



HAL
open science

Nucleic acid ligands and benzothiazole-based fluorophores: synthesis and biological applications

Mauro Safir Filho

► **To cite this version:**

Mauro Safir Filho. Nucleic acid ligands and benzothiazole-based fluorophores: synthesis and biological applications. Other. COMUE Université Côte d'Azur (2015 - 2019), 2018. English. NNT: 2018AZUR4235 . tel-03270769

HAL Id: tel-03270769

<https://theses.hal.science/tel-03270769v1>

Submitted on 25 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT

Ligands d'Acides Nucléiques et Fluorophores Dérivés de Benzothiazoles

Synthèse et Applications Biologiques

Mauro SAFIR FILHO

Institut de Chimie de Nice

**Présentée en vue de l'obtention
du grade de docteur en Chimie
d'Université Côte d'Azur**

Dirigée par: Rachid Benhida

Co-encadrée par: Anthony R. Martin

Soutenue le: 14 Décembre 2018

Devant le jury, composé de :

Dr. Ling Peng, Directeur de Recherche au
CNRS, Aix-Marseille Université (Président)

Pr. Sylvain Routier, Professeur, Université
d'Orléans (Rapporteur)

Dr. Nicolas Spinelli, Maître de Conférences,
HDR, Université Grenoble Alpes (Rapporteur)

Dr. Rachid Benhida, Directeur de Recherche au
CNRS, Université Côte d'Azur (Directeur)

Dr. Anthony R. Martin, Chargé de Recherche au
CNRS, Université Côte d'Azur (Co-directeur)

Université Côte d'Azur

École doctorale « Sciences Fondamentales et Appliquées »

Ligands d'Acides Nucléiques et Fluorophores Dérivés de Benzothiazoles Synthèse et Applications Biologiques

Jury :

Dr Ling Peng	Directeur de Recherche au CNRS Aix-Marseille Université	Président du jury
Pr. Sylvain Routier	Professeur Université d'Orléans	Rapporteur
Dr. Nicolas Spinelli	Maitre de Conférences, HDR Université Grenoble Alpes	Rapporteur
Dr. Rachid Benhida	Directeur de Recherche au CNRS Université Côte d'Azur	Directeur de thèse
Dr. Anthony R. Martin	Chargé de Recherche au CNRS Université Côte d'Azur	Co-Directeur de thèse

Nucleic Acid Ligands and Benzothiazole-Based Fluorophores

Synthesis and Biological Applications

Abstract: Interdisciplinary is an obvious feature of scientific research. Notably, the continuously expanding researches in the medical field fostered this interdisciplinarity. Within this context, the work presented here deals with fluorescence spectroscopy and nucleic acid chemistry and aims at providing new tools suited for applications in life sciences. This manuscript presents two sections. In Part I, we developed novel series of highly bright and non-toxic push-pull fluorophores, compatible with cellular applications. In fact, we describe new benzothiazole-based fluorophores that can be structurally modulated to provide dyes with tunable photophysical properties. Their structure-photophysics relationship is also reported. Next, using these fluorophores, we produced new fluorescent probes to monitor the enzymatic activity of the β -galactosidase. Of note, since β -galactosidase is a marker of cellular senescence, we also used our probes for the detection of early stages of cellular senescence using fluorescence microscopy and flow cytometry analysis. Part II of this manuscript is devoted to the chemistry of nucleic acids. First, we describe the preparation and evaluation of novel RNA ligands targeting the IIIId domain of the HCV IRES. These ligands can make a base triplet with the IIIId sequence to provide specificity, and harbour amino functions to stabilize the triplet by electrostatic interaction with the phosphodiester of IIIId. Finally, the last part concerns the development of a new strategy allowing the post-synthetic functionalization of nucleic acids at the anomeric position. To do this, we designed new phosphoramidite platforms that, once incorporated into oligonucleotide strands, can be involved in various glycosidation reactions.

Key words: fluorescence, push-pull, β -galactosidase, RNA ligands, oligonucleotides, post-synthetic functionalization

Ligands d'Acides Nucléiques et Fluorophores Dérivés de Benzothiazoles

Synthèse et Applications Biologiques

Résumé: L'interdisciplinarité est désormais inhérente à la recherche scientifique. L'essor des recherches dans le domaine médical a notamment favorisé cette interdisciplinarité. C'est dans ce contexte que s'inscrivent les travaux présentés dans ce manuscrit. En effet, ce dernier vise à la production de nouveaux outils adaptés à des problématiques des sciences de la vie. Ce manuscrit traite de spectroscopie de la fluorescence et de chimie des acides nucléiques. Il est

divisé en deux sections: Dans la partie I, nous présentons de nouveaux fluorophores push-pull très brillants, non toxiques et pouvant être utilisé en cultures cellulaires. Ces composés, basés sur des motifs benzothiazoles, peuvent être modulés structurellement afin d'obtenir les propriétés photophysiques souhaitées. D'ailleurs, pour ces composés, une relation structure-propriétés photophysiques a été établie. Ensuite, en utilisant ces fluorophores, nous avons conçu de nouvelles sondes fluorescentes pour suivre l'activité enzymatique de la β -galactosidase. Cette dernière étant un marqueur de la sénescence cellulaire, nous avons utilisé nos sondes pour détecter la sénescence, *in vitro*, par des méthodes de microscopie de fluorescence et de cytométrie en flux. La partie II de ce manuscrit est consacrée à la chimie des acides nucléiques. Tout d'abord nous y décrivons la préparation et l'évaluation de nouveaux ligands d'ARN ciblant le domaine IIIId de l'IRES du VHC. Ces ligands présentent d'une part, un motif permettant de faire un triplet de base avec la séquence IIIId afin d'apporter de la spécificité, et d'autre part, des fonctions aminées afin de stabiliser le triplet par interactions électrostatiques avec les phosphodiesteres de IIIId. Enfin, la dernière partie concerne le développement d'une nouvelle stratégie permettant la fonctionnalisation post-synthèse d'acides nucléiques en position anomère. Pour ce faire, nous avons conçu des plateformes phosphoramidites qui, une fois incorporées dans des brins oligonucléotidiques, peuvent être impliquées dans diverses réaction de glycosidation.

Mots clés: fluorescence, push-pull, β -galactosidase, ligands d'ARN, oligonucléotides, fonctionnalisation post-synthétique

Acknowledgements

I would like express my sincere thanks to:

CAPES, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil, for the fellowship (Process N° 99999.001498/2015-01) that enable me to take my PhD.

The Université Côte d'Azur for these three years of studies.

France, to be a country in which I was pretty well received and to give me conditions to have a great period of life.

Pr. S. Routier and Dr. N. Spinelli who evaluated my work and provided me with the honor to join the jury of this thesis as reviewers.

Dr. L. Peng who evaluated my work and provided me with the honor to join the jury of this thesis as president of the jury.

My thesis director, Dr. Rachid Benhida, to welcome me in his laboratories (even before the PhD) and to give me this amazing opportunity to join his research group as a PhD student. Thank you for all your help and for always being available and present for any kind of discussion. Thanks for your guidance during all these years. I can't express how much I could learn and growth as a person and as a professional here. In the same way, I also would like to say many thanks to my advisor Dr. Anthony R. Martin. Thanks for all your support, for all advices, help, solutions and especially for your friendship. It was a pleasure to work with you. I feel honored to have shared these period with a so talented researcher and see at somehow a mirror to follow during my professional life.

I would like also to acknowledge to the permanent staff of our group in the ICN, especially to the members of our team Dr. Luc Demange and Dr. Cyril Ronco. I also would like to express my gratitude to the director of the *école doctorale* (ED-SFA) Dr. Elizabeth Taffin-de-Givenchy for your help and the fast solutions. For the other members of Molecules Bioactive (MB) group, Dr. Benoit Michel, Dr. Christophe Di Giorgio and Dr. Maria Duca I also would like to acknowledge for any discussion, time or contribution in any part of these work.

I also would like to thanks Dr. Sebastien Fiorucci for all for his help and contribution with DFT calculations, with corrections and for all the time and discussions we shared. My sincere thanks to Dr. Thierry Darmanin and Dr. Guilhem Godeau for your help and friendship. Thanks, in particular to Dr. Godeau to give me the opportunity to join the Summer School in Moscow in 2017. Incredibly, it was there that I improve my French skills :).

I could not forget to thank Dr. Lionel Massi and Dr. Mark Gaysinski from the mass spectrometry and NMR platforms, respectively.

I also would like to thank Dr. Guillaume Robert (C3M), Maéva Gesson (C3M) and Sandy Giuliano (Centre Scientifique de Monaco) for their contribution and collaboration in different projects.

To all the colleagues from the MB group and from the ICN: Lorenzo, H el ene, Estelle, Nikolas Antoine, Hella, Hamza, Celine, Chlo e, Tom, C ecile, Nelli, Lou, Claudio, Janah, Antonien, Vincent, Serena, Rime, Seb, Cecile, Janwa, Sabri, Caroline, Hernando, Alexandra and Anh. We had a great time and a lot of fun. Thanks for all happy hours, for the *randonn ees*, dinners and day trips. I wish you all the best for your lives.

Especially, I would like to thank Alex and Sasha to share these three years with me every day in the lab, for the lunches, for the coffee time, for the talks and jokes. Thank you guys, it was really a pleasure. Hope to see you again in the future. I can't forget also to thank Ngoan and Guillaume for all the talks, especially during these last months. I wish you the best during your thesis. And for Pascal, thanks dude, to be a nice colleague. It has been a pleasure to work with you. I wish you success in your near future. You deserve it.

To the good friends I could meet in Nice: Rui, Sabine, Siguara, Ana, Angelo, Mathieu, Maria, Johann, Maysa, Rog erio, Sabrina, Will and Mavi. With you I could have the impression to be at home sometimes. It was really nice to share these years together. It will be so hard to forget all the Terrace parties, the dinners and the moments at the beach. Hope to see you again, anywhere, in Brazil or in any other country.

Aqui gostaria de agradecer aos professores e orientadores que de alguma forma participaram desta trajet oria. Obrigado aos professores Ione Maluf Baibich, Silvana In es Wolke e Paulo Augusto Netz pela oportunidade de trabalhar em seus laborat orios durante minha passagem pelo Instituto de Qu mica da UFRGS. Obrigado em especial ao professor Fabiano Rodembusch, pelos conselhos, pela amizade e parceria, ao grande parceiro Rodrigo Duarte, pelo companheirismo e por toda ajuda e introdu  o no universo de laborat rio de qu mica org nica e aos demais amigos do Grupo de Pesquisa em Fotoqu mica Org nica Aplicada (GPFOA) pelos anos em que estive com voc s. Devo muito deste trabalho   todos os mentores que fizeram parte da minha trajet oria acad mica.

A todos os amigos que ficaram no Brasil. Obrigado pelos anos que tive junto de voc s. Em especial aos amigos que a qu mica me deu, Gui, Maria, Nati, Rafa, Mila e Joana. Obrigado   aqueles que estiveram presentes nos momentos mais dif ceis. Saudades de voc s.

Agradeço à minha família, meus tios Joel e Ilce, aos meus avós pelo amor e carinho de sempre. Saudades.

Aos meus sogros Francisco e Cristina, por todo carinho, acolhimento e por sempre terem sido presentes. Tenho muito a agradecer por tudo que sempre fizeram por mim. À minha mais nova vó, Estela, por ser sempre tão amorosa e por realmente me fazer sentir como um neto.

Agradeço aos meus pais, Mauro e Suzane, por absolutamente tudo que fizeram por mim. Pelo amor, pelo carinho e dedicação. Por terem me oferecido a melhor educação possível e pelos princípios que me passaram. Se de alguma forma cheguei até aqui, devo muito disso à vocês, por terem sempre dado o máximo para me proporcionar e oferecer as melhores condições. Obrigado por aceitarem as minhas escolhas (as vezes um não poderia ter sido bom, mas...heheh). Saudades imensas, amo vocês. Vocês são os melhores pais do mundo.

Agradeço ao meu amor, minha grande e fiel companheira Gabriela, por todo amor, carinho e dedicação durante esses longos 8 anos. Te agradeço por ser essa pessoa tão maravilhosa e sonhadora, por ter nos embarcado nessa grande aventura. Se não fosse por ti, jamais poderíamos ter vivido tudo isso. Obrigado por estar sempre ao meu lado, por todo incentivo e ajuda que tu sempre me deu, me dá e, tenho certeza, continuará dando. Por ser o meu suporte, a minha maior motivação. Só tenho a te agradecer. Ninguém mais do que tu teve um papel tão principal na minha vida durante todos esses anos. Te amo muito, obrigado por tudo. Obrigado por todas loucuras, viagens e momentos que tivemos ao longo desses anos. Tu mais do que ninguém sabe que elas serão sempre inesquecíveis. Um beijo grande, te amo sempre. Essa conquista também é tua.

Table of Contents

General Introduction	1
Chapter 1 - Photophysical Processes and Fluorescence Sensing	7
1.1 Light Absorption.....	7
1.1.1 Electronic Transitions in Organic Molecules.....	8
1.1.2 Nature of the Electronic States.....	8
1.1.3 Franck-Condon Principle	9
1.1.4 The Beer-Lambert Law	9
1.2 Excited States and Photophysical Processes	10
1.2.1 Photophysical Processes	10
1.2.2 Insights on the Excited State Lifetime	11
1.2.3 Fluorescence Quantum Yield (Φ_f).....	12
1.3 Small-Molecule Organic Fluorophores.....	13
1.3.1 Classic Fluorophores.....	13
1.3.2 Environment Responsive Fluorophores	14
1.4 Solvatochromism	16
1.4.1 Types of Solvatochromism	16
1.5 The Intramolecular Charge Transfer (ICT) Excited State and the Solvatochromism	17
1.5.1 Intramolecular Charge Transfer (ICT) and Locally Excited (LE) States	17
1.5.2 "Planar" ICT and Twisted Intramolecular Charge Transfer (TICT) States.....	18
1.5.3 Solvent Effect and the Energy of the ICT State.....	19
1.5.4 The Role of CT States on the Radiative and Non-Radiative Decays.....	21
1.5.5 Viscosity Effects on the Excited States (LE, ICT and TICT) Interplay	21
1.6 Photoinduced Electron Transfer (PET).....	22
1.7 Fluorescence Labeling	24
1.8 Fluorescence Sensing.....	24
1.8.1 Self-Immolative Linkers	26
1.8.2 Sensing enzymatic activity - Intensity-Based Probes	27
1.8.3 Environmental Sensing Based on Push-Pull Fluorophores.....	42
1.9 Conclusions - Chapter 1	44
Chapter 2 - Development of new push-pull benzothiazole-based fluorophores: synthesis and photophysical studies	47
2.1 Introduction - Styryl-Based Fluorophores	48
2.2 Synthesis of New Styryl-Based Fluorophores	50
2.2.1 Synthesis of Benzothiazole Precursors 107 and 109	50
2.2.2 Synthesis of Benzothiazole-based Styryl Fluorophores.....	51

2.3 Photophysical Studies	53
2.3.1 Modulation of the Acceptor Moiety - Substituted Phenyl Series: Dyes 85-93	53
2.3.2 Modulation of the Acceptor Moiety - Polycyclic Aromatic Hydrocarbon Series: Dyes 94-98	60
2.3.3 Modulation of the Donor Moiety - Spectroscopic Impact of the Alkylation of the 6-Amino Functional Group	65
2.4 Benzothiazole-Based Fluorophores Bearing Extended π -Conjugated Spacer	68
2.4.1 Synthesis of Benzothiazole-Based Fluorophores Bearing Extended π -Conjugated Spacer.....	69
2.4.2 Photophysical Properties of the Benzothiazole-Based Fluorophores Bearing Extended π - Conjugated Spacer	70
2.5 Conclusion - Chapter 2	77
Chapter 3 - Development of highly sensitive fluorescent probes for the detection of β- galactosidase activity - application to the real-time monitoring of senescence in live cells.....	79
3.1 Introduction - Therapy Induced Senescence (TIS) and senescence-associated β -galactosidase (SA- β -gal)	80
3.2 Synthesis of Probes 121-124.....	82
3.3 Spectroscopic Properties	83
3.4 Kinetics	86
3.5 Stability, Selectivity and Fluorescence Emission in Cell.....	87
3.6 Limit of Detection.....	89
3.7 Fluorescence Imaging of SA- β -gal in Living Cells	90
3.8 Conclusions - Chapter 3.....	94
Conclusion Part I	95
References	96
Chapter 4 - Synthesis and Biophysical Studies of Novel RNA Ligands to Target HCV IRES III_d Loop.....	107
4.1 State of the Art	108
4.1.1 The Nucleotides and the Primary Structure of DNA and RNA	108
4.1.2 RNA Structure and Functions	109
4.1.3 RNA Targets	114
4.1.4 Hepatitis C Virus (HCV).....	115
4.1.5 RNA Ligands	117
4.1.6 Oligonucleotides (ONs)	117
4.1.7 Small Molecules.....	126
4.2 Ligands to Target the HCV IRES III _d Loop.....	136
4.2.1 Ligand Design.....	136
4.2.2 Synthesis of the Ligands 166a-j	137

4.2.3 Synthesis of Ligands 167a-c	138
4.2.4 Affinity Studies with HCV IRES III _d Loop	139
4.3 Conclusions - Chapter 4.....	142
Chapter 5 - Post-Synthetic Modification of Oligonucleotides	145
5.1 State of the Art.....	146
5.1.1 Chemical Synthesis of Oligonucleotides - Phosphoramidite Chemistry	146
5.1.2 Post-Synthetic Modification of Oligonucleotides.....	149
5.2 Post-Synthetic Functionalization at the Anomeric Position - An Innovative Approach.....	159
5.2.1 Aglycone Moiety Screening.....	160
5.2.2 The Synthesis of Phosphoramite 214	163
5.2.3 Oligonucleotide Synthesis.....	165
5.2.4 Preliminary Studies on the Post-Synthetic Functionalization of ODN1 and ODN2	167
5.3 Method Optimization and the Alternative Approach for the Orthogonal PSF	170
5.3.1 Complementary Studies on the Post-Synthetic Functionalization using the ODN1	170
5.3.2 Ortho-Alkynyl Benzoates - The Potential and Orthogonally Activable Synthon	170
5.4. Conclusions - Chapter 5.....	173
References	174
General Conclusion	183
Experimental Section	189
General Methods.....	189
Synthetic Procedures.....	195
Annexes	233

General Introduction

Interdisciplinary is nowadays an obvious feature of scientific research. Among the major research areas in life sciences, the fields in the interface between chemistry, physics and biology are still in continuous growth. Today one of the greatest challenges in the current medicinal chemistry concerns the control of the rapid evolution of human diseases, such as cancers, viral and bacterial pathologies. Within this context, the search for innovative medical tools and technologies is the central subject of an incredibly high number of researches. Therefore, the work presented herein lies at the interface between chemistry and chemical biology and aims at providing new tools and technologies useful in the context of life sciences.

This thesis manuscript discloses part of the research work developed during my PhD, in the group of Bioactive Molecules, at the Institute of Chemistry of Nice (ICN), under the supervision of Dr. Rachid Benhida and Dr. Anthony R. Martin. The work described here is compiled in two parts, and covers two different research areas. The Part I is centered on the development of fluorescent probes for biological applications and encompasses the three first chapters of this manuscript (**Chapters 1-3**). The Chapter 1 is dedicated to state the fluorescence spectroscopy field, underlining key aspects related to the photophysical properties of organic fluorophores. Different classes of fluorophores will be presented. A special section is dedicated to introduce the concept of solvato(fluoro)chromism and its relationship with the formation of polar states will be emphasized. Finally, at the end of this chapter, an overview addressing the major approaches used in fluorescence sensing applications will be provided, with emphasis on methods applied to detect enzymatic activity. Chapter 2 describes our work dealing with solvatochromic fluorophores. In fact, the synthesis and the spectroscopic properties of two novel series of benzothiazole containing push-pull fluorophores will be discussed (**Figure I1**). The first section is focused on styryl-based fluorophores (**Figure I1-A**). The second discloses the assessment of a another family of highly bright fluorescent dyes, displaying an extended π -conjugated framework (**Figure I1-B**). Of note, the effect of different structural modifications on the fluorophore core, *i.e.* different donor and acceptor moieties and the π -conjugated spacer, will be discussed and correlated with the observed spectroscopic properties.

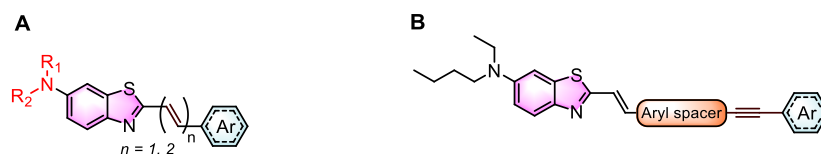


Figure 11. General structure of the benzothiazole-based push-pull fluorophores that will be discussed in **Chapter 2**.

Finally, with the understanding of the structure-photophysical relationship of our fluorophores, we proceed to the development of novel fluorescent tools by exploiting this new family of fluorogenic molecules. Thus, **Chapter 3** is focused on the valorization of these powerful dyes and their use for practical applications. Therefore, the development of four novel fluorescent probes (**Figure 12**) to monitor the enzymatic activity of the β -galactosidase (β -gal) will be described. The spectroscopic studies and the *in vitro* enzymatic investigations will be presented. Thereafter, the monitoring of senescence-associated β -galactosidase activity in living cells, using fluorescence microscopy and flow cytometry analysis will be reported.

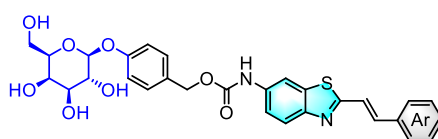


Figure 12. General structure of the fluorescent probes to report β -galactosidase activity that will be discussed in **Chapter 3**.

As aforementioned, the second part of this manuscript deals with another research area. Indeed, in this Part II, we will investigate the chemistry of nucleic acids (**Chapters 4 and 5**). Essentially, two different aspects of nucleic acid chemistry will be covered in this part. The **Chapter 4** is dedicated to the discussion concerning the development of new small-sized RNA ligands. Thereby, this chapter will strictly focus on RNA. In fact, the RNA functions, structures and its uses as a druggable target will be exposed. Examples of oligonucleotide-based and small-sized RNA ligands will be highlighted. Moreover, we will present the synthesis of two new series of multimodal ligands, designed using a triplex-based rational approach and disclose the structure-binding relationship studies against the HCV IRES III_d loop as a relevant therapeutic RNA target (**Figure 13**).

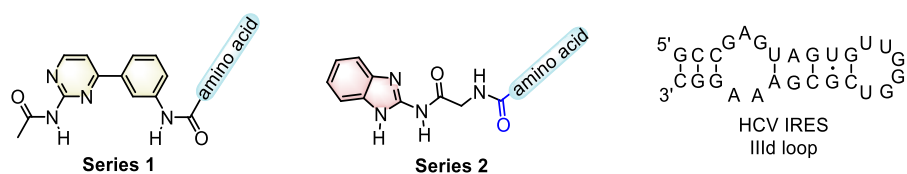


Figure I3. General structure of the reported RNA ligands to target the HCV IRES IIIId loop discussed in **Chapter 4**.

Finally, **Chapter 5** will discuss our research concerning the synthetic modification of nucleic acids. Here, we will present the development of a new methodology applied to the post-synthetic functionalization of oligonucleotides directly on the anomeric position. Thereby, at the beginning of this chapter we will underline the existing chemical approaches used to post-synthetically functionalize oligonucleotides. Afterwards, the whole strategy applied to identify potential deoxyribose platforms harbouring benzoyl aglycones units and their incorporation into DNA will be revealed. Finally, the preliminary results concerning the post-synthetic functionalization using a Lewis acid catalyzed glycosylation reaction will be discussed (**Figure I4**).

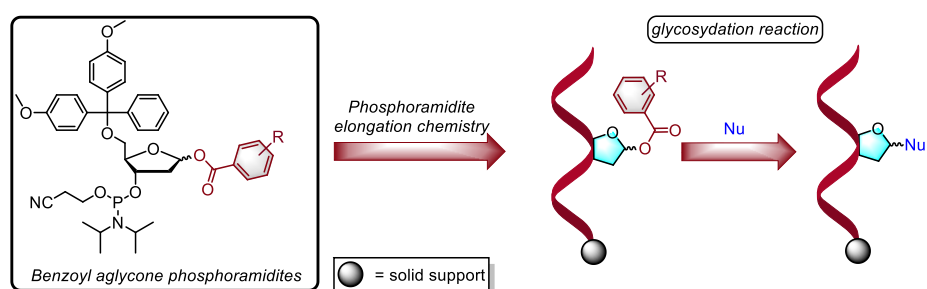


Figure I4. Synthetic approach applied for the post-synthetic functionalization of oligonucleotides explored in **Chapter 5**.

**Part I. Design, photophysical properties and sensing applications
of novel benzothiazole-based fluorophores with remarkable
environmental sensitivity**

Part I - Chapter 1

Chapter 1 - Photophysical Processes and Fluorescence Sensing

Undoubtedly, fluorescence spectroscopy is among the most powerful tools for sensing chemical species and study the dynamics of biological systems. Fluorescence provide high level of accuracy as well as temporal and spatial resolution. In addition, fluorescence imaging techniques offer a non-invasive approach to visualize biological environments with high sensitivity, requiring extremely low concentrations of the fluorescent probe. For this reasons, the design and development of novel fluorescent tools is a continuously expanding field of research. Of note, the design of organic molecules featuring high brightness, environmental sensitivity, photostability and fine-tuned photophysical properties has attracted a lot of attention. In fact, the combination of robust fluorophores with the last generations of optical devices have enabled a fast development of different areas of medical research. For instance, medical diagnosis based on fluorescence imaging and fluorescence-guided surgery are some important optical tools available in medical sciences. Furthermore, the labeling of biomolecules with fluorescent markers is still one of the most powerful and easy implemented approaches to probe and study a wide variety of processes, ranging from simple bimolecular interactions to the study of mutation and structural modifications on nucleic acids.

Along this chapter, different aspects within the fluorescence spectroscopy field will be presented and discussed, starting from the basic fundamentals associated with this physical phenomena to its uses in sensing applications.

1.1 Light Absorption

The origins of photophysical processes in organic molecules is closely related to the phenomena of light absorption. This, in turn, is one of the innumerable processes that can occur from the interaction between light and matter and is related to electronic transitions between different electronic states. When a photon of energy corresponding to the energy difference between two electronic states is absorbed, the excitation of an electron from the ground state to a more energetic excited state takes place. As a result, the total energy of the molecule increases.¹

1.1.1 Electronic Transitions in Organic Molecules

Organic molecules present six different types of electronic transitions involving the σ , n and π orbitals (**Figure 1.1**). Among these transitions, the relevant and important transitions in photophysical processes occur between n and π orbitals ($n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$) and involve the absorption of photons generally in the UV-Vis-NIR region. The excitation of an electron from the σ orbitals requires higher energy (far-UV absorption) and will not be considered here.²

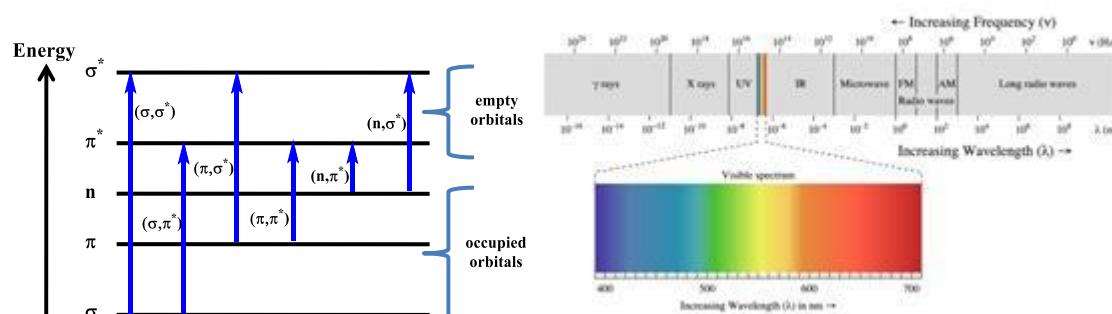


Figure 1.1. Electronic transitions in organic molecules (left). Electromagnetic spectrum (right).

1.1.2 Nature of the Electronic States

To better understand the absorption and subsequent photophysical processes it is important to consider the contribution of nuclear motions and how they correlate with the electronic structure of the molecules. The total energy of a molecule is composed of an electronic energy and a portion of kinetic and potential energy, referring to all modes of vibration and rotation of the molecules ($E = E_{el} + E_v + E_r$). In addition, these modes of vibration and rotation are also quantized. However, the energy difference between electronic states is significantly higher than the energy difference between the vibrational and rotational states. As a result, each electronic state may contain multiple and discrete vibrational (v) and rotational levels. Thus, electrons can be excited from the lowest vibrational level of its ground state ($S_0, v = 0$) to different vibrational levels of the excited state S_n (**Figure 1.2-A**). Therefore, the resulting absorption (and also the emission spectra), in the UV-Vis-NIR region, of organic molecules in solution are normally presented as band shapes.³

1.1.3 Franck-Condon Principle

The concept of vertical electronic transitions is also important to discuss. The electronic transition associated with maximum absorption (and the emission maximum) is known as the Franck-Condon (FC) transition. Since an electronic transition occurs in a sufficiently short time compared with molecular motion (10^{-15} s vs 10^{-10} - 10^{-12} s), the Franck-Condon transition occurs between the vibrational levels of the electronic states S_0 ($v = 0$) and S_{nFC} ($v = n$) that present the most similar molecular structure, *i.e.* the higher overlap between the ψ^2 of both upper and lower vibronic states (**Figure 1.2-B**).^{4,5}

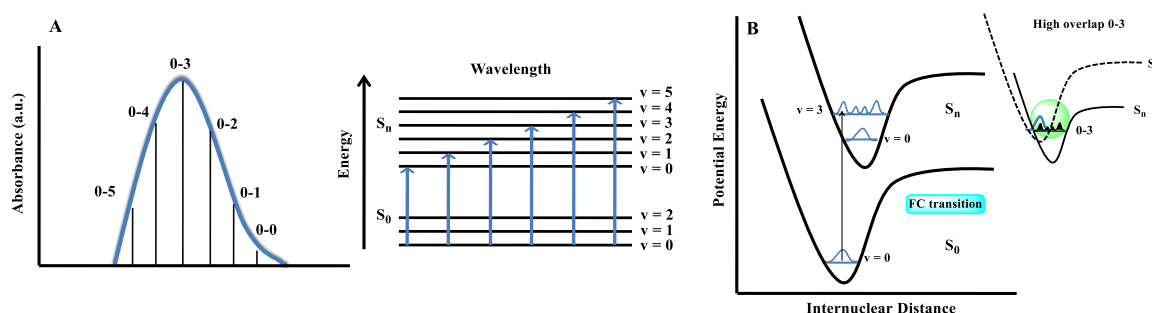


Figure 1.2. (A) Correlation between electronic transitions and absorption spectra in organic molecules. (B) Representation of Franck-Condon principle.

1.1.4 The Beer-Lambert Law

The physical parameter to describe the probability of occurrence of an electronic transition is called molar absorptivity or molar extinction coefficient (ϵ). It is characteristic of each molecule and depends on the wavelength of the incident radiation and the solvent. The absorbance (A) is the amount of light absorbed by a molecule. These two physical properties (A and ϵ) are directly related by the Beer-Lambert Law (**Equation 1.1**).³

$$A(\lambda) = \log \frac{I_{\lambda}^0}{I_{\lambda}} = \epsilon(\lambda) b C, \text{ Equation (1.1)}$$

where C is the molar concentration of the solute and b the optical path length traveled by the radiation through the sample. For allowed electronic transitions ($\pi \rightarrow \pi^*$), the values of ϵ are in the range of 10^3 to 10^5 (M.cm)⁻¹, while for $n \rightarrow \pi^*$ electronic transitions, prohibited by symmetry, the values are significantly lower, in the range of 10^2 (M.cm)⁻¹.¹

1.2 Excited States and Photophysical Processes

1.2.1 Photophysical Processes

As briefly described above, organic molecules can be excited by light passing from their ground state (S_0) to an excited electronic level (S_n). However, since the excited state is a higher energy state, the molecule tends to return to its ground state in order to minimize its energy. The deactivation of excited states may occur through several pathways, including photochemical⁶ or photophysical processes.

The photophysical processes can be separated in two classes: radiative and non-radiative. Non-radiative decay can arise from two basic mechanisms: vibrational relaxation (vr) and *internal conversion* (IC). These processes favor a rapid deactivation rates; vibrational relaxation (10^{-12} - 10^{-10} s) occurs within the same electronic state and largely include loss of energy as heat transfer to the environment (*i.e.*, solvent). *Internal conversion* (10^{-11} - 10^{-9} s) is favored when an effective overlap of the wave functions of the vibrational levels of two electronic states with the same multiplicity exists ("*vibronic coupling*"). However, it is more probable to occur between two *excited* states ($S_2 \rightarrow S_1$, for example) as the energy difference between the S_2 and S_1 is lower compared with the $S_1 \rightarrow S_0$ gap. A third mechanism, also nearly isoenergetic, known as *intersystem crossing* (ISC), occurs between states of different multiplicity (10^{-10} - 10^{-8} s). While forbidden by quantum mechanics restrictions, ISC can be promoted by the presence of internal or external heavy atoms and due to the *spin-orbit coupling*, resulting in the population of triplet excited states.²

The radiative decays encompass two types of luminescent processes: fluorescence and phosphorescence. In the fluorescence, the emission of light occurs accompanied by an electronic transition between two states with the same multiplicity (spin conservation, 10^{-12} - 10^{-7} s). Normally, with rare exceptions, it involves the relaxation from the lowest vibrational level of the lowest singlet excited state (S_1) to the ground state (S_0). By contrast, the phosphorescence emission involves a prohibited electronic transition between states of different multiplicities (without conservation of spin). Thereby, triplet excited states are long-lived states compared with singlet states and phosphorescence emission can take place during longer times (10^{-3} - 10^2 s). Below, in **Figure 1.3**, is presented the simplified Jablonski diagram, which summarizes the photophysical processes discussed in this paragraph.⁴

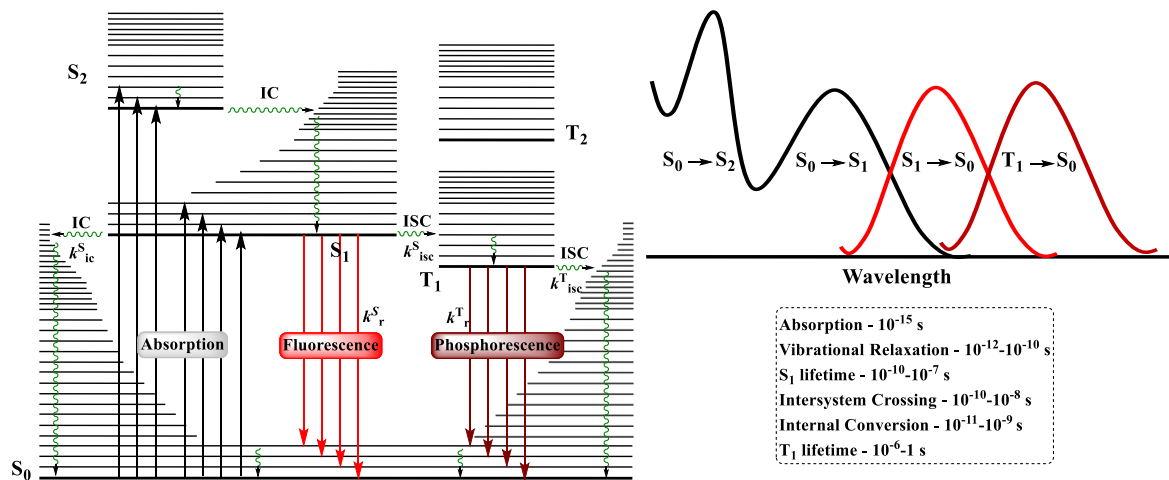


Figure 1.3. Simplified Perrin-Jablonski diagram. (Adapted from²)

This simplified Jablonski diagram also highlights some important behaviors, as: (i) in the absence of non-linear processes, the emission always involves photons of lower energy than in absorption due to energy losses resulting from vibrational relaxation; (ii) once in the excited state, all the deactivation processes compete until the molecule returns to its ground state. (iii) triplet states are always less energetic than the corresponding singlet state (according to the **Hund's rule**). Moreover, fluorescence and phosphorescence emissions take place with notorious probability from the lowest vibrational level of the lowest excited state (S_1 and T_1 , respectively - **Kasha's rule**).⁷

In addition, molecular stiffness, the incorporation of electron donating or electron-withdrawing substituents and the presence of heavy atoms are factors that may favor different pathways of deactivation.¹

1.2.2 Insights on the Excited State Lifetime

Just after the excitation of a population of molecules with a short time pulse of light, the resulting emission intensity decays exponentially with a specific time, that reflects the average lifetime of the first singlet excited state S_1 . In the Jablonski diagram (**Figure 1.3**) are highlighted the kinetic constants of the different processes that may occur in parallel and compete with the fluorescence emission once the molecules reach the S_1 state.

$k_{r(S)}$ - kinetic constant for $S_1 \rightarrow S_0$ fluorescence emission decay

$k_{ic(S)}$ - kinetic constant for $S_1 \rightarrow S_0$ internal conversion

$k_{isc(S)}$ - kinetic constant for $S_1 \rightarrow T_1$ intersystem crossing

Here, it is convenient to introduce the general non-radiative decay kinetic constant ($k_{nr(S)}$) and the global decay kinetic constant ($k_{t(S)}$)

$$k_{nr(S)} = k_{ic(S)} + k_{isc(S)}, \text{ Equation (1.2)}$$

$$k_{t(S)} = k_{r(S)} + k_{nr(S)}, \text{ Equation (1.3)}$$

It is important to highlight that several others non-radiative processes may also be present and will contribute to enhance the general non-radiative kinetic constant ($k_{nr(S)}$).

Considering that the $S_1 \rightarrow S_0$ decay is a single elementary step process, the rate of the excited state decay follows a first order kinetic that is directly proportional to the concentration of the excited molecules in a specific time t ; the proportionality factor corresponds to the global kinetic constant $k_{t(S)}$ decay.

$$\frac{dS_1}{dt} = -k_{r(S)}S_1 - k_{nr(S)}S_1 = -k_{t(S)}S_1, \text{ Equation (1.4)}$$

$$S_1(t) = S_1^0 e^{-k_{t(S)}t}, \text{ Equation (1.5)}$$

According to the kinetic constants highlighted above, the first singlet excited state lifetime τ_S can be described by the **Equation (1.6)** and represents the time during which the S_1 concentration decreases to $1/e$ of its initial value.

$$\tau_S = \frac{1}{k_{t(S)}} = \frac{1}{k_{r(S)} + k_{nr(S)}}, \text{ Equation (1.6)}$$

1.2.3 Fluorescence Quantum Yield (Φ_{fl})

Fluorescence quantum yield (Φ_{fl}) is the fraction of excited molecules that returns to the S_0 ground state after emission of a fluorescence photon. It gives the information concerning the fluorescence emission efficiency related other photophysical processes (non-radiative decays) of the excited molecule. It also correlates the rate of fluorescence emission ($k_{r(S)}$) with the rate of all photophysical processes ($k_{t(S)}$) according to the **Equation 1.7**.

$$\Phi_{fl} = \frac{k_{r(S)}}{k_{t(S)}} = k_{r(S)} \times \tau_S, \text{ Equation (1.7)}$$

1.3 Small-Molecule Organic Fluorophores

Fluorophores constitute the basis of fluorescence spectroscopy. A great diversity of fluorescent probes are already known and they can be simply divided into two main groups: intrinsic and extrinsic fluorophores. Intrinsic fluorophores are those molecules in which fluorescence occurs naturally. Aromatic amino acids, as tryptophan, tyrosine and phenylalanine, chlorophyll, flavonoids and NADH are some examples of naturally emissive compounds. By contrast, extrinsic fluorophores consist in a large group of synthetic fluorescent molecules that can be "used" to afford or modify fluorescence properties of a system.

In this section, the synthetic fluorescent dyes will be distinguished and classified in two major groups, according to the effect of environmental changes on the photophysical properties: classic (or non-environment responsive) and environment sensitive fluorophores. Both groups will be illustrated with some examples and their features will be discussed; although, more attention will be focus on the environment-sensitive fluorophores as they are central to the work presented in **Chapters 2 and 3**.

1.3.1 Classic Fluorophores

Non-environment responsive fluorophores encompass a group of extensively used fluorescent probes including non-substituted polycyclic aromatic hydrocarbons (PAH) such as pyrene (**1**) and anthracene (**2**) and heteroaromatic derivatives including fluorescein (**3**), rhodamine (**4**), bodipy (**5**), cyanine (**6**) and squaraine (**7**) derivatives (**Figure 1.4**). These dyes typically contain an extended π -conjugated system fully delocalized which gives very poor sensitivity to environmental changes.

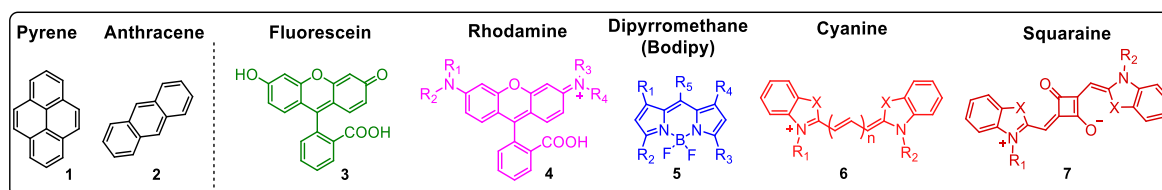


Figure 1.4. Structure of widely used classical fluorophores.

Due to their very high Φ_{fl} and long wavelengths of absorption and emission bands, fluorescein, rhodamine and bodipy dyes have found suitable applications in fluorescence sensing, imaging and biolabeling applications.^{8,9} Among these, bodipy dyes have emerged as a powerful alternative to fluorescein and rhodamines due to their extremely high brightness ($\Phi_{fl} \sim 1$; $\epsilon \geq 80,000 \text{ (M.cm)}^{-1}$), insensitivity to pH and broad range of excitation and emission wavelengths. In addition, the modification of the *meso* position also make bodipy dyes useful to design molecular rotors (**Figure 1.4**, position R₅).¹⁰ Cyanines and squaraines represent the groups of NIR dyes, displaying long-wavelength absorption and emission with extremely high extinction coefficients ($> 100,000 \text{ (M.cm)}^{-1}$). They have been used in bioconjugation and medical applications.^{11,12} Moreover, the high tendency to aggregate in the ground state (H and J aggregates) confers to these dyes a certain degree of sensitivity to environmental changes.¹³ However, this it is out from the scope of this thesis and will not be discussed.

The major drawback concerning all these probes lies in their very small Stokes shift values, which considerably enhance self-quenching effects. Moreover, the fast rate of photobleaching of fluorescein analogues, especially after bioconjugation, may also limits their applications.

1.3.2 Environment Responsive Fluorophores

The solvent effect and the local polarity have strong and complex effects on the photophysical properties of some organic dyes.¹⁴ Nevertheless, several other factors in addition to polarity may also affect the fluorescence emission spectra, the Φ_{fl} , τ_S and also the nature of the excited state of organic fluorophores. This section will be dedicated to the discussion concerning the effect of the local polarity and viscosity and how they can affect/modulate the absorption and different mechanisms once the molecule reaches the excited state. More specifically, the following discussion will focus only on push-pull dyes and no other types of environment responsive fluorophores, such as flippers or those presenting excited state intramolecular proton transfer (ESIPT), will be taken in consideration.

1.3.2.1 Push-Pull Dyes

Push-pull dyes encompass a broad group of environment sensitive organic dyes. They have been the subject of intensive research due to their wide range of potential applications

1.4 Solvatochromism

The interaction between the solute and the solvent can occur through specific and non-specific effects. Non-specific interactions (known as general solvent effect) are mainly due to the macroscopic parameters of the solvent (refractive index and dielectric constant) while specific effects involve solute-solvent interactions (for example hydrogen-bonding).³² In molecules whose dipole moments are different between the ground state (μ_g) and the excited state (μ_e), the solvent may interact and stabilize these states to different extents. Thus, as a result, changes in the spectroscopic properties (absorption or emission) may be observed. However, only molecules that are themselves polar can display pronounced sensitivity to polarity. Moreover, the higher the dipole-moment differences the larger the solvatochromic effect.^{2,33}

1.4.1 Types of Solvatochromism

Solvatochromic effects can be expressed either on the absorption or on the emission spectra and can be accompanied by changes in the absorbed or emitted colors. Organic molecules can basically present two different types of solvatochromism: positive and negative. In molecules whose dipole moment increases after the absorption of a photon ($\mu_e > \mu_g$), polar solvents will stabilize more strongly the excited state than the ground state. Hence, a bathochromic shift of the absorption or emission maxima will be observed (positive solvatochromism). In contrast, if the ground state is more polar than the excited state ($\mu_g > \mu_e$), then an hypsochromic shift in solvents with increased polarity will be observed (negative solvatochromism), since the energy barrier between the states augments (**Figure 1.7**).

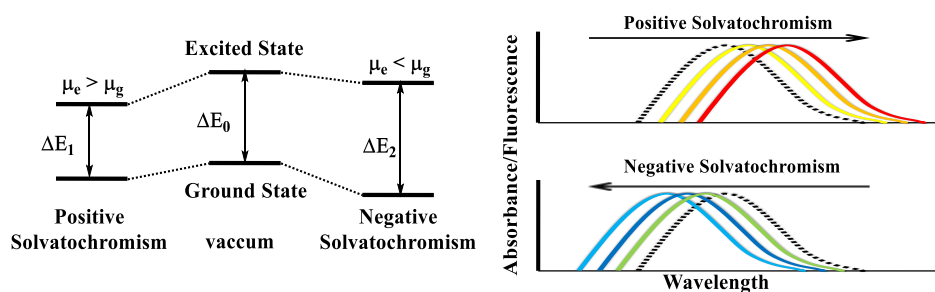


Figure 1.7. Schematic representation of positive and negative solvatochromic shift.

the excited state due to general solvent effects, but also dictates which state has the lowest energy.⁴

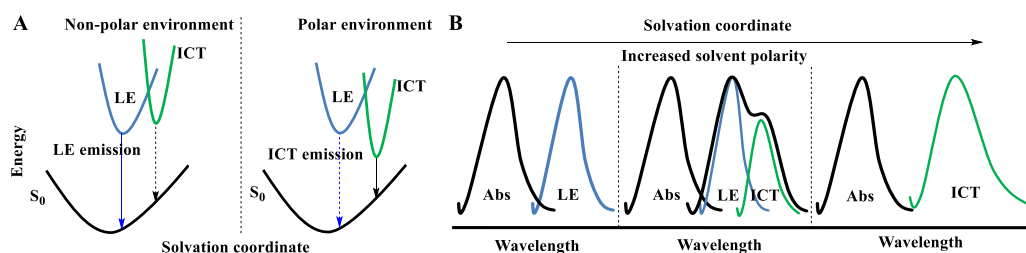


Figure 1.9. (A) Effect of solvent polarity on the energies of LE and ICT states. (B) Schematic representation of the absorption and emission spectra for systems displaying LE and/or ICT excited states.

1.5.2 "Planar" ICT and Twisted Intramolecular Charge Transfer (TICT) States

A particular case of ICT states happens when the relaxation through the charge transfer state is accompanied by internal rotation within the fluorophore (**Figure 1.10**). In the resulting TICT state, the molecule reaches a fully charge separated state with concomitant loss of donor-acceptor conjugation. As the rotation angle approximates to 90° , the molecule reaches its maximum excited state dipole-moment. Thus, TICT can also be stabilized with increased solvent polarity. Interestingly, in some cases, molecular rotation is a requirement to afford complete charge separation and the formation of charge transfer state. Therefore, constraining bond rotation by internal or external factors can strongly modulate the LE and TICT interplay.³⁶ In addition, examples of molecules that pass gradually from a "planar" ICT to a "non planar" TICT state in highly polar solvents can also be found in the literature.^{37,38}

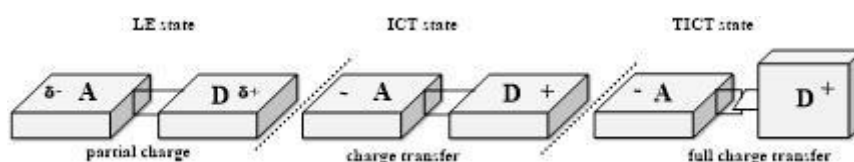
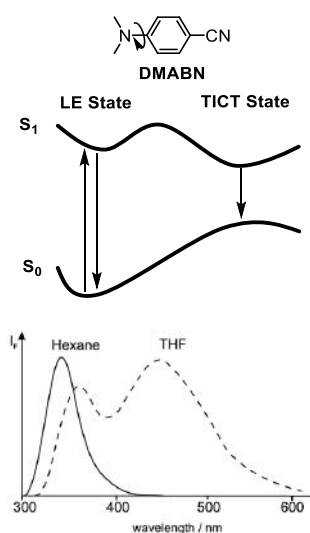


Figure 1.10. Schematic representation of LE, "planar" ICT and "non-planar" TICT states.

The first example of a system depicting ICT reported in the literature is the 4-dimethylaminobenzonitrile (**DMABN, Figure 1.11-A**) (1962).³⁹ In the excited state, charge is transferred from the dimethylamino donating group to the electron acceptor cyano moiety. Apart from its simple molecular structure, **DMABN** presents a single emission band in non-

polar solvents (*i.e.*, hexane) and a dual emission in the more polar THF; this is due to the formation of a TICT state. Since then, innumerable **DMABN** derivatives have been used and studied in the last 20 years. In particular, the group of Prof. Zacharisse have "intensively" investigated the influence of solvents on the LE/(T)ICT interplay in this class of molecules (**Figure 1.11-B**).^{36,40} Essentially, they observed that the optimal disposition of the electron donating (*N*-methyl-*N*-isopropylamine) and the electron withdrawing (cyano) groups along the benzene ring are strictly necessary to enable the formation of the charge transfer state in this class of compounds. Of note, the analogues displaying both groups *para*-oriented are favored to form TICT states when compared with the parent compounds with *meta* and *ortho* orientation.

(A)



(B)

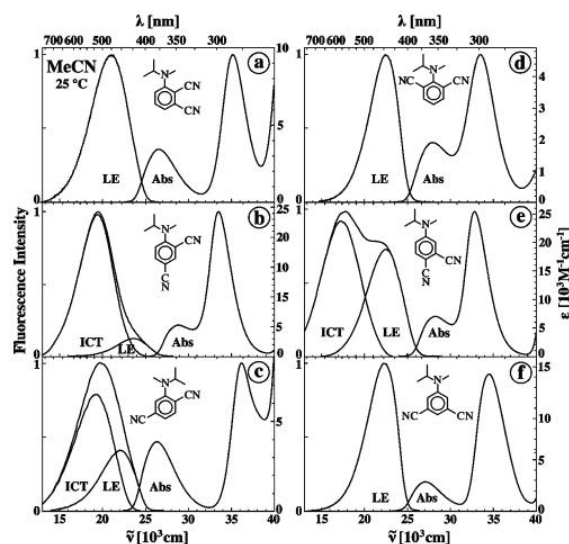


Figure 1.11. (A) Structure of DMABN and its fluorescence emission spectra in hexanes and THF. (B) DMABN analogues displaying fluorescence emission from LE and (T)ICT states.

1.5.3 Solvent Effect and the Energy of the ICT State

Solvatochromic effects, in the emission spectra, originate from the formation of the more polar charge transfer excited states. As aforementioned, polar states can be highly stabilized by specific and/or general solvent effects. Once a molecule reaches its excited state, the dipole-moment is dramatically enhanced due to the formation of an ICT state. At this point, solvent molecules, which are also dipoles, may reorient themselves in order to better stabilize the more dipolar fluorophore. As a result, the energy of the excited state became lower and

the wavelength of the emitted photons increases ($h\nu_2'$ vs $h\nu_1'$, **Figure 1.12**). Since solvent relaxation is based on dipole-dipole interactions, this effect may be strengthened as the solvent polarity augments. Thus, a gradual increase in the solvent polarity reduces the energy gap between the excited and ground states, hence causing a bathochromic shift in the emission maxima ($h\nu_2'$, $h\nu_2''$ and $h\nu_2'''$, **Figure 1.12**).

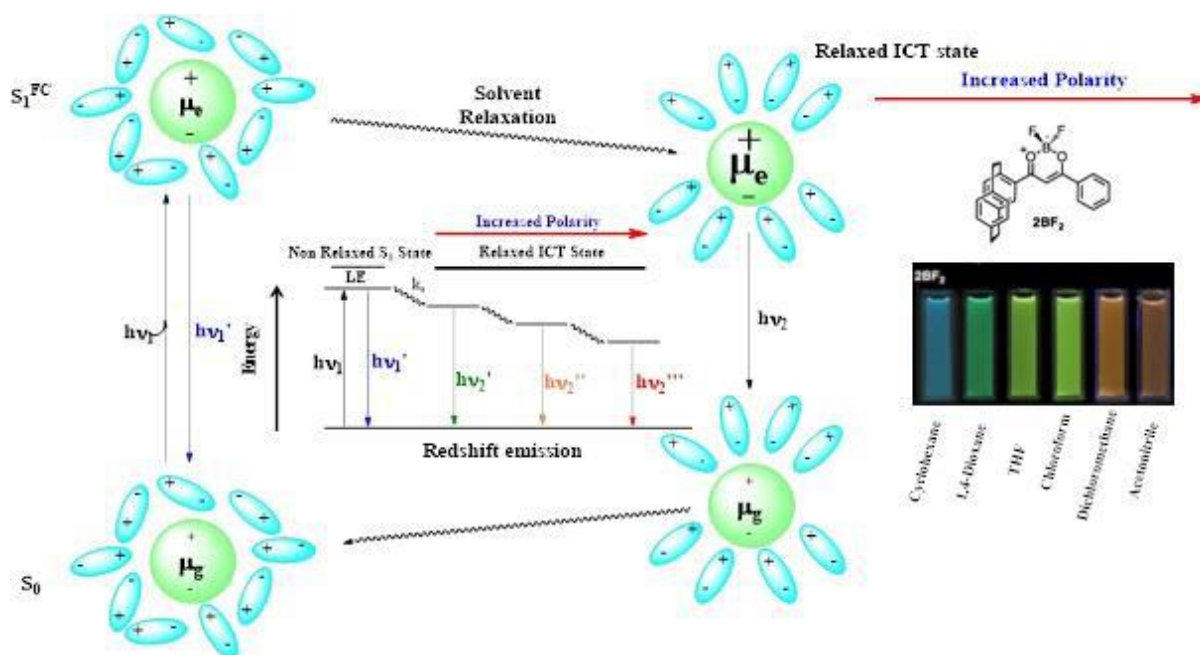


Figure 1.12. Effect of solvent polarity and dipole orientation on the stabilization of CT excited states. The positive solvatofluorochromic effect of **2BF₂** is also highlighted.

The ICT state can be evidenced by performing a photophysical study in solvents with distinct polarity. Several models have been developed.⁴¹⁻⁴³ Among them, the Lippert-Mataga equation (**Equation 1.8**)^{41,42} is one of the mostly used approach to evaluate the general solvent effect on the absorption and emission spectra. Moreover, it allows the estimation of the dipole moment differences between the excited and ground state of organic fluorophores.

$$\nu_{abs} - \nu_{em} = C + \frac{2\Delta f}{hc\alpha^3} (\mu_e - \mu_g)^2, \text{ Equation (1.8)}$$

$$\Delta f = \left(\frac{\epsilon-1}{2\epsilon+1} \right) - \frac{1}{2} \left(\frac{\eta^2-1}{2\eta^2+1} \right), \text{ Equation (1.9)}$$

Here, the ν_{abs} and ν_{em} are the absorption and emission maxima (in wavenumber), α is the radius of the Onsager's cavity (in cm), h is the Planck constant, c is the speed of light, μ_e and μ_g are the dipole moments of the excited and ground states and Δf is the Lippert parameter

(solvent orientation polarizability). This last parameter is particular for each solvent/mixture and accounts for two macroscopic parameters: the refractive index (η) and the dielectric constant (ϵ) (**Equation 1.9**) As the ICT band can be stabilized by increasing the solvent polarity, the observation of a linear relationship for the plot $(\nu_{\text{abs}} - \nu_{\text{em}})$ vs Δf or ν_{em} vs Δf is frequently used as an indicative for the formation of a charge transfer state.

1.5.4 The Role of CT States on the Radiative and Non-Radiative Decays

With some exceptions, one common observation resulting from the intramolecular charge transfer is the reduced fluorescence emission intensity in high polar media, especially in water.⁴⁴ The reduced energy gap between excited and ground state in polar environment can favors the $S_1 \rightarrow S_0$ internal conversion instead fluorescence emission, thus increasing the rate of non-radiative decay.⁴⁵ This consequence is more important when strong EDG and EWG are present or when the dipole moment differences are large.⁴⁶ A prominent quenching effect is also observed when the charge transfer state involves molecular torsion (*i.e.* TICT). In this case, the conjugation loss offers an additional non-radiative tunnel to release energy. Thus, fluorescence emission becomes significantly less probable to occur as the molecular geometry of the ground and excited states are not similar.^{37,47}

1.5.5 Viscosity Effects on the Excited States (LE, ICT and TICT) Interplay

The rates of solvent relaxation and the molecular motion are strongly dependent on the temperature and viscosity. At reduced temperatures the solvent may become more viscous, hence the time for solvent reorientation is longer.⁴ Additionally, conformational transitions become slower as the viscosity enhances. In molecules in which a twisted motion is required to the formation of charge transfer state (TICT) an increase in the solvent viscosity can reduce or inhibit the access to such polar state.³⁹ Interestingly, the suppression of the TICT state emission does not necessarily results in the enhancement of the LE band intensity (**Figure 1.13-A**, compound **23**).⁴⁸ Viscosity can also modulate the TICT \leftrightarrow ICT conversion.⁴⁹ As a general trend, an increase in the viscosity in such kind of system tends to enlarge the energy barrier to access the non-emissive twisted geometry. Therefore, the planar ICT become the major state and a large enhancement on the Φ_{fl} appears (**Figure 1.13-B**, compound **24**).

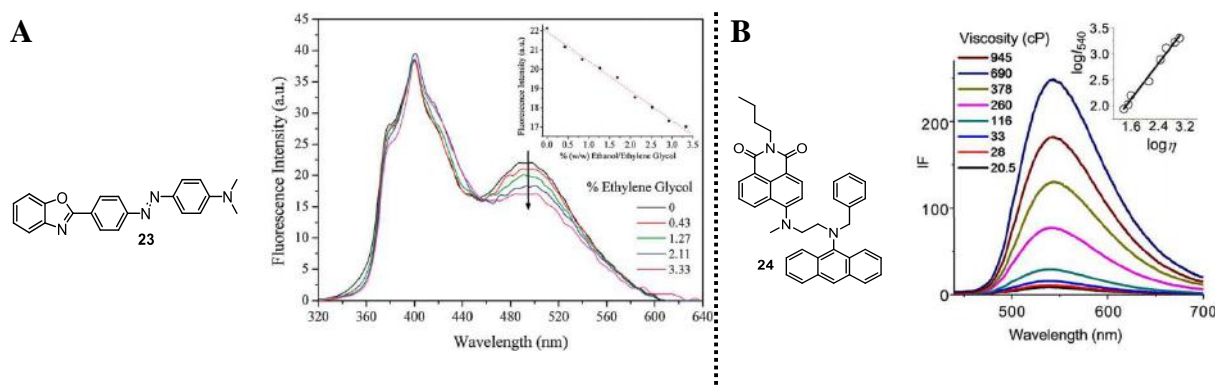


Figure 1.13. Viscosity effect on the LE-ICT-TICT emission bands.

1.6 Photoinduced Electron Transfer (PET)

As well as for the great majority of the photophysical processes, the first singlet excited state S_1 is also the main state from which competitive photochemical processes will occur.⁶ Looking at the electronic configuration of ground and excited state molecules, one can intuitively infer that the redox properties of both states are fairly different. Basically, a molecule becomes a better electron donating scaffold (a reducing agent) in the S_1 state than in the ground state. Intuitively, the vacancy created in the HOMO orbital after excitation also makes the excited species a more powerful oxidizing agent (electron acceptor scaffold). Thereby, it is expected that the processes involving electron transfer may be the primary photochemical pathway that can occur upon light absorption.⁶

In PET, an electron is transferred between an excited species and another molecule that lies in the ground state. If we consider a fluorophore in the excited state in presence of second molecule, two different electron transfer mechanisms may occur (**Figure 1.14**). If the HOMO and LUMO orbitals of a molecule are of higher energy (electron donor) than the frontier orbitals of the excited fluorophore, then the fluorophore will be reduced (reductive PET). By contrast, if the LUMO orbital of the excited fluorophore displays higher energy, the fluorophore may undergoes oxidation by transferring an electron to the LUMO orbital of an electron acceptor molecule (oxidative PET).¹

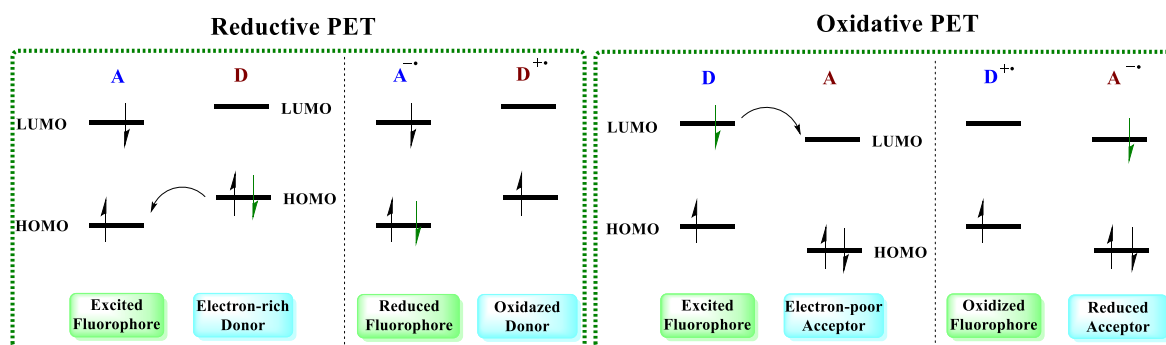


Figure 1.14. Representation of reductive and oxidative PET processes.

Electron transfer is a process that occurs through electron exchange interactions, requiring effective orbital overlap.⁴ So, short distances are required and diffusion is a key process to afford high PET efficiency. Thus, intermolecular electron transfer is normally less efficient and requires high concentration of the corresponding ground state molecule (quencher) to inhibit the fluorophore emission.⁵⁰ Of note, for the bimolecular system (**25** and MV^{+2}) highlighted in **Figure 1.15-A**, the required concentration of methylviologen (MV^{+2}) to suppress in 50% the fluorescence intensity of the naphthalimide fluorophore (**25**) ranges in $\approx 0.1M$. This large quencher concentration can be avoided by the introduction of a covalent linkage between the corresponding donor-acceptor fluorophore/quencher pair. In this case, the PET efficiency will be directly related to the distance between the donor and acceptor moieties. This is clearly illustrated in **Figure 1.15-B** (see compounds **26** vs **27**).

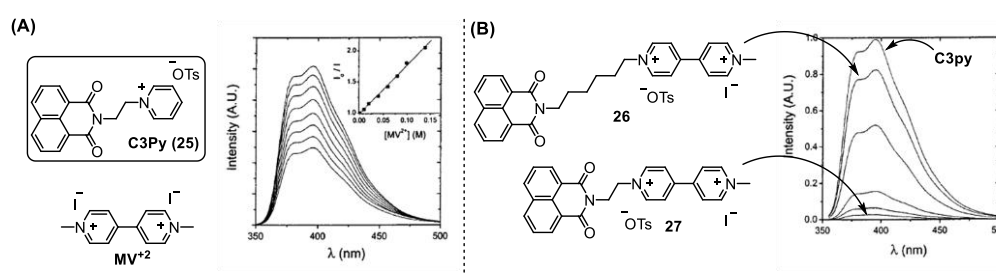


Figure 1.15. Comparison between intramolecular and intermolecular PET efficiencies.

A close look at the above electron transfer scheme clearly shows an important feature of the PET process. No matter the direction of electron transfer, PET frequently results in fluorescence suppression and the system relaxes *via* non-radiative decays. Some exceptions include the formation of charge transfer complexes (*e.g.* exciplexes) that may undergo fluorescence emission. However, they will not be discussed herein. Overall, photoinduced electron transfer has been intensively exploited in sensing applications and (bio)technology to access a great

diversity of light up fluorescent probes.^{51–53} This effect will be discussed in more details in *Section 1.8.2.3*.

1.7 Fluorescence Labeling

The labeling of biomolecules, such as nucleic acids and proteins with extrinsic organic dyes has emerged as a powerful tool to study the affinity, dynamics and the environmental changes of these macromolecules upon interaction with different substrates. In particular, due to the lack of optimal intrinsic fluorescence emission properties, such as the low Φ_{fl} and emission in the UV region, the conjugation of biomolecules with small fluorophores affords new ways to probe and investigate such kind of molecular systems. Typically, no matter the labeled substrate, the selected fluorophore should always be easily incorporated and it must not modify the natural structure of the biomolecule. Moreover, it should be preferentially attached close to the region of interest and show an excitation wavelength outside the range of the biomolecule absorption.

1.8 Fluorescence Sensing

Fluorescence based techniques encompass a broad group of versatile tools for several applications. They can provide real-time, rapid, sensitive, and specific detection of wide range of analytes, enzymes, environmental changes and biological processes and have been successfully employed as sensing, imaging and therapeutic tools.⁵⁴

Different fluorescence emission parameters can be recorded and used in sensing applications. The fluorescence intensity F ($F(\lambda)$), anisotropy and lifetime can be measured separately or simultaneously to access a broad scope of different signaling responses (**Figure 1.16**).^{4,54} Among all these fluorescence methods, approaches based on single changes in the fluorescence intensity are the most common and easily implemented in laboratories, and adaptable to complementary techniques such as fluorescence-based microscopy.

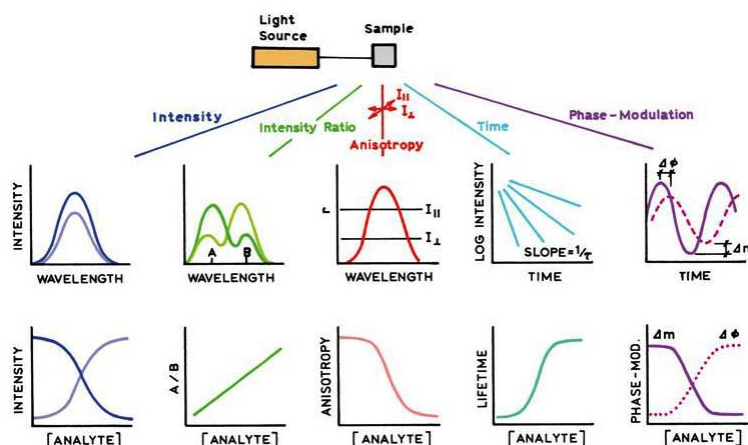


Figure 1.16. Fluorescence-based techniques.⁴

Fluorescent probes can be designed to present single or multiple emission bands. Single band probes can be used to produce systems with changeable light (ON) and dark (OFF) states which are easily observed and recorded. Systems with light-up responses (OFF-ON probes) constitute the main scope of single band probes. Depending on the Φ_{fl} enhancement in the ON state, these probes can afford very high sensitivity and display readily changeable fluorescence signaling in presence of very low level of the correspondent target analyte. However, as fluorescence intensity is apparatus-dependent, quantitative analysis can be really compromised. In contrast, systems displaying two or more emission bands can benefit from λ -ratiometry, *i.e.*, the signal can be collected as the intensity ratio between two or more emission bands. Thus, the sensing signal is only related to the analyte concentration, which is absolutely valuable for quantitative analysis.⁵⁴

No matter the field of application (chemical, enzymatic or environmental sensing, cellular tracking), the fluorescent reporter should always feature optimal spectroscopic properties such as high brightness, photo and chemo stability, optimal excitation and emission wavelengths, large Stokes shift to avoid self-quenching effects and adequate solubility and/or permeability in the probing environment.⁵⁵ In this section, a general overview presenting some of the approaches employed to design operational light-up and ratiometric probes will be described. The central discussion will be settled around intensity-based fluorescent probes applied to report a specific enzymatic activity. Hereafter, the sensing of β -galactosidase (β -gal) activity will be used as the model enzyme target, establishing a bridge with the case of study that will be presented in **Chapter 3** (*Development of highly sensitive fluorescent probes for the detection of β -galactosidase activity – application to the real-time monitoring of senescence in live cells*). The following subsection, prior to the bibliographic review on β -gal

probes, we will briefly state the concept of self-immolative linkers. They will be important part of the trigger unit in a wide range of architected sensors that will be presented in the sequence, as well as in the probes that will be reported by us in **Chapter 3**.

1.8.1 Self-Immolative Linkers

Self-immolative linkers are a special group of chemical spacers that have gained popularity in recent years. Usually, these group of spacers have a polarized system constituted by chemically stable bonds (functional groups), which are used to assembly a general target molecule (*e.g.* a fluorophore or a drug) with an specific reactive unit. Primarily, self-immolative linkers are stable under a variety of chemical conditions; however, they undergo a fast disassembly to produce the naked target molecule, in a thermodynamic and entropic driven process that are triggered by a chemical reaction with the reactive unit. Self-immolative linkers have been extensively used in the context of medicinal chemistry (drug release systems)⁵⁶ and fluorescence sensing applications.⁵⁷⁻⁵⁹ Nowadays, many drugs are administrated not in its pharmacologically active form but as a prodrug, which is inactive per se. Ideal prodrugs remains as it until they reach their site of action. There, they are metabolized (processed by enzymes) to regenerate the pharmacologically active compound. This drug-prodrug systems have been elaborated to improve the pharmacokinetics properties of the active drugs as well as their bioavailability and selective distribution within the body.⁶⁰ Based on the same idea, numerous fluorescence sensors have been designed by the assembly of a specific reactive unit (trigger motifs) with fluorescent reporters through self-immolative linkages. Essentially, this "pro-fluorophore" should feature distinct fluorescence properties (emission intensity, emission wavelength, anisotropy...) related to its naked form. Therefore, upon processing of the trigger unit, the original optical properties of the fluorophore will be restored. Especially, the use of enzyme-labile self-immolative linkers arose as powerful tools applied to the fluorescence sensing of enzymatic activity.^{61,62} Since enzyme-catalyzed reactions usually take place under mild conditions, (physiological temperature, pH 4.0 - 7.4), highly specific sensors with fast and sensible responses have been designed by using self-immolative linkers modified with an appropriate enzyme substrate. Therefore, this feature has stimulated intense research to develop novel and efficient self-immolative linkers featuring optimal kinetics and thermodynamics properties concerning their disassembly.⁵⁶ The widely used approach in fluorescence sensing is centered on linkers that trigger non-amplified self-immolative elimination, *i.e.* just a single fluorophore molecule is released upon

the dissociative activation. In **Figure 1.17-A** are highlighted some of the potential linkers already used or suitable to be employed in the preparation of molecular fluorosensors.⁶³ **Figure 1.17-B** depicts the mechanism of self disassembly of linkers 2 and 4 (**Figure 1.17-A**) considering a β -gal-mediated hydrolysis of β -galactoside units. Whatever the linker, the target molecule can be released upon enzymatic processing (hydrolysis, oxidation, reduction, nucleophilic attack) of the reactive/trigger unit. Thus, a specific and irreversible cascade decomposition is triggered, generating the free target molecule and small sized byproducts as well.

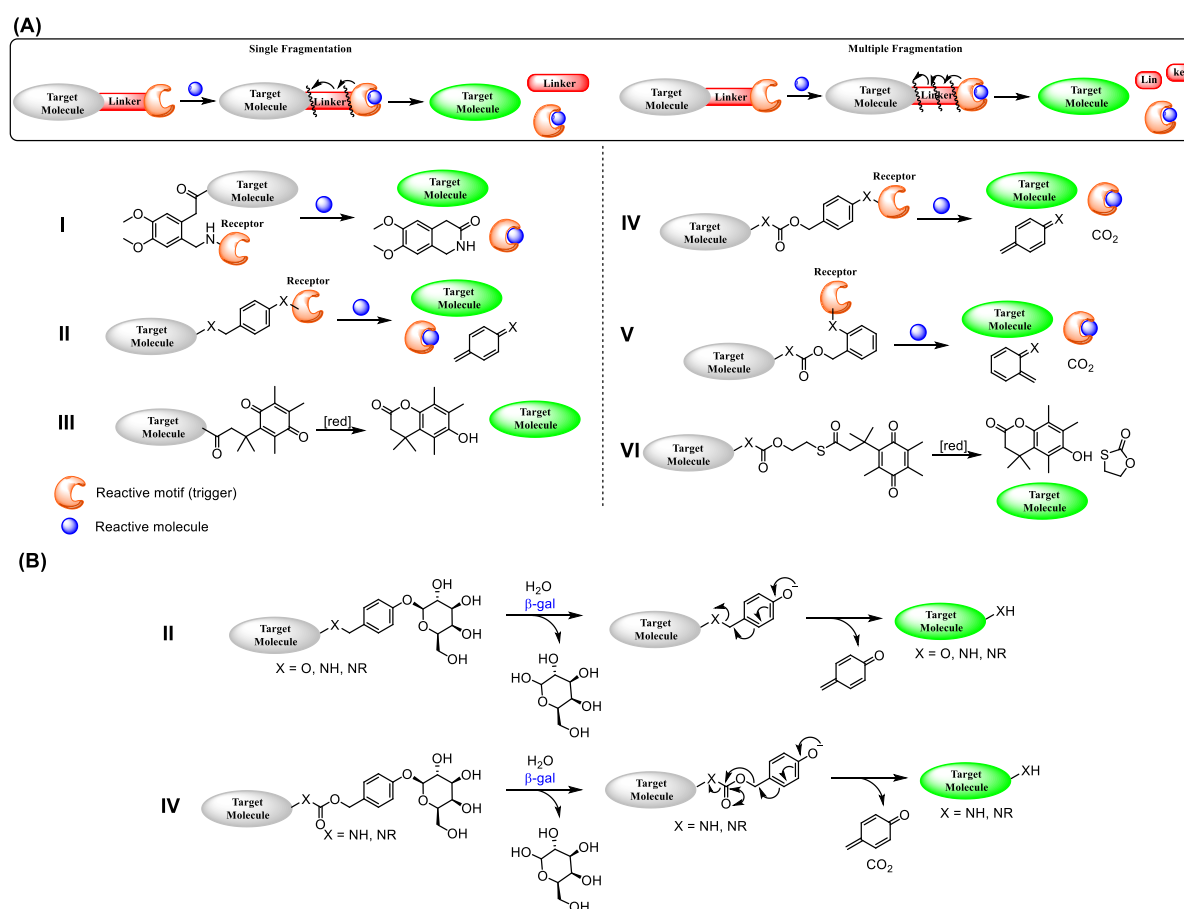


Figure 1.17. (A) Structure and fragmentation of potential self-immolative linkers. (B) Example of β -gal-mediated cleavage and the disassembly mechanism concerning the self-immolative linkers 2 and 4.

1.8.2 Sensing enzymatic activity - Intensity-Based Probes

In this section we will give an overview of the methods that can be used to design probes for enzymatic activity sensing. In particular, when reported in the literature, we will illustrate our

discussion with examples of β -gal activity sensing. Nowadays different approaches have been used to access operational probes with OFF-ON and ratiometric responses to monitor enzymatic activity. No matter how simple or complex is the molecular architecture of the probe, it always presents both: a signaling motif (*i.e.* a fluorophore) and an enzyme substrate which act as a trigger to the spectroscopic signal upon its enzymatic processing. The most simple and predictable method involves the use of probes with changeable π -conjugated systems (CCS)⁶⁴ prior and after the enzymatic recognition. Both light-up and ratiometric systems can be designed by controlling the extension of the π -conjugated core. More sophisticated systems have been constructed by taking advantage of bioluminescence and photophysical processes involved in energy transfer and fluorescence quenching or enhancement. PET, Aggregation Induced Emission (AIE), and bioluminescence-based probes have been extensively described in the context of light-up sensors. Fluorescence or Bioluminescence Resonance Energy Transfer (FRET and BRET) and excimer-based systems have been exploited for both cases (single and dual emissive probes). Hereafter, examples of systems to target β -gal activity reported in the literature will be presented; for a practical overview, these probes are gathered according to the different mechanisms of response highlighted above (CCS, ICT, PET, AIE, bioluminescence, excimer and FRET).

1.8.2.1 CCS-Based Probes

Turn-On fluorescent probes can be simply obtained by controlling the chemical structure of the probe to afford distinct photophysical properties prior/after the enzymatic processing. The enzymatic reaction with the probe induces a change in the π -conjugated system (CCS)⁶⁴ which modulates the spectroscopic properties of the probe (λ_{abs} , λ_{em} , Stokes shift and Φ_{fl}). Thereby, CCS mechanism has been used as an efficient tool to access excellent probes with highly predictable properties. The most common method involves the chemical transformation of a fluorophore with a receptor unit (enzyme substrate) to restrains the electronic conjugation of the system. Thus, when the probe exists in this low extended π -conjugated state ("pro-fluorophore"), it usually has no or weak spectroscopic signal (absorption or fluorescence emission) in a relatively long wavelength region. However, upon enzyme processing, the π -conjugated system extends to a larger scale, thus generating a spectroscopic signal in the long wavelength region. The two main approaches based on CCS

resort on: ring-closing reactions and nucleophilic addition; in this manuscript, we will presented below only the former case (ring-closing reactions).

Fluorescein, rhodamine and rhodol dyes encompass an important group of classical and highly water emissive fluorophores exploited in light-up molecular probes (**Figure 1.18**). They can be easily converted to non-emissive spirolactone analogues by the modification of their terminal hydroxyl and amino groups with a specific reactive moiety. **FDG**, **29** and **30** have been successfully employed as tools in biomedical research and cellular biology to report β -galactosidase activity. It is noteworthy that such probes display an outstandingly high S/N ratio.^{61,65–68} Another important example of xanthene-based β -galactosidase probe is the selenium substituted rhodol analogue **31**.⁶⁹ Interestingly, this probe does not change from a non-emissive to an emissive state. It changes from a non-toxic to a phototoxic cyclic form after β -gal processing. The selenium atom favors the population of triplet excited states thus acting as a photosensitizer to generate singlet oxygen. Thus, this probe has been successfully used as an activatable photosensitizer to selectively produce singlet oxygen species only in cells in which β -galactosidase is expressed.

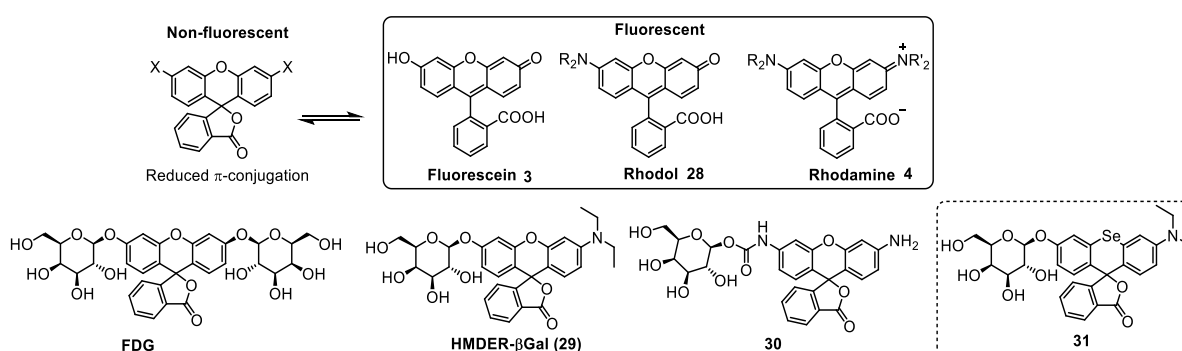


Figure 1.18. β -gal probes based on ring-closing reactions.

FDG, reported by Rotman et al.,⁶⁷ is a non-fluorescent and highly sensitive substrate for β -galactosidase. It is effective to stain cells that overexpress β -galactosidase and has been largely used in fluorescence microscopy. In the presence of β -galactosidase it undergoes a dramatic enhancement in the fluorescence emission intensity which can be used to the signaling of enzymatic activity. However, **FDG** shows poor kinetic performance due to the requirement for sequential hydrolysis of both galactopyranoside moieties and requires a strict control of the intracellular pH before staining.

The use of rhodol derivatives for the β -gal activity sensing was introduced by Nagano's groups in 2011. They identified that rhodol analogues bearing hydroxymethyl group instead of the classical carboxylate moiety are robust candidates to produce probes for different hydrolases. **HMDER- β Gal (29)**,⁶⁶ belongs to the first generation of rhodol-based β -galactosidase probes described by Nagano. It is a rhodol analogue functionalized with a β -galactopyranoside unit on the phenolic hydroxy group. At physiological pH, **HMDER- β Gal (29)** exists as its spirocyclic structure and is very weakly fluorescent ($\Phi_{fl} = 0.009$). However, upon the cleavage of the saccharide moiety, the spirocycle opens, hence leading to an extension of conjugation (CSS). This directly affects the fluorescence properties of **HMDER** as it triggers a 76-fold enhancement of the fluorescence emission intensity ($\Phi_{fl} = 0.141$). **HMDER- β Gal (29)** also proved effective in the visualization of β -gal activity in live HEK/LacZ cells. More recently, Asanuma and co-workers⁶⁵ reported a second generation of rhodol probes. After screening few amino groups connected to the xanthene scaffold, they identified **HMRef- β Gal (32)** (**Figure 1.19**) as an optimized β -gal reporter displaying 1420-fold enhancement on its fluorescence emission after enzymatic processing and remarkable kinetics compared to **HMDER** analogue. **HMRef- β Gal (32)** is a robust probe to visualize β -gal activity in several ovarian cancer cell lines. Moreover, its use in imaging diagnostics in vivo to detect metastatic cancer via fluorescence endoscopy has also been validated.

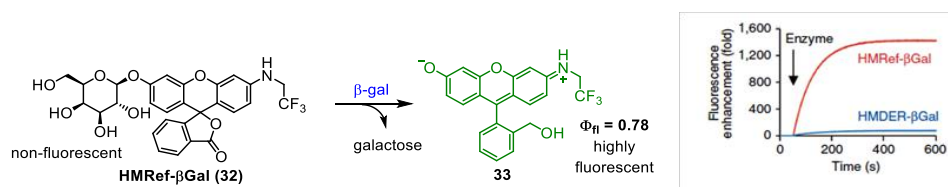


Figure 1.19. Mechanism of β -gal detection using probe **HMRef- β Gal (32)**. (figure copied from Asanuma et al.⁶⁵)

In 2013, Tung's group reported a distinct cyanine-like based probe to detect β -gal in cells (**Figure 1.20**).^{70,71} The probe **34**, a dicationic cyanine analogue, displays an aromatic π -conjugated spacer functionalized with a self-immolative 4- β -D-galactopyranosyl benzyl ether (trigger unit); this aryl spacer is linked to two sulfo-indoline moieties, one at each side. Upon enzymatic processing, a tremendous change on the molecular conjugation takes place, which redshifts, in more than 100 nm, the absorption and the emission spectra of the probe (**Figure 1.20**, conversion of **34** into **35**). Moreover, the probe displays an emission in the NIR region

upon excitation at 570 nm. Of note, the β -gal processing induces a 110-fold enhancement on the fluorescence intensity. Interestingly, this probe seems to display a distinct mechanism within the cells, rendering a different optical response. The low electron density of the polymethine chain enables a fast nucleophilic attack of thiols (e.g. glutathion, cystein...). This in turn triggers a retro-aldol type degradation of the probe hence leading to a complete degradation of the NIR reported to form a simple merocyanine fluorophore (**38**); the emission maxima of **38** is blueshifted compared with **35**.

