



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

Towards the Synthesis of Photoswitchable Pheromones Interfacing Functional Yeast

verfasst von / submitted by

Darja Isabelle Rohden, B.Sc. B.Sc.

angestrebter akademischer Grad / in partial fulfillment of the requirements for the degree of

Master of Science (MSc)

Wien, 2022 / Vienna 2022

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

UA 066 863

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Masterstudium Biologische Chemie

Betreut von / Supervisor:

Univ.-Prof. Davide Bonifazi, PhD

Der experimentelle Teil dieser Arbeit wurde in der Zeit von April 2021 bis März 2022 am Institut für Organische Chemie an der Fakultät für Chemie der Universität Wien in der Gruppe von Univ.-Prof. Davide Bonifazi, PhD durchgeführt.

Erklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen Hilfsmittel als die angegebenen benutzt habe. Die Stellen, die anderen Werken (einschließlich des Internets und anderer elektronischer Text- und Datensammlungen) dem Wortlaut oder dem Sinn nach entnommen sind, habe ich in jedem einzelnen Fall durch Angabe der Quelle bzw. der Sekundärliteratur als Entlehnung kenntlich gemacht. Weiterhin versichere ich hiermit, dass die o.g. Arbeit noch nicht anderweitig als Abschlussarbeit einer Master-Prüfung eingereicht wurde. Mir ist ferner bekannt, dass ich bis zum Abschluss des Prüfungsverfahrens die Materialien verfügbar zu halten habe, welche die eigenständige Verfassung der Arbeit belegen können.

Wien, 12. April 2022

Acknowledgements

First of all, I want to thank all the people motivating and supporting me during the course of my master's thesis. I want to thank Univ.-Prof. Davide Bonifazi, PhD for hosting me in the research group and giving me the opportunity to work on this project. Also, I want to mention and thank Dr. Laura Maggini for the scientific support. Special thanks goes to my two supervisors André Oliveira Sequeira and Stavroula-Melina Sakellakou for their support and guidance - not only in scientific questions. I really appreciate your engagement and patience. In this sense, I also want to thank Terézia Morávková, who was not only a lab mate for me, but developed into a really good friend. Thanks also to the remaining 1H17 lab crew for all the good advises and for the friendly working environment, as well as to Rúben Ferreira for his help with the photophysical studies. Further, another word of thanks is addressed to the whole Bonifazi group. Thank you for your willingness to help and for taking me in into your international family. There are other thanks which need to be given to the Core Facilities of the University of Vienna for the numerous NMR and MS spectra taken. Last but not least, special thanks goes to my friends and family that supported me during the last year. Thank you for always keeping me on track and for helping me focus on my work.

Abstract

Biological systems require a high level of control and reversibility. In order for external users and researchers to be able to interfere with these processes and program them in a desired way, one option is to incorporate responsive organic molecules. Such compounds are responding to external stimuli by changing between two distinguishable states. When this external trigger is light, the compounds are referred to as *photoswitches*. In the purpose of this work, it was attempted to synthesize one of these photoswitches, namely an azobenzene derivative that is decorated with four fluorines in the *ortho* positions of the diazo bond and a respective amino and carboxy moiety in the *para* ends of the molecule. To this extent, it was aimed to incorporate it in a peptide's sequence responsible for the mating of a yeast species, *Schizosaccharomyces pombe*. The incorporation of this azobenzene in the peptidic core was envisioned to allow switching on and off the peptide by changing its conformation, and thus its biological activity to the cognate receptor it interacts with in the surface of the cells. The system will allow the reversible activation of cellular function in the genetically engineered yeasts. The main part of this work is focusing on the synthesis and photophysical characterization of the targeted azobenzene. As scaffold, a *ortho*-tetrafluoroazobenzene was chosen for the reason being that this will allow the selective isomerization by application of light in the visible range. The *para* positions of the molecule were designed to bear an amino and a carboxyl group on the two opposite sites of the compound. These functional groups can both be amidated permitting the coupling into the backbone of the studied peptide. Photophysical studies have been conducted on the targeted species, which confirmed the suitability of this compound for applications in biological systems.

Kurzfassung

Biologische Systeme benötigen ein hohes Level an Kontrolle und Reversibilität. Um diese Prozesse gezielt beeinflussen und kontrollieren zu können, bietet sich der Einbau von organischen Verbindungen an, welche auf externe Impulse reagieren. Das Stimulieren dieser Verbindungen führt zu einem Wechsel zwischen den zwei klar definierten Zuständen des Moleküls. Handelt es sich bei dem externen Stimulus um Licht, werden die Verbindungen als *Photoswitches* bezeichnet. Diese Arbeit hatte zum Ziel, einen solchen Photoswitch zu synthetisieren. Bei dem angestrebten Molekül handelt es sich um ein Azobenzol, welches an den *ortho*-Positionen der Diazobindung mit Fluoriden substituiert ist und zusätzlich an den zwei *para*-Positionen eine Carboxyl- beziehungsweise Aminogruppe trägt. Es war beabsichtigt, das Azobenzol in die Sequenz eines Peptids einzubauen, welches für die Paarung der Hefen *Schizosaccharomyces pombe* von Bedeutung ist. Durch den Einbau des Photoswitches direkt in die Peptidstruktur war es vorgesehen, das Ein- und Ausschalten des Peptids durch eine Konformationsänderung zu ermöglichen und dadurch seine biologische Funktion und die Interaktion mit seinem transmembranären Rezeptor zu steuern. Das Design des Systems erlaubt es, Zellfunktionen in genetisch veränderten Hefen reversibel zu aktivieren. Der Hauptteil dieser Arbeit fokussiert sich auf die Synthese und photochemische Charakterisierung des angestrebten Azobenzols. Als Grundgerüst wurde ein *ortho*-Tetrafluorazobenzol gewählt, da es die selektive Isomerisierung mit Licht im sichtbaren Bereich ermöglicht. Es wurde entschieden als funktionelle Gruppen an den *para*-Positionen zusätzlich eine Amino- und eine Carboxylgruppe einzufügen, wodurch sich die Möglichkeit ergibt, mittels Amidierungen den Photoswitch direkt in das Rückgrat der Aminosäurekette einzubauen. Abschließende photochemische Untersuchungen belegten die Eignung der gewählten Verbindung für die Anwendung in einem biologischen System.

Contents

List of Abbreviations	xi
Chapter 1: Introduction	1
1.1 When a Reversible Biological Process is Disturbed: The Causes of Diabetes . .	1
1.2 Examples of Photoswitches in Reversible Biological Processes	3
1.3 Aim of the Master Thesis	7
Chapter 2: Study of P-Factor	9
2.1 Introduction	9
2.1.1 <i>Schizosaccharomyces pombe</i> as Biological System	9
2.1.2 The Mating Pheromone P-Factor and its Interaction with the Mam2 Receptor	10
2.1.3 The Genetically Modified Strains Used in This Project	15
2.2 Results and Discussion	16
2.2.1 Solid Phase Peptide Synthesis	16
2.2.1.1 General Procedure	16
2.2.1.2 Generation of P-Factor	18
2.2.2 Circular Dichroism	19
2.2.2.1 Methodological Background	19
2.2.2.2 CD Studies of P-Factor	20
2.2.3 Incubation of P-Factor with Fission Yeasts	23
2.3 Conclusions	24
Chapter 3: Azobenzenes	25
3.1 Introduction	25
3.1.1 Photochemical Characteristics and Isomerization of Azobenzene	25
3.1.2 Tailoring Azobenzenes to Biological Applications	27
3.1.3 Methods of Synthesis	29
3.1.3.1 Azo Coupling	29
3.1.3.2 Mills Reaction	30
3.1.3.3 Wallach Rearrangement	31
3.2 Results and Discussion	32
3.2.1 Targeted Synthesis: HOOC- <i>o</i> -TFAB-NH ₂	32
3.2.1.1 Strategy A: A Detour via Triazenes	33
3.2.1.2 Strategy B: Introduction of Protection Groups	38
3.2.1.3 Strategy C: Esterification of the Benzoic Acid	43
3.2.2 Attempts to Couple HOOC- <i>o</i> -TFAB-NH ₂ with Amino Acids	45
3.2.3 Synthesis of <i>o</i> -tetrafluoroazobenzene (TFAB)	49

3.2.4	Photophysical Studies	50
3.2.4.1	Study of the Triazenes	51
3.2.4.2	Study of the Azobenzenes	52
3.3	Conclusions and Perspective	57

Chapter 4: Experimental **59**

4.1	Instrumentation	59
4.2	Materials and General Methods	60
4.3	Yeast Treatment	61
4.3.1	In-vitro Assays on Microplate Reader	61
4.3.2	<i>S. pombe</i> Viability Assay	61
4.4	Solid Phase Peptide Synthesis	62
4.4.1	Peptide Synthesizer	62
4.4.2	Preparation of P-Factor	62
4.4.3	P-Factor Isolation	64
4.4.4	P-Factor Analysis, Purification & Identification	64
4.4.5	Secondary Structure Analysis of P-Factor by Circular Dichroism	65
4.5	Organic Syntheses	66
4.5.1	Synthesis of Tetrafluorotriazenes	66
4.5.2	Introduction of Protection Groups to the Aromatic Precursors	67
4.5.3	Synthesis of <i>ortho</i> -Tetrafluoroazobenzenes	69
4.5.4	Attempts to Couple <i>ortho</i> -Tetrafluoroazobenzene to Amino Acids	71
4.6	Ultraviolet Light Experiments	72
4.6.1	Absorbance Spectra of Triazenes	72
4.6.2	Isomerization Studies	72
4.6.3	Fatigue of Isomerization	73
4.7	Appendix	74
4.7.1	Protocols for Solid Phase Peptide Synthesis	74
4.7.2	ESI-HRMS and NMR Spectra	76

Bibliography	xii
-------------------------------	------------

List of Abbreviations

AO	Acridine Orange	ICL	intracellular loop
ATP	adenosine triphosphate	IR	infrared
Boc	tert-butoxycarbonyl	LCP	left-handed circularly polarized lights
CD	circular dichroism	LUMO	lowest unoccupied molecular orbital
DCC	dicyclohexylcarbodiimide	MeSO₃	methane sulfonate
DCU	dicyclohexylurea	MPEG	methoxy polyethylene glycol
DIPEA	diisopropylethylamine	MS	mass spectrometry
DMF	<i>N,N</i> -dimethylformamide	MV	measuring vessel
DMSO	dimethylsulfoxide	MW	microwave
ECL	extracellular loop	NMP	N-methylpyrrolidone
EDC	1-ethyl-3-(3-dimethylamino-propyl)carbodiimide	NMR	nuclear magnetic resonance
EDG	electron donating group	PEM	photo-elastic modulator
EMM	Edinburgh Minimal Media	PI	Propidium Iodide
ESI	electron spray ionization	PSS	photostationary state
EWG	electron withdrawing group	RCP	right-handed circularly polarized lights
Fmoc	fluorenylmethoxycarbonyl	RT	room temperature
GFP	green fluorescent protein	RV	reaction vessel
GPCR	G-protein coupled receptor	SPPS	solid phase peptide synthesis
HATU	hexafluorophosphate azabenzotriazole tetramethyl uronium	TFA	trifluoroacetic acid
HBTU	hexafluorophosphate benzotriazole tetramethyl uronium	TFAB	tetrafluoroazobenzene
HMS	sodium hydroxymethanesulfonate	TFE	tetrafluoroethylene
HOAt	1-hydroxy-7-azabenzotriazole	TFP	tetrafluorophenol
HOMO	highest occupied molecular orbital	THF	tetrahydrofuran
HPLC	high performance liquid chromatography	TIPS	triisopropylsilsilane
HRMS	high resolution mass spectrometry	TLC	thin-layer chromatography
		UV	ultraviolet

Chapter 1: Introduction

This master thesis is embedded in a larger project focusing on the exploration of a yeast-based system for the production of biopharmaceuticals *in-situ*. One possible application is the treatment of diabetes and thus, the expression of insulin. The project bears two main challenges: First, the cells have to adhere regio-selectively inside the organism to generate a cellular assembly. Second, the therapeutic production needs to be controllable reversibly. The research of the present thesis was exclusively dealing with the second part. This first chapter is starting by describing the biochemical backgrounds of diabetes and the current state-of-the-art treatment (section 1.1). It was decided to use light as trigger to reversibly induce the insulin production in the yeasts. This is why, the second part of this chapter is focusing on the advantages of light for the application in biological systems and introduces different photoswitchable molecules (section 1.2).

1.1 When a Reversible Biological Process is Disturbed: The Causes of Diabetes

Reversibility is a concept found in basically every biological systems. Some reversible processes are quite obvious and almost everyone has already witnessed them. One example is the contraction of a muscle, which goes back into relaxation. While we are able to control this process in, e.g. our limbs, some muscles work without our active participation. The most popular example is our heart, which is regulated by exactly coordinated electric signals^[1]. But the phenomenon is not limited to the animal kingdom. *Mimosa pudica* is well-known for its response to mechanical forces, or changes in temperature and light. These triggers lead to the plant rolling up its leafs as a result of osmotic movements induced by changes in the ion distributions^[2]. In general, reversible mechanisms have a specific trigger for inducing and stopping a process. These triggers can vary from determined substrate levels or temporal ligand-receptor interactions to physicochemical causes. If these processes are disturbed, diseases can develop. One popular example is diabetes, which today affects 537 million adults according to the IDF (International Diabetes Federation)¹. In a healthy human body, the increase in blood sugar concentration leads to the production of insulin in pancreatic β -cells of the islets of Langerhans and its secretion into the blood. This in turn induces the uptake of glucose into the muscle and fat tissue, and at the same time enhances glycolysis and glycogen formation (Figure 1.1). In total, this leads to a decrease in blood sugar levels^[3,4]. To understand the mechanism of this reversible process, it is important to have a closer look on the role of glucose for the insulin production (Figure 1.1). Glucose is actively transported into the pancreatic β -cells where it gets metabolised, leading to the increase of intracellular adenosine triphosphate (ATP)

¹ <https://idf.org/aboutdiabetes/what-is-diabetes/facts-figures.html>, accessed 3rd April, 2022.

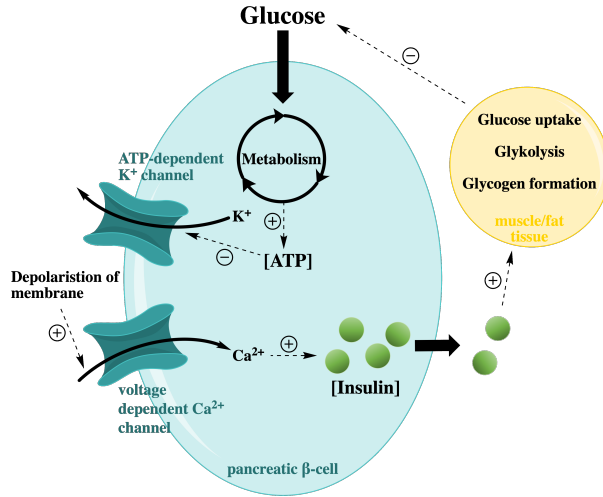


Figure 1.1: Regulation of insulin production and secretion. Dashed arrows describe activating and inhibiting relations. Bold arrows show the movement of materials across the membrane^[5,6]. This cycle is a typical example of reversibility, in which the trigger (a high glucose concentration) is depleted as a result of the process it originally induced. The underlying concept is also known as negative feedback loop^[4].

concentration and thus to a closing of ATP-dependent potassium channels. As a consequence, the cell membrane gets depolarised and charge-dependent calcium channels are activated. The increase in intracellular calcium concentration then leads to the secretion of insulin. As described above, the insulin induces several processes that result in a decrease of glucose levels and thus, less ATP formation in the pancreatic β -cells. The cell membrane repolarises and the insulin secretion stops^[3]. This cycle is a typical example of reversibility, in which the trigger (a high glucose concentration)

In diabetes, the blood sugar regulation by insulin is disturbed. Depending on the type of the disease, either the pancreatic β -cells are destroyed by the immune system due to an autoimmune defect (type I)^[7] or the body experiences an insulin resistance (type II)^[8]. The latter initially leads to an increase in insulin production to cope with the resistance. In the long term, also this type is characterised by low insulin levels as the pancreatic β -cells cannot maintain this high production. Consequently, the therapy for most patients suffering from any of these types includes, among other drugs, an external insulin supply to regulate the glucose levels^[7,8]. The method applied the most is the injection of insulin or an adhesive injection pump. Besides the physical obstacle in the everyday life, the patients can suffer from localized insulin-derived amyloidosis as a consequence of the injections^[9]. Thus, the research is focusing on more targeted approaches^[10,11]. One possibility to diminish the side effects coming from the punctual high insulin doses is to use a cellular body for the production and release of the pharmaceutical. In principle, this approach aims to restore the original reversible process by replacing the damaged part (the pancreatic β -cells) with an artificial one (the foreign cellular body). The expression of large proteins, such as insulin, in microorganisms is the standard methodology today^[12]. This is not only because of cost-effectiveness but also due to the enzymatic precision in protein folding^[13]. Currently, the pharmaceutical production is spatially and temporally separated from the delivery to the patient. An *in-situ* model, where the cellular body is incorporated in the organism and the production can be reversibly switched on and off through an external stimulus, bears the advantage of on-demand production and more steady delivery.

1.2 Examples of Photoswitches in Reversible Biological Processes

The study of a specific reversible process in a biological system often requires the spatiotemporal control of a biologically active compound. In neuroscience, for example, the reversible regulation of ion channels, and thus the neuronal firing, allows the control and exact study of specific neural circuits^[14]. Similarly, by alternately activating and deactivating specific receptors, their role in the nervous system and potentially in the development of diseases can be explored^[15]. In the mentioned cases, it is common to use small molecules that can switch between two stable states, each of them having a different influence on the studied biologically active compound^[14,15]. Such small molecules are also known as molecular switches. The alternation between the two states can be triggered by an external stimulus, e.g. temperature, electricity or light^[16]. The more complex the studied system gets, the more precise the trigger has to function in terms of space and time. Light is a highly tunable and fast stimulus and, depending on the used wavelength, it is non-invasive and orthogonal towards most elements in biological systems^[17,18]. This is why, light responsive molecular switches are getting more into the focus of research. To design photochromic compounds, also called photoswitches, for biological applications some requirements have to be met. In a review published in 2011, Beharry and Woolley^[18] highlighted four main points: (i) The molecule should effectively absorb at wavelengths between 340 and 800 nm as these was the range non-harmful for a biological system. The switch between the two states should (ii) happen with high efficiency to minimize the required light dosage and (iii) lead to an alternation in the studied biomolecule and its biological activity; and (iv) the compound should be non-toxic and stable in the biological environment in both states. Szymański *et al.*^[17] completed these requirements by "high fatigue resistance in an aqueous environment" for specific applications.

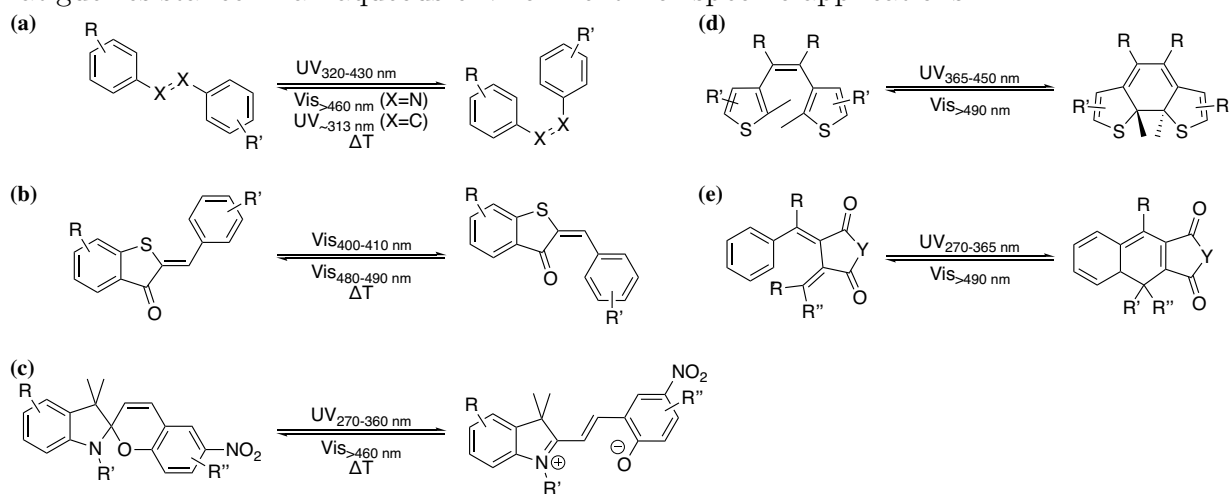


Figure 1.2: Examples of common photoswitches. (a) azobenzene ($X=N$) and stilbene ($X=C$), (b) hemithioindigo, (c) spiropyran, (d) diarylethene, (e) fulgide ($Y=O$) and fulgimide ($Y=N$).

Examples of common molecular structures of photoswitches are shown in Figure 1.2. They all consist of conjugated π -systems and thus, show absorption bands from the ultraviolet (UV) to infrared (IR) region. By irradiating the compounds at the respective wavelength of the π - π^* transition, the conversion between the two states can be induced. In case of azobenzene, stilbene and hemithioindigo, the excitation of an electron from the bonding π -orbital into the antibonding π^* -orbital leads to a breaking of the double bond. This allows rotation and thus, the interconversion of the *cis* and *trans* isomers (Figure 1.3a)^[19]. On the other hand, the irradiation can also induce unimolecular pericyclic reactions (diarylethene, fulgide and fulgimide). As example, the conversion of diarylethene is shown in Figure 1.3b. The irradiation with UV excites one electron from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO), leading to a newly generated HOMO with three nodes (Figure 1.3c). Consequently, the new σ -bond has to be formed in a conrotatory mechanism^[19], resulting in the two methyl groups showing into opposite directions in the closed isomer. Due to the higher stability of the σ -bond in respect to the broken π -bond, the cyclic states are thermally stable. Photoswitches comprising a photogenerated isomer with high thermal stability belong to the P-type category, referring to the back reaction being induced *photochemically*. In contrary, azobenzene, stilbene and spiropyran are associated with low thermal half-lives of one isomer and are categorized as T-type, in which the back reaction is *thermally* activated^[20].

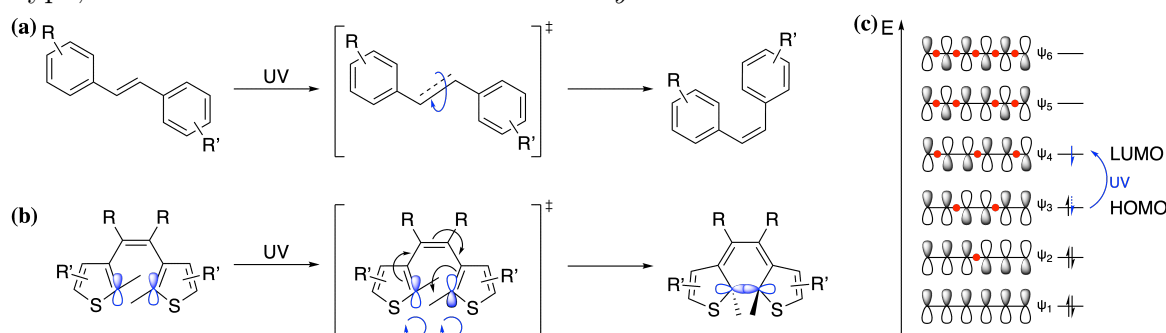


Figure 1.3: Conversion of photoswitches by UV light. **(a)** Isomerization of stilbene by rotation around the central C-C bond enabled by temporal rupture of the double bond. **(b)** Conrotatory electrocyclic reaction in diarylethene upon UV irradiation. **(c)** Representation of p-orbitals of diarylethene upon excitation of an electron by UV light.

The two states of the photoswitches typically show different characteristics. Depending on the particular chromophore, the isomerization can lead to changes in either or both geometry as well as charge distribution and polarity of the compound^[17]. This opens the opportunity for various applications in different stages of biological macromolecules, starting at the DNA level and going through to the structure and interactions of peptides and proteins (Figure 1.4)^[17]. The scope of this master thesis is dealing with peptide/protein interactions, which is why the following section will focus on the applications of photoswitches in peptides and proteins. Due to the high change in geometry

and end-to-end-distance, azobenzenes and stilbenes are perfectly suitable for structural studies of biological compounds. In 2005, Aemissegger *et al.*^[21] created a photoinducible β -hairpin by incorporating an azobenzene into the backbone of a peptide. Irradiation and isomerization to the *cis* form led to the disruption of the structure. The designed azobenzene derivative was later used by Hofmann *et al.*^[22] to investigate light-induced structural changes of AzoChignolin (Figure 1.5a). This is only one example of how the developed system can be exploited for the analysis β -hairpins in peptides and proteins. Also other chromophores can be integrated directly into the peptide backbone. One further example is the study conducted by Schober *et al.*^[23] in 2019 focusing on the conditionally cellular uptake of substrates. The researchers investigated the cell-penetrating characteristics of cyclic peptides comprising a diarylethene (Figure 1.5b). In its closed state, the molecule has a low flexibility due to the tricyclic structure. After irradiation, an electrocyclic reaction leads to the relocation of the double bonds and allows a free rotation of the thiophenes (Figure 1.3b). The researchers correlate this higher flexibility with the observed differences in cellular uptake^[23].

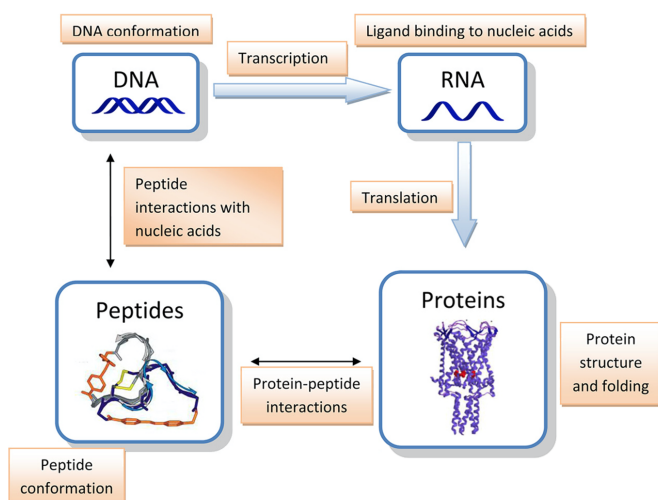


Figure 1.4: Biological processes that can be controlled by photoswitches. Figure taken from Szymański *et al.*^[17].

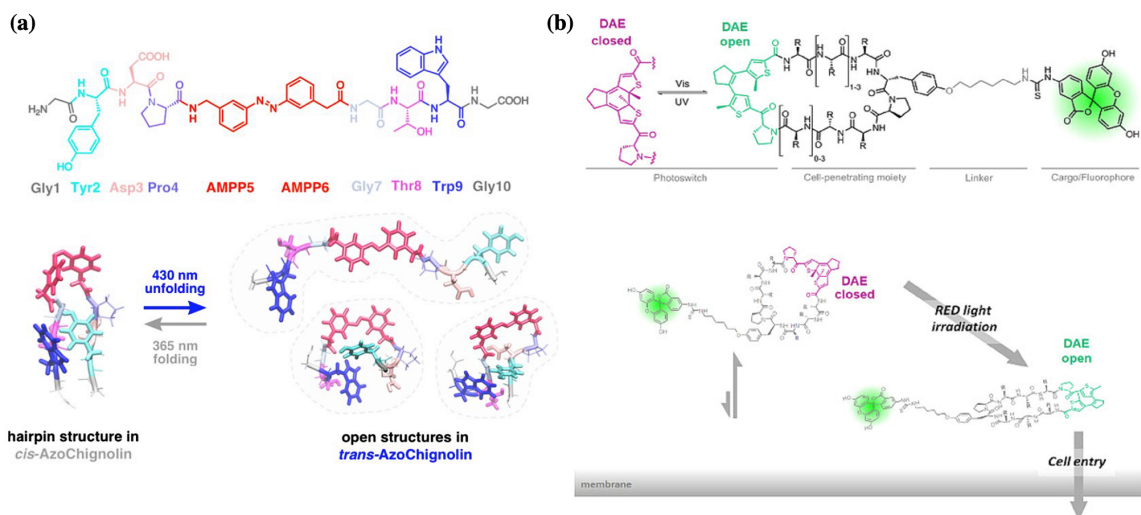


Figure 1.5: Photoswitches integrated into the peptide backbone. (a) Investigation of light-induced structural changes of AzoChignolin using azobenzene. Figure taken from Hofmann *et al.*^[22]. (b) Control of cellular uptake of cyclic peptides comprising a diarylethene. Figure taken from Schober *et al.*^[23].

Another method to introduce photoswitches into peptides is by using non-native amino acids comprising the chromophore as the side chain. Nakayama *et al.*^[24] exploited the different space requirements of *cis* and *trans* azobenzene to create spatial distance between two inter-

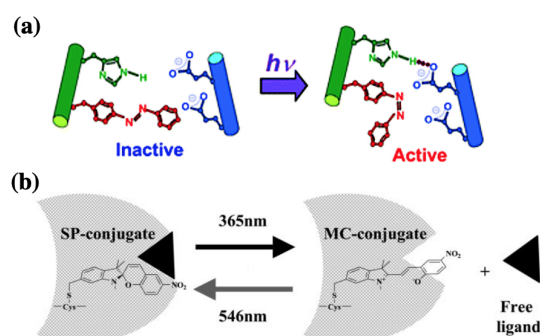


Figure 1.6: Photoswitches in amino acid side chains. (a) Azobenzene as non-native amino acid to control dimerization of *Bam*HI. Figure taken from Nakayama *et al.*^[24]. (b) Spiropyran integrated into the binding site of G-actin by orthogonal coupling to regulate binding events. Figure taken from Sakata *et al.*^[25].

acting proteins (Figure 1.6a). The investigated biomolecule is *Bam*HI, an endonuclease that is only active as dimer. By exchanging a non-functional amino acid that is located in the dimer interface with an azophenylalanine residue, the dimerization can be disturbed if the azobenzene is present *trans*^[24]. Instead of using a non-native amino acid, the integration of the chromophore into the side chain can also be achieved by orthogonal coupling. In 2005, Sakata *et al.*^[25] used thiol-active spiropyran derivatives, which they incubated with a protein containing exactly one cysteine residue, namely rabbit muscle G-actin (Figure 1.6b). This protein is a monomeric actin, which is an important subunit of the filaments in the cytoskeleton and the muscle cells^[26]. The cysteine is located in a domain involved in interactions with further actins and actin-binding proteins. The researchers used spiropyran, as it is the photoswitch comprising the highest change in dipole moment. Upon irradiation, it undergoes a heterolytic cleavage of the C-O σ -bond, leading to the development of merocyanine with a negatively charged oxygen and a positively charged nitrogen (Figure 1.2c). This change in dipole moment inside the active center was sufficient to inhibit the binding^[25]. The method can be even extended by using a photoswitch that is activated on both ends and thus, the coupling yields the formation of a bridge. One typical application for this procedure is the control of α -helices^[27,28]. By coupling an azobenzene to two residues of a peptide which are spaced i , $i+11$ in the sequence (which corresponds to three turns in the helix), the correct structure is maintained in *trans* conformation of the chromophore, while irradiation leads to a disruption of the helix (Figure 1.7)^[27,28].

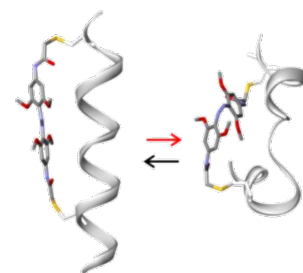


Figure 1.7: Azobenzene crosslinked to an α -helix for structural control. Figure taken from Samanta *et al.*^[27].

1.3 Aim of the Master Thesis

The Bonifazi group at the University of Vienna is working with a yeast-based system that is genetically engineered to be used as a biofactory for the production of pharmaceuticals *in-situ*. The yeasts are modified to express an insulin derivative when incubated with a specific peptide pheromone. The idea of this project is to explore the interaction of this pheromone with its respective receptor on the cell membrane and to remotely control the induction of intracellular signalling that results in insulin production (Figure 1.8). To achieve this, it is envisaged that one could induce a reversible conformational change in the peptide with only one of the conformations being able to activate the receptor. As introduced in section 1.2, integrating a chromophore into the backbone of a peptide allows to control the peptide's secondary structure with an external stimulus. For the reasons mentioned in the same section, light was chosen as the external trigger to induce the isomerization.

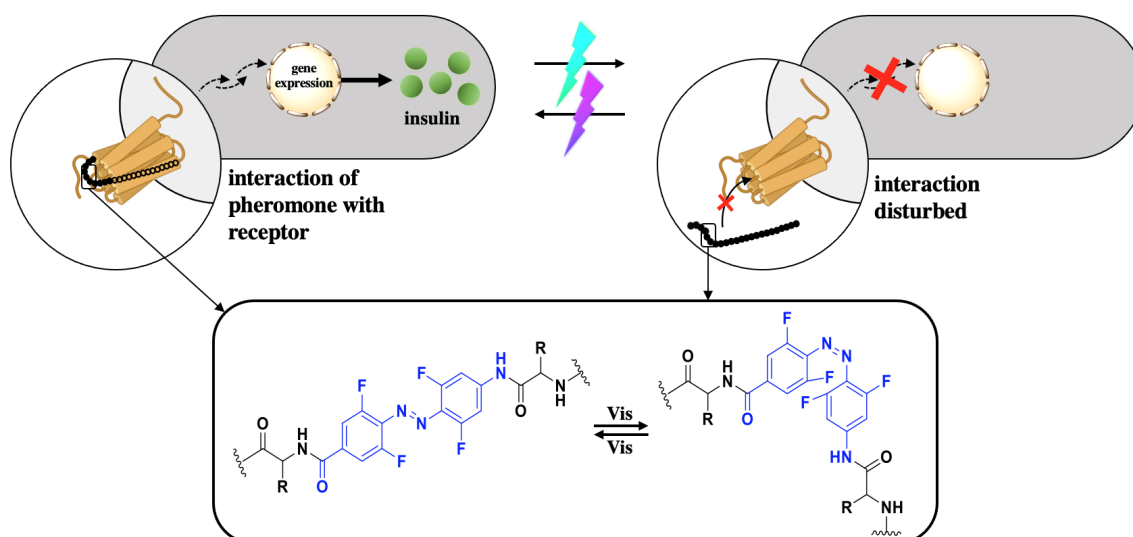


Figure 1.8: Schematic representation of the general idea of this thesis. Interaction of the pheromone with its receptor leads to insulin expression. By triggering the system with light, the isomerization of azobenzene leads to a conformational change in the peptide, which results in the disruption of the interaction.

There are some conditions that have to be considered in the choice of the right photoswitch. As it is desired to create a system that either expresses a high amount of the pharmaceutical or completely stops the production, a photoswitch with high photostationary states (PSSs) is needed. In this sense, it is also important to choose a molecule that can be selectively isomerized in both directions and where both isomers are thermodynamically stable. Furthermore, isomerization has to occur in the visible range of the spectrum, as UV light up to a wavelength of 340 nm can harm biological organisms^[18]. A photoswitch with high PSSs is azobenzene. Even though, the unmodified parent molecule does not match all of the above criteria, a lot of research was done in the last decades

focussing on the development of new classes of azobenzenes with high thermal stability and low-energy isomerization wavelengths^[29,30]. Additionally, derivatives of this molecule exhibit straightforward synthetic pathways. In total, this led to the following hypothesis:

If an appropriately modified azobenzene derivative is integrated into the backbone of a peptide, it is possible to use visible-range light to reversibly change the peptide's conformation and to control the cellular activation that induces the production of a biopharmaceutical.

This thesis will start by introducing the yeast-based biological system including the performed genetic modifications, and the interaction of the pheromone with its receptor (chapter 2). Subsequently, the synthesis and analysis of the azobenzene derivative is discussed in chapter 3.

Chapter 2: Study of P-Factor

As already mentioned in section 1.3, a yeast-based system, namely *Schizosaccharomyces pombe*, is chosen as the organism to create the reversible biological system. This chapter will begin by giving more insights into this model organism (section 2.1). A special focus will lay on the organism's mating pheromone P-factor and its receptor. The interaction between these two is intended to be used as the trigger for the insulin production. Further on, the genetic modifications to transform *S. pombe* cells into biofactories for pharmaceuticals are introduced. Afterwards, the synthesis of P-factor and its structural features in solution will be discussed in section 2.2. This part also includes a performed prestudy, in which the unmodified P-factor is incubated with the yeasts. Finally, the chapter is concluded by a short summary and perspective (section 2.3).

2.1 Introduction

2.1.1 *Schizosaccharomyces pombe* as Biological System

The biological system chosen for this project is *Schizosaccharomyces pombe* (*S. pombe*). Recently, Chin *et al.*^[31] reported the presence of this species in the gut microbiome of healthy humans. This finding indicates that the cells are not harmful for the human organism and can be suitable for an application *in-situ*. The organism is also called 'fission yeast', as they are able to asexually reproduce by a the central division of the cell ('fission')^[32]. They belong to the ascomycota inside the kingdom of fungi. The division got its name from special sacks ('asci'), in which the sexual spores are generated and stored^[33]. *S. pombe* cells exist in two haploid mating types: M and P. The cell type identity is genetically defined in the *MAT* locus. The M-cells are expressing *mat1-M* and the P-cells *mat1-P*^[33]. While both mating types have the full genetic information, only one locus is transcriptionally active and the other one is stored in a silent cassette. Like that, the cells in a colony are able to switch between the mating types if necessary^[33]. During asexual reproduction, the yeasts grow in length before they divide in two identical daughter cells by medial fission^[32]. If nutrients (especially nitrogen) become limited, an intracellular signalling leads to a decrease in intracellular cAMP levels, which in turn triggers the switch to sexual reproduction. Subsequently, the two mating types start releasing specific diffusible mating pheromones: M-factor and P-factor (Figure 2.1a). The peptides are recognized by respective receptors on the cell membrane of the opposite sex, Map3 and Mam2, respectively, and induce mating-related changes. This includes, among others, the expression of more mating pheromones and receptors, and morphological changes. Precisely, the cells start growing projections towards each other and produce agglutinins on the cell walls to enable fusion of two cells with opposite mating types. The diploid zygote enters

into meiosis and the DNA gets duplicated. Subsequently, two consecutive meiotic divisions lead to the formation of four haploid spores which are germinated (Figure 2.1b) [34].

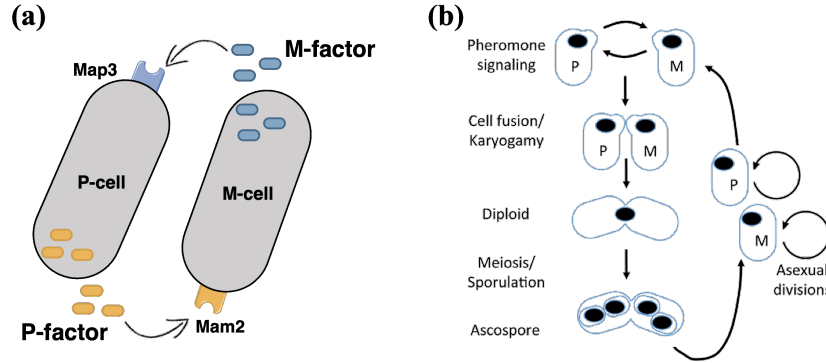


Figure 2.1: (a) Pheromone based interaction between the two mating types in sexual reproduction. (b) Life cycle of *S. pombe*. The cells can divide by mitotic asexual growths or opposite sex mating. After cell fusion, the diploid cell undergoes meiosis and sporulation to regenerate haploid cells. Figure taken from Bennett and Turgeon [33].

Not only because of their interesting life cycle, the *S. pombe* became more popular in recent years. The yeast comprises the smallest sequenced eukaryotic genome without any evidence for genome duplications. This makes this organism especially suitable for genetic studies as the gene of interest may be present in only one copy [32].

2.1.2 The Mating Pheromone P-Factor and its Interaction with the Mam2 Receptor

As described in section 2.1.1, the fission yeast can enter the sexual reproduction cycle upon nutrition depletion [35]. The two distinct cell types, M- and P-cells, start producing and secreting pheromones which are recognized at the outer cell membrane of the opposite sex. P-factor is one of the two pheromones. The peptide bears 23 amino acids and is produced by P-cells [36]. Thus, after secretion, it is recognized by a receptor on the membrane of M-cells, the Mam2 receptor (Figure 2.1a) [35]. The receptor belongs to the family of G-protein coupled receptors (GPCRs), which is the largest integral membrane protein family and consists of six classes (A-F) [37]. Although the six classes differ in their sequences, the general structure is a common characteristic among all of them (Figure 2.2).

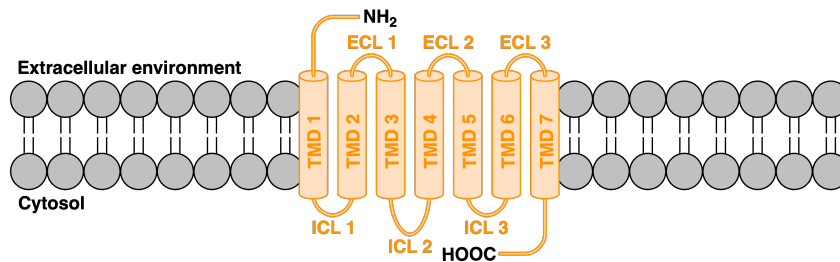


Figure 2.2: General structure of G-protein coupled receptors.

They all possess seven transmembrane domains in the shape of α -helices, which are connected by three intracellular and three extracellular loops (ICLs and ECLs)^[37,38]. The N-terminals of the receptors are exposed to the extracellular matrix and, in some classes, participate in the ligand binding. The high variety found in the ICL and ECL sequences are responsible for the family's ability to interact with a wide range of ligands. The Mam2 receptor belongs to class D - fungal mating pheromone receptors. As the name suggests, this class is only present in fungi. While for the classes A-C and F, there is already a lot of data focusing on the interactions of ligands with their receptors^[37,38], class D receptors are relatively unexplored. The main reason lays in their phylogenetic distance to the other classes. In general, structural studies of transmembrane proteins, such as GPCRs, are challenging due to the necessity of integrating them into a cell membrane. The dynamisms in the loop regions of the receptors further hamper the experiments^[38].

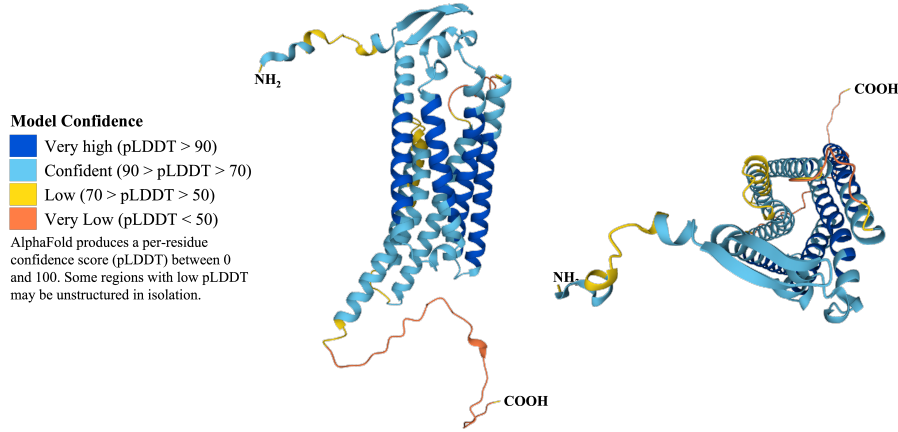


Figure 2.3: AlphaFold prediction of Mam2 conformation as uploaded on the Uniprot database³.

There is a very limited amount of data available describing the interaction between the P-factor and Mam2. A good starting point is the sequence and prediction of Mam2's structure (Figure 2.3), which is uploaded on the Uniprot database². However, the prediction has only low confidence in the NH_2 region and in the ECLs. Furthermore, the dataset does not include the ligand and thus, the exact position of the pheromone inside the receptor's cavity remains unclear. *Saccharomyces cerevisiae*, a distant relative of *S. pombe*^[36], comprises a similar mating system^[39]. The organism is one of the most used in microbiological assays and thus, there is a high density of available studies. The budding yeasts also communicate via short peptides (α -factor and **a**-factor), which are recognized by GPCRs on the membrane of the other mating-type. It is reported that the Ste2p, the receptor of α -factor, has homology with Mam2^[40]. Furthermore, a comprehensive study done by Imai and Yamamoto^[36] revealed a high similarity between the functions and precursors of P-factor and α -factor. Both peptides arise from a large precursor bearing a signal sequence and possible N-linked glycosylation sites, important for the secretion pathway in the cell. The mature pheromone sequence is present in repeats, which are separated by basic amino acid residues as cleavage signals. After posttranslational processing, both natural

² uniprot.org/uniprot/Q00619, accessed 19th March, 2022.

peptides are unmodified. Upon their interaction with the receptor on the cell membrane, they induce cell cycle arrest in the G_1 phase and changes in the cell morphology towards mating. The latter includes the expression of more mating pheromones and receptors. Additionally, proteases are expressed, leading to the degradation of the pheromone and thus, downregulation of the response^[36]. Another important similarity between the two pheromones is found in their sequence. Even though, the majority of amino acids differs and α -factor only consists of 13 amino acids (instead of 23 in case of P-factor), both bear two proline residues at the third and sixth position from the C-terminus (Table 2.1)^[36,39].

