide 46 (A) and ESI-MS spectrum of its reaction with 4-CF_3 PBA in bioconjugation conditions (B)

After confirming its reactivity towards the complexation of boronic acids, compound **42** was employed for the modification of laminin CDPGYIGSR fragment, a short oligopeptide belonging to the laminin protein.[122,123] The YIGSR fragment is the shortest sequence recognized by the 67LR, a cell surface receptor that is overexpressed in many cancer cell lines, constituting a potential targeting unit for the construction of a therapeutic bioconjugate.[124–126] Moreover, this peptide contains an *N*-terminal cysteine residue, which can be modified by the maleimide warhead of **42** without interfering with the recognition sequence, which is located on the *C*-terminus.

In order to attach 3HQ-maleimide derivative **42** to the desired peptide, a 0.1 mM solution of laminin fragment in ammonium acetate pH 7 20 mM solution was first prepared and

stirred for 30 minutes at 37 ºC with tris(2-carboxyethyl)phosphine (TCEP) in order to avoid the presence of disulphide-bridged dimers.

A



laminin fragment

m/z [M+H]$^+$: 967.2

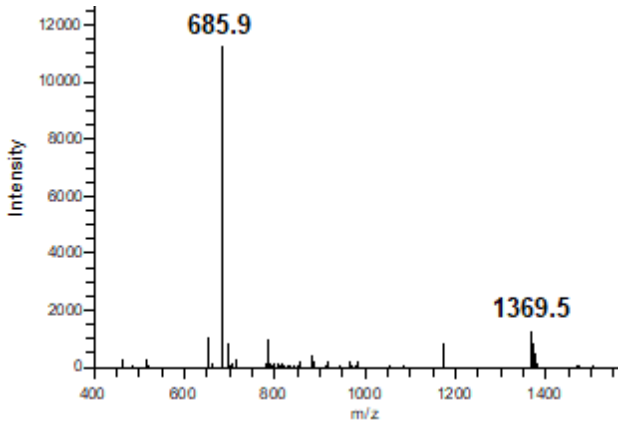
42

47

m/z [M+H]$^+$: 1369.5

m/z [M+2H]$^{2+}$: 685.9

B



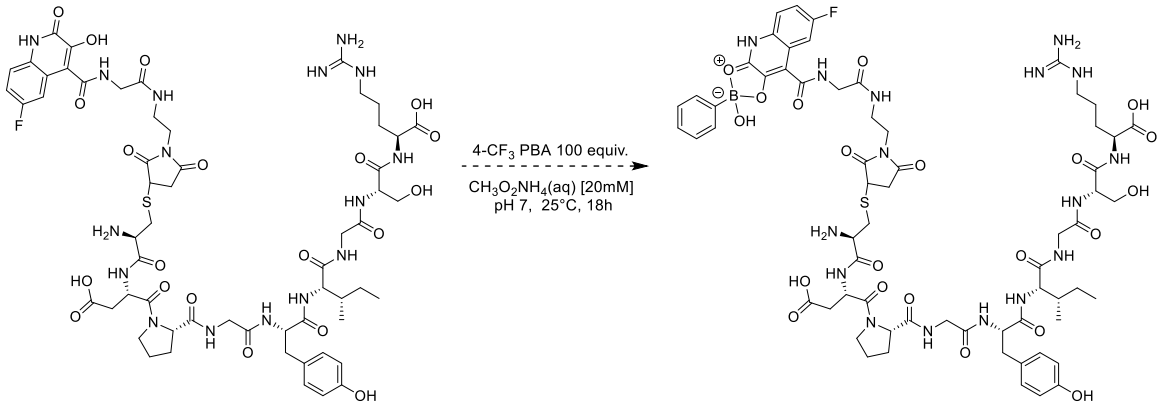
Scheme 17 – Functionalization of the laminin fragment with 42 (A) and relative ESI-MS spectrum (B)

Subsequently, 5 equivalents of compound **42** were added to the reaction, which was stirred for 2 hours, leading to complete functionalization of the peptide, as confirmed by ESI-MS (**Scheme 17**). The obtained solution of **47** was dialyzed for 12 h using a Pur-A-Lyzer™ Midi Dialysis Kit with a 1 KDa cut-off, in order to remove the remaining TCEP and the excess of compound **42** left in the reaction mixture.

The resulting 3HQ-modified peptide **47** was subsequently reacted with 4-CF_3 PBA in bioconjugation conditions. initially, 10 equivalents of BA were employed, according to the reaction conditions optimized up to this point. Regrettably, this amount of BA was not

sufficient to promote the formation of the desired product, probably due to the higher dilution in which this reaction was performed compared to the previous model reactions.

A



47

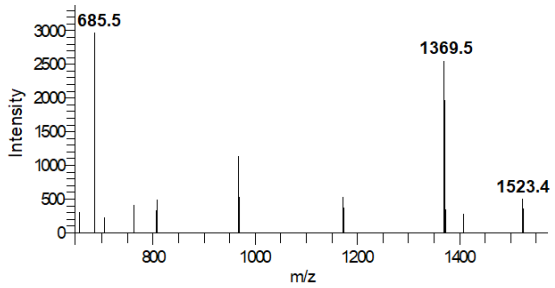
m/z $[M+H]^+$: 1369.5

m/z $[M+2H]^{2+}$: 685.9

48

m/z $[M\text{-}OH]^+$: 1523.4

B

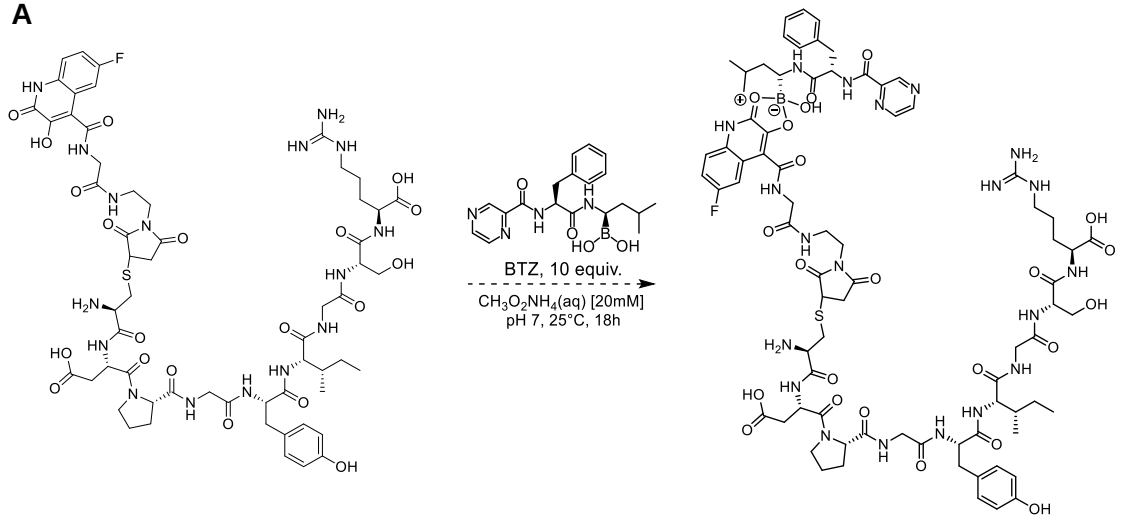


Scheme 18 – Reaction of 3HQ-laminin derivative 47 with 4-CF3 PBA (A) and relative ESI-MS spectrum (B)

With the objective of promoting the desired ligation, the reaction was repeated with 100 equivalents of 4-CF_3 PBA and with increased reaction times, up to 18 hours (**Scheme 18A**). This last protocol allowed the formation of the desired complex between the modified peptide, albeit without reaching complete conversion, as shown in **Scheme 18B**.

The decrease in reactivity observed between the model reactions and the peptide modification reaction is probably linked to the dependence of 3HQ-BA complexation on the steric constraints imposed by the amide ligand in position 4. In fact, the bulky structure of the peptidic unit is likely to impose an excessive sterical hindrance on the 3HQ system, hampering its reactivity.

A



47

m/z $[M+H]^+$: 1369.5

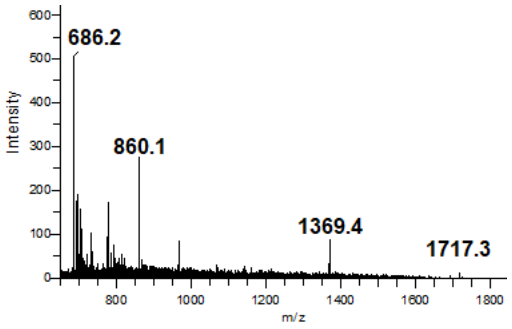
m/z $[M+2H]^{2+}$: 685.9

49

m/z $[M\text{-}OH]^+$: 1717.3

m/z $[M+H\text{-}OH]^{2+}$: 860.1

B



Scheme 19 – Reaction of 3HQ-laminin derivative 47 with Bortezomib (A) and relative ESI-MS spectrum (B)

As the 3HQ-BA interaction alone proved not to be strong enough for the construction of bioconjugates, it was envisioned that the presence of additional favourable interactions between the 3HQ-functionalized peptide and the payload could provide the additional interactions necessary to overcome the aforementioned sterical effects. Following on this concept, Bortezomib (BTZ), an FDA-approved drug for the treatment of multiple myeloma, was chosen as the payload of choice. Besides being a therapeutically useful compound, the relevance of BTZ for such application lies in its structural features. This molecule, in fact, not only presents a boronic acid moiety that can react with the 3HQ core, but also

has a pseudopeptidic structure that could promote additional interactions with the protein backbone *via* hydrogen bonding.

Therefore, 3HQ-laminin derivative **47** was reacted with 10 equivalents of BTZ in the same conditions reported previously for the reaction with 4-CF_3 PBA and stirred overnight. As shown in **Scheme 19B**, the reactivity profile of this system was similar to the one observed for the reaction with 100 equivalents of 4-CF_3 PBA, indicating that the contribution of the pseudopeptidic structure towards the formation of the product is relevant. In addition, since BTZ is an alkylic boronic acid, it has a less electron-deficient boron centre and was expected to have poorer affinity towards 3HQ complexation.

II.3 Conclusions

In this chapter, the study of the 3-hydroxy quinolinone scaffold as a potential ligand for boronic acids in bioconjugation conditions was presented. Starting from the simple core **1**, ESI-MS experiments proved that it can quickly and efficiently react with different BAs in bioconjugation conditions. Subsequently, this structure was optimized to accommodate a carboxamide substituent on position 4 of the heterocycle. Along with the evolution of the 3HQ structure, a convenient and efficient synthetic route was developed, making the construction of 4-carboxamide-3HQ libraries easily accessible (**Scheme 9**).

Taking advantage of this methodology, two series of 3HQ derivatives were synthesized, starting with benzylamide-substituted derivatives **22** and **25**. The study of the properties of these two compounds indicated that the insertion of a fluorine atom on position 7 of the 3HQ scaffold has a positive effect on BA ligation, as compound **25** showed a better reactivity profile than **22**. Surprisingly, the reactivity of the two 4-bezylamide-3HQ derivatives towards PBA proved to be worse than what observed for the starting core (compound **1**). In order to overcome this limitation, 4-CF_3-PBA was employed in the reaction, leading to a significant improvement of the system thanks to its more pronounced electron-deficient character. The second library of 3HQ derivatives was synthesized by reacting fluorinated intermediate **28** with several amino acid esters, yielding 3HQ amino acid derivatives **29**-**35**. Reaction of this second library of compounds with 4-CF_3-PBA in bioconjugation conditions indicated the importance that the sterical hindrance of the 4-carboxamide susbstituent has on the ligation. In fact, among this second series of compounds, glycine derivative **29** produced the most efficient BA ligation. On the other hand, 3HQs with more bulky substituents, such as leucine derivative **31** and phenylalanine

derivative **33** showed worse reactivity. This behaviour was confirmed by measuring the K_a of **29** and **31** towards PBA complexation by using the competitive AL titration assay. As expected, glycine derivative **29** showed a K_a almost 7-fold higher than what measured for leucine derivative **31** (**Figure 20**), confirming the importance of sterical factors in the binding of BAs.

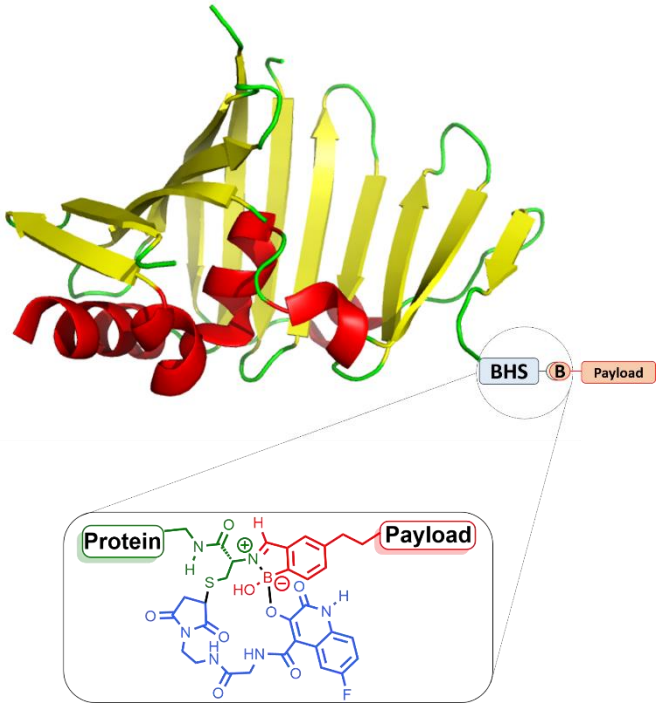
The last step of this study was the development of a 3HQ platform that could be attached to proteins, with the objective of applying the 3HQ-BA complexation towards the preparation of bioconjugates. To this end, the structure of **29** was modified, leading to the synthesis of compound **42**, which presents a glycine spacer and a maleimide warhead to react with cysteine residues.

This platform was applied to the modification of a model peptidic scaffold, the laminin fragment. Despite the successful installation of the 3HQ core on the peptide, its reactivity resulted not to be as good as observed during the development of the technology. The most probable explication for this behaviour lies on the higher sterical hindrance that a peptidic structure imposes on the 3HQ core, interfering with its reactivity.

However, as the experiments with Bortezomib show, combining the 3HQ-BA ligation with other interactions leads to improvements in the coordination profile, thus indicating a possible avenue to improve this promising bioconjugation platform.

Chapter III

3HQs as Boronic Acid Hotspots for site-selective iminoboronate formation



Abstract

Despite exhibiting inefficient boronate ester formation in the context of peptide modification, 3HQs have shown promising properties as ligands for boronic acids.

In this chapter, this scaffold was applied in the framework of iminoboronate formation, by installation of the 3HQ platform onto *N*-terminal cysteine residues.

The peptide functionalization with the 3HQ moiety enabled the site-selective formation of iminoboronates with the *N*-terminus amino group of various peptides, yielding stabilized constructs that were applied to construct fluorescent non-internalyzing bioconjugates.

A grateful acknowledgement to all collaborators involved in this chapter:

Rita Padanha (ESI-MS experiments)

Luis F. Veiros (DFT calculations)

Fábio Fernandes (*in vitro* internalization studies)

Eurico Cabrita (^{11}B-NMR studies)

Francisco Corzana Lopez (^{15}N-NMR studies)

Héctor Busto Sancirián (^{15}N-NMR studies)

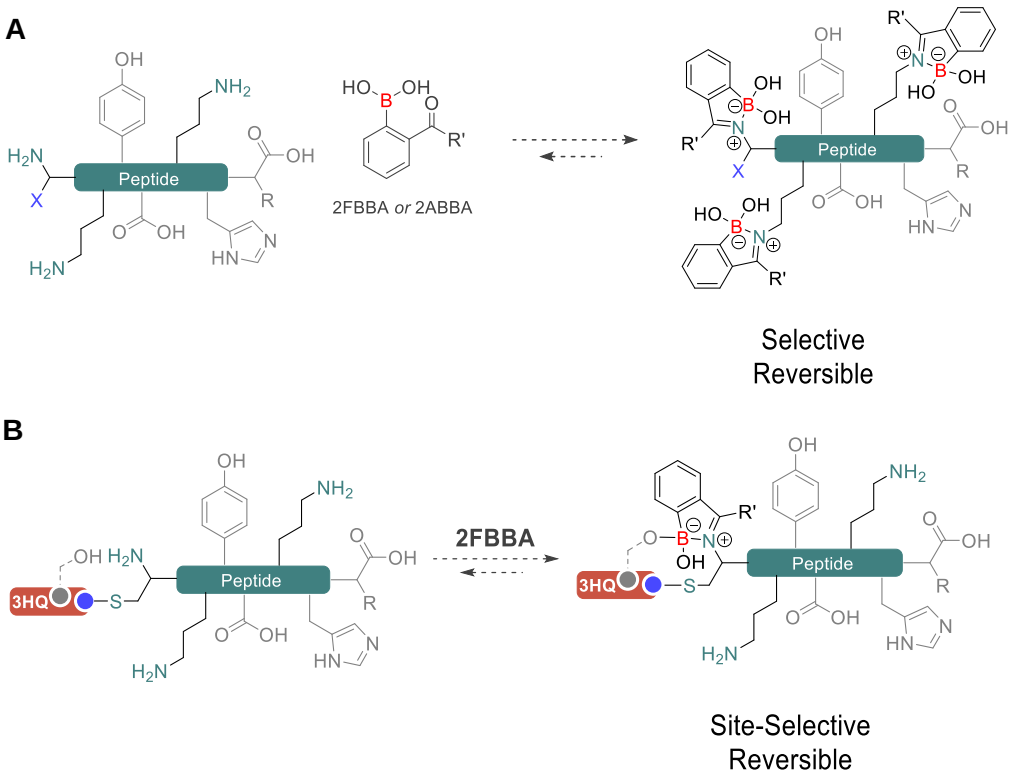
III.1 Rationale and Goals

As reported in the last chapter, the 3HQ core effectively coordinated with BAs, establishing a reversible covalent B-O bond and exhibiting K_a values higher than most carbohydrates regarding boronate ester formation. This interesting reactivity, however, was found to be strongly dependant on the sterical hindrance imposed on the system by the carboxamide substituent, leading to a loss of reactivity in the context of protein modification. In fact, when the 3HQ core was attached to a peptidic scaffold through a maleimide warhead, the resulting construct did not bind boronic acids with the same efficiency shown in model reactions. Interestingly, when employing Bortezomib as boronic acid, the reactivity profile improved compared to what observed with 4-CF_3-PBA. This result was surprising, considering that the boron atom in 4-CF_3-PBA is significantly more electron deficient than the alkylic boronic acid present in the structure of Bortezomib. A possible explanation for this improved reactivity might be that, since Bortezomib has a pseudopeptidic structure, it was able to form hydrogen bonds with the peptidic backbone of the protein fragment used, thus providing additional stabilizing interactions.

Based on these observations, we envisioned that employing the 3HQ-BA complexation in tandem with other interactions would lead to efficient and stable BA coordination in bioconjugation conditions. In order to develop this concept, the reactivity of the 3HQ core with BAs was explored to direct the formation of iminoboronates towards the *N*-terminal amine group in peptides, resulting in a synergistic functionslization based on 3HQ/iminoboronate technology.

Despite the success of the iminoboronate technology, demonstrated in many different applications,[127] this technology still faces some drawbacks, namely the lack of site-selectivity and the low stability on the construct under physiological conditions. In fact, as described in the introduction, 2-CBBAs are prone to react with exposed amino group on the surface of a protein, leading to unselective and heterogeneous modification. Also, while stimuli-responsiveness is one of the most appreciated properties of iminoboronates, their rapid hydrolysis in the presence of several natural molecules limits their application in complex biological contexts.

In this chapter we studied the possibility to use the 3HQ scaffold to improve the stability and selectivity of iminoboronate formation (**Scheme 20**).



Scheme 20 – Traditional iminoboronate formation with 2-CBBAs (A) and site-selective iminoboronate formation based on the 3HQ platform (B)

III.2 Results and discussion

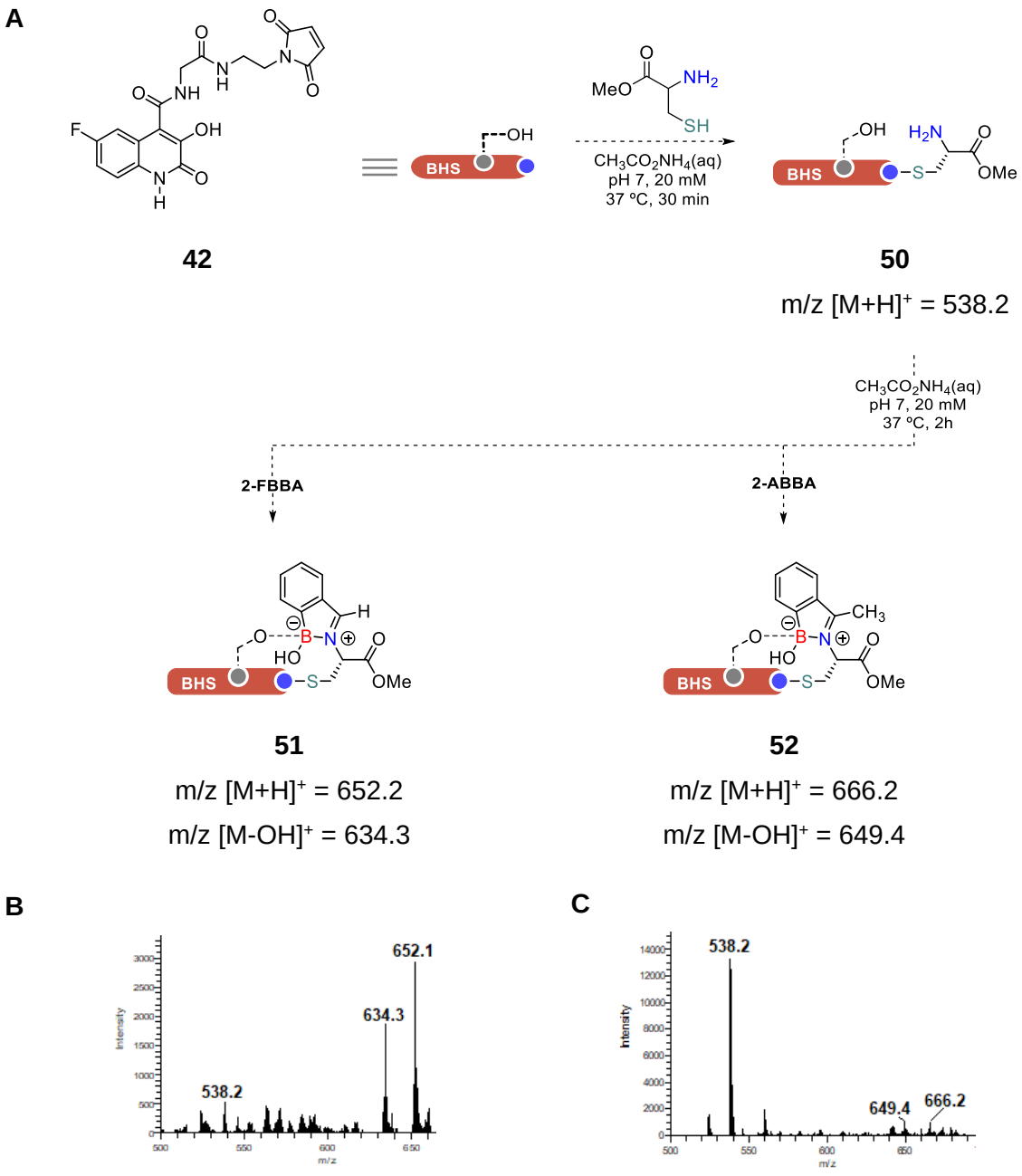
With the objective of performing the site-selective installation of iminoboronates exploring the 3HQ core as a boronic acid hot-spot (BHS), we envisioned the use of compound **42**, which has a maleimide warhead for cysteine modification, for the functionalization of *N*-terminal cysteine residues. As shown in **Scheme 20B**, this functionalization would lead to peptidic structures in which iminoboronate formation is directed onto the *N*-terminus amine residue due to the presence of the 3HQ moiety. Additionally, the 3HQ-iminoboronate adduct described in **Scheme 20B**, is expected to be more stable than classic iminoboronates due to the formation of the B-O bond between the 3HQ and the BA moiety.

III.2.1 Model reactions

In order to test this strategy, compound **42** was reacted with cysteine methyl ester in ammonium acetate pH 7 20 mM solution at 37 ºC, obtaining compound **50**, which was adopted as model for the modification of *N*-terminal cysteine residues with the 3HQ platform (**Scheme 21A**). Once complete formation of compound **50** was observed by ESI-

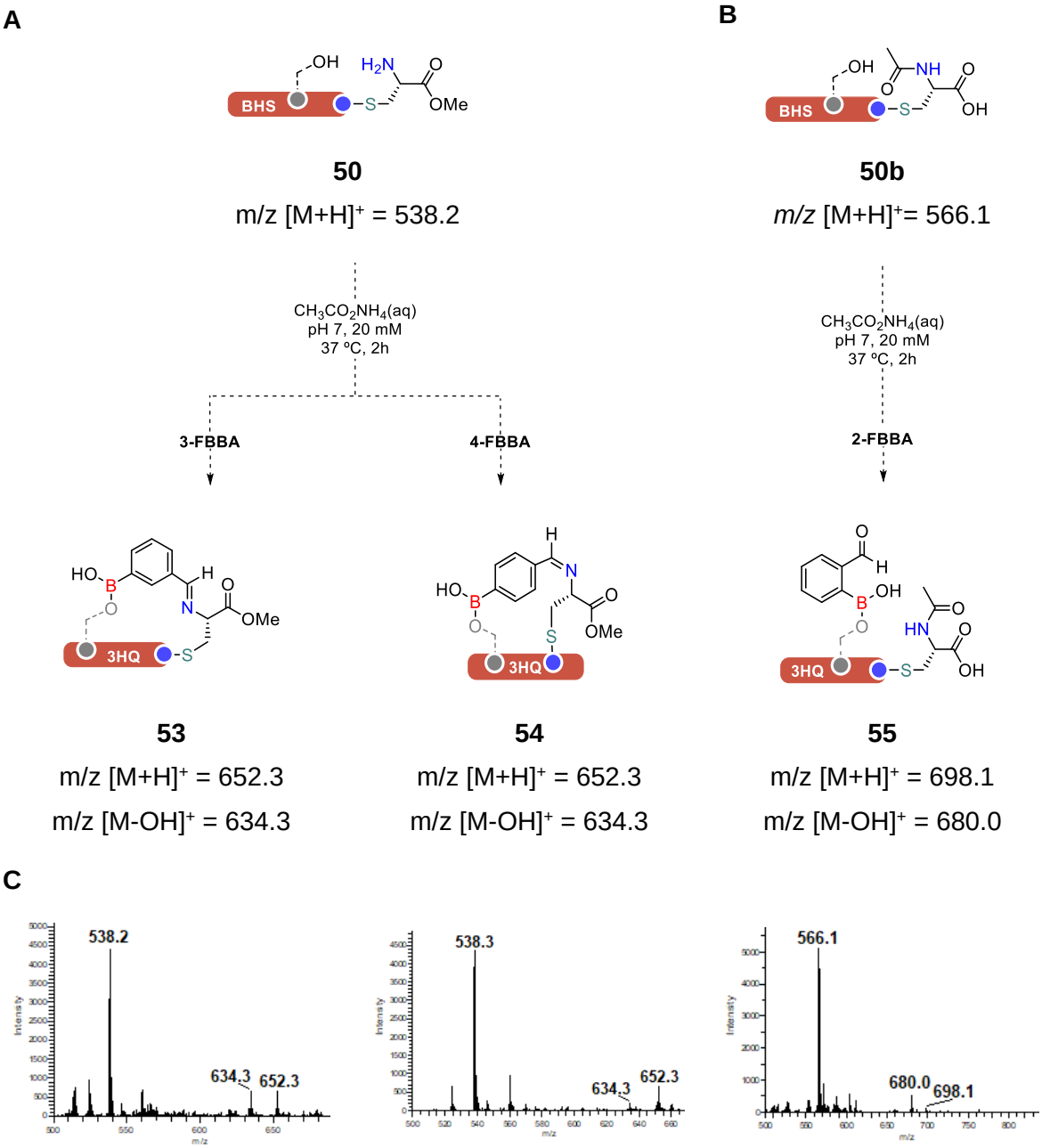
MS, 100 equiv. of 2-FBBA were added to the mixture, which, as shown by the ESI-MS spectrum in **Scheme 21B**, led to a very efficient iminoboronate formation within 2 hours.

The same reaction was then performed using 2-ABBA instead of 2-FBBA, but resulted in poor conversion, probably due to the sterical impediment of the additional methyl group (**Scheme 21C**).



Scheme 21 - Functionalization of CysOMe with 42 and subsequent HQ-iminoboronate formation with 2-CBBAs in bioconjugation conditions (A), ESI-MS spectra of the reaction of 3HQ-CysOMe 50 with 2-FBBA (B) and with 2-ABBA (C).

Moreover, negative results were obtained also when compound **50** was reacted with the other isomers of FBBA, namely the *ortho*- and *para*- substituted, as with these compounds iminoboronate formation is not possible.



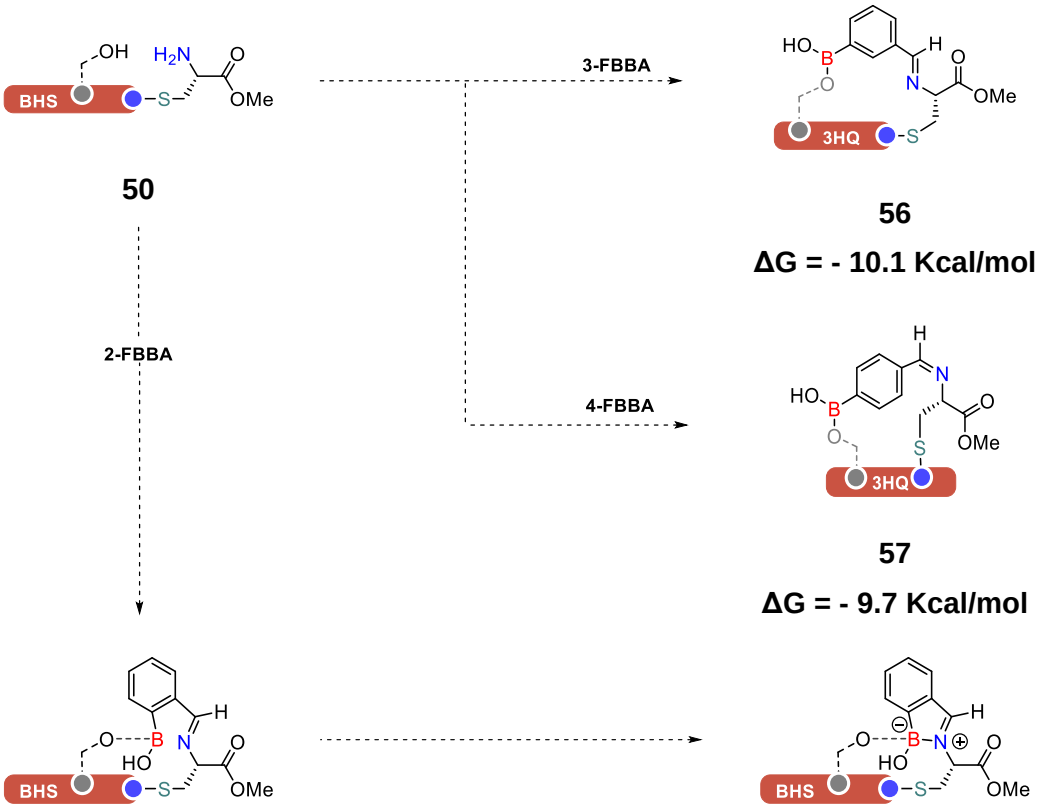
C

Scheme 22 – Reaction of intermediate 50 with FBBA isomers (A), reaction of acetylated intermediate 50b with 2FBBA (B) and relative ESI-MS spectra (C).

In fact, as shown in **Scheme 22A**, the 3- and 4- isomers of FBBA can only form imines with the *N*-terminus amine group. Similarly, the acetylation of the *N*-terminal amino group

in **50b** also limited considerably the reactivity with 2-FBBA, which supports the synergistical formation of the 3HQ-iminoboronate (**Scheme 22B**).

In order to better rationalize and quantify the importance of the iminoboronate formation in this system, DFT calculations were performed to evaluate the energy corresponding to each intermediate of the proposed reaction mechanism. As shown in **Scheme 23**, DFT calculations support that the 3HQ-iminoboronate **51** formation is an overall favourable process, corresponding to an energetic gain of -15.4 Kcal/mol. Additionally, these results support the lack of reactivity observed with the 3- and 4-FBBA isomers, as iminoboronate formation significantly contributes to the overall stabilization energy by ~ 6 Kcal/mol.



Scheme 23 – Energetic profile of the reaction of 3HQ-CysOMe intermediate 50 with the various isomers of FBBA derived from DFT calculations performed using the Gaussian 09 software package

In order to corroborate the observations collected from the ESI-MS experiments and the DFT calculations, the system was also studied by means of ^{1}H, ^{11}B and ^{15}N NMR

spectroscopy. To this end, a 5 mM solution of compound **42** was prepared in a 1:4 mixture of DMSO-*d6* and PBS pH 7.4 and reacted with ^{15}N-enriched cysteine methyl ester. The first step was followed by ^{1}H-NMR spectroscopy. The complete disappearance of the peak corresponding to the double bond of the maleimide indicated the complete conversion to 3HQ-CysOMe **50** intermediate. Next, 1 equivalent of 2-FBBA was added to the mixture, and, after 5 min of reaction at room temperature, it was analysed by using ^{1}H, ^{11}B and ^{15}N NMR.^{11}B-NMR spectroscopy is a powerful tool to study the hybridization of the boron atom, in fact, while planar trigonal sp^2 boron atoms have chemical shifts in the 20-50 ppm region, tetrahedral boronate species with sp^3 hybridization fall in the 0-10 ppm region. As iminoboronate formation causes the boron atom to shift from sp^2 to sp^3 hybridization, this experiment gives valuable insight towards confirming the presence of an iminoboronate species in solution.

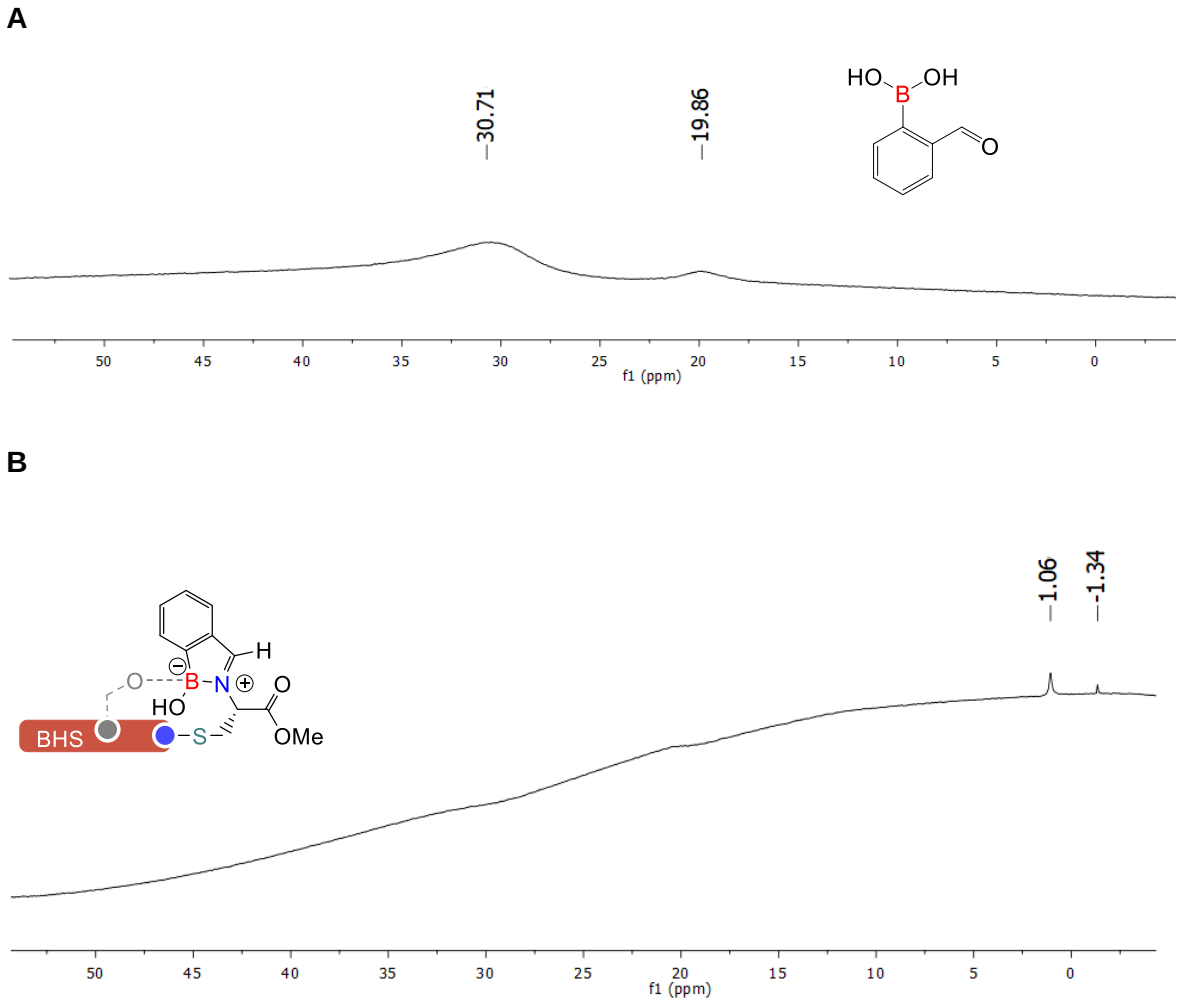


Figure 21 – ^{11}B NMR spectra of 2-FBBA (A) and of 3HQ-iminoboronated 51 (B) obtained in a 1:4 DMSO-d6:PBS solution at RT

As shown in **Figure 21A**, the spectrum of 2-FBBA alone presents two peaks at, respectively, 30.71 ppm and 19.86 ppm, in line with what is reported for uncharged trigonal boronic acid species. On the other hand, the spectrum of the reaction of **50** with 2FBBA shows two peak with, respectively, 1.06 ppm and -1.34 ppm chemical shifts. Having confirmed the complexation of the boronic acid in the system, it was necessary to assess the participation of the nitrogen atom in the ligation, by analysing the ^{15}N-NMR of the reaction. While the chemical shift of the ^{15}N-enriched cysteine methyl ester falls at 35.0 ppm, in line with what expected for a primary amine, the signal of the reaction has a chemical shift of 528.5 ppm, indicating a clear change in the electronic character of the nitrogen atom.

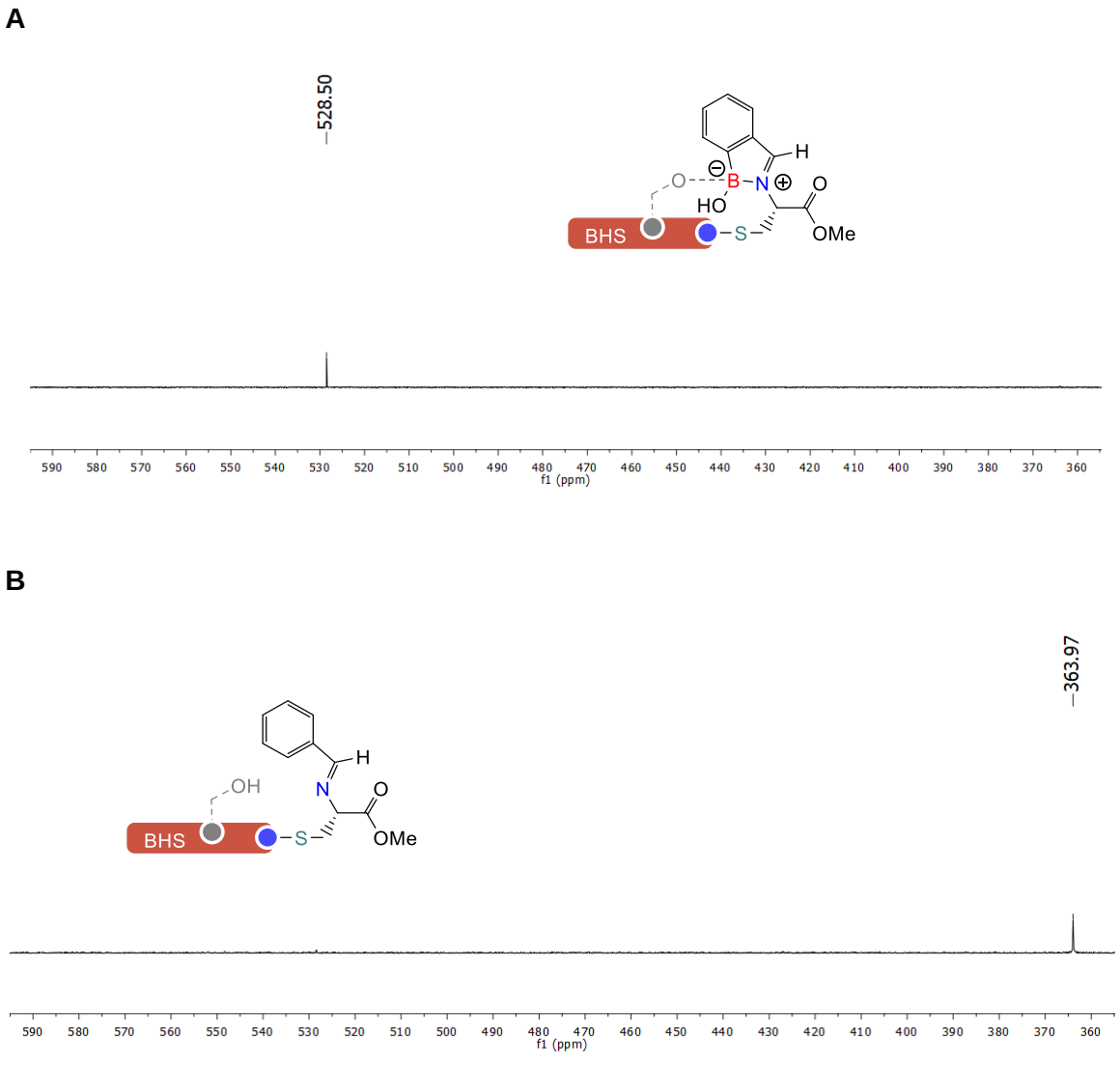


Figure 22 – ^{15}N NMR spectra of product 51 (A) and of its benzylimine analogue (B) obtained in a 1:4 DMSO-d6:PBS solution at RT

In order to confirm that the observed species involves the N-B dative bond postulated for iminoboronates and does not correspond to a classic imine, the same reaction was performed employing benzaldehyde instead of 2-FBBA. Analysis of the ^{15}N-NMR spectrum of this reaction indicates a chemical shift of 364.0 ppm for the imine nitrogen, falling in the expected region for this kind of compounds (**Figure 22**).[128] Confronting the control reaction with the 3HQ-iminoboronate formation reaction, it clearly emerged that the nitrogen is involved in this ligation not as a simple imine, but as a more electron-deficient species, suggesting the existence of the N-B dative bond.

Finally, as the iminoboronate formation was confirmed by both ^{11}B and ^{15}N NMR spectroscopy, ^{1}H NMR spectroscopy was used to confirm the involvement of the 3HQ core in the complex with 2-FBBA. In fact, both ^{11}B- and ^{15}N-NMR experiments confirmed the formation of the characteristic dative N-B bond of iminoboronates. Using water in the experiment instead of deuterium oxide allowed for the determination of the involvement of the 3HQ core in binding 2-FBBA by quantification of the hydroxyl substituents of the BA. While iminoboronates have two hydroxyl substituents on the tetrahedral boron atom, the postulated structure of the 3HQ-iminoboronate presents only one OH group on the boron centre, since the fourth coordination spot is occupied by the phenol of the 3HQ. As shown in **Figure 23**, the peak corresponding to the two hydroxyl groups of 2-FBBA shifts from 8.26 ppm to a peak at 10.07 ppm that integrates for one proton, corresponding to the OH. This kind of chemical shift is also observed in related boron-nitrogen heterocycles, like diazaborines,[129] that also possess a single OH substituent on the boron atom. In conclusion, the various NMR experiments performed on the reaction confirmed the proposed structure for compound **51**.

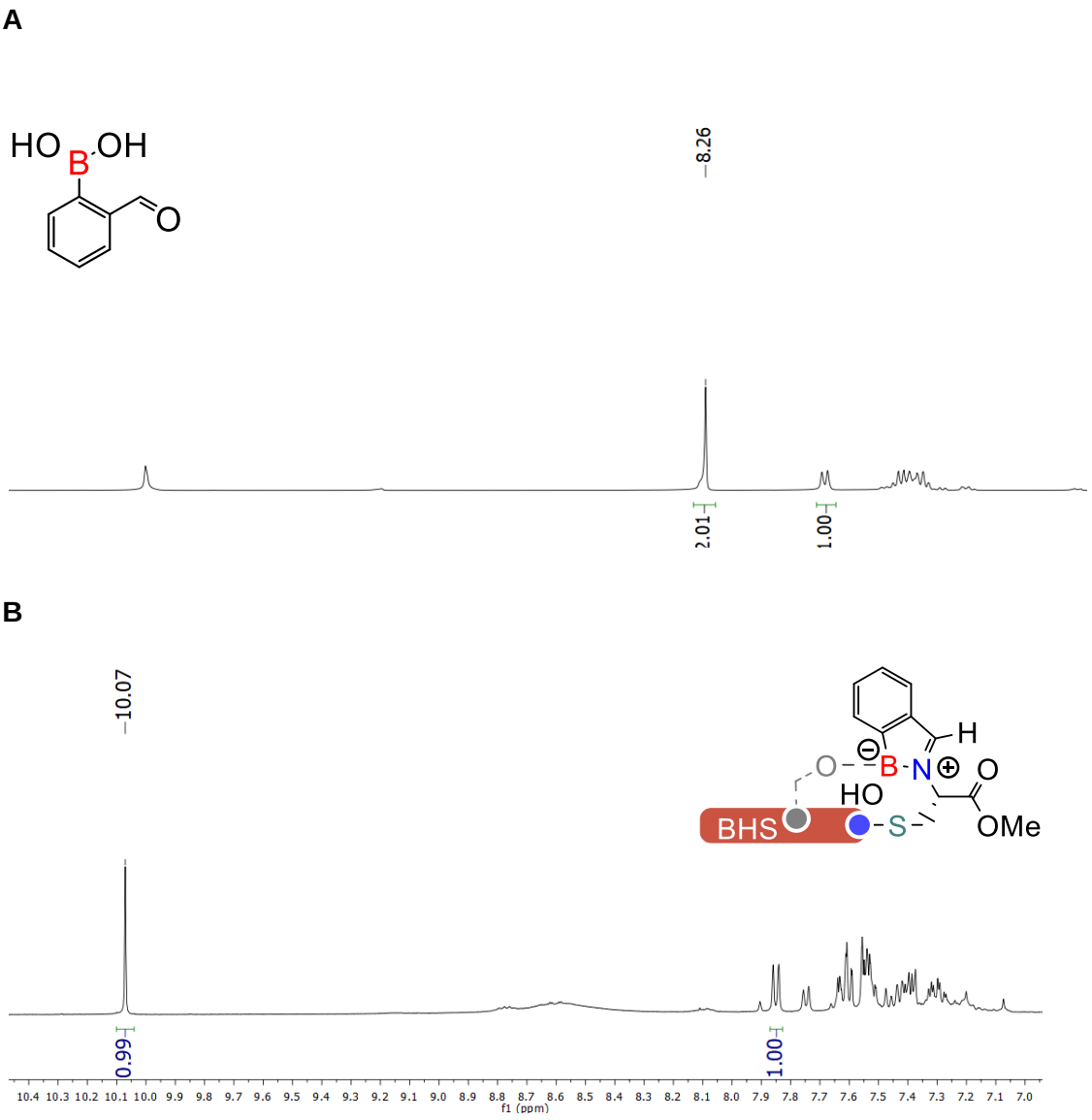
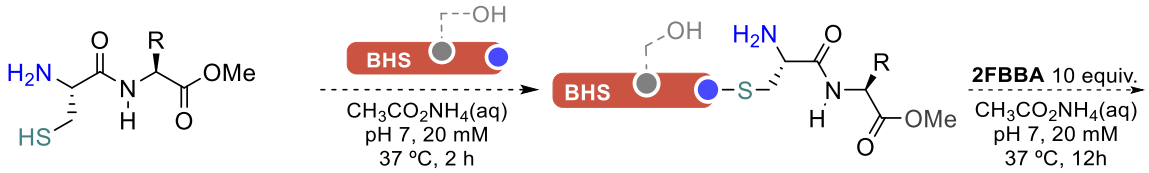


Figure 23 – ^{1}H NMR spectra of 2-FBBA (A) and of product 51 (B)

Having successfully characterized the reaction of 3HQ-iminoboronate formation on cysteine methyl ester using experimental and computational methods, the next step was to apply this methodology in the functionalization of more complex peptides.

In the context of protein modification, the peptidic units are complex structures that expose different functionalities that could interfere with the iminoboronate formation. Specifically, as the objective of this work is the development of a methodology that is site-selective towards *N*-terminal cysteines, it was imperative to ensure that the presence of other amino acid side chains on the peptidic target did not yield cross-reactivity.

A



58 R: $CH_2CH(CH_3)_2$ [Ile] **42**

59 R: CH_3 [Ala]

60 R: BnOH [Tyr]

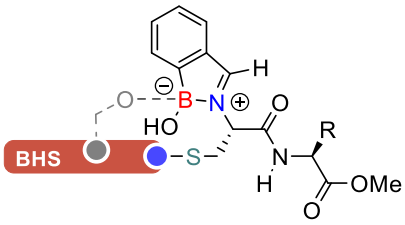
61 R: $(CH_2)_2CO_2H$ [Glu]

58a: *m/z* $[M+H]^+$= 651.2

59a: *m/z* $[M+H]^+$= 609.2

60a: *m/z* $[M+H]^+$= 701.2

61a: *m/z* $[M+H]^+$= 667.2



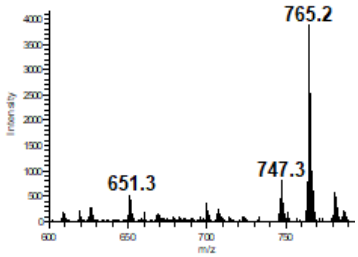
58b: *m/z* $[M+H]^+$ = 765.2; $[M-OH]^+$ = 747.3

59b: *m/z* $[M+H]^+$ = 732.2; $[M-OH]^+$ = 705.3

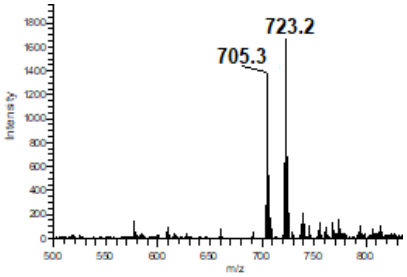
60b: *m/z* $[M+H]^+$ = 815.2; $[M-OH]^+$ = 797.3

61b: *m/z* $[M+H]^+$ = 781.3; $[M-OH]^+$ = 763.3

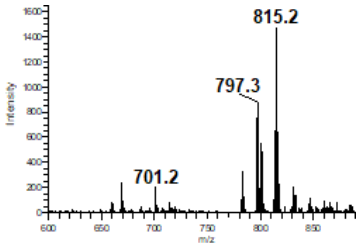
B



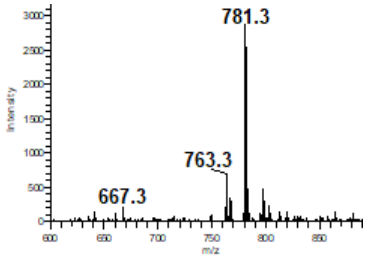
C



D



E

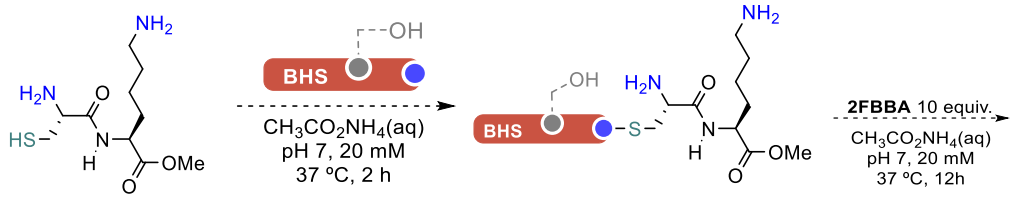


Scheme 24 – Functionalization of dipeptides 58 - 61 with BHS 42 and subsequent reaction with 2FBBA in bioconjugation conditions (A) along with spectra corresponding to the formation of products 58b (B), 59b (C), 60b (D) and 61b (E).

As a preliminary study regarding the compatibility of 3HQ-iminoboronate formation with different peptidic side chains, a small library of *N*-terminal cysteine dipeptides was employed as substrate for modification. As shown in **Scheme 24**, dipeptides **58** -**61** were first modified with BHS **42** in ammonium acetate pH 7 solution at 37 ºC, obtaining the corresponding BHS-modified dipeptides at 0.2 mM concentration. All of these constructs were subsequently reacted with 20 equiv. of 2FBBA and stirred for 12h before ESI-MS analysis. Gratifyingly, all dipeptides reacted as expected, generating the desired products with excellent conversion, despite the presence of various potentially interfering moieties.

