UNIVERSITY OF NAPLES FEDERICO II

DEPARTMENT OF PHARMACY

PhD programme in Pharmaceutical Sciences

CONFORMATIONAL AND INTERACTION STUDIES THROUGH SOLUTION NMR AND OTHER BIOPHYSICAL TECHNIQUES.

PhD student
FEDERICA SANTORO

Coordinator
Prof.
ROSARIA MELI

Tutor
Prof.
ALFONSO CAROTENUTO

XXXV CYCLE (2020-2022)

CONTENTS

Abstract 4

INTRODUCTION AND GENERAL BACKGROUND 8

1. BIOPHISICAL METHODS: techniques to study structure, properties, dynamics or function of biomolecules. 9

2. OVERVIEW ON THE AREAS OF APPLICATION 11

2.1 Conformational Analysis of Peptides 11

2.2 Ligand-DNA Interaction from NMR Spectroscopy 12

2.3 Study of Reaction Environment 13

2.4 Protein Structural Analysis 14

METHODS 15

1.NUCLEAR MAGNETIC RESONANCE 16

1.1 NMR Techniques for Conformational Analysis of Peptides: 2D Experiments and Structural Determinations 16

1.2 Ligand-Based NMR methods: Saturation Transfer Difference (STD) 22

1.3 Paramagnetism-assisted NMR techniques to investigate reaction environment 25

2. CIRCULAR DICHROISM 27

3. X-RAY CRYSTALLOGRAPHY 29

Chapter1 -CONFORMATIONAL ANALYSIS OF PEPTIDES 31

1.1 A novel β-hairpin peptide derived from the ARC repressor selectively interacts with the major groove of B-DNA 32

1.1.1 Introduction 33

1.1.2 Results and Discussion 34

1.1.3 Conclusion 44

1.1.4 Experimental section 45

1.2 Disulfide Bond Replacement with 1,4- and 1,5-Disubstituted [1,2,3]-Triazole on C-X-C Chemokine Receptor Type 4 (CXCR4) Peptide Ligands: Small Changes that Make Big Differences 49

1.2.1 Introduction 50

1.2.2 Results and discussion 52

1.2.3 Conclusion ... 64

1.2.4 Experimental section ... 65

1.3 Glucopeptides derived from myelin-relevant proteins and hyperglucosylated non-typeable *Haemophilus influenzae* bacterial adhesin cross-react with multiple sclerosis specific antibodies: a step forward in the identification of native autoantigens in multiple sclerosis. ... 73

1.3.1 Introduction ... 74

1.3.2 Results and Discussion ... 77

1.3.3 Conclusion ... 82

1.3.4 Experimental section ... 83

1.4 Synthetic short-chain peptide analogues of H1 relaxin lack affinity for the RXFP1 receptor and relaxin-like bioactivity. Clues to a better understanding of relaxin agonist design ... 85

1.4.1 Introduction ... 86

1.4.2 Results and Discussion ... 87

1.4.3 Conclusion ... 91

1.4.4 Experimental section ... 92

Chapter2 - LIGAND-DNA INTERACTION BY NMR SPECTROSCOPY ... 95

2.1 Synthesis and Characterization of Bis-Triazolyl-Pyridine Derivatives as Noncanonical DNA-Interacting Compounds ... 96

2.1.1 Introduction ... 97

2.1.2 Results and Discussion ... 100

2.1.3 Conclusion ... 110

2.1.4 Experimental section ... 111

Chapter3 - STUDY OF REACTION ENVIRONMENT BY NMR ... 116

3.1 Photo-micellar Catalyzed Synthesis of Amides from Isocyanides: Optimization, Scope, and NMR Studies of Photocatalyst/Surfactant Interactions. ... 117

3.1.1 Introduction ... 118

3.1.2 Results and Discussion ... 120

3.1.3 Conclusion ... 131

3.1.4 Experimental section ... 132

Chapter4 - PROTEIN STRUCTURAL ANALYSIS WITH X-RAY ... 142

4.1 Investigation of antitoxin-mimicking peptides and the bacterial toxin Doc binding mode 143

4.1.1 Introduction 144

4.1.2 Results and Discussion 145

4.1.3 Conclusion 148

4.1.4 Experimental Section 149

SUPPLEMENTARY MATERIAL 150

References 180

Abstract

During my PhD, I applied biophysical techniques, mainly nuclear magnetic resonance (NMR) and circular dichroism (CD), but also X-ray crystallography, aimed at investigating different systems of pharmacological and chemical interest. They can be grouped in: 1) conformational analysis of bioactive peptides; 2) interaction analysis between small-molecules and DNA; 3) molecular interactions in a micellar reaction environment; 4) bacterial protein–protein interactions. These studies are briefly described below.

1) *Conformational analysis of bioactive peptides.*

1a) I have carried on a conformational analysis of novel β-hairpin peptides derived from the transcription factor protein, ARC repressor. Transcription factors (TFs) have a remarkable role in the homeostasis of the organisms and there is a growing interest in how they recognize and interact with specific DNA sequences. In particular, ARC repressors, form dimers that insert antiparallel β-sheets into the major groove of DNA. The 3D conformational arrangement of the investigated peptides was preliminarily verified through circular dichroism (CD). The peptides, found to be compatible with a β-hairpin folding (% of β-sheet structure > 30%), were analyzed using NMR, confirming a clear tendency of all peptides to adopt the β-hairpin structure. Finally, NMR-based structure calculation was also performed on the peptide which showed the most interesting binding and selectivity properties, and the simulated annealing gave a quite well-defined β-hairpin structure with a type I' β-turn on its tip. In conclusion, we designed a novel β-hairpin peptide derived from the ARC repressor that selectively interacts with the major groove of B-DNA.

1b) I explored the solution conformation of novel cyclic peptides bearing the 1,4- and 1-5-disubstituted [1,2,3]-triazolyl moieties which replaced the disulfide bond of a known potent peptide antagonist of CXCR4. CXCR4 is a G-protein coupled receptor (GPCR), which modulates many physiological functions by interacting with its own ligand C-X-C motif chemokine 12 (CXCL12). The CXCR4/CXCL12 pair

has emerged as key players in promoting the tumor growth, invasion, angiogenesis and metastasis in more than 30 different human cancers. Several strategies to hamper the interaction between CXCR4 and CXCL12 have been investigated, leading to the development of potent CXCR4 antagonists including cyclopeptides. The investigation of these novel cyclic peptides led to the identification of an agonist and an antagonist showing high affinity and selectivity against CXCR4. Their different efficacy was rationalized by NMR and computational studies that showed slight differences in the binding mode of the two peptides, which might account for the distinct ability in modulating CXCR4.

1c) I explored the secondary structure of three glucosylated peptides derived from proteins that have been selected by a bioinformatic approach for their conformational (OMGp and RTN4R) or sequence (FAN) homology with CSF114(Glc), a synthetic antigen which is specifically recognized by antibodies (Abs) in Multiple Sclerosis (MS) patients' sera. CD measurements indicated that the glucopeptides derived from OMGp and RTN4R featured a high percentage of β-strand content, as observed in the lead compound CSF114(Glc), while the peptide derived from FAN is mainly in a random coil conformation. Notably, β-strand content paralleled glucopeptide ability in Abs recognition. Notably, the progression of MS is also linked to exogenous infectious agents expressing antigenic molecules. In this context, MS antibodies recognize a cell-surface adhesin protein of non-typeable *Haemophilus influenzae* (NTHi) termed HMW1. Interestingly, peptides derived from OMGp and RTN4R, as well as CSF114(Glc), were able to cross-react with antibodies recognizing HMW1, thus confirming the presence of a shared epitope. The structural correspondence between an exogenous protein and a physiological one, is the basis of the hypothesis that molecular mimicry triggers the breakdown of self-tolerance in MS. Moreover, my results suggest that the resemblance among the MS-specific epitopes has a significant conformational component.

1d) Alongside, I performed conformational analysis of synthetic short-chain peptide analogues of Relaxin. The peptide hormone Relaxin (RLX) holds great promise as a

cardiovascular and anti-fibrotic agent but is limited by the pharmacokinetic issues. Six low molecular weight peptides were designed with the objective to obtain RLX analogues with improved pharmacokinetic features. Their secondary structure propensity was explored by CD spectroscopy. CD spectra of the peptides showed a tendency to assume α-helical secondary structure, typical of the hormone bioactive conformation. Despite the favourable premises, none of the tested peptides revealed a substantial affinity to RLX-receptor, RXFP1, nor displayed any RLX-like biological effects. Albeit negative, these results offer additional information about the structural requirements that new peptide therapeutics shall possess to effectively behave as RXFP1 agonists and RLX analogues.

2) *Interaction analysis between small-molecules and DNA.*

I have investigated the ability of novel designed compounds to interact with noncanonical DNAs, such as G-quadruplexes (G4s) and i-motifs (iMs), whose occurrence in gene promoters, replication origins, and telomeres highlights the breadth of biological processes that they might regulate. In particular, STD NMR experiments were performed. STD NMR spectra demonstrated a selective interaction between some of the developed compounds and noncanonical DNAs. Specifically, the presence of both aromatic and aliphatic signals in the STD spectra suggests that the aromatic moieties of some compounds could interact with the G4 and iM through π-π stacking with the guanines of G-tetrads and the adenines of loops, respectively, while the side chains could interact with the backbone of the DNA molecules through electrostatic interactions. Our investigation led to the identification of novel compounds that might be used as tools to shed light on the mechanisms underlying the controversial biological roles of G4 and iM structures as well as on their intricated relationships.

3) *Molecular interactions in a micellar reaction environment*.

NMR can provide information about the structure, dynamics, and interactions of molecules in a reaction environment. Particularly, I have performed NMR experiments to study the reaction environment at the atomic level of a photo-micellar

catalyzed synthesis of amides from isocyanides. I studied the localization of the photocatalyst $[Ir(ppy)_2bpy]PF_6$ relative to the surface and the interior of two micellar systems SDS or CTAC which turned out to be the most and the least efficient reaction systems, respectively. The study was carried out determining the NOE contacts between photocatalyst and the micelles and using specific paramagnetic probes, 16-doxylstearic acid (16-doxyl) and Mn^{2+}, for the position determination. NMR analysis demonstrated that the catalyst is on average positioned on the micelle surface and can flip from the outer to inner part of it in the case of SDS while it is deeply inserted in CTAC micelles. The obtained experimental data allow a rational approach for selecting the best reaction conditions suggesting that, for an optimal catalytic efficiency, the photocatalyst must be positioned on the micelles' surface and almost free to move inside and outside the micelles.

4) *Bacterial protein–protein interactions.*

Finally, X-ray crystallography can be used to study the structure of proteins, as well as their interactions with other molecules. I applied X-ray crystallography to investigate the binding mode of antitoxin-mimicking peptides derived from the protein Phd and the bacterial toxin Doc. One of the major challenges of studying toxin-antitoxin modules is to obtain good amounts of protein. In fact, in bacterial expression systems, toxin expression severely inhibits normal growth, resulting in production of only trace amounts of the toxin of interest. To overcome this issue, I applied a protocol of co-expression of Doc with Phd peptides that allowed us to obtain sufficient amounts of complex to be used for X-ray crystallography. The structure of the Doc-Phd complex helped us to confirm the molecular details of their interaction. The next step is to obtain the structure of Doc with a higher affinity peptide, in order to proceed with the comparison of the two binding modes aimed at designing novel antitoxin compounds useful as antibacterial weapons.

INTRODUCTION AND GENERAL BACKGROUND

1. BIOPHISICAL METHODS: techniques to study structure, properties, dynamics or function of biomolecules.

Biophysics is an interdisciplinary field that combines the principles of physics and biology to study the physical and chemical properties of biological systems. There are a wide variety of biophysical techniques that are used to study the structure, properties, dynamics, and function of biomolecules, including peptides, proteins, and nucleic acids.

During my PhD, I applied biophysical techniques, mainly nuclear magnetic resonance (NMR) and circular dichroism (CD), but also cryo-electron microscopy (Cryo-EM) and X-ray crystallography, on different targets to obtain detailed information about the three-dimensional structure and conformation of biomolecules at atomic resolution.

NMR is a powerful technique that utilizes the magnetic properties of certain atomic nuclei to study the structure and dynamics of molecules. It is a widely used technique for analyzing the conformations of peptides and proteins, as well as their interactions with other molecules, such as DNA.

CD spectroscopy measures the absorption of left- and right-handed circularly polarized light by chiral molecules and is often used to study the secondary structure and conformational changes of proteins. It can be used to analyze the conformational changes of biomolecules, such as those that occur upon binding to a ligand or upon changes in the reaction environment.

Cryo-EM is a powerful imaging technique that uses electrons to produce high-resolution images of biological specimens, and is particularly useful for studying large, flexible, or transient protein complexes. Cryo-EM can be used to study the conformations of proteins in their native, functional states, as well as their interactions with other molecules.

X-ray crystallography is a technique that uses the diffraction of X-rays by a crystal to determine the three-dimensional structure of a molecule. This

technique involves growing large, ordered crystals of the protein of interest. Then X-ray diffraction patterns are used to determine the positions of the atoms within the crystal. It can be used to study the structure of proteins in various conformational states, as well as their interactions with other molecules.

Overall, these biophysical techniques are invaluable tools and are widely used in a variety of research applications.

2. OVERVIEW ON THE AREAS OF APPLICATION

2.1 Conformational Analysis of Peptides

The conformational analysis of peptides can be used to study their secondary structure. Secondary structure refers to regular, recurring arrangements in space of adjacent amino acid residues in a polypeptide chain. It is often maintained by hydrogen bonds between amide hydrogens and carbonyl oxygens of the peptide backbone.

Understanding the secondary structure of peptides can provide insight into their function and help to design and develop new drugs and therapies. One of the aims of conformational studies of peptides is the possible determination of the bioactive conformation. Bioactive conformation refers to the specific three-dimensional structure of a molecule that is necessary for it to bind to a specific target and perform its biological function. Many small molecules and biomolecules, including peptides, have a specific bioactive conformation that is required for their activity. For example, many drugs are designed to bind to specific proteins or receptors in the body and alter their activity. The ability of the drug to bind to the target and perform its function is dependent on the conformation of the molecule.

The study of the structure of proteins and peptides is generally carried out in the solid state, by diffractometric methods (X-rays) or in solution, using circular dichroism (CD) and nuclear magnetic resonance (NMR) techniques.

2.2 Ligand-DNA Interaction from NMR Spectroscopy

Ligand-DNA interactions can cause conformational changes in the DNA molecule, which can be important for regulating gene expression and other biological processes. By studying these interactions, it is possible to gain insight into the conformational changes that occur in the DNA upon ligand binding and how these changes affect the function of the DNA. Furthermore, many drugs work by binding to specific DNA sequences and modulating gene expression. Studying the interaction of these molecules with DNA, it is possible to identify new drug candidates and understand how these drugs bind to and affect the function of the DNA. It is also possible to identify new targets for gene regulation and develop strategies for modulating gene expression in a controlled manner.

NMR spectroscopy is a powerful tool that can be used to study these interactions and understand their effects on the structure and function of DNA.

2.3 Study of Reaction Environment

NMR can provide information about the structure, dynamics, and interactions of molecules in a reaction environment.

The past several decades have witnessed a tremendous growth in the application of catalytic routes to the synthesis of complex organic molecules, driven by academic and industrial discoveries of efficient, selective catalysts for a wide variety of liquid- and multiphase organic transformations. These developments coincide with major efforts in the pharmaceutical and chemical industries toward streamlining the waste disposal in an ever more ecologically aware market. In particular, the area of organocatalysis has received much attention due to its advantages in the field of green chemistry. It can be very useful allow a rational approach to the process of fine-tuning the conditions of synthetic organic chemistry procedures and NMR spectroscopy can easily fit in this context. Indeed, by analyzing the NMR spectrum of the reactants and products, it is possible to obtain detailed information into the factors that influence the reaction, such as the electronic and steric properties of the reactants, the solvent, and any catalysts that are used. By using this information, it may be possible to design more effective reaction conditions.

2.4 Protein Structural Analysis

Understanding the structure of a protein can provide important insights into its function. Studying the structure of a protein can therefore help scientists understand how it performs its specific functions within the body, such as catalyzing chemical reactions, transporting molecules, or providing structural support.

Additionally, understanding the structure of a protein can also be useful for developing new drugs. Many drugs work by targeting specific proteins in the body and modulating their function. By understanding the structure of these proteins, scientists can design drugs that specifically target and modulate their function. This can be especially useful for developing drugs to treat diseases caused by abnormal protein function, such as cancer or neurological disorders.

X-ray crystallography is the technique of election for the study of protein structure. X-ray crystallography involves growing crystals of the protein of interest and then using X-rays to determine the arrangement of atoms within the protein. This technique can provide detailed information about the three-dimensional structure of proteins, which is important for understanding their function.

METHODS

1.NUCLEAR MAGNETIC RESONANCE

1.1 NMR Techniques for Conformational Analysis of Peptides: 2D Experiments and Structural Determinations

High-resolution two-dimensional NMR spectroscopy is one of the most important tools in the elucidation of three-dimensional structure of fair-sized biomolecule. Secondary structure in peptides refers to the local, regular, and recurring patterns in the conformation of the polypeptide chain. These patterns are formed by the interactions between the peptide bonds and the side chains of the amino acids that make up the peptide. The most common secondary structures found in peptides are the α-helix and β-sheet. The α-helix is a coiled structure formed by hydrogen bonds between the peptide backbone atoms, while the β-sheet is a structure formed by hydrogen bonds between the peptide backbone atoms in adjacent strands. Other secondary structures such as β-turns and loops also exist, but they are less common.

The first step in the structural study of any molecule by NMR is to assign all the resonances related to protons and, possibly, other nuclei with magnetic moment. The main method of assigning proton resonances of peptides and proteins is called "sequential". This method was first introduced by Wüthrich *et al.*[1] and is based on the use of two types of experiments:

1. experiments that give information about relationships along bonds.
2. experiments that give information about relationships through space.

Correlation experiments, such as COSY, DQF-COSY[2], TOCSY[3] etc, are mainly used for the assignment of the residue type. These experiments allow to group the resonances that belong to the individual amino acids, since only the intraresidual relationships are shown, as well as to identify the spin classes of the amino acids to which they belong. The simplest correlation experiment is the COSY experiment (two-dimensional homonuclear

correlation spectroscopy), first described in 1976, which shows only direct correlations through bonds. This basic experiment has now been surpassed by much more sophisticated experiments, such as DQF-COSY (two-dimensional homonuclear correlation spectroscopy with double quantum filter), that have enhanced resolution. Individually, these experiments, which show only direct correlations through bonds, are of limited utility and for this reason they are performed together with experiments that also show indirect or relayed correlations, such as the COSY-RELAYED or TOCSY experiments. Once the spin systems have been identified, the search for sequential contacts between nearby spin systems can be carried out systematically. This information is obtained from the analysis of the regions of the NOESY[4] spectra that contain the cross-peaks NH-NH, NH-Hα, NH-Hβ. Overhauser effects (NOE) due to Hα-NH distances are especially useful for sequential assignment of β-sheets sections; while in the case of helical strokes, effects from short NH-NH distances are more intense and more diagnostic.

Once the sequential assignment has been carried out, the following parameters are collected: coupling constants, temperature coefficients and NOE effects. The $^{3}J_{H\alpha\text{-}NH}$ vicinal coupling constant is the easiest to determine experimentally, by measuring the anti-phase separation of cross-peak NH_i-$H_{\alpha i}$ components in COSY experiments. Generally, values of $^{3}J < 5$ Hz are indicative of helix, values of $^{3}J > 9$ are indicative of extended structures.

Furthermore, the main types of secondary structure (α-helices, β-sheets and turn) are generally associated with characteristic hydrogen bridge patterns. Indeed, the values of the temperature coefficients of the amide protons of the backbone are of particular importance for determining the presence of hydrogen bonds. Values of coefficients below 3.0 ppb/K (absolute value)

indicate protons not exposed to the solvent and therefore likely to participate in the hydrogen bond.

Finally, intramolecular NOE effects give information about interprotonic distances of neighbouring atoms in space. Each type of secondary structure is characterized by a particular scheme of NOE (short-range: sequential; medium range: maximum 4 residues; long-range: greater than 4 residues). The α-helices are characterised by a series of strong or medium intensity NOE effects such as $d_{NH\text{-}NH}(i, i+1)$, $d_{NH\text{-}NH}(i, i+2)$, $d_{H\alpha\text{-}NH}(i, i+3)$, $d_{H\alpha\text{-}NH}(i, i+4)$, $d_{H\alpha\text{-}H\beta}(i, i+3)$, and weak effects such as $d_{H\alpha\text{-}NH}(i, i+1)$ (**Figure 1**).

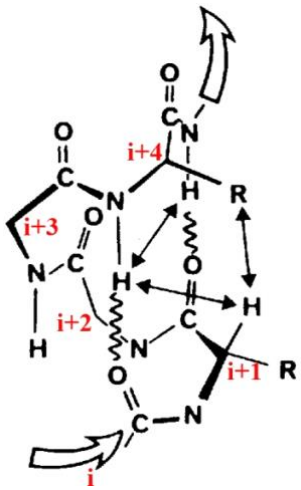


Figure 1*. Representation of a α-helix structure. The hydrogen bonds are represented by wavy lines, the observed NOE are indicated by two-pointed arrows.*

The β-sheet structures are characterized by strong $H_{\alpha i}$-NH_{i+1} and the absence of other short-range effects. Such β-sheet structures can be identified by long-range interlocking NOE effects, such as $H_{\alpha i}$-$H_{\alpha j}$ effects (**Figure 2**).

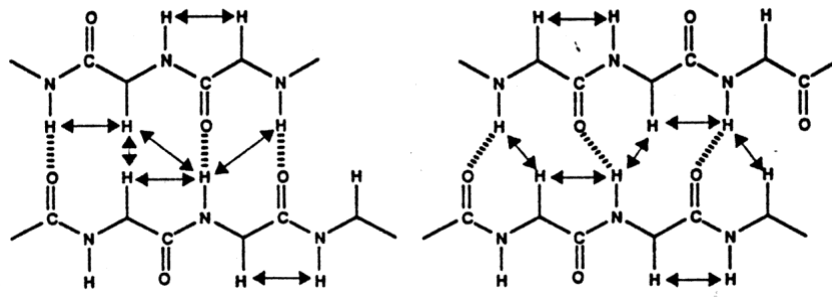


Figure 2. *Representation of a β-sheet structure. On the left is shown an antiparallel β-sheet structure, on the right a parallel β-sheet structure. The hydrogen bonds are represented with dashed lines, the observed NOE are shown with two-pointed arrows.*

For small peptides, the most common secondary structure elements are the turn of type β and type γ. A β-turn consists of 4 residues (called i, i+1, i+2, i+3) and the type of β-turn (I, II, I', II', etc) is defined by the dihedral angles and central residues (i+1, i+2) (**Figure 3**, **Table 1**).

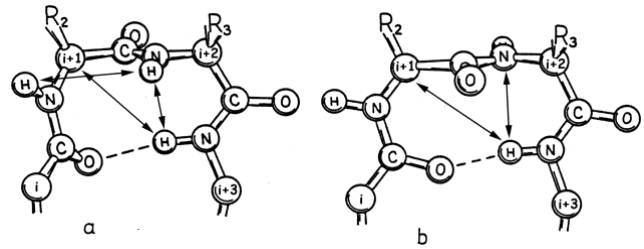


Figure 3. *Representation of a β-turn structure. On the left is a type I β-turn structure, on the right a type II β-turn structure.*

Table 1. Values of the dihedral angles Φ and ψ in the various types of β-turn.

B-Turn	***Φ****1*	***Ψ*** 1	***Φ***2	***Ψ*** 2
I	-60	-30	-90	0
I'	60	30	90	0
II	-60	120	80	0
II'	60	-120	-80	0

A turn of type β is defined by a distance $d_{H\alpha\text{-}H\alpha}(i, i+3) < 7Å$. Generally, but not always, there is a hydrogen bond between CO (i) and NH (i+3). These structures are characterized by coupling constants $^{3}J_{H\alpha\text{-}NH}$ that allow to go

back to the angle Φ relative to central residues of the turn. In addition, for β-turns the diagnostic signal is $H_{\alpha i}$-HN_{i+2} between the i+1 and i+3 residues of the turn (**Figure 4**).

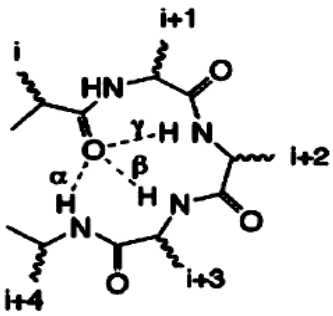


Figure 4. *Distinction between α-, β-, γ-turn*

Once the NMR data described above have been obtained, NOE effects, coupling constants and exchange rates of amide protons can be used to determine the three-dimensional structure of the molecule. The most widely used methods for the calculation of structures are based on the exploration of limited regions around an initial point or on statistical sampling of larger areas of conformational space. The first class of methods, to which belong energy minimizations and molecular dynamics simulations at room temperature, presuppose sufficiently realistic and representative starting structures of the main region of conformational space accessible to the system. For example, these conditions are met when, knowing the crystallographic structure of a molecule, we want to determine its structure in solution, and there is experimental evidence to suggest a limited distance from the solid state. Where, on the other hand, there is no *a priori* information on a given structure, it is necessary to use statistical sampling methods, such systematic mapping of the conformational space, Monte Carlo methods, or high temperature molecular dynamics simulations (the so-called simulated annealing). The procedure termed simulated annealing is a high-temperature

molecular dynamics in order to allow the peptide to escape from any relative minimum of potential energy. The system is eventually cooled slowly to a temperature of 0 K.

1.2 Ligand-Based NMR methods: Saturation Transfer Difference (STD)

Several NMR spectroscopic techniques are used to understand the binding process at a molecular level and to identify new bioactive substances. In general, two main experimental approaches exist. The first focuses on the NMR signals of the ligand, in which the different motion and NMR properties of a small molecule are exploited in its free state and in its bound state to a macromolecule; the second approach focuses on chemical-shift changes of the target protein upon binding of the ligand.

Saturation-transfer NMR spectroscopy[5] is part of the ligand-based NMR methods. The method is based on the transfer of saturation from the macromolecule to the bound ligands which in turn, by exchange, is moved into solution where it is detected.

During the saturation time (τ_{sat}) the binding site of the macromolecules is consecutively occupied by *n* ligand molecules with $n=f_{PB}*\tau_{sat}/\tau_{res}$, where f_{PB} is the fraction of occupied binding sites[6] and τ_{res} is the residence time of the ligand (**Figure 5**). This turnover is responsible for the amplification of the information resulting from the saturated macromolecules. A large excess of ligands allows for the maximum effect to be observed.

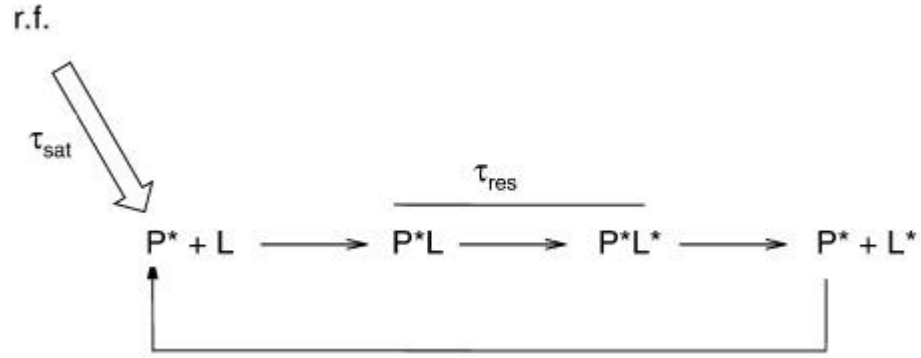


Figure 5. *The Scheme shows the situation of a solution containing ligand and macromolecule after the initial saturation time. Assuming a large excess of ligand over the target macromolecule, the rebinding of already saturated ligands can be neglected. (r.f.: radio frequency for saturation; τ_{sat}: saturation time; τ_{res}: residence time of ligand in the binding site; P*: saturated protein; L: unsaturated ligand; L*: saturated ligand.)* [6]

If a ligand is in fast exchange between the bound state and the free state on the NMR time scale a transfer of saturation between the free and the bound state. By utilizing the power of difference spectra, the method can easily be used for homonuclear spectroscopy, especially proton NMR experiments, to obtain well-resolved spectra of the ligand alone. Subtracting a spectrum in which the macromolecule is saturated from one without protein saturation produces a difference spectrum in which only the signals of the ligand(s) remain.[7](**Figure 6**)

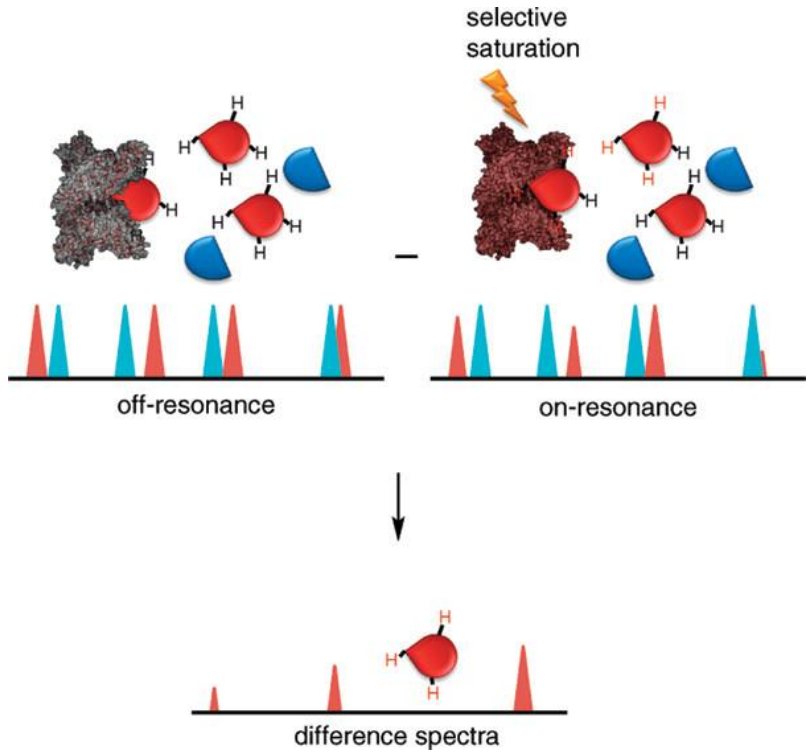


Figure 6. *Scheme of the STD-NMR experiment. The exchange between free and bound ligand allows intermolecular transfer of magnetization from the target to the bound small molecule.*[7]

The irradiation frequency is set to a value where only resonances from the protein nuclei and no resonances from ligand nuclei are located. Therefore, in the on-resonance experiment selective saturation of the signals of the protein nuclei is achieved. For the on-resonance irradiation frequency values around −1 ppm is practical because no ligand nuclei resonances are found in

this spectral region whereas the significant line width of protein signals still allows selective saturation. If the ligands show no resonance signals in the aromatic proton spectral region the saturation frequency may also be placed here or even further downfield (δ=11–12 ppm).

Finally, the intensity of the signals obtained depends mainly on two factors: the excess of ligand molecules used and the extent of ligand saturation. Working with an excess of ligands, it is possible to saturate a greater number of ligands in a few seconds, and consequently amplify the signal strength. In addition, the extent of ligand saturation depends on the residence time of the ligand in the protein binding pocket. In fact, if the binding is very strong, for example when the K_d values are below 1 nM, the saturation transfer to the ligand is not very efficient. If, on the other hand, the K_d values are 100 nM or higher, the fast exchange of the ligand between the free and bound form will lead to more ligands in solution to which saturation will be transferred.[5]

1.3 Paramagnetism-assisted NMR techniques to investigate reaction environment

NMR spectroscopy is a potentially powerful technique for the atomic visualization of dynamic conformations and of interactions between biomolecules in solution. In this context, paramagnetism-assisted NMR approaches could be applied to the characterization of the reaction environment at the atomic level, for instance in photomicellar catalyzed synthesis.

The presence of unpaired electrons in molecules makes them paramagnetic and largely affects their NMR spectra. The main effects related to the presence of a paramagnetic center, i.e., of atoms or ions with unpaired electrons, are varied (**Figure 7**). First of all, the NMR shifts are perturbed, so that the shifts measured for the paramagnetic molecule (δ) and for a diamagnetic analogue (δ^{dia}) (i.e., for the same molecule without the paramagnetic center or with the paramagnetic metal ion substituted by a diamagnetic one) differ. These differences are called hyperfine shifts (δ^{hs}). The nuclear relaxation rates of the paramagnetic molecule (R_1, longitudinal; R_2, transverse) are increased with respect to those of a diamagnetic analogue (R_1^{dia} and R_2^{dia}). This difference is called paramagnetic relaxation enhancement. Lastly, the probabilities for the different orientations of the paramagnetic molecule in a magnetic field are not the same, so a partial self-orientation occurs. This partial self-orientation is responsible for the occurrence of residual dipolar couplings ($\Delta\nu^{rdc}$), analogously to what happens when partial molecular orientation is driven by external devices. The presence of paramagnetic residual dipolar couplings affects the *J*-coupling between nuclei.[8]

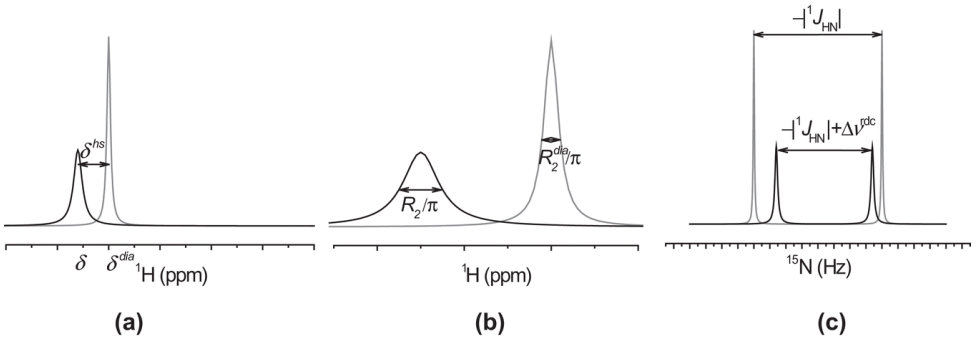


Figure 7. *The presence of a paramagnetic center affects NMR shifts (a), relaxation rates (b) and ^1J splitting (c).*[8]

The merging of micellar and photoredox catalysis represents a key issue to promote "in water" photochemical transformations. 1D and 2D NMR experiments, performed in the presence of paramagnetic probes, enable the study of the reaction environment along with the localization of the photocatalyst with respect to the micelles, thus providing experimental data that allow a rational approach for selecting the best performing photocatalyst/surfactant pairs.

To determine the position of photocatalyst in the micelle, it is possible to use spin-labels, such as 12-doxyl-stearic acid and 5-doxylstearic[9] acid as well as Mn^{2+} ions ($MnCl_2$)[10], to induce selective broadening of resonances from protons close to the paramagnetic probes. The doxyl group containing the spin-label free radicals is bound to carbon 12 or 5 of the stearic acid. 12-doxyl-stearic acid particularly broadens resonances of protons close to the center of the micelle, whereas 5-doxyl-stearic acid broadens the resonances of protons close to the micelle surface. At low concentrations, Mn^{2+} particularly affects resonances of water and the surface of micelles. The paramagnetic broadening effects of the agents on the photocatalyst resonances can be studied by comparing one-dimensional ^{1}H as well as two-dimensional TOCSY spectra in the presence and absence of the paramagnetic agents.

2. CIRCULAR DICHROISM

As full NMR or X-ray crystallographic structure determinations of peptides are demanding and time-consuming and because a large number of peptides must be examined, circular dichroism spectroscopy is frequently used as the initial screening method. It is indeed possible to obtain substantial information on the structure of peptides in solution by measuring their optical activity.

The optical activity characteristic of organic molecules is a result of the absorption of light as electrons are promoted to higher molecular orbitals. Plane-polarized light can be resolved into its two circularly polarized components: left circularly polarized light, whose electric vector rotates counterclockwise about the axis perpendicular to the direction of travel of the light beam, and right circularly polarized light, whose rotation is clockwise. A chiral compound exhibits optical activity because its absorption of left circularly polarized light is not equal to its absorption of right circularly polarized light. After passing through a chiral medium, the electric vectors describe an ellipse whose major axis lies along a new angle of rotation. The measured eccentricity of the ellipse represents the unequal absorptions of left and right circularly polarized light referred to as circular dichroism.[11]A CD spectrum plots the differential absorbance of left and right circularly polarized light vs wavelength.

Circular dichroism is useful in determining the secondary structure of peptides because we can reasonably assume that the peptide backbone contributes significantly in the far UV region and because the spectrum reflects the spatial arrangement of chiral units in the peptide chain. There are three main classes of secondary structure: the α-helix, the β-sheet, and the random coil. The α-helix produces the most distinctive CD spectrum: a very strong positive band near 192 nm, and two negative maxima of approximately equal intensity near 222 nm and 208 nm. The β-sheet exhibits a single negative band near 217 nm. The random

coil exhibits a strong negative band at 197 nm and a small positive band at 217 nm[12](**Figure. 8**).

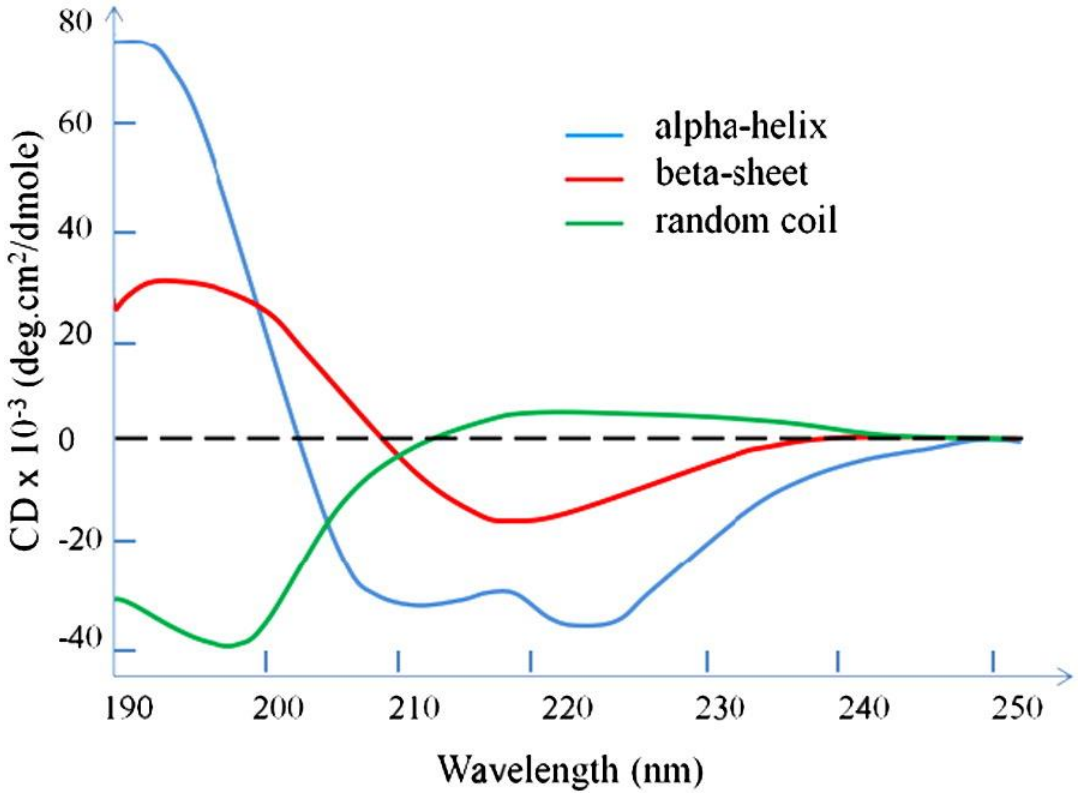


Figure 8. *Each of the three basic secondary structures of a polypeptide chain (α-helix, β-sheet and random coil) show a distinctly different characteristic CD spectrum.*[12]

3. X-RAY CRYSTALLOGRAPHY

X-ray crystallography is a technique that uses X-ray diffraction to determine the three-dimensional structure of a molecule. The basic principle of X-ray crystallography is that X-rays are scattered by the atoms in a crystal, and the pattern of the scattered X-rays can be used to determine the positions of the atoms. Its foundation principle lies in Bragg's law of X-ray diffraction by crystals, i.e. by well-ordered packing of homogenous molecules in three-dimension. Illuminated by a beam of X-ray light, the crystal can diffract the light at various angles, some of which have stronger intensity than others (**Figure 9**).[13] This kind of intensity variation at different angles can be recorded on media as a "diffraction pattern". The diffraction pattern, normally appearing as a series of sharp spots, reflects the structural arrangement of atoms within the crystal and therefore can be used to deduce the original structure of the crystal through a process known as phasing. Once the phases have been determined, the electron density map is calculated, and the atomic structure of the macromolecule can be determined by fitting the atomic coordinates into the electron density map.

To get this three-dimensional structure, several steps that falls within multiple disciplines are required. The first step is the production of highly pure macromolecule in large quantity. Then, the target molecule must pass from soluble state to a solid crystalline ordered state. Obtaining a single homogeneous crystal, that result to high quality diffraction data, represents a crucial step in the process of determining a macromolecular structure. The crystals obtained are cryo-cooled to protect them from radiation damage, and then placed into a monochromatic X-ray beam produced by an appropriate source, either a rotating anode generator or a synchrotron radiation. The waves scattered by the electrons of the macromolecules ordered in the crystal generate a diffraction spot on the screen of the detector. All the spots, regularly spaced, constitute the diffraction pattern. The information contained in each diffraction spot is characterized by

the amplitude and the phase of the structure factor characterizing the corresponding scattered wave. Once a first set of phases is estimated, a first electron density map is calculated. If this map is sufficiently interpretable, the macromolecule can be built step by step in this map. A combination of automated algorithm and manual method available through interactive graphics softwares are used, leading to a final model composed of the three-dimensional coordinates of each atom of the macromolecules. From that first built model, the diffraction intensities are calculated by Fourier transform and compared to the intensities experimentally measured. This comparison allows the step by step improvement of the model. This cyclical process is called the crystallographic refinement, alternating the search for global minimum of energy functions and manual reconstruction of the model. The final step, downstream the structure determination by X-ray diffraction, concerns the interpretation of the structure and its integration into the biological context.

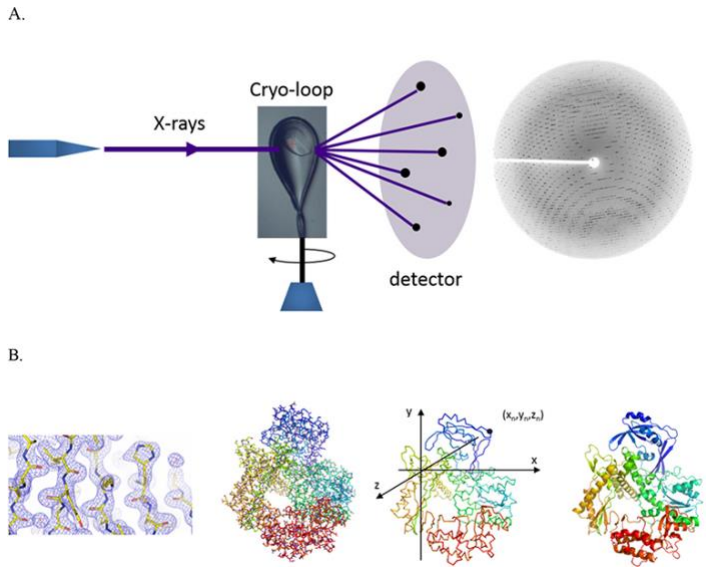


Figure 9. *(A) A monochromatic X-ray beam bombards a crystal frozen in a cryo-loop that rotates on itself. The observed diffraction spots are the result of the impact on the detector of the wave diffracted by the electrons in the crystal. (B) Electron density map of a fragment of a macromolecule is represented (left). The three-dimensional structure of a macromolecule (here a protein) is represented in three ways: all-atoms, backbone and cartoon representation.*[13]

Chapter1 -CONFORMATIONAL ANALYSIS OF PEPTIDES

1.1 A novel β-hairpin peptide derived from the ARC repressor selectively interacts with the major groove of B-DNA

A great chemical challenge consists of using the principles of DNA recognition by Transcription factors (TFs) to design minimized peptides that maintain the DNA affinity and specificity characteristics of the natural counterparts. TFs recognize DNA using a variety of structural motifs. Among those, the ribbon-helix-helix (RHH) proteins, exemplified by the MetJ and ARC repressors, form dimers that insert antiparallel β-sheets into the major groove of DNA. In this context, a peptide mimic of an antiparallel β-sheet is very attractive since it can be obtained by a single peptide chain folding in a β-hairpin structure and can be as short as 14 amino acids or less. In this chapter is presented the design of eight linear and two cyclic dodeca-peptides endowed with β-hairpins. Their DNA binding properties have been investigated using fluorescence spectroscopy together with the conformational analysis through circular dichroism and solution NMR. We found that one of our peptides, **p6**, is able to bind DNA, albeit without sequence selectivity. Notably, it shows a topological selectivity for the major groove of the DNA which is the interaction site of ARC and many other DNA-binding proteins. Moreover, we found that a type I' β-hairpin folding pattern is a favorite peptide structure for interaction with the B-DNA major groove. **p6** is a valuable lead compound for the development of novel analogs with sequence selectivity.

1.1.1 Introduction

Gene expression is regulated through the integrated action of many regulatory elements; among this constellation of components, Trascriptor Factors (TFs) have a leading role in the initiation of gene expression.[14] TFs recognize ds-DNA (double stranded DNA) using a variety of structural motifs but the most relevant contacts to DNA occur in the major groove from amino acids of α-helical regions. The bihelical DNA-binding motif (helix-turn-helix), zinc-finger domains, and the basic region-leucine zipper (bZIP) motif are part of this category. In addition, there are TFs and other DNA-binding proteins that do not rely on α-helices for specific DNA recognition, although they are less common. For instance, the ribbon-helix-helix (RHH) proteins, exemplified by the MetJ and ARC repressors, form dimers that insert antiparallel β-sheets into the major groove of DNA with the side chains on the face of the β-sheet contacting the base pairs. ARC repressor is a transcription factor abundantly expressed in the striatum human muscle and cardiac tissue that selectively interacts with the caspase recruitment domain (CARD) of caspase-2 and caspase-8, thus inhibiting cell death caused by hypoxia and hydrogen peroxidase-mediated cells H9c2, protecting cardiac muscle from postischemic cardiomyopathy.[15] Its abundance in long life cells suggests a protective action on the muscular fibers involved in apoptotic processes caused by mechanical stress and oxidative damage. Building synthetic ARC mimetics able to recognize specific double stranded DNA (ds-DNA) and the design of small molecules capable to control ARC activity, could be valuable approaches in medicinal chemistry.[16] In this context, we designed eight linear and two cyclic dodeca-peptides endowed with β-hairpins derived from the ARC repressor.

1.1.2 Results and Discussion

Design. In the design of the novel β-hairpin peptides we explored different types of structural modifications involving the native amino acid sequence of the ARC protein, primarily by conserving the key amino acids of each β-strand (e.g. QFNLR, **Table 2**) to guarantee the sequence specific interactions with the DNA.[17]A series of well-known dipeptide β-turn inducers have been incorporated between the two strands in place of Trp- Pro sequence, to promote the β-hairpin formation (**Figure 10**).[18]

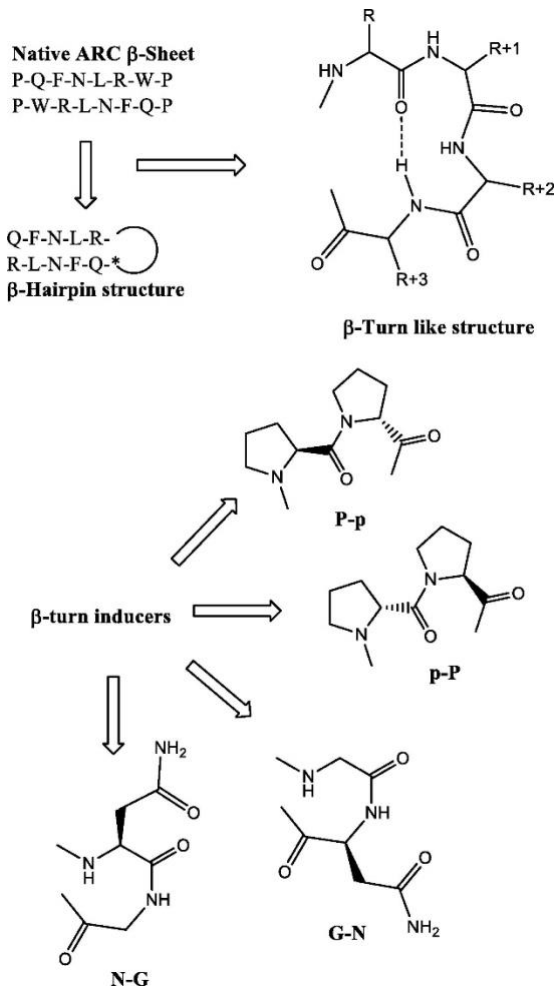


Figure 10. *Native ARC β-sheet and β-turn inducer sequences incorporated in the novel peptides.*

In particular, Asn-Gly (N-G) dipeptide is a type I' β-turn and Gly-Asn (G-N) dipeptide is a type II' β-turn inducer.[19]Also, DPro-Pro (p-P) has been used as nucleating turn of β-hairpins promoting a type II′ β-turn.[20] Furthermore, the enantiomeric Pro-DPro (P-p) sequence was attempted. One-by-one insertion of these four dipeptide β-turn inducers gave peptides **p1**–**p4** reported in **Table 2**.

Table 2. Amino acid sequences of the novel synthetic peptides.

Structure Type	Compounds	Sequences[a]
β-sheet portion	ARC protein	**P-Q-F-N-L-R-W-P//P-Q-F-N-L-R-W-P**
Linear β-hairpins	**p1**	**Q-F-N-L-R-***N-G***-Q-F-N-L-R-**NH_2
	p2	**Q-F-N-L-R-***G-N***-Q-F-N-L-R-**NH_2
	p3	**Q-F-N-L-R-***p-P***-Q-F-N-L-R-** NH_2
	p4	**Q-F-N-L-R-***P-p***-Q-F-N-L-R-** NH_2
	p5	**Q**-W-**N**-W-**R-***N-G***-Q**-W-**N**-W-**R-** NH_2
	p6	**Q**-W-**N**-W-**R-***G-N***-Q**-W-**N**-W-**R-** NH_2
	p7	**Q**-W-**N**-W-**R-***p-P***-Q**-W-**N**-W-**R-** NH_2
	p8	**Q**-W-**N**-W-**R-***P-p***-Q**-W-**N**-W-**R-** NH_2
Cyclic β-hairpins	**p9**	**Q-F-N-***c*[C**R-***N-G***-Q**C**]N-L-R-** NH_2
	p10	**Q-***c*[C**NLR-***N-G***-QFN**C**]R-** NH_2

a Lower-case letter indicates D-amino acids. In bold are shown the conserved natural residues. In italics are shown the β-turn inducers.

Then, considering that in the crystal structure of ARC repressor in complex with its DNA target site (PDB: 1PAR)[17], Phe and Leu residues are buried into the protein hydrophobic core, they were replaced with Trp residues in all peptides **p1**–**p4** to obtain a Trp-Zip like structure,[21] resulting in compounds **p5**–**p8**, respectively (**Table 2**).

To further force the folding of the Asn-Gly β-turn containing peptides into a β-hairpin secondary structure, two cysteine residues have been inserted in place of the faced residues Leu and Phe,[22] in order to close the cycle by a

disulfide bridge. In this way two novel constrained peptides have been obtained (compounds **p9–p10**). All the original sequences have been prepared as C-terminal amide to avoid undesirable unspecific ionic interactions with DNA.

Conformational analysis. The 3D conformational arrangement of such compounds was preliminarily verified through circular dichroism (CD). All peptides were analysed in phosphate buffer and secondary structure content derived using the online server BestSel.[23] Peptides **p2**, **p3**, **p4**, **p6**, and **p8** were found to be compatible with a β-hairpin folding (% of β-sheet structure > 30%) (**Figure 11**, **Table 3).** In particular, **p3** and **p6** showed the highest β-sheet content (about 50%) in the series.