Table 2.1: Amino acid sequences for the mating pheromones α -factor and P-factor.

Pheromone	Amino acid sequence
α -factor	WHWLQDKPGQPMY
P-factor	TYADFLRAYG S WNTFVNPDRLPNL

turn region

Due to its ring structure and the missing hydrogen bond donating capability, a proline is acting differently in the peptide backbone and thus, has a stronger influence on its secondary structure^[41]. Because of chemical hindrances of the side chains, the other amino acids are almost entirely present in the *trans* conformation. Proline, however, can exist in both isomeric forms (Figure 2.4a) and the *cis* form is found in approximately 5% of acyclic peptides, introducing a structural heterogeneity^[42]. Also, the surrounding amino acids strongly influence the final conformation. If, for example, a glycine is following the proline, the peptide folds into a β -turn (Figure 2.4b)^[43].

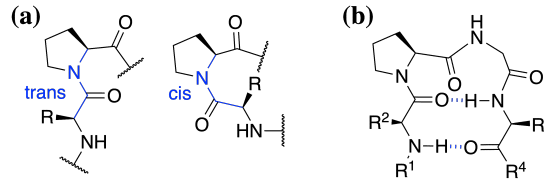


Figure 2.4: Structure of (a) *trans/cis* proline and (b) the β -turn induced by the Pro-Gly motif.

Thus, the total amount of prolines and their distance to each other is an important factor when predicting a peptide's secondary structure. Furthermore, they are often found in regions important for the regulation of a peptide's function and tend to be highly conserved in the peptide sequences^[44]. The high similarity in the function of P-factor and α -factor, and the homology of their GPCRs in combination with the analogy in the position of the two prolines leads to the assumption that also the pheromones' secondary structure and their interaction with the respective receptor are similar. This is why, the following paragraph will go into more detail of researches focusing on the structure of α -factor and its interaction with the Ste2p receptor. The findings will be used to draw conclusions about P-factor and Mam2.

When analysing the sequence of α -factor, the Pro-Gly motif is present in the C-terminal region, indicating a β -turn. Indeed, several studies reported a β -turn around these residues,

whose presence was strongly correlated with the intracellular signal transduction. The latter was shown by a loss of binding and activation in modified peptides where the turn was not present or disrupted^[45–49]. In comparison, there is no Pro-Gly motif in the P-factor sequence. Assuming a similar structure, other factors have to stabilize the turn in this pheromone. Armstrong *et al.*^[50] investigated the interactions between arginine side chains and other amino acids. The authors report a stacking phenomenon between the guanidinium group and the amide group of an asparagine side chain if they appear in spatial proximity. According to their results, the parallel stacking of arginine and asparagine occurs both in α -helices as well as around turns when the two amino acids are three positions apart from each other. The interaction leads to a stabilization of the respective structural motif. In case of P-factor, the two residues are found at position 17 and 20 (Table 2.1). One of the two enclosed amino acids is a proline, which is known to interrupt an α -helix^[44]. Therefore, the discussed findings give further evidence for a turn motif in the C-terminal region of P-factor, which is stabilized by the parallel stacking of the surrounding asparagine and arginine residues. Further studies of α -factor focused on the N-terminal region of the pheromone and revealed that this part is strongly interacting with lipid vesicles in both presence and disruption of the turn^[46,47]. This can be explained by the three aromatic residues at the N-terminus. The interaction of the pheromone with its receptor was broken down into stepwise conformational changes starting with the aromatic and basic amino acids at the N-terminus interacting with the lipid bilayer. As a consequence, the peptide is entering the cavity of Ste2p and further interactions, like ring stacking or hydrogen bonding, occur between the ligand and the receptor^[51]. A conformational change is induced, leading to the formation of the loop region and the positioning of also the C-terminus towards the inside of Ste2p (Figure 2.5a)^[52,53]. Abel *et al.*^[54] revealed that both termini are important for the successful interaction, whereby the N-terminus is more relevant for the signalling and the C-terminus for the binding (Figure 2.5b). Recent studies revealed the crystal structure of active Ste2p (Figure 2.5c)^[49]. The receptor is present as dimer^[55] and thus, interacts with two α -factors in its orthosteric binding pocket, which is located between the residues of the transmembrane domains' extracellular zone^[49].

When applying this knowledge on P-factor and the Mam2 receptor, it can be hypothesized that upon the interaction of the two, the pheromone folds into a β -turn in its C-terminal region and that both termini are entering the receptor's cavity. In 2000, Ladds and Davey^[56] investigated the importance of the C-terminal leucine. As described above, the cellular response to P-factor is regulated by a protease, namely Sxa2. The enzyme belongs to the serine carboxypeptidases and removes the terminal residue^[57,58]. The researchers report an inactivation of the peptide after exposure to the enzyme. More recent studies confirmed these findings: By using Atomic Force Microscopy, Sasuga *et al.*^[35] examined the binding of P-factor and P-factor Δ Leu to Mam2. The technique measures the force needed

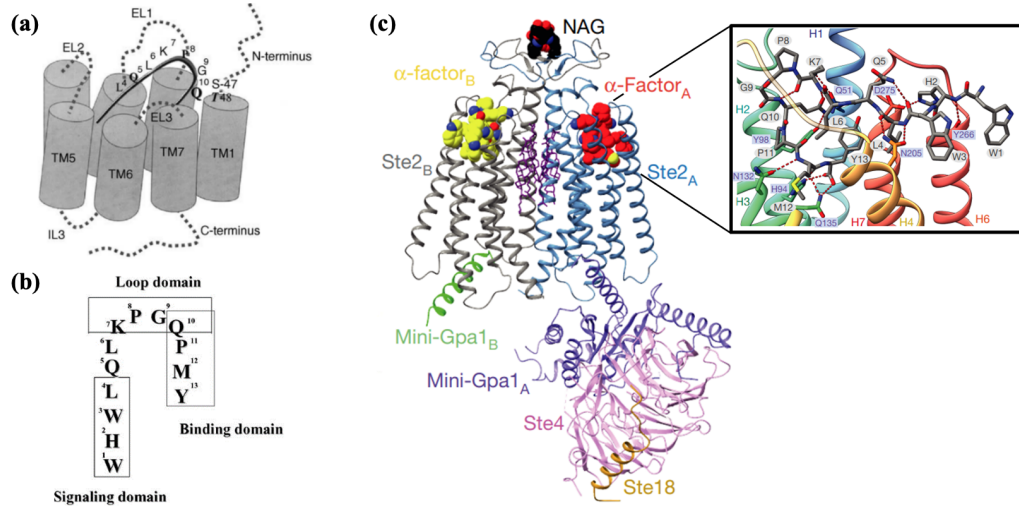


Figure 2.5: Interaction of α -factor with Ste2p. (a) Schematic representation of α -factor binding inside Ste2p. Figure taken from Lee *et al.*^[51]. (b) Domains of α -factor as proposed by Abel *et al.*^[54]. (c) Crystal structure of active Ste2p-G-protein complex upon binding of α -factor. Figure taken from Velazhahan *et al.*^[49].

to interrupt the binding between a chemical and a surface. When doing the experiment with the leucine-depleted pheromone, the observed force was similar to that of the negative control, confirming that no binding is occurring (Figure 2.6a,b). In a second experiment, the group investigated the signalling efficiency of the two analogues by using a green fluorescent protein (GFP) reporter assay. Again, the results of P-factor Δ Leu were equivalent to that of the negative control (Figure 2.6c)^[35]. In 2013, Hidaka *et al.*^[59] performed similar experiments by using P-factor analogues of different lengths with sequential deletions of the N-terminus. With decreasing length of the peptide, less GFP expression was reported (Figure 2.6d), leading to the conclusion that the lengths of the N-terminus has an influence on either or both the binding efficiency and the signalling activation^[59].

In order to further investigate the importance of specific amino acids in the sequence, a peptide library was created previously to the current work³. The designed sequences were orientated on the above stated results and included C- and N-terminal truncated peptides as well as alanine substitutions. In contrast to the previous studies, the group also examined truncated peptides, in which the terminal amino acid was retained. It was possible to confirm the importance of the length of the pheromone for the interaction with Mam2. However, in some cases, the binding efficiency could be partially restored in the presence of the terminal amino acid. Furthermore, it was shown that no signalling occurred upon substitution of one or both prolines³.

³not published, see doctoral theses of André Oliveira Sequeira and Stavroula-Melina Sakellakou (2022)

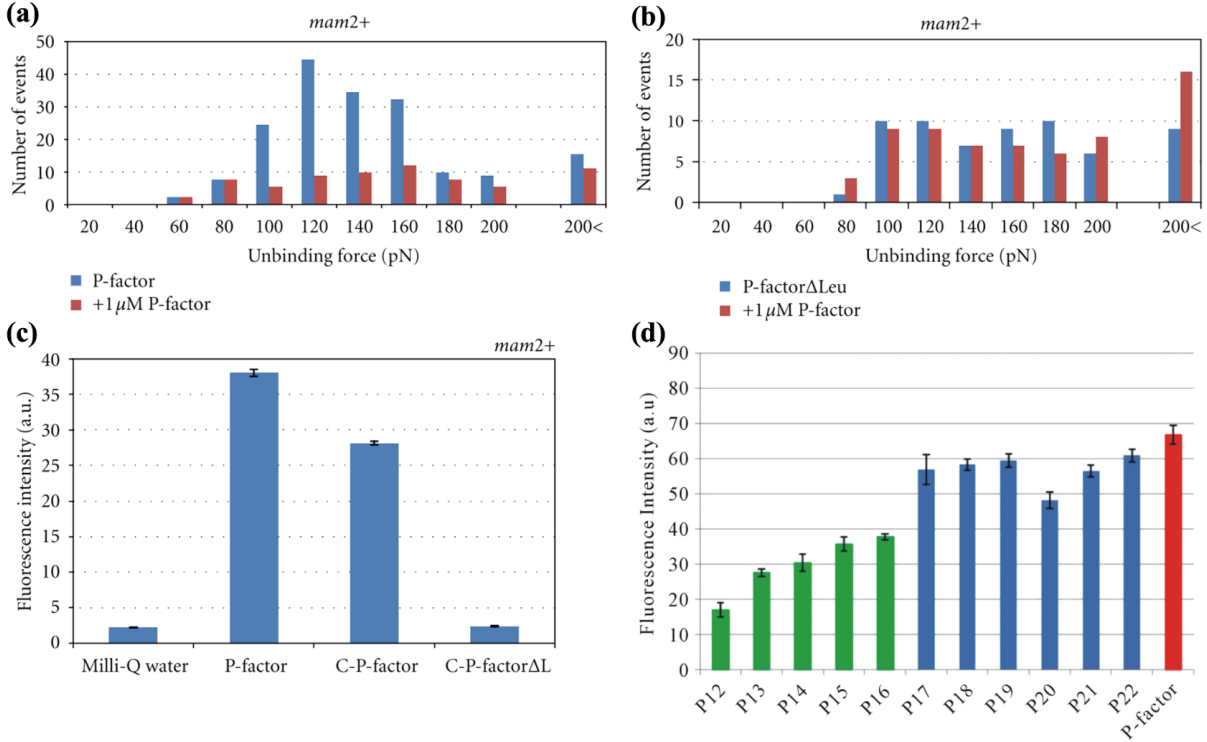


Figure 2.6: (a)-(c) Results of studies performed by Sasuga *et al.*^[35]: Binding events between mam2+ cells and (a) P-factor or (b) P-factorΔLeu. (c) GFP reporter assay for P-factor and P-factorΔLeu. (d) Results of GFP reporter assay performed by Hidaka *et al.*^[59] using peptide sequences with truncations from the N-terminus.

2.1.3 The Genetically Modified Strains Used in This Project

In order to use *S. pombe* cultures as a biofactory for pharmaceuticals, they had to be genetically engineered⁴. The aim of the project is to create a reversible biological system that can produce insulin. For this work, the long-acting insulin analogue Albulin is used, which is "obtained by direct gene fusion of a single-chain human insulin to human serum albumin"^[60], making it more stable in physiological conditions. Thus, the most important modification was the integration of plasmid pDH794 bearing the genetic information for the expression of Albulin. To quantify the production of Albulin in real-time *in-vivo*, the biopharmaceutical was tagged with GFP. As described in section 2.1.1, the intracellular response to the pheromone only occurs when the cells are experiencing a limitation in nitrogen supply. For the microbiological work, this would demand filtering and washing the cells before any experiment. To uncouple the response from the environmental conditions, the deletion of the *cyr* gene is sufficient^[36]. The last genetic modification is the inactivation of the protease Sxa2. This protein is used by the yeasts to balance the cellular response by removing the peptide's C-terminal amino acid and thus, inactivating the pheromone^[57,58]. To diminish the risk of cleaving the azobenzene-modified P-factor during the cell experiments, the respective gene was deleted^[36]. All these manipulations

⁴ Genetic modifications were performed by Prof. Damien Hermand, University of Namur, Belgium.

were combined in strain 1838, which bears the genotype *cyr::hphR sxa2::kanR* pDH794. Furthermore, strain 1831 was used in this project to examine the effective binding and interaction of (modified) P-factor with Mam2. As already mentioned in section 2.1.2, the induced cell signalling leads to the expression of different gene products including Mam2 receptor. Strain 1831 is expressing the GFP reporter gene as fusion with Mam2 and produces Albulin without GFP. The strain bears the genotype *cyr::hphR sxa2::kanR mam2-GFP-natR* pDH792.

2.2 Results and Discussion

2.2.1 Solid Phase Peptide Synthesis

The chemical synthesis of peptides has been a long-studied field in bioorganic chemistry. Before the development of the solid phase peptide synthesis (SPPS) by R.B. Merrifield in 1963^[61,62], the synthesis included laborious purification steps between the individual coupling reactions and, in most cases, gave the targeted protein in low yield and purity^[63]. By introducing a solid phase, Merrifield diminished these drawbacks, which will be further discussed in section 2.2.1.1. Over the years, several improvements and adaptations led to the method becoming less dangerous and easier to handle^[63,64]. Today, this strategy is the most chosen one to synthesize even relatively complex peptides^[63]. As already mentioned in section 2.1.2, the native pheromone has no posttranslational modifications. Imai and Yamamoto^[36] proved that the synthetically derived peptide shows the same activity as the purified one from yeast expression. Thus, it was decided to use SPPS for the synthesis of P-factor.

2.2.1.1 General Procedure

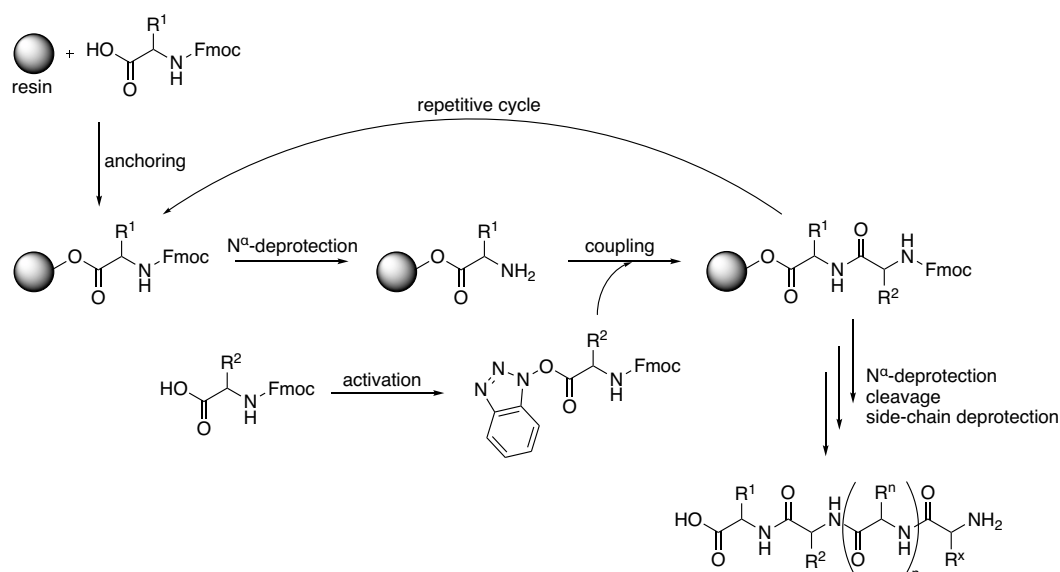


Figure 2.7: Approach of solid phase peptide synthesis. As an example, the Fmoc protection group and an activation with HATU is illustrated. Figure adapted from Fields^[64].

The general SPPS strategy starts with the first amino acid being attached to a solid resin and then coupled stepwise to the following residues (Figure 2.7). This allows the usage of the activated amino acid in excess compared to the growing peptide chain, giving higher coupling yields^[64]. To diminish side reactions, the amino acids are temporarily protected on their amino groups until their coupling to the chain is completed. A subsequent washing removes the unreacted amino acid without the need of further purification steps. The amino end is then deprotected to allow the peptide to enter in the next coupling cycle. After the last amino acid is coupled, the peptide chain is cleaved from the resin.

Another important factor to exclude side products is the protection of reactive amino acid side chains. The chosen protection groups are stable during the whole synthesis and their removal is typically performed simultaneously with the final cleavage. In general, two strategies can be distinguished in SPPS, which are named after the α -amino group protective sites. In the Boc-strategy, a tert-butoxycarbonyl (Boc) group is attached to the incoming amino acid at its amino end^[62]. The side chains are protected by benzyl based protection groups. While the Boc group can be released with trifluoroacetic acid (TFA) in CH_2Cl_2 , the side chain deprotection and the final cleavage from the resin are performed with stronger acids like HF or HBr^[64]. Thus, this strategy is not completely orthogonal and the repeated exposure to TFA during the peptide synthesis can lead to partial deprotection of the side chains, early cleavage from the resin and the development of side products^[63]. Furthermore, the use of HF requires special safety measurements, making this strategy less favoured. The alternative is the Fmoc-strategy, where the α -amino group is protected with a fluorenylmethoxycarbonyl (Fmoc) group^[65]. As this is a base labile protection group, an orthogonal strategy can be achieved by using acid labile protection groups on the amino acid side chains^[64]. This leads to both less risk of side products and milder reaction conditions.

The choice of solid support depends mainly on the likeliness of the target peptide to form aggregates. Besides the presence of hydrophobic side chains (like leucine, valine, phenylalanine or isoleucine), which tend to aggregate in the polar solvents used in SPPS, the length of the peptide is an important factor in this sense. With increasing chain length, aggregation is more probable^[66]. By choosing a resin with a lower loading (e.g. 0.5 to 1.2 mmol/g for peptides with 10 to 20 amino acids) or one that is providing a polar environment, this side-effect can be reduced. Furthermore, the available resins can be distinguished between those generating peptides with C-terminal carboxyl groups (e.g. Merrifield, 2-chlorotrityl, Wang) and those for C-terminal amides (e.g. MBHA, Rink amide, Sieber amide)^[67]. The target peptide, P-factor, has 23 amino acids with a content of approximately 50% of hydrophobic residues and terminates in a carboxyl group. Thus, a Wang resin with a loading of 0.650 mmol/g was chosen.

2.2.1.2 Generation of P-Factor

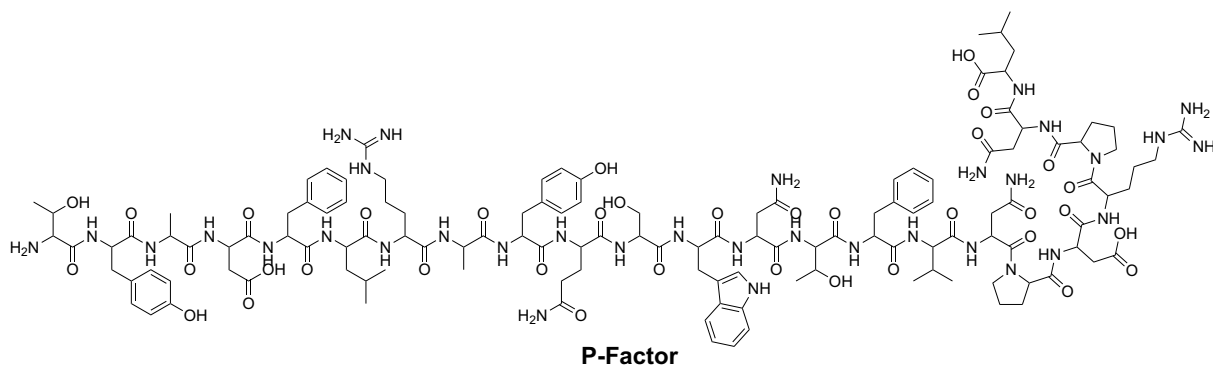


Figure 2.8: Sequence and structure of the mating pheromone P-factor.

The P-factor was synthesized using a *Focus XC* semi-automatic peptide synthesizer. For the reasons shortly discussed in the previous section, the Fmoc-strategy was chosen. As the C-terminal amino acid in the P-Factor is a leucine, a preloaded Fmoc-Leu Wang resin was used as solid support. Deprotection of the α -amino group is performed by addition of 20% piperidine in *N,N*-dimethylformamide (DMF) in three cycles. The activation of the incoming amino acid can be achieved by reacting it with an activation reagent, like hexafluorophosphate azabenzotriazole tetramethyl uronium (HATU), hexafluorophosphate benzotriazole tetramethyl uronium (HBTU) or dicyclohexylcarbodiimide (DCC) (Figure 2.9)^[62,64]. It has been reported that couplings with uronium species show more rapid kinetics than those performed with carbodiimides^[64]. The difference between HBTU and HATU is that the latter bears a nitrogen in the seventh position of the six-membered ring. The pyridine moiety is able to provide anchimeric assistance in the coupling reaction due to its lone pair on the nitrogen. Thus, the reaction with HATU is faster than that with HBTU, which is why it was chosen as activation reagent.

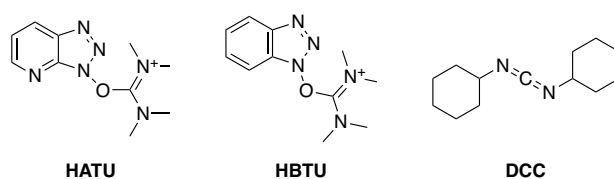


Figure 2.9: Activation agents for amidation reactions. For simplicity, the counterions are not shown.

The mechanism is illustrated in Figure 2.10. HATU exists in two isomers, the uronium salt and the iminium salt, which can be transferred into each other by a simple intramolecular mechanism. The iminium salt can be attacked by the amino acid's carboxyl group to give another uronium species. In this step, it is crucial to have a basic environment as the carboxylic acid has to be deprotonated in order to be a good nucleophile. For this purpose, diisopropylethylamine (DIPEA) is added to the reaction mixture. Subsequent attack of the azabenzotriazole-*N*-oxide anion leads to the formation of the active ester under the liberation of tetramethylurea^[68].

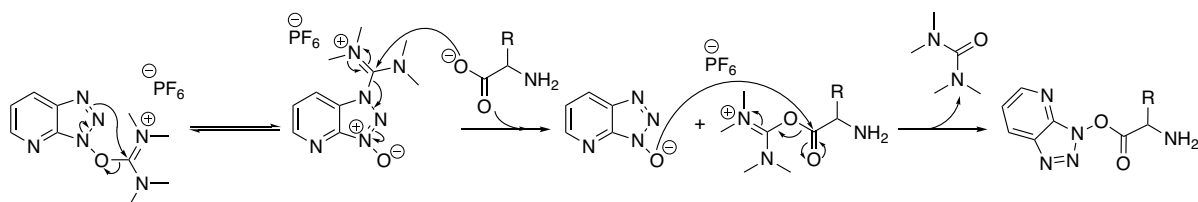


Figure 2.10: Mechanism of amino acid activation by HATU^[68].

Both the activation and the coupling are done in a solvent mixture of DMF and N-methylpyrrolidone (NMP). The deprotection and coupling were repeated until all 23 amino acids were coupled. Subsequently, the peptide is subjected to a cleavage cocktail based on TFA (compare section 4.4.3). As described in section 2.2.1, TFA can be used for the cleavage from the resin and final N $^{\alpha}$ and side chain deprotection. Further, the cocktail contains triisopropylsilane (TIPS), water and phenol. While the first two compounds work as scavengers to catch evolving electrophilic cations arising from the deprotection of the side chains, phenol protects the tyrosine and tryptophane residues from oxidation. The gained crude peptide was then purified by high performance liquid chromatography (HPLC). The HPLC chromatograms and ESI-HRMS analysis of pure P-factor can be found on the pages 64 and 65 in chapter 3.3.

2.2.2 Circular Dichroism

2.2.2.1 Methodological Background

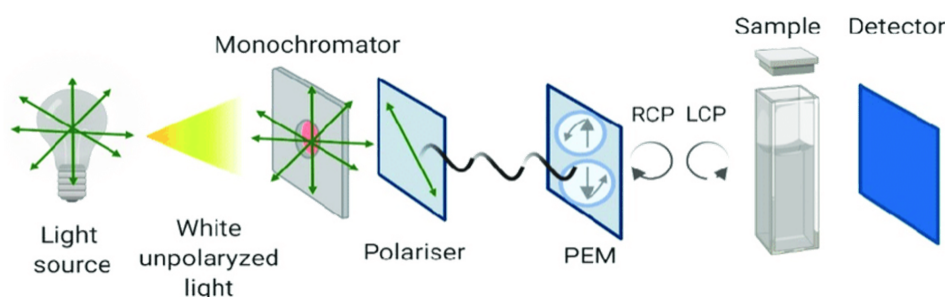


Figure 2.11: Schematic representation of a CD spectrometer. Figure taken from Pignataro *et al.*^[69].

To investigate the secondary structure of the synthesized P-factor in solution, circular dichroism (CD) spectroscopy experiments were performed. This technique uses a beam of plane-polarized light to analyse the chirality of optically active molecules^[69,70]. Peptides belong to this group of molecules as they consist of numerous amide bonds and (depending on the peptide) aromatic side chains and disulphide bridges. The plane-polarized light in the CD experiment is a superposition of left-handed and right-handed circularly polarized lights (LCP and RCP). In a CD spectrometer (Figure 2.11), the incoming beam is first polarized and then directed through a photo-elastic modulator (PEM), which produces the LCP and RCP^[69]. While in the incoming beam, both LCP and RCP have the same

intensity, the lights are absorbed differently by the optically active molecules, leading to a change of the absorption. This phenomenon can be observed by a difference in the amplitude of the polarised light vector^[70].

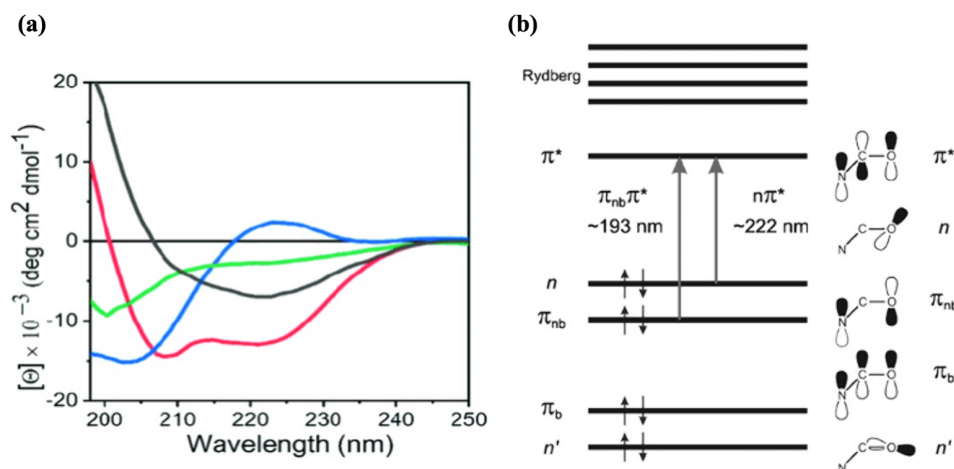


Figure 2.12: (a) Characteristic CD spectra for secondary structures: random-coil (green), α -helix (red), β -sheet (black) and polyproline (blue). Figure taken from Pignataro *et al.*^[69]. (b) the electronic transitions of an amide bond. On the right, molecular orbitals of the amide bond are shown: π_b , π_{nb} and π^* correspond to the bonding, non-bonding and antibonding π orbitals, respectively. n and n' correspond to the two lone pairs on the oxygen atom. Figure taken from Bulheller *et al.*^[70].

The amide bond in peptides shows characteristic electronic transitions (Figure 2.12b), which can also be observed in a CD spectrum: The $n \rightarrow \pi^*$ transition appears around 220 nm with a negative band. The $\pi \rightarrow \pi^*$ transition is observed as a second negative band around 210 nm and a strong positive band around 192 nm^[70]. The secondary structure closest to these characteristic transitions is the α -helix (Figure 2.12a). Only the two negative bands differ by 2 nm and appear at 222 and 208 nm. In a β -sheet, only one negative band is observed around 216 nm and the positive band is slightly shifted to 195 nm^[71]. The spectra of a β -turn is almost inverted. It shows a strong positive band at 202 to 206 nm and a negative one at 182 to 189 nm^[71,72]. In case of poly-prolines, the typical transition bands are observed around 200 to 205 nm (strong, negative) and 225 to 230 nm (weak, positive)^[71,73,74]. In the conformation of a random coil, only one negative band around 195 to 200 nm can be found^[71].

2.2.2.2 CD Studies of P-Factor

For the CD studies, 60 μM solutions of P-factor in water (pH 7.0), acetate buffer (pH 4.8) and different concentrations of tetrafluoroethylene (TFE) in water were prepared. The samples in water and acetate buffer were measured 4 times over a period of 3 hours to see if the conformation of the peptide is changing over time. The acetate buffer was chosen to mimic conditions used in the cell experiments. The yeast cells are grown and incubated in Edinburgh Minimal Media (EMM) which has a pH of 4.8. P-factor shows a

very low solubility in this media and thus, it was not possible to do the CD experiments in EMM. Instead, the acetate buffer was prepared with pH 4.8 and used alternatively. The samples containing TFE were only measured once each. TFE is known to induce α -helix formation in peptides^[75–77] and is therefore a useful tool to establish if the formation of this structural motif is possible in the given peptide sequence. The CD spectra of P-factor in both water and acetate buffer showed only one negative band with its minimum at 203 nm and 200 nm, respectively (Figure 2.14a,b). As discussed in section 2.2.2.1, this interprets to the conformation of the peptide being a random coil in these solvents. By analysing the results using the BeStSel algorithm^[78], it showed that four main structural features can be detected in the conformation of P-factor: helix, strand, turn and random coil. Indeed, the most abundant motif is the random coil, which takes up over 50% in water. The least present is the turn with only up to 5% (Table 2.2). In acetate buffer, the conformation is slightly changed. The distribution is fluctuating over time between the strand and the random coil being the most abundant structural feature. The percentage of the turn motif is also increased and fluctuates up to 15% (Table 2.3). Due to the lower pH in this solvent, some amino acid side chains might be deprotonated, inducing electrostatic interactions between the residues and thus, leading to conformational changes. When studying the TFE containing samples, the ratio of the helix motif increased with increasing TFE concentration. Already in the CD spectra, this phenomenon was detectable as the curves developed into the characteristic shape with a positive band at 194 nm and two negative bands at 221 and 208 nm (Figure 2.14c). The BeStSel algorithm later confirmed this observation as the percentage of the helix motif increased to over 40% (Table 2.4).

In conclusion, the CD studies revealed that P-factor does not adopt a stable conformation in solution. Even though the most abundant structural feature is the random coil, the peptide can be forced into more structurally ordered conformations depending on its environment. Different solvents or pH regions can influence the short-range local interactions and lead to the disruption or development of specific secondary structures. Computational studies done earlier in the group⁵ suggest an α -helix in the N-terminal region of the peptide between residue 2 and 13, and a loop structure around the proline residues in the C-terminus (residues 17 to 20, Figure 2.13). As discussed in section 2.1.2, the latter was already expected due to the high similarity to α -factor of *S. cerevisiae*^[36] and the amino acid sequence in this region^[50]. It was reported before that P-factor is establishing this conformation upon its interaction with the Mam2 receptor^[35,59]. The present results indicate that, similar to α -factor, the predicted structure is induced by the chemical environment close to or inside the receptor.

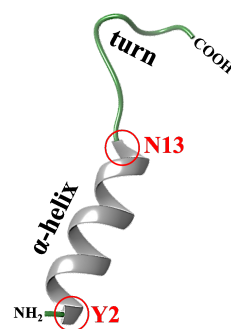


Figure 2.13: Secondary structure of P-factor as calculated in computational studies done earlier in the group⁶.

⁵ not published, see doctoral theses of André Oliveira Sequeira and Stavroula-Melina Sakellakou (2022)

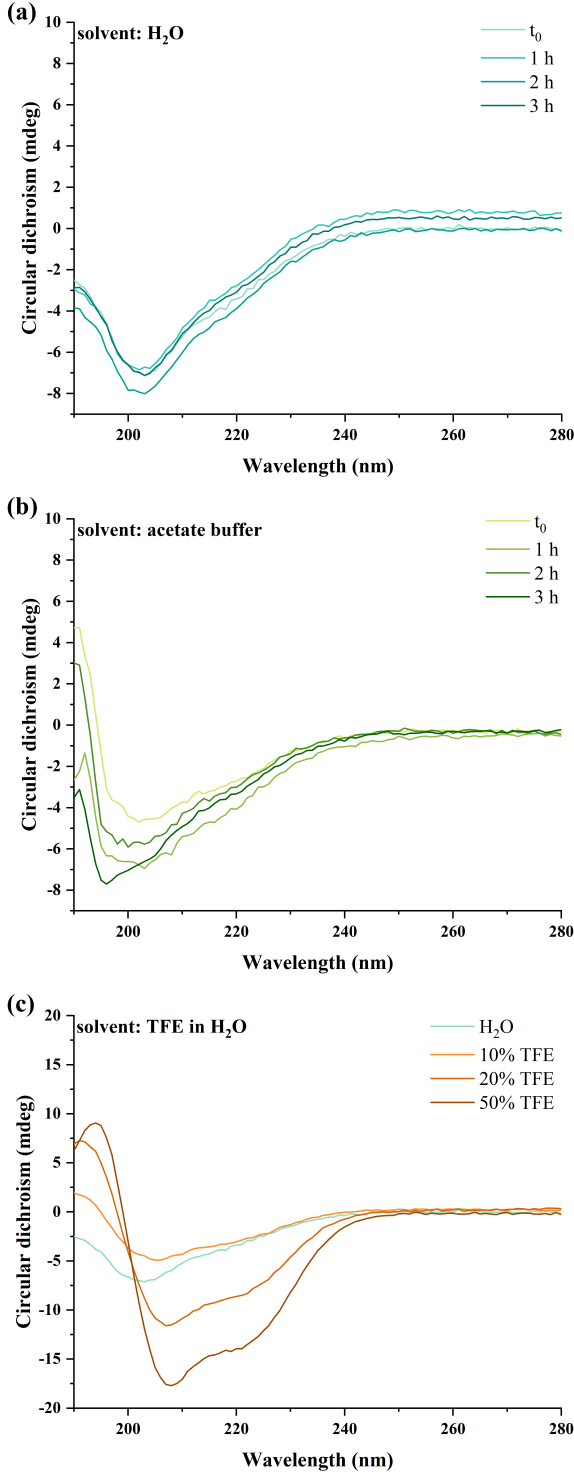


Table 2.2: Content of secondary structure motifs for P-factor in H₂O determined by BeStSel^[78].

Time	Content (%)			
	Helix	Strand	Turn	Coil
0 h	19.0	25.8	4.7	50.4
1 h	16.8	23.3	5.2	54.7
2 h	18.0	30.0	1.9	50.1
3 h	17.8	23.8	4.8	53.7

Table 2.3: Content of secondary structure motifs for P-factor in acetate buffer determined by BeStSel^[78].

Time	Content (%)			
	Helix	Strand	Turn	Coil
0 h	21.7	34.0	4.0	40.2
1 h	15.5	41.8	11.6	31.1
2 h	18.9	36.3	5.0	39.9
3 h	15.6	43.7	16.1	24.6

Table 2.4: Content of secondary structure motifs for P-factor in TFE/H₂O mixtures determined by BeStSel^[78].

Solvent	Content (%)			
	Helix	Strand	Turn	Coil
H ₂ O	19.0	25.8	4.7	50.4
10% TFE	25.2	19.2	10.5	45.2
20% TFE	39.7	22.3	-	38.0
30% TFE	41.5	18.7	-	39.8

Figure 2.14: CD analysis of P-factor in (a) H₂O, (b) acetate buffer and (c) TFE/H₂O mixtures.

2.2.3 Incubation of P-Factor with Fission Yeasts

The aim of this project was to create a reversible biological system to produce Albulin in fission yeast cells. As a prove of concept, it was important to analyse the expression of Albulin in *S. pombe* when incubated with unmodified P-factor. To quantify the Albulin expression, strain 1838 (compare section 2.1.3) was used as it contains the fusion protein Albulin-GFP. The cells were grown overnight in EMM in a shaker incubator at 37 °C and diluted into four batches of 0.2 OD_{595nm} the next morning. Two samples were treated with P-factor in a final concentration of 12 μ M. The other two batches were treated with sterile H₂O as control experiments. The cultures were incubated for another 6 hours. The EMM was then removed by centrifugation and several washings with sterile H₂O. For the fluorescence intensity measurements, the cultures were adjusted to 0.4 OD_{595nm} in H₂O and transferred to a 96 well-plate.

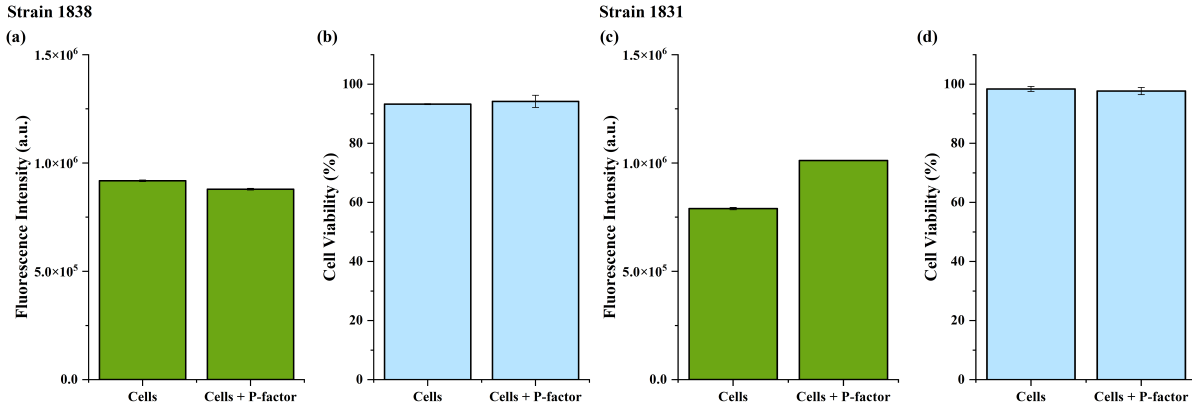


Figure 2.15: (a) GFP expression and (b) cell viability of strain 1838 after incubation with P-factor for 6 hours. (c) GFP expression and (d) cell viability of strain 1831 after incubation with P-factor for 6 hours.

The observed fluorescence intensity did not match the expected outcome of the experiment. It was assumed that the incubation with P-factor would lead to an activation of Mam2 and thus, to an expression of Albulin-GFP and an increase in fluorescence. However, the intensity decreased from 9.2×10^5 to 8.8×10^5 a.u. (Figure 2.15a). If the fusion protein is produced by the *S. pombe*, it stays inside the cell and gets diffused in the cytosol. In the presence of high amounts, it can start to form intracellular aggregates, which is assumed to result into photobleaching of the GFP. Due to the inconclusive results, the experiment was repeated with strain 1831. This culture does not express Albulin and has the reporter gene tagged to Mam2. In contrary to Albulin, the receptor gets integrated into the cell membrane, limiting intracellular aggregation. If the synthesized P-factor is effectively binding and activating the Mam2 receptor, the incubation is expected to induce the expression of Mam2-GFP and thus, to increase the fluorescence. After 6 hours of incubation at 32 °C, the intensity increased from 7.9×10^5 a.u. in the control experiment to 10.1×10^5 a.u. (Figure 2.15c).

A factor that has to be taken into account is the auto-fluorescence of *S. pombe* cells. In previous experiment done in the group, it was noticed that the organisms express a green fluorescence, which even increases upon cell death. This phenomenon can falsify the results of the fluorescence intensity measurement. Thus, a second experiment was performed in which the cells were subjected to Acridine Orange (AO) and Propidium Iodide (PI) in order to investigate the cell viability^[79]. Only AO can penetrate both dead and living cells, as it acts as an nucleic acid intercalator. The dye creates a green fluorescence in the cells. PI is a negatively charged molecule and is thereby repulsed by the cell membrane. Thus, it can only stain dead cells and gives a red fluorescence. In case of being stained by both dyes, the two fluorophores undergo a Förster resonance energy transfer and the AO emission was absorbed by the PI, leading to the fluorescence of only red light^[80]. In combination, the two dyes visualize dead and living cells in two distinct colours, making it possible to quantify the viability of the culture^[79]. In case of strain 1838, the viability of the cells at time 0 and after 6 hours was confirmed to 94% (Figure 2.15b). Strain 1831 showed a viability of 98% (Figure 2.15d). Thus, changes in the green fluorescence in the intensity measurement can be attributed to GFP.

2.3 Conclusions

The CD experiments indicated that the P-factor can be found in different conformations depending on its chemical environment. While in a neutral aqueous solution it predominantly exists as random coil, it is believed that the interaction with the Mam2 receptor will force the pheromone into a helical and turn like conformation. These experiments have to be repeated with the azobenzene-modified P-factor to examine if the modification has a significant impact on the pheromone's conformation. In a proof of concept, it was possible to show that the unmodified P-factor is able to bind Mam2. Due to probable photobleaching, it was not possible to confirm the expression of Albulin in *S. pombe*. Instead, the interaction was examined by the expression of Mam2-GFP. This was especially important as it verifies that the cells can be activated by synthetically derived P-factor. In the future work of this project, the incubation of P-factor with Mam2-GFP expressing strain 1831 can be used as a reference for experiments with the azobenzene-modified pheromones. It will be of high interest to compare the fluorescence intensity to that of incubations with the photoswitchable pheromones. The most desired outcome would be a similar intensity for one azobenzene isomer and a significantly decreased intensity (most favourable to the level of the water control) for the other one.

Chapter 3: Azobenzenes

To achieve reversibility in the biological system, it was decided to introduce a photo-switch into the backbone of P-factor. As already shortly discussed in section 1.3, an azobenzene was chosen for this purpose. This chapter is, at first, giving more theoretical background about azobenzenes (section 3.1). The section is going into more detail on the photochemical characteristics of azobenzenes in biological systems and explains known synthesis methods for symmetric and asymmetric azobenzenes. Subsequently, section 3.2 is focusing on the target compound by first discussing the study of its synthesis and then exploring its photophysical characteristics and comparing it to those of a symmetric *o*-TFAB. Additionally, an attempt to couple the synthesized azobenzene to amino acids was performed as a first step towards the integration into the P-factor. The attempted approaches are also discussed throughout section 3.2. A final section is concluding the chapter with a short summary and perspective (section 3.3).

3.1 Introduction

3.1.1 Photochemical Characteristics and Isomerization of Azobenzene

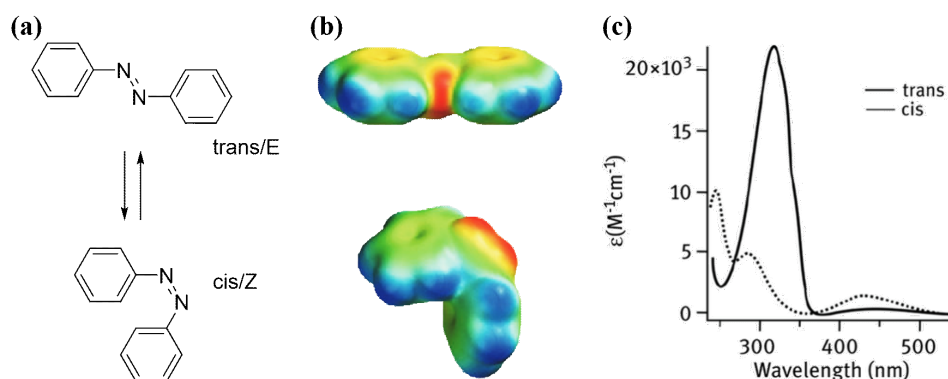


Figure 3.1: Isomeric structures of azobenzene and the corresponding absorption spectra: (a) The structures of azobenzene in *trans* and *cis* form. (b) Spacefilling models of the two isomers coloured by electrostatic potential (red-negative to blue-positive). (c) Absorption spectra of azobenzene *trans* and *cis* dissolved in EtOH. Figure taken from Beharry and Woolley^[18].

