Development of New Synthetic Approaches to Biologically-Active Peptides

Sanne J.M. Verhoork

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

June 2019

Acknowledgements

Back in November 2015 I was at a symposium when I received an email confirming that I had been offered a PhD position in Liverpool. I had never been there before, so this marked the beginning of an exciting, yet terrifying, adventure. I could not have wished for anything more. Before moving to Liverpool, I did have some prejudices. It would be the city of football hooligans. While football definitely turned out to be a big part of the culture, thankfully, the hooligan part was not. On the contrary, Scousers turned out to be lovely, warm, friendly people (though sometimes hard to understand) and the city overall was very welcoming, and I am happy to have called it home for three (and a bit) years.

Though a city, and a university, would be nothing without its people. First, I would like to express my gratitude to my director of studies, Dr. Chris Coxon. You are one of the kindest, most genuine, people I have ever met, and I truly believe I could not have had a better supervisor. Thank you for all the support, both work and non-work related.

Secondly, I would like to mention the other members of my supervisory team, Dr. Adam Sharples and Prof. Claire Stewart, as well as Prof. Lu-Yun Lian from the University of Liverpool. I would also like to express my gratitude to all the technical support staff, including Nicola Dempster, Rob Allen, Phil Salmon, and Campbell Woods.

Thank you to everyone in my research group who I have worked closely with. I will never forget the adventures of PepSquad and Broccoli Boy in Blindhoven or Maastricht (Blindhoven 2.0). I have enjoyed working with you. Also thank you to everyone who I have shared an office with during the time I have been here. You are far too many to mention by now, but the only right word to say is probably *Grazie*!

I would also like to say thank you to all my friends outside of work, Carlota, Carmen, Christian, Ecaterina, Francisco, Ingrid, Josep, Laura, Lauren, Lia, Lidia, Mafalda, Mark, Melanie, Nicoline, Stefania, and Wanda. In addition, there are a few friends who deserve a special mention. Thank you to Sophie, who can probably still recite my 3MT presentation. Thank you to Alex, for our relaxing evenings and days in playing video games or wandering around town trying to catch ‘em all. Thank you to Nathan and Alexandra, for our countless of visits to Down the Hatch and the excellent seitan, as well as providing me with cats in Liverpool. Thank you to Natalie, for our many dinners, coffees and cake. Also, thank you for helping me move house, in the absolute chaos it was! Which brings me to Ellen, thank you for being so caring and always being there to listen to me, as well as letting me live in your house (so much better than the tunnels across from the university!). A lot of things would not have been possible without you. Thank you to Jennifer, for being the evidence that life is not static and that it is possible to make new friends in the most unexpected ways and places, as well as for getting me into running, which I never thought I would enjoy. Finally, thank you to David, mostly for being an all-round mess in exactly the right way.

Tot slot, wil ik graag mijn familie noemen. Ondanks de grote afstand ben ik ontzettend dankbaar dat jullie me al deze tijd gesteund hebben en altijd vertrouwen in me hebben gehad. De afgelopen jaren zijn niet altijd even makkelijk geweest, met toch wat behoorlijke dieptepunten. Ondanks alles zijn jullie er altijd voor me geweest en ik houd ontzettend veel van jullie.

Abstract

Synthetic peptides are widely used in drug research due to their high selectivity and efficaciousness, as well as in chemical biology as tools for modulating protein-protein interactions. While there are a large number of naturally occurring peptides, which can be used for the development of new pharmaceuticals, there is also an increasing interest in rationally designed synthesised and optimised peptides. In the field of chemical biology peptides are used to understand complex biological processes, where gaining knowledge and understanding can lead to improved human health and the development of novel drug targeting strategies.

Since the synthesis of the first peptide back in 1882, there have been major changes to the peptide synthesis protocol, leading to robust Fmoc-solid phase peptide synthesis that is used in most laboratories today. However, the peptide purification protocols have been lacking behind, meaning peptide purification is still a major bottleneck in fast, efficient, environmentally friendly peptide synthesis. In this thesis, a new peptide synthesis and purification method has been developed and optimised making use of an acrylamide-scavenging approach.

Different strategies to introduce an N-terminal acrylamide cap have been explored, including the use of acryloyl chloride and acrylic anhydride, where the latter was suited for automated SPPS. It was found that replacement of piperidine by DABCO for Fmoc-deprotections was equally efficient and suitable for automated SPPS. The scavenging of acrylamide-tagged deletions was performed making use of a thiol-resin, of which the rate was determined by LCMS in peptide model systems. Upon optimisation of acrylamide-tagging and Fmoc-deprotection conditions, the new purification method was compatible with both manual and automated SPPS, requiring only minor adjustments from standard synthesis protocols. This allowed for the synthesis and purification of the diabetes peptide exenatide (39 amino acids), without making use of preparative-HPLC. An attempt was made to extend this methodology to the synthesis of cyclic thioether peptides; however, this did not provide conclusive results.

An alternative approach to cyclising peptides was explored, making use of disulfide crosslinking with hexafluorobenzene in biologically active p53 peptides. Residues of the biologically active pDI peptide that had been shown to be tolerant to substitution were replaced with cysteine (analogues), including D-cysteine, homocysteine and penicillamine. It was found that replacement of L-cysteine by any of the analogues resulted in altered binding affinity and selectivity.

Finally, the use of fluorine as a reporter for ^{19}F NMR analysis of biological molecules was investigated. Fluorinated prolines were used to explore protein-protein interactions and dynamics, specifically to study the peptidyl-prolyl isomerase cyclophilin D (CypD). First, short substrates were synthesised, including a section of Bcl2 (8 amino acids) and F_0F_1 ATPase subunit B (9 amino acids), which formed a proof of principle. Significant line broadening was observed in the ^{19}F NMR spectrum with increased protein concentration, indicating substrate binding. These results were more apparent for the Subunit B substrate, indicating it is a better binder to CypD than the Bcl2 peptide. The synthesis of another CypD substrate, a part of the poly proline motif in p53, was optimised for the incorporation of 4*S*-fluoroprolines. A very significant finding was that each proline residue had a distinct chemical shift. The *native cis-trans* ratios of each peptide was determined, after which some changes in the *cis-trans* ratios were observed in the presence of CypD. For some of the fluorinated proline residues line broadening was observed upon incubation with CypD, indicating the occurrence of protein-protein interactions.

Abbreviations

Ac	Acetyl
Acr	Acrylamide
AcrCl	Acryloyl chloride
AUC	Area under the curve
Boc	*Tert*-butyloxycarbonyl
Cbz	Carbobenzoxy
CsA	Cyclosporin A
CuAAC	Copper-(I)-catalysed alkyne-azide cycloaddition
CypD	Cyclophilin D
DABCO	1,4-Diazabicyclo[2.2.2]octane
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DCM	Dichloromethane
DIC	N,N′-Diisopropylcarbodiimide
DIPEA	N,N-Diisopropylethylamine
DMF	Dimethyl formamidine
DMPA	2,2-dimethoxy-2-phenylacetophenone
EDG	Electron-donating group
E-Factor	Environmental Impact Factor
Et_3N	Triethylamine
EWG	Electron-withdrawing group
FA	Formic Acid
FKBP	FK506 binding protein
Fmoc	9-fluorenylmethyloxycarbonyl
GSH	Glutathione
HF	Hydrogen fluoride
HFB	Hexafluorobenzene
HPLC	High-performance liquid chromatography
IEC	Ion-exchange chromatography
IM	Intra muscular
IV	Intravenous
K_d	Dissociation constant
LL	Low loading
*m*CPBA	*m*-Chloroperoxybenzoic acid
MDM2	Murine double minute 2
MDMX	Murine double minute X
Mmt	Monomethoxytrityl
mPTP	Mitochondrial permeability transition pore
MS	Mass spectroscopy
MW	Microwave
NAC	N-acetyl-L-cysteine
NAL	N_α-acetyl-L-lysine
N/D	Not determined
NMR	Nuclear magnetic resonance
NPC	Normal phase chromatography
PEG	Polyethylene glycol

PPI	Protein-protein interaction
PPIase	Peptidyl-prolyl isomerase
prep-HPLC	Preparative HPLC
PS	polystyrene
RPC	Reversed-phase chromatography
RT	Room temperature
RT	Retention time
SC	Subcutaneous
SEC	Size-exclusion chromatography
SPE	Solid phase extraction
SPPS	Solid phase peptide synthesis
SPR	Surface Plasmon Resonance
TFA	Trifluoroacetic acid
TIC	Total ion count
TIPS	Triisopropylsilane
Trt	Trityl
UV	Ultra violet

Amino acids

Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V
Homocysteine	hCys	-
Penicillamine	Pen	-

Contents

1 General introduction

In this chapter a general overview of the field of peptide research will be provided, creating context for the research aims of this thesis. The research objectives are discussed and an overview of the contents of the thesis is provided.

1.1 Peptides

Peptides are compounds consisting of two or more amino acids linked in a chain, while the carboxyl group of each acid is joined to the amino group of the next by an amide bond (-OC-NH-). The length of peptides can vary, from dipeptides consisting of only two amino acids, to polypeptides formed of up to around 50 residues. Amino acid chains larger than this are typically defined as proteins (Figure 1.1).

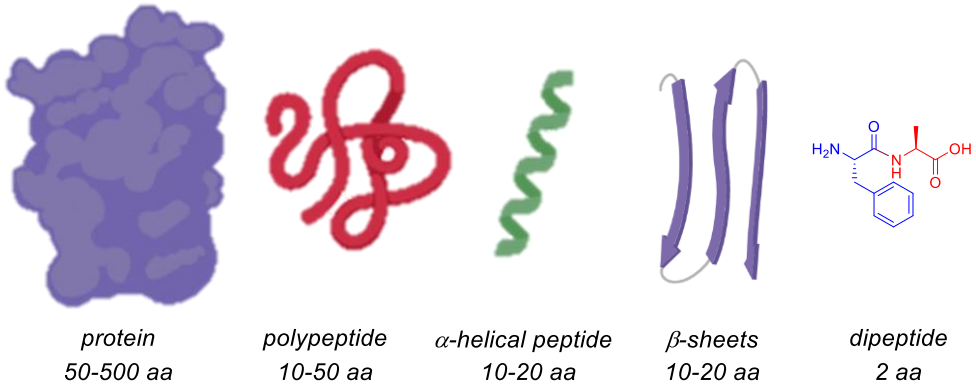


Figure 1.1: Overview of scale differences between proteins and peptides.

There is a wide range of applications for synthetic peptides: from drug discovery to material design.[1] One of the largest advantages of using peptides over proteins is the possibility to synthesise them in a highly controlled way within a laboratory. This allows for easily applicable chemical variation *e.g.* substitution of a single amino acid. Additionally, peptides can be designed and synthesised based on proteins occurring in nature (*e.g.* tumour suppressor protein p53), peptides occurring in nature (*e.g.* conotoxins), or can be a completely new structure or sequence; and can even include unnatural amino acids to perform non-native functions. These possibilities contribute to the growing interest in peptide research, including in the fields of drug research and chemical biology.

1.1.1 Peptides as therapeutic agents

The current pharmaceutical market is dominated by biologicals and small, synthetic molecules. In 2016, 7 out of 10 of the top 10 highest selling drugs were protein-based e.g. monoclonal antibodies.[2] Despite being effective in treatment of disease, there are clear downsides to using either of those drug classes for human therapy. Drugs that are considered to be biologicals are often large, expensive proteins that need to be injected (IV, SC or IM) into patients due to their particular pharmacokinetic properties. In addition, their complexity makes these macromolecules sensitive to degradation (chemical hydrolysis, denaturation and aggregation) when stored under inappropriate conditions.[3] However, due to the biological structure these drugs are often highly potent and target-specific as they mimic natural occurring biomolecules that may be aberrantly expressed or malfunctional. Small molecules on the other hand are sometimes cheaper to produce and can be readily synthesised in high scale quantities. The problem occurring with small molecules, is that they are often less specific (e.g. protein kinase inhibitors) and more likely to induce adverse effects when administered to humans. Peptides are short biomolecules that can chemically be synthesised. They are built from (mostly) natural occurring amino acids, and often resemble the active part of a protein (or biological drug). By developing peptide drugs an attempt is made to retain the beneficial properties of a biomolecule (high selectivity and high potency) yet decrease the production costs and improve the bioavailability.

As peptides are recognised for being highly selective and efficacious there is an increased interest in peptides in pharmaceutical development. Currently, over 60 peptide drugs have been approved for the United States, Europe and Japan. Another 150 are in active clinical development, and an additional 260 potential peptide drugs have been tested in human clinical trials.[4] Some peptide pharmaceuticals even reached ‘blockbuster’ status, gaining hundreds of millions in sales (Figure 1.2). A large number of naturally occurring peptides (over 7000) have been identified, which can form starting points for the development of new pharmaceuticals.[5] Naturally occurring peptides are involved in a range of crucial roles in human physiology,

including as hormones, neurotransmitters, growth factors, ion channel ligands, or anti-infectives.[6–9] With an increasing bacterial resistance for the classical antibiotics (*e.g.* penicillins), there has also been an increasing interest in the development of anti-microbial peptides.[10] As of 2015, there have been more than 60 FDA (Food and Drug Administration) approved peptide medicines on the market, which is expected to grow, with approximately 140 peptides in clinical trials and more than 500 peptides in preclinical development.[5,11]

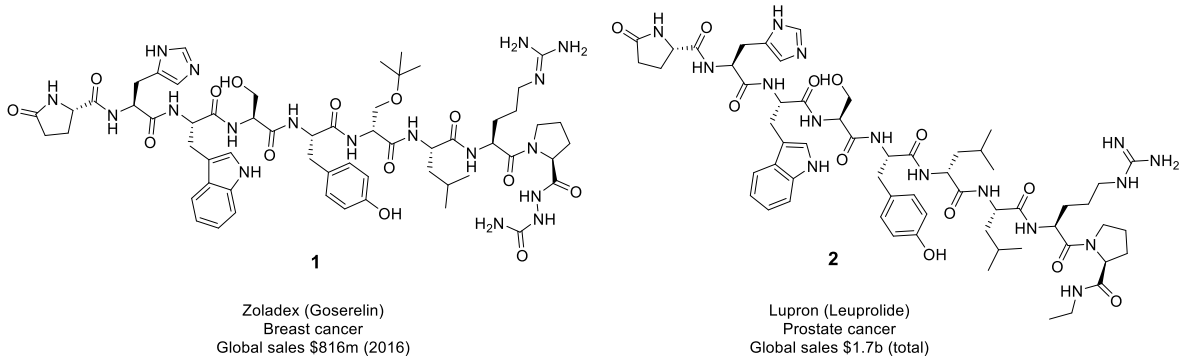


Figure 1.2: Examples of 'blockbuster' peptide drugs.

1.1.2 Peptides as tools in chemical biology

While drug development is a large part of peptide research, this is not the only area of interest. As peptides mimic naturally occurring biological structures, they become especially interesting for the targeting and study of biological processes e.g. protein-protein interactions (PPIs) occurring in health and disease. PPIs and the formation of protein complexes are important biological processes that need to function perfectly for health. In the case of disruptions of PPIs, peptides can be developed for therapeutic use e.g. designing vaccines by epitope-mapping, however, research interest is not limited to this. Proteins are large, complex molecules, and their interactions can often be difficult to understand. The overall 'shape' of a protein is determined by several factors. The first is the primary structure, which represents the order of the amino acid sequence. The secondary protein structure represents the initial folding of the protein into α-helices or β-sheets. The tertiary protein structure describes the three-dimensional shape of a single protein. The formation of protein complexes through

protein-protein interactions are a part of the quaternary protein structure. By using peptides as tools to study PPIs, complicated process can be simplified and analysed in more detail.[12] For example, instead of using two entire proteins to determine or study the interface of interaction, a peptide of the relevant peptide sequence can be synthesised. Obtaining more information on these interactions can improve human health and form leads for the development of novel drug targeting strategies.

1.2 Objectives and overview of this thesis

This research comprises two main areas of interest:

1) Optimisation of the peptide purification process
2) Application of peptides as tools for understanding protein-protein interactions

Chapter 2 explores the origins of peptide synthesis and the improvements that have been introduced over the last decades. It focuses on problems that are still ever present in current day peptide synthesis and attempts to introduce a new protocol to improved synthesis procedures. In Chapter 3 further developments are made, applying the results from model reactions to automated solid phase peptide synthesis procedures. Attempts are made to synthesise and purify large peptide sequences that are biologically relevant. Chapter 4 explores peptide cyclisation, utilising knowledge that has been obtained in the previous two chapters.

The focus of the thesis changes from synthetic development towards biological application from Chapter 5. Here, the importance of amino acid residue stereochemistry upon target engagement in stapled peptides is investigated. Chapter 6 focuses on utilising fluorination to understand protein-protein interactions and peptidylprolyl *cis-trans* isomerisation. An overall conclusion and future perspectives are presented in Chapter 7. All experimental procedures are described in Chapter 8.

1.3 References

1. De La Rica, R. & Matsui, H. Applications of peptide and protein-based materials in bionanotechnology. *Chem. Soc. Rev.* **39**, 3499–3509 (2010).

2. PharmaCompass. Top Drugs by Sales Revenue in 2015: Who Sold The Biggest Blockbuster Drugs? (2016). Available at: http://www.pharmacompass.com/radio-compass-blog/top-drugs-by-sales-revenue-in-2015-who-sold-the-biggest-blockbuster-drugs. (Accessed: 3rd February 2017)

3. Morrow, T. & Felcone, L. H. Defining the difference: What Makes Biologics Unique. *Biotechnol. Healthc.* **1**, 24–9 (2004).

4. Lau, J. L. & Dunn, M. K. Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorg. Med. Chem.* **26**, 2700–2707 (2018).

5. Fosgerau, K. & Hoffmann, T. Peptide therapeutics: current status and future directions. *Drug Discov. Today* **20**, 122–128 (2015).

6. Rotwein, P. Peptide Growth Factors. in *Endocrinology* 79–99 (Humana Press, 1997). doi:10.1007/978-1-59259-641-6_6

7. Li, K. W. Peptide Neurotransmitters and Hormones. in *eLS* (John Wiley & Sons, Ltd, 2012). doi:10.1002/9780470015902.a0000063.pub2

8. Mirabeau, O. *et al.* Identification of novel peptide hormones in the human proteome by hidden Markov model screening. *Genome Res.* **17**, 320–7 (2007).

9. Robinson, S. D. *et al.* Diversity of Conotoxin Gene Superfamilies in the Venomous Snail, Conus victoriae. *PLoS One* **9**, e87648 (2014).

10. Gordon, Y. J., Romanowski, E. G. & McDermott, A. M. A Review of Antimicrobial Peptides and Their Therapeutic Potential as Anti-Infective Drugs. *Curr. Eye Res.* **30**, 505–515 (2005).

11. Kaspar, A. A. & Reichert, J. M. Future directions for peptide therapeutics development. *Drug Discov. Today* **18**, (2013).

12. Katz, C. *et al.* Studying protein-protein interactions using peptide arrays. *Chemical Society Reviews* **40**, 2131–2145 (2011).

2 A novel purification strategy for the synthesis of difficult peptides

The ability to chemically synthesise peptides in a lab has been an invaluable development within the field of medicinal chemistry and chemical biology. This chapter explores the history of peptide synthesis, emphasising improvements that have been made. While the quality of the synthesis has improved, purification of synthetic peptides remains a bottleneck. Especially for 'difficult' *e.g.* long peptides, containing *e.g.* hard-to-couple amino acid residues, current synthesis and purification protocols are often insufficient to achieve a high-quality product. To date, little attention has been given to the optimisation of the peptide purification process and the introduction of new purification methods, which is where this chapter focuses.

Introduction

2.1 Timeline of peptide synthesis

2.1.1 Early developments

In 1882, T. Curtius synthesised and characterised benzoyl glycylglycine, which was the first N-protected dipeptide, obtained after reacting the silver salt of glycine with benzoyl chloride.[1] The first native peptide to have been chemically synthesised came 20 years later, in the year 1901, when E. Fischer and E. Fourneau published the hydrolysis of the diketopiperazine of glycine, resulting in the dipeptide glycylglycine.[2] Major limitations during the early years of peptide synthesis included the lack of availability of enantiomerically pure L-amino acids, and the absence of suitable orthogonal protection groups. The development of the carbobenzoxy (Cbz) protecting group by M. Bergmann and L. Zervas marked a new era for synthetic peptide production.[3] Using this new protection group, numerous small peptides were synthesised, such as glutathione,[4] carnosine,[5] and oxytocin.[6] Synthesis of more complex peptides became possible in 1957, when the *tert*-butyloxycarbonyl group (Boc) was introduced, which allowed for orthogonal deprotection from Cbz.[7] Up to date, the Boc protecting group is still commonly used in peptide synthesis. Alongside the developments in protection techniques, were the advancements in coupling conditions. The carbodiimide activation method was introduced in 1955, which improved coupling rates, however, high propensities of racemisation were the main drawback.[8] By introducing an activator base into the reaction mixture, the racemisation issues were largely overcome.[9] To date, a large range of activators and activator bases are available, which form an invaluable toolkit for the modern-day peptide chemist.

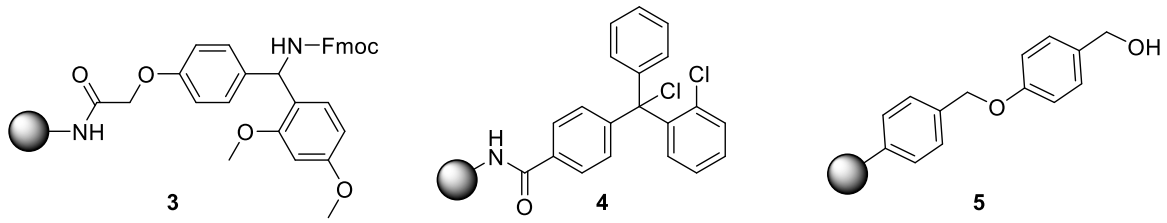
2.1.2 Solid phase peptide synthesis (SPPS)

While early stage peptide synthesis was performed in solution, introduction of solid phase peptide synthesis (SPPS) in 1963 by R.B. Merrifield revolutionised the field.[10] SPPS was developed in an attempt to overcome solubility and purification problems associated with the synthesis of longer peptides in solution. By synthesising the peptide on an insoluble solid particle, excess reagents and by-products could be washed away, without affecting the

peptide itself, which greatly improved peptide purity. Merrifield demonstrated that he could synthesise a four amino acid peptide on a partially chloromethylated polystyrene resin, using repetitive coupling and deprotection cycles. The convenience of SPPS quickly led to the development of new resins and protecting groups. Most noteworthy, is the introduction of the 9-fluorenylmethyloxycabonyl (Fmoc) protection strategy in 1970 by L.A. Carpino and G.Y. Han.[11] In contrary to Boc-synthesis, Fmoc-synthesis utilises lability to base, which allowed for acid-labile protecting groups on amino acid side chains. This also eliminated the requirement of dangerous and aggressive hydrogen fluoride (HF), which was required for resin cleavage in Boc-SPPS.[12,13]

2.1.3 Introduction of new and improved resins

The properties of the resins used is of importance for the overall quality of the peptide synthesis. A lot of focus has been put on the identity of the resin linker, which determine the functionalities at the C-terminus, most often -amide, -alcohol, or -thioester. However, for the quality (e.g. yield, purity) of the peptide synthesis (rather than the chemical properties of the peptide) the physical properties of the resin are more important. This includes resin swelling, resin loading, solvent compatibility, and peptide diffusion.[14]



Rink Amide ProTide Resin
C-terminus: Amide

Cl-TCP(Cl) ProTide Resin
C-terminus: Acid

Wang resin
C-terminus: Acid

Figure 2.1: Commonly used acid-labile SPPS resin linkers.

2.1.3.1 Resin polymer

There are three classes of resin polymers that are the most common; polystyrene (PS), PS-functionalised polyethylene glycol (PEG), and pure cross-linked PEG resins. PS resins are commonly used for the synthesis of short- to medium-length peptides, while PEG-based resins

have been shown to be more successful for the synthesis of medium-length and long peptides, as well as 'difficult' sequences.[15]

2.1.3.2 Resin swelling

The swelling of the resin refers to the expansion of size of the polymer beads, which is caused by absorption of the solvent within the porous structure. The resin swellability in dimethylformamide (DMF), diethyl ether, and dichloromethane (DCM) is the most relevant for peptide synthesis. For a resin to be optimal for the use in peptide synthesis, the resin polymer would swell considerably in DMF, which is used during amino acid coupling steps. A swollen state allows for a better availability of reaction sites.[16] Additionally, it is desirable for a resin to shrink in a different solvent (using e.g. diethyl ether), which permits all reagents to be flushed out after completion of a reaction step. The particle size of the resin beads is commonly expressed in mesh, which indicates the number of holes per inch. A higher number of mesh (more holes) refers to a smaller particle size. A correlation between mesh size and swellability properties have been demonstrated, albeit in a different polymer system.[17]

2.1.3.3 Resin loading

Resin loading indicates the number of binding sites available on the resin, and is typically expressed in mmol/g. There is a wide range of different resin loading available, some as low as 0.15 mmol/g (Cl-MPA ProTide Resin (LL), CEM Corporation), while others can reach up to a 2.0 mmol/g loading (Chlorotriphenylmethane Resin, Sigma Aldrich). It is important to select the correct resin loading, based on the peptide length and propensity to aggregate. In general, the lower the resin loading, the more complicated the synthesised peptide can be, as there will be less steric hindrance from nearby peptide chains. However, the costs of synthesis tend to be larger, as a greater amount of resin is needed to yield the same amount of peptide.

2.1.3.4 Peptide diffusion

While the wide range of linker molecules used in resins determine the C-terminus of the resin, it is also important for the peptide diffusion (release from the resin). Resin linkers are required

to be stable under peptide synthesis conditions, yet rapidly and fully cleavable upon completion of the synthesis. The first resin, designed by Merrifield, could be liberated from the polymer by saponification, or, in the case of the brominated resin, by HBr treatment.[10] Modern day SPPS linkers are still largely cleavable under relatively harsh reaction conditions (albeit milder than HF), with most cleavages taking place in a mixture of TFA/H_2O/TIPS (*e.g.* 95 : 2.5 : 2.5).

2.1.4 Microwave-assisted automated peptide synthesis

The repetitive nature of peptide synthesis forms the perfect setting for automation, which was already noted by Merrifield in 1963.[10] This quickly led to the development of the first automated peptide synthesiser in the mid-sixties.[18,19] Over the years improvements to both hardware and software have led to availability of at least a dozen different commercial solid phase peptide synthesisers.[20]

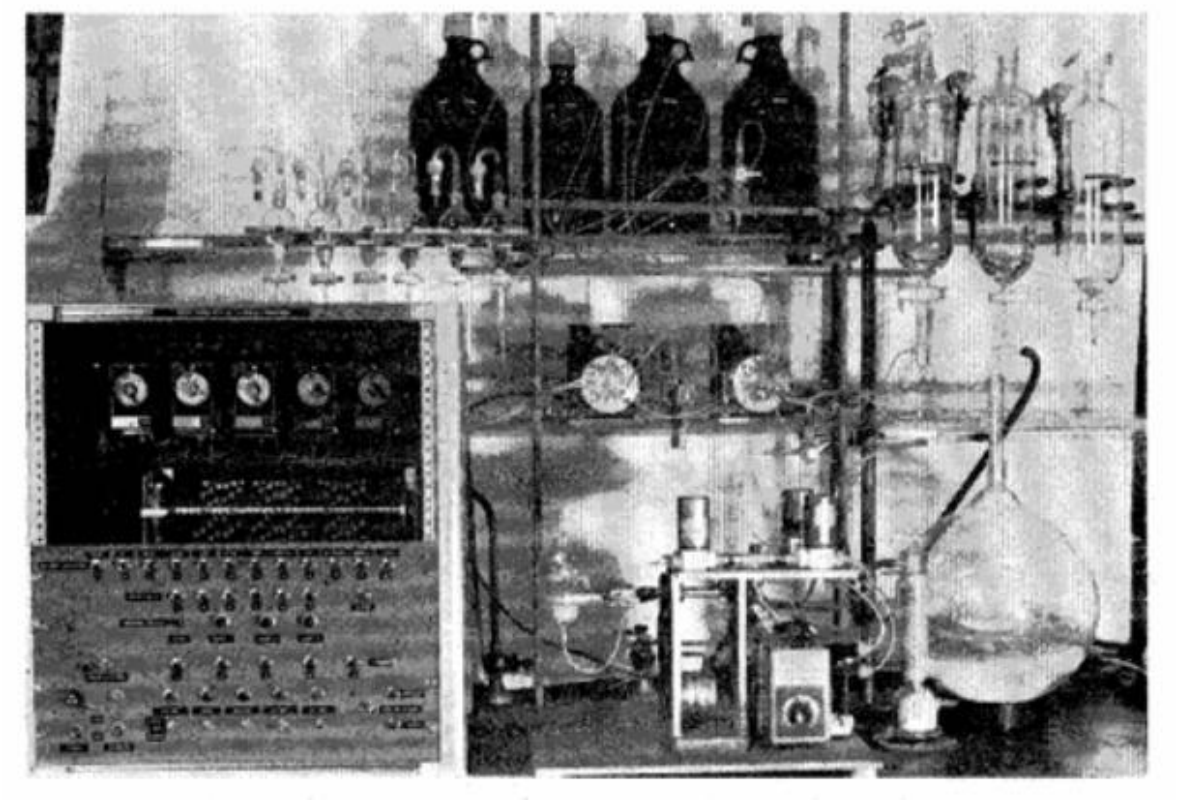


Figure 2.2: First automated peptide synthesis apparatus. Reproduced with permission from Merrifield et al., 1966. Copyright © 1966, American Chemical Society

While there was a dogma comprising the need for peptide synthesis to be performed at room temperature, Erde'lvi and Gogoll worked on a procedure using microwave radiation method. Their method employed rapid (1.5 – 2 min) microwave cycles, which heated up the reaction temperatures up to 110°C.[21] Though reaction rates were faster, the rate limited step involved transfer between fritted vessels (to remove excess solvent and reagents) and microwave resistant vials. This changed in 2005, when CEM introduced the first fully automated microwave-assisted peptide synthesiser, the Liberty™.[22] In present times there are two main automated microwave synthesisers systems on the market, produced by CEM and Biotage. The Liberty™ and successors make use of a valve-based system where amino acids and activators are added through a system of tubing and valves. In 2010, the first robotic-based peptide synthesiser was introduced by Biotage (*Syro* Wave™), which makes use of the movement of a robotic arm moving between different reagent positions and the reaction vessel. The introduction of microwave-assisted automated peptide synthesisers has greatly improved the field, by speeding up reaction times and improving synthesis quality. Advancements in this field are still being made, in attempts to improve synthesis efficiency and purity, and limit side reactions as racemisation and aspartimide formation.

2.1.5 Current limitations

While great improvements have been made with the introduction of solid phase peptide synthesis, new protection groups and diverse resins, compared to the first synthetic peptide back in 1882, there are still limitations in current days rapid automatic peptide synthesis. Despite all developments, each coupling step does not always occur to 100% completion. This leads to a decrease in quality of the peptide product as the synthesis continues, as illustrated in Figure 2.3. Incomplete coupling steps can result in truncated side product, or peptides containing one or more amino acid 'deletions' compared to the target peptide. This problem becomes increasingly significant as the target peptides become longer or more complicated.

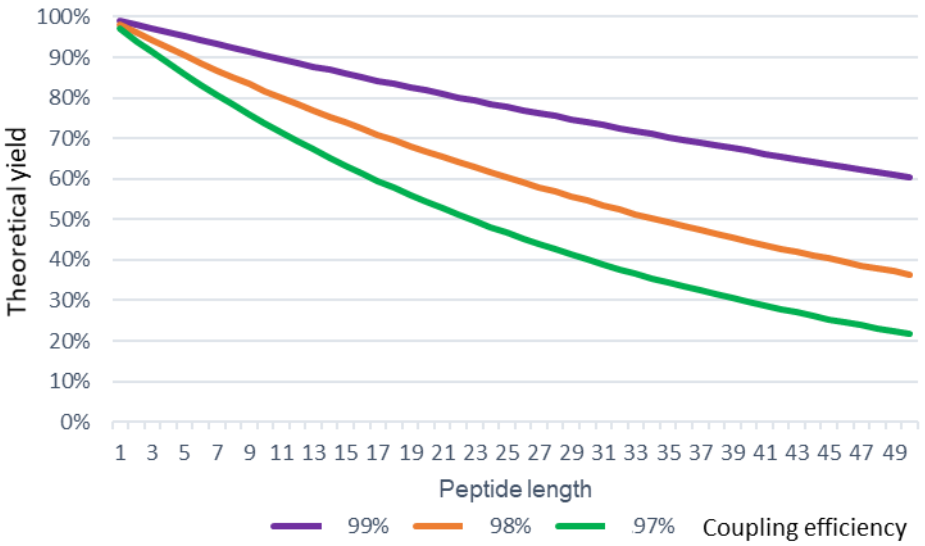


Figure 2.3: Illustration of theoretical decrease in peptide yield upon an increase of peptide length, for a variation of amino acid coupling success rates.

2.2 Improved peptide synthesis

Standard peptide synthesis procedures have undergone great improvement since the first introduction of the field. However, in the case of 'difficult' peptide sequences, where *β*-sheet formation is common and can promote aggregation during peptide synthesis, or multiple hard-to-couple residues (*e.g.* arginine) are introduced in succession, standard solid phase peptide synthesis procedures may not be enough to achieve a high-quality product. One major problem occurring during synthesis are incomplete couplings. Sometimes, despite significant efforts, the next amino acid will not couple at a 100% success rate, which could be caused by e.g. aggregation, low reactivity, or steric hindrance. This causes problems during the purification procedure, as a multitude of amino acid deletions can afford overlapping peaks in high-performance liquid chromatography (HPLC) analysis (see Section 2.3 for HPLC purification). In more detail, a single deletion could shift the retention time either forwards or backwards, depending on the properties of the side chain of the respective amino acid. For example, if a polar amino acid deletion occurs, the retention time is likely to become longer, while hydrophobic amino acid deletions will decrease the retention time. As multiple amino acids can be deleted within one peptide sequence, and multiple truncated peptides can be

formed, the final mixture of peptide and peptide impurities can become very difficult to purify. Valuable strategies to improve the isolated yield of peptides have been developed alongside new resins, activators, and activator bases, of which the most well-known ones are described here.

2.2.1 Double coupling of amino acids

While standard solid phase peptide synthesis employs cycles of single coupling an amino acid, double coupling of trickier amino acids may be considered. Under single coupling conditions, the amino acid, activator and base are reacted with the N-terminus of the peptide that is growing on resin. After the reaction, all excess reagents are flushed away, which is proceeded by the next cycle step, namely the deprotection. However, if for example only 70% of the free amine has reacted with the added reagents, this will lead to 30% of peptides containing this deletion, as the reaction site remains available for the subsequent coupling step. To overcome this problem, it is possible to double couple an amino acid, meaning that an additional coupling step is performed prior to deprotection (Figure 2.4). This pushes the completion of the reaction to a higher percentage, though it may not be effective for all peptides.

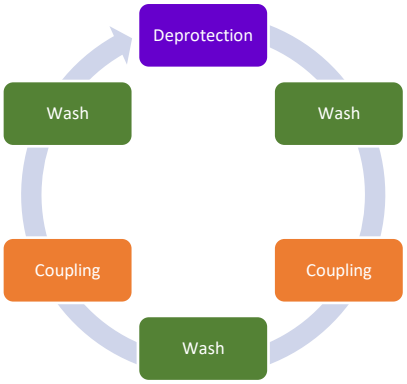


Figure 2.4: Diagram of 'double coupling' synthesis cycle.

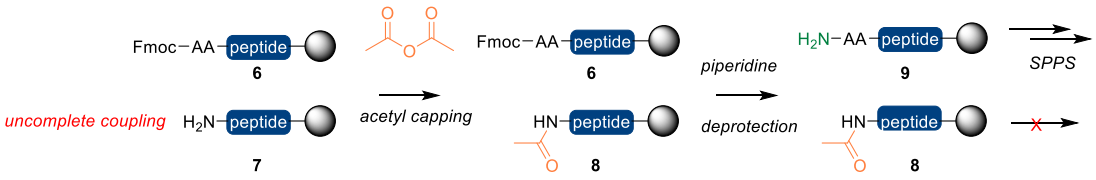
Other factors that need to be considered when utilising double (or more) couplings, are the financial and environmental drawbacks. While it may improve synthesis, double coupling also doubles the costs and the waste material for a single coupling step. If multiple double couplings are required during synthesis, these negative consequences quickly add up, and may not be the most efficient or green approach to improving peptide synthesis quality.

2.2.2 Acetyl capping

Acetyl capping (or “blocking”) can be introduced to prevent the formation of a wide variety of deletions. After each amino acid coupling step, an additional step of acetyl capping is introduced (Figure 2.5). Chemical properties of acetyl capped peptides are often sufficiently different to simplify HPLC procedures.

Figure 2.5: Diagram of ‘capping’ synthesis cycle.

The acetyl cap can be introduced by the addition of acetic anhydride in DMF to the reaction vessel, after which this can react with any free (unreacted) amine present, truncating the peptide and preventing it from further growing (Scheme 2.1). The acetyl cap is base stable and is not affected by piperidine during deprotection steps.



Scheme 2.1: Acetyl ‘caps’ can be introduced using acetic anhydride to prevent truncation of unreacted peptide.

2.2.3 Pseudo-prolines

While the previously two approaches have focused on improving coupling rates and preventing deletions, they are not directly overcoming the cause of the difficult coupling. A problem that often occurs during synthesis of longer peptides is aggregation. In this case,

some peptides (depending on the sequence) start forming a secondary structure (primarily *β*-sheets) while still on resin. This blocks access to the N-terminus of the peptide chains and makes it hard for new reagents to reach the reaction site. Introduction of a proline residue has been known to disrupt the formation of both α-helices and *β*-sheets.[23] T. Haack and M. Mutter developed a technique to use the advantage of the destabilising effects of proline, without having to alter the peptide sequence to contain a proline, by introducing 'pseudo-prolines'.[24]

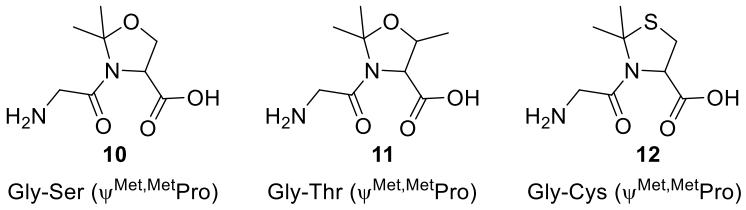
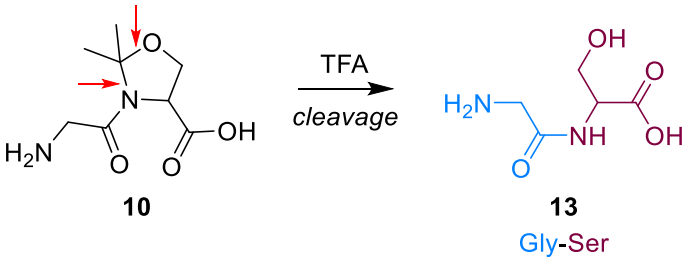


Figure 2.6: Structures of available pseudoproline residues.

Pseudo-prolines are masked versions of the original amino acid and are available as the Ser/Thr-derived oxazolidine or Cys-derived thiazolidine derivatives. They can be incorporated in peptide sequences as dipeptides (alongside the preceding amino acid), and the native sequence is restored upon cleavage by acid hydrolysis (Scheme 2.2).[25] The introduction of pseudo-prolines has been a successful approach for the synthesis of numerous peptides,[26,27] however the application is limited to a small number of amino acids.

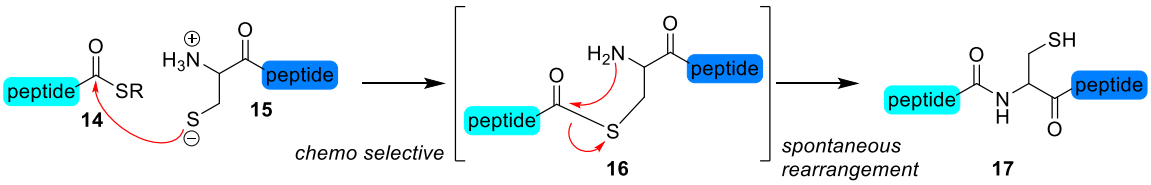


Scheme 2.2: Hydrolysis of pseudoproline to native residues glycine and serine.

2.2.4 Chemical ligation

Even though SPPS is very successful for short peptides, synthesis becomes more troublesome when a certain peptide length is reached, typically around the 30 amino acids mark (this is variable depending on the peptide sequence).[13] These limitations made SPPS less suitable for the synthesis of proteins unless it would be possible to synthesise separate

segments and ligate them together to form the full protein. This approach was first attempted by M. Schnölzer and S.B.H. Kent in 1992. They synthesised two parts of the HIV-1 protein, which were both ~50 amino acids in length, after which the amino-terminal half was reacted with the unprotected carboxyl-terminal half to give the full length protein with a pseudo-peptide bond at the ligation site.[28] The first ligation experiments gave a non-peptide bond at the site of ligation, leading to the development of a new technique which did provide a native peptide bond. 'Native chemical ligation' was introduced by Dawson *et al*. in 1994, where a chemo selective reaction would take place of two unprotected peptide segments (**14** and **15**), which initially gave a thioester-linked peptide (**16**). Spontaneous rearrangement would yield a native peptide bond at the ligation site (**17**) (Scheme 2.3).[29]



Scheme 2.3: Overview of native chemical ligation.

Chemical ligation techniques are dependent on segment terminal amino acids, between which the new peptide bond needs to be formed. Over recent years there has been a lot of development to make this technique applicable to a wider range of amino acids, including phenylalanine,[30] valine,[31] and lysine.[32]

2.3 Standard peptide purification protocols

While the previously described strategies improve the synthesis of peptides, it does not always lead to crude purities that are up to desirable standards. This leads to the need of peptide purification, which remains one of the main bottlenecks to date. Peptides are typically purified by preparative HPLC, requiring multiple injections of often small amounts of crude reaction mixture, which is separated using long and slow mobile phase gradients.

2.3.1 High-performance liquid chromatography (HPLC)

There are currently at least eight different liquid chromatography modes available for peptide analysis and purification: size-exclusion chromatography (SEC), ion-exchange chromatography (IEC), normal phase chromatography (NPC), reversed-phased chromatography (RPC), hydrophilic interaction chromatography, immobilised metal ion affinity chromatography, and biospecific/biomimetic affinity chromatography.[33] HPLC is a popular choice for peptide analysis and purification as excellent resolution can be achieved under a wide range of conditions for both closely related and distinct molecules.[34] It is possible to follow the migration of the peptides through a UV detector, which is looped in the system directly after the column. Additionally, a HPLC system can be linked up to a mass spectrometer, allowing for very accurate characterisation of complicated reaction mixtures. When HPLC is used for purification, the system is normally equipped with a fraction collector, which will trigger collection based on acquired UV and/or MS data. There is a difference between analytical and preparative HPLC scale, where the latter uses larger columns with more packing material. Analytical flow rates are typically between 0.5 – 1.3 mL/min, preparative utilises flow rates from 3 mL/min for semi-prep to 17 mL/min for prep. The injection volume and sample concentration is typically larger for prep HPLC. Translational differences can occur when switching from an analytical system to a preparative system, and some additional method development may be required.

2.3.1.1 Size-exclusion chromatography

As the name suggests, size-exclusion chromatography separates molecules based on their size. The column used for this purification method consists of a porous matrix with various pore sizes. Smaller molecules tend to have a longer retention time, as they are able to infiltrate deeper into the matrix, while larger molecules pass through the column more quickly.[35]

2.3.1.2 Ion-exchange chromatography

Ion-exchange chromatography separates molecules based on difference in their charge and can be a powerful tool for the separation of peptides. There are two types of ion-exchange

columns: anion-exchangers and cation-exchangers. The anion-exchangers have a positively charged stationary phase, and will bind negative ions as a result, while the cation-exchangers have reversed polarity.[36] While ion-exchange chromatography can be useful for separation of peptides and proteins, it does not always form the best approach to purify out a synthetic peptide from a reaction mixture, as a single deletion in a bigger peptide may not influence the overall charge enough to maintain a desirable resolution.

2.3.1.3 Reverse-phase chromatography

While normal phase chromatography makes use of a polar stationary phase, and nonpolar solvents, reversed-phase chromatography, as the name suggests, does the opposite. This is especially suitable for peptide purification, as peptides often dissolve well in polar solvents like water, methanol, and acetonitrile. Despite limitations, prep-RP HPLC has been a popular approach for peptide purification since the 1980s.[37]

2.3.2 Flash chromatography

While flash chromatography is typically used with silica gel in a normal phase solvent system, introduction of new stationary phases has made this technique suitable for peptide purification. Similar to reversed-phase HPLC, flash chromatography uses polarity to separate molecules by their retention time on the stationary phase. Pre-packed reversed-phase cartridges are readily available, making this technique easily accessible. While flash chromatography and HPLC are very similar, the main difference can be observed in the resolution, which can be explained by the use of smaller, more spherical particles in HPLC. This also means flash chromatography has a much lower back-pressure than is common during HPLC. As a result, flash chromatography is more suitable for larger scale, crude, fast purification of relatively clean reaction mixtures. Over recent years, fully automated flash chromatography systems have been developed (Biotage® Selekt, Teledyne CombiFlash NextGen 300+, Sorbtech EZ Flash System and more).

2.4 Improved peptide purification

While HPLC purification remains the standard peptide purification method to date, there have been developments to optimise those procedures. Additionally, alternative purification methods have been described.

2.4.1 Advancement of HPLC protocols

2.4.1.1 Fused-core columns

As briefly discussed in Section 2.3.2, liquid chromatography is highly dependent on the particle properties of the solid state. Smaller particle size increases the surface area, which improves resolution, however, the pressure required for solvent to pass through increases as well. Fused-Core® technology was introduced in 2006, which makes use of silica particles with a solid core and a thin, porous outer shell. This allows for improved purification results with modest back-pressures.[38,39] Since its introduction, fused-core particles have been further developed for the purification of small molecules up to proteins,[40] and specifically designed columns are commercially available (HALO®, Phenomenex®, Ascentis Express).

2.4.1.2 Ion-exchange doped reversed phase (DRP) HPLC

Reversed-phase chromatography has been very successful for the purification of peptides, however, sometimes a reaction mixture is so complex that RPC alone is not effective. New stationary phases are being developed that combine RPC with a different type of chromatography, allowing for orthogonal separation within one system. A Swiss research group introduced a new mixed material column, which consisted predominantly of reversed phase ligands, while ion exchange ligands (for electrostatic interactions) were present in doping quantities. They observed improved separation for a range of different peptides.[41]

2.4.2 Solid phase extractions (SPE)

Often overlooked is the possibility to purify peptides using solid phase extraction. SPE utilises short, pre-packed cartridges containing a modified stationary phase (usually silica based) that can provide separation based on non-polar, polar, ionic or covalent interactions. Advantages

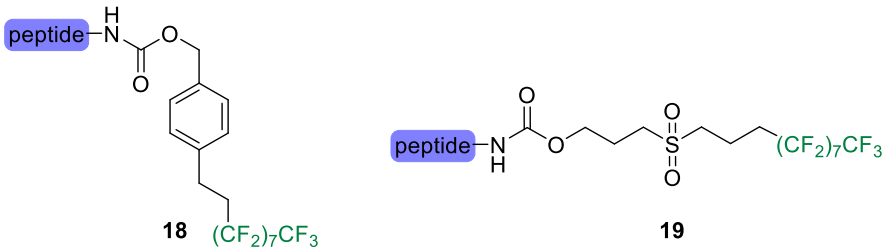
of this method are the speed, reproducibility, and efficiency, however, this method is not successful for complicated reaction mixtures, with closely related impurities.[42]

2.4.3 Catch-and-release purification

A recent development in the field of peptide chemistry, has been the introduction of 'catch-and-release' techniques. The aim of this is simple: upon completion of the synthesis a tag is attached to the desired product. This tag can then interact with a solid support, either covalently or electrostatically, which allows for flushing out of excess reagents and untagged side products. After this purification step the tag is removed, yielding the unaltered desired product. A major selling point of this purification technique is the elimination of the costly, time-consuming HPLC-purification step.

2.4.3.1 Fluorophilic tags

J.H. van Boom's group developed a reversible benzyloxycarbonyl probe derivatised with a fluorophilic tail (FZ-tag), which could be purified by fluorous chromatography. During SPPS, an acetyl capping step (Section 2.2.2) was introduced after each amino acid coupling, and the fluorous tag was introduced onto the final product at the end of the synthesis.[43] The FZ-tag, however, is acid-labile, restricting the application in Fmoc-SPPS. To overcome this problem, the same group continued to developed a fluorous methylsulfonylethoxycabonyl tag, which they named FMsc,[44] which is acid-stabile and compatible with Fmoc-SPPS.



FZ-tag FMsc-tag

Figure 2.7: Structures of fluorous tags (marked in green) attached to a peptide.

2.4.3.2 Belyntic's Peptide Easy Clean Technology

Similarly, to previously described fluorous approach is the peptide easy clean (PEC) technology developed by the German-based company Belyntic. They also follow standard SPPS procedures, with acetyl capping, and introduce a special linker on the desired peptide sequence at the end of the synthesis. However, the novelty of their procedure is introduced when the tagged peptide is immobilised on a specially introduced resin ("catch"). Impurities can then be washed away by filtration, after which the desired peptide is cleaved from the resin ("release") yielding the purified desired peptide.[45] A drawback of this technique is the reliability of the specific coupling of the linker to the desired peptide product. Insufficient coupling could decrease peptide yield.

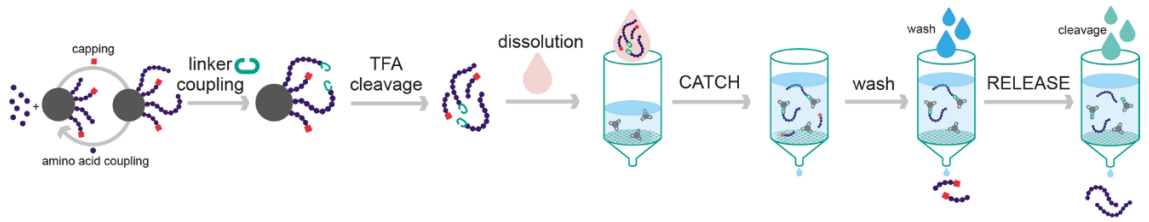


Figure 2.8: Schematic overview of PEC-technology. Reproduced with permission from Belyntic®.

2.4.3.3 Sulfotools' Clean Peptide Technology

While the two previously described methods are dependent on tagging the final product and successful acetyl capping during synthesis, the method developed by Sulfotools takes a different approach. They developed a novel capping reagent, replacing the traditional acetyl cap by a sulphonated cap (either 2-sulfoacetic acid or 4-sulfobenzoic acid). Additionally, they introduced sulphonate groups on the Fmoc-temporary protection group, as well as amino acid side chain protecting groups. This would allow for better solubility in water, avoiding the need for organic solvents during peptide synthesis.[46] A clear drawback of this technique is the lack of competitive commercial availability of the amino acid building blocks. Upon cleavage, all sulphonate groups are removed from the desired peptide, but not from the tagged impurities. The reaction mixture is passed through an anion or cation exchange affinity column, where sulphonated impurities are often well separated from the desired peptide.

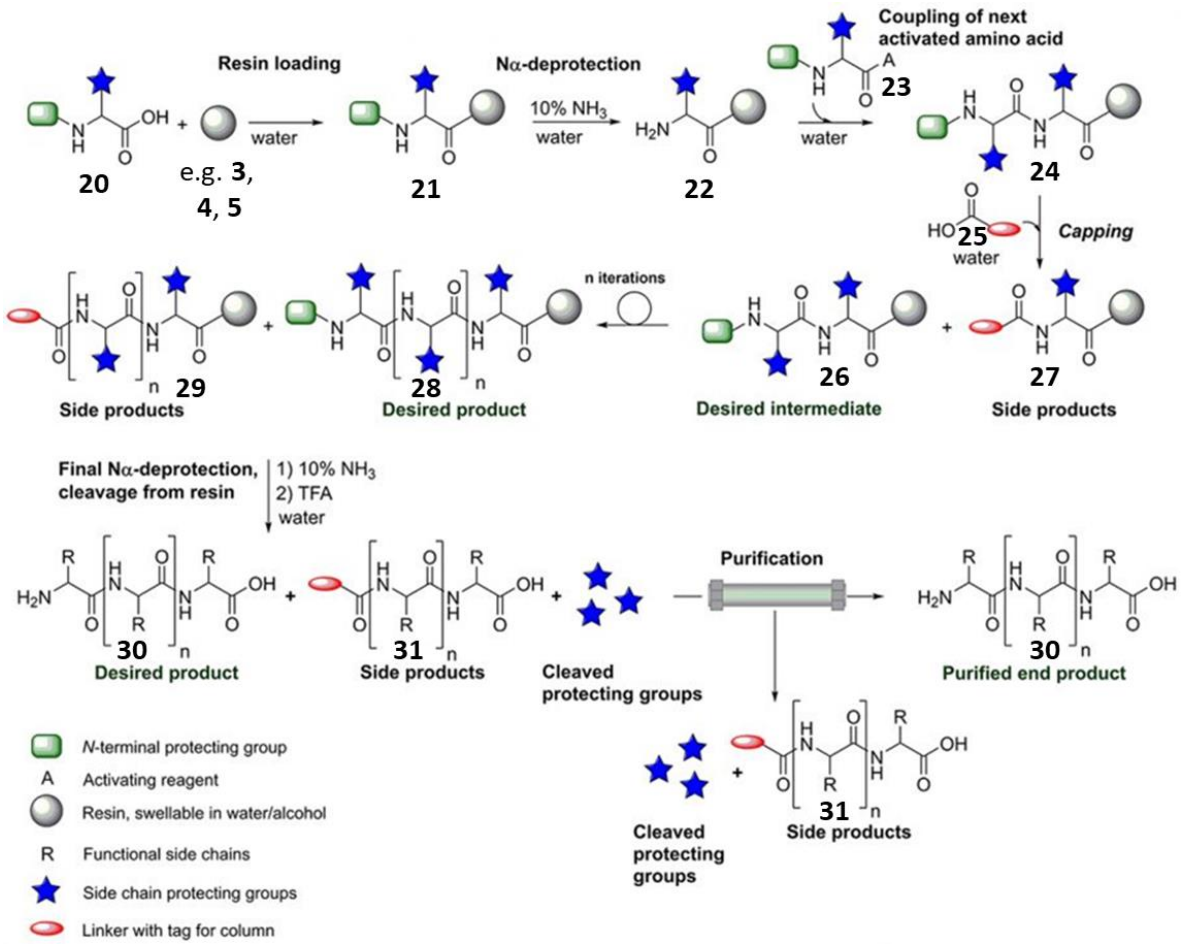


Figure 2.9: Schematic overview of Clean Peptide Technology. Reproduced with permission from Sulfotools®.

2.5 Introduction of a new peptide purification protocol

2.5.1 Aim of this chapter

Since the first introduction of peptide synthesis, and later solid phase peptide synthesis, the technologies have advanced significantly. Despite all efforts, synthesis is often not as straightforward as expected, so time-consuming, costly, ineffective purification procedures are required. In this chapter we have worked on the development and optimisation of a novel, effective, easy to use, and most notable green peptide purification protocol.

2.5.2 Acrylamide-tagging

In analogy with previously explained acetyl capping, we chose to introduce an additional capping stage during our peptide synthesis. This requires minimal adjustments to standard SPPS procedures that are currently in use by many peptide research groups. Rather than acetyl, we chose to incorporate an acrylamide-tag into our peptide synthesis because of its bifunctionally; reducing truncation and providing a covalent scavenging-tag (Figure 2.10). While other scavenging-tags could have been explored (e.g. azides or alkynes for a CuAAC "click" approach) we considered acrylamides the most cost-effective and easiest to incorporate in standard SPPS procedures.

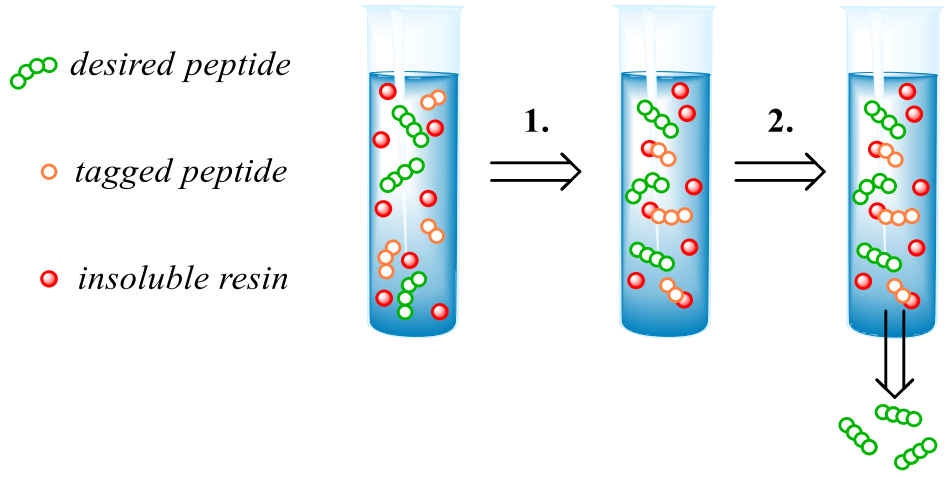


Figure 2.10: Graphic overview of peptide purification process. 1) All acrylamide-tagged impurities will covalently attach to the insoluble thiol-resin displayed in red. 2) The desired peptide, which does not have the acrylamide-tag, will remain in solution, and can be isolated through filtration.

2.5.3 Proposed advantages of our method

While our acrylamide-tagging method has some similarity with previously described catch-and-release techniques, there are important differences. Both the fluorous-tagging technique and peptide easy clean technology tag the desired final product. After purification this tag then needs to be removed in an additional step. It is important to consider that peptides that require advanced purification strategies have been difficult to synthesise. The question that remains is if a chemist would really want to modify a long and difficult peptide that has just been synthesised with great effort. Additionally, for this method to work, acetyl capping needs to be quantitative at each step to prevent any truncations from getting tagged during the final step. When the impurities are tagged instead of the final product, the reversal modification step ("release") is eliminated. Clean peptide technology (Sulfotools) does tag impurities instead of the final product, which can be an advantage, however, this purification method is then dependent on interactions with ion exchange chromatography columns. The advantage of our method is that all impurities are covalently bound to a solid support resin, fully eliminating the dependence on HPLC resolution for purification.

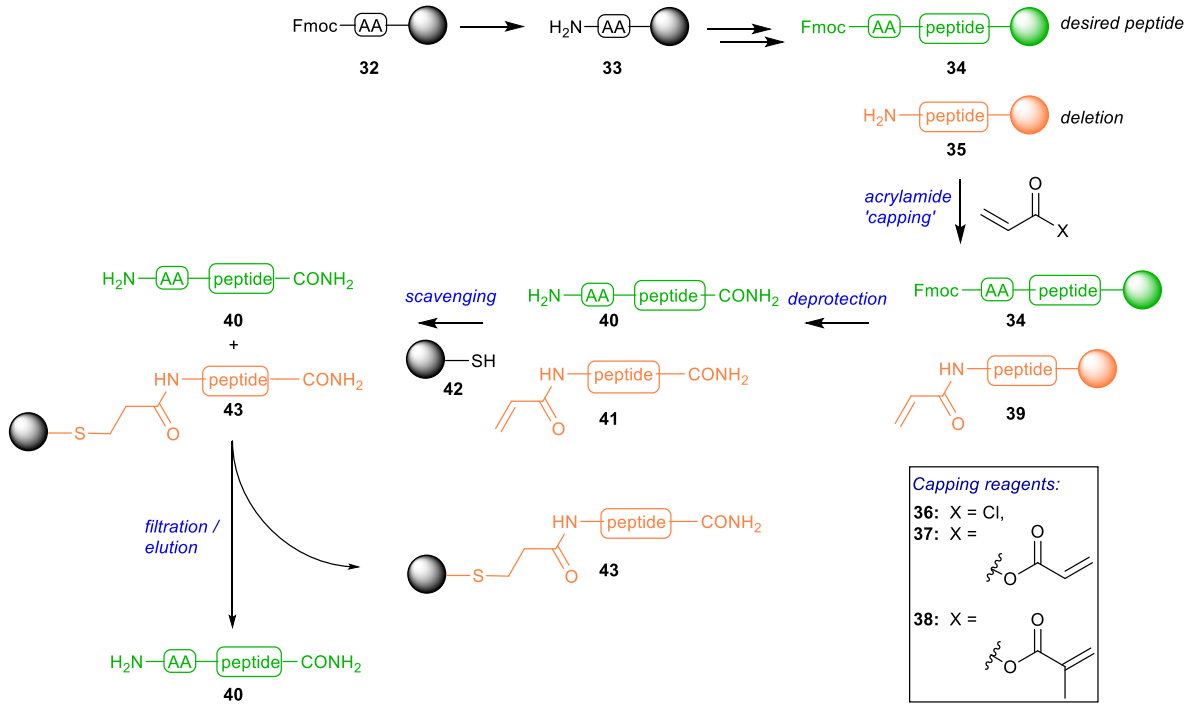
Peptide synthesis is considered to be one of the most wasteful and least green chemical processes.[47] Large amounts of excess reagents are required to yield a product that is purified by HPLC as a standard procedure, which uses large amounts of solvents for a milligram of purified material. Climate change and environmental burden of all industries is a current topic, and it is important to try and reduce the amount of waste generated for the future of our planet.[48] Our purification protocol aims to improve purified peptide yield, while decreasing the amount of waste generated dramatically.

2.5.4 Approach

We hypothesise that our acrylamide-tagging approach can be incorporated into standard SPPS methods without major alterations of the protocol, as illustrated in Scheme 2.4. The addition of an acetyl capping step, using acetic anhydride in DMF, is standard practice in most peptide labs, and formed the starting point for our development studies. As quantitative

tagging of impurities is essential to the effectiveness of this purification method, optimal conditions for the addition of a N-terminal acrylamide-tag were explored.

The purpose of the acrylamide-tag is initiating the possibility to scavenge using a thiol-based resin. Subsequently, the reactivity of peptidyl-acrylamides with model-thiols was studied to gain insight in the potential of our proposed scavenging approach. Furthermore, the selectivity for the thiol functional group was determined. Compatibility with standard SPPS procedures was also considered to be essential, as this will make our new purification protocol widely accessible for everyone working in the peptide synthesis field. With this in mind, the stability of peptidyl-acrylamides was thoroughly investigated under standard synthesis conditions. Finally, an attempt at a model purification was undertaken, to demonstrate the potential of this new purification method.



Scheme 2.4: Overview of proposed solid phase peptide synthesis route with acrylamide capping.

Results and discussion

2.6 Demonstration of observed synthesis difficulties

To demonstrate the difficulty of removing a single deletion from a mixture, two structurally similar peptides were synthesised and combined (Ac-SEPAKY-$CONH_2$ (**44**) and Ac-EPAKY-$CONH_2$ (**45**)). These peptide sequences were chosen to include a range of functional groups. The HPLC-MS spectrum of a 1:1 mixture can be seen in Figure 2.11. Even with one single deletion a small overlap between the two peaks can already be observed, complicating purification by preparative-HPLC or any other retention time-based purification method.

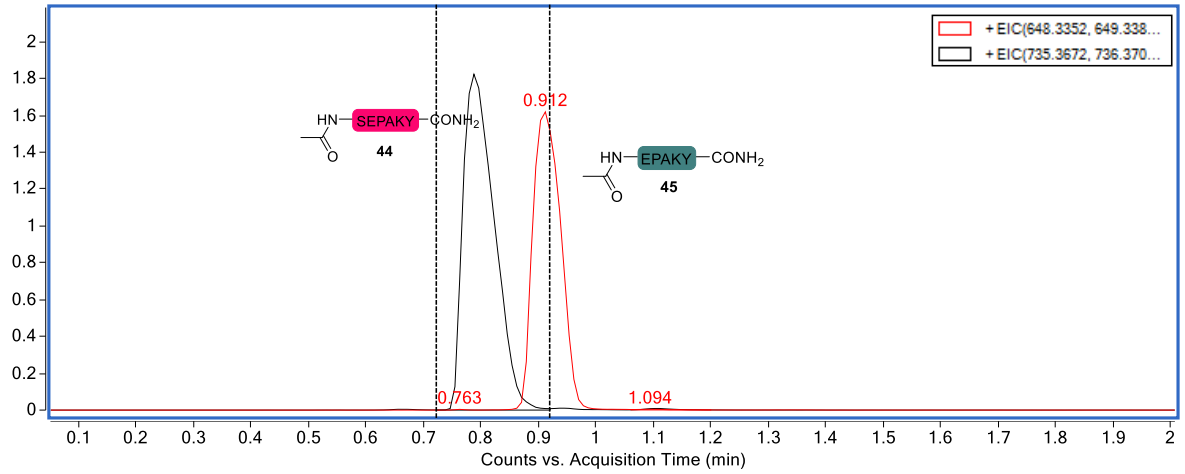


Figure 2.11: HPLC-MS find-by-formula chromatogram of the peptides Ac-SEPAKY-$CONH_2$ (**44**) and Ac-EPAKY-$CONH_2$ (**45**).

When synthesising a longer, more complex peptide, the purification process becomes even more complicated. As previously explained in Section 2.2.2, a single deletion could shift the retention time either forwards or backwards, depending on the properties of the side chain of the respective amino acid. Peptide **46**, consisting of 29 amino acids (NH_2-KPSKIPVLLRKGGRQIKIWFQNRRMKWKK-$CONH_2$), was synthesised to demonstrate the effects of different common coupling strategies. All amino acid solutions, activators and reaction times were kept constant between the different experiments, with the only differences being A) single coupling, or B) double coupling. For the standard single coupling procedure (Figure 2.12.A) there is a clear overlap of peaks visible, where the desired peptide is buried

underneath other closely related peptides. Additionally, most of the crude synthesis products appeared to be impurities.

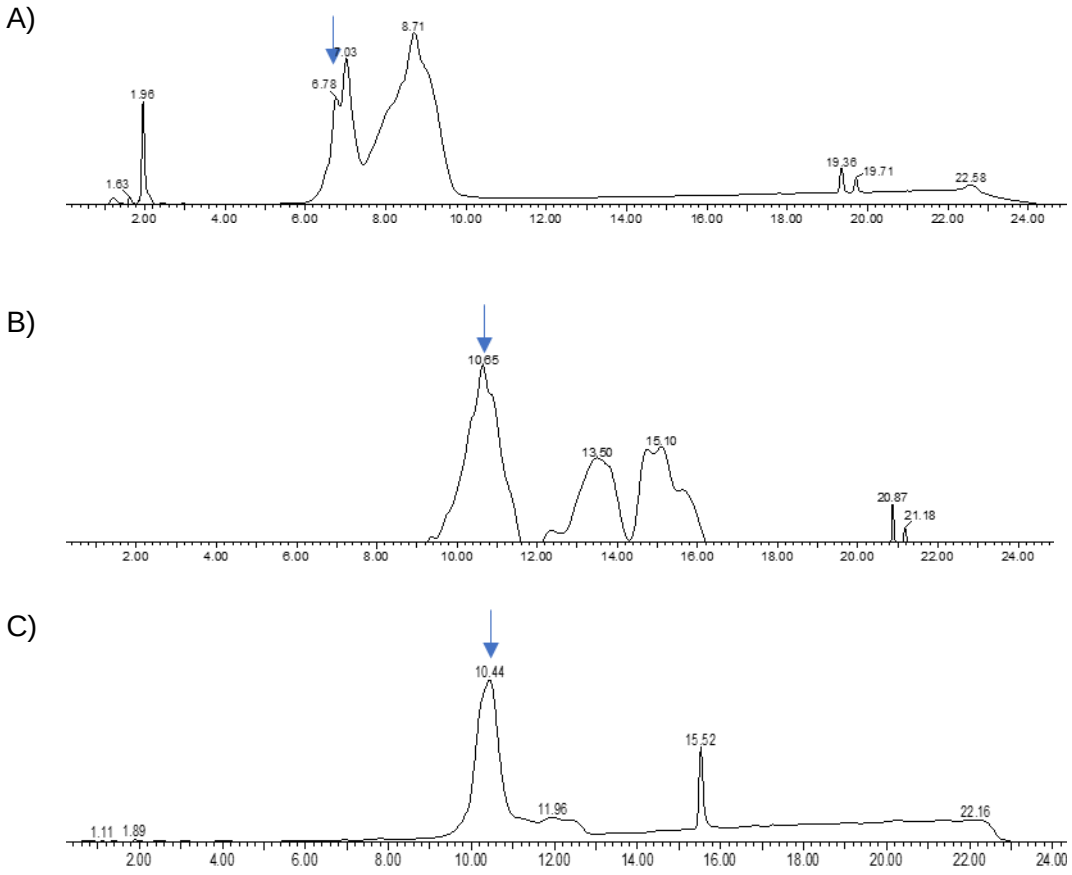


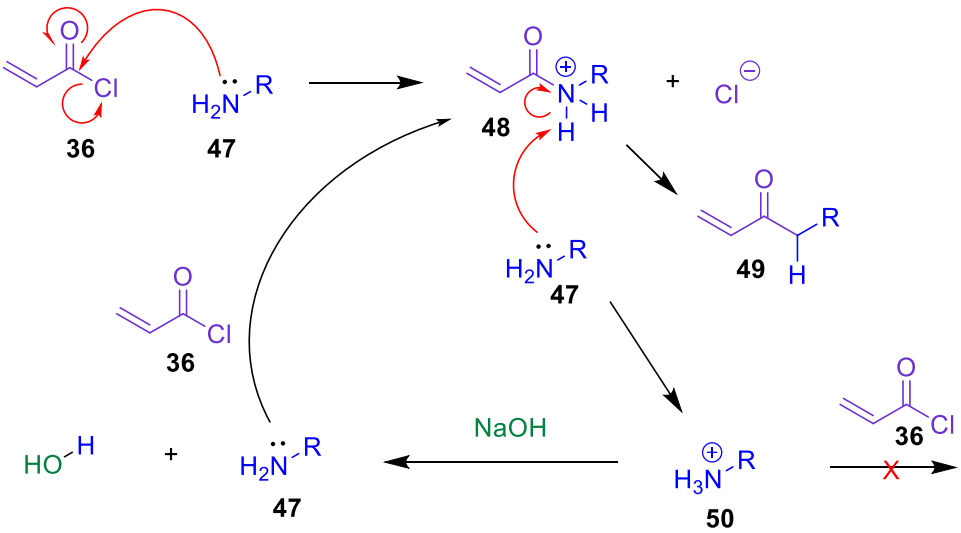
Figure 2.12: HPLC chromatograms of NH_2-KPSKIPVLLRKGGRQIKIWFQNRRMKWKK-$CONH_2$ (**46**) prepared under different conditions: A) single couplings; B) double couplings; C) double coupling after prep-HPLC. The arrow indicates the location of the target peptide. Note: Organic phase in chromatogram A is acetonitrile (0.1% FA) and is methanol (0.1%) in chromatogram B and C.

It is possible to improve peptide quality by double coupling each amino acid, or at least the amino acids that prove to be difficult. For shorter peptides this is often sufficient to improve the yield to an acceptable standard, however, for longer peptides formation of secondary structures (*e.g.* β-sheets) can obstruct successful couplings (Figure 2.12.B). While there was a slight improvement compared to single coupling conditions, double couplings still did not reach desired purity standards or even allow for successful preparative HPLC. As most

impurities have closely related retention times to the peptide of interest, impurities remain present in the peptide sample, even after preparative-HPLC (Figure 2.12.C). These results demonstrate the urgent need for new purification methods, that do not depend on HPLC-purification.

2.7 Introduction of acrylamide tags on the peptide N-terminus

In analogy with previously described catch-and-release purification techniques (Section 2.4.3), we plan to introduce a “catch”-tag at the N-terminus of our impurities, for which we have chosen acrylamides. The reaction between acryloyl chloride and an amine is a common approach to introduce an acrylamide group in small molecules.[49,50] Acrylamides have been introduced in amino acids before,[51–53] however, no literature examples could be found for peptides. The most commonly used reaction conditions in literature examples are mild, occurring at room temperature in water with sodium hydroxide, following a Schotten-Baumann reaction mechanism (Scheme 2.5). While the addition of base is not required at the beginning of the reaction, it is required to keep the amine nucleophilic by keeping the pH of the reaction mixture constant after the release of HCl.



Scheme 2.5: Schotten-Bauman Reaction mechanism of acryloyl chloride with an amine.

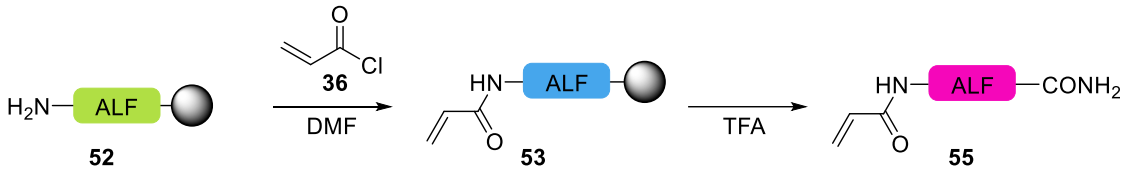
To be able to use the acrylamide-group as an impurity-tag it is important that the reaction conditions are compatible with conventional (automated) solid phase peptide synthesis conditions. The most widely used 'capping' method utilises 20% acetic anhydride in DMF, which reacts rapidly with the free amine that is left after an incomplete coupling. The reaction conditions for the acetic anhydride capping procedure were used as a starting point for the optimisation of acrylamide-capping conditions. DMF was used as a solvent, due to its compatibility with standard solid phase peptide synthesis.

Preliminary experiments were carried out using 20%, 15%, 10%, and 5% stock solutions of acryloyl chloride in DMF. Approximately 25 mg of resin-bound tripeptide (NH_2-AKY-●, **51**, 0.55 mmol/g) was used and was treated for 15 or 60 minutes, with or without DIPEA as base. The results were analysed by HPLC and showed complete conversion for all conditions. To prevent unnecessary exposure to toxic chemicals, as well as limiting chemical waste, further experiments were carried out to determine the minimal amount of acryloyl chloride required. While working with stock solutions it was noticed that acryloyl chloride in DMF starts a transparent slightly pale-yellow solution, which slowly turns dark yellow/brown over a few hours. This prompted us to work with fresh acryloyl chloride solutions for each future experiment, rather than preparing stock solutions. A non-nucleophilic tripeptide (NH_2-ALF-●, **52**) was synthesised manually on ProTide™ rink amide resin (0.55 mmol/g loading) and was used to further investigate a range of coupling conditions (Table 2.1).

Two different reaction temperatures were explored, either at room temperature, suitable for manual peptide synthesis, or at 90°C, compatible with automated microwave-assisted peptide synthesis. The coupling was only performed once, not twice as commonly used for acetyl 'capping' procedures. No base was added to any of these reactions, as the addition did not alter the conversion rates during the preliminary experiments. However, as fewer equivalents of acryloyl chloride were used, addition of base may have had a more profound effect. The percentage of acrylamide formation was determined by LC-MS upon cleavage of the peptide

from the resin. The AUC (area under the curve) of peptide **54** (NH_2-ALF-$CONH_2$) and peptide **55** (Acr-ALF-$CONH_2$) was used to calculate conversion yield.

Table 2.1: Optimisation of acryloyl chloride (AcrCl) reaction on a tripeptide (NH_2-ALF-●) (25 mg; 0.01375 mmol), conversion yields are based on UV absorbance at 215 nm as determined by LC-MS.



AcrCl in DMF concentration (molar)	Reaction time	Temperature	Conversion yield (%)
0.02 M	15 min	20°C	29
0.125 M	15 min	20°C	56
0.25 M	15 min	20°C	62
0.5 M	15 min	20°C	73
1 M	15 min	20°C	74
0.5 M	60 min	20°C	84
0.25 M	1 min	90°C (MW)	98
0.5 M	1 min	90°C (MW)	100

Reactions at room temperature showed a clear increase in conversion when the concentration was increased from 0.02 M (~1.5 eq.) up to 0.5 M (~36 eq.), while an increase from 0.5 M to 1.0 M did not make a difference. Increasing the reaction time from 15 minutes to 1 hour for 0.5 M also increased the conversion yield, as expected. The minimal improvement between 0.5 M and 1.0 M concentrations may have been the result of a built-up of HCl, which could have protonated the unreacted amine, inhibiting completion of the reaction.

The reaction between acryloyl chloride and the free N-terminal amine occurred very rapidly under microwave conditions. However, it must be noted that during the initial set-up of the experiment conditions some ‘rapid pressure build-up’ errors occurred, with the pressure in the reaction vessel exceeding 7.0 bar. When heat was applied more gently, by increasing the ramp time (time it takes for the microwave to reach the target temperature), the pressure

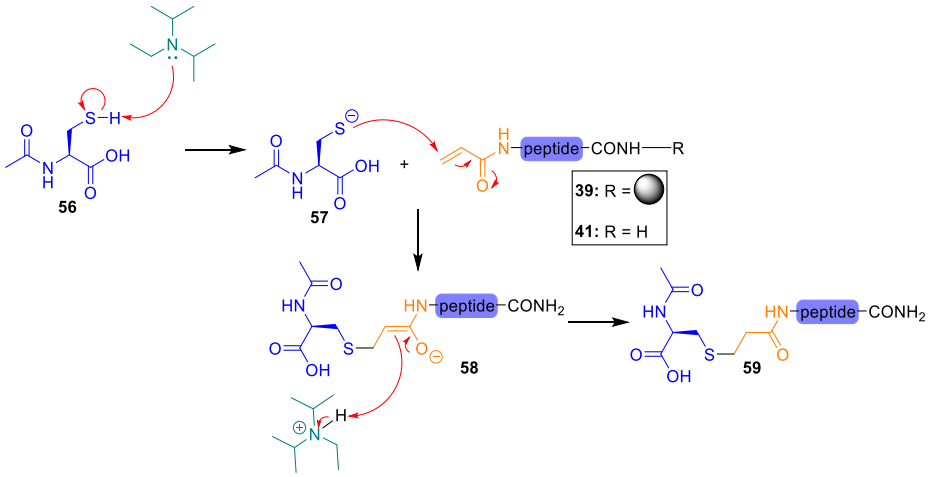
remained more stable reaching a maximum of 2.5 bar at 90°C. The observed pressure increase could be explained by exceeding the boiling point of acryloyl chloride (75°C). Thermal instability of acryloyl chloride could also be a reason for improved conversion yield under microwave conditions, as decomposition of the molecule may make it more electrophilic.

For all future experiments, 0.5 M concentration of acryloyl chloride in the microwave (90°C, 1 min) was used for acrylamide-tagging (Acr-tagging).

2.8 Modelling the reaction and selectivity of peptidyl-acrylamide with a model thiol (N-acetyl-L-cysteine)

It was planned to use a thiol-based resin for the scavenging of acrylamide-tagged impurities, it was important to gain insight in the reactivity between thiols and acrylamide-tagged peptides, as well as the selectivity for thiols over other functional groups present in amino acids. If a reaction could take place selectively between a model thiol and peptidyl-acrylamides, this would give a clear indication whether the resin-based scavenging approach is viable.

N-acetyl-L-cysteine (NAC, **56**) was used as a model thiol (Scheme 2.6) in these experiments. The model Acr-tagged peptides were prepared according to previously optimised conditions. The reaction between NAC and the Acr-peptides could be monitored by HPLC (shift in retention time) and LC-MS (change in molecular weight). To ensure complete reactivity of all Acr-tagged peptide, various amounts, but always an excess, of NAC was used in all experiments. The reaction progress was closely monitored via HPLC and LCMS. For the HPLC experiments the percentage of product and starting material was determined by comparison of the area under the curve (AUC) at 280 nm. This wavelength was chosen because of the high specific absorbance caused the tyrosine residue, while the optimal wavelength for peptide bonds (215 nm) had more background signals.



Scheme 2.6: Reaction mechanism between N-acetyl-L-cysteine and an acrylamide-tagged peptide following a Michael addition.

2.8.1 Reactivity

A three amino acid peptide **60** (Acr-AKY-$CONH_2$) was synthesised to monitor the reaction between an acrylamide-tagged peptide and N-acetyl-L-cysteine (NAC). Conjugate addition (scavenging) reactions were carried out in a mixture of water and methanol (1:1) at a 1 mg/mL peptide concentration. At several timepoints (Figure 2.13) a sample was taken and further analysed by HPLC. In each case an excess of NAC and base (DIPEA) was used, albeit at diverse equivalents compared to the amount of peptide.

Initial experiments showed a clear increase in the NAC-conjugate formation (peptide **61**) over time (Figure 2.15). However, the concomitant formation of an additional product was also observed. After analysis by LC-MS the hypothesis was formed that a cyclisation reaction between the lysine residue and the acrylamide group could potentially occur (Figure 2.14). Alternatively, a dimerisation reaction could be taking place, however, the molecular weight fell outside of the range that was collected. Further studies into cyclisation reactions were conducted and can be found in Chapter 4.

