Novel Conjugates of Porphyrins and Peptides as Light-Responsive Molecules and Materials

PAOLO DOGNINI

A thesis submitted in partial fulfilment of the requirements of Liverpool John Moores University for the degree of Doctor of Philosophy

This research programme was carried out in collaboration with, UA Doctoral Training Alliance, The University of Edinburgh and Università degli Studi di Parma

December 2023

List of contents

Abstract

The conjugation of biological molecules with functional small molecules is a common strategy in medicinal and biological chemistry to obtain compounds with improved properties and additional functionalities.

In this work, the conjugation of peptides and porphyrins was studied. The thiol-fluoride Nucleophilic Aromatic Substitution (S_NAr) was selected as conjugation tool because cysteine can be naturally present or easily introduced in peptide sequences, whilst *meso*-derivatised porphyrins with polyfluorinated rings are synthetically accessible.

Initially, in Chapter 2, the reaction was optimised using N-acetyl cysteine and hexafluorobenzene as a model scaffold. Different reaction conditions were also explored on model peptides, and their compatibility with the starting reagents was tested. Two new base/solvent combinations for peptide stapling on hexafluorobenzene were identified and one of these proved to be a robust bioconjugation tool for peptides and porphyrins as well.

Subsequently, in Chapter 3, bolaamphiphilic peptides with self-assembling behaviour were designed to include a cysteine and allow porphyrin coupling. When combined, porphyrin and self-assembling peptides can give origin to supramolecular systems with regular morphology and ordered arrangement that enhance the photochemical properties of embedded chromophores. The bola-porphyrin conjugates formed supramolecular structures under different conditions of concentration, pH, temperature, ionic strength, and time.

Finally, in Chapter 4, it was observed that fully unprotected peptides containing two or more cysteines can cyclise by reacting twice or further with the same porphyrin. Besides the advantages of constrained peptides over their linear counterparts, the integration of a functional scaffold within the cyclic architecture is a step forward for the development of novel multifunctional molecules. This technique provided controlled access to cyclic, bicyclic, and tricyclic peptide systems that could be used as mimics of natural machineries, molecular probes, light-activated drugs, and biomarker detection tools.

In conclusion, mastery of a novel bioconjugation strategy for peptides and porphyrins was demonstrated. The reaction occurs in solution with unprotected peptides and using peptide-compatible conditions, but can also be performed on-resin under specific circumstances. This indicates chemoselectivity, robustness and wide applicability. Exploiting this synthetic approach, light-responsive soft materials were obtained from the self-assembly of porphyrin and bolaamphiphilic peptides and polyfluorinated were introduced porphyrin as a versatile scaffold for the cyclisation of bioactive peptides.

Declaration

The work presented in this thesis was carried out at Liverpool John Moores University (School of Pharmacy and Biomolecular Sciences, Byrom Street City Campus), at The University of Edinburgh (EaStChem School of Chemistry, Joseph Black Building) and at the University of Parma (Department of Mathematical, Computer and Physical Sciences). Unless otherwise stated, the work is the original work of the author.

While registered as a candidate for the degree of Doctor of Philosophy, for which submission is now made, the author has not been registered as a candidate for any other award. This thesis has not been submitted in whole, or in part, for any other degree by the candidate.

Acknowledgements

Publications

- 5,10,15,20-Tetrakis(pentafluorophenyl)porphyrin as Functional Platform for Peptide Stapling and Multicyclisation

  Paolo Dognini, Talhat Chaudhry, Giulia Scagnetti, Michele Assante, George S. M. Hanson, Kehinde Ross, Francesca Giuntini, Christopher R. Coxon

  Chemistry – A European Journal, **2023**, ISSN 0947-6539

  https://doi.org/10.1002/chem.202301410

- Peptide-Tetrapyrrole Supramolecular Self-Assemblies: State of the Art

  Paolo Dognini, Christopher R. Coxon, Wendel A. Alves, Francesca Giuntini

  Molecules, **2021**, Volume 26, Issue 3, Pages 693-716, ISSN 1420-3049

  https://doi.org/10.3390/molecules26030693

- PepTherDia: database and structural composition analysis of approved peptide therapeutics and diagnostics

  Vera D'Aloisio, Paolo Dognini, Gillian A. Hutcheon, Christopher R. Coxon

  Drug Discovery Today, **2021**, Volume 26, Issue 6, Pages 1409-1419, ISSN 1359-6446

  https://doi.org/10.1016/j.drudis.2021.02.019

- Using ^{19}F NMR and two-level factorial design to explore thiol-fluoride substitution in hexafluorobenzene and its application in peptide stapling and cyclisation

  Paolo Dognini, Patrick M. Killoran, George S. Hanson, Lewis Halsall, Talhat Chaudhry, Zasharatul Islam, Francesca Giuntini, Christopher R. Coxon

  Peptide Science, **2021**, Volume 113, Issue 1, ISSN 2475-8817

  https://doi.org/10.1002/pep2.24182

Abbreviations

4F-2CN	2,3,5,6-tetrafluoroterephthalonitrile
AA	amino acid
Acm	acetamidomethyl
Ahx	6-aminohexanoic acid
AI	absorption intensities
AMP	antimicrobial peptides
ANOVA	analysis of variance
BODIPY	boron dipyrromethene
Calu-3	non-small-cell lung cancer cell line
CD	circular dichroism
Cha	cyclohexyl alanine
CLIPS	chemical linkage of peptides onto scaffolds
CuAAC	copper catalysed azide-alkyne cycloaddition
DAPI	4′,6-diamidino-2-phenylindole
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCU	dicyclohexylurea
DIC	diisopropilcarbodiimmide
DIPEA	N,N-diisopropylethylamine
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DOE	design of experiment
Dpm	1,2-diphenylmaleyl
DSSC	dye sensitised solar cells
EWG	electron-withdrawing groups
FD	factorial design
FI	fluorescence intensities
FITC	fluorescein isothiocyanate
Fmoc	9-fluorenylmethyloxycarbonyl
GSH	glutathione

HFB	hexafluorobenzene
HFIP	hexafluoroisopropanol
HPLC	high performance liquid chromatography
IUPAC	International Union of Pure and Applied Chemistry
LC-MS	liquid chromatography–mass spectrometry
LFP	laser flash photolysis
Mmt	(4-methoxyphenyl)diphenylmethyl
MPP	mitochondria-penetrating peptides
MUA	mercaptoundecanoic acid
MW	molecular weight
NAC	N-acetylcysteine
NHS	N-hydroxysuccinimide
NLS	nuclear localisation signal
NMR	nuclear magnetic resonance
Npys	3-nitro-2-pyridinesulfenyl
Nu	nucleophile
OVAT	one variable at a time
PACT	photodynamic antimicrobial chemotherapy
PAI	photoacoustic imaging
PB	phosphate buffer
PCI	photochemical internalization
PDT	photodynamic therapy
PE	petroleum ether
PEG	polyethylene glycol
PET	positron emission tomography
PG	protecting group
PN	phenalenone
PS	photosensitiser
PTT	photothermal therapy
ROS	reactive oxygen species

RT	retention time
SCCIC8	squamous cell carcinoma cells
SDS	sodium dodecyl sulphate
SEM	scanning electron microscopy
SNAr	nucleophilic aromatic substitution
SPAAC	strain-promoted azide-alkyne cycloaddition
SPACE	skin penetrating and cell entering
SPPS	solid phase peptide synthesis
StBu	tert-butylsulfenyl
STmp	trimethoxy phenylthiol
STORM	stochastic optical reconstruction microscopy
TBAB	tetra butyl ammonium bromide
tBu	tert-butyl
TCEP	(tris(2-carboxyethyl)phosphine)
TCSPC	time-correlated single photon counting
TEM	transmission electron microscopy
TFE	trifluoroethanol
Thp	2-tetrahydropyranyl
TIC	total ion count
TIRF	total internal reflection fluorescence
TOF	time of flight
TPGS	D-α-tocopherol polyethylene glycol succinate
TRIS	tris(hydroxymethyl)aminomethane
Trt	triphenylmethyl
WL	wavelength

1 Introduction

The bioconjugation of peptides and porphyrins, two important families of biologically active molecules, as well as the study of the resulting bioconjugates, have been the topic of this doctoral project and will be discussed in this thesis. The scientific background and objectives of the research project will be outlined in this first chapter.

1.1 PEPTIDES

1.1.1 Definition

Despite some discrepancy in its classification, a peptide can be broadly defined as a chemical entity comprising at least two amino acids (AA), linked by an amide bond (also called peptide bond under these circumstances), with a maximum of 50 residues and a molar mass <5000 g/mol.[1] When the size of an amino acid sequence exceeds these values, the macromolecule is considered a protein or a peptide polymer. There are 20 common and naturally occurring AAs: they all have the same general structure (**1**) but are characterised by a distinctive side chain (R) appended to the α-carbon of the carboxylic acid. AAs are usually indicated by full name, three or single letter code as reported in Figure 1.
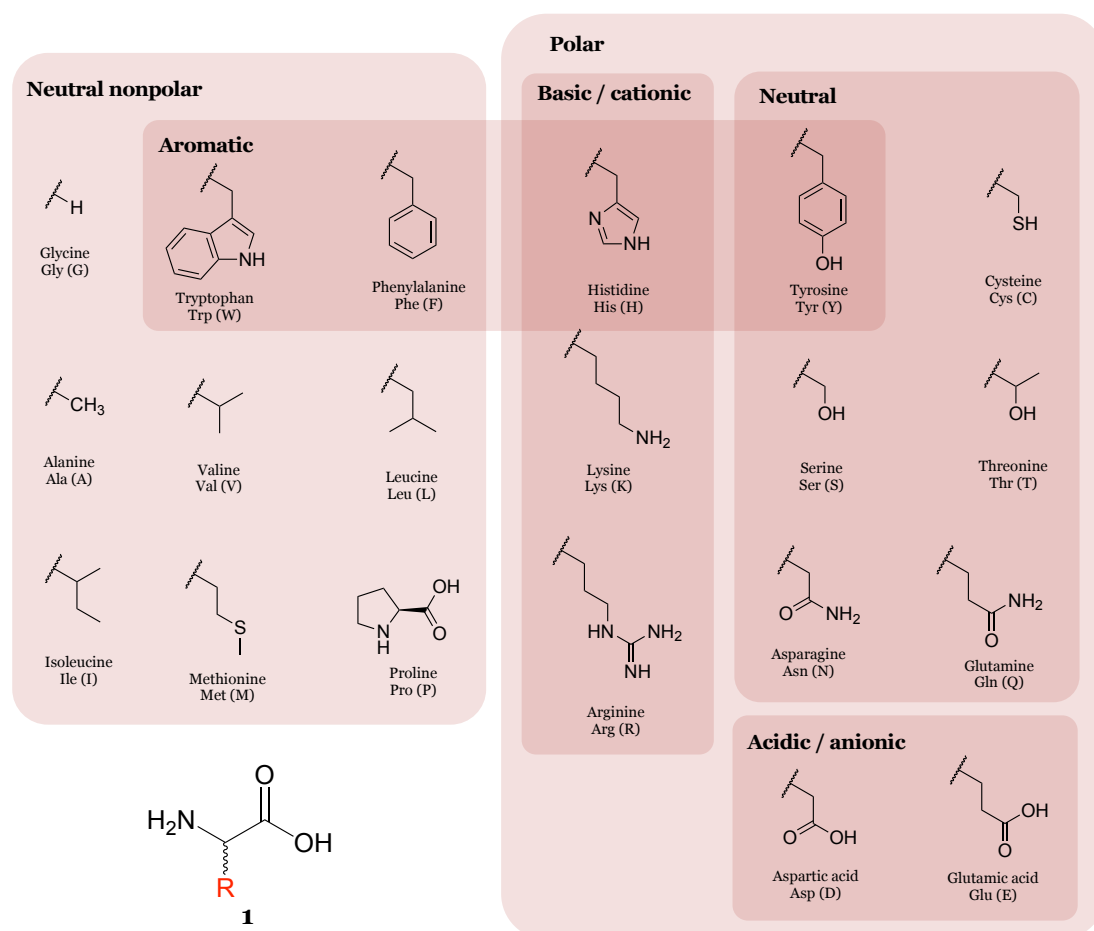
*Figure 1. General AA structure (**1**) and chemical structure of common and naturally occurring AA side chains (R), with corresponding AA full name, three or single letter code.*

Natural AAs generally contain primary amine groups and are chiral molecules, with *S* configuration of the α-carbon and overall L-configuration. Within this group, glycine is the only non-chiral AA because it has two hydrogens on the α-carbon, proline is the only secondary amine with a cyclic structure, and cysteine is the only natural AA in which the α-carbon has an *R* configuration.

AAs with a D-configuration (*e.g.* D-alanine, **2** in Figure 2), AAs with alternative side chains (*e.g.* cyclohexyl alanine, **3** in Figure 2), AAs bearing the amino group in non-α-position (*e.g.* β-AA such as β-alanine, **4** in Figure 2), or AAs presenting side chains attached to the amine nitrogen (*i.e.* peptoid monomers such as sarcosine, **5** in Figure 2) are just a few examples of non-natural AAs.

Peptide sequences can comprise all kind of AAs and are conventionally written (with single or three letter codes) or drawn (indicating the full chemical structure) with the N-terminal AA (free amine) on the left and the C-terminal AA (free carboxylic acid) on the right. Prefixes like di-, tri-, tetra-, penta- etc. can be used to indicate the number of residues. An example of a tripeptide (**6**) is given in Figure 2.
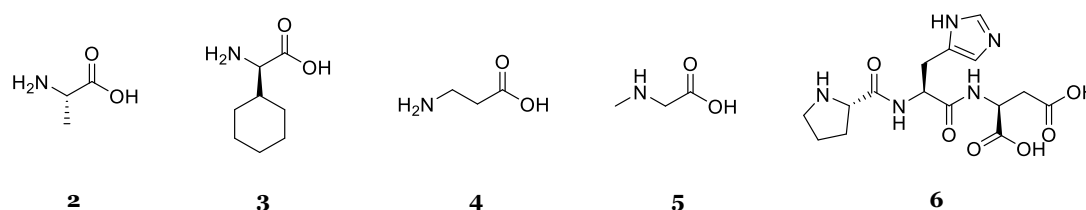
*Figure 2. Chemical structure of D-alanine (**2**), cyclohexyl alanine (**3**), β-alanine (**4**), sarcosine (**5**) and tripeptide H-ProHisAsp-OH or PHD (**6**).*

1.1.2 Synthesis

The key reaction for peptide synthesis is the formation of an amide bond between the amino group of one AA and the carboxyl group of the next. The nucleophilic addition-elimination at the carbonyl group is generally performed in solution, using a free amine and an activated carboxylic acid, for example as acyl chloride or reactive ester.[2]

However, AAs often contain multiple reactive functional groups that may interfere and generate side-products, requiring a protecting group (PG) to control the reaction outcome. Depending on specific chemical requirements, different activation and protection strategies come into play, but the main prerequisite is that carboxylic acid and amine protections are removed under different (orthogonal) conditions.[2] Therefore, the synthesis of an entire peptide requires a succession of AA deprotection, activation and coupling reaction cycles.

Solid Phase Peptide Synthesis (SPPS) was designed to perform these steps in an iterative and cyclic process.[3] The peptide attached on a solid support allows user-friendly handling and inter-step washing of deprotection/coupling side-products by resin filtration.

Exploiting recent technologies, all these steps were subsequently fully automated and paired with microwave heating to improve yields and reaction times.[4–6] A classic SPPS cycle is

represented in Figure 3A and detailed in Figure 3C, while an automated peptide synthesiser is shown in Figure 3B.
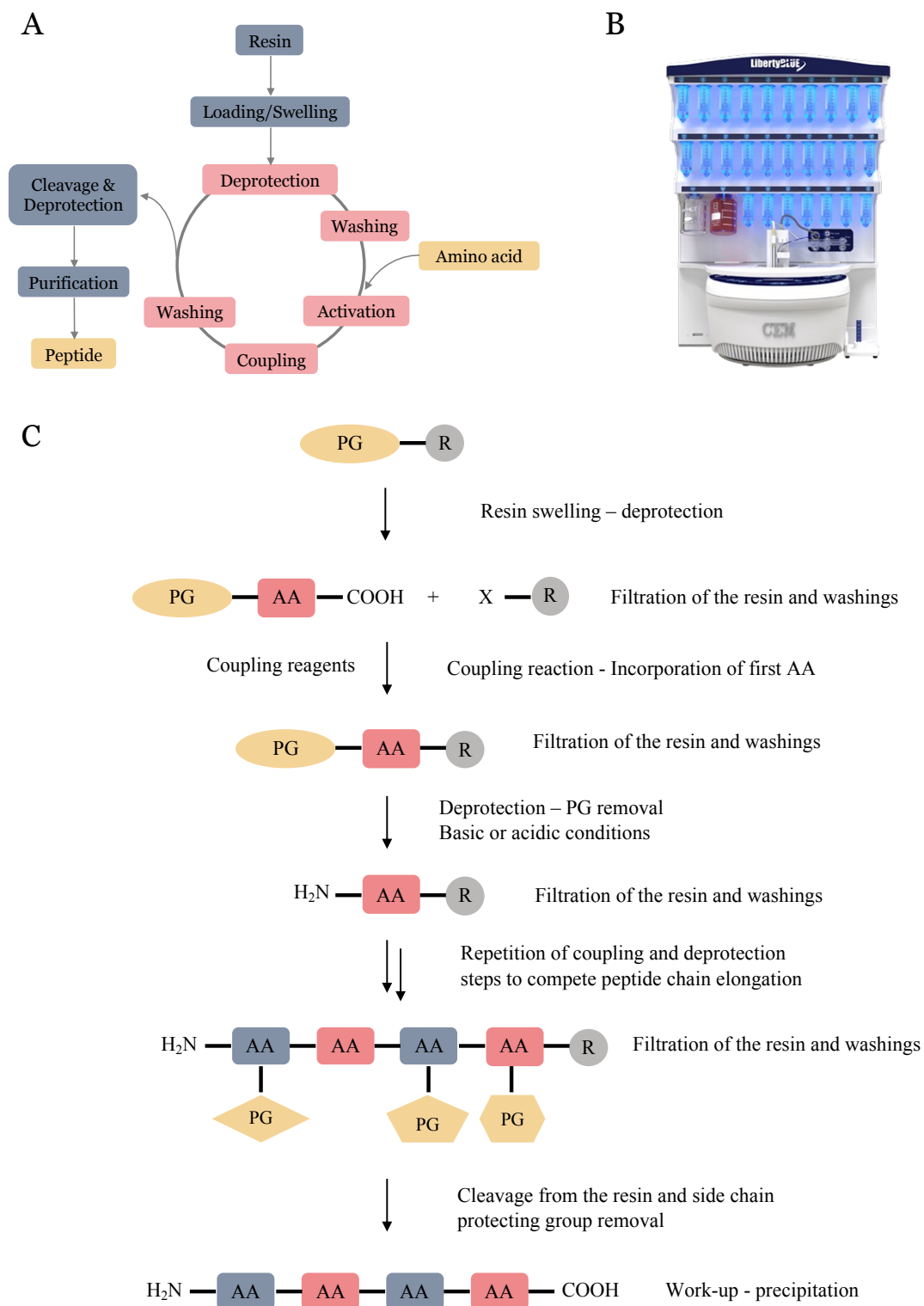
Figure 3. A) Simplified scheme for a classic SPPS cycle; B) CEM Liberty Blue Solid Phase Peptide Synthesiser; C) Detailed scheme for a classic SPPS cycle where R= resin, PG= protecting group, AA= amino acid.

Nowadays, SPPS is probably the most common approach for peptide synthesis. However, the manufacturing of the most complex and long AAs sequences still relies on recombinant technologies, while synthesis in solution is generally advised for short chains or for specific couplings only.[7] Some of the most important events in the history of peptide synthesis and development of ultimate SPPS are highlighted in Figure 4.

Year	Event
1901	•Synthesis of GlyGly dipeptide and introduction of the term **peptide** by Fischer *et al.*
1922	•Discovery of insulin and animal extraction by Banting *et al.*
1952	•Total synthesis of Oxytocin by Du Vigneaud *et al.*
1963	•Introduction of SPPS by Merrifield *et al.*
1966	•Automation of SPPS by Merrifield *et al.*
1978	•Production of recombinant insulin by Goeddel *et al.*
2003	•Introduction of Microwave-assisted SPPS by Collins *et al.*
2014	•Development of High Efficiency-SPPS methodology by Collins *et al.*
2020	•More than 100 therapeutic and diagnostic peptides on the market by D'Aloisio *et al.*

Figure 4. Timeline of peptide synthesis. For references see[1,3–6,8–11].

1.1.3 Applications

Peptides are naturally occurring compounds that, in term of complexity, properties and molecular weight, occupy the chemical space between small molecules and proteins. These biomolecules can be found in many living systems, displaying an array of biological functions.[12]

A peptide sequence with more than just a few AAs offers several possibilities of AA combinations: the variety of side chains, presenting functional group with different chemical properties, provides chemical diversity, structural variability, and a broad range of biological activities. Because their synthetic accessibility makes them appealing to modify and tune, significant efforts were directed to the synthesis and modification of natural peptides, but also to the exploration of novel sequences.

Nowadays, there are numerous classes of peptides with just as many applications. Therapeutic peptides, cyclic peptides, and self-assembling peptides are three groups that were studied in this work.

1.1.4 Therapeutic peptides and PepTherDia

The panorama of the pharmaceutical market has been traditionally dominated by small molecule drugs, with protein-based therapeutics that arose beside them as powerful medicines

for certain diseases. Peptide drugs, which combine many benefits of the two other drug classes (Figure 5), have recently found success as well.[13]
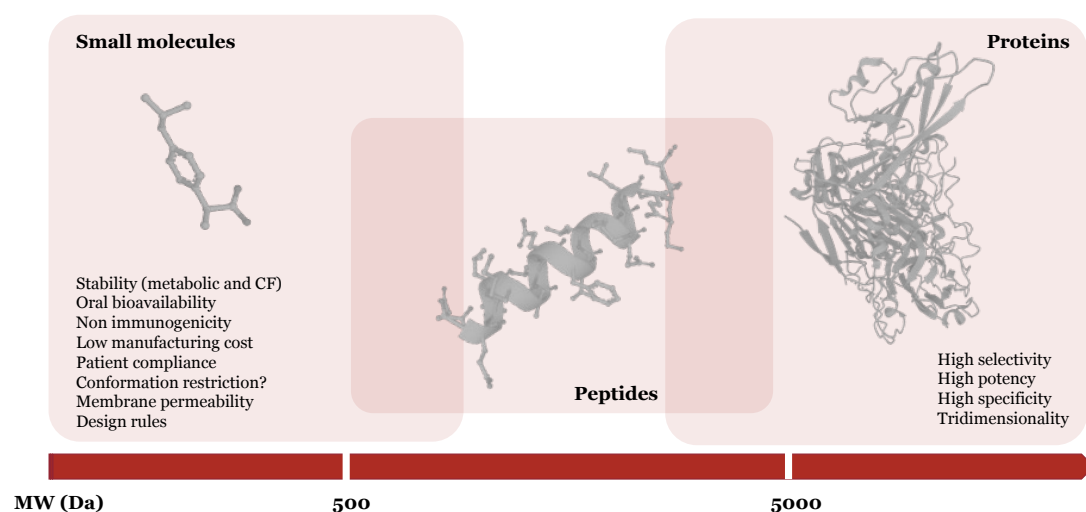
Figure 5. Graphical representation of how peptides fit in the gap between small bioactive molecules and proteins in term of MW and properties.

As a result of their bio-similarity, the pharmacokinetic of peptides typically affords rapid distribution, easy elimination and lower accumulation in tissues, low toxicity, biocompatibility, low risk of generating a dangerous immune response and overall safety.[14] Pharmacodynamically, they are often highly bioactive and highly specific. Interestingly, a short peptide can mimic the 3D structure of parts of a protein surface involved in, for example, protein-protein interactions or in epitope recognition, which are comprised of a relatively small number of AAs.[15] This could allow the modulation of all those pharmacological targets (around 70%) that are estimated to be non-druggable by both high and low molecular weight (MW) compounds.

Indeed, peptides have been used in the therapeutic field for over a century,[16] but their applications have been limited by pharmacokinetic liabilities. Low adsorption, poor physicochemical stability (short shelf-life), fast metabolism (short half-life), poor cell penetration and low oral bioavailability are tangible drawbacks and limitations that recent medicinal chemistry strategies in peptide research had to overcome. Solubility and production costs can also sometimes be challenging aspects.[17] Recently introduced techniques for peptide chemical modification, bioconjugation, formulation and administration have definitively boosted peptide applications.[18] From a manufacturing point of view, remarkable developments were introduced contextually with the increased interest in peptide drugs, and made the process scalable, cheaper, more efficient, fully automated, and – maybe most importantly – greener.[19] These progresses led to a consistent number of peptide-based drugs approved on the main pharmaceutical markets and an increasing number of those undergoing clinical trials.

In the clinic, peptides are mainly used to treat metabolic disorders, cancer, and cardiovascular diseases but also infectious diseases, pain, complications to the urinary tract, as well as

neurological, gastrointestinal, and respiratory conditions. New peptide-based drugs will be useful candidates to meet the need for treating of the above illnesses, which incidence is growing.

Despite the attention received by this class of therapeutics, there were no examples of studies that aimed at understanding the most common structural features found in successfully approved peptide drugs, nor any complete and freely accessible database that provides information about these pharmaceutical products. Therefore, a structural analysis of all the peptides for therapeutic and diagnostic applications approved in the main pharmaceutical markets was performed in collaboration with Vera D'Aloisio (PhD student at LJMU), with the aim of providing an overview of their properties. The results from this structural analysis as well as other information collected were published and made freely accessible online on PepTherDia. This database has been periodically updated and, as of June 2023, gathers 112 approved peptides with therapeutic or diagnostic applications. Some of the outcomes of this analysis that are relevant for this work will be discussed hereafter, but the reader can refer to the published manuscript for details.[1]

1.1.5 Cyclic peptides

Macrocyclic peptides are defined as peptides "bearing a ring spanning multiple amino acid residues",[20] while multicyclic peptides are similar architectures with more than one connection between different residues, thereby presenting at least two distinct rings. From the abovementioned structural analysis of approved peptide therapeutics and diagnostics, peptide cyclisation was found to be one of the most frequent structural characteristics of approved peptide drugs. 52 out of the 112 approved peptide-based therapeutics and diagnostics drugs are macrocyclic or multicyclic. Within all the cyclic peptides, the most common cyclisation strategy is a disulfide bond and rings comprising from 6 to 8 residues are the most frequent. This is confirmed by another recent publication on the topic.[21] It should not be surprising that many natural peptides present a cyclic or multicyclic structure as well.[22]

Different studies demonstrated that a cyclic structure could improve the properties of a peptide, as compared to the linear analogue. The constrained geometry enhances peptide biological stability in term of resistance to enzymatic degradation (*endo-* and *exo-*peptidases)[23,24] and extreme pH conditions.[25] Moreover, cyclisation favours a protein-like fold,[26] which is particularly useful for short AA sequences that would not normally arrange with a thermodynamically stable secondary structure.[27]

Conformational stability and 3D architecture are important features for high affinity and selective interactions because constrained peptides are more likely to achieve a better complementarity with the target and, at the same time, non-specific interactions are reduced. Finally, cyclisation might also enable an easier crossing of cell membranes and subsequent action at intracellular level by 'masking' the polar peptide backbone. [28–30]

Multicyclisation can further improve the benefits introduced by the formation of one ring, especially if the single macrocycle comprises more than 10 AAs.[31] Within these chemical

species, the example of cyclotides, a family of multicyclic peptide natural products, stands out above all. Cyclotides are plant-derived multicyclic peptides with a sequence of around 30 AA which is cyclised head-to-tail and further constrained with a cysteine knot motif between side chains (Figure 6).[32] These characteristic interlinking disulfide bridges form a rigid and well-defined structure with up to 6 smaller loops that make these peptides remarkably stable to thermal and enzymatic degradation.

Linaclotide, a 14-mer peptide with 3 disulfide bonds used as an oral drug to treat irritable bowel syndrome, is one example of a cyclotide-like peptide which is clinically employed.[33] Cyclotide stable architecture also tolerates modifications and bioactive epitopes grafting into the loops. Therefore, cyclotides are used in medicinal chemistry as scaffolds for the design of new pharmaceuticals.

For examples, starting from the framework of the cell-penetrating cyclotides MCoTI-I, a p53 modulator sequence was inserted for stabilisation and intracellular delivery. A similar strategy was also used to obtain an orally bioavailable chimeric of the cyclotide kalata B1, active as bradykinin B1 receptor antagonist for the treatment of chronic pain. Based on these models, a multicyclic architecture incorporating different peptide loops can be imagined, where each one has a specific function in the final construct. There are other examples of rationally designed bicyclic peptides, presenting a binding sequence (*e.g* to protein–tyrosine phosphatase 1B, to NEMO-IkB kinase) confined into one ring and a cell-penetrating motif in the other ring, that showed improved properties compared to the linear or macrocyclic original peptides.[34]
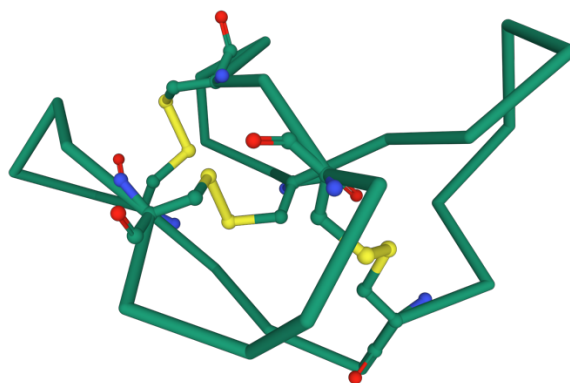
Figure 6. Image from the Protein Data Bank[35] *of the cyclotide varv F (PDB 3E4H)*[36] *crystal structure with peptide represented as backbone in green, cysteine as ball and stick and disulfide bonds in yellow.*

Obviously, these advantages of cyclic peptides come with an increased effort when designing and synthesising the molecule. Therefore, the development of new peptide cyclisation techniques has received significant attention in recent years.[37–40]

A cyclic structure can be achieved with different strategies to connect peptide head, tail, or side chains (Figure 7). The chemistry employed often overlaps with the most common bioconjugation reactions, which will be discussed later in a dedicated section (1.3).

Cyclisation is generally performed in solution under high dilution, on-resin or simultaneously to the deprotection-cleavage step, and usually there is no direct participation of the AAs involved in the interaction with the target.[41] Head-to-tail cyclisation typically involves and amide bond formation (*e.g.* directly or via Native Chemical Ligation[42]) whilst the developments in side-chain to side-chain cyclisation are diverse.
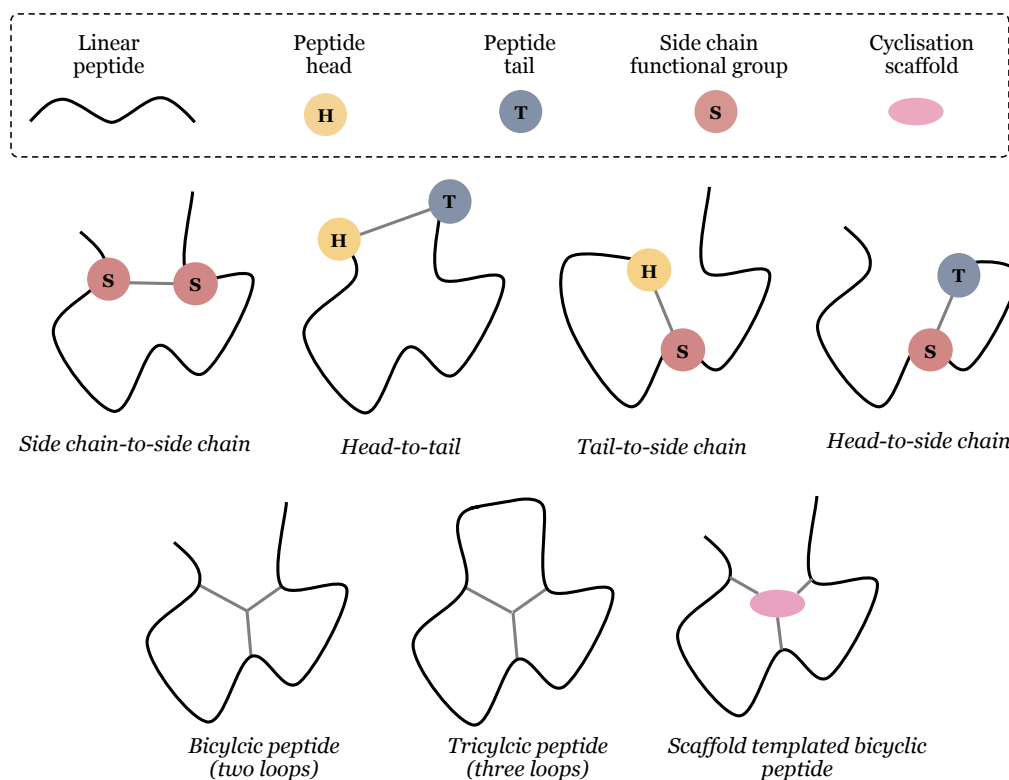
Figure 7. Graphical representation of cyclic peptide architectures.

In 1995, Miller and Grubbs reported the first example of one-component side chain-to-side chain peptide cyclisation by olefin metathesis. The reaction (Scheme 1) between alkenyl AAs of a general peptide (**7**), catalysed by ruthenium catalyst (**8**), formed a "staple" that covalently cross-linked two side chains of different AAs within an α-helical structure (**9**). The residues, usually at *i, i + 4* or *i, i + 7* distance, are stabilised by the staple, and the peptide helicity is consequently improved.[43]
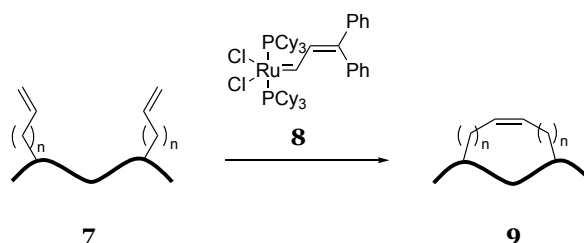
*Scheme 1. General reaction scheme for peptide cyclisation by olefin metathesis of alkenyl AA side chains. Alkene configuration in product **9** is for graphical purpose only.*

Ten years later, the first two-component strategy, which used a chemical scaffold in the role of a cyclisation core, was also applied for peptide cyclisation. This was called Chemical LInkage

of Peptides onto Scaffolds (CLIPS) technology and introduced by Timmerman and co-workers[44] as an epitope mapping tool to connect multiple loops of a linear peptide and mimic discontinuous epitopes of antibodies. Originally, an aromatic ring bearing reactive alkyl bromides as anchor points for cysteine thiols was employed as scaffold (*e.g.* 1,3-bis (bromomethyl)benzene, **11** in Scheme 2). Subsequently, CLIPS technology has been applied to a wide range of peptides, performed with more complex synthetic procedures, and implemented with various scaffolds.[34,45–50]
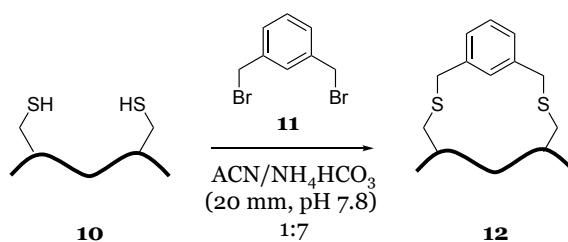
Scheme 2. General reaction scheme for the Chemical Linkage of Peptides onto Scaffolds (CLIPS).

In this context, alternative multicomponent cyclisation strategies were developed and, among these, cysteine crosslinking using polyfluoroaromatics was first described by Pentelute.[51] This peptide cyclisation strategy is particularly interesting and relevant for this work and will be discussed separately in the next chapter.

1.1.6 Self-assembling peptides

Self-assembly is a spontaneous process, during which molecules in solution organise to generate a supramolecular structure with ordered architecture.[52] This behaviour is directed by specific and noncovalent intermolecular interactions that are repeatedly established within the system. Hydrogen bonding, electrostatic and/or hydrophobic interactions, π-π stacking, and van der Waals forces can be promoted (but also reversed) by chemical conditions or physical stimuli. This process can ultimately lead to the formation of larger supramolecular organizations named assemblies, typically with rod-, fibre-, sheet-like, or spherical shape, and dimensions ranging from nanometre to micrometre scale.

The terms *co-assembly*, which denotes the assembly of two (or more) chemically independent components, and *aggregates*, irregular clusters of molecules that ultimately form particles without organisation nor order, should be distinguished from self-assembly and assemblies, respectively.[52]

Self-assembly is a common process in biological systems: many complex cellular construct and machineries are supramolecular structures made of smaller biological molecules. For examples, the lipid bilayer of cell membranes is made of phospholipids, the double and single strand helixes of nucleic acids are made of nucleotides, proteins have different levels of organisation – secondary, tertiary, and quaternary structure – originated from AA sequences.

Although nature mainly employs long AA sequences (*i.e.* proteins rather than peptides), shorter chains can also be designed to replicate natural structures and used as building blocks

for the development of artificial self-assembled materials.[53] These soft materials obtained from peptides are usually dynamic (respond to external stimuli with property changes), biodegradable, biocompatible, synthetically accessible, and modulable. A strong self-assembly tendency as well as favourable physicochemical and mechanical properties are also desirable properties for these systems.

The classification of self-assembling peptides is generally based on the AA sequences, on the shape of the formed supramolecular structures or on their applications.[54–56] Some examples of these classes are provided hereafter and represented in Figure 8. The most common self-assembling peptide sequences are:

- dipeptides and short peptides – frequently comprising residues with strong aggregation tendency such as FF, dipeptide derived from the amyloid motif KLVFF.[57]

- peptide amphiphiles – sequences combining a hydrophilic head and a hydrophobic tail.[58]

- peptide bolaamphiphiles – sequences where a hydrophobic body is within two hydrophilic ends.[59]

- cyclic peptides – in which the peculiar geometry influences the supramolecular structures obtained.[60]

- collagen-like peptides – containing the canonical tripeptide motif Gly–X–Y, where X and Y residues are generally proline (Pro) and hydroxyproline (Hyp).[61]

- derivatised peptides – such as peptides with protecting groups or conjugated with carbon chains, lipids, nucleic acids and more.[62]
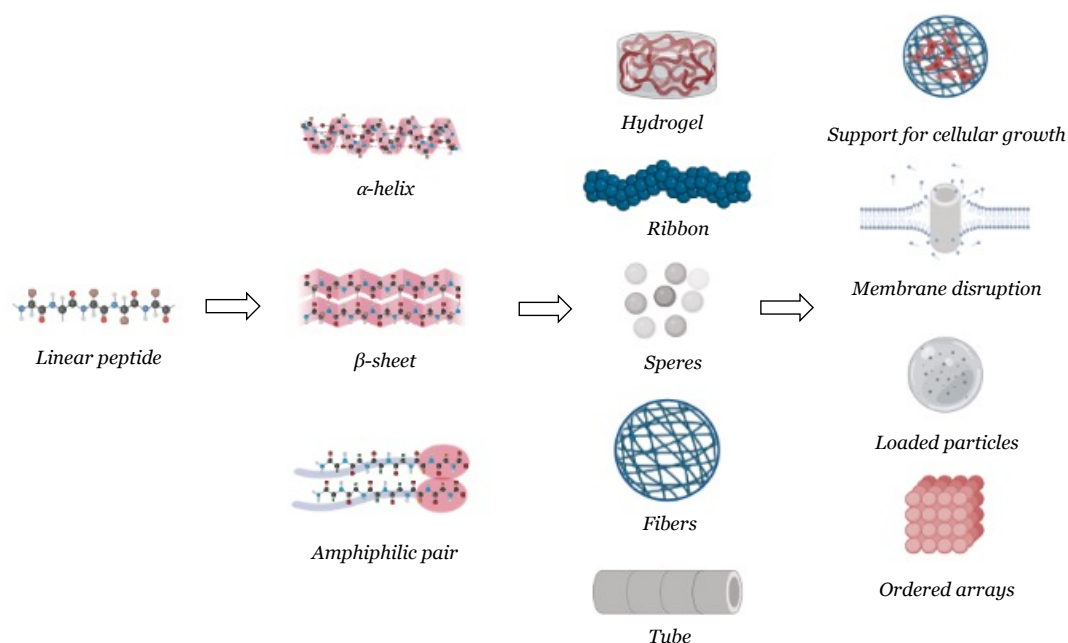
Figure 8. Self-assembly of peptides, from 1D to complex 3D architectures. Created with BioRender.[63]

The most common self-assembled architectures are helix-, sheet- and surfactant-like that further organise into 3D structures with morphologies such as spheres, fibrils, fibres, and tubes. These species were reported to have applications as bioactive materials,[64] biosensors,[65] drug delivery systems,[66] 3D scaffolds for cell culture,[67] electronic devices,[68] supports for tissue regeneration and engineering,[69] bio-printing[70] and artificial mimicking of the self-assembly of more complex natural systems as described in Section 3.

The ability of peptides to form these long-range ordered supramolecular architectures in given environments is a very interesting property of this class of chemicals which was explored and exploited in this research.[71,72]

1.2 PORPHYRINS

1.2.1 Definition

Porphyrins are heterocyclic molecules containing four pyrrole rings linked by methine bridges at their α-carbon. This peculiar chemical connectivity forms a tetrapyrrole macrocycle with square planar architecture and an 18-π electron aromatic system which is responsible for many of the interesting properties of these chemicals. The general structure of a porphyrin core, which is also called porphin (**13**), is reported in Figure 9A, and three important structural positions (α, β and *meso*), the aromatic ring and the tetradentate cavity are highlighted.
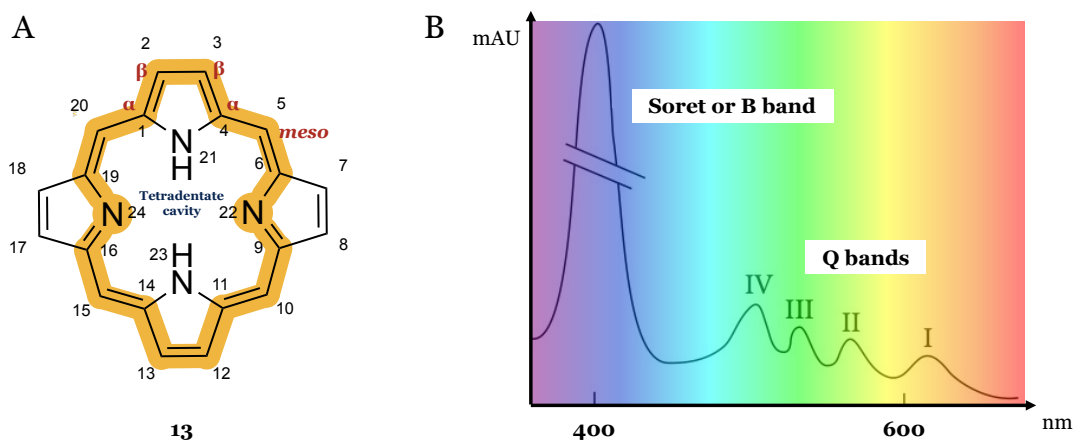
Figure 9. A) Chemical structure of porphin where the positions α, β and meso, the 18-π electron aromatic system (yellow shading) and the tetradentate cavity are indicated; B) Typical absorption spectrum for a porphyrin. Spectrum in figure 9B was modified from Josefsen et al.[73]

Porphyrins (from Greek πορφύρα or porphyra, meaning purple) are easily distinguished for their very intense colour, which varies from pink to dark purple. The UV/Vis spectrum of a porphyrin is a fingerprint for this class of chemicals and is reported in Figure 9B.

The strong absorption in the visible region of the electromagnetic spectrum is due to the extended conjugated π-system. This is characterised by the Soret band and the Q bands. The Soret band is also called B band and is the most intense absorption around 410 nm corresponding to the $S_0 \rightarrow S_2$ transition. The Q bands are several bands (I – IV) of lower intensity in the visible region between 450 and 650 nm corresponding to the $S_0 \rightarrow S_1$

transition, which are responsible for the porphyrin colour. The intensity and exact wavelength of these bands is affected by the tetrapyrrole ring substitution, and the metal ions coordinated within the central nitrogen atoms.[74]

After the energy is absorbed by the porphyrin ring to generate the excited states, the electrons can relax back to the ground state in different ways, as illustrated by the Jablonski diagram reported in Figure 10.[75] In general, energy can be dissipated via radiative or non-radiative decay. The first mechanism leads to the emission of photons by photoluminescence, while the second is determined by internal conversion through vibrations and energy release as heat.

Photoluminescence can be distinguished between fluorescence, if the transition takes place between electronic states with the same electron spin multiplicity (singlet-singlet), and phosphorescence, when intersystem crossing and electronic states with the different electron spin multiplicity (singlet-triplet) are involved.

After singlet-triplet intersystem crossing, additional mechanisms of energy dissipation from the triplet excited state are also allowed if a quencher molecule is present. Static or dynamic collision with the quencher deactivates the excited triplet state of the porphyrin. In static quenching, the excited species and the quencher form a non-emissive complex but, if there is an energy or electron transfer between the two, the quenching is dynamic. When this latter quenching pathway takes place between the porphyrin in its triplet state and oxygen (normally present in its singlet state), reactive oxygen species or singlet oxygen are generated via Type I or Type II reactions respectively. Due to this peculiar electronic behaviour, porphyrins are classified and employed as photosensitisers (PS).
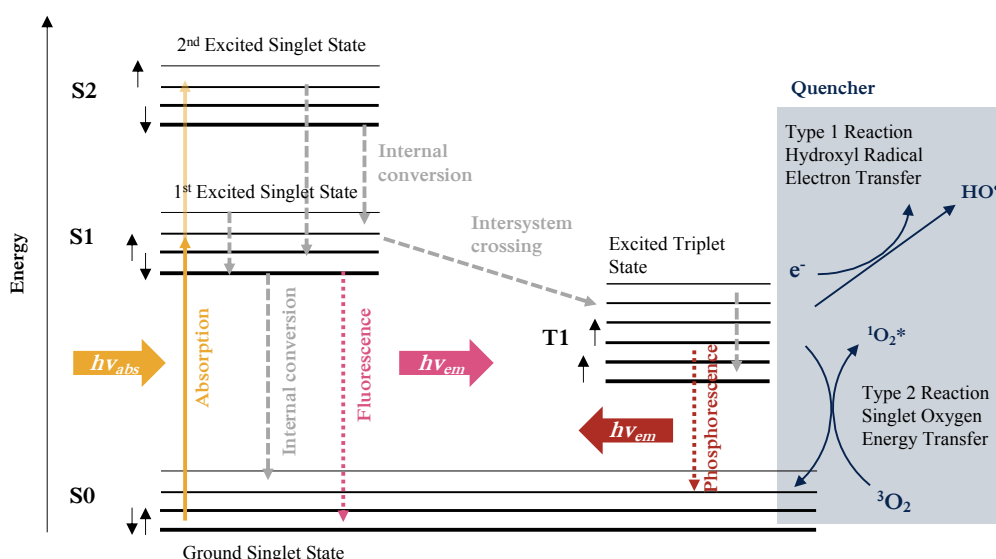
Figure 10. Simplified Jablonski diagram.

The rigid and planar aromatic architecture is also responsible for great chemical and physical stability (pH, temperature, and different environmental conditions), while the tetrapyrrole core allows different metals to be coordinated within the tetradentate cavity.

The metal ion accepts lone pair electrons from the two nitrogen atoms and contextually donates its own to the porphyrin molecule, forming delocalized π bonds and allowing an increased electron distribution within the conjugate system. This latter modification drastically changes the electronic properties of the ring and consequently the behaviour of the whole molecule, besides providing the possibility of binding of additional ligands. Usually, the UV/Vis spectrum of metalloporphyrins exhibits a shift in the absorption wavelength of the Soret band and a smaller number of Q bands, because of an increased molecular symmetry. Vice versa, when the porphyrin is asymmetric, the UV/Vis spectrum usually presents multiple Q bands.

1.2.2 Synthesis

β- and *meso*-substituted porphyrins are naturally available compounds but, since the first porphyrin synthesis performed in 1935 by Rothemund,[76] the evolution of synthetic strategies has granted laboratory access to many of these derivatives and inspired the design of new ones.

A general curved arrows mechanism for the tetrapyrrole ring formation is reported in Scheme 3. The key reaction for porphyrin synthesis is the initial electrophilic substitutions on the α positions of pyrrole (**14**) with an aldehyde (**15**). This chemical mechanism, repeated for four times, leads to the porphyrinogen ring closure via dipyrromethane (**16**) formation. Porphyrinogen (**17**) oxidation to generate the tetrapyrrole aromaticity is the final step of the process. However, the reaction can also happen polymerically to form polypyrromethanes as side products and, for this reason, the final tetrapyrrole yield is generally negatively affected.
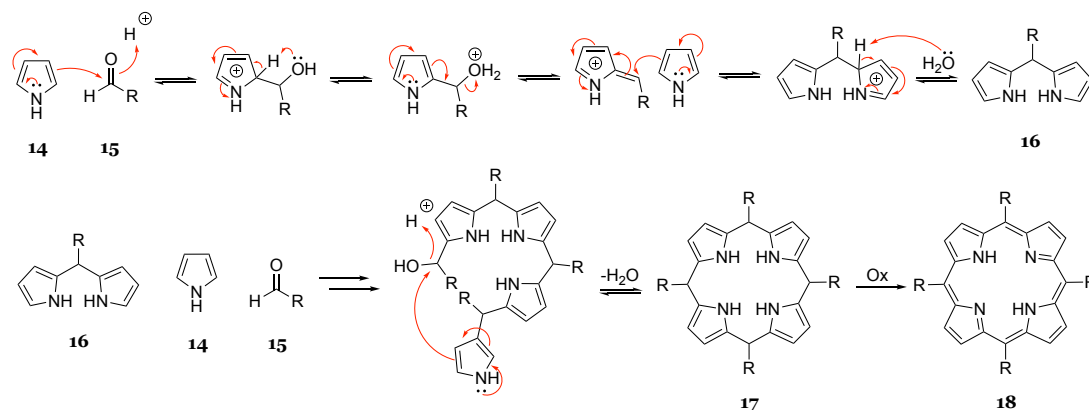
Scheme 3. Reaction scheme and mechanism for the tetrapyrrole ring formation.

In this approach, aldehyde substitution (R) determines the *meso* derivatisation of the final 5,10,15,20-tetrasubstituted porphyrin (**18**). For the synthesis of symmetric porphyrins, one aldehyde is employed and gives the same substituent on positions 5, 10, 15 and 20. The one-pot conditions reported in scheme 4 from Adler-Longo and Lindsey porphyrin synthesis are probably the most employed procedures to date.[77,78]
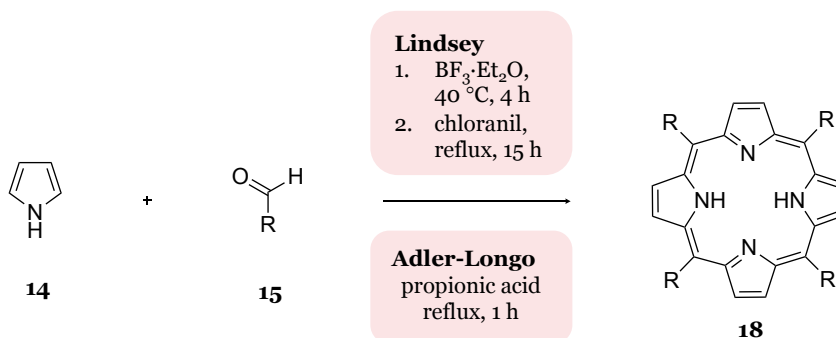
*Scheme 4. Reaction scheme and conditions from Lindsey and Adler-Longo procedures for the synthesis of 5,10,15,20-tetrasubstituted porphyrin (**18**).*

For asymmetric porphyrins (bearing different substituents on *meso* positions), more than one aldehyde can be employed in one pot following a modification of Adler-Longo procedure called Little or "mixed-aldehyde" synthesis.[79] When starting from two different aldehydes, indicated here as A and B, the obtained porphyrins can be identified as:

- A4 – if the 4 substituents are from A.

- A3B – if 3 substituents are from A and 1 substituent is from B.

- A2B2 – two isomers with 2 substituents from A and 2 substituents from B.

- AB3 – if 3 substituents are from B and 1 substituent is from A.

- B4 – if the 4 substituents are from B.

Little synthesis remains one of the most employed but, even though the modulation of A/B aldehydes ratio can direct towards a specific product, a mixture of difficult purification is usually obtained. Therefore, a higher control on the porphyrin final substitution and stereochemistry may be desirable and can be achieved, for example, by stepwise synthesis via dipyrromethene. This approach also applies for the synthesis of 5,15-disubstituted porphyrin (**19**), for which a model procedure is reported in Scheme 5.
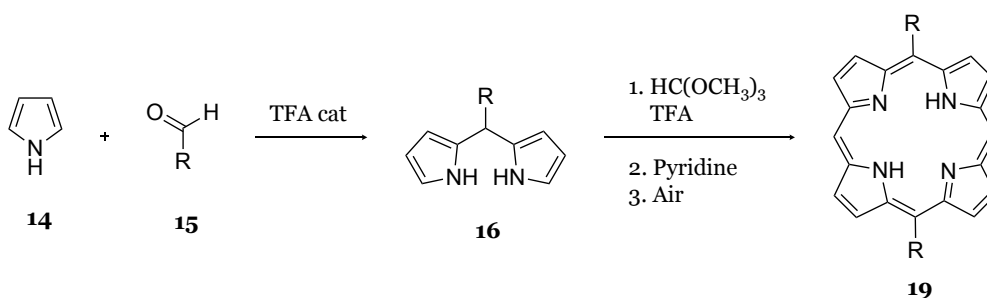
*Scheme 5. Reaction scheme and conditions for the synthesis of 5,15-bissubstituted porphyrin (**19**). [80]*

The synthesis of more complex tetrapyrroles, such as β-substituted natural porphyrin, are fascinating and much more articulated, but their presentation goes behind the scope of this work.[81]

1.2.3 Applications

Similarly to peptides, porphyrins have such a crucial role in the biochemistry of living systems that porphyrin-derived pigments are frequently called "pigments of life".[82] For example, haem (**20**), chlorophyll (**21**), vitamin B_{12} (**22**) and coenzyme F_{430} (**23**) – binding iron, magnesium, nickel, and cobalt, respectively – are porphyrins associated with different biological functions (Figure 11). Haem is the prosthetic group of haemoglobin and coordinates the transport of O_2 and CO_2 in red blood cells towards and away from the tissues. The same porphyrin is also the catalytic core of many enzymes and supports the electron-transport chains necessary for oxidative phosphorylation and ATP creation. If haem is essential for animal life, chlorophyll has the same role for plants. This reduced porphyrin is involved in the light harvesting and energy conversion processes of photosynthesis.
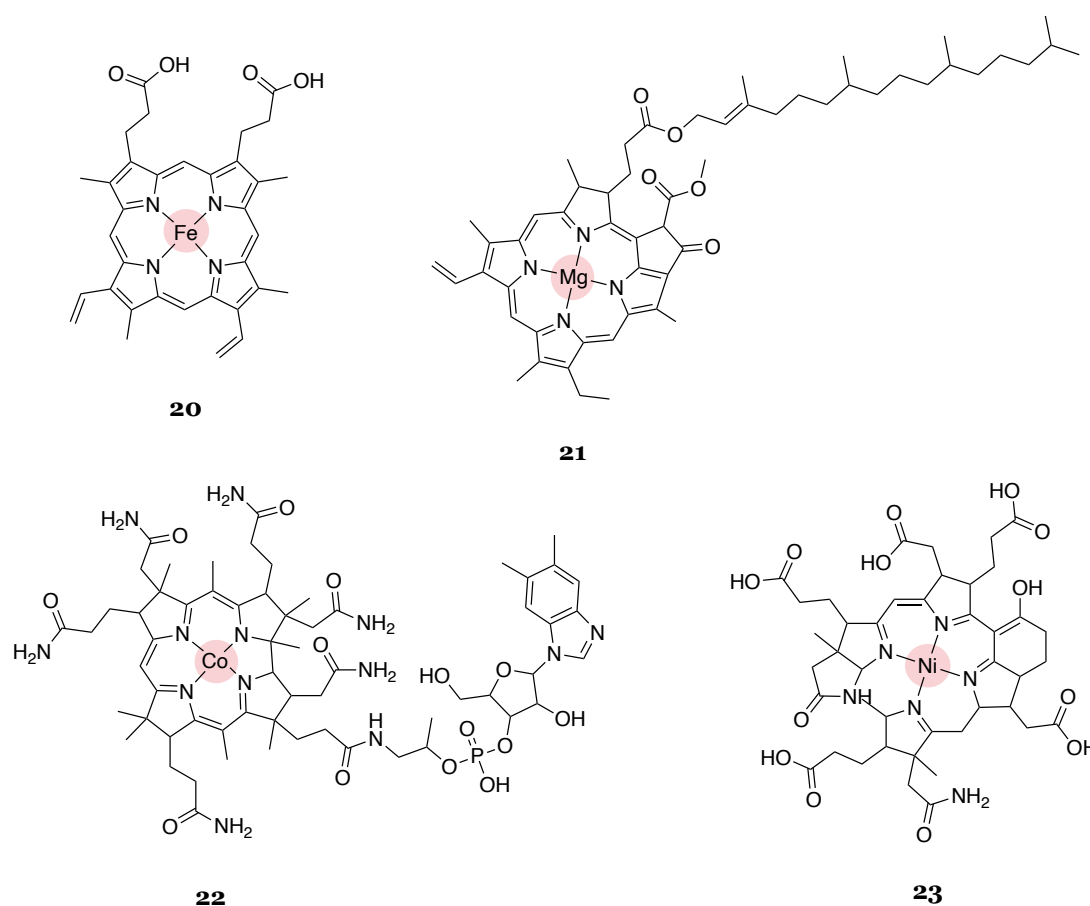
*Figure 11. Chemical structure of haem (**20**), chlorophyll (**21**), vitamin B_{12} (**22**) and coenzyme F_{430} (**23**).*

Porphyrin applications mainly derive from their tuneable electronic properties and photo-response, particularly the high yield of excited triplet state and the ability to transfer electrons or energy from this excited state to neighbouring chemical species. Natural and synthetic porphyrins as well as their metal complexes are commonly employed in artificial photosynthesis, dye sensitised solar cells (DSSC), photocatalysis, oxygen sensitivity, and many biomedical applications.

Systems for artificial photosynthesis simulate the natural process by absorbing light energy in a wide range of wavelengths, separating charges, and channelling the separated charge to a designed catalytic site where water oxidation and CO_2 reduction occur. To this end, porphyrins can be employed as both electron donor and acceptor.[83]

In the design of DSSC, where sunlight is converted into electric power, porphyrins are one of the applicable photosensitizers with the function of light-harvesting antennae.[84] They are usually deposited on the surface of a TiO2 porous layer to absorb visible light, generate excited electrons and transfer them to the semiconductor.

The electron donor and electron acceptor potential of porphyrin excited states make these molecules potent oxidants or reductants as well. Therefore, these can be also used to photocatalyse reductive or oxidative processes, for example C–C and C-heteroatom bond formation.[85]

Metalloporphyrins are widely employed as optical oxygen sensors, and offer many advantages compared to traditional methods for oxygen measurement. This technique exploits the emissive behaviour of porphyrin, the dynamic oxygen quenching of the triplet state and the proportionality between luminescence intensity, triplet lifetime and the concentration of the quencher (oxygen), as described by the Stern-Volmer equation (Equation 1). Porphyrin-based sensors are commonly applied in scientific research and beyond.[86]

$$\frac{I_0}{I} = \frac{\tau_0}{\tau} = 1 + k_q \tau_0 [O_2]$$

Equation 1. Stern-Volmer equation. I_0= luminescence intensity in the absence of oxygen; I= luminescence intensity in the presence of oxygen; τ_0= triplet lifetime in the absence of oxygen; τ= triplet lifetime in the presence of oxygen; k_q= bimolecular quenching rate constant; $[O_2]$= conc. of oxygen.

Porphyrins are also suitable for many types of medical imaging and therapies such as photodynamic therapy (PDT), photothermal therapy (PTT), photoacoustic imaging (PAI), photodynamic antimicrobial chemotherapy (PACT), photoimmunotherapy (PIT), radiotherapy, fluorescence imaging, image-guided surgery, oxygen sensing, magnetic-resonance imaging (MRI) and Positron Emission Tomography/Computed Tomography (PET/CT). These are related to three following properties: i) light-induced generation of cytotoxic reactive oxygen species (ROS), ii) fluorescence emission after irradiation and iii) metal coordination.[87,88] PDT and related techniques will be specifically discussed in the next section.

1.2.4 Porphyrin-based therapeutics

At a therapeutic level, porphyrins are mostly employed as photosensitizers for PDT. PDT is a treatment performed with a drug and a visible/near infrared light source: the drug is a photosensitizer whose electronic behaviour is explained by the Jablonski diagram previously presented in Figure 10. There is also a third indispensable element for PDT, but this is oxygen and is present in most of the biological environments.[89]

When the photosensitizer molecule absorbs visible light, the excited singlet states is populated. Via intersystem crossing, excited electrons can reach the lower triplet state. In the presence of oxygen, energy or electrons are ultimately transferred to molecular oxygen and singlet oxygen or oxygen radicals (*e.g.* superoxide radical anion) are formed respectively. These molecules are generally called ROS and are the actuators of the photodynamic process. Due to their high reactivity, ROS can damage cellular structures such as membranes, proteins, or nucleic acid via non-specific oxidation reactions to ultimately promote cell death.

Selectivity is a great advantage for this approach because the drug is administered – locally or systemically using appropriate targeted delivery strategies – and activated only in the irradiated areas. On the other hand, the main drawback is that PDT can only be used to treat target areas that are reachable with light and, despite the technological support, it is challenging to irradiate deep tissues.[90]

After PDT discovery by Raab and von Tappeiner at the beginning of the 20th century[91] and the first clinical trial in 1976,[92] Porfimer sodium - an hematoporphyrin derivative - was approved as the first commercial photosensitizer for PDT under the name Photofrin®.[93] After that, second- and third-generation photosensitizers were developed and approved to treat cancer, pre-cancerous diseases, or pathogen infections.

At the time of writing, many porphyrin-based drugs are under investigation for PDT treatment and will probably join other approved drugs such as Verteporfin (Visudyne®) and Temoporfin (Foscan®) on the market.[94] Examples of alternative therapeutic applications are iron replacement therapy (Feraheme - Ferumoxytol) and porphyria (Hemin or Fe protoporphyrin IX - Panhematin).

1.2.5 Porphyrin aggregates

Depending on the tetrapyrrole substitution and environmental conditions (particularly the solvent), porphyrins display a strong aggregative behaviour. Stacking between tetrapyrrole rings usually leads to the formation of J-aggregates (tetrapyrrole with side-by-side disposition) and/or H-aggregates (tetrapyrrole with face-to-face disposition) as represented in Figure 12.

The aggregation into J-type aggregates generate a red-shift of the absorbance bands (mainly Soret), whereas the band are blue-shifted when H-aggregates are formed. This process is sometimes associated with band broadening and decrease in absorption intensity as well. To be consistent with most literature reports, these structures are defined here as aggregates but, even if they are sometimes irregular, the term assemblies would be more appropriate when these are ordered (see section 1.1.6).

Monomer   H-aggregates   J-aggregates
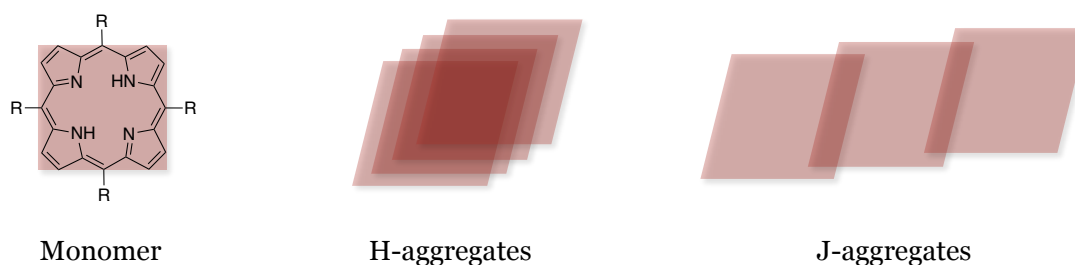
Figure 12. Porphyrin monomer disposition in H- and J-aggregates.

This strong aggregative tendency is important because the photo-response and electronic properties of the chromophores are strongly influenced by their architectural arrangement resulting from this process. Properties such as electron transfer and conductivity can be improved by aggregates formation, but some can also be negatively affected or completely lost, as the case of fluorescence or triplet formation.

Therefore, a frequent challenge in the development of functional porphyrin architectures is to organise tetrapyrroles and maintain the required functionality after assembly or aggregation. The chemical modification or conjugation of porphyrin moieties can be one of the strategies to address this issue.

1.3 BIOCONJUGATION

The formation of a covalent bond between molecules, where at least one is biologically active (*e.g.* peptides, proteins or oligonucleotides), is called bioconjugation. This is a strategy commonly employed in medicinal chemistry and chemical biology for improving and/or combining the properties of bioactive compounds, as well as obtaining new features that cannot always be obtained from the simple combination of the two separate constituents.[95]

The main requirements for a successful bioconjugation reaction can be summarised as:

- mild and biomolecule-compatible conditions. The use of organic solvents, strong oxidising or reducing agents, acid or bases, metals and high temperatures should be limited. This aspect is particularly relevant when the bioconjugation is performed under physiological conditions or involves peptides, proteins, and nucleic acid.

- fast and highly efficient, especially if performed *in vitro* or *in vivo* where product purification is not possible.

- chemoselective, and even better if site-specific. This is crucial when the strategy involves biomacromolecules such as proteins, where many functional groups and potential reacting sites are present.

- applicable across a broad range of pH and temperature values.

- generating a stable bond. However, this may not be the case when the design aim at obtaining a cleavable linker on purpose (prodrug).

- being user-friendly, from synthesis set up to product detection. For example, possibility of automation or execution on solid supports may be further advantageous.

- adhering to the criteria of green chemistry.

Most of the bioconjugation reactions developed over the years were designed for protein modification, but these can be equally employed for peptides, because both are targeting AA residues. Similar approaches are also valid for the bioconjugation of other bioactive compounds such as oligonucleotides, sugars, lipids, or probes.

Finding the ideal bioconjugation reaction that satisfies all the above-mentioned requirements is not straightforward, but medicinal chemists have provided some valid options (Figure 13). Within this assortment of synthetic bioconjugation tools – which does not include those relying on enzymatic reactions – the main distinction is between ligation at native functional groups (N and C-terminus groups or thiol, amine and alcohol on native side-chains), and reaction performed on non-natural residues.

- Amide formation
- Thiourea formation
- Native chemical ligation (NCL)
- Ester formation
- Thioether formation
- Staudinger ligation
- Cu-catalysed Azide-Alkine Cyclisation (CuAAC)
- Thiazolidine formation via propargyl amine-isothiocyanate
- Phosphoramidite coupling
- Diels Alder reaction
- Olefin metathesis
- Thiol haloacetamide
- Hydrazone formation
- Arylation chemistry
- Disulphide bond
- Maleimide coupling

Figure 13. List of the most common bioconjugation reactions.

1.3.1 Click chemistry and bioorthogonal chemistry

The 2022 Nobel Prize in Chemistry was awarded by The Royal Swedish Academy of Sciences to Carolyn R. Bertozzi, Morten Meldal and K. Barry Sharpless "for the development of click chemistry and bioorthogonal chemistry."[96]

In the early 2000, Sharpless introduced the concept of *click chemistry* to indicate a bioconjugation reaction in which molecular building blocks snap together quickly and efficiently.[97] As indicated by the authors, to be called with such name, a chemical reaction should:

- be fast, preferably with a thermodynamic driving force greater than 20 kcal/mol.

- operate under simple reaction conditions, in a solvent that is benign or easily removed.

- be highly efficient and give very high yields.

- be highly selective and stereospecific.

- be reliable, modular, and wide in scope.

- not give unwanted by-product or only inoffensive by-products that can be removed by non-chromatographic methods.

- performed with starting materials readily available.

- not be reversible, leading to products which are stable under physiological conditions.

- occur in the presence of oxygen and in water.

Shortly after, Sharpless and Meldal independently presented the copper catalysed azide-alkyne cycloaddition (CuAAC), an elegant reaction that has now become almost synonymous with click chemistry (Scheme 6).[98,99] Differently substituted alkyl and aryl azides (**24**) react at room temperature with a variety of terminal alkynes (**25**) in the presence of a Cu^I catalyst, producing the 1,4-disubstituted 1,2,3-triazole (**26**) in very high yields.
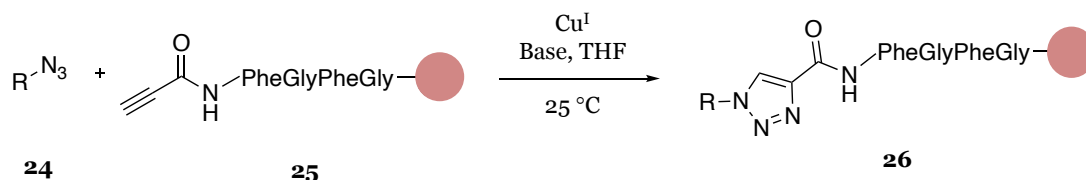
Scheme 6. General reaction scheme for CuAAC applied to a peptide system. [98]

CuAAC is experimentally straightforward, very robust, widely applicable, nearly independent of the azide/alkyne subsistent nature, proceeds to completion within hours, works with a broad range of solvents (including water) as well as pH, and has a very low interference by other functional groups. This laid the foundation to many comprehensive developments and

applications in the field of medicinal chemistry (*e.g.* drug bioconjugates), molecular biology (*e.g.* marking of genes, protein and other biomacromolecules) and material sciences.[100,101]

Click chemistry's potential was demonstrated by Bertozzi when she implemented CuAAC in living organisms. Because of the toxicity of copper ions to cells, she developed the strain-promoted azide-alkyne cycloaddition (SPAAC, Scheme 7), in which the azides (**24**) react with highly strained cycloalkynes (**27**) to give two regioisomeric triazoles (**28** and **29**) without catalyst, because the release of ring strain upon cycloaddition facilitates of the reaction progress.[102] SPAAC was employed as bioorthogonal reaction to map glycans on the surface of cells without disrupting their normal chemistry.
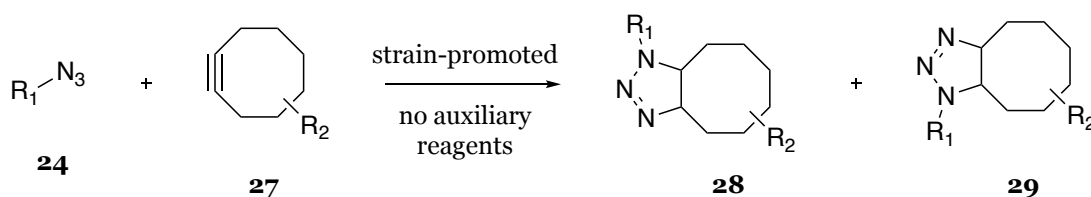
Scheme 7. General reaction scheme for SPAAC.

Alongside click chemistry, *bioorthogonal chemistry* is another important concept introduced by these laureates. The term *orthogonal* has already been employed in this manuscript when discussing the chemistry of peptide synthesis to indicate the selective installation or removal of one PG, not affecting the orthogonal ones. When this is implemented to biological systems, it indicates "reactions of functional groups that are so selective for each other that they can be ligated in a richly functionalized biological milieu".[103]

Nowadays, many reactions are considered to be 'click' reactions, mainly because of their simple synthetic procedures, fast kinetics, high yields, and physiological compatibility. Some examples are nucleophilic substitution involving para-fluoride of pentafluorophenyl groups, metal-free [3+2] cycloaddition (including but not limited to SPAAC), Michael additions of thiols with maleimide, Diels-Alder reactions, and thiol-alkene radical addition reactions. Even if they do not meet all the requirements defined by Sharpless and his co-workers, their advantage is the absence of a metal catalyst.[104]

1.3.2 Bioconjugation of peptides and porphyrins

This introduction described some of the peculiarities of peptides and porphyrins, as well as some general medicinal-chemical similarities. Consequently, the synthesis of peptide-porphyrin bioconjugates could deliver combined and advantageous properties for chemical biology and medicinal chemistry.

For example, peptides can impart to the bioconjugate a sequence-specific biological activity, improve porphyrin solubility, promote cell penetration, modulate porphyrin aggregation, or direct binding to specific targets. On the other end, porphyrins can then exploit its light-responsive behaviour and photosensitizer properties.

This combination results in various novel biomedical and material sciences applications such as PDT, PTT, PACT, photochemical internalization (PCI), drug delivery, fluorescence, and photoacoustic imaging.

Additionally, in many cases, the self-assemblies of the two combined moieties can favour a defined organisation of the porphyrin, which is helpful for mimicking biological light-harvesting systems, creating artificial ones, and transferring electron in optoelectronic devices and solar cells.

As a result, strategies to synthesise these interesting, multifunctional, and modulable bioconjugates together with related applications have been investigated and reported in literature with increasing attention and relevance. Within the bioconjugation reactions listed in the previous section, a few of them were indeed employed for peptide-porphyrin conjugation, as reported by Pathak and Wu.[105,106]

1.4 PROJECT AIMS AND OBJECTIVES

In this context, peptide-porphyrin bioconjugates were identified as powerful medicinal chemistry tools, that require further improvements in terms of bioconjugation strategies and a broader exploration in term of applications purposes. Therefore, the aim of this research project was to develop a novel bioconjugation reaction that would allow straightforward access to novel examples of these conjugates for several applications.

Specifically, the objectives were:

I.   Explore and optimise the S_NAr between hexafluorobenzene (HFB) and N-acetylcysteine (NAC) as model reaction for the conjugation of cysteine-containing peptides to HFB and, ultimately, to polyfluorinated porphyrins.

II. Apply the new peptide-porphyrin bioconjugation strategy for the coupling of self-assembling peptides and porphyrins, characterise these newly synthesised peptide-porphyrin conjugates and study their properties.

III. Apply the new peptide-porphyrin bioconjugation strategy for the porphyrin-templated cyclisation of cysteine-rich peptides, characterise these newly synthesised porphyrin-templated cyclic peptides and study their properties.

Each of the following experimental chapters (Chapters 2, 3, and 4) will correspond to one of the above objectives, and general conclusions were included afterwards (Chapter 5). The Materials and Methods section (Chapter 6) will be dedicated to experimental techniques, methodologies, synthetic procedures, as well as compound characterisation. A summary table (Table 25) listing all the peptide sequences employed for this work was also included at the beginning of Chapter 6. Finally, Chapter 7 is an appendix that includes additional experimental work not strictly related to the main research project.

2 Expanding the scope of thiol-fluoride substitution

Nucleophilic aromatic substitution (S_NAr) on polyfluoroaromatics by means of differently substituted thiols is a well-known reaction in organic chemistry, but its applications in the arylation of peptides have been relatively rarely explored. Similarly, polyfluorinated porphyrins have been conjugated to many compounds exploiting this chemistry but surprisingly few examples report conjugation to cysteine and cysteine-containing peptides. Nevertheless, this is a potentially useful bioconjugation tool because cysteine can be naturally present or easily introduced in peptide sequences and *meso*-derivatised porphyrins with polyfluorinated rings are synthetically accessible.