Azobenzenes are among the most studied photoswitchable molecules. After their discovery, the research was focussing on its suitability as synthetic colouring agent. By modifying the nature and size of the ring or integrating different substituents, researchers achieved the tuning of the molecule's colour and shade, giving the opportunity to generate cheap and reliable dyes^[18]. Thanks to this research, there is a high amount of available synthetic strategies. In 1937, Hartley^[81] observed the *trans* to *cis* isomerization for the first time. This finding changed the research interest, leading to more groups examining the

molecule's photochemistry^[82]. The two isomeric forms of an unmodified azobenzene and their absorption spectra are shown in Figure 3.1. The *trans* isomer has an intense π - π^* band around 320 nm and a weak n - π^* band around 440 nm. By irradiation near the peak wavelength at 340 nm, the molecule can be transferred into its *cis* isomer showing a

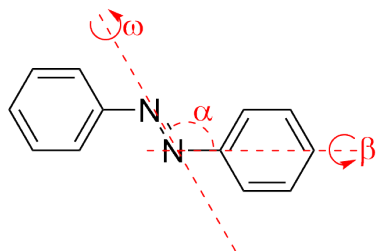


Figure 3.2: Geometry of azobenzene and angles discussed in this thesis. Figure adapted from Dokić *et al.*^[83].

decreased and slightly shifted π - π^* band around 280 nm. In contrast, the n - π^* band does not shift but shows a higher absorption than that of the *trans* isomer, giving the opportunity to use this transition band for a photoinduced isomerization from *cis* to *trans*^[18]. As the *trans* form is thermodynamically more stable^[84], the molecule will take this conformation at equilibrium in the dark. The thermal half-life of the *cis* form was determined to be 2 days^[85]. The isomerization also leads to a geometrical transformation. While the *trans* isomer is planar with a dihedral angle ω of 180° and a rotational angle β of 0° (Figure 3.2), the *cis* form is more compact with its phenyl rings rotated by 55° (β) and a dihedral angle ω of 10° ^[18,83]. Due to the nature of the diazo double bond, both conformations have an inversion angle α of 120° . The isomerization can follow four mechanisms: rotation, inversion, concerted inversion and inversion-assisted rotation (Figure 3.3)^[85]. As the name suggests, in the first mechanism, a rotation around the diazo bond occurs, affecting the dihedral angle ω while α remains stable. To allow

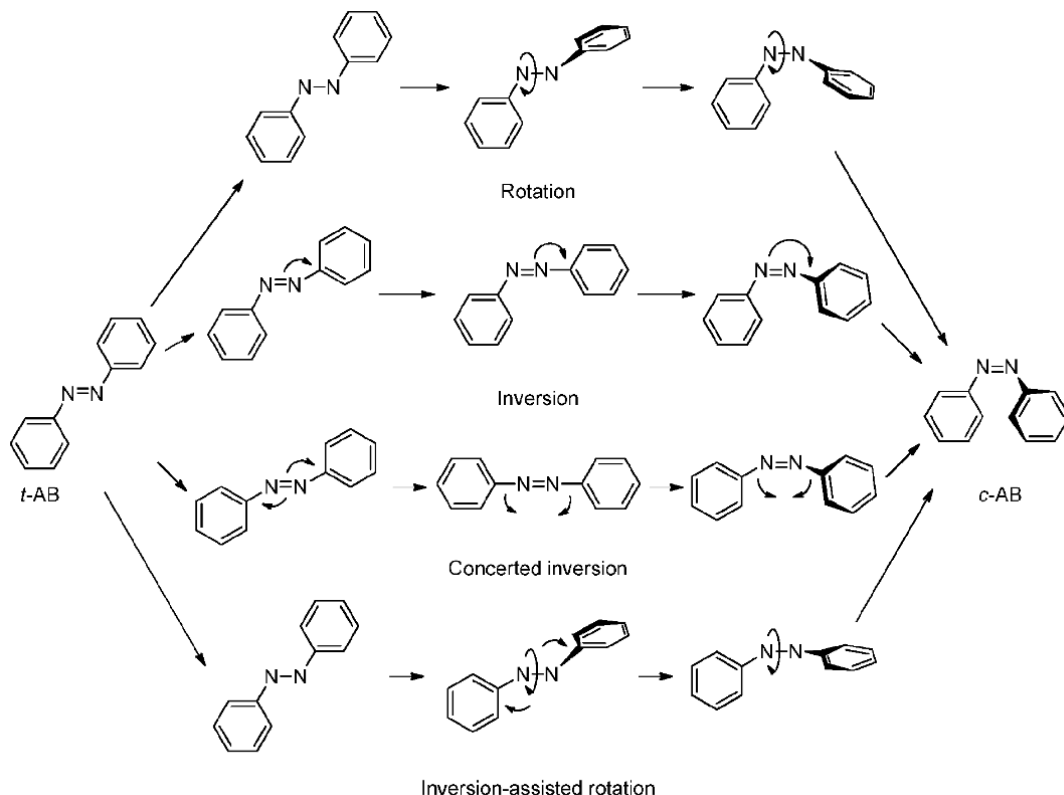


Figure 3.3: The four proposed mechanisms of photoisomerization. Figure taken from Bandara and Burdette^[85].

the rotation, the pathway requires the N=N π bond to be ruptured. During the inversion mechanism, ω stays fixed while one of the inversion angles α increases to 180° in the transition state. The concerted inversion involves both inversion angles to increase to 180° in the transition state, leading to a linear conformation. The last mechanism consists of changes in both the dihedral angle ω and one inversion angle α , whereby the more profound change is occurring in ω ^[85]. All the transition states can relax into both the *trans* and the *cis* isomer, which is why the PSS from all four pathways bear both forms.

3.1.2 Tailoring Azobenzenes to Biological Applications

Different researches indicate that the isomerization mechanism and quantum yield can be influenced by, e.g., the chosen wavelength^[86], the solvent^[83,87] and substituents on the phenyl rings^[85]. The latter is also leading to a change in the absorption spectra of the azobenzene, as it influences the electronic distribution of the molecule^[88]. As mentioned above, the non-substituted azobenzene can be isomerized to its *cis* and *trans* form by irradiation at 340 nm and 440 nm, respectively. To use the azobenzene in a biological system, isomerization wavelengths in the visible light region are necessary. The longer wavelengths can more easily enter the cells without being absorbed by biomolecules present in the organism (like NADH, nucleic acids or proteins)^[18,89]. The incorporation of an electron donating group (EDG) (e.g. amines or hydroxides) in *para* or *ortho* position red-shifts the π - π^* band leading to a partial overlap with the n- π^* band (Figure 3.4)^[88,90]. This effect can be even increased when introducing a push-pull system into the compound, creating pseudo-stilbene-like azobenzenes. In order to do so, an EDG is attached at the *para* position of one phenyl ring and an electron withdrawing group (EWG) exactly on the opposite site. Due to the more delocalized π electrons, less energy is required to induce the π - π^* transition, thus leading to higher wavelengths of the respective absorption band^[88]. However, an overlap of the transition bands results into a low quantum yield as both isomerizations are induced with similar wavelengths^[90]. Furthermore, the higher electron density in the N=N π orbital yields into a drastically decreased thermal half-life of the *cis* isomer in the range of milliseconds to minutes^[85,90,91]. To effectively control a reversible biological system, such short half-lives and the low quantum yield are not desirable^[18,90]. A different approach to change the absorption spectra is by focussing on the n- π^* transition band. In the non-substituted azobenzene both isomers absorb around 440 nm. Designing a molecule where the two isomers have distinct n- π^* bands would open the opportunity to

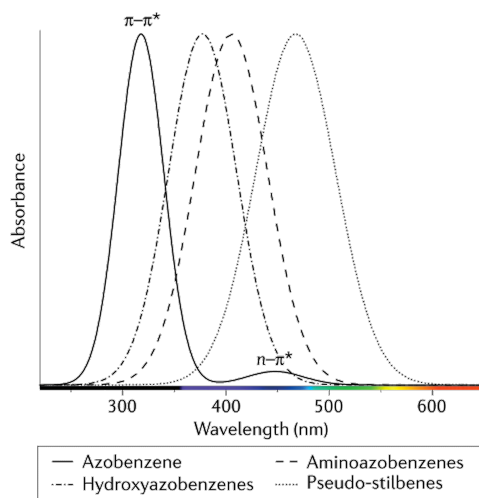


Figure 3.4: Influence of *para* substituents on the π - π transition band of azobenzene. Figure taken from Jerca *et al.*^[90].

use this transition for the isomerization in both directions^[29]. Progress in this direction was made by introducing *ortho* substituents in the phenyl rings^[18]. Dong *et al.*^[30] showed that both *o*-diaminoazobenzene and *o*-tetramethoxyazobenzene have sufficiently separated $n\text{-}\pi^*$ bands. They explained this phenomena by electronic and steric effects coming from the substituents. In case of the methoxy substituted compound, the *trans* isomer experiences electronic interactions between the methoxy groups and the lone pairs on the nitrogen in the azo bond, leading to a destabilization of the n -orbital and a distortion of the planarity. After isomerizing to *cis*, the two groups have a higher spatial distance, disrupting this interaction. Thus, only the *trans* $n\text{-}\pi^*$ band is influenced by the *ortho* substitution, leading to a red-shift compared to the *cis* isomer. Furthermore, the group determined the thermal half life of the *cis* isomer to be around 2 days, making it suitable for the application in a controlled reversible biological system. However, methoxy groups are known to not be stable in reductive environments like the intracellular matrix^[92]. Indeed, the *o*-tetramethoxyazobenzene was reduced within a few hours when subjected to glutathione^[30]. By using fluorines instead of methoxy groups this phenomena can be avoided. Bleger *et al.*^[29] report similar results for the separation of the $n\text{-}\pi^*$ transition bands (Figure 3.5) and a thermal half-life of the *cis* isomer of up to 2 years when working with an *o*-TFAB. They attribute this to a reduction in the electron density of the $\text{N}=\text{N}$ bond due to the electron withdrawing character of the fluorines, resulting in the n -orbital having a lower energy. In comparison to the methoxy-substituted azobenzene, the fluorinated compound is less distorted, as the fluorine atom is smaller than the methoxy group. Photoisomerization was achieved effectively by using visible-range light with high PSS of over 85% in both directions. The group has already studied the effects of additionally introduced EDG and EWG on the UV-vis spectra. Again, substitution with amides in *para* position (weakly electron donating) leads to a red-shift of the $\pi\text{-}\pi^*$ band, but at the same time the $n\text{-}\pi^*$ separation and the PSS decreased. The electron donating effect increases again the electron density in the $\text{N}=\text{N}$ bond, thereby counterbalancing the effect of the fluorines. On the other hand, the introduction of esters in the *para* positions even

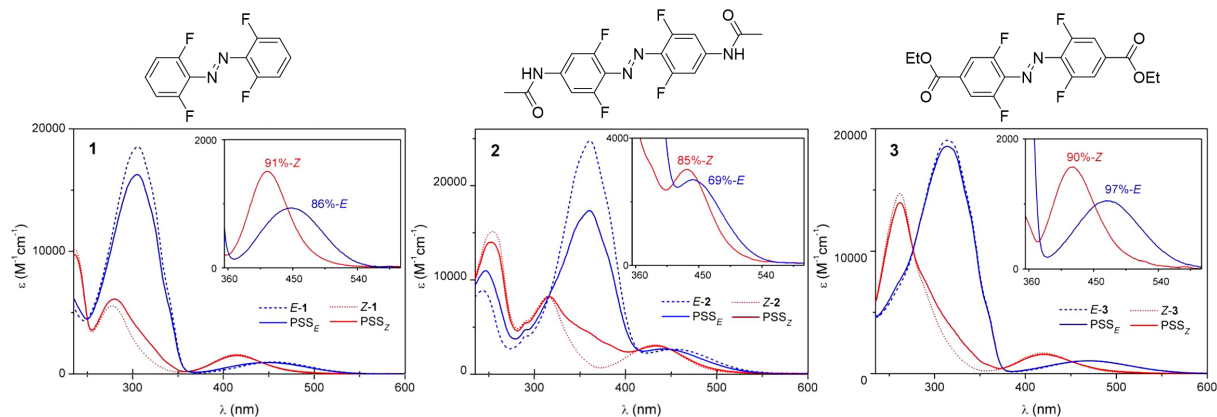


Figure 3.5: Absorption spectra of *o*-TFAB derivatives. Figure adapted from Bleger *et al.*^[29].

enhances the $n\text{-}\pi^*$ separation by further destabilizing the n -orbital. Additionally, the PSS increased to over 90% in both directions.

In conclusion, the *o*-TFAB derivatives are promising compounds to reach the aim of this thesis. Bleger *et al.*^[29] so far only investigated the effect of symmetric substitutions (Figure 3.5). However, in the course of this project, the azobenzene shall be integrated into the backbone of a peptide. In order to do so, it is necessary to introduce a carboxylic acid in the *para* position of one phenyl ring and an amino group at the opposite site (Figure 3.11, page 32). Upon the insertion into the backbone, both groups will be converted into amides.

3.1.3 Methods of Synthesis

Depending on the structure of the target azobenzene, there is a variety of synthetic approaches to be followed. The most common procedures are the azo coupling and the Mills reaction, which are of special interest for the generation of asymmetric compounds^[93,94]. The following sections will go into more detail on these strategies. Additionally, the Wallach rearrangement will be introduced, which can be used to form azobenzenes in strong acidic conditions^[95]. Further approaches include reductive and oxidative mechanisms to build the target compound from aromatic nitro derivatives or anilines. These strategies are mainly suitable for symmetric azobenzenes^[93].

3.1.3.1 Azo Coupling

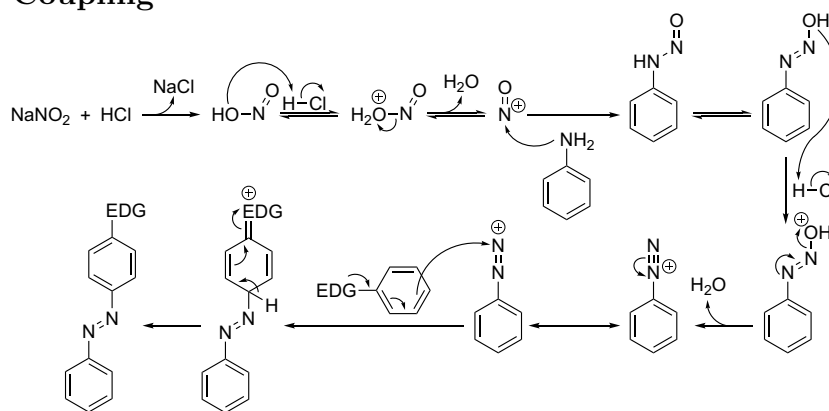


Figure 3.6: Azo coupling mechanism. Figure adapted from Merino^[93].

The standard procedure to synthesize asymmetric azobenzenes is following the azo coupling mechanism. An aniline is first converted into a diazonium salt and is then reacted with the second aromatic moiety to undergo an electrophilic aromatic substitution (Figure 3.6)^[93]. Typically NaNO_2 is used as a nitrosating agent, which is converted into a nitrosonium ion in the first step. The positively charged ion is then attacked by an aniline to give *n*-phenyl-nitrosamine, which is further protonated and, under the release of H_2O , forms an aryl diazonium salt^[19]. As the diazonium group is strongly electron attracting^[96], it can be attacked by the second aromatic species to form the azo group between the two rings. This reaction is more facilitated if an EDG is present in the ring. The azo bond is then formed preferably in the *para*, or if not available in the *ortho*, position of the

EDG. One important factor in the mechanism is the right choice of pH. The formation of the nitrosonium ion and later the diazonium salt only occur in acidic pH (Figure 3.6). On the other hand, the coupling reaction is more favoured in slightly basic conditions since, in most cases, the EDG can get protonated in acidic pH. Like that, their electron donating strength is decreased and the reaction is slowed down, or even stopped. However, if the environment is too basic, the diazonium salt is converted into the diazohydroxide^[93].

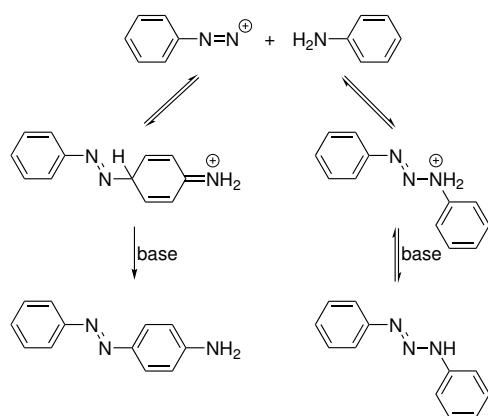


Figure 3.7: Competition between *C*- and *N*-attack. Figure adapted from Mayr *et al.*^[94].

of aromaticity. However, to convert the intermediate into the final species, a deprotonation is necessary. By maintaining a slightly acidic pH, this can be prevented and the formation of the azobenzene can be observed^[94].

A special case occurs when the second aromatic moiety is an aniline. In this case, the electrophilic aromatic substitution competes with a *N*-coupling, which is kinetically more favoured and gives a triazene as final product (Figure 3.7). The kinetic preference can be explained by the formed intermediate that has the positive charge in a more favourable position than the intermediate evolving from the *C*-attack. Furthermore, the latter leads to a disruption

3.1.3.2 Mills Reaction

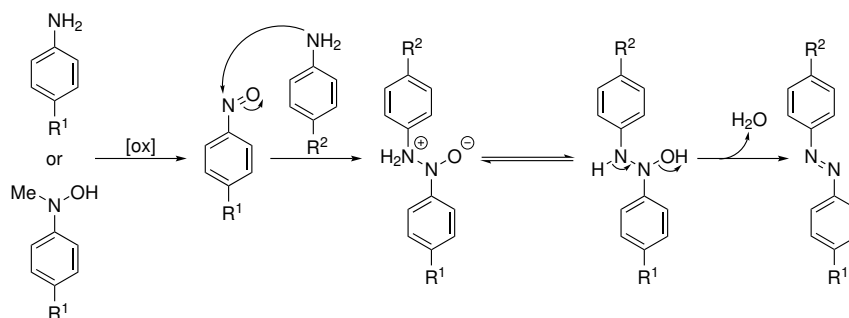


Figure 3.8: Mechanism of Mills reaction. Figure adapted from Merino^[93].

Azobenzene synthesis can be achieved following the Mills reaction conditions. It starts with the formation of an aromatic nitroso derivative by oxidation of e.g. an aniline or an aromatic methyl- hydroxylamine. A second aniline can attack the nitroso group to give the azobenzene after a dehydration. To diminish the reaction between already oxidized and non-oxidized aniline, a two-layered system can be used. While most anilines are soluble in aqueous media, the nitroso derivative will move into the organic phase^[93]. In case of the synthesis of a symmetric azobenzene, this separation is not necessary, as the reaction between the aniline and nitroso derivative of the same aromatic species is desired.

3.1.3.3 Wallach Rearrangement

Azobenzenes can also be formed from their corresponding azoxybenzene by undergoing a Wallach rearrangement. Strong acidic conditions are necessary to enable this reaction as they allow the formation of a dicationic intermediate from an only weakly basic starting material (Figure 3.9). The reaction in H_2SO_4 gives *p*-hydroxyazobenzene as main product. However, depending on the reaction conditions also *o*- and non-substituted compounds are observed^[93]. In a temperature range of 240-250 °C, the last one is generated as major

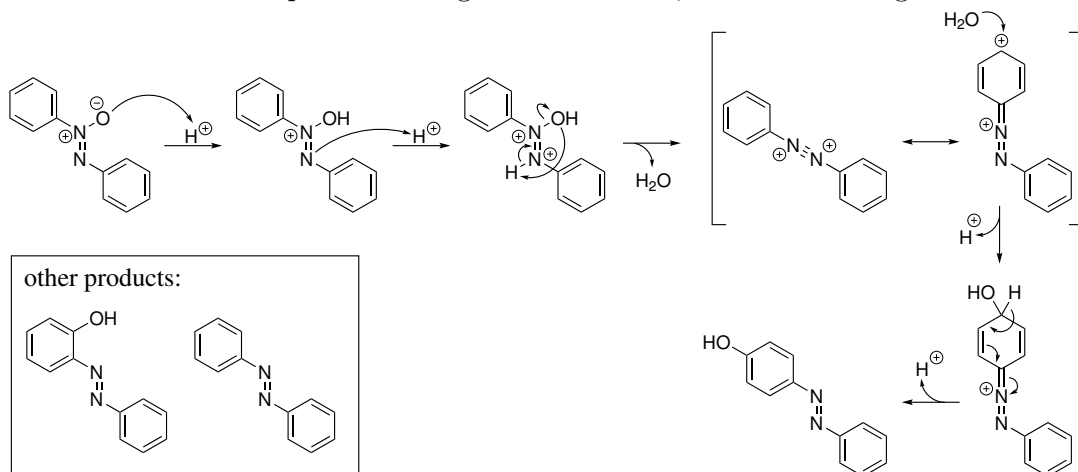


Figure 3.9: Mechanism of Wallach rearrangement to form *p*-hydroxyazobenzene.

product. Lalitha *et al.*^[95] report that also the strength of the acidic environment influences the product ratio. They achieved the generation of the non-substituted azobenzene as major product when heating azoxybenzene with the acidic zeolite HY at 120 °C. Zeolites are solid acids that can be used to effectively control the acidity of a reaction^[97]. The authors suggest a mechanism in which the protonated azoxybenzene can undergo a reduction during the extraction with CH_2Cl_2 . Traces of *p*- and *o*-hydroxyazobenzene are also observed. They propose a mechanism without the dicationic intermediate. Instead, the group describes the formation of a cyclic intermediate after protonation of the azoxybenzene (Figure 3.10).

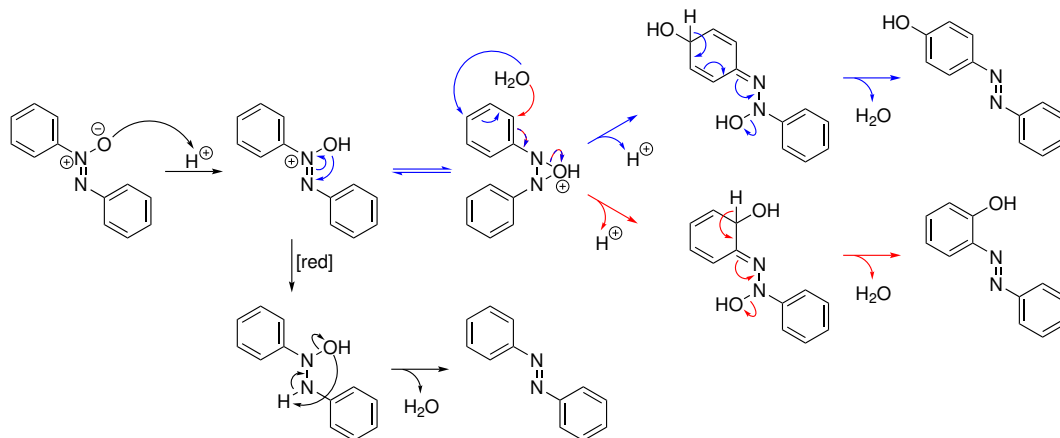


Figure 3.10: Mechanism of Wallach rearrangement when using zeolites as acids. Figure adapted from Lalitha *et al.*^[95].

3.2 Results and Discussion

3.2.1 Targeted Synthesis: HOOC-*o*-TFAB-NH₂

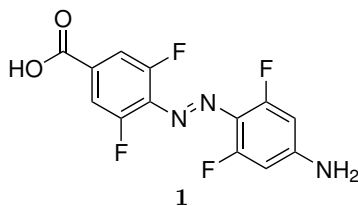


Figure 3.11: Target compound **1**: HOOC-*o*-TFAB-NH₂.

The target compound is HOOC-*o*-TFAB-NH₂ **1** (Figure 3.11). The asymmetric azobenzene bears four fluorines in *ortho*-position of the diazenyl bond and, additionally, a carboxyl and an amino group in the *para* positions. As described in section 3.1.3, the main strategy to synthesize an asymmetric azobenzene is by performing an azo coupling between an aromatic moiety as donor and a diazonium as acceptor. This gives in principle two options for the synthesis of compound **1**, which are illustrated in Figure 3.12. As the donor moiety needs to be electron rich, an electron withdrawing carboxyl group would prevent the reaction from occurring. Furthermore, forming a diazonium species from an asymmetric aminoaniline would give a mixture of isomers as the diazonium can form on either side of the aromate. Thus, the 4-amino-3,5-difluorobenzoic acid was chosen as acceptor and the 3,5-difluoroaniline as donor.

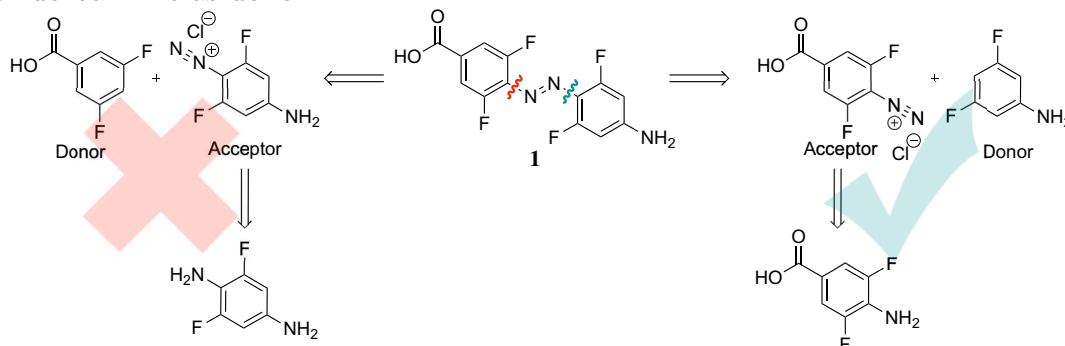
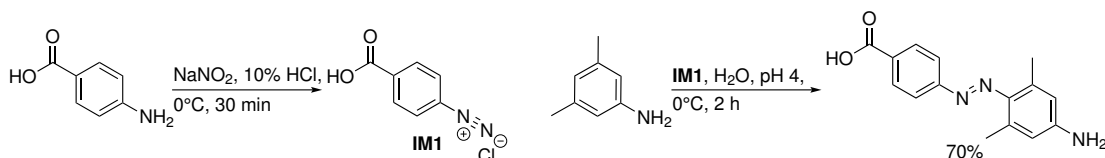
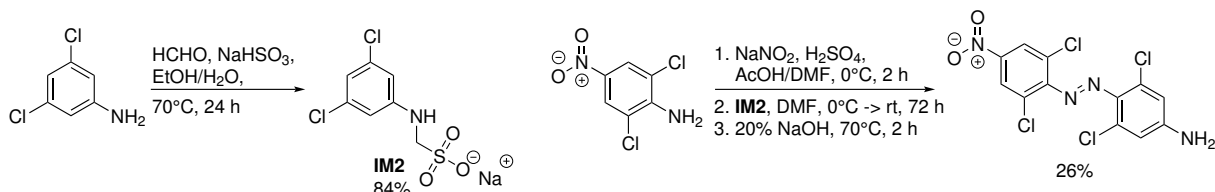


Figure 3.12: Two possible routes to synthesize HOOC-*o*-TFAB-NH₂ **1**.

In general, two procedures were used throughout this work and adapted accordingly to the chosen strategy. The first one is reported by Prestel and Möller^[98] and comprises the azo coupling between 4-aminobenzoic acid as acceptor and 3,5-dimethylaniline as donor in H₂O at pH 4 using HCl as an acid (Scheme 3.1a). To improve the solubility, a second procedure that was already applied by Samanta *et al.*^[27] was performed. The authors chose DMF as solvent, and acetic acid and H₂SO₄ as acids to couple 2,6-dichloro-4-nitroaniline and 3,5-dichloroanilinemethanesulfonate (Scheme 3.1b). While Prestel and Möller^[98] slowly added the freshly generated diazonium species to the dissolved aniline derivative, Samanta *et al.*^[27] changed the order of reactants and added the aniline species to the diazonium

(a) Synthesis Strategy by Prestel and Möller:

(b) Synthesis Strategy by Samanta *et al.*:

Scheme 3.1: Azo coupling strategies as proposed by (a) Prestel and Möller^[98], and (b) Samanta *et al.*^[27]. These procedures were used as templates and adapted throughout the thesis for the synthesis of compound **1**.

solution. Both strategies can have advantages. The diazonium is more stable at acidic pH^[93] and the aniline derivative is dissolved in neutral aqueous solution. Thus, slow addition of the diazonium species to the aniline solution ensures the reaction before the degradation of the diazonium moiety. On the other hand, the diazonium solution might have some unreacted nitrosating agent if there was no full conversion in the first step. Slow addition of the solution to an excess of 3,5-difluoroaniline can lead to the formation of another diazonium species, which undergoes azo coupling with a second aniline to give side product **2** (Figure 3.13).

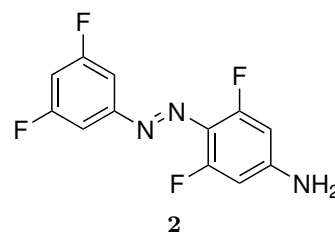
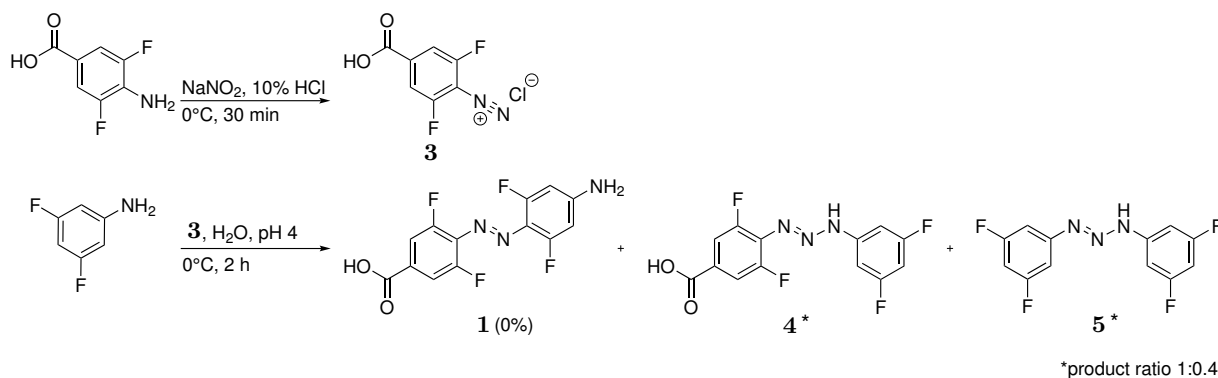


Figure 3.13: Possible side product: azobenzene **2**.

3.2.1.1 Strategy A: A Detour via Triazenes



Scheme 3.2: Synthesis strategy A1: Synthesis of azobenzene **1** from 4-amino-3,5-difluorobenzoic acid and 3,5-difluoroaniline. Target compound **1** is not observed. The procedure is adapted from Prestel and Möller^[98].

The first attempt to synthesize $\text{HOOC-}o\text{-TFAB-NH}_2$ **1** was adapted from Prestel and Möller^[98] (Scheme 3.2). The benzoic acid derivative was reacted with NaNO_2 at low temperatures in a 10% HCl aqueous solution. The formation of diazonium species **3** could be easily followed as the light brown suspension evolved into a bright yellow

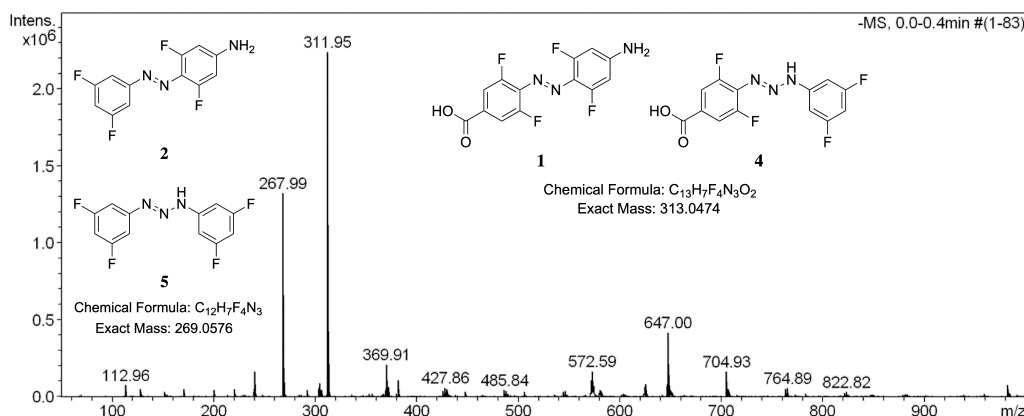


Figure 3.14: MS analysis of coupling between 4-amino-3,5-difluorobenzoic acid and 3,5-difluoroaniline. The spectra indicates the formation of azobenzenes **1** and **2**, or triazenes **4** and **5**.

low solution within 30 minutes. Addition of diazonium salt **3** to a solution of the aniline in water, a light orange solid precipitated instantly. As discussed in section 3.1.3, a mild acidic pH needs to be maintained to prevent *N*-coupling. After complete consumption of the starting materials (2 hours), the reaction was neutralized. Due

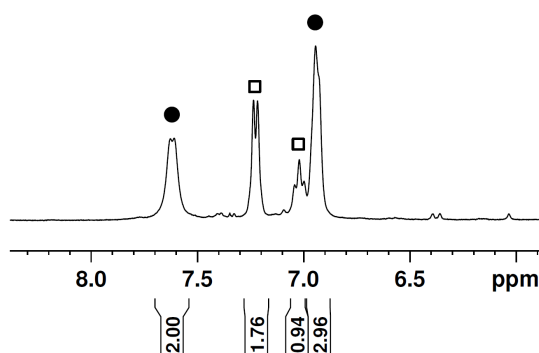
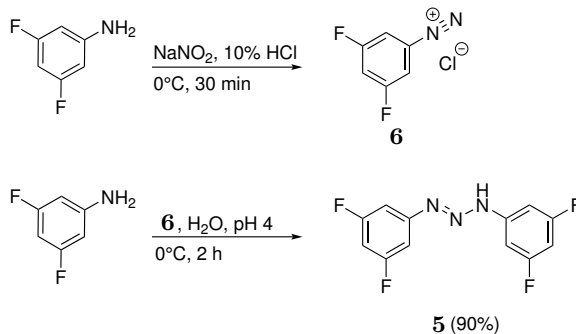
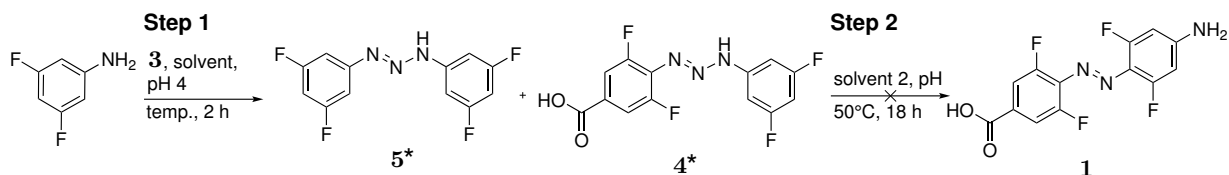


Figure 3.15: ^1H NMR spectrum showing a mixture of triazenes **4** (●) and **5** (□).

analysis revealed the formation of triazenes **4** and **5** (Figure 3.15). To get better insights and to confirm the formation of triazene **5**, the latter was synthesized using the same procedure as above but forming the diazonium from 3,5-difluoroaniline (Scheme 3.3). Both NMR and MS analysis confirmed the product formation.



Scheme 3.3: Synthesis of main side product **5**. The procedure is adapted from Prestel and Möller^[98].



Scheme 3.4: Synthesis strategy A2: Synthesis of triazene **4** and subsequent conversion into azobenzene **1**. Compound **3** is synthesized as shown in Scheme 3.2.

*Ratio of triazenes **5** and **4** as reported in Table 3.1.

Table 3.1: Optimization of step 1: *N*-coupling to give triazene **4**.

Entry	Solvent	eq (NaNO ₂)	Order	Temp. (°C)	Ratio (5 / 4)* ¹
1	H ₂ O	1	3 to aniline	0	0.4:1
2	DMF	1	3 to aniline	0	* ²
3	THF	1	3 to aniline	0	* ²
4	H ₂ O	1	aniline to 3	0	0.2:1
5	H ₂ O	0.9	3 to aniline	0	0:1
6	H ₂ O	0.9	3 to aniline	RT	0:1

*¹determined by NMR

*²not determined, presence of side product **5** confirmed by mass

As already discussed in section 3.1.3, the formation of triazenes is reversible and it is reported that in low pH and at higher temperatures azobenzenes are the thermodynamically

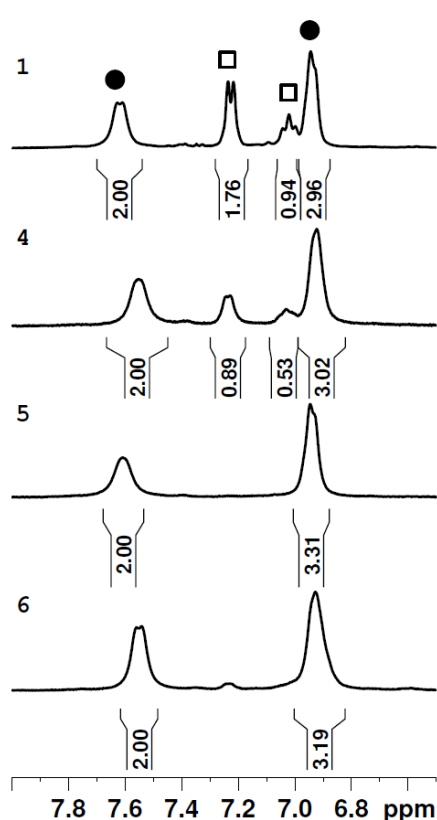


Figure 3.16: ¹H NMR spectra showing a mixture of triazenes **4** (●) and **5** (□) resulting from different reaction conditions. Numbering of the spectra refers to the entries of Table 3.1.

favoured products^[99,100]. Thus, the synthetic strategy was extended by an additional step to convert an isolated triazene into the respective azobenzene (Scheme 3.4). Before focusing on the rearrangement (step 2), the triazene synthesis needed to be optimized to diminish the formation of side product **5**. The presence of **5** means that there is still nitrosating agent in solution when the reaction mixture is added to the aniline species. In a first attempt, the solvent was changed to increase the solubility, and thus the reactivity, of the benzoic acid species. Both DMF and tetrahydrofuran (THF) led to complete dissolution but no improvement of purity (Table 3.1, entries 2 and 3). The second approach was to change the order and to slowly add the aniline derivative to diazonium species **3** (entry 4). This led to a decrease in side product formation by 50% (Figure 3.16, spectrum 4). Finally, decreasing the equivalents of NaNO₂ gave compound **4** without side product **5** in quantitative yield (entry/spectrum 5). It was further observed that the second step does not need to be

performed at 0 °C (entry/spectrum 6). The ^1H NMR spectrum still showed impurities. However, as in the thin-layer chromatography (TLC) only one spot was visible, it was decided to not do a column chromatography but to recrystallize the product. In a first attempt, the crude was dissolved in MeOH and then precipitated by addition of H_2O , which did not increase the purity. Further approaches included the crystallization from hot solvents. Neither a mixture of H_2O with MeOH or isopropanol, nor pure toluene, hexane or dioxane gave pure product. However, triazene **4** is not the final product. Thus, the crude was used in the next step, which was the transposition into target compound $\text{HOOC-}o\text{-TFAB-NH}_2$ **1** in acidic conditions^[101].

Table 3.2: Optimization of step 2: Rearrangement of triazene **4** to azobenzene **1**.

Entry	Solvent 2	pH	Conversion* ¹	1
1	H_2O	4	0%	n.d.* ²
2	H_2O	1	0%	n.d.
3	MeOH	4	17%	n.d.
4	MeOH	1	100%	n.d.

*¹determined by NMR

*²not detected

Four reactions were performed (Table 3.2): 50 mg of **4** were dissolved in 1 mL solvent (final concentration 0.16 M) and heated to 50 °C for 18 hours. The first reaction was performed in conditions oriented at the previously discussed step 1 of the synthesis strategy (entry 1). This would allow to add this second step without the need of purification or workup in between. In a second attempt, the pH was further decreased to generate a strong acidic environment (entry 2). As the triazene is better soluble in MeOH, the same reaction was repeated in this solvent (entries 3 and 4). None of the attempts gave target compound **1** (Figure 3.17). Only in MeOH, a conversion back to the starting materials was observed, confirming the reversibility of the triazene formation. However, as no azobenzene is obtained, it is assumed that the acidic conditions lead to a protonation of the aniline, consequently reducing the nucleophilicity of the aromatic core. As a higher pH would again favour the formation of a triazene over that of an azobenzene, other strategies have to be explored.

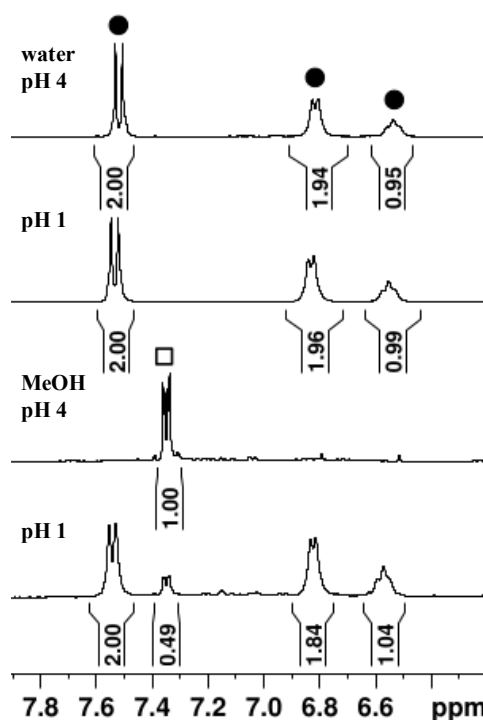
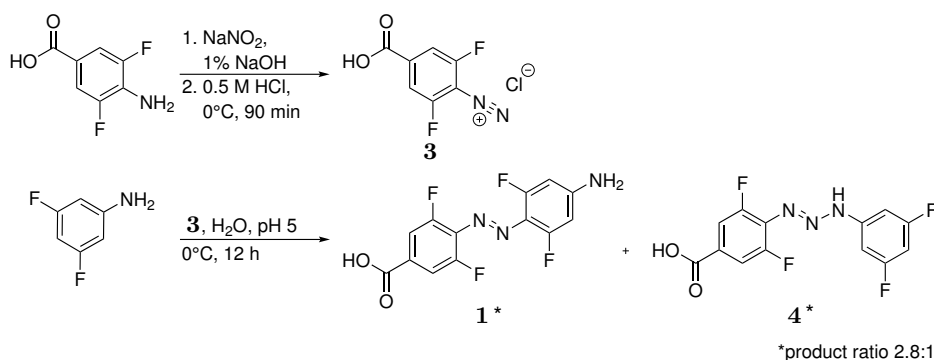


Figure 3.17: Comparison of ^1H NMR spectra during the optimization for the rearrangement from triazene **4** (●) to azobenzene **1**. There is no conversion to the azobenzene. Instead, in MeOH, the conversion back to the starting material (benzoic acid □) can be observed.



Scheme 3.5: Synthesis of target compound **1** following a patent published by Zhao *et al.*^[102].

While working on the targeted synthesis, Zhao *et al.*^[102] published a patent about the synthesis of the targeted HOOC-*o*-TFAB-NH₂ **1**. They are also performing an azo coupling in two steps by first generating a diazonium species and then coupling it to the aniline derivative. However, their procedure differs in terms of pH, reaction times and equivalents of the reaction partners^[102]. For the generation of diazonium species **3**, they first dissolve the benzoic acid derivative and NaNO₂ in 1% NaOH(aq) before acidifying with 0.5 M HCl(aq). It was assumed that this would increase the solubility and reactivity of the benzoic acid species and thus, lead to a decrease in the formation of an aniline dimer (to give triazene **5** or azobenzene **2**). Furthermore, the longer reaction time in the coupling step could lead to the formation of the thermodynamically more favored azobenzene **1**. Thus, the reaction was attempted with these conditions. After 12 hours, the formed orange solid was allowed to precipitate and then filtered. As shown in Figure 3.15, the proton peaks of triazene **4** are overlapping, therefore hiding the triplet pattern coming from the non-substituted *para* position. Since this triplet signal is indicative for the presence of triazene **4**, the NMR solvent was changed to MeOD. This solvent results in a different chemical shift of the triplet pattern making the peak distinguishable from the others. The ¹H NMR spectrum of the crude product showed a mixture of triazene **4** and azobenzene **1** in a ratio of about 2.8:1 (Figure 3.18). As it was not possible to get a sufficient separation in the TLC, it was decided to proceed with strategy B.