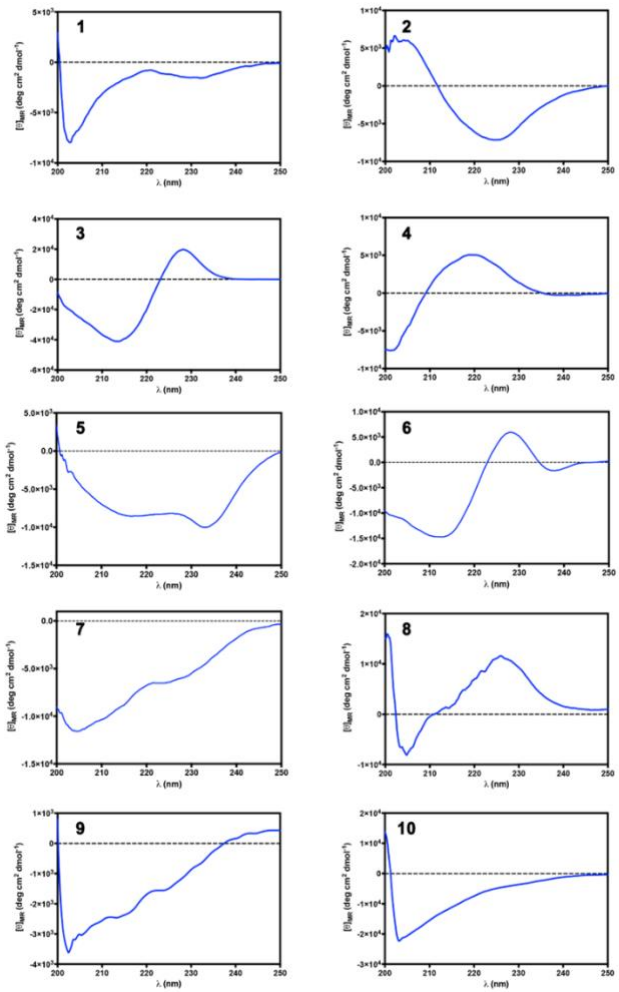


Figure 11. *Circular dichroism of* ***p1-10*** *(125 µM) in 10 mM phosphate buffer (pH 7.5) and 100 mM of NaCl.*

Table 3. Percentage of secondary structure from CD spectra.

Peptide	Helix	β-sheet	Turn	Others
p1	25.0	0.0	8.0	67.0
p2	0.0	**30.9**	18.4	50.7
p3	0.0	**51.2**	0.0	48.8
p4	0.0	**32.4**	21.8	45.7
p5	5.6	10.3	17.4	66.7
p6	13.7	**47.1**	1.9	37.3
p7	44.1	12.3	0.0	43.6
p8	2.5	**37.1**	20.5	39.8
p9	24.7	0.0	12.0	63.4
p10	68.5	0.0	0.0	36.5

Starting from CD results, NMR analysis was performed on peptides **p2**, **p3**, **p4**, **p6**, and **p8**. NMR spectra of all peptides in water solution showed a clear tendency of these peptides to adopt the β-hairpin structure. In particular, a medium range NOE $H\alpha_i$-HN_{i+2} between residue 6 and residue 8 indicated the presence of a β-turn structure about residues 6–7; strong sequential NOEs in the segments 1–5 and 8–12 indicated that those regions are in extended conformation; moreover, inter-strands NOEs between residues 2–11 and 4–9 are clearly diagnostic of an antiparallel β-sheet, thus defining an overall β-hairpin structure. Structure calculation was also performed on the peptide **p6** which, as reported below, showed the most interesting binding properties. The simulated annealing protocol based on the NMR constraints gave a quite well-defined β-hairpin structure with a type I' β-turn on its tip (**Figure 12**). The four Trp residues form a zipper structure on one side of the β-hairpin while Gln (1,8), Arg (5,12) and Asn (3) are oriented on the other side forming a polar surface.

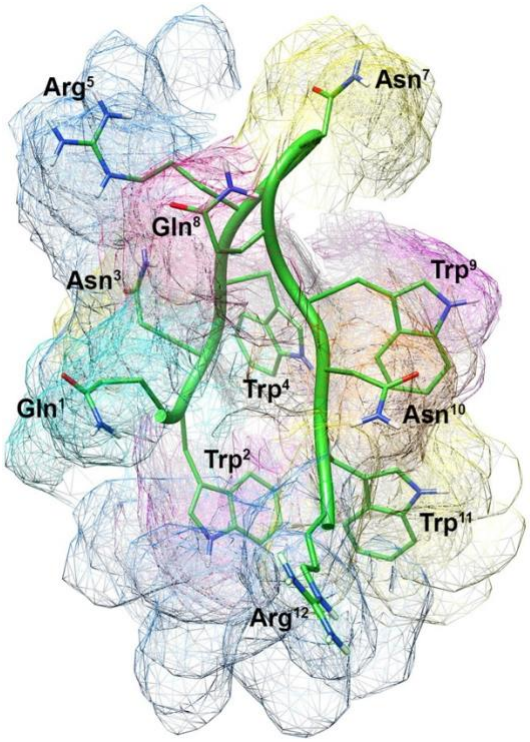


Figure 12. *Lowest energy conformer of peptide* ***p6****. Backbone is evidenced as a ribbon. Side chains of the 10 lowest energy conformers are also shown as mesh surface. Surfaces are distinguished with different colors.*

Interaction with DNA oligomers. The capacity of the selected peptides (**p2**, **p3**, **p4**, **p6**, and **p8**) to interact with ds-DNA, has been tested on the ds-oligonucleotide DNA1 containing the consensus target TAGA (ds-DNA1: 5′-GCGAG TAGA GC TTTTT GC TCTA CTCGC-3′) and DNA2 with a double mutation in the consensus sequence (ds-DNA2: 5′-GCGAG CACA GC TTTTT GC TGTG CTCGC-3′). For this purpose, methyl green (MG) displacement assays were performed. MG is a triphenylmethane dye able to bind to the DNA major groove with a stoichiometry of one molecule per thirteen nucleotides and showing a preferential binding for AT-rich regions[24,25]which is also the interaction site of the ARC repressor. MG is almost non-fluorescent when free in solution, while it becomes intensely fluorescent when bound to DNA.[24] Thus, a peptide-induced MG displacement produces a fluorescence intensity decrease which can be

monitored as a function of the peptide concentration, thus allowing the evaluation of its relative binding affinity to the DNA under examination. We first performed a qualitative analysis of the MG displacement from DNA1 and DNA2 induced by addition of a molar excess of **p2**, **p3**, **p4**, **p6**, and **p8**. Results of these experiments, reported in **Figure 13**, indicate that **p6**, which is characterized by one of the highest β-sheet content, was the most effective of the series in displacing MG from both DNA1 and DNA2.

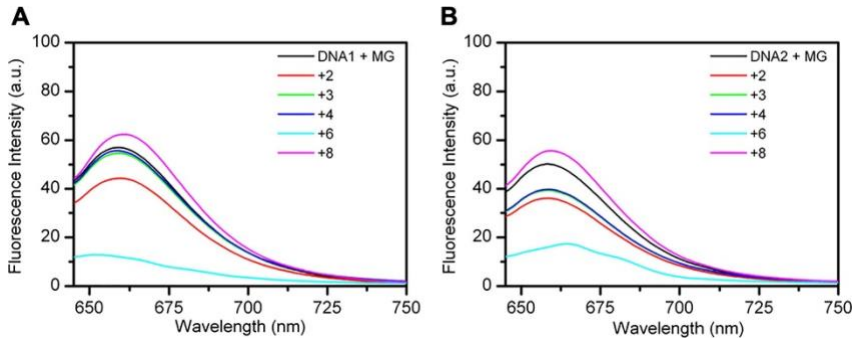


Figure 13. *Displacement of methyl green (MG) from (**A**) DNA1 and (**B**) DNA2 by the investigated peptides.*

Interestingly, **p2**, the second most effective peptide of the tested series, share with **p6** the GN turn sequence (**Table 2**) suggesting that a type I' β-turn is the best suited within a β-hairpin peptide for the interaction with the DNA major groove. Next, the affinity of **p6** for DNA1 and DNA2 was evaluated by measuring the concentration of ligand required to decrease the fluorescence of the probe by 50% (DC_{50}). Results of such experiments showed in **Figure 14** clearly revealed that the binding of **p6** does not strictly depend on the DNA sequence, since DC_{50} values of 25.5 (±1.3) and 12.8 (±0.7) μM were obtained for DNA1 and DNA2, respectively.

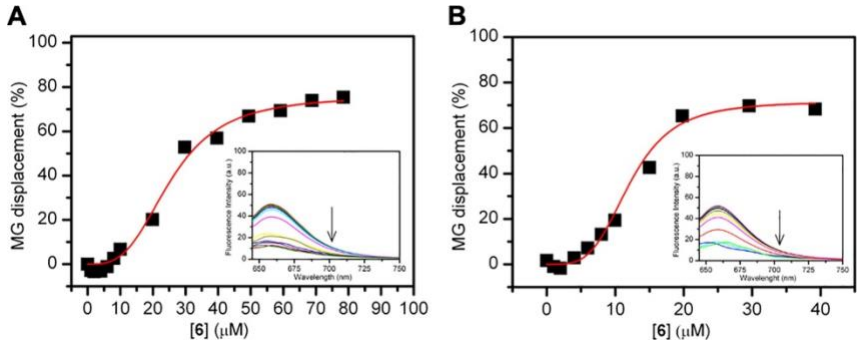


Figure. 14. *Plots of peptide* ***6****-induced MG displacement for (****A****) DNA1 and (****B****) DNA2.*

In order to exclude the possibility of the selected peptides (**p2**, **p3**, **p4**, **p6**, and **p8**) interacting with the minor groove of DNA, the ethidium bromide (EB) displacement assay was performed. EB is known to intercalate between DNA base pairs,[26]and its fluorescence intensity increases remarkably upon binding to DNA, while a marked quenching of the fluorescence is expected if the peptide competes for the same binding site of the dye. None of the investigated peptides induced a significant displacement of EB from both DNA1 and DNA2, apart a very slight effect observed for **p8** (**Figure 15).** To note that peptide **8** showed a very similar behavior with consensus DNA1 and mutated DNA2. These results would discard DNA intercalation of the peptides as the main DNA binding mode.

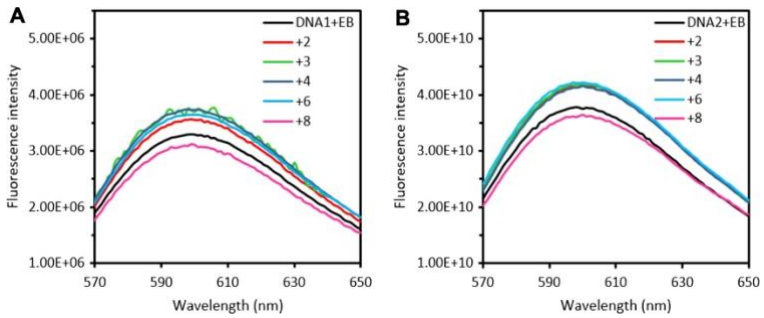


Figure 15. *Displacement of ethidium bromide (EB) from (A) DNA1 and (B) DNA2 by the investigated peptides.*

Interaction of peptide p6 with DNA1 and DNA2 by docking studies. Interaction mode of peptide **p6** with DNA1 and DNA2 was investigated by docking studies. NMR lowest energy structure of peptide **p6** was docked with both DNA1 and DNA2 models. Docking procedures using the program AUTODOCK[27]clustered 100 poses in 24 clusters (10/100 poses in the first cluster) for **p6**/DNA1 and 100 poses in 29 clusters for **p6**/DNA2 complexes (11/100 poses in the first cluster). Statistics and energy terms are reported in **Table S1** (**Supplementary Material**).

Clearly, the high number of clusters indicate that docking did not provide a unique **p6**/DNA1 or **p6**/DNA2 complex (**Figure 16 A, C**) confirming that peptide **p6** has not a sequence selectivity for DNA. Interestingly, peptide **p6** interacts with the major groove of B-DNA according to the experimental displacement results. Moreover, the polar face of the hairpin is mainly involved in the interactions with the DNA while the Trp zipper has only a scaffold stabilization function. **Figure 16B** and **16D** show details on the peptide-base contacts of the lowest energy pose of **p6**/DNA1 and **p6**/DNA2 complexes, respectively. They include charge reinforced hydrogen bonds between guanidino and phosphate groups of Arg5 and A22, and Arg12 and A4; one hydrogen bond between Gln1 and A22; one hydrogen bond between Asn10 and T24 for **p6**/DNA1 complex. Considering **p6**/DNA2 complex, charge reinforced hydrogen bonds between Arg5 and C2, Arg12 with A9 and G10; one hydrogen bond between Asn10 and T19; one hydrogen bond between Gln1 and G1 were observed. To note that the free energy of binding of **p6**/DNA1 is higher than **p6**/DNA2 complex (-4.29 vs -7.99 kcal/mol, **Table S1-Supplementary Material**). Moreover, considering the lowest energy poses, peptide **p6** does not interact within the TAGA box in the **p6**/DNA1 complex, while it displays various contacts with the corresponding CACA box in the **p6**/DNA2 complex. These findings can give the key to

understanding the observed lower affinity of peptide **p6** for DNA1 compared to DNA2.

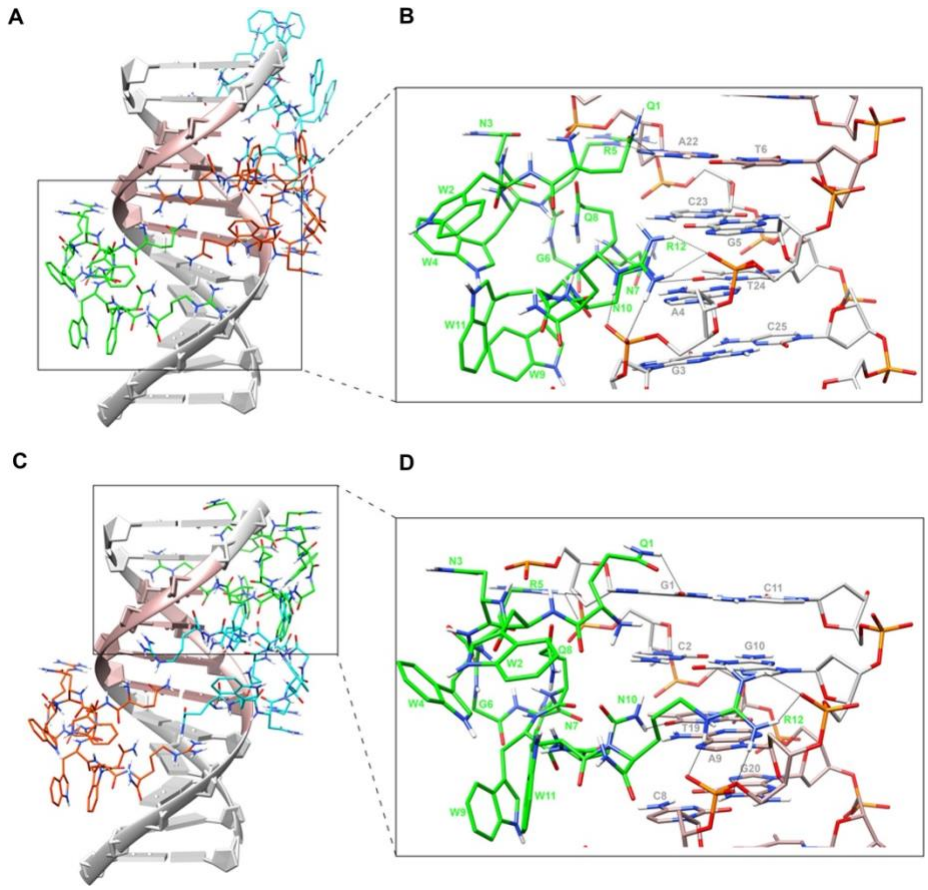


Figure 16. *(**A**) DNA1 or (**C**) DNA2 model in complex with peptide **6**. The three lowest energy poses of each docking calculation are shown (carbon atoms of the absolute lowest pose are shown in green). DNA backbone is represented as gray ribbon, bases are displayed as slab. TAGA or CACA box are evidenced in pink. (**B**) Zoom of the peptide **6** lowest energy pose interacting with DNA1 or (**D**) DNA2. Hydrogens bond are represented with dashed lines. Color codes: polar hydrogen, white; nitrogen, blue; oxygen, red; phosphorous, yellow. Non-polar hydrogens are not shown for clarity.*

1.1.3 Conclusion

The design of peptides with a well-defined structure as ARC mimetics is an appealing task in the field of medicinal chemistry due to its novelty and specificity, because the binding to the DNA is made by a β-sheet domain, which is peculiar, being alpha helix often used for this process. We have designed a series of linear and cyclic β-hairpin dodeca-peptides with the aim to mimic the 3D structure of the ARC β-sheet region interacting with DNA; among them, **p2**, **p3**, **p4**, **p6** and **p8** display a β-hairpin structure. We found that one of our peptides, **p6**, is able to bind DNA, albeit without sequence selectivity. However, it shows a topological selectivity since it binds to the major groove of the DNA which, interestingly, is the interaction site of ARC and many other DNA-binding proteins. Moreover, we found that a type I' β-harpin folding pattern is a favorite peptide structure for interaction with the B-DNA major groove. Certainly, **p6** is a valuable lead compound for the development of novel analogs with sequence selectivity.

1.1.4 Experimental section

Peptides. Peptides **1-10** were purchased from China Peptides (Shangai, China).

Circular dichroism. CD spectra were carried out on a Jasco-715 coupled with a thermostat Nestlab RTE-11, using an acquisition range: 250–190 nm; band width: 2.0 nm; resolution: 0.2 nm; accumulation: 3 scans; sensitivity: 10 mdeg; response time: 0.25 s, speed: 100 nm/min. CD measurements were made in a 2 mm cell at 20 ◦C. Peptides **p1–p10** (125 μM) were dissolved in 10 mM phosphate buffer (pH 7.5) containing 100 mM of NaCl. Secondary structure analysis was based on BestSel algorithm.[23]

NMR spectroscopy. All NMR experiments on β-hairpin dodeca-peptides were recorded on a Varian INOVA 700 MHz spectrometer equipped with a z-gradient 5 mm triple-resonance probe head. All the spectra were recorded at 298 K. The samples were prepared by dissolving the appropriate amount of peptides in 180 μL of H_2O and 20 μL of ^2H_2O (pH 5.5), to obtain a concentration 1 mM in peptide solution. One-dimensional (1D) NMR spectra were recorded in the Fourier mode with quadrature detection. The water signal was suppressed by gradient echo.[28]2D DQF-COSY,[2] TOCSY,[3]and NOESY[4]spectra were recorded in the phase-sensitive mode using the method from States.[29]Data block sizes were 2048 addresses in t2 and 512 equidistant t1 values. Before Fourier transformation, the time domain data matrices were multiplied by shifted sin2 functions in both dimensions. A mixing time of 80 ms was used for the TOCSY experiments. NOESY experiments were run with a mixing time of 100 ms. The qualitative and quantitative analyses of DQF-COSY, TOCSY, and NOESY spectra, were obtained using the interactive program package XEASY.[30] Almost complete ^1H NMR chemical shift assignments were effectively achieved for all peptides according to the Wüthrich[1]procedure via the usual systematic application of DQF-COSY,

TOCSY, and NOESY experiments with the support of the XEASY software package. (**Tables S2–S6, Supplementary Material**).

Structure calculation. NOE-based distance restraints were obtained from NOESY spectra of β-hairpin peptide **p6**. The NOESY cross peaks were integrated with the XEASY program and were converted into upper distance bounds using the CALIBA program incorporated into the program package CYANA.[31] Only NOE derived constraints were considered in the annealing procedures. An ensemble of 100 structures was generated with the simulated annealing of the program CYANA. Then, 10 structures were chosen, whose interproton distances best fitted NOE derived distances, and refined through successive steps of restrained and unrestrained energy minimization calculations using the Discover algorithm (Accelrys, San Diego, CA) and the consistent valence force field.[32] The minimization lowered the total energy of the structures; no residue was found in the disallowed region of the Ramachandran plot. The final structures were analyzed using the InsightII program (Accelrys, San Diego, CA). Molecular graphics images were realized using the UCSF Chimera package.[33]

Methyl green displacement assay. Methyl Green (MG) displacement experiments were performed at 20 ◦C on a FP-8300 spectrofluorometer (Jasco) equipped with a Peltier temperature controller accessory (Jasco PCT-818). A sealed quartz cuvette with a path length of 1 cm was used. Fluorescence spectra were recorded with excitation at 640 nm and emission from 645 to 750 nm. Excitation and emission slits were set both at 5 nm. For the assay 8.2 μg/ mL (1 μM single strand, or 27 μM per nucleotide) of pre-folded DNA1 or DNA2 targets and 100 μg/mL (312 μM per nucleotide) CT-DNA were used. DNA/MG solutions were prepared by mixing each DNA with MG (1.26 μM in the case of DNA1 and DNA2 or 15.3 μM for CT), and the corresponding mixtures were stirred at room temperature for 1 h before

use. The ability of **p2**, **p3**, **p4**, **p6**, and **p8** to displace MG from DNA1, DNA2 was evaluated by adding each peptide to a final concentration of 35 μM (450 μM for CT). Stock solutions of peptides were prepared in ultra-pure H_2O at a concentration of 2 mM. All DNA/MG/peptide mixtures were equilibrated for 3 min before spectra acquisition. For the displacement titration, different amounts of **p6** (from a 20 mM stock solution) were added to the DNA/MG solution (from 0 to 78 μM in the case of DNA1, from 0 to 39 μM for DNA2, and from 0 to 423 μM in the case of CT) followed by a 3 min equilibration time before spectra acquisition. The percentage of displacement was calculated as follows: MG displacement (%) = 100 - [(F/F_0) × 100], where F stands for the intensity of the fluorescence emission signal at 660 nm of MG bound to DNA after each ligand addition and F_0 without added ligand. The percentage of displacement was then plotted as a function of the concentration of added ligand. DC_{50} values were designed as the required concentration to displace 50% MG from each investigated DNA.

Ethidium bromide displacement assay. Ethidium bromide (EB) displacement spectra were obtained with a Jobin-Yvon Fluoromax-3 (DataMax 2.20) coupled to a Wavelength Electronics LFI-3751 temperature controller, using the following settings: increment: 1.0 nm; integration time: 0.1 s; excitation slit width: 4.0 nm; emission slit width: 6.0 nm at 20 °C. In the case of ethidium bromide, the excitation wavelength applied was 540 nm and the emission spectra were acquired from 565 to 650 nm. For each assay 1 μM of pre-folded DNA1 or DNA2 targets and 4.5 μM of EB were used, and the corresponding DNA/EB mixtures were stirred at room temperature for 1 h before use. Next, each peptide (**p2**, **p3**, **p4**, **p6**, and **p8**) was added to the corresponding DNA/dye solution to a final concentration of 15 μM. Stock solutions of peptides were 2 mM in ultra-pure H_2O.

Docking study. DNA1 and DNA2 models were built in B-DNA conformation using the Biopolymer tool of the InsightII program (Accelrys, San Diego). Only the double helix segments of the oligonucleotides were included in the models (bases 1–11 and 17–27). The poses for the **p6**/DNA complex were generated by docking the lowest energy conformers of **p6** obtained by NMR to the DNA1 and DNA2 models using the program AUTODOCK 4.2.[32] Docking statistics are reported in **Table S1** of the **Supplementary Material**. Refinement of lowest energy pose of both **p6**/DNA1 and **p6**/DNA2 complexes was achieved by *in vacuo* energy minimization with the Discover algorithm using the steepest descent and conjugate gradient methods until a RMSD of 0.05 kcal/mol per Å was reached. Energy terms in **Table S1** refer to these optimized complexes.

1.2 Disulfide Bond Replacement with 1,4- and 1,5-Disubstituted [1,2,3]-Triazole on C-X-C Chemokine Receptor Type 4 (CXCR4) Peptide Ligands: Small Changes that Make Big Differences

In this chapter of my PhD thesis, I will be described the investigation of the the structural and biological effects ensuing from the disulfide bond replacement of a potent and selective C-X-C chemokine receptor type 4 (CXCR4) peptide antagonist, with 1,4- and 1,5- disubstituted 1,2,3-triazole moieties. Both strategies produced candidates that showed high affinity and selectivity against CXCR4. Notably, when assessed for their ability to modulate the CXCL12-mediated cell migration, the 1,4-triazole variant conserved the antagonistic effect in the low-mid nanomolar range, while the 1,5-triazole one displayed the ability to activate the migration, becoming the first in class low-molecular-weight CXCR4 peptide agonist. By combining NMR and computational studies, it is provided a valuable model that highlighted differences in the interactions of the two peptidomimetics with the receptor that could account for their different functional profile. Furthermore, these findings could be translated to different GPCR-interacting peptides for the pursuit of novel chemical probes that could assist in dissecting the complex puzzle of this fundamental class of transmembrane receptors.

1.2.1 Introduction

Disulfide bond, also known as disulfide bridge, is an evolutionary conserved post-transcriptional modification involving the sulfhydryl groups of cysteine pairs, which are oxidized to form the so-called cystine moiety.[34]

The development of efficient synthetic strategies[35] offered the opportunity for considering this type of covalent bond as an approach to conformationally constraining linear sequences that would otherwise be flexible in solution.[36]

From a structural point of view, natural or synthetically engineered peptides containing one or multiple disulfide bonds show the propensity to shape in well-defined structural motifs, including β-turn,[37] β-sheet,[38] β-hairpin,[39] and α-helix structures.[40] Unfortunately, the disulfide bridge itself presents potential drawbacks, such as susceptibility to enzymes (i.e. disulfide isomerases) or to reducing agents and thiols (i.e. glutathione),[41] which can affect this covalent bond leading to structural changes with consequent loss of biological activity and/or stability.

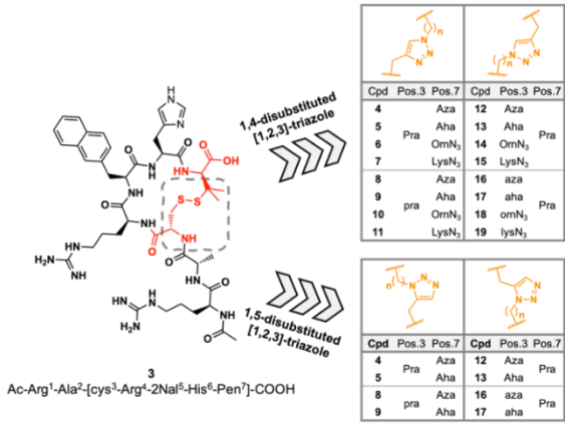
In light of these considerations, alternative cyclization strategies such as lactams,[42] dicarba-,[43] thioethers,[44] diselenides[45] and 1,4- and 1,5-disubstituted [1,2,3]-triazoles bridges,[46] have been proposed as bio-isosteric surrogates of this key structural element. Among them, the [1,2,3]-triazole scaffold gained increasing interest due to its synthetic advantages, which made this approach a good option not only for introducing conformational constrains in linear sequences[47] but also for site-specific conjugation with non-peptidic moieties (i.e. carbohydrates, drugs, metal chelating agents, etc.).[48] Recently, 1,4- and 1,5- disubstituted 1,2,3-triazoles have been employed to engineer peptides interacting with G-coupled receptors (GPCR) such as the AGRP chimeric peptide[49] and the urotensin-II peptide,[50] leading to compounds endowed with different potency and selectivity profiles.

In this chapter of my PhD thesis, it is described the effect of the disulfide/triazole replacement within a new family of cyclic heptapeptides acting as chemokine receptor 4 (CXCR4) antagonists. CXCR4 is a G-protein coupled receptor (GPCR), which modulates many physiological functions, including the bone marrow progenitor cell retention and chemotaxis, by interacting with its ligand C-X-C motif chemokine 12 (CXCL12).[51] Lately, the CXCR4/CXCL12 pair has emerged as a key player in promoting tumor growth, invasion, angiogenesis, and metastasis in more than 30 different human cancers.[52] Several strategies have been investigated to disrupt the CXCR4/CXCL12 interactions, leading to the development of potent CXCR4 antagonists including antibodies,[53] small-molecules,[54] and peptides.[55] Nonetheless, growing shreds of evidence suggest that the use of CXCR4 agonists, rather than antagonists, can be an alternative approach to negatively modulate the CXCR4/CXCL12 axis either by rapidly desensitizing the receptor or altering the CXCL12 gradient which is needed to provide efficient chemotaxis.[56] However, the majority of the available CXCR4 peptide agonists do not possess "druglike" features being characterized either by low affinity or high molecular weight structure (i.e. more than 20 amino acid residues).[57]

The following describes how the systematic substitution of the disulfide bridge of one the most active CXCR4 peptide antagonist known in the literature, peptide **3,**[58]with 1,4- and 1,5-disubstituted [1,2,3]-triazolyl moieties has led to the identification of an agonist and an antagonist that showed high affinity and selectivity against CXCR4. To this end, NMR and molecular modeling studies were performed with aim to provide a model that could disclose the molecular bases for their binding to CXCR4.

1.2.2 Results and discussion

Design and synthesis. The new series of 1,4- and 1,5-triazolic cyclic peptides were designed choosing the cyclo-heptapeptide **3** (Ac-Arg^1-Ala^2-[cys^3-Arg^4-2Nal^5-His^6-Pen^7]-COOH) (**Figure 17**) as model structure and replacing the residues involved in the cyclization (cys^3 and Pen^7) with ω-alkynyl and ω-azido amino acids. In general, it proceeded according to three lines of intervention: a) variation of the cyclization amplitude using ω-azido amino acids with increasing side chain lengths (azidoalanine or Aza, azidohomoalanine or Aha, azidornitine or OrnN_3 and azidolysine or LysN_3); b) variation of the amino acid stereochemistry replacing cys^3; c) different orientation of the [1,2,3]-triazolyl bridge relative to the peptidic backbone.



Cpd	Pos.3	Pos.7	Cpd	Pos.3	Pos.7
4	Pra	Aza	12	Aza	Pra
5		Aha	13	Aha	
6		OmN_3	14	OmN_3	
7		$LysN_3$	15	$LysN_3$	
8	pra	Aza	16	aza	Pra
9		Aha	17	aha	
10		OmN_3	18	omN_3	
11		$LysN_3$	19	$lysN_3$	

Cpd	Pos.3	Pos.7	Cpd	Pos.3	Pos.7
4	Pra	Aza	12	Aza	Pra
5		Aha	13	Aha	
8	pra	Aza	16	aza	Pra
9		Aha	17	aha	

Figure 17. *Series of 1,4- and 1,5-disubstituted [1,2,3]-triazole peptides inspired to peptide* ***3.*** *Lower-case letters indicate D-amino acids.*

The 1,4-disubstituted triazolyl-bridged series was synthesized combining the entire set of the above described modifications and led to the compounds **4-19** (**Table 4**). Conversely, the study on the 1,5- disubstituted triazole peptides focused on the sole use of Aha/aha and Aza/aza residues as ω-azido amino acids. This choice was dictated by the results of the 1,4-triazolyl tether replacement as well as by a previous report indicating Aha and Aza residues as the best options to achieve cyclization closely related to a cystine moiety

in terms of number of bonds and cyclization geometry,[59] led to the compounds **21-27** (**Table 4**).

Table 4. Sequences of the parent peptide, **3**, and synthesized peptides.

Cpd	SEQUENCE
3	Ac-Arg-Ala-[**cys**-Arg-2Nal-His-**Pen**]-COOH
4	Ac-Arg-Ala-[**Pra**-Arg-2Nal-His-**Aza**]-COOH
5	Ac-Arg-Ala-[**Pra**-Arg-2Nal-His-**Aha**]-COOH
6	Ac-Arg-Ala-[**Pra**-Arg-2Nal-His-**OrnN_3**]-COOH
7	Ac-Arg-Ala-[**Pra**-Arg-2Nal-His-**LysN_3**]-COOH
8	Ac-Arg-Ala-[**pra**-Arg-2Nal-His-**Aza**]-COOH
9	Ac-Arg-Ala-[**pra**-Arg-2Nal-His-**Aha**]-COOH
10	Ac-Arg-Ala-[**pra**-Arg-2Nal-His-**OrnN_3**]-COOH
11	Ac-Arg-Ala-[**pra**-Arg-2Nal-His-**LysN_3**]-COOH
12	Ac-Arg-Ala-[**Aza**-Arg-2Nal-His-**Pra**]-COOH
13	Ac-Arg-Ala-[**Aha**-Arg-2Nal-His-**Pra**]-COOH
14	Ac-Arg-Ala-[**OrnN_3**-Arg-2Nal-His-**Pra**]-COOH
15	Ac-Arg-Ala-[**LysN_3**-Arg-2Nal-His-**Pra**]-COOH
16	Ac-Arg-Ala-[**aza**-Arg-2Nal-His-**Pra**]-COOH
17	Ac-Arg-Ala-[**aha**-Arg-2Nal-His-**Pra**]-COOH
18	Ac-Arg-Ala-[**ornN_3**-Arg-2Nal-His-**Pra**]-COOH
19	Ac-Arg-Ala-[**lysN_3**-Arg-2Nal-His-**Pra**]-COOH
20	Ac-Arg-Ala-[**Pra**-Arg-2Nal-His-**Aza**]-COOH
21	Ac-Arg-Ala-[**Pra**-Arg-2Nal-His-**Aha**]-COOH
22	Ac-Arg-Ala-[**pra**-Arg-2Nal-His-**Aza**]-COOH
23	Ac-Arg-Ala-[**pra**-Arg-2Nal-His-**Aha**]-COOH
24	Ac-Arg-Ala-[**Aza**-Arg-2Nal-His-**Pra**]-COOH
25	Ac-Arg-Ala-[**Aha**-Arg-2Nal-His-**Pra**]-COOH
26	Ac-Arg-Ala-[**aza**-Arg-2Nal-His-**Pra**]-COOH
27	Ac-Arg-Ala-[**aha**-Arg-2Nal-His-**Pra**]-COOH

Bolded amino acids represent those involved in the study. Lower-case letters indicate D-amino acids.

Peptides were synthesized in a stepwise manner according to the ultrasound-assisted solid phase peptide synthesis (US-SPPS) protocols recently reported.[60] In general, the linear oligomers were assembled on a 2-chloro tritylchloride (2CTC) resin by iterative cycles of Fmoc-deprotection and coupling reaction steps to later undergo to N-terminal acetylation. The side chain-to-side chain cyclization was achieved performing a CuAAC or

RuAAC under different conditions according to the regioisomerism of the desired [1,2,3]-triazolyl tether. (**Figure S1 Supplementary Material**)

CXCR4 Binding Assays. The newly 1,4- and 1,5- triazole derivatives were preliminarily assessed for their capability to impair the interaction between the receptor and an anti-CXCR4 PE-conjugated antibody (Clone 12G5) in CEM-CCRF human T leukemia cells overexpressing CXCR4. At first, we evaluated the effects produced by the replacement of the cystine moiety between cys^3 and Pen^7 (**3**), with a 1,4-triazole bridging endowed with different sizes. As reported in **Table 5**, the introduction of cycles obtained by using $LysN_3$/ $lysN_3$, $OrnN_3$/$ornN_3$ led to a substantial loss of affinity, with the resulting peptides displaying IC_{50} values from sub-micromolar to micromolar range (**6**, **7**, **10**, **11**, **14**, **15**, **18**, **19**). Among the others, with the sole exception of **5** (IC_{50}= 51 ± 8 nM) and **12** (IC_{50}= 18 ± 6 nM), the peptides featuring Aza/aza (**4**, **8** and **16**) and Aha/aha (**9**, **13** and **17**), appeared to be significantly less able to affect the interaction between the anti-CXCR4 antibody and the receptor. Interestingly, the comparison between the **5**, **9** (1,4- triazole variant with Pra/Aha and pra/Aha, respectively) and **12**, **16** (1,4- triazole variant with Aza/Pra and aza/Pra, respectively) pairs indicates that unlike the parent peptides containing the disulfide bridge, the configuration of the amino acid in position 3 is determinant, with the L- preferred to the D-configuration. Likewise, also the orientation of the 1,4-triazole significantly affected the affinity as underscored by comparing the **5**, **13** (1,4- triazole variant with Pra/Aha and Aha/Pra, respectively) and **4**, **12** (1,4- triazole variant with Pra/Aza and Aza/Pra, respectively) couples of regioisomers.

Table 5. IC_{50} values (Mean ± SD) of 1,4-disubstituted 1,2,3-triazole derivatives of **3** obtained measuring the inhibition of CXCR4 binding of anti-CXCR4 PE-antibodies (Clone 12G5).

Cpd	SEQUENCE	IC_{50} (nM)
3	Ac-Arg-Ala-[**cys**-Arg-2Nal-His-**Pen**]-COOH	1.5 ± 0.5
4	Ac-Arg-Ala-[**Pra**-Arg-2Nal-His-**Aza**]-COOH	630 ± 28
5	Ac-Arg-Ala-[**Pra**-Arg-2Nal-His-**Aha**]-COOH	51 ± 8
6	Ac-Arg-Ala-[**Pra**-Arg-2Nal-His-**OrnN_3**]-COOH	2550 ± 350
7	Ac-Arg-Ala-[**Pra**-Arg-2Nal-His-**LysN_3**]-COOH	>10000
8	Ac-Arg-Ala-[**pra**-Arg-2Nal-His-**Aza**]-COOH	1550 ± 700
9	Ac-Arg-Ala-[**pra**-Arg-2Nal-His-**Aha**]-COOH	3750 ± 130
10	Ac-Arg-Ala-[**pra**-Arg-2Nal-His-**OrnN_3**]-COOH	>10000
11	Ac-Arg-Ala-[**pra**-Arg-2Nal-His-**LysN_3**]-COOH	>10000
12	Ac-Arg-Ala-[**Aza**-Arg-2Nal-His-**Pra**]-COOH	18 ± 6
13	Ac-Arg-Ala-[**Aha**-Arg-2Nal-His-**Pra**]-COOH	788 ± 85
14	Ac-Arg-Ala-[**OrnN_3**-Arg-2Nal-His-**Pra**]-COOH	371 ± 61
15	Ac-Arg-Ala-[**LysN_3**-Arg-2Nal-His-**Pra**]-COOH	>10000
16	Ac-Arg-Ala-[**aza**-Arg-2Nal-His-**Pra**]-COOH	3730 ± 890
17	Ac-Arg-Ala-[**aha**-Arg-2Nal-His-**Pra**]-COOH	610 ± 224
18	Ac-Arg-Ala-[**ornN_3**-Arg-2Nal-His-**Pra**]-COOH	262 ± 176
19	Ac-Arg-Ala-[**lysN_3**-Arg-2Nal-His-**Pra**]-COOH	409 ± 122

Bolded amino acids represent non-proteinogenic amino acids.

Concerning peptides **20**-**27**, the introduction of the 1,5-triazole bridge did not afford, in general, an improvement in the capability to compete for CXCR4 (**Table 6**). When the cyclization was achieved using Aza/aza (**24** and **26**), we observed a dramatic loss of activity (IC_{50} > 10 μM) probably as a consequence of the excessive constrain induced by such short side-chained amino acids. Among the peptides in which the 1,5-triazole bridge was obtained using Aha/aha, **21** and **23** exhibited affinity in the same nanomolar range, although differing in the configuration of the propargylalanine (Pra) (IC_{50} = 22 ± 3 nM and IC_{50} = 22 ± 5 nM, respectively). Similarly to what observed for the 1,4 series, again the comparison between peptides bearing a reverse 1,5-triazolyl bridge, such as the **21**, **25** (1,5-triazole variant with Pra/Aha and Aha/Pra, respectively) and **23**, **27** (1,5-triazole variant with pra/Aha and aha/Pra, respectively) pairs indicates a significant reduction of the capability to interact with CXCR4 (up to 100 fold), thus strengthening the concept that the inversion of this structural feature might be detrimental for the affinity retention.

Table 6. IC_{50} values (mean ± SD) of 1,5-disubstituted 1,2,3-triazole derivatives of **3** obtained measuring the inhibition of CXCR4 binding of anti-CXCR4 PE-antibodies (Clone 12G5).

Cpd	SEQUENCE	IC_{50} (nM)
3	Ac-Arg-Ala-[**cys**-Arg-2Nal-His-**Pen**]-COOH	1.5 ± 0.5
20	Ac-Arg-Ala-[**Pra**-Arg-2Nal-His-**Aza**]-COOH	NT
21	Ac-Arg-Ala-[**Pra**-Arg-2Nal-His-**Aha**]-COOH	22 ± 3
22	Ac-Arg-Ala-[**pra**-Arg-2Nal-His-**Aza**]-COOH	NT
23	Ac-Arg-Ala-[**pra**-Arg-2Nal-His-**Aha**]-COOH	22 ± 5
24	Ac-Arg-Ala-[**Aza**-Arg-2Nal-His-**Pra**]-COOH	>10000
25	Ac-Arg-Ala-[**Aha**-Arg-2Nal-His-**Pra**]-COOH	1790 ± 350
26	Ac-Arg-Ala-[**aza**-Arg-2Nal-His-**Pra**]-COOH	>10000
27	Ac-Arg-Ala-[**aha**-Arg-2Nal-His-**Pra**]-COOH	150 ± 18

NT: Not Tested compounds; Bolded amino acids represent those involved in the study.

^{125}I-CXCL12 Competition Binding Assay. Among the best performing peptides in the competition assay with anti-CXCR4 PE-conjugated antibody, we selected **12** and **21** as the most representative 1,4- and 1,5 triazole derivatives, respectively. These compounds were thus further assessed on CCRF-CEM cells in a competition binding assay using a radiolabeled version of CXCL12 as a ligand (^{125}I-CXCL12). Indeed, the two peptides **12** and **21** share the same backbone stereochemistry offering us the opportunity to dissect the sole effect of the triazolyl cyclization on the ability to compete with ^{125}I-CXCL12. As shown in **Figure 18**, peptides **12** and **21** displaced the ^{125}I-labeled CXCL12 (IC_{50} = 36.0 ± 0.3 nM and IC_{50} = 30.5 ± 7.7 nM for **12** and **21**, respectively) with an affinity comparable to the disulfide bridged cyclic peptide **3** (IC_{50} = 20 ± 2 nM). Interestingly, these results only partially recapitulated those obtained using the anti-CXCR4 antibody (**Tables 5 and 6**). Indeed, while in former testing **12** and **21** proved to be 10-fold less active than peptide **3**, in the current competition experiment they displaced the ^{125}I-CXCL12 with the same range of potency (IC_{50} into two-digit nanomolar range). The difference in affinity of **12** and **21** in the two assays could be explained by the antibody and CXCL12 having different binding sites on CXCR4, thus differentially affected by the two peptidomimetics.

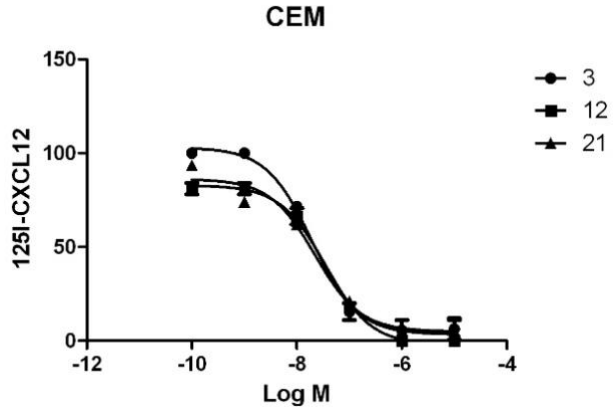


Figure 18. *Competition binding assays of **3**, **12**, and **21** with ^{125}I-CXCL12 using CCRF-CEM cells incubated with 60 pM ^{125}I-CXCL12 as a tracer and increasing concentrations (from 100 pM to 10 μM) of either **3**, **12** and **21** as competitors. The binding curve was fitted to a one-site binding model. Results showed (mean ± SD) are the average of three experiments*

Receptor Selectivity. As reported in literature, peptide **3** can selectively interact with CXCR4 sparing other chemokine receptors, such as CXCR3 and CXCR7. These two chemokine receptors allowed testing the affinity of the peptides versus a related (CXCR7) and an unrelated (CXCR3) receptors. Indeed, CXCR7 (also known as ACKR3) shares with CXCR4 the endogenous ligand CXCL12 while CXCR3 is activated upon binding of a different chemokine, namely CXCL11. Thus, **12** and **21** were tested against CXCR7 and CXCR3 for probing the effect of the disulfide bond replacement on their selectivity profile. This experiment was carried out by using CXCR3-overexpressing COLO205 human colorectal cancer cell line and the ACKR3-overexpressing MCF-7 human breast cancer cells, which were incubated with the peptides and the binding specificity was evaluated using anti-CXCR3 (anti-CXCR3 FITCantibody - R&D FAB160F clone 49801) and anti-ACKR3 (anti- ACKR3 APC-antibody - R&D FAB4227A clone 11G8) antibodies. As depicted in **Figure 19**, **12** and **21** did not show significant

binding to the two receptors up to 10 μM concentrations, indicating that the applied modifications did not affect the selectivity towards CXCR4.

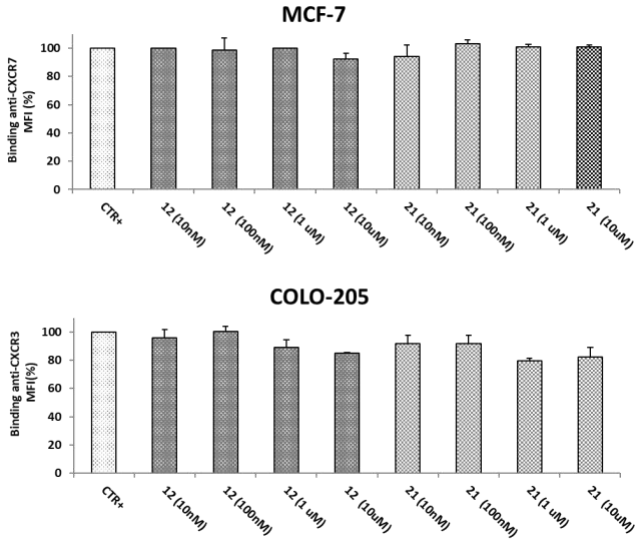


Figure 19. *Binding of **12** and **21** to MCF-7 breast cancer cells and to COLO205 colon cancer cells, overexpressing CXCR7 and CXCR3, respectively. The binding was evaluated through flow cytometry using anti-CXCR3 FITC-antibody and anti-CXCR7 APC-antibody. Data are presented as bar graphs showing mean ± SD.*

Cell Migration Experiments. The inhibition of the CXCL12-dependent migration of the cells overexpressing CXCR4 is the main functional characteristic of the CXCR4 antagonists. Thus, we evaluated the impact of **12** and **21** on the CXCL12-mediated migration of CCRF-CEM cells in transwell-based assays, including **3** as a reference compound. As depicted in Figure 5, **12** was able to inhibit CCRF-CEM migration similarly to parent peptide even at a nanomolar concentration. Surprisingly, **21**, although endowed with a comparable affinity for the receptor did not show any ability to influence the CXCL12 mediated migration questioning its capacity to exert an antagonistic effect on the receptor (**Figure 20**).

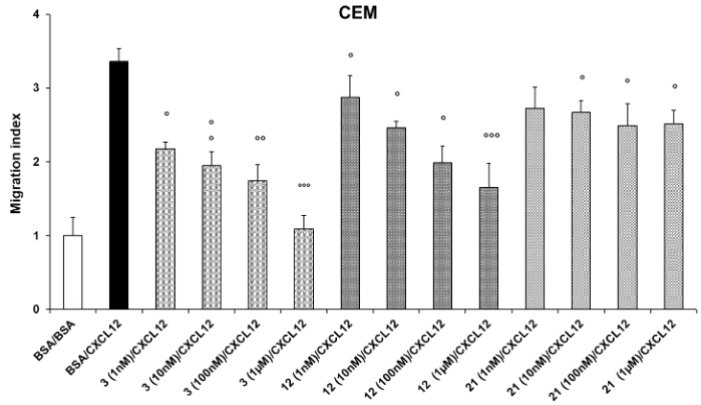


Figure 20. *Comparative CXCL12-dependent cells migration experiments on **3**, **12** and **21**. Experiments were conducted on the CCRF-CEM cells. Migrated cells on the lower surface were fixed, stained with DAPI, and counted microscopically. The results are expressed as the migration index relative to migration in the presence of BSA alone. Data are presented as bar graph showing mean ± SD.*** $p<0.001$, BSA/CXCL12 vs BSA/BSA; ° $p<0.05$, °°$p<0.01$, °°°$p<0.001$ **3** or **12**/CXCL12 or **21**/CXCL12 vs BSA/CXCL12*

To this end, we designed an additional migration experiment aimed at assessing the ability of **21** to prompt cellular mobility by activating the receptor in the absence of CXCL12. Thus, the ability of **21** to induce migration of CXCR4-overexpressing cells was evaluated and compared with that of the endogenous ligand CXCL12 and the partial agonist AMD-3100. As depicted in **Figure 21**, **21** significantly enhanced the migration of CCRF-CEM cells with efficacy and potency comparable to the full agonist CXCL12 (10 nM and 1 nM, respectively) and higher than AMD-3100 suggesting that the modifications introduced conferred agonistic properties to the peptide. The activity switching of **21** is further confirmed by its bell-shaped dose/response curve, which is characteristic of the compounds acting as CXCR4 agonists. Interestingly, to the best of our knowledge, **21** represents the first in class low-molecular weight CXCR4 peptide agonist resulting from small modifications performed on previously reported antagonist. This opens the way towards the evaluation of other known CXCR4 peptide antagonists, which could serve as scaffold for the synthesis of novel potential agonists for this relevant biological target.

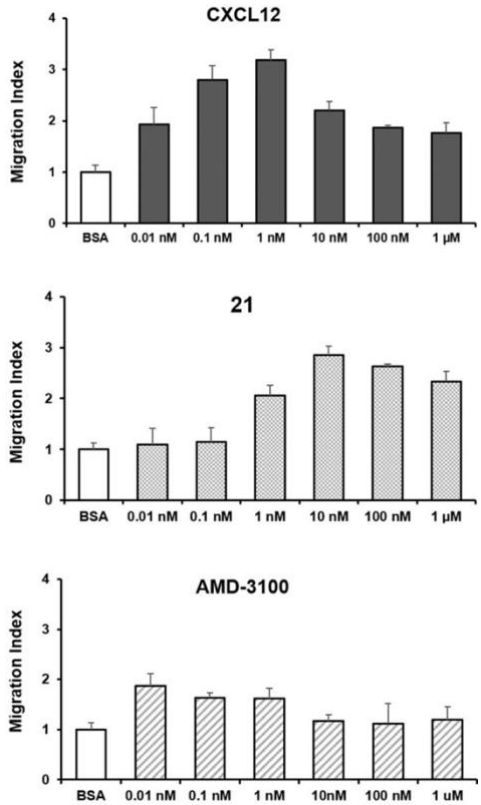


Figure 21. *Comparative CXCL12, 21, AMD-3100 dependent- cell migration experiments. Experiments were conducted on the CCRF-CEM cells. Migrated cells on the lower surface were fixed, stained with DAPi, and counted microscopically. The results are expressed as the migration index relative to migration in the presence of BSA alone. Data are presented as a bar graph showing the mean ± SD.*

NMR analysis and Molecular Modeling. To shed light on the conformation adopted by **12** and **21** and correlate this with their different biological behavior NMR experiments were first performed. Complete ^{1}H NMR chemical shift assignments (**Tables S7 and S8 Supplementary Material**) were performed according to the Wüthrich procedure.[1] DQFCOSY,[2] TOCSY,[3] and NOESY[4] experiments were carried out in a 200 mM SDS micellar solution. The use of SDS micelles to study the conformational properties of CXCR4 ligands is motivated on the basis of their interaction with a membrane receptor. For peptides acting as ligand of membrane receptors, the use of membrane mimetic media, such as SDS or dodecylphosphocoline, is suggested hypothesizing a membrane-assisted

mechanism of interactions between the peptides and their receptors.[61] Many NMR parameters and, in particular, several medium-range NOEs (**Tables S9 and S10, Supplementary Material**) indicated a well-defined conformation for both peptides. Upfield shifts of sidechain protons of the Arg^4 and NOE contacts between Arg^4 and $2Nal^5$ point to the spatial proximity of these two side chains. NMR constraints from the SDS micelle solution were used as the input data for a simulated annealing structure calculation. An ensemble of ten well-defined structures was obtained for both peptides. Peptide **12** structures showed two distorted (type IV) β-turns along residues 2-5 and 3-6. In peptide **21**, three non-canonical β - turns could be found along residues 2-5, 3-6, and 4-7. The lowest energy conformer of each peptide and their superposition with the parent peptide **3** are displayed in **Figure 22**. The comparison highlights that peptide **21** can be better superimposed with **3** than **12** (**Figure 22d** and **c**, respectively). Particularly, the side chains of the pharmacophoric residues Arg^1, Arg^4, $2Nal^5$ and His^6 occupy approximately the same spatial regions in the two peptides **3** and **21**.