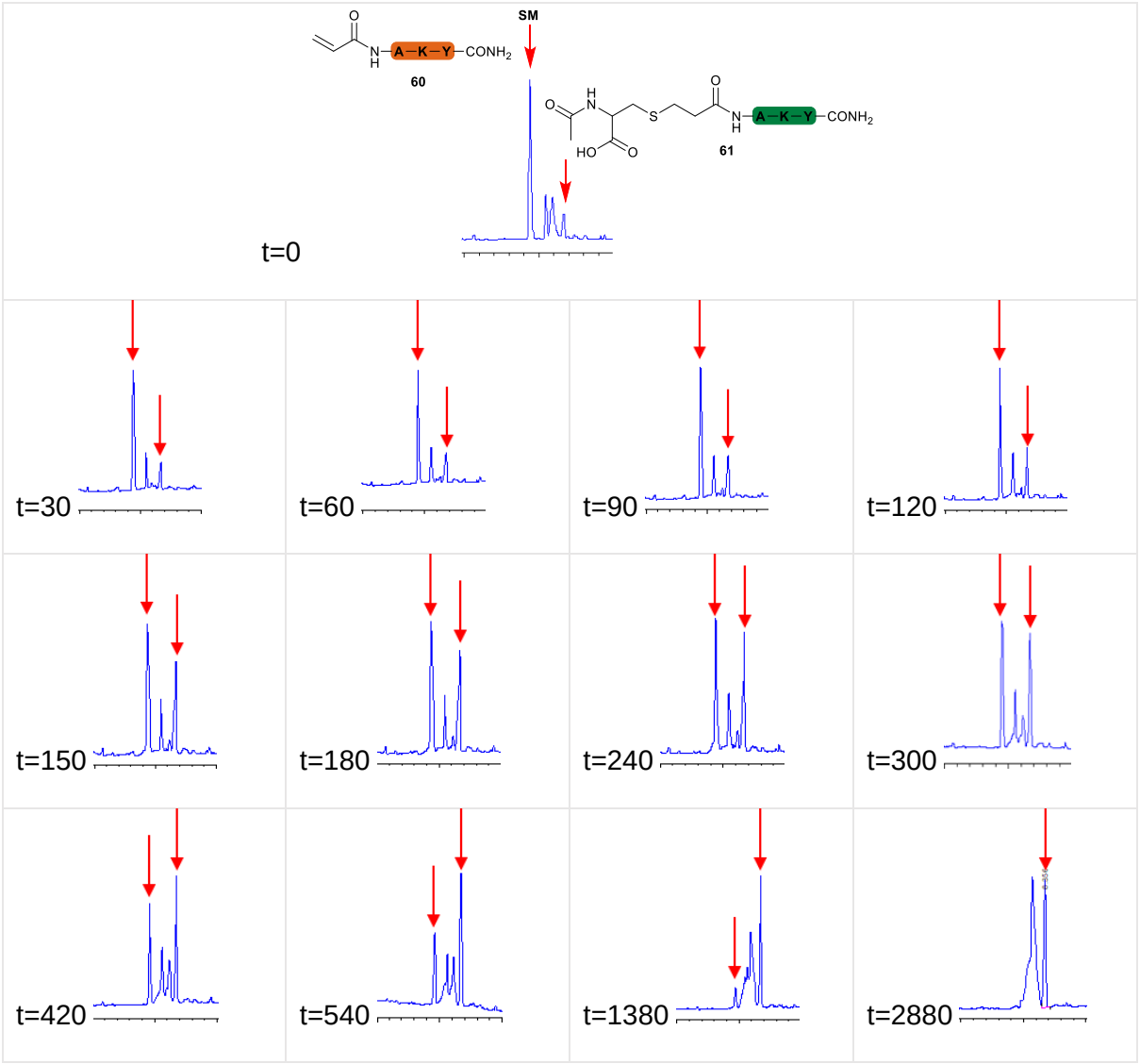


Figure 2.13: HPLC chromatogram showing product formation (peptide **61**) peak and decrease in starting material (peptide **60**). Reaction conditions are in a 1:1 solution of water and methanol at room temperature. NAC and DIPEA were added in 40 equivalents compared to peptide. Time (t) is expressed in minutes. SM = starting material.

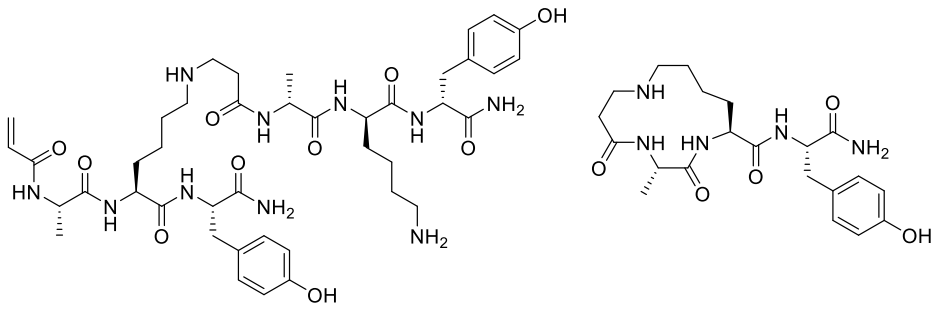


Figure 2.14: Proposed dimerised and cyclised by-products.

Increasing the amount of NAC in the reaction increased the reaction rate and reduced the amount of competing side product that was formed (Figure 2.14). While adding a great excess of one reagent can cause purification problems in solution phase synthesis, it should not cause any difficulties when the reagent (in our case the thiol) is on a solid support. It should be possible to add a great excess of thiol-resin to the reaction mixture, for faster scavenging, as demonstrated with this NAC-model.

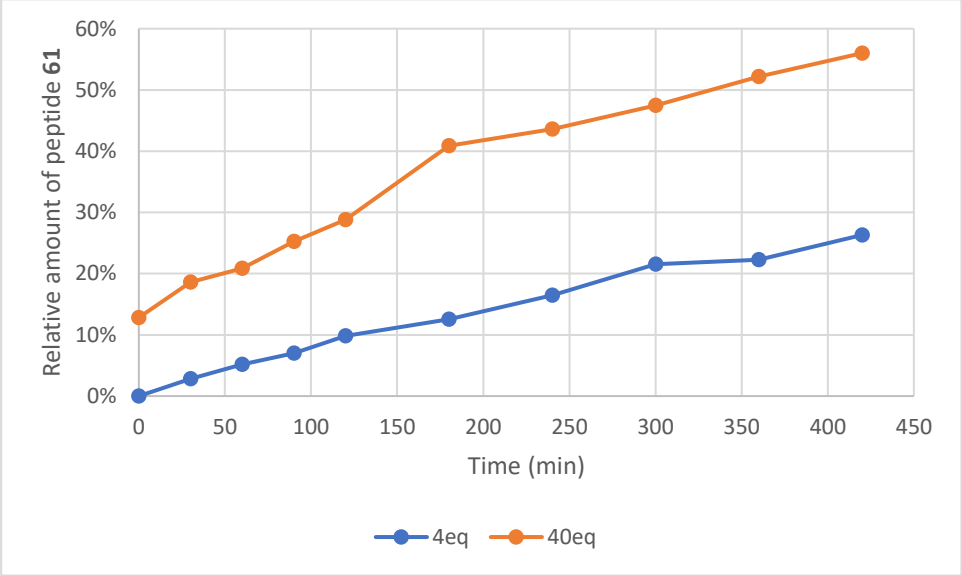


Figure 2.15: Increasing the equivalents of NAC (thiol) in the reaction mixture increases the reaction rate.

Increasing the amount of base initially increased the reaction rate, however, it does not have such a profound effect as increasing the thiol equivalents (Figure 2.16). The use of more base was expected to deprotonate the NAC thiol faster, explaining the initial increase in reactivity. However, as the amount of -SH decreases, both through reactivity and deprotonation, the effect of additional base diminishes.

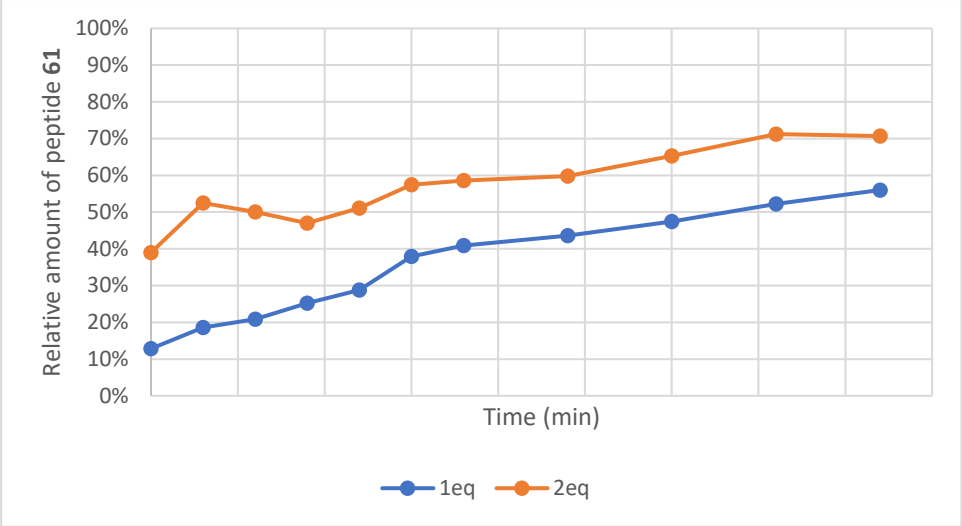


Figure 2.16: Product formation increases when more equivalents of DIPEA (base) compared to NAC (thiol) are used. The effect becomes less profound over time.

An alternative model peptide was designed to overcome any potential competitive nucleophilic side chain reactions that occurred during studies with peptide **60**. Additional experiments to determine NAC reactivity were carried out using peptide **55** (Acr-ALF-$CONH_2$). The reaction between the peptide and NAC was monitored by LC-MS both on-resin and off-resin, as well as without base and in the presence of DIPEA. Resin-based experiments were carried out to give preliminary insight in the reactivity of the peptidyl-acrylamide with thiol-resin, in which the liquid and solid state are reversed. For solid state experiments all data points were collected from the same reaction mixture. When a resin sample needed to be collected, the liquid was collected in a falcon tube, after which the resin was briefly washed and dried. A small amount of resin was taken (~ 10 mg) and the liquid was re-added to the reaction vessel. In-solution experiments were carried out in an HPLC vial, which allowed for rapid automated sampling and analysis at each time point. Single experiments were used to collect all data (time) points for each reaction condition, which was considered to provide more reliable results, as intra-experimental (*i.e.* within one dataset) manual errors, like measurement inaccuracies, are minimalised.

To confirm that base was required to promote the reaction, NAC was treated with peptide **55** with or without DIPEA present in the reaction mixture. Based on earlier experiments, an

excess of NAC (40 eq.) compared to peptide **55** was used, while DIPEA was added in equal equivalents to NAC. Resin based experiments showed no reaction in the absence of base. This was confirmed by experiments conducted in solution, supporting previously proposed base-promoted reaction mechanism.

In the presence of DIPEA the reaction occurred relatively rapidly (~ 2 hours) in solution under conventional, room temperature conditions. Interestingly, when the reaction was conducted with the peptide still on resin, the reaction occurred much slower, with approximately 3% of product formed after 24 hours.

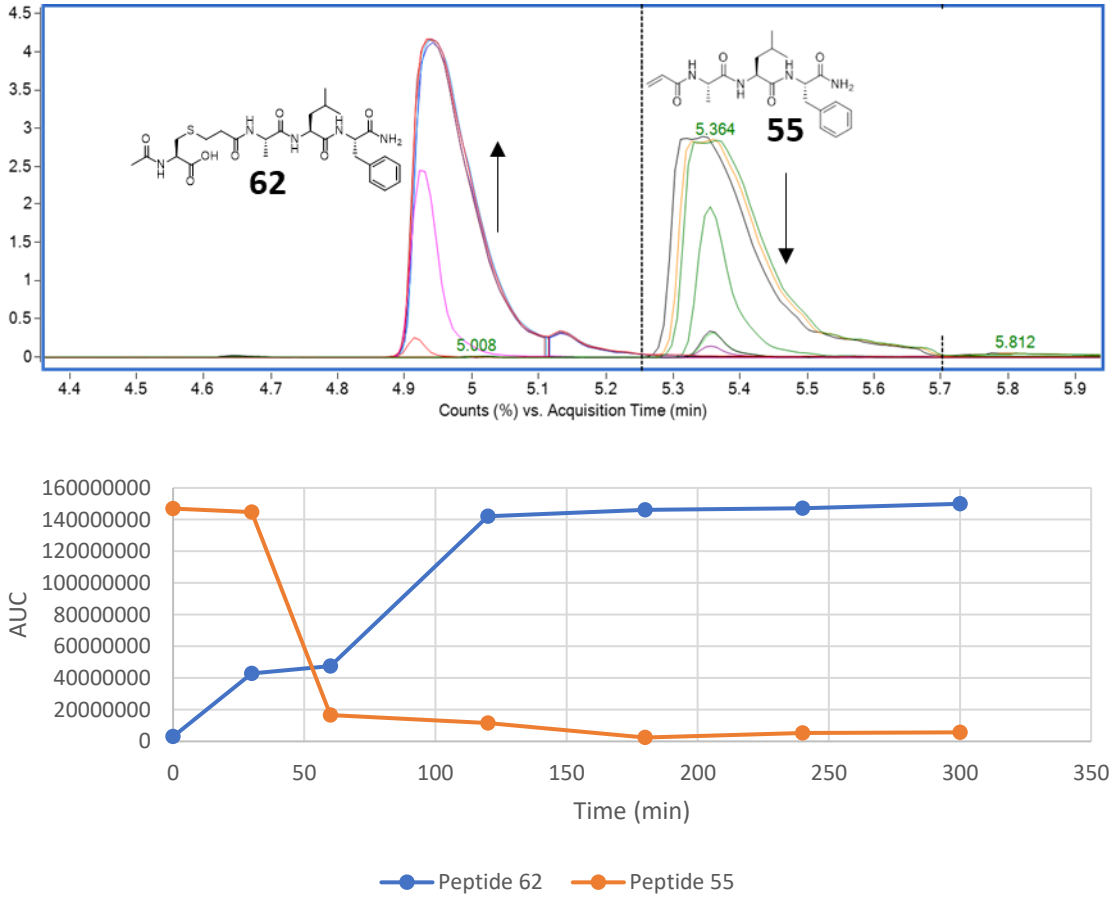


Figure 2.17: Reaction between N-acetyl-cysteine (NAC; 40 eq.) and **55** (Acr-ALF-$CONH_2$) in solution in the presence of DIPEA (40 eq.) at room temperature.

Additional on-resin experiments were carried out with microwave heating (90°C), while varying the duration of heating, to investigate whether the scavenging reaction could be accelerated

for non-heat sensitive peptides. Acr-ALF-● (**53**) and N-acetyl-L-cysteine (NAC) were loaded into a microwave vial, with DMF as solvent (3 mL), and DIPEA as base. NAC and DIPEA were present in the same equivalents (40 eq.) compared to peptide as during previous room temperature experiments. Conversion rates remained very low, not exceeding a 10% completion. Nevertheless, reaction times were much shorter compared to room temperature conditions, while still achieving slightly better conversion results.

Table 2.2: Conversion of **53** (Acr-ALF-●) to **62** (NAC-Acr-ALF-$CONH_2$) after different lengths of microwave irradiation at 90°C.

Time	Peptide 62 (%)	Peptide 53 (%)
5 min (MW)	5	95
10 min (MW)	8	92
24 hours (r.t.)	3	97

A more accurate comparison may be provided when both the peptide and NAC are in solution, as there would be no interference from the resin (*e.g.* swellability and steric hindrance). Microwave vials were loaded with 1 mg of peptide **55**, 40 eq. of NAC, 40 eq. of DIPEA, in 1 mL of water. The reaction mixture was then heated to in the microwave, for different times (Table 2.3), after which it was flash-frozen and lyophilised, prior to HPLC-MS analysis. Surprisingly, a longer reaction time did not always correlate with an increase in conjugate addition product (peptide **62**).

Table 2.3: Conversion of **55** (Acr-ALF-$CONH_2$) to **62** (NAC-Acr-ALF-$CONH_2$) after different lengths of microwave irradiation.

Time	Temperature	Peptide 62 (%)	Peptide 55 (%)
5 min	90°C	56%	44%
10 min	90°C	55%	45%
15 min	90°C	61%	39%
15 min	75°C	34%	66%
30 min	90°C	28%	72%

2.8.2 Selectivity

Acrylamide formation can occur during food processing, and has been associated with toxicity.[54] Especially the reactivity of acrylamides towards amino acids, through Michael addition has gained interest.[55] Cysteine, lysine, arginine, and serine all contain nucleophilic side chains, which could undergo a nucleophilic attack.[56] However, we hypothesise that cysteine will be more reactive than any of the other amino acids, which should lead to selectivity for the scavenging reaction over cyclisation or dimerisation reactions, if a large excess of free thiol is present. The higher reactivity of cysteine can be explained by the large higher-energy, more diffuse orbitals of the thiol-group present in cysteine, which forms an excellent soft nucleophile (Figure 2.18), while the other nucleophilic amino acids are harder nucleophiles due to more localised charge, resulting from their smaller atomic radii and greater electronegativity.[57]

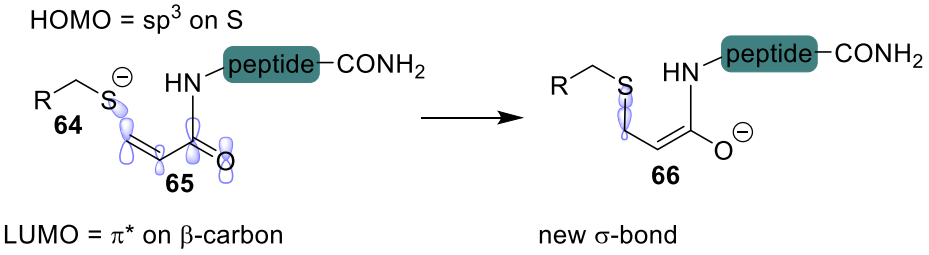
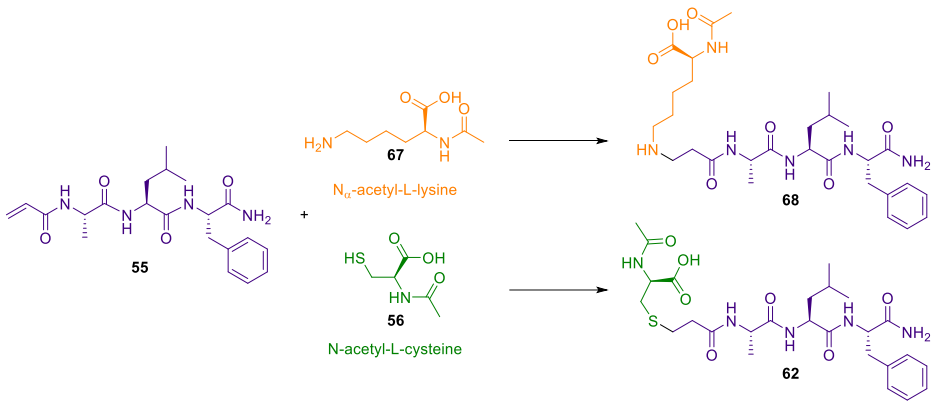


Figure 2.18: Thiols are among the best nucleophiles at undergoing conjugate additions, mainly due to the efficient orbital overlap with the LUMO of the conjugated alkene.

The selectivity for thiols over other functional groups present in amino acids was demonstrated in a competition experiment between N-acetyl-L-cysteine (NAC) and N_{α}-acetyl-L-lysine (NAL) (Scheme 2.7). The same equivalents (40 eq.) of each amino acid was added to a reaction vessel containing peptide **55**. Several reaction conditions were explored, including microwave and conventional methods, based on the reaction conditions that were used to determine thiol-reactivity. For conventional conditions, the same reaction mixture was analysed after 1 hour and after 2 hours, as previous experiments showed this period was enough for (near) completion of the reaction. Microwave conditions were also based on previous reactivity experiments, where a 15-minute reaction time achieved the highest conversion yield (61%).



Scheme 2.7: Competition experiment between N-acetyl-L-cysteine and N_{α}-acetyl-L-lysine.

After completion of the reactions the relative ratios of NAC and NAL conjugates were estimated, based on the AUC as measured by LCMS, where the assumption was made that both peptides have similar extinction coefficients and absorption maxima (215 nm). In each case, the majority of product obtained was the NAC-conjugation product, peptide **62**. These results suggest that any unwanted side reactions involving lysine can be minimised if a thiol-group is introduced in great excess.

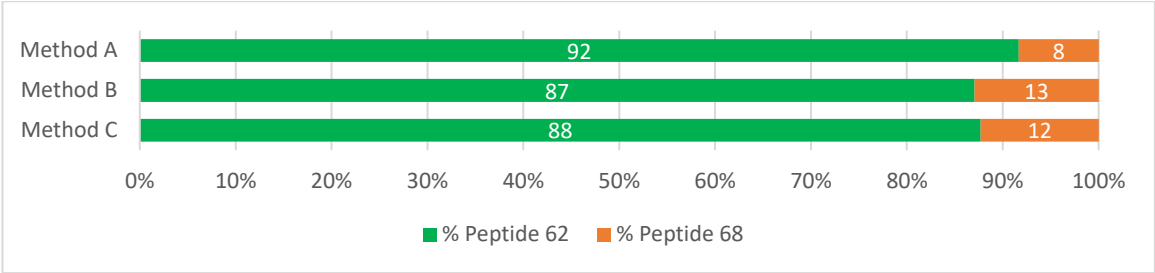


Figure 2.19: All experiments were conducted in water at a 1 mg/mL concentration. Percentages are based on the AUC as measured by UV at a wavelength of 215 nm. Method A) Microwave, 75°C, 15 min, in water; Method B) Conventional, 25°C, 1 hour; Method C) Conventional, 25°C, 2 hours.

2.9 Compatibility of acrylamides with standard solid phase synthesis techniques

Standard Fmoc-solid phase peptide synthesis only requires a few reagents; an amino acid, activator, and activator base for amino acid couplings, and a base for deprotections. A commonly used selection of these reagents comprises of DIC as activator, Oxyma as activator base, and piperidine as deprotection base (Figure 2.20). To be able to use our acrylamide-scavenging approach, it is important that the acrylamide-group remains stable through all synthesis steps, without interfering with the amino acid couplings.

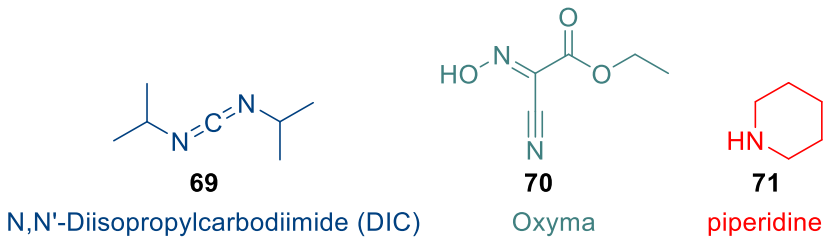
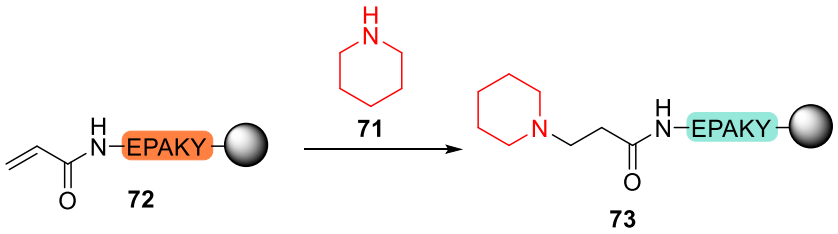


Figure 2.20: Structures of commonly used reagents during SPPS.

Exposure of an acrylamide-tagged peptide (**72**) on resin (25 mg, Acr-EPAKY-●) to coupling reagents DIC and Oxyma under standard amino acid coupling conditions (0.5 M in DMF, 1.5 minutes, 90°C, MW) did not provide any side reactions. The only reagent that was considered to potentially be problematic was piperidine (**71**), as evidence from the literature suggests it can slowly react with acrylamides under high temperature.[58–60] During automated microwave-assisted peptide synthesis, reactions are heated up to ~90°C to speed up the piperidine promoted deprotection step, which could lead to the formation of peptidyl-acrylamide-piperidine conjugates (Scheme 2.8).



Scheme 2.8: Proposed reaction between piperidine and peptidyl-acrylamide.

Acrylamide-tagged peptide **72** was synthesised (on-resin, ~25 mg) and treated with piperidine under both microwave and conventional deprotection conditions, to observe whether any piperidine adducts could be detected. Standard SPPS deprotection protocols were used, which was considered to give most representative outcome. Microwave conditions were based on the synthesis cycle programmed in the Liberty Blue™ automated peptide synthesiser, 1.5 minutes at 90°C. Conventional conditions utilised a 2x 15 minutes reaction time at room temperature. As can be seen in Figure 2.21 a piperidine adduct was observed, with a ~39% conversion under conventional conditions and even up to 97% conversion under microwave conditions, as determined by LCMS.

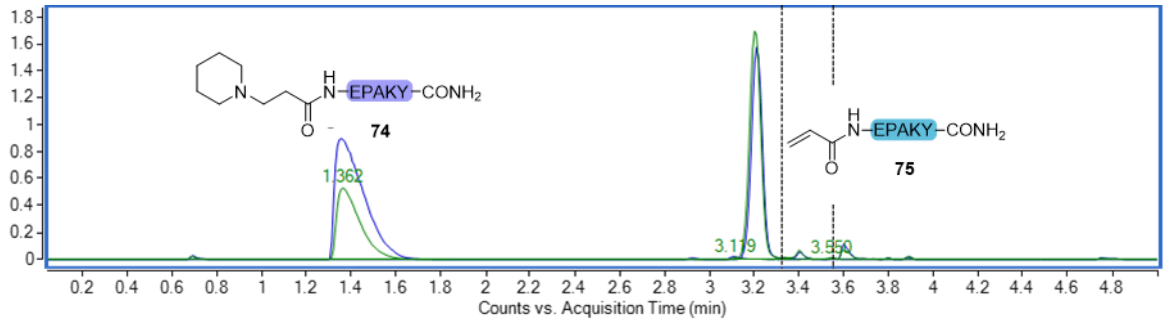
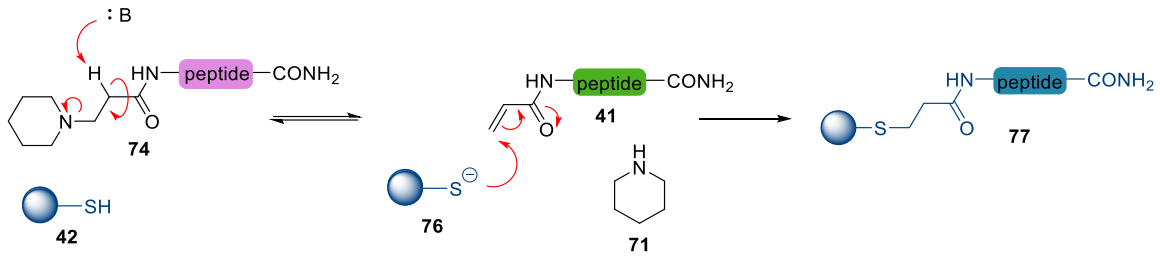


Figure 2.21: Formation of piperidine adduct under conventional (green) and microwave (blue) conditions.

Given the outcomes, piperidine was considered to be a poor choice of base for the deprotection in the presence of the acrylamide-group. If up to 97% of the acrylamide is converted to piperidine propanamide **74**, the thiol scavenger will be unable to react and isolate side products. Reoccurring deprotection cycles take place after coupling of each amino acid, which will likely push the Michael Addition of piperidine to completion by the end of the synthesis. Nevertheless, the piperidine adduct could also accelerate the speed and selectivity of the covalent attachment between the thiol and the peptide, if the piperidine can function as a good leaving group, as conjugate addition is often reversible.[61] Two different approaches were proposed and investigated.

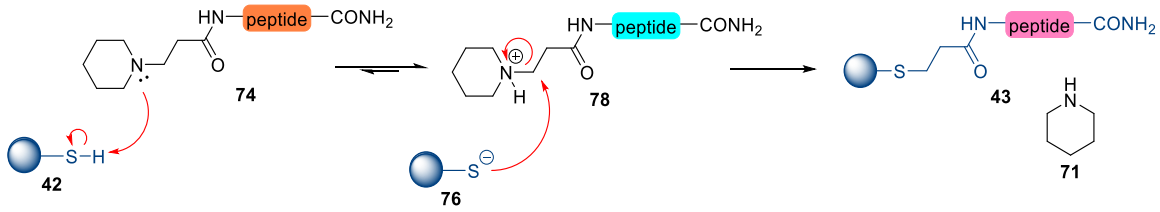
Initially, we proposed a reaction in which the addition of base was required to eliminate piperidine by removal of one of the α-protons (Scheme 2.9). This would restore the original

acrylamide group. The Michael Addition between the thiol-resin and the acrylamide-tagged peptide could then take place. Reformation of the conjugated enone system would be the driving force of this reaction, however, the neutral piperidine was considered to be a poor leaving group.



Scheme 2.9: Proposed base-catalysed substitution of the piperidine group by thiol resin.

Alternatively, we considered that the conjugate addition of piperidine to acrylamide might be reversible and, therefore, substitutable if protonated by the thiol sulphur, given that the tertiary amine is more basic (pKa ~11) than a thiol (pKa ~6.4). The now negatively charged sulphur can conceptually attack the *β*-carbon and substitute the protonated (positively charged) piperidine leaving group. The procedure could be auto-catalytic and would not require addition of an additional base to promote the reaction. A benefit of the process described here would be the functioning of piperidine as a protecting group of the acrylamide during the peptide synthesis, resulting in a suppression of potential polymerisation side products.



Scheme 2.10: Proposed auto-catalytic substitution of the piperidine group by thiol resin.

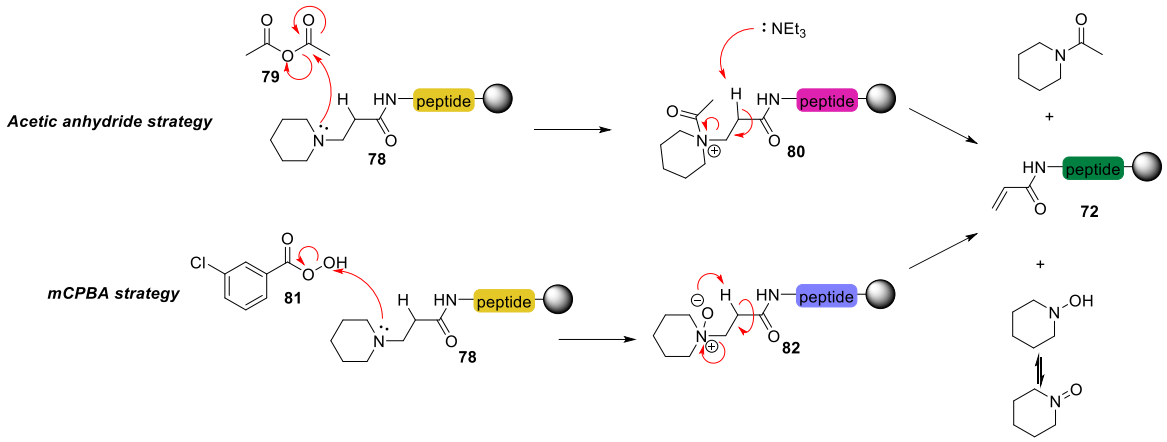
To test proposed hypothesis, the piperidine adduct (peptide **74**) was dissolved in a mixture of water and methanol (1:1, 1 mg/mL) and was treated with the thiol resin (500 mg, 1.3 mmol/g loading) in the microwave (90°C, 15 minutes). These reaction conditions correspond to approximately 430-fold excess of thiol, which should be sufficient to drive the reaction.

However, no decrease in the amount of piperidine adduct **74** was observed over time, suggesting there was no reaction through an auto-catalytic mechanism. Given that no reaction had been observed, base was added to the reaction mixture to investigate the potential of the base-catalysed substitution. DIPEA was added to the reaction mixture (226 uL, 2 eq.), followed by microwave irradiation for 15 minutes. Again, no decrease in peptide **74** was observed by LCMS. A different approach to reforming the acrylamide group, prior to reaction with the thiol resin, was required.

Two different strategies to overcome the piperidine adduct problem were proposed. The first strategy focussed on reforming the acrylamide group at the end of the peptide synthesis, by increasing the reactivity of piperidine as a leaving group. The second strategy was avoiding the use of piperidine altogether, by exploring the possibility of alternative deprotection protocols.

2.9.1 Restoring the acrylamide group

Attempts to directly substitute or eliminate the piperidine group were unsuccessful, therefore, a different approach was required. Conceptually, it should be possible to eliminate piperidine and reform the acrylamide group while the peptide is still on resin, to prevent the occurrence of reactions with unprotected peptide side chains. Two different on-resin strategies were explored: acetylation of the piperidine nitrogen using acetic anhydride (Method A), and oxidation of the piperidine nitrogen using *m*CPBA (*m*-chloroperoxybenzoic acid) (Method B). Both approaches attempt to transform the piperidine into a better leaving group, by quaternising the nitrogen at the end of the peptide synthesis. The positively charged piperidine should be a better leaving group, making it susceptible for substitution or elimination.



Scheme 2.11: Proposed strategies to restore the acrylamide group at the end of peptide synthesis.

2.9.1.1 Acetic anhydride approach (Method A)

The reformation of an acrylamide by elimination of piperidine after acetylation with acetic anhydride has been reported in a Chinese patent[62] following a two-step reaction. The reaction conditions that were described in the patent were not fully compatible with solid-phase peptide synthesis, so these were adapted. In the patent, toluene was used as a solvent, however, this was replaced by DMF in our protocol, to improve the swellability of the resin. In the patent, the reaction mixture, containing 1 eq. of acetic anhydride, was gradually raised to 60-80°C and stirred for two hours. Instead, we undertook the reaction in the microwave, heating it to 80°C for 30 minutes. This first step forms the N-acetyl-piperidine leaving group, which could conceivably be eliminated upon addition of triethylamine (1 eq.). A polymerisation inhibitor BHT was used in the protocol of the patent, however, this was not deemed necessary in our protocol, as all peptidyl-acrylamide is still attached to the resin, inhibiting polymerisation by spatial separation. Analysis was conducted by LCMS after cleavage from the resin. A 98% reformation of acrylamide (**72**) was achieved for peptide **78** (Pip-Acr-EPAKY). One limitation of using this approach to restore the acrylamide group is the lack of compatibility with free amines. As the step to quaternise the piperidine-N uses acetic anhydride, this results in an acetyl-N-terminus of the non-tagged desired peptide, if free amine is present.

2.9.1.2 mCBPA approach (Method B)

A second strategy used the oxidising agent *m*CPBA, based on a similar reaction previously reported in literature.[63] In this case a [4-(2-piperidin-1-ylethylsulfanyl)- phenyl] carbamic acid tert-butyl ester decomposed to N-hydroxypiperidine and vinyl sulfone, via triple oxidation. We hypothesised this method could also be used for restoring the acrylamide group. To achieve this, the peptide (**78**) on-resin was swollen in DCM, followed by the addition of 5 equivalents of mCPBA, which was left shaking for 2 hours at room temperature. This procedure lead to recovery of ~99% of the original acrylamide peptide from peptide **78** as determined by LC-MS.

2.9.1.2.1 Nitroso group formation

Upon employing the mCPBA approach a new product was observed when a peptide containing an unprotected amine was present, which was suspected to be the result of a nitroso-group. To confirm that *m*CPBA introduced a nitroso-group at the N-terminus, the reaction was repeated using two model peptides, Ac-ALF-● (**83**) and NH_2-ALF-● (**52**). If the proposed reaction takes place at the amine group, no change to the starting material would be observed for the acetyl-capped peptide. Indeed, analysis by LC-MS showed no change, and only the Ac-ALF-$CONH_2$ peptide was observed (found MW $[M+H]^+$ 391.2343; calculated 390.23).

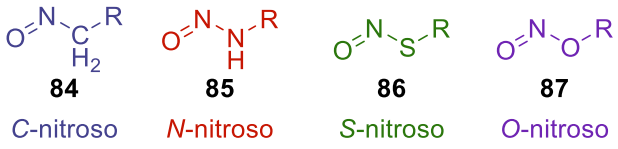
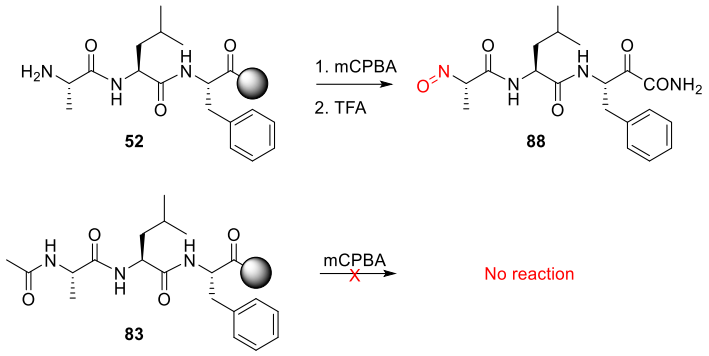


Figure 2.22: Nitroso functional group.

The reaction between NH_2-ALF-● and *m*CPBA resulted in a more complex mixture, confirming a reaction takes place at the N-terminus. The majority of peptide **52** was converted in the nitroso peptide **88** (O=N-ALF-$CONH_2$, found MW $[M+H]^+$ 363.2030; calculated 362.20).



Scheme 2.12: Model reactions for the determination of nitroso-group formation.

Functionalisation of the peptide N-terminus can be interesting for expanding their biological activity, or for chemical ligation purposes. To our best knowledge, C-Nitroso peptides have not been reported in literature to date. More commonly, the nitrosamine (**85**) ($R^1N(–R^2)–N=O$) has been reported (Figure 2.22), which can be introduced in compounds by nitrosation.[64,65] Reduction back to the amine is the most commonly described reaction for the nitroso-group.[66] Alternatively, the nitroso-group can be further oxidised to form a nitro-group.[67] Furthermore, a Diels-Alder reaction between nitroso-compounds and 1,3-dienes have been described, though this reaction may only be applicable to aromatic nitroso groups.[68]

2.9.2 Alternative approaches to Fmoc-deprotection

To overcome the problematic reaction between piperidine and acrylamide a range of different deprotection bases were explored. Piperidine contains a secondary amine making it a suitable candidate for a Michael Addition to the β-carbon of the acrylamide. By using a tertiary amine as base, the reaction between the acrylamide should be either avoided or be significantly inhibited. If a base-peptide adduct does form with a tertiary amide, the product should be unstable enough to be readily substituted for the more stable thiol-ene product.

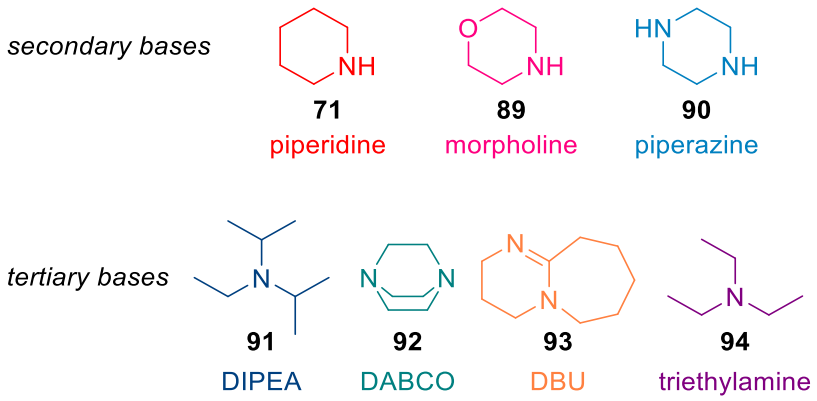


Figure 2.23: Overview of bases that could potentially be used as deprotection reagents.

Morpholine and piperazine are sometimes used during SPPS as they are cheaper and less toxic than piperidine.[69] However, morpholine also has a free NH, while piperazine even contains two, making both molecules unsuitable as piperidine replacements for our purposes. Evidence in the literature demonstrated the possibility for Fmoc-deprotection with DBU **93** (1,8-Diazabicyclo[5.4.0]undec-7-ene), DABCO **92** (1,4-diazabicyclo[2.2.2]octane), Et_3N **94**, $NaCO_2$, NaOtBu, KOtBu, and DIPEA **91** (N,N-diisopropylethylamine)[70,71]. We chose to not explore inorganic, salt-based bases, due to concern of potential build-up of precipitation within the automatic peptide synthesiser. Tertiary bases DIPEA, DABCO, DBU, and Et_3N were of interest to explore further.

Two sets of experiments were carried out: first, the compatibility of several tertiary bases with acrylamide was explored. It is of interest to find a base that does not react with the acrylamide group, but still has the capability to deprotect the base-labile Fmoc-group during synthesis with high efficiency. Therefore, secondly the deprotection potential of all bases was determined under both conventional and microwave deprotection conditions.

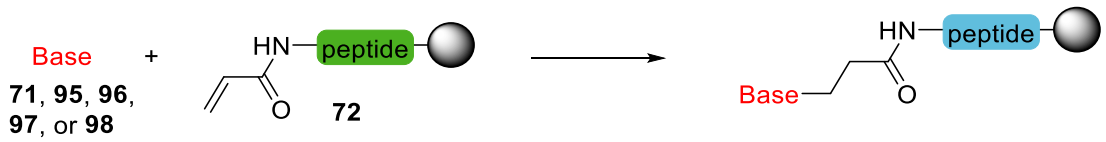
2.9.2.1 Acrylamide tolerance

A short 5 amino acid on-resin peptide **72** (Acr-EPAKY-●) was synthesised to analyse the reactivity between the N-terminal acrylamide group and the different bases that could be used as Fmoc deprotection reagents. From previous experiments it was already known that a reaction with piperidine would take place, however, the reaction was repeated with model

peptide **72** to allow for better comparison with the other bases. After reactions under microwave (90°C, 1.5 min) or conventional conditions (2x 15 min, RT) peptides were cleaved from the resin (TFA/TIPS/H_2O, 8:1:1) and analysed by LCMS. The outcomes of the reaction between acrylamide and base were estimated based on the AUC and are reported in Table 2.4.

As expected, the highest conversion to base-peptide adducts formed in the reaction with the secondary amine piperidine. The tertiary bases either reacted much more slowly with the acrylamide or the product may have been unstable and re-formed the acrylamide upon cleavage. The amount of base-adduct formed was generally acceptably low so as to continue to explore the alternative deprotection approach using each of the bases. Based on these results DBU would be the least favourable base to continue with, while DIPEA shows the most promise. Nevertheless, each base shows a considerably lower conversion than piperidine.

Table 2.4: Reaction of different bases with acrylamide-tagged peptides using a model peptide **72** (Acr-EPAKY-●). *Percentages reported are the amount of base-peptide adduct formed after a) conventional (2x 15 min, room temperature) or b) microwave (1.5 minutes, 90°C) conditions as determined by LC-MS (215 nm).



Base	Conventional[a]	Microwave[b]
20% piperidine (**71**)	39%	98%
20% DIPEA (**91**)	0%	0%
20% DABCO (**92**)	0%	0%
20% DBU (**93**)	1%	2%
20% Et_3N (**94**)	0%	0%

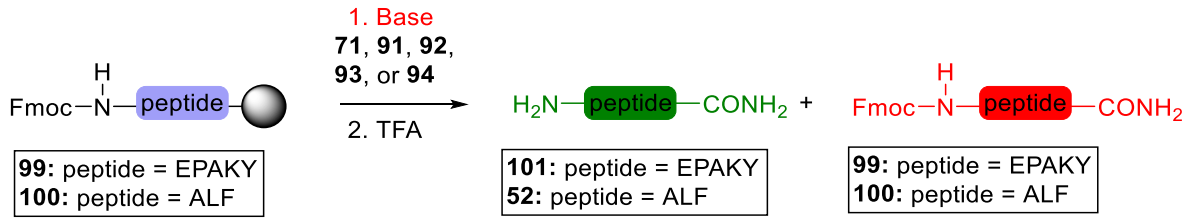
2.9.2.2 Fmoc deprotection potential

The deprotection efficiency of different bases that were found to be compatible with the acrylamide, were explored and compared to the 'standard' piperidine deprotection. All bases were prepared as 20% (*v/v*) solutions in DMF (standard for piperidine), allowing for future optimisation of the conditions if the initial results seemed promising. As with the acrylamide tolerance studies, microwave (90°C, 1.5 min) and conventional conditions (2x 15 min, RT) were investigated. Analysis was performed by LCMS, from which the AUC was used to estimate the deprotection potential.

The results (Table 2.5) show that DIPEA is not efficient, with around 2% deprotection observed. Consequently, investigations into this base were discontinued. This outcome was not entirely unexpected, and could even be considered good, as DIPEA can be used in conjunction with an activator for amino acid couplings.[72] In this case, if DIPEA would remove Fmoc-groups it could interfere with the success of clean couplings. The reason why DIPEA might be a poor deprotection base could be the shielding of the nitrogen by steric hindrance of the ethyl and isopropyl-groups.

Triethylamine showed some potential, yet with room for improvements, with initial deprotection percentages being close to 50% for microwave conditions. Attempts were made to improve the deprotection rates, by increasing Et_3N concentration, the reaction time, or by double deprotection. Interestingly, increasing the amount of base to 50% (*v/v*) Et_3N did not improve the deprotection efficiency. The other two alterations to the Et_3N deprotection protocol, doubling reaction length and performing the reaction step twice, increased the deprotection amount slightly, with the double deprotection reaching up to 82% Fmoc-removal. However, the results were deemed insufficient to continue with Et_3N as a deprotection reagent.

Table 2.5: The removal of the Fmoc-group using different bases was explored with the use of a model peptide. *Percentages reported are the amount of deprotected peptide after a) conventional (2x 15 min, room temperature) or b) microwave (1.5 minutes, 90°C) conditions, unless otherwise specified, as determined by LC-MS. Percentages reported are based on mass detection.



Base		Peptide	Conventional[a]	Microwave[b]
Piperidine (**71**)	20%	Fmoc-EPAKY-●	94%	88%
DIPEA (**91**)	20%	Fmoc-EPAKY-●	2%	2%
DBU (**93**)	20%	Fmoc-EPAKY-●	93%	94%
Et_3N (**94**)	20%	Fmoc-EPAKY-●	34%	48%
	50%	Fmoc-EPAKY-●	-	44%
	20% - 3 min	Fmoc-EPAKY-●	-	57%
	2x 20%	Fmoc-EPAKY-●	-	82%
DABCO (**92**)	20%	Fmoc-EPAKY-●	96%	93%
	20%	Fmoc-ALF-●	-	99%
	15%	Fmoc-ALF-●	-	99%
	10%	Fmoc-ALF-●	-	99%
	5%	Fmoc-ALF-●	-	>99%

DBU afforded deprotection rates comparable to piperidine deprotections under conventional conditions (93%) and slightly improved results under microwave conditions (94%). However, based on its reactivity towards the acrylamide, it was more favourable to continue optimisation of the reaction conditions for the other bases with equivalent deprotection potential.

The most promising base was DABCO, which afforded even greater deprotection rates than piperidine for peptide **99** (96% compared to 94% under conventional conditions). To confirm the deprotection potential of DABCO, additional experiments were carried out using a different model peptide (Acr-ALF-●, peptide **100**). This different model was used to ensure the reactivity

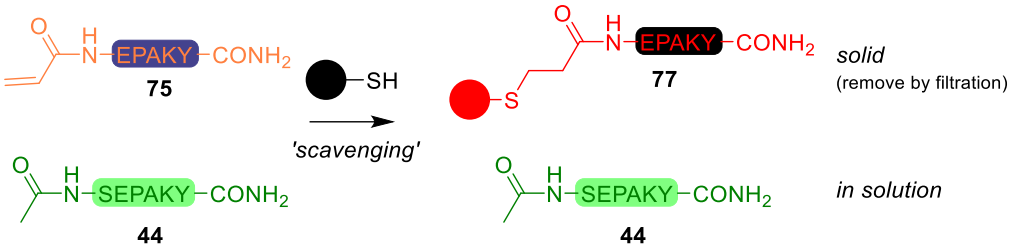
of DABCO was independent of the peptide sequence. Additionally, experiments using lower percentages of base were performed. This showed that a concentration as low as 5% (*v/v*) of DABCO in DMF could achieve excellent deprotection rates. This knowledge was taken forward in the development of the synthesis protocol for automated peptide synthesis, which can be found in Chapter 3.

2.9.3 Optimisation conclusions

While the formation of an acrylamide-piperidine adduct initially appeared to be problematic, successful strategies have been developed to overcome this. Two approaches were proposed; reforming the acrylamide group at the end of the synthesis or avoiding the use of piperidine as a deprotection reagent. Both approaches can be used for our scavenging procedure, however, there are some limitations. If the desired peptide has an N-terminal free amine, only the DABCO deprotection procedure can be applied. For acetyl-capped peptides of interest, both acrylamide restoration approaches and DABCO deprotections can be used.

2.10 Modelling of optimised approach: Separation of single deletion from desired peptide product

The purpose of introducing an acrylamide-tag on peptide impurities is for scavenging unwanted ‘deletions’ after cleavage from the solid support. We have successfully incorporated acrylamide in to our peptides and have shown its reactivity and preference towards a model thiol. In the final stage of our optimisation studies, we wanted to demonstrate how acrylamide-tagged impurities can successfully be removed from a mixture with our desired peptide, using a solid supported resin containing a thiol functional group.



Scheme 2.13: Scavenging of acrylamide-tagged peptide **75**, while acetyl-capped peptide **44** remains in solution.

To monitor the efficiency of the scavenging reaction two peptides were prepared (Ac-SEPAKY-$CONH_2$ (**44**) and Acr-EPAKY-$CONH_2$ (**75**)), of which one contains an acetyl cap and the other our new acrylamide-tag (Scheme 2.13). The peptides were mixed in a 1:1 ratio to a final concentration of 0.05 mg/mL of each peptide (10 mL, H_2O/MeOH, 1:1; 0.5 mg of each peptide). Various amounts of base and thiol-resin (ISOLUTE® Si-Thiol Biotage, 1.3 mmol/g) were added, and the different 'scavenging' experiments were monitored by LC-MS over time. The amount of Acr-tagged peptide that was scavenged was quantified using the peptide of interest (peptide **44**) as internal standard in a simple competition experiment.

For this study, DIPEA was chosen as a base due to its non-nucleophilic character, alongside being volatile, simplifying its removal from the cleaned-up peptide solution. However, as mentioned previously in Section 2.8, addition of DIPEA to the reaction appears to induce a side reaction. For scavenging purposes, it was found that increasing the amount of thiol-resin present decreases the amount of unwanted side product.

Table 2.6: Overview of explored scavenging conditions.

Method	SH-Resin		DIPEA		TIPS
	mg	mmol	µL	mmol	µL
A	50	0.065	20	0.13	-
B	500	0.65	226	1.3	-
C	500	0.65	226	1.3	20
D	500	0.65	452	2.6	-
E	1000	1.3	452	2.6	-

The results of scavenging of a single deletion in a peptide mixture are displayed in Figure 2.24. A clear decrease in peptide **75** was observed over time in all cases, while there was also a slight increase in side product formation. When using different amounts of thiol-resin a difference in scavenging rate was observed, where the 1000 mg of resin (~1300 eq.) was the only condition (Method **E**) to reach completion after 3 hours (Figure 2.25). In general, the trend follows an increase in resin increases the reaction speed, which is also assisted by an increase in the amount of base present. However, an increase in the amount of base also appeared to catalyse the side reaction, so should be added with caution.

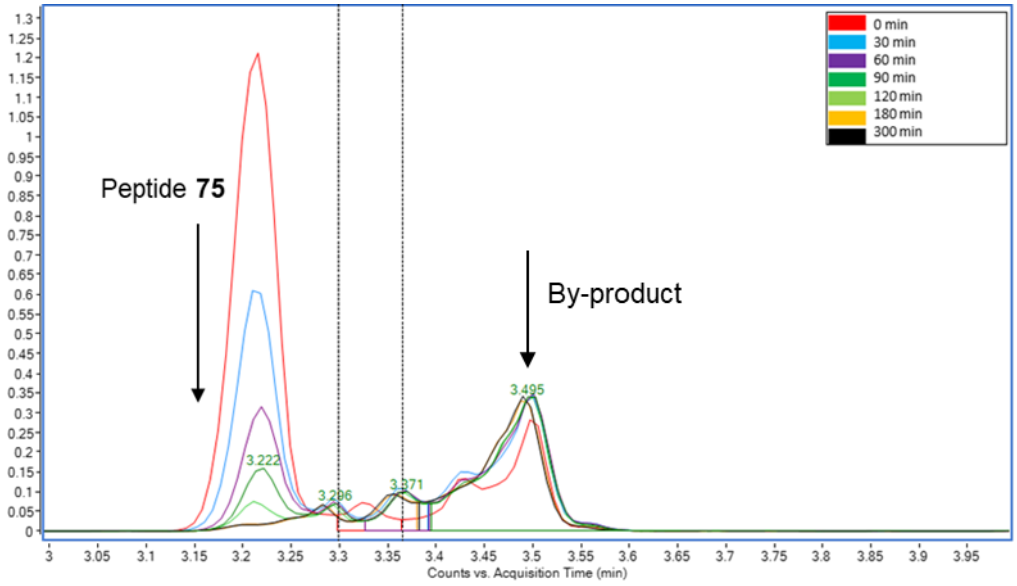


Figure 2.24: Decrease in peptide **75** after prolonged scavenging time following method **E**.

Using 1000 mg of resin for a total amount of peptide of 0.5 mg (~1300 equivalents of thiol compared to acrylamide), the scavenging reaction was expected to take place at a much faster rate than the experimentally measured 3 hours. The slower reaction rate for the thiol-resin compared to free N-acetyl-L-cysteine could be explained through a lower exposure of binding sites, because of steric hindrance from the solid support resin. A similar issue was observed when the reaction between NAC and peptide **53** was studied with the peptide still on resin. Consequently, previously calculated reaction rates between NAC and acrylamide-tagged peptide **55** and **60** are not directly translatable to thiol-resin scavenging.

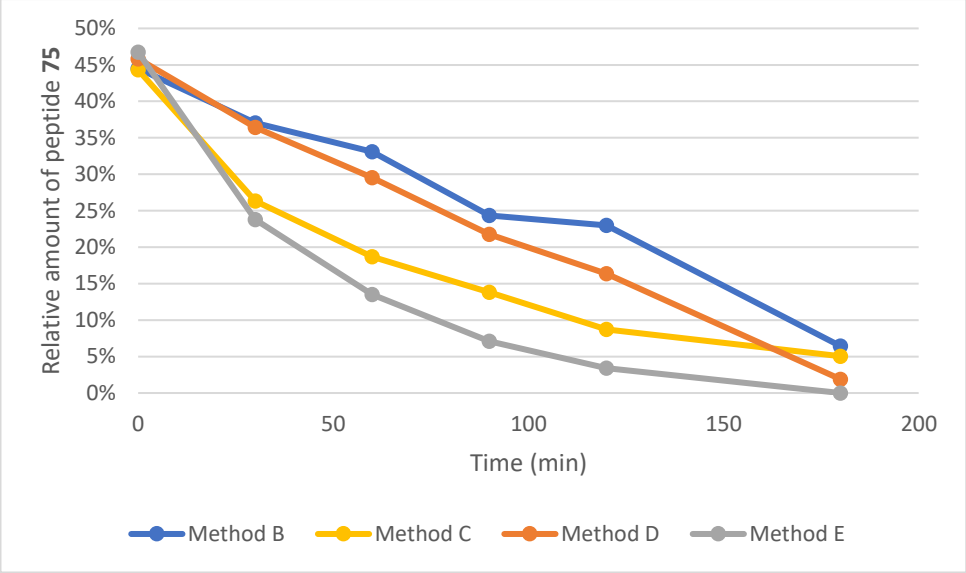


Figure 2.25: Removal of peptide **75** from reaction mixture over time.

2.11 Conclusion

In this chapter we aimed to develop and optimise a novel peptide purification protocol. We considered the applicability to standard solid phase peptide synthesis to be essential, which led to the idea of acrylamide-tagging of peptide impurities. Using acrylamide was hypothesised to be dual-functional; it would prevent the formation of truncated peptide impurities and it could function as a scavenging tag in a final purification step.

A method to 'tag' peptides with an acrylamide-group has been developed successfully. A quantitative conversion was required for effectiveness of this purification method. Reaction conditions have been determined under which a 100% conversion yield was obtained (0.25 M, 1.5 min, 90°C, MW), which were applied to all further experiments.

The possibility to scavenge acrylamide-tagged peptides was demonstrated in a model reaction with N-acetyl-L-cysteine. This provided us with more insight in the reaction progress and an unexpected side-reaction (potential cyclisation) was discovered. This prompted us to explore the selectivity of acrylamides further. A ~9-fold preference for a thiol functional group over an amine functional group was observed in competition experiments. These results suggested that the thiol-scavenging approach could be successful if a large excess of thiol was present.

Before being able to apply acrylamide-tagging to standard SPPS procedures the compatibility with commonly used reagents, including piperidine, Oxyma, and DIC, was determined. The acrylamide group was found to be stable in the presence of Oxyma and DIC, but not with piperidine, as piperidine could undergo a nucleophilic attack on the acrylamide *β*-carbon. Two strategies to overcome this problem were explored, and both were successful. Firstly, it was found to be possible to restore the acrylamide-group at the end of the peptide synthesis, using either acetic anhydride or *m*CPBA, under the condition that the desired (un-tagged) peptide does not have a free N-terminal amine. Secondly, an alternative deprotection protocol was developed to avoid the need of piperidine altogether. Deprotections using DABCO were found to be the most effective.

A first insight in the scavenging potential was obtained when one acrylamide-tagged peptide was removed from a mixture (1:1) with an acetyl-capped peptide. The scavenging rate was found to be dependent on the amount of thiol-resin present, with an increase in base also accelerating the reaction. Formation of a side product increased when more base was used, however, this could largely be supressed by using a greater excess of thiol-resin.

In the end, the amount of resin that is needed for scavenging may be dependent on the individual quality of the peptide synthesis, as well as the time available for purification. Increasing the amount of resin may increase reaction speed, but could be considered more wasteful, if not all available thiol-residues on the resin are being used. General guidelines for the amount of impurity versus the amount of resin and time necessary for purification can be estimated, though may vary slightly per individual peptide.

2.12 References

1. Curtius, T. Ueber einige neue der Hippursäure analog constituirte, synthetisch dargestellte Amidosäuren. *J. für Prakt. Chemie* **26**, 145–208 (1882).

2. Fischer, E. & Fourneau, E. Ueber einige Derivate des Glykocolls. *Berichte der Dtsch. Chem. Gesellschaft* **34**, 2868–2877 (1901).

3. Bergmann, M. & Zervas, L. Über ein allgemeines Verfahren der Peptid-Synthese. *Berichte der Dtsch. Chem. Gesellschaft (A B Ser.* **65**, 1192–1201 (1932).

4. Harington, C. R. & Mead, T. H. Synthesis of Glutathione. *Biochem. J.* **29**, 1602–1611 (1935).

5. Sifferd, R. H. & du Vigneaud, V. A new synthesis of carnosine with some observations on the splitting of the benzyl group from carbobenzoxy derivatives and from benzylthio ethers. *J. Biol. Chem.* **108**, 753–761 (1935).

6. du Vigneaud, V., Resslbr, C., Swan, J. M., Roberts, C. W. & Katsoyannis, P. G. The Synthesis of Oxytocin. *J. Am. Chem. Soc.* **76**, 3115–3121 (1954).

7. McKay, F. C. & Albertson, N. F. New Amine-masking Groups for Peptide Synthesis. *J. Am. Chem. Soc.* **79**, 4686–4690 (1957).

8. Sheehan, J. C. & Hess, G. P. A New Method of Forming Peptide Bonds. *J. Am. Chem. Soc.* **77**, 1067–1068 (1955).

9. Bodanszky, M. & Vigneaud, V. du. A Method of Synthesis of Long Peptide Chains Using a Synthesis of Oxytocin as an Example. *J. Am. Chem. Soc.* **81**, 5688–5691 (1959).

10. Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149–2154 (1963).

11. Carpino, L. A. & Han, G. Y. 9-Fluorenylmethoxycarbonyl amino-protecting group. *J. Org. Chem.* **37**, 3404–3409 (1972).

12. Sakakibara, S., Shimonishi, Y., Kishida, Y., Okada, M. & Sugihara, H. Use of anhydrous hydrogen fluoride in peptide synthesis. I. Behavior of various protective groups in anhydrous hydrogen fluoride. *Bull. Chem. Soc. Jpn.* **40**, 2164–7 (1967).

13. Jaradat, D. M. . Thirteen decades of peptide synthesis: key developments in solid phase peptide synthesis and amide bond formation utilized in peptide ligation. *Amino Acids* **50**, 39–68 (2018).

14. Arshady, R. Synthesis and characterization of polymer supports. *Makromol. Chem* **189**, 1295–1303 (1988).

15. Shelton, P. T. & Jensen, K. J. Linkers, Resins, and General Procedures for Solid-Phase Peptide Synthesis. in 23–41 (Humana Press, Totowa, NJ, 2013). doi:10.1007/978-1-62703-544-6_2

16. Sarin, V. K., Kent, S. B. H. & Merrifield, R. B. Properties of Swollen Polymer Networks. Solvation

and Swelling of Peptide-Containing Resins in Solid-Phase Peptide Synthesis. *J. Am. Chem. Soc.* **102**, 5463–5470 (1980).

17. Canal, T. & Peppas, N. A. Correlation between mesh size and equilibrium degree of swelling of polymeric networks. *J. Biomed. Mater. Res.* **23**, 1183–1193 (1989).

18. Merrifield, R. B., Stewart, J. M. & Jernberg, N. Instrument for automated synthesis of peptides. *Anal. Chem.* **38**, 1905–1914 (1966).

19. Merrifield, R. B. & Stewart, J. M. Automated peptide synthesis. *Nature* **207**, 522–523 (1965).

20. Jensen, K. J., Tofteng, P., Søren, S. & Pedersen Editors, L. *Peptide Synthesis and Applications*. (Springer Science, 2013). doi:10.1007/978-1-62703-544-6

21. Erdélyi, M. & Gogoll, A. Rapid microwave-assisted solid phase peptide synthesis. *Synthesis (Stuttg).* 1592–1596 (2002). doi:10.1055/s-2002-33348

22. M. Collins, J. & J. Collins, M. Microwave-enhanced Solid-phase Peptide Synthesis. in *Microwaves in Organic Synthesis* 898–930 (Wiley-VCH Verlag GmbH, 2008). doi:10.1002/9783527619559.ch20

23. Tolomelli, A., Ferrazzano, L. & Amoroso, R. 4-Substituted Prolines: Useful Reagents in Enantioselective Synthesis and Conformational Restraints in the Design of Bioactive Peptidomimetics. *Adricimica ACTA* **50**, 3–15 (2017).

24. Haack, T. & Mutter, M. Serine derived oxazolidines as secondary structure disrupting, solubilizing building blocks in peptide synthesis. *Tetrahedron Lett.* **33**, 1589–1592 (1992).

25. Wöhr, T. *et al.* Pseudo-prolines as a solubilizing, structure-disrupting protection technique in peptide synthesis. *J. Am. Chem. Soc.* **118**, 9218–9227 (1996).

26. Ullmann, V. *et al.* Convergent solid-phase synthesis of N-glycopeptides facilitated by pseudoprolines at consensus-sequence Ser/Thr residues. *Angew. Chemie - Int. Ed.* **51**, 11566–11570 (2012).

27. Abedini, A. & Raleigh, D. P. Incorporation of pseudoproline derivatives allows the facile synthesis of human IAPP, a highly amyloidogenic and aggregation-prone polypeptide. *Org. Lett.* **7**, 693–696 (2005).

28. Schnölzer, M. & Kent, S. B. H. Constructing proteins by dovetailing unprotected synthetic peptides: Backbone-engineered HIV protease. *Science (80-.).* **256**, 221–225 (1992).

29. Dawson, P. E., Muir, T. W., Clark-Lewis, I. & Kent, S. B. Synthesis of proteins by native chemical ligation. *Science* **266**, 776–9 (1994).

30. Crich, D. & Banerjee, A. Native Chemical Ligation at Phenylalanine. *J. Am. Chem. Soc* **129**, 10064–10065 (2007).

31. Haase, C., Rohde, H. & Seitz, O. Native chemical ligation at valine. *Angew. Chemie - Int. Ed.* **47**, 6807–6810 (2008).

32. Yang, R., Pasunooti, K. K., Li, F., Liu, X.-W. & Liu, C.-F. Dual native chemical ligation at lysine. *J. Am. Chem. Soc.* **131**, 13592–13593 (2009).

33. Boysen, R. I. & Hearn, M. T. W. HPLC of Peptides and Proteins. in *Current Protocols in Protein Science* 8.7.1-8.7.40 (John Wiley & Sons, Inc., 2001). doi:10.1002/0471140864.ps0807s23

34. Aguilar, M.-I. Basic Theory and Methodology. in *HPLC of Peptides and Proteins Methods and Protocols* 3–8 (Humana Press, 2004).

35. Irvine, G. B. Size-exclusion high-performance liquid chromatography of peptides: A review. *Anal. Chim. Acta* **352**, 387–397 (1997).

36. Alpert, A. J. & Andrews, P. C. Cation-exchange chromatography of peptides on poly(2-sulfoethyl aspartamide)-silica. *J. Chromatogr. A* **443**, 85–96 (1988).

37. Rivier, J., McClintock, R., Galyean, R. & Anderson, H. Reversed-phase high-performance liquid chromatography: preparative purification of synthetic peptides. *J. Chromatogr. A* **288**, 303–328 (1984).

38. Cunliffe, J. M. & Maloney, T. D. Fused-core particle technology as an alternative to sub-2-µm particles to achieve high separation efficiency with low backpressure. *J. Sep. Sci.* **30**, 3104–3109 (2007).

39. Destefano, J. J., Langlois, T. J. & Kirkland, J. J. Characteristics of superficially-porous silica particles for fast HPLC: some performance comparisons with sub-2-microm particles. *J. Chromatogr. Sci.* **46**, 254–60 (2008).

40. Kirkland, J. J., Schuster, S. A., Johnson, W. L. & Boyes, B. E. Fused-core particle technology in high-performance liquid chromatography: An overview. *J. Pharm. Anal.* **3**, 303–312 (2013).

41. Khalaf, R. *et al.* Doping reversed-phase media for improved peptide purification. *J. Chromatogr. A* **1397**, 11–18 (2015).

42. Herraiz, T. & Casal, V. Evaluation of solid-phase extraction procedures in peptide analysis. *J. Chromatogr. A* **708**, 209–221 (1995).

43. Filippov, D. V. *et al.* Use of benzyloxycarbonyl (Z)-based fluorophilic tagging reagents in the purification of synthetic peptides. *Tetrahedron Lett.* **43**, 7809–7812 (2002).

44. De Visser, P. C. *et al.* A novel, base-labile fluorous amine protecting group: Synthesis and use as a tag in the purification of synthetic peptides. *Tetrahedron Lett.* **44**, 9013–9016 (2003).

45. Zitterbart, R. & Seitz, O. Linker molecule and use thereof in methods for purifying peptides. (2017).

46. Knauer, S., Roese, T. M. L., Avrutina, O., Kolmar, H. & Uth, C. Method for peptide synthesis and apparatus for carrying out a method for solid phase synthesis of peptides. (2015).

47. Lawrenson, S. B., Arav, R. & North, M. The greening of peptide synthesis. *Green Chem.* **19**, 1685–1691 (2017).

48. Sheldon, R. A. The *E* factor 25 years on: The rise of green chemistry and sustainability. *Green Chem.* **19**, 18–43 (2017).

49. Chanthamath, S., Takaki, S., Shibatomi, K. & Iwasa, S. Highly stereoselective cyclopropanation of α,β-unsaturated carbonyl compounds with methyl (diazoacetoxy)acetate catalyzed by a chiral ruthenium(II) complex. *Angew. Chemie - Int. Ed.* **52**, 5818–5821 (2013).

50. Gholinejad, M., Dasvarz, N. & Nájera, C. Novel oxime-palladacycle supported on clay composite as an efficient heterogeneous catalyst for Sonogashira reaction. *Inorganica Chim. Acta* **483**, 262–270 (2018).

51. Iwakura, Y., Toda, F. & Suzuki, H. Synthesis of N-[1-(1-Substituted 2-Oxopropyl)]Acrylamides and -Methylacrylamides. Isolation and some Reactions of Intermediates of the Dakin-West Reaction. *J. Org. Chem.* **32**, 440–443 (1967).

52. Kaniewska, K., Romański, J. & Karbarz, M. Oxidation of ferrocenemethanol grafted to a hydrogel network through cysteine for triggering volume phase transition. *RSC Adv.* **3**, 23816–23823 (2013).

53. Hashidzume, A., Tanaka, A. & Sato, T. Interaction of poly(N-acryloyl-amino acids) with saccharides in aqueous media. *Polymer (Guildf).* **51**, 18–21 (2010).

54. Mottram, D. S. & Friedman, M. Symposium on the Chemistry and Toxicology of Acrylamide. *J. Agric. Food Chem.* **56**, 5983–5983 (2008).

55. Zamora, R., Delgado, R. M. & Hidalgo, F. J. Model Reactions of Acrylamide with Selected Amino Compounds. *J. Agric. Food Chem.* **58**, 1708–1713 (2010).

56. Adams, A., Hamdani, S., Lancker, F. Van, Méjri, S. & De Kimpe, N. Stability of acrylamide in model systems and its reactivity with selected nucleophiles. *Food Res. Int.* **43**, 1517–1522 (2010).

57. LoPachin, R. M. & Gavin, T. Reactions of electrophiles with nucleophilic thiolate sites: relevance to pathophysiological mechanisms and remediation. *Free Radic. Res.* **50**, 195–205 (2016).

58. Koshti, N. M., Jacobs, H. K., Martin, P. A., Smith, P. H. & Gopalan, A. S. Convenient Method for the Preparatian of Some Polyhydroxamic Acids: Michael Addition of Amines to Acrylohydroxamic Acid Derivatives. *Tetrahedron Lett.* **35**, 5157–5160 (1994).

59. Simonyan, G. S., Arutyunyan, R. S., Beileryan, N. M. & Grigoryan, E. A. Synthesis of New Surface-Active β-Amino Amides, Agents for Fixation of Dyes on Cotton Fibers, from N-

[Tri(hydroxymethyl)methyl]acrylamide. *Russ. J. Appl. Chem.* **76**, 258–260 (2003).

60. Han, C. *et al.* Synthesis and evaluation of 2-anilinopyrimidines bearing 3-aminopropamides as potential epidermal growth factor receptor inhibitors. *Eur. J. Med. Chem.* **77**, 75–83 (2014).

61. Zhong, Y., Xu, Y. & Anslyn, E. V. Studies of Reversible Conjugate Additions. *European J. Org. Chem.* **2013**, 5017–5021 (2013).

62. 王智刚郑成. Preparation method of acrylamide-based compounds. (2017).

63. Griffin, R. J. *et al.* Searching for Cyclin-Dependent Kinase Inhibitors Using a New Variant of the Cope Elimination. *J. Am. Chem. Soc.* **128**, 6012–6013 (2006).

64. Shiino, M., Watanabe, Y. & Umezawa, K. Synthesis of N-substituted N-nitrosohydroxylamines as inhibitors of Mushroom Tyrosinase. *Bioorg. Med. Chem.* **9**, 1233–1240 (2001).

65. Zhang, Y. Y. *et al.* Nitrosation of tryptophan residue(s) in serum albumin and model dipeptides. Biochemical characterization and bioactivity. *J. Biol. Chem.* **271**, 14271–9 (1996).

66. Brandhuber, B. J., Jiang, Y., Kolakowski, G. R. & Winski, S. L. PYRAZOLYL UREA, THIOUREA, GUANIDINE AND CYANOGUANIDINE COMPOUNDS AS TRKA KINASE INHIBITORS. (2014).

67. Oni, T. T., Okoro, U. C. & Ugwu, D. I. Synthesis of 1-azabenzo[a]phenoxazin-5-one and 11-amino-1,8,10-triazabenzo[a]phenoxazin-5-one and their Functionalized aryl derivatives via Mizoroki-Heck arylation methodology. *Orient. J. Chem.* **31**, 371–378 (2015).

68. Yamamoto, H. & Momiyama, N. Rich chemistry of nitroso compounds. *Chem. Commun.* 3514–3525 (2005). doi:10.1039/b503212c

69. Peterlin Masic, L. Role of Cyclic Tertiary Amine Bioactivation to Reactive Iminium Species: Structure Toxicity Relationship. *Curr. Drug Metab.* **12**, 35–50 (2011).

70. Ralhan, K., Guru Krishnakumar, V. & Gupta, S. Piperazine and DBU: a safer alternative for rapid and efficient Fmoc deprotection in solid phase peptide synthesis. *RSC Adv.* **5**, 104417–104425 (2015).

71. Takahashi, D. Method for removing FMOC group. (2014).

72. Stawikowski, M. & Fields, G. B. Introduction to Peptide Synthesis. in *Current protocols in protein science* **Chapter 18**, 1–17 (NIH Public Access, 2002).

3 Optimisation and application of acrylamide scavenging to the automated synthesis of complicated peptide drugs

This chapter explores the development of ‘acrylamide-tagging’ purification methods in an automated setting. It builds on the results of manual experiments from the previous chapter and aims to apply these findings to new synthesis method development for the automated microwave-assisted CEM Liberty Blue™ peptide synthesiser. Optimisation of the DABCO deprotection method (Chapter 2.9.2) is explored, followed by methods for ‘acrylamide-tagging’. These newly developed microwave synthesis cycles are then applied to synthesise several model peptides. After another round of optimisations, the new methods are put to the test in the synthesis and purification of several existing, complex peptide drugs.

Introduction

3.1 Green chemistry

Peptide synthesis by amide bond formation between partly protected amino acid residues has been considered among the most wasteful chemical processes. For each coupling step, at least one (N_α), protecting group is required, though often the amino acid side chain contains a protection group as well.[1] In typical procedures, coupling reagents (*e.g.* DIC and Oxyma) are also required for each coupling step, as well as a deprotection reagent (piperidine) upon completion of each step. Additionally, peptide synthesis is commonly carried out using large amounts of the toxic solvent DMF.[2]

Over recent years the concept of 'green chemistry', or sustainable chemistry, has gained ground in the chemical world. The importance of the topic is stressed by the dedication of four scientific journals: Green Chemistry (RSC), Green Chemistry Letters and Reviews (Taylor & Francis), ChemSusChem (Wiley), and ACS Sustainable Chemistry and Engineering (ACS). Green chemistry aims to reduce or eliminate the use and generation of hazardous substances, and is typically described by the following twelve principles:

1. Prevent waste
2. Atom economy
3. Less hazardous synthesis
4. Design benign chemicals
5. Benign solvents & auxiliaries
6. Design for energy efficiency
7. Use of renewable feedstocks
8. Reduce derivatives
9. Catalysis (vs stoichiometric)
10. Design for degradation
11. Real-time analysis for pollution prevention
12. Inherently benign chemistry for accident prevention

Design is considered an important aspect of green chemistry, as it includes novelty, planning and systematic conception. Careful planning is required to design chemical processes that have minimal environmental consequences, for which previously stated principles can be useful guidelines. This framework aims to be applicable to all aspects of the process life-cycle, including efficient use of raw materials, safe transformation, reduced toxicity, and biodegradability of reagents and products.[3]

The acrylamide-scavenging purification method does not address all twelve principles of green chemistry. The main issues that are being addressed are: **1)** the prevention of waste, **2)** solvent reduction and **3)** energy efficiency.

3.1.1 Waste prevention

The amount of waste generated during synthesis can be better understood by calculating the Environmental Impact Factor (E-Factor), which quantifies the amount of waste generated per kilogram of product.[4] An important factor that must be considered when calculating the *E* factor is defining raw materials or starting materials. For solid phase peptide synthesis, unprotected amino acids are the raw materials, while commercially available Fmoc-protected residues are the starting materials (in most cases). The pharmaceutical industry is associated with larger *E* values, due to the widespread use of organic solvents, and use of stoichiometric reagents rather than catalysts.[4] This is also the case for peptide chemistry, where large amounts of DMF are used for washing steps, and coupling reagents are added in excess for a single coupling step. A shortcoming of the *E* factor is that it does not take into account the environmental impact of the waste products but is solely mass-based. This has resulted into the introduction of additional measurements, including the environmental quotient (EQ),[5] Environmental Assessment Tool for Organic Synthesis (EATOS),[6] and EcoScale.[7]

3.1.1.1 The E-factor of a 7 amino acid peptide

To illustrate the scale at which waste is generated during standard solid phase peptide synthesis (using single couplings) the E-factor was calculated for a 7 amino acid peptide (**102**, NH_2-FIVALVF-$CONH_2$). Standard synthesis procedures require a large amount of reagents, including amino acids, deprotection reagents, activation reagents and solid support resin. This calculation does not take into account the amount of waste generated to obtain the starting materials. For the synthesis of 0.1 mmol of peptide **102** there is a requirement of ~160 mg of resin (0.63 mmol/g loading). All Fmoc-amino acids, prepared at 0.2 M stock solution in DMF, add up to a total of 1.236 g of amino acids and 17.5 mL of DMF.

DIC (1M in DMF; 7.35 mL) is used for the activation of each coupling step, which translates to 1.147 mL of DIC and 6.203 mL of DMF. The activator base Oxyma (1M in DMF; 3.68 mL) is required at slightly smaller quantities of 0.523 g, with 3.157 mL of DMF. Finally, a deprotection solution is required for the removal of Fmoc-protection groups after each coupling cycle. For this a 20% solution of piperidine in DMF is typically used. The Liberty Blue™ (CEM) requires slightly over 5 mL for each deprotection, bringing the total of deprotection solution to 35.21 mL (7.042 mL piperidine, 28.168 mL DMF) for this peptide (**102**).

Upon completion of the synthesis the peptide needs to be cleaved from the resin using a cleavage cocktail typically comprising TFA/TIPS/H_2O (9:0.5:0.5). If a total of 3 mL is used, another 4.289 g (4.023 g TFA, 0.116 g TIPS, 0.150 g H_2O) of reagents are consumed. This brings the total amount of reagents used (excluding solvent DMF) to **15.80 g**. However, another major source of waste during solid phase synthesis are the wash solvents, in the case of peptide synthesis often DMF. The Liberty Blue™ synthesiser uses a total of 222.21 mL of DMF for wash steps for peptide **102**. This brings the total of all DMF used during the synthesis to ~277 mL, or **260.60 g**. As mentioned earlier, these calculations have been made for the synthesis of peptide **102** (NH_2-FIVALVF-$CONH_2$), which has a molecular weight of 807.05. Assuming the synthesis was completed with a 100% efficiency of each coupling step, this leads to a yield of **0.081 g**. To calculate the E-factor the amount of waste generated is divided

by the yield. In the case of peptide **102** this leads to an E-factor of **3,424.81**. In comparison, production of the small molecule Viagra™ has an E-factor of 6.4.[4]

In other words, for each gram of peptide **102** that is synthesised, another 3424.81 grams of waste is generated. However, this number is only taking into account the synthesis of the crude product. In most cases further purification by prep-HPLC is required prior to obtaining the final peptide product. The amount of solvent required for purification is highly dependent on the quality of the crude peptide, as well as its chemical properties. For peptide **102**, a 25-minute run would most likely be sufficient to obtain adequate separation. Prep-HPLC makes use of larger columns than typically used for analytical HPLC and can withstand higher flow rates and loading amounts. A flow rate that can be used for the average prep column is 10 mL/min, which accounts for 250 mL (or roughly 250 g) of waste solvent per run in this example. The amount of peptide that can be recovered from prep-HPLC varies, so in this example we will assume one run yields 2.5 mg of purified peptide. From these numbers, an additional E-factor for the purification process can be calculated, resulting in an E-factor of **100,000**. While this number is only a crude estimation, it illustrates the need for a less wasteful peptide purification process.

To use the acrylamide-scavenging purification method, more waste is generated during the peptide synthesis stage. For each deprotection step, the standard 20% piperidine solution has been replaced by a 5% DABCO in DMF (*w/v*) solution, resulting in 1.6 g of DABCO waste and ~30.4 g of DMF waste. Additionally, after each amino acid coupling step, another acrylamide-tagging step has been introduced. This adds another 143 mg of reagent used per residue when acryloyl chloride (0.5 M) is used, adding up to a total of 1 g (or 1.12 mL) of waste for peptide **102** (and 26.88 mL of DMF). With these numbers the E-factor of the synthesis can be recalculated, bringing it to **3677.36**. While this number is slightly larger (252.55) than synthesis without acrylamide-tagging, significant waste prevention is expected to be achieved by preventing the prep-HPLC peptide purification process.

3.1.2 Atom economy

The atom economy represents the molecular weight of the product in comparison to the molecular weight of the reagents used.[8] Peptide chemistry typically has a low atom economy, due to the large number of protection groups used, as well as the use of stoichiometric coupling reagents. The atom efficiency slightly decreases for the acrylamide-tagging methods, compared to standard peptide synthesis, due to the use of additional reagents. The reduction of solvent usage is not taken into account.

3.1.3 Solvents

Making use of less hazardous solvents forms an important area of focus within sustainable chemistry in general [9], as well as peptide synthesis specifically.[10] The commonly used solvent during synthesis, DMF, is known to be reprotoxic.[11] A lot of research in the 'greening' of peptide synthesis focusses on replacing DMF by an environmentally friendlier alternative. For the purification process by HPLC relatively benign solvents as water and methanol or acetonitrile are used. Water is generally considered to be safe and not impose any hazards.[12,13] Considering this, using acrylamide-scavenging does not make a large contribution to greening peptide synthesis, as it does not replace or eliminate the need for a toxic solvent (DMF). However, once water has been used for prep-HPLC it has been contaminated with chemicals (*e.g.* formic acid or trifluoroacetic acid) and has to be treated as chemical waste, of which the disposal can still lead to high costs. Therefore, despite using solvents that are considered less hazardous, HPLC should still be avoided if possible.

3.1.4 Energy

Little attention is given to the consumption of electricity during peptide synthesis and purification. Preparative HPLC requires multiple instruments to be running for a prolonged period of time, especially when peptides are being produced and purified on a larger scale. Each run will require multiple compartments to be switched on, including a pump, UV detector, fraction collector, autosampler, degasser, and in some cases mass detector. In comparison, the acrylamide scavenging method only requires a shaker to keep the thiol scavenging-resin well suspended in the crude peptide solution. Additionally, scavenging can be achieved over

a shorter period of time compared to multiple HPLC runs. Finally, after HPLC purification the purified peptide remains dissolved in a large amount of solvent, from which it needs to be concentrated. This can be achieved by evaporation and lyophilisation. While lyophilisation is still required for the acrylamide-scavenging procedure, the solvent quantity is much lower, minimising energy usage.

3.1.5 Renewable materials

Acrylamide-scavenging requires a solid support thiol-resin to remove “tagged” peptide-impurities from solution. As a Michael addition is reversible,[14] there is potential to regenerate the thiol-resin. This possibility could be further explored in future experiments to optimise sustainability of the acrylamide-scavenging procedure. It is of importance that regenerating the thiol-resin would not require more reagents (and thus waste) than are being saved.

3.2 Application of acrylamide-scavenging to automated peptide synthesis

3.2.1 Aim of this chapter

In the previous chapter manual procedures were developed to introduce an acrylamide-tag on peptide impurities. It was demonstrated that a single (tagged) deletion could be scavenged selectively from a peptide mixture. In this chapter, previously obtained knowledge is applied to the development of automated solid phase peptide synthesis methods.

3.2.2 Approach

The effectiveness of the acrylamide-scavenging procedure is first demonstrated using controlled experiments where an attempt has been made to remove multiple deletions from a peptide mixture. An alternative scavenging reaction using a photocatalyst was also explored. Method development was undertaken to use DABCO for Fmoc-removal using a Liberty Blue™ automated synthesiser. Next, method development was undertaken for acrylamide-‘tagging’ using automated synthesis methods. Finally, the acrylamide-scavenging synthesis and purification method was used for the synthesis of model peptides and peptide drugs to demonstrate the efficiency.