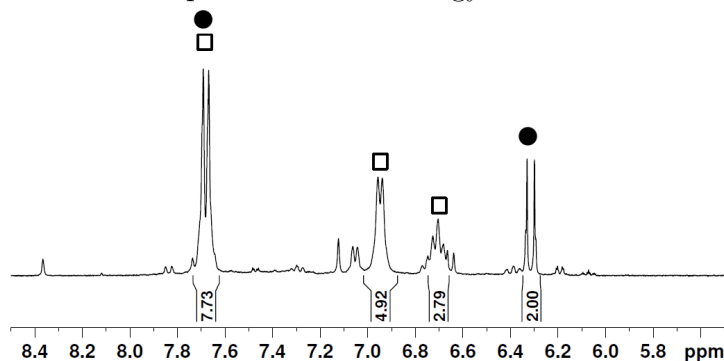
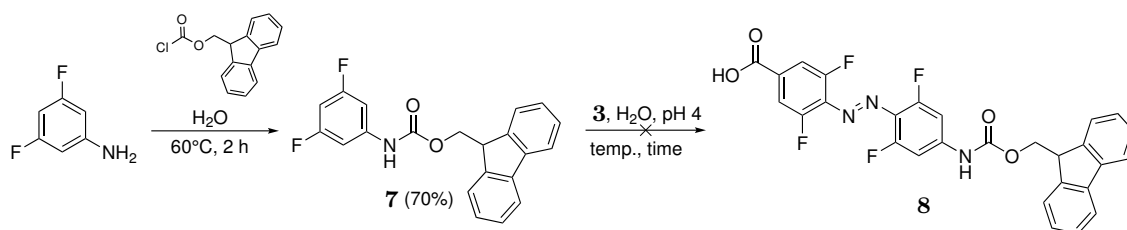


Figure 3.18: ¹H NMR spectrum showing a mixture of triazene **4** (□) and azobenzene **1** (●).

3.2.1.2 Strategy B: Introduction of Protection Groups



Scheme 3.6: Synthesis strategy B1: Protection of aniline with Fmoc before azo coupling in H₂O. Compound **3** is synthesized as shown in Scheme 3.2. The protection step is adapted from Gawande and Branco^[103], the azo coupling is adapted from the procedure proposed by Prestel and Möller^[98].

Table 3.3: Optimization parameters for synthesis strategy B1.

Entry	Temp. (°C)	Time (h)	8 ^{*1}
1	RT	2	n.d. ^{*2}
2	RT	72	n.d.
3	50	16	n.d.

^{*1}determined by MS/NMR analysis

^{*2}not detected

An easy way to prevent *N*-coupling of the aniline is to protect the amino group. As the azobenzene will be integrated into a peptide, it seems reasonable to protect the NH₂ group with a Fmoc protection group. This led to the synthesis strategy shown in Scheme 3.6. The procedure was established by Gawande and Branco^[103] and asks to react 3,5-difluoroaniline with Fmoc-Cl in water at 60°C. The group proved that the aqueous environment significantly improves yield and reaction time. The reaction can easily be followed by TLC and was complete after 2 hours. Addition of petroleum ether to a solution of the crude product in EtOAc allowed precipitation of protected aniline species **7** with 70% yield as pure product. The azo coupling was done using a similar procedure as already tested for the unprotected aniline derivative. As there is no free amino group in the second aromatic moiety, there is no need to decrease the equivalents of nitrosating agent, which is why 1 eq was used. The first attempts were again tried in H₂O, as this gave the most promising results in the triazene syntheses. However, no product formation was observed and increasing both reaction time and temperature did not change the outcome (Table 3.3).

A different approach is used by the Woolley group at the University of Toronto^[27]. Their target compound is an asymmetric *o*-tetrachloroazobenzene with a nitro and an amino group on the two *para* positions (Figure 3.19). Following their procedure (Scheme 3.7), diazonium **3** is formed over a period of 2 hours in a mixture of DMF, acetic acid and H₂SO₄. In the second step, a protected aniline derivative is dissolved in DMF and added

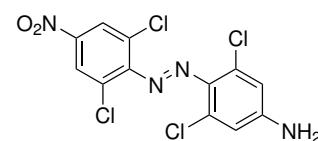
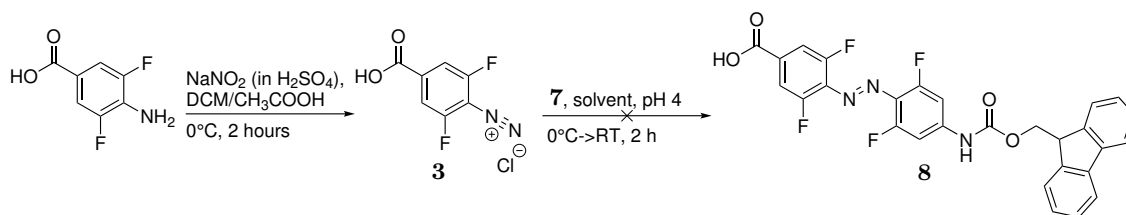


Figure 3.19: Tetrachloroazobenzene as synthesized by Samanta *et al.*^[27].



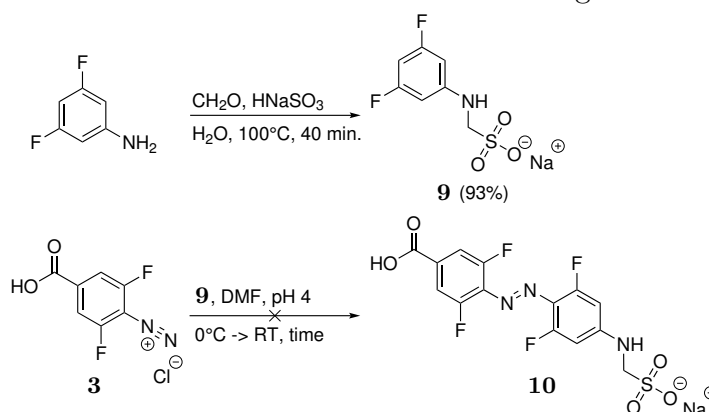
Scheme 3.7: Synthesis strategy B2: Protection of aniline with Fmoc before azo coupling in DMF or THF. Compound **7** is synthesized as shown in Scheme 3.6. The azo coupling is adapted from the procedure proposed by Samanta *et al.*^[27]

Table 3.4: Optimization parameters for synthesis strategy B2.

Entry	Solvent	eq (7)	8 *
1	DMF	0.5	n.d.* ²
2	DMF	1	n.d.
3	THF	1	n.d.

*verified by MS/NMR analysis

dropwise to the diazonium solution. After stirring for one hour at 0 °C, the reaction needs to run for 72 hours at room temperature (RT). Furthermore, they used a 2:1 ratio of diazonium **3** with respect to the aniline species. When applying this procedure, the reaction mixture turned into a dark red solution already after 2 hours. The reaction was quenched with ice-cooled H₂O, leading to the precipitation of an orange solid. MS analysis of the precipitate only indicated starting material and no traces of the product (Table 3.4, entry 1). The reaction was repeated with a 1:1 ratio and with THF as solvent. Both did not give the product (entries 2 and 3). The Fmoc protection group is an electron withdrawing substituent due to its carbamate structure. The nucleophilicity of the aromatic donor ring is thus even more decreased, which explains the low reactivity. Again, it is not possible to increase the pH as this would lead to Fmoc cleavage. It is therefore necessary to use a protection group that does not bear an electron withdrawing effect.



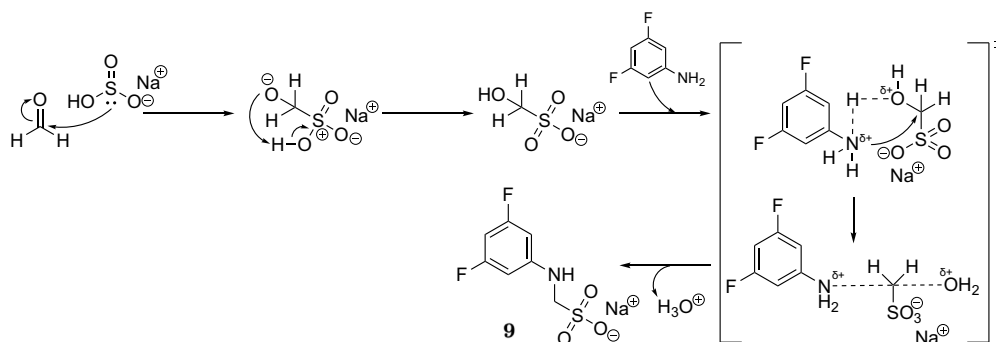
Scheme 3.8: Synthesis strategy B3: Protection of aniline as methyl sulfonate before azo coupling in DMF. Compound **3** is synthesized as shown in Scheme 3.2. The protection step is adapted from Ishikawa *et al.*^[104], the azo coupling is adapted from the procedure proposed by Samanta *et al.*^[27].

Table 3.5: Optimization parameters for synthesis strategy B3.

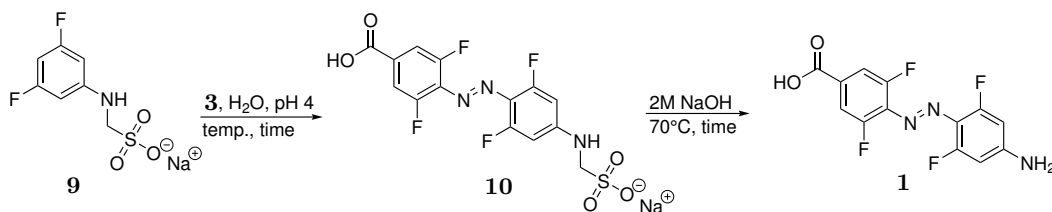
Entry	Time (h)	eq (9)	10 ^{*1}
1	72	0.5	n.d. ^{*2}
2	2	0.5	n.d.
3	2	1	n.d.
4	20	0.5	n.d.
5	20	1	n.d.

^{*1}determined by MS^{*2}not detected

A protection group that is already used in the synthesis of azobenzenes comprising an aniline is methane sulfonate (MeSO₃)^[105–107]. The synthesis plan was adapted to that shown in Scheme 3.8. Again, the first step is the protection of the aniline. The mechanism for the sulfomethylation was already investigated by Chen *et al.*^[108] and starts with the formation of sodium hydroxymethanesulfonate (HMS) from formaldehyde and NaHSO₃ (Figure 3.20)^[19]. In neutral pH, there is a high affinity between the negative charge of the sulfonate and the positively charged aniline, leading to a 4-membered cyclic intermediate. In a S_N2 reaction, a hydronium ion leaves the compound while the new bond is formed between the aniline and the MeSO₃. To perform this reaction, a

**Figure 3.20:** Mechanism of sulfomethylation as proposed by Chen *et al.*^[108].

procedure reported by Ishikawa *et al.*^[104] was used. The 3,5-difluoroaniline was reacted with NaHSO₃ and formaldehyde in water and refluxed at 100 °C for 40 minutes. It was possible to yield 93% of compound **9** after filtration and washing without further purification. With the protected aniline derivative in hand, the azo coupling was again tried in DMF following the procedure of the Woolley group, as they also used the MeSO₃ in their synthesis route (Table 3.5, entry 1)^[27]. The obtained solid was analysed by MS and did not show any formation of the product. Changing both reaction time and the ratio between donor and acceptor moiety did not improve the outcome (entries 2 to 5).



Scheme 3.9: Synthesis strategy B4: Protection of aniline as methyl sulfonate before azo coupling in H_2O . Compound **3** is synthesized as shown in Scheme 3.2, compound **9** is synthesized as shown in Scheme 3.8. The azo coupling is adapted from the procedure proposed by Prestel and Möller^[98].

Table 3.6: Optimization parameters for synthesis strategy B4.

Entry	Temp. (°C)	Time (h)	Deprotection	Conversion*	Purity*
1	RT	2	2.5 h	56%	medium
2	RT	1.5	18 h	60%	medium
3	RT	0.6	24 h	58%	low
4	RT	72	-	72%	low
5	50	18	2.5 h	83%	low
6	0	3	2.5 h	80%	medium

*determined by NMR

As an alternative, the procedure proposed by Prestel and Möller^[98] was envisaged and modified by using 2 M NaOH(aq) as deprotection (Scheme 3.9). The conditions that were previously optimized for the triazene synthesis were chosen for the first attempt (Table 3.6, entry 1). The deprotection was set up for 2.5 hours as described by Samanta *et al.*^[27]. To isolate HOOC-*o*-TFAB-NH₂ **1**, the crude reaction mixture was washed with CH_2Cl_2 at neutral pH to remove remaining unprotected 3,5-difluoroaniline. Subsequently, the product was extracted at pH 5 with EtOAc and analyzed by ¹H NMR (Figure 3.21, spectrum 1), which still showed 44% not reacted benzoic acid species.⁶ Furthermore, the analysis of the remaining aqueous phase revealed a high amount of still protected azobenzene **10**. Thus, two factors had to be optimized in this reaction: the conversion and the deprotection yields. For the latter, the deprotection time was increased to first 18 and then 24 hours (entries/spectra 2-3). While both did not increase the yield, it was noticeable that the amount of impurities in the ¹H NMR spectrum increased with time. After 24 hours, one of the product peaks almost completely disappeared, indicating the decomposition of the azobenzene - presumably by substitutions of the fluorines (spectrum 3). It was concluded that the product is not stable at high temperatures in strong basic conditions for longer periods of time and the initial deprotection time of 2.5 hours was kept, accepting the lower yield for this step. To optimize

⁶ To calculate the conversion of benzoic acid, the product doublet at 6.2 ppm (or in case of unprotected product in spectra 4 at 6.85 ppm) was chosen as a reference. The product doublet at 7.6 ppm has a bigger integral indicating the inclusion of impurities in this peak and making it not suitable as a reference peak. In spectra 3 and 5, the reference doublet is (partly) disappearing due to decomposition of the product. In these cases, the other doublet is taken as reference.

the conversion, the coupling time was increased to 72 hours, leading again to an increase in impurities without significantly improving the conversion (entry/spectrum 4). Similar results were obtained when heating the reaction to 50 °C and letting it run overnight (entry/spectrum 5). Even though the conversion was increased to 83%, the ^1H NMR spectrum showed more impurities and not identified side products. Again, the higher temperatures are assumed to lead to (partial) decomposition of the product. Thus, the reaction was tried at 0 °C for 3 hours (entry 6), but no full conversion was observed (80%). However, there were almost no side products visible in the ^1H NMR spectrum (Figure 3.21, spectrum 6). In a last attempt, the procedure reported in the patent published by Zhao *et al.*^[102] (section 3.2.1.1) was again considered. Diazonium **3** was generated following the patent's procedure and then coupled to the protected aniline species **9** as described in Table 3.6 - entry 6. As already discussed in section 3.2.1.1, the patent uses a different procedure to generate the diazonium species, which was assumed to increase the conversion of this step. Thus, the combination of the two procedures was believed to increase the total turnover. Indeed, the ^1H NMR spectrum showed less not reacted benzoic acid species (Figure 3.21, spectrum 7). However, the impurities were increased, which is why these conditions were discarded. The reaction parameters described in Table 3.6 - entry 6 showed the highest purity in the screening of conditions with a sufficient conversion. To generate a bigger batch, which can be subjected to different purification methods, a scale up in a 3.07 g scale was performed. It was possible to yield 3.53 g of crude product with a 0.09:1 ratio of benzoic acid/azobenzene **1**. This results into a higher conversion than that of the small scale (92% compared to 80%). Again, the ^1H

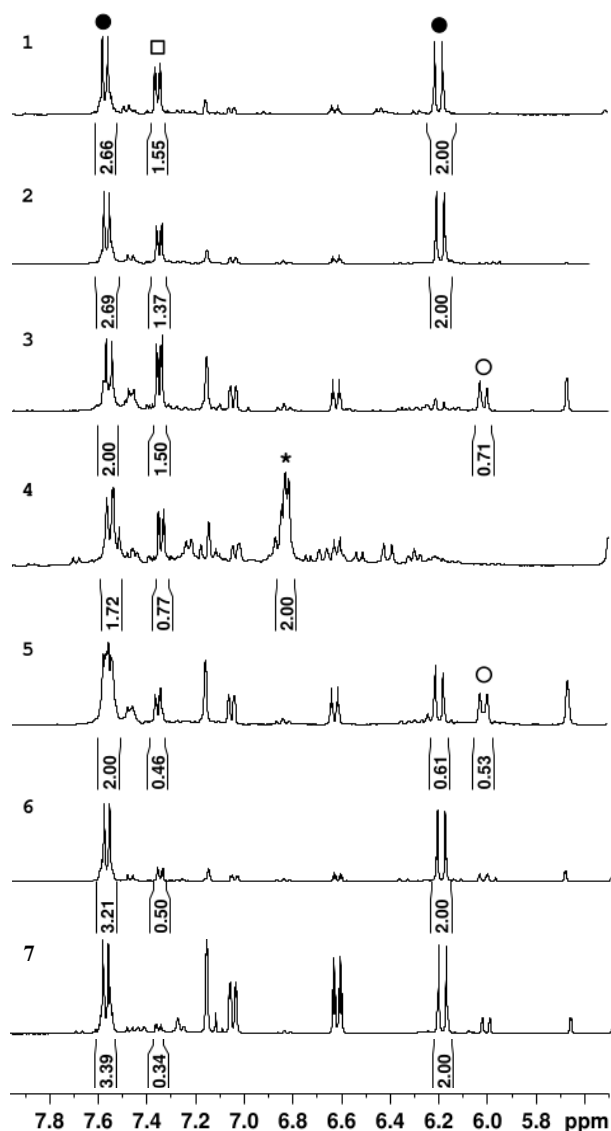


Figure 3.21: ^1H NMR spectra showing a mixture of azobenzene **1** (●) and the not reacted benzoic acid derivative (□) resulting from different reaction conditions. Decomposition of the product is indicated with ○. In spectrum 4, the product was not deprotected, giving a different chemical shift for the doublet signal (*). Numbering of the spectra refers to the entries of Table 3.6. Spectrum 7 results from the reaction including conditions from the patent published by Zhao *et al.*^[102].

NMR spectrum shows impurities in the aromatic region. Unfortunately, there was no sufficient separation in the TLC when trying different mixtures of MeOH, CH₂Cl₂, CHCl₃ and EtOAc (Figure 3.22). A first purification attempt by flash column chromatography starting with EtOAc and then slowly increasing the amount of MeOH up to 15% led to the separation of only the benzoic acid derivative. The other impurities remained in the product fractions. Similar results were obtained when precipitating the product from a EtOAc solution by addition of petroleum ether (Figure 3.23). As it was not possible to purify the product, the project was continued with strategy C.

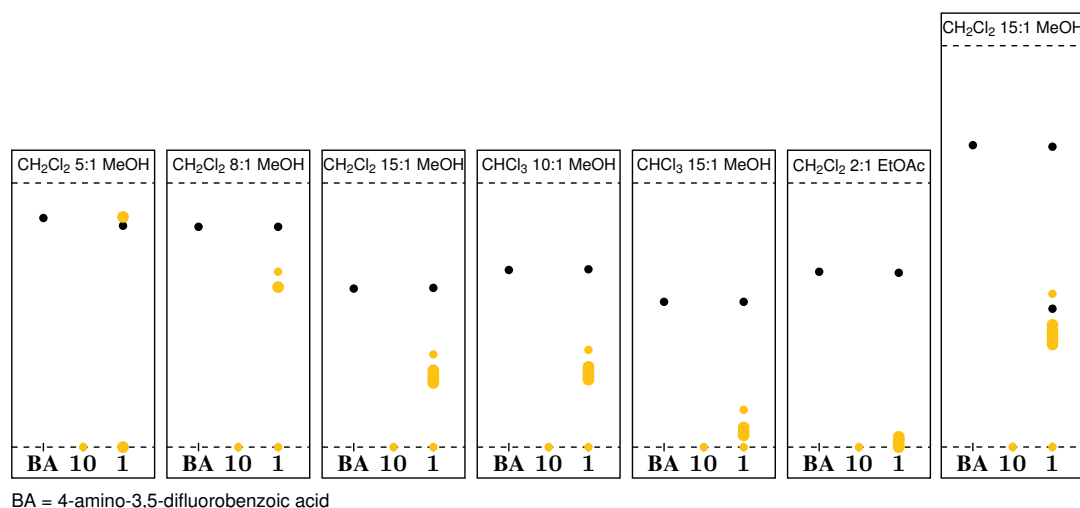
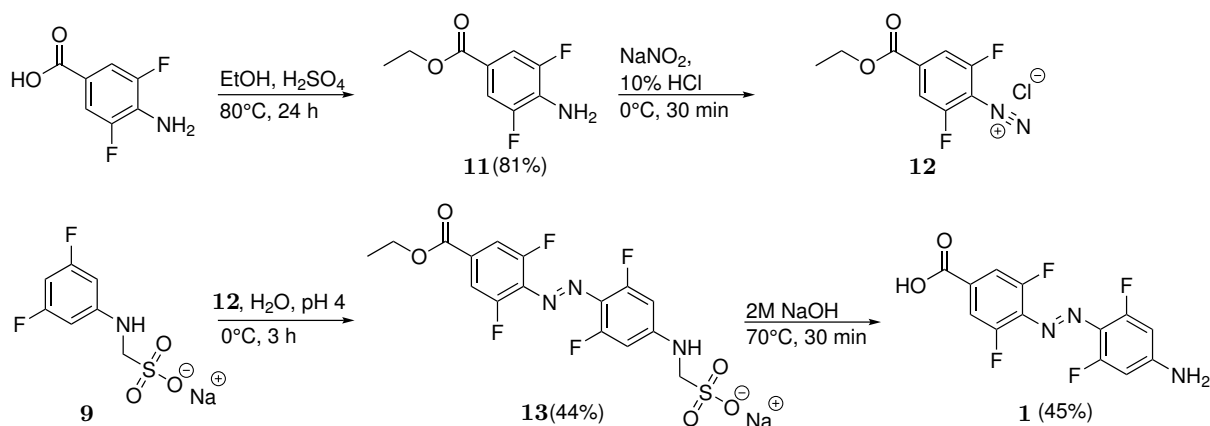


Figure 3.22: TLC of benzoic acid, protected azobenzene **10** and target compound **1** using different eluents. Black spots are visible under UV light 254 nm. Orange spots are already colored in visible light.

3.2.1.3 Strategy C: Esterification of the Benzoic Acid



Scheme 3.10: Synthesis strategy C: Esterification of benzoic ester and protection of aniline as methyl sulfonate before azo coupling in H₂O at 0 °C for 3 hours following a deprotection in 2 M NaOH at 70 °C for 30 min. Compound **9** is synthesized as shown in Scheme 3.8. The azo coupling is adapted from the procedure proposed by Prestel and Möller^[98].

The problems with the purification of HOOC-*o*-TFAB-NH₂ **1** are coming mainly from its zwitterionic character. This is why, the negatively charged benzoate was converted into an ester. The synthetic strategy therefore changed to that in Scheme 3.10. To generate the benzoic ester, a solution of 4-amino-3,5-difluorobenzoic acid in EtOH was acidified with concentrated H₂SO₄ and refluxed for 24 hours. The pure ester could be obtained with 81% yield. The azo coupling was performed following the optimized procedure as described in section 3.2.1.2, Scheme 3.9 - entry 6 but leaving out the final deprotection in NaOH. Both the ester and the methane sulfonate can be released in alkaline conditions. Before going to this step, double protected azobenzene **13** was purified by flash column chromatography on a Biotage[®] system. Using this instrument facilitated the separation as it detects the absorption spectra during the run. It was possible to generate 44% of pure protected azobenzene **13**, which was then subjected to deprotection in NaOH(aq) at 70 °C for 30 minutes. Even though there was no full conversion, higher temperatures or longer reaction times led to the decomposition of the compound. With the chosen conditions, it was possible to generate pure HOOC-*o*-TFAB-NH₂ **1** (Figure 3.23) in 45% yield, giving an overall yield of 16%.

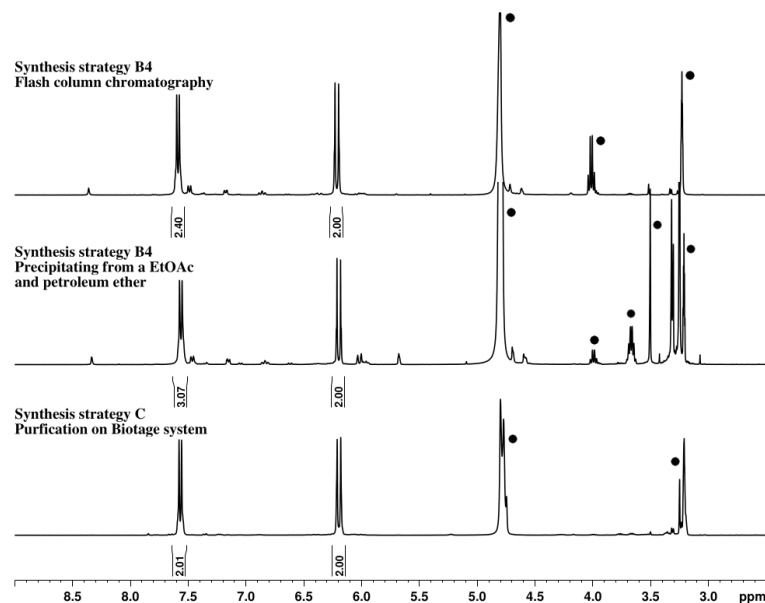


Figure 3.23: Comparison of ¹H NMR spectra from purification attempts after strategy B4 and pure product generated by strategy C. Solvent signals are indicated with ●.

3.2.2 Attempts to Couple HOOC-*o*-TFAB-NH₂ with Amino Acids

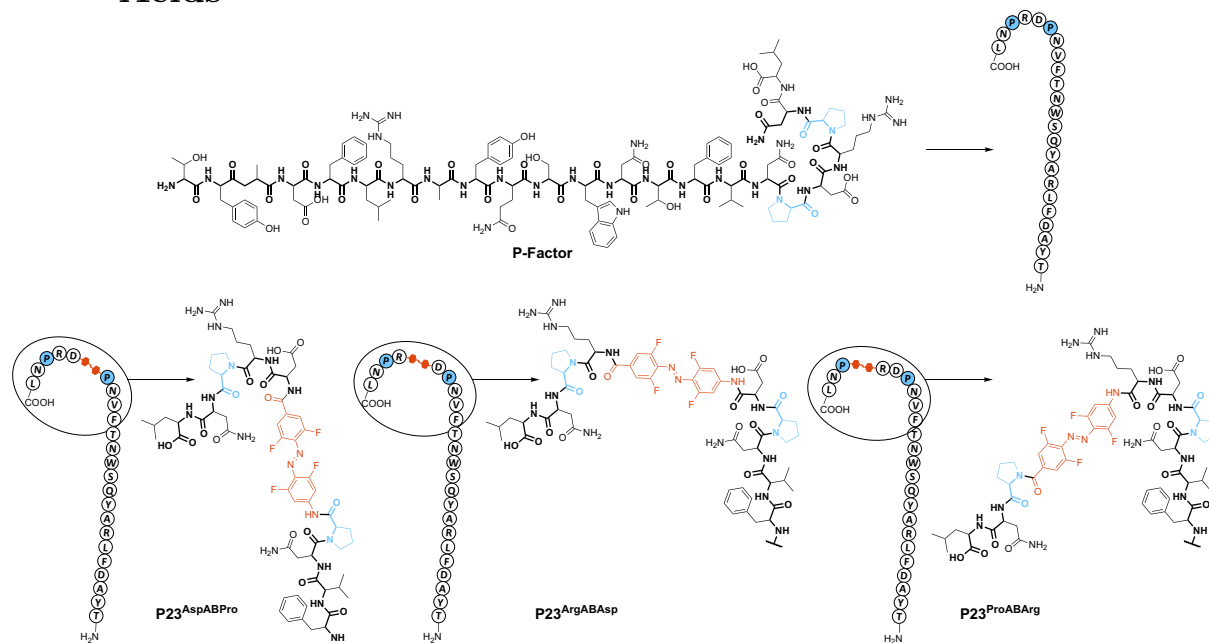


Figure 3.24: Possible positions to integrate compound **1** into the P-factor.

The synthesized HOOC-*o*-TFAB-NH₂ **1** is supposed to be integrated into the backbone of P-factor. First, it had to be decided at which exact position HOOC-*o*-TFAB-NH₂ **1** shall be integrated into the backbone. As described in section 2.1.2, the possibilities are limited, as some positions are important for either the pheromone's binding to or its interaction with the receptor. It is assumed that the formation of the loop between amino acid 15 and 23 (phenylalanine F to leucine L) is necessary for the successful activation of the receptor and, thus, the initiation of insulin production. For the generation of a reversible system, the receptor shall only be activated if the photoswitch is present in one specific isomer. Therefore, it was decided to integrate the compound into the loop region, assuming that the secondary structure is only retained in either the *cis* or the *trans* conformation. The three chosen positions are highlighted in Figure 3.24. As a first idea, treating azobenzene **1** as a normal amino acid and using it in the Fmoc synthesis strategy (described in section 2.2.1) seems like the logical first attempt. To achieve this, the amino group needs to be protected with Fmoc (Figure 3.25).

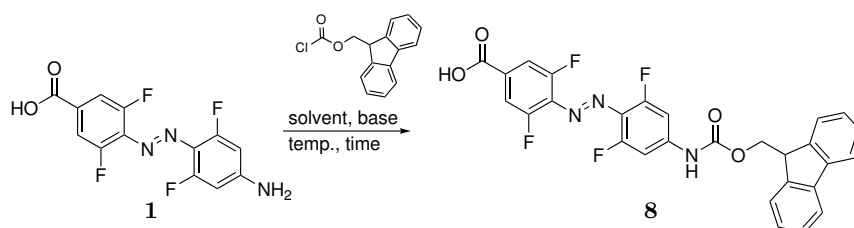


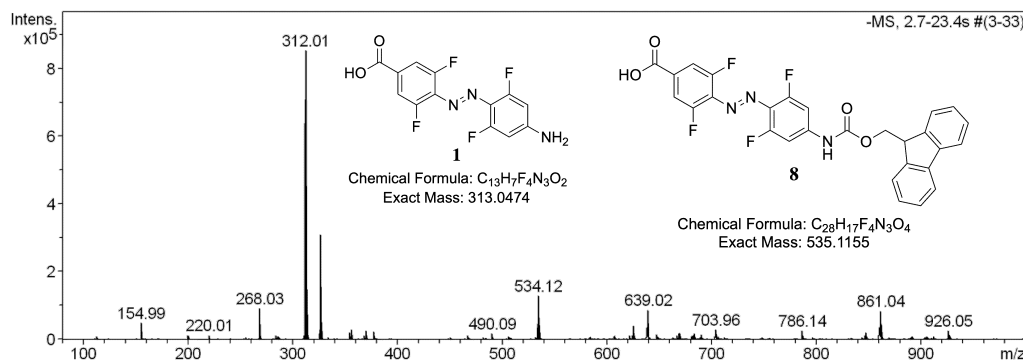
Figure 3.25: Attempts to protect azobenzene **1** with Fmoc.

Table 3.7: Optimization parameters to protect azobenzene **1** with Fmoc.

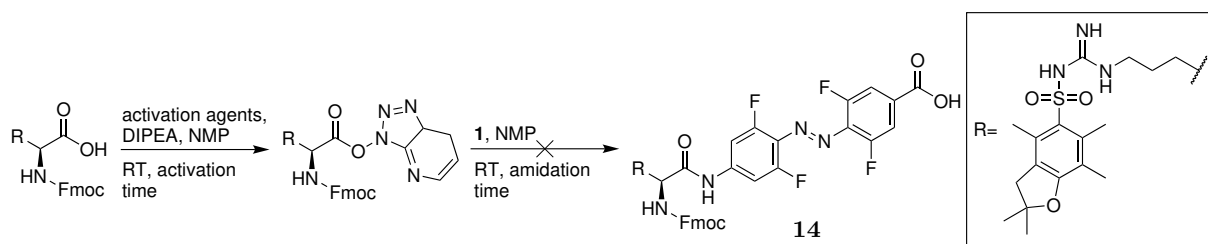
Entry	Solvent	Base	eq (Fmoc-Cl)	Temp. (°C)	Time (h)	8 ^{*1}
1	H ₂ O	-	1.2	60	2	n.d. ^{*2}
2	THF	NaHCO ₃ (aq)	1.05	0→RT	18	n.d.
3	THF	NaHCO ₃ (aq)	3.15	0→RT	0.5	traces
4	THF	NaHCO ₃ (aq)	3.15	0→RT	1.5	traces
5	THF	NaHCO ₃ (aq)	3.15	0→RT	18	traces
6	dry CH ₂ Cl ₂	pyridine	1.1	RT	18	traces
7	Dioxane	NaHCO ₃ (aq)	1.15	RT	40	n.d.

^{*1}determined by MS^{*2}not detected

Initially, the reaction was tried using the same conditions as for the Fmoc protection of 3,5-difluoroaniline (Table 3.7, entry 1, compare Scheme 3.6). Due to solubility problems, no reaction occurred. A procedure reported by Auernheimer *et al.*^[109] is using THF as solvent and adds NaHCO₃(aq) to generate mild basic conditions. This should buffer the released HCl during the synthesis^[103]. Following this procedure did not yield any product (entry 2). By using the same reaction conditions, but increasing the equivalents of Fmoc-Cl (entries 3 to 5), it was possible to see traces of azobenzene **8** (Figure 3.26). However, even by prolonging the reaction time, the yield of the protection did not improve. Assuming that the aqueous solution leads to a Fmoc cleavage, the solvent was changed to dry CH₂Cl₂ and the base to pyridine. This did again not increase the yield of the reaction (entry 6). In a final attempt, dioxane was used as solvent, which did not give any detectable amounts of the desired product (entry 7). As a consequence, this strategy was discarded.

**Figure 3.26:** MS analysis of azobenzene **1** protection with Fmoc-Cl to give compound **8**.

Prestel and Möller^[98] also integrated their synthesized azobenzene in the backbone of a peptide. Instead of introducing a protection group, the authors directly coupled the azobenzene's amino group to a Fmoc protected amino acid. As HOOC-*o*-TFAB-NH₂ **1** shall be inserted inside the loop of the peptide (Figure 3.24), its amino group needs to be coupled to the carboxyl groups of either proline, aspartate or arginine. Initially, the coupling was performed only with arginine (Scheme 3.11). In their procedure, Prestel and Möller^[98] first activated the amino acid's carboxyl group with HATU and 1-hydroxy-7-



Scheme 3.11: Attempts to couple azobenzene **1** with HOAt-activated arginine. Procedure adapted from Prestel and Möller^[98].

Table 3.8: Optimization parameters to couple azobenzene **1** with HOAt-activated arginine.

Entry	eq (1)	c (mM)	eq (DIPEA)	Activation time	Activation agents	Amidation time (h)	14 * ¹
1	1	83	2	5 min	HATU+HOAt	14	n.d.* ²
2	1	83	5	5 min	HATU+HOAt	20	n.d.
3	1	83	3	3 h	HATU+HOAt	18	n.d.
4	1	83	3	3 h	EDC+HOAt	18	n.d.
5	0.1	8.3	2	2 h	HATU+HOAt	72	n.d.

*¹determined by MS and TLC

*²not detected

azabenzotriazole (HOAt) in NMP under basic conditions. They suggested a coupling time of 14 hours after which they precipitated the product in ice-cooled water. The MS analysis revealed no product formation (Table 3.8, entry 1). The first attempt was to increase the equivalents of DIPEA (entry 2) as the activation is base-catalysed (Figure 2.10). Since this did not yield the product, the activation time was increased to 3 hours and the first activation agent was exchanged to 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) (entry 3 and 4). Both did not change the outcome. As in some of the attempts, the MS analysis indicated a coupling between two HOOC-*o*-TFAB-NH₂ **1** (Figure 3.27), the concentration of the azobenzene was reduced by factor 10 to diminish this side reaction. Again, no product formation was observed.

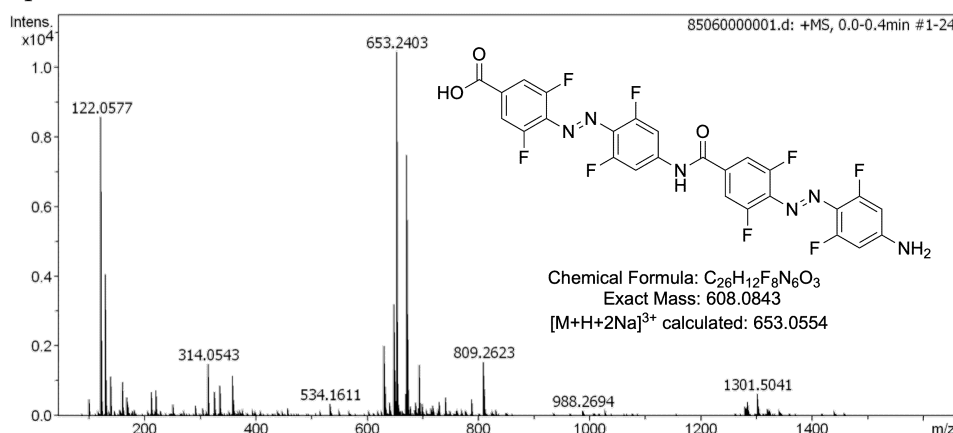


Figure 3.27: MS analysis of coupling reaction between azobenzene **1** and HOAt-activated Fmoc-Arg(Pbf)-OH to give compound **14**. The product was not observed. Instead a dimerization of **1** can be assumed.

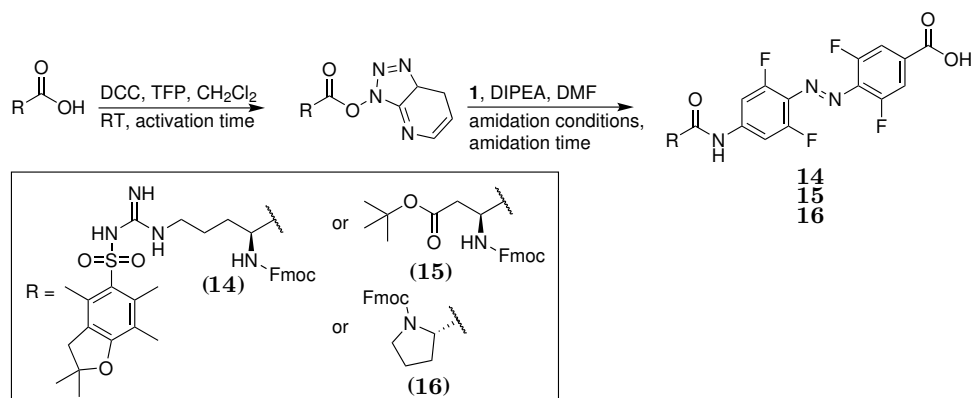


Figure 3.28: Attempts to couple azobenzene **1** with TFP-activated amino acids. Procedure adapted from Zhang *et al.*^[110].

Table 3.9: Optimization parameters to couple azobenzene **1** with TFP-activated amino acids.

Entry	amino acid	c (mM)	Activation time (h)	Amidation conditions	Amidation time (h)	14/15/16 ^{*1}
1	Arg	150	3	0 °C→RT	72	n.d. ^{*2}
2	all 3	150	3	0 °C→RT, MW	2	n.d.
3	Arg	100	3	0 °C→RT, MW	2	✓
4	Pro, Asp	100	2	0 °C→RT, MW	2	n.d.
5	Asp	100	2	0 °C→80 °C, MW	2	n.d.

^{*1}determined by MS and TLC

^{*2}not detected

An alternative procedure that was reported by Zhang *et al.*^[110] to couple the amino group of a lysine to methoxy polyethylene glycol (MPEG)-acetic acid was chosen. The group activates the carboxyl group with DCC and tetrafluorophenol (TFP) in CH₂Cl₂ for 2 hours. Addition of acetone leads to precipitation of the formed side product dicyclohexylurea (DCU), which can be filtered off. The activated carboxylic acid is then added to an ice-cooled solution of the lysine in DMF and stirred for 19 hours at RT. The procedure was adapted to the required starting materials, giving the synthesis route shown in Figure 3.28. A first attempt did not yield any detectable product (Table 3.9, entry 1). Since it is known that the amidation reaction performed in a microwave (MW) reactor increases the conversion^[111], this idea was also attempted (entry 2). To do this and to exclude that the arginine was the limiting factor, the reaction was repeated with all three amino acids for 2 hours in a MW reactor (entry 2). None of these reactions gave any product. For the next attempt, the amount of solvent was increased to a final azobenzene concentration of 0.10 M (entry 3). The MS analysis of the product indicated the formation of **14** (Figure 3.29). However, it was not possible to create a sufficient separation in the TLC as already the Fmoc-Arg(Pbf)-OH and HOOC-*o*-TFAB-NH₂ **1** have almost the same R_f. Thus, the product could not be isolated. This is why the reaction was repeated with Fmoc-Pro-OH and Fmoc-Asp(OtBu)-OH but did not yield into any product formation (entry 4). As

a last attempt, the MW reactor was set to 80 °C to increase the conversion but again without success (entry 5).

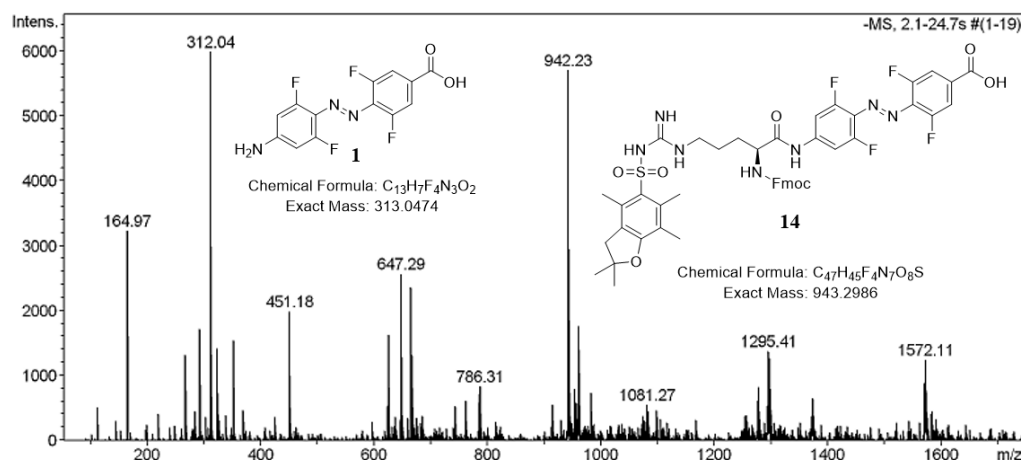


Figure 3.29: MS analysis of coupling reaction between azobenzene **1** and Fmoc-Arg(Pbf)-OH to give compound **14**.

3.2.3 Synthesis of *o*-TFAB



Scheme 3.12: Synthesis of *o*-TFAB **17** following a procedure published by Bleger *et al.*^[29].

To analyse the influence of the *para* substituents of HOOC-*o*-TFAB-NH₂ **1** on the photophysical characteristics, *o*-TFAB **17** was synthesized. This compound only bears the four fluorines in *ortho* position and no *para* substituents. Thus, it is a symmetric azobenzene and can be synthesized using an oxidative mechanism starting from 2,6-difluoroaniline. As oxidizing agent KMnO₄ supported in FeSO₄ was chosen. Following the procedure by Bleger *et al.*^[29] the starting material 3,5-difluoroaniline was refluxed with the oxidant in CH₂Cl₂ for 16 hours (Scheme 3.12). After filtering off the remaining and consumed oxidizing agents, the product was purified by flash column chromatography to give orange crystals. The oxidation of aniline has three states going over hydroxyamine **18** and nitrosobenzene **19** to nitrobenzene **20** (Figure 3.30). In general, all these states can be used to form an azobenzene depending on the conditions applied. Hydroxyamine **18** and nitrosobenzene **19** can form azoxybenzene **21** that, in the presence of an acid, can be transformed to the corresponding azobenzene undergoing a Wallach reaction (section 3.1.3.3). Another option is to transfer nitrobenzene **20** to a reductive mechanism. However, the given conditions are neither acidic nor reductive. Thus, the most probable reaction happening is the Mills reaction in which nitrosobenzene **19** reacts with a non-oxidized aniline and forms the azobenzene in a condensation reaction (section 3.1.3.2). The procedure asks to use a 3-fold excess of the oxidation agent compared to the aniline. Looking at the mechanism, only

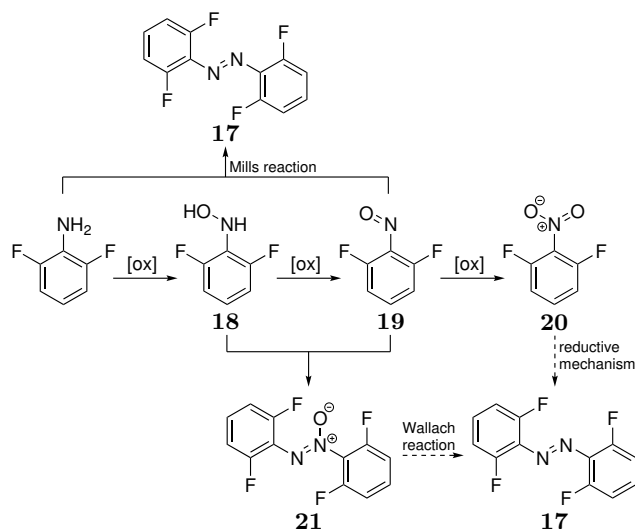


Figure 3.30: Oxidation states of 2,6-difluoroaniline and possible routes to generate **17**.

half of the starting material shall be oxidized until the second stage. This means that an equal amount of oxidant would be sufficient. By using a higher ratio, it is more likely to have more and also higher oxidized species. Even though the Wallach reaction needs acidic conditions, azoxybenzene **21** can still form, explaining the low yield of only 4%. However, no effort was taken to optimize this procedure as the generated 18 mg were enough for the photophysical studies.