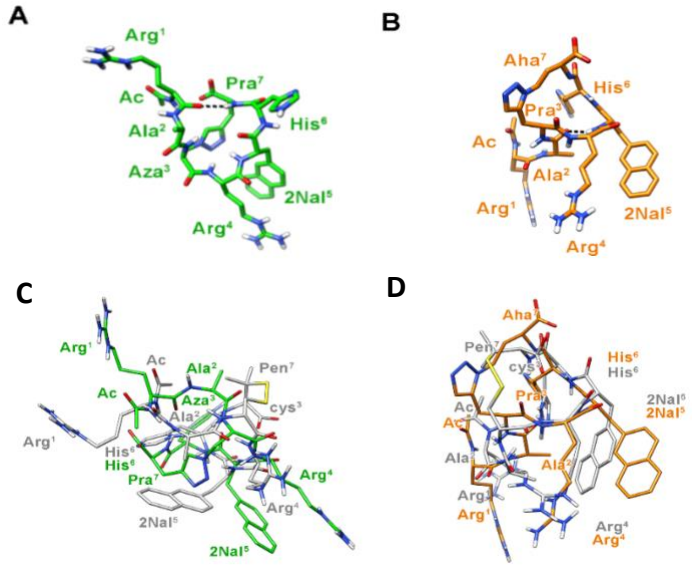


Figure 22. *NMR structures of the lowest energy conformer of **12** (orange, A) and **21** (green, B) and their superposition (C and D respectively), on the C α atoms of residues 3-7, with peptide **3** (grey sticks). Hydrogen bonds are shown as dashed black lines.*

Then, to elucidate at an atomic level the binding mode of the newly synthesized peptides to CXCR4 and provide insights into their pharmacological behavior, extensive computational studies were performed starting from the NMR structure of compounds **12** and **21**.

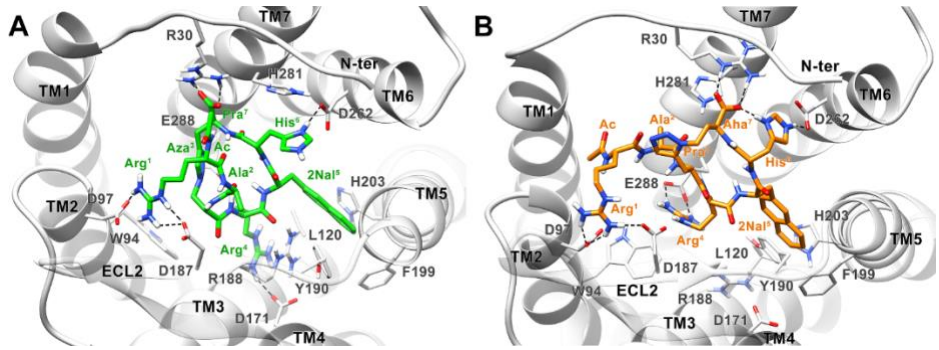


Figure 23. *(A) Binding mode of **12** (A, green sticks) and **21** (B, orange sticks) at the CXCR4 binding site (gray cartoons) obtained through MD simulations. Receptor amino acids important for peptide binding are shown as sticks. The color code for the heteroatoms is: blue for nitrogen, red for oxygen. Hydrogen bonds are displayed as dashed black lines while non-polar hydrogens are omitted for sake of clarity.*

Initially, the Glide-SP peptide docking was applied for the preliminary prediction of the binding conformations of each ligand. The top ranked docking poses were then evaluated based on the interaction schemes reported in previous computational studies on peptide CXCR4 ligands such as **3**.[58] The docking pose shows that **12** is hosted into the transmembrane (TM) bundle of CXCR4, occupying both the minor (TMS1) and the major pockets (TMS2) of the ligand binding site (**Figure 23A**). The final MD pose of **21** is similar to that of **12**, with the two peptides contacting almost the same hot spots in the CXCR4 binding site (**Figure 23B**). The major difference resides in the relative position and, consequently, in the interaction partner of Arg^4 which, in the case of **21**, forms a salt-bridge with the side chain of $E288^{7.39}$ instead of $D171^{4.60}$ as observed for **12**. This is likely due to the presence of diverse triazole bridges in the backbone of the two compounds (Aza – 1,4-triazole in **12** *vs* Aha – 1,5-triazole in **21**), which, as shown by NMR (**Figure 22**), can

influence their bioactive conformation and, thus, their amino acids side chains orientation. Probably, for the same reason, an intramolecular H-bond can be formed between the His^{6} imidazolinium group and the Aha^{7} carboxylate moiety of **21**, which can further stabilize the binding conformation of this ligand (**Figure 23B**). The interaction modes of **12** and **21** are in line with the nanomolar potency of these peptides and might also account for their distinct behavior in cell migration experiments. In particular, the salt bridge established through Arg^{4} by **12** and **21** with $D171^{4.60}$ and $E288^{7.39}$, respectively, might correlate with the different abilities of the two compounds to trigger the receptor activation. Indeed, some of the most potent peptide CXCR4 antagonists, such as CVX15, FC131, and its analogs, have been shown to directly contact $D171^{4.60}$ residue, as here found for **12** but not for **21**.[62,63] On the other hand, previous mutagenesis experiments[64] and molecular modeling studies[65] highlighted the role of $E288^{7.39}$ in the activation of CXCR4 by potent agonists such as CXCL12 and its chimeric derivatives. In particular, $E288^{7.39}$ was reported to take part in an extensive interaction network with neighboring conserved residues, such as $Y45^{1.39}$, $W94^{2.60}$ and $Y116^{3.32}$, whose changes can initiate the receptor activation process, similar to what observed in other GPCRs.[66]

1.2.3 Conclusion

In this chapter of my thesis, it is described a systematic replacement of the disulfide bond involving cys and Pen residues of the cyclic heptapeptide **3**, a potent and selective CXCR4 antagonist, with 1,4- and 1,5-disubstituted [1,2,3]-triazolyl bridges. Different structural features of the triazole bridging were investigated including the size of the cyclization ring, the orientation of the [1,2,3]-triazole with respect to the main backbone, and the stereochemistry of the amino acids involved in the cyclization. This investigation led to the identification of **12** (1,4-triazole) and **21** (1,5-triazole) showing high affinity and selectivity against CXCR4. CXCL12-mediated cell migration assay confirmed that **12** retained the antagonistic effect with a potency comparable to the parent peptide **3**. Conversely, **21** significantly activated the migration, showing a dose/response curve distinctive of the compounds acting as CXCR4 agonists. Their different efficacy was rationalized by NMR and computational studies that showed slight differences in the binding mode of the two peptides, which might account for the distinct ability of **12** and **21** in modulating CXCR4. Remarkably, the activity switching observed for **21** renders this cyclic heptapeptide the first in class low-molecular-weight peptide acting as CXCR4 agonist. Furthermore, the findings of this study support the application of the strategy described above on different natural and synthetic GPCR-interacting peptides containing disulfide bonds to obtain chemical probes that could ultimately help in dissecting the complex puzzle of this fundamental class of transmembrane receptors.

1.2.4 Experimental section

General synthetic procedure for 1,4-triazolic peptides (5-7 and 9-19). Exemplified procedure for peptide 5 2-Chloro trytylchloride resin (51 mg, 0.072 mmol, 1.40 mmol/g, 1.2 equiv. relative to the mmol of first amino acid) was swelled in anhydrous DMF over 30 min and then drained on a vacuum manifold for solid phase extraction without any further treatment. Subsequently, a solution of FmocAha-OH (22 mg, 0.06 mmol) and DIPEA (25 μL, 0.12 mmol, 2 equiv. relative to the millimoles of the first amino acid) in anhydrous DMF (1.5 mL) was added and the so obtained mixture was shaken overnight at room temperature. The residual chloride reactive groups were capped by adding a previously mixed solution of DIPEA (25 μL, 0.144 mmol) in DCM/MeOH (9:1, 3 mL) and allowing the resin to gently shake for 1 hour, before being washed with DMF (2 × 0.5 min) and DCM (2 × 0.5 min). After initial functionalization step, linear peptides were assembled on solid support by a conventional Fmoc/tBu approach and following an US-SPPS protocol.[57] In general, Fmoc-deprotection was carried out by suspending a SPPS reactor in an ultrasonic bath (SONOREX RK 52 H), and exposing the resin to the effect of ultrasounds in the presence of a 20% piperidine solution in DMF (2 × 1 min, 1.5 mL for each treatment). Likewise coupling reactions were carried out irradiating the resin with ultrasounds for five minutes in the presence of a DMF solution (1.2 mL) containing: Fmoc-amino acids (0.12 mmol, 2 equiv.), HBTU (46 mg, 0.12 mmol, 2 equiv.) and HOBt (18 mg, 0.12 mmol, 2 equiv.) as coupling reactants, and DIPEA (42 μL, 0.24 mmol, 4 equiv.) as base. Completion of the reaction steps was ascertained by Kaiser ninhydrin or TNBS tests. Between each synthetic step the resin was washed with DMF (2 × 0.5 min) and DCM (2 × 0.5 min) to wash out the residual chemicals. Once the aminoacidic sequence was assembled, Fmoc was removed as on last amino acid and the N-terminal primary amino group was acetylated shaking the resin for ten minutes with a solution of Ac_2O (11 μL,

0.12 mmol, 2 equiv.) and DIPEA (42 µL, 0.24 mmol, 4 equiv.) in DMF (1.2 mL). The so obtained resin-bound peptides were washed with DMF (2 × 0.5 min), DCM (2 × 0.5 min), and Et_2O (3 × 0.5 min) and dried exhaustively. Then the oligomers were cleaved from the solid support by treatment with a solution of TFA/TIS (95:5, 1.5 mL) for 3 h at room temperature. The resin was filtered and the crude peptides precipitated from the TFA solution, diluting to 13 mL with cold Et_2O and then centrifuged (6000 rpm × 15 min). The supernatant was carefully removed and the precipitate suspended again in 13 mL Et_2O as described above. The resulting wet solid was dried under reduced pressure and then transferred into a round-bottom flask where it was redissolved in a mixture of tBuOH/H_2O (1.5:1; 30 mL; 2 mM). To the stirring solution, it was added ascorbic acid (106 mg, 0.6 mmol, 10 equiv.) and $CuSO_4$*$10H_2O$ (37 mg, 0.15 mmol, 2.5 equiv.), complete dissolution was essential, therefore the flask was sonicated until a clear solution was observed, then the mixture was allowed to stir overnight. The organic solvent was evaporated and the turbid solution was filtered through a syringe filter (Sartorius minisart SRP 25, 25 mm, pores diameter 0.45 µm). The clear solution was then purified by preparative RP-HPLC (solvent A: water + 0.1 % TFA; solvent B: acetonitrile + 0.1 % TFA; from 10 to 60% of solvent B over 20 min, flow rate: 10 mL min−1). Fractions of interest were collected, evaporated by organic solvent, frozen, and then lyophilized. Obtained products were characterized by analytical HPLC and HRMS

Synthetic procedure for the 1,4-triazole peptides (4 and 8). Peptides with Azidoalanine (Aza/aza) were assembled initially functionalizing a 2- CTC resin with Fmoc-L-His(Trt)-OH following the procedure previously described. In detail, the 2-CTC resin (70 mg, 0.1125 mmol, 1.60 mmol/g, 1.5 equiv relative to the millimoles of the first amino acid) was swelled in anhydrous DMF over 30 min and then drained on solid-phase peptide

manifold without any further treatment. Subsequently, a solution of FmocL-His(Trt)-OH (47 mg, 0.075 mmol, 1 equiv.) and DIPEA (26 µL, 0.15 mmol, 2 equiv relative to the millimoles of the first amino acid) in anhydrous DMF (1.5 mL) was added and the so obtained mixture was shaken overnight at room temperature. As previously described, the residual chloride reactive groups were capped by adding a previously mixed solution of DIPEA (26 µL, 0.15 mmol) in DCM/MeOH (9:1, 3 mL) and allowing the resin to gently shake for 1 hour, before being washed with DMF (2 × 0.5 min) and DCM (2 × 0.5 min). By this stage the linear peptides were assembled on solid support according to the aforementioned USSPPS protocol, until final acetylation was accomplished. Intermolecular CuAAC was carried out by treating for 16 hours the resin with a solution of Fmoc-L-Aza-OH (40 mg, 0.1125 mmol, 1.5 equiv.), CuI (14 mg, 0.075 mmol, 1 equiv.), ascorbic acid (26 mg, 0.15 mmol, 2 equiv.), 2,4,6- collidine (49 µL, 0.375 mmol, 5 equiv.), and DIPEA (66 µL, 0.375 mmol, 5 equiv.) in DMF (3 mL). The resin was washed exhaustively with DMF (3 × 0.5 min) and then treated with a 0.02 M solution of potassium N,Ndiethyldithiocarbammate in DMF (4 × 15 min) to wash the solid support out of the copper catalyst. Fmoc protective group was removed from the so anchored amino acid by piperidine solution and then the full-protected peptides were cleaved from the resin by treating with 5 mL of DCM/TFE/acetic acid 8:1:1 (2 × 1 hour). The solution was collected by filtration in a round bottom flask and then evaporated under reduce pressure. The white solid was dissolved in anhydrous THF (15 mL) and then allowed to react overnight with PyAOP (86 mg, 0.165 mmol, 2.2 equiv.), 0.6 M HOAt solution in DMF (275 µL, 0.165 mmol, 2.2 equiv.) and DIPEA (114 µL, 0.66 mmol, 8.8 equiv.). The mixture was evaporated under vacuum and the so obtained crude treated with a solution of TFA/TIS (95:5, 2 mL) for 3 hours. The solution volume was reduced up to around 1 mL, transferred into a falcon tube, the crude peptide was precipitated with approximately 13 mL of cold

Et2O, and then centrifuged (6000 rpm × 15 min). The supernatant was carefully removed and the precipitate suspended again in 13 mL Et2O as described above. The resulting wet solid was dried under reduced pressure and then purified by preparative RP-HPLC (solvent A: water + 0.1 % TFA; solvent B: acetonitrile + 0.1 % TFA; from 10 to 60% of solvent B over 20 min, flow rate: 10 mL min−1). Fractions of interest were collected, evaporated by organic solvent, frozen, and then lyophilized. Obtained products were characterized by analytical HPLC and HRMS.

General synthetic procedure for 1,5-triazole peptides (20-27). Exemplified procedure for peptide 24 Preparation of 1,5-triazole derivatives was carried out using a 2-CTC resin (70 mg, 0.1125 mmol, 1.60 mmol/g, 1.5 equiv relative to the millimoles of the first amino acid). It was swelled in anhydrous DMF over 30 min and then drained on solid-phase peptide manifold without any further treatment. Subsequently, a solution of Fmoc-L-Pra-OH (25 mg, 0.075 mmol, 1 equiv.) and DIPEA (26 μL, 0.15 mmol, 2 equiv relative to the millimoles of the first amino acid) in anhydrous DMF (1.5 mL) was added and the so obtained mixture was shaken overnight at room temperature. The residual chloride reactive groups were capped by adding a previously mixed solution of DIPEA (25 μL, 0.144 mmol) in DCM/MeOH (9:1, 3 mL) and allowing the resin to gently shake for 1 hour, before being washed with DMF (2 × 0.5 min) and DCM (2 × 0.5 min). After that, the solid phase assembly of linear peptides was conducted following the US-SPPS protocol as already described above. Upon completion of these synthetic steps, the resin was washed with DMF (3 × 0.5 min), DCM (3 × 0.5 min), Et2O (3 × 0.5 min), and dried exhaustively. The resin was transferred in a microwave reactor (5 mL capacity) along with a stirring bar, sealed and swelled in anhydrous DMF (1 mL) for 15 minutes under argon flush. Then a solution of Cp*RuCl(PPh3)2 (9 mg, 0.012 mmol, 20% mol.) in anhydrous

DMF (0.5 mL) was added dropwise and the mixture stirred for further 20 minutes under argon atmosphere. The vessel was placed in a microwave reactor and heated at 65 °C for 150 minutes (330 rpm stirring speed to avoid polystyrene support stress). Completion of the RuAAC was accomplished by sampling the resin and comparing the HPLC profile of the peptides before and after the reaction was performed. Once completion was ascertained, the resin was washed exhaustively with DMF (3 × 1 min) and Et2O (3 × 1min) before cleavage and purification steps were carried out as abovementioned. Obtained products were characterized by analytical HPLC and HRMS.

12G5 anti-CXCR4-PE Competitive Binding. 5×10^5 CCRF-CEM cells were preincubated with increasing peptide concentrations (10-11 to 10-5 M) in the binding buffer (PBS 1× plus 0.2% BSA and 0.1% NaN_3) for 45 min at 37 °C, 5% CO_2 and then labeled for 30 min using an anti-CXCR4 PE-antibody (FAB170P, clone 12G5, R&D Systems, Minneapolis, MN, USA). A FACSCanto II flow cytometer (B.D. Biosciences, CA, USA) and the Diva version 6.1.1 software were used to collect and analyze the data. The half-maximal inhibitory concentration (IC_{50}) values were calculated using GraphPad Prism software.

^{125}I-CXCL12 Competitive Binding. CCRF-CEM cells ($1\cdot10^6$ cells/sample) were suspended in binding buffer (PBS containing 50 mM HEPES, 1 mM $CaCl_2$, 5 mM $MgCl_2$, 0.5% BSA, and 0.3 mM NaN_3) and incubated with 60 pM of $^{[125I]}$CXCL12 (Perkin- Elmer, 2200 Ci/mmol) plus increasing concentrations (10^{-11} to 10^{-5} M) of the respective peptide of interest (**3**, **12** and **21**). The amount of bound radioactivity was determined after 1 h at 4 °C using a γ counter and expressed as a percent of total added activity. The half-maximal inhibitory concentration (IC_{50}) values were calculated using GraphPad Prism software.

Migration Assay. To perform migration assays, 24-well trans-well chambers (Corning, NY) containing inserts with polycarbonate membranes (5 μm pore-size) were used. The upper chamber, which contained the membrane insert, was precoated with collagen (human collagen type I/III) and fibronectin (20ug/mL each) following standard protocols. CCRF-CEM (human Tlymphoblast leukemia cell line) cells were pleated in upper chambers (2.5×10^5 cells/well) in RPMI medium with 0.5% BSA in presence of dilutions of the various peptides; 100ng/mL CXCL12 was added in the lower chambers. When assessing peptides agonist activity, dilutions of the various peptides in RPMI medium (0.5% BSA) were added to the lower chambers. After incubation for 6 h at 37 °C in a humidified incubator in 5% CO_2, the non-migrating cells were removed from the upper chamber using a cotton swab, and the cells that had migrated to the lower surface of the membrane insert were fixed in 4% (w/v) paraformaldehyde in PBS and stained for 15 min with DAPI. Cells were visualized under the fluorescent microscope (Carl Zeiss, Axio Scope.A1) and counted (cells in ten randomly chosen fields). Migration was reported as migration index, the ratio between the number of cells migrating toward CXCL12 and the number of cells migrating toward BSA 0.5% medium.

NMR spectroscopy. The samples for NMR spectroscopy were prepared by dissolving the appropriate amount of peptides **12** and **21** in 180 μl of ^1H_2O (pH 5.5), 20 μl of ^2H_2O to obtain a concentration 2 mM of peptide and 200 mM of SDS-d_{25} in a 3 mm tube. All NMR spectra were acquired at 298 K on a Bruker Avance NEO 700 MHz spectrometer equipped with a Z-gradient cryoprobe. The spectra were calibrated relative to TSP (0.00 ppm) as an internal standard. Two-dimensional (2D) DQF-COSY,[2] TOCSY[3], and NOESY[4] spectra, were recorded in the phase-sensitive mode using the method from States.[29] A mixing time of 80 ms and 100 ms was used for the

TOCSY and NOESY experiments, respectively. 1D and 2D NMR spectra were processed with the Bruker TOPSPIN 4.0.8 software packages. The temperature coefficients of the amide proton chemical shifts were calculated from 1D ^{1}H NMR and 2D TOCSY experiments performed at different temperatures in the range 298-313 K by means of linear regression. Complete ^{1}H NMR chemical shift assignments (**Tables S7 and S8**) were performed for peptides according to the Wüthrich procedure[1] using the interactive program package XEASY.[30]

3D Structural Calculations. The NOE-based distance restraints were obtained from NOESY spectra collected with a 100 ms mixing time. The NOE cross-peaks were integrated with the XEASY program and were converted into upper distance bounds using the CALIBA program incorporated into the CYANA program package.[31] Only NOE derived constraints were considered in the annealing procedures. The restraints applied during the calculations are reported in **Tables S9** and **S10**. NMR-derived upper bounds were imposed as semi parabolic penalty functions with force constants of 16 Kcal mol-1 Å^{-2}. A distance maximum force constant of 250 Kcal/ mol-1 Å^{-2} was used. Cyclic peptides **12** and **21** were built using the Insight Builder module (Accelrys Software Inc., San Diego). Atomic potentials and charges were assigned using the consistent valence force field (CVFF).[32] The conformational space of compounds was sampled through 100 cycles of restrained Simulated Annealing (ε = 1r) using the Insight Discover module. Resulting conformations were then subjected to restrained Molecular Mechanics (MM) energy minimization within Insight Discover module (ε = 1r) until the maximum RMS derivative was less than 0.001 kcal/Å, using Conjugate Gradient as minimization algorithm. Finally, conformations were subjected to 100 steps of unrestrained MM Conjugate Gradient energy minimization. From the produced 100 conformations, 10

structures, whose interprotonic distances best fitted NOE derived distances, were chosen for further analysis. Molecular graphics images of the structure of **12** and **21** were realized using the UCSF Chimera package.[33]

Molecular Docking. Docking of the NMR-predicted conformations of **12** and **21** was performed in the crystal structure of CXCR4 in complex with CVX15 (PDB code: 3OE0).[62] Both the ligands and the receptor were prepared using the Protein Preparation Wizard tool, implemented in the Maestro Suite 2019.[67] Correct bond orders were assigned, missing hydrogen atoms added and all the water molecules deleted from the receptor structure. Then, a prediction of the side chains ionization and tautomeric states was performed using Epik.[68] However, for both the ligands' His^6 residue, all the three possible states were taken into account. Finally, the receptor hydrogen-bonding network was optimized, and the position of all the hydrogens minimized. For the grid generation, a virtual box of 30 Å × 30 Å × 30 Å, surrounding the inner ligand binding cavity site, was created by the means of the Receptor Grid Generator tool of Glide 8.5.[69] Docking calculations were performed employing the Glide SP-peptide protocol and the OPLS3A force field.[70] The peptides' backbone was kept fixed to preserve the NMR-predicted conformation, while an enhanced sampling of the side chains orientation was allowed. Thus, the obtained solutions were clustered based on the ligand r.m.s.d. (cutoff = 2.0 Å) and ranked according to the Glide SP scoring function.[69]

1.3 Glucopeptides derived from myelin-relevant proteins and hyperglucosylated non-typeable *Haemophilus influenzae* bacterial adhesin cross-react with multiple sclerosis specific antibodies: a step forward in the identification of native autoantigens in multiple sclerosis.

Multiple sclerosis (MS) is an inflammatory and autoimmune disorder, in which an antibody-mediated demyelination mechanism plays a critical role. MS is considered an autoimmune disease, but its aetiology is still largely unknown, although it is established that complex interactions between environmental factors and multiple genetic factors are involved. Indeed, the progression of MS is also linked to exogenous infectious agents expressing antigenic molecules, which may mimic the structure and/or conformation of endogenous mammalian extracellular membrane embedded (glyco)proteins. We prepared two glucosylated peptides derived from the human myelin proteins, i.e., Oligodendrocyte-Myelin Glycoprotein (OMGp) and Reticulon-4 receptor (RTN4R), selected by a bioinformatic approach for their conformational homology with CSF114(Glc), a synthetic antigen which is specifically recognized by antibodies in MS patients' sera. We report herein the antigenic properties of these peptides, showing on one hand that MS patient antibodies recognize the two glucosylated peptides, and on the other that these antibodies cross-react with CSF114(Glc) and with the previously described hyper-glucosylated non-typeable *H. influenzae* bacterial adhesin protein HMW1ct (Glc). These observations point to an immunological association between human and bacterial protein antigens, underpinning the hypothesis that molecular mimicry triggers breakdown of self-tolerance in MS and suggesting that RTN4R and OMGp can be considered as autoantigens.

1.3.1 Introduction

Multiple sclerosis (MS) is a complex, heterogeneous, inflammatory disorder characterized by a loss of myelin sheath surrounding the nerve axons in the central nervous system (CNS). MS is considered an autoimmune disease, but its aetiology is still largely unknown, although it is established that complex interactions between environmental factors and multiple genetic factors are involved.[71,72] At the immunological level, the hypothesis that an antibody-mediated demyelination mechanism might contribute to the immunopathology of MS is increasingly recognized.[73–76] Indeed, much effort has been made to characterize the antibody response present in MS patients in order to find early predictors for diagnosis and disease progression.[77,78] In recent years, interest has focused on minor myelin antigens among which myelin oligodendrocyte glycoprotein (MOG) is an interesting candidate autoantigen. Indeed, anti-MOG antibodies are known to play a key role in the development of demyelination in a number of in vivo and in vitro experimental systems.[79] It was demonstrated that the synthetic glucosylated myelin oligodendrocyte glycoprotein fragment (Asn^{31}(Glc)hMOG(30–50)) was able to detect autoantibodies in MS patients' sera by enzyme-linked immunosorbent assay.[80] The ability of the glucosylated sequence to detect autoantibodies in multiple sclerosis patients' sera was correlated to the N-linked glucosyl moiety.[79] Hence, the recognition properties of the molecule were optimized through the design and screening of focused libraries of glycopeptides by a "Chemical Reverse Approach", which led to the development of a specific antigenic probe, termed CSF114(Glc). An immunoenzymatic assay based on this synthetic glycopeptide was shown to identify autoantibodies in patients' sera as biomarkers of multiple sclerosis.[81,82] The glycopeptide is characterized by a β-turn structure bearing as minimal epitope a β-D-glucopyranosyl moiety linked to an Asn residue on the tip of the turn,[83,84] possibly reproducing an aberrant N-glucosylation of

myelin proteins fundamental for autoantibody recognition.[85,86] Subsequently, looking for a putative native autoantigen mimicked by CSF114(Glc), it was implemented a bioinformatic approach to screen human myelin proteins for homologies to CSF114(Glc). Since the recognition of the synthetic antigen by MS specific autoantibodies appears to be dependent on the peptide sequence and on its predominant β-turn conformation, it was performed both a sequence and a 3D alignment study.[87] This analysis yielded three sequences featuring either sequence or conformational homology with the designed glucopeptide CSF114(Glc), belonging to three relevant myelin proteins: *Factor Associated with Neutral sphingomyelinase activation protein* (FAN), *Oligodendrocyte-Myelin Glycoprotein* (OMGp) and *Reticulon-4 Receptor* (RTN4R). In a previous study,[87] the three selected sequences, modified with a β-D-glucopyranosyl moiety on an Asn residue and tested in competitive ELISA on MS patient sera, were found to cross-react with anti-CSF114(Glc) antibodies, thus demonstrating a mimicry between the designed antigenic probe and myelin proteins relevant in MS. Finally, considering that the human glycoproteome repertoire does not include the simple N-glucosylation of asparagine, i.e. the Asn(Glc) moiety, and taking into account that the progression of MS is also linked to exogenous infectious agents expressing antigenic molecules, which may mimic the structure and/or conformation of endogenous mammalian extracellular membrane embedded (glyco)proteins, attention was turned also to bacterial proteins. Accordingly, it was demonstrated that a cell-surface adhesin protein of non-typeable *Haemophilus influenzae* (NTHi) termed HMW1 and expressing a large number of N-glucosylations, is preferentially recognized by antibodies from sera of an MS patient subpopulation cross-reacting with CSF114(Glc).[88] This was the first example of an N-glucosylated native antigen that can be considered a relevant candidate for triggering pathogenic antibodies in MS, due to a possible molecular mimicry between self-molecules and an

exogenous (bacterial) antigen. Based on these evidences, the aim of the present study is to verify the possible cross-reactivity of MS patient antibodies recognizing the three (gluco)peptides selected through to the bioinformatic approach, i.e., FAN, OMGp, and RTN4R, and the recombinant hyperglucosylated C-terminal fragment of NTHi, residues 1205-1536, i.e., HMW1ct (Glc). Confirming the structural correspondence between an exogenous protein and a physiological one, is the basis of the hypothesis that molecular mimicry triggers the breakdown of self-tolerance in Multiple Sclerosis.

1.3.2 Results and Discussion

The peptides were synthesized in the native and the N-glucosylated version: FAN(635-655) (**1**), RTN4R(173-191) (**2**), OMGp(162-180) (**3**), [Asn^{641}(Glc)]FAN(635−655) (**4**), [Asn^{153}(Glc)]RTN4R(147-165) (**5**), and [Asn^{168}(Glc)]OMGp(186−204) (**6**) (**Table 7**).

Table 7*. Sequences of FAN, RTN4R, OMGp and CSF114(Glc) peptides.*

Peptide	Fragment	Sequence
1	FAN(635-655)	GITVSR**N**GSSVFTTSQDSTLK
2	RTN4R(147-165)	TFRDLG**N**LTHLFLHGNRIS
3	OMGp(162-180)	TLINLT**N**LTHLYLHNNKFT
4	[Asn^{641}(Glc)]FAN(635−655)	GITVSRN(Glc)GSSVFTTSQDSTLK
5	[Asn^{153}(Glc)]RTN4R(147-165)	TFRDLGN(Glc)LTHLFLHGNRIS
6	[Asn^{168}(Glc)]OMGp(162-180)	TLINLTN(Glc)LTHLYLHNNKFT
	CSF114(Glc)	TPRVERN(Glc)GHSVFLAPYGWMVK

The glucopeptides **1**-**6** were synthesized by MW-SPPS, following the 9-fluorenylmethoxycarbonyl (Fmoc)/tBu strategy and introducing the Fmoc-Asn[Glc(OAc)4]-OH building block during the stepwise synthesis, as previously reported.[86] All the peptides were purified to homogeneity by Reverse-Phase Flash Liquid Chromatography and analytically characterized by analytical HPLC coupled to a single quadrupole ESI-MS.

Circular dichroism. The conformation of the three glucosylated peptides **4**-**6** in water was explored by circular dichroism (CD) in a phosphate buffer solution (**Figure 24**). For comparison, peptide CSF114(Glc) was also considered in this analysis. The presence of a minimum around 200 nm, diagnostic of random coil conformation, indicated that peptides are highly flexible in this buffer. However, this minimum is more intense in the case of

peptide **4** suggesting that this is the less ordered among the investigated peptides. In the case of peptides **5**, **6**, and CSF114(Glc) it can be observed a shift of the minimum from 199 nm to 200 (for peptide **6** and CSF114(Glc)) and 201 nm (for peptide **5**) that, together with the already mentioned reduction of intensity, should point to some secondary structure stabilization. The secondary structure content (**Table 8**) was predicted based on the CD spectrum using the online server for protein secondary structure analyses BestSel.[23] As expected, all peptides have high percentage of unstructured conformation. The prediction also returned β-strand and β-turn structures that are the key elements of the β-hairpin in CSF114(Glc).[73,84] Notably, the percentage of β-strand content in the peptides is in the order **5** > **6** > **4** (**Table 8**).

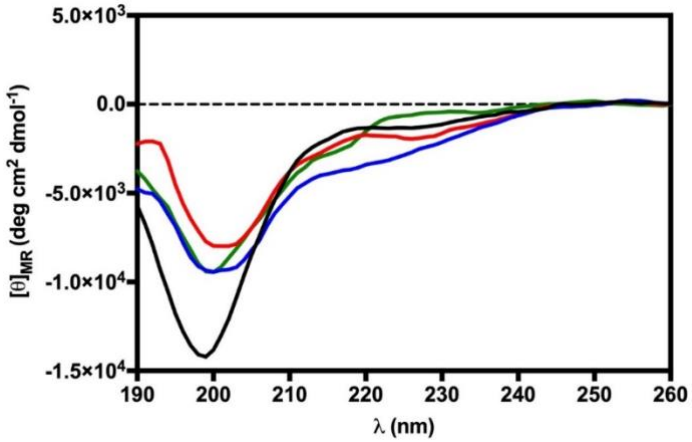


Figure 24*. CD spectra of peptide **4** (black line), **5** (red line), **6** (blue line), and CSF114(Glc) (green line) in PBS solution*

Table 8. Percentage of secondary structure from CD spectra

Peptide	α-helix	β-strand	β-turn	Random coil
4	0.0	21.2	16.4	62.4
5	1.9	33.3	19.3	45.5
6	4.5	31.0	17.9	46.6
CSF114(Glc)	2.4	32.2	16.2	49.2

Inhibition ELISA. Inhibition ELISA experiments were performed to evaluate antibody cross-reactivity between the selected N-glucosylated peptides of myelin proteins and HMW1ct(Glc) [**I(Glc)**]. The N-glucosylated peptides [Asn^{641}(Glc)]FAN(635−655) (**4**), [Asn^{153}(Glc)]RTN4R(147-165) (**5**), and [Asn^{168}(Glc)]OMGp(162-180) (**6**) were coated separately and the unglucosylated peptides **1**, **2**, **3**, the glucopeptides **4**, **5**, **6**, and the hyperglucosylated HMW1ct(Glc) protein **I(Glc)** were used as inhibitors at different concentrations with a reference MS serum diluted 1:100 (MS14), known to contain antibodies that recognize HMW1ct(Glc). The results obtained (**Figure 25**) show that antibodies that recognize the N-glucosylated peptides **4**, **5**, and **6**, used as antigens on the plate, are inhibited by the hyperglucosylated protein HMW1ct(Glc) [**I(Glc)**].

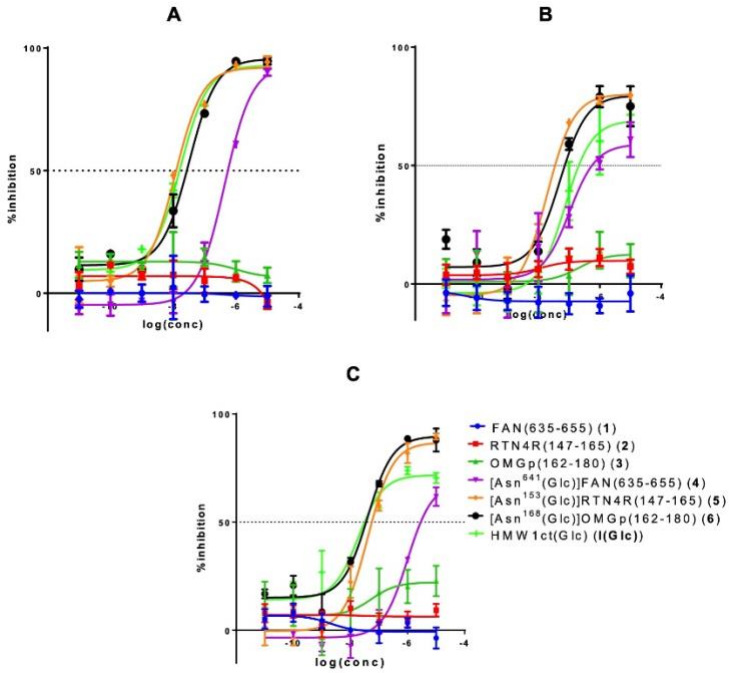


Figure 25. *Inhibition of binding of MS14 serum IgG antibodies to (A) [Asn^{641}(Glc)]FAN(635−655), (B) [Asn^{153}(Glc)]RTN4R(147-165), and (C) [Asn^{168}(Glc)]OMGp(162-180) by peptides **1**, **2**, **3**, glucopeptides **4**, **5**, **6**, and hyperglucosylated HMW1ct(Glc) protein I(Glc) at different concentrations.*

Moreover, self-cross-reactivity was observed for each one of the glucosylated peptides. On the other hand, unglucosylated peptides **1**, **2**, and **3** are not able to inhibit the antibody binding, thus confirming the fundamental role of the Asn(Glc) moiety as the minimal epitope recognized by MS sera.[73] The IC_{50} values (**Table S11, Supplementary Materials**) show that the glucosylated peptide [Asn^{641}(Glc)]FAN(635−655) (**4**) is generally the weakest inhibitor (IC_{50}: $5.0 \cdot 10^{-7}$ M), while the glucosylated [Asn^{153}(Glc)]RTN4R(147-165) (**5**) is the strongest (IC_{50}: $3.0 \cdot 10^{-8}$ M). We subsequently performed the inhibition experiments coating the plates with the hyperglucosylated adhesin HMW1ct(Glc) [**I(Glc)**] and using all the peptides as inhibitors, in order to verify if the cross-reactivity between the glucosylated peptides and the hyperglucosylated protein was observed also when the latter high affinity antigenic probe is coated on the plate, while the peptide inhibitors are in solution. We preliminarily observed that the three unglucosylated peptide sequences, lacking Asn(Glc), were not able to inhibit MS serum antibodies recognizing HMW1ct(Glc). Subsequently, as shown in **Figure 26**, we demonstrated that only the two glucosylated peptides [Asn^{153}(Glc)]RTN4R(147-165) (**5**) and [Asn^{168}(Glc)]OMGp(162-180) (**6**) were able to cross-react with Abs that recognize HMW1ct(Glc), inhibiting the antibody binding with an IC_{50} of $9.96 \cdot 10^{-7}$ M and $2.73 \cdot 10^{-7}$ M, respectively (**Table S11, Supplementary Materials**), thus confirming the presence of a shared epitope. At variance, the glucopeptide [Asn^{641}(Glc)]FAN(635−655) (**4**) was not able to inhibit antibody binding.

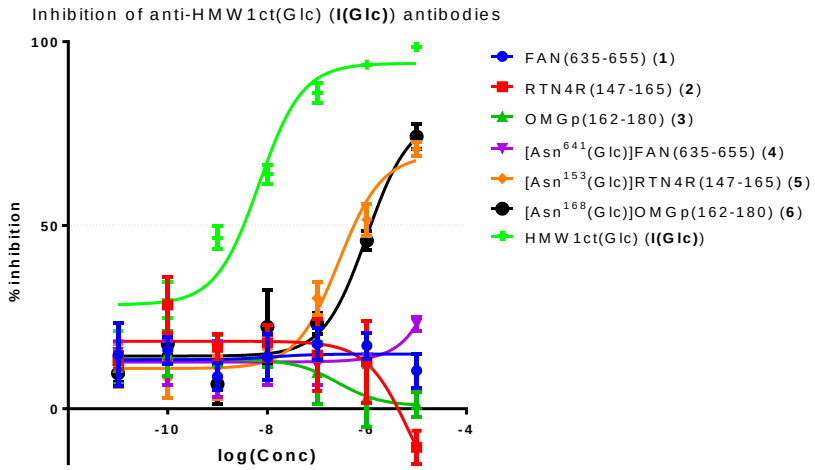


***Figure 26**. Inhibition of binding of MS14 serum IgG antibodies to HMW1ct(Glc) by peptides **1**, **2**, **3**, glucopeptides **4**, **5**, **6**, and hyperglucosylated HMW1ct(Glc) protein **I(Glc)** at different concentrations.*

Interestingly, this cross-reactivity was displayed only by the two glucopeptides **5** and **6** selected by a bioinformatic approach based on structural homology and not by the third one, i.e. [Asn^{641}(Glc)]FAN(635−655) (**4**), selected by a mere sequence homology with CSF114(Glc). Notably, CD measurements indicate that peptides **5** and **6**, as well as peptide CSF114(Glc), feature a higher percentage of β-strand content, while peptide **4** is mainly in a random coil conformation, thus confirming the possible conformational homology between these two peptides (**5** and **6**) and CSF114(Glc) and characterized by a β-hairpin structure.[73] This important observation indicates that the MS-specific epitope, present in the hyperglucosylated bacterial NTHi adhesin, is significantly recapitulated by the synthetic probe CSF114(Glc) and it is also present in the two human myelin proteins RTN4R and OMGp. Moreover, it suggests that the resemblance among the MS-specific epitopes is not limited to the Asn(Glc) moiety, but has a significant conformational component.

1.3.3 Conclusion

In conclusion, the demonstration of the presence of an epitope specifically recognized by MS patient sera and shared among CSF114(Glc), the hyperglucosylated bacterial adhesin, and the glucosylated peptides derived from two relevant human myelin proteins, i.e., RTN4R and OMGp, strongly suggests that the latter proteins may be considered as MS autoantigens. This observation adds relevant evidence in support of two essential concepts highly debated in the recent literature concerning MS, but not fully explored: the involvement of sugar moiety, i.e., N-glucosylated asparagine residues, and the contribution of infective agents (bacteria and/or viruses) to the disease aetiology. In fact, our results strongly support the immunological resemblance between human and bacterial glucosylated protein antigens, underpinning the hypothesis that molecular mimicry triggers the breakdown of self-tolerance in Multiple Sclerosis. Further experimental work is in progress to confirm this hypothesis and in particular to clarify the potential origins and the role of glucosylation of these antigens, since it is well known that this co-translational modification should be considered aberrant, when found in human proteins.

1.3.4 Experimental section

MW-assisted solid-phase peptide synthesis. Peptides were synthesized by microwave-assisted solid-phase synthe- sis (MW-SPPS) following the Fmoc/tBu strategy, using the Liberty BlueTM automated microwave peptide synthesizer (CEM Corporation, Matthews, NC, USA) following the protocol previously described.[89,90]

Circular dichroism. CD spectra of peptides **4-6** and CSF114(Glc) were recorded using quartz cells of 0.1 cm path length with a JASCO J-710 CD spectropolarimeter at 25 °C. The spectrum was measured in the 260−190 nm spectral range, 1 nm bandwidth, 8 accumulations, and 100 nm/min scanning speed. The peptides were dissolved in PBS at a concentration of 100 μM. The secondary structure content of the peptides was predicted using the online server for protein secondary structure analyses BestSel.[23]

Inhibition ELISA. Nunc-Immuno MicroWell 96 well polystyrene ELISA plates (NUNC Maxisorb, product code M9410, Merck, Milano, Italy) were coated with a solution 10 μg/mL, in pure carbonate buffer 0.05 M (pH 9.6) for peptide antigens or PBS (pH = 7.2) for protein HMW1ct(Glc), adding 100 μL/well, and incubated at 4°C overnight. Wells were washed (5×) using a an automatic Hydroflex microplate washer (Tecan Italia, Milano, Italy) using washing buffer (0.9% NaCl, 0.05% Tween 20). Nonspecific binding sites were blocked with 100 μL/well of Fetal Bovine Serum (FBS) buffer solution (10% in washing buffer) at room temperature for 1 h. Antibody apparent affinity was measured following the protocol for competitive ELISA previously reported.[91] Briefly, sera dilution was previously defined in titration curves at the semi-saturating (i.e., 50%) absorbance value of 0.7. Seven different concentrations of each synthetic peptide or protein antigenic probe were used as inhibitors. Then, sera samples at the selected dilution were incubated in parallel with increasing concentrations of the antigenic probes (range 1×10^{-11} to 1×10^{-5} M) for 1 h at room temperature. All

competitive experiments were performed in triplicate. After washes (3×), uninhibited antibodies were identified by adding 100 μL/well of alkaline phosphatase-conjugated with anti-human immunoglobulin G specific antibodies (IgG, Merck, Milano, Italy) diluted 1:3000 in FBS buffer. The microplates were then incubated 3 h at room temperature and, after washes (3×), 100 μL of substrate solution consisting of 1 mg/mL p-nitrophenyl phosphate (pNPP, Merck, Milano, Italy) and $MgCl_2$ 0.01 M in carbonate buffer (pH 9.6) were added. After approximately 30 min, the reaction was stopped with 1 M NaOH solution (50 μL/well) and the absorbance was read in a multichannel ELISA reader (Tecan Sunrise, Maennedorf, Switzerland) at 405 nm. Antibody titer values were calculated as (mean Abs of serum triplicate) - (mean Abs of blank triplicate) representing graphically the absorbance inhibition percentage. One positive and one negative serum, as references, were included in each plate for further normalization. Each experiment was performed at least twice in different days. Within-assays and between-assays coefficients of variations were below 10%. Calculated half maximal inhibitory concentration (IC_{50}) are reported for each antigenic probe. Coating conditions were set-up independently for each peptide and results are reported in the **Supplementary Material** (**Figures S2-S4**). The reference serum (MS14) was drawn from a relapsing-remitting MS patient, who had given informed consent, and was previously diagnosed after lumbar puncture, cerebrospinal fluid analysis, and MRI examination, fulfilling the established international diagnostic criteria. [92,93] This patient did not follow any specific treatment for the disease and showed high IgG antibody titers for years.

1.4 Synthetic short-chain peptide analogues of H1 relaxin lack affinity for the RXFP1 receptor and relaxin-like bioactivity. Clues to a better understanding of relaxin agonist design

The peptide hormone relaxin (RLX), also available as clinical-grade recombinant protein (serelaxin), holds great promise as a cardiovascular and anti-fibrotic agent but is limited by the pharmacokinetic issues common to all peptide drugs. In this study, by a computational modelling chemistry approach, we have synthesized and tested a set of low molecular weight peptides based on the putative receptor-binding domain of the B chain of human H1 RLX isoform, with the objective to obtain RLX analogues with improved pharmacokinetic features. Some of them were stabilized to induce the appropriate 3-D conformation by intra-chain triazolic staples, which should theoretically enhance their resistance to digestive enzymes making them suited for oral administration. Despite these favourable premises, none of these H1 peptides, either linear or stapled, revealed a sufficient affinity to the specific RLX receptor RXFP1. Moreover, none of them was endowed with any RLX-like biological effects in RXFP1-expressing THP-1 human monocytic cells and mouse NIH-3T3-derived myofibroblasts in in vitro culture, in terms of significantly relevant cAMP elevation and ERK1/2 phosphorylation, which represent two major signal transduction events downstream RXFP1 activation. This was at variance with authentic serelaxin, which induced a clear-cut, significant activation of both these classical RLX signaling pathways. Albeit negative, the results of this study offer additional information about the structural requirements that new peptide therapeutics shall possess to effectively behave as RXFP1 agonists and RLX analogues.

1.4.1 Introduction

Recombinant human H2 relaxin, later termed serelaxin, is a clinical-grade drug,[94] which was subject of several clinical trials to explore its therapeutic potential, particularly as cardiovascular and antifibrotic agent.[95-97] In humans, there are seven relaxin family peptides: human relaxin-1 (RLX-1), RLX-2, RLX-3, insulin-like peptide 3 (INSL3), INSL4, INSL5 and INSL6.[98] The relaxin family peptides produce their physiological effects by activating a group of four G protein-coupled receptors (GPCRs), relaxin family peptide receptors 1–4 (RXFP1–4).[98] Serelaxin is the recombinant equivalent of RLX-2 which acts by activating RXFP1; however, its spread has been hampered by the pharmacokinetic issues common to all peptide drugs. In fact, its short half-life (~2 h) upon i.v. administration,[99] forced it to be given parenterally in multiple daily doses, with obvious drawbacks in terms of ease of use, patients' compliance and suitability for long-term treatments.[95] To try to overcome these problems, over the years analogues of the hormone have been designed that maintain its ability to bind to and activate RXFP1[100] with the advantages of improved resistance to proteolytic catabolism, extended half-life and suitability for oral delivery.[101] RLX-2 hormone is composed of 2 peptide chains, A and B; the A chain has a stabilizing function essential for receptor affinity, while the B chain contains the receptor-binding domain, made up by an arginine cassette including Arg13, Arg17 and Ile20, located on the same face of an α-helix, typical of the hormone bioactive conformation.[98] Recombinant H1 RLX was found to have comparable RXFP1 affinity as H2 RLX,[102] as well as similar chronotropic and inotropic effects in the isolated rat heart assay. Moreover, H1 RLX showed a greater alpha-helical conformational preference in water than H2 RLX and retained cardiotropic effects even upon modification of the C-terminus of the B-chain.[103] On these grounds, we reasoned that low molecular weight, single-

chain synthetic analogues of H1 RLX would deserve to be investigated as possible RXFP1 agonists.

1.4.2 Results and Discussion

Design. Based on the putative receptor-binding domain of the H1 RLX B chain, six low molecular weight peptides were designed and synthesized. The starting point was the C-terminal truncated fragment of H1 RLX B chain H1RLX(1–23) (peptide **1** in **Table 9**), which is expected to adopt an α-helical conformation in the native hormone, on the basis of the structural data available for H2 RLX.[104] The Cys residues in positions 10 and 22 were replaced by Ser, to avoid unwanted formation of disulphide bridge and the α-helix was installed inducing i to i + 4 triazole bridge modifying, one at a time, three couples of positions not directly involved in receptor binding, i.e., 10-14, 14–18, and 17–21, respectively. Accordingly, as shown in **Table 9**, the conformational constraint in position 10–14 was obtained replacing Ser10 by Pra and Leu14 by Lys (N3), or, vice versa, Ser10 by Lys (N3) and Leu14 by Pra, obtaining, after CuAAC reaction, the stapled cyclo-peptides **2** and **3**, respectively, characterized by opposite orientation of the triazole moiety, as shown in **Table 9**. Similarly, peptides **4** and **5** bear the triazole bridge in position 14–18 and peptides **6** and **7** in position 17–21, respectively.

Table 9. H1 RLX peptide analogues.

No.	Name	Structure *
1	[Ser^{10,22}]H1 RLX B(1-23)	K-W-K-D-D-V-I-K-L-S-G-R-E-L-V-R-A-Q-I-A-I-S-G
2	[Ser^{22}] Cyclo [Pra^{10}, LysN$_3$^{14}]H1RLXB(1-23)	K-W-K-D-D-V-I-K-L - **X** - G-R-E - **Z** - V-R-A-Q-I-A-I-S-G
3	[Ser^{22}] Cyclo [LysN$_3$^{10}, Pra ^{14}]H1RLXB(1-23)	K-W-K-D-D-V-I-K-L - **Z** - G-R-E - **X** - V-R-A-Q-I-A-I-S-G
4	[Ser^{10,22}] Cyclo [Pra^{14}, LysN$_3$^{18}]H1RLXB(1-23)	K-W-K-D-D-V-I-K-L-S-G-R-E - **X** - V-R-A - **Z** - I-A-I-S-G
5	[Ser^{10,22}] Cyclo [LysN$_3$^{14}, Pra^{18}] H1RLXB(1-23)	K-W-K-D-D-V-I-K-L-S-G-R-E - **Z** - V-R-A - **X** - I-A-I-S-G
6	[Ser^{10,22}] Cyclo [Pra^{17}, LysN$_3$^{21}]H1RLXB(1-23)	K-W-K-D-D-V-I-K-L-S-G-R-E-L-V-R - **X** - Q-I-A - **Z** - S-G
7	[Ser^{10,22}] Cyclo [LysN$_3$^{17}, Pra^{21}]H1RLXB(1-23)	K-W-K-D-D-V-I-K-L-S-G-R-E-L-V-R - **Z** - Q-I-A - **X** - S-G

* **X**: Pra; **Z**: LysN$_3$

Circular Dichronism. The secondary structure propensity of the constrained H1 RLX B-chain analogues was explored by circular dichroism (CD) spectroscopy, performed both in phosphate buffer and SDS micelles (**Figure 27**), in comparison with the linear unmodified peptide [Ser^{10}, Ser^{22}] H1 RLX B (1–23) **1**. In general, CD spectra of the stapled peptides in phosphate solution (**Figure 27A**) showed a tendency to assume α-helical secondary structure, as compared to the linear reference Peptide **1** (green). Interestingly, the 3 peptides of the R series, characterized by a C-terminal oriented triazole stapling (**2**: pale green, **4**: blue, and **6**: light blue) are slightly more helical than the corresponding analogues, characterized by N-terminal oriented triazole (**3**: pale blue, **5**: brown, and **7**: light blue), as indicated by the more intense minimum at 222 nm. We subsequently performed CD spectroscopy in SDS micelles to explore the amphipathic properties of the analogues (**Figure 27B**), since the main aim of our design was to expose the key residues of H1 RLX B chain receptor binding cassette on the same face of a

helical structure. Interestingly, all peptides showed a pronounced helicity in SDS micelles, including the linear reference **1** (green), although slightly less than the stapled ones. Among the stapled peptides, **2** (pale green) and **3** (pale blue) showed the highest degree of helicity (**Figure 27B**).