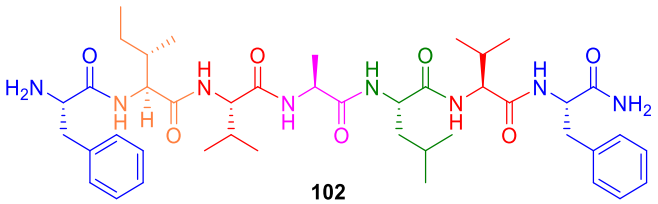
Results and discussion

3.3 Modelling of optimised approach: Separation of multiple deletions from desired peptide product

In the previous chapter it was demonstrated that a single impurity could be scavenged from a peptide mixture. However, this is not a good representation of a complete synthesis, as it is likely multiple deletions would occur in longer peptides. To demonstrate the wide applicability of the acrylamide-scavenging purification two model peptides were synthesised and multiple deletions were introduced in a controlled fashion.

3.3.1 Non-nucleophilic scavenging

For the first example, a peptide was designed that consisted of only non-nucleophilic residues. This was considered to be a good starting point, as no side reactions or possible cyclisations were to be expected. The full peptide sequence (**102**, NH_2-FIVALVF-$CONH_2$) was synthesised. Each "deletion" was then synthesised individually and acrylamide-tagged, with the exception of a single C-terminal phenylalanine and valine-phenylalanine dipeptide (**103**, Acr-IVALVF-$CONH_2$; **104**, Acr-VALVF-$CONH_2$; **105**, Acr-ALVF-$CONH_2$; **106**, Acr-LVF-$CONH_2$). A peptide mixture was prepared in water/MeOH (10 mL; 50:50), consisting of 5 mg peptide **102** and ~1 mg of each of the "deletions".



102

Figure 3.1: Structure of peptide **102**, with each amino acid residue displayed in a different colour

The peptide mixture was incubated with 1 g (1.3 mmol) of thiol-resin at room temperature. Each 30 minutes a sample was collected for direct LC-MS analysis. A 'Find By Formula' scan was performed for each of the expected peptides, after which the area under the curve (AUC) was used to calculate the composition of the mixture. No internal standard was used as

variations from *e.g.* injection inaccuracies were not expected to influence the composition of the peptide mixture.

As expected, the mixture at the start of the scavenging procedure consisted of about 45% of desired peptide (**102**). While each of the deletions was added at an amount of ~1 mg, slight variations in composition percentages were observed, with peptide **105** making up ~25% of the mixture, while peptide **103**, **104** and **106** only made up ~10% each. These differences can be explained by differences in molecular weight as the same amount in mg will represent a different amount in mmol. Another explanation of the starting variation of each deletion is manual error of weighing out very small quantities. After one hour (Figure 3.2) an increase of the percentage of peptide **102** can be observed, suggesting successful scavenging. Over time, peptide **104**, **105**, and **106** appear to have been successfully removed from the peptide mixture, or only a very small amount is detected. Peptide **103** appears to be increasing over time, before decreasing after 300 minutes (5 hours). This rise can be explained as the graph does not display the absolute amount of each peptide present, but a relative amount where the peptides are compared to each other. Additionally, the Michael addition is a reversible reaction, so it is possible some of the scavenged peptide is released back into solution over time. Overall, the purity of peptide **102** had improved, increasing from ~45% to ~80%.

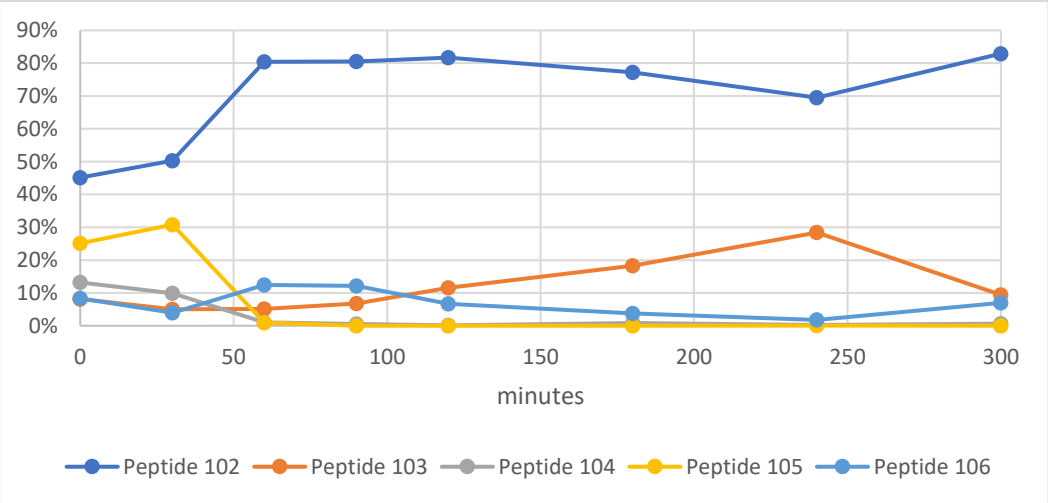


Figure 3.2: Overview of the consistency of the peptide mixture over time during the scavenging process.

3.3.2 Scavenging of a representative model peptide

While peptide **102** formed a good initial model for the acrylamide-scavenging of multiple deletions, it is not representative for the average peptide, as it does not contain any nucleophilic, acidic, or basic residues. This led to the design of peptide **107** (Figure 3.3; NH_2-TWEKMPY-$CONH_2$), which has an equal distribution of acidic and basic residues. Additionally, it contained a relatively difficult to couple residue like tryptophan.

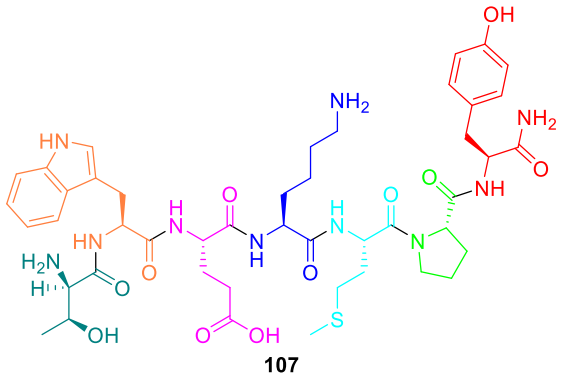


Figure 3.3: Structure of peptide **107**, with each amino acid residue displayed in a different colour

Peptide **107** and each of the deletions (**108**, Acr-WEKMPY-$CONH_2$; **109**, Acr-EKMPY-$CONH_2$; **110**, Acr-KMPY-$CONH_2$; **111**, Acr-MPY-$CONH_2$) were synthesised, capped, cleaved from the resin and lyophilised. Prior to mixing the peptides for the purpose of scavenging, each peptide was analysed by LC-MS to confirm purity. Interestingly, the acrylamide-tagging was not equally efficient for each "deletion". In future experiments, it would be beneficial to synthesise a short model peptide, with each of the 20 common amino acids at the N-terminal (*e.g.* NH_2-XAA-$CONH_2$, where X = any amino acid). This would help understand the acrylamide-coupling efficiency for each residue and whether differences occur.

For the purpose of acrylamide-scavenging in this study, three peptides were used in the mixture; peptide **107** as desired peptide (~8 mg) and peptide **109** and **110** as deletions (~4.5 and ~5.5 mg, respectively). Peptide **108** and **111** were excluded as they were not isolated as pure peptides prior to preparation of the peptide mixture. Unlike the previous experiment, the

composition of each deletion was not factored in to be equal, but it was random to represent realistic synthesis conditions. As for the experiment with peptide **102**, the peptides were dissolved in water/MeOH (10 mL; 50:50) and 1 g of thiol-resin was used for scavenging. HPLC analysis was performed prior to the scavenging procedure, clearly indicating the presence of multiple impurities (Figure 3.4.A). The peptide mixture was shaken in presence of the resin for four hours, after which the composition of the mixture was analysed by HPLC (Figure 3.4.B). This shows great improvements in the purity compared to the mixture prior to scavenging, suggesting the effectiveness of the new peptide purification method.

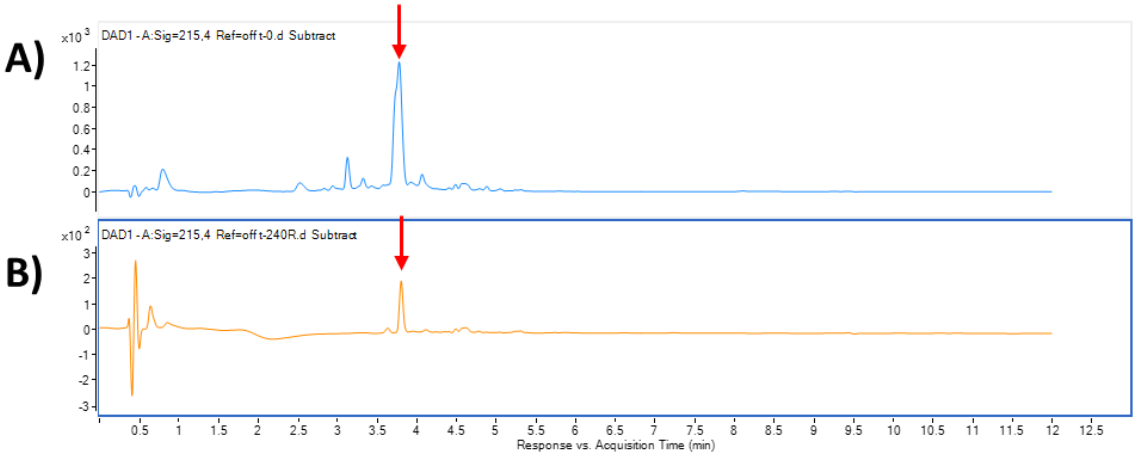
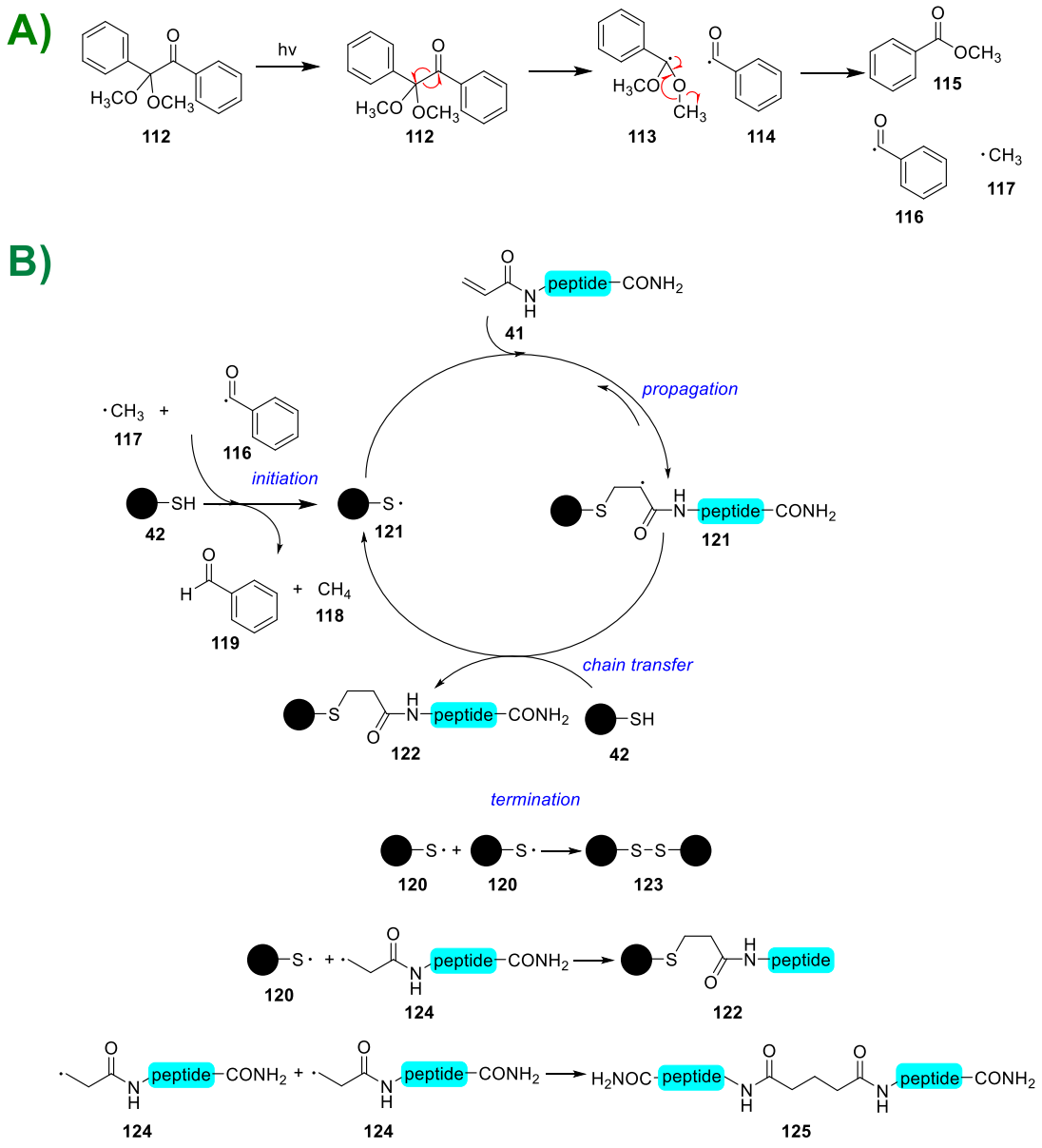


Figure 3.4: HPLC chromatogram (215 nm) of the peptide mixture before (A) and after (B) 4 hours of acrylamide-scavenging. The target peptide (**107**) is indicated by the red arrow.

3.4 Optimisation of scavenging reaction

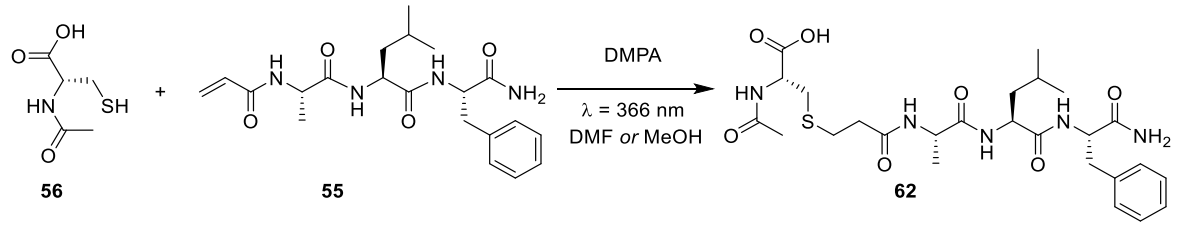
While scavenging impurities following conjugate addition has been proven possible, alternative scavenging methods remain of interest, to explore possibilities of faster reaction times, or enhanced selectivity. As such, the photocatalysed intermolecular thiol-ene reaction, also known as hydrothiolation, between the resin-thiol and peptide-acrylamide was investigated as an alternative conjugation method. This reaction was hypothesised to be more selective, preventing intramolecular cyclisation, or dimerisation. Thiol-ene or thiol-yne reactions have been used before for peptide cyclisation,[15,16] making use of 2,2-dimethoxy-2-phenylacetophenone (DMPA, **113**) as a photocatalyst (Scheme 3.1.A). The cyclisation

reactions are further described in Section 4.1. Making use of molecular photoinitiators (PI) is more efficient than direct photoinitiation and generation of thiyl radicals. The photocatalysed reaction does not require the use of a conventional conjugate acceptor and can work on any unactivated alkene or alkyne. In general there are two classes of PI, the cleavage photoinitiators, which DMPA is a part of, and hydrogen-transfer photoinitiators.[17]



Scheme 3.1: A) Radical generation of photo initiator DMPA. B) Scavenging cycle.

Radical generation from DMPA occurs after irradiation with UV light at 366 nm, yielding methyl and benzoyl radicals.[18] The thiol-ene reaction is initiated by these radicals, generating a highly reactive thiyl radical ●-S· (**120**) that can attack the peptidyl-acrylamide (**41**). After the propagation step, the newly generated ●-SCH_2C·HCONH-peptide (**121**) can undergo a chain transfer (Scheme 3.1), regenerating the thiyl radical (**120**), which can undergo another thiol-ene cycle.



56 **55** **62**

Scheme 3.2: Photocatalysed reaction between peptide **55** and NAC.

To model this approach, the reaction between NAC and peptide **55** (Acr-ALF-$CONH_2$) was performed to gain insight in the reactivity (Scheme 3.2). Two solvent options compatible with peptides were explored; methanol and DMF. Peptide **55**, N-acetyl-L-cysteine (40 eq.), and DMPA (20 eq.) were added to a glass vial. Solvent (1 mL, degassed DMF or methanol) was added and the mixture was gently shaken until all solids were dissolved. The reaction mixture was exposed to 366 nm UV light inside a reflective box for 1 hour. Appropriate work up conditions were explored.

The organic components generated upon initiation of the thiol-resin (benzaldehyde) are soluble in diethyl ether and hexane. Therefore, the following conditions were explored:

- **A.** Dilution in water, after which the organic components are removed by liquid extraction using diethyl ether or hexane. Lyophilisation of the remaining aqueous phase.
- **B.** Concentration of reaction mixture (removal of DMF) by nitrogen air stream, followed by procedure A.
- **C.** Addition of water to reaction mixture and direct lyophilisation.

No differences were observed between the different work up conditions. There was no reaction observed between peptide **55** and NAC when the reaction components were dissolved in MeOH. This suggests that degassing could be important for the reactivity, though this has not been confirmed by attempting the reaction in non-degassed DMF. A reaction between peptide **55** and NAC was observed when degassed DMF was used as solvent, however, the observed molecular weight (by LC-MS) analysis did not correspond with the expected product. The calculated molecular weight for peptide **62** is 565.69. During LC-MS analysis, the molecular weight that was observed (582.26 $[M+H]^+$) corresponded to an unexpected addition of a hydroxyl-group. It would be of interest to undertake additional model reactions between NAC and acrylamide-tagged peptide sequences, to gain understanding in whether this is a residue specific phenomenon, or if reaction takes place at the acrylamide-group.

To confirm that the reaction took place *via* a photo-catalysed radical thiol-ene reaction and not a Michael addition, reaction mixture in the dark. As expected, no reaction was observed, confirming the photocatalysed reaction mechanism.

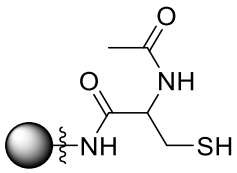
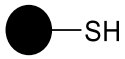
3.5 Resin properties

Thiol resins are sometimes used for the scavenging of palladium from a reaction mixture. Our experiments utilised a commercial resin (ISOLUTE® Si-Thiol Biotage) for the purpose of acrylamide scavenging. While this silica-based resin may be commercially available, it might not have the best properties for compatibility with peptides and peptide-compatible solvents.

A polymer resin may contain better properties for peptidyl-acrylamide scavenging. The porous, larger resin could provide easier access to thiol-binding sites for the acrylamide-tagged peptides. To test this hypothesis, ProTide Rink Amide resin (0.63 mmol/g loading, CEM™) was used to prepare an alternative thiol-resin (**126**) (Table 3.1). The resin was prepared by coupling one residue of Cys(Mmt), which was subsequently deprotected during short cleavage cycles (5x 2 min; 1% TFA, 5% TIPS, DCM). Using a low percentage of TFA ensured that

cysteine would not be cleaved from the resin, while it is sufficient for the removal of the Mmt protection group.[19]

Table 3.1: Overview of the different properties of a polymer-based and a silica-based thiol resin.

Properties	Polymer resin (126)	Silica resin (42)
Structure		
Loading	0.63 mmol/g	1.3 mmol/g
Swelling	Yes	No
Particle size	100-200 mesh (149-74 µm)	60 microns (µm)
Core	Porous	Solid

A mixture of peptide **44** (Ac-SEPAKY-$CONH_2$) and peptide **75** (Acr-EPAKY-$CONH_2$) in H_2O/MeOH (10 mL, 1:1) was prepared. The composition of the mixture was analysed by LC-MS, after which it was added to the polymer thiol-resin. Similar to the experiment in Section 2.11, 0.65 mmol of SH-binding sites were used, corresponding to ~1000 mg of resin. DIPEA (450 µL; 2.6 mmol) was added to catalyse the reaction. The reaction was monitored over time, with samples taken after 30, 60, 90, 120, 180, and 240 minutes for LC-MS analysis. Surprisingly, there was no significant decrease observed in the relative amount of peptide **75**, compared to peptide **44** (Figure 3.5). It is possible that the deprotection of the Mmt-group was not completely successful, which could lead to fewer thiol-binding sites being available for scavenging than expected or calculated. Alternatively, the coupling of cysteine to the resin could have been incomplete. Finally, it is possible that the binding sites in the polymer resin were not exposed enough for an effective reaction between the thiol and acrylamide. Based on these results, no further investigations into alternative thiol-resins were undertaken. The silica-based thiol-resin was used for all further experiments.

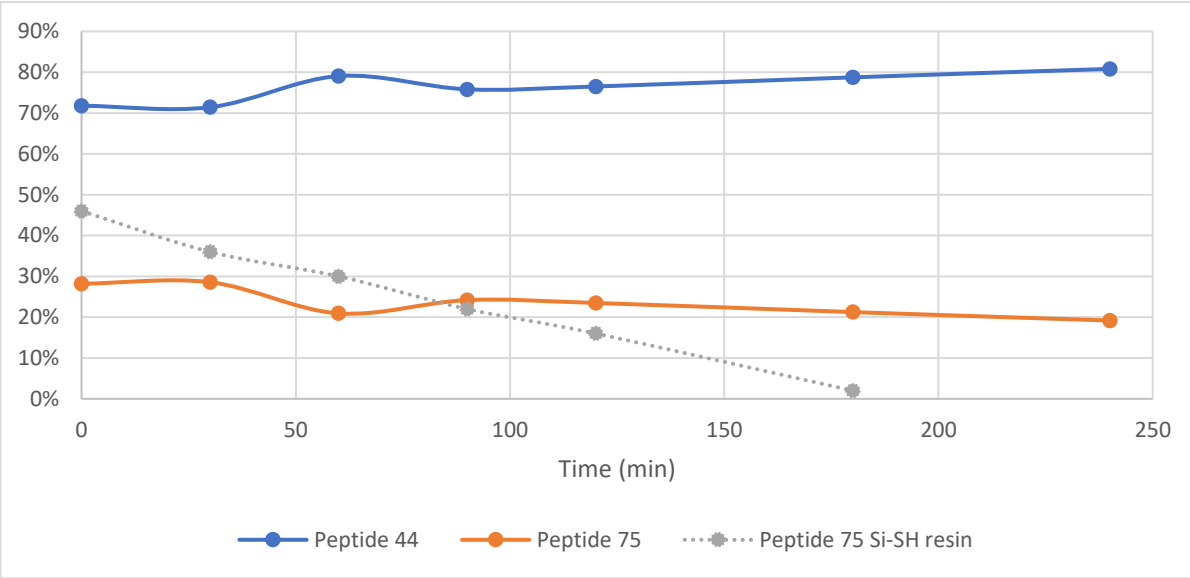


Figure 3.5: Relative amount of peptide **44** and **75** over time. The removal of peptide **75** using similar quantities of Si-SH resin is displayed in grey for comparison.

3.6 Automated peptide synthesised using DABCO

In traditional automated solid phase peptide synthesis, a solution of 20% piperidine in DMF is used for deprotection of the Fmoc-group after successful amino acid coupling. As explored in the previous chapter (Section 2.9.2), deprotection using DABCO as base works equally well. To confirm compatibility with automated microwave-assisted peptide synthesis, a short six amino acid peptide (peptide **127**, NH_2-ARQALF-$CONH_2$) was synthesised. This sequence was designed as it was not expected to take part in intramolecular reactions, yet it still contained challenging amino acid residues that would make the sequence representative for “real” peptide synthesis conditions. Prior to synthesis, all lines of the CEM Liberty Blue™ peptide synthesiser were back flushed and purged, to prevent contamination with piperidine. No instrumental complications occurred during the synthesis.

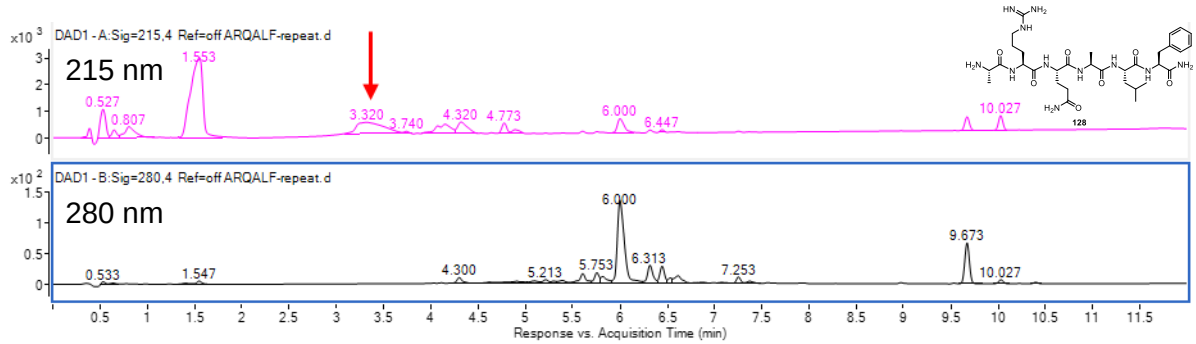


Figure 3.6: LC-MS results of initial peptide synthesis of peptide **127** using DABCO for Fmoc-deprotection. The peptide of interest is indicated by the red arrow.

An initial attempt to synthesise peptide **127** using DABCO (5% *w/v*) for deprotections on the automated peptide synthesiser did not result in satisfactory pure peptide (Figure 3.6). This could be the consequence of a scale-up miss-match; during manual testing conditions 1 mL of deprotection solution was used for ~0.01375 mmol of resin binding capacity (25 mg, 0.55 mmol/g loading). The standard deprotection method of the peptide synthesiser uses 4 mL of deprotection solution, while the standard peptide synthesis scale is 0.1 mmol (~180 mg, 0.55 mmol/g loading). In other words, during automated peptide synthesis 1 mL of deprotection solution is used for the deprotection of 0.025 mmol of peptide, which is ~two-fold higher compared to the manual test amount that was used. This, therefore, required some optimisation in order to translate into automated, full-scale synthesis.

The initial results lead to four potential options to improve the quality of the synthesis:

A. Perform the deprotection step twice (4 mL, 5%, 1.5 min, repeat)

B. Double the reaction time (3 min instead of 1.5 min)

C. Double the amount of liquid added to reaction vessel (8 mL instead of 4 mL)

D. Double the % of DABCO used (10% solution instead of 5% solution)

Options A-C were tested in the peptide synthesiser. Using 10% DABCO instead of 5% (option D) was not explored, as high concentrations of DABCO in DMF results in solubility issues. The results of the other deprotection reaction conditions are presented in Figure 3.7.

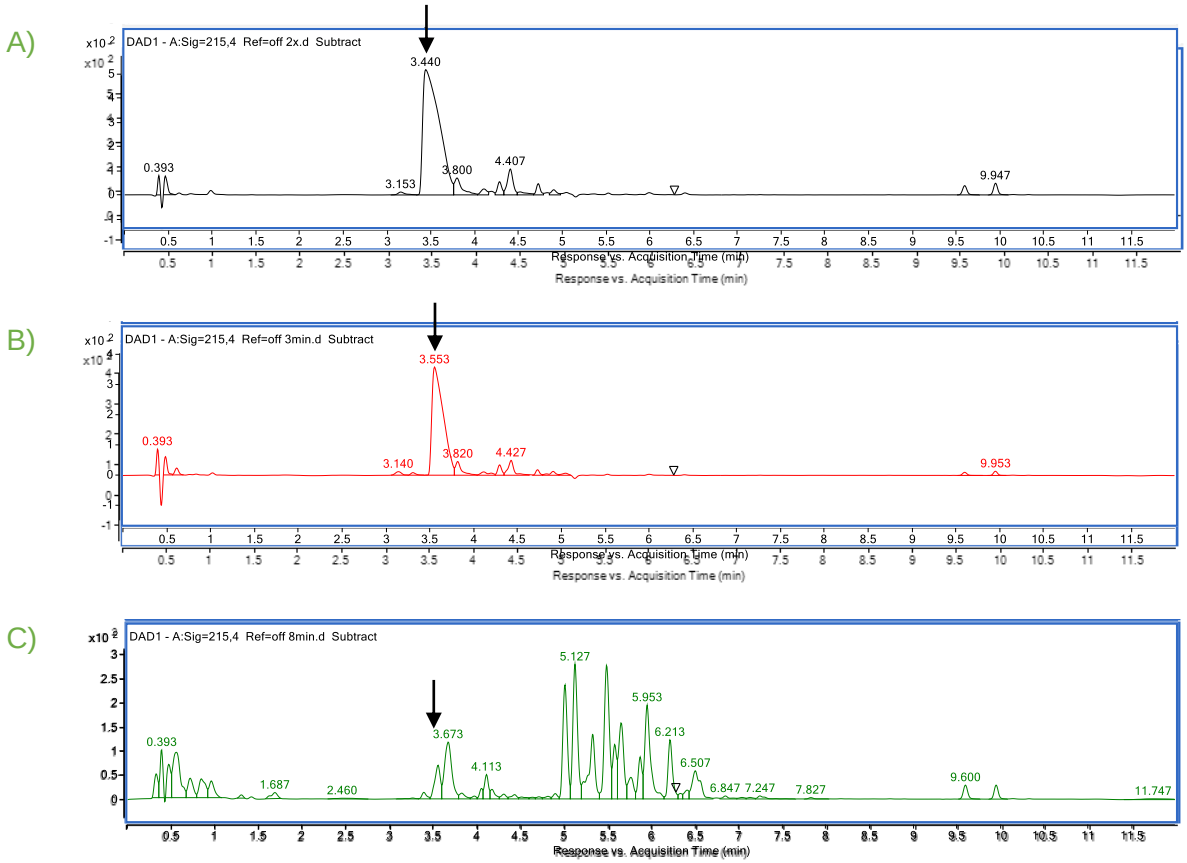


Figure 3.7: Optimisation of DABCO deprotection step. Peptide of interest (peptide **127**) is marked with an arrow. Method A) double deprotection; Method B) 3 minutes deprotection; Method C) 8 mL deprotection volume.

Surprisingly, doubling the amount of deprotection reagent leads to a drastic decrease in purity of synthesised peptide. Despite the reaction vessel being able to contain up to 30 mL of liquid, the decrease in purity could be explained if insufficient drainage occurs after the deprotection step. The double deprotection and double reaction time conditions showed significantly improved results. When using a 3 minutes deprotection, a crude purity of ~84% was achieved. These results (5% DABCO in DMF, 3 min deprotection) are comparable with piperidine deprotection and was considered sufficient to continue future experiment with.

3.7 Automated peptide synthesised using acryloyl chloride

During normal synthesis acryloyl chloride will only be used to react with residual free amine resulting from incomplete amino acid coupling steps. Ideally, this reaction would never take place, as preferably there would be a 100% coupling efficiency. However, to ensure the effectiveness of the acryloyl chloride coupling, the reaction was first carried out on a peptide containing a free terminal-NH_2 before applying it to regular peptide synthesis.

In Section 2.7 a 0.5 M solution of acryloyl chloride in DMF (40 µL/mL; 1 mL) was used for successful acrylamide-tagging at a ~0.01375 mmol scale. Using the peptide synthesiser at 0.1 mmol scale resulted in a scale increase of 7.27-fold compared to the manual experimental conditions. Standard synthesis amino acid couplings use a total volume of 4 mL per coupling, which was used as a starting point for the development of the acrylamide-tagging reaction conditions. A preliminary experiment was carried out where a 5 amino acid peptide (Acr-RQALF-$CONH_2$, peptide **128**) was synthesised as normal (standard piperidine SPPS), with a final acrylamide-tagging step upon completion of the synthesis. Amino acid coupling microwave methods (one cycle subsequently: Cycle **A**; 75°C, 5 sec.; 80°C, 20 sec.; 86°C, 8 sec.; 90°C, 120 sec.) or conventional coupling conditions (Cycle **B**; 25°C, 300 sec.) were used.

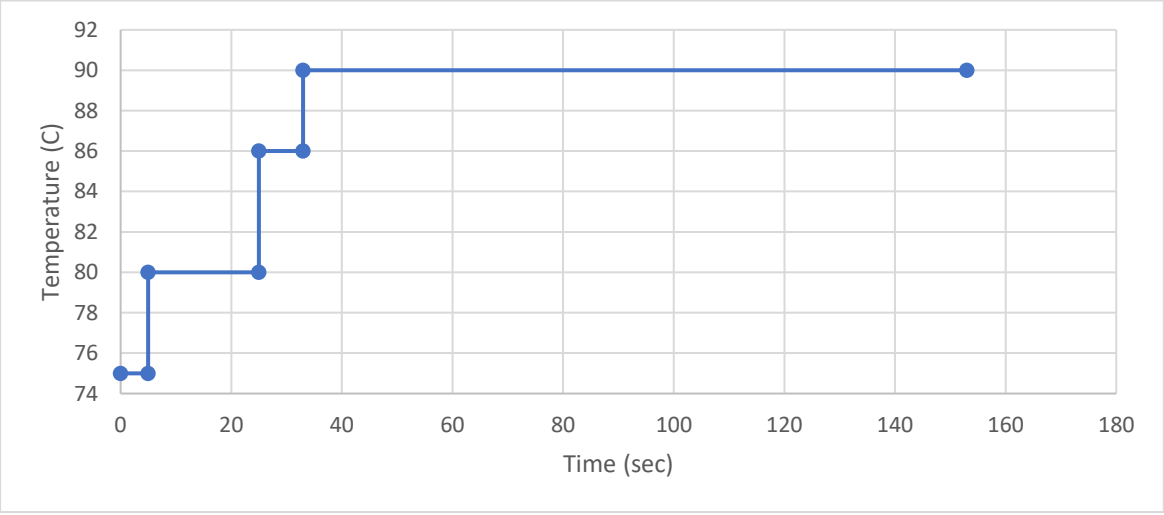


Figure 3.8: Visual overview of temperature changes following microwave cycle **A**.

No abnormal changes in pressure were observed and synthesis proceeded as normal. Based on previous manual experiments a completion of >50% of the reaction would be expected under microwave conditions. Theoretically, to reach completion of the reaction, 8 mL of acryloyl chloride solution should be added to the reaction vessel. Nevertheless, it is unlikely such reaction conditions are required, as only a small percentage of the resin should contain free amine resulting from incomplete amino acid coupling in normal situations.

Under both conditions two products were formed, where the main product was identified as the desired acrylamide-peptide. The other molecular weights found (m/z = 688; 1375) likely corresponds to dimer formation (Figure 3.9). The amount of dimer formed was significantly larger for the conventional coupling conditions (~32%), compared to the microwave conditions (~16%). This suggests the reaction time could be an important factor. Subsequently, two additional microwave cycles were programmed with shorter reaction times.

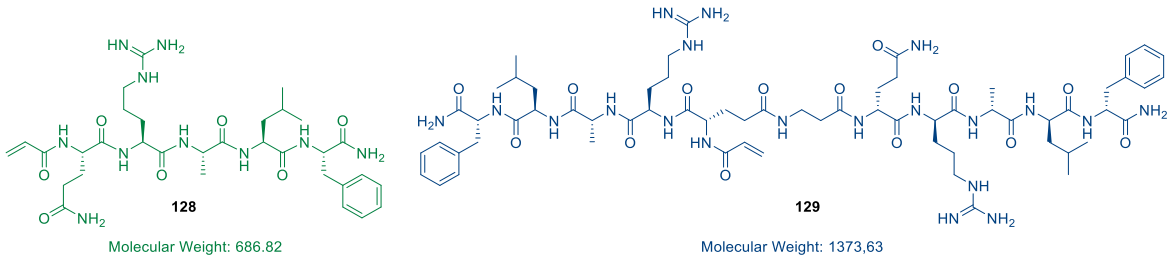


Figure 3.9: Structures of target peptide **128** and potential dimer **129**.

The first microwave cycle had a faster build up to the final temperature (Cycle **C**; 75°C, 5 sec.; 80°C, 5 sec.; 86°C, 5 sec.; 90°C, 45 sec.) and the total reaction time comprised 1 minute. A second, shorter microwave heating method was also programmed (Cycle **D**; 75°C, 5 sec.; 80°C, 5 sec.; 86°C, 5 sec.; 90°C, 15 sec.), where the acrylamide-tagging cycle was conducted twice, with a wash step in between. The second method was employed to investigate whether a side-product (*e.g.* HCl) was formed during the reaction, which could catalyse the dimerisation process.

Surprisingly, the shorter microwave cycles did not lead to an improvement in single product formation. An overview of product- and dimer-formation for all tested microwave cycles can

be found in Figure 3.10. The lack of improvement for the shorter microwave cycles can potentially be explained by the faster temperature build-up. This can result in temperature overshoots which might promote dimerisation.

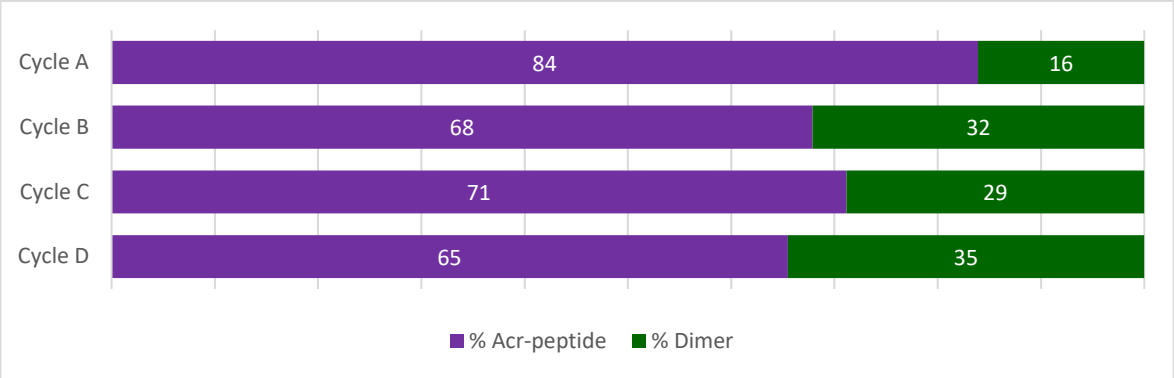


Figure 3.10: Comparison of formation of desired Acr-peptide product versus dimerisation under different microwave cycle conditions.

The influence of temperature could be investigated by conducting additional experiments with either a faster, or slower temperature build-up. However, the results of the standard (153 sec.) microwave cycle was considered sufficient to continue future experiments with. Under expected reaction conditions most of the peptide would have an Fmoc-amino acid at the end of the peptide chain, rather than a free amine resulting from incomplete couplings. Additionally, it may be unlikely that two unreacted peptide chains are in close enough proximity to form dimers (Figure 3.11). Therefore, before continuing to optimise the acrylamide-tagging step, a full peptide synthesis using DABCO deprotections and acrylamide-tagging was performed.

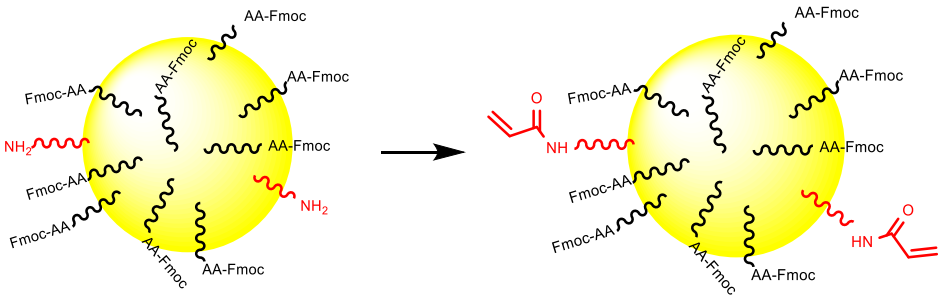
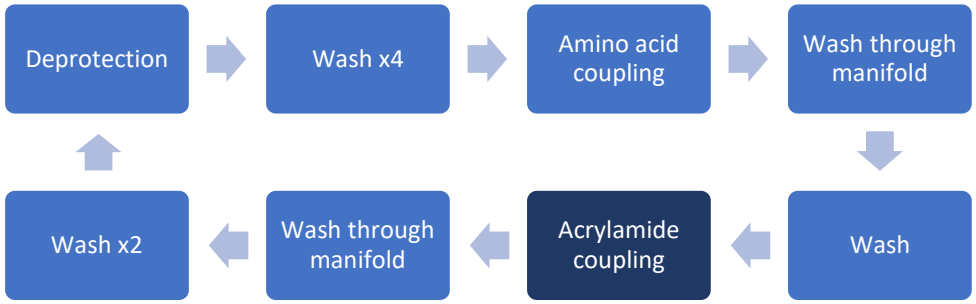


Figure 3.11: Illustration of peptide chains on resin during acrylamide-tagging step.

3.8 Automated peptide synthesis with DABCO and acrylamide-tagging

Before applying the acrylamide-tagging technique to full length drug peptides, a short model peptide **127** (NH_2-ARQALF-$CONH_2$) consistent with previous experiments was synthesised. DABCO was used for deprotections, using a 5% concentration and 3 minutes reaction times. Acrylamide-tagging was performed using the standard (153 sec.) microwave cycle (Cycle A), as described previously. An overview of the microwave cycles and an overview of the full synthesis, and possible tagged side-products, are displayed in Figure 3.12.

A)



B)

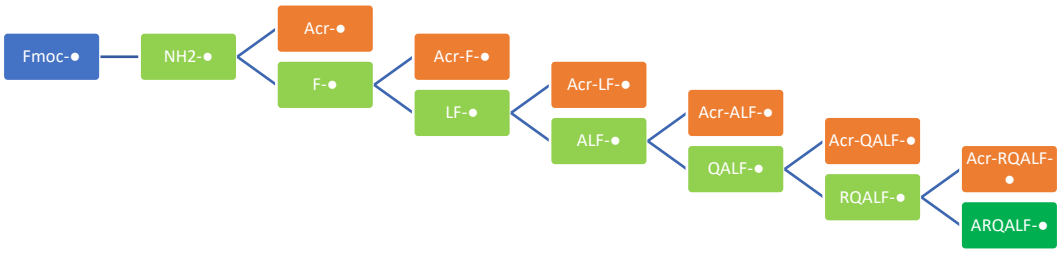


Figure 3.12: Schematic overview of full peptide synthesis procedure. A) Microwave cycles that were used; B) Full pathway with potential acrylamide-tagged side products.

No complications were observed in the automatic peptide synthesiser during synthesis. However, it must be noted that at the end of the synthesis the reaction vessel had partly turned yellow. This could be a consequence from HCl forming during the acrylamide coupling step. This was not observed during initial 'final' acrylamide-tagging studies, which could indicate incomplete removal from the reaction vessel. To overcome this problem additional washing steps could be inserted after the acrylamide-tagging step. Alternatively, delayed addition of a small amount of base could quench the reactivity, and thus improve the quality of synthesis by eliminating the formation of side products.

A first attempt to synthesise peptide **127** using this method resulted in a range of products. The major deletions that were present were arginine (R) and its neighbouring amino acid, the N-terminal alanine. This was to be expected due to the bulky properties of the Pbf protecting group on the arginine side chain. The quality of the synthesis was poorer than expected, however, this provided an excellent opportunity to test the thiol-scavenging procedures. Insufficient drainage and residual HCl could have influenced the synthesis procedure. Despite the wide variation of products present, all except the final product, were acrylamide-tagged.

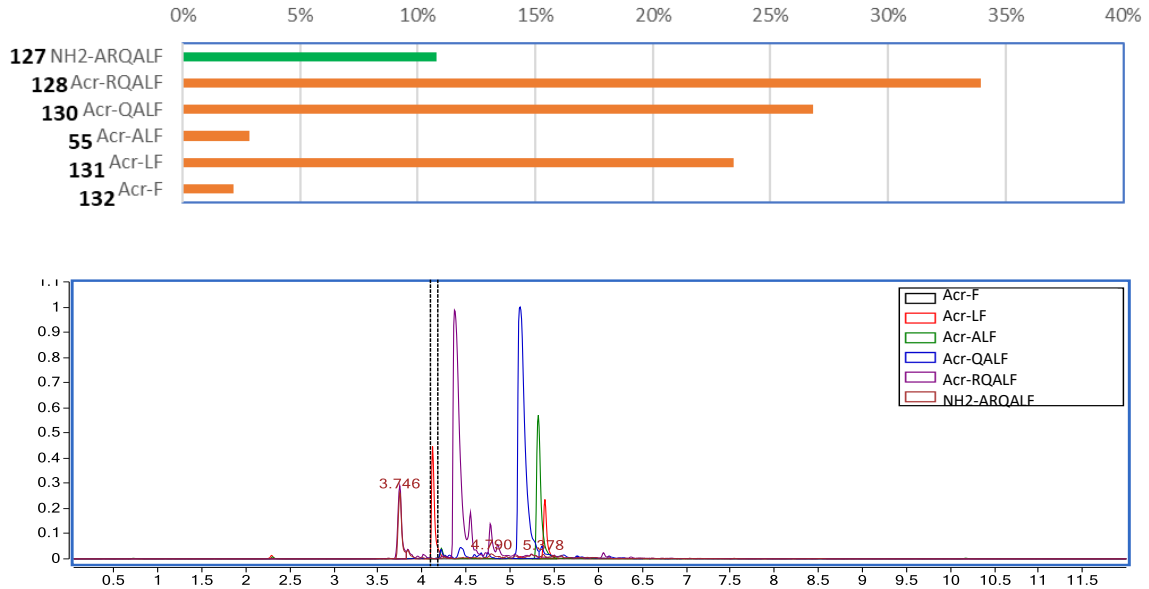
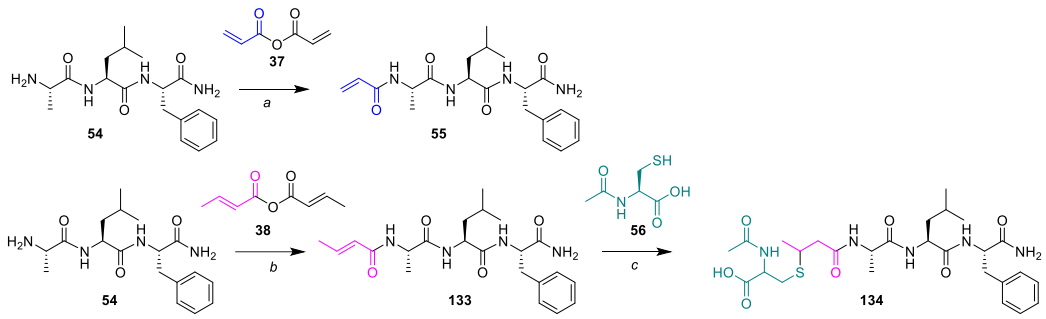


Figure 3.13: Product distribution overview of first synthesis of peptide **127** using DABCO and acrylamide-tagging methods.

3.9 Anhydride approach

While the use of acryloyl chloride was sufficient during manual peptide synthesis, its use was more complicated in the automated peptide synthesiser. Despite the inclusion of multiple washing steps, it is likely the formation of hydrochloric acid was affecting the synthesis. This could be observed by a change of colour of the reaction vessel from transparent to pale yellow. Additionally, acryloyl chloride is volatile and toxic, making it a bad combination to work with on larger scale and prolonged duration. An alternative, milder, approach to introduce the acrylamide was required. This led us to explore the use of acrylic anhydride, a direct analogue of acetic anhydride. Additionally, due to commercial availability, we explored crotonic anhydride, which contains a terminal methyl-group, thus reducing reactivity of the alkene (Scheme 3.3).



Scheme 3.3: Introduction of acrylamide groups with a) acrylic anhydride and b) crotonic anhydride.

Introduction of the methyl-acrylamide group was attempted following similar reaction conditions that were found to be successful for acryloyl chloride based synthesis. A three amino acid peptide was synthesised on-resin (25 mg, 0.68 mmol/g loading) and was treated with crotonic anhydride or acrylic anhydride under microwave (90°C, 1.5 minutes) or conventional conditions (2x15 minutes, room temperature). In each case, a full, clean conversion to the (methyl-)acrylamide peptide was observed.

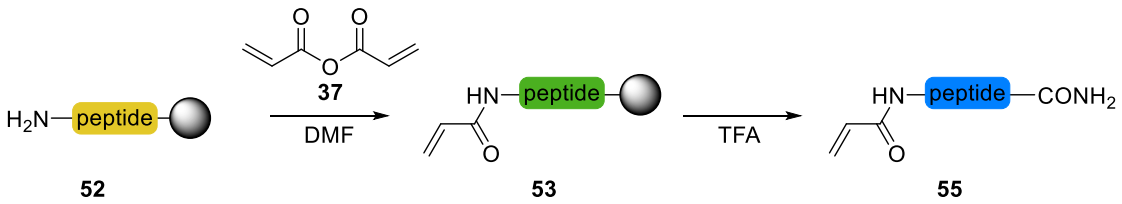
To further investigate the effect of the methyl group, peptide **134** was incubated with N-acetyl-L-cysteine for 15 minutes at 75°C in the microwave. No conversion to the NAC-peptide conjugate was observed, possibly due to steric hindrance or inductive electron-donating effects of the methyl group.

3.10 Automated peptide synthesis with acrylic anhydride

Introducing an N-terminal acrylamide group using acryloyl chloride did not provide the desired clean results, as some potential dimerisation was observed. Following these results, manual introduction of the acrylamide group has been performed using acrylic anhydride, which provided promising results. To ensure sufficient reactivity in the automated synthesiser, a short three amino acid model peptide **55** (Acr-ALF-$CONH_2$, 0.1 mmol scale) was synthesised.

The concentrations of acrylic anhydride in DMF used were based on the required concentrations as previously determined in manual coupling experiments of acryloyl chloride and acrylic anhydride. An initial solution was prepared of 5% acrylic anhydride in DMF (0.25 M), which was diluted to 2.5% (0.125 M) and 1% (0.05 M) solutions for further experiments (Table 3.2).

Table 3.2: Optimisation of acrylic anhydride reaction on tripeptide **52** (NH_2-ALF-●) (160 mg, 0.1 mmol), conversion yields are based on UV absorbance at 215 nm.



Acrylic anhydride in DMF concentration (molar)	Reaction time	Temperature	Conversion yield (%)
0.25 M	1.5 min	90°C (MW)	100
0.125 M	1.5 min	90°C (MW)	100
0.05 M	1.5 min	90°C (MW)	100

LC-MS analysis confirmed complete synthesis of the acrylamide-tagged peptide (**55**) under all three tested conditions. The 1% acrylic anhydride solution was considered to be cost-effective enough, while still maintaining an easily handleable volume, and was used for further acrylamide-tagging synthesis experiments.

3.11 Optimised protocol for automatic peptide synthesis with acrylamide-tagging

A second attempt at synthesising model peptide **127** (NH_2-ARQALF-$CONH_2$) using automated peptide synthesis was performed. All amino acid residues were coupled using standard single coupling procedures. Deprotection of Fmoc-protecting groups was achieved using DABCO (5% *w/v* in DMF). Acrylamide-tagging was performed with acrylic anhydride (1% in DMF). The same procedure was followed as described in Figure 3.12.A.

Upon completion of the synthesis, the peptide was cleaved off resin using a cleavage cocktail comprising of TFA/TIPS/H_2O (3 mL; 9 : 0.5 : 0.5). The peptide was precipitated in cold diethyl ether, spun down in the centrifuge and lyophilised. The crude product was then analysed by LC-MS (Figure 3.14), confirming synthesis of the correct peptide in good purity (>90%). These results suggest that the acrylamide-tagging synthesis procedure could subsequently be applied to longer, more complex peptides.

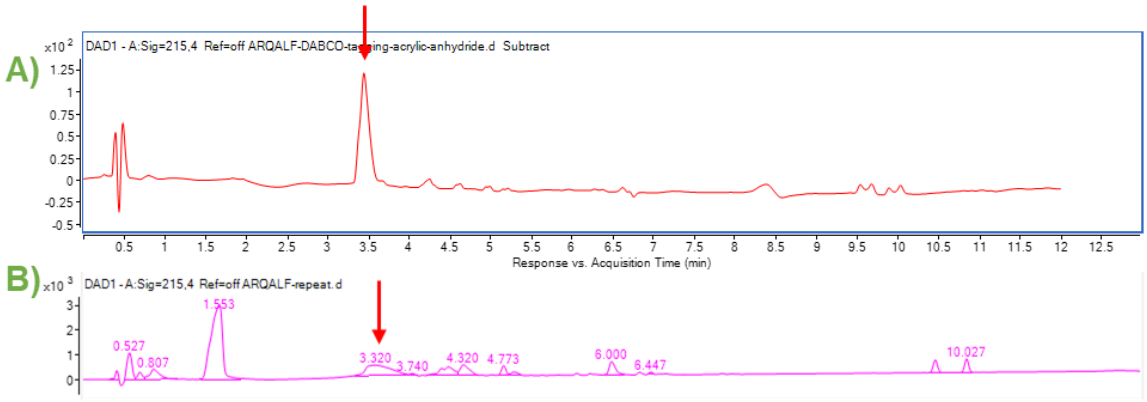


Figure 3.14: A) HPLC chromatogram (215 nm) of crude synthesis of peptide **127** following optimised acrylamide-tagging conditions, with the red arrow depicting the target peptide; B) HPLC chromatogram (215 nm) of peptide **127** following acrylamide-tagging synthesis prior to optimisation.

3.12 Application of acrylamide-tagging synthesis procedure to relevant biological target peptides: p53 sequence

The applicability of the acrylamide-tagging synthesis and purification procedure to obtain biologically relevant peptides was first tested with peptide **135** (NH_2-DPGPDEAPRMPEAAPPVAPAPAAPT-$CONH_2$), which is a part of the p53 protein (for more information see Chapter 6). The peptide was synthesised following initial optimisation (Chapter 6.4), leading to double couplings of proline and subsequent residues. After each amino acid coupling step, acrylamide-tagging using acrylic anhydride (1% in DMF) was introduced. Fmoc-removal was performed using DABCO (5% *w/v* in DMF).

3.12.1 Direct cleavage and resin transfer

Upon completion of the peptide sequence on resin, an attempt was made to directly purify the peptide mixture. Approximately 80 mg of the peptide on-resin (0.05 mmol) was used and cleaved using a cleavage cocktail comprising of TFA/TIPS/H_2O (4 mL; 9 : 0.5 : 0.5). Additionally, 1000 mg of thiol-resin was added to cleavage mixture, to investigate whether an acid catalysed Michael addition between the acrylamide-tagged peptide impurities and the thiol-resin could occur. If this would be possible, this would decrease purification time and reduce the amount of reagents required for purification. After allowing for cleavage and scavenging time (four hours) the peptide was precipitated in cold diethyl ether and lyophilised. The crude peptide was analysed by LC-MS, showing that the correct peptide was synthesised, however, that direct scavenging did not occur (Figure 3.15), suggesting base-catalysis is required.

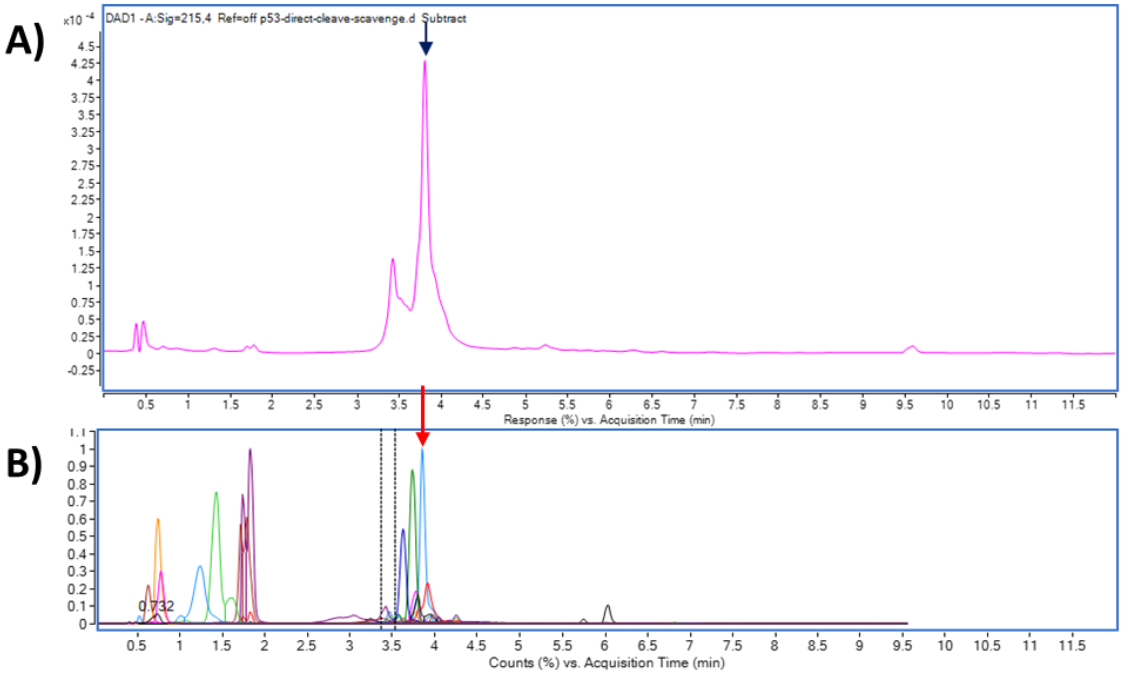


Figure 3.15: A) LC-MS chromatogram (215 nm) of crude peptide **135** after cleavage and direct scavenging attempt in absence of base. B) Molecular weight scan of expected peptides (target peptide and acrylamide-tagged impurities); target peptide depicted with red arrow. All products are scaled to maximum.

3.12.2 Base-catalysed scavenging

Upon completion of the peptide synthesis, cleavage, precipitation and lyophilisation an attempt was made to purify peptide **135** using previously optimised base-catalysed scavenging conditions. Crude peptide **135** (10 mg) was incubated with 1000 mg of thiol-resin (2 eq. DIPEA, 450 µL). As peptide **135** consists of 25 amino acids, a total 25 likely "targets" was to be expected, consisting of one desired peptide and up to 24 impurities (acrylamide-tagged peptides). A sample was taken from the scavenging mixture after 30, 60, 90, 120, 180 and 240 minutes, which was analysed by LC-MS. A 'Find by Formula' scan was conducted to identify how many of these targets could be found. Over time, the number of targets qualified (*i.e.* peptide side products detected) was seen to be decreasing (Figure 3.16), suggesting impurities are successfully being scavenged from the reaction mixture. Interestingly, after 3 hours a slight increase was observed, possibly as an effect from the reversibility of the Michael addition.

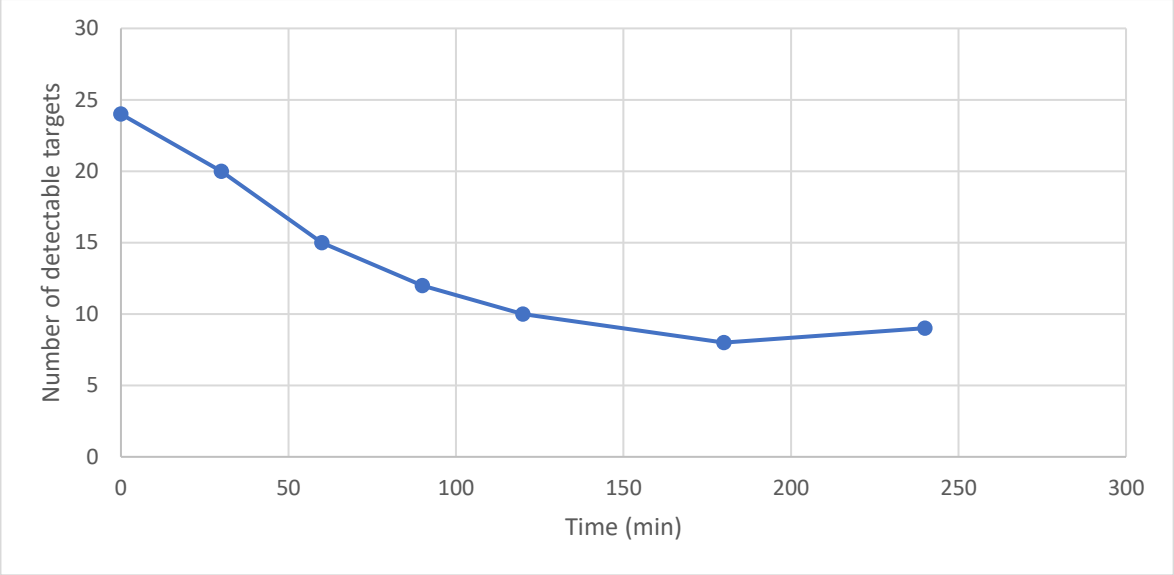


Figure 3.16: Detection of target peptides.

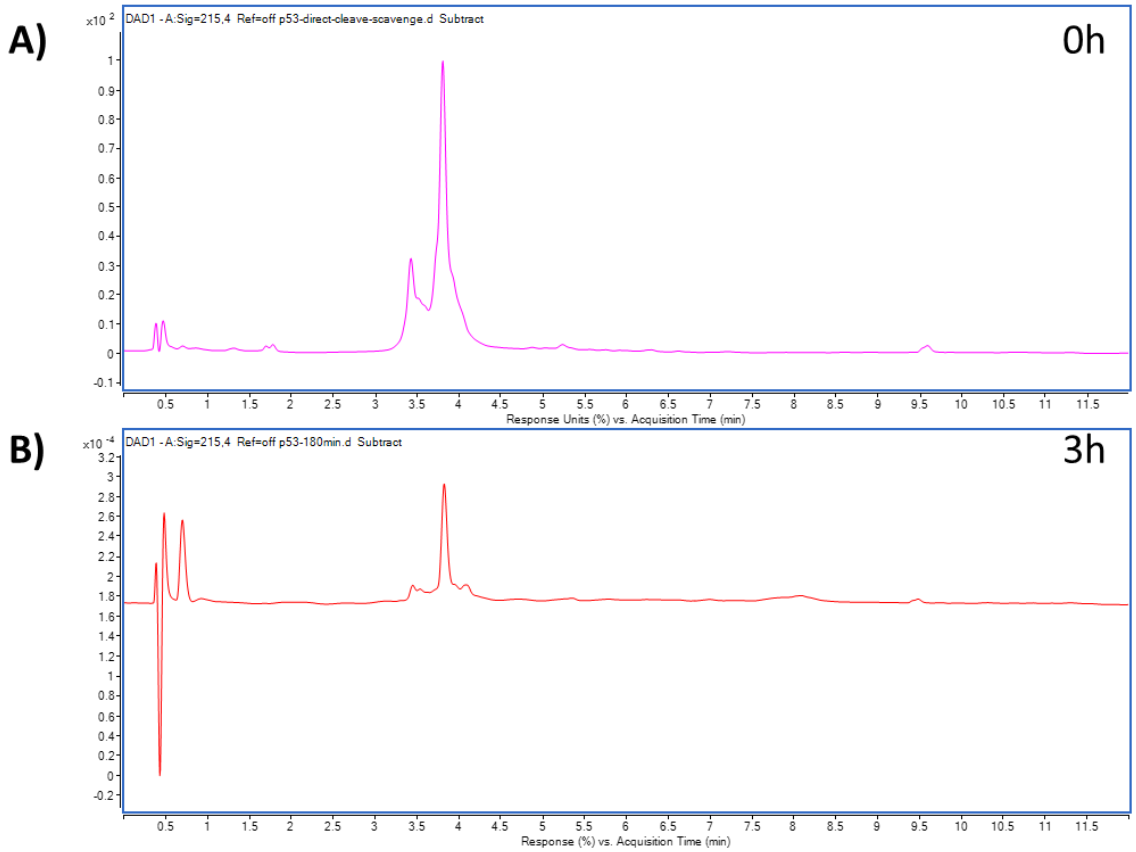


Figure 3.17: HPLC chromatogram of the scavenging of peptide **135** in presence of base (DIPEA 2 eq.) at A) 0h and B) 3h.

3.12.3 Resin refresh

As possible reversibility of the scavenging reaction appeared to start occurring after 3 hours, an attempt was made to push the scavenging to completion by refreshing the thiol-resin. Crude peptide **135** (10 mg) was incubated with 1000 mg of thiol-resin (2 eq. DIPEA, 450 µL) as before, but after three hours the resin was replaced with 1000 mg of fresh thiol-resin. Unexpectedly, refreshing the resin did not significantly further clean up the peptide mixture (Figure 3.18 B and C). Perhaps refreshment of base would have been required as well, to further push the scavenging to completion, however, this was not attempted.

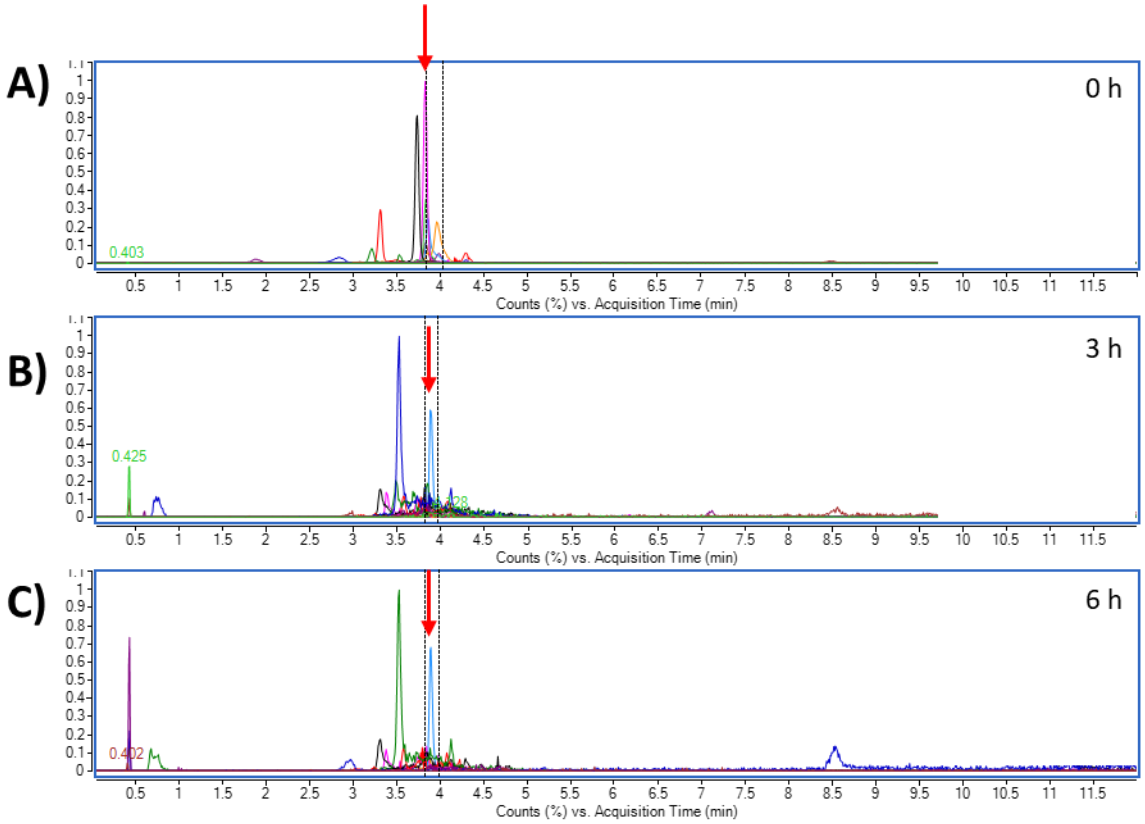


Figure 3.18: Find by Formula results of scavenging experiment of p53 peptide **135** after A) 0h; B) 3h; C) 6h. Each peptide is scaled to maximum. The target peptide is marked by the red arrow.

3.13 Application of acrylamide-tagging synthesis procedure to relevant biological target peptides: Exenatide

As a final test to demonstrate the efficiency of the acrylamide-scavenging method a complicated biologically relevant peptide drug was synthesised. Exenatide is a glucagon-like peptide-1 receptor agonist (GLP-1 receptor agonist) and is used for the treatment of diabetes. It is sold in the UK under the brand name Bydureon, as 2 mg powder and solvent for prolonged-release suspension for injection pre-filled pen (AstraZeneca UK Ltd), recommended as a single weekly dosage, at a cost of £73.36 [20]. If the production and purification process could be simplified, the costs could decrease.

3.13.1 Standard synthesis procedure

Exenatide is a relatively large peptide consisting of a total of 39 amino acids (NH_2-HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS-$CONH_2$, **136**), with a molecular weight of 4184.03. To demonstrate how difficult it is to prepare, it has been synthesised at 0.1 mmol scale (0.63 mmol/g loading, ProTide Rink amide resin) using only single amino acid couplings. No optimisation of the synthesis conditions were explored. The resulting crude peptide can be seen in Figure 3.19. As can be observed, there is no single sharp peak with a good resolution. Instead, there seems to be a range of peaks that are all have closely related retention times, causing an overlap of multiple peptide products. This makes it extremely difficult to purify peptide **136** by prep-HPLC.

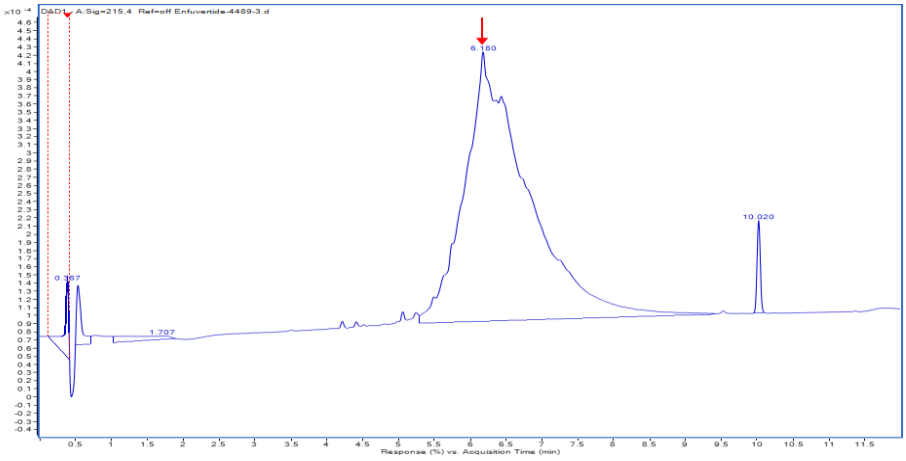


Figure 3.19: HPLC chromatogram of exenatide. The desired peptide (**136**) is indicated by the arrow.

3.13.2 Acrylamide-scavenging procedure

Exenatide was then prepared using the acrylamide-tagging synthesis procedure, still using only single couplings. The only difference with the previous procedure is the use of (5%) DABCO for Fmoc-removal (instead of 20% piperidine) and the introduction of acrylamide-tagging steps. The resulting crude peptide was analysed by LC-MS (Figure 3.20.A). As expected, the crude peptide was of poor quality, and purification by prep-HPLC would be complicated.

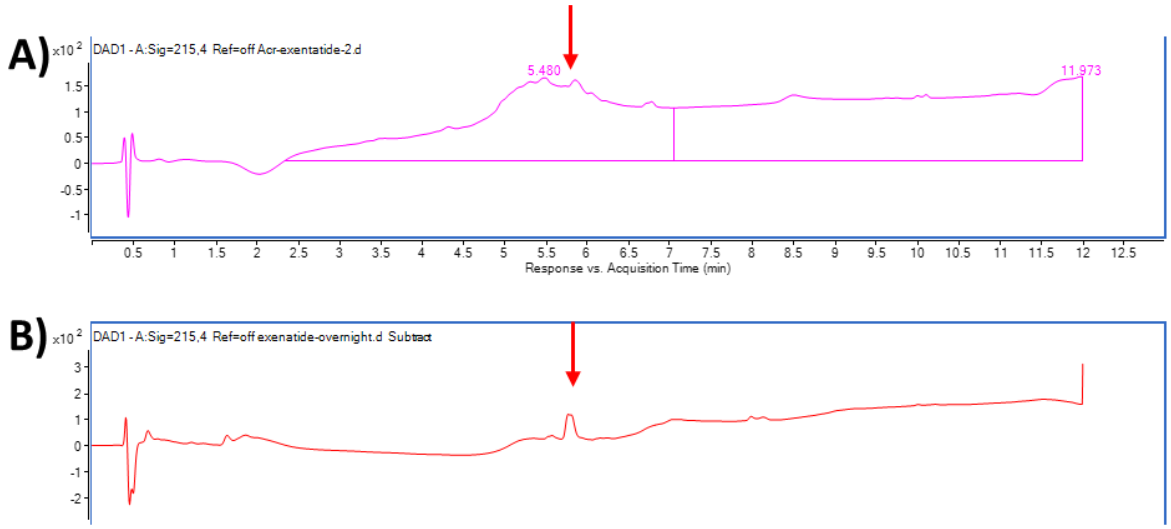


Figure 3.20: A) HPLC chromatogram of crude exenatide peptide **136** synthesised following acrylamide-tagging procedure. B) HPLC chromatogram (215 nm) of exenatide (**136**) after acrylamide-scavenging.

The crude exenatide (**136**) (10 mg) was dissolved in water/MeOH (10 mL; 1:1) and incubated with 1000 mg of thiol-resin in the presence of DIPEA (452 µL). The mixture was shaken over night (~16 hours), followed by LC-MS analysis. An improvement of the HPLC chromatogram was observed (Figure 3.20.B). For medical grade purity, additional prep-HPLC would still be required, however, it would be a lot easier to achieve after the acrylamide-scavenging step. In addition, the synthesis procedure can still be optimised (*e.g.* by introducing double couplings) to achieve a higher purity peptide after acrylamide-scavenging. The power of the acrylamide-scavenging approach has been demonstrated with this complicated example peptide.

3.14 Conclusion

This chapter built on previous optimisation of the development of a procedure for a new peptide synthesis and purification method, which incorporated ‘acrylamide-tagging’ and ‘acrylamide-scavenging’. New automated coupling procedures were successfully developed, after which an ‘acrylamide-tagging’ step could be introduced into standard automated SPPS. Additionally, alternative deprotection of the Fmoc-group, utilising DABCO instead of piperidine was successfully incorporated. The final optimised conditions employed 3 minute deprotections using 5% DABCO in DMF for Fmoc-removal and 1% acrylic anhydride in DMF for acrylamide-tagging. The applicability to the synthesis and purification of large and complicated peptide drugs has been demonstrated with two examples, p53 and exenatide.

The ‘acrylamide-scavenging’ purification method has wide potential applicability to the field of peptide drug research. Firstly, by applying this synthesis and purification method during scale-up and purification of large quantities HPLC may be avoided. However, acrylamide-purification is not just limited to large quantities. It can also be applied during parallel synthesis for the generation of peptide libraries, commonly used in drug discovery. When a library is synthesised the crude purity of each sample may vary, which could lead to unreliable results in biological assessment. It would be beneficial if peptide libraries were purified, however, using conventional HPLC purification would require a lot of time (optimisation of HPLC methods) and solvents. When a peptide library scan is undertaken, quick results are often required so more research can be undertaken following the identification of a lead molecule. Making use of acrylamide-scavenging purification could improve biological assay results without the time loss associated with prep-HPLC.

In conclusion, the acrylamide-scavenging purification method has been further optimised for automated SPPS. The first successful scavenging experiments on representative drug peptides have been undertaken, clearly demonstrating the potential of this method in the future of peptide synthesis and research.

3.15 References

1. Albericio, F. Orthogonal protecting groups for Nα-amino and C-terminal carboxyl functions in solid-phase peptide synthesis. *Biopolymers* **55**, 123–139 (2000).

2. Goodman, M. *Synthesis Of Peptides And Peptidomimetics*. (2004).

3. Anastas, P. & Eghbali, N. Green Chemistry: Principles and Practice. *Chem. Soc. Rev.* **39**, 301–312 (2010).

4. Sheldon, R. A. The *E* factor 25 years on: the rise of green chemistry and sustainability. *Green Chem.* **19**, 18–43 (2017).

5. Sheldon, R. A. Consider the environmental quotient. *Chemtech* **25**, 38–47 (1994).

6. Eissen, M. & Metzger, J. O. Environmental Performance Metrics for Daily Use in Synthetic Chemistry. *Chem. - A Eur. J.* **8**, 3580 (2002).

7. Van Aken, K., Strekowski, L. & Patiny, L. EcoScale, a semi-quantitative tool to select an organic preparation based on economical and ecological parameters. *Beilstein J. Org. Chem.* **2**, 3 (2006).

8. Trost, B. Selectivity: A Key to Synthetic Efficiency. *Science (80-.).* **219**, 245–250 (1991).

9. Welton, T. Solvents and sustainable chemistry. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **471**, 20150502 (2015).

10. Jad, Y. E., Kumar, A., El-Faham, A., De La Torre, B. G. & Albericio, F. Green Transformation of Solid-Phase Peptide Synthesis. *ACS Sustainable Chemistry and Engineering* **7**, 3671–3683 (2019).

11. Kennedy, G. L. Toxicology of dimethyl and monomethyl derivatives of acetamide and formamide: a second update. *Crit. Rev. Toxicol.* **42**, 793–826 (2012).

12. Kerton, F. & Marriott, R. *Alternative Solvents for Green Chemistry*. (Royal Society of Chemistry, 2013). doi:10.1039/9781849736824

13. Breslow, R. Hydrophobic effects on simple organic reactions in water. *Acc. Chem. Res.* **24**, 159–164 (1991).

14. Krenske, E. H., Petter, R. C. & Houk, K. N. Kinetics and Thermodynamics of Reversible Thiol Additions to Mono- and Diactivated Michael Acceptors: Implications for the Design of Drugs That Bind Covalently to Cysteines. *J. Org. Chem.* **81**, 11726–11733 (2016).

15. Aimetti, A. A., Shoemaker, R. K., Lin, C.-C. & Anseth, K. S. On-resin peptide macrocyclization using thiol–ene click chemistry. *Chem. Commun.* **46**, 4061 (2010).

16. Tian, Y. *et al.* Stapling of Unprotected Helical Peptides via Photo-Induced Intramolecular Thiol-Yne Hydrothiolation. *Chem. Sci.* **7**, 1–6 (2016).

17. Hoyle, C. E., Lee, T. Y. & Roper, T. Thiol-enes: Chemistry of the past with promise for the future. *J. Polym. Sci. Part A Polym. Chem.* **42**, 5301–5338 (2004).

18. Bordoni, A. V., Lombardo, M. V. & Wolosiuk, A. Photochemical radical thiol–ene click-based methodologies for silica and transition metal oxides materials chemical modification: a mini-review. *RSC Adv.* **6**, 77410–77426 (2016).

19. Merck. Protocols for the Fmoc SPPS of Cysteine-containing Peptides. Available at: https://www.sigmaaldrich.com/technical-documents/protocols/chemistry/fmoc-spps-cysteine-peptides.html.

20. Bydureon 2 mg powder and solvent for prolonged-release suspension for injection in pre-filled pen - Summary of Product Characteristics (SmPC) - (eMC). Available at: https://www.medicines.org.uk/emc/product/3650/smpc. (Accessed: 24th May 2019)

4 A Novel Peptide Cyclisation *via* Thiol-Acrylamide Conjugate Addition

During the development of the acrylamide-scavenging procedures (Chapters 2 and 3), it was noted that there were possible peptide cyclisation side-products resulting from the nucleophilic attack of sidechain heteroatoms at the acrylamide. This is likely to be a significant hurdle to overcome when thiols (e.g. cysteine) are present, yet should be slower in the presence of amines. Peptide cyclisation protocols are useful as they can be used to synthesise novel peptides that mimic, for example surface features of naturally occurring proteins. In this chapter, the chemoselectivity and conditions-optimisation are investigated for the side-reactions mentioned, as a deliberate way to perform peptide cyclisation through thiol-acrylamide conjugation.

Introduction

4.1 Cyclic peptides in chemical biology and medicinal chemistry

There is a growing interest in peptides as therapeutic agents due to potent biological activity, high target-specificity, little non-specific binding to molecular structures other than desired target, low accumulation in tissues (but better tissue penetration than proteins), predictable metabolism, and lower toxicity than proteins and small molecules. Additionally, they are often very potent and have lower manufacturing costs (small peptides vs. recombinant proteins).[1,2] However, in isolation linear peptides do not always maintain their native conformation and binding ability, as they lack the structural reinforcement of the rest of the protein.[3] Cyclising or stapling a peptide can reinforce the native peptide structure. As a result, cyclic peptides exhibit enhanced bioavailability and proteolytic stability in comparison to their linear counterparts.[4] These factors make structurally-constrained peptides especially attractive in the targeting and modulation of protein-protein interactions.[5,6] While early stapling methods involved proteogenic amino acids, like lactam formation,[7–9] the hydrocarbon staple has made the greatest impact on the field.[10–13] Since then, a wider range of synthetic methods have been developed, including cycloaddition,[14] Pd-catalysed cross-coupling,[15–18] and cysteine-based stapling (e.g. perfluoroaryl stapling).[19,20] Some of these methods require unnatural amino acids or cross-linkers, which can increase the manufacturing costs on a large scale.

Disulfide bridges are widely prevalent in nature and in synthetic medicines (*e.g.* in oxytocin, octreotide, vasopressin) between two or more cysteine residues (Figure 4.1). Disulfide bridges are sensitive to reducing conditions and scrambling of desired disulfide connectivity, limiting their potential as therapeutics, emphasising the need for alternative synthetic approaches to disulfide mimetics.[21] Pharmacokinetic properties of disulfides have been enhance by replacement of the -S-S- with more robust groups e.g. thioethers (carbetocin).