Aditionally, peptide **62**, which presents a lysine side chain, was reacted in the same conditions used in the same model reaction. This dipeptide has two amino groups and for that reason, upon reaction with 2FBBA, can generate iminoboronate both at the *N*-terminus and on the ε amine of the lysine, potentially leading to a heterogeneous mixture of mono- and di-substituted peptides. Despite this, as shown in, the ESI-MS spectrum of the reaction between **62a** and 2FBBA shows only the mono-substituted product.

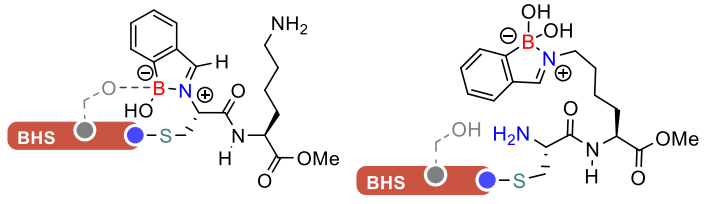
A



62 **42** **62a**

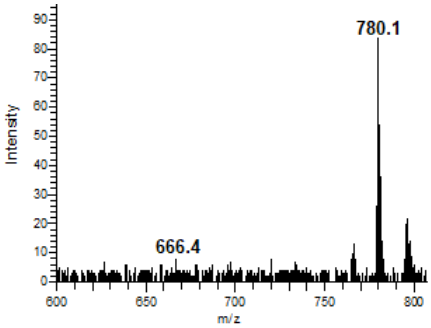
m/z $[M+H]^+$ = 666.4

B



62b **62c**

m/z $[M+H]^+$ = 780.1 m/z $[M-OH]^+$ = 780.1



Scheme 25 – Functionalization of Cys-Lys dipeptide 62 with BHS 42 and subsequent reaction with 2FBBA (A) along with ESI-MS spectrum of the reaction (B)

As shown in **Scheme 25**, two mono-substituted products are possible in this reaction, **62b** and **62c**, MS-MS fragmentation experiments were performed, with the objective of elucidating the site of the modification. The fragmentation of the signal with *m/z*=780.1 yielded a fragment with *m/z*=575.2, which can be assigned to radical cation **62d**, shown in **Figure 24**. This fragmentation pattern suggests that the iminoboronate is preferentially formed at the *N*-terminus and supports the importance of the 3HQ in directing the iminoboronate formation.

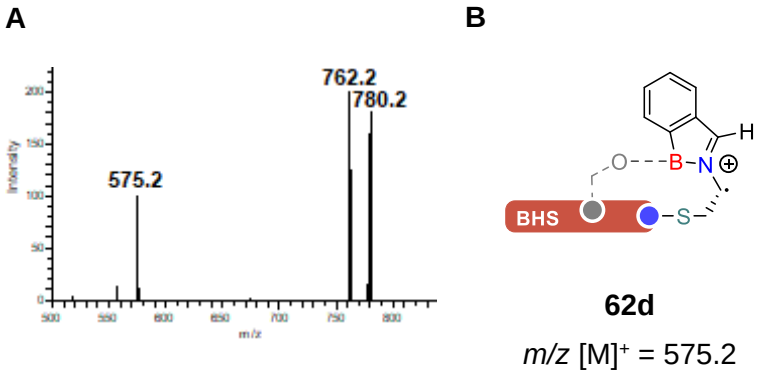
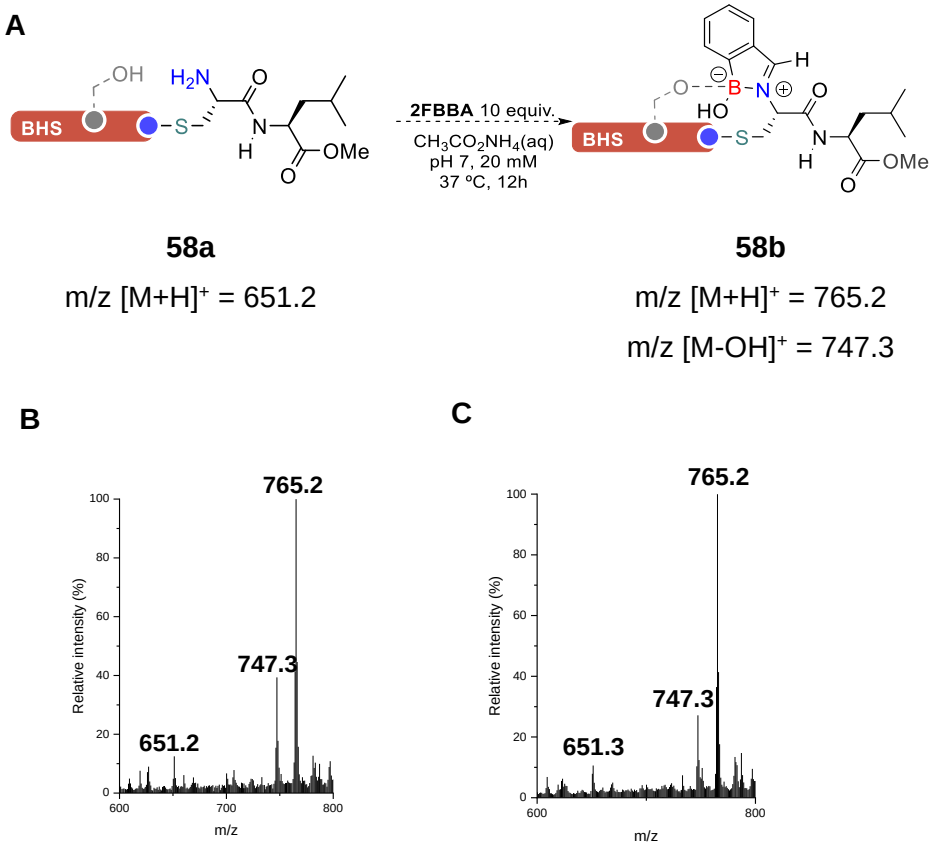


Figure 24 – MS-MS fragmentation spectrum of the peak with *m/z*=780.2 (A) and proposed structure of the radical cation fragment corresponding to *m/z*=575.2 (B).

Once etablished that the 3HQ-iminoboronate formation is not only compatible with different amino acid side chains, but is also site-selective towards *N*-terminal cysteine residues, the next issue to address was the stability of the 3HQ-iminoboronate assembly. In fact, as mentioned before, the susceptibility of iminoboronates to be hydrolysed in the presence of carbohydrates limits their use in many biological settings. Since the affinity of 3HQ derivatives towards BA complexation was measured to be significantly higher than the one of different biologically-occurring diols, 3HQ-iminoboronated were expected to show an improved stability profile in the presence of several biomolecules. In order to test this hypothesis, the Cys-Ile 3HQ-iminoboronate **58** was incubated in different conditions of pH (7.0 and 4.5) and in the presence of different biomolecules, such as glucose, bovine serum albumin (BSA) and glutathione (GSH). ESI-MS spectra of the obtained solutions were recorded over time in order to obtain a qualitative analysis of the behaviour of 3HQ-iminoboronates in the aforementioned conditions. As shown in **Scheme 26**, the ESI-MS spectrum of product **58b** remained unchanged during a 4 weeks incubation time in ammonium acetate solution pH 7 at 37 ºC, indicating remarkable stability in neutral aqueous conditions.

A second solution, analogous to the one used in the previous experiment, was treated with acetic acid, until reaching pH 4.5, then stirred at 37 ºC. After 32 hours in these conditions, the ESI-MS spectrum of this reaction did not show a significant decomposition of the construct, compared to the starting solution, indicating good tolerance of the 3HQ-iminoboronate towards acidic conditions (**Figure 25A**).



Scheme 26 – 3HQ-iminoboronate formation on Cys-Leu dipeptide 58 (A) and relative ESI-MS spectra after 12 hours (B) and 4 weeks (C) of incubation at pH 7 and 37 ºC.

Similarly to what was observed in acidic conditions, the presence of either glucose or BSA did not significantly affect the ESI-MS spectrum of the reaction, indicating good tolerance of the 3HQ-iminoboronate towards these endogenous biomolecules in a 32h period (**Figure 25B and C**). On the other hand, the presence of glutathione significantly affected the reaction profile, leading to an increase in the intensity of signal corresponding to compound **58b** after 8h of incubation, suggesting that GSH mediates the hydrolysis of the complex (**Figure 25D**). These stability studies indicated an overall improved stability profile for 3HQ-iminoboronates compared to traditional iminoboronates, which would have been

hydrolysed in all the conditions reported. Importantly, despite showing increased stability, 3HQ-iminoboronates are still susceptible to hydrolysis promoted by GSH, hence they did not lose the valuable reversibility and stimulus-responsiveness that characterize iminoboronate technology. GSH, in particular, is a valuable stimulus to trigger the release of a payload, as it is known to be present in abnormally high concentrations inside cancer cells, making the 3HQ-iminoboronate technology a useful platform for selective payload release in tumoural environment.

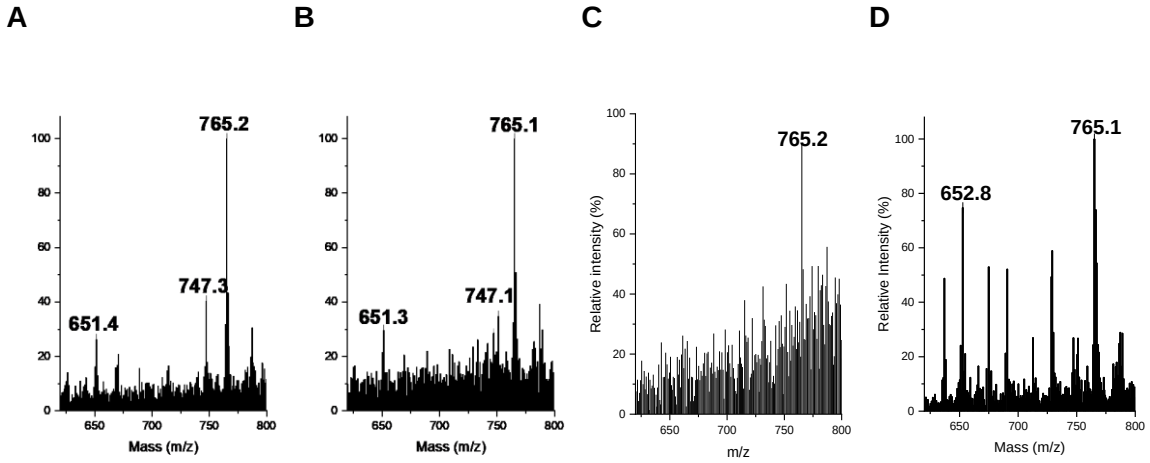


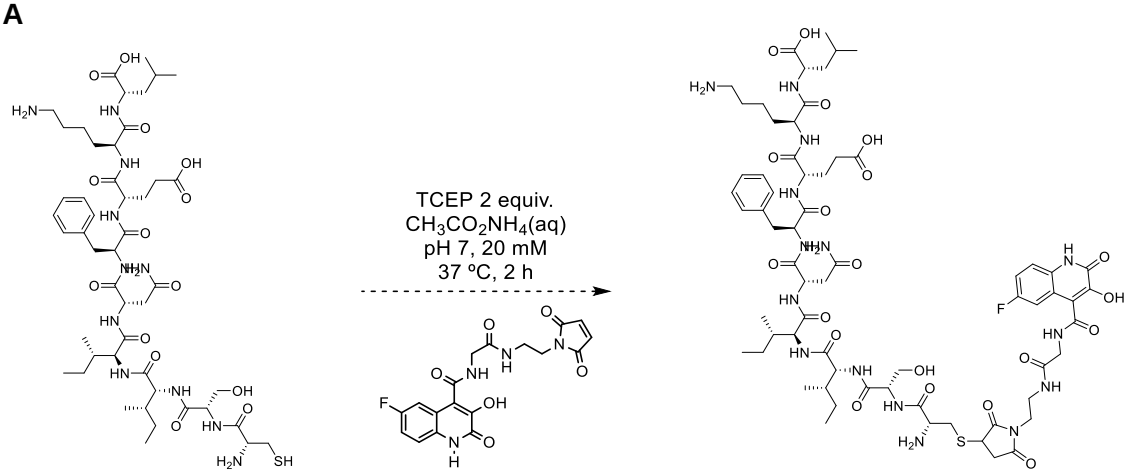
Figure 25 – ESI-MS spectra of compound 58b incubated at pH 7 for 32h (A), with 10 eq of glucose for 32h (B), with 10% BSA for 32h (C) and with 10 eq of GSH for 8h (D)

III.2.2 Peptide modification

After characterizing the 3HQ-iminoboronate formation from a structural and mechanistic point of view with model peptides, this technology was then applied to more complex peptides. We initiated this study by evaluating the site-selectivity of the technology in the reaction with a c-Ovalbumin fragment that exhibits a competing lysine residue.

This peptide was first modified with the BHS **42** in ammonium acetate pH 7, yielding compound the corresponding 3HQ-modified peptide **63** (**Scheme 27**). This derivative was then reacted at 37 ºC with 10 equivalents of 2FBBA for 12h, yielding the 3HQ-iminoboronate **64** with peak to peak conversion.

A



c-Ovalbumin fragment

m/z $[M+H]^+$: 1067.3

m/z $[M+2H]^{2+}$: 534.1

42

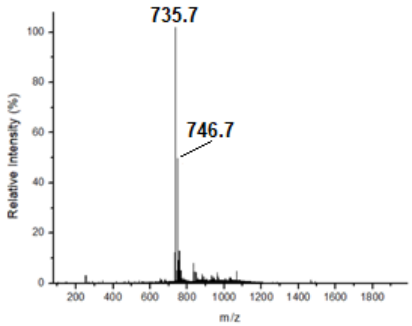
63

m/z $[M+H]^+$ = 1468.6

m/z $[M+2H]^{2+}$ = 735.7

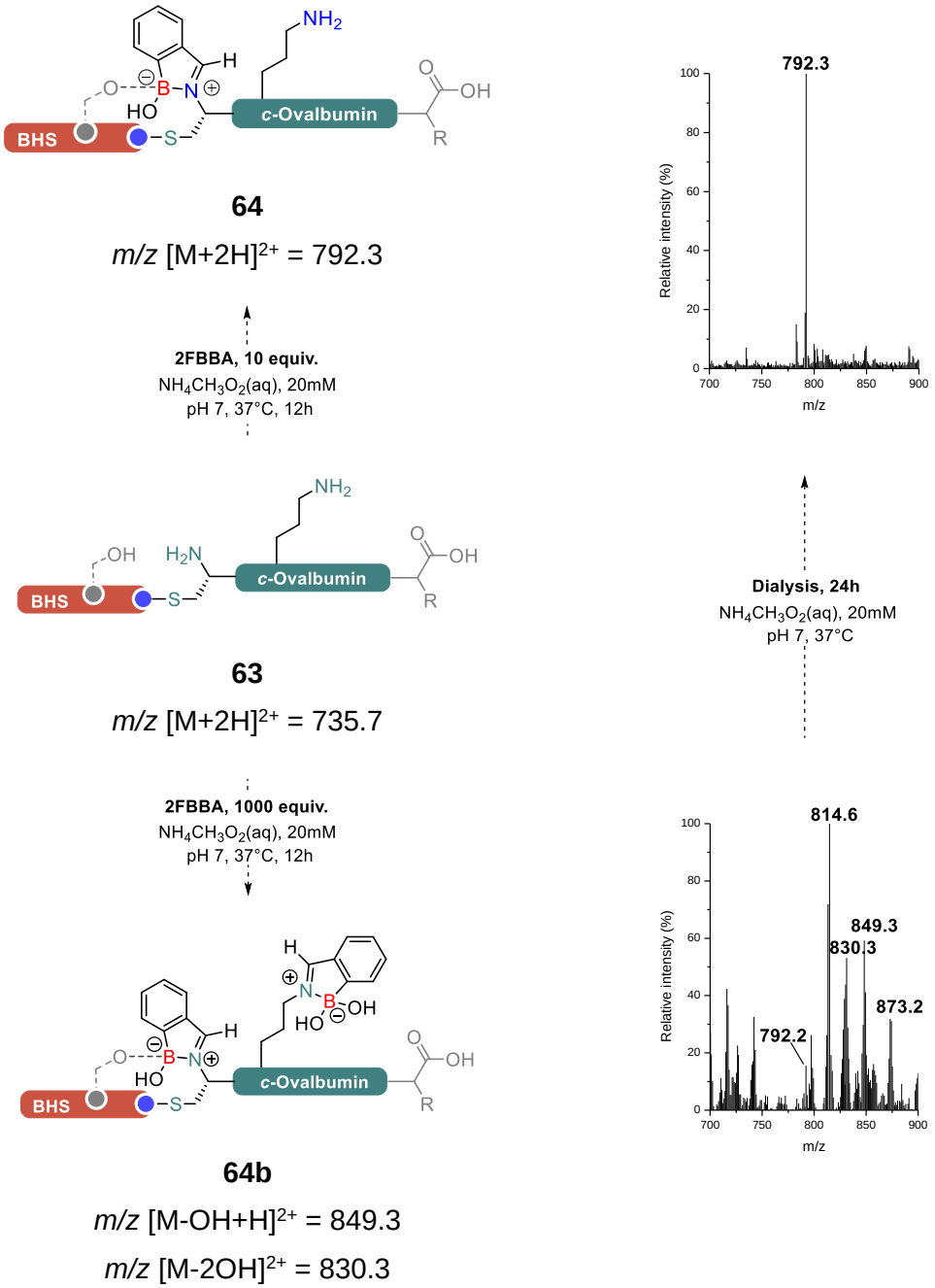
m/z $[M+H+Na]^{2+}$ = 746.7

B



Scheme 27 – Modification of the c-Ovalbumin fragment with 3HQ-BHS 42 (A) and relative ESI-MS spectrum (B)

Notably, it was possible to remove the excess of 2FBBA from the reaction mixture by dialysis, employing a Pur-A-Lyzer™ Midi Dialysis Kit with a 1 KDa cut-off. In fact, the resulting solution of **64** was dialyzed in a 1000-fold volume of ammonium acetate pH 7 over 24h, which guaranteed the removal of the excess of BA from the reaction mixture. The fact that product **64** could be recovered intact indicates that the 3HQ-BHS approach leads to remarkably stable adducts, which can resist to dilution. The stability of the 3HQ-iminoboronates in dialisys conditions is a feature that few BA-based bioconjugation strategies possess, and constitutes a valuable property, allowing for easy and straightforward purification of the conjugates based on this platform.



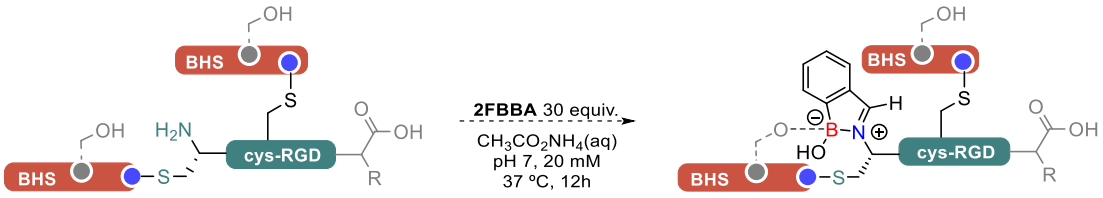
Scheme 28 – Reaction of 3HQ-c-Ovalbumin derivative 63 with 10 or 1000 eq of 2FBBA, along with ESI-MS spectra corresponding to the resulting products 64 and 64b

When 3HQ-c-Ovalbumin **63** was reacted with 1000 equivalents of 2FBBA, it was possible to obtain the double-modified peptide **64b**, bearing both the 3HQ-iminoboronate on the *N*-terminal cysteine along with a classic iminoboronate on the lysine side-chain. Interestingly, after dialyzing compound **64b** for 24h in the same conditions applied before, the only

product that could be detected by ESI-MS was **64** (**Scheme 28**). This result further supported the already observed improved stability of 3HQ-iminoboronates.

The data gathered from the experiments with the c-Ovalbumin confirmed the reactivity and properties that emerged in the model experiments with cysteine dipeptides, indicating the suitability of the 3HQ technology for the construction of homogeneous conjugates by site-selective iminoboronate installation.

A



65

m/z $[M+H]^+$: 1587.5

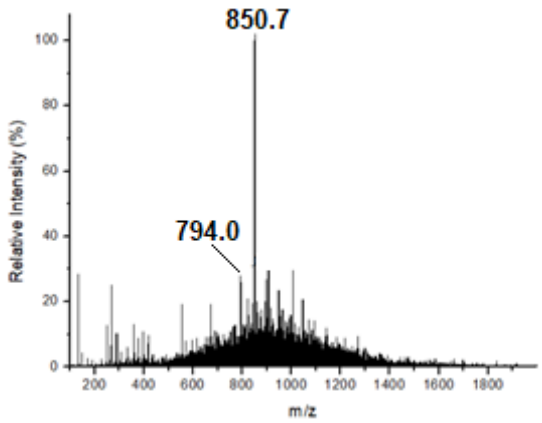
m/z $[M+2H]^{2+}$: 794.0

66

m/z $[M+H]^+$ = 1700.4

m/z $[M+2H]^{2+}$ = 850.7

B



Scheme 29 – Reaction of the 3HQ-functionalized cys-RGD peptide 65 with 2FBBA (A) along with the ESI-MS spectrum of the reaction (B).

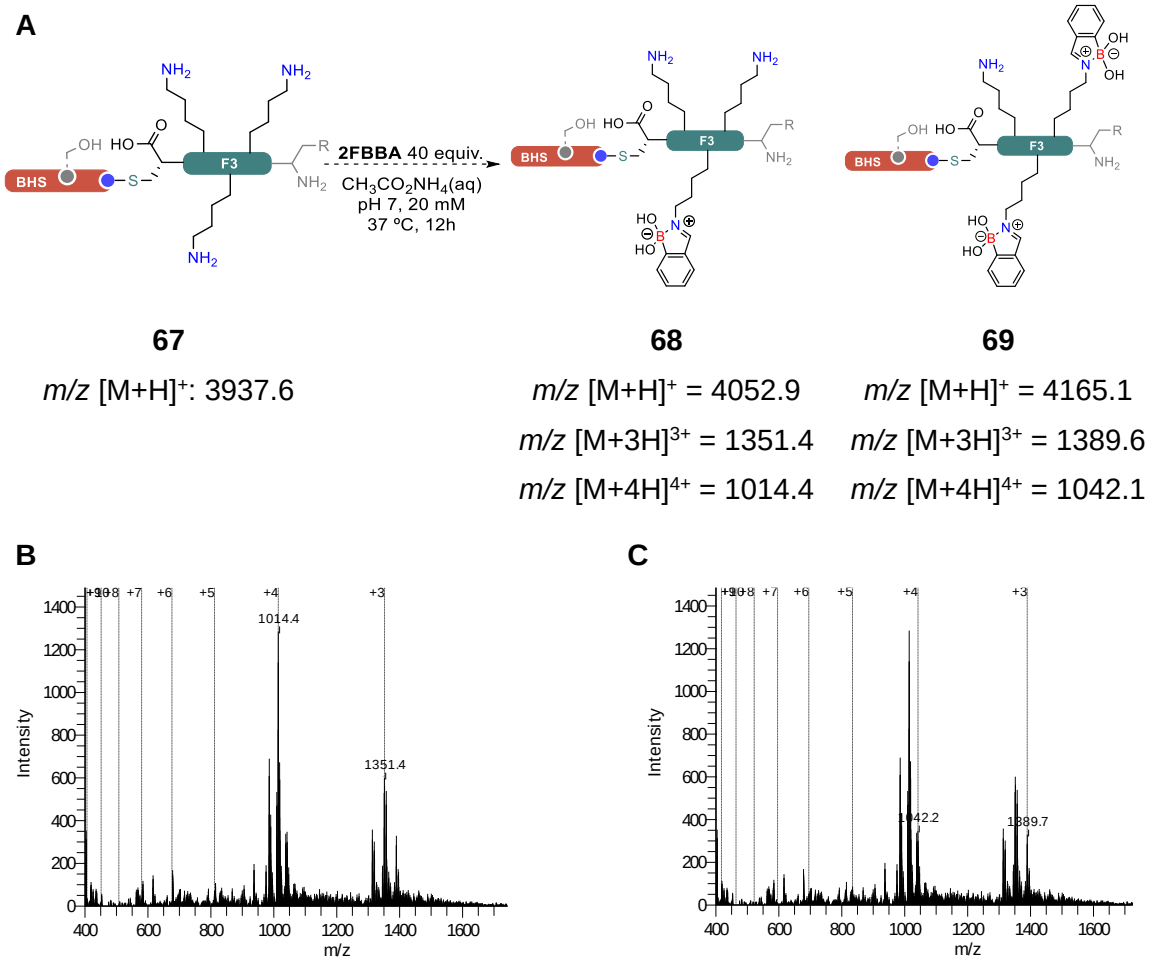
Subsequently, other peptides were employed as targets for the 3HQ-mediated iminoboronate formation, with the objective of validating the applicability of this platform towards *N*-terminal cysteine modification. With this objective, a derivative of the RGD peptide, featuring an *N*-terminal cysteine, was selected as the next peptide to modify with this technology. The RGD sequence (Arg-Gly-Asp) is a short peptidic chain that has been employed as a targeting moiety in the framework of drug delivery in oncology. In fact, this sequence is recognized and internalized by various receptors of the integrin family, which

are often overexpressed on the surface of cancer cells.[130] Expanding on the known RGD motif, the sequence CGGCGGDGR was adopted for this experiment, exhibiting both *N*-terminal and in-chain cysteine residues.

Following the same protocol described before, the 3HQ hot spot **42** was inserted on the cys-RGD peptide. As expected, both cysteine residues reacted with compound **42**, resulting in a peptide with a double 3HQ modification, compound **65**. Upon incubation of this derivative with 2FBBA, only one product was formed, despite the presence of two 3HQ moieties on the peptide, the product formed corresponded to the addition of only one molecule of 2FBBA upon the structure, confirming that the 3HQ core only allows BA ligation on *N*-terminal cysteine residues (**Scheme 29**).

Finally, the anti-nucleolin F3 peptide was tested as a substrate for 3HQ modification, with the objective of confirm the specificity of our approach. In fact, this peptide features a *C*-terminal cysteine residue onto which the 3HQ core can be installed, but is not expected to form 3HQ-iminoboronates. However, two lysine residues are present in the proximity of the *C*-terminal cysteine of the F3 peptide, potentially leading to the formation of non-specific 3HQ-stabilized iminoboronates. This peptide has been used in various applications as a targeting unit, as it is recognized and internalized by the nucleolin receptor, which is up-regulated in several cancer cell lines.[131]

Straightforward modification of the *C*-terminal cysteine with 3HQ boron hot-spot **42** was carried out following the same protocol adopted for the modification of the previously reported peptides, yielding F3 derivative **67**. Reaction of 3HQ-peptide **67** with 2FBBA led to the formation of two products, corresponding to the addition of either one or two 2FBBA units onto the peptidic scaffold. As the F3 peptide presents several lysine residues, it was expected for it to be able to form iminoboronates upon reaction with 2FBBA, however, the role of the 3HQ core in these product had to be clarified. In fact, one of the possibilities regarding the products observed was that the 3HQ boron hot-spot could have participated in the formation of stabilized iminoboronates with the ε amino group of one of the lysine residues present on the structure. However, after dialyzing the reaction mixture for 24h in the same conditions adopted earlier, only 3HQ-modified peptide **67** could be detected in solution. This result confirmed once again the site-selectivity of this approach because, even though **67** contained both the 3HQ boron hot-spot as well as several exposed amino groups, the formation of stabilized 3HQ-iminoboronates was not detected as it is limited uniquely to *N*-terminal cysteine residues.



Scheme 30 – Reaction of 3HQ-F3 peptide 67 with 2FBBA (A) and ESI-MS spectrum of the reaction, with deconvolution of the peaks corresponding to product 68 (B) and 69 (C).

After the reported experiments with model peptides, it was demonstrated how the 3HQ boron hot-spot platform is a useful tool for the construction of homogenous conjugates, combining the stimulus-responsiveness of iminoboronates with a more robust scaffold

III.2.3 Construction of a fluorescent bioconjugate

Having developed a reliable methodology for the site-selective installation of stabilized iminoboronates on *N*-terminal cysteines, we next applied this technology in the construction of a fluorescent bioconjugate. Given the characteristics of the 3HQ-based BHS, namely its ability to release a boronated payload as a response to the presence of glutathione, it constitutes a potentially useful platform for the delivery of small molecules in the tumoural environment through the targeting of a non-internalizing receptor.

In the context of therapeutic bioconjugates, internalizing receptors have predominantly been the target of choice for the delivery of cargo to cells, as receptor-mediated internalization constitutes a convenient pathway for ensuring that the payload reaches the intracellular environment. This approach, however, has the intrinsic limitation of relying on the expression of the receptors of choice to be effective, meaning that, in the heterogeneous population of cells that constitute the tumoural tissue, the ones that do not express the targeted receptor will not receive the payload. Another issue with this approach is target accessibility, as targeting units are often large entities and have a limited ability to reach cells that are buried in the malignant tissue to deliver a payload which relies on receptor-mediated internalization in order to elicit its effect.

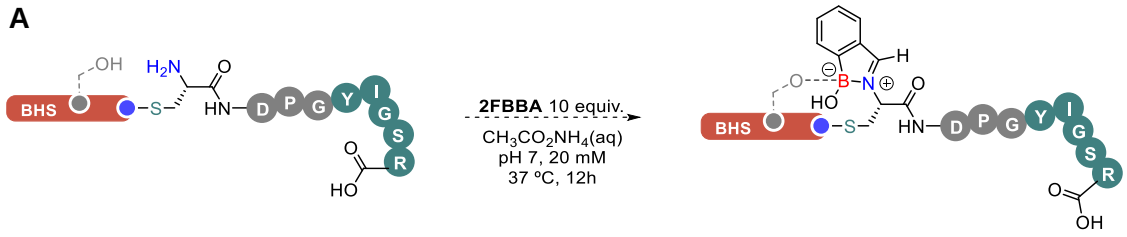
Both of these issues can be addressed by employing cleavable bioconjugates that target non-internalizing receptors overexpressed in the desired location.[132] This approach consists in using a targeting unit that is selectively recognized by a non-internalizing surface receptor overexpressed in the diseased tissue of choice. In order for this strategy to be successful, however, the conjugate has to be able to release the payload in response to a stimulus that is present in the diseased area and, crucially, the payload must be able to spontaneously enter cells. The clear advantage of this approach is that, as the payload is able to spontaneously internalize cells and it is released in the diseased tissue, even cells that do not express the receptor that is being targeted, will uptake the therapeutic entity. This so-called “bystander effect” allows to efficiently deliver the active molecule to a heterogeneous population of cells that, while all part of the same malignancy, might express different surface receptors, as in the case of many tumours. Additionally, as the payload is a small molecule with good cell-permeating capabilities, target accessibility is less of an issue in this case, as this system does not rely on receptor-mediated internalization of large biomolecules.

Although this approach is very appealing, it is difficult to design conjugates with these properties, as they need to meet specific requirements. Namely, the payload needs to be able to passively permeate cells without the help of a receptor and it must be linked to the targeting unit in a reversible manner that allows its release in the diseased tissue.

The boron hot-spot technology described in this chapter presents all the characteristics that are needed to build a non-internalizing bioconjugate, as the 3HQ-iminoboronate unit has shown good stability in biological conditions, while being reversible in the presence of glutathione, which is abundant in tumour tissue. Also, boronated payloads have

remarkable cell-penetrating capabilities, as the transient formation of boronate esters with the glycans on the cell surface promotes their internalization.[70,71]

A



47

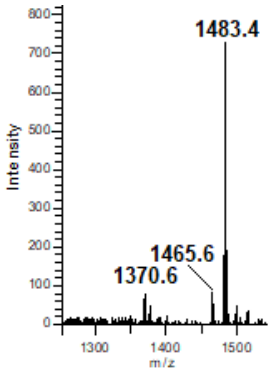
m/z $[M+H]^+$= 1370.6

70

m/z $[M+H]^+$= 1483.4

m/z $[M-OH]^+$= 1465.6

B



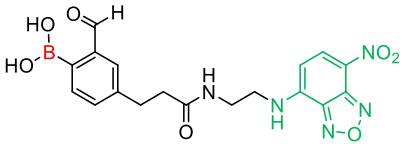
Scheme 31 – Reaction of 3HQ-laminin 47 with 2FBBA (A) and relative ESI-MS spectrum (B)

The targeting unit of choice for this application was the CDPGYIGSR laminin fragment, already employed as a model peptide in the previous chapter. This short peptide belongs to the laminin protein and contains the YIGSR sequence, which is the minimal motif necessary for the recognition by the 67LR receptor. As previously mentioned, the 67LR is a non-internalizing cell surface receptor that is overexpressed in several cancer cell lines, it is involved in cellular motility and metastasis, constituting a biomarker of poor prognosis.[122]

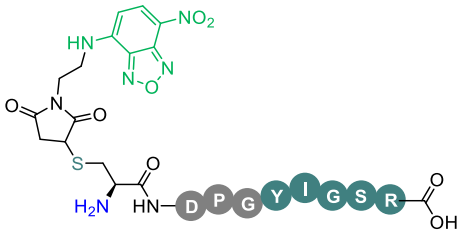
With this in mind, a proof of concept for the validity of this approach was devised by using a fluorescent 2FBBA derivative as the payload to attach to the laminin fragment *via* 3HQ-iminoboronate formation.

As a preliminary study, 3HQ-modified laminin fragment **47** was reacted with 2FBBA in bioconjugation conditions, readily affording derivative **70** (**Scheme 31**). This derivative was

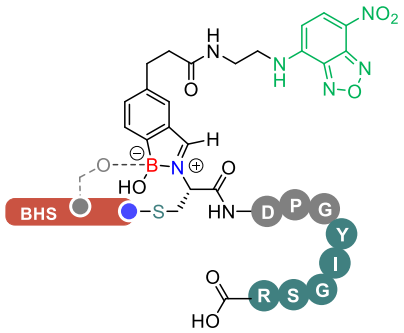
purified by dialysis in the same conditions applied before, allowing for the removal of the excess of 2FBBA without hydrolysing the assembly. Subsequently, fluorescent 2FBBA-NBD derivative **71** was employed in the same reaction, yielding 3HQ-laminin-NBD bioconjugate **73**. In order to better appreciate the importance of having a cleavable ligation between the laminin fragment and the fluorescent payload, derivative **73** was also prepared by using standard maleimide chemistry to produce a covalently linked analogue of **72**. These two peptide conjugates have the same payload and the same targeting unit, as the only difference between them is the fact that the 3HQ-iminoboronate linkage of **72** is reversible, while the thioether ligation of **73** is remarkably more stable.



71



72



73

Scheme 32 – Structure of NBD-2FBBA fluorescent probe 71, 3HQ-laminin-NBD reversible conjugate 73 and of covalent NBD-laminin conjugate 72

Conjugates **73** and **72** were used *in vitro* to evaluate their ability to promote the internalization of the fluorescent probe inside HT29 colon cancer cells, which are known to overexpress the 67LR laminin receptor. 2FBBA fluorescent probe **71** was also included in this assay, in order to confirm the ability of boronated compounds to passively penetrate cell membranes efficiently.The three selected compounds were incubated with HT-29 colon cancer cells for 1h, then the culture medium was washed and substituted with fresh one and then each well was analysed by confocal microscopy, yielding the images reported in **Figure 26**.

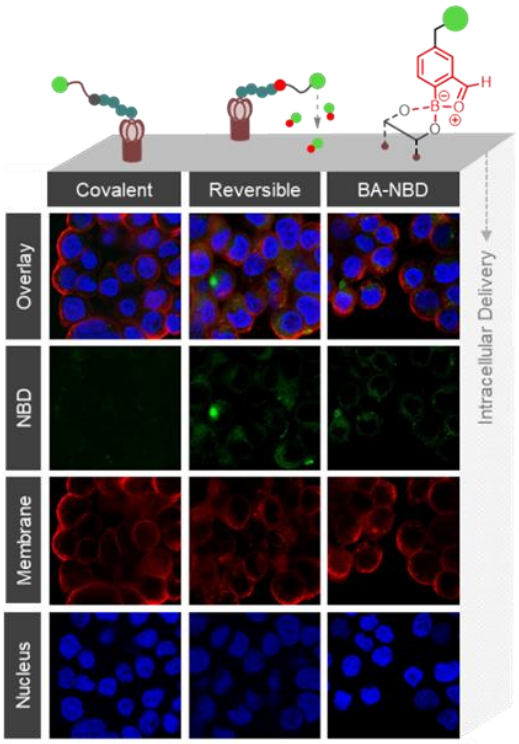


Figure 26 – Confocal microscopy images of the incubation of compounds 71**,** 73 **and** 72 **in HT29 colon cancer cells.**

The images obtained by confocal microscopy show the nucleus and the membrane of the cells marked, respectively, with Hoechst 33258 (blue) and WGA–A647 conjugate (red), while the fluorescence signal of the NBD probe was visualized in green. As shown in **Figure 26**, NBD-BA probe **71** was spontaneously able to permeate the cells, resulting in a green fluorescence emission inside the cellular compartment. The comparison between the images regarding incubation of conjugates **73** and **72** showed a very clear pattern, as only the experiments with reversible conjugate **73** yielded internalization of the fluorescent probe. Moreover, quantification of the fluorescence inside the cells showed that the ones treated with **73** had an intensity 3.5 times superior to the ones treated with only borylated NBD **71** (**Table 2**).

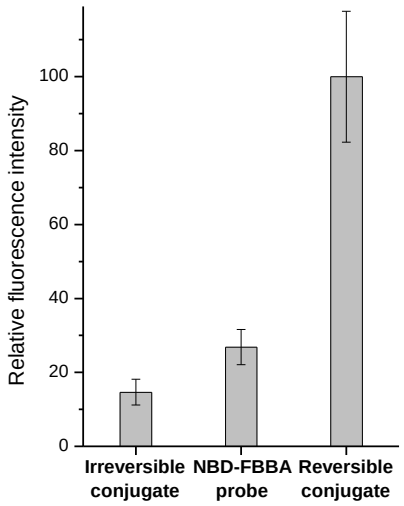


Table 2 – Relative fluorescence intensity measured inside the cells incubated with: Irreversible conjugate 72, NBD-FBBA fluorescent probe 71 and reversible conjugate 73.

Altogether these results confirmed the behaviour that was expected from each compound. First, the spontaneous internalization of borylated compounds was confirmed by the experiments involving compound **71**. Comparison of the results obtained with the two laminin conjugates clearly indicates that only the cleavable construct is able to promote the internalization of the probe, as the 67LR is a non-internalizing receptor. In addition to this, the fluorescent intensity obtained inside the cells treated with reversible conjugate **73** was significantly higher than the one obtained for the probe alone. This result supports the importance of using the laminin peptidic fragment as a targeting unit to increase the local concentration of the payload.

III.3 Conclusions

In this chapter, the maleimide-3HQ derivative **42** was applied as a boron hot-spot for the site-selective installation of stabilyzed iminoboronates on *N*-terminal cysteine residues.

This system was first tested in test reactions using cysteine methyl ester as model compound, confirming the expected reactivity *via* both ESI-MS and NMR experiments. Our approach was then applied to the modification of a small library of *N*-terminal cysteine dipeptides, yielding the expected iminoboronate products in the presence of different side chains that could have interfered with the reaction. Notably, one of these dipeptides contained a lysine unit which, with its ε amino group, could have given cross reactivity with

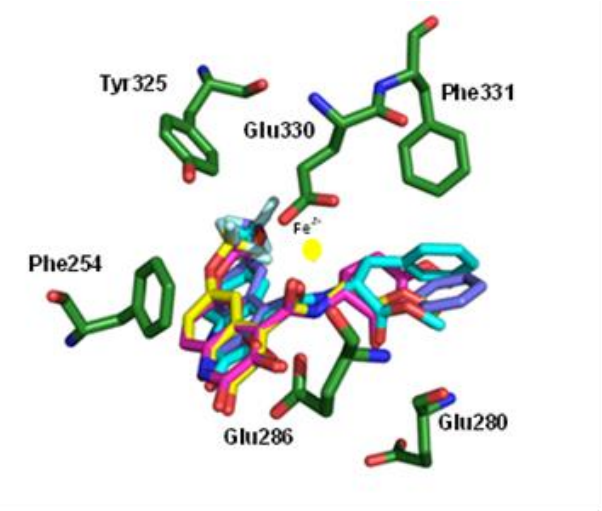
2FBBA. Gratifyingly, only the 3HQ-iminoboronate product was formed in this case and its structure was confirmed by MS-MS fragmentation experiments. Moreover, the 3HQ-iminoboronate product of the Cys-Ile dipeptide was employed for the assessment of the stability of the assembly in various conditions. Incubation of this compound at pH 7 and 4.5 had no significant effect on its stability, as also the presence of glucose or BSA did not promote hydrolysis of the 3HQ-iminoboronate after 32h. However, the presence of GSH in the reaction medium was found to promote the partial hydrolysis of the assembly after 8h of incubation. These results suggested that the iminonoboronates formed *via* our BHS platform have remarkable stability in aqueous media, but are still cleaved in the presence of GSH, which is an important biomarker of tumour cells.

The use of our BHS platform was then extended to more complex peptidic fragments, resulting in the formation of the stabilized 3HQ-iminoboronate on the *N*-terminal cysteine residue of both c-Ovalbumin and Laminin fragment. Importanly, despite the presence of a lysine residue in the structure of the c-Ovalbumin peptide, installation of the BHS unit allowed for the site-selective formation of iminoboronate only on the *N*-terminus of the peptide. Additionally, the 3HQ-iminoboronates formed on peptides proved to be stable in dialysis conditions, which is a remarkable feature for this kind of constructs, as it allows for an easy and straightforward purification of the conjugates obtained *via* this methodology.

Finally, this methodology was applied for the construction of a reversible fluorescent bioconjugate targeting the 67LR receptor, which successfully delivered its fluorescent payload inside HT29 cancer cells. This result allowed not only to confirm the applicability of our methodology in the construction of functional conjugates, but also led to the validation of the 67LR as a potential target for non-internalizing conjugates for cancer therapy.

Chapter IV

3HQs as New Human Phenylalanine Hydroxylase Multivalent Modulators



Abstract

Phenylketonuria (PKU) is an inherited metabolic disease, arising from mutation of the gene coding for the phenylalanine hydroxylase protein (PAH). As this protein is responsible for the conversion of phenylalanine (L-Phe) to tyrosine, its mutations lead to the toxic accumulation of L-Phe, which is the cause of PKU.

In order to rescue a mutant protein and restore its activity, two main approaches are usually adopted, the development of pharmacological or activity chaperones. While pharmacological chaperones (PCs) have the function of binding the misfolded protein and helping it to properly fold by stabilizing the native conformation, activity chaperones (ACs) are aimed at increasing the enzymatic activity of the mutated protein.

In this framework, 3HQs were investigated as potential multivalent modulators, that is, molecules able to combine both PC and AC activity by stabilizing the active conformation and increasing the catalytic activity of PAH.

With this objective, three series of 3HQ derivatives were synthesized and tested in various experiments in order to extrapolate structure-activity relationship between the 3HQ structure and the effect on PAH.

A grateful acknowledgement to all collaborators involved in this chapter:

Roberta Paterna (development of the 3HQ synthesis)

Catarina Silverio (3HQ synthesis)

Prof. Ana Paula Leandro (Studies on hPAH)

Raquel R. Lopes, João Leandro (hPAH production and purification)

Catarina S. Tomé, Miguel Teixeira, Pedro Sousa, João B. Vicente (Assays on hPAH)

Rita C. Guedes (Docking studies)

IV.1 Rationale and Goals

3HQs exhibit a pseudopeptidic structure that is able to coordinate with enzyme metal cofactors. Because of this, this scaffold has the potential to be employed to design molecules that can interact with metalloproteins and elicit a therapeuthical effect.

In this part of the work, 3HQs were tested as modulators of the Human Phenylalanine Hydroxylase (hPAH) enzyme, whose misregulation is the cause of phenylketonuria (PKU), an inherited metabolic disorder.

Inherited metabolic disorders (IMDs), in general, arise from the mutation of genes that code for enzymes or transporters involved in the intermediary metabolism, leading to loss or alteration of activity of these proteins. More than 50% of IMDs are caused by missense mutations that might lead to an impaired catalytic activity and/or incorrect folding of the variant proteins. Incorrect folding of the protein of interest does not only cause its loss of activity, but also leads to premature degradation, and a decrease in the protein intracellular concentration.

In the context of pharmacological treatment of the disease, one of the most common approaches relies on the design of small molecular inhibitors of the enzymatic target that causes the pathology. This kind of strategy is employed for the treatment of many conditions and it is based on the optimization of a molecular structure in order to obtain a compound with high affinity for the target protein, which blocks its activiy upon binding.

However, when the disease is caused by a structural aberration of a protein, the approach is necessarily different. In fact, as the objective in these conditions is to restore the native properties of a mutated protein, small molecular chaperons are needed as treatments. These molecular chaperones are usually devised to either restore the native folding of the protein (pharmacological chaperones) or to enhance and protect its activity (activity chaperones).[133,134]

The development of this kind of chaperones is important in regards to phenylalanine hydroxylase, a fundamental enzyme in amino acid metabolism that is prone to loss of activity and misfolding due to the high incidence of mutations associated to its gene. In fact, as of 2019, more than 1000 mutations have been identified in the PAH gene, among which around 62% are missense mutations.[135] This enzyme catalyses the conversion of L-Phenylalanine (L-Phe) to L-Tyrosine (L-Tyr), regulating the concentration of L-Phe in the blood. Deficient activity of the phenylalanine hydroxylase enzyme is the cause of

Phenylketonuria, the most common inherited metabolic disorder of amino acid metabolism, affecting one in 10.000 new-borns.[136] This condition was first reported in 1934 by Dr. Asbjørn Følling, a Norwegian physician that detected phenylpiruvic acid in urine samples of two brothers suffering from physical impairments. Dr. Følling postulated, correctly, that the accumulation of phenylpiruvic acid in the urines was caused by an inherited malfunction of the phenylalanine metabolism.[137]

When untreated, PKU can lead to brain damage, progressive psychomotorial impairments, epilepsy, neurological and behavioural problems, arising from the toxic effect of L-Phe accumulation in the blood, called hyperphenylalaninemia.[138]

The severity of PAH deficiency is classified based on the blood concentration of L-Phe, as the baseline levels of this amino acid is 120 μmol/L, values above this threshold and up to 600 μmol/L are indicative of mild hyperphenylalaninemia. Between 600 and 1200 μmol/L of L-Phe the condition is identified as mild PKU, while, above the 1200 μmol/L value it is indicated as classical PKU.[139] Currently, the only therapeutic approach approved for PKU consists in a lifelong dietary restriction of L-Phe, hence, patients are forced into a low protein diet, which often leads to poor patient quality of life, malnutrition and psychosocial complications.

Human PAH (hPAH) is a non-heme iron-dependent enzyme that hydroxylates L-Phe into L-Tyr in the presence of molecular oxygen and the (6R)- L-erythro-5,6,7,8-tetrahydrobiopterin (BH_4) cofactor (**Figure 27A**). This enzyme is constituted as a dimer of dimers, in which each monomer is organized in three domains: The *N*-terminal regulatory domain, the catalytic domain and the *C*-terminal domain, which promotes the assembly of the monomers. The catalytic site contains the iron centre, coordinated by a triad of amino acids (His285, His290 and Glu330) and three water molecules. The hydrophobic and negatively-charged pocket that harbours the iron centre is also where the natural substrate, L-Phe, and the cofactor BH_4 bind during the hydroxylation reaction. As this enzyme is responsible for the homeostasis of the L-Phe concentration in the blood, it is physiologically pre-activated by allosteric binding of this substrate. In fact, L-Phe is able to bind to the interface established by dimerization of the regulatory domains of two monomers, stabilizing the “activated” form of the homotetramer.[140–142] The transition to the activated state consists in the structural rotation of the N-terminal regulatory domain, exposing the active site and facilitating the entrance of L-Phe in the catalytic centre. When the enzyme is not activated by L-Phe, the regulatory domain blocks the entrance to the active site, lowering the affinity of PAH for its substrate. In contrast to what observed for

L-Phe, the binding of the BH_4 cofactor results in the stabilization of the resting state of the protein, lowering its activity. Interestingly, the bound conformation adopted by PAH in the presence of BH_4 presents an increased resistance to unfolding and degradation, leading to a \potential mechanism for the stabilization of misfolded hPAH variants[143,144]. However, the use of BH_4 as pharmacological chaperone did not yield satisfactory results in patients, even employing high dosages (20 mg/Kg body weight). However, even though not all phenotypes of PAH mutation responded to it, the mode of action of BH_4 remains an important starting point for the design of PCs for the stabilization of PAH.[145,146]

Interestingly, the iron centre localized in the catalytic site of hPAH has never been explored to design hPAH modulators but constitutes a valuable binding site for both ACs and PCs. As reported in the previous chapters, 3-Hydroxy Quinolinones are efficient metal chelators, which, in this context, could be employed for the coordination of the iron centre of hPAH. Moreover, this bicylic scaffold shares some structural features with BH_4 and can be easily functionalized to insert a phenylalanine amide motif on position 4 of the scaffold. Therefore, it was envisioned that, by producing 3HQ derivatives bearing a phenylalanine substituent on position 4, it could be possible to obtain a molecule that combine the stabilizing action of BH_4, along with the activating effect of L-Phe on a scaffold that can also interact with the iron core of the active site. (**Figure 27B**).

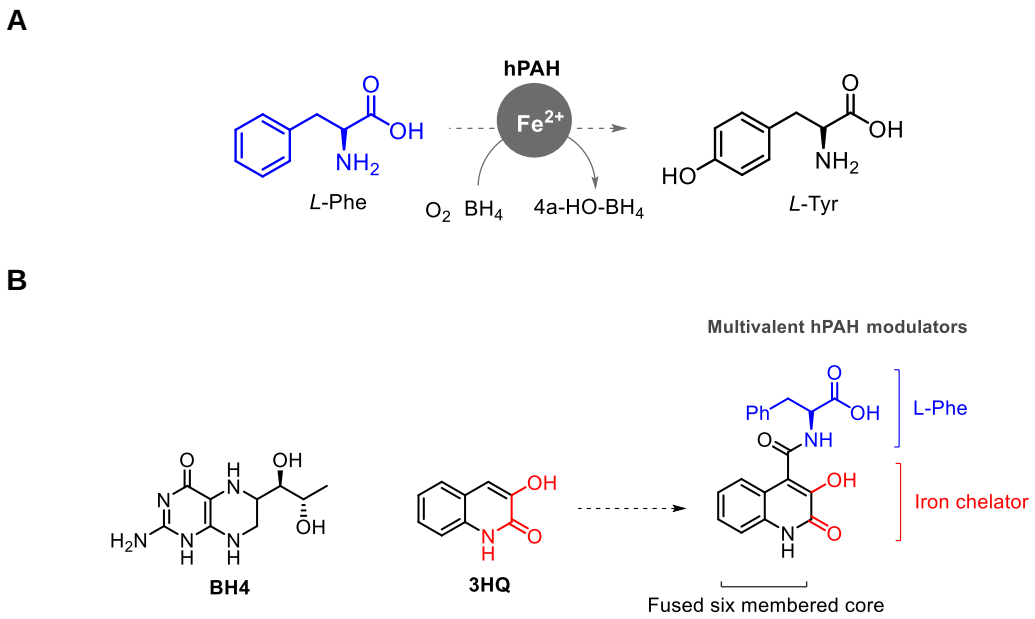
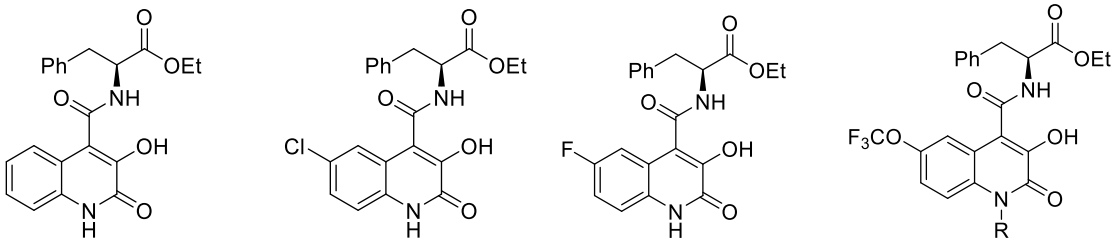


Figure 27 – Catalytic hydroxylation of L-Phe by hPAH (A). Structure of the BH_4 cofactor, of the basic 3HQ core and of the envisioned 3HQ-based multivalent hPAH modulators.

IV.2 Results and discussion

Taking advantage of the synthetic methodologies already described in the previous chapters, a small library of 3HQ derivatives has been synthesized to be tested as both pharmacological and activity chaperones. The first series that was synthesized and tested, Series 1, is comprised of 3HQ derivatives bearing the L-Phe motif in position 4 of the scaffold, in order to mimic the natural substrate of hPAH (**Scheme 33**). The first four compounds of Series 1 differ from each other based on the nature of the substituent on position 7 of the heterocycle. As the electronic character of the bicyclic system is influenced by the properties of this substituent, electron-withdrawing substituents were chosen to occupy position 7 of the scaffold in order to increase its tendency towards iron chelation. Additionally, Series 1 also contained an *N*-benzylated analogue of compound **76,** 3HQ **77**, with the purpose of exploring the chemical space relative to the substitution on the amide present in position 1 of the 3HQ core.