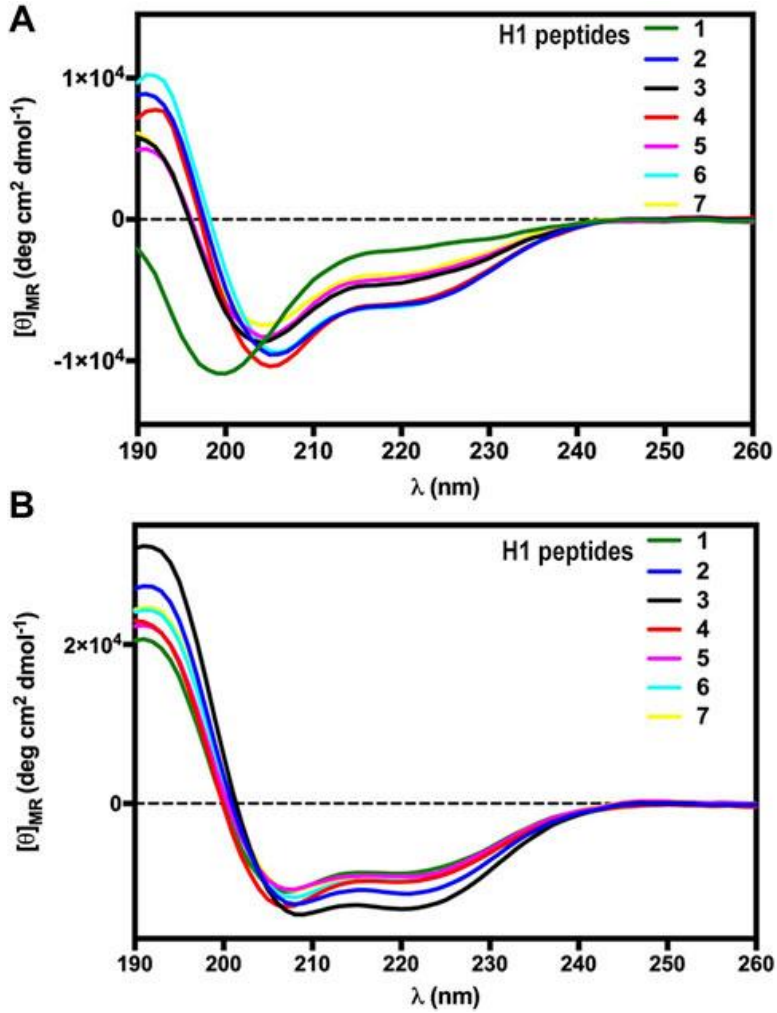


Figure 27. *CD spectra of the H1 RLX stapled analogues (**2–7**) and the linear reference peptide **1** in phosphate buffer (**A**) and in SDS micelles (**B**).*

Receptor binding assay and biological activity. Despite the favourable premises, none of the tested H1 peptides revealed a substantial affinity to RXFP1 (**Figure 28**) nor displayed any RLX-like biological effects, in terms

of significantly relevant cAMP elevation and ERK1/2 phosphorylation in RXFP1-expressing cells (**Figure 29**). This was at variance with authentic serelaxin, which induced a clear-cut, significant activation of both the classical signaling pathways downstream RXFP1. Moreover, B7-33, the known RXFP1 agonist designed on the receptor-binding domain of H2 RLX,[105] was also able to induce a significant, although less prominent, cAMP response.

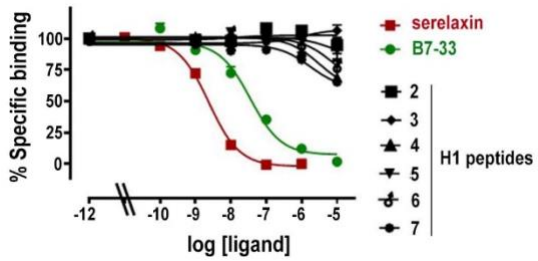


Figure 28. *Competition binding curves for serelaxin, B7-33 and H1 peptides in HEK-293T stably expressing RXFP1-BP. Data are mean ± SEM from 2 independent experiments. The competition binding data were fitted using a one site binding model.*

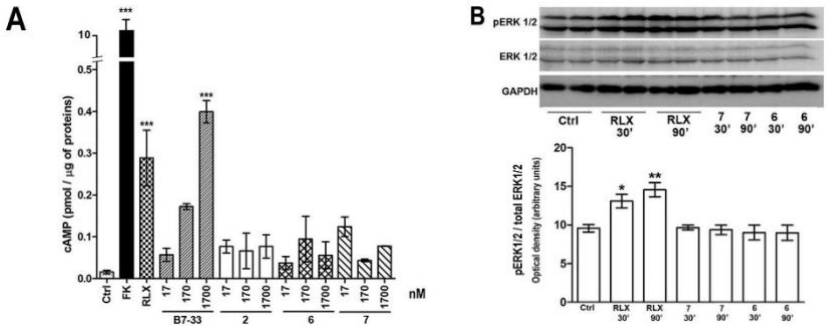


Figure 29*. Assessment of RXFP1 signaling pathway activation. (A) cAMP generation in human monocytic THP-1 cells: both serelaxin (RLX) and peptide B7-33 induce a statistically significant elevation of cAMP, while none of the tested H1 peptides yielded a significant elevation of cAMP at any tested concentration. The adenylate cyclase activator forskolin (FK) was used to assess maximal cAMP yield. (B) ERK1/2 phosphorylation in mouse myofibroblasts: serelaxin induced a statistically significant increase in pERK1/2 at both 30 and 90 min, while none of the tested H1 peptides did. Values are mean ± SEM of 3 independent experiments. Significance of differences: ∗ p < 0.05, ∗∗p < 0.01, ∗∗∗p < 0.001.*

1.4.3 Conclusion

The observation that this single-chain, linear analogue of H2 RLX maintains some efficacy in activating RXFP1, while the synthesized H1 RLX single-chain peptides are not effective, suggests that differences between H1 and H2 RLX B-chain sequences, particularly in the N-terminal portion distal to the common receptor binding cassette, play a crucial role in receptor activation. In addition, we hypothesize that the triazole conformational constraint, devised to stabilize the bioactive conformation, plays on the contrary a negative effect, which does not favor the correct ligand-receptor interaction. Albeit negative, the results of this study offer additional information about the structural requirements that new peptide therapeutics shall possess to effectively behave as RXFP1 agonists and RLX analogues. A more detailed conformational analysis, based on nuclear magnetic resonance (NMR) and molecular modelling, will enable the refinement of the design of the constrained analogues, changing the position and/or the extension of the triazole bridge, potentially leading to a better replica of the native bioactive conformation.

1.4.4 Experimental section

Peptides. Serelaxin (batch B917056/1/1, prepared by Boehringer-Ingelheim Inc.) was kindly donated by the Relaxin RRCA Foundation, Florence, Italy. Short-chain peptide analogues of H1 RLX (**Table 9**) and the known RXFP1 agonist B7-33[104] were synthesized as previously described.[86,106]

CD spectroscopy. CD spectra were recorded using quartz cells of 0.1 cm path length with a JASCO J-710 CD spectropolarimeter at 25 °C. All spectra were measured in the 260-190 nm spectral range, 1 nm bandwidth, 4 accumulations and 100 nm/min scanning speed. Each peptide was dissolved in water to prepare a 1 mM peptide stock solution. The CD spectra were performed in PBS and in SDS (20 mM) using a peptide concentration of 100 μM.

Cell culture for ligand-receptor binding assays. Human embryonic kidney cells (HEK-293T) stably expressing RXFP1- BP)[107] were used for competition binding experiments using europium-labelled H2 relaxin, as described.[108] This construct contains the ectodomain of human RXFP1 fused to the single transmembrane and cytoplasmic region of CD8. This ectodomain-only construct has a higher affinity for single chain H2 relaxin peptides and is used as a surrogate assay for binding affinity. Assays were performed in whole cell 96 well plates assays with a single concentration of the europium-labeled H2 relaxin (1 nm) in the presence or absence of increasing concentrations of the competing peptides. Non-specific binding was assessed using 1 μM of serelaxin. Each concentration point was assessed in triplicate in 2 independent experiments. The binding data were analyzed using GraphPad Prism 9 and expressed as mean ± SEM. They were fitted using a one site binding model.

Cell culture for cAMP assay. The RXFP1 downstream cAMP pathway was tested on human monocytic THP-1 cells (ECACC, Salisbury, UK) cultured in suspension in RPMI medium containing 10% foetal bovine serum,

0.05mm 2- mercaptoethanol, 250 U/ml penicillin G and 250 μg/ml streptomycin, in a 5% CO_2 atmosphere at 37°C. Cells (5×10^5) were placed in a 24-well plate, added with IBMX (100 μM) to prevent cAMP catabolism and then incubated with the different stimulants. As positive controls, either serelaxin (17 nm) and B7- 33 (17, 170, 1700 nm) were used. The H1 peptides **4**, **6** and **7**, selected among those which had shown a minimal ability to bind RXFP1 in the competition binding experiments, were also used at 17, 170 and 1700 nm final concentration. The cAMP assay was carried out 15min after RXFP1 stimulation, coinciding with the second sustained cAMP surge.[109,110] The adenylate cyclase activator forskolin (100 μM) was used to determine the maximal cAMP levels. Triplicate cAMP measurements were performed by a Direct cAMP ELISA Kit (Enzo, Milan, Italy). Results were calculated using a 4 parameter logistic (4PL) curve fitting program, as suggested by the manufacturer. The values of cAMP were normalized by the amount of proteins, measured by the micro-BCA Protein Assay Kit (Pierce, IL, United States) and expressed as mg/ml. For statistical purposes, 3 independent experiments were performed.

Cell culture for ERK1/2 phosphorylation assay. The RXFP1 downstream ERK1/2 phosphorylation pathway was tested on myofibroblasts using a method based on that described in literature,[104] with minor modifications. Mouse NIH-3T3 fibroblasts (ATCC, Manassas, VA, United States) expressing RXFP1[97] were cultured in DMEM supplemented with 10% fetal bovine serum (FBS), 2 mm glutamine, 250 U/ml penicillin G and 250 μg/ml streptomycin, in a humidified atmosphere with 5% CO_2 at 37°C. They were induced to myofibroblasts in DMEM containing 2% FBS and 2 ng/ml human TGF-β1 (PeproTech, Rocky Hill, NJ, United States) for 12, 24, 48 and 72 h. [97] Myofibroblasts were then incubated for 30 or 90 min with either serelaxin (17 nM), or B7-33 (1.7 μM), or the H1 peptides 6 and 7 (1.7 μM), selected among those which had shown a minimal ability to bind RXFP1 in the

competition binding experiments. Untreated myofibroblasts were used as controls. To verify fibroblast-to-myofibroblast transition and ERK1/ 2 phosphorylation, cells were lysed in cold buffer composed of (mM): 10 Tris/HCl pH 7.4, 10 NaCl, 1.5 $MgCl_2$, 2 Na_2EDTA, 1% Triton X-100, added with 10x Sigmafast Protease Inhibitor Cocktail tablets and 100 mM Na_3VO_4 to inhibit endogenous phosphatases. Total protein content was measured spectrophotometrically using micro-BCA Protein Assay Kit (Pierce). Sixty μg of total proteins were electrophoresed by SDS–PAGE and blotted onto PVDF membranes (Millipore, Bedford, MA, United States). The membranes were incubated overnight at 4°C with: goat polyclonal anti-α-smooth muscle actin (α-sma, 1:1,500, AbCam, Cambridge, UK); rabbit polyclonal anti-pERK1/2 (1:1,000, Cell Signaling, Milan, Italy), rabbit polyclonal anti-ERK1/2 (1: 1,000, Cell Signaling), mouse monoclonal anti-GAPDH (1:2000; Invitrogen, Waltham, MA, United States), assuming GAPDH as control invariant protein. Specific bands were detected using appropriate peroxidase-labeled secondary antibodies (1:15,000; Vector, Burlingame, CA, United States) and enhanced chemiluminescent substrate (ECL, Sigma-Aldrich). Densitometric analysis of the bands was performed using Scion Image Beta 4.0.2 software (Scion Corp., Frederick, MD, United States) and the values normalized to GAPDH. For statistical purposes, 3 independent experiments were performed.

Chapter2 - LIGAND-DNA INTERACTION BY NMR SPECTROSCOPY

2.1 Synthesis and Characterization of Bis-Triazolyl-Pyridine Derivatives as Noncanonical DNA-Interacting Compounds

Besides the well-known double-helical conformation, DNA is capable of folding into various noncanonical arrangements, such as G-quadruplexes (G4s) and i-motifs (iMs), whose occurrence in gene promoters, replication origins, and telomeres highlights the breadth of biological processes that they might regulate. Particularly, previous studies have reported that G4 and iM structures may play different roles in controlling gene transcription. Anyway, molecular tools able to simultaneously stabilize/destabilize those structures are still needed to shed light on what happens at the biological level. A set of 31 bis-triazolyl-pyridine derivatives was generated by combining a multicomponent reaction and a click chemistry functionalization. The newly synthesised compounds were screened by circular dichroism for their ability to interact with different G4 and/or iM DNAs and to affect the thermal stability of these structures. All the compounds were then clustered through multivariate data analysis, based on such capability. The most promising compounds were subjected to a further biophysical and biological characterization, leading to the identification of two molecules simultaneously able to stabilize G4s and destabilize iMs, both in vitro and in living cells.

2.1.1 Introduction

The canonical double helix is the most widely recognized genomic DNA structure. However, DNA is structurally dynamic and able to adopt a number of alternative secondary structures, such as cruciforms, triplexes, G-quadruplexes, and i-motifs.[111-113] G-quadruplex (G4) structures, four-stranded helical complexes that can arise from guanine-rich sequences, are among the most extensively studied noncanonical DNA secondary structures.[114] The core scaffold of a G4 is the G-tetrad (**Figure 30**), a cyclic array of four guanines connected by eight Hoogsteen hydrogen bonds. Two or more G-tetrads can self-associate into vertical stacks giving rise to G4s, which are extremely stable structures, although the stability of a G4 structure depends on many factors, including the length and sequence composition of the G-rich motif, the nature of the binding cations, and the folding topology.[115-118] Indeed, G4s can adopt a wide variety of topologies, depending on the combinations of strand orientation and length and composition of the loops (the intervening sequences between the G-rich repeats).[119,120] DNA sequences that can fold into G4 structures are broadly distributed across the human genome and mainly located at telomeres, gene promoters, and introns.[121,122] Due to such a high genomic prevalence, G4s have been speculated to play a role in several biological processes. Indeed, they may offer telomere protection and block replication fork progression,[114] while in gene promoters, they may represent a physical obstacle in the way of the RNA polymerase machinery thus affecting the expression of the downstream gene.[123] The strand complementary to a G4-forming DNA motif is a C-rich sequence that can fold in another four-stranded structure called the i-motif (iM or i-DNA).[124] This is composed of two head-to-tail, intercalated, parallel-stranded duplexes held together by hemi-protonated cytosine-cytosine ($C\text{-}C^{+}$) base pairs (**Figure 30**).

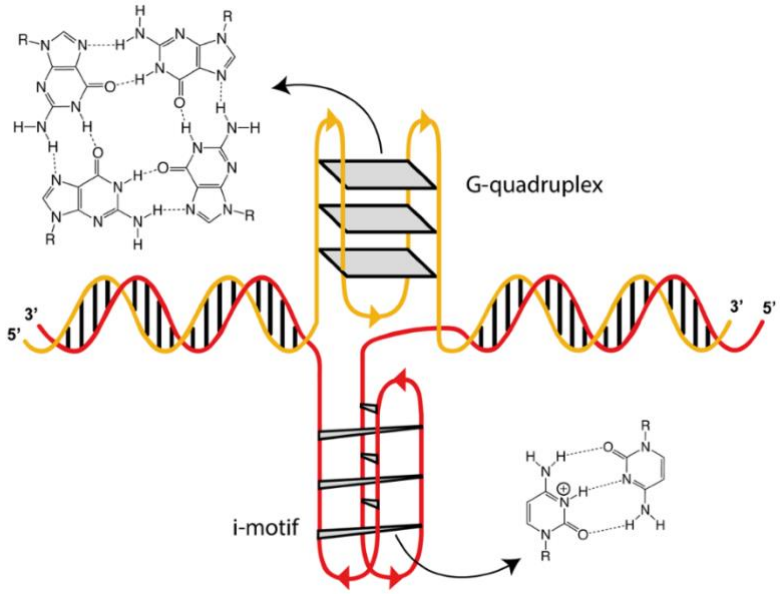


Figure 30. *Noncanonical DNA secondary structures: G-quadruplex (G4) and i-motif (iM).* ***Top left****: a G-tetrad;* ***bottom right****: a hemi-protonated cytosine-cytosine (C-C$^+$) base pair.*

Very recently, evidence for iM formation in cells has been provided by NMR experiments[125] and by the discovery of an antibody that specifically binds iM structures in the nuclei of human cells.[126] Analogously to G4s, the presence of iM structures may affect gene transcription and telomere biology.[112] Interestingly, the formation of iM structures is cell-cycle dependent, peaking at the late G1 phase,[126] whereas G4 formation is maximal during the S phase.[127] This suggests that iMs and G4s might play different roles in regulating gene expression and transcription. Many studies, focused on disentangling the biological roles of G4 and iM structures, demonstrate that such structures can be considered as molecular targets in cancer therapies,[124,128,129] leading to significant therapeutic advantages in the treatment of cancer. For example, direct evidence of the role of DNA G4s in gene regulation has been provided, showing that oncogene promoter activity is generally repressed using G4s stabilizers, while destabilization of G4s generally leads to increased gene transcription.[130] On the other hand, constant

efforts are still underway to fully decipher the specific biological functions of iM structures. In this contest, the scientific community has also made big efforts to find small molecules able to recognize G4 and iM structures. More than a thousand compounds that bind and stabilize G4 structures have been found so far. Many of them display considerable selectivity for G4s over single-stranded and double-stranded DNA, and some of them are also able to discriminate among distinct G4 folding topologies. Compared to the well-documented examples of G4 ligands, the discovery of specific iM-targeting compounds is much more limited.[131,132] The design of small molecules able to interact with G4s is usually based on the following requisites: (i) the presence of a (hetero)-aromatic system that gives π-stacking interactions with planar G-tetrads and flanking bases, (ii) a V-shaped form to maximize the interaction with the G4, (iii) two or three cationic side chains able to participate in ionic interactions with the phosphate backbone of loops and grooves of G4 structures. On the other hand, loop binding and recognition of hemi-protonated C-C^+ base pairing seem to be the most favoured binding modes of iM ligands.[133,134] Therefore, inspired by pyridostatin, a potent G4-binder based on an *N*,*N*′-bis(quinolinyl)pyridine-2,6-dicarboxamide scaffold[135-137] which has also been shown to interact with iM DNA (although without significant effects on the thermal stability of such structure),[138] a library of bis-triazolyl-pyridine derivatives has been generated. Then, the ability of such compounds of interacting with G4 and iM structures has been investigated both from a biophysical and biological point of view. The synthesized bis-triazolyl-pyridines were preliminarily screened for their ability to interact with G4 and/or iM DNAs through circular dichroism (CD) spectroscopy and clustered based on their capability to affect the thermal stability of the investigated DNA sequences by employing the Principal Component Analysis (PCA) technique.–The ability of the best-selected compounds to interact with noncanonical DNAs was further investigated by

using NMR and fluorescence spectroscopies. Finally, the proficiency of such compounds to modulate G4 and iM structures formation in living cells was assessed.

2.1.2 Results and Discussion

The library of 31 bis-triazolylpyridines consist of: series a **15a**–**24a**, series b **15b**–**21b**, series c **15c**–**21c**, series d **15d**–**21d**, containing one, two, three or four methylene groups, respectively The synthetic route followed for the preparation of the putative noncanonical DNA-interacting compounds is depicted in **Figure 31**, the synthesis details are reported in the published article.[139]

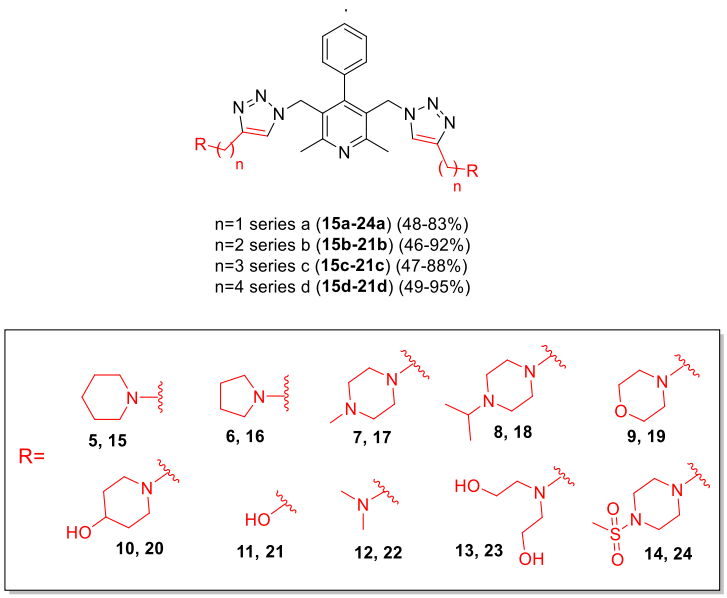


Figure 31. *The library of 31 bis-triazolylpyridines iazolylpyridines*

Circular Dichroism Screening. The 31 synthesized bis-triazolyl-pyridines were preliminarily screened for their ability to interact with G4 and iM DNAs through circular dichroism (CD) spectroscopy.[140] Various G-rich sequences that can form parallel, antiparallel, and hybrid G4 structures and their iM-

forming C-rich counterparts were used in these experiments. In particular, we employed the G-/C-rich motifs from the promoter regions of BCL-2 (*Bcl-2 G4* and *Bcl-2 iM*)[141,142] and c-MYC (*c-Myc G4* and *c-Myc iM*) oncogenes[141,143], as well as the G-rich 23-mer (*Tel*$_{23}$ *G4)* and the C-rich 22-mer (*hTeloC iM*) truncations of the human telomeric DNA sequence.[144-146] Moreover, a 27-mer hairpin-duplex-forming sequence (*Hairpin*) was also used to evaluate the selectivity of the compounds for noncanonical DNAs over duplex DNA. The DNA stabilizing/destabilizing properties of the 31 bis-triazolyl-pyridineswere evaluated by CD-melting experiments measuring the compound-induced change in the apparent melting temperature ($\Delta T_{1/2}$) either of G4, iM, or duplex structures. Considering a large number of $\Delta T_{1/2}$ values to be compared (31 compounds tested on 7 DNA sequences give a total of 217 values of $\Delta T_{1/2}$), a multivariate data analysis tool, such as Principal Component Analysis (PCA)[147] , was employed to explore the data set. In this way, similarities and differences in the behavior of the investigated compounds could be detected and correlated to their structural features. *Hairpin* and *Tel*$_{23}$ *G4* were excluded from the analysis since their thermal stability was not significantly affected by any of the tested compounds ($|\Delta T_{1/2}| < 4$ °C). Therefore, the PCA was performed including two G4-forming sequences (*Bcl-2 G4* and *c-Myc G4*) and three iM-forming sequences (*hTeloC iM*, *c-Myc iM*, and *Bcl-2 iM*) for a total of 5 DNA sequences. After PCA, we the compounds able to affect the thermal stability of both iM and G4 structures (**15c** and **18c**) or exclusively of the iMs (**20a** and **23a**) were selected. In addition, **21d** was also employed to have a comparison with a compound unable to affect the stability of any of the investigated DNAs.

Nuclear Magnetic Resonance (NMR) Studies. The spectral regions of the imino and aromatic protons of *c-Myc G4* and *hTeloC iM* in the absence and presence of compounds (2 molar equiv) are shown in **Figure 32** and **Figure 33**, respectively. According to the literature, the *c-Myc G4* sequence forms, under the experimental conditions used, a single G4 conformation characterized by 12 well-resolved imino proton peaks, corresponding to the 12 guanines involved in the three G-tetrad planes.[140] On the other hand, *hTeloC iM* folds in an iM structure characterized by 3 well-resolved imino proton peaks that correspond to the 6 intercalated C-C$^+$pairs.[146]

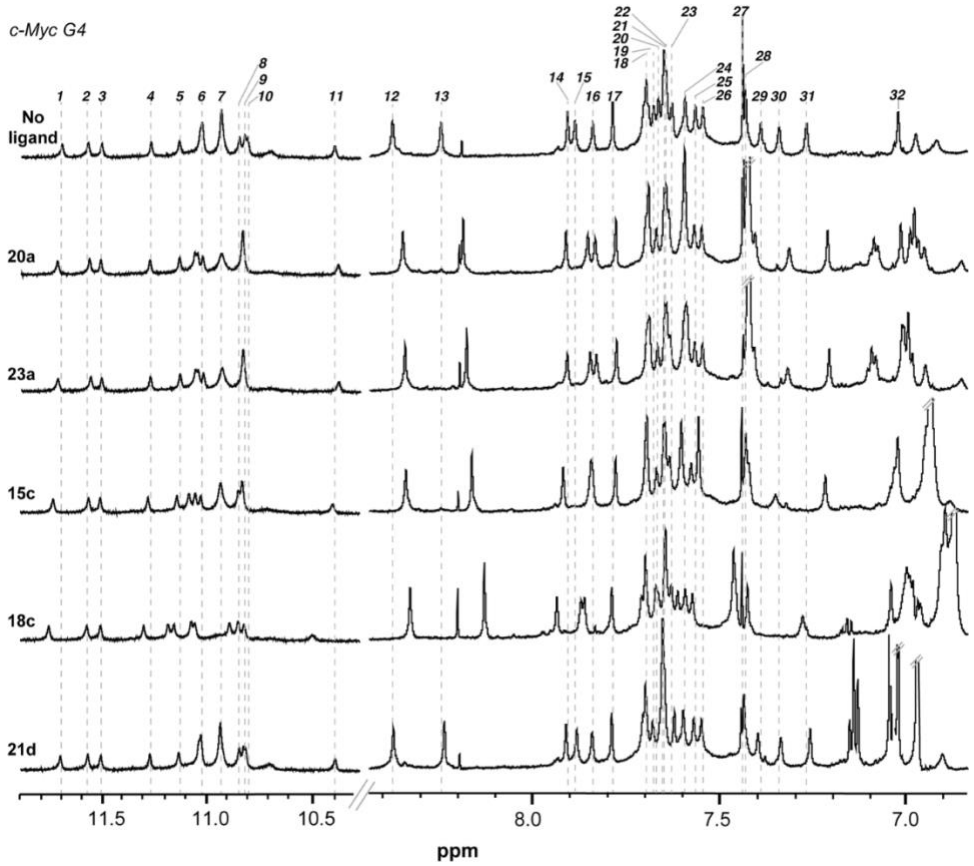


Figure 32. *^{1}H-NMR spectra of c-Myc G4 in the absence and presence of compounds **20a**, **23a**, **15c**, **18c,** and **21d**.*

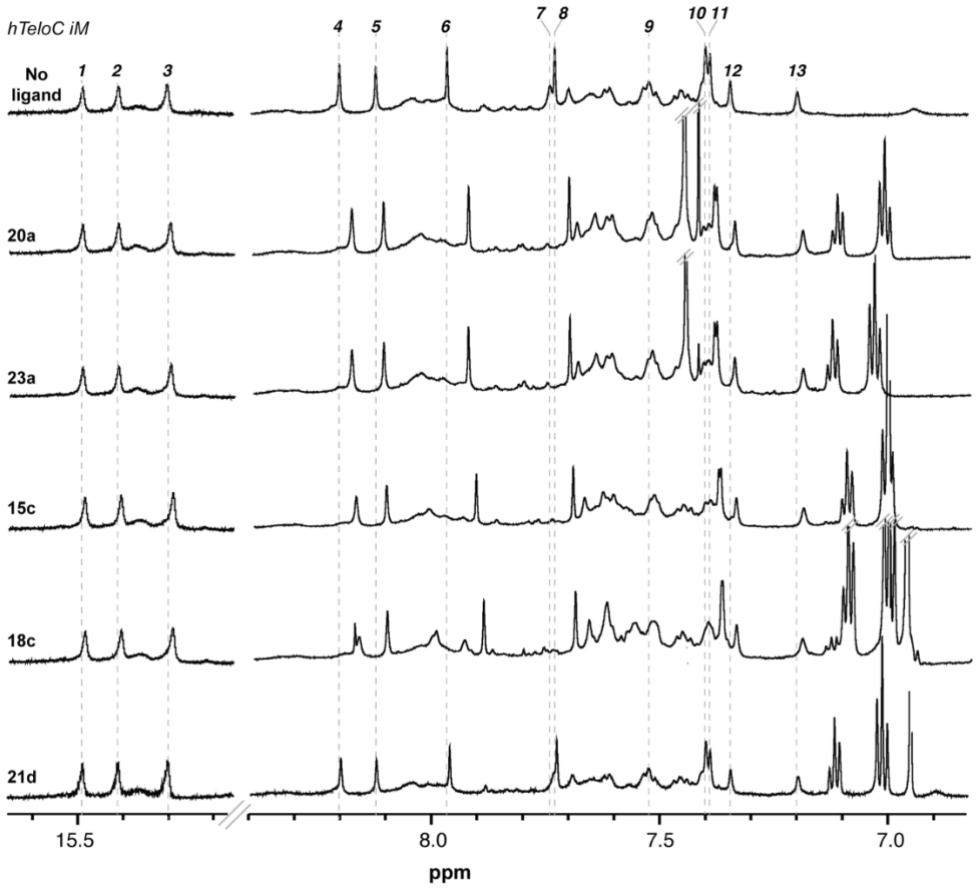


Figure 33. *^{1}H-NMR spectra of hTeloC iM in the absence and presence of compounds **20a**, **23a**, **15c**, **18c**, and **21d**.*

No relevant shift of both the imino and aromatic signals of *c-Myc G4* and *hTeloC iM* was observed upon the addition of **21d** (**Figure 32 and 33**). Conversely, **15c**, **18c**, **20a** and **23a** mainly affected the imino protons of *c-Myc G4*, as well as the aromatic protons of both *c-Myc G4* and *hTeloC iM* (**Figure 32 and 33**). As for *c-Myc G4*, not negligible chemical shift changes were observed for the imino (H1) protons of G16 belonging to the 5′ G-tetrad (G7−G11−G16−G20), as well as for those of guanines involved in the 3′ G-tetrad (G9−G13−G18−G22). At the same time, the most perturbed aromatic protons turned out to be those of A15, located in the double nucleotide loop (T14A15) connecting G13 and G16, and of A25 which stacks over the 3′ G-

tetrad, specifically, with residues G9 and G22. Overall, these results suggest a common end-stacking binding mode for **15c**, **18c**, **20a**, and **23a** to the 3′ G-tetrad of *c-Myc G4*, which is also combined with loop binding. Conversely, the imino proton signals of *hTeloC iM* were not affected by compounds **15c**, **18c**, **20a**, and **23a**. However, significant chemical shift changes of aromatic protons were observed especially for those residues located in the loops of *hTeloC iM*, and particularly for A5, A6, A11, A12, A17, and A18. These observations suggest interaction with iM loops, which is one of the most common mechanisms of interaction proposed for iM-targeting compounds.[134]

Saturation-Transfer Difference (STD) NMR Analysis. In order to confirm the binding of compounds **15c**, **18c**, **20a** and **23a** to *c-Myc G4* and *hTeloC iM*, a ligand-based NMR approach was used. Particularly, Saturation Transfer Difference NMR (STD NMR) experiments were performed in buffer solutions containing a large excess of compounds (250 μM) with respect to the DNA (20 μM).[5] STD NMR is based on the magnetization transfer by the Nuclear Overhauser Effect (NOE) from the macromolecule to the ligand. STD NMR spectra are the result of the subtraction of a 1D ^{1}H spectrum where the macromolecule protons are selectively saturated (on-resonance spectrum) from a reference spectrum in which the macromolecule is not saturated (off-resonance spectrum). The resulting difference STD spectrum will only show the ligand signals interacting with the macromolecule. The STD spectra of **15c**, **18c**, **20a** and **23a** are shown in **Figures 33-36.** The proton resonances for **15c**, **18c**, **20a** and **23a,** observed in presence of both *c-Myc G4* and *hTeloC iM*, demonstrate once again an interaction between the investigated compounds and the DNAs. In the STD spectra, the presence of both aromatic and aliphatic signals for **15c**, **18c**, **20a** and **23a** suggests that the aromatic moieties of these compounds could

interact with the G4 and iM through π-π stacking with the guanines of G-tetrads and the adenines of loops, respectively, while the side chains could interact with the backbone of the DNA molecules through electrostatic interactions. The very low signal intensity of the protons of **21d** in the STD spectra (**Figure S5, Supplementary Materials**), after normalization against the 1D ^{1}H NMR spectra, shows that this compound does not interact with the DNA sequences under investigation.

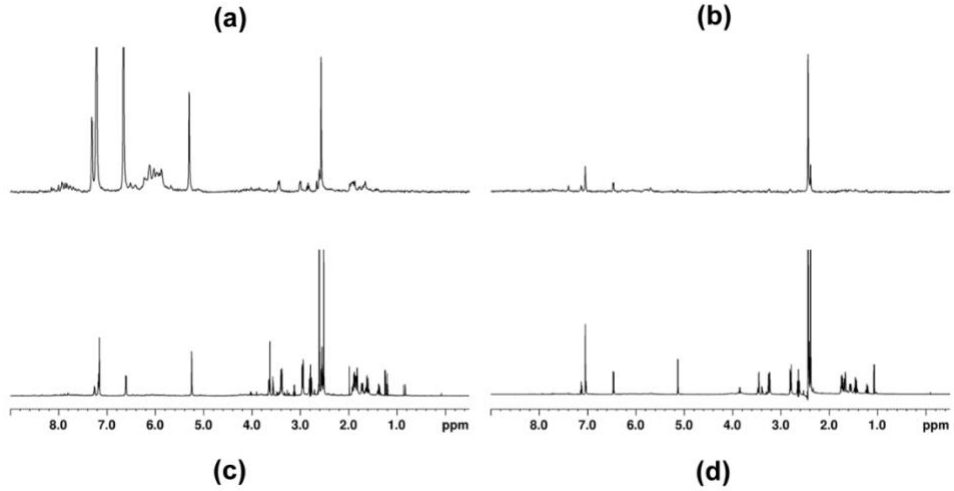


***Figure 33**. 1D ^{1}H NMR spectra (**c** and **d**) and STD spectra (**a** and **b**) of compound **15c** (0.250 mM) in presence of c-Myc G4 (**a** and **c**) and hTeloC iM (**b** and **d**).*

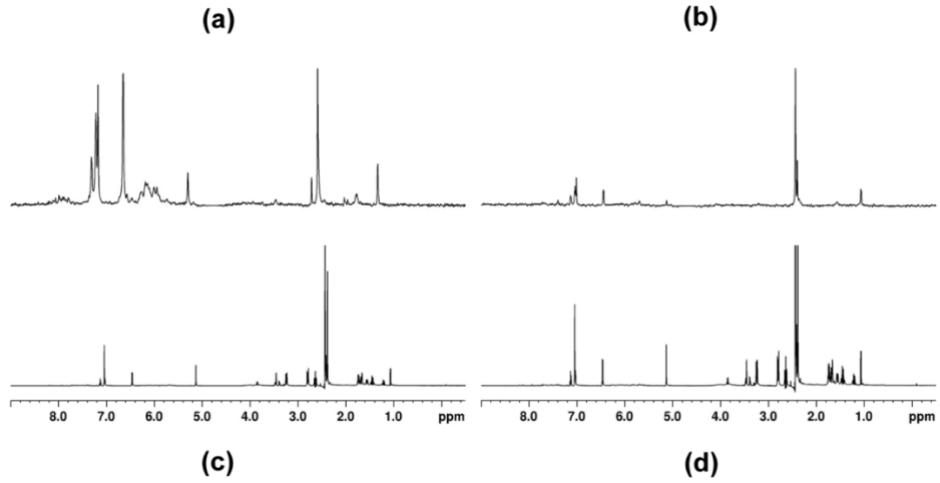


***Figure 34**. 1D ^{1}H NMR spectra (**c** and **d**) and STD spectra (**a** and **b**) of compound **18c** (0.250 mM) in presence of c-Myc G4 (**a** and **c**) and hTeloC iM (**b** and **d**).*

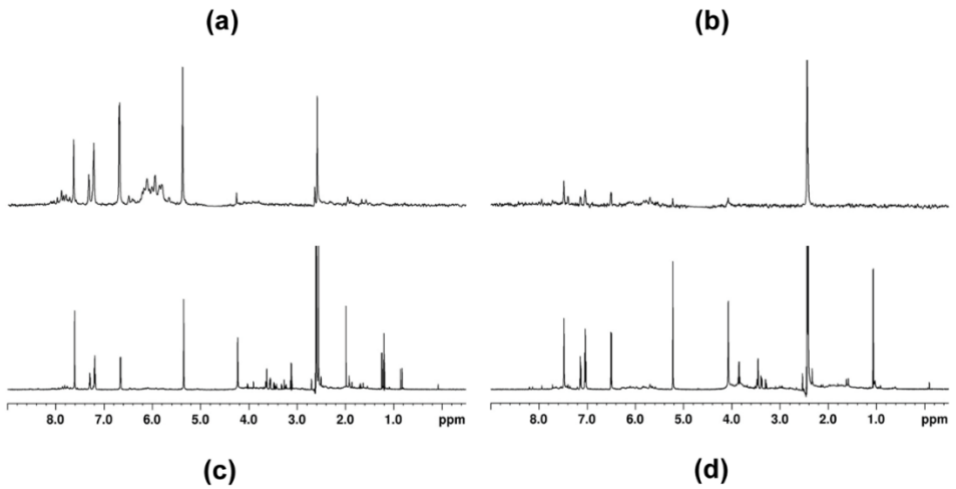


***Figure 35**. 1D ^{1}H NMR spectra (**c** and **d**) and STD spectra (**a** and **b**) of compound **20a** (0.250 mM) in presence of c-Myc G4 (**a** and **c**) and hTeloC iM (**b** and **d**).*

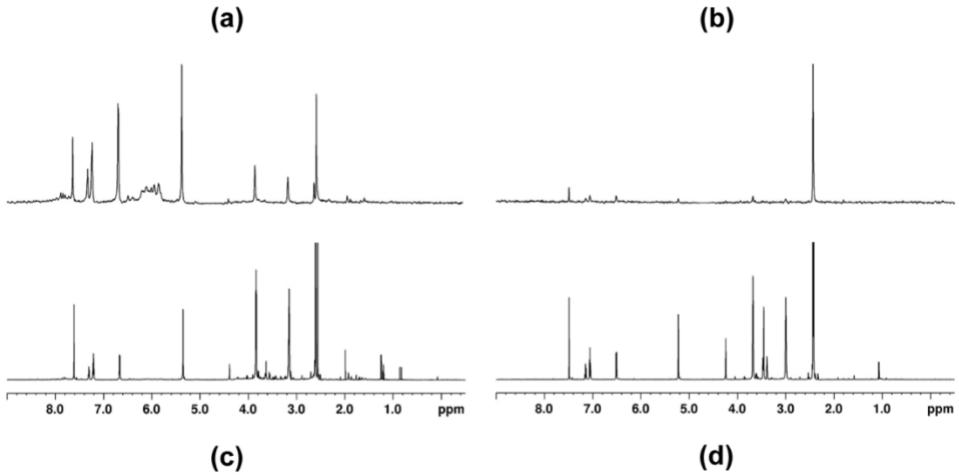


***Figure 36**. 1D ^{1}H NMR spectra (**c** and **d**) and STD spectra (**a** and **b**) of compound **23a** (0.250 mM) in presence of c-Myc G4 (**a** and **c**) and hTeloC iM (**b** and **d**).*

Evaluation of Ligand Affinity by Fluorescence Experiments. To obtain quantitative data regarding the affinity of **15c**, 1**8c**, **20a** and **23a** for *c-Myc G4*, *c-Myc iM*, and *hTeloC iM*, fluorescence titration experiments were performed. Fluorescence emission spectra of the compounds in the absence and presence of increasing amounts of oligonucleotides were recorded. Binding constants (K_b) are reported in **Table 10**. The results of the interpolation analysis indicated that in those cases the data could not be fitted with a single binding site, but they were well fitted with two binding sites per DNA molecule. Compound **18c**, which displayed the strongest affinity for the G4 structure of *c-Myc*, also exhibited the least affinity for the iM structure of both *c-Myc* and *hTeloC*. On the other hand, compounds **15c**, **20a** and **23a** showed higher affinity values for *c-Myc iM* and *hTeloC iM* rather than to *c-Myc G4*.

Table 10. *Binding constant (K_b) values obtained by fitting the fluorescence titration curves to an independent and equivalent binding sites model as previously described* [145].

Compound	K_b ($\times 10^6$ M^{-1}) [1]		
	c-Myc G4	***c-Myc iM***	***hTeloC iM***
20a	2.3	3.1	3.1
23a	2.2	3.2	3.1
15c	2.5	3.4	3.2
18c	3.1	2.8	2.6

[1] The error on the K_b values is ± 0.2.

Furthermore, fluorescence titration experiments were carried out to investigate the interaction of **21d** with a G4 (*c-Myc G4*) and an iM (*c-Myc iM*) structure (**Figure S6a,b-Supplementary Materials**). The obtained curves were then fitted giving K_b values of 1.1 (±0.2) × 10^6 and 1.6 (±0.2) × 10^6 M^{-1} for *c-Myc G4* and *c-Myc iM*, respectively, and stoichiometry of 2:1 in both cases. These results confirm the lower (although not null) affinity of this compound for these targets. Finally, the same experiment was performed for **18c** with a C-rich single-stranded oligonucleotide [d(CT)$_{15}$] (**Figure S6c-Supplementary Materials**), giving a 1:1 stoichiometry and a K_b of 4.6 (±0.4)

$\times 10^6$ M^{-1}, and thus suggesting that the decrease in melting temperature of the iMs could also be due to preferred binding to the unfolded DNA, which would shift the folding equilibrium.

Effect of the Compounds on G4/iM Structures in Living Cells. Starting from the results of the biophysical assays, the selected molecules (**15c, 18c, 20a, 23a** and **21d** as a control) were tested for their capability to destabilize the iM and/or stabilize the G4 structures also in the cells. To this aim, U2OS human osteosarcoma cells, one of the models already used by Zeraati and colleagues to validate their anti-iM antibody (iMab),[126] were treated for 24 h with 2 μM of each compound and the iM structures were evaluated by immunofluorescence (IF) microscopy. As a positive control, cells were maintained in 8% of CO_2 for 2.5 h, a condition that is known to favor the formation of intracellular iM structures.[126] In accordance with the previously reported results, all the investigated molecules, except for **21d**, succeeded in destabilizing the iMs as demonstrated by the significant reduction in the iM structures following the treatments (**Figure 37a-Figure S7, Supplementary Materials**). IF experiments highlighted that the compounds differ from each other in terms of potency in destabilizing the iMs, with **15c** and **18c** that—inducing a reduction of about 70% and 60% in the iMs fluorescence signal, respectively—turned out to be the most effective compounds. Notably, the major efficacy of **15c** and **18c** in destabilizing iMs well correlates with their capability to stabilize G4 structures (**Figure 37b Figure S8, Supplementary Materials**). Indeed, the dual activity of **15c** and **18c**, already observed at the biophysical level, is in line with the literature data showing that, under certain conditions, the formation of G4s and iM structures can be reciprocally influenced.[149] Altogether, our data confirmed the ability of the selected compounds to affect G4/iM structures also in the complexity of a cellular model, demonstrating the biological translatability of the biophysical results.

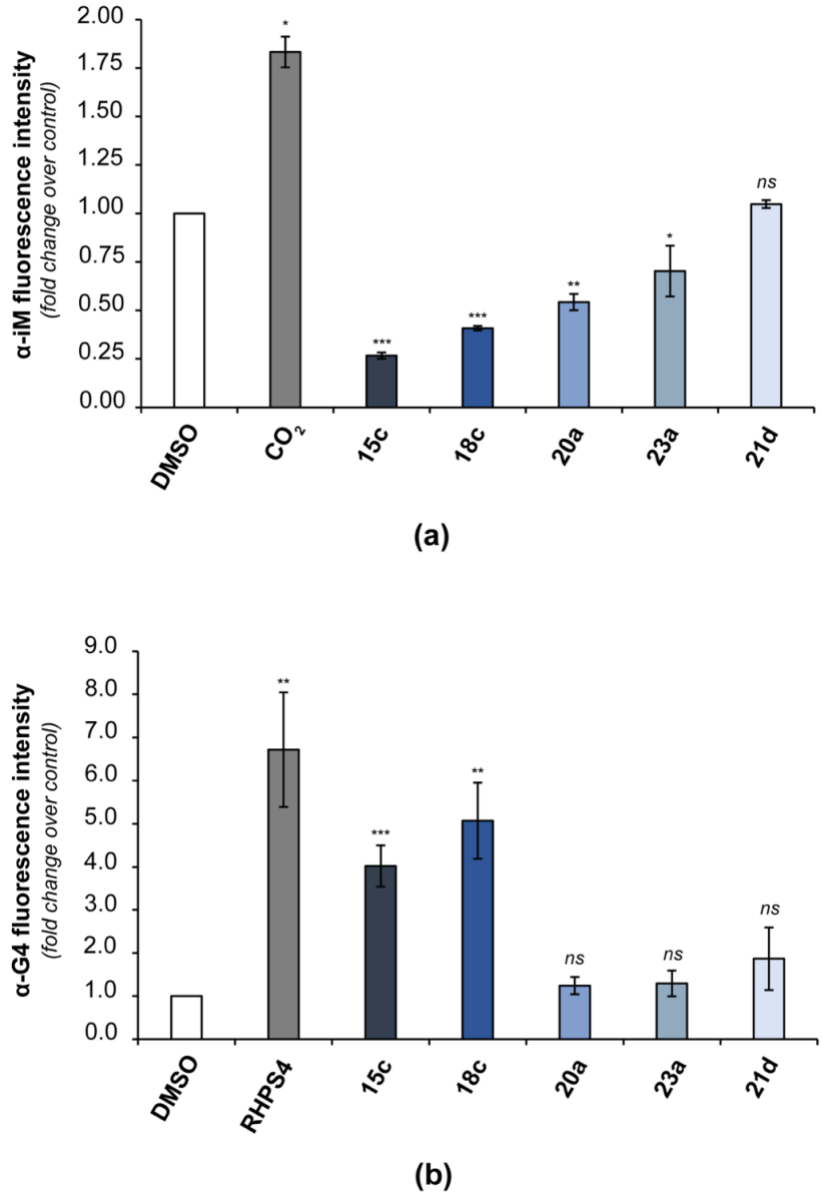


Figure 37. *Quantitative analysis of fluorescence intensity of anti-iM (**a**) and anti-G4 (**b**) signal. For each panel, 25 cells/condition were analysed by ImageJ software and the results were expressed as fold change of fluorescence intensity over the negative control (DMSO). Results are referred to the images shown in Figures S6 and S7, respectively. Histograms show the mean ± SD of three independent experiments. Statistical significance was calculated using unpaired student t-tests on GraphPad Prism 6 (***p<0.001; **p<0.01; *p<0.05).*

2.1.3 Conclusion

In the present study, we rationally designed a new library of potential G4/iM-targeting compounds according to the main structural features required to design effective noncanonical DNA-interacting compounds. Then, using multiple biophysical techniques, we dissected the profile of each compound in terms of effects on G4/iM structures. Finally, in order to translate the biophysical data into a more complex environment, we probed the capability of the selected compounds to affect the formation of G4/iM structures in U2OS cells. Interestingly, our biophysical screening led to the identification of compounds **15c** and **18c** as candidates capable of concomitantly stabilizing the G4 structures and destabilizing the iM ones, and compounds **20a** and **23a** that selectively affected the thermal stability of the iMs, showing negligible effects on G4 structures. NMR and fluorescence titration experiments provided insights into the interaction mode and affinity of the selected compounds towards the noncanonical DNA structures under investigation. Noteworthy, the results of the biophysical study found confirmation also at the biological level. Indeed, analyses performed by confocal microscopy showed that the same compounds were able to destabilize iM structures in U2OS cells, the best-characterized cell model used so far to study iMs formation. Intriguingly, **15c** and **18c** were able to effectively destabilize iM structures by simultaneously stabilizing G4s, so resulting—also in the complexity of the cellular environment—the most effective compounds among those synthesized and tested. Our investigation led to the identification of new compounds that might be used as tools to shed light on the mechanisms underlying the controversial biological roles of G4 and iM structures and their intricated relationship.

2.1.4 Experimental section

Circular dichroism. The samples for CD measurements were prepared by dissolving *Bcl-2 G4*, *c-Myc G4*, and *Tel*$_{23}$ *G4* in 25 mM potassium phosphate buffer (pH 7.0), and *Bcl-2 iM*, *c-Myc iM*, and *hTeloC iM* in 10 mM sodium phosphate buffer (pH 5.0),[150] and the *Hairpin* in 10 mM sodium phosphate buffer (pH 7.0). Circular dichroism (CD) experiments (spectra and melting) were performed on a Jasco J-815 spectropolarimeter (JASCO Inc., Tokyo, Japan) equipped with a PTC-423S/15 Peltier temperature controller, using a quartz cuvette with a path length of 0.1 cm. Data was obtained from samples at 15–20 μM final oligonucleotide concentration, in the absence and presence of 10 molar equivalents of each compound (2.5, 5, or 10 mM in 100% DMSO). CD spectra were recorded at 5 and 90 °C for *Bcl-2 iM*, *c-Myc iM*, and *hTeloC iM*, and at 20 °C and 100 °C for *Bcl-2 G4*, *c-Myc G4*, *Tel*$_{23}$ *G4*, and *Hairpin*. Each spectrum was acquired in a wavelength range of 220–360 nm, averaged over three scans, and subtracted from the buffer baseline. The scanning speed was set to 100 nm/min, with a 4 s response time, 1 nm data pitch, and 2 nm bandwidth. CD melting experiments were performed at 1 °C/min heating rate, in the 5–90 °C temperature range for *Bcl-2 iM*, *c-Myc iM*, and *hTeloC iM*, and in the 20–100 °C temperature range for *Bcl-2 G4*, *c-Myc G4*, *Tel*$_{23}$ *G4*, and *Hairpin*. Changes in the CD signal were followed at the wavelengths of the maximum CD intensity: 264 nm for *Bcl-2 G4* and *c-Myc G4*, 287 nm for *Tel*$_{23}$ *G4*, 288 nm for *Bcl-2 iM*, *c-Myc iM*, and *hTeloC iM*. As for *Hairpin*, CD melting curves were recorded by following the change in the CD signal at 251 nm, the wavelength of the minimum intensity value of the respective CD spectrum. All CD melting curves were normalized between 0 and 1. The apparent melting temperatures ($T_{1/2}$) were mathematically calculated by using the curve fitting function in Origin 7.0 software (OriginLab, Northampton, MA, USA). The $T_{1/2}$ values of the DNAs alone are: *Bcl-2 G4* = 59.1 (±0.9) °C; *c-Myc G4* = 78.6 (±0.4) °C; *Tel*$_{23}$ *G4* =

55 (±0.7) °C; *Bcl-2 iM* = 62.8 (±0.9) °C; *c-Myc iM* = 73.6 (±0.8) °C; *hTeloC iM* = 51.5 (±0.2) °C; *Hairpin* = 71.1 (±0.9) °C. $\Delta T_{1/2}$ values correspond to the difference between the DNA melting temperatures with and without compounds

Principal Component Analysis. The $\Delta T_{1/2}$ values obtained from the CD experiments were collected in a data matrix consisting of 31 rows (compounds) and 7 columns (DNA sequences). Before PCA, data was converted to absolute values, then mean-centered and scaled to unit variance (autoscaling). Autoscaling employs the standard deviation as a scaling factor thus giving all the $\Delta T_{1/2}$ values the same chance to affect the model. Moreover, data relative to the *Hairpin* and *Tel$_{23}$ G4* sequences were removed before the analysis because of the very small variation of the $\Delta T_{1/2}$ values (<4 °C). PCA was then computed by using PLS Toolbox 8.6.1 (Eigenvector Research Inc., Wenatchee, WA, USA) in Matlab R2015b (The Mathworks Inc., Natick, MA, USA) environment. The pK_a values for each compound were calculated by using the Percepta Software version 14.3.0 (ACD/Labs, Toronto, ON, Canada)

NMR spectroscopy. 1D ^{1}H-NMR Experiments The samples for nuclear magnetic resonance (NMR) spectroscopy were prepared at 100–150 μM final oligonucleotide concentration, in 250 μL (H_2O/D_2O 9:1) buffer solution (10 mM sodium phosphate buffer (pH 5.0) for *hTeloC iM*, and 10 mM potassium phosphate buffer (pH 5.0) for *c-Myc G4*). DNA/compound mixtures were obtained by adding 2 equivalents of the compounds directly to the DNA solution inside the NMR tube,[151] with a final DMSO concentration of 6%. NMR experiments were performed by employing a 700 MHz Bruker spectrometer. In particular, 1D ^{1}H-NMR spectra were recorded at 10 °C with a 'zgesgp' pulse-program (a gradient-based excitation sculpting using 180° water-selective pulses), including the following parameters: 128 scans, spectral width 17,241 Hz, delay 3 s, receiver gain 101, and 32 k points. All

NMR spectra were calibrated by centering the water signal at 4.69 ppm (at 10 °C and pH 5.0). NMR data processing was done by using the vendor software TOPSPIN 4.0.7 (Bruker Biospin Gmbh, Rheinstetten, Germany). ***NMR spectroscopy. STD NMR Experiments.*** In the 1D ^{1}H ligand-based STD NMR experiments, each compound (250 μM), previously solubilized in DMSO-d_6, was added to a DNA solution (20 μM, 600 μL) in 100% of deuterium oxide containing 10 mM phosphate buffer, pH 5.0. All the spectra were acquired at 298 K with a Bruker AVANCE NEO NMR spectrometer (Bruker Biospin Gmbh, Rheinstetten, Germany) operating at 700 MHz (^{1}H Larmor frequency), equipped with a 5 mm TCI 3 channels HCN cryo-probe head, optimized for ^{1}H sensitivity. The spectrometer was also equipped with a SampleCase (autosampler) for NMR screening. The spectra were processed with the vendor software TOPSPIN 4.0.7 (Bruker Biospin Gmbh, Rheinstetten, Germany). Saturation Transfer Difference spectra were acquired with 1024 scans, with on-resonance irradiation at 5.95 and 9.12 ppm for selective saturation of *c-Myc G4* and *hTeloC iM* resonances, respectively, and off-resonance irradiation at 40 ppm for reference spectra. A train of 40 Gaussian-shaped pulses of 50 ms (with 1 ms delay between pulses) was used, for a total saturation time of 4 s. The saturation width of the used radiofrequency pulses was 100 Hz. The STD spectra were obtained by internal subtraction of the saturated spectrum from the reference spectrum by phase cycling, with a spectral width of 20 ppm, relaxation delay 1.0 s, 16 k data points for acquisition, and 65 k for transformation.