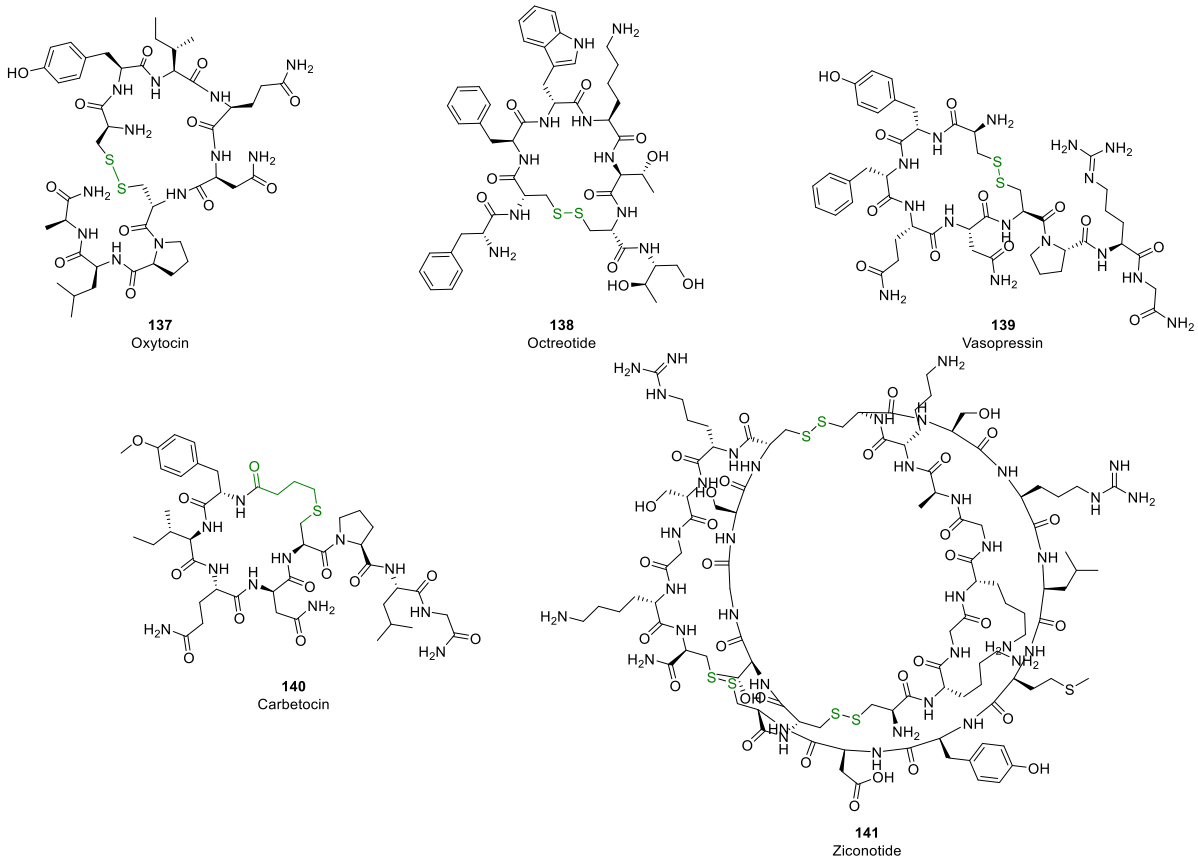
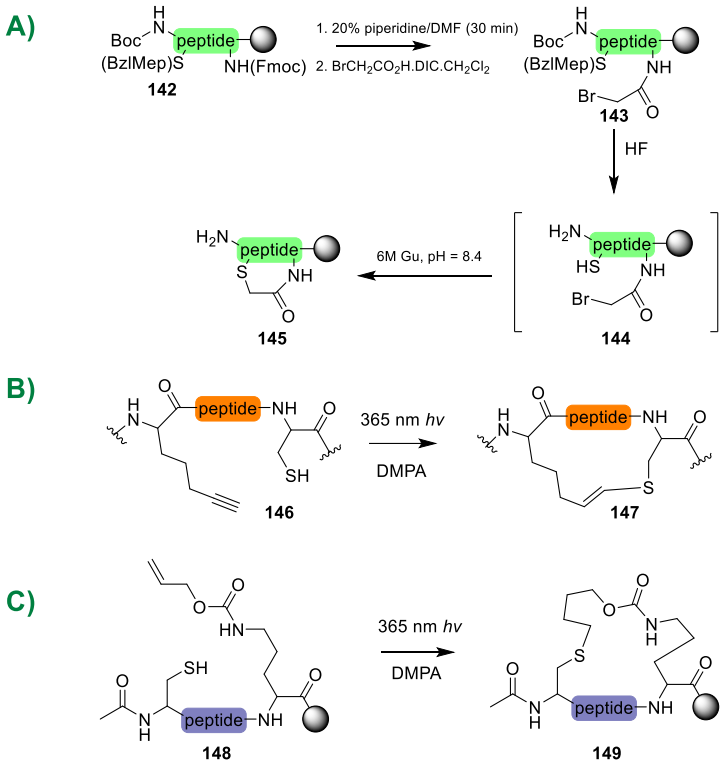


Figure 4.1: Structures of disulfide or thio-ether containing peptides.

Cysteine-disulfide crosslinking is broadly undertaken via two approaches; intramolecular and two-component crosslinking (Scheme 4.1, Figure 5.5). The two-component approach utilises two cysteine residues and a bifunctional electrophilic ‘linker’-molecule to form a macrocycle, which has been widely explored[22–29] allowing finetuning for the best peptide properties. For further reading, see Chapter 5. Additionally, cysteine replacements and alternative stereochemistry can be used for an optimal peptide shape.[30] While there has been great focus on two-component reactions, there has been less development in the field of one-component (intramolecular) cyclisation reactions.

4.1.1 One-component peptide cyclisation

Thioether cyclisation was developed by Bunel and Dawson, linking a cysteine with a bromoacetylated ornithine residue (Orn) in solution[31] (Scheme 4.1.A). Alternatively, commercially available alloc-lysine can react with cysteine in the presence of 2,2-dimethoxy-2-phenylacetophenone (DMPA) while exposed to 365 nm light (Scheme 4.1.B). Advantages of this technique include the use of commercially available amino acids and the short reaction time (20 – 60 min).[32] Two more photo-induced intramolecular cyclisation methods have been developed since,[33,34] limiting intramolecular thio-stapling to unnatural amino acid and photo-selective methods.



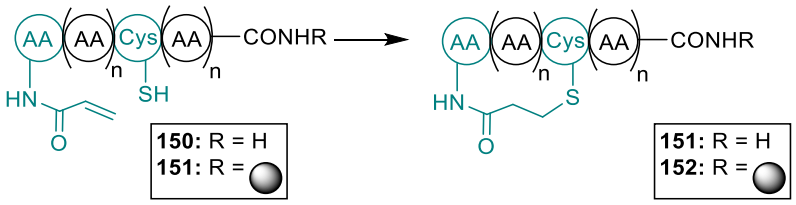
Scheme 4.1: Examples of one component stapling reactions by A) thio-ether ligation;[31] B) thiol-yne hydrothiolation;[34] and C) thiol-ene click chemistry.[32]

4.2 Introduction of a novel one-component cyclisation method

4.2.1 Proposed approach

During previous work (Chapters 2 and 3), a side product was observed following thiol-acrylamide peptide scavenging. It was hypothesised that this could be the result of an intramolecular cyclisation reaction between nucleophilic sidechain residues and the acrylamide group. It was proposed, that this could be taken advantage of for directed cyclisation reactions using cysteine residues. In this chapter, we aim to develop a new protocol for peptide cyclisation *via* thiol-acrylamide conjugate addition.

Using methods developed for the purpose of acrylamide-scavenging, acrylamide-groups can be introduced into designed model peptides. To limit potential side reactions and to optimise cyclisation reaction conditions, initial model peptides designed for this approach contained largely non-nucleophilic sidechains, with the exception of a cysteine residue (Scheme 4.2). Additionally, differences between on-resin and in-solution cyclisations were explored. Finally, an attempt was made to synthesise cyclic peptides based on literature.

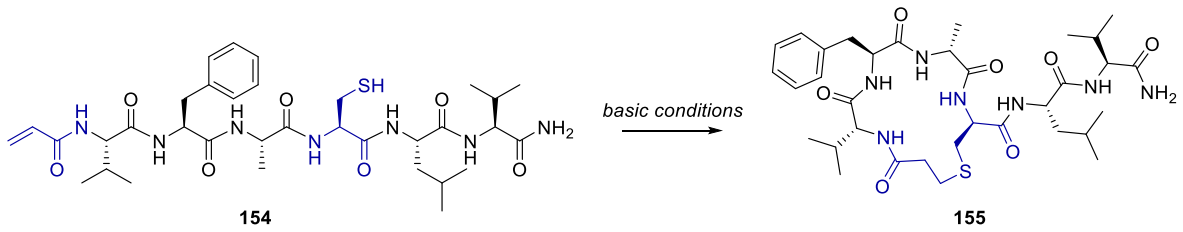


Scheme 4.2: Schematic overview of the cyclisation reaction. AA: amino acid.

Results and discussion

4.3 Synthesis of cyclic model peptide **155**

The reactivity of acrylamide with cysteine, and other nucleophilic amino acids, has thoroughly been discussed in the previous two chapters. To explore the scope of a cyclisation reaction, peptide **154** (Acr-VFACLV-$CONH_2$) was synthesised using standard automated SPPS, followed by manual coupling of the acrylamide-group using acryloyl chloride. This sequence was selected as an initial model as it contained only cysteine as a nucleophilic amino acid (Scheme 4.3). The peptide was cleaved from the resin using cleavage cocktail comprising of TFA/TIPS/H_2O (3 mL, 9 : 0.5 :0.5), precipitated in cold diethyl ether and lyophilised.



Scheme 4.3: Cyclisation of linear peptide **154** to cyclic peptide **155**.

4.3.1 Base-catalysed cyclisation

An initial cyclisation reaction was attempted under conventional and microwave conditions. As previously demonstrated in Section 2.8 the reaction between acrylamide and a thiol can require several hours. For the room temperature conditions, the peptide (1 mg) was treated with 25 mM DIPEA in DMF (1 mL) for 4 hours. Afterwards, the peptide was precipitated in cold diethyl ether and analysed using LC-MS. In an attempt to speed up the reaction, the analogous reaction was conducted under microwave conditions (90°C for 15 min).

As the molecular weight of the linear and cyclic peptide are identical, identification was dependent on LC retention time shifts. For comparison, peptide **154** was also analysed directly after cleavage from the resin. In this case, a retention time of 6.41 minutes was observed (Figure 4.2). Upon treatment under microwave conditions a shift of retention time could be observed to 5.97 minutes. Similarly, under the room temperature cyclisation conditions two peaks can be observed at ~5.5 and ~6 minutes. This could suggest that the new peptide peak

is formed with a retention time of 5.97 minutes corresponds to the cyclic peptide. However, the LC signals are very weak, and the experiment would have to be repeated at a higher concentration to obtain more significant results. These results do suggest that cyclisation under microwave conditions provides a cleaner peptide product, however, more experiments are required.

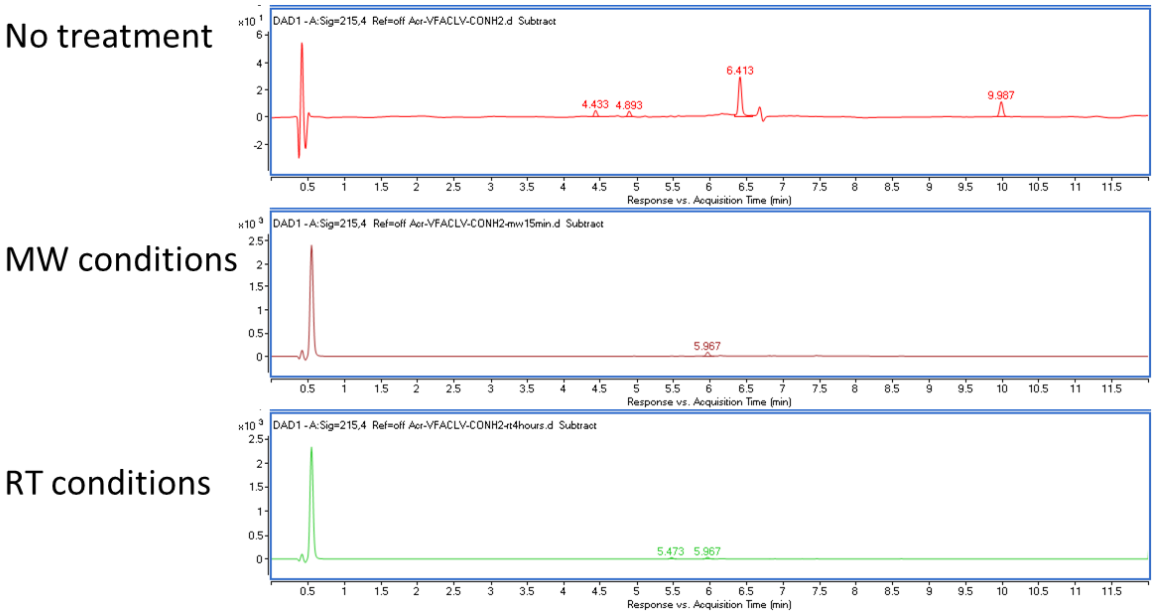
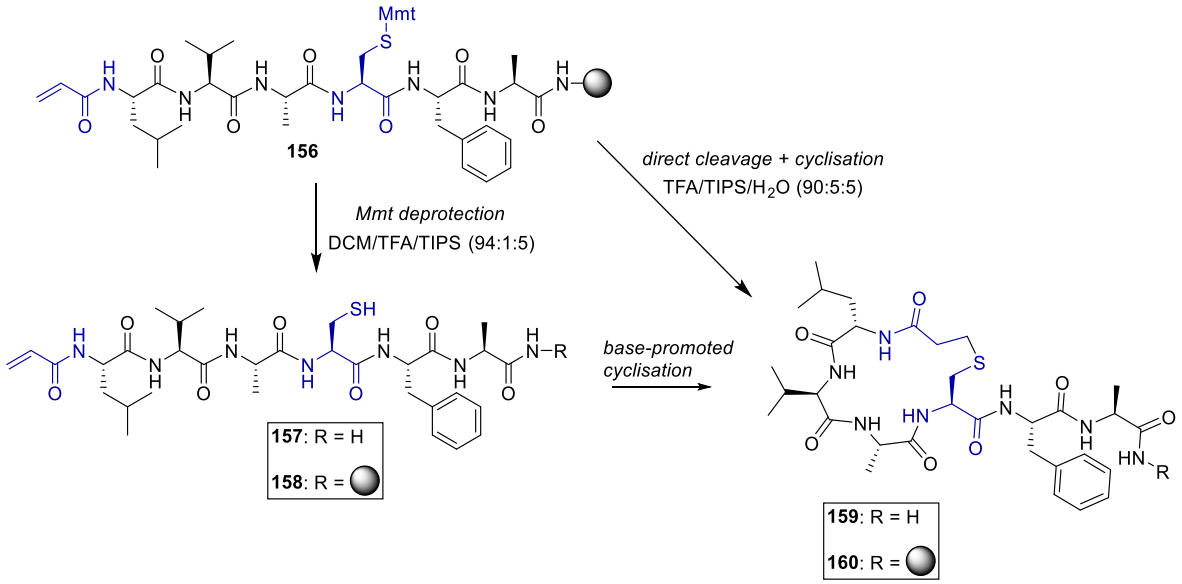


Figure 4.2: HPLC chromatograms of attempted cyclisation of peptide **147**.

4.4 Synthesis of cyclic model peptide **159**

In a second peptide the Cys(Trt) building block was replaced with Cys(Mmt), which allowed for orthogonal deprotection of the cysteine residue and thus cyclisation on resin. A new sequence Acr-LVACFA-$CONH_2$ (peptide **150**) was designed, again containing only non-nucleophilic amino acid residues. The new sequence was used to ensure the previous lack of significant results was not sequence specific. Two sets of experiments were carried out using this peptide sequence, investigating acid- and base-catalysed cyclisation.



Scheme 4.4: Cyclisation of peptide **156**.

4.4.1 Acid-promoted cyclisation

As the reaction requirements for the cyclisation were unknown, a first attempt was made to look into an acid-promoted Michael addition. Previous hypothesis entailed a peptide containing a cysteine residue where cleavage from the resin was expected to deliver the linear, fully-deprotected peptide, which could be subsequently cyclised under basic conditions. However, enhancement of the electrophilic character of the β-carbon of the acrylamide group under acidic (cleavage) conditions makes the acrylamide group an excellent Michael acceptor for a thiol-residue (cysteine).[35] Since the conjugate addition should be possible under acidic conditions, single step cleavage-cyclisation during the deprotection phase of the synthesis was considered to be achievable. Cyclisation was expected to be enhanced by the use of

Cys(Mmt), as the Mmt protecting group would be the first to be removed, allowing the cyclisation to take place before any other cleavage would take place. Peptide **156** was synthesised on resin, after which an attempt was made at direct cleavage and cyclisation using normal cleavage cocktail (TFA/TIPS/H_2O, 90 : 5 : 5), while shaking at room temperature. A sample of the cleavage mixture was analysed by LC-MS at the start of the cleavage and then every subsequent hour (in total 5 samples). It was expected that a shift from a linear peptide to a cyclic one would be observed over time. After one hour of treatment with TFA, a peak started to emerge with a retention time between 6.00 – 6.05 minutes, corresponding to a mass-to-charge ratio of 676.35 $[M+H]^+$, which equates to the calculated molecular weight of both the linear (**157**) and cyclic (**160**) peptide. After two hours of attempted cyclisation, there is another peak visible at a retention time of 6.71 minutes, corresponding to a mass-to-charge ratio of 675.34 $[M+H]^+$. While the calculated molecular weight for both the cyclic product and the linear product are the same, a difference of one proton has been observed in literature by Aimetti *et al*, where the linear product has one proton more than its cyclic counterpart (see Figure S2 of reference 32). This could suggest that this is also the case for peptide **157** and **159**, in which case the peak at 6.00-6.05 minutes corresponds to linear peptide **157**, and the peak at 6.71 minutes to cyclic peptide **159**. However, the peak at 6.71 is not very well defined, which could be a consequence of instability of the cyclic product in an acidic environment. The peaks at a retention time of 4.53 and 9.96 minutes could not be reasonably assigned to a peptide product.

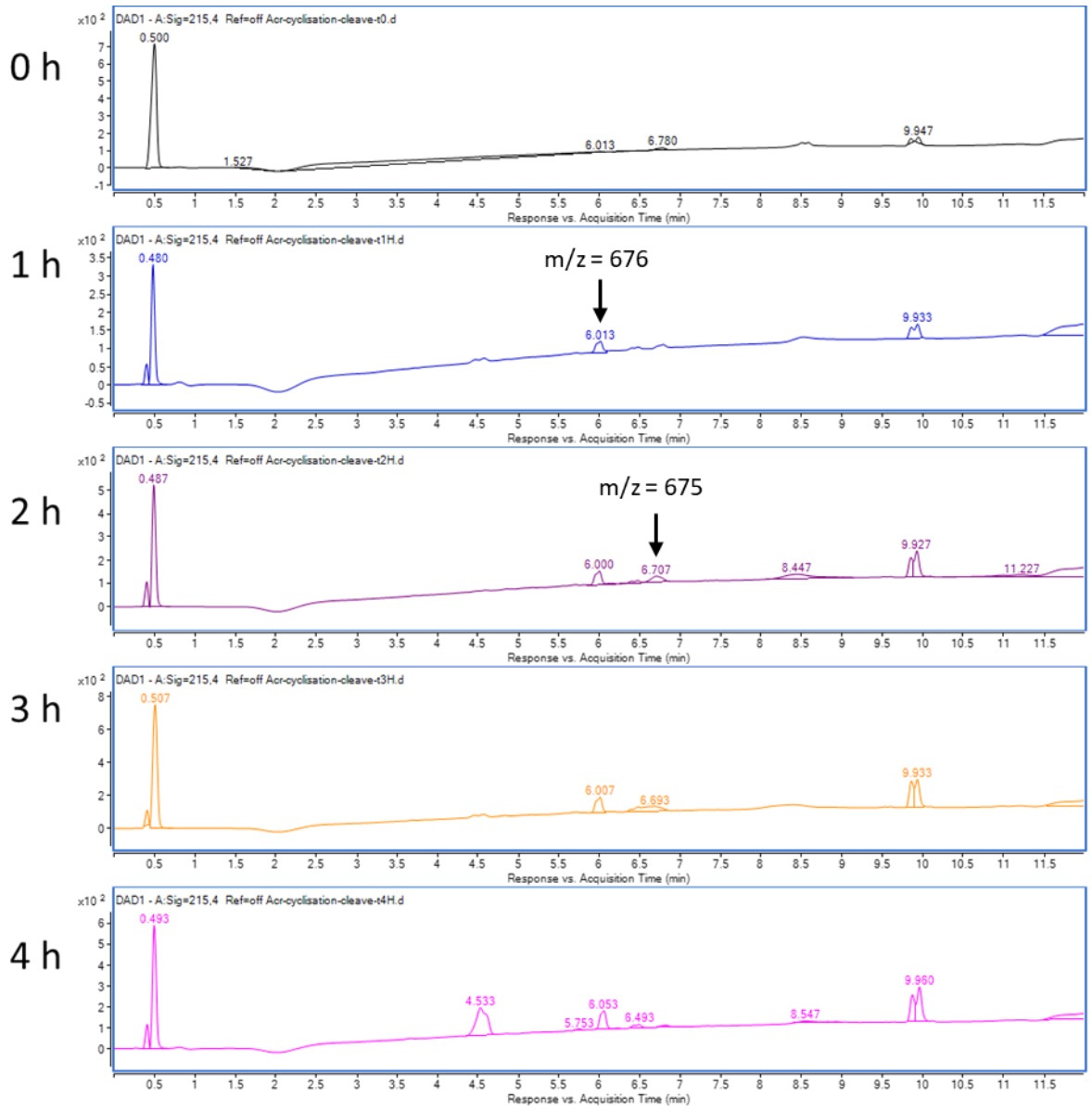


Figure 4.3: HPLC chromatograms (215 nm) monitoring cleavage of peptide **156** from the resin and potential cyclisation reaction. The peak at 0.49 minutes corresponds to the TFA present in the sample mixture.

4.4.2 Base-promoted on-resin cyclisation

Subsequently, the base-promoted on resin cyclisation of peptide **156** was attempted using a low loading resin (0.19 mmol/g loading) to avoid potential dimerisation. To allow for the cyclisation to take place, the cysteine protecting group (Mmt) was first removed using a mixture of DCM/TFA/TIPS (94 : 1 : 5; ~5 mL, 5 x 2 min). Upon removal of Mmt, recognisable by the orange coloured deprotection solution, the resin (**158**) was incubated in a solution of

25 mM DIPEA in DMF (3 mL). A sample of resin (~25 mg) was collected hourly (1 – 4 h), and the DIPEA solution was replaced with a fresh solution. A final sample was collected 24 hours after initiation of the cyclisation reaction. Each sample was then cleaved from the resin over 2 hours (TFA/TIPS/H_2O, 9 : 0.5 : 0.5; 1 mL), precipitated in cold diethyl ether, dissolved in H_2O/MeOH (1:1; 1 mL) for analysis by LC-MS (Figure 4.3).

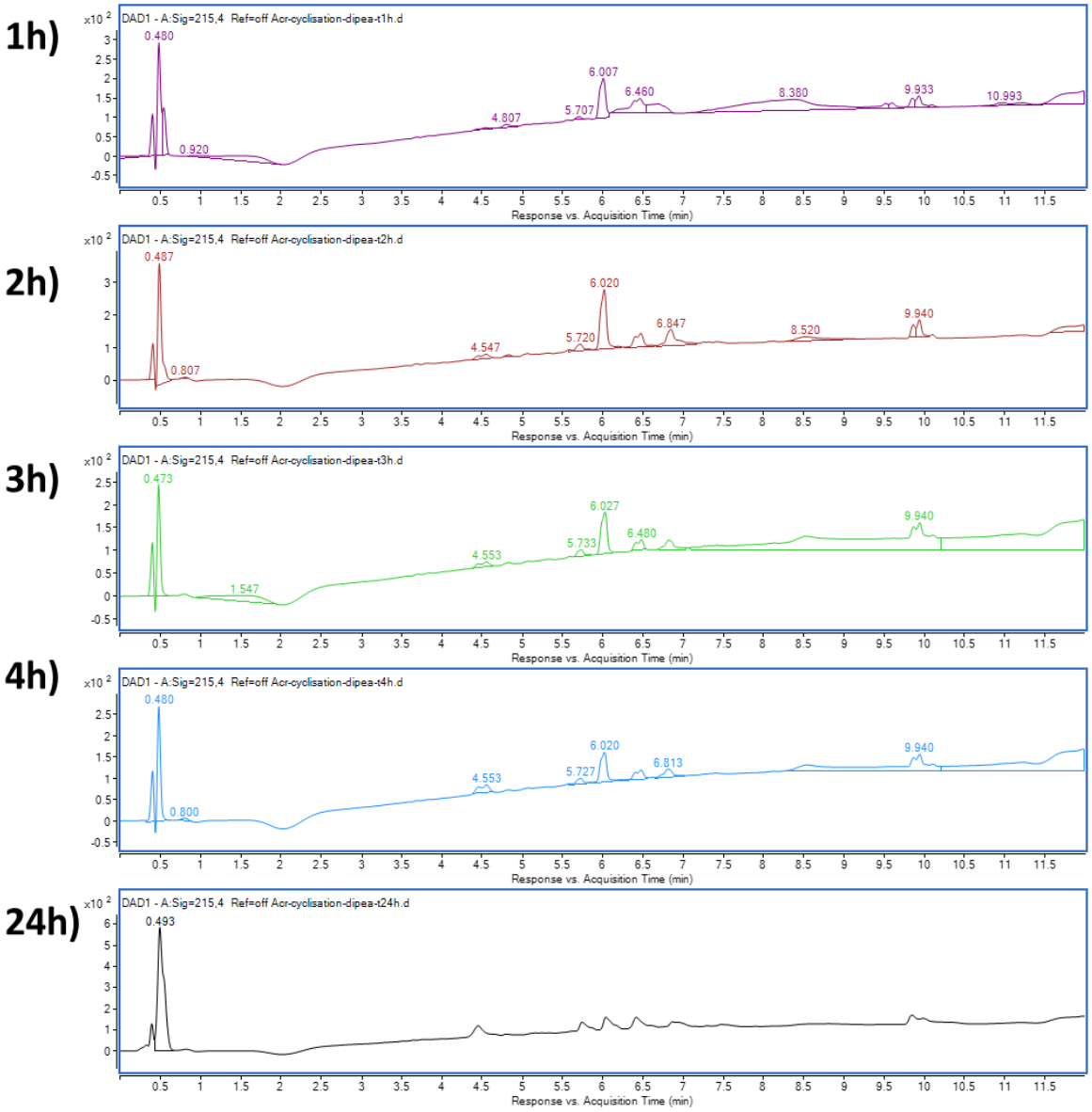


Figure 4.4: HPLC chromatograms (215 nm) monitoring potential base-catalysed cyclisation reaction of peptide **158**.

Similarly to the acid-catalysed experiment, a peak in the LC chromatogram can be observed with a retention time of 6.02 minutes, which remains present in the first four hours (and has the mass-to-charge ratio of 676.35 – corresponding to the [M+H$^+$] for the linear product). After two hours of attempted based-catalysed cyclisation two additional product peaks started to appear, at a retention time of 6.48 and 6.84 minutes. Interestingly, the molecular weights found for these peaks are 676.35 and 675.35, respectively. This could suggest the occurrence of racemisation for the former, and possible cyclisation for the latter peptide product. However, no products have been isolated for further analysis, so this cannot be stated with certainty. Surprisingly, after incubation of the peptide on-resin for 24 hours, the peak shapes of the peptide products appear less distinguishable by HPLC, which could be a consequence due to partial hydrolysis.

4.5 Investigation into possible dimerisation

Previous attempts to cyclise peptide **154** or peptide **156** did not provide clear evidence that a cyclisation reaction would take place. As higher molecular weight species (double the calculated number) were also detected, it was suggested that dimerisation of peptides could be a significant product. There are several applications for peptide dimerisation. For example, it could be used to carry a peptide drug into a cell, when the peptide drug is dimerised with a cell-penetrating peptide sequence. The peptide drug could be released due to the reversible nature of the Michael addition. After further developments, it could potentially be possible to direct a dimerisation reaction by adjusting two short peptide sequences to be compatible towards each other. In other words, one peptide sequence could be more positively charged, while the other peptide contains more negatively charged residues in the area of dimerisation.

An experiment was designed to test the possibility of dimerisation under the conditions explored earlier. Two peptides (**154** and **157**) with distinguishable molecular weights were synthesised. The chemical composition of both peptides was similar (largely hydrophobic and in the same proportion) so no bias towards either hetero- or homodimerisation was expected.

Only if dimerisation would take place, the molecular weight species matching the hetero dimer peptide would be detectable. This would be distinguishable from other higher molecular weight species that could be detected upon homodimerisation.

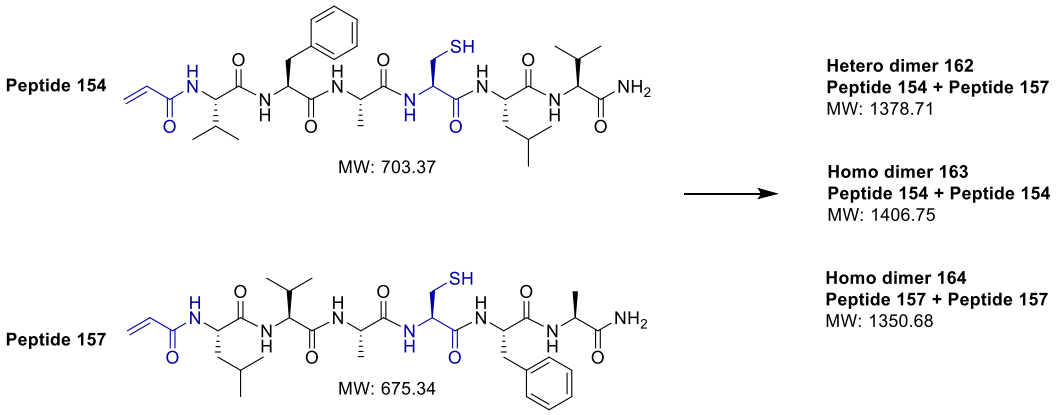


Figure 4.5: Structures of peptide **154** and peptide **157** and the resulting molecular weights if dimerisation takes place.

Dimerisation was attempted in solution phase. Peptide **154** and **157** were mixed at a concentration of 1 mg / mL each in a total volume of 1 mL 25 mM DIPEA in DMF. Two different reaction conditions were explored; **1)** conventional conditions: 15 minutes shaking at room temperature, and **2)** microwave conditions: 1 minute at 90°C.

In both cases, LC-MS analysis did not provide satisfactory data by UV (215 nm) analysis, with no peptide being detected. However, in the MS TIC scan, multiple peaks did appear, with the main peaks corresponding to mass-to-charge ratios of 659.32, 687.35, 1373.67, and 1429.73 after treatment under microwave conditions (Figure 4.5). Room temperature conditions provided similar results. These mass-to-charge ratios do not correspond to any of the calculated molecular weights (Figure 4.5). To gain better understanding in potential dimerisation, future studies are required. Conducting experiments at a higher concentration may provide results that are easier to interpret.

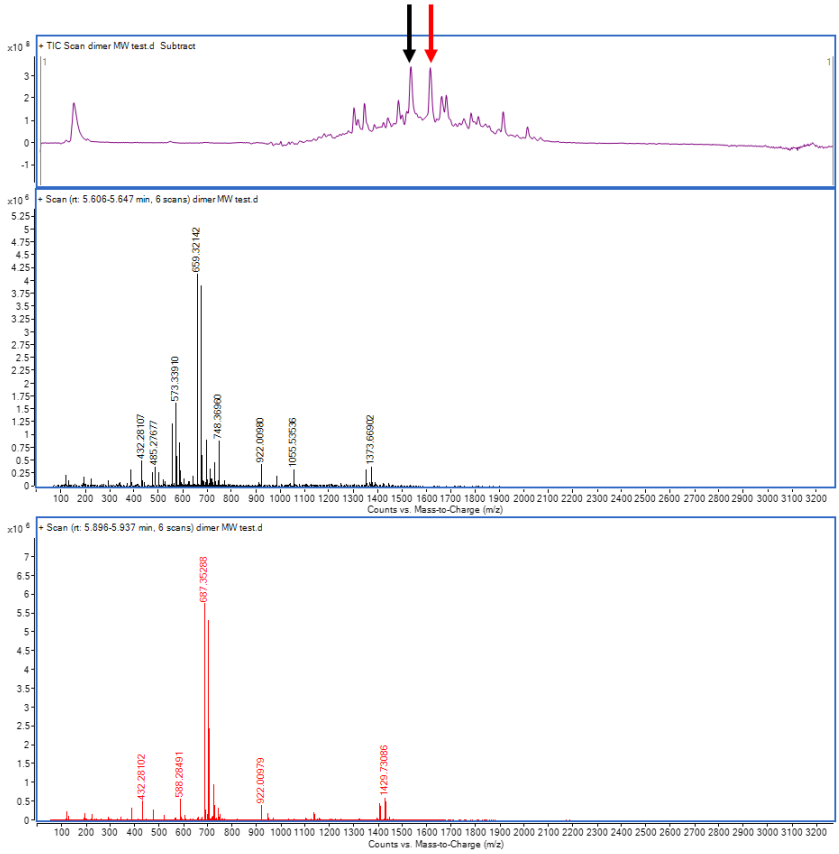
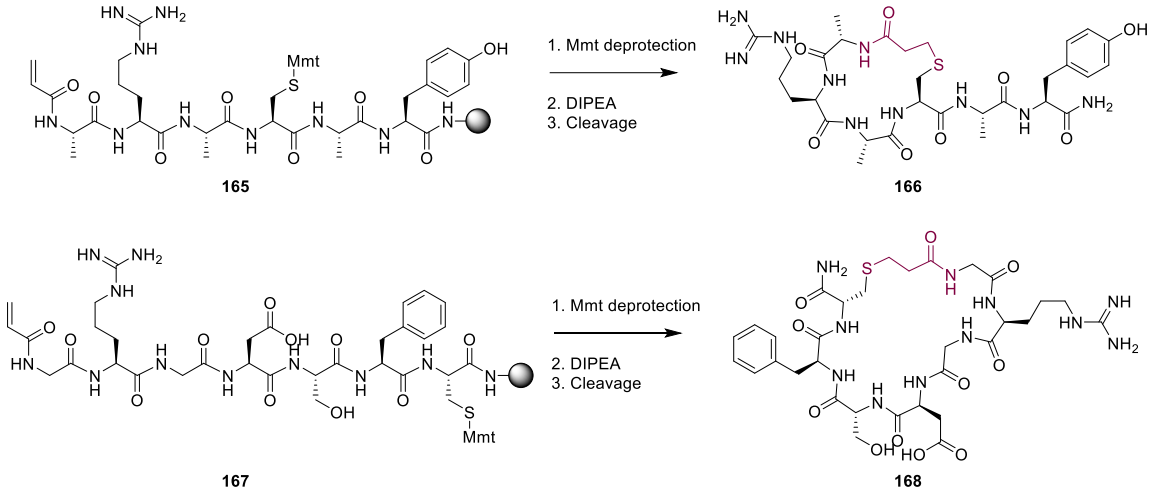


Figure 4.6: MS results of attempted dimerisation using peptide **154** and peptide **157**. The colour of the arrow equals the colour of the accompanying MS spectrum.

4.6 Synthesis of known cyclic peptides using on-resin thiol-acrylamide conjugate addition

Given the disappointing results obtained and to eliminate the possibility that the sequences selected were naturally difficult to cyclise, two alternative peptide sequences: Peptide **165** (Acr-ARACAY-●) based on the work of Tian *et al.* [34] and peptide **167** (Acr-GRGDSFC-●) based on the work of Aimetti *et al.* [36] were utilised for subsequent experiments. These sequences had been demonstrated before (though using different cyclisation procedures) to undergo cyclisation.



Scheme 4.5: Cyclisation of peptide **165** and peptide **167**.

Based on the previous cyclisation attempts, the cyclisation of peptide **165** and peptide **167** was monitored over multiple days, rather than multiple hours, as the reactions appeared to proceed at a slow rate. Peptides were synthesised using standard automated SPPS procedures, with an extra step of N-terminal acrylamide-capping using acrylic anhydride in DMF (1%). Cysteine residues were introduced with an Mmt protecting group, allowing for orthogonal deprotection preceding the cyclisation experiments, using DCM/TFA/TIPS (94 : 1 : 5, 5x 2 min, 2 mL).

Prior to attempted cyclisation a small portion of the resin (~25 mg) was collected for direct cleavage. Remaining peptide on-resin was incubated in 25 mM DIPEA in DMF. Samples (~25 mg resin) were collected after 24, 48, 72, and 96 hours, accounting for a total of 5 samples. All peptides were cleaved from the resin for 2 hours using a cleavage cocktail comprising of TFA/TIPS/H_2O (9: 0.5 : 0.5, 300 µL).

Peptide **166** did not provide a clear HPLC chromatogram upon analysis, with 10 different peaks being detected. The calculated MW (706.8) can be detected in the peaks at a retention time of 1.6, 1.7, 3, and 3.6 minutes when the crude peptide is directly analysed upon cleavage. Interestingly, only the chromatogram after 72 hours of incubation with base appeared to have

one main peptide product (Figure 4.7). Overall, this peptide was not considered to be a suitable, clear model peptide to analyse the potential cyclisation reaction.

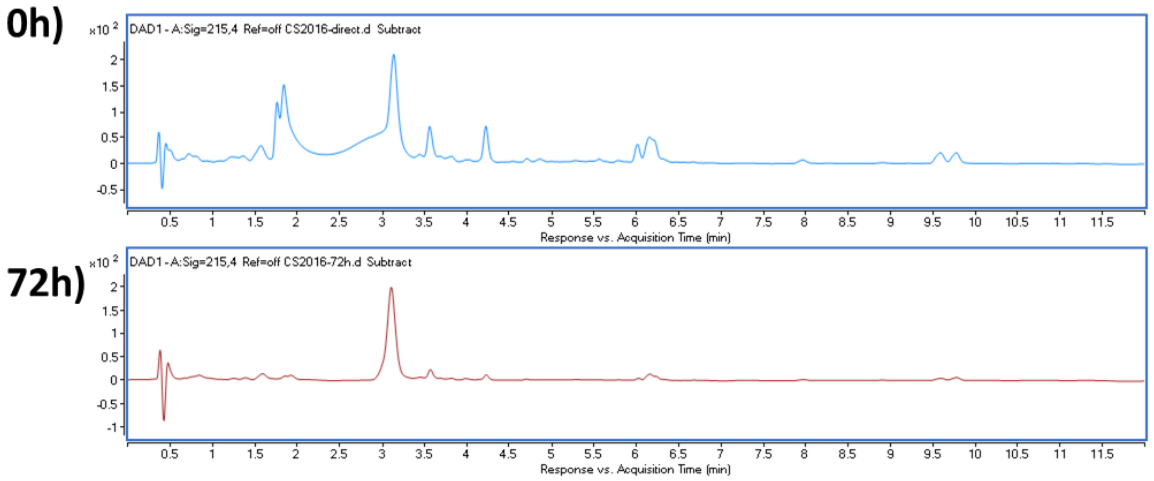


Figure 4.7: HPLC chromatogram (215 nm) of peptide **166** directly after cleavage and 72 hours after incubation with base (25 mM DIPEA in DMF).

Peptide **168** appeared to be a better model peptide, with only three distinct peaks being identifiable upon direct cleavage (Figure 4.8). The peaks at a retention time of 3.29-3.33 and 3.90-3.91 minutes correspond to the predicted molecular weight of 793.85. The third peak at a retention time of 4.44-4.45 is most likely related to the correct peptide sequence, however, not fully deprotected and still containing a tBu-group.

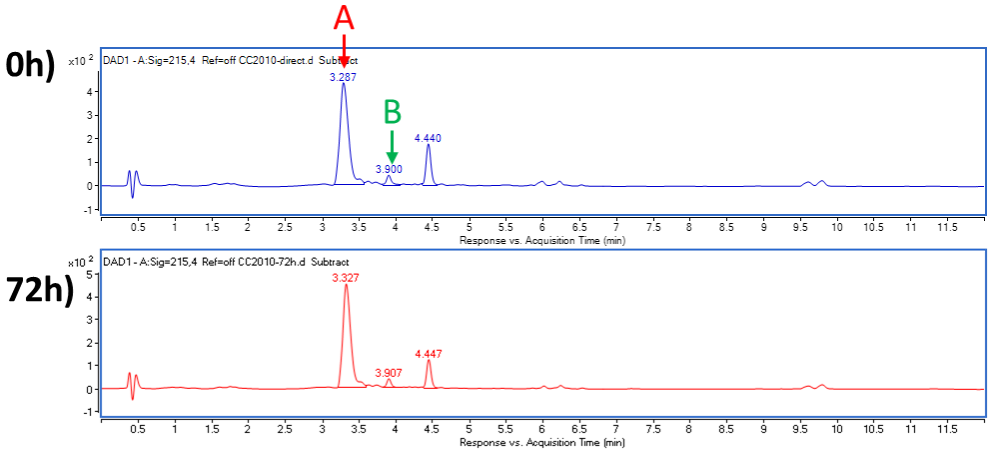


Figure 4.8: HPLC chromatogram (215 nm) of peptide **168** directly after cleavage and 72 hours after incubation with base (25 mM DIPEA in DMF).

The distribution of the two peaks corresponding to the correct molecular weight (Peak A and Peak B) was monitored over time. The area under the curve (AUC) of each peak was calculated. The relative occurrence of each peptide was calculated from the sum of Peak A and B. Over time, there does not appear to be a major shift from Peak A towards Peak B and the peptide represented by Peak A remains the major species (Figure 4.8). As it is unknown which peak would represent the possibly cyclic peptide, it is not possible to draw any conclusions from these results on the cyclisation process.

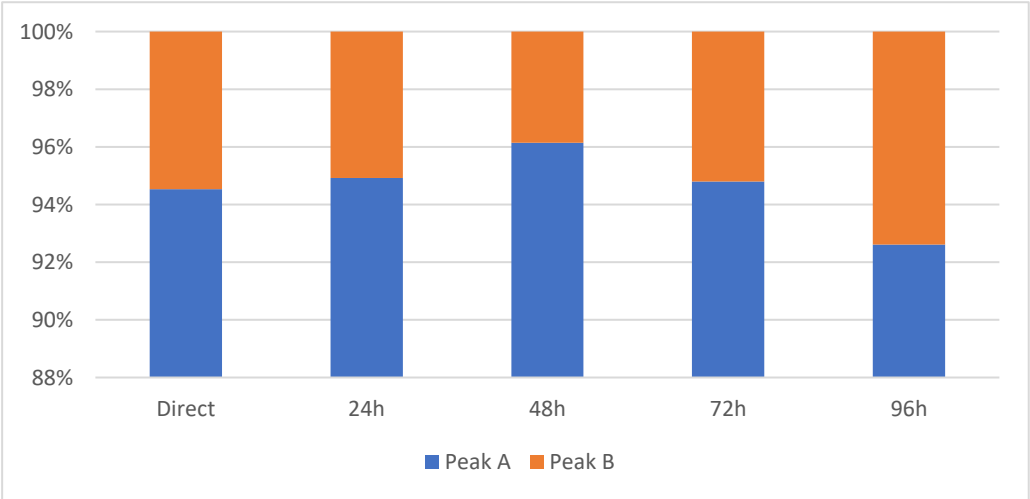


Figure 4.9: Distribution of occurrence (relative AUC) between Peak A and Peak B of HPLC chromatograms (215 nm) of peptide **168**.

4.7 Conclusion

Peptide cyclisation forms an interesting area of study as it allows for nearly endless possibilities for the development of novel cyclisation methods. Cyclic peptides are of great interest in the field of medicinal chemistry, as they often display superior proteolytic stability compared to their linear counterparts.[6] To make cyclic peptides more readily available and diverse, new cyclisation methods should be developed.

In this chapter, preliminary data has suggested that it is possible to cyclise peptides *via* thiol-acrylamide conjugate addition. However, an effective protocol has not yet been developed. For more conclusive results, more experiments are required. These could be conducted at a larger scale to improve the analysis conditions. Once reaction requirements and controllability are fully understood, there is potential for further studies, including ring-size explorations and expansions to acrylamide containing amino acid side chains.

4.8 References

1. Acar, H. *et al.* Self-assembling peptide-based building blocks in medical applications. *Adv. Drug Deliv. Rev.* **110**–**111**, 65–79 (2017).

2. Vlieghe, P., Lisowski, V., Martinez, J. & Khrestchatisky, M. Synthetic therapeutic peptides: science and market. *Drug Discov. Today* **15**, 40–56 (2010).

3. Lau, Y. H., De Andrade, P., Wu, Y. & Spring, D. R. Peptide stapling techniques based on different macrocyclisation chemistries. *Chemical Society Reviews* **44**, 91–102 (2015).

4. Millward, S. W., Fiacco, S., Austin, R. J. & Roberts, R. W. Design of cyclic peptides that bind protein surfaces with antibody-like affinity. *ACS Chem. Biol.* **2**, 625–634 (2007).

5. Nevola, L. & Giralt, E. Modulating protein–protein interactions: the potential of peptides. *Chem. Commun.* **51**, 3302–3315 (2015).

6. Ahrens, V. M., Bellmann-Sickert, K. & Beck-Sickinger, A. G. Peptides and peptide conjugates: therapeutics on the upward path. *Future Med. Chem.* **4**, 1567–1586 (2012).

7. Ye, Y., Gao, X., Liu, M., Tang, Y. & Tian, G. Studies on the synthetic methodology of head to tail cyclization of linear peptides. *Lett. Pept. Sci.* **10**, 571–579 (2003).

8. Yang, L. & Morriello, G. Solid phase synthesis of 'head-to-tail' cyclic peptides using a sulfonamide 'safety-catch' linker: The cleavage by cyclization approach. *Tetrahedron Lett.* **40**, 8197–8200 (1999).

9. Romanovskis, P. & Spatola, A. F. Preparation of head-to-tail cyclic peptides via side-chain attachment: implications for library synthesis. *J. Pept. Res.* **52**, 356–374 (1998).

10. Miller, S. J., Grubbs, R. H. & Beckman, M. Synthesis of Conformationally Restricted Amino Acids and Peptides Employing Olefin Metathesis. *J. Am. Chem. Soc.* **117**, 5855–5856 (1995).

11. Blackwell, H. E. & Grubbs, R. H. Highly Efficient Synthesis of Covalently Cross-Linked Peptide Helices by Ring-Closing Metathesis. *Angew. Chemie Int. Ed.* **37**, 3281–3284 (1998).

12. Scott J. Miller, Helen E. Blackwell, and & Grubbs*, R. H. Application of Ring-Closing Metathesis to the Synthesis of Rigidified Amino Acids and Peptides. (1996). doi:10.1021/JA961626L

13. Kim, Y.-W. & Verdine, G. L. Stereochemical effects of all-hydrocarbon tethers in i,i+4 stapled peptides. *Bioorg. Med. Chem. Lett.* **19**, 2533–2536 (2009).

14. Scrima, M. *et al.* CuI catalyzed azide-alkyne intramolecular i-to-(i+4) side-chain-to-SideChain cyclization promotes the formation of Helix-Like secondary structures. *European J. Org. Chem.* **2010**, 446–457 (2010).

15. Willemse, T., Schepens, W., Vlijmen, H., Maes, B. & Ballet, S. The Suzuki–Miyaura Cross-Coupling as a Versatile Tool for Peptide Diversification and Cyclization. *Catalysts* **7**, 74 (2017).

16. Noisier, A. F. M. & Brimble, M. A. C-H functionalization in the synthesis of amino acids and peptides. *Chem. Rev.* **114**, 8775–8806 (2014).

17. Feng, Y. & Chen, G. Total synthesis of celogentin C by stereoselective C-H activation. *Angew. Chemie - Int. Ed.* **49**, 958–961 (2010).

18. Tang, J., He, Y., Chen, H., Sheng, W. & Wang, H. Synthesis of bioactive and stabilized cyclic peptides by macrocyclization using C(sp 3)–H activation. *Chem. Sci.* **8**, 4565–4570 (2017).

19. Spokoyny, A. M. *et al.* A Perfluoroaryl-Cysteine S_NAr Chemistry Approach to Unprotected Peptide Stapling. *J. Am. Chem. Soc.* **135**, 5946–5949 (2013).

20. Fairlie, D. P. & Dantas de Araujo, A. Review stapling peptides using cysteine crosslinking. *Biopolymers* **106**, 843–852 (2016).

21. Sokolowska, I., Ngounou Wetie, A. G., Woods, A. G. & Darie, C. C. Automatic determination of disulfide bridges in proteins. *J. Lab. Autom.* **17**, 408–16 (2012).

22. Timmerman, P., Beld, J., Puijk, W. C. & Meloen, R. H. Rapid and quantitative cyclization of multiple peptide loops onto synthetic scaffolds for structural mimicry of protein surfaces. *ChemBioChem* **6**, 821–824 (2005).

23. Zhang, F., Sadovski, O., Xin, S. J. & Woolley, G. A. Stabilization of folded peptide and protein structures via distance matching with a long, rigid cross-linker. *J. Am. Chem. Soc.* **129**, 14154–14155 (2007).

24. Wang, Y. *et al.* Application of Thiol-yne/Thiol-ene Reactions for Peptide and Protein Macrocyclizations. *Chem. - A Eur. J.* **23**, 7087–7092 (2017).

25. Vinogradova, E. V., Zhang, C., Spokoyny, A. M., Pentelute, B. L. & Buchwald, S. L. Organometallic palladium reagents for cysteine bioconjugation. *Nature* **526**, 687–691 (2015).

26. Muppidi, A. *et al.* Rational Design of Proteolytically Stable, Cell-Permeable Peptide-Based Selective Mcl-1 Inhibitors. *J. Am. Chem. Soc.* **134**, 14734–14737 (2012).

27. Wang, Y. & Chou, D. H. C. A Thiol-Ene Coupling Approach to Native Peptide Stapling and Macrocyclization. *Angew. Chemie - Int. Ed.* **54**, 10931–10934 (2015).

28. Grison, C. M. *et al.* Double quick, double click reversible peptide “stapling”. *Chem. Sci.* **8**, 5166–5171 (2017).

29. Zou, Y. *et al.* Convergent diversity-oriented side-chain macrocyclization scan for unprotected polypeptides. *Org Biomol Chem* **12**, 566–573 (2014).

30. Verhoork, S. J. M. *et al.* Tuning the binding affinity and selectivity of perfluoroaryl-stapled peptides by cysteine-editing. *Chem. - A Eur. J.* (2018). doi:10.1002/chem.201804163

31. Brunel, F. M. & Dawson, P. E. Synthesis of constrained helical peptides by thioether ligation:

application to analogs of gp41. *Chem. Commun.* **0**, 2552 (2005).

32. Aimetti, A. A., Shoemaker, R. K., Lin, C.-C. & Anseth, K. S. On-resin peptide macrocyclization using thiol–ene click chemistry. *Chem. Commun.* **46**, 4061 (2010).

33. Hoppmann, C., Schmieder, P., Heinrich, N. & Beyermann, M. Photoswitchable Click Amino Acids: Light Control of Conformation and Bioactivity. *ChemBioChem* **12**, 2555–2559 (2011).

34. Tian, Y. *et al.* Stapling of Unprotected Helical Peptides via Photo-Induced Intramolecular Thiol-Yne Hydrothiolation. *Chem. Sci.* **7**, 1–6 (2016).

35. Singh, R. & Goswami, T. Acid catalyzed 1, 2 Michael addition reaction: a viable synthetic route in designing fullerene core starlike macromolecule. *J. Phys. Org. Chem.* **21**, 225–236 (2008).

36. Aimetti, A. A., Shoemaker, R. K., Lin, C.-C. & Anseth, K. S. On-resin peptide macrocyclization using thiol-ene click chemistry. *Chem. Commun. (Camb).* **46**, 4061–3 (2010).

5 Exploring the Scope of Cysteine-Editing in Thiol-Perfluoroaryl Peptide Stapling

This chapter explores novel stapled peptide inhibitors of the interaction between MDM2/MDMX and the well-known tumour suppressor protein p53. The reactivation of p53 through disruption of the p53-MDM2/MDMX protein-protein interaction has been validated as a method of cancer treatment and is thus well studied. There are multiple stapled peptides that have been developed, most of which make use of the alkene metathesis stapling method. Stereochemistry may play an important role in the conformation of these peptides, as the key amino acids need to be pointing in the correct direction for optimal affinity. The thiol-fluoroaryl stapling method has been optimised for ring size, but not yet for cysteine stereochemistry. Here, the p53-MDM2 protein-protein interaction was selected as a model system, in which to probe cysteine-replacement. We investigated the replacement of L-cysteine with combinations of D-cysteine, homocysteine and penicillamine to examine the effect of i) stereochemistry, ii) cysteine homologation, and iii) adding steric bulk at the beta-carbon.

The work in this chapter has been published:

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.

Available at: https://doi.org/10.1002/chem.201804163

Introduction

5.1 The MDM2-p53 protein-protein interaction

Protein-protein interactions (PPIs) are commonly defined as physical contacts with molecular docking between proteins, occurring in cells or living organisms. This interaction should be specific, not accidental, and have a specific biological purpose.[1] The p53 tumour suppressor is a major regulator of the cell cycle and is activated in response to genotoxic stress, inducing cell cycle arrest and apoptosis in afflicted cells.[2] Loss of functional p53 through gene mutation or deletion is implicated in approximately half of all tumours.[3,4] The murine double minute 2 (MDM2) protein (also known as HDM2 in humans) is an important negative regulator of the tumour suppressor p53. In many p53 wild-type tumours, including brain, breast and soft tissue tumours,[5] overexpression of MDM2 leads to silencing of p53 and prevents its activation in response to cellular stress *e.g.* ionising radiation of DNA-lesions.[6] The interaction between MDM2 and p53 affects p53 in at least three different ways. Firstly, the transcriptional activity of p53 is inhibited by the binding of MDM2 to its transactivation domain.[7] Secondly, the E3 ubiquitin ligase function of MDM2 promotes p53 degradation.[8,9] Thirdly, MDM2 continuously moves between the nucleus and the cytoplasm, promoting the export of p53 from the nucleus upon binding.[10] In addition to interacting with p53, MDM2 can form a heterodimer with its homolog MDMX (also known as MDM4, or HDMX and HDM4 in humans), which leads to enhanced p53 ubiquitination activity.[11] Besides, MDMX is a negative regulator of p53 and overexpression of this protein can cause p53 silencing, like overexpression of MDM2.[12,13]

In the case of MDM2/MDMX overexpression, the disruption of the PPI between p53 and MDM2/MDMX proteins can reactivate p53 and induce cell cycle arrest and apoptosis in tumour cells.[14] There are several strategies under investigation that have validated this, including decreasing the cellular concentration of MDM2 with antisense oligodeoxynucleotides,[15] blocking the MDM2-mediated ubiquitination of p53,[16] utilising mimics of the negative regulator of MDM2 p14AR,[17] and finally directly disrupting the interaction using molecules that bind at the interface between p53 and MDM2.

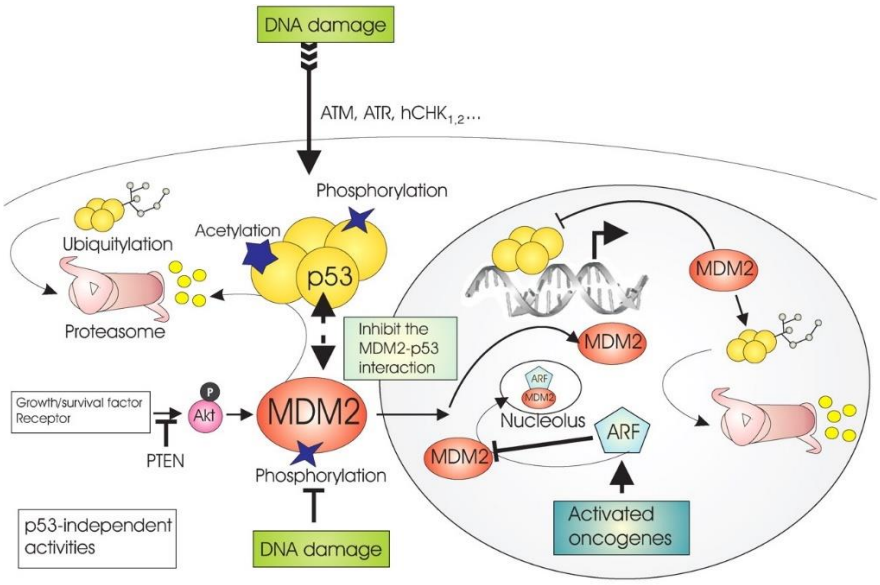


Figure 5.1: Inhibiting the p53–MDM2 interaction: an important target for cancer therapy.[18] Reproduced with permission from American Association for Cancer Research©.

The interaction between MDM2/MDMX and p53 takes places at the hydrophobic cleft of MDM2, which binds the amphipathic α-helix in the p53 N-terminal transactivation domain. The interaction is underpinned by hydrophobic and aromatic interactions of three key amino acids of p53, Phe[19], Trp[23], and Leu[26].[19] These three amino acids have been found to be invariant across species,[20] further supporting the functional importance of these residues.

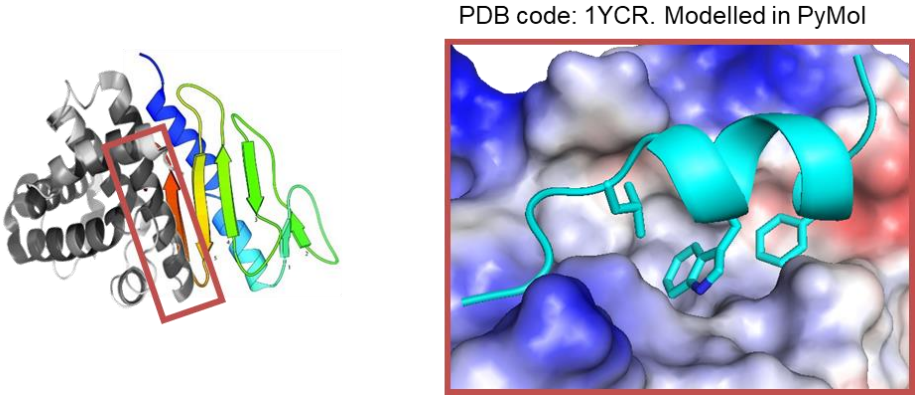


Figure 5.2: Protein-protein interaction between p53 N-terminal transactivation domain and hydrophobic cleft of MDM2.

Attempts to reactivate p53 by disrupting its interaction using small molecules, including spiro[isoindole-1,5'-isoxazolidin]-3(2H)-ones,[21] 1,5-disubstituted tetrazoles,[22] tri-substituted pyrroles,[23] and nutlins,[24] have been made.[25,26] While these molecules (Figure 5.3) show a wide range of chemical diversity, they have all been designed to mimic the previously mentioned key amino acid residues of p53. Although being successful at disrupting the interaction between p53 and MDM2, these small molecules often fail to disrupt the interaction with MDMX despite the structural similarity between MDM2 and MDMX.[27] When MDM2 alone is silenced, chemoresistance can occur for the therapeutic treatment through MDMX overexpression.[28] A potent dual inhibitor of MDM2/MDMX is required for effective p53 reactivation, however research into small molecules is very limited to date, due to distinct structural differences between both binding pockets.[29–31]

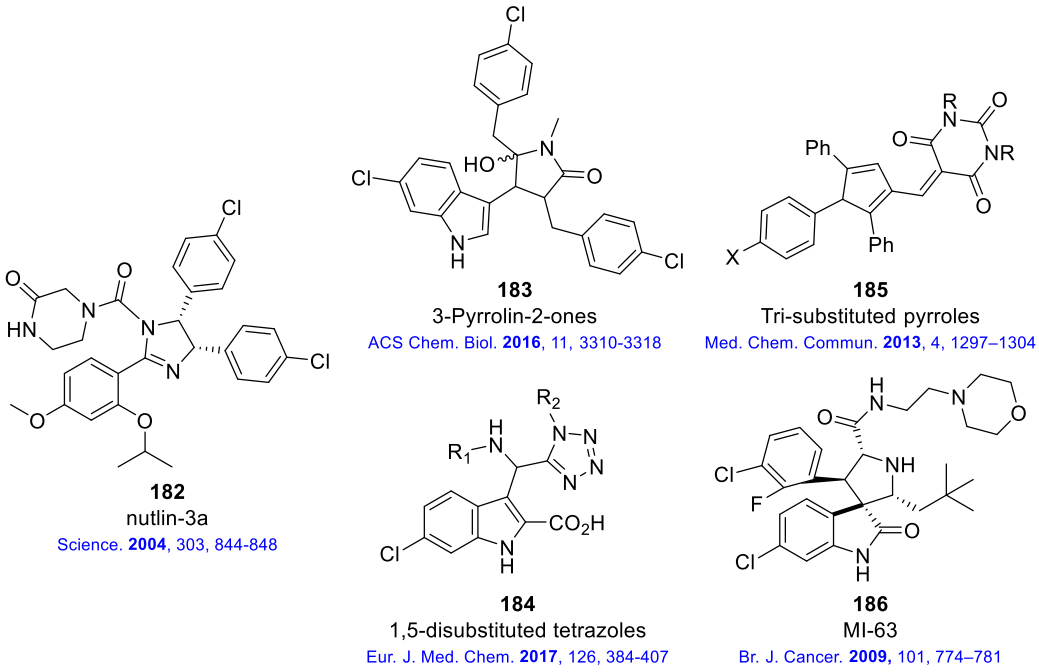


Figure 5.3: Examples of small molecule inhibitors of the MDM2/p53 protein-protein interactions.

Alternatively, peptides can form potential therapeutics able to modulate protein-protein interactions. Peptides may have several advantages over small molecules, as they are able to form 'native-like' interactions with the protein target. They are also capable of interacting over a larger surface area, which is often required for effective PPI disruption. Nevertheless, native peptides often suffer from low proteolytic stability and have a short circulating plasma

half-life (typically around 3-6 minutes).[32] However, the stability can be enhanced by several chemical modifications, including cyclisation, PEGylation,[33] lipidation,[34] N-methylation,[35] C-/N-terminal modifications,[36] peptidomimetics (e.g. peptoids or aza-peptides),[37,38] or using D-amino acids.[39] The introduction of a crosslinker or "staple" greatly improves the pharmacokinetic and conformational properties in some cases and makes peptides more suitable as potential therapeutics.[40] Advantages of peptide stapling include increased proteolytic stability, longer half-life time, conformational stability, improved absorption, and improved biological activity.[41–43]

The effectiveness of stapled α-helical peptides has already demonstrated in other PPI examples, including the NOTCH transactivation complex,[44] BH3-Bcl2 interaction,[45] and Ras-Sos interaction.[46]

Multiple different peptide stapling approaches have already been explored to inhibit the MDM2-p53 interaction, making this an ideal model system in which to study modification of the staple and its influence on the overall peptide stability into the optimal bio-active α-helical conformation. These SAH designs are based on the binding of p53 in the hydrophobic cleft on the N-terminal surface of MDM2 as previously described (Figure 5.2).

5.2 Hydrocarbon-stapled peptides

The stability of an α-helical conformation is largely dependent on hydrogen-bond formation in the protein or peptide backbone between a C=O and the NH of the *i+4* residue.[47] Ring-closing metathesis can also be used to replace this hydrogen bond by a C-C linkage (hydrogen bond surrogates), increasing helical stability.[48,49] The most common method used for peptide stapling typically employs the use of the all-hydrocarbon (alkene) linker developed by Grubbs and Blackwell.[50–52] This method originally used O-allyl serine residues, which were metathesised together. More recent improvements to the method employs the incorporation of unnatural (non-proteinogenic) α,α-disubstituted alkene-containing amino acids in to the peptide sequence, however, monosubstituted α-pentenylglycine was also shown to be a suitable substrate.[53,54] After synthesis of the linear peptide on-resin, the peptides are treated with Grubbs catalyst to form a macrocyclic peptide. Hydrogenation can then also be carried out prior to cleavage from the resin, to prevent problems with *E* + *Z* isomers (Figure 5.4). Advantages of this on-resin synthesis approach are the selectivity of reaction, limiting the amount of dimerisation or even polymerisation. The optimal positions to introduce a staple can be determined by staple scanning methodologies.[55] After determination of the position in the sequence, the alkene chain length and stereochemistry can be optimized, as explored by the Verdine group.[56,57] The optimal length between the residues that are to be linked is based on one of two helical turns, aligning the space where the linker fits. Likewise, stereochemistry was found to be important for improved alignment. Often a combination of (*R*)- and (*S*)-alkenyl amino acids stapled in a *i, i+4* (one turn) or *i, i+7* (two turns) fashion gives the most α-helical peptide conformation.

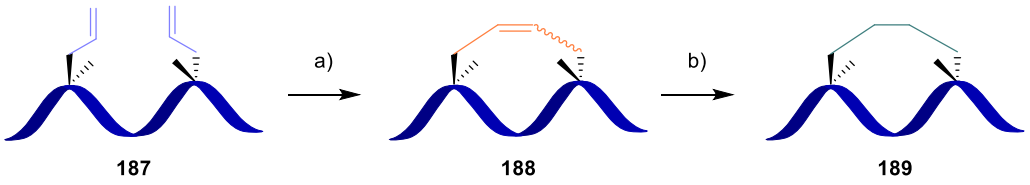


Figure 5.4: Schematic overview of hydrocarbon stapling; a) Grubbs Ring Closing metathesis; b) hydrogenation.

5.2.1 SAH-p53-8

The first series of stapled peptides to modulate the HDM2-p53 protein-protein interaction were developed in 2007 and were based on the Ac-LSQETFSDLWKLLPEN-NH_2 wild type p53 sequence. The only hydrocarbon stapled peptide that had the ability to induce cell death was SAH-p53-8 (Ac-QSQQTF(R_8)NLWRLL(S_5)QN-NH2), where (R_8) is an *R*-octenylalanine, and (S_5) *S*-pentenylalanine.[58] This peptide consisted of the highest percentage (85%) of α-helicity of this series, indicating that this could be an important property for MDM2 affinity in this specific case. In further experiments a 25-fold binding preference for HDMX compared to HMD2 was found for this peptide, which further enhances p53 reactivation in addition to the effect from HMD2 binding. SAH-p53-8 was also found to be more effective than the leading small molecule inhibitor Nutlin 3a,[59] confirming how promising peptides are for modulating protein-protein interactions. Computational studies of this peptide revealed that the carbon staple contributed to ca. 10% of the surface contact area between the peptide and MDM2. It has been proposed that this protects the hydrogen bond between Trp^{23} and Leu^{54} from solvent competition, which would improve the binding interaction.[60]

5.2.2 ATSP-7041

Following the initial success of the first series of stapled peptide, a new series was developed in 2013 by Chang *et al*.[28] These peptides were based on a sequence (pDI: Ac-LTFEHYWAQLTS-NH_2) identified by phage display, and was structurally distinct from the previously reported stapled peptide inhibitors.[61,62] ATSP-7041 (Ac-LTF(R_8)EYWAQ(Cba)(S_5)SAA-NH_2, where Cba is L-2-amino-4-cyanobutyric acid) was found to remain in the biologically active conformation and possessed favourable drug-like properties including efficient cell penetration, specific high-affinity binding to both target proteins, and excellent stability *in vivo*.[28] Clinical evaluation of this peptide has been announced, and the cellular uptake, intracellular targeting selectivity and p53- dependent cytotoxicity have been validated.[63] This peptide is currently in clinical trials by Aileron therapeutics, under the name of ALRN-6924. The peptide was well tolerated during the human phase 1 trial, and studies have now entered phase 2a clinical trials.[64]

5.2.3 Staplins

The staplins, also called sMTides, were inspired by a linear p53 peptide that was also identified by phage display (Ac-TSFAEYWNLLS-NH_2).[65] During development, asparagine was replaced with alanine as it (Asn) was found to disrupt the α-helical conformation and was not critical for binding to MDM2. The resulting stapled peptide sequence (sMTide-02/Staplin: Ac-TSF(R_8)EYWALL(S_5)-NH_2) had improved affinity for MDMX and increased p53 reactivation compared to the parent sequence. Further improvements were made by replacing the tryptophan residue by a 6-Cl modified (D- or L-)tryptophan.[66] This lead to another potent inhibitor sMTide-02A/Staplin-2 (Ac-TSF(R_8)EY(L-6Cl-W)ALL(S_5)-NH_2).[67,68]

5.3 Cysteine-stapled peptides

Disulfide bridges between two cysteine residues are prevalent in nature in both large proteins and small peptides. One role is to control the conformation/topology of complex secondary and tertiary structures. Peptides can contain one or more disulfide bonds that are essential for the peptide shape and function.[69] However, disulfide bridged peptides are sensitive to reducing conditions, limiting their potential as pharmaceuticals.

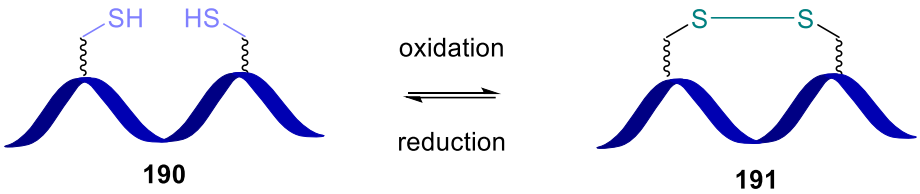


Figure 5.5: Disulfide formation is reversible.

There has been great interest in utilising the high nucleophilicity of the cysteine thiol-group in stapling approaches alternative to the natural disulfide bridge. Where natural disulfides are sensitive to reduction by glutathione (GSH) in cells, stapled peptides have been demonstrated to be more stable.[70] In general, two stapling approaches exist up to date; 1-component stapling where an intramolecular reaction takes place between a specially designed residue and a cysteine, and 2-component stapling, where an intermolecular reaction takes place between two cysteine residues and a linker molecule (Figure 5.6).

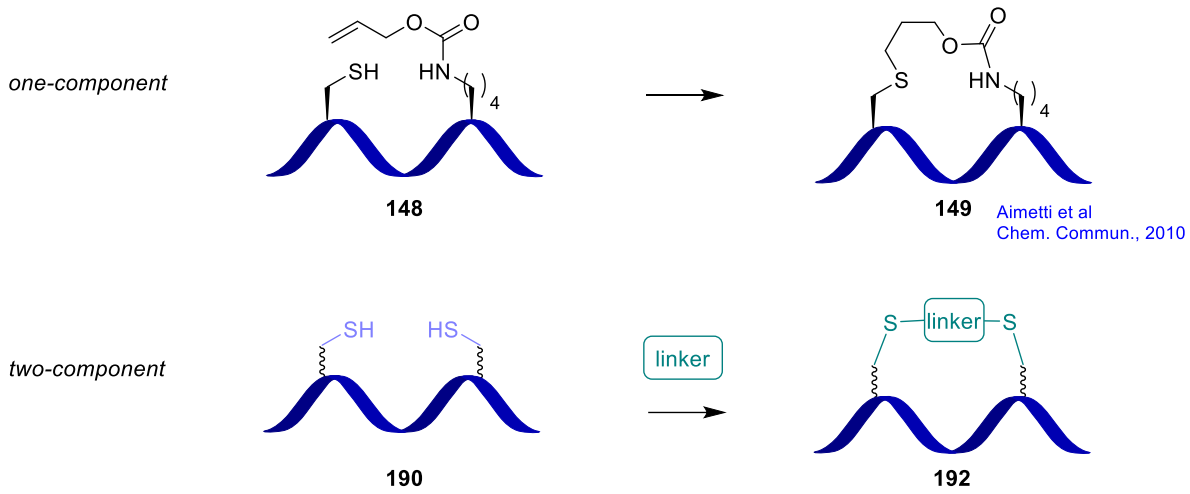


Figure 5.6: Cysteine stapling approaches

There are many two-component cysteine linkers available, which has recently been reviewed by Fairlie and Dantas de Araujo.[71] In 2005, Timmerman *et al.* published a method to cyclise cysteine-containing peptides with bromomethyl-benzene derivatives. Depending on the linker molecule this would result in a single, double, or triple loop peptide.[72] However, peptide stapling is often performed on α-helical peptides with the aim to constrain the peptide in this shape. The first thiol cross-linker was designed for an *i, i+11* motif by the Woolley group in 2007. Their crosslinker contained two sulfonate groups to increase water solubility and counteract the hydrophobic effect from the introduction of a rigid crosslinker.[73] More cysteine-linkers have been developed since, which are summarised in Figure 5.7 based on their stapling motif.[74–77] Worth mentioning is the dibromomaleimide-linker that has been developed by the Wilson group in 2017. In addition to cross-linking cysteines this linker can further be functionalised with an alkyne group on the nitrogen, allowing for further 'click' chemistry reactions.[78]

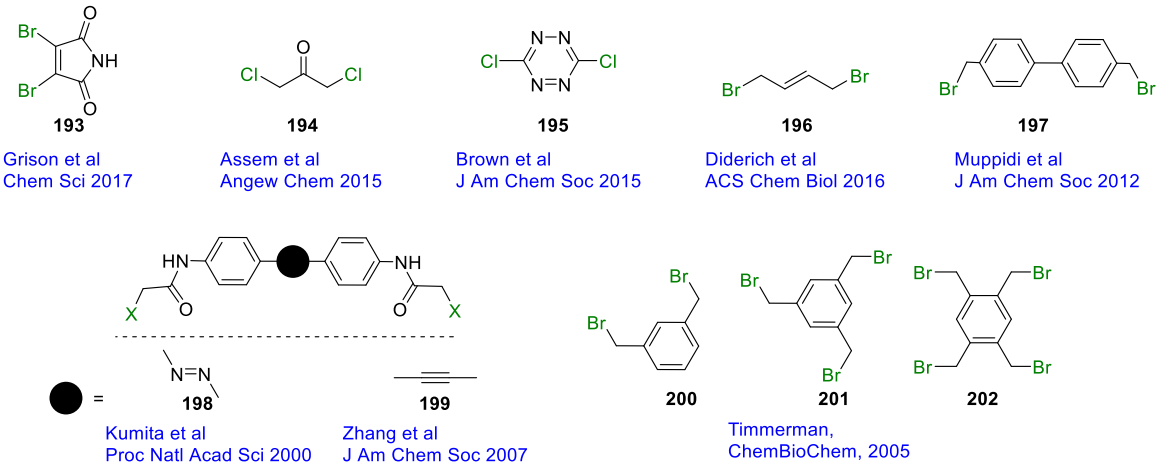
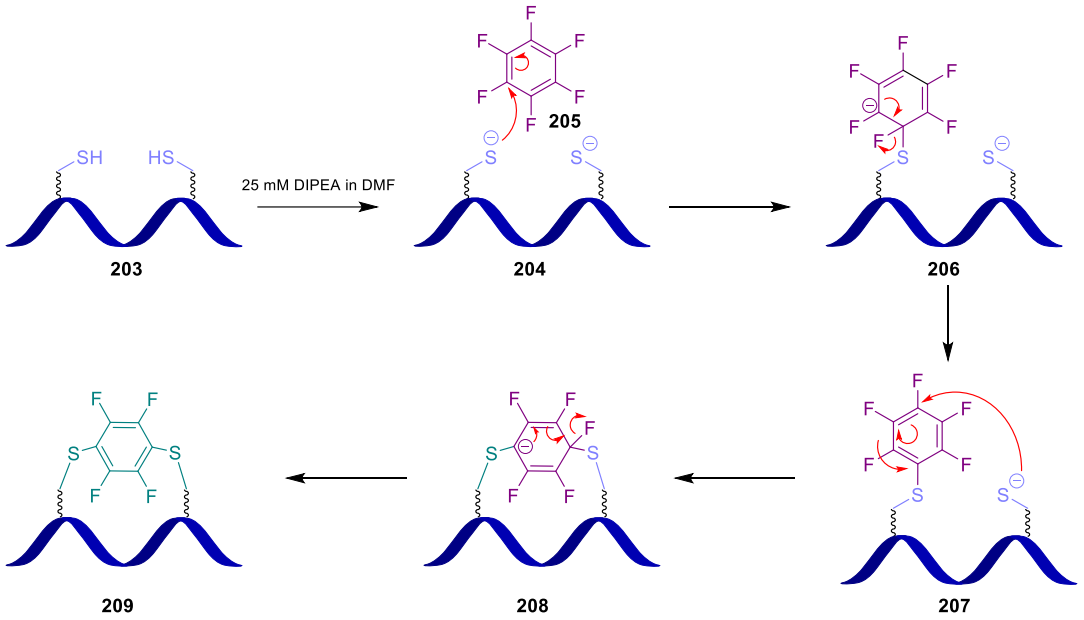


Figure 5.7: Linker molecules for two-component thiol-stapling reactions.

5.3.1 Perfluoro-aryl peptide stapling

Perfluoroaromatic reagents such as hexafluorobenzene and decafluorobiphenyl have been shown to react with sulfur-centred nucleophiles including cysteine, displaying a 1,4- and 4,4-disubstitution pattern respectively, at the perfluoroaromatic ring(s) (Scheme 5.1). A likely advantage of perfluoroaryl-stapling over disulfide bridge formation is the increased cellular stability under reducing conditions making the peptides more suitable for biological targets. This stapling method has been employed in the design of an HIV-1 capsid assembly peptide with improved receptor-binding, cell permeability, and proteolytic stability compared to the non-stapled peptide.[70] Concurrently, reactions of related perfluoro-heteroaromatic reagents e.g. pentafluoropyridine with amino acids containing alcohol sidechains has been reported.[79–81] Perfluoroaryl-stapling methods are potentially widely applicable to bioconjugation and peptide stapling. Advantageously, it does not require: 1) the use of synthetic amino acids; 2) protection of other amino acid residues; or 3) complicated chemical procedures for introduction of the cross-linking motif. Additionally, introduction of a bi-fluoroaryl staple has recently been shown to enhance blood-brain barrier transport compared to the non-macrocyclic analogue.[82]



Scheme 5.1: Perfluoroaryl-stapling of cysteines.

Further advancements for the application of this peptide stapling technique have been developed as it was less efficient for the macrocyclization of longer and unstructured peptides. Enzymatic Glutathione S-transferase (GST)-catalysed reactions have been introduced to overcome this problem by building in a GSH motif. In addition, increased regioselectivity was observed for a peptide containing more than two cysteines, thus improving the applicability to a wider variety of peptide sequences.[83] The applicability of the perfluoroaryl-stapling technique has been demonstrated in cell biological assays by the synthesis of an analog of H2 relaxin, which exhibited increased hydrophobicity compared to its wildtype.[84]

Besides cysteine arylation, a method utilising the N-arylation of lysine has been reported. The advantage of lysine arylation over cysteine arylation is the chemical stability of the N-C (aryl) bond. However, the reaction rate observed for lysine is lower than that of cysteine arylation, matching the lower nucleophilicity of lysine. Decreased degradation of a lysine-stapled model peptide was also observed compared to its cysteine equivalent.[85]

5.4 Exploring the scope of cysteine-peptide stapling

5.4.1 Aim of this chapter

While the distance between the cysteine residues has been explored in a diversity-oriented side-chain macrocyclisation scan for unprotected peptides,[86] little emphasis has been put on the nature of cysteine residues, with the exception of the introduction of homocysteine.[78]

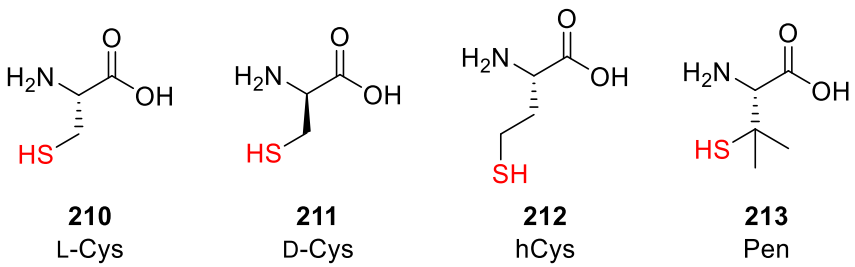
5.4.2 Approach

In analogy to the hydrocarbon approach, where a combination of *(R)*- and *(S)*-stereochemistry is used, we have explored the consequences of the replacement of L-cysteine by D-cysteine for stereochemical effects, homocysteine for a potential increase in flexibility, and penicillamine for beta-carbon steric effects. This was expected to play an important role in target binding affinity and selectivity.

Results and discussion

5.5 Synthesis of stapled α-helical peptides

The molecular interaction between the tumour suppressor protein p53 and MDM2 has been well characterised and validated as a drug target in various stapled peptide systems (see Section 5.1). For this reason, it formed an excellent model system to explore our cysteine-editing approach. The twelve amino acid peptide NH_2-LTFEHYWAQLTS-$CONH_2$ (**214**, pDI peptide) identified by phage display[62] used in the design of our SAHs was used as a starting point for our studies. The effect of pDI peptide on the MDM2-p53 interaction has been previously explored using photochemical stapling approaches[87] and lysine nitrogen arylation stapling approaches.[85] However, in both examples the stereochemistry of the residues involved in the peptide staple was not considered. Three key amino acid residues, Phe, Trp, and Leu, have been identified as being important for MDM2-binding and biological activity, where the spatial positioning of these residues is crucial.[19] Theoretically, positioning of the staple could influence the position of the key residues in space, by slightly deforming the α-helical shape, depending on how well the staple 'fits' between the two stapled cysteine residues.



210 L-Cys | **211** D-Cys | **212** hCys | **213** Pen

Figure 5.8: Cysteine analogues used in this study.

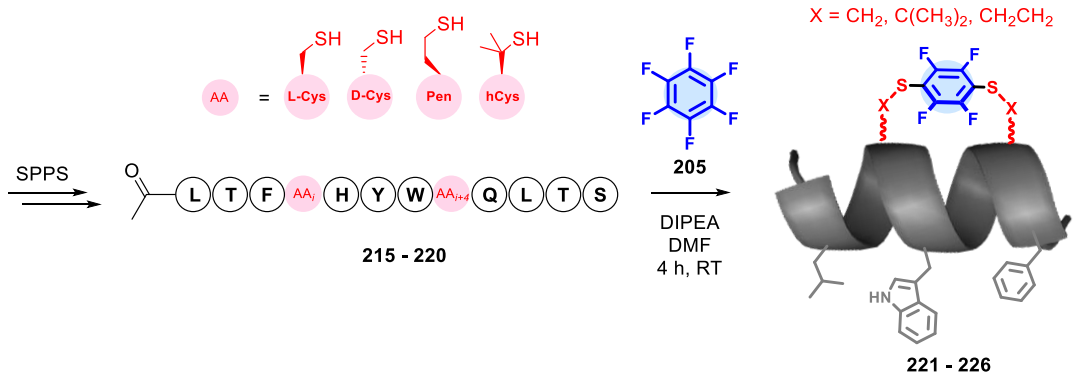
Residues from the pDI peptide that had been shown to be tolerant to substitution,[87] e.g. Ala and Glu, were replaced with cysteine (analogues), with a relative spacing of four amino acids (*i, i+4*) in between, corresponding to equivalent positions at neighbouring turns on the α-helix. Peptides were synthesised on ProTide Rink amine resin (0.55 mmol/g) following standard automated microwave-assisted Fmoc-solid phase peptide synthesis conditions, affording the

following general sequence: Ac-LTF(**AA**$_i$)HYW(**AA**$_{i+4}$)QLTS-CONH$_2$, where **AA** represents a single cysteine analogue residue (Figure 5.8).