Series 1



74 **75** **33** **76** R=H **77** R=Bn

Scheme 33 – First series of 3HQ derivatives tested as hPAH modulators

IV.2.1 Effect of 3HQ derivatives on hPAH activity and thermal stability and evolution of the 3HQ library

The delicate role of chaperones predicates that this kind of molecules interact with the target protein, stabilizing its structure without interfering with its catalytic activity. Hence, it is fundamental to strike a good balance regarding the binding affinity of these compounds towards hPAH. The first test to be performed, in this context, was to measure the catalytic activity of isolated hPAH in the presence of the various 3HQ derivatives. However, as this enzyme has the ability to be activated by its substrate, L-Phe, a correct evaluation of the

effect of each 3HQ derivative on the activity of hPAH must take this phenomenon into account. In order to do so, each compound was tested in three conditions: non-activated, compound-activated or compound/substrate activated (**Figure 28**).

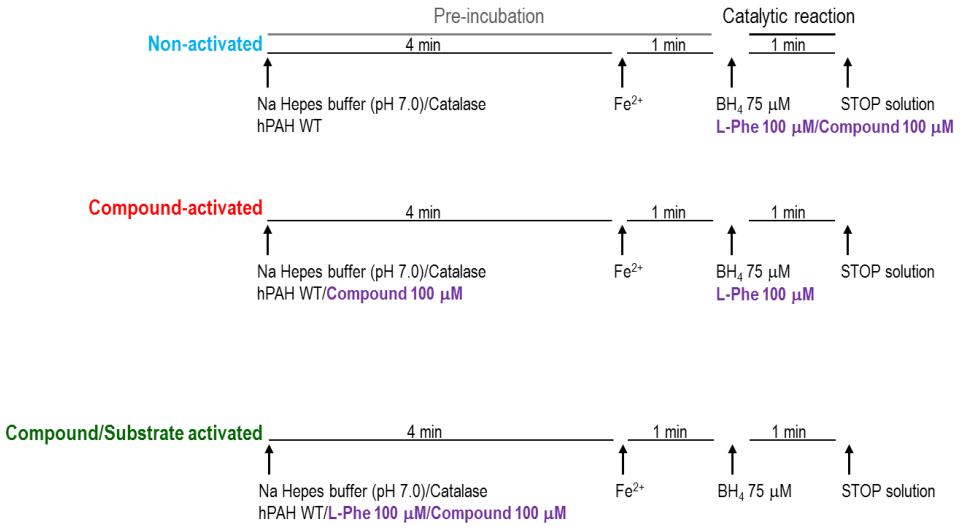


Figure 28 – Protocols for the three assays performed to evaluate the effect of 3HQ derivatives on the activity of hPAH

In the assay performed in non-activated conditions, hPAH was incubated in buffer for 4 minutes, then the iron cofactor was added and, after a minute, BH_4 and L-Phe were added, along with the 3HQ in exam. This experiment excluded any possible pre-activation effect that the presence of either the 3HQ or of L-Phe might have on the catalytic activity. The second assay, compound-activated, was designed to study the effect that pre-complexation with each 3HQ derivative had regarding the activity of the enzyme. In this case, hPAH was incubated for four minutes with the 3HQ of choice, then the iron cofactor was added, followed, one minute later, by the addition of BH_4 and L-Phe. Finally, the third protocol, compound/substrate activated, was performed by incubating both L-Phe and the 3HQ with hPAH before adding Fe and BH_4 with the same sequencing utilized in the other assays. This triad of assays was designed in order to better evaluate the contribution that each 3HQ could give to the structure of hPAH upon interaction, control assays were also performed using 1% of DMSO as it was used as vehicle for each compound.

In control experiments, recombinant hPAH presented an enzymatic activity of 1,504 ± 66 nmol $Tyr.min^{-1}.mg^{-1}$. As expected, pre-incubation with DMSO did not result in enzyme activation (1,620 ± 67 nmol $Tyr.min^{-1}.mg^{-1}$; $P = 0.1$), while incubation with L-Phe 100 µM

resulted in a 1.7-fold activation (2,590 ± 130 nmol Tyr.min^{-1}.mg^{-1}), confirming the activating effect that the substrate has on hPAH.

As observed in **Figure 29**, only compound **76** was able to increase the activity of hPAH in non-activated conditions (1.1-fold). On the other hand, compounds **74**, **75** and **33** inhibited hPAH activity, resulting in, respectively 48, 25 and 56% of activity compared to the control. Moreover, the compound-activated assay for compound **76** showed that this compound had the lowest impact on the activity of the protein in the compound-activated (69% activity) and compound/substrate-activated conditions (57%) when compared to the control.

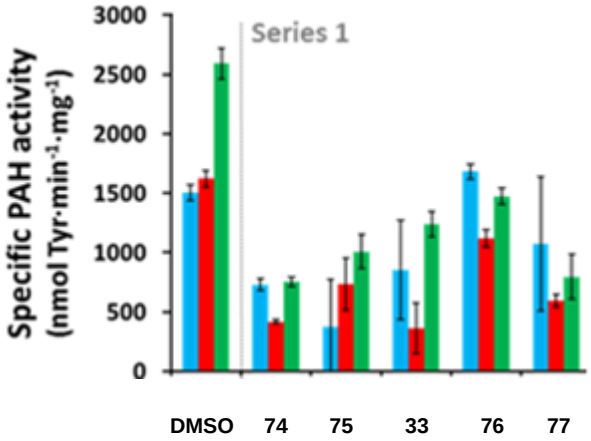


Figure 29 – Activity of hPAH in the presence of 3HQ derivatives from series 1 in three different conditions: Non-activated (blue); Compound-activated (red) and Compound/Substrate-activated (green).

Importantly, when comparing the results from the compound-activated and compound/substrate-activated assays, it became evident that these compounds do not inhibit the activation of hPAH by its substrate. In fact, as shown in **Figure 29,** the enzymatic activity of hPAH measured in the presence of the compounds from this series (red bars) always increased when L-Phe was also present in the assay (green bars). The activity improvement in the presence of L-Phe indicates that all of the compounds from series 1 allow the activation of hPAH in the presence of its substrate, which is a fundamental aspect of the biological role of this enzyme.

As the enzymatic activity assays gave insight about the potential of the various 3HQ derivatives as activity chaperones, the next step was to assess their ability to stabilize the structure of hPAH, acting as pharmacological chaperones as well.

In order to do so, the thermostability of hPAH was measured after incubation with each 3HQ derivative of series 1. This assay was performed by slowly ramping the temperature of the protein sample and registering the temperature at which the denaturation of each subunit occurs. In the case of hPAH, two points of denaturation can be observed: 43.4 ± 0.6 °C, corresponding to the denaturation of the regulatory domain (T_{m1}) and 53.5 ± 0.4 °C, relative to the denaturation of the catalytic domain (T_{m2}). Increase in the temperature of denaturation of each subunit are observed when they interact with a ligand that stabilizes them, as in the case of L-Phe. In fact, as this substrate binds to both subunits, in the presence of 1 mM of L-Phe the values of T_{m1} and T_{m2} raise to 51.0 ± 0.2 °C and 58.4 ± 0.04 °C, respectively.

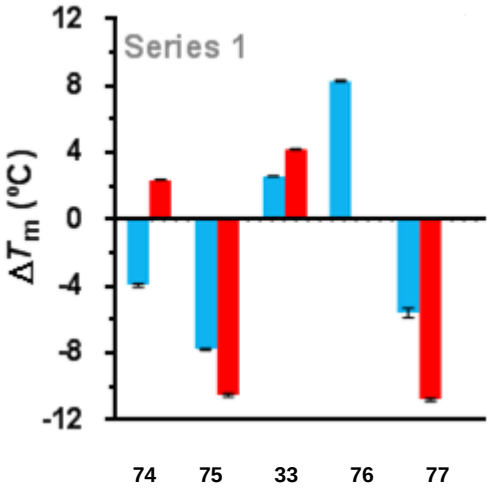


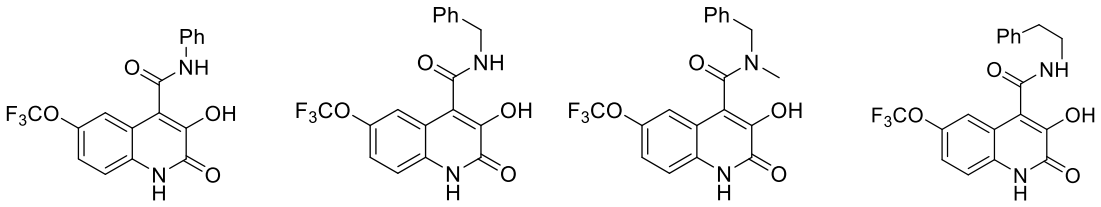
Figure 30 - Change in the thermal stability of hPAH in the presence of 3HQs from series 1. The mid-point of thermal denaturation (Tm) is indicated in blue for the regulatory domain (Tm_1) and in red for the catalytic one (Tm_2)

Testing the first series of 3HQs for their ability to enhance hPAH thermostability, different effects were observed, indicating different potential points of interaction for each compound (**Figure 30**). Compound **74** had a destabilizing effect on the regulatory domain (ΔT_{m1} = -3.9 ± 0.1 °C), while stabilizing the catalytic domain (ΔT_{m2} = 2.4 ± 0.1 °C), suggesting a possible interaction with this region of the protein. On the other hand, incubation with 3HQ **75** had a negative effect on the stability of both domains (ΔT_{m1} = -7.7 °C and ΔT_{m2} = -10.5 °C). Both compounds **35** and **76** produced an increase in the stability of the regulatory domain (ΔT_{m1} = +2.6 °C and ΔT_{m1} = +8.3 °C, respectively) when incubated with hPAH. Moreover, compound **35** also stabilized the catalytic domain (ΔT_{m2} = +4.2 °C). Surprisingly, despite showing similar behaviour to **76** in the activity assays, its benzylated

derivative, compound **77** had a negative impact on the thermal stability of both domains. Overall, compound **76** was the best-performing one among the 3HQs of series 1, given its lower impact on the catalytic activity of hPAH, while not interfering with the L-Phe activation mechanism. Moreover, this compound showed a positive effect on the stabilization of the regulatory domain of hPAH, hence it was chosen as the starting point for the synthesis of a second series of 3HQ derivatives.

The second series of 3HQ derivatives was designed starting from the 7-OCF_3-3HQ scaffold of compound **76**, adding different carboxamide derivatives in position 4 of the heterocycle. In order to gain insight on the effects of the position and flexibility of the aromatic ring of the L-Phe derivatives of series 1, compounds of series 2 were synthesized with aromatic substituents of different size and branching on the 4-carboxamide position (**Scheme 34**).

Series 2



78 **79** **80** **81**

Scheme 34 – Series 2 of 3HQ derivatives, based on the –OCF_3 3HQ scaffold

Compounds **78-81** were tested in the same assays performed for the compounds of series 1, with the objective of elucidating the effect of different aromatic substituents. When tested for their effect on enzymatic activity, compounds **78-81** showed to inhibit hPAH in all conditions (**Figure 31**).

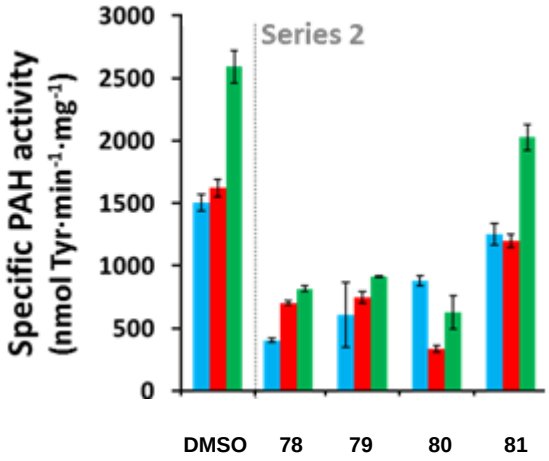


Figure 31 - Activity of hPAH in the presence of 3HQ derivatives from series 2 in three different conditions: Non-activated (blue); Compound-activated (red) and Compound/Substrate-activated (green).

On the other hand, the presence of compound **81** caused very little inhibition of the catalytic activity and allowed for a 1.7-fold activation of the enzyme in the presence of L-Phe. This result represented an improvement over the performance of compound **76**, which was the best candidate from series 1 and can be rationalized with the fact that the phenylethanamide moiety of compound **81** is a structural analogue of the side chain of L-Phe. As shown in **Figure 32**, when compounds of series 2 were submitted to thermostability assays, they did not yield any stabilization effect, as compounds **78**-**81** did not elicit any significant change of stability on either domain ($|\Delta T_m| \leq 1.3$ °C). Compound **78** fared the worst of series 2 in this assay, as it strongly decreased the thermal stability of both the regulatory and the catalytic domain ($\Delta T_{m1} = -4.7$ °C and $\Delta T_{m2} = -11.2$ °C).

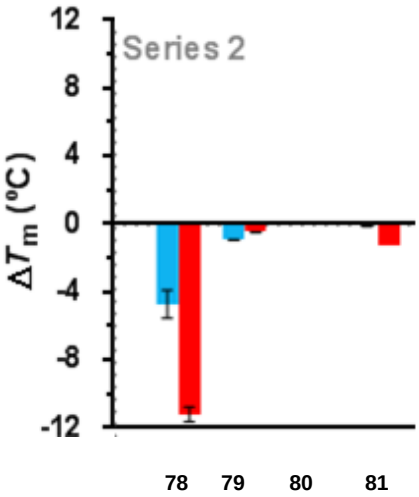
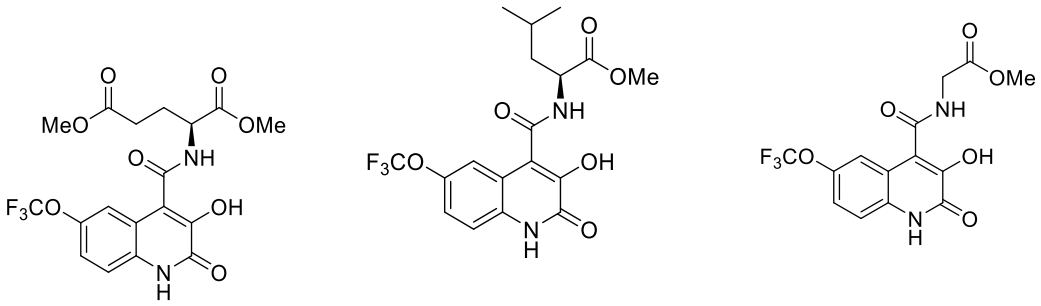


Figure 32 - Change in the thermal stability of hPAH in the presence of 3HQs from series 2. The mid-point of thermal denaturation (Tm) is indicated in blue for the regulatory domain (Tm_1) and in red for the catalytic one (Tm_2)

Among the compounds of series 2, compound **81** was the one that yielded the most positive results in these preliminary tests, as it showed the best effects on the enzymatic activity so far. However, while it performed similarly to its structural analogue, compound **76**, regarding the effect on enzymatic activity, it did not yield any stabilizing effect on hPAH. As the only difference between these two compounds is the lack of the peptidic backbone in compound **81**, the contribution of an amino acid substituent on position 4 of the 3HQ scaffolds was the next factor to examine. To this end, a third series of 3HQ derivatives was synthesized, comprised of derivatives bearing different non-aromatic amino acid substituents on the amide handle (**Scheme 35**).

Series 3



82 **83** **84**

Scheme 35 - Series 3 of 3HQ derivatives, bearing different amino acid substituents

As shown in **Figure 33**, this series interacted favourably with hPAH, as the enzymatic activity was maintained between 40% and 60% for each sample in the non-activated conditions. Moreover, all compounds from series 3 allowed for the activation of hPAH by L-Phe, as they afforded significant activation in compound/substrate activated conditions. Interestingly, compound **84** also had the effect of activating the enzyme *per se*, as in compound-activated conditions it afforded a 1.6-fold increase in enzymatic activity.

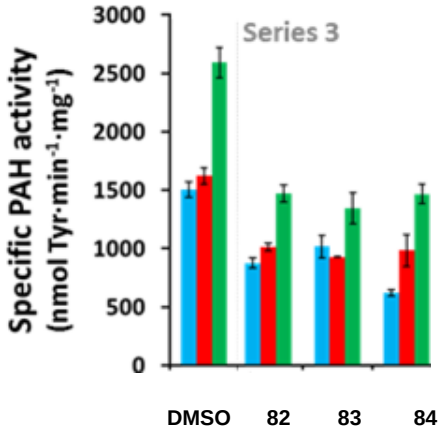


Figure 33 - Activity of hPAH in the presence of 3HQ derivatives from series 3 in three different conditions: Non-activated (blue); Compound-activated (red) and Compound/Substrate-activated (green).

Regarding the effect of the third series of 3HQ derivatives on hPAH thermal stability, only compound **84** elicited an effect, stabilizing the regulatory domain (ΔTm1 = 4.0 ± 0.1 ºC), while the other two derivatives had no significant effect on either domain (**Figure 34**)

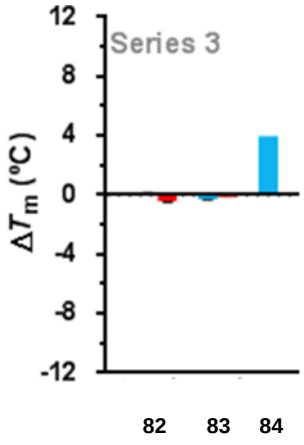


Figure 34 - Change in the thermal stability of hPAH in the presence of 3HQs from series 3. The mid-point of thermal denaturation (Tm) is indicated in blue for the regulatory domain (Tm_1) and in red for the catalytic one (Tm_2)

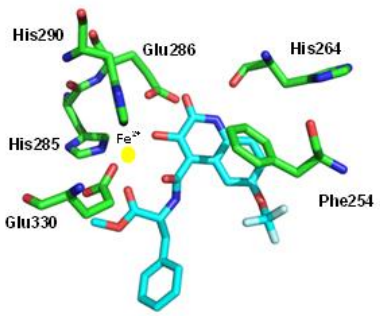
The preliminary screening performed on the various 3HQ derivatives synthesized in this study indicated that, given their effect on hPAH activity and/or stabilization, compounds **76**, **81**, **82** and **84** represented the most promising candidates for docking studies. These compounds were chosen based on the consideration that they did not inhibit the enzymatic activity upon incubation with hPAH, they allowed further activation by L-Phe and had a stabilizing effect on the protein.

IV.2.2 Molecular docking of 76, 81, 82 and 84 to hPAH

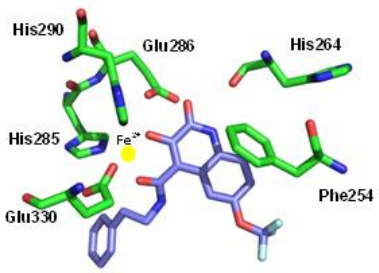
As compounds **76**, **81**, **82** and **84** showed the most interesting properties as pharmacological and/or activity chaperones, they were submitted to molecular docking simulations in order to gain insight about their interaction with hPAH. This study was performed by docking the four selected 3HQs in the active site of hPAH modelled on the crystallographic structure available for this enzyme (PDB ID: 3PAH). In the crystal structure adopted, the protein shows the Fe(II) catalytic centre to which adrenaline, a reversible inhibitor, is bound through its two hydroxyl groups.[147]

The first computational analysis performed on the selected compounds was to dock them inside the catalytic domain, in order to gain insight on their preferred conformation as well as their potential interaction with the catalytic iron centre. The poses obtained in this way are shown in **Figure 35**, in which the Fe(II) centre is represented in yellow, surrounded by the three residues that directly interact with it (His285, His290 and Glu330). Besides these three residues, the docking images also show the other amino acids present in the active site that establish interactions with the 3HQ compounds in exam.

A)



B)



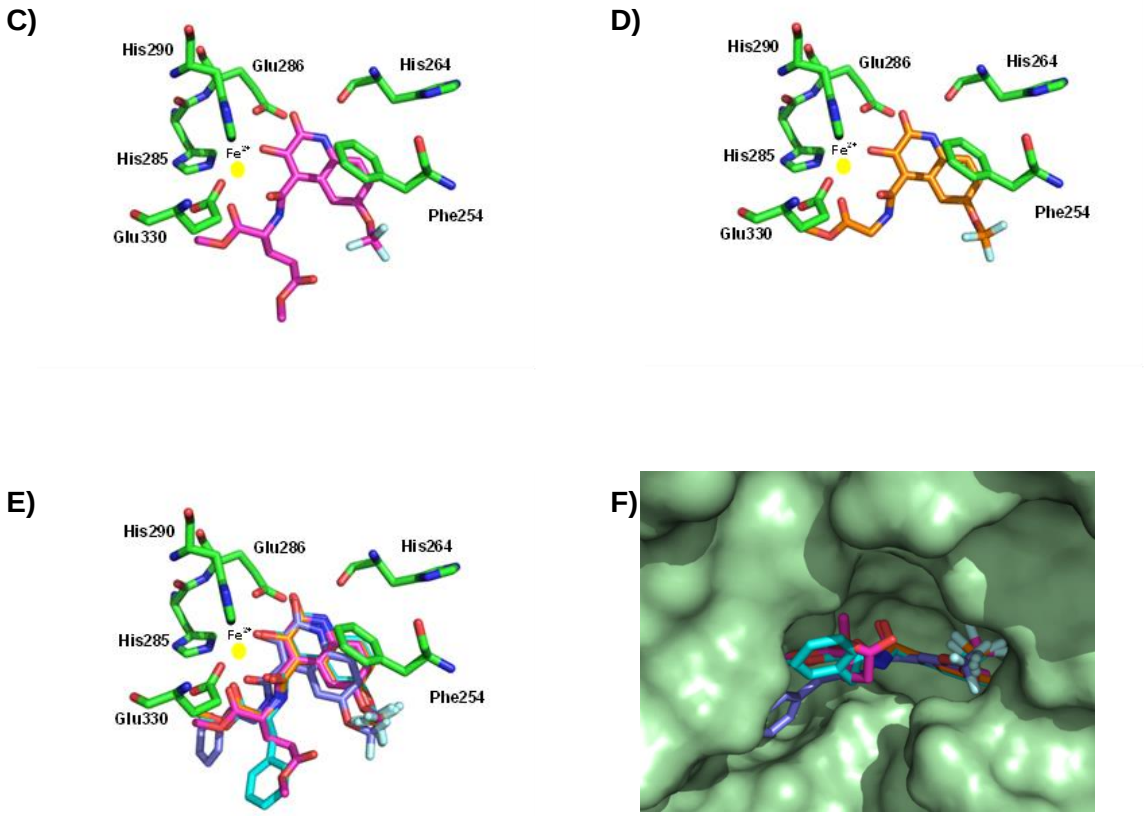


Figure 35 - Best docking poses of selected 3HQ derivatives into the active site of hPAH: structures for compound 76 (A), 81 (B), 82 (C) and 84 (D), along with superimposition of the 4 compounds in the active site (E) and on the hPAH surface (F).

As shown in **Figure 35**, the four selected 3HQ derivatives adopt similar conformations in the hPAH active site, with the hydroxyl group of the quinoline ring oriented towards the active site iron at a distance of 2.24, 1.98 and 2.15 Å for compound **76**, **81** and **82**, respectively. On the other hand, in the case of compound **84** only the amide side chain on position 4 of the heterocycle was found close to the iron centre, at distances of 1.73 and 2.25 Å. Interestingly, while the bicyclic cores of these 3HQs resulted to be overlapped in the obtained poses, compound **81** showed a different orientation regarding the amide substituent, probably due to the lack of an amino acid motif in this position, which is present in the other tested compounds.

The poses observed for compounds **76**, **81** and **82** are structurally similar to what observed in the case of incubation of hPAH with catecholamine derivatives (dopamine, adrenaline and noradrenaline), which are known to be reversible inhibitors for this enzyme. In fact,

also the compounds belonging to this class coordinate the Fe(II) centre through their OH groups. Additionally, catecholamine derivatives also establish hydrogen bonds with the side chains of Tyr325 and Glu330, this is an important interaction, which is fundamental for the recognition of such molecules and is at the base of their strong inhibitory action .[147] Interestingly, among the tested compounds, only compound **82** showed the ability to establish similar interactions, as its phenolic OH group was found at 2.66 Å of the side chain of Glu330.

Additionally, the poses obtained for compounds **76**, **81**, **82** and **84** were compared to the positioning of the BH_4 cofactor inside the catalytic pocket, with the objective of determining if the structural similarity between the 3HQ heterocycle and BH_4 leads to similar binding poses. In this second docking study, the protein structure was derived by the crystal structure of the ternary complex of hPAH with BH_4 and norleucine as substrate analogue (PDB ID: 1MMT). As shown in **Figure 36**, the overlap of the structures of the 3HQ derivatives in exam matches the positioning of BH_4, confirming the hypothesis that the 3HQ scaffold could act as a mimic of the BH_4 cofactor.

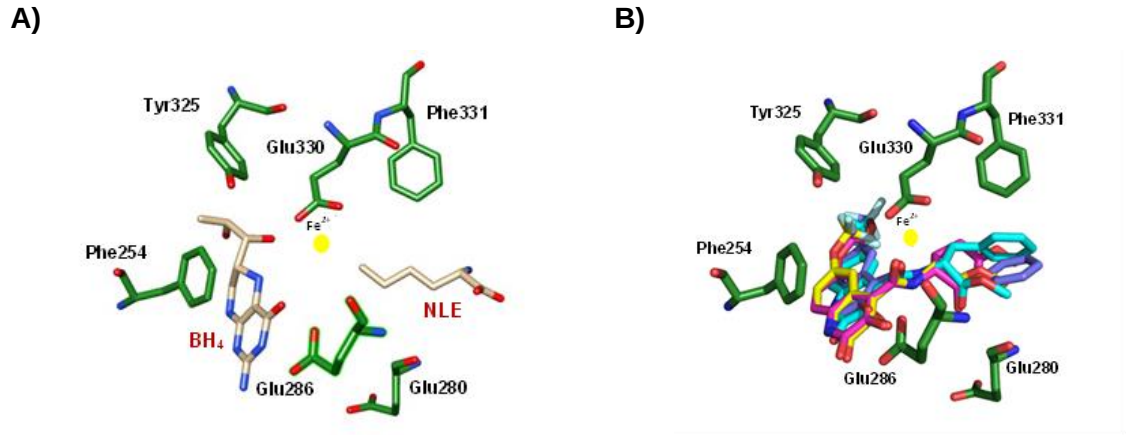


Figure 36 – Crystal structure of the catalytic site of hPAH with BH4 cofactor and norleucine (NLE) substrate analogue (PDB ID: 1MMT) (A) and superimposition of the best docking poses of 3HQ derivatives 76, 81, 82 and 84 in the catalytic centre of hPAH (PDB ID: 1MMT).

As shown in **Figure 36A**, the BH_4 cofactor mainly interacts with Phe254, Leu249, Ser251 and Gly247 inside the catalytic pocket of hPAH. Analogously, the most favourable placements of the 3HQ derivatives in exam indicated the presence of π-stacking interactions with Phe254 and the formation of a hydrogen bond with Leu249, but not with Ser251 and Gly247.

While the poses of the 3HQ derivatives resulted to be structurally similar to the orientation of BH_4 inside the catalytic site of hPAH, the same was not observed regarding the pose of

the substrate analogue. In fact, the pose of NLE in the 1MMT crystal structure showed that the fundamental residues for substrate recognition and binding are His285, Ser349, Arg270 and Thr278. Among these residues, only His285 was found to be able to form interaction with the selected 3HQs, as the distance between this residue and the various compounds docked was found to be around 2.5 Å. This observation is correlated to the fact that 3HQs have a significantly lower affinity for the active site of hPAH compared to the natural substrate, L-Phe, which is a positive factor as it indicates that the presence of these compounds does not impede the desired enzymatic activity.

The estimated binding energies regarding the interaction between the 3HQs and both crystal structures of the active site of hPAH (3PAH and 1MMT) calculated with Goldscore fitness function showed that compound **84** presented the lowest affinity values (scores 69.72 and 64.63), while compound **81** showed the best affinity (scores 78.13 and 78.51).

IV.3 Conclusions

Among the 12 3HQ derivatives synthesized, compounds **76**, **81**, **82** and **84** were the ones that gave the most promising results as pharmacological and activity chaperones. This kind of biological activity is a delicate matter, as it requires the chaperone molecule to interact with the target protein without blocking its function. In this regard, the affinity of the molecules is a fundamental factor, as it needs to be sufficient to efficiently bind to the target protein, while not interacting with it so strongly that the natural substrate is no longer allowed inside the protein. In this context, the four selected 3HQs showed to have a good affinity profile towards hPAH, as compound **76** did not show any inhibitory effect, compound **81** yielded a 17% decrease in activity and both compounds **82** and **84** caused a ≈50% activity loss. As described in the first paragraph of the chapter, the activity of hPAH can be enhanced in the presence of its natural substrate, L-Phe, following an activation mechanism that allows this enzyme to guarantee the homeostasis of L-Phe by increasing its activity in response to a surge in the concentration of its substrate. This mechanism is fundamental in the activity of hPAH and, as such, its preservation is crucial. Gratifyingly, pre-incubation of hPAH with each of the 4 selected 3HQ derivatives allowed further activation of the enzyme by L-Phe, indicating that the presence of the compounds did not compromise the conformational changes that are needed for the substrate-promoted activation mechanism of hPAH. Moreover, the selected compounds elicited the highest

response to L-Phe activation, namely 78% for compound **81**, while compounds **76**, **82** and **84** yielded ≈60% activation.

The activity data indicate that the binding of these 3HQs to hPAH falls in the optimal affinity range for their use as chaperone molecules, as their interaction is efficient enough to cause a stabilizing effect on the structure, as demonstrated by the thermal stability assays, while not inhibiting L-Phe complexation.

This behaviour can be rationalized by comparing the data emerged from the docking studies with what is known about the binding mode of potent inhibitors, such as catecholamines. As shown in the previous paragraph, the four selected 3HQ derivatives were found to bind the Fe(II) centre in the catalytic pocket of hPAH through the interaction of the OH group of the 3HQ motif, adopting a conformation that is analogous to the one found in the crystal structures obtained by the complexation of hPAH with adrenaline. Crucially, while adrenaline and other catecholamine derivatives are able to chelate the Fe(II) atom with two hydroxyl groups, the 3HQ derivatives only establish one interaction with the iron centre, namely through the OH in position 3 of the scaffold, which leads to a looser binding of 3HQs to hPAH.

In conclusion, some of the 3HQ derivatives presented in this chapter showed interesting properties as multivalent modulators for hPAH. In fact, they showed good results regarding the stabilization of the protein folding, while not blocking the enzymatic activity and preserving the substrate-triggered activation of the protein.

Chapter V

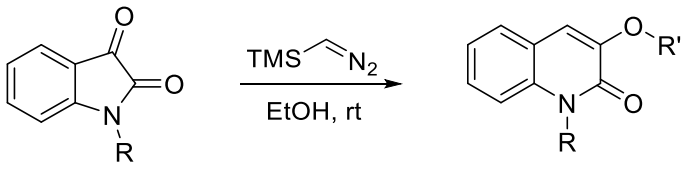
Supporting Information

V.1. Experimental section of Chapter II

V.1.2 Synthesis and structural characterization

Dry reaction solvents, as dichloromethane and acetonitrile were dried over calcium hydride under nitrogen or argon atmosphere. Tetrahydrofuran was dried in a mixture of metallic sodium and benzophenone under nitrogen or argon atmosphere. DMSO and dimethylformamide (DMF) were purchased from Acros (Extra Dry over Molecular Sieve, AcroSeal®). Other reaction solvents were used without any purification. Compounds indicated as commercially available were purchased from habitual commercial sources and used without further purification. Thin layer chromatography was performed using Merck silica gel 60F254 aluminium plates and visualized using UV light and in a phosphomolybdic acid solution. In column chromatography it was silica gel 60 M purchased from MN (Ref. 815381). NMR spectra were recorded in a Bruker AMX 400 and in a Bruker Fourier 300 using $CDCl_3$ and $(CD_3)2SO$ as deuterated solvents. All coupling constants are expressed in Hz and chemical shifts (δ) in ppm. Multiplicities are given as: s (singlet), d (doublet), dd (double doublet), t (triplet), q (quartet), and m (multiplet). Fourier transform infrared spectroscopy (FTIR) was done in Perkin Spectrum Two. UV spectra were traced in Thermo Scientific Evolution 201 UV-visible spectrophotometer. Low resolution mass spectrometry (LRMS) was performed in a mass spectrometer (Micromass Quattro Micro API, Waters, Ireland) with a Triple Quadrupole (TQ) and with an electrospray ion source (ESI) operating in positive mode. It was also used an ion trap mass analyser (Thermo Scientific LCQ Fleet Ion Trap LC/MS) equipped with an electrospray interface. High resolution mass spectrometry (HRMS) was performed in a mass spectrometer (LTQ Orbitrap XL) using an ESI source operating in positive mode. Elemental analysis (E.A) was made in a Flash 2000 CHNS-O analyser (ThermoScientific, UK).

- **Synthesis of simple 3-hydroxy-quinolinones 1, 6, 7**



1: R=H, R'=H

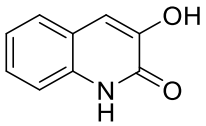
4: R=H, R'=CH_3

5: R=CH_3, R'=H

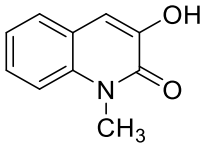
Scheme 36 – Synthetic scheme for the preparation of simple 3HQ cores

The basic HQ cores were prepared according to literature procedures starting from commercially available Isatin and N-Methyl Isatin.[116]

General procedure for compounds 1, **5:** In a flame-dried, round-bottomed flask, the proper isatin derivative (0.999 mmol), (diazomethyl)trimethylsilane 2M solution in diethyl ether (0.500 mL, 0.999 mmol) and triethylamine (0.279 mL, 1.998 mmol) were dissolved in absolute ethanol (5mL) and placed under nitrogen atmosphere, stirring at room temperature for 15 h. The resulting suspension was then filtered, the solid collected, washed with cold ethanol and dried in vacuum. No further purification steps were required as the product was recovered as a pure solid after filtration.

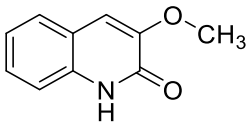


3HQ 1. yellow solid, 81% yield, ^{1}H-NMR spectrum matches what described in the literature.[116] **^{1}H NMR (300 MHz, ($CDCl_3$):** 12.01 (br s, 1H), 9.46 (br s, 1H), 7.49 (d, J = 7.9 Hz, 1H), 7.27 (m, 2H), 7.12 (ddd,J = 7.9, 6.8, 1.7 Hz, 1H), 7.08 (s, 1H).



3HQ 5. grey solid, 75% yield, ^{1}H-NMR spectrum matches what described in the literature.[148] **^{1}H NMR (300 MHz, $CDCl_3$):** *δ* 7.52 (dd, J=7.8, 1.5 Hz, 1H), 7.46 (m, 1H), 7.36 (d,J=8.3 Hz, 1H), 7.27 (td,J=8.4, 1.2 Hz, 1H), 7.13 (bs, 1H), 7.12 (s, 1H),

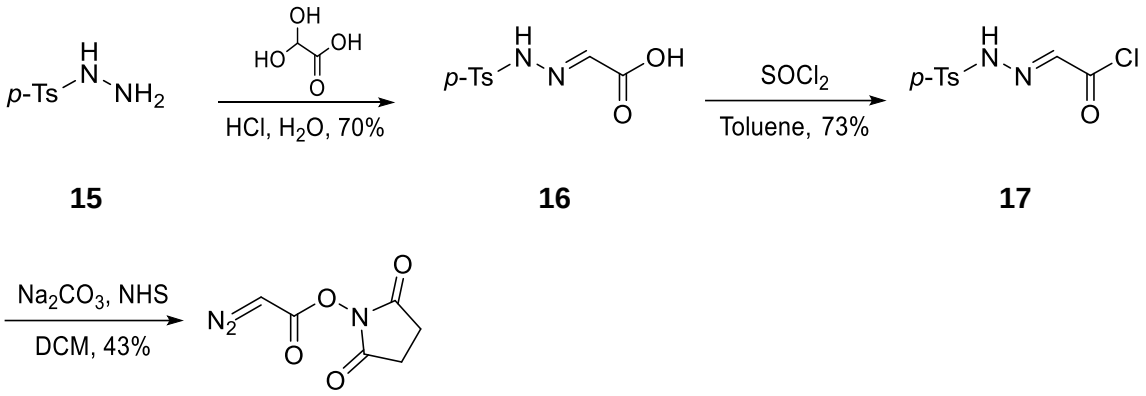
3.83 (s, 3H)



3HQ 4. In a flame-dried, round-bottomed flask, isatin (147 mg, 0.999 mmol), (diazomethyl)trimethylsilane 2M solution in diethyl ether (1 mL, 1.998 mmol) and triethyl amine (0.279 mL, 1.998 mmol) were dissolved in ethanol (5mL) and placed under nitrogen atmosphere, stirring at room temperature for 15 h. The resulting suspension was then filtered, the solid collected, washed with cold ethanol and dried in vacuum. No further purification steps were required as the product was recovered as a pure yellow solid after filtration with 78% yield.

The NMR spectrum matches what is reported in the literature.[149] **^{1}H NMR (300 MHz, $CDCl_3$):** δ 10.92 (bs, 1 H), 7.50 (d, J = 7.32 Hz, 1 H), 7.38 (m, 2 H), 7.22 (t, J = 6.83 Hz, 1 H), 7.00 (s, 1 H), 3.98 (s, 3 H).

- **Synthesis of diazo-NHS intermediate 18**

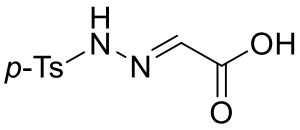


15 **16** **17**

18

Scheme 37 – Synthetic route towards diazo-NHS intermediate 18

Diazo succinimide **18** was synthesized following literature procedures, starting from commercially available Glyoxylic acid and *p*-toluenesulfonylhydrazine.



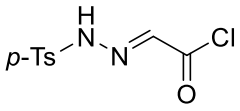
(E)-2-(2-tosylhydrazono)acetic acid 16. This reaction has been performed according to literature procedures.[150]

2,2-dihydroxyacetic acid (0.989 g, 10.74 mmol) was dissolved in water (9.9 ml) at 65 °C under magnetic stirring; to this solution, a suspension of *p*-toluenesulfonylhydrazine (2 g, 10.74 mmol) in 2.5 M aqueous HCl (6.01 ml, 15.04 mmol), previously heated to 65 °C, was added at once. The resulting suspension was left under magnetic stirring for 15 minutes at 65 °C, after which it was left to cool down at room temperature and then refrigerated overnight. The resulting crude was filtered and washed with cold water, then left to air dry for 2 days, followed by an overnight period under high vacuum.

The resulting white solid crude was dissolved in ethyl acetate and precipitated with hexane, then filtered and washed with a cold mixture of Hexane: Ethyl Acetate 2:1. the resulting white solid was dried at high vacuum overnight. No further purification was needed, as the product was obtained as a white solid with 70% isolated yield.

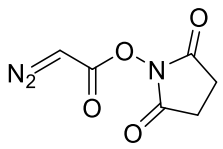
^{1}H-NMR spectrum and melting point match the ones described in literature.[150]

^{1}H NMR (300 MHz, $CDCl_3$): δ 11.00 (br s, 1H), 10.05 (br s, 1H), 7.78 (d, 2H, J = 8.3Hz), 7.40 (d, 2H, J = 8.5Hz), 7.29 (s, 1H), 2.39(s, 3H).



(E)-2-(2-tosylhydrazono)acetyl chloride 17. Compound **17** was prepared according to literature-known procedures.[151] In a flame-dried flask, (E)-2-(2-tosylhydrazono)acetic acid (1.33 g, 5.49 mmol) was suspended in freshly distilled toluene (6.46 ml), then thionyl chloride (0.801 ml, 10.98 mmol) was slowly added dropwise to the flask. The reaction was then heated to reflux under argon atmosphere. The reaction was stopped after 2 hours by quickly cooling the flask to room temperature with a water bath. The reaction solution was filtered through celite, then the filtrate was concentrated at reduced pressure, yielding a solid crude. The crude was dissolved in warm dry toluene, then hexane was added until the formation of precipitate was observed, which was filtered and washed with cold hexane. No further purification was performed, as the product was obtain as pale yellow

crystals, with melting point corresponding to what indicated in the literature (mp: 104-108 °C).[151] Yield: 83%



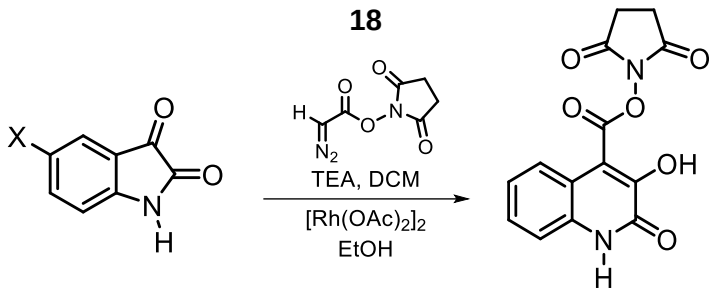
NHS – Diazoacetate 18. Compound **18** was prepared according to literature procedures.[120]

(E)-2-(2-tosylhydrazono)acetyl chloride (2 g, 7.67 mmol) was dissolved in 20 mL of freshly distilled DCM and slowly added dropwise over 1h to a stirred suspension of *N*-hydroxy succinimide (0.971 g, 8.44 mmol) and Na_2CO_3 (1.220 g, 11.51 mmol) in 15 mL DCM stirring at 0 °C under argon atmosphere. After the addition was complete, the reaction was kept under stirring at 0 °C for 1 hour and then 3 hours at room temperature, after which the reaction mixture was put centrifuged in plastic test tubes. After centrifugation, the surnatant was filtered through celite and the filtrate was concentrated at reduced pressure, yielding a pale green solid. The crude was dissolved in a minimal amount of DCM and then precipitated by dropwise addiction of hexane, the resulting solid was collected and washed with cold hexane, yielding the product as a pale green solid with 43% yield.

The NMR spectra of the product match what is reported in the literature.[120]

^{1}H NMR (300 MHz, $CDCl_3$): δ 5.21 (s, 1H); 2.84 (s, 4H).

- **Synthesis of 3HQ-NHS derivatives 21 and 28**

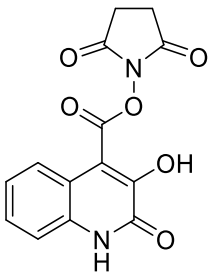


19 X=H
27 X=F

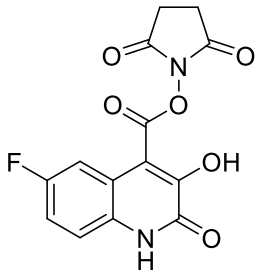
21 X=H
28 X=F

Scheme 38 – Synthetic route towards 3HQ-NHS derivatives 21 and 28

4-NHS-3HQs **21** and **28** were prepared adapting a procedure previously developed in our group.[119] In a thick-walled vial for centrifuge, 0.606 mmol of the appropriate isatin derivative were dissolved, along with NHS-Diazoacetate **18** (150 mg, 0.818 mmol) in dry DCM (3 ml) and stirred with dry triethylamine (16.88 µl, 0.121 mmol) at room temperature under Argon atmosphere. When complete conversion of the starting isatin was observed, the solvent was evaporated at reduced pressure, then the resulting crude was dissolved in absolute Ethanol (3 ml), and then diacetoxyrhodium (2.68 mg, 6.06 µmol) was added to the mixture, causing a rapid bubbling of the reaction mixture, with consequent formation of a beige precipitate. The reaction was stirred for 30 minutes, then it was centrifuged, and the solid was collected by decantation. The resulting precipitate was washed with ethanol and then twice with diethyl ether, yielding the corresponding NHS-3-hydroxyquinolin-2(1H)-one derivatives as pure solids.



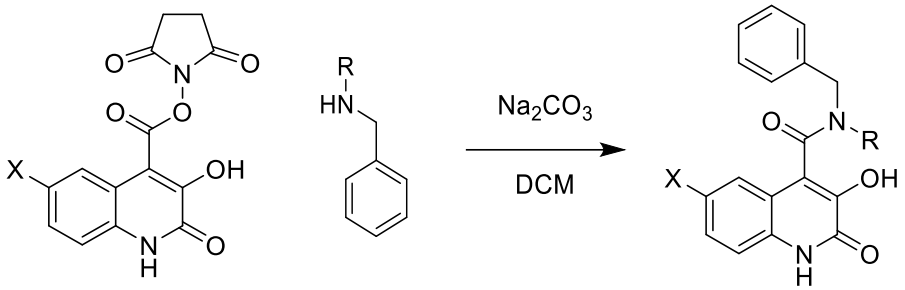
4-NHS-3-HQ 21. Grey solid, yield 60%. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.53 (s, 1H), 11.16 (s, 1H), 7.80 (dd, J = 8.1, 1.3 Hz, 1H), 7.46 – 7.33 (m, 2H), 7.26 (ddd, J = 8.3, 6.8, 1.6 Hz, 1H), 2.91 (d, J = 6.4 Hz, 4H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 170.72, 161.64, 157.68, 147.46, 133.31, 127.90, 123.72, 123.48, 116.98, 116.14, 111.04, 26.12. **LRMS (m/z)**: calculated for $[M+H]^+$ $C_{14}H_{11}N_2O_6$: 303.1; found: 303.2



4-NHS-3-HQ 28. Grey solid, yield 76%. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.36 (s, 1H), 7.74 (dd, J = 11.1, 2.5 Hz, 1H), 7.32 (dd, J = 8.9, 5.3 Hz, 1H), 7.23 – 7.12 (m, 1H), 2.89 (s, 4H).**^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 173.22, 170.83, 161.38, 158.33 (d, J = 237.8 Hz),

157.68, 150.42, 129.66, 118.62, 117.91 (d, J = 8.6 Hz), 115.05 (d, J = 24.1 Hz), 109.08 (d, J = 25.7 Hz), 26.11. **LRMS (m/z)**: calculated for $[M+H]^+$ $C_{14}H_{10}FN_2O_6$: 321.1; found: 321.2.

- **Synthesis of 3HQ Benzylamide derivatives 22, 25 and 40**



21 X=H

28 X=F

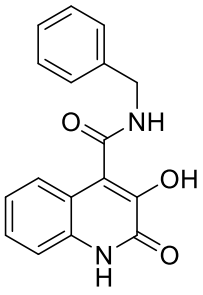
22 X=H, R=H

25 X=F, R=H

40 X=F, R=CH_3

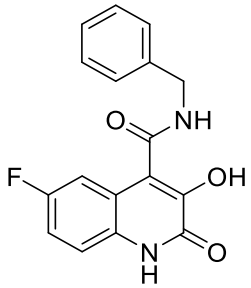
Scheme 39 – General route towards benzyl amide 3HQs 22, 25 and 40

In a thick-walled glass vial, 0.110 mmol of the appropriate amine and Na_2CO_3 (106 mg, 1.000 mmol were mixed in 1 mL of dry DCM, then 0.1 mmol of the appropriate 4-NHS-3HQ was added to the mixture, which was stirred overnight under argon atmosphere at room temperature. After 16h of reaction, the solvent was evaporated at reduced pressure, the resulting crude was suspended in 0.5 mL of water, and then treated with aqueous HCl 1M until it reached pH 4. The resulting suspension was centrifuged and the surnatant was removed and the resulting solid was washed again with water, then dried at high vacuum.

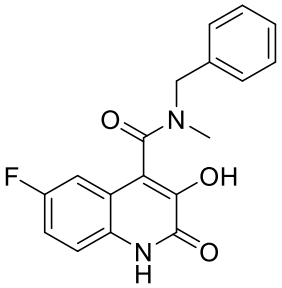


3HQ 22. Grey solid, yield 52%. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.21 (s, 1H), 9.82 (s, 1H), 9.02 (t, J = 6.0 Hz, 1H), 7.44 – 7.22 (m, 8H), 7.15 (dt, J = 8.0, 3.8 Hz, 1H), 4.51 (d, J = 6.0 Hz, 2H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 164.90, 158.81, 142.46, 139.68, 133.58,

128.73, 127.68, 127.25, 127.14, 124.24, 122.82, 121.39, 118.84, 115.64, 42.68. **LRMS (ESI)**: 295.12 ($M+H^+$). **Exact mass**: 318.30026 ($M+Na^+$).



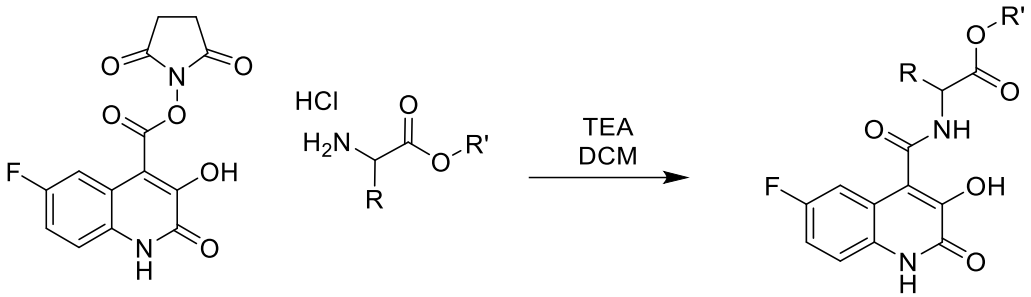
3HQ 25. Grey solid, yield 92%. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** *δ* 12.28 (s, 1H), 10.09 (s, 1H), 9.02 (t, J = 6.0 Hz, 1H), 7.43 – 7.29 (m, 5H), 7.23 (ddd, J = 11.6, 9.0, 4.9 Hz, 2H), 7.04 (dd, J = 10.3, 2.7 Hz, 1H), 4.50 (d, J = 6.0 Hz, 2H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** *δ* 164.65, 159.58, 158.57, 156.44, 144.39, 139.65, 130.09, 128.74, 127.72, 127.30, 120.12 (d, J = 9.4 Hz), 119.89 (d, J = 3.7 Hz), 117.43 (d, J = 8.9 Hz), 114.72 (d, J = 23.4 Hz), 109.20 (d, J = 24.3 Hz), 42.76. **LRMS (ESI):** 313.21 ($M+H^+$). **Exact mass:** 313.09771 ($M+H^+$).



3HQ 40. Grey solid, yield 96%. **^{1}H-NMR (DMSO-d6):** *δ* 12.30 (bs, 1H); 10.25 (bs, 1H), 7.41-7.00 (m, 7H), 6.97-6.88 (dd, 1H); 2 rotamers (1:1.6 ratio) Bn CH2: [4.88-4.63 (dd, 1.2H), 4.4 (s, 0.77H)]; N-CH3: [2.93 (s, 1.05H) 2.79 (s, 1.86 H).

^{13}C-NMR (DMSO-d6): *δ* 165.53, 159.68, 158.11, 143.17, 137.58, 136.9, 130.58, 129.03, 128.88, 128.16, 127.76, 119.36, 117.65, 115.34, 115.02, 108.94, 108.62, [53.91, 49.76], [35.26, 32.15]. **Exact mass:** 325.1010 $(M-H)^-$.

- **Synthesis of 3HQ amino acid derivatives 29-35**



28

29: R=H, R'=CH_3

30: R=CH_3, R'=CH_3

31: R=$CH_2CH(CH_3)_2$, R'=CH_3

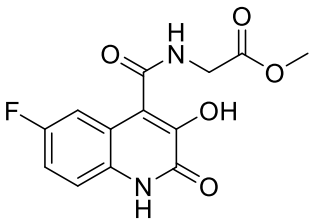
32: R=Ph, R'=CH_3

33: R= CH_2Ph, R'= CH_2CH_3

34: R=CH_2OH, R'=CH_3

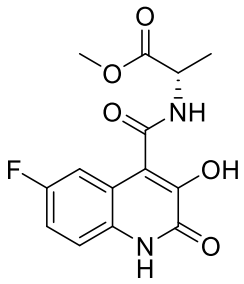
35: R=CH_2SH, R'=CH_3

In a thick-glassed vial, 0.749 mmol of the appropriate amino acid ester were suspended in dry DCM (3.5 mL) and stirred with dry triethylamine (125 µl, 0.899 mmol) until the solution became clear, then intermediate **28** (200 mg, 0.625 mmol) was added and the mixture was stirred overnight at room temperature under argon atmosphere. The solvent was subsequently evaporated at reduced pressure, then the crude was suspended in 1 mL of water, to which aqueous HCl 1M was slowly added dropwise until pH 4 was reached. The resulting suspension was centrifuged, the surnatant was removed and the resulting solid was washed again with water, then collected and dried in high vacuum.

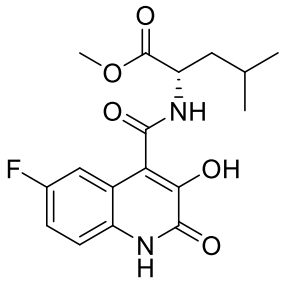


Glycine-3-HQ 29. Grey solid, yield 72%. **^{1}H NMR (300 MHz, $(CD_3)_2SO$): *δ*** 12.27 (s, 1H), 10.17 (d, J = 29.8 Hz, 1H), 9.04 (s, 1H), 7.45 (dd, J = 10.5, 2.7 Hz, 1H), 7.32 (dd, J = 8.9, 5.1 Hz, 1H), 7.23 (td, J = 8.6, 2.8 Hz, 1H), 4.06 (s, 2H), 3.72 (s, 3H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** *δ* 170.21, 164.87, 157.98, 157.63 (d, J = 236.9 Hz), 143.53, 129.67, 119.68

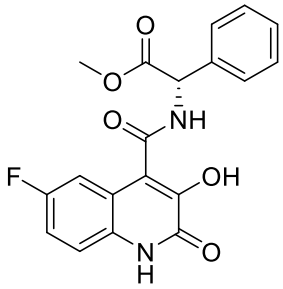
(d, J = 9.5 Hz), 119.46 (d, J = 3.5 Hz), 116.83 (d, J = 8.9 Hz), 114.40 (d, J = 24.3 Hz), 109.21 (d, J = 24.2 Hz), 51.88, 40.80. **LRMS (ESI):** 295.03 $(M+H)^+$. **Exact mass:** 293.05801 $(M-H)^-$.