This work started with the investigation of the reaction between hexafluorobenzene (HFB, **30**) and N-acetyl cysteine (NAC, **31**), used as model system for screening different base-solvent combinations with a One Variable At a Time (OVAT) approach (I in Figure 14). Contextually, Lewis Halsall performed a full factorial design (FD) coupled with ^{19}F NMR to explore the parameters that predominantly affect product outcomes.[107] The complexity of the system was then increased by introducing cysteine-containing peptides (II in Figure 14), and this optimisation resulted in novel conditions for fast HFB-peptide stapling and cyclisation. Finally, the same conditions turned out to be suitable to replace HFB with polyfluorinated porphyrins and develop a novel peptide-porphyrin bioconjugation strategy (III in Figure 14).
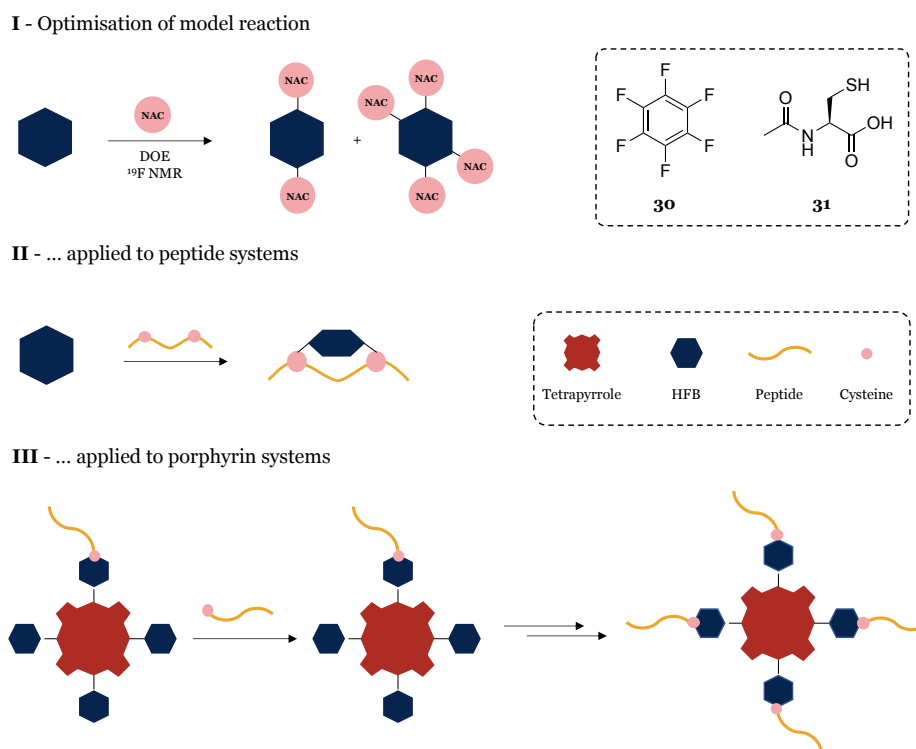

*Figure 14. Graphical representation of the research work described in this chapter. Chemical structure of hexafluorobenzene (**30**) and N-acetyl cysteine (**31**).*

2.1 NUCLEOPHILIC AROMATIC SUBSTITUTION

Arylation chemistries are example of a toolbox developed for small molecules that respect most of the requirements for a bioconjugation strategy and can be successfully employed as such. Among the transformations of aromatic rings, S_NAr exploits the reactivity of electron-deficient aryl halides with various nucleophilic groups encountered in biomolecules.

As illustrated in Scheme 8, during the S_NAr reaction, a nucleophile (Nu) attacks an electron-poor aromatic molecule (**32**), resulting in the substitution of a leaving group. This reaction proceeds *via* an addition-elimination mechanism with the intermediate formation of the Jackson-Meisenheimer complex (**33**), a reactive negatively charged adduct between the arene and the nucleophile.[108] The rate-determining step for this reaction is the initial attack of the nucleophile to aromatic ring, which disrupts the aromaticity. Therefore, the introduction of electron-withdrawing groups (EWG) that stabilise the negative charge of the intermediate, activates the aromatic ring and increases the reaction rate.

32   **33**   **34**
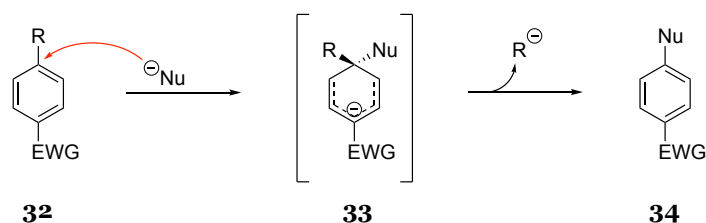
Scheme 8. Reaction scheme for the Jackson-Meisenheimer complex formation.

Because the leaving group loss is not rate-determining, highly electronegative atoms such as fluorine are often used to increase the ring electrophilicity and promote the nucleophilic attack, despite they would behave as poor leaving groups at a later stage. This effect is amplified when more EWGs are present: in polyfluoroaromatics, the reactivity is dominated by S_NAr and fluoride displacement can take place with a wide range of nucleophiles.[109]

2.1.1 Organic Fluorine and Cysteine

The extreme electronegativity and small size of the fluorine atom determines its great reactivity in organic chemistry and the high strength of the C-F bond.[110] Despite the limited presence of organic fluorine in natural products, fluorination is a common strategy in medicinal chemistry for the modulation of different properties such as molecular conformation, binding affinity, pKa, lipophilicity, target interaction, membrane permeability (particularly the central nervous system), metabolic stability and bioavailability. Radioactive ^{18}F labelling has also been successfully exploited for positron emission tomography (PET). Therefore, there are many fluorine-containing drugs and diagnostics used in medicine.[111]

The study of fluorine-containing molecules is favoured by the aid of ^{19}F nuclear magnetic resonance (^{19}F NMR). This is a fast, easily interpretable, quantifiable, and reproducible technique. ^{19}F NMR allows for easier characterisation of chemicals compared to ^{1}H NMR, because there are no background fluorine resonances from solvents or side components of the sample and therefore less spectral crowding. This aspect makes a big difference when

analysing biological samples and promoted biological application of ^{19}F NMR such as metabolic studies, binding studies, and structural analysis of biomacromolecules.[112]

This practical spectroscopic technique has been a useful tool for the development of this work and could also be for future applications. As will be discussed in this chapter, reaction monitoring and optimisation of the S_NAr on polyfluoroaromatics was performed using ^{19}F NMR because the number and symmetry of resonances is a direct indicator of the ring substitution degree on the molecule, hence of product formation. Furthermore, fluorine signal integration is quantitative and allows for accurate quantification of the relative amount of each product.
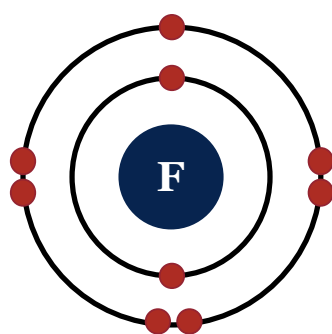
- Atomic number : 9
- Atomic mass: 18.998403 u
- Electronic configuration: $1s^2\ 2s^2\ 2p^5$
- Atom radius: 147 pm
- Electronegativity: 3.98
- C-F bond energy: 485 kj/mol
- C-F bond length: 135 pm

Figure 15. Chemical properties of the fluorine atom. Data are from [113,114].

The peculiar free thiol on cysteine side chain is predominantly in the deprotonated form at basic pH (pKa ~9), and it is more acidic than other common nucleophiles found in proteins and peptides, such as the hydroxyl group of serine (pKa ~13) or the phenol group of tyrosine (pKa ~10), even if these values are variable depending on the local environment. Compared to oxygen and nitrogen, the lower charge density and higher polarisability of sulfur make it form bonds with a more covalent nature.

The consequent higher reactivity of thiolates towards electrophiles, such as maleimides and alkyl halides, combined with the lower frequency of cysteines (1.5% of relative abundance in proteins), favours chemoselective cysteine modification in the presence of other amino acid side chains and make this AA one the most convenient players in bioconjugation. Disulfide formation, alkyl halide substitution, and maleimide coupling are the main strategies employed for the chemoselective derivatisation of cysteines.
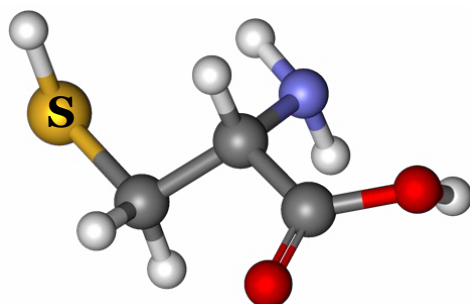
- Thiol pka ~ 9
- Relative abundance in proteins: 1.5%
- Sulfur electronegativity: 2.58
- S-S bond energy: 226 kj/mol
- S-S bond length: 205 pm
- C-S bond energy: 272 kj/mol
- C-S bond length: 182 pm

Figure 16. Chemical properties of sulfur and cysteine. Data are from [95,114]

2.1.2 Thiol-fluoride nucleophilic aromatic substitution on HFB

HFB is the most elementary perfluoroaromatic system and, besides a few other applications,[115–117] it has been frequently employed as substrate for thiol-fluoride substitution.
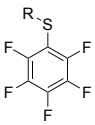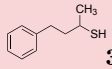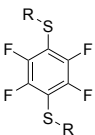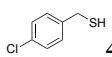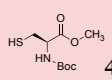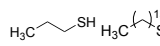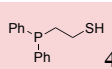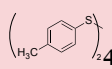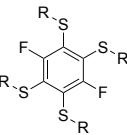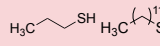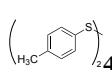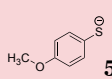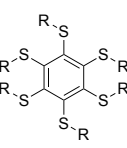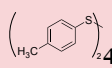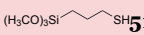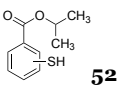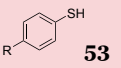
The examples of reactions between HFB and different thiols summarized in Table 1 show how the level of thiol-fluoride substitution can potentially be controlled by careful selection of the reagents' nucleophilicity and the reaction conditions. When HFB reacts with an initial thiol nucleophile, all six positions are equivalent. However, when monosubstituted pentafluorobenzene (**35**) reacts with a second thiol, the *para* substitution is favoured over *ortho* and *meta*. This is explained by the higher electron density and sulfur contribution to intermediate anion stabilisation in *ortho* and *para* position, as represented by the mesomeric structures highlighted in Scheme 9, which cannot be drawn when the substitution happens on *meta*. Discrimination between *ortho* and *para* position, both stabilising the intermediate anion, can be mainly ascribed to steric effects.

A preferred second substitution in the *para* position promptly affords disubstitution products with a 1,4-regiosubstitution pattern (**36**). For this reason, there are only a few approaches reported to selectively obtain mono-substituted pentafluorophenyl, while disubstitution is generally performed under mild conditions at room temperature. Similarly, tri- and penta-substituted products are rarely isolated in significant quantities. Stronger bases, longer reaction times, higher temperatures and a greater amount of thiol can orient the reaction toward 1,2,4,5-tetrasubstitution (**37**), while 1,2,3,4,5,6-hexasubstitution (**38**) can be achieved with even harsher and vigorous conditions. It should be noted that aromatic thiolates (*i.e.* thiophenol or derivatives) are stronger nucleophiles than aliphatic thiols and their higher reactivity is reflected by the reaction conditions.
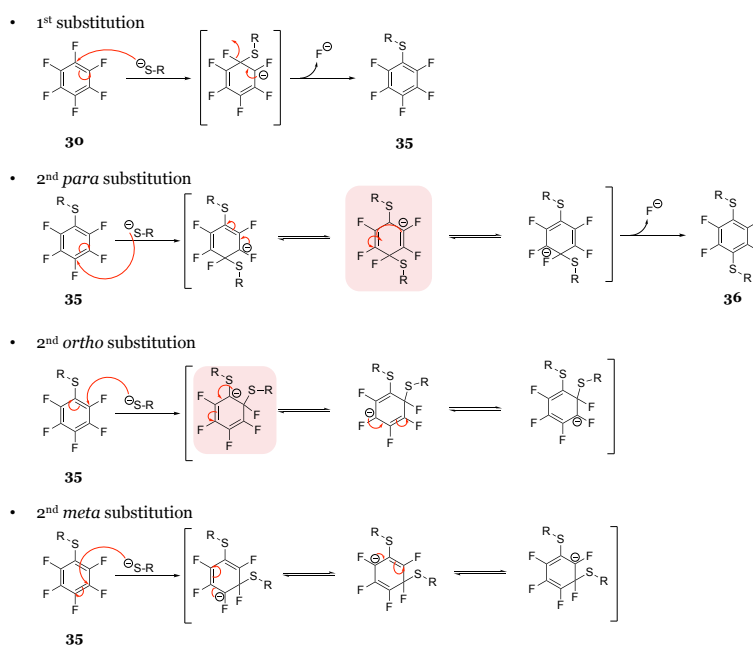
Scheme 9. Mechanism for S$_N$Ar regioselectivity on HFB.

Table 1. Summary of reactions between HFB and different thiols reported in the literature. Refer to Chapter 8 for abbreviations.

Product	Conditions	Thiol	Thiol structure	Ref
35	K_2CO_3, DMF, RT, 4.5 h	1 eq.	**39**	[118]
	THF, -40 °C	1 eq.	**40**	[119]
36	NaH, DMF, RT, overnight	2 eq.	**41**	[120]
	TRIS base, DMF, RT, 3-5 h	2 eq.	**42**	[121]
	TRIS base, DMF, RT, 4.5 h	2 eq.	**43**	[51]
	Cs_2CO_3, NBu_4Cl, THF, RT, 1-3 days	2 eq.	**44** **45**	[122]
	Cs_2CO_3, NBu_4Cl, THF, RT, 18 h	2 eq.	**46**	[123]
	Et_2O, RT, 2 days	2 eq.	**47**	[124]
	$(RhH(PPH_3)_4$, dppBz, PPh_3, PhCl, 80 °C, 4 h	1 eq.	**48**	[125]
37	DMI/EtOH (1:1), 80 °C, 75 h	4 eq.	**40**	[126]
	Cs_2CO_3, NBu_4Cl, THF, RT, 3-14 days	4 eq.	**44** **49**	[122]
	$(RhH(PPH_3)_4$, dppBz, PPh_3, PhCl, 80 °C, 12 h	2 eq.	**48**	[125]
	DMI, 40 °C, 72 h	4 eq.	**50**	[127]
	HMPA, 0 °C, 0.1 h	4 eq.	**47**	[128]
38	$(RhH(PPH_3)_4$, dppBz, PPh_3 , PhCl, 80 °C, 48 h	3 eq.	**48**	[125]
	HMPA, 0 °C, 1.5 h	10 eq.	**47**	[128]
	NaH, THF, RT, 24 h	6 eq.	**51**	[129]
	Cs_2CO_3, DMI, 35-50 °C, 3-6 days	10 eq. 15 eq.	**52**	[130]
	NaH, DMI, RT, 6 h - 3 days	18 eq.	**53**	[131]

2.2 S$_N$Ar BETWEEN HEXAFLUOROBENZENE & N-ACETYL CYSTEINE

Fluoride substitution on polyfluoroaromatics can be performed with nucleophiles such as alcohols, amines, thiols and selenols that are frequently present in natural and bioactive compounds. It is generally stoichiometric, efficient, fast, and performed with a broad range of solvents, depending on the reagent solubility. The diagnostic nature of these substrates using ^{19}F NMR can be exploited to monitor reaction outcomes. However, most of the reported procedures for thiol-fluoride substitution reviewed in the introduction are not ideal when working with peptides. For example, the use of temperatures higher than 50 °C, metal catalysis and a large excess of thiols should preferably be avoided for bioconjugation reactions involving molecules with low chemical stability. Therefore, subject to proper upgrades regarding chemoselectivity and biocompatibility, this can be considered a valuable bioconjugation tool.

New bases and solvents to perform this chemical transformation were explored, with the final aim of obtaining selective conversion to either di- or tetra-thiol-substituted fluorobenzenes under mild and 'peptide-friendly' conditions.

As starting point for these preliminary studies, a simple and repeatable experiment combined with a fast detection method were necessary. To meet these requirements, the model reaction between HFB (**30**) and NAC (**31**) was studied using ^{19}F NMR (Figure 17A).
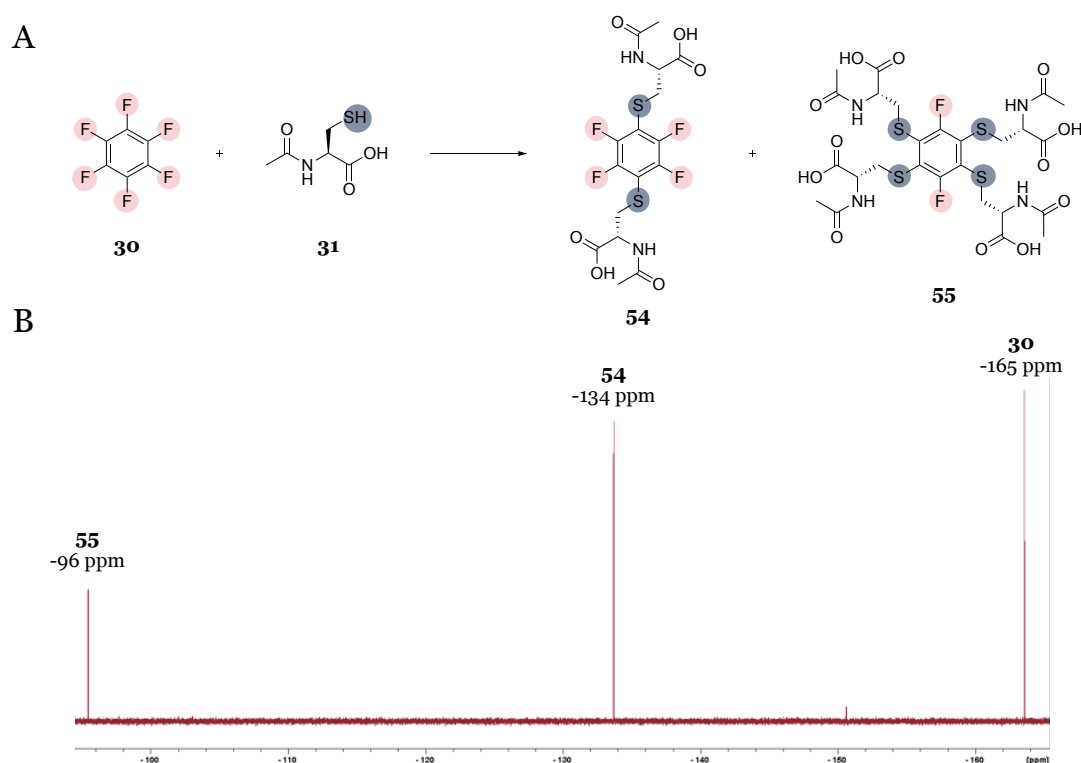
Figure 17. A) Reaction scheme for HFB-NAC model reaction; B) Typical NMR signals for HFB, disubstitution product and tetrasubstitution product.

NAC and HFB are both accessible, cheap, and easy to handle. The use of NAC is a good model for cysteine-containing peptides and the acetylated amine reproduces secondary amides of the peptide backbone. Most importantly, HFB is a good model for more complicated polyfluorinated aromatic rings. This fluorinated species contains 6 equivalent fluorine nuclei with a singlet signal at -165 ppm (Figure 17B). The same applies for the two major products: the disubstituted (**54**) and tetrasubstituted (**55**) rings give two singlets with chemical shifts of -134 ppm and -96 ppm respectively (Figure 17B). Relative quantification of the main reaction products can, therefore, be performed by integrating the signals and dividing by the number of fluorine nuclei.

2.2.1 Optimisation of a chemical process

In different research and development areas, the optimisation of a process is usually performed by changing the variable parameters individually and evaluating the outcome each time. This approach, called One Variable At a Time (OVAT), gives very specific information, but only for the conditions in which the experiments have been performed.

As a better alternative, a powerful statistical technique named Design of Experiment (DOE) can be employed. With a limited number of experiments, DOE provides high-quality information in the whole experimental domain.

The difference between these two approaches can be effectively represented as in Figure 18, where the cubes are the experimental domain, and the blue/red dots are the experiments. In addition, DOE reflects the interactions between controllable factors, while OVAT would only be valid if the variables are totally independent from each other.[132]
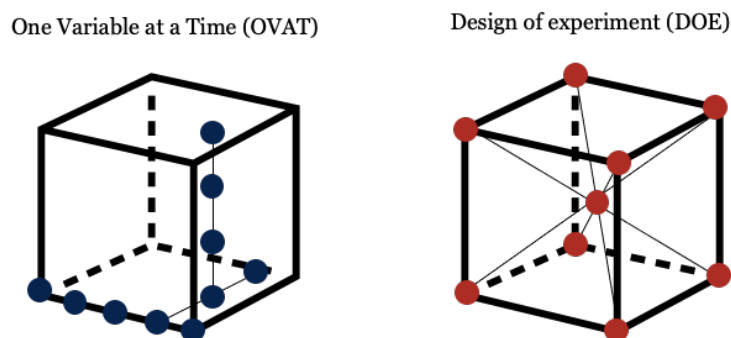
Figure 18. Graphical representation of the experimental space difference between OVAT and DOE approaches.

Therefore, DOE is an advantageous tool with respect to time, money and relevancy of the results and can be usefully applied in the chemistry field. For example, time, temperature, pH, ionic strength, volumes, and concentrations are just a few variables that play a role within, or may affect the outcome of, a standard laboratory reaction or chemical process. Consequently, these factors are frequently considered during optimisation processes and their mutual interaction is highly likely. DOE allows their individual evaluation, in addition to the interactions between them to be observed.

There are different types of DOEs, depending on the process under investigation, number of variables, stage of the optimisation and resources availability. A factorial design (FD) in which each variable has two levels is the simplest experimental design. This is usually referred as 2k FD, where k is the number of variables under study and the number of experiments is equal to 2k. As the number of variables or the number of levels for each variable increases, the total number of experiments necessary to complete a full FD drastically increases as well.

Contrary to the OVAT approach, in which one variable is changed while the others are kept constant, in a FD every variable is changed and combined with different values of the other variables. This allows estimation of the interactions between variables and analysis of the effect of changing one variable in relation to the concurrent values of the other variables.

2.2.2 OVAT screening

When planning the S_NAr between HFB and NAC, the first aspects to be defined are the nature of the employed base and solvent. These two variables were initially screened with an OVAT approach (Figure 19), while other parameters such as reaction time, reagents ratio, and temperature were considered in a second and more comprehensive optimisation stage.
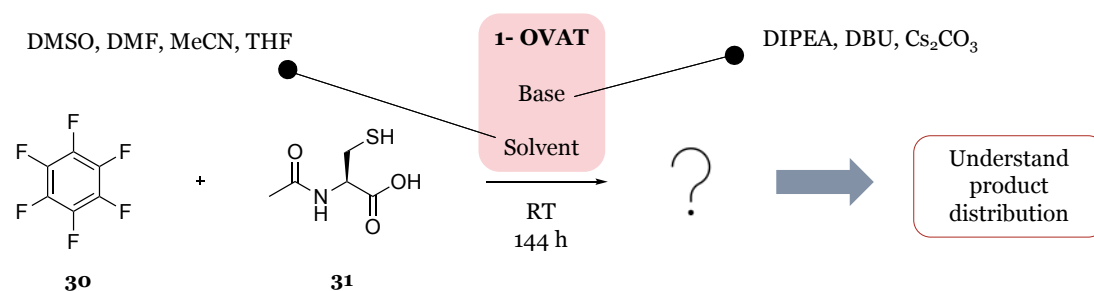
Figure 19. Reaction and parameters investigated during the initial OVAT screening.

A base is generally employed in this chemical transformation to generate a more reactive anionic nucleophile. N,N-diisopropylethylamine (DIPEA), 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) and Cs_2CO_3 were the explored bases, to include examples of molecules with different basic strength, steric hindrance and organic/inorganic nature.

Contextually, an aprotic medium is preferred because a protic solvent could interact with the base as well as promote hydrogen bonding between anions, inhibiting their nucleophilic behaviour.

The solvent should also preferably be polar as this better solvates the charged intermediate. When computationally simulating the solvent effects on the S_NAr mechanism, the Jackson-Meisenheimer complex has been reported to be more stable in a polar aprotic solvent, rather than in a polar protic one.[133]

Dimethyl sulfoxide (DMSO), dimethylformamide (DMF), acetonitrile (MeCN) and tetrahydrofuran (THF) were selected within the polar aprotic solvents typically used for these reactions. They have different properties in terms of dielectric constant, viscosity, and boiling point reported in Table 2.[134]

Table 2. Chemical-physical characteristics of the solvent employed during the initial OVAT screening.

Solvent	Dielectric constant (20 °C)	Viscosity (cP, 20°C)	Boiling point (°C)
THF	7.6	0.55	66 °C
MeCN	37.50	0.37	82 °C
DMF	36.70	0.92	153 °C
DMSO	46.7	2.24	189 °C

The reaction in Scheme 10 was performed at room temperature with i) HFB/NAC molar ratio of 1:4, ii) HFB/base molar ratio of 1:20, and iii) monitoring of the level of substitution after 4 h, 24 h and 144 h.
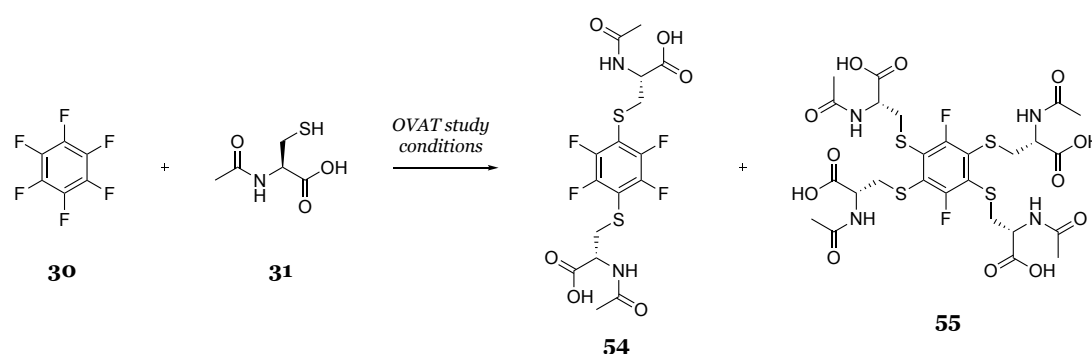
Scheme 10. Reaction scheme for di and tetra substituted products.

The results of this initial screening are reported in Table 3 as percentages of unreacted HFB and conversion to di- or tetrasubstituted product at the longest time point (144 h).

Table 3. Results of the initial OVAT screening are presented as percentage of product **30**, **54** *and* **55** *under different reaction conditions.*

	DIPEA				DBU				Cs₂CO₃			
	THF	MeCN	DMF	DMSO	THF	MeCN	DMF	DMSO	THF	MeCN	DMF	DMSO
% of 30	100	19	23	30	58	-	33	43	30	20	22	24
% of 54	-	62	77	70	42	40	29	32	53	60	78	52
% of 55	-	19	-	-	-	60	38	25	17	20	-	24

From a solvent point of view, THF was probably the least effective for this reaction: no substitution products were observed to form using DIPEA and, even if using DBU (42% of **54**) and Cs₂CO₃ (53% of **54** and 17% of **55**) was slightly better, conversion to disubstituted and tetrasubstituted products was lower compared with the other conditions. MeCN performed generally well, was best when combined with DBU (40% of **54** and 60% of **55**) and exhibited a decrease in conversion to tetrasubstituted product when the base was Cs₂CO₃ (20 % of **55**) or DIPEA (19 % of **55**). Cs₂CO₃ in DMF led to the formation of a gel and precludes the reaction,

whose outcome was better if the solvent was combined with DBU or DIPEA. DMF/DIPEA could be a valid choice if aiming at disubstitution (77% of **55**). DMSO worked well with every base, yielding a clean disubstitution with DIPEA (70% of **55**) and suggesting promising level of tetrasubstitution if Cs_2CO_3 was employed (52% of **54** and 24% of **55**).

From a base point of view, DIPEA (pK_a of conjugate acid in H_2O is 10.75[135]) gave the lowest reactivity and slowest conversion, which however, may allow the selective isolation of disubstituted product. This is interesting as the combination DMF/DIPEA is one of the most employed in perfluoroaryl peptide stapling (see section 2.3). DBU, being a stronger (pKa in H_2O is 11.9[136]) and more hindered base than DIPEA, promoted a faster reaction and usually a higher amount of tetrasubstitution product. Cs_2CO_3 (pKa of conjugate acid in H_2O is 10.33[135]) was difficult to solubilise in DMF, but when employed in combination with an appropriate solvent seemed to promote higher order substitution, similarly to DBU.

Overall, the general trend was that solvents with higher polarity and stronger bases promoted the substitution. Differences in solvent boiling point and viscosity did not appear to have a significant effect in the reactivity but might be relevant in subsequent steps such as reaction workup. In conclusion, selective conversion to disubstituted or tetrasubstituted HFB could potentially be controlled by appropriate choice of the reagents and conditions employed. 1,4-disubstitution (**54** in Figure 20) could be obtained cleanly using combinations of DMF/DIPEA and DMSO/DIPEA, while 1,2,4,5-tetrasubstitution (**55** in Figure 20) could be obtained using MeCN/DBU and DMSO/Cs_2CO_3. However, this higher substitution product was only obtained in moderate yield and further optimisation was required.
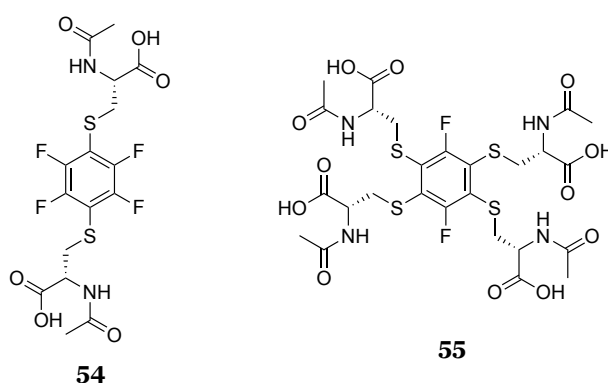
Figure 20. Chemical structure of di and tetra substituted products.

Some critical aspects of these experiments that require attention can be identified as HFB volatility, uncontrolled precipitate formation, and reaction sampling.

HFB relative volatility (boiling point= 80 °C), in combination with the small volumes employed and long reaction times, makes the exact reagents ratio difficult to control. The reaction vessel should be selected in accordance with the reaction mixture volume to have the lowest head-space volume possible. This should be sealed, especially for reactions performed over a long time or those under heating. Eventual loss of unreacted HFB would obviously affect the final reaction analysis and yield calculation.

In these model reactions, a precipitate was often observed, sometimes with yellow shades, usually when using caesium carbonate as the base. The precipitate was considered to likely be a caesium salt of the fluoride anion (*e.g.* CsF) that precipitates due to poor solubility, especially in solvents less polar than DMSO. This was confirmed by the presence of a signal with a chemical shift of -115/-120 ppm that has been observed previously in ^{19}F NMR.[137] From our experiments, the chemical shift of CsF was variable and could interfere with the analysis of reaction outcome and yield calculation if not properly accounted for. For example, dilution of the reaction mixture with water or organic solvents can shift its chemical shift up to -140 ppm. Literature data confirm that the F- chemical shifts are strongly solvent dependent.[138]

The identification of CsF as a side product was also supported by control reactions. By removing one of the reagents (HFB, base, or thiol) from the previously described reaction mixture, this peak was not observed. This signal is also not observed after work-up by HCl addition, in a reaction mixture without the nucleophile (solution of HFB and base), or without the base (solution of HFB and thiol). It was not possible to relate this event to any other experimental observation, chemical process or product observed. No undefined compounds, other than the ones related to reagents and products under investigation, were detected after LC-MS analysis. These experimental observations are in accordance with the hypothesis of fluoride ion detection, which is also supported by literature data. [137,138]

As consequence of the previous point, sampling of a non-homogeneous reaction mixture could lead to variable analytical results. Sampling may be performed on the supernatant only, on the precipitate only, or on a homogeneous mixture after complete solubilisation of the precipitate. In this case, the solvent employed for dilution of the reaction mixture influences the NMR analysis. For example, a sample of a DMSO reaction diluted with DMSO gives a different spectrum compared to the same sample diluted with water or HCl solution after work-up.

2.2.3 OVAT screening of water-based conditions

Separately to the above screening of organic solvents, water-based reaction conditions were also explored, following a similar OVAT approach (Figure 21). This option would be preferable in terms of compatibility with biomolecules, costs, environmental impact, safety, and impurity profiles. These conditions would also benefit hydrophilic peptides but could be challenging for more hydrophobic AA sequences or organic compounds and reagent with low water solubility.
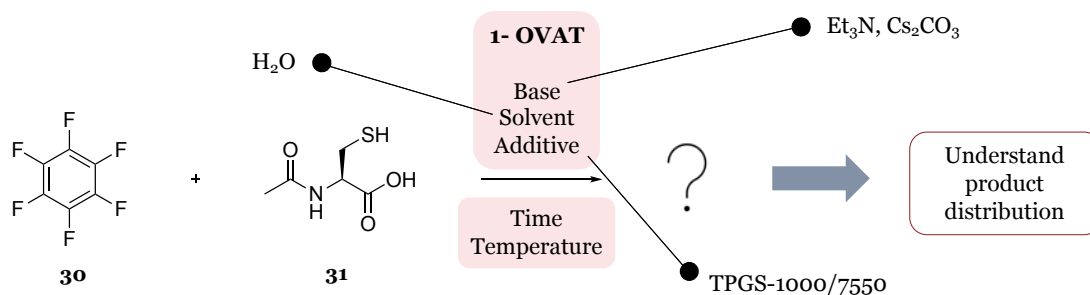
Figure 21. Reaction and parameters investigated during the OVAT screening of water-based conditions.

Despite the lack of literature examples regarding HFB reactions in water, the work from Liu *et al*.[139] described how 2,3,5,6-tetrafluoroterephthalonitrile (4F-2CN, **56**) reacted with glutathione (GSH, **57**) in phosphate buffer (PB), 100 mM, pH 7.4, at 37 °C to give the di- (**58**) or tetrasubstituted (**59**) product in a few hours (Scheme 11).
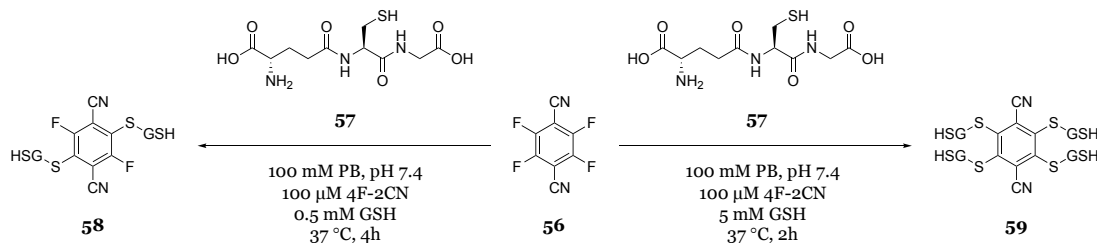
*Scheme 11. Reaction scheme and water-based conditions for 4F-2CN (**56**) substitution with GSH (**57**) employed by Liu and co-workers.*[139]

As a starting point, this procedure was reproduced using HFB (instead of 4F-2CN) and cysteine, NAC (**31**) or GSH (**32**). No substitution was observed in phosphate buffered saline (PBS), nor in H_2O plus base (data not shown). This different outcome is probably ascribable to the different solubility of 4F-2CN and HFB in water ($logP_{HFB}$= 2.63 and $logP_{4F-2CN}$= 1.81, calculated using 2022 Molinspiration Cheminformatics[140]).

Considering the advantages that an aqueous medium could offer, further time and resources were allocated toward the identification of a possible solution. It is reported that small amounts of a nanomicelle-forming amphiphile in water can provide a lipophilic medium in which reactions could take place.[141,142] This procedure has been applied to a broad range of chemical processes employing water-insoluble organic substrates, to improve their effectiveness in water at room temperature.

If the low solubility of HFB in water was precluding NAC from reacting, this may be solved by a reaction mixture comprising a commercially available surfactant that, upon dissolution in water, spontaneously forms micellar reactors. D-α-Tocopherol polyethylene glycol succinate (TPGS) 1000 (**60**) and TPGS 750-M (**61**) – made by a lipophilic tocopherol (or vitamin E) moiety (Figure 22, red box), a succinic acid linker and a hydrophilic polyethylene glycol (PEG) moiety of variable length (Figure 22, blue box), PEG-1000 or MPEG-750 (n~15) – were chosen for this purpose.[143]
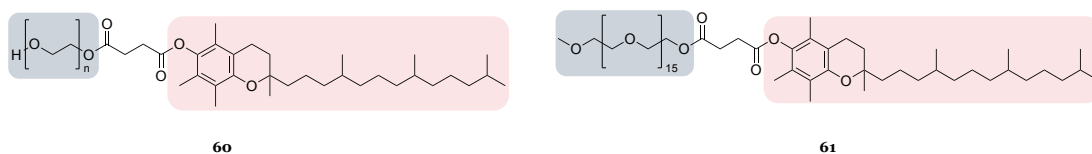
*Figure 22. Chemical structure of TPGS-1000 (**60**) and TPGS-750M (**61**).*

Reactions were performed with 5 equivalents of nucleophile (both Cys and NAC), 5 or 10 equivalents of base, and 2% wt surfactant, at room temperature or at 45 °C.[141] The experimental conditions, adapted from Isley *et al*.,[141] and the corresponding product distribution (in %) are reported in Table 4. Direct sampling and analysis of the reaction

mixture was performed as previously described, because a complete workup procedure would lead to HFB loss and without unreacted HFB in the RM, the percentage of conversion cannot be calculated.

*Table 4. Results of the initial OVAT screening of water-based conditions. Percentages of starting reagent (**30**) or products (**54/55**) were calculated from NMR spectra as described in this section.*

Thiol	Surfactant	Base	Temp	Time	% of 30	% of 54	% of 55
Cys (5 eq)	TPGS-1000	Cs_2CO_3 (5 eq)	rt	15 h	95	5	-
	TPGS-1000	Et_3N (5 eq)	rt	24 h	95	5	-
	TPGS-1000	Et_3N (5 eq)	45	20 h	90	10	traces
	TPGS-1000	Et_3N (5 eq)	45	96 h	25	70	5
	TPGS-750M	Et_3N (5 eq)	rt	72 h	65	35	-
	TPGS-750M	Et_3N (5 eq)	45	96 h	20	75	5
NAC (5 eq)	TPGS-750M	Et_3N (5 eq)	rt	72 h	75	25	traces
	TPGS-750M	Cs_2CO_3 (10 eq)	rt	72 h	45	50	5

A limited amount of product was formed, probably because the TPGS aqueous suspension with lipophilic micelles could solubilise HFB, but NAC would still be in the water phase preventing the complete interaction of the reagents. Trends seemed to indicate an increased reactivity with heating and a higher amount of base, as well as improvements with TPGS-750M over TPGS 1000, as reported for other S_NAr on different substrates.[141] However, as previously noted, heating could affect HFB evaporation and its final concentration.

Sodium dodecyl sulfate (SDS, surfactant) and tetra butyl ammonium bromide (TBAB, phase transfer catalyst) were subsequently investigated in the procedure but no reactions were observed.

Overall, these results indicated that that, contrary to a water-only reaction mixture, supplement of TPGS-based micelles can lead to product formation with possible margins of improvement. On the other side, these are slower, more expensive (because of the surfactant) and lower yielding conditions compared to those employing organic solvents, which appeared to suit best our final scope and were preferred for further developments.

2.2.4 Factorial design optimisation

After identifying DMF/DIPEA, MeCN/DBU and DMSO/Cs_2CO_3 as the three best base-solvent combinations to be applied for peptide stapling using HFB, DMSO/Cs_2CO_3 was selected as the most promising for further detailed optimisation of HFB-NASC model reaction and future application (Figure 23).

This system is the association of a highly polar and aprotic solvent, that enable reagents and products to remain in solution, and a strong base, that increase reaction speed as well as formation of higher-order substitution products (tetrasubstitution).

For further optimisation of these conditions, a 2-level 5 factorial full design space was planned.[107] Based on the observations described in section 2.2.2, 32 (2^5) individual experiments were performed from the combination of low (min) and high (max) values for the following variables:

- reaction time (4/168 h);

- temperature (25/50 °C);

- solvent volume (5/10 mL);

- molar equivalent of base (8/20 eq);

- molar equivalent of thiol (4/10 eq).
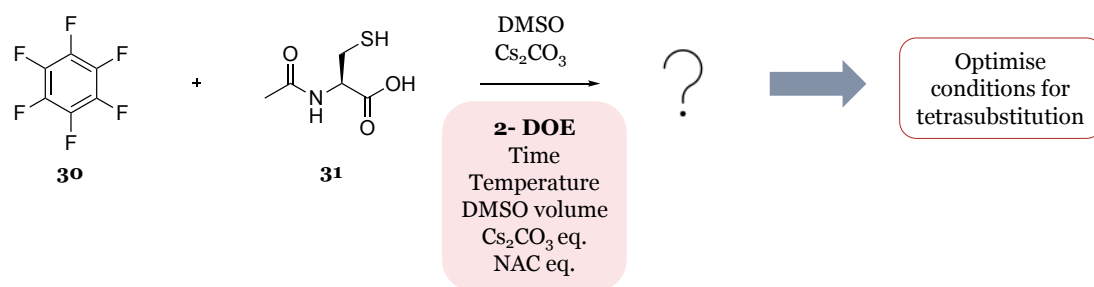
Figure 23. Reaction and parameters investigated during the DOE optimisation.

With a small series of experiments and a quick statistical analysis, this DOE approach allowed the exploration of the influence of each variable and provided indication on the relevancy of their interactions on product outcome. The results were analysed and are discussed below with the final aim of understanding what factors can favour tetrasubstitution over disubstitution.

Opposite effects and trends would obviously reflect the influence of each factor in promoting 1,4-disubstitution (graphs are reported in section 6.2).

Data were processed with Minitab while analysis of variance (ANOVA) led to P-values of the means <0.05, as indication of statistically relevant results.

A first look at the main effect plots (Figure 24) suggested time and base concentration as the major factors that influenced HFB substitution. In these graphs, a steep slope of the line connecting the mean outcome for high and low values of the variables was indication of their significant influence on the experiment. Increasing the values of time and base concentration would likely drive the reaction towards tetrasubstitution. Quite surprisingly, a lower-to-insignificant effect on product selectivity was reported for the molar equivalents of thiol, reaction temperature and reaction concentration – the line slope is shallow.
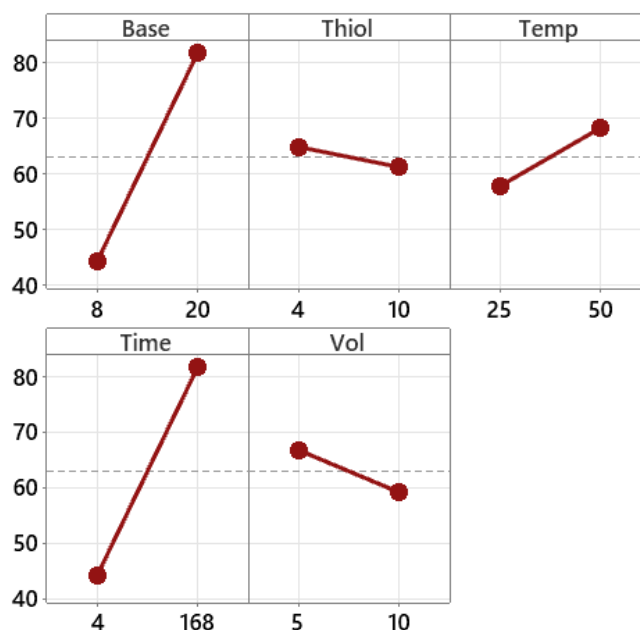
Figure 24. Main effect plots. On the x axes are reported the low/high values explored for each variable and on the y axes the percentage of conversion to tetrasubstitution.

Main effects plots were used to express the mean effect of each individual variable on product conversion and to establish general trends, such as the tendency of these reactions to favour tetrasubstitution over disubstitution but did not show their relationship and were not detailed enough to indicate the optimised conditions for tetrasubstitution.

Factor	Name	Unit
A	Base	8/20 eq.
B	Thiol	4/10 eq.
C	Temp	25/50 °C
D	Time	4/168 h
E	Vol	5/10 mL
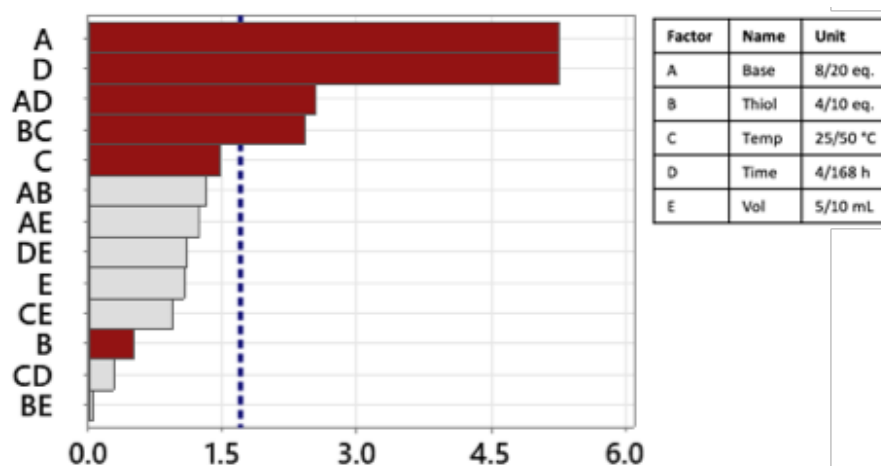
Figure 25. Pareto chart. The values on the x axis indicate the influence on the reaction of each factor from the y axis. The blue dotted line indicates the significance level.

The general observations from main effects plots were confirmed by the Pareto chart (Figure 25), which contextually highlighted the importance of combinations between base-time and thiol-temperature.

It is important to emphasise that, within a reaction system, each condition not only influences the product but contextually affects the other conditions as well. Therefore, a chemical system can only be superficially investigated using an OVAT approach. For example, thiol and

temperature had a low influence on the reaction outcome on their own, but their interaction was relevant.

Furthermore, the assumption that some variables were not a strong indicator of reaction outcome individually does not imply that they did not influence the reaction, but on the contrary a more complex interaction involving multiple factors and leading to intricate effects may take place.
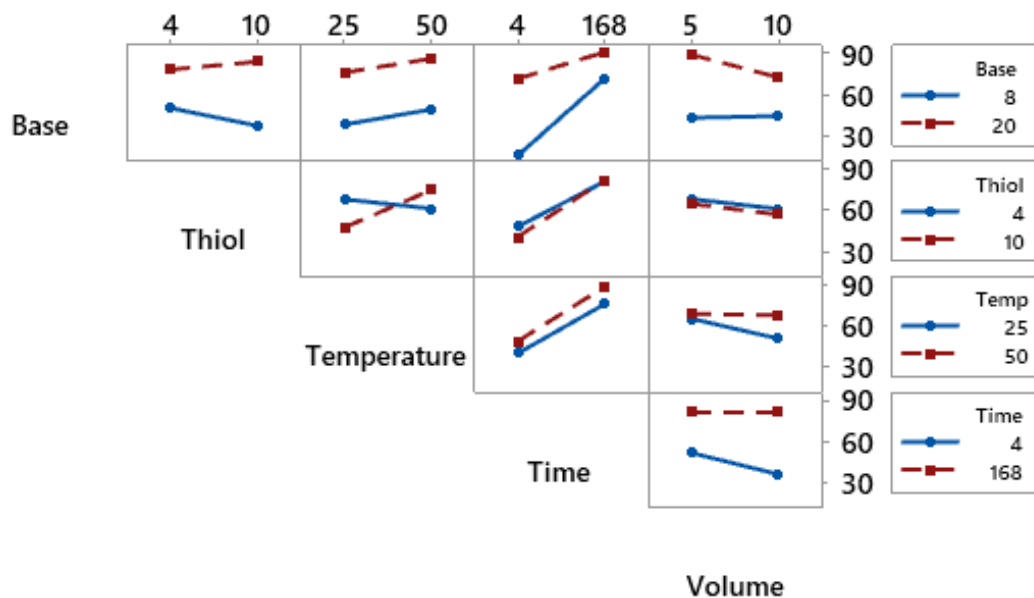
Figure 26. Interaction plots. On the x axes are reported the two values explored for each variable and on the y axes the percentage of conversion to tetrasubstitution.

The interaction plots (Figure 26) are a different representation of the above observations and can help in visualising the effect of every condition on each other. In these graphs, the distance between each line and the relationship between their slope should be considered. A larger distance between two lines corresponds to a greater impact of one variable on the outcome, while the smaller difference in slopes indicates a smaller interaction.

Overall, these results suggested that a greater amount of base and longer reaction times generally afforded a higher proportion of tetrasubstitution over disubstitution. However, practical conditions may vary from optimised conditions depending on specific requirements.

2.3 PEPTIDE STAPLING WITH HFB

When two deprotected cysteines are present within a general peptide sequence (**62**), these undergo 1,4-disubstitution on HFB (**30**) to give the cyclic structure **63**, generally without higher order substitution products (Scheme 12). Such transformations were typically performed employing the peptide in DMF with the addition of a base (Table 5). To date, no other examples employing different base-solvent combinations or alternative methodologies are reported.
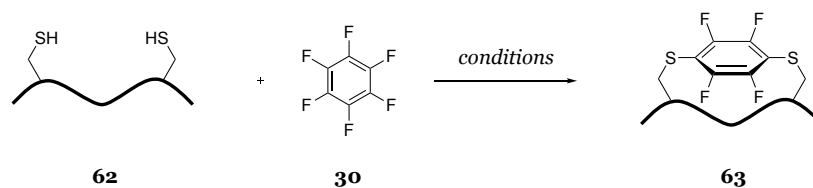
Scheme 12. General reaction scheme for HFB-templated peptide cyclisation. Refer to the example reported in Table 5 for peptide sequences.

Table 5. Summary of reactions conditions for HFB-templated peptide cyclisation reported in the literature.

Conditions	Peptide Sequence	Ref
TRIS-base, DMF, rt, 4.5 h	YCGGGCAL YCERSCNMK ITFCDLLCYYGKKK CNLLCEAKKLNDAQAPK	[51]
TRIS-base, DMF, rt, 2 h	From *i,i + 1* [IKFTNGLCCLYESKR] to *i,i + 14* [CIKFTNGLLYESKRC]	[144]
1. TRIS-base, DMF, rt, 30 min 2. GST, TCEP, pH=8, rt, 5 min	E(StBu)CG(GLKAG)$_x$C x= 2, 3, 4	[145]
TRIS-base, DMF, rt, 18 h	CDEETGEC CEETGC CDPETGEC CLDPETGEFC	[146]
TRIS-base, DMF, rt	HX$_1$EGCX$_2$TSCX$_3$X$_4$ HX$_1$EGTX$_2$TCDX$_3$X$_4$C	[147]
Et$_3$N, DMSO/DMF (1:1), rt, 14 h	CGNKRTRGC	[148]
DIPEA, DMF, RT, 4 h	YCGGGCAL FKACGKGCA	[149]
DIPEA, DMF, RT, 4 h	LTFCHYWCQLTS	[150]
TRIS-base, TCEP, DMF, rt, 2 h	WGKGCGKGUGKGCW	[50]
DMF, Gn·HCl, TRIS, PPh$_3$, RT, 24 h	ZLYSALANKCCHVGCTKRSLARFC	[151]

After introducing this stapling technique, the Pentelute Group made extensive use of polyfluoroaromatics for the modification of different bioactive peptides. It is worth mentioning the application of this strategy for i) protein tagging via regioselective cysteine arylation in water, both interface and non-interface;[152,153] ii) decafluorobiphenyl-templated macrocyclic peptides for enhanced blood brain barrier (BBB) penetration (Scheme 13);[154,155] iii) macrocyclization of non-helical peptides by cysteine crosslinking at different distances (from *i, i + 1* to *i, i + 14*);[144] iv) lysine crosslinking with perfluoroaryl reagents which are more

electron-deficient than HFB (*e.g.* diperfluorophenyl sulfone);[156] v) synthesis of the challenging A-chain of H2 relaxin after intramolecular disulfide replacement between Cys10 and Cys15 with HFB;[151] vi) macrocyclization of peptides up to 40 AA catalysed by glutathione S-transferase (GST);[145,157] vii) enzyme-independent regioselective bioconjugation via 'π-clamp' approach[158–160] and other studies to understand the peptide sequence effect on the propensity to react with perfluoroaromatics.[161,162]
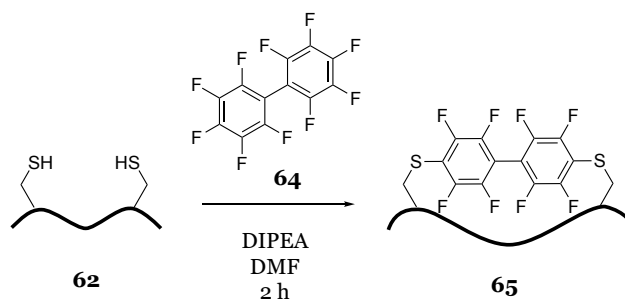
Scheme 13. General reaction scheme for decafluorobiphenyl-templated peptide cyclisation.

Pentafluoropyridine is another perfluoroheteroaromatic system that, exploiting S$_N$Ar reactivity, is employed in peptide chemistry as a tagging and protecting group for phenols (Scheme 14).[163] Given its higher reactivity towards nucleophiles compared to HFB, pentafluoropyridine can react with cysteine thiols, tyrosine hydroxyl groups and lysine side chains. Careful selection of solvents can increase chemoselectivity of thiol-fluoride substitution in reactions between pentafluoropyridine and peptides presenting these functional groups.[149]
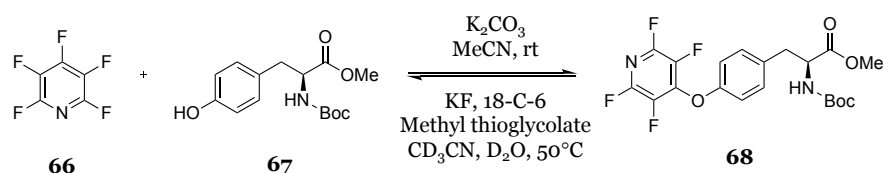
Scheme 14. Reaction scheme for pentafluoropyridine peptide tagging.

Figure 27 provides an overview of applications of HFB and similar polyfluoroaromatics in the peptide field. This is confirmation of the efficiency, chemoselectivity and robustness of the thiol-fluoride S$_N$Ar. These reactions also respect most of the requirements of click chemistry, being quantitative, having a wide functional group tolerance, and offering modulability to obtain chemoselectivity. However, they are not always as fast as click reactions.
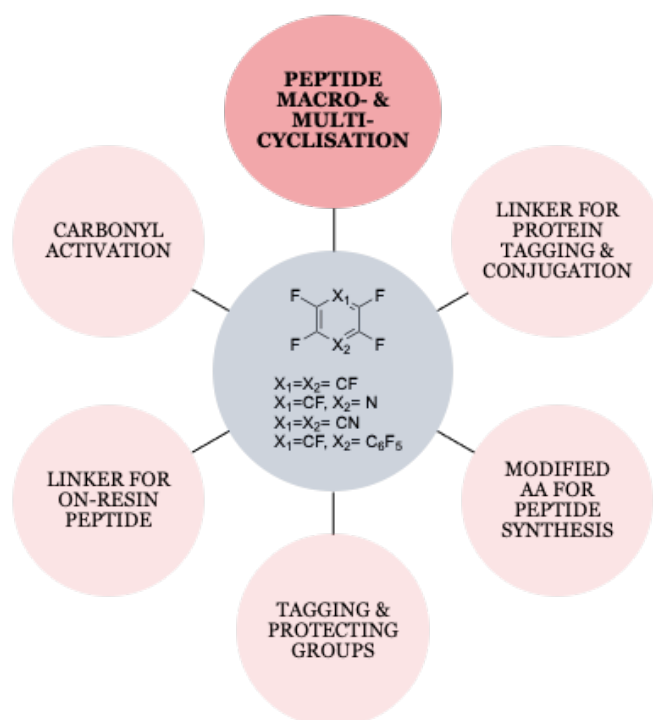
Figure 27. Perfluoroaromatic and perfluoroheteroaromatic compounds as tool for peptide bioconjugation, modification, stapling and cyclisation.

2.3.1 Disubstitution in solution

As the next step of the work, the outcomes of base-solvent combination screenings and factorial design optimisation studies were applied to the exploration of HFB-templated di-cysteine containing peptide stapling.

Peptides employed for the cyclization experiments on HFB, as well as the rest of the peptides used in this work, were synthesised with a microwave-assisted automated solid phase peptide synthesiser as C-terminal amides and at a 0.1 mmol scale, using Rink amide Pro-Tide resin and the conditions of Fmoc chemistry described in Scheme 15.

MeCN and DMSO proved to be effective in our model system and may provide alternative conditions for different applications, whereas this transformation is typically performed using DMF. These solvents could offer improved reagent solubility as well as cleaner and faster reaction. These additional features acquire even more relevance if aiming at achieving a higher order substitution for a multicyclisation strategy, or when performing the reaction on different substrates which are not soluble in DMF.
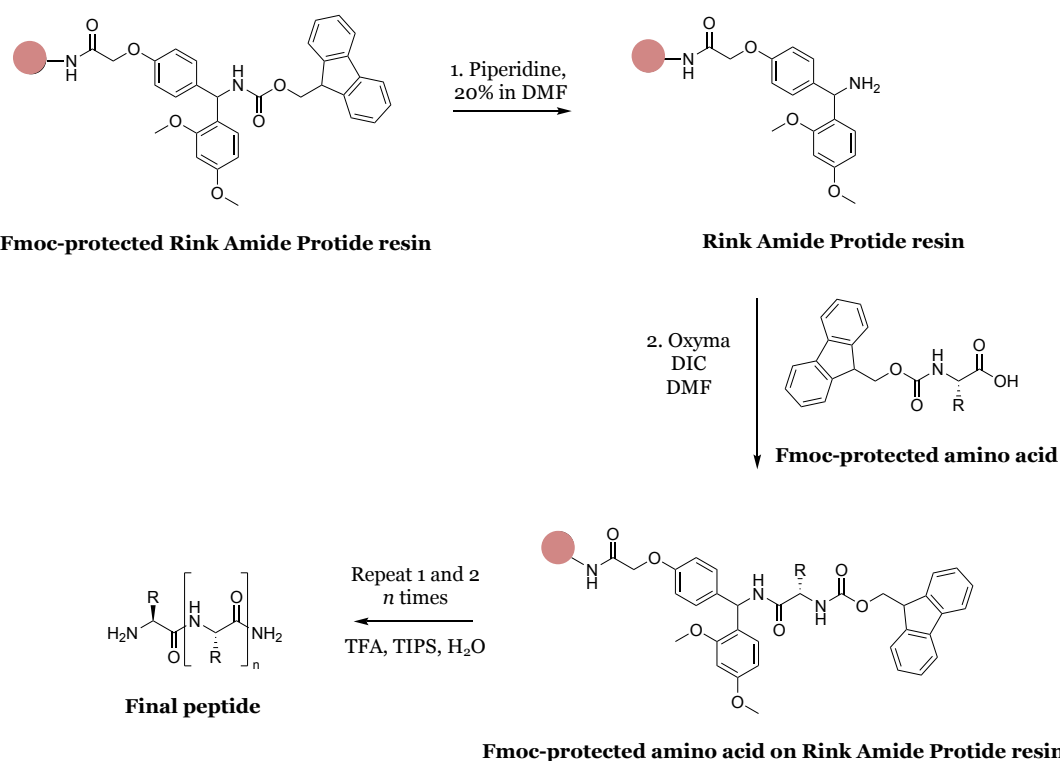
Scheme 15. *Reaction scheme and conditions for the Fmoc-SPPS of peptides amides used in this work.*

To test this hypothesis and explore applications in peptide stapling and cyclisation, the combinations MeCN/DBU and DMSO/Cs$_2$CO$_3$ were applied to the model di-cysteine containing peptide system **69** (Scheme 16 and Table 6). This sequence is derived from literature[51] and has been extensively used for similar studies.
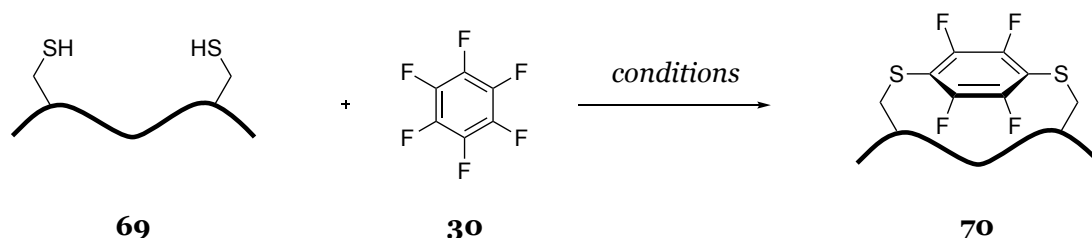
Scheme 16. *General reaction scheme for HFB-templated peptide cyclisation.*

Table 6. *Reactions conditions for HFB-templated peptide cyclisation explored in this work.*

Conditions	Peptide	C-terminus	Product
Cs$_2$CO$_3$, DMSO, rt, 1 min	H-YCGGGCAL-NH$_2$ (**69**)	Free peptide	**70**
DBU, ACN, rt, 1 h	H-YCGGGCAL-NH$_2$ (**69**)	Free peptide	**70**
DIPEA, DMF, rt, 18 h	H-YCGGGCAL-NH$_2$ (**69**)	Resin	**70**

Due to the early stage of reaction development and the increased complexity compared to the model system HFB/NAC, ^{19}F NMR analysis was used in combination with LC-MS for reaction

monitoring and product characterisation. In this case, the small reaction scale led to a diluted sample and residual TFA from peptide cleavage in the crude reaction mixture led to a strong fluorine signal, making ^{19}F NMR analysis more difficult.

By sampling the reaction mixture at different time points (1 min, 15 min, 30 min, 45 min, 1 h, 2 h, 4 h, 8 h, 24 h), a kinetic profile of product formation was produced (Figure 28). The cyclisation reaction between **69** and HFB turned out to be faster than HFB/NAC substitution and HFB-peptide stapling previously reported in DMF/DIPEA, hence the inclusion of particularly short time points.
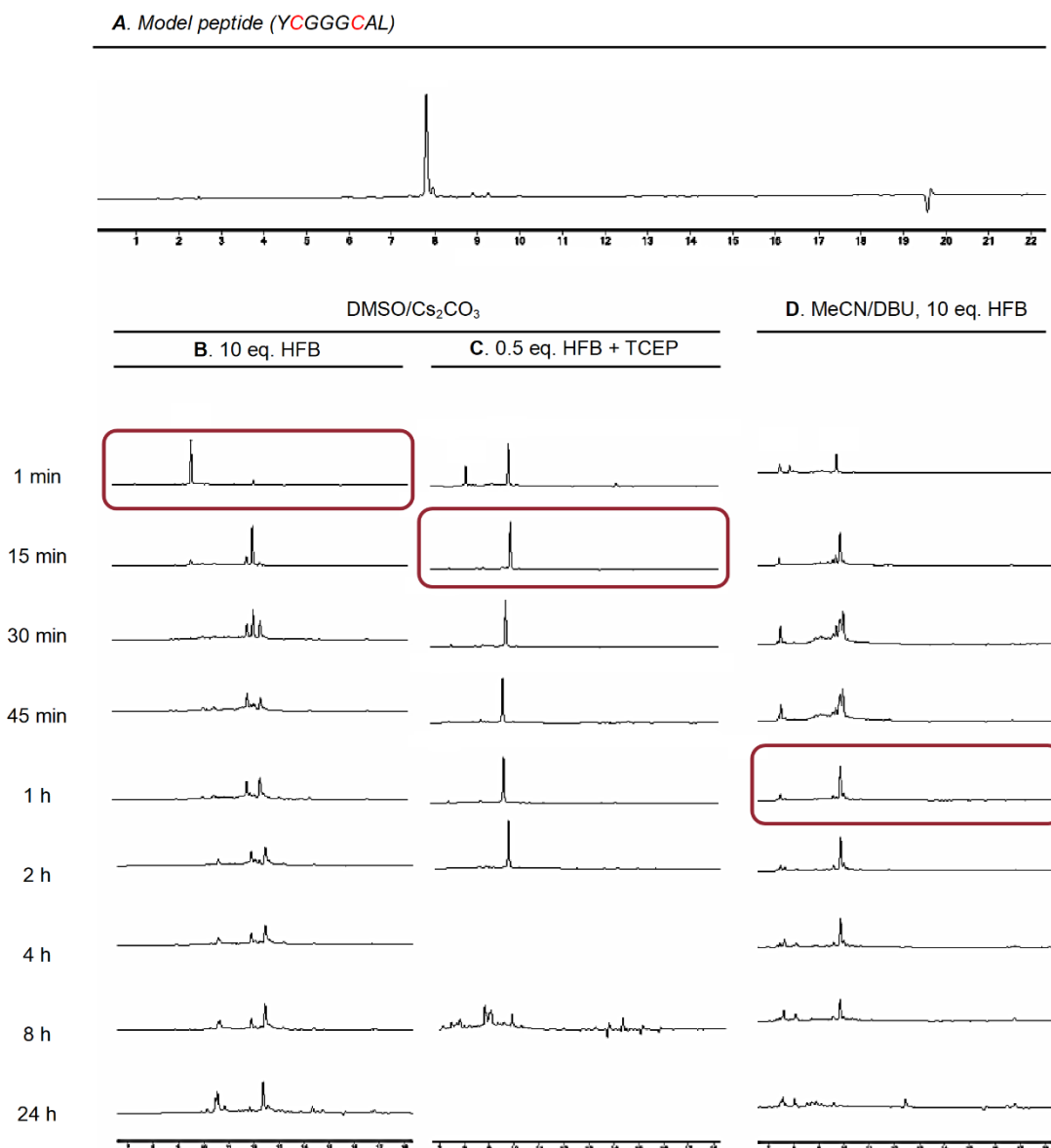
*Figure 28. A) HPLC chromatogram of model peptide **69**; Time courses of the reaction between peptide **69** and B) HFB (10 eq.) in DMSO/Cs₂CO₃; C) HFB (0.5 eq.) in DMSO/Cs₂CO₃ plus TCEP; D) HFB (10 eq.) in MeCN/DBU.*

When the reaction was performed in DMSO/Cs$_2$CO$_3$, complete conversion of the linear peptide into the HFB-templated cyclic form occurred almost instantaneously (<1 min, Figure

28). However, only traces of this cyclic product were detected after 15 min under the same conditions, which is presumably subject to β-elimination (as described in the next section) and degradation into unidentified products.

Due to the unexpectedly high speed of product formation, this transformation could be defined a click reaction.[97] If the same process was performed in MeCN/DBU, a slower but nevertheless complete conversion of the starting peptide into the cyclic product was also observed (within approximately 1 h, Figure 28).

The peptide stapling reactions are likely to happen faster than those of the model system with NAC due to a proximity effect related to the intramolecular nature of the second substitution step. Following longer reaction times under the DMSO/Cs$_2$CO$_3$ conditions, unidentifiable side products were observed to form, probably from oxidation or dimerisation reactions.

Other presumed degradation products might be originated from the E1cb base-mediated β-elimination reaction on substituted cysteines (Scheme 17).[164]
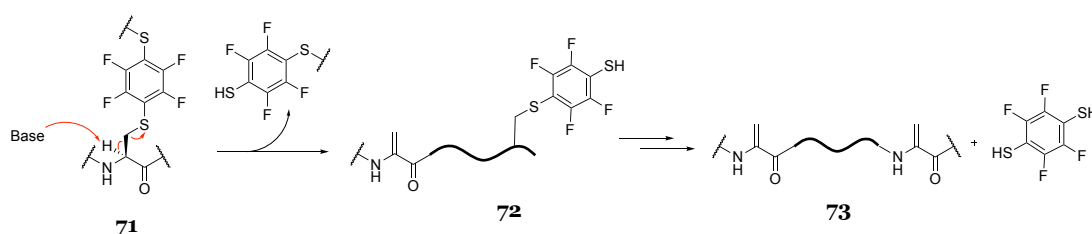
Scheme 17. Mechanism for the E1cb base-mediated β-elimination reaction.

This side reaction, probably occurring shortly after the formation of the cyclic product (between 1 and 15 minutes), is proposed to produce a dehydroalanine residue and reopen the cycle. The *m/z* ratios corresponding to monounsaturated conjugate or diunsaturated free peptide (**74** and **75** in Figure 29) were observed in small quantities, but they were not isolated nor fully characterised.
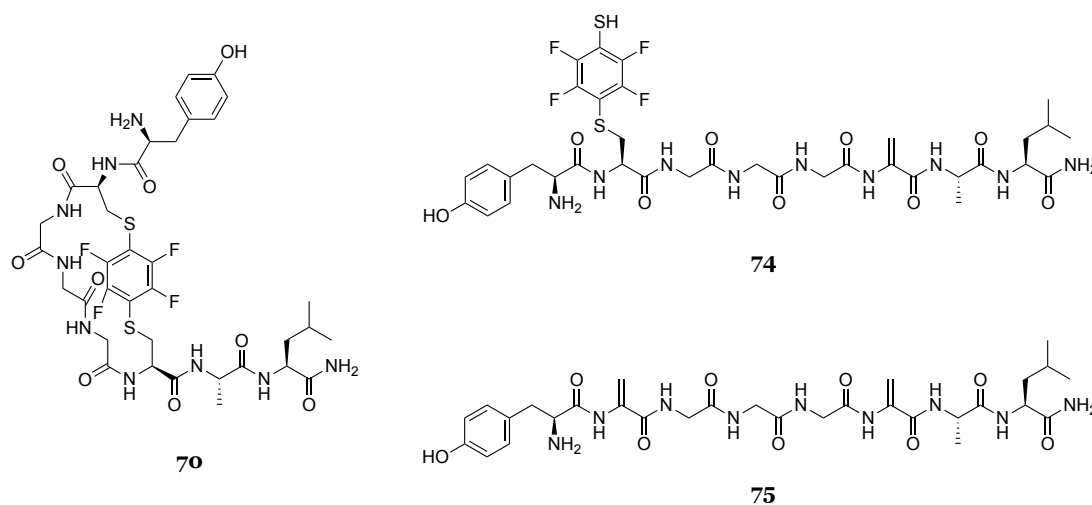
Figure 29. Chemical structure of HFB-cyclisation product (**70**), mono-unsaturated conjugate (**74**) or di-unsaturated free peptide (**75**) from peptide **69**.

In the case of the monosubstituted side product, the MW is the same as the cyclic product of interest, but the retention time is different. In both DMSO and MeCN, side products were not ascribed to any higher order substitution of HFB (*i.e.* tetrasubstitution leading to bicyclic species), and these were not observed even after prolonged reaction times. However, a larger amount of peptide could lead to the formation of such multicyclic systems.

Subsequently, the reaction in DMSO/Cs$_2$CO$_3$ was repeated with i) a lower amount of HFB (around 0.5 eq., Figure 28), to mitigate the high conversion rate and ii) the addition of tris(2-carboxyethyl)phosphine (TCEP), a non-thiol reducing agent that might limit the formation of side products. The combination of these factors afforded the cyclic product rapidly and cleanly as before but, in this case, the product remained stable for at least 2 h. Again, no tetrasubstitution products were observed.

As a general remark, the reaction worked well under both conditions, was completed in under 15 min even with a sub-stoichiometric amount of HFB, and TCEP played an important role in maintaining the stability of the product.

2.3.2 Disubstitution on-resin

Despite there being examples of single-component cyclisation carried out on solid-supported peptides, two-component reactions are generally performed in solution to minimise potential site isolation (which happens when the reactive moieties are not able to react with each other because they are attached to the support) and by-products formation on-resin.[165]

However, solid phase reactions offer undeniable advantages such as user-friendly handling, less laborious workup/easier purification, possibility of automation and further on-resin modification. This results in a potential greener, faster, cheaper, and possibly higher yielding process. Therefore, the possibility to perform the perfluoroaryl stapling of a di-cysteine containing peptide on-resin was also explored.

Unfortunately, the optimised cyclisation conditions were not suitable for polystyrene supports because the employed solvents MeCN and DMSO precluded swelling of the resin. Therefore, the efficiency of the standard solution-phase procedure (DMF/DIPEA) was investigated for the cyclisation of a di-cysteine containing peptide on-resin.[51,149,150]

After selective on-resin Mmt deprotection of **76** (Mmt-protected peptide **69**) cysteine residues, the cyclic product **70** was successfully obtained after overnight shaking of the resin with a reaction mixture made of HFB and DIPEA in DMF at room temperature (Scheme 18). Although DMF/DIPEA are not the conditions identified to be the best for fast cyclisation in solution, these worked well for on-resin procedure and are a valid option for this chemistry.
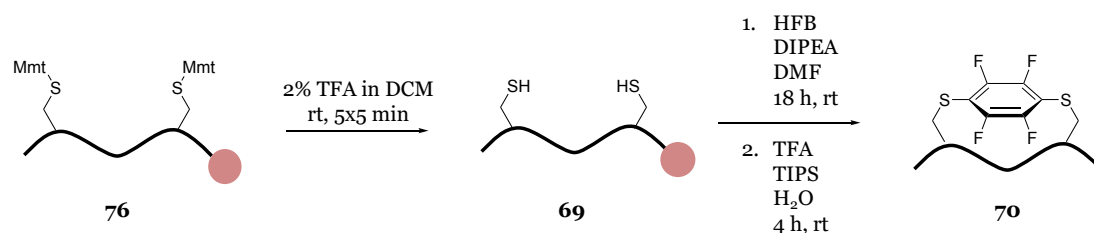
*Scheme 18. Reaction scheme and conditions for on-resin HFB-templated cyclisation of peptide **69**.*

2.3.3 Cyclisation of SPACE peptide

As this was the first example of HFB-mediated cyclisation on solid phase, this methodology was repeated with a longer and more complex peptide sequence to explore its applicability. The Skin Penetrating and Cell Entering (SPACE) peptide (**77**) is a peptide that can translocate across the stratum corneum into the epidermis and dermis layers of the skin, likely via the endocytic pathways, and penetrate skin cells.[166]

Acetylated SPACE peptide (**78**) was successfully cyclised with HFB following the same on-resin procedure described for the model peptide **69**. After selective deprotection of Mmt-protected acetylated SPACE peptide (**79**), the final cyclic product **80** was obtained using DMF/DIPEA reaction mixture (Scheme 19). This is evidence that the two-component HFB-peptide cyclisation technique may adapt well to changes in the amino acid sequence, length and inter-thiol spacing.
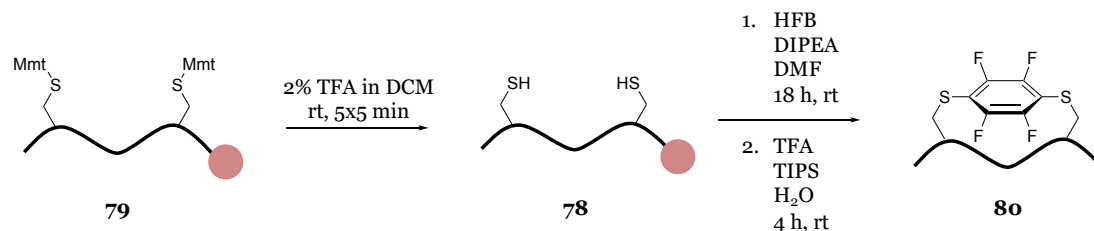
*Scheme 19. Reaction scheme and conditions for on-resin HFB-templated cyclisation of peptide **80**.*

To verify if the biological activity of this peptide was retained after replacing the disulfide bridge with a 1,4-dithiol substituted tetrafluorobenzene, the uptake of the fluorescently labelled SPACE-HFB cyclic conjugate was monitored (Figure 30).