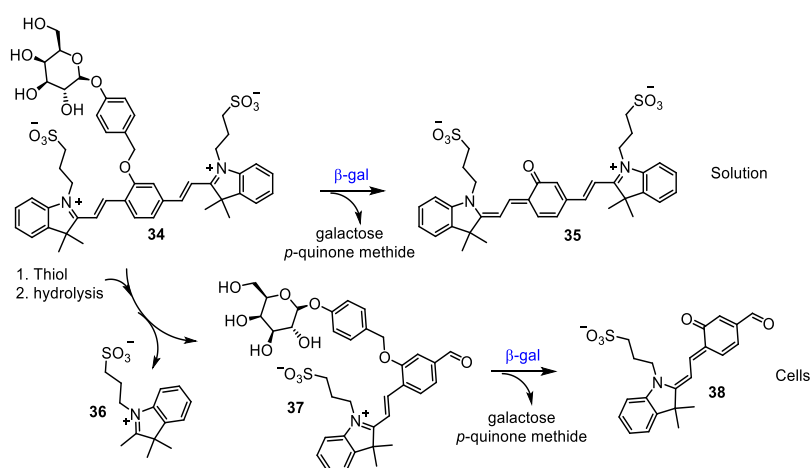


Figure 1.20. β -gal probe **34** reported by the Tung's groups.

Fluorosensors can also be developed by modifying functional groups attached to the extremities of organic fluorophores; of note these groups are usually the electron-donor or -acceptor motifs. Indeed, the acylation of electron donating moieties (OH, NH₂, NHR) can be used to modulate the extension of conjugation (or intramolecular charge transfer) by reducing the electron donating ability of these groups. Many light-up and ratiometric probes, with lots of different molecular designs and fluorescent reporters have been described by exploiting this approach.

Nagano's group also reported a novel sila-fluorescein presenting a modified phenyl group which limits the formation of spirocyclic form **40**.⁷² This sila-fluorescein behave differently from its parent compound as it displayed an impressive redshift in the absorption maxima when in its anionic form in comparison with its neutral structure. The derivatization of this novel fluorescein analogue with a β -D-galactose unit enabled the assessment of a novel light-up probe **39** with high brightness to detect β -gal activity in HEK293/LacZ cells with noteworthy imaging contrast (**Figure 1.21**).⁷²

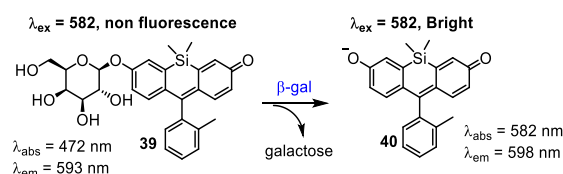


Figure 1.21. Sila-fluorescein based β -Gal probe **39** exploiting CCS approach: functional group modifications.

Using a similar approach (modification of end sites of the probe), Zhang and co-workers⁷³ reported a light-up NIR fluorescent probe **41** to detect endogenous senescence associated β -galactosidase activity in a model senescent cells (**Figure 1.22**). Apart from the optimal spectroscopic range for excitation and emission (> 680 nm) and the ability to detect β -gal concentrations as low as 0.1 nM, the probe displayed only 12-fold enhancement after enzymatic processing (conversion of compound **41** into **42**, **Figure 1.22**).

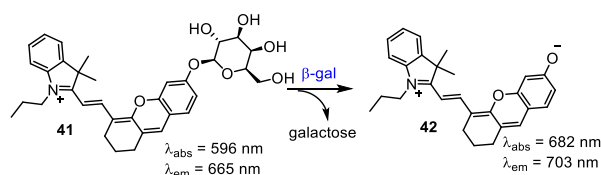


Figure 1.22. β -Gal probe **41** based on the CCS approach: functional group modifications.

In 2004, exploiting the 7-hydroxy-9*H*-(1,3-dichloro-9,9-dimethylacridin-2-one (DDAO) fluorophore (**44**), Tung⁷⁴ and collaborators demonstrated that the known probe **43**⁷⁵ could be an useful fluorescent sensor for *in vivo* imaging of β -galactosidase activity (**Figure 1.23**). **44** displays a far red fluorescence emission signal. Light-up response was obtained when the probe was selectively excited in the NIR region (λ_{abs} **43** = 465 vs 646 for **44**). Moreover, using red fluorescence real-time imaging, the β -gal activity was readily detectable in 9L glioma cells transfected with LacZ gene (β -gal) in comparison with the native 9L glioma cells.

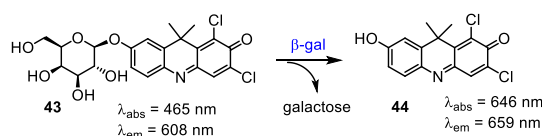


Figure 1.23. β -Gal probe **43** based on the modulation of π -conjugated system: functional group modifications.

1.8.2.2 ICT-Based Probes

Nowadays, with the growing interest on push-pull dyes, especially due to their non linear optical properties, the great majority of the reported probes used to detect and visualize β -galactosidase activity have been designed exploiting different fluorescent dyes displaying intramolecular charge transfer excited states.

Recently, two groups reported the use of naphthalimide fluorophores to report β -gal activity (**Figure 1.24**). Zhang⁷⁶ described the **NI- β Gal (45)**, which is a two photons absorption probe, to visualize *E. coli* β -galactosidase activity *in cellulo* (live glioblastoma C6 cells) and *in vivo* (U-87 MG tumors in mice), both transfected with LacZ gene (β -gal). The probe displayed a 680-fold enhancement of the fluorescence intensity ratio, between the yellow and blue emission, after enzymatic processing (conversion of **45** into **46**, **Figure 1.24-A**). Moreover, compared with non-transfected tumor cells, a 10-fold higher fluorescence signal was observed in mice imaging in U-87 MG glioblastoma tumors that were beforehand transfected with Lacz gene (β -gal). In a different work, Huang and co-authors⁷⁷ reported the lysosome-targeting **FC- β Gal (47)** probe to detect β -gal activity in ovarian SKOV-3 cells *via* light-up or dual excitation ratiometric imaging. **47** displays a detection limit as low as 4.0×10^{-5} U/mL. Glycosidic bond cleavage of **47** triggers a 1000-fold enhancement of the fluorescence intensity signal in the yellow spectral region (conversion of **47** into **48**, **Figure 1.24-B**). Nevertheless, apart from the versatility and the facilities to chemically modify the naphthalimide scaffold, these fluorescent dyes show only low to moderate brightness due to the reduced extinction coefficient (ϵ) values (10000 - 15000 (M.cm)⁻¹).

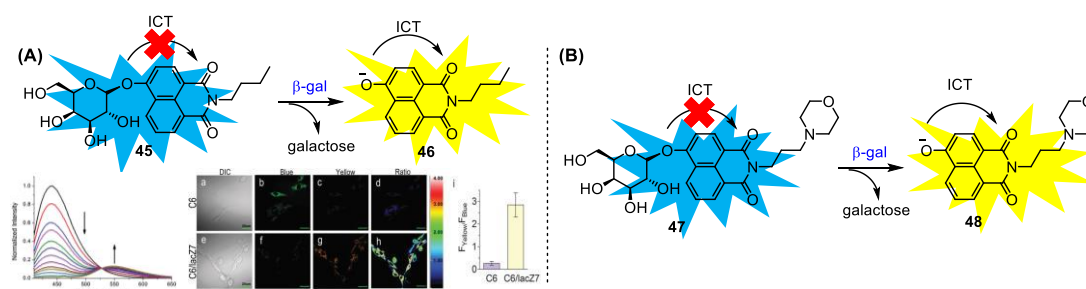


Figure 1.24. β -galactosidase probes (**Ni- β Gal (45)** and **FC- β Gal (47)**) based on naphthalimide push-pull fluorophores with ratiometric responses. (Images were copied from Zhang *et al.*⁵⁹).

The use of push-pull fluorophores displaying emission in the far red-NIR spectral region have also been described in the recent years, thus allowing the non-invasive *in vivo* visualization of β -gal activity (**Figure 1.25**). Gu and co-workers⁷⁸ designed the probe **49** to detect such activity by assembling a β -D-galactose substrate directly on the phenolic hydroxyl group of a dicyanomethylene-4H-pyran (DCM) chromophore (**50**). The probe displays a simultaneous turn-ON and ratiometric signaling in response to β -galactosidase activity (conversion of **49** into **50**, **Figure 1.25**), however with relatively low fluorescence enhancement (14-fold and 34-fold considering the ratiometric $I_{\text{red}}/I_{\text{green}}$ and turn-ON I_{red} modes, respectively). Nevertheless, probe **49** could report β -gal activity in ovarian cancer cells (OVCAR-3 cells) and HEK293T cells transfected with LacZ gene. In addition, it proved to be an excellent tool for the real-time *in vivo* (in mice) high-resolution three-dimensional bioimaging of β -gal activity in colorectal tumors.

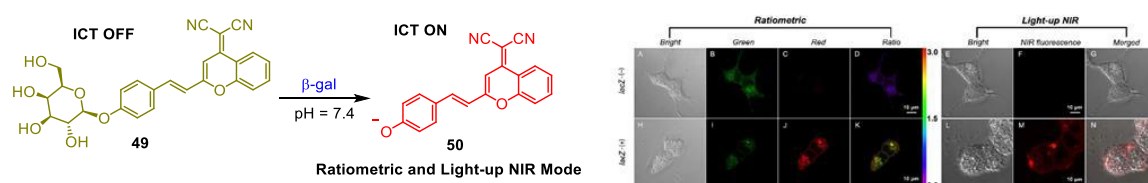


Figure 1.25. NIR ICT-based probe **49** to report β -galactosidase activity *in vivo*. (images were copied from Gu *et al.*⁷⁸).

Using a parent NIR fluorophore, Kim's group reported a ratiometric system applied to the *in vivo* imaging of stimulated β -gal activity in hepatocellular carcinoma HepG2 cells and in xenograft mouse models.⁷⁹ This probe, **51**, displays a NIR ratiometric (I_{665}/I_{615}) signaling when excited at 580 nm in response to β -galactosidase processing with low background interference and no pH dependence on the emitted fluorescent signal (**Figure 1.26**).

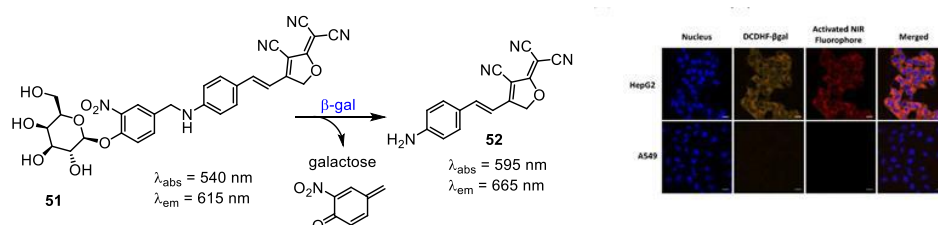


Figure 1.26. NIR ICT-based probe **51** to report β -galactosidase activity *in vivo*. (images were obtained from Kim *et al.*⁷⁹).

Using the push-pull 6-(benzo[*d*]thiazol-2'-yl)-2-(methylamino)-naphthalene fluorophore (**54**) displaying significant two-photon excitation, the group of Kim⁸⁰ reported the ratiometric probe **53** to detect and visualize senescence-associated β -Gal activity in live cells and aged tissues (**Figure 1.27**). The reported fluorophore was assembled to a galactoside sugar, through an electron deficient benzylcarbamate self-immolative linker. With the presence of the carbamate linkage, **53** displays a blue emission when excited at 376 nm. However, after β -gal processing, the naked amino group that is restored increases the ICT character and redshifts the emission wavelength to the yellow spectral region. The ratio of emission intensities is increased by 120-fold after enzymatic processing. Probe **53** undergoes fast β -gal mediated hydrolysis and can detect β -gal levels as low as 0.25 nM. Moreover, this probe was effective to monitor middle stages of cellular senescence in human diploid fibroblast (HDFs) senescent model cells.

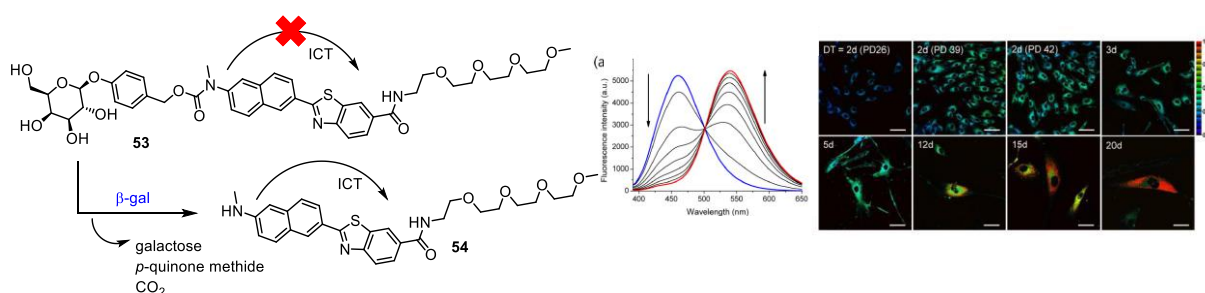


Figure 1.27. Ratiometric probe **53** reported by Kim's group. (Images were copied from Lee *et al.*⁸⁰).

1.8.2.3 PET-Based Probes

Light-up fluorescent probes based on PET systems became popular during the last decades. Molecular systems bearing PET donors and/or acceptors has enable the production of a large library of versatile, selective and effective probes for sensing numerous chemicals. Heretofore, sensors based on the reductive electron transfer process encompass the majority of PET-based probes. This includes a large number of fluorescent reporters for metal cations,⁸¹ hydrogen peroxide,⁸² methylglyoxal,⁸³ hypochloric acid,⁸⁴ enzymes⁶⁸ (enzymatic processes) and many other analytes. Common receptors are electron-rich aromatic rings (phenols, anisoles, anilines), aliphatic amines and chelating groups. In this case, the electron donating ability of the receptor can be disrupted by stabilizing its HOMO orbital *via* metal/proton complexation or by a specific chemical modification. Thus, the electron transfer

process will be blocked and the fluorescence emission will be restored. By contrast, probes based on oxidative PET normally use electron deficient groups as quenchers, including 2,4-(dinitrophenyl)sulphonamide (DNBS),⁵⁷ maleimide,⁸⁵ quinones⁸⁶, pyridines and certain organic cations as imidazolium, and methylviologene (paraquat).

Heretofore, only two probes based on PET mechanism have been described to report and visualize β -galactosidase activity (compounds **55** and **57**, **Figure 1.28**). A third probe, **58**, has been developed analogously, but used as a selective photosensitizer for photodynamic therapy (PDT).⁸⁷ Based on the 2-Me TokyoGreen fluorophore (**56**), Nagano's group reported two powerful probes based on reductive PET process. **55**⁸⁸ and **57**,⁸⁹ are fluorescein analogues bearing 2-methyl-4-alkoxy phenylgroups; both bears a single galactose unit directly connected to the hydroxy group of the fluorescein core through a glycosidic bond. Differently from classical fluorescein dyes, **55** does not adopt the non fluorescent spirocyclic structure. It displays an extremely weak fluorescence emission (quenching *via* reductive PET) in the absence of β -galactosidase. However, upon enzymatic processing, a dramatic enhancement on the fluorescence intensity is observed (420-fold, $\Phi_{fl} = 0.84$). Conversely to other PET-based systems, in this case, the disruption of the PET process is due to a change in the electronic properties of the acceptor fluorescein. In addition, **55** displays much faster kinetics compared with the fluorescein analogue **FDG**. Moreover, **57**, a second generation of Tokyo Green β -gal reporter showed impressive results in the visualization of β -galactosidase activity in cells and *in vivo*. The probe features the same photophysical profile as **55**. However, the introduction an esterase substrate enabled the formation of a hydrophilic product, compared with **55**, which can be well retained in cells, even after washing. This improvement allowed for the *in vivo* visualization of intraperitoneal tumors that express β -gal.

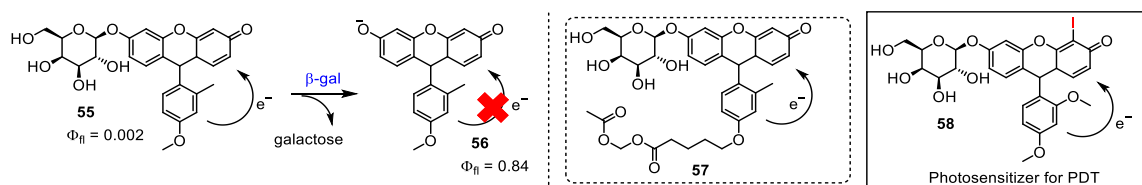


Figure 1.28. Reductive PET-based probes to report β -gal activity in cancer cells.

1.8.2.4 AIE-Based Probes

Aggregation induced emission (AIE) systems have been largely employed as a tool for chemo/biological sensing. AIE luminogens (AIEgens) display negligible fluorescence emission in solution but they strongly fluoresce in the aggregate/solid state as a result of the restricted molecular motion that inhibits non-radiative decays.⁹⁰ Usually, probes based on AIE can be obtained by conjugating hydrophilic moieties to AIEgens (with or without a linker) to afford water soluble conjugates; this, in turn, results in the suppression of the fluorescence emission (OFF state). Thus, after processing of the receptor (*e.g.* hydrolysis of a glycosidic bond), the hydrophobic AIEgen is released and aggregates, hence leading to the enhancement of the fluorescence signal (ON state). The first generation of probes were developed based on electrostatic interactions. Nowadays, AIE probes are designed by introducing a diversity of recognition moieties including sugars, oligonucleotides and peptides.⁹¹

Recently, AIE sensors have been reported to detect β -galactosidase activity exploiting tetraphenylethylene (TPE) AIEgens as well as other examples of solid-state emissive fluorophores resorting to an ESIPT mechanism (**Figure 1.29**).^{59,92,93} In 2015, Liu's group⁹³ reported the first AIE-based β -galactosidase probe exploiting the bis- β -galactopyranoside-modified salicylaldehyde azine fluorophore **SA- β Gal (59)**. However, apart from its very high enhancement of fluorescence emission intensity (820-fold), the probe displays relatively slow kinetics and low to moderate limit of detection (0.014 U/mL). In another work, a TPE-based probe with improved kinetic performance was designed by Jiang and co-workers.⁵⁹ The reported probe, **61**, presents a single galactoside unit, which is linked to a modified TPE AIEgen through a self-immolative benzyl linker to afford a fast β -galactosidase processing. The probe proved efficient in the detection of β -gal activity in ovarian OVCAR-3 cancer cells and, as expected, displayed faster kinetics than the previous AIE-based probe reported by Liu's group.

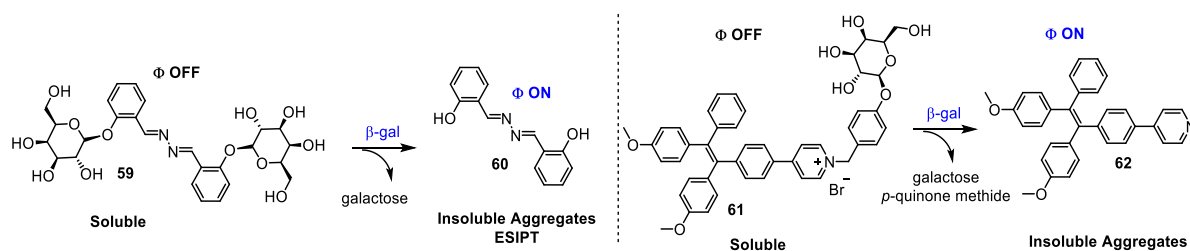


Figure 1.29. AIE-based light-up probes to target β -galactosidase activity.

1.8.2.5 Bioluminescence-Based Probes

The use of bioluminescence has emerged as an important tool in fluorescence sensing and imaging of biological systems. Bioluminescent species are able to emit light by the luciferase-catalyzed oxidation of small molecules, known as luciferin substrates (**Figure 1.30**). The major advantage of bioluminescent probes is that there are no requirements for sources of excitation (as lasers) as light is produced by a biochemical reaction. Thus, no background emission due to cellular autofluorescence should be observed. However, this technique suffers from a relatively weak luminescent signal and very poor spatial resolution.

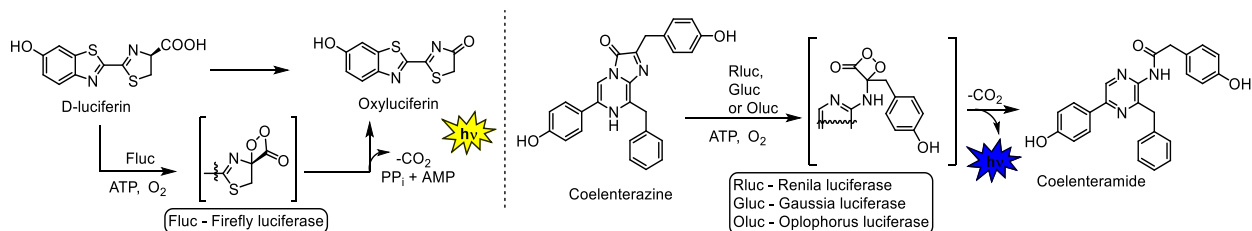


Figure 1.30. Structure and mechanism of light producing reaction between D-luciferin and Coelenterazine substrates and luciferase enzymes.

The chemical characterization of natural luciferase substrates (*i.e.* luciferins) remains extremely complicated. Heretofore, only two native luciferins have been routinely used in sensing applications, D-luciferin and coelenterazine,^{94,95} which restrict the imaging to only two color channels. Indeed, they respectively lead to the emission of yellow and blue signals ($\lambda_{em} = 560$ and 480 nm, respectively). Recently, significant efforts have been done to overcome this issue. A broad range of modified luciferin substrates have subsequently been reported. To date, the modification of the luciferin emitter structure granted access to a library of luciferin substrates with enhanced brightness and covering all the visible spectrum.⁹⁵

Luciferin-based probes have been designed by incorporating specific moieties onto a luciferin substrate. The direct modification of D-luciferin functional groups (hydroxy or carboxylic acid) enabled the design of sensors of small analytes^{96,97} and enzymes⁹⁸, either in cells and in tissues (**Figure 1.31**). Essentially, these modified luciferins are not substrates of luciferase. However, cleavage of this reactive unit releases the free luciferin that is further processed by luciferase in the light-producing reaction.

The D-luciferin-galactoside conjugate **Lugal (63)**,⁹⁸ reported by Blau's group (**Figure 1.31**) was successfully used to report β -gal activity in cells and *in vivo*. This conjugate is composed of D-luciferin where a galactose unit is connected through a beta-glycosidic bond to the hydroxy function. Probe **63** exhibits a better kinetic profile for imaging than luciferin itself. It displays a 30- to 200-fold brighter bioluminescent emission in mice models injected with cells expressing both β -galactosidase and Firefly luciferase (Fluc) compared with those expressing Fluc only. Moreover, **63** shows a 30 to 60-fold higher signal-to-noise ratio in subcutaneous tissues compared with **43** (used as control, *Section 1.8.2.1*). This ratio raises up a 150-fold when dipper tissue (*in vivo*) are analyzed.

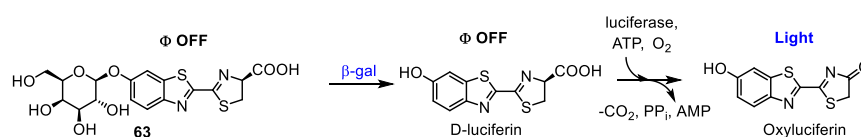


Figure 1.31. Example of bioluminescent-based probe to detect β -gal activity.

1.8.2.6 Excimer Based Probes

Lozano-Torres and co-workers⁹⁹ reported a two-photon fluorescent turn-ON probe **64** based on excimer emission to detect senescence-associated β -gal activity in senescent SK-MEL-103 (human melanoma) cells and *in vivo* (**Figure 1.32**). The probe is composed of 3 moieties, the 4-bromonaphthalimide core, a galactose (which is acetylated), connected through a *N*-glycosidic bond to a L-histidine methyl ester linker. Probe **64** displays negligible fluorescence emission ($\Phi_{fl} = 0.002$). However, upon β -gal processing, an intense emission in the yellow spectral region appears (**65**, $\Phi_{fl} = 0.458$, 286-fold, **Figure 1.32**). Interestingly, the presence of the galactoside moiety seems to inhibit the association between the excited and ground state chromophore and the excimer formation. However, the naked fluorophore **65** can assembly to form the respective excited state dimer. *In cellulo* imaging revealed a 10-fold

enhancement on the fluorescence emission intensity in SK-MEL-103 cells treated with Palbociclib (senescence inducer) when compared to non senescent control cells. Moreover, this enhancement was even higher (15-fold) when comparing mice tumors treated and non treated with the senescence inducing drug (**Figure 1.32**).

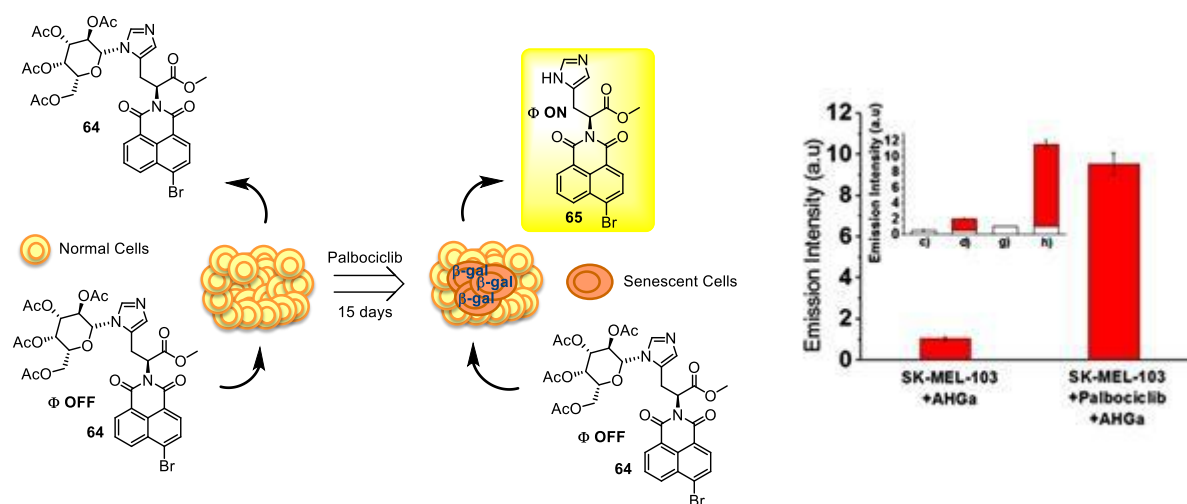


Figure 1.32. SA- β -gal detection using the turn-ON excimer-based fluorescent probe **64**.

1.8.2.7 Resonance Energy Transfer-Based Probes

Resonance Energy Transfer (RET) is a non-radiative mechanisms that involves energy transfer between a donor and a second molecule, acting as acceptor. Substantial spectral overlap between emission spectra of the fluorophore and absorption of the acceptor is necessary for an efficient energy transfer. Fluorescence resonance energy transfer (FRET) are amongst the most common and important tools in chemical sensing¹⁰⁰, especially to probe and study biomolecular systems as oligonucleotides and proteins.^{101–103} Both light-up and ratiometric systems have been designed based on RET. On the one hand, in turn-ON approaches, the acceptor moiety is selected in order to provide an effective quenching of the fluorophore emission (**Figure 1.33-A**). Azo dyes are typically used as fluorescence suppressors.^{104,105} Other designs are also reported, by combining PET and FRET¹⁰⁶ or using 7-nitrobenz-2-oxa-1,3-diazole (NBD) chromophore as quencher.^{107,108} On the other hand, ratiometric probes can be easily designed by connecting two different fluorophores with appropriate emission and absorption wavelengths (**Figure 1.33-B**).^{109,110}

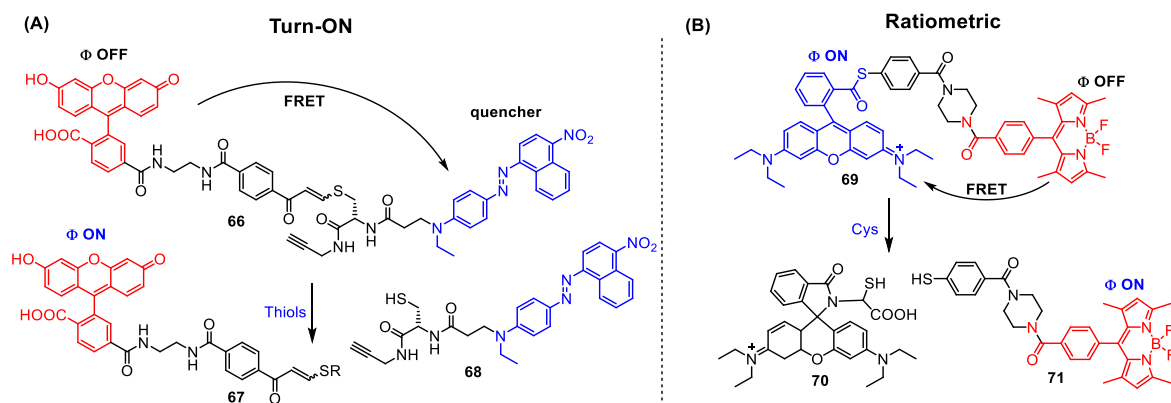


Figure 1.33. Examples of different molecular arrangements in FRET-based fluorescent probes. Turn-ON (A) and ratiometric (B) probes.

Heretofore, only one probe exploiting FRET mechanism had been developed for β -gal sensing. Nagano and co-workers¹⁰⁹ reported a FRET-based system for the fluorimetric detection of β -galactosidase and its simultaneous labeling (compound **72**, **Figure 1.34**). They described an original approach by connecting a coumarin donor and a fluorescein acceptor group through a modified benzyl carbamate linker bearing a β -galactopyranoside moiety. Once the system is placed in presence of β -gal, the enzymatic processing of the β -galactose substrate triggers the cleavage of the carbamate bond tethered to the fluorescein containing arm and the formation of a *p*-quinone methide active site (conversion of **72** into **73**, **Figure 1.34**). Two important features can be observed subsequently to this β -gal processing. First, a change in the fluorescence emission (as FRET is no more existent). Second, the liberated *p*-quinone methide-appended coumarin will react with a nucleophilic residue of the enzyme which results in the labeling of this protein (**75**, **Figure 1.34**).

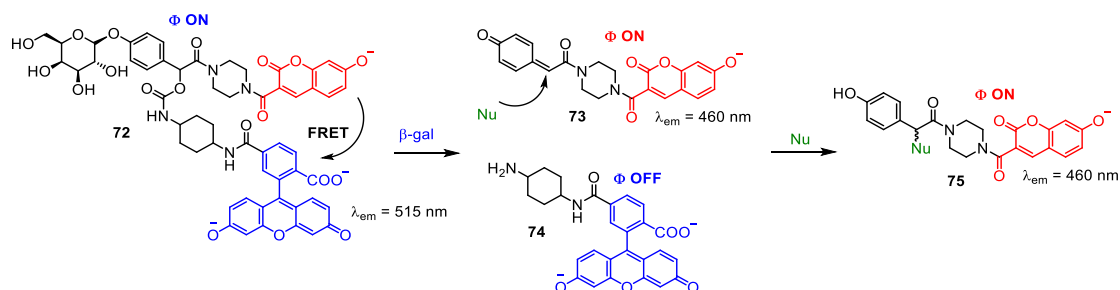


Figure 1.34. FRET-based system to the fluorescence labeling and sensing of β -galactosidase.

1.8.3 Environmental Sensing Based on Push-Pull Fluorophores

Differently from chemical or enzyme sensing, the detection of environmental changes can only be achieved by using environment responsive fluorophores. The main physicochemical properties that have been probed by fluorescence sensing approaches include changes in local viscosity, polarity, water content and pH. Fluorophores displaying excited state intramolecular charge transfer provide suitable spectroscopic changes in response to all these parameters. In addition, when used as fluorescent labels, solvatochromic dyes can detect the main types of biomolecular interactions as the ones involving nucleic acids, membranes and proteins.¹¹¹ Below we present a brief overview of the approaches that are used to probe environmental changes using solvatochromic fluorophores, with focused applications in cellular biology and fluorescence labeling.

1.8.3.1 Viscosity Probes

Molecular rotors encompass the major group of viscosity probes (**Figure 1.35**). The high dependence on the Φ_{fl} with viscosity make TICT the key fluorescence mechanisms explored to probe viscosity changes. **CCVJ (76)**, is one of the most famous example of molecular rotor, employed to study viscosity of membranes and the rigidity of binding sites of proteins.⁴ Nowadays, novel examples of molecular rotors have been reported, including merocyanine derivatives such as **77**,¹¹² used to visualize changes on mitochondrial viscosity and the dioxaborine-carbetocin conjugate **78**³⁸ used to target oxytocin receptor and visualize the cellular surface with excellent signal-to-noise ratio.

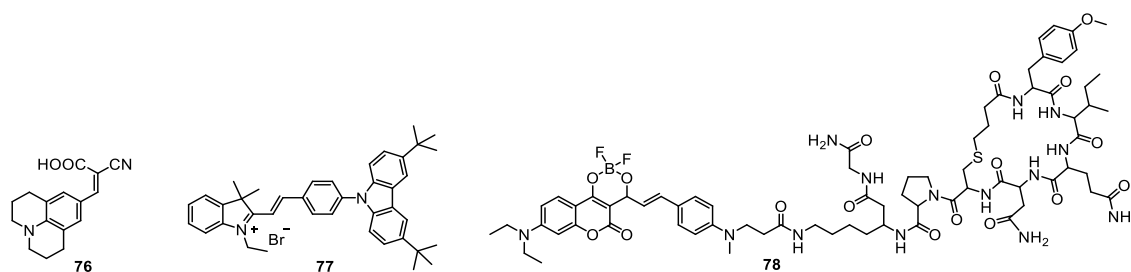


Figure 1.35. Environment sensitive molecular rotors.

1.8.3.2 Membrane Probes

Push-pull fluorophores also display suitable properties to probe different domains in cellular membranes. Although there are a large library of fluorescent membrane probes available, environment responsive dyes can afford additional information about the heterogeneity of cellular membranes.^{21,113} For instance, Klymchenko and co-workers recently reported a push-pull Nile Red zwitterion analogue **79** with redshifted absorption and emission maxima related to the parent **Laurdan** (**80**) (**Figure 1.36**).¹¹³ The reported Nile Red probe **79** presents such an important solvatochromic sensitivity that it enabled the ratiometric imaging of liquid-ordered (Lo) and liquid-disordered (Ld) phases in model membranes. It also permits to visualize the maturation of endosomes and to follow conformational changes of membrane proteins.

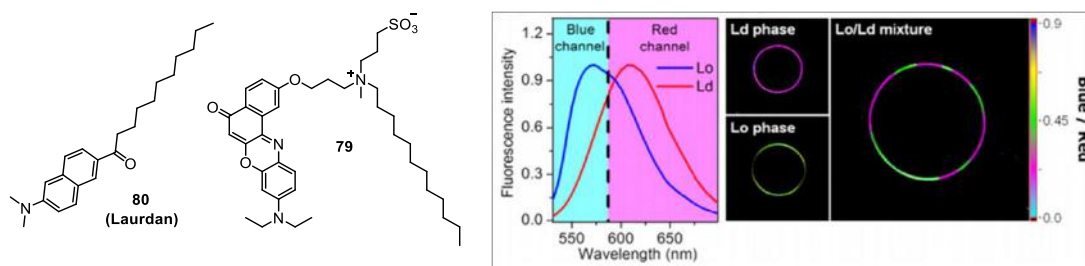


Figure 1.36. Membrane probe based on zwitterion Nile Red push-pull dye **79**.

1.8.3.3 Labeling Applications

The use of push-pull fluorophores for protein and oligonucleotide labeling have emerged as a powerful alternative to methods based on FRET and anisotropy. In fact, interaction of biomolecular partners basically results in the reduction of the local water concentration, thus decreasing the environment polarity. As a consequence, a blue shift on the position of the emission band and an enhancement on the Φ_{fl} may be used to monitor molecular interactions (**Figure 1.37**).^{54,111} Using single push-pull fluorophore, Sainlos and co-workers could probe the PDZ domains of specific protein targets PSD-95 and Shank3 by using C-terminal peptide probes modified with 4-dimethylamino phthalimide (4-DMAP) fluorophore (**Figure 1.37-A**).¹¹⁴ Riedl et al., reported the synthesis of modified cytidine triphosphate nucleotide tailored with 4-aminophthalimide (API) fluorophore, its incorporation into DNA sequences and the

study of DNA-protein interaction (**Figure 1.37-B**).^{115,116} In a different fashion, Dziuba and collaborators also studied DNA-protein association by using DNA sequences modified with push-pull fluorene dyes (**Figure 1.37-C**).¹¹⁷

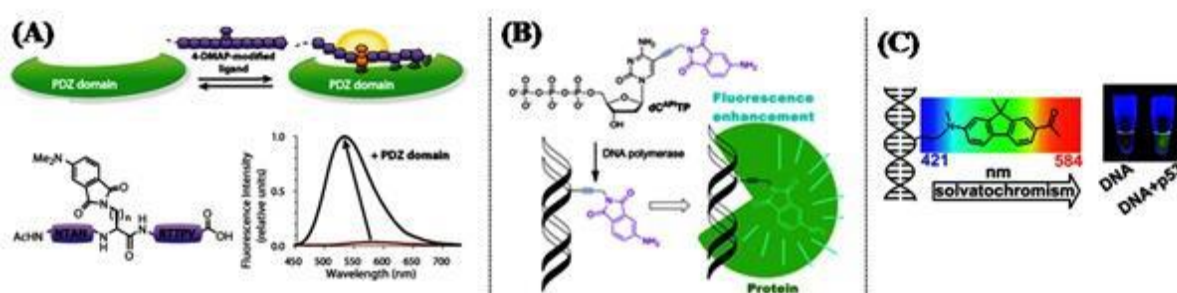


Figure 1.37. Examples of Push-pull fluorophores as single fluorescent labels to study biomolecular interactions.

1.9 Conclusions - Chapter 1

Chapter 1 was devoted to introduce and highlight some important features concerning fluorescence spectroscopy. The base concepts related to the phenomenon of fluorescence (photophysical processes, (T)ICT, electron transfer, fluorescence sensing) were introduced and discussed in certain detail as they will be fundamental for the understanding of the next two chapters of this thesis manuscript (**Chapters 2** and **3**). In the initial part (*sections 1.1* and *1.2*), the general remarks, related to the light absorption and the main photophysical processes that are triggered in an organic molecule upon absorption of a photon, were highlighted. All the different electronic transitions were summed up. The concept of vertical transitions (Franck-Condon) and the Law of Beer-Lambert were introduced as well. The major photophysical processes that take place upon light absorption (IC, ISC, fluorescence and phosphorescence emission) were briefly discussed with the support of Jablonski diagram. Therefore, the interlink between these processes with important properties as the fluorescence quantum yield and the fluorescence lifetime could then be noticed.

Taking into account the main thematic of **Chapter 2** (*Development of new push-pull benzothiazole-based fluorophores: synthesis and photophysical studies*), the *sections 1.3* to *1.5* were dedicated to present and establish a parallel among classic and environment responsive fluorophores. Few and important remarks concerning classic fluorescent dyes (fluorescein, rhodamine), that will appear along this manuscript, were also highlighted. Next,

the main discussion was dealing with solvatofluorochromic dyes. The definition of positive and negative solvatochromism was initially appointed. Then, the formation of intramolecular charge transfer (ICT and TICT) excited states in intrinsically polar compounds was discussed. The effect on the band shape of the emission spectra, associated with the interplay between Locally Excited (LE) and (T)ICT states resultant from changes in the environment polarity was shown. The effect of solvent polarity on the energy of the ICT state and how it is correlated with changes in the emitted color was demonstrated as well (solvatofluorochromism). Moreover, the influence of the environmental viscosity to modulate the conversion among LE-(T)ICT states was also highlighted. Although it is not related to none of the following chapters, it underlies an important feature of some fluorescent dyes displaying strong charge transfer excited state, which was briefly highlighted in *section 1.8.3.1*, in the context of environmental sensing (viscosity probes).

Short highlights concerning photoinduced electron transfer (PET) processes (*section 1.6*) and fluorescence labelling (*section 1.7*) were also provided during this introduction. Thereafter, the last part of this **Chapter 1** (*section 1.8* - fluorescence sensing) was dedicated to contextualize the use of fluorescence spectroscopy in sensing applications. As previously introduced, the main body of the **Chapter 3** (*Development of novel fluorescent probes for the detection of β -galactosidase activity*) of this manuscript is centered on the sensing of β -galactosidase activity, in the context of cellular senescence. Therefore, we dedicated this section to present a bibliographic review to show the majority of the fluorescent probes developed to report β -gal activity *in vitro*, *in cellulo* and *in vivo*. Initially, a short panorama considering the different fluorescence-based techniques used in sensing application was presented. Then, the central discussion was focused on the context of intensity-based probes. Probes with different architectures, exploiting different mechanism of sensing were described. Of note, the description of the main advantages and limitations of each probe was also given; they will be useful for comparison purposes in **Chapter 3**. Finally, short remarks concerning the use of solvatofluorochromic dyes in the environment sensing (membrane probes, viscosity probes) as well as in labeling, to report bimolecular interactions were shown.

Altogether, the introduction presented in **Chapter 1** gave a broad overview of the fluorescence phenomenon, the methods in which it is commonly used and several examples of applications.

Chapter 2 - Development of new push-pull benzothiazole-based fluorophores: synthesis and photophysical studies

Chapter Outline

The design and development of powerful fluorescent tools is a continuously expanding field of research.¹¹¹ Within this area, the quest for evermore bright, robust, fine-tuned and sensitive environment-responsive dyes has attracted a lot of attention. For this purpose, much effort has been devoted to the development of new environment sensitive fluorescent dyes with improved photophysical properties. However, it is noteworthy that heretofore only a small number of studies have been done to well establish a comprehensive structure-photophysics relationship within a family of fluorophores.^{46,118–120} In this chapter, we will present a new family of push-pull fluorophores bearing 6-amino functionalized benzothiazoles as electron donating groups. Firstly, the synthesis of novel and highly solvatochromic styryl dyes will be described (**Figure 2.1**). We decided to focus this work on these dyes where "A" is a σ -acceptor as just few studies have been reported for this type of framework. Then, a comprehensive photophysical study will be presented, to highlight the structure-photophysics relationship concerning these novel fluorophores. The influence of the three types of modifications of the central 2-styryl-6-aminobenzothiazole will be evaluated: (1) the alkylation of the 6-amino function (D); (2) decoration of the phenyl ring of the styryl moiety with various electron-withdrawing (EWG) and electron-donating groups (EDG) (A); and (3) the replacement of this simple phenyl motif with larger frameworks as polycyclic aromatic hydrocarbons (*e.g.* naphthyl, anthryl, and pyrenyl.) (**Figure 2.1**).

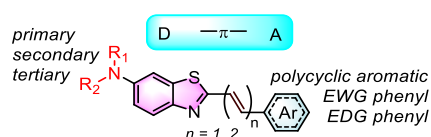


Figure 2.1. General structure of styryl-based push-pull dyes studied in this **Chapter 2**.

In the second part of this chapter, with the aim to design dyes with tuned spectroscopic properties, we will thus complement these preliminary studies and describe the synthesis and the photophysical characterization of a second series of 6-aminobenzothiazole-containing push-pull fluorophores bearing an extended π -conjugated framework (**Figure 2.2**). Two types

of modification will be presented and discussed: the acceptor group (**Ar** in blue, **Figure 2.2**) and the π -conjugated linker (**Aryl spacer**, **Figure 2.2**).

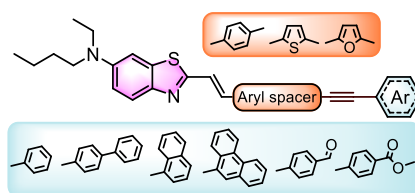


Figure 2.2. General structure of the push-pull fluorophore bearing extended π -conjugated scaffold.

2.1 Introduction - Styryl-Based Fluorophores

Styryl-based fluorophores constitute a common class of molecules exhibiting a D- π -A framework and encompass a broad variety of applications.¹²¹ Nevertheless, while a range of diverse styryl-based dyes have been reported,^{122–128} including derivatives of benzothiazoles,^{129–132} studies focused on neutral, highly emissive and solvatofluorochromic styryl-based benzothiazole scaffolds remain scarce in the literature.^{18,133–135} In 2013, Herner *et al.*¹⁸ described a class of fluorogenic dyes adaptable for direct bioconjugation via azido-alkyne cycloaddition albeit with relative reduced quantum yield. Ono and co-workers¹³⁴ reported two methylenemalononitrile-containing benzothiazole push-pull dyes with high affinity to probe $\alpha\beta$ amyloid aggregates in the human brain. In a different work, Liu and collaborators¹³⁵ described a solvatofluorochromic dye bearing a basic pyridine acceptor as a potential probe for the visualization of changes in the intracellular pH. Finally, Bashmakova *et al.*¹³³ investigated the photophysical behavior of three new solvatofluorochromic styryl benzothiazole-based dyes by steady-state, time-resolved fluorescence techniques and computational calculations. General scaffolds investigated in these aforementioned studies are presented **Figure 2.3**. Thus, there is still a need for a more comprehensive study devoted to the establishment of a robust structure-photophysics relationship in this peculiar class of dyes. This may prove helpful in the design of new fluorescent tools harboring optimized photophysical properties.

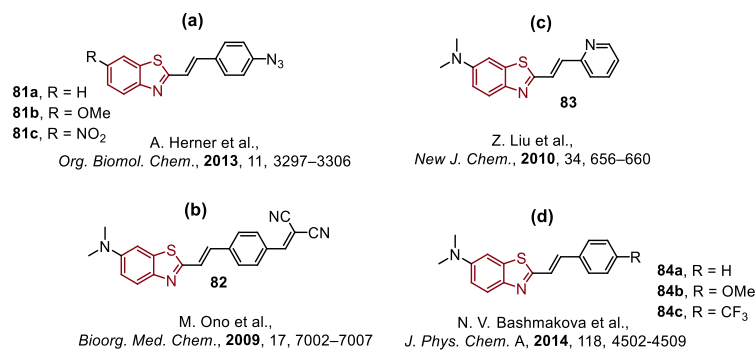


Figure 2.3. Benzothiazole-based styryl fluorophores.

In line with the studies mentioned above that aim at accessing new fluorescent probes for practical applications, we will describe herein the synthesis and the photophysical studies of 20 styryl-based push-pull dyes derived from 6-aminobenzothiazoles (**Figure 2.4**). For practical reasons, these 20 fluorophores have been divided into 3 categories according to three modifications on the fluorophore core: blue: *Substituted Phenyl Series* (dyes **85-93**); green: *Polycyclic Aromatic Hydrocarbon (PAH) Series* (dyes **94-98**); and red: *N-Substituted Series* (dyes **99-104**).

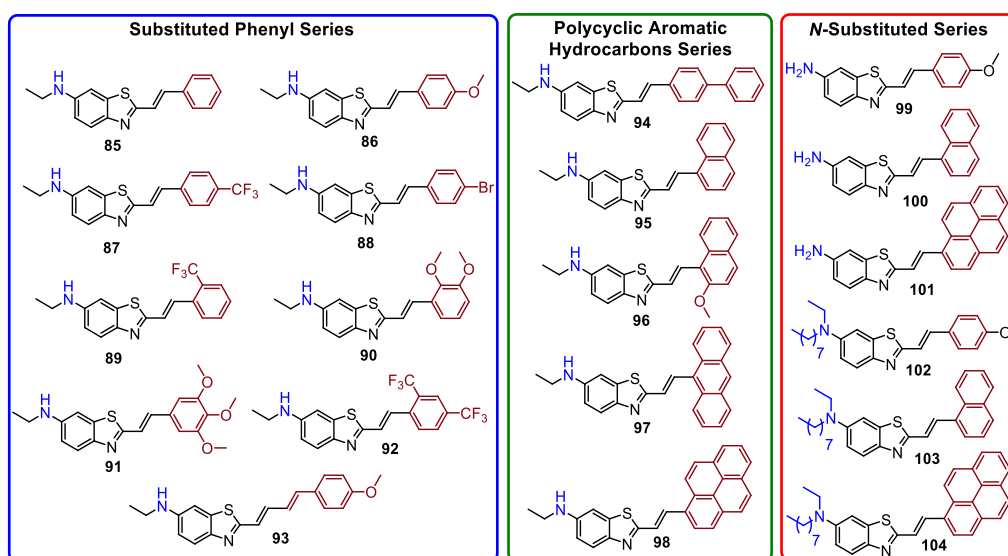
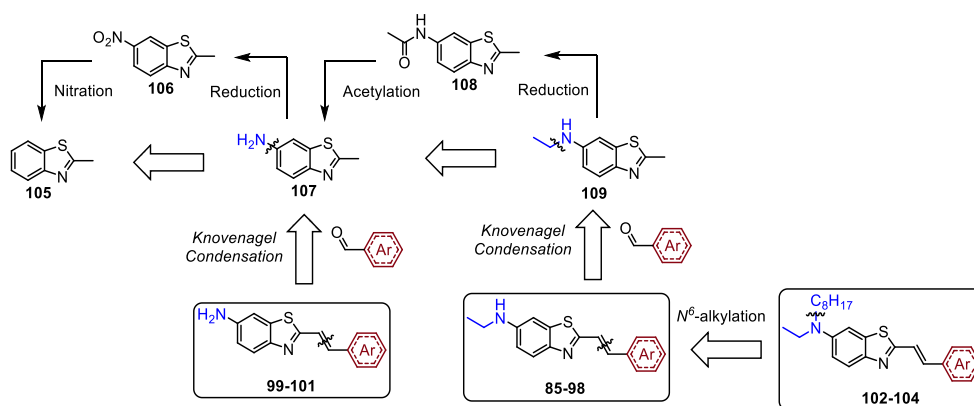


Figure 2.4. Structure of the first family of styryl benzothiazole fluorophores studied in the Chapter 2.

2.2 Synthesis of New Styryl-Based Fluorophores

The synthetic strategy envisaged to access these series of styryl fluorophores is highlighted in the *Scheme 2.1*. The different families of dyes were prepared through a linear synthetic route, in a practical protocol using classical methodologies. Initially, the commercially available 2-methylbenzothiazole (**105**) was converted to the key benzothiazole intermediates **107** and **109** over two to four steps. A conventional functional group transformation (regioselective nitration followed by reduction) was applied to prepare the 6-amino-2-methylbenzothiazole (**107**). Thus, the mono alkylated intermediate **109** was prepared using a two steps acetylation-reduction pathway. Finally, the dyes **85-98** and **99-101** series were synthesized using *via* a Knoevenagel condensation between **109** and **107** and a series of aromatic aldehydes (phenyl derivatives functionalized with EDG or EWG and derivatives containing larger frameworks as polycyclic aromatic hydrocarbons). Further alkylation of the dyes bearing the N^6 -ethylamino group (**86**, **95** and **98**) afforded the respective fluorophores **102**, **103** and **104**. For synthetic convenience, a series of fluorophores bearing the secondary amino group (N^6 -ethylamino) were used to probe the influence of different acceptor groups (dyes **85-98**).

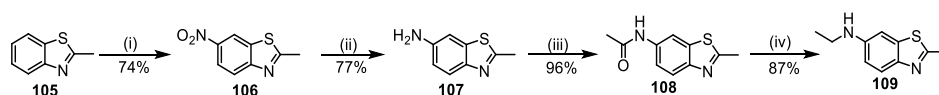


Scheme 2.1. Retrosynthetic pathway for the synthesis of the styryl-based fluorophores.

2.2.1 Synthesis of Benzothiazole Precursors **107** and **109**

The precursor 6-amino-2-methylbenzo[*d*]thiazole (**107**) was synthesized starting from commercially available 2-methylbenzothiazole (**105**) in two steps and 57% overall yield (*Scheme 2.2*). Firstly, 2-methylbenzothiazole was converted to the respective 2-methyl-6-nitrobenzo[*d*]thiazole (**106**) by nitration at 0°C using a stoichiometric amount of sodium nitrate in concentrated sulfuric acid as the nitrating mixture. The nitro derivative **106** was

obtained in good yield (74%) as light yellow needles after recrystallization from ethanol.¹⁸ Next, the treatment of a methanolic solution of **106** at room temperature with zinc powder and ammonium formate furnished the respective amino derivative **107** (77%). Subsequent acylation of the amino group of **107** using 1.1 equivalents of acetyl chloride in dry DMF and in the presence of DIPEA afforded the *N*-acetamido derivative **108** in excellent yield (96%). The synthesis of the key intermediate **109**, *N*-ethyl-2-methylbenzo[*d*]thiazol-6-amine, was accomplished by reduction of **108** using lithium aluminum hydride at room temperature for 12 h. Compound **109** was obtained in high yield (87%) after purification by flash column chromatography (*Scheme 2.2*). This route was performed on a multigram scale to provide enough material for the subsequent steps.

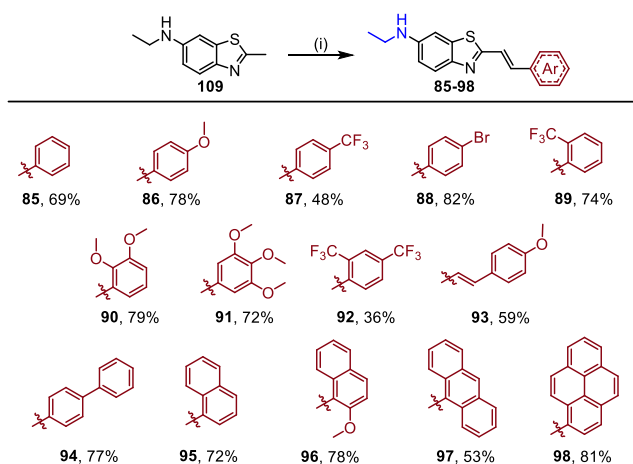


Scheme 2.2. Synthesis of the benzothiazole precursors **107** and **109**. Reagents and conditions.

- (i) NaNO_3 , H_2SO_4 , 0°C , 2 h, 74%. (ii) Zn, NH_4HCO_2 , methanol, r.t., overnight, 77%. (iii) AcCl , DIEA, DMF, $0 - 20^\circ\text{C}$, 1 h, 96%. (iv) LiAlH_4 , THF, r.t., 12 h, 87%.

2.2.2 Synthesis of Benzothiazole-based Styryl Fluorophores

The synthetic route towards dyes **85-98** involves the aldol condensation between various aromatic aldehydes and the intermediate **109** that bears an activated methyl group on its 2-position (*Scheme 2.3*). The aromatic aldehydes were chosen to modulate the steric and electronic properties of the final dyes. For this family of compounds, either the plain benzaldehyde, or benzaldehydes that are mono/poly substituted with EWG ($-\text{CF}_3$, Br) or EDG ($-\text{OMe}$), were used. Cinnamaldehyde bearing *p*-OMe EDG was selected to investigate the influence of the length of the π -conjugated bridge. Bulkier aromatic aldehydes, such as 1-naphthaldehyde, were also employed. Importantly, we included polycyclic aromatic aldehydes able to promote excimer formation (9-anthryl and 1-pyrenyl radicals) in this study.

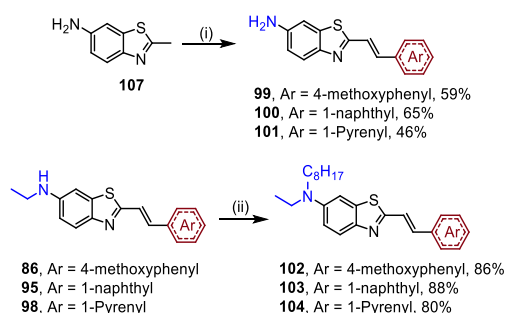


Scheme 2.3. Synthesis of benzothiazole-based styryl fluorophores bearing *N*-ethylamino group (**85-98**). Reagents and conditions. (i) aromatic aldehyde (1.3 equiv.), KOH, DMF, r.t., sonication, 1 h.

The preparation of **85-98** was performed at room temperature under ultrasonic irradiation, in *N,N*-dimethylformamide using a stoichiometric amount of potassium hydroxide. It is worth mentioning that attempts to use milder bases such as tertiary amines or sodium acetate did not afford any product. On the contrary, potassium *tert*-butoxide behaved similarly to hydroxide and proved efficient in the activation of the 2-methyl group of **109**. Thus, using KOH as the base, all dyes (**85-98**) were isolated in satisfactory yield (36-82%) after purification by chromatography. Finally, to ensure the high purity required for photophysical studies, **85-98** were further recrystallized from methanol (12-49% yield).

Next, the completion of this family of dyes was pursued by synthesizing two more series in which the 6-amino function of precursor **107** either remains as a primary amine or is converted to the tertiary *N*-ethyl-*N*-octyl-6-amino derivative. These two series enable the evaluation of the inductive effect of the alkyl substituents on the donating ability of the amino group attached to the benzothiazole scaffold (**Scheme 2.4**). Dyes bearing the primary amino group (**99-101**) were accessed in a similar way to dyes **85-98** but starting from the benzothiazole precursor **107** instead of **109**. For the second series, the tertiary amino derivatives (**102**, **103** and **104**) were obtained in excellent yields (80-88%) by an alkylation of the corresponding dye (**86**, **95** and **98**) using 1-iodooctane in the presence of potassium carbonate (**Scheme 2.4**). Interestingly, no byproducts corresponding to the quaternarization of the endocyclic nitrogen in the thiazole ring was observed. Hence, this simple and straightforward transformation could also be employed in the design and preparation of

valuable spectroscopic tools; for instance, this easy entry to dye modification could be used in the tailoring of membrane probes or biolabeling experiments.



Scheme 2.4. Synthesis of the fluorophores bearing primary and tertiary amino groups.

Reagents and conditions. (i) aromatic aldehyde (1.3 equiv.), KOH, DMF, r.t., sonication, 1 h.

(ii) 1-iodooctane (3 equiv.), K_2CO_3 , acetonitrile, $80^\circ C$, 24 h.

2.3 Photophysical Studies

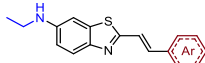
2.3.1 Modulation of the Acceptor Moiety - Substituted Phenyl Series: Dyes 85-93.

With these structurally characterized dyes hands we proceeded to the investigation of their photophysical properties in solutions of cyclohexane (Cyhex), toluene, 1,4-dioxane, ethyl acetate (EA), acetonitrile, dimethyl sulfoxide (DMSO) and methanol. The photophysical properties obtained for **85-93**, in these solvents, are summarized in **Table 2.1**.

To evaluate the influence of the phenyl substituents on the spectroscopic behavior of the dyes, we varied these substituents while keeping the donor group (ethylamino-) on the benzothiazole scaffold unchanged. Dye **85**, containing an unsubstituted, phenyl ring was chosen as the reference. All the dyes **86-93** containing substituted phenyl rings absorbed more strongly than **85** ($27.1 - 48.3 \times 10^3 M^{-1}.cm^{-1}$ vs. $25.4 \times 10^3 M^{-1}.cm^{-1}$, respectively), with **86** (*p*-OMe) presenting the highest hyperchromic effect of this series. Additionally, all dyes presented absorption bands in the visible range (384 - 412 nm in acetonitrile), with ϵ values in agreement with $\pi-\pi^*$ electronic transitions. Large Stokes shifts in polar solvents, *ca.* 5751-7605 cm^{-1} in acetonitrile, were also observed. The relatively weak solvent dependence in the absorption maxima of **85-93** indicates that the stabilization energy for the ground S_0 and the Franck-Condon state of S_1 are quite similar. In contrast, they showed gradually increased

bathochromic shifts in their emission maxima when going from the least polar aprotic solvent to the most polar solvent (**Table 2.1, Figure 2.5**). Using **85** as the reference, dyes containing **Table 2.1**. Relevant photophysical data from UV-visible absorption and fluorescence emission of dyes **85-93**.

Varied acceptor (substituted phenyl)/Fixed donor (N⁶-ethylamine)



Dye ($\epsilon \times 10^3$, ($M^{-1}.cm^{-1}$))	λ , $\Delta\lambda_{ST}^{[a]}$, $\Phi_{fl}^{[b]}$	Cyhex	Toluene	1,4- Dioxane	EA	DMSO	MeCN	MeOH
85 (25.4)	λ_{abs}	380	385	387	385	400	384	389
	λ_{em}	445	467	474	486	530	510	519
	$\Delta\lambda_{ST}$	3843	4560	4742	5398	6132	6434	6439
	Φ_{fl}	0.81	0.83	0.73	0.61	0.67	0.63	0.57
86 (48.3)	λ_{abs}	380	386	387	384	399	385	390
	λ_{em}	447	464	469	476	514	495	509
	$\Delta\lambda_{ST}$	3944	4355	4517	5033	5607	5772	5994
	Φ_{fl}	0.58	0.75	0.79	0.74	0.78	0.64	0.61
87 (27.1)	λ_{abs}	394	400	398	400	417	398	405
	λ_{em}	463	484	497	521	575	549	548
	$\Delta\lambda_{ST}$	3782	4338	5005	5806	6589	6911	6443
	Φ_{fl}	0.63	0.67	0.74	0.47	0.78	0.52	0.52
88 (36.8)	λ_{abs}	389	394	395	394	409	393	398
	λ_{em}	456	471	482	501	546	525	529
	$\Delta\lambda_{ST}$	3777	4149	4569	5420	6134	6397	6222
	Φ_{fl}	0.56	0.66	0.57	0.56	0.87	0.50	0.52
89 (34.4)	λ_{abs}	389	394	396	395	415	395	401
	λ_{em}	462	483	499	518	572	549	547
	$\Delta\lambda_{ST}$	4062	4677	5212	6011	6614	7101	6656
	Φ_{fl}	0.78	0.64	0.70	0.68	0.85	0.59	0.68
90 (34.3)	λ_{abs}	379	386	388	387	404	386	393
	λ_{em}	449	468	475	487	531	512	521
	$\Delta\lambda_{ST}$	4113	4539	4720	5306	5920	6375	6251
	Φ_{fl}	0.53	0.65	0.63	0.62	0.79	0.45	0.52
91 (41.1)	λ_{abs}	384	390	390	388	403	390	393
	λ_{em}	450	470	474	484	526	507	519
	$\Delta\lambda_{ST}$	3819	4364	4544	5112	5802	5917	6177
	Φ_{fl}	0.50	0.83	0.74	0.61	0.76	0.73	0.53
92 (38.1)	λ_{abs}	407	414	415	416	438	412	422
	λ_{em}	487	516	533	565	630	600	591
	$\Delta\lambda_{ST}$	4036	4775	5334	6339	6958	7605	6776
	Φ_{fl}	0.80	0.83	0.76	0.57	0.61	0.55	0.65
93 (62.4)	λ_{abs}	401	409	409	406	421	405	408
	λ_{em}	473	491	498	507	549	528	536
	$\Delta\lambda_{ST}$	3796	4083	4369	4906	5538	5751	5853
	Φ_{fl}	0.44	0.57	0.63	0.48	0.68	0.46	0.45

^[a]Stokes shifts ($\Delta\lambda_{ST}$) are reported in cm^{-1} . ^[b] Φ_{fl} were measured using quinine sulfate in an aqueous 0.1 M HClO₄ solution ($\lambda_{exc} = 350$ nm, $\Phi_{fl} = 0.59$) or fluorescein in 0.1 M NaOH aq. solution ($\lambda_{exc} = 475$ nm, $\Phi_{fl} = 0.89$) as the references.¹³⁶

para-substituted phenyl rings (**86-888**), presented similar absorption spectra, whereas the emission maxima was slightly blue-shifted for dye **86** that bears an EDG (*p*-OMe), and red-

shifted for dyes containing EWGs (**87**, *p*-CF₃ and **88**, *p*-Br), in polar solvents. The Φ_{fl} for dyes **85-93** were in the 0.44 - 0.87 range for all the tested solvents; no strong dependence on the solvent polarity was observed. In this series, dye **92** that features the 2,4-bis-(trifluoromethyl)phenyl pattern exhibited the most accentuated solvent sensitivity as indicated by the largest red-shift of its emission maxima ($\Delta\lambda_{\text{ST}} = 7605 \text{ cm}^{-1}/188 \text{ nm}$, in acetonitrile) (**Figure 2.5**). Additionally, the *ortho* or *para* positioning of the -CF₃ substituent (**87** and **89**) does not significantly affect either the absorption or emission maxima, while increasing the number of OMe groups (dye **90** and **91**), regardless of their location, resulted in a moderate bathochromic shift in the emission maxima in polar solvents, compared with **86**.

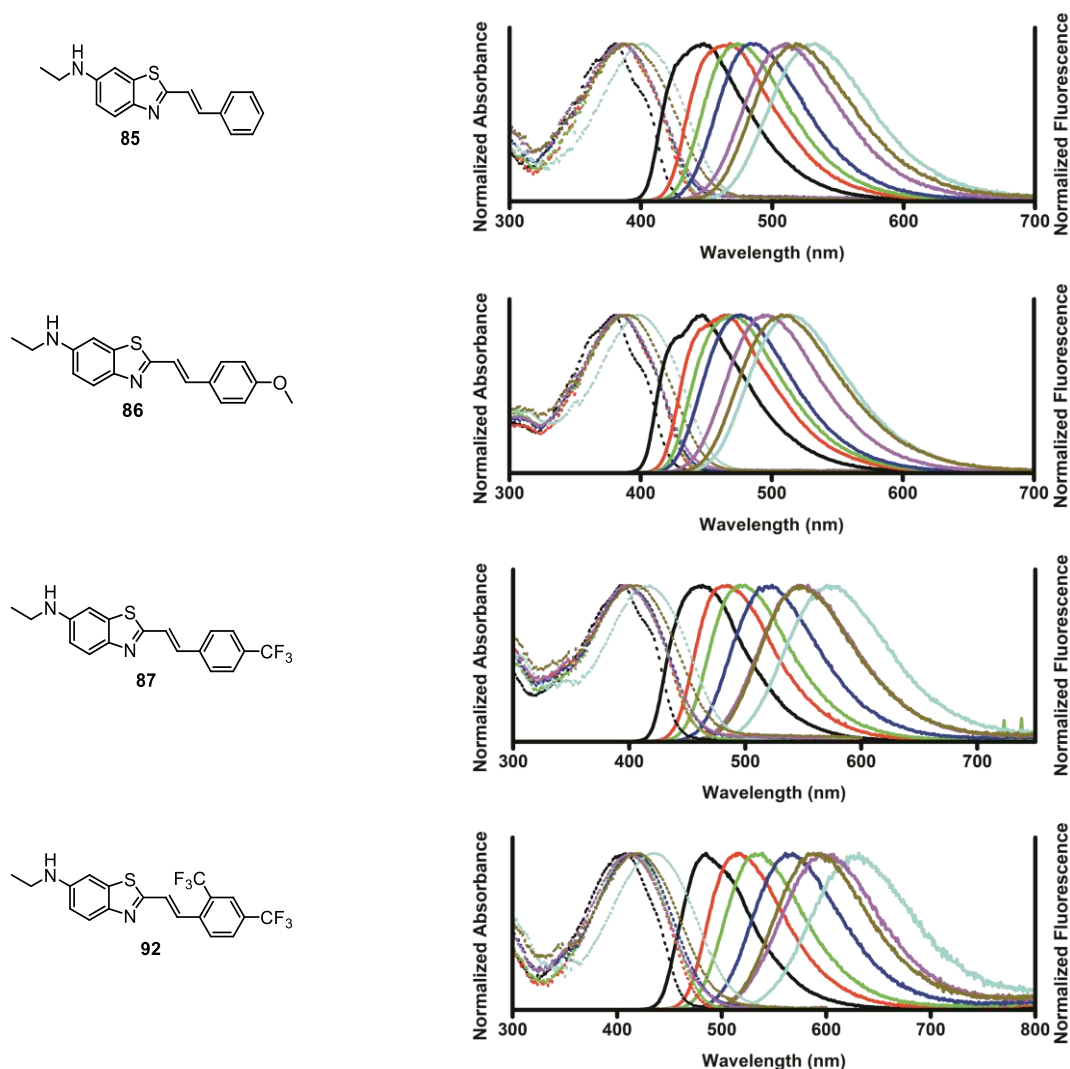


Figure 2.5. Normalized absorption and fluorescence emission spectra of dyes **85**, **86**, **87** and **92** in cyclohexane (black), toluene (red), 1,4-dioxane (green), ethyl acetate (blue), MeCN (magenta), DMSO (cyan) and MeOH (dark yellow).

The influence of the alkene bridge length between the benzothiazole and the phenyl ring on the spectroscopic properties was evaluated by comparing dyes **86** and **93** (Figure 2.6); the latter bears an additional vinyl unit. The absorption maximum for **93** was equally red-shifted, by about 20 nm, in all the tested solvents. This effect was accompanied with an increase in the molar extinction coefficient ϵ (+29%, compared with **86**), as a consequence of its higher degree of conjugation. Moreover, it is interesting to note that the λ_{abs} redshift, related to the additional vinyl unit, is significantly lower compared to similar cationic¹³⁷ and neutral^{34,138} polymethine dyes. In contrast with the absorption maximum of **93**, the fluorescence emission spectra showed some degree of solvent dependency, especially for the most polar media. In fact, while the λ_{em} difference between **93** and **86** is only 26 nm in cyclohexane, it reached 35 nm in DMSO.

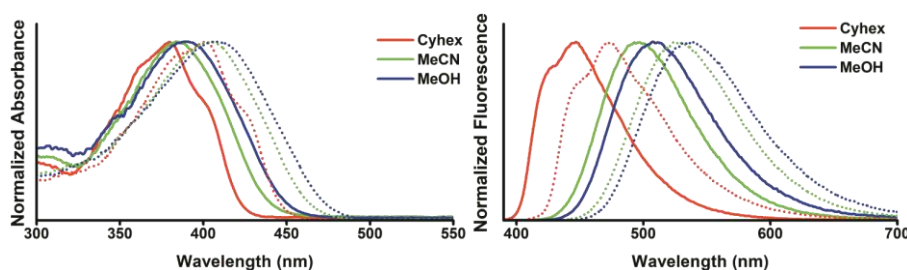


Figure 2.6. Normalized absorption (left) and fluorescence emission spectra (right) of dyes **86** (solid) and **93** (dotted) in different solvents.

The sensitivity of the dyes to solvent polarity was evaluated by plotting the Stokes shifts vs Reichardt's $E_{\text{T}}(30)$ scale (Figure 2.7). This latter quantitative parameter accounts for the dielectric constant of the solvent and its hydrogen-bond donor ability.⁴³ Among the first series of dyes (**85-93**), **92** showed the highest sensitivity (highest slope, Figure 2.7). As a general trend, more structured emission bands were observed in cyclohexane. This may be due to some degree of contribution from a locally excited (LE) state emission.³⁶ However, the large positive solvatochromism and the proportionally redshifted emission maxima for the dyes along the range of polarity strongly suggest the formation of a relaxed ICT state in solvents of increasing polarity. Moreover, the fact that only a single and broad structureless emission band is observed for all the dyes also supports the hypothesis of a fluorescence emission from an ICT state. In addition, the relatively intense fluorescence emission, even in polar environments, hints that the geometry of the ICT excited state seems to be close to the ground state and contributions associated to the formation of a non-coplanar TICT state may

not be extensively involved. Differences between the dipole moments ($\mu_e - \mu_g$) of the ground and excited states of the dyes was estimated by plotting the Stokes shifts as a function of the Lippert-Mataga orientation polarizability (Δf) parameter^{41,42} (**Figure 2.7** and **Table 2.2**). Values of 10.3, 9.9, 11.7 and 13.4 D, for $\mu_e - \mu_g$, were obtained for dyes **85**, **86**, **87** and **92**, respectively. The increase in the excited state dipole moment is accomplished by a higher level of charge separation between the donor (6-ethylamino) and the acceptor (aryl group). Hence, since the donor group remained unchanged in this series, the largest charge separation, corresponding to the highest $\mu_e - \mu_g$ difference, was observed for **92** that features the most electron-poor arene (2,4-bis-CF₃) on the acceptor moiety.

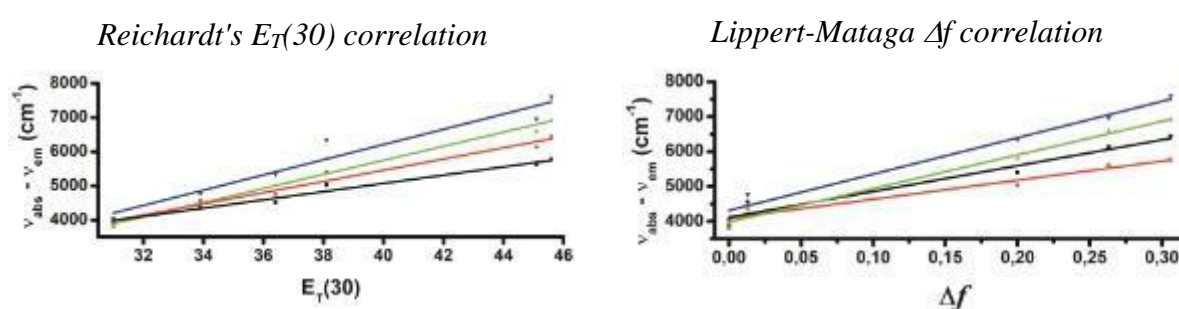


Figure 2.7. Dependence of the Stokes shift on the empirical polarity parameters $E_T(30)$ and Δf for dyes **85** (red), **86** (black), **87** (green) and **92** (blue). Correlation factors R^2 are, respectively, 0.96, 0.97, 0.95 and 0.93 for $E_T(30)$ and 0.93, 0.95, 0.98 and 0.97 relative to Δf . The linear plots were obtained with all the tested solvents except methanol.

Table 2.2. Dipole moment differences between excited and ground states ($\mu_e - \mu_g$), for dyes **85-93**.

<i>Varied acceptor (substituted phenyl)/Fixed donor (N⁶-ethylamine)</i>									
Dye	85	86	87	88	89	90	91	92	93
$(\mu_e - \mu_g)^{[a]}$	10.3	9.9	11.7	11.4	12.2	10.5	11.0	13.4	10.9

^[a]values are given in Debye (D).

To complete the spectroscopic study, the dyes were further investigated by DFT calculations at the MPW1PW91/6-311+G(d,p) level of theory using Gaussian03. Optimized ground-state geometries were found to be close to planarity for all dyes except for those presenting *ortho*-substituents (**89**, **90** and **92**), where the phenyl ring appeared slightly twisted respectively to

the 2-vinylbenzothiazole motif. The calculations revealed that the less energetic electronic transitions occur between the frontier molecular orbitals HOMO and LUMO. As a general trend, the highest occupied molecular orbital (HOMO) is partially localized on the benzothiazole donor and around the vinyl bridge while the lowest unoccupied molecular orbital (LUMO) is spread across the bridge and the phenyl ring (**Figure 2.8**). Compared with other dyes, the LUMO of **92** features a higher electronic density on its acceptor motif (2,4-*bis*-(CF₃)). This strongly supports the large ICT character that was observed for this dye. Furthermore, for **85-87** and **92**, time-dependent (TD)-DFT calculations in cyclohexane and DMSO yielded their corresponding calculated absorption maxima associated to the first electronic transition S₀→S₁. Compared with the experimental values, the calculated absorption maxima followed the same trend and were only slightly red-shifted.

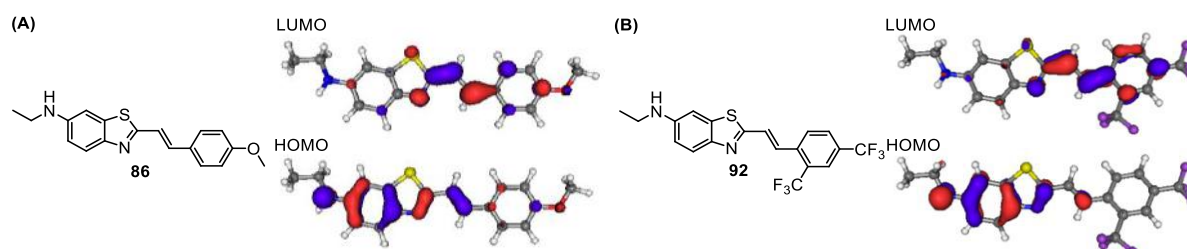


Figure 2.8. Representation of HOMO and LUMO, for dyes **86** (A) and **92** (B), obtained from optimized ground state geometries.

Table 2.3. Time-dependent TD-DFT calculations data.

Dye	Solvent	λ_{abs}	$\lambda_{\text{abs}}(\text{calc.})$	S ₀ →S ₁ transition energy (eV)
85	Cyclohexane	380	408	3.0367
	DMSO	400	418	2.9648
86	Cyclohexane	380	406	3.0510
	DMSO	399	415	2.9864
87	Cyclohexane	394	431	2.8731
	DMSO	417	441	2.8062
92	Cyclohexane	407	450	2.7549
	DMSO	438	463	2.6767

Finally, to overcome the challenges in predicting the emission maxima and Stokes shifts by DFT calculations, the correlation between λ_{em} and $\Delta\lambda_{\text{ST}}$ (in wavelength) for dyes containing *para*-substituted phenyl rings was performed using the σ_{p} Hammett substituent parameter model.¹³⁹ Dye **85** was selected as reference ($\sigma_{\text{p}} = 0$), while dyes **86**, **87** and **88** were selected as representative dyes containing EDG and EWG. **Figure 2.9** depicts the plots of emission

maxima (λ_{em}) vs σ_p for five different solvents. Excellent correlations for the λ_{em} were achieved for all solvents ($R^2 = 0.94 - 0.98$) except for those of low polarity (cyclohexane and toluene). In parallel, linear trend between Stokes shift and σ_p values were obtained exclusively for solvents with moderate and high polarity (1,4-dioxane to methanol), with R^2 values ranging from 0.80 to 0.98 (see *Annex section, Annex 1*). To verify our prediction model, the emission maxima and the Stokes shift for the reported dye **85** were calculated for five solvents using the respective curves of linear regression obtained for each solvent. The values were slightly shifted compared to the experimental ones (**Table 2.4**), with errors in an acceptable range (± 3 nm for both λ_{em} and $\Delta\lambda_{ST}$).

Table 2.4. Prediction study based on Hammett parameter model for dye **85**.

Solvent	λ_{em} (experimental/predicted)	$\Delta\lambda_{ST}$ (experimental/predicted)
1,4-Dioxane	474/476	87/86
Ethyl acetate	486/489	101/101
DMSO	530/532	130/129
MeCN	510/511	126/124
MeOH	519/520	130/127

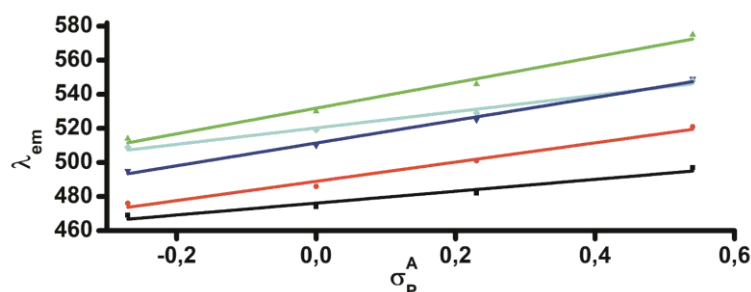


Figure 2.9. Dependence the emission maxima of **85**, **86**, **87** and **88** on the Hammett values (σ_p) in methanol (cyan), acetonitrile (blue), dimethyl sulfoxide (green), ethyl acetate (red) and 1,4-dioxane (black). Correlation factors R^2 are, respectively, 0.94, 0.98, 0.98, 0.99 and 0.98.

To sum up, all the dyes from this first series were mostly insensitive to polarity in the ground state (weak solvatochromic effect on the absorption band). By contrast, a strong positive solvatofluorochromism was verified for all fluorophores, with more pronounced effect for dyes bearing EWG attached to the acceptor phenyl group (Stokes shift up to 192 nm was obtained for dye **92**). This spectroscopic observation is in agreement with the formation of an excited state with a strong ICT character. Interestingly, the influence of the alkene bridge

length between the benzothiazole and the phenyl ring on the spectroscopic properties of these dyes is significantly less pronounced compared with some parent polymethine dyes. Moreover, the perfect correlation between Hammett parameter (σ_p) and Stokes shifts and λ_{em} for dyes containing para-substituted phenyl rings, could provide a complementary tool to predict photophysical properties of other members of this family of fluorescent probes.

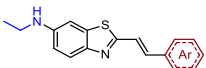
2.3.2 Modulation of the Acceptor Moiety - Polycyclic Aromatic Hydrocarbon Series: Dyes 94-98

To get further insight into the effect of the acceptor moiety on the photophysical properties of the dyes, different polycyclic aromatic hydrocarbons were evaluated using the same benzothiazole scaffold while keeping the 6-ethylamino function unchanged. The corresponding data are presented in **Table 2.5**. **Figure 2.10** depicts the UV-visible and fluorescence emission spectra of dyes **94-98** in seven representative solvents.

Similarly to the phenyl series (**85-93**), dye **85** was again selected as the reference. All dyes containing polycyclic aromatic groups, **94-98**, showed stronger absorption than **85** ($28.9 - 54.6 \times 10^3 \text{ M}^{-1} \cdot \text{cm}^{-1}$ vs. $25.4 \times 10^3 \text{ M}^{-1} \cdot \text{cm}^{-1}$, in dioxane). The absorption maxima for compounds containing up to three fused phenyl rings (**95-97**) exhibited a moderate bathochromic shift (9-20 nm in acetonitrile) compared with **85**. In contrast, replacing the phenyl ring with a bulky 1-pyrenyl core (dye **98**) strongly red-shifted the maximum (~ 40 nm in all the tested solvents) as a result of its higher degree of conjugation. Furthermore, this series, **94-98**, showed little dependency on the solvent polarity in their λ_{abs} , thus supporting the hypothesis of a weak charge transfer character in their ground state. In contrast with their absorption maxima, the fluorescence emissions were affected in a similar manner to the phenyl derivatives **85-93**. Indeed, a high degree of solvent polarity dependency on λ_{em} was observed ($\Delta\lambda_{ST}$ up to 9103 cm^{-1} in acetonitrile).

Table 2.5. Relevant photophysical data from UV-visible absorption and fluorescence emission of dyes **94-98**.