3.2.4 Photophysical Studies

The following sections focus on the photophysical characteristics of the synthesized triazenes and azobenzenes. By also studying the triazenes, the UV can be used in the future as an additional indication of purity. To compare the absorption spectra, the molar absorption coefficient ϵ_λ was calculated using the Lambert-Beer's law^[112]:

$$A = -\log(T) = \log \frac{I_0}{I} = \epsilon_\lambda * c * d \quad (3.1)$$

with: A = absorbance
 T = transmission
 I_0, I = intensity before/after passing the sample
 ϵ_λ = molar absorption coefficient ($\text{M}^{-1}\text{cm}^{-1}$)
 c = concentration (M)
 d = optical path (cm)

3.2.4.1 Study of the Triazenes

Similar to an azobenzene, the triazene can be found as either *cis* or *trans* form with the latter being the more thermodynamically stable one. It is reported that the *cis* isomer shows a lower absorption and can be formed by irradiation at the peak wavelength^[113]. To investigate the two synthesized triazenes **5** and **4**, 0.1 mM solutions of both compounds in MeOH were prepared and the UV-Vis spectra acquired. The two molecules give very similar spectra with absorption peaks at 235 nm ($\epsilon_\lambda = 9.75 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$) and 350 nm ($\epsilon_\lambda = 10.37 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$), or 232 nm ($\epsilon_\lambda = 9.32 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$) and 348 nm ($\epsilon_\lambda = 13.42 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$), respectively (Figure 3.31). Subsequently, both cuvettes were irradiated at 348 nm for 45 minutes to induce isomerization to the *cis* isomer. It was expected to see a difference in the absorption spectra, i.e. a decrease in the peak area. However, the spectra acquired directly after the irradiation completely overlapped with those taken before (Figure 3.31).

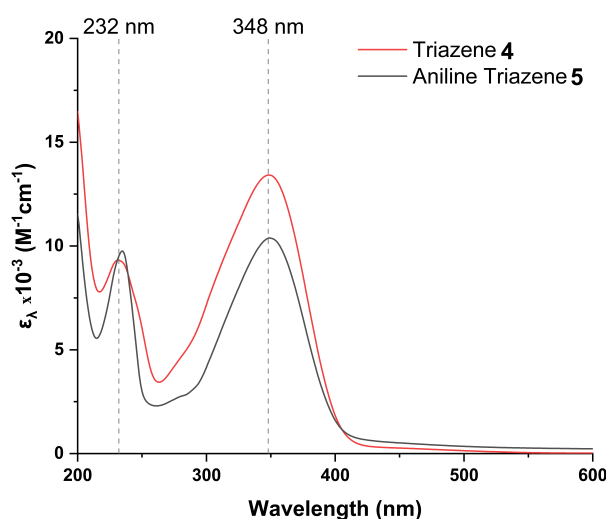


Figure 3.31: UV-Vis spectra of **4** and **5**.

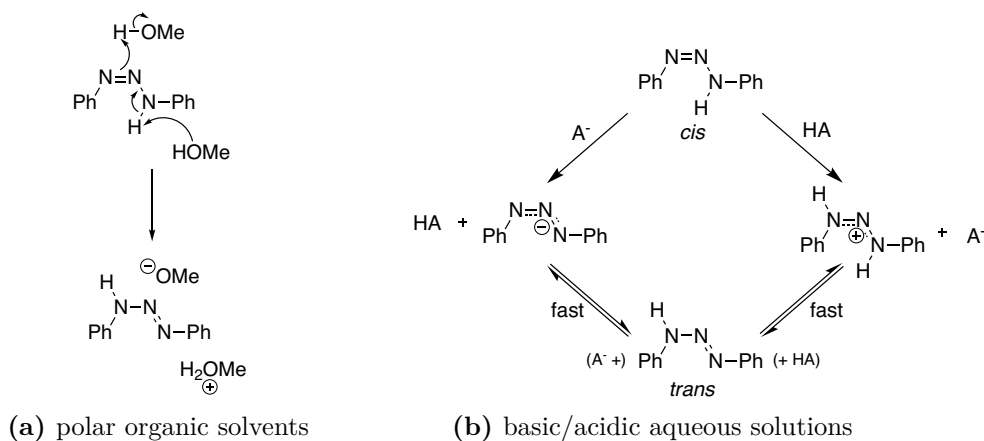


Figure 3.32: Thermodynamic isomerization of triazenes from *cis* to *trans*. Figure adapted from Barra and Chen^[114].

The conversion mechanism between the two isomers is solvent-dependent. In polar organic solvents (such as MeOH), a solvent-assisted 1,3-hydrogen shift is proposed^[114]. In this case, the protic solvent molecule can act as both a proton donor and acceptor and catalyses the isomerization as shown in Figure 3.32a. A similar mechanism takes place in aqueous solution. Depending on the pH, the process can be described as either acid-catalysed or base-catalysed and resembles that of the enol-ketonization (Figure 3.32b)^[114]. In both mechanisms the first step (ionization) is the rate limiting step. These transitions are

reported as fast processes, which happen in the time range of millisecond to seconds depending on the polarity of the organic solvent or the pH of the aqueous solution^[113,114]. During the experiment, the cuvette needed to be manually transferred from the irradiation lamp to the spectrophotometer (3 meters). Furthermore, the spectra acquisition time is 60 seconds. In total, this time exceeds the time needed to thermally relax the *cis* isomer back to *trans*, explaining why no difference in the absorption spectra was observed.

3.2.4.2 Study of the Azobenzenes

As already discussed in the beginning of the chapter (section 3.1), azobenzenes have two isomeric structures, *cis* and *trans*, from which the latter is the thermodynamically more stable one. Depending on the substitutions in the aromatic rings, the absorption bands for the π - π^* and the n - π^* transitions can be either distinct from each other or partially/fully overlapped. Furthermore, the substituents are reported to have an influence on the thermal half-life of the *cis* isomer^[88]. The scaffold of the target compound was designed in order to achieve both a separation of the n - π^* bands in *cis* and *trans* isomers and an increase in the thermal half-life. In a first experiment, the effects of the fluorines in *ortho* position were examined. In order to do so, a 0.5 mM solution of thermally relaxed azobenzene **17-*trans*** in MeOH was prepared and a UV-Vis spectra was acquired in a spectrophotometer using an optical path of 0.1 cm. Two distinct bands around 303 nm ($\epsilon_\lambda = 18.78 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$) and 456 nm ($\epsilon_\lambda = 0.94 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$) could be assigned to the π - π^* and n - π^* transitions, respectively (Figure 3.33a). A comparison with the spectra of non-substituted azobenzene shows an unusual red-shift in the n - π^* band. As already discussed in section 3.1, Dong *et al.*^[30] reported a similar phenomenon with an *o*-tetramethoxyazobenzene, which they explained by the *trans* isomer's non-planar conformation coming from the methoxy's steric and electronic effects. The less intense shift in the *o*-tetrafluoro compounds was justified by Bleger *et al.*^[29] by the less distorted planarity in the *trans* isomer. The right edge of the n - π^* absorption band (505 nm) was chosen to induce isomerization of the molecule into its *cis* form. The cuvette was irradiated for 20 minutes (Figure 3.33a). The π - π^* transition band showed a slight blue shift to 280 nm with a lower absorption ($\epsilon_\lambda = 7.56 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$). The n - π^* band showed a blue-shift of 44 nm after irradiation and appeared at 412 nm with an absorption coefficient ϵ_λ of $1.68 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$. This outcome is of great interest as a non-substituted azobenzene does not show a significant shift in the n - π^* band upon isomerization. Both conformations have their maximum at 440 nm. The two distinct n - π^* bands in the *ortho* substituted compound opens the opportunity to use this transition band to induce selective isomerization between the *trans* and *cis* isomers. As the bands lay in the visible light area, they are more favourable for the *in-vivo* use.

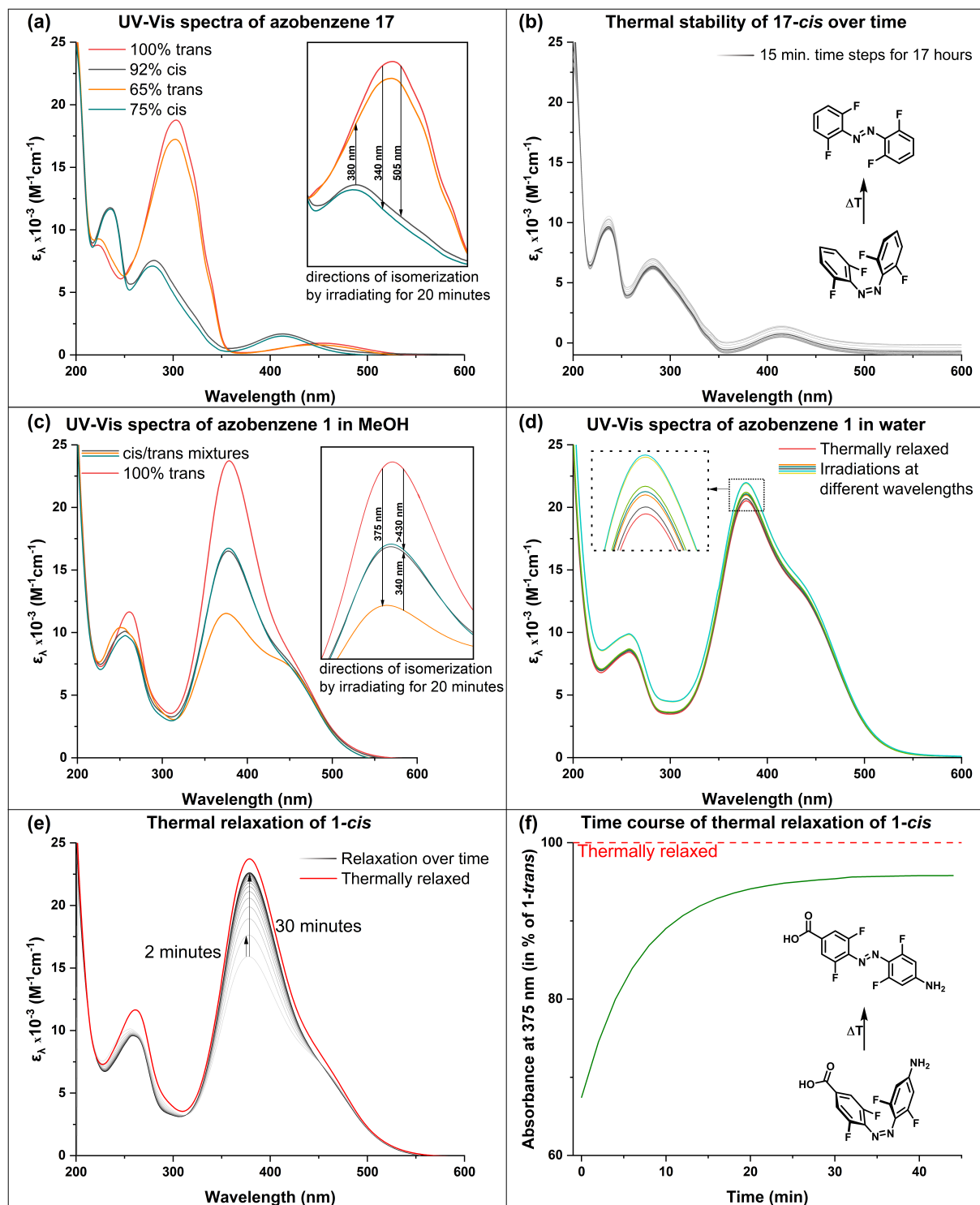


Figure 3.33: UV-Vis spectra of **17** and **1** in different isomerization states: (a) Isomerization of **17** with three wavelength (505 nm, 380 nm, 340 nm) for 20 minutes after thermal relaxation in dimethylsulfoxide (DMSO). (b) Thermal stability of **17-cis** in the dark at 25 °C. Isomerization of **1** in MeOH (c) and H₂O (d) with different wavelengths for 20 minutes after thermal relaxation in DMSO. It was not possible to calculate the PSS as the ¹H NMR spectra of both isomers are overlapping. e, f Thermal relaxation of **1-cis** at RT in the dark.

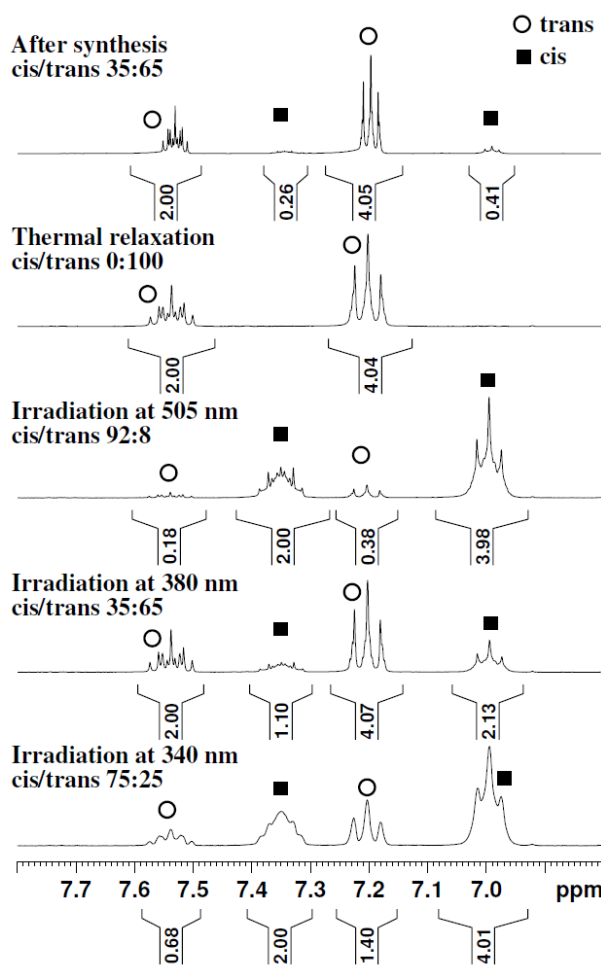


Figure 3.34: ^1H NMR spectra of **17** in different isomerization states. The spectra were acquired directly after the synthesis, after the thermal relaxation in DMSO and after irradiation at different wavelengths.

Additionally to the UV-Vis spectra, a ^1H NMR spectrum was taken. As the two isomers have different chemical shifts, the PSS can be calculated. For the irradiation at 505 nm a PSS of 92% *cis* was determined (Figure 3.34). Subsequently, the sample was irradiated at the left edge of the new $n\text{-}\pi^*$ absorption band (380 nm) to convert the azobenzene back into its *trans* isomer. Only a PSS of 65% *trans* could be obtained. Finally, the irradiation lamp was set to 340 nm (right edge of the *trans* isomer's $\pi\text{-}\pi^*$ band) to induce again the isomerization to *cis*. The calculation of the PSS gave 75% *cis*, meaning that the irradiation at the $n\text{-}\pi^*$ band gives a higher isomerization yield. Comparing the total spectra of the *trans* and *cis* isomers, this outcome could have been foreseen. While the *cis* shows no absorption at 505 nm, 340 nm still lays in its $\pi\text{-}\pi^*$ band. To analyse the stability of the *cis* isomer in polar solvents, the sample was again irradiated at 505 nm to reach a high PSS. The cuvette was then placed into a temperature controlled spectrophotometer and kept at 25 °C for 17 hours in the dark. Spectra were acquired every 15 minutes and overlaid (Figure 3.33b). The isomer is stable in these conditions in the monitored time period. These results are consistent with literature, where the compound was assigned to a half-life of 700 days in similar conditions^[29]. The researchers explain this high thermal stability by the higher activation barrier that needs to be overcome in order to form **17-trans** from **17-cis**. Unfortunately, it was not possible to perform the described experiment in more physiological conditions. The yeast cultivation and the human physiological temperature are 32 °C or 37 °C, respectively. Thus, it would have been more favourable to dissolve the azobenzene in H_2O and to monitor the *cis* stability at these temperatures. However, the azobenzene has a low solubility in H_2O . It was tried to do the experiment in MeOH at 37 °C but the solvent evaporated overnight, falsifying the results. It was proceeded by analysing the effect of the *para* substituents on the UV-Vis spectra and the isomerization. For this, a 0.5 mM solution of thermally relaxed azobenzene **1-trans** in MeOH and a 0.05 mM solution in H_2O were prepared. Spectra of both samples were

acquired in quartz cuvettes. The MeOH solution was again measured with an optical path of 0.1 cm and the H₂O solution in a 1.0 cm cuvette. The path lengths were chosen due to the different solubilities of the compound in the two solvents and the resulting different concentrations of the solutions. In contrast to the spectra of **17**, that of the *para* substituted azobenzene **1** shows overlapping of the π - π^* and n- π^* absorption bands in both MeOH and H₂O (Figure 3.33c, d). Both spectra show a maximum at 379 nm with molar absorption coefficients ϵ_λ of $23.73 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$ and $20.51 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$, respectively. The peaks were assigned to the π - π^* absorption bands. The short irregularity on the right side of the peaks indicate the n- π^* bands at approx. 452 nm and 436 nm, respectively. The two samples were subsequently irradiated at different wavelengths (510, 490, 450, 430, 375 and 340 nm) for 20 minutes each. While the sample dissolved in H₂O did not show any significant difference in the absorption spectra, irradiation at 375 nm led to a decreased π - π^* band of the MeOH probe with 49% of its original absorbance ($\epsilon_\lambda = 11.54 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$). Also, the maximum slightly shifted to 375 nm. All the other wavelengths led to an molar absorption coefficient of $16.51 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$ (70%). With the decreasing intensity of the π - π^* band, the n- π^* band became more visible. Comparing the spectra before and after irradiation, it shows that both *trans* and *cis* isomer have almost the same π - π^* and n- π^* absorption bands. Again, ¹H NMR spectra (in both MeOD and D₂O) were taken directly after irradiation, but no difference was observed (Figure 3.35). This could have two reasons: Either the two isomers have exactly the same chemical shift or the *cis* thermally relaxes back to the *trans* in a short time. To investigate this further, the MeOH sample was irradiated again at 375 nm and placed in the spectrophotometer. A series of spectra was acquired every 2 minutes (Figure 3.33e, f). The results showed that within 20 minutes the intensity returns to 95% of that of the *trans*. As the ¹H NMR spectrum was taken within 10 minutes after the irradiation, it should have shown both isomers. Since this was not the case, it can be concluded that both isomers have the same chemical shift. It was therefore not possible to calculate the PSS.

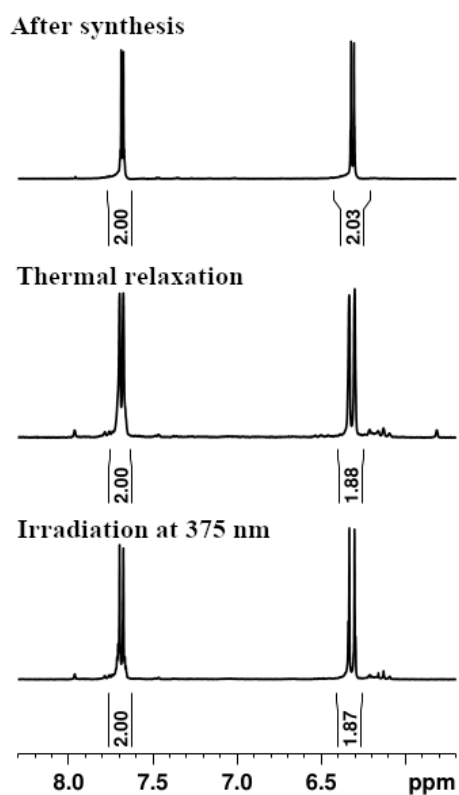


Figure 3.35: ¹H NMR spectra of **1**. The spectra were acquired directly after the synthesis, after the thermal relaxation in DMSO and after irradiation at 375 nm. No differences can be observed.

The reason behind the short half-life of the *cis* isomer lays in the two *para* substituents, leading to the formation of a push-pull system. Azobenzenes of this kind are reported with a red-shift of the π - π^* absorption band and really short half-lives of the *cis* form^[88]. In the case of a pseudo-stilbene (a push-pull azobenzene without *ortho* substitution), the half-life shortens to milliseconds and when analysing an azopyrimidine, researchers were able to define a half-life of only 40 ns^[115]. As already described in section 3.1.2, a low thermal stability can also be desirable depending on the application. However, for a controlled reversible system, it is necessary to selectively form both isomers in a stable and high PSS by irradiation at specific wavelengths. The low thermal stability combined with the overlapping absorption bands are not favourable for this application. Nevertheless, the results of the present study cannot directly be translated to the azobenzene as it would be present in the P-factor. By integrating the molecule into the backbone of the peptide, the two *para* substituents will transform into amides (Figure 3.36). Even though the amide group is still electron withdrawing or donating, depending on its rotation, this effect is weaker than that of carboxyls and amines. Therefore, the push-pull system will be weaker and should have less influence on both the π - π^* absorption band and the *cis* half-life.

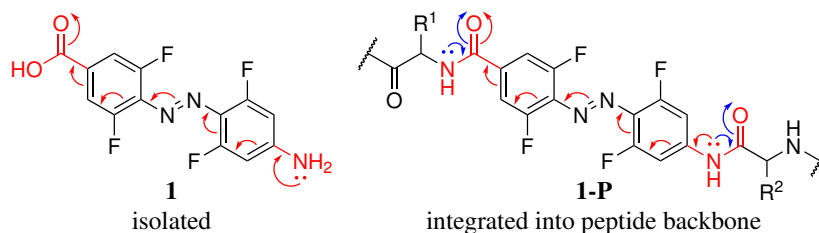


Figure 3.36: Push-pull system of isolated azobenzene **1** compared to that in the peptide. Red arrows show the direction of the push-pull system. Blue arrows indicate the counteracting forces induced by the amide bonds.

As it is envisaged to use the designed photoswitch for the repeated activation and deactivation of a biofactory, it is of high interest to be able to induce the isomerization infinitely many times without the photoswitch experiencing any decrease in isomerization yield. To show that the two azobenzenes **1** and **17** are fulfilling this requirement, new samples of thermally relaxed compounds were prepared as above and subjected to 8 cycles of repeated irradiations for 30 minutes each. Compound **17** was irradiated alternately at 505 and 380 nm (Figure 3.37a). For compound **1**, two samples were prepared. While one was irradiated alternately at 375 and 340 nm, the other was only irradiated at 375 nm and then allowed to relax thermally in the dark (Figure 3.37b). Both samples did not show any major changes in the isomerization yield during the investigated period.

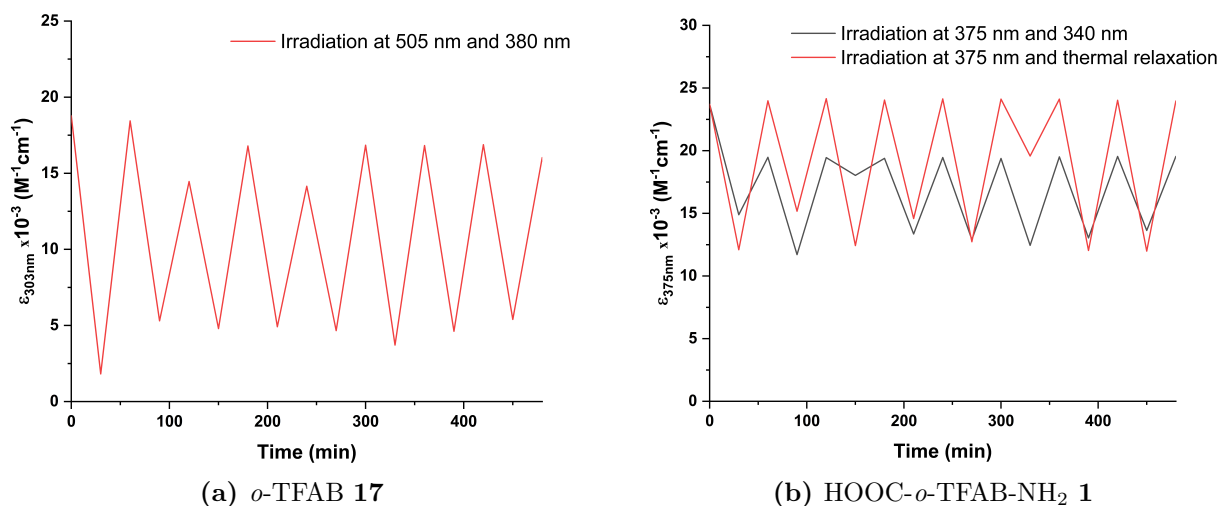


Figure 3.37: Fatigue graph of isomerization of azobenzenes **17** and **1**.

3.3 Conclusions and Perspective

During this Master's thesis, it was possible to generate target compound HOOC-*o*-TFAB-NH₂ **1** in an overall yield of 16%. Due to the low electron density in the aromatic system and the presence of an aniline, the standard azo coupling mechanism led to the formation of a triazene as main product. The introduction of an electron donating protection group allowed the directed synthesis of the targeted azobenzene. Photophysical studies revealed that an *ortho* substitution with fluorines leads to a sufficient separation of the $n\text{-}\pi^*$ bands of the two isomers and also prolongs the thermal stability of the *cis* isomer. However, target compound HOOC-*o*-TFAB-NH₂ **1** bears a push-pull system due to an EWG in one and an EDG in the opposite *para* position of the diazo bond. As a consequence, the thermal stability is shortened to minutes and the spectra of the two isomers show nearly identical $\pi\text{-}\pi^*$ and $n\text{-}\pi^*$ transition bands. Three strategies were attempted towards the integration of azobenzene **1** into the peptide. It was neither possible to protect the compound's amino group with Fmoc nor to isolate the photoswitch coupled to an amino acid.

It was hypothesized that the integration of a modified azobenzene derivative into the backbone of a peptide would give the opportunity to reversibly and selectively change the peptide's conformation and thus, control the interaction with its cellular receptor. Hence, it was aimed to create an azobenzene with both isomerization wavelengths in the visible light region and two thermally stable isomers. Consistent with literature, it was possible to confirm the suitability of a TFAB scaffold for this application. As the amidation of final compound **1** was not achieved, the photophysical evaluation of the photoswitch in the peptide backbone was not conducted during the course of this master thesis.

The screening of reaction conditions revealed that the azo coupling has to be performed at 0 °C. So far, the reaction was quenched after 2.5 hours, as after this time, no progression

was observed. This resulted in a yield of only 44%. It is envisaged to optimize the azo coupling by extending the reaction time. To maintain the required low temperatures, usage of a cryostat should be considered. During the final deprotection, decomposing of the compound was observed by ^1H NMR spectroscopy. It was assumed that this was caused by hydroxylation of the aryl fluorides in the strong basic aqueous conditions. Thus, future work will focus on alternative strategies for this final step, including the extension of reaction time in lower temperatures to diminish the hydroxylation, and the exploration of non-aqueous conditions. The synthesized azobenzene **1** has a reduced reactivity in its amino group due to the electron withdrawing effect of both the fluorides and the carboxyl group in the opposite *para* position. This complicates the amidation of this functional group. It will therefore be of interest to investigate reaction conditions that are able to increase the nucleophilicity of the mildly active amine.

The push-pull system in the final compound was counteracting the advantages of the TFAB scaffold concerning the photophysical properties. However, it is assumed that this effect will be weakened after incorporation of the compound into the peptide backbone due to the amidation of both *para* positions. A suggested way to characterize a similar system is to prepare an azobenzene with an acylated amino and an amidated carboxyl terminus. In this way, the delocalization of the electrons can be reduced, which will hopefully diminish the push-pull effect.

Chapter 4: Experimental

4.1 Instrumentation

Thin layer chromatography (TLC) was conducted on pre-coated aluminum sheets with 0.20 mm *Machevery-Nagel* Alugram SIL G/UV254 with fluorescent indicator UV254 or UV366. **Column chromatography** was carried out using *Merck Gerduran* silica gel 60 (particle size 63-200 μm). In indicated cases, the separation is done on a *Biotage*[®] Isolera[™] Spektra LS flash system. If not indicated, the separation is done by manual column chromatography. **Solid phase peptide syntheses** were performed on a semi-automatic *FOCUS XC* peptide synthesizer, coming with a computer/control system from *aapptec*. **High performance liquid chromatography** (HPLC) analyses and purifications were performed on the following HPLC system: Analytical HPLC: *Agilent 1260 Infinity* liquid chromatograph system. The column used was a *Varian Pursuit C18*, 5 μm , 250x4.6 mm column. Preparative HPLC: *Agilent 1260 Infinity* liquid chromatograph system. The column used was a *Varian Pursuit C18*, 5 μm , 250x21.2 mm column. 0.1% TFA in H_2O and 0.1% TFA in CH_3CN were used as eluents in all cases. **Lyophilisation** was performed on a *Christ* Freeze Dryer ALPHA 2-4 LDplus, connected to a *Vaccumbrand* Chemistry-HYBRID-pump. The ice condenser was approximately at -85°C and the vacuum around 2×10^{-3} mbar. **Nuclear magnetic resonance** (NMR) ^1H , ^{13}C , ^{19}F , HSQC, HMBC, and NOESY spectra were obtained on a 700 MHz NMR (*Bruker BioSpin* AV III HD 700), 600 MHz NMR (*Bruker BioSpin* AV III 600) or 400 MHz NMR (*Bruker BioSpin* AV NEO 400). Chemical shifts were reported in ppm according to tetramethylsilane using the solvent residual signal as an internal reference (DMSO- d_6 : ^1H = 2.50 ppm, ^{13}C = 39.52 ppm, MeOD: ^1H = 3.31 ppm, ^{13}C = 49.15 ppm). Coupling constants (J) were given in Hz and were averaged. Resonance multiplicity was described as s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), br (broad signal), dd (doublet of doublets), td (triplet of doublets) or tt (triplet of triplets). Carbon spectra were acquired with a complete decoupling for the proton. Signals observed in ^1H spectra at 1.27 ppm and 0.8 ppm are traces of grease. **Melting points** (mp) were measured on a *Gallenkamp* apparatus in open capillary tubes and have not been corrected. **Infrared spectra** (IR) were recorded on a *Shimadzu* IR Affinity 1S FTIR spectrometer in ATR mode with a diamond mono-crystal. Selected absorption bands are reported in wavenumber (cm^{-1}). Transmittance is reported as weak (w, $>85\%$), medium (m, $>70\%$) or strong (s, $<70\%$). **Ultraviolet-Visible absorption spectra** (UV-Vis) were recorded in quartz cuvettes using an *Agilent-Cary* 5000 UV-Vis-NIR spectrophotometer. Temperature-controlled measurements were done using a *JASCO* V-750 UV-Vis spectrophotometer. All spectra were acquired against the respective solvent of the sample as background. **Irradiation** of the compounds was performed using a *LOT* MSH-300 monochromator system. **Mass spectrometry analyses**

of the purified peptide was carried out by electron spray ionization (ESI)-high resolution mass spectrometry (HRMS). ESI-HRMS spectra (m/z 50-1900) were obtained on a maXis UHR ESI-Qq-TOF mass spectrometer (*Bruker Daltonics*, Bremen, Germany) in positive and/or negative ion mode by direct infusion. The sum formulas of the detected ions were determined using *Bruker Compass DataAnalysis* 4.1 based on the mass accuracy ($\Delta m/z \leq 5$ ppm) and isotopic pattern matching (SmartFormula algorithm). **Centrifugations** were done in a refrigerated centrifuge (*Eppendorf* 5840 R) at 4 °C and 4000 rpm. **Yeast incubation** was performed in an *Eppendorf* Innova 42 incubator at 32 °C with 180 rpm. **Cell viability assays** were performed with fluorescence detection in an *LUNA-II YF* automated yeast cell counter. **Fluorescence Intensity Measurements** were carried out in a *VICTOR Nivo* Multimode Microplate Reader equipped with excitation filter 480/30 nm and emission filter 530/30 nm. **Circular Dichroism** (CD) spectra were recorded on a *AppliedPhotophysics* Chirascan-plus CD spectrometer using a bandwidth of 1.0 nm. The spectra were acquired between 280 and 190 nm with 1.0 nm steps.

4.2 Materials and General Methods

Chemicals were purchased from *Sigma Aldrich*, *Acros Organics*, *Fluorochem*, *TCI*, *aapptec*, *carbosynth*, and *ABCR*, and were used as received. Resins for solid phase synthesis were purchased from *Sigma Aldrich* and *aapptec*. Solvents were purchased from *Sigma Aldrich* and *Acros Organics*. Deuterated solvents were purchased from *Eurisotop*. Anhydrous DMF was purchased from *Acros Organics*. Low temperature baths (0 °C) were prepared using a mixture of ice in H₂O. The inert atmosphere was maintained using argon-filled balloons equipped with a syringe and needle that was used to penetrate the silicon stoppers used to close the flasks' necks. The addition of liquid reagents was done by means of dried plastic syringes. After extraction and washings, the organic phase containing the target compound was dried using MgSO₄ as drying agent.

For **peptide synthesis and HPLC purification**, all materials were obtained from commercial suppliers: HPLC grade acetonitrile (CH₃CN), water, piperidine, trifluoroacetic acid (TFA), triisopropylsilyl silane (TIPS), N-methylpyrrolidone (NMP), *N,N*-dimethylformamide (DMF), diisopropylethylamine (DIPEA), fluorenylmethoxycarbonyl (Fmoc) amino acids, and hexafluorophosphate azabenzotriazole tetramethyl uronium (HATU) from *Sigma Aldrich*, *Fisher Scientific*, *Carbosynth*, *Fluorochem*, *Acros* or *AGTC Bioproducts* and all the resins from *Sigma Aldrich*.

4.3 Yeast Treatment

4.3.1 In-vitro Assays on Microplate Reader

***S. pombe* treatment:** Overnight cultures of the *S. pombe* strains were grown in EMM in a shaker incubator (180 rpm, 32 °C). The cultures were then brought to 0.2 OD_{595nm} by diluting with EMM (final volume 5.0 mL). To each batch, 167 μ L of P-factor solution (1.0 mg in 1.0 mL of HPLC grade sterile water, final concentration 12 μ M) were added and the cultures again placed in the shaker incubator for 6 hours. To completely remove EMM, the yeast cultures were subjected four times to precipitation by centrifugation (5 min, 2000 rpm) and resuspension in sterile HPLC grade water. The pellets were finally resuspended in 5.0 mL water and adjusted to 0.4 OD_{595nm}. 400 μ L of these suspensions were added to a 96 well-plate for the measurement.

Measurement parameters: For reproducibility purposes, each *S. pombe* strain was examined in replicates and the mean of each measurement was calculated. The parameters of the measurements are presented in Table 4.1a.

Table 4.1: Experimental parameters for in-vitro assays.

(a) Microplate reader		(b) Automated cell counter	
Parameter	Value	Parameter	Value
Excitation filter	480/30 nm	Dilution factor	1.11
Emission filter	530/30 nm	Min. fluorescence object size	1 μ m
Measurement direction	top	Min. fluorescence object size	60 μ m
Measurement time (ms)	1000	Size gating	1-60 μ m
Z-focus (mm)	10.9	Exposure green fluorescence	7
Excitation spot size (mm)	0.5	Exposure red fluorescence	3
Emission spot size (mm)	1	Light green	13
Flash energy (mJ)	10	Light red	13
PMT HV (V)	612		

4.3.2 *S. pombe* Viability Assay

***S. pombe* treatment:** The overnight culture (grown as described in section 4.3.1) was directly subjected to four cycles of precipitation by centrifugation (5 min, 5000 rpm) and resuspension in Cell Dilution Buffer II (*BioCat GmbH*) to increase cell membrane permeability and thus enable dye uptake. The final yeast suspension was adjusted to 0.1 OD_{595nm}. To 15 μ L of the diluted cell culture 5 μ L Acridine Orange/Propidium Iodide mixture (*BioCat GmbH*) were added and the resulting suspension was incubated for 10 min in a static incubator (32 °C). Finally, 10 μ L of the suspension were applied on a LUNATM 2-Channel Cell Counting Slide and placed in the cell counter.

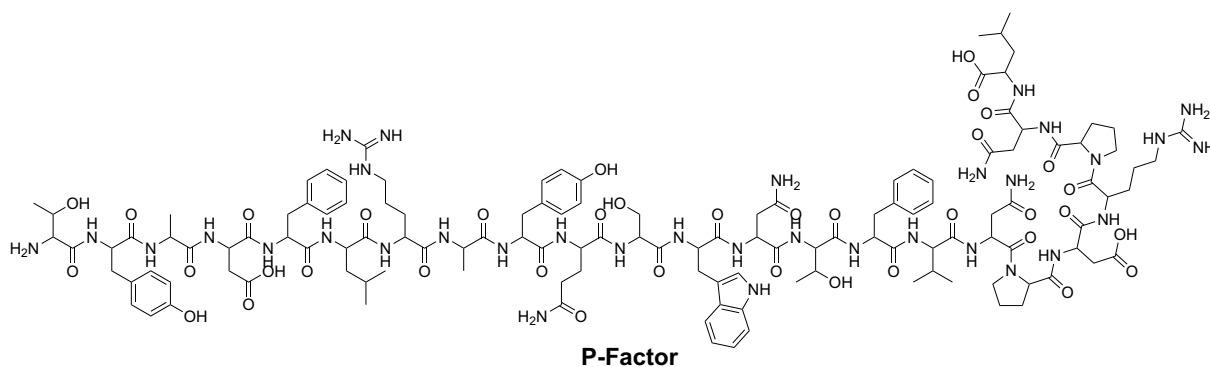
Measurement parameters: Upon insertion of the slide the fluorescently labelled cells are visualized by manually adjusting the focus. The adjusted measurement parameters are presented in Table 4.1b.

4.4 Solid Phase Peptide Synthesis

4.4.1 Peptide Synthesizer

For the peptide synthesis a semi-automatic solid phase peptide synthesizer (Focus XC from *aapptec*) was used. The reactions take place in its two 50.0 mL reaction vessels (RVs) which are equipped with glass frits at the bottom that keep the resin in the vessel during the waste removal by filtration. The amino acids (dissolved in NMP) are stored in 24 90.0 mL amino acid vials before being taken for reaction. All reagents (except piperidine) are measured in one of the two measuring vessels (MVs) before being transferred to the RV. The amounts delivered are controlled by software commands with *ALL* corresponding to the whole volume from the amino acid vial and *LS1*, *LS2* and *LS3* corresponding to 2.5 mL, 5.0 mL and 7.5 mL, respectively. The latter volumes are recognized by a photocell. Finally, the instrument is connected to six solvent bottles containing NMP and DMF as washing solvents, a stock solution of the Fmoc-deprotecting agent piperidine, HATU as coupling reagent and DIPEA. The instrument is constantly pressurized with Ar which was set to 40 psi at the cylinder regulator (source) and adjusted to 11 psi using the instrument regulator. During the pre-activation of the amino acid in the MV, the solution is mixed by an argon flush before being added to the RV and coupled to the resin.

4.4.2 Preparation of P-Factor



The P-factor was synthesized on a 200 mg scale. Fmoc-Leu Wang resin (100-200 mesh, crosslinked with 1% DVB, 0.650 mmol/g loading) was used as a solid support (Figure 4.1). The concentrations and quantities of the used amino acids and reactants are shown in Table 4.2.

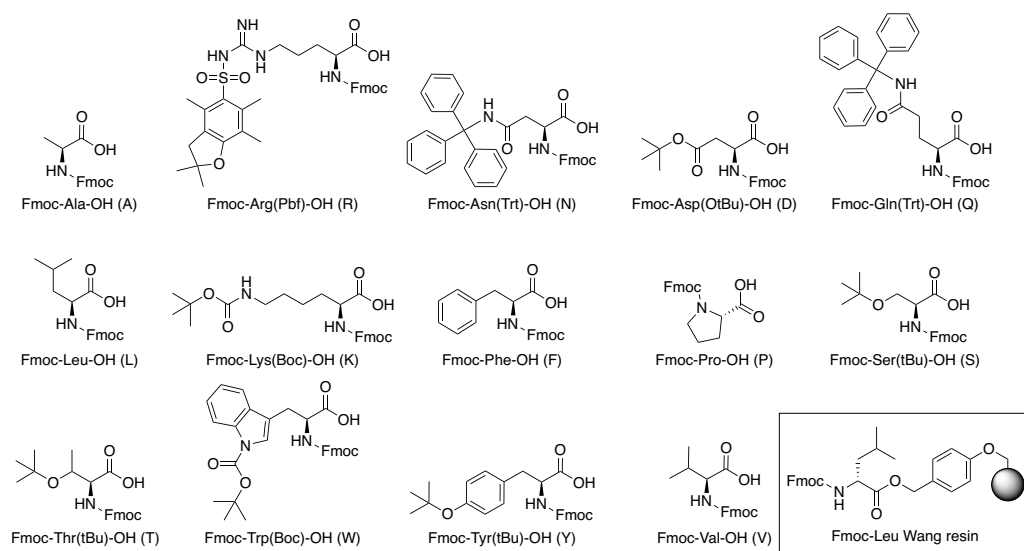


Figure 4.1: Commercial amino acids and Wang resin used in SPPS.

Table 4.2: Quantities and concentrations used in SPPS for a 200 mg scale (resin loading: 0.650 mmol/g).

Reagents/ Solvents	c (M)	V (mL)	n (mmol)	eq	M (g/mol)	Mass (mg)
Resin	-	-	0.125	1	-	200
Fmoc-amino acid-OH	0.26	2.4	0.624	5	*	*
HATU	0.30	1.8	0.540	4.5	380.23	205
DIPEA	2.0	1	2	16	129.25	259

*depends on amino acid

The synthesis was performed at RT and under protection of Ar. Four major protocols were run (exact software commands are shown in the appendix):

Protocol 1: Swell and Coupling The resin is first swollen in 6.0 mL NMP for 60 minutes and then the N-terminal Fmoc protection group of the preloaded amino acid is removed by addition of 20% piperidine (3 cycles). To couple the second amino acid, it is first activated in a mixture of 0.3 M HATU in DMF and 2.0 M DIPEA in NMP. The solution is then added to the RV and the coupling is performed for 25 minutes.

Protocol 2: Coupling and Deprotection The N-terminal Fmoc protection group is removed by addition of 20% piperidine (3 cycles). The amino acid that is next in the sequence is pre-activated in a mixture of 0.3 M HATU in DMF and 2.0 M DIPEA in NMP before being added to the RV. The coupling takes place for 25 minutes. This protocol is performed for the amino acids 2-22.

Protocol 3: Final Coupling The final coupling is performed similarly to protocol 2. After finishing the sequence, the resin was stored as suspension in NMP in an external vial. It was kept at 4 °C until proceeding with protocol 4 for which it was transferred back to the RV.

Protocol 4: Final Deprotection The last N-terminal Fmoc protection group is removed by addition of 20% piperidine (3 cycles). Finally, the resin is washed with DMF and NMP.

4.4.3 P-Factor Isolation

The resin was first washed with 50.0 mL CH_2Cl_2 and then treated with 5.0 mL cleavage cocktail consisting of TFA, phenol, H_2O and TIPS in a ratio of 8.8:0.5:0.5:0.2 for 4 hours^[116]. The successful cleavage from the resin was indicated by a color change from red/orange to yellow. To remove the resin, the solution was filtered through glass wool and the solvent evaporated by applying an Ar stream. The remaining yellow oil was treated with 30.0 mL cold Et_2O , giving a white precipitate that was pelleted by centrifugation (2000 rpm, 5 min). To remove any traces of cleavage cocktail, the pellet was resuspended with another 30.0 mL cold Et_2O and centrifuged again. This step was repeated one more time before drying the pellet in an Ar stream and storing it at 4 °C.

4.4.4 P-Factor Analysis, Purification & Identification

The isolated crude P-factor was analyzed by analytical reverse phase HPLC with 0.1% TFA in H_2O (A) and 0.1% TFA in CH_3CN (B) as eluents. The flow rate was set to 0.75 mL/min and the gradient reported in Table 4.3 applied. The detection wavelengths chosen were $\lambda_1=254$ nm and $\lambda_2=210$ nm^[117]. Subsequently, the crude peptide was purified by preparative reverse phase HPLC using the same eluents as above but with a 10.0 mL/min flow rate. The desired fractions were pooled and sublimed. The purity of P-factor was confirmed by analytical reverse phase HPLC and ESI-HRMS. HPLC retention time: 15.191 min. MS (ESI-HRMS) $\text{C}_{128}\text{H}_{183}\text{N}_{35}\text{O}_{36}$ requires 2787.3405; found 1395.1790 $[\text{M}+2\text{H}]^{2+}$, 930.4559 $[\text{M}+3\text{H}]^{3+}$, 698.0942 $[\text{M}+4\text{H}]^{4+}$, which resolves to 2790.358 ± 0.0036 [M].