Following the synthesis of Mmt-protected SPACE peptide, 6-aminohexanoic acid (Ahx) was incorporated at the N-terminus of the SPACE peptide, as a spacer to separate the peptide from the fluorophore. Alloc-protected lysine was inserted after the spacer and employed for on-resin conjugation of fluorescein isothiocyanate (FITC).

At the end of the sequence, coupling of another Ahx residue and acetyl capping were performed, to obtain AcAhxK(Alloc)AhxACTGSTQHQCG-NH$_2$ (peptide **81**). The final Ahx residue did not have a specific function in this experiment but was generally employed by Talhat Chaudhry (PhD student at LJMU)[167] to provide spacing for further bioconjugation. Peptide **81** was initially Mmt-deprotected (step 1) and cyclised with HFB (step 2) on-resin.
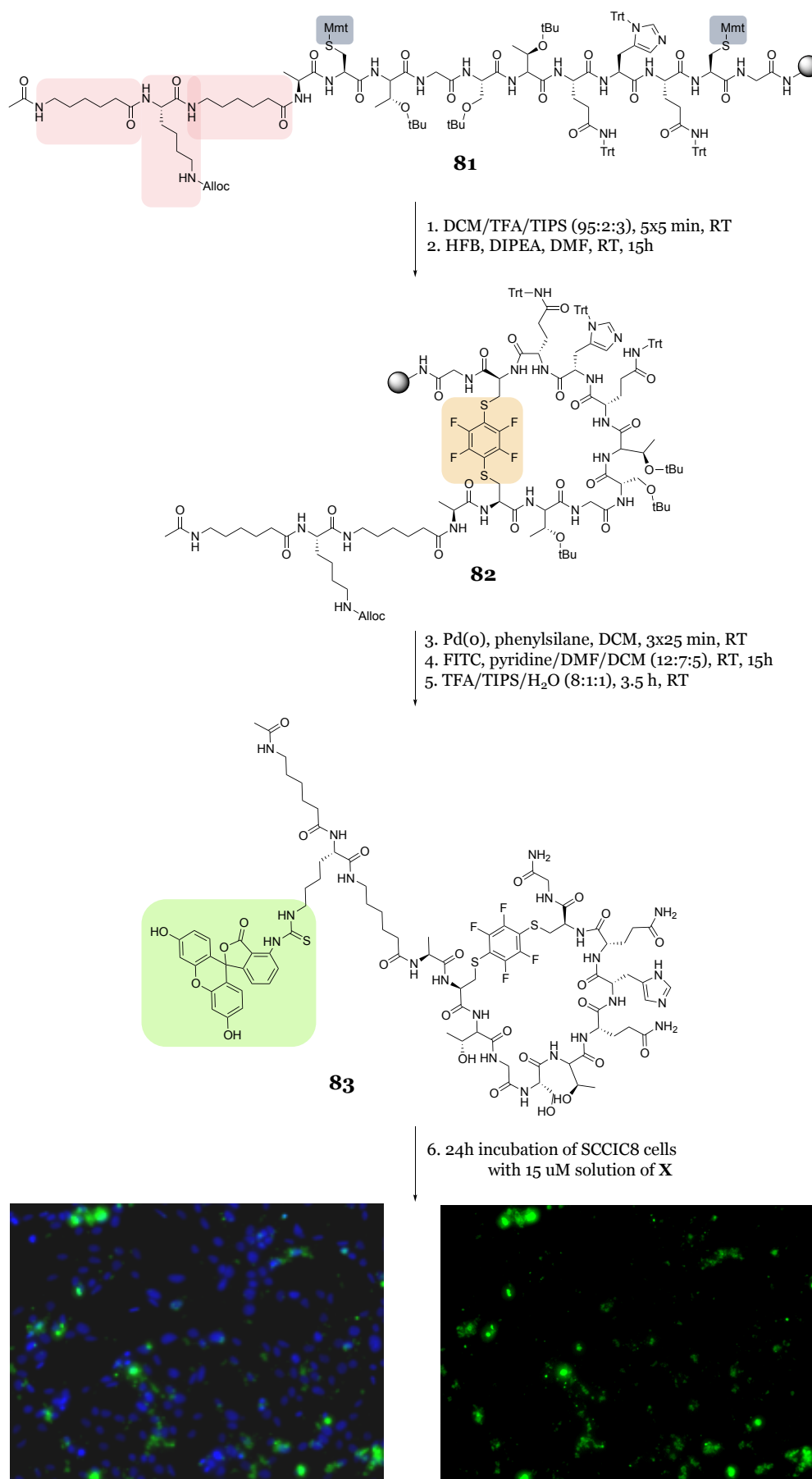
*Figure 30. Scheme for the synthesis and biological testing of conjugate **83**.*

After HFB cyclisation, conjugates **82** was Alloc-deprotected (step 3) and conjugated with FITC (step 4) on-resin. FITC labelling was executed by Talhat Chaudhry as detailed in the Material and Method section and gave the final product **83**.

Finally, conjugate **83** was cleaved (step 5) and employed for biological testing (step 6). After treatment, the uptake into SCCIC8 cells was monitored by fluorescent microscopy with a Leica microscope. These cellular uptake studies, also performed by Talhat Chaudhry, demonstrated that the HFB-stapled and fluorescently labelled SPACE peptide (**83**) maintained the original skin cell penetrating properties after modifications (Figure 30). As future experiment, the comparison of these results with those obtained from SPACE peptide labelled with a standard probe would benefit this work.

A rather complex macromolecule was obtained with a straightforward procedure made of fast couplings, easy washing steps, and isolation of a reasonably clean product as dry pellet after cleavage, highlighting the advantages of solid supported synthesis.

2.4 PEPTIDE MULTICYCLISATION

Cysteine residues can also be exploited for the multicyclisation of peptides, not only by multiple disulphide bridges - as for cyclotides - but also with two component strategies. Timmerman and co-workers applied their CLIPS technology (Scheme 20) to cysteine rich peptides (**84** and **87**) and synthesised bicyclic (**86**) and tricyclic (**89**) systems, using tri- (**85**) and tetrabromomethyl benzene (**88**) respectively. This is a two-component strategy that exploits multiple reactive sites on the same scaffold.
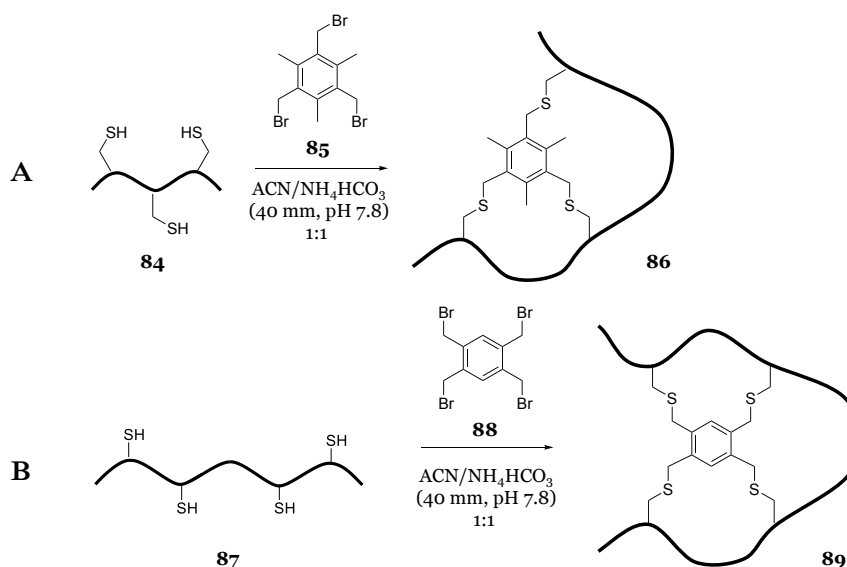
Scheme 20. General reaction scheme for peptide multicyclisation using CLIPS technology.

Slightly different approaches can also be employed, as showed by another work from the Pentelute Group.[168] They prepared three-components bicyclic systems by double decafluorobiphenyl macrocyclisation of two different cysteine pairs (Scheme 21A) as well as five-components bicyclic systems via decafluorobiphenyl primary substitution with three separate cysteines and secondary substitution with 1,3,5-benzenetrithiol (Scheme 21B).
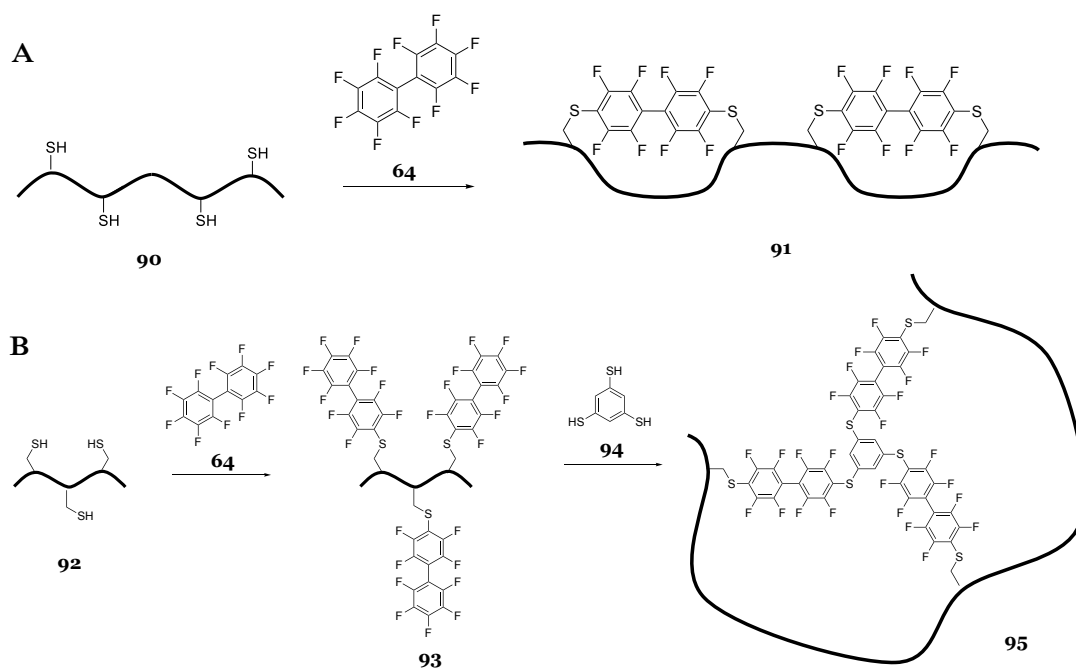
Scheme 21. General reaction scheme for peptide multicyclisation using polyfluorinated compounds.

The polyfluoroaromatic reagent 4F-2CN (**56**) was also used by Liu et al for multicyclisation purposes (Scheme 22). In the reaction named *precisely regulated and efficient locking of linear peptide* (PROP-locking),[139] 4 thiols (2 cysteine and 2 penicillamine) of the same peptide (**96**) react with the four fluorine atoms on the ring of **56** to give a regiocontrolled multicyclic peptide system (**97**).
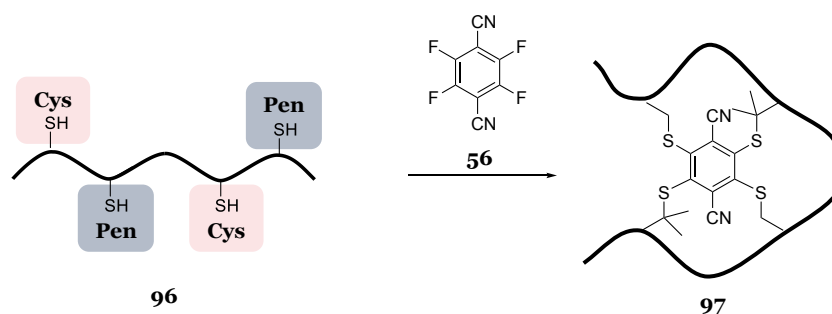
Scheme 22. General reaction scheme for peptide multicyclisation using 4F-2CN.

After the cyclisation of a di-cysteine containing peptide on HFB, four positions remain conceptually available for substitution, making HFB a potential scaffold for stepwise multicyclisation or post-stapling introduction of additional functionalities. Table 1 (at page 39), as well as our optimisation studies, indicated that higher order substitutions are possible using non-amino acid thiols. However, this chemistry typically requires conditions that are not always compatible with peptides, such as higher temperature, longer reaction times and/or stronger inorganic bases due to the reduced reactivity of the fluorine atoms following the first two substitutions. This may be one of the reasons why, despite the established applications of HFB for peptide stapling, there are no examples of HFB-templated peptide multicyclisation.

2.4.1 One-step tetrasubstitution

After the successful cyclisation reactions of di-cysteine-containing peptides described above, the implementation of the same conditions for the multicyclisation of a tetra-cysteine containing peptide was the next objective. The sequence WG**C**K**C**GGKGG**C**GKGGG**C**GW (**98**) was selected from the publication of Liu and co-workers: this group previously cyclised this model peptide around 4F-2CN (**56**), providing the only example of a peptide multicyclisation strategy on a single polyfluorinated ring to date.[139]

Unfortunately, all the previously developed conditions – on-resin reaction in DMF/DIPEA, as well as reaction in DMSO/Cs_2CO_3 or MeCN/DBU solutions (Table 7) – appeared to fail to afford the tetrasubstituted (multicyclic) peptide. Analysis of the crude products were of difficult interpretation and were poorly informative due to HPLC and LC-MS chromatograms showing a high number of unresolved peaks and undefined by-products (Figure 31).
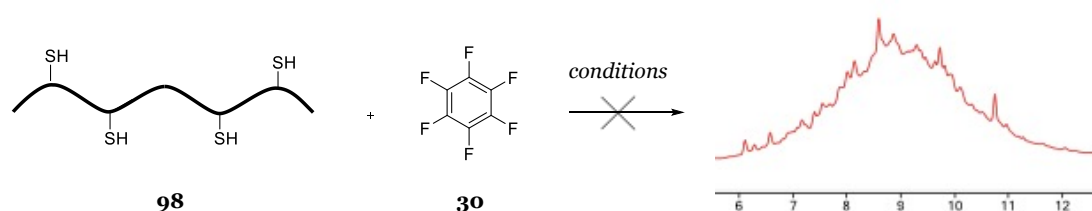
Figure 31. General reaction scheme for one-step HFB-templated peptide multicyclisation and HPLC chromatogram of unsuccessful reaction.

Table 7. Reactions conditions for one-step HFB-templated peptide multicyclisation explored in this work.

Conditions	Peptide Sequence	N-term
Cs_2CO_3, DMSO, rt, 1 min	WGCKCGGKGGCGKGGGCGW (**98**)	H (Free peptide)
DBU, MeCN, rt, 1 h	WGCKCGGKGGCGKGGGCGW (**98**)	H (Free peptide)
DIPEA, DMF, rt, 18 h	WGCKCGGKGGCGKGGGCGW (**98**)	Resin

The mixture of different chemical species was originally ascribed to the low purity of the starting crude peptide following on-resin SPPS and cleavage. However, all the efforts in the optimisation of **98** synthesis to obtain a crude peptide with higher purity provided no tangible benefit. Purification was also difficult as the impurities were not appropriately resolved by HPLC and the isolated product yield not compatible with further experiments. It was also difficult to determine if side products originated during coupling steps or cleavage procedure. Regardless the use of different thiol scavengers in the cleavage cocktail, such as dithiothreitol or thioanisole, unidentifiable *m/z* ratios probably related to cysteine oxidation were the main impurities observed. Therefore, a crude peptide of acceptable purity (42%) for chemical modification was employed to continue the work. Differently from our standard microwave-

assisted automated SPPS method, the same peptide was synthesised by Wu and co-workers using flow SPPS, making the experimental conditions not directly comparable.

The impure reaction product may also be ascribed to the high number of cysteines in the sequence that increase the oxidation sensitivity of the peptide when in strong basic conditions or in DMSO. This was not confirmed because it was not possible to identify any side product, but it was noted that the reaction conditions employed by Wu and co-workers were milder and water-based, which cannot be applied to HFB due to its poor solubility compared to 4F-2CN.

Additionally, the simultaneous reaction of four unprotected thiolates with six available positions on the HFB ring can potentially lead to the formation of 12 multicyclic regioisomers (Figure 32), and therefore a mixture of species in the final product. 4F-2CN has only four fluorine atoms on the ring and therefore four possible reaction sites. As explained by the authors, there are 6 possible multicyclic regioisomers (Figure 32) and consequently half of the products that reaction on HFB can give.
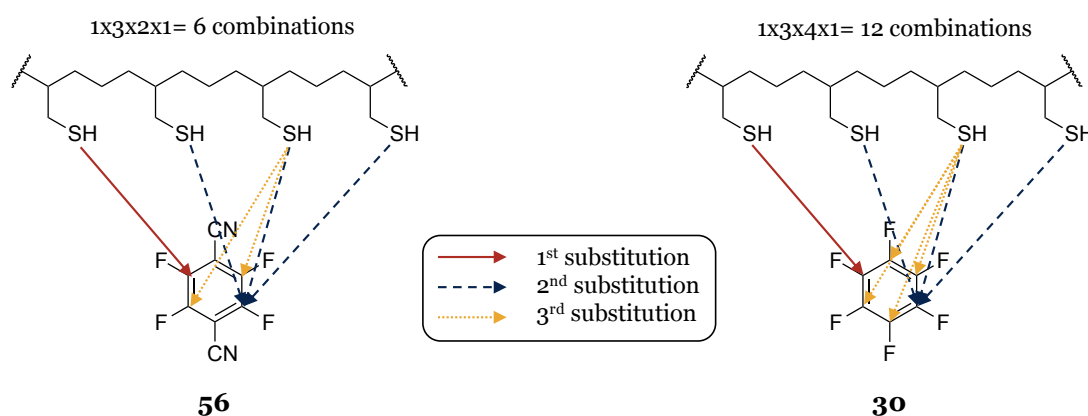
Figure 32. Graphical representation of the possible regioisomers combinations when a tetra-cysteine containing peptide reacts with 4F-2CN or HFB.

Liu *et al* tackled the regioisomer problem by introducing a combination of cysteine and penicillamine in the peptide sequence. These are two AAs bearing thiols with different nucleophilicity, which react in one pot with 4F-2CN and only give 2 regioisomers.

2.4.2 Two-step tetrasubstitution

An alternative strategy against multiple regioisomers formation could be the implementation of two orthogonal protecting groups for cysteines, to selectively deprotect a pair of thiols at a time.

Initially, two thiols would be deprotected on-resin and cyclised with HFB under DMF/DIPEA conditions. This step would mirror what was previously described for di-cysteine containing peptide and give the formation of one intermediate product only. Performing the first step on-resin would also make the intermediate isolation easier and the protected peptide on solid support more stable.

Subsequently, the cyclized peptide could either be deprotected on-resin and the final multicyclisation step performed on solid support or cleaved from the polymer to complete the substitution in solution (Figure 33).
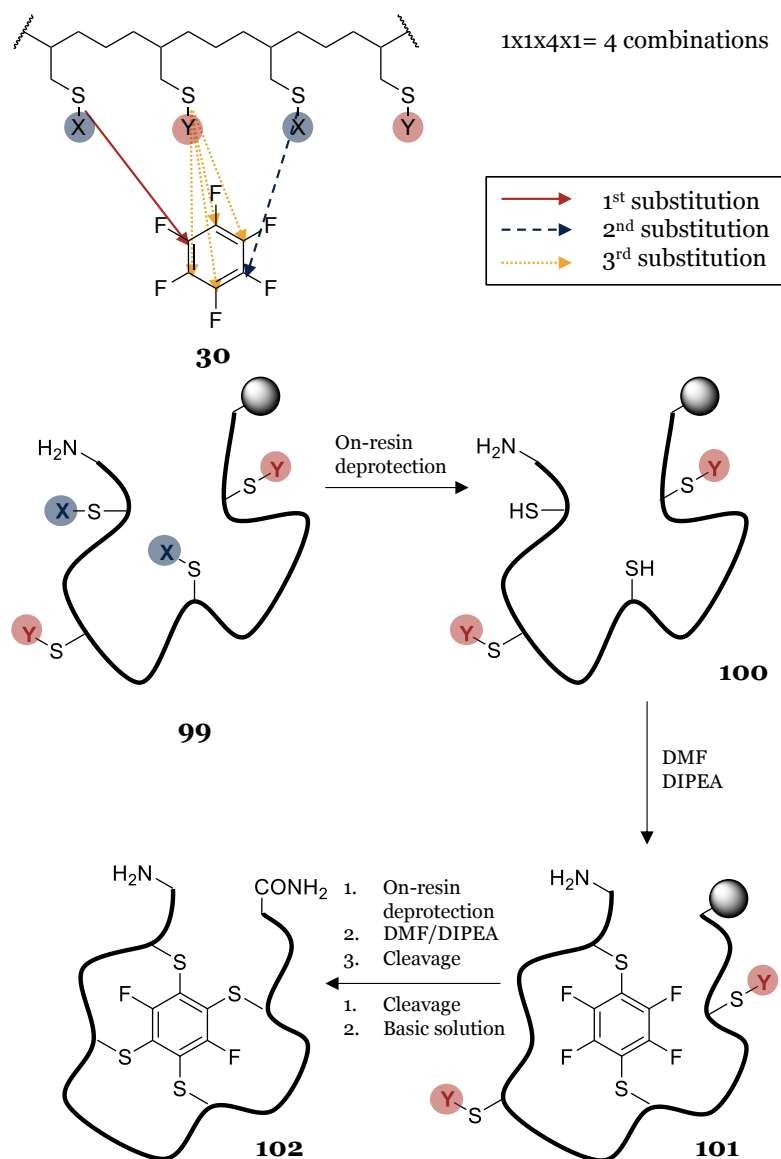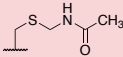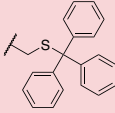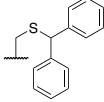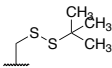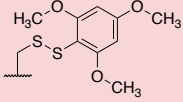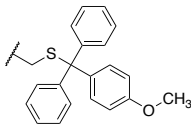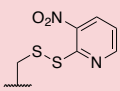
Figure 33. A) Graphical representation of the possible regioisomers combinations when a peptide containing 4 cysteine residues protected with two orthogonal protecting groups. B) planned reaction scheme for the synthesis of HFB-templated multicyclic peptide.

There are several commercially available options for side-chain protection of cysteines in Fmoc-SPPS but picking two orthogonal PGs suitable for the planned synthetic route was not straightforward. Within the same sequence, a pair of cysteines should be suitable for mild, safe, and selective on-resin deprotection without the simultaneous formation of disulfide bridges, while the other pair should be easily deprotected at a later stage, on-resin or during the cleavage step.

Table 8 lists commercially available PGs for Fmoc-cysteine-SH at the time this work was developed.[169]

Table 8. Summary of commercially available cysteine with different side-chain protection.[169]

Protecting group (Abbreviation)	Chemical structure	Cleavage reagents
Acetamidomethyl (Acm) Stable to TFA. I_2 or Tl^{3+} → removal + disulfide bond.		Hg^{2+}, Ag^+, I_2, Tl^{3+}, CH_3SiCl_3/PhSOPh, RSCl, Pd(Cl)$_2$
Tert-butyl (tBu) Stable to TFA and I_2. CH_3SiCl_3/PhSOPh → removal + disulfide bond. DPDS in acid → Cys(Pys) with disulfides and Cys(Acm).		Hg(II), HF at 20°C, TFA/DMSO, CH_3SiCl_3/PhSOPh
Triphenylmethyl (Trt) TFA cleavage → sulfhydryl peptide Routine use in Fmoc SPPS.		Hg^{2+}, Ag^+, I_2, Tl^{3+}, TFA
1,2-diphenylmaleyl (Dpm) Stable during Mmt removal with 2% TFA. TFA cleavage → sulfhydryl peptide Routine use in Fmoc SPPS.		Hg^{2+}, Ag^+, I_2, Tl^{3+}, TFA
2-tetrahydropyranyl (Thp) TFA cleavage → sulfhydryl peptide. Routine use in Fmoc SPPS. Alternative to Trt, less racemisation and β-elimination.		Hg^{2+}, Ag^+, I_2, Tl^{3+}, TFA
Tert-butylsulfenyl (StBu) Stable to TFA, without thiol scavengers. In combination with Acm selectively forms disulfides. Removal is slow → STmp for on-resin deprotection.		RSH, R_3P
Trimethoxy phenylthiol (STmp) Moderately stable to TFA without thiol scavengers. With Mmt, selectively forms of two disulfides on-resin. Better cleavage yield than StBu, but more expensive.		RSH, R_3P
(4-methoxyphenyl)diphenylmethyl (Mmt) Selectively removed in 2% TFA → peptide on solid phase On-resin disulfide formation or Cys thiol modification.		2% TFA in DCM, Hg^{2+}, Ag^+, I_2, Tl^{3+}
3-nitro-2-pyridinesulfenyl (Npys) Stable to TFA without thiol scavengers. Activates thiols for disulfide bond formation and selective preparation of mixed disulfides. Expensive.		RSH, R_3P

The combination of trityl (triphenylmethyl, Trt) and methoxytrityl ((4-methoxyphenyl) diphenylmethyl, Mmt) PG was initially chosen and peptide **103** (H-WGC(Mmt)KCGGKGGC(Mmt)GKGGGCGW-NH$_2$) was synthesised. Mmt is more acid-sensitive than Trt because the -OMe group stabilises the trityl cation formed after the deprotection. Hence, Mmt can be removed with very low acid concentration (around 2-3%) in a short time. However, Trt was not stable in such conditions and the selective removal of S-Mmt over S-Trt was not achieved. These PGs are incompatible and with no practical use as orthogonal protection because the acid-concentration stability window is too narrow.

Trt was subsequently combined with tert-butylsulfenyl (StBu) and peptide **104** (H-WGC(StBu)KCGGKGGC(StBu)GKGGGCGW-NH$_2$) was obtained. It is reported that the removal of StBu is highly sequence dependent and could be slow and difficult to perform completely. For these reasons, this is generally performed in solution after peptide cleavage, but deprotection on-resin was required for this strategy.[170]

The StBu deprotection mechanism is a reduction of the cysteine thiol achievable with reducing agents, usually tertiary phosphines (*e.g.* Ph$_3$P[171] or TCEP[145]), or by thiol-exchange with the simultaneous oxidation of another thiol. The second method was chosen because it is less hazardous and more suitable for on-resin procedure.

To optimise this difficult PG removal, the different thiols from Figure 34 – β-mercaptoethanol (**105**),[172] dithiothreitol (**106**),[25] acetamidothiophenol (**107**),[26] ethanedithiol (**108**),[27] thioglycolic acid (**109**),[173] gluthatione (**57**),[41] – were screened for deprotecting peptide **110**.
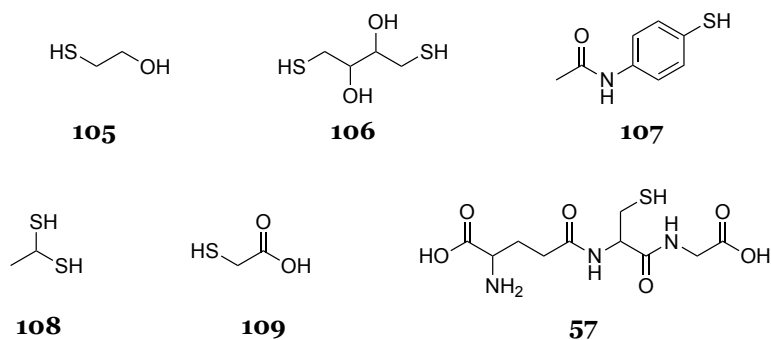
*Figure 34. Chemical structure of β-mercaptoethanol (**105**), dithiothreitol (**106**), acetamidothiophenol (**107**), ethanedithiol (**108**), thioglycolic acid (**109**), and gluthatione (**57**).*

The pool of results was disappointing, as no condition afforded complete StBu deprotection with a single treatment in a reasonable time. β-mercaptoethanol gave the best balance between high deprotection yield and clean crude peptide. Contextually, DIPEA was preferred to help the on-resin StBu deprotection of WGC(StBu)KCGGKGGC(StBu)GKGGGCGW-NH$_2$ (**110**), while N-methyl morpholine is generally used for this purpose.[174] Finally, the partially deprotected **111** was obtained by double treatment with excess of β-mercaptoethanol with 10% DIPEA in DMF for 48 h.

The first substitution step with HFB was then performed and only the monosubstituted product (**112**) was observed within a mixture of unidentified products (Figure 35).

Unfortunately, the second part of the process, carried out in solution after cleavage, led again to uninterpretable LC-MS results and the experiment could not be considered successful.
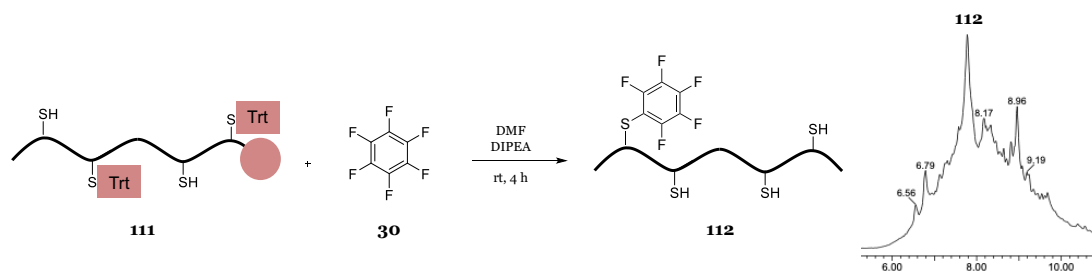
Figure 35. Reaction scheme and HPLC chromatogram of intermediate product 112.

To evaluate the possibility that the encountered problems were related to the first partial thiol deprotection, STmp was identified as a valuable alternative to StBu. STmp is more sensitive to thiol exchange and on-resin deprotection is more effective compared to StBu.

However, after the synthesis of peptide **113**, this strategy was quickly abandoned because the results were not significantly different from StBu and the Fmoc-Cys-STmp building block is more expensive than Fmoc-Cys-StBu.[170] Finally, the Trt-Acm pair was also unsuccessfully tested, because Acm on-resin deprotection of peptide **114** with $PdCl_2$ in DMF did not work.[175]

Overall, HFB-templated peptide multicyclisation products were not obtained by the two-step strategy, under any of the conditions tested and despite evaluating several combinations of thiol protecting groups, which are summarised in Table 9.

In addition to the above considerations, difficulties in achieving tetrasubstitution could also be related to a lower reactivity of the four fluorine atoms after the first substitution, most likely due to steric hindrance and structural constraints.

Table 9. Summary of the cysteine side-chain protection pairs explored in this work.

First	Second	Comments
Mmt	**Trt**	Not orthogonal. Complete Mmt removal with low amount of TFA partially remove Trt as well.
Trt	**StBu/STmp**	StBu removal is difficult, but the protecting group orthogonality is valid and the on-resin deprotection conditions are compatible and selective.
Trt	**STmp**	STmp is a better but more expensive alternative to StBu. Its deprotection is easier but the overall procedure was not improved.
Trt	**Acm**	Acm deprotection on-resin did not work.
Mmt	**Dpm**	See orthogonal deprotection conditions from Table 8. To be tested in future work.

2.4.3 Future work on hexafluorobenzene

There is probably one more combination of thiol PGs to be tested if pursuing HFB-templated peptide multicyclisation following this approach: a pair of thiols could be orthogonally protected with Mmt, which allows initial on-resin deprotection at low acid concentration, and

an acid stable (not Trt nor STmp) thiol PG that can be removed during or after cleavage could be used for the second cysteine pair. This strategy would suggest combinations such as Mmt/Dpm or Mmt/StBu.

Alternatively, the strategy from Liu et al. employing thiols with different nucleophilicity could be experimented on HFB.[139] However, due to the chemical differences between HFB and 4F-CN previously highlighted, its success is not straightforward.

It would also be informative to test the reactivity of the four remaining fluorine atoms after the first cyclisation step, using thiols that are not part of the same peptide sequence. This could prove the theoretical feasibility of the multicyclisation procedure and indicate a potential technique for post-stapling introduction of additional functionalities.

A completely different peptide sequence of easier synthetic accessibility or purification could be of expedient implementation.

As described earlier in this chapter, the results from the DOE optimisation can be used to control thiol-fluoride di- and tetrasubstitution. After synthesis and isolation of the substitution products between HFB (**30**) and differently substituted cysteine (**115**), a potential scaffold for branched peptide or multicyclic systems could be obtained (Scheme 23A). This scaffold could be used, for example, for the synthesis of dendrimer architecture with fluorobenzene core (**117**).

Finally, this thiol-fluoride S_NAr chemistry could also be employed in reverse: polyfluorinated amino acids, which are coupled during SPPS and part of a fluorinated peptide (**118**), may serve as anchor points for bioconjugation of thiol (**119**) as in Scheme 23B.
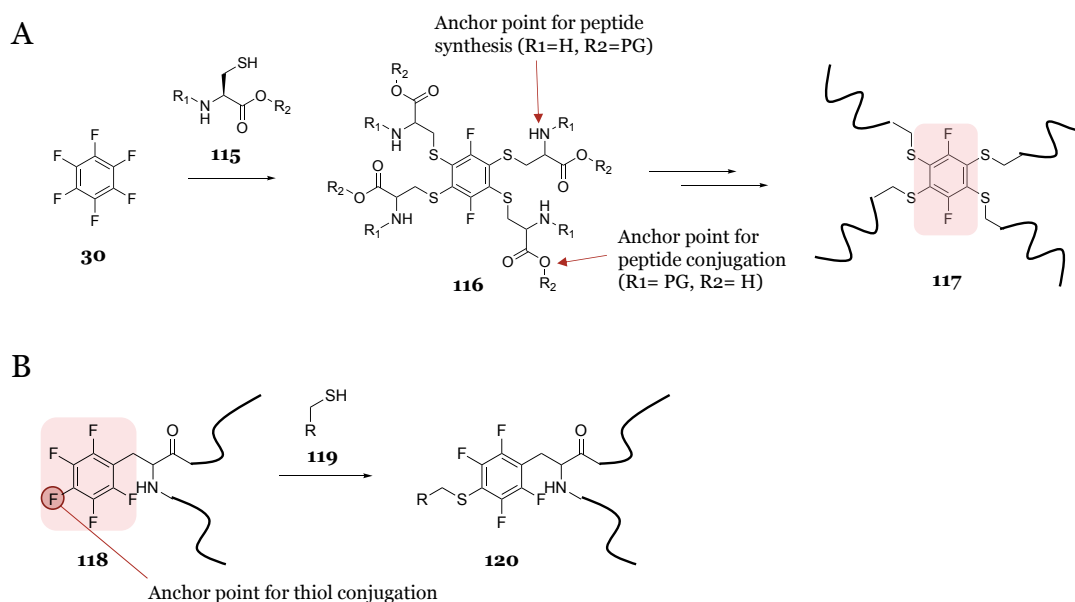
Scheme 23. General reaction scheme for A) use of HFB as scaffold for branched peptide; B) use of polyfluorinated amino acids as anchor points for thiol bioconjugation. R1 and R2 are different protection/activation for cysteine. R is a generic carbon substituent.

2.5 PORPHYRIN-PEPTIDE BIOCONJUGATION

2.5.1 Polyfluorinated porphyrins

The properties of porphyrins and the advantages of their bioconjugation with peptides have been described earlier. The starting point for the development of novel peptide-porphyrin conjugates is the optimisation of the bioconjugation reaction and the fast, selective, and peptide-compatible thiol-fluoride S_NAr could be suitable for this scope. The main requirements for this application are a peptide sequence containing one or more cysteine residues and the tetrapyrrole *meso* positions substituted with polyfluorinated ring(s).

5,10,15,20-tetrakis(pentafluorophenyl)porphyrin (**121** in Figure 36) is a tetrasubstituted symmetric porphyrin and is reported in literature as a versatile scaffold for the conjugation of nucleophiles.[176] The *meso*-pentafluorophenyl substituent is analogous to the extensively explored structure and chemistry of HFB, where the *para*-fluoride is the reactive site. Porphyrin **121** exhibits three typical ^{19}F NMR signals with a 2:1:2 integration ratio and a chemical shift of -138, -153 and -162 ppm (Figure 36), corresponding to the 8 *ortho*, 4 *para* and 8 *meta* fluorine atoms, respectively.
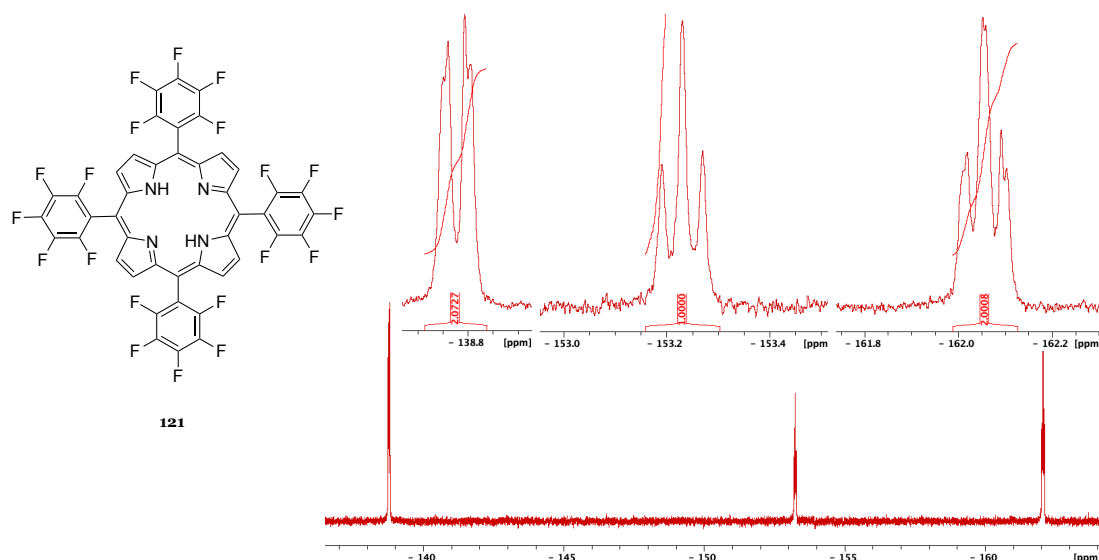
*Figure 36. Chemical structure of 5,10,15,20-tetrakis(pentafluorophenyl)porphyrin (**121**) and its typical ^{19}F NMR spectrum.*

Examples of **121** conjugation with cyclodextrins, sugars, nucleotides, porphyrins, and other reactive moieties *via* amine/hydroxy/thiol S_NAr were reviewed by Costa and co-workers.[177] However, cysteine and cysteine peptides have been rarely coupled following this strategy so far.[178]

2.5.2 S_NAr on polyfluorinated porphyrins

An initial study of the reactivity between **121** and simple thiols was performed following the approach previously employed for HFB. DMF/DIPEA, a common solvent-base combination employed in the literature for peptide-porphyrin conjugation,[179] was initially compared with the combination of DMSO and Cs_2CO_3 following the interesting observations from earlier work

(Scheme 24 and Table 10). Unfortunately, the combination of MeCN and DBU could not be included in this study due to the low solubility of **121** in MeCN.
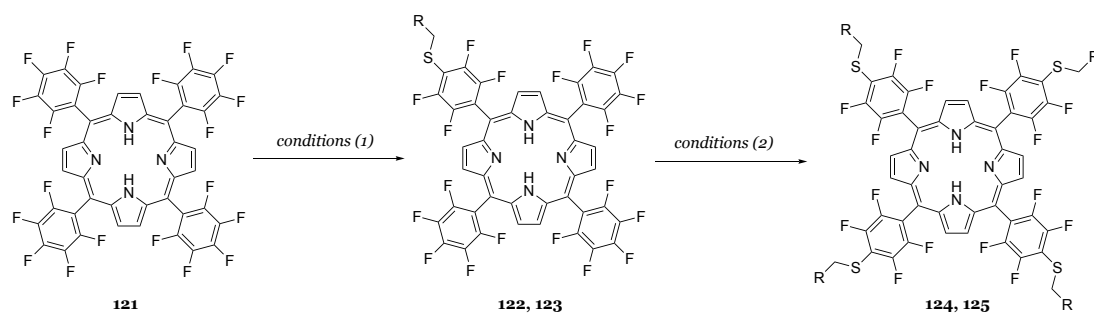
*Scheme 24. General reaction scheme for thiol substitution on **121**.*

*Table 10. Reactions conditions for thiol substitution on **121** explored in this work. Conversion rates obtained by NMR analysis are displayed in Figure 38.*

Conditions	Thiol (R)	Product
1. NAC (1 eq), Cs$_2$CO$_3$, DMSO, rt 2. + 4 eq NAC	NAC (**31**)	**122**
1. NAC (1 eq), DIPEA, DMF, rt 2. + 4 eq NAC	NAC (**31**)	**124**
1. GSH (1 eq), Cs$_2$CO$_3$, DMSO, rt 2. + 4 eq GSH	GSH (**57**)	**123**
1. GSH (1 eq), DIPEA, DMF, rt 2. + 4 eq GSH	GSH (**57**)	**125**

Substitution on **121** was only observed to take place in the *para* position and, as the reaction proceeds with the formation of single or multiple substituted products, two signals corresponding to *ortho* and *meta* fluorine atoms of substituted *meso*-tetrafluorophenyl groups appear around -134 and -139 ppm. The fully tetrasubstituted product contains four pairs of equivalent *ortho* fluorine nuclei and four pairs of equivalent *meta* fluorine nuclei (Figure 37). This allows easy reaction monitoring and products quantification by ¹⁹F NMR as for HFB.

Despite HFB and relative di- and tetrasubstitution products have very distinct chemical shift, **121** mono-, di-, and trisubstitution products are difficult to differentiate. Their spectra are made from the combination of signals from *meso*-tetrafluorophenyl groups and *meso*-pentafluorophenyl groups, which have very similar chemical shift regardless the substitution of the remaining rings. Therefore, the main information from ¹⁹F NMR relates to the percentage of reacted porphyrin and not the level of substitution. To this end, the decreasing signal at -162 ppm and the increasing signal at -132 ppm can be set as reference of unreacted and reacted pentafluorophenyl groups, respectively. To determine the level of substitution, these experiments were combined with HPLC and LC-MS analysis at a later stage.

Fully tetrasubstituted **121**

Partially substituted **121**

121
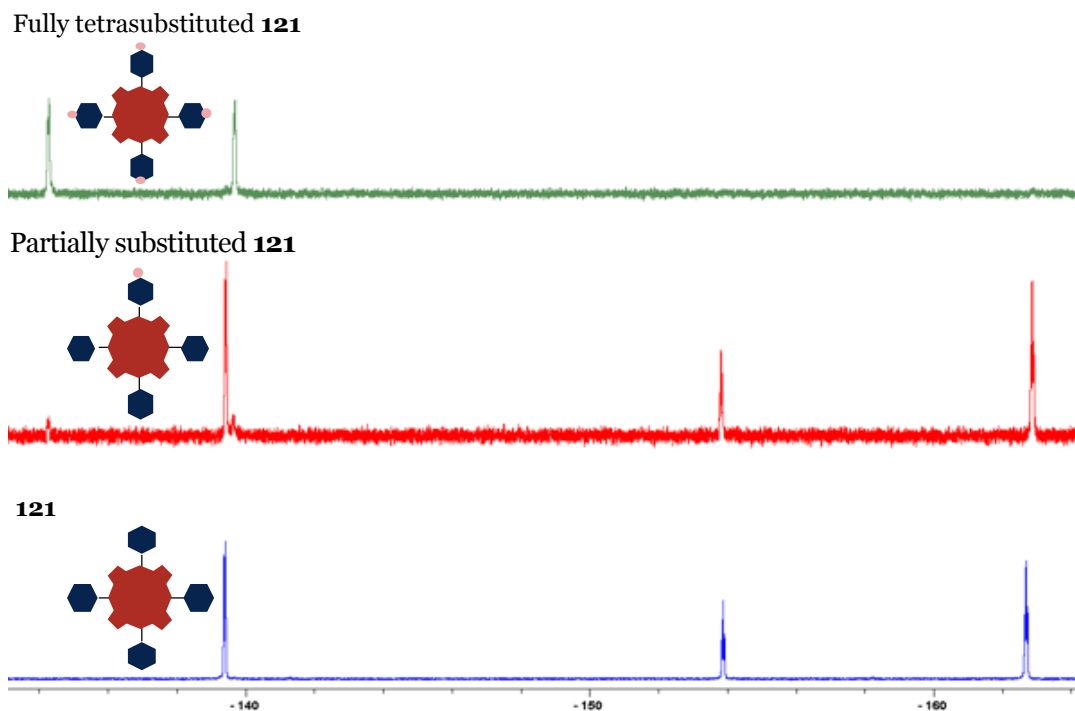
*Figure 37. NMR spectra of **121**, partially substituted **121**, and fully tetrasubstituted **121** with graphical representation of corresponding products obtained with conditions from Table 10.*

In the reaction between **121** and NAC (1:1), the solvent-base combination of DMSO and Cs_2CO_3 led to the formation of the NAC-monosubstituted porphyrin (**122**) within 15 min and then increased over time (red in Figure 38). **122** was the main product in a mixture of multiply substituted minor products not distinguishable by ^{19}F NMR. Two days after the addition of 4 more equivalents of thiol, 100% of tetrasubstituted porphyrin **124** was obtained.

The combination DMF/DIPEA showed a similar trend but with a slightly slower reaction onset: typically, monosubstitution appeared only after 1 h and then increased with a similar rate (yellow in Figure 38). However, fully tetrasubstituted product were more difficult to achieve, requiring one week after the addition of NAC in excess (timepoint not included in Figure 38).

Hence, DMSO/Cs_2CO_3 favoured a slightly faster reaction onset and significantly promoted a faster tetrasubstitution over DMF/DIPEA.

Reactions with glutathione (GSH) were faster than NAC but trends were similar, and DMF/DIPEA remained slightly slower than DMSO/Cs_2CO_3. GSH monosubstitution on **121** (conjugate **123**) appeared after 15 min in DMSO/Cs_2CO_3 (grey in Figure 38) and after 30 min in DIPEA/DMF (blue in Figure 38).

Interestingly, after further addition of 4 equivalents of GSH, complete tetrasubstitution of **121** (**125**) was achieved in two days in DMSO/Cs_2CO_3 as observed for NAC, and much faster than using DMF/DIPEA that requires one week (timepoint not included in Figure 38).
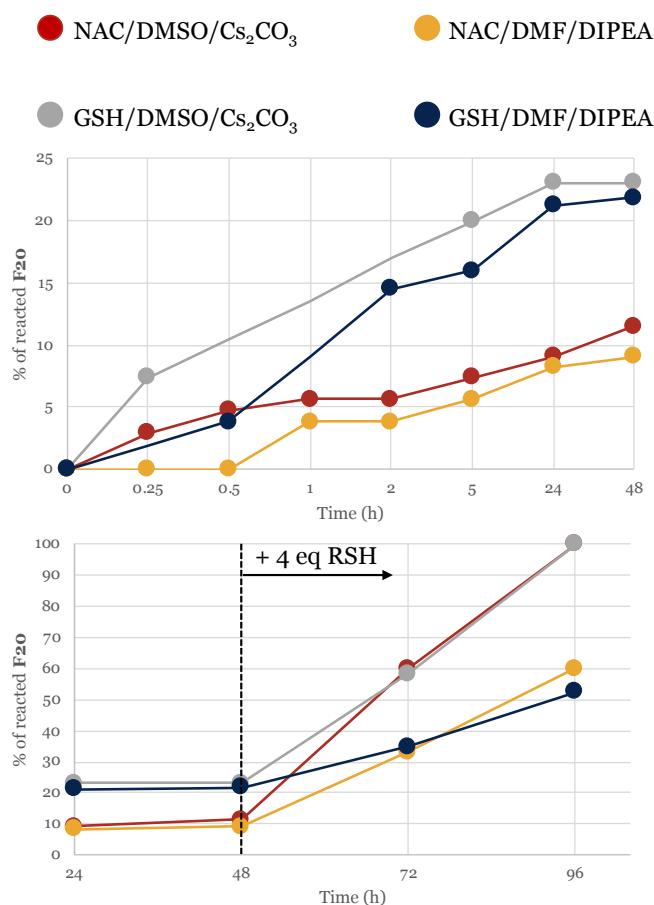
*Figure 38. Kinetic plots for reaction between **121** and NAC or GSH in* DMSO/Cs₂CO₃ *or DMF/DIPEA.*

Once again, DMSO/Cs₂Co₃ proved to be particularly effective for this chemistry and the better solubility of **121** in DMSO may have additionally favoured this condition over DMF. This initial study confirmed the reactivity of **121** towards different thiols and allowed this work to proceed with the conjugation of more complex peptide systems.

2.5.3 Preliminary studies on peptide-porphyrin conjugation

The ability to control regioselectivity of thiol-fluoride substitution is a topic of concern for bioconjugation. In this case, regioselectivity is reflected in coupling the porphyrin on the desired residue of the peptide sequence but also preventing multiple conjugations when the peptide or the tetrapyrrole ring display more than one group with comparable reactivity.

For this reason, most of the reported examples of peptide-porphyrin bioconjugates take advantages of asymmetrically functionalised porphyrins. This favours the regioselectivity of the peptide conjugation but requires synthesis of A3B type porphyrins, obtained with more difficult and lower yielding procedures if compared to symmetric analogues. A4 porphyrins can sometimes be regioselectivity modified after their synthesis and before the conjugation, to have one reaction site only. Coupling with non-natural AA, previously included in the peptide sequence, is another approach that can grant regiocontrol.[179]

In this project, **121** is a symmetric porphyrin with four potential conjugation sites for nucleophiles via S_NAr. Initially, conjugation was pursued with CAAARGD (**126**) as a model peptide presenting a cysteine residue at the N-terminus. This peptide sequence was chosen for the cell permeable activity of the RGD motif[180], for its overall good solubility and for being a short and accessible sequence. Completely deprotected **126** was successfully coupled to **121** in solution, employing both DMSO/Cs_2CO_3 and DIPEA/DMF conditions (Scheme 25 and Table 11). As previously observed with NAC, reactions in solution worked well but have the drawback of giving multiple higher-order substitution products, alongside monoconjugation (**128**).
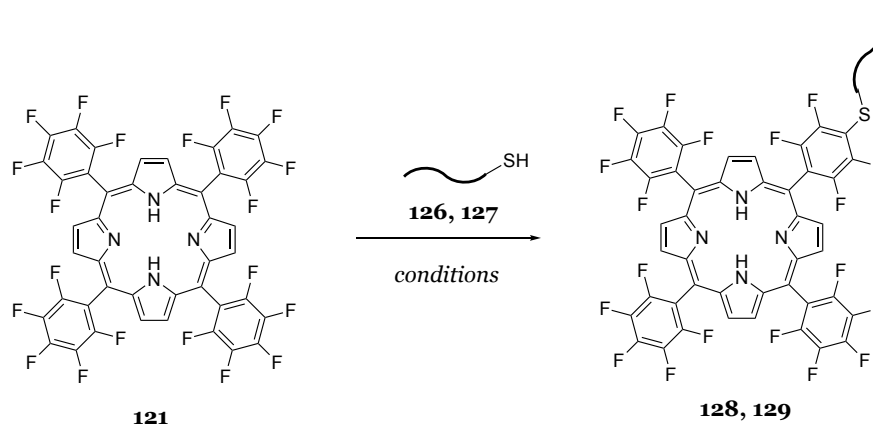
*Scheme 25. General reaction scheme for peptide substitution on **121**.*

*Table 11. Reactions conditions, peptide sequences and products from studies of peptide substitution on **121** explored in this work.*

Conditions	Peptide Sequence	N-term	Product
Cs_2CO_3, DMSO, rt	CAAARGD (**126**)	H (Free peptide)	**128**
DIPEA, DMF, rt	CAAARGD (**126**)	H (Free peptide)	**128**
DIPEA, DMF, rt	CAAARGD (**126**)	Resin	**128**
DCM, DBU	CAAARGD (**126**)	Resin	**128**
DIPEA, DMF	CGGKLVFF (**127**)	Resin	**129**
DIPEA, NMP	CGGKLVFF (**127**)	Resin	**129**
DIPEA, DCM	CGGKLVFF (**127**)	Resin	Not analysed
DIPEA, DCE	CGGKLVFF (**127**)	Resin	Not analysed

Further studies indicated that it may be possible to promote monosubstitution on a symmetrical porphyrin such as **121**. In particular, the reagent ratio can be altered by increasing the molar equivalents of porphyrin to favour monosubstitution (without altering the reaction scale). However, when performing the reaction in solution, this approach can be sometimes limiting because it increases the volume of solvent required to solubilise the

porphyrin and impact on the final workup and solvent removal. Meanwhile, switching to a conjugation on solid phase can make the use of porphyrin excess more affordable. This, along with the reaction set-up and morphology, may favour a lower degree of conjugation. In both cases, unreacted **121** can be recovered, but this is more practical when filtering the resin at the end of the reaction. Porphyrin stability to Fmoc-SPPS cleavage procedure in TFA was tested, and the tetrapyrrole ring was stable in these conditions, with and without the presence of the insoluble polymer (see section 6.10.10). When the peptide was not attached to the solid support, no interaction between **121** and the solid support were noted.

When pursuing **126** conjugation on-resin, the spacing between arginine and cysteine was set to 3 AAs. This parameter turned out to be very important, since on-resin reaction between **121** and CGGRGD (two glycine residues as spacer) performed at the beginning of this study did not work, and the starting peptide was found unaffected after cleavage. Steric hindrance of arginine protecting groups could possibly make the thiol unavailable for reaction. This hypothesis was partially supported by the failure of on-resin acetylation of this peptide using the standard procedure 20% Ac_2O in DMF. There would be no need for further spacing if the thiol reaction site is incorporated at the end of a carbon chain: this strategy was explored but soon abandoned because of unsuccessful coupling of mercaptoundecanoic acid (MUA, **130** in Figure 39) to the arginine residue.
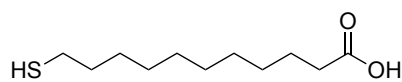
130

Figure 39. Chemical structure of mercaptoundecanoic acid (MUA).

However, when employing peptide **126** and 3 Ala for spacing, the monoconjugated product (**128**) was obtained from on-resin S_NAr on **121** (Scheme 25). When performed in DCM/DBU, the reaction was fast and the porphyrin was easier to recover, but most of the time the purity of the crude was negatively affected. DMF/DIPEA gave a cleaner crude but with the price of a slower, and frequently uncomplete, reaction.

The sequence KLVFF[181,182] was subsequently chosen to expand the scope of this study and explore the conjugation of **121** with a different peptide. After addition of a Gly-Gly spacer and a N-terminal cysteine, the final peptide to be employed resulted CGGKLVFF (**127**). The S_NAr on **121** was attempted with this peptide directly on-resin (Scheme 25), using DIPEA in different solvents (DMF, NMP, DCM, DCE) as reported in Table 11. From preliminary analysis, it seemed that DMF led to both porphyrin mono and disubstitution (**129** and **131** in Figure 40 respectively), while NMP favoured monoconjugation. The regioselectivity of the disubstitution was not extensively explored, but the two disubstitution products were both obtained and observable by LC-MS at different retention times. The solvents DCM and DCE were studied as well, but difficulties when trying to solubilise the product prevented the analysis. However,

a resin colour change from light yellow to purple was observed and this is a visual indication of the porphyrin being attached to the peptide.
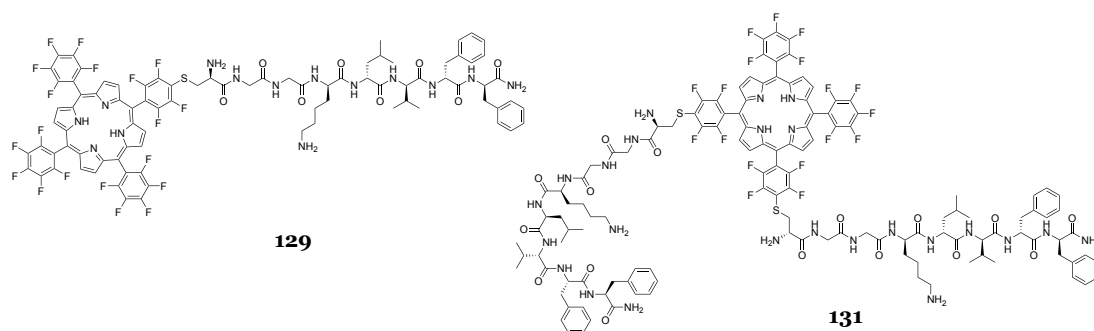
*Figure 40. Chemical structure of mono (**129**) and disubstitution (**131**) product from the reaction between **121** and peptide **127**.*

Due to the amyloidogenic origin of the KLVFF sequence,[183,184] it was very difficult to completely dissolve the final reaction crude product, made of conjugation products and the unreacted starting peptide, at a suitable concentration for HPLC and LC-MS analysis. In addition, these species have a high tendency to self-assemble and did not prove to be suitable for these initial investigations. Therefore, despite some insights of possible product formation and after different attempts to characterise these poorly soluble conjugates, the studies were continued with different peptides.

2.5.4 Practical aspects of peptide-porphyrin conjugation

Reaction set up, work up and analytical procedures are detailed in the Materials and Methods section, but some practical aspects which turned out to be crucial for the development of a peptide-porphyrin conjugation strategy will be presented here.

When working up the DMSO/Cs$_2$CO$_3$ reaction, quenching with TFA combined with precipitation in Et$_2$O had a major impact on the applicability of this solvent/base combination. This procedure allows the residual unreacted peptide and peptide-porphyrin conjugation products to be precipitates as a pellet, while free **121** remained in the supernatant solution. This is also useful because, after Et$_2$O evaporation, **121** remained in DMSO and can be recovered unaffected by precipitation and silica gel filtration.

Alternative workup procedures, employing different acids (such as formic acid) or different precipitation methods did not perform well. The reaction set up has been consequently adjusted to the necessities of the above procedure.

The reaction mixture should be as concentrated as possible (*i.e.* the minimum amount of DMSO necessary to solubilise the porphyrin and the peptide) and should not contain an excessive amount of base to quench. These precautions would allow one to contain the volume of TFA and Et$_2$O employed for the workup. Compared to initial trials, reactions were routinely performed in sealed vials for easiness of creating a nitrogen atmosphere.

From the analytical point of view, large aromatic and fluorinated molecules such as **121** are generally not well ionised in mass spectroscopy. AA or peptide conjugation may help, but ionization could still be challenging. Therefore, it was difficult to identify the porphyrin species from the total ion count (TIC) of LC-MS analysis. For this purpose, time of flight (TOF) based MS systems turned out to be better than quadrupole-based ones.

As previously mentioned, analytical complications can also be faced with ^{19}F NMR when the complexity of the system increases, the reaction crude gets more complex to analyse, and the signal of residual TFA from peptide cleavage and reaction quenching is strong.

Finally, the low solubility of the porphyrin in solvents compatible with HPLC and LC-MS sample preparation (H_2O, MeOH, MeCN) sometimes provided an additional challenge, which was solved by porphyrin coupling with hydrophilic species.

On the other end, the good absorption at 400 nm of porphyrin-related products made the task of distinguishing conjugated and non-conjugated product easier using UV detection at multiple wavelengths.

Disubstitution products were detected during the synthesis of **122**, **123**, **128** and **129**, but no experimental evidence was gathered to distinguish between the two regioisomers. Presumed assignments can be made based on Wu *et al.*[106]

2.6 CONCLUSION AND FUTURE WORK

In conclusion, the initial application of an OVAT approach has provided a detailed understanding of the reactivity of HFB towards NAC. The importance of base and solvent effects was established, and new conditions that afford selective preparation of either 1,4-di or 1,2,4,5-tetra thiol-fluoride substitution products were developed with a 2-level factorial design study in combination with ^{19}F NMR.

This introduced new base-solvent combinations (DMSO/Cs_2CO_3 and MeCN/DBU) that permit rapid (<1 min and <1 h, respectively), clean and selective peptide stapling under peptide-compatible conditions. The products obtained using these procedures are of equal if not better crude purity than those obtained with previously reported peptide stapling approaches (and reported in Table 5 at page 51), by comparison of the published data.

The two-component on-resin perfluoroaryl peptide stapling is achievable with high crude conversion, using conditions that were previously only used in solution. This reaction, performed on solid support, is compatible with different peptide sequences and was effective for HFB-templated cyclisation of SPACE peptide.

The activity of this peptide was retained after replacing the normally occurring disulfide bridge with a 1,4-dithiol substituted tetrafluorobenzene.

Despite the promising results obtained from stapling and macrocyclisation in solution and on-resin, the multicyclisation of a tetracysteine peptide around HFB was not completed using

such conditions. A diminished reactivity after the second substitution (most likely due to steric hindrance), the lack of regioselectivity and the peptide instability were challenges that were not fully addressed, even after the introduction of orthogonal protecting group for cysteine thiols and a stepwise deprotection/cyclisation approach. However, exploiting the acquired knowledge on HFB, a procedure for mono or multiple conjugation of symmetric tetrasubstituted porphyrins was developed. Once again, this strategy exploited the bioconjugation of aryl fluoride and thiol-containing peptides in solution or on solid support via S$_N$Ar.

The peptide-porphyrin conjugation strategies developed in the first part of this work were exploited as the starting point for the development of peptide-porphyrin self-assembly nanostructures (Chapter 3) and for the introduction of porphyrin-templated peptide stapling and multicyclisation strategies (Chapter 4).

3 Peptide-porphyrin self-assemblies

The great reactivity of polyfluoroaromatics towards thiols makes them excellent substrates to which cysteine-containing peptides can be coupled, as outlined in the previous chapter. The feasibility and straightforward conditions of the S_NAr reaction on polyfluorinated porphyrins were presented, and a peptide-porphyrin bioconjugation strategy that has never been reported so far was introduced.[185]

As will be described in this chapter, the knowledge acquired on this thiol-fluoride substitution was applied for the synthesis of peptide-porphyrin bolaamphiphiles, setting the basis for the development of novel light responsive materials made by self-assembling bioconjugates.

Bolaamphiphilic peptides are self-assembling molecules characterised by hydrophilic residues at both ends of a lipophilic chain.[59] To allow porphyrin conjugation, a cysteine residue was introduced in the middle of the bolaamphiphilic sequence KA_6K.

The self-assembly of these conjugates was studied at different concentrations, pH values, ionic strengths, temperatures, and times. Steady-state and time-resolved spectroscopy methods, as well as microscopy techniques, were used to study the progresses of the assembly and the photophysical properties of the final assemblies.

Supramolecular architectures with regular morphologies and defined arrangement can be generated by the self-assembly of peptide-porphyrin conjugates,[62] and the light-responsive properties of embedded chromophores consequently affected. These soft materials can potentially be employed as efficient photo-catalysts, antennas for dye-sensitized solar cells, biosensors, and light-triggered therapeutic agents.

Graphical representations of the conjugates obtained from bolaamphiphilic peptides and polyfluorinated porphyrins investigated in this work are reported in Figure 41.
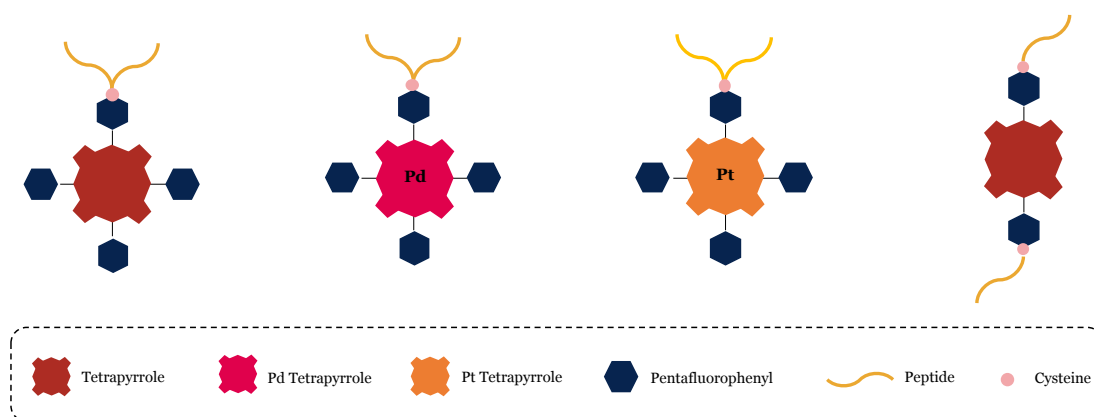
Figure 41. Graphical representations of the conjugates obtained from bolaamphiphilic peptides and polyfluorinated porphyrins investigated in this part of the work.

3.1 BOLAAMPHIPHILIC MOLECULES

The term *bolaamphiphile* (and related adjective *bolaamphiphilic*) derives from *bola*, the Spanish and Portuguese noun for ball. The *bolas* are a type of throwing weapon made of two weights at the ends of a cord, which is used to capture animals by trapping their legs (Figure 42A). In chemistry, *bolaamphiphile* indicates a molecule that has a hydrophobic core (the cord) and hydrophilic extremities (the weights). Therefore, compounds bearing this general structure of hydrophilic-hydrophobic-hydrophilic blocks can be described as a bolaamphiphilic molecules. These species should be distinguished from amphiphilic molecules, which are characterised by two main moieties only, one hydrophobic and one hydrophilic.

This section of the work was focused on bolaamphiphilic peptides and their conjugation with porphyrins. When porphyrins are part of a bolaamphiphilic structure, the tetrapyrrole is generally *meso*-substituted with phenyl groups and represents the hydrophobic block, while two or more para-phenyl substituents confer the hydrophilicity (Figure 42B). These structures have a high tendency to self-assemble, where the porphyrin rings usually promote the stronger interactions and stack with differently organized dispositions. The exemplar bolaamphiphilic porphyrin (**132**) reported in Figure 42C was presented by Parquette research group.[186–188]
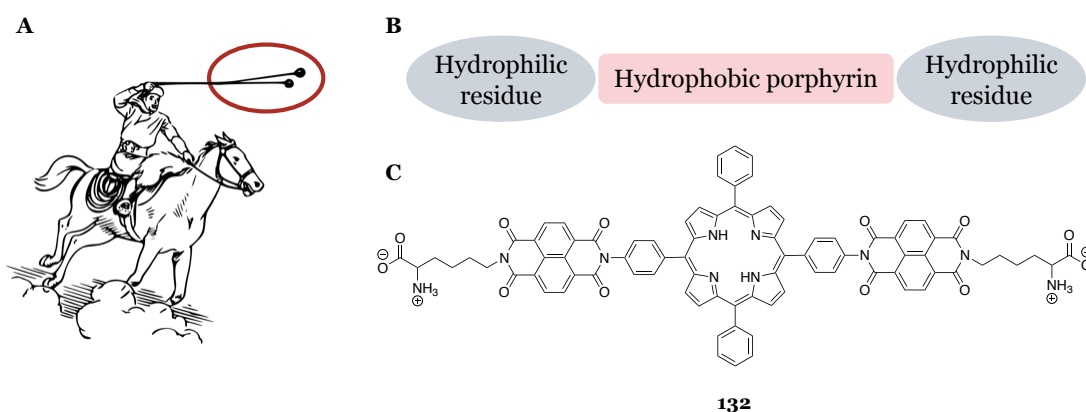
*Figure 42. A) Representation of bolas being thrown for hunting purposes; B) General structure of a bolaamphiphilic porphyrin; C) Chemical structure of a bolaamphiphilic porphyrin (**132**) reported by Parquette.*[186–188]

3.1.1 Bolaamphiphilic peptides

Bolaamphiphilic peptides are self-assembling molecules characterised by hydrophilic AA(s) at both ends of a chain of lipophilic AAs. Differently from other classes of self-assembling peptides, there are few examples of AA sequences reported in literature that respect this definition. Bolaamphiphilic peptides totally made by natural AA residues were principally reported by Chen's, [189,190] Hamley's [191,192] and Xiao's [193] research groups, while there are some other sporadic examples of bolaamphiphilic peptide hybrids.

Chen's bolaamphiphilic peptides contain a hydrophobic core made by a variable number of continuous alanine residues and different hydrophilic ends afforded by AA with cationic,

anionic, and neutral hydrophilic side chains (general structure **133** in Figure 43). When dissolved in water, these peptides self-assembled into small nanospheres, nanorods or nanofibres with lengths up to micrometres. These macromolecular structures were stable to heat (up to 80°C) but could reversibly disassemble into smaller substructures upon sonication. Effects of variation in hydrophobic core and hydrophilic heads on the assembly were investigated by the authors, showing that the assembled nanostructures were mainly affected by N- and C-terminal AAs, especially in relation to pH changes, while the number of AA in the middle section appeared to be less important if the hydrophobic contribution is maintained.[189,190]

n= 4, 5, 6, 7, 8
R= Arg, Lys, Glu, Asp, Ser side chains

133
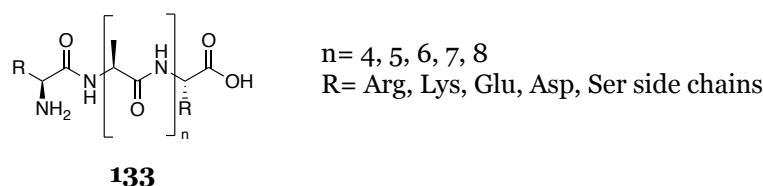
Figure 43. General structure of Chen's group bolaamphiphilic peptides.

Hamley's general bolaamphiphilic structure is slightly more complex than Chen's, but it has less AA variability. In their work, the authors employed bolaamphiphilic peptides with a fixed hydrophobic core made by four leucine residues within two phenylalanine residues, and arginine (peptide **134**) or glutamic acid (peptide **135**) were the hydrophilic ends (Figure 44). Both arginine and glutamic acid derivatives organise into β-sheet architectures, probably because this is driven by the hydrophobic core (particularly the phenylalanine residues) and not the hydrophilic ends. These structures were specifically characterised, and solutions of the arginine derivative were used to coat polystyrene films for cell culture applications.[191,192]

134, R= Arg
135, R= Glu
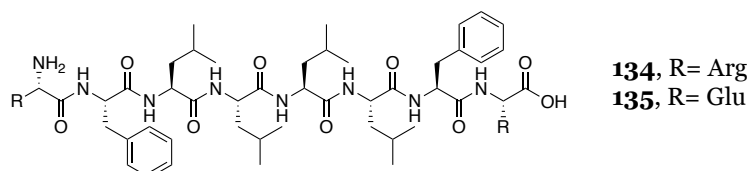
Figure 44. Chemical structures of Hamley's group bolaamphiphilic peptides.

Finally, Yao and co-workers developed a collagen-like bolaamphiphilic peptide. Starting from the GlyProHyp and ProProGly triple helical motifs, which are relatively hydrophobic, they appended two aspartic acids at both ends, functioning as hydrophilic moieties (peptide **136** and **137** in Figure 45).[193] The authors obtained self-assembled nanospheres under a broad range of pH conditions, notwithstanding the content of collagen AAs but requiring terminal charged acidic residues to trigger the self-assembly.

136

137
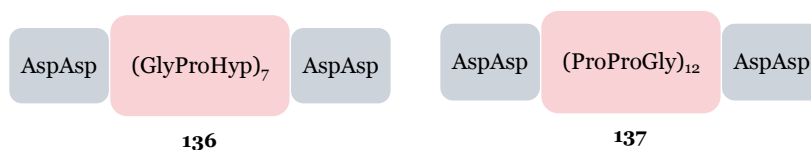
Figure 45. General structures of Xiao's group bolaamphiphilic peptides.

Alternatively, bolaamphiphilic peptide hybrids that generally have hydrophilic AAs at both ends and a carbon-based non-peptidic linker were also synthesised. Some examples are reported in Figure 46.[59,194]
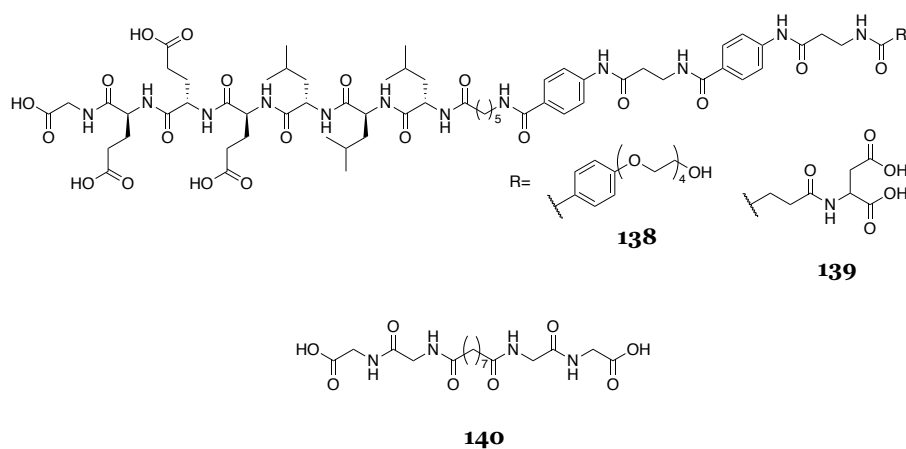
Figure 46. Chemical structures of bolaamphiphilic peptide hybrids.

3.2 BOLA-BIOCONJUGATES

The covalent and non-covalent association of porphyrins with self-assembling peptides has been explored by different research groups and resulted in the generation of differently organised soft materials with application in enzyme mimetics, catalysts, light harvesting arrays, optical switches, molecular photonic wires, and phototherapy.

This peculiar research field have been extensively reviewed,[62,195] identifying a great potential in such systems and room for improvement and progress. For example, a novel bioconjugation reaction could allow access to alternative self-assembling systems made by variable peptides and porphyrins, consequently opening the way to different supramolecular structures and applications. In addition, most of the examples comprising metal porphyrins were obtained from Zn porphyrins, with a few other metals such as Co, Fe, and Sn, sporadically employed.

The interest in the study of peptide, porphyrin, and self-assembling systems as independent topics is widespread, and research works that combine all of them acquired even more relevance. Therefore, the optimised thiol-fluoride S_NAr reaction was exploited with the aim of coupling newly designed bolaamphiphilic peptides to polyfluorinated porphyrins, studying the self-assembling behaviour of these bioconjugates and characterising the electronic properties of the assembled structures. This workflow provided knowledge over the handling and behaviour of these materials, set the basis for further studies in this field and indicated potential applications.