Varied acceptor (PAH)/Fixed donor (N⁶-ethylamine)



Dye ($\epsilon \times 10^3$, $\text{M}^{-1}\cdot\text{cm}^{-1}$)	λ , $\Delta\lambda_{\text{ST}}^{\text{[a]}}$, $\Phi_{\text{fl}}^{\text{[b]}}$	Cyhex	Toluene	1,4- Dioxane	EA	DMSO	MeCN	MeOH
85 (25.4)	λ_{abs}	380	385	387	385	400	384	389
	λ_{em}	445	467	474	486	530	510	519
	$\Delta\lambda_{\text{ST}}$	3843	4560	4742	5398	6132	6434	6439
	Φ_{fl}	0.81	0.83	0.73	0.61	0.67	0.63	0.57
94 (44.7)	λ_{abs}	391	398	399	396	413	396	401
	λ_{em}	464	479	487	507	557	531	533
	$\Delta\lambda_{\text{ST}}$	4024	4249	4529	5528	6259	6420	6176
	Φ_{fl}	0.63	0.77	0.60	0.48	0.79	0.52	0.55
95 (37.3)	λ_{abs}	383	392	395	393	412	393	399
	λ_{em}	475	491	502	519	571	548	548
	$\Delta\lambda_{\text{ST}}$	5057	5143	5396	6177	6758	7197	6814
	Φ_{fl}	0.37	0.47	0.45	0.51	0.82	0.54	0.70
96 (36.8)	λ_{abs}	386	396	397	394	414	396	402
	λ_{em}	478	494	503	513	559	537	536
	$\Delta\lambda_{\text{ST}}$	4986	5009	5308	5887	6265	6630	6218
	Φ_{fl}	0.42	0.30	0.48	0.47	0.75	0.43	0.48
97 (28.9)	λ_{abs}	400	408	407	405	419	404	412
	λ_{em}	557	571	581	600	674	639	628
	$\Delta\lambda_{\text{ST}}$	7047	6996	7358	8025	9029	9103	9139
	Φ_{fl}	0.47	0.48	0.40	0.36	0.36	0.33	0.31
98 (54.6)	λ_{abs}	418	428	427	425	443	425	429
	λ_{em}	508	526	531	558	621	591	580
	$\Delta\lambda_{\text{ST}}$	4238	4353	4586	5608	6470	6608	6068
	Φ_{fl}	0.68	0.71	0.73	0.60	0.65	0.59	0.63

^[a]Stokes shifts ($\Delta\lambda_{\text{ST}}$) are reported in cm^{-1} . ^[b] Φ_{fl} were measured using quinine sulfate in an aqueous 0.1 M HClO_4 solution ($\lambda_{\text{exc}} = 350 \text{ nm}$, $\Phi_{\text{fl}} = 0.59$) or fluorescein in 0.1 M NaOH aq. solution ($\lambda_{\text{exc}} = 475 \text{ nm}$, $\Phi_{\text{fl}} = 0.89$) as the references.¹³⁶

Dye **97**, containing the 9-anthryl scaffold, showed the broadest absorption band with the lowest ϵ value among the polycyclic aromatic hydrocarbon derivatives. Additionally, the large shoulder around 350 nm may indicate a lack of planarity and that the resulted absorption band could contain a high contribution from a twisted anthracene moiety (**Figure 2.10**). This hypothesis was supported by the optimized DFT structure of **97** (**Figure 2.11**). It showed both a non-coplanar orientation of the 9-anthryl and vinylbenzothiazole motifs as well as a high electronic density located on the anthryl framework in the frontier orbitals.

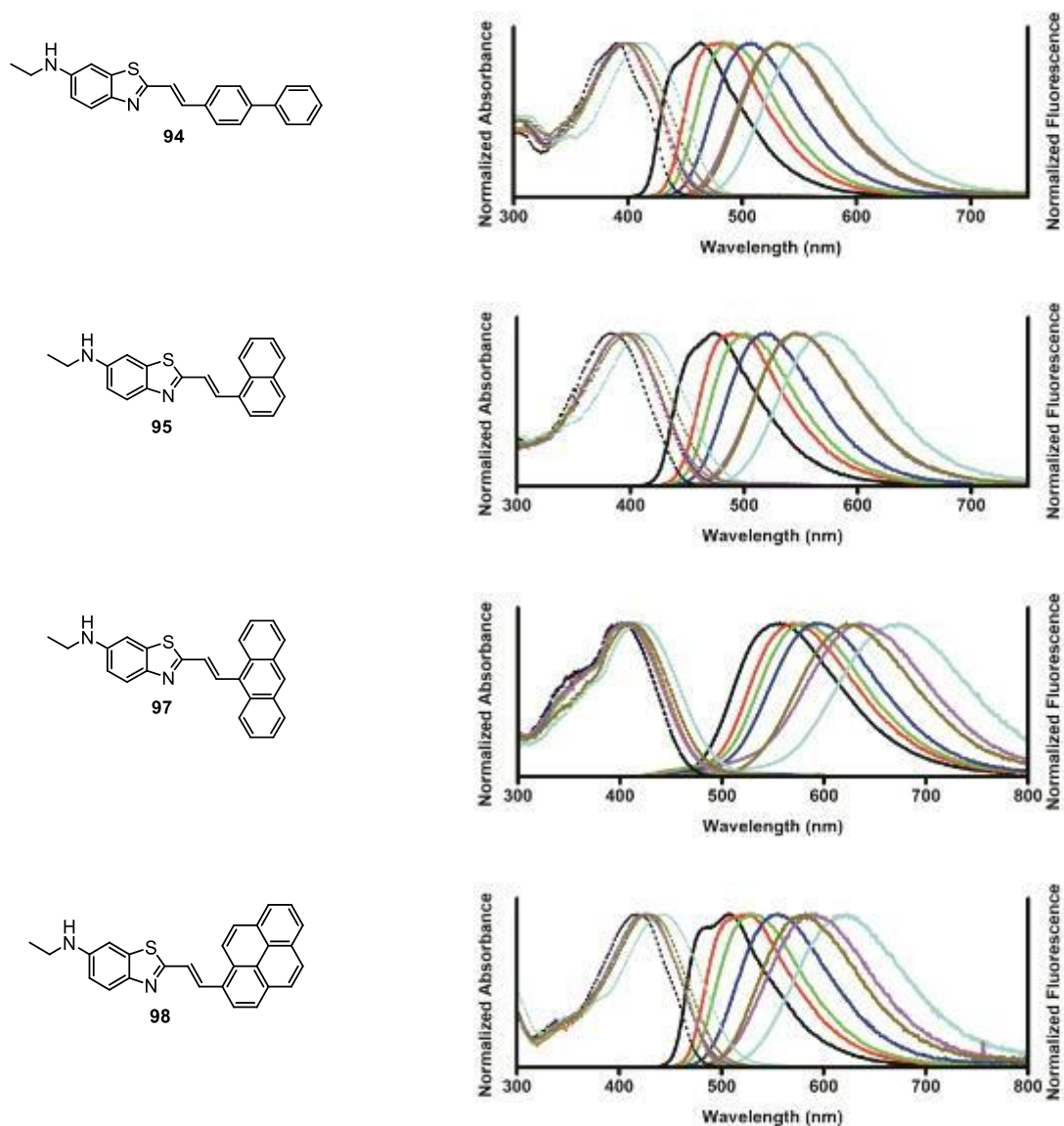


Figure 2.10. Normalized absorption and fluorescence emission spectra of dyes **94**, **95**, **97** and **98** in cyclohexane (black), toluene (red), 1,4-dioxane (green), ethyl acetate (blue), MeCN (magenta), DMSO (cyan) and MeOH (dark yellow).

Compared with the other dyes of the same series, **97** presented distinct emission behavior with the lowest Φ_{fl} in polar solvents. It displayed an impressive mega-Stokes shift in DMSO (*ca.* 9029 cm^{-1} /255 nm). In addition, the intriguing highly redshifted and broad structureless emission band observed in cyclohexane ($\Delta\lambda_{ST} = 7047 \text{ cm}^{-1}$ /157 nm) may suggest that a highly dipolar state is also formed in non-polar environments. Based on photophysical studies concerning derivatives of 4-(9-anthryl)-*N,N*-dimethylaniline (ADMA) and analogues,^{140,141} we suspect that the strong electron acceptor character of the anthracene group in the excited state may favor an intramolecular electron transfer mechanism from the 6-*N*-ethylamino

group to the anthryl radical, thus yielding an highly dipolar structure ($D^+ \cdot A^-$) with reduced fluorescence emission.

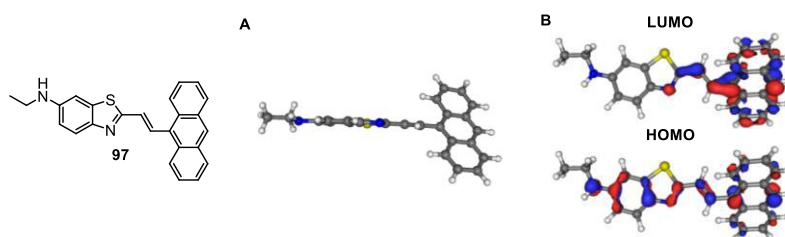


Figure 2.11. (A) Optimized DFT structure of **97**. (B) Frontier orbitals obtained from optimized ground state geometry.

As a general observation, all dyes bearing polycyclic aromatic hydrocarbons on the acceptor side (**94-98**) present more pronounced levels of solvent sensitivity than **85** ($\Delta\lambda_{ST}$ comparison, **Table 2.5**). The Φ_{fl} for dyes **94-98** were in the 0.30 - 0.82 range for all the tested solvents and, as well as pointed for the phenyl series (**85-93**), no strong dependence on the solvent polarity was also verified. In addition, similarly for the phenyl series, the attachment of an electron donating -OMe group has a detrimental effect on the polarity sensitivity (**95** vs **96**); in all solvents, **95** displays larger Stokes shifts values than its parent dye **96**.

As for the first series (substituted phenyl), a linear correlation between the Stokes shifts of dyes **94-96** and **98** and the solvent polarity ($E_T(30)$) (**Figure 2.12**) was evidenced in this second series, except for **97** in which the reduced slope does not correlate with the photophysical data. This last observation may indicate that emission takes place from a distinct and highly polar CT state (*e.g.* involving an electron transfer process). The calculated values of $\mu_e - \mu_g$ for dyes **94-96** and **98** (11.7, 10.6, 10.2 and 12.4 D, respectively, **Table 2.6**) are in agreement with the spectroscopic observation and support the strong charge transfer character for these dyes in the excited state.

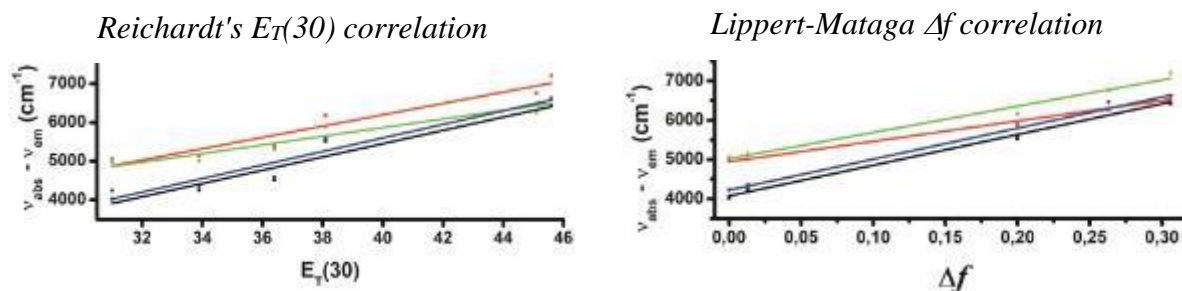


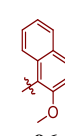
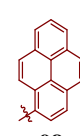


Figure 2.12. Dependence of the Stokes shift on the empirical polarity parameters $E_T(30)$ and Δf for dyes **94** (red), **95** (black), **96** (green) and **98** (blue). Correlation factors R^2 are, respectively, 0.96, 0.97, 0.95 and 0.93 for $E_T(30)$ and 0.99, 0.98, 0.99 and 0.98 relative to Δf . The linear plots shown were obtained with all the tested solvents except methanol.

Table 2.6. Dipole moment differences between excited and ground states ($\mu_e - \mu_g$) for dyes **94**, **95**, **96** and **98**.

<i>Varied acceptor (PAH)/Fixed donor (N⁶-ethylamine)</i>				
				
Dye	94	95	96	98
$(\mu_e - \mu_g)^{[a]}$	11.7	10.6	10.2	12.4

^[a]Values are given in Debye (D).

Finally, further investigations to highlight a putative excimer emission for the pyrene-containing dye, **98**, was performed. Surprisingly, no strong evidence associated with an excited state dimer emission was observed regardless of the broad range of concentrations or the different solvents that were probed (**Figure 2.13**). Nevertheless, these observations do not rule out the possibilities of designing dyes derived from **98** as fluorescent ratiometric monomer-excimer emitting probes.

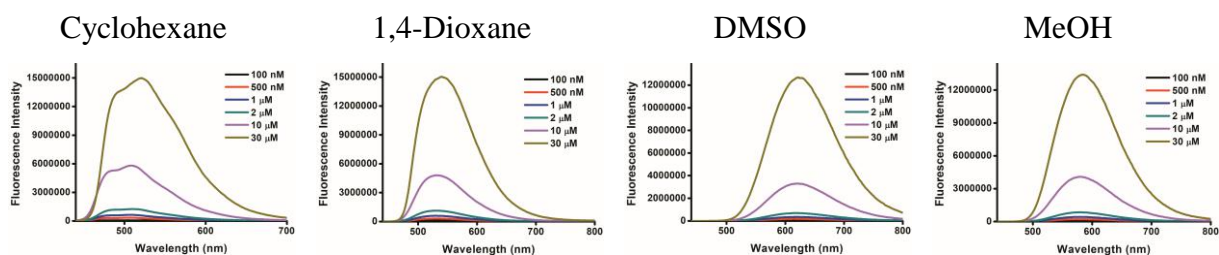


Figure 2.13. Fluorescence emission spectra of dye **98** in different concentrations.

2.3.3 Modulation of the Donor Moiety - Spectroscopic Impact of the Alkylation of the 6-Amino Functional Group

The last part of our study aimed at evaluating the effect of alkyl radicals attached to the 6-amino group of the benzothiazole ring on the spectroscopic properties of the dyes. For this purpose, the parent dyes bearing either a primary (**99**, **100** and **101**), a secondary (**86**, **95** and **98**) or a tertiary amine at the 6-position of the benzothiazole (**102**, **103** and **104**) were compared.

Figure 2.14 depicts the UV-Vis absorption and fluorescence emission spectra for these nine dyes in cyclohexane and acetonitrile. Additional photophysical data are reported in the **Table 2.7**.

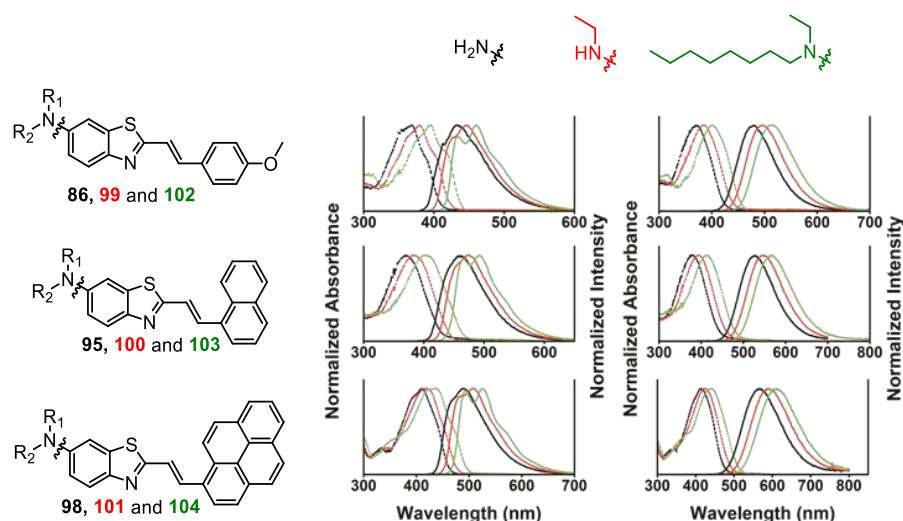


Figure 2.14. Normalized absorption (dashed) and fluorescence emission (solid) spectra of dyes **86**, **99** and **102** (top), **95**, **100** and **103** (middle) and **98**, **101** and **104** (bottom) in cyclohexane (left) and acetonitrile (right).

As expected and in agreement with their positive inductive effect, increasing the number of alkyl substituents on the 6-amino group resulted in a gradual red-shift of both absorption and emission maxima (~ 10 to 15 nm per alkyl substituent). In addition, more structured emission bands in cyclohexane are also observed for dyes containing dialkyl substituted amines. It may be a consequence of the increased hydrophobicity and the loss of the hydrogen bond formation ability, thereby reducing the intermolecular dye-dye interactions. The main spectroscopic properties (ϵ , Φ_{fl} , and $\Delta\lambda_{ST}$) were moderately different among the dye families (bearing primary, secondary and tertiary amines). Dyes **99** and **100**, bearing the primary

amine, showed reduced Φ_{fl} and higher ϵ values, in comparison with their analogues **102** and **103**, respectively. In contrast, dyes **101** and **104** presented similar Φ_{fl} values, especially in polar solvents, with a higher extinction coefficient for **104**. Linear trends between the Stokes shifts and the solvent polarity parameters $E_T(30)$ and Δf were also obtained. The transition dipole moments were directly correlated with the number of alkyl substituent on the amino group, with dyes bearing the tertiary amine presenting larger $\mu_e - \mu_g$ values (Table 2.8). Finally, the attachment of a second alkyl substituent on the 6-amino position (**102**, **103** and **104**) does not significantly reduce the brightness ($\epsilon \cdot \Phi_{fl}$), when compared to the analogue containing the secondary *N*-ethylamine EDG (**86**, **95** and **98**).

Table 2.7. Relevant photophysical data from UV-visible absorption and fluorescence emission of dyes containing primary (**99-101**) and tertiary amines (**102-104**).

Fixed acceptor (Ar)/Varied donor (6-amino group, $R_1 = R_2 = H$; $R_1 = Et$, $R_2 = nOctyl$)

Dye ($\epsilon \times 10^3$, $M^{-1} \cdot cm^{-1}$)	λ , $\Delta\lambda_{ST}^{[a]}$, $\Phi_{fl}^{[b]}$	Cyhex	Toluene	1,4- Dioxane	EA	DMSO	MeCN	MeOH
99 (37.7)	λ_{abs}	368	374	379	374	388	371	371
	λ_{em}	433	449	463	466	501	479	496
	$\Delta\lambda_{ST}$	4079	4466	4786	5279	5813	6077	6793
	Φ_{fl}	0.25	0.33	0.45	0.29	0.52	0.41	0.30
102 (33.1)	λ_{abs}	395	401	400	396	410	400	405
	λ_{em}	459	480	480	487	525	515	421
	$\Delta\lambda_{ST}$	3529	4104	4166	4718	5342	5582	5497
	Φ_{fl}	0.69	0.80	0.76	0.66	0.80	0.63	0.61
100 (34.4)	λ_{abs}	370	378	380	382	401	380	381
	λ_{em}	461	480	492	507	559	527	532
	$\Delta\lambda_{ST}$	5335	5621	5990	6454	7048	7340	7449
	Φ_{fl}	0.37	0.25	0.28	0.38	0.73	0.38	0.25
103 (29.7)	λ_{abs}	403	413	413	410	425	413	417
	λ_{em}	466	497	514	531	584	566	559
	$\Delta\lambda_{ST}$	3354	4092	4757	5558	6406	6545	6091
	Φ_{fl}	0.49	0.66	0.45	0.69	0.76	0.63	0.64
101 (54.2)	λ_{abs}	410	417	417	416	432	414	417
	λ_{em}	488	508	517	539	608	566	567
	$\Delta\lambda_{ST}$	3898	4295	4638	5485	6700	6486	6344
	Φ_{fl}	0.73	0.69	0.61	0.55	0.76	0.49	0.47
104 (65.9)	λ_{abs}	438	443	441	439	453	440	443
	λ_{em}	527	547	545	569	630	610	590
	$\Delta\lambda_{ST}$	3855	4291	4327	5204	6202	6333	5624
	Φ_{fl}	0.65	0.80	0.54	0.58	0.72	0.59	0.43

^[a]Stokes shifts ($\Delta\lambda_{ST}$) are reported in cm^{-1} . ^[b] Φ_{fl} were measured using quinine sulfate in an aqueous 0.1 M HClO₄ solution ($\lambda_{exc} = 350$ nm, $\Phi_{fl} = 0.59$) or fluorescein in 0.1 M NaOH aq. solution ($\lambda_{exc} = 475$ nm, $\Phi_{fl} = 0.89$) as the references.¹³⁶

Table 2.8. Dipole moment differences between excited and ground states ($\mu_e - \mu_g$), in Debye for dyes **99-104** (primary and tertiary amino series).

<i>Fixed acceptor (Ar)/Varied donor (6-amino group)</i>							
Dye	99	100	101	102	103	104	
$(\mu_e - \mu_g)^{[a]}$	9.2	10.0	11.7	11.6	14.8	14.2	

^[a]Values are given in Debye (D).

To sum up, a series of styryl push-pull fluorophores, bearing only σ -acceptors and exhibiting a strong environment sensitivity were prepared in a practical synthetic route. While the absorption of these fluorophores was negligibly affected by polarity, all the reported dyes exhibited significant variations in their emission spectra in response to solvent polarity. The large positive solvatochromism observed in the emission spectra and the evidences obtained from computation calculations indicates the formation of ICT excited states. In addition, these fluorophores exhibit intense emission in polar and non-polar solvents, contrasting with the push-pull fluorophores bearing π -acceptor groups, which are, in their great majority, highly quenched in polar environments.

Nevertheless, despite all the positive and interesting findings, an improvement on the absorption properties of this family of fluorophores (*e.g.* redshift on the λ_{abs} through the visible region) remains an important point to evaluate and access dyes featuring advanced spectroscopic properties for (bio)applications. As a remark, with exception of dyes **98**, **99** and **104**, bearing the 1-pyrenyl framework, which exhibit intense absorption bands in the visible region ($\lambda_{\text{abs}} = 410 - 453 \text{ nm}$; $\epsilon = 54.2 - 65.9 \text{ (M.cm)}^{-1}$), the great majority of the reported fluorophores display absorption bands in the limit between ultraviolet and visible regions. To circumvent this, a simple extension on the π -conjugated core, without changing the major features of the fluorophore backbone, could be sufficient to displace the absorption maxima to longer wavelengths.

2.4 Benzothiazole-Based Fluorophores Bearing Extended π -Conjugated Spacer

In continuation to of efforts to produce innovative environment responsive fluorophores, we decided to investigate a new molecular platform. To obtain dyes with improved photophysical properties (as highlighted above), a second series of push-pull fluorophores has been developed. This second series was designed aiming to access fluorophores with improved absorption properties (λ_{abs} and ϵ) compared with the styryl series presented before. Therefore, this new family of dyes was engineered to contain an extended alkene-aryl-alkyne spacer as the π -conjugated linker to assembly both, the modified 6-*N*-ethyl-*N*-butylaminobenzothiazole (D) and the terminal aryl (A) motifs (**Figure 2.15**). Phenyl, phenyl rings bearing electron withdrawing groups (4-methylbenzoate and 4-formylphenyl) as well as polycyclic aromatic hydrocarbons (9-phenanthryl, 4-biphenyl and 1-naphthyl) were selected as acceptor moieties to build up the large charge transfer character of the dye excited state. In parallel, we expected that the replacement of the single vinyl linker of the previous family with an extended π -conjugated spacer could be beneficial in two aspects: (i) to redshift the absorption maxima and (ii) enlarge the degree of charge separation between the donor and the acceptor groups in the excited state to enhance the dye sensitivity to solvent polarity. This study also includes the comparison of the electronic effect of three distinct aryl spacers (1,4-substituted phenyl and 2,5 substituted furane and thiophene) on the spectroscopic properties of the dyes bearing the -COOMe acceptor group (**Figure 2.15**). Finally, we also performed a comprehensive and comparative photophysical study concerning this second series by UV-vis absorption and steady-state fluorescence emission spectroscopy.

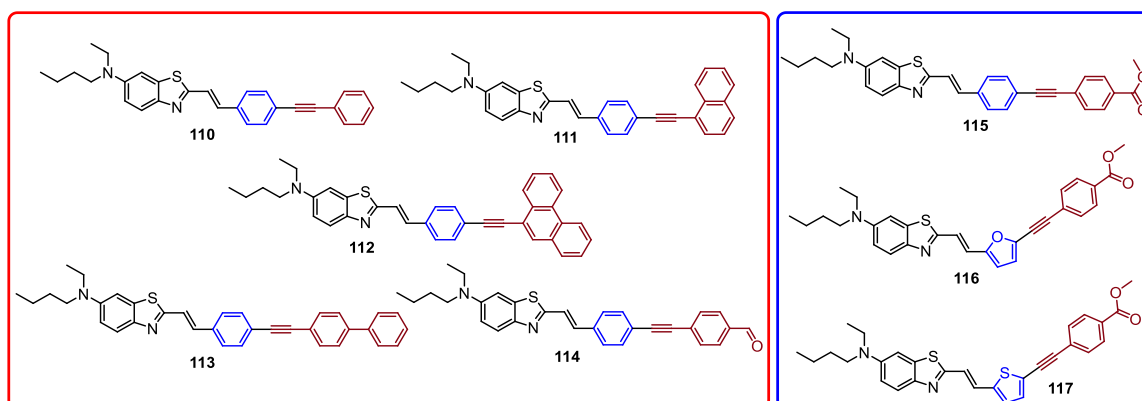
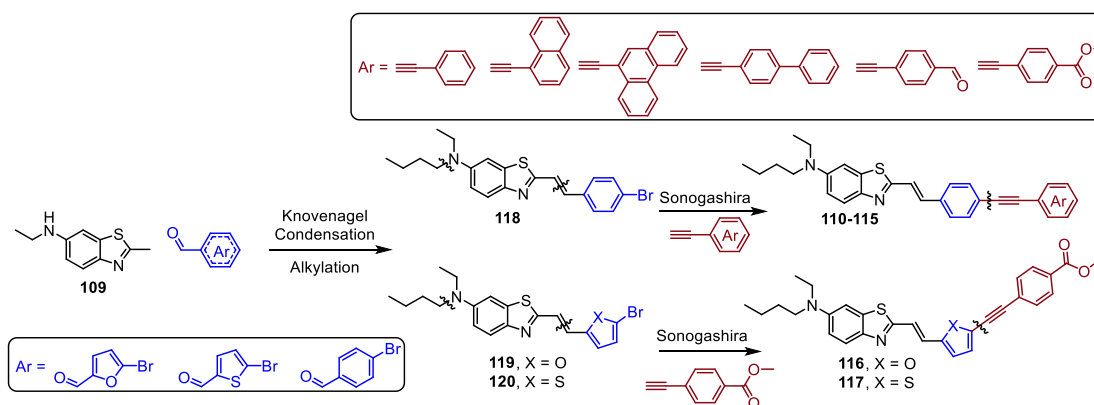


Figure 2.15. Chemical structure of the benzothiazole-based push-pull dyes **110-117** studied in this section.

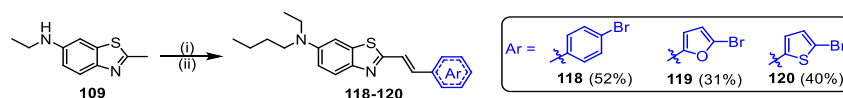
2.4.1 Synthesis of Benzothiazole-Based Fluorophores Bearing Extended π -Conjugated Spacer

The linear and three steps synthetic route used to access the dyes **110-117** is outlined in **Scheme 2.5**. Following a similar protocol as previously used, three precursor dyes (**118-120**) were firstly synthesized through a classical Knoevenagel condensation using aromatic aldehydes (4-bromobenzaldehyde, 5-bromo-2-thiophenecarboxaldehyde and 5-bromo-2-furanicarboxaldehyde) with the benzothiazole derivative **109**. Then, the N^6 -ethylamino groups were directly alkylated with 1-iodobutane. Finally, these three intermediates (**118-120**) were subjected to a series of Sonogashira cross-coupling reactions with six different alkynylarenes to afford the fluorophores **110-117**.



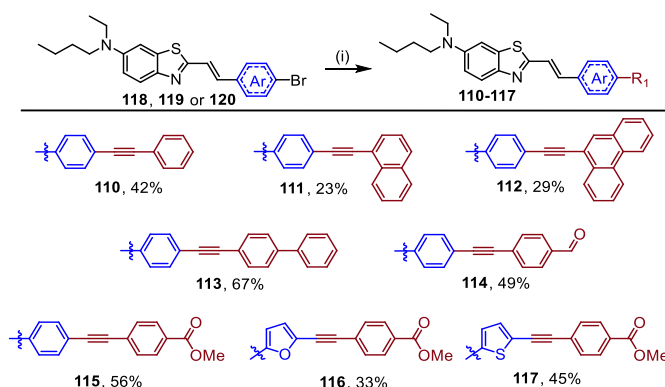
Scheme 2.5. synthetic pathway for the synthesis of fluorophores **110-117**.

The preparation of **118-120** was performed via a *one-pot* two steps synthetic route (**Scheme 2.6**). Firstly, the activated benzothiazole **109** was stirred in presence of 1.3 equivalents of the corresponding aldehyde at room temperature, in *N,N*-dimethylformamide, using a stoichiometric amount of potassium hydroxide. After complete conversion of the starting material, the base was filtered off and 1-iodobutane and stoichiometric amount of potassium carbonate were added. The mixture was then heated to 80°C overnight to afford the corresponding dyes **118-120** in satisfactory yield (31-52 %) and purity after purification by chromatography.



Scheme 2.6. Synthesis of the precursor bromo containing fluorophores **118-120**. Reagents and conditions. (i) aromatic aldehyde (1.3 equiv.), KOH, DMF, r.t., overnight; (ii) 1-iodobutane, K₂CO₃, DMF, 80°C, overnight.

With the bromide precursor dyes **118-120** in hands we proceeded to their connection to various aryl scaffolds via a Sonogashira cross-coupling reaction. Arylacetylene structures used for the coupling include phenylacetylene, substituted phenylacetylenes (4-biphenyl, 4-formyl and 4-carboxymethyl), and larger frameworks such as 1-ethynynaphthalene and 9-ethylphenanthrene. The target dyes, **110-117**, were obtained in moderate yields (23-67 %) after purification by preparative thin layer chromatography (**Scheme 2.7**).



Scheme 2.7. Synthesis of dyes **110-117**. Reagents and conditions. (i) aryl alkyne (1.5 equiv.), Pd(PPh)₄ (20 mol%), CuI (15 mol%), DMF:TEA (1:1), 50°C, 12 h.

2.4.2 Photophysical Properties of the Benzothiazole-Based Fluorophores Bearing Extended π -Conjugated Spacer

With these structurally characterized new compounds in hands (**110-117**) we proceeded to the investigation of their photophysical properties in solutions of cyclohexane, 1,4-dioxane, ethyl acetate, acetonitrile, dimethyl sulfoxide and methanol. **Figure 2.16** depicts the UV-Vis absorption and fluorescence emission spectra for dyes **110**, **111**, **113** and **115**. **Table 2.9** summarizes the main photophysical data. Additional spectra are presented in the *Annex section* (**Annex 2**).

It can be observed that all dyes present two intense absorption bands located in the visible (~ 419 - 446 nm) and in the near ultraviolet (~ 318 - 356 nm) regions. These may be associated with the lowest energetic $S_0 \rightarrow S_1$ and to upper excited states $S_1 \rightarrow S_n$ electronic transitions, respectively. In addition, high molar extinction coefficient values ($\sim 24.6 - 39.7 \times 10^4 \text{ M}^{-1} \cdot \text{cm}^{-1}$ and $35.0 - 53.3 \times 10^4 \text{ M}^{-1} \cdot \text{cm}^{-1}$) are observed for both absorption band maxima, in agreement with $\pi-\pi^*$ transitions. Interestingly, both absorption bands present very weak charge transfer character as a moderate solvatochromic effect is evidenced for all dyes when the solvent polarity is increased ($\sim 17 \text{ nm}$).

Regarding the absorption spectra, the maxima in the near-UV (λ_{abs1} , **Table 2.9**) region is bathochromically shifted for molecules bearing 1-naphthyl and 9-phenanthryl scaffolds (**111** and **112**) when compared with dyes **110** (phenyl) and **113-115** (4-biphenyl, 4-formylphenyl and 4-methylbenzoate, respectively); by contrast, the main absorption band present similar maxima for the six derivatives (λ_{abs2} , **Table 2.9**). In addition, we compared, the absorption spectra of dye **110** with its parent styryl dye **85**; The maximum wavelength associated with the lowest energetic electronic transition (λ_{abs2}) of dye **110** was redshifted from 35 to 40 nm compared with **85**.

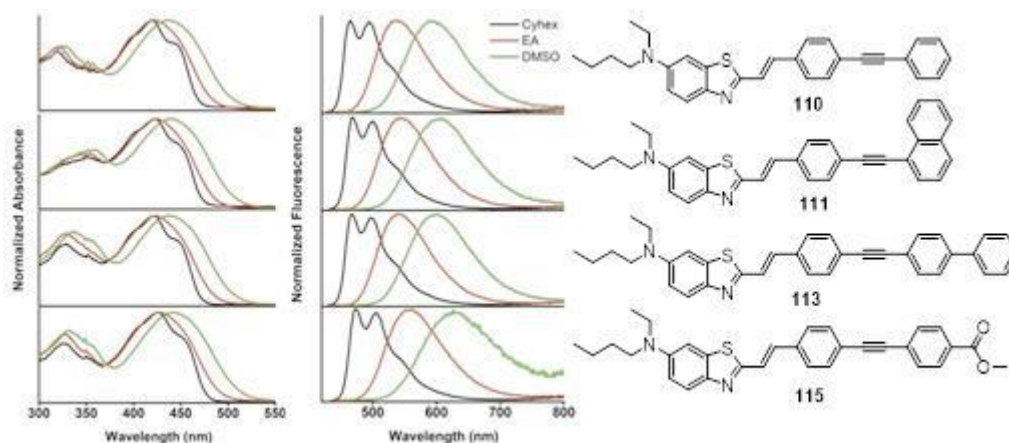
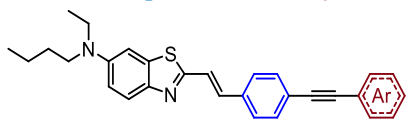


Figure 2.16. UV-Vis absorption and fluorescence emission spectra in solution of the dyes **110**, **111**, **113** and **115**.

Table 2.9. Relevant photophysical data from UV-visible absorption and fluorescence emission of dyes **110-115**.

Fixed spacer/Varied Aryl



Dye	Solvent	ϵ_1/ϵ_2	$\lambda_{\text{abs1}}/\lambda_{\text{abs2}}$	λ_{em}	$\Delta\lambda_{\text{ST}}^{\text{[a]}}$	$\Phi_{\text{fl}}^{\text{[b]}}$
110	Cyclohexane	27.5/42.2	318/419	465/495	3664	0.53
	1,4-Dioxane	28.0/40.4	321/424	516	4205	0.76
	Ethyl acetate	27.4/38.8	322/421	539	5200	0.64
	DMSO	24.6/35.0	326/436	593	6072	0.93
	MeCN	28.2/38.9	322/426	578	6173	0.63
	MeOH	26.2/35.8	322/429	563	5548	0.58
111	Cyclohexane	25.5/47.3	342/422	468/500	3696	0.60
	1,4-Dioxane	27.1/46.3	349/427	518	4114	0.71
	Ethyl acetate	26.8/44.4	344/424	545	5236	0.70
	DMSO	25.6/40.0	356/439	603	6195	0.94
	MeCN	28.7/44.7	345/426	584	6350	0.76
	MeOH	28.0/41.2	350/429	567	5673	0.54
112	Cyclohexane	23.8/40.4	346/424	470/503	3704	0.86
	1,4-Dioxane	24.4/39.0	348/427	523	4298	0.92
	Ethyl acetate	25.5/40.1	350/425	548	5281	0.56
	DMSO	25.9/36.1	355/440	609	6306	0.88
	MeCN	27.4/39.0	350/430	590	6306	0.79
	MeOH	27.2/38.9	352/432	573	5696	0.65
113	Cyclohexane	36.2/53.3	327/421	467/498	3672	0.75
	1,4-Dioxane	37.8/52.2	331/426	516	4094	0.75
	Ethyl acetate	39.4/49.8	327/423	541	5156	0.73
	DMSO	35.4/43.4	337/438	601	6262	0.99
	MeCN	39.7/47.9	329/426	581	6192	0.74
	MeOH	36.5/44.6	330/428	569	5789	0.78
114	Cyclohexane	32.0/46.9	327/431	479/513	3708	0.65
	1,4-Dioxane	33.0/44.4	331/434	534	4314	0.87
	Ethyl acetate	37.2/47.6	330/429	574	5888	0.63
	DMSO	33.8/41.2	338/446	611	6054	<0.001
	MeCN	39.7/47.2	330/432	590	6198	<0.001
	MeOH	32.1/38.2	332/433	563	5332	0.14
115	Cyclohexane	32.5/49.7	326/426	474/505	3672	0.82
	1,4-Dioxane	33.9/46.6	328/430	525	4208	0.95
	Ethyl acetate	35.1/46.8	327/427	558	5498	0.80
	DMSO	33.9/42.2	332/442	630	6751	0.31
	MeCN	39.5/46.9	328/431	604	6645	0.37
	MeOH	37.3/44.1	328/432	583	5995	0.26

^[a]Stokes shifts ($\Delta\lambda_{\text{ST}}$) are reported in cm^{-1} . ^[b] Φ_{fl} were measured using quinine sulfate in an aqueous 0.1 M HClO_4 solution ($\lambda_{\text{exc}} = 350 \text{ nm}$, $\Phi_{\text{fl}} = 0.59$) or fluorescein in 0.1 M NaOH aq. solution ($\lambda_{\text{exc}} = 475 \text{ nm}$, $\Phi_{\text{fl}} = 0.89$) as the references.¹³⁶

The fluorescence emission spectra for each dye were obtained using the relevant absorption maxima (λ_{abs1} and λ_{abs2}) as the excitation wavelengths. Differently from the absorption spectra, all dyes present remarkable redshifted emission maxima when going from the least apolar aprotic solvent to the most polar solvent (**Table 2.9**, **Figure 2.16**). This positive solvatochromism emphasizes the large charge transfer character of the dyes excited states. Among all dyes, compound **115**, bearing the strong 4-methylbenzoate EWG, presents the highest solvent sensitivity, with $\Delta\lambda_{\text{ST}}$ values up to 6751 cm^{-1} in DMSO. The range of emission for the dyes bearing non functionalized aryl acceptors were virtually the same, considering all probed solvents. Additionally, whatever the excitation wavelength (λ_{abs1} or λ_{abs2}), the same emission spectra was obtained (*Annex Section*, **Annex 3**). Thus suggesting that the rate of internal conversion from the upper S_n excited states to S_1 is significantly fast and fluorescence emission takes place only from the relaxed S_1 excited state. High Φ_{fl} values were generally observed for all dyes, except for the molecules bearing π -electron acceptor groups (**114** and **115**). In such cases, an abrupt decrease in the emission intensity was observed in solvents with increased polarity. It is known that the reduced energy gap between the excited and ground states in polar solvents may favored internal conversion, hence creating a non-radiative tunnel to release energy. However, the particular behavior observed for dye **114**, bearing the strong electron withdrawing formyl group perhaps indicates that the formation of a non emissive twisted intramolecular charge transfer state (TICT) may be present. Essentially, changes on the molecular geometry associated with an internal rotation within the fluorophore can usually provide an increase in the rate of the non-radiative decays, hence causing the fluorescence quenching. Many examples of solvatochromic fluorophores displaying either weak fluorescence emission or an huge drop of the Φ_{fl} values in high polar solvents have been connected with the formation of a putative TICT excited state.^{142,143} Finally, it should also be noticed that an anomalous emission in methanol (blue shifted compared to ethyl acetate) was observed for dye **114**. This may be due to a conversion of the formyl group into an (hemi)acetal moiety, thus reducing its acceptor character.

The positive solvatochromism was further investigated using the Lippert-Mataga model (**Figure 2.17** and **Annex 4**). The Stokes shift was plotted as a function of the orientation polarizability (Δf) in aprotic solvents. In all cases, linear relationships were obtained. These results support the hypothesis of an increase in the dipole moment of the dyes in the excited states.

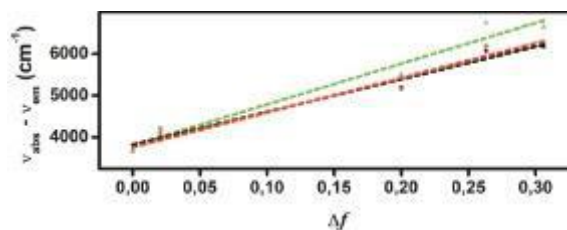
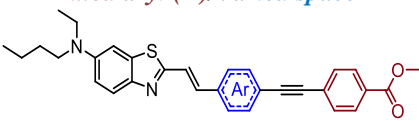





Figure 2.17. Dependence of the Stokes shift ($\lambda_{\text{abs}2} - \lambda_{\text{em}}$, in cm^{-1}) on the Lippert-Mataga orientation polarizability parameter for dyes **110** (black), **113** (red) and **115** (green). Correlation factors R^2 are, respectively, 0.97, 0.97 and 0.96.

To complement this study, the dyes **116** and **117**, containing the 2,5-substituted furane and thiophene groups were analyzed to verify the effect of different (hetero)aryl scaffolds on the π -conjugated spacer. Similarly to dyes **110-115**, compounds **116** and **117** also present the two characteristic absorption bands in the visible and near-UV regions. The visible band appeared bathochromically shifted compared with the analog dye **115** (439 - 455 nm for **116**; 447 - 464 nm for **117**, **Table 2.10**, **Figure 2.18**). In contrast, the band located in the near-UV region was largely affected by replacing the six member phenyl ring by the five-membered heteroaromatic rings. While **115** presents an absorption maxima centered around 325 nm ($\epsilon \sim 32.5 - 39.5 \text{ M}^{-1} \cdot \text{cm}^{-1}$), a significant red shift was observed for the dyes **116** and **117**; up to 43 nm in DMSO for the thiophene-containing derivative **117**. Otherwise, both dyes show weaker absorption associated to the more energetic $S_0 \rightarrow S_n$ transition when compared with dye **115** ($\epsilon \sim 21.7 - 25.4 \text{ M}^{-1} \cdot \text{cm}^{-1}$ for **116**; 20.6 - 29.5 $\text{M}^{-1} \cdot \text{cm}^{-1}$ for **117**). Concerning the fluorescence emission spectra, the same trend was also observed. The replacement of the phenyl ring by a thiophene moiety was also accompanied by a bathochromic shift ($\sim 30 \text{ nm}$) in the emission maxima for all solvents. Moreover, whatever the aryl group used to link the donor and the acceptor moieties, a strong reduction in the fluorescence emission intensity was observed in high polar solvents, especially in protic solvents. Of note, the three dyes (**115-117**) present absorption maxima suited for selective excitation with a blue diode laser (445 nm).

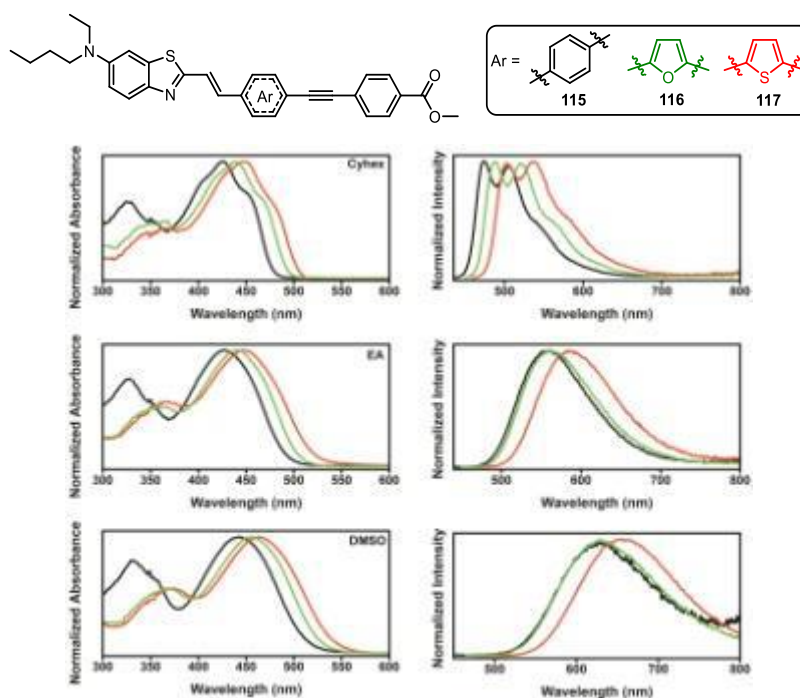
Table 2.10. Relevant photophysical data from UV-visible absorption and fluorescence emission of dyes **115-117**.

Fixed aryl (A)/Varied spacer



Dye	Solvent	ϵ_1/ϵ_2	$\lambda_{\text{abs1}}/\lambda_{\text{abs2}}$	λ_{em}	$\Delta\lambda_{\text{ST}}^{[a]}$	$\Phi_{\text{fl}}^{[b]}$
115 Ar = 	Cyclohexane	32.5/49.7	326/426	474/505	3672	0.82
	1,4-Dioxane	33.9/46.6	328/430	525	4208	0.95
	Ethyl acetate	35.1/46.8	327/427	558	5498	0.80
	DMSO	33.9/42.2	332/442	630	6751	0.31
	MeCN	39.5/46.9	328/431	604	6645	0.37
	MeOH	37.3/44.1	328/432	583	5995	0.26
116 Ar = 	Cyclohexane	21.7/47.9	354/439	489/522	3621	0.66
	1,4-Dioxane	24.8/49.6	362/444	535	3830	0.82
	Ethyl acetate	24.4/47.8	358/440	563	4965	0.82
	DMSO	23.2/42.0	364/455	630	6105	0.24
	MeCN	25.4/46.4	358/441	607	6201	0.29
	MeOH	22.5/42.2	358/446	584	5298	0.14
117 Ar = 	Cyclohexane	20.6/47.8	362/447	503/535	3679	0.78
	1,4-Dioxane	24.6/49.3	368/451	556	4187	0.79
	Ethyl acetate	29.5/51.7	366/448	586	5256	0.62
	DMSO	23.1/42.1	375/464	656	6307	0.43
	MeCN	25.8/47.8	367/451	633	6375	0.43
	MeOH	24.9/43.8	364/453	614	5788	0.38

^[a]Stokes shifts ($\Delta\lambda_{\text{ST}}$) are reported in cm^{-1} . ^[b] Φ_{fl} were measured using quinine sulfate in an aqueous 0.1 M HClO_4 solution ($\lambda_{\text{exc}} = 350 \text{ nm}$, $\Phi_{\text{fl}} = 0.59$) or fluorescein in 0.1 M NaOH aq. solution ($\lambda_{\text{exc}} = 475 \text{ nm}$, $\Phi_{\text{fl}} = 0.89$) as the references.¹³⁶

**Figure 2.18.** UV-Vis absorption and fluorescence emission spectra of dyes **115** (black), **116** (green) and **117** (red) in three representative solvents (Cyhex, EA and DMSO).

The push-pull interplay in the ground and excited states of dyes **110-117** was further investigated by studying their absorption and emission spectra upon addition of trifluoroacetic acid (TFA) in 1,4-dioxane solutions. For this purpose, **115** was selected as the benchmark fluorophore. The protonation is expected to take place at the 6*N*-butyl-*N*-ethyl-6-amino group, thus reducing its electron-donor character. Experimentally, the addition of TFA led to the disappearance of the two absorption bands, that were centered at 328 and 430 nm and the appearance of a new maxima at 367 nm, (**Figure 2.19**). This may indicate a moderate decrease in the charge transfer in the ground state. By contrast, this acidification resulted in a nearly complete quench of the ICT fluorescence emission band, with the concomitant appearance of a new band located around 425 nm.

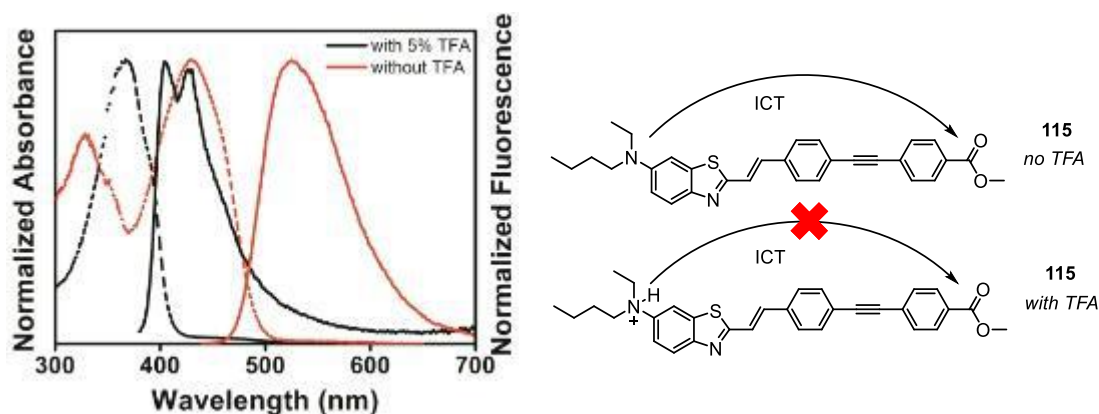


Figure 2.19. Normalized UV-vis absorption (dash) and fluorescence emission (solid) spectra for dye **115** in absence (red) and presence (black) of 5% of trifluoroacetic acid in 1,4-dioxane.

Finally, taking into account that solvatochromic dyes exhibiting excited states with strong ICT character are normally quenched in water, hydration studies were performed to investigate the sensitivity of our dyes to H₂O. Dyes **113** and **115** were selected to probe the H₂O effect on dyes bearing σ and π acceptors. **Figure 2.20** depicts the emission spectra of dyes **113** and **115** in 1,4-dioxane with increased amount of water (0-50%, v:v). Firstly, it is possible to observe that as the water content augments, the emission maxima are gradually redshifted. This is a direct consequence of the polarity increase. Additionally, this bathochromic shift is always accompanied by a substantial drop in the fluorescence intensity (both for **113** and **115**). Among the two probes, **115** seems to be considerably more sensitive to hydration. Both molecules present a significant reduction on their fluorescence emission

intensities in presence of 10% of H₂O (18 and 25% of reduction for **113** and **115**, respectively). However, while **113** emission is reduced by 45% in mixtures containing more than 50% of water, the emission of **115** was almost completely suppressed in such conditions. Furthermore, a perfect linear relationship between fluorescence intensity and water content was obtained for dye **115**.

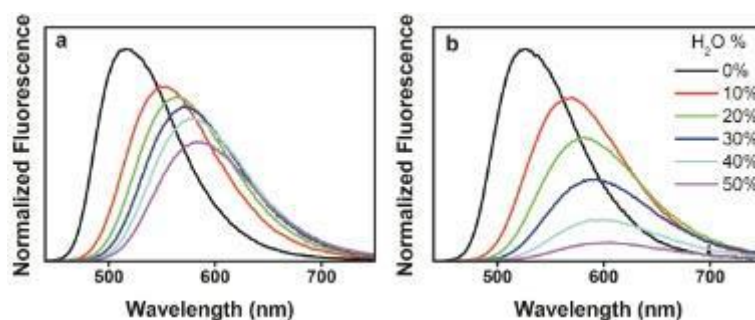


Figure 2.20. Fluorescence emission spectra of dyes **113** (a) and **115** (b) in incremental addition of H₂O in 1,4-dioxane (0 - 50%, v:v). Conditions: dye concentration, 1 μ M.

2.5 Conclusion - Chapter 2

In the first section of this chapter, we presented the efficient preparation of a series of 20 styryl-based solvatofluorochromic dyes starting from 6-amino-2-methylbenzothiazole. The dyes were synthesized in satisfactory overall yield, using a simple and practical synthetic route starting from commercially available and inexpensive materials. The ease of *N*⁶-alkylation renders this position ideal as a prospective anchoring point for further (bio)conjugation. Spectroscopic studies coupled with DFT calculations enabled the understanding of the relationship between structure and photophysical properties of the presented dyes in terms of electronic properties of the acceptors (dyes **85–98**) and inductive effects associated with multiple alkylation of the 6-amino group (dyes **99–104**). The dyes proved to be highly emissive in polar and non-polar solvents. They are characterized by the formation of an intramolecular charge transfer excited states as corroborated by their large positive solvatochromic emission bands. High Stokes shifts, up to 192 and 255 nm, could be accessed for the phenyl and the polycyclic aromatic series, respectively. This highlights that the commonly used π -acceptors groups can be efficiently replaced by purely σ -acceptors to broadly tune the emission wavelength towards the red region. Finally, for dyes containing para-substituted phenyl rings, the perfect correlation between Hammett parameter (σ_p) and

Part I - Chapter 2

Stokes shifts and λ_{em} provides a complementary approach to anticipate the photophysical properties of other members of this family of fluorescent probes. Using this correlation and the subtle changes brought by the 6-amino alkylation, specific fluorescent tools that absorb and emit in a particular spectral range can be easily designed.

In the second section of this chapter, we presented a second series of fluorophores harbouring an extended π -conjugated scaffold (**110-117**). This second series could be accessed starting from a bromo-containing styryl derivative via a single Sonogashira cross coupling reaction in moderate yields. Interestingly, these dyes were characterized by two intense absorption bands (UV and visible) and single and redshifted emission spectra. The photophysical studies highlighted that the dyes bearing π -electron acceptor groups were highly quenched in high polar environment, suggesting the formation of TICT states.

Chapter 3 - Development of highly sensitive fluorescent probes for the detection of β -galactosidase activity - application to the real-time monitoring of senescence in live cells

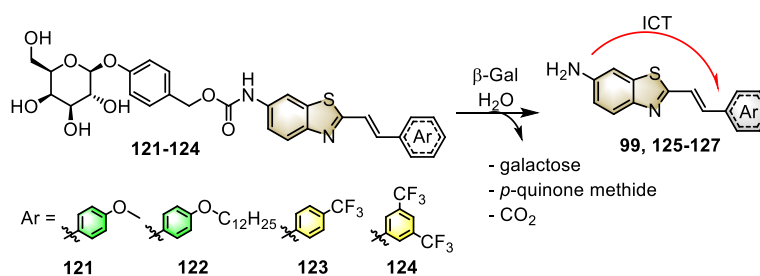
Chapter Outline

Fluorescence-based microscopy are non-invasive and non-destructive techniques that emerged as powerful tools for many areas of research. In particular, the combination of fine-tuned, photostable and bright dyes with the last generations of optical devices have allowed the fast development of different areas of biomedical research, from medical diagnostics to *in cellulo* and *in vivo* real-time imaging to study composition and dynamics of cellular systems.⁵⁵

Cell senescence is a phenomenon deeply related to cellular aging. It is a cytostatic (or dormant) state in which cells cease to divide. Characteristically, senescent cells display increased β -galactosidase activity compared with normal cells. This phenomenon, which is known as senescence-associated β -galactosidase activity (SA- β -gal) results from the overexpressed and accumulated levels of endogenous β -gal proteins produced in the lysosome of senescent cells. Although the β -gal expression is not a requirement to enter in a senescent state, the importance of this SA- β -gal activity has been reinforced as it is considered as an important marker used to detect cellular senescence. Thereby, the research of novel β -gal assays, based on optical techniques and applied to cellular senescence, intensified during the last years.

In this context, this chapter will states our work concerning the development of new and engineered fluorescent probes for the monitoring and visualization of enzymatic activity of β -galactosidase during cellular senescence. Firstly, a short introductory section will disclose the concepts of cellular senescence, β -galactosidase activity as well as the use of therapy induced senescence (TIS) in medicinal chemistry. In the second part, the synthesis of four new fluorescent probes, **121-124**, will be presented (*Scheme 3.1*). The probes were designed by combining solvatofluorochromic styryl-benzothiazole fluorophores (presented in **Chapter 2**) with an enzyme-labile β -D-galactopyranoside (trigger moiety), through a self-immolative benzyl carbamate spacer (*Scheme 3.1*). Here, we anticipate that β -gal-mediated hydrolysis of the β -galactoside substrate would trigger the decomposition of the carbamate linkage and

liberate the naked amino group, thereby restoring the ICT character of the excited state, redshifting the emission maxima (*Scheme 3.1*). Importantly, we envisioned the access of probes with distinct photophysical properties (emission in different spectroscopic regions). Thus, four different phenyl derivatives harbouring EDG and EWG were selected as the acceptor moieties (*p*-OMe, *p*-OC₁₂H₂₅, *p*-CF₃ and 3,5-bis-CF₃) in order to modulate the emission maxima of these probes. Finally, following the preparation of these four probes, we will discuss their spectroscopic characterization and report the *in vitro* (enzymatic) and *in cellulo* evaluation of these dyes using fluorescence microscopy and flow cytometry methods.



Scheme 3.1. β -gal probes **121-124** investigated in this work.

3.1 Introduction - Therapy Induced Senescence (TIS) and senescence-associated β -galactosidase (SA- β -gal)

β -galactosidase (β -gal) is a glycoside hydrolase enzyme responsible for the hydrolysis of β -galactosides into monosaccharides by breaking the glycosidic bond. It has been widely used as the key marker enzyme for *in vivo* tumor growth monitoring,¹⁴⁴ transcriptional regulation,¹⁴⁵ gene expression¹⁴⁶ and importantly for specific characterization of cellular senescence.¹⁴⁷ Cell senescence is a persistent cytostasis which results from a non-lethal stress. It is characterized by a morphological and biochemical phenotype that includes an abnormally accumulated β -gal activity called senescence-associated β -galactosidase (SA- β -gal). This SA- β -gal activity is a widely accepted marker of cellular senescence.¹⁴⁷

In recent years, therapy-induced senescence (TIS) attracted a lot of attention from both academic and industrial research groups interested in the development of new anti-cancer therapies. In fact, TIS appeared as a new potential mean to treat cancer by inducing the entrance of cancer cells into a cytostatic state.¹⁴⁸ One of the main assets of TIS, as reported in several studies,^{149,150} is that cancer cells resistant to chemotherapy remain sensitive to the induction of senescence. This suggests that TIS could be widely used regardless of the

resistance state of the tumors. Furthermore, activation of the immune system by senescent cells is involved in tumor clearance.¹⁵¹ Thus, TIS may provide an effective way to fight cancer. Unfortunately, only a few compounds have been identified as senescence inducers. Indeed, a limiting factor in the efficient discovery of senescence inducers has been the lack of powerful methods for high-throughput screening.

The most common method to visualize SA- β -gal is the widely used **X-Gal** assay.¹⁴⁷ **X-Gal**, 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (**Figure 3.1**), is a synthetic substrate of β -gal which yields a blue precipitate after its processing. In parallel, different β -galactosidase fluorescent probes have also been developed (*Section 1.8.2, Chapter 1*). However, most of the fluorophores used for this purpose exploit the conventional coumarin, fluorescein¹⁵² or the fluoresceine-rhodamine hybrid rhodol dyes as fluorescent reporters (*e.g.* the commercially available 4-methyl-umbelliferyl β -D-galactopyranoside - **4-MUG** - and fluorescein di- β -D-galactopyranoside - **FDG** - derivatives; **Figure 3.1**). Although featuring extreme sensitivity due to the brightness of the generated fluorescent products, these probes present slow kinetic response and sharp Stokes shift. Moreover, they are also unsuitable for live-cell imaging since they need to be used on fixed cells that underwent a pH-modulation step before staining.

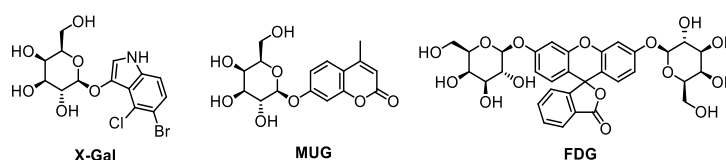


Figure 3.1. Chemical structures of commercially available probes used for the visualization and detection SA- β -gal activity.

In the last years, sustained efforts have been produced to overcome some of these issues. However, despite the significant advances with the design of many novel fluorescent probes to the detect of β -gal activity (**Chapter 1, Section 1.8.2**), no reports concerning the real-time monitoring and detection of early to middle stages of cellular senescence have been considered. In 2014, Kim and co-workers⁸⁰ partially addressed this issue. With their blue/yellow ratiometric fluorescent probe **53**, they could report and monitor cellular changes during middle stages of senescence *via* confocal microscopy in model senescent cells (see **Figure 3.2**). Therefore, inspired by Kim's work, and based on our studies dealing with amino containing styryl-benzothiazole fluorophores, we present herein the synthesis, photophysical

characterization and *in vitro* and *in cellulo* evaluation of our novel probes **121-124** to report SA- β -gal activity.

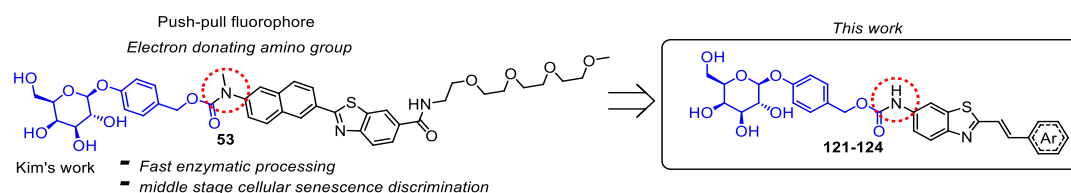
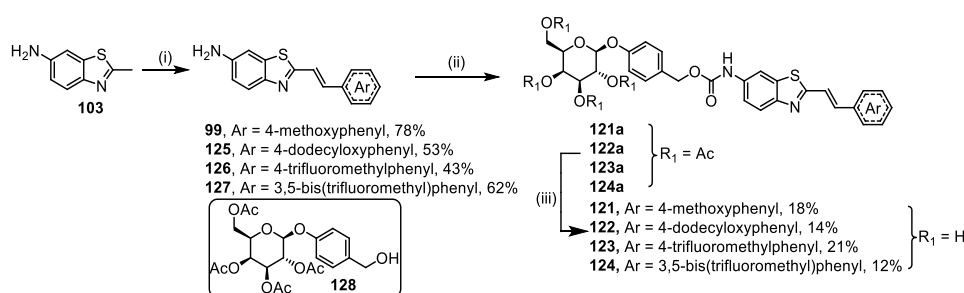


Figure 3.2. Molecular structure of β -gal probes reported by Kim's group and by us.

3.2 Synthesis of Probes 121-124

The synthetic route to the probes **121-124** is depicted on *Scheme 3.2*. Starting from the 2-methyl-6-aminobenzothiazole **107**, the corresponding dye reporters **99**, **125-127** were obtained in one step *via* the condensation reaction of **107** with the respective substituted benzaldehyde. These fluorescent dyes intermediates were obtained in a moderate to excellent yields (43 - 78%) Then, the preparation of acetylated **121a-124a** probes was performed in *one-pot* two-step fashion. Firstly, fluorophores **99**, **125-127** were converted to the corresponding carbamoyl chlorides by treatment with an equimolar amount of triphosgene, in presence of triethylamine. After complete conversion of the starting material, the intermediate was treated with the benzyl alcohol **128** (*Scheme 3.2*),¹⁸ to afford the acetylated probes **121a-124a**. Finally, probes **121-124** were obtained in moderate yields (12-21%) after cleavage of the acetyl groups in presence of sodium methoxide in methanol.



Scheme 3.2. Synthesis of probes **121-124**. Reagents and conditions: (i) benzaldehyde derivative (1.2 equiv.), KOH (1 equiv.), DMF, r.t., 12 h. (ii) a) triphosgene (1 equiv.), triethylamine (3 equiv.), 1,4-dioxane, r.t., 1 h. b) **128** (1 equiv.), r.t., overnight. (iii) MeONa (5 equiv.), MeOH, r.t., overnight.

3.3 Spectroscopic Properties

Prior to the investigation concerning the β -gal activity detection, we briefly evaluated the photophysical properties of the four fluorescent reporters (**99**, **125-127**) and the corresponding probes **121-124** in four different systems (1,4-dioxane, ethyl acetate, dimethylsulfoxide and PBS containing 20% of DMSO; **Figure 3.3**, **Table 3.1**). Interestingly, two different sets of behavior could be observed within this probe family. Firstly, dyes **99**, **125-127** displayed intense emission in the visible spectrum when excited on the near-UV region (~ 380 nm). The emission maxima showed gradual redshift in solvents with increased polarity. High Φ_{fl} values were obtained in organic solvents, with no strong dependency of the solvent polarity. Dyes **126** and **127**, bearing the trifluoromethylated groups, show higher level of solvent dependency compared with dyes **99** and **125**. Their emission maxima (λ_{em}) were 35 nm bathochromically shifted (560 vs. 525 nm) in mixtures of PBS/DMSO (8/2) compared with dyes **99** and **125**. It is known that fluorescent probes displaying ICT emission are usually quenched in water. However, dyes **99** and **126** are still relatively bright in highly aqueous environment ($\Phi_{fl} = 0.32$ and 0.30 , respectively in PBS/DMSO 8/2).

Conversion of the fluorescent reporters **99**, **125-127** to the corresponding probes (**121-124**) by the introduction of the carbamate moiety had different effects. In fact, the acylation of the $-NH_2$ group of dyes **99**, **124-127** strongly modulated their absorption and fluorescence emission maxima. The absorption maxima of **121-124** were blueshifted ($\lambda_{abs} = 361-370$ nm) compared with their respective dye reporters **99**, **125-127** ($\lambda_{abs} = 375-410$ nm). The same tendency was observed on the fluorescence emission maxima. Probes **121-124** displayed emission in the blue spectral region ($\lambda_{em} = 426-480$ nm), with no significant degree of solvent dependency. Interestingly, probes **121** and **122** bearing the donating groups on the acceptor side (p -OMe and p -OC₁₂H₂₅) were highly quenched in all probed solvents ($\Phi_{fl} < 0.005$). By contrast, probes containing the fluorinated groups (**123** and **124**) were found to display high fluorescence quantum yields ($\Phi_{fl} = 0.21-0.41$). However, they were poorly soluble in PBS buffer solution where they presumably form highly emissive aggregates. The formation of aggregates **123** and **124** was hypothesized as sequential acquisitions of fluorescence emission spectra of both probes revealed a continuous and time-dependent drop of the emission intensity as the putative aggregates did sediment. This observation was further confirmed as the centrifuged solution (13000 rpm, 10 min) was not fluorescent anymore; the aggregates being at the bottom of the cuvette.

Lastly, probe **122** and its corresponding dye reporter **125** were found to be too poorly emissive and soluble to be investigated further. Thereby, these first observations allowed us to potentially access a new set of ratiometric (**123** and **124**) and light-up (**121**) probes.

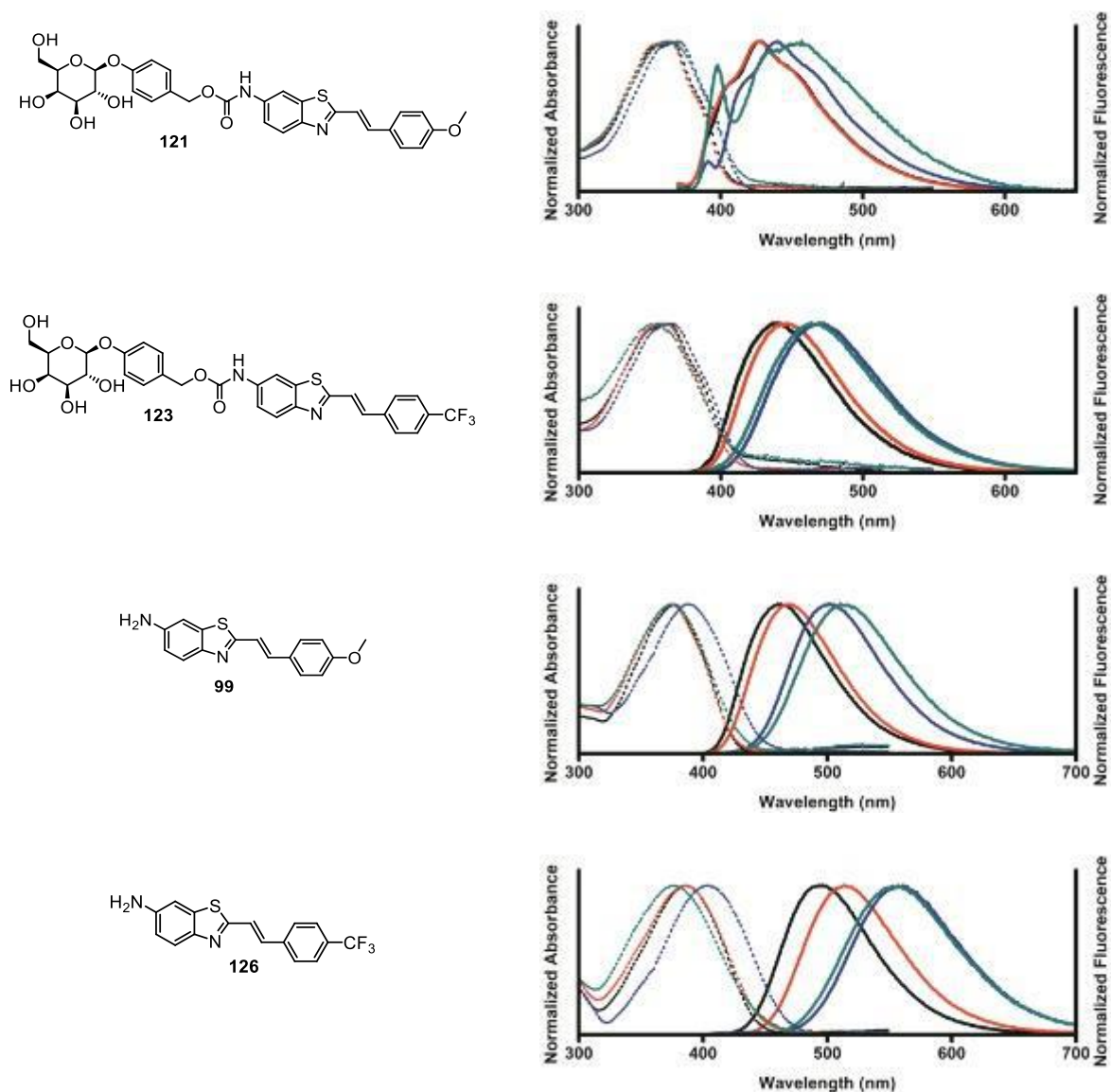


Figure 3.3. Normalized absorption and fluorescence emission spectra of probes **121** and **123** and the respective reporter fluorophores **99** and **126**, respectively, in 1,4-dioxane (black), ethyl acetate (red), DMSO (blue) and mixture PBS/DMSO (8/2) (dark cyan).

Table 3.1. Main photophysical data from the UV-visible absorption and steady-state fluorescence emission analysis of dyes **99**, **125-127** and probes **121-124**.

Compound	Solvent	λ_{abs} (nm)	λ_{em} (nm)	$\Delta\lambda_{\text{ST}}$ (nm/cm ⁻¹)	$\epsilon \times 10^3$ (M.cm) ⁻¹	Φ_{fl} ^[a]
99	1,4-dioxane	379	463	84/4786	37.7	0.45
	Ethyl acetate	374	466	92/5278	34.2	0.29
	DMSO	388	501	113/5813	31.0	0.52
	PBS	376	514	138/7140	29.4	0.32
125	1-4,dioxane	375	462	87/5021	37.0	0.59
	Ethyl acetate	375	470	95/5390	38.2	0.57
	DMSO	389	502	113/5786	35.4	0.61
	PBS	-	501	-	-	-
126	1-4,dioxane	385	494	109/5731	27.3	0.60
	Ethyl acetate	385	513	128/6480	28.5	0.52
	DMSO	403	560	157/6956	23.0	0.58
	PBS	376	555	179/8577	26.2	0.30
127	1-4,dioxane	391	508	117/5890	24.2	0.52
	Ethyl acetate	391	531	140/6743	26.7	0.45
	DMSO	410	577	167/7059	21.0	0.41
	PBS	380	565	185/8616	20.2	0.14
121	1-4,dioxane	361	427	66/4281	36.2	<0.005
	Ethyl acetate	362	426	64/4150	37.1	<0.005
	DMSO	370	439	69/4247	34.8	<0.005
	PBS	364	455	91/5494	33.4	<0.005
122	1-4,dioxane	363	428	65/4183	37.3	<0.005
	Ethyl acetate	363	428	65/4183	38.1	<0.005
	DMSO	370	439	69/4247	34.7	<0.005
	PBS	-	473	-	-	-
123	1-4,dioxane	361	440	79/4973	33.5	0.31
	Ethyl acetate	361	445	84/5228	30.2	0.41
	DMSO	366	469	103/6000	28.2	0.47
	PBS	354	466	112/6789	30.5	0.21
124	1-4,dioxane	363	440	77/4820	29.9	0.38
	Ethyl acetate	363	458	95/5714	33.2	0.39
	DMSO	366	480	114/6489	28.3	0.41
	PBS	-	475	-	-	-

^[a]Fluorescence quantum yields were measured using quinine sulfate in aqueous 0.1M HClO₄ solution ($\Phi_{\text{fl}} = 0.54$, $\lambda_{\text{exc}} = 350$ nm) as reference.

Next, we started to investigate the response of compounds **121**, **123** and **124** as effective reporters of β -gal activity. As a preliminary screening, fluorescence emission spectra of the probes in simulated physiological conditions (PBS buffer, pH = 7.4) were recorded in absence and after treatment with 1 U/mL β -gal for 10 min (**Figure. 3.4**).

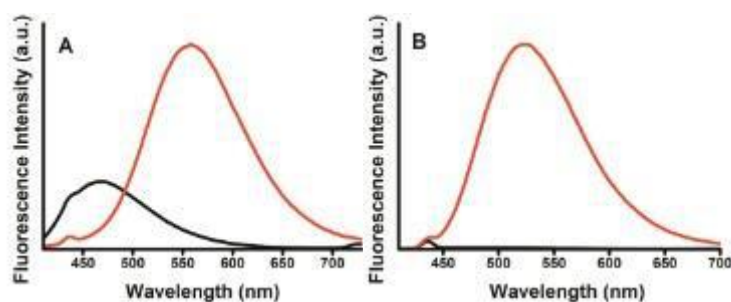


Figure 3.4. Probes **123** (A) and **121** (B) in absence (black) and presence of 1 U/mL of β -galactosidase (red) after 10 min in PBS. Conditions: probe concentration $\sim 0.2 \mu\text{M}$; $\lambda_{\text{exc}} = 380 \text{ nm}$.

Here, the expected light-up and two-band ratiometric system could be confirmed concerning the non-fluorinated **121** and fluorinated probes **123** and **124**, respectively. In fact, in the absence of enzyme, probes **123** and **124** display their characteristic fluorescence emission in the blue region (470 nm), corresponding to the β -D-galactopyranoside derived benzyl carbamate protected form (**123** and **124**). However, after 10 min in presence of 1 U/mL of β -gal, the emission at 470 nm was completely suppressed and the new yellow band centered at 560 nm appeared (fluorophores **126** and **127**). Nevertheless, while the reporter dye **126** displayed strong enhancement on the fluorescence emission intensity (12-fold related to **123**, **Figure 3.4**), compound **127** was poorly emissive compared to probe **124**.

In contrast, after treatment with β -gal, **121** displayed a spectacular 206-fold increase in the fluorescence emission intensity in the green spectral region (525 nm, fluorophore **99**). Thus, these findings suggest that **121** afforded the higher β -gal sensitivity among our four probes.

3.4 Kinetics

The kinetics of the β -gal-mediated conversion of probes **121** (to **99**) and **123** (to **126**) was investigated by analyzing the fluorescence intensity variations of the newly formed ICT bands (525 nm and 560 nm for **121** and **123**, respectively) as a function of time. The response of our probes was compared to the commercially available and commonly employed **FDG**. The emission spectra at 525 nm of **121** treated with 0.05 U/mL of β -gal in PBS buffer (pH 7.4, 37°C) continuously increased according to the progress of the enzymatic hydrolysis. Emission intensity reached its maximum after 18 min of reaction (**Figure 3.5-A**). Probe **123** displayed a similar kinetic response. The emission spectra of **123** increased gradually at 560 nm with a concomitant decrease at 470 nm (**Figure 3.5-B**). The ratio of the emission

intensities ($F_{\text{yellow}}/F_{\text{blue}}$) at 470 nm (F_{blue}) and 560 nm (F_{yellow}) increased by 66-fold upon β -gal processing of the probe; the maximum was reached after 20 min. Pseudo-first order kinetic constants were also accessed (**121**, 0.0045 s^{-1} and **123**, 0.0036 s^{-1}). It should be noted that the kinetic responses of both probes were much faster than that of **FDG** (**Figure 3.6**). Indeed, to attain similar kinetics, the **FDG** probe should be treated by 4 U/mL of a β -gal; *i.e.* a 80-fold increase in the enzyme concentration. This comparative experiment highlights the high β -gal sensitivity of probes **121** and **123** compared with **FDG**.

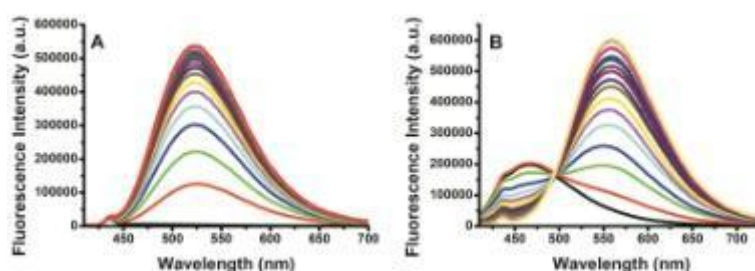


Figure 3.5. Fluorescence emission spectra of probes **121** (A) and **123** (B) in the presence of 0.05 U/mL of β -gal at different times (from 0-25 min). Conditions: probe concentration $\approx 0.2 \mu\text{M}$; $\lambda_{\text{exc}} = 380 \text{ nm}$.

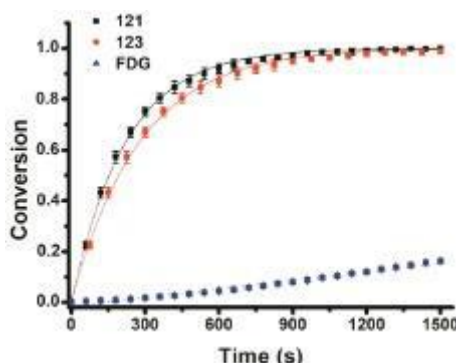


Figure 3.6. Comparison of the time course of the fluorescence intensity between probes **121**, **123** and **FDG** ($0.2 \mu\text{M}$) in PBS buffer (pH 7.4, 37°C) after addition of β -gal (0.05 U/mL).

3.5 Stability, Selectivity and Fluorescence Emission in Cell

Considering the superior sensitivity of probe **121** to report β -gal activity, we decided to study its applicability to biological systems (*e.g.* (live)-cell imaging). Thus, we started by evaluating its stability in PBS at different pH values (4.0, 6.0, 7.4 and 9.0). **121** showed high stability in moderately acidic (pH 4.0 and 6.0) and basic (pH 9.0) media as almost no residual

emission at 525 nm, related to the appearance of the non-galactosylated dye reporter **99**, was observed after 96 hours. (**Figure 3.7**).

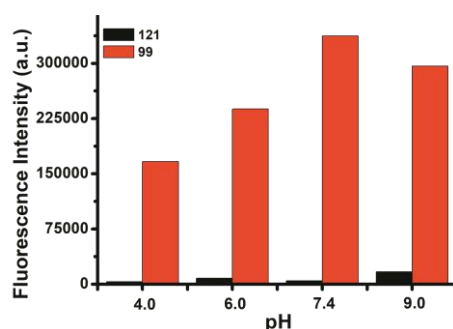


Figure 3.7. Fluorescence emission intensity of **121** (black) and reporter fluorophore **99** (red) in different pH after incubation during 96 hours. Condition: dye concentration $\approx 0.2 \mu\text{M}$; $\lambda_{\text{exc}} = 380 \text{ nm}$

Considering the fluorescence intensity of dye **99**, stronger emission was observed in neutral and basic conditions, as expected, while a moderate quenching was observed at pH 4.0. Moreover, **99** displayed high photostability (**Figure 3.8**).

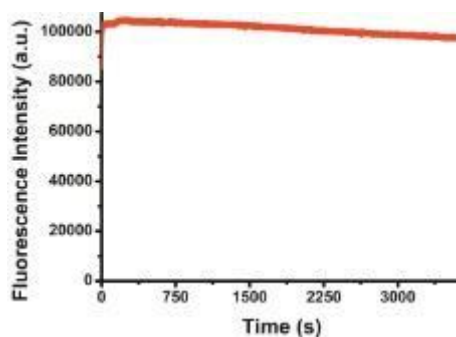


Figure 3.8. Fluorescence emission intensity of dye **99** under continuous irradiation in PBS solution during 1 hour. Condition: probe concentration, $0.2 \mu\text{M}$; $\lambda_{\text{exc}} = 405 \text{ nm}$.

Fluorescence emission spectra in A375 melanoma cell lines of **121** and its corresponding non-protected dye **99** were acquired (**Figure 3.9**). It was clearly observed that probe **121** displays exactly the same fluorescence profile as observed in the *in vitro* assays (**Figure 3.4 vs Figure 3.9**). In addition, no bioconversion of **121** into its fluorescent reporter **99** was evidenced. To ascertain the selective bioconversion of **121** by β -galactosidase enzyme, we treated **121** with increasing amounts of A375 cell lysates, supplemented or not with β -gal (**Figure 3.10**). No fluorescence signals (*i.e.* conversion of **121** into **99**) were recorded when

the A375 lysate was used without β -gal supplementation. Thus, this result strongly supports the fact that **121** is selectively processed by β -gal and that no interference with other cell constituents is occurring.

Therefore, probe **121** seems robust enough to be employed in the detection of β -galactosidase activity under biological conditions.

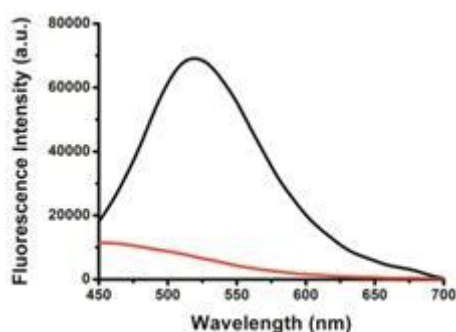


Figure 3.9. Fluorescence emission spectra of probe **121** (red) and the respective dye **99** (black) in A375 melanoma cells. Condition: probe concentration, 1.0 μ M; λ_{exc} = 405 nm.

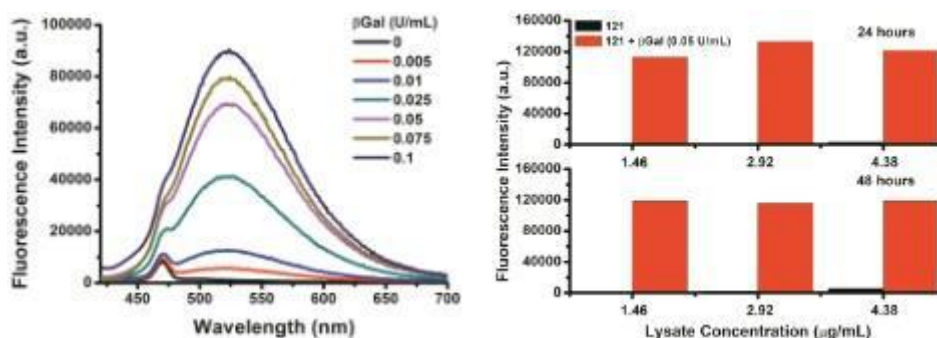


Figure 3.10. Fluorescence emission spectra of probe **121** in presence of cell lysate (1.46 μ g/mL) with different amounts of β -gal. Condition: probe concentration, 0.2 μ M; λ_{exc} = 405 nm (left). Fluorescence emission intensities recorded at λ_{em} = 525 nm of probe **121** in presence of different amount of cell lysate. Condition: probe concentration, 0.2 μ M; λ_{exc} = 405 nm (right).