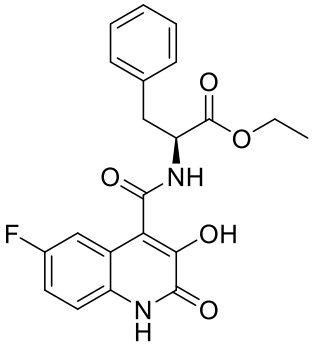
Alanine-3-HQ 30. Grey solid, yield 71%. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.25 (s, 1H), 10.02 (s, 1H), 8.99 (d, J = 6.5 Hz, 1H), 7.34 (ddd, J = 9.0, 6.1, 3.8 Hz, 2H), 7.23 (td, J = 8.6, 2.8 Hz, 1H), 4.47 (p, J = 7.5 Hz, 1H), 3.72 (s, 3H), 1.37 (d, J = 7.3 Hz, 3H).**^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 173.38, 164.66, 158.42, 158.07 (d, J = 237.1 Hz), 143.71, 130.15, 120.20 (d, J = 4.5 Hz), 120.17, 117.30 (d, J = 8.7 Hz), 114.88 (d, J = 24.2 Hz), 109.55 (d, J = 24.6 Hz)., 52.44, 48.45, 16.95. **LRMS (ESI):** 309.04 $(M+H)^+$. **Exact mass:** 307.0738 $(M-H)^-$.



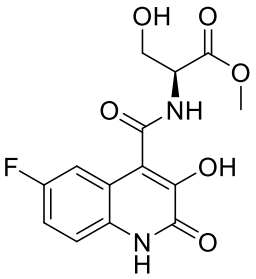
Leucine-3-HQ 31. Grey solid, yield 71%. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.27 (s, 1H), 9.90 (s, 1H), 9.05 (s, 1H), 7.41 – 7.15 (m, 3H), 4.45 (d, J = 22.3 Hz, 1H), 3.71 (s, 3H), 1.82 – 1.47 (m, 3H), 0.91 (d, J = 6.4 Hz, 6H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 173.30, 165.00, 158.47, 158.04 (d, J = 237.0 Hz), 144.02, 130.08, 120.21 (d, J = 9.5 Hz), 120.02, 117.29 (d, J = 8.9 Hz), 114.79 (d, J = 24.5 Hz), 109.42 (d, J = 24.6 Hz), 52.44, 51.10, 24.69, 23.32, 21.58. **LRMS (ESI):** 351.11 $(M+H)^+$. **Exact mass:** 349.12054 $(M-H)^-$.



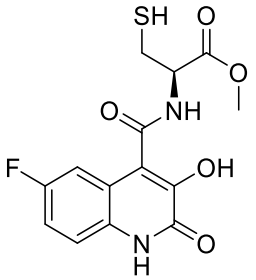
Phanylglycine-3-HQ 32. Grey solid, yield 76%. ^{1}H NMR (300 MHz, $(CD_3)_2SO$): δ 12.24 (s, 1H), 10.01 (s, 1H), 9.41 (d, J = 6.6 Hz, 1H), 7.36 (m, 8H), 5.63 (d, J = 6.6 Hz, 1H), 3.70 (s, 3H). ^{13}C NMR (75 MHz, $(CD_3)_2SO$): δ 171.36, 164.82, 158.39, 158.06 (d, J = 237.1 Hz), 143.87, 135.89, 130.16, 129.09, 128.89, 128.62, 120.25 (d, J = 9.6 Hz), 119.89 (d, J = 3.5 Hz), 117.28 (d, J = 8.9 Hz), 114.82 (d, J = 24.3 Hz), 109.49 (d, J = 24.7 Hz), 57.24, 52.83. **LRMS (ESI):** 371.03 $(M+H)^+$. **Exact mass:** 369.08901 $(M-H)^-$.



Phanylalanine-3-HQ 33. Grey solid, yield 84%. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.08 (s, 1H), 9.83 (s, 1H), 7.55 – 7.17 (m, 7H), 7.15 – 7.00 (m, 1H), 4.65 (dt, J = 8.7, 6.1 Hz, 1H), 4.10 (q, J = 7.1 Hz, 2H), 3.04 (qd, J = 13.7, 7.4 Hz, 2H), 1.15 (t, J = 7.1 Hz, 3H).**^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 171.75, 164.68, 158.40, 157.98 (d, J = 237.3 Hz), 143.80, 137.63, 130.09, 129.67, 128.69, 127.04, 120.06 (d, J = 2.8 Hz), 119.92, 117.26 (d, J = 9.1 Hz), 114.88 (d, J = 24.2 Hz), 109.49 (d, J = 24.6 Hz), 61.20, 54.46, 36.87, 14.38. **LRMS (ESI):** 399.01 $(M+H)^+$. **Elemental analysis** calculated (%) for C21H19FN2O5 + 0.7 H_2O: C, 61.37; H, 5.00; N, 6.82; Found: C, 61.06; H, 5.09; N, 7.21.

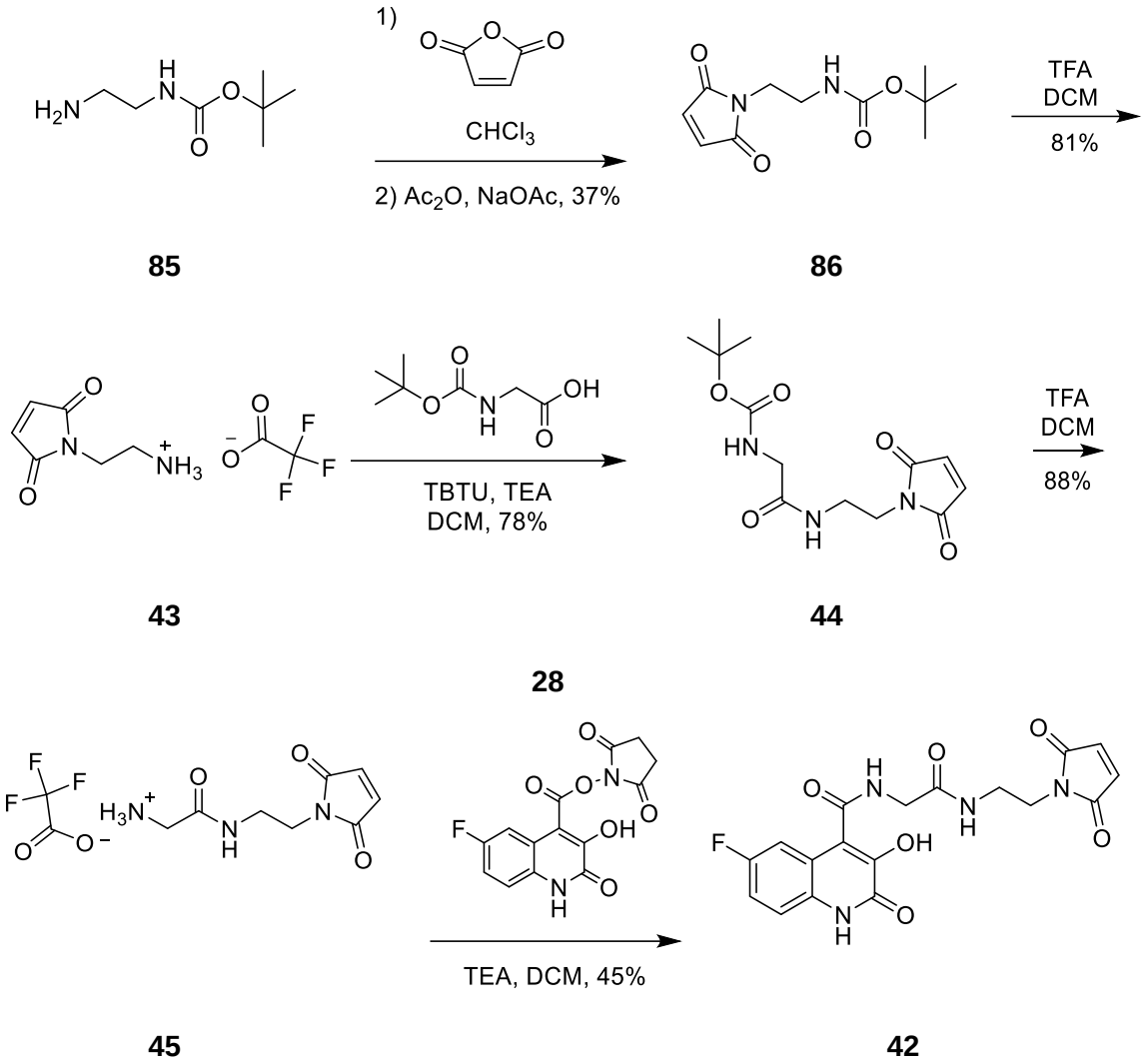


Serine-3-HQ 34. Grey solid, yield 32%. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.25 (s, 1H), 10.08 (s, 1H), 8.82 (d, J = 7.1 Hz, 1H), 7.40 (dd, J = 10.5, 2.7 Hz, 1H), 7.32 (dd, J = 8.9, 5.1 Hz, 1H), 7.23 (td, J = 8.6, 2.7 Hz, 1H), 4.55 (dt, J = 6.6, 4.4 Hz, 1H), 3.79 (dt, J = 9.5, 4.9 Hz, 2H), 3.72 (s, 3H).**^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 171.28, 164.92, 158.39, 158.06 (d, J = 236.9 Hz), 143.83, 130.15, 120.15 (d, J = 9.6 Hz), 119.97 (d, J = 3.4 Hz), 117.26 (d, J = 8.9 Hz), 114.85 (d, J = 24.6 Hz), 109.74 (d, J = 25.0 Hz), 61.44, 55.69, 52.44. **LRMS (ESI):** 324.89 $(M+H)^+$. **Exact mass:** 246.16579 (Hydrolysed amide + Na) $^+$.



Cysteine-3-HQ 35. Grey solid, yield 26%. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.31 (s, 1H), 10.11 (s, 1H), 9.20 (s, 1H), 7.42 – 7.10 (m, 3H), 4.79 (d, J = 8.7 Hz, 1H), 3.74 (s, 3H), 3.28 (dd, J = 14.1, 4.6 Hz, 1H), 3.09 (dd, J = 13.9, 9.4 Hz, 1H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 171.21, 164.99, 158.34, 158.06 (d, J = 237.2 Hz), 143.95, 130.11, 119.95 (d, J = 9.6 Hz), 119.75 (d, J = 3.4 Hz), 117.34 (d, J = 9.1 Hz), 114.96 (d, J = 24.8 Hz), 109.56 (d, J = 24.6 Hz),61.96, 52.85, 52.13. **LRMS (ESI):** 341.03 $(M+H)^+$. **Exact mass:** 339.20033 $(M-H)^+$.

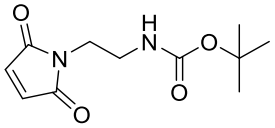
- Synthesis of 3HQ boron hotspot **42**



85 **86** **43** **44** **28** **45** **42**

Scheme 40 – Synthetic route towards 3HQ BHS 42

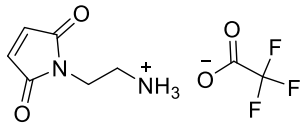
Compound **44** was prepared by coupling commercially available Boc-Glycine and maleimide **43**, which was prepared following the procedures reported in literature.[152] Deprotection of **44** yielded linker **45**, that was reacted with NHS-HQ **28** to afford compound 42**.**



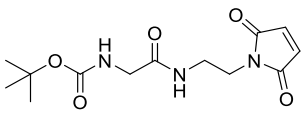
Boc-diethylamine-maleimide 86. This compound was prepared according to literature-known procedures[152] starting from commercially available tert-butyl (2-

aminoethyl)carbamate. Tert-butyl (2-aminoethyl) carbamate (2.54 g, 15.85 mmol) was dissolved in chloroform (40.9 ml) and slowly added dropwise to a stirring solution of maleic anhydride (1.340 g, 13.67 mmol) in chloroform (13.63 ml). The reaction was left stirring overnight, then the solvent evaporated at reduced pressure. The crude intermediate was dissolved in Acetic anhydride (40.9 ml), along with sodium acetate (6.17 g, 75 mmol) and stirred at 100 °C for 90 minutes, then it was poured in cold water (136 ml) and stirred for 1h. Then, the reaction mixture was extracted with ethyl acetate, the organic layer was washed with a $NaHCO_3$ saturated aqueous solution, then dried over sodium sulphate and concentrated at reduced pressure, yielding a light brown solid crude. The crude was purified by flash chromatography with a hexane/ethyl acetate 7/3 eluent mixture, yielding the product as a white solid with 31% yield. ^{1}H-NMR spectrum matches the one described in literature.[152]

^{1}H NMR (300 MHz, $CDCl_3$): δ 6.70 (s, 2H); 4.73 (s, 1H); 3.72 – 3.62 (m, 2H), 3.35 (q, J = 5.8 Hz, 2H); 1.41 (s, 9H).



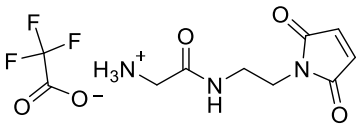
Diethylamine-maleimide TFA salt 43. This compound was prepared according to literature-known procedures.[152] Boc-diethylamine-maleimide **86** (1 g, 4.16 mmol) was dissolved in DCM (10.00 ml) at 0 °C and then Trifluoroacetic acid (5.00 ml, 64.9 mmol) was added. The reaction was stirred for 1h, then the solvent was removed by evaporation and then the reaction mixture was re-dissolved in MeOH (3 ml), subsequently, diethyl ether (50.0 ml) was added dropwise, promoting precipitation of the product, which was filtered and dried at reduced pressure. The product was isolated as a white solid, obtained with 81% yield, which was used without further purification.



Boc-Glycine-maleimide linker 44. In a flame-dried Schlenk flask, under argon atmosphere, Boc-Glycine (207 mg, 1.180 mmol) was dissolved in DCM and stirred for 10 minutes with triethylamine (0.197 mL, 1.416 mmol), then TBTU (455 mg, 1.416 mmol) was added and the mixture was stirred for 30 additional minutes. Subsequently,

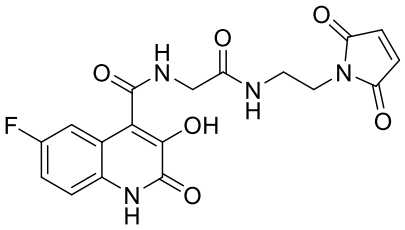
diethylamine-maleimide TFA salt **43** (600 mg, 2.361 mmol) was added to the mixture, along with triethylamine (0.428 mL, 3.07 mmol). The reaction was stirred for 2 days at room temperature under argon atmosphere. The reaction mixture was then dried at reduced pressure, and the crude was purified by column chromatography on silica gel with AcOEt as eluent, yielding the product as a white solid with 75% yield.

^{1}H NMR (300 MHz, $(CD_3)_2SO$): δ 6.71 (s, 2H), 6.52 (s, 1H), 5.10 (s, 1H), 3.79 – 3.65 (m, 4H), 3.47 (q, J = 5.1 Hz, 2H), 1.45 (s, 9H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ170.86, 170.04, 156.00, 134.23, 80.33, 44.28, 38.80, 37.37, 28.31. **HRMS (m/z)**: calculated for $[M\text{-}tButyl+OH]^-$ $C_9H_{12}N_3O_6$: 258.0726; found: 257.9588



Glycine-maleimide TFA salt 45. Boc-Glycine-maleimide linker **44** (88 mg, 0.296 mmol) was dissolved in DCM (650 µL) at 0 °C and then Trifluoroacetic acid (365 µL, 4.74 mmol) was added to the mixture. The reaction was stirred for 1h, after which the solvent was removed by evaporation and then the reaction mixture was re-dissolved in MeOH (200 µL). Diethyl ether (3.2 mL) was then added dropwise to promote precipitation of the product. The solid was recovered by centrifugation of the crude mixture, washed with diethyl ether and dried at high vacuum, yielding the product as a white solid with 88% yield. The product is used without further purification, even though some small impurities were observed.

^{1}H NMR (300 MHz, $(CD_3)_2SO$): δ 6.73 (s, 2H), 6.43 (s, 1H), 5.05 (s, 1H), 3.78 – 3.69 (m, 4H), 3.53 – 3.46 (m, 2H), 1.47 (s, 9H).



3HQ-BHS 42. In a thick-walled vial glycine-maleimide TFA salt **45** (50 mg, 0.161 mmol) was dissolved in DCM (446 µl), then triethylamine (13.44 µl, 0.096 mmol) was added and the mixture was stirred until it became clear. Subsequently, F-NHS-3-hydroxyquinolin-2(1H)-one **20** (25.7 mg, 0.080 mmol) was added to the mixture, which was stirred

overnight. After 16h of reaction, the solvent was evaporated at reduced pressure, the crude was treated with 1.5 mL of water, then the suspension was treated with 0.05 M aqueous HCl solution until neutral pH was achieved, resulting in the precipitation of the product, which was collected after centrifugation and filtration as a grey solid with 45% yield.

^{1}H NMR (300 MHz, $(CD_3)_2SO$): *δ* 12.24 (s, 1H), 10.10 (bs, 1H), 8.70 (s, 1H), 8.09 (s, 1H), 7.52 (d, J = 13.1 Hz, 1H), 7.30 (dd, J = 8.8, 5.1 Hz, 1H), 7.22 (td, J = 8.6, 2.2 Hz, 1H), 7.00 (s, 2H), 3.78 (d, J = 5.7 Hz, 2H), 3.53 – 3.45 (m, 4H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** *δ* 207.00, 165.01, 159.66, 158.35, 156.53, 143.72, 134.96, 130.17, 120.21 (d, J = 25.1 Hz), 117.22 (d, J = 7.9 Hz), 114.96 (d, J = 24.3 Hz), 110.30, 109.90 (d, J = 8.3 Hz), 46.11, 31.14, 14.51. **HRMS (m/z)**: 385.2925 $(M\text{-}OH)^-$

V.1.3 ESI-MS experiments

The reactions of HQs with model boronic acids were all performed following the following protocol:

22 µL of the appropriate boronic acid solution (100 mM in DMF) were first diluted with 65 µL of ammonium acetate buffer pH 7, then 10 µL of the corresponding HQ solution (22 mM in DMF) were added to the mixture, which was stirred for 5 minutes. After this time, a 10 µL aliquot of the reaction was uptaken and diluted with 190 µL of ammonium acetate buffer and injected into the mass spectrometer for analysis.

Reaction proportions: 1 eq. HQ + 10 eq. BA.

Reaction concentration: 3.4 mM

Injection concentration: 0.18 mM

- **Model reactions of basic 3HQ cores**

A

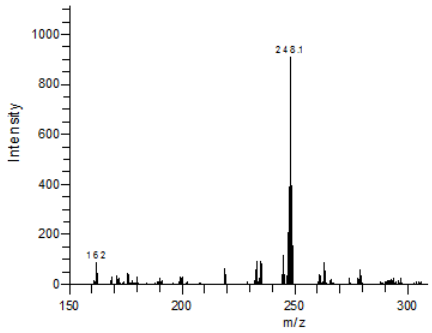
1

m/z [M+H]$^+$: 162.0

2

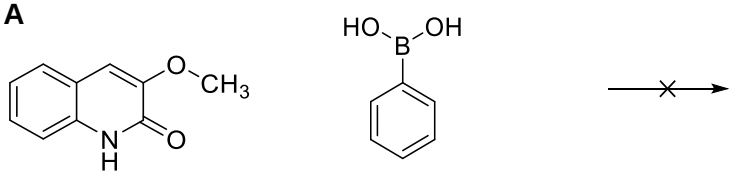
m/z [M-OH]$^+$: 248.1

B



Scheme 41 – Reaction of basic 3HQ core 1 with PBA (A) and relative ESI-MS spectrum (B)

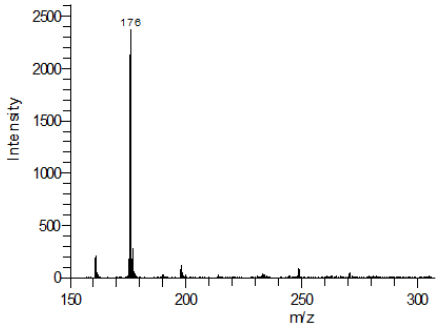
A



4

m/z $[M+H]^+$: 176.0

B



Scheme 42 – Reaction of 3HQ 4 with PBA (A) and relative ESI-MS spectrum (B)

A

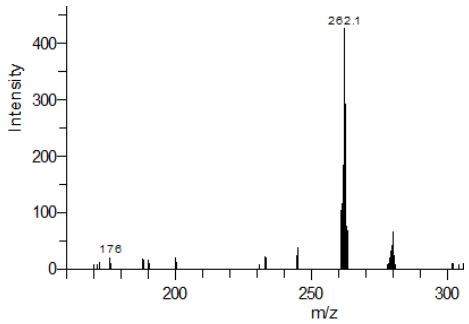
5

m/z $[M+H]^+$: 176.0

87

m/z $[M-OH]^+$: 262.1

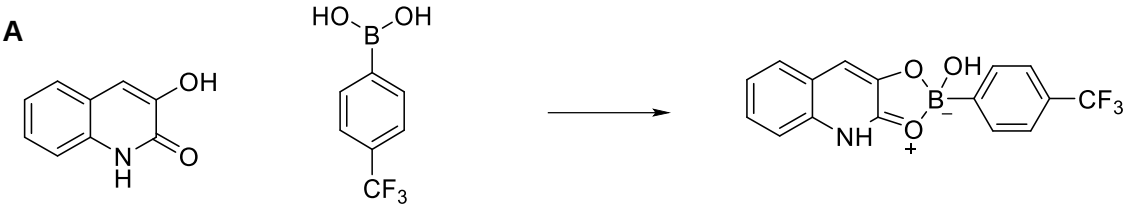
B



Scheme 43 – Reaction of 3HQ 5 with PBA (A) and relative ESI-MS spectrum (B)

- **Reactions of basic 3HQ 1 with different boronic acid derivatives**

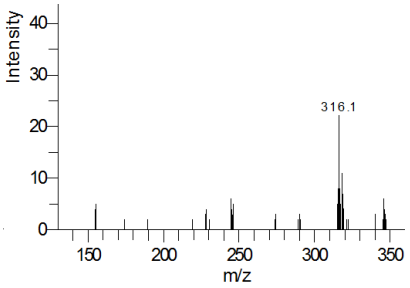
A



1 **6**

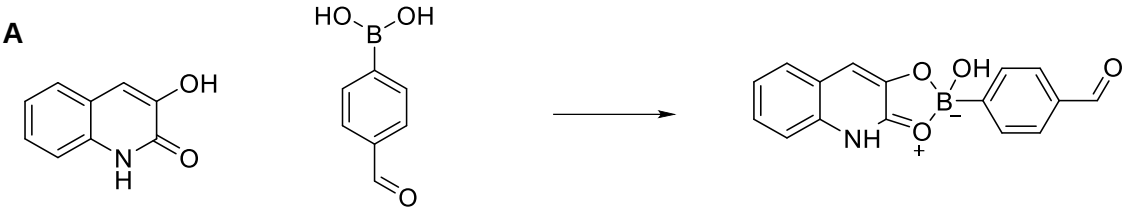
m/z [M+H]$^+$: 162.0 *m/z* [M-OH]$^+$: 316.1

B



Scheme 44 – Reaction of simple 3HQ core 1 with 4-CF_3-PBA (A) and relative ESI-MS spectrum (B)

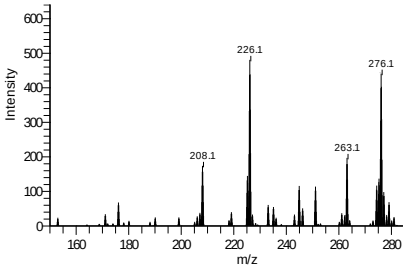
A



1 **7**

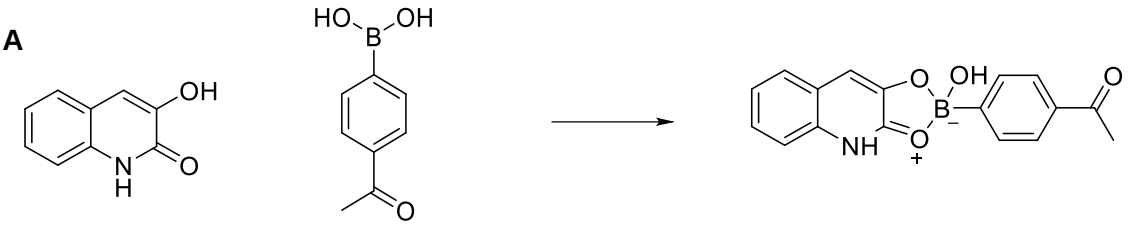
m/z [M+H]$^+$: 162.0 *m/z* [M-OH]$^+$: 276.1

B



Scheme 45 – Reaction of simple 3HQ core 1 with 4-FBBA (A) and relative ESI-MS spectrum (B)

A



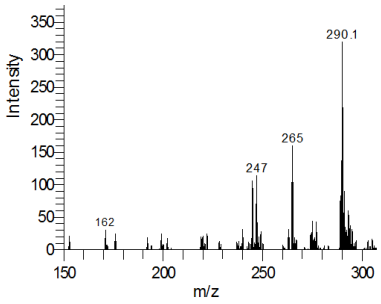
1

m/z [M+H]$^+$: 162.0

8

m/z [M-OH]$^+$: 290.1

B



Scheme 46 – Reaction of simple 3HQ core 1 with 4-FBBA (A) and relative ESI-MS spectrum (B)

A

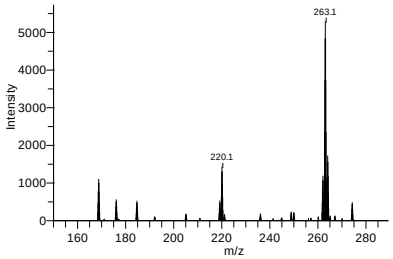
1

m/z [M+H]$^+$: 162.0

9

m/z [M-OH]$^+$: 263.1

B



Scheme 47 – Reaction of simple 3HQ core 1 with 3-NH_2-PBA (A) and relative ESI-MS spectrum (B)

A

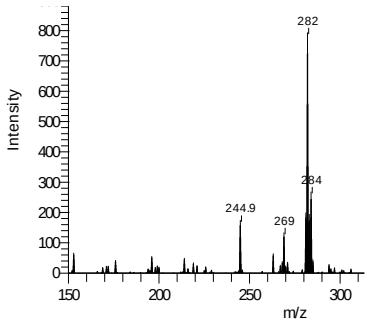
1

m/z [M+H]$^+$: 162.0

10

m/z [M-OH]$^+$: 282.0

B



Scheme 48 – Reaction of simple 3HQ core 1 with 3-Cl-PBA (A) and relative ESI-MS spectrum (B)

A

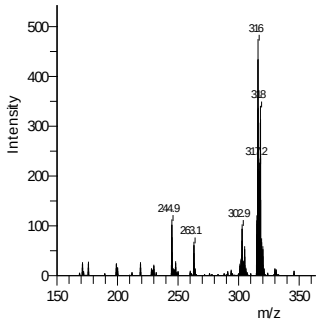
1

m/z [M+H]$^+$: 162.0

11

m/z [M-OH]$^+$: 316.0

B



Scheme 49 – Reaction of simple 3HQ core 1 with 3-Cl-PBA (A) and relative ESI-MS spectrum (B)

A

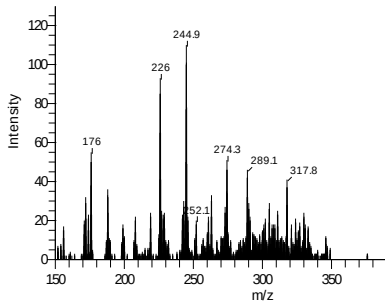
1

m/z [M+H]$^+$: 162.0

11

m/z [M-OH]$^+$: 252.1

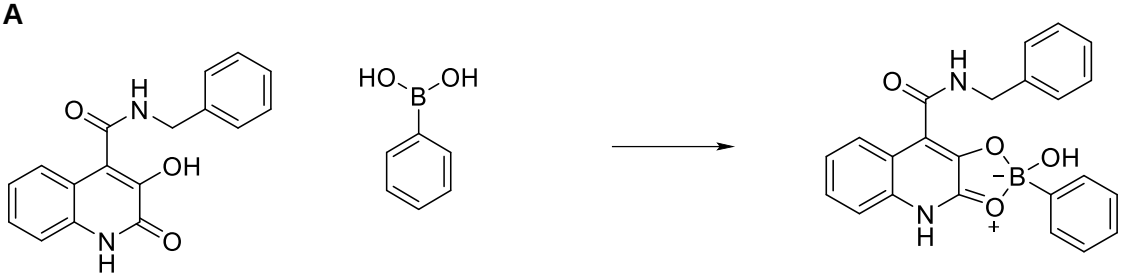
B



Scheme 50 – Reaction of simple 3HQ core 1 with cyclohexyl-BA (A) and relative ESI-MS spectrum (B)

- **Reaction of 3HQ-benzylamide derivatives**

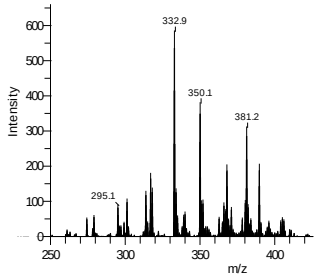
A



22 **23**

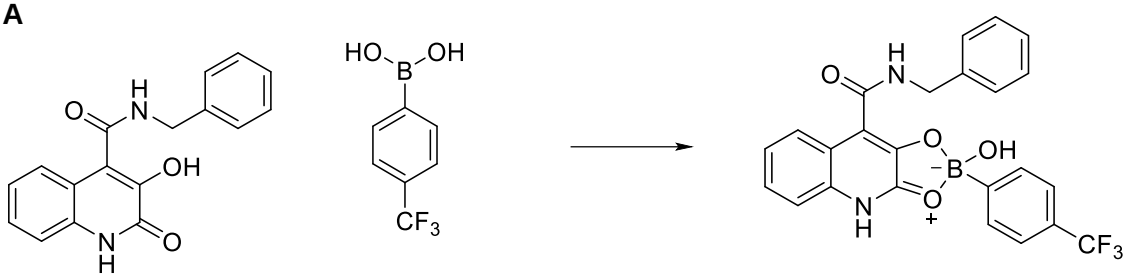
m/z [M+H]$^+$: 295.1 *m/z* [M-OH]$^+$: 381.2

B



Scheme 51 – Reaction of 3HQ-benzylamide derivative 22 with PBA (A) and relative ESI-MS spectrum (B)

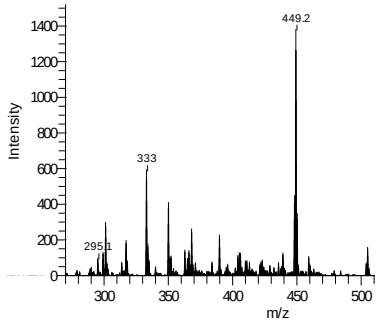
A



22 **24**

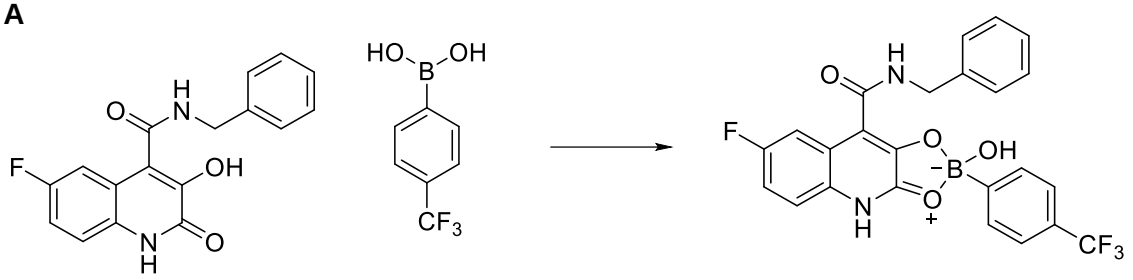
m/z [M+H]$^+$: 295.1 *m/z* [M-OH]$^+$: 449.2

B



Scheme 52 - Reaction of 3HQ-benzylamide derivative 22 with 4-CF_3-PBA (A) and relative ESI-MS spectrum (B)

A



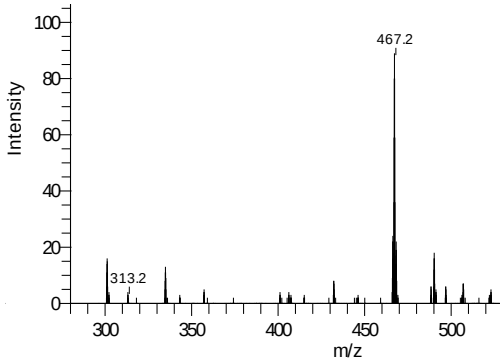
25

m/z $[M+H]^+$: 295.1

24

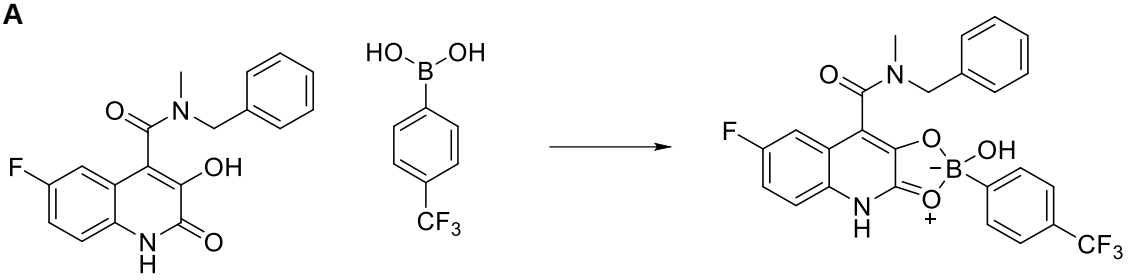
m/z $[M\text{-}OH]^+$: 467.2

B



Scheme 53 - Reaction of 3HQ-benzylamide derivative 25 with 4-CF_3-PBA (A) and relative ESI-MS spectrum (B)

A



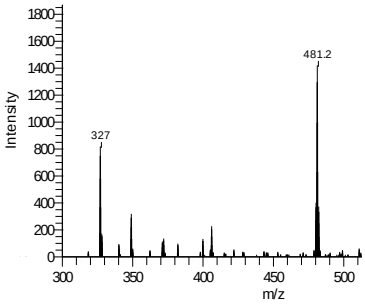
40

m/z [M+H]$^+$: 327.0

41

m/z [M-OH]$^+$: 481.2

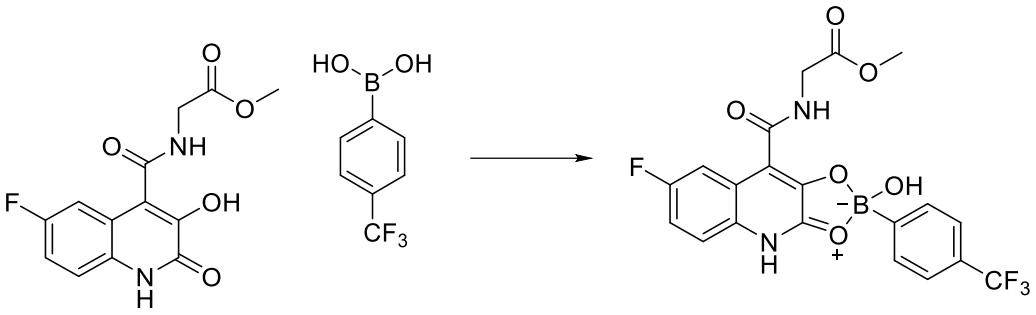
B



Scheme 54 - Reaction of 3HQ-methylbenzylamide derivative 40 with 4-CF_3-PBA (A) and relative ESI-MS spectrum (B)

- **Reaction of 3HQ amino acid derivatives with CF_3-PBA**

A



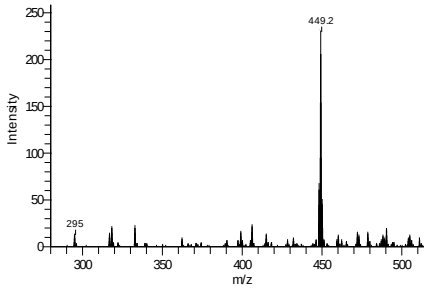
29

m/z [M+H]$^+$: 295.0

36

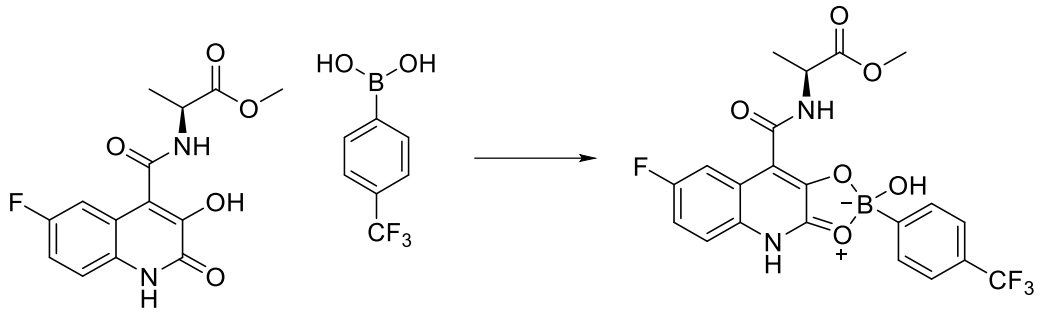
m/z [M-OH]$^+$: 449.2

B



Scheme 55 - Reaction of 3HQ-glycine derivative 29 with 4-CF_3-PBA (A) and relative ESI-MS spectrum (B)

A



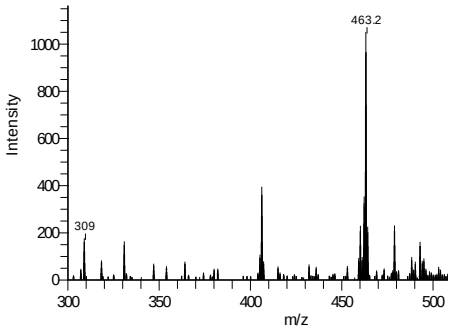
30

m/z $[M+H]^+$: 309.0

88

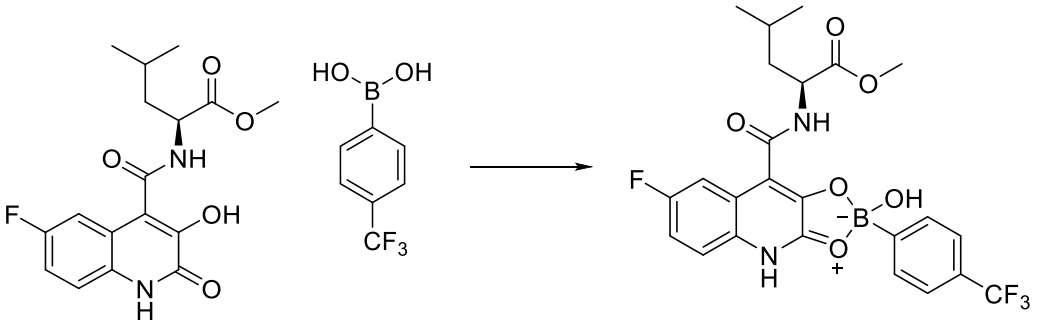
m/z $[M-OH]^+$: 463.2

B



Scheme 56 - Reaction of 3HQ-alanine derivative 30 with 4-CF_3-PBA (A) and relative ESI-MS spectrum (B)

A



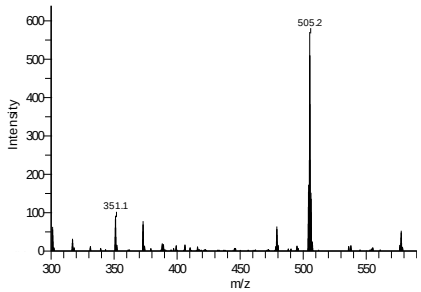
31

m/z [M+H]$^+$: 351.1

37

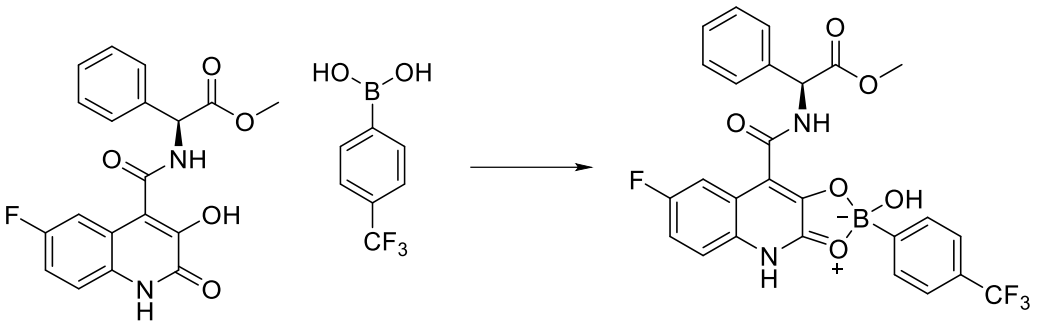
m/z [M-OH]$^+$: 505.2

B



Scheme 57 - Reaction of 3HQ-leucine derivative 31 with 4-CF_3-PBA (A) and relative ESI-MS spectrum (B)

A



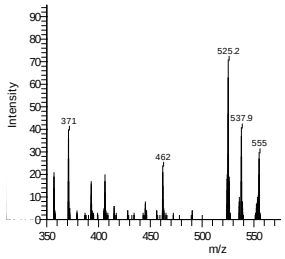
32

m/z [M+H]$^+$: 371.0

89

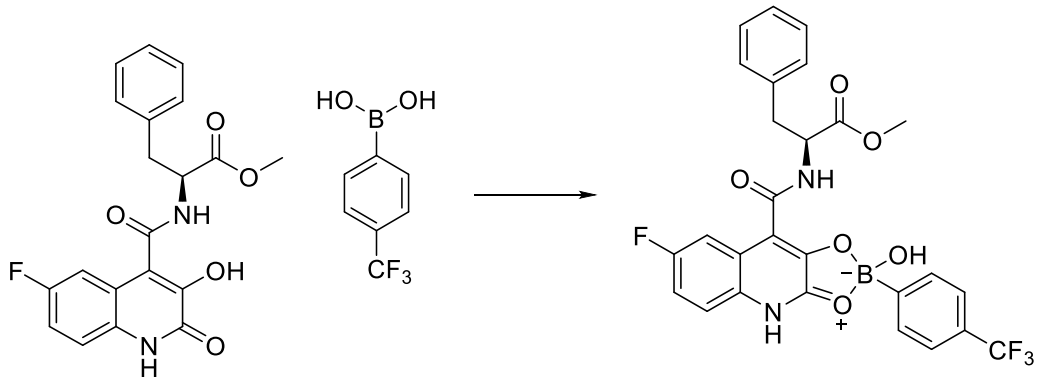
m/z [M-OH]$^+$: 525.2

B



Scheme 58 - Reaction of 3HQ-phenylglycine derivative 32 with 4-CF_3-PBA (A) and relative ESI-MS spectrum (B)

A



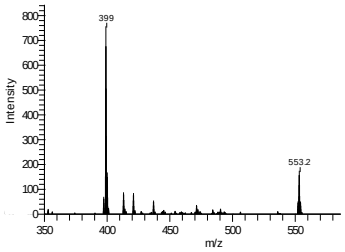
33

m/z $[M+H]^+$: 399.0

38

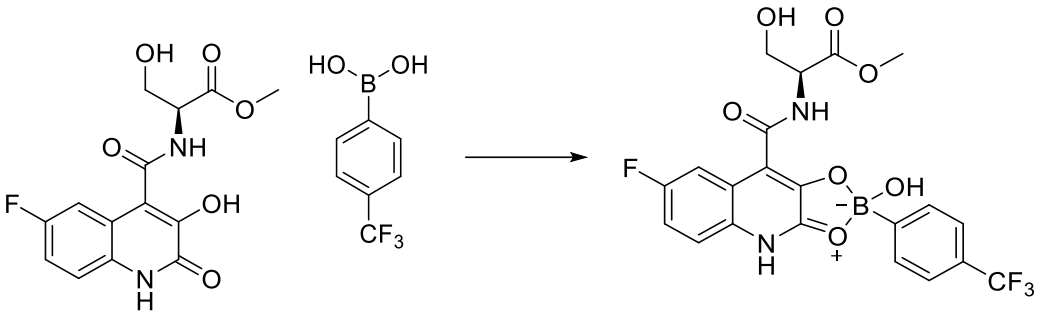
m/z $[M-OH]^+$: 553.2

B



Scheme 59 - Reaction of 3HQ-phenylalanine derivative 33 with 4-CF_3-PBA (A) and relative ESI-MS spectrum (B)

A



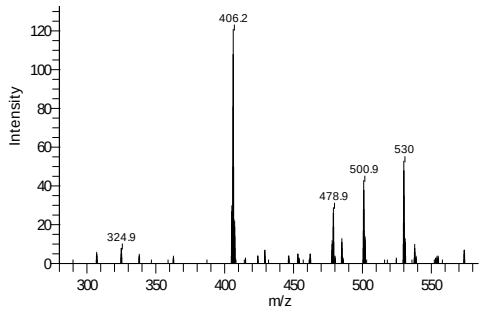
34

m/z $[M+H]^+$: 324.9

90

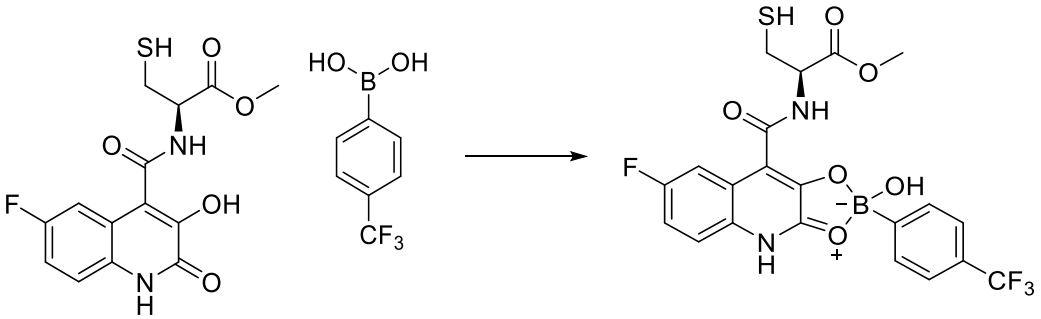
m/z $[M\text{-}OH]^+$: 478.9

B



Scheme 60 - Reaction of 3HQ-serine derivative 34 with 4-CF_3-PBA (A) and relative ESI-MS spectrum (B)

A



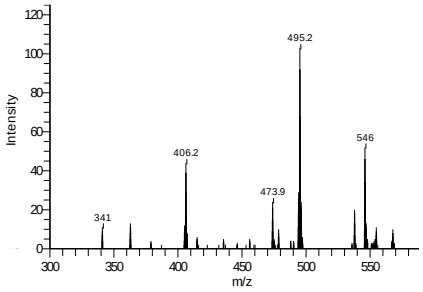
35

m/z $[M+H]^+$: 341.0

91

m/z $[M\text{-}OH]^+$: 495.2

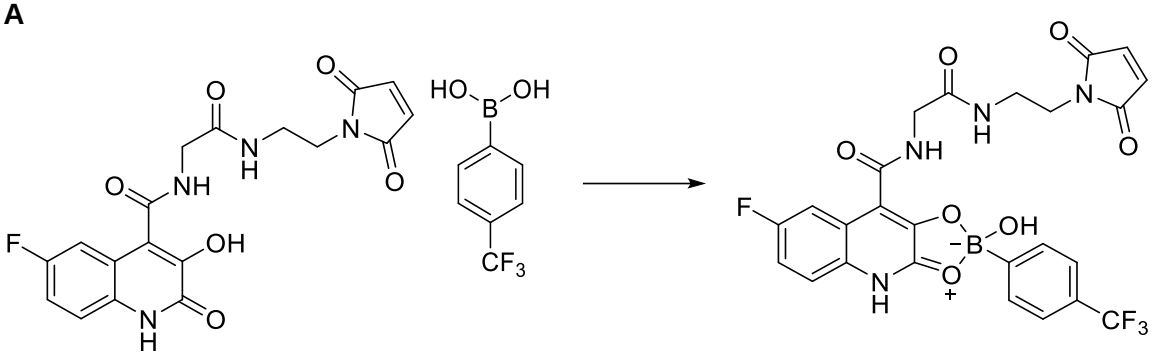
B



Scheme 61 - Reaction of 3HQ-cysteine derivative 35 with 4-CF_3-PBA (A) and relative ESI-MS spectrum (B)

- **Reaction of 3HQ-boron hotspot 42 with 4-CF_3-PBA**

A



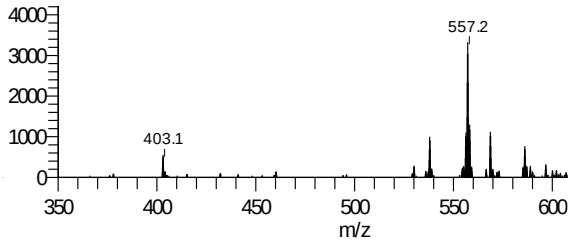
42

m/z $[M+H]^+$: 403.1

43

m/z $[M-OH]^+$: 557.2

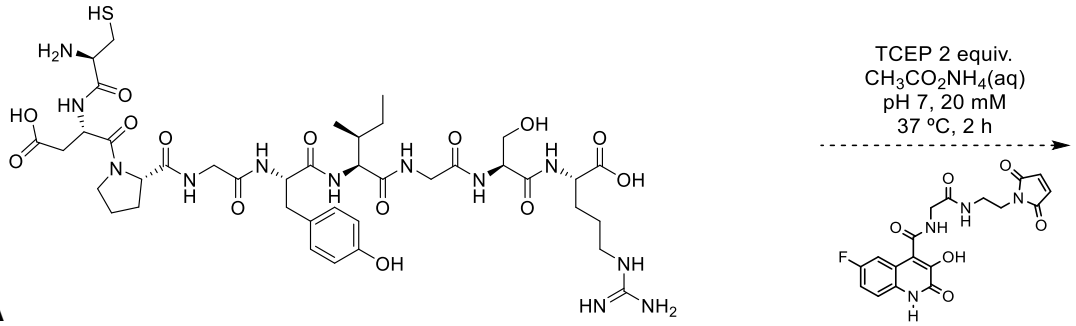
B



Scheme 62 - Reaction of 3HQ- boron hotspot 42 with 4-CF3-PBA (A) and relative ESI-MS spectrum (B)

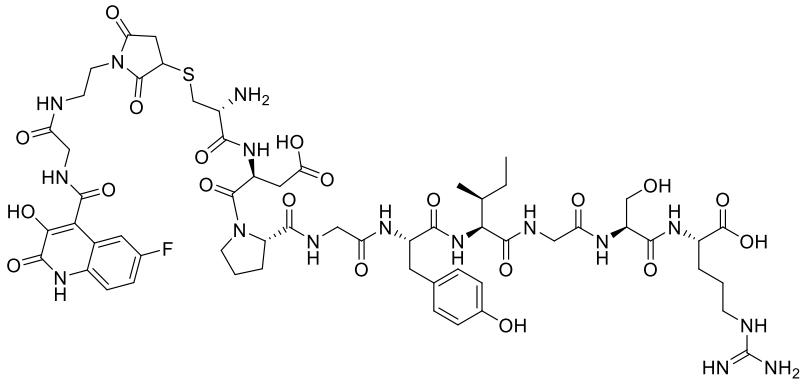
- **Functionalization of Laminin with 42 and subsequent reactions with BAs**

A



CDPGYIGSR laminin fragment **42**

m/z $[M+H]^+$: 967.2

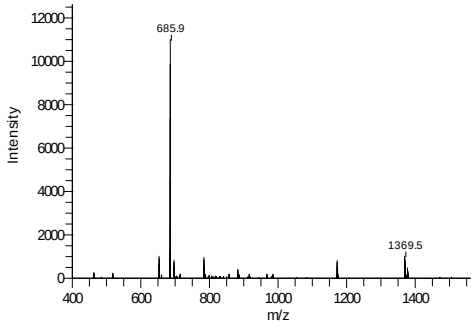


47

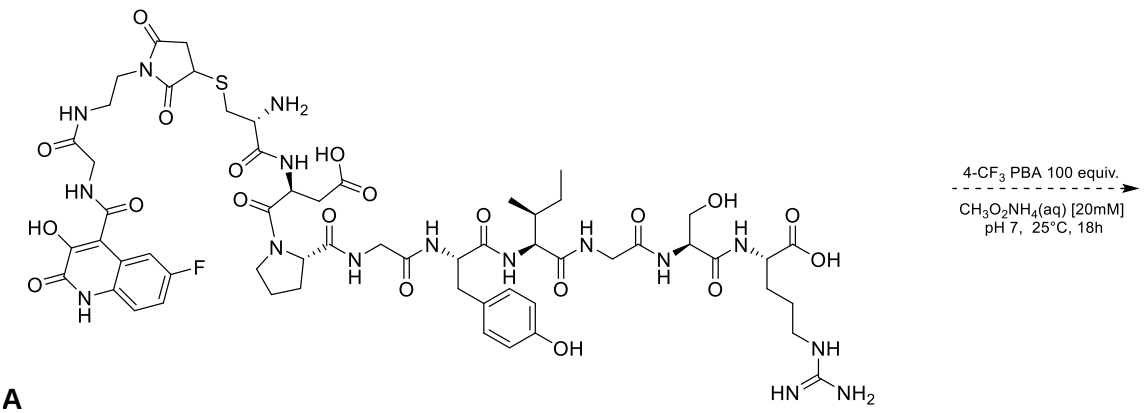
m/z $[M+H]^+$: 1369.5

m/z $[M+2H]^{2+}$: 685.9

B



Scheme 63 – Functionalization of the laminin fragment with 3HQ-boron hotspot 42 (A) and relative ESI-MS spectrum (B)

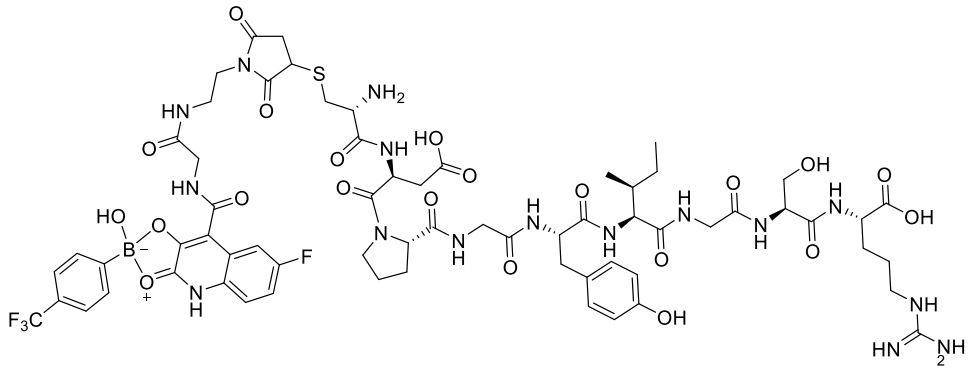


A

47

m/z $[M+H]^+$: 1369.5

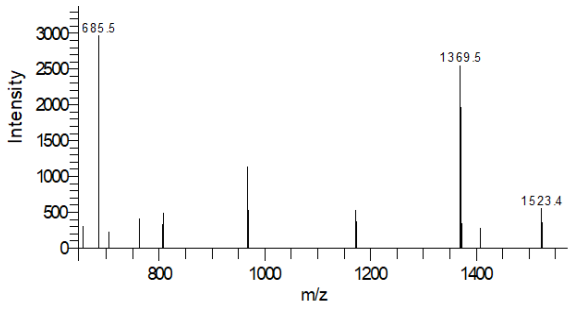
m/z $[M+2H]^{2+}$: 685.9



48

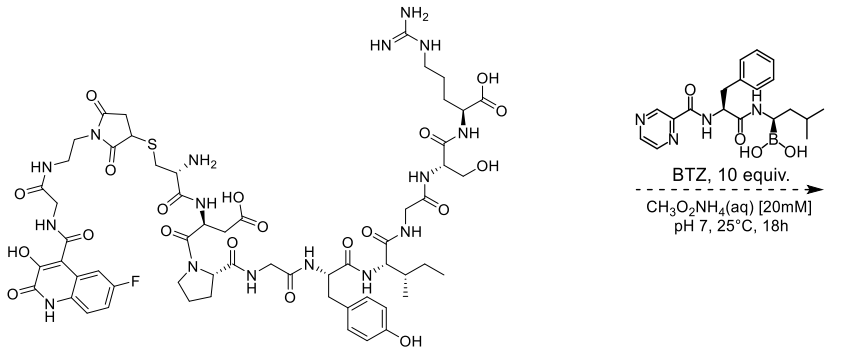
m/z $[M-OH]^+$: 1523.4

B



Scheme 64 – Reaction of 3HQ-modified laminin fragment 47 with 4-CF3 PBA (A) and relative ESI-MS spectrum (B)

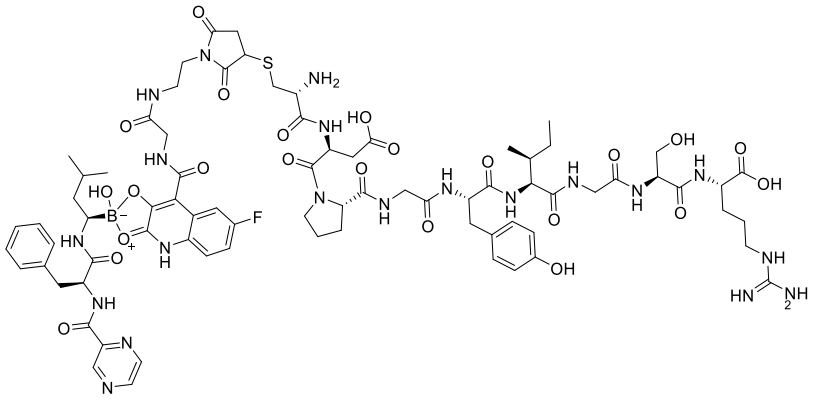
A



47

m/z $[M+H]^+$: 1369.5

m/z $[M+2H]^{2+}$: 685.9

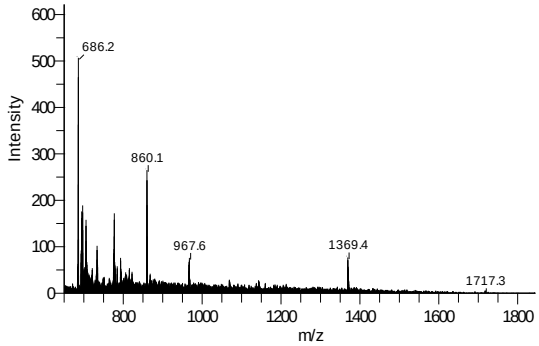


49

m/z $[M-OH]^+$: 1717.3

m/z $[M+H-OH]^{2+}$: 860.1

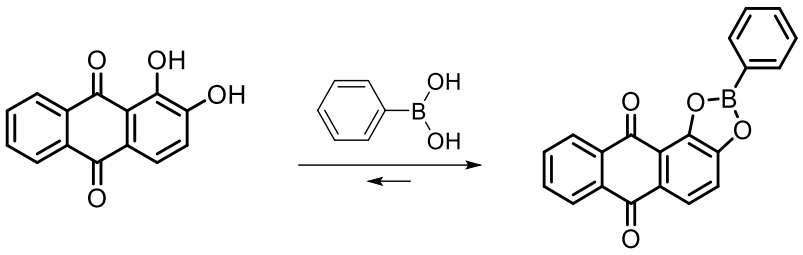
B



Scheme 65 – Reaction of 3HQ-modified laminin fragment 47 with Bortezomib (A) and relative ESI-MS spectrum (B)

V.1.4 Determination of K_a *via* fluorimetric competitive titration with Alizarin

- **Titration of Alizarin with PBA**



Alizarin — **Not Fluorescent**

Alizarin-PBA complex — **Fluorescent**

Scheme 66 – Complexation equilibrium of Alizarin with PBA

The fluorimetric titration of Alizarin with PBA was performed according to the procedure reported by Wang and co-workers .[73]

A 50 µM solution of Alizarin in ammonium acetate (20mM, pH 7)/DMF 1:4 mixture was titrated with a 50 mM solution of PBA in the same solvent mixture. The PBA solution also contained Alizarin 50 µM in order to balance the dilution that occurs during the titration.