3.2.1 Design of the bolaamphiphilic peptide

The bolaamphiphilic peptide sequence was designed to be water soluble, to comprise only natural AA, to include a cysteine residue as anchor point for the bioconjugation, and to be easy to synthesise. Therefore, Chen's peptide general structure was chosen as the starting point,

being simpler and more water soluble than Hamley's and Xiao's. Also, rather than extended sheets, fibres and spheres were considered more appropriate for initial studies.

Within the array of Chen's bolaamphiphiles variations (Figure 43), six alanine residues were chosen as hydrophobic core because the authors demonstrated the higher self-assembly tendency of KA_6K peptide over longer and shorter chains. It was assumed that, in such a sequence, the influence of an extra cysteine would have been less likely to change the overall peptide self-assembly behaviour.

Regarding the hydrophilic ends, peptides with both acidic (glutamic acid) and basic (arginine and lysine) residues were synthesised, while residues with alcohol side chains were not prepared. Ultimately, arginine and lysine derivatives were preferred over glutamic acid equivalents because of the better solubility in water and in acidified water solutions, which are more commonly used during standard analytical procedures.

As described in section 2.3.1, most of the peptides employed in this work had an amidated C-terminus. However, the bolaamphiphilic peptides and their derivatives were synthesised with a free acid at the C-terminus using a preloaded Wang resin as described in Scheme 26. The microwave-assisted automated solid phase peptide synthesiser, the 0.1 mmol scale and the rest of Fmoc chemistry conditions remained the same as for the peptide amides.

Wang resin preloaded with Fmoc-Lys(Boc)

1. Piperidine, 20% in DMF

Wang resin preloaded with Lys(Boc)

2. Oxyma DIC DMF

Fmoc-protected amino acid

Fmoc-protected dipeptide on Wang resin

Repeat 1 and 2 *n* times

TFA, TIPS, H_2O (95:3:2)

Final peptide
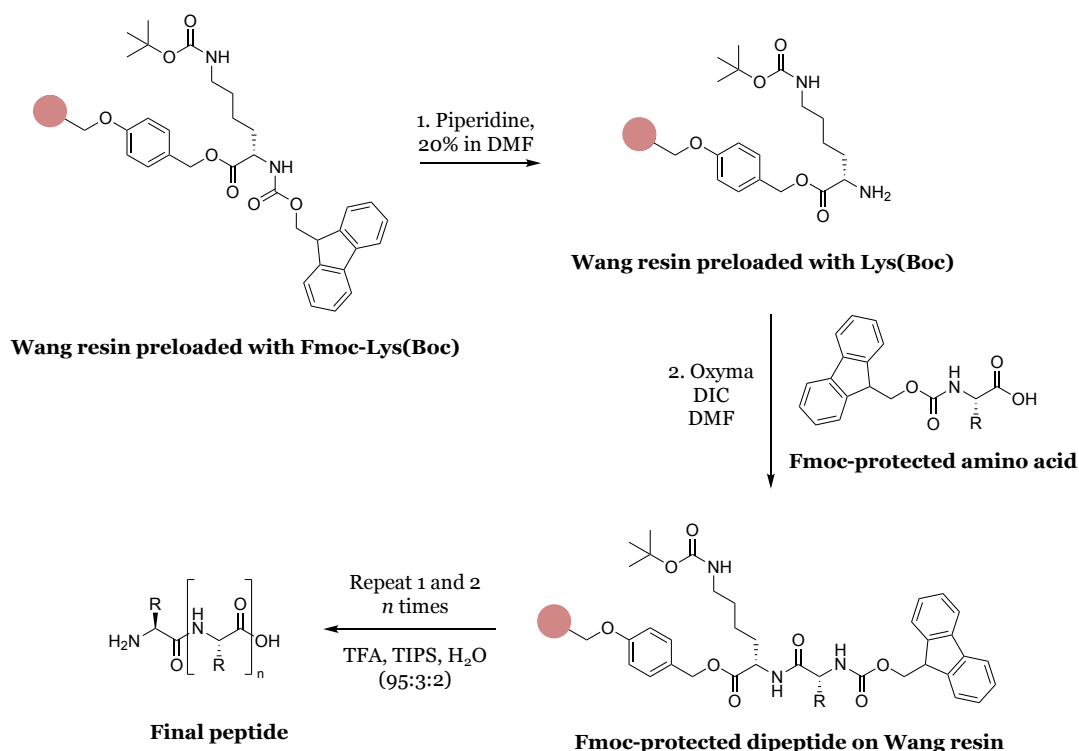
Scheme 26. Reaction scheme and conditions for the Fmoc-SPPS of peptides acids used in this work.

Despite a reasonably easy synthesis of the peptide RA_6R, when a cysteine was inserted in the middle of the hydrophobic core to obtain RA_3CA_3R, the addition of the N-terminal arginine proved to be very challenging. Only around half of the sites reacted after the first coupling and, even after repeated coupling steps, the amount of complete peptide was not improved.

Between arginine and lysine peptides, the latter residues was preferred because the final cysteine derivative granted a better synthetic accessibility.

The selected starting peptide KA$_6$K (**141**) from Qiu *et al.*[189] was therefore modified to obtain KA$_3$CA$_3$K (**142**), whose structures are fully reported in Figure 47. Even after cysteine insertion, the core of peptide **142** can still be considered hydrophobic. It should be noted that Rink amide resin was employed for the synthesis of most peptides mentioned in this work, that consequently presented an amidated C-terminus. As mentioned above, only for the synthesis of bolaamphiphilic peptides, preloaded Wang resins were used and carboxylic acid at the C-terminus were obtained after cleavage. This was necessary to mimic the peptides described in the literature and because the peptide/conjugate charge balance is an important parameter to favour or disfavour the assembly.
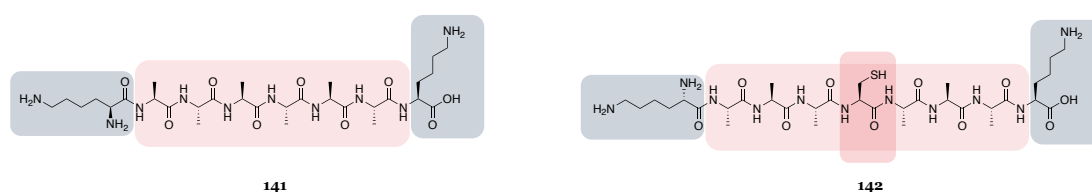
*Figure 47. Chemical structure of peptides KAAAAAAK (**141**) and KAAACAAAK (**142**).*

3.2.2 Synthesis of bolaamphiphilic conjugates

The covalent combination of one peptide and one porphyrin (*i.e.* monoconjugation) was identified as the preferred conjugate design to begin the self-assembly studies. Therefore, coupling of the modified peptide **142**, presenting an anchor point for conjugation as part of the bolaamphiphilic structure, to **121** via thiol-fluoride S$_N$Ar was explored (Figure 48A).

Despite the previous reactions developed on peptides presenting a cysteine at the N-terminus, the reactivity of peptide **142** had to be reassessed because the cysteine was in the middle of the sequence.

The conjugation was initially attempted on-resin because of the collected evidence suggesting that this reaction setup could favour monosubstitution. However, the substitution yield was low, probably because the cysteine in the middle of the peptide chain was not easily accessible.

The reaction was subsequently performed with the deprotected peptide in solution, using DMSO/Cs$_2$CO$_3$, with the addition of TCEP and N$_2$ atmosphere to prevent oxidation, at room temperature and under magnetic stirring. Most importantly, an excess (2 equivalents) of **121** was employed to limit the formation of higher order substitution products. Multiple substitutions could not be completely avoided but conjugate **143** was successfully obtained with minor presence of side products (Figure 48B).

For the above and the following peptide-porphyrin conjugates, reaction were performed at a 0.01 mmol scale following the procedure described in section 6.10 and yields are reported in section 6.10.20 as part of their characterization.

A

B
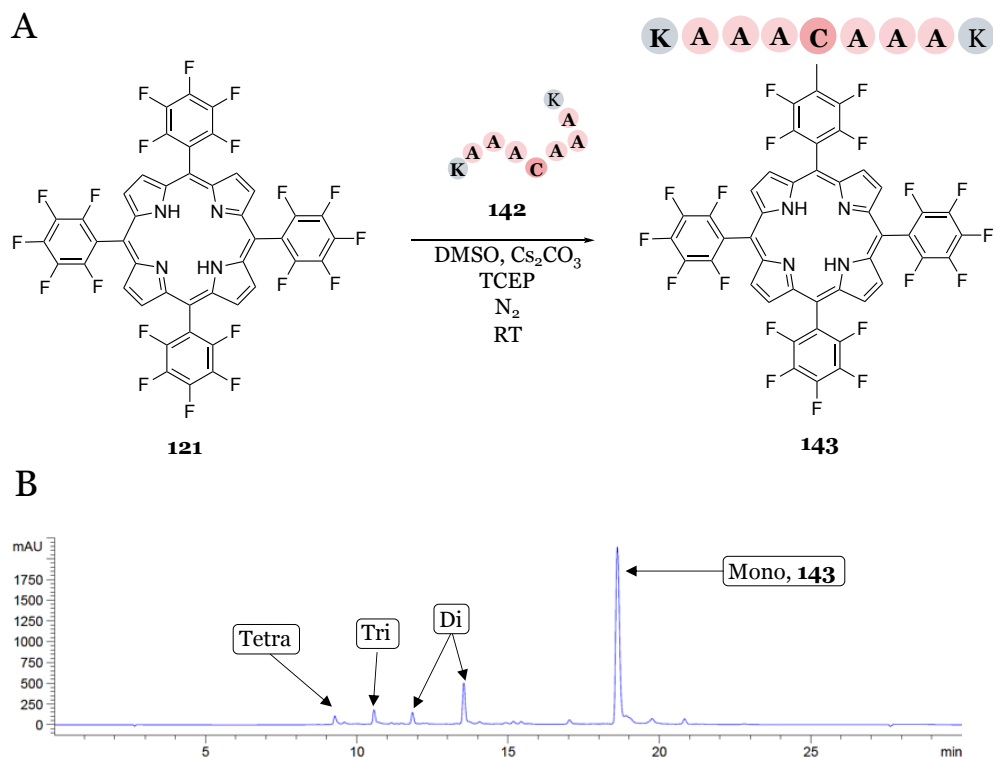
*Figure 48. A) Reaction scheme for the synthesis of conjugate **143**; B) HPLC chromatogram acquired at 400 nm of the crude product from the above reaction.*

Metal derivatives of **121**, specifically palladium 5,10,15,20-tetrakis(pentafluorophenyl) porphyrin (**144** in Figure 49) and platinum 5,10,15,20-tetrakis(pentafluorophenyl)porphyrin (**145** Figure 49), were also synthesised. These porphyrins are characterised by very strong phosphorescence, high triplet quantum yield, long triplet lifetime. The electron-donor substituents of *meso*-aryl groups make them more stable to oxidation and photodegradation than other similar Pd/Pt tetrapyrroles. For these reasons, **144** and **145** are widely employed as oxygen sensors.[196] These porphyrins were also coupled to peptide **142** following the same strategy developed on **121**, obtaining conjugates **146** and **147** with comparable product distribution and similar purity (Figure 49).
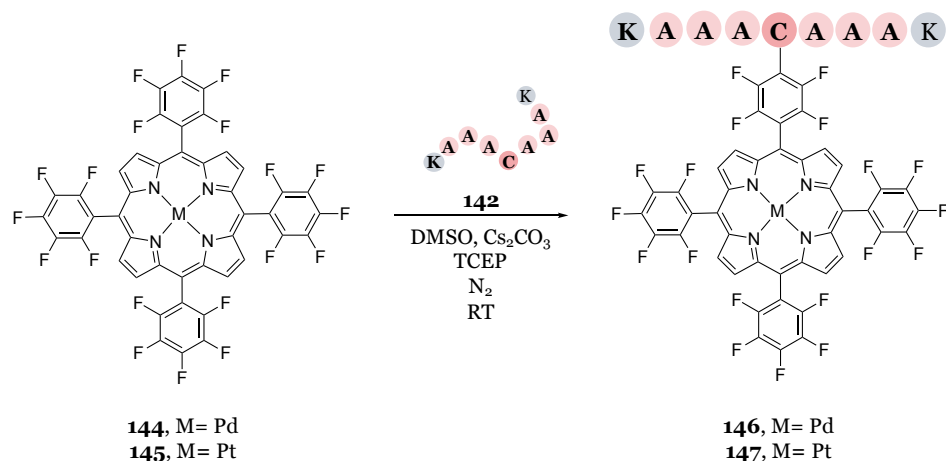
*Figure 49. Reaction scheme for the synthesis of conjugates **146** and **147**.*

Multiple conjugation designs were excluded from this study because **121** disubstitution leads to two regioisomers whose controlled formation, identification and isolation are particularly difficult. Moreover, tri or tetrasubstitution on **121** would lead to a high MW compound and an overall less appealing bioconjugate.

On the contrary, the simplification of the system is preferable. With this purpose in mind, 5,15-tetrakis(pentafluorophenyl)porphyrin (**148** Figure 50) was introduced in the bolaamphiphilic design. Porphyrin **148** is a polyfluorinated tetrapyrrole with two pentafluorophenyl groups in position 5 and 15. The reactivity toward S_NAr is comparable to **121**, but only two substitutions on the two *para* positions are possible. The A2B2 geometry of **148** requires a multi-step synthesis which is more complex and lower yielding compared to **121**. This is probably one of the reasons why **148** is not as much studied as other tetrapyrroles.

In this work, **121** has been used as principal scaffold for peptide conjugation, while **148** was only employed for the synthesis of specific conjugates and as chemical tool to support certain experiments. When **148** is employed for thiol conjugation, only two conjugation sites are available, leading to a controllable disubstitution. However, complete double substitution of **148** with peptide **142** turned out be more challenging than previous conjugations. Moreover, the crude had lower purity and the products were very sensitive to oxidation. Therefore, this reaction and conjugate were not developed further.
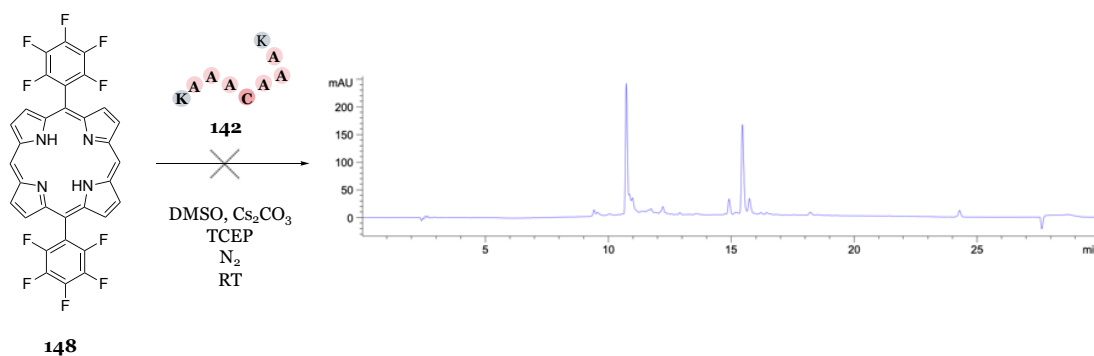
*Figure 50. Reaction scheme for the substitution of **148** with peptide **142** and HPLC chromatogram acquired at 400 nm of the crude product of unsuccessful reaction.*

In line with the same simplification strategy, the peptide **142** was shortened to CAAK (peptide **149**). Technically, this AA sequence is not bolaamphiphilic but, when the double substitution product is obtained, the overall molecule (conjugate **150** in Figure 51) is a hybrid bolaamphiphile.

Reaction on-resin was attempted but turned out not to be suitable for this purpose. In this case, although the cysteine at the N-terminus was free to react, only the monoconjugation product was obtained, while the second substitution necessary to obtain the bolaamphiphile was not achieved. However, the substitution took place cleanly following the procedure in solution and using an excess of peptide (4 equivalents). Even if solubility of **148** in DMSO was not as good as **121**, the solubility of the final conjugate in this solvent was drastically improved.
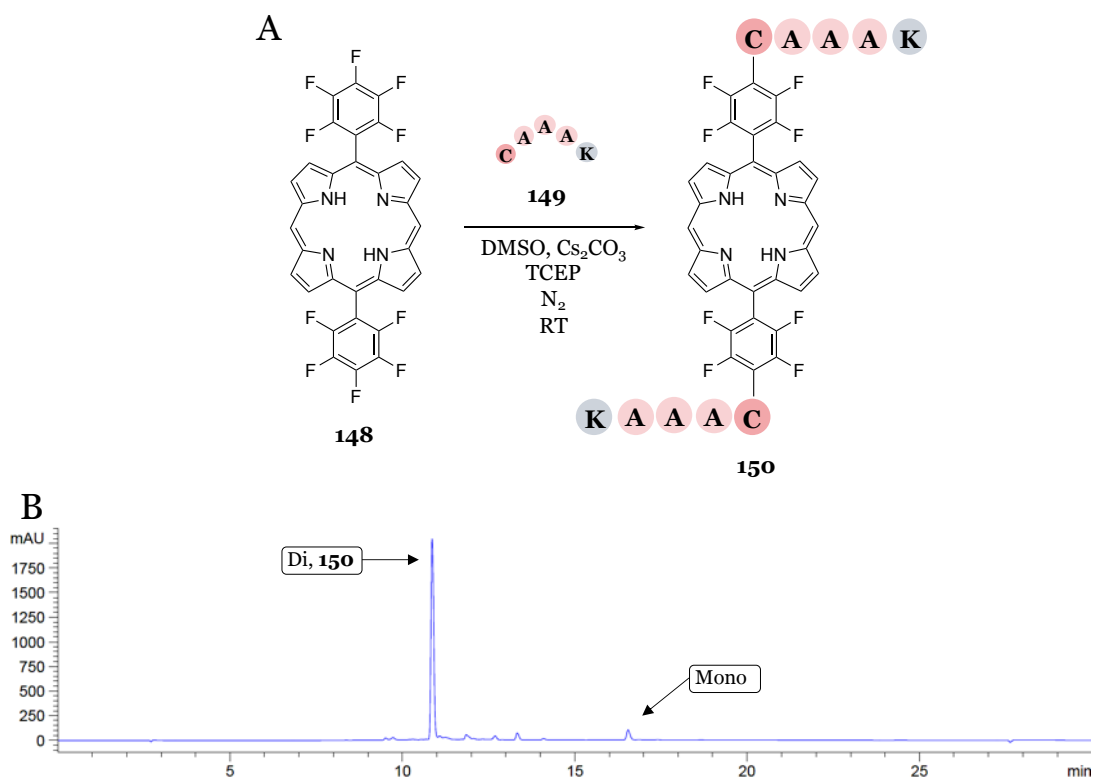
*Figure 51. A) Reaction scheme for the synthesis of **150**; B) HPLC chromatogram acquired at 400 nm of the crude product from the above reaction.*

Glutamic acid derivatives of conjugates **143** and **150** were also designed and synthesised, following the same procedure. However, as previously experienced for the free peptides, handling, analysis and purification were harder compared to lysine counterparts, and satisfactory characterization was not achieved. Therefore, these species were temporarily set aside and will be investigated in future studies.

3.2.3 The "solvent switch" method

Self-assembly can be induced with different external stimuli: the "solvent switch" (or "two solvents") method is the most common procedure, although changes to pH, ionic strength, and temperature are also used. In the first method, the conjugate is initially solubilised in the minimum amount of a chaotropic "good solvent", which can vary depending on the chemical properties of the species under investigation.

Subsequently, the concentrated solution is diluted with a "bad solvent", which must be miscible with the solvent employed in the previous step. Incubation at a certain temperature, after eventual sonication or heating, completes the procedure. After dilution with the bad solvent and incubation, different nanoarchitectures can be formed via self-assembly.

Spectroscopical and microscopy techniques are the standard techniques for characterisation of these systems. Due to the presence of a chromophore within the system, spectroscopy can be informative about the kinetic of the assembly process and the interaction of the porphyrin ring with surrounding molecules, while microscopy allows morphological characterisation.

Examples of good-bad solvent pairs employed in literature to perform the solvent switch method are reported in Table 12.[62] The final concentration of the good solvent is usually around 1-2% v/v and in any case not higher than 5-10% v/v; most of the time further addition of good solvent can reverse the assembly.

Table 12. Summary of good-bad solvent pairs reported in literature for the self-assembly of peptide-porphyrin systems.

Good solvent	Bad solvent
TFE, HFIP, DMSO	H_2O, MeOH, EtOH, THF, CH_3CN
DCM	Heptane, n-Hexane, c-Hexane
HFIP	Hexane
H_2O	MeOH, EtOH
CH3CN	Toluene
THF	Hexane, H_2O

Conjugates **143**, **146**, **147**, **150** are soluble in polar organic solvents such as DMSO, DMF and fluorinated solvents (hexafluoroisopropanol (HFIP) and trifluoroethanol (TFE)), quite soluble in MeOH, ACN and acidified water, and poorly soluble in water only. They are not soluble in apolar organic solvents such as hexane, DCM, toluene. With the aim of exploring the assembly of these conjugates in water-based conditions (*i.e.* H_2O employed as bad solvent, a good system for biological applications), HFIP and DMSO were initially employed as good solvents.

These solvents were chosen for their high solubility power, which allows preparation of concentrated stock solution, and for being water miscible. Concentrated HFIP/DMSO solutions were subsequently diluted with the bad solvent (H_2O) to have 1% v/v of good solvent in the final solution.

3.2.4 Preliminary assembly experiments

Before proceeding with a more extensive study, preliminary assembly experiments were performed to develop experimental procedures as well as understand main variable effects on assembly outcome. Straightforward information about the morphology and macromolecular organisation can be obtained with Transmission Electron Microscopy (TEM).

TEM uses electrons that are transmitted through the sample to generate an image with very high spatial resolution, which is informative of on the inner structure of the sample. In TEM, liquid samples are placed on a small carbon-coated copper grid and allowed to absorb for a few minutes. After sample loading, solutions of inorganic compounds comprising heavy atoms are used as staining reagents. These stains, being electron dense and opaque for electrons, create a contrast between the background and the species deposited on the grid, allowing their visualisation.[197]

Interestingly, the polyfluorinated porphyrins (**121**, **144**, **145**, **148**) has an electron density which scatters the electron beam, and no staining is required to visualise this tetrapyrrole with

TEM. Therefore, the peptide-porphyrin assembled supramolecular architectures were visualised as dark structures on a bright background. Other studies on peptide-porphyrin assemblies employed carbon-coated copper grids without staining, confirming that the tetrapyrrole ring has enough electron density to dissipate the microscope's electron beam.[198]

Similarly, an interesting publication from Sonzini and co-workers showed that the coupling of a single 3,5-bis(trifluoromethyl)-benzoic acid in a peptide can have the same effect.[199] A stain-free TEM analysis is a useful feature, because the sample preparation procedure is simplified and the risk of staining reagents interfering with the sample morphology is avoided.

Initially, a HFIP stock solutions of conjugate **143** was diluted with water, to reach a final concentration of 0.5 mM conjugate and 10% v/v of good solvent. The samples were observed with TEM after different incubation periods at room temperature (Figure 52). One hour after preparation, short and wavy fibres could be immediately observed. These structures got thicker after one day and longer after one week of incubation. It is interesting to note the micrometric length (ca. 9 µm) of a single fibre and its ribbon-like folding after 7 days of incubation.
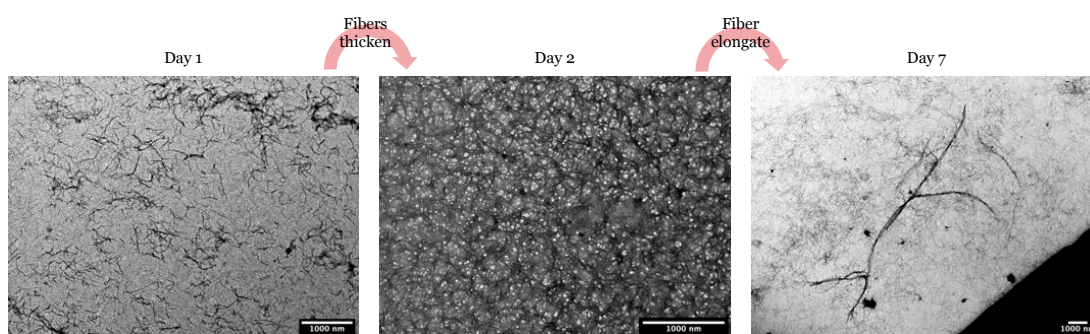
*Figure 52. TEM images of the self-assembly of conjugate **143** (0.5 mM with 10% v/v of HFIP in H₂O) after different incubation times.*

Under the same conditions, conjugate **150** showed a slightly different morphology (Figure 53). Its tendency was to form straight and thin rods/needle, rather than the curvy and thicker fibres observed before for conjugate **143**. These were found to stack together and elongate with time, forming thick bundles but without significant modification of their individual size.
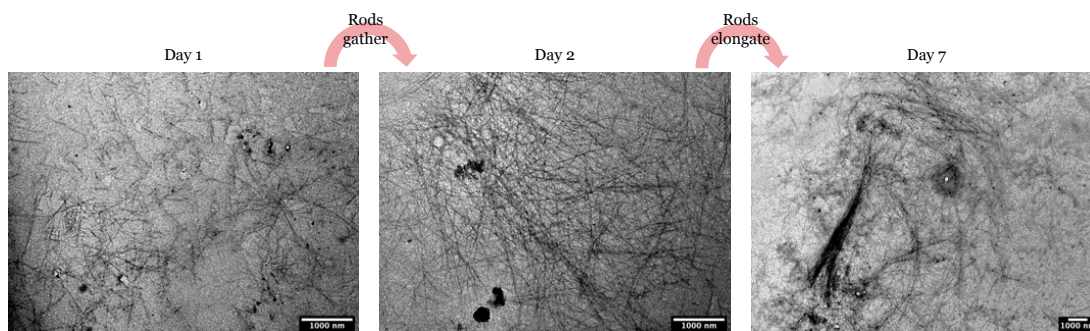
*Figure 53. TEM images of the self-assembly of conjugate **150** (0.5 mM with 10% v/v of HFIP) after different incubation times.*

When starting from a DMSO stock solution, a similar behaviour was observed for both conjugates **143** and **150** under the same dilution and incubation conditions presented for experiment based on HFIP stock solutions.

However, in these examples, the final concentration of conjugate (0.5 mM) and percentage of good solvent (10% v/v) were above average for a standard self-assembly investigation. In a following experiment, stock solutions of conjugate **143** in HFIP or DMSO were diluted with water to a final concentration of 50 μM conjugate and 1% v/v of good solvent. Fibres were not observed under these more diluted conditions, but rather irregular micelles and aggregates that requires further investigation were noted.

From these preliminary experiments, the use of HFIP or DMSO did not have a significant influence on the morphology but, more interestingly and at the same time not surprisingly, the conjugate concentration had a more important role. Similarly, the percentage of good solvent in the diluted solution could influence the assembly as well, but its value should be kept as low as possible.

Compared to DMSO, HFIP has a lower solvent cut off, and therefore less influence on spectroscopic/absorbance measurements. Between the two, HFIP is more volatile, making the solute recovery easier but at the same time handling and storage more difficult. HFIP is also more toxic to cells, which makes the system not suitable for biological applications. Finally, HFIP is acidic: this changes the solution pH and could potentially affect the stability of metallo-porphyrin by removing the metal from the tetrapyrrole cavity. Therefore, DMSO was preferred as good solvent for further studies, but HFIP can still be considered as a valid alternative for specific applications.

3.2.5 Different conditions of self-assembly

After choosing the good and bad solvents for the solvent switch method, there are other variables that can influence the self-assembly process. Some of these are, for example, conjugate concentration, pH value and ionic strength of the solution, as well as temperature and time of incubation.

There are published works in which the influence of some conditions were individually evaluated on the self-assembling process of peptides alone or in combination with porphyrins.[200,201] Understanding the influence of each variable can elucidate the properties of the assembled systems, enhance their stability, and allow control over the kinetics of the self-assembly process and outcome, such as formation of specific nanostructures.

For example, Xing and co-workers were able to tune the self-assembling properties of a peptide amphiphile by using different isosteres of a solvent-exposed carboxylic acid (carboxylic acid, sulfonic acid, phosphoric acid, methyl sulfonamide, methyl sulfonamide, trifluoromethyl sulfonamide, and cyanamide) and contextually modulating variables such as pH, ionic strength, surface charge and temperature. Depending on the isostere pKa value, the peptide had different ionisations states and ultimately formed nanostructures with different

morphologies. All the peptides assembled into ribbons but headgroups were found to influence their width and twisting depending on the conditions. These parameters are also strictly related to the thermal stability of the nanostructures.[202]

However, in all these studies, each variable and its effects has always been assessed individually. Recalling the observations about OVAT and DOE approaches employed for the optimisation of chemical reactions (section 2.2.1), the same experimental procedure was applied to the self-assembly process.

Firstly, UV/vis absorbance and fluorescence emission analysis were selected as the response for each experiment: the variation in absorption/emission intensities and wavelength values are indicative of the occurring assembly process and informative about the engagements of porphyrins and tetrapyrroles in general.[203–205] These properties are also quickly measurable using a plate reader in combination with a 96-well plate, requiring small volumes, and therefore, a small amount of conjugate solution for each experiment. More specific analyses requiring higher solution volumes were performed on selected experiments/conditions at a later stage.

A full factorial design was set up to include the following variables with corresponding low and high values:

- concentration of conjugate in the final solution: 10 μM – 100 μM
- pH of water solution ("bad solvent"): 7.4 – 12
- ionic strength of water solution ("bad solvent"): 10 mM – 100 mM
- temperature of incubation: 5 °C – 37 °C

The amount of DMSO in the final solution was always 1% v/v.

The lowest value of conjugate concentration was chosen because of the instrument lower limit of detection when using 100 μL of solution in each well. The highest concentration value was set 10 times higher to give an absorbance of around 1 AU. This concentration range comprises most of the procedures employed in the reviewed self-assembly experiments. Molecules with a high tendency to self-assemble must retain this behaviour even in diluted conditions and the conjugate concentration should not be too high to prevent non-ordered aggregation.

An excitation range from 300/350 nm to 600 nm was set to include all the porphyrin bands plus margin for eventual wavelength shifts. The presence of DMSO and the plate cut-off prevented this detection to go to lower values. The emission detection ranged from 600 nm to 740 nm (instrument maximum) and included the two porphyrin emission bands.

Peptides **141** and **142** have an isoelectric point of around 10 (calculated with Bachem online peptide analysing tool[206]). At physiological pH, these peptides have an overall positive charge (+2) due to three protonated amines (two lysine residues and the N-terminus) and a deprotonated C-terminus. At basic pH, their overall charge is negative (-1), due to the neutral amines and the deprotonated acid. At acidic pH, the C-terminus as well as the amines are protonated, and there are three positive charges (+3) with no negative charge.[190] The

physiological pH (phosphate buffer at pH=7.4) was included in the DOE to investigate the behaviour in systems mimicking the biological environment.

The basic pH (10 mM NaOH aqueous solution with pH=12) was preferred to the acidic pH because the protonation/deprotonation of the lysine residues could play an important role in the assembly process and was worthwhile investigating. Specifically, shifting from neutral to basic pH and deprotonating the lysine residues could have been informative about their role in the assembly. Differently, shifting from neutral to acid pH and protonating the C-terminus would have informed about the influence of the acid protonation/deprotonation which was considered less relevant at this stage.

Regarding the ionic strength (IS), 10 mM was the lower salt concentration necessary to prepare a phosphate buffer, while the higher IS value was set considerably greater and also including physiological IS (variable, but usually between 100 mM and 200 mM[207]).

The higher temperature value of 37 °C was set to meet physiological condition, but even higher values could lead to peptide degradation and solvent evaporation during incubation. On the other hand, the low temperature of 5 °C was chosen to explore the broader range possible without solution freezing. For practical reasons, this corresponds to the temperature of a usual laboratory fridge.

From the combinations of these variables, the 19 experiments (2^4 + midpoint in triplicate) reported in Table 13 were planned.

*Table 13. List of experiments for the DOE study of conjugates **143** and **150** self-assembling under different conditions.*

Experiment	Conc. (μM)	pH	I.S. (mM)	Temp (*C)
1	100	12	10	5
2	100	12	200	5
3	100	12	10	37
4	100	12	200	37
5	100	7.4	10	5
6	100	7.4	200	5
7	100	7.4	10	37
8	100	7.4	200	37
9	10	12	10	5
10	10	12	200	5
11	10	12	10	37
12	10	12	200	37
13	10	7.4	10	5
14	10	7.4	200	5
15	10	7.4	10	37
16	10	7.4	200	37
17	55	9.7	105	21
18	55	9.7	105	21
19	55	9.7	105	21

Each experiment was prepared and analysed after 10 min, 2 h, 24 h, 82 h, 168 h of incubation at the specified temperature. Sampling at different time points introduces time as another variable of the study.

The choice of the incubation timepoints was mainly related to practical reasons. Dilution, sampling, and analysis of the first timepoint were performed at room temperature and required a few minutes, while an incubation longer than one week would have excessively expanded the experiment duration. This full factorial design was performed on conjugate **143** (in triplicate) and conjugate **150** (in duplicate).

3.2.6 Self-assembly of conjugate **143**

Absorption and emission spectra for each experiment and at every timepoint were the output of the DOE study. These can be collectively visualised to have a representation of the assembly process. Starting with conjugate **143**, three examples of absorption and emission spectra variation over time are reported in Figure 54.

For the DMSO control, a 50 μM solution of conjugate **143** in 100% DMSO was incubated at room temperature: no changes of absorption/fluorescence intensities and wavelengths were measured over time. This means that the conjugate does not self-assemble in the good solvent but rather stays in solution as single molecule. On the contrary, when the intensity decreases and/or shifts, this is indication that the assembly process is taking place and that the chromophores are part of a supramolecular structure.
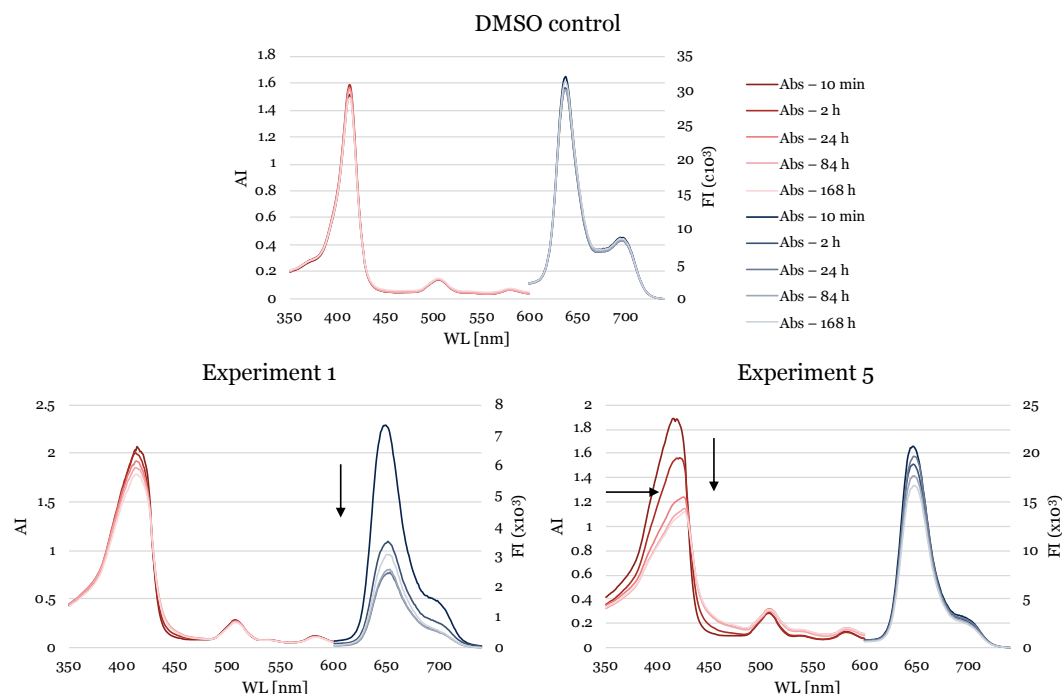
*Figure 54. Absorption and emission spectra of conjugate **143** from DMSO control, conditions of DOE experiment 1 and conditions of DOE experiment 5, at different timepoints as indicated in the legend. The arrows show absorbance/fluorescence intensity decrease and λ_{max} shift with time.*

In experiment 1, a 100 μM solution of conjugate **143** was diluted with aqueous NaOH solution, IS = 10 mM, pH = 12 and incubated at 5° C. Small changes in UV/vis absorption but visible changes in emission were observed from spectroscopical analysis. Differently, significant changes in both intensities and wavelength of the absorption Soret band took place in experiment 5, where a 100 μM solution of conjugate **143** was diluted with PB, IS = 10 mM, pH = 7.4 and incubated at 5° C. In these latter conditions, the shape of the Soret band changes with the formation of a new red-shifted band while variations in emission are less marked. All the remaining absorption and emission data are reported in the material and method section.

A qualitative comparison between these results can help understanding the process. To visualise the results, absorption intensities (AI) at a fixed wavelength (415 nm) were normalised and plotted against time for each experiment. The obtained graph (Figure 55) was informative and helpful to compare the different assembly conditions. A similar behaviour was observed when λ_{max} variations over time were considered.
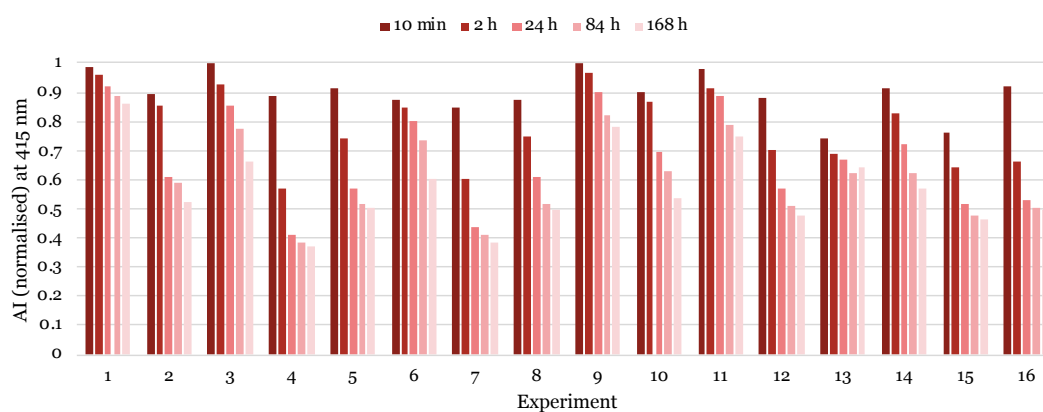
*Figure 55. Absorption intensities at 415 nm obtained from conjugate **143** under the conditions of DOE experiments 1-16 and at different timepoints. Absorption intensities were normalised for concentration.*

As noted by the similar profile of 100 μM and 10 μM experiment pairs (1-9, 2-10, 3-11, …), conjugate concentration did not play a drastic role. AI at 415 nm decreased in all the experiments but was mostly affected when the conjugate was in a neutral pH solution. In this case, IS seems to have lower impact and a low temperature only slows down the process. The variation of AI was lighter for the experiments in basic pH and low IS, regardless the concentration and the incubation temperature (*e.g.* experiments 1, 3, 9, 11). Under these conditions, but with higher IS, the AI was more affected (*e.g.* experiments 2, 4, 10, 12).

A significant shift in the Soret λ_{max} and contextual change in the band shape were also observed over time. As can be noted from Figure 54, this is probably due to fading of the original free porphyrin band and appearance of a signal related to new species absorbing at lower energy. This could be explained by the arrangement of single chromophores into ordered arrays during the assembly process. Specifically, a red shift of the Soret band is usually ascribed to J-

type aggregates. On the other end, Q bands were less affected, with negligible wavelength shift, no shape change, and less pronounced intensity decrease.

The shift of Soret λ_{max} values is another parameter of interesting visualisation and analysis, bearing in mind the previous observations on absorbance intensity variation. The values of λ_{max} versus time were plotted for 100 µM samples (experiments 1-8) and 10 µM samples (experiments 9-16) and reported in Figure 56A and 56B, respectively.

At first glance, the two graphs are reasonably similar, as confirmation of the small effect of a different conjugate concentration. Experiments at basic pH and low IS, which showed small absorbance intensity variations, did not show relevant shift in the Soret band. On the contrary, all the conditions that promoted decrease in absorption intensity also promoted shift in the Soret band.

The shift was more or less rapid, depending on the experiment conditions, but usually reached the same endpoint (*i.e.* between 425 nm and 428 nm). Experiments 2-6 and 10-14, which were performed under the same conditions, are exceptions because the trend varied depending on the concentration. At high IS and basic pH, higher conjugate concentration promoted a shift in the Soret band, similarly to neutral solutions. Meanwhile, no shift is observed with lower conjugate concentration, where there is a clear difference between pH 12 and 7.4.
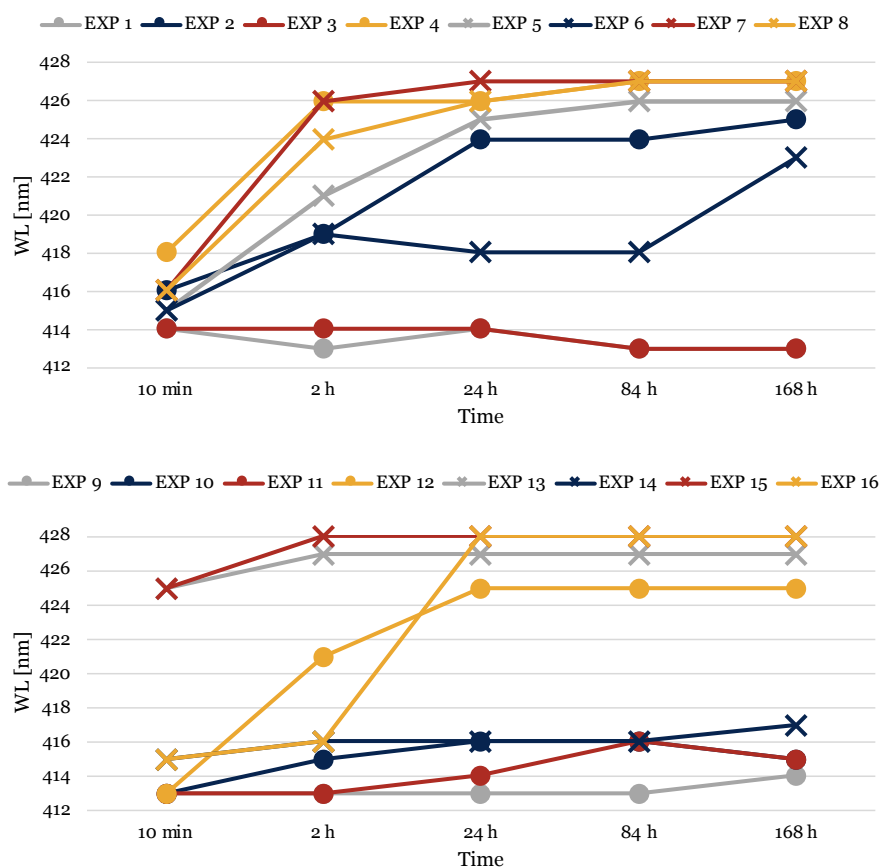
Figure 56. Variation of λ_{max} (Soret band wavelength) during the self-assembly process of conjugate 143 under the conditions of DOE experiments 1-8 (A) and 9-16 (B) at different incubation times.

The intensity plot in Figure 57 can also be generated for emission intensities at a fixed wavelength (650 nm, one of the two porphyrin emission bands) over time. During the assembly of conjugate **143**, significant and fast fluorescence quenching was observed in most of the experiments. This is indication of a close arrangement between porphyrins that allows intermolecular electronic interactions between excited singlet states.

Fluorescence quenching is in accordance with absorbance red-shift, because this spectroscopical behaviour is typical of J-aggregation and suggest that hydrophobic and/or π–π interactions play a role in the assembly. Porphyrin aggregation and fluorescence quenching are generally associated with branching factors such as singlet-singlet annihilation, electron, and energy transfer between the aromatic moieties, but were not investigated further.[208]

For experiments 1, 2, 4, 9 and 12, there was an unusual rise in intensity after a strong and fast quenching. This may be ascribed to instrument instability or rearrangement in the assembly. Since the fluorescence was not completely quenched, there might be free emitting monomers in solution which did not assemble.
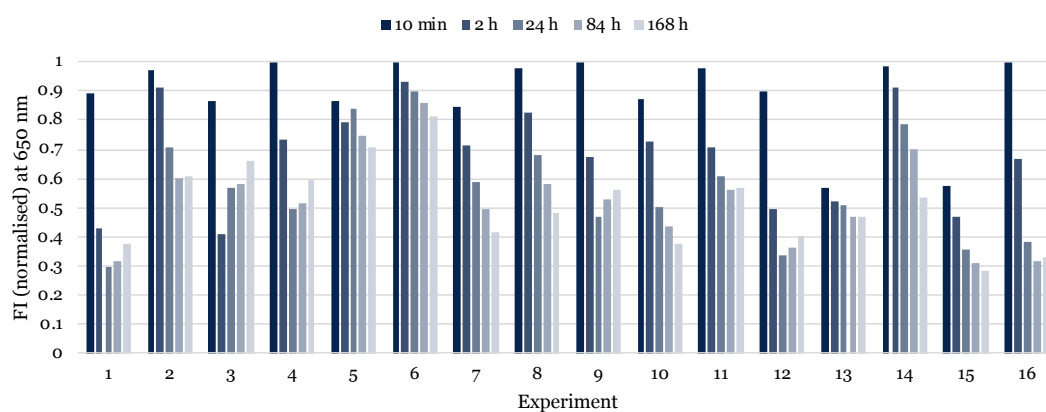
*Figure 57. Emission intensities at 650 nm obtained from conjugate **143** under the conditions of DOE experiments 1-16 at different timepoints. Emission intensities were normalised for concentration and solvent.*

Contrary to absorption bands, fluorescence emission bands did not undergo significant wavelength shift. Therefore, plotting fluorescence intensities at each λ_{max} over time would not be informative. Variations of band shape were also not observed.

After collecting the experimental data, Minitab software can be used to construct a model that describes the relationship between the response and critical factors of a DOE study. If the model is adequate, this can be used to optimise and control the process variables. Absorbance intensities recorded at 415 nm and normalised for concentration were therefore loaded into a Minitab model design as response value for each DOE experiment. However, the software could not adequately describe the relationship between the response and the factors because a curvature was identified in the data.

Therefore, 12 extra experiments (reported in Table 32, section 6.11) were added to the design to model the curvature. After a quadratic model was fitted with these additional points, this was able to explain 84.25% of the variation in absorbance values. The Pareto chart (Figure 58), main effect plots (Figure 59) and interaction plots (Figure 60) generally confirmed the qualitative observations previously discussed. Interestingly, the variable "concentration" appeared to be the one determining the curvature in the model and would require further investigation.

This analysis was not performed with measurements of fluorescence emission as they appeared to be more variable and less stable than absorbance ones. Overall, this process is probably too complex to be fully interpreted and reliably controlled with such model and the current knowledge.

For the purpose of this thesis and related discussion, DOE spectroscopic data were presented using various plots, but direct observation and comparison of full spectra reported in section 6.11 is probably the best way to analyse these results.
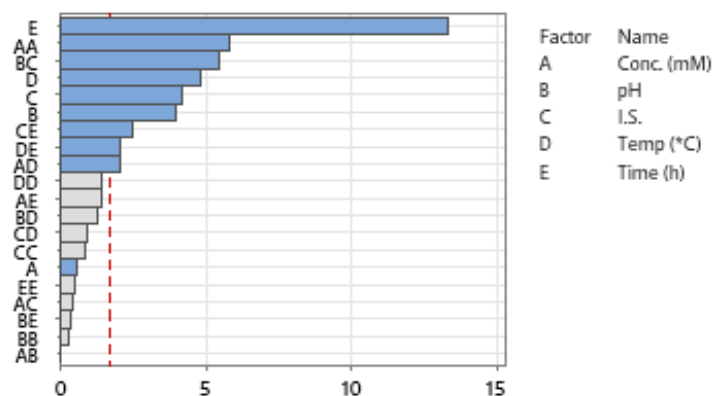
*Figure 58. Pareto Chart for the quadratic model describing the influence of each factor on the assembly of conjugate **143**. The blue bars represent the terms that were included in the model and a longer bar means more influence on absorbance, while grey bars are for terms that were removed. There is a relationship between absorbance and the factors in the model at the 0.1 level of significance (red dotted line).*
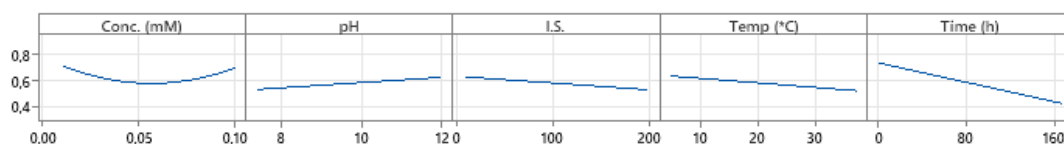
*Figure 59. Main effect plot for the quadratic model describing the assembly of conjugate **143**. This chart describes how the absorbance at 415 nm changes with the change of a single factors.*
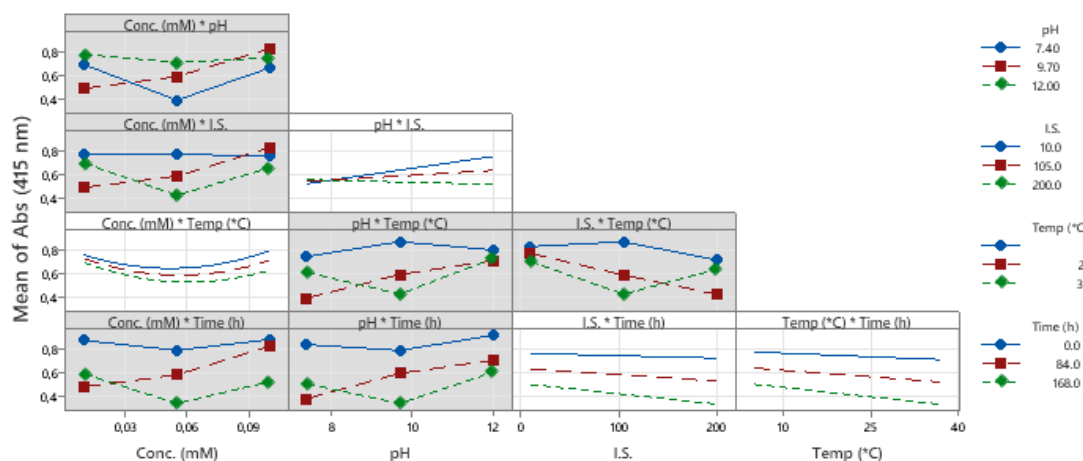
*Figure 60. Interactions plot for the quadratic model describing the assembly of conjugate **143**. This chart describes how the absorbance at 415 nm changes with a change of the two factors. A grey background represents a term that was removed from the model because not statistically significant.*

Contextually to absorbance and emission measurements, some of the experiments were also sampled at different timepoints and analysed with TEM and Scanning Electron Microscopy (SEM) microscopy. Compared to the previously described TEM, SEM creates an image by detecting reflected electrons with a lower spatial resolution. As a result, SEM can only provide information on the sample surface.

Spheres were the prevalent structure (Figure 61A and 61B) observed in the samples prepared at pH= 7.4 (experiments 5-8, 13-16). These were regular nanoparticles with size between 100 and 200 nm. However, due to their relatively small dimension, they could not be distinguished clearly using SEM, where only clusters of nanospheres were observed (Figure 61C). The DLS of the same sample is reported in Figure 62, as a confirmation of the particle size range observed with TEM and SEM.
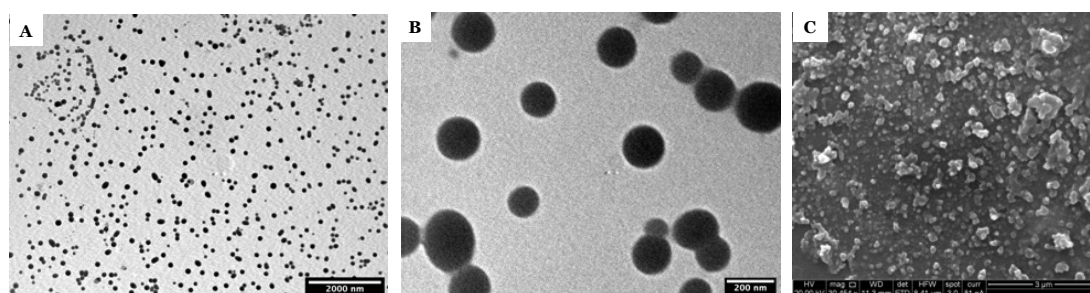
Figure 61. TEM (A and B) and SEM (C) images of a sample from DOE experiment 16, after one week of incubation.
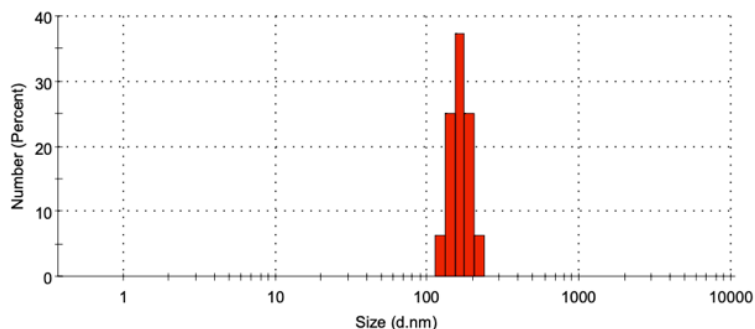
Figure 62. DLS of a sample from DOE experiment 16 after one week incubation.

At pH= 12, spheres were observed only in samples with 100 μM conjugate, IS= 200 mM and 5 °C incubation temperature. However, compared to the nanostructures observed at pH= 7.4, these were smaller and often organised with a branched disposition (Figure 63A and 63B). These slight variations may be due to a different assembly rate and peptide charge. In the remaining experiments, straight fibres or irregular structures were observed (Figure 63C).
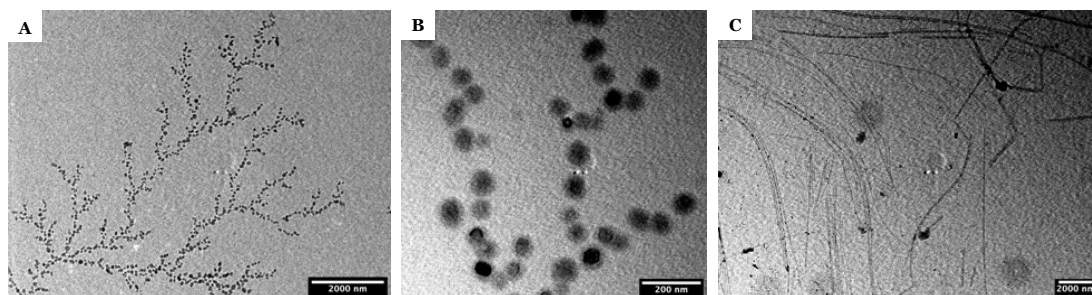
Figure 63. TEM images of samples from DOE experiments 2 (A and B) and 12 (C).

In the DMSO control, the TEM grid was mostly empty, with just a few dark, small, and irregular dots spread around the analysed surface (Figure 64A). This observation is compatible with the absence of assemblies in these conditions. Overall, 10 μM seems to be a good concentration for electron microscopy analysis of conjugate **143** assemblies, while irregular structures were formed with the higher 100 μM concentration (in Figure 64B and 62C). These can be irregular aggregates of conjugates and/or agglomerates of pre-assembled nanostructures.
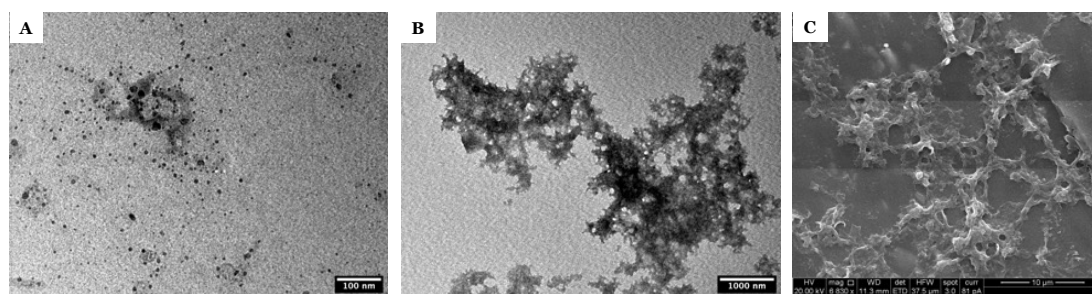
Figure 64. A) TEM images of a sample from DMSO control solution; B) TEM and C) SEM images from DOE experiments at 100 μM concentration.

In conclusion, the self-assembly was generally favoured in neutral pH. The IS effect was more relevant in basic pH, when a high ionic concentration stimulated the assembly which is otherwise less favoured. A higher temperature speeded up the process but did not change the result in term of spectroscopical properties variation and nanostructure morphology. On the contrary, a slower assembly at lower temperature may favour an ordered architecture. The conjugate concentration played a minor role for the self-assembly of **143**, meaning that the process take place equally even at the preferred lower concentrations.

A correlation between wavelength shift and nanoparticles formation was observed. At basic pH, nanospheres were only observed at higher conjugate concentration. Their formation is associated with a wavelength shift which was comparable to the one measured at neutral pH, condition which is likewise characterised by nanoparticle formation. Meanwhile, when fibres and irregular structures are formed, there was minor decrease in Soret intensity and no wavelength shift.

Finally, both the peptide and the porphyrin moieties probably contributed to the assembly. The pH dependent behaviour implies a contribution of the peptide, which was stronger when ionised. However, the marked variation in absorbance and emission indicated that the porphyrin rings are interacting and playing a role in the process as well.

3.2.7 Self-assembly of conjugate 150

A full factorial design was also performed to study the self-assembly of conjugate **150**, following the procedure described for conjugate **143** and under the same conditions (experiments 1-19). Examples of conjugate **150** absorption and emission spectra in 100% DMSO, in PB pH = 7.4, IS = 10 mM after incubation at 5 °C (experiment 5), and in PB, pH= 7.4, IS= 10 mM after incubation at 37 °C (experiment 8) are reported in Figure 65. The remaining absorption and emission scansions are reported in the section 6.11.

The DMSO control confirmed the absence of any assembly process in DMSO solution, as commented for conjugate **143** earlier. After dilution with water, the difference between the absorption spectrum of conjugate **150** in non-assembly conditions (DMSO control) and the one in assembly conditions (water-based solutions) is considerable.

Particularly, in all the assembly conditions tested, a new band appears at 330 nm from the first time point (10 minutes) onwards. The intensity of this unusual band is comparable to the original Soret band at 400 nm. In this case, H-type arrangements of porphyrin rings could be the origin of this marked and blue-shifted band.[209,210]

A small shoulder was also generated in water-based conditions around 450 nm, maybe ascribable to co-existing J-type aggregates in smaller percentage. Therefore, the fast formation of both aggregates, as well as the presence of some residual free porphyrin, provoked wavelength shift and band broadening, and was responsible of the big changes in the absorption spectrum. This immediate changes in the UV/vis spectrum indicate a rapid assembly, which may be related to a less ordered architecture and explain the presence of

multiple bands. At later time points, the spectral changes were limited: only a slight intensity decrease and negligible wavelength shift were observed.
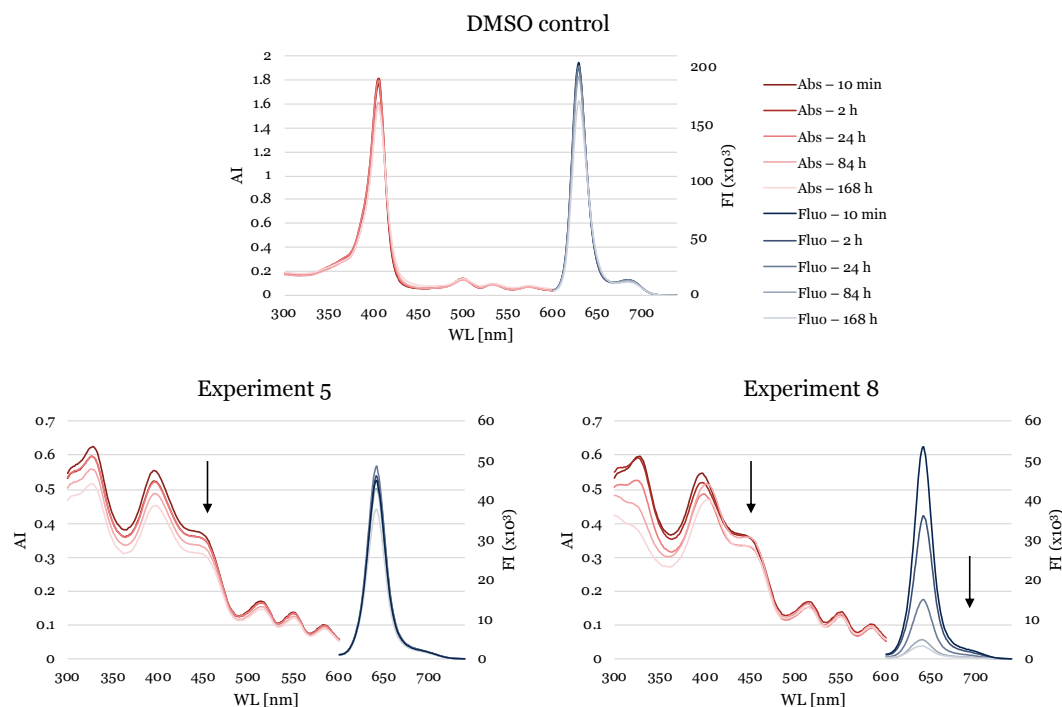
*Figure 65. Absorption and emission spectra of conjugate **150** from DMSO control, conditions of DOE experiment 5 and conditions of DOE experiment 8, at different timepoints as indicated in the legend. The arrows show absorbance/fluorescence intensity decrease with time.*

This small spectral variation between the initial and final timepoints was confirmed when the intensity of absorption at 330 or 400 nm was plotted over time for each experiment (Figure 66). At 330 nm, the most pronounced absorbance intensity variations are those measured from experiment incubated at 37 °C (experiments 3, 4, 7, 8, 11, 12, 15, 16), and even more at pH 12. However, if compared to **143** profiles, variations between different timepoints of the same sample were limited.

The same comment was applicable for absorbance intensities at 400 nm: except for experiments 9 and 11, which have a low conjugate concentration in NaOH 10 mM solution and started from a higher intensity, variations were even smaller.

Fluorescence emission spectra did not change in shape, but intensities were quenched (Figure 67) and behaved similarly to what observed for conjugate **143** (Figure 57). Prominent FI quenching was measured in experiments with incubation at 37 °C, the same conditions that also led to the most noticeable decrease in absorption intensity. Other conditions did not cause large variations, but sometimes an unusual increase of FI at intermediate time points was observed.

As anticipated earlier, after the immediate generation of new bands, spectra were not subject to further wavelength shift and, for this reason, plotting λ_{max} over time was not informative.
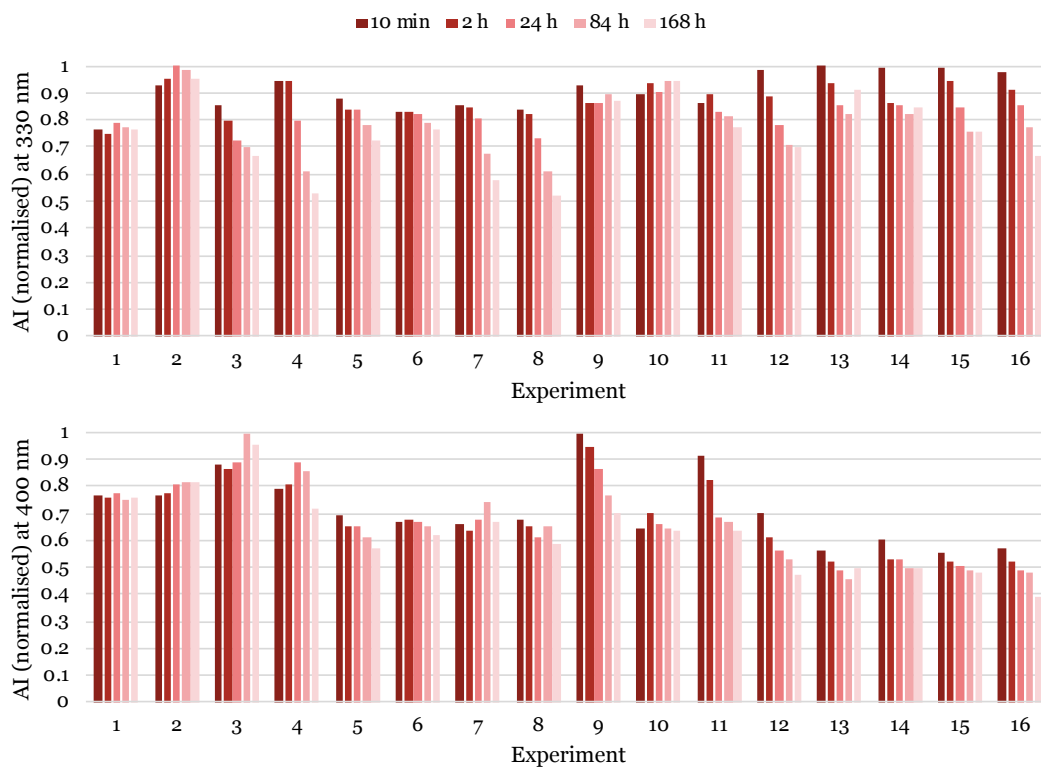
*Figure 66. Absorption intensities at 330 nm (A) and 400 nm (B) obtained from conjugate **150** under the conditions of DOE experiments at different timepoints. Absorption intensities were normalised for concentration.*
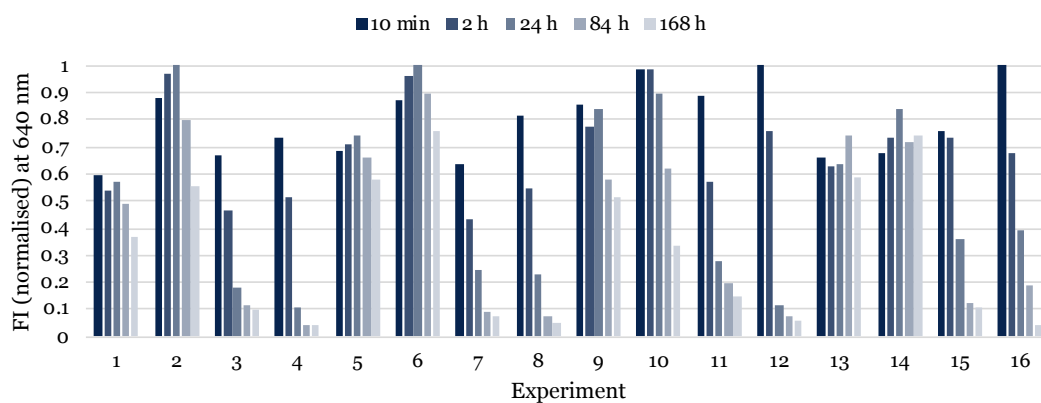
*Figure 67. Emission intensities at 650 nm obtained from conjugate **150** under the conditions of DOE experiments 1-16 at different timepoints. Emission intensities were normalised for concentration and solvent.*

When Minitab analysis was performed on conjugate **150** absorbance data as previously described in section 3.2.6 for conjugate **143**, the quadratic model was able to explain 92.49% of the variation in absorbance values.

The Pareto Chart, Main Effect Plots, and Interaction Plots for conjugate **150** are reported in Figure 68, 69 and 70 respectively.
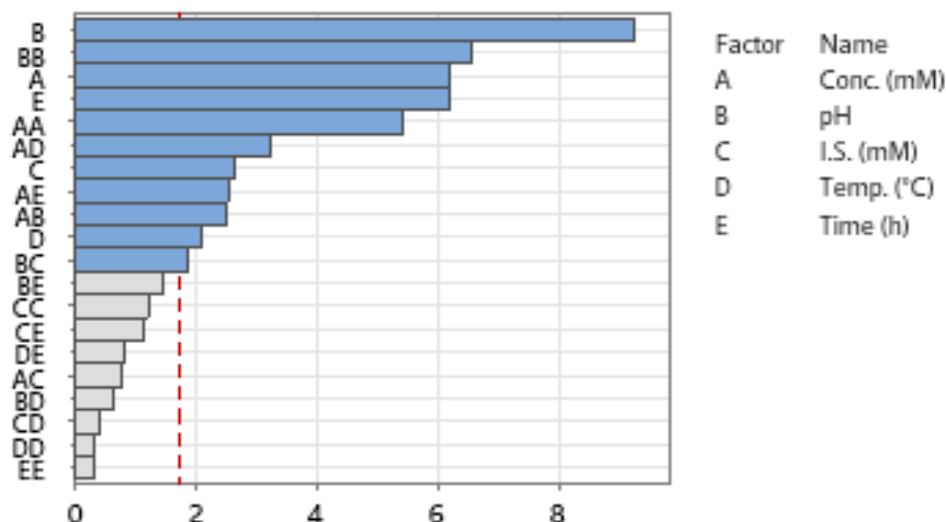
Figure 68. Pareto Chart for the quadratic model describing DOE for the assembly of conjugate **150**. The blue bars represent the terms that are included in the model, while grey bars are for terms that were removed. There is a relationship between absorbance and the factors in the model at the 0.1 level of significance (red dotted line), while a longer bar means more influence on absorbance.
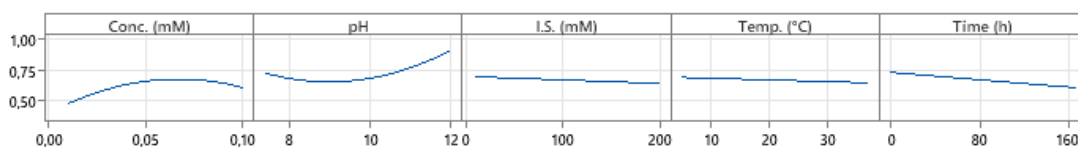
Figure 69. Main effect plot for the quadratic model describing DOE for the assembly of conjugate **150**. This chart describes how the absorbance at 400 nm changes with the change of a single factors.
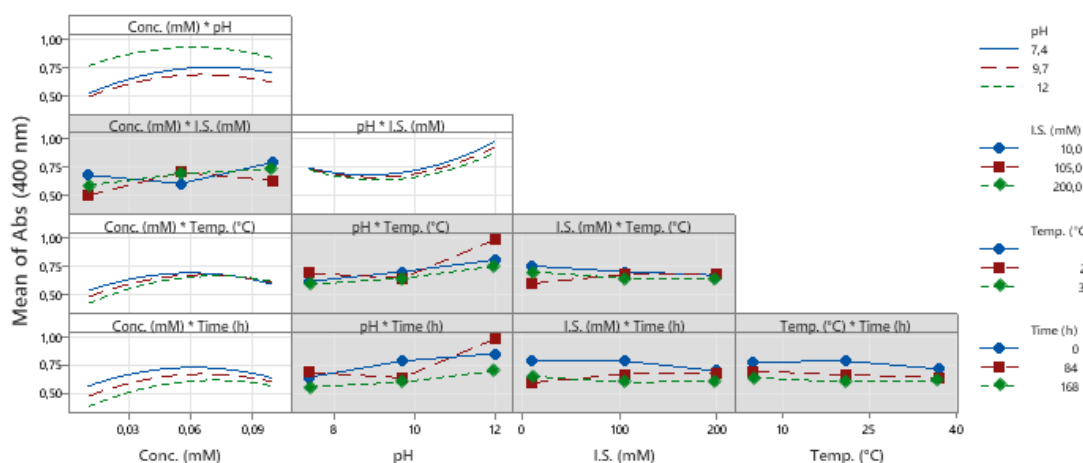
Figure 70. Interactions plot for the quadratic model describing DOE for the assembly of conjugate **150**. This chart describes how the absorbance at 400 nm changes with a change of the two factors. A grey background represents a term that was removed from the model because not statistically significant.

From this analysis, pH, ionic strength, and concentration were the factors that most influenced the assembly. Aside from the smaller influence of time, these results are slightly in

contrast with previous comments. The main divergence for this analytical process is that the model does not include the starting point conditions and characteristics, in addition to the limitation previously described.

When performing microscopy analysis of **150** assemblies, grids obtained with lower conjugate concentration experiments (10 µM) appeared poorly populated. Therefore, the pictures reported below were obtained from experiments with high conjugate concentration (100 µM).

At basic pH values and low IS, flake-like structures as well as irregular aggregates were observed with TEM (Figure 71A). However, some spheres were formed – surrounded by irregular structures – by increasing the IS under the same pH conditions (Figure 71B).

At neutral pH, spheres became the widespread structure, which size is generally smaller than 100 nm with some larger exceptions (Figure 71C). Interestingly, spheres can be observed even if the sampling is performed shortly after dilution of the DMSO stock solution with the aqueous phase (Figure 71D). This is consistent with the sudden change observed in absorption spectrum.

Despite the overall morphology and size being similar, these nanoparticles did not appear to have a shape as regular as those formed from **143**, maybe due to the rapid assembly. SEM analysis of conjugate **150** assemblies was also performed, but results are not presented because were not informative.
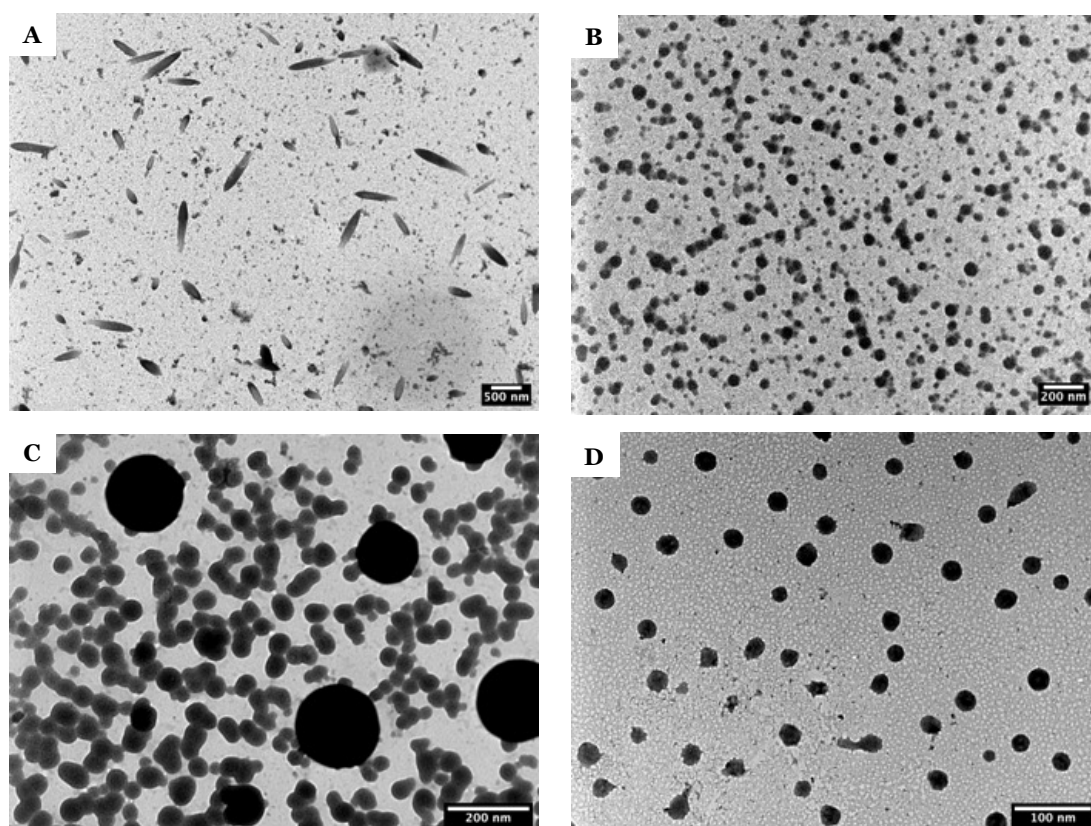
Figure 71. TEM images of samples from DOE experiments 1 (A), 4 (B), 7 (C) and immediately after preparation (D).