3.6 Limit of Detection

To further evaluate the response sensitivity of probe **121** to the presence of β -gal, fluorescence titration experiments were performed. As shown in **Figure 3.11**, in the absence of β -gal, probe **121** exhibited very weak emission at 525 nm. With the increase of β -gal concentration from 0 to 0.05 U/mL, the emission band at 525 nm gradually rose. Notably, the

plot fitted linearly with a correlation coefficient of 0.99 in the β -gal concentration range of 0-0.05 U/mL, indicating that **121** is a competent probe for the quantitative detection of low levels of β -galactosidase. The detection limit was determined to be as low as $4.62 \pm 0.46 \times 10^{-5}$ U/mL, ($3\sigma/\kappa$, where σ is the standard deviation of blank measurements and κ is the slope) which was in the same range or considerably lower (1000-fold) than the previously reported light-up probes.^{59,77} From these findings, we inferred that probe **121** displayed readily detectable changes in its fluorescence signature at very low levels of β -gal (5×10^{-4} U/mL).

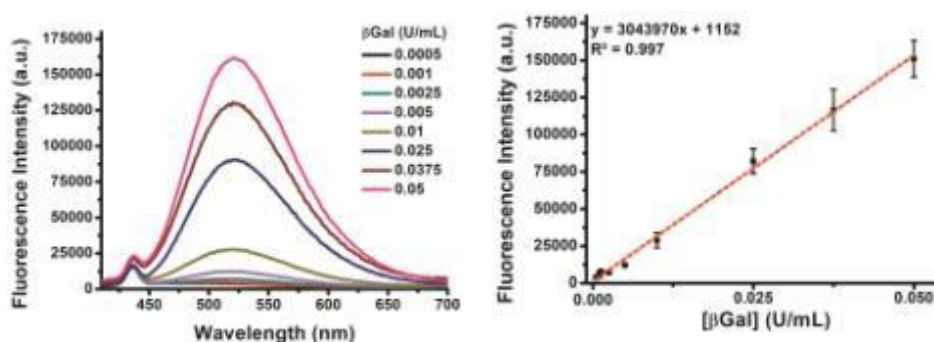


Figure 3.11. Fluorescence emission spectra of **121** in presence of different concentrations of β -gal (0, 0.0005, 0.001, 0.00125, 0.0025, 0.005, 0.01, 0.025, 0.0375, 0.05 U/mL) (left). Fluorescence emission intensity of **121** recorded at $\lambda_{\text{em}} = 525$ nm after 10 min in different concentrations of β -galactosidase in PBS (pH = 7.4) for three independent experiments. Condition: dye concentration ≈ 0.2 μM ; $\lambda_{\text{exc}} = 380$ nm (right).

3.7 Fluorescence Imaging of SA- β -gal in Living Cells

Therapy-induced senescence (TIS) represents a novel challenge in cancer drug discovery. This increasing interest in senescence-inducer compounds requires the exploitation of new tools for the development of high-throughput screening. Therefore, we next investigated whether **121** probe could be used to detect the endogenous senescence-associated β -galactosidase (SA- β -gal) on different induced-senescence cancer models.

Two cell lines were chosen for senescence-induced model imaging, including human melanoma A375 cells and human colorectal adenocarcinoma HT-29 cells. SA- β -gal is overexpressed endogenously in senescent cells, including therapy-induced senescence (such

as hydroxyurea (HU), H₂O₂, or others). For this study, we chose hydroxyurea to induce cell senescence and assess the ability of our probe **121** to monitor SA- β -gal in cells.

Thus, the two cell lines aforementioned were treated with HU for 96 h, washed with PBS and finally incubated with 0.1 μ M **121** for 30 min at 37°C. Then, the cells were washed and fixed with paraformaldehyde (PFA) before observation on confocal microscope. Importantly, compared with **X-gal** or **FDG** reporters that require a pH modulation step before staining, **121** can be used directly on cells; hence simplifying the staining process. Senescent cells displayed a remarkable increase in cell size (about 10-fold increase for HT-29 senescent-cells). This observation was correlated with the appearance of a strong fluorescence signal when compared with the non-senescent cells (**Figure 3.12**). In addition, the cytotoxicity of the fluorescent reporter has been evaluated and showed no toxicity; A375 showed 100% and 74.5% of cell viability after 48 h of incubation in presence of 0.1 μ M and 1.0 μ M of compound, respectively. Therefore, these results successfully demonstrated that our probes can be applied to visualize endogenous β -galactosidase activity in living cells.

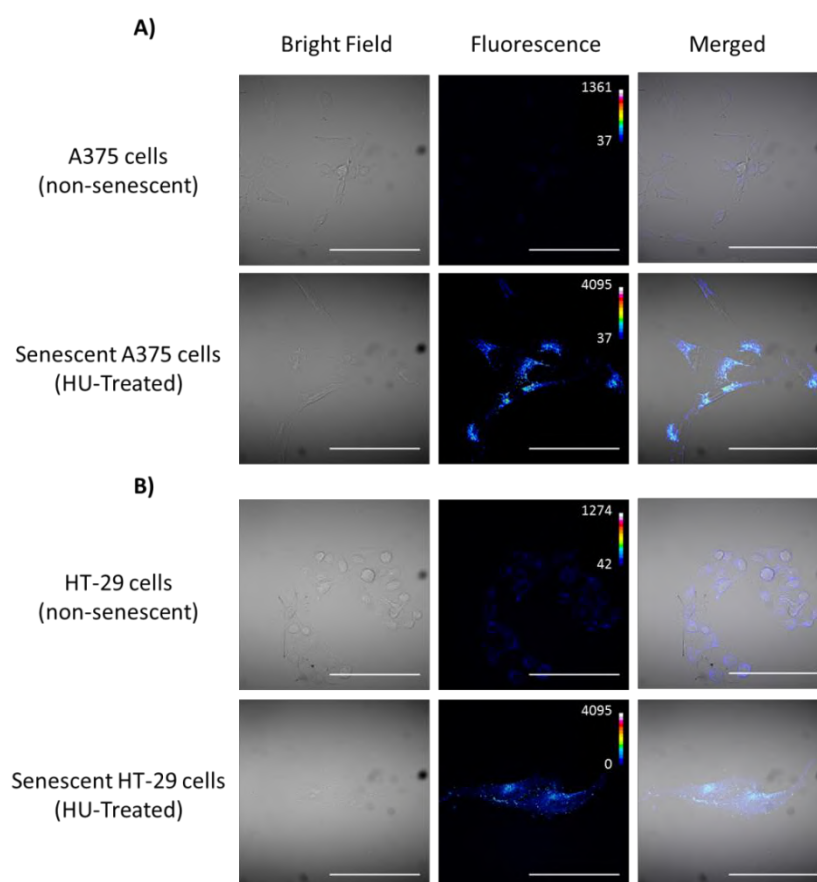


Figure 3.12. Fluorescence images of normal (HU non-treated) and senescent cells (50 μ M HU) incubated with 0.1 μ M **121**. A) human melanoma A375 cells; B) human colorectal

adenocarcinoma HT-29. Images were acquired using a confocal microscope with a 63X NA 1.4 objective; scale bars = 100 μm .

In addition to the experiment presented on **Figure 3.13**, where we monitored the SA- β -gal activity within already senescent cancer cells, we used **121** to monitor the establishment of senescence in live cell (A375 cell line, up to 72 h). This live-cell experiment could be performed because of the non-toxicity of our fluorescent reporter. In this experiment, we treated A375 cells with both **121** and HU and monitored the emergence of SA- β -gal activity by fluorescence time-lapse microscopy. As expected, HU-induced senescence was visually corroborated by the progressive increase in cell size from to 24 h to 72 h (**Figure 3.13-A**).

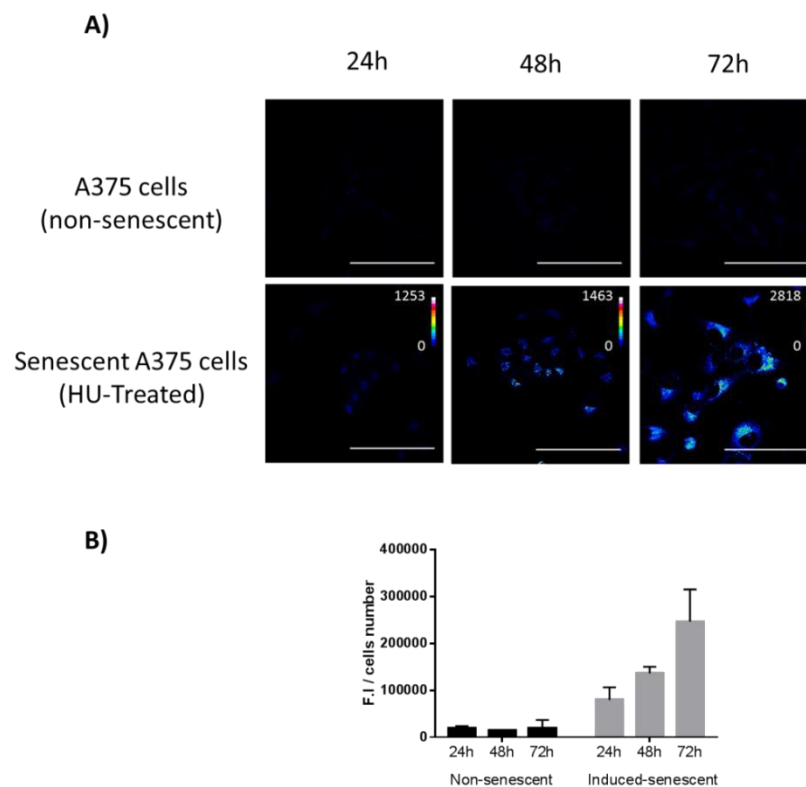


Figure 3.13. Fluorescence time-lapse microscopy of senescence establishment in human melanoma A375 cells. (A) images of A375 non-senescent (non-treated) and induced-senescent cells (50 μM HU) incubated with 0.1 μM **121** at various times (24, 48 and 72 h). Images were acquired using a confocal microscope with a 63X NA 1.4 objective; scale bars = 100 μm . (B) Quantification of fluorescence intensity of A375 senescence-induced cells treated with 0.1 μM **121**.

This observation was closely correlated with the increase of the fluorescence intensity (**Figure 3.13-B**). Conversely, normal A375 cells (HU non-treated, *i.e.* non-senescent), did not exhibit either morphological changes or fluorescence intensity enhancement.

Altogether these results support that, **121** allows the detection of minor increases in SA- β -gal activity, thus demonstrating the ability of our probe to monitor and detect, in real-time, the senescence at early stages in *live cells* (without pH modulation and interference of the senescence inducer treatment, HU).

Next, we evaluated the compatibility of **121** with flow cytometry analyses. For this purpose, we treated A375 cells with HU at 0, 10 and 50 μ M, and assess the applicability of **121** in the quantification of the proportion of senescent cells. After 72 h of HU treatment, the cells were collected, incubated with 0.1 μ M of **121** and subsequently analyzed by Fluorescence-activated cell sorting (FACS). As expected, a dose-dependent increase in size (forward scatter, FSC) and granularity (side scatter, SSC) was observed and closely correlated with an increase of fluorescence signal (72.5% at 10 μ M and 86.9% at 50 μ M, **Figure 3.14**).

This demonstrates the applicability of our probes in flow cytometry analyses, a widely used tools on biomedical and clinical research; hence **121** could be used in the development of a high-throughput screening protocol using flow cytometry.

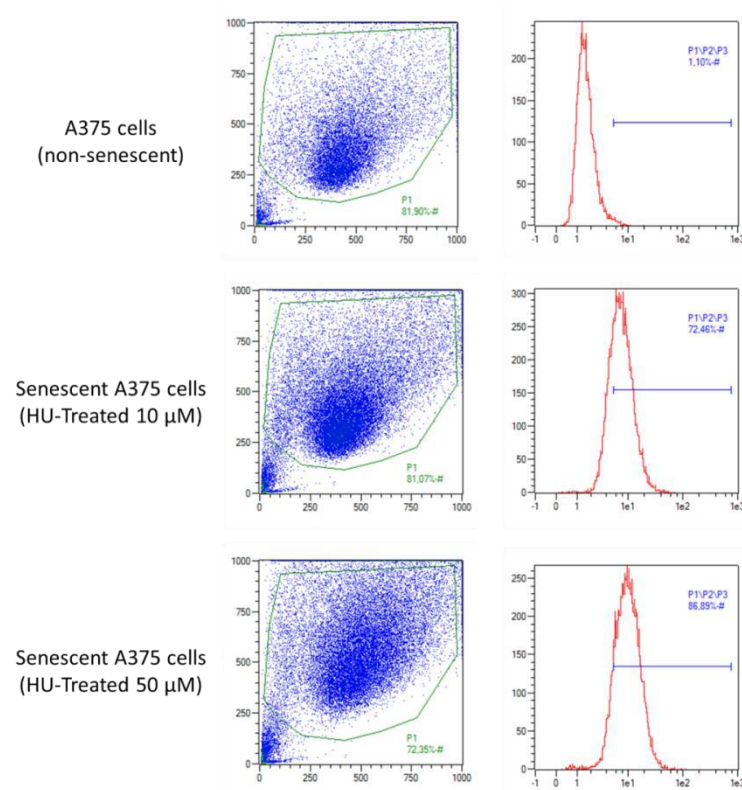


Figure 3.14. Flow cytometry analysis. SSC/FSC parameters and fluorescence signal have been evaluated by flow cytometry in normal and induced-senescent cells by hydroxyurea (HU) at various doses (0, 10 and 50 μM).

3.8 Conclusions - Chapter 3

In this study, we designed new small molecule light-up and ratiometric fluorescent probes for the detection of β -gal activity either in enzymatic or in *in vitro* experiment in living cells. Among the four probes that were prepared, **121**, **123** and **124** presented a remarkable enhancement on the kinetic response compared with the other reported β -gal probes (**4-MUG** and **FDG**). In particular, because of its non-toxicity, stability to hydrolytic degradation, cell permeability and turn-on fluorescence property, probe **121** appeared as a powerful and user-friendly fluorescent tool. Indeed, **121** can be used either in fixed or live cells and does not require a pH modulation step. It demonstrated an extremely high sensitivity ($\text{LOD} \approx 4.62 \times 10^{-5}$ U/mL in PBS), together with a high performance in the detection and visualization of minor increases of SA- β -gal activity in cells. Moreover, **121** allowed the real-time monitoring of the senescence appearance (TIS in A375 cancer cell line), *via* time-lapse fluorescence microscopy. Applicability of **121** in FACS analysis was also ascertained. Thus, it would be compatible with the development of a high-throughput screening for drug-discovery application in TIS

Conclusion Part I

In the first part of this manuscript we presented and discussed our major findings concerning the synthesis, the characterization and the sensing application of a new family of fluorescent dyes. Important concepts about environment sensitive fluorophores (push-pull dyes, ICT, TICT states, solvato(fluoro)chromism) as well as fluorescence sensing approaches were established on the first chapter. Thereafter, in the second chapter, two families of dyes displaying emission from intramolecular charge transfer (ICT) excited states could be deeply studied. Firstly, by using the simple and easily accessible 6-aminofunctionalized benzothiazole heterocycles, we showed the straightforward synthesis of a library of fine-tuned and highly emissive solvatochromic styryl-based fluorophores. The study and analysis of this family of dyes, using electronic spectroscopy techniques (absorption and fluorescence emission in the UV-visible region) and TD-DFT calculations enabled us to understand their particular structure–photophysics relationship. Three types of modifications of the central 2-styryl-6-aminobenzothiazole could be clearly evaluated (2 kinds of acceptor modulations and the alkylation of the donor amino group).

Subsequently, to complement and improve the absorption properties of the styryl family of dyes, a small group of fluorophores harboring an extended π -conjugated spacer were also investigated. Interestingly, these dyes feature two intense absorption bands (near-UV and visible regions) and a single, water sensitive and highly redshifted emission band compared with the styryl analogues. The formation of non-emissive TICT state could also be highlighted for fluorophores bearing the strong π -electron withdrawing formyl group on the acceptor side.

Finally, in **Chapter 3**, we focused on the valorization of our styryl fluorophores. Four probes were developed to monitor the activity of the β -galactosidase enzyme (β -gal), *in vitro* and in *living cells*. With the established structure-photophysical relationship concerning our benzothiazole-based fluorophores, probes harboring distinct spectroscopic properties (*e.g.* emission wavelength) were designed. Satisfactorily, two sets of fluorescent probes (*i.e.* light-up and ratiometric), with fast kinetic responses and high sensitivity could be obtained. Moreover, this fast and sensitive response allowed the detection of senescence-associated β -galactosidase (SA- β -gal) activity in *live cells*, using fluorescence microscopy and flow cytometry analysis

References

- (1) Wardle, B. *Principles and Applications of Photochemistry*, Wiley, Hoboken, 1st edn, 2009.
- (2) Valeur, B. *Molecular Fluorescence Principles and Applications*, Wiley-VCH, Weinheim, 2nd edn, 2001.
- (3) Pavia, D. L.; Lampman, G. M.; Kriz, G. S.; Vyvyan, J. R. *Introduction to Spectroscopy*, Brooks/Cole, Belmont, 4th edn, 2009.
- (4) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy Principles of Fluorescence Spectroscopy*, Springer, New York, 3rd edn, 2006.
- (5) Franck, J.; Dymond, E. *Trans. Faraday Soc.* **1926**, 26, 536–542.
- (6) Dadashi-Silab, S.; Doran, S.; Yagci, Y. *Chem. Rev.* **2016**, 116, 10212–10275.
- (7) Kasha, M. *Discuss. Faraday Soc.* **1950**, 9, 14–19.
- (8) Beija, M.; Afonso, C. A. M.; Martinho, J. M. G. *Chem. Soc. Rev.* **2009**, 38, 2410–2433.
- (9) Loudet, A.; Burgess, K. *Chem. Rev.* **2007**, 107, 4891–4932.
- (10) Kuimova, M. K.; Yahiolglu, G.; Levitt, J. A.; Suhling, K. *J. Am. Chem. Soc.* **2008**, 50, 6672–6673.
- (11) Yuan, L.; Lin, W.; Zheng, K.; He, L.; Huang, W. *Chem. Soc. Rev.* **2013**, 622–661.
- (12) Sun, W.; Guo, S.; Hu, C.; Fan, J.; Peng, X. *Chem. Rev.* **2016**, 7768–7817.
- (13) Collot, M.; Kreder, R.; Tatars, A. L.; Patsenker, L. D.; Mely, Y.; Klymchenko, A. *Chem. Commun.* **2015**, 51, 17136–17139.
- (14) Cabot, R.; Hunter, C. A. *Chem. Soc. Rev.* **2012**, 41, 3485–3492.
- (15) Lu, Y. J.; Wang, Z. Y.; Hu, D. P.; Deng, Q.; Huang, B. H.; Fang, Y. X.; Zhang, K.; Wong, W. L.; Chow, C. F. *Dye. Pigment.* **2015**, 122, 94–102.
- (16) Cheniour, M.; Gueyrard, D.; Goekjian, P. G.; Granjon, T.; Marcillat, O. *RSC Adv.* **2016**, 6, 5547–5557.
- (17) Lord, S. J.; Lee, H. D.; Samuel, R.; Weber, R.; Liu, N.; Conley, N. R.; Thompson, M. A.; Twieg, R. J.; Moerner, W. E. *J. Phys. Chem. B* **2010**, 114, 14157–14167.
- (18) Herner, A.; Nikić, I.; Kállay, M.; Lemke, E. A.; Kele, P. *Org. Biomol. Chem.* **2013**, 11, 3297.
- (19) Eördögh, Á.; Steinmeyer, J.; Peewasan, K.; Schepers, U.; Wagenknecht, H. A.; Kele, P. *Bioconjug. Chem.* **2016**, 27, 457–464.

- (20) Kuo, C.; Hochstrasser, R. M. *J. Am. Chem. Soc.* **2011**, *133*, 4664–4667.
- (21) Niko, Y.; Didier, P.; Mely, Y.; Konishi, G.; Klymchenko, A. S. *Sci. Rep.* **2016**, *6*, 18870.
- (22) Ni, Y.; Kannadorai, R. K.; Yu, S. W.-K.; Chang, Y.-T.; Wu, J. *Org. Biomol. Chem.* **2017**, *15*, 4531–4535.
- (23) Zhu, L.; Younes, A. H.; Yuan, Z.; Clark, R. J. *J. Photochem. Photobiol. A Chem.* **2016**, *311*, 1–15.
- (24) Achelle, S.; Rodríguez-López, J.; Bureš, F.; Robin-Le Guen, F. *Dye. Pigment.* **2015**, *121*, 305–311.
- (25) Imahori, H.; Umeyama, T.; Ito, S. *Acc. Chem. Res.* **2009**, *42*, 1809–1818.
- (26) Lin, Y.; Li, Y.; Zhan, X. *Chem. Soc. Rev.* **2012**, *41*, 4245–4272.
- (27) Szlapa, A.; Kula, S.; Blaszkiewicz, U.; Grucela, M.; Schab-Balcerzak, E.; Filapek, M. *Dye. Pigment.* **2016**, *129*, 80–89.
- (28) Xiao, P.; Frigoli, M.; Dumur, F.; Graff, B.; Gimes, D.; Fouassier, J. P.; Laleve, J. *Macromolecules* **2014**, *47*, 106–112.
- (29) Zhu, Y.; Wang, X. *Dye. Pigment.* **2013**, *97*, 222–229.
- (30) Wei, R.; Xu, Z.; Liu, X.; He, Y.; Wang, X. *J. Mater. Chem. C* **2015**, *3*, 10925–10933.
- (31) Bureš, F. *RSC Adv.* **2014**, *4*, 58826–58851.
- (32) Reichardt, C. *Solvents and solvent effects in organic chemistry*, Wiley-VCH, Weinheim, 3rd edn, 2003.
- (33) Marini, A.; Munnoz-Losa, A.; Biancardi, A.; Mennucci, B. *J. Phys. Chem. B* **2010**, *114*, 17128–17135.
- (34) Kulinich, A. V.; Ishchenko, A. A. *Russ. Chem. Rev.* **2009**, *78*, 141–164.
- (35) Kulinich, A. V.; Mikitenko, E. K.; Ishchenko, A. A. *Phys. Chem. Chem. Phys.* **2016**, *18*, 3444–3453.
- (36) Galievsky, V. A.; Druzhinin, S. I.; Demeter, A.; Mayer, P.; Kovalenko, S. A.; Senyushkina, T. A.; Zachariasse, K. A. *J. Phys. Chem. A* **2010**, *114*, 12622–12638.
- (37) Baladi, T.; Granzhan, A.; Piguel, S. *Eur. J. Org. Chem.* **2016**, *2016*, 2421–2434.
- (38) Karpenko, I. A.; Niko, Y.; Yakubovskiy, V. P.; Gerasov, A. O.; Bonnet, D.; Kovtun, Y. P.; Klymchenko, A. S. *J. Mater. Chem. C* **2016**, *4*, 3002–3009.
- (39) Grabowski, Z. R.; Rotkiewicz, K.; Rettig, W. *Chem. Rev.* **2003**, *103*, 3899–4031.
- (40) Galievsky, V. A.; Druzhinin, S. I.; Demeter, A.; Kovalenko, S. A.; Senyushkina, T.; Mayer, P.; Zachariasse, K. A. *J. Phys. Chem. A* **2011**, *115*, 10823–10845.
- (41) Lippert, E. *Z. Naturforsch.* **1955**, *10*, 541–545.

- (42) Mataga, N.; Kaifu, Y.; Koizumi, M. *Bull. Chem. Soc. Jpn.* **1956**, *29*, 465–470.
- (43) Reichardt, C. *Chem. Rev.* **1994**, *94*, 2319–2358.
- (44) Loving, G. S.; Sainlos, M.; Imperiali, B. *Trends Biotechnol.* **2010**, *28*, 73–83.
- (45) Niko, Y.; Kawauchi, S.; Konishi, G. I. *Chem. Eur. J.* **2013**, *19*, 9760–9765.
- (46) Shaya, J.; Fontaine-Vive, F.; Michel, B. Y.; Burger, A. *Chem. Eur. J.* **2016**, *22*, 10627–10637.
- (47) Sasaki, S.; Drummen, G. P. C.; Konishi, G. I. *J. Mater. Chem. C* **2016**, *4*, 2731–2743.
- (48) Santos, F. D. S.; Descalzo, R. R.; Gonçalves, P. F. B.; Benvenuti, E. V.; Rodembusch, F. S. *Phys. Chem. Chem. Phys.* **2012**, *14*, 10994–11001.
- (49) Liu, T.; Liu, X.; Spring, D. R.; Qian, X.; Cui, J.; Xu, Z. *Sci. Rep.* **2014**, *4*, 1–7.
- (50) Le, T. P.; Rogers, J. E.; Kelly, L. A. *J. Phys. Chem. A* **2000**, *104*, 6778–6785.
- (51) Zhang, W.; Ma, Z.; Du, L.; Li, M. *Analyst* **2014**, *139*, 2641–2649.
- (52) Escudero, D. *Acc. Chem. Res.* **2016**, *49*, 1816–1824.
- (53) Daly, B.; Ling, J.; De Silva, A. P. *Chem. Soc. Rev.* **2015**, *44*, 4203–4211.
- (54) Demchenko, A. P. *Introduction to fluorescence sensing*, Springer, New York, 2nd edn, 2015.
- (55) Kobayashi, H.; Ogawa, M.; Alford, R.; Choyke, P. L.; Urano, Y. *Chem. Rev.* **2010**, *110*, 2620–2640.
- (56) Blencowe, C. A.; Russell, A. T.; Greco, F.; Hayes, W.; Thornthwaite, D. W. *Polym. Chem.* **2011**, *2*, 773–790.
- (57) Zhang, J.; Yu, B.; Ning, L.; Zhu, X.; Wang, J.; Chen, Z.; Liu, X.; Yao, X.; Zhang, X.; Zhang, H. *European J. Org. Chem.* **2015**, *2015*, 1711–1718.
- (58) Yang, Q.; Jia, C.; Chen, Q.; Du, W.; Wang, Y.; Zhang, Q. *J. Mater. Chem. B* **2017**, *5* (10), 2002–2009.
- (59) Jiang, G.; Zeng, G.; Zhu, W.; Li, Y.; Dong, X.; Zhang, G.; Fan, X.; Wang, J.; Wu, Y.; Tang, B. Z. *Chem. Commun.* **2017**, *53*, 4505–4508.
- (60) Rautio, J.; Kumpulainen, H.; Heimbach, T.; Oliyai, R.; Oh, D.; Järvinen, T.; Savolainen, J. *Nat. Rev. Drug Discov.* **2008**, *7*, 255–270.
- (61) Burke, H. M.; Gunnlaugsson, T.; Scanlan, E. M. *Chem. Commun.* **2015**, *51*, 10576–10588.
- (62) Razgulin, A.; Ma, N.; Rao, J. *Chem. Soc. Rev.* **2011**, *40*, 4186–4216.
- (63) Grether, U.; Waldmann, H. *Chem. Eur. J.* **2001**, *7*, 959–971.
- (64) Shi, W.; Ma, H. *Chem. Commun.* **2012**, *48*, 8732–8744.
- (65) Asanuma, D.; Sakabe, M.; Kamiya, M.; Yamamoto, K.; Hiratake, J.; Ogawa, M.;

- Kosaka, N.; Choyke, P. L.; Nagano, T.; Kobayashi, H.; Urano, Y. *Nat. Commun.* **2015**, *6*, 1–7.
- (66) Kamiya, M.; Asanuma, D.; Kuranaga, E.; Takeishi, A.; Sakabe, M.; Miura, M.; Nagano, T.; Urano, Y. *J. Am. Chem. Soc.* **2011**, *133*, 12960–12963.
- (67) Rotman, B.; Zderic, J. A.; Edelstein, M. *Proc. Natl. Acad. Sci. U.S.A.* **1963**, *50*, 1–6.
- (68) Sakabe, M.; Asanuma, D.; Kamiya, M.; Iwatate, R. J.; Hanaoka, K.; Terai, T.; Nagano, T.; Urano, Y. *J. Am. Chem. Soc.* **2013**, *135*, 409–414.
- (69) Ichikawa, Y.; Kamiya, M.; Obata, F.; Miura, M.; Terai, T.; Komatsu, T.; Ueno, T.; Hanaoka, K.; Nagano, T.; Urano, Y. *Angew. Chem. Int. Ed.* **2014**, *53*, 6772–6775.
- (70) Han, J.; Han, M. S.; Tung, C. H. *Mol. Biosyst.* **2013**, *9*, 3001–3008.
- (71) Redy-Keisar, O.; Kisin-Finfer, E.; Ferber, S.; Satchi-Fainaro, R.; Shabat, D. *Nat. Protoc.* **2014**, *9*, 27–36.
- (72) Egawa, T.; Koide, Y.; Hanaoka, K.; Komatsu, T.; Terai, T.; Nagano, T. *Chem. Commun.* **2011**, *47*, 4162–4164.
- (73) Zhang, J.; Li, C.; Dutta, C.; Fang, M.; Zhang, S.; Tiwari, A.; Werner, T.; Luo, F. T.; Liu, H. *Anal. Chim. Acta* **2017**, *968*, 97–104.
- (74) Tung, C.-H.; Zeng, Q.; Shah, K.; Kim, D.-E.; Schellingerhout, D.; Weissleder, R. *Cancer Res.* **2004**, *64*, 1579–1583.
- (75) Corey, P. F.; Trimmer, R. W.; Biddlecom, W. G. *Angew. Chem. Int. Ed.* **1991**, *1*, 1646–1648.
- (76) Zhang, X. X.; Wu, H.; Li, P.; Qu, Z. J.; Tan, M. Q.; Han, K. L. *Chem. Commun.* **2016**, *52*, 8283–8286.
- (77) Huang, J.; Li, N.; Wang, Q.; Gu, Y.; Wang, P. *Sens. Actuators, B* **2017**, *246*, 833–839.
- (78) Gu, K.; Xu, Y.; Li, H.; Guo, Z.; Zhu, S.; Zhu, S.; Shi, P.; James, T. D.; Tian, H.; Zhu, W. H. *J. Am. Chem. Soc.* **2016**, *138*, 5334–5340.
- (79) Kim, E. J.; Kumar, R.; Sharma, A.; Yoon, B.; Kim, H. M.; Lee, H.; Hong, K. S.; Kim, J. S. *Biomaterials* **2017**, *122*, 83–90.
- (80) Lee, H. W.; Heo, C. H.; Sen, D.; Byun, H. O.; Kwak, I. H.; Yoon, G.; Kim, H. M. *Anal. Chem.* **2014**, *86*, 10001–10005.
- (81) Carter, K. P.; Young, A. M.; Palmer, A. E. *Chem. Rev.* **2014**, *114*, 4564–4601.
- (82) Reja, S. I.; Gupta, M.; Gupta, N.; Bhalla, V.; Ohri, P.; Kaur, G.; Kumar, M. *Chem. Commun.* **2017**, *53*, 3701–3704.
- (83) Wang, T.; Douglass, E. F.; Fitzgerald, K. J.; Spiegel, D. A. *J. Am. Chem. Soc.* **2013**, *135*, 12429–12433.

- (84) Venkatesan, P.; Wu, S. P. *Analyst* **2015**, *140*, 1349–1355.
- (85) Matsumoto, T.; Urano, Y.; Shoda, T.; Kojima, H.; Nagano, T. *Org. Lett.* **2007**, *9*, 3375–3377.
- (86) Dias, G. G.; King, A.; De Moliner, F.; Vendrell, M.; Da Silva Júnior, E. N. *Chem. Soc. Rev.* **2018**, *47*, 12–27.
- (87) Yogo, T.; Urano, Y.; Kamiya, M.; Sano, K.; Nagano, T. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 4320–4323.
- (88) Urano, Y.; Kamiya, M.; Kanda, K.; Ueno, T.; Hirose, K.; Nagano, T. *J. Am. Chem. Soc.* **2005**, *127*, 4888–4894.
- (89) Kamiya, M.; Kobayashi, H.; Hama, Y.; Koyama, Y.; Bernardo, M.; Nagano, T.; Choyke, P. L.; Urano, Y. *J. Am. Chem. Soc.* **2007**, *129*, 3918–3929.
- (90) Mei, J.; Leung, N. L. C.; Kwok, R. T. K.; Lam, J. W. Y.; Tang, B. Z. *Chem. Rev.* **2015**, *115*, 11718–11940.
- (91) Gao, M.; Tang, B. Z. *ACS Sensors* **2017**, *2*, 1382–1399.
- (92) Yang, W.; Zhao, X.; Zhang, J.; Zhou, Y.; Fan, S.; Sheng, H.; Cao, Y.; Hu, Y. *Dye. Pigment.* **2018**, *156*, 100–107.
- (93) Peng, L.; Gao, M.; Cai, X.; Zhang, R.; Li, K.; Feng, G.; Tong, A.; Liu, B. *Optoelectron. Adv. Mater.* **2010**, *4*, 1166–1169.
- (94) Jiang, T.; Du, L.; Li, M. *Photochem. Photobiol. Sci.* **2016**, *15*, 466–480.
- (95) Rathbun, C. M.; Prescher, J. A. *Biochemistry* **2017**, *56*, 5178–5184.
- (96) Ke, B.; Wu, W.; Wei, L.; Wu, F.; Chen, G.; He, G.; Li, M. *Anal. Chem.* **2015**, *87*, 9110–9113.
- (97) Hemmi, M.; Ikeda, Y.; Shindo, Y.; Nakajima, T.; Nishiyama, S.; Oka, K.; Sato, M.; Hiruta, Y.; Citterio, D.; Suzuki, K. *Chem. Asian J.* **2018**, *13*, 648–655.
- (98) Wehrman, T. S.; von Degenfeld, G.; Krutzik, P. O.; Nolan, G. P.; Blau, H. M. *Nat. Methods* **2006**, *3*, 295–301.
- (99) Lozano-Torres, B.; Galiana, I.; Rovira, M.; Garrido, E.; Chaib, S.; Bernardos, A.; Muñoz-Espín, D.; Serrano, M.; Martínez-Máñez, R.; Sancenón, F. *J. Am. Chem. Soc.* **2017**, *139*, 8808–8811.
- (100) Yuan, L.; Lin, W.; Zheng, K.; Zhu, S. *Acc. Chem. Res.* **2013**, *46*, 1462–1473.
- (101) Barthes, N. P. F.; Gavvala, K.; Bonhomme, D.; Dabert-Gay, A. S.; Debayle, D.; Mély, Y.; Michel, B. Y.; Burger, A. *J. Org. Chem.* **2016**, *81*, 10733–10741.
- (102) Sekar, R. B.; Periasamy, A. *J. Cell Biol.* **2003**, *160*, 629–633.
- (103) Margineanu, A.; Chan, J. J.; Kelly, D. J.; Warren, S. C.; Flatters, D.; Kumar, S.; Katan,

- M.; Dunsby, C. W.; French, P. M. W. *Sci. Rep.* **2016**, *6*, 28186.
- (104) Shiu, H. Y.; Chong, H. C.; Leung, Y. C.; Wong, M. K.; Che, C. M. *Chem. Eur. J.* **2010**, *16*, 3308–3313.
- (105) Leriche, G.; Budin, G.; Darwich, Z.; Weltin, D.; Mély, Y.; Klymchenko, A. S.; Wagner, A. *Chem. Commun.* **2012**, *48*, 3224–3226.
- (106) Su, D.; Teoh, C. L.; Sahu, S.; Das, R. K.; Chang, Y. T. *Biomaterials* **2014**, *35*, 6078–6085.
- (107) Wei, C.; Wei, L.; Xi, Z.; Yi, L. *Tetrahedron Lett.* **2013**, *54*, 6937–6939.
- (108) Jia, X.; Chen, Q.; Yang, Y.; Tang, Y.; Wang, R.; Xu, Y.; Zhu, W.; Qian, X. *J. Am. Chem. Soc.* **2016**, *138*, 10778–10781.
- (109) Komatsu, T.; Kikuchi, K.; Takakusa, H.; Hanaoka, K.; Ueno, T.; Kamiya, M.; Urano, Y.; Nagano, T. *J. Am. Chem. Soc.* **2006**, *128*, 15946–15947.
- (110) Long, L.; Lin, W.; Chen, B.; Gao, W.; Yuan, L. *Chem. Commun.* **2011**, *47*, 893–895.
- (111) Klymchenko, A. S. *Acc. Chem. Res.* **2017**, *50*, 366–375.
- (112) Zhou, K.; Ren, M.; Deng, B.; Lin, W. *New J. Chem.* **2017**, *41*, 11507–11511.
- (113) Kucherak, O. A.; Oncul, S.; Darwich, Z.; Yushchenko, D. A.; Arntz, Y.; Didier, P.; Me, Y.; Klymchenko, A. S. *J. Am. Chem. Soc.* **2010**, *132*, 4907–4916.
- (114) Sainlos, M.; Iskenderian, W. S.; Imperiali, B. *J. Am. Chem. Soc.* **2009**, *131*, 6680–6682.
- (115) Riedl, J.; Pohl, R.; Ernsting, N. P.; Ors, P.; Hocek, M. *Chem. Sci.* **2012**, *3*, 2797–2806.
- (116) Riedl, J.; Me, P.; Pohl, R.; Orsa, P.; Fojta, M.; Hocek, M. *J. Org. Chem.* **2012**, *77*, 8287–8293.
- (117) Dziuba, D. D.; Pospisil, P.; Matyasovský, J.; Brynda, J.; Nachtigallová, D.; Rulisek, L.; Pohl, R.; Hof, M.; Hocek, M. *Chem. Sci.* **2016**, *7*, 5775–5785.
- (118) Peng, T.; Yang, D. *Org. Lett.* **2010**, *12*, 496–499.
- (119) Strehmel, B.; Sarker, A. M.; Detert, H. *ChemPhysChem* **2003**, *4*, 249–259.
- (120) Morales, A. R.; Frazer, A.; Woodward, A. W.; Ahn-White, H. Y.; Fonari, A.; Tongwa, P.; Timofeeva, T.; Belfield, K. D. *J. Org. Chem.* **2013**, *78*, 1014–1025.
- (121) Deligeorgiev, T.; Vasilev, A.; Kaloyanova, S.; Vaquero, J. J. *Color. Technol.* **2010**, *126*, 55–80.
- (122) Yalçın, E.; Achelle, S.; Bayrak, Y.; Seferoğlu, N.; Barsella, A.; Seferoğlu, Z. *Tetrahedron Lett.* **2015**, *56*, 2586–2589.
- (123) Tayade, R. P.; Sekar, N. *Dye. Pigment.* **2016**, *128*, 111–123.
- (124) Umape, P. G.; Gawale, Y.; Sekar, N. *J. Fluoresc.* **2014**, *24*, 1087–1098.

- (125) Yadav, U. N.; Kumbhar, H. S.; Deshpande, S. S.; Sahoo, S. K.; Shankarling, G. S. *RSC Adv.* **2015**, *5*, 42971–42977.
- (126) Jędrzejewska, B.; Ośmiałowski, B.; Zalesny, R. *Photochem. Photobiol. Sci.* **2016**, *15*, 117–128.
- (127) Telore, R. D.; Sekar, N. *Dye. Pigment.* **2016**, *129*, 1–8.
- (128) Szukalski, A.; Parafiniuk, K.; Haupa, K.; Goldeman, W.; Sahraoui, B.; Kajzar, F.; Mysliwiec, J. *Dye. Pigment.* **2017**, *142*, 507–515.
- (129) Vidya, B.; Iniya, M.; Sivaraman, G.; Sumesh, R. V.; chellappa, D. *Sens. Actuators, B* **2017**, *242*, 434–442.
- (130) Matsumura, K.; Ono, M.; Kitada, A.; Watanabe, H.; Yoshimura, M.; Iikuni, S.; Kimura, H.; Okamoto, Y.; Ihara, M.; Saji, H. *J. Med. Chem.* **2015**, *58*, 7241–7257.
- (131) Naeem, K. C.; Subhakumari, A.; Varughese, S.; Nair, V. C. *J. Mater. Chem. C* **2015**, *3*, 10225–10231.
- (132) Fan, L.; Nan, M.; Ge, J.; Wang, X.; Lin, B.; Zhang, W.; Shuang, S.; Dong, C. *New J. Chem.* **2018**, *42*, 13479–13485.
- (133) Bashmakova, N. V.; Shaydyuk, Y. O.; Levchenko, S. M.; Masunov, A. E.; Przhonska, O. V.; Bricks, J. L.; Kachkovsky, O. D.; Slominsky, Y. L.; Piryatinski, Y. P.; Belfield, K. D.; Bondar, M. V. *J. Phys. Chem. A* **2014**, *118*, 4502–4509.
- (134) Ono, M.; Hayashi, S.; Kimura, H.; Kawashima, H.; Nakayama, M.; Saji, H. *Bioorg. Med. Chem.* **2009**, *17*, 7002–7007.
- (135) Liu, Z.; Zhang, C.; He, W.; Qian, F.; Yang, X.; Gao, X.; Guo, Z. *New J. Chem.* **2010**, *34*, 656–660.
- (136) Würth, C.; Grabolle, M.; Pauli, J.; Spieles, M.; Resch-genger, U. *Nat. Protoc.* **2013**, *8*, 1535–1550.
- (137) Kundu, K.; Knight, S. F.; Willett, N.; Lee, S.; Taylor, W. R.; Murthy, N. *Angew. Chem. Int. Ed.* **2009**, *48*, 299–303.
- (138) Kulinich, A. V.; Ishchenko, A. A.; Chibisov, A. K.; Zakharova, G. V. *J. Photochem. Photobiol. A Chem.* **2014**, *274*, 91–97.
- (139) Hansch, C.; Leo, A.; Taft, R. W. *Chem. Rev.* **1991**, *91*, 165–195.
- (140) Martin, M. M.; Plaza, P.; Changenet-barret, P. *J. Phys. Chem. A* **2002**, *106*, 2351–2358.
- (141) Okada, T.; Fujita, T.; Kubota, M.; Masaki, S.; Ide, R.; Sakata, Y.; Mlsuiw, S. *Chem. Phys. Lett.* **1972**, *14*, 563–568.
- (142) Mata, G.; Luedtke, N. W. *Org. Lett.* **2013**, *15*, 2462–2465.

- (143) Cao, C.; Liu, X.; Qiao, Q.; Zhao, M.; Yin, W.; Mao, D.; Zhang, H.; Xu, Z. *Chem. Commun.* **2014**, *50*, 15811–15814.
- (144) Brunner, N.; Thompson, E. W.; Spang-Thomsen, M.; Rygaard, J.; Dano, K.; Zwiebel, J. A. *Eur. J. Cancer* **1992**, *28A*, 1989–1995.
- (145) Fu, Y.; Xiao, W. *Methods Mol. Biol.* **2006**, *313*, 257–264.
- (146) Serebriiskii, I. G.; Golemis, E. A. *Anal. Biochem.* **2000**, *285*, 1–15.
- (147) Dimri, G. P.; Lee, X.; Basile, G.; Acosta, M.; Scott, G.; Roskelley, C.; Medrano, E. E.; Linskens, M.; Rubelj, I.; Pereira-Smith, O. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9363–9367.
- (148) Ewald, J. A.; Desotelle, J. A.; Wilding, G.; Jarrard, D. F. *J. Natl. Cancer Inst.* **2010**, *102*, 1536–1546.
- (149) Schwarze, S. R.; Fu, V. X.; Desotelle, J. A.; Kenowski, M. L.; Jarrard, D. F. *Neoplasia* **2005**, *7*, 816–823.
- (150) Chang, B. D.; Broude, E. V.; Dokmanovic, M.; Zhu, H.; Ruth, A.; Xuan, Y.; Kandel, E. S.; Lausch, E.; Christov, K.; Roninson, I. B. *Cancer Res.* **1999**, *59*, 3761–3767.
- (151) Coppé, J.-P.; Desprez, P.-Y.; Krtolica, A.; Campisi, J. *Annu. Rev. Pathol. Mech. Dis.* **2010**, *5*, 99–118.
- (152) Debacq-Chainiaux, F.; Erusalimsky, J. D.; Campisi, J.; Toussaint, O. *Nat. Protoc.* **2009**, *4*, 1798–1806.

**Part II. Nucleic acid chemistry: the targeting of RNA structures
by using small-sized ligands and the newest synthetic strategy
applied to the post-synthetic functionalization of oligonucleotides**

**Chapter 4 - Synthesis and Biophysical Studies of Novel RNA
Ligands to Target HCV IRES III_d Loop**

Chapter Outline

Ribonucleic acids (RNA) are biopolymers responsible for a wide number of biological processes. For instance, coding RNAs carry the genetic information that contains the amino acid sequence that will be translated into proteins. Additionally, noncoding RNAs have recently been recognized to play a key role in many other biological processes such as gene regulation, tumorigenesis and viral translation. Thereby, RNA is nowadays considered as a potential biological target for a great number of human diseases.

In this chapter, we will focus on the development of new small-sized RNA ligands to target non-coding RNAs. In the recent years, many studies have been demonstrated that viral translation can be effectively inhibited by targeting noncoding and structured regions of the viral RNAs using different ligands. As a consequence, the viral genomic RNA has emerged as a potent and alternative target to discover new antiviral therapies. Here, we are particularly interested in the targeting of the HCV IRES III_d loop, which have been considered as a potential therapeutic target to overcome HCV infection.

Thus, in this chapter we will initially present a bibliographic review emphasizing the RNA functions, structures and the families of ligands that are known to interact with it. Then, a description concerning the synthesis of two novel series of multimodal ligands, designed using a rational combination of molecular recognition elements and electrostatic/H-bonding interactions to strengthen the RNA-ligand complex stability, will be provided. Finally, the biophysical study concerning the interaction of the ligands with HCV IRES III_d loop RNA model will be discussed.

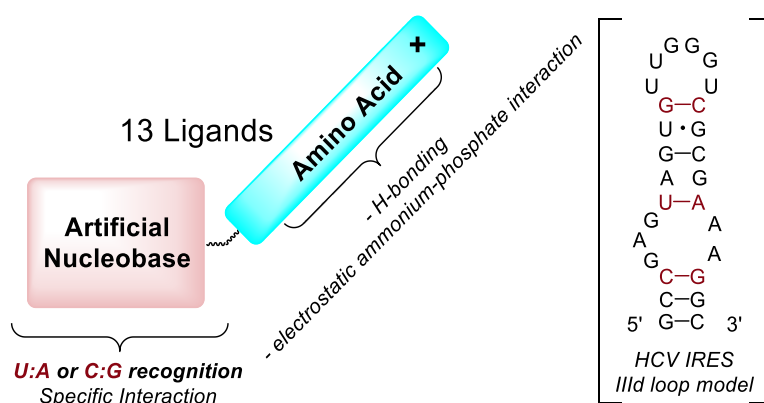


Figure 4.1. General representation of the ligand structures to target RNA secondary structures.

4.1 State of the Art

4.1.1 The Nucleotides and the Primary Structure of DNA and RNA

DNA and RNA are negatively charged biopolymers constituted by a chain of nucleotide building blocks. Each nucleotide is composed by three fundamental moieties: a (deoxy)ribose sugar, a phosphate group and one nucleobase connected to the 1-position (anomeric position) of the sugar by *N*-glycosidic bond in β -configuration. The chemical structure of DNA and RNA differs only by two chemical modifications: the sugar and the nucleobases that compose each nucleotide. DNA features a 2-deoxyribose saccharide and the canonic thymine (T), adenine (A), guanine (G) and cytosine (C) as nucleobases. Non-canonical bases such as 5-substituted cytosine (5-methyl and 5-hydroxymethyl) can also be found. In contrast, RNA contains a ribose as sugar and the canonical bases A, G, C and U (uracil), instead of T (**Figure 4.2**). RNA can also feature a broad variety of "non-canonical" nucleobases. The inosine, pseudouridine, dihydrouridine, m7-guanosine are some examples. However, apart from the chemical similarity, DNA and RNA possess huge structural and functional differences.¹

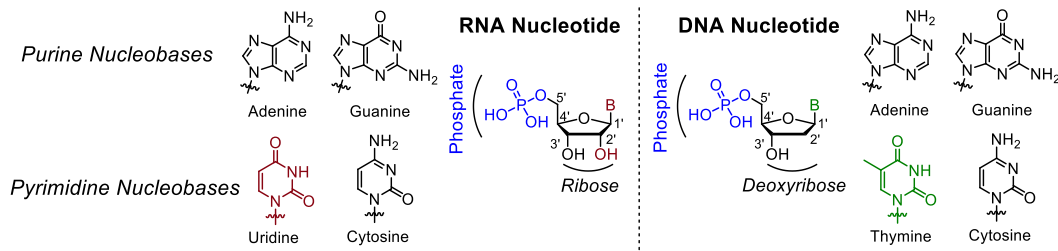


Figure 4.2. General structure of RNA and DNA nucleotides.

The structure of nucleic acids is usually divided in different levels of organization. The linear nucleotide sequence defines the primary structure of NAs. The nucleotides are connected to together through 5',3'-phosphodiester bonds (**Figure 4.3-A**). The secondary structure is dictated by the interactions between the bases within the same NA strand or with another strand. Essentially, associative base pairing between the canonical and complementary pyrimidine-purine U:A (T:A) and C:G base pairs are formed *via* Watson-Crick hydrogen bond interactions and stabilized through strong stacking forces among the nucleobases (**Figure 4.3-B**).^{2,3} The secondary structure and the upper organization levels will be further discussed in the next section, in the context of ribonucleic acids.

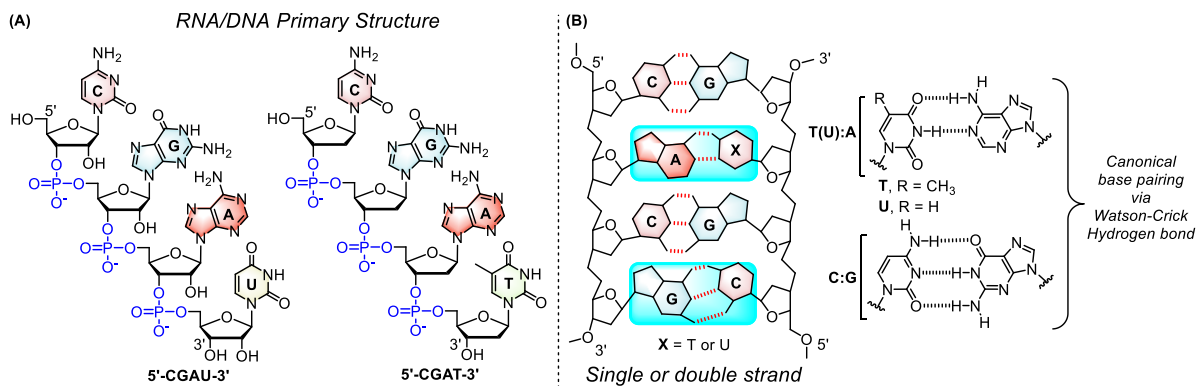


Figure 4.3. (A) DNA/RNA primary structure. (B) The canonical purine-pyrimidine base pair through Watson-Crick hydrogen bonds.

4.1.2 RNA Structure and Functions

RNA is a multifaceted biomolecule, essential for numerous biological processes. It is implicated in the transcription of genetic information (coding RNAs), but also regulates numerous processes inside the cell (non-coding RNAs).⁴ Overall, RNA is a key molecule in the transcription and translation processes. Transcription is the first step of gene expression pathway, in which the genetic information encoded into DNA molecules is copied

(transcribed) into RNA molecules. This process is enzymatically triggered and regulated by an RNA polymerase, which promotes both, the separation of the two strands of the DNA helix and the polymerization of RNA nucleotides. The RNA polymerase uses the single stranded DNA sequence as a template to produce a RNA strand identical with the coding DNA strand (except that T units are replaced with U). Transcription can produce both, RNA precursors for functional RNAs products (RNA genes), or mRNAs, that are information carriers coding for the synthesis of proteins. Essentially, mRNAs contain the codon sequence, which will be translated (decoded) in the ribosome to produce specific polypeptides (proteins) as the final functional gene product (**Figure 4.4**).³

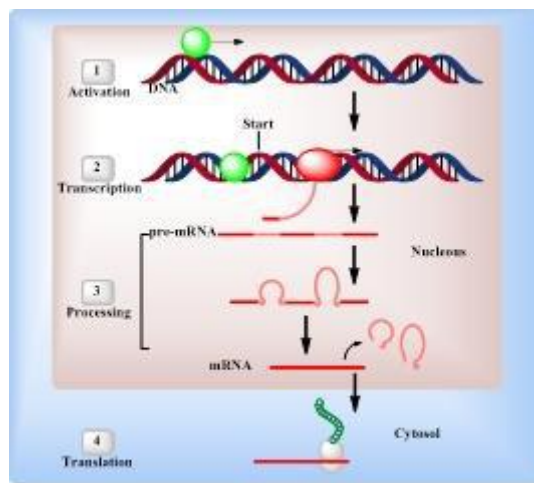


Figure 4.4. Simplified mechanism of transcription and protein translation.

In addition to its role in gene translation, RNA is also involved in the control and regulation of multiple biological processes; this is due to its ability to adopt a great diversity of tridimensional conformations. In addition to its primary structure (nucleotidic sequence), the secondary and tertiary structures afford many additional functionalities, which enable RNA to interact with a variety of molecules. Moreover, RNA can regulate gene expression during all steps of the process, from the DNA-RNA transcription step to post-translational modification. It also catalyzes biochemical transformations and transport molecules within the cell.

Differently from DNA, that normally exists as two strands (two paired DNA strands) associated to each other *via* Watson-Crick interactions to form anti-parallel double helical structures, RNA is usually occurring as a single strand polynucleotide. Moreover, alike proteins, RNA can fold into a large diversity of tridimensional structures, as a result of the base pairing between complementary regions. Thus, within these RNA 3D-structures both

single and double-stranded domains are present. The double-stranded domains result from the anti-parallel association through Watson-Crick H-bonds. This motif adopt an helical shape in which two grooves are present; the major groove and the minor groove, according to their relative size (**Figure 4.5**). In these structures, the bases stacks at the center of the helix while the negatively charged phosphodiester backbone are oriented on the outside of the helix. By contrast, the non-complementary regions of the RNA sequences form single-stranded domains which results in the formation of three other types of secondary structured domains: loops, bulges and junctions (**Figure 4.5**). Among the secondary structures, stem loops (or hairpin loops) are considered the most common motifs. In this structure, the double helical tract is the stem, while the unpaired nucleotide sequence in the extremity of the helix forms the loop. Internal loops and bulges are formed due to the separation of the double helical tract on both strands (internal loop) or in only one strand (bulge) by unpaired nucleotides. A mismatch is a special type of internal loop in which a single nucleotide is unpaired on each strand. Junctions are single stranded regions connecting double stranded domains. The most common are the 3-way and 4-way junctions (**Figure 4.5**).^{1,5}

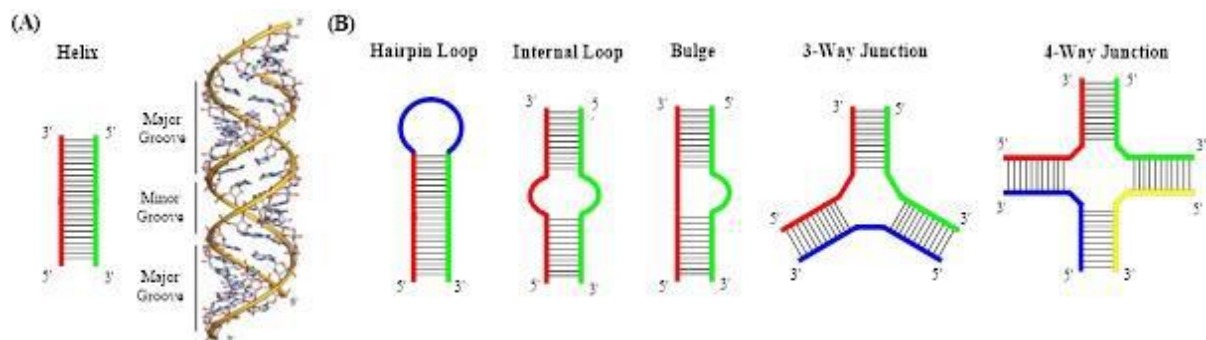


Figure 4.5. RNA secondary structures. (A) Schematic representation of an RNA helix and the 3D structure of an A-form helix. (B) Schematic representation of RNA loops, bulges and junctions.

The interaction between RNA secondary structures elements lead to the formation of RNA tertiary structures. The RNA domains (either single or double stranded structures) are associated through van der Waals forces, specific Watson-Crick hydrogen bonds and/or unusual pairs (like Wobble base pair)⁶ involving hairpin loops or internal bulges forming other three dimensional levels of organization. Pseudoknots and kissing-loops are two motifs with remarkable functions that are formed by the association of two unpaired strands (**Figure 4.6**). Pseudoknots are important functional elements in RNA structure, found in several

classes of RNA, and possess a diversity of key functions in enzymes (ribozymes)⁷ and viruses.^{8,9} They are formed when nucleotides from a hairpin loop pairs with a single stranded region outside of the hairpin, hence forming an helical segment. The second motif, "kissing hairpins" are formed via the base-pairing of the unpaired nucleotides of 2 different loops. This yields loop-loop complexes which are recognized to be essential for genomic RNA dimerization, regulation of gene expression and viral replication.^{10,11}

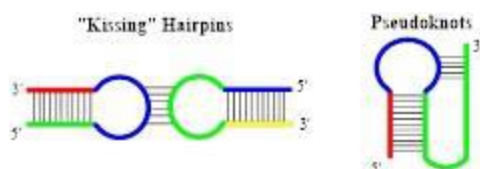


Figure 4.6. Pseudoknot and loop-loop complexes structures.

The triple helix organization is another important structure that results from the association of single stranded region and an helical structure. Although it has been discovered over 50 years ago, the diversity of biological roles played by RNA triplexes just started to be elucidated.¹² Essentially, a pre-formed Watson-Crick base paired RNA stem may accommodate a third strand to form a new triplex structure, without disrupting the preformed duplex. These base triplets are formed mainly with the Watson-Crick base pairs in the major groove of the helix. However, in certain cases, they can also be established by minor groove interactions. The single stranded sequence associates to the purine base of the preformed helix by tertiary interactions via Hoogsteen or reversed Hoogsteen contacts (**Figure 4.7**).¹³ Triple helix formation obeys precise rules imposed by several structural constraints.¹⁴ It requires an homopurine-homopyrimidine sequence in the target double stranded RNA. Hoogsteen hydrogen bonds constitutes the major element for the formation of base triplets, although the electronic repulsion between the phosphates, the pH and entropic barriers are some constraints involved in the triplex helix formation. Standard major groove base triplets found in natural RNAs include the canonical U.A:U and C.G:C and the triplet U.G:C (where . represents the hydrogen bonds on the Hoogsteen edge and : represents the Watson-Crick base pairing, **Figure 4.7**).^{15,16}

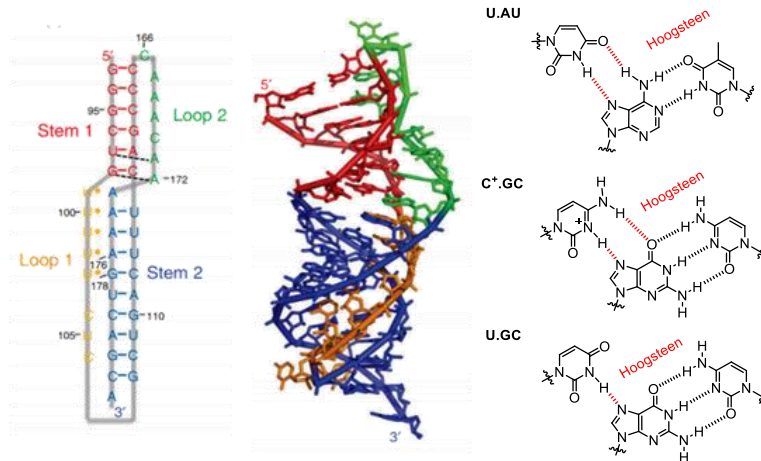


Figure 4.7. RNA triple helix observed in the H-type pseudoknot structure found in human telomerase RNA and the corresponding standard RNA base triplets.

A large variety of RNAs have already been identified in the past and in the recent years. In addition to mRNAs, important as carriers of genetic information, transfer RNA (tRNA) and ribosomal RNA (rRNA) are the other two important groups of RNAs involved in the protein translation step. tRNAs carry the relevant amino acid to the ribosomal machinery during protein translation. Each tRNA is connected to one of the 20 natural amino acids on its terminal 3'-OH group. Basically, tRNAs contains the complementary 3 nucleotide sequence (anticodon) to match the codon sequence of the reading mRNA.¹⁷ On the other hand, rRNAs are the major components of the ribosomes, representing more than 50 % of all RNAs present in eukaryotic organisms. It associates with ribosomal proteins forming a ribonucleoprotein complex. The ribosome is divided into two subunit (the small 40S subunit and the large 60S), that works as a single machine to translate mRNAs into polypeptides. The small 40S unit binds to the target mRNA while the large 60S unit, a peptidyl-transferase unit, binds to the aminoacyl-tRNA (tRNA charged with an amino acid) and catalyzes the formation of the peptide linkage between two consecutive amino acids according to the protein sequence encoded in the mRNA.¹⁸

RNA molecules can also perform specific and/or catalytic functions, in association or not with proteins. Ribozymes is a group of RNA molecules featuring enzymatic activity. They play fundamental roles in many processes, especially during replication and the processing of mRNAs.

Recently, new important groups of noncoding RNAs have been identified. Among them, small noncoding RNAs, such as miRNAs and siRNA, and long noncoding RNAs (lncRNAs) are some examples of ribonucleic acids that have been strongly implicated in the regulation

of gene expression. LncRNAs are noncoding RNAs with nucleotide sequences longer than 200 nucleotides. They have been identified to affect the regulation of gene expression in both transcriptional and post-transcriptional fashions.^{19,20} On the one hand, small ncRNAs (up to 25 nucleotides) are known to direct post-transcriptional regulation of gene expression by binding to conserved sequences within the 3'-untranslated region (3'-UTR) of target mRNAs.²¹ This association can result in the translational repression or mRNA degradation, which leads to the inhibition of the correspondent gene expression. miRNAs and siRNAs have also been associated with a broad range of essential biologic processes, as cellular growth and differentiation, apoptosis, and modulation of host response to viral infection. Moreover, considering their role in the regulation of gene expression, abnormal levels (under or overexpression) of ncRNAs have been identified as key points on the development and progress of certain pathologies, including several aggressive cancers.^{22,23}

4.1.3 RNA Targets

The different kinds of RNAs found in eucaryotic and procaryotic cells play particular and specific functions in the cell. Due to their ability to fold multiple 3D structures, RNAs can interact with a large number of chemical substrates, such as proteins, nucleic acids and small molecules. Thus, RNA is being considered as an important therapeutic target. It is noteworthy that, the resolution of the RNA secondary and tertiary structures provides pivotal information to establish either modes of action and biological targets and, thereby, it can be used as a tool to design ligands for a specific RNA sequence.

Nowadays, a large number of RNAs have been recognized as potential biological targets. Among them, the most studied includes tRNAs, rRNAs, coding and noncoding regions of mRNAs, ribozymes and miRNAs. Several strategies have been explored with the aim to identify and access more active, selective and specific ligands, with ability to inhibit the main biological role in which the target RNA is involved. In this work, we are interested in viral RNAs, more specifically, the IIIId sub domain of the Internal Ribosome Entry Site (IRES) of the Hepatitis C Virus (HCV).

4.1.4 Hepatitis C Virus (HCV)

According with the World Health Organization (WHO), HCV affects about 170 millions people worldwide hence representing a major health problem. It is a main cause of chronic hepatitis and may lead to hepatocellular carcinoma and liver cirrhosis.²⁴

Hepatitis C virus (HCV) is a positive-sense single-stranded RNA virus belonging to the *flaviviridae* family that is expressed and replicated in the cytoplasm. The HCV genome is composed of one RNA strand with approximately 9600 bases length and the several strains can be divided into 7 genotypes and 67 subtypes. The translation of HCV produces a precursor polyprotein containing 3010-3033 amino acids, that is further hydrolyzed by viral and cellular proteases to lead to structural proteins (the protein) core C and both envelope proteins E1 and E2) and functional viral proteins (p7, NS2, NS3, NS4A, NS4B, NS5A, NS5B).²⁵

While the sequence of HCV RNA varies among the reading frame region (coding region), the non-translated sequences at the 5' and 3' ends, called UTRs (Untranslated Regions), are highly conserved among different HCV genotypes and strongly implicated in viral replication. These UTRs are, in addition, highly structured. They adopt several hairpin structures, thus forming stem-loop subdomains (**Figure 4.8**). The 3'-end of the 3'-UTR region consists of highly conserved "stem-loop" structures that are involved in the formation of the replication initiation complex by interacting with the viral polymerase (NS5B) and cellular factors. By contrast, the 5'-NTR region plays a double role in the replication of the virus. Indeed, it participates in the translation of the genome since the initiation of translation is controlled by an Internal Ribosomal Entry Site (IRES) sequence which occupies most of the 5'UTR region (regions II to IV, **Figure. 4.8**).

HCV IRES is a highly conserved sequence of the HCV genome and composed of three folding domains (II-IV, **Figure 4.8**) that contribute to its activity. It allows the recruitment of the ribosomal machinery together with a restricted number of eukaryotic initiation factors (eIFs) to promote the translation of the viral RNA in a 5'-cap-independent pathway.²⁶ In particular, domains III and IV are essential for the viral translation. Indeed, the domain IV contains the initiator codon AUG while domain III (more precisely the stem-loop structure III_d) presents the binding site for the 40S ribosome subunit to promote the initiation of the translation. Although domain II itself does not bind to the small ribosomal subunit, it also remains essential for IRES activity.^{26,27}

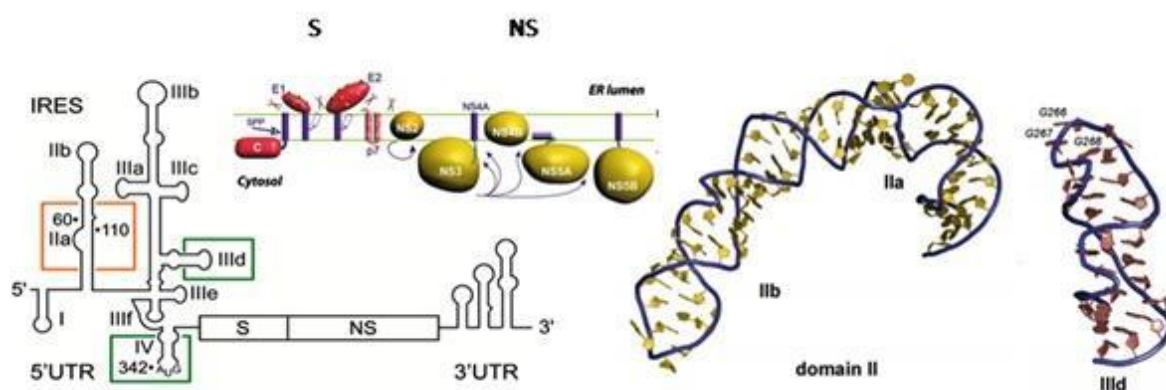


Figure 4.8. HCV Genome (left) and crystal structure of IRES IIa-b and IIId sub domains (right).

Until recently, treatments against HCV were restricted to ribavirin and (pegylated)-interferon, while exhibiting limited efficacy. In the 2010s, a remarkable number of direct acting antivirals (DAAs) started to supplement the market with new anti-HCV drugs.²⁸ The great majority of these drugs being designed to inhibit viral proteins activity. In 2013, the FDA approval of sofosbuvir (**129**) (**Figure 4.9**), a nucleotide inhibitor of the NS5B viral polymerase, revolutionized HCV treatment. It yields impressive sustained virological response (SVR) and limited adverse effects.²⁹ Nevertheless, this treatment is contraindicated for patients suffering severe renal impairments and its efficacy on HCV genotype 3 is not satisfactory.³⁰

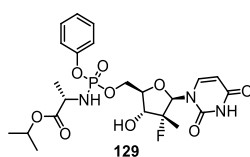


Figure 4.9. Chemical structure of Sofosbuvir.

Thus, new complementary approaches to overcome HCV infection are still needed. Some of them have been exploiting the inhibition of both, mi-RNA 122³¹ or HCV translation by targeting its Internal Ribosomal Entry Site (IRES).^{26,32} Many efforts have been devoted to the development of new drugs able to inhibit the HCV translation by hampering its IRES functions. In particular, the discovery of ligands targeting the IRES IIId loop is highly relevant as this domain is of utmost importance for the interaction with the 40S ribosomal subunit. In addition, as more structural information emerges and the model of HCV IRES-

mediated translation becomes more refined, IRES sites are likely to pave the new wave of RNA targets in the context of viral diseases. Known ligands for HCV IRES encompass antisense oligonucleotides and RNA aptamers as well as small-sized organic molecules; they will be presented in the next section.

4.1.5 RNA Ligands

In the last years, the aim to access new interferents with ability to inhibit specific RNA functions for therapeutic purposes bolstered the research focusing on the development of novel and more potent RNA ligands. In this context, a large variety of ligands featuring original structures have been reported. More precisely, three main approaches have been explored to access potent ligands: oligonucleotides, peptides and small-sized interferents.^{33,34} In this section, it will be presented the state of the art concerning RNA ligands, with an emphasis, when available, on structures targeting the HCV IRES domains. The bibliographic review will be focused on oligonucleotide and small-sized ligands.

4.1.6 Oligonucleotides (ONs)

Oligonucleotides are short fragments of nucleic acids consisting of a few tens of nucleotides. ONs display the ability to hybridize with a genomic DNA and RNAs presenting complementary sequences. Through this hybridization process ONs may modify gene expression.³³ Thus, they have been extensively explored as RNA ligands, essentially due to the extremely high affinity and specificity related to the target RNA sequence. Despite their high therapeutic potential for a wide variety of diseases, the use of ONs remains limited, especially due to the high rate of degradation by nucleases and difficulties to access more efficient vectorization.

Few years ago, apart from some encouraging results, only few examples of therapeutic oligonucleotide agents had shown satisfactory results in clinical trials and were available on the market. However, the fast development of oligonucleotide-based therapies abruptly change this scenery and place them in the spotlights of the newest discovered drugs, specially to treat extremely rare diseases. The first antisense oligonucleotide agent Patisiran (Anylam pharmaceuticals) had been recently approved by the FDA (2018) for the treatment of familial amyloid polyneuropathy (FAP), an hereditary neurodegenerative disease.³⁵

Apart from this, the lack of stability, the mode of administration, or the high risk of side effects are still some of the issues to overcome in ON-based therapies. Chemical modifications on nucleotide motifs (ribofuranose, phosphates) have emerged as a powerful solutions to increase stability to nucleases processing. The application of ONs also suffers from delivery issues. Limitations to cross cellular and nuclear membranes and avoid fast processment by the reticuloendothelial or lysosomal systems are other obstacles to circumvent before the ON reaches its target. Intensive research has been made in this context. Physical (sonoporation, magnetofection) and chemical (conjugation with cell penetrating peptides-CPP, small peptides, cationic lipids and polymers) methods have been developed to ensure more efficient delivery systems.³³ Nevertheless, vectorization methods are out from the scope of this thesis and will not be discussed.

In the following paragraphs, an overview concerning oligonucleotide-based RNA ligands is presented, with a focus on the use of antisense approaches. Other nucleic acid-based approaches (aptamers, ribozymes, siRNA, shRNA) will be briefly highlighted. Examples of each group of ON ligands will be linked to context of HCV.

4.1.6.1 Antisense Oligonucleotides (ASOs)

The first report concerning the inhibition of viral replication using oligonucleotides dates from 1978. Since then, the reports of Zamecnik and Stephenson^{36,37} have been stimulating the quest for antisense oligonucleotides.³⁸ Essentially, antisense oligonucleotides are designed to match specifically with a target RNA. The pairing of both strands, the oligonucleotide and the target RNA sequence can provide two different mechanisms of RNA inhibition: the physical blockage of translation or by a fast degradation of the mRNA through ribonuclease H (RNases H) processing.

The activity of ASOs is dependent of a diversity of factors such as their efficiency to enter in the cell, the stability of the formed complex with the target RNA and their resistance to enzymatic processing. Standard RNA and DNA ASOs normally display poor uptake and incredibly short intracellular and serum lifetimes. Thus, in the recent years, a great diversity of modified ONs has emerged as alternative to circumvent this low potency of "natural" RNA and DNA oligonucleotides. Modifications on the sugar and phosphate backbones have proven to be positive, increasing the resistance of ASOs against enzymatic degradation. First generation of ASOs bearing phosphorothioate, boranophosphate, oxepane, and fluoro-arabino (FANA) linkages have been described.³⁹ They successfully activate RNase H upon binding to

the target mRNA. Other ONs featuring modified sugar scaffolds, including 2'-*O*-(2-methoxyethyl) and 2'-OMe RNAs or locked nucleic acids (LNA) (second generation), peptide nucleic acids (PNA) and phosphoramidate morpholino oligomers (PMO) (third generation) are presented below.

4.1.6.1.1 Phosphorothioate-Modified Oligonucleotides

The use of modified phosphate backbone has emerged as a powerful alternative to the native phosphodiester linkages. Variants on the phosphate backbone have been developed to alter chemical properties of standard ONs and therefore overcome the major challenges associated to the use of oligonucleotides *in vivo*: the delivery to the interior of the cell through the plasma membrane, which is mostly impermeable to polar molecules and minimize intracellular nuclease degradation.

Phosphorothioate oligonucleotides, referred to as S-oligonucleotides, is today the most widely-used phosphate backbone variant incorporated in oligonucleotides (**Figure 4.10**). One important feature concerning the O to S replacement on the phosphate backbone comes with the generation of a stereogenic α -phosphorus center (**Figure 4.10**), resulting in formation of a mixture of diastereoisomeric oligonucleotides.⁴⁰ This property can assign for different effects. Essentially, stereochemistry highly affects the pharmacologic activity of oligonucleotides as it does for small drugs. In a mixture of S-oligonucleotides isomers, it is logic to expect that only few diastereoisomers are responsible to contribute for its activity. The great majority of them, in fact, lack from sufficient hybridization ability to the target RNA sequence, mainly due to conformational limitations. In addition, it had already been demonstrated that, in certain cases, the phosphorothioate linkages with S configuration can impute enhanced lipophilicity and metabolic stability when compared with the R isomers. Thereby, the control of the stereochemistry of S-oligonucleotides has been recognized to be important to improve their biological activity. Recently, methods that enable the synthesis of enantiomerically pure S-oligonucleotides have been described.⁴⁰

The use of phosphorothioate linkages have been providing a remarkable improvement on the ON stability against a variety of extra and intracellular nucleases. Essentially, this modified phosphorothioate oligodeoxynucleotides promote target RNA degradation on the hybridization sequence via cellular RNase H enzymes.⁴¹

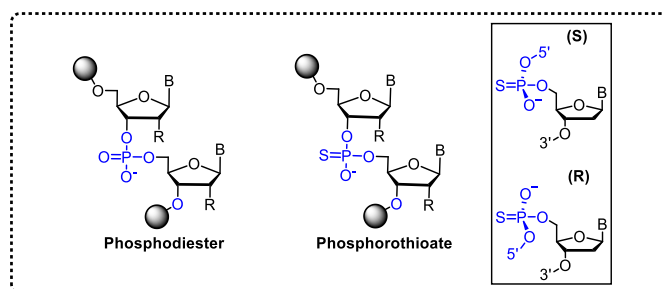


Figure 4.10. Structure of phosphorothioate oligonucleotides.

By the end of 20th century, FDA approved the Fomiversen (ISIS 301012), a first-generation antisense agent and is used for the local treatment of retinitis cytomegalovirus in patients suffering from acquired immunodeficiency syndrome (AIDS).⁴² During the same period, ISIS Pharmaceutical also reported sequential studies and clinical trials with a family of phosphorothioate ASOs as active therapy to inhibit HCV translation. Two oligonucleotides, ISIS 6095 (5'-GCCTTTCGCGACCCAACACT-3') and ISIS 6547 (5'-GTGCTCATGGTGCACGGTCT-3'), containing 20 nucleotide sequences, were reported. They were effective to inhibit HCV polyprotein translation by targeting the stem-loop of IRES III_d subdomain (ISIS 6095) and to sequences containing the AUG initiation codon (ISIS 6547). Both oligonucleotides caused concentration-dependent reductions in HCV RNA and protein levels, with IC₅₀ in the nanomolar range (100 to 200 nM).⁴³

The use of another ASO, ISIS 14803, a parent oligonucleotide bearing 5-methyl-cytosine residues with the same nucleotide sequence as ISIS 6547 was also reported. When evaluated in HCV-IRES-infected mouse model, ISIS 14803 showed potency comparable to that of the parental oligonucleotide ISIS 6547.^{41,44} However, clinical trials (phase II) demonstrated weak plasma HCV RNA reduction in patients treated during 4 week. This insufficient potency, which was also accompanied by side effects led to discontinuation of the development effort of this ASO in 2005.^{45,46}

4.1.6.1.2 Locked Nucleic Acids (LNA)

Locked nucleic acids have emerged as powerful therapeutic tools for antisense approaches. Recently, Santaris Pharma a/s, a Danish pharmaceutical company arose as the unique company to have the rights to explore LNA technology. LNA are ONs containing a modified RNA nucleotides, which presents an extra bridge to connect the 2' oxygen and 4' carbon on the ribose platform. This structure locks the ribose in the 3'-*endo* conformation (**Figure 4.11**),

often found in A-form duplexes. This locked ribose can afford enhanced base stacking and backbone pre-organization and improve the hybridization properties of the oligonucleotide.⁴⁷

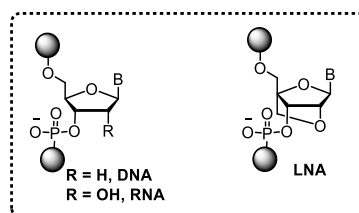


Figure 4.11. Structure of a LNA nucleotide.

Developed by Santaris Pharma a/s to treat chronic HCV, Miravirsen is a 15-nucleotide locked nucleic acid-modified antisense oligonucleotide bearing phosphorothioate linkages. It targets a non-coding miRNA of the host cell and is currently in clinical trials (phase II). It displays high affinity and specificity to target the 5'-UTR of the liver expressed mature miR-122. Miravirsen can sequester and thus inhibit miRNA-122, which is important for the stability and propagation of HCV RNA. Basically, it binds to two target sites (S1 and S2) in the 5'-UTR region of the HCV genome, and forms a miRNA-122–HCV duplex that enhances HCV stability to nucleolytic degradation or from host innate immune responses.³¹ Miravirsen shown low nanomolar *in vitro* affinity to miRNA-122.⁴⁸ *In vivo* analysis has already shown that Miravirsen can provide long-lasting viral suppression and significant virological responses without evidence of resistant mutations at the two miR-122 binding sites of the 5'-UTR of the HCV genome.^{49,50}

4.1.6.1.3 The 2'-O modified Nucleic acids

In analogy with LNAs, the introduction of modifications on the 2'-O position of the ribose also locks the sugar in a C3'-endo conformation (**Figure 4.12**).⁵¹ This kind of modification also proved to be efficient in antisense therapies. Recently, FDA approved the use of Mipomersen (ISIS301012, January 2013), a second-generation gapmer antisense oligonucleotide of 20 nucleotides length with phosphorothioates backbone. It inhibits the translation of the apolipoprotein B, which is a target in the treatment of homozygous familial hypercholesterolemia. Mipomersen binds to the mRNA of apolipoprotein b and activates its degradation by the RNase H. In addition to its phosphorothioate modification, it also features

the 5-methylation of all of its cytosine and uracil nucleobases. Moreover the ribose unit of the five first and five last nucleotides are modified by a 2'-*O*-(2-methoxyethyl) groups.⁵²

Another important example of a 2'-*O*-methyl phosphorothioate oligonucleotide it was developed BioMarin to treat Duchenne Muscular Dystrophy (DMD). This antisense agent did not reach the market as the phase 3 clinical trials (vs placebo) in more than 300 patients failed to demonstrate sufficient clinical benefit. Moreover, this trials also highlighted a moderate toxicity of Drisapersen, thus leading to the discontinuation of its.³³

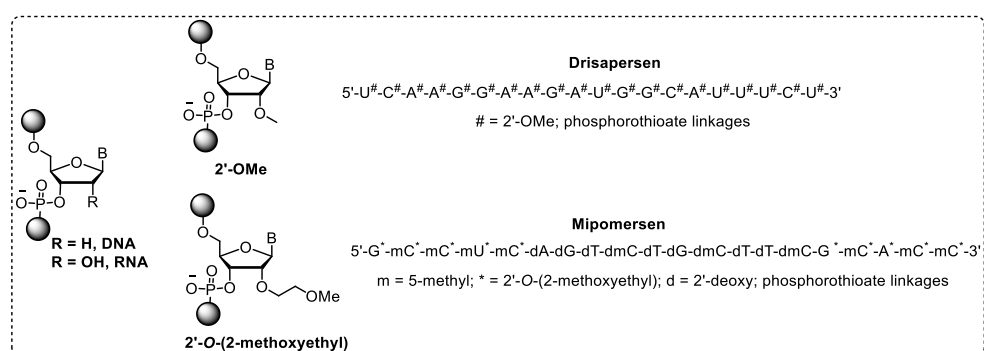


Figure 4.12. Structure of 2'-*O* modified ribonucleotides. The primary sequence of Drisapersen and Mipomersen oligonucleotides are also highlighted.

Concerning the development of 2'-*O*-alkylated ASOs against HCV, ISIS studied the antisense agent ISIS 11155 (5'-TTTAGGATTCGTGCTCATGG-3'), a 2'-*O*-(2-methoxyethyl) phosphorothioate oligonucleotide. ISIS 11155 demonstrated improved hybridization affinity to target RNA sequences spanning the AUG initiation codon and a concentration-dependent reduction in HCV core protein expression in H8Ad17c cells, with IC₅₀ of 62 nM. The results were comparable to the phosphorothioate-containing ON analogue of the same sequence.⁴³

Two other groups reported their advances in targeting IRES domains of HCV using 2'-*O*-methyloligoribonucleotides. The group of J. J. Toulmé⁵³ investigated a series of 2'-*O*-methyloligoribonucleotides as inhibitors of HCV IRES by targeting the domain IIIId. *In vitro* studies performed in rabbit reticulocyte lysates demonstrated the high potency of short ONs sequences (from 14 to 17-nucleotides) to inhibit HCV translation in a low nanomolar range (IC₅₀ < 10 nM). It acts by displacing the formation of 40S-IRES initiation complex, which is indispensable for the HCV translation. In a different work, Berry and co-workers⁵⁴ validated the HCV IRES pseudoknot located close to subdomain IV as a potential druggable target using 2'-*O*-Me ASOs. Their *in vitro* inhibitory activity for the IRES-mediated HCV translation reached IC₅₀ values ranging from 140 to 400 nM.

4.1.6.1.4 Peptide Nucleic Acid (PNA)

Peptide nucleic acid (PNA) are artificially synthesized nucleic acid analogues bearing DNA or RNA bases, in which the pentose scaffold is replaced by *N*-(2-aminoethyl)-glycine units which serve as the backbone (**Figure 4.13**). The purine and pyrimidine bases are attached to the backbone by carboxymethyl linkages.⁵⁵ The PNA skeleton is neutral. Thereby PNA/DNA duplexes are more stable than a conventional DNA/DNA due to the absence of electrostatic repulsion.^{56,57} PNAs are relatively flexible compared to DNA or RNA. They are well known for their ability to invade DNA duplexes and have been considered effective antisense agents in cultured cells.^{55,58} Nevertheless, their ability to invade highly structured RNA sequences inside cells is still not established.

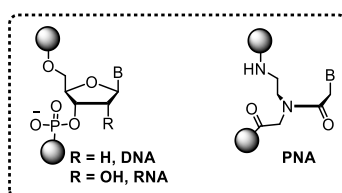


Figure 4.13. Structure of PNA nucleotide.