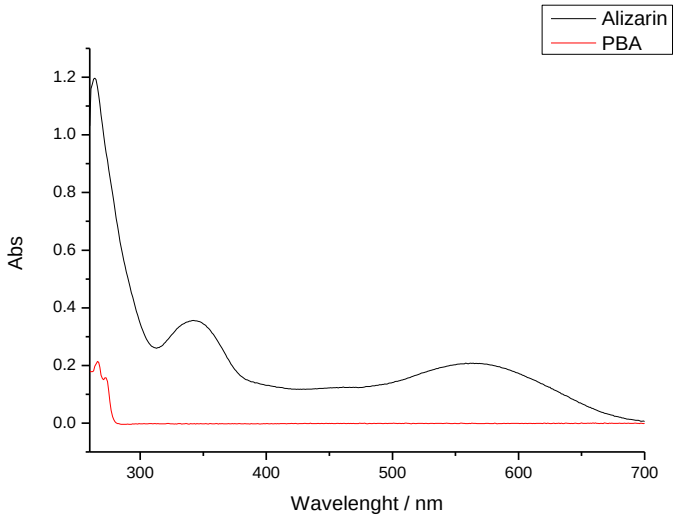


Figure 37 – UV-VIS spectra of Alizarin (black) and PBA (red)

The UV-VIS spectrum of the titration is reported in **Figure 38**, showing the formation of the PBA-Alizarin complex in a concentration-dependant fashion.

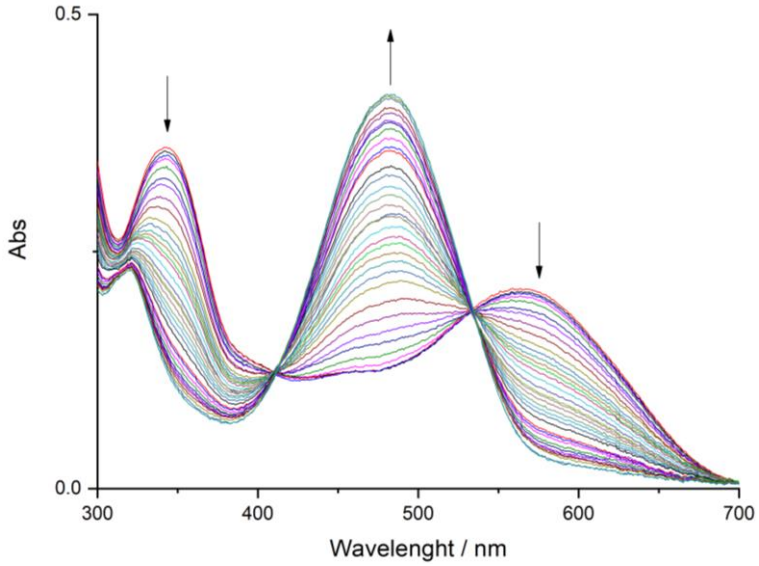


Figure 38 – UV-VIS spectrum of the titration of Alizarin with PBA

The UV-VIS spectrum of the titration presents two isosbestic points, at λ=420 nm and at λ=530 nm, both of which are optimal wavelength for fluorescence excitation. The first of these points, λ=420 nm, was used for exciting the system in the fluorescence titration experiments, yielding the graph in **Figure 39**, showing the concentration-dependant increase of fluorescence of the system due to the formation of the Alizarin-PBA complex.

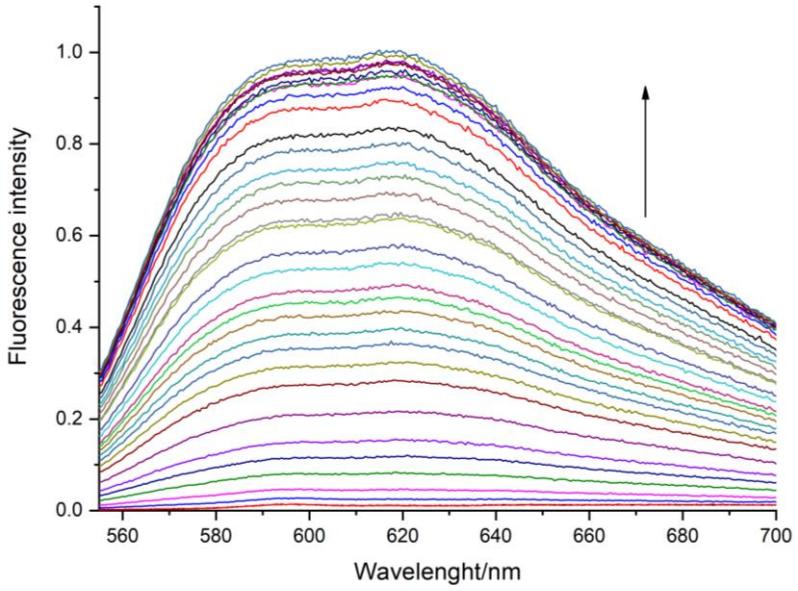


Figure 39 – Fluorescence spectrum of the titration fo Alizarin with PBA

The emission values at the λ=620 nm maximum were used for the calculation of the K_a of AL-PBA complexation in the chosen solvent mixture. The fluorescence intensity data were used to determine the K_a following the Benesi-Hildebrand method, which starts from **Equation 1** that states the relationship between the fluorescence intensity (I_f) and the K_a.

$$\Delta I_f = \frac{(\Delta k p_0 K_a)[L][I_0]}{1 + K_a[L]}$$

Equation 1 - Relationship between fluorescence intensity and association constant

The double reciprocal of **Equation 1**, **Equation 2** yields the Benesi-Hildebrand equation, from which the K_a can be extrapolated as the quotient of the intercept and the slope in a plot of 1/[PBA] vs. $1/\Delta I_f$.

$$\frac{1}{\Delta I_f} = (\Delta k p_0 I_0 K_a)^{-1} \frac{1}{[L]} + (\Delta k p_0 I_0)^{-1}$$

Equation 2 – Benesi-Hildebrand equation

Following this protocol, the fluorescence intensity at 620 nm was plotted in a Benesi-Hildebrand plot (**Figure 40**) to determine the K_a for Alizarin-PBA complexation in this solvent mixture, indicated as $K_{AL.}$

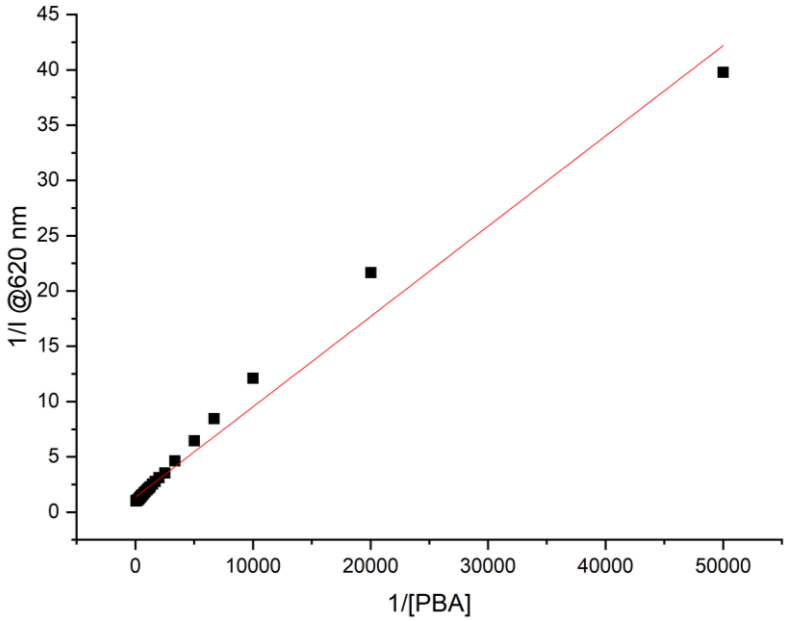


Figure 40 – Benesi-Hildebrand plot for the titration of Alizarin with PBA

The mathematical treatment of the data of the Benesi-Hildebrandt plot is reported in **Table 3** , which indicates a K_{AL} value of 1224 ± 31 M^{-1} for the Alizarin-PBA complexation in the solvent mixture used.

Equation parameters			
Equation	$y = a + bx$		
Pearson's r	0.990580337205665		
Adj. R-Square	0.980624384607107		
K_a calculations			
		Value	Standard Error
1/I	Intercept	1.35904791805114	0.203063928674616
1/I	Slope	$8.1670724116526E^{-4}$	$2.06121649249865E^{-5}$
Calculated K_{AL}	1224 ± 31 M^{-1}		

Table 3 – Equation parameters and mathematical treatment of the Benesi-Hildebrandt plot for the titration of Alizarin with PBA

- **Competitive titration of the PBA-Alizarin system with compounds 29 and 31**

The determination of the K_a for the ligation of compounds **29** and **31** with PBA was determined by competitive fluorescence titration of an equimolar solution of AL and PBA.

The starting equimolar solution of PBA/AL 100 µM in ammonium acetate (20mM, pH 7)/DMF 1:4 mixture was titrated with a 100 mM solution of the appropriate HQ derivative in the same solvent mixture. The HQ solution also contained AL and PBA 100 µM in order to balance the dilution that occurs during the titration.

The fluorescence intensity at a fixed wavelenght was plotted against the concentration of HQ, showing a clear concentration-dependant fluorescence decrease. From these data, it was possible to extrapolate the K_a for the association of our compounds and PBA, following the protocol described by Tirelli and co-workers.[153]

The relationship between K_a of the HQ and K_{AL} is expressed as $K_a = K_{AL} \times K_{exc}$

Where K_{exc} is the exchange constant relative to the competitive equilibrium between AL and the appropriate HQ to bind PBA, which can be expressed as **Equation 3**.

$$K_{exc} = \frac{([\mathrm{HQ} - \mathrm{PBA}]\ [\mathrm{AL}])}{([\mathrm{AL} - \mathrm{PBA}]\ [\mathrm{HQ}])}$$

Equation 3 – Definition of K_{exc}

Using the data from the previous titration of AL with PBA, it is possible to determine [AL-PBA] at point 0 of the titration, where there is no HQ. This value is set as 1 and the fluorescence intensity of the other points of the titration are normalized according to it, yielding the values of [AL-PBA] for each point in the curve. From this value, it is possible to extrapolate the other values for each point, adopting **Equation 4** to describe the system equilibrium.

$$[\mathrm{HQ} - \mathrm{PBA}] = BA_{tot} - [\mathrm{AL} - \mathrm{PBA}]$$

Equation 4 – Definition of the distribution of BA in the system

The aforementioned data treatment yields a set of K_{exc}, one for each point of the titration, the average of these is the K_{exc} used to obtain the desired K_a.

It must be noted that the mathematical method used in this protocol starts from the assumption that there is no free boronic acid in solution (**Equation 4**), which results in divergent K_a values for the first points of the titration. Considering this factor, only the points of the titration where there are 20 equivalents or more of HQ compared to AL were considered, this approach resulted in consistent K_{exc} value across the curve.

Competitive titration of the AL-PBA system with compound 29

The determination of the K_a values of 3HQ **29** was performed according to the general procedure described above.

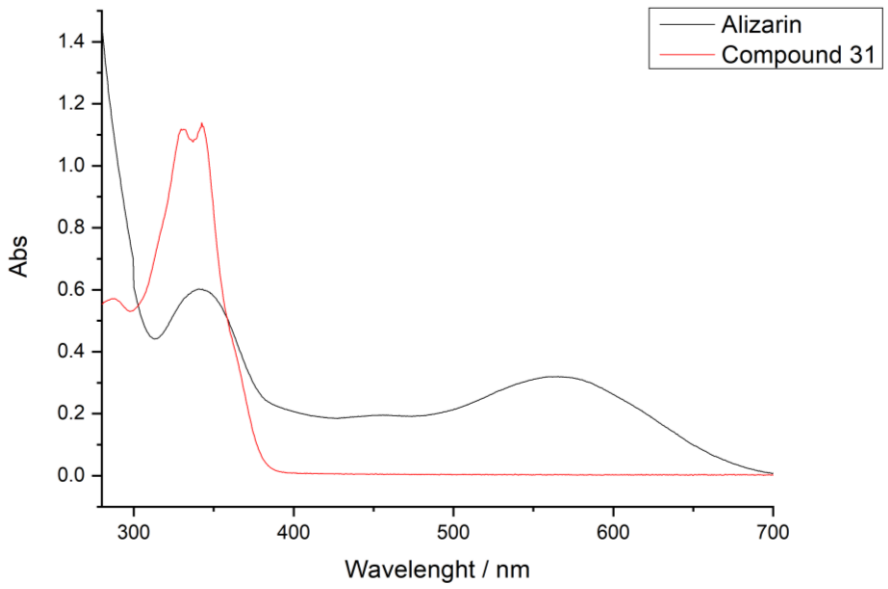


Figure 41 – UV-VIS spectrum of Alizarin and 3HQ 29

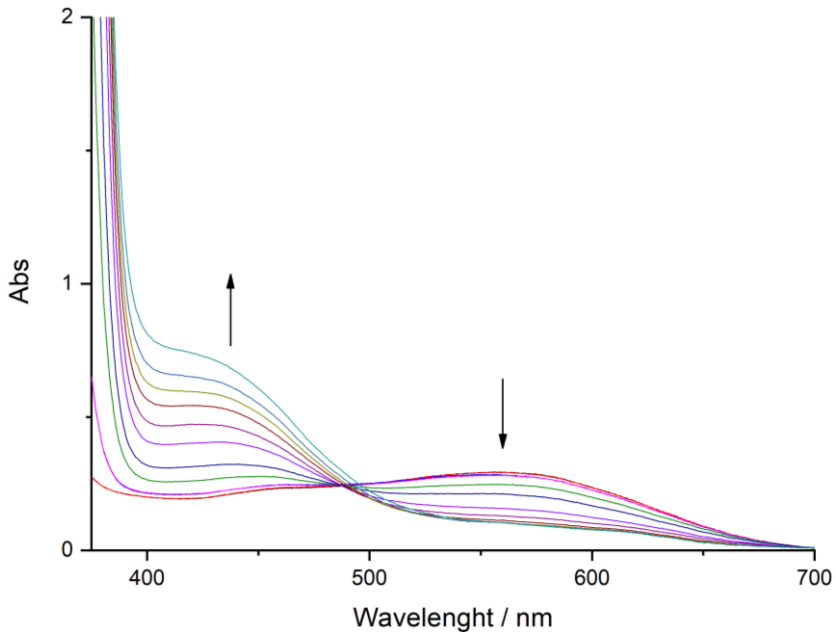


Figure 42 – UV-VIS spectrum of the titration of the PBA-Alizarin system with 3HQ 29

The fluorescence spectrum of the titration was recorded with λ_{exc}= 490 nm, as it is the point of the UV-VIS spectrum where the absorbance suffers the least variations. It has to be noted that, in this point, there is a slight increase in absorption from the reaction mixture, this could influence the titration by artificially augmenting the fluorescence in the final points of the titration, resulting in a minor understatement of the actual K_a value.

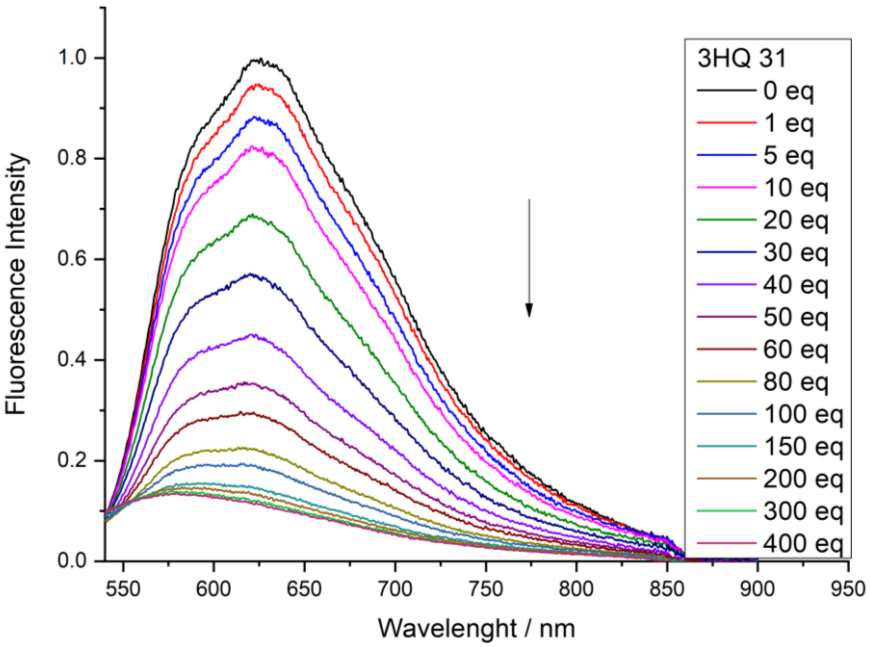


Figure 43 – Fluorescence spectrum of the titration of the PBA-Alizarin system with 3HQ 29

The fluorescence intensity at 626 nm was used to normalize the fluorescence intensity values and perform the K_{exc} calculations , according to the general procedure stated above.

K_{exc}= 0.57064 M^{-1}

K_a= 699 ± 17 M^{-1}

Competitive titration of the AL-PBA system with compound 31

The determination of the K_a values of 3HQ **31** was performed according to the general procedure depicted above.

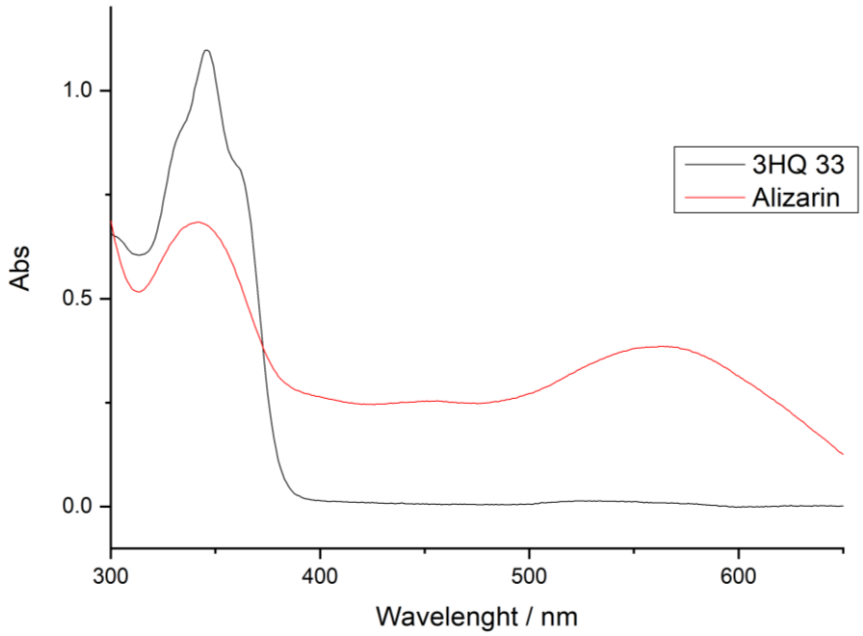


Figure 44 – UV-VIS spectra of 3HQ 31 and Alizarin

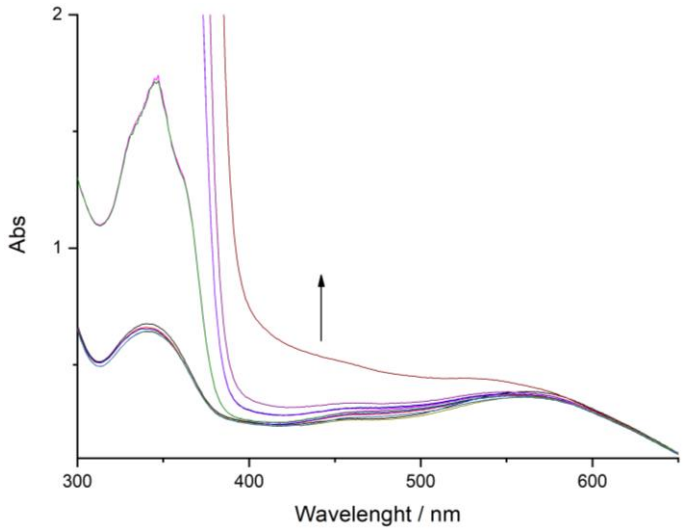


Figure 45 – UV-VIS of the competitive titration of the PBA-Alizarin system with 3HQ 31

The fluorescence spectrum of the titration was recorded with λ_{exc}= 565 nm, as it is the point of the UV-VIS spectrum where the absorbance suffers the least variations. It has to be noted that, in this point, there is a slight increase in absorption from the reaction mixture, this could influence the titration by artificially augmenting the fluorescence in the final points of the titration, resulting in a minor understatement of the actual K_a value.

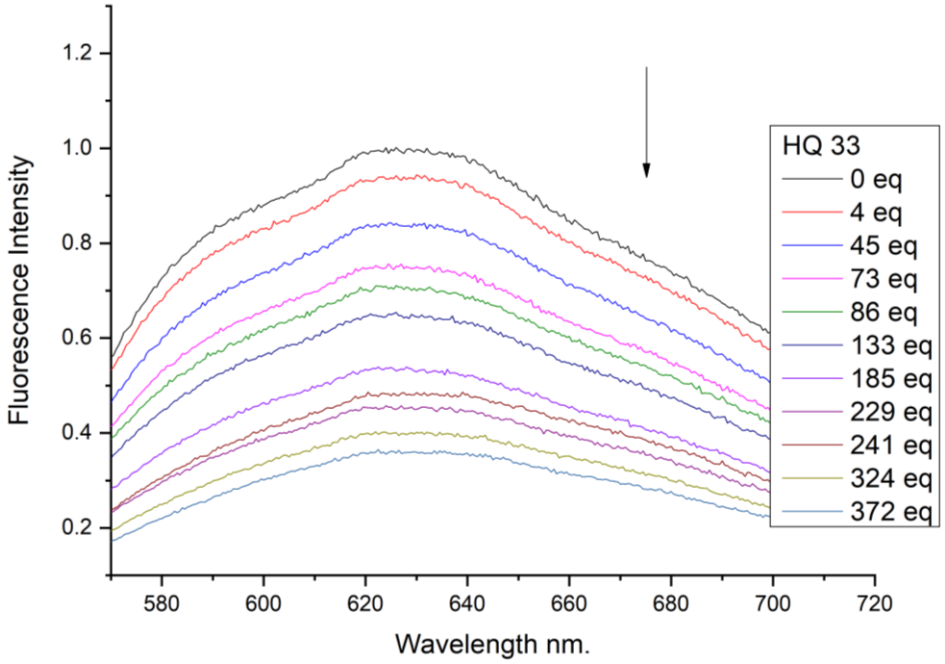


Figure 46 – Fluorescence spectrum of the titration of the PBA-Alizarin system with 3HQ 31

The fluorescence intensity at 620 nm was used to normalize the fluorescence intensity values and perform the K_{exc} calculations , according to the general procedure stated above.

K_{exc}= 0.09609 M^{-1}

K_a= 117.7 ±3 M^{-1}

V.1.5 DFT calculations

All calculations were performed using the GAUSSIAN 09 software package,[154] and the M06-2X functional, without symmetry constraints. That is a hybrid meta-GGA functional developed by Truhlar and Zhao,[155] and it was shown to perform very well for main-group kinetics, providing a good description of long range effects such as van der Waals interactions or π-π stacking. The optimized geometries were obtained with a standard 6-31G(d,p)[156] basis set. The electronic energies obtained at that level (E_{b1}) were converted to free energy at 298.15 K and 1 atm (G_{b1}) by using zero point energy and thermal energy corrections based on structural and vibration frequency data calculated at the same level. Solvent effects (water) were accounted for in all calculations by means of the Polarisable Continuum Model (PCM) initially devised by Tomasi and coworkers[157–159] with radii and non-electrostatic terms of the SMD solvation model, developed by Truhlar *et al*.[160]

Atomic coordinates of the optimized species

H_2O

H	0.309606	0.555609	0.407703
O	0.359183	-0.390182	0.219479
H	-0.355188	-0.554827	-0.408983

PBA

C	0.495678	0.195758	-1.079988
B	1.920883	0.430385	1.091386
O	2.990134	1.076049	0.527941
H	2.883552	1.298814	-0.405375
C	0.618774	0.025896	0.307869
O	1.993812	0.123528	2.424658
H	2.821146	0.424064	2.826629
C	-0.471414	-0.521452	1.002968
H	-0.401456	-0.665465	2.077929
C	-1.643753	-0.882881	0.340846
H	-2.475920	-1.303578	0.897606
C	-1.745225	-0.703662	-1.039421
H	-0.746670	-0.026444	-2.825109
C	-0.672421	-0.165152	-1.750664
H	1.319691	0.609983	-1.657477
H	-2.656209	-0.983941	-1.560198

4-(trifluoromethoxy)-PBA

C	1.052908	0.434380	-1.095336
B	2.481071	0.671136	1.078607
O	3.544989	1.320798	0.515804
F	-3.427083	-1.207841	-0.974282
H	3.441376	1.547686	-0.416894
C	1.174734	0.265214	0.292870
O	2.548400	0.355288	2.407708
F	-2.872377	0.132663	-2.577510
H	3.373969	0.652263	2.816321
C	0.089067	-0.279955	0.991611
F	-2.194671	-1.913601	-2.603103
H	0.158964	-0.422595	2.065919
C	-1.086542	-0.644355	0.337844
H	-1.918328	-1.064440	0.893467
C	-1.176675	-0.460876	-1.039230
H	-0.194308	0.209274	-2.839317
C	-0.111827	0.074933	-1.764510
H	1.874201	0.847648	-1.675460
C	-2.414569	-0.859918	-1.786108

29

N	1.490693	-1.872254	2.300477
O	2.691566	0.065919	2.461740
H	-1.497807	1.609429	0.812917
O	1.040713	1.346138	0.788853
N	-1.764645	1.181524	-0.065468
H	-2.947118	2.819539	-0.519765

C	-1.170467	0.038367	-0.442457	C	0.458738	2.991283	-3.012927
O	-1.341178	-0.458722	-1.561815	**39**			
O	-2.190649	2.900968	-3.209327	N	-1.991067	-3.138758	-0.233750
C	-0.278438	-0.619804	0.569289	C	-0.488712	3.170299	1.688831
O	-0.376066	2.486750	-1.954862	O	-3.367135	-2.340016	1.389019
C	0.774372	0.060457	1.089546	B	-2.335800	1.592954	0.701448
C	1.729675	-0.576487	2.007528	F	3.315282	1.561898	4.086086
H	0.997965	-4.392908	2.886325	N	-2.028701	1.776922	-0.975858
C	0.426284	-2.617439	1.809512	C	-0.960301	1.859024	1.497971
H	-0.929703	-5.733482	2.011019	H	-0.751340	3.453814	-1.042820
C	0.276643	-3.953686	2.203913	O	-2.779940	0.179542	0.747113
F	-2.731114	-4.810067	0.398378	C	-0.996100	0.818180	-1.367422
H	-3.323297	1.369376	-1.456974	O	0.022530	1.206806	-1.883966
C	-0.786112	-4.697206	1.725559	F	3.596514	3.297391	2.829917
H	-2.305584	-2.350310	-0.224731	O	-2.225475	2.407922	-3.639945
C	-1.687084	-4.079316	0.859060	C	-1.331886	-0.569253	-1.012894
H	1.855441	1.599470	1.262160	C	-0.204568	0.812825	2.044019
C	-1.570155	-2.772433	0.450686	F	2.503119	3.476108	4.677501
H	2.145785	-2.323720	2.935553	O	-1.364185	4.459645	-3.306082
C	-2.509246	1.957894	-1.028006	C	-2.208239	-0.765611	0.025421
C	-0.485556	-2.009541	0.925498	H	-0.541996	-0.216113	1.946768
C	-1.681433	2.486007	-2.184330	C	-2.578204	-2.137797	0.459213
H	0.312390	2.395984	-3.915681	C	0.983267	1.049926	2.736652
H	1.478778	2.891484	-2.648019	H	1.550500	0.222656	3.150376
H	0.221406	4.038587	-3.206750	H	-0.897348	-5.114827	-1.564321

C	-1.098167	-2.991034	-1.289002
C	1.427075	2.358923	2.891573
H	0.692557	-4.873597	-3.488918
H	1.042236	4.447354	2.509532
C	-0.592517	-4.134744	-1.918304
C	0.692900	3.427317	2.374786
F	1.480183	-2.594461	-4.438770
H	-1.052945	4.016462	1.301115
C	2.701864	2.661625	3.618016
H	-2.491591	3.811848	-0.957294
C	0.282296	-4.007327	-2.981965
H	0.461527	-0.612832	-3.174772
C	0.626380	-2.722782	-3.393742
C	0.150335	-1.577948	-2.797365
H	-2.237515	-4.084282	0.052416
C	-1.738024	3.153906	-1.398150
C	-0.735024	-1.698549	-1.709349
C	-1.801997	3.275889	-2.902955
H	-0.835925	3.978543	-5.261600
H	-1.024734	5.713430	-4.857212
H	-2.473389	4.676824	-5.053588
C	-1.434307	4.714913	-4.723139
O	-3.444700	2.413884	0.995549
H	-2.918321	1.477116	-1.403000
H	-3.205214	3.229285	1.451861

36

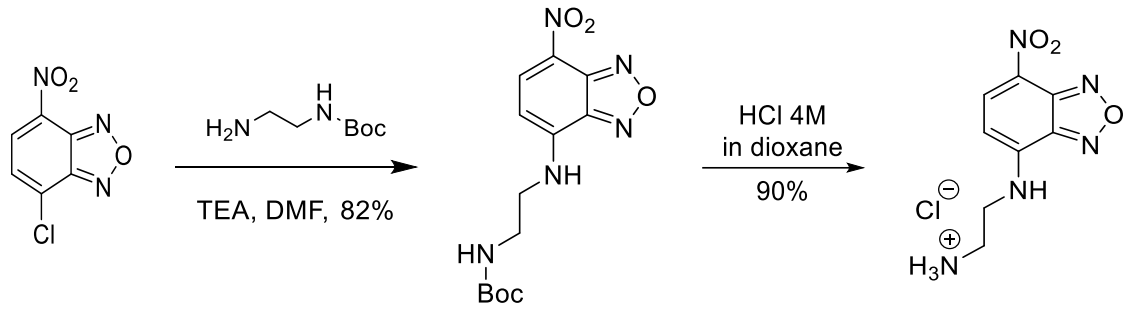
N	-0.980713	-2.465365	-0.075060
C	1.605544	0.857592	4.063314
H	-0.672264	-3.394137	0.215009
H	-3.377791	1.736894	0.339499
O	-0.313059	-1.505249	1.921696
F	5.824523	2.169654	2.325740
B	-0.343979	-0.036842	2.543959
N	-2.993428	2.176748	-0.488034
C	1.164481	0.442847	2.800621
O	-1.146986	-0.018948	3.707028
H	-2.630453	4.196406	-0.952674
H	-2.055715	-0.286531	3.514903
O	-0.985323	0.712169	1.374878
C	-2.167065	1.489264	-1.288728
O	-1.736291	1.958106	-2.349807
F	5.357586	2.627351	4.387212
O	-4.155439	4.461456	-2.975272
C	-1.798179	0.107921	-0.845648
C	2.087600	0.459858	1.744716
F	6.001289	0.642790	3.840316
O	-4.719357	2.329609	-2.555708
C	-1.254399	-0.121368	0.379895
H	1.773485	0.145439	0.750613
C	-0.831804	-1.450532	0.750599

C	3.403869	0.871919	1.932057	H	-4.178074	3.844800	-0.162140
H	4.104356	0.879098	1.102662	C	-2.297236	-3.283670	-3.403328
H	-1.374279	-4.394340	-1.809899	H	-2.929614	0.069281	-3.374662
C	-1.555059	-2.302917	-1.337340	C	-2.719585	-2.011061	-3.791462
C	3.810471	1.281678	3.202299	C	-2.575137	-0.889582	-3.012556
H	-2.436301	-4.129469	-4.067273	C	-3.456044	3.482902	-0.897065
H	3.244132	1.596499	5.258349	C	-1.969769	-1.016310	-1.744339
C	-1.712035	-3.424170	-2.160607	C	-4.131695	3.480930	-2.255225
C	2.918639	1.276044	4.273158	H	-6.134846	2.989045	-3.934036
F	-3.306125	-1.881045	-5.003345	H	-5.739147	1.241369	-3.921357
H	0.914414	0.858356	4.901458	H	-4.584855	2.416500	-4.626766
C	5.234762	1.682292	3.433832	C	-5.336307	2.249715	-3.852584

V.2. Experimental section of Chapter III

V.2.2. Synthesis and structural characterization

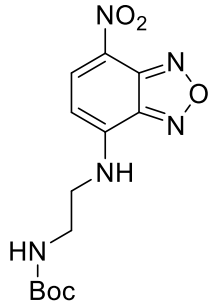
- **Synthesis of ehylenodiamine-NBD 93**



92 **93**

Scheme 67 – Synthesis of ehylenodiamine-NBD 93

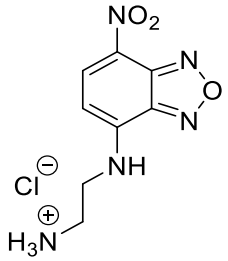
Compound **93** was prepared by attaching Boc-ethylenodiamine to commercially available NBD-chloride, followed by deprotection in acidic medium.



92

tert-butyl (2-((7-nitrobenzo[c][1,2,5]oxadiazol-4-yl)amino)ethyl)carbamate 92. Compound **92** was prepared according to literature procedures.[161] In a flame-dried schlenk tube, 4-chloro-7-nitrobenzo[c][1,2,5]oxadiazole (500 mg, 2.506 mmol) was dissolved in dry DMF (10 ml), then tert-butyl (2-aminoethyl)carbamate (435 µL, 3.01 mmol) was added, followed by the addition of freshly distilled triethylamine (349 µL, 2.506 mmol). the reaction was stirred at room temperature under argon atmosphere for 3.5 hours in the dark. Subsequently, the crude was dissolved in NH4Cl saturated solution, then extracted 3 times with EtOAc. the resulting organic phase was washed with water and brine, then dried over sodium sulphate, filtered and concentrated at reduced pressure. No further purification was needed as the 1H-NMR spectrum indicates the good purity of the sample, in line with what reported in the literature. Yield: 82%

^{1}H NMR (300 MHz, $CDCl_3$): δ 8.48 (d, J = 8.7 Hz, 1H), 8.01 (br s, 1H),6.16 (d, J = 8.6 Hz, 1H), 5.09 (br s, 1H), 3.59 (br s, 4H), 1.46 (s, 9H).



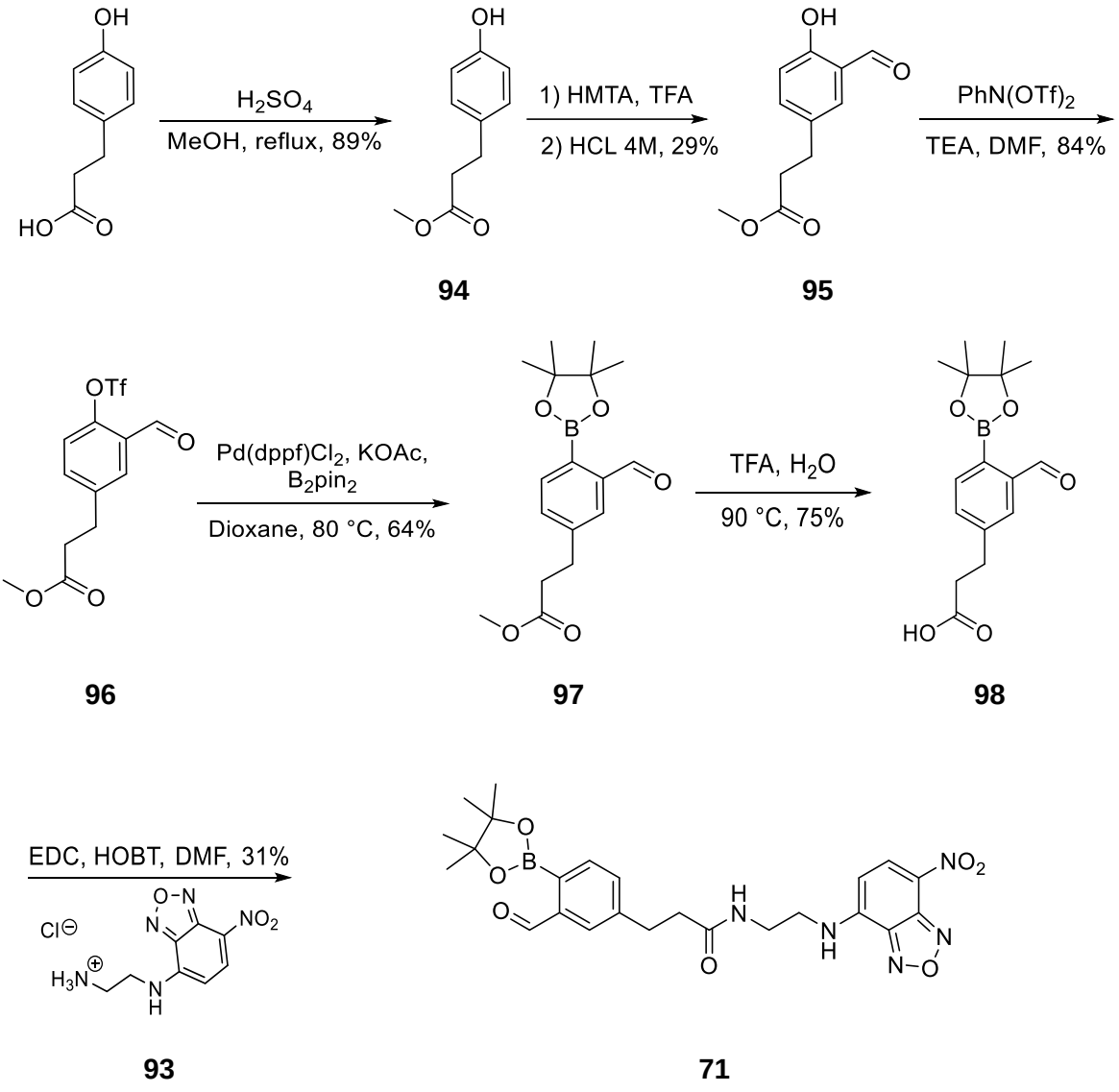
93

N1-(7-nitrobenzo[c][1,2,5]oxadiazol-4-yl)ethane-1,2-diamine 93. Compound **93.** was prepared following literature procedures.[161] Compound **92** (100 mg, 0.309 mmol) was dissolved in HCl 4M solution in dioxane (2.320 ml) and stirred overnight under argon

atmosphere in the dark. The reaction mixture was dried at reduced pressure, yielding a dark brown oil, to which MTBE was added, causing the precipitattion of a brown solid, which was collected after centrifugation and washing with MTBE and dried at high vacuum. No further purification was needed as the 1H-NMR spectrum indicates the good purity of the sample, in line with what reported in the literature. Yield: 90%

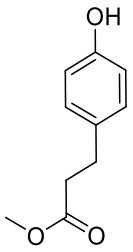
^{1}H NMR (300 MHz, $(CD_3)_2SO$): δ 9.36 (bs, 1H), 8.56 (d, J = 8.8 Hz, 1H), 8.15 (bs, 3H), 6.54 (d, J = 9.0 Hz, 1H), 3.80 (m, 2H), 3.14 (q, J = 5.9 Hz, 2H).

- **Synthesis of 2FBBA-NBD probe 71**



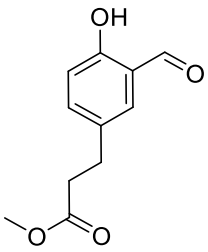
Scheme 68 – Synthetic route towards 2FBBA-NBD probe 71

Compound **71** was prepared by coupling carboxylic acid-functionalized 2-FBBA **98** with the previously synthesized amino-NBD compound **93**.



94

Methyl 3-(4-hydroxyphenyl)propanoate 94. 3-(4-hydroxyphenyl)propanoic acid (10 g, 60.2 mmol) was dissolved in MeOH (100ml). Concentrated H_2SO_4 (32 µl, 0.600 mmol) was then added and the reaction mixture was heated to reflux for 2 h. After the elapsed time, the reaction mixture was concentrated under vacuo and the residue was dissolved in EtOAc and washed with NaHCO3 saturated solution, brine and then dried over anhydrous Na2SO4, filtered and concentrated under vacuo providing the methyl 3-(4-hydroxyphenyl)propanoate as a yellow oil, with ^{1}H-NMR spectrum in according to what reported in the literature.[162] Yield 89%. **^{1}H NMR (300 MHz, ($CDCl_3$):** δ 7.08 – 7.02 (m, 2H), 6.79 – 6.69 (m, 2H), 3.67 (s, 3H), 2.88 (t, J = 7.8 Hz, 2H), 2.60 (t, J = 7.8 Hz, 2H).

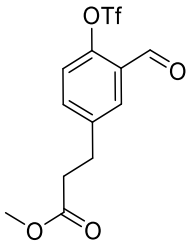


95

Methyl 3-(3-formyl-4-hydroxyphenyl)propanoate 95. Compound **94** (9.6 g, 53.3 mmol) and hexamethylenetetramine (9.54 g, 68.0 mmol) were refluxed in trifluoroacetic acid (76 ml) under Argon atmosphere for 18 hours. The solution was poured into aqueous HCl 4 M (76 ml). After stirring at room temperature for 3 hours, the mixture was neutralized with Na_2CO_3 saturated solution and subsequently extracted with CH_2Cl_2. After drying over Na_2SO_4, the solvent was removed and the crude was purified by column chromatography

on silica gel with a 4:1 Hexane:EtOAc mixture as eluent. The pure product appears as a yellow oil, obtained with 29% and ^{1}H-NMR spectrum in accord with what reported in the literature.[163]

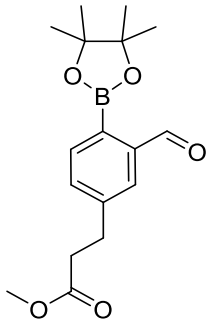
1H NMR (300 MHz, ($CDCl_3$): δ 10.87 (s, 1H), 9.86 (bs, 1H), 7.46 – 7.34 (m, 2H), 6.92 (d, J = 9.3 Hz, 1H), 3.66 (s, 3H), 2.93 (t, J = 7.5 Hz, 2H), 2.62 (t, J = 7.5 Hz, 2H).



96

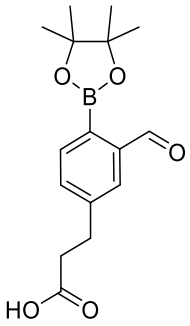
Methyl 3-(3-formyl-4-(((trifluoromethyl)sulfonyl)oxy)phenyl)propanoate 96. Compound **96** was prepared following an already reported procedure.[104] To a solution of methyl 3-(3-formyl-4-hydroxyphenyl)propanoate **94** (1 g, 4.80 mmol) in DMF (11.52 ml), triethylamine (2 ml, 14.41 mmol) was added and the solution stirred for 1 h. Then, N-Phenyltrifluoromethanesulfonimide (2.75 g, 7.68 mmol) was added portionwise and the reaction stirred for another 3 h. Water was then added and the mixture extracted with MTBE. The organic layer was washed with water, LiCl (aq), brine, dried over Na_2SO_4, concentrated and purified by column chromatography on silica gel, with a mixture of hexane:MTBE 7:3 as eluent to give the desired product as a yellow oil. Yield: 84%.

^{1}H NMR (300 MHz, ($CDCl_3$): δ 10.20 (d, J = 0.6 Hz, 1H), 7.80 (d, J = 2.4 Hz, 1H), 7.55 (dd, J = 8.5, 2.4 Hz, 1H), 7.30 (d, J = 8.5 Hz, 1H), 3.64 (s, 3H), 3.02 (t, J = 7.5 Hz, 2H), 2.66 (t, J = 7.5 Hz, 2H). **^{13}C NMR (75 MHz, $CDCl_3$):** δ 186.69, 172.62, 148.35, 142.09, 136.00, 130.56, 128.37, 122.60, 118.68 (q, J = 320.6 Hz), 51.89, 34.86, 29.99. **LRMS (m/z):** Calculated for $[M-H]^-$ $C_{12}H_{10}F_3O_6S$: 339.0, found: 339.0. **Elemental Anaysis**: Calculated for $C_{12}H_{11}F_3O_6S$: C, 42.36; H, 3.26; S, 9.42. Found: C, 42.32; H, 3.25; S, 9.23.



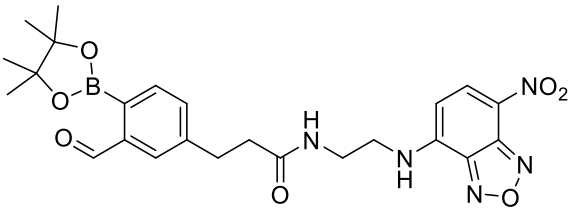
97

Methyl 3-(3-formyl-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl) propanoate 97. Compound **97** was prepared following an already reported procedure.[164] To a solution of bis(pinacolato)diborane (0.970 g, 3.82 mmol) in dry dioxane (10.08 ml) under Ar, was added dry potassium acetate (0.865 g, 8.82 mmol). After degassing for 15 min, $Pd(dppf)Cl_2 \cdot CH_2Cl_2$ (0.240 g, 0.294 mmol) and methyl 3-(3-formyl-4-(((trifluoromethyl)sulfonyl)oxy)phenyl) propanoate (1 g, 2.94 mmol) were added to the reaction mixture. The mixture was stirred at 80 °C for 30 min. The reaction was quenched by adding ice-water (6.72 ml). The resulting mixture was extract with 50% EtOAc/hexanes. The extract was washed with brine, dried and concentrated to dryness. The residue was purified by chromatography on silica gel with a 3:7 EtOAc:Hexane mixture as eluent to give the desired product as a yellow oil with 64% yield. **^{1}H NMR (300 MHz, $CDCl_3$):** δ 10.57 (s, 1H), 7.84 – 7.77 (m, 2H), 7.44 (dd, J = 7.7, 1.8 Hz, 1H), 3.66 (s, 3H), 3.02 (t, J = 7.7 Hz, 2H), 2.65 (t, J = 7.7 Hz, 2H), 1.38 (s, 12H). **^{13}C NMR (75 MHz, $CDCl_3$):** 194.85, 172.95, 143.73, 141.69, 136.15, 133.16, 127.32, 84.40, 51.77, 35.18, 30.74, 24.91. The signal corresponding to the carbon directly bound to boron is not visible due to the quadrupolar relaxation of the B nucleus. **LRMS (m/z)**: Calculated for $[M+H]^+$ C17H23BO5: 319.2, found: 319.3.