Overall, these results indicated a very quick assembly in all conditions of pH, IS, conjugate concentration, time, and temperature of incubation. This behaviour led to an instant and marked modification in the absorption spectrum morphology, while further changes during incubation were levelled. The spectral changes appeared to be the same regardless the morphology of the nanostructures. The small responsiveness to pH and the lower sensibility to ionic strength may indicate that, in this system, the peptide has lower influence in the assembly process which is rather directed by the porphyrin. Neutral pH and 10 µM ionic strength gave the best results in term of nanoparticle regular morphology and size distribution, as confirmed from DLS analysis (Figure 72).
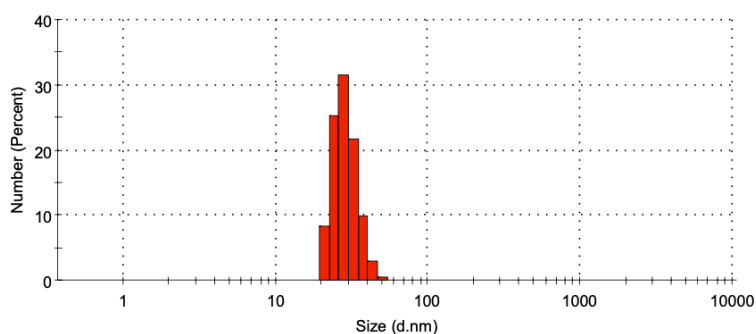
*Figure 72. DLS of a sample from the self-assembly of conjugate **150** in PB, pH = 7.4, IS = 10 mM, after incubation at room temperature for one week.*

3.2.8 Photophysical properties

The properties and consequent applications of a photosensitizer depend on its electronic behaviour when excited with light irradiation. Steady-state techniques employed for the analysis of DOE experiments were useful to have a first understanding of these systems, but time-resolved absorption and emission spectroscopy can better elucidate electronic transitions and give more information about molecular organisation. Therefore, besides early steady-state absorption and emission analysis, fluorescence anisotropy, time-correlated single photon count and laser flash photolysis measurements were also performed at the University of Parma with the collaboration of Prof Cristiano Viappiani.

Time-correlated single photon count (TCSPC) is employed for the measurement of fluorescence decay lifetime and provides more information than simple steady state fluorescence emission scan. For example, time-resolved data may reveal different decay times for the same fluorescent molecule depending on the chemical environments or the influence of other components. For this study, the mean value of each fluorescent decay lifetime was extrapolated from the analysis, as more sophisticated studies are required to identify the origin of each component.

Laser Flash Photolysis (LFP) is a technique for studying transient absorption of chromophores. A short and intense light pulse from a laser source excites the sample and creates short-lived excited states. These intermediates are associated with temporal changes

in absorption characteristics, which can be recorded using a spectrally continuous lamp in a single beam absorption spectrometer.

Fluorescence Anisotropy (FA), which describes the polarization of the fluorescence emission, was also analysed. Upon excitation with polarized light, the emission from a sample can be randomly oriented or polarized. Fluorophores with rate of rotational diffusion faster than the rate of emission have a depolarized emission and anisotropy close to zero. However, when fluorophores are conjugated to macromolecules or part of nano/micro-assemblies, the rotational diffusion gets slower. If the orientation of the fluorophore becomes significant and affects the emission polarisation, the measured anisotropy values are between +1 and -1.

The photophysical properties of non-assembled species were initially measured from conjugates solutions in the good solvent. The behaviour of the conjugates in the good solvent can be considered as a "starting point" for comparison with the results obtained from the respective assembled system after addition of the bad solvent.

As previously explained when the Jablonski diagram and the Stern-Volmer equation were discussed, the presence of molecular oxygen influences the electronic behaviour of a photosensitiser. Therefore, photophysical measurements can have variable results depending on the oxygen concentration in the sample. In this work, the analyses were generally performed in oxygenated solvents but, when indicated, these were also repeated after deoxygenation of the sample with 15 minutes of nitrogen flow.

The measured parameters for a 5 μM solution of conjugates **143** and **150** in DMSO are summarised in Table 14. These values are comparable with reported literature data on **121**, albeit measured in other solvents.[211] Less spectroscopical information, especially regarding time-resolved techniques, are available for **148**.

*Table 14. Photophysical characterisation of conjugates **143** and **150** in DMSO solution. All analysis were performed on oxygenated samples.*

	Absorbance	Fluorescence	TCSPC	LFP
143	412, 506, 581 nm Intensity of Soret: 0.8	Ex: 500 nm 639, 704 (6×10^5) nm	Ex: 500 nm Em: 700 nm 13.61 ns	Abs: 450 nm 1.8 us
150	405, 499, 533, 573, 625 nm Intensity of Soret: 0.75	Ex: 500 nm 628 (8×10^5), 693 nm	Ex: 500 nm Em: 630 nm 12.3 ns	Abs: 450 nm 1.2 us

A single condition from DOE studies was chosen for further investigation of the chromophore properties in the assembled state. Specifically, DMSO stock solutions of conjugates **143** and **150** were diluted with PB, pH=7.4, IS=10 mM to a final concentration of 10 μM. These conditions from DOE experiments were chosen because showed to induce the self-assembly as well as nanoparticles formation. Solutions were incubated at 5° C and individually analysed in quartz cuvette right after dilution (t1, analysis time between each technique to be

considered), after 24 h (t2) and after one week (t3) of incubation. The fourth time point (t4) was measured after 3 more days of incubation at room temperature, because the system may not have reached the final assembly state after one week at 5°C. The results from these analyses are reported in Table 15.
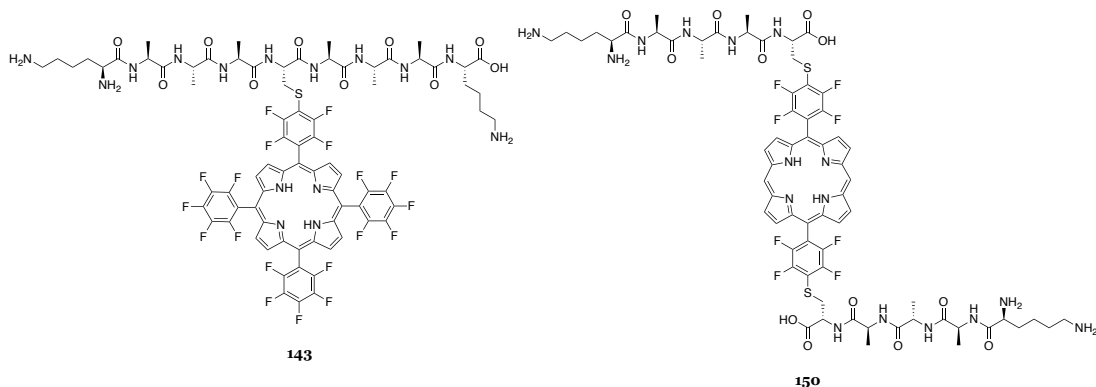
*Figure 73. Chemical structure of conjugates **143** and **150**.*

Absorption and fluorescence scan measurements confirmed the results obtained from the analysis of DOE experiments performed with the plate reader (see section 6.11). Often, a decrease in fluorescence intensity goes together with a decrease in fluorescence lifetime. Both conjugates showed a fluorescence decay much shorter than the corresponding DMSO value, and the lifetime also decreased over time. These are multicomponent fluorescence lifetimes that may be ascribed to self-quenching of porphyrin with different conformation or arrangements in the nanoparticles, but only the mean value is reported in Table 15.

Finally, no transient absorption at 450 nm was observed at all timepoints. Since aggregated porphyrins have many ways to dissipate energy, if there is an alternative path to the one leading to the triplet state, that is usually preferred. These photophysical properties are consistent with the formation of porphyrin aggregates.

*Table 15. Photophysical characterisation of conjugates **143** and **150** in assembly conditions. All analysis were performed on oxygenated samples. t1= after dilution, t2= after 24 h at 5 °C, t3= after one week at 5 °C, t4= after t3 and 3 more days at room temperature.*

	Absorbance	Fluorescence	TCSPC	LFP
143	t1 – 412, 509, 585 nm	Ex: 509	Ex: 500 nm	Abs: 450 nm
	t2 – 414, 509, 585 nm	t1 – 654, 709 (2×10^5) nm	Em: 700 nm	Not detected
	t3 – 415, 509, 586 nm	t2 – 654, 708 (1×10^5) nm	t1 – 5.8 ns	
	t4 – 420, 510, 586 nm	t3 – 654 (6×10^4), 708 nm	t2 – 4.0 ns	
	Intensity of Soret:	t4 – 655 (3×10^4) nm	t3 – 3.3 ns	
	0.65 → 0.44 → 0.2 → 0.14		t4 – 2.8 ns	
150	t1 – 328, 396, 446, 513, 548, 583, 635 nm	Ex: 513	Ex: 500 nm	Abs: 450 nm
	t2 – 327, 397, 446, 513, 549, 583, 636 nm	t1 – 642 (6×10^4), 705 nm	Em: 642 nm	Not detected
	t3 – 326, 398, 448, 515, 549, 583, 637 nm	t2 – 642 (6×10^4), 705 nm	t1 – 3.6 ns	
	t4 – 324, 402, 447, 514, 549, 583, 636 nm	t3 – 642 (3×10^4), 705 nm	t2 – 2.8 ns	
	Intensity of Soret:	t4 – 642 (7×10^3) nm	t3 – 1.8 ns	
	0.25 → 0.22 → 0.19 → 0.2		t4 – 2.4 ns	

As expected, no significant excitation anisotropy signals were recorded for DMSO control solutions. When in assembly conditions, only a weak signal of fluorescence anisotropy was detected for both **143** and **150** at all timepoints. Theoretically, the observed nanoparticles should have slowed down the porphyrin rotational diffusion and give defined anisotropy values. However, the anisotropy analysis was not informative and the signal weakness could be due to the quenched fluorescence emission of assembled porphyrin or the small particle size.

3.2.9 Metal derivatives of bolaamphiphilic conjugate 150

The bolaamphiphilic metal derivatives **146** and **147** (Figure 74) were also characterized with the above-mentioned spectroscopy techniques.
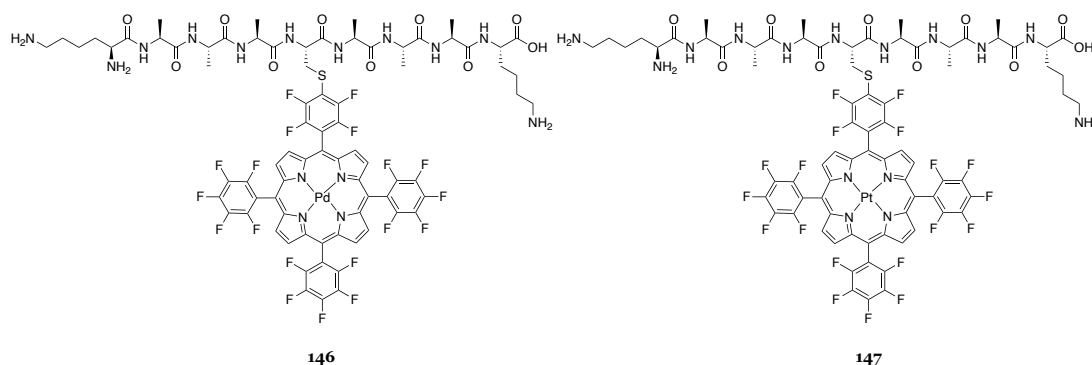
*Figure 74. Chemical structure of conjugates **146** and **147**.*

In DMSO solution, the UV/vis spectra of conjugates **146** and **147** was not affected over time, confirming that the assembly was not taking place under those conditions. However, the fluorescence emission was almost completely quenched and not sensitive to the presence of oxygen (*i.e.* no variations after deoxygenation). Triplet lifetimes were also surprisingly short (almost below the instrument detection limit), with and without oxygen. This means that the energy decayed with a mechanism which is not related to the presence of oxygen (*i.e.* not via triplet formation).

Further experiments proved that this behaviour was not related to the conjugate concentration but rather to the solvent employed. If the porphyrins somehow interact in DMSO, these data could not be considered as starting point for non-assembled conditions. Therefore, the solvent for these analyses was changed to methanol, in which Pd/Pt porphyrins **144** and **145** and related conjugates **146** and **147** are more soluble if compared to free base porphyrin **121**. The analysis reported in Table 16 were performed on 5 μM solution of **146** and **147** in MeOH + 1% v/v DMSO.

UV/vis spectra were consistent with the properties of the starting porphyrins **144** and **145**.[212–215] For these metal derivatives, the absorption of the Soret band was more intense than free base conjugate at the same concentration.

*Table 16. Photophysical characterisation of conjugates **146** and **147** in MeOH solution. All analysis were performed on oxygenated samples unless indicated by "N_2 – ".*

	Absorbance	**Fluorescence**	**TCSPC**	**LFP**
146	403, 517, 550 nm Intensity of Soret: 1.1	Ex: 517 nm 673 (5×10^2) nm N_2 – 674 (2×10^4), 740 nm	Ex: 500 nm Em: 675 nm 6.8 ns N_2 – 11.1 ns	Abs: 450 nm 300 ns N_2 – 19 μs
147	389, 505, 538 nm Intensity of Soret: 1.2	Ex: 539 nm 651 (2×10^4), 714 nm N_2 – 652 (1×10^5), 717 nm	Ex: 500 nm Em: 650 nm 34.2 ns N_2 – 21.3 ns	Abs: 450 nm 350 ns N_2 – 3 μs

Fluorescence emission in oxygenated solvent was extremely quenched and barely observable for palladium derivative **146**, while it was quenched but detectable for the platinum analogue **147**. Deoxygenation with N_2 flow for 10-15 minutes increase their emission of two orders and one order of magnitude respectively.

Fluorescence lifetime measurements were affected by the strong phosphorescence of these species. The phosphorescence was limited in oxygenated solvent, but this emission improved and the interference was greater when oxygen was removed. Phosphorescence emission, which is longer than fluorescence emission, gave a long-lasting background signal over all the time frame of fluorescence decay (100 ns). This signal interfered with the fitting and therefore the final lifetime value.

In oxygenated solvent, triplet excited states were very short and barely observable with transient absorption analysis. Only a small signal that quickly decayed could be detected. However, when the oxygen was removed, the triplet lifetime became much longer.

Published experimental data regarding **144** described long lifetime values for fluorescence and triplet state, both in the order of hundreds of μs, in the absence of oxygen at room temperature.[214] For **145**, reported lifetimes are shorter and only reached tenths of μs.[213] The decay of these long-lived excited states are usually further extended by immobilization on solid supports, such as silica or polystyrene, while at atmospheric oxygen concentration are orders of magnitude shorter.

Assembly of metal conjugates **146** and **147** was not studied using the full factorial design but only under a single assembly condition. DMSO stock solutions of conjugates **146** and **147** were diluted with PB, pH=7.4, IS=10 mM to a final conjugate concentration of 10 μM, and the solutions were analysed with spectroscopic (Table 17) and microscopy techniques.

Absorption scan measurements showed an intensity decrease of the Soret band as previously observed for the free base porphyrin. Unexpectedly, and differently from conjugate **143**, no wavelength shift occurred during incubation at 5° C, but almost 10 nm shifts were observed for both **146** and **150** after incubation at room temperature.

*Table 17. Photophysical characterisation of conjugates **146** and **147** in assembly conditions. All analysis were performed on oxygenated samples unless indicated by "N_2-". t1= after dilution, t2= after 24 h at 5 °C, t3= after one week at 5 °C, t4= after t3 and 3 more days at room temperature. *analysed but problems with fitting*

	Absorbance	Fluorescence	TCSPC	LFP
146	t1 – 402, 521, 555 nm	Ex: 521 nm	Ex: 500 nm	Abs: 450 nm
	t2 – 403, 521, 555 nm	t1 – 675, 800 (10^3) nm	Em: 675 nm	t1 – 20 ns, 1.9 µs, 150 µs
	t3 – 404, 521, 554 nm	t2 – 675, 800 (10^3) nm	t1 – 3.9 ns	t2 – 500 ns
	t4 – 411, 523, 557 nm	t3 – 676, 800 ($2x10^3$) nm	t2 – 4.4 ns	t3 – no
	Intensity of Soret:	t4 – too low	t3 – 4.8 ns	t4 – Not analysed
	0.85 → 0.65 → 0.35 → 0.25	N_2 – no changes	t4*	
147	t1 – 388, 508, 542 nm	Ex: 508 nm	Ex: 500 nm	Abs: 450 nm
	t2 – 388, 509, 542 nm	t1 – 656 ($4x10^4$), 742 nm	Em: 656 nm	t1 – <10 ns, 2 µs, 60 µs
	t3 – 388, 509, 542 nm	t2 – 655 ($2x10^4$), 742 nm	t1 – 27.9 ns	t2 – 16 ns, 3/6 µs, 60 µs
	t4 – 397, 512, 545 nm	t3 – 655 ($3x10^4$), 741 nm	t2 – 30.2 ns	t3 – no
	Intensity of Soret:	t4 – 654 ($1.5x10^4$) nm	t3 – 20.6 ns	t4 – Not analysed
	1.2 → 1 → 0.85 → 0.2	N_2 – no changes	t4*	

In oxygenated conditions, fluorescence emission intensities had small variations because fluorescence was significantly quenched. When the solvent was deoxygenated, the emission had minimal variation compared to the previous conditions. Therefore, differently to MeOH solution, the conjugates lost the response to oxygen.

Fluorescence lifetime was again difficult to measure precisely due to the strong phosphorescence interference. The longer phosphorescence lifetime generated a high and constant background signal that masked the fluorescence emission, making it apparently longer than it probably was.

When measuring transient absorptions, a slightly lower assembly rate for metal conjugates allowed us to detect triplets at t1 and t2 for both **146** and **147**, which were ascribed to the unassembled species. This may be related to the different solubility compared to free base, because metal derivatives are more water soluble, and the assembly could have been consequently slower. The observed triplets had three components: a first transient which was fast (a few ns) and below instrument resolution (Ultrafast Time-Resolved Laser Spectroscopy are required for picosecond species), an intermediate transient lasting a few µs, and a particularly long third transient of tenths of µs. However, these three transients progressively disappeared when the assembly process proceeded (from t3 sample onwards).

If there is no observable triplet after prolonged incubation, these systems cannot be used as photosensitisers for PDT nor for oxygen sensing. In other words, if the triplet is not involved in the energy decay, there is no possibility that singlet oxygen is generated (which, for example, a PDT requirement) or emission properties change as a response of different oxygen concentration in the solution (which is, for example, an oxygen sensor requirement).

3.2.10   Photoacoustic analysis

Fluorescence emission was quenched in most of the samples – probably because of transitions such as singlet-singlet annihilation or electron and energy transfer between the aromatic moieties – and triplet states were not observed. Therefore, energy must not have been dissipated via intersystem crossing to the triplet state but rather via alternative paths, probably as heat release or other internal conversions. On this basis, self-assembled species could be a useful tool for photothermal therapy and photoacoustic imaging.[216,217]

Photoacoustic spectroscopy is a technique that measures the sound wave generated in a sample after laser excitation. This acoustic signal is created by the thermal expansion derived from the release of absorbed energy as heat by the species in the sample. After comparison with a reference compound, which releases 100% of the energy absorbed as heat, the percentage of absorbed energy released as heat by the sample under investigation can be calculated.[218] Photothermal therapy or photoacoustic imaging are medical application of photosensitizers that exploits this biophysical principle.

Photoacoustic analyses were performed on a 100 μM solution of conjugate **143** in PB, pH = 7.4, IS = 10 mM, after incubation at room temperature for one week. Because of the instrument settings and the laser employed, an absorption value of around 0.15 upon excitation at 532 nm was required and determined the high conjugate concentration. Right after dilution, the absorption at this wavelength for the sample prepared as above was 0.156 (dotted line in Figure 75A). A solution of Brilliant Black BN, functioning as reference compound, was prepared in the same solvent to have a comparable absorbance (0.164).[219]

However, after one week of incubation, the absorption at 532 nm of the same solution of conjugate **143** almost doubled intensity (0.2991) due to a pronounced scattering effect (full line in Figure 75A). The scattering could be a consequence of nanoparticles aggregation, probably promoted by electrostatic attraction between deprotonated C-terminal acid and protonated amines. Although this phenomenon was already marginally observed with some conjugate **143** solutions (less with samples from conjugates **146**, **147**, and **150**), it was particularly emphasised in this experiment by the higher conjugate concentration.

Scattering was problematic because the AI value at 532 nm directly affected processing of photoacoustic analysis data. A function describing the scattering intensity as inversely proportional to the wavelength provided a correction factor. This correction factor was used to obtain a correct absorption value (0.130) which is closer to the actual absorption at 432 nm (Figure 75B).

The excitation of a porphyrin can sometimes be associated with a change of molecular structural volume, possibly related to an interaction of pyrrolic nitrogen with water.[220] This event would influence the generated sound wave and consequently the photoacoustic results. Therefore, photoacoustic measurements were repeated at different temperatures (20 °C, 15° C, 10 °C, 7 °C and 5 °C) to evaluate the presence the presence and take into consideration the contribution of volume variations. While decreasing the temperature, the eventual volume

variation will be reduced and its influence on the analysis will decrease. Photoacoustic measurements at different temperature are reported in Figure 76.
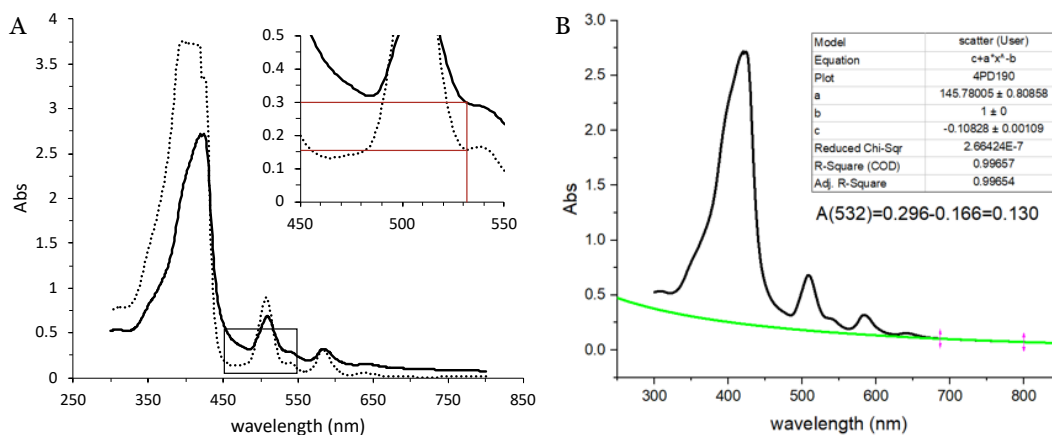
Figure 75. A) Absorption spectrum of a 100 μM solution of conjugate 143 in PB, pH = 7.4, IS = 10 mM, before (dotted line) and after (full line) incubation; B) function describing the scattering effect (in green) on the spectrum of the sample prepared for photoacoustic analysis after incubation.

After absorbance correction and results elaboration, it was estimated that **143** assemblies released between 70 and 80% of the energy absorbed as heat. To complete the energy balance (20/30% to 100%), the remaining energy must be released or stored via alternative processes.

Considering the consistent fluorescence quenching, the influence of fluorescence emission would be just a few percentage points and not relevant for the overall energy balance. Intersystem crossing is another electronic process to be excluded from this balance because no triplets were observed from laser flash photolysis measurements. A significant change of volume was also excluded after performing the measurements at different temperatures. One of the remaining options is that the energy was chemically accumulated in the system, but further studies are necessary to completely elucidate this process.
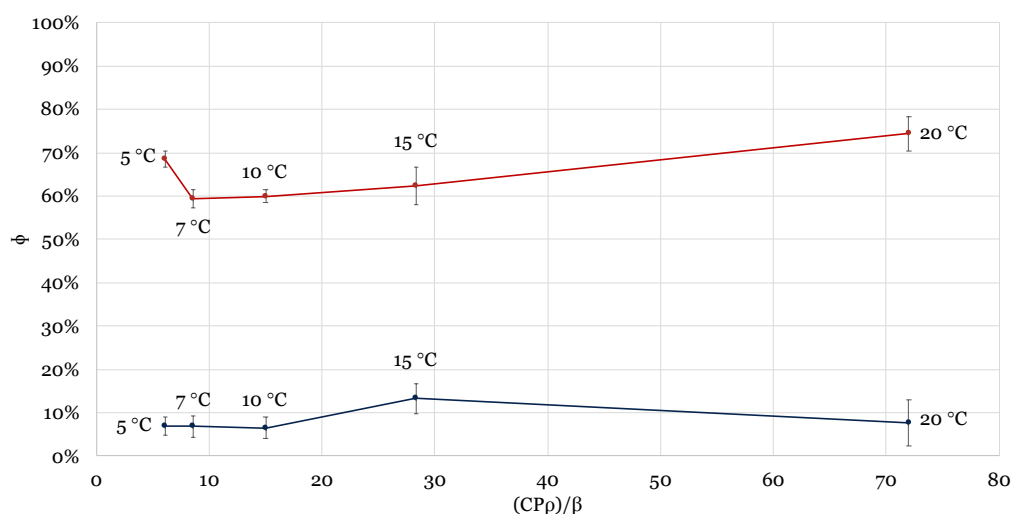
Figure 76. Results of photoacoustic analysis on 100 mM solution of 143 in PB, pH = 7.4, IS = 10 mM, after incubation at room temperature for one week. Red and blue lines correspond to different transient.

3.2.11   Hybrids assemblies

Only a few research works have explored the transient absorption of porphyrin embedded in self-assembled systems with peptides. In the co-assembled fibres designed by Fry and co-workers, the triplet was not observed when the chromophore and the peptide were used in a 1:6 ratio.[200] However, by increasing the amount of peptide (up to 1:60), the spatial separation between neighbouring porphyrin was amplified, intersystem crossing was facilitated, and formation of the triplet excited species was ultimately allowed.

Outdistancing porphyrins to prevent stacking and maintain their properties could be a useful approach to tune the properties of the self-assembled system. To achieve this, the simultaneous assembly of peptides and conjugates into the same architecture has to be induced.

When these hybrid systems co-assemble, a higher number of peptides should build the main structure, while the less abundant peptide-porphyrin conjugate would fit within other peptides. With the co-assembly process, light sensitive materials could be obtained with the advantage of using a lower amount of conjugate, which is more expensive to prepare compared to the peptide.

Before proceeding with the experiments on hybrid systems, the self-assembly of the original peptide KAAAAAAK (**141**) and the modified peptide KAAACAAAK (**142**) were individually studied with TEM and DLS (Figure 77).
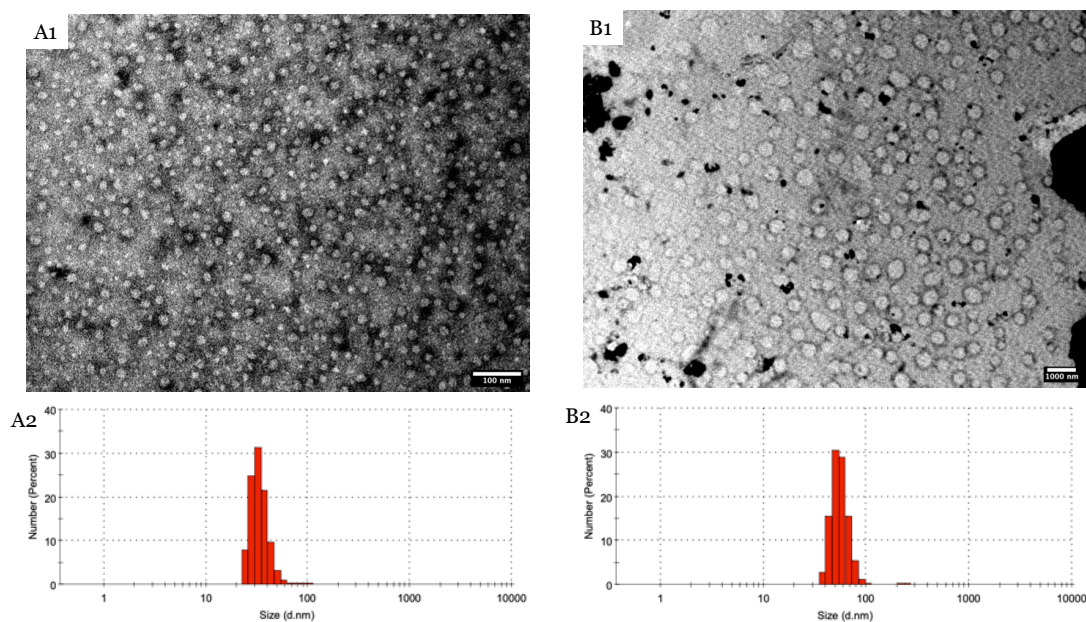
*Figure 77. A) TEM image (1) and DLS (2) of a sample from peptide **141** self-assembly in PB, pH = 7.4, IS = 10 mM after 24h of incubation; B) TEM (1) image and DLS (2) of sample from peptide **142** self-assembly in PB, pH = 7.4, IS = 10 mM after 24h of incubation.*

Samples were prepared reproducing conditions previously applied for conjugates: a peptide stock solution in DMSO was diluted with PB, pH = 7.4, IS = 10 mM to have 50 μM of peptide final concentration and 1% (v/v) of DMSO. After 24 h of incubation, regular spheres were

observed for both peptides (Figure 77A1 and 77B1). These structures appeared white on a darker background due to staining with phosphotungstic acid.

The nanoarchitectures differs from the fibres obtained from the self-assembly of peptide **141** originally described by Qiu *et al.*,[189] probably because of the differences in experimental procedures and sample preparation. Moreover, spheres formed from peptide **142** are bigger and less regular than those from peptide **141** (hundreds of nm vs tens of nm). Microscopy observations were confirmed by DLS analysis for both samples (Figure 77A2 and 77BS).

After understanding the self-assembly behaviour of the individual peptides, mixed solutions of peptide **141** and conjugates **143** or **150** with different ratios and concentrations were prepared from DMSO stock solutions. For instance, analysed peptide/conjugate ratios were 100:1, 10:1 and 1:1. These mixed stock solutions were then diluted with the aqueous phase to induce the co-assembly and have final concentrations of 100 μM or 10 μM of peptide and 10 μM or 1 μM of conjugate.

It is challenging to prove whether peptides and conjugates are co-assembling to form a structure together, or self-assembling independently to form different structures. Theoretically, if the co-assembly was not taking place, it would be foreseeable to spot the same structures observed for conjugate **143** self-assembly when using TEM grids without stain.

Meanwhile, after staining, the self-assembled structures of peptide should also be visible as white nanoparticles as reported in Figure 77. In practice, when peptide **141** and conjugate **143** were diluted with PB, pH= 7.4, IS= 10 mM to a final concentration of 100 μM and 1 μM respectively (ratio 100:1), spheres were the widespread structures observed but these had a different morphology and size if compared to previous experiments.
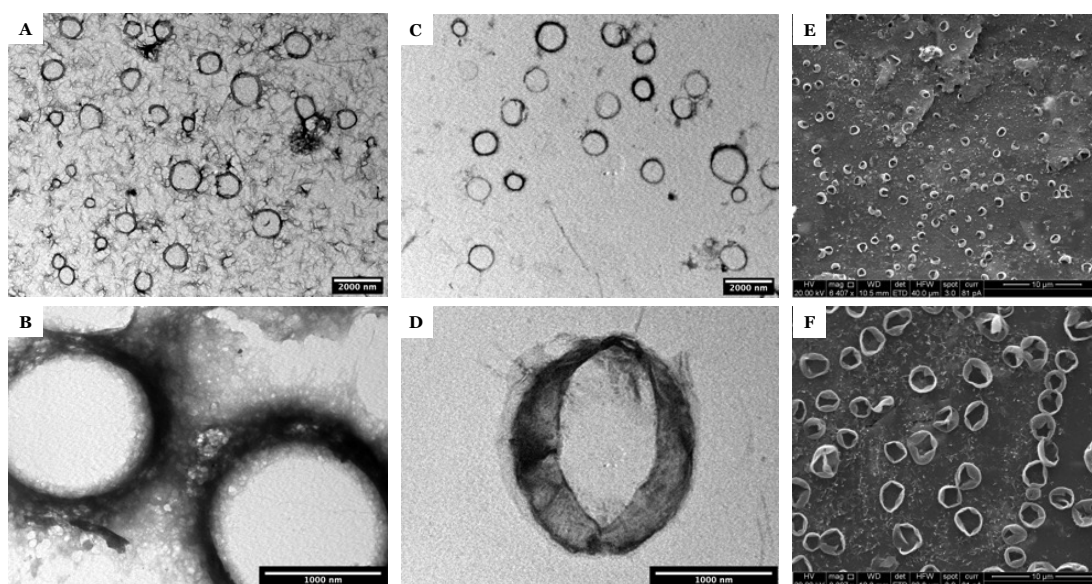
Figure 78. A-F) TEM image of peptide-conjugate hybrids assemblies; E-F) SEM image of peptide-conjugate hybrids assemblies.

Firstly, the spheres appeared on TEM grids as thin rings with a dark border and a clear (possibly empty) inside (Figures 78A-F). The border appeared dark before and after staining, indicating the porphyrin presence in that area. Some of these structures were surrounded by fibres, interconnecting different spheres. No white nanoparticles appeared after staining.

Secondly, the spheres had variable sizes which ranged from hundreds of nanometres to a few micrometres. This morphology is confirmed with SEM, even if spheres tended to be slightly bigger in size from this analysis. The main structures observed were microspheres, which seemed to be empty and blasted, probably because of solvent evaporation during sample preparation (Figure 78F).[221] Smaller spheres with the same ring morphology were observed as well (Figure 79A-F).
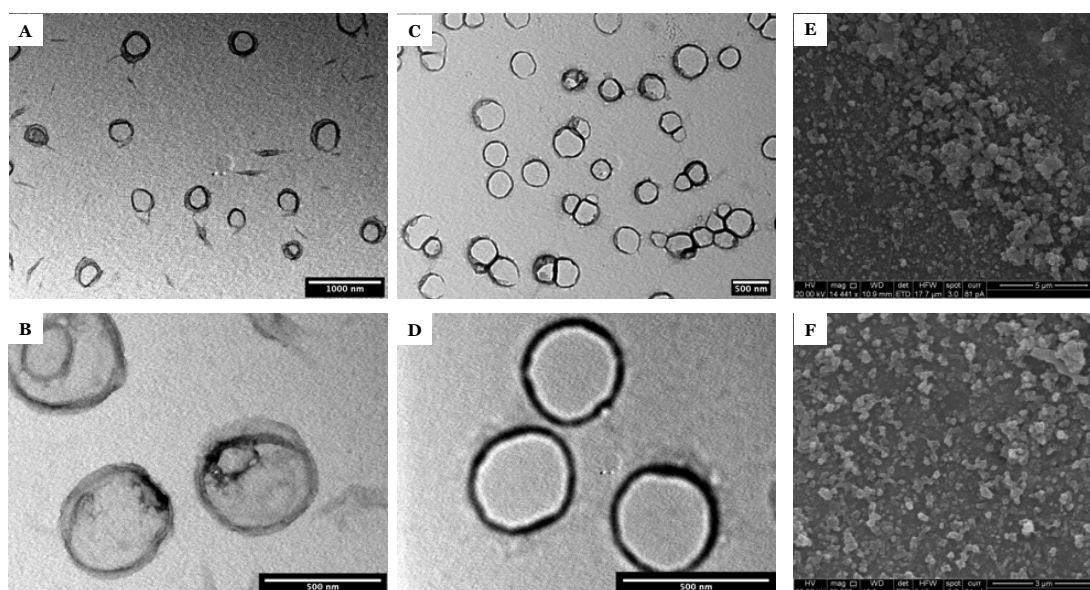
Figure 79. A-D) TEM image of peptide-conjugate hybrids assemblies; E-F) SEM image of peptide-conjugate hybrids assemblies.

The presence of only one monodisperse population of spheres with micrometric size was also a sign of potential co-assembly (Figure 80). Without co-assembly, two different populations of nanoparticles would be expected, one from peptide nanoparticles and one from conjugate nanoparticles.
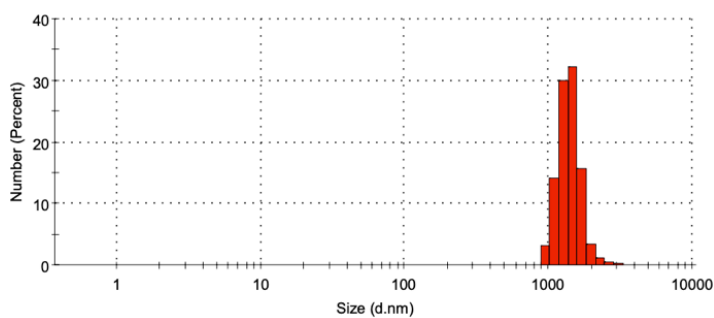
Figure 80. DLS of a sample from peptide 141 and conjugate 143 in PB, pH= 7.4, IS= 10 mM with a final concentration of 100 μM and 1 μM respectively (ratio 100:1).

At basic pH, no spheres were observed for peptide **141** and conjugate **143** co-assembly. Meanwhile, in this case, peptide nanoparticles (white) appeared after staining (Figure 81A).

Conjugate **150** co-assembly with peptide **141** in PB, pH = 7.4, IS = 10 mM to a final ratio of 100:1 (Figure 81B and 81C) resembled previous co-assembly results under the same conditions.
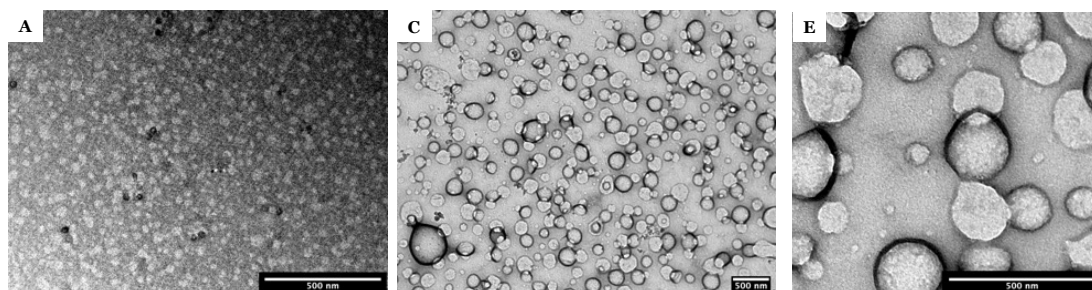
*Figure 81. TEM images of samples from A) peptide **141** and conjugate **143** in NaOH, pH= 12, IS= 10 mM with a final concentration of 100 μM and 1 μM respectively; B-C) peptide **141** and conjugate **150** in PB, pH= 7.4, IS= 10 mM with a final concentration of 100 μM and 1 μM respectively.*

Photophysical characterisation of these co-assembled systems was performed as previously described for the self-assembled systems made by conjugates only. The results are reported in Table 18, as well as details of the specific peptide/conjugate ratio employed in each experiment.

*Table 18. Photophysical characterisation of hybrids systems from peptide **141** and conjugates **143** in assembly conditions. All analysis were performed on oxygenated samples. t1= after dilution, t2= after 24 h at 5 °C, t3= after one week at 5 °C, t4= after t3 and 3 more days at room temperature.*

141/143	Absorbance	Fluorescence	TCSPC	LFP
100:1 μM	t1 – 414, 507, 585 nm t2 – 418, 508, 586 nm t3 – 419, 509, 587 nm t4 – 427, 508, 584 nm Intensity of Soret: 0.08 → 0.05 → 0.04 → 0.03	Ex: 414 nm t1 – 654, 710 (1×10^5) nm t2 – 653, 710 (8×10^4) nm t3 – 654, 710 (4×10^4) nm t4 – 654 (1×10^5), 708 nm	Ex: 500 nm Em: 654 nm t1 – 5.1 ns t2 – 5.3 ns t3 – 5.0 ns t4 – 4.9 ns	Ex: 535 nm Em: 450 nm Not detected
100:10 μM	t1 – 413, 508, 585 nm t2 – 413, 509, 586 nm t3 – 413, 509, 586 nm t4 – 430, 511, 586 nm Intensity of Soret: 0.6 → 0.44 → 0.35 → 0.15	Ex: 509 nm t1 – 654, 709 (1.5×10^5) nm t2 – 654 (8×10^4), 709 nm t3 – 654 (6.5×10^4), 710 nm t4 – 655 (2.5×10^4), 709 nm	Ex: 500 nm Em: 654 nm t1 – 4.5 ns t2 – 4.3 ns t3 – 4.1 ns t4 – 2 ns	Ex: 535 nm Em: 450 nm Not detected
10:10 μM	t1 – 413, 508, 586 nm t2 – 413, 508, 585 nm t3 – 413, 509, 585 nm t4 – 425, 510, 586 nm Intensity of Soret: 0.65 → 0.46 → 0.33 → 0.12	Ex: 509 nm t1 – 654, 709 (1.4×10^5) nm t2 – 654 (6×10^4), 709 nm t3 – 654 (6×10^4), 709 nm t4 – 655 (2×10^4), 709 nm	Ex: 500 nm Em: 654 nm t1 – 4.5 ns t2 – 3.9 ns t3 – 3.9 ns t4 – 2.1 ns	Ex: 535 nm Em: 450 nm Not detected

Results were generally aligned to those obtained for **143** assemblies (Table 15) and there was not much difference when varying the peptide/porphyrin ratios. The Soret band was considerably red shifted (up to 430 nm), fluorescence intensities were quenched, lifetimes shortened, and no triplets were observed. The chromophores did not preserve their properties indicating that the porphyrin rings are not outdistanced within the systems, or the co-assembly did not take place at all.

3.2.12   Fluorescence Microscopy

Fluorescence microscopy is frequently used for biological applications due to the advantages over light and transmission microscopy. After labelling the cellular and subcellular structures of interest with fluorescent probes, these become visible as bright points on a dark background, emitting only at certain wavelength ranges. Consequently, an image with significantly improved contrast is generated and additional information from differently emitting labels (hence with distinct colour) can be obtained.

Compared to fluorescence microscopy, TEM has the advantage of a higher magnification, but most of the time this technique lacks the contrast between background and samples (from here the necessity of staining reagents as described) as well as the possibility to differentiate between different structures within the sample (black and white only). The apex of fluorescence microscopy development was reached with the introduction of high-resolution techniques, where the emission of a single molecule can be detected under specific conditions.

The design of fluorophores, appropriate tuning of their photophysical properties, and their selective bioconjugation to target molecules are fundamental aspects for high-resolution imaging.[222] Fluorescence microscopy is not only useful for cell imaging but also for the study of nanoarchitectures, and its combination with TEM is a very powerful approach.[223,224]

In a STochastic Optical Reconstruction Microscopy (STORM) microscope, the laser can be employed in Total Internal Reflection Fluorescence (TIRF) mode and adjusted to hit the sample carrier glass at a critical angle. In TIRF mode, only fluorophores within the first 150–250 nm above the glass sample interface are exited, avoiding most of the background fluorescence, and this setting is particularly good for studying nanoparticles which are attached on the glass surface.

In a preliminary experiment, performed with a STORM microscope in TIRF mode, a sample from the self-assembly of 1 μM solution of conjugate **143** in PB, pH = 7.4, IS = 10 mM was studied. Fluorescent supramolecular structures (probably nanosphere) were successfully observed (Figure 82), but their exact size and morphology cannot be defined because these structures are below instrument resolution and therefore smaller than a pixel size (350 nm).

These are promising results because fluorescence microscopy techniques can be practical, alone or in combination with other microscopy methods, for supporting the description of these systems in future works.
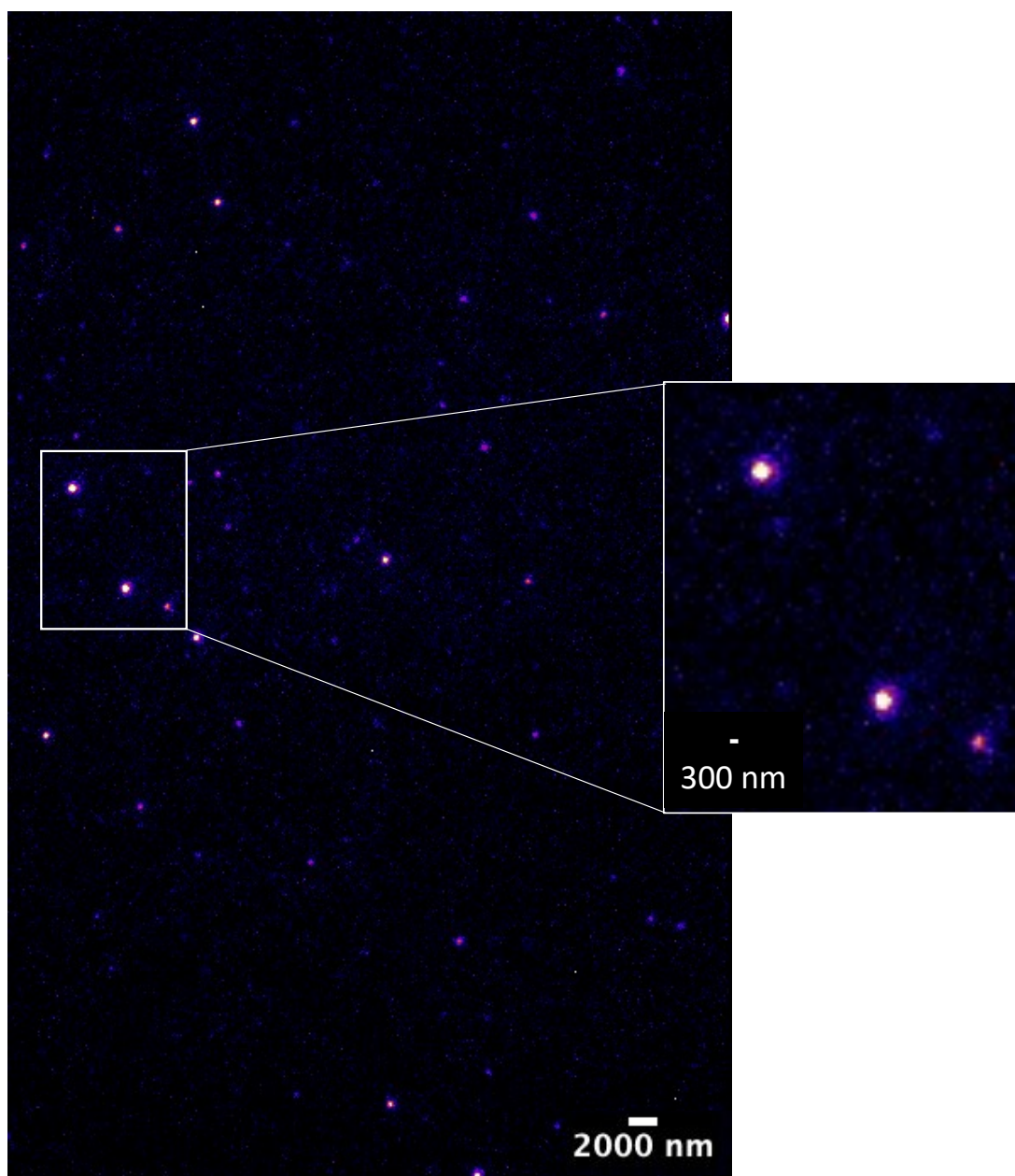
Figure 82. Fluorescence microscope image of the self-assembled nanoparticle of conjugate **143**.

3.3 CONCLUSION AND FUTURE WORK

In conclusion, the successful application of the thiol-fluoride S$_N$Ar for the synthesis of self-assembling peptide-porphyrin conjugates was presented in this chapter. The aim of these studies on self-assembled nanoarchitectures was to set the basis for further developments in this research field.

The assembly of these novel bioconjugates was studied under different conditions, conjugate concentrations and properties of the aqueous phase. Different spectroscopy (steady state and time resolved) as well as microscopy (TEM, SEM, fluorescence) techniques were used and nanospheres were the most common structures observed. Overall, following the analysis of the results, two of the many possible arrangements of the conjugates within the nanoparticles were hypothesised and represented in Figure 83.
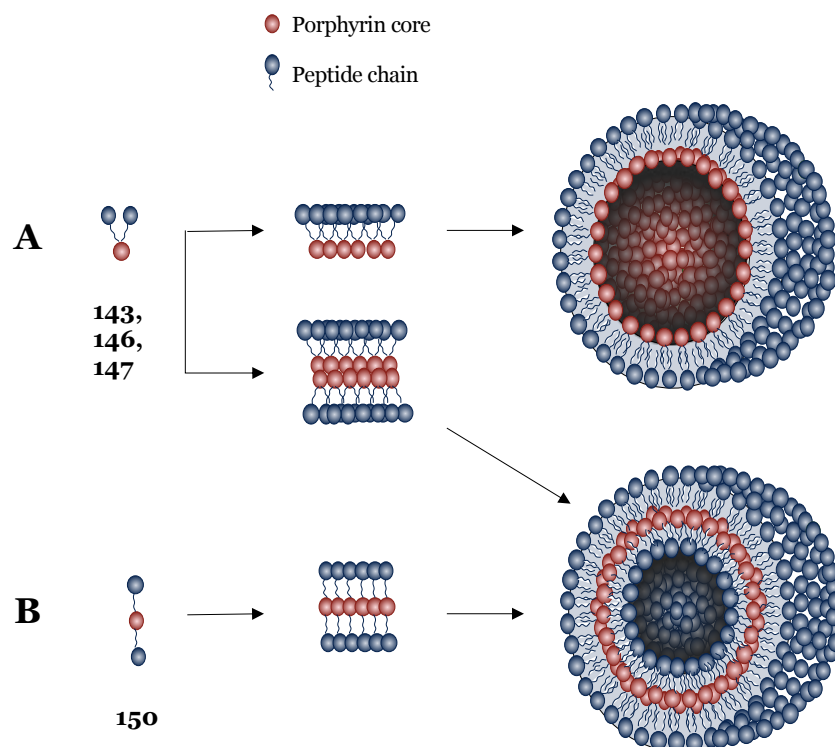
Figure 83. Hypothesis of nanoparticle architectures.

To conclude the study of conjugates **143, 146, 147, 150** self-assembling, time-resolved experiments using the other conditions included in the DOE, such as those in basic pH or with higher IS, could be performed. As future work, completely different conditions of self-assembly, such as acidic pH (*e.g.* 3.8), can be eventually included in the study as well.

Circular dichroism analyses are undergoing with collaborators, and these are usually very informative when describing these systems.

Photoacoustic analyses and fluorescence microscopy were only preliminarily executed but could be notably improved and informative. After a complete characterisation, the stability of the nanoparticles should also be evaluated, especially when the conjugates assemble in strong basic conditions or when incubated at higher temperatures.

After completing the studies of supramolecular systems obtained from conjugates **143, 146, 147, 150**, changes to the peptide and conjugate design should be considered. Small changes to the peptide sequence can be initially planned, for example introducing an amide C-terminus to prevent NH_4^+ and COO^- aggregation. Subsequently, bigger variations, such as peptides with acidic heads or with completely different bolaamphiphilic sequences, could be explored to expand the plethora of self-assembling systems.

Finally, the co-assembly of hybrids systems also requires more studies and, due to the lack of sufficient information, arrangements were not hypothesised for these species. It will be important to confirm if there was co-assembly between peptides and conjugates, and test different peptide/conjugate ratios under different assembly conditions.

4 A new scaffold for peptide cyclisation

In the final part of this research work, the complexity of the peptide-porphyrin bioconjugation strategy was further increased. The thiol-fluoride S_NAr reaction, performed on porphyrin **121** with peptides containing multiple cysteine residues, is presented in this chapter. Exploiting this method, the tetrapyrrole ring was explored as photoactive scaffold for peptide cyclisation.

A variety of fully-deprotected di-cysteine containing peptides with different inter-thiol distances were cyclised using porphyrin **121**, in solution and under peptide-compatible conditions, demonstrating the generality, chemo-selectivity, and wide applicability of this method. A similar methodology was applied to porphyrin **148** and, combining the results from reactions with porphyrin **121**, important insights regarding the geometrical constraints and regioselectivity of the cyclisation were obtained. The remaining para-positions of macrocyclic conjugates are available for further functionalisation: these could be exploited for thiol conjugation or two-step multicyclisation of peptides. Similarly, with appropriate cysteine-rich amino acid sequences, one-step peptide multicyclisation was also possible.

The SPACE peptide (**77**), previously employed for HFB-templated macrocyclisation (Section 2.3.3), was cyclised with porphyrin **121** and tested for cytotoxicity and cellular uptake on squamous cell carcinoma cells (SCCIC8). The resulting water-soluble conjugate retained cell penetration properties and the porphyrin retained intrinsic fluorescence, without influencing cell viability. The possibility to 'dial in' the peptide sequence and exploit the photoactive tetrapyrrolic scaffold for multiple functionalisation, metal coordination and light-responsive properties opens the way to a wide range of applications in biological systems.
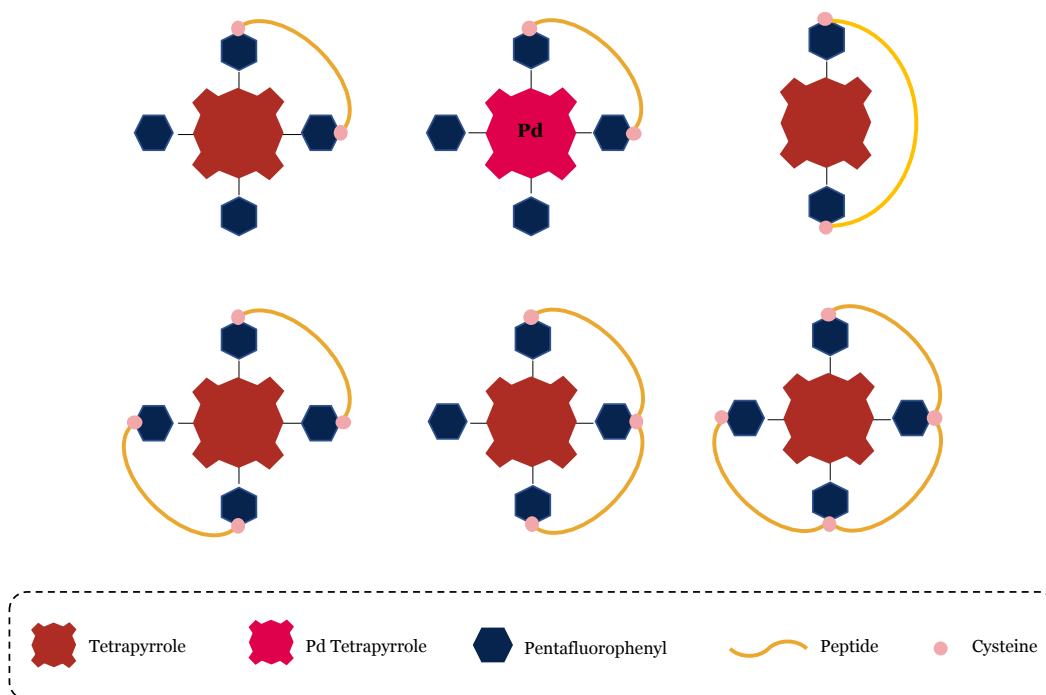
Figure 84. Graphical representations of the conjugates obtained from cysteine-rich peptides and polyfluorinated porphyrins investigated in the final part of the work.

4.1 PEPTIDE CYCLISATION WITH FUNCTIONAL SCAFFOLDS

In Chapter 1, the increasing interest in cyclic peptides was described. This is reflected in the constant development of chemical strategies for peptide macro and multicyclisation.[37–40] Most commonly, peptides are cyclised using one-component strategies: this chemistry involves only the peptide and exploits the reaction between functional groups that are part of the same AA sequence. Carboxylic acids (Asp, Glu, C-terminus), amines (Lys, N-terminus), thiols (Cys) or non-natural AAs designed to include reactive moieties (olefin, azide, alkyne) are key players in these methods.

When a two-component strategy is employed, this usually combines the peptide and a small molecule, such as HFB or the other examples previously presented.[51] These non-peptide scaffolds are commercially available or readily accessible via chemical synthesis and have a good reactivity towards AA side chains. Ultimately, their role is purely structural, because they only provide the support or linkage on to which the peptide is assembled.

Recent studies presented the integration of a functional scaffold within a peptide cyclic architecture as an enrichment in the development of novel peptide-based multifunctional molecules. A functional scaffold is a molecule that has an intrinsic functionality, which is exploitable after the construction of the peptide architecture. Example of functional scaffold which were reported as side chain cross-linkers for peptide cyclisation are: i) intrinsically fluorescent species such as Ir(III) complexes (**151**),[225,226] isoindole (**152**),[227] and boron dipyrromethene (**153,** BODIPY);[228] ii) photoresponsive molecules such as 3,3'-bis(sulfonato)-4,4'-bis(acetamido)azobenzene (**154**),[229] and tetrazine (**155**);[230] iii) catalytically active substrate such as His-Pd(II) complex (**156**).[231] Figure 85 shows how these molecules are included in the peptide architecture.
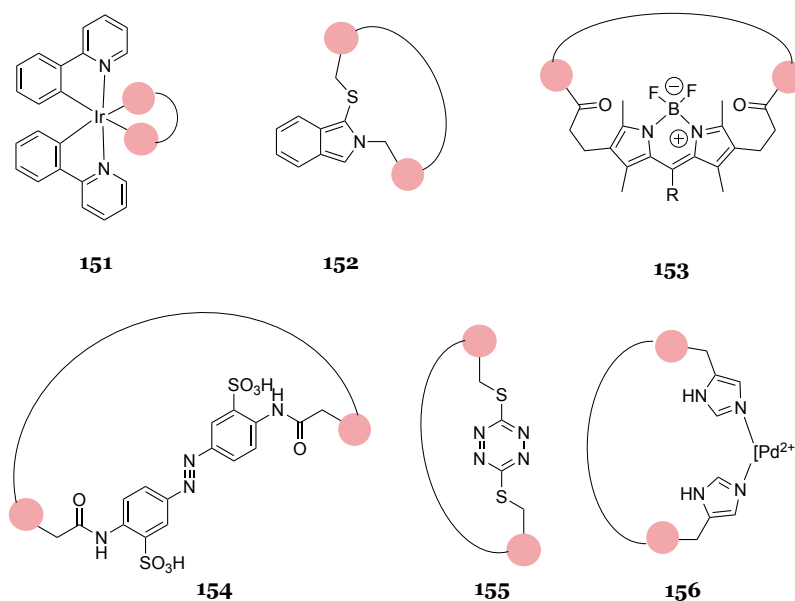
*Figure 85. Chemical structure of peptides cyclised with Ir(III) complex (**151**), isoindole (**152**), boron dipyrromethene (**153**), 3,3'-bis(sulfonato)-4,4'-bis(acetamido)azobenzene (**154**), tetrazine B (**155**), and His-Pd(II) complex (**156**).*

Polyfluorinated porphyrins are accessible via chemical synthesis and have a reactivity towards nucleophiles via S_NAr under mild conditions, but they also have light-responsive properties ascribable to the tetrapyrrole ring. Therefore, it was hypothesised that porphyrin **121** could be explored as novel, advanced, and functional scaffold for templating peptide macro and multicyclisation.

4.2 STRAPPED PORPHYRINS

In complex porphyrin systems such as heme-containing enzymes, the three-dimensional architecture and microenvironment surrounding the porphyrin ring are pivotal for the stability and reactivity of the construct and can influence its photophysical and photochemical properties. The self-assembling systems described earlier are confirmation of this general concept but, if the correct organization of the chromophores cannot be achieved with their spontaneous arrangement, this can be obtained by chemically engineering the molecule.

In this context, Lehn[232] and Sauvage[233] were the pioneers in the synthesis of porphyrins derivatised with chains spanning over the tetrapyrrole ring, called 'strapped' porphyrins (Figure 86). Since then, a lot of effort has been dedicated to the development and applications of these systems, with the main purpose of resembling tetrapyrrolic cofactors within a protein architecture.[234]
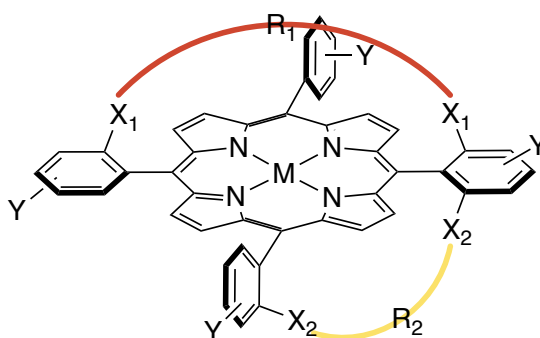
157

Figure 86. General structure of a strapped porphyrin.

Two different synthetic approaches can be used to access this peculiar tetrapyrrole architecture: post-strapping, when the porphyrin scaffold is synthesised first and then functionalised with a strap afterwards, or pre-strapping, when the strapped system is directly synthesised by using a pre-strapped aldehyde pair.[235] Some examples of reported porphyrin straps are summarised in Table 19.

As can be appreciated from these systems, the strap is usually made of aliphatic or aromatic carbon chains. Contextually, the porphyrin generally has asymmetrically substituted *meso* phenyl groups: this allows selective connection of *ortho* or *meta* positions when following a post-strapping procedure.[234] So far, strapped porphyrins were usually employed as heme-enzyme mimics and catalysts, or to build more complex macromolecular systems such as molecular cages.[236,237]

Table 19. Summary of some of the different porphyrin straps reported in literature. Refer to Figure 86 for $R_{1/2}$ and $X_{1/2}$. Because the focus is on the strapping chemistry, other substituents of meso-phenyl groups (Y) or coordinated metals (M) were not specified.

X1	R1	X2	R2	Ref
H	-	O	$(CH_2)_6$; CH_2-Ph-CH_2	[238]
H	-	Proline	$(CH_2)_2$; $(CH)_2$; Ph	[239]
H	-	O	$(CH_2)_{12}$	[240]
O	$(CH_2)_{12}$	H	-	[240]
H	-	SCO	$(CH_2)_{7/8/10}$; Ph; CH_2-Ph-CH_2	[235]
O	CH_2-Ph-CH_2	H	-	[241]
H	-	O	CH_2-Ph-CH_2	[241]
O	$(CH_2)_{11}$-S-S-$(CH_2)_{11}$	H	-	[242]

4.3 PORPHYRIN-TEMPLATED PEPTIDE CYCLISATION

On one hand, a peptide system has the spatial flexibility and chemical reactivity to be employed for the synthesis of strapped porphyrins. On the other hand, by implementing a porphyrin in a cyclic peptide architecture, the tetrapyrrole would act as scaffold but also as photoactive moiety. This could attractively combine the advantages of peptide cyclisation and porphyrin strapping, allowing access to a new chemical space.

Peptide-porphyrin cyclic architectures are interesting macromolecules that could have potential application as natural enzyme mimics, photo-functional systems, catalysts, sensors, probes, diagnostic agents, molecular cages and many more. A peptide can cooperate with the functionality of the porphyrin with his biological activity or by mimicking epitopes, reactive sites, or cavities of natural proteins.

However, despite the promising field and the advanced studies in both peptide and porphyrin chemistry, there are only two examples of porphyrin strapping made with a peptide backbone, and both were obtained with a post-strapping method.

Karpishin and co-workers reported the first example of peptide-porphyrin cyclic system (**158**). The authors cyclised a 22-residue α-helical peptide over the face of a zinc porphyrin by coordination of two histidines, at around 16 Å, with two copper ions (Figure 87).[243]

Six alanine residues separate the two inner histidine residues, creating a spacing between the two coordination sites at *i, i+7*. This construct was designed to have a conformationally stable peptide with AAs spanning above the porphyrin metal centre. The authors planned the manipulation of these key residues to enable the creation of a substrate-binding cavity that resemble native haem protein and enzymes for oxygenative catalysis.

158
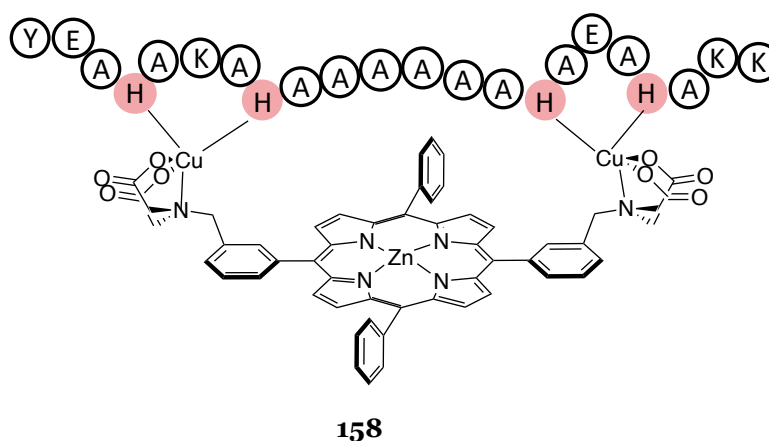
Figure 87. Chemical structure of the cyclic peptide synthesised by Karpishin and co-workers.

Shortly after, Geier and Sasaki obtained a peptide-porphyrin covalent conjugate (**159**) from the reaction of two cysteine residues of an α-helical peptide with two bromoacetamide linkers, previously inserted on meta positions of a manganese tetraphenylporphyrin (Figure 88).[244–246] Similarly to the previous example, this system mimics the cytochrome P450 monooxygenase family and was used as a catalyst to oxidize alkenes in the presence of iodosylbenzene and under different reaction conditions.

The spacing between the amino acids involved in the cyclisation is *i, i+11*, with the hydrophobic side chains projected toward the porphyrin ring to form the oxidation substrate binding site. Most interestingly, the peptide influenced the outcome of the catalysis when compared to surface-bound Mn(III)tetraphenylporphyrin as control.

159
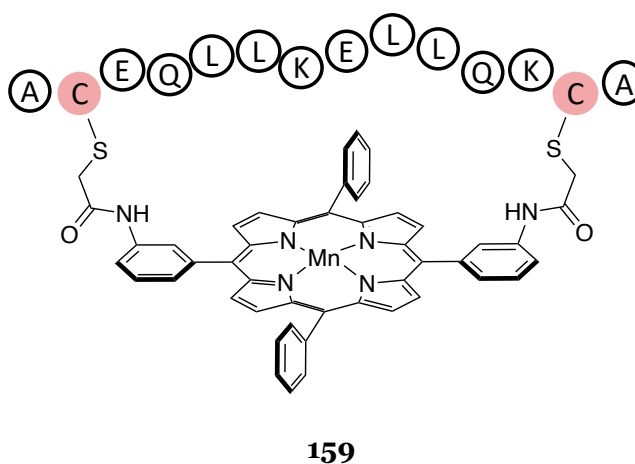
Figure 88. Chemical structure of the cyclic peptide synthesised by Geier and Sasaki.

Despite these promising studies, little is known about the compatibility between porphyrin and peptides with different lengths, different AA sequences and different properties. Moreover, frequently employed A_2B_2 asymmetric porphyrins are usually harder to isolate if compared to tetra symmetric porphyrin, whereas an easier access to the tetrapyrrolic scaffold would be an advantage for further applicability. Symmetric porphyrins may also allow cyclisation with different and more flexible geometries, albeit probably at the expenses of

selectivity. The explored bioconjugation reactions in the field of strapped porphyrins are also limited, and a broader range of available techniques could benefit the research in this area.

In this context, porphyrin **121** has four equivalent reactive sites for thiol-fluoride substitution and can conceptually act as a platform for peptide macro- and multicyclisation through reaction with peptides containing two or more cysteine residues.

4.3.1 Peptide cyclisation using porphyrin **121**

PDT is often used for the topical treatment of skin diseases. Its application is promoted by the non-invasiveness and good patient compliance but limited by poor penetration of photosensitizers into the skin (see section 1.2.4). Skin penetrating peptides can be chemically conjugated to other molecules and act as a cargo for transdermal drug delivery.[247]

As part of an effort to design skin-penetrating porphyrins for photodynamic applications, the Skin Penetrating and Cell Entering (SPACE) peptide (**77**)[166,248] was used for conjugation with porphyrin **121** under Cs_2CO_3/DMSO conditions.

SPACE peptide, previously employed for HFB-templated macrocyclisation, has good solubility in both water and DMSO, it is easily synthesised with single coupling SPPS, it has a known biological activity (skin cell-penetration) in its disulfide form and, most importantly, contains two cysteines located at *i, i+8* positions. Treatment of peptide **77** with **121** in Cs_2CO_3/DMSO at room temperature afforded a peptide-porphyrin cyclic product.

As shown in Figure 89, two potential cyclic products can be generated from the reaction between a peptide displaying two cysteines and porphyrin **121**. The product resulting from the displacement of the *para*-fluorides on the *meso*-aryl rings in positions 5 and 15 will be referred to as the 5-15 cyclisation product. Meanwhile, the one resulting from the same displacement reaction on the rings in 5-10 *meso* positions is defined as the 5-10 cyclisation product.

121
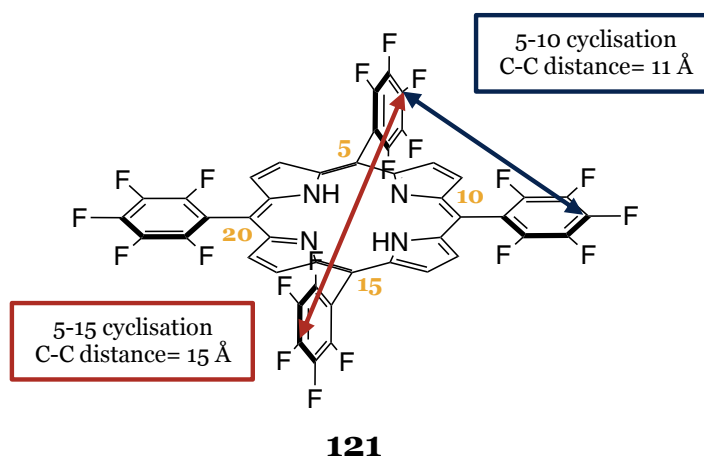
Figure 89. Graphical representation of 5-10 (in blue) and 5-15 (in red) cyclisation.

Computational modelling of porphyrin **121** using the semiempirical method XTB[249,250] were performed by Michele Assante (PhD student at LJMU). Distances of around 15 Å and 11 Å

between the 4-position carbon atoms of pentafluorophenyl groups in the 5-15 and 5-10 positions were respectively estimated.

Using the same approach, the inter-thiol distance within the SPACE peptide sequence was estimated to be approximately 23 Å, which was predicted to allow both 5-10 and 5-15 cyclisation to occur in line with previous studies from Karpishin and Geier.[243,244]

The studies described hereafter were planned and performed to identify the minimal inter-cysteine distance that would direct the cyclisation to occur at the 5-10 positions. A controlled 5-10 cyclisation would be preferable because allows discrete control of the regioselectivity during peptide stapling and multicyclisation, and this geometry may be achieved by selection of the appropriate distance between the thiols.

4.3.2 Cyclisation of a peptide library on porphyrin 121

Starting from the amino acid sequence of peptide **77**, a library of peptides displaying pair of cysteines at different distances was synthesised to explore the influence of thiol position on the reactivity with **121** and product formation.

By removing one AA at a time **77**, the *i, i+8* inter-thiol distance was sequentially decreased down to *i, i+1* (peptides **165-172**) to understand the optimal inter-thiol length for 5-10 regioselectivity. Similarly, by including additional alanine residues to the SPACE peptide, this distance was also increased up to *i, i+9* (peptide **164**) and *i, i+10* (peptide **163**) to verify eventual reactivity changes with longer peptides as well.

The SPACE peptide control (**173**) – corresponding to the scrambled AA sequence of peptide **77** – was included in this library as a control for biological purposes (see section 4.3.8). As the work developed, the library was further extended to contextually evaluate whether this reaction is sequence specific, selective against nucleophiles commonly present in natural AA, and compatible with peculiar residues such as proline.

The additional peptides were not derived from peptide **77** but completely different sequences were selected instead. Specifically, the cyclic region (residues 32–43) of the 53-mer human epidermal growth factor (hEGF, peptide **174**),[251] the apamin-derived peptide MiniAp-3 (peptide **175**)[252] and the M13 peptide, an analogue of the peptide 5-15 portan-10 (peptide **176**) were synthesised.[154,155]

Peptides **174-176** are bioactive sequences, and their cyclisation may be particularly interesting for potential biological applications as well. In particular, the cyclisation of brain-penetrating peptides such as MiniAp-3 and M13 with a polyfluorinated porphyrin could enhance their blood-brain barrier (BBB) penetration, in line with Pentelute's work on perfluoroarene-based peptide macrocycles.[154,155]

With further computational calculations on the whole peptide library, a cysteine spacing between *i, i+4* and *i, i+5* in unstructured peptides was estimated to be long enough to bridge the shorter distance of 11 Å for 5-10 cyclisation. All the peptide sequences and sulfur-sulfur distances are listed in Table 20.

*Table 20. Library of peptides employed for porphyrin-templated macrocyclisation experiments. Purity values by HPLC for each crude peptide are reported in section 6.9.2. Peptide **176** was purified.*

Peptide	Amino acid sequence	Cys-Cys distance	S-S distance (Å)
77	ACTGSTQHQCG-NH$_2$	*i, i+8*	22.59
163	ACATGSTQHQACG-NH$_2$	*i, i+10*	28.08
164	ACATGSTQHQCG-NH$_2$	*i, i+9*	26.27
165	ACTSTQHQCG-NH$_2$	*i, i+7*	21.29
166	ACSTQHQCG-NH$_2$	*i, i+6*	19.94
167	ACTQHQCG-NH$_2$	*i, i+5*	14.81
168	ACQHQCG-NH$_2$	*i, i+4*	14.67
169	ACHQCG-NH$_2$	*i, i+3*	10.31
170	ACACG-NH$_2$	*i, i+2*	8.23
171	ACHCG-NH$_2$	*i, i+2*	7.97
172	ACCG-NH$_2$	*i, i+1*	6.83
173	ACTHGQTQSCG-NH$_2$	*i, i+8*	24.49
174	NCVVGYIGERCQ-NH$_2$	*i, i+9*	25.27
175	CKAPETALC-NH$_2$	*i, i+8*	22.16
176	AGYLLGKINLKACAALAKKCL-NH$_2$	*i, i+7*	N.A.