Few works have been reporting the use of PNAs as antisense agents to inhibit HCV translation. The group of Olivier Hantz⁸ studied the *in vitro* (cultured cells and in rabbit reticulocyte lysate assay) ability of short PNAs (6-10 nucleotides length) to inhibit HCV translation by targeting the IRES subdomains III_d and IV. High specificity for the target sequences were reported (no activity for sequences bearing a single base deletion) and a potent inhibitory activity of IRES was achieved (IC₅₀ ranging between 20-40 nM). Negative control experiments also demonstrate that no inhibition was obtained by using phosphorothioate oligonucleotides of same sequence; hence supporting the importance of the PNA backbone for a potent inhibitory activity.

Nulf and Corey reported a comparative study between the intracellular activity of PNA and chimeric LNA antisense oligonucleotides. This was performed on several IRES domains.⁵⁹ Their key findings was that the HCV IRES translation is susceptible to be inhibited by antisense PNAs when targeting a wide range of sequences (stem-loop in domains II and III_d and pseudoknot in domain IV). Both, PNA and LNA showed a dose-dependent inhibitory

activity, with EC₅₀ values ranging between 50-150 nM. Nevertheless, the use of PNAs for *in vivo* inhibition of HCV IRES remains undisclosed.

4.1.6.1.5 Phosphoramidate Morpholino Oligomers (PMO)

PMOs (morpholinos) are synthetic nucleic acids in which the pentose scaffold is replaced by a morpholine moiety. They have been considered as promising third-generation antisense molecules that possess favorable hybridization, nuclease stability, and toxicity profiles.⁶⁰ Structurally, morpholinos also possess nucleobases attached to the morpholine scaffold. These morpholine rings are linked to each other by a different phosphorodiamidate linkages instead of phosphodiester (Figure 4.14). The replacement of the anionic phosphate groups with the uncharged phosphorodiamidates eliminates the ionization in the usual range of physiological pH; thus morpholinos are uncharged in the organisms or cells. PMOs are most commonly used as single-stranded oligonucleotides, so that heteroduplexes can be formed between the morpholino strand and a complementary DNA/RNA sequences through Watson-Crick base pairing.⁶¹ Moreover, PMOs do not inhibit gene expression by activation of RNase H, they rather stop protein translation through steric blockade.⁶⁰

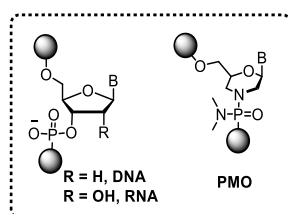


Figure 4.14. General structure of the phosphoramidate morpholino nucleotide unit.

Morpholinos have been used in biological studies. Efficacy, low toxicity, sequence stability were verified in different works.^{62,63} Moreover, it has been evidenced that gene expression in embryos of Zebrafish could be modified using PMOs. Recently, FDA approved the therapeutic use of Eteplirsen, a 30-mer PMO to treat DMD.⁶⁴

Morpholinos have also been evaluated in the context of HCV IRES inhibition function. A PMO complementary to the region of the starting AUG codon of the HCV IRES has been reported by Jubin and co-workers. It reduces by more than 80% the protein translation in rabbit reticulocyte lysates.⁶⁵ In a different study, the group of McCaffrey ascertained the use PMOs to block HCV translation *in vivo*. They reported the potential therapeutic application

of a 20-mer PMO (5'-GAGATTTACGATTCATGATC-3') to inhibit HCV IRES activity and protein translation in preclinical mouse models. Inhibitory activities greater than 95 % after six days treatment were achieved.⁶⁶

4.1.6.2 Aptamers

Aptamers are RNA or DNA oligonucleotides able to selectively bind to specific target molecules. They can be generated against different targets such as small organic molecules, peptides, proteins or nucleic acids.⁶⁷ Usually, aptamers are characterized by high affinities for their target and remarkable specificity. Easy to synthesize, they can be converted into interesting tools including diagnostic techniques by grafting different functional scaffolds.

Aptamers can be identified using libraries containing up to 10^{15} different sequences by an *in vitro* combinatorial selection method called SELEX (*systematic evolution of ligands by exponential enrichment*). Most aptamers are composed of RNA, because of their ability to adopt varied and complex structures, which allows to create specific domain with distinct geometries on their surfaces, allowing to bind different ligands. The SELEX approach only needs a library of ONs in which the sequences on the 3' and 5' extremities are known and the internal sequence is aleatory.^{68,69}

RNA aptamers have been designed to target all HCV IRES domains. The subdomain III_d have been identified as the most targetable region; aptamers with activity comparable with ASOs have been identified in *in vitro* assays. Nevertheless, no consistent *in vivo* studies have been described until this moment. Kikuchi and co-workers strongly contributed for this field. They reported the inhibition of the IRES-dependent HCV translation using aptamers to target the apical loop of subdomain II⁷⁰ and domains III_d and IV.^{67,71} Their aptamers display low nanomolar dissociation constants ($4.4 \leq K_d \leq 130$ nM).

4.1.6.3 Short interfering RNAs (siRNAs) and short hairpin RNAs (shRNAs) - RNA Interference Mechanism

RNA interference (RNAi) is one of the most important cellular pathways that regulates gene expression in live organisms.⁷² It is a sequence-specific cellular post-transcriptional gene silencing (PTGS) mechanism that also acts as an intrinsic defense against viral pathogens.⁷³ Just after its first report in mammalian cells (2001), intense research on synthetic siRNAs and shRNAs started to be developed to access RNAi-based therapies for cancers, tumors and viral

pathologies.^{23,74} The simplified mechanism of RNA interference is presented on **Figure 4.15**. The pre-formed shRNA (or dsRNA) is initially processed by Dicer (step 1), to yield a double stranded siRNA which is loaded in the RNA-induced silencing complex (RISC) (step 2). Then, the antisense strand directs RISC to the endogenous complementary mRNA (step 3) and eventually causes its degradation through its nuclease activity (step 4). The overall result is the inhibition of the protein translation that was encoded in the targeted mRNA.

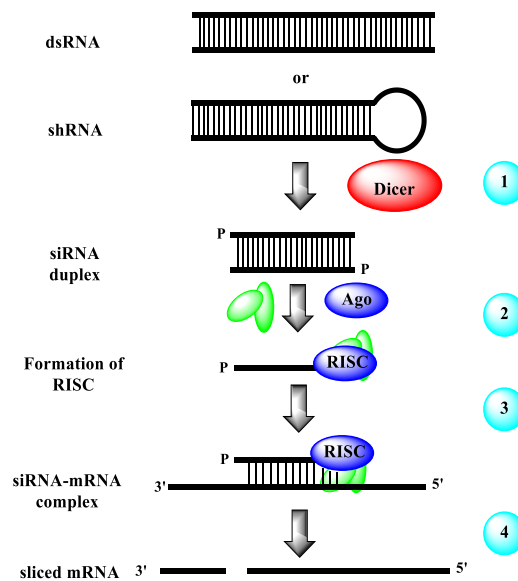


Figure 4.15. Simplified mechanism of RNA interference. Steps 1-4 are highlighted in the text.

The targeting of HCV IRES sequences by siRNA and shRNA have shown efficiency for *in vitro*^{75,76} and *in vivo*⁷⁷ inhibition of viral translation by the RNA interference mechanism. Nevertheless, similarly to other oligonucleotide-based therapeutic approaches, a more effective method for their delivery inside the cells remains an important limitation.

4.1.7 Small Molecules

The targeting of RNA with small molecules has emerged as a potential and complementary alternative to the use of oligonucleotides. In this context, small-sized RNA ligands have been developed⁷⁸ and studied to explore four major classes of targets: mRNA, antibacterial, antiviral and ncRNAs. Within each of these RNA targets, several approaches have been exploited to achieve the desired biological effect, as inhibiting RNA-protein interactions or preventing protein production by binding to a particular RNA sequence. Heretofore, the great

majority of the ligands have been discovered after classical screening of libraries of molecules. Nevertheless, in the recent years special groups of RNA binders, displaying high specificity and affinity *in vitro* and *in vivo* to different RNA targets have been identified by using new drug design approaches.⁷⁹⁻⁸³

Thereafter, in this section, we will provide a bibliographic overview of the different groups of RNA ligands. Initially, the use aminoglycosides as RNA binders will be stated. Later, the state the art concerning the small molecule ligands that targets HCV IRES domains will be presented. Finally, we will introduce the concept of the triplex-based rational design. This is a smart approach that have been successfully used to produce innovative RNA binders. Moreover, it will be the central part of our studies dealing with HCV ligands, which will be presented later on this chapter.

Nowadays, ligands targeting the bacterial ribosomes are the great majority of clinically used RNA ligands. The relative expression of RNAs in the cell constitutes the main barrier to access specific ligands to act in different RNA targets. Antibiotics are the major class of small-sized RNA ligands available on the market, many of them are targeting rRNA, the most abundant RNA inside the cell. However, considering that non-coding RNAs represents only a minor fraction of all expressed RNAs (< 5%), the design of ligands to target ncRNAs in live organisms remains a tremendous challenge.

Optimal RNA ligands should feature a combination of different parameters in order to interact and be accommodated in specific "pockets" or domains within the target RNA. A delicate equilibrium between the ability to form electrostatic, hydrogen bonds, van der Waals or π -stacking interactions is fundamental to achieve optimal binding affinity and selectivity towards certain RNA targets.^{84,85}

4.1.7.1 Aminoglycosides

Aminoglycosides are one of the major class of antibiotics. Streptomycin, discover in the first half of 1940s (isolated from soil bacteria *Streptomyces griseus*) is the first example reported to date. Aminoglycosides interact with the bacterial rRNA, and impede the protein translation step. They consist of structures containing two or more aminosugar units that are connected to each other by glycosidic linkages (**Figure 4.16**). In several aminoglycosides, the molecular structure is built up from a central 2-deoxystreptamine core (2-DOS, **126**), onto which different sugar or aminosugar scaffolds are seembled. Structurally, they can be divided into

three classes: the monosubstituted 2-DOS, and the 4,6 and 4,5-disubstituted derivatives. Neamine (**128**) and nebramine are the regular aminoglycoside motifs that display antibiotic activity whereas 2-DOS itself does not show any biological activity.^{34,86}

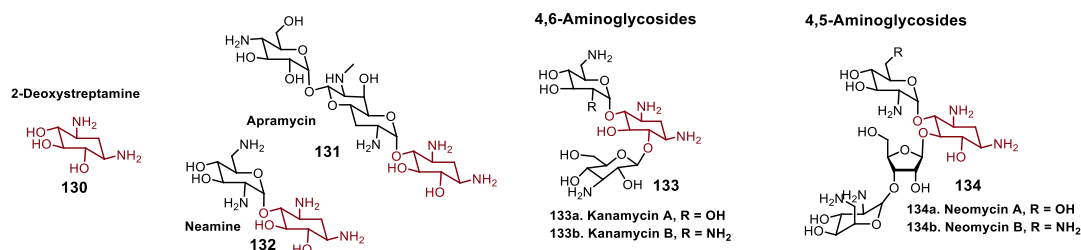


Figure 4.16. Examples of aminoglycoside antibiotics.

Aminoglycosides are known to be strong RNA binders. In 1987, Moazed and Noller⁸⁷ showed that many aminoglycosides antibiotics bind directly to the bacterial 16S rRNA A-site, disrupting the protein synthesis by guiding the incorporation of wrong amino acid in the sequence on the translated polypeptide. Aminoglycosides are polycationic molecules under physiological conditions, displaying the ability to interact with RNA via electrostatic interactions and hydrogen bonding. However, while they bind to RNA structures with extremely high binding affinities, their greatest limitation is related to the lack of selectivity for specific RNA sequences. This is mainly a consequence of their prominent non-specific mode of interaction which is dominated by electrostatic forces.

Recently, it has been demonstrated that both natural and semi-synthetic aminoglycosides can also interact with different ncRNA such as pre-miRNA and noncoding regions of viral RNAs (**Figure 4.17**).^{88,89} Aminoglycosides have also been explored as pharmacophore scaffolds to rationally design new ligands using a triplex-based rational design (*Section 4.1.7.3*). Thus, many efforts have been done to access synthetically modified aminoglycosides. The introduction of side functionalizations (**135**),⁹⁰ the guanidilation of amino groups (**136**),⁹¹ the conjugation of two or more aminoglycosides through a covalent linker (**137**)⁹² or the replacement of 2-DOS or aminosugars by mimetic scaffolds (**138**)⁹³ are some of approaches that were used to access more potent antibiotics. These approaches are also employed in the context of the antibioresistance issue and the conception of viral translation inhibitors. (**Figure 4.17**).

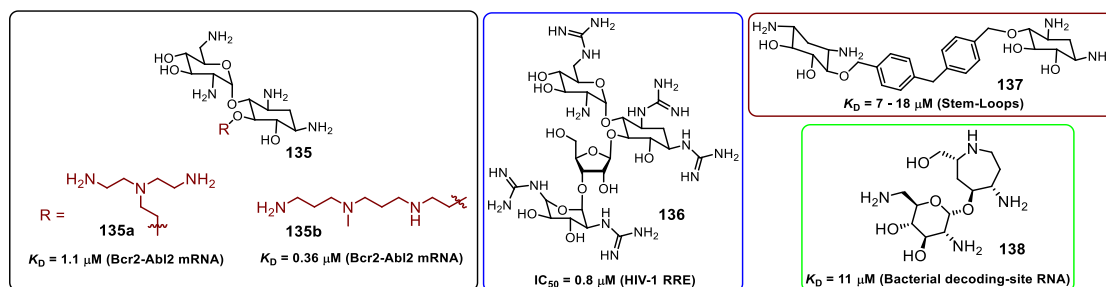


Figure 4.17. Structure of synthetic aminoglycosides.

In a consistent work, Wong's group⁹⁴ reported a new library of 4,6-linked tobramycin (**139**) analogues with various aminosugars connected to the 6-position of 2-DOS (**Figure 4.18**). Potent aminoglycosides were identified after screening several important RNA targets present in the bacterial, viral or human RNAs. Many analogues proved to have high affinity and selectivity to different domains of HCV IRES, to the ribosome 16S A-site and to human oncogenic mRNAs sequences, with low to submicromolar dissociation constants.

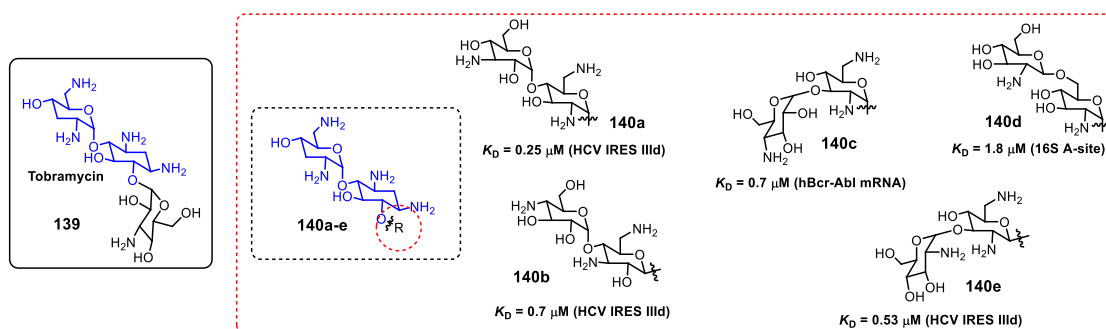


Figure 4.18. Structure of synthetic tobramycin analogues with affinity to different RNA targets.

4.1.7.2 HCV IRES Small-Sized Ligands

Small-sized ligands to target specific domains of HCV are seldom found in the scientific literature. Heretofore, the known ligands have been identified by classical screening against the different IRES domains. The majority of ligands display high affinity to the HCV IRES subdomain IIa (micro to submicromolar range). Two research groups identified potential molecules to bind to subdomains IIIId and IIIe in the micro molar range after high throughput screening. In 2004, the first group, Gooding and co-authors,⁹⁵ identified 7 potential compounds with low micromolar affinity ($K_d < 50 \mu\text{M}$) to bind to the subdomain IIIe after the screening of more than 300,000 molecules using flow injection mass spectrometry (FIA-MS). The second group, Bough and co-workers,⁹⁶ also identified ligands with high affinity to

the IIIId subdomain; submicromolar dissociation constants were reported. However, the chemical structure of these ligands is still undisclosed.

Amino-substituted phenazines and biarylguanidines were identified as inhibitors of IRES activity *in vitro*. Wang and co-workers⁹⁷ reported that amino-substituted phenazine analogues act as moderate HCV IRES inhibitors after molecular screening and SAR developments (**Figure 4.19**). However, apart from the moderate activity against HCV IRES, the compounds lacked selectivity and were moderately toxic, and no improvement on the inhibitory activity was achieved after conventional SAR studies. In a similar approach, Jefferson and collaborators⁹⁸ disclosed biarylguanidines as powerful *in vitro* inhibitors of HCV IRES. Two hits (compound **142** and **143**) displaying micromolar inhibitory activity were identified after high throughput screening and were subjected to hit-to-lead optimization. This enabled the access to biarylguanidine ligands of enhanced activity (**144**, $IC_{50} = 2.1 \mu\text{M}$, **Figure 4.19**).

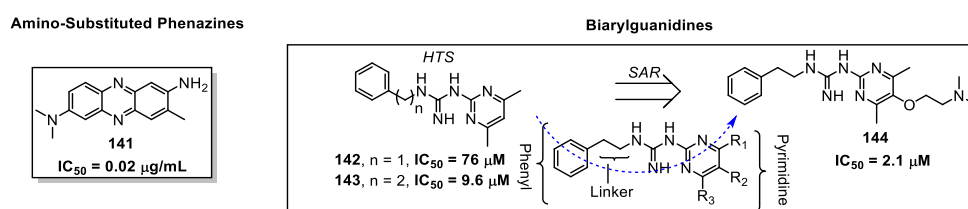


Figure 4.19. Structure of the phenazine and biarylguanidines identified to inhibit HCV IRES activity.

During the last 15 years, two important families of RNA ligands derived from 2-aminobenzimidazole and DAP (3,5-diamino piperidine) dipeptides have been discovered and extensively studied. In 2005, the group of Seth (ISIS Pharmaceuticals)⁹⁹ identified the first 2-aminobenzimidazole analogue **145** as a potent ligand and pharmacophore to target HCV IRES subdomain IIa with micromolar affinity ($K_d \approx 100 \mu\text{M}$). Tedious SAR optimization led to the discovery of the analogues **147** and **148**, displaying submicromolar *in vitro* affinities to the same RNA target and low micromolar inhibition of HCV replicon ($EC_{50} < 5 \mu\text{M}$). Of note, no cytotoxicity in Huh-7 cells (human hepatocyte-derived carcinoma cell line) was observed (**Figure 4.20**).¹⁰⁰

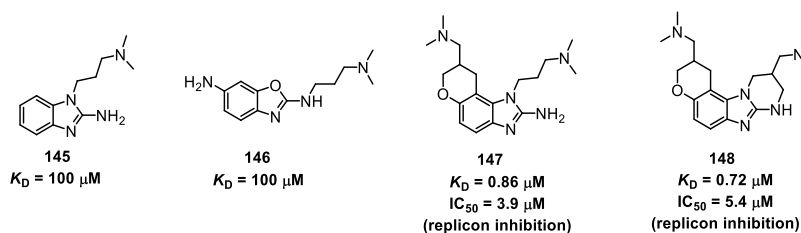


Figure 4.20. Structure of Seth's benzimidazole ligands of HCV IRES IIa subdomain.

Benzimidazole analogues are selective inhibitors of IRES-driven viral translation. Hermann's group^{101,102} demonstrated that compound **147** is able to bind to the bulge of subdomain IIa (**Figure 4.21-A**) and induce sufficient conformational changes to disrupts the IRES function and block the viral protein synthesis. Moreover, the SAR analysis for these families of benzimidazoles were perfectly correlated with the crystal structure of the RNA complex of the IRES IIa subdomain with benzimidazole **147** (**Figure 4.21-B**). This permitted the elucidation of key aspects to optimize molecular structure of new generation of ligands. Derivatives lacking the amino group on the 2-position of the benzimidazole or the *N,N*-dimethylaminopropyl substituent on *N1* were not able to interact with the target RNA (compound **149**, **Figure 4.21-D**). Indeed, the crystal structures underpins the fact that both groups are involved in the binding. Moreover, analogues bearing basic substituents on the 6-position displayed higher affinities.

Apart from the high activity, the laborious synthetic route, with only 10 % overall yield to access the tetracyclic benzimidazole **148** remains a great limitation to further development and implementation of these benzimidazole analogues in the industry.

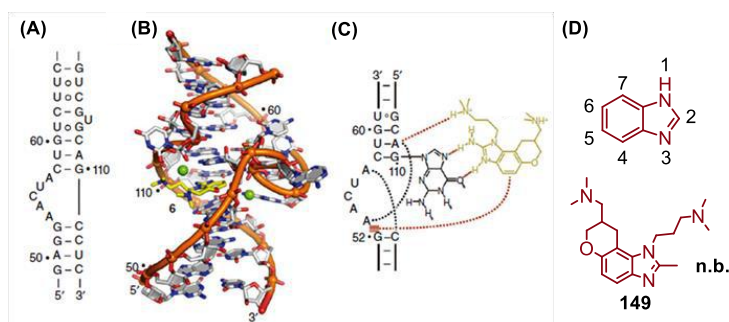


Figure 4.21. Binding interaction between benzimidazole **147** and HCV IRES IIa subdomain. (A) Genomic sequence of HCV IRES subdomain IIa. (B) Crystal structure of benzimidazole-HCV IRES subdomain IIa complex. (C) Representation of the modes of interaction between **147** and HCV IRES IIa subdomain. (D) Positions on the benzimidazole core; structure of the analogue of **147** which lacks the amino group on the 2-position.

The second class of ligands to target the IIa subdomain was originally reported by the group of Hermann. Carnevali and co-workers¹⁰³ described the synthesis and activity of dipeptide containing the 3,5-dimethylamino piperidine core (DAP), a mimetic analogue of 2-deoxystreptamine (2-DOS). Two potential ligands derived from basic aminoacids were identified (**150** and **151**, **Figure 4.22-A**). They display inhibitory activity (75% at 10 μM) against HCV replicon and low micromolar binding affinities to the target RNA sequence. Differently from the benzimidazole ligands, the DAP derivatives interact mainly *via* a strong electrostatic component. In fact, both DAP dipeptides bind in competition with the structural Mg^{2+} ions. Nevertheless, the exact mechanism of IRES inhibition remains undisclosed, although it might involve the displacement of magnesium cations, thus locking the flexible RNA in a L-shaped conformation which prevents ribosome release (**Figure 1.22-B**).

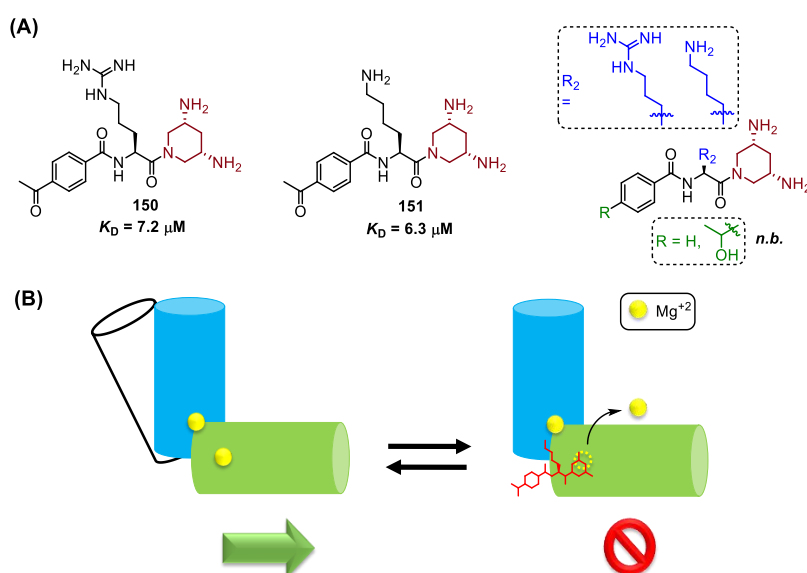


Figure 4.22. (A) Structure of DAP containing dipeptide ligands for HCV IRES subdomain IIa. (B) Putative mode of action of DAP dipeptide ligands bond to the IIa subdomain.

(Adapted from Carnevali *et al.*¹⁰³).

In 2014, new benzimidazole and benzoxazole derivatives were reported. Two structure-binding activity relationship (SBR) studies were described by Hermann group.^{104,105} Series of *N*-1-aryl-substituted 2-aminobenzimidazole and benzoxazole derivatives (**Figure 4.23**) were synthesized and their binding affinity against the HCV IRES IIa subdomain was evaluated. However, only moderate to weak affinities were observed considering both series of ligands

and no clear structure-binding relationship could be drawn concerning the benzimidazole analogues.

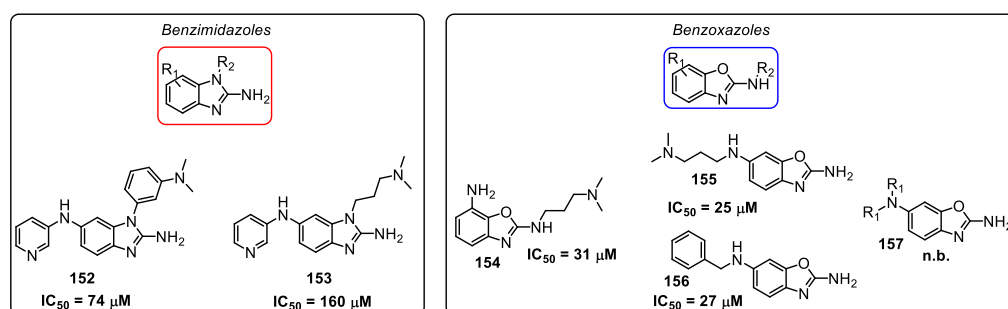


Figure 4.23. Structure of the selected benzimidazole and benzoxazole ligands targeting the HCV IRES IIa subdomain.

4.1.7.3 RNA Ligands Developed Using a Triplex-Based Rational Design

In the last years, our group also contributed to this field by implementing a new approach to design RNA ligands exploiting a triplex-based rational design. Multimodal ligands were designed by combining specific artificial nucleobase motifs as site specific recognition elements and functional groups, as aminoglycosides or amino acids, to strengthen the ligand affinity to the target RNA. Ligands with high binding affinity and selectivity were reported to target noncoding sequences of viral RNAs and pre-miRNAs.

Artificial nucleobases are heteroaromatic scaffolds able to form triplets with the Watson-Crick T(U):A and C:G base pairs. They have been intensively reported in the context of antineoplastic therapies, to design triplex-forming oligonucleotides (TFOs), by targeting base pair mismatch in oligopurine-oligopyrimidine double helical DNAs to enhance DNA triplex stability. There are many reported motifs already described, with the ability to form nucleobase triplets via one, two or three Hoogsteen hydrogen bonds (**Figure 4.24**).¹⁰⁶ While the majority of these artificial nucleobases featuring three Hoogsteen hydrogen bonds with the canonical Watson-Crick base pairs have been developed to target C:G base pairs, our group addressed and solve some limitations by reporting the S-nucleobase. This artificial nucleobase, bearing a phenylthiazole scaffold, is able to form three Hoogsteen H-bonds to recognize T:A base pair inversion with T_m values close to those of canonical triplexes.^{107,108}

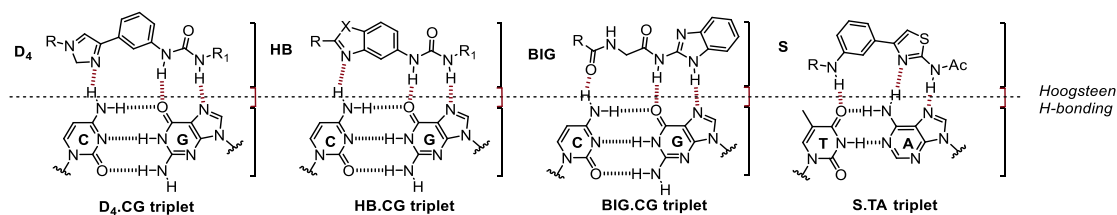


Figure 4.24. Some artificial nucleobases featuring three Hoogsteen H-bonds with the canonical T(U):A and C:G base pairs.

Two sets of ligands were designed to target the stem-bulge of HIV-1 TAR RNA (**Figure 4.25**). A series of ligands with low micromolar affinity and high specificity to HIV-1 TAR RNA were firstly reported in 2010.¹⁰⁹ The ligands were designed by assembling the S-nucleobase with series deoxyribose platforms containing lysine and arginine residues connected to the 3' and 5' positions. No influence from the anomeric configuration was observed (**Figure 4.25-A**, compounds **158a-d**). More recently, several amino acid residues were directly connected to the S-nucleobase scaffold in series containing one or two residues. Ligands with sub micromolar binding affinities to the same RNA target were reported. In addition, the ligands were effective to inhibit HIV-1 replication in model cells, with low to sub micromolar IC₅₀ values and negligible cytotoxicity (**Figure 4.25-B**, series of compounds **159**, **160** and **161**).¹¹⁰

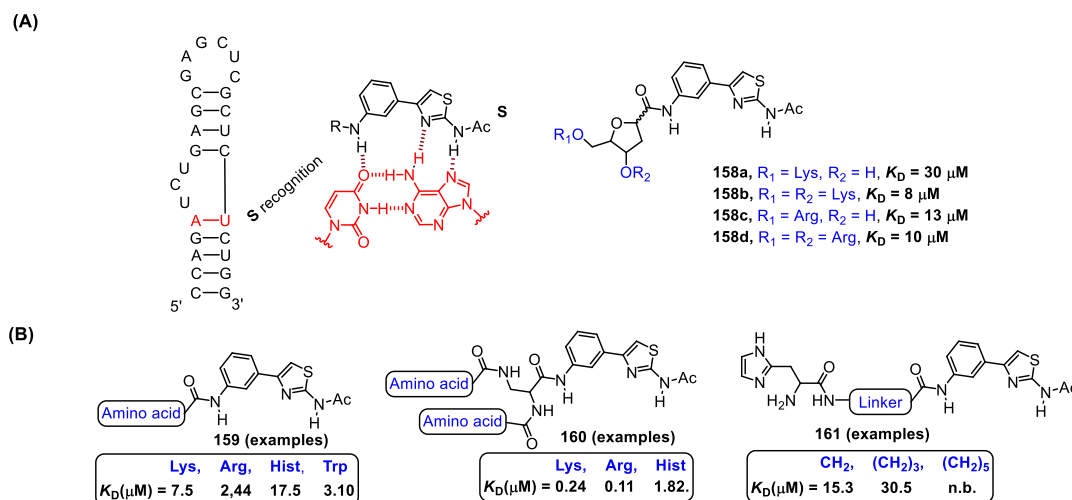


Figure 4.25. Ligands derived from S-nucleobase to target HIV-1 TAR RNA.

Ligands to inhibit the expression of oncogenic miRNAs have also been developed using similar approach (**Figure 4.26**). A first series of aminoglycoside-artificial nucleobase conjugates were designed to target pre-miRNA-372 and 373 by the assembling neomycin to

different artificial nucleobase scaffolds (**Figure 4.26-A**).¹¹¹ Ligands displaying low nanomolar binding affinities to pre-miRNA-372 were obtained. In addition, efficient *in vitro* inhibition of pre-miRNA-372 processing by Dicer into the corresponding oncogenic miRNA-372 was observed ($2.42 < IC_{50} < 25.2 \mu M$). Moreover, the neomycin-S-nucleobase conjugate also inhibits the proliferation of human gastric adenocarcinoma AGS cell line in a dose dependent manner.

In a complementary work, ligands with improved biological activity were synthesized by assembling three aminoglycoside scaffolds (neomycin, neamine and 2-DOS) to modified S-nucleobase analogues through different linkers (**Figure 4.26-B**). The increase of the aminoglycoside size strongly enhances both, the binding affinity (K_D) and Dicer inhibition activity (IC_{50}) values.¹¹²

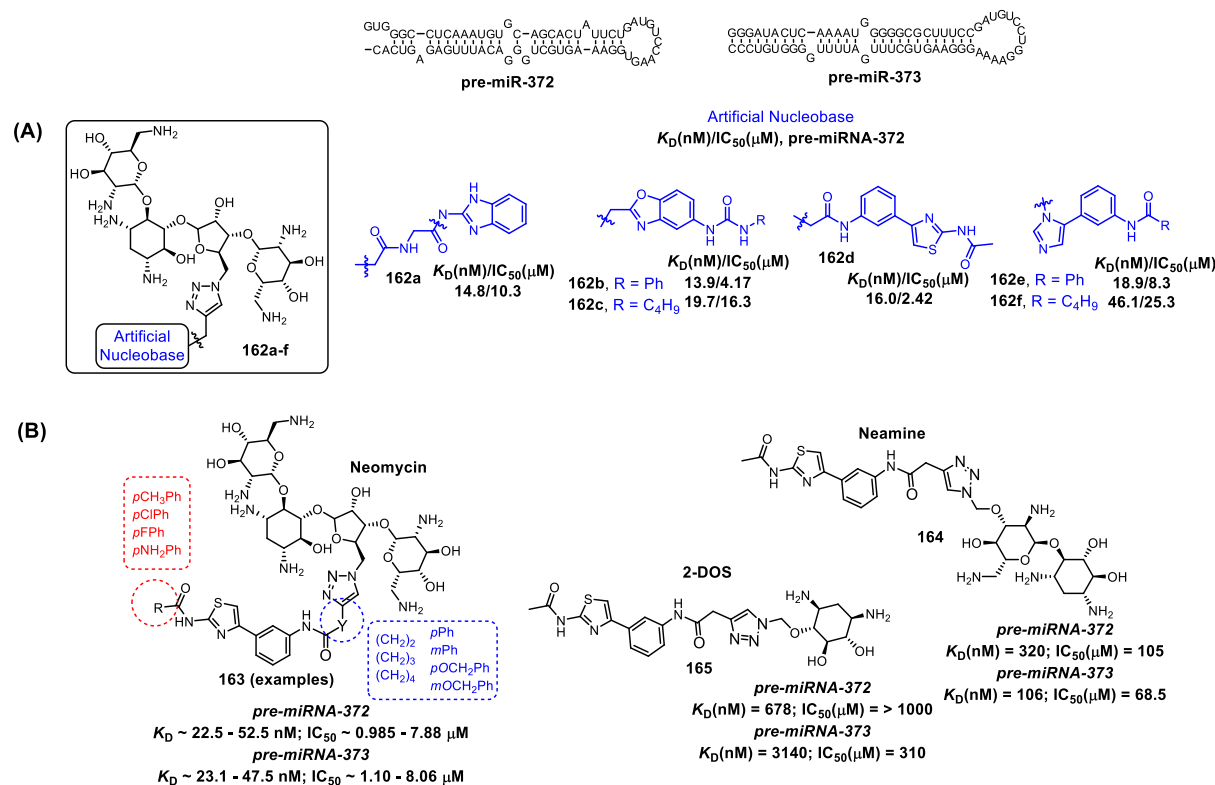


Figure 4.26. RNA ligands based on aminoglycosides-artificial nucleobase conjugates. (A) Neomycin-artificial nucleobase conjugates. (B) S-nucleobase analogues conjugated with 2-DOS, neamine and neomycin aminoglycoside.

4.2 Ligands to Target the HCV IRES III_d Loop

4.2.1 Ligand Design

In continuation of our efforts to produce innovative RNA ligands using our triplex-based rational design, we present herein the synthesis of two new artificial nucleobase scaffolds, **166** and **167** (**Figure 4.27**). These compounds were designed to feature adequate hydrogen bonding donors and acceptors to enable the molecular recognition of U:A and C:G base pairs, respectively. Moreover, the tethering of these structures to various amino acids will be also reported (**Figure 4.27**). It is worth noting that while artificial nucleobases **166** and **167** were designed to feature three hydrogen bonds to increase the specificity (recognition of U:A vs C:G), while the amino acid counterparts are critical for high affinity binding (H-bonding and electrostatic ammonium-phosphate interaction). Finally, we will disclose the binding affinity of our novel structures to a model of HCV IRES III_d domain as a potential RNA target.

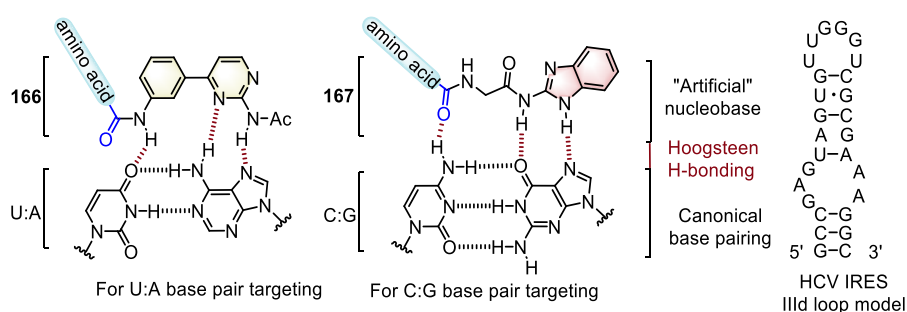


Figure 4.27. Structures of **166**, **167** and HCV IRES III_d loop model.

The design of the platforms **166** and **167** was achieved by analogy with the existing structures of artificial nucleobases known to establish Hoogsteen base pairing in triplex-forming oligonucleotides.¹⁰⁶ Thus, the molecular structure of **166** was drawn in analogy to the artificial nucleobase “S”, already reported above, where the thiazole ring is replaced by a pyrimidine moiety (**Figure 4.28**). It is worth noting that the aminopyrimidine system should account for more H-bonding ability and π -stacking interactions along the RNA helix, hence modulating the binding properties. Similarly to **166**, the structure of **167** is analogous to a benzimidazolylglycine compound **BIG** (**Figure 4.24**, Section 4.1.7.3) that was reported by Sasaki and Maeda¹¹³ and further studied by our group in aminoglycoside conjugates.¹¹¹

Notably, the BIG moiety have not been linked to amino acids. Thus, we herein complement this study and will use these data for comparative purpose. This should allow us to identify which base pair is best suited for IRES IIIId loop targeting using our multimodal ligands.

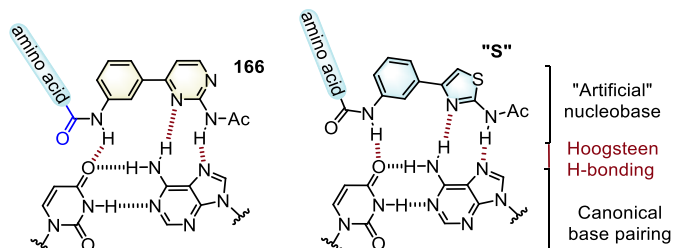
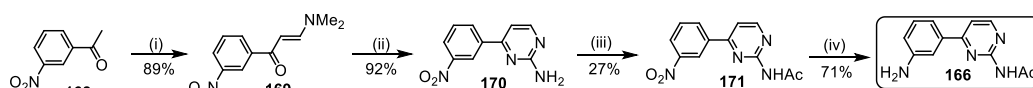


Figure 4.28. Structures of the artificial nucleobase **166** and the analogue **S**-nucleobase

4.2.2 Synthesis of the Ligands **166a-j**

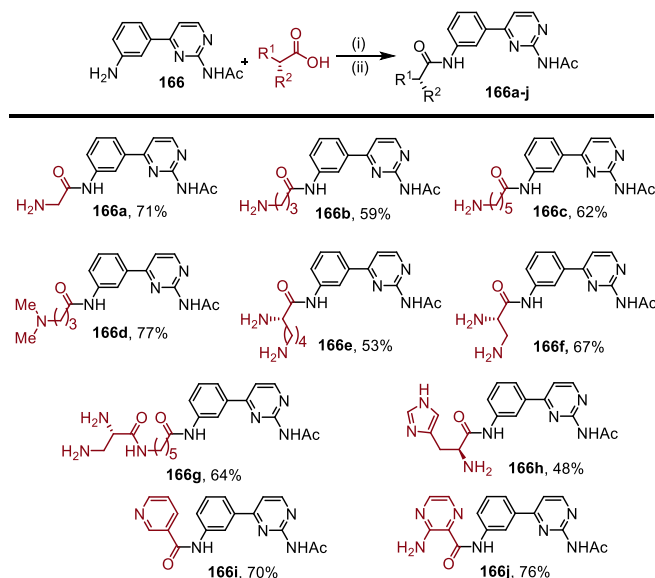
The first step of our study consists of the synthesis of platforms **166** and **167** and their derivatization with various amino acids. Thus, we first undertook the preparation of **166**, following a linear 4-step strategy (*Scheme 4.1*, steps i–iv) starting from *m*-nitroacetophenone (**168**). Firstly, **168** was homologated into the corresponding enaminone **169** using the *N,N*-dimethylamine dimethylacetal reagent. Next, the construction of the 2-aminopyrimidine ring was performed by cyclization of **170** with guanidine in *n*BuOH to afford the biaryl compound **4.5**. Then, the free amino function of **170** was acylated using acetyl chloride in anhydrous pyridine. This acylation provided **171** in moderate yields (27%) probably owing to the low nucleophilicity of the 2-aminopyrimidine moiety. Finally, compound **166** was obtained after reduction of the nitro function into an amino group that will serve as anchoring point for various amino acids tethering.



Scheme 4.1. Synthetic pathway for the preparation of **166**. Reagents and conditions. (i) *N,N*-dimethylformamide dimethylacetal, xylene, 140°C, 12 h; (ii) guanidine hydrochloride, *n*BuOH, 100°C, 24 h; (iii) AcCl, pyridine, 0-50°C, 12 h; (iv) NaBH₄, Pd/C, MeOH, r.t., 3 h.

With compound **166** in hands, we proceeded to its connection and unmasking to various amino acids through their carboxylic acid function. These structures, **166a-j**, were obtained in decent to good yields (48-77%) and are summarized on *Scheme 4.2*. The variety of connected

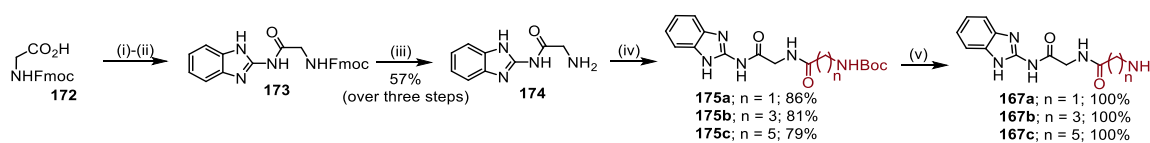
amino acids was selected to evaluate the structure-binding relationship in terms of (1) chain length between the carboxylic acid and the amine, (2) number and pK_a of amino functions, (3) presence of an α -amino group or an aromatic ring.



Scheme 4.2. Synthesis of ligands **166a-j**. Reagents and conditions: (i) *N*-Boc-protected amino acid, HBTU, HOBT, DIEA, DMF, 40°C, 24 h; (ii) TFA, CH₂Cl₂, r.t., overnight. (Yields are isolated overall yields over the 2 steps).

4.2.3 Synthesis of Ligands **167a-c**

Following the preparation of the first family of ligands, we proceeded to the synthesis of the second series (**167a-c**) designed for C:G base pair targeting (**167**). As aforementioned, this second set of compounds was prepared for comparative purpose with the first family (**166a-j**). The synthesis of **167a-c**, was performed in a 4-step sequence. Starting from the *N*-Fmoc-glycine (**172**), after the formation of its corresponding acid chloride (**Scheme 4.3**, step i), the amide bond was established with the 2-aminobenzimidazole and the *N*-Fmoc was subsequently unmasked, thus affording **174** in a 57% overall yield (3 steps). Next, the free amine was engaged in a peptide coupling with three different *N*-Boc protected amino acids to yield **175a-c**. Finally, acidic treatment prompted the *N*-Boc deprotection and furnished **167a-c** in 79-86% yields over two steps (**Scheme 4.3**).



Scheme 4.3. Synthesis of ligands **167a-c**. Reagents and conditions. (i) thionyl chloride, THF/CH₂Cl₂, 60°C, 0.5 h; (ii) 2-aminobenzimidazole, DMAP, DIEA, DMF, 100°C, 24 h; (iii) piperidine, DMF, r.t., 2 h. (iv) *N*-Boc-protected amino acid, HBTU, HOBt, DIEA, DMF, 40°C, 24 h; (v) TFA, CH₂Cl₂, r.t., overnight. Yields are isolated overall yields over the 2 steps.

4.2.4 Affinity Studies with HCV IRES III_d Loop

With these fully characterized new compounds in hands (**166a-j** and **167a-c**), we proceeded to their evaluation as ligands of the RNA model of HCV IRES III_d loop presented on **Figure 4.28**. We used for this study a classical fluorescence-based assay for quantification of RNA-ligand interactions (IRES fragment labeled at its 5'-end with a fluorescent reporter Alexa488).

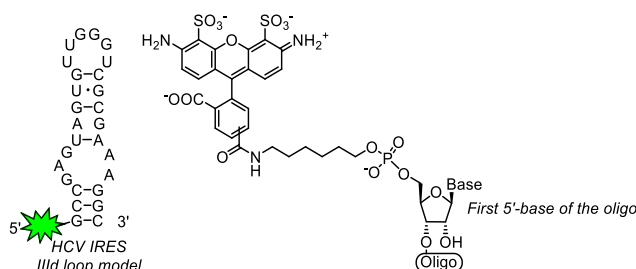


Figure 4.28. Structure of the 5'-Alexa488 modified HCV IRES III_d loop RNA model used in this binding assay.

Dissociation constant ($^{20^\circ\text{C}}K_d$) of each ligand was determined by plotting the fluorescence intensity as a function of the ligand concentrations (**Figure 4.28**). Neomycin, a well-known non-specific RNA binder was used as a positive control. Based on these results concerning compounds **166a-j** (**Table 4.1**), some structure-binding relationship (SBR) could be drawn. We observed that the nature of the linked amino acid dramatically influences the $^{20^\circ\text{C}}K_d$ with values ranging from 4.3 to 863.4 μM ; neomycin exhibiting a $^{20^\circ\text{C}}K_d$ of 8.5 μM . Interestingly, the nucleobase **166** with no appended amino acids, or the lysine amino acid alone are unable to bind IRES RNA (**Table 4.1**, entries 15 & 16), demonstrating therefore that both domains (nucleobase and amino acid) cooperate in binding to RNA. The first trend of this SBR

concerns the length of the hydrocarbon chain between carbonyl and terminal amine of the amino acid moiety. Its lengthening from one to five methylene units correlates with a 73-fold decrease in $^{20^\circ\text{C}}K_d$ (Table 4.1, entries 2-4). Additionally, switching from a primary terminal amine to a more basic tertiary dimethylamino function is also beneficial in terms of binding (Table 4.1, entry 3 vs entry 5). The number of amino group and H-bond donor/acceptor is also highly important as their increase favors the binding (entries 6-8). Nevertheless, this last point must be in balance with an appropriate chain length for an optimal effect. The use of histidine or other aromatic amino acids (entries 9-11) displayed limited efficacy. Finally, concerning the three compounds belonging to the 2-aminobenzimidazole series (167a-c), their binding was too weak to be measured. In the present study, the ligand exhibiting the highest affinity for IRES III d loop is **166e** with a $^{20^\circ\text{C}}K_d$ of 4.3 μM , twofold more potent than neomycin. While it obviously does not compare favorably with oligonucleotide ligands,^{53,65,114} this K_d value falls in the same range of other small molecule ligands of IRES IIa.^{99,103}

Table 4.1. K_d values and thermodynamic parameters for ligand/IRES III d interaction.

Entry	Ligand	Number of amines ^a	$^{20^\circ\text{C}}K_d$ (μM) ^b	ΔG° (KJ/mol)	ΔH° (KJ/mol)	TAS° (KJ/mol)
1	Neomycin	6	8.5	-28.1	-39.8 ± 4.4	11.7 ± 4.4
2	166a	1	863.4	n.a. ^c	n.a. ^c	n.a. ^c
3	166b	1	220.5	n.a. ^c	n.a. ^c	n.a. ^c
4	166c	1	11.8	-27.7	-3.1 ± 0.4	$+24.6 \pm 0.4$
5	166d	1	36.9	-25.2	-8.0 ± 1.1	$+17.2 \pm 1.1$
6	166e	2	4.3	-30.0	-25.7 ± 2.6	$+4.3 \pm 2.7$
7	166f	2	10.1	-27.8	-40.3 ± 3.1	-12.5 ± 3.2
8	166g	2	16.8	-26.5	-33.8 ± 2.4	-7.3 ± 2.4
9	166h	2	19.7	-25.9	-48.5 ± 5.1	-22.6 ± 5.2
10	166i	1	659.3	n.a. ^c	n.a. ^c	n.a. ^c
11	166j	1	232.3	n.a. ^c	n.a. ^c	n.a. ^c
12	167a	1	n.b. ^d	n.a. ^c	n.a. ^c	n.a. ^c
13	167b	1	n.b. ^d	n.a. ^c	n.a. ^c	n.a. ^c
14	167c	1	n.b. ^d	n.a. ^c	n.a. ^c	n.a. ^c
15	166	0	n.b. ^d	n.a. ^c	n.a. ^c	n.a. ^c
16	Lysine	2	n.b. ^d	n.a. ^c	n.a. ^c	n.a. ^c

^aNumber of amino functions, *i.e.* sites of protonation on the amino acid side chain. ^b K_d measurements were performed at 20°C in a 20 mM HEPES buffer (pH = 7.4) containing 20 mM of NaCl, 140 mM of KCl and 3 mM of MgCl₂ - K_d values are reported with an uncertainty of $\pm 10\%$.^{110c} n.a.: not accessible. ^dn.b.: no binding at > 30°C.

Following the determination of the K_d of each ligand for IRES IIIId at 20°C, we complemented these datasets by measuring the equilibrium constants at various temperature ranging from 5 to 36 °C; thus, we could extract the corresponding thermodynamic parameters from the Van't Hoff equation (**Table 4.1**). It appears that the Gibbs free energy variations (ΔG°) fall in the same range for all the ligands ($\Delta\Delta G^\circ_{\max} = 4.1$ kJ/mol, **Table 4.1**); although the enthalpic and entropic contributions may vary drastically. In fact, compounds exhibiting the longest flexible alkylamino side chains tend to bind their RNA target in an entropically driven fashion (*i.e.* **166c-e**). This could be due to some degrees of residual motility of the formed complex and/or desolvation contribution.¹¹⁵ Counterintuitively, **166g** binding seems to be enthalpically driven. This could be explained by the presence of two amines and one amide function at the end of **166g** side chain, which might lock the conformation of the ligand at its binding site, hence minimizing the entropic contribution. Finally, ligands displaying short side chain bearing H-bond donors/acceptors (**166f** & **166h**) feature the largest enthalpic contribution.

Having evaluated the affinity of these new ligands for IRES IIIId RNA, we selected the best ligand (**166e**) and its related congener (**166c**) to study their specificity towards the unique U:A base pair located at the stem-bulge junction. For this purpose, we used a mutated sequence of the IRES IIIId (Mut-IIIId) in which the U:A base pair is replaced by a C:G (**Figure 4.29**, **Table 4.2**).

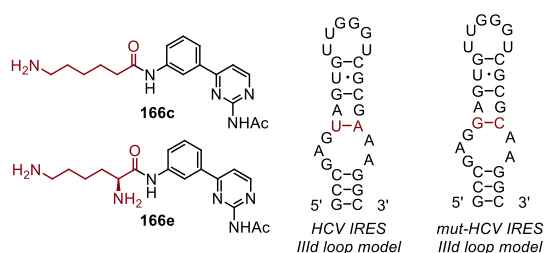


Figure 4.29. Ligands **166c** and **166e** and the structure of the wild and mutated sequences of HCV IRES IIIId loop.

Importantly, we found that the binding of **166e** to Mut-IIIId was drastically decreased with more than 100-fold compared with the wild-type sequence (^{20°C} K_d values 4.3 vs 554.4 μ M, **Table 4.2**, entry 1), proving therefore the high selectivity for U:A over C:G. The same behavior was observed using the ligand **166c** since the corresponding ^{20°C} K_d was too high to be determined (**Table 4.2**, entry 2). These pieces of evidence confirmed again that the ligands derived from **166** effectively target the U:A base pair of IRES IIIId. To further support this

result, we assessed the ligand **167c** against Mut-IIIId. While **167c** does not bind at all to the wild-type sequence, it displayed a $^{20^\circ\text{C}}K_d$ value of 41.4 μM against Mut-IIIId (Table 4.2, entry 3). Since several C:G base pairs are present in Mut-IIIId and the wild-type IRES IIIId sequences, this last observation suggests that our ligands feature some degree of site specificity and are tailored to target the stem-bulge junction where the U:A base pair is located.

Finally, we evaluated the ability of our ligands **166e** and **166c** to bind the IRES IIIId loop sequence in the presence of natural tRNA competitors (Table 4.2). To do so, we measured the $^{20^\circ\text{C}}K_d$ value of **166e** and **166c** against IRES IIIId in the presence of a 50-fold excess of a mixture of natural tRNA. This competition experiment revealed that **166e** and **166c** are able to selectively target the IIIId loop of IRES since only a marginal increase of the corresponding K_d values was observed (2.4- and 1.4-fold, respectively).

Table 4.2. K_d determination to probe the site-selectivity and the specificity of the ligands.

Entry	Ligand	$^{20^\circ\text{C}}K_d$ (μM) ^a		
		IRES IIIId	Mut-IRES IIIId	IRES IIIId + tRNA ^b
1	166c	11.8	n.b. ^c	17.1
2	166e	4.3	554.4	10.5
3	166c	n.b. ^c	41.4	-

^a K_d measurements were performed at 20°C in a 20 mM HEPES buffer (pH = 7.4) containing 20 mM of NaCl, 140 mM of KCl and 3 mM of MgCl_2 - K_d values are reported with an uncertainty of $\pm 10\%$.¹¹⁰ ^bMeasured in the presence of a 50-fold excess of a mixture of natural tRNA. ^cn.b.: no binding.

4.3 Conclusions - Chapter 4

In conclusion, we reported herein two new series of rationally design multimodal RNA ligands featuring new artificial nucleobases. These compounds proved easily accessible in only 4 to 5 steps. We presented their ability to bind to a model HCV IRES IIIId loop RNA and compared their potency with neomycin. It is noteworthy that this study discloses the first structures of small-molecule ligands of IRES IIIId loop. Low micromolar dissociation constant could be obtained with a twofold lower value for our best ligand (**166e**) compared with neomycin. Moreover, it seems that its binding is both enthalpically and entropically favored ($T\Delta S^\circ = + 4.3$ kJ/mol). This study validates our rational design strategy and could

now be used to design new series of multimodal RNA ligands and paves the way towards more potent compounds.

Importantly, the understanding of the structure-binding relationship for this class of ligands clarified important features that could be helpful to access novel ligands with higher affinities. With the knowledge concerning the contribution between both, the enthalpic and entropic components to the binding, we could envisage the design of a second generation of molecules by performing punctual modifications on the ligand structure. Heretofore, two putative modifications on the core of ligands **166c** or **166e** could be highly beneficial to strengthen the binding interaction with the target RNA (**Figure 4.30**). In fact, an increase of the rigidity of the binding motif could account for an entropic gain (enhance the entropic component of the binding) whereas the bis-alkylation of the terminal amino groups may also strengthen the binding as well, in accordance with the tendency observed for the binding constants (tertiary amines provide stronger affinities when compared with the analogue ligand bearing primary amine (**166b** vs **166d**)).

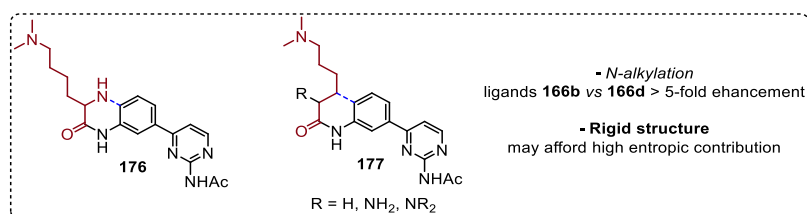


Figure 4.30. Potential modifications to achieve ligands with improved binding affinities.

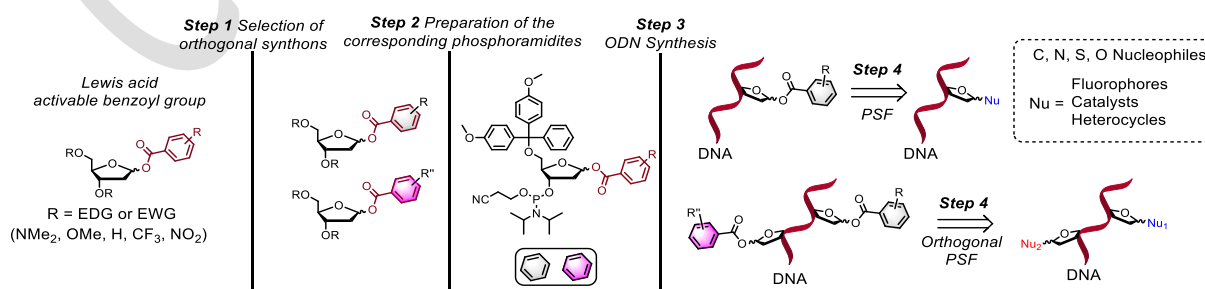
CONFIDENTIEL

Chapter 5 - Post-Synthetic Modification of Oligonucleotides

Chapter Outline

While oligonucleotides have been considered as important therapeutic agents to target DNA and RNA structures (**Chapter 4**), they have also been extensively employed in many other technological applications. Oligonucleotides have been used to study and detect mutation and structural modifications on nucleic acids. especially, the labeling of oligonucleotides with fluorescent markers have emerged as a powerful alternative to probe and study a wide variety of processes using fluorescence-based assays. Moreover, modified oligonucleotides have also found useful applications in chemistry, for instance as catalytic support (*on DNA* catalysis). Thus, it is of highest interest to develop new methodologies applied to the synthesis of modified oligonucleotides.

Two different synthetic routes are used to produce chemically modified oligonucleotides: the *ante* and post-synthetic strategies. In this chapter, we will explore the development of a new methodology applied to the post-synthetic functionalization of oligonucleotides. Essentially, we will focus on the direct functionalization on the anomeric position of deoxyribose platforms (**Scheme 5.1**). Therefore, the following discussion will highlight all the different steps of this project: (1) the screening and selection of suitable deoxyribose platforms with orthogonal reactivity on the anomeric position, to be incorporated in DNA via phosphoramidite chemistry; (2) the preparation of the phosphoramidite synthons using the identified deoxyribose scaffolds; (3) the incorporation of these platforms within oligonucleotide strands using an automated solid-phase synthesis; (4) the set on of a post-synthetic functionalization process.



Scheme 5.1. Schematic representation of our methodology of PSF of oligonucleotides.

Hereafter, prior to the presentation of our results, a brief overview concerning the synthesis of oligonucleotides via phosphoramidite chemistry and the reported methodologies available to post-synthetically functionalize nucleic acids we be provided.

5.1 State of the Art

5.1.1 Chemical Synthesis of Oligonucleotides - Phosphoramidite Chemistry

The earliest methods described for the synthesis of oligonucleotides date from 1950s and were based on solution-phase synthesis using H-phosphonate and phosphotriester chemistry (Michaelson and Todd^{116,117}) and the Khorana's phosphodiester approach.¹¹⁸ Later, in 1960s and 1970s, Letsinger's group¹¹⁹ developed the phosphotriester method. The main progress of this method, compared with the phosphodiester approach, comes with the protection of the phosphate moiety with 2-cyanoethyl group. It substantially reduces side reactions on the free oxygen on the internucleosidic phosphate. This method also implemented the use of more efficient coupling agents, hence reducing dramatically the length of the polynucleotide synthesis.

In 1970s, the use of the more reactive phosphorus (III) derivatives started to be investigated. In 1975, Letsinger reported the phosphite triester-based method using 3'-*O*-chlorophosphites synthons for the formation of internucleosidic linkages.¹²⁰ Six years later, in the early 1980s, based on phosphitetriester chemistry, Caruthers' group implemented for the first time the solid-phase synthesis of oligonucleotides.¹²¹ Finally, in the same year, the same research group described the phosphoramidite synthons, a prospective class of key intermediates to produce oligodeoxynucleotides.¹²² These new and reactive intermediates greatly tuned the use of the solid-phase technology through the automated nucleic acid synthesis. Moreover, the use 2-cyanoethyl *N,N*-diisopropylphosphoramidites was established as the regular synthon used in the automated phosphoramidite elongation chemistry.

The nucleoside phosphoramidites are sufficiently stable to be stored at low temperatures (-18°C), under an inert atmosphere for several months and can be prepared in multigram scale. Their easy preparations and their high reactivity bolstered their use as reference commodity reagents for the solid phase synthesis of nucleic acids. They only become unstable and reactive upon diisopropylamine protonation. In the **Figure 5.1** are shown the main phosphoramidite synthons used for regular DNA synthesis. Of note, both, the nucleobase and

the phosphoramidites bear base labile protecting groups (acyl and cyanoethyl groups, respectively). This combination facilitates the final deprotection of the nucleic acid in a single step. Finally, to prevent undesired side reactions during the oligonucleotide elongation, the 5'-position of the (deoxy)ribose backbone may be protected. However, differently from the base labile protecting groups used on the phosphoramidite and the nucleobases, the 5'-OH are protected with an orthogonal and acid labile 4,4'-dimethoxytrityl (DMTr) group.

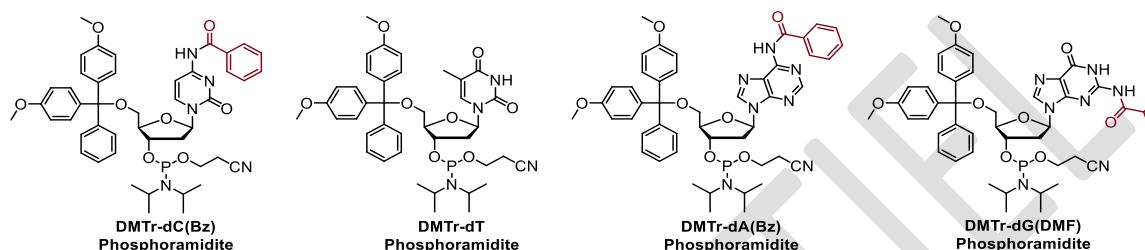
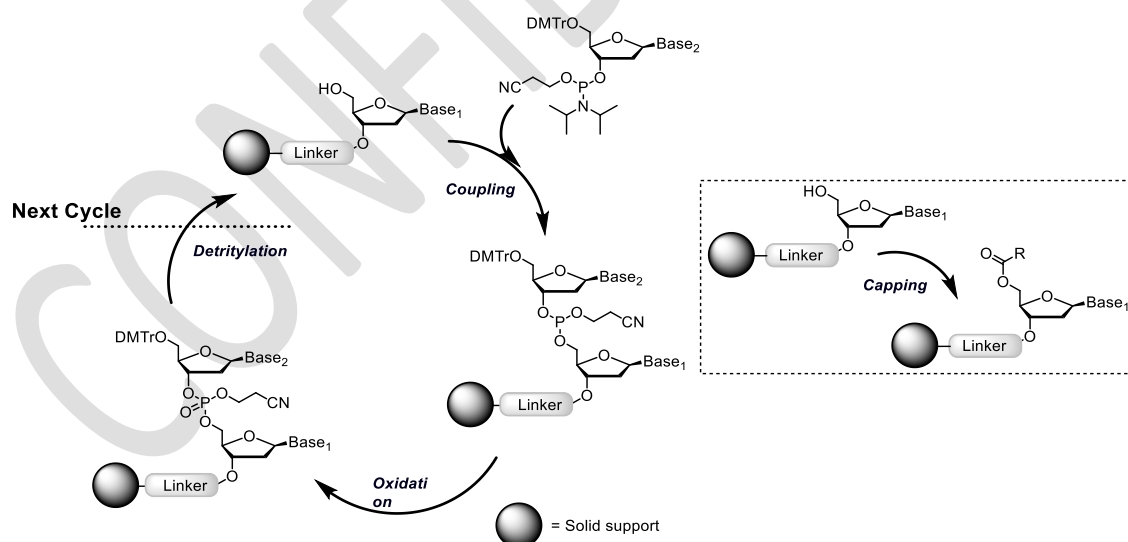


Figure 5.1. Examples of regular and commercially available phosphoramidite synthons.

Differently from the DNA biosynthesis, the phosphoramidite chemistry proceeds in the 3'- to 5'- direction (*Scheme 5.2*). The nucleic acid chain is grown step by step by incorporating one nucleotide per cycle iteration. The cycle of synthesis is presented below and is composed of four main steps: detritylation, coupling, capping and oxidation.



Scheme 5.2. Solid-phase oligonucleotide synthesis using the phosphoramidite method.

5.1.1.1 Detritylation

At the beginning of DNA synthesis, the first protected nucleoside is already bounded to the solid-support by its 3'-hydroxy function. The first step is the removal of the 5'-ODMT_r protecting group of the solid-support-linked nucleoside. The trityl deprotection occurs relatively fast in presence of trichloroacetic acid (TCA, 3% wt. in dichloromethane). The main advantage of DMTr protecting group are related to the formation of the 4,4'-dimethoxytritylium cations, which presents a strong orange color that is used to monitor coupling efficiency after each nucleotide incorporation. After DMTr removal, the supported nucleoside features a 5'-OH group ready for coupling with next residue.

5.1.1.2 Activation and Coupling

The naked hydroxyl group on the 5'- position of the supported (oligo)nucleotide is free to react with the incoming phosphoramidite monomer. Prior to the coupling, the phosphoramidite nucleoside is mixed with a weak acid (usually 1H-tetrazole derivative), which promotes the protonation of the diisopropylamino moiety followed by its substitution with the tetrazole to form an activated adduct. Such intermediate then reacts with the hydroxy group on the 5'- position of the growing oligonucleotide chain to form a phosphite triester linkage. Considering the coupling of the regular nucleoside phosphoramidite, the yield is generally higher than 98% under standard conditions. For modified phosphoramidite synthons produced in the laboratory, the coupling step needs to be optimized and perhaps needs longer coupling times or multiple coupling cycles to achieve high conversion yields (> 98%).

5.1.1.3 Capping

High coupling yields (> 99.5 %) are normally expected in each cycle. However, even using the most optimized chemistry and fresh reagents, it is still not possible to obtain quantitative incorporation of the incoming phosphoramidite on the growing nucleic acid. The unreacted 5'-hydroxyl groups thus need to be masked to inhibit undesirable reactions during the following cycles. Thereby, a capping step is performed to avoid deletion mutations, *i.e.*, the synthesis of oligonucleotides missing one or multiple nucleotides, which can result in a complex mixture of "shortmers" (n-1, n-2, n-3, etc.). Capping is normally performed using a mixture of two different solutions: an anhydride (acetic, Ac₂O or phenoxyacetic, PAC₂O anhydrides) and *N*-methylimidazole (NMI), both as THF solutions containing up to 10% of

pyridine, to ensure that the pH remains basic (prevent detritylation). Thus the capping step consists in the acylation of the unreacted 5'-OH groups to yield their unreactive carboxylic ester counterparts.

5.1.1.4 Oxydation

The last step corresponds to the oxydation of the phosphitetriester internucleosidic linkage into the most stable phosphate triester. Iodine, in presence of water and pyridine is conventionally used for the oxydation. This reaction is very fast and is finished in few seconds (about 10 to 15s in standard procedures). The resultant phosphate triester is, in fact, the DNA backbone containing the 2-cyanoethyl protecting group.

5.1.1.5 Support Cleavage and Deprotection

Once the oligonucleotide synthesis is complete, the phosphates linkages are still protected with the 2-cyanoethyl groups while the nucleobases are protected with basic labile groups (dmf, Bz and Ac). The universal support usually uses base-labile ester linkages (frequently succinyl ester) to connect the 3'- terminal hydroxyl group of the first nucleoside. The support cleavage is normally carried out in mild conditions, in presence of concentrated aqueous ammonia solution at room temperature.

The released oligonucleotide can be heated to 50-60°C during a couple of hours (or kept at room temperature for a longer time, *ca.* 24 h) to afford the complete deprotection of the nucleobases and the phosphates, to access the expected oligonucleotide. Importantly, prior to the ammonia treatment, the oligonucleotide can be obtained as its 5'-OH or 5'-ODMTr derivative depending on the application of the final detritylation step. In the later case, the presence of the DMTr group could ease the HPLC purification of the oligonucleotide, however, the detritylation step will need to be performed (in solution).

5.1.2 Post-Synthetic Modification of Oligonucleotides

Oligonucleotides can be synthesized by two different synthetic routes: *ante*- and *post*-synthetic approaches (**Figure 5.2**). In an *ante*-synthetic fashion, the desired functionalities must be introduced directly on the phosphoramidite synthon prior to the nucleic acid synthesis. It is the main approach when a specific modification is desired, no matter how

many times it would be incorporated or in which oligonucleotide sequence. However, when the incorporation of multiple functionalities/modifications are the main goal, an *ante*-synthetic route will require the tedious preparation of all the different phosphoramidite synthons, in parallel. In this case, post-synthetic approaches can clearly overcome this issue. Basically, in a post-synthetic fashion, the aim is to introduce a modified nucleotide containing a reactive group on the desired oligonucleotidic sequence. Once the synthesis of the precursor nucleic acid is finished, it can be modified with multiple functionalities by using a specific chemical transformation. Thus, a single phosphoramidite synthon is required and the multistep and time-consuming synthetic route to access a diversity of modified phosphoramidites can be avoided. In addition, post-synthetic functionalization also allows the implementation of mild conditions to introduce sensible functionalizations (*i.e.* groups unstable under the oligonucleotide synthesis) or to perform conjugation between nucleic acids and other biomolecules (*i.e.* nucleic acid-protein conjugation).^{123,124}

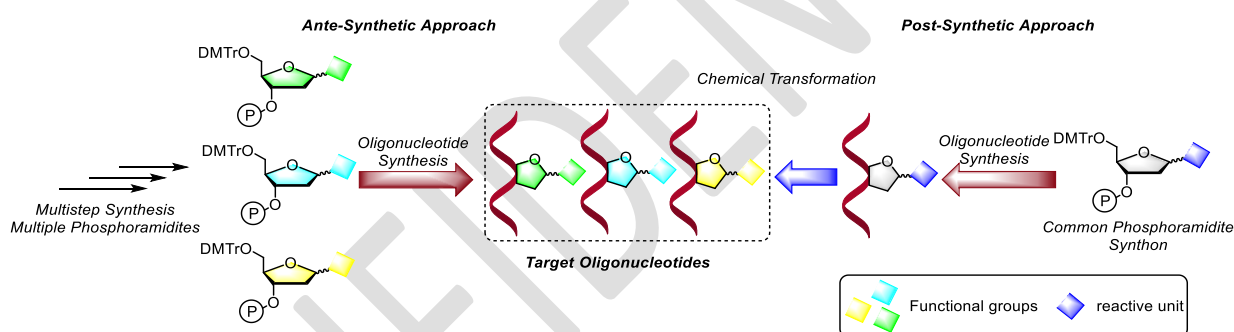


Figure 5.2. *Ante*-synthetic vs post-synthetic approaches.

The post-synthetic transformations can be performed in two different ways: prior to the release from the support (*on support*) or afterwards (*in solution*) (**Figure 5.3**). The choice of one or another approach should be carefully selected. When performed *on support*, the final functionalities need to be stable under the conditions required for the final deprotection/release step. Intuitively, if the precursor oligonucleotide ready for post-synthetic functionalization is unstable under the final deprotection condition, the *on support* approach remains as the unique alternative. When performed *in solution*, the scope of chemical transformation can be increased as more sensitive functionalities may be introduced since no further deprotection is not required.

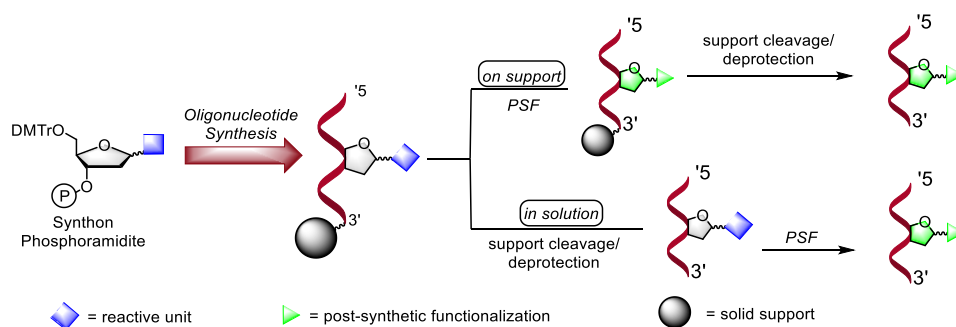


Figure 5.3. On support vs in solution PSF of oligonucleotides.

Heretofore, the major reported methodologies explored in the context of post-synthetic modification of nucleic acids are based on click reactions such as copper catalyzed azide-alkyne cycloaddition,^{125–131} thiol-ene chemistry,^{123,132–134} Diels-Alder/cycloaddition,¹³⁵ oxime ligation,¹³⁶ amination,¹³⁷ reductive amination,¹²⁹ amidation,¹²⁷ photoclick processes,¹³⁸ as well as photo-cross-linking.¹³⁹ Nevertheless, these post-synthetic transformations also encompass palladium catalysed reactions such as Suzuki, Sonogashira and Stille cross-coupling reactions.¹⁴⁰ To date, despite this wide compilation of methodologies, only few approaches to introduce two or more functionalities using orthogonal chemical transformations have been reported.^{127,129,134,141}

In addition, the most common positions explored to perform such kind of modifications are on the nucleobases, the internucleosidic phosphates, the 5'-terminal hydroxy group or on the sugar skeleton, whereas researches focusing on the anomeric position have not been reported yet (**Figure 5.4**). Thus, it still promising to focus on the post-synthetic functionalization of nucleic acid on the anomeric positions. Prior to the presentation of our results, a brief overview to state the post-synthetic functionalization of nucleic acids on different nucleosidic motifs will be exposed. Selected examples of single, multiple and orthogonal transformations will be surveyed.

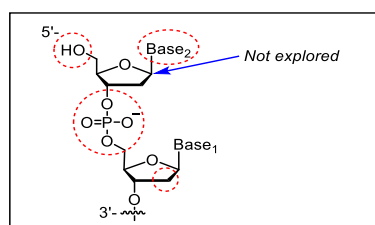


Figure 5.4. Most common positions explored for the PSF of oligonucleotides.

5.1.2.1 Nucleobase Functionalization

Post-synthetic functionalizations exploiting modified nucleobases have been the major approach to modify oligonucleotides. For instance, Panattoni and co-workers¹²⁶ reported the use of ONs containing a 5-modified thymidine nucleotide bearing flexible hydrophilic and hydrophobic alkyne linkers to perform post-synthetic conjugation of DNA *via* CuAAC reactions. Both, the triphosphate based polymerase synthesis and the phosphoramidite chemistry were employed for the synthesis of the corresponding DNA sequences. Fast and effective cycloaddition reactions were reported. Moreover, the use of hydrophilic peglated linker proved to increase twice the reaction rate compared with the octadiynyl-linked DNA sequences (**Figure 5.5**).

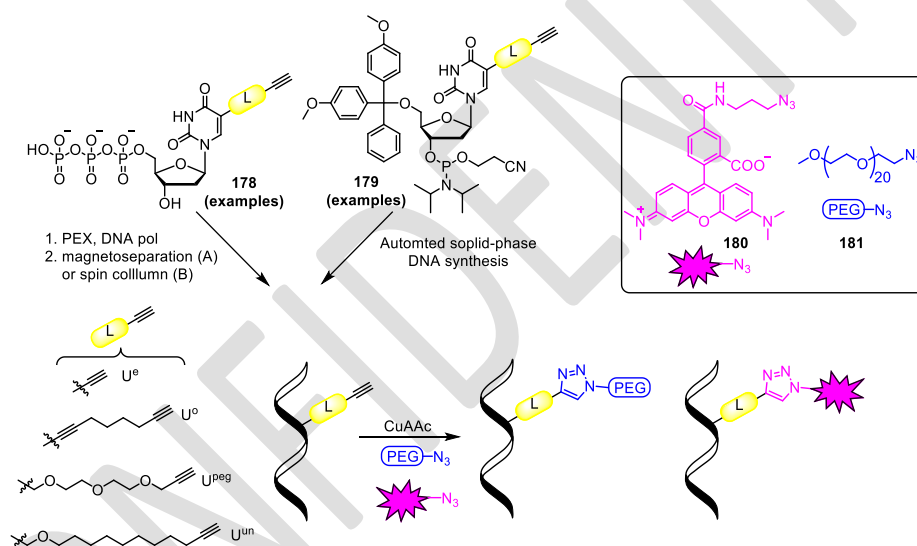


Figure 5.5. PSF of ODNs using CuAAC cycloaddition reactions on alkynyl modified nucleobase.

The use of photoclick reactions have been reported as well. Arndt and co-workers¹³⁸ recently demonstrated the use of a push-pull substituted diaryltetrazole nucleotides as a prospective tool for PSF of NAs. The connected tetrazole dye on the 5 position of a deoxyuridine nucleotide was used to conjugate a cyanine (Sulfo-Cy3, **182**) fluorophore *via* a photoactivable tetrazole-maleimide cyclization, activated by LED irradiation at 365 nm (**Figure 5.6**). Improved kinetic effect was also reported for this type of photoclick conjugation when compared with classical CuAAC reactions.

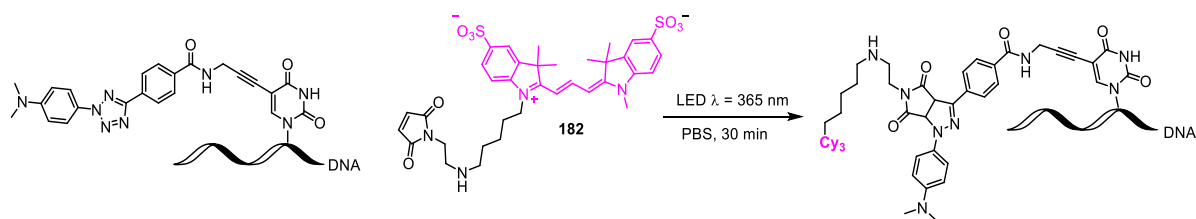


Figure 5.6. PSF of oligodeoxynucleotides via photoclick reaction.

In a different work, using the simple and commercially available 5-iododeoxyuridine modified oligonucleotides, Lercher et al.¹⁴⁰ described the use of palladium catalyzed Suzuki–Miyaura cross-coupling reaction under mild and aqueous conditions (TRIS buffer, 37.8°C) using unprotected nucleotides. The straightforward methodology was applied to a diversity of boronic esters (**183–186**, **Figure 5.7**) containing different functionalities, enabling the synthesis of modified ONs, with reduced formation of the corresponding deiodinated byproduct (**Figure 5.7**).

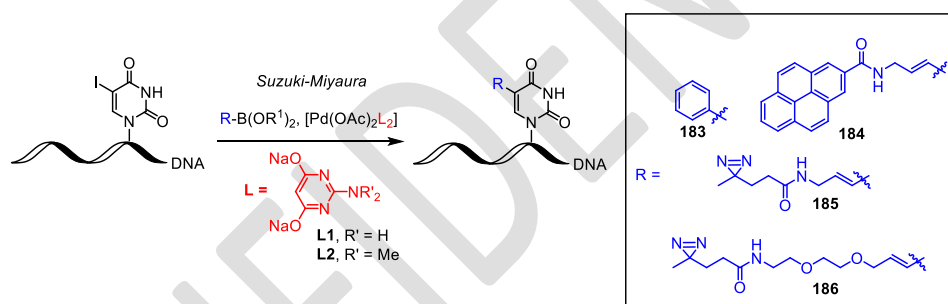


Figure 5.7. PSF of ODNs using Suzuki-Miyaura cross coupling reactions.

Based on 1-4 Michael's addition, Dadov and co-workers¹³² reported the use of two types of modified cytidine nucleotides bearing acrylamide or vinylsulfonamide as reactive sites for direct conjugation of thiols containing amino acids to DNA (**Figure 5.8**). They demonstrate that the vinylsulfonamide containing ONs are able to readily react with thiols under physiological conditions. This high reactivity was successfully applied to the cross-linking of DNA with several constructs of p53 protein. Meanwhile, the high compatibility of the vinylsulfonamide modified cytidine triphosphates with the polymerase-catalyzed primer extension (PEX) synthesis also makes this platform suitable to design DNA methyltransferase inhibitors (*e.g.* DNA-modifying proteins bearing free cysteine amino acids).

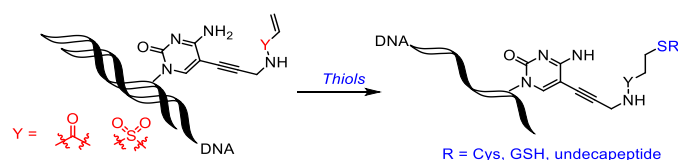


Figure 5.8. 1,4-thiol-ene-like Michael reactions applied to the PSF of oligodeoxynucleotides.

PSF *via* Huisgen 1,3-dipolar cycloaddition reactions using alkenes and nitrile oxides have also been reported. Essentially, Gutmiedl¹³⁶ described the synthesis and DNA incorporation of a deoxynucleotide platform containing a terminal norbornene scaffold as an active site for post-synthetic labeling. The *on DNA* cycloaddition reactions performed *on support* or *in solution* afforded mainly *exo*-2-isoxazolines products (**Figure 5.9**). The authors also highlighted that only 10 equivalents of the corresponding nitrile oxide are required, supporting a suitable transformation for the labeling of NAs with high cost markers. Moreover, the slow rate of conversion between the nitrile oxide and alkynes, when subjected to the same reaction conditions, also enable to envisage the bioorthogonal transformation using CuAAC and the norbornene/nitrile oxide cycloaddition reactions.

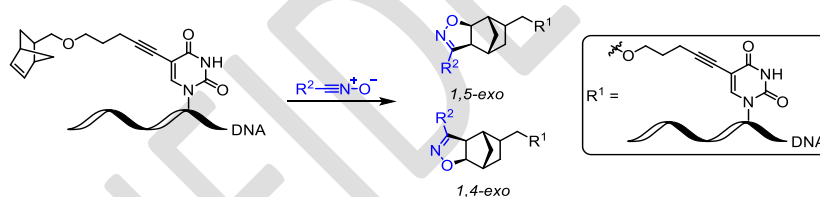


Figure 5.9. PSF using Huisgen 1,3-dipolar cycloaddition reaction on ODNs bearing 5-modified deoxyuridine nucleobases.

5.1.2.2 Phosphate Functionalization

The PSF on the phosphate backbone have also been investigated. Several fluorescent labeled oligonucleotides used in biophysical assays have been synthesized using phosphate modifications, especially on their 5'- or 3'-ends, by using modified phosphoramidite linkers (**Figure 5.10**) (**Chapter 4**, labeled HCV IRES IIIId loop).

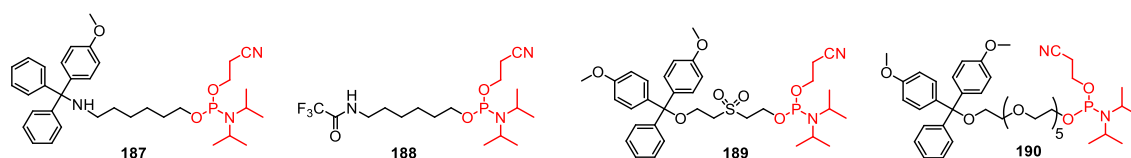


Figure 5.10. Common phosphoramidite linkers.