98

3-(3-formyl-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)propanoic acid 98. Compound **98** was prepared following a procedure already reported by our group.[104] In a round bottom flask, under inert atmosphere, compound **97** (382 mg, 1.201 mmol) was stirred in a solution of trifluoroacetic acid (1.850 ml, 24.01 mmol) and Water (8 ml) at 90 °C for 2h, after which the solvent was concentrated in vacuo. The crude was suspended toluene (15 mL) and evaporated until dryness for 3 times. Then the residue was suspended in Hexane, triturated and decanted. The precipitate was washed 3 times with hexane, and the filtrate was evaporated until dryness, affording the desired product as a white solid with 75% yield. **^{1}H NMR (300 MHz, $CDCl_3$):** δ 10.82 (bs, 1H), 10.55 (s, 1H), 7.85 – 7.78 (m, 2H), 7.44 (dd, J = 7.7, 1.8 Hz, 1H), 3.01 (t, J = 7.7 Hz, 2H), 2.69 (t, J = 7.6 Hz, 2H), 1.37 (s, 12H). **^{13}C NMR (75 MHz, $CDCl_3$):** δ 194.94, 178.46, 143.45, 141.74, 136.22, 133.19, 127.41, 84.48, 35.11, 30.43, 24.95. The signal corresponding to the carbon directly bound to boron is not visible due to the quadrupolar relaxation of the B nucleus. **LRMS (m/z):** calculated for $[M+H]^+$ C16H22BO5: 305.2, found: 305.6. **Elemental Analysis (%):** Calculated for C16H21BO5: C, 63.18; H, 6.96; B, 3.55; O, 26.30. Found: C, 63.06; H, 7.01.



71

3-(3-formyl-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-N-(2-((7-nitrobenzo[c][1,2,5]oxadiazol-4-yl)amino)ethyl)propanamide 71. Compound **71** was prepared adapting a procedure reported in literature.[161] In a flame-dried Schlenk tube,

under argon atmosphere, compound 31 (63 mg, 0.207 mmol) was dissolved in dry DMF (1 mL) at 0 °C and triethylamine (57.7 µl, 0.414 mmol) was added dropwise. Subsequently, HOBT (38.1 mg, 0.249 mmol), EDC (47.7 mg, 0.249 mmol) and compound 26 (81 mg, 0.311 mmol) were added and the reaction was stirred at room temperature in the dark, periodically checking by TLC (DCM:MeOH 95:5). After 5h the reaction was complete and the mixture was poured into an aqueous HCl solution (0.1 M, 10 mL) and extracted with CH2Cl2. The combined organic phase was washed with saturated Na_2CO_3 and brine then dried over magnesium sulphate. The organic phase was purified via column chromatography on silica gel (eluent: DCM:MeOH 95:5) with no success, as the product appeared to degradate in silica. Other procedures, such as preparative TLC, column chromatography on alumina or precipitation never afforded the pure product, similarly to what observed by other groups preparing NBD-based probes.[161] The product was dissolved in acetonitrile and isolated through semi-preparative HPLC on C-18 reverse phase column. Eluent mixture ACN:H_2O starting with a 1:9 ratio and reaching 9:1 ratio after 25 minutes of run. Total run time: 30 min. The peak corresponding to the product was detected at 23.1 minutes, the resulting fractions were concentrated at reduced pressure. The isolated compound resulted to be the pinacol hemiacetal derivative of the desired product, which was used for the preparation of the Laminin conjugate, as both pinacol groups of the molecule resulted to be hydrolized in conjugation conditions. Yield: 31%

^{1}H NMR (300 MHz, $CDCl_3$): δ 8.43 (d, J = 8.5 Hz, 1H), 7.80 (s, 1H), 7.56 – 7.47 (m, 2H), 7.09 (dd, J = 7.6, 1.7 Hz, 1H), 6.49 (s, 1H), 6.05 (d, J = 8.5 Hz, 1H), 6.02 (bs, 1H, overlapped with the doublet at 6.05) 3.61 – 3.52 (m, 2H), 3.48 (m, 2H), 3.01 (t, J = 7.1 Hz, 2H), 2.53 (t, J = 7.1 Hz, 2H), 1.32 (s, 12H), 1.30 (s, 6H), 1.23 (s, 6H). **^{13}C NMR (75 MHz, $CDCl_3$)**: δ 195.18, 174.78, 145.13, 144.20, 142.86, 136.58, 136.22, 135.35, 133.63, 127.95, 126.44, 124.87, 123.51, 98.14, 84.47, 83.75, 82.68, 75.07, 24.86, 24.54, 22.21. The signal corresponding to the carbon directly bound to boron is not visible due to the quadrupolar relaxation of the B nucleus. **HRMS (m/z):** Calculated for $[M+Na]^+$ C30H40BN5NaO8: 632.2868, found: 632.2861.

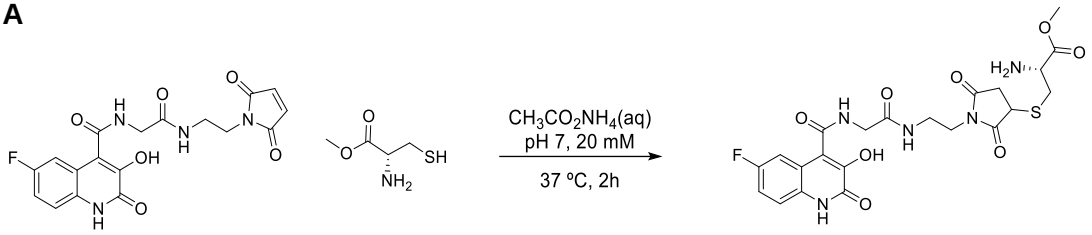
V.2.3. ESI-MS experiments

3HQ BHS **42** (100 mM in DMF) was diluted to 200µM in ammonium acetate 20 mM pH 7 end stirred for 1h with 1eq of either CysOMe or *N*-Ac-Cys (both 50 mM in water).

Subsequently, 100 eq of the appropriate BA (100 mM in DMF) and the mixture was stirred for 2h at 37°C before recording the ESI-MS spectrum of the reaction.

- **Reaction of BHS 42 with Cysteine methyl ester**

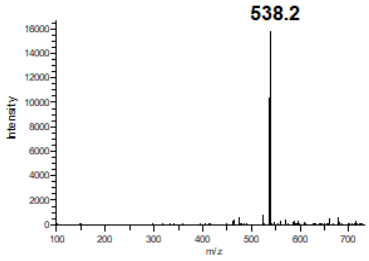
A



42 **CysOMe** **50**

m/z $[M+H]^+$= 403.1 *m/z* $[M+H]^+$= 538.2

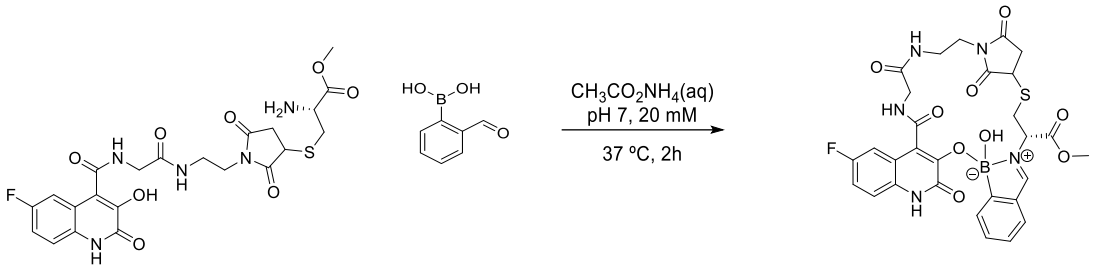
B



Scheme 69 - Reaction of 3HQ boron hotspot 42 with cysteine methyl ester in bioconjugation conditions (A) and relative ESI-MS spectrum (B)

- **Reaction of BHS-CysOMe adduct 50 with CBBAs**

A



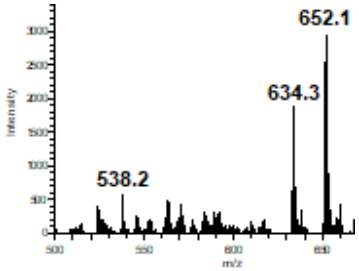
50 **2-FBBA** **51**

m/z $[M+H]^+$= 538.2

m/z $[M+H]^+$ = 652.2

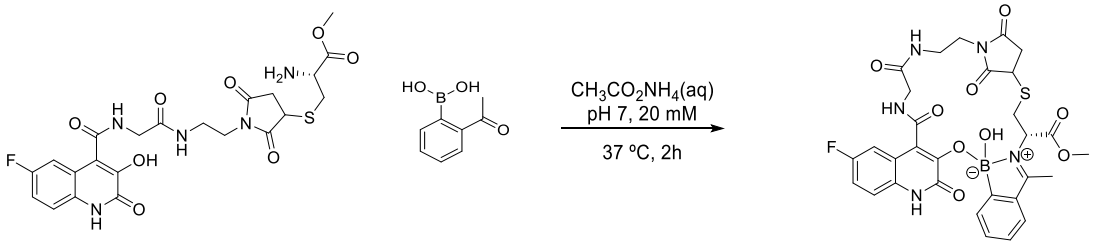
m/z $[M-OH]^+$ = 634.3

B



Scheme 70 – Reaction of BHS-CysOMe adduct 50 with 2-FBBA in bioconjugation conditions (A) and relative ESI-MS spectrum (B)

A



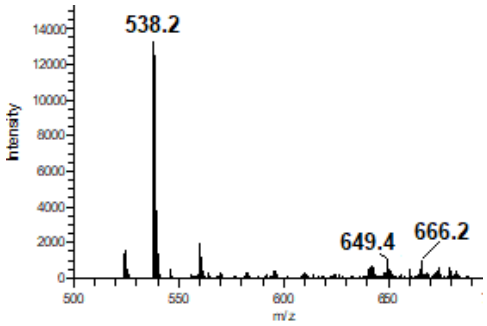
50 **2-ABBA** **52**

m/z [M+H]$^+$= 538.2

m/z [M+H]$^+$ = 666.2

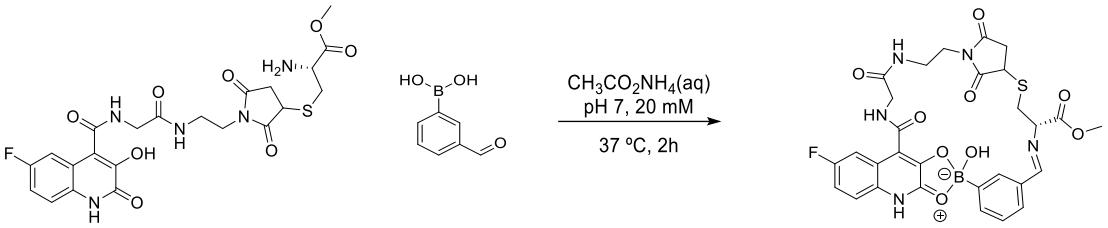
m/z [M-OH]$^+$ = 649.4

B



Scheme 71 – Reaction of BHS-CysOMe adduct 50 with 2-ABBA in bioconjugation conditions (A) and relative ESI-MS spectrum (B)

A



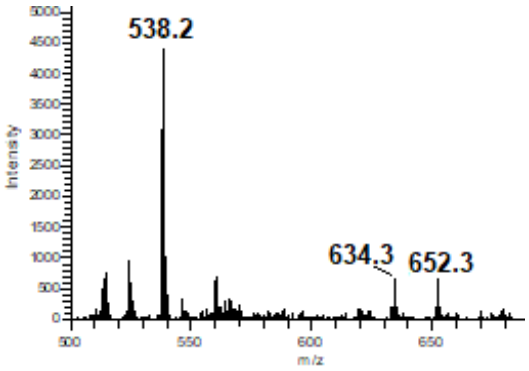
50 **3-FBBA** **53**

m/z [M+H]$^+$= 538.2

m/z [M+H]$^+$ = 652.3

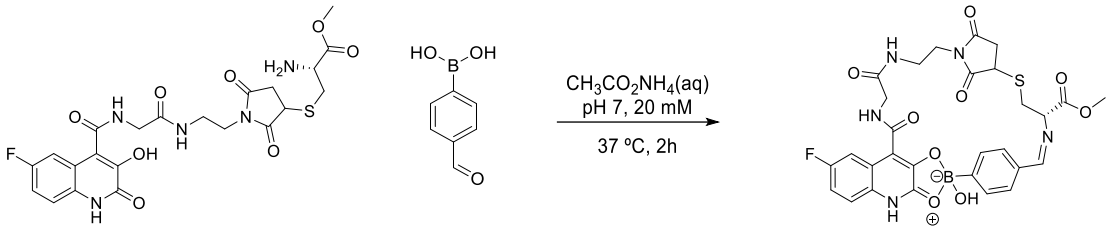
m/z [M-OH]$^+$ = 634.3

B



Scheme 72 – Reaction of BHS-CysOMe adduct 50 with 3-FBBA in bioconjugation conditions (A) and relative ESI-MS spectrum (B)

A



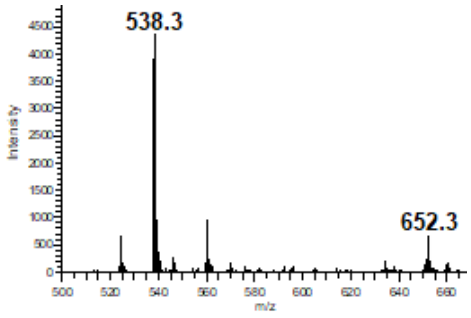
50 **4-FBBA** **54**

m/z $[M+H]^+$= 538.2

m/z $[M+H]^+$ = 652.3

m/z $[M-OH]^+$ = 634.3

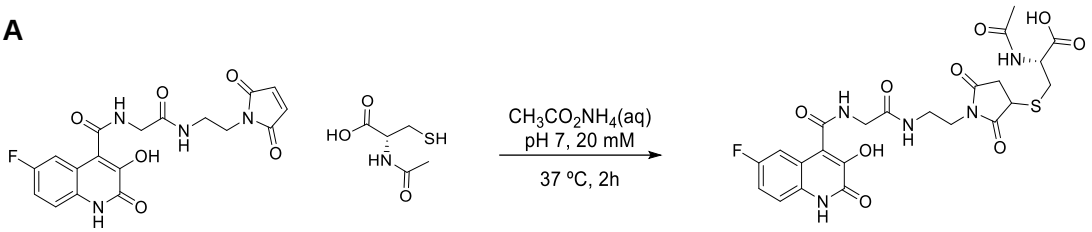
B



Scheme 73 – Reaction of BHS-CysOMe adduct 50 with 4-FBBA in bioconjugation conditions (A) and relative ESI-MS spectrum (B)

- **Reaction of BHS 42 with *N*-Acetyl Cysteine and subsequent reaction with 2-FBBA**

A



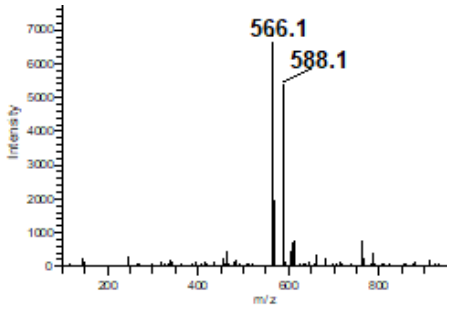
42 **N-Ac Cys** **50b**

m/z $[M+H]^+$= 403.1

m/z $[M+H]^+$= 566.1

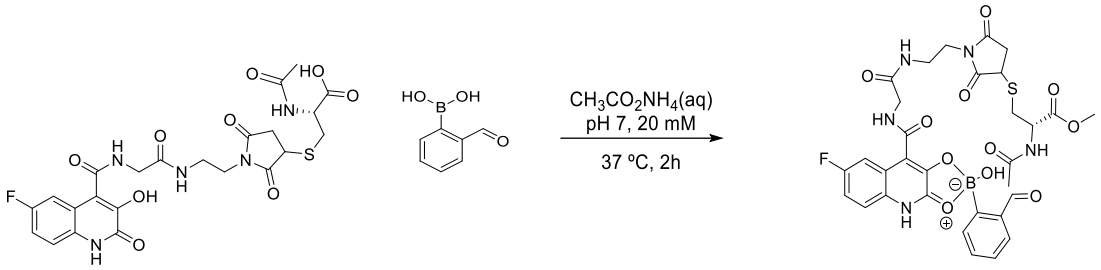
m/z $[M+Na]^+$= 588.1

B



Scheme 74 - Reaction of 3HQ boron hotspot 42 with *N*-Acetyl cysteine in bioconjugation conditions (A) and relative ESI-MS spectrum (B)

A



50b **2-FBBA** **55**

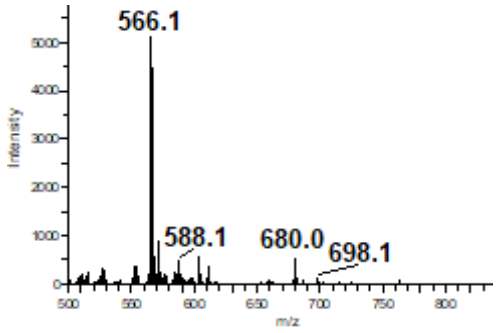
m/z $[M+H]^+$= 566.1

m/z $[M+Na]^+$= 588.1

m/z $[M+H]^+$ = 698.1

m/z $[M-OH]^+$ = 680.0

B

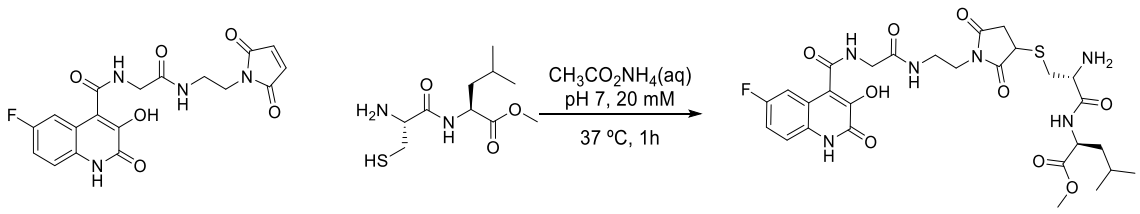


Scheme 75 – Reaction of BHS-*N*-Acetyl Cys adduct 50b with 2-FBBA in bioconjugation conditions (A) and relative ESI-MS spectrum (B)

- **Reactions of BHS 42 with *N*-terminal Cysteine dipeptides and 2-FBBA**

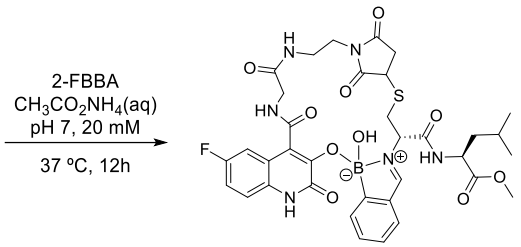
BHS **42** (100 mM in DMF) was diluted to 200µM in ammonium acetate 20 mM pH 7 end stirred for 1h with 1eq of the appropriate dipeptide (50 mM in water). Subsequently, 20 eq of 2-FBBA (100 mM in DMF) and the mixture was stirred for 12h at 37°C before recording the ESI-MS spectrum of the reaction.

A



42 **58** **58a**

m/z $[M+H]^+$= 403.1 *m/z* $[M+H]^+$= 651.2

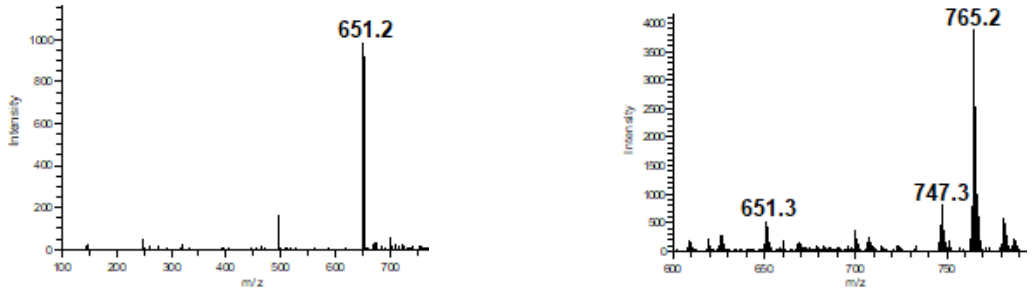


58b

m/z $[M+H]^+$ = 765.2

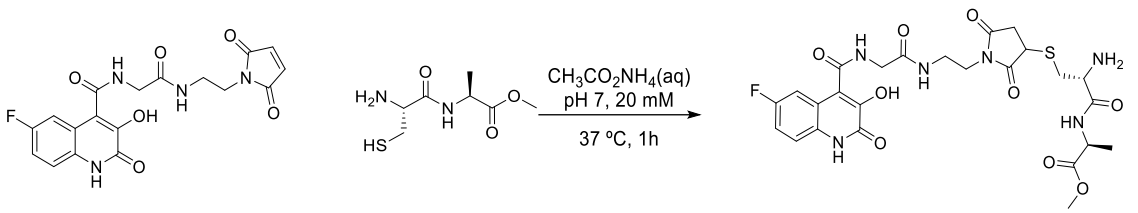
m/z $[M-OH]^+$ = 747.3

B **C**



Scheme 76 Functionalization of dipeptide 58 with BHS 42 and subsequent reaction with 2-FBBA in bioconjugation conditions (A) along with ESI-MS spectrum of the first step of the reaction (B) and of 3HQ-iminoboronate formation (C)

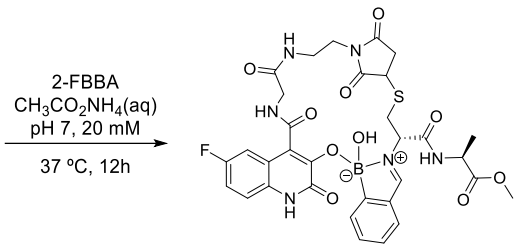
A



42 **59** **59a**

m/z $[M+H]^+$= 403.1

m/z $[M+H]^+$= 609.2

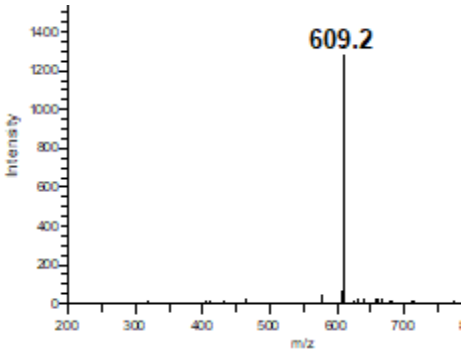


59b

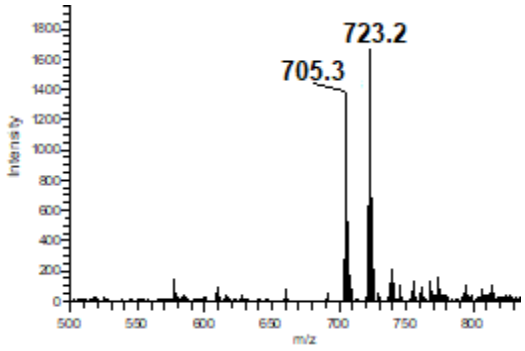
m/z $[M+H]^+$ = 732.2

m/z $[M-OH]^+$ = 705.3

B

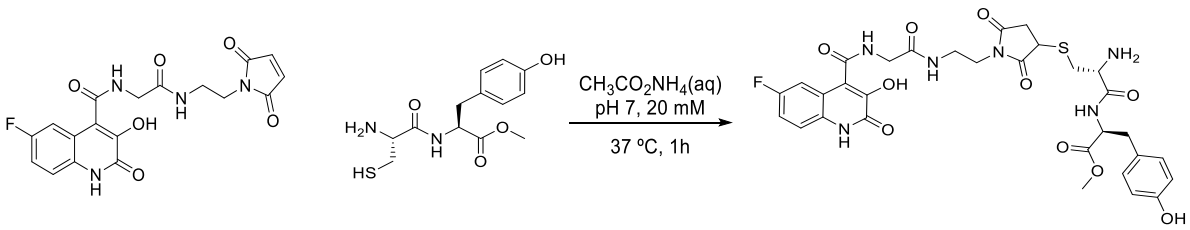


C



Scheme 77 Functionalization of dipeptide 56 with BHS 42 and subsequent reaction with 2-FBBA in bioconjugation conditions (A) along with ESI-MS spectrum of the first step of the reaction (B) and of 3HQ-iminoboronate formation (C)

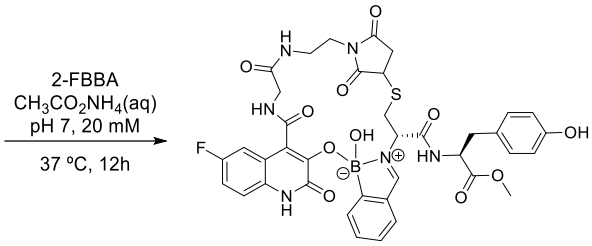
A



42 **60** **60a**

m/z $[M+H]^+$= 403.1

m/z $[M+H]^+$= 701.2

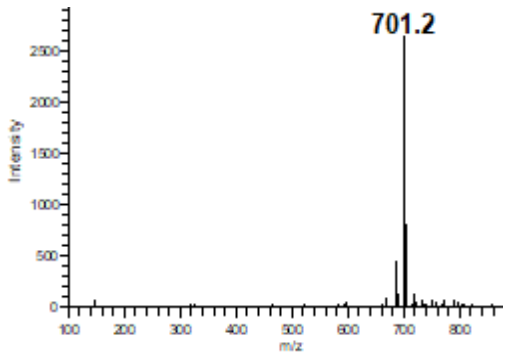


60b

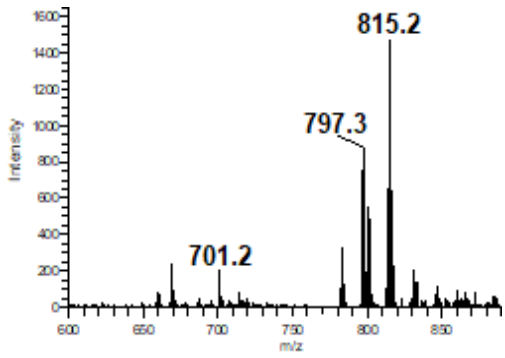
m/z $[M+H]^+$ = 815.2

m/z $[M-OH]^+$ = 797.3

B

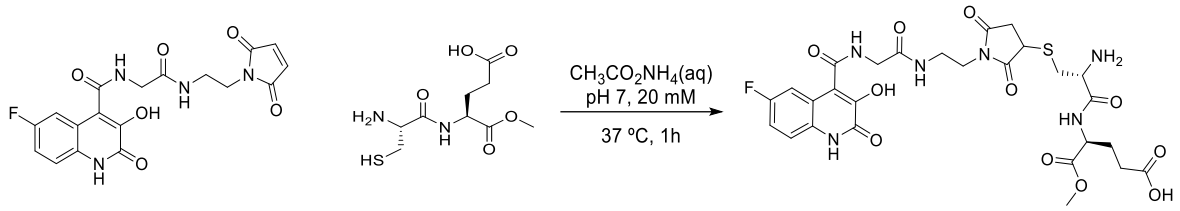


C



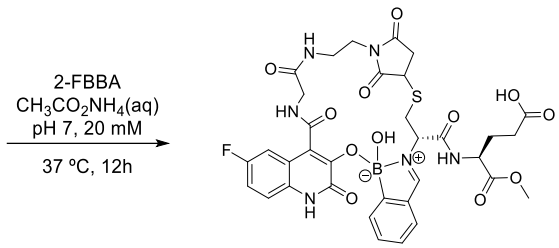
Scheme 78 Functionalization of dipeptide 60 with BHS 42 and subsequent reaction with 2-FBBA in bioconjugation conditions (A) along with ESI-MS spectrum of the first step of the reaction (B) and of 3HQ-iminoboronate formation (C)

A



42 **61** **61a**

m/z $[M+H]^+$= 403.1 *m/z* $[M+H]^+$= 667.2

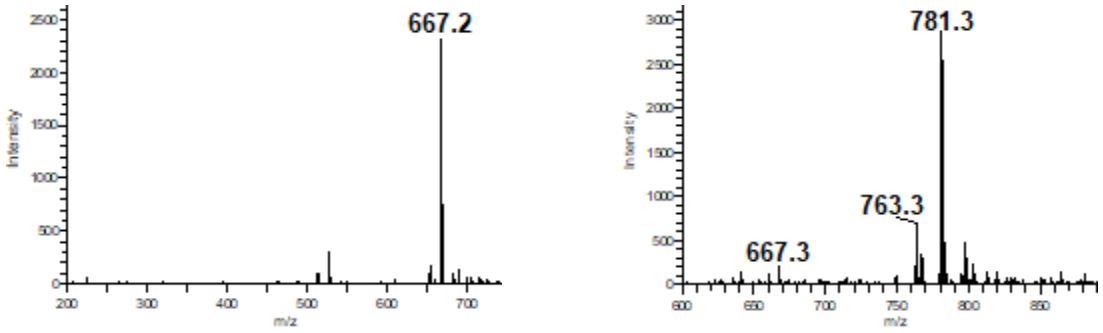


61b

m/z $[M+H]^+$ = 781.3

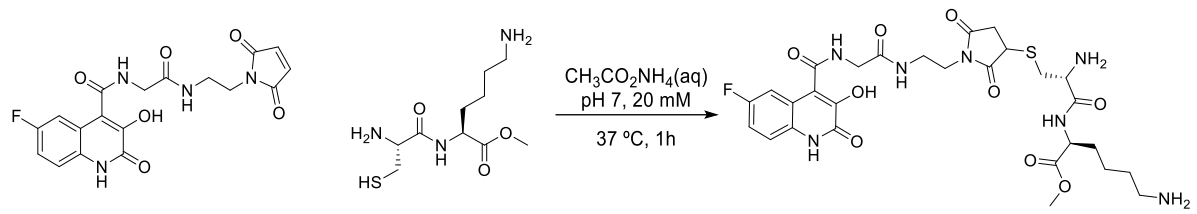
m/z $[M-OH]^+$ = 763.3

B **C**



Scheme 79 Functionalization of dipeptide 61 with BHS 42 and subsequent reaction with 2-FBBA in bioconjugation conditions (A) along with ESI-MS spectrum of the first step of the reaction (B) and of 3HQ-iminoboronate formation (C)

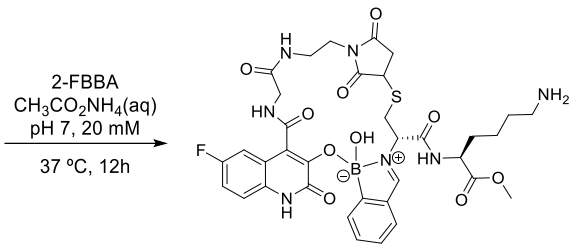
A



42 **62** **62a**

m/z $[M+H]^+$= 403.1

m/z $[M+H]^+$= 666.3

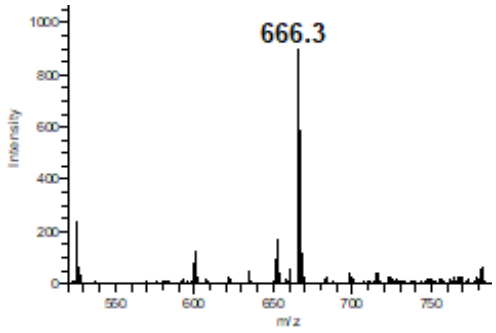


62b

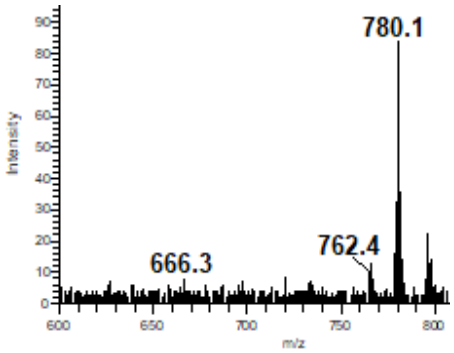
m/z $[M+H]^+$ = 780.1

m/z $[M-OH]^+$ = 762.4

B



C

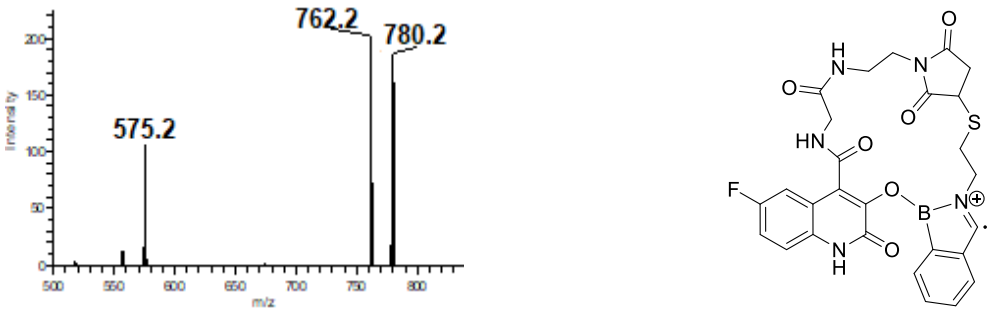


Scheme 80 Functionalization of dipeptide 62 with BHS 42 and subsequent reaction with 2-FBBA in bioconjugation conditions (A) along with ESI-MS spectrum of the first step of the reaction (B) and of 3HQ-iminoboronate formation (C)

- **MS-MS fragmentation of 62b**

The peak at *m/z* = 780.1 was subject to MS-MS fragmentation, yielding the following spectrum:

A **B**

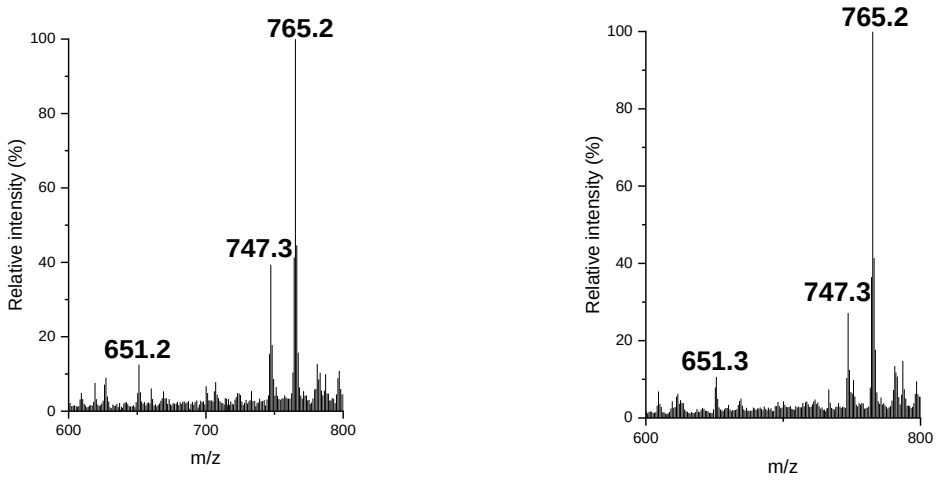


Scheme 81 – MS-MS spectrum of the fragmentation of 62b (A) and proposed structure of the fragment obtained (B)

- **Stability of 58b at pH 7**

A 200 µM solution of adduct **58b**, resulting from the previously reported reaction, was stirred at 37°C and monitored by ESI-MS after 4 weeks.

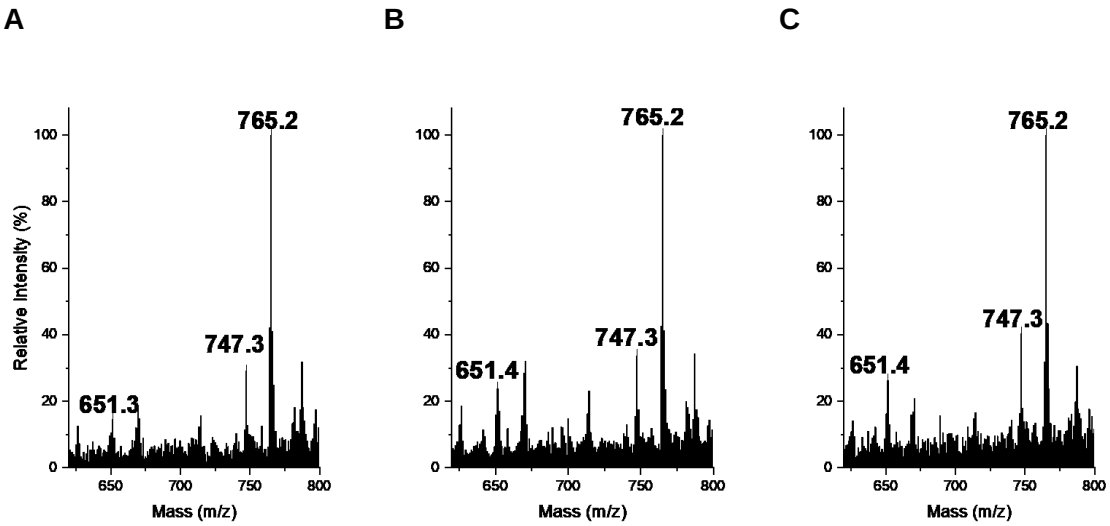
A **B**



Scheme 82 – ESI-MS spectra of 58b at pH 7 after 12h (A) and 4 weekes (B)

- **Stability of of 58b at pH 4.5**

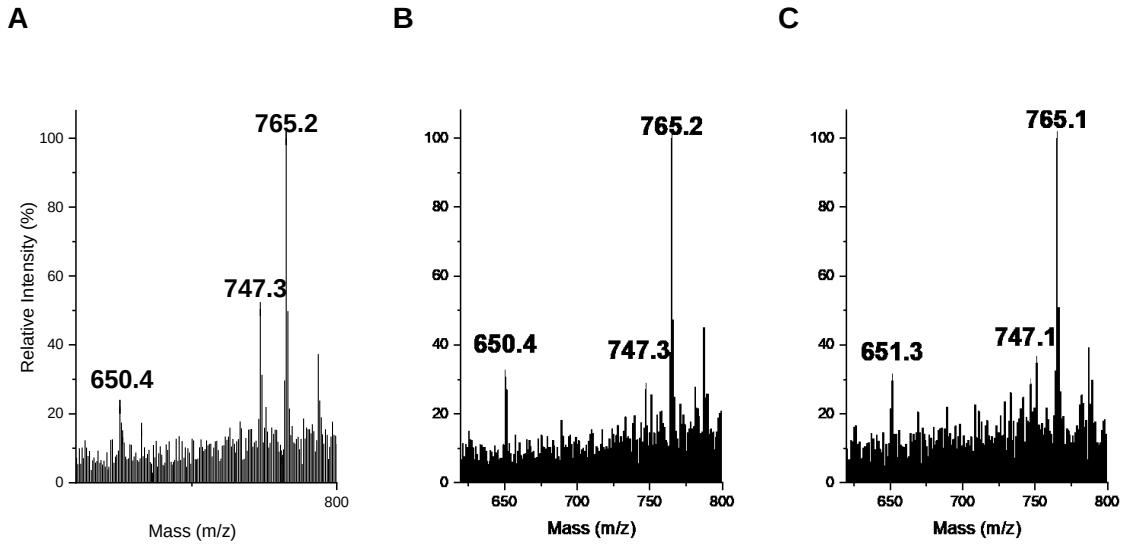
The pH of a 200 µM solution of adduct **58b**, resulting from the previously reported reaction, was adjusted to 4.5 with a 100 µM solution of acetic acid and stirred at 37°C and monitored by ESI-MS.



Scheme 83 – ESI-MS spectra of 58b at pH 4.5 after 8h (A), 24h (B) and 32h (C)

- **Stability of 58b in the presence of 10 eq of glucose**

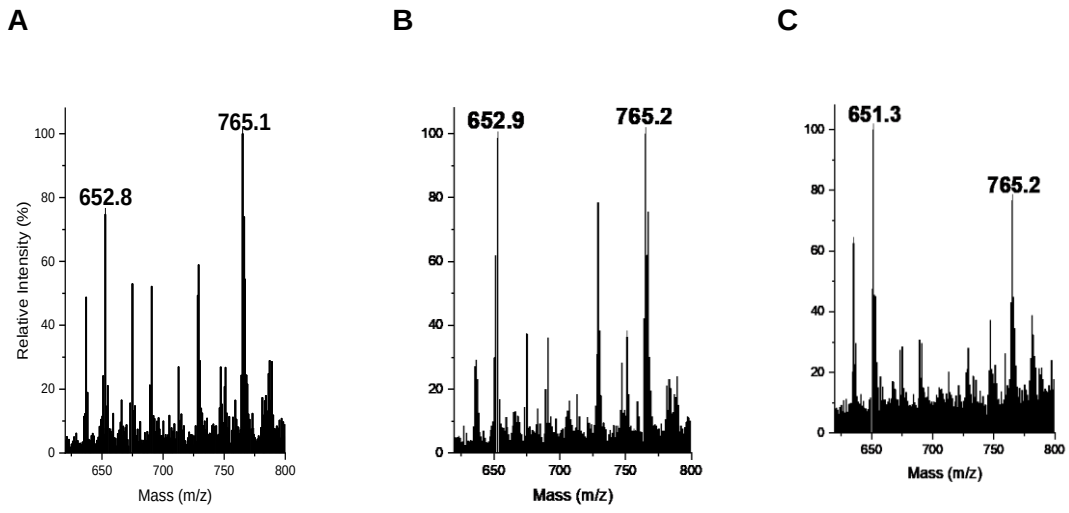
To a 200 µM solution of **58b**, 10 eq of Glucose were added (100 mM stock solution in water) and the resulting mixture was stirred at 37°C and monitored by ESI-MS.



Scheme 84 – ESI-MS spectra of 58b with 10 eq of glucose after 8h (A), 24h (B) and 32h (C)

- **Stability of 58b in the presence of 10 eq of glutathione (GSH)**

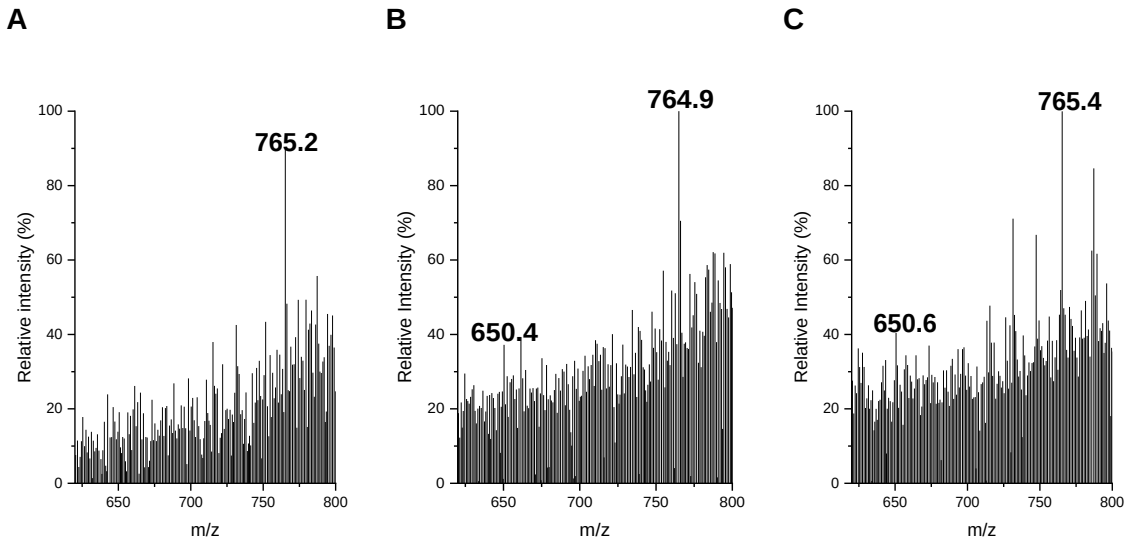
To a 200 µM solution of adduct **58b**, 10 eq of glutathione were added (100 mM stock solution in water), the pH of the reaction was adjusted to 7 using a 100 µM solution of ammonia in water. The resulting mixture was stirred at 37°C and monitored by ESI-MS.



Scheme 85 – ESI-MS spectra of 58b with 10 eq of GSH after 8h (A), 24h (B) and 32h (C)

- **Stability of 58b in the presence of 10% of bovine serum albumin (BSA)**

To a 200 µM solution of adduct **58b**, 0.1 eq of bovine serum albumin were added (1 mg/mL stock solution in water), the pH of the reaction was adjusted to 7 using a 100 µM solution of ammonia in water. The resulting mixture was stirred at 37°C and monitored by ESI-MS.



Scheme 86 – ESI-MS spectra of 58b with 10% of BSA after 8h (A), 24h (B) and 32h (C)

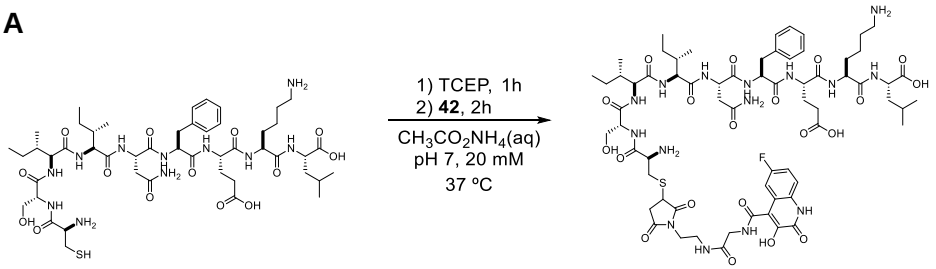
- **Reactions of BHS 7 with model peptides and 2FBBA**

Each of the model peptides (C-Ovalbumin, Cys-RGD and F3-Cys) were prepared as stock solutions 1mg/mL in water, then diluted to 500 µM in ammonium acetate 20 mM pH 7. The functionalization with BHS 7 was performed with different conditions, optimized for each peptide.

- **Functionalization of C-Ovalbumin with BHS 42 and reaction with 2FBBA**

To a 500 µM solution in ammonium acetate 20 mM pH7 of C-Ovalbumin, 2 eq of TCEP (50 mM in water) were added and the mixture was stirred 1h at 37 °C. Subsequently, 5 eq of BHS **42** (100 mM in DMF) were added and the mixture was stirred 2h until complete functionalization of the peptide.

A



C-Ovalbumin fragment

m/z $[M+H]^+$ = 1067.3

m/z $[M+2H]^{2+}$ = 534.1

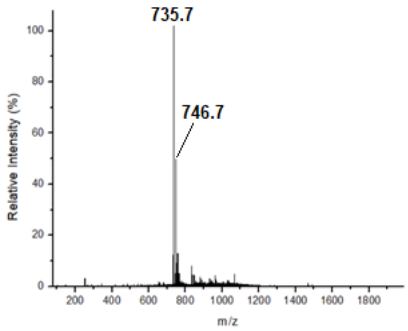
63

m/z $[M+H]^+$ = 1468.6

m/z $[M+2H]^{2+}$ = 735.7

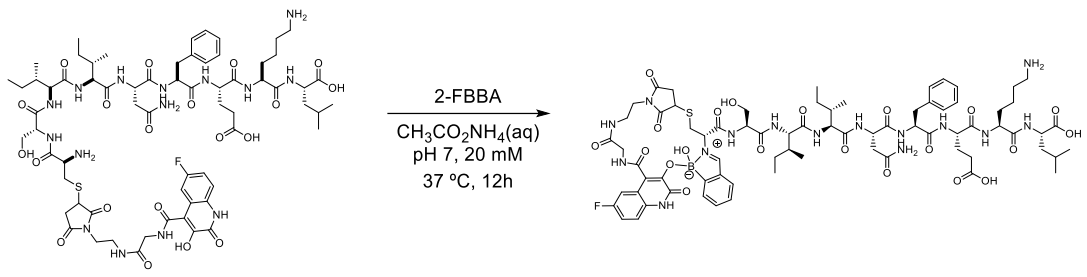
m/z $[M+H+Na]^{2+}$ = 746.7

B



Scheme 87 – Functionalization of C-Ovalbumin fragment with BHS 42 (A) and relative ESI-MS spectrum (B)

The solution obtained in the previous step was diluted to 200 µM with ammonium acetate 20 mM pH 7, then 10 eq of 2FBBA (100 mM in DMF) were added and it was stirred at 37 °C for 12h.



63

m/z $[M+H]^+$ = 1468.6

m/z $[M+2H]^{2+}$ = 735.7

m/z $[M+H+Na]^{2+}$ = 746.7

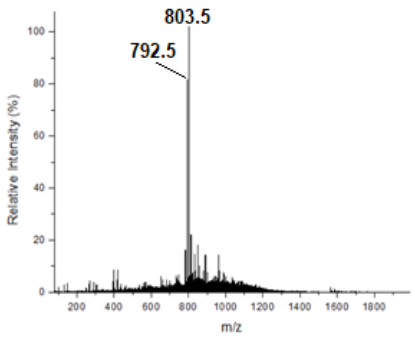
64

m/z $[M+H]^+$ = 1583.5

m/z $[M+2H]^{2+}$ = 792.3

m/z $[M+H+Na]^{2+}$ = 803.5

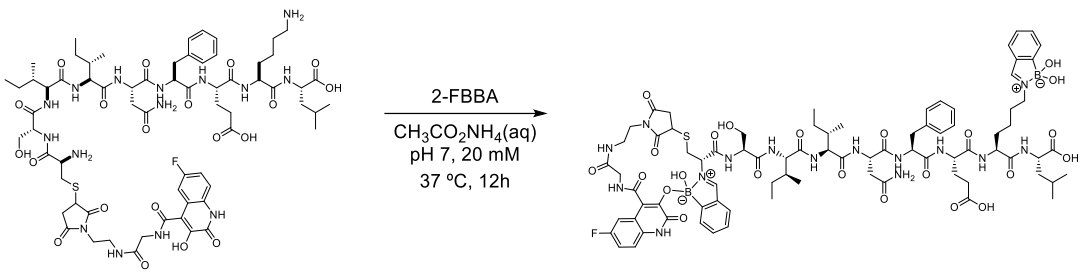
B



Scheme 88 – Reaction of C-Ovalbumin-BHS derivative 63 with 2-FBBA (A) and relative ESI-MS spectrum (B)

- **Double iminoboronate formation on C-Ovalbumin-BHS 63**

The solution of C-Ovalbumin-BHS **63** obtained in the previuos step was stirred overnight with 1000 eq of 2FBBA (1 M stock solution in DMF) and then analyzed by ESI-MS.



63

m/z $[M+H]^+$ = 1468.6

m/z $[M+2H]^{2+}$ = 735.7

m/z $[M+H+Na]^{2+}$ = 746.7

64b

m/z $[M+H]^+$ = 1714.5

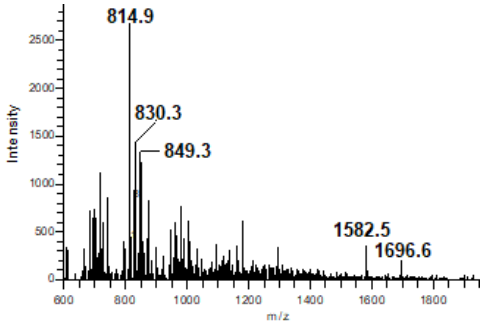
m/z $[M\text{-}OH]^+$ = 1696.6

m/z $[M+H\text{-}OH]^{2+}$ = 849.3

m/z $[M\text{-}2OH]^{2+}$ = 830.3

m/z $[M+H\text{-}3OH]^{2+}$ = 814.9

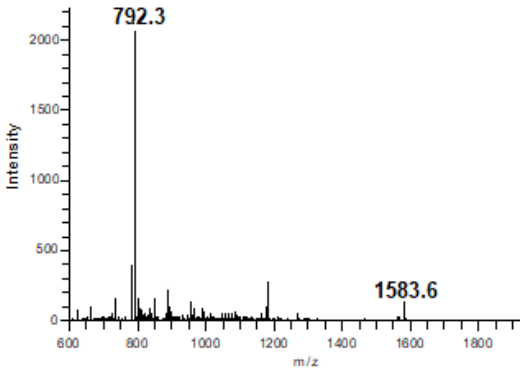
B



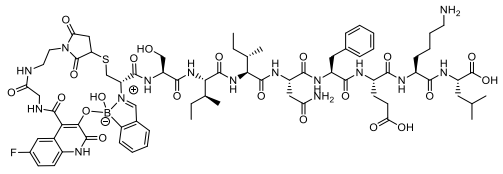
Scheme 89 – Reaction of C-Ovalbumin-BHS derivative 63 with 1000 eq of 2-FBBA (A) and relative ESI-MS spectrum (B)

The double iminoboronate solution was then dialyzed for 24h in ammonium acetate pH solution, yielding the following ESI-MS spectrum, which indicates the presence of only the C-Ovalbumin BHS iminoboronate.

A



B



64

m/z $[M+H]^+$ = 1583.5

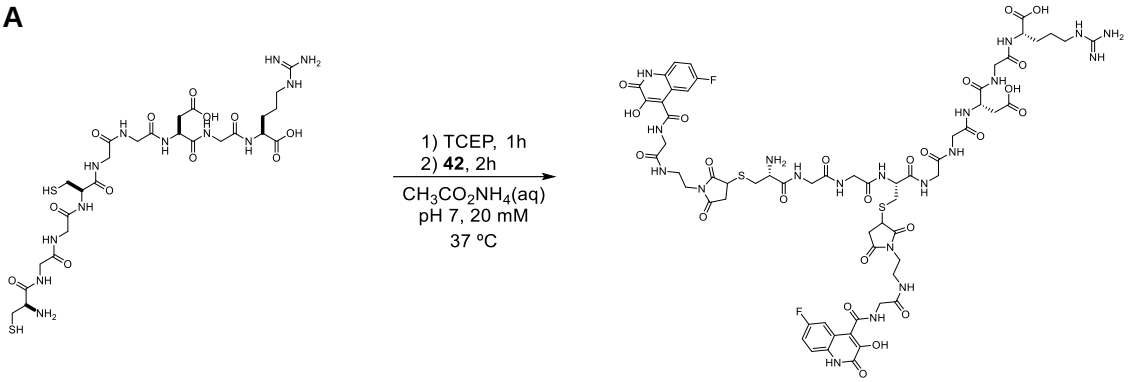
m/z $[M+2H]^{2+}$ = 792.3

Scheme 90 – ESI-MS spectrum of the solution of 64b after 24h of dialysis in ammonium acetate pH 7 20 Mm (A) and structure of C-Ovalbumin 3HQ-iminoboronate 64 (B)

- **Functionalization of Cys-RGD with BHS 42 and reaction with 2FBBA**

To a 500 µM solution in ammonium acetate 20 mM pH7 of Cys-RGD, 4 eq of TCEP (50 mM in water) were added, along with 10 eq of BHS **42** (100 mM in DMF) and the mixture was stirred 4h until complete functionalization of the peptide on both cysteine residues.