Under the Cs$_2$CO$_3$/DMSO cyclisation conditions (described in detail in the next section), cyclic products (conjugates **163-176** in Scheme 27) were observed for all the peptides aside from the one with the shortest inter-thiol distance of *i, i+1* (peptide **172**).
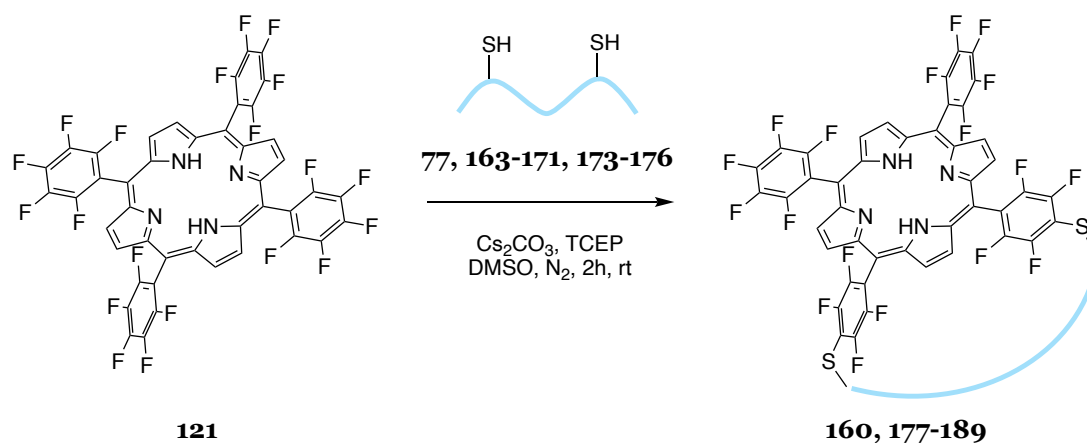
Scheme 27. General reaction scheme for the porphyrin-templated cyclisation of bis-cysteine peptides.

The proportion (percentage by HPLC peak area) of the monocyclic products relative to the population of the other peptide-porphyrin species formed within the reaction crude – including bicyclic products and porphyrin dimers identified by LC-MS, as well as other unidentified side products (Figure 90) – is reported in Table 21 as a measure of reaction outcome.

Monoconjugate   Monocyclic   Bicyclic   Porphyrin dimer
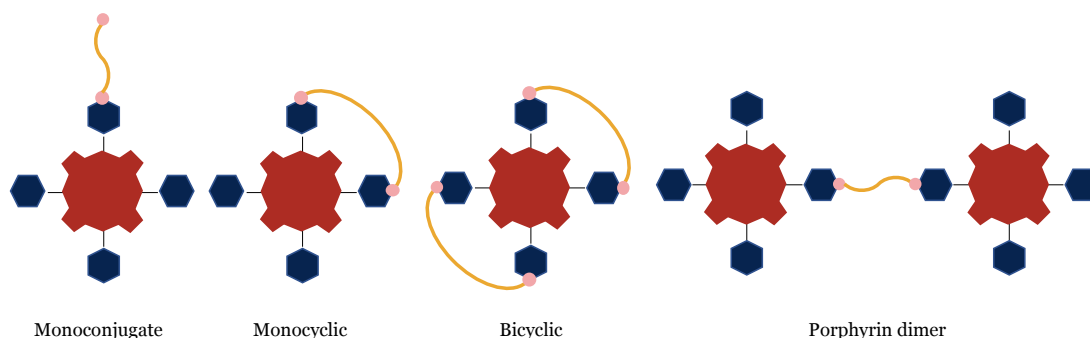
Figure 90. Graphical representation of peptide-porphyrin species identified from the reaction between 121 and peptides 77, 163-176.

Conversion to monocyclic species was particularly high for short and intermediate inter-thiol distances (*i.e.* from *i, i+8* to *i, i+2*), while it was slightly lower for longer distances (*e.g.* peptides **163**, **164**, **174**). For distances longer than *i, i+8*, cysteine residues were probably more easily accessible for intermolecular reaction, as confirmed by the higher percentage of porphyrin dimers. The exceptionally high conversion percentage of peptide **176** to monocyclic product **189** can be ascribed to the purified peptide employed as starting reagent (and not crude as for all other examples). In addition, when the inter-thiol distance was shorter, an increasing percentage of bicyclic species was observed, with an anticipated opposite trend for porphyrin dimers formation.

The formation of unidentified side products tended to be lower for inter-thiol distances between *i, i+5* and *i, i+3,* suggesting better overall reactivity. Different conjugation products (mono, di and tri) were observed for peptide **172**, having the shortest inter-thiol distance of *i, i+1*, and these were all included under the field "other".

Cyclic species arising from peptides with shorter inter-thiol distances are likely to exhibit a 5-10 geometry, because the thiols cannot spatially reach the 5-15 position, as supported by the computational calculations. In the case of peptides with a higher number of residues between thiols, their length could potentially allow the cyclisation with 5-15 geometry, but the 5-10 isomer may still be sterically and statistically favoured.

For all peptides, one main peak attributable to the 5-10 macrocyclisation product was observed by HPLC/LC-MS. However, for peptides with longer inter thiol distances such as peptides **77**, **164** and **165**, a second peak with negligible intensity but with same *m/z* ratio was also encountered and ascribed to the presumably less favoured 5-15 isomer of macrocyclisation.

Table 21. Peptides and corresponding conjugation products from the reaction with 121.

Peptide	Cys-Cys distance	Monocyclic (product ID)	Bicyclic	Dimer	Other
77	*i, i+8*	46% (**160**)	4%	10%	40%
163	*i, i+10*	35% (**177**)	1%	12%	52%
164	*i, i+9*	36% (**178**)	4%	10%	50%
165	*i, i+7*	50% (**179**)	4%	6%	40%
166	*i, i+6*	56% (**180**)	8%	3%	33%
167	*i, i+5*	65% (**181**)	14%	4%	17%
168	*i, i+4*	65% (**182**)	6%	4%	25%
169	*i, i+3*	65% (**183**)	22%	2%	11%
170	*i, i+2*	6% (**184**)	46%	1%	47%
171	*i, i+2*	43% (**185**)	17%	5%	35%
172	*i, i+1*	-	-	-	100%
173	*i, i+8*	47% (**186**)	3%	7%	43%
174	*i, i+9*	36% (**187**)	3%	14%	47%
175	*i, i+8*	44% (**188**)	4%	2%	50%
176	*i, i+7*	72% (**189**)	2%	6%	20%

With the purpose of demonstrating the wide applicability of this technique, conjugates **190** and **191** were also successfully obtained from palladium porphyrin **144** without noticeable changes in the reactivity compared with metal-free porphyrin **121** (Scheme 28 and Table 22).
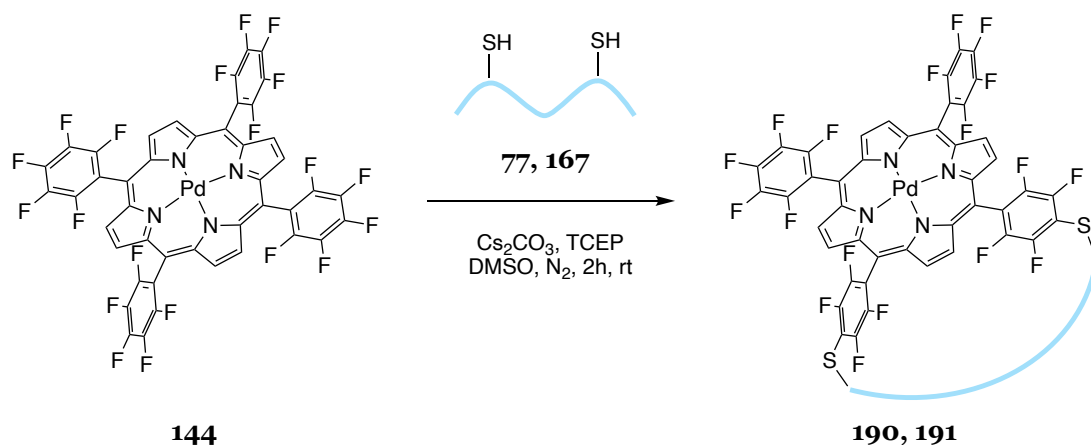
Scheme 28. General reaction scheme for the palladium porphyrin-templated cyclisation of bis-cysteine peptides.

Table 22. Peptides and corresponding conjugation products from the reaction with porphyrin 144.

Peptide	Cys-Cys distance	Monocyclic	Bicyclic	Dimer	Other
77	*i, i+8*	47% (**190**)	2%	17%	34%
167	*i, i+5*	63% (**191**)	4%	5%	28%

Except for **176**, all the reactions were performed using crude synthesised peptides (cleaved from resin but non-purified) and nevertheless proceeded successfully. This is demonstration of the versatility and robustness of this conjugation methodology and could be an advantage for rapid synthesis of peptide macrocycle libraries.

The demonstrated success of this macrocyclisation reaction when performed with a broad range of peptides and with a metal porphyrin was further proof of its generality, evidence that this cyclisation is not sequence specific and example of the multifunctionality of the tetrapyrrolic scaffold.

4.3.3 Reaction kinetics

The peptide-porphyrin cyclic product **160** was readily obtained following reaction between equimolar equivalents of **121** and **77** at room temperature (Scheme 29). To simplify the workup procedure, the minimum volume of DMSO that dissolved all reagents was employed. A nitrogen atmosphere and a thiol-free reducing agent such as tris (2-carboxyethyl) phosphine (TCEP) was introduced to limit the possible formation of side/oxidation products, including inter- and intramolecular peptide disulfides. These conditions did not differ much from those employed for simpler monoconjugation reactions.
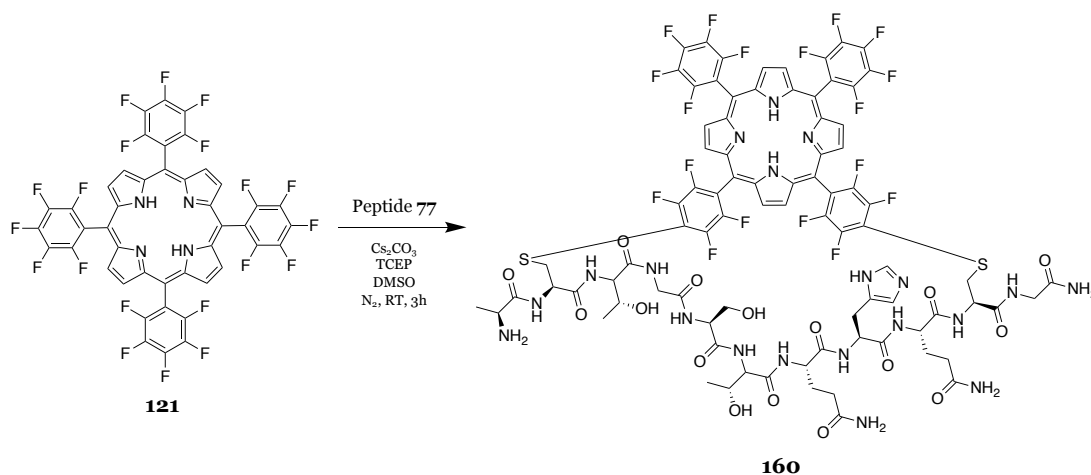
*Scheme 29. Reaction scheme and conditions for the synthesis of conjugate **160**.*

The reaction was monitored by HPLC and LC-MS analysis by sampling the reaction mixture at different timepoints. Setting the UV-Vis detector at 400 nm allowed detection of porphyrin-related species (Figure 91A), while 215 nm was used as wavelength for free peptide monitoring (Figure 91B).

Substitution products were firstly observed after 15 min and increased over time, until one main peak corresponding to the monocyclic product was detected at final timepoints (2 h and 3 h). A second peak with *m/z* ratio corresponding to macrocyclisation was also spotted, but it had negligible intensity (approximately 1:40 ration with the main peak). In the final reaction mixture, within other unidentifiable species, the bicyclic conjugate (**161** in Figure 92) and porphyrin dimer (**162** in Figure 92, *i.e.* two porphyrins reacted with two cysteine residues from the same peptide) were also identified.
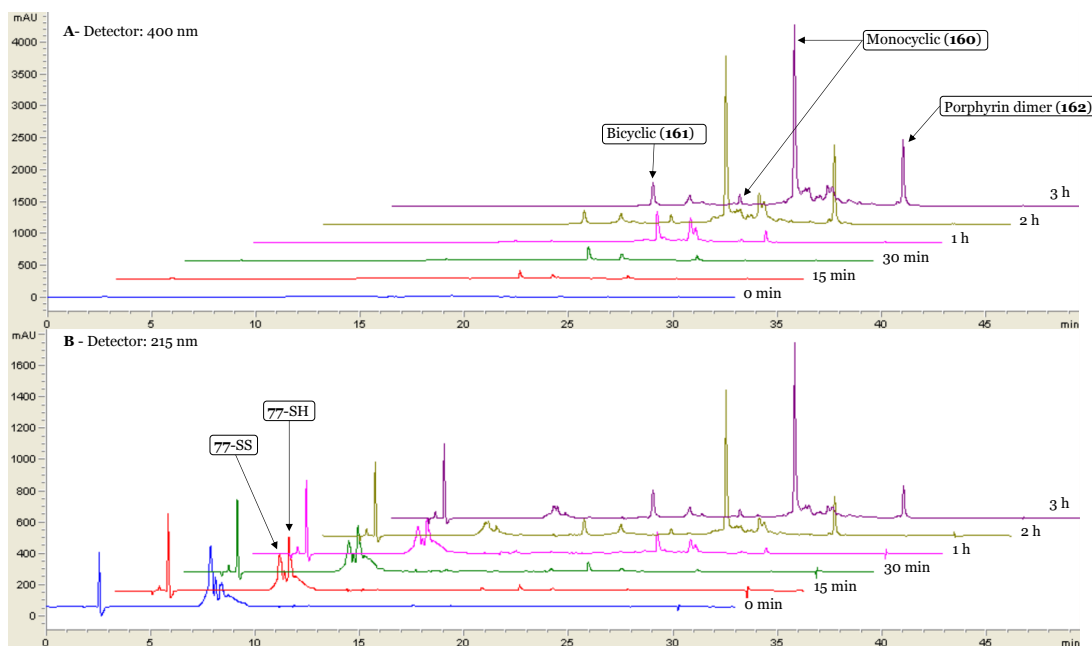
Figure 91. Time course of the porphyrin-templated cyclisation of SPACE peptide monitored with HPLC at 400 nm (A) and 215 nm (B).

The peptide-porphyrin species identified from the reaction mixture were graphically represented in Figure 90 and 92.
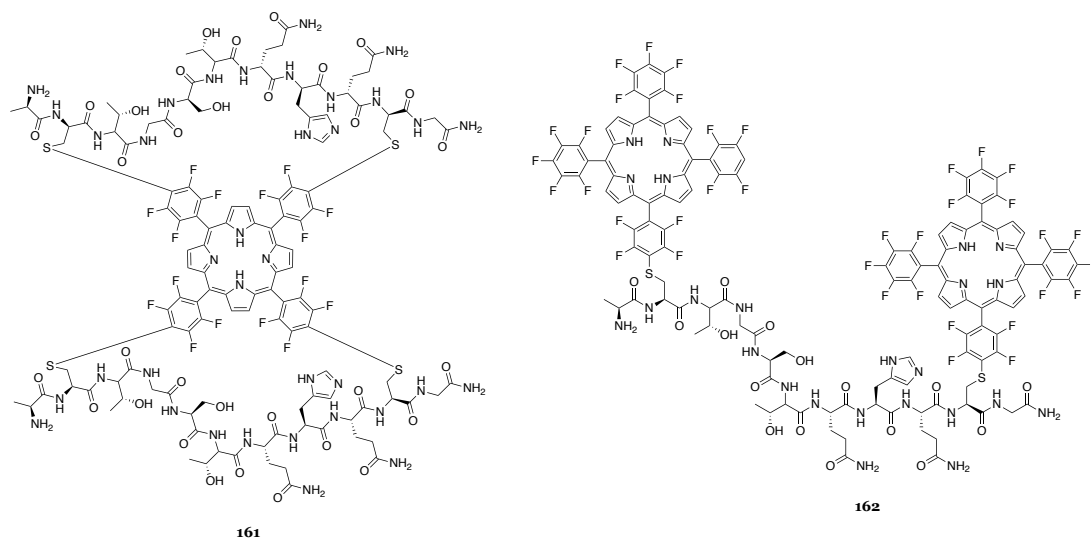
*Figure 92. Chemical structures of bicyclic conjugates (**161**) and porphyrin dimer (**162**).*

It should be noted that porphyrin dimers presented two chromophores and their UV/vis signal at 400 nm had theoretically double the absorption intensity compared to other species. This is confirmed by the different ratio between the same peaks observed from chromatograms acquired at 215 nm. The monosubstituted non-cyclic conjugate was briefly observed at early timepoints (15 min) but not in the final mixture. This intermediate was short-lived in DMSO/Cs$_2$CO$_3$ but can be easily isolated if using DMF/DIPEA, conditions that had always afforded slower reactivity in these studies.

The biggest improvement in product formation was registered between 1 h and 2 h (from 30% to 45% of the crude), without significant differences between 2 h and 3h. After 3 h, no starting peptide was observed at 215 nm, indicating complete consumption of the starting reagent (Figure 91B). Peptide **77** was employed as crude for this reaction: a few impurities were initially present, and these could have generated the other unidentifiable side products.

The stability of the peptide in Cs_2CO_3/DMSO was verified using a control experiment, performed by dissolving previously purified **77** in this reaction mixture without porphyrin **121**. The mixture was sampled and analysed as described before (Figure 93). At the first timepoint, around 50% of oxidised peptide (**77-SS**) was immediately formed and could be detected at shorter retention times (8 min vs 8.5 min). This was explained by the slower solubilisation of TCEP in DMSO compared to peptide and base.

With time, the reducing agent also dissolved, and reduction of the disulfide originated back the starting reactive peptide (from **77-SS** to **77-SH)**. After 3 h, reduced peptide (**77-SH)** was still observed as main component of the mixture along with a few degradation products that were not identified. Therefore, no significant degradation of **77** – but only its reversible disulfide form – was overall observed by HPLC/LC-MS.
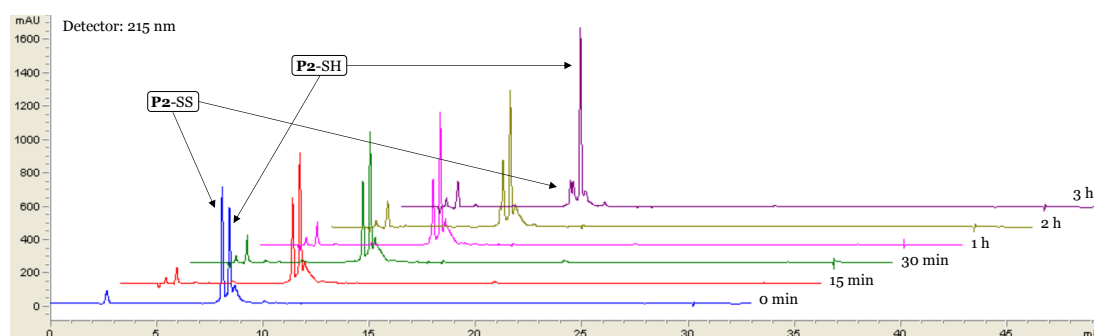
Figure 93. Time course of the SPACE peptide (77) stability experiment monitored with HPLC at 215 nm.

If the initial peptide oxidation (ascribable to slow TCEP dissolution) took place in a complete reaction mixture, this could have prevented thiols reactivity, and therefore, be a rate-determining factor that requires further investigation. Addition of peptide to a DMSO mixture with pre-dissolved TCEP could avoid initial peptide oxidation in favour of a faster reactivity. Detailed reaction conditions and experimental procedures are discussed in Chapter 6.

To better understand the 5-10 vs 5-15 formation, the reaction kinetics was further explored using porphyrin **121**, different peptides (peptides **77, 166-168**) and porphyrin/peptide ratios of 2:1 and 1:2 (Figure 94).
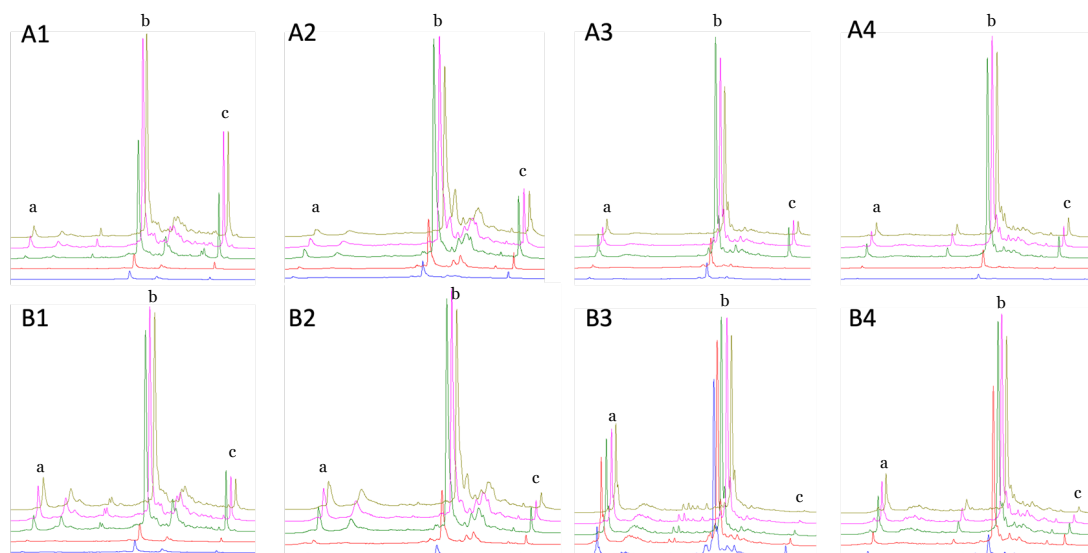
*Figure 94. Reaction time course for the cyclisation of peptides **77** (1), **166** (2), **167** (3), **168** (4) with porphyrin **121**. Porphyrin/peptide ratios are 2:1 (A) or 1:2 (B). Samplings were performed after 15 min (blue), 30 min (red), 1 h (green), 2 h (magenta), 3 h (olive). a= homobicylcic, b= monocyclic, c= peptide-bridged porphyrin dimers.*

Firstly, the di-substituted monocyclic product was observed to occur rapidly and cleanly in all cases, as previously observed for peptide **77**. When porphyrin **121** was treated with excess of peptide (Figure 94B1-4), direct conversion to a homobicyclic conjugate, in which two molecules of the same peptide reacted twice with the porphyrin (Figure 95), was improved compared with conditions using a lower molar ratio of peptide.
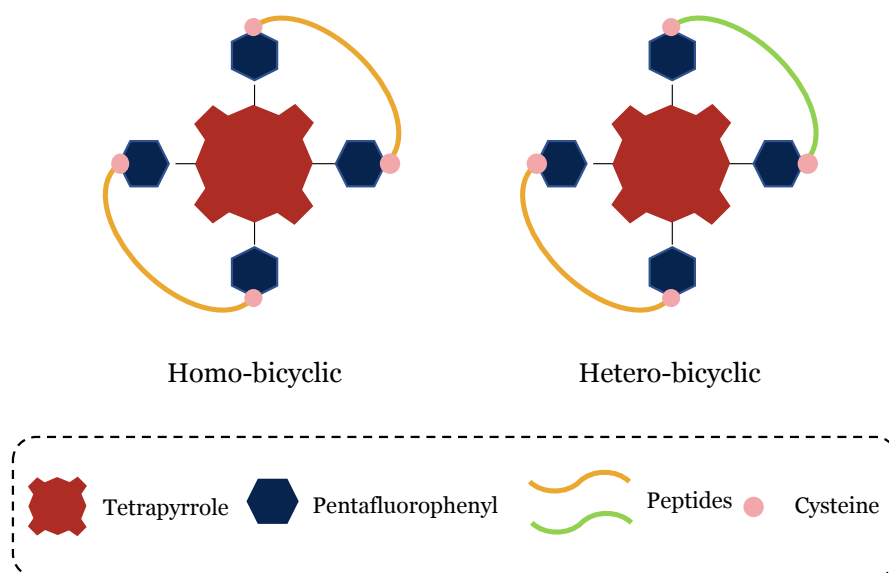
Figure 95. Graphical representation of homobicyclic and heterobicyclic conjugates.

The homobicyclic product was more prevalent for a shorter inter-thiol separation such as *i, i+5,* while greater inter-thiol spacing led to a smaller amount of the homobicyclic species. Peptide-bridged porphyrin dimers were observed in small quantities, but these species were comprehensibly more abundant when using 2:1 porphyrin-peptide ratios (Figure 94A1-4).

4.3.4 NMR studies of cyclic products

To gain further insight into how the number of inter-thiol residues was related to preferential 5-10 and 5-15 substitution products, main products from the cyclisation of porphyrin **121** with peptides **77** and **166-168** were purified and analysed by NMR. The NMR studies were executed with the support of Sam Hanson (PhD student at The University of Edinburgh).

As previously shown, unreacted pentafluorophenyl groups exhibit three ^{19}F NMR signals at -136, -151 and -161 ppm with an integration pattern of 2:1:2. After the cyclisation, two signals, corresponding to *ortho* and *meta* fluorine atom pairs on substituted tetrafluorobenzene, appears between -135 and -140 ppm. ^{19}F NMR experiments confirmed the double substitution on two pentafluorophenyl groups and these signals showed a complex multiplicity that can be ascribed to a loss of symmetry conferred by the peptide strapping. The spectra below show the ^{19}F NMR of conjugate **160** (Figure 96) and conjugate **181** (Figure 97) cyclic products.
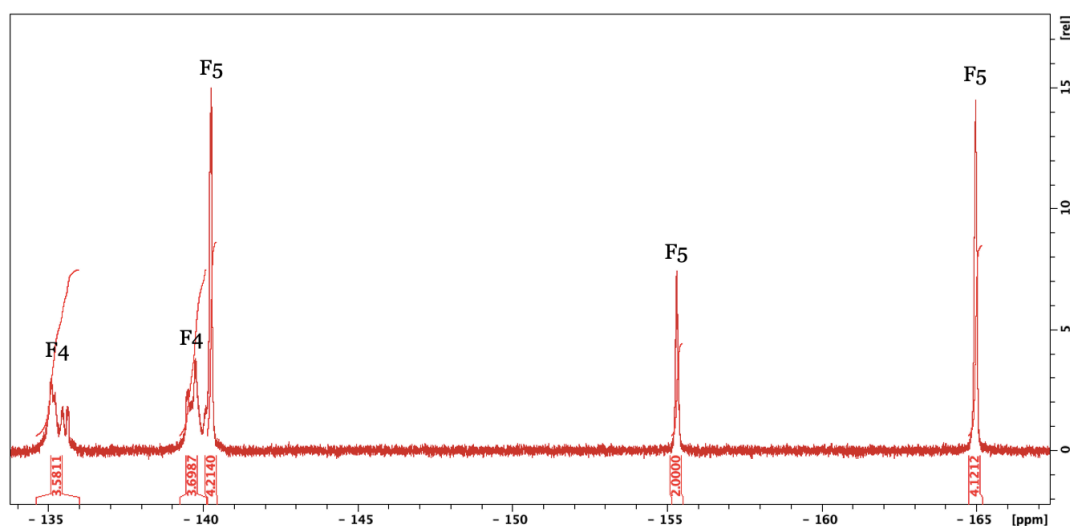
*Figure 96. ^{19}F NMR spectrum of conjugate **160**. F4= signals from substituted tetrafluorobenzene; F5= signals from meso-pentafluorophenyl groups.*
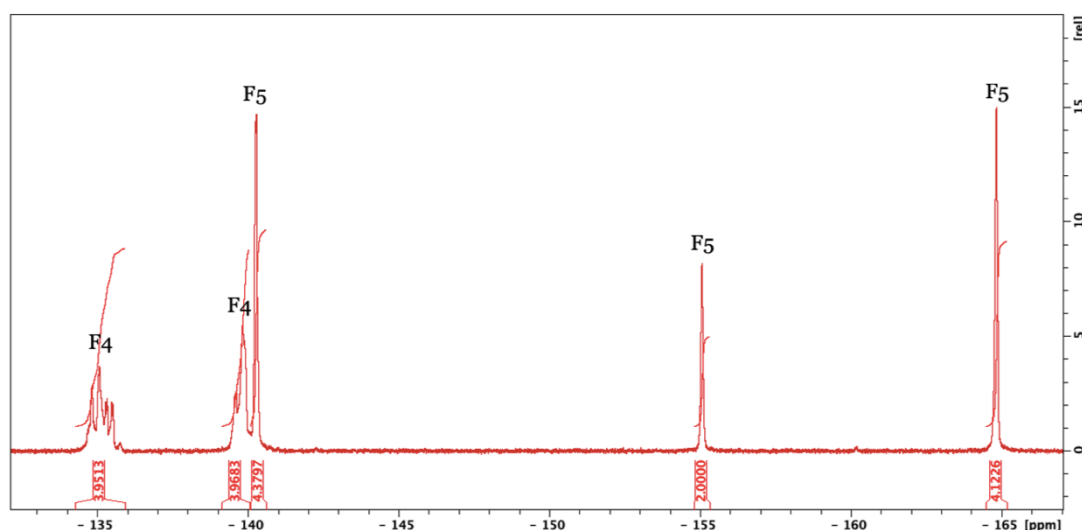
*Figure 97. ^{19}F NMR spectrum of conjugate **181**. F4= signals from substituted tetrafluorobenzene; F5= signals from meso-pentafluorophenyl groups.*

These experiments did not show a clear difference between substitution products originated from peptides with inter thiol length from *i, i+8 to i, i+4*, indicating that the same cyclisation geometry likely dominated, and supporting the hypothesis that the 5-10 isomer was the preferred cyclisation product.

The electrons of the tetrapyrrole aromatic macrocycle generate a strong ring current and an induced magnetic field. This is stronger in the centre of the ring and causes the extreme shielding of the N-H in ^{1}H NMR (Figure 98). When analysing a porphyrin, the pyrrolic nitrogen peak is shifted to -2 ppm from a usual chemical shift of 7.1 ppm for liquid pyrrole.[253]
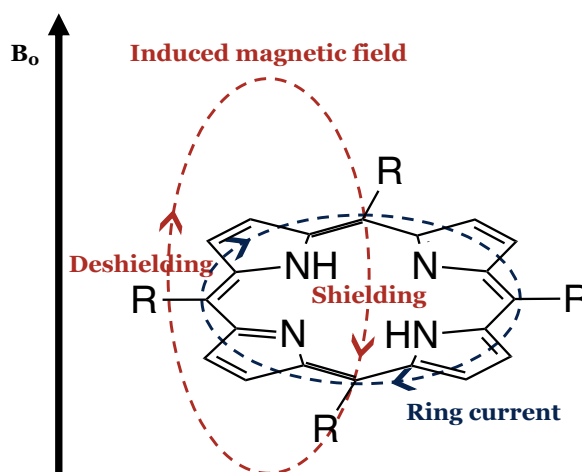
Figure 98. Graphical representation of the tetrapyrrole ring current.

The same effect applies to any proton belonging to a strap which spans across the face of a porphyrin, as already reported.[235,240] Therefore, ^{1}H NMR was also considered as a potential tool to understand the position of the peptide chain and the geometry of this system. However, beyond the spectra complexity, the comparison between ^{1}H NMR of peptide **77** and conjugate **160** did not show the expected proton shifts.

Both these NMR experiments require a more detailed investigation but could potentially support the regioselectivity study and isomer determination.

4.3.5 Cyclisation on porphyrin 148

Since HPLC, LC-MS and NMR analyses were not informative with regards to 5-10 vs 5-15 cyclisation, the cyclisation of key peptides was also performed with porphyrin **148**. Compared to porphyrin **121**, this scaffold can only afford the cyclic product with 5-15 geometry because lacks the two *meso* pentafluorophenyl groups in position 10 and 20.

In this case, the reaction time was exceptionally extended to 5h, because of the lower solubility of porphyrin **148** in DMSO, and to make sure that, even in case of a slower kinetic compared to porphyrin **121**, the main products could be observed. As indicated in Scheme 30 and Table 23, cyclic products (conjugates **192** and **193**) were obtained from peptides **77** (SPACE peptide) and **166**.
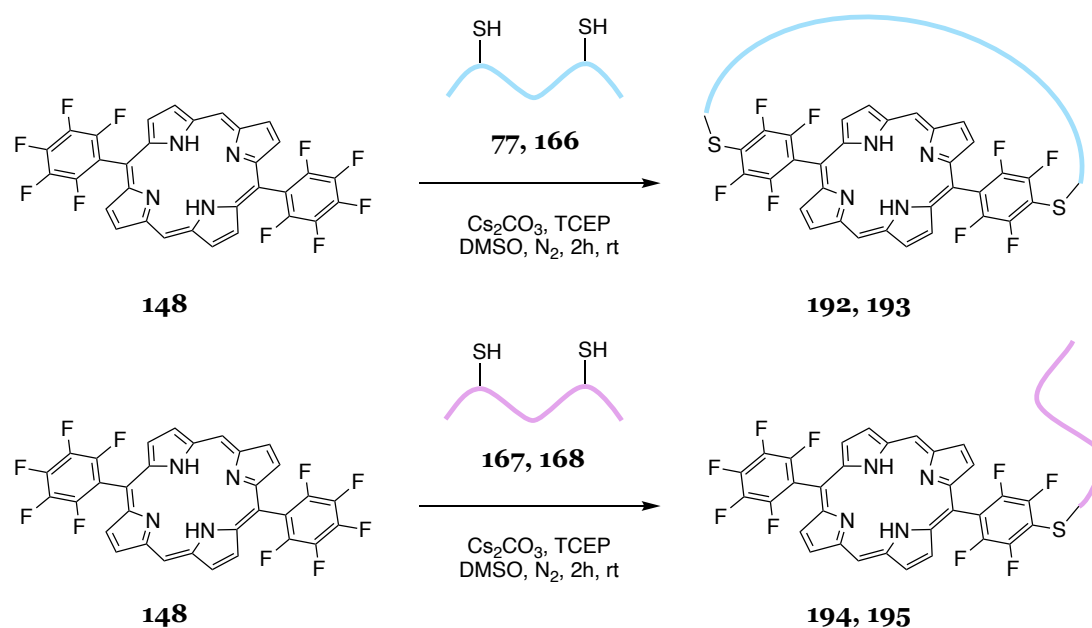
*Scheme 30. General reaction scheme for the cyclisation of bis-cysteine peptides with porphyrin **148**.*

*Table 23. Peptides and corresponding conjugation products from the reaction with porphyrin **148**.*

Peptide	Cys-Cys distance	Monocyclic	Conjugate	Other
77	*i, i+8*	55% (**192**)	-	45%
166	*i, i+6*	10% (**193**)	40%	50%
167	*i, i+5*	-	50% (**194**)	50%
168	*i, i+4*	-	60% (**195**)	40%

The reactivity of peptide **77** (*i, i+8*) with porphyrin **148** is comparable to what previously observed for porphyrin **121**, leading to a clean conversion and good product distribution (Figure 99). Differently, peptide **166** (*i, i+6*) did not react well with porphyrin **148**, probably because of the shorter inter-thiol distance: despite the product of interest was observed by HPLC with a corresponding peak of good intensity, crude purity and percentage of conversion were lower than previous examples.
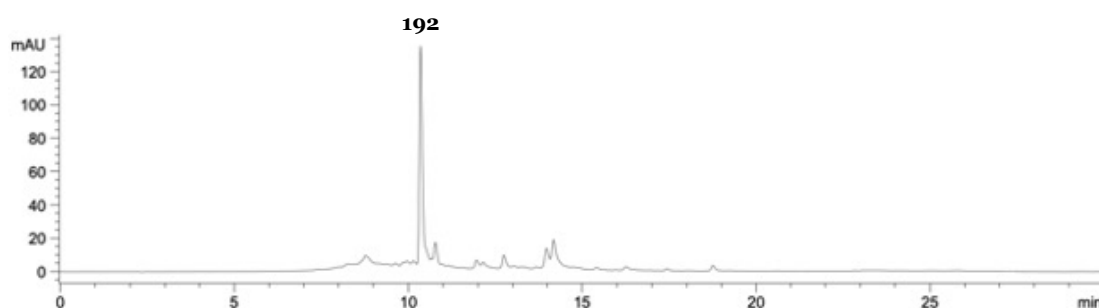
*Figure 99. HPLC chromatogram of crude conjugate **192** acquired at 400 nm.*

On the contrary, peptides **167** (*i, i+5*) and **168** (*i, i+4*), which were previously successfully cyclised around **121**, did not give any cyclic species but only mono-conjugation products (**194** and **195**).

NMR analysis was not performed on porphyrin **148** cyclisation products (**192-195**) because the reactions were executed on a small scale (see section 6.10.5) and the reaction crude was not enough for purification and NMR analysis. These were analysed as crude by HPLC and LC-MS only.

Overall, these results indicated that *i, i+2* is the minimal inter-thiol distance that allows 5-10 cyclisation while *i, i+6* is the minimal inter-thiol distance that could also permit a forced 5-15 cyclisation. The observation that only a single (major) monocyclisation product is normally obtained may further support the preferred 5-10 geometry, which is probably statistically and kinetically favoured. However, the potential of peptides with inter-thiol distance longer than *i, i+6* to react in 5-15 positions is not a guarantee that the ring closure with 5-10 geometry is obtained and further experiment which were beyond the scope of this work will be necessary to define this aspect of the reaction.

4.3.6 Two-step multicyclisation

Time course studies and detailed analyses of cyclisation side products suggested that, following the successful installation of a first bis-cysteine peptide on porphyrin **121,** it was possible to install a second bis-cysteine peptide onto the same porphyrin scaffold. Therefore, the controlled formation of bicyclic species was also evaluated.

Initially, the isolated monocyclic product **160** was treated with a second batch of peptide **77** (*i, i+8*) following the usual Cs_2CO_3/DMSO procedure, but no further substitution was observed (Scheme 31).
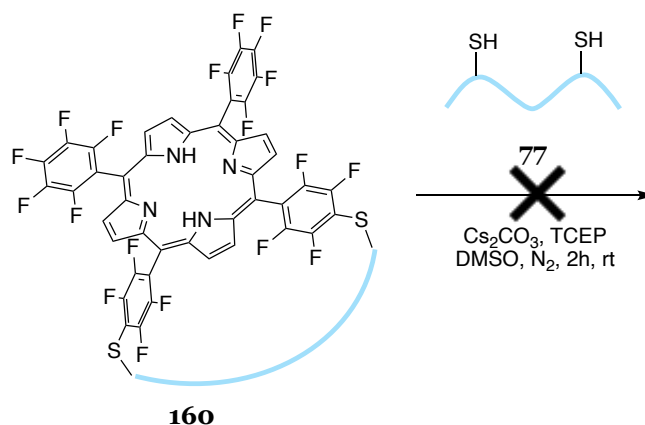
Scheme 31. Reaction scheme for the attempted two-step multicyclisation on conjugate **160**.

However, treatment of **160** with an excess of 4-acetamidethiophenol (**196**) under the same conditions afforded the tetra-substituted product **197** (Scheme 32) and complete conversion of the starting reagent **160**. This indicated that the remaining *para*-fluorine atoms are not

deactivated by the initial substitutions, and that the lower reactivity may be ascribed to steric hindrance or any other factor occurring when two peptides were employed.
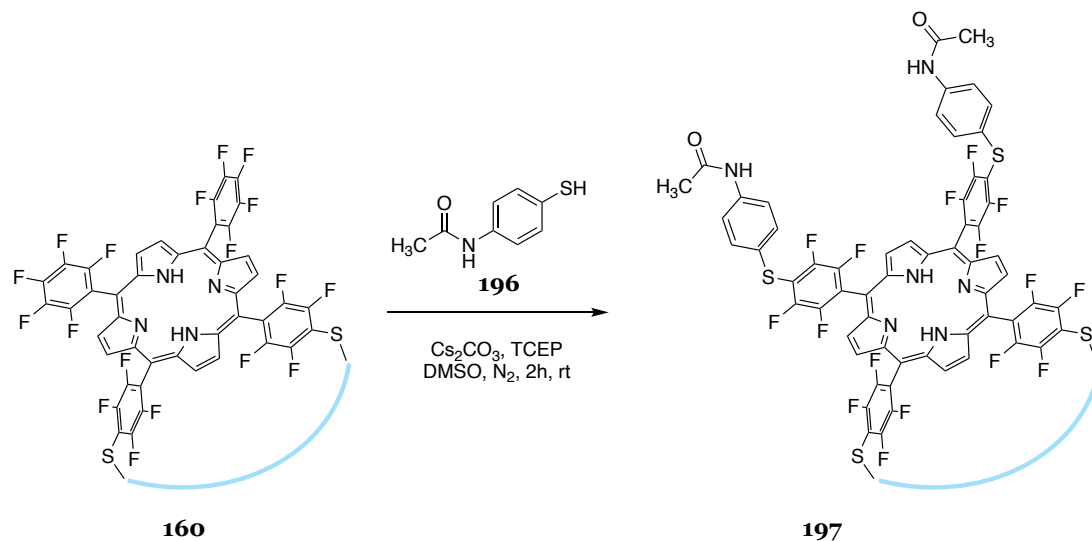
*Scheme 32. General reaction scheme for the conjugation of 4-acetamidethiophenol (**196**) on conjugate **160**.*

Conversely, when the crude 5-10 disubstituted conjugate **181** was treated with an excess of the *i, i+5* peptide (**167**), complete conversion to the homobicyclic tetra-substituted product **198** was obtained (Scheme 33). As a demonstration of the stepwise modularity of these systems, monocyclic **181** was also treated with an excess of a different bis-cysteine peptide (peptides **77** and **168**) to obtain heterobicyclic peptides (conjugates **199** and **200**, respectively).
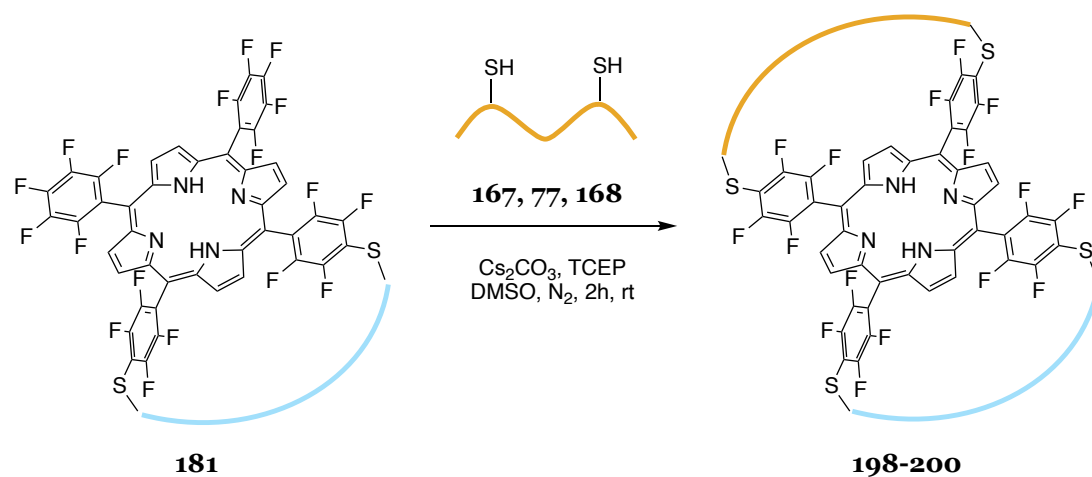
*Scheme 33. General reaction scheme for the two-step multicyclisation on conjugate **X**.*

^{19}F NMR of crude conjugate **198** confirmed substitution of all *para* fluoride atoms, with almost complete disappearance of non-substituted porphyrin signals (Figure 100).

Aside from these explanatory examples, the reaction on monocyclic species can potentially be performed with any other thiol (small molecule or peptide) to further functionalise the porphyrin scaffold. Likewise, with this modular approach, two different peptide sequences

with short inter-thiol distance can be employed in a two-step multicyclisation strategy to obtain multicyclic and multifunctional species, maintaining regiocontrol over the cyclisation.
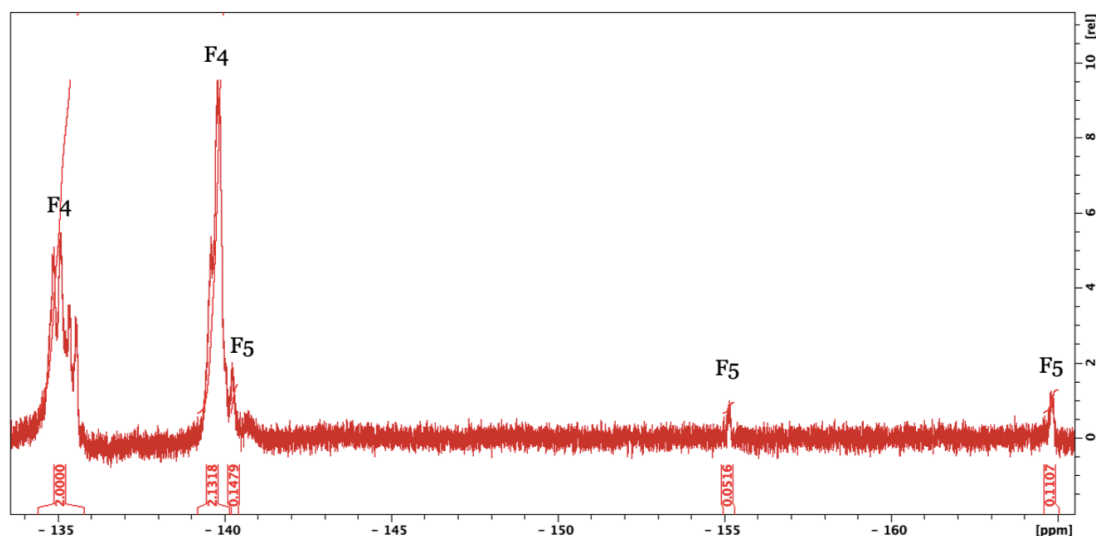
*Figure 100. ^{19}F NMR spectrum of crude conjugate **198**. F4= signals from substituted tetrafluorobenzene; F5= signals from meso-pentafluorophenyl groups.*

4.3.7 One-step multicyclisation

To explore the extension of these reactions to peptides containing more than two cysteine residues, the longer peptides **201-203** were synthesised. As the *i, i+5* inter-thiol distance previously provided fast, selective and clean conversion to 5-10 cyclic products, the CTQHQC motif was selected as a prototype length for multicyclisation.

Peptides **201-203** were then employed for reaction with porphyrin **121** as detailed in Scheme 34 and Table 24.

*Table 24. Peptides and corresponding conjugation products from the reaction with porphyrin **121**.*

Peptide	Sequence	Cys-Cys distance	Multicyclic	Other
201	ACTQHQCTQHQCG-NH2	*i, i+5*	30% (**204**)	70%
202	ACTQHQCTQHQCTQHQCG-NH2	*i, i+5*	20% (**205**)	80%
203	ACSTQHQCSTQHQCSTQHQCG-NH$_2$	*i, i+6*	34% (**206**)	66%

The bicyclic conjugate **204** was obtained in one pot following the procedure previously described. The product distribution is slightly different from the macrocyclisation reaction, but the multicyclic specie could be clearly observed by HPLC and LC-MS (Figure 101).
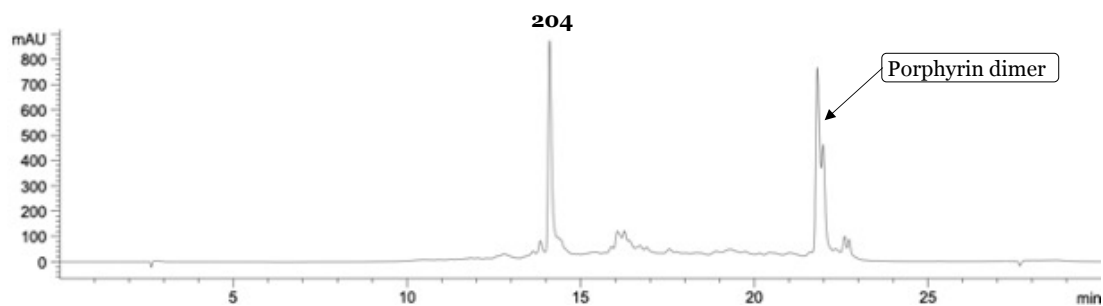
*Figure 101. HPLC chromatogram of crude conjugate **204** acquired at 400 nm.*

The tricyclic product (conjugate **205**, Scheme 34) could also be synthesised from peptide **202** and after an extended reaction time of 5 hours. A clearly detectable amount of multicyclic product was finally obtained as part of a mixture of regioisomers, side products and some intermediates from incomplete cyclisation (Figure 102). In this latter case, a greater distance between cysteine favoured the complete multicyclisation as demonstrated with peptide **203** and its tricyclic product **206**.
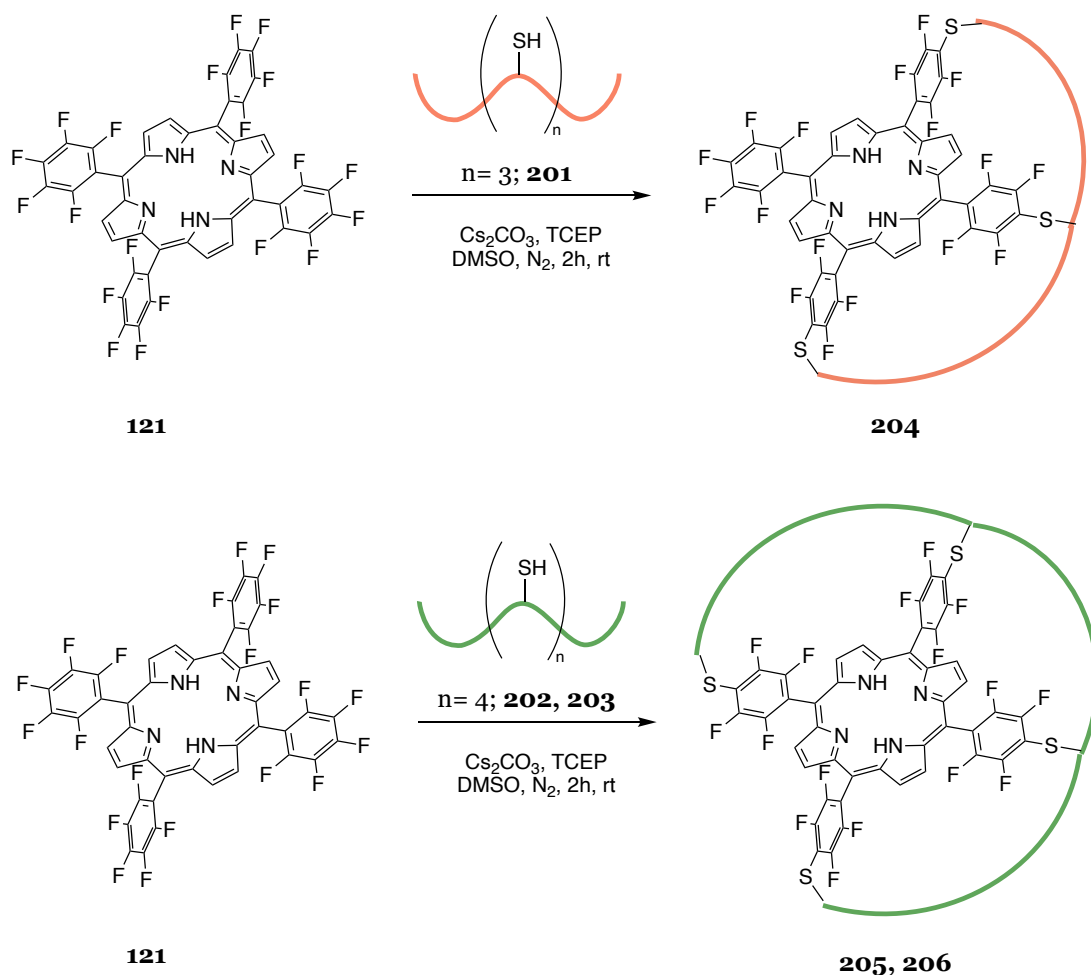
Scheme 34. General reaction scheme for one-step porphyrin-templated multicyclisation of cysteine-rich peptides.

These significant results allow direct access to multicyclic peptide systems from peptides with multiple cysteine residues and are key findings for the future development of porphyrin-

templated cyclic peptides. This is also further confirmation of the level of control acquired over this chemical process.
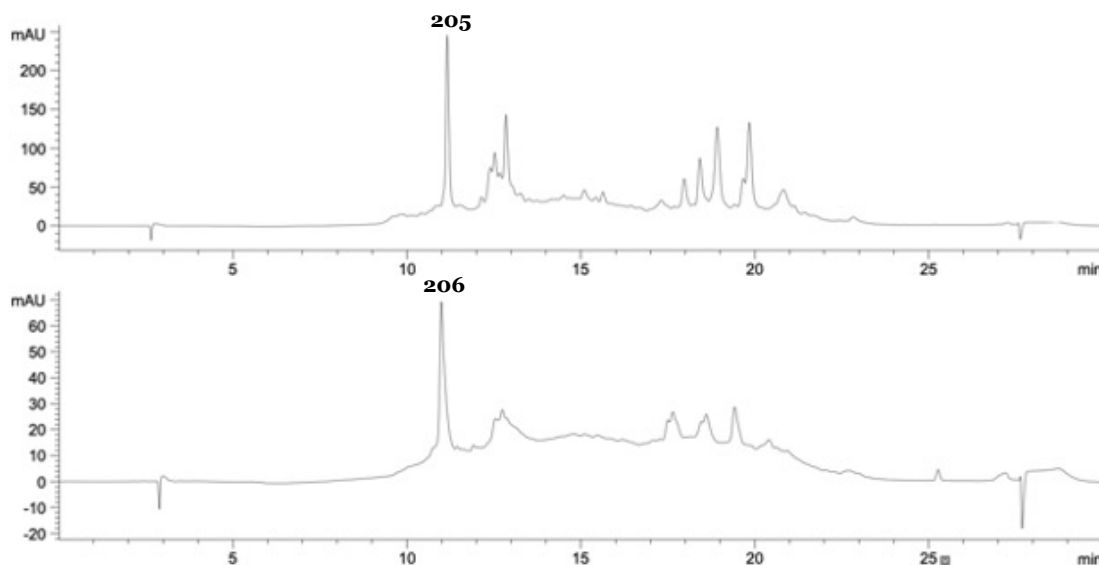
*Figure 102. HPLC chromatogram of crude conjugates **205** (A) and **206** (B).*

4.3.8 Biological activity of conjugate **160**

Finally, the water-soluble conjugate **160** (Figure 103A) was biologically tested by Talhat Chaudhry and Giulia Scagnetti (PhD students at LJMU) to determine whether the skin cell penetrating properties of the original SPACE peptide were retained after porphyrin-templated cyclisation.[166,248]

Cellular uptake of conjugate **160** was tested in SCCIC8 cells[254] and cell imaging was performed by fluorescence microscopy. Differently from previous studies on self-assembled nanoparticles, a confocal microscope was employed. This analysis was facilitated by the intrinsic fluorescence of the conjugates in the red region, which is one of the properties conferred by the integration of the tetrapyrrole core within the peptide architecture. Conjugate **160** characteristic absorption and emission spectra are shown in Figure 103B.
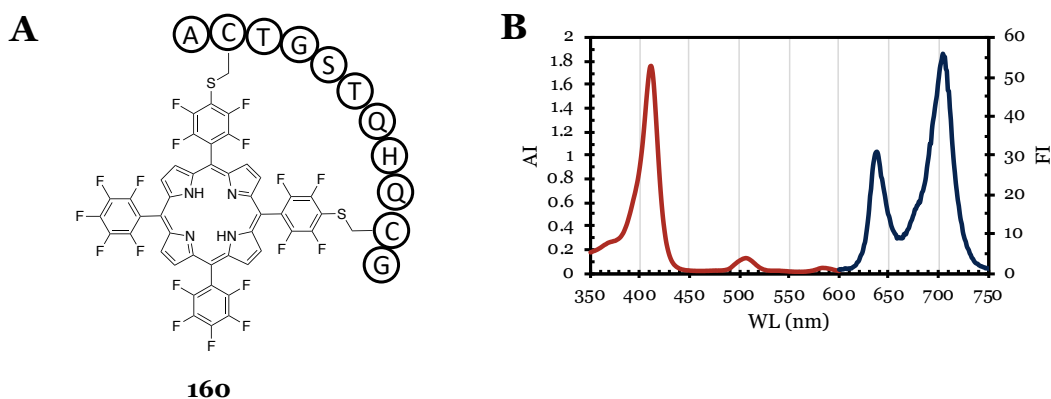
*Figure 103. A) Simplified chemical structure of conjugate **160**; B) Absorption (red) and emission (blue) spectra of conjugate **160**.*

After 24h treatment with **160,** nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI). Subsequent analysis of cells showed significant uptake of **160** in the cytoplasm. Figure 104A shows the overlapping between DAPI specific emission in the blue range and porphyrin emission in the red range. In Figure 104B, only the emission in the red range was selected. Since DAPI emission was observed in both ranges, the overlapping between red and blue emissions makes the nuclei purple in Figure 104A and red in Figure 104B. This analysis indicated the presence of the peptide-porphyrin conjugate in the cytoplasm, and therefore, suggested that the cyclic conjugate retained the skin cell penetrating properties of the free peptide. Specifically, there might also be a specific localization in the spots with more intense emission which could be ascribable to **160** accumulations in subcellular structures such as lysosomes. However, this hypothesis requires further investigation. As a control experiment, the same test and analysis were performed for the uptake of conjugate **186** – the cyclized version of the SPACE peptide control. Uptake of conjugate **186** in SCCIC8 cells was only slightly weakened compared to **160** (Figure 104C and 104D). This suggested that the uptake of the conjugates was related to the presence of certain AAs, regardless a specific order of these residues in the sequence. However, another control experiment was performed with a different cell line, and non-small-cell lung cancer cell line (Calu-3) were treated with conjugate **160,** which was clearly not taken-up. As can be observed in Figure 104E and 104F, only the DAPI stained nuclei are detected, at DAPI specific wavelength as well as at the porphyrin emission range. Thus, conjugate **160** cyclisation with **121** did not preclude the cellular uptake and the conjugate localized in SCCIC8 cells in a cell specific manner.
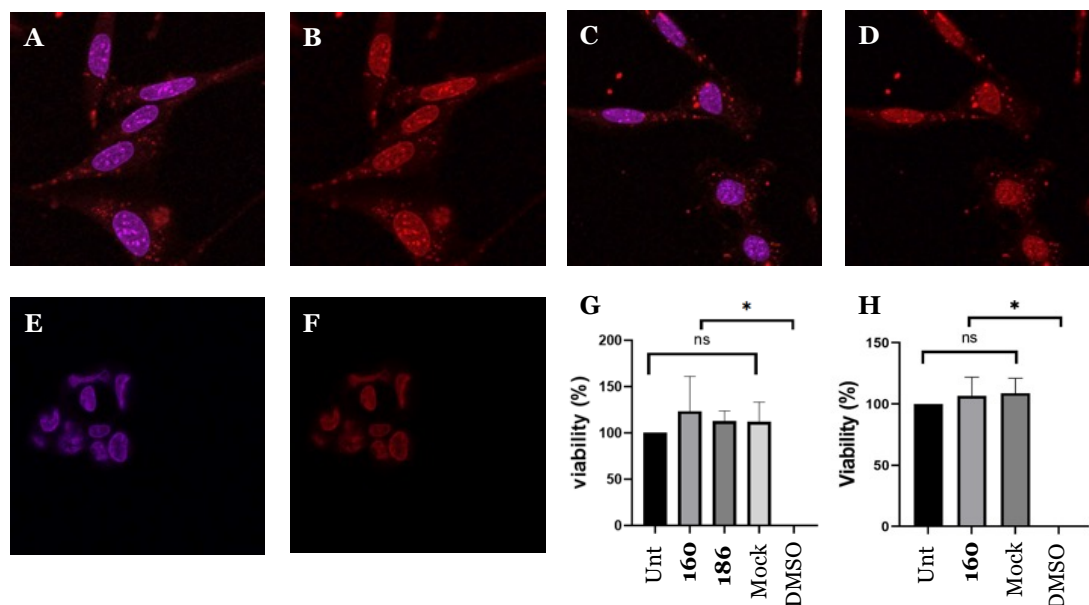
*Figure 104. Intracellular localisation (red fluorescence) of conjugate **160** (A and B) and conjugate **186** (C and D) in SCCIC8 cells. No uptake of conjugate **160** in Calu 3 cells (E and F). Magnification is 63 X. Cell viability (%) of SCCIC8 cells (G) and Calu 3 cells (H).*

Similarly to what has already been noted for HFB-templated SPACE peptide, the uptake results obtained with conjugate **160** could be compared with those from SPACE peptide labelled with a standard probe as an additional control experiment for future work.

Furthermore, cytotoxicity studies were performed with conjugates **160** and **186**, using working concentration of 10 μM. These experiments showed that conjugate **160** did not induce cytotoxicity in SCCIC8 nor Calu-3 cells (Figure 104G and 104H). This is a positive result because porphyrin cytotoxic activity could be selectively obtained only after specific irradiation, whenever necessary.

4.4 CONCLUSION AND FUTURE WORK

The aim of this work was to introduce porphyrin **121** as a widely applicable platform for the cyclisation of cysteine-containing peptides, relying on a previously developed robust conjugation technique. As an alternative approach to reported porphyrin strappings and peptide cyclisation tools, this method can be performed on a tetra-substituted symmetric porphyrin with any peptide sequence presenting – naturally or after specific modification – two cysteine at defined distances.

Its application to a variety of peptides demonstrate that the chemistry is not sequence specific and that the thiolate-fluoride $S_N Ar$ reaction is chemoselective. Important insights into the geometrical constraints of the macrocyclisation, related to the minimum number of AA required to reach the *para* positions of 5-15 and 5-10 *meso* phenyl groups, were contextually investigated with the aim of selectively obtaining the 5-10 cyclic product by modulation of the inter-thiol length. If there is at least one amino acid between the cysteine residues, the peptide can cyclise in the 5-10 position. When the inter-cysteine distance is extended to five or more residues, the peptide could potentially react in the 5-15 position as well, but the 5-10 cyclic product may still be preferred.

The ability of the conjugation product between porphyrin **121** and SPACE peptide **77** to permeate skin cells, without cytotoxicity, confirmed that the peptide retained its original cell penetrating activity. In addition, porphyrin-mediated macrocyclization allowed the development of a fluorescent peptide-porphyrin conjugate. Therefore, this strategy joins and concurrently improves the individual properties of these two classes of biologically active compounds.

A more specific characterization of these system will be necessary to help understand the preferred geometry for peptides with longer inter-thiol distance, for example employing x-ray and CD studies. Future work could also comprise the addition of peptides to the library and, depending on the sequences, execution of more specific biological assays. Despite being a difficult to synthesise $A_2 B_2$ porphyrin, **148** is an interesting scaffold that could be worth investigating further. NMR studies on this substrate could help understand the geometry of the 5-15 system.

The modulation of the porphyrin properties by possible incorporation of a metal in the tetrapyrrole core adds supplementary value for the potential applications of the final porphyrin-templated cyclic peptides. This strategy offers the possibility to modulate the photoactive tetrapyrrolic scaffold and differently exploit its properties, in combination with those of the peptide. Different metal derivatives of **121** could be employed in future studies to expand the range of applications, not only as biologically active species but also as catalyst or light-responsive materials.

After the cyclisation of a single peptide, cyclic conjugates **177-189** offer the possibility of further functionalisation. This can be exploited to conjugate single thiols or to access multicyclic species by two-step reaction with another peptide. Monocyclic peptide-porphyrin system, which are already multifunctional, could be further functionalised by conjugating thiols of non-peptidic nature, such as cyclodextrins to improve hydrophilicity or phosphonium cations for mitochondrial uptake.

Similarly, the two-step synthesis of multicyclic peptides could be used to combine AA sequences with complementary activity on the same scaffold. For example, a system comprising one loop with a cellular delivery sequence and another loop with a subcellular localising peptide could be designed.

Cysteine-rich peptides can also be wrapped around the porphyrin scaffold in a one-step multicyclisation procedure. This multicyclisation tool is highly promising but requires further optimisation. This optimisation would include performing the reaction using model peptides with different length as well as exploring conditions to push the multicyclisation to completion and limit the formation of regioisomers. Once optimised, this strategy will allow grafting of multiple active AA sequences in a single peptide and its subsequent one-step porphyrin-templated multicyclisation to obtain an incredibly versatile and multifunctional construct.

As a further development, it seems feasible to plan the coupling of the N and C-termini of this tricyclic species to perform head-to-tail cyclisation and form a fourth loop.

Finally, since the stability to pH, temperature and serum should be improved for cyclic and ever more multicyclic peptides, this would be an interesting point to evaluate that will further enrich the potential of these systems.

5 General conclusion

The thiol-fluoride nucleophilic aromatic substitution has always been a valuable tool in organic chemistry. In this work, novel reaction conditions and unexplored uses of this chemistry for peptide bioconjugation were studied. The diverse and successful results presented here are further proof of this reaction potential and set the basis for additional improvement and applications of resulting products.

In Chapter 2, starting from the conditions reported in literature, initial examination and subsequent optimisation of the thiol-fluoride nucleophilic aromatic substitution were performed on model systems. This provided a detailed understanding and practical mastery of this chemical process, from general laboratory procedures to specific reagent performances.

Specifically, DMSO/Cs2CO3 was identified as a new and most promising condition for the bioconjugation of cysteine-containing non-protected peptides with polyfluoroaromatics. In addition, the possibility to choose between different base/solvent combinations (*e.g.* DMSO/Cs2CO3, DMF/DIPEA, MeCN/DBU), combined with the modulation of reaction conditions (*e.g.* dilution, reagent ratio, temperature), provided control over product formation (*i.e.* different levels of thiol-fluoride substitution). This extended knowledge of the thiol-fluoride substitution and availability of multiple reaction conditions can be useful to exploit this chemistry on alternative substrate, giving access to new chemical species for advanced applications.

This preliminary work was also key for the development a novel peptide-porphyrin bioconjugation tool. Cysteine-containing peptides of different length and sequences were coupled - once or multiple times - to a polyfluorinated porphyrin using the new and appositely developed experimental procedure based on DMSO/Cs2CO3 reaction conditions.

The first objective was therefore achieved in full, in respect of the thiol-fluoride $S_N Ar$ optimization and its application for peptide bioconjugation with HFB and porphyrin. On a side note, HFB-templated peptide multicyclisation was not achieved using this technique and will require further work. As evidence of the applicability of this new bioconjugation tool, thiol-fluoride nucleophilic aromatic substitution between peptide and porphyrin was employed in two research fields of high interest for the scientific community: synthesis of peptide-porphyrin self-assembling systems and peptide cyclisation/porphyrin strapping.

In Chapter 3, novel self-assembling molecules were synthesised combining modified KA6K bolaamphiphilic peptides and polyfluorinated porphyrins via thiol-fluoride $S_N Ar$. These systems mostly assembled into nanometric and micrometric nanoparticles that

were clearly observed with different microscopy techniques. The good solvent, pH and ion strength of the bad solvent, concentration of conjugate in the final solution, as well as time and temperature of incubation were the different self-assembly conditions explored, but only time, pH and ionic strength appeared to have a significant effect on the process. Both the peptide and porphyrin moieties contributed to the assembly, and the tetrapyrrolic part of the molecule probably promoted stacking interactions that also led to quenching of the light induced properties typical of J- and H-aggregates formation from free tetrapyrrole rings.

These new bolaamphiphilic systems fit perfectly as new examples of self-assembling species for peptide-porphyrin supramolecular architectures. This field is under constant improvement and the combination of polyfluorinated porphyrins with bolaamphiphilic peptides is a significant novelty for the development of new systems and applications.

The second objective was also successfully accomplished, because the assembly of the conjugates was showed and explored in detail. However, for future applications, the assembly should not quench the properties of the porphyrin. Optimization of the hybrids systems obtained from the co-assembly process could be a solution.

In Chapter 4, the porphyrin-templated cyclisation of any peptide that includes cysteine residues in appropriate number and positions was also performed with this technique. Natural peptides or appositely designed amino acid sequences can be macrocyclised or multicyclised (in one or two steps) around a porphyrin scaffold to obtain 3D macromolecules that combine the properties of the chosen peptide and the porphyrin. To achieve these significant results, peptides with inter-thiol distances ranging from *i, i+10* to *i, i+1* were tested, and the *i, i+5* motif demonstrated to be the most appropriate for regioselective macrocyclisation and one-step multicyclisation.

As an additional advantage, the porphyrin-templated cyclisation did not interfere with the biological activity of the tested peptide and examples of skin-penetrating and fluorescent peptides were obtained.

Overall, this technique is a valid extension of the existing methods for peptide cyclisation as well as porphyrin strapping, with many possibilities to be further upgraded or implemented in other architectures. The third objective exceeded expectations: not only macrocyclic species were obtained, but also peptide multicyclisation was observed. These studies will ease access to novel self-assembly systems, cyclic architectures and any other macromolecules obtained from the bioconjugation of peptides and porphyrins that could be potentially applied as antibody mimics, therapeutic agents (PDT, PTT, PACT), detection tools for health screening and early diagnosis of disease (^{19}F Magnetic Resonance Imaging (MRI), or nuclear medicine (after tetrapyrrole metal/radioisotope insertion), or models of complex natural machineries.

6 Materials and Methods

After an initial table to summarise all the peptide sequences employed in this work (Table 25), this chapter will comprise a description of the scientific methodologies applied, comprising technical details and specific experimental procedures, as well as materials employed and compound characterisation.

Table 25. Summary of the peptide sequences employed in this work.

Peptide	Amino Acid Sequence	Reference
69	H-YCGGGCAL-NH$_2$	[51]
76	H-YC(Mmt)GGGC(Mmt)AL-NH$_2$	
77	H-ACTGSTQHQCG-NH$_2$	[166]
78	Ac-ACTGSTQHQCG-NH$_2$	
79	Ac-AC(Mmt)TGSTQHQC(Mmt)G-NH$_2$	
81	Ac-AhxK(Alloc)AhxACTGSTQHQCG-NH$_2$	
98	H-WGCKCGGKGGCGKGGGCGW-NH$_2$	[139]
103	H-WGC(Mmt)KCGGKGGC(Mmt)GKGGGCGW-NH$_2$	
104	H-WGC(StBu)KCGGKGGC(StBu)GKGGGCGW-NH$_2$	
110	H-FAC(StBu)AA-NH$_2$	
113	H-WGC(STmp)KCGGKGGC(STmp)GKGGGCGW-NH$_2$	
114	H-WGC(Acm)KCGGKGGC(Acm)GKGGGCGW-NH$_2$	
126	H-C(Mmt)AAARGD-NH$_2$	[180]
127	H-C(Mmt)GGKLVFF-NH$_2$	[181,182]
141	H-KAAAAAAK-OH	[189,190]
142	H-KAAACAAAK-OH	
149	H-CAAAK-OH	
163	H-ACATGSTQHQACG-NH$_2$	
164	H-ACATGSTQHQCG-NH2	
165	H-ACTSTQHQCG-NH$_2$	
166	H-ACSTQHQCG-NH$_2$	
167	H-ACTQHQCG-NH$_2$	
168	H-ACQHQCG-NH$_2$	

169	H-ACHQCG-NH$_2$	
170	H-ACACG-NH$_2$	
171	H-ACHCG-NH$_2$	
172	H-ACCG-NH$_2$	
173	H-ACTHGQTQSCG-NH$_2$	[166]
174	H-NCVVGYIGERCQ-NH$_2$	[251]
175	H-CKAPETALC-NH$_2$	[252]
176	H-AGYLLGKINLKACAALAKKCL-NH$_2$	[154,155]
201	H-ACTQHQCTQHQCG-NH$_2$	
202	H-ACTQHQCTQHQCTQHQCG-NH$_2$	
203	H-ACSTQHQCSTQHQCSTQHQCG-NH$_2$	

6.1 GENERAL INFORMATION

All the reagents were purchased from commercial suppliers and used directly as indicated in the appropriate experimental procedures.

All Fmoc L-amino acids with standard side-chain protecting groups (Fmoc-Ala-OH, Fmoc-Cys(Trt)-OH, Fmoc-Gly-OH, Fmoc-His(Trt)-OH, Fmoc-Leu-OH, Fmoc-Gln(Trt)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Thr(tBu)-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Val-OH, Fmoc-Ile-OH, Fmoc-Glu(tBu)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Pro-OH), Rink Amide Pro-Tide resin, Fmoc-Lys(Boc)-Wang resin and Oxyma Pure were purchased from CEM UK Ltd (Buckingham, UK).

Trifluoroacetic acid (TFA, HPLC grade), N,N-diisopropylcarbodiimide (DIC), 2,3,4,5,6-pentafluoro benzaldehyde and silica gel were purchased from Fluorochem UK Ltd (Hadfield, UK). Acetonitrile (MeCN, both HPLC and MS grade), formic acid (MS grade), dimethylformamide (DMF), boron trifluoride etherate (BF$_3$•Et$_2$O), chloranil, petroleum ether, ethyl acetate (EtOAc), methanol (MeOH) and dichloromethane (CH$_2$Cl$_2$) were purchased from Fisher Scientific UK Ltd (Loughborough, UK).

N,N-Diisopropylethylamine (DIPEA), triethylamine (Et$_3$N), piperidine, hexafluorobenzene (HFB), triisopropylsilane (TIPS), sodium sulfate anhydrous (Na$_2$SO$_4$), sodium bicarbonate (NaHCO$_3$) dry dimethyl sulfoxide (DMSO), caesium carbonate (Cs$_2$CO$_3$), sodium hydroxide (NaOH), sodium chloride (NaCl), diethyl ether (Et$_2$O), tris(2-carboxyethyl)phosphine (TCEP), pyridine, deuterated chloroform (CDCl$_3$), deuterated methanol (CD$_3$OD), TPGS, pyrrole, Brilliant Black Bn, and the resazurin sodium salt were purchased from Sigma Aldrich Company Ltd (Gillingham, UK).

Gibco Dulbecco's Modified Eagle Medium (DMEM), New Born Calf Serum (NBCS) and TrypLE Express Enzyme (1×) without phenol red were purchased from Thermofisher UK Ltd (Runcorn, UK) as well as the 4′,6-diamidino-2-phenylindole (DAPI) containing fluoromount G mounting medium.