Fluorescence Titrations Fluorescence titration experiments were performed at 25 °C on an FP-8300 spectrofluorometer (Jasco, Easton, MD, USA) equipped with a Peltier temperature controller system (Jasco PCT-818, Jasco, Easton, MD, USA). A 1 cm path length, sealed quartz cuvette was used. Titrations were carried out by stepwise addition (5 μL) of *c-Myc G4*, *c-Myc iM*, *hTeloC iM*, or a single-stranded DNA [d(CT)$_{15}$] solution (150–200 μM)

to a cell containing a fixed concentration of a compound solution (2.5–3.5 μM) in the appropriate buffer (10 mM potassium phosphate buffer (pH 5.0) for *c-Myc G4*, and 10 mM sodium phosphate buffer (pH 5.0) for *c-Myc iM*, *hTeloC iM*, and single-stranded DNA). The excitation wavelength was set at 275 nm, and emission spectra were recorded in the wavelength range of 285–600 nm. Both excitation and emission slit widths were set at 5 nm. After each DNA addition, the solution was stirred and allowed to equilibrate for 5 min before spectrum acquisition. The fraction of bound compound (α) at each point of the titration was calculated following the changes of fluorescence intensity at the maximum of intensity (307 nm). The compound concentration was corrected for dilution effects resulting from the change in volume due to DNA solution addition. It should be pointed out that the excitation of compounds at 275 nm led to a certain filter effect due to DNA absorption at this wavelength, which in part contributes to the observed quenching of compound fluorescence independent of binding. Titration curves were obtained by plotting α versus the DNA concentration. The equilibrium dissociation constant (K_b) and the stoichiometry of interaction were estimated by fitting the resulting curve to an independent and equivalent binding site model as previously described.[152] The experiments were repeated in duplicate, and the results are presented as the mean ± S.D. ***Immunofluorescence (IF) Microscopy Experiments***. Human osteosarcoma cells (U2OS) were purchased from American Type Culture Collection (ATCC) and were grown in Dulbecco's Modified Eagle Medium (DMEM, Euroclone, Milan, Italy) supplemented with 10% Fetal Bovine Serum (FBS), 2 mM L-glutamine, and antibiotics, at 37 °C, in a 5% CO_2-95% air atmosphere. As a positive control for the stabilization of iM structures, U2OS were cultured in the presence of 8% of CO_2 for 2.5 h, as reported elsewhere.[126] As for G-quadruplex structures stabilization, treatment with 1 μM RHPS4 for 24 h was used as a positive control.[153] All compounds were

dissolved at a concentration of 10 mM in DMSO. Cells were incubated for 24 h with the compounds, at a final concentration of 2 µM. cells were fixed in 4% formaldehyde in phosphate-buffered saline (PBS), at RT, for 10 min, permeabilized in 0.25% Triton X-100 in PBS 1×, at RT, for 5 min, and incubated with blocking solution (3% FBS in PBS 1×) at RT for 1 h. For immuno-labeling, cells were incubated at RT, for 2 h, with the flag-tagged antibody that recognizes iM structures. Successively, cells were washed three times with PBS 1× and incubated with the antibody anti-Flag (Sigma-Aldrich, St. Louis, MO, USA, #F7425) for 1 h. Finally, after three washes in PBS 1×, cells were incubated with the antibody anti-rabbit IgG (H + L), F(ab′)2 Fragment (Alexa Fluor 555 Conjugate) (Cell Signaling, Danvers, MA, USA, #4413S) for 1 h. For G4 structures staining, after the permeabilization, cells were incubated with the specific antibody (Anti-DNA/RNA G-quadruplex (BG4), Absolute Antibody, Oxford, UK, #Ab00174-1.1) at RT, for 2 h, washed three times with PBS 1×, and incubated with the antibody anti-mouse IgG (H + L), F(ab')2 Fragment (Alexa Fluor 555 Conjugate) (Cell Signaling Technology, Danvers, MA, USA, #4409S) for 1 h. For all the IF experiments, nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI, Sigma-Aldrich, St. Louis, MO, USA, #D9542), and fluorescence signals were recorded by using a Zeiss LSM 880 with airyscan (Zeiss, Jena, Germany), 63× magnification. IF experiments were quantified by Image J (version 1.53e, National Institutes of Health, NIH, Bethesda, MD, USA). 25 cells were screened for each condition and the results were expressed as fold change of fluorescence intensity (anti-G4 or anti-iM signal) over the negative control (DMSO). Histograms show the mean ± SD of three independent experiments. Statistical significance was calculated using unpaired student *t*-tests on Prism 6 (GraphPad, San Diego, CA, USA).

Chapter3 - STUDY OF REACTION ENVIRONMENT BY NMR

3.1 Photo-micellar Catalyzed Synthesis of Amides from Isocyanides: Optimization, Scope, and NMR Studies of Photocatalyst/Surfactant Interactions.

The merging of micellar and photoredox catalysis represents a key issue to promote “in water” photochemical transformations. In this chapter of the thesis is presented a photomicellar catalyzed synthesis of amides from *N*-methyl-*N*-alkyl aromatic amines and both aliphatic and aromatic isocyanides. The mild reaction conditions enabled a wide substrate scope and a good functional groups tolerance, as further shown in the late-stage functionalization of complex bioactive scaffolds. Furthermore, solution 1D and 2D NMR experiments performed, for the first time, in the presence of paramagnetic probes enabled the study of the reaction environment at the atomic level along with the localization of the photocatalyst with respect to the micelles, thus providing experimental data to drive the identification of optimum photocatalyst/surfactant pairing.

3.1.1 Introduction

Aqueous micellar catalysis, beyond "on water" and "with water" reactions,[154-157] has been established as a huge opportunity to perform *in water* reactions, thus drawing synthetic organic chemistry procedures closer to Nature's mild reaction conditions.[158-160] The aggregation of surfactants to micellar nanoreactors offers indeed the possibility of mimicking enzymes' hydrophobic pockets, which provide suitable reaction vessels for lipophilic substrates. By overcoming the need for large amounts of organic solvents, while preserving their efficiency and selectivity in promoting organic transformations, micellar catalysis enables the reduction of organic waste along with the prevention of water stream contamination arising from the use of water miscible solvents.[161] Despite the far-reaching advances in the application of micellar catalysis to a wide range of key synthetic transformations (spanning from condensations, oxidations, deprotections, multicomponent reactions, C−C and C heteroatom bond-forming processes, and even metal-catalyzed cross-couplings),[160,162] very few reports about visible-light promoted reactivities of radical precursors in water are so far available.[163-165] Water has been indeed mostly reported as a cosolvent, in order to achieve complete dissolution of reactants/reagents for homogeneous conditions or as the oxygen source.[166-168] Recently, Lipshutz developed an amphoteric polyethylene glycol ubiquinol succinate (PQS)-attached photocatalyst able to self-aggregate in water into nanomicelles, whereas the covalently bound Iridium complex promotes, upon visible light excitation, the formation of open shell species from suitable precursors (**Figure 38a**).[169] If substantial advantages are associated with this catalytic system, such as the possibility to recycle the metal catalyst several times, the necessity of a

covalent functionalization of the surfactants to introduce the active photocat alytic moiety could limit a broad application of this approach.

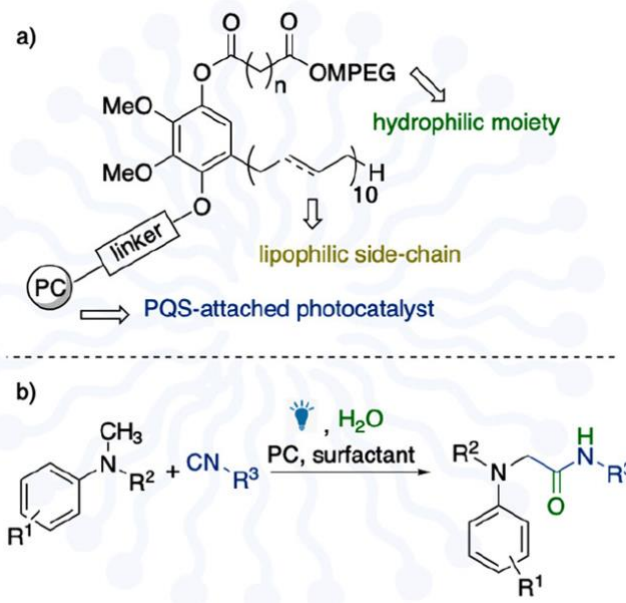


Figure 38. *a) Lipshutz PQS-attached photocatalyst; b) "in water" formation of amides from isocyanides herein developed.*

This observation is even more relevant when new chemical transformations are involved if considering the necessity to screen different photocatalysts endowed with different redox potentials, i.e., either stronger or poorer oxidants/reductants, as well as different surfactants to achieve optimum reaction conditions. To this end, the commercial availability of a wide range of surfactants covalently modified with photoredox active catalysts would expedite the spreading of micellar systems as reaction media in visible-light-triggered transformations. Meanwhile, with the hope that such an array of photoactive micelles will be soon available, we wondered if the switch from organic solvents to aqueous micellar solutions could be feasible, at least for those transformations involving water either as a cosolvent or a substrate. It is worth noting that, in the 1980s, the influence of micelles on photochemical reactivity was of great interest due to the possibility of modulating the reaction outcome via cage, preorientational, polarity, and counterion

effects.[170,171] In light of these considerations,[172,173] we sought to investigate the opportunities and the challenges of synthesizing amides from isocyanides and *N*-methyl-*N*-alkyl aromatic amines in water (**Figure 38b**).

3.1.2 Results and Discussion

Synthesis of Amides. The model reaction was reported in 2013 by Rueping et al. and required acetonitrile as the optimum solvent system, whereas water was used as a substrate in a 10 equiv amount.[174,175] A preliminary test reaction involving *N*,*N*-dimethylaniline **1** and toluenesulfonylmethyl isocyanide (TosMIC) **2**, performed in 2% TPGS-750-M (0.15 M) at room temperature and under irradiation with 16 W blue light-emitting diodes (LEDs) for 48 h, gave the desired product **3** in a fair 30% yield, proving that a water-based solvent system could be feasible for this transformation (**Figure 39**).

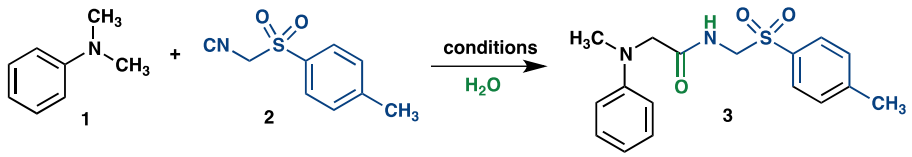


Figure 39. *Test reaction involving N,N-dimethylaniline **1** and toluenesulfonylmethyl isocyanide (TosMIC) **2**, performed in 2% TPGS-750-M (0.15 M) at room temperature and under irradiation with 16 W blue light-emitting diodes (LEDs) for 48 h, gave the desired product **3** in a 30% yield.*

To identify optimum in water reaction conditions, we performed the reaction in the presence of different micellar media (**Figure 40**) and different photoredox catalysts and by changing temperature, light source potency, and equivalents of starting materials **1** and **2** (**Table S12-Supplementary Materials**). Although irradiation with 30 W blue LEDs slightly increased the yield, the beneficial micellar effect was apparent if compared with the reaction performed in pure water. An increase of the reaction temperature to ~50 °C (fan switched off) led to detrimental results both in TPGS-750-M and in pure water, as well as the addition of Na_2CO_3 as a base.

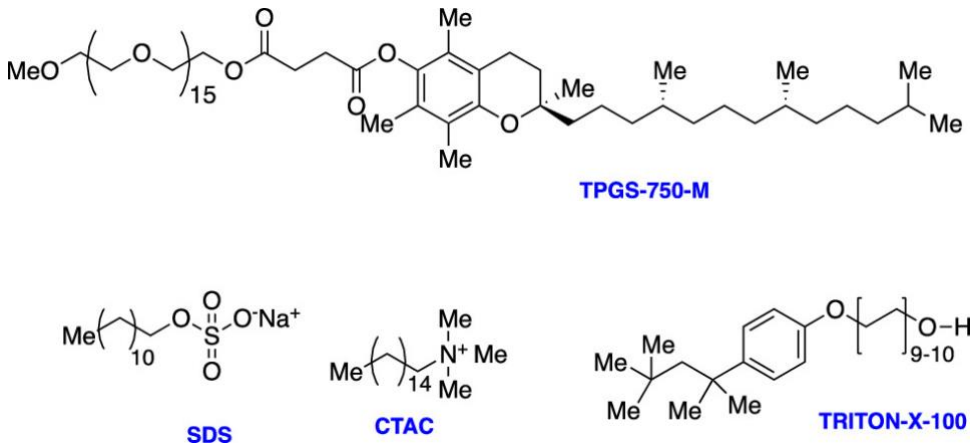


Figure 40. *Structures of the surfactants tested.*

Interestingly, an increment of *N*,*N*-dimethylaniline equivalents from 1 to 2 together with a lower amount of isocyanide **2** (1 equiv) led to a good 68% yield. Decreasing **1** to 1.5 equiv was detrimental, although, again, the micellar medium was shown to be key for the success of the reaction. A further screening of both metal-based and organic photocatalysts did not produce better results, whereas a survey of different micellar media resulted in the identification of a 2% solution of sodium dodecyl sulfate (SDS) as the best solvent system, enabling a yield comparable to that of acetonitrile. When these reaction conditions were applied to a 10-fold mmol scale (3 mmol), they proved to still be able to provide the desired product, although in a moderate 20% yield. Finally, further attempts to use an organic photocatalyst as well as reducing the catalyst loading were unproductive. On the basis of the reaction mechanism proposed by Rueping et al.,[175] the presence of molecular oxygen (the reaction was performed open flask) was key to provide an electron donor able to regenerate the iridium photocatalyst.

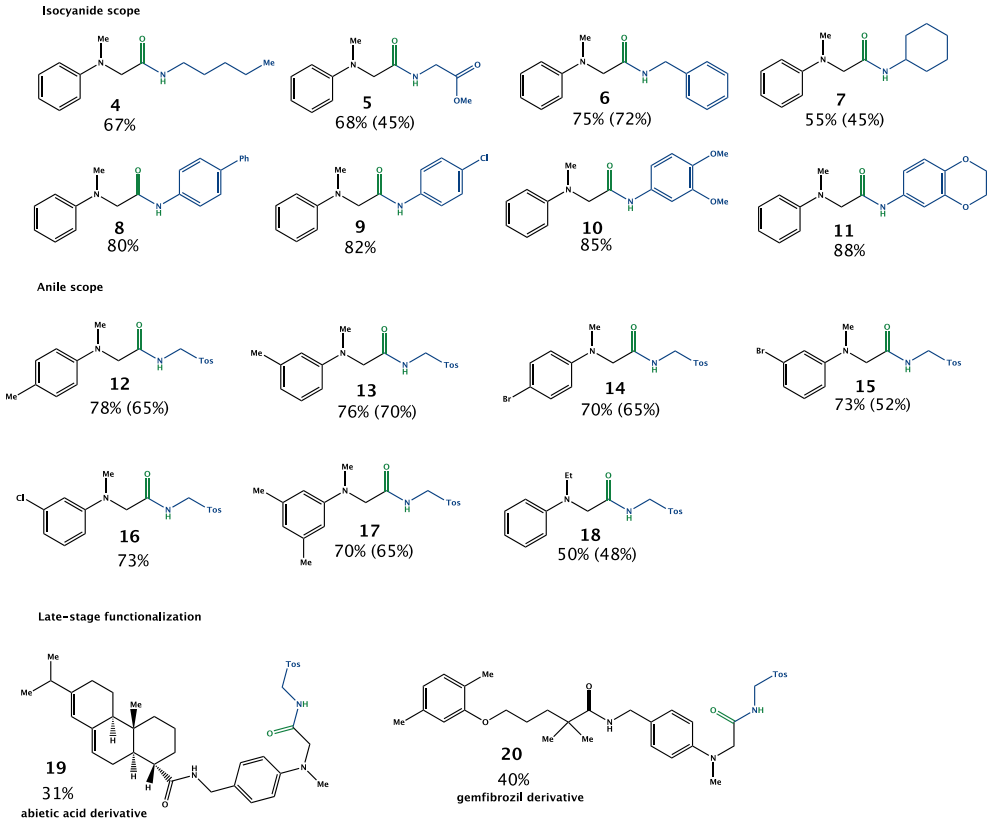
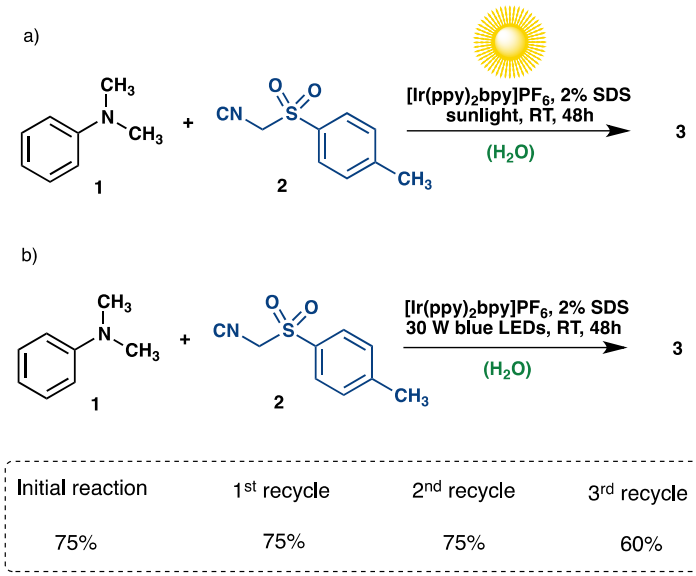


Figure 41. *Reaction scope (standard reaction conditions standard reaction conditions as in Table S12-supplementary, entry 24, performed on a 0.3 mmol scale; yields in parentheses refer to the reaction performed in acetonitrile as reported in ref 175).*

With the aim of proving the generality of these in water optimum reaction conditions, the substrate scope was investigated by reacting *N*,*N*-dimethylaniline **1** with different aliphatic isocyanides (yields in parentheses refer to the reaction performed in acetonitrile as reported in ref 175; **Figure 41**). Both primary (**4**−**6**, **Figure 41**) and secondary (**7**, **Figure 41**) isocyanides were competent substrates with yields spanning from 55 to 75%. In addition, less nucleophilic aromatic isocyanides, regardless the presence of either electron-with- drawing (**8** and **9**, **Figure 41**) or electron-donating (**10** and **11**, **Figure 41**) substituents, smoothly afforded the corresponding amides in 80−88% yields. Similarly, the aromatic amine scope was good, with yields

close to or higher than those obtained when using acetonitrile both for electron-poor (**14**–**16**, **Figure 41**) and electron-rich (**12**, **13**, **17**, **Figure 41**) anilines. Finally, nonsymmetrical *N*-ethyl- *N*-methylaniline regioselectively afforded amide **18** in a good 50% yield (**Figure 41**). Noteworthy, late-stage functionalization of abietic acid and gemfibrozil-derived *N*,*N*-dimethylanilines was accomplished in 31 and 40% yields, respectively. To further demonstrate the robustness of this methodology, the reaction was performed under sunlight irradiation, providing the target compound **3** in a good 60% yield (**Scheme 1**). Additionally, the micellar aqueous mixture can be recycled with no or slight change in the yield (**Scheme 1**).



Initial reaction	1st recycle	2nd recycle	3rd recycle
75%	75%	75%	60%

Scheme 1. *a) Natural sunlight induced reaction; b) Recycle performed by introducing all components to the aqueous reaction mixture from the previous reaction, after extraction with EtOAc.*

Nuclear Magnetic Resonance (NMR) Studies. To investigate whether a rational approach for selecting the best performing photocatalyst/surfactant pairs could be feasible, the interactions occurring between the photocatalyst and SDS and cetyltrimethylammonium chloride (CTAC) 2% solutions (i.e., the most and the least efficient micellar systems) were investigated at the atomic level via solution NMR techniques.[176] 1D ^{1}H NMR spectra of

[Ir(ppy)$_2$bpy]PF$_6$ in pure water and in the presence of micelles (**Figure 42**) were very different, indicating that the interaction between the catalyst and the micelles occurred in both cases. While the spectrum of the catalyst in pure water showed low intensity and broad signals (**Figure 42a**), probably due to its poor solubility and, as a consequence, to aggregation phenomena, both the proton spectra acquired in SDS and CTAC 2% solutions showed narrow and well-resolved signals (**Figure 42 b,c**), being similar to each other and similar to that obtained in organic solvents (e.g., CD_2Cl_2).[177]

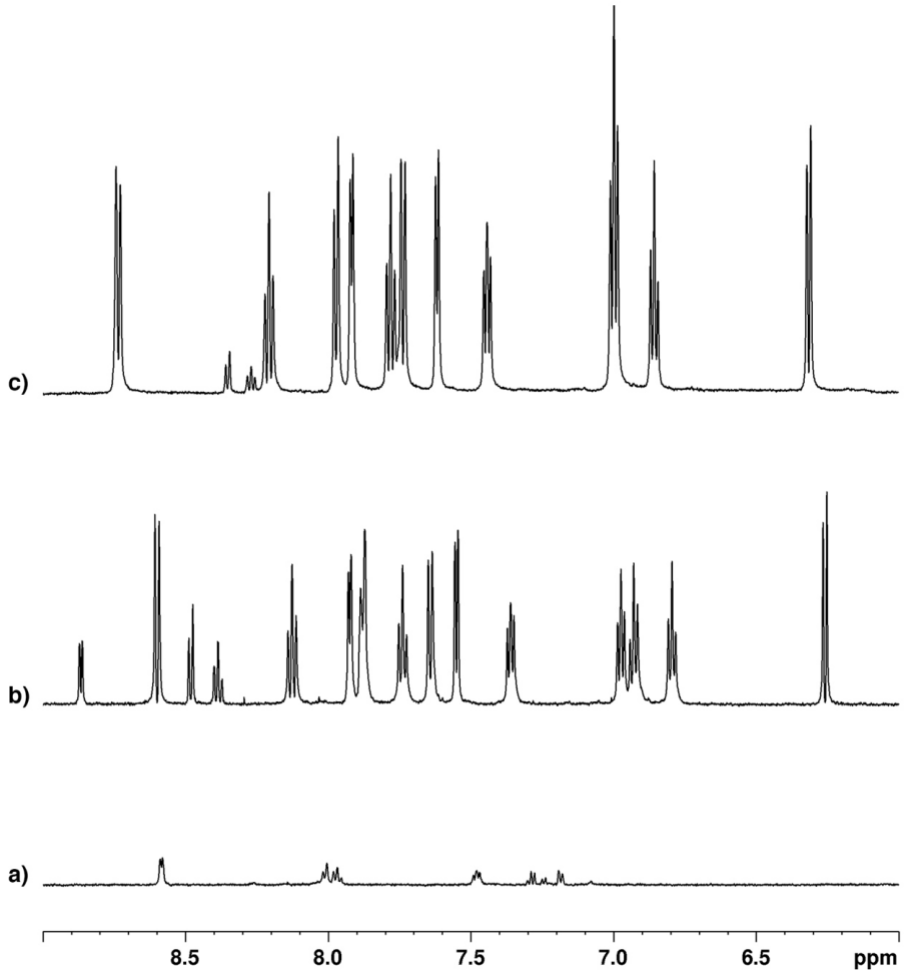


Figure 42. *^{1}H NMR of [Ir(ppy)$_2$bpy]PF$_6$ in water (a), in SDS (b) and CATC (c) solutions*

Complete assignment of the proton signals of $[Ir(ppy)_2bpy]PF_6$ in the presence of micelles was achieved via the analysis of 2D $^1H-^1H$ COSY,[2,178] TOCSY,[3] and NOESY[4] spectra (**Tables S13** and **S14**, **Supplementary Materials**). Furthermore, the localization of $[Ir(ppy)_2bpy]PF_6$ relative to the surface and the interior of SDS or CTAC micelles was studied using paramagnetic probes: 16-doxylstearic acid (16-doxyl) and Mn^{2+}. It is well-known, indeed, that unpaired electrons could lead to dramatically accelerated longitudinal and transverse relaxation rates of protons in spatial proximity via highly efficient spin and electron relaxation.[10] Therefore, these paramagnetic probes were expected to cause the broadening of the NMR signals and a decrease in resonance intensities either for a molecule outside the micelle (Mn^{2+}) or deeply buried in the micelle (16-doxyl).[179,180] Accordingly, ^1H NMR spectra of $[Ir(ppy)_2bpy]PF_6$, in micellar solutions, in the presence of increasing concentrations of the spin-labels (**Figure 43a−h** and **a′−h′**), with all other conditions kept constant, were recorded. The signal intensities of $[Ir(ppy)_2bpy]PF_6$ in SDS 2% solution showed significant reduction of all signals after the addition of both the Mn^{2+} and the 16-doxyl at the highest concentrations (**Figure 43 f−h** and **g′,h′**). This result suggested that the catalyst is, on average, positioned on the micelle surface and can flip from the outer to inner part of it so that it can feel both the external (Mn^{2+}) and internal (16-doxyl) paramagnetic effects.

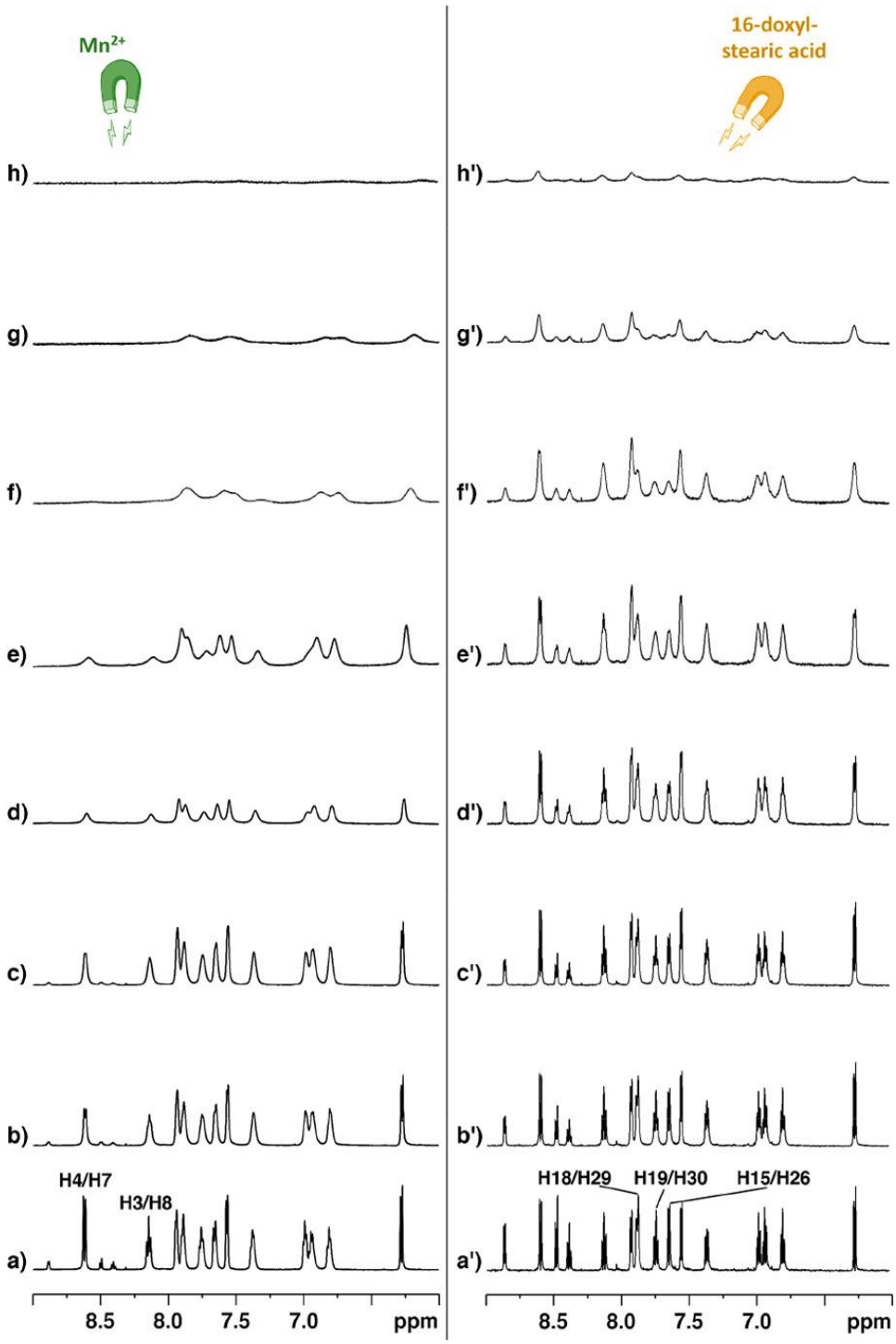


Figure 43. *^{1}H NMR spectra of $[Ir(ppy)_2bpy]PF_6$, in SDS 2% solution (a,a') and in the presence of 0.1 mM (b,b'), 0.2 mM (c,c'), 0.5 mM (d,d'), 1.0 mM (e,e'), 2.0 mM (f,f'), 3.0 mM (g,g'), and 5.0 mM (h,h') of Mn^{2+} (left) and 16-doxylstearic acid (right) probes.*

More detailed investigation revealed that protons H4/H7 and H3/H8 were the most affected by Mn^{2+}, whereas protons H15/H26, H19/ H30, and H18/H29 were the most affected by 16-doxyl even at the lower concentration used for these paramagnetic probes (**Figure 44**). Considering the same catalyst $[Ir(ppy)_2bpy]PF_6$ in CTAC micelles (**Figure 45a−h** and **a′−h′**), a generalized reduction of the signals' intensity was observed upon the addition of 16-doxyl (**Figure 45d′−h′**), while the addition of Mn^{2+} reduced the signal intensity of 30% at most (**Figure 45b−h**). These findings demonstrated that $[Ir(ppy)_2bpy]PF_6$ tended to be deeply inserted in CTAC micelles, probably because of the electronic repulsion between the positively charged metal center of the photocatalyst and the CTAC quaternary ammonium head.

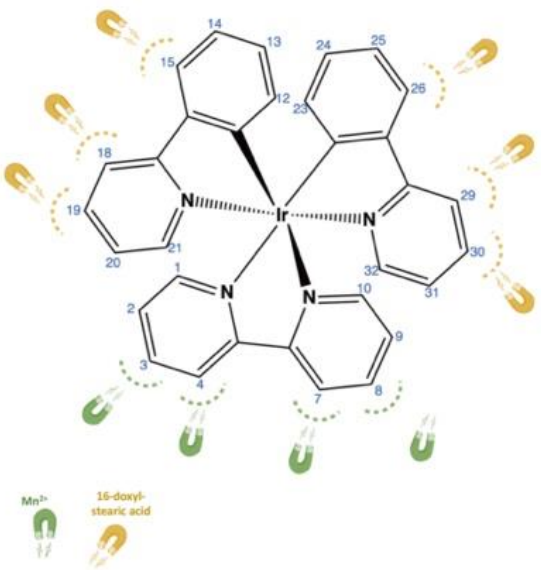


Figure 44. *Schematic representation of most affected protons of $[Ir(ppy)_2bpy]PF_6$ by Mn^{2+} (green) and 16-doxylstearic acid (yellow) in SDS 2% solution.*

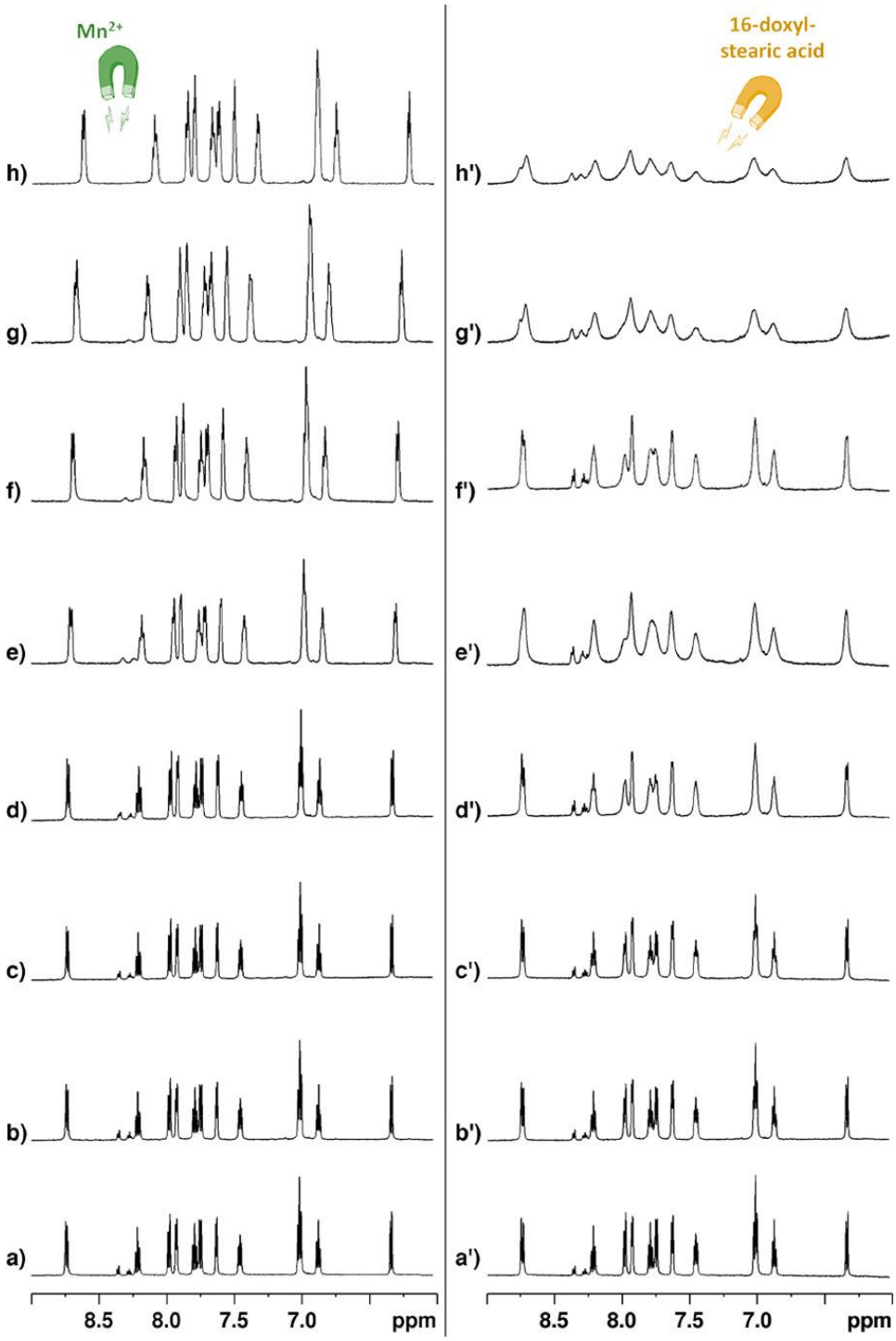


Figure 45. *^{1}H NMR spectra of $[Ir(ppy)_2bpy]PF_6$, in CTAC 2% solution (a,a') and in the presence of 0.1 mM (b,b'), 0.2 mM (c,c'), 0.5 mM (d,d'), 1.0 mM (e,e'), 2.0 mM (f,f'), 3.0 mM (g,g'), and 5.0 mM (h,h') of Mn^{2+} (left) and 16-doxylstearic acid (right) probes*

To further investigate the catalyst/micelle relative orientation, we acquired NOESY[4] spectra of the catalyst together with (not deuterated) SDS and CTAC 2% solutions (**Figure 46**). As shown in **Figure 46a**, NOE cross-peaks corresponding to intermolecular contacts are observable in the $[Ir(ppy)_2bpy]$-PF_6−CTAC system only from the main chain CH_2 of CTAC with all the catalyst protons, thus confirming that $[Ir(ppy)_2bpy]PF_6$ is deeply buried into these micelles. Considering the $[Ir(ppy)_2bpy]PF_6$/SDS system (**Figure 46b**), crosspeaks attributable to intermolecular contacts are observed between the catalyst and different proton signals of SDS (αCH_2, βCH_2, and main chain CH_2), confirming that the catalyst is positioned on the micelle surface in the presence of SDS micelles. In particular, catalysts H3/H8 and H20/H31 interact selectively with SDS αCH_2 and H4/H7, whereas H1/H10, H18/H29, H15, H26, H21/H3, and H12/H23 interact with SDS α and βCH_2, as reported in **Figure 46b**. The main chain CH_2 of SDS interacts with all of the catalyst protons, indicating a degree of movement freedom.

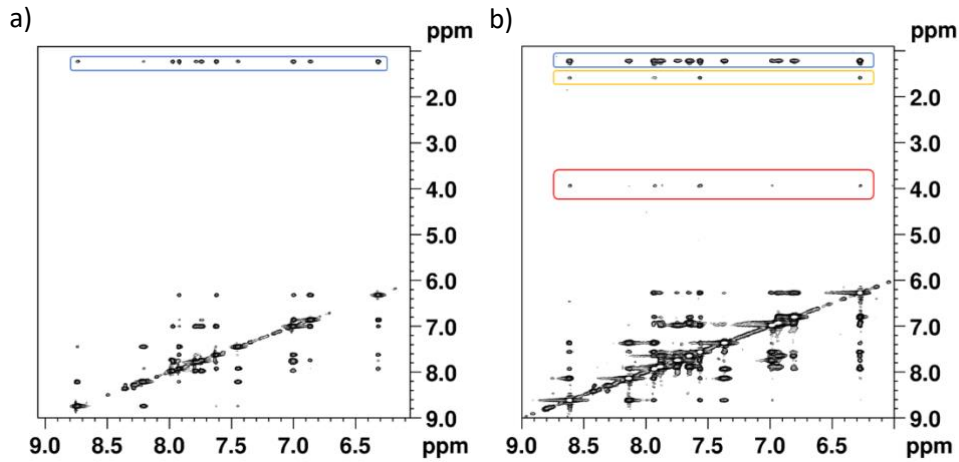


Figure 46. *$^1H-^1H$ 2D NOESY spectra of $[Ir(ppy)_2bpy]PF_6$ in CTAC (a) and (b) SDS 2% solutions. Different signal strips are evidenced as boxes: blue, main chain CH_2; red, αCH_2; yellow, β CH_2.*

Paramagnetic probes and NOE results are consistent and allowed us to establish the average relative disposition of the [Ir(ppy)$_2$bpy]PF$_6$/SDS system, as reported in **Figure 47**. In this model the 2,2′-bipyridine moiety is positioned mostly outside, whereas the 2-phenylpyridinates are located inside the micelle. Overall, the obtained experimental data strongly suggest that, for an optimal catalytic efficiency, the photocatalyst must be positioned on the micelles' surface and almost free to move inside and outside the micelles. Consistently, the latter conditions could be enabled by micellar systems with a reverse polarity with the respect to the photocatalyst as in the case of [Ir(ppy)$_2$bpy]PF$_6$ and SDS.

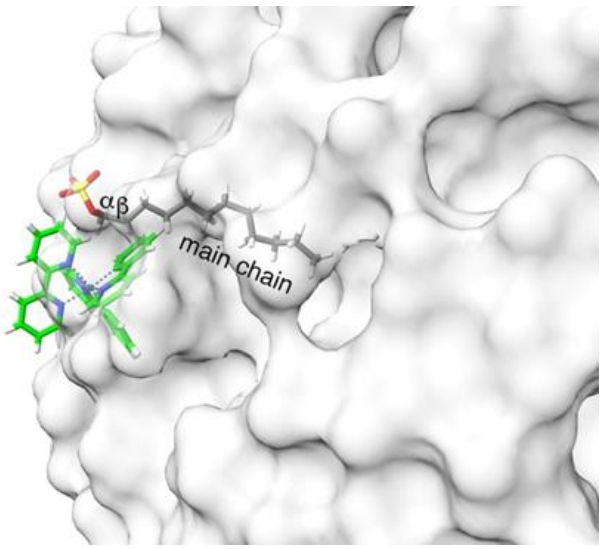


Figure 47. *Graphic representation of [Ir(ppy)$_2$bpy]PF$_6$ bound with SDS micelle. Catalyst atoms are reported with the following color codes: carbon, green; hydrogen, white; nitrogen, blue. The micelle is reported as a gray surface. For the sake of clarity, only one SDS molecule is shown (color code: carbon, gray; hydrogen, white;oxygen, red; sulfur, yellow).*

3.1.3 Conclusion

In conclusion, optimum photomicellar catalytic conditions for the synthesis of amides starting from readily available *N*-methyl- *N*-alkyl aromatic amines and both aliphatic and aromatic isocyanides have been developed. The substrate scope is broad with a good functional group tolerance as further highlighted in late-stage functionalization of drugs and natural compounds. Moreover, this transformation can be performed under sunlight irradiation without any decrease in the yield. The possibility to recycle the micellar solvent system, as shown for up to four cycles, while preserving conversion and yields of the amide target compounds, further supports the greenness of the developed conditions. Furthermore, in order to provide information at the atomic level about the interactions between the photocatalyst and the micelles as well as the photocatalyst localization with respect to the latter, solution 1D and 2D NMR experiments were performed in the presence of paramagnetic probes. To our knowledge, this represents the first application of such probes to the characterization of a (photo)micellar reaction environment. Such studies revealed a reverse polarity principle, according to which a negatively charged surfactant such as SDS could provide the localization of a positively charged photocatalyst on the micelles' surface, while a cation surfactant such as CTAC will cause a burying of the photocatalyst deeply into the micelles, probably preventing its interaction with light. We are confident that the rationalization of the optimum photocatalyst/ surfactant pairing could drive the exploration and the optimization of future photomicellar-catalyzed reactions, thus promoting the merging of these intrinsically green chemical approaches underlying the flourishing of in water photoredox catalytic transformations.

3.1.4 Experimental section

General Synthetic Methods. Commercially available reagents and solvents were used without further purification. When necessary, the reactions were performed in oven-dried glassware under a positive pressure of dry nitrogen. Photochemical reactions were carried out using a PhotoRedOx Box (EvoluChemTM) with a 30 W blue LED (EvoluChemTM, model: HCK1012-01-008, wavelength 450 nm, LED: CREE XPE). A holder suitable for 4 mL scintillation vials (45 × 14.7 mm) was fitted with a Schlenk flask: this allows a fixed sample placement distance from the light source. All NMR spectra were obtained with a Bruker Avance NEO 400 or 700 MHz instrument. Experiments for structure elucidation were performed in CDCl3 at 25 °C with a RT-DR-BF/ 1H-5 mm-OZ SmartProbe. High-resolution ESI-MS spectra were obtained on a Thermo LTQ Orbitrap XL mass spectrometer. The spectra were recorded by infusion into the ESI source using MeOH as the solvent. Chemical shifts (δ) are reported in parts per million (ppm) relative to the residual solvent peak. Column chromatography was performed on silica gel (70−230 mesh ASTM) using the reported eluents. Thin layer chromatography (TLC) was carried out on 5 × 20 cm plates with a layer thickness of 0.25 mm (silica gel 60 F254) to monitor the reaction using UV as the revelation methods.

General procedure for the preparation of compounds 3-20 (Method A). To a 4 mL colourless glass vial equipped with a magnetic stir bar were added the isocyanide (0.3 mmol, 1 equiv.), the photocatalyst $[Ir(ppy)_2bpy]PF_6$ (0.003 mmol, 1% mol), and the aniline derivative (0.6 mmol, 2 equiv.). Then 2 mL of the micellar solution (2% weight/volume concentration) (0.15 M) was added into the reaction vial and the resulting mixture was stirred open flask in a PhotoRedOx Box, under 30 W blue LED irradiation, at room temperature for 48 h. After the completion of the reaction, as monitored by TLC (Hexane/EtOAc 7:3 and 5:5), the reaction mixture was extracted with

EtOAc (x3), the collected organic layers were washed with brine (x1), dried over dry Na_2SO_4, filtered and evaporated to dryness to give a reaction crude which was purified as specified for each compound.

2-(Methyl(phenyl)amino)-*N*-(tosylmethyl)acetamide 3.[175] The title compound **3** was prepared following the general procedure method A. The crude material was purified by trituration with Hexane/EtOAc (2:1) or column cromatography (Hexane/EtOAc 8:2 to 5:5). Yield: 74.8 mg, 75%; white amorphous solid; R_f = 0.4 (TLC Hexane/EtOAc 5:5). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.72 (d, J = 8.3 Hz, 2H), 7.37-7.26 (m, 5H), 6.89 (t, J = 7.3 Hz, 1H), 6.68 (d, J = 7.9 Hz, 2H), 4.69 (d, J = 6.9 Hz, 2H), 3.75 (s, 2H), 2.99 (s, 3H), 2.46 (s, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 170.3, 149.0, 145.6, 133.8, 130.0, 129.5,128.8, 119.3, 113.4, 59.8, 58.6, 40.1, 21.8. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{17}H_{21}N_2O_3S^+$ = 333.1267; found: 333.1279. **2-(Methyl(phenyl)amino)-*N*-pentylacetamide 4.**[181] The title compound **4** was prepared following the general procedure method A. The crude material was purified by column cromatography (Hexane/EtOAc 8:2 to 6:4). Yield: 47.1 mg, 67%; colourless oil; R_f = 0.4 (TLC Hexane/EtOAc 6:4). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.31-7.23 (m, 2H), 6.84 (t, J = 7.3 Hz, 1H), 6.77-6-70 (m, 2H), 6.56 (brs, 1H), 3.85 (s, 2H), 3.27 (q, J = 9.4 Hz, 2H), 3.00 (s, 3H), 1.49-1.41 (m, 2H), 1.34-1.16 (m, 4H), 0.85 (t, J = 7.2 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 170.2, 149.3, 129.4, 118.6, 113.1, 59.0, 39.7, 39.2, 29.3, 29.0, 22.3, 14.0. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{14}H_{23}N_2O^+$ = 235.1805; found: 235.1808.

Methyl *N*-methyl-*N*-phenylglycylglycinate 5.[175] The title compound **5** was prepared following the general procedure method A. The crude material was purified by column cromatography (Hexane/EtOAc 8:2 to 6:4). Yield: 48.2 mg, 68%; colourless oil; R_f = 0.4 (TLC Hexane/EtOAc 6:4). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.31-7.25 (m, 2H), 7.05 (brs, 1H), 6.85 (t, J = 7.3 Hz, 1H), 6.81-6.75 (m, 2H), 4.07 (d, J = 5.8 Hz, 2H), 3.91 (s, 2H), 3.73 (s, 3H), 3.04

(s, 3H).^{13}C NMR (100 MHz, $CDCl_3$): δ 171.1, 170.0, 149.3, 129.4, 118.9, 113.4, 58.8, 52.4, 40.8, 39.8. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{12}H_{17}N_2O_3^+$ = 237.1234; found: 237.1243.

***N*-benzyl-2-(methyl(phenyl)amino)acetamide 6.** [175]The title compound **6** was prepared following the general procedure method A. The crude material was purified by column cromatography (Hexane/EtOAc 8:2 to 6:4). Yield: 57.2 mg, 75%; colourless oil; R_f = 0.4 (TLC Hexane/EtOAc 6:4). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.32-7.23 (m, 5H), 7.22-7.17 (m, 2H), 6.91 (brs, 1H, NH), 6.84 (t, J = 7.3 Hz, 1H), 6.77-6.71 (m, 2H), 4.49 (d, J = 6.0 Hz, 2H), 3.92 (s, 2H), 3.00 (s, 3H).^{13}C NMR (100 MHz, $CDCl_3$): δ 170.4, 149.3, 138.0, 129.4, 128.7, 127.5, 127.5, 118.8, 113.3, 59.0, 43.1, 39.9. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{16}H_{19}N_2O^+$ = 255.1492; found: 255.1504.

***N*-cyclohexyl-2-(methyl(phenyl)amino)acetamide 7.** [175] The title compound **7** was prepared following the general procedure method A. The crude material was purified by column cromatography (Hexane/EtOAc 8:2 to 6:4). Yield: 40.6 mg, 55%; colourless oil; R_f = 0.4 (TLC Hexane/EtOAc 6:4). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.31-7.23 (m, 2H), 6.84 (t, J = 7.3 Hz, 1H), 6.77-6-71 (m, 2H), 6.45 (brs, 1H), 3.86-3-81 (m, 3H), 2.99 (s, 3H), 1.90-1.80 (m, 2H), 1.72-1.53 (m, 4H), 1.40-1.34 (m, 2H), 1.17-0.99 (m, 2H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 169.3, 150.1, 129.4, 118.7, 113.3, 59.2, 47.8, 39.7, 33.0, 25.4, 24.8. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{15}H_{23}N_2O^+$ = 247.1805; found: 247.1815.

***N*-([1,1'-biphenyl]-4-yl)-2-(methyl(phenyl)amino)acetamide 8.** The title compound **8** was prepared following the general procedure method A (with 3 equiv. of *N,N*-dimethylaniline for 72h). The crude material was purified by column cromatography (Hexane/EtOAc 8:2 to 6:4). Yield: 75.9 mg, 80%; pale-brown solid; R_f = 0.4 (TLC Hexane/EtOAc 6:4). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.47 (brs, 1H, NH), 7.63 – 7.58 (m, 2H), 7.58-7.53 (m, 4H), 7.45-7.40 (m,2H), 7.36-7.29 (m, 2H), 6.91 (t, J = 7.3 Hz, 1H), 6.88 – 6.84 (m, 2H),

3.98 (s, 4H), 3.10 (s, 3H).^{13}C NMR (100 MHz, $CDCl_3$): δ 168.8, 149.5, 140.5, 137.5, 136.5, 129.6, 128.8, 127.6, 127.2, 126.9, 120.2, 119.5, 113.8, 60.1, 40.1. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{21}H_{21}N_2O^+$ = 317.1648; found: 317.1662.