Table 4.3: Gradient used for peptide analysis.

Time (min)	Eluent A	Eluent B
0.00	80.00	20.00
25.00	60.00	40.00
35.00	0.00	100.00
40.00	80.00	20.00

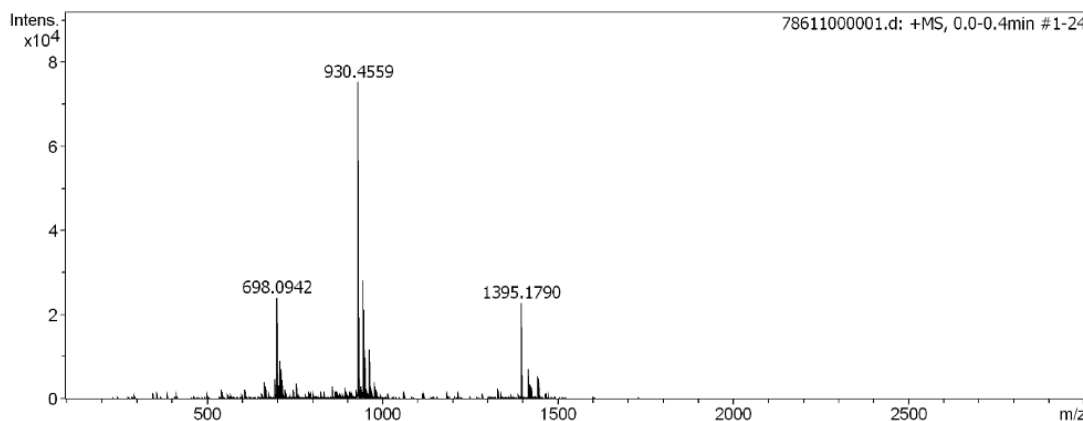


Figure 4.2: ESI-HRMS analysis of pure P-factor.

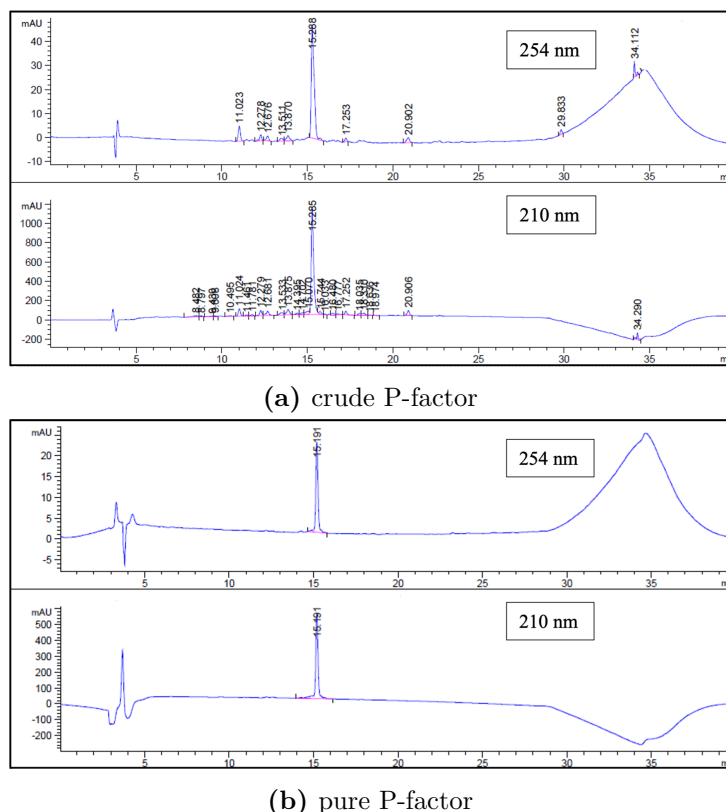


Figure 4.3: HPLC chromatograms of (a) crude and (b) pure P-factor as observed at $\lambda_1=254$ nm and $\lambda_2=210$ nm. At 35 min, there is an increase of absorbance due to the percentage of acetonitrile in the column (100% CH_3CN + 0.1% TFA).

4.4.5 Secondary Structure Analysis of P-Factor by Circular Dichroism

The secondary structure of P-factor in solution was investigated by CD. Different solvents were chosen for the experiments: HPLC grade water at neutral pH as well as solutions of TFE in HPLC grade water (10%, 20% and 50%). Additionally, a 0.1 M acetate buffer (pH 4.8) was prepared with the concentrations shown in Table 4.4. After dissolving the salt and the acid in 40.0 mL distilled H_2O , the pH was adjusted to 4.8 using HCl and NaOH. The buffer was then brought to an endvolume of 50.0 mL with distilled H_2O . All solvents were degassed before usage.

Table 4.4: Preparation of 50.0 mL acetate buffer.

Reagent	Amount (mg)	c (mM)
Sodium acetate	239.56	58
Acetic acid	124.9	42

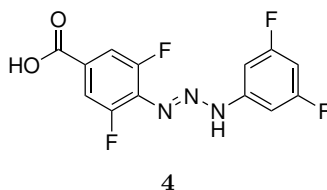
Solutions of 0.5 mg P-factor in 3.0 mL solvent were prepared (final concentration $60.0 \mu\text{M}$) and sonicated to facilitate solvation. $200 \mu\text{L}$ of these solutions were applied on a 0.5 mm quartz cuvette. The cuvette was placed into the CD spectrophotometer and the absorbance acquired against the respective solvent as background. Each sample is measured in a

sequence with 5 repeats and the average of these spectra is determined. After each measurement, the cuvette was cleaned with H₂O and EtOH. Except for the probes with TFE, each sample was measured once per hour over a period of 3 or 4 hours, giving 4 or 5 measurements each, respectively.

4.5 Organic Syntheses

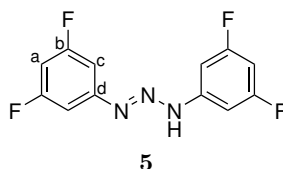
4.5.1 Synthesis of Tetrafluorotriazenes

4-(3-(3,5-difluorophenyl)triaz-1-en-1-yl)-3,5-difluorobenzoic acid **4**



In a 10 mL round-bottomed flask, 4-amino-3,5-difluorobenzoic acid (55.0 mg, 0.318 mmol, 1 eq) was suspended in 1.5 mL of 10% HCl(aq) and cooled to 0 °C. NaNO₂ (19.7 mg, 0.286 mmol, 0.9 eq) was dissolved in 0.30 mL water and added dropwise under vigorous stirring. After 30 minutes, a clear yellow solution of the diazonium salt was obtained and kept at 0 °C. A solution of 3,5-difluoroaniline (41.0 mg, 0.318 mmol, 1 eq) in 2.5 mL water was prepared in a 25 mL round-bottomed flask. To this solution, the diazonium salt was added in portions. After every addition, the reaction mixture was brought to pH 4 using 2.0 M NaOH(aq). The solution was stirred for another hour at RT. The mixture was neutralized with saturated NaHCO₃(aq) and the resulting orange suspension centrifuged. The solvent was decanted and the solid freeze dried to give 71.0 mg of a light orange solid. Despite several purification attempts, it was not possible to obtain the title compound **4** as pure product. MS (ESI-HRMS) C₁₃H₇O₂F₄N₃ calc. for [M-H]⁻ 312.0402; found 312.0404, calc. for [2M+Na-2H]⁻ 647.0694; found 647.0704.

1,3-bis(3,5-difluorophenyl)triaz-1-ene **5**

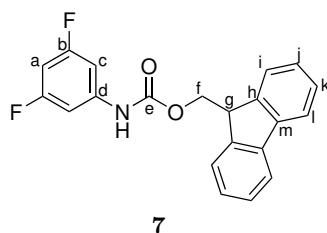


In a 25 mL round-bottomed flask, a suspension of 3,5-difluoroaniline (100.0 mg, 0.775 mmol, 1 eq) in 7.4 mL of 10% HCl(aq) was cooled to 0 °C. NaNO₂ (48.0 mg, 0.697 mmol, 0.9 eq) was dissolved in 1.4 mL water and added dropwise under vigorous stirring. After 30 minutes, a clear yellow solution of the diazonium salt was obtained and kept at 0 °C. A second portion of 3,5-difluoroaniline (100.0 mg, 0.775 mmol, 1 eq) was suspended in 12.0 mL water in a 50 mL round-bottomed flask. To this solution, the diazonium salt was added in

portions over a period of 1 hour. After every addition, the reaction mixture was brought to pH 4 using 2.0 M NaOH(aq). The solution was stirred for another hour at RT. The mixture was neutralized with saturated NaHCO₃(aq) and the resulting orange suspension centrifugated. The solvent was decanted and the solid freeze dried to give 188.0 mg of the title compound **5** as a light orange solid (0.698 mmol, 90%). mp 169-172 °C (decomp.). ¹H NMR (600 MHz, MeOD) δ_H : 7.04–7.00 (m, 4H, CH_c), 6.69 (t, J = 9.0 Hz, 2H, CH_a). ¹³C NMR (151 MHz, MeOD) δ_C : 165.9 (C_b or C_d*), 164.3 (C_b or C_d*), 102.0 (C_c), 100.6 (C_a). ¹⁹F NMR (376 MHz, MeOD) δ_F : -111.1. FTIR (ATR) ν (cm⁻¹): 3324 (m, C-N), 3095 (w, C-H_{ar}), 1596 (s, N=N), 1479 (s, C-C_{ar}), 1243 (s, C-F). MS (ESI-HRMS) C₁₂H₇F₄N₃ calc. for [M-H]⁻ 268.0503; found 268.0502, calc. for [2M+Na-2H]⁻ 559.0896; found 559.0912. *the carbons b and d are not distinguishable in the spectra.

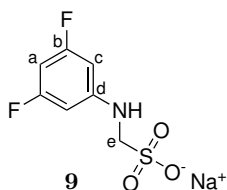
4.5.2 Introduction of Protection Groups to the Aromatic Precursors

(9H-Fluoren-9-yl)methyl (3,5-difluorophenyl)carbamate **7**

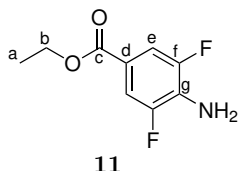


7

In a 25 mL round-bottomed flask, 9H-fluorenylmethoxycarbonyl chloride (0.87 g, 6.71 mmol, 1 eq) and 3,5-difluoroaniline (2.08 g, 8.02 mmol, 1.2 eq) were dissolved in 9.0 mL water. The reaction was stirred at 60 °C for 2 hours. The product was filtered off and washed with water. The precipitate was dried on the freeze dryer. The crude solid was dissolved in EtOAc (5.0 mL) and precipitated by adding petroleum ether (45.0 mL). After centrifugation and decantation of the supernatant, the solid was washed once with petroleum ether (15.0 mL). The remaining solvent was evaporated *in vacuo* to give compound **7** as white crystals (1.65 g, 4.70 mmol, 70%). mp 181-184 °C. ¹H NMR (600 MHz, MeOD) δ_H : 7.81 (d, J = 7.5 Hz, 2H, CH_i), 7.70 (d, J = 7.5 Hz, 2H, CH_j), 7.4 (t, J = 7.4 Hz, 2H, CH_k), 7.33 (td, J = 7.5, 1.1 Hz, 2H, CH_j), 7.09 (d, J = 6.8 Hz, 2H, CH_c), 6.55 (tt, J = 9.2, 2.3 Hz, 1H, CH_a), 4.52 (d, J = 6.5 Hz, 2H, CH_f), 4.27 (t, J = 6.5 Hz, 1H, CH_g), NH not visible due to fast exchange rate with MeOD. ¹³C NMR (151 MHz, MeOD) δ_C : 165.6 (C_d), 164.0 (C_b), 155.2 (C_e), 145.1 (C_h), 142.7 (C_m), 128.9 (C_k), 128.2 (C_j), 126.1 (C_i), 121.0 (C_l), 102.3 (C_c), 98.5 (C_a), 67.8 (C_f), 48.3 (C_g). ¹⁹F NMR (376 MHz, MeOD) δ_F : -111.6. FTIR (ATR) ν (cm⁻¹): 3331 (m, C-N), 3062 (w, C-H_{ar}), 2979 (w, C-H), 1704 (s, C=O), 1603 (m, N-H), 1443 (m, C-C_{ar}), 1245 (s, C-F). MS (ESI-HRMS) C₂₁H₁₅O₂F₂N calc. for [M+Na]⁺ 374.0963; found 374.0962, calc. for [2M+Na]⁺ 725.2031; found 725.2026.

((3,5-Difluorophenyl)amino)methanesulfonate **9**

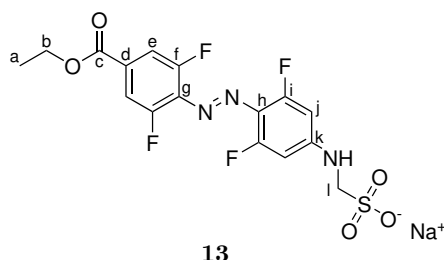
In a 100 mL round-bottomed flask, NaHSO_3 (19.0 g, 183 mmol, 1.47 eq) was dissolved in 45.0 mL water. To this, formaldehyde (13.8 mL of a 38% aqueous solution, 185 mmol, 1.49 eq) and 3,5-difluoroaniline (16.0 g, 124 mmol, 1 eq) were added under vigorous stirring. The reaction was heated to 100 °C reflux for 40 minutes and then allowed to cool down to RT. The evolving white solid was filtered off and washed with a small amount of EtOH (36.0 mL) to obtain 28.1 g of compound **9** as a white powder (183 mmol, 93%). mp >260 °C. ^1H NMR (400 MHz, MeOD) δ_{H} : 6.42–6.35 (m, 2H, CH_c), 6.11 (tt, $J = 9.3, 2.2$ Hz, 1H, CH_a), 4.27 (s, 2H, CH_e), NH not visible due to fast exchange rate with MeOD. ^{13}C NMR (176 MHz, MeOD) δ_{C} : 164.0 (C_b), 150.2 (C_d), 95.4 (C_c), 91.2 (C_a), 60.4 (C_e). ^{19}F NMR (376 MHz, MeOD) δ_{F} : -113.2. FTIR (ATR) ν (cm^{-1}): 3350 (m, C-N), 3092 (w, C- H_{ar}), 1596 (s, N-H), 1484 (m, C- C_{ar}), 1362 (w, S=O), 1247 (m, C-F). MS (ESI-HRMS) $\text{C}_7\text{H}_6\text{O}_3\text{F}_2\text{NS}$ calc. for $[\text{M-H}]^-$ 222.0041; found 222.0042, calc. for $[\text{2M+Na-2H}]^-$ 466.9976; found 466.9979.

Ethyl 4-amino-3,5-difluorobenzoate **11**

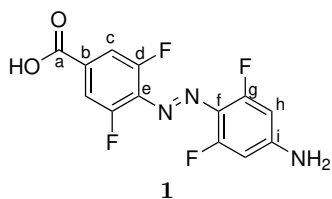
A 25 mL round-bottomed flask was charged with a solution of 4-amino-3,5-difluorobenzoic acid (100.0 mg, 0.578 mmol, 1 eq) in 5.0 mL EtOH. It was acidified with 3 drops of concentrated H_2SO_4 and heated to 80 °C reflux for 24 hours. The reaction was neutralized using saturated $\text{NaHCO}_3(\text{aq})$ and extracted with CH_2Cl_2 (3 x 15.0 mL). The organic layer was dried over MgSO_4 , filtered and the solvent evaporated *in vacuo* to give the title compound **11** as a brown precipitate (94.0 mg, 0.467 mmol, 81%). mp 65–67 °C. ^1H NMR (700 MHz, MeOD) δ_{H} : 7.47–7.41 (m, 2H, CH_e), 4.30 (q, $J = 7.1$ Hz, 2H, CH_b), 1.35 (t, $J = 7.1$ Hz, 3H, CH_a), NH not visible due to fast exchange rate with MeOD. ^{13}C NMR (176 MHz, MeOD) δ_{C} : 166.8 (C_c), 151.8 (C_f), 132.1 (C_g), 117.5 (C_d), 113.3 (C_e), 62.0 (C_b), 14.6 (C_a). ^{19}F NMR (376 MHz, MeOD) δ_{F} : -134.8. FTIR (ATR) ν (cm^{-1}): 3362 (s, C-N), 2984 (m, C- H_{ar}), 2916 (w, C-H), 1767 (s, C=O), 1631 (s, N-H), 1468 (m, C- C_{ar}), 1221 (s, C-F). MS (ESI-HRMS) $\text{C}_9\text{H}_9\text{O}_2\text{F}_2\text{N}$ calc. for $[\text{M+Na}]^+$ 224.0494; found 224.0491.

4.5.3 Synthesis of *ortho*-Tetrafluoroazobenzenes

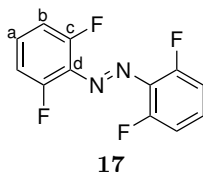
((4-((4-(ethoxycarbonyl)-2,6-difluorophenyl)diazenyl)-3,5-difluorophenyl)amino)-methanesulfonate **13**



In a 10 mL round-bottomed flask, benzoate **11** (58.1 mg, 0.289 mmol, 1 eq) was dissolved in a solution of 1.1 mL CH₃CN and 0.3 mL conc. HCl, and cooled to 0 °C. NaNO₂ (19.9 mg, 0.289 mmol, 1 eq) was dissolved in 0.3 mL water and added dropwise under vigorous stirring. After 30 minutes, a clear yellow solution of the diazonium salt was obtained and kept at 0 °C. A solution of methanesulfonate **9** (70.8 mg, 0.289 mmol, 1 eq) in 2.3 mL water was prepared in a 25 mL round-bottomed flask at 0 °C. To this suspension, the diazonium salt was added dropwise. The dark red reaction mixture was stirred at 0 °C for 30 minutes before another 10.0 mg of methanesulfonate **9** (0.041 mmol, 0.14 eq) were added. The reaction was stirred another 2 hours until no further reaction progress was observed by TLC. The reaction was extracted with EtOAc, dried over MgSO₄, filtered and the solvent evaporated *in vacuo* to give 80.0 mg of crude product. The crude was purified by flash column chromatography (*Biotage*[®] system) using a gradient of MeOH in EtOAc from 0 to 80% to give 58.0 mg of the title compound **13** as a red solid (0.127 mmol, 44%). mp 227-229 °C (decomp.). ¹H NMR (600 MHz, MeOD) δ_H : 7.70–7.67 (m, 2H, CH_e), 6.55 (d, *J* = 12.8 Hz, 2H, CH_j), 4.41 (q, *J* = 7.1 Hz, 2H, CH_b), 4.36 (s, 2H, CH_l), 1.41 (t, *J* = 7.1 Hz, 3H, CH_a), NH not visible due to fast exchange rate with MeOD. ¹³C NMR (151 MHz, MeOD) δ_C : 165.4 (C_c), 160.0 (C_i), 156.1 (C_f), 154.6 (C_k), 137.0 (C_g), 132.1 (C_d), 124.1 (C_h), 114.4 (C_e), 97.4 (C_j), 63.0 (C_b), 60.8 (C_l), 14.5 (C_a). ¹⁹F NMR (376 MHz, MeOD) δ_F : -118.2, -123.9. FTIR (ATR) ν (cm⁻¹): 3441 (m, C-N), 3081 (w, C-H_{ar}), 2983 (w, C-H), 1716 (s, C=O), 1629 (s, N-H), 1573 (s, N=N), 1476 (w, C-H), 1434 (m, C-C_{ar}), 1359 (s, S=O), 1237 (s, C-F). MS (ESI-HRMS) C₁₆H₁₂O₅F₄N₃S calc. for [M-H]⁻ 434.0439; found 434.0438.

4-((4-amino-2,6-difluorophenyl)diazenyl)-3,5-difluorobenzoic acid **1**

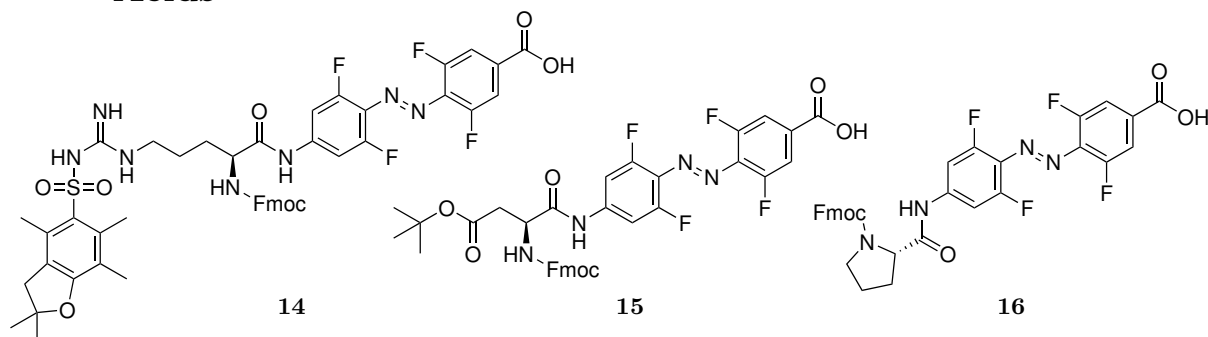
In a 100 mL round-bottomed flask, compound **13** (500.0 mg, 1.09 mmol, 1 eq) was dissolved in 65.0 mL 5% NaOH(aq) and heated to 70 °C for 30 minutes. The reaction was allowed to reach RT and acidified to pH 1 using 10% HCl(aq). The product was extracted with EtOAc (3x 50.0 mL), dried over MgSO₄, filtered and the solvent evaporated *in vacuo*. The crude was purified by flash column chromatography (*Biotage*[®] system) using MeOH in EtOAc from 0 to 20% to give 153 mg of the final product **1** as a red solid (0.488 mmol, 45%). mp 221-224 °C (decomp.). ¹H NMR (400 MHz, MeOD) δ_H : 7.68 (q, $J = 3.5$ Hz, 2H, CH_c), 6.32 (q, $J = 2.2$ Hz, 2H, CH_h), NH not visible due to fast exchange rate with MeOD. ¹³C NMR (151 MHz, MeOD) δ_C : 167.8 (C_a), 160.3 (C_g), 156.6 (C_i), 156.0 (C_d), 136.6 (C_e), 133.7 (C_b), 123.4 (C_f), 114.5 (C_c), 97.6 (C_h). ¹⁹F NMR (376 MHz, MeOD) δ_F : -118.4, -124.6. FTIR (ATR) ν (cm⁻¹): 3374 (s, C-N), 3067 (w, C-H_{ar}), 1710 (s, C=O), 1631 (s, N-H), 1573 (s, N=N), 1476 (m, C-C_{ar}), 1218 (s, C-F). MS (ESI-HRMS) C₁₃H₇O₂F₄N₃ calc. for [M-H]⁻ 312.0402; found 312.0401, calc. for [2M+Na-2H]⁻ 647.0696; found 647.0712, calc. for [M+H]⁺ 314.0547; found 314.0547, calc. for [M+Na+H]²⁺ 336.0367; found 336.0368, calc. for [2M+2Na-H]⁺ 671.0660; found 671.0665.

1,2-bis(2,6-difluorophenyl)diazene **17**

A 100 mL round-bottomed flask was charged with 2,6-difluoroaniline (500.0 mg, 3.87 mmol, 2 eq). To this, a solution of freshly ground KMnO₄ (2.0 g, 12.7 mmol) and FeSO₄·7H₂O (1.95 g, 7.01 mmol) in 20.0 mL CH₂Cl₂ was added and the mixture was refluxed at 40 °C for 16 hours. The reaction was allowed to reach RT and filtered through celite. The filtrate was dried over MgSO₄ and the solvent evaporated *in vacuo*. The crude was purified by flash column chromatography (heptane/EtOAc 9:1) to give 18.0 mg of title compound **17** as orange crystals (0.070 mmol, 4%). The compound was obtained in a mixture of *cis* and *trans* in a ratio of 1:1.86 as determined by ¹H NMR. The peaks were assigned to the respective isomer after isomerization studies (compare section 4.6.2). mp 68-71 °C. ¹H NMR (700 MHz, MeOD) (*trans* isomer) δ_H : 7.54 (tt, $J = 8.5, 5.9$ Hz, 2H, CH_a), 7.23-7.18 (m, 4H, CH_b). ¹³C NMR (151 MHz, MeOD) (*trans* isomer) δ_C : 156.8 (C_c), 133.3 (CH_a), 133.0 (C_d), 113.8 (C_b). ¹⁹F NMR (376 MHz, MeOD) δ_F : -122.0 (*cis*)

-123.6 (*trans*). FTIR (ATR) ν (cm⁻¹): 3105 (w, C-H_{ar}), 1586 (s, N=N), 1471 (m, C-C_{ar}), 1244 (s, C-F). MS (ESI-HRMS) C₁₂H₆F₄N₂S calc. for [M+Na]⁺ 277.0359; found 277.0358.

4.5.4 Attempts to Couple *ortho*-Tetrafluoroazobenzene to Amino Acids



In the following procedures the amounts are given for the arginine coupling to yield **14**. For aspartate (**15**) and proline couplings (**16**), similar equivalents and concentrations were used and have to be adapted accordingly. The procedures written below constitute the basis methods which were further adapted as described in section 3.2.2.

Procedure a In a 10 mL round-bottomed flask, a solution of Fmoc-Arg(Pbf)-OH (207 mg, 0.319 mmol, 1 eq), HATU (121 mg, 0.319 mmol, 1 eq) and HOAt (69.5 mg, 0.511 mmol, 1.6 eq) in 2.6 mL NMP was prepared. After complete dissolution, 111 μ L DIPEA (0.639 mmol, 2 eq) were dropped into the reaction and it was stirred for 5 minutes at RT. Azobenzene **1** (100.0 mg, 0.319 mmol, 1 eq) was dissolved in 1.3 mL NMP and added to the stirred solution. After 14 hours, the reaction was poured into 25.5 mL of ice-cooled water to precipitate the product, which was then filtered and washed with cold water. The solid was redissolved in EtOAc and dried over MgSO₄. The solvents were then evaporated *in vacuo*. MS or TLC did not confirm the presence of the product but still showed starting materials.

Procedure b A 10 mL round-bottomed flask was charged with a solution of Fmoc-Arg(Pbf)-OH (162 mg, 0.250 mmol, 1 eq) and TFP (49.8 mg, 0.300 mmol, 1.2 eq) in 1.25 mL CH₂Cl₂. To this, DCC (51.6 mg, 0.250 mmol, 1 eq) were added and the reaction was stirred at RT for 2 hours. The mixture was diluted with 12.5 mL acetone and filtered to remove DCU. The filtrate was concentrated *in vacuo* and dissolved in 0.625 mL DMF to be used in the next step. In a 10 mL round-bottomed flask, azobenzene **1** (94.0 mg, 0.300 mmol, 1.2 eq) and 174 μ L DIPEA (1.00 mmol, 4 eq) were dissolved in 2.0 mL DMF and cooled to 0 °C. To this, the activated amino acid was added and the reaction stirred at RT for 72 hours. The mixture was poured into 12.5 mL brine, extracted with CH₂Cl₂ and washed with 1.0 M HCl(aq) and brine. The organic phase was dried over MgSO₄ and the solvent evaporated *in vacuo*. MS or TLC did not confirm the presence of the product but still showed starting materials.

4.6 Ultraviolet Light Experiments

4.6.1 Absorbance Spectra of Triazenes

1.57 mg of triazene **4** and 1.35 mg of triazene **5** were each dissolved in 5.0 mL MeOH to reach a final concentration of 1.0 mM. The samples were transferred into a quartz cuvette with an optical path of 1 cm and placed in the spectrophotometer. The cuvettes were irradiated at the two peaks (232 nm and 348 nm) for 45 minutes each and another spectra was acquired after each irradiation period.

4.6.2 Isomerization Studies

Thermal relaxation To form 100% of the *trans* isomers 12.0 mg of azobenzene **17** and 40.0 mg of **1** were each dissolved in 2.0 mL DMSO and heated to 100 °C reflux overnight under protection from light. The solvent was removed by freeze drying. NMR confirmed the absence of the *cis* isomer.

Isomerization of 17 1.27 mg of thermally relaxed azobenzene **17** were dissolved in 5.0 mL MeOH and again diluted by factor 2 to reach a final concentration of 0.5 mM. The sample was transferred into a quartz cuvette with an optical path of 1 mm and placed in the spectrophotometer. After acquiring the spectra, the cuvette was irradiated at 505 nm, 380 nm and 340 nm for 25 minutes each. After each irradiation period another spectrum was acquired. To define the PSS, NMR spectra were also taken after each irradiation period.

Isomerization of 1 1.57 mg of thermally relaxed azobenzene **1** were dissolved in 5.0 mL MeOH and again diluted by factor 2 to reach a final concentration of 0.5 mM. The sample was transferred into quartz cuvettes with an optical path of 1 mm and placed in the spectrophotometer. After acquiring the spectra, the cuvette was irradiated at 510 nm, 490 nm, 450 nm, 430 nm, 375 nm and 340 nm 20 minutes each. After each irradiation period another spectrum was acquired. As only the irradiation at 375 nm led to a significant change in the absorption spectra, a NMR spectra was only taken after this irradiation period. To mimic more physiological conditions, another 1.57 mg of thermally relaxed azobenzene **1** were dissolved in 20.0 mL HPLC grade H₂O and again diluted by factor 5 to reach a final concentration of 0.05 mM. The experiment was repeated as stated above but by using a quartz cuvettes with an optical path of 1 cm.

Stability of *cis* isomer After irradiating the above 0.5 mM solution of compound **17** at 505 nm to form a high PSS of the *cis* isomers, the sample was placed in the spectrophotometer and heated to 25 °C for 17 hours. A spectra was acquired every 15 minutes. As the relaxation for compound **1** happens within a few minutes, the above described MeOH sample of this compound was irradiated at 375 nm for 30 minutes and placed in the

spectrophotometer at RT (25 °C). A spectra was acquired every 2 minutes for a period of 44 minutes.

4.6.3 Fatigue of Isomerization

A sample of 0.50 mM solution azobenzene **17** in MeOH was prepared as described in section 4.6.2. After acquiring the UV spectra, the cuvette was irradiated at 505 nm for 30 minutes and another spectra was measured. The sample was then irradiated at 380 nm for 30 minutes before the next measurement. These steps were repeated another 8 times. Azobenzene **1** was treated in a similar manner. The irradiation wavelengths were 375 nm and 340 nm. Additionally, a second sample was prepared as described. Instead of irradiating at 340 nm, it was allowed to relax thermally between the irradiations at 375 nm (30 minutes).

4.7 Appendix

4.7.1 Protocols for Solid Phase Peptide Synthesis

Protocol 1

****Cycle 1: Swell and Coupling****

******* Start Cycle 1 *******

1. Fill RV with 6.00 mL from NMP
2. Mix RV via Mech / N2 Mix for 30.00 min
3. Empty RV to waste 1
4. Fill RV with 6.00 mL from NMP
5. Mix RV via Mech / N2 Mix for 30.00 min
6. Empty RV to waste 1
7. Fill RV with 6.00 mL from Pip.
8. Mix RV via Mech / N2 Mix for 4.00 min
9. Empty RV to waste 1
10. Fill RV with 6.00 mL from Pip.
11. Mix RV via Mech / N2 Mix for 4.00 min
12. Empty RV to waste 1
13. Fill RV with 6.00 mL from Pip.
14. Mix RV via Mech / N2 Mix for 4.00 min
15. Empty RV to waste 1
16. Wash RV with 15.00 mL from DMF 4 times and hold for 0.20 Seconds
17. Wash Line with 15.00 mL from DMF to Waste 1
18. Wash MV with 15.00 mL from DMF to Waste 1
19. Fill MV from HATU with 1.80 ml
20. Fill MV with ASN-AA1 (ALL AA)
21. Fill MV from DIEA with 1.00 ml
22. Mix MV via Nitrogen Mix for 1.00 min
23. Fill RV from MV (ALL)
24. Wash MV with 15.00 mL from DMF to Waste 1
25. Wash Line with 15.00 mL from DMF to Waste 1
26. Mix RV via Mech / N2 Mix for 25.00 min
27. Empty RV to waste 1
28. Fill RV with 6.00 mL from NMP
29. Mix RV via Mechanical Mix for 1.00 min
30. Empty RV to waste 1
31. Wash RV with 15.00 mL from DMF 2 times and hold for 0.20 Seconds

Protocol 2

****Cycle 2: RV – T****

******* Start Cycle 2 *******

32. Fill RV with 6.00 mL from Pip.
33. Mix RV via Mech / N2 Mix for 4.00 min
34. Empty RV to waste 1
35. Fill RV with 6.00 mL from Pip.
36. Mix RV via Mech / N2 Mix for 4.00 min
37. Empty RV to waste 1
38. Fill RV with 6.00 mL from Pip.
39. Mix RV via Mech / N2 Mix for 4.00 min

40. Empty RV to waste 1
41. Wash RV with 15.00 mL from DMF 4 times and hold for 0.20 Seconds
42. Wash Line with 15.00 mL from DMF to Waste 1
43. Wash MV with 15.00 mL from DMF to Waste 1
44. Fill MV from HATU with 1.80 ml
45. Fill MV with PRO-AA2 (ALL AA)
46. Fill MV from DIEA with 1.00 ml
47. Mix MV via Nitrogen Mix for 1.00 min
48. Fill RV from MV (ALL)
49. Wash MV with 15.00 mL from DMF to Waste 1
50. Wash Line with 15.00 mL from DMF to Waste 1
51. Mix RV via Mech / N2 Mix for 25.00 min
52. Empty RV to waste 1
53. Fill RV with 6.00 mL from NMP
54. Mix RV via Mechanical Mix for 1.00 min
55. Empty RV to waste 1
56. Wash RV with 15.00 mL from DMF 2 times and hold for 0.20 Seconds

Protocol 3

Cycle 3: RV – T

**** Start Cycle 3 ****

57. Fill RV with 6.00 mL from Pip.
58. Mix RV via Mech / N2 Mix for 4.00 min
59. Empty RV to waste 1
60. Fill RV with 6.00 mL from Pip.
61. Mix RV via Mech / N2 Mix for 4.00 min
62. Empty RV to waste 1
63. Fill RV with 6.00 mL from Pip.
64. Mix RV via Mech / N2 Mix for 4.00 min
65. Empty RV to waste 1
66. Wash RV with 15.00 mL from DMF 4 times and hold for 0.20 Seconds
67. Wash Line with 15.00 mL from DMF to Waste 1
68. Wash MV with 15.00 mL from DMF to Waste 1
69. Fill MV from HATU with 1.80 ml
70. Fill MV with ARG-AA3 (ALL AA)
71. Fill MV from DIEA with 1.00 ml
72. Mix MV via Nitrogen Mix for 1.00 min
73. Fill RV from MV (ALL)
74. Wash MV with 15.00 mL from DMF to Waste 1
75. Wash Line with 15.00 mL from DMF to Waste 1
76. Mix RV via Mech / N2 Mix for 25.00 min
77. Empty RV to waste 1
78. Fill RV with 6.00 mL from NMP
79. Mix RV via Mechanical Mix for 1.00 min
80. Empty RV to waste 1
81. Wash RV with 15.00 mL from DMF 2 times and hold for 0.20 Seconds
82. Wash RV with 15.00 mL from NMP 2 times and hold for 0.20 Seconds
83. Fill RV with 5.00 mL from NMP

** End of Synthesis**

Storage

Protocol 4

****Cycle 4: Final Deprotection****

****** Start Cycle 4 ******

84. Fill RV with 6.00 mL from Pip.
85. Mix RV via Mech / N2 Mix for 4.00 min
86. Empty RV to waste 1
87. Fill RV with 6.00 mL from Pip.
88. Mix RV via Mech / N2 Mix for 4.00 min
89. Empty RV to waste 1
90. Fill RV with 6.00 mL from Pip.
91. Mix RV via Mech / N2 Mix for 4.00 min
92. Empty RV to waste 1
93. Wash RV with 15.00 mL from DMF 4 times and hold for 0.20 Seconds
94. Fill RV with 6.00 mL from NMP
95. Mix RV via Mechanical Mix for 1.00 min
96. Empty RV to waste 1

4.7.2 ESI-HRMS and NMR Spectra

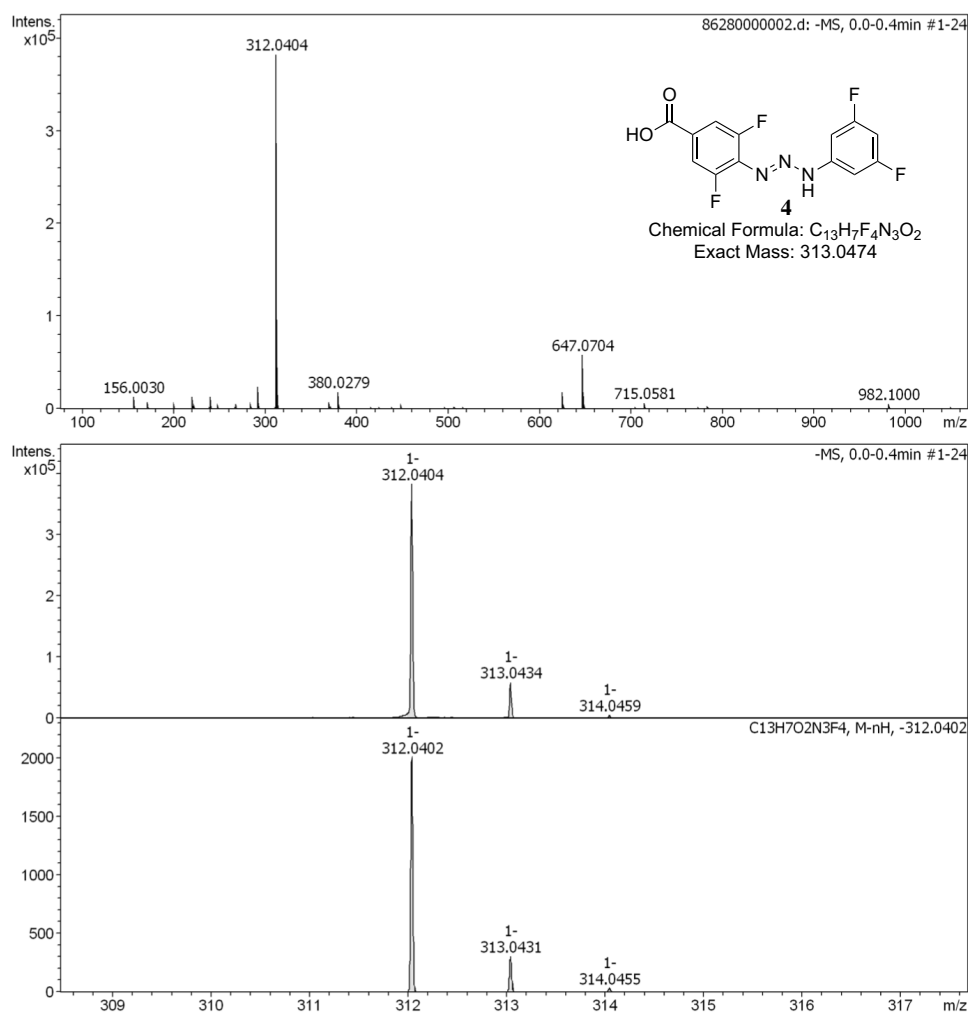


Figure 4.4: ESI-HRMS spectra of **4** ($C_{13}H_7O_2F_4N_3$) in negative mode: calc. for $[M-H]^-$ 312.0402; found 312.0404, calc. for $[2M+Na-2H]^-$ 647.0694; found 647.0704.

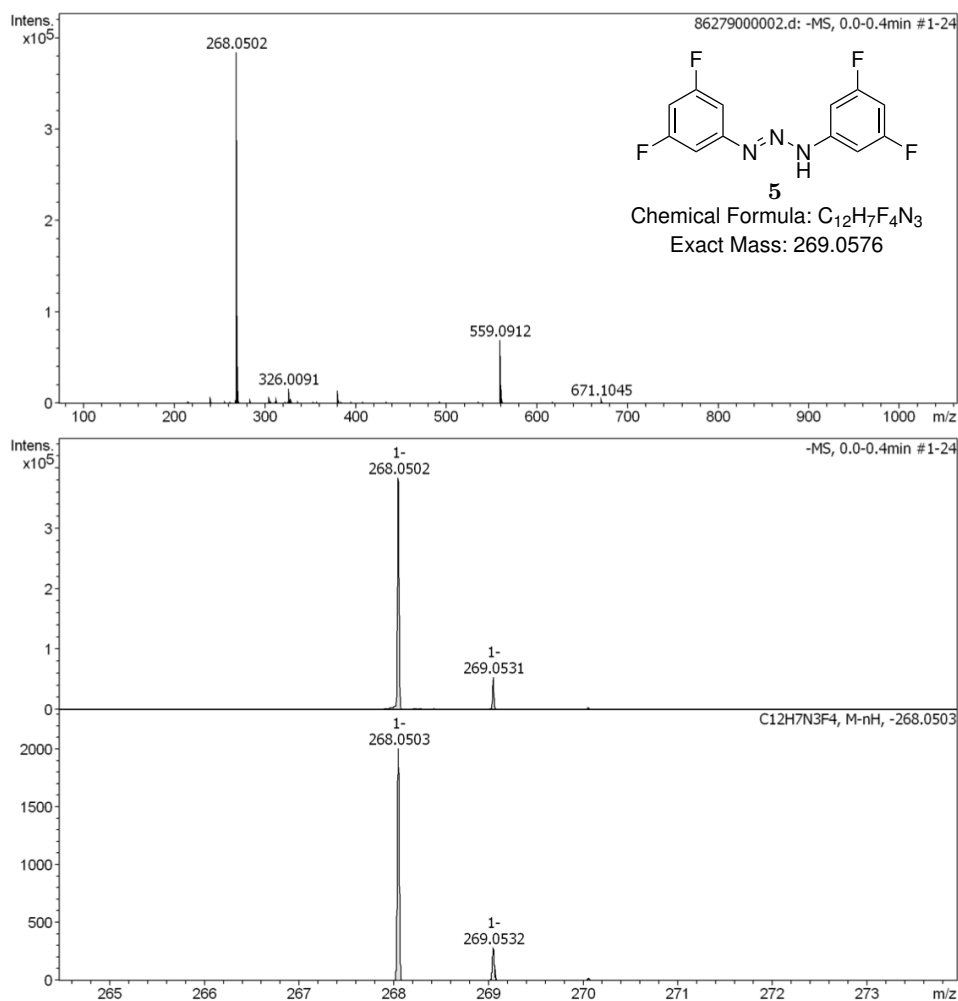
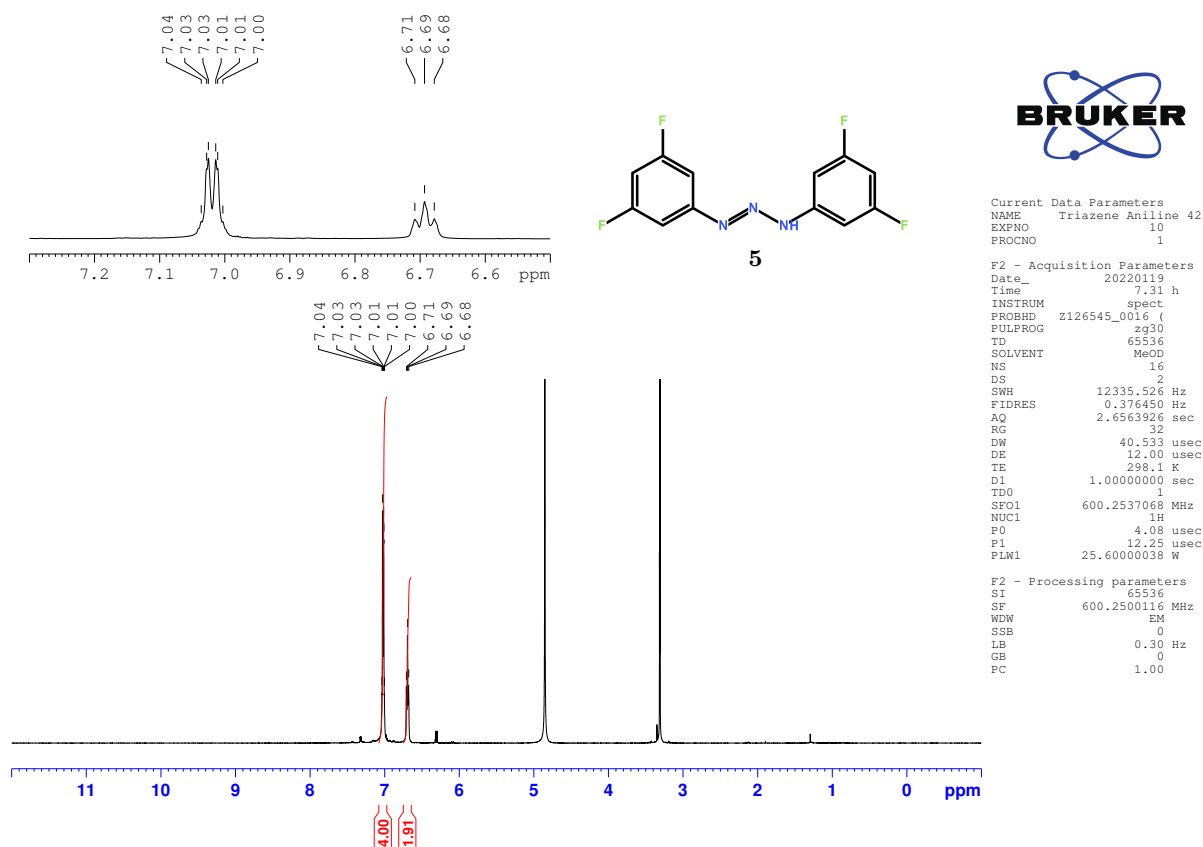
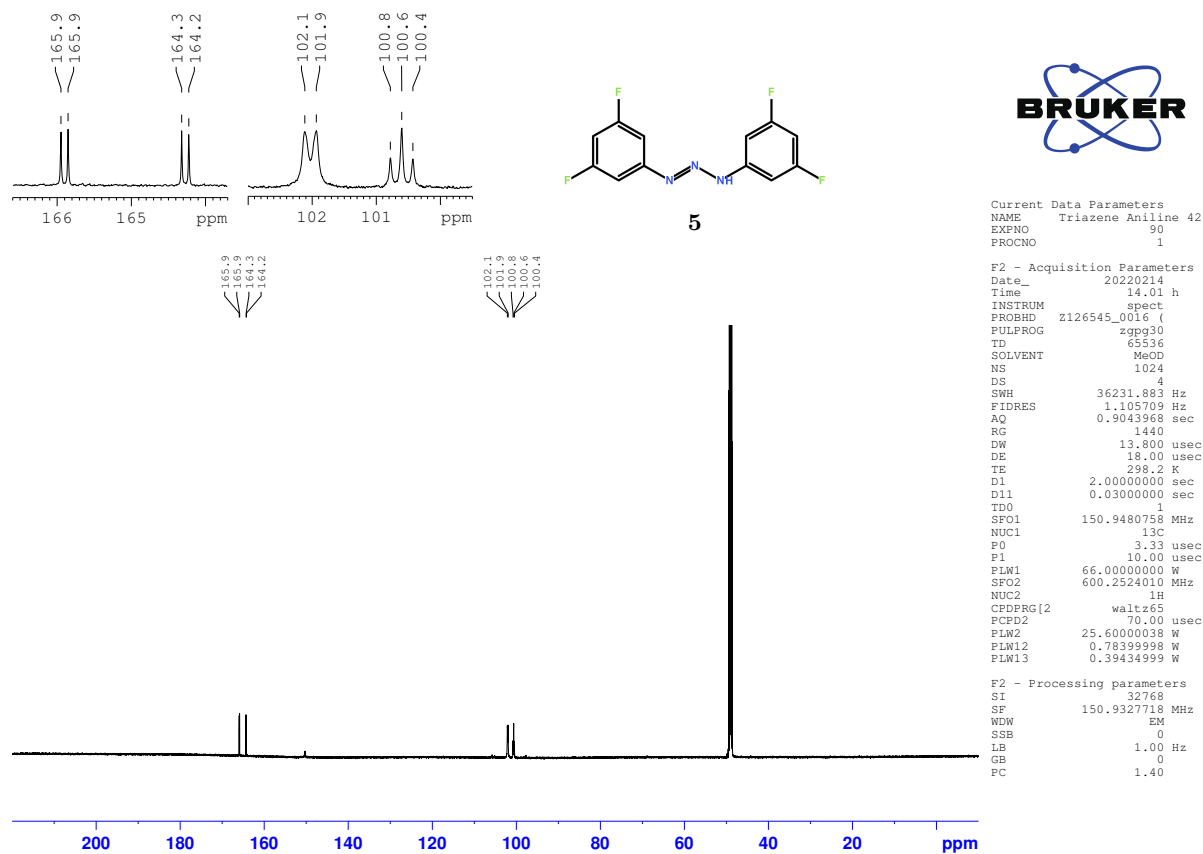


Figure 4.5: ESI-HRMS spectra of **5** ($C_{12}H_7F_4N_3$) in negative mode: calc. for $[M-H]^-$ 268.0503; found 268.0502, calc. for $[2M+Na-2H]^-$ 559.0896; found 559.0912.