Table 5.1: Crude purity of linear peptides Ac-LTF(AA$_i$)HYW(AA$_{i+4}$)QLTS-CONH$_2$ as determined by LC-MS (AUC at 215 nm). Note: hCys = homocysteine; Pen = penicillamine.

Compound	AA$_i$	AA$_{i+4}$	Crude purity
pDI peptide (214)	L-Glu	L-Ala	68%
Peptide 215	L-Cys	L-Cys	57%
Peptide 216	L-Cys	D-Cys	70%
Peptide 217	D-Cys	L-Cys	68%
Peptide 218	D-Cys	D-Cys	76%
Peptide 219	L-hCys	L-hCys	52%
Peptide 220	L-Pen	L-Pen	94%

The crude purity of all synthesised peptides ranged between 52-94% (Table 5.1). Surprisingly, peptide **220**, containing two penicillamine residues, yielded the highest crude purity, despite increased steric hindrance from its side chain. However, this could be explained through disfavouring the formation of a secondary peptide structures (e.g. β-sheet formation) on resin.[88] The introduced steric hindrance from the penicillamine residues could prevent this formation, thus improving the quality of the peptide synthesis. Likewise, the earlier introduction of D-cysteine in peptide **216** and peptide **218** could lead to destabilising effects of the secondary structure, explaining the slightly higher crude purity for those peptides. All crude purities were deemed acceptable for advancement to the stapling reaction without prior purification steps.



Scheme 5.2: Synthesis of perfluoroaryl-stapled peptides.

Perfluoroaryl-stapling was performed based on reaction conditions developed by Pentelute and co-workers, using hexafluorobenzene (**205**, HFB) as a cross-linking reagent without the need for reduction of any possible disulfide bridges (none are observed by LCMS). Their original method utilised TRIS buffer,[70] however, we found DIPEA in DMF to be equally effective, with the advantage of being volatile, simplifying isolation of the stapled peptides. All reactions took place under relatively mild conditions (Scheme 5.2, 25 mM DIPEA in DMF, room temperature 4 h), and were isolated in moderate yields (46-75%, Table 5.2). There are two main factors that are expected to influence the efficiency of the stapling reaction, helicity and nucleophilicity. The different cysteine analogues can influence the α-helicity, which in turn can alter the alignment between the two thiol-groups that are to be stapled. A better cysteine alignment, is expected to have a better conversion to the stapled product. Alternatively, if the structure is disrupted and no longer an α-helix, the peptide may be more flexible, which could also enhance a macrocyclisation reaction. Nucleophilicity could also play a role in the reactivity of the thiol-groups of each cysteine analogue, with penicillamine being the least nucleophilic and homocysteine the most nucleophilic. Surprisingly, penicillamine containing peptide **226**, had the largest conversion, suggesting nucleophilicity is not the main driving force of the 1,4-disubstitution reaction.

Table 5.2: Crude purity of perfluoroaryl-stapled peptides as determined by LC-MS (AUC at 215 nm).

Compound	AA_i	AA_{i+4}	Crude conversion
Peptide 221	L-Cys	L-Cys	46%
Peptide 222	L-Cys	D-Cys	60%
Peptide 223	D-Cys	L-Cys	71%
Peptide 224	D-Cys	D-Cys	53%
Peptide 225	L-hCys	L-hCys	67%
Peptide 226	L-Pen	L-Pen	75%

All stapled peptides **221**-**226**, peptide **215** (L-L linear), and pDI peptide (**214**) were purified by preparative-HPLC. Stapled peptides showed poor solubility in water, complicating the purification process. Purification was optimised by gradually increasing the starting percentage of the organic component of the mobile phase. Preparation of prep-HPLC samples proved difficult as the peptide would start to precipitate over time if left in solution. The optimal dissolution solvent for sample preparation was a 9:1 ratio of acetonitrile and water. The solubility problems of the peptide required a fast (short) prep HPLC method that still separated the crude peptide mixture sufficiently for accurate fraction collection. Initial purification was performed using a 10% to 100% MeOH (0.1% FA) gradient over 25 minutes at a flow rate of 3 mL/min. The first three minutes were kept isocratic at 10% MeOH so the crude peptide mixture could settle on the column. A gradient was applied from the 3rd minute until the 18th, increasing the organic phase to 100%. Elution was kept at 100% MeOH to ensure complete removal of the sample from the column. The system was then allowed to re-equilibrate back to 10% MeOH, ending the run at 25 minutes. UV spectra were recorded concomitantly at a wavelength of 215 nm, 254 nm, and 280 nm. It was chosen to continue purification using 280 nm as the background signals (UV absorbance of the mobile phase) were less apparent, leading to more accurate triggering of the fraction collector.

This method allowed for adequate purification of the L- or D-cysteine containing peptides, however precipitation would occur in the sample vials after multiple injections. The peptides all eluted after 20 minutes, indicating the requirement for a high percentage for the organic solvent. To improve and speed up purification, new methods with higher organic phase (MeOH, 0.1% FA) starting percentages were developed, including 60-100%, 70-100%, and 80-100% gradient methods. For purification of the D- or L-cysteine containing peptides the 70-100% method was most appropriate, while the 80-100% method was more suitable for the homocysteine and penicillamine containing peptides, as they were more hydrophobic. All peptides were obtained in excellent purity (>95%).

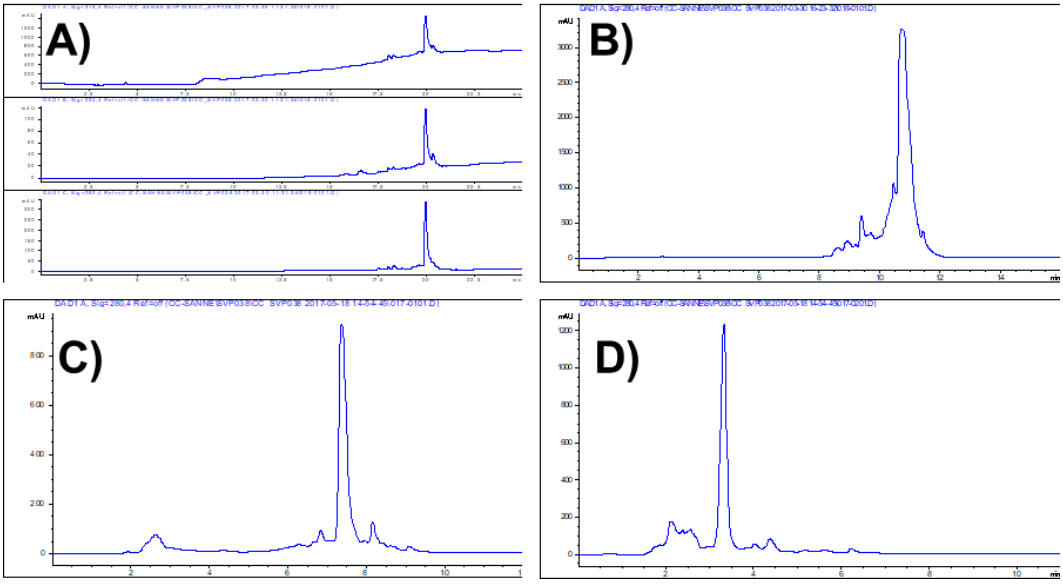


Figure 5.9: Peptide purification optimisation process. HPLC chromatograms of peptide **222**. (Ac-LTF(L-Cys)HYW(D-Cys)QLTS-$CONH_2$, HFB-stapled). A) 0-3 minutes 10% MeOH, 5-20 minutes gradient to 100% MeOH, 20-22 minutes 100% MeOH, 23-25 re-equilibrate to 60% MeOH. Total run time 25 minutes. B) 0-5 minutes 60% MeOH, 5-10 minutes gradient to 100% MeOH, 10-12 minutes 100% MeOH, 12-16 re-equilibrate to 60% MeOH. Total run time 16 minutes. C) 0-2 minutes 70% MeOH, 2-8 minutes gradient to 100% MeOH, 8-9 minutes 100% MeOH, 9-12 minutes re-equilibrate to 70% MeOH. Total run time 12 minutes. D) 0-2 minutes 80% MeOH, 2-7 minutes gradient to 100% MeOH, 7-9 minutes 100% MeOH, 9-11 minutes re-equilibrate to 80% MeOH. Total run time 11 minutes.

5.6 Biological evaluation

Following the successful synthesis, stapled peptides **221**-**226** and non-stapled reduced disulfide **215**, were measured for their binding affinity for both MDM2 and MDMX, to evaluate the effect of cysteine editing on the dissociation constant (K_d). All biological assays were performed in collaboration with Professor Martin Noble and Dr. Claire Jennings from the Northern Institute for Cancer Research at Newcastle University, however for completion, they will be discussed here briefly. Surface Plasmon Resonance (SPR) was used for the determination of the binding affinity of peptide **215**, and peptides **221**-**226** for MDM2 and MDMX. During SPR the angle of reflection of a light source is measured (Figure 5.10), which changes upon binding of the compound of interest to the target protein (in this case GST-MDM2 (17-125) and GST-MDMX (17-125)). Glutathione S-transferase (GST) fusion protein constructs for MDM2 and MDMX were used to enable affinity purification of the proteins. pDI peptide and small molecule nutlin 3a were used as positive controls. Dissociation constants (K_d) of the peptides from the protein constructs were measured to determine binding affinity (Table 5.3).

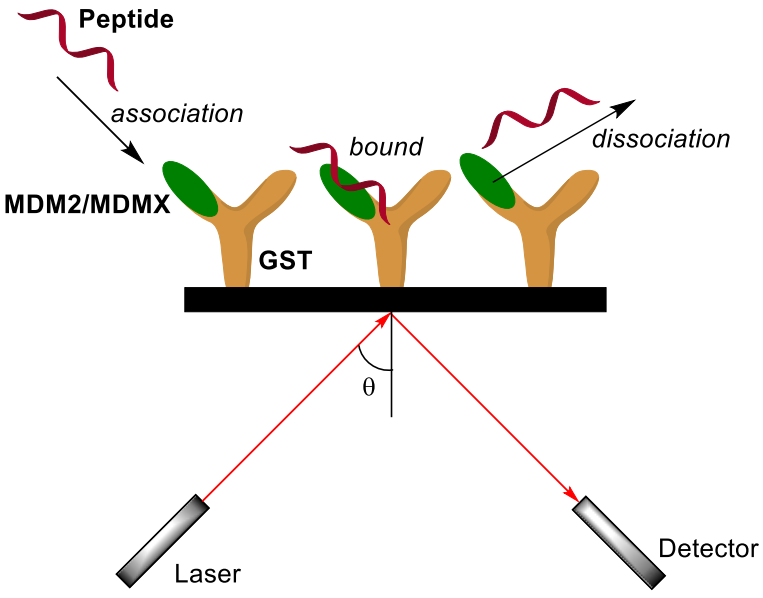
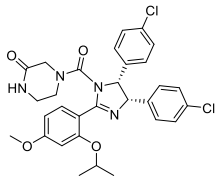
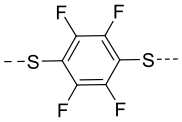


Figure 5.10: Schematic overview of surface plasmon resonance (SPR).

Table 5.3: Dissociation constants (K_d) for peptides **215**, **221**-**226** (Ac-LTF(AA_i)HYW(AA_{i+4})QLTS-$CONH_2$) binding to MDM2/MDMX attained from an SPR assays and comparison with control inhibitor nutlin 3a and pDI peptide. Note: hCys = homocysteine, Pen = Penicillamine. Chi2 = The Chi Square distribution, the distribution of the sum of squared standard normal derivatives; RU = response units.

Compound	AA_i	AA_{i+4}	Cross-linker	K_d / μM		Chi² (RU²)	
				MDM2	MDMX	MDM2	MDMX
PDI peptide (214)	L-Glu	L-Ala	-	0.04	0.02	1.36	0.82
Nutlin 3a (182) 	-	-	-	0.17	ND	0.24	ND
Peptide 215	L-Cys	L-Cys	-	0.18	0.18	1.49	0.40
Peptide 221	L-Cys	L-Cys		1.02	0.44	2.37	0.81
Peptide 222	L-Cys	D-Cys		0.41	0.37	3.08	0.99
Peptide 223	D-Cys	L-Cys		1.70	0.40	3.69	6.49
Peptide 224	D-Cys	D-Cys		0.22	0.06	1.71	0.86
Peptide 225	L-hCys	L-hCys		0.14	0.15	4.07	2.91
Peptide 226	L-Pen	L-Pen		16.90	9.89	10	8.17

A higher K_d indicates faster dissociation from the ligand-protein (MDM2 or MDMX) complex and thus lower affinity. All perfluoro-aryl stapled peptides displayed low-to-moderate micromolar affinities, albeit higher dissociation constants are observed than for PDI peptide. However, in a biological setting, the stapled peptides **221**-**226** could show enhanced activity through increased stability.[70] Indeed, the proteolytic stability of peptide **221** versus unstapled pDI peptide (**214**) was found to be improved after treatment with α-chymotrypsin (with a half-life of 230 ± 20 min and 58 ± 6 min respectively), as measured using HPLC with UV detection at 280 nM of the remaining peptide (Durham University).

The primary objective of this study was to demonstrate how cysteine-editing affects biological activity (target affinity and selectivity) in the MDM2/MDMX-p53 model system. Indeed, differences in affinity can be observed between the different peptides, as well as differences in selectivity for MDM2/MDMX. Reversing stereochemistry of both L-cysteines to D-cysteines increased the affinity for both MDM2 (~5-fold) and MDMX (>7-fold).

Noteworthy is the drastic decrease in affinity of peptide **226** (Pen), compared to the peptides lacking the gem-dimethyl group at the β-position. Interestingly, this is in stark contrast with cysteine homologation (hCys) results, with peptide **225** achieving the highest affinity for MDM2 of all synthesised stapled peptides **221**-**226**. The observed SAR may be related to the geometric constraints imposed by the cysteine analogue, and the respective positioning of the perfluoroaryl cross-linker. Peptide **226** could have lower affinity because the methyl-groups block close contact with the receptor. In line with that would be the increased affinity of peptide **225**, where the staple may have less steric or electronic hindrance with the receptor. These possibilities were further studied by *in silico* modelling (Chapter 5.6.1).

Alternatively, the α-helical structure of peptide **226** could be distorted, and not actually be in an α-helical conformation at all. The synthesis of perfluoroaryl stapled peptide **226** was surprisingly efficient (75%) compared to the other stapled peptides, which could indicate a macrocyclic shape instead of an α-helical stapled structure. Overall, the helicity of the peptide is likely to be important for the relative positions of key amino acids side chains, Phe, Trp, and Leu.

Differences in selectivity for MDM2 and MDMX could also be observed. Where unstapled peptide **215** has equipotent affinity for each isoform (K_d = 0.18), introduction of the perfluoroaryl-linker biases selectivity more than 2-fold towards MDMX. Interestingly, differences in selectivity are not observed for each of the peptides, as hCys containing peptide **225** displays similar affinities for each isoform (K_d = 0.14 and 0.15). In general, MDMX appeared to be more tolerant to ‘cysteine-editing’ than MDM2.

5.6.1 *In silico* modelling

Computational modelling experiments were carried out in collaboration with Dr. Andrew Leach from Liverpool John Moores University. Homology modelling was used based on the sequence and chain B from the structure of human MDM2 in complex with the reported high affinity pDI peptide (PDB code: 3G03). Molecular dynamics simulations were performed to obtain the free peptides **215**-**220** and stapled peptides **221**-**226**. To predict the efficiency in which the peptides could bind to MDM2, the geometries of the key amino acid residues were calculated (Figure 5.11).

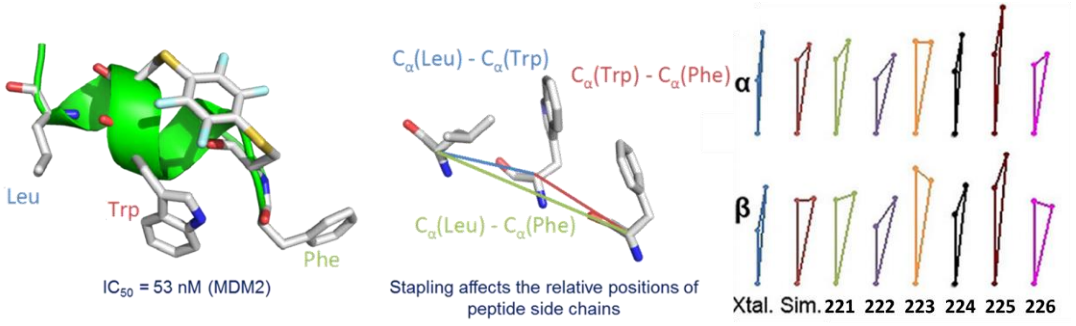


Figure 5.11: Model of perfluoroaryl-stapled peptide **221** and the geometries of the three key amino acid residues. Xtal= from X-ray structure 3G03, Sim = simulated pDI peptide.

The distances between the C_α and C_β of Phe, Trp, and Leu were calculated for each of the peptides, and were compared with the X-ray structure of pDI peptide bound to MDM2, and its simulated model. The initial analysis showed that the geometries of the triangles were rather similar because C_α-C_β bonds are pointing in the same directions, indicating the importance of the positioning for a tight interaction with the receptor. To compare how well the geometries would be retained by the free stapled peptides, the average distances were compared with those observed in the published pDI-MDM2 complex, making use of the root-mean-square deviation (RMSD). A lower RDSM indicates a better fit to the model. The double D-Cys-containing stapled peptide **224** was predicted to retain the required pharmacophoric arrangement better than even the native peptide (RMSD = 1.0). The next best was predicted to be compound **222** (RMSD=1.2), followed by **226** and **221**. Peptide **224** was correctly identified as the best of the analogues in which only stereochemistry is varied, however, the

structural variations in which methylation or homologation have been introduced do not seem to be correctly ranked, indicating the presence of other interactions influencing affinity.

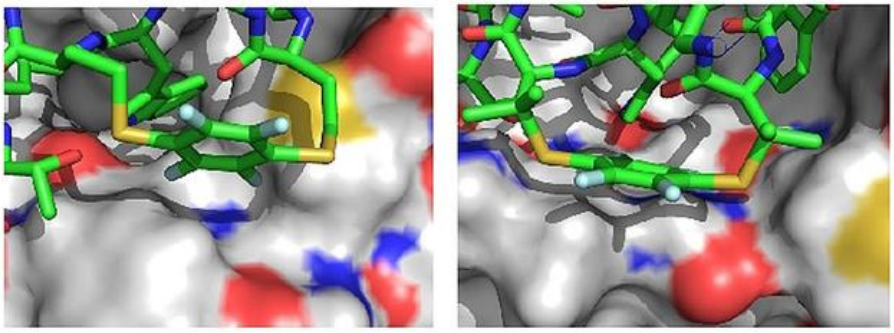


Figure 5.12: Simulated models of peptide **225** and **226**, providing possible interactions with MDM2.

Additional studies were performed to provide insight in hCys peptide **225** and Pen peptide **226**, where the peptide – MDM2 complex was simulated to investigate any extra contacts made by these two linkers that could explain the observed binding. The final snapshot is shown in Figure 5.12 and reveals that the linker of peptide **225** can form additional hydrophobic interactions with the protein. For peptide **226**, one of the methyl groups intrudes with the positioning of the peptide, preventing continuous contact of the linker. Overall, these insights help to explain the differential measured binding affinities following cysteine-editing.

5.7 Conclusion

In this chapter we have demonstrated that the conformational properties of a peptide can be modified by the nature (size and stereochemistry) of the thiol groups that are to be cross-linked. We have shown that this can influence biological activity (target affinity and selectivity) within the p53-MDM2/MDMX model system. These findings can be expanded to other two-component disulfide stapling approaches (*e.g.* a different cross-linker) and one-component thiol-ene reactions. Additionally, a similar cysteine-editing approach can be introduced during the development of SAH peptides for different protein-protein interactions. Furthermore, potential additional *in vitro* stabilising effects are of interest, as non-canonical amino acids are typically more resistant to proteolytic degradation.[39]

This work can also be expanded to the synthesis of non SAH-systems, like cyclotides (disulfide rich mini proteins). While in this case the nature of the cysteine will not be important for helicity, it can still influence overall peptide or protein conformation. This can impact biological activity as demonstrated in the SAH model system.

5.8 References

1. De Las Rivas, J. & Fontanillo, C. Protein-protein interactions essentials: key concepts to building and analyzing interactome networks. *PLoS Comput. Biol.* **6**, e1000807 (2010).

2. Lane, D. P. p53, guardian of the genome. *Nature* **358**, 15–16 (1992).

3. Nigro, J. M. *et al.* Mutations in the p53 gene occur in diverse human tumour types. *Nature* **342**, 705–708 (1989).

4. Hollstein, M., Sidransky, D., Vogelstein, B. & Harris, C. C. p53 mutations in human cancers. *Science* **253**, 49–53 (1991).

5. Toledo, F. & Wahl, G. M. Regulating the p53 pathway: in vitro hypotheses, in vivo veritas. *Nat. Rev. Cancer* **6**, 909–923 (2006).

6. Oliner, J. D., Kinzler, K. W., Meltzer, P. S., George, D. L. & Vogelstein, B. Amplification of a gene encoding a p53-associated protein in human sarcomas. *Nature* **358**, 80–83 (1992).

7. Momand, J., Zambetti, G. P., Olson, D. C., George, D. & Levine, A. J. The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. *Cell* **69**, 1237–45 (1992).

8. Wu, X., Bayle, J. H., Olson, D. & Levine, A. J. The p53-mdm-2 autoregulatory feedback loop. *Genes Dev.* **7**, 1126–32 (1993).

9. Chène, P. Inhibition of the p53-MDM2 Interaction: Targeting a Protein-Protein Interface. *Mol. Cancer Res.* **2**, (2004).

10. Roth, J., Dobbelstein, M., Freedman, D. A., Shenk, T. & Levine, A. J. Nucleo-cytoplasmic shuttling of the hdm2 oncoprotein regulates the levels of the p53 protein via a pathway used by the human immunodeficiency virus rev protein. *EMBO J.* **17**, 554–64 (1998).

11. Vousden, K. H. & Lu, X. Live or let die: the cell's response to p53. *Nat. Rev. Cancer* **2**, 594–604 (2002).

12. Danovi, D. *et al.* Amplification of Mdmx (or Mdm4) Directly Contributes to Tumor Formation by Inhibiting p53 Tumor Suppressor Activity. *Mol. Cell. Biol.* **24**, 5835–5843 (2004).

13. Shadfan, M., Lopez-Pajares, V. & Yuan, Z.-M. MDM2 and MDMX: Alone and together in regulation of p53. *Transl. Cancer Res.* **1**, 88–89 (2012).

14. Yee, K. S. & Vousden, K. H. Complicating the complexity of p53. *Carcinogenesis* **26**, 1317–1322 (2005).

15. Yu, R. Z. *et al.* Pharmacokinetics and pharmacodynamics of an antisense phosphorothioate oligonucleotide targeting Fas mRNA in mice. *J. Pharmacol. Exp. Ther.* **296**, 388–95 (2001).

16. Yang, Y. *et al.* Small molecule inhibitors of HDM2 ubiquitin ligase activity stabilize and activate

p53 in cells. *Cancer Cell* **7**, 547–559 (2005).

17. Midgley, C. A. *et al.* An N-terminal p14ARF peptide blocks Mdm2-dependent ubiquitination in vitro and can activate p53 in vivo. *Oncogene* **19**, 2312–2323 (2000).

18. Moll, U. M. & Petrenko, O. The MDM2-p53 interaction. *Mol. cancer Res.* **1**, 1001–1008 (2003).

19. Kussie, P. H. *et al.* Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* **274**, 948–53 (1996).

20. Lin, J., Chen, J., Elenbaas, B. & Levine, A. J. Several hydrophobic amino acids in the p53 amino-terminal domain are required for transcriptional activation, binding to mdm-2 and the adenovirus 5 E1B 55-kD protein. *Genes Dev.* **8**, 1235–46 (1994).

21. Giofrè, S. V *et al.* Synthesis of spiro[isoindole-1,5'-isoxazolidin]-3(2H)-ones as potential inhibitors of the MDM2-p53 interaction. *Beilstein J. Org. Chem.* **12**, 2793–2807 (2016).

22. Surmiak, E. *et al.* Rational design and synthesis of 1,5-disubstituted tetrazoles as potent inhibitors of the MDM2-p53 interaction. *Eur. J. Med. Chem.* **126**, 384–407 (2017).

23. Blackburn, T. J. *et al.* Diaryl- and triaryl-pyrrole derivatives: inhibitors of the MDM2-p53 and MDMX-p53 protein-protein interactionsdaggerElectronic supplementary information (ESI) available: Experimental details for compound synthesis, analytical data for all compounds and in. *Medchemcomm* **4**, 1297–1304 (2013).

24. Vassilev, L. T. *et al.* In Vivo Activation of the p53 Pathway by Small-Molecule Antagonists of MDM2. *Science (80-.).* **303**, (2004).

25. Shaikh, M. F. *et al.* Emerging Role of MDM2 as Target for Anti-Cancer Therapy: A Review. *Ann. Clin. Lab. Sci.* **46**, 627–634 (2016).

26. Wang, S., Zhao, Y., Aguilar, A., Bernard, D. & Yang, C.-Y. Targeting the MDM2-p53 Protein-Protein Interaction for New Cancer Therapy: Progress and Challenges. *Cold Spring Harb. Perspect. Med.* **7**, a026245 (2017).

27. Shvarts, A. *et al.* Isolation and Identification of the Human Homolog of a New p53-Binding Protein, Mdmx. *Genomics* **43**, 34–42 (1997).

28. Chang, Y. S. *et al.* Stapled α−helical peptide drug development: A potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy. *Proc. Natl. Acad. Sci.* **110**, E3445–E3454 (2013).

29. Graves, B. *et al.* Activation of the p53 pathway by small-molecule-induced MDM2 and MDMX dimerization. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 11788–93 (2012).

30. Lee, J. H. *et al.* Novel Pyrrolopyrimidine-Based α-Helix Mimetics: Cell-Permeable Inhibitors of Protein−Protein Interactions. *J. Am. Chem. Soc.* **133**, 676–679 (2011).

31. Soares, J. *et al.* DIMP53-1: a novel small-molecule dual inhibitor of p53-MDM2/X interactions

with multifunctional p53-dependent anticancer properties. *Mol. Oncol.* **11**, 612–627 (2017).

32. Fosgerau, K. & Hoffmann, T. Peptide therapeutics: current status and future directions. *Drug Discov. Today* **20**, 122–128 (2015).

33. Lawrence, P. B. & Price, J. L. How PEGylation influences protein conformational stability. *Curr. Opin. Chem. Biol.* **34**, 88–94 (2016).

34. Ward, B. P. *et al.* Peptide lipidation stabilizes structure to enhance biological function. *Mol. Metab.* **2**, 468–479 (2013).

35. Chatterjee, J., Gilon, C., Hoffman, A. & Kessler, H. N-Methylation of Peptides: A New Perspective in Medicinal Chemistry. *Acc. Chem. Res.* **41**, 1331–1342 (2008).

36. Marino, G., Eckhard, U. & Overall, C. M. Protein Termini and Their Modifications Revealed by Positional Proteomics. *ACS Chem. Biol.* **10**, 1754–1764 (2015).

37. Qvit, N., Rubin, S. J. S., Urban, T. J., Mochly-Rosen, D. & Gross, E. R. Peptidomimetic therapeutics: scientific approaches and opportunities. *Drug Discov. Today* **22**, 454–462 (2017).

38. Vagner, J., Qu, H. & Hruby, V. J. Peptidomimetics, a synthetic tool of drug discovery. *Curr. Opin. Chem. Biol.* **12**, 292–296 (2008).

39. Feng, Z. & Xu, B. Inspiration from the mirror: D-amino acid containing peptides in biomedical approaches. *Biomol. Concepts* **7**, 179–187 (2016).

40. Jamieson, A. & Robertson, N. Regulation of protein-protein interactions using stapled peptides. *Reports Org. Chem.* **5**, 65 (2015).

41. Acar, H. *et al.* Self-assembling peptide-based building blocks in medical applications. *Adv. Drug Deliv. Rev.* **110**–**111**, 65–79 (2017).

42. Vlieghe, P., Lisowski, V., Martinez, J. & Khrestchatisky, M. Synthetic therapeutic peptides: science and market. *Drug Discov. Today* **15**, 40–56 (2010).

43. Lau, Y. H., De Andrade, P., Wu, Y. & Spring, D. R. Peptide stapling techniques based on different macrocyclisation chemistries. *Chemical Society Reviews* **44**, 91–102 (2015).

44. Moellering, R. E. *et al.* Direct inhibition of the NOTCH transcription factor complex. *Nature* **462**, 182–8 (2009).

45. Walensky, L. D. *et al.* Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. *Science* **305**, 1466–70 (2004).

46. Patgiri, A., Yadav, K. K., Arora, P. S. & Bar-Sagi, D. An orthosteric inhibitor of the Ras-Sos interaction. *Nat. Chem. Biol.* **7**, 585–7 (2011).

47. Qian, H. & Schellman, J. Helix-coil theories: a comparative study for finite length polypeptides. *J. Phys. Chem.* **96**, 3987–3994 (1992).

48. Wang, D., Chen, K., Kulp, J. L. & Arora, P. S. Evaluation of biologically relevant short α-helices stabilized by a main-chain hydrogen-bond surrogate. *J. Am. Chem. Soc.* **128**, 9248–9256 (2006).

49. Liu, J., Wang, D., Zheng, Q., Lu, M. & Arora, P. S. Atomic structure of a short a-helix stabilized by a main chain hydrogen-bond surrogate. *J. Am. Chem. Soc.* **130**, 4334–4337 (2008).

50. Miller, S. J., Grubbs, R. H. & Beckman, M. Synthesis of Conformationally Restricted Amino Acids and Peptides Employing Olefin Metathesis. *J. Am. Chem. Soc.* **117**, 5855–5856 (1995).

51. Blackwell, H. E. & Grubbs, R. H. Highly Efficient Synthesis of Covalently Cross-Linked Peptide Helices by Ring-Closing Metathesis. *Angew. Chemie Int. Ed.* **37**, 3281–3284 (1998).

52. Scott J. Miller, Helen E. Blackwell, and & Grubbs*, R. H. Application of Ring-Closing Metathesis to the Synthesis of Rigidified Amino Acids and Peptides. (1996). doi:10.1021/JA961626L

53. Yeo, D. J., Warriner, S. L. & Wilson, A. J. Monosubstituted alkenyl amino acids for peptide 'stapling'. *Chem. Commun. (Camb).* **49**, 9131–3 (2013).

54. Nomura, W. *et al.* Cell-permeable stapled peptides based on HIV-1 integrase inhibitors derived from HIV-1 Gene products. *ACS Chem. Biol.* **8**, 2235–2244 (2013).

55. Jamieson, A. G., Boutard, N., Sabatino, D. & Lubell, W. D. Peptide Scanning for Studying Structure-Activity Relationships in Drug Discovery. *Chem. Biol. Drug Des.* **81**, 148–165 (2013).

56. Schafmeister, C. E., Po, J. & Verdine, G. L. An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides. *J. Am. Chem. Soc* **122**, 5891–5892 (2000).

57. Kim, Y.-W. & Verdine, G. L. Stereochemical effects of all-hydrocarbon tethers in i,i+4 stapled peptides. *Bioorg. Med. Chem. Lett.* **19**, 2533–2536 (2009).

58. Bernal, F., Tyler, A. F., Korsmeyer, S. J., Walensky, L. D. & Verdine, G. L. Reactivation of the p53 Tumor Suppressor Pathway by a Stapled p53 Peptide. *J. Am. Chem. Soc.* **129**, 2456–2457 (2007).

59. Bernal, F. *et al.* A Stapled p53 Helix Overcomes HDMX-Mediated Suppression of p53. *Cancer Cell* **18**, 411–422 (2010).

60. Baek, S. *et al.* Structure of the Stapled p53 Peptide Bound to Mdm2. *J. Am. Chem. Soc.* **134**, 103–106 (2012).

61. Hu, B., Gilkes, D. M. & Chen, J. Efficient p53 Activation and Apoptosis by Simultaneous Disruption of Binding to MDM2 and MDMX. *Cancer Res.* **67**, 8810–8817 (2007).

62. Phan, J. *et al.* Structure-based design of high affinity peptides inhibiting the interaction of p53 with MDM2 and MDMX. *J. Biol. Chem.* **285**, 2174–83 (2010).

63. Wachter, F. *et al.* Mechanistic validation of a clinical lead stapled peptide that reactivates p53 by dual HDM2 and HDMX targeting. *Oncogene* **36**, 2184–2190 (2017).

64. Danial, N. N. *et al.* Dual role of proapoptotic BAD in insulin secretion and beta cell survival. *Nat. Med.* **14**, 144 (2008).

65. Li, C. *et al.* Systematic Mutational Analysis of Peptide Inhibition of the p53–MDM2/MDMX Interactions. *J. Mol. Biol.* **398**, 200–213 (2010).

66. García-Echeverría, C., Chène, P., J. J. Blommers, M. & Furet, P. Discovery of Potent Antagonists of the Interaction between Human Double Minute 2 and Tumor Suppressor p53. *J. Med. Chem* **43**, 3205–3208 (2000).

67. Brown, C. J. *et al.* Stapled Peptides with Improved Potency and Specificity That Activate p53. *ACS Chem. Biol.* **8**, 506–512 (2013).

68. Tan, B. X. *et al.* Assessing the Efficacy of Mdm2/Mdm4-Inhibiting Stapled Peptides Using Cellular Thermal Shift Assays. *Sci. Rep.* **5**, 12116 (2015).

69. Sokolowska, I., Ngounou Wetie, A. G., Woods, A. G. & Darie, C. C. Automatic determination of disulfide bridges in proteins. *J. Lab. Autom.* **17**, 408–16 (2012).

70. Spokoyny, A. M. *et al.* A Perfluoroaryl-Cysteine S_NAr Chemistry Approach to Unprotected Peptide Stapling. *J. Am. Chem. Soc.* **135**, 5946–5949 (2013).

71. Fairlie, D. P. & Dantas de Araujo, A. Review stapling peptides using cysteine crosslinking. *Biopolymers* **106**, 843–852 (2016).

72. Timmerman, P., Beld, J., Puijk, W. C. & Meloen, R. H. Rapid and quantitative cyclization of multiple peptide loops onto synthetic scaffolds for structural mimicry of protein surfaces. *ChemBioChem* **6**, 821–824 (2005).

73. Zhang, F., Sadovski, O., Xin, S. J. & Woolley, G. A. Stabilization of folded peptide and protein structures via distance matching with a long, rigid cross-linker. *J. Am. Chem. Soc.* **129**, 14154–14155 (2007).

74. Wang, Y. *et al.* Application of Thiol-yne/Thiol-ene Reactions for Peptide and Protein Macrocyclizations. *Chem. - A Eur. J.* **23**, 7087–7092 (2017).

75. Vinogradova, E. V., Zhang, C., Spokoyny, A. M., Pentelute, B. L. & Buchwald, S. L. Organometallic palladium reagents for cysteine bioconjugation. *Nature* **526**, 687–691 (2015).

76. Muppidi, A. *et al.* Rational Design of Proteolytically Stable, Cell-Permeable Peptide-Based Selective Mcl-1 Inhibitors. *J. Am. Chem. Soc.* **134**, 14734–14737 (2012).

77. Wang, Y. & Chou, D. H. C. A Thiol-Ene Coupling Approach to Native Peptide Stapling and Macrocyclization. *Angew. Chemie - Int. Ed.* **54**, 10931–10934 (2015).

78. Grison, C. M. *et al.* Double quick, double click reversible peptide “stapling”. *Chem. Sci.* **8**, 5166–5171 (2017).

79. Hudson, A. S., Hoose, A., Coxon, C. R., Sandford, G. & Cobb, S. L. Synthesis of a novel tetrafluoropyridine-containing amino acid and tripeptide. *Tetrahedron Lett.* **54**, 4865–4867 (2013).

80. Gimenez, D. *et al.* The application of perfluoroheteroaromatic reagents in the preparation of modified peptide systems. *Org. Biomol. Chem.* **15**, 4086–4095 (2017).

81. Gimenez, D. *et al.* 2,2,2-Trifluoroethanol as a solvent to control nucleophilic peptide arylation. *Org. Biomol. Chem.* **15**, 4081–4085 (2017).

82. Fadzen, C. *et al.* Perfluoroarene-Based Peptide Macrocycles to Enhance Penetration Across the Blood-Brain Barrier. *J. Am. Chem. Soc.* jacs.7b09790 (2017). doi:10.1021/jacs.7b09790

83. Zhang, C., Dai, P., Spokoyny, A. M. & Pentelute, B. L. Enzyme-Catalyzed Macrocyclization of Long Unprotected Peptides. *Org. Lett.* **16**, 3652–3655 (2014).

84. Lühmann, T. *et al.* A perfluoroaromatic abiotic analog of H2 relaxin enabled by rapid flow-based peptide synthesis. *Org. Biomol. Chem.* **14**, 3345–3349 (2016).

85. Lautrette, G., Touti, F., Lee, H. G., Dai, P. & Pentelute, B. L. Nitrogen Arylation for Macrocyclization of Unprotected Peptides. *J. Am. Chem. Soc.* **138**, 8340–8343 (2016).

86. Zou, Y. *et al.* Convergent diversity-oriented side-chain macrocyclization scan for unprotected polypeptides. *Org Biomol Chem* **12**, 566–573 (2014).

87. Madden, M. M. *et al.* Synthesis of cell-permeable stapled peptide dual inhibitors of the p53-Mdm2/Mdmx interactions via photoinduced cycloaddition. *Bioorg. Med. Chem. Lett.* **21**, 1472–1475 (2011).

88. Tickler, A. K. & Wade, J. D. Overview of Solid Phase Synthesis of “Difficult Peptide” Sequences. in *Current Protocols in Protein Science* **50**, 18.8.1-18.8.6 (2007).

6 Fluorinated Prolines as Conformational Tools for the Understanding of CypD and p53 Interactions

As discussed in the previous chapter, the tumour suppressor protein p53 forms an important target and topic in cancer research. While interactions between p53 and specific other proteins (*e.g.* MDM2) are widely investigated and validated, the exact mechanism of interaction with others is often unknown. Cyclophilin D (CypD) is a peptidyl-prolyl isomerase (PPIase) and has been shown to interact with p53, however the exact nature of this interaction is not fully understood. This chapter aims to explore the use of fluorinated proline residues to gain insights in the p53-CypD interaction, making use of ^{19}F NMR.

A review paper on which work in this chapter is based has been published:

S.J.M. Verhoork, P.M. Killoran, C.R. Coxon, *Biochemistry*, **2018**, 5743, 6132–6143.

Available at: https://doi.org/10.1021/acs.biochem.8b00787

Introduction

6.1 Peptidyl-prolyl isomerases

Peptidyl-prolyl isomerases (PPIases) play a role in the interconversion of the proline peptide bond between the *cis* and *trans* isoforms. The proline peptide bond is the only peptide bond that naturally occurs in equilibrium between both *cis* and *trans* isoforms, whereas all other peptide bonds naturally occur as *trans* almost exclusively. Rotation around the peptide bond is slow and requires high activation energy due to the partial double bond between N and C=O.[1] PPIases function by catalysing this conversion, which may be a strategy to turn molecular switches off or on.[2] Additionally, this process has been considered to be important for protein folding and assembly of multidomain proteins.[3] This sparks interest from a biomolecular and pharmaceutical perspective, as PPIases could potentially be involved in more processes in the body than currently known. Additionally, misfolding through incorrect function of PPIases could form an interesting drug target, making it critical to understand the function of this class of enzymes better.

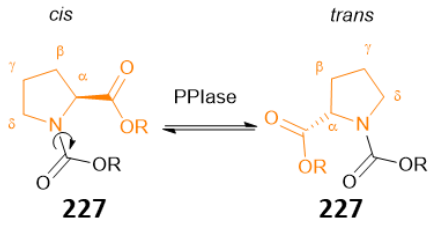


Figure 6.1: Peptidyl-prolyl isomerase activity. PPIases can lower the energy barrier between *cis*- and *trans*-isoforms, promoting isomerism.

Up to date, three classes of PPIases are known; FK506 binding proteins (FKBPs), parvulins, and cyclophilins. Despite having a communal isomerism catalysis function, these classes appear to have little in common, both structurally, and in their ligands. The main structural property that all PPIases have in common is a central β-sheet in the catalytic domains.[4] PPIases have been associated with multiple pathophysiological conditions, including

cardiovascular disease, cancer, and Alzheimer's disease. This research focuses only on the functioning of cyclophilins, more specifically Cyclophilin D.

6.1.1 Cyclophilins

Cyclophilins can occur as a single-domain protein, or as an oligomeric protein complex. This has been suggested to play a role in the recognition of and interaction with protein substrates.[5] For a PPIase to function with high efficiency it needs to bind one isoform (*cis* or *trans*) with higher affinity over the other. This would allow for the binding, conversion, and release of the substrates.[6] The human family of cyclophilins consists of 7 major proteins hCypA (also called hCyp-18a), hCypB (also called hCyp-22/p), hCypC, hCypD, hCypE, hCyp40, and hCypNK.[7]

Interestingly, most studies relating to cyclophilins have been performed using a cyclic peptide immunosuppressant called Cyclosporin A (CsA). Especially cyclophilin A forms a major target of PPIase activity inhibition by potent binding to the active site by CsA.[8] The immunosuppressive activity of the CsA-CypA is caused through inhibition of the phosphatase activity of calcineurin (CN), which is a calcium-dependent serine/threonine phosphatase that promotes the synthesis of T cell lymphokines.[9] In addition to immunosuppressive inhibition, non-immunosuppressive PPIase inhibition of cyclophilins is also a desired outcome in drug development. Inhibition of CypD has been suggested as mitochondrial protection strategy in both acute and chronic degenerative diseases.[10,11]

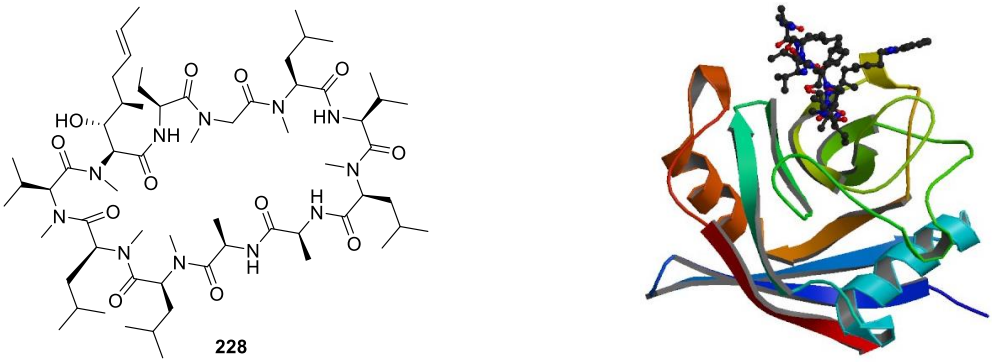


Figure 6.2: Structure of immunosuppressant cyclosporin A and crystal structure of CypD in complex with CsA analogue, JW47 (PDB: 5AE0).

Cyclophilins are abundant and can be found in most cellular compartments of most tissues. CypA, Cyp40 and CypNK can be found in the cytosol, whereas CypB and CypC are targeted to the ER protein secretory pathway[12] and CypE is localised in the nucleus.[13] Cyclophilin D (CypD) is localised in the mitochondria.[14,15] The exact physiological role of CypD has not been established, however, models exist of the involvement in the formation of a non-specific channel (mPTP) formed in the inner mitochondrial membrane (IMM) through interactions with other proteins,[5] including PiC, F_0F_1 ATPase subunit, Bcl2, and p53. CypD promotes the open conformation of the MPT pore through interaction with the mitochondrial adenine nucleotide translocator (ANT).[6,16] This open conformation allows the transport of solutes up to 1.5 kDa across the inner mitochondrial membrane, which results in mitochondrial swelling, depolarisation, and release of various intermembrane proteins.[17] Good functioning of mitochondria is essential for cell survival. Firstly, mitochondria are responsible for the production of energy for the cell in the form of ATP.[18] Secondly, they also play an important role in the execution of apoptosis and regulated necrosis.[19,20] Mitochondrial dysfunction has been associated with different diseases, including several neurological diseases,[21,22] amyotrophic lateral sclerosis, psychiatric disorders, stroke and ischemia.[5] In addition to its involvement in mitochondrial function, CypD has been reported to have an anti-apoptotic effect and has been found to be overexpressed in various tumours.[23]

6.1.2 Poly-proline motif in p53

One protein that has been identified as a substrate for CypD is p53. As previously described in Chapter 5, the tumour suppressor p53 plays an important role in approximately half of all tumours,[24,25] making it a well-studied protein. Despite all efforts, relatively little is known of the structural properties of the p53 poly-proline motif (amino acids 64–92; PRMPEAAPPVAPAPAAPTPAAPAPAPSWP), however, the conserved PXXP motifs are considered to be important for growth suppression.[26] Deletion of these motifs can impair p53-mediated apoptosis, without affecting its ability to induce growth arrest.[27] To the best of our knowledge, there is currently only one data entry of the p53 poly-proline region in the protein

data base (Figure 6.3, PDB: 2LY4). There is no full sequence data entry in the PDB of the complete p53 protein, indicating the importance of the generation of new structural data.

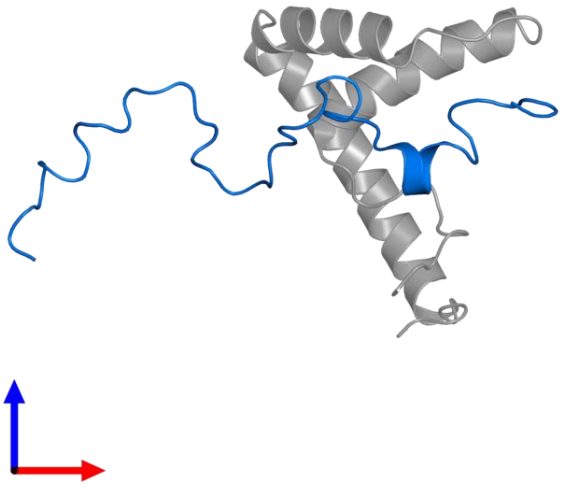


Figure 6.3: Cellular tumour antigen p53 in assembly 1. This protein is highlighted and viewed from the front. (PDB: 2LY4)

While it is known that CypD binds to p53, the exact position and nature of this interaction is unknown. It would be of interest to determine the *cis-trans* ratios of the individual peptidyl-prolyl bonds of each proline residue in the polyproline motif. This information can be used for better understanding of the p53 protein structure, as well as function as a template of monitoring PPIase activity by CypD. Strategies to obtain this knowledge experimentally are discussed in Section 6.2.

6.1.3 Interaction between Cyclophilin D and p53

p53 is well known for its central role as cellular stress sensor that responds to signals including DNA damage, oxidative stress, and ischemia, which can result in apoptosis via transcription-dependent and independent mechanisms.[28] The role of p53 in necrosis (the death of most or all cells in an organ or tissue due to disease, injury, or failure of the blood supply) is less understood. In ischemic tissues, necrosis is dependent on Cyclophilin D (CypD), the key regulator of the mitochondrial permeability transition pore (mPTP). Under normal conditions, the mPTP remains closed, however, it is opened under stress, resulting in an ion influx into the mitochondria. This disrupts the proton gradient that is normally present across the

mitochondrial membranes and shuts down oxidative phosphorylation and ATP production, which are crucial for healthy cell functioning. It is unclear how CypD becomes activated, however, the opening of the mPTP is believed to be dependent on its prolyl isomerase activity.[29] Upon oxidative stress, p53 can undergo a physical interaction with CypD, triggering the opening of the mPTP, thereby inducing necrotic cell death.[30]

Understanding the functioning of the p53 poly-proline motif can also be relevant for other protein-protein interactions p53 may be involved in, as changes induced by CypD could change the protein structure of p53. For example, the poly-proline motif has been found to play an important role in the susceptibility of negative regulation by MDM2. Proteins lacking this region made p53 more accessible to ubiquitination and nuclear export, as well as MDM2-mediated regulation compared to wildtype p53 protein.[31] Studies have been done to see what the effects are of removing the poly-proline motif, which leads to impaired apoptotic activity.[32] Taken together, the importance of the functioning of p53 has been highlighted, whereas the role of CypD in this is still poorly understood.

6.2 Fluorinated prolines as conformational tools and reporters

Proline is an unconventional amino acid in comparison to the other standard amino acids, as it is the only one of which the side chain forms a part of the protein backbone. This results in unique properties that are essential in protein and peptide function and three-dimensional structure.[33] Proline has two main conformational equilibria: 1) amide-bond *cis-trans* isomerism and 2) *endo/exo* pyrrolidine ring puckering (Figure 6.4).[34] Ring puckering and *cis-trans* isomerisation can both influence the formation of α-helices and β-sheets and are thus very important for protein structure.[35,36] Additionally, prolines have also been shown to affect the folding kinetics of proteins.[37]

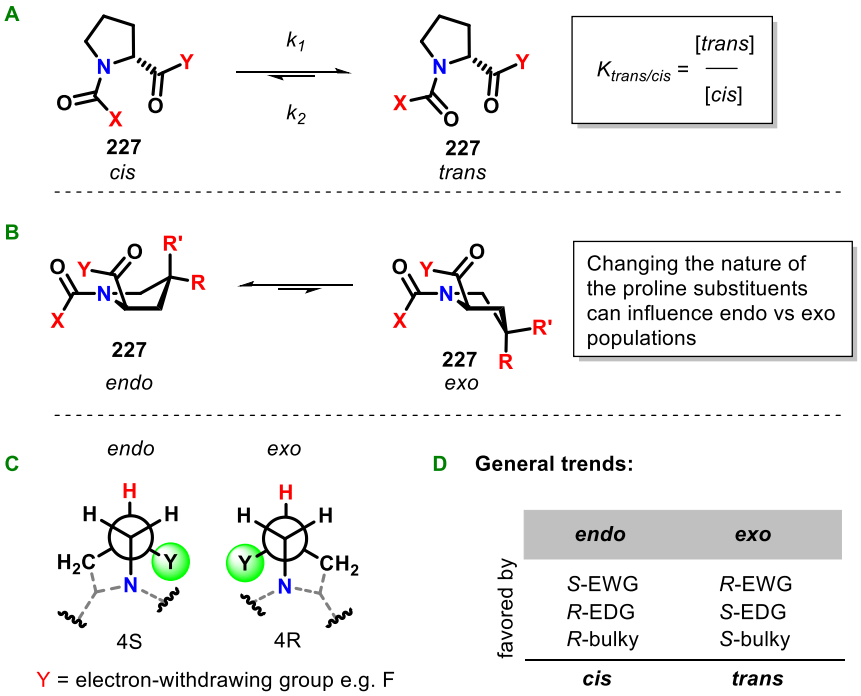


Figure 6.4: (A) cis–trans conformer isomerization within peptidyl–proline motifs. k_1 and k_2 are rate constants for cis-to-trans and trans-to-cis conformational changes, respectively. (B) Interconversion between endo and exo ring-puckered forms is influenced by proline ring substituents. (C) Newman projections showing the gauche arrangement about an N–C–C–Y axis in a 4-substituted proline and the influence on ring pucker. (D) General trends in ring pucker and X–Pro amide conformer preference (EWG, electron-withdrawing group; EDG, electron-donating group). From reference 38.

6.2.1 Fluorinated prolines

Introduction of fluorine has been popular in medicinal and pharmaceutical chemistry, with an increase in occurrence from 20 to ~30% of drugs in recent years.[39,40] The advantages of incorporating fluorine include improved metabolic stability,[41] altered physicochemical properties, and increased binding affinity. Despite the prevalence of fluorine in medicinal chemistry, relatively little research has been conducted in the role of fluorine in peptide and protein chemistry. However, in addition to conformational effects that are induced by incorporation of fluorine (discussed later), fluorination of compounds can be a useful tool for studying (protein-ligand) interactions using ^{19}F NMR.[42–44] Fluorinated prolines have the potential to be exceptionally useful tools in chemical biology to modulate protein or peptide stability and folding kinetics.

6.2.1.1 Conformational effects

Hydroxyprolines and 4-fluoroprolines are by far the most common proline analogues that have been reported (Figure 6.5). 4-Hydroxoproline occurs in nature[45] and can be used as a precursor for the synthesis of 4-fluoroprolines.[46] Depending on which substituent is present at the 4-position, different steric and stereo-electronic effects can influence the proline-peptide conformation. For electron-withdrawing substituents at the 4-position, *S*-stereochemistry strongly favours an *endo* pucker, while *R*-stereochemistry favours an *exo* pucker, which is reversed for non-electron-withdrawing groups and/or large bulky substituents.[38] Using this information, it is potentially possible to tune proline conformation within synthetic peptides, opening up possibilities for stimulus-responsive conformational switching.

Substitution of proline with fluorine (4*R*- or 4*S*-fluoroproline) affects the spontaneous isomerisation of the peptidyl-prolyl bond, with both the rate and the equilibrium constants being affected.[47] Overall, the incorporation of fluorine has been driven by two main factors: **1)** powerful conformational effects of fluorine and **2)** the possibility for ^{19}F to be used as an NMR reporter for structural and stability studies. The conformational effects of fluorine substitution at the 4-position (and 3-position) can be explained by the adoption of a *gauche* (60° torsion

angle) conformation between the amide nitrogen and the *vicinal* fluorine atom in an N-C-C-Z (Z = N, O, or F) system (Figure 6.4.C).[48]

6.2.2 4-Fluoroprolines

A wide range of fluorinated prolines have been synthesised and utilised for chemical biology studies, including 3-fluoroprolines,[49,50] 3,4-difluoroprolines,[51] trifluoromethylprolines,[52,53] fluorinated methanoprolines,[54] pentafluoro phenyl-hydroxyprolines,[55] and perfluoro-*tert*-butyl-hydroxyprolines[56] (Figure 6.5). The advantages and applications of each of these are discussed in reference 38. Here, only the properties of 4-fluoroprolines are discussed, as they are the only type of fluorinated prolines that are used for the research in this chapter.

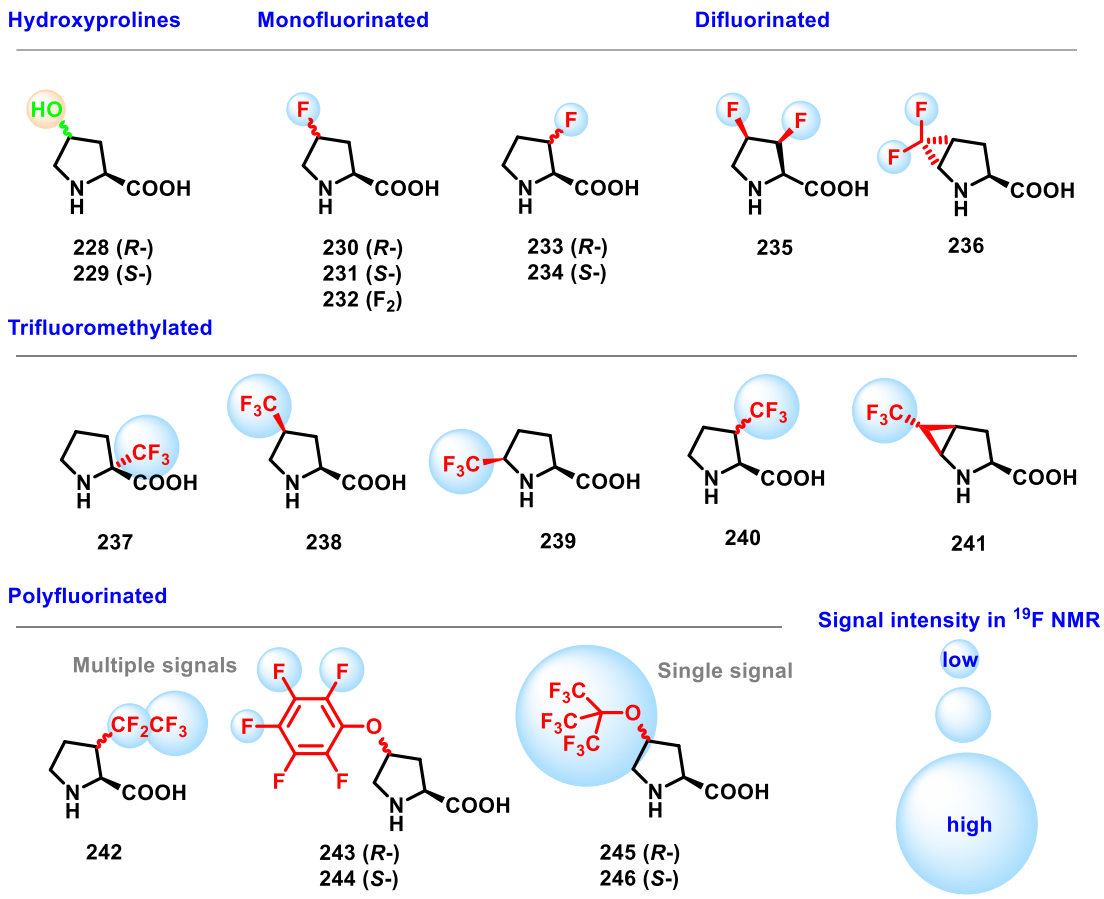


Figure 6.5: Various reported fluorinated proline derivatives used in peptides and proteins and an indication of their signal strength as reporters for ^{19}F NMR (not to scale). Adapted with permission from reference 38.

The commercial availability of L-, D-, and disubstituted Fmoc-4-fluoroprolines allows for convenient laboratory synthesis of short peptides. For longer peptides and proteins, 4-fluoroprolines have also been incorporated via total synthesis using native chemical ligation. Incorporation into proteins (collagen[57] and ubiquitin[58]) has been shown to improve thermal stability, however, there are also examples in which biological activity is lost upon incorporation of 4-fluoroprolines (proctolin[59]). When using 4-fluoroproline as a ^{19}F NMR reporter, it is important to keep these structural, conformational and biological effects in mind.

The previously described *gauche* effect forms an important factor in *cis-trans* isomerism and *endo*- and *exo*-ring puckering for 4-fluoroprolines. In the case of substitution at the 4-position with two fluorine atoms (4,4-difluoroproline (**232**)), the electronic effects counter interact, leading to behaviour similar to that of the native unsubstituted proline, based on similar $K_{trans/cis}$ equilibrium ratios.[58,60] By inserting fluorine atoms at the 4-position, the double-bond character of the prolyl amide bond is reduced, lowering the energy barrier for *trans-cis* isomerisation ($\Delta G^{\ddagger}_{300\,K}$) in the order Pro (88.4 kJ/mol) > 4*R*-Flp (87.4 kJ/mol) ~ 4*S*-Flp (86.3 kJ/mol) > F_2Pro (83.3 kJ/mol) and, therefore, increasing the isomerisation rate in the Ac-(F)Pro-OMe model system.[38,61] In general, this means that when using 4*R*- or 4*S*-fluorinated prolines as reporters for ^{19}F NMR, some of the protein and peptide properties will be altered, limiting the usage for unbiased experiments.

6.2.3 Fluorinated prolines and ^{19}F NMR to study dynamic biological processes

Different fluorinated prolines offer differing degrees of *cis* conformer stabilisation and provide a useful tool kit for probing the conformational and stability properties of a protein or peptide. An overview of fluorinated prolines accompanied by their *cis-trans* ratios is presented in Figure 6.6. While fluorination can have an effect on *cis-trans* ratios, the preceding amino acid can also play an important role, with aromatic residues favouring the *cis* isomer slightly in most cases.[62] The *cis*-peptidyl−prolyl-bond population of a Xaa-Pro can be stabilised by local CH(Pro)···π interactions when **1)** Xaa is aromatic[63] or **2)** Xaa is Pro, Gly or aromatic (Aro) in Xaa-Pro-Aro motifs[64,65] in nonfluorinated systems.

Despite limitations in the form of conformational bias, there is still advantage in studying fluorinated proline containing proteins or peptides, as these large biomolecules are difficult to study by the conventional chemical methods of ^{1}H and ^{13}C NMR. There is a significant advantage in the wide chemical shift range and characteristic signals for *cis-trans* isomers of ^{19}F NMR (Table 6.1). ^{19}F NMR spectra can often be easily interpreted, due to the low number of atoms present (e.g., one ^{19}F NMR signal for proline compared with seven ^{1}H).[66] Natural occurrence of fluorine in biomolecules and in nature is very low,[67] thus making fluorine a perfect reporter atom for ^{19}F NMR. While ^{19}F NMR studies of proteins with their ligands can provide valuable insights into ligand-target interactions and dynamics, the sensitivity of detection can be limited in complex mixtures or at low concentrations.[68–71] The interference of fluorine with the natural characteristics of a peptide/protein should be considered, however, can also be advantageous as fluorinated prolines favouring one isoform over the other are useful tools for measuring kinetics of isomerisation, and in turn by slowing down isomerisation, can be a useful probe for enzyme mechanism and function.[38]

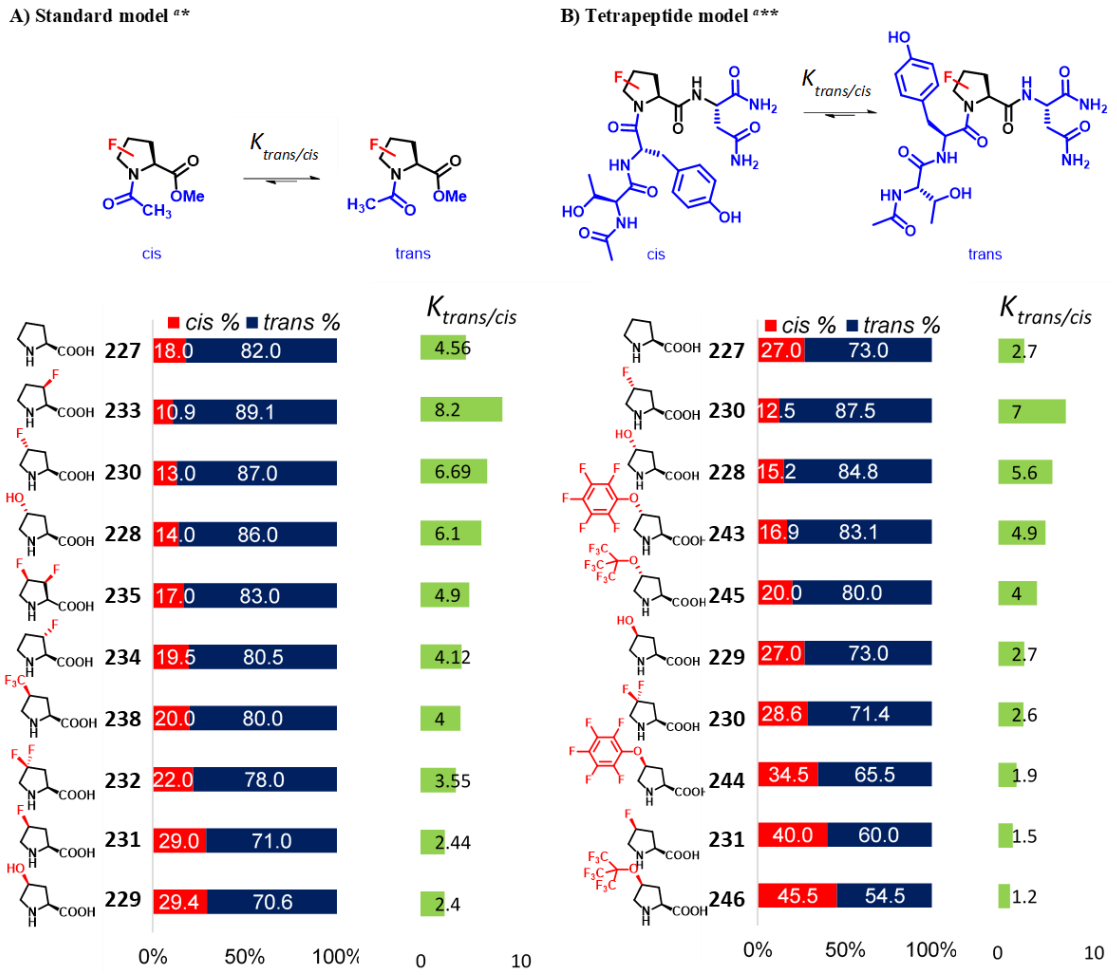
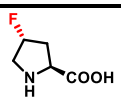
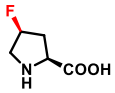
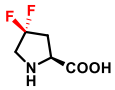
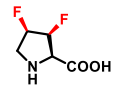
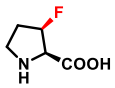
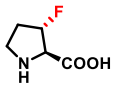
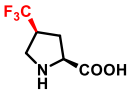
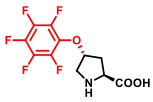
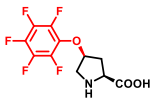
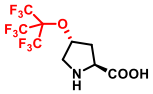
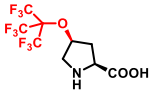


Figure 6.6: Graphical ranking of differential %*cis* and *trans* conformers (blue/grey) depending on nature of the fluorinated proline in A) the standard AcProOMe* and B) a tetrapeptide model** showing the similar trends but quantitative differences between a simple model and a short peptide sequence. The corresponding $K_{trans/cis}$ (ratio of peptide with *trans* amide bond compared to peptide with *cis* amide bond as determined by NMR) is also presented graphically (yellow). Notes: [a] Where more than one value is reported in different publications, the values are generally in good agreement with values shown. * Data obtained Ref 50 (D_2O at 37 °C), Ref 72 (D_2O at 25 °C), Ref 51 (D_2O at 25 °C) and Ref 53 (D_2O at 25 °C). ** Data obtained in 25 mM NaCl, 5 mM phosphate [pH 4], 9:1 H_2O/D_2O at 298K Ref 55.

Table 6.1: Reported Characteristic ^{19}F NMR Chemical Shifts for *trans* and *cis* Conformers in Model Systems As Indicated[a]

Fluorinated proline		Ac-(F)Pro-OMe		Ac-TY-(F)PN-NH_2	
		Trans (ppm)	***Cis (ppm)***	***Trans (ppm)***	***Cis (ppm)***
230		-177.85 [72, A]	-177.79	-177.1 [57, D]	-177.9
231		-173.26 [72, A]	-173.25	-173.2 [57, D]	-173.5
232		-98.4 (pro-*S*) [72, A] -102.2 (pro-*R*)	-95.5 (pro-*S*) -104.7(pro-*R*)	-100.6 [57, D]	-103.2, -103.9, -94.8, -95.4
235		-203.3 (F^4) [72, A] -210.4 (F^3)	-200.3 (F^4) -208.5 (F^3)	-	-
233		-108.5 [50, B]	-106.5	-	-
234		-97.5 [50, B]	-98.6	-	-
238		-71.0 [53, C]	-71.3	-	-
243		-	-	-163.2 [55, D] -161.9 -156.5	-163.6 -162.1 -156.1
244		-	-	-163.4 [55, D] -162.4 -156.3	-163.6 -162.9 -156.3
245		-	-	-70.7 [55, D]	-70.8
246		-	-	-70.0 [55, D]	-70.7

[a]NMR conditions: (A) D_2O at 25 °C, (B) 9:1 H_2O/D_2O mixture at 35 °C (no exact values provided in ref [50]; chemical shift values in the table are estimated from raw NMR spectra), (C) D_2O at 27 °C, and (D) 5 mM phosphate buffer (pH 4) with 25 mM NaCl in a 9:1 H_2O/D_2O mixture at 23 °C.

6.3 Fluoroproline positional-scan for understanding peptidyl-prolyl bond dynamics

6.3.1 Aim of this chapter

The interaction between Cyclophilin D and its substrates is still poorly understood. While there is evidence that CypD can bind to p53, the site of interaction has not yet been identified. It is also known that CypD can function as a peptidyl-prolyl isomerase, stimulating enhanced isomerisation between *cis*- and *trans*-peptidyl-prolyl bonds. To gain further understanding in protein folding, and misfolding, it is of importance to know where the interaction between CypD and p53 takes place, and if CypD affects one or more proline bonds.

6.3.2 Approach

As previously described, fluorinated prolines can be used as tools for understanding protein function and dynamics. Prior to working with the polyproline motif of p53, two other short CypD substrates were investigated, as this would provide a simpler model before moving to the more complicated p53 peptide. To locate the possible site of isomerisation for p53, a section of the polyproline motif of the p53 protein will be synthesised (peptide **135**, NH_2-D**P**G**P**DEA**P**RM**P**EAA**PP**VA**P**A**P**AA**P**T-$CONH_2$). A fluoroproline positional-scan will be performed, in which each native proline residue is sequentially replaced by a *4S*-fluoroproline. This fluorinated proline has been chosen as it closely resembles *cis-trans* ratios of native proline (13% *cis* vs 18% *cis*, respectively). While 4,4-difluoroproline may be even closer related to native proline (with 22% *cis*), it was chosen not to use this, as it has been reported that 4,4-difluoroprolines can be more structurally disruptive than the monofluorinated prolines.[73]

The *cis-trans* ratios for each peptide will be determined using ^{19}F NMR. Subsequently, each peptide will be incubated with a known amount of CypD enzyme to observe if changes in *cis-trans* ratios occur (as observed by ^{19}F NMR). As a negative control, a known inhibitor of CypD (Cyclosporine A) will then be added to observe if natural *cis-trans* ratios are restored, if changes occurred upon addition of the enzyme.

Results and discussion

6.4 Peptidyl-prolyl isomerism in short CypD-binding peptides

It was hypothesised that due to spectral crowding and complex proton resonance assignment, that the introduction of a ^{19}F label would provide a facile means of measuring prolyl bond conformational status. The benefits of ^{19}F NMR are the spin ½ nucleus having 100% natural abundance, comparable sensitivity to ^{1}H, having high sensitivity to the local environment and can be decoupled from proton resonances, thereby overcoming spectral crowding and providing distinct resonances for *cis* and *trans* conformers and relatively simple spectra. These features provide an excellent reporter to study dynamic biological systems.[38]

Initially, to test this hypothesis, relatively simple models were explored and single proline-containing peptide sequences identified as CypD substrates (Prof. Lu Yun Lian, Director of the NMR Centre for Structural Biology, The University of Liverpool) were investigated. The peptides sequences NH_2-KKYGPFVAD-$CONH_2$ (**247**) and NH_2-ELYGPSMR-$CONH_2$ (**248**) derive from sections of the F_1F_0 ATPase subunit B and Bcl2 proteins respectively and contain possible isomerisable proline residues.

Mitochondrial F_0F_1-ATP synthase is a large multisubunit complex organised into a catalytic part (F1) and a membranous moiety (F0) linked by central and peripheral stalks. The enzyme catalyses the synthesis of ATP in the presence of oxygen, whereas it hydrolyses ATP during anoxic conditions. Diffusion of protons drives the rotation of the subunits in either way for ATP synthesis or ATP hydrolysis. Cyclophilin D decreases the enzymatic activity of the complex by associating to the oligomeric forms of the ATP synthase and interacting with OSCP, subunit d, and subunit b of the lateral stalk.[74]

Bcl2 is an anti-apoptotic mitochondrial protein that functions by limiting the release of Cytochrome C (CytC) from mitochondria during apoptosis.[23,75] CytC can be released from mitochondria through the MPT pore of which the formation is stimulated by overexpression of CypD. For this reason, interactions between Bcl2 and CypD and their effects on apoptosis are

of interest. Indeed, interaction between Bcl2 and CypD has been shown by several techniques including co-immunoprecipitation, pulldown, and mammalian two-hybrid assays.[23]

Traditionally, proline *cis* and *trans* conformations can be quantified and identified using ^{1}H NMR, however, this can lead to very complex NMR spectra, causing difficulty in clear assignment and quantification. As such, peptides SubB[4(*R*)FPro] (**249**), SubB[4(*S*) FPro] (**250**), Bcl2[4(*R*)FPro] (**251**) and Bcl2[4(*S*)FPro] (**252**) were prepared incorporating (*R*)- or (*S*)-4-fluoroprolines in place of the isomerisable proline residues. The assignment of the ^{19}F NMR signals to the *cis* and *trans* isomers was based on reported conformational preferences of the specific fluoroprolines, as well as the consistency with reported chemical shifts. The value $K_{trans/cis}$ is the equilibrium constant between *trans* and *cis* conformers; this value is greater than 1 since the *trans* conformer is generally favoured over the *cis* conformer in peptide and proteins. Fluorination imposes modest stereoelectronic effects upon the prolines. SubB[4(*R*)FPro] (**249**) has a preference for the Cγ *exo* pucker, shifting the equilibrium in slight favour of the *trans* conformer, whilst the proline of SubB[4(*S*)FPro] (**250**) favours the Cγ *endo* pucker, increasing the *cis* population although the *trans* conformer still dominates. Both SubB 4(*R*) and 4(*S*) peptides were substrates for CypD showing that the difference in ring pucker conformation and *cis-trans* has minimal effect on the isomerisation activity. In the case of the Bcl2 peptide, the 4(*R*) FPro (**251**) very strongly favours the *trans* conformer, with over 94% of the population adopting this conformation; the 4(*S*) conformation push the equilibrium in favour of the *cis* conformation, increasing from 6% in 4(*R*) to 16% in the 4(*S*) peptide (**252**) (Figure 6.7).

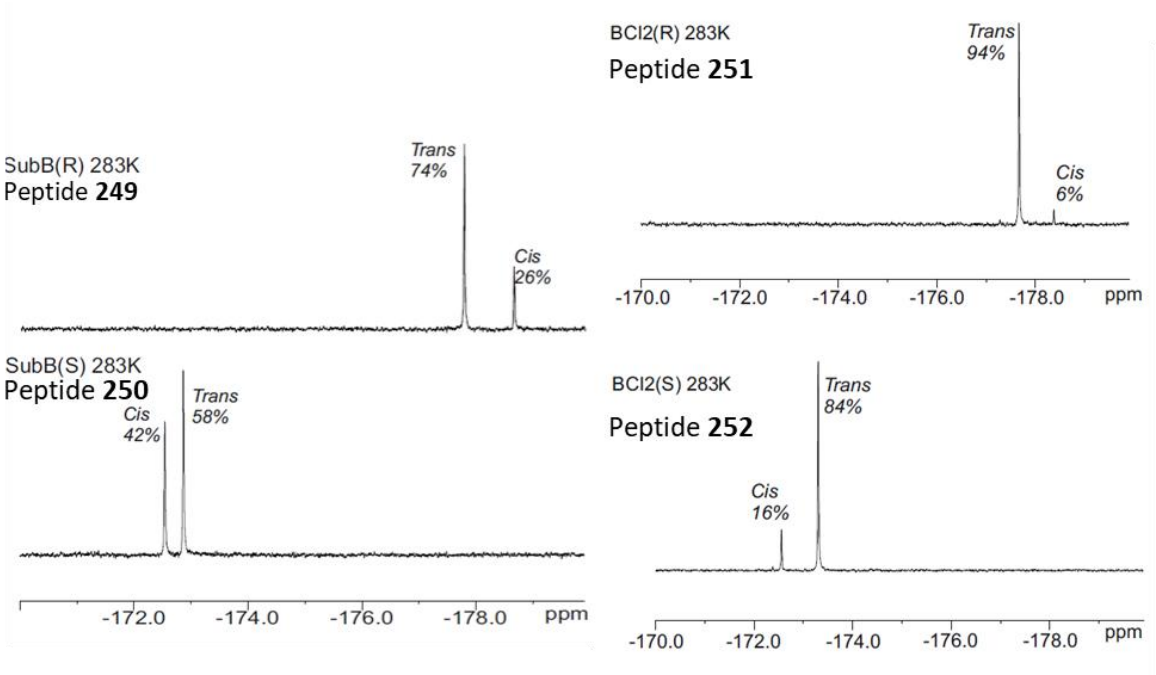


Figure 6.7: Native *cis-trans* ratios of fluorinated SubB (**249** and **250**) and Bcl2 (**251** and **252**) peptides.

In each case, addition of CypD leads to significant line-broadening in the ^{19}F NMR spectra, which correlates with an increase in the rate of dynamic exchange between the *cis* and *trans* conformers although the relative population between the *cis* and *trans* conformers remains unchanged. This continued to be the case for SubB[4(*S*)FPro] (**250**) upon further titration of CypD in increasing concentrations (200:1 – 22.5:1; peptide:CypD) with continued signal broadening (Figure 6.8). This indicates clearly that peptidyl prolyl isomerases merely increase the rates of isomerisation between the two conformers rather than converting *cis* to *trans* conformer (or vice versa) per se. This is consistent with the observation that upon addition of CsA, the *cis-trans* population ratio and line-breadth reverts to the equilibrium distribution found in the absence of the isomerase.

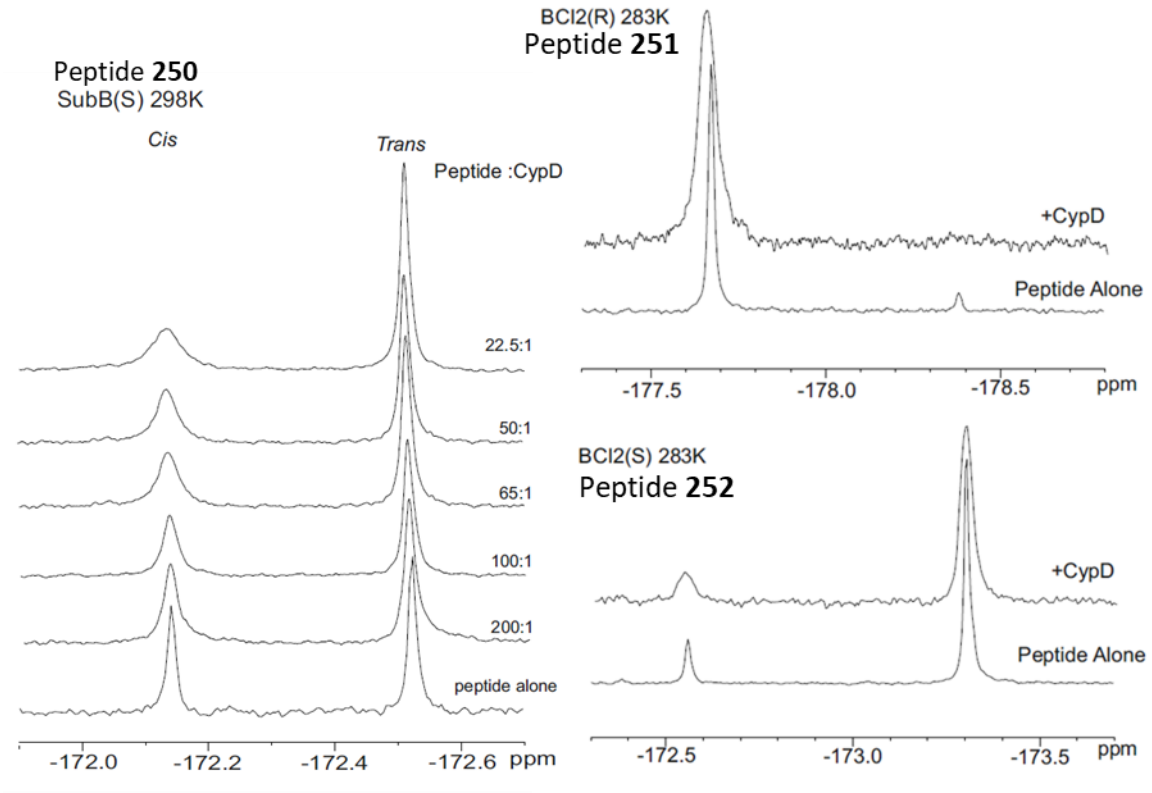


Figure 6.8: ^{19}F NMR experiments to determine peptidyl-prolyl isomerase activity of CypD for peptide **250**, **251**, and **252**.

6.5 Optimisation of the synthesis of peptide **135**

Upon the successful experiments utilising short CypD substrates an attempt was made to synthesise native polyproline containing p53 peptide **135** (Figure 6.9). Medium length peptides like peptide **135** (25 amino acids) can be difficult to synthesise. For this reason, prior to including expensive fluoroprolines in the peptide sequence, the synthesis was optimised using only 'regular' proline residues.

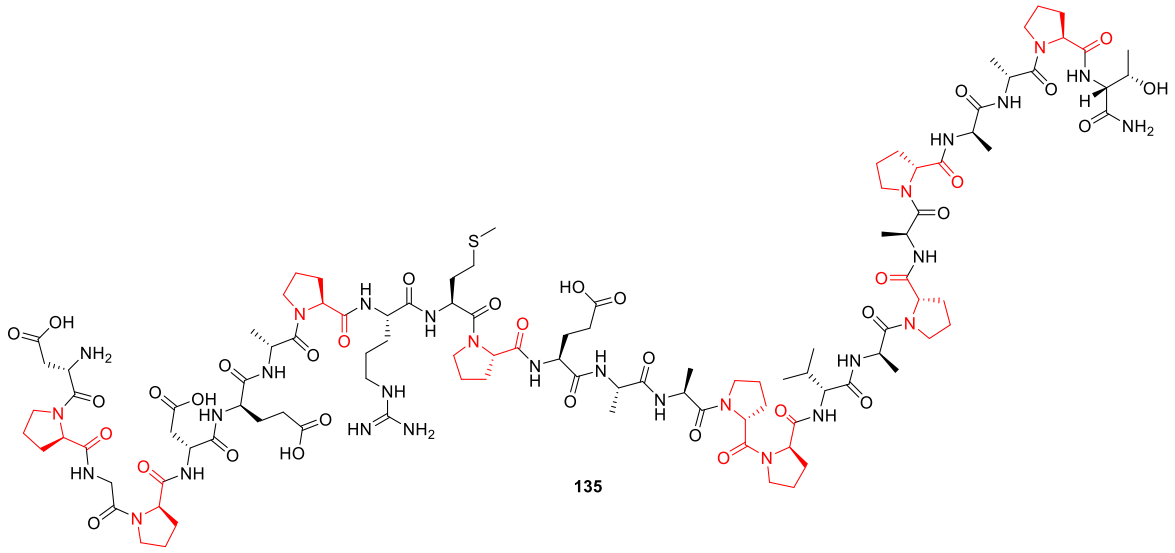


Figure 6.9: Structure of peptide **135** (NH_2-D**P**G**P**DEA**P**RM**P**EAA**PP**VA**P**A**P**AA**P**T-$CONH_2$) with all proline residues marked in red.

Peptide **135** was synthesised by automated SPPS using a Liberty Blue™ Microwave-assisted synthesiser (CEM), following standard procedures (single couplings, deprotections with 20% piperidine). Synthesis using only single amino acid couplings did yield the desired peptide, however, the quality of the crude synthesis product was too low to directly use for further experiments, or to be able to purify by prep-HPLC (Figure 6.10.A). In an attempt to improve the quality, all proline residues and subsequent amino acids were double coupled. The resulting peptide product was much cleaner and would be suitable for prep-HPLC purification (Figure 6.10.B).