***N*-(4-chlorophenyl)-2-(methyl(phenyl)amino)acetamide 9.**[182] The title compound **9** was prepared following the general procedure method A. The crude material was purified by column chromatography (Hexane/EtOAc 8:2 to 6:4). Yield: 67.6 mg, 82%; whitish solid; R_f = 0.4 (TLC Hexane/EtOAc 6:4). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.43 (s, 1H), 7.52-7.44 (m, 2H), 7.35-7.26 (m, 4H), 6.91 (t, J = 7.3 Hz, 1H), 6.85-6.80 (m, 2H), 3.95 (s, 2H), 3.08 (s, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 168.9, 149.6, 135.9, 129.4, 129.7, 129.2, 121.3, 119.7, 113.9, 60.2, 40.3. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{15}H_{16}ClN_2O^+$ = 275.0946; found: 275.0955.

***N*-(3,4-dimethoxyphenyl)-2-(methyl(phenyl)amino) acetamide 10.** The title compound **10** was prepared following the general procedure method A (with 3 equiv. of *N,N*-dimethylaniline and 2% of photocatalyst for 72h). The crude material was purified by column chromatography (Hexane/EtOAc 8:2 to 6:4). Yield: 76.6 mg, 85%; yellow solid; R_f = 0.4 (TLC Hexane/EtOAc 6:4). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.33 (s, 1H), 7.35-7.28 (m, 3H), 6.93-6.87 (m, 2H), 6.86-6.82 (m, 2H), 6.79 (d, J = 8.6 Hz, 1H), 3.95 (s, 2H), 3.89 (s, 3H), 3.85 (s, 3H), 3.08 (s, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 168.6, 149.7, 149.2, 146.2, 131.0, 129.6, 119.5, 113.9, 112.0, 111.4, 104.9, 60.2, 56.2, 56.1, 40.2. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{17}H_{21}N_2O_3^+$ = 301.1547; found: 301.1556.

***N*-(2,3-dihydrobenzo[b][1,4]dioxin-6-yl)-2-(methyl(phenyl) -amino)acetamide 11.** The title compound **11** was prepared following the general procedure method A (with 3 equiv. of *N,N*-dimethylaniline and 2% of photocatalyst for 72h). The crude material was purified by column chromatography (Hexane/EtOAc 8:2 to 6:4). Yield: 78.8 mg, 88%; yellow

solid; R_f = 0.4 (TLC Hexane/EtOAc 6:4). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.26 (s, 1H), 7.33-7.27 (m, 2H), 7.17 (d, J = 2.5 Hz, 1H), 6.92-6.85 (m, 2H), 6.84-6.76 (m, 3H), 4.25-4.20 (m, 4H), 3.93 (s, 2H), 3.06 (s, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 168.5, 149.6, 143.6, 140.8, 131.0, 129.6, 119.4, 117.3, 113.8, 113.7, 109.9, 64.5, 64.4, 60.0, 40.2. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{17}H_{19}N_2O_3^+$ = 299.1390; found: 299.1399.

2-(Methyl(p-tolyl)amino)-N-(tosylmethyl)acetamide 12.[175] The title compound **12** was prepared following the general procedure method A. The crude material was purified by column chromatography (Hexane/EtOAc 8:2 to 6:4). Yield: 81.0 mg, 78%; whitish solid; R_f = 0.4 (TLC Hexane/EtOAc 5:5). ^{1}H NMR (700 MHz, $CDCl_3$): δ 7.75-7.70 (m, 2H), 7.38 (s, 1H), 7.36-7.32 (m, 2H), 7.12-7.07 (m, 2H), 6.65-6.55 (m, 2H), 4.69 (d, J = 6.8 Hz, 2H), 3.70 (s, 2H), 2.95 (s, 3H), 2.46 (s, 3H), 2.29 (s, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 170.5, 147.0, 145.6, 133.8, 130.0, 129.9, 128.9, 128.8, 113.8, 59.8, 58.9, 40.4, 21.8, 20.3. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{18}H_{23}N_2O_3S^+$ = 347.1424; found: 347.1435.

2-(Methyl(m-tolyl)amino)-*N*-(tosylmethyl)acetamide 13. [175] The title compound **13** was prepared following the general procedure method A. The crude material was purified by column chromatography (Hexane/EtOAc 8:2 to 6:4). Yield: 79.0 mg, 76%; whitish solid; R_f = 0.4 (TLC Hexane/EtOAc 5:5). ^{1}H NMR (700 MHz, $CDCl_3$): δ 7.75-7.70 (m, 2H), 7.36-7.29 (m, 3H), 7.17 (t, J = 7.8 Hz, 1H), 6.71 (d, J = 7.1 Hz, 1H), 6.52 (s, 1H), 6.48 (d, J = 8.1 Hz, 1H), 4.68 (d, J = 6.8 Hz, 2H), 3.74 (s, 2H), 2.97 (s, 3H), 2.45 (s, 3H), 2.34 (s, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 170.4, 149.1, 145.5, 139.4, 133.8, 130.0,129.3, 128.8, 120.2, 114.2, 110.6, 59.9, 58.6, 40.1, 21.8, 21.8. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{18}H_{23}N_2O_3S^+$ = 347.1424; found: 347.1437.

-((4-Bromophenyl)(methyl)amino)-*N*-(tosylmethyl) acetamide 14.[183] The title compound **14** was prepared following the general procedure method A.

The crude material was purified by column chromatography (Hexane/EtOAc 8:2 to 6:4). Yield: 86.4 mg, 70%; whitish solid; R_f = 0.4 (TLC Hexane/EtOAc 5:5). ^{1}H NMR (700 MHz, $CDCl_3$): δ 7.76-7.72 (m, 2H), 7.40-7.30 (m, 4H), 7.18 (s, 2H), 6.55-6.50 (m, 2H), 4.69 (d, J = 6.8 Hz, 2H), 3.74 (s, 2H), 2.98 (s, 3H), 2.47 (s, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 169.8, 147.9, 145.7, 133.7, 132.2, 130.1, 130.0, 128.8, 128.8, 115.0, 111.5, 59.8, 58.4, 40.2, 21.8. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{17}H_{20}BrN_2O_3S^+$ = 411.0373; found: 411.0387.

2-((3-Bromophenyl)(methyl)amino)-*N*-(tosylmethyl)-acetamide 15.[175] The title compound **15** was prepared following the general procedure method A. The crude material was purified by trituration with Hexane/EtOAc (2:1). Yield: 90.1 mg, 73%; whitish solid; R_f = 0.4 (TLC Hexane/EtOAc 5:5). ^{1}H NMR (700 MHz, $CDCl_3$): δ 7.75-7.70 (m, 2H), 7.38-7.33 (m, 2H), 7.17 (t, J = 6.6 Hz, 1H), 7.13 (t, J = 8.1 Hz, 1H), 7.00 (d, J = 7.8 Hz, 1H), 6.82 (s, 1H), 6.57 (dd, J = 8.3, 2.2 Hz, 1H), 4.68 (d, J = 6.8 Hz, 2H), 3.76 (s, 2H), 2.98 (s, 3H), 2.46 (s, 3H). ^{13}C NMR (176 MHz, $CDCl_3$): δ 169.6, 150.2, 145.7, 133.7, 130.7, 130.1, 128.8, 123.7, 122.0, 116.6, 111.9, 59.8, 58.0, 39.9, 21.8. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{17}H_{20}BrN_2O_3S^+$ = 411.0373; found: 411.0389.

2-((3-Chlorophenyl)(methyl)amino)-*N*-(tosylmethyl)-acetamide 16. The title compound **16** was prepared following the general procedure method A. The crude material was purified by trituration with Hexane/EtOAc (2:1). Yield: 80.3 mg, 73%; pale-brown solid; R_f = 0.4 (TLC Hexane/EtOAc 5:5). ^{1}H NMR (700 MHz, $CDCl_3$): δ 7.75-7.70 (m, 2H), 7.38-7.33 (m, 2H), 7.19 (t, J = 8.1 Hz, 1H), 7.14 (t, J = 6.5 Hz, 1H), 6.86 (dd, J = 7.9, 1.0 Hz, 1H), 6.65 (t, J = 2.0 Hz, 1H), 6.53 (dd, J = 8.4, 2.4 Hz, 1H), 4.69 (d, J = 6.8 Hz, 2H), 3.76 (s, 2H), 2.99 (s, 3H), 2.46 (s, 3H). ^{13}C NMR (176 MHz, $CDCl_3$): δ 169.6, 150.0, 145.7, 135.4, 133.7, 130.5, 130.0, 128.8, 119.1, 113.3, 111.4,

59.8, 58.1, 40.0, 21.8. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{17}H_{20}ClN_2O_3S^+$ = 367.0878; found: 367.0893.

2-((3,5-Dimethylphenyl)(methyl)amino)-*N*-(tosylmethyl)acetamide 17. [175] The title compound **17** was prepared following the general procedure method A. The crude material was purified by trituration with Hexane/EtOAc (2:1). Yield: 75.7 mg, 70%; whitish solid; R_f = 0.4 (TLC Hexane/EtOAc 5:5). ^{1}H NMR (700 MHz, $CDCl_3$): δ 7.74-7.69 (m, 2H), 7.34-7.29 (m, 2H), 6.55 (s, 1H), 6.33 (s, 2H), 4.68 (d, J = 6.9 Hz, 2H), 3.73 (s, 2H), 2.95 (s, 3H), 2.45 (s, 3H), 2.30 (s, 6H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 169.5, 148.2, 144.5, 138.2, 132.9, 128.9, 127.8, 120.2, 110.4, 58.8, 57.6, 39.0, 20.8, 20.7. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{19}H_{25}N_2O_3S^+$ = 361.1580; found: 361.1592.

2-(Ethyl(phenyl)amino)-*N*-(tosylmethyl)acetamide 18. [175] The title compound **18** was prepared following the general procedure method A. The crude material was purified by column cromatography (Hexane/EtOAc 8:2 to 6:4). Yield: 52 mg, 50%; white solid; R_f = 0.4 (TLC Hexane/EtOAc 5:5). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.73-7.68 (m, 2H), 7.36-7.26 (m, 5H), 6.88 (t, J = 7.3 Hz, 1H), 6.70-6.65 (m, 2H), 4.67 (d, J = 6.9 Hz, 2H), 3.75 (s, 2H), 3.43 (q, J = 7.1 Hz, 2H), 2.47 (s, 3H), 1.19 (t, J = 7.1 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 170.5, 147.4, 145.5, 133.9, 129.9, 129.6, 128.8, 119.0, 113.7, 59.9, 55.4, 46.5, 21.8, 11.5. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{18}H_{23}N_2O_3S^+$ = 347.4525; found: 347.4534.

(1*R*,4*aR*,4*bR*,10*aR*)-7-isopropyl-1,4*a*-dimethyl-*N*-(4-(methyl(2-oxo-2((tosylmethyl)amino)ethyl)amino)benzyl)-1,2,3,4,4*a*,4*b*,5,6,10,10*a*-decahydrophenanthrene-1-carboxamide 19. The title compound **19** was prepared following the general procedure method A (using 0.15 mmol of isocyanide). The crude material was purified by column cromatography (Hexane/EtOAc 8:2 to 5:5). Yield: 30 mg, 31%; white solid; R_f = 0.4 (TLC Hexane/EtOAc 5:5). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.75-7.70 (m, 2H), 7.36-7.31 (m, 2H), 7.20-7.15 (m, 2H), 6.65-6.60 (m, 2H), 5.94 (t, J = 5.0 Hz, 1H),

5.74 (s, 1H), 5.32 (d, J = 4.5 Hz, 1H), 4.68 (d, J = 6.9 Hz, 2H), 4.38 (d, J = 5.6 Hz, 1H), 4.31 (d, J = 5.1 Hz, 1H), 3.74 (s, 2H), 2.98 (s, 3H), 2.46 (s, 3H), 2.21 (s, 1H), 2.10-1.79 (m, 10H), 1.62-1.50 (m, 2H), 1.27-1.22 (m, 5H), 1.00 (dd, J = 6.8, 3.3 Hz, 6H), 0.82 (s, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 178.2, 170.1, 148.4, 145.6, 145.3, 135.6, 133.8, 130.0, 129.4, 129.2, 128.8, 122.4, 120.4, 113.6, 59.9, 58.5, 51.0, 46.4, 45.7, 43.3, 40.1, 38.3, 37.7, 34.9, 34.7, 27.4, 25.4, 22.5, 21.8, 21.4, 20.8, 18.3, 17.1, 14.1. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{38}H_{52}N_3O_4S^+$ = 646.3673; found: 646.3701.

5-(2,5-Dimethylphenoxy)-2,2-dimethyl-*N*-(4-(methyl(2-oxo-2-((tosylmethyl)amino)ethyl)amino)benzyl)pentanamide 20. The title compound **20** was prepared following the general procedure method A (using 0.12 mmol of isocyanide). The crude material was purified by column cromatography (Hexane/EtOAc 8:2 to 5:5). Yield: 28.5 mg, 40%; colour-less oil; R_f = 0.4 (TLC Hexane/EtOAc 5:5). ^{1}H NMR (400 MHz, $CDCl_3$): 7.73-7.68 (m, 2H), 7.36-7.27 (m, 3H), 7.20-7.15 (m, 2H), 6.99 (d, J = 7.4 Hz, 1H), 6.65-6.60 (m, 4H), 5.92 (t, J = 5.0 Hz, 1H), 4.66 (d, J = 6.9 Hz, 2H), 4.36 (d, J = 5.5 Hz, 2H), 3.91 (t, J = 5.7 Hz, 2H), 3.73 (s, 2H), 2.96 (s, 3H), 2.45 (s, 3H), 2.30 (s, 3H), 2.14 (s, 3H), 1.80-1.67 (m, 2H), 1.22 (s, 6H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 177.3, 170.1, 156.9, 148.4, 145.6, 136.5, 133.8, 130.3, 130.0, 129.3, 129.1, 128.8, 123.5, 120.8, 113.6, 112.1, 68.0, 59.9, 58.4, 43.0, 41.9, 40.1, 37.5, 25.6, 25.2, 21.8, 21.4, 15.8. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{33}H_{44}N_3O_5S^+$ = 594.2996; found: 594.3016.

(1*R*,4*aR*,4*bR*,10*aR*)-*N*-(4-(dimethylamino)benzyl)-7-isopropyl-1,4*a*-dimethyl-1,2,3,4,4*a*,4*b*,5,6,10,10*a*-decahydrophenan-threne-1-carboxamide 21. The title compound **21** was prepared by reacting 4-(dimethylamino)benzylamine dihydrochloride with abietic acid (1.1 equiv.) in the presence of EDC HCl (1.1 equiv.), DIPEA (5 equiv.) and DMAP (5 mol%) in DCM (0.15 M) at room temperature overnight. The crude material was purified by column cromatography (Hexane/EtOAc 90:10 to 85:15).

Yield: 131 mg, 20%; amorphous solid; R_f = 0.4 (TLC Hexane/EtOAc 8:2). ^{1}H NMR (700 MHz, DMSO-*d*$_6$): δ 7.89 (t, *J* = 5.7 Hz, 1H), 7.05-7.00 (m, 2H), 6.68-6.63 (m, 2H), 5.71 (s, 1H), 5.30 (d, *J* = 5.1 Hz, 1H), 4.18-4.08 (m, 2H), 2.85 (s, 6H), 2.25-2.16 (m, 1H), 2.10-1.36 (m, 12H), 1.17 (s, 3H), 1.15-1.08 (m, 2H), 0.97 (dd, *J* = 6.8, 1.5 Hz, 6H), 0.76 (s, 3H).^{13}C NMR (176 MHz, DMSO-*d*$_6$): δ 177.7, 149.9, 144.7, 135.3, 128.4, 128.3, 122.9, 121.2, 112.8, 50.9, 45.8, 45.1, 42.4, 40.8, 38.2, 37.6, 34.7, 34.6, 27.4, 25.3, 22.4, 21.8, 21.2, 18.4, 17.4. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{29}H_{43}N_2O^+$ = 435.33699; found: 435.3381.

***N*-(4-(dimethylamino)benzyl)-5-(2,5-dimethylphenoxy)-2,2-dimethylpentanamide 22.** The title compound **22** was prepared by reacting 4-(dimethylamino)benzylamine dihydrochloride with gemfibrozil (1.1 equiv.) in the presence of EDC HCl (1.1 equiv.), DIPEA (5 equiv.) and DMAP (5 mol%) in DMF (0.6 M) at room temperature overnight. The crude material was diluted with brine and the product extracted with ethyl acetate (x5). The organic phase was further washed with brine solution (x5), dried over sodium sulfate, filtered, and evaporated under reduced pressure. The resulting mixture was purified by column cromatography (Hexane/EtOAc 90:10 to 80:20). Yield: 93 mg, 13.5%; amorphous solid; R_f = 0.8 (TLC Hexane/EtOAc 5:5). ^{1}H NMR (700 MHz, $CDCl_3$): δ 7.17-7.13 (m, 2H), 6.99 (d, *J* = 7.4 Hz, 1H), 6.73-6.68 (m, 2H), 6.66 (d, *J* = 7.5 Hz, 1H), 6.61 (s, 1H), 5.80 (brs, 1H, NH), 4.34 (d, *J* = 5.3 Hz, 2H), 3.91 (t, *J* = 6.1 Hz, 2H), 2.94 (s, 6H), 2.30 (s, 3H), 2.15 (s, 3H), 1.80-1.73 (m, 2H), 1.72-1.69 (m, 2H), 1.22 (s, 6H). ^{13}C NMR (176 MHz, $CDCl_3$): δ 177.1, 156.9, 136.5, 130.3, 128.9, 123.5, 120.7, 112.0, 68.0, 43.3, 41.9, 37.6, 25.6, 25.1, 21.4, 15.8. HRMS–ESI: *m/z* $[M + H]^+$ calcd for $C_{24}H_{35}N_2O_2^+$ = 383.5555; found: 383.5572.

NMR Spectroscopy. The samples for NMR spectroscopy were prepared by dissolving the appropriate amount of $[Ir(ppy)_2bpy]PF_6$ in 0.54 mL of CTAC and SDS (either deuterated or not) 2% aqueous solutions and 0.06 mL of

$^{2}H_2O$ (pH 5.0) to obtain a concentration 1.5 mM of $[Ir(ppy)_2bpy]PF_6$. NMR experiments were recorded on a Bruker Avance NEO 600 MHz spectrometer equipped with a z-gradient 5 mm triple-resonance probe head. For interaction studies between the micelles and $[Ir(ppy)_2bpy]PF_6$, DQF-COSY,[2,178] TOCSY,[3] and NOESY[4] spectra were recorded in the phase-sensitive mode using the method from States.[29] TOCSY experiments were acquired with a mixing time of 80 ms and a data block size of 4096 addresses in t2 and 256 equidistant t1 values. NOESY experiments were run with a mixing time of 500 ms. The proton, DQF-COSY, TOCSY, and NOESY spectra were processed and analyzed with the Bruker TOPSPIN 4.1.1 software packages.

Chapter4 - PROTEIN STRUCTURAL ANALYSIS WITH X-RAY

4.1 Investigation of antitoxin-mimicking peptides and the bacterial toxin Doc binding mode

Toxin inhibitors would potentially be able to significantly reduce persister formation, which in a cotreatment with antibiotic would allow for a more complete clearance of the bacterial infection and prevention of infection recurrence. The bacterial toxin Doc has been associated with the persistence of *Salmonella* in macrophages, enabling its survival upon antibiotic exposure. Toxin-antitoxin (TA) modules are promising yet underexplored targets for overcoming antibiotic failure. In literature, the molecular details are reported about the relationship of Doc with the antitoxin Phd, in particular residues have been discriminated that stabilize the toxin-antitoxin (TA) complex from those essential for inhibiting the activity of the toxin. However, the mechanisms by which the toxin neutralization is achieved is yet to be revealed. The goal of this part of my thesis is to achieve structural information about the bioactive conformation of Phd and his role on Doc toxin neutralization mechanisms.

4.1.1 Introduction

In the search of novel antimicrobial therapeutics, toxin-antitoxin (TA) modules are promising yet underexplored targets for overcoming antibiotic failure. The bacterial toxin Doc has been associated with the persistence of Salmonella in macrophages, enabling its survival upon antibiotic exposure. Before validation of Doc TA systems as therapeutic targets, detailed characterization of the interaction between toxins and antitoxins is required in order to develop toxin inhibitors. Toxin inhibitors would potentially be able to significantly reduce persister formation, which in a cotreatment with antibiotic would allow for a more complete clearance of the bacterial infection and prevention of infection recurrence. Previous studies[184] have provided insight into the molecular details of the Phd-Doc relationship, distinguishing antitoxin residues that stabilize the TA complex from those that are essential for inhibiting toxin activity. Specifically, by using synthetic peptides mimicking the Phd primary toxin binding-domain as chemical tools the role of specific residues and regions of the antitoxin on Doc toxin neutralization has been evaluated. Substitution of some residues led to poor or no inhibition of Doc, despite the formation of high-affinity complexes with picomolar dissociation constants, suggesting that toxin neutralization is achieved by mechanisms beyond high affinity interactions. Our goal is to achieve structural information about the interaction of Phd with Doc through X-ray crystallography in order to have more information about its bioactive conformation and role on Doc toxin neutralization mechanisms.

4.1.2 Results and Discussion

Doc-Phd Complex Expression. One of the major challenges of studying TA modules is the purification of active toxins.[185] In bacterial expression systems, toxin expression severely inhibits normal growth, resulting in production of only trace amounts of the toxin of interest. To overcome these issues, common strategies include inactivation of the toxin by site-directed mutagenesis or, in the case of type II TA systems (i.e. the antitoxin is a protein), coexpression with the antitoxin.[186,187] So we started with the coexpression of Doc with the C-terminal antitoxin Phd peptide, Phd$^{52\text{-}73}$, since the C-terminal domain of Phd (amino acids 52–73) is the minimal unit of Phd required for Doc inhibition (**Figure 48**).[184,188]

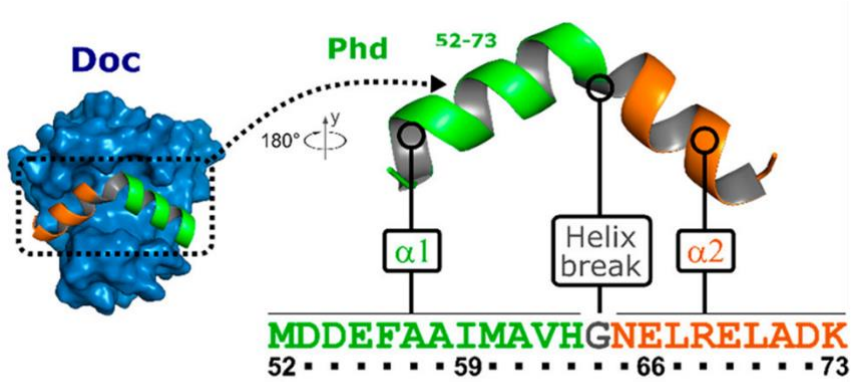


Figure 48. *Doc bound to the C-terminal domain of Phd. The sequence and bound conformation of the Phd C-terminal domain are shown on the right, two helices (α1 and α2, respectively, in green and orange) are separated by Gly64 (grey).*[184]

X-ray crystallography.

Once the Phd$^{52-73}$ -Doc complex has been purified, the sample has been concentrated to 5,5 mg/mL, the highest possible concentration without causing aggregation or precipitation. This is because crystallization requires bringing the macromolecule to supersaturation. After that, the sample was introduced in solution with precipitating agents to promote the nucleation of protein crystals, which resulted in three-dimensional crystals growing from the solution (**Figure 49**).

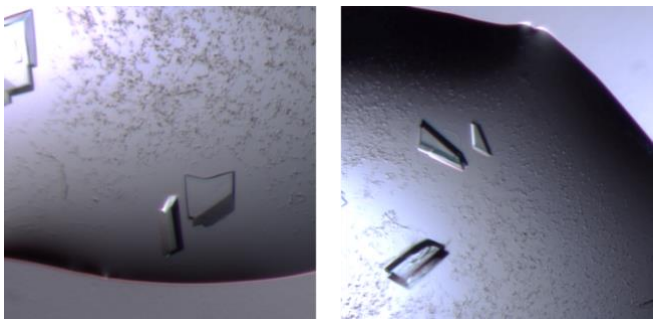


Figure 49. *Crystals formation in a 40 mM Potassium Phosphate Monobasic, 16% (w/v) PEG800, 20% (v/v) Glycerol solution.*

Using a special loop, we fished the crystals and sent it to the Diamond Light Source in Oxford to perform diffraction experiments. After the diffraction data collection and the refinement of the built structure, we created the crystal structure of the complex (**Figure 50**).

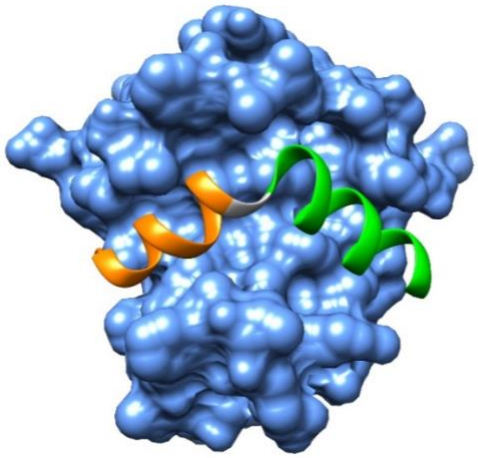


Figure 50. *Three-dimensional structure of Doc-Phd$^{52-73}$ complex.*

In a previous studies,[184] it was developed a homology model of *S.* Typhimurium Doc (Doc_{STm}) bound to C-terminal residues 52–73 of S. Typhimurium Phd (Phd_{STm}). The model was based on the published crystal structure (PDB: 3K33) [188] of Doc_{P1} in complex with Phd from *E. coli* P1 phage (Phd_{P1}). Our crystal structure validates the model, confirming the binding of the truncated version of Phd_{STm} but also the suggested α-helical conformation, where two α-helices (**Figure 50**, in green and orange) are separated by a structural "kink" at Gly64 (**Figure 50**, in gray). Furthermore, it emerged that replacing aspartate 53 of Phd with alanine resulted in a gain of affinity of more than 5-fold when compared to Phd^{52-73}. [184] The hypothesis is an increased affinity due to an enhanced helicity of the peptide, which favors the bound conformation. Actually, Ala possesses the highest helix propensity among all natural amino acids while Asp (deprotonated) is known as one of the poorest helix inducers. [189] Therefore, to test if there is a greater helicity for Phd-D53A compared with Phd^{52-73}, the same co-expression protocol was performed with Phd-D53A, in order to obtain the structure of the complex and to compare it to the Doc- Phd^{52-73} structure. Crystals haven't formed yet, but further tests are underway with different conditions to try to facilitate the process.

4.1.3 Conclusion

One of the major challenges in toxin-antitoxin module research is the purification of the toxin protein. Indeed, in bacterial expression systems, toxin expression severely inhibits normal growth and produces only trace amounts of the desired protein. The protocol of co-expression of Doc with Phd peptides allowed us to obtain sufficient amounts of complex to be used for X-ray crystallography. The obtaining of the crystallographic structure of the complex allowed us to confirm the binding mode of the Doc- Phd^{52-73} module, to date suggested only by in vitro models. Furthermore, once the Doc structure with the peptide Phd-D53A is also obtained, it will then be possible to proceed with the comparison of the two bonding modes. In fact, even if the specific residues and regions of the antitoxin that have a role on Doc toxin neutralization have been identified, the bioactive conformation and the bonding mode are yet to be revealed. Therefore, comparing the structures of the complex of Doc with Phd^{52-73} and Doc with peptides with poor or no inhibition of Doc, but forming high-affinity complexes with picomolar dissociation constants, can reveal new details in the mechanisms on Doc toxin neutralization.

4.1.4 Experimental Section

Expression and purification of proteins. His-tagged Doc protein was co-expressed with Phd$^{52-73}$ or Phd-D53A in BL21(DE3) cells with Kanamycin and Ampicillin resistance via heat shock, using pET28_Doc-His^{6} and pet14b_peptide vectors provided by Dr Anna Barnard of Faculty of Natural Sciences, Department of Chemistry, of Imperial College. 1L of BL21(DE3)-complex was grown in LB media at 37°C, 200rpm and induced to express overnight with 0.5 mM IPTG. The complex was purified using nickel-column affinity chromatography. BL21(DE3) cells were lysed in lysis buffer (20 mM Tris pH 8, 300 mM NaCl, protease inhibitor) and underwent a wash step and four elution steps: wash (20 mM Tris pH 8, 300 nM NaCl, 20 mM Imidazole), elution 1 (20 mM Tris pH 8, 300 nM NaCl, 40 mM Imidazole), elution 2 (20 mM Tris pH 8, 300 nM NaCl, 60 mM Imidazole), elution 3 (20 mM Tris pH 8, 300 nM NaCl, 80 mM Imidazole) and elution 4 (20 mM Tris pH 8, 300 nM NaCl, 200 mM Imidazole). Eluted Doc-Phd complexes were further purified by SEC through a Superdex 200 Increase 10/300 GL gel tube using a degassed buffer (20 mM Tris pH 8, 150 nM NaCl) and the AKTA pure SEC system.

X-ray crystallography. A Mosquito robot (TTP Labtech Ltd., Melbourn, UK) was used to deposit nL volumes of liquid to set up crystal screens at protein concentrations of 5,5 mg/mL. The phd-Doc complex was crystallized in 40 mM Potassium Phosphate Monobasic, 16% (w/v) PEG800. Glycerol (20%) was used as a cryoprotectant. Diffraction data were collected at the Diamond Light Source in Oxford. Coot[190] was used to fit the model to the electron density and the phenix.refine[191] programs were used for refinement. Figure panels containing crystal structures (**Figure 50**) were created using the UCSF Chimera program.[33]

SUPPLEMENTARY MATERIAL

Table S1. Docking statistics and energy terms ($\Delta G_{AD4.2}$) of **p6** with DNA1 and DNA2.

Complex	Clusters	Size[a]	DG_{bind}[b]	Electr[c]	H-Bond[c] + VdW[c] + Desolv[c]	Tors[c]
6/DNA1	24	10	-4.29	-5.07	-4.73	5.51
6/DNA2	29	11	-7.99	-5.09	-8.41	5.51

[a]Most populated cluster size. [b]DG_{bind}: free energy of binding. [c]Energy terms contributing to the AutoDock4.2 scoring function. Electr: electrostatic; H-Bond: H-Bonding; VdW: Van der Waals; Desolv: desolvatation; Tors: torsional entropy. All terms are given in Kcal/mol.

Table S2. [1]H NMR Resonance Assignments[a] of **p2** in water solution

residue	NH ($^3J_{aN}$, -Dd/DT)[b]	C[a]H	C[b]H	Others
Gln[1]		3.84	1.95, 1.99	2.21 (γ); 7.53,6.84(ε)
Phe[2]	8.78 (6.9, 10.1)	4.50	2.92	7.11 (δ); 7.19(ε); 7.16(ζ)
Asn[3]	8.44 (7.9, 4.9)	4.50	2.60, 2.50	6.82, 7.52(δ)
Leu[4]	8.25 (6.7, 8.2)	4.09	1.46	0.74, 0.79(δ)
Arg[5]	8.30 (7.3, 5.7)	4.16	1.61, 1.71	1.46 (γ); 3.00 (δ); 7.08(ε)
Gly[6]	8.22 (6.1, 5.2)	3.80		
Asn[7]	8.32 (7.1, 6.0)	4.50	2.61	7.54, 6.88 (δ)
Gln[8]	8.40 (6.9, 6.1)	4.06	1.71, 1.81	1.95, 2.02 (γ); 6.78, 7.37 (ε)
Phe[9]	8.13 (7.5, 5.6)	4.42	2.84, 3.00	7.08 (δ); 7.19 (ε), 7.14(ζ)
Asn[10]	8.20 (7.9, 4.3)	4.48	2.52, 2.65	6.82, 7.53 (δ)
Leu[11]	8.16 (6.8, 6.9)	4.12	1.49	0.73, 0.79 (δ)
Arg[12]	8.20 (7.5, 5.8)	4.12	1.61, 1.72	1.46(γ); 3.04 (δ) 7.09 (ε)

[a] Obtained at 280 K, pH = 5.5, with TSP (δ 0.00 ppm) as reference shift. Chemical shifts are accurate to ±0.02 ppm.

[b] $^3J_{aN}$ coupling constants in Hz. -Dd/DT = temperature coefficients (ppb/K) calculated in the range 280-298 K. Further signal of C-terminal amide: 6.78, 7.37 ppm.

Table S3. ^{1}H NMR Resonance Assignments[a] of **p3** in water solution

residue	NH ($^{3}J_{aN}$, -Dd/DT)[b]	C^{a}H	C^{b}H	Others
Gln[1]		3.99	2.09, 2.12	2.36 (γ); 7.00,7.67 (ε)
Phe[2]	8.92 (7.9, 9.2)	4.68	3.03, 3.12	7.26 (δ); 7.36 (ε)
Asn[3]	8.66 (7.9, 4.9)	4.68	2.64, 2.76	6.95, 7.66 (δ)
Leu[4]	8.28 (7.2, 8.3)	4.32	1.64	1.45 (γ); 0.86, 0.92 (δ)
Arg[5]	8.45 (8.7, 5.5)	4.75	1.68, 1.80	1.58, 1.62 (γ); 3.18 (δ); 7.24 (ε)
DPro[6]		4.72	2.32	1.97, 2.04 (γ); 3.67 (δ)
Pro[7]		4.37	2.25	1.89, 2.05 (γ);3.68, 3.87 (δ)
Gln[8]	8.28 (7.0, 5.8)	4.51	1.94, 2.08	2.32 (γ); 6.90, 7.45 (ε)
Phe[9]	8.16 (7.7, 6.3)	4.50	2.92, 3.09	7.23 (δ); 7.34 (ε)
Asn[10]	8.41 (7.5, 4.4)	4.59	2.65, 2,78	6.97, 7.68 (δ)
Leu[11]	8.39 (7.4, 7.2)	4.26	1.61	0.89, 0.95 (δ)
Arg[12]	8.52 (8.6, 5.9)	4.27	1.62	1.50(γ); 3.11 (δ); 7.23 (ε)

[a] Obtained at 280 K, pH = 5.5, with TSP (δ 0.00 ppm) as reference shift. Chemical shifts are accurate to ±0.02 ppm.

[b] $^{3}J_{aN}$ coupling constants in Hz. -Dd/DT = temperature coefficients (ppb/K) calculated in the range 280-298 K. Further signal of C-terminal amide: 7.22, 7.54 ppm.

Table S4. ^{1}H NMR Resonance Assignments[a] of **p4** in water solution

residue	NH ($^{3}J_{aN}$, -Dd/DT)[b]	C^{a}H	C^{b}H	Others
Gln[1]		3.98	2.09	2.35(γ); 7.25, 7.68 (ε)
Phe[2]	8.91 (6.9, 8.2)	4.65	3.06	7.25(δ); 7.34(ε)
Asn[3]	8.58 (7.9, 5.4)	4.63	2.62, 2.73	6.96, 7.65(δ)
Leu[4]	8.30 (7.3, 9.3)	4.27	1.58	1.53 (γ); 0.84, 0.91(δ)
Arg[5]	8.42 (7.1, 5.1)	4.56	1.66, 1.78	1.61 (γ); 2.94, 3.11 (δ); 7.21(ε)
Pro[6]		4.67	1.86, 2.31	1.99, 2.06 (γ); 3.59, 3.79 (δ)
DPro[7]		4.39	1.89, 2.19	1.72, 1.97 (γ); 3.59, 3.88 (δ)
Gln[8]	7.80 (8.3, 3.9)	4.30	1.90, 2.05	2.15, 2.22 (γ); 6.92, 7.62 (ε)
Phe[9]	8.44 (7.8, 8.5)	4.63	2.95, 3.10	7.22(δ); 7.32 (ε)
Asn[10]	8.57 (7.8, 5.6)	4.66	2.67, 2.78	6.96, 7.69(δ)
Leu[11]	8.43 (7.1, 8.2)	4.27	1.59	0.88, 0.94(δ)
Arg[12]	8.37 (7.6, 6.0)	4.27	1.76, 1.86	1.62 (γ); 3.18(δ)7.24 (ε)

[a] Obtained at 280 K, pH = 5.5, with TSP (δ 0.00 ppm) as reference shift. Chemical shifts are accurate to ±0.02 ppm.

[b] $^{3}J_{aN}$ coupling constants in Hz. -Dd/DT = temperature coefficients (ppb/K) calculated in the range 280-298 K. Further signal of C-terminal amide: 7.22, 7.52 ppm.

Table S5. [1]H NMR Resonance Assignments[a] of **p6** in water solution

residue	NH ($^3J_{aN}$, -Dd/DT)[b]	C[a]H	C[b]H	Others
Gln[1]		3.81	1.91	2,18(g); 6.80, 7.50(ε)
Trp[2]	8.83 (6.4, 8.1)	4.58	3.00	6.95(δ); 7.38, 9.94 (ε);6.97, 7.26 (z); 7.04 (η)
Asn[3]	8.33 (bs, 10.6)	4.52	2.36	7.38, 6.70(δ)
Trp[4]	8.00 (bs, 12.1)	4.21	2.81, 2.90	7.00(δ); 7.31,9.97(ε); 6.96, 7.26 (ζ); 7.02 (η)
Arg[5]	7.88 (7.7, 4.4)	3.87	1.25, 1.40	0.94(g); 2.67(δ); 6.83 (ε)
Gly[6]	7.60 (5.9, 7.8)	3.49, 3.62		
Asn[7]	8.17 (7.4, 4.9)	4.33	2.55	6.82, 7.46 (δ)
Gln[8]	8.12 (bs, 4.6)	4.00	1.66, 1.72	1.89(g); 6.71, 7.28(ε)
Trp[9]	8.07 (6.9, 8.2)	4.45	2.92, 3.01	6.94 (δ); 7.30, 9.91(ε); 6.95, 7.25 (ζ); 7.03 (η)
Asn[10]	8.12 (7.8, 9.4)	4.50	2.33	6.71, 7.32(δ)
Trp[11]	7.95 (bs, 11.1)	4.20	2.89	6.97 (δ); 7.38, 9.96(ε); 7.00,7.21 (ζ); 7.05 (η)
Arg[12]	7.61(8.2, 1.1)	3.84	1.22, 1.41	0.95 (g); 2.78 (δ), 6.89 (ε)

[a] Obtained at 280 K, pH = 5.5, with TSP (δ 0.00 ppm) as reference shift. Chemical shifts are accurate to ±0.02 ppm.

[b] $^3J_{aN}$ coupling constants in Hz. bs: broad signal. -Dd/DT = temperature coefficients (ppb/K) calculated in the range 280-298 K. Further signal of C-terminal amide: 6.78, 6.50 ppm.

Table S6. ^{1}H NMR Resonance Assignments[a] of **p8** in water solution

residue	NH ($^{3}J_{aN}$, -Dd/DT)[b]	C^{a}H	C^{b}H	Others
Gln^{1}		3.90	1.97, 2,01	2,25(g); 6.78, 7.45 (ε)
$\bar{T}$rp^{2}	8.67 (5.8, 9.7)	4.60	3.10, 3.13	7.05(δ); 7.53, 10.0 (ε); 7.06, 7.37 (z); 7.14 (η)
Asn^{3}	8.22 (6.9, 3.7)	4.51	2.38, 2.42	7.45, 6.83(δ)
Trp^{4}	7.92(7.5, 6.3)	4.24	3.00	7.06 (δ); 7.37, 10.0(ε); 6.99, 7.33 (ζ); 7.11 (η)
Arg^{5}	7.56 (7.2, 3.5)	4.23	1.24, 1.46	1.23(g); 2.77(δ); 6.90 (ε)
Pro^{6}		4.05	1.69, 2.14	1.88(g); 3.33, 3.38(δ)
DPro^{7}		4.29	1.82, 2.08	1.65, 1.85(g); 3.43, 3.60 (δ)
Gln^{8}	7.81 (8.2, 6.6)	4.28	1.82, 1.97	2.07(g); 6.82 ,7.46(ε)
Trp^{9}	8.27 (7.1, 10.1)	4.50	3.06	7.03(δ); 7.48, 10.0(ε); 7.04, 7.40(ζ); 7.16(η)
Asn^{10}	8.09 (7.9, 3.5)	4.41	2.26, 2.32	6.75, 7.33 (δ)
Trp^{11}	7.77 (6.9, 6.3)	4.24	3.02, 3.13	7.14(δ); 7.42, 10.1(ε); 7.02, 7.36(ζ); 7.13(η)
Arg^{12}	7.70 (8.1, 2.0)	3.88	1.29, 1.51	0.98 (g), 2.80 (δ), 6.96 (ε)

[a] Obtained at 280 K, pH = 5.5, with TSP (δ 0.00 ppm) as reference shift. Chemical shifts are accurate to ±0.02 ppm.

[b] $^{3}J_{aN}$ coupling constants in Hz. -Dd/DT = temperature coefficients (ppb/K) calculated in the range 280-298 K. Further signal of C-terminal amide: 6.93, 6.68 ppm.

Figure S1. In the case of a 1,4-triazolic bridge, the linear peptides were released from the solid support prior being cyclized in solution. The reaction was carried out in an hydroalcoholic medium using $CuSO_4*10H_2O$ and ascorbic acid as catalytic tandem and monitoring the conversion from linear to cyclic status by retention time shifts. Despite this synthetic approach was straightforward for most of the peptides, it did not work for those containing an azidoalanine (Aza) as first amino acid (**4** and **8**), for which the LC-MS analysis of the linear peptides revealed the presence of two main products ascribable to the replacement of azido group by the piperidine (data not shown). We reasoned that this pitfall could be triggered by piperidine only when Aza amino acid is anchored as tritylic ester and might evolve through a two-steps process: an initial elimination of the azido group in beta position followed by a nucleophilic attack of the piperidine on the so-formed dehydroalanine. To overcome this hurdle, we set a modified SPPS strategy conceptually distinctive from the previous described. Specifically, histidine was loaded as first amino acid, while the 1,4-disubstituted [1,2,3]-triazole was built via CuAAC reaction between the propargylalanine (Pra) already incorporated on the 2CTC resin-bound hexapeptide and the Fmoc-protected azidoalanine using copper iodide as catalyst. (FigureS1).[192] Finally, after Fmoc-deprotection, the triazole-branched peptide was released from the solid support by a mild acidic cleavage and then cyclized through a lactamization using PyAOP as coupling reagent. Final removal of residual protective groups on amino acid side chains, afforded the desired crude peptides **4** and **8** which were then purified by preparative HPLC. The synthesis of 1,5-triazolyl-bridged peptides (**20**-**27**) was carried out according to the FigureS1. The linear peptides were assembled out following the US-SPPS procedure and the final cyclization was attained on the resin-bound oligomer. The RuAAC was performed in a microwave reactor (T = 65 °C) in anhydrous DMF using $Cp*RuCl(PPh_3)_2$ as catalyst and the conversion was monitored by retention time shifts.[193] Similarly to what observed for the 1,4-series, once again the presence of Aza as first resin-bound amino acid (**20** and **22**) was not tolerated. Hence, we proceeded with the same approach applied for **4** and **8**, but unfortunately it did not succeed for peptides **20** and **22** probably as consequence of the harsher conditions required for the RuAAC with respect to CuAAC, which produced a complex mixture of side products.

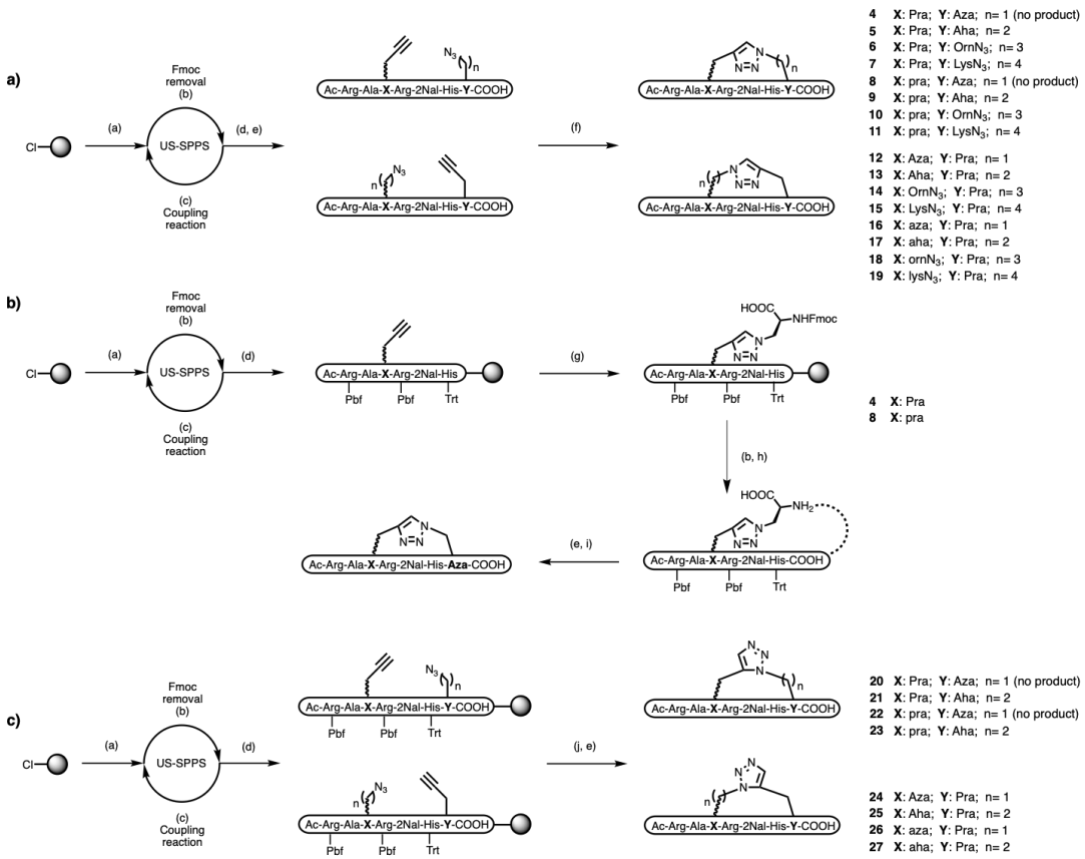


Figure S1. a) Fmoc-AA-OH, DIPEA, DMF dry, overnight; b) piperidine 20% in DMF, 2 x 1 min, US irradiation; c) Fmoc-AA-OH, HBTU, HOBt, DIPEA, DMF, 5 min, US irradiation; d) Ac_2O, DIPEA, DMF, 10 min; e) TFA, TIS (95:5), 3 h; f) $CuSO_4*10H_2O$, ascorbic acid, tBuOH/H_2O (1,5:1); g) Fmoc-Aza-OH, CuI, ascorbic acid, 2,4,6-collidine, DIPEA, DMF, 16h; h) DCM, TFE, acetic acid (8:1:1), 3 h; i) PyAOP, HOAt, DIPEA, DMF, 16 h; j) $Cp*RuCl(PPh_3)_2$, DMF dry, Ar atm, 65 °C, 150 min, microwave irradiation.

Table S7. NMR Resonance Assignments[a] of Peptide **12**.

residue	NH ($^3J_{\alpha N}$, $-\Delta\delta/\Delta T$)[b]	C^aH	C^bH	others
Arg[1]	8.12 (bs, 6.3)	4.27	1.79, 1.85	1.68(g); 3.22(d); 7.19(e)
Ala[2]	8.37 (bs, 6.5)	4.38	1.34	
Aza[3*]	7.92 (bs, 4.1)	4.99	4.93, 5.09	
Arg[4]	8.43 (ol, 3.9)	4.00	1.13, 1.36	0.46, 0.66(g); 2.26(d); 6.53(e)
2Nal[5**]	8.20 (7.1, 4.0)	5.01	3.45, 3.60	7.57, 7.83(d); 7.40, 7.84 (e); 7.36, 7.81 (z), 7,83(η)
His[6]	8.44 (ol, 7.0)	4.54	3.32, 3.48	7.36(d); 8.61(e)
Pra[7***]	8.25 (bs, 7.8)	4.65	3.25, 3.36	7,81 (d)

[a] Obtained at pH = 5.5, with TSP (δ 0.00 ppm) as reference shift. Chemical shifts are accurate to ±0.02 ppm.

[b] $^3J_{\alpha N}$ coupling constants in Hz.

$\Delta\delta/\Delta T$ = temperature coefficients (ppb/K) calculated in the range 298-313 K. Further signals: CH_3CO, 2.04 ppm.

ol: overlapped signal

bs: broad signal

* azidoalanine

** 2-naphthylalanine

*** propargylalanine

Table S8. NMR Resonance Assignments[a] of Peptide **21.**

residue	NH ($^3J_{\alpha N}$, $-\Delta\delta/\Delta T$)[b]	$C^{\alpha}H$	$C^{\beta}H$	Others
Arg[1]	8.15 (5.8, 5.2)	4.20	1.80, 1.87	1.69(g); 3.20(δ); 7.12(ε)
Ala[2]	8.35 (6.2, 5.7)	4.26	1.37	
Pra[3*]	8.37 (6.5, 5.3)	4.62	3.27, 3.53	7.75(δ)
Arg[4]	8.11 (5.8, 4.1)	3.86	1.07, 1.26	0.71, 0.79(g); 2.24, 2.40(δ);
2Nal[5**]	8.08 (6.3, 5.9)	4.80	3.07, 3.49	7.48, 7.79(δ); 7.38, 7.81(ε); 7.78, 7.79 (z), 7.80(η)
His[6]	8.19 (5.8, 4.0)	4.63	3.15, 3.34	7.32(δ); 8.56(ε)
Aha[7***]	8.07 (6.9, 2.3)	4.63	3,33	3,16 (g)

[a] Obtained at pH = 5.5, with TSP (δ 0.00 ppm) as reference shift. Chemical shifts are accurate to ±0.02 ppm.

[b] $^3J_{\alpha N}$ coupling constants in Hz.

$-\Delta\delta/\Delta T$ = temperature coefficients (ppb/K) calculated in the range 298-313 K. Further signals: CH_3CO, 2.05 ppm.

* propargylalanine

** 2-naphthylalanine

*** azidohomoalanine

Table S9. NOE derived Upper Limit Constraints of Peptide **12**.