Among the various methods of PSF at the internucleosidic phosphate linkages, Shang and co-workers¹³⁰ reported an interesting approach based on the use of alkynyl-phosphoramidite nucleosides. They described an efficient approach to produce amino-acid and peptide conjugate oligodeoxynucleotides (ODNs) with improved stability and cell uptake. The direct modification on the phosphate backbone of regular nucleotides using a general azido-alkyne cycloaddition enable the assessment of modified triazolylphosphonate analogs (TP ODNs) at any pre-selected internucleotidic linkage (**Figure 5.11**). The reported TP ODNs revealed a remarkable improvement in their physicochemical properties when compared with natural phosphodiester DNAs by changing the net negative charge, thus enhancing the nucleic acid stability towards nuclease degradation. Moreover, this non-charged triazolylphosphonate also enhanced the cell membrane permeability, thus dispensing the use of artificial carriers.

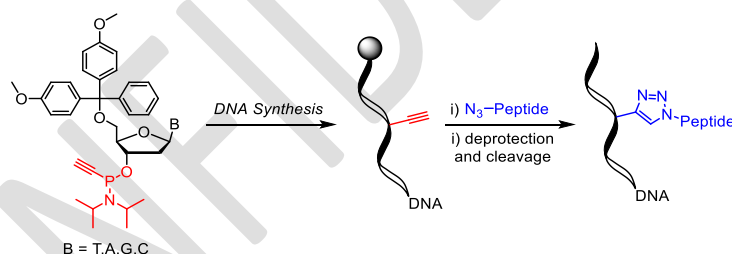


Figure 5.11. PSF on the internucleosidic phosphonates using CuAAC cycloaddition transformations.

Meyer et al.¹³⁴ described the synthesis of modified oligonucleotides using a thiol-ene Michael-type addition methodology (TMTAs). Precursor oligonucleotides bearing 5'-terminal thioacetatehexyl phosphate modifications were synthesized by using phosphoramidites linkers. These S-containing oligonucleotides were firstly subjected to a set of PSF using acrylamide derivatives in a TMTA transformation, under basic conditions (**Figure 5.12**). The authors also demonstrate a linear two-steps synthetic route to introduce both, a thioacetate and an alkyne motif to perform orthogonal PSF with multiple labels using the phosphoramidite linker **194**. Moreover, the deprotection and release of the oligonucleotides from the solid support occur under the same reaction conditions as the one used for the thiol-

ene addition. Thus, final thiol-ene modification and oligonucleotide recovery can be performed in a unique step.

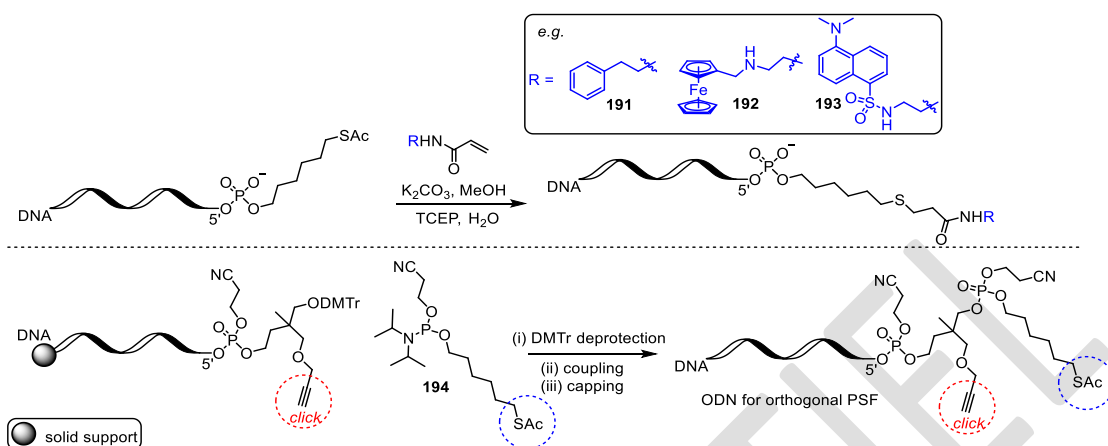


Figure 5.12. Thiol-ene based PSF of oligonucleotides on the 5'-terminal phosphate backbone.

5.1.2.3 Ribose Functionalization

The design of modified (deoxy)ribose platforms displaying a reactive group for PSF of NAs have also been reported. The functionalization on the 2' and 4'-positions are currently the most studied while a single work has been reported exploiting the anomeric position (1'-position).¹²⁵ The most common approach focused on the use of click processes based on CuAAC reactions. For example, Berndl and co-workers¹²⁸ reported the bioconjugation of phenoxazinium and coumarin fluorophores (**195** and **196**) on two types of alkynyl-modified oligonucleotides containing either a non-nucleosidic linker or an uridine nucleotide modified at its 2'-position (**Figure 5.13**).

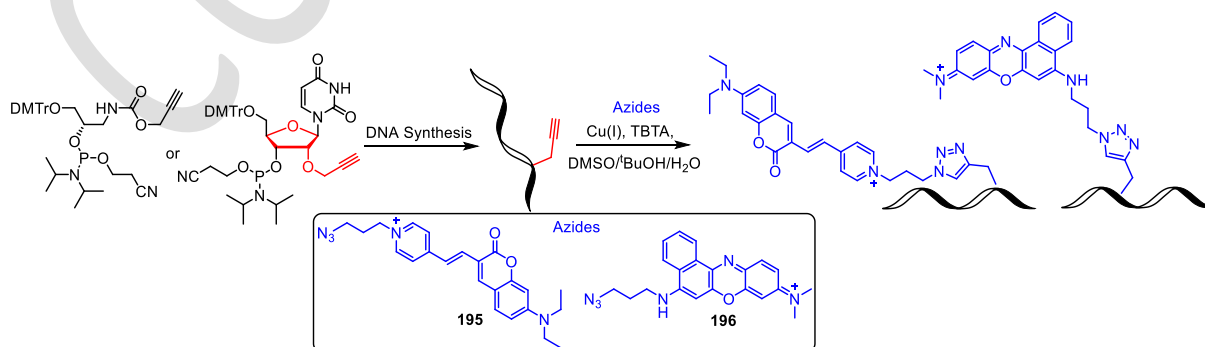


Figure 5.13. PSF of ODNs bearing a 2'-O propargyl substituted uridine nucleotide using CuAAC cycloaddition.

Using the same chemical approach, Hari¹⁴² and Obika¹⁴³ pioneered the use of 1-deoxy-1-ethynyl- β -D-ribofuranose platform to performed PSF through CuAAC reactions. Then Kitamura and co-workers¹²⁵ took advantage on this methods to produce oligonucleotide analogues using azidomethyladenine and azidomethylthymine (**197** and **198**, **Figure 5.14**).

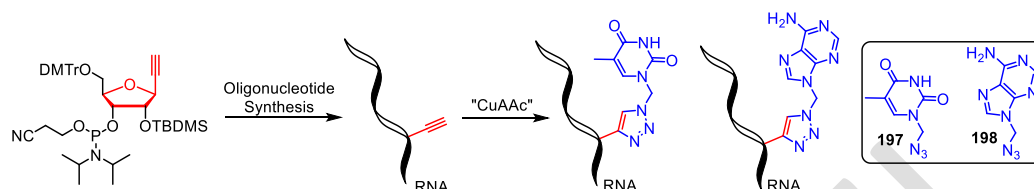


Figure 5.14. PSF of RNA oligomers bearing the 1-alkynyl modified ribose platform via CuAAC cycloaddition.

5.1.2.4 Orthogonal PSF

Several groups developed methods to perform sequential and selective post synthetic functionalization of oligonucleotides. It appears that all the reported methods for an orthogonal PSF of nucleic acids are based on two approaches: The first one is based on chemoselectivity; *i.e.* two different reactive functions displaying orthogonal reactivity are used for post-synthetic functionalization purposes. The second approach uses a single function that can be selectively converted into different products using adapted reaction conditions.

Exemple of the first approach: Gutmiedl and co-workers¹⁴¹ described the synthesis of oligodeoxynucleotides and long PCR amplicons bearing modified nucleobases carrying styrene and alkyne motifs. Sequential 1,3-Dipolar cycloaddition reactions were performed enabling the orthogonal PSF of DNA (**Figure 5.15**).

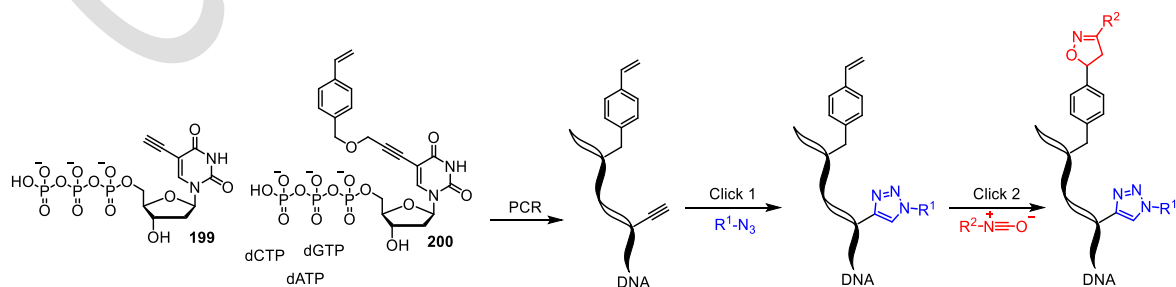


Figure 5.15. Orthogonal PSF of ODNs bearing styrene and alkyne motifs *via* 1,3-Dipolar cycloaddition reactions.

The fast reaction between the styrene scaffolds and nitrile oxides enable the easy functionalization with 3,5-disubstituted isoxazolines in a copper free approach, while triazolyl modifications could be synthesized in the sequence via classical CuAAC reactions either *on support* or *in solution*.

Another example of this first approach: A mild and orthogonal protocol based on CuAAC cycloaddition and amidation reactions was developed for PSF using 2'-O-methyl ester and 2'-O-propargyl nucleoside scaffolds to access modified siRNAs. The method was successfully used to build up oligonucleotides with structural diversity, by installing up to five pentylamines (**201**) and five azido *N*-acetylgalactosamine (**202**) derivatives along the 2'-positions of the oligonucleotides (**Figure 5.16**).¹²⁷

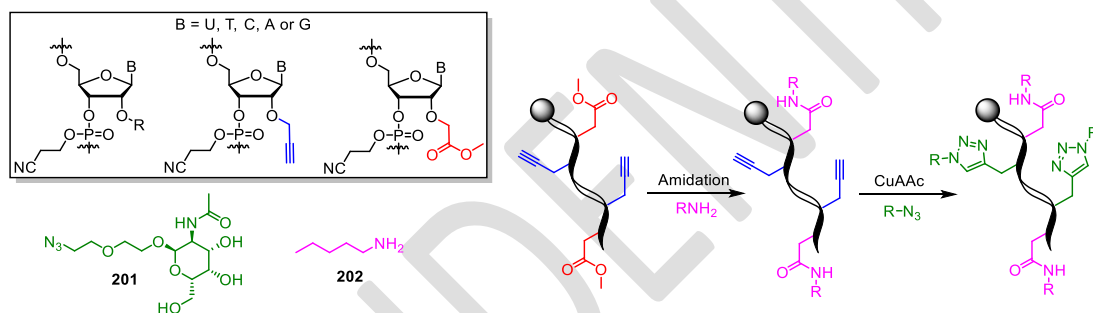


Figure 5.16. Orthogonal PSF of ODN using CuAAC cycloaddition and amidation reactions.

In a different work, based in the second approach, Rao and collaborators¹²⁹ described a protocol for PSF of ONs *via* either CuAAC or Staudinger reactions (**Figure 5.17**). Using a simple azido group, modified oligonucleotides displaying substituted triazoles or free amino groups could be easily produced using simple reaction conditions.

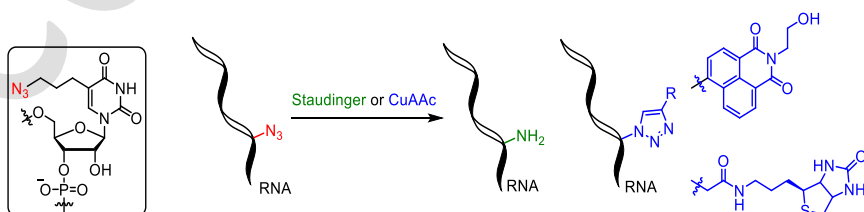


Figure 5.17. Orthogonal PSF of ONs using CuAAC cycloaddition or Staudinger reactions.

5.2 Post-Synthetic Functionalization at the Anomeric Position - An Innovative Approach

In complementation of the large variety of methods already described for the post-synthetic modifications of nucleic acids, we initiated the development of an innovative methodology to perform direct functionalization of oligonucleotides on their anomeric position. To our knowledge, the chemical transformation directly on the 1'-position using a post-synthetic approach has not been reported in the literature. Thus, two main objectives will be explored here:

- I) the development and validation of a robust methodology allowing for the PSF of deoxyribose units at their anomeric position.;
- II) The search for different deoxyribose platforms of anomeric orthogonal reactivity. Thus different modifications could be introduced on a single strand in a sequential and selective fashion.

In this project, we decided to explore a Lewis acid catalyzed nucleophilic substitution on the anomeric position of the deoxyribose moiety (**Figure 5.18**) as a putative glycosidation reaction. This general concept will be firstly explored by studying a family of different deoxyribose platforms harbouring variously substituted benzoyl aglycones. In fact, the presence of electron donating or withdrawing substituents on the phenyl core of the Bz group should allow the modulation of the nucleofuge ability of the aglycone moiety. Here, we expect that an orthogonal activation of the electronically tuned benzoyl groups could be achieved using the suitable Lewis acids. Essentially, benzoyl groups bearing electron withdrawing groups could be potentially activated by using weaker Lewis acids when compared with those groups displaying EDG.

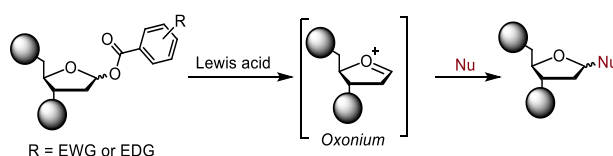
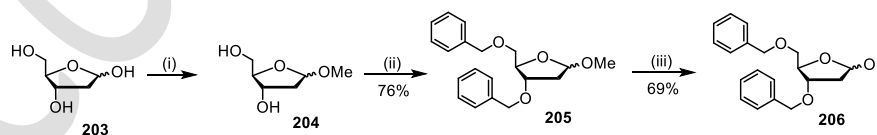


Figure 5.18. Glycosidation reactions adapted to the PSF of oligonucleotides.

5.2.1 Aglycone Moiety Screening

5.2.1.1 Synthesis of model deoxyribose platforms bearing 1'-O-benzoyl groups at the anomeric position

To assess the feasibility of the proposed PSF methodology, we first investigated what kind of benzoyl aglycone would be suited for our purpose, *i.e.* which 1'-O-benzoyl groups could be incorporated into DNA sequences (compatible with phosphoramidite chemistry) and could be efficiently activated under mild conditions. For this purpose, deoxyribose derivatives containing both 5'- and 3'-OH groups protected as benzyl ethers were synthesized and modified in their 1'-position with benzoyl derivatives **207a-207g** in a linear four steps synthetic route, starting from D-2-deoxyribose (**Schemes 5.3** and **5.4**). Firstly, the D-2-deoxyribose (**203**) was treated with methanol, in presence of catalytic amount of acetyl chloride to produce the corresponding acetal **204** in quantitative yield. The reaction, known as Fisher glycosidation, was kept no more than 60 min to avoid the isomerization and formation of the more stable 6-membered pyranoside analogue. Then, the intermediate **5.2** was refluxed overnight in THF in the presence of potassium hydroxide and an excess of benzyl chloride. This protection step yielded the compound **205** in 76% yield after purification by column chromatography. Finally, the acetal **205** was dissolved in an acetic acid/water (4/1) mixture and heated under 60 °C during 48 h to afford the key intermediate **206** as a yellowish oil after purification by chromatography in 69% yield. This intermediate **206** is suitably protected on its 3' and 5' positions and available for derivatization at its anomeric position.

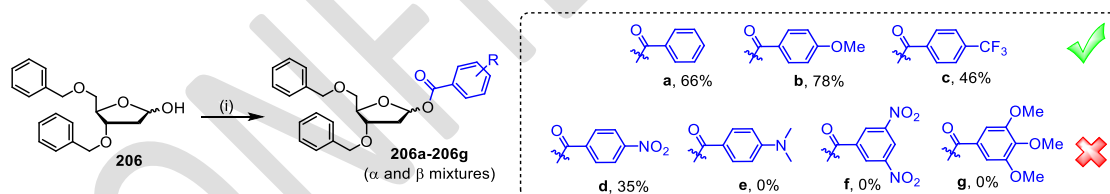


Scheme 5.3. Synthesis of deoxyribose intermediate **206**. Reagents and conditions. (i) MeOH, AcCl, r.t., 45 min. (ii) BnCl (7 equiv.), KOH (4.5 equiv.), THF, reflux. (iii) AcOH/H₂O (8/2), 60°C, 48 h.

With compound **206** in hands we could proceed to the preparation of a small family of glycoside derivatives harbouring various benzoyl-based aglycones. For this purpose **206** was solubilized in dry pyridine and treated with 1.5 equivalents of the corresponding benzoyl

chlorides, at room temperature, until complete consumption of **206** (TLC monitoring) (*Scheme 5.4*). In this diversification process, seven different benzoyl chlorides were used. They encompass phenyl cycle non-substitution (**207a**) or substitution with *p*-OMe (**207b**), *p*-CF₃ (**207c**), *p*-NO₂ (**207d**), *p*-NMe₂ (**207e**), 3,5-(NO₂)₂ (**207f**) and 3,4,5-(OMe)₃ (**207g**).

Compounds **207a**, **207b** and **207c** exhibiting a plain benzoyl core or a *p*-OMe and *p*-CF₃ substitution pattern could be readily accessed in 66, 78 and 46% yield, respectively. Unexpectedly, when benzoyl groups containing strong electron donating substituents were used (**207e** and **207g**) the corresponding glycoside derivative could not be obtained (**207e** or proved unstable upon storage (**207g**, decomposes in few hours). Conversely, using the 3,5-dinitrobenzoyl chloride, the corresponding glycoside **207f** is readily formed (< 1 h, TLC monitoring), however, the product was too sensitive to silica gel to be chromatographed (even when triethylamine is added to the eluent). Therefore, **207f** could not be isolated. Finally, concerning the highest reactivity (lower stability) of the nitro derivatives in acidic media, we also evaluated the stability of the derivative **207d** under the detritylation condition applied to nucleic acid synthesis (3% wt. TCA in dichloromethane). In this case, complete degradation of the material was observed in less than 5 min, demonstrating the non-compatibility of the nitro-decorated benzoyl glycoside derivative with the phosphoramidite elongation chemistry of nucleic acids.



Scheme 5.4. Synthesis of benzoyl derivatives **207a-207g**. Reagents and conditions. (i) corresponding benzoyl chloride (1.5 equiv.), pyridine, r.t.

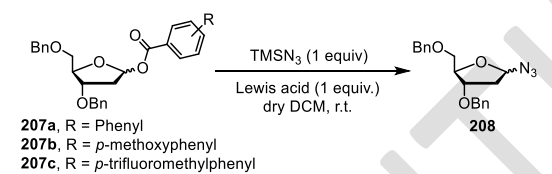
Altogether, we produced 3 different benzoyl glycosides having aglycones of putatively different nucleofuge properties.

5.2.1.2 Lewis Acids Screening

With the three derivatives in hands (**207a-207c**), we proceeded to the screening of different reaction conditions, aiming to access: i) suitable conditions to perform an efficient glycosidation reaction using these glycosyl donors; ii) possible orthogonal activation protocol

of this three different platforms (use of different Lewis acids). For this preliminary screening, mixtures of the α and β anomers of **207a-207c** were treated with different Lewis acids in the presence of azidotrimethylsilane (TMSN_3) as the model nucleophile. This benchmark reaction was performed in dry dichloromethane, using an equimolar concentration of the glycosyl donor, TMSN_3 and Lewis acid, at room temperature, during 1 h. The reaction outcome was first assessed by TLC monitoring (**Table 5.1**).

Table 5.1. Screening of several Lewis acid activators for the benchmark reaction between **207a-207c** and TMSN_3 .



Entry	Lewis Acid	Product (Y/N)*	Entry	Lewis Acid	Product (Y/N)*	Entry	Lewis Acid	Product (Y/N)*
207a			207b			207c		
1	$\text{BF}_3 \cdot \text{Et}_2\text{O}$	Y	12	$\text{BF}_3 \cdot \text{Et}_2\text{O}$	Y	23	$\text{BF}_3 \cdot \text{Et}_2\text{O}$	Y
2	$\text{Bi}(\text{OTf})_3$	Y	13	$\text{Bi}(\text{OTf})_3$	Y	24	$\text{Bi}(\text{OTf})_3$	Y
3	$\text{In}(\text{OTf})_3$	Y	14	$\text{In}(\text{OTf})_3$	Y	25	$\text{In}(\text{OTf})_3$	Y
4	FeCl_2	N	15	FeCl_2	N	26	FeCl_2	N
5	FeCl_3	N	16	FeCl_3	N	27	FeCl_3	N
6	ZnCl_2	Y	17	ZnCl_2	Y	28	ZnCl_2	Y
7	SnCl_2	Y	18	SnCl_2	Y	29	SnCl_2	Y
8	CoCl_2	N	19	CoCl_2	N	30	CoCl_2	N
9	NiCl_2	N	20	NiCl_2	N	31	NiCl_2	N
10	$\text{Cu}(\text{MeCN})_4$	N	21	$\text{Cu}(\text{MeCN})_4$	N	32	$\text{Cu}(\text{MeCN})_4$	N
11	CuCl_2	N	22	CuCl_2	N	33	CuCl_2	N

*Y - Product revealed by TLC; N - product not observed by TLC.

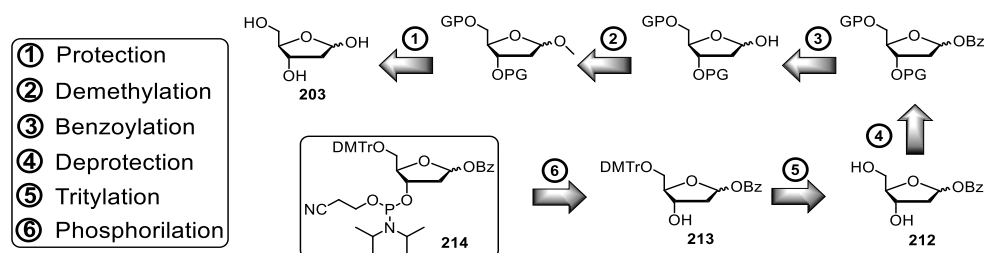
From this screening, $\text{BF}_3 \cdot \text{Et}_2\text{O}$, $\text{Bi}(\text{OTf})_3$ and $\text{In}(\text{OTf})_3$ proved to be suitable candidates to activate the three glycoside donors (**207a-207c**). ZnCl_2 and SnCl_2 could also promote the formation of the target product **208**. However, their extremely low solubility in organic solvents limits their further applicability to catalyze reactions on oligonucleotides. Unfortunately, no matter the benzoyl moiety used (**207a-207c**), no orthogonal conditions could be found in this preliminary screening considering these eleven Lewis acids employed. We presuppose that the reduced pK_a differences between the three corresponding benzoic

acids (*p*-trifluoromethylbenzoic acid, benzoic acid and *p*-methoxybenzoic acid) may not be high enough ($\Delta(\text{p}K_a) \approx 1$) to afford the selectivity.

For all conditions where the target product was formed, degradation of the starting material was also observed albeit to a small extent. It is also important to highlight that anhydrous salts were used to avoid hydrolysis of the starting material. When $\text{FeCl}_2 \cdot 4\text{H}_2\text{O}$ was employed as source of Iron(II), the target product could be observed, however a large amount of starting material degradation was observed (entries 4, 15 and 26). In addition, the use of acetonitrile as a solvent do not provide any significant changes compared with dichloromethane. However, knowing that acetonitrile is a viable solvent for this transformation is a valuable information as it solubilizes efficiently most of the Lewis acids and could provide some control of the anomery of the final product. Importantly, increasing the TMSN_3 concentration (2 equivalents) was beneficial for the reaction: lower amounts of degraded starting material was observed (TLC analysis). Finally, concerning this first and preliminary assay, we decided to select the simplest derivative containing the plain benzoyl group (**207a**) to proceed further towards the preparation of DNA strands incorporating this platform.

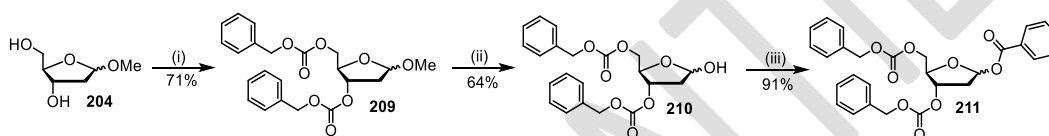
5.2.2 The Synthesis of Phosphoramite **214**

In order to obtain oligonucleotides suitable for post-synthetic modification, our first step consisted in the synthesis of the phosphoramidite pendant of **207a**. Starting from D-2-deoxyribose, the envisaged retrosynthetic pathway to access the respective phosphoramidite derivative **214** is presented in the *Scheme 5.5*. Differently from the model compounds applied to study the Lewis acid activation, here we used a benzyloxycarbonate (Cbz) to protect both hydroxyl groups on the 3' and 5' positions of the sugar framework. We decided to switch the protection strategy as Cbz groups can be readily removed while benzyl ethers can sometimes be tedious to cleave off.



Scheme 5.5. Retrosynthetic pathway towards the target phosphoramidite **214**.

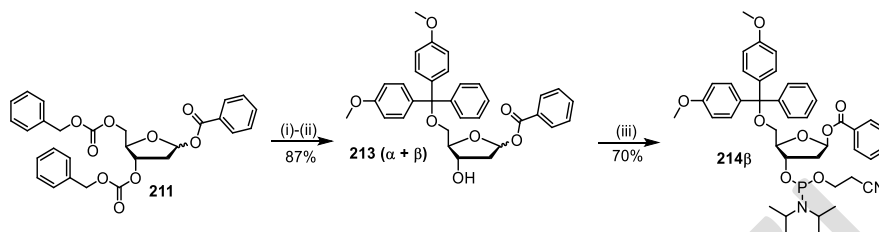
Starting from the acetal **204**, the intermediate 3',5'-di-*O*-Cbz-1'-OBz-D-2-deoxyribose (**211**) was synthesized after three steps, in 41% overall yield (*Scheme 5.6*). Firstly, **204** was dissolved in DCM and treated with an excess (4.5 equiv.) of benzylchloroformate, in the presence of *N,N*-(dimethylamino)pyridine (DMAP) as the base to afford the 3',5'-bis-*O*-Cbz protected D-2-deoxyribose **209** as a colorless oil after purification (71%). Next, the compound **209** was dissolved in an acetic acid/water (4/1, v/v) mixture and heated at 60°C for 48 h to afford the key intermediate **210** (64% yield) as a white solid after chromatography. Finally, the acylation step was performed at room temperature, in dry pyridine, using 1.5 equivalents of benzoyl chloride. The corresponding 1'-*O*-benzoylated intermediate **211** was obtained in 91% yield as a mixture of α and β anomers.



Scheme 5.6. Synthesis of intermediate **211**. Reaction and conditions. (i) CbzCl (4.5 equiv.), DMAP (6 equiv.), DCM, r.t. (ii) AcOH/H₂O (8/2), 60°C, 48 h. (iii) BzCl (1.2 equiv.), pyridine, 0 - r.t., overnight.

Once with the suitable aglycone installed, the compound **211** had to be functionalized to produce the phosphoramidite reagent required for the oligonucleotide synthesis. This was performed *via* a three steps synthetic pathway (*Scheme 5.7*). Initially, the intermediate **211** was subjected to hydrogenolysis, to remove the Cbz protecting groups using H₂ and Pd/C, in absolute ethanol containing 1% of TEA. Satisfactorily, the reaction was completed in less than 2 h (TLC monitoring). Then, the crude **212** was filtered over Celite, in order to remove the metal catalyst and co-evaporated three times with dry pyridine to remove all the traces of ethanol. The product **212** was used in the following step without purification. Next, the 3',5'-di-OH-1'-OBz-D-2-deoxyribose (**212**) was solubilized in dry pyridine, and treated with a slight excess of 4,4'-dimethoxytrityl chloride (DMTrCl, 1.2 equiv.), at room temperature, overnight, to afford the 1'-OBz-5'-trytilated intermediate **213** ($\alpha + \beta$), in 87% yield after chromatography. During the column chromatography of **213**, a pure fraction of the β anomer could be isolated (**213 β**). This pure anomer was subsequently converted to the corresponding phosphoramidite after treatment with 1.2 equivalents of 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite and using 3 equivalents of freshly distilled *N,N*-

diisopropylethylamine (DIEA) in dry DCM. The phosphoramidite **213 β** was obtained after purification on silica gel column chromatography in 70% yield. It exists as a mixture of 2 diastereoisomers as the phosphorous atom is a chirogenic centre.

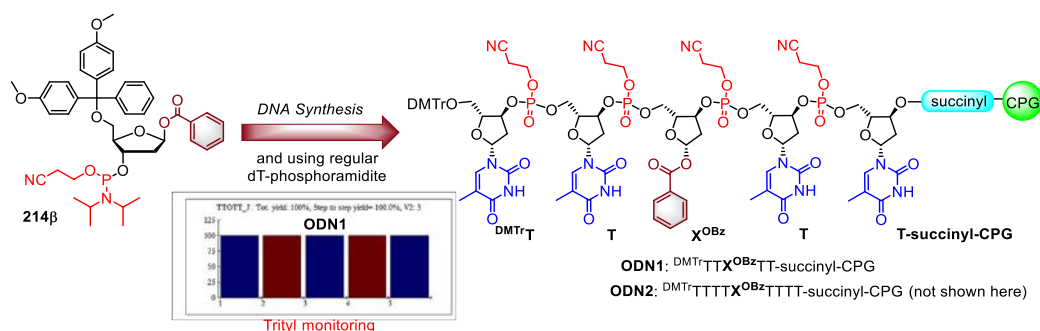


Scheme 5.7. Synthesis of phosphoramidite **214 β** . Reaction and conditions. (i) H₂, Pd/C (10%), TEA (1%), EtOH, r.t., 2 h. (ii) DMTrCl (1.2 equiv.), pyridine, r.t., overnight. (iii) (using pure β anomer of **213**) 2-Cyanoethyl *N,N*-diisopropylchlorophosphoramidite (1.2 equiv.), DIEA (3 equiv.), DCM, r.t., 1h.

5.2.3 Oligonucleotide Synthesis

Following the synthesis of the phosphoramidite reagent **214 β** , we proceed to the next step of our project: the incorporation of our aglycone platform into oligonucleotide sequences. Initially, to verify the feasibility of our approach, we projected the synthesis of two different oligonucleotide sequences. These sequences, displaying poly(thymine) strands with 5 or 9 nucleosides, in which one T is replaced with our aglycone platform, were selected to check the compatibility of our phosphoramidite reagent with the regular phosphoramidite elongation chemistry.

The oligonucleotide synthesis was then performed in an automated oligonucleotide synthesizer (1 μ M scale) using the reagent **214 β** together with the regular dT-phosphoramidite. Pleasantly, phosphoramidite **214 β** exhibited a high coupling efficiency during its incorporation in the oligonucleotide sequences **ODN1** and **ODN2** (coupling step > 95%, for 10 independent experiments, considering **ODN1** with DMTr OFF mode) (**Scheme 5.8**). In fact, **214 β** disclose similar coupling yields when compared with the commercially available dT-phosphoramidite used to prepare these two sequences.



Scheme 5.8. Incorporation of the aglycone moiety into two nucleic acid sequences (**ODN1** and **ODN2**).

The apparent success concerning the incorporation of our phosphoramidite platform **214β** in the both oligonucleotide strands (**ODN1** and **ODN2**) was confirmed by the MALDI-ToF mass spectrometry analysis of the supported oligonucleotides (**Figure 5.19**). In this case, the laser of the MALDI apparatus induces both the release of the oligonucleotide from the solid support (*via* fragmentation at phosphate junctions) and its desorption/ionization. On the MALDI-ToF analysis of **ODN1** (**Figure 5.19**), it is possible to observe the appearance of two important fragments. The most relevant corresponds to the $DMTrTTX^{OBz}T$ fragment (at $m/z = 1748$ Da; NB: a lack default of calibration during MALDI-ToF MS analysis resulted in about 5 Da bias between calculated and observed m/z values). This fragment corresponds to the (n-1)-oligonucleotide where cleavage at the first 3'-end phosphate junction occurred during MALDI-ToF MS analysis.

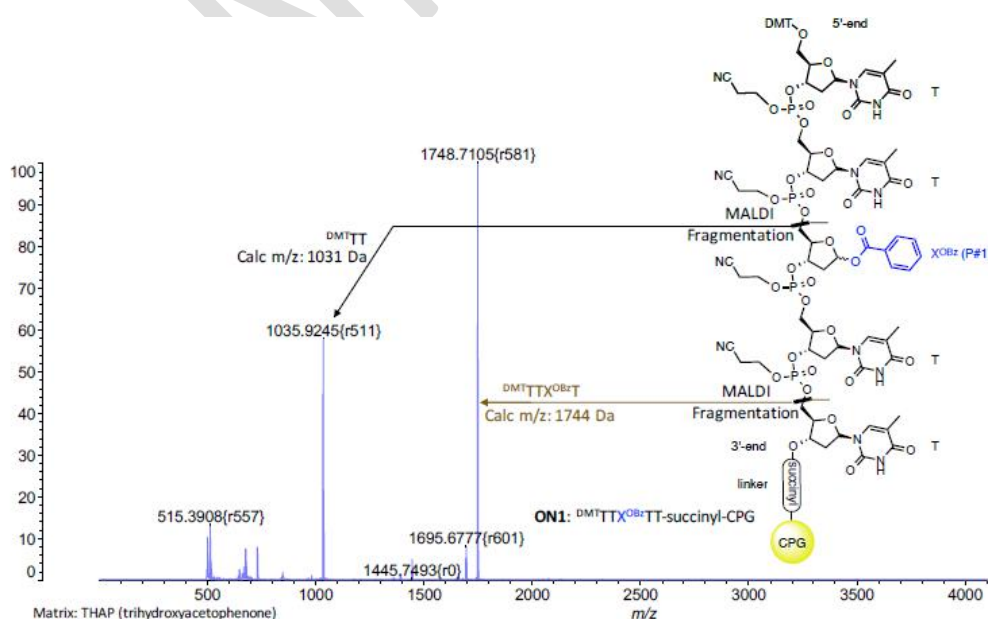
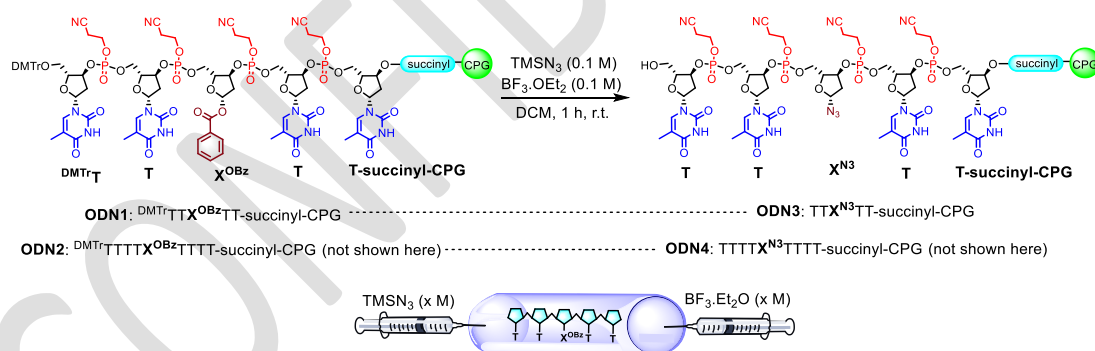


Figure 5.19. MALDI-ToF MS analysis of **ODN1** in negative mode.

5.2.4 Preliminary Studies on the Post-Synthetic Functionalization of ODN1 and ODN2

Following the synthesis of the both oligonucleotide sequences, **ODN1** and **ODN2**, we start to investigate the feasibility to use glycosidation reactions to post-functionalize both ON sequences. For this purpose, we set up a first trial-reaction using TMSN_3 (the same model nucleophile used during the Lewis acid screening on the nucleosidic benzoyl aglycone model compounds, *section 5.2.1.2*) in the presence $\text{BF}_3 \cdot \text{OEt}_2$ (which was also selected from our previous screening) as the Lewis acid. Considering the base-lability of the benzoyl aglycone moiety, the PSF reactions were performed while the ONs were still grafted onto the solid support (CPG). To perform this reaction, the support-bonded oligonucleotides **ODN1** and **ODN2** were placed in a DNA synthesis column. Afterwards, the Lewis acid and the glycosyl acceptor (nucleophile) were pumped together inside the column (**Scheme 5.9**). For this preliminary assay, dichloromethane solutions of both, $\text{BF}_3 \cdot \text{Et}_2\text{O}$ (0.1 M) and TMSN_3 (0.1 M) were pumped inside the cell and the reaction was kept at room temperature during 1 hour. Finally, the reagents were removed from the cell, the oligonucleotide was extensively washed with anhydrous acetonitrile, dried and directly analyzed by MALDI-ToF mass spectrometry.



Scheme 5.9. First trial post-synthetic functionalization reaction on **ODN1** and **ODN2**.

Pleasingly, the MALDI-ToF MS analysis clearly indicates that **ODN1** was converted into the expected product containing the azido function (**ODN3**, X^{N_3} residue) in replacement of the OBz group (X^{OBz} , **Figure 5.20**). Of note, the acidic condition used to perform this post-synthetic glycosidation reaction ($\text{BF}_3 \cdot \text{OEt}_2$) also triggers the cleavage of the 5'-end DMTr group during the course of the reaction.

The MALDI-ToF MS analysis of the crude reaction revealed a signal at m/z 1368 Da, which corresponds to the fragment $\text{TTX}^{\text{N}_3}\text{T}$, and is characteristic of the product **ODN3**. In addition,

the spectrum also features a signal that corresponds to the fragment $\text{TTX}^{\text{OBz}}\text{T}$ (m/z 1448 Da) stemming from unconverted starting material (**ODN1**) albeit with the loss of the 5'-DMTr group. Finally, minor signals at m/z 1394 and 1342 correspond to $\text{TTX}^{\text{OBz}}\text{T}$ with the loss of one or two cyanoethyl moiety (CNE) from phosphate groups during the MS analysis. (NB: For the same reason as aforementioned, on spectrum **Figure 5.20**, m/z values are also 5 Da-biased).

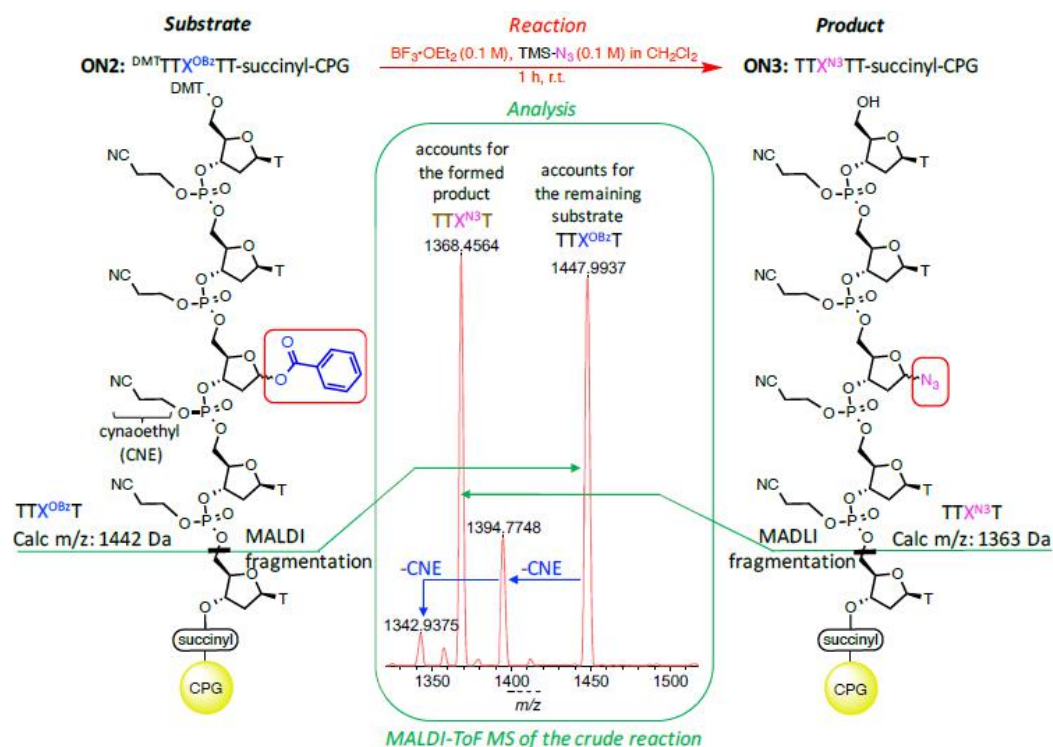


Figure 5.20. Reaction and MALDI-ToF MS analysis of the test reaction between **ODN1** and TMSN_3 , in presence of $\text{BF}_3 \cdot \text{OEt}_2$.

The post-synthetic glycosylation reaction was also ascertained for the longer oligonucleotide strand **ODN2** ($\text{DMTrTTT}\text{TTX}^{\text{OBz}}\text{TTTT-succinyl-CPG}$). Similarly as observed for **ODN1** ($\text{DMTrTTX}^{\text{OBz}}\text{TT-succinyl-CPG}$), the reaction of **ODN2** with TMSN_3 in the presence of $\text{BF}_3 \cdot \text{OEt}_2$ also afforded the expected azido product **ODN4**. Indeed, the MALDI-ToF MS spectra of the crude reaction (**Figure 5.21**) also features the same pattern as observed for **ODN1**; m/z signals corresponding to the expected product (**ODN4**) and to the unconverted starting oligonucleotide (**ODN2**). The first set, in pink (m/z : 2796, 2742 and 2689) corresponds to the fragmentation of the expected product **ODN4**.

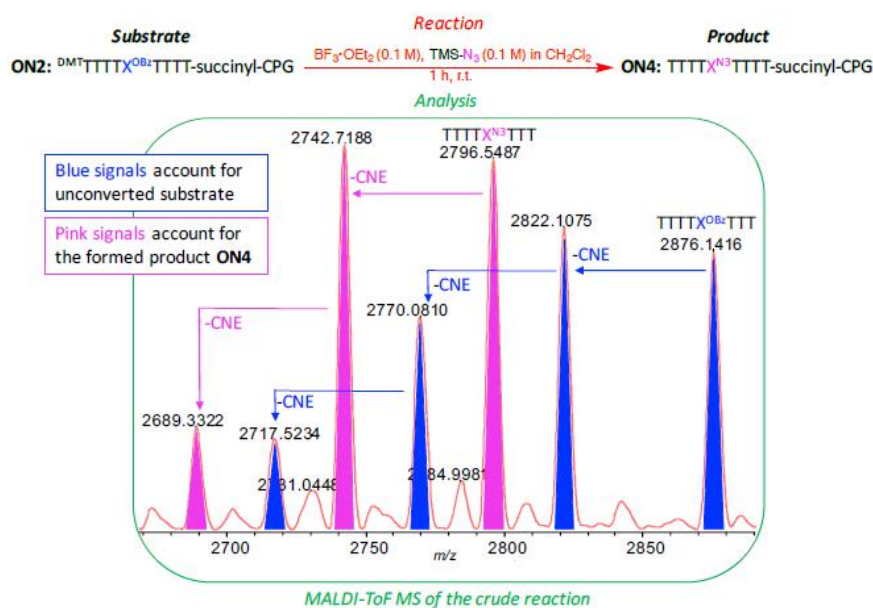


Figure 5.21. Reaction and MALDI-ToF MS analysis of the trial Fisher glycosidation reaction between **ODN2** and TMSN_3 in the presence of $\text{BF}_3 \cdot \text{OEt}_2$.

In a similar manner, **ODN4** is also observed as its (n-1)-oligomer $\text{TTTTX}^{\text{N}_3}\text{TTT}$ with and without the loss of cyanoethyl moieties (CNE). Likewise, unconverted starting material **ODN2**, which has lost its 5'-DMTr, is observed as its (n-1)-oligomer $\text{TTTTX}^{\text{OBz}}\text{TTT}$ with and without the loss of CNEs (blue set of signals; m/z 2876, 2822, 2770, 2717).

Altogether, these preliminary results strongly support the feasibility of this post-synthetic glycosidation reaction to functionalize oligonucleotides. Indeed, we could also show that the phosphoramidite platforms can be readily prepared and used in the chemical synthesis of oligonucleotides. Despite the success of this approach, the orthogonal activation using electronically tuned benzoyl aglycones could not be ascertained yet. Currently, we are working to generate a second 2'-deoxyribose platform to further explore the orthogonal post-synthetic functionalization. This ongoing studies will be briefly contextualized in the next section.

5.3 Method Optimization and the Alternative Approach for the Orthogonal PSF

5.3.1 Complementary Studies on the Post-Synthetic Functionalization using the **ODN1**

Once the viability of our methodology was ascertained, we initiated a more complete investigation in order to identify optimized conditions to perform this glycosylation reaction. Thus, a preliminary screening of the reactions conditions, using **ODN1** and the model reaction (TMSN₃ in presence of Lewis acid) is currently ongoing. Essentially, four parameters are being evaluated: the reaction time (1, 2, 3 and 4 h), the solvent (dichloromethane or acetonitrile), the Lewis acid (BF₃.OEt₂, Bi(OTf)₃ and In(OTf)₃) and the reagent (Lewis acid and TMSN₃) concentration (0.1 M and 1 M). Meanwhile, preliminary studies were also started to investigate the possibility to use O-nucleophiles. Therefore, similar reactions, using *p*-methoxyphenol as the glycosyl acceptor, are also being evaluated.

5.3.2 *Ortho*-Alkynyl Benzoates - The Potential and Orthogonally Activable Synthons

Alternatively to the use of oxophilic Lewis acids that promote the activation of the anomeric position, we thought to explore *ortho*-alkynyl benzoates as a potential and compatible group to design an orthogonally activatable 2'-deoxyribose platform for glycosylation reaction. During the last decade, several groups have been exploring the carbophilicity of gold catalysts to notably activate alkyne functions and to perform the synthesis of iso-coumarine derivatives¹⁴⁴ or to activate electrophiles for nucleophilic substitution reactions.^{145,146} Recently, two groups described the use *o*-alkynyl benzoates in sugar scaffolds. Ding and co-workers reported the use of *o*-alkynyl benzoates as an efficient and easily removed protecting group for the 2'-position to design a series of substituted ribonucleosides (**Figure 5.22**).¹⁴⁷ In different works, the group of Yu demonstrated the high efficiency of glycosyl *o*-alkynyl benzoates as donors for glycosylation reactions.¹⁴⁸⁻¹⁵² In their seminal work (2008), the glycosidic coupling of glycosyl *o*-hexynyl benzoate with alcohols was successfully described by using PPh₃AuOTf (**Figure 5.23-A**).¹⁵² More recently, with the use of a similar methodology, the synthesis of β -rhamnopyranosides, with good to excellent anomery control was also described using Ph₃PAuBAR₄^F as the catalyst (**Figure 5.23-B**).¹⁵¹ Thus, *o*-

alkynylbenzoates may be a potential alternative to achieve synthons with orthogonal activation.

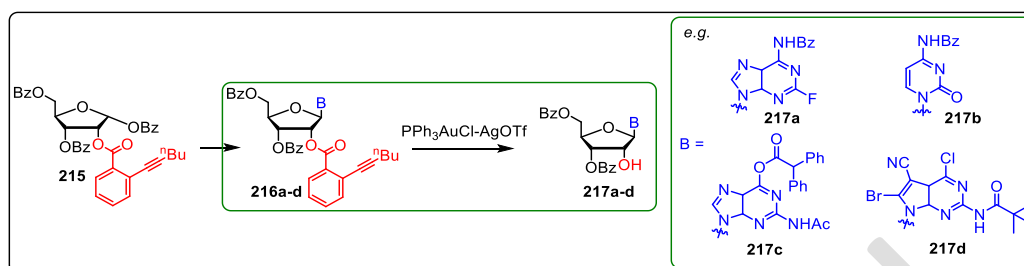


Figure 5.22. The use of *o*-hexynyl benzoate as the 2'-OH protecting group for the synthesis of modified nucleosides reported by Ding.

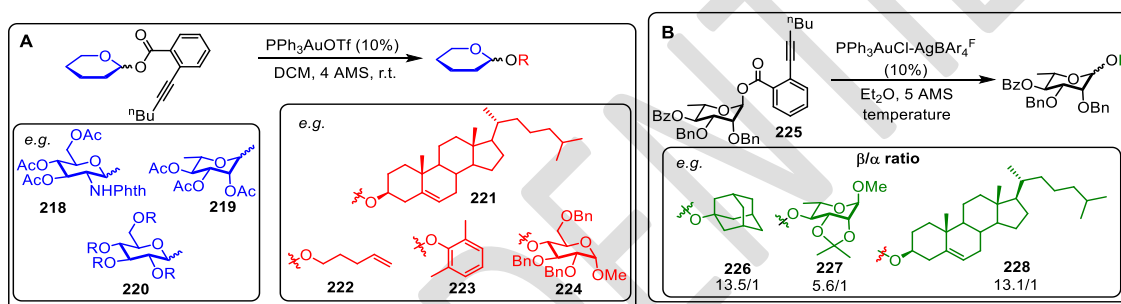


Figure 5.23. Examples of the use of glycosyl *o*-alkynyl benzoates as donors for glycosylation reactions.

Therefore, based on the aforementioned works, especially the ones dealing with glycosidation reactions on 6-membered saccharides, we envisaged to explore the same chemistry on 2'-deoxyribose platforms. Within this concept, we plan to access a 2'-deoxyribose glycoside framework bearing an ortho-alkynyl benzoate aglycone moiety. In fact, using this approach, we expect that activation of the anomeric position (formation of the *oxonium* intermediate) in this second platform can be reached using the conventional or an adapted version of the gold catalyst systems reported by Yu (**Figure 5.24**).

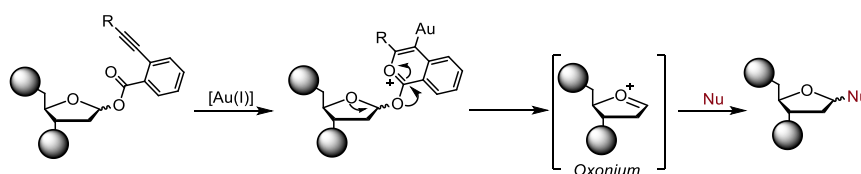


Figure 5.24. Proposed mechanism for the activation of the anomeric position with Au(I) catalyst.

To ascertain this hypothesis, we firstly verified the stability of the our model compound **5.5a** in a benchmark reaction using TMSN_3 and the gold catalyst $\text{PPh}_3\text{AuNTf}_2$ (Gagosz's catalyst). Pleasingly, the starting material remained unaffected after 1 h of reaction using 1 equivalent of TMSN_3 and 10 % of gold catalyst. Thus, taking into account this preliminary information, it appears that 1'-OBz and 1'-OBz(hexynyl) glycosyl donor could be orthogonally activated. Thus, we planned to access, in a similar manner as used for the first series of benzoyl derivatives (**207a-207g**), the model deoxyribose platform **229** (**Figure 5.25**) displaying an *ortho*-alkynyl benzoate scaffold. Compound **229** will be used to verify the validity of this hypothesis and to set on optimized conditions for its activation. These optimized conditions will be subsequently used for the PSF of oligonucleotides. Finally, the phosphoramidite synthon corresponding to **229** will be synthesized and used to incorporate this second D-2-deoxyribose platform into oligonucleotide sequences. With two platform orthogonal phosphoramidites in hands we should be able to perform either simple PSF of ODN (**Figure 5.26-A**, using 5.Ph2 for example) or orthogonal transformation in ODNs sequences bearing both 2'-deoxyribose platforms (**Figure 5.26-B**).

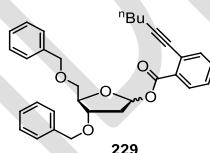


Figure 5.25. Structures of the model compound **229** bearing the *ortho*-hexynyl benzoate scaffold.

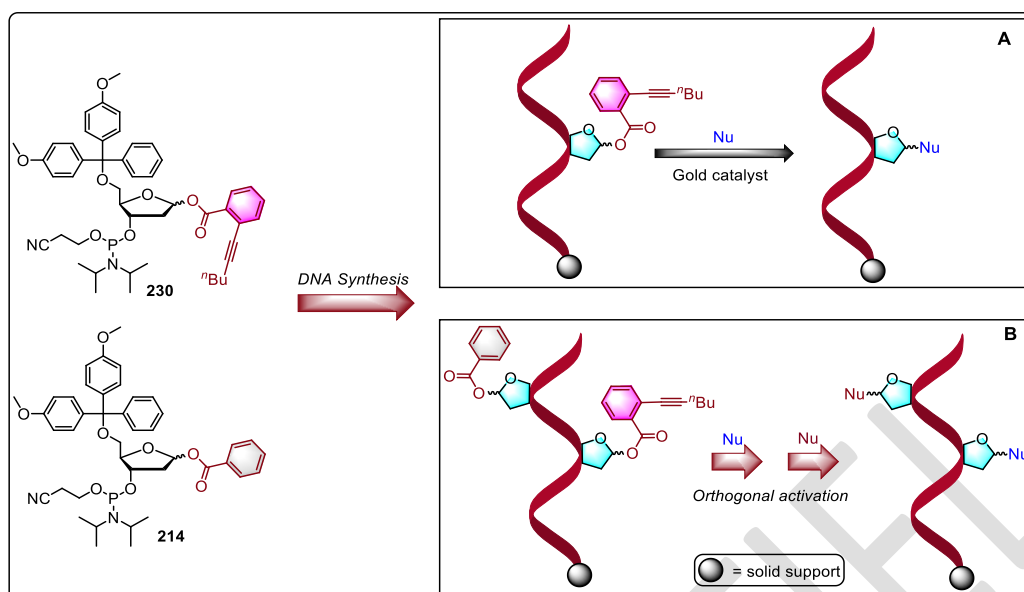


Figure 5.26. Schematic representation concerning the use of the benzoyl aglycone platform **230** to perform PSF of nucleic acids. (A) simple PSF. (B) orthogonal PSF using both platforms **214** and **230**.

5.4. Conclusions - Chapter 5

In summary, in this chapter we disclosed the development of a new methodology to post-synthetically functionalize oligonucleotides. We described and validated the first methodology applied to the direct PSF on the anomeric position, based on a glycosidation reaction. In this work, a Lewis acid activatable 1'-*O*-benzoyl-2'-deoxyribose platforms could be identified and efficiently incorporated into oligonucleotides using phosphoramidite elongation chemistry. Similar coupling efficiency (> 95%) with the regular dT-phosphoramidite synthon were achieved, as well as high stability of the benzoyl aglycone moiety to the acidic condition used during all cycles of trityl deprotection. Moreover, the preliminary studies concerning the PSF using our modified oligonucleotides ascertained the feasibility of our method. Finally, further studies to optimize the reaction conditions as well as the use of different nucleophiles are currently in progress.

References

- (1) Belmont, P.; Constant, J. F.; Demeunynck, M. *Chem. Soc. Rev.* **2001**, *30*, 70–81.
- (2) Watson, J. D.; Crick, F. C. H. *Nature* **1953**, *171*, 737–738.
- (3) Lodish, H.; Berk, A.; Matsudaira, P.; Kaiser, C. A.; Krieger, M.; Scott, M. P.; Zipursky, L.; Darnell, J. *Molecular Cell Biology*, New York, Freeman, W. H. & Company, 5th edn, 2003; Vol. 1.
- (4) Mortimer, S. A.; Kidwell, M. A.; Doudna, J. A. *Nat. Rev. Genet.* **2014**, *15*, 469–479.
- (5) Lim, C. S.; Brown, C. M. *Front. Microbiol.* **2018**, *8*, 2582.
- (6) Varani, G.; McClain, W. H. *EMBO Rep.* **2000**, *1*, 18–23.
- (7) Chen, J. J.-L.; Greider, C. W. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 8080–8085.
- (8) Alotte, C.; Martin, A.; Caldarelli, S. A.; Di Giorgio, A.; Condom, R.; Zoulim, F.; Durantel, D.; Hantz, O. *Antiviral Res.* **2008**, *80*, 280–287.
- (9) van Batenburg, F. H.; Gultyaev, A. P.; Pleij, C. W. *Nucleic Acids Res.* **2001**, *29*, 194–195.
- (10) Cao, S.; Xu, X.; Chen, S. J. *RNA* **2014**, *20*, 835–845.
- (11) Singh, A.; Sethaphong, L.; Yingling, Y. G. *Biophys. J.* **2011**, *101*, 727–735.
- (12) Conrad, N. K. *Wiley Interdiscip. Rev. RNA* **2014**, *5*, 15–29.
- (13) Hoogsteen, K. *Acta Crystallogr.* **1963**, *16*, 907–916.
- (14) Thuong, N. T.; Hélène, C. *Angew. Chem. Int. Ed.* **1993**, *32*, 666–690.
- (15) Qiao, F.; Cech, T. R. *Nat. Struct. Mol. Biol.* **2008**, *15*, 634–640.
- (16) Devi, G.; Zhou, Y.; Zhong, Z.; Toh, D. F. K.; Chen, G. *Wiley Interdiscip. Rev. RNA* **2015**, *6*, 111–128.
- (17) Schimmel, P. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 45–58.
- (18) Schmeing, T. M.; Ramakrishnan, V. *Nature* **2009**, *461*, 1234–1242.
- (19) Kung, J. T. Y.; Colognori, D.; Lee, J. T. *Genetics* **2013**, *193*, 651–669.
- (20) Zhao, X.-Y.; Xiong, X.; Liu, T.; Mi, L.; Peng, X.; Rui, C.; Guo, L.; Li, S.; Li, X.; Lin, J. D. *Nat. Commun.* **2018**, *9*, 2986.
- (21) Choudhuri, S. J. *Biochem. Mol. Toxicol.* **2010**, *24*, 195–216.
- (22) Iorio, M. V.; Croce, C. M. *Carcinogenesis* **2012**, *33*, 1126–1133.
- (23) He, L.; Hannon, G. J. *Nat. Rev. Genet.* **2004**, *5*, 522–531.
- (24) Webster, D. P.; Klenerman, P.; Dusheiko, G. M. *Lancet* **2015**, *385*, 1124–1135.

- (25) Smith, D. B.; Bukh, J.; Kuiken, C.; Muerhoff, A. S.; Rice, C. M.; Stapleton, J. T.; Simmonds, P. *Hepatology* **2014**, *59*, 318–327.
- (26) Lukavsky, P. J. *Virus Res.* **2009**, *139*, 166–171.
- (27) Mauger, D. M.; Golden, M.; Yamane, D.; Williford, S.; Lemon, S. M.; Martin, D. P.; Weeks, K. M. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 3692–3697.
- (28) Martinello, M.; Hajarizadeh, B.; Grebely, J.; Dore, G. J.; Matthews, G. V. *Nat. Rev. Gastroenterol. Hepatol.* **2018**, *15*, 412–424.
- (29) Gentile, I.; Maraolo, A. E.; Buonomo, A. R.; Zappulo, E.; Borgia, G. *Expert Opin. Drug Discov.* **2015**, *10*, 1363–1377.
- (30) Nakamura, M.; Kanda, T.; Haga, Y.; Sasaki, R.; Wu, S.; Nakamoto, S.; Yasui, S.; Arai, M.; Imazeki, F.; Yokosuka, O. *World J. Hepatol.* **2016**, *8*, 183–190.
- (31) Janssen, H. L.; Reesink, H. W.; Lawitz, E. J.; Zeuzem, S.; Rodriguez-Torres, M.; Patel, K.; van der Meer, A. J.; Patick, A. K.; Chen, A.; Zhou, Y.; Persson, R.; King, B. D.; Kauppinen, S.; Levin, A. A.; Hodges, M. R. *N. Engl. J. Med.* **2013**, *368*, 1685–1694.
- (32) Dibrov, S. M.; Parsons, J.; Carnevali, M.; Zhou, S.; Rynearson, K. D.; Ding, K.; Garcia Sega, E.; Brunn, N. D.; Boerneke, M. A.; Castaldi, M. P.; Hermann, T. *J. Med. Chem.* **2014**, *57*, 1694–1707.
- (33) Geinguenaud, F.; Guenin, E.; Lalatonne, Y.; Motte, L. *ACS Chem. Biol.* **2016**, *11*, 1180–1191.
- (34) Thomas, J. R.; Hergenrother, P. J. *Chem. Rev.* **2008**, *108*, 1176–1224.
- (35) Adams, D.; Gonzalez-Duarte, A.; O’Riordan, W. D.; Yang, C.-C.; Ueda, M.; Kristen, A. V.; Tournev, I.; Schmidt, H. H.; Coelho, T.; Berk, J. L.; Lin, K.-P.; Vita, G.; Attarian, S.; Planté-Bordeneuve, V.; Mezei, M. M.; Campistol, J. M.; Buades, J.; Brannagan, T. H.; Kim, B. J.; Oh, J.; Parman, Y.; Sekijima, Y.; Hawkins, P. N.; Solomon, S. D.; Polydefkis, M.; Dyck, P. J.; Gandhi, P. J.; Goyal, S.; Chen, J.; Strahs, A. L.; Nochur, S. V.; Sweetser, M. T.; Garg, P. P.; Vaishnav, A. K.; Gollob, J. A.; Suhr, O. B. *N. Engl. J. Med.* **2018**, *379*, 11–21.
- (36) Stephenson, M. L.; Zamecnik, P. C. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 285–288.
- (37) Zamecnik, P. C.; Stephenson, M. L. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 280–284.
- (38) Chan, J. H. P.; Lim, S.; Wong, W. S. F. *Clin. Exp. Pharmacol. Physiol.* **2006**, *33*, 533–540.
- (39) Mutso, M.; Nikonov, A.; Pihlak, A.; Žusinaite, E.; Viru, L.; Selyutina, A.; Reintamm, T.; Kelve, M.; Saarma, M.; Karelson, M.; Merits, A. *PLoS One* **2015**, *10*, e0128686.
- (40) Iwamoto, N.; Butler, D. C. D.; Svrzikapa, N.; Mohapatra, S.; Zlatev, I.; Sah, Di. W. Y.; Meena; Standley, S. M.; Lu, G.; Apponi, L. H.; Frank-Kamenetsky, M.; Zhang, J. J.; Vargeese, C.; Verdine, G. L. *Nat. Biotechnol.* **2017**, *35*, 845–851.
- (41) McHutchison, J. G.; Patel, K.; Pockros, P.; Nyberg, L.; Pianko, S.; Yu, R. Z.; Andrew

- Dorr, F.; Kwoh, T. J. *J. Hepatol.* **2006**, *44*, 88–96.
- (42) Andrei, G.; De Clercq, E.; Snoeck, R. *Infect. Disord. Drug Targets* **2009**, *9*, 201–222.
- (43) Hanecak, R.; Brown-Driver, V.; Fox, M.; Azad, R.; Furusako, S.; Nozaki, C.; Ford, C.; Sasmor, H.; Anderson, K. *J. Virol.* **1996**, *70*, 5203–5212.
- (44) Zhang, H.; Hanecak, R.; Brown-Driver, V.; Azad, R.; Conklin, B.; Fox, M. C.; Anderson, K. P. *Antimicrob. Agents Chemother.* **1999**, *43*, 347–353.
- (45) Soler, M.; McHutchison, J. G.; Kwoh, T. J.; Dorr, F. A.; Pawlotsky, J. M. *Antivir Ther* **2004**, *9*, 953–968.
- (46) Georgopapadakou, N. *Expert Opin. Investig. Drugs* **2007**, *16*, 1–10.
- (47) Campbell, M. A.; Wengel, J. *Chem. Soc. Rev.* **2011**, *40*, 5680–5689.
- (48) Gebert, L. F. R.; Rebhan, M. A. E.; Crivelli, S. E. M.; Denzler, R.; Stoffel, M.; Hall, J. *Nucleic Acids Res.* **2014**, *42*, 609–621.
- (49) Ottosen, S.; Parsley, T. B.; Yang, L.; Zeh, K.; Van Doorn, L. J.; Van Der Veer, E.; Raney, A. K.; Hodges, M. R.; Patick, A. K. *Antimicrob. Agents Chemother.* **2015**, *59*, 599–608.
- (50) Laxton, C.; Brady, K.; Moschos, S.; Turnpenney, P.; Rawal, J.; Pryde, D. C.; Sidders, B.; Corbau, R.; Pickford, C.; Murray, E. J. *Antimicrob. Agents Chemother.* **2011**, *55*, 3105–3114.
- (51) Lind, K. E.; Mohan, V.; Manoharan, M.; Ferguson, D. M. *Nucleic Acids Res* **1998**, *26*, 3694–3799.
- (52) Henry, S.; Stecker, K.; Brooks, D.; Monteith, D.; Conklin, B.; Bennett, C. F. *J. Pharmacol. Exp. Ther.* **2000**, *292*, 468–479.
- (53) Tallet-Lopez, B.; Aldaz-Carroll, L.; Chabas, S.; Dausse, E.; Staedel, C.; Toulme, J. J. *Nucleic Acids Res.* **2003**, *31*, 734–742.
- (54) Berry, K. E.; Wagharay, S.; Doudna, J. A. *RNA* **2010**, *16*, 1559–1569.
- (55) Doyle, D. F.; Braasch, D. A.; Simmons, C. G.; Janowski, B. A.; Corey, D. R. *Biochemistry* **2001**, *40*, 53–64.
- (56) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497–1500.
- (57) Egholm, M. *Nature* **1993**, *366*, 529–531.
- (58) Liu, Y.; Braasch, D. A.; Nulf, C. J.; Corey, D. R. *Biochemistry* **2004**, *43*, 1921–1927.
- (59) Nulf, C. J.; Corey, D. *Nucleic Acids Res.* **2004**, *32*, 3792–3798.
- (60) Paul, S.; Caruthers, M. H. *J. Am. Chem. Soc.* **2016**, *138*, 15663–15672.
- (61) Nan, Y.; Zhang, Y. J. *Front. Microbiol.* **2018**, *9*, 750.
- (62) Blum, M.; De Robertis, E. M.; Wallingford, J. B.; Niehrs, C. *Dev. Cell* **2015**, *35*, 145–149.

- (63) Hudziak, R. M.; Barofsky, E.; Barofsky, D. F.; Weller, D. L.; Huang, S.-B.; Weller, D. *Antisense Nucleic Acid Drug Dev.* **1996**, *6*, 267–272.
- (64) Aartsma-Rus, A.; Krieg, A. M. *Nucleic Acid Ther.* **2017**, *27*, 1–3.
- (65) Jubin, R.; Vantuno, N. E.; Kieft, J. S.; Murray, M. G.; Doudna, J. a; Lau, J. Y.; Baroudy, B. M. *J. Virol.* **2000**, *74*, 10430–10437.
- (66) McCaffrey, A. P.; Meuse, L.; Karimi, M.; Contag, C. H.; Kay, M. A. *Hepatology* **2003**, *38*, 503–508.
- (67) Kikuchi, K.; Umehara, T.; Nishikawa, F.; Fukuda, K.; Hasegawa, T.; Nishikawa, S. *Biochem. Biophys. Res. Commun.* **2009**, *386*, 118–123.
- (68) Ladju, R. B.; Pascut, D.; Massi, M. N.; Tiribelli, C.; Sukowati, C. H. C. *Oncotarget* **2018**, *9*, 2951–2961.
- (69) Sun, H.; Zhu, X.; Lu, P. Y.; Rosato, R. R.; Tan, W.; Zu, Y. *Mol. Ther. Nucleic Acids* **2014**, *3*, e182.
- (70) Kikuchi, K.; Umeharal, T.; Fukudal, K.; Hwang, J.; Kuno, A.; Hasegawa, T.; Nishikawa, S.; Resources, B.; Science, A. I. **2003**, *270*, 263–270.
- (71) Kikuchi, K.; Umehara, T.; Fukuda, K.; Kuno, A.; Hasegawa, T.; Nishikawa, S. *Nucleic Acids Res.* **2005**, *33*, 683–692.
- (72) Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C. *Nature* **1998**, *391*, 806–811.
- (73) Neema Agrawal, P. V. N. Dasaradhi, Asif Mohmmmed, Pawan Malhotra, Raj K. Bhatnagar, and S. K. M. *Mikrobiyol. Bul.* **2004**, *38*, 285–294.
- (74) Lee, C. H.; Kim, J. H.; Lee, S. W. *World J. Gastroenterol.* **2013**, *19*, 8949–8962.
- (75) Hamazaki, H.; Takahashi, H.; Shimotohno, K.; Miyano-Kurosaki, N.; Takaku, H. *Nucleosides, Nucleotides Nucleic Acids* **2006**, *25*, 801–805.
- (76) Korf, M.; Jarczak, D.; Beger, C.; Manns, M. P.; Krüger, M. *J. Hepatol.* **2005**, *43*, 225–234.
- (77) Moon, J.-S.; Lee, S.-H.; Kim, E.-J.; Cho, H.; Lee, W.; Kim, G.-W.; Park, H.-J.; Cho, S.-W.; Lee, C.; Oh, J.-W. *PLoS One* **2016**, *11*, e0146710.
- (78) Hermann, T. *Wiley Interdiscip. Rev. RNA* **2016**, *7*, 726–743.
- (79) Velagapudi, S. P.; Gallo, S. M.; Disney, M. D. *Nat. Chem. Biol.* **2014**, *10*, 291–297.
- (80) Sztuba-Solinska, J.; Shenoy, S. R.; Gareiss, P.; Krumpe, L. R. H.; Grice, S. F. J. Le; Keefe, B. R. O.; Schneekloth, J. S.; Le Grice, S. F. J.; O’Keefe, B. R.; Schneekloth, J. S. *J. Am. Chem. Soc.* **2014**, *136*, 8402–8410.
- (81) Bottini, A.; De, S. K.; Wu, B.; Tang, C.; Varani, G.; Pellicchia, M. *Chem. Biol. Drug Des.* **2015**, *86*, 663–673.
- (82) Lee, M. K.; Bottini, A.; Kim, M.; Bardaro, M. F.; Zhang, Z.; Pellicchia, M.; Choi, B.

- S.; Varani, G. *Chem. Commun.* **2014**, *50*, 368–370.
- (83) Su, Z.; Zhang, Y.; Gendron, T. F.; Bauer, P. O.; Chew, J.; Yang, W.; Fostvedt, E.; Jansen-west, K.; Belzil, V. V.; Desaro, P.; Johnston, A.; Overstreet, K.; Boeve, B. F.; Dickson, D.; Floeter, M. K.; Traynor, B. J.; Morelli, C.; Ratti, A.; Silani, V.; Rademakers, R.; Brown, R. H.; Rothstein, J. D.; Boylan, K. B.; Petrucelli, L.; Disney, M. D. *Neuron* **2014**, *83*, 1–8.
- (84) Gallego, J.; Varani, G. *Acc. Chem. Res.* **2001**, *34*, 836–843.
- (85) Warner, K. D.; Hajdin, C. E.; Weeks, K. M. *Nat. Rev. Drug Discov.* **2018**, *17*, 547–558.
- (86) Becker, B.; Cooper, M. A. *ACS Chem. Biol.* **2012**, *8*, 105–115.
- (87) Moazed, D.; Noller, H. F. *Nature* **1987**, *327*, 389–394.
- (88) Wang, Y.; Hamasaki, K.; Rando, R. R. *Biochemistry* **1997**, *36*, 768–779.
- (89) Tran, T. P. A.; Vo, D. D.; Di Giorgio, A.; Duca, M. *Bioorg. Med. Chem.* **2015**, *23*, 5334–5344.
- (90) Sucheck, S. J.; Greenberg, W. A.; Tolbert, T. J.; Wong, C. H. *Angew. Chem. Int. Ed.* **2000**, *39*, 1080–1084.
- (91) Luedtke, N. W.; Baker, T. J.; Goodman, M.; Jolla, L.; July, R. V. *J. Am. Chem. Soc.* **2000**, *122*, 12035–12036.
- (92) Liu, X.; Thomas, J. R.; Hergenrother, P. J. *J. Am. Chem. Soc.* **2004**, *126*, 9196–9197.
- (93) Barluenga, S.; Simonsen, K. B.; Littlefield, E. S.; Ayida, B. K.; Vourloumis, D.; Winters, G. C.; Takahashi, M.; Shandrick, S.; Zhao, Q.; Han, Q.; Hermann, T. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 713–718.
- (94) Liang, F. Sen; Wang, S. K.; Nakatani, T.; Wong, C. H. *Angew. Chem. Int. Ed.* **2004**, *43*, 6496–6500.
- (95) Gooding, K. B.; Higgs, R.; Hodge, B.; Stauffer, E.; Heinz, B.; McKnight, K.; Phipps, K.; Shapiro, M.; Winkler, M.; Ng, W. L.; Julian, R. K. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 884–892.
- (96) Baugh, C.; Shaohui Wang; Bin Li; Appleman, J. R.; Thompson, P. A. *J. Biomol. Screen.* **2009**, *14*, 219–229.
- (97) Wang, W.; Prévile, P.; Morin, N.; Mounir, S.; Cai, W.; Siddiqui, M. A. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1151–1154.
- (98) Jefferson, E. A.; Seth, P. P.; Robinson, D. E.; Winter, D. K.; Miyaji, A.; Osgood, S. A.; Swayze, E. E.; Risen, L. M. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5139–5143.
- (99) Seth, P. P.; Miyaji, A.; Jefferson, E. A.; Sannes-Lowery, K. A.; Osgood, S. A.; Propp, S. S.; Ranken, R.; Massire, C.; Sampath, R.; Ecker, D. J.; Swayze, E. E.; Griffey, R. H. *J. Med. Chem.* **2005**, *48*, 7099–7102.
- (100) Liu, S.; Nelson, C. A.; Xiao, L.; Lu, L.; Seth, P. P.; Davis, D. R.; Hagedorn, C. H.

- Antiviral Res.* **2011**, *89*, 54–63.
- (101) Dibrov, S. M.; Ding, K.; Brunn, N. D.; Parker, M. A.; Bergdahl, B. M.; Wyles, D. L.; Hermann, T. *Proc. Natl. Acad. Sci. U.S.A* **2012**, *109*, 5223–5228.
- (102) Parsons, J.; Castaldi, M. P.; Dutta, S.; Dibrov, S. M.; Wyles, D. L.; Hermann, T. *Nat. Chem. Biol.* **2009**, *5*, 823–825.
- (103) Carnevali, M.; Parsons, J.; Wyles, D. L.; Hermann, T. *ChemBioChem* **2010**, *11*, 1364–1367.
- (104) Ding, K.; Wang, A.; Boerneke, M. A.; Dibrov, S. M.; Hermann, T. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 3113–3117.
- (105) Rynearson, K. D.; Charrette, B.; Gabriel, C.; Moreno, J.; Boerneke, M. A.; Dibrov, S. M.; Hermann, T. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 3521–3525.
- (106) Malnuit, V.; Duca, M.; Benhida, R. *Org. Biomol. Chem.* **2011**, *9*, 326–336.
- (107) Guianvarc'h, D.; Fourrey, J. L.; Maurisse, R.; Sun, J. S.; Benhida, R. *Bioorg. Med. Chem.* **2003**, *11*, 2751–2759.
- (108) Guianvarc'h, D.; Benhida, R.; Fourrey, J. L.; Maurisse, R.; Sun, J. S. *Chem. Commun.* **2001**, *1*, 1814–1815.
- (109) Duca, M.; Malnuit, V.; Barbault, F.; Benhida, R. *Chem. Commun.* **2010**, *46*, 6162–6164.
- (110) Joly, J. P.; Mata, G.; Eldin, P.; Briant, L.; Fontaine-Vive, F.; Duca, M.; Benhida, R. *Chem. Eur. J.* **2014**, *20*, 2071–2079.
- (111) Vo, D. D.; Staedel, C.; Zehnaker, L.; Benhida, R.; Darfeuille, F. *ACS Chem. Biol.* **2014**, *9*, 711–721.
- (112) Vo, D. D.; Tran, T. P. A.; Staedel, C.; Benhida, R.; Darfeuille, F.; Di Giorgio, A.; Duca, M. *Chem. Eur. J.* **2016**, *22*, 5350–5362.
- (113) Wang, Q.; Contag, C. H.; Ilves, H.; Johnston, B. H.; Kaspar, R. L. *Mol. Ther.* **2005**, *12*, 562–568.
- (114) Hanecak, R.; Brown-Driver, V.; Fox, M.; Azad, R.; Furusako, S.; Nozaki, C.; Ford, C.; Sasmor, H.; Anderson, K. *J. Virol.* **1996**, *70*, 5203–5212.
- (115) Klebe, G. *Nat. Rev. Drug Discov.* **2015**, *14*, 95–110.
- (116) Michelson, A. M.; Todd, A. R. *J. Label. Comp. Radiopharm.* **1955**, *2206*, 2632–2638.
- (117) Webb, R. F. **1965**, 3291–3296.
- (118) Gilham, P. T.; Khorana, H. G. *J. Am. Chem. Soc.* **1958**, *80*, 6212–6222.
- (119) Letsinger, R. L.; Ogilvie, K. K. *J. Am. Chem. Soc.* **1969**, *91*, 3350–3355.
- (120) Letsinger, R. L.; Finnan, J. L.; Heavner, G. A.; Lunsford, W. B. *J. Am. Chem. Soc.* **1975**, *97*, 3278–3279.

- (121) Matteucci, M. D.; Caruthers, M. H. *J. Am. Chem. Soc.* **1981**, *103*, 3185–3191.
- (122) Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett.* **1981**, *22*, 1859–1862.
- (123) Eritja, R.; Pons, A.; Escarcellar, M.; Giralt, E.; Albericio, F. *Tetrahedron* **1991**, *47*, 4113–4120.
- (124) Moulton, H. M.; Nelson, M. H.; Hatlevig, S. A.; Reddy, M. T.; Iversen, P. L. *Bioconjug. Chem.* **2004**, *15*, 290–299.
- (125) Kitamura, Y.; Asakura, R.; Terazawa, K.; Shibata, A.; Ikeda, M.; Kitade, Y. *Bioorg. Med. Chem. Lett.* **2017**, *27*, 2655–2658.
- (126) Panattoni, A.; Pohl, R.; Hocek, M. *Org. Lett.* **2018**, *20*, 3962–3965.
- (127) Zewge, D.; Butora, G.; Sherer, E. C.; Tellers, D. M.; Sidler, D. R.; Gouker, J.; Copeland, G.; Jadhav, V.; Li, Z.; Armstrong, J.; Davies, I. W. *Bioconjug. Chem.* **2018**, *29*, 1859–1865.
- (128) Berndl, S.; Herzig, N.; Kele, P.; Lachmann, D.; Li, X.; Wolfbeis, O. S.; Wagenknecht, H. A. *Bioconjug. Chem.* **2009**, *20*, 558–564.
- (129) Rao, H.; Sawant, A. A.; Tanpure, A. A.; Srivatsan, S. G. *Chem. Commun.* **2012**, *48*, 498–500.
- (130) Shang, S.; Monfregola, L.; Caruthers, M. H. *Signal Transduct. Target. Ther.* **2016**, *1*, 16019.
- (131) El-Sagheer, A. H.; Brown, T. *Chem. Soc. Rev.* **2010**, *39*, 1388–1405.
- (132) Dadová, J.; Orság, P.; Pohl, R.; Brázdová, M.; Fojta, M.; Hocek, M. *Angew. Chem. Int. Ed.* **2013**, *52*, 10515–10518.
- (133) Alam, M. R.; Ming, X.; Nakagawa, O.; Jin, J.; Juliano, R. L. *Bioorg. Med. Chem.* **2013**, *21*, 6217–6223.
- (134) Meyer, A.; Vasseur, J. J.; Morvan, F. *Eur. J. Org. Chem.* **2013**, *3*, 465–473.
- (135) Schoch, J.; Wiessler, M.; Jäschke, A. *J. Am. Chem. Soc.* **2010**, *132*, 8846–8847.
- (136) Gutmiedl, K.; Wirges, C. T.; Ehmke, V.; Carell, T. *Org. Lett.* **2009**, *60*, 2596–2599.
- (137) Leszczynska, G.; Sadowska, K.; Sierant, M.; Sobczak, M.; Nawrot, B.; Sochacka, E. *Org. Biomol. Chem.* **2017**, *15*, 5332–5336.
- (138) Arndt, S.; Wagenknecht, H. A. *Angew. Chem. Int. Ed.* **2014**, *53*, 14580–14582.
- (139) Shigdel, U. K.; Zhang, J.; He, C. *Angew. Chem. Int. Ed.* **2008**, *47*, 90–93.
- (140) Lercher, L.; McGouran, J. F.; Kessler, B. M.; Schofield, C. J.; Davis, B. G. *Angew. Chem. Int. Ed.* **2013**, *52*, 10553–10558.
- (141) Gutmiedl, K.; Fazio, D.; Carell, T. *Chem. - A Eur. J.* **2010**, *16* (23), 6877–6883.
- (142) Hari, Y.; Nakahara, M.; Pang, J.; Akabane, M.; Kuboyama, T.; Obika, S. *Bioorg. Med. Chem.* **2011**, *19*, 1162–1166.

- (143) Nakahara, M.; Kuboyama, T.; Izawa, A.; Hari, Y.; Imanishi, T.; Obika, S. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 3316–3319.
- (144) Saikia, P.; Gogoi, S. *Adv. Synth. Catal.* **2018**, *360*, 2063–2075.
- (145) Aikawa, H.; Kaneko, T.; Asao, N.; Yamamoto, Y. *Beilstein J. Org. Chem.* **2011**, *7*, 648–652.
- (146) Asao, N.; Aikawa, H.; Tago, S.; Umetsu, K. *Org. Lett.* **2007**, *9*, 4299–4302.
- (147) Ding, H.; Li, C.; Zhou, Y.; Hong, S.; Zhang, N.; Xiao, Q. *RSC Adv.* **2017**, *7*, 1814–1817.
- (148) Zhang, Q.; Sun, J.; Zhu, Y.; Zhang, F.; Yu, B. *Angew. Chem. Int. Ed.* **2011**, *50*, 4933–4936.
- (149) Yu, B. *Acc. Chem. Res.* **2018**, *51*, 507–516.
- (150) Li, W.; Yu, B. *Chem. Soc. Rev.* **2018**, Advance Article.
- (151) Zhu, Y.; Shen, Z.; Li, W.; Yu, B. *Org. Biomol. Chem.* **2016**, *14*, 1536–1539.
- (152) Li, Y.; Yang, Y.; Yu, B. *Tetrahedron Lett.* **2008**, *49*, 3604–3608.

General Conclusion

This manuscript highlighted part of the work that was realized in the group of *Molecules Bioactives* along the last three years, with focus on two main fields of research.

Concerning the Part I of this work, in **Chapter 1** was presented the state the art of photophysical processes and fluorescence sensing in order to introduce the work described in **Chapters 2** and **3**. **Chapter 2** highlighted the design and development of two libraries of novel push-pull fluorophores derived from 6-aminobenzothiazoles. Initially, a library of 20 styryl-based push-pull dyes (**85-104**) were prepared by using an efficient and practical synthetic route. A series of aryl scaffolds, from substituted phenyl rings containing electron donating and withdrawing groups to polycyclic aromatic derivatives, was surveyed. We also explored and evaluated the inductive effect of *N*-alkyl substituted benzothiazoles using three different arrangements (**Figure C1**). Moreover, we could study the photophysical properties of all dyes using UV-vis absorption and steady-state fluorescence emission measurements in solution. Complementary information were obtained thanks to the TD-DFT calculations. This study highlighted a clear structure-photophysics relationship for this class of fluorophores. The dyes presented high brightness, absorption bands in the visible range (~ 370 - 453 nm) and large solvatochromism comprising all the visible spectrum, as a consequence of the strong intramolecular charge transfer (ICT) nature of their excited state.