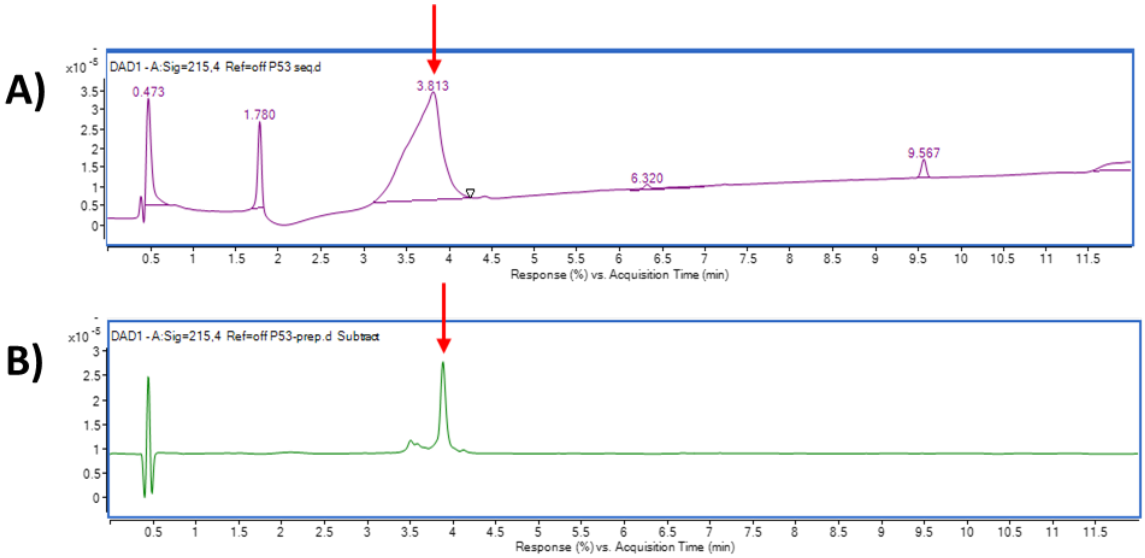


Figure 6.10: HPLC chromatogram (215 nm) of crude peptide **135** using A) standard synthesis procedures; or B) optimised synthesis procedures. The desired peptide product (**135**) is indicated with the red arrow.

6.5.1 Incorporation of 4*S*-Fluoroproline residues

After the development of an effective automated synthesis strategy of peptide sequence **135**, the more expensive 4-*cis*-fluoroproline was incorporated into the peptide sequences. In all cases, the peptides were synthesised by automated SPPS using standard procedures (single couplings, 20% piperidine deprotections), except for the fluoroproline and the subsequent amino acid residue, which were coupled manually (Figure 6.11). Prior to the introduction of fluoroproline, a test cleave was performed, to ensure the quality of the preceding sequence was preserved. After each manual coupling, completion of the reaction was determined by Kaiser test. Where necessary, the remaining peptide sequence was synthesised using automated peptide synthesis following standard procedures as before. An attempt was made to purify peptides by prep-HPLC, however, this did not improve the purity as much as expected and yielded only small quantities. For this reason, it was chosen to continue the ^{19}F NMR experiments using the crude fluorinated peptides.

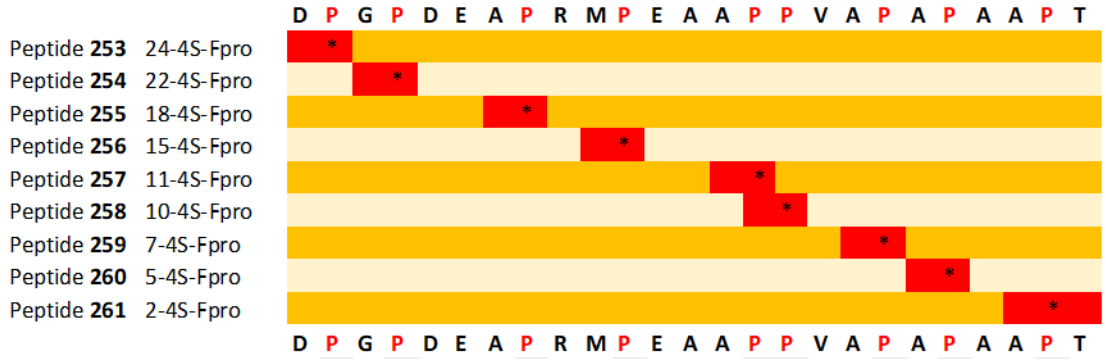


Figure 6.11: Synthesis overview of fluoroproline-scan p53 peptides. Red indicates manual coupling. Each fluorinated proline is marked with an *.

6.6 ^{19}F NMR of fluorinated p53 peptide sequences

As each p53 peptide sequence (peptide **253** - **261**) contained one 4*S*-fluoroproline at a distinct position, this allowed for extensive ^{19}F NMR studies to gain insight in the *cis-trans* preference and isomerisation of each individual proline residue. Each peptide was expected to give two fluorine signals, corresponding to each isoform. Peptide samples were prepared at 5 mg/mL in phosphate buffer [250 mM K_2HPO_4, 250 mM KH_2PO_4, 100 mM NaCl, 0.5 mM EDTA, 1 mM DDT; pH 6.8]. All NMR spectra were calibrated and normalised using the TFA signals. All ^{19}F peptide signals fall in the range between -171.5 and -174.5 ppm (Figure 6.12), which is consistent with reported data in the literature where chemical shifts of -173.26 (*trans*) and -173.25 (*cis*) are reported for Ac-(FPro)-OMe[51] and -173.2 (*trans*) and -173.5 (*cis*) for Ac-TY-(FPro)N-NH_2.[57] Interestingly, the signal resolutions are different for each fluorinated peptide, with for example **258** showing a large resolution, and **253** a very small resolution. While each fluorinated proline residue is expected to mainly report the *cis-trans* ratios at its local site, it is also possible the *cis-trans* ratios of nearby non-fluorinated proline residues influence the signals, which could explain the larger resolution for some peptides.

25” - **NH_2-DPGPDEAPRMPEAAPPVAPAPAAPT-$CONH_2$** – 1”

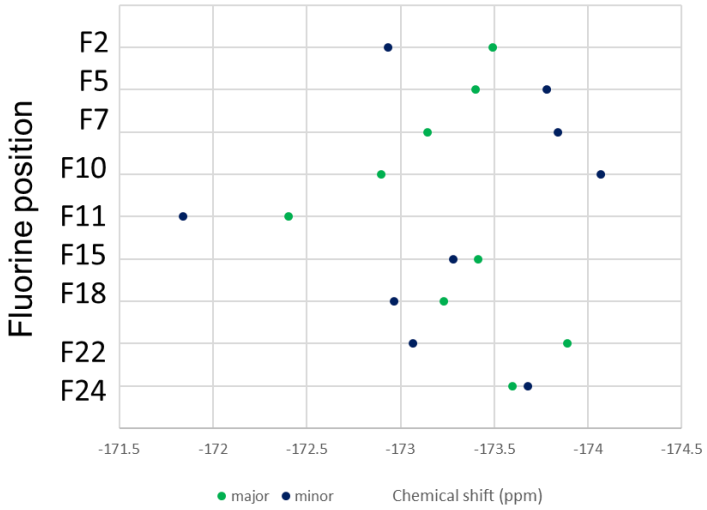


Figure 6.12: Visualisation of the chemical shifts of the minor and major ^{19}F signals for each fluorinated proline containing peptide.

Based on literature values, there appears to be a trend that the *cis* conformation typically has a lower chemical shift value than the *trans* conformer (Table 6.1). This would suggest that four of the proline residues in the poly-proline p53 sequence occur mainly as the *cis* conformer (Figure 6.11). It should be noted that three of these residues are in relatively close proximity to each other (5, 7, and 10), which could suggest a helical (possibly polyproline I helix) tertiary structure. Five of the peptides were found to be predominantly in *trans* conformation, ranging between 75-96% *trans*. As previously shown in Figure 6.6, Ac-Pro-OMe (**227**) occurs as 18% *cis*, while the substitution with 4*S*-F (**231**) increases the *cis* species to 29%. Similarly, in the tetrapeptide model, the native proline peptide (**227**) occurs as 27% *cis*, while the 4*S*-F-substituted (**231**) occurs as 40% *cis*. This shows that the substitution with fluorine can cause a change in *cis*-*trans* ratios of 11-13%, which could also be the case for the fluorinated p53 peptides **253**-**261**. As mentioned previously in Section 6.2, 4*R*- or 4*S*-substitution can affect the isomerisation of the peptidyl-prolyl bond, affecting

both the rate and equilibrium constants,[47] which could have influenced the reported cis-trans ratios in Figure 6.11.

Table 6.2: Determined ^{19}F Chemical shifts for the major and minor conformers in each individual NH_2-DPGPDEAPRMPEAAPPVAPAPAAPT-$CONH_2$ sequence.

P53 sequences	major (ppm)	minor (ppm)	Major conformer
Peptide 253 (FPro-24)	-173.6	-173.7	*Cis*
Peptide 254 (FPro-22)	-173.9	-173.1	*Trans*
Peptide 255 (FPro-18)	-173.2	-173.0	*Trans*
Peptide 256 (FPro-15)	-173.4	-173.3	*Trans*
Peptide 257 (FPro-11)	-172.4	-171.8	*Trans*
Peptide 258 (FPro-10)	-172.9	-174.1	*Cis*
Peptide 259 (FPro-7)	-173.1	-173.8	*Cis*
Peptide 260 (FPro-5)	-173.4	-173.8	*Cis*
Peptide 261 (FPro-2)	-173.5	-172.9	*Trans*

To be certain that the reported *cis-trans* ratios are correct, or at be certain whether a proline residue occurs as mainly *cis*, it would be beneficial to obtain a crystal structure of the non-fluorinated p53 peptide (**135**) as additional data. Alternatively, more complex NMR studies could be performed (*e.g.* 2D NOESY or HOESY (^{19}F-^{1}H)),[76] after which the ^{19}F signals for the fluorinated peptides would be easier to interpret. Additionally, it is important to state that by introducing a fluorinated proline, the natural *cis-trans* ratios can be influenced, which could make 4*S*-FPro less reliable as a reporter.[38]

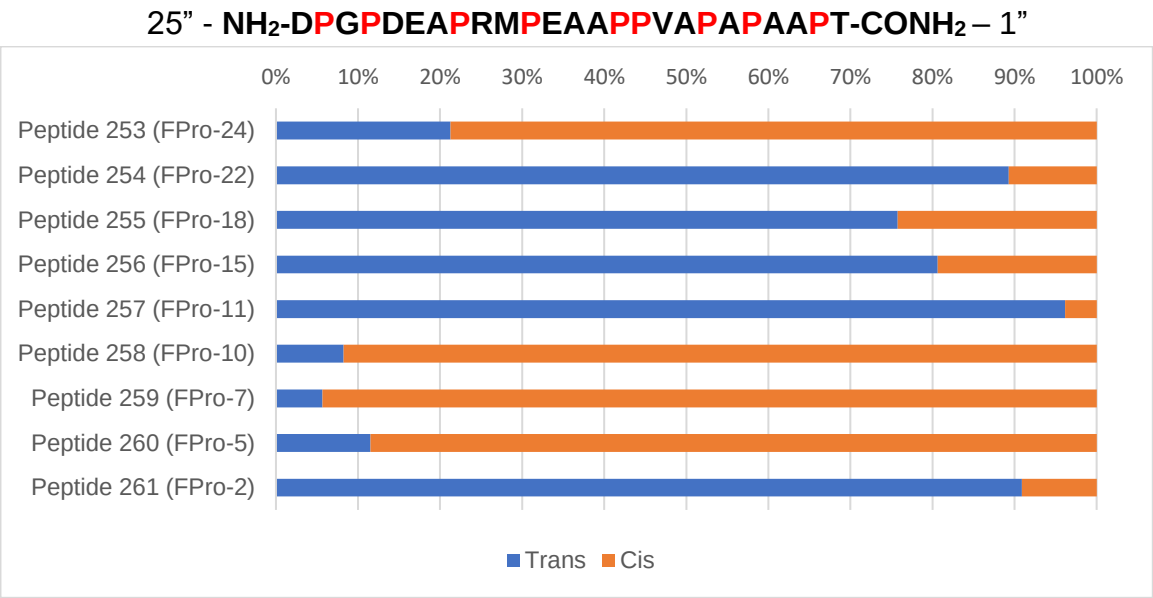


Figure 6.13: *Cis-trans* ratios of individual fluoroproline p53 peptides based on literature chemical shift patterns.

A very significant observation was that each ^{19}F-p53 sequence produced a fluorine signal at a distinct ppm chemical shift, which allowed for an overlay of all the spectra (Figure 6.14). These results suggest that it should be possible to build a single sequence in which each proline residue is replaced by a fluoroproline. However, conformational effects induced by fluorine might cause slight changes in chemical shift when multiple fluoroprolines are incorporated. If a p53 peptide sequence, or full protein, is synthesised using only fluorinated proline residues in the poly-proline motif it could potentially be used to build a 3D model using NMR, which could also be useful for structure conformation, interactions and whole single protein dynamics.

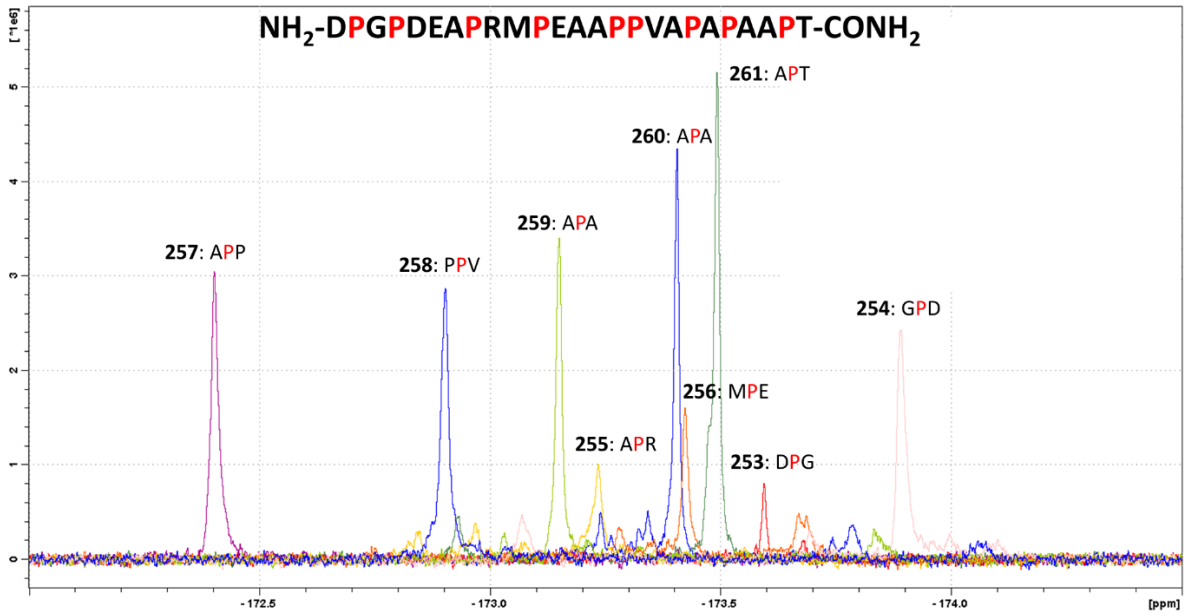


Figure 6.14: Overlay of ^{19}F NMR signals of individual p53 peptide sequences.

6.7 Monitoring PPIase activity of CypD by ^{19}F NMR

As it is known that CypD can bind p53,[30] it was of interest to determine if there would be a site of prolyl-peptidyl isomerism of p53 caused by CypD. Fluorinated p53 peptide sequences (**253-261**) were incubated with CypD and the *cis-trans* ratio was monitored by ^{19}F NMR. As CypD functions as a PPIase, it is expected to change the *cis-trans* ratios of the proline residues it interacts possibly by lowering the energy barrier for interconversion. The recombinant purified CypD was added at a 1:200 dilution (0.5%, 0.0001 mmol), meaning that for each protein molecule, 200 molecules of peptide substrate (0.002 mmol) were available. This was considered to be a suitable starting concentration, as it is unknown at which proline residue binding and thus altered isomerisation would take place. Therefore, by using a low concentration of protein to start with, the amount of (expensive) protein that is used for proline positions that CypD does not bind to is minimised. For residues that show a change in *cis-trans* ratios upon addition of the protein, higher concentration experiments can be performed in future titration studies (*e.g.* 1:100, 1:50, 1:20, 1:10, 1:1).

As previously described, cyclosporin A (CsA) is a well-known active site competitive inhibitor of the cyclophilins. Given the competitive nature of CsA's inhibition of CypD, it was considered

that it would displace bound peptide substrates and restore natural cis-trans ratios. Each peptide was monitored by ^{19}F NMR to determine its native *cis-trans* ratios as described in Section 6.5. Subsequently, CypD was added to each of the fluorinated peptides in a 1:200 dilution, which was then analysed by ^{19}F NMR. Finally, CsA was added to each sample, in a 1:1 fashion compared to peptide (0.002 mmol), which should ensure inhibition of all protein due to the high affinity of CsA for CypD (K_D=13.4 nM[77]). The effect of inhibition on the *cis-trans* ratios was monitored by ^{19}F NMR, as before. The *cis-trans* ratios for peptide **253** (24-FPro) in the presence of CypD was not recorded.

The results of these experiments are shown in Figure 6.13 and Table 6.3. For most of the proline residues a change could be observed in the *cis-trans* ratios upon addition of the protein, with the exception of 2-FPro (**261**) and 10-FPro (**258**). Three of the peptides (**254**, **256**, and **260**) displayed behaviour according to expectations, where the addition of CypD lowered the *major* conformer *%* (- 1-3%), and addition of CsA restored and even increased the *major* conformer *%* (+ 1-2%). Similarly, peptide **255** (18-FPro) had a 2% decrease in the major *trans* conformer upon addition of CypD, and a 5% increase compared to the native peptide *cis-trans* ratio after inhibition with CsA. Finally, for peptide **259** and **257** the addition of CsA did not appear to restore the original *cis-trans* ratios. Overall, the main changes in *cis-trans* ratios appear to occur in peptides **255**, **257**, and **259**.

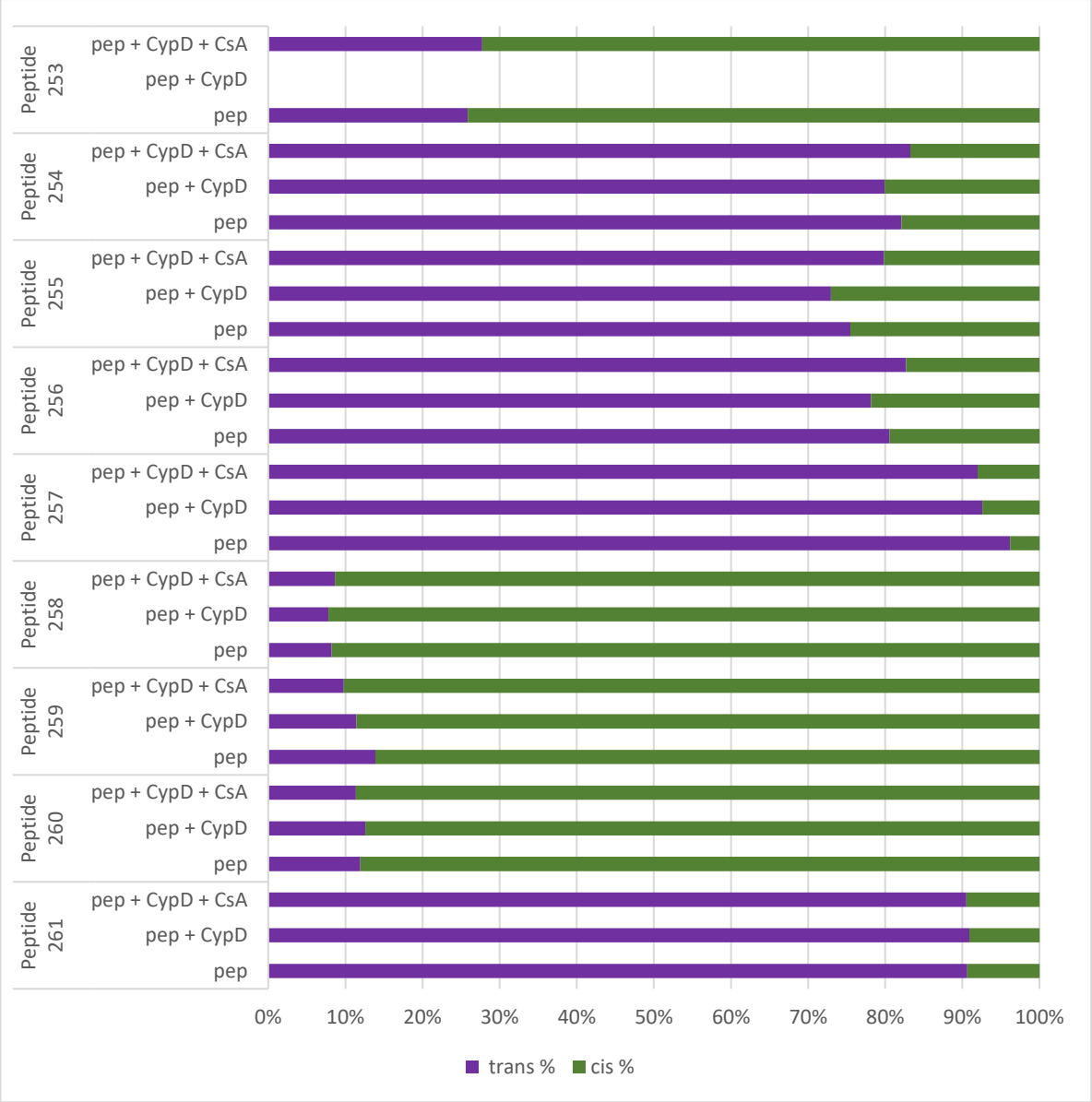


Figure 6.15: Visual overview of determined *cis-trans* ratios by ^{19}F NMR for 1) peptide only; 2) peptide + CypD; and 3) peptide + CypD + CsA.

Table 6.3: Calculated % of *cis* and *trans* conformers for each of the fluorinated peptide sequences. Results show the % for the peptide alone, in combination with CypD, and in combination with both CypD and CsA.

		trans	cis
Peptide **253**	pep	26%	74%
	pep + CypD	0%	0%
	pep + CypD + CsA	28%	72%
Peptide **254**	pep	82%	18%
	pep + CypD	80%	20%
	pep + CypD + CsA	83%	17%
Peptide **255**	pep	75%	25%
	pep + CypD	73%	27%
	pep + CypD + CsA	80%	20%
Peptide **256**	pep	81%	19%
	pep + CypD	78%	22%
	pep + CypD + CsA	83%	17%
Peptide **257**	pep	96%	4%
	pep + CypD	93%	7%
	pep + CypD + CsA	92%	8%
Peptide **258**	pep	8%	92%
	pep + CypD	8%	92%
	pep + CypD + CsA	9%	91%
Peptide **259**	pep	14%	86%
	pep + CypD	11%	89%
	pep + CypD + CsA	10%	90%
Peptide **260**	pep	12%	88%
	pep + CypD	13%	87%
	pep + CypD + CsA	11%	89%
Peptide **261**	pep	91%	9%
	pep + CypD	91%	9%
	pep + CypD + CsA	91%	9%

While differences in *cis-trans* ratios can be observed upon addition of CypD and CsA, it is important to take into account the reliability and sensitivity of the NMR. Interpretation of the data can be difficult due to a high signal-to-noise ratio, which can be the case at lower concentrations (Figure 6.14). In addition, *cis-trans* ratios are based on manual integration of the signals, which can be prone to error due to personal interpretation. Additionally, assigning which peak represents the *cis* species and which peak represents the *trans* species can be difficult. Nevertheless, there is still valuable information to be obtained from crude ^{19}F NMR

data. Interaction between a protein and a ligand has been associated with line broadening of the NMR signals.[78] This phenomenon can be observed for peptide **259**, where line broadening appears to occur when the peptide is in solution with CypD (Figure 6.16). While the resolution is not ideal, this technique can provide a first insight into which proline would be affected by CypD induced prolyl-peptidyl isomerism. In the case of the fluorinated p53 peptides, line broadening appears to occur at position 5 (**260**), 7 (**259**) and 11 (**257**). This could form a starting point for more research into these residues, by resynthesising them, followed by purification and ^{19}F NMR studies at higher concentrations of both peptide and enzyme.

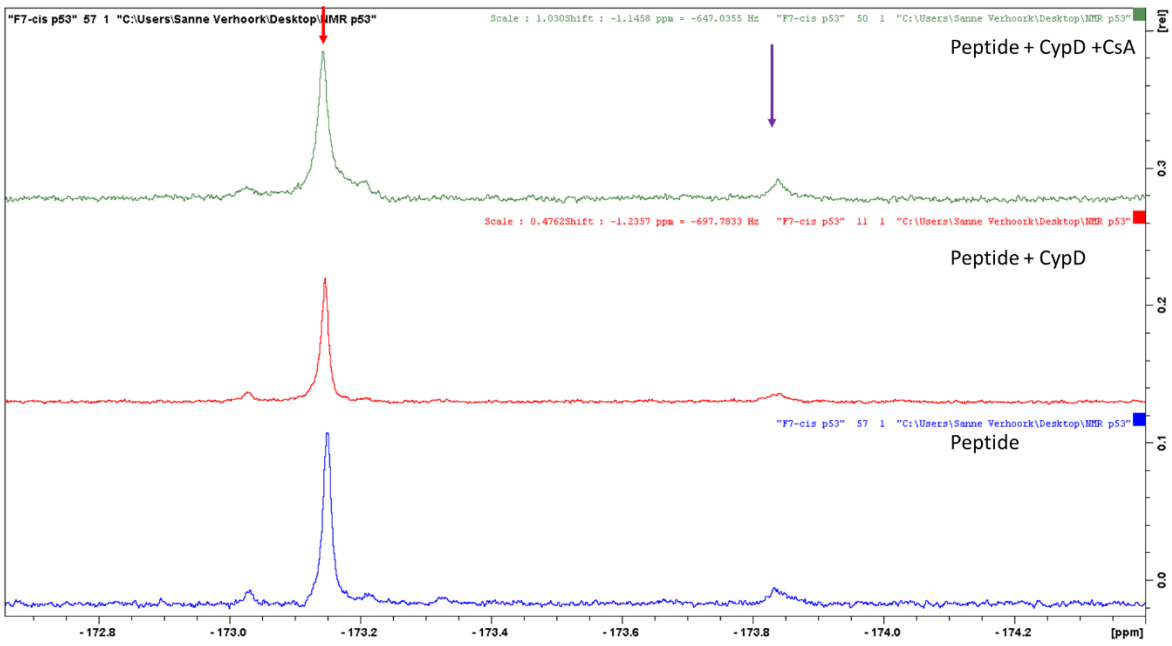


Figure 6.16: NMR signals of peptide **259**; peptide **259** + CypD; peptide **259** + CypD + CsA. The *cis* peptide peak is indicated with the red arrow, the *trans* peptide peak with the purple arrow.

6.8 Conclusion

In this chapter the interaction between CypD and p53 was thoroughly investigated using fluorinated substrates. Prior to studies of the polyproline containing p53 peptide (**135**), short CypD substrates based on the F_1F_0 ATPase subunit B and Bcl2 proteins were synthesised and analysed for their *cis-trans* ratios and interaction with CypD, demonstrating the effectiveness of ^{19}F NMR. Subsequently, synthesis of p53 poly proline sequence **135** was successfully optimised, after which fluorinated proline residues could be incorporated at distinct positions, yielding nine 4*S*-fluorinated sequences. Following isolation of each crude peptide, native *cis-trans* ratios were determined using ^{19}F NMR, which showed that proline 5, 7, 10 and 24 were primarily in the *cis* isoform, whereas residue 2, 11, 15, 18, and 22 occurred primarily in the *trans* isoform. A significant finding was that each fluorinated proline residue had a distinct chemical shift (ppm), which opens up possibilities for future research utilising a fully fluorinated p53 peptide sequence.

Finally, each peptide was incubated with CypD and CsA to determine if changes in *cis-trans* ratios would occur, which was found to be the case for seven out of nine peptides. The main changes appear to occur at residue 7 (**259**), 11 (**257**), and 18 (**255**). In addition to *cis-trans* ratios, it was also of interest to observe if line broadening would occur in the presence of CypD. This was found to be the case for peptide **260** (FPro-5), **259** (FPro-7), and **257** (FPro-11). Combining these results, it shows that peptide **259** (FPro-7) and **257** (FPro-11) could be isomerised by CypD. This could be confirmed in future studies using purified peptides at higher concentrations.

6.9 References

1. Cheng, H. N. & Bovey, F. A. Cis-Trans equilibrium and kinetic studies of acetyl-L-proline and glycyl-L-proline. *Biopolymers* **16**, 1465–1472 (1977).

2. Matouschek, A., Rospert, S., Schmid, K., Glick, B. S. & Schatz, G. Cyclophilin catalyzes protein folding in yeast mitochondria. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 6319–23 (1995).

3. Göthel, S. F. & Marahiel, M. A. Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts. *Cell. Mol. Life Sci.* **55**, 423–436 (1999).

4. Rostam, M. A. *et al.* Peptidyl-prolyl isomerases: Functionality and potential therapeutic targets in cardiovascular disease. *Clin. Exp. Pharmacol. Physiol.* **42**, 117–124 (2015).

5. Schiene-Fischer, C. Multidomain Peptidyl Prolyl cis/trans Isomerases. *Biochim. Biophys. Acta - Gen. Subj.* **1850**, 2005–2016 (2015).

6. Schmidpeter, P. A. M. & Schmid, F. X. Prolyl isomerization and its catalysis in protein folding and protein function. *J. Mol. Biol.* **427**, 1609–1631 (2015).

7. Wang, P. & Heitman, J. The cyclophilins. *Genome Biol.* **6**, 226 (2005).

8. Takahashi, N., Hayano, T. & Suzuki, M. Peptidyl-prolyl cis-trans isomerase is the cyclosporin A-binding protein cyclophilin. *Nature* **337**, 473–5 (1989).

9. Liu, J. *et al.* Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell* **66**, 807–815 (1991).

10. Halestrap, A. P. Mitochondria and reperfusion injury of the heart-A holey death but not beyond salvation. *Journal of Bioenergetics and Biomembranes* **41**, 113–121 (2009).

11. Di Lisa, F. & Bernardi, P. A CaPful of mechanisms regulating the mitochondrial permeability transition. *Journal of Molecular and Cellular Cardiology* **46**, 775–780 (2009).

12. Galat, A. Peptidylprolyl cis/trans isomerases (immunophilins): biological diversity--targets--functions. *Curr. Top. Med. Chem.* **3**, 1315–47 (2003).

13. Mi, H., Kops, O., Zimmermann, E., Jäschke, A. & Tropschug, M. A nuclear RNA-binding cyclophilin in human T cells. *FEBS Lett.* **398**, 201–205 (1996).

14. Andreeva, L., Heads, R. & Green, C. J. Cyclophilins and their possible role in the stress response. *Int. J. Exp. Pathol.* **80**, 305–15 (1999).

15. Hamilton, G. S. & Steiner, J. P. Immunophilins: Beyond Immunosuppression. *J. Med. Chem.* **41**, 5119–5143 (1998).

16. Hardy, J. & Selkoe, D. J. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* **297**, 353–356 (2002).

17. Haworth, R. A. & Hunter, D. R. The Ca2+-induced membrane transition in mitochondria: I. The protective mechanisms. *Arch. Biochem. Biophys.* **195**, 460–467 (1979).

18. Nieminen, A.-L. Apoptosis and necrosis in health and disease: role of mitochondria. *Int. Rev. Cytol.* **224**, 29–55 (2003).

19. Baines, C. P. Role of the mitochondrion in programmed necrosis. *Front. Physiol.* **1 NOV**, 156 (2010).

20. Jeong, S.-Y. & Seol, D.-W. The role of mitochondria in apoptosis. *BMB Rep.* **41**, 11–22 (2011).

21. Chan, D. C. Mitochondria: Dynamic Organelles in Disease, Aging, and Development. *Cell* **125**, 1241–1252 (2006).

22. Kang, J. S. *et al.* Docking of Axonal Mitochondria by Syntaphilin Controls Their Mobility and Affects Short-Term Facilitation. *Cell* **132**, 137–148 (2008).

23. Giorgio, V. *et al.* Cyclophilin D modulates mitochondrial F0F1-ATP synthase by interacting with the lateral stalk of the complex. *J. Biol. Chem.* **284**, 33982–33988 (2009).

24. Nigro, J. M. *et al.* Mutations in the p53 gene occur in diverse human tumour types. *Nature* **342**, 705–708 (1989).

25. Hollstein, M., Sidransky, D., Vogelstein, B. & Harris, C. C. p53 mutations in human cancers. *Science* **253**, 49–53 (1991).

26. Walker, K. K. & Levine, A. J. Identification of a novel p53 functional domain that is necessary for efficient growth suppression. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 15335–40 (1996).

27. Venot, C. *et al.* The requirement for the p53 proline-rich functional domain for mediation of apoptosis is correlated with specific PIG3 gene transactivation and with transcriptional repression. *EMBO J.* **17**, 4668–79 (1998).

28. Vousden, K. H. & Lu, X. Live or let die: the cell's response to p53. *Nat. Rev. Cancer* **2**, 594–604 (2002).

29. Kroemer, G., Galluzzi, L. & Brenner, C. Mitochondrial Membrane Permeabilization in Cell Death. *Physiol. Rev.* **87**, 99–163 (2007).

30. Fayaz, S. M. & Rajanikant, G. K. Modelling the molecular mechanism of protein–protein interactions and their inhibition: CypD–p53 case study. *Mol. Divers.* **19**, 931–943 (2015).

31. Berger, M., Vogt Sionov, R., Levine, A. J. & Haupt, Y. A role for the polyproline domain of p53 in its regulation by Mdm2. *J. Biol. Chem.* **276**, 3785–90 (2001).

32. Venot, C. *et al.* The requirement for the p53 proline-rich functional domain for mediation of apoptosis is correlated with specific PIG3 gene transactivation and with transcriptional repression. *EMBO J.* **17**, 4668–4679 (1998).

33. MacArthur, M. W. & Thornton, J. M. Influence of proline residues on protein conformation. *J. Mol. Biol* **218**, 397–412 (1991).

34. DeRider, M. L. *et al.* Collagen Stability: Insights from NMR Spectroscopic and Hybrid Density Functional Computational Investigations of the Effect of Electronegative Substituents on Prolyl Ring Conformations. *J. Am. Chem. Soc.* **124**, 2497–2505 (2002).

35. Cremer, D. & Pople, J. A. General definition of ring puckering coordinates. *J. Am. Chem. Soc.* **97**, 1354–1358 (1975).

36. Chang, D.-K., Cheng, S.-F., Trivedi, V. D. & Lin, K.-L. Proline Affects Oligomerization of a Coiled Coil by Inducing a Kink in a Long Helix. *J. Struct. Biol.* **128**, 270–279 (1999).

37. Dugave, C. & Demange, L. Cis−Trans Isomerization of Organic Molecules and Biomolecules: Implications and Applications. *Chem. Rev.* **103**, 2475–2532 (2003).

38. Verhoork, S. J. M., Killoran, P. M. & Coxon, C. R. Fluorinated prolines as conformational probes and ^{19}F NMR reporters in peptide and protein chemistry. *Biochemistry* **57**, 6132–6143 (2018).

39. Purser, S. *et al.* Fluorine in medicinal chemistry. *Chem. Soc. Rev.* **37**, 320–330 (2008).

40. Zhou, Y. *et al.* Next Generation of Fluorine-Containing Pharmaceuticals, Compounds Currently in Phase II−III Clinical Trials of Major Pharmaceutical Companies: New Structural Trends and Therapeutic Areas. *Chem. Rev.* **116**, 422–518 (2016).

41. Park, B. K., Kitteringham, N. R. & O'Neill, P. M. Metabolism of fluorine-containing drugs. *Annu. Rev. Pharmacol. Toxicol.* **41**, 443–470 (2001).

42. Mishra, N. K., Urick, A. K., Ember, S. W. J., Schonbrunn, E. & Pomerantz, W. C. Fluorinated aromatic amino acids are sensitive ^{19}F NMR probes for bromodomain-ligand interactions. *ACS Chem. Biol.* **9**, 2755–2760 (2014).

43. Richards, K. L., Rowe, M. L., Hudson, P. B., Williamson, R. A. & Howard, M. J. Combined ligand-observe ^{19}F and protein-observe 15N,1H-HSQC NMR suggests phenylalanine as the key Δ-somatostatin residue recognized by human protein disulfide isomerase. *Sci. Rep.* **6**, 19518 (2016).

44. Marsh, E. N. G. & Suzuki, Y. Using ^{19}F NMR to Probe Biological Interactions of Proteins and Peptides. *ACS Chem. Biol.* **9**, 1242–1250 (2014).

45. Krane, S. M. The importance of proline residues in the structure, stability and susceptibility to proteolytic degradation of collagens. *Amino Acids* **35**, 703–710 (2008).

46. Chorghade, M. S. *et al.* Practical syntheses of 4-fluoroprolines. *J. Fluor. Chem.* **129**, 781–784 (2008).

47. Golbik, R. *et al.* Peptidyl Prolyl cis/trans-Isomerases: Comparative Reactivities of Cyclophilins, FK506-Binding Proteins, and Parvulins with Fluorinated Oligopeptide and Protein Substrates. *Biochemistry* **44**, 16026–16034 (2005).

48. Briggs, C. R. S. *et al.* The observation of a large gauche preference when 2-fluoroethylamine and 2-fluoroethanol become protonated. *Org. Biomol. Chem.* **2**, 732–740 (2004).

49. Thomas, C. A., Talaty, E. R. & Bann, J. G. 3S-Fluoroproline as a probe to monitor proline isomerization during protein folding by 19F-NMR. *Chem. Commun.* **0**, 3366 (2009).

50. Kim, W., Hardcastle, K. I. & Conticello, V. P. Fluoroproline flip-flop: Regiochemical reversal of a stereoelectronic effect on peptide and protein structures. *Angew. Chem., Int. Ed.* **45**, 8141–8145 (2006).

51. Hofman, G.-J. *et al.* Minimising conformational bias in fluoroprolines through vicinal difluorination. *Chem. Commun.* **54**, 5118–5121 (2018).

52. Jlalia, I. *et al.* Synthesis of an MIF-1 analogue containing enantiopure (S)-α- trifluoromethyl-proline and biological evaluation on nociception. *Eur. J. Med. Chem.* **62**, 122–129 (2013).

53. Kubyshkin, V. *et al.* γ-(S)-Trifluoromethyl proline: evaluation as a structural substitute of proline for solid state ^{19}F-NMR peptide studies. *Org. Biomol. Chem.* **13**, 3171–3181 (2015).

54. Mykhailiuk, P. K. *et al.* Synthesis of trifluoromethyl-substituted proline analogues as ^{19}F NMR labels for peptides in the polyproline II conformation. *Angew. Chem., Int. Ed.* **47**, 5765–5767 (2008).

55. Pandey, A. K., Naduthambi, D., Thomas, K. M. & Zondlo, N. J. Proline editing: A general and practical approach to the synthesis of functionally and structurally diverse peptides. Analysis of steric versus stereoelectronic effects of 4-substituted prolines on conformation within peptides. *J. Am. Chem. Soc.* **135**, 4333–4363 (2013).

56. Tressler, C. M. & Zondlo, N. J. (2 S,4 R)- and (2 S,4 S)-Perfluoro- tert -butyl 4-hydroxyproline: Two conformationally distinct proline amino acids for sensitive application in ^{19}F NMR. *J. Org. Chem.* **79**, 5880–5886 (2014).

57. Bretscher, L. E., Jenkins, C. L., Taylor, K. M., DeRider, M. L. & Raines, R. T. Conformational stability of collagen relies on a stereoelectronic effect. *J. Am. Chem. Soc.* **123**, 777–778 (2001).

58. Crespo, M. D. & Rubini, M. Rational Design of Protein Stability: Effect of (2S,4R)-4-Fluoroproline on the Stability and Folding Pathway of Ubiquitin. *PLoS One* **6**, e19425 (2011).

59. Kitamoto, T., Ozawa, T., Abe, M., Marubayashi, S. & Yamazaki, T. Incorporation of fluoroprolines to proctolin: Study on the effect of a fluorine atom toward peptidic conformation. *J. Fluor. Chem.* **129**, 286–293 (2008).

60. Wedemeyer, W. J., Welker, E. & Scheraga, H. A. Proline cis-trans isomerization and protein folding. *Biochemistry* **41**, 14637–14644 (2002).

61. Renner, C. *et al.* Fluoroprolines as Tools for Protein Design and Engineering. *Angew. Chem., Int. Ed.* **40**, 923–925 (2001).

62. Thomas, K. M., Naduthambi, D. & Zondlo, N. J. Electronic Control of Amide cis−trans Isomerism via the Aromatic−Prolyl Interaction. *J. Am. Chem. Soc.* **128**, 2216–2217 (2006).

63. Reimer, U. *et al.* Side-chain effects on peptidyl-prolyl cis/trans isomerisation. *J. Mol. Biol.* **279**, 449–460 (1998).

64. Ganguly, H. K., Kaur, H. & Basu, G. Local control of cis-peptidyl-prolyl bonds mediated by CH···π interactions: the Xaa-Pro-Tyr motif. *Biochemistry* **52**, 6348–57 (2013).

65. Sarkar, S. K., Young, P. E., Sullivan, C. E. & Torchia, D. A. Detection of cis and trans X-Pro peptide bonds in proteins by ^{13}C NMR: application to collagen. *Proc. Natl. Acad. Sci. U. S. A.* **81**, 4800–4803 (1984).

66. Cobb, S. L. & Murphy, C. D. ^{19}F NMR applications in chemical biology. *J. Fluor. Chem.* **130**, 132–143 (2009).

67. Harper, D. B. & O'Hagan, D. The fluorinated natural products. *Nat. Prod. Rep.* **11**, 123–133 (1994).

68. Lee, Y.-J. *et al.* Genetically encoded fluorophenylalanines enable insights into the recognition of lysine trimethylation by an epigenetic reader. *Chem. Commun.* **52**, 12606–12609 (2016).

69. Sarker, M. *et al.* Tracking Transitions in Spider Wrapping Silk Conformation and Dynamics by ^{19}F Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **55**, 3048–3059 (2016).

70. Abboud, M. I. *et al.* ^{19}F-NMR Reveals the Role of Mobile Loops in Product and Inhibitor Binding by the São Paulo Metallo-β-Lactamase. *Angew. Chem., Int. Ed.* **56**, 3862–3866 (2017).

71. Kenward, C., Shin, K., Sarker, M., Bekkers, C. & Rainey, J. K. Comparing and Contrasting

Fluorotryptophan Substitutions for ^{19}F Membrane Protein NMR Spectroscopy. *Biophys. J.* **112**, 499a (2017).

72. Lin, Y.-J. & Horng, J.-C. Impacts of terminal (4R)-fluoroproline and (4S)-fluoroproline residues on polyproline conformation. *Amino Acids* **46**, 2317–2324 (2014).

73. Lummis, S. C. R. *et al.* Cis-trans isomerization at a proline opens the pore of a neurotransmitter-gated ion channel. *Nature* **438**, 248–252 (2005).

74. Adams, J. M., Cory, S. & Science. The Bcl-2 protein family: arbiters of cell survival. *Science (80-.).* **281**, 1322–1326 (1998).

75. Galluzzi, L., Kepp, O. & Kroemer, G. Mitochondria: master regulators of danger signalling. *Nat Rev Mol Cell Biol* **13**, 780–788 (2012).

76. Marion, D. An introduction to biological NMR spectroscopy. *Mol. Cell. proteomics* **12**, 3006–30025 (2013).

77. Gutiérrez-Aguilar, M. & Baines, C. P. Structural mechanisms of cyclophilin D-dependent control of the mitochondrial permeability transition pore. *Biochimica et Biophysica Acta - General Subjects* **1850**, 2041–2047 (2015).

78. Kleckner, I. R. & Foster, M. P. An introduction to NMR-based approaches for measuring protein dynamics. *Biochimica et Biophysica Acta - Proteins and Proteomics* **1814**, 942–968 (2011).

7 Conclusions and future work

This thesis explored the synthesis and purification methods of biologically active peptides, as well as methods for evaluation of its biological activity.

A new synthesis and purification method was successfully developed for the synthesis of difficult to synthesise biologically active peptides, making use of acrylamide-tagging steps after each amino acid coupling step. The methods were optimised for both manual and automated solid phase peptide synthesis using several short model peptides (3-6 amino acids), prior to the synthesis of longer, more complex biologically active peptides. The effectiveness of the new synthesis-purification method was demonstrated using a 25 amino acid p53 sequence and the drug exenatide (39 amino acids). The importance of efficient purification methods is stressed by the large quantities of waste generated using the more traditional HPLC purification protocols. The acrylamide-scavenging procedure forms a powerful tool to avoid HPLC purification, making peptide synthesis less wasteful and a greener process.

Upon the development of the acrylamide-scavenging method, the hypothesis was formed that N-terminal acrylamide could also be used for an intramolecular thiol-acrylamide conjugate addition. However, experiments did not provide conclusive results. Future experiments could be undertaken to determine optimal reaction conditions and macrocycle ring-size.

An alternative approach to macrocyclisation of peptides can be undertaken by perfluoroaryl—stapling between thiol-bearing residues, which stabilises a peptide in an α-helical conformation. It was found that the nature and stereochemistry of the cross-linked cysteine analogues can influence biological activity (target affinity and selectivity). These findings are important as there are many reported ways to cross-link disulfides, however, the influence of stereochemistry, homologation or beta-carbon steric hindrance had not yet been considered. The obtained knowledge is directly translatable to other research in the field of thiol-based peptide-stapling. Additionally, this knowledge can be expanded to the synthesis of non α-

helical systems, for example disulfide rich mini-proteins *e.g.* cyclotides, as well as other peptides for different biological targets.

Finally, the uses of fluorine as a reporter in peptide and protein chemistry has been demonstrated to be a high sensitivity probe for dynamic biological processes at the single molecule level. The power of fluorination and ^{19}F NMR has been shown using substrates for the peptidyl-prolyl isomerase CypD. Fluorinated proline residues were successfully incorporated in Bcl2, F_1F_0 ATPase subunit B, and p53 sequences. The *cis-trans* ratios for each of the peptides could be determined by ^{19}F NMR and binding to CypD could be observed by line-broadening in the spectra. Additionally, it was shown that in a poly-proline sequence, each proline had a distinct chemical shift, raising interest to synthesise peptides containing multiple fluorinated prolines.

In conclusion, new synthetic approaches have been successfully developed, which can now be applied to the synthesis of new complex biologically active peptides. Additionally, methods have been developed to use peptides as tools for understanding protein-protein interactions and other biological processes. The new synthetic and biological evaluation strategies can be applied to a wide range of peptides and proteins, providing valuable tools for future research.

7.1 Future work

7.1.1 Acrylamide-scavenging

The work in this thesis has set a basis for future experiments. While significant evidence was found for the functioning of the acrylamide-scavenging peptide purification, additional experiments should be carried out. First, previously conducted experiments should be repeated to demonstrate reproducibility. This will allow for any errors that have been introduced by measurements to be identified. For example, in Figure 3.2 an increase in 'deletion' peptide **103** was observed over time, while it should have been scavenged from the mixture. By repeating the experiment in the exact same way as before, this observation can be better understood. In line with this is the observation of a side product that has been

observed during the model scavenging studies (e.g. during experiments using peptide **60**, Figure 2.13). While this side product formation has formed a starting point for future studies into peptide cyclisation, the product has not been isolated for further analysis. If the side product is isolated it can be further studies by NMR to get full characterisation.

Further studying the optimal pH conditions for the scavenging reaction can from another approach to prevent side product formation. The pKa of a thiol (8.00 for cysteine, 9.52 for N-acetyl-L-cysteine) is lower than that of lysine (pKa 12.48). Finding the right pH balance between promoting thiol-based reactivity while preventing amine reactivity can possibly lead to cleaner scavenging reactions.

Another experiment that should be conducted is following a more quantitative approach. During the experiments described in this thesis the yields of the purification (acrylamide-scavenging) procedure were not determined, as the quality (purity) of the purification method was considered to be more important. However, it would be beneficial to gain insight in the quantities of peptide obtained post-purification. A known amount of crude peptide can be added to the thiol-resin for scavenging, after which the remaining solution can be lyophilised to yield the purified peptide. The amount and purity of this can be determined, which will give further insights into the efficiency of the acrylamide-scavenging method. Additionally, a decrease in signal intensity in some of the LC-MS results was observed (e.g. Figure 3.4, 3.17). This could indicate interaction of the non-tagged peptide with the thiol-resin. To confirm whether this takes place or not, a known amount of non-tagged, pure, desired peptide can be added to the thiol-resin for a known amount of time. The remaining solution can then be lyophilised and analysed. This will give insights in any product loss that might occur.

To ensure efficiency of purification it is important that the acrylamide-tagging functions equally well for each amino acid. An experiment should be performed on a short model (tri-)peptide at which the N-terminus is varied to be all the twenty standard amino acids (e.g. NH_2-XAA-●, where X = any amino acid). The coupling efficiency of the acrylamide-group and the resulting purity of the peptides can then be analysed by LC-MS. Conducting this experiment should

help understanding the observed changes in purity of the deletions from peptide **107** (NH_2-TWEKMPY-$CONH_2$), where the crude synthesis of peptide **108** (Acr-WEKMPY-$CONH_2$) and **111** (Acr-MPY-$CONH_2$) displayed impurities when analysed by LC-MS. Additionally, work can be undertaken to make the acrylamide-scavenging purification method applicable to cysteine-bearing peptides. This can be done by experimenting with orthogonal cysteine protection groups. There are multiple orthogonal protecting groups available for thiols, including *p*-methylbenzyl (Meb; HF, $MeSiCl_3$ cleavable), acetamidomethyl (Acm; I_2, DTNP, Tl(III), Hg(II) cleavable), which can be explored. It is of importance that removal of this protection group can take place after cleavage from the resin and acrylamide-scavenging purification without introducing new impurities.

Furthermore, the photochemical approach that was introduced in Section 3.4 can be further investigated. During the initial experiment a molecular weight that corresponded to the addition of a hydroxyl-group was observed. Upon repetition of the photochemical experiment following the same procedure as before (using degassed DMF), the resulting product can be isolated and further studied by NMR. This can confirm if the gain in molecular weight indeed corresponds to an additional hydroxyl-group and if so, it can help identify where this addition takes place. If the reactivity is better understood, an altered photochemical procedure can be developed where this side reaction can be avoided.

Finally, the synthesis of Exenatide can be further optimised by introducing double couplings. Making use of double couplings in combination with the acrylamide-tagging and scavenging approach, the need to HPLC could possibly be eliminated. If this has led to desirable drug discovery grade purity (> 95%), the acrylamide-scavenging experiments will be further extended to the synthesis and purification of more complex and diverse biologically active peptides. Lastly, upon completion of optimisation, a more standardised procedure will be developed, allowing for parallel synthesis and purification, allowing for quick and clean synthesis of new peptide libraries.

7.1.2 Peptide cyclisation

Peptide cyclisation *via* thiol-acrylamide conjugate addition was attempted in Chapter 4. While preliminary data suggested that it was possible that cyclisation took place, this has not been confirmed. More experiments are required, including repetition of experiments already conducted. It would be of interest to isolate the obtained peptide products that are described in Section 4.3 and 4.4. Upon isolation the peptides can be further analysed by NMR which could provide better understanding in the products that were observed by LC-MS.

To gain further insight into the optimal ring size, it would also be of interest to perform a 'macrocyclisation-scan'. For this, a cysteine residue would be introduced at different locations in an otherwise identical peptide sequence. Each peptide can then be cyclised under the same reaction conditions, after which the yield of each cyclisation reaction is compared. It could be that some ring sizes are cyclised more efficiently than others.

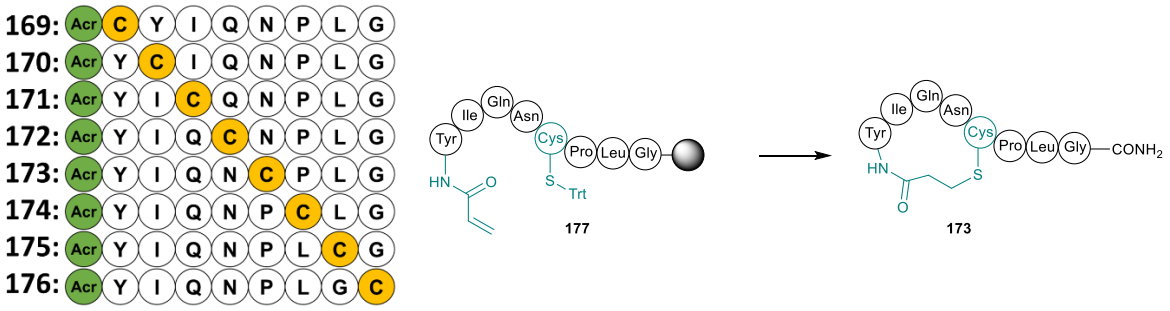
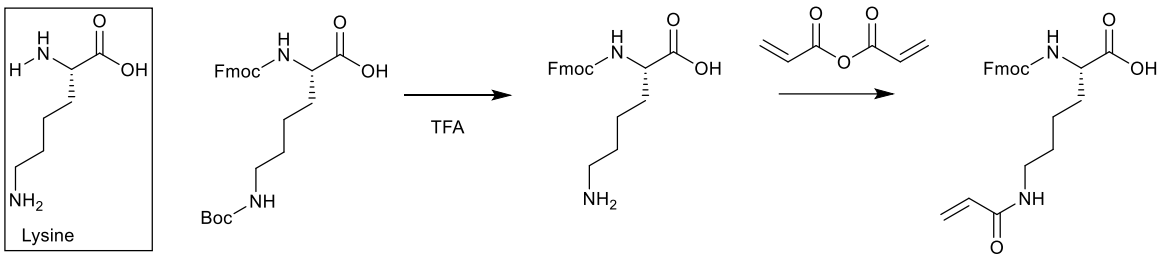


Figure 7.1: Example of macrocyclisation scan using Oxytocin peptides.

To investigate the substrate scope of this novel macrocyclisation strategy, oxytocin **139** (NH_2-CYIQNCPLG-$CONH_2$), a small cyclic peptide consisting of nine amino acids and one disulfide bridge between Cys-1 and Cys-6 and could be used as a model peptide. The N-terminal cysteine can be replaced by an acrylamide-tag, keeping the original location of the disulphide bond in place for peptide **173**. To probe for peptide ring size tolerance, a series of 8 peptides could be designed, featuring *i,i+1* through to *i,i+8* ring sizes (Figure 4.10).

Once a procedure for N-terminal-acrylamide to thiol cyclisation has been optimised, this technique could be expanded to different amino acids. For example, acrylamide could be coupled to the amine group in the lysine side chain (Scheme 4.6), resulting in a new Fmoc-

Lys(Acr)-OH building block, which could be incorporated in a range of peptides. Using this building block, an acrylamide group could be introduced at any place within a peptide sequence, increasing the possibilities for more diverse peptide stapling or cyclisation.



Scheme 7.1: Proposed synthesis route for Fmoc-Lys(Acr)-OH.

7.1.3 Fluorinated reporters

The work using fluorine as a reporter to study biological complexes will be continued. In future experiments, the acrylamide-scavenging synthesis and purification method can be used for the synthesis of complex fluorine-containing biologically active peptides. The p53 poly-proline sequences that showed potential isomerisation in the presence of CypD (**259** and **257**) will be resynthesised, purified and retested for their biological activity at higher enzyme concentration, similarly to the experiments in Section 6.4. Furthermore, the *cis-trans* ratios for peptide **253** in the presence of CypD was not recorded. This experiment will be repeated to complete the data presented in Figure 6.15.

For the work in this thesis, the *cis-trans* ratios for each peptide have only been determined by ^{19}F NMR, however, this is not sufficient to assign the *cis* or *trans* conformer with absolute certainty. To do this, peptides of interest (**259** and **257**) will be resynthesised and purified (> 95%) by prep-HPLC, and then analysed in 2D NOESY NMR experiments.

Additionally, using fluorine as a reporter to understand peptide and protein conformation will be expanded to using other amino acid residues (*e.g.* fluorophenylalanine). The advantage of using a different amino acid as reporter, instead of proline, is the elimination of conformational bias that is introduced by the fluorine-ammonium *gauche* effect for 3- or 4-fluoroprolines. To further understand the *cis-trans* preferences of proline, a study will be conducted that looks at

the effect of each standard amino acid neighbouring proline on either side (Figure 7.2). *Cis-trans* ratios will be determined by ^{19}F NMR, and will be used to gain more insight in protein folding and dynamics.

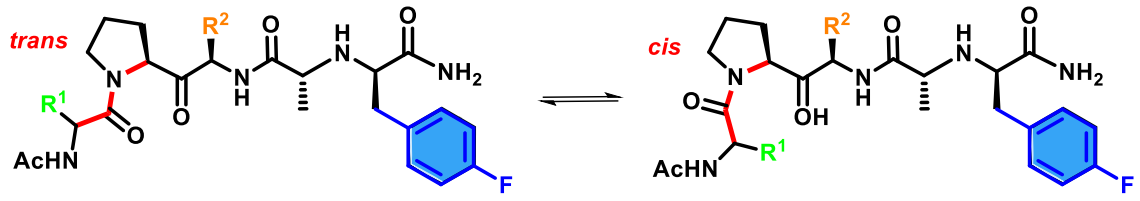


Figure 7.2: Fluorophenylalanine to study the effect of amino acid **X** in sequences X-Pro-Ala-Ala-FPhe and Tyr-Pro-X-Ala-FPhe on measured *cis/trans* ratio in our model peptides, respectively.

8 Experimental

8.1 Materials

All Fmoc L-amino acids (CEM), Fmoc D-amino acids (Merck), Rink Amide ProTide resin (100-200 mesh, 0.55 mmol/g; CEM), diisopropylcarbodiimide (DIC; Sigma), Oxyma Pure™ (CEM), *N*.*N*'-dimethylformamide 99%+ extra pure (DMF; Fisher UK), diisopropylethylamine (DIPEA; Merck Millipore) and piperidine (Sigma) were purchased from commercial suppliers and used directly as indicated in the appropriate experimental procedures. PyBOP was purchased from Novabiochem. All fluorinated Fmoc L-amino acids and acrylic anhydride were purchased from Fluorochem. Thiol-resin was purchased from Biotage. All other chemicals and reagents (acryloyl chloride, crotonic anhydride, 1,4-Diazabicyclo[2.2.2]octane (DABCO), 1,8-Diazabicyclo[5.4.0]undec-7-ene (DBU), 2,2-dimethoxy-2-phenylacetophenone (DMPA), Fmoc-Cys(Mmt)-OH, Fmoc-Pen-OH, Hexafluorobenzene, N-acetyl-L-Cysteine, N-acetyl-L-Lysine, Trifluoroacetic acid (TFA), triisopropylsilane (TIPS) were purchased from Sigma Aldrich and solvents (HPLC grade) were purchased from Fisher Scientific.

8.2 General peptide synthesis procedures

8.2.1 General procedure A: Manual solid phase peptide synthesis

Rink amide ProTide resin (200 mg, 0.78 mmol/g loading; 0.156 mmol) was added to a 10 mL polypropylene fritted syringe. Fmoc deprotection occurred in the presence of piperidine (4 mL, 20% *v/v*; in DMF; 587 eq.) at room temperature for 2x 15 minutes. SPPS couplings employed Fmoc amino acid (0.624 mmol; 4 eq.), PyBOP (0.624 mmol; 4 eq.) and DIPEA (1.248 mmol; 8 eq.) in 7.5 mL DMF. Standard coupling procedures employed single coupling of each amino acid (1 hour, room temperature). Upon completion of each coupling or deprotection step the resin was washed with 2x 5 mL DMF and 1x 5 mL diethyl ether.

8.2.2 General procedure B: Microwave-assisted manual solid peptide synthesis

Rink amide ProTide resin (180mg, 0.56 mmol/g loading; 0.1 mmol) was added to a 10 mL microwave vial. Fmoc deprotection occurred in the presence of piperidine (4 mL, 20% *v/v*; in DMF; 587 eq.) at 90°C for 2 minutes with microwave heating using a Discover® SP Microwave Synthesiser (CEM). Upon completion of a reaction step the content of the microwave vial were flushed through a fritted polypropylene syringe to dispose of excess reagent. Subsequently the remaining solid resin was washed with 2x 10 mL DMF and then re-added to the microwave vial for the next reaction step. SPPS couplings employed Fmoc amino acid stock solutions (2.5 mL, 0.2 M in DMF; 5 eq.), DIC (1 mL, 1 M in DMF; 10 eq.) and Oxyma (0.5 mL, 1 M in DMF; 5 eq.). Standard coupling procedures employed single coupling of each amino acid (5 min, 90 °C).

8.2.3 General procedure C: Microwave-assisted automated solid phase peptide synthesis

Linear peptide sequences were prepared using automated Fmoc-SPPS methods on a Liberty Blue microwave-assisted peptide synthesiser (CEM). Solid phase synthesis was conducted using Rink amide ProTide resin (180 mg, 0.56 mmol/g loading; 0.1 mmol) unless otherwise stated. SPPS was conducted employing the required Fmoc amino acids (0.2 M in DMF; 5 eq.); with DIC (1 M stock solution in DMF; 10 eq.), Oxyma Pure (1M stock solution, 5 eq.) and piperidine (20% *v/v* in DMF; 587 eq., 4 mL) as activator, and deprotection, respectively. Standard coupling procedures employed double coupling of each amino acid (2.5 min, 90 °C).

Amino acids bearing thermally-sensitive protecting groups *e.g.,* Fmoc-L-Cys(Trt)-OH, Fmoc-D-Cys(Trt)-OH, Fmoc-Pen(Trt)-OH, Fmoc-hCys(Trt)-OH and Fmoc-His(Boc)-OH were coupled under milder conditions (50°C for 10 min).

8.2.4 General procedure D: Cleavage of peptides from resin

Peptides were cleaved from the resin as the C-terminal amide by treatment with a cleavage cocktail (5 mL; comprising TFA, TIPS and water (9 : 0.5 : 0.5 *v/v*) with regular shaking at room temperature for 4 h. Peptides were precipitated from cleavage solutions by dropwise addition into cold diethyl ether followed by centrifugation. The resulting pellet was successively suspended in cold diethyl ether and centrifuged twice further. The solids obtained were dissolved in water/MeCN (depending upon solubility), frozen and lyophilised.

8.2.5 General Procedure E: Kaiser test for coupling completion

Two stock solutions were prepared; A) 500 mg ninhydrin in 10 mL ethanol and B) 8 g phenol in 2 mL ethanol. A small sample of resin (1-2 mg) was added to a 1.5 mL Eppendorf tube, after which 5 drops of solution A and solution B were added and mixed. The tube was heated to 100°C for the duration of 5 minutes. The colour of the resin beads was then judged by eye, where blue represented incomplete amino acid couplings and yellow/transparent successful couplings.

8.2.6 General procedure F: Acetyl-capping of peptide N-terminus

Following on-resin synthesis of the appropriate sequence, N-terminal capping was performed using Ac_2O/DMF (20% *v/v*, 2 x 15 min) with shaking at room temperature, prior to cleavage from the resin.

8.2.7 General procedure G: Manual conventional acryloyl chloride N-capping

On-resin peptide (ProTide, resin loading 0.55 mmol/g; 25 mg) was loaded in a 1.5 mL centrifuge tube. One of the following concentrations of acryloyl chloride was added to each experiment:

A. 1 M in DMF
B. 0.5 M in DMF
C. 0.25 M in DMF
D. 0.125 M in DMF
E. 0.02 M in DMF

The reaction was carried out at room temperature for 15 minutes with shaking at 600 rpm. Upon completion of each experiment, peptides were cleaved following general procedure D and analysed using HPLC method J.

8.2.8 General procedure H: Manual microwave-assisted acryloyl chloride N-capping

On-resin peptide (ProTide, resin loading 0.55 mmol/g; 25 mg) was loaded in a

10 mL microwave vial. Reactions were carried out using a Discover® SP Microwave Synthesiser (CEM) for 1 min at 90°C for the following concentrations of acryloyl chloride:

A. 0.5 M in DMF
B. 0.25 M in DMF

Upon completion of each experiment, peptides were cleaved following general procedure D and analysed using HPLC method P.

8.2.9 General procedure I: Manual conventional acrylic anhydride N-capping

On-resin peptide (ProTide, resin loading 0.68 mmol/g; 25 mg) was loaded in a 1.5 mL centrifuge tube. Acrylic anhydride (0.5 mmol; 70 mg; 29 eq.) in DMF (1 mL) was added to the resin. The mixture was shaken at room temperature for 15 minutes. The resin was washed with 1 mL DMF and fresh acrylic anhydride (0.5 mmol; 70 mg µL; 29 eq.) in DMF (1 mL) was added and the mixture was shaken for 15 min.

8.2.10 General procedure J: Manual microwave-assisted acrylic anhydride N-capping

On-resin peptide (ProTide, resin loading 0.68 mmol/g; 25 mg) was loaded in a 10 mL microwave vial. Acrylic anhydride (0.5 mmol; 70 mg; 29 eq.) in DMF (1 mL) was added to the resin. The mixture was heated to 90°C for 1.5 min using a Discover® SP Microwave Synthesiser (CEM).

8.2.11 General procedure K: Automated microwave-assisted acrylic anhydride N-capping

A stock solution of acrylic anhydride in DMF (1%) was added to amino acid position 27 on a Liberty Blue microwave-assisted peptide synthesiser (CEM). For the acrylamide capping step, the standard amino acid coupling cycle was used, using bottle 27 for reagent, activator and activator base.

8.3 General analytical procedures: HPLC, LCMS and NMR analysis

High Performance Liquid Chromatography

Analytical and semi-preparative HPLC employed an Agilent 1200 Series HPLC comprising a diode-array detector (215/280 nm) and G1364C fraction collector (semi-prep only). Specific details are provided below:

8.3.1 General procedure L: Analytical HPLC I

Analytical purity of the peptides was determined by high performance liquid chromatography (HPLC) with an XBridge® Peptide BEH C18 130 Å 5 µm column (4.6 x 150 mm), measuring UV absorbance at 215 nm, 254 nm and 280 nm. Samples were prepared by dissolving the peptide in water and acetonitrile (MeCN) with, if necessary, a drop of formic acid (FA). The HPLC was set up with a two-component system of H_2O/0.1% FA, and MeOH/0.1% FA for a run time of 25 min at a flow rate of 1 mL/min. The elution method was set up as follows: 0-5 min 90% H_2O/0.1% FA, from the fifth minute a gradient was applied to 100% MeOH/0.1%FA at 20 min, after which the system could equilibrate back to 90% H_2O/0.1%FA at 25 min.

8.3.2 General procedure M: Analytical HPLC II

Peptides were solubilised in MeCN and water and separated using a Phenomenex C18 analytical HPLC column (3.6 µm particle size, 4.6 × 150 mm column) with a binary eluent system comprising MeOH / H_2O (18 min gradient: 5-95% with 0.1% formic acid) as mobile phase. Operating pressures were in the range of 2000-3000 PSI.

8.3.3 General procedure N: Semi-preparative HPLC

Semi-preparative HPLC purification of peptides. Crude samples (25 mg/mL; 40 µL injection) were separated using a XBridge® Peptide BEH C18 Prep 130 Å 5µm column (10 x 150 mm) (Waters) with a binary eluent system comprising MeOH and H_2O (with 0.1% *v/v* formic acid) as mobile phase. Operating pressures were around 2000 PSI. Isolated pure peptides were concentrated *in vacuo* to remove organic volatiles and the aqueous solutions were then flash frozen (liquid N_2) and lyophilised.

Mass spectrometry

8.3.4 General procedure O: Low-resolution LCMS analysis

Samples were analysed using a Waters TQD mass spectrometer, using an XBridge (Waters, Manchester, UK) C18 analytical column (5 µm particle size, 4.6 × 150 mm) with a binary eluent system comprising MeOH / H_2O (18 min gradient: 5-95% with 0.1% formic acid) as mobile phase. Operating pressures were in the range of 2000-3000 PSI. Electrospray ionisation mass spectrometry was conducted in positive ion mode (*m/z* range: 600 – 1900) using a cone voltage of 50 V, desolvation temperature of 400°C and source temperature of 100 °C.

8.3.5 General procedure P: High-resolution (accurate) mass spectrometry

Samples were analysed using an Agilent 1260 Infinity II LC system with Agilent 6530 Accurate-Mass QToF spectrometer, using an Agilent ZORBAX Eclipse Plus C18 Rapid Resolution HD analytical column (1.8 µm particle size, 2.1 × 50 mm) with a binary eluent system comprising MeOH / H_2O (12 min gradient: 1-99% with 0.1% formic acid) as mobile phase. Operating pressures were in the range of 2000-3000 PSI. Electrospray ionisation mass spectrometry was conducted in positive ion mode (*m/z* range: 50 – 3200) using a fragmentor voltage of 150 V, gas temperature of 325 °C (flow 10 L/min) and sheath gas temperature of 400 °C (flow 11 L/min). Reference ions were Purine (121.0509) and Hexakis(1H,1H,3H-tetrafluoropropoxy)phosphazine (922.0098) (API-TOF Reference Mass Solution Kit, Agilent). Exact mass measurements of the products were based on the protonated molecules $[M+H]^+$ and were detected as sodiated adducts $[M+Na]^+$.

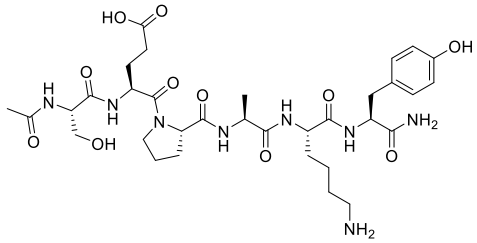
Nuclear Magnetic Resonance

8.3.6 General procedure Q: ^{19}F NMR

All NMR spectra were recorded using a Bruker 600 MHz Ascend NMR at 298 K. Peptide samples were prepared at 5 mg/mL in phosphate buffer [250 mM K_2HPO_4, 250 mM KH_2PO_4, 100 mM NaCl, 0.5 mM EDTA, 1 mM DDT; pH 6.8]. Peptides were analysed by ^{19}F NMR with protein decoupling. Residual TFA (-76.55 ppm) was used as standard for ^{19}F NMR. A total number of 512 scans was recorded for each sample.

8.4 Synthesis of peptides

Peptide 44 - Ac-SEPAKY-CONH$_2$

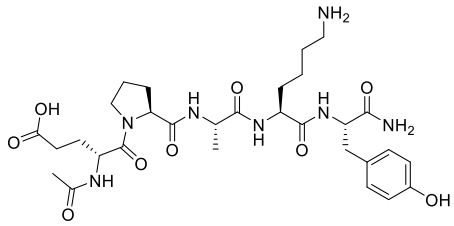


HPLC RT: 1.443 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{33}H_{50}N_8O_{11}$: 734.81; found. 735.3672

Peptide **44** was synthesised following standard automated SPPS procedure C. Upon completion of the synthesis they were acetyl-capped following procedure F, prior to peptide cleavage following general procedure D. Peptides were individually analysed following general procedure P.

Peptide 45 - Ac-EPAKY-CONH$_2$



HPLC RT: 0.91 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{30}H_{45}N_7O_9$: 647.73; found. 648.3352.

Peptide **45** was synthesised following standard automated SPPS procedure C. Upon completion of the synthesis they were acetyl-capped following procedure F, prior to peptide cleavage following general procedure D. Peptides were individually analysed following general procedure P.

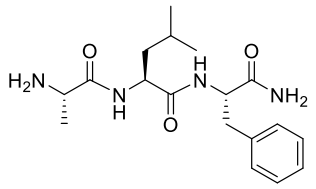
Peptide 46 - NH_2-KPSKIPVLLRKGGRQIKIWFQNRRMKWKK-$CONH_2$

HPLC RT: 10.44 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{168}H_{284}N_{54}O_{33}S$: 3620.53; found: N/D.

Peptide **46** was synthesised following SPPS general procedure C employing single or double amino acid couplings. Upon completion of the synthesis the peptide was cleaved from the resin following general procedure D. The crude double-coupled peptide 46 was purified using general semi-preparative HPLC procedure K. All peptides were analysed using general procedure O.

Peptide 54 - NH_2-ALF-$CONH_2$

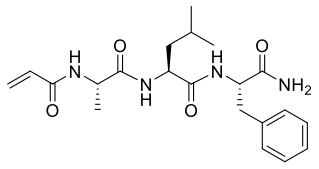


HPLC RT: 3.89 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{18}H_{28}N_4O_3$: 348.45; found 349.2245.

Peptide **54** was synthesised following standard automated SPPS procedure C. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide 55 - Acr-ALF-$CONH_2$



HPLC RT: 5.15 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{21}H_{30}N_4O_4$: 402.50; found 403.2348.

Peptide **55** was prepared according to the following procedures:

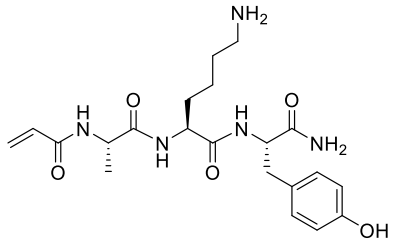
Procedure i: Peptide **55** was synthesised following standard SPPS procedure C. Acrylamide capping was performed following general procedure G. The peptide was cleaved from the resin following procedure D. Analysis was performed following procedure P.

Procedure ii: Peptide **55** was synthesised following standard SPPS procedure C. Acrylamide capping was performed following general procedure H. The peptide was cleaved from the resin following procedure D. Analysis was performed following procedure P.

Procedure iii: Peptide **55** was synthesised following standard SPPS procedure C. Acrylamide capping was performed following general procedure I. The peptide was cleaved from the resin following procedure D. Analysis was performed following procedure P.

Procedure iv: Peptide **55** was synthesised following standard SPPS procedure C. Acrylamide capping was performed following general procedure J. The peptide was cleaved from the resin following procedure D. Analysis was performed following procedure P.

Peptide 60 – Acr-AKY-$CONH_2$



HPLC RT: 7.212 min

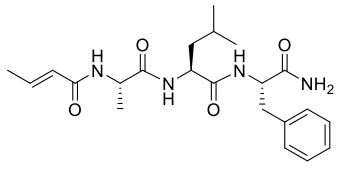
LC-MS (ESI+), $[M+H]^+$ calc. $C_{21}H_{31}N_5O_5$: 433.51; found N/D.

Peptide **60** was prepared according to the following procedures:

Procedure i: Peptide **60** was synthesised following standard SPPS procedure C. Acrylamide capping was performed following general procedure G. The peptide was cleaved from the resin following procedure D. Analysis was performed following procedure P.

Procedure ii: Peptide **60** was synthesised following standard SPPS procedure C. Acrylamide capping was performed following general procedure H. The peptide was cleaved from the resin following procedure D. Analysis was performed following procedure P.

Peptide 133 - MeAcr-ALF-$CONH_2$



HPLC RT: 5.21 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{22}H_{32}N_4O_4$: 416.52; found: 417.25309

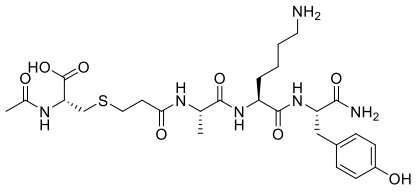
Peptide **133** was prepared according to the following procedures:

Procedure i: On-resin peptide (ProTide, resin loading 0.68 mmol/g; Peptide 52, NH_2-ALF-●, 25 mg; 0.017 mmol) was loaded in a 1.5 mL centrifuge tube. Crotonic anhydride (0.5 mmol; 74.12 µL; 29 eq.) in DMF (1 mL) was added to the resin. The mixture was shaken at

room temperature for 15 minutes. The resin was washed with 1 mL DMF and fresh crotonic anhydride (0.5 mmol; 74.12 µL; 29 eq.) in DMF (1 mL) was added and the mixture was shaken for 15 min.

Procedure ii: On-resin peptide (ProTide, resin loading 0.68 mmol/g; Peptide 52, NH_2-ALF-●, 25 mg; 0.017 mmol) was loaded in a 10 mL microwave vial. Crotonic anhydride (0.5 mmol; 74.12 µL; 29 eq.) in DMF (1 mL) was added to the resin. The mixture was heated to 90°C for 1.5 min using a Discover® SP Microwave Synthesiser (CEM).

Peptide 61 – NAC-Acr-AKY-CONH$_2$

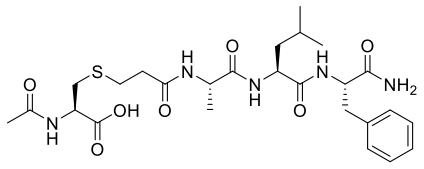


HPLC RT: 8.346 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{26}H_{40}N_6O_8S$: 596.70; found N/D.

Peptide **60** (10 mg; 0.023 mmol) was added to a 1.5 mL centrifuge tube. N-acetyl-L-cysteine (177 mg; 0.92 mmol; 40 eq.) and DIPEA (60 µL; 0.92 mmol; 40 eq.) in DMF (1 mL) was added to the resin. Samples (50 µL) were collected every 30 minutes for HPLC analysis following procedure M.

Peptide 62 – NAC-Acr-ALF-CONH$_2$



HPLC RT: 4.789 min

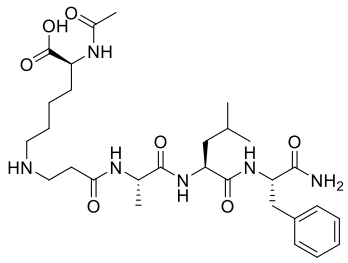
LC-MS (ESI+), $[M+H]^+$ calc. $C_{26}H_{39}N_5O_7S$: 565.69; found 566.2653.

Peptide **62** was prepared according to the following procedures:

Procedure i: On-resin peptide **55** (60 mg; 0.0336 mmol) was added to a 1.5 mL centrifuge tube. N-acetyl-L-cysteine (219.33 mg; 1.344 mmol; 40 eq.) in DMF (1 mL) was added to the resin. Samples (~10 mg) were collected after 30, 60, 120, 180, 240, and 1440 min. The peptides were cleaved from the resin following procedure D and analysed following procedure P.

Procedure ii: Peptide **55** (1 mg, 0.0025 mmol), NAC (40 eq., 16.2 mg, 0.1 mmol), and DMPA (20 eq., 12.8 mg, 0.05 mmol) were added to a glass reaction vial. Solids were dissolved in degassed DMF (1 mL). The reaction mixture was then exposed to 366 nm UV light for 1h.

Peptide 68 - NAL-Acr-ALF-CONH$_2$

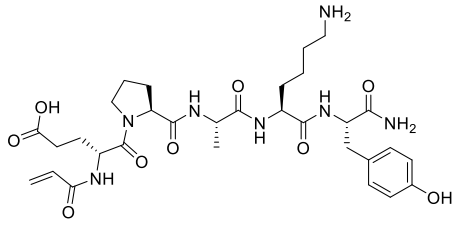


HPLC RT: 4.32 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{29}H_{46}N_6O_7$: 590.72; found 591.3499.

Peptide **55** (1 mg) was added to a 1.5 mL centrifuge tube. N-acetyl-L-lysine (18.68 mg; 0.099 mmol; 40 eq.) and DIPEA (17.3 µL; 0.099 mmol; 40 eq.) in DMF (1 mL) was added to the peptide. The reaction was heated to 90°C for 15 minutes using a Discover® SP Microwave Synthesiser (CEM). The peptide was analysed by HPLC following procedure P.

Peptide **75 - Acr-EPAKY-CONH$_2$**



HPLC RT: 1.362 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{31}H_{45}N_7O_9$: 659.74; found. 660.3342

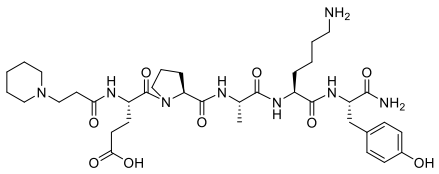
Peptide **75** was prepared according to the following procedures:

Procedure i: Peptide **75** was synthesised following standard SPPS procedure C. Acrylamide capping was performed following general procedure G. The peptide was cleaved from the resin following procedure D. Analysis was performed following procedure P.

Procedure ii: Peptide **75** was synthesised following standard SPPS procedure C. Acrylamide capping was performed following general procedure H. The peptide was cleaved from the resin following procedure D. Analysis was performed following procedure P.

Procedure iii: Peptide **75** was synthesised following standard SPPS procedure C. Acrylamide capping was performed following general procedure K. The peptide was cleaved from the resin following procedure D. Analysis was performed following procedure P.

Peptide 74 - Pip-Acr-EPAKY-CONH$_2$

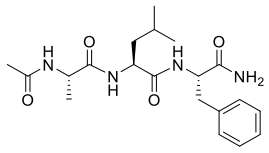


HPLC RT: 1.362 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{36}H_{56}N_8O_9$: 744.89; found. 745.4243

On-resin peptide **75** (25 mg) was added to a microwave vial containing 1 mL of 20% piperidine in DMF. The reaction was heated to 85°C for 90 seconds using a Discover® SP Microwave Synthesiser (CEM). The peptide was cleaved from the resin following procedure D and analysed by HPLC following procedure P.

Peptide 83 - Ac-ALF-CONH$_2$

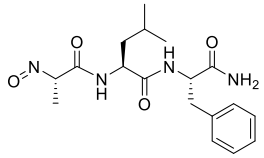


HPLC RT: 4.807 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{20}H_{30}N_4O_4$: 390.23; found 391.2343

Peptide **83** was synthesised following standard automated SPPS procedure C. Upon completion of the synthesis they were acetyl-capped following procedure F, prior to peptide cleavage following general procedure D. Peptides were individually analysed following general procedure P.

Peptide 88 - O=N-ALF-CONH$_2$

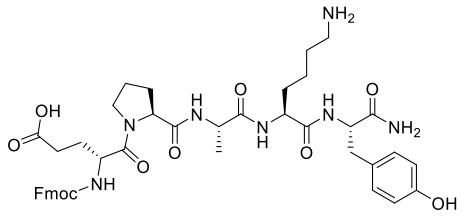


HPLC RT: 5.333 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{18}H_{26}N_4O_4$: 362.20; found 363.2030.