A



Cys-RGD

m/z $[M+H]^+$ = 781.3

m/z $[M+2H]^{2+}$ = 391.1

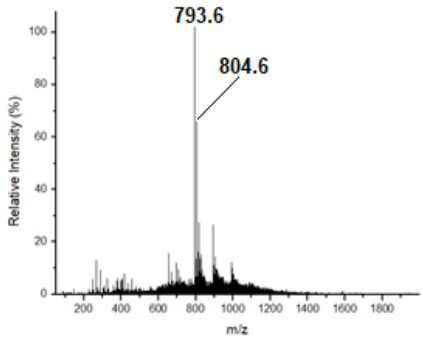
65

m/z $[M+H]^+$ = 1586.5

m/z $[M+2H]^{2+}$ = 793.6

m/z $[M+H+Na]^{2+}$ = 804.6

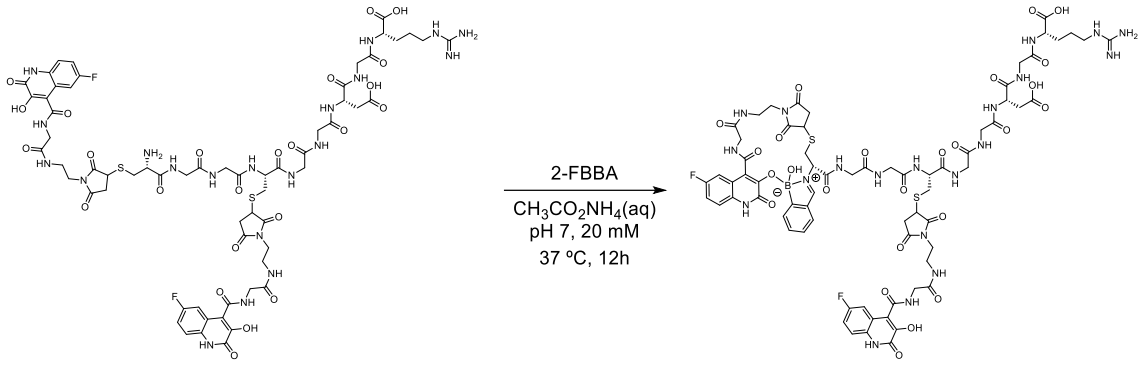
B



Scheme 91 – Functionalization of the Cys-RGD peptide with BHS 42 (A) and relative ESI-MS spectrum (B)

The solution obtained in the previous step was diluted to 200 µM with ammonium acetate 20 mM pH7, then 40 eq of 2FBBA (100 mM in DMF) were added and the mixture was stirred at 37 °C for 12h.

A



65

m/z $[M+H]^+$ = 1586.5

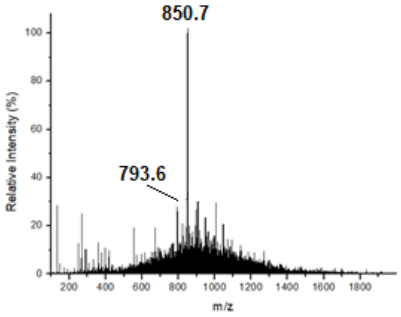
m/z $[M+2H]^{2+}$ = 793.6

66

m/z $[M+H]^+$ = 1700.4

m/z $[M+2H]^{2+}$ = 850.7

B

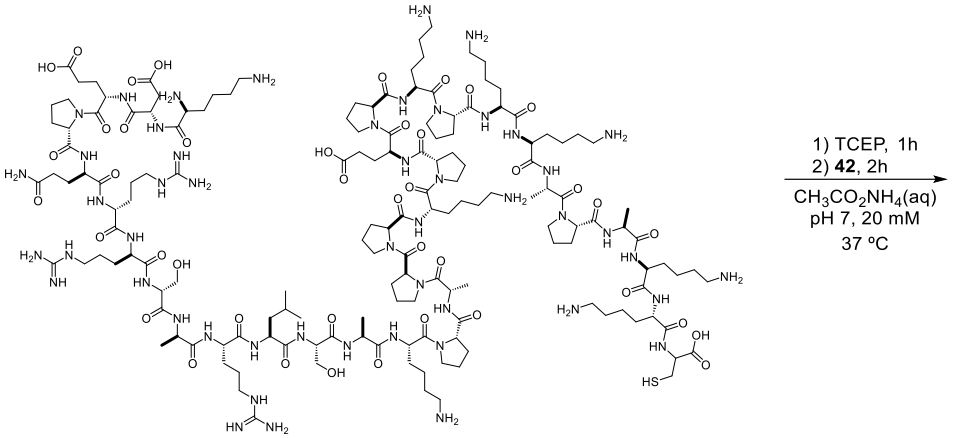


Scheme 92 – Reaction of 3HQ-Cys-RGD derivative 65 with 2-FBBA (A) and relative ESI-MS spectrum (B)

- **Functionalization of F3-Cys with BHS 42 and reaction with 2FBBA**

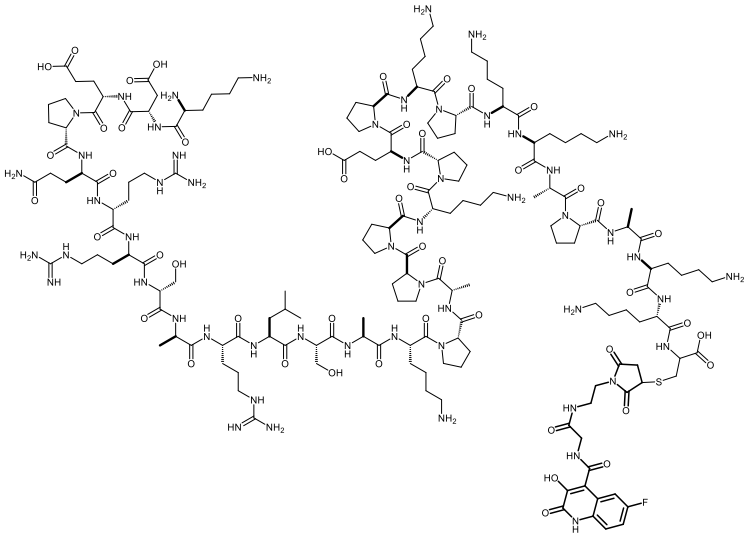
To a 200 µM solution in ammonium acetate 20 mM pH7 of F3-Cys, 5 eq of TCEP (50 mM in water) were added, along with 20 eq of BHS **42** (250 mM in DMF) and the mixture was stirred 4h until complete functionalization of the peptide.

A



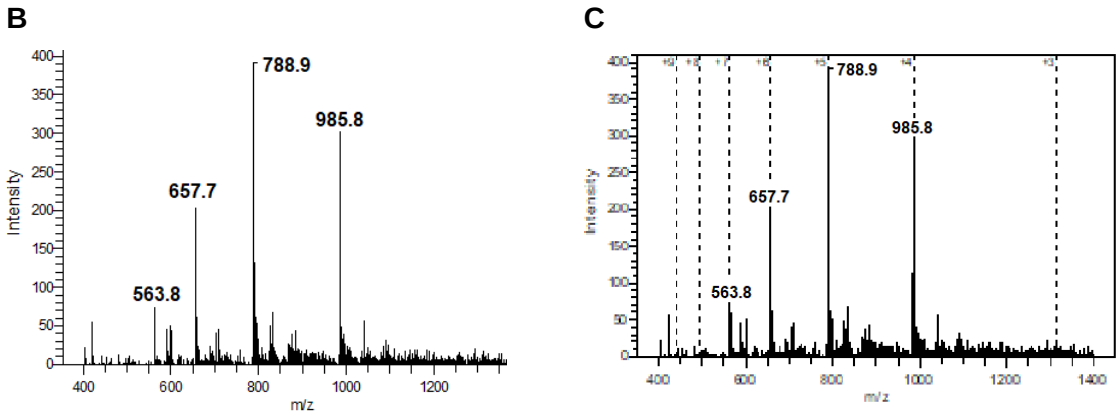
F3-Cys

m/z $[M+H]^+$ = 3536.2



67

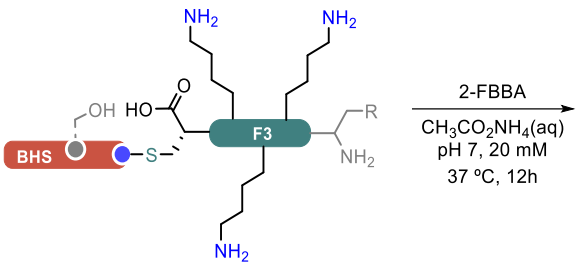
m/z $[M+H]^+$ = 3937.6



Scheme 93 Functionalization of F3-Cys peptide with BHS 42 (A) along with relative ESI-MS spectrum (B) and deconvolution of the peaks belonging to 67 (C)

The solution obtained in the previous step was mixed with 30 eq of 2FBBA (30 mM in DMF) and stirred at 37 °C for 12h.

A

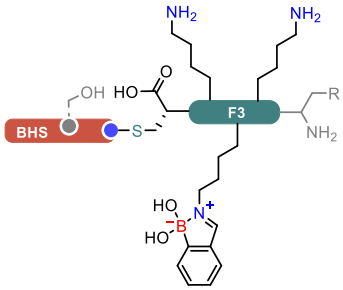


67

m/z $[M+H]^+$: 3937.6

m/z $[M+3H]^{3+}$ = 1313.5

m/z $[M+4H]^{4+}$ = 985.5

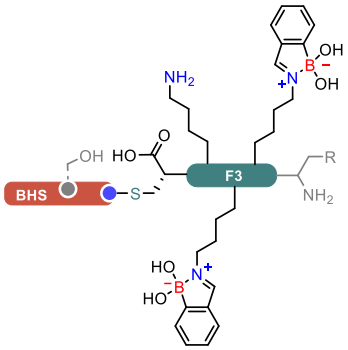


68

m/z $[M+H]^+$ = 4052.9

m/z $[M+3H]^{3+}$ = 1351.4

m/z $[M+4H]^{4+}$ = 1014.4



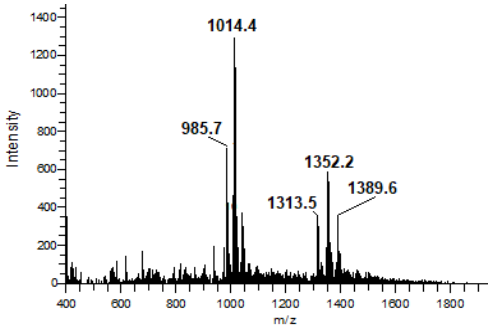
69

m/z $[M+H]^+$ = 4165.1

m/z $[M+3H]^{3+}$ = 1389.6

m/z $[M+4H]^{4+}$ = 1042.1

B

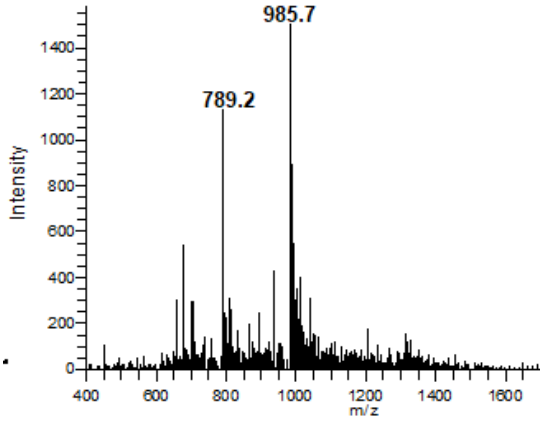


Scheme 94 – Reaction of F3-Cys-BHS 67 with 2FBBA (A) and relative ESI-MS spectrum (B)

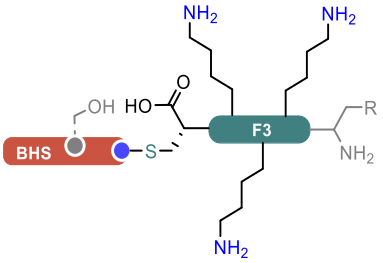
The deconvolution of the spectrum reported above shows the presence of the starting material (m/z = 3937.6), the mono iminoboronate derivative (m/z = 4052.9) and the double iminoboronate product (m/z = 4165.1).

The reaction mixture was dialyzed for 12h in ammonium acetate pH7, leading to the complete disassembly of both iminoboronate products, indicating that they are classic iminoboronates formed on the various lysine residues present on the F3 peptide.

A



B



99

m/z $[M+H]^{+}$: 3937.6

m/z $[M+4H]^{4+}$ = 985.5

m/z $[M+5H]^{5+}$ = 789.2

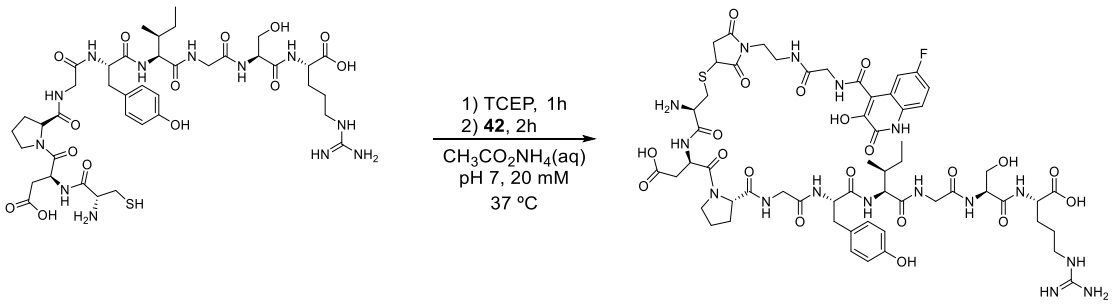
Scheme 95 –ESI-MS spectrum of the reaction mixture after 12h of dialysis (A) along with structure of the starting F3-Cys-BHS 67

Functionalization of Laminin with BHS 42 and conjugation with 2FBBAs

To a 500 µM solution in ammonium acetate 20 mM pH7 of Laminin, 1.5 eq of TCEP (50 mM in water) were added, along with 5 eq of BHS **42** (100 mM in DMF) and the mixture was stirred until complete functionalization of the peptide.

-

A



Laminin fragment

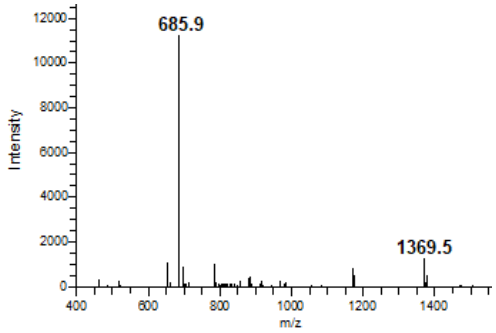
m/z $[M+H]^+$: 967.2

47

m/z $[M+H]^+$: 1369.5

m/z $[M+2H]^{2+}$ = 685.9

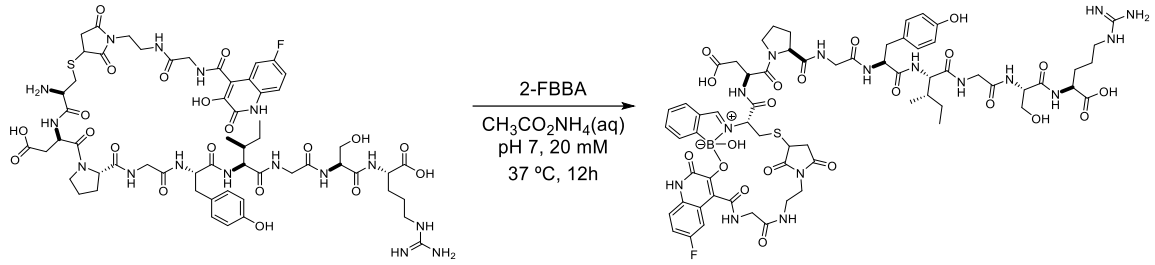
B



Scheme 96 – Functionalization of Laminin fragment peptide with BHS 42 (A) and relative ESI-MS spectrum (B)

The solution obtained in the previous step was mixed with 10 eq of 2FBBA (100 mM in DMF) and stirred at 37 °C for 12h.

A



47

m/z $[M+H]^+$: 1369.5

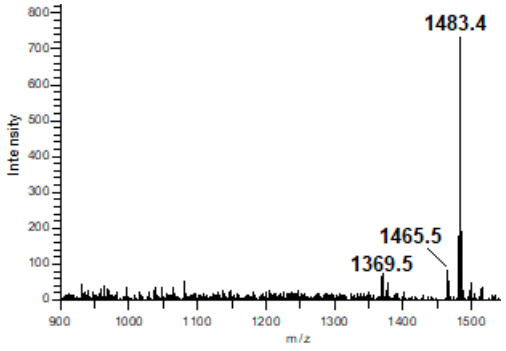
m/z $[M+2H]^{2+}$ = 685.9

70

m/z $[M+H]^+$: 1483.4

m/z $[M-OH]^+$ = 1465.6

B

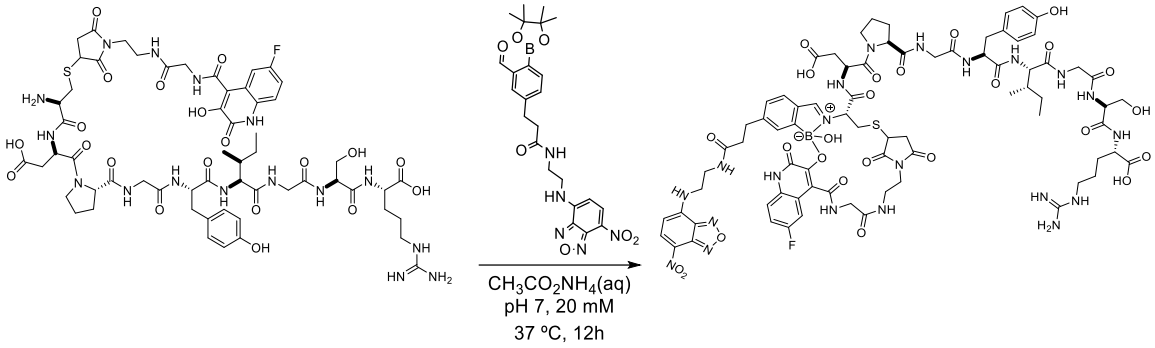


Scheme 97 – Reaction of Laminin-BHS 47 with 2FBBA (A) along with relative ESI-MS spectrum (B)

- **Functionalization of BHS-Laminin 47 with 2FBBA-NBD 71**

The solution of BHS-Laminin **47** obtained in the previous step was mixed with 50 eq of 2FBBA-NBD probe **71** (100 mM in DMF) and stirred at 25 °C for 12h.

A



47

m/z $[M+H]^+$: 1369.5

m/z $[M+2H]^{2+}$ = 685.9

71

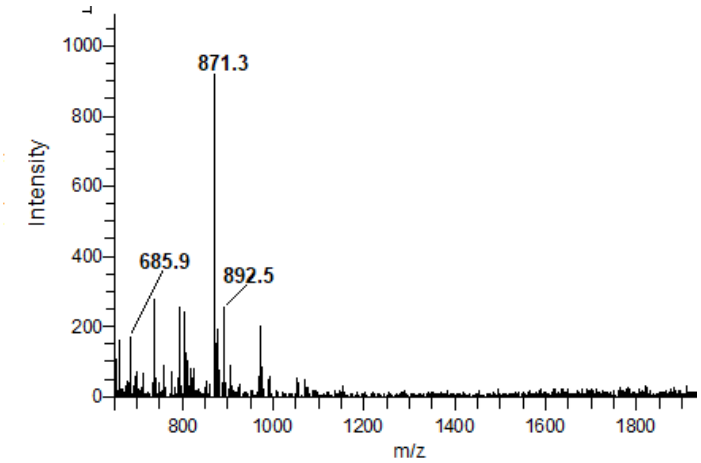
73

m/z $[M+H]^+$: 1760.6

m/z $[M+H+Na]^{2+}$ = 892.5

m/z $[M+H-OH]^{2+}$ = 871.3

B



Scheme 98 - Reaction of BHS-Laminin 47 with 2FBBA-NBD 71 (A) and relative ESI-MS spectrum (B)

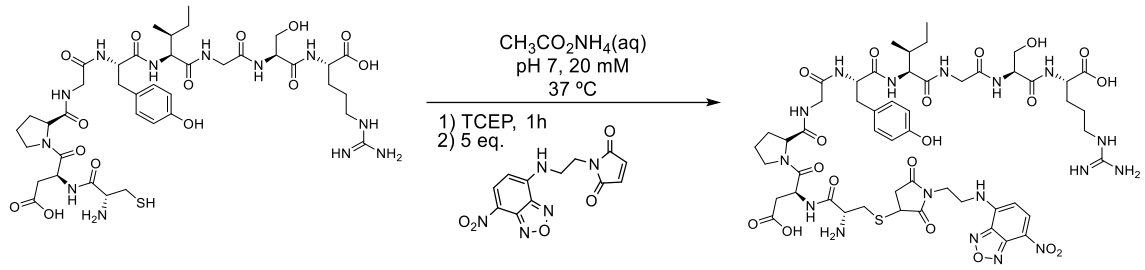
The resulting reaction mixture was dialyzed 48h with ammonium acetate pH 7.

- **Functionalization of Laminin with maleimide-NBD 101**

The thioether-linked Laminin-NBD conjugate was obtained using a maleimide-NBD derivative already synthesized in our group,[165] compound **101**, which was readily available for this work.

To a 500 µM solution in ammonium acetate 20 mM pH7 of Laminin, 1.5 eq of TCEP (50 mM in water) were added, along with 5 eq of maleimide-NBD **100** (100 mM in DMF) and the mixture was stirred 5h until complete functionalization of the peptide. After complete conjugation, the reaction mixture was dialyzed for 48h with ammonium acetate solution pH7.

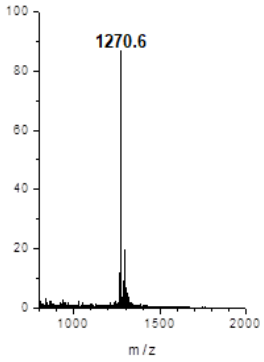
A



Laminin fragment **101** **72**

m/z $[M+H]^+$: 967.2 *m/z* $[M+H]^+$: 1270.6

B



Scheme 99 – Reaction of Laminin fragment with NBD-maleimide derivative 102 (A) along with relative ESI-MS spectrum (B)

V.2.4. Computational studies

All calculations were performed using the GAUSSIAN 09 software package,[154] and the M06-2X functional, without symmetry constraints. That is a hybrid meta-GGA functional developed by Truhlar and Zhao,[155] and it was shown to perform very well for main-group kinetics, providing a good description of long range effects such as van der Waals interactions or π-π stacking. The optimized geometries were obtained with a standard 6-31G(d,p)[156] basis set. The electronic energies obtained at that level (E_{b1}) were converted to free energy at 298.15 K and 1 atm (G_{b1}) by using zero point energy and thermal energy corrections based on structural and vibration frequency data calculated at the same level. Solvent effects (water) were accounted for in all calculations by means of the Polarisable Continuum Model (PCM) initially devised by Tomasi and coworkers[157–159] with radii and non-electrostatic terms of the SMD solvation model, developed by Truhlar *et al*.[160]

Atomic coordinates of the optimized species

H_2O

H	0.309606	0.555609	0.407703
O	0.359183	-0.390182	0.219479
H	-0.355188	-0.554827	-0.408983

50a

C	-5.570270	0.203704	1.976483
C	-5.775793	-1.175170	2.052728
C	-4.731818	-2.013038	1.712896
C	-3.509815	-1.466000	1.303025
C	-3.314321	-0.069567	1.225758
C	-4.386065	0.774303	1.579103
N	-2.461681	-2.322555	0.956011
C	-1.289811	-1.860010	0.568525
C	-1.043285	-0.441310	0.514321
C	-2.028641	0.445739	0.813724
C	-1.778914	1.907446	0.582640
O	-2.444783	2.538771	-0.241966
N	-0.763683	2.428590	1.295795
C	-0.155020	3.705936	0.991609
F	-6.595351	1.016135	2.318558
C	1.200527	3.600535	0.303829
O	1.862086	4.633337	0.115084
N	1.620274	2.383591	-0.057930
C	2.957035	2.133150	-0.567687
C	4.014700	2.363267	0.517107
N	5.277674	1.756235	0.130759
C	5.590260	0.442324	0.423640
C	6.874772	0.076259	-0.296568
C	7.393094	1.414180	-0.834165
C	6.242907	2.374507	-0.643589

O	6.150019	3.514562	-1.061523	H	-0.000388	4.278674	1.908888
O	4.903746	-0.278478	1.125764	H	-0.824369	4.276118	0.344709
S	6.457198	-1.039327	-1.686341	H	1.039886	1.571650	0.131773
C	6.041293	-2.550602	-0.749096	H	3.160489	2.768757	-1.434423
C	5.469870	-3.614002	-1.697325	H	2.985638	1.087621	-0.890148
C	5.344033	-4.892688	-0.877073	H	3.702673	1.894811	1.452998
O	4.225031	-4.934736	-0.164377	H	4.186397	3.426578	0.689902
O	6.201714	-5.754798	-0.840177	H	7.566873	-0.426915	0.381342
N	4.210352	-3.339291	-2.382651	H	7.695830	1.399892	-1.883652
O	-0.285489	-2.560386	0.188621	H	8.238220	1.776161	-0.241532
O	0.195829	-0.230644	0.088236	H	5.337489	-2.308512	0.049790
H	-2.615681	-3.331067	0.983994	H	6.958907	-2.941404	-0.299993
C	3.441400	-2.427790	-1.926076	H	6.215419	-3.840060	-2.464226
C	2.160691	-2.106293	-2.600954	C	4.056791	-6.080059	0.691124
C	1.025602	-1.684136	-1.876549	H	3.682406	-1.835174	-1.042824
C	-0.130768	-1.393367	-2.621955	H	-1.033603	-1.070400	-2.105489
C	-0.178072	-1.500648	-4.008689	H	-1.091985	-1.259014	-4.543161
C	0.958110	-1.913640	-4.705653	H	0.942691	-1.992389	-5.788402
C	2.116562	-2.214282	-4.000808	H	3.012223	-2.519443	-4.532596
B	0.880403	-1.554037	-0.273671	H	1.861735	-1.640712	1.439448
O	2.036307	-1.782094	0.498592	H	3.095257	-5.937993	1.179669
H	-6.734769	-1.564080	2.375966	H	4.862270	-6.113544	1.426845
H	-4.846228	-3.091594	1.761020	H	4.053230	-6.991955	0.092047
H	-4.285834	1.853836	1.543581	**51**			
H	-0.299676	1.824334	1.966867	C	-5.557769	0.102011	-3.387826

C	-5.991526	-1.212135	-3.212650	O	-0.647522	0.790021	2.289445
C	-5.163312	-2.095378	-2.545089	S	1.591371	-1.407028	3.439485
C	-3.919840	-1.657898	-2.067324	C	0.697642	-2.216095	2.067434
C	-3.495334	-0.330947	-2.255831	C	1.721624	-2.612019	0.988768
C	-4.345969	0.563990	-2.933880	C	1.358913	-3.877550	0.220630
N	-3.084862	-2.538869	-1.398131	O	0.812509	-4.792654	1.004858
C	-1.869287	-2.233713	-0.877001	O	1.643190	-4.068234	-0.945297
C	-1.436078	-0.832097	-1.034068	N	1.976507	-1.491585	0.085102
C	-2.207500	0.054436	-1.719932	O	-1.164214	-3.086528	-0.314132
C	-1.664934	1.422687	-2.023816	O	-0.263670	-0.506297	-0.471585
O	-1.470499	1.763892	-3.196466	C	2.972924	-0.700486	0.310379
N	-1.413533	2.207638	-0.965782	C	3.006165	0.473663	-0.553724
C	-0.768039	3.499313	-1.076868	C	1.919133	0.441345	-1.440009
F	-6.377182	0.962542	-4.039581	C	1.777819	1.489466	-2.343342
C	0.301873	3.741210	-0.025013	C	2.686354	2.554721	-2.314653
O	0.675877	4.907061	0.188568	C	3.749272	2.575747	-1.404521
N	0.784068	2.685038	0.634260	C	3.930078	1.517594	-0.513869
C	1.680513	2.831120	1.766612	B	1.014426	-0.864878	-1.145835
C	0.939944	3.089887	3.087276	O	0.878205	-1.718637	-2.272758
N	0.688803	1.858273	3.832069	H	-6.958413	-1.520500	-3.594889
C	0.001722	0.772091	3.320199	H	-5.465100	-3.126029	-2.385870
C	0.257812	-0.421484	4.217761	H	-4.062140	1.598007	-3.099171
C	0.797468	0.206879	5.502996	H	-3.390723	-3.503554	-1.291698
C	1.235664	1.588226	5.078308	H	-1.662908	1.855520	-0.046158
O	1.924404	2.368841	5.709209	H	-1.490962	4.314590	-0.981268

H	-0.303260	3.574498	-2.064517	H	3.711994	-0.911524	1.082896
H	0.440515	1.756196	0.412432	H	0.955478	1.503816	-3.054276
H	2.353773	3.666479	1.562897	H	2.563053	3.384026	-3.006259
H	2.284609	1.922563	1.846824	H	4.437072	3.415417	-1.396845
H	-0.009521	3.599980	2.901564	H	4.753371	1.511339	0.195518
H	1.546342	3.715456	3.743232	H	0.593019	-2.610325	-2.028344
H	-0.636802	-1.031561	4.346707	H	2.673494	-2.885379	1.468606
H	1.624742	-0.330130	5.970841	H	0.070551	-6.666004	1.170472
H	-0.001524	0.321500	6.242117	H	1.411918	-6.509226	-0.008877
H	-0.031806	-1.524219	1.645239	H	-0.220793	-5.896073	-0.421690
H	0.188443	-3.098953	2.452672				
C	0.501169	-6.052894	0.381870				

V.2.5. *In vitro* internalization experiments

- **HT29 Cell culture**

Human colorectal cancer cell lines HT-29 was obtained from ECACC (European Collection of Authenticated Cell Cultures). Cells were maintained at 37 ºC with 5% CO2 in DMEM-Dulbecco's Modified Eagle Medium (Gibco® by Life Technologies), supplemented with 10% of heat-inactivated Fetal Bovine Serum (Gibco® by Life Technologies), 100 IU/ml penicillin and 100 mg/ml streptomycin (PenStrep, Invitrogen).

- **Live-cell microscopy**

Cells were cultured on µ-Slide 8 well IBIDI chambers (ibidi®) and left to adhere overnight before being treated with compounds 13, 14 and 32 at 2µM concentration for 1h in serum-free DMEM. After this incubation, extracellular probe was removed by washing with serum-free DMEM and cells were simultaneously labeled with Wheat germ agglutinin–Alexa Fluor 647 (WGA-A647) conjugate (5.0 µg/mL) and Hoechst 33342 (2 µg/mL) to allow for the identification of plasma membrane and nucleus, respectively.

Confocal laser scanning microscopy images were acquired using a Leica TCS SP5 confocal inverted microscope (DMI6000) with a 63x water-immersion apochromatic objective (1.2 numerical aperture). NBD excitation was achieved using a 488 nm Ar laser, and fluorescence was collected between 500 and 570 nm. Alexa-647 excitation was performed at 633 nm using a He-Ne laser and fluorescence emission was collected between 640-750 nm. Two-photon excitation images of Hoechst 33258 fluorescence were acquired in the same setup using a Ti:sapphire laser (Mai Tai, Spectra-Physics, Darmstadt, Germany) set to 800 nm as the excitation source.

- **Fluorescence quantification**

The fluorescence intensity of each well was measured by collecting the fluorescence only of the NBD probe channel. Each cell was identified as a distinct region of interest (ROI) and its fluorescence was quantified, then the mean of these measurements was used as the fluorescence value for each sample. In order to account for the auto fluorescence of the cells, the mean fluorescence intensity of the blank was subtracted from the mean value of each well, thus obtaining the adjusted fluorescency mean.

Blank cells

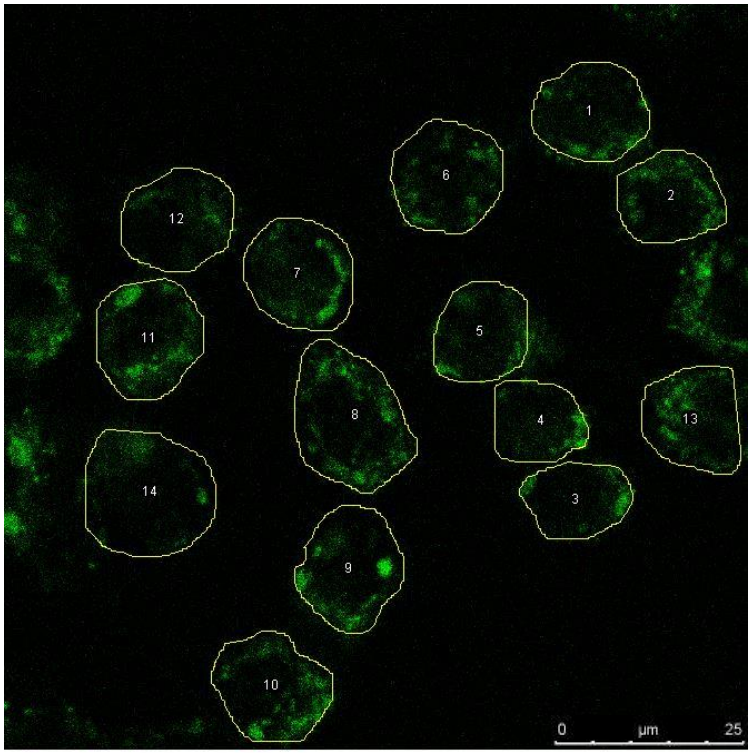


Figure 47 - Confocal microscopy image of untreated cells and ROI areas

ROI	Intensity
1	1.28
2	1.50
3	1.19
4	1.36
5	1.30
6	1.14
7	1.47
8	1.43
9	1.48
10	1.59
11	1.74
12	0.90
13	1.51
14	0.78

Mean	Adj. Mean
1.33	0

Cells incubated with NBD probe 71

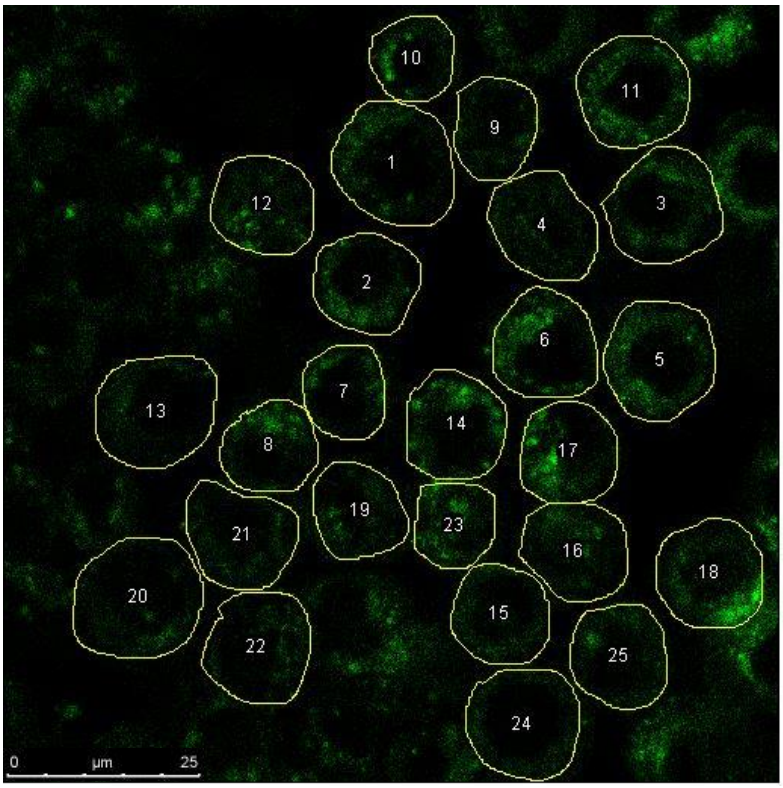


Figure 48 - Confocal microscopy image of cells treated with probe 71 and ROI areas

ROI	Intensity	ROI	Intensity	ROI	Intensity
1	2.57	**9**	2.64	**18**	4.25
2	2.86	**10**	2.95	**19**	2.67
3	2.97	**11**	3.64	**20**	2.32
4	2.80	**12**	3.26	**21**	2.51
5	3.50	**13**	2.06	**22**	2.30
6	3.42	**15**	2.79	**23**	4.25
7	2.71	**16**	3.20	**24**	2.16
8	3.28	**17**	3.87	**25**	2.85

Mean	Adj. Mean
3.01	1.68

Cells incubated with Laminin-NBD reversible conjugate 73

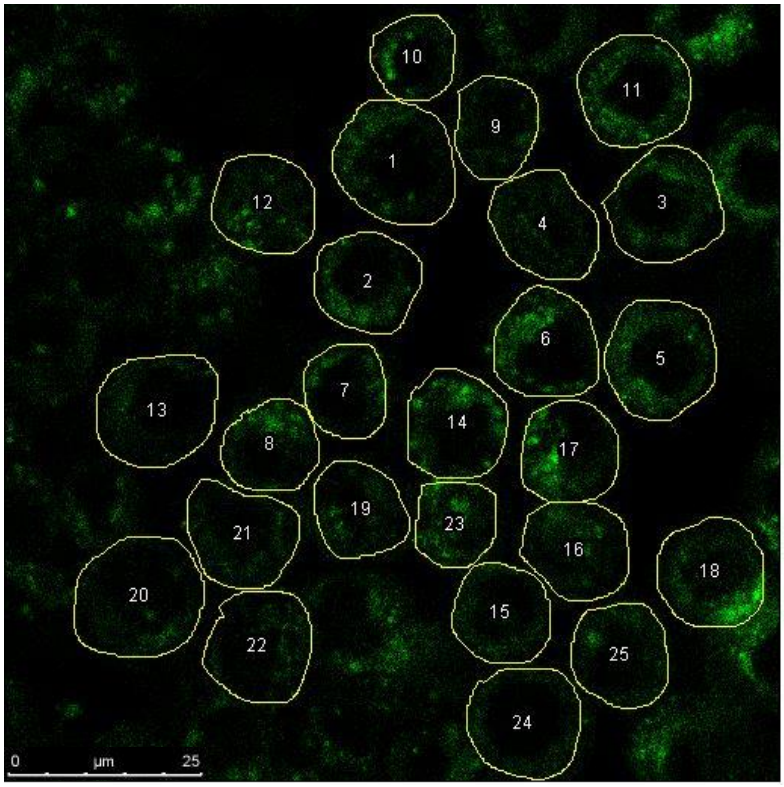


Figure 49 - Confocal microscopy image of cells treated with conjugate 73 and ROI areas

ROI	Intensity	ROI	Intensity	ROI	Intensity
1	15.26	**8**	10.59	**16**	8.41
2	10.87	**9**	8.80	**17**	9.49
3	9.88	**10**	9.82	**18**	8.54
4	10.89	**11**	10.73	**19**	9.22
5	10.78	**12**	10.75		
6	9.28	**13**	14.13		
7	9.00	**14**	7.29		

Mean	Adj. Mean
10.30	8.97

Cells incubated with Laminin-NBD covalent conjugate 72

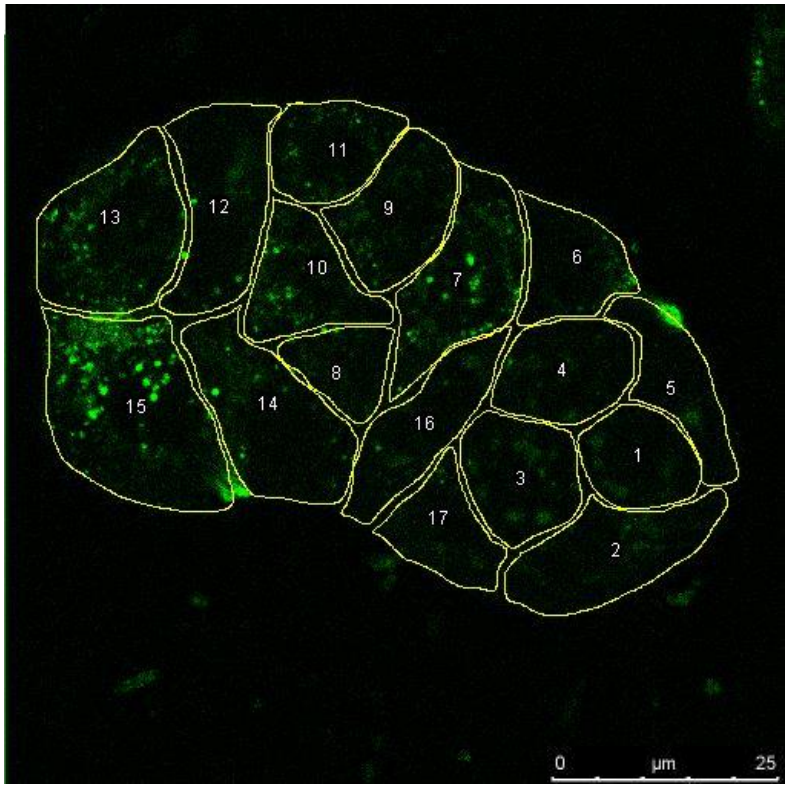


Figure 50 - Confocal microscopy image of cells treated with conjugate 72 and ROI areas

ROI	Intensity	ROI	Intensity	ROI	Intensity
1	1.14	**7**	2.10	**13**	1.87
2	1.11	**8**	1.45	**14**	1.54
3	1.27	**9**	1.40	**16**	2.78
4	1.20	**10**	1.88	**17**	1.21
5	1.45	**11**	1.51		
6	1.28	**12**	1.31		

Data treatment

Sample	Mean fluorescence	Adj. mean fluorescence	Sample fluorescence/ blank fluorescence
Blank	1.33	0	1
71	3.01	1.68	2.26
73	10.30	8.97	7.74
72	1.51	0.18	1.14

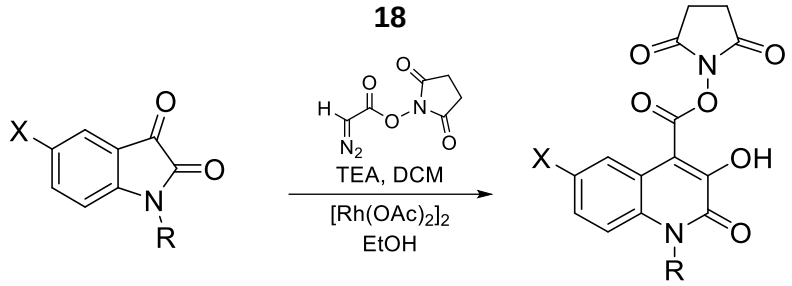
The results for compound **71** show that the probe is able to passively internalize with good efficiency thanks to its BA moiety. When the probe is covalently linked to the targeting peptide, little difference is observed compared to the blank sample, as the 67LR receptor does not promote internalization and the probe cannot passively internalize, since it is irreversibly bound to the peptide. Finally, the reversible conjugate **73** produced a significant increase in fluorescence inside the treated cells, due to its accumulation on the cell surface, followed by payload release, which produced a high concentration of NBD-FBBA probe outside the cell, which, as seen for entry **71**, can passively internalize.

V.3. Experimental section of Chapter IV

V.3.1. Synthesis and structural characterization

The synthesis of compounds **74**-**84** was carried out following the general synthetic route described in **Chapter II**.

- **Synthesis of 4-NHS-3HQ derivatives 106-108**

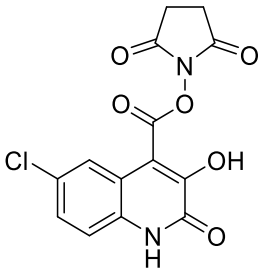


103 R=H X=Cl
104 R=H X=OCF_3
105 R=Bn X=OCF_3

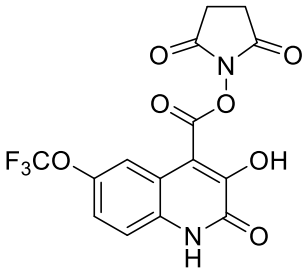
106 R=H X=Cl
107 R=H X=OCF_3
108 R=Bn X=OCF_3

Scheme 100 – Synthetic route towards Synthesis of 4-NHS-3HQ derivatives 106-108

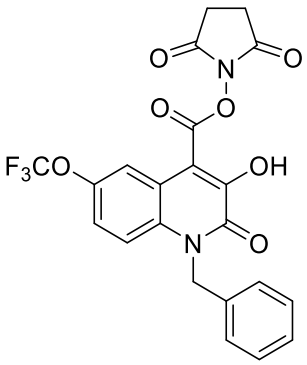
4-NHS-3HQs **106-108** were prepared adapting a procedure previously developed in our group.[119] In a thick-walled vial for centrifuge, 0.606 mmol of the appropriate Isatin were dissolved, along with NHS-Diazoacetate (150 mg, 0.818 mmol) in dry DCM (3 ml) and stirred with dry triethylamine (16.88 µl, 0.121 mmol) at room temperature under Argon atmosphere. When complete conversion of the starting isatin was observed, the solvent was evaporated at reduced pressure, then the resulting crude was dissolved in absolute ethanol (3 ml), and then diacetoxyrhodium (2.68 mg, 6.06 µmol) was added to the mixture, causing a rapid bubbling of the reaction mixture, with consequent formation of a beige precipitate. The reaction was stirred for 30 minutes, then it was centrifuged, and the solid was collected by decantation. The resulting precipitate was washed with ethanol and then twice with diethyl ether, yielding the corresponding NHS-3-hydroxyquinolin-2(1H)-one derivatives as pure solids.



4-NHS-3-HQ. **106** Grey solid, yield 50%. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.59 (s, 1H), 7.92 (d, J = 2.1 Hz, 1H), 7.43 (d, J = 2.2 Hz, 1H), 7.35 (d, J = 8.7, 1H), 2.92 (s, 4H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 170.34, 160.85, 157.25, 149.27, 131.46, 127.13, 126.96, 122.44, 118.32, 117.48, 108.77, 25.69. **LRMS (m/z)**: calculated for $[M+H^+]$ $C_{14}H_{10}ClN_2O_6$: 337.0, found: 337.3.



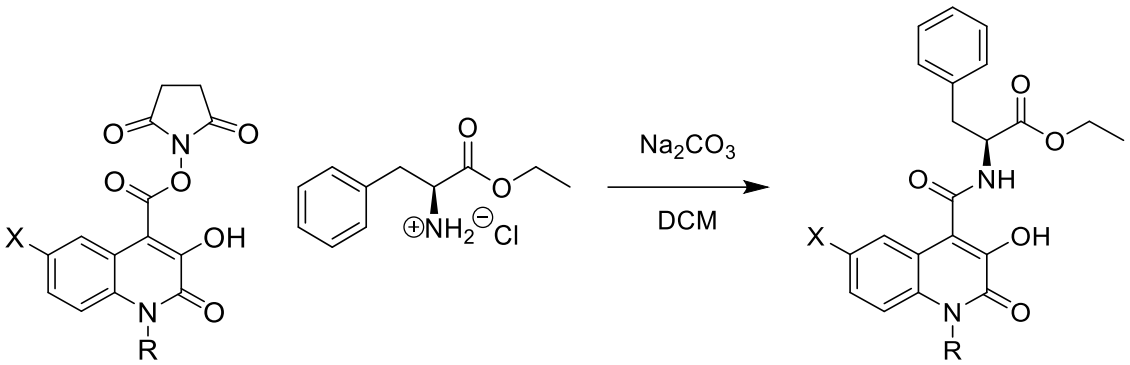
4-NHS-3-HQ 107. Grey solid, yield 97% **^{1}H NMR (300 MHz, DMSO)** δ 12.67 (s, 1H), 7.85 (s, 1H), 7.42 (s, 2H), 2.90 (s, 4H); **^{13}C NMR (75 MHz, DMSO)** δ 170.31, 160.77, 157.19, 149.00, 143.59, 131.77, 121.88, 120.63, 118.48, 117.65 (d, J = 9.84 Hz), 115.75, 109.39, 25.69. **Elemental analysis** calculated (%) for $C_{15}H_9F_3N_2O_7$ + 0,3 H_2O: C, 46.00; H, 2.48; N, 7.15, found (%): C, 46.66; H, 2.70; N, 7.21.



4-NHS-3-HQ 108. Grey solid, yield 80%. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 7.93 (s, 1H), 7.54 (d, J = 9.3 Hz, 1H), 7.44 (dd, J = 9.3, 1.8 Hz, 1H), 7.38 – 7.30 (m, 2H), 7.26 (t, J = 5.8 Hz, 3H), 5.64 (s, 2H), 2.93 (s, 4H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 173.37, 161.56,

158.92, 146.88, 143.72, 136.69, 133.07, 129.23, 127.74, 126.90, 123.10, 121.01, 120.81 (d, J=255.6 Hz), 119.95, 118.99, 117.08, 112.17, 46.113, 25.65. **LRMS (m/z)**: calculated for $[M+H^+]$ $C_{22}H_{16}F3N_2O_7$: 477.1, found: 477.3.

- **Synthesis of 3HQ-phenylalanine derivatives 74 - 77**

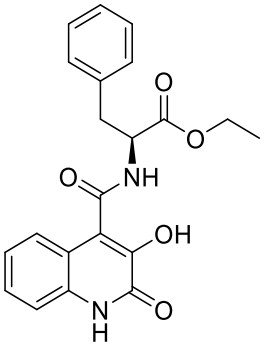


21 R=H X=H
106 R=H X=Cl
107 R=H X=OCF_3
108 R=Bn X=OCF_3

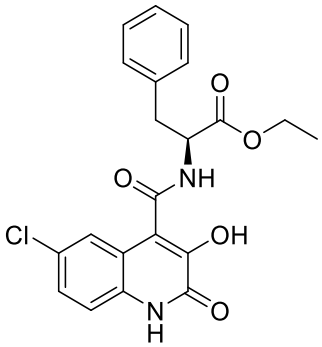
74 R=H X=H
75 R=H X=Cl
76 R=H X=OCF_3
77 R=Bn X=OCF_3

Scheme 101 – Synthetic route towards 3HQ-phenylalanine derivatives 74 - 77

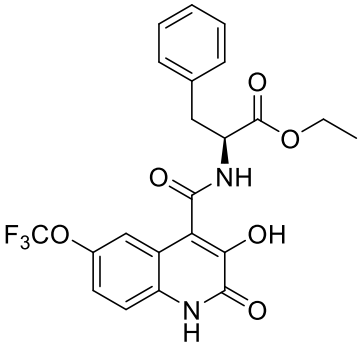
Phenylalanine ethyl ester hydrochloride (50 mg, 0.22 mmol, 1.1 eq) and Na_2CO_3 (212 mg, 2 mmol, 10 eq) were added to a stirred solution of the appropriate 4-NHS-3HQ (0.2 mmol) in dry DCM (1 mL). The mixture was stirred overnight at room temperature after which the volatiles were evaporated under reduced pressure and the crude mixture acidified with HCl (2 N). The precipitate was then centrifuged and collected by filtration. Finally, the isolated solid was thoroughly washed with H_2O to furnish the desired compounds.



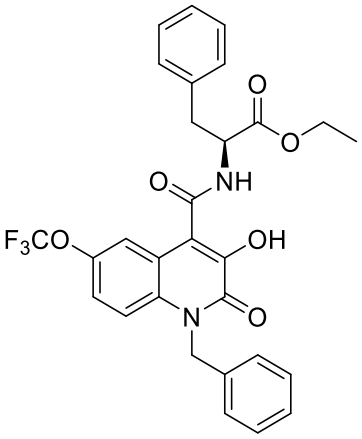
Phe-3HQ 74 Grey solid, yield 93%. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.03 (s, 1H), 9.71 (s, 1H), 7.60 – 6.82 (m, 9H), 4.86 – 4.50 (m, 2H), 4.09 (q, J = 7.1 Hz, 2H), 3.04 (ddd, J = 22.8, 13.9, 7.6 Hz, 1H), 1.15 (t, J = 7.1 Hz, 3H).**^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 171.66, 165.56, 160.02, 142.03, 137.36, 132.34, 129.37, 128.38, 126.62, 125.66, 124.12, 123.31, 122.08, 121.71, 119.50, 114.94, 60.64, 54.00, 14.08. **LRMS (m/z)**: calculated for [M+Na^+] $C_{21}H_{20}N_2O_5Na$: 403.1, found: 403.7. **Elemental analysys** calculated (%) for $C_{21}H_{20}N_2O_5$ + 0.5 H_2O: C, 61.10; H, 5.76; N, 6.79; Found: C, 60.87; H, 5.56; N, 6.65.