6.2 DESIGN OF DOE STUDIES AND MINITAB ANALYSIS

Minitab software was used to support the design of DOE studies, generate the experiment list, and perform a statistical analysis on the data. Once the variables and corresponding high and low values were established as described in sections 2.2.4, Minitab was used to generate a series of experiments (S_NAr reactions conditions) resulting from the unique combination of the variable and values. Minitab was also employed to process the results from each test (percentage of conversion from NMR analysis) and determine the effectiveness of each parameter in the outcome (conversion to 1,4-disubstitution or 1,2,4,5-tetrasubstitution product). Using Minitab, the main effects of each condition as well as the interactions between each condition were displayed graphically with graph such main effects plots, interaction plots and pareto Chart.
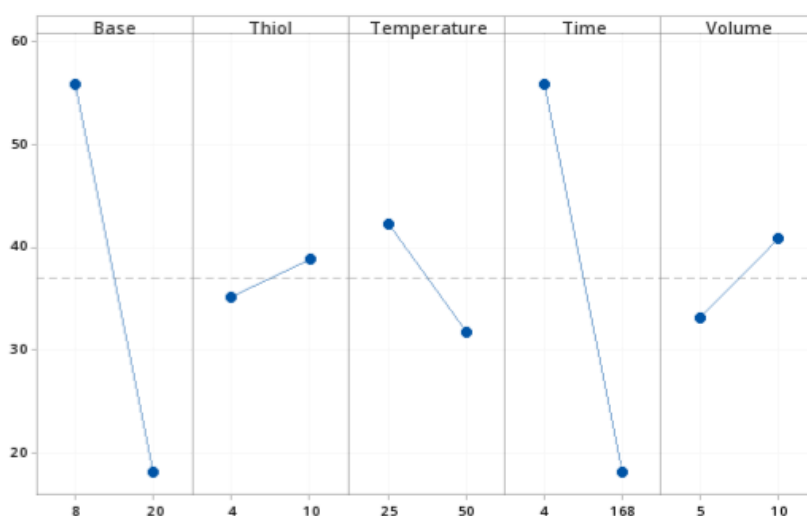
Figure 105. Main effect plots for 1,4-disubstitution. On the x axes are reported the two values explored for each variable and on the y axes the percentage of conversion to tetrasubstitution.

6.3 HPLC AND MASS SPECTROMETRY

High-resolution LC-MS characterisation was conducted using systems A, equipped with columns A, and a binary eluent system comprising H_2O and MeCN with 0.1 % formic acid as mobile phases and following gradient A. Electrospray ionisation mass spectrometry was conducted in positive ion mode (*m/z* range: 50 – 3200) using a cone voltage of 150 V, gas temperature of 325 °C (flow 10 L/min) and sheath gas temperature of 400 °C (flow 11 L/min).

Low-resolution LC-MS characterisation was conducted using systems B or C, equipped with columns B or C, and a binary eluent system comprising H_2O and MeCN with 0.1 % formic acid as mobile phases and following gradient B or C. Electrospray ionisation mass spectrometry was conducted in positive ion mode (*m/z* range: 600 – 1700) using a cone voltage of 35 V, gas temperature of 300 °C and source temperature of 100 °C. MS grade solvents (MeCN, formic acid) were purchased from Fisher Scientific and the employed ultrapure H_2O was purified with a Milli-Q® Reference Water Purification System.

Porphyrins **121**, **144**, **145**, and **148** were analysed by direct infusion in system C and using the above settings. Negative ion mode was exceptionally used for compound **144**.

HPLC analysis was conducted using systems D or E, equipped with columns C or D, and a binary eluent system comprising H_2O and MeCN with 0.1 % TFA as mobile phases and following gradients C or D.

Conjugates employed for biological assays were purified using an Agilent Infinity 1260 prep-HPLC equipped with Waters X-Bridge Peptide BEH C18 OBD Prep Column (130 A, 5 μM, 19 × 100 mm), a binary eluent system comprising H_2O and MeCN with 0.1 % TFA acid as mobile phases and following gradient E.

Peptides and conjugates for NMR or self-assembling studies were purified using a Teledyne ISCO CombiFlash NextGent 300+ System equipped with a pre-packed 5.5 mg RediSep Rf Gold C18 Column (20-40 μm) purchased from Teledyne ISCO (Lincoln, NE, USA) and a binary eluent system comprising H_2O and MeCN following gradient E.

Systems (A-E):

A.   Agilent 1260 Infinity II LC + 6530 Accurate-Mass QToF spectrometer

B.   Waters 2695 Separations Module + 996 Photodiode Array Detector + Micromass Quattro Premier Mass Spectrometer

C.   Waters 2695 Separations Module + 996 Photodiode Array Detector + Micromass LCT Premiere Mass Spectrometer

D.   Agilent 1100 Series HPLC

E.   Agilent 1200 Series HPLC

Columns (A-D):

A.   Agilent ZORBAX SB-C18 Stable Bond Analytical (5 μm, 4.6 × 150 mm)

B.   Agilent Poroshell 120 EC-C18 (2.7 μm, 3.0 × 50 mm)

C.   Waters X-Bridge Peptide BEH C18 (5 μM, 130 A, 4.6 × 150 mm)

D.   Phenomenex Gemini-NX C18 (5 μm, 110 A, 150 × 4.60 mm)

Gradients (A-F):

Table 26. Method A, 5 μL injection, 25 °C, detection at 280 nm.

Time (min)	% A (H_2O)	% B (MeCN)	Flow rate (mL/min)
0	95	5	0.8
2	95	5	0.8
18	5	95	0.8
21	5	95	0.8
23	95	5	0.8
25	95	5	0.8

Table 27. Method B, 2 μL injection, 25 °C, DAD detection between 200 nm and 600 nm.

Time (min)	% A (H₂O)	% B (MeCN)	Flow rate (mL/min)
0	90	10	0.6
10	0	100	0.6
11	90	10	0.6
13	90	10	0.6

Table 28. Method C, 10 μL injection, 25 °C, detection at 215 nm.

Time (min)	% A (H₂O)	% B (MeCN)	Flow rate (mL/min)
0	99	1	0.8
3	99	1	0.8
20	0	100	0.8
21	99	1	0.8
25	99	1	0.8

Table 29. Method D, 10 μL injection, 25 °C, detection at 400 nm.

Time (min)	% A (H₂O)	% B (MeCN)	Flow rate (mL/min)
0	90	10	0.8
20	0	100	0.8
24	0	100	0.8
25	90	10	0.8
30	90	10	0.8

Table 30. Method E, up to 700 μL injection, room temperature, detection at 400 nm.

Time (min)	% A (H₂O)	% B (MeCN)	Flow rate (mL/min)
0	70	30	8.0
20	0	100	8.0
21	70	30	8.0
25	70	30	8.0

Table 31. Method F, variable injection volume, room temperature, detection at 215 nm.

Time (min)	% A (H_2O)	% B (MeCN)	Flow rate (mL/min)
0	90	10	18.0
5	0	100	18.0
8	0	100	18.0

6.4 NUCLEAR MAGNETIC RESONANCE

All ^{1}H, ^{13}C and ^{19}F NMR spectra were recorded on a Bruker Ascend 600 Avance lll NMR, equipped with 5 mm RT probe PA BBO z-gradient, operating at 300 K and different transmitter frequencies depending on the nuclei (*i.e.* [1]H = 600.1 MHz, [13]C = 150.9 MHz, [19]F = 564.7 MHz).

Samples were dissolved with 0.6 mL of the indicated solvents and transferred in high-resolution NMR tubes.

The instrument was controlled with Topspin control PC version 3.6.2 while data analysis was carried out using TopSpin 4.1.3. Instruments and software are from Bruker UK Ltd (Coventry, UK).

For the description of compound characterization, the number of protons represented by the signal and coupling constants (*J*) were included as needed and the following abbreviations were used peak multiplicity: 's'= singlet, 'd'=doublet, 't'= triplet, dd= doublet of doublets, dt= doublet of triplets, td= triplet of doublets, and 'm'= multiplet (to denote complex pattern).

6.5 SPECTROSCOPIC TECHNIQUES

For the below-mentioned spectroscopic techniques, 500 µL of peptide-porphyrin conjugate solution (**143**, **146**, **147**, and **150**) – prepared and stored following the different self-assembly conditions described in section 3.2.5 and 6.11 – were sampled after shaking and transferred in a 1 cm quartz cuvette. Alternatively, porphyrins **121**, **144**, **145**, and **148** were dissolved in MeOH at specific concentrations as defined in previous sections.

Absorption spectra were measured using a Jasco V-650 spectrophotometer from Jasco Europe scanning between 800 nm and 300 nm. A blank sample made with the appropriate diluent was always acquired to set the baseline before analysis of the actual sample.

Steady-state fluorescence emission and anisotropy spectra were measured with a SF5 spectrofluorometer from Edinburgh Instruments Ltd. (Livingston, UK) exiting at the λ_{max} of the analysed porphyrin, acquiring between 550 nm and 800 nm and applying a 2 cm bandwidth.

Fluorescence decays were measured with a FLS920 time-correlated single-photon counting system (TCSPC) from Edinburgh Instruments Ltd. (Livingston, UK) with pulsed LED excitation at 500 nm, operating at 5 MHz repetition rate and 100 kHz photons. The Instrument Response Function was measured after each analysis. The $\chi2$ value of around 1.0–1.5 and visual inspection of residuals were considered for evaluating the quality of the fitting.

Transient absorption measurements were performed with the single-shot laser flash photolysis setup described by Abbruzzetti *et al.*[255] and using a 450 nm laser.

Photoacoustic analyses were performed with the setup described by Abbruzzetti *et al.*[219] The self-assembly sample solution were prepared at a 100 μM concentration to obtain a 0.1-0.2 absorbance intensity at 532 nm. The reference solution was prepared dissolving Brilliant Black Bn (dye content of ca. 60%) in PBS pH= 7.4 to obtain an absorbance intensity comparable to the self-assembly sample solution at 532 nm.

All these spectroscopy experiments were performed at the University of Parma.

6.6 MICROSCOPY TECHNIQUES

6.6.1 Scanning Electron Microscopy

50 μL of peptide (**141** and **142**) or peptide-porphyrin conjugate solution (**143**, **146**, **147**, and **150**) – after preparation and incubation as described in section 3.2.5 and 6.11 – were placed on top of an aluminium pin stub and left to evaporate at room temperature overnight.

The stub was then coated with gold prior to analysis on a FEI Inspect S model with accelerating voltage of 0.1-30 kv range and a nominal magnification up to 100 kx.

Acquired images were analysed with ImageJ 1.51 software.

6.6.2 Transmission Electron Microscopy

5 μL of peptide (**141** and **142**) or peptide-porphyrin conjugate solution (**143**, **146**, **147**, and **150**) – after preparation and incubation as described in section 3.2.5 and 6.11 – were placed directly on a copper grid covered by Formvar/Carbon film (400 mesh), let to absorb for 5 min and then the excess solution was removed with a filter paper.

For peptides, 10 μL of negative staining reagent (1% phosphotungstic acid solution) were added afterwards, let to absorb for 5 min and then the excess solution was removed by capillarity with a filter paper.

For conjugates, no stain was necessary.

In both cases, grids were dried in air for at least 15 min before analysis.

The grids were observed with a FEI Morgagni 268 microscope, operating at 100 kV and under high vacuum, with a magnification up to 140 kx. Images were recorded using iTEM FEI imaging software via a Megaview II digital camera.

Acquired images were analysed with ImageJ 1.51 software.

6.6.3 Fluorescence Microscopy

40 μL of peptide-porphyrin conjugate solution (**143**, **146**, **147**, and **150**) – after preparation and incubation as described in section 3.2.5 and 6.11 – were placed on a glass slide to fill a channel created with a glass cover. After 30 minutes, the sample solution was washed away with 200 μL buffer and the channel was sealed.

The glass slides were observed with a Nanoimager microscope from ONI (Oxford, UK) at the University of Parma using TIRF mode, camera exposure of 100 ms and laser excitation at 10%.

Acquired images were analysed with ImageJ 1.51 software.

6.7 DENSITY LIGHT SCATTERING

200 μL of peptide (**141** and **142**) or peptide-porphyrin conjugate solution (**143**, **146**, **147**, and **150**) – after preparation and incubation as described in section 3.2.5 and 6.11 – were sampled after shaking and transferred in a reduced-volume plastic cuvette.

Sizes of peptide and peptide-porphyrin self-assemblies were determined at room temperature by dynamic light scattering, using a Zetasizer instrument (Malvern Ltd., United Kingdom).

The dispersant was set as water, the equilibration time was set to 15 seconds, and – using manual measurements settings – number of runs and run duration were set to 10 and 3 seconds respectively.

Data were processed using a Zetasizer software version 8.02 (Malvern Ltd., United Kingdom).

6.8 PORPHYRIN SYNTHESIS AND CHARACTERIZATION

6.8.1 5,10,15,20-tetrakis(pentafluorophenyl)porphyrin (121)
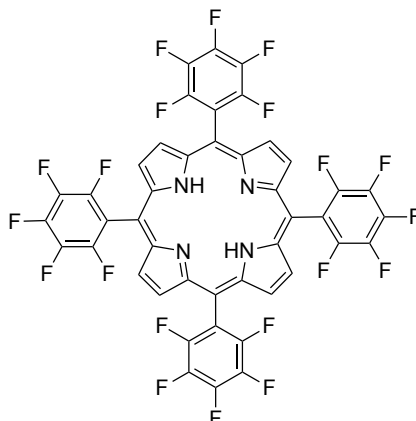
In a 3-neck round bottom flask, CH_2Cl_2 (600 mL) was stirred for 5 min under N_2 flow. 2,3,4,5,6-Pentafluorobenzaldehyde (2.5 g, 12.7 mmol) and freshly distilled pyrrole (1 mL, 14.4 mmol) were transferred into the flask and the mixture stirred under N_2 at room temperature for 30 min. $BF_3 \cdot Et_2O$ (0.5 mL, 4 mmol) was added, and the reaction heated to 40 °C and protected from light. After 4 h, chloranil (3.5 g, 14 mmol) was added and the reaction heated to reflux overnight and protected from light. After 18 h, the reaction mixture was cooled to room temperature and washed with a saturated solution of $NaHCO_3$ (3 × 200 mL). The organic phase was dried over anhydrous Na_2SO_4, filtered, and evaporated to dryness under reduced pressure. The crude product was then loaded on a silica gel column and filtered using petroleum ether/CH_2Cl_2 (6:4). Evaporation of the eluate afforded a purple solid, which was crystallised from CH_2Cl_2/MeOH to obtain 5,10,15,20-tetrakis(pentafluorophenyl) porphyrin as purple crystals (1.4 mg, 1.44 mmol, 9%).[78]

Formula: $C_{44}H_{10}F_{20}N_4$

Calculated Exact Mass: 974.06

Calculated Molecular Weight: 974.56

^{1}H NMR (CDCl$_3$) δ, ppm: 8.92 (s, 8H, β-H), -2.90 (s, 2H, NH$_{int}$)

^{13}C NMR (CDCl$_3$) δ, ppm: 147.36, 145.69, 138.44, 136.76, 115.53, 103.66

^{19}F NMR (CDCl$_3$) δ, ppm: -136.55 (dd, J_o = 7 Hz, J_p = 25 Hz, F$_o$) -151.26 (t, 4F, J_o = 20.5 Hz, F$_p$), -161.30 (dt, 8F, J_p = 8 Hz, J_o = 23 Hz, F$_m$)

m.p. (°C): > 300

UV/vis (MeOH, nm) λ$_{max}$ (Soret): 408; Q bands: 505, 585

ESI-MS (*m/z*): 975.065 [M+H]$^+$

This characterisation is in agreement with previous literature data.[256]

6.8.2 5,15-bis(pentafluorophenyl)porphyrin (148)

2,3,4,5,6-pentafluorobenzaldehyde (1.6 g, 8 mmol) was transferred into an oven-dried round bottom flask and conditioned with argon flow for a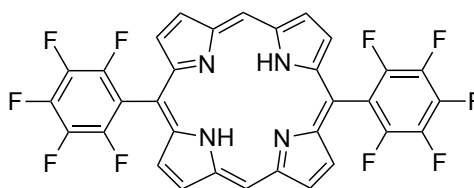 few minutes before the addition of freshly distilled pyrrole (111 mL, 1.6 mol). Finally, TFA (100 μL, 0.8 mmol) was added under magnetic stirring. After 10 min, the reaction was quenched with NaOH (75 mL, 0.2 M). EtOAc (80 mL) was added to the reaction mixture and the organic phase washed with H_2O (3 × 100 mL) and brine (100 mL), dried over Na_2SO_4 and evaporated to dryness under reduced pressure. The obtained yellow oil was loaded onto a silica gel column and separated by flash-chromatography using $EtOAc/CH_2Cl_2/Et_3N$ (8:2:1). After evaporation of the eluate, the product was crystallised from hexane to obtain (pentafluorophenyl)dipyrromethene (**226**) as a white solid (1.52 g, 4.9 mmol, 60%).[257]

In a 3-neck round bottom flask, equipped with a condenser, a dropping funnel, an argon inlet adaptor, and a magnetic stirring bar, (pentafluorophenyl)dipyrromethane (480 mg, 1.54 mmol) was dissolved in CH_2Cl_2 (420 mL). Trimethylorthoformate (12 mL, 110.8 mmol) was added in one portion under argon flow. A solution of TFA (4.5 mL, 58.5 mmol) in CH_2Cl_2 (150 mL) was added dropwise over 30 min at room temperature and in the dark. The reaction mixture was stirred at room temperature for 4 h before quenching the reaction with pyridine (16 mL, 0.2 mmol). After 19 h, compressed air was bubbled into the reaction mixture for 15 min and the reaction stirred for 4 h more at room temperature, in the air and light. The mixture was then evaporated to dryness and the crude product was purified with silica gel flash chromatography using CH_2Cl_2/petroleum ether (7:3) as mobile phase. The eluate was evaporated and the product crystallised from CH_2Cl_2/MeOH yielding 5,15-bis(pentafluorophenyl)porphyrin as dark solid (55 mg, 0.085 mmol, 5.5%).[80]

Formula: $C_{32}H_{12}148N_4$

Calculated Exact Mass: 642.09

Calculated Molecular Weight: 642.46

^{1}H NMR (CDCl$_3$) δ, ppm: 10.39 (s, 2H), 9.49 (d, J = 4.5 Hz, 4H), 9.00 (d, J = 4 Hz, 4H), -3.25 (s, 2H)

^{19}F NMR (CDCl$_3$) δ, ppm: -136.65 (dd, J_o = 8 Hz, J_p = 24 Hz, F_o) -152.26 (t, 4F, J_o = 21 Hz, F_p), -161.84 (dt, 8F, J_p = 8.5 Hz, J_o = 22 Hz, F_m)

m.p. (°C): > 300

UV-Vis (CH$_2$Cl$_2$, nm) λ$_{max}$ (Soret): 403; Q bands: 500, 533, 575, 629

ESI-MS (m/z): 643.387 [M+H]$^+$

These characterisations are in agreement with previous literature data.[80,257]

6.8.3 Palladium 5,10,15,20-tetrakis(pentafluorophenyl)porphyrin (**144**)
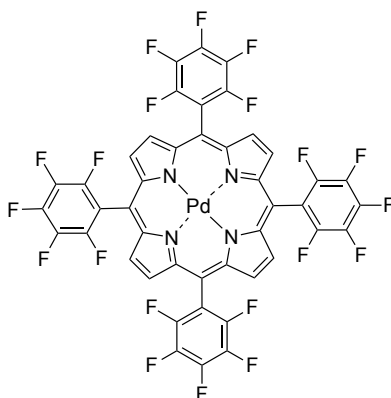
In a round bottom flask, equipped with a condenser, a $CaCl_2$ guard tube, and a magnetic stirring bar, 5,10,15,20-tetrakis(pentafluorophenyl)porphyrin (200 mg, 0.2 mmol) and $PdCl_2(PhCN)_2$ (230 mg, 0.6 mmol) were dissolved in PhCN (420 mL) under inert atmosphere. The reaction mixture was heated to reflux and stirred at room temperature for 24 h. The mixture was then cooled to room temperature and the product was precipitated by addition of cold MeOH. The crude product was purified using silica gel flash chromatography using CH_2Cl_2/PE (7:3) as mobile phase. The eluate was evaporated, and the product crystallised from CH_2Cl_2/MeOH/H_2O yielding Palladium 5,10,15,20-tetrakis(pentafluorophenyl)porphyrin (**144**) as red crystals (186 mg, 0.173 mmol, 87%).[214]

Formula: $C_{44}H_8F_{20}N_4Pd$

Calculated Exact Mass: 1077.95

Calculated Molecular Weight: 1078.96

^{1}H NMR (CDCl$_3$) δ, ppm: 8.88 (s, 8H, β-H)

^{13}C NMR (CDCl$_3$) δ, ppm: 29.7, 105.55, 115.22, 131.26, 141.83

^{19}F NMR (CDCl$_3$) δ, ppm: -136.41 (dd, J_o =7.5 Hz, J_p =23.5 Hz, F$_o$), -151.18 (t, 4F, J_o =21 Hz, F$_p$), -161.22 (dt, 8F, J_p =7.5 Hz, J_o =22.5 Hz, F$_m$)

m.p. (°C): > 300

UV-Vis (MeOH, nm) λ$_{max}$ (Soret): 405; Q bands: 518, 551

ESI-MS (*m/z*): 1113.5883 [M+Cl]$^-$

This characterisation is in agreement with previous literature data.[214,215]

6.8.4 Platinum 5,10,15,20-tetrakis(pentafluorophenyl)porphyrin (145)
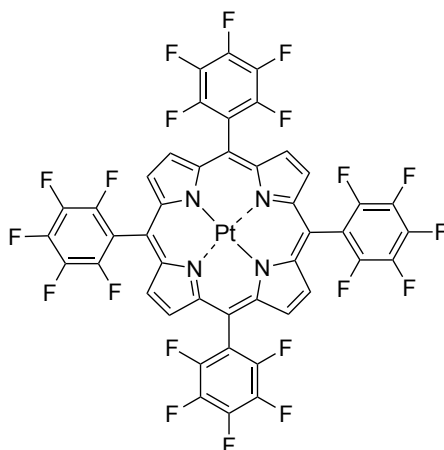
$PtCl_2(PhCN)_2$ was prepared by adding anhydrous $PtCl_2$ (100 mg, 1 mmol) in small portions of benzonitrile (3 mL) previously heated at 100 °C. 10 minutes after the last addition, the solution was filtered while still hot and the filtrate was allowed to cool down to room temperature before precipitating the product of interest with slow addition of petroleum ether (30 mL).The precipitate was filtered, washed with petroleum ether (3x1 mL) and dried in vacuum to obtain $PtCl_2(PhCN)_2$ as a pale-yellow powder (188 mg, 0.45 mmol, 45%).

5,10,15,20-Tetrakis(pentafluorophenyl)porphyrin (100 mg, 0.1 mmol), $PtCl_2(PhCN)_2$ (125 mg, 0.3 mmol) and Na(OAc) (0.5 mmol, 41 mg) were placed in a 15 mL microwave tube and dissolved in chlorobenzene (8 mL) under magnetic stirring and N_2 flow. After 10 min, the tube was placed in a CEM Microwave and heated at 132 °C for 30 min. The reaction mixture was then transferred to a round bottom flask, concentrated under reduced pressure and filtered on silica gel using hexane/CH_2Cl_2 as mobile phases. The filtered solution was then evaporated to dryness and the red solid redissolved in CH_3CN and a few drops of HCl 37%. This solution was stirred at 60 °C for 1 h. When at room temperature, H_2O was added to precipitate **145** which was recovered by filtration with filter paper as red powder (65 mg, 0.06 mmol, 60%).[214]

Formula: $C_{44}H_8F_{20}N_4Pt$

Calculated Exact Mass: 1167.01

Calculated Molecular Weight: 1167.63

^1H NMR (CDCl$_3$) δ, ppm: 8.83 (s, 8H, β-H)

^{13}C NMR (CDCl$_3$) δ, ppm: 29.5, 106.1, 114.8, 130.7, 141.4

^{19}F NMR (CDCl$_3$) δ, ppm: -136.36 (dd, J_o =7.4 Hz, J_p =23.3 Hz, F$_o$), -151.03 (t, 4F, J_o =20.7 Hz, F$_p$), -161.06 (dt, 8F, J_p =7.0 Hz, J_o =22.3 Hz, F$_m$)

m.p. (°C): > 300

UV-Vis (solvent, nm) (MeOH, nm) λ$_{max}$ (Soret): 390; Q bands: 510, 542

This characterisation is in agreement with previous literature data.[212,213]

6.9 SOLID PHASE PEPTIDE SYNTHESIS

6.9.1 General procedure for SPPS

Each linear peptide sequence was prepared using automated Fmoc-solid phase peptide synthesis methods on a CEM Liberty Blue microwave-assisted peptide synthesizer.

Solid-phase synthesis was conducted on a 0.1 mmol scale using Rink amide Pro-Tide resin (179 mg, 0.56 mmol/g loading), employing the required Fmoc-L-amino acids (0.2 M in DMF, 5 eq.), DIC (1 M in DMF; 10 eq.) as activator, Oxyma Pure (1 M in DMF, 5 eq.) as racemisation suppressor, and piperidine (20% v/v in DMF; 4 mL) as deprotection reagent. Standard coupling procedures employed single coupling of each amino acid (2.5 minutes, 90 °C) and Fmoc-deprotection (2 min, 90 °C). Racemisation-prone amino acids bearing thermally-sensitive protecting groups, for example Fmoc-Cys(Trt)-OH, were coupled under milder conditions (10 min, 50 °C).

Following on-resin synthesis of the appropriate sequence, the resin was transferred to 10 mL syringes with frits and shrunk by washing with Et_2O. Finally, peptides were cleaved from the resin as the C-terminal amide by treatment with a cleavage cocktail of TFA, TIPS and H_2O (8:1:1 v/v) with shaking at room temperature for 4 h. Peptides were precipitated from cleavage solutions by dropwise addition into cold Et_2O followed by centrifugation. The resulting pellet was successively suspended in cold Et_2O, vortexed and centrifuged twice further.

The solids obtained after Et_2O removal and its complete evaporation were dissolved in H_2O, MeCN or MeOH with 0.1% TFA v/v (or a mixture of these solvents, depending on the solubility) and analysed by HPLC and LC-MS as described below. Unless otherwise stated in the main text (peptide **176**), peptides were used as crude for successive reactions without further chromatographic purification.

Peptides **141**, **142** and **149** were synthesised starting with Fmoc-Lys(Boc)-Wang resin and following the procedure described above to obtain the final peptide with carboxylic acid C-terminus.

6.9.2 Characterisation of peptides

Peptide **69**

Sequence: H-TyrCysGlyGlyGlyCysAlaLeu-NH2

Formula: $C_{30}H_{47}N_9O_9S_2$

Calculated Exact Mass: 741.29

Calculated Molecular Weight: 741.88

ESI-MS (*m/z*): 1483.59326 [2M+H]+, 764.28445 [M+Na]+, 742.30260 [M+H]+

RP-HPLC RT (System A, Column A, Method A): 7.8 min

Purity by analytical HPLC (crude): 90%

Peptide 77

Sequence: H-AlaCysThrGlySerThrGlnHisGlnCysGly-NH2

Formula: $C_{40}H_{66}N_{16}O_{16}S_2$

Calculated Exact Mass: 1090.43

Calculated Molecular Weight: 1091.18

ESI-MS (*m/z*): 1091.7626 [M+H]$^+$, 546.3910 [M+2H]$^{2+}$

RP-HPLC RT (System E, Column D, Method C): 7.1 min

Purity by analytical HPLC (crude): 80%

Peptide 78

Sequence: AcAlaCysThrGlySerThrGlnHisGlnCysGly-NH2

Formula: $C_{42}H_{68}N_{16}O_{17}S_2$

Calculated Exact Mass: 1132.44

Calculated Molecular Weight: 1133.22

ESI-MS (*m/z*): 1133.09 [M+H]$^+$

RP-HPLC RT (System B, Column C, Method D): 2 min

Peptide 98

Sequence: H-TrpGlyCysLysCysGlyGlyLysGlyGlyCysGlyLysGlyGlyGlyCysGlyTrp-NH$_2$

Formula: $C_{72}H_{109}N_{25}O_{19}S_4$

Calculated Exact Mass: 1755.72

Calculated Molecular Weight: 1757.06

ESI-MS (*m/z*): 1757.72527 [M+H]$^+$, 878.86831 [M+2H]$^{2+}$, 586.24951 [M+3H]$^{3+}$, 439.94057 [M+4H]$^{4+}$

RP-HPLC RT (System E, Column D, Method C): 9.4 min

Purity by analytical HPLC (crude): 42%

Peptide 104

Sequence: H-TrpGlyCys(StBu)LysCysGlyGlyLysGlyGlyCys(StBu)GlyLysGlyGlyGlyCysGlyTrp-NH$_2$

Formula: $C_{80}H_{125}N_{25}O_{19}S_6$

Calculated Exact Mass: 1931.79

Calculated Molecular Weight: 1933.40

ESI-MS (*m/z*): 966.9777 [M+2H]$^{2+}$, 644.9695 [M+3H]$^{3+}$, 484.2317 [M+4H]$^{4+}$

RP-HPLC RT (System E, Column D, Method C): 11.2 min

Purity by analytical HPLC (crude): 47%

Peptide 114

Sequence: H-TrpGlyCys(Acm)LysCysGlyGlyLysGlyGlyCys(Acm)GlyLysGlyGlyGlyCysGlyTrp-NH$_2$

Formula: $C_{78}H_{119}N_{27}O_{21}S_4$

Calculated Exact Mass: 1897.80

Calculated Molecular Weight: 1899.22

ESI-MS (m/z): 1266.9023 $[2M+3H]^{3+}$, 949.9133 $[M+2H]^{2+}$, 633.6168 $[M+3H]^{3+}$, 475.4559 $[M+4H]^{4+}$

RP-HPLC RT (System E, Column D, Method C): 9.3 min

Purity by analytical HPLC (crude): 35%

Peptide *110*

Sequence: H-PheAla(StBu)CysAlaAla-NH$_2$

Formula: $C_{25}H_{40}N_6O_5S_2$

Calculated Exact Mass: 568.25

Calculated Molecular Weight: 568.75

ESI-MS (m/z): 1137.6108 $[2M+H]^+$, 569.2343 $[M+H]^{2+}$

RP-HPLC RT (System C, Column C, Method D): 8.9 min

Peptide *126*

Sequence: H-CysAlaAlaAlaArgGlyAsp-NH$_2$

Formula: $C_{24}H_{43}N_{11}O_9S$

Calculated Exact Mass: 661.30

Calculated Molecular Weight: 661.74

ESI-MS (m/z): 662.3265 $[M+H]^+$, 331.6495 $[M+2H]^+$

RP-HPLC RT (System C, Column C, Method D): 1.5 min

Peptide *127*

Sequence: H-CysGlyGlyLysLeuValPhePhe-NH$_2$

Formula: $C_{42}H_{64}N_{10}O_8S$

Calculated Exact Mass: 868.46

Calculated Molecular Weight: 869.10

ESI-MS (m/z): 879.4152 $[M+H]^+$

RP-HPLC RT (System A, Column A, Method A): 10.7 min

Peptide *141*

Sequence: H-LysAlaAlaAlaAlaAlaAlaLys-OH

Formula: $C_{30}H_{56}N_{10}O_9$

Calculated Exact Mass: 700.42

Calculated Molecular Weight: 700.84

ESI-MS (m/z): 701.5891 $[M+H]^+$, 351.2871 $[M+2H]^{2+}$

RP-HPLC RT (System D, Column D, Method C): 2.3 min

Purity by analytical HPLC (crude): 95%

Peptide 142

Sequence: H-LysAlaAlaAlaCysAlaAlaAlaLys-OH

Formula: $C_{18}H_{61}N_{11}O_{10}S$

Calculated Exact Mass: 803.43

Calculated Molecular Weight: 803.98

ESI-MS (m/z): 804.4306 $[M+H]^+$, 402.7319 $[M+2H]^{2+}$

RP-HPLC RT (System D, Column D, Method C): 2.3 min

Purity by analytical HPLC (crude): 65%

Peptide 149

Sequence: H-CysAlaAlaAlaLys-OH

Formula: $C_{18}H_{34}N_6O_6S$

Calculated Exact Mass: 462.23

Calculated Molecular Weight: 462.57

ESI-MS (m/z): 463.2142 $[M+H]^+$

RP-HPLC RT (System D, Column D, Method C): 2.3 min

Purity by analytical HPLC (crude): 85%

Peptide 163

Sequence: H-AlaCysAlaThrGlySerThrGlnHisGlnAlaCysGly-NH$_2$

Formula: $C_{46}H_{76}N_{18}O_{18}S_2$

Calculated Exact Mass: 1232.50

Calculated Molecular Weight: 1233.34

ESI-MS (m/z): 1234.0126 $[M+H]^+$, 617.5123 $[M+2H]^{2+}$, 412.0273 $[M+3H]^{3+}$

RP-HPLC RT (System D, Column D, Method C): 7.6 min

Purity by analytical HPLC (crude): 58%

Peptide 164

Sequence: H-AlaCysAlaThrGlySerThrGlnHisGlnCysGly-NH$_2$

Formula: C$_{43}$H$_{71}$N$_{17}$O$_{17}$S$_2$

Calculated Exact Mass: 1161.47

Calculated Molecular Weight: 1162.26

ESI-MS (*m/z*): 1162.5264 [M+H]$^+$, 581.7865 [M+2H]$^{2+}$

RP-HPLC RT (System D, Column D, Method C): 7.4 min

Purity by analytical HPLC (crude): 58%

Peptide 165

Sequence: H-AlaCysThrSerThrGlnHisGlnCysGly-NH$_2$

Formula: C$_{38}$H$_{63}$N$_{15}$O$_{15}$S$_2$

Calculated Exact Mass: 1033.41

Calculated Molecular Weight: 1034.13

ESI-MS (*m/z*): 1034.5149 [M+H]$^+$, 517.7452 [M+2H]$^{2+}$

RP-HPLC RT (System D, Column D, Method C): 7.2 min

Purity by analytical HPLC (crude): 59%

Peptide 166

Sequence: H-AlaCysSerThrGlnHisGlnCysGly-NH$_2$

Formula: C$_{34}$H$_{56}$N$_{14}$O$_{13}$S$_2$

Calculated Exact Mass: 932.36

Calculated Molecular Weight: 933.03

ESI-MS (*m/z*): 933.6769 [M+H]$^+$, 467.3278 [M+2H]$^{2+}$

RP-HPLC RT (System D, Column D, Method C): 4.8 min

Purity by analytical HPLC (crude): 82%

Peptide 167

Sequence: H-AlaCysThrGlnHisGlnCysGly-NH$_2$

Formula: C$_{31}$H$_{51}$N$_{13}$O$_{11}$S$_2$

Calculated Exact Mass: 845.33

Calculated Molecular Weight: 845.95

ESI-MS (*m/z*): 846.7586 [M+H]$^+$, 423.8545 [M+2H]$^{2+}$

RP-HPLC RT (System D, Column D, Method C): 4.2 min

Purity by analytical HPLC (crude): 80%

Peptide 168

Sequence: H-AlaCysGlnHisGlnCysGly-NH$_2$

Formula: $C_{27}H_{44}N_{12}O_9S_2$

Calculated Exact Mass: 744.28

Calculated Molecular Weight: 744.84

ESI-MS (m/z): 745.4860 [M+H]$^+$, 373.2356 [M+2H]$^{2+}$

RP-HPLC RT (System D, Column D, Method C): 3.1 min

Purity by analytical HPLC (crude): 91%

Peptide 169

Sequence: H-AlaCysHisGlnCysGly-NH$_2$

Formula: $C_{22}H_{36}N_{10}O_7S_2$

Calculated Exact Mass: 616.22

Calculated Molecular Weight: 616.71

ESI-MS (m/z): 1233.9468 [2M+H]$^+$, 617.5123 [M+H]$^+$, 309.2538 [M+2H]$^{2+}$

RP-HPLC RT (System D, Column D, Method C): 2.9 min

Purity by analytical HPLC (crude): 84%

Peptide 170

Sequence: H-AlaCysAlaCysGly-NH$_2$

Formula: $C_{14}H_{26}N_6O_5S_2$

Calculated Exact Mass: 422.14

Calculated Molecular Weight: 422.52

ESI-MS (m/z): 845.3962 [2M+H]$^+$, 423.1778 [M+H]$^+$

RP-HPLC RT (System D, Column D, Method C): 3.3 min

Purity by analytical HPLC (crude): 70%

Peptide 171

Sequence: H-AlaCysHisCysGly-NH$_2$

Formula: $C_{17}H_{28}N_8O_5S_2$

Calculated Exact Mass: 488.16

Calculated Molecular Weight: 488.58

ESI-MS (m/z): 489.3728 [2M+H]$^+$

RP-HPLC RT (System D, Column D, Method C): 2.7 min

Purity by analytical HPLC (crude): 72%

Peptide 172

Sequence: H-AlaCysCysGly-NH$_2$

Formula: C$_{11}$H$_{21}$N$_5$O$_4$S$_2$

Calculated Exact Mass: 351.10

Calculated Molecular Weight: 351.44

ESI-MS (m/z): 1054.7456 [3M+H]$^+$, 703.5029 [2M+H]$^+$, 352.2367 [M+H]$^+$

RP-HPLC RT (System D, Column D, Method C): 2.8 min

Purity by analytical HPLC (crude): 94%

Peptide 173

Sequence: H-AlaCysThrHisGlyGlnThrGlnSerCysGly-NH$_2$

Formula: C$_{40}$H$_{66}$N$_{16}$O$_{16}$S$_2$

Calculated Exact Mass: 1090.43

Calculated Molecular Weight: 1091.18

ESI-MS (m/z): 1091.6542 [M+H]$^+$, 546.3252 [M+2H]$^{2+}$

RP-HPLC RT (System D, Column D, Method C): 7.2 min

Purity by analytical HPLC (crude): 84%

Peptide 174

Sequence: H-AsnCysValValGlyTyrIleGlyGluArgCysGln-NH$_2$

Formula: C$_{55}$H$_{90}$N$_{18}$O$_{17}$S$_2$

Calculated Exact Mass: 1338.62

Calculated Molecular Weight: 1339.55

ESI-MS (m/z): 1340.0433 [M+H]$^+$, 670.4722 [M+2H]$^{2+}$, 447.3465 [M+3H]$^{3+}$

RP-HPLC RT (System D, Column D, Method C): 10.2 min

Purity by analytical HPLC (crude): 75%

Peptide 175

Sequence: H-CysLysAlaProGluThrAlaLeuCys-NH$_2$

Formula: $C_{38}H_{67}N_{11}O_{12}S_2$

Calculated Exact Mass: 933.44

Calculated Molecular Weight: 934.14

ESI-MS (m/z): 934.4890 [M+H]$^+$, 467.7361 [M+2H]$^{2+}$

RP-HPLC RT (System D, Column D, Method C): 9.3 min

Purity by analytical HPLC (crude): 85%

Peptide 176

Sequence: H-AlaGlyTyrLeuLeuGlyLysIleAsnLeuLysAlaCysAlaAlaLeuAlaLysLysCysLeu-NH$_2$

Formula: $C_{98}H_{173}N_{27}O_{23}S_2$

Calculated Exact Mass: 2160.26

Calculated Molecular Weight: 2161.75

ESI-MS (m/z): 1081.6011 [M+2H]$^{2+}$, 721.4022 [M+3H]$^{3+}$, 541.3055 [M+4H]$^{4+}$

RP-HPLC RT (System D, Column D, Method C): 12.0 min

Purity by analytical HPLC (crude): 82%

Peptide 201

Sequence: H-AlaCysThrGlnHisGlnCysThrGlnHisGlnCysGly-NH$_2$

Formula: $C_{54}H_{86}N_{22}O_{19}S_3$

Calculated Exact Mass: 1442.56

Calculated Molecular Weight: 1443.60

ESI-MS (m/z): 1443.7104 [M+H]$^+$, 722.2967 [M+2H]$^{2+}$, 481.8657 [M+3H]$^{3+}$

RP-HPLC RT (System D, Column D, Method C): 7.4 min

Purity by analytical HPLC (crude): 85%

Peptide 202

Sequence: H-AlaCysThrGlnHisGlnCysThrGlnHisGlnCysThrGlnHisGlnCysGly-NH$_2$

Formula: $C_{77}H_{121}N_{31}O_{27}S_4$

Calculated Exact Mass: 2039.79

Calculated Molecular Weight: 2041.25

ESI-MS (m/z): 1021.3347 [M+2H]$^{2+}$, 680.9044 [M+3H]$^{3+}$, 511.1588 [M+4H]$^{4+}$

RP-HPLC RT (System D, Column D, Method C): 7.6 min

Purity by analytical HPLC (crude): 89%

*Peptide **203***

Sequence: H-AlaCysSerThrGlnHisGlnCysSerThrGlnHisGlnCysSerThrGlnHisGlnCysGly-NH$_2$

Formula: C$_{86}$H$_{136}$N$_{34}$O$_{33}$S$_4$

Calculated Exact Mass: 2300.89

Calculated Molecular Weight: 2302.48

ESI-MS (*m/z*): 1152.7425 [M+2H]$^{2+}$, 768.8237 [M+3H]$^{3+}$, 576.8010 [M+4H]$^{4+}$

RP-HPLC RT (System D, Column D, Method C): 7.6 min

Purity by analytical HPLC (crude): 72%

6.10 CONJUGATION REACTIONS

6.10.1 *S$_N$Ar on HFB with NAC for OVAT experiments*

N-acetylcysteine (28.3 mg), 1.5 mL of solvent (THF, CH$_3$CN, DMF, DMSO), the base (150 µL of DIPEA, 281.8 mg of Cs$_2$CO$_3$, or 130 µL DBU) and HFB (5 µL) were added to a 1.5 mL screw top vial and the mixture was stirred at room temperature for 4 h using stirrer bars and magnetic stirrer-plates.

The reaction outcomes were obtained by direct analysis of the reaction mixture (without reaction work-up) with ^{19}F NMR to measure the percentage of unreacted HFB and substitution products (conjugates **54** and **55**).

6.10.2 *S$_N$Ar on HFB with NAC for OVAT experiments in H$_2$O*

N-acetylcysteine (106 mg), 1.5 mL of TPGS-1000 or TPGS-750M in H$_2$O, the base (90 µL of Et$_3$N or 211.5 of Cs$_2$CO$_3$) and HFB (15 µL) were added to a 1.5 mL screw top vial and the mixture was stirred at room temperature or 45°C using stirrer bars and magnetic stirrer-plates.

The reaction outcomes were obtained by direct analysis of the reaction mixture (without reaction work-up) with ^{19}F NMR to measure the percentage of unreacted HFB and substitution products (conjugates **54** and **55**).

6.10.3 *S$_N$Ar on HFB with NAC for DOE experiments*

N-acetylcysteine (113 mg or 282 mg), Cs$_2$CO$_3$ (450 mg or 1129 mg), DMSO (5 mL or 10 mL), and HFB (20 µL) were added to a 10 mL fritted syringe and the mixture was stirred at room temperature or 50 °C for 4 h or 168 h using stirrer bars and magnetic stirrer-plates. When stirring at 50 °C was required, a water bath with temperature monitor was used.

The reaction outcomes were obtained by direct analysis of the reaction mixture (without reaction work-up) with ^{19}F NMR to measure the percentage of unreacted HFB and substitution products (conjugates **54** and **55**).

6.10.4 *Peptide cyclisation with HFB in solution*

The crude peptide **69** (10 mg, 0.0135 mmol), 1.5 mL of solvent (CH$_3$CN or DMSO), Cs$_2$CO$_3$ (88 mg, 20 mmol) or DBU (40 mg, 0.27 mmol), eventually TCEP (0.05 mmol, 14.3 mg) and finally HFB (50 mg,

0.27 mmol) were added to a 1.5 mL screw top vial and the mixture was stirred at room temperature for 24 h, using stirrer bars and magnetic stirrer-plates, with sampling at intermediate timepoints.

The reaction outcomes were obtained by direct analysis of the reaction mixture (without reaction work-up) with ^{19}F NMR to measure the percentage of unreacted HFB and substitution product (conjugate **70**). LC-MS analysis was also performed after dilution of 20 μL of reaction mixture with 1 mL H_2O + 0.1% TFA.

6.10.5   *Peptide cyclisation with HFB on-resin*

After SPPS of peptide **76**, the mass of dry resin corresponding to 0.0135 mmol of peptide was weighed directly in an empty 10 mL fritted-syringe and swollen for 5-10 minutes in 2 mL of DMF.

Mmt-protected cysteine residues were selectively deprotected on-resin with 5x5 mL of a solution comprising CH_2Cl_2:TFA:TIPS (95:2:3), for 5 min at room temperature. The cleavage solution was discarded, and the resin washed successively with CH_2Cl_2 and DMF.

A solution of HFB (45 uL, 0.4 mmol) in DIPEA (1 mL) was drawn up into the syringe containing the swollen peptide resin and shaken at room temperature for 18 h. Subsequently, the reaction mixture was discarded, the resin washed with CH_2Cl_2 and DMF and shrunk with Et_2O.

The cyclic product was finally cleaved and analysed following the procedure described in sections 6.4 and 6.8 respectively to obtain conjugate **70**.

6.10.6   *Peptide-FITC Conjugation*

Mmt deprotection and HFB cyclisation of peptide **81** were performed following the procedures described in section 6.10.4. After that, the Lys(Alloc) on the spacer was deprotected using Pd(0) and phenylsilane in CH_2Cl_2, for 3x25 min at room temperature. FITC was then conjugated to the free amine of intermediate **82** in the presence of pyridine, DMF and CH_2Cl_2 (12:7:5), overnight at room temperature.

The uptake of final product **83** into SCCIC8 cells was monitored by fluorescent microscopy with a Leica microscope following 24h incubation of cells with 15 uM solution of conjugate.

6.10.7   *S_NAr on porphyrin **121** with NAC or GSH*

5,10,15,20-tetrakis(pentafluorophenyl)porphyrin (**121**, 15 mg, 0.015 mmol), N-acetylcysteine or GSH (0.015/0.075 mmol), and Cs_2CO_3 (39 mg, 0.12 mmol) or DIPEA (21 μL, 0.12 mmol) were added to a 5 mL screw top vial and dissolved in DMSO or DMF (3 mL). The mixture was stirred at room temperature using stirrer bars and magnetic stirrer-plates.

The reaction outcomes were obtained by direct analysis of the reaction mixture (without reaction work-up) with ^{19}F NMR at specific timepoints (15 min, 30 min, 1 h, 2 h, 5 h, 24 h, 48 h) to measure the percentage of unreacted HFB and substitution products (conjugates **122-125**).

6.10.8   *Peptide conjugation with porphyrin **121** on resin*

After SPPS of peptides **126** and **127**, the mass of dry resin corresponding to 0.05 mmol of peptide was weighed directly in an empty 10 mL fritted-syringe and swollen for 5-10 minutes in CH_2Cl_2.

Methoxy-trityl protected cysteine residues were selectively deprotected on-resin with 5x5 mL of a solution comprising CH_2Cl_2:TFA:TIPS (95:2:3), for 5 min at room temperature. The cleavage solution was discarded, and the resin washed successively with CH_2Cl_2.

A solution of porphyrin **121** (200 mg, 0.2 mmol) and DBU (150 μL, 1 mmol) in 10 mL CH_2Cl_2 was drawn up into the syringe containing the swollen peptide resin and shaken at room temperature for 20 min. Subsequently, the reaction mixture was blown down and the porphyrin recovered after CH_2Cl_2 evaporation, silica gel filtration and crystallisation with MeOH. The resin was washed with CH_2Cl_2 and shrunk with Et_2O.

The conjugated products (conjugates **128** and **129**) were finally cleaved and analysed following the procedure described above.

The same procedure was also applied using DIPEA (100 μL, 1 mmol) as base and DMF (10 mL) as solvent.

6.10.9 Porphyrin compatibility with Rink Amide Resin

To test the compatibility of 5,10,15,20-tetrakis(pentafluorophenyl)porphyrin (**121**) with on-resin peptide conjugation, the reaction described in section 6.10.7 was performed following the same procedure but without the addition of the base. Under these conditions, the peptide and the porphyrin did not react. After resin washing, the resin had the original aspect. After cleavage, the peptide was found unaffected by analysing the crude pellet by HPLC and LC-MS. The porphyrin was also found unaffected in the discarded reaction mixture analysed by MS analysis.

6.10.10 Porphyrin stability during TFA cleavage

To test the compatibility of 5,10,15,20-tetrakis(pentafluorophenyl)porphyrin (**121**) with the cleavage conditions, **121** was mixed with the standard cleavage cocktail (TFA:TIPS:H_2O, 8:1:1) and shaken at room temperature for 4 h. After this, the mixture was dried and analysed by MS and the porphyrin was unaffected.

6.10.11 Peptide conjugation with porphyrin 121 in solution

The crude peptide (0.01 mmol), porphyrin **121** (20 mg, 0.02 mmol), Cs_2CO_3 (26 mg, 0.08 mmol), and TCEP (12 mg, 0.04 mmol) were combined in a 5 mL glass vial with lid and rubber septum. The reaction atmosphere was conditioned with N_2 flow for a few minutes before adding dry DMSO (4 mL). After stirring the mixture at room temperature for 2.5 h, the reaction was quenched by addition of TFA (2 mL).

Products were precipitated from the solution by addition of Et_2O followed by centrifugation. The resulting pellet was successively suspended in Et_2O and centrifuged twice further for complete porphyrin and DMSO removal.

The solids obtained after drying in air were dissolved in H_2O and/or CH_3CN with 0.1% TFA v/v (or a mixture of these solvents, depending on the solubility) and analysed by HPLC and LC-MS as described below.

This procedure was applied for peptides **126, 127** and **142** to obtain conjugates **128, 129** and **143.**

6.10.12 Peptide conjugation with metalloporphyrin in solution

The crude peptide **142** (0.01 mmol), metalloporphyrin (0.02 mmol), Cs_2CO_3 (26 mg, 0.08 mmol), and TCEP (12 mg, 0.04 mmol) were combined in a 5 mL glass vial with lid and rubber septum. The reaction atmosphere was conditioned with N_2 flow for a few minutes before adding dry DMSO (4 mL). After stirring the mixture at room temperature for 2.5 h, the reaction was quenched by addition of TFA (2 mL).

Products were precipitated from the solution by addition of Et_2O followed by centrifugation. The resulting pellet was successively suspended in Et_2O and centrifuged twice further for complete porphyrin and DMSO removal.

The solids obtained after drying in air were dissolved in H_2O and/or CH_3CN with 0.1% TFA v/v (or a mixture of these solvents, depending on the solubility) and analysed by HPLC and LC-MS as described below.

This procedure was applied for porphyrin **144** and **145** to obtain conjugates **146** and **147.**

6.10.13 Peptide conjugation with porphyrin **148** in solution

The crude peptide **149** (0.04 mmol), porphyrin **148** (6 mg, 0.01 mmol), Cs_2CO_3 (26 mg, 0.08 mmol), and TCEP (11.5 mg, 0.04 mmol) – were combined in a 5 mL glass vial with lid and rubber septum. The reaction atmosphere was conditioned with N_2 flow for a few minutes before adding dry DMSO (4 mL). After stirring the mixture at room temperature for 5 h, the reaction was quenched by addition of TFA (2 mL).

Products were precipitated from the solution by addition of Et_2O followed by centrifugation. The resulting pellet was successively suspended in Et_2O and centrifuged twice further for complete porphyrin and DMSO removal.

Conjugate **150** was obtained as solid obtained after pellet drying in air, dissolved in H_2O and MeCN with 0.1% TFA v/v and analysed by HPLC and LC-MS as described below.

6.10.14 Peptide monocyclisation with porphyrin **121**

The crude peptide (0.01 mmol), porphyrin **121** (9.75 mg, 0.01 mmol), Cs_2CO_3 (26 mg, 0.08 mmol), and TCEP (11.5 mg, 0.04 mmol) – were combined in a 5 mL glass vial with lid and rubber septum. The reaction atmosphere was conditioned with N_2 flow for a few minutes before adding dry DMSO (3 mL). After stirring the mixture at room temperature for 2 h, the reaction was quenched by addition of TFA (1.5 mL).

Products were precipitated from the solution by addition of Et_2O followed by centrifugation. The resulting pellet was successively suspended in Et_2O and centrifuged twice further for complete porphyrin and DMSO removal.

The solids obtained after drying in air were dissolved in H_2O and/or MeCN with 0.1% TFA v/v (or a mixture of these solvents, depending on the solubility) and analysed by HPLC and LC-MS as described below.

This procedure was applied for peptides **77, 163-176** to obtain monocyclic conjugates **160, 177-189.**

6.10.15 Peptide monocyclisation with porphyrin 144

The crude peptide (0.01 mmol), porphyrin **144** (0.01 mmol), Cs_2CO_3 (26 mg, 0.08 mmol), and TCEP (11.5 mg, 0.04 mmol) – were combined in a 5 mL glass vial with lid and rubber septum. The reaction atmosphere was conditioned with N_2 flow for a few minutes before adding dry DMSO (3 mL). After stirring the mixture at room temperature for 2 h, the reaction was quenched by addition of TFA (1.5 mL).

Products were precipitated from the solution by addition of Et_2O followed by centrifugation. The resulting pellet was successively suspended in Et_2O and centrifuged twice further for complete porphyrin and DMSO removal.

The solids obtained after drying in air were dissolved in H_2O and/or MeCN with 0.1% TFA v/v (or a mixture of these solvents, depending on the solubility) and analysed by HPLC and LC-MS as described below.

This procedure was applied for peptides **77** and **167** to obtain monocyclic conjugates **190** and **191** respectively.

6.10.16 Peptide monocyclisation with porphyrin 148

The crude peptide (0.002 mmol), porphyrin **148** (1.3 mg, 0.002 mmol), Cs_2CO_3 (5.2 mg, 0.016 mmol), and TCEP (2.3 mg, 0.008 mmol) – were combined in a 5 mL glass vial with lid and rubber septum. The reaction atmosphere was conditioned with N_2 flow for a few minutes before adding dry DMSO (3 mL). After stirring the mixture at room temperature for 5 h, the reaction was quenched by addition of TFA (1.5 mL).

Products were precipitated from the solution by addition of Et_2O followed by centrifugation. The resulting pellet was successively suspended in Et_2O and centrifuged twice further for complete porphyrin and DMSO removal.

The solids obtained after drying in air were dissolved in H_2O and/or MeCN with 0.1% TFA v/v (or a mixture of these solvents, depending on the solubility) and analysed by HPLC and LC-MS as described below.

This procedure was applied for peptides **77, 166-168** to obtain conjugates **192-195.**

6.10.17 Synthesis of conjugate 197

The crude peptide-porphyrin conjugate **160** (0.01 mmol), 4-acetamidethiophenol (**196**, 6.7 mg, 0.04 mmol), Cs_2CO_3 (5.2 mg, 0.016 mmol), and TCEP (2.3 mg, 0.008 mmol) – were combined in a 5 mL glass vial with lid and rubber septum. The reaction atmosphere was conditioned with N_2 flow for a few minutes before adding dry DMSO (3 mL). After stirring the mixture at room temperature for 5 h, the reaction was quenched by addition of TFA (1.5 mL).

Products were precipitated from the solution by addition of Et_2O followed by centrifugation. The resulting pellet was successively suspended in Et_2O and centrifuged twice further for complete DMSO removal.

Conjugate **197** was obtained as solid obtained after pellet drying in air, dissolved in H_2O and MeCN with 0.1% TFA v/v and analysed by HPLC and LC-MS as described below.

6.10.18 Two-step peptide multicyclisation with porphyrin **121**

The crude peptide-porphyrin conjugate **181** (0.001 mmol), crude peptide (0.002 mmol), Cs_2CO_3 (5.2 mg, 0.016 mmol), and TCEP (2.3 mg, 0.008 mmol) – were combined in a 5 mL glass vial with lid and rubber septum. The reaction atmosphere was conditioned with N_2 flow for a few minutes before adding dry DMSO (3 mL). After stirring the mixture at room temperature for 5 h, the reaction was quenched by addition of TFA (1.5 mL).

Products were precipitated from the solution by addition of Et_2O followed by centrifugation. The resulting pellet was successively suspended in Et_2O and centrifuged twice further for complete DMSO removal.

The solids obtained after drying in air were dissolved in H_2O and/or MeCN with 0.1% TFA v/v (or a mixture of these solvents, depending on the solubility) and analysed by HPLC and LC-MS as described below.

This procedure was applied for peptides **167**, **77**, or **168** to obtain conjugates **198-200.**

6.10.19 One-step peptide multicyclisation with porphyrin **121**

The crude peptide (0.002 mmol), porphyrin **121** (9.75 mg, 0.002 mmol), Cs_2CO_3 (5.2 mg, 0.016 mmol), and TCEP (2.3 mg, 0.008 mmol) – were combined in a 5 mL glass vial with lid and rubber septum. The reaction atmosphere was conditioned with N_2 flow for a few minutes before adding dry DMSO (3 mL). After stirring the mixture at room temperature for 5 h, the reaction was quenched by addition of TFA (1.5 mL).

Products were precipitated from the solution by addition of Et_2O followed by centrifugation. The resulting pellet was successively suspended in Et_2O and centrifuged twice further for complete porphyrin and DMSO removal.

The solids obtained after drying in air were dissolved in H_2O and/or MeCN with 0.1% TFA v/v (or a mixture of these solvents, depending on the solubility) and analysed by HPLC and LC-MS as described below.

This procedure was applied for peptides **201**, **202** and **203** to obtain conjugates **204-206**.

6.10.20 Characterisation of peptide conjugates

*Conjugate **54***
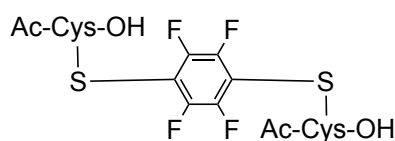
Formula: $C_{16}H_{16}F_4N_2O_6S_2$
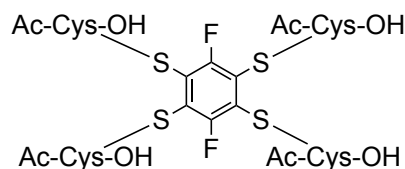
Calculated Exact Mass: 472.04

Calculated Molecular Weight: 472.43

ESI-MS (m/z): 473.04691 [M+H]+

RP-HPLC RT (System A, Column A, Method A): 10.9 min

^{19}F NMR (DMSO) δ, ppm: -132.244 (s, 4F)

Conjugate 55

Formula: $C_{26}H_{32}F_2N_4O_{12}S_4$
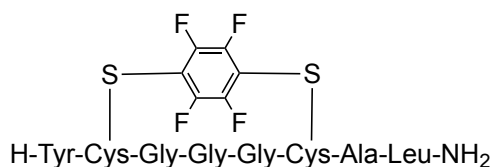
Calculated Exact Mass: 758.09

Calculated Molecular Weight: 758.79

ESI-MS (m/z): 759.09531 [M+H]+

RP-HPLC RT (System A, Column A, Method A): 8.8 min

^{19}F NMR (DMSO) δ, ppm: -95.4383 (s, 2F)

Conjugate 70

Formula: $C_{36}H_{45}F_4N_9O_9S_2$

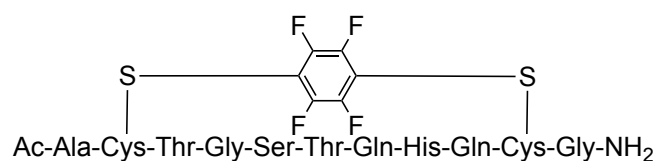Calculated Exact Mass:887.271

Calculated Molecular Weight: 887.92

ESI-MS (m/z): 1755.55000 [2M+H]+, 888.27915 [M+H]+

RP-HPLC RT (System A, Column A, Method A): 9.6 min

Purity by analytical HPLC (crude): 80%

^{19}F NMR (H$_2$O) δ, ppm: -137.247 (s, 4F)

Conjugate 80

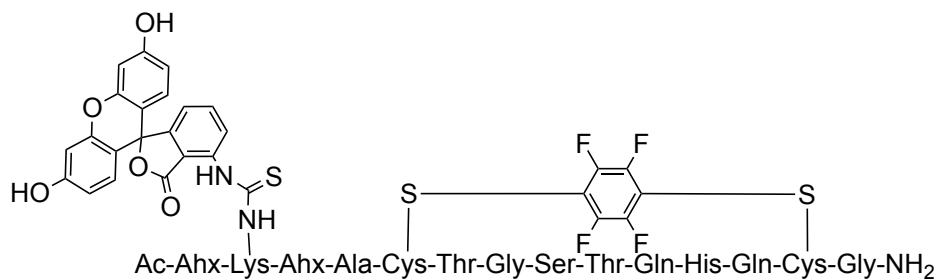Formula: $C_{48}H_{66}F_4N_{16}O_{17}S_2$

Calculated Exact Mass: 1278.42

Calculated Molecular Weight: 1279.26

ESI-MS (m/z): 1279.14 [M+H]$^+$, 640.32 [M+2H]$^{2+}$

RP-HPLC RT (System B, Column C, Method D): 7.5 min

Conjugate **83**

Ac-Ahx-Lys-Ahx-Ala-Cys-Thr-Gly-Ser-Thr-Gln-His-Gln-Cys-Gly-NH$_2$

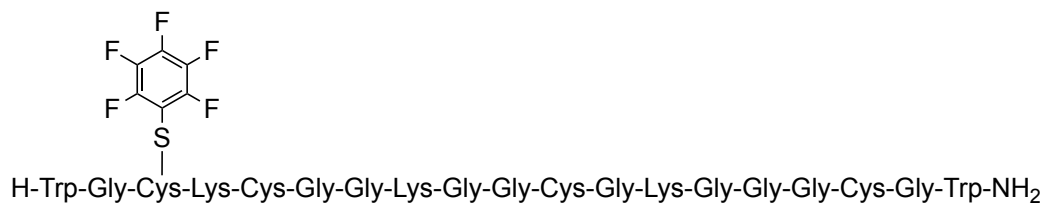Formula: $C_{87}H_{111}F_4N_{21}O_{25}S_3$

Calculated Exact Mass: 2021.72

Calculated Molecular Weight: 2023.14

ESI-MS (m/z): 1012.3915 [M+2H]$^{2+}$, 675.2648 [M+3H]$^{3+}$

RP-HPLC RT (System B, Column C, Method D): 9.5 min

Conjugate **112**

H-Trp-Gly-Cys-Lys-Cys-Gly-Gly-Lys-Gly-Gly-Cys-Gly-Lys-Gly-Gly-Gly-Cys-Gly-Trp-NH$_2$
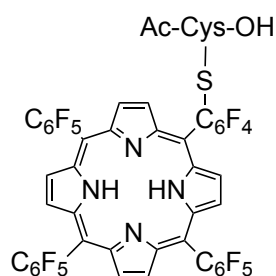
Formula: $C_{78}H_{108}F_5N_{25}O_{19}S_4$

Calculated Exact Mass: 1921.71

Calculated Molecular Weight: 1923.11

ESI-MS (m/z): 961.1253 [M+2H]$^{2+}$, 641.0817 [M+3H]$^{3+}$, 481.0593 [M+4H]$^{4+}$

RP-HPLC RT (System C, Method C, Column A): 7.7 min

*Conjugate **122***

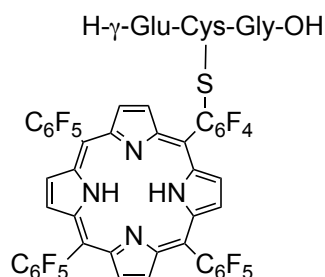

Formula: C$_{49}$H$_{18}$F$_{19}$N$_5$O$_3$S

Calculated Exact Mass: 1117.08

Calculated Molecular Weight: 1117.74

ESI-MS (*m/z*): 1118.0883 [M+H]+

RP-HPLC RT (System A, Column A, Method A): 13.8 min

^{19}F NMR (DMSO) δ, ppm: -133.662 (m, 2F), -138.814 (m, 6F), -139.042 (m, 2F), -153.209 (m, 3F), -162.264 (m, 6F)

*Conjugate **123***

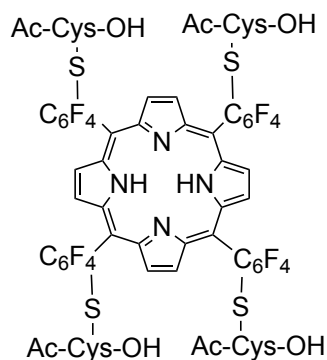Formula: C$_{54}$H$_{28}$F$_{19}$N$_7$O$_6$S

Calculated Exact Mass: 1261.14

Calculated Molecular Weight: 1261.87

ESI-MS (*m/z*): 2524.2781 [2M+H]$^+$, 1262.1422 [M+H]$^+$

RP-HPLC RT (System A, Column A, Method A): 11.7 min

^{19}F NMR (DMF) δ, ppm: -134.481 (m, 2F), -139.661 (m, 8F), -155.033 (m, 3F), -163.859 (m, 6F)

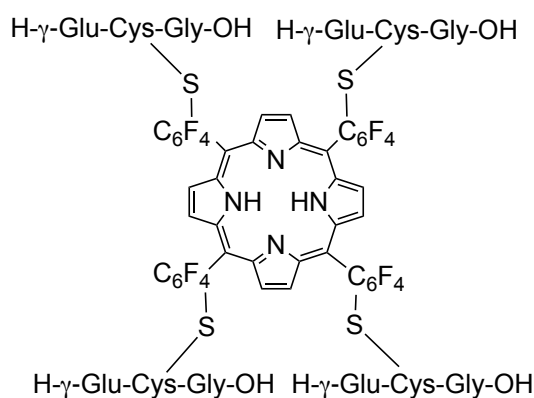*Conjugate 124*

Formula: $C_{64}H_{42}F_{16}N_8O_{12}S_4$

Calculated Exact Mass:1546.15

Calculated Molecular Weight: 1547.30

ESI-MS (*m/z*): 1547.2081 [M+H]+

RP-HPLC RT (System A, Column A, Method A): 10.0 min

^{19}F NMR (DMSO) δ, ppm: -134.278 (m, 8F), -139.687 (m, 8F)

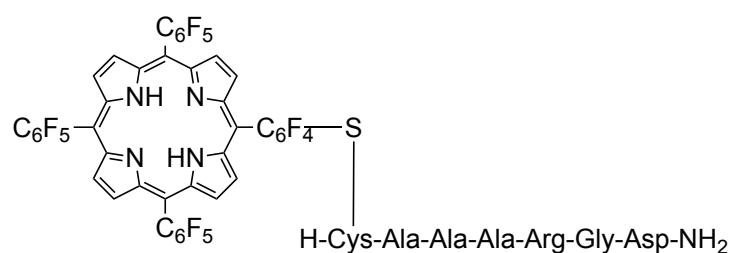*Conjugate 125*

Formula: $C_{84}H_{74}F_{16}N_{16}O_{24}S_4$

Calculated Exact Mass: 2122.37

Calculated Molecular Weight: 2123.82

ESI-MS (*m/z*): 2124.3772 [M+H]$^+$, 1062.6939 [M+2H]$^{2+}$, 708.7985 [M+3H]$^{3+}$, 531.8511 [M+4H]$^{4+}$

RP-HPLC RT (System A, Column A, Method A): 5.8 min

^{19}F NMR (DMF) δ, ppm: -134.021 (m, 8F), -139.417 (m, 8F)

Conjugate 128

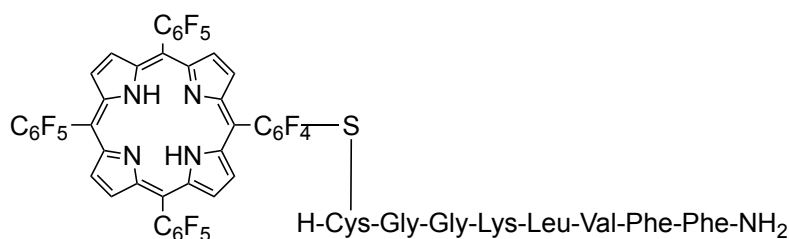Formula: $C_{68}H_{52}F_{19}N_{15}O_9S$

Calculated Exact Mass: 1615.35

Calculated Molecular Weight: 1616.29

ESI-MS (*m/z*): 1616.5819 [M+H]$^+$, 808.7606 [M+2H]$^{2+}$, 539.5410 [M+3H]$^{3+}$

RP-HPLC RT (System D, Column C, Method D): 20.4 min

^{19}F NMR (DMSO) δ, ppm: -134.52 (m, 2F), -139.53 (m, 8F), -154.49 (m, 3F), -163.56 (m, 6F)
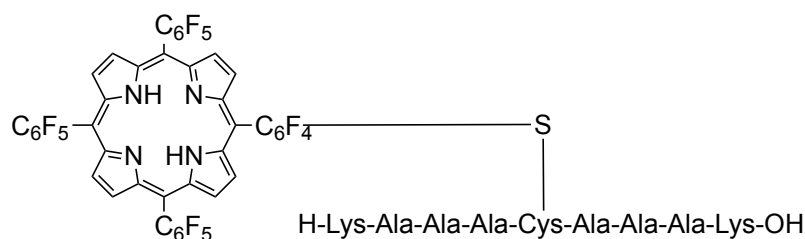
Conjugate 129

Formula: $C_{86}H_{73}F_{19}N_{14}O_8S$

Calculated Exact Mass: 1822.52

Calculated Molecular Weight: 1823.65

ESI-MS (*m/z*): 912.28725 [M+2H]$^{2+}$

RP-HPLC RT (System A, Column A, Method A): 9.4 min

Conjugate 143

Formula: $C_{77}H_{70}F_{19}N_{15}O_{10}S$

Calculated Exact Mass: 1757.48

Calculated Molecular Weight: 1758.53

ESI-MS (*m/z*): 880.2432 [M+2H]$^{2+}$, 587.1621 [M+3H]$^{3+}$

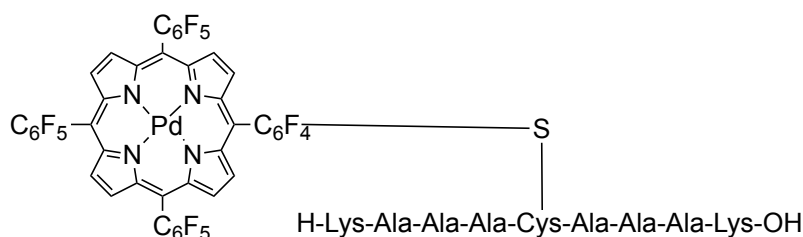RP-HPLC RT (System E, Column C, Method D): 18.5 min

Purity by analytical HPLC (after purification): 93%

Yield (after purification): 35%

^{19}F NMR (CD$_3$OD) δ, ppm: -135.00 (m, 2F), -139.5 (m, 2F), -141.14 (m, 6F), -155.18 (m, 3F), -164.83 (m, 6F)

UV-Vis (DMSO, nm) λ$_{max}$ (Soret): 412; Q bands: 508, 589, 652

Conjugate *146*

Formula: C$_{77}$H$_{68}$F$_{19}$N$_{15}$O$_{10}$PdS

Calculated Exact Mass: 1861.37

Calculated Molecular Weight: 1862.94

ESI-MS (*m/z*): 933.2424 [M+2H]$^{2+}$, 622.1742 [M+3H]$^{3+}$

RP-HPLC RT (System E, Column C, Method D): 17.5 min

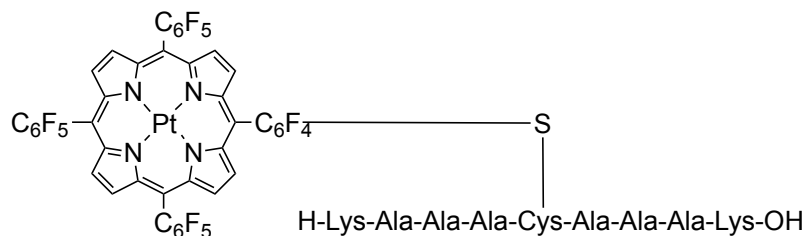Purity by analytical HPLC (after purification): 85%

Yield (after purification): 33%

^{19}F NMR (CD$_3$OD) δ, ppm: -134.78 (m, 2F), -139.52 (m, 2F), -141.16 (m, 6F), -155.03 (m, 3F), -164.62 (m, 6F)

UV-Vis (DMSO, nm) λ$_{max}$ (Soret): 395; Q bands: 514, 545

Conjugate *147*

Formula: C$_{77}$H$_{68}$F$_{19}$N$_{15}$O$_{10}$PtS

Calculated Exact Mass: 1950.43

Calculated Molecular Weight: 1951.60

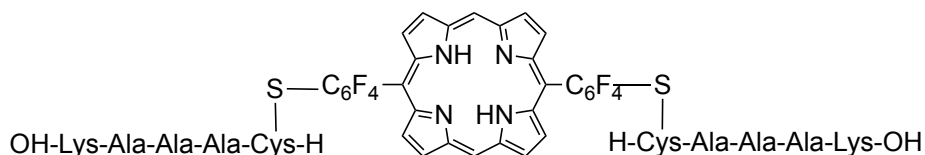ESI-MS (m/z): 977.3113 [M+2H]$^{2+}$, 651.8523 [M+3H]$^{3+}$

RP-HPLC RT (System E, Column C, Method D): 18.5

Purity by analytical HPLC (after purification): 91%

Yield (after purification): 34%

UV-Vis (DMSO, nm) λ_{max} (Soret): 390; Q bands: 509, 541

Conjugate 150

Formula: $C_{68}H_{78}F_8N_{16}O_{12}S_2$

Calculated Exact Mass: 1526.53

Calculated Molecular Weight: 1527.58

ESI-MS (m/z): 765.1876 [M+2H]$^{2+}$, 510.1238 [M+3H]$^{3+}$

RP-HPLC RT (System E, Column C, Method D): 10.8 min

Purity by analytical HPLC (after purification): 90%

Yield (after purification): 25%

^{19}F NMR (CD$_3$OD) δ, ppm: -132.04 (m, 4F), -139.09 (m, 4F)

UV-Vis (DMSO, nm) λ_{max} (Soret): 402; Q bands: 504, 533, 580, 630

Conjugate 160

Formula: $C_{84}H_{74}F_{18}N_{20}O_{16}S_2$

Calculated Exact Mass: 2024.47

Calculated Molecular Weight: 2025.73

ESI-MS (m/z): 1013.8258 [M+2H]$^{2+}$, 689.8533 [M+H+2Na]$^{3+}$, 676.1896 [M+3H]$^{3+}$

RP-HPLC RT (System D, Column C, Method D): 15.9 min

Yield (after purification): 18%

^{19}F NMR (CD$_3$OD) δ, ppm: -135.1 (m, 4F), -139.9 (m, 8F), -155.2 (m, 2F), -164.9 (m, 4F)

UV-Vis (DMSO, nm) λ_{max} (Soret): 411; Q bands: 507, 586, 650

Conjugate 177

C$_6$F$_5$ C$_6$F$_5$

C$_6$F$_4$ C$_6$F$_4$

H-Ala-Cys-Ala-Thr-Gly-Ser-Thr-Gln-His-Gln-Ala-Cys-Gly-NH$_2$

Formula: $C_{90}H_{84}F_{18}N_{22}O_{18}S_2$

Calculated Exact Mass: 2166.55

Calculated Molecular Weight: 2167.89

ESI-MS (m/z): 1085.0803 [M+2H]$^{2+}$, 723.3773 [M+3H]$^{3+}$
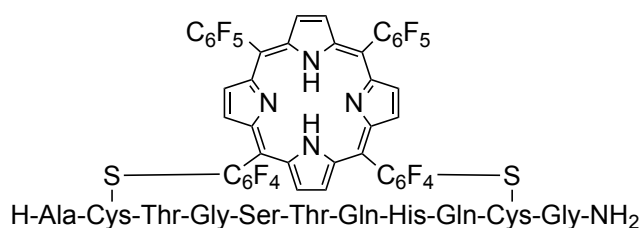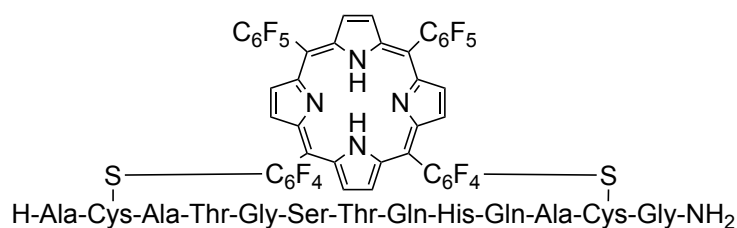
RP-HPLC RT (System D, Column C, Method D): 15.9 min