Figure 4.6: ^1H NMR spectra of **5** in MeOD.Figure 4.7: ^{13}C NMR spectra of **5** in MeOD.

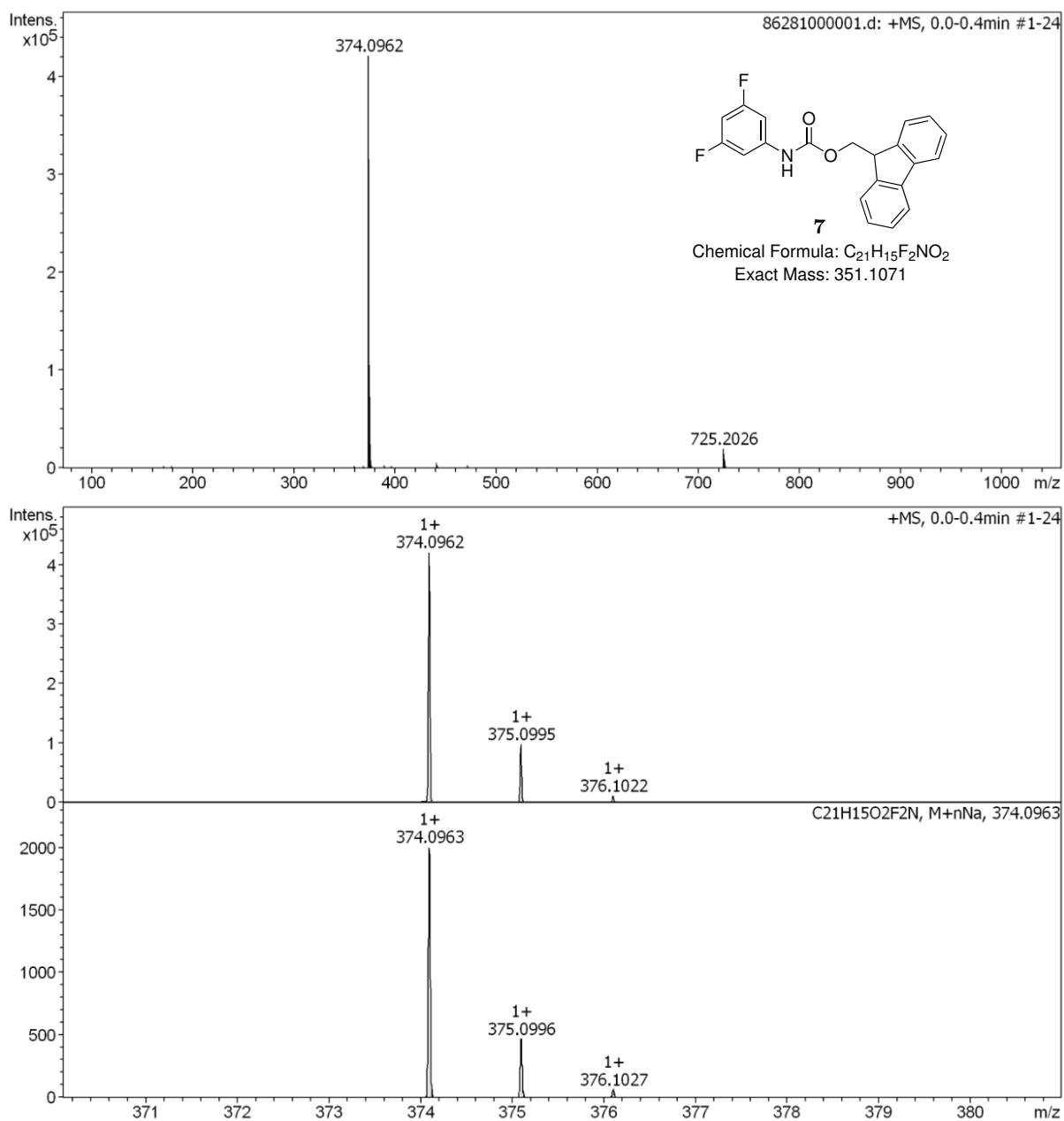
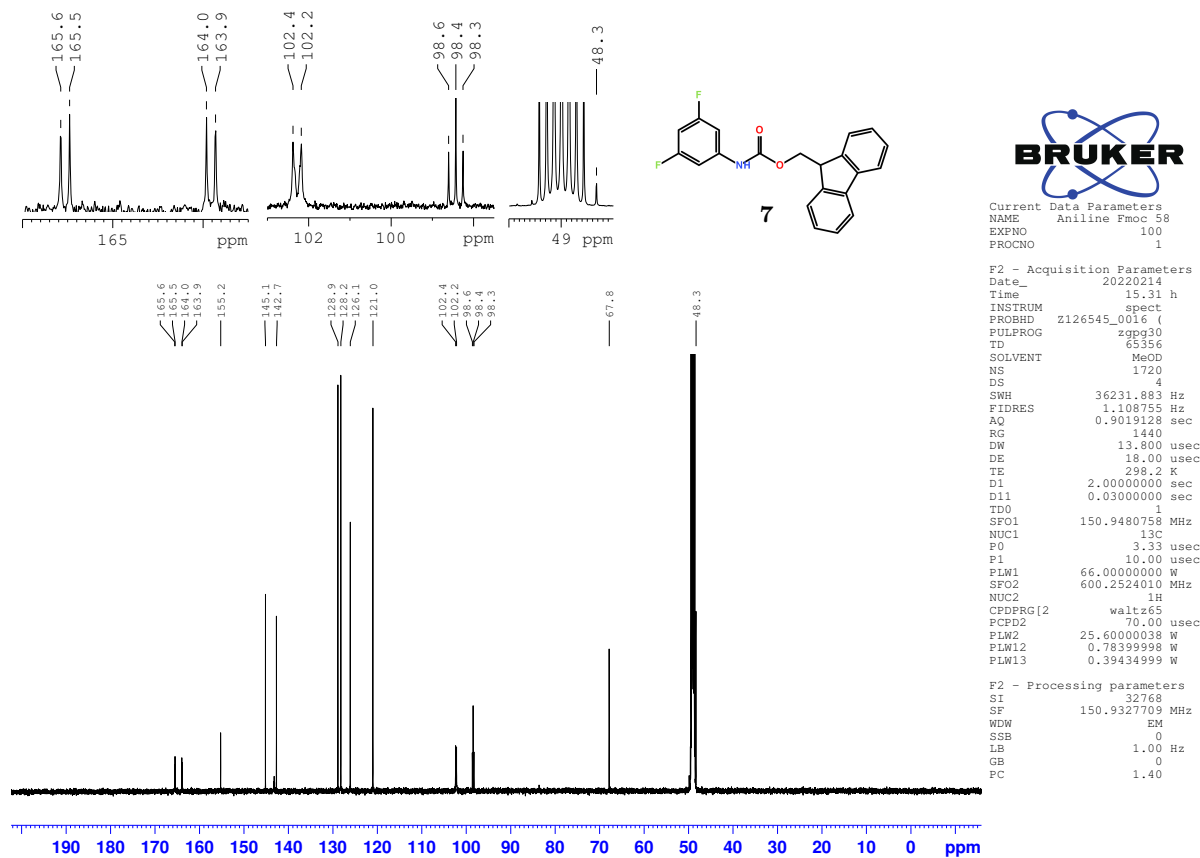
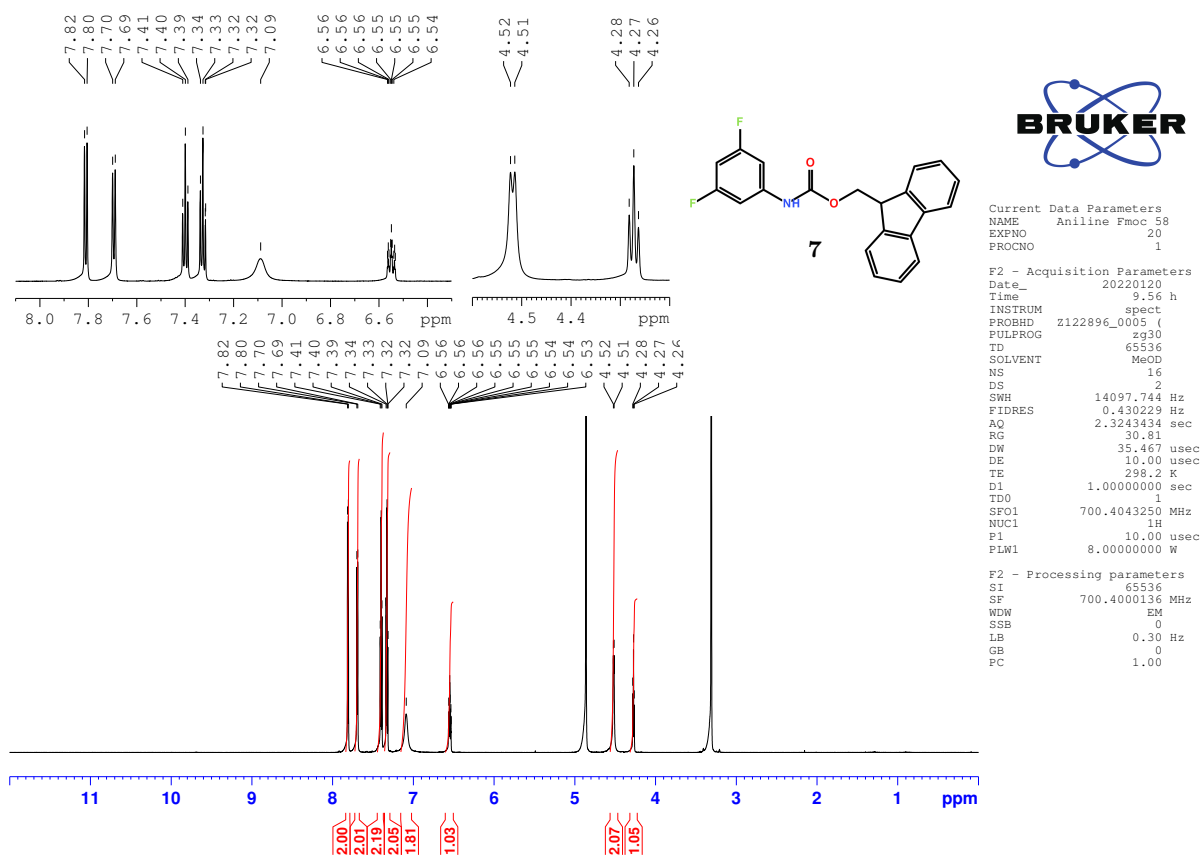


Figure 4.8: ESI-HRMS spectra of **7** ($C_{21}H_{15}O_2F_2N$) in positive mode: calc. for $[M+Na]^+$ 374.0963; found 374.0962, calc. for $[2M+Na]^+$ 725.2031; found 725.2026.



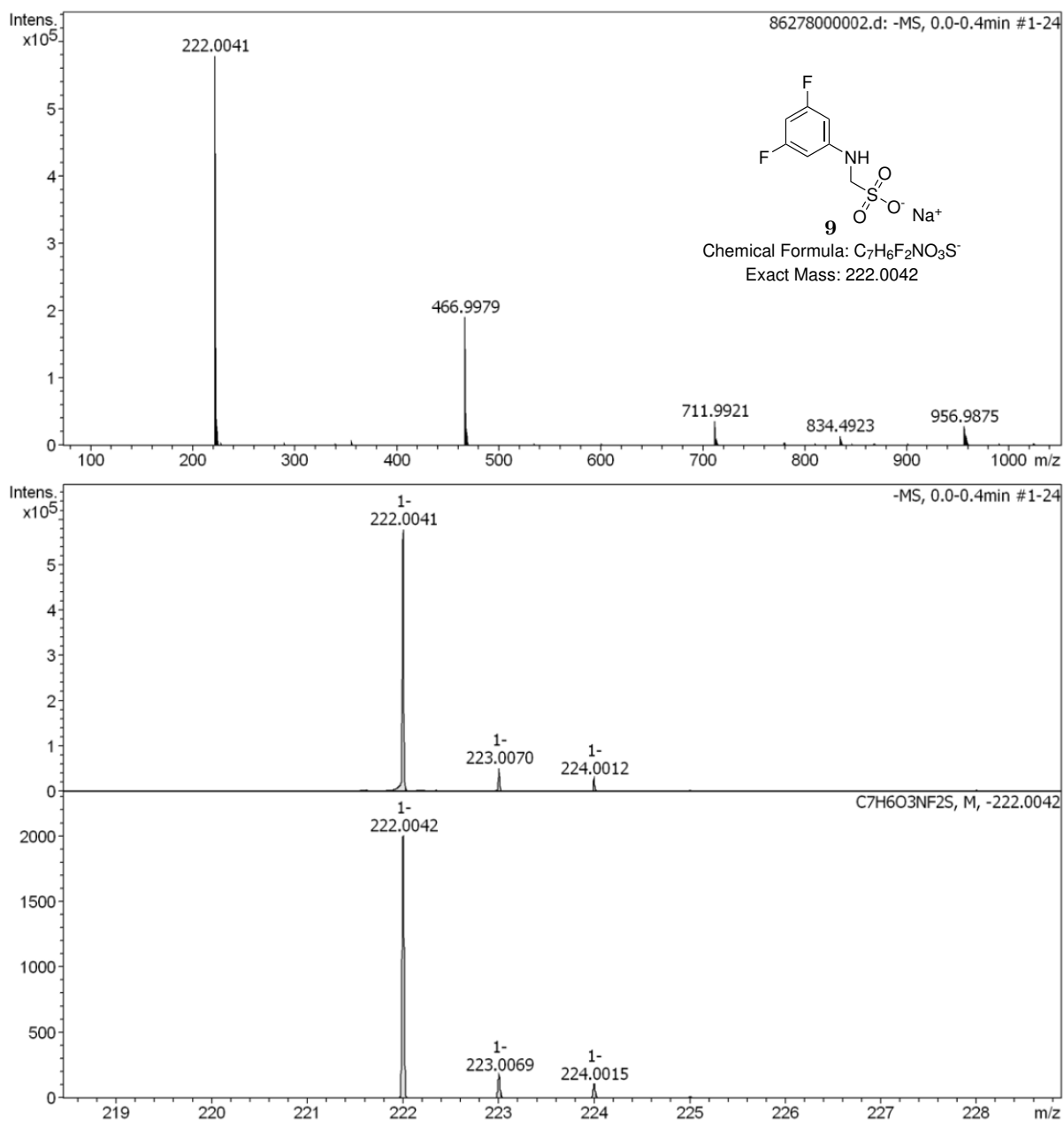
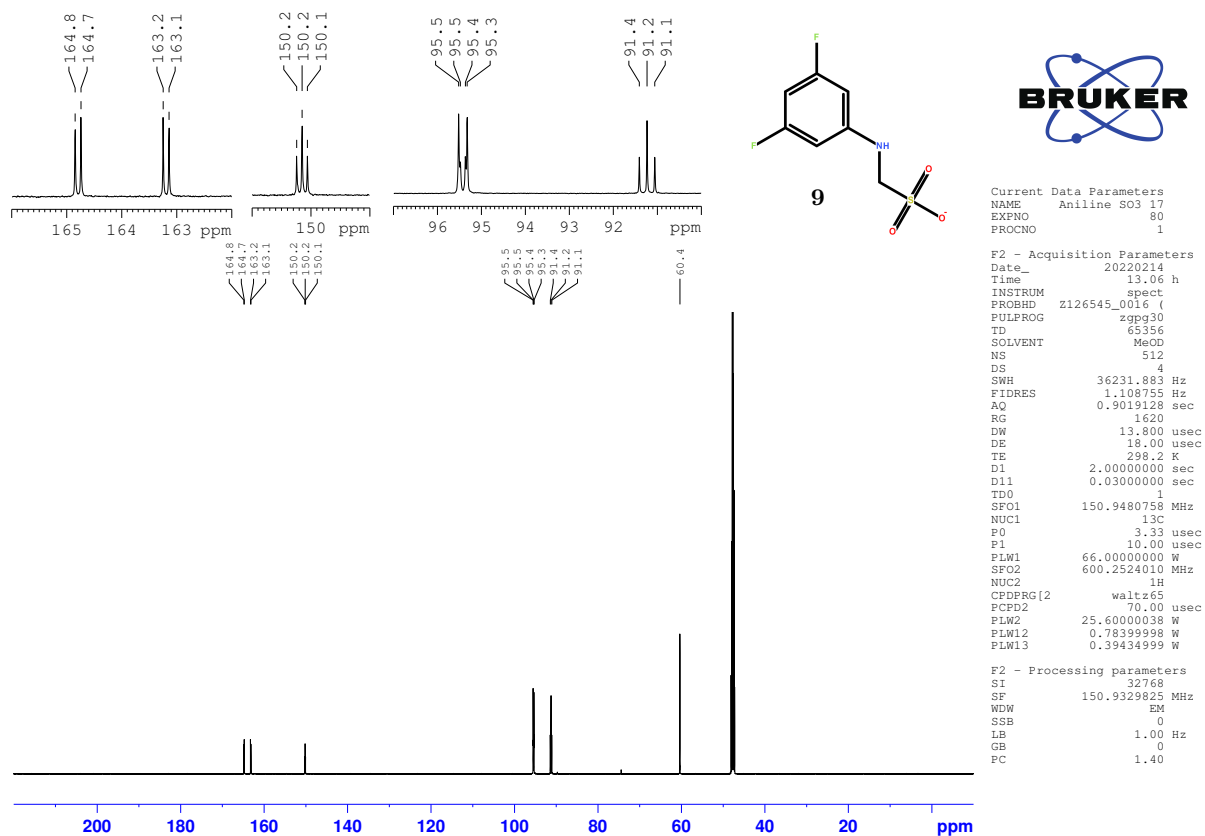
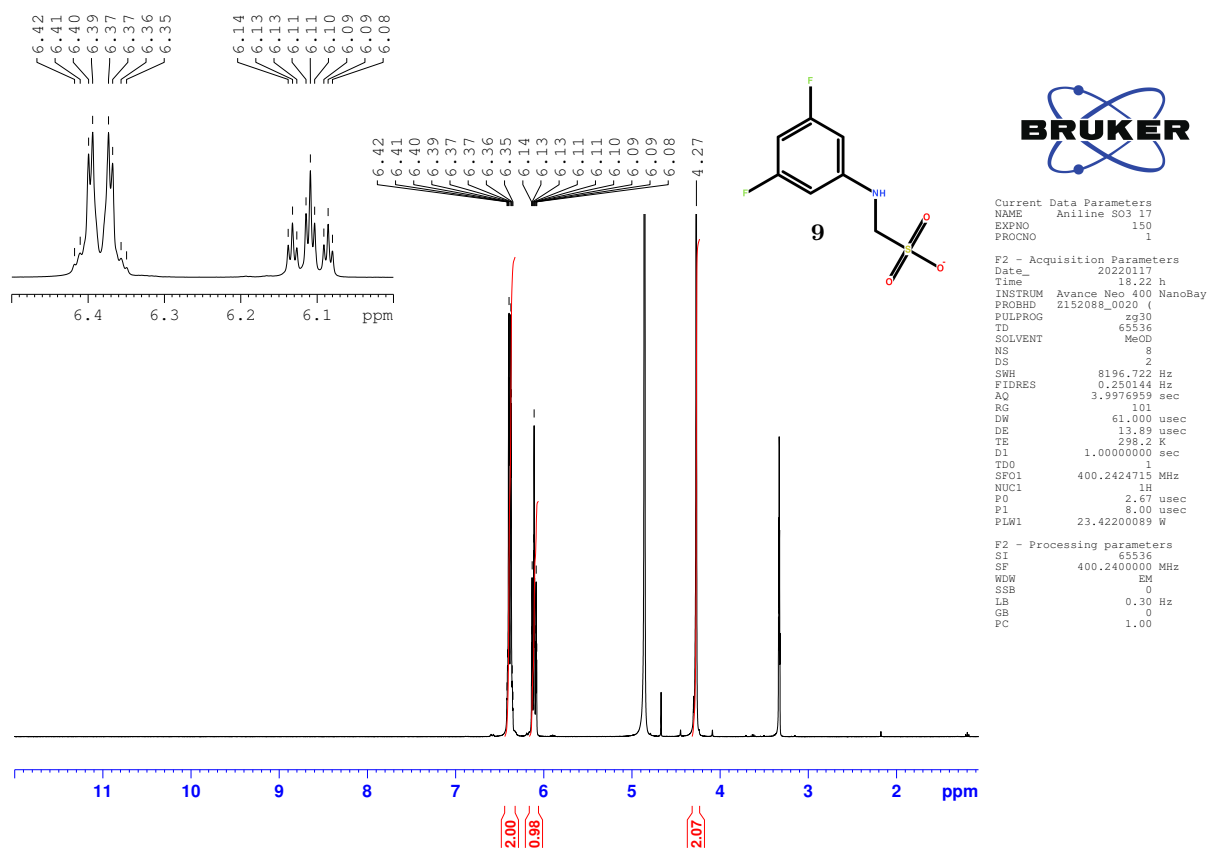


Figure 4.11: ESI-HRMS spectra of **9** ($C_7H_6O_3F_2NS$) in negative mode: calc. for $[M-H]^-$ 222.0041; found 222.0042, calc. for $[2M+Na-2H]^-$ 466.9976; found 466.9979.



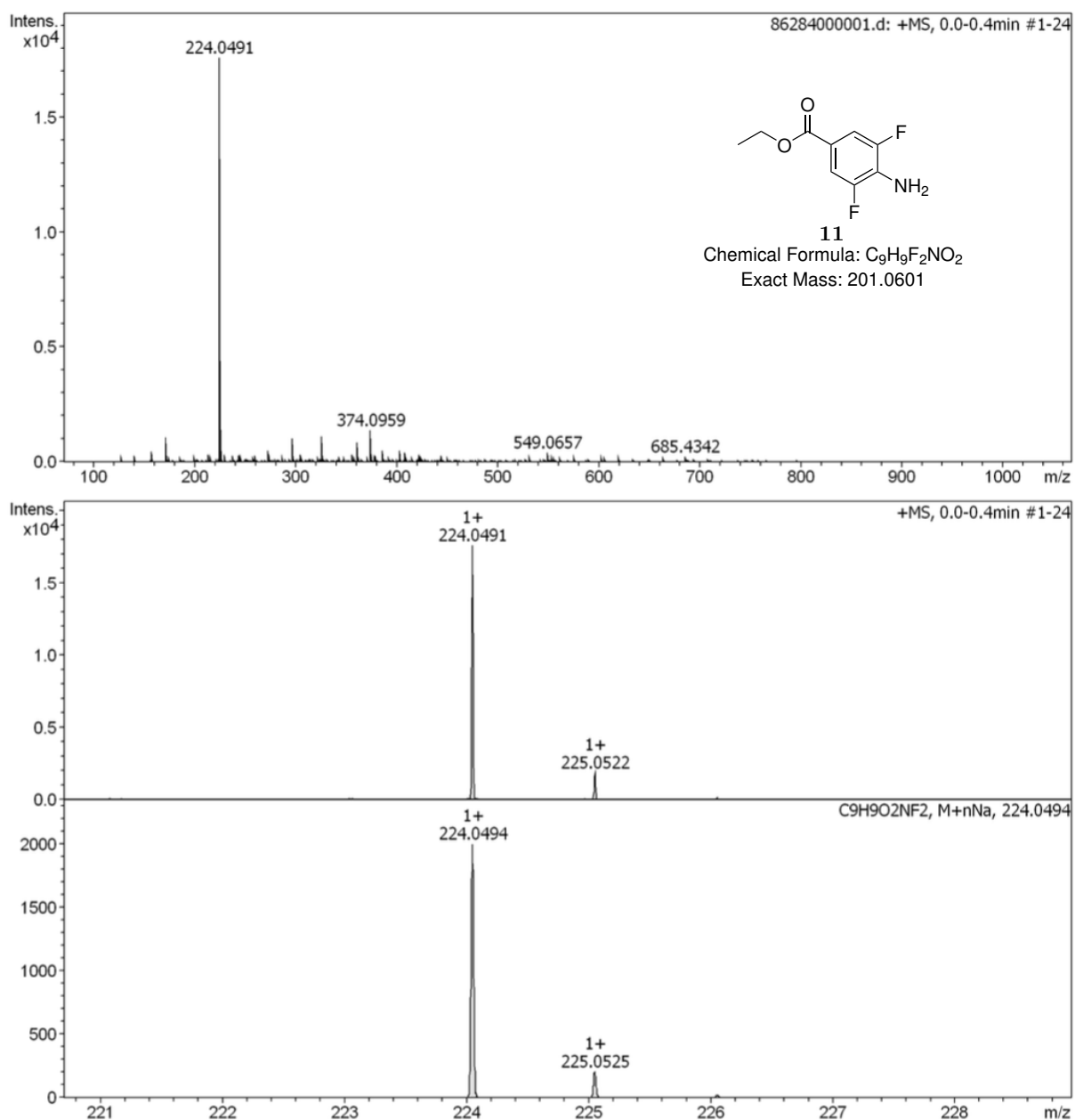
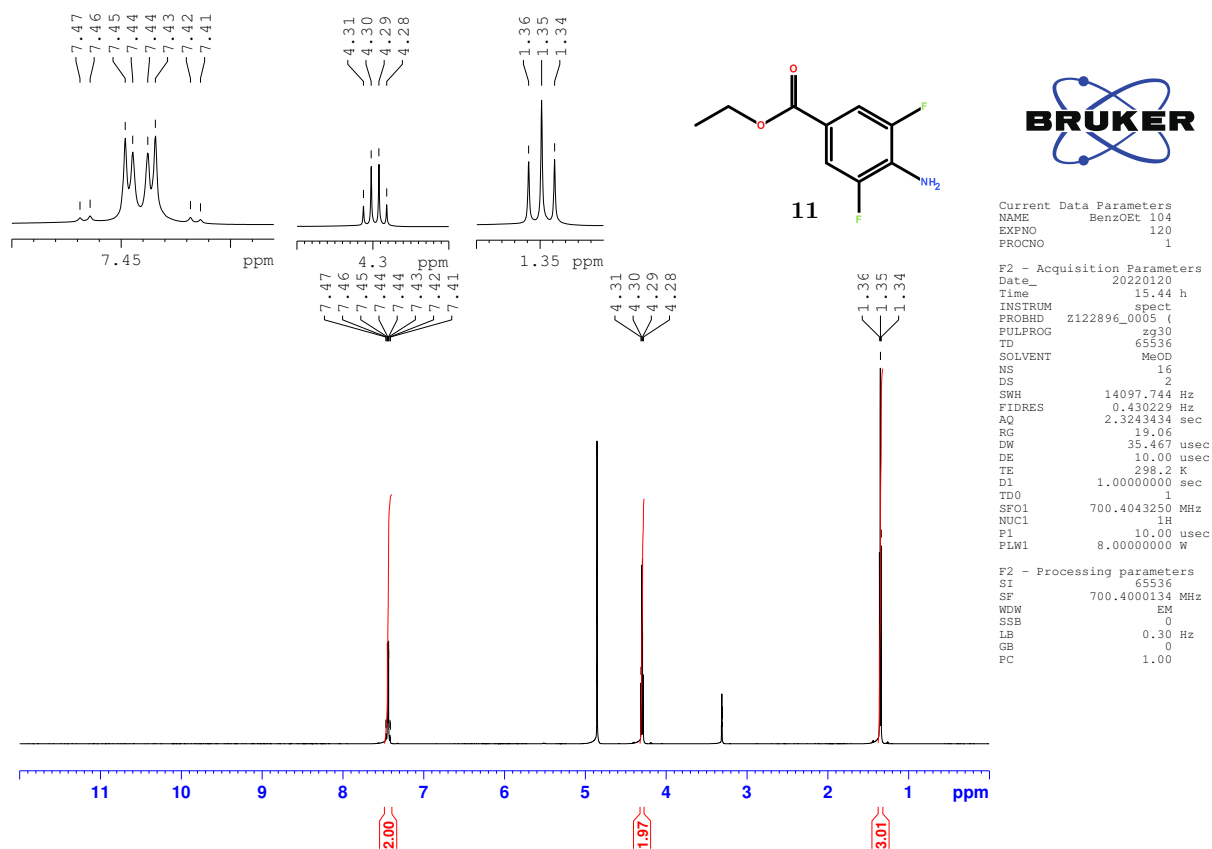
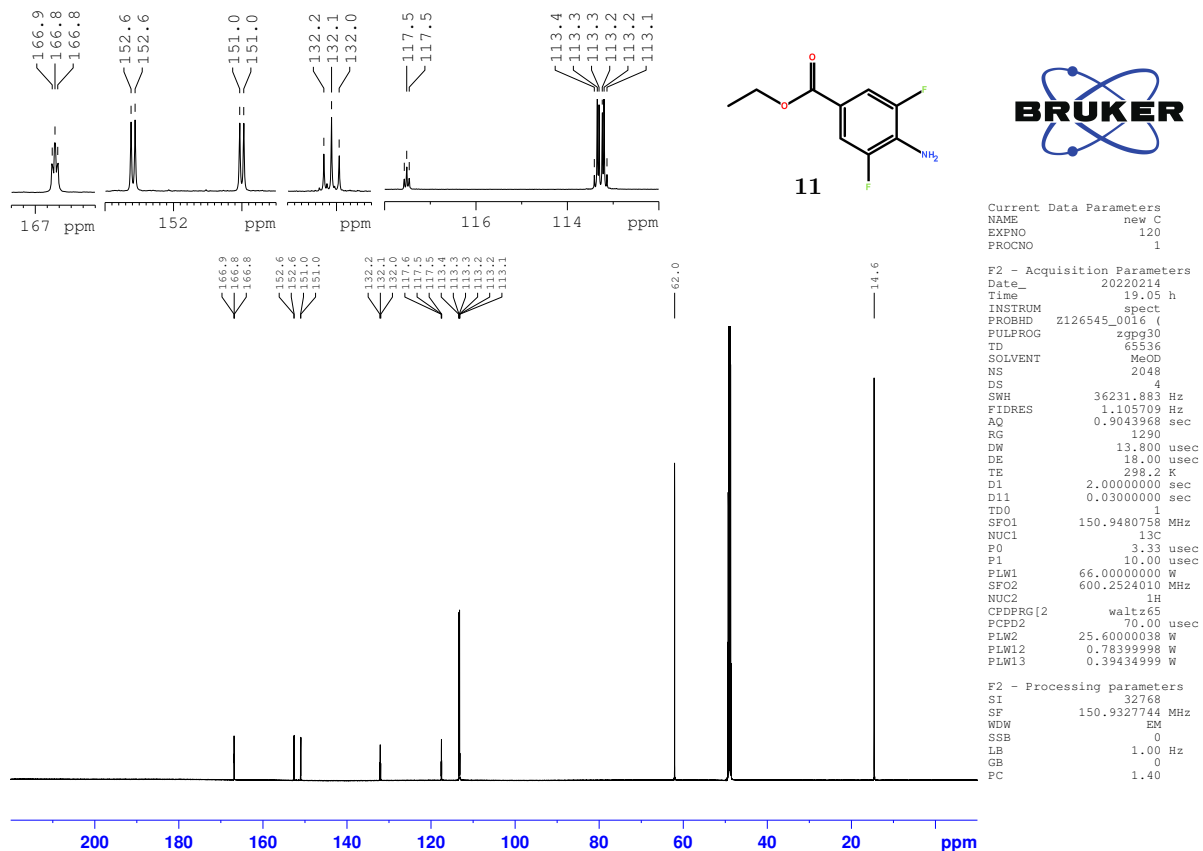


Figure 4.14: ESI-HRMS spectra of **11** ($C_9H_9O_2F_2N$) in positive mode: calc. for $[M+Na]^+$ 224.0494; found 224.0491.

Figure 4.15: ^1H NMR spectra of 11 in MeOD.Figure 4.16: ^{13}C NMR spectra of 11 in MeOD.

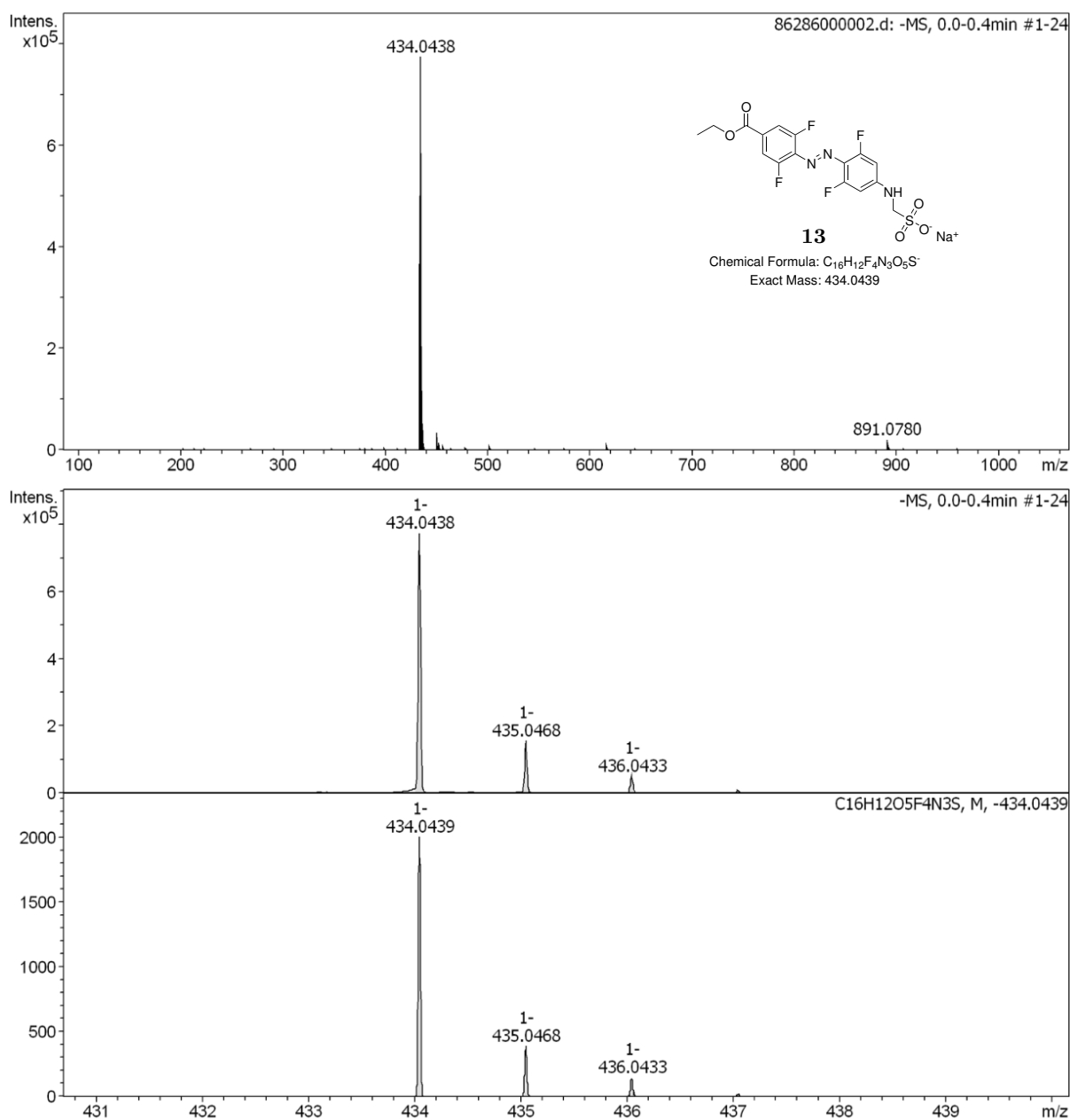
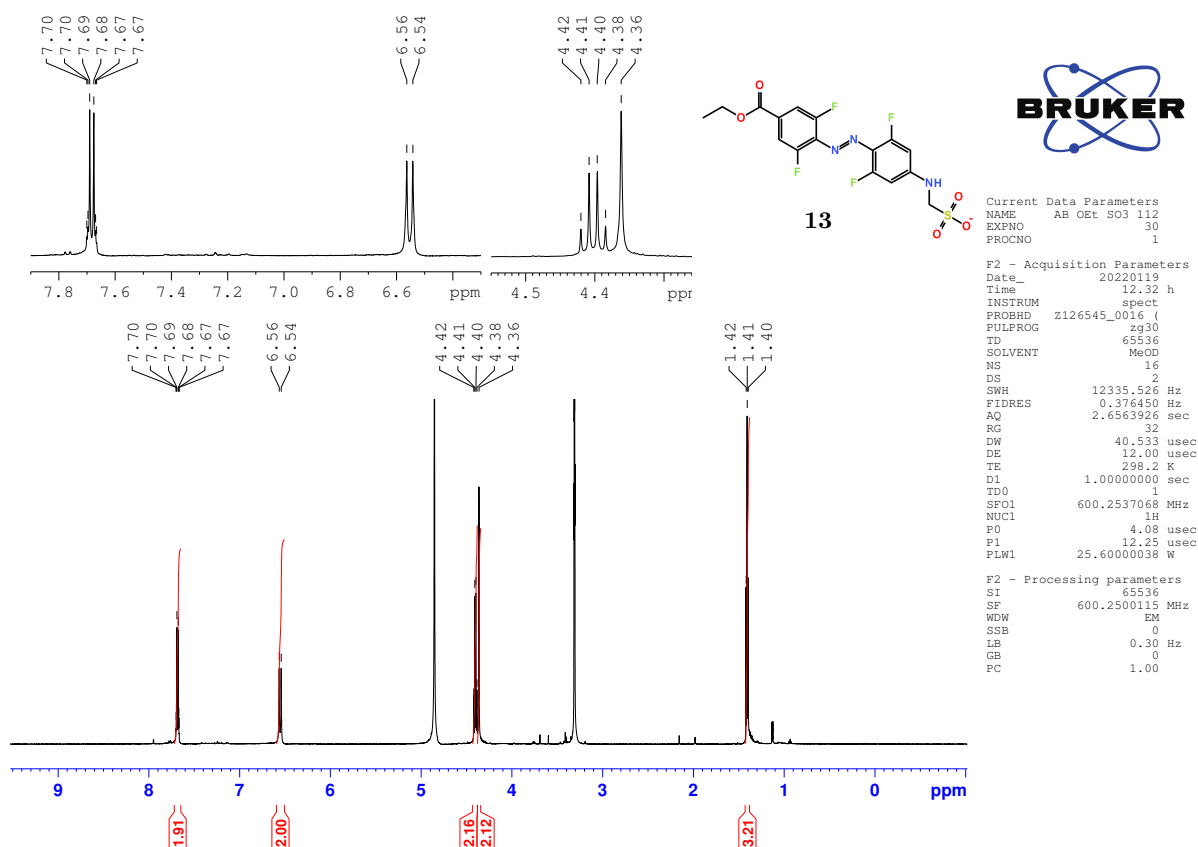
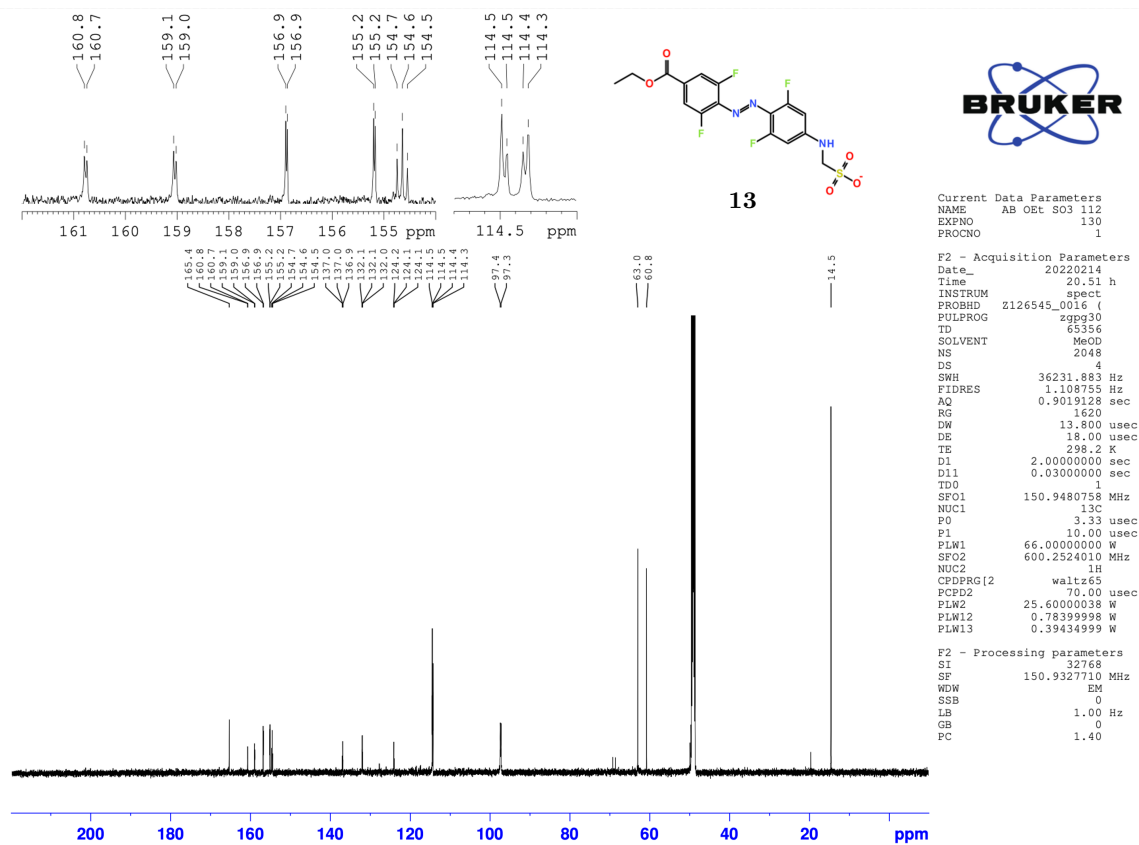


Figure 4.17: ESI-HRMS spectra of **13** ($C_{16}H_{12}O_5F_4N_3S$) in negative mode: calc. for $[M-H]^-$ 434.0439; found 434.0438.