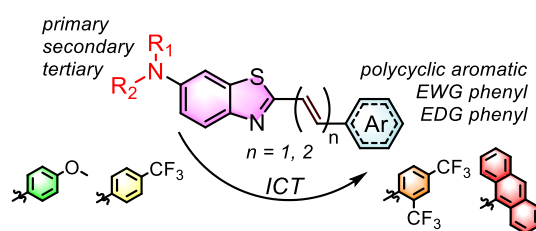


Figure C1. Styryl-based push-pull fluorophores.

In the second section, a series of fluorophores bearing an extended π -conjugated scaffold (**110-116**) could be successfully accessed, *via* Sonogashira cross coupling reactions, in reasonable yields (**Figure C2**). Compared with the first series of styryl fluorophores, these dyes were characterized by two intense absorption bands, one located in the near-UV and the second in the visible spectral region and a single, water sensitive and highly solvatochromic emission band. The photophysical studies also highlighted that dyes bearing π -electron acceptor groups were quenched in high polar environment, suggesting the formation of TICT

states. In addition and in analogy with the styryl-based series, all the other dyes were highly emissive in both polar and non-polar solvents. Moreover, these dyes were evaluated for *in cellulo* applications and investigations concerning their cellular localization are currently being assessed; preliminary co-localization studies had already shown that these dyes tends to accumulate inside lipid droplets within the cell. Importantly, while the noticed fluorescence quenching in water media greatly accounts for the very low background, intense fluorescence emission coming from the lipidic compartments was observed. Further studies are currently being performed to better investigate this localization.

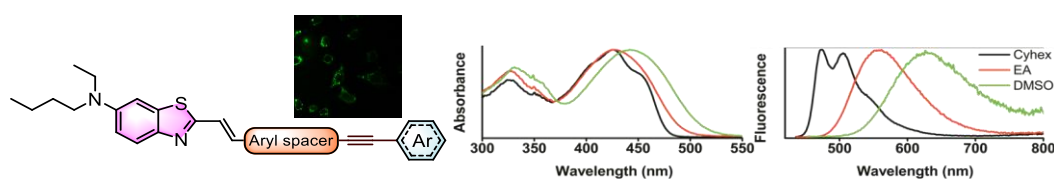


Figure C2. Benzothiazole-based push-pull fluorophores with extended π -conjugated system

In **Chapter 3** we described the development of four novel fluorescent probes (**121-124**) to monitor the activity of the β -galactosidase enzyme (β -gal), *in vitro* and in living cells. Initially, the synthesis and the screening of four different fluorophores displaying distinct photophysical properties enabled us to access new light-up and two-band ratiometric reporters. The four probes exhibited an extremely fast response and over 200-fold fluorescence enhancement (**121**, **Figure C3**) following the enzymatic cleavage of the β -D-galactopyranoside unit. This rapid and extremely sensitive response allowed the detection of senescence-associated β -galactosidase (SA- β -gal) activity; a widely used biomarker of cellular senescence. More importantly, **121** also enabled us to monitor, in real-time, the emergence of senescence in *live cells*, *i.e.* the phenotypic transformation from normal to senescent cell. These findings underpin the fact that **121** may find useful applications in biomedical research. Importantly, **121** is suitable for epifluorescence and confocal microscopies, and flow cytometry techniques, which are among the most common analytical tools in biology. The use of probe **121** in model senescent tissues as well as in tumors are envisaged in the near future. In addition, this initial study focusing on the turn-ON probe **121** can be further extended; notably, the use of **123**, for the ratiometric visualization of SA- β -gal, will be assessed.

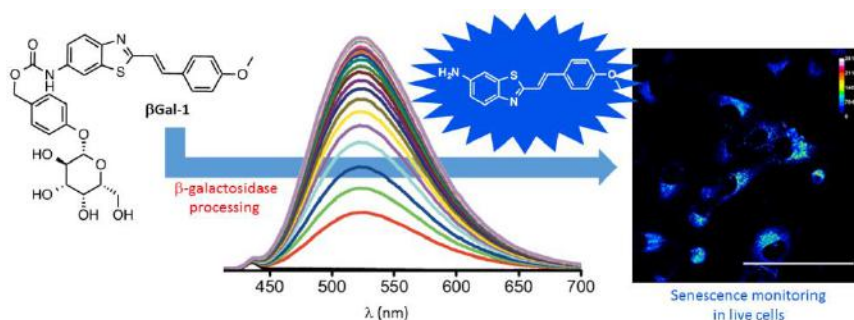


Figure C3. β -gal probes to monitor SA- β -gal activity in live cells.

In the Part II of this manuscript, we presented and described our work dealing with different aspects of nucleic acid chemistry. In fact, in **Chapter 4** we briefly described the major features concerning RNAs structures and associated them with the prospective use of RNAs as potential druggable targets (either with small-sized ligands or oligonucleotides). Thereafter, we reported the synthesis of two new artificial nucleobase scaffolds, featuring adequate hydrogen bonding donors and acceptors for the molecular recognition of U:A and C:G base pairs, respectively, using a triplex-based rational design strategy. The tethering of these structures to various amino acids and the assessment of these artificial nucleobase aminoacid conjugates as RNA ligands against a model of HCV IRES III_d domain were successfully reported. It is noteworthy that in this work we disclose the first structures of small-molecule ligands of IRES III_d loop. In addition, low micromolar dissociation constant could be obtained with a twofold lower value for our best ligand (**Figure C4**) compared with neomycin. Finally, good selectivity to the HCV IRES III_d loop could also be ascertained by competition studies in presence of large excess of natural tRNA. Thus, this study validates our original and rational design strategy and can now be used to develop new and more powerful series of RNA ligands.

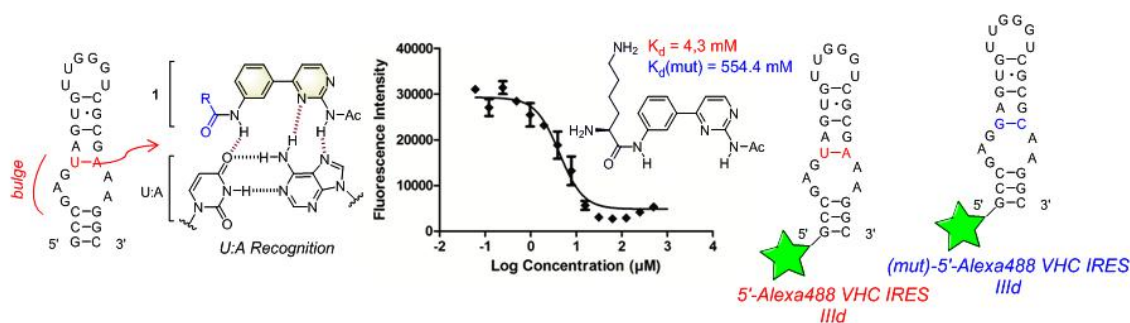


Figure C4. RNA ligands to target HCV IRES III_d loop.

Finally, in **Chapter 5**, we disclose our efforts towards the development of the first methodology applied to the post-synthetic functionalization of oligonucleotides directly at the anomeric position of deoxyribose platforms. We could demonstrate the validity of our method, which could be extremely beneficial and a powerful synthetic route to introduce a great variety of nucleophiles using a classic Fisher glycosidation reaction.

More precisely, following a preliminary screening to evaluate and select suitable synthons to be incorporated on DNA, we could identify two introduce 2'-deoxyribose platforms bearing an activatable benzoyl unit connected to their anomeric position. One of these platform was successfully introduced into oligonucleotides sequences *via* the classical phosphoramidite elongation chemistry. Good compatibility of our phosphoramidite building block with the oligonucleotide synthesis was observed; great stability to all cycles of trityl deprotection and high coupling yields (> 95%) were achieved. Currently, the screening of multiple reaction conditions to enable the post-synthetic functionalization of oligonucleotide harbouring this platform is ongoing. Nevertheless, we presented preliminary results that ascertained the feasibility of our strategy. Notably, the first PSF of an oligonucleotide at one of its anomeric position could be performed.

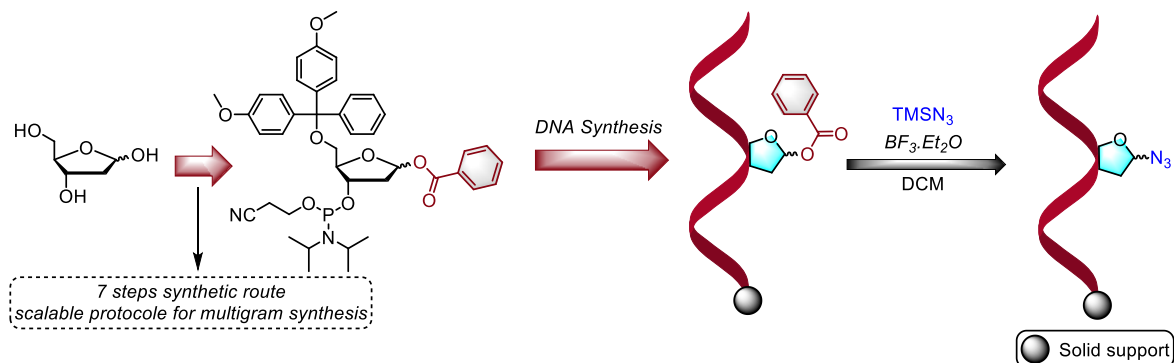


Figure C5. Post-synthetic functionalization of oligonucleotides.

This work leads to three publications in peer-reviewed journals and presentations as part of international conferences. These works are listed below.

Bibliographic Production:

- **SAFIR FILHO, M.**; DAO, P.; GESSON, M.; MARTIN, A. R.; BENHIDA, R. Development of highly sensitive fluorescent probes for the detection of β -galactosidase activity – application to the real-time monitoring of senescence in live cells. *Analyst*, **2018**, 143, 2680-2688.
- **SAFIR FILHO, M.**; FIORUCCI, S.; MARTIN, A. R.; BENHIDA, R. Design, synthesis and photophysical studies of styryl-based push–pull fluorophores with remarkable solvatochromism. *New. J. Chem.*, **2017**, 41, 13760-13772.
- **SAFIR FILHO, M.**; MARTIN, A. R.; BENHIDA, R. Assessment of new triplet forming artificial nucleobases as RNA ligands directed towards HCV IRES III_d loop. *Bioorganic & Medicinal Chemistry Letters*, **2017**, 27, 1780-1783.

Contributed Presentation & Abstracts in Conferences:

- **SAFIR FILHO, M.**; DAO, P.; MARTIN, A. R.; BENHIDA, R. Benzothiazole-Based Styryl Dyes: Spectroscopic Studies and Senescent Cells Detection. *25th Young Research Fellow Meeting*, March/2018, Orléans, France. (Oral Presentation).
- **SAFIR FILHO, M.**; MARTIN, A. R.; BENHIDA, R. Rational design and synthesis of innovative RNA ligands to target HCV internal ribosome entry site. *2nd International Conference on Pharmaceutical Chemistry*, October/2017, Barcelona, Spain. (Oral Presentation)
- **SAFIR FILHO, M.**; MARTIN, A. R.; BENHIDA, R. Rational design of innovative RNA ligands to target HCV internal ribosome entry site. *5èmes Journées Méditerranéennes des Jeunes Chercheurs*, October/2017, Montpellier, France. (Oral Presentation)
- **SAFIR FILHO, M.**; MARTIN, A. R.; BENHIDA, R. Rational design and synthesis of new triplet forming artificial nucleobases as RNA ligands to target HCV IRES III_d loop. *Journée ICN-Galderma*, March/2017, Nice, France. (Oral Presentation)
- **SAFIR FILHO, M.**; MARTIN, A. R.; BENHIDA, R. Rational Design and Synthesis of New RNA Ligands to Target Oncogenic RNA. *8ème Colloque du Cancéropôle PACA*, March/2017, Marseille, France. (Poster)

- **SAFIR FILHO, M.**; MARTIN, A. R.; BENHIDA, R. Design and Synthesis of New RNA ligands to Target HCV Internal Ribosome Entry Site. *4èmes Journées Méditerranéennes des Jeunes Chercheurs*, October/2016, Nice, France. (Poster)

Other Events:

- Summer School: *Modern Physicochemical Methods in Chemistry and Pharmacy (147 h)*, July/2017, Moscow, Russia, RUDN University.

Experimental Section

General Methods

All chemical reagents were obtained from commercial sources (Aldrich, Across, Alfa Aesar) and were used as supplied. No further purification were performed

All the chemical reactions were monitored simultaneously by LCMS and thin-layer chromatography (TLC, Merck silica gel 60 F254 plates) and visualized both by UV radiation (254 & 365 nm) and by spraying with a relevant staining agent followed by a subsequent warming with a heat gun. Column chromatography was performed with flash silica gel (40–63 mm). All NMR spectra (^1H , ^{13}C , ^{19}F or ^{31}P) were recorded on Bruker Advance Spectrometers (200 or 400 MHz). ^1H NMR (200 and 400 MHz), $^{13}\text{C}\{^1\text{H}\}$ NMR (50 and 101 MHz), $^{19}\text{F}\{^1\text{H}\}$ (188 and 377 MHz) and $\text{P}\{^1\text{H}\}$ (162 MHz) spectra were obtained with samples dissolved in CDCl_3 , $\text{DMSO-}d_6$, $\text{CD}_2\text{Cl}_2-d_2$, $\text{CD}_3\text{OD-}d_4$ and D_2O with the solvent residual signals as internal references: CHCl_3 ($^1\text{H} = 7.26$ ppm, $^{13}\text{C} = 77.16$), $\text{DMSO-}d_6$ ($^1\text{H} = 2.50$ ppm, $^{13}\text{C} = 39.52$), $\text{CD}_2\text{Cl}_2-d_2$ ($^1\text{H} = 5.32$ ppm, $^{13}\text{C} = 53.84$), $\text{CD}_3\text{OD-}d_4$ ($^1\text{H} = 3.31$ ppm, $^{13}\text{C} = 49.00$), and D_2O ($^1\text{H} = 4.79$ ppm). Chemical shifts (δ) are given in ppm to the nearest 0.01 (^1H) or 0.1 ppm (^{13}C). The coupling constants (J) are given in Hertz (Hz). The signals are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, m = multiplet, dd = doublet of doublets, br s = broad singlet), coupling constants (J) and integration. Dye syntheses were performed in a Branson 5510 ultrasound bath. Low resolution mass spectra (MS) were recorded on an Esquire 3000 Plus apparatus with ESI in both positive and negative modes. High Resolution Mass Spectra (HRMS) were recorded on a ThermoFisher Q Exactive (ESIMS) at a resolution of 140 000 at m/z 200. Density functional theory (DFT) and time-dependent DFT (TD-DFT) calculations were performed using the Gaussian0356 software at the MPW1PW91/6-311+G(d,p) level of theory for all atoms without any symmetry restrictions. Optimized structures of the ground state were checked via a calculation of vibrational frequencies, ensuring that no imaginary frequencies were present. The vertical excitation energies are determined using the conventional TD-DFT procedure on the ground state geometries. The effects of solvents were taken into account by the polarizable continuum model (PCM).

Chapter 2

All solvents for absorption and fluorescence experiments were of spectroscopic grade. Absorption spectra were recorded on a Cary 4 spectrophotometer (Varian) using quartz cells of 1 cm path length. Fluorescence spectra were recorded on FluoroMax 4.0 spectrofluorometer (Jobin Yvon, Horiba). The wavelengths corresponding to the absorption maxima were used as the excitation wavelengths. Otherwise, the respective excitation wavelength is properly indicated. Stock solutions of the dyes were prepared using 1,4-dioxane (all dyes of **Chapter 2**). The samples used for spectroscopic measurements contained 0.1% or less v/v of the stock solvent. UV-vis absorption spectra were recorded using a dye concentration of *ca.* 10^{-5} M. Fluorescence emission spectra were recorded using a dye concentration of *ca.* $\sim 10^{-6}$ M ($\sim A \leq 0.05$ for quantum yield analysis).

Chapter 3

β -gal studies

Stock solutions of the dyes were prepared using 1,4-dioxane or DMSO (β -gal probes **121-124**). The samples used for spectroscopic measurements contained 0.1% or less v/v of the stock solvent. Fluorescence emission spectra in presence of β -galactosidase were recorded in PBS buffer solutions, at 37°C, with probe concentration ≈ 0.2 μ M. Fluorescence emission spectra in cells were recorded on a synergy H1 microplate reader (Biotek) with $\lambda_{\text{exc}} = 405$ nm as excitation wavelength.

Cell experiments

Human melanoma A375 cells and human colorectal adenocarcinoma HT-29 cells were obtained from American Type Culture Collection (ATCC, Manassas, VA, USA). A375 and HT-29 cells were cultured in Dubelcco's Modified Eagle's Medium (DMEM) containing

10% fetal bovine serum (FSB) and 5% penicillin/streptomycin at 37°C in a 5% CO₂ incubator.

Live cells and fixed cells fluorescence imaging

For fixed imaging, glass cover slips were placed into 6-well plate and A375 and HT-29 were plated at 8×10⁴ cells/well and incubated at 37°C in a 5% CO₂ incubator overnight. On the next day, the cells were incubated with 50 μM hydroxyurea (HU) to induce senescence. Cells without any treatment were used as negative controls. After 96h, cells were washed with PBS before incubated 30 min with 0.1 μM **121** at 37°C. Samples were washed with PBS and fixed with paraformaldehyde 4% before the coverslips were mounted onto a microscope slide with Fluoromount aqueous mounting medium (Sigma Aldrich) before fluorescence imaging.

For live cells imaging, A375 cells were plated on a 4-well μ-slides (Ibidi) at 4×10⁴ cells/well and incubated at 37°C in a 5% CO₂ incubator overnight. On the next day, 0.1 μM **βGal-1** was added before the induction of senescence using 50 μM HU, then fluorescence imaging was performed every 24h.

Confocal fluorescent imaging was performed using a confocal microscope Nikon A1R scan head on a Nikon Eclipse Ti stand, using a Plan Apo 63X oil 1.4 NA objective. The excitation channel used was a LASER diode 405 nm and the emission signal was collected with an internal PMT detector equipped with a 450/50 filter set, and the transmission signal was collected on an external PMT. Fluorescence intensity quantification was done using Fiji.

Flow cytometry analysis

A375 cells were plated on 12-well plates at 8×10⁴ cells/well and incubated at 37°C in a 5% CO₂ incubator overnight. On the next day, the cells were treated with HU to induce senescence. After 72h, trypsinized cells were washed with PBS 2% BSA and incubated 30 min with 0.1 μM **121** at 37°C. After washes, cells were analyzed by FACS on a MACSQuant VYB (Myltenyi).

Chapter 4

Binding assay with HCV IRES III_d loop

The buffer solution was prepared in Milli-Q water and filtered through 0.22 mm Millipore filters (GP ExpressPLUS membrane). A small aliquot (100 mL) was first filtered and then discarded to avoid any contaminants that might be leached from the filter. All standard fluorescence measurements were performed in buffer I (20 mM HEPES, pH 7.4, containing 20 mM NaCl, 140 mM KCl, and 3 mM MgCl₂, at 25°C).

Ligand solutions were prepared as serial dilutions in buffer I at a concentration four times higher than the desired final concentration to allow for the subsequent dilution before the addition of the RNA solution. An automated pipetting system (epMotion 5075, Eppendorf) was used in order to perform these binding studies on 384-well plates (Greiner bio-one). Refolding of the RNA was performed by using a thermocycler (ThermoStat Plus, Eppendorf) as follows: the 5'-Alexa488 HCV IRES III_d loop RNA (5'-Alexa-GCCGAGUAGUGUUGGGUCGCGAAAGGC-3', 40 nmol) was diluted with buffer I to a concentration of 200 nM (1mL), denatured by heating to 90°C for 2 min, cooled to 4°C for 10 min, then incubated at 25°C for 15 min. After refolding, the RNA was diluted to a working concentration of 20 nM through addition of the appropriate amount of buffer I. After addition of each ligand (30 µL) on the 384-well plates in 15 dilutions (from 1 mM to 61 nM as final concentrations) and in duplicate, the RNA solution (30 µL) was added to each well containing ligand to give a final concentration of 10 nM. The fluorescence was measured on a GeniosPro apparatus (Tecan) with an excitation filter of (485±10) nm and an emission filter of (535±15) nm. Each point was measured five times with a 500 µs integration time and averaged. Binding was allowed to proceed overnight at 5°C to achieve equilibrium. To study the temperature dependence, the plates were incubated after overnight equilibrium at different temperatures ranging from 5°C to 36°C. Neomycin was used as a control because its ability to bind RNA has already studied using several methods.

Data analysis

Binding data were analyzed using Graphpad Prism 5 software. Unless otherwise stated, binding profiles were well modeled by using a simple model and assuming one-to-one stoichiometry.

Biding affinity and selectivity studies

Biding affinity and selectivity studies were performed by standard fluorescence measurements using the same procedure reported above. For the selectivity analysis, the fluorescent HCV IRES III_d RNA model was mixed with 50 equivalents of natural tRNA and was titrated with the ligands **1c** and **1e** at 20°C. For the binding site studies, a mutated fluorescent HCV IRES III_d RNA (5'Alexa-GCCGAGGAGUGUUGGGUCGCGCAAGGC-3') was used and titrated with the respective ligands. Stock solutions of tRNA were prepared in water and quantified using an extinction coefficient of 9640 cm⁻¹ M⁻¹ per base.

Chapter 5

Oligonucleotide synthesis

DNA synthesis was carried out on an K&A DNA/RNA synthesizer H-8 using standard phosphoramidite chemistry. Reagents and controlled pore glass (DMTr-dT-CPG, succinyl linker, 1 μmol) were purchased from ChemGenes Corporation and Link Technologies. The DNA strands were synthesized using a standard protocol with PAC₂O as the capping reagent. The CPG vial was flushed with acetonitrile at the end of the oligonucleotide synthesis. The supported nucleic acids were dried and used for subsequent modifications without any purification or stored at -20°C.

The CPG-supported oligonucleotides were analyzed by MALDI-ToF mass spectrometry using 2',4',6'-trihydroxyacetophenone (THAP) as matrix. The THAP stock was prepared as a saturated solution in water:acetonitrile 1:1 and mixed with the supported oligonucleotides on the MALDI-ToF plate. MALDI-ToF analysis was performed on a Shimadzu AXIMA Assurance Linear MALDI-TOF Mass Spectrometer using negative mode.

Post-synthetic functionalization

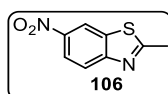
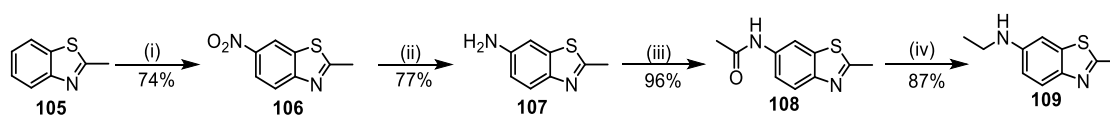
The supported oligonucleotides **ODN1** or **ODN2** were placed in a DNA synthesis column. To this system, solutions of TMSN₃ (0.1 M) and BF₃.OEt₂ (0.1 M) were pumped together inside

the column. The reaction was kept during 1 hour, at room temperature. Finally, the columns were flushed with acetonitrile, dried and analyzed by MALDI-ToF mass spectrometry.

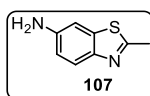
Synthetic Procedures

Chapter 2

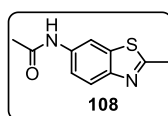
Synthesis of benzothiazole intermediates **107** and **109**



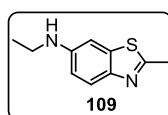
(i) In a 250 mL round-bottom flask were added 29.8 g (200 mmol) of 2-methylbenzothiazole and 100 mL of concentrated sulfuric acid. The mixture was cooled to 0°C and a solution of 17.0 g (200 mmol) of sodium nitrate in 100 mL of concentrated sulfuric acid was slowly added while keeping the temperature of the reaction media below 5 °C. After completion of the addition, the reaction was further stirred for 2 hours at 0°C before pouring the mixture into 1L of crushed ice. The precipitate formed was filtered off, and washed with cold water and ethanol. It was subsequently recrystallized from methanol to afford 28.7 g of light yellow solid. Yield: 74 %. ¹H NMR (200 MHz, CDCl₃) δ (ppm) 8.73 (d, *J* = 2.3 Hz, 1H), 8.29 (dd, *J* = 9.0, 2.3 Hz, 1H), 7.99 (d, *J* = 9.0 Hz, 1H), 2.89 (s, 3H). ¹³C NMR (50 MHz, CDCl₃) δ(ppm) 173.4, 157.2, 144.8, 136.1, 122.7, 121.6, 118.1, 20.8.



(ii) To a solution of 19.4 g (100 mmol) of **106** in 200 mL of methanol, were successively added 6.5 g of zinc powder (100 mmol) and 25.4 g (400 mmol) of ammonium formate. The reaction was stirred overnight at room temperature. Then, the crude mixture was filtered over celite, the filtrate was diluted with ethyl acetate and extracted with water (3 × 200 mL). The organic phase was concentrated in *vacuo* and the crude residue was purified by flash chromatography on silica gel, using a 3/7 (v/v) mixture of cyclohexane/ethyl acetate as the eluent, to afford 12.6 g of compound **107** as a grey powder. Yield: 77 %. ¹H NMR (200 MHz, CDCl₃) δ (ppm) 7.68 (d, *J* = 8.6 Hz, 1H), 7.02 (d, *J* = 2.1 Hz, 1H), 6.76 (dd, *J* = 8.6, 2.3 Hz, 1H), 3.80 (br s, 2H), 2.73 (s, 3H). ¹³C NMR (50 MHz, CDCl₃) δ (ppm) 162.7, 146.8, 144.1, 137.3, 122.7, 115.2, 105.9, 19.9.

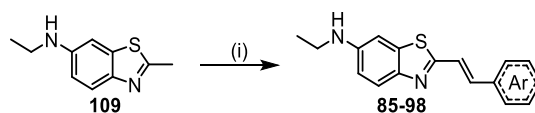


(iii) In a 100 ml round-bottom flask, 4.9 g (30 mmol) of **103**, 6.1 g (60 mmol) of triethylamine and 0.3 g (3 mmol) of DMAP were dissolved in 30 mL of anhydrous DMF. The mixture was cooled to 0°C in an ice bath before the dropwise addition of acetyl chloride (2.35 mL, 33 mmol). After completion of the addition, the ice bath was removed and the mixture was stirred at r.t. for 1 hour. The solvent was removed under reduced pressure and the residue was dissolved in 100 mL of ethyl acetate, washed with water (3×100 mL) and brine (100 mL). The organic layer was dried over magnesium sulfate, filtered and the volatiles were removed in *vacuo*. Finally, the crude residue was purified by flash chromatography on silica gel, using cyclohexane/ethyl acetate (6/4, v/v) as the eluent, to afford 5,92 g of pure compound **104** as light yellow solid. Yield: 96 %. Mp: 139-140 °C. ¹H NMR (400 MHz, CDCl₃) δ (ppm) 8.39 (s, 1H), 8.08 (s, 1H), 7.82 (d, *J* = 8.7 Hz, 1H), 7.27 (d, *J* = 8.7 Hz, 1H), 2.81 (s, 3H), 2.21 (s, 3H). ¹³C NMR (50 MHz, CDCl₃) δ (ppm) 168.9, 166.6, 150.1, 136.6, 135.1, 122.3, 118.8, 112.6, 24.6, 20.1. HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₁₀H₁₁N₂OS 207.0587; Found 207.0588.

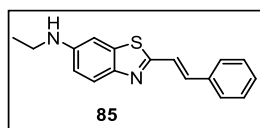


(iv) A 250 mL round-bottom flask was charged with 5.0 g (24 mmol) of **104**, and 100 mL of dry THF, under an inert atmosphere (Ar). To this mixture were slowly added 3.1 g (4 equiv.) of LiAlH₄ and the reaction was subsequently stirred overnight. The unreacted LiAlH₄ was quenched via the dropwise addition of acetone. The crude was neutralized with 1 M HCl and filtered over celite. The filtrate was concentrated to dryness, dissolved in 150 mL of ethyl acetate and washed with brine (3×150 mL). The organic layer was dried over magnesium sulfate, filtered and the volatiles were removed in *vacuo*. The crude was purified by flash chromatography on silica gel, using cyclohexane/ethyl acetate (8/2, v/v) as the eluent, to provide 4.0 g of precursor **105** as a grey solid. Yield: 87%. Mp: 107-108 °C. ¹H NMR (200 MHz, CDCl₃) δ (ppm) 7.68 (d, *J* = 8.7 Hz, 1H), 6.90 (d, *J* = 2.3 Hz, 1H), 6.69 (dd, *J* = 8.7, 2.4 Hz, 1H), 3.69 (br s, 1H), 3.15 (q, *J* = 7.1 Hz, 2H), 2.72 (s, 3H), 1.25 (t, *J* = 7.1 Hz, 3H). ¹³C NMR (50 MHz, CDCl₃) δ (ppm) 161.7, 146.3, 145.8, 137.7, 122.6, 113.9, 102.2, 38.8, 19.8, 14.8. HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₁₀H₁₃N₂S 193.0794; Found 193.0796.

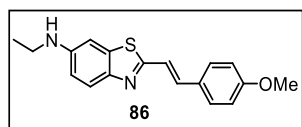
General procedure for the synthesis of dyes **85-98**



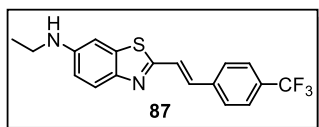
(i) Precursor **109** (1 equiv.) in DMF (0.2 M), potassium hydroxide (1 equiv.) and aldehyde (1,3 equiv.) were subjected to ultrasonic irradiation for 1 hour at room temperature. The solvent was removed in *vacuo*, and the crude was purified by flash column chromatography on silica gel using cyclohexane/ethyl acetate mixtures as the eluents. Finally, recrystallization from methanol afforded the desired dyes as bright solids.



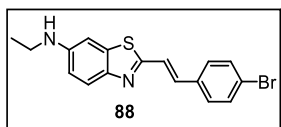
The crude product was purified using cyclohexane/ethyl acetate (8/2, v/v) as the eluent. Yield 69 %. Mp: 119-120 °C. ¹H NMR (200 MHz, CDCl₃) δ (ppm) 7.75 (d, *J* = 8.8 Hz, 1H), 7.55 (dd, *J* = 8.0, 1.5 Hz, 2H), 7.45 – 7.30 (m, 5H), 6.95 (d, *J* = 2.3 Hz, 1H), 6.74 (dd, *J* = 8.8, 2.4 Hz, 1H), 3.68 (br s, 1H), 3.21 (q, *J* = 7.1 Hz, 2H), 1.30 (t, *J* = 7.1 Hz, 3H). ¹³C NMR (50 MHz, CDCl₃) δ (ppm) 162.2, 146.9, 146.4, 136.8, 135.9, 135.5, 129.0, 127.2, 123.5, 122.7, 114.7, 101.9, 38.8, 14.9. HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₁₇H₁₇N₂S 281.1107; Found 281.1108.



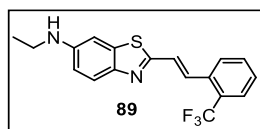
The crude product was purified using cyclohexane/ethyl acetate (75/25, v/v) as the eluent. Yield 78 %. Mp: 134-135 °C. ¹H NMR (400 MHz, CDCl₃) δ (ppm) 7.73 (d, *J* = 8.8 Hz, 1H), 7.48 (d, *J* = 8.7 Hz, 2H), 7.30 (d, *J* = 16.2 Hz, 1H), 7.22 (d, *J* = 16.2 Hz, 1H), 6.94 (d, *J* = 2.3 Hz, 1H), 6.91 (d, *J* = 8.8 Hz, 2H), 6.73 (dd, *J* = 8.8, 2.3 Hz, 1H), 3.83 (s, 4H), 3.20 (q, *J* = 7.1 Hz, 2H), 1.29 (t, *J* = 7.1 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ (ppm) 162.6, 160.4, 146.8, 146.5, 136.6, 135.2, 128.7, 128.6, 123.3, 120.6, 114.5, 114.4, 102.0, 55.5, 38.8, 14.9. HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₁₈H₁₉N₂OS 311.1213; Found 311.1213.



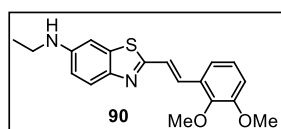
The crude product was purified using cyclohexane/ethyl acetate (8/2, v/v) as the eluent. Yield 48 %. Mp: 140-141 °C. ¹H NMR (200 MHz, CDCl₃) δ (ppm) 7.76 (d, *J* = 8.8 Hz, 1H), 7.63 (s, 4H), 7.44 (d, *J* = 16.3 Hz, 1H), 7.32 (d, *J* = 16.3 Hz, 1H), 6.96 (d, *J* = 2.3 Hz, 1H), 6.77 (dd, *J* = 8.8, 2.3 Hz, 1H), 3.85 (br s, 1H), 3.23 (q, *J* = 7.1 Hz, 2H), 1.31 (t, *J* = 7.1 Hz, 3H). ¹³C NMR (50 MHz, CDCl₃) δ (ppm) 161.2, 147.0, 146.5, 139.4, 137.1, 133.4, 130.4 (q, *J* = 32.6 Hz), 127.3, 125.9 (q, *J* = 3.8 Hz), 125.0, 123.8, 115.0, 101.9, 38.9, 14.8. HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₁₈H₁₆N₂F₃S 349.0981; Found 349.0983.



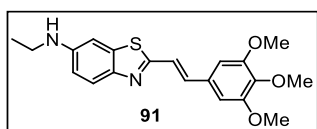
The crude product was purified using cyclohexane/ethyl acetate (8/2, v/v) as the eluent. Yield 82 %. Mp: 178-179 °C. ¹H NMR (400 MHz, CDCl₃) δ (ppm) 7.75 (d, *J* = 8.8 Hz, 1H), 7.51 (d, *J* = 7.2 Hz, 2H), 7.40 (d, *J* = 7.2 Hz, 2H), 7.33 (d, *J* = 16.1 Hz, 1H), 7.26 (d, *J* = 16.2 Hz, 1H), 6.96 (s, 1H), 6.75 (d, *J* = 8.8 Hz, 1H), 3.96 (br s, 1H), 3.22 (q, *J* = 7.0 Hz, 2H), 1.31 (t, *J* = 7.0 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ (ppm) 161.7, 146.9, 146.5, 136.9, 134.9, 133.9, 132.2, 128.6, 123.6, 123.4, 122.9, 114.8, 101.9, 38.9, 14.9. HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₁₇H₁₆N₂BrS 359.0212; Found 359.0219.



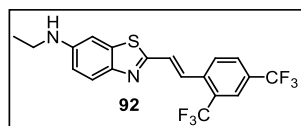
The crude product was purified using cyclohexane/ethyl acetate (8/2, v/v) as the eluent. Yield 74 %. Mp: 129-130 °C. ¹H NMR (400 MHz, CDCl₃) δ (ppm) 7.79 (d, *J* = 7.9 Hz, 1H), 7.77 (d, *J* = 8.8 Hz, 1H), 7.70 – 7.64 (m, 2H), 7.57 (t, *J* = 7.7 Hz, 1H), 7.41 (t, *J* = 7.7 Hz, 1H), 7.36 (d, *J* = 16.0 Hz, 1H), 6.96 (d, *J* = 2.3 Hz, 1H), 6.76 (dd, *J* = 8.8, 2.4 Hz, 1H), 3.99 (br s, 1H), 3.22 (q, *J* = 7.1 Hz, 2H), 1.31 (t, *J* = 7.1 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ (ppm) 161.5, 147.1, 146.6, 137.1, 134.9, 132.2, 130.8 (q, *J* = 1.9 Hz), 128.5, 128.1 (q, *J* = 30.0 Hz), 127.3, 126.7, 126.3 (q, *J* = 5.7 Hz), 124.4 (d, *J* = 273.9 Hz), 123.8, 114.9, 102.0, 38.9, 14.8. HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₁₈H₁₆N₂F₃S 349.0981; Found 349.0983.



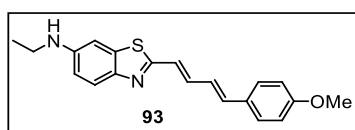
The crude product was purified using cyclohexane/ethyl acetate (75/25, v/v) as the eluent. Yield 79 %. Mp: 69-70 °C. ¹H NMR (200 MHz, CDCl₃) δ (ppm) 7.74 (d, *J* = 8.8 Hz, 1H), 7.64 (d, *J* = 16.5 Hz, 1H), 7.41 (d, *J* = 16.4 Hz, 1H), 7.23 (dd, *J* = 7.9, 1.3 Hz, 1H), 7.07 (t, *J* = 8.0 Hz, 1H), 6.94 (d, *J* = 2.2 Hz, 1H), 6.88 (dd, *J* = 8.1, 1.4 Hz, 1H), 6.73 (dd, *J* = 8.8, 2.3 Hz, 1H), 3.89 (s, 3H), 3.88 (s, 3H), 3.21 (q, *J* = 7.1 Hz, 2H), 1.29 (t, *J* = 7.1 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ (ppm) 162.8, 153.2, 147.6, 146.9, 146.4, 136.8, 130.2, 130.1, 124.4, 124.2, 123.5, 118.6, 114.5, 112.7, 101.9, 61.4, 55.9, 38.8, 14.8. HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₁₉H₂₁N₂O₂S 341.1318; Found 341.1321.



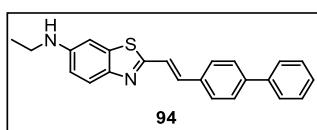
The crude product was purified using cyclohexane/ethyl acetate (7/3, v/v) as the eluent. Yield 72 %. Mp: 50-51 °C. ^1H NMR (200 MHz, CDCl_3) δ (ppm) 7.65 (d, $J = 8.8$ Hz, 1H), 7.23 (d, $J = 16.2$ Hz, 1H), 7.14 (d, $J = 16.3$ Hz, 1H), 6.87 (d, $J = 2.2$ Hz, 1H), 6.70 – 6.63 (m, 3H), 3.82 (s, 6H), 3.80 (s, 3H), 3.13 (q, $J = 7.1$ Hz, 2H), 1.22 (t, $J = 7.1$ Hz, 3H). ^{13}C NMR (50 MHz, CDCl_3) δ (ppm) 161.9, 153.5, 146.9, 146.3, 139.0, 136.7, 135.2, 131.6, 123.4, 122.2, 114.6, 104.2, 101.9, 61.1, 56.2, 38.8, 14.8. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{20}\text{H}_{23}\text{N}_2\text{O}_3\text{S}$ 371.1424; Found 371.1431.



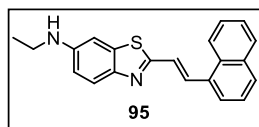
The crude product was purified using cyclohexane/ethyl acetate (85/15, v/v) as the eluent. Yield 36 %. Mp: 122-123 °C. ^1H NMR (400 MHz, CDCl_3) δ (ppm) 7.93 (s, 1H), 7.88 (d, $J = 8.2$ Hz, 1H), 7.83 – 7.73 (m, 2H), 7.62 (d, $J = 16.1$ Hz, 1H), 7.41 (d, $J = 16.0$ Hz, 1H), 6.93 (d, $J = 1.9$ Hz, 1H), 6.75 (dd, $J = 8.8, 2.1$ Hz, 1H), 3.89 (br s, 1H), 3.21 (q, $J = 7.1$ Hz, 2H), 1.30 (t, $J = 7.1$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ (ppm) 160.4, 147.5, 146.4, 138.6, 137.4, 130.2 (q, $J = 33.6$ Hz), 129.0 (q, $J = 3.4$ Hz), 128.9, 128.7, 128.4 (q, $J = 30.8$ Hz), 127.8, 124.1, 123.5 (q, $J = 274.3$ Hz), 123.6 - 123.4 (m), 123.3 (q, $J = 272.3$ Hz), 115.1, 101.5, 38.7, 14.8. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{19}\text{H}_{15}\text{N}_2\text{F}_6\text{S}$ 417.0855; Found 417.0856.



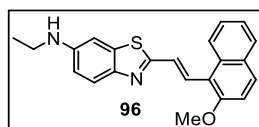
The crude product was purified using cyclohexane/ethyl acetate (8/2, v/v) as the eluent. Yield 59 %. Mp: 148-149 °C. ^1H NMR (400 MHz, CDCl_3) δ (ppm) 7.72 (d, $J = 8.7$ Hz, 1H), 7.42 (d, $J = 7.3$ Hz, 2H), 7.14 (dd, $J = 15.0, 10.4$ Hz, 1H), 6.94 – 6.82 (m, 5H), 6.78 – 6.72 (m, 2H), 3.83 (s, 3H), 3.79 (br s, 1H), 3.22 (q, $J = 6.7$ Hz, 2H), 1.30 (t, $J = 6.7$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ (ppm) 162.3, 160.0, 146.8, 146.6, 136.8, 136.4, 136.1, 129.7, 128.3, 125.9, 125.0, 123.3, 114.6, 114.4, 101.9, 55.5, 38.8, 14.9. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{20}\text{H}_{21}\text{N}_2\text{OS}$ 337.1369; Found 337.1370.



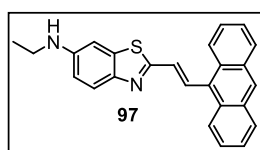
The crude product was purified using cyclohexane/ethyl acetate (8/2, v/v) as the eluent. Yield 77 %. Mp: 194-195 °C. ^1H NMR (400 MHz, CDCl_3) δ (ppm) 7.76 (d, $J = 8.8$ Hz, 1H), 7.68 – 7.59 (m, 6H), 7.63 (s, 6H), 7.46 (t, $J = 7.3$ Hz, 2H), 7.42 – 7.33 (m, 3H), 6.97 (s, 1H), 6.75 (d, $J = 8.8$ Hz, 1H), 3.81 (br s, 1H), 3.23 (q, $J = 6.9$ Hz, 2H), 1.31 (t, $J = 7.0$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ (ppm) 162.2, 147.0, 146.5, 141.7, 140.5, 136.9, 135.0, 134.9, 129.0, 127.7, 127.7, 127.6, 127.1, 123.5, 122.7, 114.7, 101.9, 38.8, 14.9. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{23}\text{H}_{21}\text{N}_2\text{S}$ 357.1420; Found 357.1421.



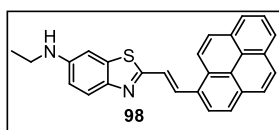
The crude product was purified using cyclohexane/ethyl acetate (8/2, v/v) as the eluent. Yield 72 %. Mp: 116-117 °C. ^1H NMR (400 MHz, CDCl_3) δ (ppm) 8.26 (d, $J = 8.1$ Hz, 1H), 8.16 (d, $J = 15.9$ Hz, 1H), 7.84 (m, 4H), 7.55 (m, 3H), 7.44 (d, $J = 15.8$ Hz, 1H), 6.99 (s, 1H), 6.77 (d, $J = 8.7$ Hz, 1H), 3.89 (br s, 1H), 3.24 (q, $J = 6.6$ Hz, 1H), 1.32 (t, $J = 6.9$ Hz, 1H). ^{13}C NMR (101 MHz, CDCl_3) δ (ppm) 162.2, 147.0, 146.5, 136.9, 133.9, 133.4, 132.3, 131.3, 129.4, 128.9, 126.6, 126.2, 125.8, 125.3, 124.3, 123.7, 123.6, 114.7, 102.0, 38.9, 14.9. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{21}\text{H}_{19}\text{N}_2\text{S}$ 331.1263; Found 331.1264.



The crude product was purified using cyclohexane/ethyl acetate (8/2, v/v) as the eluent. Yield 78 %. 74-75 °C. ¹H NMR (400 MHz, CDCl₃) δ (ppm) 8.27 (d, *J* = 8.6 Hz, 1H), 7.96 (d, *J* = 16.4 Hz, 1H), 7.84 - 7.76 (m, 3H), 7.69 (d, *J* = 16.4 Hz, 1H), 7.53 (t, *J* = 7.7 Hz, 1H), 7.39 (t, *J* = 7.4 Hz, 1H), 7.32 (d, *J* = 9.0 Hz, 1H), 6.99 (s, 1H), 6.76 (d, *J* = 8.7 Hz, 1H), 4.03 (s, 3H), 3.91 (br s, 1H), 3.23 (q, *J* = 7.1 Hz, 2H), 1.32 (t, *J* = 7.1 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ (ppm) 163.8, 156.0, 146.8, 146.6, 136.7, 132.6, 130.4, 129.3, 129.1, 128.7, 128.4, 127.2, 123.9, 123.7, 123.5, 118.2, 114.5, 113.0, 102.2, 56.4, 38.9, 14.9. HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₂₂H₂₁N₂OS 361.1369; Found 361.1369.



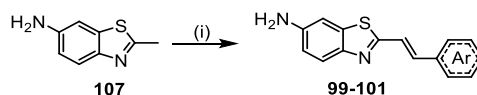
The crude product was purified using cyclohexane/ethyl acetate (75/25, v/v) as the eluent. Yield 53 %. Mp: 180-181 °C. ¹H NMR (400 MHz, CDCl₃) δ (ppm) 8.45 (s, 1H), 8.42 – 8.34 (m, 2H), 8.27 (d, *J* = 16.4 Hz, 1H), 8.04 - 8.02 (m, 2H), 7.82 (d, *J* = 8.8 Hz, 1H), 7.53 – 7.48 (m, 4H), 7.29 (d, *J* = 16.4 Hz, 1H), 7.05 (d, *J* = 2.1 Hz, 1H), 6.81 (dd, *J* = 8.8, 2.2 Hz, 1H), 4.17 (br s, 1H), 3.27 (q, *J* = 7.1 Hz, 2H), 1.34 (t, *J* = 7.1 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ (ppm) 161.9, 146.8, 146.7, 136.9, 132.2, 131.6, 131.3, 130.8, 129.7, 128.9, 127.7, 126.2, 125.8, 125.5, 123.8, 114.9, 102.4, 39.1, 14.9. HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₂₅H₂₁N₂S 381.1420; Found 381.1422.



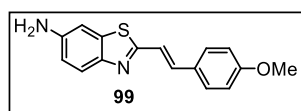
The crude product was purified using cyclohexane/ethyl acetate (7/3, v/v) as the eluent. Yield 81 %. Mp: 179-180 °C. ¹H NMR (400 MHz, CDCl₃) δ (ppm) 8.51 (d, *J* = 9.3 Hz, 1H), 8.45

(d, $J = 15.9$ Hz, 1H), 8.34 (d, $J = 8.1$ Hz, 1H), 8.21 – 8.16 (m, 4H), 8.09 – 8.00 (m, 3H), 7.81 (d, $J = 8.7$ Hz, 1H), 7.62 (d, $J = 15.9$ Hz, 1H), 7.01 (d, $J = 2.2$ Hz, 1H), 6.79 (dd, $J = 8.7, 2.3$ Hz, 1H), 4.00 (br s, 1H), 3.25 (q, $J = 7.1$ Hz, 2H), 1.33 (t, $J = 7.1$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ (ppm) 162.5, 146.9, 146.8, 137.0, 132.2, 131.9, 131.6, 131.0, 130.2, 129.0, 128.3, 128.1, 127.7, 126.3, 125.8, 125.6, 125.4, 125.2, 125.1, 125.0, 123.8, 123.6, 122.9, 114.8, 102.2, 38.9, 14.9. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{27}\text{H}_{21}\text{N}_2\text{S}$ 405.1420; Found 405.1420.

General procedure for the synthesis of dyes 99-101.

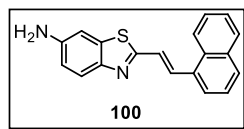


(i) **107** (1 equiv.) in DMF (0.2 M), potassium hydroxide (1 equiv.), and the aldehyde (1,3 equiv.) were subjected to ultrasonic irradiation for 1 hour at room temperature. The solvent was removed in *vacuo*, and the residue suspended in water. After acidification to pH 1 using concentrated HCl, the aqueous layer was extracted with ethyl acetate to remove the aldehyde in excess. The water layer was neutralized with NaOH (1M) and extracted with ethyl acetate. The organic layer was dried over magnesium sulfate, filtered and the volatiles were removed in *vacuo*. The crude was purified by flash column chromatography on silica gel using cyclohexane/ethyl acetate mixtures as the eluents. Finally, recrystallization from methanol afforded the desired dyes as bright solids.

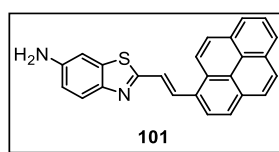


The crude product was purified using cyclohexane/ethyl acetate (8/2, v/v) as the eluent. Yield 59 %. Mp: 182-183 °C. ^1H NMR (400 MHz, CDCl_3) δ (ppm) 7.74 (d, $J = 8.6$ Hz, 1H), 7.49 (d, $J = 8.7$ Hz, 2H), 7.32 (d, $J = 16.2$ Hz, 1H), 7.22 (d, $J = 16.2$ Hz, 1H), 7.08 (d, $J = 2.2$ Hz, 1H), 6.92 (d, $J = 8.7$ Hz, 2H), 6.80 (dd, $J = 8.6, 2.3$ Hz, 1H), 3.84 (s, 5H). ^{13}C NMR (101

MHz, CDCl₃) δ (ppm) 163.6, 160.6, 147.5, 144.7, 136.2, 135.8, 128.7, 128.6, 123.4, 120.5, 115.7, 114.5, 105.8, 55.5. HRMS (ESI) m/z : [M + H]⁺ Calcd for C₁₆H₁₅N₂OS 283.0899; Found 283.0900.

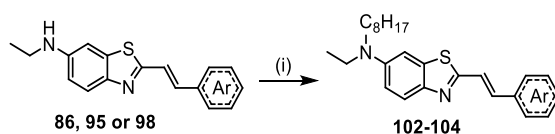


The crude product was purified using cyclohexane/ethyl acetate (8/2, v/v) as the eluent. Yield 65 %. Mp: 194-195 °C. ¹H NMR (400 MHz, CDCl₃) δ (ppm) 8.25 (d, J = 8.4 Hz, 1H), 8.19 (d, J = 15.9 Hz, 1H), 7.88 (t, J = 8.8 Hz, 2H), 7.81 (t, J = 9.1 Hz, 2H), 7.61 - 7.50 (m, 3H), 7.44 (d, J = 16.0 Hz, 1H), 7.12 (s, 1H), 6.84 (d, J = 8.6 Hz, 1H), 3.88 (br s, 2H). ¹³C NMR (101 MHz, CDCl₃) δ (ppm) 163.1, 147.5, 145.0, 136.5, 133.9, 133.2, 132.9, 131.4, 129.5, 128.9, 126.7, 126.2, 125.8, 125.2, 124.4, 123.8, 123.6, 115.9, 105.8. HRMS (ESI) m/z : [M + H]⁺ Calcd for C₁₉H₁₅N₂S 303.0950; Found 303.0951.

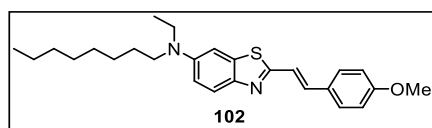


The crude product was purified using cyclohexane/ethyl acetate (7/3, v/v) as the eluent. Yield 46 %. Mp: 225 °C (decomp.). ¹H NMR (400 MHz, DMSO) δ (ppm) 8.66 (d, J = 9.4 Hz, 1H), 8.61 (d, J = 8.3 Hz, 1H), 8.48 (d, J = 15.9 Hz, 1H), 8.40 – 8.31 (m, 4H), 8.23 (d, J = 1.7 Hz, 2H), 8.12 (t, J = 7.6 Hz, 1H), 7.81 (d, J = 15.9 Hz, 1H), 7.70 (d, J = 8.7 Hz, 1H), 7.12 (d, J = 2.1 Hz, 1H), 6.83 (dd, J = 8.7, 2.1 Hz, 1H), 5.58 (d, 2H). ¹³C NMR (101 MHz, DMSO) δ (ppm) 159.9, 147.7, 145.2, 136.2, 131.1, 130.9, 130.4, 130.3, 129.5, 128.2, 128.1, 127.8, 127.4, 126.5, 125.8, 125.5, 125.4, 124.9, 124.2, 123.9, 123.8, 123.1, 122.6, 115.1, 103.6. HRMS (ESI) m/z : [M + H]⁺ Calcd for C₂₅H₁₇N₂S 377.1107; Found 377.1113.

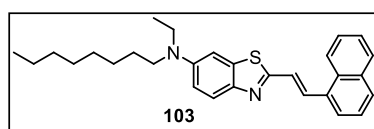
General procedure for the synthesis of dyes 102-104.



(i) 6-ethylamino substituted dye **86**, **95** or **98** (1 equiv.) in acetonitrile (0.2 M), 1-iodooctane (3 equiv.) and potassium carbonate (3 equiv.) were heated to reflux during 24 hours in a sealed vial. The solvent was removed under reduced pressure and the crude was purified by flash column chromatography on silica gel, using cyclohexane/ethyl acetate mixtures as the eluents. Recrystallization from methanol afforded the desired dyes as bright solids.

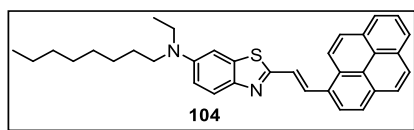


The crude product was purified using cyclohexane/ethyl acetate (95/5, v/v) as the eluent. Yield 86 %. Mp: 56-57 °C. ¹H NMR (400 MHz, CDCl₃) δ (ppm) 7.74 (d, *J* = 9.1 Hz, 1H), 7.47 (d, *J* = 8.7 Hz, 2H), 7.27 (d, *J* = 16.2 Hz, 1H), 7.20 (d, *J* = 16.2 Hz, 1H), 6.96 (d, *J* = 2.4 Hz, 1H), 6.90 (d, *J* = 8.7 Hz, 2H), 6.82 (dd, *J* = 9.1, 2.5 Hz, 1H), 3.82 (s, 3H), 3.41 (q, *J* = 7.0 Hz, 2H), 3.28 (t, *J* = 7.8 Hz, 2H), 1.64 – 1.58 (m, 2H), 1.33 – 1.24 (m, 10H), 1.18 (t, *J* = 7.0 Hz, 3H), 0.88 (t, *J* = 6.8 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ (ppm) 162.1, 160.4, 146.6, 145.1, 136.9, 134.9, 128.9, 128.6, 123.1, 120.8, 114.5, 112.7, 102.1, 55.5, 51.1, 45.6, 31.9, 29.6, 29.5, 27.7, 27.3, 22.8, 14.2, 12.5. HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₂₆H₃₅N₂OS 423.2465; Found 423.2465.



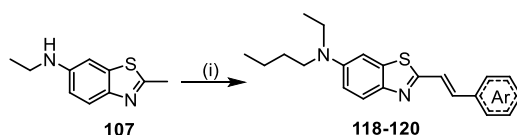
The crude product was purified using cyclohexane/ethyl acetate (9/1, v/v) as the eluent. Yield 88 %. Mp: 44-45 °C. ¹H NMR (200 MHz, CDCl₃) δ (ppm) 7.75 (d, *J* = 8.8 Hz, 1H), 7.55 (dd, *J* = 8.0, 1.5 Hz, 2H), 7.45 – 7.30 (m, 5H), 6.95 (d, *J* = 2.3 Hz, 1H), 6.74 (dd, *J* = 8.8, 2.4 Hz, 1H), 3.68 (s, 1H), 3.21 (q, *J* = 7.1 Hz, 2H), 1.30 (t, *J* = 7.1 Hz, 3H). ¹³C NMR (50 MHz, CDCl₃) δ (ppm) 162.2, 146.9, 146.4, 136.8, 135.9, 135.5, 129.0, 127.2, 123.5, 122.7, 114.7,

101.9, 38.8, 14.9. HRMS (ESI) m/z : $[M + H]^+$ Calcd for $C_{29}H_{35}N_2S$ 443.2515; Found 443.2516.

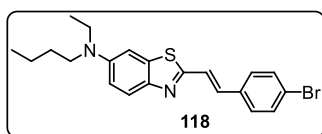


The crude product was purified using cyclohexane/ethyl acetate (9/1, v/v) as the eluent. Yield 80 %. Mp: 130-131 °C. 1H NMR (400 MHz, $CDCl_3$) δ (ppm) 8.50 (d, $J = 9.3$ Hz, 1H), 8.41 (d, $J = 15.9$ Hz, 1H), 8.33 (d, $J = 8.1$ Hz, 1H), 8.20 - 8.14 (m, 4H), 8.08 - 7.99 (m, 3H), 7.85 (d, $J = 9.0$ Hz, 1H), 7.62 (d, $J = 15.9$ Hz, 1H), 7.04 (d, $J = 1.9$ Hz, 1H), 6.89 (dd, $J = 9.0, 2.1$ Hz, 1H), 3.45 (q, $J = 6.9$ Hz, 2H), 3.40 - 3.28 (t, $J = 7.8$ Hz), 1.73 - 1.59 (m, 2H), 1.37 - 1.32 (m, 10H), 1.23 (t, $J = 7.0$ Hz, 3H), 0.91 (t, $J = 5.9$ Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ (ppm) 161.8, 146.9, 145.3, 137.3, 131., 131.7, 131.6, 131.0, 130.2, 128.9, 128.2, 127.9, 127.6, 126.3, 125.8, 125.5, 125.4, 125.2, 125.2, 125.0, 123.8, 123.5, 122.9, 112.9, 102.0, 51.1, 45.6, 31.9, 29.7, 29.5, 27.7, 27.4, 22.8, 14.2, 12.6. HRMS (ESI) m/z : $[M + H]^+$ Calcd for $C_{35}H_{37}N_2S$ 517.2672; Found 517.2674.

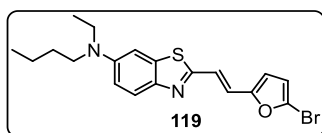
General procedure for the synthesis of dyes **118-120**



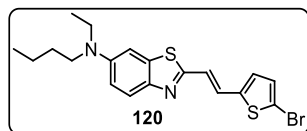
(i) Precursor **109** (1 equiv.) in DMF (0.2 M), potassium hydroxide (1 equiv.) and corresponding aldehyde (1.3 equiv.) were stirred overnight at room temperature. After complete conversion of the starting material by TLC, the base was filtered off and 3 equivalents of 1-iodobutane and stoichiometric amount of potassium carbonate were added. The mixture was heated to 80°C overnight. The solvent was removed in *vacuo*, and the crude was purified by flash column chromatography on silica gel using cyclohexane/ethyl acetate mixtures as the eluents. Finally, recrystallization from methanol afforded the desired dyes in good yield and satisfactory purity for the following steps.



The crude product was purified using cyclohexane/ethyl acetate (9/1, v/v) as the eluent. Yield 52 %. ^1H NMR (400 MHz, CDCl_3) δ 7.78 (d, $J = 9.1$ Hz, 1H), 7.51 (d, $J = 8.5$ Hz, 2H), 7.40 (d, $J = 8.5$ Hz, 2H), 7.34 (d, $J = 16.2$ Hz, 1H), 7.25 (t, $J = 16.2$ Hz, 1H), 6.98 (d, $J = 2.4$ Hz, 1H), 6.86 (dd, $J = 9.1, 2.5$ Hz, 1H), 3.43 (q, $J = 7.1$ Hz, 2H), 3.35 (t, $J = 7.5$ Hz, 2H), 1.66 – 1.62 (m, 2H), 1.42 – 1.34 (m, 2H), 1.20 (t, $J = 7.0$ Hz, 3H), 0.98 (t, $J = 7.3$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 161.0, 146.8, 145.1, 137.2, 135.1, 133.6, 132.2, 128.6, 123.5, 123.5, 122.8, 112.9, 101.8, 50.8, 45.6, 29.8, 20.5, 14.2, 12.5. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{21}\text{H}_{24}\text{BrN}_2\text{S}$ 415.0844; Found 415.0840.



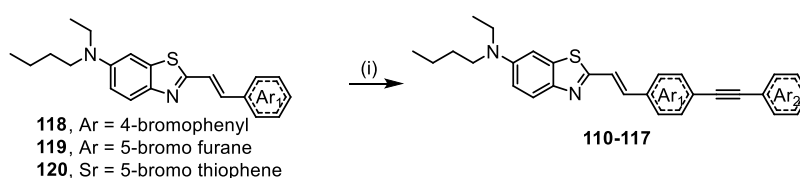
The crude product was purified using cyclohexane/ethyl acetate (9/1, v/v) as the eluent. Yield 31 %. ^1H NMR (400 MHz, CDCl_3) δ 7.77 (d, $J = 9.1$ Hz, 1H), 7.19 (d, $J = 15.9$ Hz, 1H), 7.05 (d, $J = 16.0$ Hz, 1H), 6.97 (s, 1H), 6.86 (d, $J = 8.9$ Hz, 1H), 6.44 (d, $J = 3.3$ Hz, 1H), 6.38 (d, $J = 3.4$ Hz, 1H), 3.43 (q, $J = 7.0$ Hz, 2H), 3.32 (t, $J = 7.8$ Hz, 2H), 1.65 – 1.58 (m, 2H), 1.43 – 1.34 (m, 2H), 1.20 (t, $J = 7.0$ Hz, 3H), 0.97 (t, $J = 7.3$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 160.5, 154.1, 146.7, 145.3, 137.4, 123.6, 123.4, 121.1, 120.7, 114.1, 113.4, 113.0, 101.9, 50.8, 45.7, 29.8, 20.5, 14.2, 12.5. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{19}\text{H}_{22}\text{BrN}_2\text{OS}$ 405.0636; Found 405.0629.



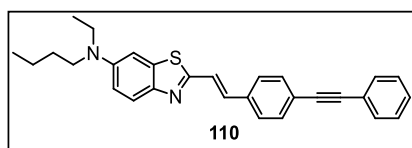
The crude product was purified using cyclohexane/ethyl acetate (9/1, v/v) as the eluent. Yield 40 %. ^1H NMR (400 MHz, CDCl_3) δ 7.76 (d, $J = 9.1$ Hz, 1H), 7.32 (d, $J = 15.9$ Hz, 1H), 7.04 (d, $J = 15.9$ Hz, 1H), 6.98 (d, $J = 3.8$ Hz, 1H), 6.96 (d, $J = 2.5$ Hz, 1H), 6.91 (d, $J = 3.9$ Hz, 1H), 6.85 (dd, $J = 9.1, 2.5$ Hz, 1H), 3.43 (q, $J = 7.1$ Hz, 2H), 3.32 (t, $J = 7.7$ Hz, 2H), 1.65 – 1.58 (m, 2H), 1.44 – 1.34 (m, 2H), 1.20 (t, $J = 7.1$ Hz, 3H), 0.98 (t, $J = 7.3$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 160.4, 146.8, 145.1, 143.1, 137.2, 130.9, 128.3, 126.9, 123.4, 122.4,

113.5, 112.9, 101.8, 50.8, 45.6, 29.8, 20.5, 14.2, 12.5. HRMS (ESI) m/z : $[M + H]^+$ Calcd for $C_{19}H_{22}BrN_2S_2$ 421.0408; Found 421.0404.

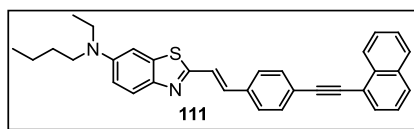
General procedure for the synthesis of dyes 110-117



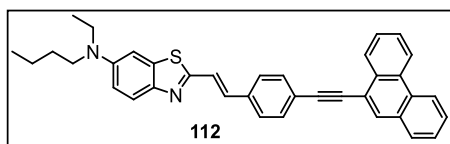
(i) Precursor dye (**118-120**) (0.1 mmol), $Pd(PPh)_4$ (20 mol%), CuI (10 mol%) and corresponding aryl alkyne (1.5 equiv) were suspended in DMF/TEA (1:1, 0.05 M) and heated to $50^\circ C$ overnight. The mixture was diluted with ethyl acetate (5 mL) and filtered on celite and extracted with 0.1 M solution of EDTA disodium salt (3 x 5 mL). The organic phase was dried with $MgSO_4$ and concentrated in *vacuo*. The crude was purified by preparative thin layer chromatography using cyclohexane/ethyl acetate mixtures as eluent to afford dyes **110-117** as bright solids.



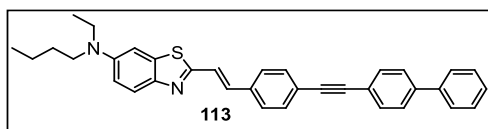
The crude product was purified using cyclohexane/ethyl acetate (9/1, v/v) as the eluent. Yield 32 %. 1H NMR (400 MHz, CD_2Cl_2) δ 7.74 (d, $J = 9.1$ Hz, 1H), 7.56 – 7.54 (m, 6H), 7.41 – 7.37 (m, 4H), 7.32 (d, $J = 16.2$ Hz, 1H), 7.02 (d, $J = 2.5$ Hz, 1H), 6.87 (dd, $J = 9.1, 2.5$ Hz, 1H), 3.44 (q, $J = 7.0$ Hz, 2H), 3.34 (t, $J = 7.8$ Hz, 2H), 1.67 – 1.61 (m, 2H), 1.45 – 1.35 (m, 2H), 1.20 (t, $J = 7.0$ Hz, 3H), 0.98 (t, $J = 7.4$ Hz, 3H). ^{13}C NMR (101 MHz, CD_2Cl_2) δ 161.1, 147.2, 145.4, 137.6, 136.4, 134.1, 132.4, 131.9, 128.8, 127.4, 124.0, 123.7, 123.6, 123.5, 113.1, 102.1, 91.2, 89.7, 51.0, 45.8, 30.0, 20.7, 14.2, 12.5. HRMS (ESI) m/z : $[M + H]^+$ Calcd for $C_{29}H_{29}N_2S$ 437.2051; Found 421.2046.



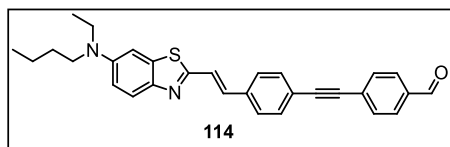
The crude product was purified using cyclohexane/ethyl acetate (9/1, v/v) as the eluent. Yield 42 %. ^1H NMR (400 MHz, CDCl_3) δ 8.45 (d, $J = 8.3$ Hz, 1H), 7.87 (t, $J = 8.3$ Hz, 2H), 7.79 (t, $J = 8.2$ Hz, 2H), 7.67 – 7.60 (m, 3H), 7.58 – 7.53 (m, 3H), 7.47 (dd, $J = 8.1, 7.3$ Hz, 1H), 7.41 (d, $J = 16.2$ Hz, 1H), 7.32 (d, $J = 16.2$ Hz, 1H), 6.99 (d, $J = 2.5$ Hz, 1H), 6.87 (dd, $J = 9.1, 2.5$ Hz, 1H), 3.44 (q, $J = 7.1$ Hz, 2H), 3.33 (t, $J = 7.8$ Hz, 2H), 1.67 – 1.59 (m, 2H), 1.43 – 7.35 (m, 2H), 1.21 (t, $J = 7.0$ Hz, 3H), 0.99 (t, $J = 7.3$ Hz, 3H). ^{13}C NMR (101 MHz, CDCl_3) δ 161.2, 146.8, 145.1, 137.3, 136.1, 134.2, 133.4, 132.2, 130.6, 129.0, 128.5, 127.1, 126.9, 126.6, 126.3, 125.4, 123.7, 123.6, 123.5, 120.9, 112.9, 101.8, 94.5, 89.3, 50.8, 45.6, 29.8, 20.5, 14.2, 12.5. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{33}\text{H}_{31}\text{N}_2\text{S}$ 487.2208; Found 487.2205.



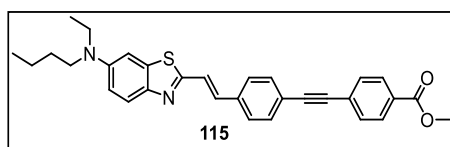
The crude product was purified using cyclohexane/ethyl acetate (9/1, v/v) as the eluent. Yield 23 %. ^1H NMR (400 MHz, CD_2Cl_2) δ 8.74 – 8.70 (m, 1H), 8.69 (d, $J = 8.2$ Hz, 1H), 8.57 – 8.54 (m, 1H), 8.11 (s, 1H), 7.91 (d, $J = 7.4$ Hz, 1H), 7.76 – 7.67 (m, 6H), 7.65 – 7.60 (m, 3H), 7.41 (d, $J = 16$ Hz, 1H), 7.34 (d, $J = 16.2$ Hz, 1H), 7.02 (d, $J = 2.4$ Hz, 1H), 6.86 (dd, $J = 9.1, 2.5$ Hz, 1H), 3.43 (q, $J = 7.0$ Hz, 2H), 3.33 (t, $J = 7.8$ Hz, 2H), 1.66 – 1.62 (m, 2H), 1.42 – 1.34 (m, 2H), 1.19 (t, $J = 7.0$ Hz, 3H), 0.98 (t, $J = 7.3$ Hz, 3H). ^{13}C NMR (101 MHz, CD_2Cl_2) δ 161.1, 147.2, 145.4, 137.6, 136.6, 134.1, 132.5, 132.3, 131.6, 131.4, 130.7, 130.5, 129.0, 128.1, 127.6, 127.6, 127.5, 127.4, 127.2, 124.1, 123.7, 123.6, 123.3, 123.0, 119.8, 113.1, 102.1, 94.4, 89.5, 51.0, 45.8, 30.1, 30.0, 20.8, 14.2, 12.5. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{37}\text{H}_{33}\text{N}_2\text{S}$ 537.2364; Found 537.2362.



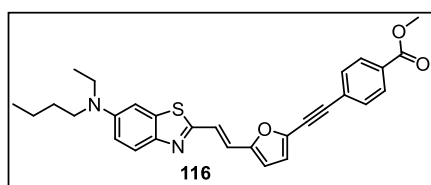
The crude product was purified using cyclohexane/ethyl acetate (9/1, v/v) as the eluent. Yield 57 %. ^1H NMR (400 MHz, CD_2Cl_2) δ 7.74 (d, $J = 9.1$ Hz, 1H), 7.65 – 7.58 (m, 6H), 7.58 (m, 4H), 7.47 (t, $J = 7.6$ Hz, 2H), 7.49 – 7.36 (m, 2H), 7.32 (d, $J = 16.2$ Hz, 1H), 7.02 (d, $J = 2.5$ Hz, 1H), 6.87 (dd, $J = 9.1, 2.5$ Hz, 1H), 3.44 (q, $J = 7.0$ Hz, 2H), 3.34 (t, $J = 7.8$ Hz, 1H), 1.67 – 1.59 (m, 2H), 1.45 – 1.37 (m, 2H), 1.20 (t, $J = 7.0$ Hz, 3H), 0.98 (t, $J = 7.4$ Hz, 3H). ^{13}C NMR (101 MHz, CD_2Cl_2) δ 161.1, 147.2, 145.4, 141.5, 140.6, 137.6, 136.4, 134.1, 132.4, 132.3, 129.3, 128.1, 127.5, 127.4, 127.3, 124.0, 123.7, 123.6, 122.4, 113.1, 102.1, 91.2, 90.4, 51.0, 45.8, 30.0, 20.8, 14.2, 12.5. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{35}\text{H}_{33}\text{N}_2\text{S}$ 513.2364; Found 513.2359.



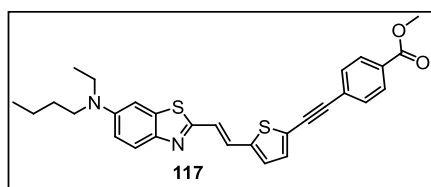
The crude product was purified using cyclohexane/ethyl acetate (8.5/1.5, v/v) as the eluent. Yield 52 %. ^1H NMR (400 MHz, CD_2Cl_2) δ 10.01 (s, 1H), 7.87 (d, $J = 8.4$ Hz, 2H), 7.74 (d, $J = 9.1$ Hz, 1H), 7.70 (d, $J = 8.2$ Hz, 2H), 7.58 (s, 4H), 7.40 (d, $J = 16.2$ Hz, 1H), 7.32 (d, $J = 16.2$ Hz, 1H), 7.02 (d, $J = 2.5$ Hz, 1H), 6.87 (dd, $J = 9.1, 2.6$ Hz, 1H), 3.44 (q, $J = 7.0$ Hz, 2H), 3.34 (t, $J = 7.8$ Hz, 2H), 1.66 – 1.61 (m, 2H), 1.44 – 1.35 (m, 2H), 1.20 (t, $J = 7.1$ Hz, 3H), 0.98 (t, $J = 7.4$ Hz, 3H). ^{13}C NMR (101 MHz, CD_2Cl_2) δ 191.7, 160.9, 147.2, 145.4, 137.6, 137.1, 136.0, 133.9, 132.6, 132.5, 129.9, 129.7, 127.4, 124.4, 123.6, 122.9, 113.1, 102.0, 93.5, 90.32, 51.0, 45.8, 30.0, 20.7, 14.2, 12.5. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{30}\text{H}_{29}\text{N}_2\text{OS}$ 465.2001; Found 465.1999.



The crude product was purified using cyclohexane/ethyl acetate (8/2, v/v) as the eluent. Yield 48 %. ^1H NMR (400 MHz, CD_2Cl_2) δ 8.02 (dd, $J = 8.5, 1.8$ Hz, 2H), 7.74 (d, $J = 9.1$ Hz, 1H), 7.62 (dd, $J = 8.5, 1.8$ Hz, 2H), 7.58 (s, 4H), 7.40 (d, $J = 16.2$ Hz, 1H), 7.32 (d, $J = 16.2$ Hz, 1H), 7.02 (d, $J = 2.5$ Hz, 1H), 6.87 (dd, $J = 9.1, 2.6$ Hz, 1H), 3.91 (s, 3H), 3.44 (q, $J = 7.1$ Hz, 2H), 3.34 (t, $J = 7.6$ Hz, 2H), 1.66 – 1.59 (m, 2H), 1.44 – 1.35 (m, 2H), 1.20 (t, $J = 7.0$ Hz, 3H), 0.98 (t, $J = 7.4$ Hz, 3H). ^{13}C NMR (101 MHz, CD_2Cl_2) δ 166.7, 160.9, 147.2, 145.4, 137.6, 136.9, 133.9, 132.6, 131.9, 130.2, 129.9, 128.1, 127.4, 124.3, 123.6, 123.1, 113.1, 102.0, 92.6, 90.4, 52.5, 51.0, 45.8, 30.0, 20.7, 14.2, 12.5. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{31}\text{H}_{31}\text{N}_2\text{O}_2\text{S}$ 495.2106; Found 495.2105.



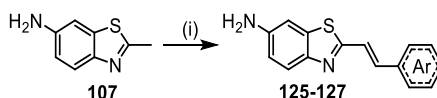
The crude product was purified using cyclohexane/ethyl acetate (8/2, v/v) as the eluent. Yield 33 %. ^1H NMR (400 MHz, CD_2Cl_2) δ 8.03 (dd, $J = 8.5, 1.8$ Hz, 2H), 7.73 (d, $J = 9.1$ Hz, 1H), 7.62 (dd, $J = 8.5, 1.7$ Hz, 2H), 7.29 (d, $J = 16.0$ Hz, 1H), 7.12 (d, $J = 16.0$ Hz, 1H), 7.01 (d, $J = 2.2$ Hz, 1H), 6.87 (dd, $J = 9.1, 2.5$ Hz, 1H), 6.81 (d, $J = 3.5$ Hz, 1H), 6.58 (d, $J = 3.5$ Hz, 1H), 3.91 (s, 3H), 3.44 (q, $J = 7.0$ Hz, 2H), 3.33 (t, $J = 7.4$ Hz, 2H), 1.66 – 1.58 (m, 2H), 1.44 – 1.35 (m, 2H), 1.19 (t, $J = 7.0$ Hz, 3H), 0.98 (t, $J = 7.3$ Hz, 3H). ^{13}C NMR (101 MHz, CD_2Cl_2) δ 166.6, 160.3, 153.8, 147.2, 145.6, 137.8, 137.3, 131.6, 130.5, 129.9, 126.9, 123.6, 122.8, 120.9, 119.3, 113.2, 112.9, 102.0, 94.6, 82.7, 54.4, 52.6, 51.0, 45.8, 30.0, 20.7, 14.2, 12.5. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{29}\text{H}_{29}\text{N}_2\text{O}_3\text{S}$ 485.1899; Found 485.1897.



The crude product was purified using cyclohexane/ethyl acetate (8/2, v/v) as the eluent. Yield 45 %. ^1H NMR (400 MHz, CD_2Cl_2) δ 8.02 (dd, $J = 8.5, 1.7$ Hz, 2H), 7.72 (d, $J = 9.1$ Hz, 1H), 7.60 (dd, $J = 8.5, 1.6$ Hz, 2H), 7.41 (d, $J = 15.9$ Hz, 1H), 7.26 (d, $J = 3.8$ Hz, 1H), 7.18 – 7.12 (m, 2H), 7.00 (d, $J = 2.5$ Hz, 1H), 6.86 (dd, $J = 9.1, 2.5$ Hz, 1H), 3.91 (s, 3H), 3.44 (q, $J = 7.0$ Hz, 2H), 3.33 (t, $J = 7.5$ Hz, 2H), 1.66 – 1.62 (m, 2H), 1.44 – 1.35 (m, 2H), 1.19 (t, $J = 7.0$ Hz, 3H), 0.98 (t, $J = 7.3$ Hz, 3H). ^{13}C NMR (101 MHz, CD_2Cl_2) δ 166.7, 160.2, 147.2, 145.4, 144.1, 137.7, 134.1, 131.6, 130.3, 129.9, 128.6, 127.6, 126.9, 123.6, 123.6, 123.2, 113.2, 102.0, 94.5, 86.1, 52.5, 51.0, 45.8, 30.0, 20.7, 14.2, 12.5. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{29}\text{H}_{29}\text{N}_2\text{O}_2\text{S}_2$ 501.1670; Found 501.1665.

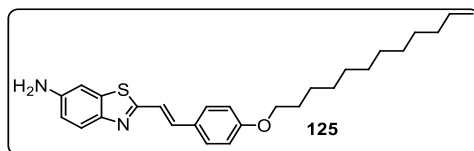
Chapter 3

General procedure for the synthesis of fluorophores 86, 125-127.

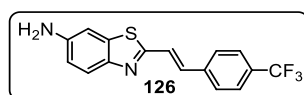


(i) 2-methylbenzo[d]thiazol-6-amine (**107**, 2.0 mmol, 1 equiv.) in DMF (0.2 M), potassium hydroxide (2.0mmol, 1 equiv.), and the aldehyde (2.6mmol, 1.3 equiv.) were stirred for 12 hours at room temperature. The solvent was removed in *vacuo*, and the residue suspended in water. After acidification to pH 1 using concentrated HCl, the aqueous layer was extracted with ethyl acetate to remove the aldehyde in excess. The aqueous layer was neutralized with NaOH (1M) and extracted with ethyl acetate. The organic layer was dried over magnesium sulfate, filtered and the volatiles were removed in *vacuo*. The crude was purified by flash

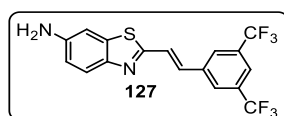
column chromatography on silica gel using cyclohexane/ethyl acetate (Cy/EA) mixtures as the eluent. Finally, recrystallization from methanol afforded the desired fluorophores, **125-127**, as bright solids.



Starting from 328mg of **107**, compound **125** was obtained in 53% yield after purification on silica gel column chromatography (eluent: Cy/EA - 80/20) and recrystallization from methanol. ¹H NMR (400 MHz, CDCl₃) δ 7.73 (d, *J* = 8.6 Hz, 1H), 7.47 (d, *J* = 8.2 Hz, 2H), 7.32 (d, *J* = 16.2 Hz, 1H), 7.21 (d, *J* = 16.2 Hz, 1H), 7.07 (s, 1H), 6.91 (d, *J* = 8.1 Hz, 2H), 6.80 (d, *J* = 8.6 Hz, 1H), 3.98 (t, *J* = 6.4 Hz, 2H), 3.83 (s, 2H), 1.87 – 1.76 (m, 2H), 1.47 – 1.44 (m, 2H), 1.32 – 1.27 (m, 16H), 0.88 (t, *J* = 6.4 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 163.7, 160.2, 147.5, 144.7, 136.2, 135.9, 128.7, 128.4, 123.4, 120.3, 115.7, 115.0, 105.8, 68.3, 32.1, 29.8, 29.8, 29.7, 29.7, 29.5, 29.5, 29.4, 26.2, 22.8, 14.3. HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₂₇H₃₇N₂OS 437.2627; Found 437.2622.



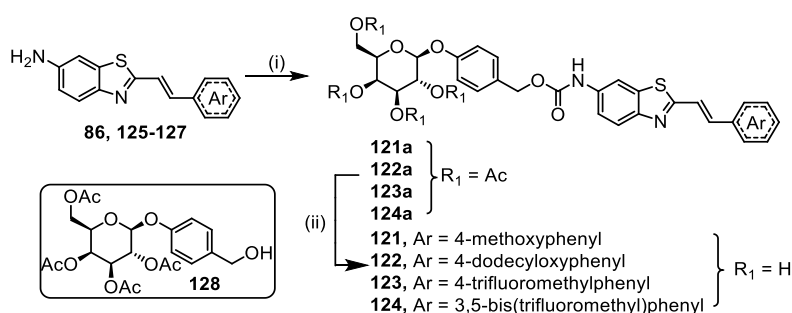
Starting from 328 mg of **107**, compound **126** was obtained in 43% yield after purification on silica gel column chromatography (eluent: Cy/EA - 80/20) and recrystallization from methanol. ¹H NMR (400 MHz, CDCl₃) δ 7.78 (d, *J* = 8.7 Hz, 1H), 7.64 (s, 4H), 7.43 (d, *J* = 16.3 Hz, 1H), 7.37 (d, *J* = 16.3 Hz, 1H), 7.10 (d, *J* = 2.1 Hz, 1H), 6.84 (dd, *J* = 8.7, 2.3 Hz, 1H), 3.90 (s, 2H). ¹⁹F NMR (376 MHz, CDCl₃) δ -62.66. ¹³C NMR (101 MHz, CDCl₃) δ 162.1, 147.4, 145.3, 139.3, 136.6, 133.9, 130.6 (q, *J* = 33 Hz), 127.3, 126.0 (q, *J* = 4 Hz), 124.9, 124.2 (q, *J* = 272 Hz), 123.9, 116.0, 105.6. HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₁₆H₁₂F₃N₂S 321.0673; Found 321.0669.



Starting from 328 mg of **107**, compound **127** was obtained in 62% yield after purification on silica gel column chromatography (eluent: Cy/EA - 85/15) and recrystallization from

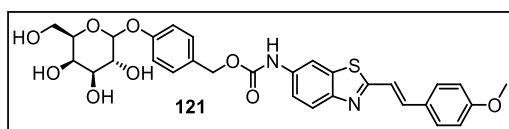
methanol. ^1H NMR (400 MHz, CDCl_3) δ 7.96 (s, 2H), 7.81 – 7.79 (m, 2H), 7.48 (d, J = 16.2 Hz, 1H), 7.40 (d, J = 16.3 Hz, 1H), 7.10 (d, J = 2.2 Hz, 1H), 6.85 (dd, J = 8.7, 2.3 Hz, 1H), 3.92 (br s, 2H). ^{19}F NMR (376 MHz, CDCl_3) δ -62.40. ^{13}C NMR (101 MHz, CDCl_3) δ 161.0, 147.4, 145.5, 138.1, 136.8, 132.5 (q, J = 33 Hz), 131.9, 126.9 – 126.8 (m), 126.2, 124.2, 123.3 (q, J = 273 Hz), 122.2 – 122.0 (m), 116.2, 105.5. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{17}\text{H}_{11}\text{F}_6\text{N}_2\text{S}$ 389.0547; Found 389.0545.