Phe-3HQ 75 Grey solid, yield 87%. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.31 (s, 1H), 9.10 (s, 1H), 7.46 – 7.10 (m, 8H), 4.63 (dd, J= 14.2, 8.0 Hz, 2H), 4.12 (q, J = 7.0 Hz, 2H), 3.04 (ddd, J = 23.1, 13.9, 7.6 Hz, 1H), 1.17 (t, J = 7.1 Hz, 3H).**^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 171.60, 164.51, 164.47, 158.42, 143.73, 137.49, 132.07, 129.54, 128.60, 126.98, 126.84, 126.73, 123.22, 120.18, 119.57, 117.28, 61.06, 54.44, 14.36 **LRMS (m/z)**: calculated for [M+Na^+] $C_{21}H_{19}N_2O_5ClNa$: 436.4, found: 436.6. **Elemental analysys** calculated (%) for $C_{21}H_{19}N_2O_5Cl$ + 1.1 H_2O: C, 58.03; H, 4.92; N, 6.45; Found: C, 58.24; H, 4.87; N, 6.25.

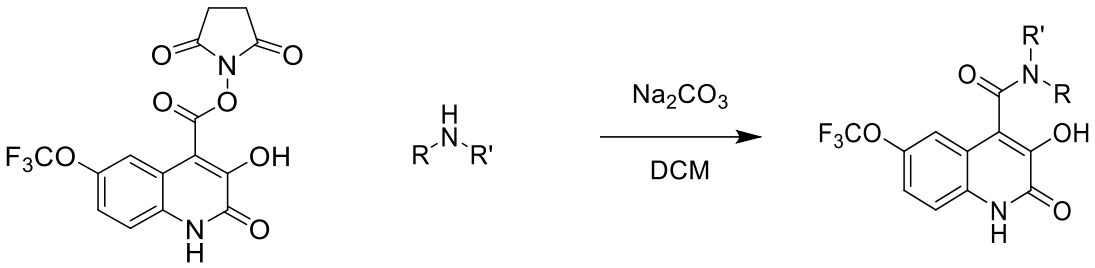


Phe-3HQ 76 Grey solid, yield 57%. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.27 (d, J = 5.9 Hz, 1H), 11.58 (s, 1H), 9.21 (s, 1H), 7.39 – 7.07 (m, 7H), 6.86 (d, J = 6.8 Hz, 1H), 4.63 (q, J = 6.6 Hz, 1H), 4.01 (q, J = 6.9 Hz, 2H), 3.08 – 2.92 (m, 2H), 1.08 (t, J = 7.1 Hz, 3H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 172.40, 169.28, 164.39, 163.01, 143.07, 137.43, 129.28, 128.28, 127.90, 126.48, 124.95, 115.88, 115.05, 113.28, 103.79, 60.06, 53.86, 37.95, 14.02. **LRMS (m/z)**: Calculated for $[M+H^+]$ $C_{22}H_{20}F_3N_2O_6$: 465.1, found: 465.0. **Elemental analysis:** calculated (%) for $C_{22}H_{19}F_3N_2O_6$ + 3 H_2O: C, 54.77; H, 4.39; N, 5.81; found (%):C, 54.51; H, 3.71; N, 6.18



Phe-3HQ 77 Grey solid, yield 80%. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 10.32 (s, 1H), 9.21 (s, 1H), 7.48 (d, J = 9.3 Hz, 1H), 7.38 – 7.07 (m, 12H), 5.63 (q, J = 16.2 Hz, 2H), 4.67 (dd, J = 14.7, 7.5 Hz, 1H), 4.12 (q, J = 7.1 Hz, 2H), 3.06 (ddd, J = 23.1, 13.9, 7.5 Hz, 2H), 1.17 (t, J = 7.1 Hz, 3H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 172.05, 169.40, 164.33, 163.13, 143.65, 143.15, 137.18, 131.90, 131.10, 129.72, 129.15, 128.27, 127.95, 127.36, 124.94, 122.12, 115.95, 115.18, 113.39, 103.91, 77.11, 76.24, 54.99, 14.37. **HRMS (m/z)**: Calculated for $[M+Na^+]$ $C_{29}H_{25}F_3N_2O_6Na$: 555.17370, found: 555.17270.

- **Synthesis of 4-carboxamide 3HQ derivatives 78-81**

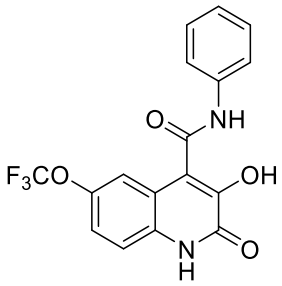


107

78 R=H, R'= Ph
79 R=H, R'= Bn
80 R=CH_3, R'= Bn
81 R=H, R'= CH_2Bn

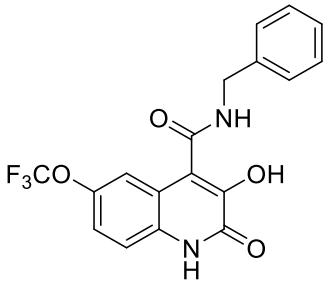
Scheme 102 – Synthetic route towards 4-carboxamide 3HQ derivatives 78-81

In a thick-glassed vial, 0.749 mmol of the appropriate amine were dissolved in dry DCM (3.5 mL) and stirred with Na_2CO_3 (795 mg, 7.5 mmol) until the solution became clear, then intermediate **107** (241 mg, 0.625 mmol) was added and the mixture was stirred overnight at room temperature under argon atmosphere. The solvent was subsequently evaporated at reduced pressure, then the crude was suspended in 1 mL of water, to which aqueous HCl 1M was slowly added dropwise until pH 4 was reached. The resulting suspension was centrifuged, the surnatant was removed and the resulting solid was washed again with water, then collected and dried in high vacuum

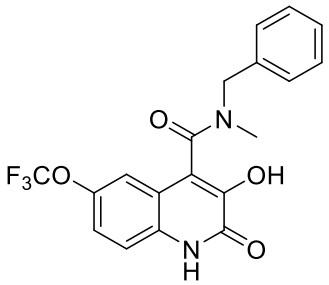


Phenylamide 3HQ 78. Grey solid, 90% yield. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.49 (s, 1H), 10.59 (s, 1H), 7.74 (d, J = 7.6 Hz, 2H), 7.47 – 7.29 (m, 5H), 7.13 (t, J = 7.4 Hz, 1H); **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 162.47, 158.28, 143.93, 143.31, 138.82, 132.05, 128.92, 123.97, 121.87, 120.06, 119.72, 119.49, 119.14, 117.18, 115.54. **LRMS (m/z)**: calculated for [M+H$^+$] $C_{17}H_{12}F_3N_2O_4$: 365.1, found: 365.0. **Elemental analysis**

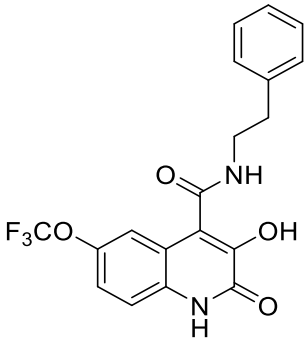
calculated (%) for $C_{17}H_{11}F_3N_2O_4$ + 0.5 Et_2O: C, 56.86; H, 4.02; N, 6.98; found (%): C, 56.66; H, 3,85; N, 6.63.



Benzylamide 3HQ 79. Grey solid, 86% yield. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.43 (s, 1H), 10.25 (s, 1H), 9.09 (t, J = 6.1 Hz, 1H), 7.50 – 7.18 (m, 8H), 4.51 (d, J = 6.0 Hz, 2H); **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 163.90, 158.23, 143.54, 143.19, 139.20, 132.01, 128.28, 127.22, 126.89, 121.86, 119.96, 119.78, 119.39, 117.04, 115.62, 42.52. **LRMS (m/z)**: calculated for [M+H$^+$] $C_{18}H_{14}F_3N_2O_4$: 379.1, found: 379.0. **Elemental analysis calculated** (%) for $C_{18}H_{13}F_3N_2O_4$ + 1,50 H_2O : C, 53.34; H, 3.98; N, 6.91; found (%): C, 53.65; H, 3.64; N, 6.96.

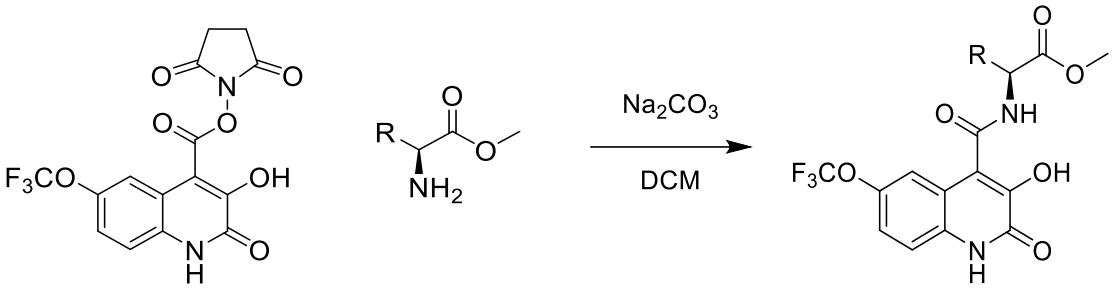


***N*-methyl Benzylamide 3HQ 80**. Grey solid, 89% yield. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** [mix of rotamers] δ 12.45 (s, 0.61H), 12.45 (s, 0.39H), 10.47 (s, 0.35H), 10.45 (s, 0.65H), 7.49 – 7.04 (m, 8H), 4.77 (dd, J = 50.2, 14.8 Hz, 1.3H), 4.43 (dd, J = 35.2, 15.6 Hz, 0.70H), 2.94 (s, 1H), 2.80 (s, 2H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 170.12, 164.57, 143.17, 138.15, 128.55, 127.51, 126.84, 125.20, 124.85, 122.07, 118.69, 115.34, 111.93, 111.18, 110.92, 53.86, 49.29, 34.89, 31.80; **HRMS (m/z)**: calculated for [M+Na$^+$] $C_{19}H_{16}F_3N_2O_4Na$: 415.08760, found: 415.08725.



Phenylethylamide 3HQ 81. Grey solid, 90% yield. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.42 (s, 1H), 8.71 (d, J = 6.9 Hz, 1H), 7.42 (d, J = 8.9 Hz, 1H), 7.37 – 7.17 (m, 6H), 7.14 (d, J = 2.6 Hz, 1H), 3.54 (q, J = 6.8 Hz, 2H), 2.84 (t, J = 7.2 Hz, 2H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 163.80, 158.35, 143.93, 143.13, 139.28, 131.85, 128.72, 128.27, 126.11, 121.87, 119.58, 119.49, 118.48, 116.88, 115.91, 35.06. **LRMS (m/z)**: calculated for [M+H$^+$] $C_{19}H_{16}F_3N_2O_4$: 393.3, found: 393.1. **Elemental analysis** calculated (%) for $C_{19}H_{15}F_3N_2O_4 + H_2O$: C, 55.61; H, 4.18; N, 6.83; found (%): C, 55.38; H, 3.84; N, 6.75.

- **Synthesis of 4-carboxaminoacid 3HQ derivatives 82-84**



107

82 R=$(CH_2)_2CO_2CH_3$

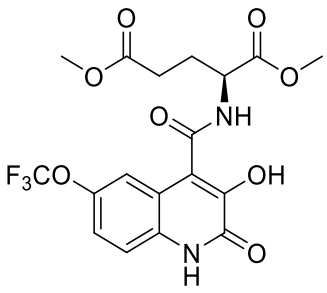
83 R=$CH_2CH(CH_3)_2$

84 R=H

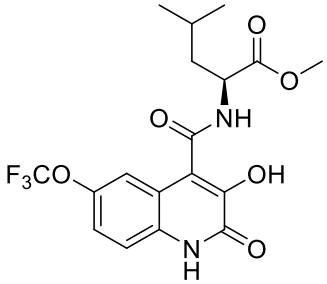
Scheme 103 – Synthetic route towards 4-carboxamide 3HQ derivatives 78-81

In a thick-glassed vial, 0.749 mmol of the appropriate amino acid ester were suspended in dry DCM (3.5 mL) and stirred with Na_2CO_3 (795 mg, 7.5 mmol),then intermediate **107** (241 mg, 0.625 mmol) was added and the mixture was stirred overnight at room temperature under argon atmosphere. The solvent was subsequently evaporated at reduced pressure, then the crude was suspended in 1 mL of water, to which aqueous HCl 1M was slowly added dropwise until pH 4 was reached. The resulting suspension was centrifuged, the

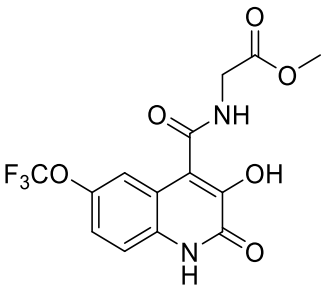
surnatant was removed and the resulting solid was washed again with water, then collected and dried in high vacuum.



Glutamic acid dimethyl ester 3HQ 82. Grey solid, 57% yield. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.43 (s, 1H), 10.21 (s, 1H), 9.04 (d, J = 7.0 Hz, 1H), 7.59 – 7.27 (m, 3H), 4.48 (d, J = 2.4 Hz, 1H), 3.70 (s, 3H), 3.60 (s, 3H), 2.50 (2H, overlapped with solvent peaks), 2.18 – 1.81 (m, 2H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 172.40, 171.53, 164.04, 157.90, 143.31, 142.97, 131.74, 119.75, 119.26, 119.17, 118.27, 116.67, 115.74, 51.76, 51.25, 51.19, 29.32, 25.33; **LRMS (m/z)**: calculated for [M+H$^+$]: $C_{18}H_{18}F_3N_2O_8$: 447.1, found: 447.0. **Elemental analysis** calculated (%) for: $C_{18}H_{17}F_3N_2O_8 + H_2O$: C, 44.82; H, 4.39; N, 5.81; O, 33.17; found (%):C, 44.44; H, 3.99; N, 6.05.



Leucine methyl ester 3HQ 83. Grey solid, 70% yield. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.40 (s, 1H), 10.12 (s, 1H), 8.98 (s, 1H), 7.38 (s, 3H), 4.47 (s, 1H), 3.69 (s, 3H), 1.83 – 1.44 (m, 3H), 0.91 (s, 6H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 174.05, 169.40, 164.33, 163.13, 143.10, 127.95, 124.94, 122.12, 115.95, 115.18, 113.39, 103.91, 51.59, 50.13, 24.54; **LRMS (m/z)**: calculated for [M+Na$^+$] $C_{18}H_{19}F_3N_2O_6Na$: 439.1, found: 439.0. **Elemental analysis** calculated (%) for: $C_{18}H_{19}F_3N_2O_6$ + 0,5 H_2O: C, 50.83; H, 4.74; N, 6.59; found (%):C, 49.46; H, 4.10; N, 6.25.



Glycine methyl ester 3HQ 84 Grey solid, 88% yield. **^{1}H NMR (300 MHz, $(CD_3)_2SO$):** δ 12.43 (s, 1H), 10.23 (s, 1H), 9.14 – 8.90 (m, 1H), 7.60 (m, 1H), 7.56 – 7.27 (m, 2H), 4.19 – 3.95 (m, 2H), 3.75 (s, 3H). **^{13}C NMR (75 MHz, $(CD_3)_2SO$):** δ 170.05, 164.64, 158.13, 143.51, 143.25, 131.98, 121.90, 119.95, 119.45, 118.51, 116.87, 116.28, 51.82, 40.89. **HRMS (m/z)**: calculated for [M-H$^+$] $C_{14}H_{10}F_3N_2O_6$: 359.04960, found: 359.04977.

V.3.2. Production and purification of recombinant hPAH

Recombinant hPAH was expressed in *E. coli* Top10 cells in fusion with a hexa-histidyl peptide to allow purification by immobilized metal affinity chromatography (IMAC) as described by Leandro and co-workers.[166] After IMAC purification, hPAH tetramers were isolated by size-exclusion chromatography (SEC) on a HiLoad Superdex 200 HR column (1.6 x 60 cm, GE Healthcare) equilibrated with 20 mM Na-HEPES buffer containing 200 mM NaCl, pH 7.0 (SEC buffer), at a flow rate of 0.7 mL·min-1, at 4°C.

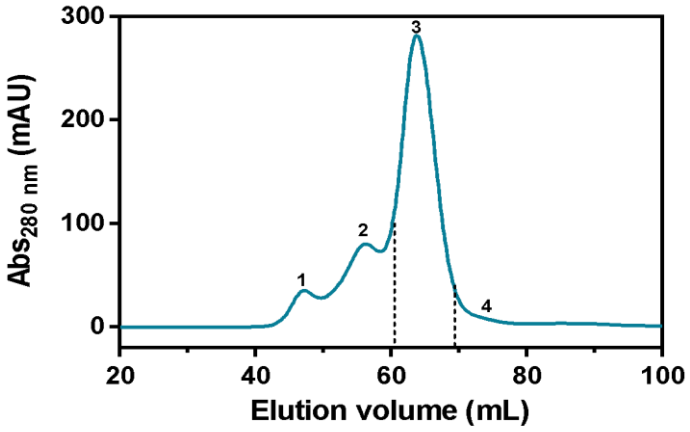


Figure 51 - Chromatographic profile of recombinant hPAH, obtained by size exclusion chromatography (SEC)

The resulting chromatogram, reported in **Figure 51**, presents 4 peaks: Peak 1 represents higher-order oligomeric forms (eluted in the void volume; 44-47 mL), peak 2 a presumably octameric form (52-56 mL), peak 3 the tetramer (~220 kDa; 56-66 mL) and peak 4 the dimer (~110 kDa; 69-71 mL). The apparent molecular mass of the enzyme forms were estimated using the elution position of standard molecular mass markers as a reference (not shown). The two dashed lines indicate the collected tetramers utilized for the assays.

V.3.3. Enzymatic activity assays

The hPAH activity was measured as previously described in the literature.[167] Briefly, the enzymatic reaction was performed in a 200 µL final volume, containing 100 mM Na-HEPES, pH 7, 0.1 mg.mL-1 catalase, 5 µg (final concentration of 0.112 µM of tetramer) of recombinant hPAH tetramers and 100 µM $(NH_4)2Fe(II)SO_4$. Three experimental conditions were used to study the compounds' modulatory effect, namely: Non-activated, Compound-activated, and Compound/Substrate-activated (**Figure 28**, **Chapter IV**), using 100 µM L-Phe and 100 µM of tested compound. The reaction was started by addition of 75 µM BH4 (prepared in 5 mM ascorbic acid). In the Non-activated assay, the L-Phe substrate and tested compounds were added together with 75 µM BH4 at the start of the hydroxylation reaction. In the Compound-activated condition, hPAH was pre-incubated with the tested compound for 4 min at 25°C and the reaction was started by addition of BH4 together with L-Phe. In the Compound/Substrate-activated condition, hPAH was pre-incubated simultaneously with L-Phe and the tested compound (4 min, 25°C). Appropriate control reactions were performed where the substrate and/or the compound were omitted.

After a 1-min incubation, the reaction was stopped and the amount of L-Tyr produced was quantified by HPLC. The specific hPAH enzymatic activity is expressed as nmol of L-Tyr produced during 1 min per mg of protein (nmol L-Tyr.min-1-mg-1). Three independent assays were always performed.

V.3.4. Thermal stability assays

Thermal denaturation profiles of recombinant hPAH were obtained by differential scanning fluorimetry (DSF) in a C1000 Touch thermal cycler equipped with a CFX96 optical reaction module (Bio Rad) as described in the literature.[167] Briefly, DSF assays were performed using 100 µg.mL-1 hPAH tetramers (final concentration of 0.45 µM of tetramer), SYPROTM Orange (Invitrogen; 5000-fold commercially available stock solution) at a 2.5-

x final concentration and ramping the temperature from 20 to 90°C, at 1 °C·min^{-1}. Fluorescence data were acquired using the FRET channel. Data were processed using CFX Manager Software V3.0 (Bio-Rad) and the GraphPad Prism 6. Temperature scan curves were fitted to a biphasic dose-response function and the melting temperature (Tm) values were obtained from the midpoint of the first (Tm1) and second transitions (Tm2) corresponding, respectively, to the thermal denaturation of the regulatory and catalytic domain.[168,169] The Tm values in the presence of 3HQ were compared to those obtained in the presence of 1% DMSO (vehicle control) and a change in the Tm was considered significant when the $|\Delta Tm| \geq 2$ °C.[170]

V.3.5. Docking studies

To study and rationalize the influence of different 3-hydroxyquinolin-2(1H)-one substituents on hPAH molecular recognition, compounds **76**, **81**, **82**, and **84** were submitted to molecular docking simulations. Compounds **76**, **81**, **82**, and **84** were built, protonated at pH = 7.4 and geometry optimized using MOE 2019.01.[171] All molecules were submitted to a conformational search and energy minimization as implemented in MOE 2019.1. Conformers were created by a systematic search generating rotatable bond angle combinations. Each dihedral angle combination was subject to an energy minimization by means of the Amber12:EHT force field. The global minimum of each molecule was submitted to docking simulations. For the docking simulation studies, the target models of hPAH were designed starting from Protein Data Bank (PDB) crystallographic structures 1MMT and 3PAH. The original PDB entries were selected taking into account the source organism (Homo sapiens), the best available X-ray resolution and the co-crystallized ligands. Before being used in the docking simulations, all the structures were submitted to the Protein Preparation tool of MOE 2019.01 software package, to remove the co-crystallized inhibitors, add hydrogen atoms at pH = 7.4, assign the correct protonation and tautomeric states, missing residues, and to remove all crystallographic water molecules.

All the docking simulations (non-covalent) were performed using GOLD 5.2 program (CCDC).[172] The hPAH binding site was defined at the iron centre in the protein catalytic site, and the binding site radius was set to 20 Å. No constraints were used in any calculation. Flexible ligand sampling was considered in the docking procedure. All other parameters were set to defaults for the GOLD docking process. Molecular docking studies

were then performed using the Goldscore scoring function from GOLD 5.2 software package and each ligand was subjected to 1000 docking runs. The docking protocol was validated for all the four structures prepared re-docking the crystallographic ligands and their poses were reproducible with root-mean-square deviations (RMSD´s) below 1.5 Å.

The calculated distances (Å) of the selected 3HQs to the catalytic non-heme iron and its ligands (His285, His290 and Glu330), and to specific amino acid residues that interact with catecholamines reversible inhibitors (Phe254, Tyr325 and Glu330) are reported in **Table 4**. The compounds were docked into the catalytic centre of hPAH (PDB ID: 3PAH)

	Fe^{2+}	**Phe254**	**His285**	**His290**	**Tyr325**	**Glu330**
76	2.24 (OH) 2.25 2.34	3.16	OH-H 2.7	HO-NH 1.99	OC-HO 2.99	CO-H3C 2.26
81	1.98 (OH) 1.73	3.13	OH-H 2.49	CO-NH 1.8	OC-HO 4.52	CO-OC 2.68
82	2.15 (OH) 2.26 2.38	3.04	OH-H 2.57	CO-NH 1.9	OC-HO 3.05	CO-H3C 2.14
84	1.73 2.25	3.54	CO-H 1.95	CO-NH 2.4	OC-HO 5.19	HO-O 2.66

Table 4 - Calculated distances (Å) of the selected 3HQs to the catalytic non-heme iron and its ligands and to specific amino acid residues that interact with catecholamines

The calculated distances (Å) of the selected 3HQs to specific amino acid residues, involved in cofactor (BH_4) and substrate (l-Phe) binding are shown in . The compounds were docked into the catalytic centre of hPAH (PDB ID: 3PAH)

		76	**81**	**82**	**84**
BH_4					
	Gly247	C(ar)-OC 4.6	C(ar)-OC 5.65	C(ar)-OC 5.09	C(ar)-OC 5.04
	Leu249	O -(HN) 2.69 F-(OC) 2.27	O -(HN) 4.62 F-(OC) 2.21	O -(HN) 3.39 F-(OC) 2.41	O -(HN) 2.84 F-(OC) 2.04
	Ser251	F -(OC) 4.76 F-OH 4.53	F -(OC) 4.91 F-OH 3.73	F -(OC) 4.85 F-OH 3.95	F -(OC) 4.92 F-OH 4.84
	Phe254	C(ar)-C(ar) 3.16	C(ar)-C(ar) 2.86	C(ar)-C(ar) 2.83	C(ar)-C(ar) 3.10
	Glu286	1.73	1.70	1.70	2.30
L-Phe					
	His285	OH-H 2.7	OH-H 2.49	OH-H 2.57	CO-H 1.95
	Ser349	-	-	-	-
	Arg270	-	-	-	-
	Thr278	-	-	-	-
	Tyr128	-	-	-	
	Pro281	CO-H2C 3.11 HO- H2C 3.13	CO-H2C 3.22 HO- H2C 3.13 CH-N 2.73	CO-H2C 3.37 HO- H2C 3.08	CO-H2C 3.36 HO- H2C 2.35
	Ala 322	-	-	-	-
	Trp326	C(ar)C(ar) 3.5-4.2	-	C(ar)-OCH3 2.5, 2.9	C(ar)-OC 3.17

	Phe331	H3C-H 3.7	-	H3C-H 3.3	-
	Ser250	F -(OC) 3.03 F -(NH) 3.35	F -(OC) 3.26 F -(NH) 4.44	F -(OC) 2.73 F -(NH) 3.87	F -(OC) 2.43 F -(NH) 3.35
	Gly346	H-CH3 3.4	-	H-CH3 3.20	
	Ser350	-	-	-	-

Table 5 - Calculated distances (Å) of the selected 3HQs to specific amino acid residues, involved in cofactor (BH_4) and substrate (l-Phe) binding

Bibliographical references

[1] L. Ciani, S. Ristori, *Expert Opin. Drug Discov.* **2012**, *7*, 1017–1027.

[2] N. S. Hosmane, *Boron Science: New Technologies and Applications*, Taylor & Francis, Boca Raton, **2012**.

[3] D. B. Diaz, A. K. Yudin, *Nat. Chem.* **2017**, *9*, 731–742.

[4] W. Yang, X. Gao, B. Wang, *Med. Res. Rev.* **2003**, *23*, 346–368.

[5] "The Merck Index Online," can be found under https://www.rsc.org/Merck-Index, **2019**.

[6] S. J. Baker, J. W. Tomsho, S. J. Benkovic, *Chem. Soc. Rev.* **2011**, *40*, 4279–4285.

[7] E. Frankland, B. F. Duppa, *Justus Liebigs Ann. Chem.* **1860**, *115*, 319–322.

[8] D. G. Hall, in *Boronic Acids*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, FRG, **2006**, pp. 1–99.

[9] P. V. Ramachandran, *Future Med. Chem.* **2013**, *5*, 611–612.

[10] J. W. B. Fyfe, A. J. B. Watson, *Chem* **2017**, *3*, 31–55.

[11] S. Lascano, K.-D. Zhang, R. Wehlauch, K. Gademann, N. Sakai, S. Matile, *Chem. Sci.* **2016**, *7*, 4720–4724.

[12] N. Fujita, S. Shinkai, T. D. James, *Chem. – An Asian J.* **2008**, *3*, 1076–1091.

[13] R. Smoum, A. Rubinstein, V. M. Dembitsky, M. Srebnik, *Chem. Rev.* **2012**, *112*, 4156–4220.

[14] R. F. Barth, Z. Zhang, T. Liu, *Cancer Commun.* **2018**, *36*, 1–7.

[15] K. Ono, *Ther. Radiol. Oncol.* **2018**, *2*, 1–7.

[16] K. Nedunchezhian, N. Aswath, M. Thiruppathy, S. Thirugnanamurthy, *J. Clin. Diagnostic Res.* **2016**, *10*, 1–4.

[17] P. R. Hamann, L. M. Hinman, I. Hollander, C. F. Beyer, D. Lindh, R. Holcomb, W. Hallett, H. R. Tsou, J. Upeslacis, D. Shochat, et al., *Bioconjug. Chem.* **2002**, *13*, 47–58.

[18] P. D. Senter, E. L. Sievers, *Nat. Biotechnol.* **2012**, *30*, 631–637.

[19] K. A. Poon, K. Flagella, J. Beyer, J. Tibbitts, S. Kaur, O. Saad, J. H. Yi, S. Girish, N. Dybdal, T. Reynolds, *Toxicol. Appl. Pharmacol.* **2013**, *273*, 298–313.

[20] S. Parakh, A. C. Parslow, H. K. Gan, A. M. Scott, *Expert Opin. Drug Deliv.* **2016**, *13*, 401–419.

[21] J. Anderl, H. Faulstich, T. Hechler, M. Kulke, in (Ed.: L. Ducry), Humana Press, Totowa, NJ, **2013**, pp. 51–70.

[22] I. L. Medintz, H. Uyeda, E. R. Goldman, H. Mattoussi, *Quantum Dot Bioconjugates for Imaging, Labelling and Sensing*, **2006**.

[23] T. Tamura, I. Hamachi, *J. Am. Chem. Soc.* **2019**, *141*, 2782–2799.

[24] J. M. Chalker, G. J. L. Bernardes, Y. A. Lin, B. G. Davis, *Chem. - An Asian J.* **2009**, *4*, 630–640.

[25] J. M. J. M. Ravasco, H. M. F. Faustino, A. Trindade, P. M. P. Gois, *Chem. - A Eur. J.* **2019**, *25*, 43–59.

[26] N. Stephanopoulos, M. B. Francis, *Nat. Chem. Biol.* **2011**, *7*, 876–884.

[27] E. A. Hoyt, P. M. S. D. Cal, B. L. Oliveira, G. J. L. Bernardes, *Nat. Rev. Chem.* **2019**, DOI https://doi.org/10.1038/s41570-019-0079-1.

[28] G. Leriche, L. Chisholm, A. Wagner, *Bioorganic Med. Chem.* **2012**, *20*, 571–582.

[29] A. Dal Corso, S. Cazzamalli, R. Gébleux, M. Mattarella, D. Neri, *Bioconjug. Chem.* **2017**, *28*, 1826–1833.

[30] W. K. Chow, O. Y. Yuen, P. Y. Choy, C. M. So, C. P. Lau, W. T. Wong, F. Y. Kwong, *RSC Adv.* **2013**, *3*, 12518–12539.

[31] B. C. Das, P. Thapa, R. Karki, C. Schinke, S. Das, S. Kambhampati, S. K. Banerjee, P. Van Veldhuizen, A. Verma, L. M. Weiss, et al., *Future Med. Chem.* **2013**, *5*, 653–676.

[32] P. C. Trippier, C. McGuigan, *Medchemcomm* **2010**, *1*, 183–198.

[33] “Protein Data Bank,” can be found under https://www.rcsb.org/, **n.d.**

[34] E. Kupperman, E. C. Lee, Y. Cao, B. Bannerman, M. Fitzgerald, A. Berger, J. Yu, Y. Yang, P. Hales, F. Bruzzese, et al., *Cancer Res.* **2010**, *70*, 1970–1980.

[35] W. W. Bachovchin, W. Y. L. Wong, S. Farr-jones, A. B. Shenvi, C. A. Kettner, *Biochemistry* **1988**, *27*, 7689–7697.

[36] E. Tsilikounas, C. A. Kettner, W. W. Bachovchin, *Biochemistry* **1993**, *32*, 12651–12655.

[37] E. Tsilikounas, C. Kettner, W. W. Bachovchin, *Biochemistry* **1992**, *31*, 12839–12846.

[38] R. Bone, D. Frank, C. A. Kettner, D. A. Agard, *Biochemistry* **1989**, *28*, 7600–7609.

[39] J. Adams, M. Kauffman, *Cancer Invest.* **2004**, *22*, 304–311.

[40] H. Avet-Loiseau, N. J. Bahlis, W. J. Chng, T. Masszi, L. Viterbo, L. Pour, P. Ganly, A. Palumbo, M. Cavo, C. Langer, et al., *Blood* **2017**, *130*, 2610–2618.

[41] B. E. Elewski, R. Aly, S. L. Baldwin, R. F. González Soto, P. Rich, M. Weisfeld, H.

Wiltz, L. T. Zane, R. Pollak, *J. Am. Acad. Dermatol.* **2015**, *73*, 62–69.

[42] T. Akama, S. J. Baker, Y.-K. Zhang, V. Hernandez, H. Zhou, V. Sanders, Y. Freund, R. Kimura, K. R. Maples, J. J. Plattner, *Bioorg. Med. Chem. Lett.* **2009**, *19*, 2129–2132.

[43] S. J. Hecker, K. R. Reddy, M. Totrov, G. C. Hirst, O. Lomovskaya, D. C. Griffith, P. King, R. Tsivkovski, D. Sun, M. Sabet, et al., *J. Med. Chem.* **2015**, *58*, 3682–3692.

[44] G. Springsteen, B. Wang, *Tetrahedron* **2002**, *58*, 5291–5300.

[45] M. P. Curran, K. McKeage, *Drugs* **2009**, *69*, 859–888.

[46] M. Shirley, *Drugs* **2016**, *76*, 405–411.

[47] J. P. M. Antonio, G. D. V Farias, F. M. F. Santos, R. Oliveira, P. M. S. D. Cal, P. M. P. Gois, in *Non-covalent Interact. Synth. Des. New Compd.*, John Wiley & Sons, Ltd, **2016**, pp. 23–48.

[48] N. Aykin-Burns, I. M. Ahmad, Y. Zhu, L. W. Oberley, D. R. Spitz, *Biochem. J.* **2010**, *418*, 319–335.

[49] L. Wang, S. Xie, L. Ma, Y. Chen, W. Lu, *Eur. J. Med. Chem.* **2016**, *116*, 84–89.

[50] D. Lee, S. Park, S. Bae, D. Jeong, M. Park, C. Kang, W. Yoo, M. A. Samad, Q. Ke, G. Khang, et al., *Sci. Rep.* **2015**, *5*, 1–13.

[51] L. Du, M. Li, S. Zheng, B. Wang, *Tetrahedron Lett.* **2008**, *49*, 3045–3048.

[52] G. Cao, W. Mao, J. Han, Z. Zhao, M. Gao, X. Xu, S. Wang, H. Ye, C. Chu, *Bioorg. Chem.* **2018**, *81*, 362–366.

[53] K. Xu, L. He, X. Yang, Y. Yang, W. Lin, *Analyst* **2018**, *143*, 3555–3559.

[54] D. Andina, J. C. Leroux, P. Luciani, *Chem. - A Eur. J.* **2017**, *23*, 13549–13573.

[55] Y. Liu, L. Bai, Y. Li, Y. Ni, C. Xin, C. Zhang, J. Liu, Z. Liu, L. Li, W. Huang, *Sensors Actuators, B Chem.* **2019**, *279*, 38–43.

[56] L. Zhang, M. Qian, J. Xia, L. Chen, H. Cui, Y. Xia, L. Xu, J. Wang, X. Peng, *J. Photochem. Photobiol. A Chem.* **2019**, *370*, 12–17.

[57] A. E. Hargrove, R. N. Reyes, I. Riddington, E. V Anslyn, J. L. Sessler, *Org. Lett.* **2010**, *12*, 4804–4807.

[58] L. Du Ý, N. Ni Ý, M. Li, B. Wang, *Tetrahedron Lett.* **2010**, *51*, 1152–1154.

[59] X. Q. Zhan, B. Y. Su, H. Zheng, J. H. Yan, *Anal. Chim. Acta* **2010**, *658*, 175–179.

[60] E. V. Lampard, A. C. Sedgwick, X. Sun, K. L. Filer, S. C. Hewins, G. Kim, J. Yoon, S. D. Bull, T. D. James, *ChemistryOpen* **2018**, *7*, 262–265.

[61] Y. Liu, J. Niu, J. Nie, F. Meng, W. Lin, *New J. Chem.* **2017**, *41*, 3320–3325.

[62] D. Srikun, E. W. Miller, D. W. Domaille, C. J. Chang, *J. Am. Chem. Soc.* **2008**, *130*, 4596–4597.

[63] A. R. Lippert, T. Gschneidtner, C. J. Chang, *Chem. Commun.* **2010**, *46*, 7510–7512.

[64] S. Lu, C. Jia, H. Huang, J. Tang, Y. Han, *J. Fluoresc.* **2015**, *26*, 121–127.

[65] E. Lindberg, N. Winssinger, *ChemBioChem* **2016**, *1*, 1612–1615.

[66] J. Xu, J. Zhai, Y. Xu, J. Zhu, Y. Qin, D. Jiang, *Analyst* **2016**, *141*, 2380–2383.

[67] X. Shi, Z. Wang, X. Wang, Y. Chen, Z. Lu, *Anal. Chem.* **2017**, *89*, 5278–5284.

[68] J. L. M. Jourden, S. M. Cohen, *Angew. Chemie Int. Ed.* **2010**, *49*, 6795–6797.

[69] E. J. Kim, S. Bhuniya, H. Lee, H. M. Kim, C. Cheong, S. Maiti, K. S. Hong, J. S. Kim, *J. Am. Chem. Soc.* **2014**, *136*, 13888–13894.

[70] G. A. Ellis, M. J. Palte, R. T. Raines, *J. Am. Chem. Soc.* **2012**, *134*, 3631–3634.

[71] K. A. Andersen, T. P. Smith, J. E. Lomax, R. T. Raines, *ACS Chem. Biol.* **2016**, *11*, 319–323.

[72] S. D. Bull, M. G. Davidson, J. M. H. van den Elsen, J. S. Fossey, A. T. A. Jenkins, Y.-B. Jiang, Y. Kubo, F. Marken, K. Sakurai, J. Zhao, et al., *Acc. Chem. Res.* **2013**, *46*, 312–326.

[73] G. Springsteen, B. Wang, *Tetrahedron* **2002**, *58*, 5291–5300.

[74] H. G. Kuivila, A. H. Keough, E. J. Soboczenski, *J. Org. Chem.* **1954**, *19*, 780–783.

[75] A. Finch, P. J. Gardner, P. M. McNamara, G. R. Wellum, *J. Chem. Soc. A Inorganic, Phys. Theor.* **1970**, *0*, 3339–3345.

[76] W. L. A. Brooks, C. C. Deng, B. S. Sumerlin, *ACS Omega* **2018**, *3*, 17863–17870.

[77] J. P. Lorand, J. O. Edwards, *J. Org. Chem.* **1959**, *24*, 769–774.

[78] W. Zhai, X. Sun, T. D. James, J. S. Fossey, *Chem. – An Asian J.* **2015**, *10*, 1836–1848.

[79] T. D. James, M. D. Phillips, S. Shinkai, *Boronic Acids in Saccharide Recognition*, The Royal Society Of Chemistry, **2006**.

[80] D. G. Hall, *Structure, Properties, and Preparation of Boronic Acid Derivatives. Overview of Their Reactions and Applications*, John Wiley & Sons, Ltd, **2006**.

[81] B. Pappin, M. J. Kiefel, T. A. Houston, in *Compr. Stud. Glycobiol. Glycotechnol.*, IntechOpen, **2012**, pp. 37–54.

[82] X. Wu, Z. Li, X.-X. Chen, J. S. Fossey, T. D. James, Y.-B. Jiang, *Chem. Soc. Rev.* **2013**, *42*, 8032–8048.

[83] X. Wu, X.-X. Chen, Y.-B. Jiang, *Analyst* **2017**, *142*, 1403–1414.

[84] W. J. Ramsay, H. Bayley, *Angew. Chemie Int. Ed.* **2018**, *57*, 2841–2845.

[85] Y.-L. Yang, Y.-P. Lee, Y.-L. Yang, P.-C. Lin, *ACS Chem. Biol.* **2014**, *9*, 390–397.

[86] T. L. Halo, J. Appelbaum, E. M. Hobert, D. M. Balkin, A. Schepartz, *J. Am. Chem. Soc.* **2009**, *131*, 438–439.

[87] E. Montanari, A. Gennari, M. Pelliccia, L. Manzi, R. Donno, N. J. Oldham, A. Macdonald, N. Tirelli, *Bioconjug. Chem.* **2018**, *29*, 2550–2560.

[88] M. E. Kimple, A. L. Brill, R. L. Pasker, *Curr. Protoc. protein Sci.* **2013**, *73*, Unit-9.9.

[89] M. Rosenberg, J. L. Wiebers, P. T. Gilham, *Biochemistry* **1972**, *11*, 3623–3628.

[90] P. R. Westmark, B. D. Smith, *J. Pharm. Sci.* **1996**, *85*, 266–269.

[91] R. Tuytten, F. Lemière, W. Van Dongen, E. Witters, E. L. Esmans, R. P. Newton, E. Dudley, *Anal. Chem.* **2008**, *80*, 1263–1271.

[92] A. R. Martin, J.-J. Vasseur, M. Smietana, *Chem. Soc. Rev.* **2013**, *42*, 5684–5713.

[93] M. Naito, T. Ishii, A. Matsumoto, K. Miyata, Y. Miyahara, K. Kataoka, *Angew. Chemie Int. Ed.* **2012**, *51*, 10751–10755.

[94] E. Brustad, M. L. Bushey, J. W. Lee, D. Groff, W. Liu, P. G. Schultz, *Angew. Chemie Int. Ed.* **2008**, *47*, 8220–8223.

[95] C. C. Liu, A. V Mack, E. M. Brustad, J. H. Mills, D. Groff, V. V Smider, P. G. Schultz, *J. Am. Chem. Soc.* **2009**, *131*, 9616–9617.

[96] F. Wang, W. Niu, J. Guo, P. G. Schultz, *Angew. Chemie Int. Ed.* **2012**, *51*, 10132–10135.

[97] Z. J. Chen, W. Ren, Q. E. Wright, H. W. Ai, *J. Am. Chem. Soc.* **2013**, *135*, 14940–14943.

[98] B. Akgun, D. G. Hall, *Angew. Chemie Int. Ed.* **2016**, *55*, 3909–3913.

[99] P. M. S. D. Cal, J. B. Vicente, E. Pires, A. V. Coelho, L. F. Veiros, C. Cordeiro, P. M. P. Gois, *J. Am. Chem. Soc.* **2012**, *134*, 10299–10305.

[100] P. Akkapeddi, S.-A. Azizi, A. M. Freedy, P. M. S. D. Cal, P. M. P. Gois, G. J. L. Bernardes, *Chem. Sci.* **2016**, *7*, 2954–2963.

[101] J. M. McFarland, M. B. Francis, *J. Am. Chem. Soc.* **2005**, *127*, 13490–13491.

[102] V. Raindlová, R. Pohl, M. Hocek, *Chem. - A Eur. J.* **2012**, *18*, 4080–4087.

[103] B. M. Chapin, P. Metola, V. M. Lynch, J. F. Stanton, T. D. James, E. V Anslyn, *J. Org. Chem.* **2016**, *81*, 8319–8330.

[104] P. M. S. D. Cal, R. F. M. Frade, V. Chudasama, C. Cordeiro, S. Caddick, P. M. P. Gois, *Chem. Commun.* **2014**, *50*, 5261–5263.

[105] P. M. S. D. Cal, R. F. M. Frade, C. Cordeiro, P. M. P. Gois, *Chem. – A Eur. J.* **2015**, *21*, 8182–8187.

[106] G. Akçay, M. A. Belmonte, B. Aquila, C. Chuaqui, A. W. Hird, M. L. Lamb, P. B. Rawlins, N. Su, S. Tentarelli, N. P. Grimster, et al., *Nat. Chem. Biol.* **2016**, *12*, 931–936.

[107] A. Bandyopadhyay, J. Gao, *J. Am. Chem. Soc.* **2016**, *138*, 2098–2101.

[108] R. M. R. M. Lopes, A. E. Ventura, L. C. Silva, H. Faustino, P. M. P. Gois, *Chem. - A Eur. J.* **2018**, *24*, 12495–12499.

[109] K. G. Cunningham, G. G. Freeman, *Biochem. J.* **1953**, *53*, 328–332.

[110] J. H. Birkinshaw, M. Luckner, Y. S. Mohammed, K. Mothes, C. E. Stickings, *Biochem. J.* **1963**, *89*, 196–202.

[111] M. Luckner, K. Mothes, *Tetrahedron Lett.* **1962**, *3*, 1035–1039.

[112] M. Y. Wei, R. Y. Yang, C. L. Shao, C. Y. Wang, D. S. Deng, Z. G. She, Y. C. Lin, **2011**, *47*, 322–325.

[113] C. B. M. Poulie, L. Bunch, *ChemMedChem* **2013**, *8*, 205–215.

[114] C. Ballatore, D. M. Huryn, A. B. Smith, *ChemMedChem* **2013**, *8*, 385–395.

[115] H. Y. Sagong, A. Parhi, J. D. Bauman, D. Patel, R. S. K. Vijayan, K. Das, E. Arnold, E. J. Lavoie, **2013**, *2*, 8–11.

[116] A. J. Duplantier, S. L. Becker, M. J. Bohanon, K. A. Borzilleri, B. A. Chrunyk, J. T. Downs, L. Y. Hu, A. El-Kattan, L. C. James, S. Liu, et al., *J. Med. Chem.* **2009**, *52*, 3576–3585.

[117] G. Springsteen, B. Wang, *Chem. Commun.* **2001**, 1608–1609.

[118] N. R. Candeias, R. Paterna, P. M. P. Gois, *Chem. Rev.* **2016**, *116*, 2937–2981.

[119] R. Paterna, V. André, M. T. Duarte, L. F. Veiros, N. R. Candeias, P. M. P. Gois, *European J. Org. Chem.* **2013**, 6280–6290.

[120] M. P. Doyle, A. V. Kalinin, *J. Org. Chem.* **1996**, *61*, 2179–2184.

[121] A. Gennari, C. Gujral, E. Hohn, E. Lallana, F. Cellesi, N. Tirelli, D. Chimica, I. Chimica, G. N. Politecnico, V. Mancinelli, et al., **n.d.**, 1–10.

[122] J. Nelson, N. V. McFerran, G. Pivato, E. Chambers, C. Doherty, D. Steele, D. J. Timson, *Biosci. Rep.* **2008**, *28*, 33.

[123] R. Gopalakrishna, U. Gundimeda, S. Zhou, H. Bui, A. Davis, T. McNeill, W. Mack, *Biochem. Biophys. Res. Commun.* **2018**, *495*, 230–237.

[124] V. Cioce, V. Castronovo, B. M. Shmookler, S. Garbisa, W. F. Grigioni, L. A. Liotta,

M. E. Sobel, *J. Natl. Cancer Inst.* **1991**, *83*, 29–36.

[125] V. P. Terranova, C. N. Rao, T. Kalebic, I. M. Margulies, L. a Liotta, *Proc. Natl. Acad. Sci.* **1983**, *80*, 444–8.

[126] M. Karpatová, E. Tagliabue, V. Castronovo, A. Magnifico, E. Ardini, D. Morelli, D. Belotti, M. I. Colnaghi, S. Ménard, *J. Cell. Biochem.* **1996**, *60*, 226–234.

[127] S. Cambray, J. Gao, *Acc. Chem. Res.* **2018**, *51*, 2198–2206.

[128] Bruker Corporation, **2012**.

[129] J. P. M. António, L. M. Gonçalves, R. C. Guedes, R. Moreira, P. M. P. Gois, *ACS Omega* **2018**, *3*, 7418–7423.

[130] A. D. Corso, L. Pignataro, L. Belvisi, C. Gennari, *Curr. Top. Med. Chem.* **2016**, *16*, 314–329.

[131] M. S. Bhojani, R. Ranga, G. D. Luker, A. Rehemtulla, B. D. Ross, M. E. Van Dort, *PLoS One* **2011**, *6*, e22418.

[132] G. Casi, D. Neri, *Mol. Pharm.* **2015**, *12*, 1880–1884.

[133] C. L. Beaulieu, M. E. Samuels, S. Ekins, C. R. Mcmaster, A. M. Edwards, A. R. Krainer, G. G. Hicks, B. J. Frey, K. M. Boycott, A. E. Mackenzie, **2012**, 1–11.

[134] L. Matalonga, **2017**, 177–193.

[135] N. Blau, W. Yue, B. Perez, “PAHvdb: Phenylalanine Hydroxylase Gene Locus-Specific Database,” can be found under http://www.biopku.org/home/pah.asp, **2019**.

[136] N. Blau, F. J. Van Spronsen, H. L. Levy, **2010**, 1417–1427.

[137] S. E. Christ, S. E. Christ, **2010**, *5213*.

[138] J. J. Mitchell, Y. J. Trakadis, C. R. Scriver, *Genet. Med.* **2011**, *13*, 697.

[139] F. Guttler, *Acta Paediatr. Scand. Suppl.* **1980**, *280*, 1–80.

[140] E. C. Arturo, K. Gupta, A. Héroux, L. Stith, P. J. Cross, E. J. Parker, P. J. Loll, E. K. Jaffe, *Proc. Natl. Acad. Sci.* **2016**, *113*, 2394–2399.

[141] S. P. Meisburger, A. B. Taylor, C. A. Khan, S. Zhang, P. F. Fitzpatrick, N. Ando, *J. Am. Chem. Soc.* **2016**, *138*, 6506–6516.

[142] D. Patel, J. Kopec, F. Fitzpatrick, T. J. McCorvie, W. W. Yue, *Sci. Rep.* **2016**, *6*, 23748.

[143] E. K. Jaffe, *Mol. Genet. Metab.* **2017**, *121*, 289–296.

[144] A. Martínez, A. Calvo, K. Teigen, A. Pey, *Chapter 3 Rescuing Proteins of Low Kinetic Stability by Chaperones and Natural Ligands*, **2008**.

[145] A. L. Pey, B. Pérez, L. R. Desviat, M. A. Martínez, C. Aguado, H. Erlandsen, A. Gámez, R. C. Stevens, M. Thórólfsson, M. Ugarte, et al., *Hum. Mutat.* **2004**, *24*, 388–399.

[146] D. Dunja, A. Katharina, **2010**, *19*, 2039–2049.

[147] H. Erlandsen, T. Flatmark, R. C. Stevens, E. Hough, *Biochemistry* **1998**, *37*, 15638–15646.

[148] B. Alcaide, P. Almendros, C. Aragoncillo, G. Gómez-Campillos, M. Arnó, L. R. Domingo, *Chempluschem* **2012**, *77*, 563–569.

[149] P. J. Manley, M. T. Bilodeau, *Org. Lett.* **2004**, *6*, 2433–2435.

[150] M. Mukherjee, A. K. Gupta, Z. Lu, Y. Zhang, W. D. Wulff, *J. Org. Chem.* **2010**, *75*, 5643–5660.

[151] C. J. Blankley, F. J. Sauter, H. O. House, J. H. Ham, R. E. Ireland, *Org. Synth.* **1969**, *49*, 22.

[152] F. Liu, A. S. Y. Ni, Y. Lim, H. Mohanram, S. Bhattacharjya, B. Xing, *Bioconjug. Chem.* **2012**, *23*, 1639–1647.

[153] A. Gennari, C. Gujral, E. Hohn, E. Lallana, F. Cellesi, N. Tirelli, *Bioconjug. Chem.* **2017**, *28*, 1391–1402.

[154] J. . F. D. J. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenb, *Gaussian Inc., Wallingford, …* **2010**.

[155] Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.* **2008**, *120*, 215–241.

[156] Y. Zhao, D. G. Truhlar, *Chem. Phys. Lett.* **2011**, *502*, 1–13.

[157] M. Cossi, V. Barone, B. Mennucci, J. Tomasi, *Chem. Phys. Lett.* **1998**, *286*, 253–260.

[158] J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev.* **2005**, *105*, 2999–3094.

[159] B. Mennucci, J. Tomasi, *J. Chem. Phys.* **1997**, *106*, 5151–5158.

[160] A. V Marenich, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem. B* **2009**, *113*, 6378–6396.

[161] B. Bernardim, P. M. S. D. Cal, M. J. Matos, B. L. Oliveira, N. Martínez-Saéz, I. S. Albuquerque, E. Perkins, F. Corzana, A. C. B. Burtoloso, G. Jiménez-Osés, et al., *Nat. Commun.* **2016**, *7*, 1–9.

[162] V. Rauniyar, D. G. Hall, *J. Org. Chem.* **2009**, *74*, 4236–4241.

[163] Z.-N. Sun, H.-L. Wang, F.-Q. Liu, Y. Chen, P. K. H. Tam, D. Yang, *Org. Lett.* **2009**, *11*, 1887–1890.

[164] D. Ding, Y. Zhao, Q. Meng, D. Xie, B. Nare, D. Chen, C. J. Bacchi, N. Yarlett, Y.-K. Zhang, V. Hernandez, et al., *ACS Med. Chem. Lett.* **2010**, *1*, 165–169.

[165] H. Faustino, M. J. S. A. Silva, L. F. Veiros, G. J. L. Bernardes, P. M. P. Gois, *Chem. Sci.* **2016**, *7*, 5052–5058.

[166] C. Nascimento, J. Leandro, P. R. Lino, L. Ramos, A. J. Almeida, I. T. de Almeida, P. Leandro, *Appl. Biochem. Biotechnol.* **2010**, *162*, 192–207.

[167] F. Montalbano, J. Leandro, G. D. V. F. Farias, P. R. Lino, R. C. Guedes, J. B. Vicente, P. Leandro, P. M. P. Gois, *RSC Adv.* **2014**, *4*, 61022–61027.

[168] M. Thórólfsson, B. Ibarra-Molero, P. Fojan, S. B. Petersen, J. M. Sanchez-Ruiz, A. Martínez, *Biochemistry* **2002**, *41*, 7573–7585.

[169] S. W. Gersting, K. F. Kemter, M. Staudigl, D. D. Messing, M. K. Danecka, F. B. Lagler, C. P. Sommerhoff, A. A. Roscher, A. C. Muntau, *Am. J. Hum. Genet.* **2008**, *83*, 5–17.

[170] G. A. Senisterra, P. J. Finerty, Jr, *Mol. BioSyst.* **2009**, *5*, 217–223.

[171] C. C. G. Inc., **2012**.

[172] G. Jones, P. Willett, R. C. Glen, A. R. Leach, R. Taylor, *J. Mol. Biol.* **1997**, *267*, 727–748.