Conjugate 178

C$_6$F$_5$ C$_6$F$_5$

C$_6$F$_4$ C$_6$F$_4$

H-Ala-Cys-Ala-Thr-Gly-Ser-Thr-Gln-His-Gln-Cys-Gly-NH$_2$
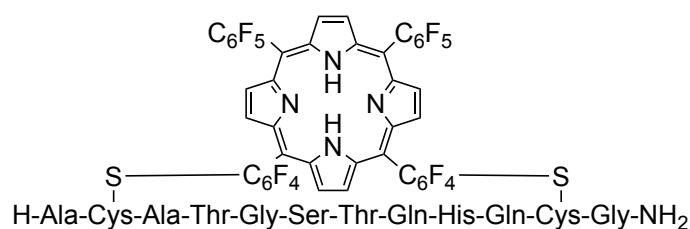
Formula: $C_{87}H_{79}F_{18}N_{21}O_{17}S_2$

Calculated Exact Mass: 2095.51

Calculated Molecular Weight: 2096.81

ESI-MS (m/z): 1049.3163 [M+2H]$^{2+}$, 714.2386 [M+H+2Na]$^{3+}$, 699.8887 [M+3H]$^{3+}$

RP-HPLC RT (System D, Column C, Method D): 15.5 min

Conjugate 179

C$_6$F$_5$ C$_6$F$_5$

C$_6$F$_4$ C$_6$F$_4$

H-Ala-Cys-Thr-Ser-Thr-Gln-His-Gln-Cys-Gly-NH$_2$
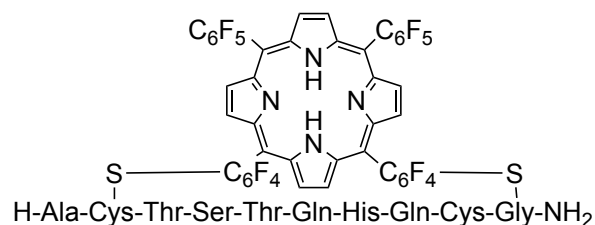
Formula: $C_{82}H_{71}F_{18}N_{19}O_{15}S_2$

Calculated Exact Mass: 1967.45

Calculated Molecular Weight: 1968.68

ESI-MS (m/z): 985.28886 [M+2H]$^{2+}$, 670.8729 [M+H+2Na]$^{3+}$, 657.1962 [M+3H]$^{3+}$

RP-HPLC RT (System D, Column C, Method D): 15.8 min

*Conjugate **180***

C_6F_5 C_6F_5

C_6F_4 C_6F_4

S S

H-Ala-Cys-Thr-Ser-Thr-Gln-His-Gln-Cys-Gly-NH$_2$

Formula: $C_{78}H_{64}F_{18}N_{18}O_{13}S_2$

Calculated Exact Mass: 1866.41

Calculated Molecular Weight: 1867.57

ESI-MS (*m/z*): 934.7472 [M+2H]$^{2+}$, 636.8464 [M+H+2Na]$^{3+}$, 623.4749 [M+3H]$^{3+}$

RP-HPLC RT (System E, Column C, Method D): 18.1 min

^{19}F NMR (CD$_3$OD) δ, ppm: -135.2 (m, 4F), -140.0 (m, 8F), -155.2 (m, 2F), -164.9 (m, 4F)

Yield (after purification): 24%

*Conjugate **181***

C_6F_5 C_6F_5

C_6F_4 C_6F_4

S S

H-Ala-Cys-Thr-Gln-His-Gln-Cys-Gly-NH$_2$
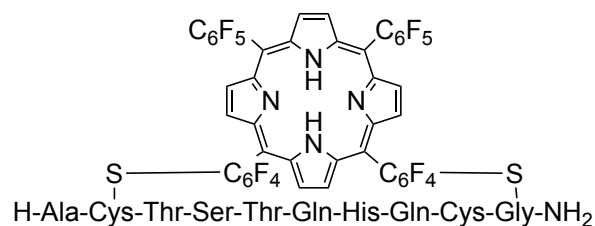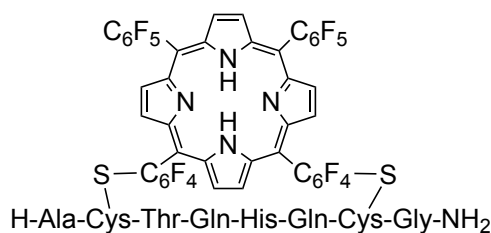
Formula: $C_{75}H_{59}F_{18}N_{17}O_{11}S_2$

Calculated Exact Mass: 1799.37

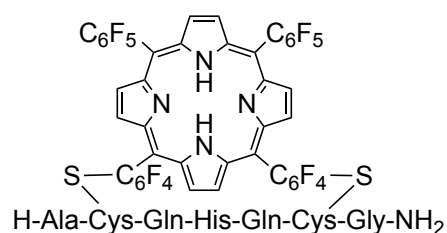Calculated Molecular Weight: 1780.50

ESI-MS (*m/z*): 891.0790 [M+2H]$^{2+}$, 608.4000 [M+H+2Na]$^{3+}$, 594.3957 [M+3H]$^{3+}$

RP-HPLC RT (System E, Column C, Method D): 18.3 min

^{19}F NMR (CD$_3$OD) δ, ppm: -135.2 (m, 4F), -140.1 (m, 8F), -155.2 (m, 2F), -164.9 (m, 4F)

Yield (after purification): 25%

Conjugate *182*

H-Ala-Cys-Gln-His-Gln-Cys-Gly-NH$_2$

Formula: $C_{71}H_{52}F_{18}N_{16}O_9S_2$

Calculated Exact Mass: 1678.33

Calculated Molecular Weight: 1679.39

ESI-MS (m/z): 1679.8125 [M+H]$^+$, 840.3854 [M+2H]$^{2+}$, 574.2904 [M+H+2Na]$^{3+}$

RP-HPLC RT (System D, Column C, Method D): 16.4 min

^{19}F NMR (CD$_3$OD) δ, ppm: -135.1 (m, 4F), -140.2 (m, 8F), -155.1 (m, 2F), -164.9 (m, 4F)

Yield (after purification): 39%

Conjugate *183*

H-Ala-Cys-His-Gln-Cys-Gly-NH$_2$
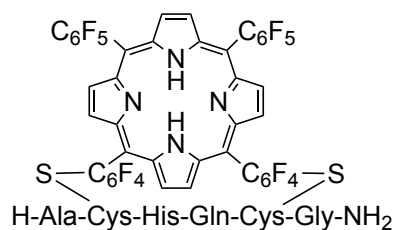
Formula: $C_{66}H_{44}F_{18}N_{14}O_7S_2$

Calculated Exact Mass: 1550.27

Calculated Molecular Weight: 1551.26

ESI-MS (m/z): 1552.0144 [M+H]$^+$, 776.4614 [M+2H]$^{2+}$, 532.0419 [M+H+2Na]$^{3+}$, 518.0013 [M+3H]$^{3+}$

RP-HPLC RT (System X, Column C, Method D): 17.1 min

Conjugate *184*

H-Ala-Cys-His-Cys-Gly-NH$_2$
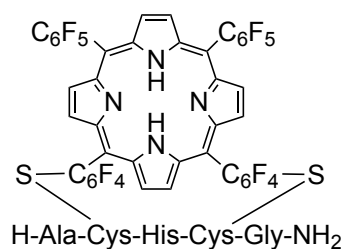
Formula: $C_{58}H_{34}F_{18}N_{10}O_5S_2$

Calculated Exact Mass: 1356.19

Calculated Molecular Weight: 1357.07

ESI-MS (m/z): 1357.5035 [M+H]$^+$, 905.7087 [2M+3H]$^{3+}$, 679.2785 [M+2H]$^{2+}$

RP-HPLC RT (System D, Column C, Method D): 19.9 min

*Conjugate **185***

H-Ala-Cys-His-Cys-Gly-NH$_2$
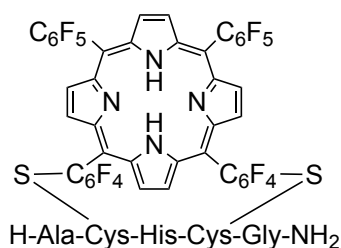
Formula: $C_{61}H_{36}F_{18}N_{12}O_5S_2$

Calculated Exact Mass: 1422.21

Calculated Molecular Weight: 1423.13

ESI-MS (m/z): 1423.8884 [M+H]$^+$, 712.4102 [M+2H]$^{2+}$, 488.9788 [M+H+2Na]$^{3+}$, 475.2771 [M+3H]$^{3+}$

RP-HPLC RT (System D, Column C, Method D): 18.5 min

*Conjugate **186***
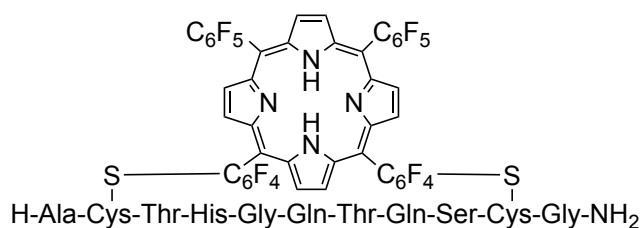
H-Ala-Cys-Thr-His-Gly-Gln-Thr-Gln-Ser-Cys-Gly-NH$_2$

Formula: $C_{84}H_{74}F_{18}N_{20}O_{16}S_2$

Calculated Exact Mass: 2024.47

Calculated Molecular Weight: 2025.73

ESI-MS (m/z): 1014.0622 [M+2H]$^{2+}$, 690.0623 [M+H+2Na]$^{3+}$, 676.3733 [M+3H]$^{3+}$

RP-HPLC RT (System D, Column C, Method D): 15.6 min

Yield (after purification): 10%

*Conjugate **187***

H-Asn-Cys-Val-Val-Gly-Tyr-Ile-Gly-Glu-Arg-Cys-Gln-NH$_2$
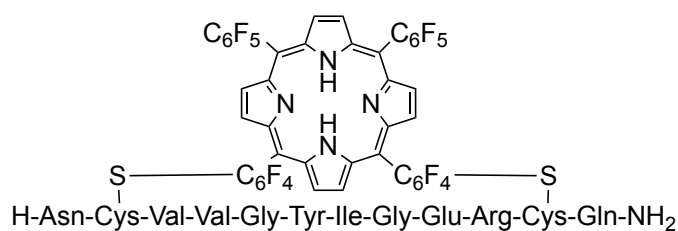
Formula: $C_{99}H_{98}F_{18}N_{22}O_{17}S_2$

Calculated Exact Mass: 2272.66

Calculated Molecular Weight: 2274.10

ESI-MS (m/z): 1516.6774 [2M+3H]$^{3+}$, 1137.8462 [M+2H]$^{2+}$, 758.9305 [M+3H]$^{3+}$

RP-HPLC RT (System D, Column C, Method D): 16.1 min

Conjugate 188

H-Cys-Lys-Ala-Pro-Glu-Thr-Ala-Leu-Cys-NH$_2$
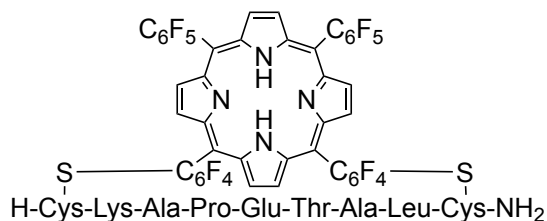
Formula: $C_{82}H_{75}F_{18}N_{15}O_{12}S_2$

Calculated Exact Mass: 1867.49

Calculated Molecular Weight: 1868.69

ESI-MS (m/z): 935.3064 [M+2H]$^{2+}$, 637.5211 [M+H+2Na]$^{3+}$, 623.5334 [M+3H]$^{3+}$

RP-HPLC RT (System D, Column C, Method D): 17.1 min

Conjugate 189

H-Ala-Gly-Tyr-Leu-Leu-Gly-Lys-Ile-Asn-Leu-Lys-Ala-Cys-Ala-Ala-Leu-Ala-Lys-Lys-Cys-Leu-NH$_2$

Formula: $C_{142}H_{181}F_{18}N_{31}O_{23}S_2$

Calculated Exact Mass: 3094.31

Calculated Molecular Weight: 3096.30

ESI-MS (m/z): 1549.1902 [M+2H]$^{2+}$, 1032.8458 [M+3H]$^{3+}$, 744.8635 [M+4H]$^{4+}$, 617.0869 [M+5H]$^{5+}$

RP-HPLC RT (System D, Column C, Method D): 16.9 min

Conjugate 190

H-Ala-Cys-Ser-Thr-Gln-His-Gln-Cys-Gly-NH$_2$
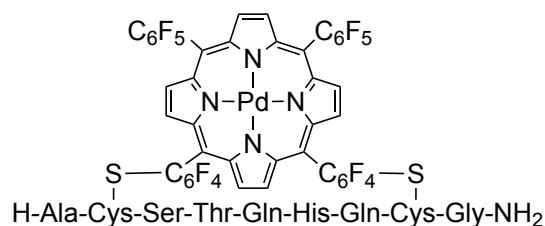
Formula: $C_{84}H_{72}F_{18}N_{20}O_{16}PdS_2$

Calculated Exact Mass: 2128.36

Calculated Molecular Weight: 2130.36

ESI-MS (m/z): 1065.2397 $[M+2H]^{2+}$, 725.1728 $[M+H+2Na]^{3+}$, 710.5242 $[M+3H]^{3+}$

RP-HPLC RT (System E, Column C, Method D): 18.3 min

Conjugate 191
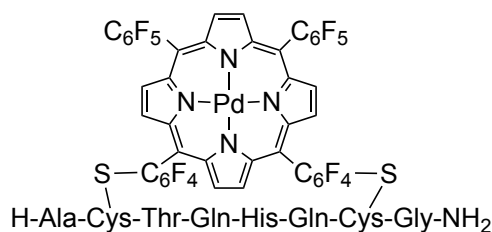
Formula: $C_{75}H_{57}F_{18}N_{17}O_{11}PdS_2$

Calculated Exact Mass: 1883.26

Calculated Molecular Weight: 1884.90

ESI-MS (m/z): 942.6487 $[M+2H]^{2+}$

RP-HPLC RT (System D, Column C, Method D): 18.8 min

Conjugate 192

Formula: $C_{72}H_{76}F_8N_{20}O_{16}S_2$

Calculated Exact Mass: 1692.51

Calculated Molecular Weight: 1693.63

ESI-MS (m/z): 847.5228 $[M+2H]^{2+}$, 579.0417 $[M+H+2Na]^{3+}$, 565.3703 $[M+3H]^{3+}$

RP-HPLC RT (System D, Column C, Method D): 10.3 min

Conjugate 193
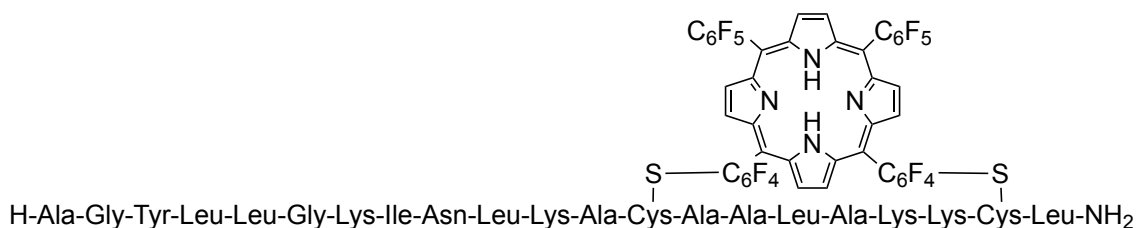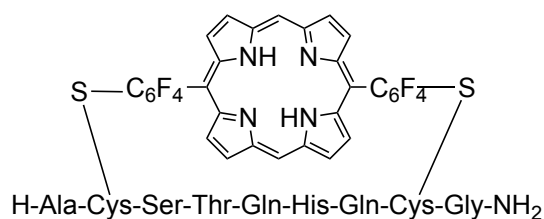
Formula: $C_{66}H_{66}F_8N_{18}O_{13}S_2$

Calculated Exact Mass: 1534.44

Calculated Molecular Weight: 1535.47

ESI-MS (m/z): 768.2599 $[M+2H]^{2+}$, 526.1618 $[M+H+2Na]^{3+}$

RP-HPLC RT (System E, Column C, Method D): 13.07

*Conjugate **194***

S—C_6F_4 ... C_6F_5

H-Ala-Cys-Thr-Gln-His-Gln-Cys-Gly-NH$_2$
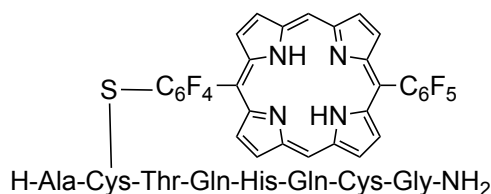
Formula: $C_{63}H_{62}F_9N_{17}O_{11}S_2$

Calculated Exact Mass: 1467.41

Calculated Molecular Weight: 1468.40

ESI-MS (m/z): 1469.0731 $[M+H]^+$, 735.0255 $[M+2H]^{2+}$, 490.6806 $[M+3H]^{3+}$

RP-HPLC RT (System E, Column C, Method D): 15.9/16.4 min

*Conjugate **195***

S—C_6F_4 ... C_6F_5

H-Ala-Cys-Gln-His-Gln-Cys-Gly-NH$_2$
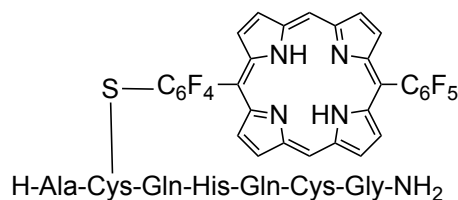
Formula: $C_{59}H_{55}F_9N_{16}O_9S_2$

Calculated Exact Mass: 1366.36

Calculated Molecular Weight: 1367.30

ESI-MS (m/z): 1367.9575 $[M+H]^+$, 684.4943 $[M+2H]^{2+}$, 456.6762 $[M+3H]^{3+}$

RP-HPLC RT (System E, Column C, Method D): 13.9/14.4 min

*Conjugate **197***

CH$_3$-CO-NH-Ph—S–C_6F_4 ... C_6F_4—S—Ph-NH-CO-CH$_3$

S—C_6F_4 ... C_6F_4—S

H-Ala-Cys-Ser-Thr-Gln-His-Gln-Cys-Gly-NH$_2$
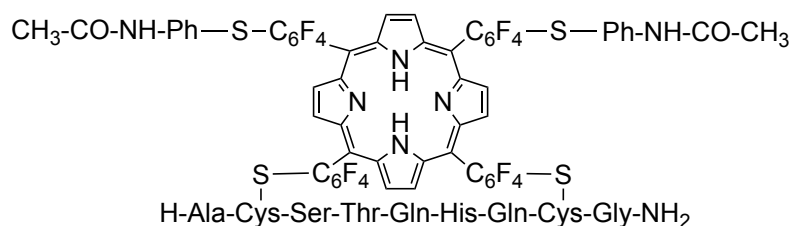
Formula: $C_{100}H_{90}F_{16}N_{22}O_{18}S_4$

Calculated Exact Mass: 2318.54

Calculated Molecular Weight: 2320.17

ESI-MS (m/z): 1547.4374 [2M+3H]$^{2+}$, 1160.8646 [M+2H]$^{2+}$, 773.8849 [M+3H]$^{3+}$

RP-HPLC RT (System E, Column C, Method D): 14.7 min

Conjugate 198

H-Ala-Cys-Thr-Gln-His-Gln-Cys-Gly-NH$_2$
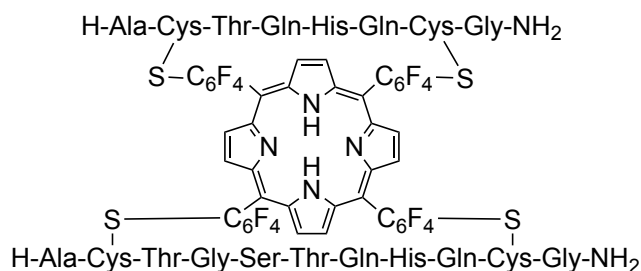
Formula: $C_{106}H_{108}F_{16}N_{30}O_{22}S_4$

Calculated Exact Mass: 2584.69

Calculated Molecular Weight: 2586.43

ESI-MS (m/z): 1293.8453 [M+2H]$^{2+}$, 863.3235 [M+3H]$^{3+}$, 658.0114 [M+3H+Na]$^{4+}$, 647.7198 [M+4H]$^{4+}$

RP-HPLC RT (System E, Column C, Method D): 8.4 min

^{19}F NMR (CD$_3$OD) δ, ppm: -135.0 (m, 8F), -139.7 (m, 8F)

Yield (after purification): 27%

Conjugate 199

H-Ala-Cys-Thr-Gln-His-Gln-Cys-Gly-NH$_2$
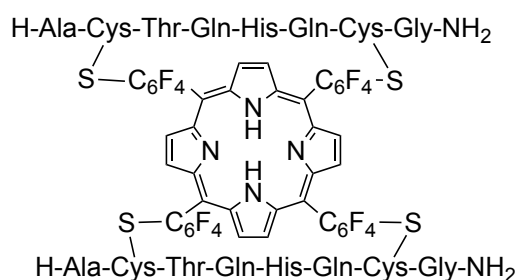
Formula: $C_{115}H_{123}F_{16}N_{33}O_{27}S_4$

Calculated Exact Mass: 2829.79

Calculated Molecular Weight: 2831.67

ESI-MS (m/z): 944.9680 [M+3H]$^{3+}$, 708.9626 [M+4H]$^{4+}$

RP-HPLC RT (System E, Column C, Method D): 10.6 min

*Conjugate **200***
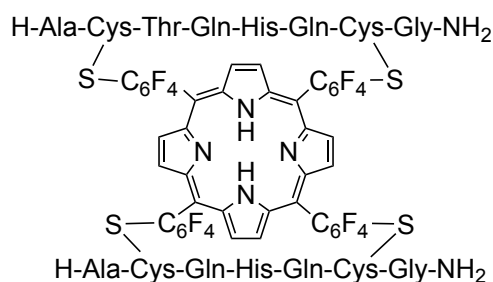
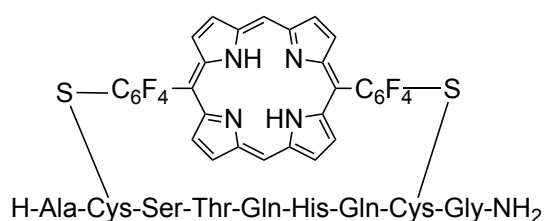Formula: $C_{102}H_{101}F_{16}N_{29}O_{20}S_4$

Calculated Exact Mass: 2483.64

Calculated Molecular Weight: 2485.33

ESI-MS (m/z): 829.1957 [M+3H]$^{3+}$

RP-HPLC RT (System E, Column C, Method D): 10.6 min

*Conjugate **204***
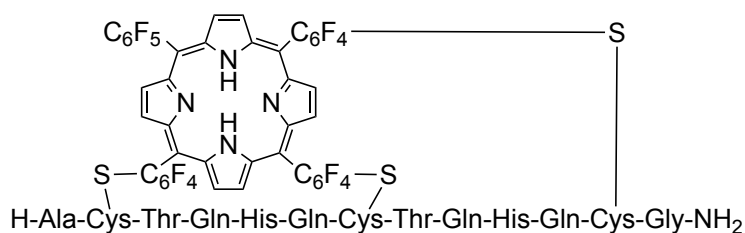
Formula: $C_{98}H_{93}F_{17}N_{26}O_{19}S_3$

Calculated Exact Mass: 2356.60

Calculated Molecular Weight: 2358.14

ESI-MS (m/z): 1179.8256 [M+2H]$^{2+}$, 786.9054 [M+3H]$^{3+}$, 610.9289 [M+2H+2Na]$^{4+}$, 600.4277 [M+3H+Na]$^{4+}$

RP-HPLC RT (System E, Column C, Method D): 14.1 min

^{19}F NMR (CD$_3$OD) δ, ppm:

Yield (after purification): 15%

*Conjugate **205***
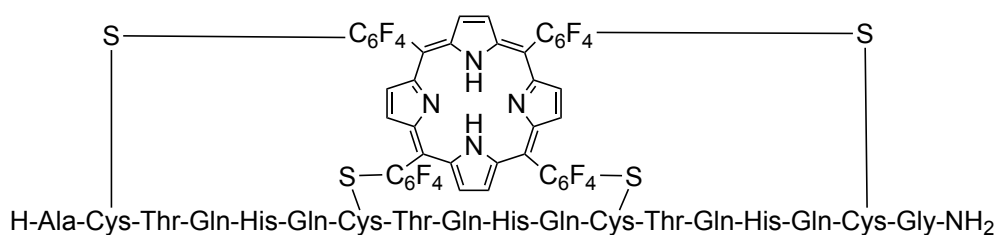
Formula: $C_{121}H_{127}F_{16}N_{35}O_{27}S_4$

Calculated Exact Mass: 2933.83

Calculated Molecular Weight: 2935.78

ESI-MS (*m/z*): 1468.8330 [M+2H]$^{2+}$, 979.2777 [M+3H]$^{3+}$, 745.2099 [M+3H+Na]$^{4+}$, 734.9792 [M+4H]$^{4+}$

RP-HPLC RT (System E, Column C, Method D): 11.2 min

^{19}F NMR (CD$_3$OD) δ, ppm:

Yield (after purification): 28%

*Conjugate **206***

H-Ala-Cys-Thr-Gln-His-Gln-Cys-Thr-Gln-His-Gln-Cys-Thr-Gln-His-Gln-Cys-Gly-NH$_2$
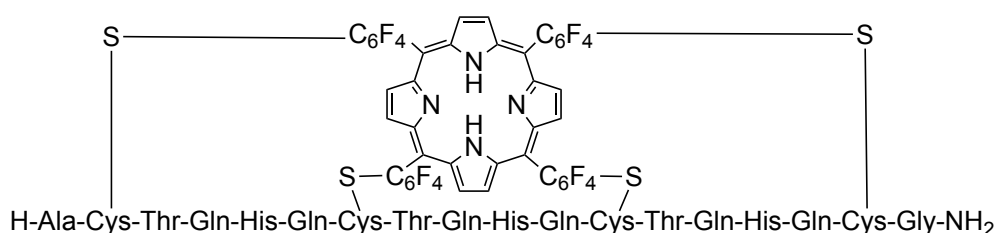
Formula: C$_{130}$H$_{142}$F$_{16}$N$_{38}$O$_{33}$S$_4$

Calculated Exact Mass: 3194.92

Calculated Molecular Weight: 3197.01

ESI-MS (*m/z*): 1067.2273 [M+3H]$^{3+}$, 800.6884 [M+4H]$^{4+}$, 640.5292 [M+5H]$^{5+}$

RP-HPLC RT (System E, Column C, Method D): 11.0 min

6.11 SELF-ASSEMBLING STUDIES

Peptide-porphyrin conjugates **143** and **150** (as solid and after purification) were dissolved in DMSO to obtain a stock solution 10 mM, subsequently stored in freezer.

Depending on the required concentration of the final sample, a DMSO solution of conjugate (the stock solution or diluted solutions obtained from the stock) was diluted with water or different water-based buffers to have a final concentration of DMSO always equal to 1% v/v.

Phosphate buffer 10 mM was prepared by dissolving 538 mg of Na$_2$HPO$_4$.12H$_2$O (7.5 mM) and 60 mg of NaH$_2$PO4 (2.5 mM) in 200 mL of ultrapure H$_2$O and the pH adjusted to 7.4. Aqueous solution of NaOH 10 mM was prepared by dissolving 80 mg of solid NaOH in 200 mL of ultrapure H$_2$O to obtain a pH of 12. Aqueous solution of NaOH pH=9.7 was prepared by diluting around 1-2 mL of 10 mM NaOH pH=12 in 100 mL of ultrapure H$_2$O. For all solutions, ionic strength was adjusted by addition of NaCl (0.59 g/100 mL → 100 mM or 1.17 g/100 mL → 200 mM).

After incubation for a specific time and at defined temperature, 100 μL of self-assembly solution were sampled and transferred in a 96-well plate (BD Falcon 96 clear microplate with UV/Vis transparent film) for analysis. Finally, the plate was read on a CLARIOstar plate reader (BMG Labtech, Aylesbury, UK), to record the absorbance (scan between 350 nm and 600 nm, resolution 1 nm, settling time 0.5 s,

22 flashes/well) and fluorescence (excitation wavelength of 415 nm or 405 nm, emission wavelength of 600/740 nm, top optic, resolution 1 nm, settling time 0.5 s, 40 flashes/well) data.

Table 32. Additional experiment required for fitting a non-linear model on Minitab.

Experiment*	Conc. (μM)	pH	I.S. (mM)	Temp (*C)
20	10	9.7	105	21
21	100	9.7	105	21
22	50	9.7	105	21
23	50	9.7	105	21
24	50	9.7	105	21
25	50	9.7	105	21
26	50	9.7	105	21
27	50	9.7	105	5
28	50	9.7	105	37
29	50	12	105	21
30	50	9.7	200	21
31	50	9.7	10	21
32	50	7.4	105	21

6.12 COMPUTATIONAL STUDIES

Computational studies were performed by Michele Assante during his PhD at LJMU.

For each of the 15 peptides (**77**, **163-175**), up to ten conformers were generated from a SMILES string with ETDKGv3[258] and implemented in the RDKit python package.[259] ETDKGv3 was not able to generate any conformer for peptide **176**.

Each generated conformer was optimized with XTB[249,250] with the keyword 'ohess' using the default option for accuracy. Generalised Born model with Solvent Accessible Surface Area for non-polar interaction (GBSA)[260] for DMSO was used as implicit solvent model. The average of the three-dimensional Euclidean distance was calculated, for each set of conformers, considering the distance between the two sulphur atoms (S) of the peptide on the XTB-optimised geometry.

The porphyrin **121** structure was generated in a similar fashion with ETDKGv3 and optimised to a minimum with XTB using the same options for the peptides.

Two distances between the 4-position carbon of pentafluorophenyl groups were considered: one between adjacent groups (15-20, 20-5, 5-10, 15-10) and one between opposite groups (10-20, 15-5). The average distances were 10.9 Å for the adjacent groups and 15.5Å for the opposite groups.

6.13 CELLULAR STUDIES

Cell culture, cell viability, and cellular uptake studies were performed by Talhat Chaudhry and Giulia Scagnetti during their PhD at LJMU.

The SCCIC8 and Calu-3 cells were maintained in DMEM supplemented with 10% NBCS and MEM supplemented with 10% FBS and 1% Penicillin/Streptomycin, respectively. Cells were passaged once they reached 70-80% confluency.

For cell viability experiments, cells were seeded at 1×10^4 per well of a 96-well plate and maintained in complete media for 24 h. Subsequently, the cells were treated with **160** and **186** (10 μM), and the respective controls. The cells were incubated for a further 24 h. After this period the treated culture medium was removed, and cells were washed with 200 μL PBS. The cells were treated with resazurin (0.1 mg/mL, 100 μL) and incubated in the dark at 37 °C for 4 h.

Finally, the plate was read on a CLARIOstar plate reader (BMG Labtech, Aylesbury, UK), fluorescence was measured at an excitation wavelength of 545 nm and an emission wavelength of 600 nm. Each experiment was performed in triplicate.

The experiments were tested for normality and statistical analysis was carried out using one-way ANOVA with Dunnett's multiple comparisons test; levels of significance: $*p < 0.05$, compared to untreated control. All statistical tests were performed with GraphPad prism 9.

For cellular uptake experiments, SCCIC8 and Calu-3 cells were seeded at 10×10^4 on glass cover slips in a 6-well plate for 24 h. Subsequently, cells were treated with **160** and **186** (10 μM) and incubated for a further 24 h. Prior to visualisation, the culture media was removed, the cells were washed with PBS and fixed using ice-cold MeOH for 7 min at -20°C. After fixation, the cells were successively washed with PBS. Cells were mounted with 4′,6-diamidino-2-phenylindole (DAPI) containing fluoromount G mounting medium.

Images were collected using a Zeiss LSM 710 Confocal microscope using oil immersion objective 63×. Images of the DAPI stained nucleus were captured using excitation wavelength 405 nm (blue laser), emission wavelengths 410-496 nm and laser power 0.8%. The cellular uptake of **160** and **186** was captured using excitation wavelength 405 nm (blue laser) and emission wavelengths 565-695 nm and laser power 5%.

References

[1]   V. D'Aloisio, P. Dognini, G. A. Hutcheon, C. R. Coxon, *Drug Discov. Today* **2021**, *26*, 1409–1419.

[2]   J. Clayden, N. Greeves, S. Warren, *Organic Chemistry*, 2nd Ed., Oxford University Press, **2012**.

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

[4]   R. B. Merrifield, J. M. Stewart, N. Jernberg, *Anal. Chem.* **1966**, *38*, 1905–1914.

[5]   J. M. Collins, M. J. Collins, R. C. Steorts, *Am. Pept. Symp.* **2003**, *71*, 361–366.

[6]   J. M. Collins, K. A. Porter, S. K. Singh, G. S. Vanier, *Org. Lett.* **2014**, *16*, 940–943.

[7]   M. W. Pennington, B. Zell, C. J. Bai, *Med. Drug Discov.* **2021**, *9*, 100071.

[8]   M. Goodman, W. Cai, N. C. Smith, *J. Pept. Sci.* **2003**, *9*, 594–603.

[9]   F. G. Banting, C. H. Best, J. B. Collip, J. J. R. Macleod, E. C. Noble, *Am. J. Physiol. Content* **1922**, *62*, 162–176.

[10]   V. du Vigneaud, C. Resslbr, J. M. Swan, C. W. Roberts, P. G. Katsoyannis, *J. Am. Chem. Soc.* **1954**, *76*, 3115–3121.

[11]   D. V. Goeddel, D. G. Kleid, F. Bolivar, *Proc. Natl. Acad. Sci. U. S. A.* **1979**, *76*, 106–110.

[12]   M. A. F. Hayashi, F. Ducancel, K. Konno, *Int. J. Pept.* **2012**, 757838.

[13]   D. J. Craik, D. P. Fairlie, S. Liras, D. Price, *Chem. Biol. Drug Des.* **2013**, *81*, 136–147.

[14]   T. Uhlig, T. Kyprianou, F. G. Martinelli, C. A. Oppici, D. Heiligers, D. Hills, X. R. Calvo, P. Verhaert, *EuPA Open Proteomics* **2014**, *4*, 58–69.

[15]   V. Azzarito, K. Long, N. S. Murphy, A. J. Wilson, *Nat. Chem.* **2013**, *5*, 161–173.

[16]   F. G. Banting, C. H. Best, J. B. Collip, J. J. R. Macleod, E. C. Noble, *Am. J. Physiol. Content* **1922**, *62*, 559–580.

[17]   J. L. Lau, M. K. Dunn, *Bioorg. Med. Chem.* **2018**, *26*, 2700–2707.

[18]   A. C. L. Lee, J. L. Harris, K. K. Khanna, J. H. Hong, *Int. J. Mol. Sci.* **2019**, *20*, 1–21.

[19]   V. Martin, P. H. G. Egelund, H. Johansson, S. Thordal Le Quement, F. Wojcik, D. Sejer Pedersen, *RSC Adv.* **2020**, *10*, 42457–42492.

[20]   A. A. Vinogradov, Y. Yin, H. Suga, *J. Am. Chem. Soc.* **2019**, *141*, 4167–4181.

[21]   H. Zhang, S. Chen, *RSC Chem. Biol.* **2022**, *3*, 18–31.

[22]   M. A. Abdalla, L. J. McGaw, *Molecules* **2018**, *23*.

[23]   R. Tugyi, G. Mezö, E. Fellinger, D. Andreu, F. Hudecz, *J. Pept. Sci.* **2005**, *11*, 642–649.

[24]   J. Matsoukas, V. Apostolopoulos, H. Kalbacher, A. M. Papini, T. Tselios, K. Chatzantoni, T. Biagioli, F. Lolli, S. Deraos, P. Papathanassopoulos, A. Troganis, E. Mantzourani, T. Mavromoustakos, A. Mouzaki, *J. Med. Chem.* **2005**, *48*, 1470–1480.

[25]   S. J. Bogdanowich-Knipp, S. Chakrabarti, T. D. Williams, R. K. Dillman, T. J. Siahaan, *J. Pept. Res.* **1999**, *53*, 530–541.

[26]   T. A. Hill, N. E. Shepherd, F. Diness, D. P. Fairlie, *Angew. Chem. Int. Ed.* **2014**, *53*, 13020–13041.

[27]   L. Y. Chan, D. J. Craik, N. L. Daly, *Sci. Rep.* **2016**, *6*, 1–13.

[28]   P. G. Dougherty, A. Sahni, D. Pei, *Chem. Rev.* **2019**, *119*, 10241–10287.

[29]   J. S. Choi, S. H. Joo, *Biomol. Ther.* **2020**, *28*, 18–24.

[30]   A. A. Rzepiela, L. A. Viarengo-Baker, V. Tatarskii, R. Kombarov, A. Whitty, *J. Med. Chem.* **2022**, *65*, 10300–10317.

[31]   W. Liu, C. Wu, *Chin. Chem. Lett.* **2018**, *29*, 1063–1066.

[32]   S. J. De Veer, M. W. Kan, D. J. Craik, *Chem. Rev.* **2019**, *119*, 12375–12421.

[33]   W. D. Chey, A. J. Lembo, B. J. Lavins, S. J. Shiff, C. B. Kurtz, M. G. Currie, J. E. MacDougall, X. D. Jia, J. Z. Shao, D. A. Fitch, M. J. Baird, H. A. Schneier, J. M. Johnston, *Am. J. Gastroenterol.* **2012**, *107*, 1702–1712.

[34]   C. A. Rhodes, D. Pei, *Chem. - A Eur. J.* **2017**, *23*, 12690–12703.

[35]   "RCSB Protein Data Bank (RCSB PDB)," can be found under https://www.rcsb.org/

[36]   C. K. Wang, H. Shu-Hong, J. L. Martin, T. Sjögren, J. Hajdu, L. Bohlin, P. Claeson, U. Göransson, K. J. Rosengren, J. Tang, T. Ning-Hua, D. J. Craik, *J. Biol. Chem.* **2009**, *284*, 10672–10683.

[37]   X. Li, S. Chen, W. D. Zhang, H. G. Hu, *Chem. Rev.* **2020**, *120*, 10079–10144.

[38]   C. Bechtler, C. Lamers, *RSC Med. Chem.* **2021**, *12*, 1325–1351.

[39]   R. Wills, V. Adebomi, M. Raj, *ChemBioChem* **2021**, *22*, 52–62.

[40]   B. Li, Z. Wan, H. Zheng, S. Cai, H. W. Tian, H. Tang, X. Chu, G. He, D. S. Guo, X. S. Xue, G. Chen, *J. Am. Chem. Soc.* **2022**, *144*, 10080–10090.

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

[42]   P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. H. Kent, *Science* **1994**, *266*, 776–779.

[43]   S. J. Miller, R. H. Grubbs, *J. Am. Chem. Soc.* **1995**, *117*, 5855–5856.

[44]   P. Timmerman, J. Beld, W. C. Puijk, R. H. Meloen, *ChemBioChem* **2005**, *6*, 821–824.

[45]   P. Lin, H. Yao, J. Zha, Y. Zhao, C. Wu, *ChemBioChem* **2019**, 1514–1518.

[46]   G. J. J. Richelle, M. Schmidt, H. Ippel, T. M. Hackeng, J. H. van Maarseveen, T.

Nuijens, P. Timmerman, *ChemBioChem* **2018**, *19*, 1934–1938.

[47]   G. J. J. Richelle, S. Ori, H. Hiemstra, J. H. van Maarseveen, P. Timmerman, *Angew. Chem. Int. Ed.* **2018**, *57*, 501–505.

[48]   D. E. Streefkerk, M. Schmidt, J. H. Ippel, T. M. Hackeng, T. Nuijens, P. Timmerman, J. H. Van Maarseveen, *Org. Lett.* **2019**, *21*, 2095–2100.

[49]   P. R. Werkhoven, M. Elwakiel, T. J. Meuleman, H. C. Quarles Van Ufford, J. A. W. Kruijtzer, R. M. J. Liskamp, *Org. Biomol. Chem.* **2016**, *14*, 701–710.

[50]   Y. Yin, Q. Fei, W. Liu, Z. Li, H. Suga, C. Wu, *Angew. Chem. Int. Ed.* **2019**, *58*, 4880–4885.

[51]   A. M. Spokoyny, Y. Zou, J. J. Ling, H. Yu, Y. S. Lin, B. L. Pentelute, *J. Am. Chem. Soc.* **2013**, *135*, 5946–5949.

[52]   R. G. Jones, C. K. Ober, P. Hodge, P. Kratochvíl, G. Moad, M. Vert, *Pure Appl. Chem.* **2013**, *85*, 463–492.

[53]   A. Levin, T. A. Hakala, L. Schnaider, G. J. L. Bernardes, E. Gazit, T. P. J. Knowles, *Nat. Rev. Chem.* **2020**, *4*, 615–634.

[54]   Z. Luo, S. Zhang, *Chem. Soc. Rev.* **2012**, *41*, 4736–4754.

[55]   M. Rad-Malekshahi, L. Lempsink, M. Amidi, W. E. Hennink, E. Mastrobattista, *Bioconjug. Chem.* **2016**, *27*, 3–18.

[56]   B. B. Gerbelli, S. V. Vassiliades, J. E. U. Rojas, J. N. B. D. Pelin, R. S. N. Mancini, W. S. G. Pereira, A. M. Aguilar, M. Venanzi, F. Cavalieri, F. Giuntini, W. A. Alves, *Macromol. Chem. Phys.* **2019**, *220*, 1–22.

[57]   S. Gupta, I. Singh, A. K. Sharma, P. Kumar, *Front. Bioeng. Biotechnol.* **2020**, *8*, 1–36.

[58]   A. Dehsorkhi, V. Castelletto, I. W. Hamley, *J. Pept. Sci.* **2014**, *20*, 453–467.

[59]   R. C. Claussen, B. M. Rabatic, S. I. Stupp, *J. Am. Chem. Soc.* **2003**, *125*, 12680–12681.

[60]   I. W. Hamley, *Angew. Chemie - Int. Ed.* **2014**, *53*, 6866–6881.

[61]   D. Gottlieb, S. A. Morin, S. Jin, R. T. Raines, *J. Mater. Chem.* **2008**, *18*, 3865–3870.

[62]   P. Dognini, C. R. Coxon, W. A. Alves, F. Giuntini, *Molecules 2021,* **2021**, *26*, 693.

[63]   "BioRender," can be found under https://www.biorender.com/

[64]   J. Li, R. Xing, S. Bai, X. Yan, *Soft Matter* **2019**, *15*, 1704–1715.

[65]   M. Lian, X. Chen, Y. Lu, W. Yang, *ACS Appl. Mater. Interfaces* **2016**, *8*, 25036–25042.

[66]   T. Fan, X. Yu, B. Shen, L. Sun, *J. Nanomater.* **2017**

[67]   Z. Yang, H. Xu, X. Zhao, *Adv. Sci.* **2020**, 7.

[68]   M. Reches, E. Gazit, *Science* **2003**, *300*, 625–627.

[69]   R. Ravichandran, M. Griffith, J. Phopase, *J. Mater. Chem. B* **2014**, *2*, 8466–8478.

[70]   J. Chen, X. Zou, *Bioact. Mater.* **2019**, *4*, 120–131.

[71]   R. Fairman, K. S. Åkerfeldt, *Curr. Opin. Struct. Biol.* **2005**, *15*, 453–463.

[72]   M. Zelzer, R. V. Ulijn, *Chem. Soc. Rev.* **2010**, *39*, 3351–3357.

[73]   L. B. Josefsen, R. W. Boyle, *Met. Based. Drugs* **2008**

[74]   J. K. Platten, P. Costesèque, *Eur. Phys. J. E* **2004**, *15*, 235–239.

[75]   A. Jabłoństski, *Nature* **1933**, *131*, 839–840.

[76]   P. Rothemund, *J. Am. Chem. Soc.* **1935**, *57*, 2010–2011.

[77]   A. D. Adler, F. R. Longo, J. D. Finarelli, J. Goldmacher, J. Assour, L. Korsakoff, *J. Org. Chem.* **1967**, *32*, 476.

[78]   J. S. Lindsey, I. C. Schreiman, H. C. Hsu, P. C. Kearney, A. M. Marguerettaz, *J. Org. Chem.* **1987**, *52*, 827–836.

[79]   R. G. Little, *J. Heterocycl. Chem.* **1981**, *18*, 129–133.

[80]   Z. Abada, L. Ferrié, B. Akagah, A. T. Lormier, B. Figadre, *Tetrahedron Lett.* **2011**, *52*, 3175–3178.

[81]   M. O. Senge, N. N. Sergeeva, K. J. Hale, *Chem. Soc. Rev.* **2021**, *50*, 4730–4789.

[82]   A. R. Battersby, *Nat. Prod. Rep.* **2000**, *17*, 507–526.

[83]   M. R. Wasielewski, *J. Org. Chem.* **2006**, *71*, 5051–5066.

[84]   Ö. Birel, S. Nadeem, H. Duman, *J. Fluoresc.* **2017**, *27*, 1075–1085.

[85]   R. Costa e Silva, L. O. da Silva, A. de Andrade Bartolomeu, T. J. Brocksom, K. T. de Oliveira, *Beilstein J. Org. Chem.* **2020**, *16*, 917–955.

[86]   M. H. Gehlen, *J. Photochem. Photobiol. C Photochem. Rev.* **2020**, *42*, 100338.

[87]   H. Huang, W. Song, J. Rieffel, J. F. Lovell, *Front. Phys.* **2015**, *3*, 1–15.

[88]   P. Fathi, Di. Pan, *Nanomedicine* **2020**, *15*, 2493–2515.

[89]   T. J. Dougherty, C. J. Gomer, B. W. Henderson, G. Jori, D. Kessel, M. Korbelik, J. Moan, Q. Peng, *JNCI J. Natl. Cancer Inst.* **1998**, *90*, 889–905.

[90]   J. H. Correia, J. A. Rodrigues, S. Pimenta, T. Dong, Z. Yang, *Pharm.* **2021**, *13*, 1332.

[91]   J. D. Spikes, *J. Photochem. Photobiol. B Biol.* **1991**, *9*, 369–371.

[92]   J. F. Kelly, M. E. Snell, *J. Urol.* **1976**, *115*, 150–151.

[93]   R. Baskaran, J. Lee, S. G. Yang, *Biomater. Res.* **2018**, *22*, 1–8.

[94]   D. S. Wishart, Y. D. Feunang, A. C. Guo, E. J. Lo, A. Marcu, J. R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, N. Assempour, I. Iynkkaran, Y. Liu, A. MacIejewski, N.

Gale, A. Wilson, L. Chin, R. Cummings, Di. Le, A. Pon, C. Knox, M. Wilson, *Nucleic Acids Res.* **2018**, *46*, D1074–D1082.

[95] C. Sornay, V. Vaur, A. Wagner, G. Chaubet, *An Overview of Chemo- and Site-Selectivity Aspects in the Chemical Conjugation of Proteins*, **2022**.

[96] "The Nobel Prize," can be found under https://www.nobelprize.org/

[97] H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021.

[98] C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057–3064.

[99] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2002**, *41*, 2596–2599.

[100] J. E. Moses, A. D. Moorhouse, *Chem. Soc. Rev.* **2007**, *36*, 1249–1262.

[101] J. Kaur, M. Saxena, N. Rishi, *Bioconjug. Chem.* **2021**, *32*, 1455–1471.

[102] N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047.

[103] D. H. Dube, C. R. Bertozzi, *Curr. Opin. Chem. Biol.* **2003**, *7*, 616–625.

[104] C. Remzi Becer, R. Hoogenboom, U. S. Schubert, *Angew. Chem. Int. Ed.* **2009**, *48*, 4900–4908.

[105] P. Pathak, M. A. Zarandi, X. Zhou, J. Jayawickramarajah, *Front. Chem.* **2021**, *9*, 1–30.

[106] Y. Wu, H. F. Chau, Y. H. Yeung, W. Thor, H. Y. Kai, W. L. Chan, K. L. Wong, *Angew. Chem. Int. Ed.* **2022**, *202207532*, 1–8.

[107] L. A. Halsall, *Using Design of Experiments (DoE) to Optimise a Reaction between Hexafluorobenzene and n-Acetylcysteine, B.Sc. Applied Chemistry, Final Year Research Project*, **2019**.

[108] A. J. J. Lennox, *Angew. Chem. Int. Ed.* **2018**, *57*, 14686–14688.

[109] L. A. Wall, W. J. Pummer, J. E. Fearn, J. M. Antonucci, *J. Res. Natl. Bur. Stand. Sect. A Phys. Chem.* **1963**, *67A*, 481.

[110] D. O'Hagan, *Chem. Soc. Rev.* **2008**, *37*, 308–319.

[111] E. P. Gillis, K. J. Eastman, M. D. Hill, D. J. Donnelly, N. A. Meanwell, *J. Med. Chem.* **2015**, *58*, 8315–8359.

[112] S. L. Cobb, C. D. Murphy, *J. Fluor. Chem.* **2009**, *130*, 132–143.

[113] D. O'Hagan, *Chem. Soc. Rev.* **2008**, *37*, 308–319.

[114] "The Royal Society of Chemistry," can be found under https://www.rsc.org/periodic-table/element/9/fluorine

[115] N. N. Vorozhtsov, V. E. Platonov, G. G. Yakobson, *Bull. Acad. Sci. USSR Div. Chem. Sci.* **1963**, *12*, 1389.

[116] R. P. Mason, W. Rodbumrung, P. P. Antich, *NMR Biomed.* **1996**, *9*, 125–134.

[117] A. Lattanzi, C. de Fusco, A. Russo, A. Poater, L. Cavallo, *Chem. Commun.* **2012**, *48*, 1650–1652.

[118] S. E. Denmark, A. J. Cresswell, *J. Org. Chem.* **2013**, *78*, 12593–12628.

[119] I. P. Beletskaya, G. A. Artamkina, A. Y. Mil'Chenko, P. K. Sazonov, M. M. Shtern, *J. Phys. Org. Chem.* **1996**, *9*, 319–328.

[120] J. Ponce González, M. Edgar, M. R. J. Elsegood, G. W. Weaver, *Org. Biomol. Chem.* **2011**, *9*, 2294–2305.

[121] O. Jacobson, X. Yan, Y. Ma, G. Niu, D. O. Kiesewetter, X. Chen, *Bioconjug. Chem.* **2015**, *26*, 2016–2020.

[122] S. Niembro, A. Vallribera, M. Moreno-Mañas, *New J. Chem.* **2008**, *32*, 94–98.

[123] R. D. Kennedy, C. W. MacHan, C. M. McGuirk, M. S. Rosen, C. L. Stern, A. A. Sarjeant, C. A. Mirkin, *Inorg. Chem.* **2013**, *52*, 5876–5888.

[124] T. Umemoto, L. M. Garrick, N. Saito, *Beilstein J. Org. Chem.* **2012**, *8*, 461–471.

[125] M. Arisawa, T. Suzuki, T. Ishikawa, M. Yamaguchi, *J. Am. Chem. Soc.* **2008**, *130*, 12214–12215.

[126] M. Mayor, J.-M. Lehn, *Helv. Chim. Acta* **1997**, *80*, 2277–2285.

[127] H. Wu, W. Chi, G. Baryshnikov, B. Wu, Y. Gong, D. Zheng, X. Li, Y. Zhao, X. Liu, H. Ågren, L. Zhu, *Angew. Chem. Int. Ed.* **2019**, *58*, 4328–4333.

[128] L. Testaferri, M. Tingoli, M. Tiecco, *J. Org. Chem.* **1980**, *45*, 4376–4380.

[129] B. Gierczyk, G. Schroeder, *Mendeleev Commun.* **2009**, *19*, 75–77.

[130] M. Villa, M. Roy, G. Bergamini, M. Gingras, P. Ceroni, *Dalton Trans.* **2019**, *48*, 3815–3818.

[131] J. N. Lowe, D. A. Fulton, S. H. Chiu, A. M. Elizarov, S. J. Cantrill, S. J. Rowan, J. F. Stoddart, *J. Org. Chem.* **2004**, *69*, 4390–4402.

[132] R. Leardi, *Anal. Chim. Acta* **2009**, *652*, 161–172.

[133] O. Acevedo, W. L. Jorgensen, *Org. Lett.* **2004**, *6*, 2881–2884.

[134] Sigma-Aldrich, "Physical Properties of Solvents," can be found under https://www.sigmaaldrich.com/deepweb/assets/sigmaaldrich/marketing/global/doc uments/614/456/labbasics_pg144.pdf

[135] D. D. Perrin, *Pure Appl. Chem.* **1969**, *20*, 133–236.

[136] R. Srivastava, *J. Mol. Catal. A Chem.* **2007**, *264*, 146–152.

[137] M. Hudlicky, *J. Fluor. Chem.* **1985**, *28*, 461–472.

[138] K. O. Christe, W. W. Wilson, *J. Fluor. Chem.* **1990**, *46*, 339–342.

[139] W. Liu, Y. Zheng, X. Kong, C. Heinis, Y. Zhao, C. Wu, *Angew. Chem. Int. Ed.* **2017**, *56*, 4458–4463.

[140] Molinspiration Cheminformatics 2023, "Calculation of Molecular Properties and Bioactivity Score," can be found under https://www.molinspiration.com/cgi-bin/properties

[141] N. A. Isley, R. T. H. Linstadt, S. M. Kelly, F. Gallou, B. H. Lipshutz, *Org. Lett.* **2015**, *17*, 4734–4737.

[142] N. R. Lee, F. Gallou, B. H. Lipshutz, *Org. Process Res. Dev.* **2017**, *21*, 218–221.

[143] B. H. Lipshutz, S. Ghorai, A. R. Abela, R. Moser, T. Nishikata, C. Duplais, A. Krasovskiy, R. D. Gaston, R. C. Gadwood, *J. Org. Chem.* **2011**, *76*, 4379–4391.

[144] Y. Zou, A. M. Spokoyny, C. Zhang, M. D. Simon, H. Yu, Y. S. Lin, B. L. Pentelute, *Org. Biomol. Chem.* **2014**, *12*, 566–573.

[145] C. Zhang, P. Dai, A. M. Spokoyny, B. L. Pentelute, *Org. Lett.* **2014**, *16*, 3652–3655.

[146] R. J. Steel, M. A. O'Connell, M. Searcey, *Bioorg. Med. Chem. Lett.* **2018**, *28*, 2728–2731.

[147] J. E. Swedberg, C. I. Schroeder, J. M. Mitchell, T. Durek, D. P. Fairlie, D. J. Edmonds, D. A. Griffith, R. B. Ruggeri, D. R. Derksen, P. M. Loria, S. Liras, D. A. Price, D. J. Craik, *Eur. J. Med. Chem.* **2015**, *103*, 175–184.

[148] A. C. Conibear, S. Chaousis, T. Durek, K. Johan Rosengren, D. J. Craik, C. I. Schroeder, *Biopolymers* **2016**, *106*, 89–100.

[149] D. Gimenez, C. A. Mooney, A. Dose, G. Sandford, C. R. Coxon, S. L. Cobb, *Org. Biomol. Chem.* **2017**, *15*, 4086–4095.

[150] S. J. M. Verhoork, C. E. Jennings, N. Rozatian, J. Reeks, J. Meng, E. K. Corlett, F. Bunglawala, M. E. M. Noble, A. G. Leach, C. R. Coxon, *Chem. Eur. J.* **2019**, *25*, 177–182.

[151] T. Lühmann, S. K. Mong, M. D. Simon, L. Meinel, B. L. Pentelute, *Org. Biomol. Chem.* **2016**, *14*, 3345–3349.

[152] E. D. Evans, Z. P. Gates, Z. Y. J. Sun, A. J. Mijalis, B. L. Pentelute, *Biochemistry* **2019**, *58*, 1343–1353.

[153] S. Tuang, D. Dieppa-Matos, C. Zhang, C. R. Shugrue, P. Dai, A. Loas, B. L. Pentelute, *Chem. Commun.* **2021**, *57*, 3227–3230.

[154] C. M. Fadzen, J. M. Wolfe, C. F. Cho, E. A. Chiocca, S. E. Lawler, B. L. Pentelute, *J. Am. Chem. Soc.* **2017**, *139*, 15628–15631.

[155] C. M. Fadzen, J. M. Wolfe, W. Zhou, C. F. Cho, N. Von Spreckelsen, K. T. Hutchinson,

Y. C. Lee, E. A. Chiocca, S. E. Lawler, O. H. Yilmaz, S. J. Lippard, B. L. Pentelute, *J. Med. Chem.* **2020**, *63*, 6741–6747.

[156] G. Lautrette, F. Touti, H. G. Lee, P. Dai, B. L. Pentelute, *J. Am. Chem. Soc.* **2016**, *138*, 8340–8343.

[157] C. Zhang, A. M. Spokoyny, Y. Zou, M. D. Simon, B. L. Pentelute, *Angew. Chem. Int. Ed.* **2013**, *52*, 14001–14005.

[158] C. Zhang, M. Welborn, T. Zhu, N. J. Yang, M. S. Santos, T. Van Voorhis, B. L. Pentelute, *Nat. Chem.* **2016**, *8*, 120–128.

[159] P. Dai, C. Zhang, M. Welborn, J. J. Shepherd, T. Zhu, T. Van Voorhis, B. L. Pentelute, *ACS Cent. Sci.* **2016**, *2*, 637–646.

[160] P. Dai, J. K. Williams, C. Zhang, M. Welborn, J. J. Shepherd, T. Zhu, T. Van Voorhis, M. Hong, B. L. Pentelute, *Sci. Rep.* **2017**, *7*, 1–11.

[161] E. D. Evans, B. L. Pentelute, *Org. Biomol. Chem.* **2019**, *17*, 1862–1868.

[162] E. D. Evans, B. L. Pentelute, *ACS Chem. Biol.* **2018**, *13*, 527–532.

[163] W. D. G. Brittain, S. L. Cobb, *Org. Biomol. Chem.* **2019**, *17*, 2110–2115.

[164] A. M. Webster, C. R. Coxon, A. M. Kenwright, G. Sandford, S. L. Cobb, *Tetrahedron* **2014**, *70*, 4661–4667.

[165] Y. H. Lau, P. De Andrade, Y. Wu, D. R. Spring, *Chem. Soc. Rev.* **2015**, *44*, 91–102.

[166] T. Hsu, S. Mitragotri, *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 15816–15821.

[167] T. Chaudhry, *Peptide Conjugates for MicroRNA-21 Inhibition in Glioblastoma Multiforme*, Liverpool John Moores University, **2023**.

[168] J. M. Wolfe, C. M. Fadzen, R. L. Holden, M. Yao, G. J. Hanson, B. L. Pentelute, *Angew. Chem. Int. Ed.* **2018**, *57*, 4756–4759.

[169] R. J. Spears, C. McMahon, V. Chudasama, *Chem. Soc. Rev.* **2021**, *50*, 11098–11155.

[170] T. M. Postma, M. Giraud, F. Albericio, *Org. Lett.* **2012**, *14*, 5468–5471.

[171] N. J. C. M. Beekman, W. M. M. Schaaper, G. I. Tesser, K. Dalsgaard, S. Kamstrup, J. P. M. Langeveld, R. S. Boshuizen, R. H. Meloen, *J. Pept. Res.* **1997**, *50*, 357–364.

[172] M. Pakkala, C. Hekim, P. Soininen, J. Leinonen, H. Koistinen, J. Weisell, U. H. Stenman, J. Vepsäläinen, A. Närvänen, *J. Pept. Sci.* **2007**, *13*, 348–353.

[173] O. S. Gudmundsson, G. M. Pauletti, W. Wang, D. Shan, H. Zhang, B. Wang, R. T. Borchardt, *Pharm. Res.* **1999**, *16*, 7–15.

[174] M. Góngora-Benítez, J. Tulla-Puche, M. Paradís-Bas, O. Werbitzky, M. Giraud, F. Albericio, *Biopolymers* **2011**, *96*, 69–80.

[175] S. K. Maity, M. Jbara, S. Laps, A. Brik, *Angew. Chem. Int. Ed.* **2016**, *55*, 8108–8112.

[176] K. M. kadish, C. Araullo-McAdams, B. C. Han, M. M. Franzen, *J. Am. Chem. Soc.* **1990**, *112*, 8364–8368.

[177] J. I. T. Costa, A. C. Tomé, M. G. P. M. S. Neves, J. A. S. Cavaleiro, *J. Porphyr. Phthalocyanines* **2011**, *15*, 1116–1133.

[178] G. Q. Jin, J. X. Wang, J. Lu, H. Zhang, Y. Yao, Y. Ning, H. Lu, S. Gao, J. L. Zhang, *Chem. Sci.* **2023**, *14*, 2070–2081.

[179] F. Giuntini, C. M. A. Alonso, R. W. Boyle, *Photochem. Photobiol. Sci.* **2011**, *10*, 759–791.

[180] S. L. Bellis, *Biomaterials* **2011**, *32*, 4205–4210.

[181] T. A. Enache, A. M. Chiorcea-Paquim, A. M. Oliveira-Brett, *Anal. Chem.* **2018**, *90*, 2285–2292.

[182] S. Lazzaro, N. Ogrinc, L. Lamont, G. Vecchio, G. Pappalardo, R. M. A. Heeren, *Anal. Bioanal. Chem.* **2019**, *411*, 6353–6363.

[183] R. Azriel, E. Gazit, *J. Biol. Chem.* **2001**, *276*, 34156–34161.

[184] E. Gazit, *FASEB J.* **2002**, *16*, 77–83.

[185] P. Dognini, P. M. Killoran, G. S. Hanson, L. Halsall, T. Chaudhry, Z. Islam, F. Giuntini, C. R. Coxon, *Pept. Sci.* **2021**, *113*, e24182.

[186] P. Behrens, A. M. Glaue, *Monatsh. Chem.* **2002**, *133*, 1405–1412.

[187] M. Morisue, T. Morita, Y. Kuroda, *Org. Biomol. Chem.* **2010**, *8*, 3457–3463.

[188] S. Tu, S. H. Kim, J. Joseph, D. A. Modarelli, J. R. Parquette, *J. Am. Chem. Soc.* **2011**, *133*, 19125–19130.

[189] F. Qiu, Y. Chen, C. Tang, Q. Zhou, C. Wang, Y. K. Shi, X. Zhao, *Macromol. Biosci.* **2008**, *8*, 1053–1059.

[190] F. Qiu, C. Tang, Y. Chen, *J. Pept. Sci.* **2018**, *24*, 1–9.

[191] E. R. Da Silva, M. N. M. Walter, M. Reza, V. Castelletto, J. Ruokolainen, C. J. Connon, W. A. Alves, I. W. Hamley, *Biomacromolecules* **2015**, *16*, 3180–3190.

[192] E. R. Da Silva, W. A. Alves, V. Castelletto, M. Reza, J. Ruokolainen, R. Hussain, I. W. Hamley, *Chem. Commun.* **2015**, *51*, 11634–11637.

[193] L. Yao, M. He, D. Li, H. Liu, J. Wu, J. Xiao, *New J. Chem.* **2018**, *42*, 7439–7444.

[194] R. Djalali, J. Samson, H. Matsui, *J. Am. Chem. Soc.* **2004**, *126*, 7935–7939.

[195] H. Wei, J. Min, Y. Wang, Y. Shen, Y. Du, R. Su, W. Qi, *J. Mater. Chem. B* **2022**, 9334–9348.

[196] S. M. Borisov, P. Lehner, I. Klimant, *Anal. Chim. Acta* **2011**, *690*, 108–115.

[197] M. Ohi, Y. Li, Y. Cheng, T. Walz, *Biol. Proced. Online* **2004**, *6*, 23–34.

[198] F. Li, D. Liu, T. Wang, J. Hu, F. Meng, H. Sun, Z. Shang, P. Li, W. Feng, W. Li, X. Zhou, *J. Solid State Chem.* **2017**, *252*, 86–92.

[199] S. Sonzini, S. T. Jones, Z. Walsh, O. A. Scherman, *Analyst* **2015**, *140*, 2735–2740.

[200] H. C. Fry, L. A. Solomon, B. T. Diroll, Y. Liu, D. J. Gosztola, H. M. Cohn, *J. Am. Chem. Soc.* **2020**, *142*, 233–241.

[201] G. Charalambidis, E. Georgilis, M. K. Panda, C. E. Anson, A. K. Powell, S. Doyle, D. Moss, T. Jochum, P. N. Horton, S. J. Coles, M. Linares, D. Beljonne, J. V. Naubron, J. Conradt, H. Kalt, A. Mitraki, A. G. Coutsolelos, T. S. Balaban, *Nat. Commun.* **2016**, *7*.

[202] H. Xing, S. M. Chin, V. R. Udumula, M. Krishnaiah, N. R. De Almeida, C. Huck-iriart, S. Picco, S. R. Lee, G. Zaldivar, K. A. Jackson, M. Tagliazucchi, S. I. Stupp, M. Conda-sheridan, **2021**, *22*, 3274-3283.

[203] O. Ohno, Y. Kaizu, H. Kobayashi, *J. Chem. Phys.* **1993**, *99*, 4128–4139.

[204] V. V. Serra, S. M. Andrade, M. G. P. M. S. Neves, J. A. S. Cavaleiro, S. M. B. Costa, *New J. Chem.* **2010**, *34*, 2757–2765.

[205] H. Isago, *Chem. Commun.* **2003**, *3*, 1864–1865.

[206] "Bachem Peptide Calculator," can be found under https://www.bachem.com/knowledge-center/peptide-calculator/

[207] A. Stein, J. P. Whitlock, M. Bina, *Proc. Natl. Acad. Sci. U. S. A.* **1979**, *76*, 5000–5004.

[208] A. S. D. Sandanayaka, Y. Araki, T. Wada, T. Hasobe, *J. Phys. Chem. C* **2008**, *112*, 19209–19216.

[209] N. C. Maiti, S. Mazumdar, N. Periasamy, *J. Phys. Chem. B* **1998**, *102*, 1528–1538.

[210] D. D. Kim, J. Heo, S. Ham, H. Yoo, C. H. Lee, H. Yoon, D. Ryu, D. D. Kim, W. D. Jang, *Chem. Commun.* **2011**, *47*, 2405–2407.

[211] J. C. P. Grancho, M. M. Pereira, M. da G. Miguel, A. M. R. Gonsalves, H. D. Burrows, *Photochem. Photobiol.* **2002**, *75*, 249.

[212] K. I. Yamashita, N. Katsumata, S. Tomita, M. Fuwa, K. Fujimaki, T. Yoda, D. Hirano, K. I. Sugiura, *Chem. Lett.* **2015**, *44*, 492–494.

[213] S. W. Lai, Y. J. Hou, C. M. Che, H. L. Pang, K. Y. Wong, C. K. Chang, N. Zhu, *Inorg. Chem.* **2004**, *43*, 3724–3732.

[214] W. P. To, Y. Liu, T. C. Lau, C. M. Che, *Chem. Eur. J.* **2013**, *19*, 5654–5664.

[215] A. F. De Moraes Filho, P. M. Gewehr, J. M. Maia, D. R. Jakubiak, *Sensors (Switzerland)* **2018**, *18*, 1–17.

[216] Q. Zou, M. Abbas, L. Zhao, S. Li, G. Shen, X. Yan, *J. Am. Chem. Soc.* **2017**, *139*, 1921–

1927.

[217] C. Ren, Z. Wang, X. Zhang, J. Gao, Y. Gao, Y. Zhang, J. J. Liu, C. Yang, J. J. Liu, *Chem. Eng. J.* **2021**, *405*, 127008.

[218] D. C. Dumitras, D. C. Dutu, C. Matei, A. M. Magureanu, M. Petrus, C. Popa, *J. Optoelectron. Adv. Mater.* **2007**, *9*, 3655–3701.

[219] S. Abbruzzetti, C. Viappiani, D. H. Murgida, R. Erra-Balsells, G. M. Bilmes, *Chem. Phys. Lett.* **1999**, *304*, 167–172.

[220] T. Gensch, C. Viappiani, S. E. Braslavsky, *J. Am. Chem. Soc.* **1999**, *121*, 10573–10582.

[221] A. Dasgupta, D. Das, *Langmuir* **2019**, *35*, 10704–10724.

[222] J. Vangindertael, R. Camacho, W. Sempels, H. Mizuno, P. Dedecker, K. P. F. Janssen, *Methods Appl. Fluoresc.* **2018**, *6*.

[223] S. Pujals, N. Feiner-Gracia, P. Delcanale, I. Voets, L. Albertazzi, *Nat. Rev. Chem.* **2019**, *3*, 68–84.

[224] T. Andrian, P. Delcanale, S. Pujals, L. Albertazzi, *Nano Lett.* **2021**, *21*, 5360–5368.

[225] X. Ma, J. Jia, R. Cao, X. Wang, H. Fei, *J. Am. Chem. Soc.* **2014**, *136*, 17734–17737.

[226] S. Ji, X. Yang, X. Chen, A. Li, D. Yan, H. Xu, H. Fei, *Chem. Sci.* **2020**, *11*, 9126–9133.

[227] M. Todorovic, K. D. Schwab, J. Zeisler, C. Zhang, F. Bénard, D. M. Perrin, *Angew. Chem. Int. Ed.* **2019**, *58*, 14120–14124.

[228] Y. Wu, H. F. Chau, W. Thor, K. H. Y. Chan, X. Ma, W. L. Chan, N. J. Long, K. L. Wong, *Angew. Chem. Int. Ed.* **2021**, *60*, 20301–20307.

[229] L. Nevola, A. Martín-Quirõs, K. Eckelt, N. Camarero, S. Tosi, A. Llobet, E. Giralt, P. Gorostiza, *Angew. Chem. Int. Ed.* **2013**, *52*, 7704–7708.

[230] S. P. Brown, A. B. Smith, *J. Am. Chem. Soc.* **2015**, *137*, 4034–4037.

[231] S. Learte-Aymamí, C. Vidal, A. Gutiérrez-González, J. L. Mascareñas, *Angew. Chem. Int. Ed.* **2020**, *59*, 9149–9154.

[232] A. Hamilton, J. M. Lehn, J. L. Sessler, *J. Am. Chem. Soc.* **1986**, *108*, 5158–5167.

[233] J. A. Faiz, V. Heitz, J. P. Sauvage, *Chem. Soc. Rev.* **2009**, *38*, 422–442.

[234] K. Sugiyasu, S. Ogi, M. Takeuchi, *Polym. J. 2014 4610* **2014**, *46*, 674–681.

[235] C. Roche, Q. Luo, G. Gil-Ramírez, H. W. Jiang, D. R. Kohn, Y. Xiong, A. L. Thompson, H. L. Anderson, *J. Org. Chem.* **2017**, *82*, 7446–7462.

[236] J. P. Collman, V. J. Lee, C. J. Kellen-Yuen, X. Zhang, J. A. Ibers, J. I. Brauman, *J. Am. Chem. Soc.* **1995**, *117*, 692–703.

[237] J. A. A. W. Elemans, R. J. M. Nolte, *Chem. Commun.* **2019**, *55*, 9590–9605.

[238] C. Donohoe, A. Charisiadis, S. Maguire, B. Twamley, F. A. Schaberle, L. C. Gomes-da-Silva, M. O. Senge, *Eur. J. Org. Chem.* **2023**, *26*, e202201453.

[239] B. Boitrel, V. Baveux-Chambenoît, P. Richard, *Eur. J. Org. Chem.* **2001**, 4213–4221.

[240] M. Urbani, T. Torres, *Chem. Eur. J.* **2014**, *20*, 16337–16349.

[241] D. Reddy, T. K. Chandrashekar, *J. Chem. Soc., Dalt. Trans.* **1992**, 619–625.

[242] J. E. Redman, J. K. M. Sanders, *Org. Lett.* **2000**, 2, 4141–4144.

[243] T. B. Karpishin, T. A. Vannelli, K. J. Glover, *J. Am. Chem. Soc.* **1997**, *119*, 9063–9064.

[244] R. R. Geier, T. Sasaki, *Tetrahedron Lett.* **1997**, *38*, 3821–3824.

[245] G. R. Geier, T. Sasaki, *Tetrahedron* **1999**, *55*, 1859–1870.

[246] G. R. Geier, T. P. Lybrand, T. Sasaki, *Tetrahedron* **1999**, *55*, 1871–1880.

[247] M. Champeau, S. Vignoud, L. Mortier, S. Mordon, *J. Photochem. Photobiol. B Biol.* **2019**, *197*, 111544.

[248] M. Chen, M. Zakrewsky, V. Gupta, A. C. Anselmo, D. H. Slee, J. A. Muraski, S. Mitragotri, *J. Control. Release* **2014**, *179*, 33–41.

[249] S. Grimme, C. Bannwarth, P. Shushkov, *J. Chem. Theory Comput.* **2017**, *13*, 1989–2009.

[250] C. Bannwarth, S. Ehlert, S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.

[251] H. Hossein-Nejad-Ariani, E. Althagafi, K. Kaur, *Sci. Rep.* **2019**, *9*, 1–10.

[252] B. Oller-Salvia, M. Sánchez-Navarro, S. Ciudad, M. Guiu, P. Arranz-Gibert, C. Garcia, R. R. Gomis, R. Cecchelli, J. García, E. Giralt, M. Teixidõ, *Angew. Chem. Int. Ed.* **2016**, *55*, 572–575.

[253] M. Sánchez-Sánchez, T. Blasco, *Chem. Commun.* **2000**, 491–492.

[254] P. Cammareri, A. M. Rose, D. F. Vincent, J. Wang, A. Nagano, S. Libertini, R. A. Ridgway, D. Athineos, P. J. Coates, A. McHugh, C. Pourreyron, J. H. S. Dayal, J. Larsson, S. Weidlich, L. C. Spender, G. P. Sapkota, K. J. Purdie, C. M. Proby, C. A. Harwood, I. M. Leigh, H. Clevers, N. Barker, S. Karlsson, C. Pritchard, R. Marais, C. Chelala, A. P. South, O. J. Sansom, G. J. Inman, *Nat. Commun. 2016 71* **2016**, *7*, 1–14.

[255] S. Abbruzzetti, C. Viappiani, J. R. Small, L. J. Libertini, E. W. Small, *J. Am. Chem. Soc.* **2001**, *123*, 6649–6653.

[256] A. W. van der Made, E. J. H. Hoppenbrouwer, R. J. M. Nolte, W. Drenth, *Recl. des Trav. Chim. des Pays-Bas* **1988**, *107*, 15–16.

[257] B. J. Littler, M. A. Miller, C.-H. Hung, R. W. Wagner, D. F. O'shea, P. D. Boyle, J. S. Lindsey, *J. Org. Chem.* **1999**, *64*, 1391–1396.

[258] S. Wang, J. Witek, G. A. Landrum, S. Riniker, *J. Chem. Inf. Model.* **2020**, *60*, 2044–

2058.

[259] "RDKit: Open-Source Cheminformatics Software," can be found under http://www.rdkit.org/

[260] W. Clark Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.