General procedure for the synthesis of probes 121-124.

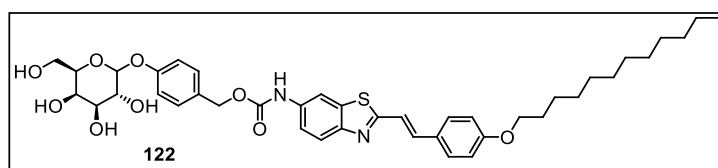


(i) The probe precursor (0.7 mmol, **86**, **125-127**, 1 equiv.) was dissolved in dry 1,4-dioxane (20 mL) with 3 equivalents of triethylamine. This solution was added dropwise, over a period of 10 minutes, to a solution containing an equimolar amount of triphosgene (0.7 mmol, 1 equiv.) in dry 1,4-dioxane (5 mL). The mixture was stirred at room temperature until complete conversion of the starting material (analyzed by LC/MS, sample diluted in methanol). Then, benzyl alcohol **128** was added (0.7 mmol, 1 equiv.) and the mixture was stirred overnight at room temperature. Subsequently, the solids were filtered off, the crude filtrate was adsorbed onto silica and purified by flash column chromatography (Cy/EA as eluting mixtures) to afford the respective tetra-*O*-acetylated precursors (**121a-124a**).

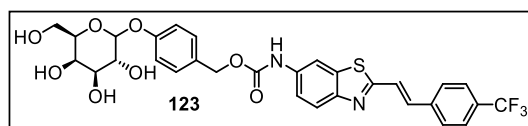
(ii) Finally, the acetylated compounds, **121a-124a**, were treated overnight with sodium methoxide (3.5 mmol, 5 equiv.) in methanol (0.1 M). The precipitates were collected by filtration, washed with methanol and dichloromethane to afford the final **121-124** probes as colorless solids.



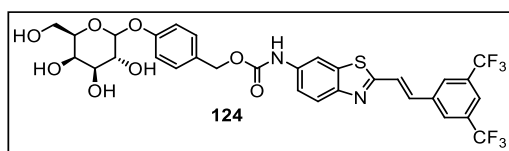
Starting from 200 mg of **86**, **121** was obtained in 18% yield. Intermediate **121a** was purified on silica gel column chromatography using Cy/EA (0-50 % EA) as the eluent mixture. ^1H NMR (400 MHz, DMSO-*d*₆) δ 10.02 (s, 1H), 8.22 (s, 1H), 7.83 (d, J = 8.6 Hz, 1H), 7.71 (d, J = 7.4 Hz, 2H), 7.54 – 7.37 (m, 5H), 7.05 – 6.99 (m, 4H), 5.15 (brs, 1H), 5.11 (s, 2H), 4.87 (brs, 1H), 4.83 (d, J = 5.9 Hz, 1H), 4.64 (brs, 1H), 4.54 (brs, 1H), 3.80 (s, 3H), 3.69 (s, 1H), 3.56 – 3.41 (m, 5H). ^{13}C NMR (101 MHz, DMSO-*d*₆) δ 165.1, 160.3, 157.4, 153.5, 149.1, 136.9, 136.6, 134.7, 129.9, 129.6, 129.2, 127.9, 122.4, 119.6, 118.1, 116.2, 114.4, 110.0, 100.9, 75.5, 73.3, 70.3, 68.1, 65.7, 60.4, 55.3. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{30}\text{H}_{31}\text{N}_2\text{O}_9\text{S}$ 595.1750; Found 595.1747.



Starting from 305mg of **125**, **122** was obtained in 14% yield. Intermediate **122a** was purified on silica gel column chromatography using Cy/EA (0-50 % EA) as the eluent mixture. ^1H NMR (400 MHz, DMSO-*d*₆) δ 10.02 (s, 1H), 8.22 (s, 1H), 7.83 (d, J = 8.7 Hz, 1H), 7.69 (d, J = 8.1 Hz, 2H), 7.53 – 7.46 (m, 2H), 7.41 – 7.37 (m, 3H), 7.01 (dd, J = 28.6, 8.2 Hz, 4H), 5.16 (br s, 1H), 5.12 (s, 2H), 4.88 (br s, 1H), 4.84 (d, J = 7.5 Hz, 1H), 4.64 (br s, 1H), 4.52 (br s, 1H), 4.00 (t, J = 6.0 Hz, 2H), 3.69 (s, 1H), 3.58 – 3.54 (m, 3H), 3.46 – 3.50 (m, J = 4.7 Hz, 1H), 3.41 – 3.39 (m, 1H), 1.73 – 1.67 (m, 2H), 1.40 – 1.39 (m, 2H), 1.35 – 1.24 (m, 16H), 0.84 (t, J = 6.5 Hz, 3H). ^{13}C NMR (101 MHz, DMSO-*d*₆) δ 165.1, 159.8, 157.4, 153.5, 149.1, 136.8, 136.6, 134.7, 129.9, 129.6, 129.2, 127.7, 122.4, 119.5, 118.1, 116.2, 114.8, 109.9, 75.5, 73.3, 70.2, 68.1, 67.6, 65.7, 60.4, 31.3, 29.0, 29.0, 28.9, 28.9, 28.7, 28.7, 28.6, 25.5, 22.1, 13.9. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{41}\text{H}_{53}\text{N}_2\text{O}_9\text{S}$ 749.3472; Found 749.3467.



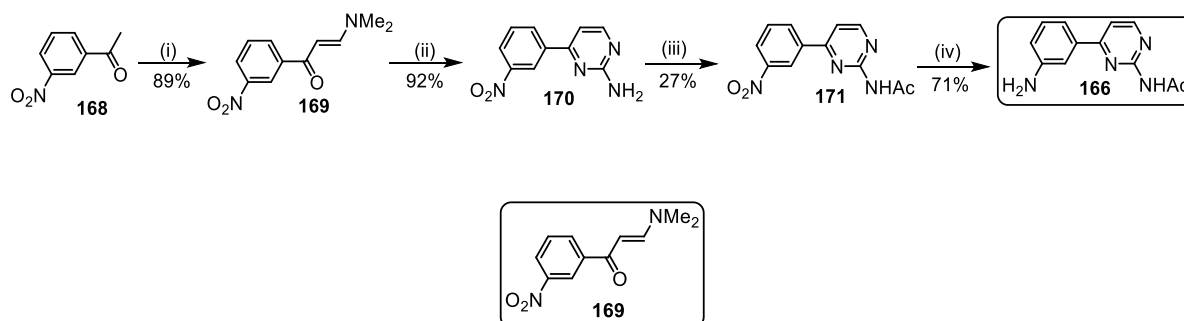
Starting from 224mg of **126**, **123** was obtained in 21% yield. Intermediate **123a** was purified on silica gel column chromatography using Cy/EA (0-50 % EA) as the eluent mixture. ^1H NMR (400 MHz, DMSO-*d*₆) δ 10.07 (s, 1H), 8.28 (s, 1H), 7.99 (d, J = 8.2 Hz, 2H), 7.91 (d, J = 8.9 Hz, 1H), 7.82 – 7.71 (m, 3H), 7.67 (d, J = 16.3 Hz, 1H), 7.51 (dd, J = 8.9, 2.0 Hz, 1H), 7.39 (d, J = 8.6 Hz, 2H), 7.05 (d, J = 8.6 Hz, 2H), 5.17 (d, J = 5.1 Hz, 1H), 5.12 (s, 2H), 4.86 (d, J = 5.7 Hz, 1H), 4.84 (d, J = 7.7 Hz, 1H), 4.64 (t, J = 5.4 Hz, 1H), 4.50 (d, J = 4.6 Hz, 1H), 3.70 (t, J = 3.9 Hz, 1H), 3.60 – 3.52 (m, 3H), 3.52 – 3.45 (m, 1H), 3.44 – 3.37 (m, 1H). ^{19}F NMR (376 MHz, DMSO-*d*₆) δ -61.06. ^{13}C NMR (101 MHz, DMSO-*d*₆) δ 163.9, 157.4, 153.5, 149.0, 139.4, 137.3, 135.2, 134.7, 129.9, 129.5, 128.9 (q, J = 32 Hz), 125.7 (q, J = 3 Hz), 124.6, 124.2 (q, J = 271 Hz), 122.9, 118.3, 116.2, 109.9, 100.9, 75.5, 73.3, 70.3, 68.1, 65.8, 60.4. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for C₃₀H₂₈F₃N₂O₈S 633.1518; Found 633.1514.



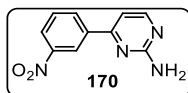
Starting from 271mg of **127**, **124** was obtained in 12% yield. Intermediate **124a** was purified on silica gel column chromatography using Cy/EA (0-50 % EA) as the eluent mixture. ^1H NMR (400 MHz, DMSO-*d*₆) δ 10.08 (s, 1H), 8.52 (s, 2H), 8.30 (d, J = 1.5 Hz, 1H), 8.05 (s, 1H), 8.00 (d, J = 16.3 Hz, 1H), 7.92 (d, J = 8.9 Hz, 1H), 7.80 (d, J = 16.3 Hz, 1H), 7.52 (dd, J = 8.9, 2.1 Hz, 1H), 7.39 (d, J = 8.7 Hz, 2H), 7.05 (d, J = 8.6 Hz, 2H), 5.16 (s, 1H), 5.13 (brs, 2H), 4.86 (brs, 1H), 4.84 (d, J = 7.7 Hz, 1H), 4.64 (br s, 1H), 4.51 (brs, 1H), 3.70 (d, J = 2.7 Hz, 1H), 3.59 – 3.53 (m, 3H), 3.51 – 3.46 (m, 1H), 3.42 – 3.39 (m, 1H). ^{19}F NMR (377 MHz, DMSO-*d*₆) δ -61.29. ^{13}C NMR (101 MHz, DMSO) δ 163.7, 157.4, 153.5, 149.0, 138.3, 137.5, 135.4, 133.0, 130.9 (q, J = 33 Hz), 129.9, 129.5, 128.0 (q, J = 2 Hz), 125.8, 123.3 (q, J = 273 Hz), 123.0, 121.9 – 121.8 (m), 118.4, 116.2, 110.0, 100.9, 75.5, 73.3, 70.3, 68.1, 65.8, 60.4, 40.2, 39.9, 39.7, 39.5, 39.3, 39.1, 38.9. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for C₃₁H₂₇F₆N₂O₈S 701.1392; Found 701.1388.

Chapter 4

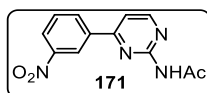
Synthesis of compound **166**



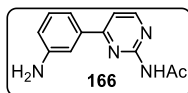
(i) In a 250 mL round bottom flask was added 16.5 g (100 mmol) of 3-nitroacetophenone, 26.5 mL of N,N-Dimethylformamide dimethyl acetal (200 mmol) and 100 mL of xylene. The mixture was heated to 140°C for 12 hours. Then the reaction mixture was cooled to room temperature and the formed precipitate was filtered off, washed with petroleum ether and diethyl ether followed by purification by chromatography using Ethyl acetate/Cyclohexene (1/3) to afford 19.6 g of a yellow solid. Yield: 89%. ¹H NMR (200 MHz, DMSO-*d*₆) δ 8.59 (t, 1H), 8.42 – 8.23 (m, 1H), 7.81 (d, *J* = 12.1 Hz, 1H), 7.70 (t, *J* = 7.9 Hz, 1H), 5.89 (d, *J* = 12.1 Hz, 1H), 3.17 (s, 1H), 2.95 (s, 1H). ¹³C NMR (50 MHz, DMSO-*d*₆) δ 182.8, 155.2, 147.9, 141.6, 133.4, 129.8, 125.1, 121.5, 90.3, 44.7, 37.3.



(ii) In a 250 mL round bottom flask was added 17.6 g (80 mmol) of compound **169**, 11.5 g of guanidine hydrochloride (120 mmol), 6 g (120 mmol) of sodium hydroxide. The mixture was suspended in 100 mL of n-butanol and heated to reflux during 24 hours. After, the reaction was cooled to room temperature and the formed precipitate was filtered in vacuum, washed with water and methanol to afford 15.9 g of a light grey powder. Yield: 92%. ^1H NMR (200 MHz, DMSO- d_6) δ 8.88 (s, 1H), 8.46 (d, $J = 7.9$ Hz, 1H), 8.39 (d, $J = 5.1$ Hz, 1H), 8.31 (dd, $J = 8.0, 1.8$ Hz, 1H), 7.77 (t, $J = 8.0$ Hz, 1H), 7.25 (d, $J = 5.2$ Hz, 1H), 6.90 (s, 2H). ^{13}C NMR (50 MHz, DMSO- d_6) δ 163.9, 161.1, 159.8, 148.3, 138.7, 132.9, 130.4, 124.9, 121.0, 105.9.



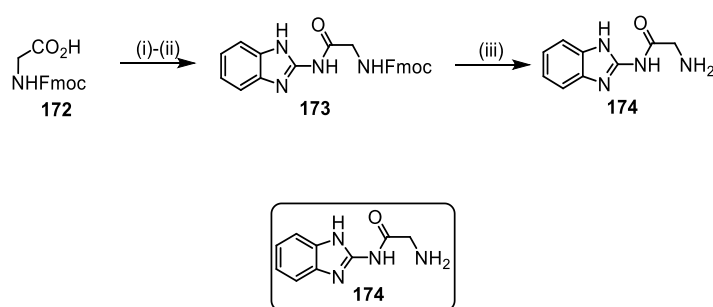
(iii) In a 250 mL round bottom flask was added 13.0 g (60 mmol) of compound **170** and 100 mL of pyridine. The reaction mixture was cooled to 0°C and it was slowly added 14.2 mL (200 mmol) of acetyl chloride. After the addition, the reaction was heated to 50°C during 12 hours. After that, the volatiles were removed under reduced pressure and the crude was washed several times with water to remove the formed salts. The residue was suspended in hot ethanol, filtered in vacuum and the filtrate was slowly cooled to room temperature to afford 4.2 g of compound **171** as grey crystals. Yield: 27%. ^1H NMR (200 MHz, DMSO- d_6) δ 10.73 (s, 1H), 9.02 (s, 1H), 8.80 (d, $J = 4.9$ Hz, 1H), 8.63 (d, $J = 7.0$ Hz, 1H), 8.41 (d, $J = 8.2$ Hz, 1H), 7.93 - 7.82 (m, 2H), 2.26 (s, 3H). ^{13}C NMR (101 MHz, DMSO- d_6) δ 169.1, 161.4, 160.0, 158.0, 148.5, 137.7, 133.3, 130.6, 125.6, 121.6, 112.1, 24.8. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{12}\text{H}_{11}\text{N}_4\text{O}_3$ 259.0831; Found: 259.0825



(iv) 4.0 g (15.5 mmol) of compound **171** and 500 mg of Pd/C were suspended in 60 mL of methanol. To this mixture was slowly added 1.9 g (51.5 mmol) of NaBH_4 . The reaction was stirred during 3 hours at room temperature followed by filtration over celite. The filtrate was concentrated in vacuum and after was diluted in 100 mL of ethyl acetate and extracted with saturated solution of NaHCO_3 followed by purification by column chromatography using

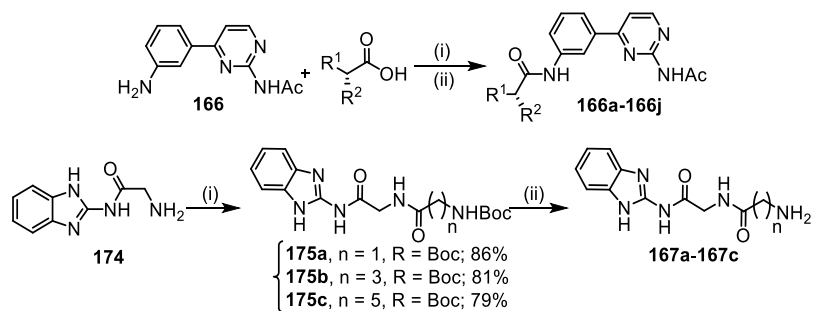
cyclohexane/ethyl acetate (1/1) to afford 2.5 g of a white solid. Yield: 71%. ^1H NMR (400 MHz, DMSO- d_6) δ 10.47 (s, 1H), 8.63 (d, $J = 5.3$ Hz, 1H), 7.52 (d, $J = 5.3$ Hz, 1H), 7.36 (t, $J = 1.8$ Hz, 1H), 7.28 (d, $J = 7.8$ Hz, 1H), 7.17 (t, $J = 7.8$ Hz, 1H), 6.74 (dd, $J = 7.9, 1.5$ Hz, 1H), 5.29 (s, 2H), 2.28 (s, 3H). ^{13}C NMR (101 MHz, DMSO- d_6) δ 169.4, 164.8, 158.9, 157.9, 149.2, 136.6, 129.4, 116.7, 114.7, 112.1, 111.6, 24.9. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{12}\text{H}_{11}\text{N}_4\text{O}_3$ 229.1089; Found: 229.1084.

Synthesis of compound **4.9**



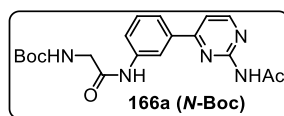
(i)-(iii) In a 25 mL round bottom flask was added 5g (16.8 mmol) of *N*-Fmoc-glycine and it was suspended in 10 mL of a mixture of DCM/THF (4/6). After was added 2 mL of thionyl chloride and the mixture was refluxed during 30 min. The volatiles were removed in reduced pressure, followed by co-evaporation with hexane to remove the excess of thionyl chloride. The crude was dissolved in 20 mL of anhydrous DMF, and it was added 4.6 mL of DIPEA (2 eq), 160 mg of DMAP (1,6mmol) and 1.3 g of 2-aminobenzimidazole (16.8 mmol). The reaction was heated at 100°C during 24 hours. After, the solvent was removed in reduced pressure and the crude was dissolved in 10 mL of 20% solution of piperidine in DMF. The mixture was stirred for 2 hours at room temperature. The formed precipitate was filtered off, washed with methanol to afford 1.8 g of compound **174** as white powder. Yield: 57%. ^1H NMR (200 MHz, MeOD- d_4) δ 7.48 – 7.44 (m, 2H), 7.19 – 7.14 (m, 2H), 3.54 (s, 2H). ^{13}C NMR (101 MHz, DMSO) δ 173.3, 146.3, 120.9, 114.2, 45.2. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_9\text{H}_{11}\text{N}_4\text{O}$ 191.0933; Found: 191.0928.

General procedure for the synthesis of compounds 166a-166j and 167a-167c

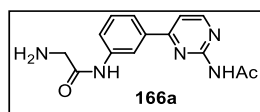


(i) In an oven dried 25 mL round bottom flask purged with argon was added 0.75 mmol of *N*-Boc protected amino acid, 0.75 mmol of HBTU (284 mg), 0.1 mmol of HOBt (13.5 mg), 2.25 mmol of DIPEA and 5 mL of DMF. The mixture was stirred at room temperature for 30 minutes followed by the addition of 0.5 mmol (114 mg) of precursor **166** or **174**. The temperature was raised to 40 °C and the reaction was kept during 24 h. After, the solvent was removed in reduced pressure and the reaction crude was purified by recrystallization from methanol to afford the *N*-Boc derivatives of **166a-166j** and **167a-167c** as white crystals, that were characterized by ¹H NMR and HRMS.

(ii) In a 10 mL round bottom flask was added 0.1 mmol of *N*-Boc protected aminoconjugates, 4 mL of dichloromethane, 5 % of water and 1 mL of trifluoroacetic acid. The mixture was stirred at room temperature overnight. The solvent was removed by reduced pressure followed by co-evaporation with 50 mL of toluene (three times). Then the crude was dissolved in methanol/ethyl acetate and was precipitated after addition of diethyl ether to afford the final ligands in quantitative yield.

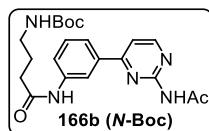


Yield: 71%. ¹H NMR (200 MHz, MeOD-*d*₄) δ 8.62 (d, *J* = 5.4 Hz, 1H), 8.45 (s, 1H), 7.90 (d, *J* = 7.8 Hz, 1H), 7.67 (d, *J* = 8.0 Hz, 1H), 7.60 (d, *J* = 5.3 Hz, 1H), 7.46 (t, *J* = 7.9 Hz, 1H), 3.90 (s, 2H), 2.35 (s, 3H), 1.48 (s, 9H). HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₁₉H₂₄N₅O₄ 386.1828; Found: 386.1822.

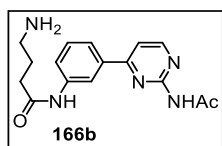


¹H NMR (200 MHz, MeOD-*d*₄) δ 8.63 (d, *J* = 5.3 Hz, 1H), 8.29 (s, 1H), 7.86 (d, *J* = 7.7 Hz, 1H), 7.67 - 7.62 (m, 2H), 7.52 (t, *J* = 7.9 Hz, 1H), 3.98 (s, 2H), 2.31 (s, 3H). ¹³C NMR (50

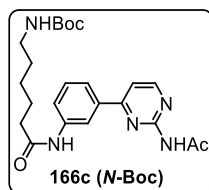
MHz, MeOD-*d*₄) δ 173.6, 166.2, 166.0, 159.9, 158.01, 138.9, 137.6, 130.9, 124.8, 124.2, 119.9, 114.1, 42.1, 24.7. HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₁₄H₁₆N₅O₂ 286.1304; Found: 286.1298.



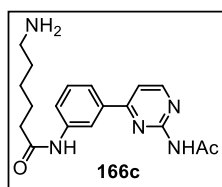
Yield: 59%. ¹H NMR (200 MHz, MeOD-*d*₄) δ 8.62 (d, *J* = 5.3 Hz, 1H), 8.45 (s, 1H), 7.89 (d, *J* = 7.6 Hz, 1H), 7.67 (d, *J* = 8.3 Hz, 1H), 7.60 (d, *J* = 5.4 Hz, 1H), 7.45 (t, *J* = 7.9 Hz, 1H), 3.14 (t, *J* = 6.7 Hz, 2H), 2.44 (t, *J* = 7.5 Hz, 2H), 2.35 (s, 3H), 1.92 – 1.80 (m, 2H), 1.43 (s, 9H). HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₂₁H₂₈N₅O₄ 414.2141; Found: 414.2137.



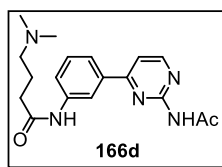
¹H NMR (200 MHz, MeOD-*d*₄) δ 8.63 (d, *J* = 5.3 Hz, 1H), 8.28 (s, 1H), 7.84 (d, *J* = 7.7 Hz, 1H), 7.66 - 7.60 (m, 2H), 7.49 (t, *J* = 7.9 Hz, 1H), 3.09 (t, *J* = 7.4 Hz, 2H), 2.60 (t, *J* = 7.2 Hz, 2H), 2.32 (s, 3H), 2.13 – 1.99 (m, 2H). ¹³C NMR (50 MHz, MeOD-*d*₄) δ 173.6, 173.6, 166.2, 159.9, 158.2, 139.6, 137.6, 130.7, 124.5, 124.4, 120.0, 114.0, 40.1, 34.3, 24.8, 23.9. HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₁₆H₂₀N₅O₂ 314.1617; Found: 314.1612.



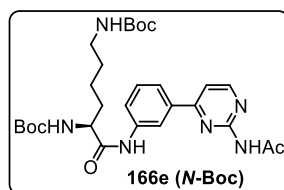
Yield: 62%. ¹H NMR (200 MHz, MeOD-*d*₄) δ 8.63 (d, *J* = 5.2 Hz, 1H), 8.45 (s, 1H), 7.89 (d, *J* = 7.7 Hz, 1H), 7.67 (d, *J* = 7.9 Hz, 1H), 7.60 (d, *J* = 5.3 Hz, 1H), 7.45 (t, *J* = 7.9 Hz, 1H), 3.04 (t, *J* = 5.1 Hz, 2H), 2.42 (t, *J* = 7.1 Hz, 2H), 2.35 (s, 3H), 1.81 – 1.67 (m, 2H), 1.56 – 1.41 (m, 13H). HRMS (ESI) *m/z*: [M + H]⁺ Calcd for C₂₃H₃₂N₅O₄ 442.2454; Found: 442.2448.



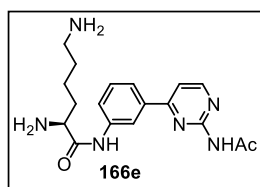
^1H NMR (200 MHz, MeOD- d_4) δ 8.63 (s, 1H), 8.25 (s, 1H), 7.85 (d, J = 7.3 Hz, 1H), 7.63 (d, J = 6.3 Hz, 2H), 7.50 (t, J = 7.4 Hz, 1H), 3.00 (t, J = 7.4 Hz, 2H), 2.47 (t, J = 7.1 Hz, 2H), 2.31 (s, 3H), 1.79 - 1.66 (m, 2H), 1.55 - 1.42 (m, 2H). ^{13}C NMR (50 MHz, MeOD- d_4) δ 175.5, 173.7, 166.2, 159.6, 157.9, 139.5, 137.3, 130.7, 124.6, 124.5, 120.1, 114.1, 40.3, 37.3, 27.8, 26.6, 25.8, 24.8. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{18}\text{H}_{24}\text{N}_5\text{O}_2$ 342.1930; Found: 342.1925.



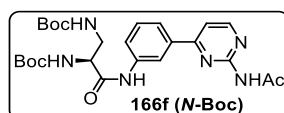
^1H NMR (400 MHz, TFAd) δ 8.78 (s, 1H), 8.68 (d, J = 6.7 Hz, 1H), 8.10 (d, J = 6.7 Hz, 1H), 8.06 (d, J = 7.7 Hz, 1H), 7.69 (d, J = 8.2 Hz, 1H), 7.62 (t, J = 7.9 Hz, 1H), 3.37 (dd, J = 13.3, 6.7 Hz, 2H), 3.03 (d, J = 5.1 Hz, 6H), 2.85 (t, J = 6.9 Hz, 2H), 2.56 (s, 3H), 2.34 - 2.26 (m, 2H). ^{13}C NMR (101 MHz, TFAd) δ 179.5, 176.6, 175.9, 153.1, 148.1, 139.6, 135.6, 132.8, 131.0, 129.0, 124.8, 115.7, 60.7, 45., 35.4, 24.7, 22.3. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{18}\text{H}_{24}\text{N}_5\text{O}_2$ 342.1930; Found: 342.1925.



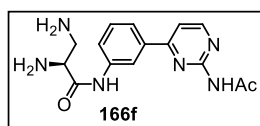
Yield: 53%. ^1H NMR (200 MHz, MeOD- d_4) δ 8.63 (d, J = 5.3 Hz, 1H), 8.46 (s, 1H), 7.90 (d, J = 7.6 Hz, 1H), 7.69 (d, J = 8.3 Hz, 1H), 7.61 (d, J = 5.3 Hz, 1H), 7.46 (t, J = 7.9 Hz, 1H), 4.17 (t, J = 6.3 Hz, 1H), 3.05 (t, J = 5.1 Hz, 2H), 2.36 (s, 3H), 1.84 - 1.69 (m, 2H), 1.59 - 1.41 (m, 22H). HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{28}\text{H}_{41}\text{N}_6\text{O}_6$ 557.3088; Found: 557.3089



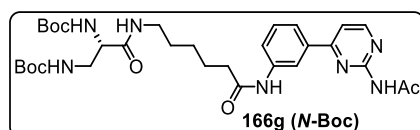
^1H NMR (200 MHz, MeOD- d_4) δ 8.67 (d, J = 5.4 Hz, 1H), 8.36 (s, 1H), 7.93 (d, J = 7.7 Hz, 1H), 7.70 - 7.67 (m, 2H), 7.57 (t, J = 7.9 Hz, 1H), 4.16 (t, J = 6.5 Hz, 1H), 3.01 (t, J = 7.6 Hz, 3H), 2.32 (s, 3H), 2.12 - 2.00 (m, 2H), 1.84 - 1.70 (m, 2H), 1.64 - 1.52 (m, 2H). ^{13}C NMR (50 MHz, MeOD- d_4) δ 173.7, 168.9, 166.1, 159.9, 158.1, 138.7, 137.7, 130.9, 125.3, 124.6, 120.3, 114.2, 54.7, 40.1, 31.8, 27.7, 24.7, 22.6. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{18}\text{H}_{25}\text{N}_6\text{O}_2$ 357.2039; Found: 357.2034.



Yield: 67%. ^1H NMR (400 MHz, MeOD- d_4) δ 8.64 (d, J = 5.2 Hz, 1H), 8.47 (s, 1H), 7.92 (d, J = 7.2 Hz, 1H), 7.68 (d, J = 7.1 Hz, 1H), 7.61 (d, J = 5.2 Hz, 1H), 7.47 (t, J = 7.7 Hz, 1H), 4.32 (t, J = 8.0 Hz, 1H), 3.52 - 3.41 (m, 2H), 2.36 (s, 3H), 1.47 (s, 9H), 1.40 (s, 9H). HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{25}\text{H}_{35}\text{N}_6\text{O}_6$ 515.2618; Found: 515.2617.

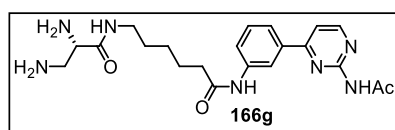


^1H NMR (200 MHz, MeOD- d_4) δ 8.64 (d, J = 5.4 Hz, 1H), 8.34 (s, 1H), 7.91 (d, J = 7.8 Hz, 1H), 7.72 - 7.64 (m, 2H), 7.55 (t, J = 7.9 Hz, 1H), 4.24 (t, J = 6.8 Hz, 1H), 3.64 - 3.40 (m, 2H), 2.31 (s, 3H). ^{13}C NMR (50 MHz, MeOD- d_4) δ 173.6, 168.5, 166.0, 159.9, 157.9, 138.6, 137.5, 130.9, 125.3, 124.7, 120.3, 114.1, 53.1, 41.8, 24.7. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{15}\text{H}_{19}\text{N}_6\text{O}_2$ 315.1569; Found: 315.1564.

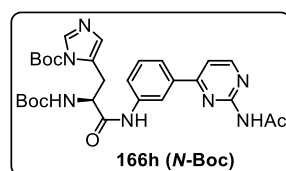


Yield: 64%. ^1H NMR (400 MHz, MeOD- d_4) δ 8.63 (d, J = 4.8 Hz, 1H), 8.45 (s, 1H), 7.89 (d, J = 7.7 Hz, 1H), 7.68 (d, J = 7.9 Hz, 1H), 7.61 (d, J = 4.9 Hz, 1H), 7.46 (t, J = 7.9 Hz, 1H), 4.11 (t, 1H), 3.24 - 3.17 (m, 4H), 2.42 (t, J = 7.2 Hz, 2H), 2.36 (s, 3H), 1.76 - 1.73 (m, 2H),

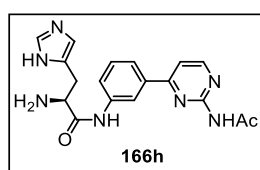
1.59 – 1.56 (m, 2H), 1.45 - 1.43 (m, 20H). HRMS (ESI) m/z : $[M + H]^+$ Calcd for $C_{31}H_{46}N_7O_7$ 628.3459; Found: 628.3458.



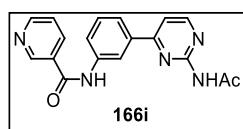
1H NMR (200 MHz, MeOD- d_4) δ 8.65 (d, $J = 5.4$ Hz, 1H), 8.46 (s, 1H), 7.90 (d, $J = 8.0$ Hz, 1H), 7.69 - 7.61 (m, 2H), 7.47 (t, $J = 7.9$ Hz, 1H), 4.03 (t, $J = 5.8$ Hz, 1H), 3.25 – 3.17 (m, 4H), 2.45 (t, $J = 7.1$ Hz, 2H), 2.36 (s, 3H), 1.80 – 1.40 (m, 6H). ^{13}C NMR (101 MHz, MeOD) δ 174.7, 174.6, 167.6, 158.3, 158.3, 140.6, 138.8, 130.3, 124.0, 123.7, 119.9, 107.8, 61.6, 53.2, 40.4, 37.7, 29.9, 27.4, 26.4, 14.4. HRMS (ESI) m/z : $[M + H]^+$ Calcd for $C_{21}H_{30}N_7O_3$ 428.2410; Found: 428.2405.



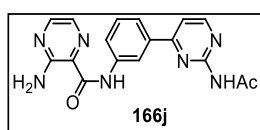
Yield: 48%. 1H NMR (200 MHz, MeOD- d_4) δ : 8.63 (d, $J = 5.3$ Hz, 1H), 8.43 (s, 1H), 8.16 (s, 1H), 7.91 (d, $J = 7.8$ Hz, 1H), 7.67 (d, $J = 8.0$ Hz, 1H), 7.61 (d, $J = 5.4$ Hz, 1H), 7.47 (t, $J = 7.9$ Hz, 1H), 7.34 (s, 1H), 4.50 (t, $J = 7.8$ Hz, 1H), 3.12 - 2.92 (m, 2H), 1.58 (s, 9H), 1.42 (s, 9H). HRMS (ESI) m/z : $[M + H]^+$ Calcd for $C_{28}H_{36}N_7O_6$ 566.2727; Found: 566.2723.



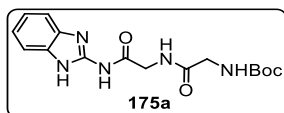
1H NMR (200 MHz, MeOD- d_4) δ 8.65 (d, $J = 5.3$ Hz, 1H), 8.59 (s, 1H), 8.29 (s, 1H), 7.91 (d, $J = 7.2$ Hz, 1H), 7.67 - 7.50 (m, 3H), 7.43 (s, 1H), 4.41 (t, $J = 6.9$ Hz, 1H), 3.47 (d, $J = 6.9$ Hz, 2H), 2.32 (s, 3H). ^{13}C NMR (50 MHz, MeOD- d_4) δ 173.6, 167.6, 165.9, 160.0, 158.1, 138.5, 137.7, 135.8, 130.9, 128.5, 125.4, 124.6, 120.3, 119.1, 114.1, 54.1, 27.9, 24.7. HRMS (ESI) m/z : $[M + H]^+$ Calcd for $C_{18}H_{19}N_7O_2$ 366.1678; Found: 366.1673.



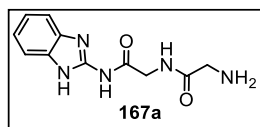
^1H NMR (200 MHz, $\text{D}_2\text{O}/\text{TFAd}$) δ 8.72 (s, 1H), 8.49 (d, $J = 8.0$ Hz, 1H), 8.40 (d, $J = 5.6$ Hz, 1H), 8.05 (d, $J = 6.7$ Hz, 1H), 7.91 (s, 1H), 7.69 – 7.62 (m, 1H), 7.47 (d, $J = 6.7$ Hz, 1H), 7.26 (d, $J = 8.1$ Hz, 1H), 7.02 (t, $J = 7.9$ Hz, 1H), 1.78 (s, 1H). ^{13}C NMR (50 MHz, $\text{D}_2\text{O}/\text{TFAd}$) δ 179.5, 175.8, 153.1, 149.9, 148.4, 138.8, 137.2, 135.9, 133.7, 133.1, 132.6, 131.0, 129.5, 124.6, 120.4, 115.9, 24.8. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{18}\text{H}_{16}\text{N}_5\text{O}_2$ 334.1304; Found: 334.1298.



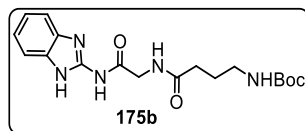
^1H NMR (200 MHz, $\text{D}_2\text{O}/\text{TFAd}$) δ 7.71 (d, $J = 6.7$ Hz, 1H), 7.58 (s, 1H), 7.20 (d, $J = 3.6$ Hz, 1H), 7.09 - 7.01 (m, 3H), 6.91 (d, $J = 8.6$ Hz, 1H), 6.58 (t, $J = 8.1$ Hz, 1H), 1.46 (s, 3H). ^{13}C NMR (101 MHz, $\text{D}_2\text{O}/\text{TFAd}$) δ 179.5, 175.9, 153.2, 150.0, 148.3, 138.9, 137.2, 135.9, 133.7, 133.1, 132.6, 131.0, 129.5, 124.5, 115.9, 24.9. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{17}\text{H}_{16}\text{N}_7\text{O}_2$ 350.1365; Found: 350.1368.



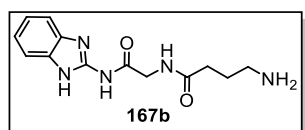
Yield: 86%. ^1H NMR (200 MHz, $\text{MeOD}-d_4$) δ 7.48 - 7.44 (m, 2H), 7.19 – 7.15 (m, 2H), 4.15 (s, 2H), 3.81 (s, 2H), 1.46 (s, 9H). HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{16}\text{H}_{22}\text{N}_5\text{O}_4$ 348.1672; Found: 348.1668.



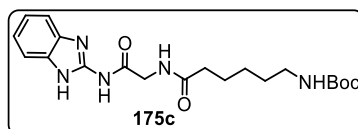
^1H NMR (200 MHz, $\text{MeOD}-d_4$) δ 7.51 - 7.46 (m, 2H), 7.25 – 7.19 (m, 2H), 4.23 (s, 2H), 3.82 (s, 2H). ^{13}C NMR (50 MHz, $\text{MeOD}-d_4$) δ 171.2, 168.1, 147.9, 135.7, 124.8, 123.5, 114.9, 112.3, 43.9, 41.5. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{11}\text{H}_{14}\text{N}_5\text{O}_2$ 248.1147; Found: 248.1142.



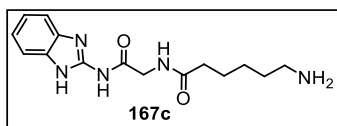
Yield: 81%. ^1H NMR (200 MHz, MeOD- d_4) δ 7.48 - 7.43 (m, 2H), 7.19 - 7.14 (m, 2H), 4.12 (s, 2H), 3.12 (t, $J = 6.7$ Hz, 2H), 2.34 (t, $J = 7.3$ Hz, 2H), 1.87 - 1.75 (m, 2H), 1.44 (s, 9H). HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{18}\text{H}_{26}\text{N}_5\text{O}_4$ 376.1985; Found: 376.1981.



^1H NMR (200 MHz, MeOD- d_4 /TFAd) δ 7.66 - 7.59 (m, 2H), 7.51 - 7.48 - 7.42 (m, 2H), 4.26 (s, 2H), 3.06 (t, $J = 7.5$ Hz, 2H), 2.52 (t, $J = 7.2$ Hz, 2H), 2.08 - 1.94 (m, 2H). ^{13}C NMR (50 MHz, MeOD- d_4 /TFAd) δ 176.3, 170.9, 162.4, 161.7, 160.6, 159.8, 143.9, 129.5, 126.7, 125.2, 119.5, 114.1, 113.8, 108.1, 44.1, 40.2, 33.2, 24.0. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{13}\text{H}_{18}\text{N}_5\text{O}_2$ 276.1460; Found: 276.1455.



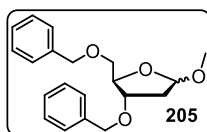
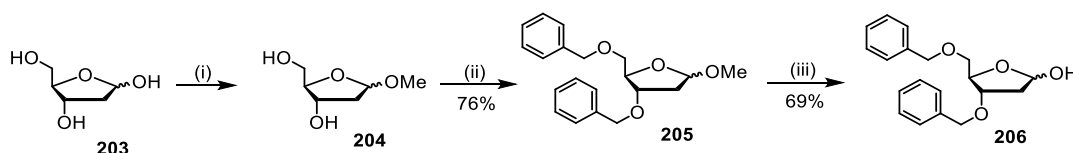
Yield: 79%. NMR (200 MHz, MeOD- d_4) δ 7.48 - 7.44 (m, 2H), 7.19 - 7.14 (m, 2H), 4.11 (s, 2H), 3.04 (t, $J = 6.3$ Hz, 2H), 2.32 (t, $J = 7.3$ Hz, 2H), 1.74 - 1.60 (m, 2H), 1.56 - 1.42 (m, 13H). HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{20}\text{H}_{30}\text{N}_5\text{O}_4$ 404.2298; Found: 404.2302.



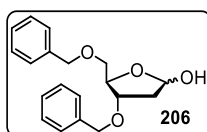
^1H NMR (200 MHz, MeOD- d_4 /TFAd) δ 7.64 - 7.57 (m, 2H), 7.47 - 7.40 (m, 2H), 4.20 (s, 2H), 2.93 (t, $J = 7.4$ Hz, 2H), 2.38 (t, $J = 7.2$ Hz, 2H), 1.77 - 1.62 (m, 2H), 1.51 - 1.36 (m, 2H). ^{13}C NMR (50 MHz, MeOD- d_4 /TFAd) δ 177.3, 171.1, 161.6, 160.8, 160.0, 159.2, 144.7, 129.8, 126.7, 125.2, 119.5, 114.2, 113.8, 108.1, 44.3, 40.6, 36.2, 28.2, 26.9, 25.9. HRMS (ESI) m/z : $[\text{M} + \text{H}]^+$ Calcd for $\text{C}_{15}\text{H}_{22}\text{N}_5\text{O}_2$ 304.1773; Found: 304.1768.

Chapter 5

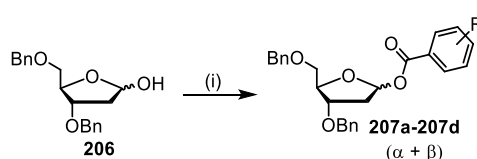
Synthesis of intermediate **206**



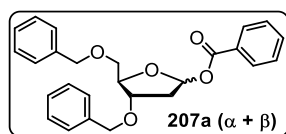
Compound **203** (5g, 38 mmol), was dissolved in dry THF (50 mL) and treated with BnCl (30.6 mL, 265 mmol) in presence of KOH (6.5 g, 114 mmol). The reaction was heated at reflux during 24 hours. Next, the reaction was diluted with 100 mL of DCM and extracted 3 times with water and washed with brine. The solvent was removed in *vacuo* and the crude was purified over silica gel, eluting with Cy/EA (0-20 % EA) to get compound **205** as a colorless oil. Yield: 76 % (9.5 g). ^1H NMR (200 MHz, CDCl_3) δ 7.35 – 7.31 (m, 10H), 5.13 – 5.07 (m, 1H), 4.61 – 4.46 (m, 2H), 4.31 – 4.23 (m, 1H), 4.20 – 3.96 (m, 1H), 3.55 – 3.52 (m, 2H), 3.42 (s, 1H), 3.32 (s, 1H), 2.31 – 1.99 (m, 2H). ^{13}C NMR (101 MHz, CDCl_3) δ 138.3, 138.2, 138.1, 138.1, 128.4, 128.4, 127.9, 127.7, 127.7, 127.7, 127.7, 105.5, 105.2, 82.9, 82.2, 80.0, 78.6, 73.5, 73.3, 72.1, 71.6, 71.6, 70.2, 55.2, 55.0, 39.4, 38.9.



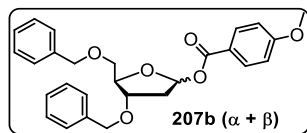
Compound **205** (9 g, 27.4 mmol) was dissolved in a mixture of acetic acid and water (8/2) and heated to 60°C during 48 hours. The solvent was evaporated in *vacuo*, and co-evaporated with toluene (100 mL) three times. The crude was purified by flash column chromatography in silica gel, using Cy/EA (60/40) as the eluent to afford compound **206** as a colorless oil. Yield: 69% (5.9 g). ¹H NMR (400 MHz, CDCl₃) δ 7.36 – 7.26 (m, 10H), 5.55 (t, *J* = 3.4 Hz, 0.4H), 5.48 (s, 0.59H), 4.62 – 4.47 (m, 5H), 4.26 (m, 0.65H), 4.13 – 4.11 (m, 0.55H), 3.67 – 3.35 (m, 3H), 2.23 – 2.21 (m, 1H), 2.13 – 2.11 (m, 1H). ¹³C NMR (101 MHz, CDCl₃) δ 138.1, 137.9, 137.5, 137.3, 128.7, 128.6, 128.6, 128.5, 128.1, 128.0, 128.0, 127.9, 127.8, 127.8, 127.7, 99.5, 99.5, 83.3, 82.6, 80.1, 79.8, 73.8, 73.6, 71.6, 71.4, 70.5, 70.5, 41.9, 39.2.



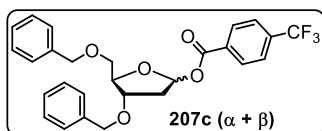
(i) Compound **206** (0.5 g, 1.6 mmol) was dissolved in dry pyridine. The solution was cooled to 0°C in ice bath and treated with the corresponding benzoyl chloride (2.4 mmol). After complete addition of the acylating agent, the reaction was stirred at room temperature until complete conversion of the starting material (TLC monitoring). The solvent was removed in *vacuo*, co-evaporated with toluene (3 x 10 mL) and the crude was purified by flash column chromatography on silica gel, eluting with Cy/EA (0-20 % EA) to afford the corresponding benzoyl aglycone as a colorless oil.



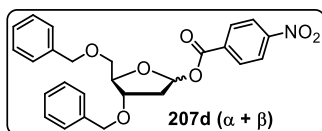
Yield: 66%. ¹H NMR (200 MHz, CDCl₃) δ 8.10 – 8.04 (m, 1H), 7.97 – 7.92 (m, 1H), 7.59 – 7.48 (m, 1H), 7.44 – 7.28 (m, 10H), 6.65 – 6.62 (m, 1H), 4.63 – 4.49 (m, 4.5H), 4.42 – 4.35 (m, 1H), 4.26 – 4.21 (m, 0.5H), 3.64 – 3.59 (m, 2H), 2.51 – 2.39 (m, 2H). ¹³C NMR (50 MHz, CDCl₃) δ 166.0, 165.6, 138.1, 138.1, 133.2, 133.1, 130.4, 130.3, 130.0, 129.8, 128.6, 128.5, 128.5, 128.5, 128.4, 127.9, 127.8, 127.8, 127.8, 127.7, 99.6, 99.4, 85.2, 84.4, 79.1, 79.1, 73.7, 73.5, 71.9, 71.5, 70.9, 70.2, 39.2, 38.3.



Yield: 78%. ^1H NMR (200 MHz, CDCl_3) δ 8.01 (d, $J = 9.0$ Hz, 1H), 7.90 (d, $J = 9.0$ Hz, 1H), 7.39 – 7.26 (m, 10H), 6.86 (dd, $J = 9.0, 0.6$ Hz, 2H), 6.63 – 6.59 (m, 1H), 4.57 – 4.50 (m, 4.5H), 4.40 – 4.36 (m, 1H), 4.25 – 4.19 (m, 0.5H), 3.85 (d, $J = 0.8$ Hz, 1H), 3.64 – 3.58 (m, 2H), 2.49 – 2.37 (m, 2H). ^{13}C NMR (50 MHz, CDCl_3) δ 165.7, 165.3, 163.6, 163.6, 138.2, 138.1, 132.1, 131.9, 128.6, 128.5, 128.5, 128.5, 127.9, 127.8, 127.8, 127.8, 122.8, 122.5, 113.7, 113.6, 99.3, 99.0, 85.0, 84.2, 79.1, 79.1, 73.7, 73.5, 71.9, 71.4, 70.9, 70.2, 55.6, 55.5, 39.2, 38.3.

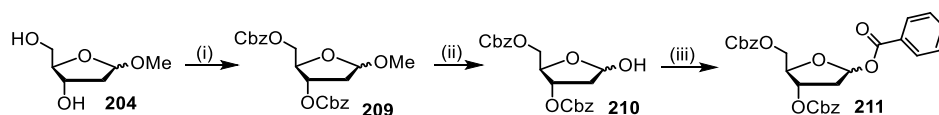


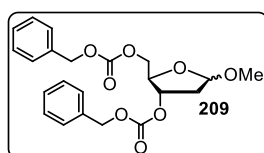
Yield: 46%. ^1H NMR (400 MHz, CDCl_3) δ 8.16 (d, $J = 8.1$ Hz, 1H), 8.03 (d, $J = 8.1$ Hz, 1H), 7.64 (t, $J = 8.5$ Hz, 2H), 7.38 – 7.27 (m, 10H), 6.67 – 6.64 (m, 1H), 4.60 – 4.55 (m, 4.5H), 4.46 – 4.39 (m, 1H), 4.28 – 4.25 (m, 1H), 3.67 – 3.58 (m, 2H), 2.57 – 2.39 (m, 2H). ^{13}C NMR (101 MHz, CDCl_3) δ 164.8, 164.5, 138.0, 137.9, 133.7, 133.4, 130.4, 130.2, 128.6, 128.6, 128.6, 128.5, 128.0, 127.9, 127.9, 127.9, 127.8, 127.8, 127.8, 127.7, 125.5 – 125.4 (m), 100.0, 99.9, 85.5, 84.5, 79.0, 78.6, 73.7, 73.5, 72.1, 71.5, 70.7, 70.2, 39.2, 38.4.



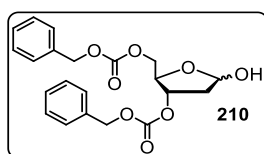
Yield: 36%. ^1H NMR (400 MHz, CD_2Cl_2) δ 8.24 – 8.10 (m, 3H), 8.08 – 8.00 (m, 1H), 7.43 – 7.18 (m, 10H), 6.58 (dd, $J = 5.7, 2.4$ Hz, 1H), 4.61 – 4.44 (m, 4H), 4.43 – 4.31 (m, 1.3H), 4.25 – 4.22 (m, 0.6H), 3.65 – 3.52 (m, 0.4H), 2.58 – 2.33 (m, 1.6H).

Synthesis of compound **211**



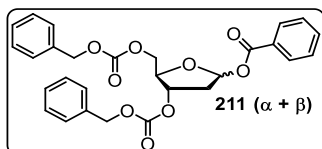


(i) Compound **204** (5g, 33.7 mmol), was dissolved in dry DCM (50 mL) and treated with CbzCl (22 mL, 152 mmol) in presence of DMAP (24.6 g, 202 mmol). The reaction was stirred at room temperature during 24 hours. The reaction was diluted with 100 mL of DCM and extracted 3 times with saturated NaHCO₃ solution, washed with brine and dried with magnesium sulfate. The solvent was removed in *vacuo* and the crude was purified over silica gel, eluting with Cy/EA (0-20 % EA) to get compound **209** as a colorless oil. Yield: 71 % (9.9 g). ¹H NMR (400 MHz, CDCl₃) δ 7.41 – 7.33 (m, 10H), 5.22 – 5.16 (m, 6H), 4.85 (t, *J* = 2.9 Hz, 1H), 3.94 – 3.84 (m, 2H), 3.36 (s, 1H), 2.21 (td, *J* = 12.4, 3.4 Hz, 1H), 1.98 (dd, *J* = 12.7, 4.7 Hz, 1H). ¹³C NMR (50 MHz, CDCl₃) δ 154.9, 154.2, 135.1, 135.1, 128.7, 128.6, 128.5, 128.4, 127.7, 127.1, 98.6, 71.4, 69.9, 69.9, 65.4, 60.5, 55.2, 30.9.



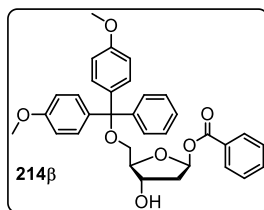
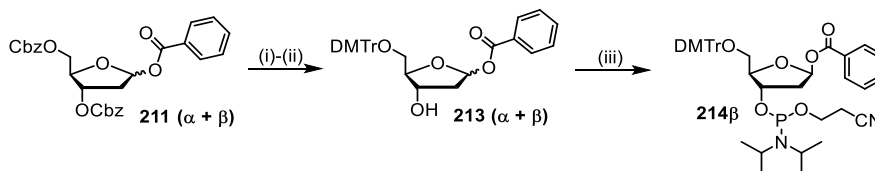
(ii) Compound **209** (6 g, 14.4 mmol) was dissolved in a mixture of acetic acid and water (8/2) and heated to 60°C during 48 hours. The solvent was evaporated in *vacuo*, and co-evaporated with toluene (100 mL) three times. The crude was purified by flash column chromatography in silica gel, using Cy/EA (60/40) as the eluent to afford compound **210** as a white solid. Yield: 64 % (3.7 g). ¹H NMR (200 MHz, CDCl₃) δ 7.39 (d, *J* = 6.6 Hz, 10H), 5.62 (dd, *J* = 4.8, 3.5 Hz, 0.42H), 5.56 (d, *J* = 4.7 Hz, 0.56H), 5.17 – 5.06 (m, 5H), 4.52 – 4.21 (m, 3H), 2.94 (br s, 1H), 2.39 – 2.10 (m, 2H). ¹³C NMR (50 MHz, CDCl₃) δ 155.1, 155.0, 154.7,

154.6, 135.1, 134.9, 128.8, 128.8, 128.8, 128.7, 128.6, 128.5, 128.5, 99.4, 98.9, 81.7, 81.5, 78.6, 78.2, 70.2, 70.1, 70.0, 68.4, 67.3, 40.0, 39.6.



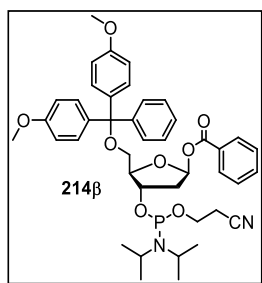
(iii) Compound **210** (3.5 g, 8.7 mmol) was dissolved in dry pyridine. The solution was cooled to 0 °C (ice bath) and then 1.2 mL (13.0 mmol) of benzoyl chloride was added dropwise. The reaction was stirred overnight at room temperature. The solvent was removed in *vacuo*, and the crude was purified by flash column chromatography on silica gel, eluting with Cy/EA (0-20 % EA) to afford **211** as a colorless oil. Yield: 91 % (4.0 g, $\alpha + \beta$). ¹H NMR (200 MHz, CDCl₃) δ 8.08 – 8.00 (m, 2H), 7.59 – 7.50 (m, 1H), 7.42 – 7.33 (m, 12H), 6.67 – 6.63 (m, 1H), 5.41 – 4.99 (m, 5H), 4.60 – 4.31 (m, 3H), 2.78 – 2.45 (m, 2H). ¹³C NMR (50 MHz, CDCl₃) δ 165.6, 165.5, 154.9, 154.9, 154.9, 154.6, 135.1, 135.0, 135.0, 134.8, 133.5, 133.4, 129.9, 129.9, 129.9, 129.7, 128.9, 128.8, 128.8, 128.8, 128.7, 128.7, 128.6, 128.5, 128.5, 128.4, 128.4, 98.9, 98.6, 83.4, 82.5, 77.5, 77.2, 70.3, 70.1, 70.1, 69.9, 67.2, 67.1, 38.6, 38.6.

Synthesis of Phosphoramidite **212**



(i)-(ii) to a solution of **211** (3.8 g, 7.5 mmol) in absolute ethanol, were added TEA (1 %, v/v) and Pd/C (10 %). The mixture was stirred at room temperature under H₂ atmosphere until the reaction was judged complete (TLC analysis, 2 hours). The reaction was then filtered over celite and the solvent was removed in *vacuo*. The crude (with intermediate **212**) was co-evaporated three times with pyridine and used for the following steps without further purification. Yield: (1.78 g, quantitative yield).

The crude of step (i) (1.78 g, 7.5 mmol) was dissolved in dry pyridine and DMTrCl (3.42 g, 9 mmol) was added in three different portions each 15 minutes. The reaction was stirred at room temperature overnight. The solvent was removed in *vacuo*, and the crude was purified by flash column chromatography on silica gel, eluting with Cy/EA (8/2 with 1% of TEA) to afford compound **213** as a yellow oil. Yield: 87 % (3.52 g). ¹H NMR (400 MHz, CD₂Cl₂) (β anomer) δ 8.03 (dd, *J* = 8.2, 1.1 Hz, 2H), 7.62 – 7.58 (m, 1H), 7.47 – 7.43 (m, 4H), 7.34 – 7.32 (m, 6H), 7.25 – 7.22 (m, 1H), 6.86 (dd, *J* = 9.0, 0.7 Hz, 4H), 6.65 (d, *J* = 5.0 Hz, 1H), 4.38 – 4.33 (m, 2H), 3.79 (s, 6H), 3.26 – 3.18 (m, 2H), 2.63 – 2.56 (m, 1H), 2.24 (d, *J* = 14.3 Hz, 2H), 1.62 (br s, 1H). ¹³C NMR (101 MHz, CD₂Cl₂) δ 165.9, 159.1, 145.3, 136.3, 136.2, 133.6, 130.6, 130.4, 129.9, 128.9, 128.4, 128.3, 127.2, 113.5, 113.5, 100.2, 88.2, 86.6, 73.1, 64.2, 55.6, 55.6, 41.6.



(iii) To a solution of a of **213** (1.08 g, 2 mmol, β anomer) in dry DCM (20 ml) it was added 3 equivalents of fresh distilled DIEA (1.1 mL, 6 mmol). Then, 2-Cyanoethyl-*N,N* diisopropylchlorophosphoramidite (0.58 mL, 2.4 mmol) was added dropwise and the reaction was stirred until complete conversion of the starting material (1 h). The solvent was removed in *vacuo* and the crude was purified by flash column chromatography eluting with Cy/EA (9/1 with 1% of TEA) to afford the phosphoramidite **214β** as with a solid. Yield: 70 % (1.05 g). ¹H NMR (400 MHz, CD₂Cl₂) δ 8.08 (d, *J* = 7.7 Hz, 2H), 7.59 (t, *J* = 7.4 Hz, 1H), 7.48 – 7.44 (m, 4H), 7.36 – 7.29 (m, 6H), 7.26 – 7.21 (m, 1H), 6.87 – 6.84 (m, 4H), 6.68 – 6.64 (m, 1H), 4.58 – 4.53 (m, 1H), 4.50 – 4.46 (m, 1H), 3.79 (d, *J* = 2.1 Hz, 6H), 3.67 – 3.59 (m, 4H),

3.37 – 3.28 (m, 1H), 3.17 – 3.12 (m, 1H), 2.65 – 2.46 (m, 2H), 2.42 – 2.32 (m, 2H), 1.18 – 1.05 (m, 12H). ^{31}P NMR (162 MHz, CD_2Cl_2) δ 148.85, 148.59.

Annexes

Annex 1

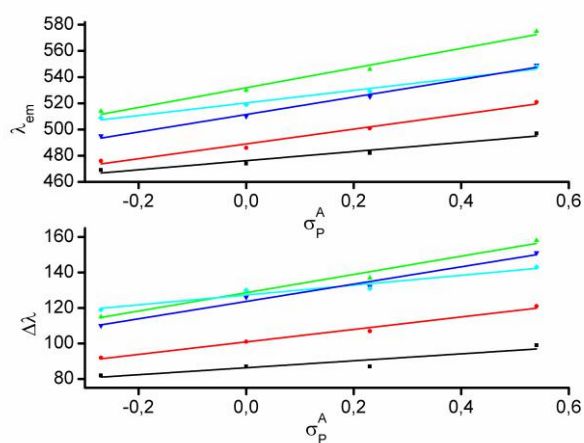
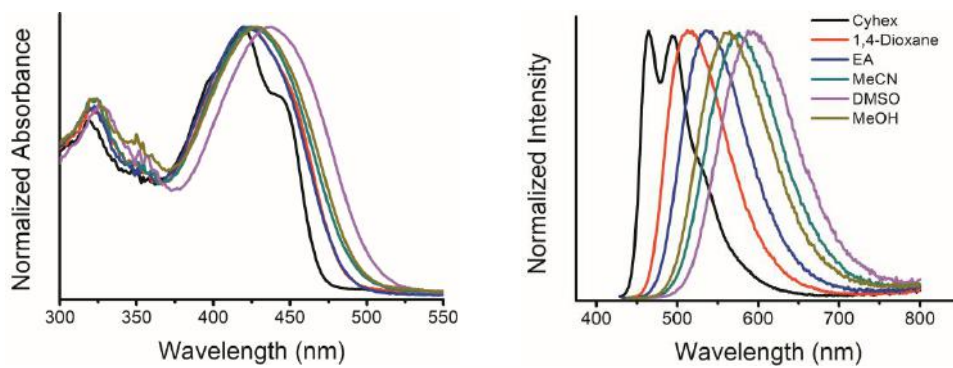
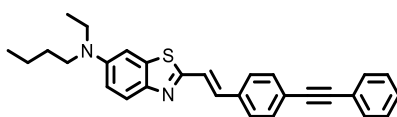


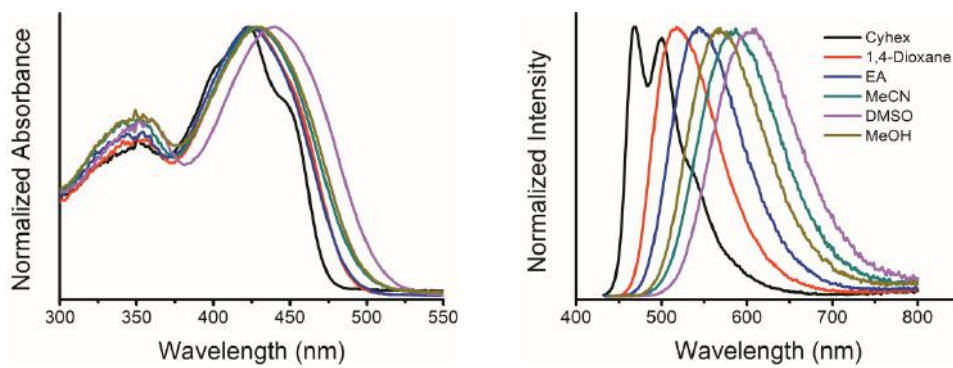
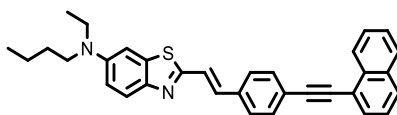
Figure A1. Dependence of the emission maxima (up) and Stokes shift (down) of **85**, **86**, **87** and **88** on the Hammett values (σ_p) in methanol (cyan), acetonitrile (blue), dimethyl sulfoxide (green), ethyl acetate (red) and 1,4-dioxane (black). Correlation factors R^2 are, respectively, 0.94, 0.98, 0.98, 0.99 and 0.98 related to λ_{em} and 0.80, 0.98, 0.97, 0.97 and 0.92 related to $\Delta\lambda_{ST}$.

Annex 2

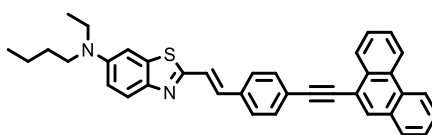
110

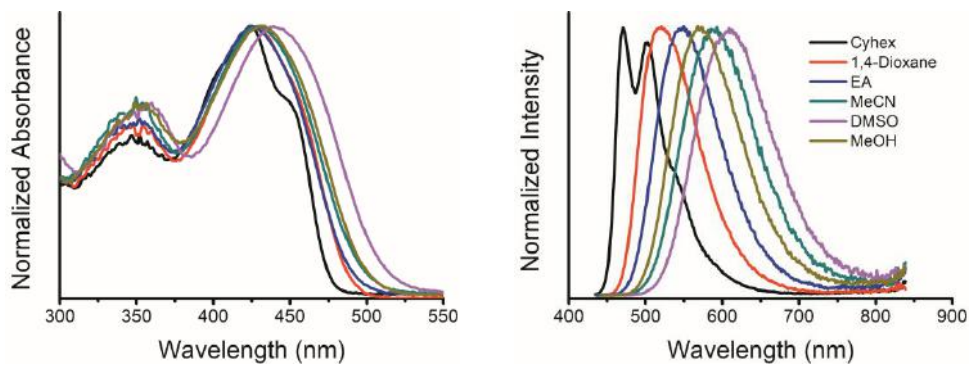


111

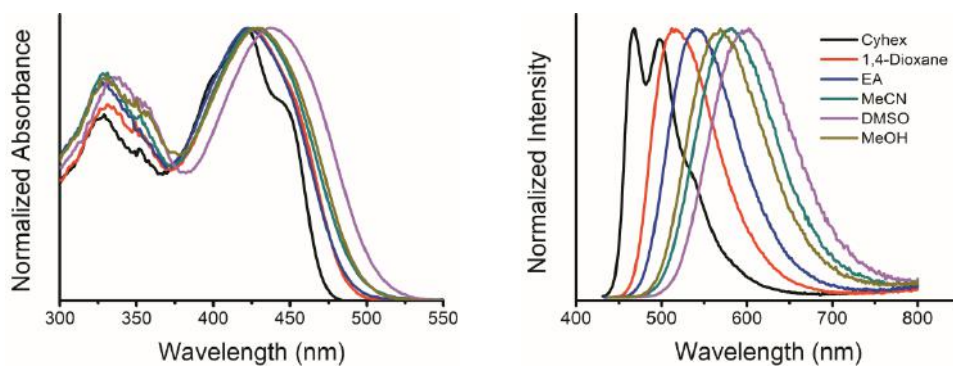
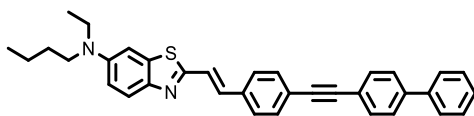


112

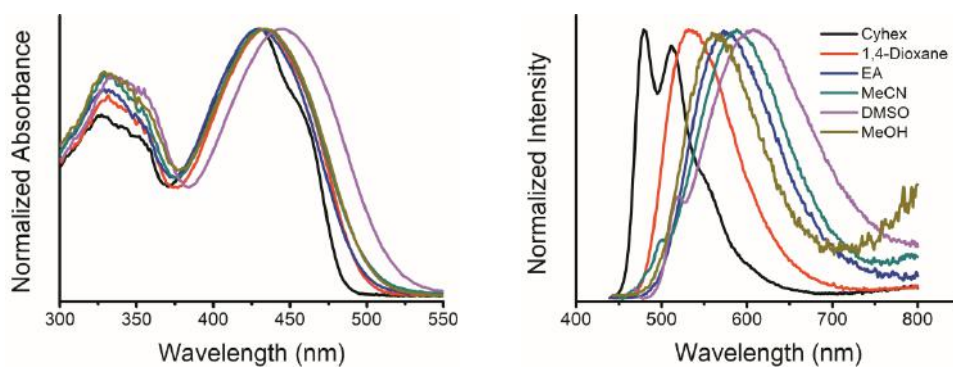
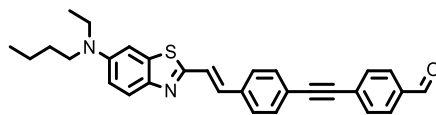




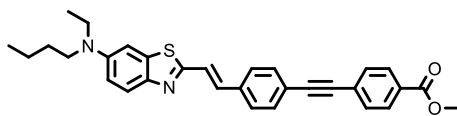
113

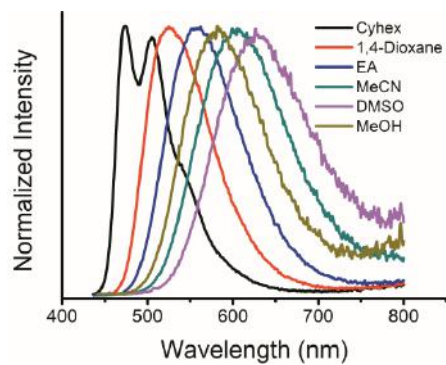
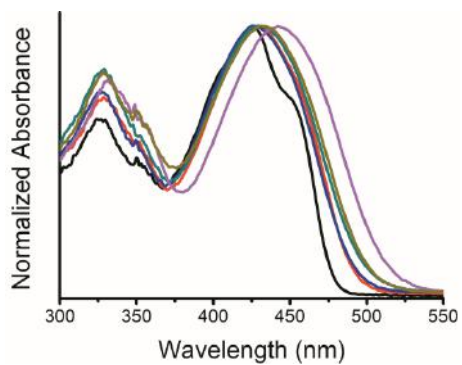


114

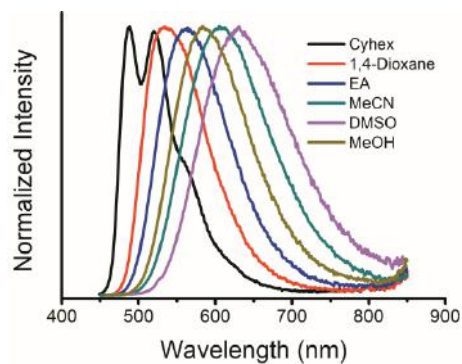
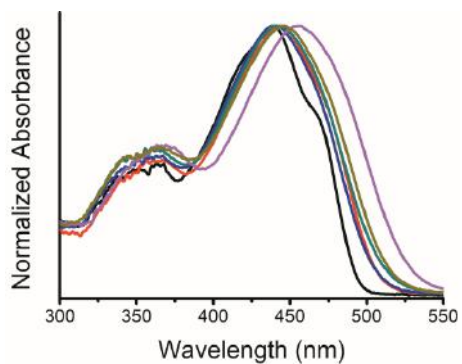
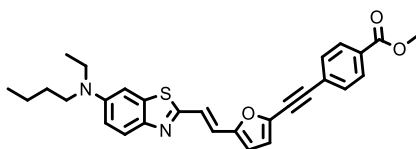


115





116



117

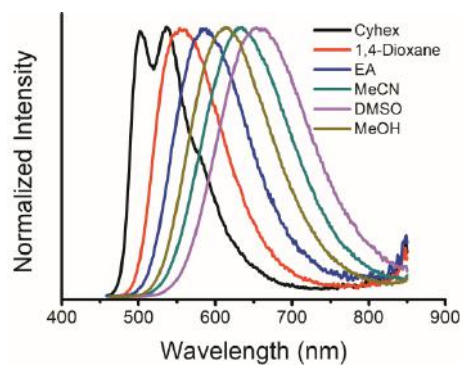
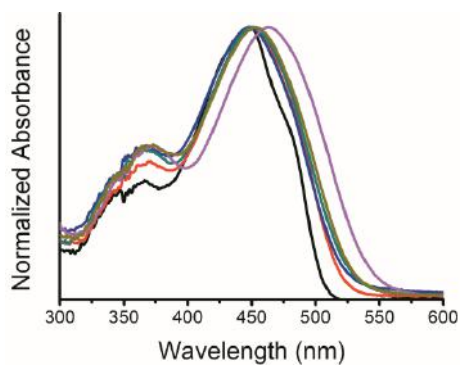
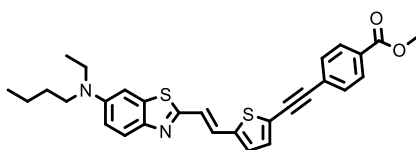
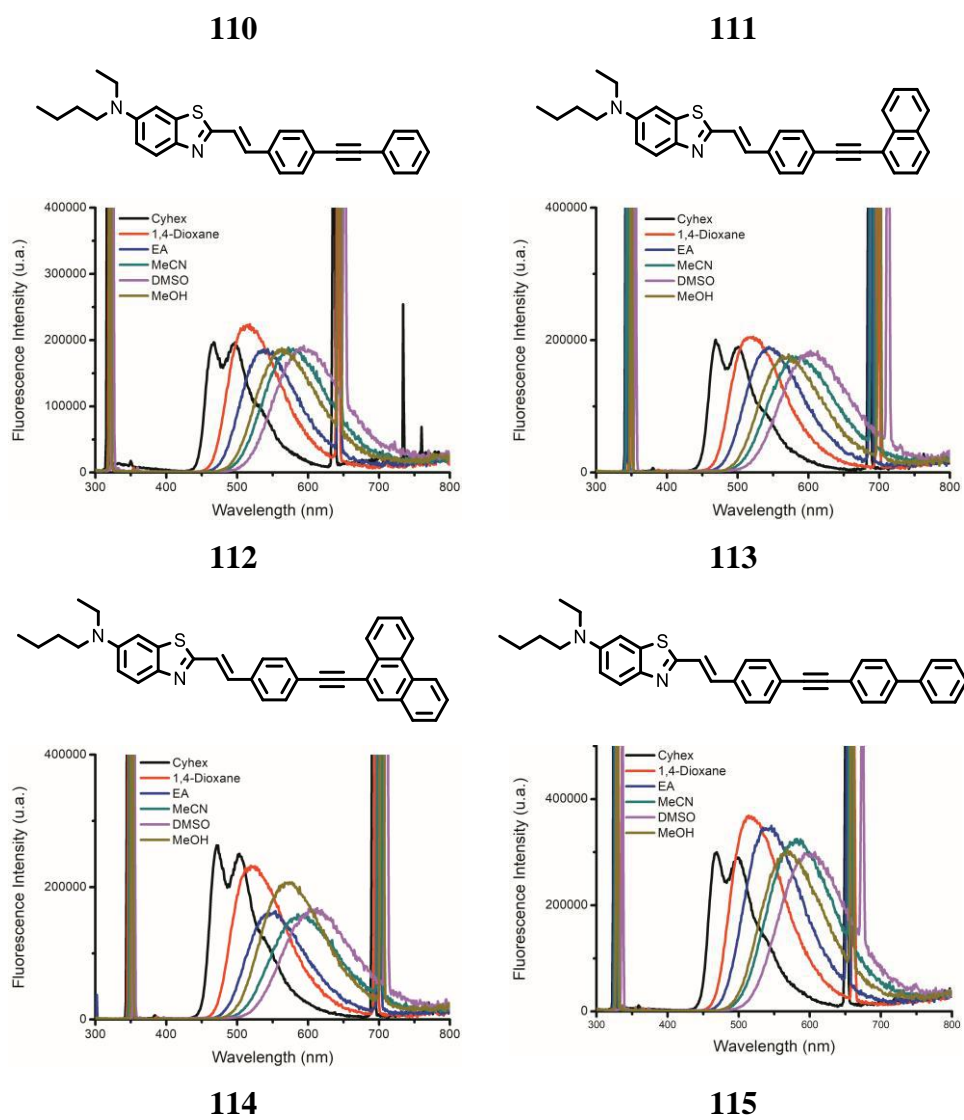
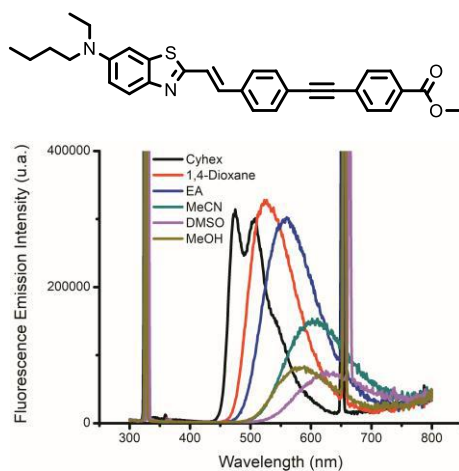
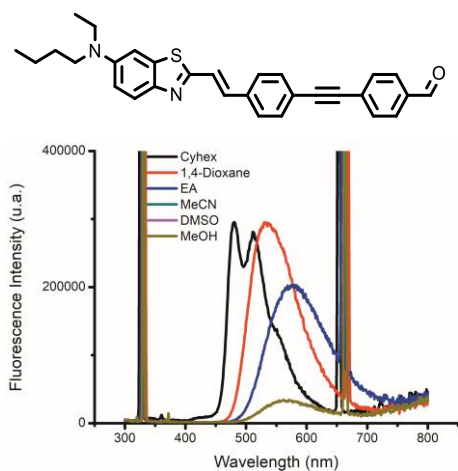


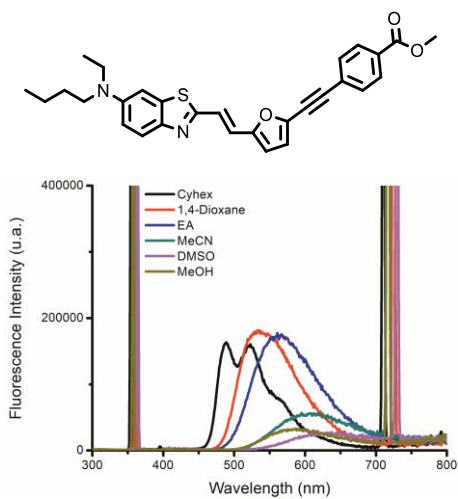
Figure A2. Absorption and fluorescence emission spectra of dyes **110-117** in different solvents. Condition: dye concentration $\approx 1 \mu\text{M}$.

Annex 3





116



117

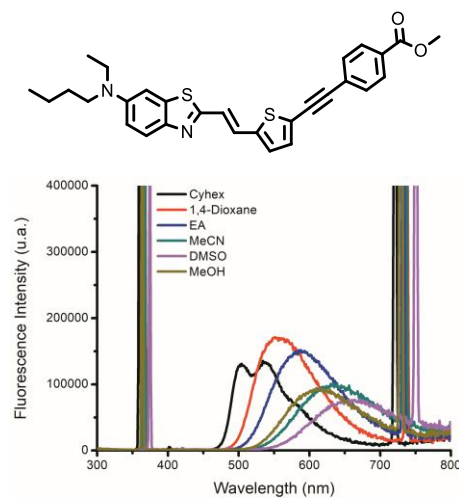
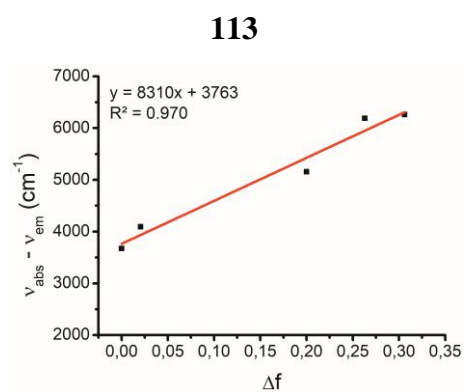
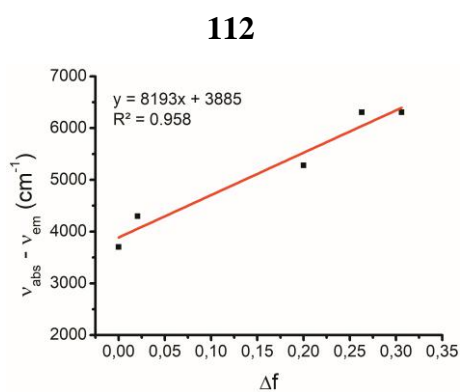
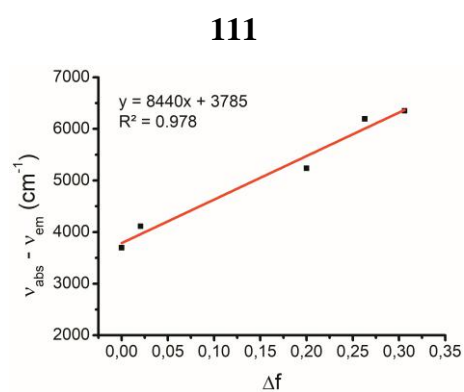
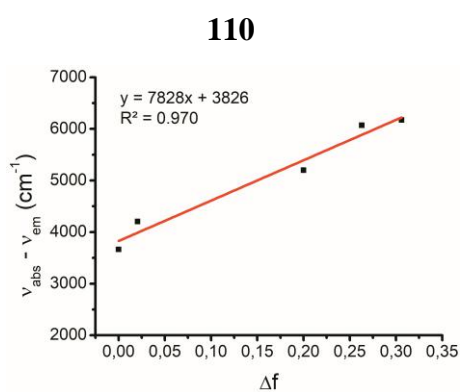


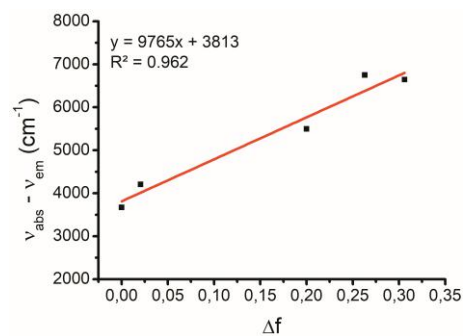
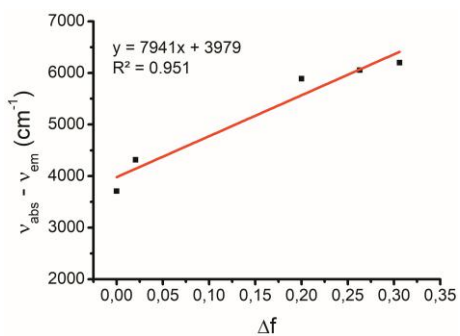
Figure A3. Fluorescence emission spectra of dyes **110-117** in different solvents after excitation in the near-UV band. Condition: dye concentration $\approx 1 \mu\text{M}$.

Annex 4



114

115



116

117

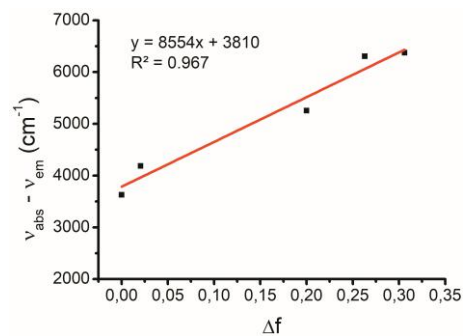
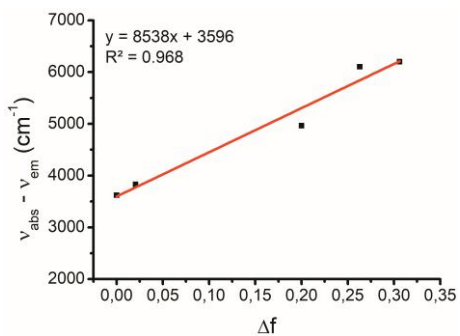


Figure A4. Correlation between Stokes shift ($\nu_{\text{abs}} - \nu_{\text{em}}$, in cm^{-1}) and the solvent polarizability parameter (Δf) according to the Lippert-Mataga model.

Nucleic Acid Ligands and Benzothiazole-Based Fluorophores

Synthesis and Biological Applications

Abstract: Interdisciplinary is an obvious feature of scientific research. Notably, the continuously expanding researches in the medical field fostered this interdisciplinarity. Within this context, the work presented here deals with fluorescence spectroscopy and nucleic acid chemistry and aims at providing new tools suited for applications in life sciences. This manuscript presents two sections. In Part I, we developed novel series of highly bright and non-toxic push-pull fluorophores, compatible with cellular applications. In fact, we describe new benzothiazole-based fluorophores that can be structurally modulated to provide dyes with tunable photophysical properties. Their structure-photophysics relationship is also reported. Next, using these fluorophores, we produced new fluorescent probes to monitor the enzymatic activity of the β -galactosidase. Of note, since β -galactosidase is a marker of cellular senescence, we also used our probes for the detection of early stages of cellular senescence using fluorescence microscopy and flow cytometry analysis. Part II of this manuscript is devoted to the chemistry of nucleic acids. First, we describe the preparation and evaluation of novel RNA ligands targeting the III_d domain of the HCV IRES. These ligands can make a base triplet with the III_d sequence to provide specificity, and harbour amino functions to stabilize the

triplet by electrostatic interaction with the phosphodiester of IIIId. Finally, the last part concerns the development of a new strategy allowing the post-synthetic functionalization of nucleic acids at the anomeric position. To do this, we designed new phosphoramidite platforms that, once incorporated into oligonucleotide strands, can be involved in various glycosidation reactions.

Key words: fluorescence, push-pull, β -galactosidase, RNA ligands, oligonucleotides, post-synthetic functionalization