On resin peptide **52** (25 mg) was added to a 1.5 mL centrifuge tube. mCPBA (20 mg) in DCM (5 mL) was added to the resin. The mixture was shaken at room temperature for 2 hours. The resin was washed using 2x 1 mL DMF and 1x 1 mL diethyl ether. The peptide was cleaved from the resin following procedure D and analysed following procedure P.

Peptide 99 - Fmoc-EPAKY-CONH$_2$

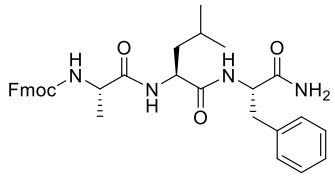


HPLC RT: 5.393 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{43}H_{53}N_7O_{10}$: 827.94; found. 828.3927

Peptide **99** was prepared following general procedure C, <u>without</u> final N-terminal deprotection. The peptide was used for on-resin Fmoc-deprotection experiments. The peptide was cleaved from the resin following procedure D and analysed following procedure P.

Peptide 100 - Fmoc-ALF-CONH$_2$

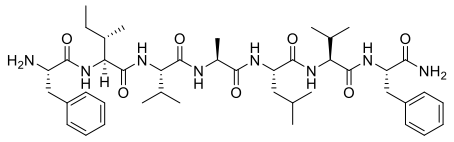


HPLC RT: 7.959 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{33}H_{38}N_4O_5$: 570.69; found. 571.2927

Peptide **100** was prepared following general procedure C, <u>without</u> final N-terminal deprotection. The peptide was used for on-resin Fmoc-deprotection experiments. The peptide was cleaved from the resin following procedure D and analysed following procedure P.

Peptide 102 - NH_2-FIVALVF-$CONH_2$

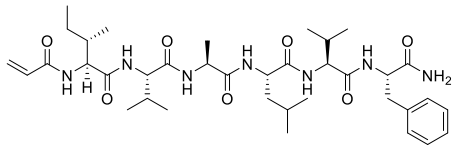


HPLC RT: 5.580 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{43}H_{66}N_8O_7$: 807.05; found. 807.5127

Peptide **102** was synthesised following standard automated SPPS procedure C. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide **103 - Acr-IVALVF-$CONH_2$**

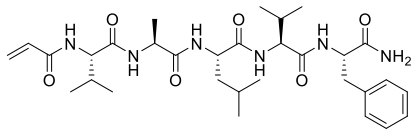


HPLC RT: 7.227 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{37}H_{59}N_7O_7$: 713.92; found. 714.4549

Peptide **103** was synthesised following standard automated SPPS procedure C. A final N-terminal acrylamide step was included following general procedure K. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide 104 - Acr-VALVF-CONH$_2$

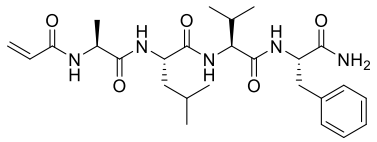


HPLC RT: 6.320 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{31}H_{48}N_6O_6$: 600.76; found. 601.3708

Peptide **104** was synthesised following standard automated SPPS procedure C. A final N-terminal acrylamide step was included following general procedure K. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide 105 - Acr-ALVF-CONH$_2$

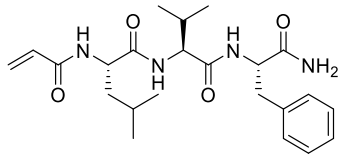


HPLC RT: 5.600 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{26}H_{39}N_5O_5$: 501.63; found. 502.3024

Peptide **105** was synthesised following standard automated SPPS procedure C. A final N-terminal acrylamide step was included following general procedure K. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide 106 - Acr-LVF-CONH$_2$

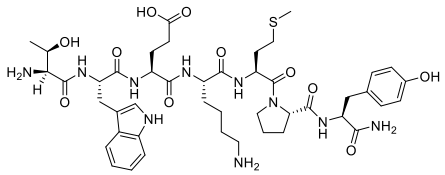


HPLC RT: 5.792 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{23}H_{34}N_4O_4$: 430.55; found. 431.2653

Peptide **106** was synthesised following standard automated SPPS procedure C. A final N-terminal acrylamide step was included following general procedure K. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide 107 - NH$_2$-TWEKMPY-CONH$_2$

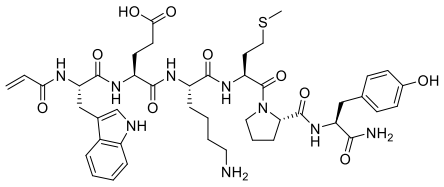


HPLC RT: 0.907 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{45}H_{64}N_{10}O_{11}S$: 953.13; found. 953.4623

Peptide **107** was synthesised following standard automated SPPS procedure C. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide 108 - Acr-WEKMPY-CONH$_2$

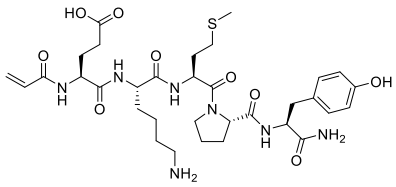


HPLC RT: 3.467 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{44}H_{59}N_9O_{10}S$: 906.07; found. 906.41808

Peptide **108** was synthesised following standard automated SPPS procedure C. A final N-terminal acrylamide step was included following general procedure K. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide 109 - Acr-EKMPY-CONH$_2$

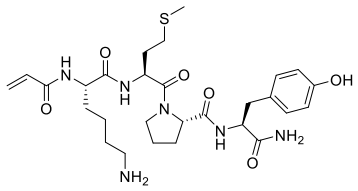


HPLC RT: 0.480 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{33}H_{49}N_7O_9S$: 719.86; found. 720.33843

Peptide **109** was synthesised following standard automated SPPS procedure C. A final N-terminal acrylamide step was included following general procedure K. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide 110 - Acr-KMPY-CONH$_2$

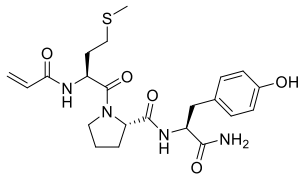


HPLC RT: 3.660 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{28}H_{42}N_6O_6S$: 590.74; found. 591.29684

Peptide **110** was synthesised following standard automated SPPS procedure C. A final N-terminal acrylamide step was included following general procedure K. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide **111 - Acr-MPY-CONH$_2$**

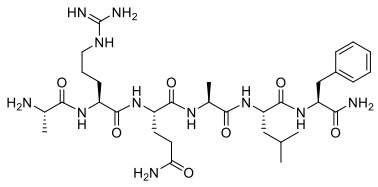


HPLC RT: 1.438 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{22}H_{30}N_4O_5S$: 462.57; found. 463.201

Peptide **111** was synthesised following standard automated SPPS procedure C. A final N-terminal acrylamide step was included following general procedure K. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide 127 – NH_2-ARQALF-$CONH_2$



HPLC RT: 5.079 min

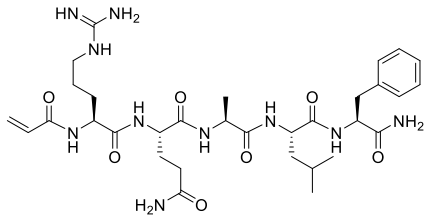
LC-MS (ESI+), $[M+H]^+$ calc. $C_{32}H_{53}N_{11}O_7$: 703.85; found. 704.4204

Peptide **127** was prepared according to the following procedures:

Procedure i: Peptide **127** was synthesised following standard automated SPPS procedure C. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Procedure ii: Peptide **127** was synthesised following standard automated SPPS procedure C, where piperidine was replaced by DABCO (5% in DMF) for Fmoc-deprotections. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide 128 – Acr-RQALF-$CONH_2$

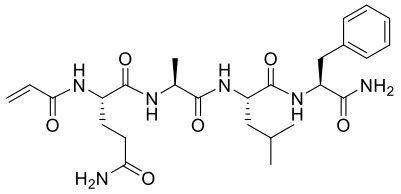


HPLC RT: 4.280 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{32}H_{50}N_{10}O_7$: 686.82; found: 687.40071.

Peptide **128** was synthesised following standard automated SPPS procedure C. A final N-terminal acrylamide step was included following general procedure K. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide 130 - Acr-QALF-CONH$_2$

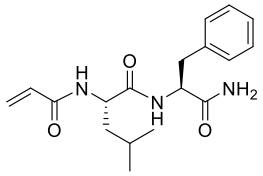


HPLC RT: 4.584 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{26}H_{38}N_6O_6$: 530.63; found. 531.2935

Peptide **130** was synthesised following standard automated SPPS procedure C. A final N-terminal acrylamide step was included following general procedure K. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide 131 – Acr-LF-CONH$_2$

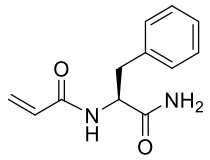


HPLC RT: 4.529 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{18}H_{25}N_3O_3$: 331.42; found. N/D

Peptide **131** was synthesised following standard automated SPPS procedure C. A final N-terminal acrylamide step was included following general procedure K. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide 132 – Acr-F-$CONH_2$



HPLC RT: 0.975 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{12}H_{14}N_2O_2$: 218.26; found. N/D

Peptide **132** was synthesised following standard automated SPPS procedure C. A final N-terminal acrylamide step was included following general procedure K. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

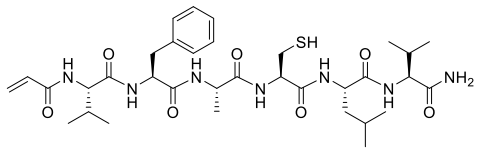
Peptide 136 – NH_2-HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS-$CONH_2$

HPLC RT: 5.46 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{184}H_{282}N_{50}O_{60}S$: 4184.03; found. N/D

Peptide **136** was synthesised following standard automated SPPS procedure C where piperidine was replaced by DABCO (5% in DMF) for Fmoc-deprotections. After each amino acid coupling step, an acrylamide-tagging step was included following general procedure K, prior to Fmoc-deprotection. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide 154 - Acr-VFACLV-CONH$_2$

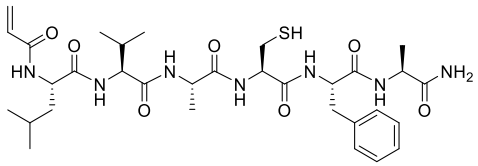


HPLC RT: 6.464 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{34}H_{53}N_7O_7S$: 703.90; found 704.3739

Peptide **154** was synthesised following standard SPPS procedure C. Cysteine was introduced bearing a Monomethoxytrityl (Mmt) protecting group, allowing for orthogonal deprotection. N-terminal acrylamide was introduced following standard procedure K. The peptide was cleaved from the resin following general procedure D. Analysis was performed according to general procedure P.

Peptide 157 - Acr-LVACFA-CONH$_2$

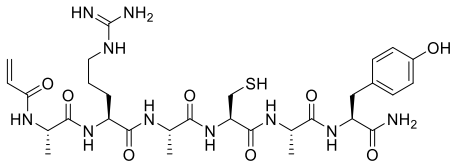


HPLC RT: 6.013 min

LC-MS (ESI+), [M+H]+ calc. $C_{32}H_{49}N_7O_7S$: 675.85; found 676.35.

Peptide **157** was synthesised following standard SPPS procedure C. Cysteine was introduced bearing a Monomethoxytrityl (Mmt) protecting group, allowing for orthogonal deprotection. N-terminal acrylamide was introduced following standard procedure K. The peptide was cleaved from the resin following general procedure D. Analysis was performed according to general procedure P.

Peptide 165 - Acr-ARACAY-CONH$_2$

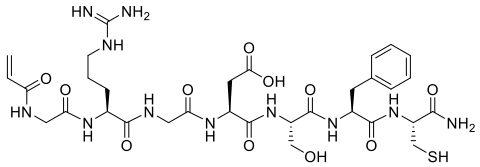


HPLC RT: 1.76 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{30}H_{46}N_{10}O_8S$: 706.82; found 707.33

Peptide **165** was synthesised following standard SPPS procedure C. Cysteine was introduced bearing a Monomethoxytrityl (Mmt) protecting group, allowing for orthogonal deprotection. N-terminal acrylamide was introduced following standard procedure K. The peptide was cleaved from the resin following general procedure D. Analysis was performed according to general procedure P.

Peptide 167 - Acr-GRGDSFC-CONH$_2$

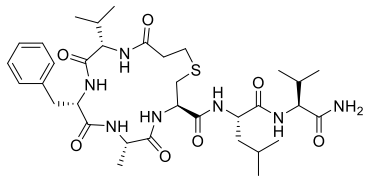


HPLC RT: 3.287 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{32}H_{47}N_{11}O_{11}S$: 793.85; found 794.33

Peptide **167** was synthesised following standard SPPS procedure C. Cysteine was introduced bearing a Monomethoxytrityl (Mmt) protecting group, allowing for orthogonal deprotection. N-terminal acrylamide was introduced following standard procedure K. The peptide was cleaved from the resin following general procedure D. Analysis was performed according to general procedure P.

Peptide 155 - Acr-VFACLV-CONH$_2$ (cyclic)



HPLC RT: 5.97 min

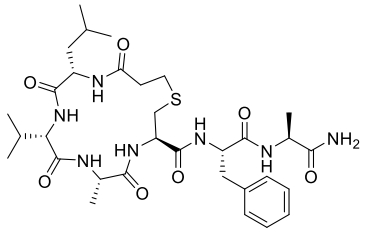
LC-MS (ESI+), $[M+H]^+$ calc. $C_{34}H_{53}N_7O_7S$: 703.90; found 704.3805

Peptide **155** was prepared according to the following procedures:

Procedure i: Peptide **154** (1 mg) was added to a 1.5 mL centrifuge tube. The peptide was treated with 25 mM DIPEA in DMF (1 mL) for 4 hours. The peptide was precipitated in cold diethyl ether, centrifuged, dissolved in H_2O/MeOH (1:1) and analysed following general procedure P.

Procedure ii: Peptide **154** (1 mg) was added to a 10 mL microwave vial. The peptide was treated with 25 mM DIPEA in DMF (1 mL) at 90°C for 15 minutes. The peptide was precipitated in cold diethyl ether, centrifuged, dissolved in H_2O/MeOH (1:1) and analysed following general procedure P.

Peptide 159 - Acr-LVACFA-CONH$_2$ (cyclic)

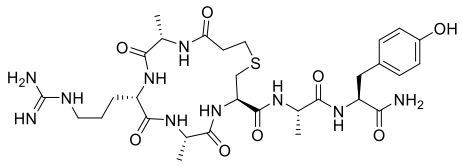


HPLC RT: 6.71 min

LC-MS (ESI+), [M+H]+ calc. $C_{32}H_{49}N_7O_7S$: 675.85; found 675.34.

From on-resin peptide **158** (100 mg; 0.019 mmol) cysteine was orthogonally deprotected prior to cyclisation. Cleavage cocktail (2 mL; comprising DCM, TIPS and TFA (94 : 5 : 1 *v/v*) was added to the resin and was shaken for 2 minutes at room temperature. The cleavage cocktail was then drained from the resin and the procedure was repeated another four times. After completion of 5 cleavage cycles the resin was washed with DCM (10 mL) and dried under vacuum. The cyclisation reaction was attempted by shaking the on-resin peptides in a solution of DIPEA (165 µL; 0.95 mmol; 50 eq.) in DMF (5 mL) over time at 37°C. Resin samples (~10 mg) were collected after 0, 30, 60, 120, and 240 minutes. The peptide was cleaved from the resin following general procedure D and analysed using procedure P.

Peptide **166 - Acr-ARACAY-CONH$_2$ (cyclic)**

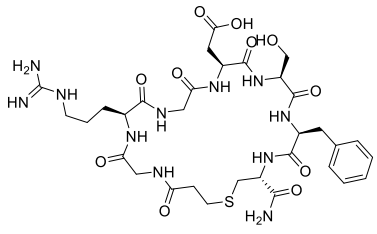


HPLC RT: 3.56 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{30}H_{46}N_{10}O_8S$: 706.82; found 706.32

From on-resin peptide **165** (180 mg; 0.1 mmol) cysteine was orthogonally deprotected prior to cyclisation. Cleavage cocktail (2 mL; comprising DCM, TIPS and TFA (94 : 5 : 1 *v/v*) was added to the resin and was shaken for 2 minutes at room temperature. The cleavage cocktail was then drained from the resin and the procedure was repeated another four times. After completion of 5 cleavage cycles the resin was washed with DCM (10 mL) and dried under vacuum. The cyclisation reaction was attempted by shaking the on-resin peptides in a solution of 25 mM DIPEA in DMF (5 mL) over time at 37°C. Resin samples (~10 mg) were collected after 24, 48, 72 and 96 hours. The peptide was cleaved from the resin following general procedure D and analysed using procedure P.

Peptide 167 - Acr-GRGDSFC-CONH$_2$ (cyclic)

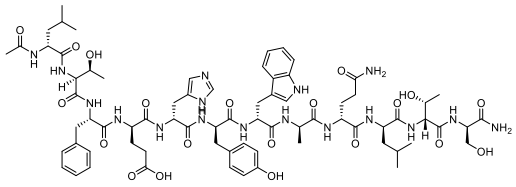


HPLC RT: 3.90 min

LC-MS (ESI+), $[M+H]^+$ calc. $C_{32}H_{47}N_{11}O_{11}S$: 793.85; found 794.323

From on-resin peptide **166** (180 mg; 0.1 mmol) cysteine was orthogonally deprotected prior to cyclisation. Cleavage cocktail (2 mL; comprising DCM, TIPS and TFA (94 : 5 : 1 *v/v*) was added to the resin and was shaken for 2 minutes at room temperature. The cleavage cocktail was then drained from the resin and the procedure was repeated another four times. After completion of 5 cleavage cycles the resin was washed with DCM (10 mL) and dried under vacuum. The cyclisation reaction was attempted by shaking the on-resin peptides in a solution of 25 mM DIPEA in DMF (5 mL) over time at 37°C. Resin samples (~10 mg) were collected after 24, 48, 72 and 96 hours. The peptide was cleaved from the resin following general procedure D and analysed using procedure P.

Peptide 214 - NH_2-LTFEHYWAQLTS-$CONH_2$

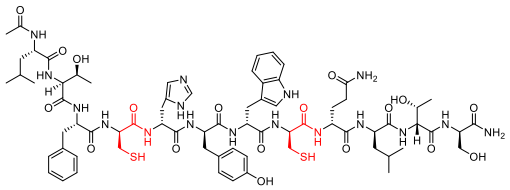


HPLC RT: 16.765 min. Crude purity: 68%.

HPLC-MS (ESI +H) calc. $C_{73}H_{101}N_{17}O_{20}$: 1535.74088; found 1536.7431.

Peptide **214** was prepared using standard automated Fmoc-SPPS procedure C, subsequently followed by peptide cleavage and lyophilisation following general procedure D. Peptide **214** was purified by semi-preparative HPLC method N. The purified peptide was analysed using procedure P.

Peptide 215 - Ac-LTF(L-Cys)HYW(L-Cys)QLTS-$CONH_2$

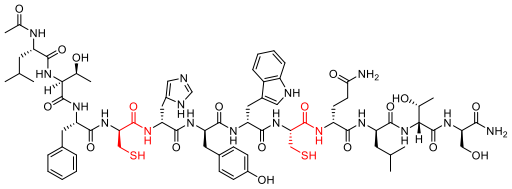


HPLC RT: 16.935. Crude purity: 57%.

HPLC-MS (ESI +H) calc. $C_{71}H_{99}N_{17}O_{18}S_2$: 1541.67954.

Peptide **215** was prepared using standard automated Fmoc-SPPS procedure C, subsequently followed by peptide cleavage and lyophilisation following general procedure D. Peptide **215** was purified by semi-preparative HPLC method N. The purified peptide was analysed using procedure P. The crude disulfide peptide was used in stapling experiments without further chromatographic purification.

Peptide 216 - Ac-LTF(L-Cys)HYW(D-Cys)QLTS-CONH$_2$

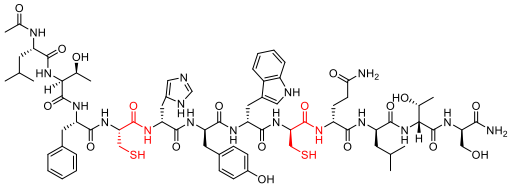


HPLC RT: 11.608 min. Crude purity: 70%.

HPLC-MS (ESI +H) calc. $C_{71}H_{99}N_{17}O_{18}S_2$: 1541.68; found 1542.69.

Peptide **216** was prepared using standard automated Fmoc-SPPS procedure C, subsequently followed by peptide cleavage and lyophilisation following general procedure D. The crude disulfide peptide was analysed following procedure P and used in stapling experiments without further chromatographic purification.

Peptide 217 - Ac-LTF(D-Cys)HYW(L-Cys)QLTS-CONH$_2$

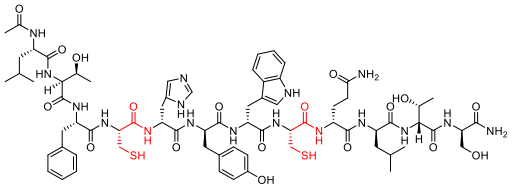


HPLC RT: 11.624 min. Crude purity: 68%.

HPLC-MS (ESI +H) calc. $C_{71}H_{99}N_{17}O_{18}S_2$: 1541.68; found1542.69.

Peptide **217** was prepared using standard automated Fmoc-SPPS procedure C, subsequently followed by peptide cleavage and lyophilisation following general procedure D. The crude disulfide peptide was analysed following procedure P and used in stapling experiments without further chromatographic purification.

Peptide 218 - Ac-LTF(D-Cys)HYW(D-Cys)QLTS-CONH$_2$

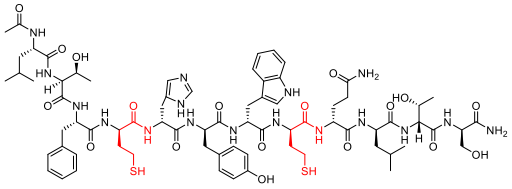


HPLC RT: 11.520 min. Crude purity: 76%.

HPLC-MS (ESI +H) calc. $C_{71}H_{99}N_{17}O_{18}S_2$: 1541.68; found 1542.66.

Peptide **218** was prepared using standard automated Fmoc-SPPS procedure C, subsequently followed by peptide cleavage and lyophilisation following general procedure D. The crude disulfide peptide was analysed following procedure P and used in stapling experiments without further chromatographic purification.

Peptide 219 - Ac-LTF(hCys)HYW(hCys)QLTS-CONH$_2$

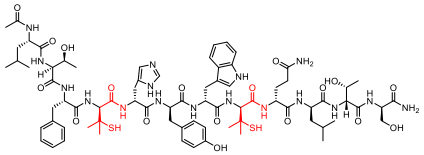


HPLC RT: 20.803 min. Crude purity: 52%.

HPLC-MS (ESI +H) calc. $C_{73}H_{103}N_{17}O_{18}S_2$: 1570.85; found 1570.65.

Peptide **219** was prepared using standard automated Fmoc-SPPS procedure C, subsequently followed by peptide cleavage and lyophilisation following general procedure D. The crude disulfide peptide was analysed following procedure P and used in stapling experiments without further chromatographic purification.

Peptide 220 - Ac-LTF(Pen)HYW(Pen)QLTS-CONH$_2$

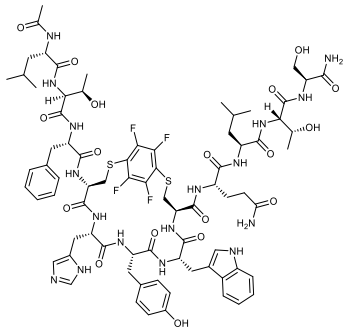


HPLC RT: 19.752 min. Crude purity: 94%.

HPLC-MS (ESI +H) calc. $C_{75}H_{107}N_{17}O_{18}S_2$: 1598.90; found 1599.63.

Peptide **220** was prepared using standard automated Fmoc-SPPS procedure C, subsequently followed by peptide cleavage and lyophilisation following general procedure D. The crude disulfide peptide was analysed following procedure P and used in stapling experiments without further chromatographic purification.

Peptide 221 - Ac-LTF(L-Cys)HYW(L-Cys)QLTS-CONH$_2$ (HFB stapled)

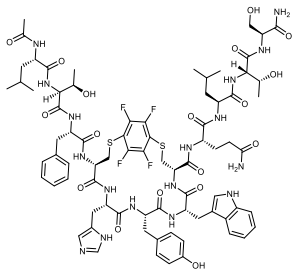


HPLC RT: 10.584 min. Crude purity: 46%.

HPLC-MS (ESI +H) calc. $C_{77}H_{97}F_4N_{17}O_{18}S_2$: 1688.84; found 1688.6612.

To a centrifuge tube containing solid crude peptides **215** (approx. 50 mg) was added DIPEA stock solution (25 mM in DMF, 1.0 mL), followed by hexafluorobenzene (20 eq., 0.66 mmol). The resulting mixture was shaken at room temperature for 4.5 h. After precipitation in Et_2O, the resulting crude stapled peptide was redissolved in a mixture of water/MeCN (exact amounts depended on solubility). Crude samples were purified using semi-preparative HPLC method N.

Peptide 222 - Ac-LTF(D-Cys)HYW(L-Cys)QLTS-CONH$_2$ (HFB stapled)

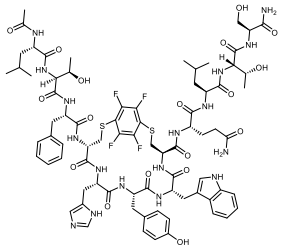


HPLC RT: 18.726 min. Crude purity: 71%.

HPLC-MS (ESI +H) calc. $C_{77}H_{97}F_4N_{17}O_{18}S_2$: 1688.84; found 1689.6642.

To a centrifuge tube containing solid crude peptides **216** (approx. 50 mg) was added DIPEA stock solution (25 mM in DMF, 1.0 mL), followed by hexafluorobenzene (20 eq., 0.66 mmol). The resulting mixture was shaken at room temperature for 4.5 h. After precipitation in Et_2O, the resulting crude stapled peptide was redissolved in a mixture of water/MeCN (exact amounts depended on solubility). Crude samples were purified using semi-preparative HPLC method N.

Peptide 223 - Ac-LTF(L-Cys)HYW(D-Cys)QLTS-CONH$_2$

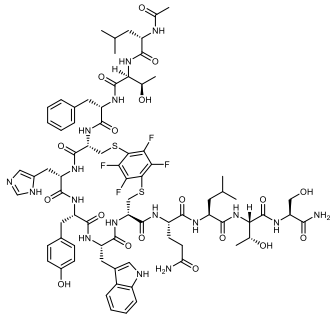


HPLC RT: 19.133 min. Crude purity: 60%.

HPLC-MS (ESI +H) calc. $C_{77}H_{97}F_4N_{17}O_{18}S_2$: 1688.84; found 1689.6646.

To a centrifuge tube containing solid crude peptides **217** (approx. 50 mg) was added DIPEA stock solution (25 mM in DMF, 1.0 mL), followed by hexafluorobenzene (20 eq., 0.66 mmol). The resulting mixture was shaken at room temperature for 4.5 h. After precipitation in Et_2O, the resulting crude stapled peptide was redissolved in a mixture of water/MeCN (exact amounts depended on solubility). Crude samples were purified using semi-preparative HPLC method N.

Peptide 224 - Ac-LTF(D-Cys)HYW(D-Cys)QLTS-CONH$_2$

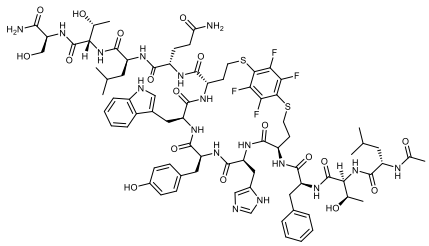


HPLC RT: 18.601 min. Crude purity: 53%.

HPLC-MS (ESI +H) calc. $C_{77}H_{97}F_4N_{17}O_{18}S_2$: 1688.84; found 1689.6636

To a centrifuge tube containing solid crude peptides **218** (approx. 50 mg) was added DIPEA stock solution (25 mM in DMF, 1.0 mL), followed by hexafluorobenzene (20 eq., 0.66 mmol). The resulting mixture was shaken at room temperature for 4.5 h. After precipitation in Et_2O, the resulting crude stapled peptide was redissolved in a mixture of water/MeCN (exact amounts depended on solubility). Crude samples were purified using semi-preparative HPLC method N.

Peptide 225 - Ac-LTF(hCys)HYW(hCys)QLTS-CONH$_2$

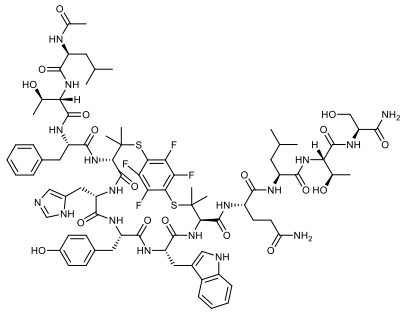


HPLC RT: 19.371 min. Crude purity: 67%.

HPLC-MS (ESI +H) calc. $C_{79}H_{101}F_4N_{17}O_{18}S_2$: 1716.89; found 1716.6957.

To a centrifuge tube containing solid crude peptides **219** (approx. 50 mg) was added DIPEA stock solution (25 mM in DMF, 1.0 mL), followed by hexafluorobenzene (20 eq., 0.66 mmol). The resulting mixture was shaken at room temperature for 4.5 h. After precipitation in Et_2O, the resulting crude stapled peptide was redissolved in a mixture of water/MeCN (exact amounts depended on solubility). Crude samples were purified using semi-preparative HPLC method N.

Peptide 226 - Ac-LTF(Pen)HYW(Pen)QLTS-CONH$_2$

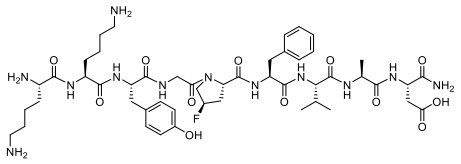


HPLC RT: 18.49 min. Crude purity: 75%.

HPLC-MS (ESI +H) calc. $C_{81}H_{105}F_4N_{17}O_{18}S_2$: 1744.95; found 1744.7251.

To a centrifuge tube containing solid crude peptides **220** (approx. 50 mg) was added DIPEA stock solution (25 mM in DMF, 1.0 mL), followed by hexafluorobenzene (20 eq., 0.66 mmol). The resulting mixture was shaken at room temperature for 4.5 h. After precipitation in Et_2O, the resulting crude stapled peptide was redissolved in a mixture of water/MeCN (exact amounts depended on solubility). Crude samples were purified using semi-preparative HPLC method N.

Peptide 249 - NH$_2$-KKYG(4R-Flp)FVAD-CONH$_2$

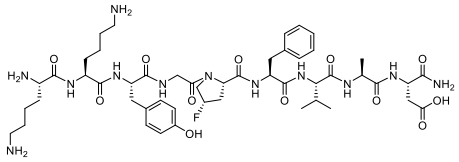


HPLC RT: 09.008 min.

HPLC-MS (ESI +H) calc. $C_{49}H_{73}FN_{12}O_{12}$: 1041.19; found 1041.18

Peptide **249** was synthesised following manual solid phase peptide synthesis procedure A, using general procedure E to monitor coupling efficiency. Peptides were purified using semi-preparative-HPLC method K and analysed using general procedure L and O.

Peptide 250 - NH_2-KKYG(4S-Flp)FVAD-$CONH_2$

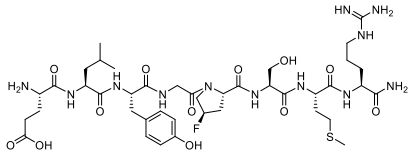


HPLC RT: 09.461 min.

HPLC-MS (ESI +H) calc. $C_{49}H_{73}FN_{12}O_{12}$: 1041.19; found 1041.24.

Peptide **250** was synthesised following manual solid phase peptide synthesis procedure A, using general procedure E to monitor coupling efficiency. Peptides were purified using semi-preparative-HPLC method K and analysed using general procedure L and O.

Peptide 251 - NH_2-ELYG(4R-Flp)SMR-$CONH_2$

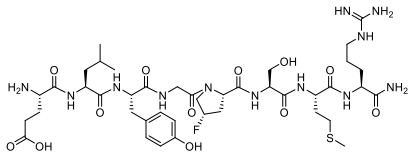


HPLC RT: 09.520 min.

HPLC-MS (ESI +H) calc. $C_{49}H_{73}FN_{12}O_{12}$: 969.10; found 969.15.

Peptide **251** was synthesised following manual solid phase peptide synthesis procedure A, using general procedure E to monitor coupling efficiency. Peptides were purified using semi-preparative-HPLC method K and analysed using general procedure L and O.

Peptide 252 - NH_2-ELYG(4S-Flp)SMR-$CONH_2$



HPLC RT: 09.112 min.

HPLC-MS (ESI +H) calc. $C_{49}H_{73}FN_{12}O_{12}$: 969.10; found 969.15.

Peptide **252** was synthesised following manual solid phase peptide synthesis procedure A, using general procedure E to monitor coupling efficiency. Peptides were purified using semi-preparative-HPLC method K and analysed using general procedure L and O.

Peptide 135 - NH_2-DPGPDEAPRMPEAAPPVAPAPAAPT-$CONH_2$

HPLC RT: 3.813 min.

HPLC-MS (ESI +H) calc. $C_{106}H_{165}N_{29}O_{34}S$: 2421.72; found 2422.1775.

Peptide **135** was prepared according to the following procedures:

Procedure i: Peptide **135** was prepared using standard automated Fmoc-SPPS general procedure C. Each proline and subsequent residue were double coupled. Upon completion of the synthesis the peptides were cleaved from the resin and lyophilisation following general procedure D. Peptides were analysed using HPLC-MS following general procedure P.

Procedure ii: Peptide **135** was synthesised following standard automated SPPS procedure C, where piperidine was replaced by DABCO (5% in DMF) for Fmoc-deprotections. After each amino acid coupling step, an acrylamide-tagging step was included following general procedure K, prior to Fmoc-deprotection. Upon completion of the synthesis peptides were cleaved following general procedure D. Peptides were analysed following general procedure P.

Peptide 253 - NH_2-D(4*S*-FPro)GPDEAPRMPEAAPPVAPAPAAPT-$CONH_2$

HPLC RT: 3.907 min.

HPLC-MS (ESI +H) calc. $C_{106}H_{164}FN_{29}O_{34}S$: 2439.71; found 2439.1770.

^{19}F NMR (600 MHz, phosphate buffer) δ -173.6, -173.7

Peptide **253** was prepared using standard automated Fmoc-SPPS general procedure C combined with manual amino acid couplings of the fluorinated and subsequent residue following general procedure A. Each proline and subsequent residue were double coupled. Upon completion of the synthesis the peptides were cleaved from the resin and lyophilisation following general procedure D. Peptides were analysed using HPLC-MS following general procedure P and by ^{19}F NMR following procedure Q.

Peptide 254 - NH_2-DPG(4*S*-FPro)DEAPRMPEAAPPVAPAPAAPT-$CONH_2$

HPLC RT: 3.881 min.

HPLC-MS (ESI +H) calc. $C_{106}H_{164}FN_{29}O_{34}S$: 2439.71; found 2439.1762.

^{19}F NMR (600 MHz, phosphate buffer) δ -173.1, -173.9

Peptide **254** was prepared using standard automated Fmoc-SPPS general procedure C combined with manual amino acid couplings of the fluorinated and subsequent residue following general procedure A. Each proline and subsequent residue were double coupled. Upon completion of the synthesis the peptides were cleaved from the resin and lyophilisation following general procedure D. Peptides were analysed using HPLC-MS following general procedure P and by ^{19}F NMR following procedure Q.

Peptide 255 - NH_2-DPGPDEA(4*S*-FPro)RMPEAAPPVAPAPAAPT-$CONH_2$

HPLC RT: N/D min.

HPLC-MS (ESI +H) calc. $C_{106}H_{164}FN_{29}O_{34}S$: 2439.71; found N/D.

^{19}F NMR (600 MHz, phosphate buffer) δ -173.0, -173.2

Peptide **255** was prepared using standard automated Fmoc-SPPS general procedure C combined with manual amino acid couplings of the fluorinated and subsequent residue following general procedure A. Each proline and subsequent residue were double coupled. Upon completion of the synthesis the peptides were cleaved from the resin and lyophilisation following general procedure D. Peptides were analysed using HPLC-MS following general procedure P and by ^{19}F NMR following procedure Q.

Peptide 256 - NH_2-DPGPDEAPRM(4*S*-FPro)EAAPPVAPAPAAPT-$CONH_2$

HPLC RT: 3.872 min.

HPLC-MS (ESI +H) calc. $C_{106}H_{164}FN_{29}O_{34}S$: 2439.71; found 2439.1742.

^{19}F NMR (600 MHz, phosphate buffer) δ -173.3, -173.4

Peptide **256** was prepared using standard automated Fmoc-SPPS general procedure C combined with manual amino acid couplings of the fluorinated and subsequent residue following general procedure A. Each proline and subsequent residue were double coupled. Upon completion of the synthesis the peptides were cleaved from the resin and lyophilisation following general procedure D. Peptides were analysed using HPLC-MS following general procedure P and by ^{19}F NMR following procedure Q.

Peptide 257 - NH_2-DPGPDEAPRMPEAA(4*S*-FPro)PVAPAPAAPT-$CONH_2$

HPLC RT: 4.077 min.

HPLC-MS (ESI +H) calc. $C_{106}H_{164}FN_{29}O_{34}S$: 2439.71; found 2440.1902.

^{19}F NMR (600 MHz, phosphate buffer) δ -171.8, -172.4

Peptide **257** was prepared using standard automated Fmoc-SPPS general procedure C combined with manual amino acid couplings of the fluorinated and subsequent residue following general procedure A. Each proline and subsequent residue were double coupled. Upon completion of the synthesis the peptides were cleaved from the resin and lyophilisation following general procedure D. Peptides were analysed using HPLC-MS following general procedure P and by ^{19}F NMR following procedure Q.

Peptide 258 - NH_2-DPGPDEAPRMPEAAP(4*S*-FPro)VAPAPAAPT-$CONH_2$

HPLC RT: 3.938 min.

HPLC-MS (ESI +H) calc. $C_{106}H_{164}FN_{29}O_{34}S$: 2439.71; found 2439.1879.

^{19}F NMR (600 MHz, phosphate buffer) δ -172.9, -174.1

Peptide **258** was prepared using standard automated Fmoc-SPPS general procedure C combined with manual amino acid couplings of the fluorinated and subsequent residue following general procedure A. Each proline and subsequent residue were double coupled. Upon completion of the synthesis the peptides were cleaved from the resin and lyophilisation following general procedure D. Peptides were analysed using HPLC-MS following general procedure P and by ^{19}F NMR following procedure Q.

Peptide 259 - NH_2-DPGPDEAPRMPEAAPPVA(4*S*-FPro)APAAPT-$CONH_2$

HPLC RT: 3.886 min.

HPLC-MS (ESI +H) calc. $C_{106}H_{164}FN_{29}O_{34}S$: 2439.71; found 2439.1763.

^{19}F NMR (600 MHz, phosphate buffer) δ -173.1, -173.8

Peptide **259** was prepared using standard automated Fmoc-SPPS general procedure C combined with manual amino acid couplings of the fluorinated and subsequent residue following general procedure A. Each proline and subsequent residue were double coupled. Upon completion of the synthesis the peptides were cleaved from the resin and lyophilisation following general procedure D. Peptides were analysed using HPLC-MS following general procedure P and by ^{19}F NMR following procedure Q.

Peptide 260 - NH_2-DPGPDEAPRMPEAAPPVAPA(4*S*-FPro)AAPT-$CONH_2$

HPLC RT: 3.805 min.

HPLC-MS (ESI +H) calc. $C_{106}H_{164}FN_{29}O_{34}S$: 2439.71; found 2439.1733.

^{19}F NMR (600 MHz, phosphate buffer) δ -173.4, -173.8

Peptide **260** was prepared using standard automated Fmoc-SPPS general procedure C combined with manual amino acid couplings of the fluorinated and subsequent residue following general procedure A. Each proline and subsequent residue were double coupled. Upon completion of the synthesis the peptides were cleaved from the resin and lyophilisation following general procedure D. Peptides were analysed using HPLC-MS following general procedure P and by ^{19}F NMR following procedure Q.

Peptide 261 - NH_2-DPGPDEAPRMPEAAPPVAPAPAA(4*S*-FPro)T-$CONH_2$

HPLC RT: 3.904 min.

HPLC-MS (ESI +H) calc. $C_{106}H_{164}FN_{29}O_{34}S$: 2439.71; found 2439.1762.

^{19}F NMR (600 MHz, phosphate buffer) δ -172.9, -173.5

Peptide **261** was prepared using standard automated Fmoc-SPPS general procedure C combined with manual amino acid couplings of the fluorinated and subsequent residue following general procedure A. Each proline and subsequent residue were double coupled. Upon completion of the synthesis the peptides were cleaved from the resin and lyophilisation following general procedure D. Peptides were analysed using HPLC-MS following general procedure P and by ^{19}F NMR following procedure Q.

Synthesis of polymer thiol-resin (126)

ProTide resin (0.63 mmol/g; 1.032 g) was deprotected using 20% piperidine in DMF (50 mL *v/v*) for 2x 30 min. Fmoc-Cys(Mmt)-OH (2.6 mmol; 1.6 g; 4 eq.), PyBOP (2.6 mmol; 1.35 g; 4 eq.), and DIPEA (5.2 mmol; 905 µL; 8 eq.) in DMF (20 mL) were added to the resin for 1 h at room temperature. The resin was deprotected and subsequently acetyl capped following general procedure F. The Mmt protecting group was removed using 5x 10 mL DCM/TFA/TIPS (94:1:5) for 2 min at room temperature.

8.5 Stability experiments

8.5.1 Acrylamide stability in solution

N-acrylamide capped peptide **55** was dissolved in a mixture of H_2O/MeOH (1:1; 1 mg/mL). The peptide was split between 4 separate HPLC vials (2.48 µmol; 1 mL in each) and one of the following conditions was applied:

A. No base
B. DIPEA (17.3 µL, 99.2 µmol, 40 eq.)
C. DABCO (11.13 mg, 99.2 µmol, 40 eq.)
D. Et_3N (13.8 µL; 99.2 µmol, 40 eq.)

LC-MS data was obtained after 0, 3, 6, and 24 hours following general procedure P.

8.5.2 Acrylamide stability under conventional reaction conditions

Acrylamide-capped peptide **72** (25 mg) was added to a 1.5 mL centrifuge tube. The peptide was treated with one of the following basic conditions:

A. 20% piperidine in DMF (1 mL)
B. 20% DIPEA in DMF (1 mL)
C. 20% Et_3N in DMF (1 mL)
D. 20% DABCO in DMF (1 mL)
E. 20% DBU in DMF (1 mL)

The reaction mixture was shaken for 2x 15 minutes at room temperature, after which the resin was washed with 1 mL DMF. The peptides were cleaved from the resin following procedure D and analysed following procedure P.

8.5.3 Acrylamide stability under microwave reaction conditions

Acrylamide-capped peptide **72** (25 mg) was added to a 10 mL microwave vial. The peptide was treated with one of the following basic conditions:

A. 20% piperidine in DMF (1 mL)
B. 20% DIPEA in DMF (1 mL)
C. 20% Et_3N in DMF (1 mL)
D. 20% DABCO in DMF (1 mL)
E. 20% DBU in DMF (1 mL)

The reaction mixture was heated to 90°C for 1.5 minutes using a Discover® SP Microwave Synthesiser (CEM)., after which the resin was washed with 1 mL DMF. The peptides were cleaved from the resin following procedure D and analysed following procedure P.

8.6 Fmoc-deprotection strategies

8.6.1 Fmoc-deprotection under conventional reaction conditions

Fmoc-protected peptide **99** (25 mg) was added to a 1.5 mL centrifuge tube. The peptide was treated with one of the following basic conditions:

A. 20% piperidine in DMF (1 mL)
B. 20% DIPEA in DMF (1 mL)
C. 20% Et_3N in DMF (1 mL)
D. 20% DABCO in DMF (1 mL)
E. 20% DBU in DMF (1 mL)

The reaction mixture was shaken for 2x 15 minutes at room temperature, after which the resin was washed with 1 mL DMF. The peptides were cleaved from the resin following procedure D and analysed following procedure P.

8.6.2 Fmoc-deprotection under microwave reaction conditions

Fmoc-protected peptide **99** (25 mg) was added to a 10 mL microwave vial. The peptide was treated with one of the following basic conditions:

A. 20% piperidine in DMF (1 mL)
B. 20% DIPEA in DMF (1 mL)
C. 20% Et_3N in DMF (1 mL)
D. 20% DABCO in DMF (1 mL)
E. 20% DBU in DMF (1 mL)

The reaction mixture was heated to 90°C for 1.5 minutes using a Discover® SP Microwave Synthesiser (CEM)., after which the resin was washed with 1 mL DMF. The peptides were cleaved from the resin following procedure D and analysed following procedure P.

8.6.3 Fmoc-deprotection using DABCO

Fmoc-protected peptide **100** (25 mg) was added to a 10 mL microwave vial. The peptide was treated with one of the following basic conditions:

A. 20% DABCO in DMF(1 mL)
B. 15% DABCO in DMF (1 mL)
C. 10% DABCO in DMF (1 mL)
D. 5% DABCO in DMF (1 mL)

The reaction mixture was heated to 90°C for 1.5 minutes using a Discover® SP Microwave Synthesiser (CEM)., after which the resin was washed with 1 mL DMF. The peptides were cleaved from the resin following procedure D and analysed following procedure P.

8.7 Thioether elimination

8.7.1 Thioether elimination using acetic anhydride

On-resin peptide **75** (ProTide resin, loading 0.55 mmol/g; ~25 mg) containing an N-terminal piperidine-acrylamide adduct was added to a 10 mL microwave vial, to which ~5 mL DMF was added. Acetic anhydride (5 µL; 2 eq.) was added to the solution, and was heated in the microwave for 30 minutes at 80°C. Triethylamine (5 µL; 2 eq.) was added to the mixture, following further microwave irradiation for 60 minutes at 100°C. The resin was filtered through a fritted syringe, then washed 3x with 10 mL DMF and 1x with 10mL diethyl ether, followed by peptide cleavage (Procedure D). The crude peptide was used for LC-MS analysis following procedure P.

8.7.2 Thioether elimination using mCPBA

On-resin peptide **75** (ProTide resin, loading 0.55 mmol/g; ~25 mg) containing an N-terminal piperidine-acrylamide adduct was added to a fritted polypropylene syringe containing ~5 mL DCM and mCPBA (purity 60%, 20 mg; 5 eq.). The mixture was shaken (600 rpm) at room temperature for 2 hours, after which the resin was washed with DCM (3 x 10 mL) and diethyl ether (1 x 10 mL). Subsequently the peptide was cleaved from the resin following standard procedure D. The crude peptide was used for LC-MS analysis following procedure P.

8.8 Biological evaluation of MDM2 peptides (**214**, **215**, **221**-**226**)

8.8.1 Protein expression and purification (Newcastle – NICR)

PCR amplified DNA encoding MDM2 17–125 and MDMX 22–111 was cloned into pGEX 6P-1 to allow expression as GST fusion proteins. Recombinant DNA was transformed into *E. coli* BL21(DE3) pLysS for expression in lysogeny broth (LB) media at 18°C and 160 rpm following isopropyl β-D-1thiogalactopyranoside (IPTG) induction (100 mM) at an OD_{600} of 0.6–1.0. Following overnight incubation, cultures were harvested, sonicated and the fusion proteins were purified from the resulting cell lysate using GST-affinity chromatography and gel filtration, using a HiLoad® 26/60 Superdex 75 column equilibrated in modified HEPES-buffered saline [mHBS; 20 mM HEPES, 100 mM NaCl, 5 mM dithiothreitol (DTT), pH 7.4]. Fractions containing purified protein were pooled, concentrated and frozen at -80°C before performing the SPR analysis.

8.8.2 Surface plasmon resonance (SPR) binding experiments (Newcastle – NICR)

SPR experiments were performed on a Biacore S200 instrument (GE Healthcare) at 25°C using a Series S carboxyl-derivatised sensor chip (CM5) prepared for capture of GST, GST-MDM2 and GST-MDMX following immobilisation of anti-GST antibody by using GST-capture and amine coupling kits (GE Healthcare). A running buffer containing 10 mM HEPES, pH 7.4, 150 mM NaCl, 0.05% Tween 20 and 3 mM ethylenediaminetetraacetic acid (EDTA) was used during sensor chip preparation. Polyclonal goat anti-GST antibody (30 µg/mL) was prepared in the immobilisation buffer (10 mM sodium acetate, pH 5.0) and was immobilised to the CM5 chip through amine coupling after injection of *N'*-(3-dimethylaminopropyl)-*N*-ethylcarbodiimide/*N*-hydroxysuccinimide (EDC/NHS; 1:1) onto the sensor chip surface for a contact time of 840 sec and at a flow rate of 5 µL/min. The antibody was injected onto the activated surface for 600 sec at 5 µL/min and the unreacted groups were then deactivated by injection of ethanolamine for 420 sec at 10 µL/min. High affinity sites were blocked by injecting recombinant GST (5 µg/mL in running buffer) twice for 300 sec at 5 µL/min prior to regenerating the sensor surface through injection of regeneration solution (10 mM glycine-HCl, pH 2.1) for 120 sec at 10 µL/min.

Analytes were dissolved in 100% DMSO to 20 mM and an Echo 550 acoustic dispenser (Labcyte) was used to dispense 12 concentration-response points for each analyte, as 1 µL droplets into a 384-well microplate (Greiner). The droplets were diluted (1:100) with 99 µL of running buffer supplemented with 1 mM DTT prior to analysis. As for the analytes, 100% DMSO was also acoustically dispensed and diluted as a 12-point concentration series to allow for solvent correction during the analysis. Freshly thawed GST or GST-fusion protein diluted in running buffer was applied to reference or test channels, respectively for 600 sec or until satisfactory response unit levels (>1000 RU) had been achieved. Binding experiments were performed at a flow rate of 30 µL/min using multicycle kinetics with injection of analytes over the captured ligand for a contact time of 60 sec, followed by a dissociation period of 1500 sec, with multi-channel data collection at 10 Hz. During multicycle analysis, running buffer was modified with 1 mM DTT and 1% (*v/v*) DMSO to match with analyte composition. Data evaluation was performed using Biacore S200 Evaluation software (Version 1.0, Build 20) with binding curves fit to a1:1 (Langmuir) interaction model for evaluation of kinetic and affinity parameters, following solvent correction and reference channel subtraction.

Appendices

- 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.

- S.J.M. Verhoork, P.M. Killoran, C.R. Coxon, *Biochemistry*, **2018**, 5743, 6132–6143.

ChemPubSoc Europe

DOI: 10.1002/chem.201804163

CHEMISTRY
A European Journal
Communication

Stapled Peptides |*Hot Paper*|

Tuning the Binding Affinity and Selectivity of Perfluoroaryl-Stapled Peptides by Cysteine-Editing

Sanne J. M. Verhoork,[a] Claire E. Jennings,[b] Neshat Rozatian,[c] Judith Reeks,[b] Jieman Meng,[b] Emily K. Corlett,[c] Fazila Bunglawala,[a] Martin E. M. Noble,[b] Andrew G. Leach,[a] and Christopher R. Coxon*[a]

Abstract: A growing number of approaches to "staple" α-helical peptides into a bioactive conformation using cysteine cross-linking are emerging. Here, the replacement of L-cysteine with "cysteine analogues" in combinations of different stereochemistry, side chain length and beta-carbon substitution, is explored to examine the influence that the thiol-containing residue(s) has on target protein binding affinity in a well-explored model system, p53–MDM2/MDMX, which is constituted by the interaction of the tumour suppressor protein p53 and proteins MDM2 and MDMX, which regulate p53 activity. In some cases, replacement of one or more L-cysteine residues afforded significant changes in the measured binding affinity and target selectivity of the peptide. Computationally constructed homology models indicate that some modifications, such as incorporating two D-cysteine residues, favourably alter the positions of key functional amino acid side chains, which is likely to cause changes in binding affinity, in agreement with measured surface plasmon resonance data.

Linear, unstructured peptide sequences often suffer from low proteolytic stability when excised from their parent proteins, limiting their development as potential therapeutics. Stapled α-helical peptides (SAHs) are a highly-promising class of therapeutic agents, which are designed to mimic an α-helical motif of a protein, and have superior proteolytic stability in vivo over the equivalent unconstrained peptide.[1] The most common method of peptide stapling employs the use of the all-hydrocarbon (alkene) linker developed by Grubbs and Blackwell,[2–4] and pioneered by the Verdine group. This strategy is used to stabilise a peptide α-helix and can often deliver impressive biological activity through steric constraint of a bio-active conformation.[5] Using the alkene metathesis approach requires the incorporation of α,α-disubstituted alkene-containing amino acids into the peptide sequence. Typically, these building blocks are either purchased at significant expense or can be obtained by multistep synthesis using, for example, nucleophilic glycine equivalents.[6] The standard all-hydrocarbon stapling approach typically incorporates a combination of *R,S*- or *S,R*-α,α′-disubstituted alkenyl amino acids with optimised chain length.[7,8] It is well known that linker length, linker orientation and linker type of the stapled peptides can affect the binding properties.[9,10] Typically, the rationale for inclusion of an additional α-methyl group was to overcome the perceived destabilising effect upon helical conformation by introducing D-amino acids.[7] However, mono-substituted α-alkenyl amino acids have been shown to be similarly effective.[11] The importance of stereochemical effects on the helical character and, thus, biological activity have been clearly demonstrated in the alkene metathesis *i*, *i* + 4 peptide stapling approach.[12]

The recent literature has shown a significant surge of interest in the two-component chemoselective cross-linking of peptides through, for example, "double-click" copper(I)-catalysed alkyne–azide cycloaddition (CuAAC) chemistry.[13–15] Most often, however, cysteine thiol residues have provided an excellent handle for peptide stapling, driven mainly by the ease and relatively low cost of obtaining the linear pre-stapled peptide. This topic has been recently reviewed by Fairlie.[16] Examples of thiol cross-linking include the use of dibromomaleimide,[17] dichloroacetone,[18] 1-,4-dichlorotetrazine,[19] 1,2,4,5-tetrabromodurene,[20] α,α′-dibromo-*m*-xylene,[21] *trans*-1,4-dibromo-2-butene and *cis*-1,4-dichloro-2-butene[22] and perfluoroaryl reagents.[23] Although significant attention has focussed on the nature of the cross-linking electrophile, comparatively little, if any, attention has focussed on the cysteine residues, with the single exception of introducing homocysteine.[17] The distance between the cysteine residues has been explored and optimised, albeit in non-helical systems,[24] yet the stereochemistry of cysteine has not been taken into account, in terms of the consequences on biological activity and the position of key amino acid side chain residues. By drawing analogy with the traditional all-hydrocarbon approach, we investigated the replacement of L-cysteine (cysteine-editing) with selected combi-

[a] *S. J. M. Verhoork, F. Bunglawala, Dr. A. G. Leach, Dr. C. R. Coxon*
School of Pharmacy and Biomolecular Sciences
Liverpool John Moores University
James Parsons Building, Byrom St, Liverpool, L3 3AF (UK)
E-mail: c.r.coxon@ljmu.ac.uk

[b] *Dr. C. E. Jennings, Dr. J. Reeks, J. Meng, Prof. M. E. M. Noble*
Northern Institute for Cancer Research, Newcastle University
Paul O'Gorman Building, Medical School, Framlington Place
Newcastle upon Tyne, NE2 4HH (UK)

[c] *N. Rozatian, E. K. Corlett*
Department of Chemistry, Durham University
South Road, Durham, DH1 3LE (UK)

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/chem.201804163.

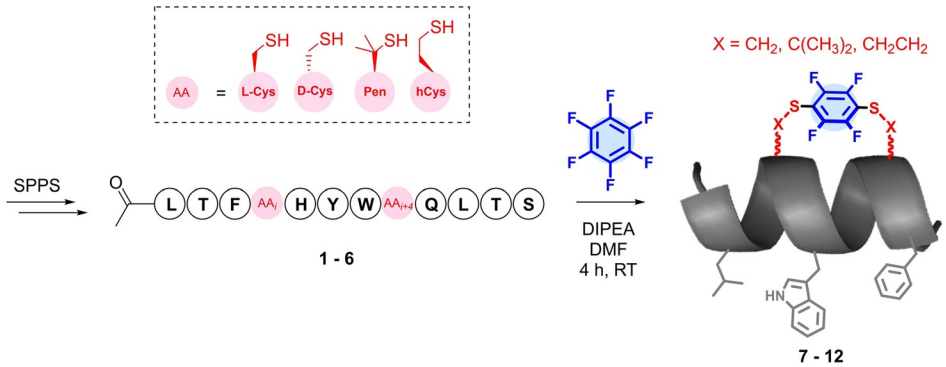
nations of D-cysteine, homocysteine and penicillamine to examine the effect of 1) stereochemistry, 2) cysteine homologation, and 3) beta-carbon substitution. We considered that the outcomes will be directly important to the cysteine-stapling work of other groups, as highlighted above.

The p53–MDM2 and p53–MDMX protein–protein interactions were selected as model systems in which to study cysteine-editing, due to their well-characterised interactions and the availability of published known stapled α-helical peptide inhibitors.[13,14,25–27] The p53 tumour suppressor is a major regulator of the cell cycle and is activated in response to genotoxic stress resulting from oncogenic signalling and exposure to, for example, ionising radiation and carcinogenic agents.[28] Through its role as a transcription factor, p53 induces cell cycle arrest and apoptosis in afflicted cells,[29] fulfilling a critical role in the maintenance of healthy functioning of cells and the avoidance of malignancy. MDM2 (murine double minute 2) and MDMX (also known as MDM4) inhibit the transcriptional activity of p53 and promote its degradation.[30] In many p53 wild-type tumours, overexpression of MDM2 and/or MDMX leads to silencing of p53 and prevents its activation in response to cellular stress,[31] making the p53-MDM2 / MDMX protein–protein interaction one of the most widely researched targets for cancer therapy. A 12-amino-acid peptide LTFEHYWAQLTS (PDI peptide) identified by phage display[32] was reported to disrupt the p53–MDM2 protein–protein interaction and has previously been stapled using non-hydrocarbon techniques.[33,34] This served as a test-bed for diversification in our studies. The key features of this peptide (and indeed, the p53 protein) that promote the biological activity are the three amino acids, Phe, Trp and Leu, with the positions being important for activity. These residues were retained and other selected residues that were previously reported to be tolerant to substitution[33] were replaced with cysteine analogues in the PDI sequence with the standard relative spacing of four (*i*, *i*+4) amino acids, corresponding to equivalent positions at neighbouring turns on the α-helix. Seven different combinations were synthesised using solid-phase peptide synthesis on Rink amide resin to afford the C-terminal amide of the form Ac-LTF(**AA**$_i$)HYW(**AA**$_{i+4}$)QLTS (Table 1). Stapling was performed using the cross-linking reagent hexafluorobenzene (Scheme 1), as demonstrated by Pentelute and co-workers,[23] due to an ongoing interest in related reactions in our laboratory. In each case the thiol cross-linking occurred cleanly under relatively mild conditions [25 mM *N*,*N*-diisopropylethylamine (DIPEA) in DMF, room temperature, $<$4.5 h].[35,36]

Table 1. Dissociation constants (K_d) for peptides (**1**, **7**–**12**) binding to MDM2/MDMX attained from an SPR assay and comparison with control inhibitor nutlin-3 a and PDI peptide. hCys = homocysteine, Pen = penicillamine. ND = not determined.

Peptide	AA$_i$	AA$_{i+4}$	Cross-linker	K_d [μM]		Chi^2 [RU^2]	
				MDM2	MDMX	MDM2	MDMX
PDI	L-Glu	L-Ala	–	0.04	0.02	1.36	0.82
Nutlin-3 a	–	–	–	0.17	ND	0.24	ND
1	L-Cys	L-Cys	–	0.18	0.18	1.49	0.40
7	L-Cys	L-Cys		1.02	0.44	2.37	0.81
8	L-Cys	D-Cys		0.41	0.37	3.08	0.99
9	D-Cys	L-Cys		1.70	0.40	3.69	6.49
10	D-Cys	D-Cys		0.22	0.06	1.71	0.86
11	L-hCys	L-hCys		0.14	0.15	4.07	2.91
12	L-Pen	L-Pen		16.90	9.89	10	8.17

The binding affinity of the synthesised p53-mimicking peptides **1** and **7**–**12** were examined by measuring the dissociation constants (K_d) for their interactions with MDM2 (amino acids 17–125) and MDMX (amino acids 22–111) expressed as glutathione S-transferase (GST) fusion protein constructs, using surface plasmon resonance (SPR; Table 1, Figure 1 and Supporting Information). The binding affinities were measured for the PDI peptide and the *cis*-imidazoline small molecule, nutlin-3 a, as positive binding controls and to validate the SPR approach against a biochemical homogenous time-resolved fluorescence assay (HTRF) assay (see Supporting Information). The SPR evaluation identified a number of highly potent tetrafluorobenzene-cross-linked SAHs with low-to-moderate micromolar affinities for MDM2 and MDMX, as measured by K_d values. In general, the perfluoroaryl-stapled peptides had higher K_d (lower affinity) than the phage-displayed PDI peptide; however, the additional proteolytic stability gained from this modification[23] (see Supporting Information) may offset the sacrifice in affinity. In fact, introduction of L-cysteine residues into the non-stapled PDI analogue (**1**) decreased equally binding affinity for both MDM2 and MDMX compared to PDI. In any case, the primary purpose of this study was for comparison of the



Scheme 1. Synthesis of perfluoroaryl-stapled peptides.

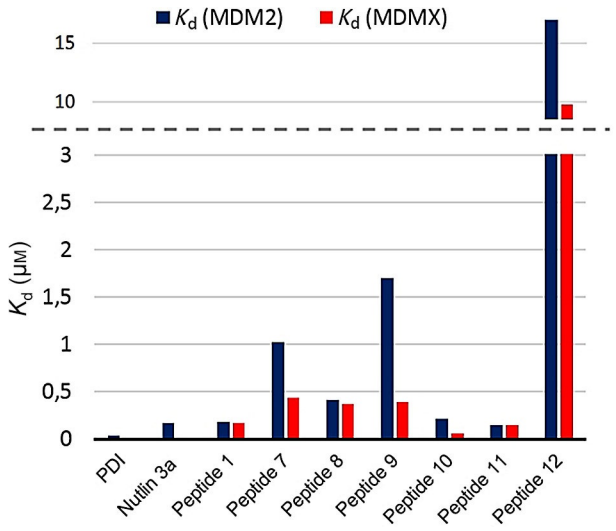


Figure 1. Comparison of binding affinities (K_d) for stapled peptides containing alternative cysteine analogues.

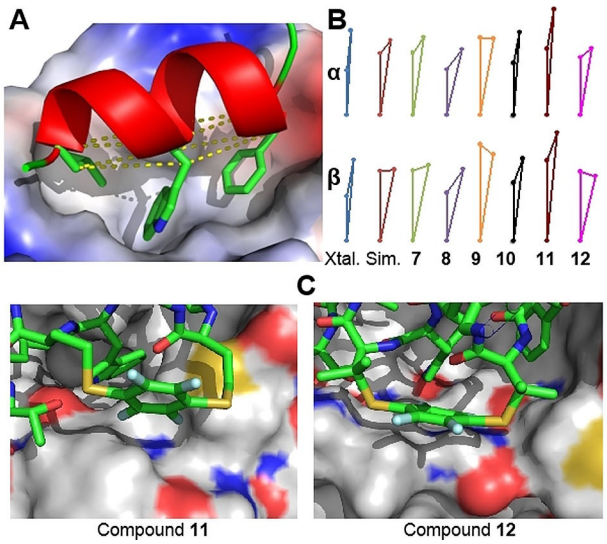


Figure 2. A) X-ray structure of MDM2–PDI peptide complex (PDB: 3G03) with C_α and C_β triangles shown with dotted yellow lines. B) Comparison of simulated cross-linked peptide Phe, Trp and Leu side chain positions. Xtal = from X-ray structure 3G03, Sim = simulated PDI peptide. C) Simulated models of **11** and **12**, providing possible interactions with MDM2.

parent **7** (L-Cys, L-Cys) with peptides with different cysteine analogues comprising the cross-link. One particularly interesting outcome is that most of the stapled peptides [**7** (>2-fold), **9** (>4-fold), **10** (>3.5-fold)] had a generally higher affinity for MDMX compared with MDM2, whereas PDI and **1** were equipotent for each isoform, albeit with lower K_d values. In general, changing a single cysteine stereochemistry from L- to D- at *i* or *i*+4 positions was well tolerated by both MDM2 and MDMX; however, MDMX appeared generally more tolerant to "cysteine-editing" than MDM2. Inversion of both L-*i* and -*i*+4 α-carbon substituents to the D-configuration (**10**) significantly enhanced the affinity for both MDMX (≈7-fold) and MDM2 (≈5-fold) compared to **7**. This has particular importance because peptides comprising D-amino acids are typically more resistant to proteolytic degradation than their canonical counterparts.

Most notably, branching of L-cysteine at the β-position with geminal dimethyl groups [L-penicillamine (L-Pen), **12**] exhibited a significantly higher K_d than all other analogues, indicating lower affinity for both MDM2 (>16-fold vs. $i_{L\text{-}Cys}$, $i+4_{L\text{-}Cys}$, **7**) and MDMX (≈22-fold vs. $i_{L\text{-}Cys}$, $i+4_{L\text{-}Cys}$, **7**). This may suggest that the L-Pen-containing peptide is significantly distorted from a well-defined α-helix or presents destabilising interactions. In stark contrast, homologation of the *i*, *i*+4 cross-linker through incorporation of homocysteine (**11**) appeared to be much better tolerated than L-cysteine, affording around a seven-fold lower K_d (7-fold higher affinity) for MDM2 and three-fold lower for MDMX versus the parent **7**. This was around equipotent with the non-stapled **1**, which itself was also non-selective for either MDM2 or MDMX. Overall, the observed structure–activity relationships (SARs) may be related to the geometric constraints imposed by the cysteine analogue and the cross-linker, and the resulting impact upon the helicity of the peptide and the relative positions of key amino acid side chains, Phe, Trp and Leu.

In light of these observations, we employed an in silico modelling approach (Figure 2) to understand the structural and conformational consequences of cysteine replacement. The apparent α-helicity of **7** was initially measured using circular dichroism but produced poor results (see Supporting Information) due to the absorbance of the fluoroaryl moiety at 222 nm, which was consistent with previous literature reports.[17] Molecular dynamics simulations were performed starting from homology models for the free peptides **1**–**6**. The homology model used the sequence and chain B from the structure of human MDM2 in complex with the reported high affinity PDI peptide [Protein Data Bank (PDB) code: 3G03], in which the Phe, Trp and Leu residues provide key points of interaction with MDM2 (Figure 2A). To benchmark simulations of the cysteine replacement peptides, this native ligand was also simulated in the same way as described below for the cross-linked complexes. In each simulation, the resulting geometries were assessed for their ability to place the three key binding residues in appropriate positions for MDM2 interaction. This used an analogy to pharmacophore triplets, a commonly used description in chemoinformatics that uses the three-dimensional positioning of three pharmacophoric points (usually key interactions, such as hydrogen-bonding or hydrophobic groups) as a descriptor and benefits from the ease of analysis and understanding of the geometry of triangles. In this case, two triangles were used to describe how modification of the peptide alters the positions of Phe, Trp and Leu side chains: one triangle formed from the three C_α atoms and the second from the three C_β atoms, shown as dashed lines in Figure 2A. This initial analysis revealed that the two triangles are rather similar because the C_α–C_β bonds are all pointed in approximately the same direction and this is clearly an important part of how the peptide forms tight interactions with the receptor.

The simulations of the cross-linked peptides involved two stages: 1) the peptide with the two cysteines in their free thiol form was simulated (in MOE using default settings; MOE = Mo-

lecular Operating Environment) starting from each of the nine possible rotamers of the cysteine (arising from rotation about the C_α–C_β bond in each of the cysteines); 2) the rotamer that positioned the two sulfur atoms closest to the separation between the two atoms in the cyclised adduct (6.37 Å) was selected and the linker was introduced by editing the molecule in MOE. The edited structure was energy minimised and a second simulation performed. In both stages, the default settings in MOE were employed. This entails use of the Nosé–Poincaré–Andersen (NPA) algorithm, using the AMBER10 forcefield with implicit solvation (with interior and exterior dielectric constants of 1 and 80, respectively). An initial 100 ps of equilibration was followed by 500 ps of production of which the second half (last 250 ps) was used in the analysis (reported every 0.5 ps giving a total of 500 data points from each simulation). The simulations can be summarised succinctly by considering the average values of each side of the triangle equivalent to those shown in Figure 2B.

When the average distances were compared with those observed in the published MDM2 complex, a root-mean-square deviation (RMSD) could be computed to permit an overall comparison of how well the free stapled peptide retains the geometry required for complex formation. This suggested that the double D-Cys-containing stapled peptide **10** (RMSD = 1.0 Å) would retain the required pharmacophoric arrangement better than even the native peptide, which adopts a slightly different geometry when free from the receptor. The next best was predicted to be compound **8** (RMSD = 1.2 Å), followed by **12** and **7**. The simulations correctly identified **10** as the best of the analogues in which only stereochemistry is varied, whereas compound **9** can be considered the least suitable by this measure. Although this is in good agreement with the SPR data for stereoisomeric peptides, the approach does not appear able to correctly rank the structural variations in which methylation or homologation have been introduced, indicating that other factors also govern their interaction and affinity with the protein. To provide insight into these two structures, molecular editing in MOE was used to convert the native complex to the stapled form for compounds **11** and **12**. The complex was then simulated to investigate any extra contacts made by these two linkers that could explain the observed binding. The final snapshot is shown in Figure 2C and reveals that **11** is able to lay its linker on a hydrophobic part of the receptor surface. Although compound **12** is also able to form some hydrophobic interactions, the shape of this linker is not amenable to making continuous contact because of the protrusion of one of the methyl groups in the penicillamine. These are particularly close to the side chain of Met62, which is in a more constrained environment when compound **12** is bound. Overall, these insights help to explain the differential measured binding affinities following cysteine-editing.

This work has demonstrated that the conformational properties of a stapled peptide, and thus the biological activity, can be modified by the nature (size and stereochemistry) of the thiol groups to be cross-linked and, indeed, the combination of these with a suitable cross-linker. This has clear implications in the tuning of binding affinity and/or target selectivity in two-component disulfide stapling of α-helical peptides and provides an important new tool in this rapidly growing area.

Experimental Section

Materials: All Fmoc L- and D-amino acids (CEM), Rink Amide ProTide resin (CEM), diisopropylcarbodiimide (DIC; Apollo Scientific), Oxyma Pure™ (CEM), N.N′-dimethylformamide (DMF; Fisher Scientific), diisopropylethylamine (DIPEA; Merck Millipore) and piperidine (Merck Millipore) were purchased from commercial suppliers and used directly as indicated in the appropriate experimental procedures. All other reagents [hexafluorobenzene, trifluoroacetic acid (TFA), triisopropylsilane (TIPS)] were purchased from Sigma Aldrich and solvents (HPLC grade) were purchased from Fisher Scientific. Nutlin-3a was purchased from NewChem Technologies Ltd, Newcastle upon Tyne, UK.

Solid-phase peptide synthesis (SPPS) of peptides 1–6 and PDI peptide: Each linear thiol-containing peptide sequence was prepared using automated Fmoc-SPPS methods on a Liberty Blue microwave-assisted peptide synthesiser (CEM). Solid-phase synthesis was conducted using Rink amide ProTide resin (180 mg, 0.56 mmol g^{-1} loading; 0.1 mmol), employing the required Fmoc amino acids (0.2 M in DMF, 5 equiv.); with DIC (1 M stock solution in DMF; 10 equiv.), Oxyma Pure (1 M stock solution, 5 equiv) and piperidine (20% *v/v* in DMF; 587 equiv., 4 mL) as the activator and deprotection agent, respectively. Standard coupling procedures employed double coupling of each amino acid (2.5 min, 90 °C). Amino acids bearing thermally-sensitive protecting groups, for example, Fmoc-L-Cys(Trt)-OH, Fmoc-D-Cys(Trt)-OH, Fmoc-Pen(Trt)-OH, Fmoc-hCys(Trt)-OH and Fmoc-His(Boc)-OH were coupled under milder conditions (50 °C for 10 min). Following on-resin synthesis of the appropriate sequence, N-terminal capping was performed using Ac_2O/DMF (20% *v/v*, 2×15 min) with shaking at room temperature. Finally, peptides were cleaved from the resin as the C-terminal amide by treatment with a cleavage cocktail [5 mL; comprising TFA, TIPS and water (9:0.5:0.5 *v/v*)] with regular shaking at room temperature for 4 h. Peptides were precipitated from cleavage solutions by dropwise addition into cold diethyl ether followed by centrifugation. The resulting pellet was successively suspended in cold diethyl ether and centrifuged twice further. The solids obtained were dissolved in water/MeCN (depending on the solubility), frozen and lyophilised. PDI and peptide **1** were purified by semi-preparative HPLC (see Supporting Information). The crude disulfide peptides **1**–**6** were used without further chromatographic purification.

Synthesis of perfluoroaryl-stapled peptides 7–12: To a centrifuge tube containing the solid crude peptide **1**–**6** (approx. 50 mg) was added DIPEA stock solution (25 mM in DMF, 1.0 mL), followed by hexafluorobenzene (20 equiv., 0.66 mmol). The resulting mixture was shaken at room temperature for 4.5 h. After precipitation in Et_2O (as above), the resulting crude stapled peptide was redissolved in a mixture of water/MeCN (exact amounts depended on solubility). Crude samples were purified using semi-preparative HPLC. Stapled peptides were characterised by analytical HPLC and high-resolution mass spectrometry (see Supporting Information).

Protein expression and purification: PCR amplified DNA encoding MDM2 17–125 and MDMX 22–111 was cloned into pGEX 6P-1 to allow expression as GST fusion proteins. Recombinant DNA was transformed into *E. coli* BL21(DE3) pLysS for expression in lysogeny broth (LB) media at 18 °C and 160 rpm following isopropyl β-D-1-thiogalactopyranoside (IPTG) induction (100 mM) at an OD_{600} of 0.6–1.0. Following overnight incubation, cultures were harvested,

sonicated and the fusion proteins were purified from the resulting cell lysate using GST-affinity chromatography and gel filtration, using a HiLoad® 26/60 Superdex 75 column equilibrated in modified HEPES-buffered saline [mHBS; 20 mM HEPES, 100 mM NaCl, 5 mM dithiothreitol (DTT), pH 7.4]. Fractions containing purified protein were pooled, concentrated and frozen at −80 °C before performing the SPR analysis.

Surface plasmon resonance (SPR) binding experiments: SPR experiments were performed on a Biacore S200 instrument (GE Healthcare) at 25 °C using a Series S carboxyl-derivatised sensor chip (CM5) prepared for capture of GST, GST-MDM2 and GST-MDMX following immobilisation of anti-GST antibody by using GST-capture and amine coupling kits (GE Healthcare). A running buffer containing 10 mM HEPES, pH 7.4, 150 mM NaCl, 0.05% Tween 20 and 3 mM ethylenediaminetetraacetic acid (EDTA) was used during sensor chip preparation. Polyclonal goat anti-GST antibody (30 $\mu g\,mL^{-1}$) was prepared in the immobilisation buffer (10 mM sodium acetate, pH 5.0) and was immobilised to the CM5 chip through amine coupling after injection of *N*′-(3-dimethylaminopropyl)-*N*-ethylcarbodiimide/*N*-hydroxysuccinimide (EDC/NHS; 1:1) onto the sensor chip surface for a contact time of 840 s and at a flow rate of 5 $\mu L\,min^{-1}$. The antibody was injected onto the activated surface for 600 s at 5 $\mu L\,min^{-1}$ and the unreacted groups were then deactivated by injection of ethanolamine for 420 s at 10 $\mu L\,min^{-1}$. High affinity sites were blocked by injecting recombinant GST (5 $\mu g\,mL^{-1}$ in running buffer) twice for 300 s at 5 $\mu L\,min^{-1}$ prior to regenerating the sensor surface through injection of regeneration solution (10 mM glycine-HCl, pH 2.1) for 120 s at 10 $\mu L\,min^{-1}$.

Analytes were dissolved in 100% DMSO to 20 mM and an Echo 550 acoustic dispenser (Labcyte) was used to dispense 12 concentration-response points for each analyte, as 1 μL droplets into a 384-well microplate (Greiner). The droplets were diluted (1:100) with 99 μL of running buffer supplemented with 1 mM DTT prior to analysis. As for the analytes, 100% DMSO was also acoustically dispensed and diluted as a 12-point concentration series to allow for solvent correction during the analysis. Freshly thawed GST or GST-fusion protein diluted in running buffer was applied to reference or test channels, respectively for 600 s or until satisfactory response unit levels (>1000 RU) had been achieved. Binding experiments were performed at a flow rate of 30 $\mu L\,min^{-1}$ using multicycle kinetics with injection of analytes over the captured ligand for a contact time of 60 s, followed by a dissociation period of 1500 s, with multi-channel data collection at 10 Hz. During multicycle analysis, running buffer was modified with 1 mM DTT and 1% (*v*/*v*) DMSO to match with analyte composition. Data evaluation was performed using Biacore S200 Evaluation software (Version 1.0, Build 20) with binding curves fit to a 1:1 (Langmuir) interaction model for evaluation of kinetic and affinity parameters, following solvent correction and reference channel subtraction.

Acknowledgements

We thank The Royal Society for a Research Grant (RG160956) and Liverpool John Moores University for a generous fully funded Ph.D. scholarship for S.J.M.V., which supported this work. Studentships (NR and EKC, 2014) were funded through the Biophysical Sciences Institute, Durham University or Liverpool John Moores University (FB, 2016). J.R. and C.J. were funded by Astex Pharmaceuticals, Cambridge, UK.

Conflict of interest

The authors declare no conflict of interest.

Keywords: cancer · computational modelling · cysteine · hexafluorobenzene · peptides

[1] N. S. Robertson, A. G. Jamieson, *Rep. Org. Chem.* **2015**, *5*, 65 – 74.
[2] S. J. Miller, R. H. Grubbs, *J. Am. Chem. Soc.* **1995**, *117*, 5855 – 5856.
[3] S. J. Miller, H. E. Blackwell, R. H. Grubbs, *J. Am. Chem. Soc.* **1996**, *118*, 9606 – 9614.
[4] H. E. Blackwell, R. H. Grubbs, *Angew. Chem. Int. Ed.* **1998**, *37*, 3281 – 3284; *Angew. Chem.* **1998**, *110*, 3469 – 3472.
[5] Examples include: a) L. D. Walensky, A. L. Kung, I. Escher, T. J. Malia, S. Barbuto, R. D. Wright, G. Wagner, G. L. Verdine, S. J. Korsmeyer, *Science* **2004**, *305*, 1466 – 1470; b) L. D. Walensky, K. Pitter, J. Morash, K. J. Oh, S. Barbuto, J. Fisher, E. Smith, G. L. Verdine, S. J. Korsmeyer, *Mol. Cell* **2006**, *24*, 199 – 210.
[6] Examples include: a) R. M. Williams, P. J. Sinclair, D. Zhai, D. Chen, *J. Am. Chem. Soc.* **1988**, *110*, 1547 – 1557; b) R. M. Williams, M. N. Im, *J. Am. Chem. Soc.* **1991**, *113*, 9276 – 9286; c) Y. N. Belokon, V. I. Bakhmutov, N. I. Chernoglazova, K. A. Kochetkov, S. V. Vitt, N. S. Garbalinskaya, V. M. Belikov, *J. Chem. Soc. Perkin Trans. 1* **1988**, 305 – 312; d) A. E. Sorochinsky, J. L. Aceña, H. Moriwaki, T. Sato, V. A. Soloshonok, *Amino Acids* **2013**, *45*, 691 – 718; e) A. E. Sorochinsky, J. L. Aceña, H. Moriwaki, T. Sato, V. A. Soloshonok, *Amino Acids* **2013**, *45*, 1017 – 1033; f) J. L. Aceña, A. E. Sorochinsky, V. A. Soloshonok, *Amino Acids* **2014**, *46*, 2047 – 2073.
[7] C. E. Schafmeister, J. Po, G. L. Verdine, *J. Am. Chem. Soc.* **2000**, *122*, 5891 – 5892.
[8] Y. W. Kim, T. N. Grossmann, G. L. Verdine, *Nat. Protoc.* **2011**, *6*, 761.
[9] T. E. Speltz, S. W. Fanning, C. G. Mayne, C. Fowler, E. Tajkhorshid, G. L. Greene, T. W. Moore, *Angew. Chem. Int. Ed.* **2016**, *55*, 4252 – 4255; *Angew. Chem.* **2016**, *128*, 4324 – 4327.
[10] K. Hu, H. Geng, Q. Zhang, Q. Liu, M. Xie, C. Sun, W. Li, H. Lin, F. Jiang, T. Wang, Y. D. Wu, *Angew. Chem. Int. Ed.* **2016**, *55*, 8013 – 8017; *Angew. Chem.* **2016**, *128*, 8145 – 8149.
[11] D. J. Yeo, S. L. Warriner, A. J. Wilson, *Chem. Commun.* **2013**, *49*, 9131 – 9133.
[12] Y. W. Kim, G. L. Verdine, *Bioorg. Med. Chem. Lett.* **2009**, *19*, 2533 – 2536. Note: there are many other examples that could be cited.
[13] Y. H. Lau, P. de Andrade, S.-T. Quah, M. Rossmann, L. Laraia, N. Skold, T. J. Sum, P. J. E. Rowling, T. L. Joseph, C. Verma, M. Hyvonen, L. S. Itzhaki, A. R. Venkitaraman, C. J. Brown, D. P. Lane, D. R. Spring, *Chem. Sci.* **2014**, *5*, 1804 – 1809.
[14] Y. H. Lau, P. de Andrade, N. Skold, G. J. McKenzie, A. R. Venkitaraman, C. Verma, D. P. Lane, D. R. Spring, *Org. Biomol. Chem.* **2014**, *12*, 4074 – 4077.
[15] Y. H. Lau, Y. Wu, P. de Andrade, W. R. J. D. Galloway, D. R. Spring, *Nat. Protoc.* **2015**, *10*, 585 – 594.
[16] D. P. Fairlie, A. D. de Araujo, *J. Pept. Sci.* **2016**, *106*, 843 – 852.
[17] C. M. Grison, G. M. Burslem, J. A. Miles, L. K. A. Pilsl, D. J. Yeo, Z. Imani, S. L. Warriner, M. E. Webb, A. J. Wilson, *Chem. Sci.* **2017**, *8*, 5166 – 5171.
[18] N. Assem, D. J. Ferreira, D. W. Wolan, P. E. Dawson, *Angew. Chem. Int. Ed.* **2015**, *54*, 8665 – 8668; *Angew. Chem.* **2015**, *127*, 8789 – 8792.
[19] S. P. Brown, A. B. Smith, *J. Am. Chem. Soc.* **2015**, *137*, 4034 – 4037.
[20] L. E. J. Smeenk, N. Dailly, H. Hiemstra, J. H. van Maarseveen, P. Timmerman, *Org. Lett.* **2012**, *14*, 1194 – 1197.
[21] A. D. de Araujo, H. N. Hoang, W. M. Kok, F. Diness, P. Gupta, T. A. Hill, R. W. Driver, D. A. Price, S. Liras, D. P. Fairlie, *Angew. Chem. Int. Ed.* **2014**, *53*, 6965 – 6969; *Angew. Chem.* **2014**, *126*, 7085 – 7089.
[22] P. Diderich, D. Bertoldo, P. Dessen, M. M. Khan, I. Pizzitola, W. Held, J. Huelsken, C. Heinis, *ACS Chem. Biol.* **2016**, *11*, 1422 – 1427.
[23] A. M. Spokoyny, Y. Zou, J. J. Ling, H. Yu, Y. S. Lin, B. L. Pentelute, *J. Am. Chem. Soc.* **2013**, *135*, 5946 – 5949.
[24] 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.
[25] F. Bernal, A. F. Tyler, S. J. Korsmeyer, L. D. Walensky, G. L. Verdine, *J. Am. Chem. Soc.* **2007**, *129*, 2456 – 2457.