0	ACE	QH	1 ARG HN	3.43
0	ACE	QH	1 ARG HA	5.57
0	ACE	QH	2 ALA HN	6.12
1	ARG	HN	1 ARG HB2	3.14
1	ARG	HN	1 ARG HB3	3.14
1	ARG	HN	1 ARG QB	2.72
1	ARG	HN	1 ARG QG	4.98
1	ARG	HN	1 ARG QD	6.38
1	ARG	HN	2 ALA HN	3.45
1	ARG	HN	3 AZA HN	4.26
1	ARG	HA	1 ARG HB2	3.02
1	ARG	HA	1 ARG HB3	3.02
1	ARG	HA	2 ALA HN	3.60
1	ARG	HA	3 AZA HN	4.47
1	ARG	HB2	1 ARG HE	4.51
1	ARG	HB2	2 ALA HN	3.42
1	ARG	HB3	1 ARG HE	4.51
1	ARG	HB3	2 ALA HN	3.42
1	ARG	QB	3 AZA HN	4.55
1	ARG	QG	2 ALA HN	5.79
1	ARG	QG	3 AZA HN	6.38
1	ARG	QD	2 ALA HN	6.38
2	ALA	HN	2 ALA HA	2.83
2	ALA	HN	2 ALA QB	3.49
2	ALA	HN	3 AZA HN	3.33
2	ALA	HA	3 AZA HN	2.79
2	ALA	HA	4 ARG HN	4.54

2	ALA	QB	3	AZA	HN	4.40
3	AZA	HN	3	AZA	HA	2.90
3	AZA	HN	3	AZA	HB1	3.27
3	AZA	HN	3	AZA	HB2	3.45
3	AZA	HN	4	ARG	HN	3.73
3	AZA	HN	5	BNA	HN	3.92
3	AZA	HA	4	ARG	HN	3.62
3	AZA	HB1	4	ARG	HN	3.24
3	AZA	HB1	5	BNA	HN	4.03
3	AZA	HB1	7	PRA	HD2	3.79
3	AZA	HB2	4	ARG	HN	3.08
3	AZA	HB2	5	BNA	HN	3.64
4	ARG	HN	4	ARG	HB2	3.30
4	ARG	HN	4	ARG	HB3	3.30
4	ARG	HN	4	ARG	QB	2.86
4	ARG	HN	4	ARG	QG	4.16
4	ARG	HN	4	ARG	QD	6.38
4	ARG	HN	5	BNA	HN	2.93
4	ARG	HA	4	ARG	HB2	2.93
4	ARG	HA	4	ARG	HB3	2.93
4	ARG	HA	4	ARG	QG	3.70
4	ARG	HA	4	ARG	QD	5.79
4	ARG	HA	4	ARG	HE	5.50
4	ARG	HA	5	BNA	HN	3.41
4	ARG	HA	5	BNA	HD2	5.00
4	ARG	HB2	4	ARG	HE	5.13
4	ARG	HB2	5	BNA	HN	3.90
4	ARG	HB2	5	BNA	HD2	4.42

4	ARG	HB3	4 ARG	HE 5.13
4	ARG	HB3	5 BNA	HN 3.90
4	ARG	HB3	5 BNA	HD2 4.42
4	ARG	QB	5 BNA	HD2 3.78
4	ARG	HG2	5 BNA	HD1 5.50
4	ARG	HG3	5 BNA	HD1 5.50
4	ARG	QG	5 BNA	HN 4.62
4	ARG	QG	5 BNA	HD2 4.05
4	ARG	QD	5 BNA	HZ3 6.38
4	ARG	QD	5 BNA	HE4 6.38
4	ARG	QD	5 BNA	HD2 6.10
4	ARG	HE	5 BNA	HZ3 5.50
4	ARG	HE	5 BNA	HE4 5.50
4	ARG	HE	5 BNA	HD2 5.41
5	BNA	HN	5 BNA	HA 2.99
5	BNA	HN	5 BNA	HB2 3.47
5	BNA	HN	5 BNA	HB3 3.99
5	BNA	HN	5 BNA	HD2 2.86
5	BNA	HA	6 HIS	HD2 4.94
5	BNA	HA	7 PRA	HN 3.98
5	BNA	HB2	5 BNA	HD1 3.79
5	BNA	HB2	5 BNA	HD2 2.80
5	BNA	HB3	5 BNA	HD1 3.52
5	BNA	HB3	5 BNA	HD2 3.77
5	BNA	HB3	6 HIS	HN 3.33
5	BNA	HB3	6 HIS	HD2 5.16
5	BNA	HB3	7 PRA	HN 4.20
6	HIS	HN	6 HIS	HB2 3.67

```
6 HIS  HN      6 HIS  HB3     3.36
6 HIS  HN      6 HIS  HD2     4.66
6 HIS  HN      7 PRA  HN      3.42
6 HIS  HA      6 HIS  HB2     2.86
6 HIS  HA      6 HIS  HB3     2.90
6 HIS  HA      6 HIS  HD2     3.21
6 HIS  HA      7 PRA  HN      2.99
6 HIS  HB2     6 HIS  HD2     3.89
6 HIS  HB3     6 HIS  HD2     3.73
7 PRA  HN      7 PRA  HB2     3.42
7 PRA  HN      7 PRA  HB3     3.42
7 PRA  HA      7 PRA  HB2     2.93
7 PRA  HA      7 PRA  HB3     2.93
```

ACE: acetyl group

BNA: β-naphthylalanine (2-naphthylalanine)

Table 10. NOE derived Upper Limit Constraints of Peptide **21**.

0	ACE	QH	1 ARG HN	3.43
0	ACE	QH	1 ARG HA	4.92
0	ACE	QH	2 ALA HN	5.07
0	ACE	QH	3 PRA HN	6.12
1	ARG	HN	1 ARG HA	2.86
1	ARG	HN	1 ARG HB2	3.14
1	ARG	HN	1 ARG HB3	3.14
1	ARG	HN	1 ARG QB	2.63
1	ARG	HN	1 ARG QG	4.42
1	ARG	HN	1 ARG QD	6.38
1	ARG	HN	1 ARG HE	5.50
1	ARG	HN	2 ALA HN	3.05
1	ARG	HA	1 ARG QB	2.65
1	ARG	HA	1 ARG QD	6.25
1	ARG	HA	2 ALA HN	3.50
1	ARG	HA	2 ALA QB	6.50
1	ARG	HA	3 PRA HN	3.36
1	ARG	HB2	1 ARG HE	4.11
1	ARG	HB2	2 ALA HN	3.48
1	ARG	HB3	1 ARG HE	4.11
1	ARG	HB3	2 ALA HN	3.48
1	ARG	QB	1 ARG HE	3.52
1	ARG	QB	2 ALA HN	3.04
1	ARG	QG	2 ALA HN	6.19
1	ARG	QG	3 PRA HN	6.38
2	ALA	HN	2 ALA HA	2.77
2	ALA	HN	2 ALA QB	3.61

```
2 ALA  HA      4 ARG  HN      4.32
2 ALA  QB      3 PRA  HN      4.36
2 ALA  QB      4 ARG  HN      5.23
2 ALA  QB      4 ARG  HA      6.53
2 ALA  QB      4 ARG  QB      5.98
2 ALA  QB      4 ARG  QG      5.13
2 ALA  QB      4 ARG  HE      6.00
3 PRA  HN      3 PRA  HA      2.65
3 PRA  HN      3 PRA  HB2     3.61
3 PRA  HN      3 PRA  HB3     3.61
3 PRA  HN      3 PRA  QB      3.04
3 PRA  HN      4 ARG  HN      3.02
3 PRA  HN      5 BNA  HN      4.07
3 PRA  HA      3 PRA  QB      2.65
3 PRA  HA      4 ARG  HN      3.47
3 PRA  HB2     4 ARG  HN      4.14
3 PRA  HB3     4 ARG  HN      4.14
3 PRA  HD2     6 HIS  HN      5.50
3 PRA  ND1     7 AHA  CG      4.00
4 ARG  HN      4 ARG  HA      2.90
4 ARG  HN      4 ARG  HB2     3.05
4 ARG  HN      4 ARG  HB3     3.05
4 ARG  HN      4 ARG  QB      2.64
4 ARG  HN      4 ARG  HG2     4.11
4 ARG  HN      4 ARG  HG3     4.11
4 ARG  HN      4 ARG  QG      3.52
4 ARG  HA      4 ARG  HB2     2.93
4 ARG  HA      4 ARG  HB3     2.93
```

4	ARG	HA	4 ARG	HG2	4.01
4	ARG	HA	4 ARG	HG3	4.01
4	ARG	HA	4 ARG	HD2	5.16
4	ARG	HA	4 ARG	HD3	5.16
4	ARG	HA	4 ARG	QD	4.32
4	ARG	HA	5 BNA	HN	3.17
4	ARG	HA	5 BNA	HD2	5.34
4	ARG	HB2	4 ARG	HE	4.79
4	ARG	HB2	5 BNA	HN	3.64
4	ARG	HB2	5 BNA	HD2	4.17
4	ARG	HB3	4 ARG	HE	4.79
4	ARG	HB3	5 BNA	HN	3.64
4	ARG	HB3	5 BNA	HD2	4.17
4	ARG	QB	5 BNA	HN	3.17
4	ARG	QB	5 BNA	HD2	3.61
4	ARG	HG2	4 ARG	HE	3.95
4	ARG	HG2	5 BNA	HD2	4.79
4	ARG	HG3	4 ARG	HE	3.95
4	ARG	HG3	5 BNA	HD2	4.79
4	ARG	QG	4 ARG	HE	3.44
4	ARG	QG	5 BNA	HN	5.34
4	ARG	QG	5 BNA	HD2	4.06
4	ARG	QD	5 BNA	HD2	4.36
4	ARG	HE	5 BNA	HD2	4.85
5	BNA	HN	5 BNA	HA	2.90
5	BNA	HN	5 BNA	HB2	2.86
5	BNA	HN	5 BNA	HB3	3.92
5	BNA	HN	5 BNA	HD1	5.01

5	BNA	HN	5	BNA	HD2	2.65
5	BNA	HN	6	HIS	HD2	3.95
5	BNA	HA	6	HIS	HN	3.17
5	BNA	HB2	5	BNA	HD1	3.79
5	BNA	HB2	5	BNA	HD2	2.40
5	BNA	HB3	5	BNA	HD1	3.55
5	BNA	HB3	5	BNA	HD2	3.69
6	HIS	HN	6	HIS	HB2	4.14
6	HIS	HN	6	HIS	HB3	4.07
6	HIS	HN	7	AHA	HN	3.64
6	HIS	HA	6	HIS	HB3	2.83
6	HIS	HA	6	HIS	HD2	3.55
6	HIS	HB3	6	HIS	HD2	3.92
7	AHA	HN	7	AHA	HA	2.83
7	AHA	HN	7	AHA	QG	6.31

ACE: acetyl group

BNA: β-naphthylalanine (2-naphthylalanine)

Table S11. Calculated IC_{50} values of anti-[Asn^{641}(Glc)]FAN(635−655), anti-[Asn^{153}(Glc)]RTN4R(147-165), anti-[Asn^{168}(Glc)]OMGp(162-180) or anti-HMW1ct(Glc) protein IgG antibodies of MS 14 serum to peptides 4-6 and protein I(Glc). Values are reported as 95% confidence interval for the calculated mean IC50 ± the standard error (SEM).

Antigen coated	Inhibitor	IC_{50}(IgG)
[Asn^{641}(Glc)]FAN(635−655) (4)	[Asn^{641}(Glc)]FAN(635−655) (**4**)	$(4.35 \pm 0.06)\cdot10^{-7}$
	[Asn^{153}(Glc)]RTN4R(147-165) (**5**)	$(1.17 \pm 0.09)\cdot10^{-8}$
	[Asn^{168}(Glc)]OMGp(162-180) (**6**)	$(3.21 \pm 0.09)\cdot10^{-8}$
	HMW1ct(Glc) (**I(Glc)**)	$(1.68 \pm 0.08)\cdot10^{-8}$
[Asn^{153}(Glc)]RTN4R(147-165) (5)	[Asn^{641}(Glc)]FAN(635−655) (**4**)	$(1.04 \pm 0.33)\cdot10^{-7}$
	[Asn^{153}(Glc)]RTN4R(147-165) (**5**)	$(1.93 \pm 0.07)\cdot10^{-8}$
	[Asn^{168}(Glc)]OMGp(162-180) (**6**)	$(5.03 \pm 0.26)\cdot10^{-8}$
	HMW1ct(Glc) (**I(Glc)**)	$(7.38 \pm 0.13)\cdot10^{-8}$
[Asn^{168}(Glc)]OMGp(162-180) (6)	[Asn^{641}(Glc)]FAN(635−655) (**4**)	$(9.81 \pm 0.16)\cdot10^{-7}$
	[Asn^{153}(Glc)]RTN4R(147-165) (**5**)	$(4.10 \pm 0.11)\cdot10^{-8}$
	[Asn^{168}(Glc)]OMGp(162-180) (**6**)	$(4.01 \pm 0.15)\cdot10^{-8}$
	HMW1ct(Glc) (**I(Glc)**)	$(1.89 \pm 0.24)\cdot10^{-8}$

HMW1(Glc) (I(Glc))	[Asn^{153}(Glc)]RTN4R(147-165) (**5**)	$(2.73 \pm 0.17)\cdot 10^{-7}$
	[Asn^{168}(Glc)]OMGp(162-180) (**6**)	$(9.96 \pm 0.23)\cdot 10^{-7}$
	HMW1ct(Glc) (**I(Glc)**)	$(6.82 \pm 0.21)\cdot 10^{-9}$

Figure S2. Peptides were dissolved in buffer carbonate (pH=9.8) or PBS (pH 7.2) independently. Plates were also coated using buffer without peptide to evaluate the influence of peptide in the signals obtained. Polystyrene 96-well ELISA plates were coated with 100 mL/well of a 10 mg/mL solution of synthetic peptide antigens **1-6** diluted in tested buffers. After overnight incubation at 4 °C, plates were washed (3×) using washing buffer. Nonspecific binding sites were blocked with 100 mL/well of fetal bovine serum buffer (10% FBS in washing buffer) or 5% BSA buffer at room temperature for 1 h. Blocking buffer was removed, and plates were incubated overnight at 4 °C with buffer as blank, an expected positive and negative sera (diluted 1:100 in 10% FBS buffer or 2.5% BSA buffer, 100 mL/well). After three washes, plates were treated with 100 mL/well of anti-human IgG or IgM alkaline phosphatase-conjugated specific antibodies diluted in FBS buffer 1:3000 (IgG) and 1:200 (IgM) for all tested antigens. After 3 h of incubation at room temperature and washes (3×), 100 mL of substrate buffer (1 mg/ml *p*NPP, $MgCl_2$ 0.01M in carbonate buffer, pH 9.6) was added to each well. Colorimetric reaction was carried out adding 100 µl of substrate reaction solution (1 mg/ml *p*NPP, $MgCl_2$ 0.01M in carbonate buffer, pH 9.6) to each well and plates were read at 405 nm using a TECAN plate reader. After 30 min, the reaction was stopped with 1 M NaOH solution (50 mL/well) and the absorbance was read in a multichannel ELISA reader (Tecan Sunrise, Männedorf, Switzerland) at 405 nm. Antibody titer values were calculated as (mean Abs of serum duplicate) - (mean Abs of blank duplicate) representing graphically in Figures **S1**-**S3** the calculated mean values.

		coating buffer 1			coating buffer 2			coating buffer 1			coating buffer 2			
	<>	1	2	3	4	5	6	7	8	9	10	11	12	
peptide coated	A	blank	C+	C-	blank	C+	C-	blank	C+	C-	blank	C+	C-	blocking 1
	B	blank	C+	C-	blank	C+	C-	blank	C+	C-	blank	C+	C-	
	C	blank	C+	C-	blank	C+	C-	blank	C+	C-	blank	C+	C-	blocking 2
	D	blank	C+	C-	blank	C+	C-	blank	C+	C-	blank	C+	C-	
no peptide	E	blank	C+	C-	blank	C+	C-	blank	C+	C-	blank	C+	C-	blocking 1
	F	blank	C+	C-	blank	C+	C-	blank	C+	C-	blank	C+	C-	
	G	blank	C+	C-	blank	C+	C-	blank	C+	C-	blank	C+	C-	blocking 2
	H	blank	C+	C-	blank	C+	C-	blank	C+	C-	blank	C+	C-	
		IgG						IgM						

Figure S2. Samples distribution on the ELISA plate during the set-up of the coating conditions.

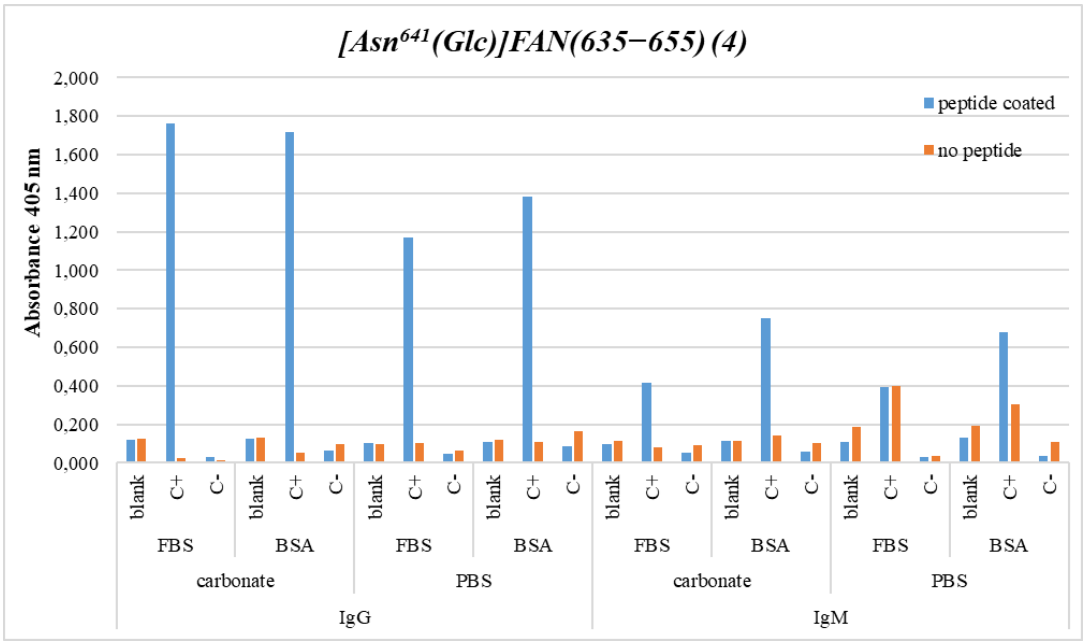


Figure S3. Coating results of [Asn^{641}(Glc)]FAN(635−655) (**4**).

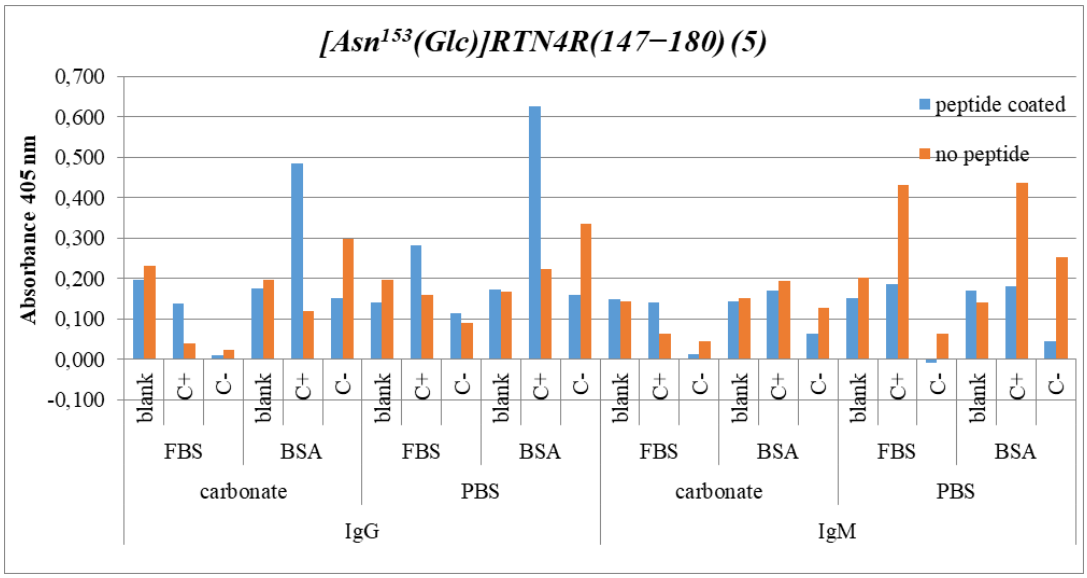


Figure S4. Coating results of [Asn^{153}(Glc)]RTN4R(147−165) (**5**).

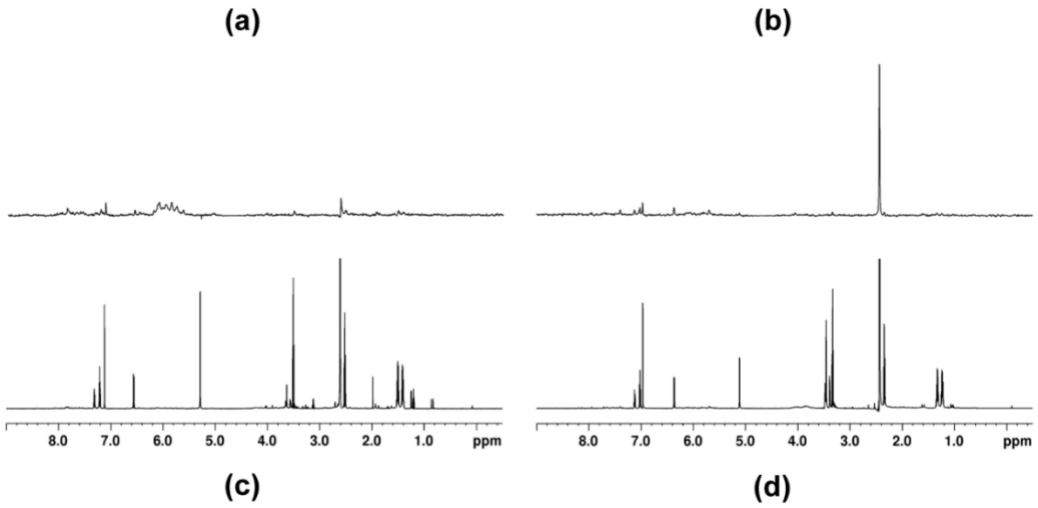
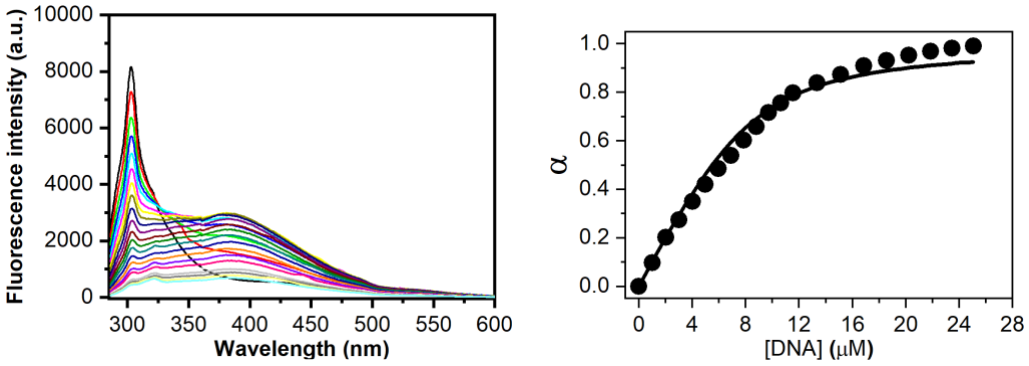
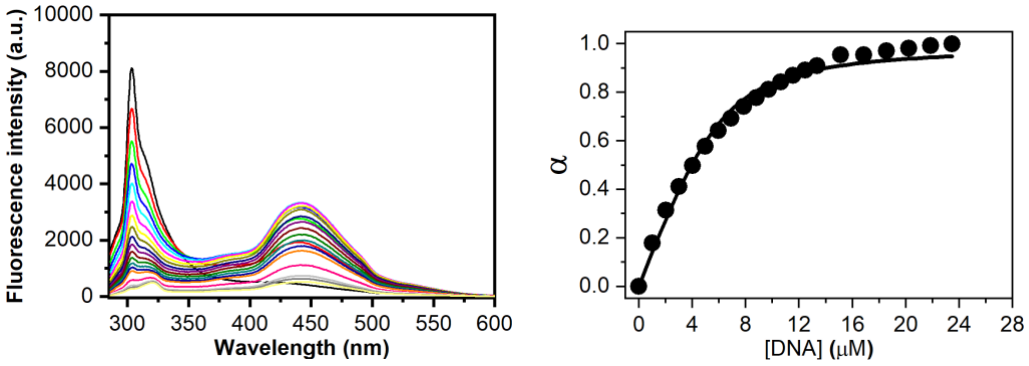


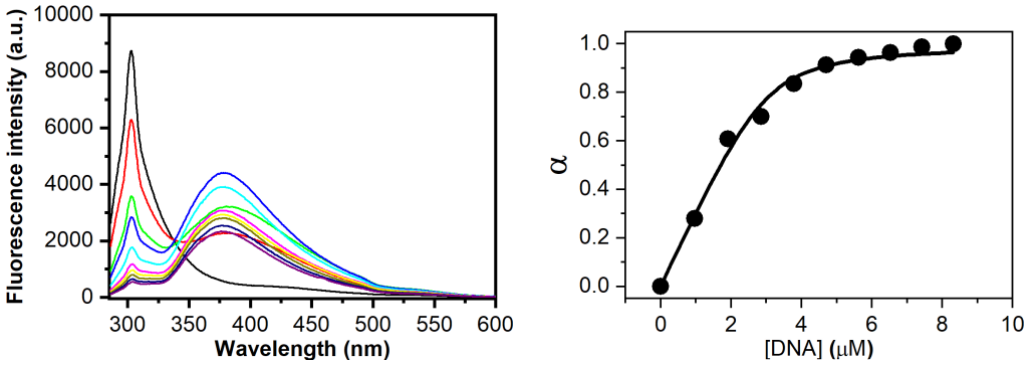
Figure S5. 1D ^{1}H NMR spectra (**c** and **d**) and STD spectra (**a** and **b**) of compound **21d** (0.250 mM) in presence of *c-Myc G4* (**a** and **c**) and *hTeloC iM* (**b** and **d**).



a)



b)



c)

Figure S6. Representative fluorescence emission spectra of (a,b) 21d, and (c) 18c (2.5–3.5 μM) in the absence and presence of stepwise additions (5 μL) of (a) c-Myc G4, (b) c-Myc iM, or (c) single-stranded DNA d(CT)15 (150–200 μM), at 25 °C. Insets show the titration curves obtained by plotting the fraction of bound compound (α) versus DNA concentration.

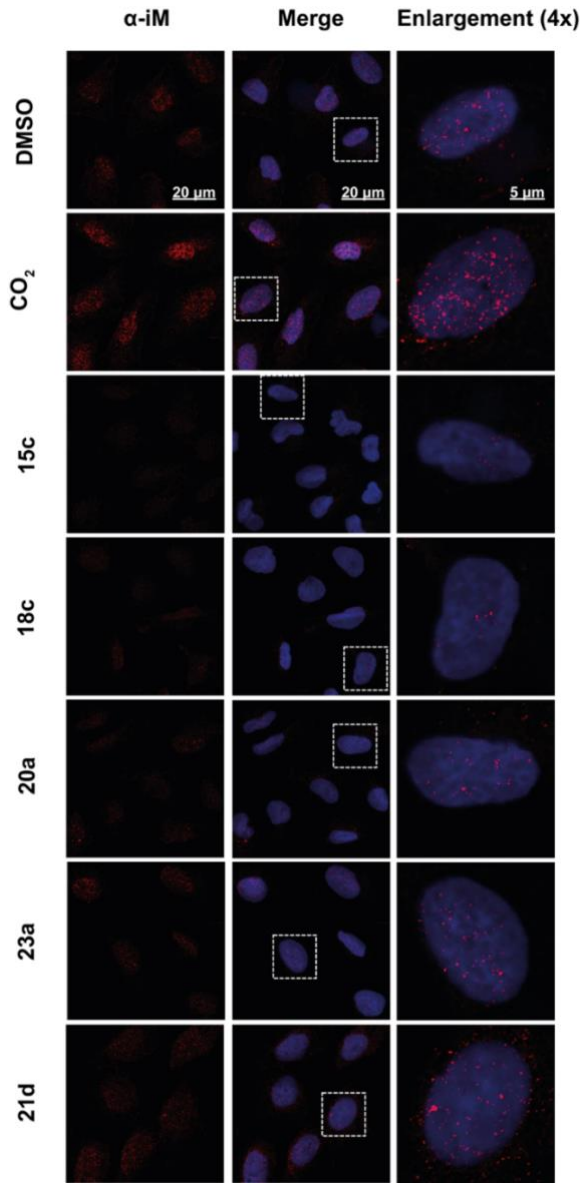


Figure S7. Biological evaluation of iM-destabilizing activity of the selected compounds. Immunofluorescence analysis of iM structures in U2OS cells treated for 24 h with 2 µM of the selected compounds or an equivalent amount of DMSO (negative control). As positive control, cells were maintained for 2.5 h in an atmosphere with 8% of CO_2. Representative images of confocal sections (63x) used to detect iM structures are shown. Left panels: iM structures (red) detected by anti-iM antibody (α-iM). Middle panels: merged images showing iM structures (red) and DAPI counterstained nuclei (blue). Right panels: 4x enlargements from the pictures in middle panels. The experiment was performed in triplicate and at least 9 fields/experiment were evaluated for each condition. Scale bars are reported in the figures.

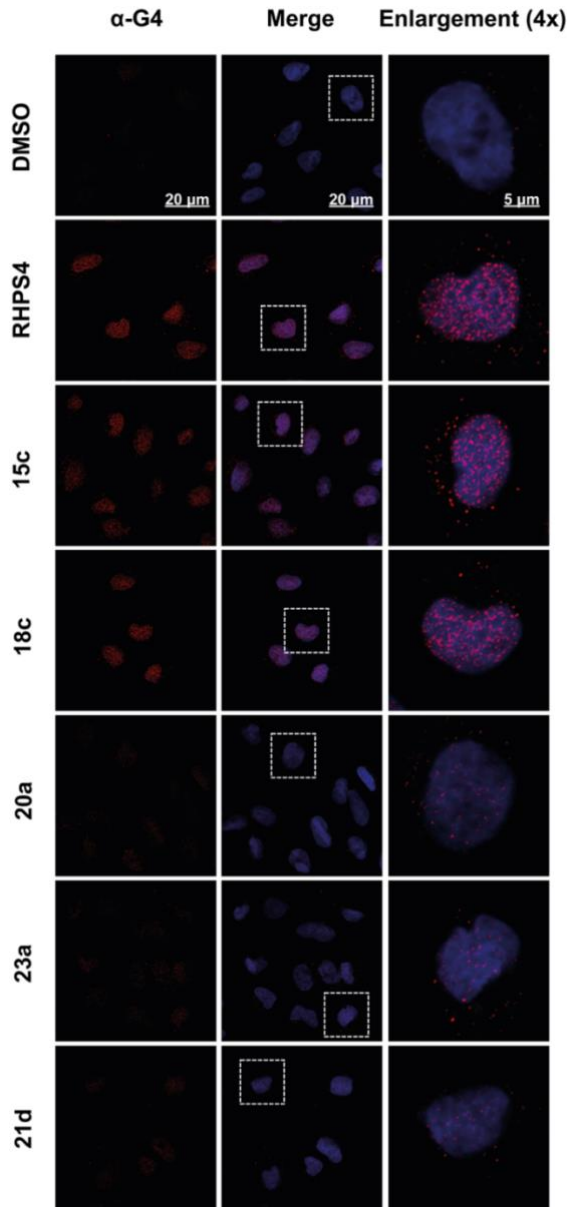


Figure S8. Biological evaluation of G4-stabilizing activity of the selected compounds. Immunofluorescence analysis of G4 structures in U2OS cells treated for 24 h with 2 µM of the selected compounds or an equivalent amount of DMSO (negative control). As positive control, cells were treated for 24 h with 1 µM RHPS4. Representative images of confocal sections (63x) used for detection of G4 structures are shown. Left panels: G4 structures (red) detected by anti-G4 antibody (α-G4). Middle panels: merged images showing G4 structures (red) and DAPI counterstained nuclei (blue). Right panels: 4x enlargements from the pictures in middle panels. The experiment was performed in triplicate and at least 9 fields/experiment were evaluated for each condition. Scale bars are reported in the figures.

Table S12. Optimization of reaction conditions

Entry	Eq. **1**	Eq. **2**	PC (mol%)	Solvent (0.15 M)	T	Watt	Yield %
1	1	1.5	$[Ir(ppy)_2bpy]PF_6$ (1)	MeCN/H_2O (10 eq)	rt	16	70 (litt 70)
2	1	1.5	$[Ir(ppy)_2bpy]PF_6$ (1)	2% TPGS-750-M	rt	16	30 [b]
3	1	1.5	$[Ir(ppy)_2bpy]PF_6$ (1)	2% TPGS-750-M	rt	30	40 [b]
4	1	1.5	$[Ir(ppy)_2bpy]PF_6$ (1)	H_2O	rt	30	20 [b]
5	1	1.5	$[Ir(ppy)_2bpy]PF_6$ (1)	2% TPGS-750-M	50°C	30	15 [b]
6	1	1.5	$[Ir(ppy)_2bpy]PF_6$ (1)	H_2O	50°C	30	12 [b]
7[a]	1	1.5	$[Ir(ppy)_2bpy]PF_6$ (1)	2% TPGS-750-M	rt	30	ND
8	2	1	$[Ir(ppy)_2bpy]PF_6$ (1)	2% TPGS-750-M	rt	30	68 (50[b])
9	1.5	1	$[Ir(ppy)_2bpy]PF_6$ (1)	2% TPGS-750-M	rt	30	38[b]
10	2	1	$[Ir(ppy)_2bpy]PF_6$ (1)	H_2O	rt	30	22[b]
11	2	1	*fac*-$Ir(ppy)_3$ (1)	2% TPGS-750-M	rt	30	25[b]
12	2	1	$Ru(bpy)_3(PF_6)_2$	2% TPGS-750-M	rt	30	6.5[b]
13	2	1	Rose Bengal (1)	2% TPGS-750-M	rt	30	20[b]
14	2	1	Eosin Y (1)	2% TPGS-750-M	rt	30	17[b]
15	2	1	$[Mes\text{-}Acr]BF_4$ (1)	2% TPGS-750-M	rt	30	7[b]
16	2	1	Eosin Y (5)	2% TPGS-750-M	rt	30	20[b]
17	2	1	$[Ir(ppy)_2bpy]PF_6$ (1)	1% TPGS-750-M	rt	30	12[b]
18	2	1	$[Ir(ppy)_2bpy]PF_6$ (1)	2% SPGS-550-M	rt	30	33[b]
19	2	1	$[Ir(ppy)_2bpy]PF_6$ (1)	5% TPGS-750-M	rt	30	32[b]
20	2	1	$[Ir(ppy)_2bpy]PF_6$ (1)	2% Triton X-100	rt	30	31[b]
21	2	1	Eosin Y (5)	2% Triton X-100	rt	30	30[b]
22	2	1	$[Ir(ppy)_2bpy]PF_6$ (1)	2% CTAC	rt	30	15[b]
23	2	1	Eosin Y (5)	2% CTAC	rt	30	30[b]
24	**2**	**1**	**$[Ir(ppy)_2bpy]PF_6$ (1)**	**2% SDS**	**rt**	**30**	**75(64[b])**
25	2	1	Eosin Y (5)	2% SDS	rt	30	18[b]
26	2	1	$[Ir(ppy)_2bpy]PF_6$ (0.5)	2% SDS	rt	30	25 [b]

[a]in presence of Na_2CO_3 (1 eq.); [b]NMR yield.

Table S13. NMR Resonance Assignments of $[Ir(ppy)_2bpy]PF_6$ in SDS 2% solution

Atom Chemical shifts

H1 7.92

H2 7.36

H3 8.12

H4 8.6

H7 8.6

H8 8.12

H9 7.36

H10 7.92

H12 6.25

H13 6.79

H14 6.92

H15 7.64

H18 7.88

H19 7.74

H20 6.97

H21 7.55

H23 6.25

H24 6.79

H25 6.92

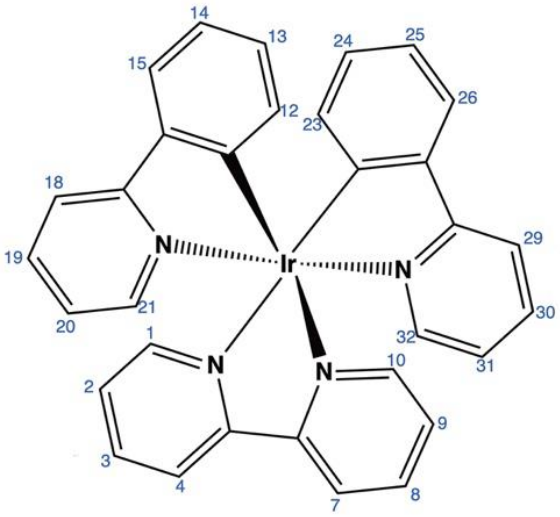
H26 7.64

H29 7.88

H30 7.74

H31 6.97

H32 7.55



Obtained at pH = 5.0, with TSP (δ 0.00 ppm) as reference shift. Chemical shifts are accurate to ±0.02 ppm.

Table S14. NMR Resonance assignments of [Ir(ppy)2bpy]PF6 in CTAC 2% solution

Atom Chemical shifts

H1 7.92

H2 7.39

H3 8.16

H4 8.68

H7 8.68

H8 8.16

H9 7.39

H10 7.92

H12 6.26

H13 6.8

H14 6.95

H15 7.68

H18 7.87

H19 7.73

H20 6.95

H21 7.57

H23 6.26

H24 6.8

H25 6.95

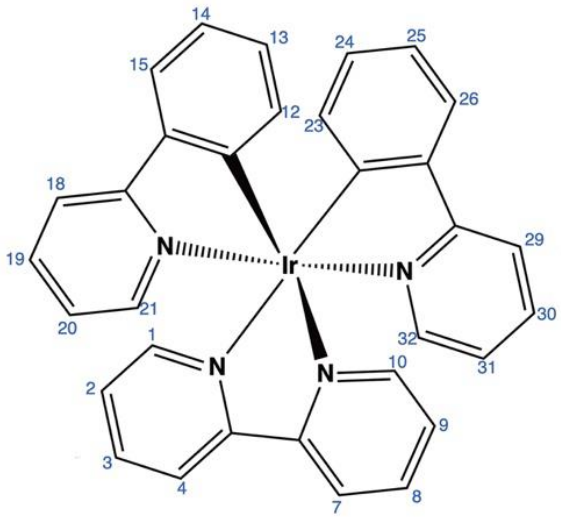
H26 7.68

H29 7.87

H30 7.73

H31 6.95

H32 7.57



Obtained at pH = 5.0, with TSP (δ 0.00 ppm) as reference shift. Chemical shifts are accurate to ±0.02 ppm.

References

(1) Wüthrich, K. NMR with Proteins and Nucleic Acids. *Europhys. News* **1986**, *17* (1), 11–13. https://doi.org/10.1051/EPN/19861701011.

(2) Piantini, U.; Sorensen, O. W.; Ernst, R. R. Multiple Quantum Filters for Elucidating NMR Coupling Networks. *Biochem. Biophys. Res. Commun* **1982**, *104* (3), 6801–6802.

(3) Braunschweiler, L.; Ernst, R. R. Coherence Transfer by Isotropic Mixing: Application to Proton Correlation Spectroscopy. *J. Magn. Reson.* **1983**, *53* (3), 521–528. https://doi.org/10.1016/0022-2364(83)90226-3.

(4) Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. Investigation of Exchange Processes by Two-dimensional NMR Spectroscopy. *J. Chem. Phys.* **2008**, *71* (11), 4546. https://doi.org/10.1063/1.438208.

(5) Mayer, M.; Meyer, B.; Park, K. C.; Meunier, S. J.; Zanini, D.; Roy, R.; Lett, C.; Romanowska, A.; Meyer, B.; Mayer, D.-C. M. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy**. ± *7386; c) W Chem. Soc. Chem. Commun* **1999**, *38* (12), 5907–5908. https://doi.org/10.1002/(SICI)1521-3773(19990614)38:12.

(6) Meyer, B.; Peters, T. NMR Spectroscopy Techniques for Screening and Identifying Ligand Binding to Protein Receptors. *Angewandte Chemie - International Edition*. February 24, 2003, pp 864–890. https://doi.org/10.1002/anie.200390233.

(7) Viegas, A.; Manso, J.; Nobrega, F. L.; Cabrita, E. J. Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for Ligand Screening and Characterization of Protein Binding. *J. Chem. Educ.* **2011**, *88* (7), 990–994. https://doi.org/10.1021/ED101169T/SUPPL_FILE/ED101169T_SI_002.DOC.

(8) Parigi, G.; Luchinat, C. Chapter 1: NMR Consequences of the Nucleus-

Electron Spin Interactions. In *New Developments in NMR*; Royal Society of Chemistry, 2018; Vol. 2018-January, pp 1–41. https://doi.org/10.1039/9781788013291-00001.

(9) Brown, L. R.; Bösch, C.; Wüthrich, K. Location and Orientation Relative to the Micelle Surface for Glucagon in Mixed Micelles with Dodecylphosphocholine: EPR and NMR Studies. *Biochim. Biophys. Acta* **1981**, *642* (2), 296–312. https://doi.org/10.1016/0005-2736(81)90447-8.

(10) Saviello, M. R.; Malfi, S.; Campiglia, P.; Cavalli, A.; Grieco, P.; Novellino, E.; Carotenuto, A. New Insight into the Mechanism of Action of the Temporin Antimicrobial Peptides. *49*, 1477–1485. https://doi.org/10.1021/bi902166d.

(11) Adler, A. J.; Greenfield, N. J.; Fasman, G. D. *[27] Circular Dichroism and Optical Rotatory Dispersion of Proteins and Polypeptides*.

(12) Wei, Y.; Thyparambil, A. A.; Latour, R. A. Protein Helical Structure Determination Using CD Spectroscopy for Solutions with Strong Background Absorbance from 190 to 230 Nm. *Biochim. Biophys. Acta - Proteins Proteomics* **2014**, *1844* (12), 2331–2337. https://doi.org/10.1016/J.BBAPAP.2014.10.001.

(13) Mayer, C.; Mayer, C. X-Ray Diffraction in Biology: How Can We See DNA and Proteins in Three Dimensions? *X-ray Scatt.* **2017**. https://doi.org/10.5772/64999.

(14) Spitz, F.; Furlong, E. E. M. Transcription Factors: From Enhancer Binding to Developmental Control. *Nat. Rev. Genet. 2012 139* **2012**, *13* (9), 613–626. https://doi.org/10.1038/nrg3207.

(15) Stefanucci, A.; Mosquera, J.; Vázquez, E.; Mascareñas, J. L.; Novellino, E.; Mollica, A. Synthesis, Characterization, and DNA Binding Profile of a Macrocyclic β-Sheet Analogue of ARC Protein. *ACS Med. Chem. Lett.* **2015**, *6* (12), 1220–1224. https://doi.org/10.1021/ACSMEDCHEMLETT.5B00363/ASSET/IMAGES/

LARGE/ML-2015-00363Y_0001.JPEG.

(16) Pazos, E.; Mosquera, J.; Vázquez, M. E.; Mascareñas, J. L. DNA Recognition by Synthetic Constructs. *ChemBioChem* **2011**, *12* (13), 1958–1973. https://doi.org/10.1002/CBIC.201100247.

(17) Raumann, B. E.; Rould, M. A.; Pabo, C. O.; Sauer, R. T. DNA Recognition by β-Sheets in the Arc Represser–Operator Crystal Structure. *Nat. 1994 3676465* **1994**, *367* (6465), 754–757. https://doi.org/10.1038/367754a0.

(18) Wu, L.; Mcelheny, D.; Setnicka, V.; Hilario, J.; Keiderling, T. A. Role of Different β-Turns in β-Hairpin Conformation and Stability Studied by Optical Spectroscopy. *Proteins Struct. Funct. Bioinforma.* **2012**, *80* (1), 44–60. https://doi.org/10.1002/PROT.23140.

(19) Pastor, M. T.; Ló Pez De La Paz, M.; Lacroix, E.; Serrano, L.; Pé Rez-Payá, E. *Combinatorial Approaches: A New Tool to Search for Highly Structured-Hairpin Peptides*; 2001. www.pnas.orgcgidoi10.1073pnas.012583999.

(20) Stefanucci, A.; Costante, R.; Macedonio, G.; Dvoracsko, S.; Mollica, A. *Cysteine-, Methionine-and Seleno-Cysteine-Proline Chimeras: Synthesis and Their Use in Peptidomimetics Design*; 2016.

(21) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. Tryptophan Zippers: Stable, Monomeric β-Hairpins. *Proc. Natl. Acad. Sci.* **2001**, *98* (10), 5578–5583. https://doi.org/10.1073/PNAS.091100898.

(22) Dostál, L.; Misselwitz, R.; Welfle, H. Arc Repressor-Operator DNA Interactions and Contribution of Phe10 to Binding Specificity. *Biochemistry* **2005**, *44* (23), 8387–8396. https://doi.org/10.1021/BI0476073.

(23) Micsonai, A.; Wien, F.; Kernya, L.; Lee, Y. H.; Goto, Y.; Réfrégiers, M.; Kardos, J. Accurate Secondary Structure Prediction and Fold Recognition for Circular Dichroism Spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (24), E3095–E3103. https://doi.org/10.1073/PNAS.1500851112/SUPPL_FILE/PNAS.201500851SI.PDF.

(24) Krey, A. K.; Hahn, F. E. Studies on the Methyl Green-DNA Complex and Its Dissociation by Drugs. *Biochemistry* **1975**, *14* (23), 5061–5067. https://doi.org/10.1021/BI00694A005/ASSET/BI00694A005.FP.PNG_V03.

(25) Kim, S. K.; Nordén, B. Methyl Green. A DNA Major-Groove Binding Drug. *FEBS Lett.* **1993**, *315* (1), 61–64. https://doi.org/10.1016/0014-5793(93)81133-K.

(26) Vardevanyan, P. O.; Antonyan, A. P.; Parsadanyan, M. A.; Davtyan, H. G.; Karapetyan, A. T. The Binding of Ethidium Bromide with DNA: Interaction with Single- and Double-Stranded Structures. *Exp. Mol. Med. 2003 356* **2003**, *35* (6), 527–533. https://doi.org/10.1038/emm.2003.68.

(27) Morris, G. M.; Ruth, H.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30* (16), 2785–2791. https://doi.org/10.1002/JCC.21256.

(28) Hwang, T. L.; Shaka, A. J. Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients. *J. Magn. Reson. - Ser. A* **1995**, *112* (2), 275–279. https://doi.org/10.1006/JMRA.1995.1047.

(29) States, D. J.; Haberkorn, R. A.; Ruben, D. J. *A Two-Dimensional Nuclear Overhauser Experiment with Pure Absorption Phase in Four Quadrants**; 1982; Vol. 48.

(30) Bartels, C.; Xia, T. he; Billeter, M.; Güntert, P.; Wüthrich, K. The Program XEASY for Computer-Supported NMR Spectral Analysis of Biological Macromolecules. *J. Biomol. NMR* **1995**, *6* (1), 1–10. https://doi.org/10.1007/BF00417486/METRICS.

(31) Güntert, P.; Buchner, L. Combined Automated NOE Assignment and Structure Calculation with CYANA. *J. Biomol. NMR* **2015**, *62* (4), 453–471. https://doi.org/10.1007/S10858-015-9924-9.

(32) Maple, J. R.; Dinur, U.; Hagler, A. T. Derivation of Force Fields for

Molecular Mechanics and Dynamics from Ab Initio Energy Surfaces. *Proc. Natl. Acad. Sci.* **1988**, *85* (15), 5350–5354. https://doi.org/10.1073/PNAS.85.15.5350.

(33) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25* (13), 1605–1612. https://doi.org/10.1002/JCC.20084.

(34) Góngora-Benítez, M.; Tulla-Puche, J.; Albericio, F. Multifaceted Roles of Disulfide Bonds. Peptides as Therapeutics. *Chem. Rev.* **2014**, *114* (2), 901–926. https://doi.org/10.1021/CR400031Z/ASSET/IMAGES/CR400031Z.SOCIAL.JPEG_V03.

(35) BARANY, G.; KNEIB-CORDONIER, N.; MULLEN, D. G. Solid-Phase Peptide Synthesis: A Silver Anniversary Report*. *Int. J. Pept. Protein Res.* **1987**, *30* (6), 705–739. https://doi.org/10.1111/J.1399-3011.1987.TB03385.X.

(36) Hughes, A. B. Amino Acids, Peptides, and Proteins in Organic Chemistry.

(37) Berger, N.; Li, F.; Mallick, B.; Brüggemann, J. T.; Sander, W.; Merten, C. Solution and Solid State Conformational Preferences of a Family of Cyclic Disulphide Bridged Tetrapeptides. *Biopolymers* **2017**, *107* (1), 28–34. https://doi.org/10.1002/BIP.22986.

(38) Almeida, A. M.; Li, R.; Gellman, S. H. Parallel β-Sheet Secondary Structure Is Stabilized and Terminated by Interstrand Disulfide Cross-Linking. *J. Am. Chem. Soc.* **2012**, *134* (1), 75–78. https://doi.org/10.1021/JA208856C/SUPPL_FILE/JA208856C_SI_001.PDF.

(39) Santiveri, C. M.; León, E.; Rico, M.; Jiménez, M. A. Context-Dependence of the Contribution of Disulfide Bonds to Beta-Hairpin Stability. *Chemistry* **2008**, *14* (2), 488–499. https://doi.org/10.1002/CHEM.200700845.

(40) Snellman, A.; Tuomisto, A.; Koski, A.; Latvanlehto, A.; Pihlajaniemi, T. The Role of Disulfide Bonds and α-Helical Coiled-Coils in the Biosynthesis of Type XIII Collagen and Other Collagenous Transmembrane Proteins. *J. Biol. Chem.* **2007**, *282* (20), 14898–14905. https://doi.org/10.1074/JBC.M609605200.

(41) Kourra, C. M. B. K.; Cramer, N. Converting Disulfide Bridges in Native Peptides to Stable Methylene Thioacetals. *Chem. Sci.* **2016**, *7* (12), 7007–7012. https://doi.org/10.1039/C6SC02285E.

(42) Hargittai, B.; Solé, N. A.; Groebe, D. R.; Abramson, S. N.; Barany, G. Chemical Syntheses and Biological Activities of Lactam Analogues of α-Conotoxin SI. *J. Med. Chem.* **2000**, *43* (25), 4787–4792. https://doi.org/10.1021/JM990635C.

(43) D'Addona, D.; Carotenuto, A.; Novellino, E.; Piccand, V.; Reubi, J. C.; Di Cianni, A.; Gori, F.; Papini, A. M.; Ginanneschi, M. Novel Sst5-Selective Somatostatin Dicarba-Analogues: Synthesis and Conformation-Affinity Relationships. *J. Med. Chem.* **2008**, *51* (3), 512–520. https://doi.org/10.1021/JM070886I/SUPPL_FILE/JM070886I-FILE002.PDF.

(44) Knerr, P. J.; Tzekou, A.; Ricklin, D.; Qu, H.; Chen, H.; Van Der Donk, W. A.; Lambris, J. D. Synthesis and Activity of Thioether-Containing Analogues of the Complement Inhibitor Compstatin. *ACS Chem. Biol.* **2011**, *6* (7), 753–760. https://doi.org/10.1021/CB2000378/SUPPL_FILE/CB2000378_SI_001.PDF.

(45) Arai, K.; Takei, T.; Okumura, M.; Watanabe, S.; Amagai, Y.; Asahina, Y.; Moroder, L.; Hojo, H.; Inaba, K.; Iwaoka, M. Preparation of Selenoinsulin as a Long-Lasting Insulin Analogue. *Angew. Chemie Int. Ed.* **2017**, *56* (20), 5522–5526. https://doi.org/10.1002/ANIE.201701654.

(46) White, A. M.; de Veer, S. J.; Wu, G.; Harvey, P. J.; Yap, K.; King, G. J.;

Swedberg, J. E.; Wang, C. K.; Law, R. H. P.; Durek, T.; Craik, D. J. Application and Structural Analysis of Triazole-Bridged Disulfide Mimetics in Cyclic Peptides. *Angew. Chemie Int. Ed.* **2020**, *59* (28), 11273–11277. https://doi.org/10.1002/ANIE.202003435.

(47) Knuhtsen, A.; Whitmore, C.; McWhinnie, F. S.; McDougall, L.; Whiting, R.; Smith, B. O.; Timperley, C. M.; Green, A. C.; Kinnear, K. I.; Jamieson, A. G. α-Conotoxin GI Triazole-Peptidomimetics: Potent and Stable Blockers of a Human Acetylcholine Receptor. *Chem. Sci.* **2019**, *10* (6), 1671–1676. https://doi.org/10.1039/C8SC04198A.

(48) Quigley, N. G.; Tomassi, S.; Saverio Di Leva, F.; Di Maro, S.; Richter, F.; Steiger, K.; Kossatz, S.; Marinelli, L.; Notni, J. Click-Chemistry (CuAAC) Trimerization of an Avβ6 Integrin Targeting Ga-68-Peptide: Enhanced Contrast for In-Vivo PET Imaging of Human Lung Adenocarcinoma Xenografts. *Chembiochem* **2020**, *21* (19), 2836. https://doi.org/10.1002/CBIC.202000200.

(49) Tala, S. R.; Singh, A.; Lensing, C. J.; Schnell, S. M.; Freeman, K. T.; Rocca, J. R.; Haskell-Luevano, C. 1,2,3-Triazole Rings as a Disulfide Bond Mimetic in Chimeric AGRP-Melanocortin Peptides: Design, Synthesis, and Functional Characterization. *ACS Chem. Neurosci.* **2018**, *9* (5), 1001–1013. https://doi.org/10.1021/ACSCHEMNEURO.7B00422/SUPPL_FILE/CN7B00422_SI_001.PDF.

(50) Pacifico, S.; Kerckhoffs, A.; Fallow, A. J.; Foreman, R. E.; Guerrini, R.; McDonald, J.; Lambert, D. G.; Jamieson, A. G. Urotensin-II Peptidomimetic Incorporating a Non-Reducible 1,5-Triazole Disulfide Bond Reveals a Pseudo-Irreversible Covalent Binding Mechanism to the Urotensin G-Protein Coupled Receptor. *Org. Biomol. Chem.* **2017**, *15* (21), 4704–4710. https://doi.org/10.1039/C7OB00959C.