Figure 4.18: ^1H NMR spectra of **13** in MeOD.Figure 4.19: ^{13}C NMR spectra of **13** in MeOD.

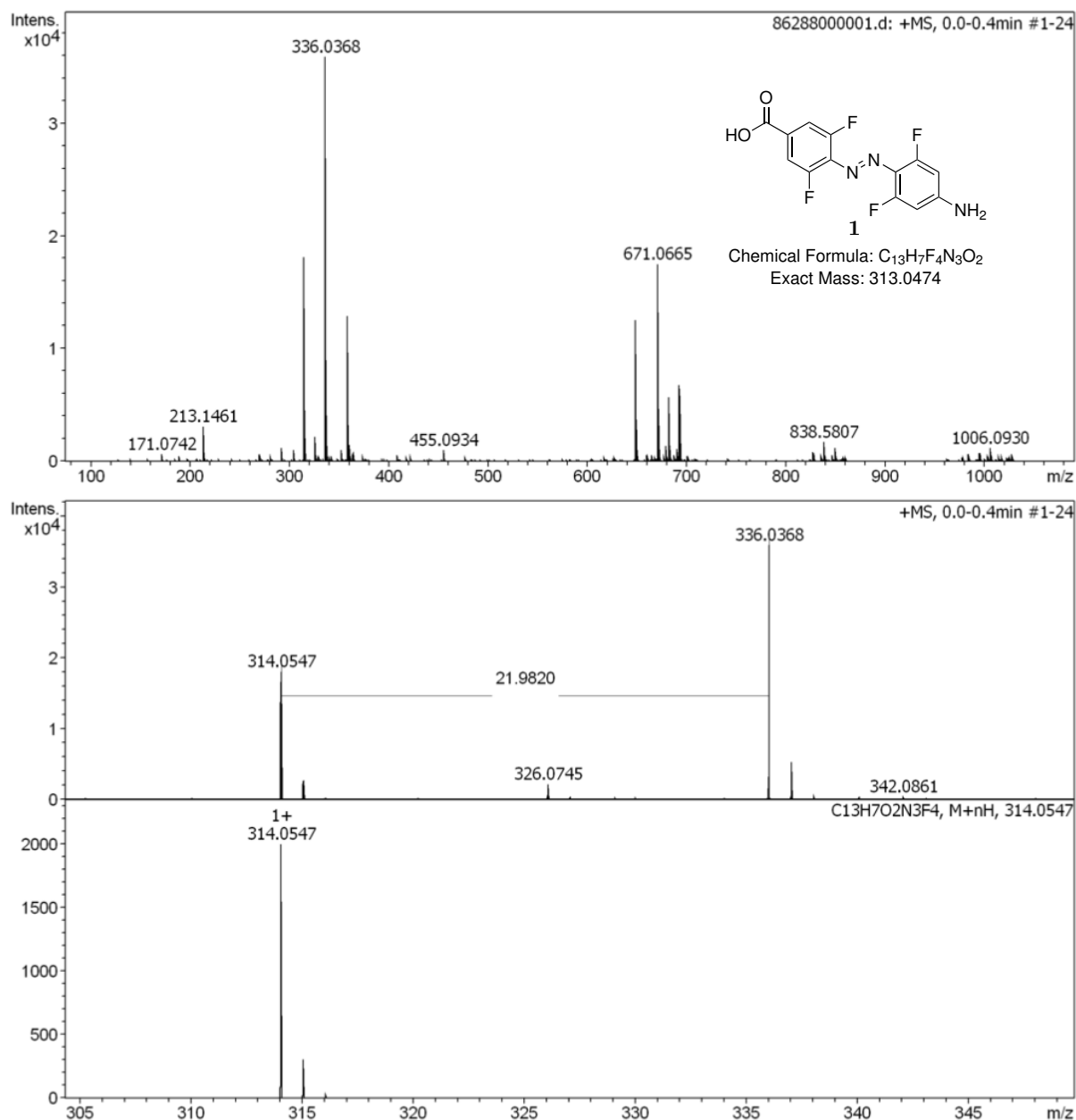


Figure 4.20: ESI-HRMS spectra of **1** ($C_{13}H_7O_2F_4N_3$) in positive mode: calc. for $[M+H]^+$ 314.0547; found 314.0547, calc. for $[M+Na+H]^{2+}$ 336.0367; found 336.0368, calc. for $[2M+2Na-H]^+$ 671.0660; found 671.0665.

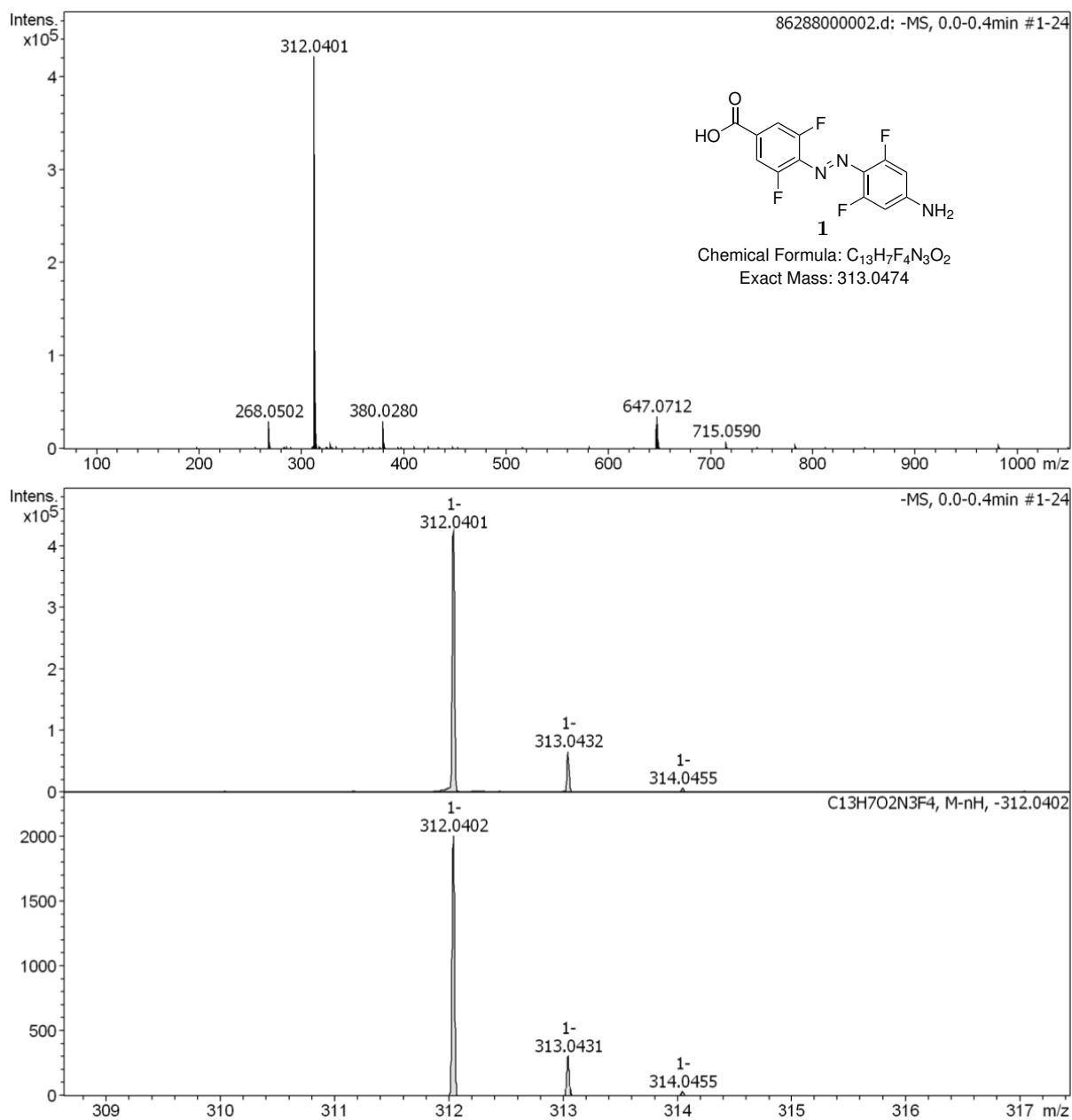
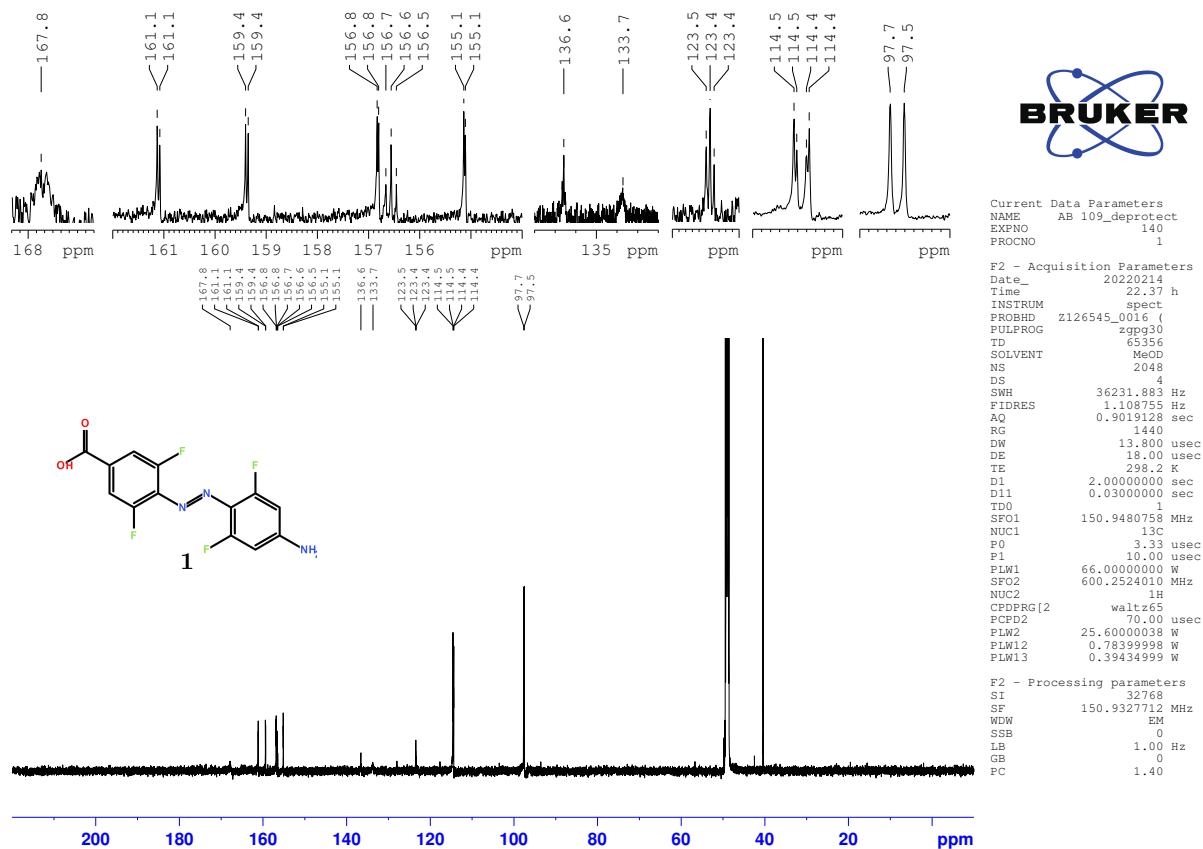
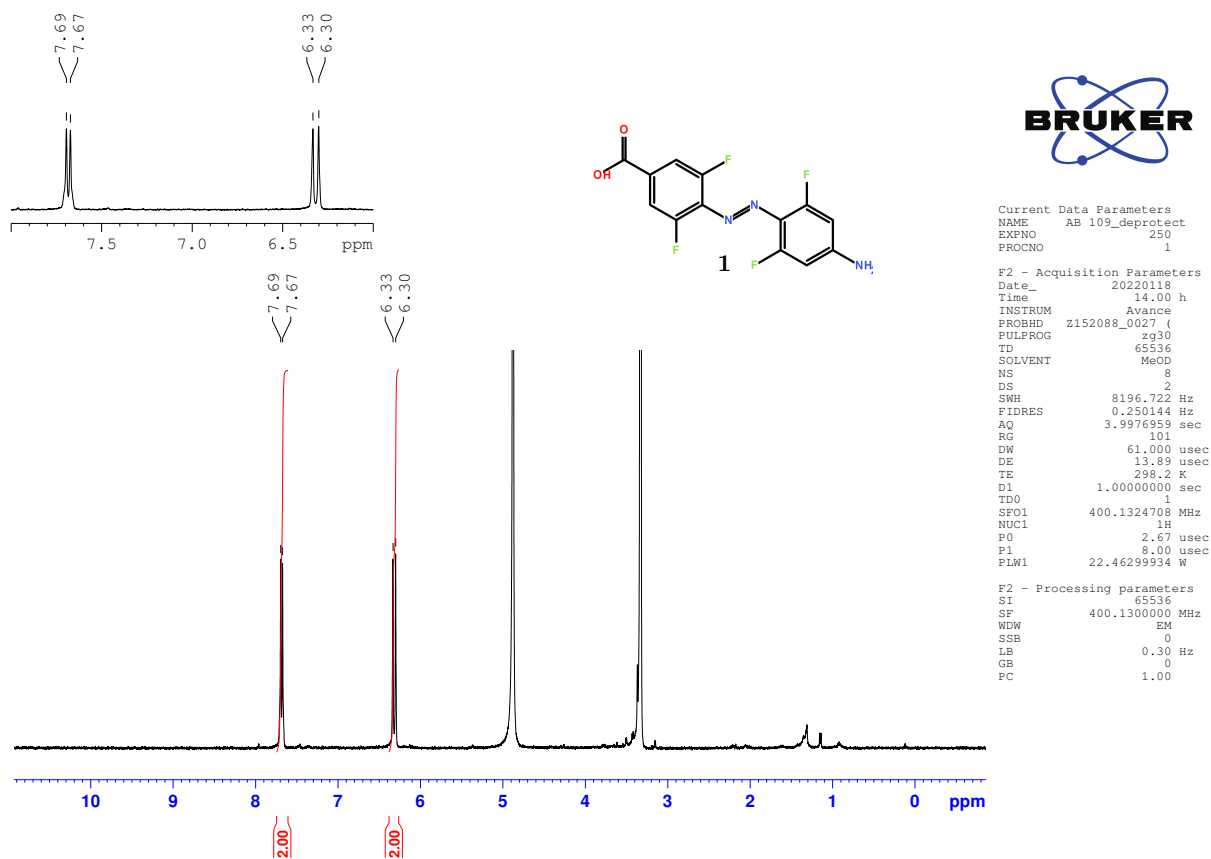


Figure 4.21: ESI-HRMS spectra of **1** ($C_{13}H_7O_2F_4N_3$) in negative mode: calc. for $[M-H]^-$ 312.0402; found 312.0401, calc. for $[2M+Na-2H]^-$ 647.0696; found 647.0712.



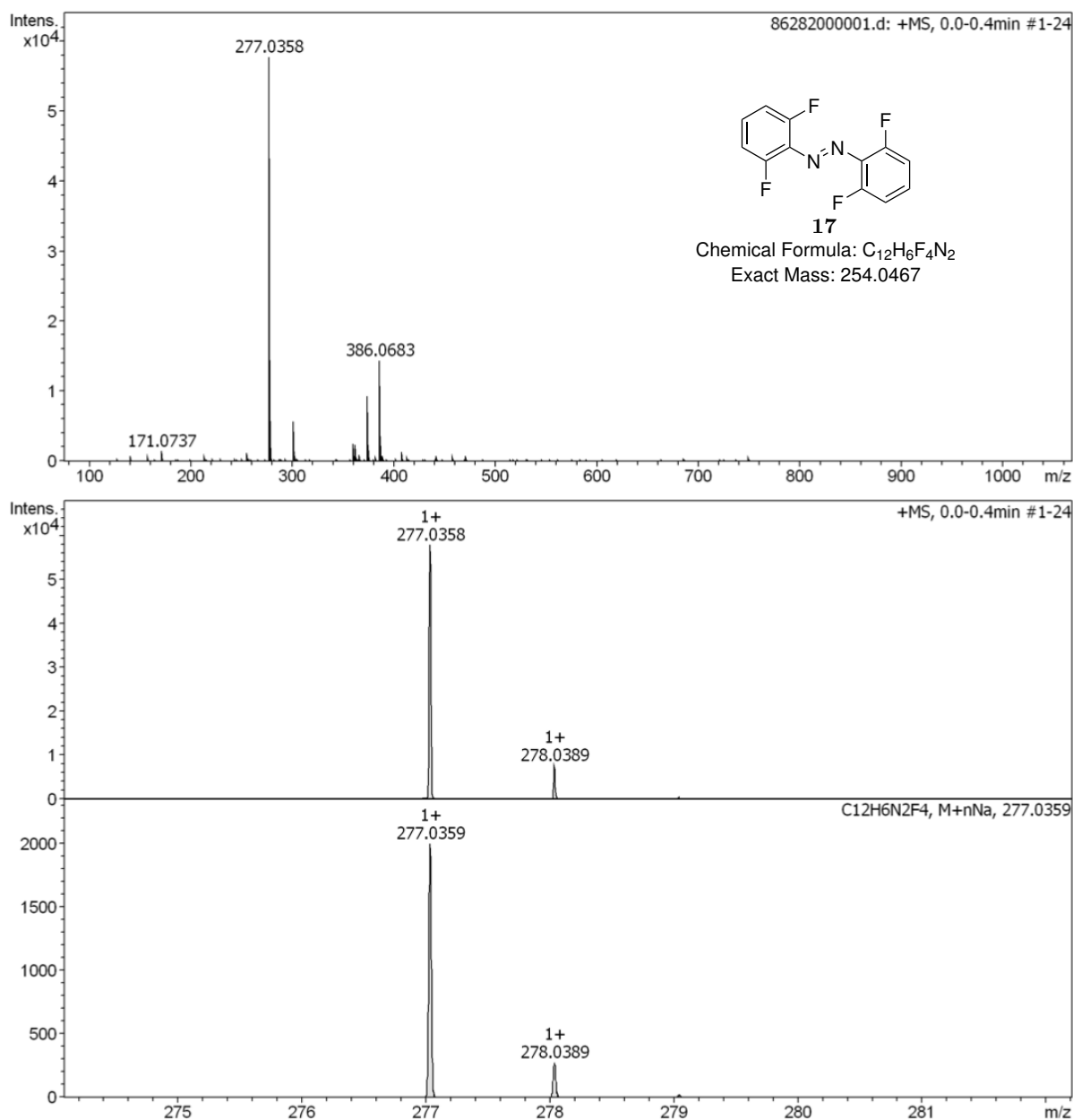


Figure 4.24: ESI-HRMS spectra of **17** ($C_{12}H_6F_4N_2S$) in positive mode: calc. for $[M+Na]^+$ 277.0359; found 277.0358.

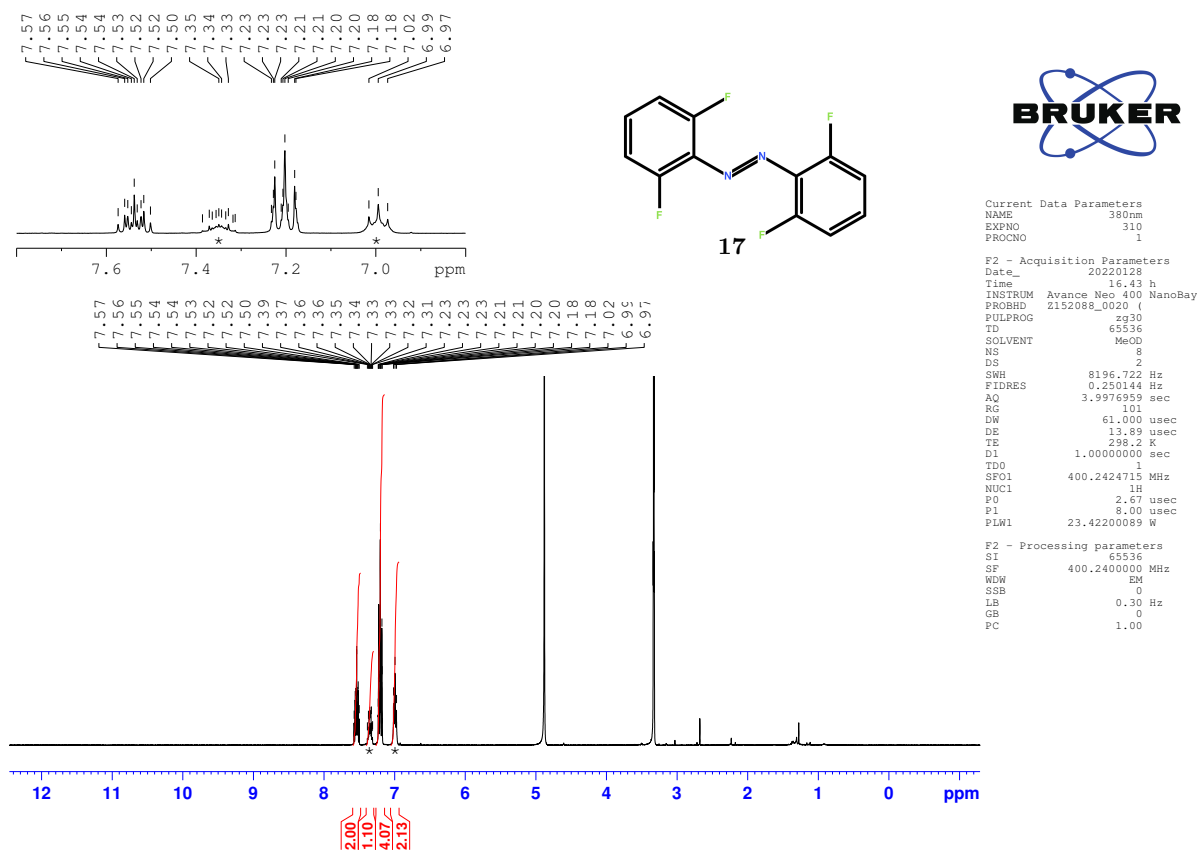


Figure 4.25: ^1H NMR spectra of **17** in MeOD. The *cis* isomer is indicated by *.

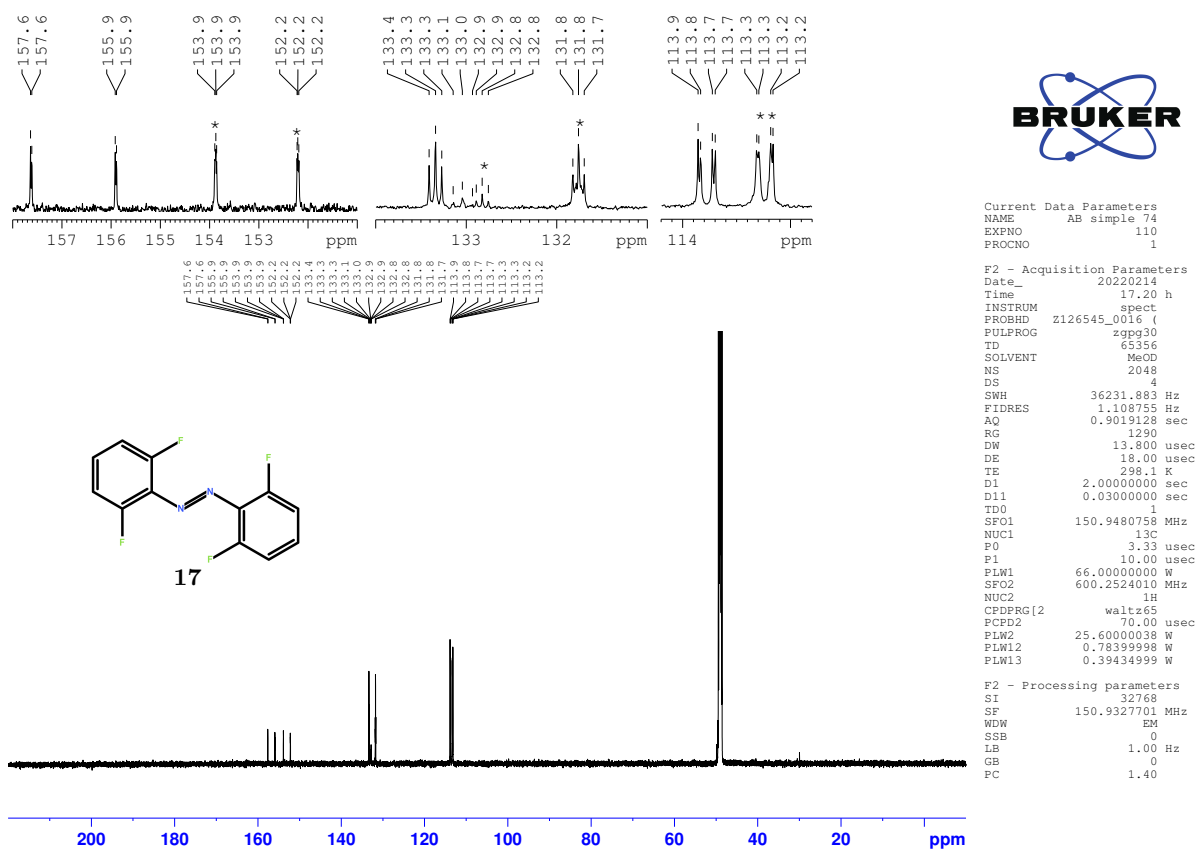


Figure 4.26: ^{13}C NMR spectra of **17** in MeOD. The *cis* isomer is indicated by *.

Bibliography

- [1] M. E. Silverman, D. Grove, C. B. Upshaw, *Circulation* **2006**, *113*, 2775–2781.
- [2] G. Roblin, P. Fleurat-Lessard, *Planta* **1987**, *170*, 242–248.
- [3] P. Rorsman, M. Braun, *Annu. Rev. Physiol.* **2013**, *75*, 155–179.
- [4] R. C. Turner, R. R. Holman, D. Matthews, T. D. Hockaday, J. Peto, *Metabolism* **1979**, *28*, 1086–1096.
- [5] P. Maechler, *Diabetes Res. Clin. Pract.* **2017**, *131*, 149–153.
- [6] H. Kang, K. Han, J. Jo, J. Kim, M. Y. Choi, *Front. Biosci.* **2008**, *13*, 6421–6431.
- [7] M. A. Atkinson, G. S. Eisenbarth, A. W. Michels, *Lancet* **2014**, *383*, 69–82.
- [8] M. Kasuga, *J. Clin. Invest.* **2006**, *116*, 1756–1760.
- [9] M. R. Nilsson, *Amyloid* **2016**, *23*, 139–147.
- [10] Ansarullah, C. Jain, F. F. Far, S. Homberg, K. Wißmiller, F. G. von Hahn, A. Raducanu, S. Schirge, M. Sterr, S. Bilekova, J. Siehler, J. Wiener, L. Oppenländer, A. Morshedi, A. Bastidas-Ponce, G. Collden, M. Irmeler, J. Beckers, A. Feuchtinger, M. Grzybek, C. Ahlbrecht, R. Feederle, O. Plettenburg, T. D. Müller, M. Meier, M. H. Tschöp, Ü. Coskun, H. Lickert, *Nature* **2021**, *590*, 326–331.
- [11] W. Hu, C. Jiang, M. Kim, Y. Xiao, H. J. Richter, D. Guan, K. Zhu, B. M. Krusen, A. N. Roberts, J. Miller, D. J. Steger, M. A. Lazar, *Genes Dev.* **2022**, *36*, 300–312.
- [12] T. Brautaset, S. Valla, *Microorganisms* **2019**, *7*, 355.
- [13] D. Mattanovich, P. Branduardi, L. Dato, B. Gasser, M. Sauer, D. Porro, *Methods Mol. Biol.* **2012**, *824*, 329–358.
- [14] M. Banghart, K. Borges, E. Isacoff, D. Trauner, R. H. Kramer, *Nat. Neurosci.* **2004**, *7*, 1381–1386.
- [15] J. Levitz, C. Pantoja, B. Gaub, H. Janovjak, A. Reiner, A. Hoagland, D. Schoppik, B. Kane, P. Stawski, A. F. Schier, D. Trauner, E. Y. Isacoff, *Nat. Neurosci.* **2013**, *16*, 507–516.
- [16] B. L. Feringa, W. R. Browne, *Molecular Switches*, Wiley-VCH, Weinheim, 2nd ed., **2011**.
- [17] W. Szymański, J. M. Beierle, H. A. Kistemaker, W. A. Velema, B. L. Feringa, *Chem. Rev.* **2013**, *113*, 6114–6178.
- [18] A. A. Beharry, G. A. Woolley, *Chem. Soc. Rev.* **2011**, *40*, 4422–4437.
- [19] J. Greeves, C. Nick, S. Warren, *Organic Chemistry*, Oxford University Press, New York, 2nd ed., **2012**.
- [20] D. J. Van Dijken, P. Kovaříček, S. P. Ihrig, S. Hecht, *J. Am. Chem. Soc.* **2015**, *137*, 14982–14991.
- [21] A. Aemissegger, V. Kräutler, W. F. Van Gunsteren, D. Hilvert, *J. Am. Chem. Soc.* **2005**, *127*, 2929–2936.
- [22] S. M. Hofmann, C. V. Frost, T. Podewin, M. Gailer, E. Weber, M. Zacharias, W. Zinth, A. Hoffmann-Röder, *J. Phys. Chem. B* **2020**, *124*, 5113–5121.
- [23] T. Schober, I. Wehl, S. Afonin, O. Babii, A. Iampolska, U. Schepers, I. V. Komarov, A. S. Ulrich, *ChemPhotoChem* **2019**, *3*, 384–391.
- [24] K. Nakayama, M. Endo, T. Majima, *Chem. Commun.* **2004**, 2386–2387.
- [25] T. Sakata, Y. Yan, G. Marriott, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 4759–4764.
- [26] W. Kabsch, H. G. Mannherz, D. Suck, E. F. Pai, K. C. Holmes, *Nature* **1990**, *347*, 37–44.
- [27] S. Samanta, A. A. Beharry, O. Sadowski, T. M. McCormick, A. Babalhavaeji, V. Tropepe, G. A. Woolley, *J. Am. Chem. Soc.* **2013**, *135*, 9777–9784.

- [28] D. G. Flint, J. R. Kumita, O. S. Smart, G. A. Woolley, *Chem. Biol.* **2002**, *9*, 391–397.
- [29] D. Bleger, J. Schwarz, A. M. Brouwer, S. Hecht, *J. Am. Chem. Soc.* **2012**, *134*, 20597–20600.
- [30] M. Dong, A. Babalhavaeji, S. Samanta, A. A. Beharry, G. A. Woolley, *Acc. Chem. Res.* **2015**, *48*, 2662–2670.
- [31] S. F. Chin, P. I. H. Megat Mohd Azlan, L. Mazlan, H. M. Neoh, *Gut Pathog.* **2018**, *10*, 29.
- [32] S. L. Forsburg, N. Rhind, *Yeast* **2006**, *23*, 173–183.
- [33] R. J. Bennett, B. G. Turgeon, *Microbiol. Spectr.* **2016**, *4*, 1–28.
- [34] J. Davey, *Yeast* **1998**, *14*, 1529–1566.
- [35] S. Sasuga, R. Abe, O. Nikaido, S. Kiyosaki, H. Sekiguchi, A. Ikai, T. Osada, *J. Biomed. Biotechnol.* **2012**, 1–8.
- [36] Y. Imai, M. Yamamoto, *Genes Dev.* **1994**, *8*, 328–338.
- [37] S. Basith, M. Cui, S. J. Macalino, J. Park, N. A. Clavio, S. Kang, S. Choi, *Front. Pharmacol.* **2018**, *9*, 128.
- [38] A. Ciancetta, D. Sabbadin, S. Federico, G. Spalluto, S. Moro, *Trends Pharmacol. Sci.* **2015**, *36*, 878–890.
- [39] F. Naider, J. M. Becker, *Peptides* **2004**, *25*, 1441–1463.
- [40] K. Kitamura, C. Shimoda, *EMBO J.* **1991**, *10*, 3743–3751.
- [41] G. Fischer, *Chem. Soc. Rev.* **2000**, *29*, 119–127.
- [42] L. Vitagliano, R. Berisio, A. Mastrangelo, L. Mazzarella, A. Zagari, *Protein Sci.* **2009**, *10*, 2627–2632.
- [43] G. D. Rose, L. M. Glerasch, J. A. Smith, *Adv. Protein Chem.* **1985**, *37*, 1–109.
- [44] J. W. Schymkowitz, F. Rousseau, L. S. Itzhaki, *J. Mol. Biol.* **2000**, *301*, 199–204.
- [45] L. A. Jelicks, F. R. Naider, P. Shenbagamurthi, J. M. Becker, M. S. Broido, *Biopolymers* **1988**, *27*, 431–449.
- [46] L. A. Jelicks, M. S. Broido, J. M. Becker, F. R. Naider, *Biochemistry* **1989**, *28*, 4233–4240.
- [47] J. S. Gounarides, M. S. Broido, J. M. Becker, F. R. Naider, *Biochemistry* **1993**, *32*, 908–917.
- [48] C.-B. Xue, A. McKinney, H.-F. Lu, Y. Jiang, J. M. Becker, F. Naider, *Int. J. Pept. Protein Res.* **2009**, *47*, 131–141.
- [49] V. Velazhahan, N. Ma, G. Pándy-Szekeres, A. J. Kooistra, Y. Lee, D. E. Gloriam, N. Vaidehi, C. G. Tate, *Nature* **2021**, *589*, 148–153.
- [50] C. T. Armstrong, P. E. Mason, J. L. Anderson, C. E. Dempsey, *Sci. Rep.* **2016**, *6*, 1–10.
- [51] B. K. Lee, S. Khare, F. Naider, J. M. Becker, *J. Biol. Chem.* **2001**, *276*, 37950–37961.
- [52] Y. L. Zhang, H. R. Marepalli, H. F. Lu, J. M. Becker, F. Naider, *Biochemistry* **1998**, *37*, 12465–12476.
- [53] K. Wakamatsu, A. Okada, T. Miyazawa, Y. Masui, S. Sakakibara, T. Higashijima, *Eur. J. Biochem.* **1987**, *163*, 331–338.
- [54] M. G. Abel, Y. L. Zhang, H. F. Lu, F. Naider, J. M. Becker, *J. Pept. Res.* **1998**, *52*, 95–106.
- [55] G. Ladds, K. Davis, A. Das, J. Davey, *Mol. Microbiol.* **2005**, *55*, 482–497.
- [56] G. Ladds, J. Davey, *Mol. Microbiol.* **2000**, *36*, 377–390.
- [57] G. Ladds, E. M. Rasmussen, T. Young, O. Nielsen, J. Davey, *Mol. Microbiol.* **1996**, *20*, 35–42.
- [58] Y. Giga-Hama, H. Kumagai, *Biotechnol. Appl. Biochem.* **1999**, *30*, 235–244.

- [59] S. Hidaka, O. Nikaido, S. Kiyosaki, A. Ikai, T. Osada, *J. Surf. Eng. Mater. Adv. Technol.* **2013**, *3*, 36–42.
- [60] A. Duttaroy, P. Kanakaraj, B. L. Osborn, H. Schneider, O. K. Pickeral, C. Chen, G. Zhang, S. Kaithamana, M. Singh, R. Schulingkamp, D. Crossan, J. Bock, T. E. Kaufman, P. Keavey, M. Carey-Barber, S. K. Krishnan, A. Garcia, K. Murphy, J. K. Siskind, M. A. McLean, S. Cheng, S. Ruben, C. E. Birse, O. Blondel, *Diabetes* **2005**, *54*, 251–258.
- [61] R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.
- [62] B. Merrifield, *Science* **1986**, *232*, 341–347.
- [63] M. Verlander, *Int. J. Pept. Res. Ther.* **2007**, *13*, 75–82.
- [64] G. B. Fields, *Curr. Protoc. Mol. Biol.* **2003**, *63*, 11.15.1–11.15.9.
- [65] L. A. Carpino, G. Y. Han, *J. Am. Chem. Soc.* **1970**, *92*, 5748–5749.
- [66] L. K. Mueller, A. C. Baumruck, H. Zhdanova, A. A. Tietze, *Front. Bioeng. Biotechnol.* **2020**, *8*, 1–17.
- [67] J. A. Moss, *Curr. Protoc. Protein Sci.* **2005**, *40*, 18.7.1–18.7.19.
- [68] E. I. Vrettos, N. Sayyad, E. M. Mavrogiannaki, E. Stylos, A. D. Kostagianni, S. Papas, T. Mavromoustakos, V. Theodorou, A. G. Tzakos, *RSC Adv.* **2017**, *7*, 50519–50526.
- [69] M. F. Pignataro, M. G. Herrera, V. I. Doderio, *Molecules* **2020**, *25*, 1–35.
- [70] B. M. Bulheller, A. Rodger, J. D. Hirst, *Phys. Chem. Chem. Phys.* **2007**, *9*, 2020–2035.
- [71] N. J. Greenfield, *Methods Enzymol.* **2004**, *383*, 282–317.
- [72] S. Brahms, J. Brahms, *J. Mol. Biol.* **1980**, *138*, 149–178.
- [73] R. W. Woody, *J. Am. Chem. Soc.* **2009**, *131*, 8234–8245.
- [74] D. D. Jenness, C. Sprecher, W. C. Johnson, *Biopolymers* **1976**, *15*, 513–521.
- [75] P. Khandelwal, S. Seth, R. V. Hosur, *Eur. J. Biochem.* **1999**, *264*, 468–478.
- [76] J. Kemmink, T. E. Creighton, *Biochemistry* **1995**, *34*, 12630–12635.
- [77] H. Reiersen, A. R. Rees, *Protein Eng. Des. Sel.* **2000**, *13*, 739–743.
- [78] A. Micsonai, F. Wien, L. Kernya, Y. H. Lee, Y. Goto, M. Réfrégiers, J. Kardos, *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, E3095–E3103.
- [79] H. L. Bank, *Vitr. Cell. Dev. Biol.* **1988**, *24*, 266–273.
- [80] L. L. Y. Chan, D. J. Laverty, T. Smith, P. Nejad, H. Hei, R. Gandhi, D. Kuksin, J. Qiu, *J. Immunol. Methods* **2013**, *388*, 25–32.
- [81] G. S. Hartley, *Nature* **1937**, *140*, 281–281.
- [82] J. Griffiths, *Chem. Soc. Rev.* **1972**, *1*, 481–493.
- [83] J. Dokić, M. Gothe, J. Wirth, M. V. Peters, J. Schwarz, S. Hecht, P. Saalfrank, *J. Phys. Chem. A* **2009**, *113*, 6763–6773.
- [84] A. R. Dias, M. E. Minas Da Piedade, J. A. Martinho Simões, J. A. Simoni, C. Teixeira, H. P. Diogo, Y. Meng-Yan, G. Pilcher, *J. Chem. Thermodyn.* **1992**, *24*, 439–447.
- [85] H. M. Bandara, S. C. Burdette, *Chem. Soc. Rev.* **2012**, *41*, 1809–1825.
- [86] E. M. Tan, S. Amirjalayer, S. Smolarek, A. Vdovin, F. Zerbetto, W. J. Buma, *Nat. Commun.* **2015**, *6*, 1–7.
- [87] J. Garcia-Amorós, A. Sánchez-Ferrer, W. A. Massad, S. Nonell, D. Velasco, *Phys. Chem. Chem. Phys.* **2010**, *12*, 13238–13242.
- [88] M. Baroncini, J. Groppi, S. Corra, S. Silvi, A. Credi, *Adv. Opt. Mater.* **2019**, *7*, 1900392.
- [89] S. Soltani, A. Ojaghi, F. E. Robles, *Biomed. Opt. Express* **2019**, *10*, 487.

- [90] F. A. Jerca, V. V. Jerca, R. Hoogenboom, *Nat. Rev. Chem.* **2021**, *6*, 51–69.
- [91] R. J. Mart, R. K. Allemann, *Chem. Commun.* **2016**, *52*, 12262–12277.
- [92] A. A. Beharry, O. Sadovski, G. A. Woolley, *J. Am. Chem. Soc.* **2011**, *133*, 19684–19687.
- [93] E. Merino, *Chem. Soc. Rev.* **2011**, *40*, 3835–3853.
- [94] H. Mayr, M. Breugst, A. R. Ofial, *Angew. Chemie - Int. Ed.* **2011**, *50*, 6470–6505.
- [95] A. Lalitha, K. Pitchumani, C. Srinivasan, *J. Mol. Catal. A Chem.* **2000**, *160*, 429–435.
- [96] A. F. Hegarty in *Chem. Diazonium Diazo Groups, Vol. 2* (Ed.: S. Patai), John Wiley & Sons, Ltd, Cork, **2010**, Chapter 12, pp. 511–591.
- [97] A. Corma, A. Martinez, *Adv. Mater.* **1995**, *7*, 137–144.
- [98] A. Prestel, H. M. Möller, *Chem. Commun.* **2016**, *52*, 701–704.
- [99] M. Sainsbury, *Aromatic Chemistry*, Oxford University Press, New York, 4th ed., **1992**.
- [100] E. Léonard, F. Mangin, C. Villette, M. Billamboz, C. Len, *Catal. Sci. Technol.* **2016**, *6*, 379–398.
- [101] Y. Ogata, Y. Nakagawa, M. Inaishi, *Bull. Chem. Soc. Jpn.* **1980**, 2853–2854.
- [102] J. Zhao, X. Yang, H. Yang, S. Li, J. Li, B. Xing, Shanxi Datong University, CN113173865A, **2021**.
- [103] M. B. Gawande, P. S. Branco, *Green Chem.* **2011**, *13*, 3355–3359.
- [104] N. Ishikawa, M. J. Namkung, T. L. Fletcher, *J. Org. Chem.* **1965**, *30*, 3878–3882.
- [105] F. Mirazizi, A. Bahrami, K. Haghbeen, H. Shahbani Zahiri, M. Bakavoli, R. L. Legge, *J. Enzym. Inhib. Med. Chem.* **2016**, *31*, 1162–1169.
- [106] J. Park, J. Koh, *Dye. Pigment.* **2009**, *82*, 347–352.
- [107] S. Biradar, R. Kasugai, H. Kanoh, H. Nagao, Y. Kubota, K. Funabiki, M. Shiro, M. Matsui, *Dye. Pigment.* **2016**, *125*, 249–258.
- [108] H.-J. Chen, J. M. Barendt, R. C. Haltiwanger, T. G. Hill, A. D. Norman, *Phosphorous Sulfur Relat. Elem.* **1981**, *11*, 295–301.
- [109] J. Auernheimer, C. Dahmen, U. Hersel, A. Bausch, H. Kessler, *J. Am. Chem. Soc.* **2005**, *127*, 16107–16110.
- [110] L. Zhang, C. R. Robertson, B. R. Green, T. H. Pruess, H. S. White, G. Bulaj, *J. Med. Chem.* **2009**, *52*, 1310–1316.
- [111] S. A. Palasek, Z. J. Cox, J. M. Collins, *J. Pept. Sci.* **2007**, *13*, 143–148.
- [112] W. Mäntele, E. Deniz, *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **2017**, *173*, 965–968.
- [113] R. Tabone, M. Barra, *Dye. Pigment.* **2011**, *88*, 180–186.
- [114] M. Barra, N. Chen, *J. Org. Chem.* **2000**, *65*, 5739–5744.
- [115] J. Garcia-Amorós, M. Díaz-Lobo, S. Nonell, D. Velasco, *Angew. Chemie* **2012**, *124*, 12992–12995.
- [116] L. A. Carpino, H. Shroff, S. A. Triolo, E. S. M. Mansour, H. Wenschuh, F. Albericio, *Tetrahedron Lett.* **1993**, *34*, 7829–7832.
- [117] S. Prasad, I. Mandal, S. Singh, A. Paul, B. Mandal, R. Venkatramani, R. Swaminathan, *Chem. Sci.* **2017**, *8*, 5416–5433.