[26] Y. S. Chang, B. Graves, V. Guerlavais, C. Tovar, K. Packman, K.-H. To, K. A. Olson, K. Kesavan, P. Gangurde, A. Mukherjee, T. Baker, K. Darlak, C. Elkin, Z. Filipovic, F. Z. Qureshi, H. Cai, P. Berry, E. Feyfant, X. E. Shi, J. Horstick, D. A. Annis, A. M. Manning, N. Fotouhi, H. Nash, L. T. Vassilev, T. K. Sawyer, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E3445 – E3454.
[27] C. J. Brown, S. T. Quah, J. Jong, A. M. Goh, P. C. Chiam, K. H. Khoo, M. L. Choong, M. A. Lee, L. Yurlova, K. Zolghadr, T. L. Joseph, *ACS Chem. Biol.* **2013**, *8*, 506 – 512.
[28] D. P. Lane, *Nature* **1992**, *358*, 15 – 16.
[29] K. S. Yee, K. H. Vousden, *Carcinogenesis* **2005**, *26*, 1317 – 1322.
[30] X. Wu, J. H. Bayle, D. Olson, A. J. Levine, *Genes Dev.* **1993**, *7*, 1126 – 1132.
[31] B. Vogelstein, D. Lane, A. J. Levine, *Nature* **2000**, *408*, 307 – 310.
[32] J. Phan, Z. Li, A. Kasprzak, B. Li, S. Sebti, W. Guida, E. Schönbrunn, J. Chen, *J. Biol. Chem.* **2010**, *285*, 2174 – 2183.
[33] M. M. Madden, A. Muppidi, Z. Li, X. Li, J. Chen, Q. Lin, *Bioorg. Med. Chem. Lett.* **2011**, *21*, 1472 – 1475.
[34] G. Lautrette, F. Touti, H. G. Lee, P. Dai, B. L. Pentelute, *J. Am. Chem. Soc.* **2016**, *138*, 8340 – 8343.
[35] Note: the reported method employed a 50 mM solution of tris(hydroxymethyl)-aminomethane (TRIS) base. Yet, we found that DIPEA performed equally well and offered the additional advantage of being volatile and, therefore, easily removed by evaporation and allowing for direct analysis by electrospray ionisation mass spectrometry.
[36] No disulfide reduction was required prior to the nucleophilic aromatic substitution reaction.

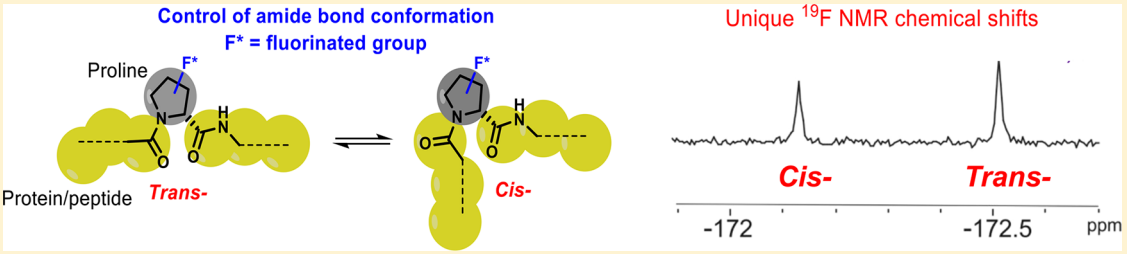
Manuscript received: August 15, 2018
Accepted manuscript online: September 26, 2018
Version of record online: November 27, 2018

Biochemistry
Perspective
Cite This: Biochemistry 2018, 57, 6132−6143
pubs.acs.org/biochemistry

Fluorinated Prolines as Conformational Tools and Reporters for Peptide and Protein Chemistry

Sanne J. M. Verhoork, Patrick M. Killoran, and Christopher R. Coxon*

School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Byrom Street Campus, Liverpool L3 3AF, U.K.



ABSTRACT: Amide bonds at the proline nitrogen are particularly susceptible to rotation, affording *cis* and *trans* isomers. Installation of a stereochemically defined electron-withdrawing fluorine atom or fluorinated groups has the power to influence the *cis*−*trans* conformational preferences of the amide bond in X−(F)Pro (where X = any other amino acid). Advantageously, this also provides a sensitive reporter for ^{19}F nuclear magnetic resonance (NMR) studies of protein conformation, interactions, and dynamics. We deliberately use the term "fluorinated prolines" as an all-encompassing term to describe proline analogues containing one or more fluorine atoms and to avoid confusion with the more well-known 4-fluoroprolines. This review presents a critical discussion of the growing repertoire of fluorinated prolines that have been described and, importantly, provides a comparison of their uses and relative influence on amide-bond conformation and discusses the significant potential of using ^{19}F NMR as a tool to probe conformational changes in polypeptides.

Of the 20 proteinogenic amino acids, proline is the only amino acid that has a side chain that forms a part of the protein backbone. This results in unique properties that are essential in peptide and protein three-dimensional structure.[1] Peptide backbone amide bonds are usually approximately planar and can have up to a >1000-fold preference for a *trans* (apart in space) arrangement, yet proline allows both *trans* and *cis* (together in space) forms to coexist (e.g., X and Y groups in Figure 1A). Proline has two main conformational equilibria: (i) amide-bond *cis*−*trans* isomerism (Figure 1A) and (ii) *endo*/*exo* pyrrolidine ring pucker (Figure 1B). The pyrrolidine ring preferentially adopts an approximate C^{γ}-*endo* or C^{γ}-*exo* pucker (Figure 1B) and occurs as a mixture in solution [66:34 *endo*:*exo* for Ac-Pro-OMe at 25 °C as derived from nuclear magnetic resonance (NMR) peak intensities].[2,3]

Ring puckering[4] and *trans*−*cis* isomerization at proline residues can influence the formation of α-helices or β-sheets as secondary protein structures through the introduction of a "kink" in the protein backbone structure.[5] However, if proline is present inside a secondary structure, it can also destabilize α-helices through the reduced hydrogen bonding capacity of the backbone nitrogen.[6,7] The presence of proline residues is often highly conserved in active-transport and ion channel proteins, where they are located in the middle of transmembrane helices, suggesting functional or structural importance.[8,9] In addition to conformational effects, prolines have also been shown to affect the folding kinetics of proteins.[10] In fully folded proteins, an X−Pro (X = any other amino acid) peptidyl−prolyl bond is ∼95% in the *trans* form and only ∼5% in the *cis* form.[11] Proline isomerization is often the rate-determining step in protein folding due to an ∼80 kJ/mol energy barrier for interconversion.[7,12,13] As prolines are important for influencing overall protein or peptide stability and folding kinetics, it forms an interesting strategy for controlling ring puckering and isomerization and for studying these events through modification or replacement of the pyrrolidine ring.

STERIC CONTROL OF PROLINE CONFORMATION

An improvement in protein stability can be advantageous for mitigating cold chain storage and transport problems associated with biologics or for understanding the functional role and dynamics of a protein. A number of conformationally biased proline analogues have been used to probe protein conformation and stability. Modifications to proline include N-alkylation or incorporation of heteroatoms into the pyrrolidine ring 4-position.[14] Sterically hindered 5-*tert*-butylproline (Tbp, **5**)[15] and 5,5-dimethylproline (Dmp, **4**) lock L-proline in a very high proportion of the *cis* conformation in peptides and proteins.[16] Various ring sizes have been explored through pipecolic acid (Pip, **2**) and azetidine-2-carboxylic acid (Aze, **3**). In general, constrained prolines provide a range of *trans*:*cis* preferences (Figure 2) in model peptide systems as measured using NMR spectroscopy,[14] and are, therefore, very useful

Received: July 24, 2018
Revised: September 25, 2018
Published: October 2, 2018

DOI: 10.1021/acs.biochem.8b00787
Biochemistry 2018, 57, 6132−6143

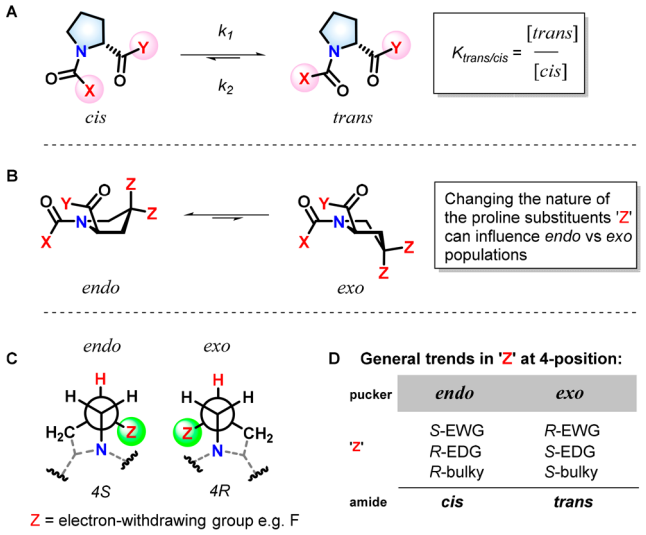


Figure 1. (A) *cis*–*trans* conformer isomerization within peptidyl–proline motifs. k_1 and k_2 are rate constants for *cis*-to-*trans* and *trans*-to-*cis* conformational changes, respectively. (B) Interconversion between *endo* and *exo* ring-puckered forms is influenced by proline ring substituents. (C) Newman projections showing the *gauche* arrangement about an N–C–C–Z axis in a 4-substituted proline and the influence on ring pucker. (D) General trends in ring pucker and X–Pro amide conformer preference (EWG, electron-withdrawing group; EDG, electron-donating group).

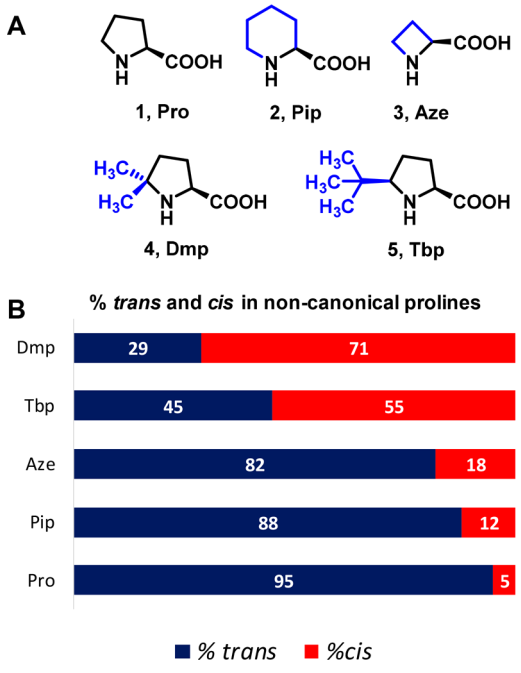


Figure 2. (A) Examples of noncanonical constrained proline "analogues" used in protein and peptide systems. (B) *cis*/*trans* preference determined in model systems. Numerical data are as reported in ref 17.

tools for understanding the stability, folding, and dynamics of peptides and proteins.

FLUORINATED PROLINES

Fluorine has enjoyed significant attention in medicinal and pharmaceutical chemistry, with an increase in occurrence from 20 to ~30% of drugs in recent years.[18,19] Introduction of fluorine has been associated with improved metabolic stability,[20] altered physicochemical properties, and increased binding affinity. Despite the prevalence of fluorine in medicinal chemistry, the main focus of research has been in small molecules, with multiple reviews describing the applicability of fluorine in this area.[18,21,22] However, the use of fluorine in peptide and protein chemistry has been explored less[23,24] and has possibly been limited by the availability of commercial fluorinated Fmoc-/Boc-protected amino acid building blocks compatible with conventional solid phase peptide synthesis or methodologies to directly and specifically fluorinate peptides. A variety of approaches are available for the synthesis of fluorinated α, β, and cyclic fluorinated amino acids, which are described in detail in an excellent review by Qiu and Qing.[25]

Fluorine-labeled amino acids have been useful in studying protein–ligand interactions using ^{19}F NMR. For example, fluorinated tyrosine and tryptophans were incorporated using recombinant expression techniques into the bromodomains of Brd4, BrdT, and BPTF to measure binding of small molecule ligands.[26] Conversely, fluorine has been incorporated into a peptide ligand to elucidate the ligand recognition and specificity of human protein disulfide isomerase using ^{19}F ligand-observe in conjunction with ^{15}N–^{1}H HSQC protein-observe NMR methods.[27] For additional reading on ^{19}F NMR as a tool for studying protein–peptide interactions, readers are directed to the review by Marsh and Suzuki.[28]

A number of reviews[23,29] have described the use of fluorine in peptide and protein chemistry but with little specific focus on fluorinated proline analogues, where this review will focus. Fluorinated prolines [(F)Pro, not limited to 4-fluoroprolines] have the potential to be exceptionally useful tools in chemical biology to modulate protein or peptide stability and improve folding kinetics. Introduction of electronegative fluorine greatly influences the kind of interactions that can be formed between proline and its natural binding partners. While organic fluorine (C–F) is a weaker hydrogen-bond acceptor than corresponding alcohols or carbonyls are,[30,31] its introduction has been found to provide new non-native interactions (see 4-fluoroprolines).[32] Fluorine is also very electronegative and, therefore, is strongly electron-withdrawing, which can weaken the hydrogen-bond accepting ability of neighboring groups and potentially weaken intermolecular interactions. However, the high electronegativity of fluorine makes it a good candidate for orthogonal multipolar or dipole interactions,[33] e.g., with backbone carbonyl groups of other proteins.[34]

CONFORMATIONAL EFFECTS

Proline contains a conformationally restricted five-membered pyrrolidine ring, which is puckered (nonplanar). Energy minima correspond to the two approximate envelope conformations with C$^{\gamma}$ *endo* and *exo* puckered[32] (Figure 1B), and their relative populations are to a significant extent dictated by the nature and the stereochemistry (*S* or *R*) of substituents at proline β (3) and γ (4) positions. 4-Hydroxyprolines (**6a** and **6b**) and 4-fluoroprolines (**7a**–**c**) are by far the most common proline analogues that have been reported (Figure 3). 4-Hydroxyproline occurs in nature[35] and can be used as a precursor for the synthesis of 4-fluoroprolines.[36] Zondlo's group performed a comprehensive study in which the 4-position substituent on hydroxyproline (Hyp) was varied to include electron-withdrawing/donating and different sized groups to compare the effect of steric versus stereoelectronic effects on the proline–peptide conformation.[37] For

DOI: 10.1021/acs.biochem.8b00787
Biochemistry 2018, 57, 6132–6143

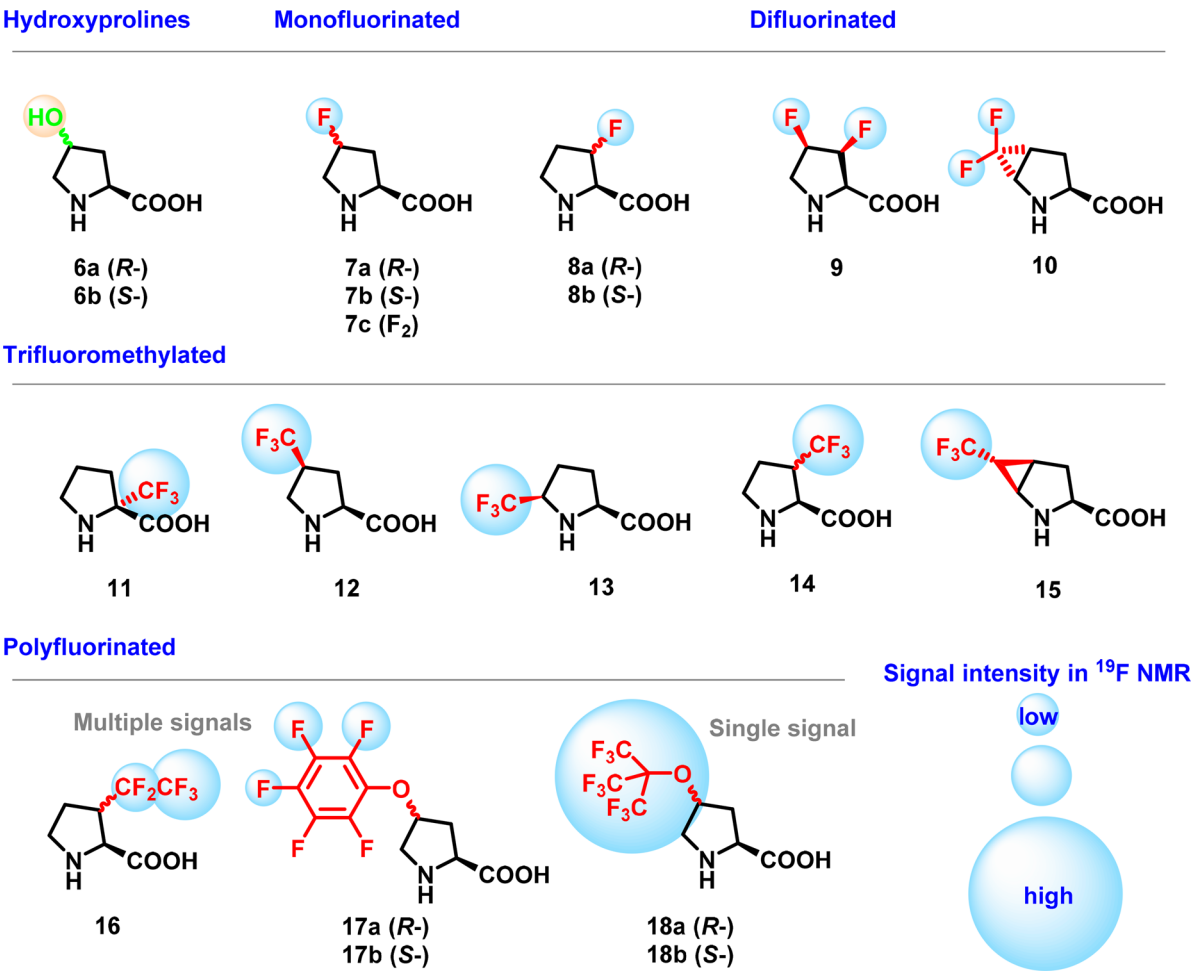


Figure 3. Various reported fluorinated proline derivatives used in peptides and proteins and an indication of their signal strength as reporters for ^{19}F NMR (not to scale).

electron-withdrawing substituents at the 4-position, *S*-stereochemistry strongly favors an *endo* pucker, while *R*-stereochemistry favors an *exo* pucker (Figure 1C). However, this is reversed for non-electron-withdrawing groups and/or large bulky substituents (Figure 1D). This demonstrated that the proline conformation is potentially tunable within synthetic peptides postsynthesis, opening up possibilities for stimulus-responsive conformational switching.

The preference for a specific ring-pucker conformation following substitution at the 3- or 4-position of the pyrrolidine ring with fluorine (or another electron-withdrawing group)[38−40] can be explained by the adoption of a *gauche* (60° torsion angle) conformation between the amide nitrogen and the *vicinal* fluorine atom in an N−C−C−Z (Z = N, O, or F) system (Figure 1C). The ring-pucker preference is, therefore, reversed by switching the stereochemistry of the electron-withdrawing group. Note that the fluorine-ammonium *gauche* effect has been thoroughly explained elsewhere.[41] In general, *exo* pucker is considered to lead to a stronger preference for the *trans*-amide rotamer and *endo* pucker a weaker *trans* preference [larger population of *cis* amide (Figure 1D)]. Studies have indicated that the *exo* ring pucker in model Ac-(F)Pro-OMe systems stabilizes the *trans* amide through an enhanced n → π^* interaction between the amide oxygen and the ester carbonyl carbon.[42] To examine the relationship, a series of conformationally locked proline analogues with a fluorinated methylene bridge (5-fluoromethanoprolines) were synthesized to mimic the "locked" *endo* and *exo* puckered prolines. This showed that, in these systems at least, substituents on proline analogues can affect the observed *trans*−*cis* preference in ways independent of puckering and n → π^* stabilizing interaction.[43] Mykhailiuk and co-workers probed the individual through-bond effect of the electron-withdrawing group at the 4-position in 2,4-methanoproline models by measuring the pyrrolidine nitrogen pK_a and *cis*−*trans* isomerization while not being confounded by ring puckering.[44] They concluded that the electron-withdrawing fluorine has a significant effect upon the kinetics but little effect upon the thermodynamics of amide-bond isomerism. For example, in 4-fluoro-2,4-methanoprolines, the barrier to rotation is reduced by ∼3 kJ/mol, leading to a rotation rate enhancement of ∼4−5-fold, as compared with that of the nonsubstituted analogue, which likely translates into the analogous proline system.

The well-explored proline substitution with fluorine (4*R*- or 4*S*-fluoroproline) affects the spontaneous isomerization of the peptidyl−prolyl bond, with both the rate and the equilibrium constants found to be affected.[45] For this reason, there is increasing motivation for the use of fluoroprolines to probe conformational, structural, and stability properties in rational protein or peptide design. Overall, the incorporation of fluorine has been driven by two main factors: (i) powerful

DOI: 10.1021/acs.biochem.8b00787
Biochemistry 2018, 57, 6132−6143

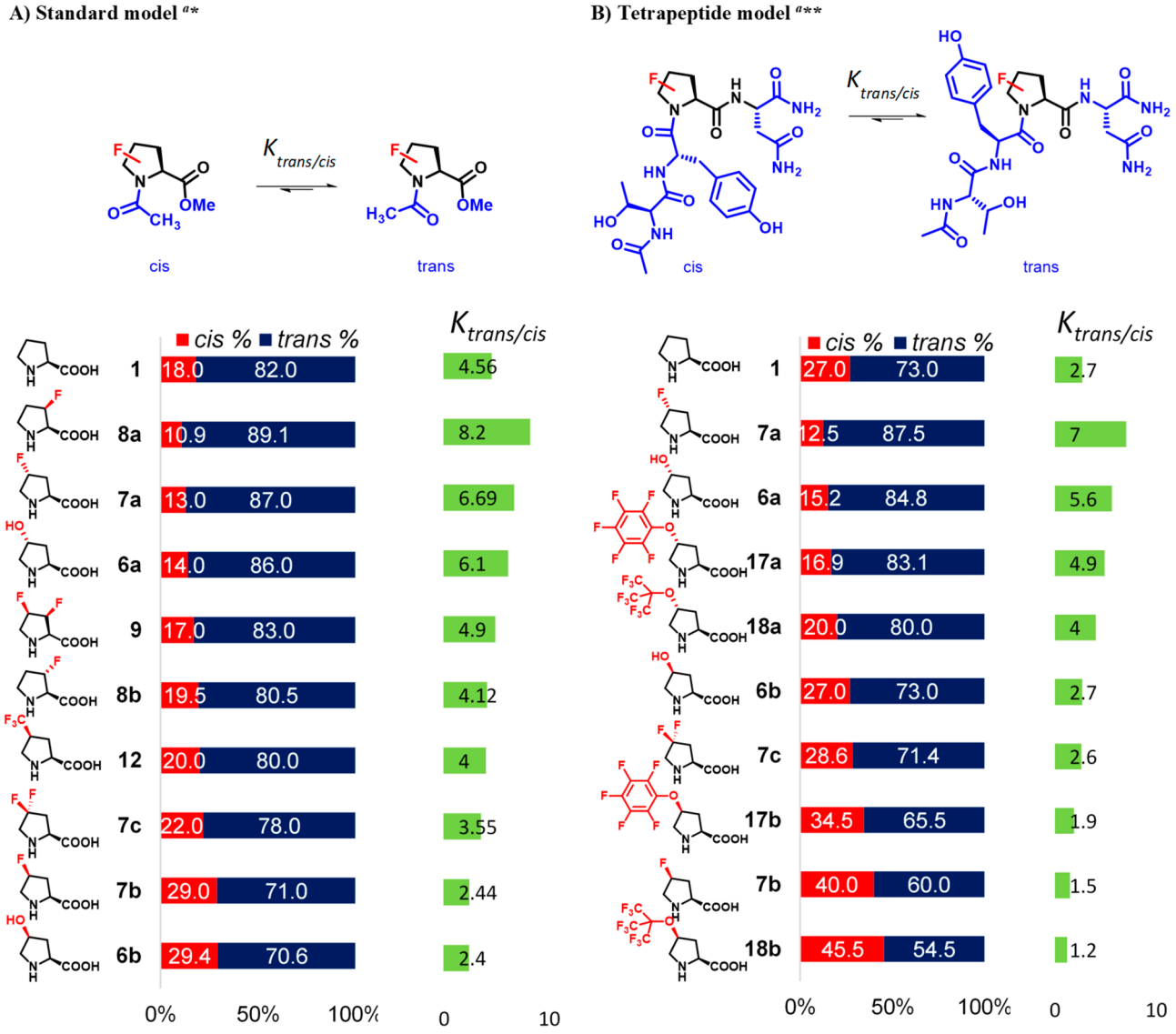


Figure 4. Graphical ranking of differential % *cis* and *trans* conformers (red/blue) depending on the nature of the fluorinated proline in (A) the standard Ac-(F)Pro-OMe* and (B) a tetrapeptide model showing similar trends but quantitative differences between a simple model and a short peptide sequence. The corresponding $K_{trans/cis}$ (ratio of the peptide with a *trans* amide bond compared to the peptide with a *cis* amide bond as determined by NMR) is also presented graphically (green). Footnote a indicates that when more than one value is reported in different publications, the values are generally in good agreement with the values shown. One asterisk indicates data obtained from refs 55 (D_2O at 37 °C), 57 (D_2O at 25 °C), 50 (D_2O at 25 °C), and 65 (D_2O at 25 °C). Two asterisks indicate data obtained in 25 mM NaCl with 5 mM phosphate (pH 4) in a 9:1 H_2O/D_2O mixture at 298 K.[37]

conformational effects of fluorine and (ii) the possibility for ^{19}F to be used as an NMR reporter for structural and stability studies. A number of fluorinated prolines have been reported, for which the conformational effects are typified by 4-fluoroprolines. The synthesis and main uses of 4-fluoroprolines have been thoroughly described recently in an excellent review by Newberry and Raines.[46]

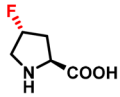
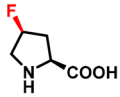
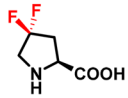
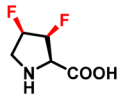
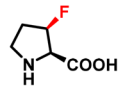
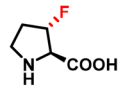
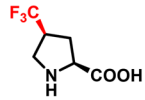
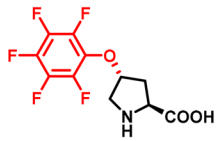
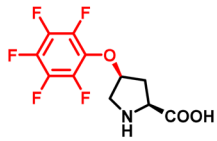
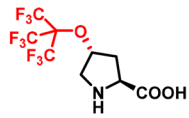
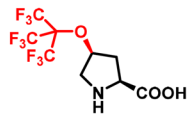
4-FLUOROPROLINES

The use of fluoroprolines as a conformational probe in chemical biology requires the artificial introduction of these through either (i) solid phase peptide synthesis or (ii) recombinant protein technologies. The first report of using synthesized fluorinated prolines in peptide or protein systems was in 1965 from Bernhard Witkop's group, incorporating tritiated *R*- and *S*-4-fluoroproline into collagen via *Escherichia coli* or guinea pig granulomata homogenates and studying enzymatic hydroxylation.[47]

The commercial availability of L-, D-, and disubstituted Fmoc-4-fluoroprolines (**7a−c**, respectively) allows for convenient laboratory synthesis of short peptides, e.g., proctolin [Arg-Tyr-Leu-(F)Pro-Thr, where (F)Pro = 4*S*-/4*R*-fluoroproline], a neurotransmitter found in motor neurons of insects and the postural motor neurons of crayfish.[48,49] However, this is generally limited to polypeptides of <40 amino acids, without relying on chemical ligation approaches. 4-Fluoroproline analogues have also been incorporated into proteins via protein total synthesis using native chemical ligation. By incorporation of 4-fluoroprolines with different *trans/cis* amide preferences, Torbeev and Hilvert showed that isomerization of Pro32 of β2-microglobulin from its native *cis* to a non-native *trans* conformation triggered misfolding and subsequent

DOI: 10.1021/acs.biochem.8b00787
Biochemistry 2018, 57, 6132−6143

Table 1. Reported Characteristic ^{19}F NMR Chemical Shifts for *trans* and *cis* Conformers in Model Systems As Indicated[a]

Fluorinated proline		Ac-(F)Pro-OMe		Ac-TY-(F)PN-NH_2	
		Trans (ppm)	*Cis (ppm)*	*Trans (ppm)*	*Cis (ppm)*
7a		-177.85 [57, A]	-177.79	-177.1 [40, D]	-177.9
7b		-173.26 [57, A]	-173.25	-173.2 [40, D]	-173.5
7c		-98.4 (pro-*S*) [57, A] -102.2 (pro-*R*)	-95.5 (pro-*S*) -104.7 (pro-*R*)	-100.6 [40, D]	-103.2, -103.9, -94.8, -95.4
9		-203.3 (F^4) [57, A] -210.4 (F^3)	-200.3 (F^4) -208.5 (F^3)	-	-
8a		-108.5 [55, B]	-106.5	-	-
8b		-97.5 [55, B]	-98.6	-	-
12		-71.0 [65, C]	-71.3	-	-
17a		-	-	-163.2 [37, D] -161.9 -156.5	-163.6 -162.1 -156.1
17b		-	-	-163.4 [37, D] -162.4 -156.3	-163.6 -162.9 -156.3
18a		-	-	-70.7 [77, D]	-70.8
18b		-	-	-70.0 [77, D]	-70.7

[a]NMR conditions: (A) D_2O at 25 °C, (B) 9:1 H_2O/D_2O mixture at 35 °C (no exact values provided in ref 55; chemical shift values in the table are estimated from raw NMR spectra), (C) D_2O at 27 °C, and (D) 5 mM phosphate buffer (pH 4) with 25 mM NaCl in a 9:1 H_2O/D_2O mixture at 23 °C.

amyloid assembly.[50] 4-Fluoroprolines have been incorporated into proteins, e.g., ubiquitin[51] and collagen,[40] where the 4*R*-fluoroproline was shown to increase the thermal stability compared to that of the wild-type protein. While there are several examples of introduction of 4-fluoroprolines biosynthetically into a protein, e.g., barstar C40A/C82A/P27A in *E. coli*, this can present many challenges.[11,52] It is important to consider the additional effects of the introduction of 4-fluoroproline into peptides or proteins for biological studies. For example, even though the 4*S*-fluoroproline proctolin insecticide peptide had the wild-type isoform peptidyl−proline bond, only the 4*R*-fluoroproline peptide showed biological activity toward *tortrix* insects, suggesting that the ring puckering can have a stronger influence on affinity than *cis*−*trans* isomerization.

The *gauche* effect is eliminated when two fluorine atoms are substituted at the proline 4-position [4,4-difluoroproline (**7c**)] as they counteract their electronic effects, leading to behavior similar to that of native unsubstituted proline, based on similar $K_{trans/cis}$ equilibrium ratios.[11,53] However, interestingly, it has been reported that this can be more structurally disruptive than the monofluorinated prolines.[17] Inserting fluorine atoms at the

DOI: 10.1021/acs.biochem.8b00787
Biochemistry 2018, 57, 6132−6143

proline 4-position was observed to reduce the double-bond character of the prolyl amide, lowering the energy barrier for *trans*−*cis* isomerization ($\Delta G^{\ddagger}_{300\ K}$) in the order Pro (88.4 kJ/mol) > 4*R*-Flp (87.4 kJ/mol) ∼ 4*S*-flp (86.3 kJ/mol) > F2Pro (83.3 kJ/mol) and, therefore, increasing the isomerization rate in the Ac-(F)Pro-OMe model system.[7] These examples show that fluorinated prolines can stabilize or destabilize peptides or proteins compared to the wild types. Overall, when 4*R*- or 4*S*-fluoroproline is used, the properties of the peptide or protein are altered compared to those with normal proline,[13] limiting the possibilities of using this as an unbiased reporter amino acid. Polyproline peptides can adopt either type I (PPI) helices with all *cis* amide bonds or type II (PPII) helices with all *trans* amide bonds. Therefore, the inclusion of conformationally biased fluoroproline has been used to stabilize these specific conformations. For example, a C-terminal $(4R\text{-FPro})_3$ motif stabilizes PPII helices, whereas a $(4S\text{-FPro})_3$ triplet destabilizes PPII helices.[54]

3-FLUOROPROLINES

Several groups have sought to design alternative fluorinated prolines as native proline replacements. These developments have been largely drawn from the desire to use fluorinated probes to study protein structure and dynamics. A fluorine substituent at the proline ring 3-position (**8a** and **8b**) is also able to adopt a *gauche* conformation with respect to the ring nitrogen, which provides control over ring puckering akin to that of the 4-fluoroprolines. The *trans* and *cis* conformers of the Ac-(3-F)Pro-OMe model peptides produce distinct ^{19}F resonances separated by 0.8 ppm in 3*S*-fluoroproline and ∼2 ppm for 3*R*-fluoroproline.[55] The $K_{trans/cis}$ equilibrium constants and relative populations of *cis* and *trans* amide for each are summarized in Figure 4. The measured $K_{trans/cis}$ for 3*S*-fluoroproline is comparable with that of proline in Ac-Pro-OMe at pH 7.4 and 37 °C. Interestingly, a higher entropic barrier for both *cis*−*trans* and *trans*−*cis* isomerization was measured by Eyring analysis for 3*R*-fluoroproline compared to the 3*S* form, indicating that a *syn*-configured fluorine atom may sterically hinder rotation around the C−N bond. X-ray diffraction studies demonstrated that the 3-fluoroproline series display the same C^{γ} ring-pucker preferences as the corresponding 4-fluoroprolines; e.g., both Ac-(3*R*)-(F)Pro-OMe and Ac-(4*R*)-(F)Pro-OMe favor *exo*, both with high *trans*:*cis* amide preferences (Figure 4).[52] This was supported by density functional theory calculations.[56] Only a relatively small impact upon $K_{trans/cis}$ is observed in the 3-fluoroprolines as compared with 4-fluoroprolines, making them useful tools that have been biosynthetically incorporated into an elastin peptide.[52] To date, there have been no reports of utilizing 3-fluoroprolines in protein systems.

3,4-DIFLUOROPROLINES

One of the main challenges remaining in this field is obtaining non-invasive probes, nonproteinogenic fluorinated amino acids that behave like their cognate counterpart, i.e., do not interfere with protein folding kinetics. While the 4,4-difluoroproline exhibits similar energy *exo* and *endo* puckers, the geminal CF_2 with diastereotopic fluorine atoms may complicate or even distort the ^{19}F NMR spectrum. A different approach to avoid conformational distortion by the introduction of fluorine at, e.g., the 4-position, thereby maintaining "natural" proline ring pucker and amide conformations, is to introduce a second fluorine atom at the vicinal 3-position (**9**) as recently reported by Linclau and co-workers (Figure 4).[57] The main advantage of this approach is the relative improvement in the NMR spectrum, which exhibits smaller $^3J_{F-F}$ coupling constants and different chemical shifts compared with those of geminal CF_2 (Table 1). This provides the opportunity for doubly labeling of proteins and peptides, e.g., polyprolines using different native-like fluorinated prolines with different characteristic ^{19}F reporter signals.

TRIFLUOROMETHYLPROLINES

The trifluoromethylation of proline is attractive as it provides a higher signal-to-noise ratio for use in ^{19}F NMR studies. Introduction of an α-trifluoromethyl substituent at proline (**11**) can, however, adversely affect the peptide backbone conformation and electronics, making α-trifluoromethylprolines difficult substrates for standard solid phase peptide synthesis, because of a lower nucleophilicity at the amino group,[58] and also altering the folding compared to that of proline.[59] Despite the difficulties, Brigaud and co-workers utilized this building block to prepare a tripeptide α-TfmPro-Leu-Gly, MIF-1 analogue for evaluation as an antinociceptive agent with enhanced analgesic activity versus the parent sequence.[60] However, it seems that incorporation into a protein system would present significant challenges to overcome.

The direct trifluoromethyl analogue of 4-fluoroproline, 4-trifluoromethylproline [TfmPro (**12**)], has been synthesized via several routes,[61−64] including by Ulrich and co-workers.[65] Unlike other fluorinated prolines, Ulrich also showed TfmPro, in the Ac-TfmPro-OMe model system, to be significantly more "native proline-like" with regard to (i) its backbone conformational propensities, (ii) its $K_{trans/cis}$ equilibrium constant (4.0 for TfmPro and 4.8 for Pro at 300 K), (iii) its activation energy for *cis* → *trans* rotation (81.8 kJ/mol for TfmPro and 84.5 kJ/mol for Pro), and (iv) its incorporation into cyclic peptide gramicidin S, which showed virtually no structural perturbation by circular dichroism as compared with the natural product. The synthesis of other trifluoromethyl-containing proline derivatives, e.g., 5-CF_3 (**13**), has been reported.[66,67] Very recently, the synthesis of novel *R*- and *S*-3-CF_3 (**14a** and **14b**, respectively) and 3-CF_2CF_3 (**16a** and **16b**) prolines were reported.[68] However, these have not been incorporated into proteins at the time of submission.

FLUORINATED METHANOPROLINES

To address a perceived lack of conformational rigidity of 4-trifluoromethylproline, Ulrich et al. designed and synthesized trifluoromethyl-3,4- and -4,5-methanoproline.[69] The use of these in peptide synthesis may be limited by a low total yield in the reported synthesis of 3,4-systems and the high acid sensitivity of some of the 4,5-epimers. However, trifluoromethyl-4,5-methanoproline **15** was incorporated specifically into polyproline II (PPII), forming the cell-penetrating "sweet arrow peptide" (VRLPPPVRLP**X**PVRLPPP, where **X** = trifluoromethyl-4,5-methanoproline). The resulting peptide was found to be compatible with solid phase peptide synthesis, to be pH-stable and racemization-resistant during the synthesis, and to stabilize the PPII helical structure. To study the reasons for the excellent stabilizing properties of trifluoromethyl-4,5-methanoproline (**15**), the more electron-withdrawing *trans*-4,5-difluoromethanoproline (**10**) was synthe-

DOI: 10.1021/acs.biochem.8b00787
Biochemistry 2018, 57, 6132−6143

sized.[70] However, this was found to be unstable upon N-deprotection. Despite being itself unsuitable for SPPS, it was shown to be possible to incorporate this into peptides using a rather long semisolution phase route, as a stable label for ^{19}F NMR structure analysis.[71] Unfortunately, for each of the methanoproline systems, no data have yet been reported to indicate *trans*:*cis* preferences or equilibrium constants.

PENTAFLUOROPHENYL-HYDROXYPROLINE

Pentafluorophenyl-hydroxyproline[37] (**17a** and **17b**) contains five fluorine atoms. However, this produces a ^{19}F NMR spectrum that is more complex than those of more symmetrical fluorinated groups. While these strongly favor a *trans* conformation, so far such groups have not yet been exploited to study protein systems. However, the corresponding perfluorinated reagents have the distinct advantage that they are particularly suited to direct chemoselective "tagging" of synthetic and possibly native peptide/protein side chains rather than using SPPS, with similar functionalities having been incorporated by the groups of Cobb and Pentelute into peptides at side chains other than hydroxyproline, e.g., cysteine and lysine.[72−76] Therefore, they may be excellent candidates for "proline editing" on a protein.

PERFLUORO-*TERT*-BUTYL-HYDROXYPROLINE

Perfluoro-*tert*-butyl-hydroxyproline (**18a** and **18b**) was designed to further enhance the signal-to-noise ratio in ^{19}F NMR studies.[77] Initially, the perfluoro-*tert*-butyl group was directly added to hydroxyproline as part of a peptide, via a so-called "proline editing" approach.[34] In the model peptide system Ac-TYXN-NH_2 (X is either 4*R*-perfluoro-*tert*-butyl-hydroxyproline or 4*S*-perfluoro-*tert*-butyl-hydroxyproline), 4*S*-perfluoro-*tert*-butyl-hydroxyproline exhibited an impressive *trans*:*cis* conformational preference ($K_{trans/cis}$ = 1.2) compared to that of the corresponding proline peptide ($K_{trans/cis}$ = 2.7). Subsequently, the Boc- and Fmoc-(2*S*,4*R*)- and -(2*S*,4*S*)-perfluoro-*tert*-butyl hydroxyprolines were synthesized and utilized successfully in SPPS in the preparation of an alanine-rich α-helical peptide and a proline-rich sequence. The coupling reaction of these fluorinated prolines was slow compared with that with standard amino acids and was attributed to steric hindrance associated with the perfluoro-*tert*-butyl group. The 4*R*-$OC(CF_3)_3$ derivative was found to promote α-helicity more than proline itself, whereas 4*S*-$OC(CF_3)_3$ discouraged helix formation. Both were found to lower the propensity to form a polyproline helix. However, increased steric demand, significant modulation of pyrrolidine ring electronics, and high lipophilicity may significantly affect their utility as probes for protein structure and dynamics.

A COMPARISON OF CONFORMATIONAL EFFECTS

It is clear that different fluorinated prolines offer differing degrees of *cis* conformer stabilization and provide a useful tool kit for probing the conformational and stability properties of a protein or peptide. There has previously been no collective assessment of the many published effects of wider proline fluorination on the relative *trans*:*cis* ratio. Using the reported $K_{trans/cis}$ values, it was possible for us to calculate the expected % *trans* and % *cis* to provide a valuable comparative reference resource. These data have been collated from a number of references that agree (in some cases cite each other) in Figure 4. Helpfully, most groups have chosen to study the Ac-(F)Pro-OMe model peptide. However, some of the data shown were derived from an alternative Ac-TYPN-NH_2 model peptide. Where comparison is possible, there is general agreement between the two systems, with *S*-stereochemistry fluorinated groups favoring a higher proportion of the *cis* conformation than their enantiomeric partner in both model systems. However, the same fluorinated proline promotes the *cis* conformer to different extents in the different systems. For example, 4*S*-fluoroproline in Ac-(F)Pro-OMe exists at ∼29% *cis*, whereas in Ac-TY(F)PN-NH_2, it promotes a significantly larger population of 40% *cis* due to the preceding amino acid, with aromatic residues favoring the *cis* isomer slightly in most cases.[78] More widely, it is known that local CH(Pro)···π interactions can stabilize the *cis*-peptidyl−prolyl-bond population of a Xaa-Pro when (i) Xaa is aromatic[79] or (ii) Xaa is Pro, Gly or aromatic (Aro) in Xaa-Pro-Aro motifs[80,81] in nonfluorinated systems. Other notable examples within Ac-TYPN-NH_2 include 4-perfluoro-*tert*-butyl- and 4-pentafluorophenyl-hydroxyprolines, which display significant $K_{trans/cis}$ preferences for *cis* amides in the *S*-configuration, while the *R*-enantiomers seem to behave more like the parent hydroxyproline. The equilibrium constants derived for each fluorinated proline may not be directly comparable as they are measured in different systems and conditions; however, this provides a rough guide to the relative ranking of each.

POTENTIAL APPLICATIONS OF FLUORINATED PROLINES AND ^{19}F NMR TO STUDY DYNAMIC BIOLOGICAL PROCESSES

Peptides and proteins are large biomolecules that are difficult to study by the conventional chemical methods of ^{1}H and ^{13}C NMR. For example, protein dynamics (e.g., isomerization and folding) can be difficult to study as the substrate and product have the same chemical constitution and exist in dynamic equilibrium. However, the ratio between *cis* and *trans* proline species was originally probed using a 4-position ^{13}C-labeled proline in collagen. While the clear advantage of ^{13}C labeling is the lack of structural and conformational perturbation compared with that of fluorinated prolines (Figure 4), the significant drawback of this approach is the requirement for an expensive NMR cryoprobe to enhance the low sensitivity of ^{13}C nuclei.[82] ^{19}F NMR has a significant advantage in that it has a very wide chemical shift range and characteristic signals for *cis*/*trans* isomers. While there are some small discrepancies between reported ^{19}F NMR chemical shift values, we have provided referenced examples in Table 1. Advantageously, the small number of atoms present (e.g., one ^{19}F NMR signal for proline compared with seven for ^{1}H) means the ^{19}F spectrum of a protein containing multiple fluorinated amino acids can be well-resolved and readily interpreted.[83] The frequency of natural occurrence of fluorine in biomolecules and in nature is very low,[84] thus making fluorine a perfect reporter atom for NMR investigations. ^{19}F NMR studies of proteins with their ligand provide valuable insight into ligand−target interactions and dynamics.[85−88] In some cases, considerable signal overlap makes measuring *cis*−*trans* inversion kinetics challenging in 4-fluoroproline systems. However, the newer derivatives (Figure 3) discussed above have provided alternative ways to distinguish conformers. Fluorinated amino acids, and by proxy peptides and proteins, allow sensitive detection in complex media, biological samples, and even microorganisms. However, the sensitivity of the detection can be limited by the

DOI: 10.1021/acs.biochem.8b00787
Biochemistry 2018, 57, 6132−6143

signal intensity of the F-containing analyte. Moreover, in a biological system, only very low concentrations may be achievable. As such, the use of proline derivatives containing several symmetrical fluorine nuclei, e.g., trifluoromethylprolines (**11**–**14**), methanoprolines (**15**) (three ^{19}F nuclei), or perfluoro-*tert*-butyl-hydroxyprolines (**18**) [nine ^{19}F nuclei (Figure 5)], is exciting as they offer enhanced sensitivity.

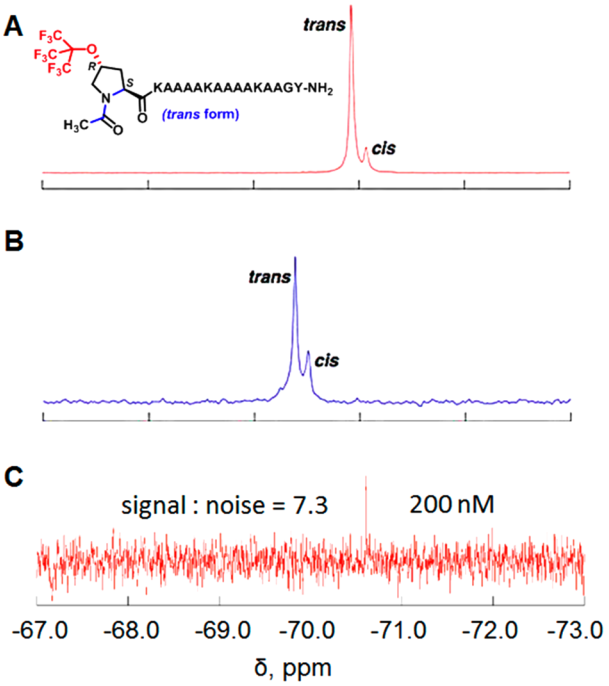


Figure 5. ^{19}F NMR spectra of Ac-XKAAAAKAAAAKAAGY-NH_2 peptides with (A) X = 2*S*,4*R*-Hyp(C_4F_9) (structure shown) and (B) X = 2*S*,4*S*-Hyp(C_4F_9). The peaks indicate the *trans* (major) vs *cis* (minor) Ac–Pro amide bonds. Experiments were performed on a 600 MHz (^{19}F, 564.5 MHz) NMR spectrometer at 298 K in a 90% H_2O/10% D_2O mixture with 5 mM phosphate buffer (pH 4) and 25 mM NaCl. (C) ^{19}F NMR spectrum of 200 nM Ac-(2*S*,4*R*-Hyp(C_4F_9))-KAAAAKAAAAKAAGY-NH_2 in D_2O with 5 mM phosphate buffer (pH 4) and 25 mM NaCl (128 scans with a 7 ppm sweep width). Reproduced with permission from ref 77. Copyright 2016 American Chemical Society.

Zondlo showed that peptides containing **18** could be used to visualize several different conformations and were detectable at a low concentration of 200 nM (Figure 5), demonstrating the utility of **18** as a biological probe.[77] This can be particularly useful if coupled with the use of stopped-flow ^{19}F NMR techniques that will allow sensitive and time-dependent measurement of folding and unfolding rates of proteins containing several fluorinated proline residues.[89]

While it should be considered that introduction of a fluorine can interfere with the natural characteristics of a peptide/protein (through, e.g., stabilized conformation, altered electronegativity, or increased hydrophobicity),[45] fluorinated prolines favoring one isoform over the other are useful tools for measuring kinetics of isomerization and, in turn, by slowing isomerization, can be a useful probe for enzyme mechanism and function.

As such, the combination of conformational and kinetic control and provision of a sensitive NMR reporter group gives fluorinated proline derivatives exceptional potential as tools for measuring peptide and protein conformation. Several examples mentioned previously have used ^{19}F NMR for easier and more accurate determination of the populations of *cis* and *trans* conformations in fluorinated proline models, and this should be applicable as a tool to study dynamic biological processes. For example, variable-temperature inversion transfer solid state ^{19}F NMR was used[55] to determine the kinetic and thermodynamic parameters of *cis*–*trans* isomerization of 3-fluoroproline models (Ac-3F-Pro-OMe). Another solid state ^{19}F NMR study investigated the alignment, conformation, and mobility of the peptide gramicidin S containing *trans*-4,5-difluoromethanoproline[71] (**10**) and 4-trifluoromethylproline[65] (**12**) in lipid membranes. The former was the first report to employ a two-spin ^{19}F system for structural studies of peptides. In Torbeev and Hilvert's work incorporating 4,4-difluoroproline (**7c**) at proline 32 of β2-macroglobulin, solution state ^{19}F NMR studies (Figure 6) showed that the F_2Pro32 amide

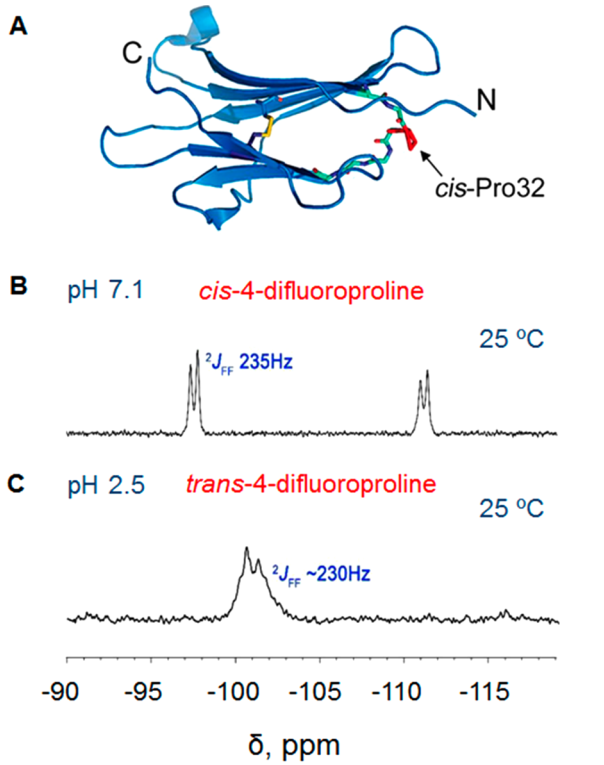


Figure 6. (A) Native structure of β2m showing the BC loop (cyan) containing *cis*-Pro32 (red) (based on Protein Data Bank entry 2YXF). (B) ^{19}F NMR (565 MHz) spectrum of folded [F_2Pro32]β2m protein (0.57 mM). (C) Spectrum recorded for [F_2Pro32]β2m protein (0.4 mM) under denaturing conditions [3.8 mM HCl with 100 mM NaCl in a 90% H_2O/10% D_2O mixture (pH 2.5)]. Reproduced with permission from ref 50. Copyright 2013 National Academy of Sciences.

adopts a *cis* conformation in the folded protein but a *trans* conformation when denatured.[50] This demonstrated that the protein conformation can be observed and measured using ^{19}F NMR. To the best of our knowledge, there have been few reports of using solution state ^{19}F NMR to study the folding and dynamics of a functional protein system containing fluorinated proline derivatives, in particular, upon interaction with a ligand or following post-translational modification-promoted conformational change. This is despite the development of tools for this specific purpose. It is unclear why the field has yet to progress beyond this point, given that this is

DOI: 10.1021/acs.biochem.8b00787
Biochemistry 2018, 57, 6132–6143

conceptually achievable as other researchers have performed similar studies using, e.g., ^{19}F-labeled phenylalanine, in which the enzyme-catalyzed *cis*-to-*trans* isomerization rate of [4-fluoro-Phe]bradykinin was measured by magnetization transfer ^{19}F NMR.[90] Indeed, it is likely that fluorinated prolines would provide a more accurate reporter of *cis*−*trans* interchange because of their proximity to the site of isomerization.

CONCLUSIONS AND FUTURE DIRECTIONS

An increasing number of fluorinated proline analogues have been reported, principally for the purpose of controlling conformation via changing the ring pucker of the proline and/or the amide-bond conformation. Some even behave as more proline-like by not interfering with natural *cis*−*trans* isomerism. These now provide the as yet underexplored opportunity to probe *cis*−*trans* isomerization and its effect in protein systems in real time using ^{19}F NMR. However, there is still a lack of proline analogues that confer the large *cis* populations that are required by protein chemists while simultaneously providing a useful reporter group. The future should see efforts directed toward preparing more conformationally biased proline analogues with the goal of approaching 100% *cis*, perhaps including fluorinated derivatives of 5,5-dimethylprolines (**4**) and 5-*tert*-butylprolines (**5**), which will also provide a high-sensitivity reporter. These will provide a comprehensive tool kit with a wide range of *cis*/*trans* preferences that can be inserted into proteins and peptides for residue-specific time-resolved probing of conformation, folding, and stability even at extremes of *trans*:*cis* populations.

Other future developments will likely use novel fluorinated prolines with distinct conformational properties to understand new protein systems or for the *de novo* design of proteins with new functions. This may be achievable through increasingly sophisticated recombinant technologies or through protein total chemical synthesis. However, while the *trans*:*cis* ratios provided by model systems are informative from the perspective of being able to rank the effects of various proline modifications, the effects are likely to be more subtle in a protein in which the *gauche* effect is overpowered by the many hundreds or thousands of interactions that promote the folding of a protein into a tertiary structure. Furthermore, much of this work using simple model systems has neglected to fully rationalize the effect of neighboring amino acids upon the observed *trans*:*cis* ratio. Therefore, a more thorough investigation of how amino acid X in an "X−(F)Pro" system affects the observed conformational properties is required as a small step toward understanding the longer distance effects in proteins.

Despite the preparation of highly fluorinated prolines, the rather low sensitivity of ^{19}F NMR remains a problem for working in biological systems. ^{19}F NMR signal enhancement methods such as biocompatible hyperpolarization transfer techniques,[91] which can be used to enhance the signal-to-noise ratio in NMR and magnetic resonance imaging (MRI) experiments by several orders of magnitude, can be employed to obtain a higher sensitivity at concentrations lower than the current state of the art at 200 nM. This will be required for fluorinated proline-containing peptides and proteins to be used in therapeutic drug monitoring and diagnostic medical imaging, such as MRI. It is known from ^{18}F positron emission tomography imaging experiments that *R*- and *S*-4-fluoroprolines are differentially taken up through the blood−brain barrier.[92] It is also possible that different conformers of therapeutic proteins and peptides will have differential adsorption and distribution characteristics. ^{19}F MRI will, therefore, become a useful tool for studying the behavior of fluorinated conformationally biased proteins and peptides in a complex living system.

Overall, the value of being able to use ^{19}F NMR to measure dynamics and protein properties in real time is yet to be fully realized and undoubtedly holds much promise. This is due to the natural scarcity of organic fluorine as well as its wide chemical shift range and high sensitivity to local environments, making analysis relatively simple and achievable in real time. Moreover, newer NMR technologies will continue to enable more precise measurements and require lower fluorinated protein concentrations, increasing the future utility of these tools in protein and peptide chemical biology.

AUTHOR INFORMATION

Corresponding Author

*E-mail: c.r.coxon@ljmu.ac.uk. Telephone: +44 151 904 6344.

ORCID

Christopher R. Coxon: 0000-0002-3375-3901

Funding

The authors thank Liverpool John Moores University for a generous fully funded Ph.D. scholarship for S.J.M.V. and the Engineering and Physical Sciences Council (EPSRC), UK (EP/R020299/1), for a postdoctoral fellowship for P.M.K.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) MacArthur, M. W., and Thornton, J. M. (1991) Influence of proline residues on protein conformation. *J. Mol. Biol. 218*, 397−412.

(2) DeRider, M. L., Wilkens, S. J., Waddell, M. J., Bretscher, L. E., Weinhold, F., Raines, R. T., and Markley, J. L. (2002) Collagen Stability: Insights from NMR Spectroscopic and Hybrid Density Functional Computational Investigations of the Effect of Electronegative Substituents on Prolyl Ring Conformations. *J. Am. Chem. Soc. 124*, 2497−2505.

(3) Kitamoto, T., Ozawa, T., Abe, M., Marubayashi, S., and Yamazaki, T. (2008) Incorporation of fluoroprolines to proctolin: Study on the effect of a fluorine atom toward peptidic conformation. *J. Fluorine Chem. 129*, 286−293.

(4) Cremer, D., and Pople, J. A. (1975) General definition of ring puckering coordinates. *J. Am. Chem. Soc. 97*, 1354−1358.

(5) Chang, D.-K., Cheng, S.-F., Trivedi, V. D., and Lin, K.-L. (1999) Proline Affects Oligomerization of a Coiled Coil by Inducing a Kink in a Long Helix. *J. Struct. Biol. 128*, 270−279.

(6) Tolomelli, A., Ferrazzano, L., and Amoroso, R. (2017) 4-Substituted Prolines: Useful Reagents in Enantioselective Synthesis and Conformational Restraints in the Design of Bioactive Peptidomimetics. *Aldrichimica Acta 50*, 3−15.

(7) Renner, C., Alefelder, S., Bae, J. H., Budisa, N., Huber, R., and Moroder, L. (2001) Fluoroprolines as Tools for Protein Design and Engineering. *Angew. Chem., Int. Ed. 40*, 923−925.

(8) Brandl, C. J., and Deber, C. M. (1986) Hypothesis about the function of membrane-buried proline residues in transport proteins. *Proc. Natl. Acad. Sci. U. S. A. 83*, 917−921.

(9) Barlow, D. J., and Thornton, J. M. (1988) Helix geometry in proteins. *J. Mol. Biol. 201*, 601−619.

(10) Dugave, C., and Demange, L. (2003) Cis−Trans Isomerization of Organic Molecules and Biomolecules: Implications and Applications. *Chem. Rev. 103*, 2475−2532.

(11) Wedemeyer, W. J., Welker, E., and Scheraga, H. A. (2002) Proline cis-trans isomerization and protein folding. *Biochemistry 41*, 14637−14644.

DOI: 10.1021/acs.biochem.8b00787
Biochemistry 2018, 57, 6132−6143

(12) Fischer, G., and Schmid, F. X. (1990) The Mechanism of Protein Folding. Implications of in Vitro Refolding Models for de Novo Protein Folding and Translocation in the Cell. *Biochemistry 29*, 2205−2212.
(13) Schmid, F. X., Mayr, L. M., Mücke, M., and Schönbrunner, E. R. (1993) Prolyl isomerases: role in protein folding. *Adv. Protein Chem. 44*, 25−66.
(14) Kern, D., Schutkowski, M., and Drakenberg, T. (1997) Rotational barriers of cis/trans isomerization of proline analogues and their catalysis by cyclophilin. *J. Am. Chem. Soc. 119*, 8403−8408.
(15) Halab, L., and Lubell, W. D. (1999) Use of steric interactions to control peptide turn geometry. Synthesis of type VI β-turn mimics with 5-tert-butylproline. *J. Org. Chem. 64*, 3312−3321.
(16) An, S. S. A., Lester, C. C., Peng, J. L., Li, Y. J., Rothwarf, D. M., Welker, E., Thannhauser, T. W., Zhang, L. S., Tam, J. P., and Scheraga, H. A. (1999) Retention of the cis proline conformation in tripeptide fragments of bovine pancreatic ribonuclease A containing a non-natural proline analogue, 5,5-dimethylproline. *J. Am. Chem. Soc. 121*, 11558−11566.
(17) Lummis, S. C. R., Beene, D. L., Lee, L. W., Lester, H. A., Broadhurst, R. W., and Dougherty, D. A. (2005) Cis-trans isomerization at a proline opens the pore of a neurotransmitter-gated ion channel. *Nature 438*, 248−252.
(18) Purser, S., Moore, P. R., Swallow, S., Gouverneur, V. V., Sciberras, D., Reines, S. A., Petty, K. J., Ögren, M., Antoni, G., Långström, B., Eskola, O., Scheinin, M., Solin, O., Majumdar, A. K., Constanzer, M. L., Battisti, W. P., Bradstreet, T. E., Gargano, C., Hietala, J., Watt, A. P., and Castro, J. L. (2008) Fluorine in medicinal chemistry. *Chem. Soc. Rev. 37*, 320−330.
(19) Zhou, Y., Wang, J., Gu, Z., Wang, S., Zhu, W., Aceña, J. L., Soloshonok, V. A., Izawa, K., and Liu, H. (2016) Next Generation of Fluorine-Containing Pharmaceuticals, Compounds Currently in Phase II−III Clinical Trials of Major Pharmaceutical Companies: New Structural Trends and Therapeutic Areas. *Chem. Rev. 116*, 422−518.
(20) Park, B. K., Kitteringham, N. R., and O'Neill, P. M. (2001) Metabolism of fluorine-containing drugs. *Annu. Rev. Pharmacol. Toxicol. 41*, 443−470.
(21) Shah, P., and Westwell, A. D. (2007) The role of fluorine in medicinal chemistry. *J. Enzyme Inhib. Med. Chem. 22*, 527−540.
(22) Gillis, E. P., Eastman, K. J., Hill, M. D., Donnelly, D. J., and Meanwell, N. A. (2015) Applications of Fluorine in Medicinal Chemistry. *J. Med. Chem. 58*, 8315−8359.
(23) Jäckel, C., and Koksch, B. (2005) Fluorine in peptide design and protein engineering. *Eur. J. Org. Chem. 2005*, 4483−4503.
(24) Salwiczek, M., Nyakatura, E. K., Gerling, U. I. M., Ye, S., and Koksch, B. (2012) Fluorinated amino acids: compatibility with native protein structures and effects on protein-protein interactions. *Chem. Soc. Rev. 41*, 2135−2171.
(25) Qiu, X.-L., and Qing, F.-L. (2011) Recent Advances in the Synthesis of Fluorinated Amino Acids. *Eur. J. Org. Chem. 2011*, 3261−3278.
(26) Mishra, N. K., Urick, A. K., Ember, S. W. J., Schonbrunn, E., and Pomerantz, W. C. (2014) Fluorinated aromatic amino acids are sensitive ^{19}F NMR probes for bromodomain-ligand interactions. *ACS Chem. Biol. 9*, 2755−2760.
(27) Richards, K. L., Rowe, M. L., Hudson, P. B., Williamson, R. A., and Howard, M. J. (2016) Combined ligand-observe ^{19}F and protein-observe 15N,1H-HSQC NMR suggests phenylalanine as the key Δ-somatostatin residue recognized by human protein disulfide isomerase. *Sci. Rep. 6*, 19518.
(28) Marsh, E. N. G., and Suzuki, Y. (2014) Using ^{19}F NMR to Probe Biological Interactions of Proteins and Peptides. *ACS Chem. Biol. 9*, 1242−1250.
(29) Yoder, N. C., and Kumar, K. (2002) Fluorinated amino acids in protein design and engineering. *Chem. Soc. Rev. 31*, 335−341.
(30) Shimoni, L., and Glusker, J. P. (1994) The geometry of intermolecular interactions in some crystalline fluorine-containing organic compounds. *Struct. Chem. 5*, 383−397.
(31) Dunitz, J. D., and Taylor, R. (1997) Organic Fluorine Hardly Ever Accepts Hydrogen Bonds. *Chem. - Eur. J. 3*, 89−98.
(32) Holzberger, B., Obeid, S., Welte, W., Diederichs, K., and Marx, A. (2012) Structural insights into the potential of 4-fluoroproline to modulate biophysical properties of proteins. *Chem. Sci. 3*, 2924−2931.
(33) Bissantz, C., Kuhn, B., and Stahl, M. (2010) A Medicinal Chemist's Guide to Molecular Interactions. *J. Med. Chem. 53*, 5061−5084.
(34) Pollock, J., Borkin, D., Lund, G., Purohit, T., Dyguda-Kazimierowicz, E., Grembecka, J., and Cierpicki, T. (2015) Rational Design of Orthogonal Multipolar Interactions with Fluorine in Protein-Ligand Complexes. *J. Med. Chem. 58*, 7465−7474.
(35) Krane, S. M. (2008) The importance of proline residues in the structure, stability and susceptibility to proteolytic degradation of collagens. *Amino Acids 35*, 703−710.
(36) Chorghade, M. S., Mohapatra, D. K., Sahoo, G., Gurjar, M. K., Mandlecha, M. V., Bhoite, N., Moghe, S., and Raines, R. T. (2008) Practical syntheses of 4-fluoroprolines. *J. Fluorine Chem. 129*, 781−784.
(37) Pandey, A. K., Naduthambi, D., Thomas, K. M., and Zondlo, N. J. (2013) Proline editing: A general and practical approach to the synthesis of functionally and structurally diverse peptides. Analysis of steric versus stereoelectronic effects of 4-substituted prolines on conformation within peptides. *J. Am. Chem. Soc. 135*, 4333−4363.
(38) Sonntag, L.-S., Schweizer, S., Ochsenfeld, C., and Wennemers, H. (2006) The "Azido Gauche Effect" Implications for the Conformation of Azidoprolines. *J. Am. Chem. Soc. 128*, 14697−14703.
(39) Hentzen, N. B., Smeenk, L. E. J., Witek, J., Riniker, S., and Wennemers, H. (2017) Cross-Linked Collagen Triple Helices by Oxime Ligation. *J. Am. Chem. Soc. 139*, 12815−12820.
(40) Bretscher, L. E., Jenkins, C. L., Taylor, K. M., DeRider, M. L., and Raines, R. T. (2001) Conformational stability of collagen relies on a stereoelectronic effect. *J. Am. Chem. Soc. 123*, 777−778.
(41) Briggs, C. R. S., Slawin, A. M. Z., Allen, M. J., O'Hagan, D., Goeta, A. E., Tozer, D. J., and Howard, J. A. K. (2004) The observation of a large gauche preference when 2-fluoroethylamine and 2-fluoroethanol become protonated. *Org. Biomol. Chem. 2*, 732−740.
(42) Newberry, R. W., and Raines, R. T. (2017) The n→π* Interaction. *Acc. Chem. Res. 50*, 1838−1846.
(43) Krow, G. R., Shoulders, M. D., Edupuganti, R., Gandla, D., Yu, F., Sonnet, P. E., Sender, M., Choudhary, A., Debrosse, C., Ross, C. W., Carroll, P., and Raines, R. T. (2012) Synthesis of 5-fluoro-and 5-hydroxymethanoprolines via lithiation of N-BOC-methanopyrrolidines. Constrained C γ-exo and C γ-endo Flp and Hyp conformer mimics. *J. Org. Chem. 77*, 5331−5344.
(44) Mykhailiuk, P. K., Kubyshkin, V., Bach, T., and Budisa, N. (2017) Peptidyl-Prolyl Model Study: How Does the Electronic Effect Influence the Amide Bond Conformation? *J. Org. Chem. 82*, 8831−8841.
(45) Golbik, R., Yu, C., Weyher-Stingl, E., Huber, R., Moroder, L., Budisa, N., and Schiene-Fischer, C. (2005) Peptidyl Prolyl cis/trans-Isomerases: Comparative Reactivities of Cyclophilins, FK506-Binding Proteins, and Parvulins with Fluorinated Oligopeptide and Protein Substrates. *Biochemistry 44*, 16026−16034.
(46) Newberry, R. W., and Raines, R. T. (2016) 4-Fluoroprolines: Conformational Analysis and Effects on the Stability and Folding of Peptides and Proteins. *Top. Heterocycl. Chem. 48*, 1−25.
(47) Gottlieb, A. A., Fujita, Y., Udenfriend, S., and Witkop, B. (1965) Incorporation of cis- and trans-4-Fluoro-L-prolines into Proteins and Hydroxylation of the trans Isomer During Collagen Biosynthesis. *Biochemistry 4*, 2507−2513.
(48) Anderson, M. S., Halpern, M. E., Keshishian, H., Carlson, J., Johansen, J., Kankel, D., Mcginnis, W., and Wyman, R. (1988) Identification of the Neuropeptide Transmitter Proctolin in Drosophila Larvae: Characterization of Muscle Fiber-Specific Neuromuscular Endings. *J. Neurosci. 8*, 242−255.
(49) Kitamoto, T., Ozawa, T., Abe, M., Marubayashi, S., and Yamazaki, T. (2008) Incorporation of fluoroprolines to proctolin:

DOI: 10.1021/acs.biochem.8b00787
Biochemistry 2018, 57, 6132−6143

Study on the effect of a fluorine atom toward peptidic conformation. *J. Fluorine Chem. 129*, 286−293.
(50) Torbeev, V. Y., and Hilvert, D. (2013) Both the cis-trans equilibrium and isomerization dynamics of a single proline amide modulate 2-microglobulin amyloid assembly. *Proc. Natl. Acad. Sci. U. S. A. 110*, 20051−20056.
(51) Crespo, M. D., Rubini, M., and Budisa, N. (2011) Rational Design of Protein Stability: Effect of (2S,4R)-4-Fluoroproline on the Stability and Folding Pathway of Ubiquitin. *PLoS One 6*, No. e19425.
(52) Kim, W., Hardcastle, K. I., and Conticello, V. P. (2006) Fluoroproline flip-flop: Regiochemical reversal of a stereoelectronic effect on peptide and protein structures. *Angew. Chem., Int. Ed. 45*, 8141−8145.
(53) Shoulders, M. D., Kamer, K. J., and Raines, R. T. (2009) Origin of the stability conferred upon collagen by fluorination. *Bioorg. Med. Chem. Lett. 19*, 3859−3862.
(54) Lin, Y.-J., and Horng, J.-C. (2014) Impacts of terminal (4R)-fluoroproline and (4S)-fluoroproline residues on polyproline conformation. *Amino Acids 46*, 2317−2324.
(55) Thomas, C. A., Talaty, E. R., and Bann, J. G. (2009) 3S-fluoroproline as a probe to monitor proline isomerization during protein folding by ^{19}F-NMR. *Chem. Commun.*, 3366−3368.
(56) Hodges, J. A., and Raines, R. T. (2005) Stereoelectronic and steric effects in the collagen triple helix: Toward a code for strand association. *J. Am. Chem. Soc. 127*, 15923−15932.
(57) Hofman, G.-J., Ottoy, E., Light, M. E., Kieffer, B., Kuprov, I., Martins, J. C., Sinnaeve, D., and Linclau, B. (2018) Minimising conformational bias in fluoroprolines through vicinal difluorination. *Chem. Commun. 54*, 5118−5121.
(58) Chaume, G., Lensen, N., Caupène, C., and Brigaud, T. (2009) Convenient synthesis of N-Terminal Tfm-Dipeptides from unprotected enantiopure α-Tfm-Proline and α-Tfm-alanine. *Eur. J. Org. Chem. 2009*, 5717−5724.
(59) Koksch, B., Ullmann, D., Jakubke, H. D., and Burger, K. (1996) Synthesis, structure and biological activity of α-trifluoromethyl-substituted thyreotropin releasing hormone. *J. Fluorine Chem. 80*, 53−57.
(60) Jlalia, I., Lensen, N., Chaume, G., Dzhambazova, E., Astasidi, L., Hadjiolova, R., Bocheva, A., and Brigaud, T. (2013) Synthesis of an MIF-1 analogue containing enantiopure (S)-α- trifluoromethyl-proline and biological evaluation on nociception. *Eur. J. Med. Chem. 62*, 122−129.
(61) Del Valle, J. R., and Goodman, M. (2002) Stereoselective synthesis of Boc-protected cis and trans-4-trifluoromethylprolines by asymmetric hydrogenation reactions. *Angew. Chem., Int. Ed. 41*, 1600−1602.
(62) Qiu, X., and Qing, F. (2002) Synthesis of Boc-protected cis- and trans-4-trifluoromethyl-d-prolines. *J. Chem. Soc., Perkin Trans.*, 2052−2057.
(63) Qiu, X.-l., and Qing, F.-l. (2002) Practical synthesis of Boc-protected cis-4-trifluoromethyl and cis-4-difluoromethyl-L- prolines. *J. Org. Chem. 67*, 7162−7164.
(64) Nadano, R., Iwai, Y., Mori, T., and Ichikawa, J. (2006) Divergent chemical synthesis of prolines bearing fluorinated one-carbon units at the 4-position via nucleophilic 5-endo-trig cyclizations. *J. Org. Chem. 71*, 8748−8754.
(65) Kubyshkin, V., Afonin, S., Kara, S., Budisa, N., Mykhailiuk, P. K., and Ulrich, A. S. (2015) γ-(S)-Trifluoromethyl proline: evaluation as a structural substitute of proline for solid state ^{19}F-NMR peptide studies. *Org. Biomol. Chem. 13*, 3171−3181.
(66) Lubin, H., Pytkowicz, J., Chaume, G., Sizun-Thomé, G., and Brigaud, T. (2015) Synthesis of Enantiopure *trans* −2,5-Disubstituted Trifluoromethylpyrrolidines and (2 *S*, 5 *R*)-5-Trifluoromethylproline. *J. Org. Chem. 80*, 2700−2708.
(67) Kondratov, I. S., Dolovanyuk, V. G., Tolmachova, N. A., Gerus, I. I., Bergander, K., Fröhlich, R., and Haufe, G. (2012) Reactions of β-alkoxyvinyl polyfluoroalkyl ketones with ethyl isocyanoacetate and its use for the synthesis of new polyfluoroalkyl pyrroles and pyrrolidines. *Org. Biomol. Chem. 10*, 8778.
(68) Tolmachova, N. A., Kondratov, I. S., Dolovanyuk, V. G., Pridma, S. O., Chernykh, A. V., Daniliuc, C. G., and Haufe, G. (2018) Synthesis of new fluorinated proline analogues from polyfluoroalkyl β-ketoacetals and ethyl isocyanoacetate. *Chem. Commun. 54*, 9683−9686.
(69) Mykhailiuk, P. K., Afonin, S., Palamarchuk, G. V., Shishkin, O. V., Ulrich, A. S., and Komarov, I. V. (2008) Synthesis of trifluoromethyl-substituted proline analogues as ^{19}F NMR labels for peptides in the polyproline II conformation. *Angew. Chem., Int. Ed. 47*, 5765−5767.
(70) Kubyshkin, V. S., Mykhailiuk, P. K., Afonin, S., Ulrich, A. S., and Komarov, I. V. (2012) Incorporation of cis- and trans −4,5-Difluoromethanoprolines into polypeptides. *Org. Lett. 14*, 5254−5257.
(71) Kubyshkin, V. S., Mykhailiuk, P. K., Afonin, S., Grage, S. L., Komarov, I. V., and Ulrich, A. S. (2013) Incorporation of labile *trans*-4,5-difluoromethanoproline into a peptide as a stable label for ^{19}F NMR structure analysis. *J. Fluorine Chem. 152*, 136−143.
(72) Hudson, A. S., Hoose, A., Coxon, C. R., Sandford, G., and Cobb, S. L. (2013) Synthesis of a novel tetrafluoropyridine-containing amino acid and tripeptide. *Tetrahedron Lett. 54*, 4865−4867.
(73) Gimenez, D., Dose, A., Robson, N. L., Sandford, G., Cobb, S. L., and Coxon, C. R. (2017) 2,2,2-Trifluoroethanol as a solvent to control nucleophilic peptide arylation. *Org. Biomol. Chem. 15*, 4081−4085.
(74) Gimenez, D., Mooney, C. A., Dose, A., Sandford, G., Coxon, C. R., and Cobb, S. L. (2017) The application of perfluoroheteroaromatic reagents in the preparation of modified peptide systems. *Org. Biomol. Chem. 15*, 4086−4095.
(75) Spokoyny, A. M., Zou, Y., Ling, J. J., Yu, H., Lin, Y. S., and Pentelute, B. L. (2013) A Perfluoroaryl-Cysteine S_NAr Chemistry Approach to Unprotected Peptide Stapling. *J. Am. Chem. Soc. 135*, 5946−5949.
(76) Lautrette, G., Touti, F., Lee, H. G., Dai, P., and Pentelute, B. L. (2016) Nitrogen Arylation for Macrocyclization of Unprotected Peptides. *J. Am. Chem. Soc. 138*, 8340−8343.
(77) Tressler, C. M., and Zondlo, N. J. (2014) 2 S,4 R)- and (2 S,4 S)-Perfluoro- tert -butyl 4-hydroxyproline: Two conformationally distinct proline amino acids for sensitive application in ^{19}F NMR. *J. Org. Chem. 79*, 5880−5886.
(78) Thomas, K. M., Naduthambi, D., and Zondlo, N. J. (2006) Electronic Control of Amide cis−trans Isomerism via the Aromatic−Prolyl Interaction. *J. Am. Chem. Soc. 128*, 2216−2217.
(79) Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M., and Fischer, G. (1998) Side-chain effects on peptidyl-prolyl cis/trans isomerisation. *J. Mol. Biol. 279*, 449−460.
(80) Ganguly, H. K., Kaur, H., and Basu, G. (2013) Local control of cis-peptidyl-prolyl bonds mediated by CH···π interactions: the Xaa-Pro-Tyr motif. *Biochemistry 52*, 6348−57.
(81) Ganguly, H. K., Majumder, B., Chattopadhyay, S., Chakrabarti, P., and Basu, G. (2012) Direct Evidence for CH···π Interaction Mediated Stabilization of Pro- *cis* Pro Bond in Peptides with Pro-Pro-Aromatic motifs. *J. Am. Chem. Soc. 134*, 4661−4669.
(82) Sarkar, S. K., Young, P. E., Sullivan, C. E., and Torchia, D. A. (1984) Detection of cis and trans X-Pro peptide bonds in proteins by ^{13}C NMR: application to collagen. *Proc. Natl. Acad. Sci. U. S. A. 81*, 4800−4803.
(83) Cobb, S. L., and Murphy, C. D. (2009) ^{19}F NMR applications in chemical biology. *J. Fluorine Chem. 130*, 132−143.
(84) Harper, D. B., and O'Hagan, D. (1994) The fluorinated natural products. *Nat. Prod. Rep. 11*, 123−133.
(85) Lee, Y.-J., Schmidt, M. J., Tharp, J. M., Weber, A., Koenig, A. L., Zheng, H., Gao, J., Waters, M. L., Summerer, D., and Liu, W. R. (2016) Genetically encoded fluorophenylalanines enable insights into the recognition of lysine trimethylation by an epigenetic reader. *Chem. Commun. 52*, 12606−12609.
(86) Sarker, M., Orrell, K. E., Xu, L., Tremblay, M. L., Bak, J. J., Liu, X. Q., and Rainey, J. K. (2016) Tracking Transitions in Spider

DOI: 10.1021/acs.biochem.8b00787
Biochemistry 2018, 57, 6132−6143

Wrapping Silk Conformation and Dynamics by ^{19}F Nuclear Magnetic Resonance Spectroscopy. *Biochemistry 55*, 3048−3059.
(87) Abboud, M. I., Hinchliffe, P., Brem, J., Macsics, R., Pfeffer, I., Makena, A., Umland, K. D., Rydzik, A. M., Li, G. B., Spencer, J., Claridge, T. D. W., and Schofield, C. J. (2017) ^{19}F-NMR Reveals the Role of Mobile Loops in Product and Inhibitor Binding by the São Paulo Metallo-β-Lactamase. *Angew. Chem., Int. Ed. 56*, 3862−3866.
(88) Kenward, C., Shin, K., Sarker, M., Bekkers, C., and Rainey, J. K. (2017) Comparing and Contrasting Fluorotryptophan Substitutions for ^{19}F Membrane Protein NMR Spectroscopy. *Biophys. J. 112*, 499a.
(89) Hoeltzli, S. D., and Frieden, C. (1995) Stopped-flow NMR spectroscopy: real-time unfolding studies of 6-^{19}F-tryptophan-labeled Escherichia coli dihydrofolate reductase. *Proc. Natl. Acad. Sci. U. S. A. 92*, 9318−22.
(90) London, R. E., Davis, D. G., Vavrek, R. J., Stewart, J. M., and Handschumacher, R. E. (1990) Bradykinin and its Gly^6 analog are substrates of cyclophilin: a fluorine-19 magnetization transfer study. *Biochemistry 29*, 10298−10302.
(91) Plaumann, M., Bommerich, U., Trantzschel, T., Lego, D., Dillenberger, S., Sauer, G., Bargon, J., Buntkowsky, G., and Bernarding, J. (2013) Parahydrogen-Induced Polarization Transfer to ^{19} F in Perfluorocarbons for ^{19} F NMR Spectroscopy and MRI. *Chem. - Eur. J. 19*, 6334−6339.
(92) Langen, K.-J., Hamacher, K., Bauer, D., Bröer, S., Pauleit, D., Herzog, H., Floeth, F., Zilles, K., and Coenen, H. H. (2005) Preferred Stereoselective Transport of the D-isomer of *cis* −4-[^{18} F]fluoro-proline at the Blood−Brain Barrier. *J. Cereb. Blood Flow Metab. 25*, 607−616.

DOI: 10.1021/acs.biochem.8b00787
Biochemistry 2018, 57, 6132−6143