(51) Burger, J. A.; Kipps, T. J. CXCR4: A Key Receptor in the Crosstalk between Tumor Cells and Their Microenvironment. *Blood* **2006**, *107* (5),

1761–1767. https://doi.org/10.1182/BLOOD-2005-08-3182.

(52) Chatterjee, S.; Behnam Azad, B.; Nimmagadda, S. The Intricate Role of CXCR4 in Cancer. *Adv. Cancer Res.* **2014**, *124*, 31–82. https://doi.org/10.1016/B978-0-12-411638-2.00002-1.

(53) Kuhne, M. R.; Mulvey, T.; Belanger, B.; Chen, S.; Pan, C.; Chong, C.; Cao, F.; Niekro, W.; Kempe, T.; Henning, K. A.; Cohen, L. J.; Korman, A. J.; Cardarelli, P. M. BMS-936564/MDX-1338: A Fully Human Anti-CXCR4 Antibody Induces Apoptosis in Vitro and Shows Antitumor Activity in Vivo in Hematologic Malignancies. *Clin. Cancer Res.* **2013**, *19* (2), 357–366. https://doi.org/10.1158/1078-0432.CCR-12-2333/85249/AM/BMS-936564-MDX-1338-A-FULLY-HUMAN-ANTI-CXCR4.

(54) Grande, F.; Giancotti, G.; Ioele, G.; Occhiuzzi, M. A.; Garofalo, A. An Update on Small Molecules Targeting CXCR4 as Starting Points for the Development of Anti-Cancer Therapeutics. *Eur. J. Med. Chem.* **2017**, *139*, 519–530. https://doi.org/10.1016/J.EJMECH.2017.08.027.

(55) Karpova, D.; Bräuninger, S.; Wiercinska, E.; Krämer, A.; Stock, B.; Graff, J.; Martin, H.; Wach, A.; Escot, C.; Douglas, G.; Romagnoli, B.; Chevalier, E.; Dembowski, K.; Hooftman, L.; Bonig, H. Mobilization of Hematopoietic Stem Cells with the Novel CXCR4 Antagonist POL6326 (Balixafortide) in Healthy Volunteers-Results of a Dose Escalation Trial. *J. Transl. Med.* **2017**, *15* (1), 1–12. https://doi.org/10.1186/S12967-016-1107-2/TABLES/3.

(56) Pelus, L. M.; Bian, H.; Fukuda, S.; Wong, D.; Merzouk, A.; Salari, H. The CXCR4 Agonist Peptide, CTCE-0021, Rapidly Mobilizes Polymorphonuclear Neutrophils and Hematopoietic Progenitor Cells into Peripheral Blood and Synergizes with Granulocyte Colony-Stimulating Factor. *Exp. Hematol.* **2005**, *33* (3), 295–307. https://doi.org/10.1016/J.EXPHEM.2004.11.008.

(57) Mona, C. E.; Besserer-Offroy, É.; Cabana, J.; Leduc, R.; Lavigne, P.; Heveker, N.; Marsault, É.; Escher, E. This Is the Accepted (Postprint)

Version of the Following Design, Synthesis, and Biological Evaluation of CXCR4 Ligands Corresponding Authors. *Org Biomol Chem* **2016**. https://doi.org/10.1039/c6ob01484d.

(58) Di Maro, S.; Trotta, A. M.; Brancaccio, D.; Di Leva, F. S.; La Pietra, V.; Ieranò, C.; Napolitano, M.; Portella, L.; D'Alterio, C.; Siciliano, R. A.; Sementa, D.; Tomassi, S.; Carotenuto, A.; Novellino, E.; Scala, S.; Marinelli, L. Exploring the N-Terminal Region of C-X-C Motif Chemokine 12 (CXCL12): Identification of Plasma-Stable Cyclic Peptides As Novel, Potent C-X-C Chemokine Receptor Type 4 (CXCR4) Antagonists. *J. Med. Chem.* **2016**, *59* (18), 8369–8380. https://doi.org/10.1021/ACS.JMEDCHEM.6B00695/SUPPL_FILE/JM6B00695_SI_001.PDF.

(59) Empting, M.; Avrutina, O.; Meusinger, R.; Fabritz, S.; Reinwarth, M.; Biesalski, M.; Voigt, S.; Buntkowsky, G.; Kolmar, H. "Triazole Bridge": Disulfide-Bond Replacement by Ruthenium-Catalyzed Formation of 1,5-Disubstituted 1,2,3-Triazoles. *Angew. Chemie Int. Ed.* **2011**, *50* (22), 5207–5211. https://doi.org/10.1002/ANIE.201008142.

(60) Merlino, F.; Tomassi, S.; Yousif, A. M.; Messere, A.; Marinelli, L.; Grieco, P.; Novellino, E.; Cosconati, S.; Di Maro, S. Boosting Fmoc Solid-Phase Peptide Synthesis by Ultrasonication. *Org. Lett.* **2019**, *21* (16), 6378–6382. https://doi.org/10.1021/ACS.ORGLETT.9B02283/SUPPL_FILE/OL9B02283_SI_001.PDF.

(61) Sargent, D. F.; Schwyzer, R. Membrane Lipid Phase as Catalyst for Peptidereceptor Interactions (Electrostatic Accumulation/Sequential Binding Steps/Macroscopic Binding Characteristics/Catalysis by Micelles/Reconstitution Criteria). **1986**, *83*, 5774–5778.

(62) Wu, B.; Chien, E. Y. T.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; Hamel, D. J.; Kuhn, P.; Handel, T. M.; Cherezov, V.; Stevens, R. C. Structures of the CXCR4

Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists. *Science (80-.).* **2010**, *330* (6007), 1066–1071. https://doi.org/10.1126/SCIENCE.1194396/SUPPL_FILE/WU.SOM.PDF.

(63) Zachariassen, Z. G.; Karlshøj, S.; Haug, B. E.; Rosenkilde, M. M.; Våbenø, J. Probing the Molecular Interactions between CXC Chemokine Receptor 4 (CXCR4) and an Arginine-Based Tripeptidomimetic Antagonist (KRH-1636). *J. Med. Chem.* **2015**, *58* (20), 8141–8153. https://doi.org/10.1021/ACS.JMEDCHEM.5B00987/SUPPL_FILE/JM5B00987_SI_001.PDF.

(64) Thiele, S.; Mungalpara, J.; Steen, A.; Rosenkilde, M. M.; Våbenø, J. Determination of the Binding Mode for the Cyclopentapeptide CXCR4 Antagonist FC131 Using a Dual Approach of Ligand Modifications and Receptor Mutagenesis. *Br. J. Pharmacol.* **2014**, *171* (23), 5313–5329. https://doi.org/10.1111/BPH.12842.

(65) Smith, E. W.; Nevins, A. M.; Qiao, Z.; Liu, Y.; Getschman, A. E.; Vankayala, S. L.; Kemp, M. T.; Peterson, F. C.; Li, R.; Volkman, B. F.; Chen, Y. Structure-Based Identification of Novel Ligands Targeting Multiple Sites within a Chemokine-G-Protein-Coupled-Receptor Interface. *J. Med. Chem.* **2016**, *59* (9), 4342–4351. https://doi.org/10.1021/ACS.JMEDCHEM.5B02042.

(66) Wescott, M. P.; Kufareva, I.; Paes, C.; Goodman, J. R.; Thaker, Y.; Puffer, B. A.; Berdougo, E.; Rucker, J. B.; Handel, T. M.; Doranz, B. J. Signal Transmission through the CXC Chemokine Receptor 4 (CXCR4) Transmembrane Helices. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (35), 9928–9933. https://doi.org/10.1073/PNAS.1601278113.

(67) Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. https://doi.org/10.1007/s10822-013-9644-8.

(68) Shelley, J. C.; Anuradha, A. E.; Ae, C.; Ae, L. L. F.; Greenwood, J. R.;

Mathew, A. E.; Ae, R. T.; Uchimaya, M. Epik: A Software Program for PK a Prediction and Protonation State Generation for Drug-like Molecules. https://doi.org/10.1007/s10822-007-9133-z.

(69) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739–1749. https://doi.org/10.1021/JM0306430/SUPPL_FILE/JM0306430_S.PDF.

(70) Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.; Friesner, R. A. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* **2016**, *12* (1), 281–296. https://doi.org/10.1021/ACS.JCTC.5B00864/SUPPL_FILE/CT5B00864_SI_001.ZIP.

(71) McFarland, H. F.; Martin, R. Multiple Sclerosis: A Complicated Picture of Autoimmunity. *Nat. Immunol. 2007 89* **2007**, *8* (9), 913–919. https://doi.org/10.1038/ni1507.

(72) Noseworthy, J. H.; Lucchinetti, C.; Rodriguez, M.; Weinshenker, B. G. Multiple Sclerosis. *https://doi.org/10.1056/NEJM200009283431307* **2000**, *343* (13), 938–952. https://doi.org/10.1056/NEJM200009283431307.

(73) Lolli, F.; Mulinacci, B.; Carotenuto, A.; Bonetti, B.; Sabatino, G.; Mazzanti, B.; D'Ursi, A. M.; Novellino, E.; Pazzagli, M.; Lovato, L.; Alcaro, M. C.; Peroni, E.; Pozo-Carrero, M. C.; Nuti, F.; Battistini, L.; Borsellino, G.; Chelli, M.; Rovero, P.; Papini, A. M. An N-Glucosylated Peptide Detecting Disease-Specific Autoantibodies, Biomarkers of Multiple Sclerosis. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (29), 10273–10278. https://doi.org/10.1073/PNAS.0503178102/SUPPL_FILE/03178SUPPTEXT

.HTML.

(74) Levin, L. I.; Munger, K. L.; O'Reilly, E. J.; Falk, K. I.; Ascherio, A. Primary Infection with the Epstein-Barr Virus and Risk of Multiple Sclerosis. *Ann. Neurol.* **2010**, *67* (6), 824–830. https://doi.org/10.1002/ANA.21978.

(75) Berger, T.; Reindl, M. Antibody Biomarkers in CNS Demyelinating Diseases – a Long and Winding Road. *Eur. J. Neurol.* **2015**, *22* (8), 1162–1168. https://doi.org/10.1111/ENE.12759.

(76) Fraussen, J.; Claes, N.; de Bock, L.; Somers, V. Targets of the Humoral Autoimmune Response in Multiple Sclerosis. *Autoimmun. Rev.* **2014**, *13* (11), 1126–1137. https://doi.org/10.1016/J.AUTREV.2014.07.002.

(77) Fraussen, J.; Vrolix, K.; Martinez-Martinez, P.; Losen, M.; De Baets, M. H.; Stinissen, P.; Somers, V. B Cell Characterization and Reactivity Analysis in Multiple Sclerosis. *Autoimmun. Rev.* **2009**, *8* (8), 654–658. https://doi.org/10.1016/J.AUTREV.2009.02.030.

(78) Somers, K.; Govarts, C.; Stinissen, P.; Somers, V. Multiplexing Approaches for Autoantibody Profiling in Multiple Sclerosis. *Autoimmun. Rev.* **2009**, *8* (7), 573–579. https://doi.org/10.1016/J.AUTREV.2009.02.001.

(79) Carotenuto, A.; Maria D'ursi, A.; Nardi, E.; Papini, A. M.; Rovero, P. Conformational Analysis of a Glycosylated Human Myelin Oligodendrocyte Glycoprotein Peptide Epitope Able To Detect Antibody Response in Multiple Sclerosis. **2001**. https://doi.org/10.1021/jm010811t.

(80) Mazzucco, S.; Matà, S.; Vergelli, M.; Fioresi, R.; Nardi, E.; Mazzanti, B.; Chelli, M.; Lolli, F.; Ginanneschi, M.; Pinto, F.; Massacesi, L.; Papini, A. M. A Synthetic Glycopeptide of Human Myelin Oligodendrocyte Glycoprotein to Detect Antibody Responses in Multiple Sclerosis and Other Neurological Diseases. *Bioorganic Med. Chem. Lett.* **1999**, *9* (2), 167–172. https://doi.org/10.1016/S0960-894X(98)00698-2.

(81) Lolli, F.; Mazzanti, B.; Pazzagli, M.; Peroni, E.; Alcaro, M. C.; Sabatino, G.; Lanzillo, R.; Morra, V. B.; Santoro, L.; Gasperini, C.; Galgani, S.; D'Elios,

M. M.; Zipoli, V.; Sotgiu, S.; Pugliatti, M.; Rovero, P.; Chelli, M.; Papini, A. M. The Glycopeptide CSF114(Glc) Detects Serum Antibodies in Multiple Sclerosis. *J. Neuroimmunol.* **2005**, *167* (1–2), 131–137. https://doi.org/10.1016/J.JNEUROIM.2005.05.016.

(82) Papini, A.M.; Chelli, M.; Rovero, P.; Lolli, F. Glycopeptides, Their Preparation and Use in the Diagnosis or Therapeutic Treatment of Multiple Sclerosis, June 19, 2003.

(83) Carotenuto, A.; D′ursi, A. M.; Mulinacci, B.; Paolini, I.; Lolli, F.; Papini, A. M.; Novellino, E.; Rovero, P. Conformation-Activity Relationship of Designed Glycopeptides as Synthetic Probes for the Detection of Autoantibodies, Biomarkers of Multiple Sclerosis. **2006**. https://doi.org/10.1021/jm060117j.

(84) Carotenuto, A.; Alcaro, M. C.; Saviello, M. R.; Peroni, E.; Nuti, F.; Papini, A. M.; Novellino, E.; Rovero, P. Designed Glycopeptides with Different-Turn Types as Synthetic Probes for the Detection of Autoantibodies as Biomarkers of Multiple Sclerosis. https://doi.org/10.1021/jm800391y.

(85) Nuti, F.; Paolini, I.; Cardona, F.; Chelli, M.; Lolli, F.; Brandi, A.; Goti, A.; Rovero, P.; Papini, A. M. Fmoc-Protected Iminosugar Modified Asparagine Derivatives as Building Blocks for Glycomimetics-Containing Peptides. *Bioorganic Med. Chem.* **2007**, *15* (12), 3965–3973. https://doi.org/10.1016/J.BMC.2007.04.007.

(86) Nuti, F.; Peroni, E.; Real-Fernández, F.; Bonache, M. A.; Chevalier-Isaad, A. Le; Chelli, M.; Lubin-Germain, N.; Uziel, J.; Rovero, P.; Lolli, F.; Papini, A. M. Posttranslationally Modified Peptides Efficiently Mimicking Neoantigens: A Challenge for Theragnostics of Autoimmune Diseases. *Biopolymers* **2010**, *94* (6), 791–799. https://doi.org/10.1002/BIP.21456.

(87) Pandey, S.; Alcaro, M. C.; Scrima, M.; Peroni, E.; Paolini, I.; Di Marino, S.; Barbetti, F.; Carotenuto, A.; Novellino, E.; Maria Papini, A.; Maria, A. D.; Rovero, P. Designed Glucopeptides Mimetics of Myelin Protein Epitopes As

Synthetic Probes for the Detection of Autoantibodies, Biomarkers of Multiple Sclerosis. **2012**. https://doi.org/10.1021/jm301031r.

(88) Walvoort, M. T. C.; Testa, C.; Eilam, R.; Aharoni, R.; Nuti, F.; Rossi, G.; Real-Fernandez, F.; Lanzillo, R.; Brescia Morra, V.; Lolli, F.; Rovero, P.; Imperiali, B.; Papini, A. M. Antibodies from Multiple Sclerosis Patients Preferentially Recognize Hyperglucosylated Adhesin of Non-Typeable Haemophilus Influenzae. *Sci. Reports 2016 61* **2016**, *6* (1), 1–12. https://doi.org/10.1038/srep39430.

(89) Nuti, F.; Real Fernandez, F.; Sabatino, G.; Peroni, E.; Mulinacci, B.; Paolini, I.; Di Pisa, M.; Tiberi, C.; Lolli, F.; Petruzzo, M.; Lanzillo, R.; Morra, V. B.; Rovero, P.; Papini, A. M. A Multiple N-Glucosylated Peptide Epitope Efficiently Detecting Antibodies in Multiple Sclerosis. *Brain Sci 2020*, 453. https://doi.org/10.3390/brainsci10070453.

(90) Rizzolo, F.; Testa, C.; Lambardi, D.; Chorev, M.; Chelli, M.; Rovero, P.; Papini, A. M. Conventional and Microwave-Assisted SPPS Approach: A Comparative Synthesis of PTHrP(1-34)NH 2. *J. Pept. Sci.* **2011**, *17* (10), 708–714. https://doi.org/10.1002/psc.1395.

(91) Real Fernández, F.; Di Pisa, M.; Rossi, G.; Auberger, N.; Lequin, O.; Larregola, M.; Benchohra, A.; Mansuy, C.; Chassaing, G.; Lolli, F.; Hayek, J.; Lavielle, S.; Rovero, P.; Mallet, J. M.; Papini, A. M. ari. Antibody Recognition in Multiple Sclerosis and Rett Syndrome Using a Collection of Linear and Cyclic N-Glucosylated Antigenic Probes. *Pept. Sci.* **2015**, *104* (5), 560–576. https://doi.org/10.1002/BIP.22677.

(92) Polman, C. H.; Reingold, S. C.; Banwell, B.; Clanet, M.; Cohen, J. A.; Filippi, M.; Fujihara, K.; Havrdova, E.; Hutchinson, M.; Kappos, L.; Lublin, F. D.; Montalban, X.; O'Connor, P.; Sandberg-Wollheim, M.; Thompson, A. J.; Waubant, E.; Weinshenker, B.; Wolinsky, J. S. Diagnostic Criteria for Multiple Sclerosis: 2010 Revisions to the McDonald Criteria. *Ann. Neurol.* **2011**, *69* (2), 292–302. https://doi.org/10.1002/ANA.22366.

(93) Thompson, A. J.; Banwell, B. L.; Barkhof, F.; Carroll, W. M.; Coetzee, T.; Comi, G.; Correale, J.; Fazekas, F.; Filippi, M.; Freedman, M. S.; Fujihara, K.; Galetta, S. L.; Hartung, H. P.; Kappos, L.; Lublin, F. D.; Marrie, R. A.; Miller, A. E.; Miller, D. H.; Montalban, X.; Mowry, E. M.; Sorensen, P. S.; Tintoré, M.; Traboulsee, A. L.; Trojano, M.; Uitdehaag, B. M. J.; Vukusic, S.; Waubant, E.; Weinshenker, B. G.; Reingold, S. C.; Cohen, J. A. Diagnosis of Multiple Sclerosis: 2017 Revisions of the McDonald Criteria. *Lancet Neurol.* **2018**, *17* (2), 162–173. https://doi.org/10.1016/S1474-4422(17)30470-2.

(94) Stults, J. T.; Bourell, J. H.; Canova-Davis, E.; Ling, V. T.; Laramee, G. R.; Winslow, J. W.; Griffin, P. R.; Rinderknecht, E.; Vandlen, R. L. Structural Characterization by Mass Spectrometry of Native and Recombinant Human Relaxin. *Biomed. Environ. Mass Spectrom.* **1990**, *19* (11), 655–664. https://doi.org/10.1002/BMS.1200191105.

(95) Bani, D. Recombinant Human H2 Relaxin (Serelaxin) as a Cardiovascular Drug: Aiming at the Right Target. *Drug Discov. Today* **2020**, *25* (7), 1239–1244. https://doi.org/10.1016/J.DRUDIS.2020.04.014.

(96) Samuel, C. S.; Bennett, R. G. Relaxin as an Anti-Fibrotic Treatment: Perspectives, Challenges and Future Directions. *Biochem. Pharmacol.* **2022**, *197*, 114884. https://doi.org/10.1016/J.BCP.2021.114884.

(97) Sassoli, C.; Nistri, S.; Chellini, F.; Bani, D. Human Recombinant Relaxin (Serelaxin) as Anti-Fibrotic Agent: Pharmacology, Limitations and Actual Perspectives. *Curr. Mol. Med.* **2021**, *22* (3), 196–208. https://doi.org/10.2174/1566524021666210309113650.

(98) Bathgate, R. A. D.; Halls, M. L.; van der Westhuizen, E. T.; Callander, G. E.; Kocan, M.; Summers, R. J. Relaxin Family Peptides and Their Receptors. *Physiol. Rev.* **2013**, *93* (1), 405–480. https://doi.org/10.1152/PHYSREV.00001.2012/ASSET/IMAGES/LARGE/Z9J0011326410012.JPEG.

(99) Chen, S. A.; Perlman, A. J.; Spanski, N.; Peterson, C. M.; Sanders, S. W.; Jaffe, R.; Martin, M.; Yalcinkaya, T.; Cefalo, R. C.; Chescheir, N. C.; Menard, M. K.; Mordenti, J. The Pharmacokinetics of Recombinant Human Relaxin in Nonpregnant Women After Intravenous, Intravaginal, and Intracervical Administration. *Pharmaceutical Research: An Official Journal of the American Association of Pharmaceutical Scientists*. 1993, pp 834–838. https://doi.org/10.1023/A:1018901009062.

(100) Park, J.-I.; Semyonov, J.; Yi, W.; Chang, C. L.; Hsu, S. Y. T. Regulation of Receptor Signaling by Relaxin A Chain Motifs. *J. Biol. Chem.* **2008**, *283* (46), 32099–32109. https://doi.org/10.1074/jbc.m806817200.

(101) Praveen, P.; Kocan, M.; Valkovic, A.; Bathgate, R.; Hossain, M. A. Single Chain Peptide Agonists of Relaxin Receptors. *Mol. Cell. Endocrinol.* **2019**, *487*, 34–39. https://doi.org/10.1016/J.MCE.2019.01.008.

(102) Bathgate, R. A.; Ivell, R.; Sanborn, B. M.; Sherwood, O. D.; Summers, R. J. International Union of Pharmacology LVII: Recommendations for the Nomenclature of Receptors for Relaxin Family Peptides. **2006**. https://doi.org/10.1124/pr.58.1.9.

(103) Wade, J. D.; Lin, F.; Salvatore, D.; Otvos, L.; Tregear, G. W. Synthesis and Characterization of Human Gene 1 Relaxin Peptides. *Biomed. Pept. Proteins Nucleic Acids* **1996**, *2* (2), 27–32.

(104) Hossain, M. A.; Kocan, M.; Yao, S. T.; Royce, S. G.; Nair, V. B.; Siwek, C.; Patil, N. A.; Harrison, I. P.; Rosengren, K. J.; Selemidis, S.; Summers, R. J.; Wade, J. D.; Bathgate, R. A. D.; Samuel, C. S. A Single-Chain Derivative of the Relaxin Hormone Is a Functionally Selective Agonist of the G Protein-Coupled Receptor, RXFP1. *Chem. Sci.* **2016**, *7* (6), 3805–3819. https://doi.org/10.1039/C5SC04754D.

(105) Hossain, M. A.; Rosengren, K. J.; Samuel, C. S.; Shabanpoor, F.; Chan, L. J.; Bathgate, R. A. D.; Wade, J. D. The Minimal Active Structure of Human Relaxin-2. *J. Biol. Chem.* **2011**, *286* (43), 37555–37565.

https://doi.org/10.1074/JBC.M111.282194.

(106) D'Ercole, A.; Sabatino, G.; Pacini, L.; Impresari, E.; Capecchi, I.; Papini, A. M.; Rovero, P. On-Resin Microwave-Assisted Copper-Catalyzed Azide-Alkyne Cycloaddition of H1-Relaxin B Single Chain 'Stapled' Analogues. *Pept. Sci.* **2020**, *112* (4), e24159. https://doi.org/10.1002/PEP2.24159.

(107) Halls, M. L.; Bond, C. P.; Sudo, S.; Kumagai, J.; Ferraro, T.; Layfield, S.; Bathgate, R. A. D.; Summers, R. J. Multiple Binding Sites Revealed by Interaction of Relaxin Family Peptides with Native and Chimeric Relaxin Family Peptide Receptors 1 and 2 (LGR7 and LGR8). *J. Pharmacol. Exp. Ther.* **2005**, *313* (2), 677–687. https://doi.org/10.1124/JPET.104.080655.

(108) Shabanpoor, F.; Bathgate, R. A. D.; Belgi, A.; Chan, L. J.; Nair, V. B.; Wade, J. D.; Hossain, M. A. Site-Specific Conjugation of a Lanthanide Chelator and Its Effects on the Chemical Synthesis and Receptor Binding Affinity of Human Relaxin-2 Hormone. *Biochem. Biophys. Res. Commun.* **2012**, *420* (2), 253–256. https://doi.org/10.1016/J.BBRC.2012.02.141.

(109) Nguyen, B. T.; Yang, L.; Sanborn, B. M.; Dessauer, C. W. Phosphoinositide 3-Kinase Activity Is Required for Biphasic Stimulation of Cyclic Adenosine 3,5-Monophosphate by Relaxin. *Mol. Endocrinol.* **2003**, *17*, 1075–1084. https://doi.org/10.1210/me.2002-0284.

(110) Bani, D.; Nistri, S.; Cinci, L.; Giannini, L.; Princivalle, M.; Elliott, L.; Bigazzi, M.; Masini, E. A Novel, Simple Bioactivity Assay for Relaxin Based on Inhibition of Platelet Aggregation. *Regul. Pept.* **2007**, *144* (1–3), 10–16. https://doi.org/10.1016/J.REGPEP.2007.05.004.

(111) Bacolla, A.; Wells, R. D. Non-B DNA Conformations as Determinants of Mutagenesis and Human Disease. *Mol. Carcinog.* **2009**, *48* (4), 273–285. https://doi.org/10.1002/MC.20507.

(112) Assi, H. A.; Garavís, M.; González, C.; Damha, M. J. I-Motif DNA: Structural Features and Significance to Cell Biology. *Nucleic Acids Res.* **2018**, *46* (16), 8038–8056. https://doi.org/10.1093/NAR/GKY735.

(113) Gajarský, M.; Živković, M. L.; Stadlbauer, P.; Pagano, B.; Fiala, R.; Amato, J.; Tomáška, L.; Šponer, J.; Plavec, J.; Trantírek, L. Structure of a Stable G-Hairpin. *J. Am. Chem. Soc.* **2017**, *139* (10), 3591–3594. https://doi.org/10.1021/JACS.6B10786/SUPPL_FILE/JA6B10786_SI_002.ZIP.

(114) Spiegel, J.; Adhikari, S.; Balasubramanian, S. The Structure and Function of DNA G-Quadruplexes. *Trends Chem.* **2020**, *2* (2), 123–136. https://doi.org/10.1016/J.TRECHM.2019.07.002.

(115) Lane, A. N.; Chaires, J. B.; Gray, R. D.; Trent, J. O. Stability and Kinetics of G-Quadruplex Structures. *Nucleic Acids Res.* **2008**, *36* (17), 5482–5515. https://doi.org/10.1093/nar/gkn517.

(116) Bhattacharyya, D.; Arachchilage, G. M.; Basu, S. Metal Cations in G-Quadruplex Folding and Stability. *Front. Chem.* **2016**, *4* (SEP). https://doi.org/10.3389/FCHEM.2016.00038/FULL.

(117) Largy, E.; Mergny, J. L.; Gabelica, V. Role of Alkali Metal Ions in G-Quadruplex Nucleic Acid Structure and Stability. *Met. Ions Life Sci.* **2016**, *16*, 203–258. https://doi.org/10.1007/978-3-319-21756-7_7.

(118) Pagano, B.; Mattia, C. A.; Cavallo, L.; Uesugi, S.; Giancola, C.; Fraternali, F. Stability and Cations Coordination of DNA and RNA 14-Mer G-Quadruplexes: A Multiscale Computational Approach. *J. Phys. Chem. B* **2008**, *112* (38), 12115–12123. https://doi.org/10.1021/JP804036J.

(119) Guédin, A.; Gros, J.; Alberti, P.; Mergny, J. L. How Long Is Too Long? Effects of Loop Size on G-Quadruplex Stability. *Nucleic Acids Res.* **2010**, *38* (21), 7858–7868. https://doi.org/10.1093/NAR/GKQ639.

(120) Bugaut, A.; Biochemistry, S. B.-; 2008, undefined. A Sequence-Independent Study of the Influence of Short Loop Lengths on the Stability and Topology of Intramolecular DNA G-Quadruplexes. *ACS Publ.* **2008**, *47* (2), 689–697. https://doi.org/10.1021/bi701873c.

(121) Marsico, G.; Chambers, V. S.; Sahakyan, A. B.; McCauley, P.; Boutell, J.

M.; Antonio, M. Di; Balasubramanian, S. Whole Genome Experimental Maps of DNA G-Quadruplexes in Multiple Species. *Nucleic Acids Res.* **2019**, *47* (8), 3862–3874. https://doi.org/10.1093/NAR/GKZ179.

(122) Zizza, P.; Cingolani, C.; Artuso, S.; Salvati, E.; Rizzo, A.; D'Angelo, C.; Porru, M.; Pagano, B.; Amato, J.; Randazzo, A.; Novellino, E.; Stoppacciaro, A.; Gilson, E.; Stassi, G.; Leonetti, C.; Biroccio, A. Intragenic G-Quadruplex Structure Formed in the Human CD133 and Its Biological and Translational Relevance. *Nucleic Acids Res.* **2016**, *44* (4), 1579–1590. https://doi.org/10.1093/NAR/GKV1122.

(123) Kim, N. The Interplay between G-Quadruplex and Transcription. *Curr. Med. Chem.* **2017**, *26* (16), 2898–2917. https://doi.org/10.2174/0929867325666171229132619.

(124) Brown, S.; Pharmaceuticals, S. K.-; 2021, undefined. The I-Motif as a Molecular Target: More than a Complementary DNA Secondary Structure. *mdpi.com* **2021**, *14* (2), 1–25. https://doi.org/10.3390/ph14020096.

(125) Dzatko, S.; Krafcikova, M.; Hänsel-Hertsch, R.; Fessl, T.; Fiala, R.; Loja, T.; Krafcik, D.; Mergny, J. L.; Foldynova-Trantirkova, S.; Trantirek, L. Evaluation of the Stability of DNA I-Motifs in the Nuclei of Living Mammalian Cells. *Angew. Chemie - Int. Ed.* **2018**, *57* (8), 2165–2169. https://doi.org/10.1002/ANIE.201712284.

(126) Zeraati, M.; Langley, D. B.; Schofield, P.; Moye, A. L.; Rouet, R.; Hughes, W. E.; Bryan, T. M.; Dinger, M. E.; Christ, D. I-Motif DNA Structures Are Formed in the Nuclei of Human Cells. *Nat. Chem. 2018 106* **2018**, *10* (6), 631–637. https://doi.org/10.1038/s41557-018-0046-3.

(127) Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. Quantitative Visualization of DNA G-Quadruplex Structures in Human Cells. *Nat. Chem. 2013 53* **2013**, *5* (3), 182–186. https://doi.org/10.1038/nchem.1548.

(128) Kosiol, N.; Juranek, S.; Brossart, P.; Heine, A.; Paeschke, K. G-Quadruplexes: A Promising Target for Cancer Therapy. *Mol. Cancer* **2021**,

20 (1). https://doi.org/10.1186/S12943-021-01328-4.

(129) Amato, J.; Iaccarino, N.; Randazzo, A.; Novellino, E.; Pagano, B. Noncanonical DNA Secondary Structures as Drug Targets: The Prospect of the i-Motif. *ChemMedChem* **2014**, *9* (9), 2026–2030. https://doi.org/10.1002/CMDC.201402153.

(130) Verma, A.; Yadav, V. K.; Basundra, R.; Kumar, A.; Chowdhury, S. Evidence of Genome-Wide G4 DNA-Mediated Gene Expression in Human Cancer Cells. *Nucleic Acids Res.* **2009**, *37* (13), 4194–4204. https://doi.org/10.1093/NAR/GKN1076.

(131) Shu, B.; Cao, J.; Kuang, G.; Qiu, J.; Zhang, M.; Zhang, Y.; Wang, M.; Li, X.; Kang, S.; Ou, T. M.; Tan, J. H.; Huang, Z. S.; Li, D. Syntheses and Evaluation of New Acridone Derivatives for Selective Binding of Oncogene C-: Myc Promoter i-Motifs in Gene Transcriptional Regulation. *Chem. Commun.* **2018**, *54* (16), 2036–2039. https://doi.org/10.1039/C8CC00328A.

(132) Brown, R. V; Wang, T.; Reddy Chappeta, V.; Wu, G.; Onel, B.; Chawla, R.; Quijada, H.; Camp, S. M.; Chiang, E. T.; Lassiter, Q. R.; Lee, C.; Phanse, S.; Turnidge, M. A.; Zhao, P.; Garcia, J. G. N.; Gokhale, V.; Yang, D.; Hurley, L. H. The Consequences of Overlapping G-Quadruplexes and i-Motifs in the Platelet-Derived Growth Factor Receptor β Core Promoter Nuclease Hypersensitive Element Can. *ACS Publ.* **2017**, *139* (22), 7456–7475. https://doi.org/10.1021/jacs.6b10028.

(133) Satpathi, S.; Sappati, S.; Das, K.; Hazra, P. Structural Characteristics Requisite for the Ligand-Based Selective Detection of i-Motif DNA. *Org. Biomol. Chem.* **2019**, *17* (21), 5392–5399. https://doi.org/10.1039/C9OB01020C.

(134) Masoud, S. S.; Nagasawa, K. I-Motif-Binding Ligands and Their Effects on the Structure and Biological Functions of I-Motif. *Chem. Pharm. Bull.* **2018**, *66* (12), 1091–1103. https://doi.org/10.1248/CPB.C18-00720.

(135) Rodriguez, R.; Müller, S.; Yeoman, J. A.; Trentesaux, C.; Riou, J. F.;

Balasubramanian, S. A Novel Small Molecule That Alters Shelterin Integrity and Triggers a DNA-Damage Response at Telomeres. *J. Am. Chem. Soc.* **2008**, *130* (47), 15758–15759. https://doi.org/10.1021/JA805615W.

(136) Müller, S.; Sanders, D. A.; Di Antonio, M.; Matsis, S.; Riou, J. F.; Rodriguez, R.; Balasubramanian, S. Pyridostatin Analogues Promote Telomere Dysfunction and Long-Term Growth Inhibition in Human Cancer Cells. *Org. Biomol. Chem.* **2012**, *10* (32), 6537–6546. https://doi.org/10.1039/C2OB25830G.

(137) Di Fonzo, S.; Amato, J.; D'Aria, F.; Caterino, M.; D'Amico, F.; Gessini, A.; Brady, J. W.; Cesàro, A.; Pagano, B.; Giancola, C. Ligand Binding to G-Quadruplex DNA: New Insights from Ultraviolet Resonance Raman Spectroscopy. *Phys. Chem. Chem. Phys.* **2020**, *22* (15), 8128–8140. https://doi.org/10.1039/D0CP01022G.

(138) Pagano, A.; Iaccarino, N.; Abdelhamid, M. A. S.; Brancaccio, D.; Garzarella, E. U.; Di Porzio, A.; Novellino, E.; Waller, Z. A. E.; Pagano, B.; Amato, J.; Randazzo, A. Common G-Quadruplex Binding Agents Found to Interact with i-Motif-Forming DNA: Unexpected Multi-Target-Directed Compounds. *Front. Chem.* **2018**, *6*, 281. https://doi.org/10.3389/FCHEM.2018.00281/BIBTEX.

(139) Quagliata, M.; Nuti, F.; Real-fernandez, F.; Kirilova, K. K.; Santoro, F.; Carotenuto, A.; Papini, A. M.; Rovero, P. Glucopeptides Derived from Myelin-relevant Proteins and Hyperglucosylated Non-typeable. 1–17.

(140) Pagano, B.; Cosconati, S.; Gabelica, V.; Petraccone, L.; De Tito, S.; Marinelli, L.; La Pietra, V.; Saverio di Leva, F.; Lauri, I.; Trotta, R.; Novellino, E.; Giancola, C.; Randazzo, A. State-of-the-Art Methodologies for the Discovery and Characterization of DNA G-Quadruplex Binders. *Curr. Pharm. Des.* **2012**, *18* (14), 1880–1899. https://doi.org/10.2174/138161212799958332.

(141) Dai, J.; Hatzakis, E.; Hurley, L. H.; Yang, D. I-Motif Structures Formed in

the Human c-MYC Promoter Are Highly Dynamic-Insights into Sequence Redundancy and I-Motif Stability. *PLoS One* **2010**, *5* (7). https://doi.org/10.1371/JOURNAL.PONE.0011647.

(142) Kendrick, S.; Akiyama, Y.; Hecht, S. M.; Hurley, L. H. The I-Motif in the Bcl-2 P1 Promoter Forms an Unexpectedly Stable Structure with a Unique 8:5:7 Loop Folding Pattern. *J. Am. Chem. Soc.* **2009**, *131* (48), 17667–17676. https://doi.org/10.1021/JA9076292.

(143) Ambrus, A.; Chen, D.; Dai, J.; Jones, R.; Biochemistry, D. Y.-; 2005, undefined. Solution Structure of the Biologically Relevant G-Quadruplex Element in the Human c-MYC Promoter. Implications for G-Quadruplex Stabilization. *ACS Publ.* **2005**, *44* (6), 2048–2058. https://doi.org/10.1021/bi048242p.

(144) Dai, J.; Carver, M.; Yang, D. Polymorphism of Human Telomeric Quadruplex Structures. *Biochimie* **2008**, *90* (8), 1172–1183. https://doi.org/10.1016/J.BIOCHI.2008.02.026.

(145) Pagano, B.; Amato, J.; Iaccarino, N.; Cingolani, C.; Zizza, P.; Biroccio, A.; Novellino, E.; Randazzo, A. Looking for Efficient G-Quadruplex Ligands: Evidence for Selective Stabilizing Properties and Telomere Damage by Drug-like Molecules. *ChemMedChem* **2015**, *10* (4), 640–649. https://doi.org/10.1002/CMDC.201402552.

(146) Phan, A. T.; Guéron, M.; Leroy, J. L. The Solution Structure and Internal Motions of a Fragment of the Cytidine-Rich Strand of the Human Telomere. *J. Mol. Biol.* **2000**, *299* (1), 123–144. https://doi.org/10.1006/JMBI.2000.3613.

(147) Wold, S.; Esbensen, K.; Geladi, P. Principal Component Analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2* (1–3), 37–52. https://doi.org/10.1016/0169-7439(87)80084-9.

(148) Giancola, C.; Pagano, B. Energetics of Ligand Binding to G-Quadruplexes. *Top. Curr. Chem.* **2013**, *330*, 211–242. https://doi.org/10.1007/128-2012-

347.

(149) King, J. J.; Irving, K. L.; Evans, C. W.; Chikhale, R. V.; Becker, R.; Morris, C. J.; Peña Martinez, C. D.; Schofield, P.; Christ, D.; Hurley, L. H.; Waller, Z. A. E.; Iyer, K. S.; Smith, N. M. DNA G-Quadruplex and i-Motif Structure Formation Is Interdependent in Human Cells. *J. Am. Chem. Soc.* **2020**, *142* (49), 20600–20604. https://doi.org/10.1021/JACS.0C11708.

(150) Iaccarino, N.; Di Porzio, A.; Amato, J.; Pagano, B.; Brancaccio, D.; Novellino, E.; Leardi, R.; Randazzo, A. Assessing the Influence of PH and Cationic Strength on I-Motif DNA Structure. *Anal. Bioanal. Chem.* **2019**, *411* (28), 7473–7479. https://doi.org/10.1007/S00216-019-02120-6.

(151) Amato, J.; Pagano, A.; Cosconati, S.; Amendola, G.; Fotticchia, I.; Iaccarino, N.; Marinello, J.; De Magis, A.; Capranico, G.; Novellino, E.; Pagano, B.; Randazzo, A. Discovery of the First Dual G-Triplex/G-Quadruplex Stabilizing Compound: A New Opportunity in the Targeting of G-Rich DNA Structures? *Biochim. Biophys. Acta - Gen. Subj.* **2017**, *1861* (5), 1271–1280. https://doi.org/10.1016/J.BBAGEN.2016.11.008.

(152) Musumeci, D.; Amato, J.; Zizza, P.; Platella, C.; Cosconati, S.; Cingolani, C.; Biroccio, A.; Novellino, E.; Randazzo, A.; Giancola, C.; Pagano, B.; Montesarchio, D. Tandem Application of Ligand-Based Virtual Screening and G4-OAS Assay to Identify Novel G-Quadruplex-Targeting Chemotypes. *Biochim. Biophys. Acta - Gen. Subj.* **2017**, *1861* (5), 1341–1352. https://doi.org/10.1016/J.BBAGEN.2017.01.024.

(153) Yang, S. Y.; Chang, E. Y. C.; Lim, J.; Kwan, H. H.; Monchaud, D.; Yip, S.; Stirling, P. C.; Wong, J. M. Y. G-Quadruplexes Mark Alternative Lengthening of Telomeres. *NAR Cancer* **2021**, *3* (3). https://doi.org/10.1093/NARCAN/ZCAB031.

(154) Narayan, S.; Muldoon, J.; Finn, M. G.; Fokin, V. V; Kolb, H. C.; Sharpless, K. B. "On Water": Unique Reactivity of Organic Compounds in Aqueous Suspension. *Wiley Online Libr.* **2005**, *44* (21), 3275–3279.

https://doi.org/10.1002/anie.200462883.

(155) Hayashi, Y. In Water or in the Presence of Water? *Angew. Chemie - Int. Ed.* **2006**, *45* (48), 8103–8104. https://doi.org/10.1002/ANIE.200603378.

(156) Blackmond, D. G.; Armstrong, A.; Coombe, V.; Wells, A. Water in Organocatalytic Processes: Debunking the Myths. *Angew. Chemie - Int. Ed.* **2007**, *46* (21), 3798–3800. https://doi.org/10.1002/ANIE.200604952.

(157) Chanda, A.; Fokin, V. V. Organic Synthesis "on Water." *Chem. Rev.* **2009**, *109* (2), 725–748. https://doi.org/10.1021/CR800448Q/ASSET/CR800448Q.FP.PNG_V03.

(158) Lipshutz, B. H.; Ghorai, S.; Cortes-Clerget, M. The Hydrophobic Effect Applied to Organic Synthesis: Recent Synthetic Chemistry "in Water." *Chem. – A Eur. J.* **2018**, *24* (26), 6672–6695. https://doi.org/10.1002/CHEM.201705499.

(159) Lipshutz, B. H. When Does Organic Chemistry Follow Nature's Lead and "Make the Switch"? *J. Org. Chem.* **2017**, *82* (6), 2806–2816. https://doi.org/10.1021/ACS.JOC.7B00010.

(160) La Sorella, G.; Strukul, G.; Scarso, A. Recent Advances in Catalysis in Micellar Media. *Green Chem.* **2015**, *17* (2), 644–683. https://doi.org/10.1039/C4GC01368A.

(161) Sheldon, R. A. Metrics of Green Chemistry and Sustainability: Past, Present, and Future. *ACS Sustain. Chem. Eng.* **2018**, *6* (1), 32–48. https://doi.org/10.1021/ACSSUSCHEMENG.7B03505.

(162) Handa, S.; Wang, Y.; Gallou, F.; Lipshutz, B. H. Sustainable Fe-Ppm Pd Nanoparticle Catalysis of Suzuki-Miyaura Cross-Couplings in Water. *Science (80-.).* **2015**, *349* (6252), 1087–1091. https://doi.org/10.1126/SCIENCE.AAC6936.

(163) Bu, M.; Lu, G.; Jiang, J.; Technology, C. C.-C. S. &; 2018, undefined. Merging Visible-Light Photoredox and Micellar Catalysis: Arylation Reactions with Anilines Nitrosated in Situ. *pubs.rsc.org*.

(164) Giedyk, M.; Narobe, R.; Weiss, S.; Touraud, D.; Catalysis, W. K.-N.; 2020, undefined. Photocatalytic Activation of Alkyl Chlorides by Assembly-Promoted Single Electron Transfer in Microheterogeneous Solutions. *nature.com*.

(165) Santos, M. S.; Cybularczyk-Cecotka, M.; König, B.; Giedyk, M. Minisci C−H Alkylation of Heteroarenes Enabled by Dual Photoredox/Bromide Catalysis in Micellar Solutions**. *Chem. - A Eur. J.* **2020**, *26* (66), 15323–15329. https://doi.org/10.1002/CHEM.202002320.

(166) Jana, S.; Verma, A.; Kadu, R.; science, S. K.-C.; 2017, undefined. Visible-Light-Induced Oxidant and Metal-Free Dehydrogenative Cascade Trifluoromethylation and Oxidation of 1, 6-Enynes with Water. *pubs.rsc.org*.

(167) Pelliccia, S.; Alfano, A. I.; Luciano, P.; Novellino, E.; Massarotti, A.; Tron, G. C.; Ravelli, D.; Giustiniano, M. Photocatalytic Isocyanide-Based Multicomponent Domino Cascade toward the Stereoselective Formation of Iminofurans. *J. Org. Chem.* **2020**, *85* (4), 1981–1990. https://doi.org/10.1021/ACS.JOC.9B02709.

(168) Zou, Y.; Guo, W.; Liu, F.; Lu, L.; Chen, J.; Chemistry, W. X.-G.; 2014, undefined. Visible-Light-Induced Photocatalytic Formyloxylation Reactions of 3-Bromooxindoles with Water and DMF: The Scope and Mechanism. *pubs.rsc.org*.

(169) Bu, M.; Cai, C.; Gallou, F.; chemistry, B. L.-G.; 2018, undefined. PQS-Enabled Visible-Light Iridium Photoredox Catalysis in Water at Room Temperature. *pubs.rsc.org*.

(170) Ramamurthy, V. Micellar Control of Photochemical Reactions. *Proc. Indian Acad. Sci. - Chem. Sci.* **1984**, *93* (4), 635–646. https://doi.org/10.1007/BF02863282.

(171) Fendler, J. H. Microemulsions, Micelles, and Vesicles as Media for Membrane Mimetic Photochemistry. *J. Phys. Chem.* **1980**, *84* (12), 1485–

1491. https://doi.org/10.1021/J100449A012.

(172) Cannalire, R.; Amato, J.; Summa, V.; Novellino, E.; Tron, G. C.; Giustiniano, M. Visible-Light Photocatalytic Functionalization of Isocyanides for the Synthesis of Secondary Amides and Ketene Aminals. *J. Org. Chem.* **2020**, *85* (21), 14077–14086. https://doi.org/10.1021/ACS.JOC.0C01946.

(173) Russo, C.; Cannalire, R.; Luciano, P.; Brunelli, F.; Tron, G. C.; Giustiniano, M. Visible-Light Photocatalytic Ugi/Aza-Wittig Cascade towards 2-Aminomethyl-1,3,4-Oxadiazole Derivatives. *Synth.* **2021**, *53* (23), 4419–4427. https://doi.org/10.1055/A-1543-3924.

(174) Rueping, M.; Vila, C.; Bootwicha, T. Continuous Flow Organocatalytic C-H Functionalization and Cross-Dehydrogenative Coupling Reactions: Visible Light Organophotocatalysis for Multicomponent Reactions and C-C, C-P Bond Formations. *ACS Catal.* **2013**, *3* (7), 1676–1680. https://doi.org/10.1021/CS400350J.

(175) Rueping, M.; Vila, C. Visible Light Photoredox-Catalyzed Multicomponent Reactions. *Org. Lett.* **2013**, *15* (9), 2092–2095. https://doi.org/10.1021/OL400317V.

(176) Schönhoff, M. NMR Studies of Sorption and Adsorption Phenomena in Colloidal Systems. *Curr. Opin. Colloid Interface Sci.* **2013**, *18* (3), 201–213. https://doi.org/10.1016/J.COCIS.2013.03.004.

(177) Schneider, G.; Bolink, H.; Constable, E.; … C. E.-D.; 2014, undefined. Chloride Ion Impact on Materials for Light-Emitting Electrochemical Cells. *pubs.rsc.org*.

(178) Marion, D.; Wüthrich, K. Application of Phase Sensitive Two-Dimensional Correlated Spectroscopy (COSY) for Measurements of 1H-1H Spin-Spin Coupling Constants in Proteins. *Biochem. Biophys. Res. Commun.* **1983**, *113* (3), 967–974. https://doi.org/10.1016/0006-291X(83)91093-8.

(179) Brown, L.; Bösch, C.; (BBA), K. W.-B. et B. A.; 1981, undefined. Location

and Orientation Relative to the Micelle Surface for Glucagon in Mixed Micelles with Dodecylphosphocholine EPR and NMR Studies. *Elsevier*.

(180) Lindberg, M.; Jarvet, J.; Langel, Ü.; Biochemistry, A. G.-; 2001, undefined. Secondary Structure and Position of the Cell-Penetrating Peptide Transportan in SDS Micelles as Determined by NMR. *ACS Publ.* **2001**, *40* (10), 3141–3149. https://doi.org/10.1021/bi0008985.

(181) Guerrero, I.; Segundo, M. S.; Correa, A. Iron-Catalyzed C(Sp3)–H Functionalization of N,N-Dimethylanilines with Isocyanides. *Chem. Commun.* **2018**, *54* (13), 1627–1630. https://doi.org/10.1039/C7CC09872C.

(182) Zhou, H.; Lu, P.; Gu, X.; Li, P. Visible-Light-Mediated Nucleophilic Addition of an α-Aminoalkyl Radical to Isocyanate or Isothiocyanate. *Org. Lett.* **2013**, *15* (22), 5646–5649. https://doi.org/10.1021/OL402573J.

(183) Ye, X.; Xie, C.; Huang, R.; Liu, J. Direct Synthesis of α-Amino from N-Alkyl Amines by the Copper-Catalyzed Ugi-Type Reaction. *Synlett* **2012**, *2012* (03), 409–412. https://doi.org/10.1055/S-0031-1290319.

(184) De Castro, G. V.; Worm, D. J.; Grabe, G. J.; Rowan, F. C.; Haggerty, L.; De La Lastra, A. L.; Popescu, O.; Helaine, S.; Barnard, A. Characterization of the Key Determinants of Phd Antitoxin Mediated Doc Toxin Inactivation in Salmonella. *ACS Chem. Biol.* **2022**, *17* (6), 1598–1606. https://doi.org/10.1021/ACSCHEMBIO.2C00276/ASSET/IMAGES/MEDIUM/CB2C00276_M001.GIF.

(185) Sterckx, Y.; Gieter, S. De; Zorzini, V.; … S. H.-P. expression and; 2015, undefined. An Efficient Method for the Purification of Proteins from Four Distinct Toxin–Antitoxin Modules. *Elsevier*.

(186) Castro-Roa, D.; Garcia-Pino, A.; … S. D. G.-N. chemical; 2013, undefined. The Fic Protein Doc Uses an Inverted Substrate to Phosphorylate and Inactivate EF-Tu. *nature.com*.

(187) Cruz, J.; Rothenbacher, F.; … T. M.-J. of B.; 2014, undefined. Doc Toxin Is a Kinase That Inactivates Elongation Factor Tu. *ASBMB*.

(188) Garcia-Pino, A.; Balasubramanian, S.; Wyns, L.; Gazit, E.; De Greve, H.; Magnuson, R. D.; Charlier, D.; van Nuland, N. A. J.; Loris, R. Allostery and Intrinsic Disorder Mediate Transcription Regulation by Conditional Cooperativity. *Cell* **2010**, *142* (1), 101–111. https://doi.org/10.1016/j.cell.2010.05.039.

(189) Pace, C.; journal, J. S.-B.; 1998, undefined. A Helix Propensity Scale Based on Experimental Studies of Peptides and Proteins. *Elsevier*.

(190) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **2010**, *66* (Pt 4), 486–501. https://doi.org/10.1107/S0907444910007493.

(191) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards Automated Crystallographic Structure Refinement with Phenix.Refine. *Acta Crystallogr. D. Biol. Crystallogr.* **2012**, *68* (Pt 4), 352–367. https://doi.org/10.1107/S0907444912001308.

(192) Tornøe, C. W.; Christensen, C.; Meldal, M. Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal Alkynes to Azides. *J. Org. Chem.* **2002**, *67* (9), 3057–3064. https://doi.org/10.1021/JO011148J/SUPPL_FILE/JO011148J_S.PDF.

(193) Wrobel, M.; Aubé, J.; König, B. Parallel Solid-Phase Synthesis of Diaryltriazoles. *Beilstein J. Org. Chem. 8115* **2012**, *8* (1), 1027–1036. https://doi.org/10.3762/BJOC.8.115.